



Influence of biomolecular environment on fluorophore photophysics

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Aiswaria Prakash

from Kochi, India

Düsseldorf, December 2022

from the Institute for Molecular Physical Chemistry,
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Printed and published with the support of the
German Academic Exchange Service (DAAD)

Supervisor: Prof. Dr. Claus Seidel

Mentor: Prof. Dr. Lutz Schmitt

Date of the oral examination: 24.02.2023

Declaration

I declare under oath that I have produced my thesis independently and without any undue assistance by third parties under consideration of the 'Principles for the Safeguarding of Good Scientific Practice at Heinrich Heine University Düsseldorf'.

The experimental work for this thesis was conducted at the Heinrich Heine University, Düsseldorf.

Dedicated to the little one

Abstract

Fluorescent labels are widely used in biochemistry to study the conformational dynamics and intermolecular interactions in biomolecules. Understanding the basic photophysical mechanisms of these fluorescent labels and their environment dependence can improve the design and accuracy of biophysical techniques. In the presented work, the influence of biomolecular environment on the fluorescent properties of covalently-linked organic dye labels are investigated and the potential of fluorescent probes to sense the local and global biomolecular environments demonstrated.

To reveal the distinct photophysical signatures of different RNA micro-environments, a systematic study of the fluorescence properties of commonly used organic fluorophores, Alexa488 and Cy5, in different RNA environments are carried out. The results are corroborated by accessible volume (AV) and molecular dynamics (MD) simulations to provide structural insights on potential interaction modes. In addition to sensing the local environment, the ability of fluorescent labels to shed light on the tertiary and global structures of RNA is demonstrated. The influence of the protein environment on dye properties is explored with Alexa488 and Alexa647 dyes labeled on the bacteriophage T4 Lysozyme. Combining spectroscopic studies and MD simulations, a detailed insight into the origin, extent and modes of dye-protein interactions is achieved. AV calculations are employed to determine the role of positively charged amino acids in dye trapping and in providing a quenching environment. AV also reveals the manner in which protein dynamics affect dye properties. Detection of substrate binding and protein denaturation are reported as illustrative cases of fluorescence sensing.

Förster resonance energy transfer (FRET) is a fluorescence-based technique that employs a pair of fluorescent labels to report on the structural dynamics of a biomolecule. The insights obtained from the photophysical studies of dyes are discussed in the context of FRET measurements and guidelines are provided to avoid or minimize detrimental effects of dye-biomolecular interactions. Those interactions which significantly alter the dye distribution are incorporated into the label description for FRET dyes and these modifications proven to significantly improve the accuracy of FRET-derived structural models. Thus, in future fluorescence-based studies, the knowledge acquired from the present study regarding the environmental sensitivity of dyes in different biomolecular environments can be used to improve the label or model design, the experiment plan or the interpretation of observations.

Contents

| | |
|---|-----------|
| Declaration | i |
| Abstract | iii |
| Contents | iv |
| Chapter 1. Introduction | 1 |
| 1.1 Fluorescently labeled biomolecules | 1 |
| 1.1.1 Sites for fluorophore labeling and coupling chemistry | 1 |
| a) Nucleic acid labeling sites | 2 |
| b) Protein labeling sites | 4 |
| c) Labeling chemistry | 5 |
| 1.1.2 Influence of environment on the fluorescent properties of coupled dyes | 7 |
| a) Absorption spectral shifts | 7 |
| b) Anisotropy changes | 7 |
| c) Fluorescence quenching | 9 |
| d) Fluorescence enhancement | 15 |
| 1.1.3 Application: Fluorescent labels as reporters of biomolecular environment. | 20 |
| a) Utilizing the environment sensitivity of fluorescence quenching | 20 |
| b) Interpreting the environment from fluorescence enhancement | 23 |
| c) Dye labels in other notable fluorescence-based techniques | 24 |
| 1.2 Motivation and objectives of present study | 26 |
| Chapter 2. Theoretical background | 30 |
| 2.1 Basics of fluorescence | 30 |
| a) Radiative and non-radiative processes | 30 |
| b) Fluorescence lifetime and quantum yield | 31 |
| c) Fluorescence anisotropy | 32 |
| d) Fluorescence quenching by photoinduced electron transfer | 33 |
| e) Association-induced fluorescence enhancement | 34 |
| 2.2 Time resolved fluorescence measurement | 35 |
| 2.3 Fluorescence correlation spectroscopy (FCS) | 37 |
| 2.4 Förster resonance energy transfer (FRET) | 39 |
| Chapter 3. Sensing the local environment and global conformation of RNA by flexibly coupled fluorescent labels | 42 |
| 3.1 Abstract | 43 |
| 3.2 Introduction | 43 |
| 3.3 Results | 47 |

| | | |
|--------|---|-----|
| 3.3.1 | Photophysics of Cy5 are sensitive to the local environment | 48 |
| 3.3.2 | Cy5 labels close to the 3'-end stack to the terminal base pair | 51 |
| 3.3.3 | Cy5 labels interact with the exposed bases at the junction | 54 |
| 3.3.4 | Site-specific photophysics of Alexa488 are dominated by quenching | 55 |
| 3.3.5 | End-stacked Alexa488 fluorophores are efficiently quenched by PET | 57 |
| 3.3.6 | Predicting photophysical properties from accessible volume simulations..... | 60 |
| 3.3.7 | Terminal base pair accessibility predicts Cy5 stacking interactions..... | 61 |
| 3.3.8 | The behavior of Cy5 labels depends on the labeling strategy..... | 63 |
| 3.3.9 | PET quenching of Alexa488 correlates with terminal guanine accessibility..... | 64 |
| 3.3.10 | MD simulations explore RNA-dye interaction modes | 65 |
| 3.3.11 | Origin of PET quenching for internal labels | 68 |
| 3.3.12 | Effect of the primary sequence | 69 |
| 3.3.13 | Sensing the global structure | 70 |
| 3.3.14 | Detecting RNA shape and size by polarization-resolved FCS. | 72 |
| 3.4 | Discussion | 74 |
| 3.5 | Conclusion and Outlook | 83 |
| 3.6 | Materials and methods..... | 83 |
| 3.7 | Supporting information | 100 |

Chapter 4. Fluorescent dyes probe the local environment of protein T4 Lysozyme:

| | | |
|-------|---|------------|
| | Exploring with experiments and simulations | 144 |
| 4.1 | Introduction..... | 145 |
| 4.2 | Results and Discussion..... | 147 |
| 4.2.1 | Alexa488 labels exhibit environment-sensitive photophysical properties | 148 |
| 4.2.2 | Simulation studies predict probable dye-protein interactions | 152 |
| | a) MD simulations allow a closer look at interaction modes | 153 |
| | b) AV simulations predict changes in fluorescence parameters..... | 157 |
| 4.2.3 | The photophysics of coupled-Alexa647 dyes are influenced by the T4L local environment | 161 |
| 4.2.4 | The source of steric restriction unraveled by AV simulations..... | 165 |
| 4.2.5 | Singly labeled dyes act as reporters of local environment and global events.... | 167 |
| | a) Information about local environment | 167 |
| | b) Detecting protein unfolding and denaturation..... | 167 |
| | c) Detecting substrate binding | 168 |
| 4.3 | Conclusions and outlook | 170 |
| 4.4 | Materials and methods..... | 170 |
| 4.5 | Supporting information | 179 |

| | | |
|-------------------|---|------------|
| Chapter 5. | Developing dye models for accurate FRET label description | 195 |
| 5.1 | Introduction | 196 |
| 5.1.1 | A review on the dye description of organic fluorescence labels..... | 196 |
| 5.1.2 | Prior experimental input and its role in FRET experiments | 200 |
| 5.2 | Results and Discussion | 201 |
| 5.2.1 | dsRNA rulers: Experimental vs predicted distance $\langle R_{DA} \rangle_E$ | 201 |
| 5.2.2 | Improving dye model based on prior knowledge | 207 |
| 5.3 | Conclusions | 209 |
| 5.4 | Materials and Methods | 210 |
| Chapter 6. | Summary | 212 |
| Chapter 7. | Additional methods | 217 |
| | Frequently used abbreviations | 228 |
| | Acknowledgements | 229 |
| | References..... | 230 |
| | Published paper..... | 253 |

Chapter 1. Introduction

1.1 Fluorescently labeled biomolecules

Fluorescence is a powerful approach for studying the conformational dynamics and intra- and intermolecular interactions occurring in biomolecules. Fluorescently labeled biomolecules have been widely used in the mechanistic studies of proteins (transcription factors¹, chromatin², enzyme³ etc.) and nucleic acids (DNA⁴, Holliday junction⁵), cell imaging (of single RNA molecules⁶, multiplexed RNA⁷, cancer cells⁸, proteins⁹), diagnostic assays (probing nucleic acid hybridization¹⁰, oligonucleotide structure¹¹, viral nucleic acid¹², disease¹³, serum proteins¹⁴) etc. Fluorescence has multiple parameters that can be exploited to yield information about the biomolecule, namely fluorescence lifetime and anisotropy, absorption and emission spectra and fluorescence quantum yield and brightness^{15,16}. Additionally, when two or more fluorophores are attached to a biomolecule, for example, in Förster resonance energy transfer (FRET), we obtain additional information via the dipolar coupling¹⁵. In this popular biophysical technique, distance-dependent energy transfer occurs between site-specifically labeled donor and acceptor fluorescent labels^{17,18}. Fluorescence correlation spectroscopy (FCS) is another widely-used technique where the conformational flexibility, structural dynamics and the hydrodynamic transport properties in biomolecules are determined from the fluctuating signal (brightness) of a fluorophore diffusing through a confocal volume¹⁹⁻²¹. In this way, various distinct fluorescence-based techniques can be applied to monitor biomolecular processes spanning over a time range from femtoseconds to seconds. The attractive features of fluorescence spectroscopy can be effectively utilized by the selection of an appropriate fluorescence technique²² for the required application, the choice and design of suitable fluorescent probes and a thorough understanding of their photophysical properties.

1.1.1 Sites for fluorophore labeling and coupling chemistry

Fluorescent labeling can be performed at different sites in a nucleic acid or protein and the chemistry of dye-biomolecule coupling is diverse. The choice of labeling sites and chemistry has to be made with a focus on the requirements of the study. A general overview of the possible sites and methods of dye attachment are described below.

a) Nucleic acid labeling sites

Since the nucleobases in nucleic acids are intrinsically non-fluorescent, they are replaced with a chemically modified fluorescent nucleobase analogue to introduce fluorescent nature. These modifications result in isomorphous, expanded, chromophoric or extended nucleosides²³. While isomorphous nucleoside analogues contain small substituents or additional heteroatoms on the natural nucleoside, expanded nucleosides contain additional aromatic rings on the natural nucleoside and chromophoric nucleosides replace the natural nucleoside analogues with bulky, aromatic chromophores. This thesis deals exclusively with extended modifications where small molecule organic fluorophores are covalently coupled to the nucleosides via linkers. This is accomplished either through groups present naturally in nucleic acids (like amino, hydroxyl, phosphate groups) or through some other reactive groups (primary amines, thiols, azides etc.) introduced specifically for targeted coupling²⁴. Covalent fluorescent labeling can occur at various sites on a nucleic acid, namely the phosphate backbone, the sugar moiety or the nucleobases²⁵⁻³³(Fig. 1.1).

Fluorescent dye coupling at the backbone can occur at phosphate groups at the terminal sites, 5' or 3' end^{24,34,35} or at the internucleotidic phosphate group. In some cases, the phosphate group itself is modified to a phosphorothioate group by substituting sulfur for one of the non-bridging oxygens of the phosphate group³⁶⁻⁴⁰. Of the different possibilities for backbone labeling, the 5' labeling appears more frequently in applications^{36,41,42} than 3' labeling⁴³⁻⁴⁵ or internucleotidic labeling^{39,40,46-48}. The prevalence of 5' labels in comparison to 3' labels could be because most of the biological applications of chemically synthesized oligonucleotides require molecules with free 3'-hydroxyl functions and therefore the common automated solid phase oligonucleotide synthesizers anchor the 3' end of an oligonucleotide to a support such that this end is cleaved only during the final deprotection, while the 5' end is available for modifications⁴⁵. DNA backbone dye labeling has been used in the literature to control the orientation of the dye in FRET studies. Ranjit et al. made use of this strategy by rigidly attaching the Cy5 fluorophore to the DNA backbone by internucleotidic modifications⁴⁸. These rigid attachments reduce the uncertainties in donor-acceptor distance associated with the flexible linkers (commonly referred to as the kappa square uncertainty in FRET⁴⁹) while also preventing undesirable

dye-DNA interactions⁴⁸. Iqbal et al. took control over the relative dye orientation by labeling the cyanine dyes on the 5' terminus where they are expected to predominantly stack in the manner of an additional base pair⁵⁰.

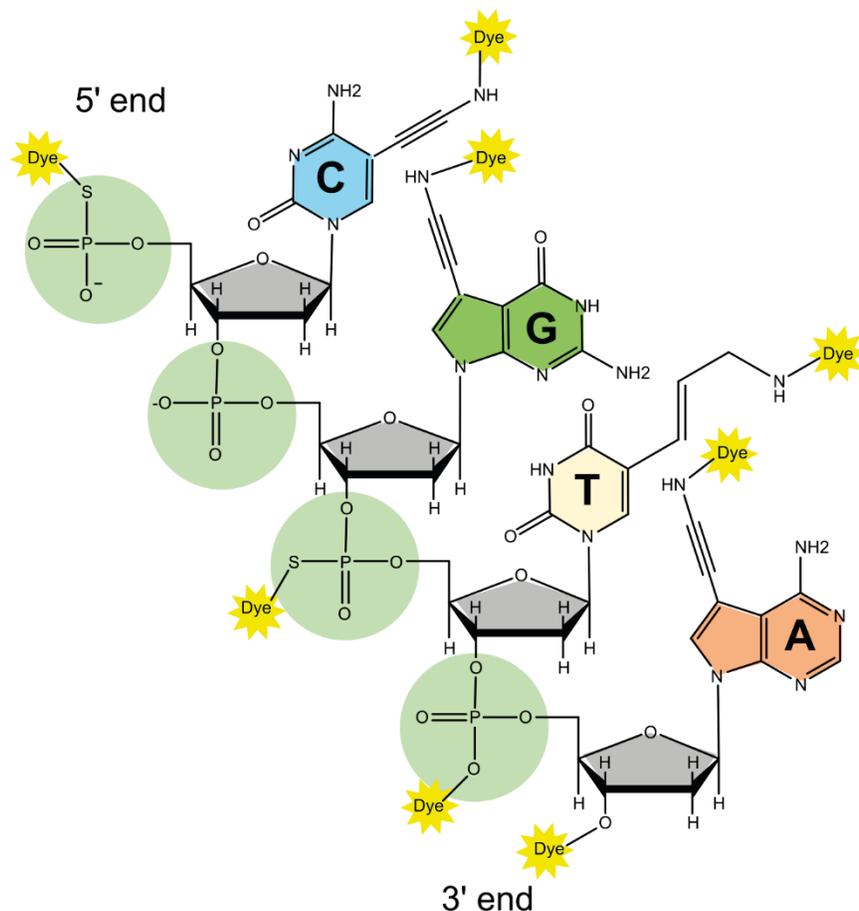


Figure 1.1. Dye labeling sites on the nucleic acid backbone, sugar and modified nucleobase.

Strategies for dye coupling to any nucleobase have been developed where the fluorescent functionalization of the nucleobases is preceded by its chemical modification. This advancement was made possible in the wake of the fluorescence-based DNA sequencing approach which required dye labeling at every position in the DNA sequence^{25,51}. The most ideal attachment site in a B-form DNA is position 5 for pyrimidine bases and position 7 for purines (more specifically, 7-deazapurines), because these positions do not affect the Watson–Crick base pairing and are well-accommodated in the major groove of the DNA duplex. Additionally, the C5 modified pyrimidines and C7 modified 7-deazapurines are favored in DNA polymerase-catalysed nucleobase

polymerization^{25,51}. For the common RNA A-form helix, the situation is expected to be different due to the difference in the groove geometry^{52,53}.

Compared to dye labeling at the backbone and nucleobases, the covalent fluorescent labeling at the sugar moiety is less common. Nevertheless, the 2'(3')N-methylantraniloyl (Mant)⁵⁴ and 2',3'-O-(2,4,6-trinitrophenyl) (TNP)⁵⁵ derivatives have found applications in structural and mechanistic studies of nucleotide binding sites, catalytic sites⁵⁶⁻⁵⁹ etc. In fluorescence-based DNA sequencing, the fluorescent dyes have also been attached to the sugar of dideoxynucleoside triphosphates to act as chain terminators^{27,60}.

b) Protein labeling sites

Fluorescent probes in proteins can be intrinsic or extrinsic. Intrinsic probes are naturally occurring amino acids like Tryptophan (Trp) or Tyrosine(Tyr) which are intrinsically fluorescent⁶¹⁻⁶⁴. Extrinsic fluorophores fall into four main categories: fluorescent proteins, peptide tags, nanoparticle labels and small organic fluorescent molecules. Fluorescent proteins (e.g. Green Fluorescent Protein⁶⁵) are 'built-in' genetic labels approx. 25 kDa in size, fused through recombinant cloning to the fluorescent protein at the N or C terminus^{22,66}. Peptide and protein tags (e.g. SNAP tag⁶⁷) are fluorescently labeled synthetic peptides and special proteins, respectively, which are attached to the target protein^{22,68-70}. Nanoparticle labels such as quantum dots are semiconductor nanocrystals with a reactive core that controls their optical properties (e.g. CdSe core) with high quantum yields and high extinction coefficients coupled to the protein of interest^{71,72}.

In this thesis, I employ small organic fluorescent molecules as fluorescent labels. Organic fluorophores generally have a wide spectral range, small size, good photostability and adequate brightness²². These organic fluorescent molecules are covalently coupled to an amino acid in the protein (Fig. 1.2). The side-chain amine group of lysine and the thiol group of cysteine are suitable handles with which commercially available dye derivatives can be coupled to. Cysteine labeling is highly popular due to this amino acid's relative rarity in the protein and the ease of its introduction into a specific site by site-directed mutagenesis without perturbing the protein's function. Thiol-reactive dye derivatives such as maleimide, conjugations such as haloacetyl, alkyl halide, and disulfide interchange derivatives may be used^{73,74}.

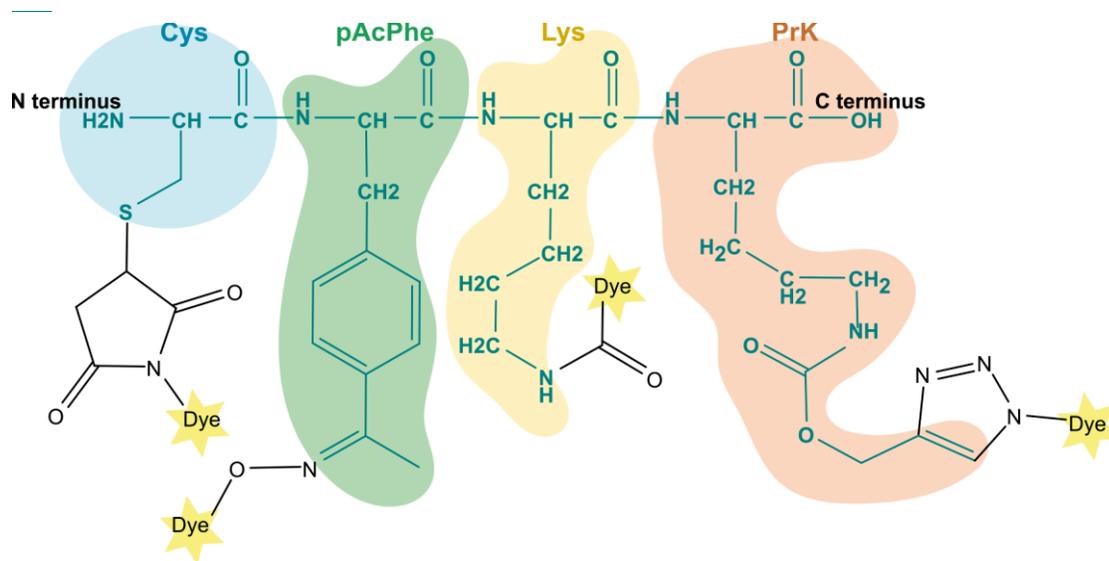


Figure 1.2. Illustration of dye labeling sites on a peptide. Examples of natural (Cys, Lys) and unnatural (pAcPhe, PrK) amino acids used for covalent dye attachment are shown (indicated by shading). The various functionalizations commonly used for dye attachment are also shown (black).

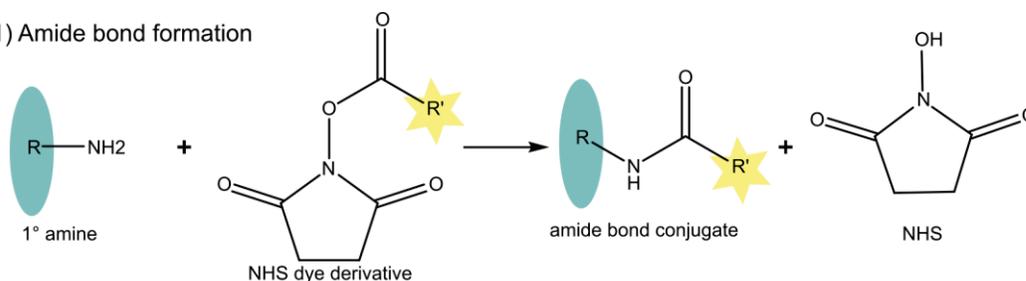
An orthogonal route to achieve site-specific dual-labeling of proteins with dyes is using amber suppression technology to genetically introduce the unnatural amino acid (UAA) p-acetylphenylalanine (pAcPhe). The genetically encoded unnatural ketone amino acid is then labeled with a hydroxylamine-derivative of the desired fluorophore⁷⁵. Other UAAs like homopropargylglycine, propargyllysine (PrK) or ethynylphenylalanine that provide alkynyl sites for attachment of azido dye derivatives may also be genetically incorporated^{76,77}. Similarly, there are UAAs with ring-strained alkynes or alkenes that react, respectively, with dyes containing azide or tetrazine groups⁷⁸.

c) Labeling chemistry

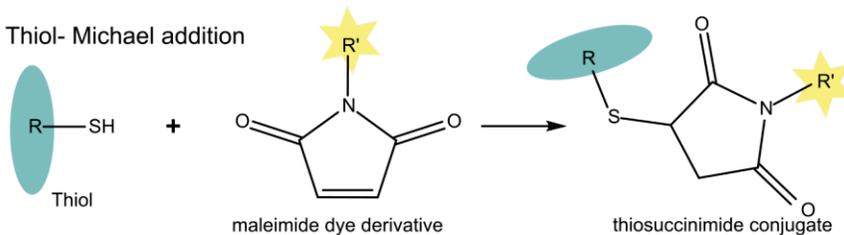
The common coupling chemistries for covalent dye attachment can be described by four basic coupling functionalizations (Scheme 1): (1) Amide bond formation: between amino group in modified nucleobases, in Lys residues or at the N terminus of the protein and the N-hydroxysuccinimide (NHS) derivative of the dye⁷⁹. (2) Michael addition: between a thiol group in the biomolecule and the maleimide derivative of the dye^{34,73}. (3) Click chemistry: cycloaddition between azido and alkynyl groups in the biomolecule and the dye^{34,80-82}. (4) Oxime formation: between the ketone-bearing amino acid, p-acetylphenylalanine, in proteins and hydroxyl amine functionalized dye^{75,83}.

In several practical applications like FRET, PET etc., the fluorescent dye is attached via a linker to the biomolecule. The linker serves as an extended arm that allows larger spatial separation between the labeling site and fluorophore^{25,30,51,84} to allow unrestricted dye movement. Dyes that are covalently attached by chemical linkers are constrained to certain domains on the nucleic acids. Therefore, the effective local concentration of a dye molecule in the immediate vicinity of a particular region of the DNA molecule is increased, reaching effective concentrations of up to 1 M depending on the length of the linker. Thus, dye-DNA interactions, which would otherwise not occur if the dye was not conjugated, are manifested⁸⁵.

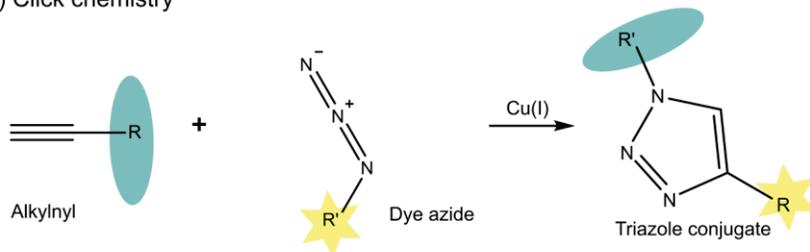
(1) Amide bond formation



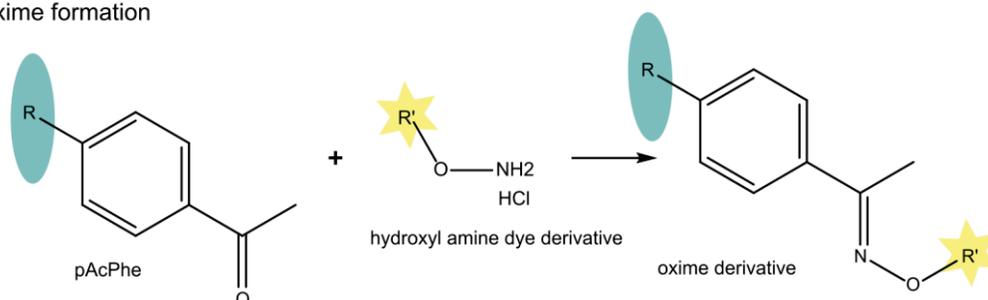
(2) Thiol- Michael addition



(3) Click chemistry



(4) Oxime formation



Scheme 1.1. Common strategies for dye labeling on nucleic acids and proteins

1.1.2 Influence of environment on the fluorescent properties of coupled dyes

When the covalently coupled fluorescent dyes are labeled at a particular site in a nucleic acid or protein, they interact with suitable functional groups in their vicinity. These interactions may be electrostatic⁸⁶⁻⁸⁸, hydrophobic^{89,90} or π -stacking⁹¹⁻⁹⁶ and they modulate the fluorescence properties of the dyes. Some of the general photophysical changes observed are described below.

a) Absorption spectral shifts

When dyes are terminally labeled at nucleic acids, π -stacking interaction of the dye and terminal nucleobase causes electronic coupling between them, and a ground state complex is formed, which results in a characteristic shift of the absorption spectra of the dye. Kupstat et al.⁹⁷ reported a bathochromic shift of 8 nm for the oxazine dye ATTO680 and 7-31 nm shift for ATTO610 due to conjugation with the DNA at the helix terminus. Heinlein et al.⁹⁸ observed bathochromic shifts of 4-6 nm upon addition of deoxynucleotide monophosphates for the absorption maxima of the rhodamine dyes, Rh6G and TMR, and the oxazine dyes, MR121 and JA242. In studies conducted by Li et al.⁹⁹, TMR showed a red shift of about 15 nm in 1 M guanosine. However, investigations from Ranjit and Levitus¹⁰⁰ caution the assignment of bathochromic shifts to ground state complex formation. The 3 nm red shift they observed for the cyanine dye, Cy3B, in concentrated solutions of dGMP could be assigned to changes in the refractive index of the solution when high concentrations of nucleotides are dissolved in the buffer; thus proving that the observed shifts need not necessarily be due to dye-biomolecular interactions.

b) Anisotropy changes

Fluorescence anisotropy is found to be highly sensitive to dye interactions with the biomolecule. When molecular rotations occur during the excited state lifetime of a dye, its fluorescence will be depolarized, resulting in a reduction of anisotropy over time. However, when dye-biomolecular interactions occur, free rotations of the dye are hindered, bringing about an increased anisotropy¹⁰¹. For example, in fluorescein⁴⁴ and MR121⁹⁸ dyes, high anisotropy values were measured when labeled to DNA hairpins due to dye interaction with a guanosine residue at the end of the isolated DNA strand. The interaction between the Elongation factor like-1 GTPase (EFL1) and the Shwachman-

Diamond Syndrome (SBDS) protein was reflected in the fluorescence anisotropy changes of the dye 4',5'-bis(1,3,2-dithioarsolan-2-yl) fluorescein labeled at the C-terminus of the recombinant SBDS protein and helped uncover quantitative and mechanistic information on the binding mode between these proteins¹⁰².

A change in anisotropy can arise from changes in the local segmental motion of the dye as well as changes in the global motion. For the oxazine dye, ATTO610, conjugation with double stranded DNA (dsDNA) leads to delayed fluorescence anisotropy due to a complete loss of segmental mobility⁹⁷. For the dyes Alexa488, Oregon Green 488 and Oregon Green 514 when conjugated to biomolecules, changes in steady-state and time-resolved fluorescence anisotropy occur and are assigned to changes in both local and global dye motions¹⁰¹.

Sindbert et al.¹⁰³ described dsDNA and dsRNA internally labeled with Alexa488 and Cy5 using the wobbling-in-cone (WIC) model. The xanthene chromophore of Alexa488 is perpendicular to the linker. Therefore, Alexa488 anisotropy can sense both the linker wobbling and the dye rotations. Even if the linker is completely stiff, the dye can rotate about the linker, resulting in a low residual anisotropy. The cyanine dye is attached parallel to the linker such that the anisotropy senses predominantly the linker motions, as the dye rotation about the linker hardly affects the fluorescence depolarization and if the linker is stiff, a very high residual anisotropy results. The experimental anisotropy decays also showed that the dye motion is more restricted for the RNA since the groove is deeper and narrower in RNA than in DNA.

Anisotropy changes in cyanine dyes have been observed in other studies too. Steffen et al. reported an increased amplitude of the global anisotropy due to sulfonated Cy3 and Cy5 dyes stacking with the nucleic acid⁹⁰. They also demonstrated that the extent of motional restriction on the dyes by the macromolecular environment can be deduced from dynamic anisotropy measurements. Sanborn et al. showed that Cy3 dye, on interaction with single stranded DNA (ssDNA) experiences lack of rotational freedom when linked to the terminal 5' phosphate. This was evidenced by the slow anisotropy component >2.5 ns which dominates 88% of the anisotropy decay⁴².

c) Fluorescence quenching

When fluorophores with appropriate redox potentials interact with the guanine nucleobase in nucleic acids or an amino acid like Tryptophan in proteins within the van der Waals radius, fluorescence quenching is seen to occur¹⁰⁴. Oxazines (e.g. MR121 and ATTO 655¹⁰⁵), rhodamines (e.g. Rh6G¹⁰⁶, TAMRA¹⁰⁷, Alexa488, Alexa555 and Alexa594¹⁰⁸), fluoresceins (e.g. FAM¹⁰⁹) and BODIPY(e.g. BODIPY® FL¹¹⁰) dyes are some of the most common commercial fluorophores seen to undergo fluorescence quenching.

Table 1.1. Redox properties of selected fluorophores measured vs NHE

| Free Fluorophore | E_{red} [V] | E_{ox} [V] | $E_{0,0}$ [eV] | Solvent | Reference |
|---|--------------------|-------------------|----------------|--------------|--------------------------------|
| 5-FAM | -0.47 | | 2.46 | Aqueous | Torimura et al. ¹¹¹ |
| BODIPY FL | -0.83 | | 2.43 | Aqueous | Torimura et al. ¹¹¹ |
| BODIPY R6G | -0.22 | | 2.31 | Aqueous | Torimura et al. ¹¹¹ |
| Texas Red | -0.88 | | 2.08 | Aqueous | Torimura et al. ¹¹¹ |
| MR121 | -0.26 ^b | 1.64 ^b | 1.90 | Acetonitrile | Heinlein et al. ⁹⁸ |
| R6G | -0.71 ^b | 1.63 ^b | 2.27 | Acetonitrile | Heinlein et al. ⁹⁸ |
| Cy5 | -0.64 ^b | 1.06 ^b | 1.88 | Acetonitrile | Heinlein et al. ⁹⁸ |
| 7-methoxycoumarin (C-3H) | -1.62 | 2.04 | 3.48 | Acetonitrile | Seidel et al. ³⁶ |
| 7-methoxycoumarin-3-carboxylic acid ethyl ester (C-3Es) | -1.16 | 2.20 | 3.35 | Acetonitrile | Seidel et al. ³⁶ |
| 3-cyano-7-methoxycoumarin (C-3CN) | -1.00 | 2.30 | 3.26 | Acetonitrile | Seidel et al. ³⁶ |
| Carbostyryl-124 (Ca-124) | -2.24 | 1.19 | 3.30 | DMF | Seidel et al. ³⁶ |
| Coumarin-120 (C-120) | -1.9 | 1.38 | 3.22 | DMF | Seidel et al. ³⁶ |
| RhodamineB (RhB) | -0.56 ^c | 1.18 ^c | 2.19 | Acetonitrile | Fries, J.R. ¹¹² |

E_{ox} and E_{red} are the reduction potentials of the oxidized electron donor and the electron acceptor, $E_{0,0} = (E_{abs} - E_{em})/2$, is the energy of the singlet-singlet transition of the fluorophore and NHE: normal hydrogen electrode. ^b Potentials measured against saturated calomel electrode (SCE) and converted to NHE by adding 0.24 V. ^c Potentials measured against Ag/AgCl and converted to NHE by adding 0.20 V. Dye structures in Fig. 1.4.

The main mechanism for fluorescence quenching is photoinduced electron transfer (PET) where an electron transfer occurs between an electron donor and an electron acceptor when they are within the van der Waals radius. The energetics of PET are explained

further in chapter 2.1 (d). Usually, the quencher is an amino acid or nucleobase and the electron acceptor is the organic fluorophore. Among nucleobases, guanine is the most prominent PET quencher due its low E_{ox} and thus high propensity to donate an electron^{36,113}, while in proteins, it is tryptophan¹⁰⁴. The sequence of quenching efficiency for the nucleobases was found by Seidel et.al.³⁶. If the nucleobases are reduced, the sequence of the quenching efficiency was found to be $G < A < C < T \leq U$; and the sequence reversed when the nucleobases are oxidized. A summary of redox properties of common dyes and nucleobases is given in Tables 1.1 and 1.2.

PET quenching can occur transiently through molecular collisions (dynamic quenching), or by ground-state complex formation (static quenching). Since the ground state complexes formed by static quenching are non-fluorescent or very weakly fluorescent, the predominantly observed fluorescence is from the uncomplexed fluorophore, leaving the fluorescence lifetime usually unperturbed^{104,114}. For this reason, although dynamic and static quenching reduce fluorescence intensity, most often, only the dynamic quenching reduces fluorescence lifetime.

Table 1.2. Redox properties of nucleoside/amino acid quenchers measured vs NHE

| Nucleoside/ amino acid | E_{red} [V] | E_{ox} [V] | Solvent | Reference |
|---------------------------|---------------|--------------|--------------|------------------------------|
| dG | -2.43 | 1.48 | Acetonitrile | Seidel et al. ³⁶ |
| dT | -2.18 | 2.11 | Acetonitrile | Seidel et al. ³⁶ |
| dC | -2.35 | 2.14 | Acetonitrile | Seidel et al. ³⁶ |
| dA | -2.52 | 1.96 | Acetonitrile | Seidel et al. ³⁶ |
| dU | -2.07 | ≥ 2.39 | Acetonitrile | Seidel et al. ³⁶ |
| Tryptophan | | 0.81;0.64 | Aqueous | Doose et al. ¹⁰⁴ |
| Tyrosine | | 0.7 | Aqueous | Doose et al. ¹⁰⁴ |
| <i>p</i> -Cresol | <-2.16 | 1.64 | Acetonitrile | Seidel et al. ¹¹⁵ |

E_{ox} and E_{red} are the reduction potentials of the oxidized electron donor and the electron acceptor, NHE: normal hydrogen electrode, ^b Potentials measured against saturated calomel electrode (SCE) and converted to NHE by adding 0.24 V.

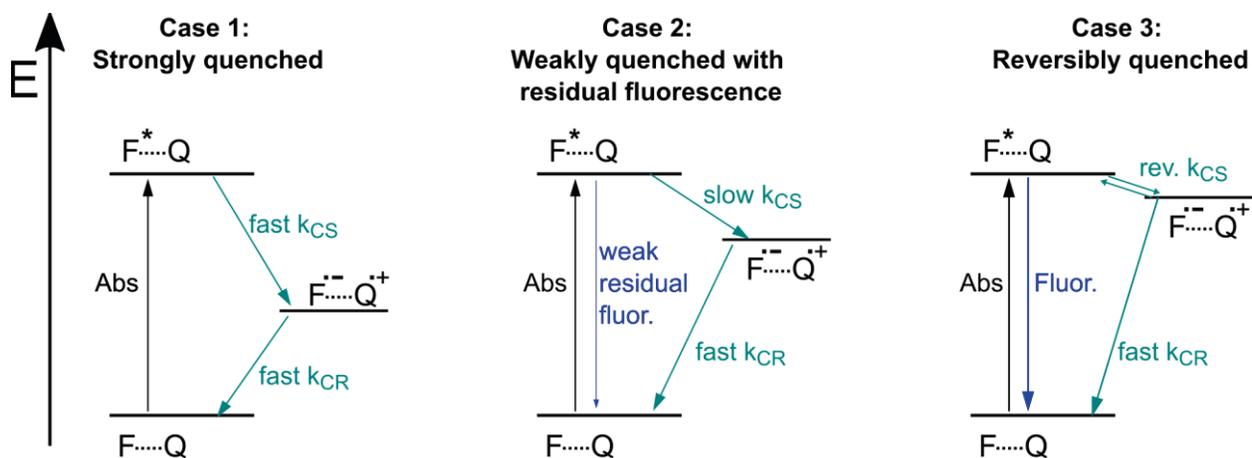
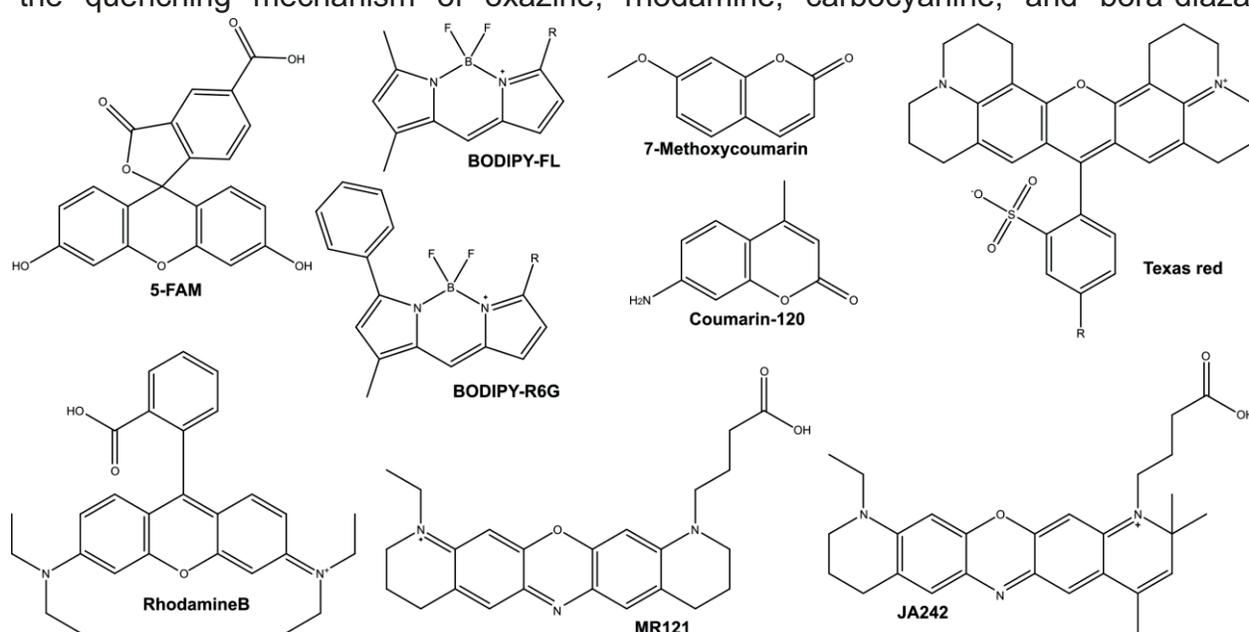


Figure 1.3. Fluorescence of charge transfer complexes. The ground state complex $[F \cdots Q]$ on absorption gets excited to $[F^* \cdots Q]$, the excited ground state complex. This can form the charge transfer complex $[F^{\cdot-} \cdots Q^{\cdot+}]$ by charge separation with a rate constant k_{CS} . The CT-complex then decays non-radiatively by charge recombination with a rate constant k_{CR} . Case 1: The CT complex is fully quenched due to fast k_{CS} and k_{CR} . Case 2: Weak residual fluorescence results from slow k_{CS} . Case 3: Little to no fluorescence quenching. CT-complex formation is reversible (indicated as rev. k_{CS}) with fluorescence lifetime is enhanced.

The fluorescence properties of a statically quenched dye depend upon the rate constant of formation of the charge transfer (CT) complex from the excited state ground state complex (Figure 1.3). **Case 1:** The CT-complex is very stable and the rate constant of charge separation (k_{CS}) is high. In this case, the backward reaction from the CT-complex to the excited ground state complex is highly unlikely. The energy of the CT-complex is then dissipated non-radiatively via charge recombination (CR) to the ground state complex, resulting in strong quenching with the formation of a dark complex, eg. R6G-guanine conjugate. Polar environments favor the stabilization of the charge separated state and, in turn, favor dark complexes in static quenching. This prompt fluorescence quenching prevents us from observing the ground state complex in the fluorescence spectra. However, since the electronic structure of the free dye is different from that of the dye-quencher ground state complex, the complex is readily observable by absorption spectroscopy. **Case 2:** The rate constant of charge separation and thereby the formation of the CT-complex is much slower, resulting in weak quenching. The excited state ground state complex can undergo a weak residual fluorescence¹⁰⁴ (k_{res}) resulting in a short fluorescence lifetime component of approx. hundreds of ps. The energy of the CT-complex is not high enough for the reversible transition. **Case 3:** When the dye is a poor

electron acceptor, eg. C-120-adenine conjugate, the energy of CT-complex is very close to or even higher than that of the excited ground state complex. k_{CS} is very slow and there is a high probability for the CT-complex to return back to the excited ground state, resulting in reversible quenching. When the excited state then emits by fluorescence, the lifetime is enhanced. Such cases with no quenching or reversible quenching occurs more frequently in a non-polar solvent.

For the different dyes, the contribution of static and dynamic quenching depends on the dye-quencher combination. Chen et al. identified four quenchers for AL488: tryptophan (Trp), tyrosine (Tyr), histidine (His), and methionine (Met)¹⁰⁸. While tryptophan quenches by dynamic and static pathways, PET quenching by His is mostly static. In another study, steady-state and time-resolved fluorescence measurements were performed to elucidate the quenching mechanism of oxazine, rhodamine, carbocyanine, and bora-diaza-



indacene dyes by amino acids. ATTO 655, ATTO 680, and the oxazine derivative MR 121 are quenched almost exclusively by tryptophan by static quenching; while rhodamine, fluorescein, and bora-diaza-indacene derivatives are quenched by tyrosine residues also. For the two bora-diaza-indacene dyes, BODIPY-FL and BODIPY630/650, dynamic quenching predominates. Meanwhile, the quenching of carbocyanine dyes by PET were found to be negligible¹¹⁶.

Figure 1.4. Structures of fluorophores discussed. 5-FAM¹¹¹, BODIPY-FL¹¹¹, BODIPY-R6G¹¹¹, 7-methoxycoumarin³⁶, Coumarin-120³⁶, Texas red¹¹¹, Rhodamine B¹¹², MR121¹⁰⁴ and JA242⁹⁸.

Quenching via PET is also controlled by the steric access of the dye to the quenching moiety. For the BODIPY-R6G probe, it has been shown that, while the free dye is strongly quenched by guanine, the DNA-labeled dye only exhibits slight quenching upon hybridization to a complementary strand containing guanine due to the steric hindrance of the bulky phenyl groups present in BODIPY-R6G¹¹¹. For the rhodamine dye Alexa546 the presence of the negatively charged sulfonates and the bulky chlorine atoms in the phenyl group prevent the dye from interacting with the nucleobases at a distance short enough for electron transfer to be efficient¹⁰⁰. It was seen that in a DNA-hairpin, the dye MR121 adopts an end-capped conformation because of its symmetrical planar structure, resulting in efficient PET. In contrast, the quenching efficiency for the label JA242 is much lower due to the three additional methyl groups which impede coplanar stacking in an endcapped conformation⁹⁸. This feature is useful when studying folding-unfolding transitions in proteins. Neuweiler et al. investigated folding dynamics of the 20-residue mini-protein Trp-cage where a single Trp residue is buried in a small hydrophobic core and shielded from the solvent¹¹⁷. When MR121 dye is labeled on the surface of the protein, in the folded state of the protein, the dye is shielded from quenching interactions with the “caged” Trp residue in the hydrophobic core, whereas in the denatured state, Trp is solvent-exposed and quenches MR121 fluorescence.

Chen et al.¹⁰⁷ showed that the PET interaction between guanosine residues in DNA and Tetramethyl Rhodamine dye resulted in quenching when these guanosine residues were close to the terminal of the DNA where the dye is attached. Quenching was seen to be higher when the G residues were in the complementary strand near the dye. Ohya et al.¹¹⁸ studied the effect of duplex formation on eosin and tetramethylrhodamine attached to the 5'-terminal of DNA oligos. They also came to the conclusion that guanine residues on the complementary DNA strand near the dyes had a major influence on the quenching. Similar results were obtained by Berger¹¹⁹ who studied guanine quenching of 5' terminal Rh6G label coupled to dsDNA by a C6 linker with the guanine at different locations. At the terminus, the quenching was higher when guanine was on the complementary strand compared to the same strand as the label. Guanines at 2 and 3 bases from the terminal label had drastically lower quenching comparatively. The effective distance of guanine

quenching was up to 3 base pairs from the label. When a shorter linker was used, the dye was seen to be quenched more by the guanine on the strand to which it was attached.

In addition to nucleobases like guanine and amino acids like Trp, quenching can also be caused by extrinsically introduced quenchers like Dabcyl, BHQ-1 and QSY-7.¹²⁰ It was seen that when both a reporter and a quencher dye were present, static quenching, involving a reporter–quencher ground-state complex, occurs. Here, FRET and static quenching frequently occur together. Marras et al.¹²¹ proved that even though spectral overlap between the emission spectrum of the fluorophore and the absorption spectrum of the quencher is a prerequisite for efficient FRET, it is not a significant factor determining PET quenching efficiency. As long as the fluorophore and the quencher come within van der Waals contact distance of each other, quenching efficiencies stay very high even in the absence of spectral overlap.

While PET generally requires van der Waals contact, a long-distance quenching effect has also been reported^{98,113,122}. It has been postulated that electron-transfer reactions do not require intimate molecular contact and may well occur over larger distances. The charge transfer may occur along the nucleic acid helix¹²³⁻¹²⁵ or the proteins¹²⁶ which act as bridges through which the electrons of the fluorophore can be transferred^{127,128}. The assumption is that a molecular band is created due to electronic interactions in the biomolecule where the electronic states are delocalized over the length of the molecule and aid in conduction¹²⁹.

In synthetic DNA hairpins, GG or GGG sequences have been shown to be very shallow hole traps and reversible hole transport between a guanine donor and such sequences ensues.¹¹³ Additionally, energy studies showed that electron transfer across peptides are influenced by the magnitude and direction of the dipole generated by the peptide secondary structure. Thus the nature of the secondary structure of the peptide also plays an active role in PET¹³⁰.

In addition to PET, fluorescence quenching by proton-coupled electron transfer (PCET) mechanism³⁶ is also noteworthy. Here, electron and proton transfers occur either in a concerted manner, or sequentially, and results in a stable complex. Non-radiative transition from this state contributes to the overall quenching of the fluorophore.

d) Fluorescence enhancement

For cyanine dyes such as Cy3, Dy547, Cy5, Alexa647 or Thiazole orange^{103,131-134} interactions with biomolecules can instead lead to fluorescence enhancement. The common feature among all these dyes is that they are composed of two rings interconnected by carbon–carbon double bonds that undergo cis–trans isomerization¹³⁵. The cause of the enhancement in fluorescence has been shown to be a decrease in the efficiency of photoisomerization, and it occurs in conditions where the dye is sterically constrained. This phenomenon is known in literature under different names: when the nucleic acid-coupled dye undergoes fluorescent enhancement on interaction with a protein, it was named “Protein-induced fluorescence enhancement” (PIFE)^{135,136}. When the dye interaction occurs with a nucleic acid, it was called “Nucleic acid induced fluorescence enhancement” (NAIFE)⁹⁰. For the specific case of RNA, the name “RNA-induced fluorescence enhancement” (RIFE)⁹⁰ was also proposed.

In addition to steric restriction, fluorescence enhancement in cyanine dyes have been demonstrated by a completely different mechanism that occurs in the presence of D₂O medium¹³⁷. When the dyes are measured in H₂O or protonated alcohols¹³⁸, fluorescence quenching can occur due to efficient energy transfer (FRET) from the excited state of the dye to the harmonics and combination bands of OH vibrational modes of water. The solvent-assisted quenching rate constant in H₂O for the Cy5 dye is $1.7 \times 10^8 \text{ s}^{-1}$. Therefore, replacing protonated solvents by deuterated ones enhances the dye fluorescence. It should be noted that the extent of enhancement depends upon the dye’s accessibility to water molecules. If the dye is in an environment where it is shielded from the solvent, the enhancement is expected to be minimal.

Fluorescent enhancement on steric restriction was first observed for the fluorophore Cy3 labeled on ssDNA close to the binding site of *E. coli* SSB protein^{132,139}. Another good example for PIFE is the case of the Cy3 labels attached at the 3’ and 5’ of a ssDNA⁴³. Here, it was seen that a difference in the steric hindrance induced by the bound hRPA protein at the 3’ and 5’ ends causes a difference in fluorescence lifetimes in the corresponding labels. The different photophysical responses at either ends of the same helix showed that the protein environment close to the cyanine dye can modulate the rate constant of cis-trans photoisomerization and thus its fluorescence properties.

The attachment of a bulky substituent such as DNA can also produce a decrease in the isomerization rate constant, and a concomitant increase in the fluorescence quantum yield (QY) and lifetime⁴². Studies show that steric constraints for the dye occur not only due to protein binding, but also due to stacking with the helix end, in a nick, gap or overhang region in a nucleic acid¹⁴⁰, giving rise to a phenomenon called stacking-induced fluorescence increase (SIFI). The fluorescence enhancement in all these cases are due to association of the dye or dye-labeled nucleic acid with a biomolecule. Therefore, in this thesis, the term association-induced fluorescence enhancement (AIFE) is coined, encompassing PIFE, NAIFE, RIFE and SIFI, where the fluorescence enhancement is caused due to dye interaction with the biomolecule.

To elucidate the photophysical mechanism behind PIFE¹³⁶, Stennett et al. used complexes of DNA and Klenow fragment and a combination of time-resolved fluorescence and transient spectroscopy techniques. They proved that the enhancement of fluorescence correlates with a decrease in the efficiency of photoisomerization, and occurs when the dye is sterically constrained. Widengren and Schwille¹⁴¹ attempted to further explain the kinetics of this process by using fluorescence correlation spectroscopy (FCS) to study the fluorescence fluctuations of Cy5 under different excitation conditions. They successfully arrived at the rate constants of the isomerization processes (forward and back-isomerization) and the triplet transitions.

There are several factors influencing the AIFE mechanism like temperature, viscosity of medium, dye linker properties and geometrical features of the nucleic acid, the nucleobase sequence and the state of nucleic acid hybridization.

Temperature: The fluorescent properties of Cyanine dyes attached to a nucleic acid are temperature dependent because photoisomerization is an activated process that competes with radiative decay from the excited state^{131,142}. As a result, increasing the temperature favors photoisomerization, thereby causing a drop in fluorescence^{42,143,144}. The activation energy for photoisomerization can be obtained from¹⁴⁵:

$$\ln[\phi_f^{-1}(T) - \phi_{f,max}^{-1}] = \ln A - E_a/RT \quad (1.1)$$

where ϕ_f is the temperature-dependent fluorescence quantum yield of Cy3, $\phi_{f,max}$ is the maximum fluorescence efficiency that would be measured in the absence of

photoisomerization, A is a pre-exponential Arrhenius factor that depends on the microscopic friction and the radiative lifetime of fluorescence, and E_a represents the activation energy for isomerization. The effect of temperature must, however, be taken in conjunction with other factors like hybridization and melting. For instance, when Cy3 was labeled to a DNA, it was seen that with an increase in temperature, the dsDNA melted to the ssDNA and the fluorescence QY increased by 40%¹⁴⁶.

Viscosity: Isomerization depends strongly on the viscosity of the medium^{42,134,143,147}. At higher viscosity the isomerization is hindered, increasing the energy barrier for the non-radiative decay pathway and thus the quantum yield. The dependence of the rate constant of isomerization (k_{ISO}) on the viscosity is given by:

$$k_{ISO} = f(\eta) \exp\left(-E_0/RT\right), \quad (1.2)$$

where η is the solvent viscosity coefficient and E_0 was found to be solvent independent¹⁴³.

Dye linker: The dye linker influences the dye behavior by (1) determining the accessibility (reach) of the dye and (2) modulating the dye wobbling motion within the constraints of the nucleic acid (NA) groove. In the former case, the linker lengths affect the fluorescent properties of cyanine dyes by allowing or preventing stacking at the helix end^{42,140,146} due to its accessibility. In the latter case, the fluorescence anisotropies are altered by virtue of the length and flexibility of a linker. Sindbert et al.¹⁰³ studied Cy5 dyes with linkers of different lengths and rigidity. In RNA, if the linker is stiff and short, the wobbling motions within the groove are negligible because the major groove of the RNA is very deep and narrow. The major groove of DNA is wider, and thus even a small linker wobbling motion is observed. For a highly flexible linker, the linker wobbling is limited by the opening angle of the groove. Linker lengths can also affect dye-dye interactions. For a multiply labeled oligonucleotide, the Cy3-Cy3 interactions were seen to be favored at higher labeling densities when longer linkers are used, as opposed to short linkers¹⁴⁸.

Nucleic acid geometry: It was seen that although Cy3 interacts more strongly with the purine bases, such interactions require some NA flexibility for the dye to be able to adopt a proper stacked conformation with a nucleobase. Hence the rigid structure of poly(dA) and other oligonucleotides containing repeats of purines can prevent strong Cy3 DNA interactions¹⁴². The geometry of dye-NA interaction is different depending on the type of

NA⁵⁰. X ray crystal structures of Cyanine dyes on RNA showed that the long axis of the fluorophore is approximately parallel to that of the terminal base pair⁹². The fluorophores were shown to be stacked on the terminal base pair such that their indole nitrogen atoms lie on the major groove side. On DNA, however, NMR studies of Cy5 show the fluorophore rotated 180° about its long axis, so that the indole nitrogen atoms are on the minor groove side of the terminal base pair⁹¹. Another effect of nucleic acid geometry arises from the dependence of fluorescence lifetime on persistence length. Due to the longer persistence length of RNA, slightly lower lifetimes were obtained for sCy3–RNA oligonucleotides compared to the corresponding DNA conjugates⁹⁰. The orientation of the protein with respect to the DNA-labeled Cy3 was also seen to affect the magnitude of the fluorescence enhancement⁴³.

Nucleobase sequence: The fluorescence of cyanine dyes covalently attached to DNA exhibits sequence-dependent properties as established by different experimental and simulation - based studies. The fluorescence signal and lifetime of the free Cy3 dye was found to increase by addition of the nucleobases; the changes being more dramatic in the case of the purines, dGMP and dAMP¹⁴⁵. In later studies, similar observations were made for Cy3 covalently attached to the 5' terminus of DNA duplexes where the extent of stacking and hence the fluorescent properties of Cy3 was influenced by the identity of the terminal base pair^{94,142}, with the purines (dGMP, dAMP) exhibiting a sharper increase in fluorescence efficiency and lifetime than the pyrimidines (dTMP, dCMP). This is because the stacking interactions are stronger for the purines due to their more extended conjugated system. Asymmetrical cyanine dyes like Thiazole orange (TO) behave similar to symmetrical cyanines¹³³ in that these dyes also have a lower affinity for pyrimidines than purines.

Molecular dynamics simulations, characterizing the effect of the terminal base pair on the conformational dynamics of Cy3 attached to the 5' end of a duplex DNA, showed that Cy3 has significantly less affinity for the TA base pair than for AT, GC and CG¹⁴⁹. In a different study, the dynamics of the dyes, Cy3 and Cy5 attached to either C or G at the 3' or 5' ends of an RNA duplex¹⁵⁰ was investigated. 3' attached dyes were shown to explore more space in contrast to 5' attached dyes that often stacked on the ends of the RNA duplex. On the 3' end, Cy5 showed conformational freedom only when attached to cytosine.

These studies demonstrate the sensitivity of AIFE to the immediate environment on nucleic acids.

Hybridization: Hybridization to the target DNA significantly affects the fluorescence intensity of the probes. For singly labeled DNA, the quantum yield of Cy3 was shown to decrease upon going from ssDNA to dsDNA^{42,146}. Morten et al. reported that a covalently attached Cy3 dye undergoes strong enhancement of fluorescence intensity and lifetime upon going from the ssDNA to the annealed hairpin¹⁴⁰. The results of hybridization studies, though seemingly contradictory, could be explained by considering the linker lengths which were much shorter in the previous studies^{42,146} potentially preventing helix-end stacking. In a study involving a series of multiply labeled fluorescent oligonucleotides, the change in fluorescence upon duplex formation ranged from -35% to $+20\%$ ¹⁴⁸. This increase or decrease of fluorescence upon duplex formation depends upon several factors like labeling density, fluorophore charge and linker length. Probes that are labeled at low density, have short linkers, or are labeled with Cy3NOS (as opposed to Cy3), exhibited less fluorescence when hybridized.

The fluorescence properties of a covalently attached Cy3 dye was studied in a DNA hairpin for the influence of a nearby nick, gap or single-strand overhang¹⁴⁰. The dye was shown to undergo strong enhancement of fluorescence intensity and lifetime when stacked in a nick, gap or overhang region in the duplex region of a DNA hairpin because the dye interaction with the nick or gap increases the stability of the closed (hybridized) conformation of the hairpin. An interesting observation of nucleic acid systems undergoing AIFE was that a localized shielding of their negative charges promote interactions with the fluorophore. Thus cations like Magnesium(II) sequester some negative charges of the phosphate backbone and thus favor dye-nucleic acid interactions⁹⁰.

Although we describe fluorescence enhancement in all these cases, quenching is also possible by a PIFE-like mechanism, called PIFQ (protein-induced fluorescence quenching). This is dependent on the initial state of DNA-Dye complex. For example, if the dye in the initial state is more restricted than in the final state, fluorescence quenching results. For instance, Rashid et al.¹⁴⁷ showed that if an oligo has considerable secondary structure or strong specific interactions that restrict the dye's photoisomerization in the

initial state which is subsequently disrupted in the presence of ssDNA-binding protein RPA or a complementary strand, there is a relative increase in the dye photoisomerization which decreases the fluorescence relative to the initial state.

1.1.3 Application: Fluorescent labels as reporters of biomolecular environment.

(The text in this section 1.1.3 was written in conjugation with Dr. Anders Barth, and was originally drafted as part of a planned review article.)

Different aspects of the environment-sensitivity of fluorophores have been exploited to report on the features of the biomolecule. A few of them are discussed below.

a) Utilizing the environment sensitivity of fluorescence quenching

Fluorescence quenching of coupled dyes by nucleic acids has found widespread application in sensing and as a reporter for structural dynamics. The large structural changes upon hybridization of single strands or melting of duplexes have a profound effect on the photophysical properties of attached fluorophores. The presence of guanines in single-stranded nucleic acids results in strong dynamic quenching by PET if they are in reach of the fluorophore. In this case, hybridization results in a recovery of fluorescence due to the reduced accessibility of guanines in the duplex¹¹. On the contrary, if guanines are present on the complementary strand of the duplex, hybridization may result in a reduction of fluorescence¹¹. Commonly, experiments are designed such that the fluorophore is initially quenched and lights up upon hybridization¹⁰. An example for such fluorogenic probes are molecular beacons¹⁵¹ that consist of a stem-loop structure that places the terminally attached fluorophore close to a quencher on the other end of the strand. Upon binding to the target sequence, the fluorophore is separated from the quencher and the fluorescence is recovered.

Instead of introducing a specific quencher molecule, fluorogenic probes based on DNA-hairpins using terminal guanines as quenchers have been widely applied. These “smart probes”¹⁵² have the advantage that a double functionalization of the nucleic acid is not needed. A detailed characterization of the quenching of the oxazine dye MR121 in such stem-loops has been performed by Heinlein et al.⁹⁸. The study revealed a nearly completely quenched ground-state complex and showed that the quenching efficiency is maximized through the addition of short single-stranded overhangs and the use of the

guanine derivative 7-deazaguanosine as a better electron donor. Higher sensitivity for smart probes in the nanomolar range was achieved by single molecule burst analysis measurements on a confocal microscope using the oxazine dye JA242¹⁵². Torimura et al.¹¹¹ measured the reduction of fluorescence for the dyes 5-FAM, BODIPY-FL and TAMRA upon hybridization to a guanine-containing complementary strand and used this effect for the detection of specific DNA sequences¹¹⁰. Fluorescence quenching by guanines has also been applied to monitor polymerase chain reaction (PCR) assays. Using primers labeled with the dye fluorescein, the progress of the PCR could be monitored in real-time through the reduction of fluorescence upon binding to guanine-rich amplicons¹⁵³. In a different approach, self-quenched fluorescently labeled primers have been used for quantitative and multiplexed detection¹⁵⁴. Like smart probes or molecular beacons, the primers were designed as hairpins that place the fluorophores FAM and JOE close to a guanine, allowing the reaction to be followed by the increase in fluorescence upon formation of the PCR product.

Quenching by photoinduced electron transfer has also been used as a reporter for the structural dynamics of nucleic acids. Using ensemble measurements, the kinetics of the hairpin ribozyme have been monitored through the quenching of a fluorescein label on the 3' end of the substrate by a specific guanine on the 5' end of the ribozyme⁴⁴. Early single-molecule studies using burst analysis of freely diffusing molecules in solution have investigated the quenching of TMR linked to dsDNA by nearby guanines^{155,156}. The detection of populations of single molecules with different fluorescence lifetimes allowed to assign the multiple fluorescence lifetime components observed in ensemble experiments to distinct conformational states. The large modulation of the fluorescence intensity provided by PET quenching makes it ideal for use with fluorescence correlation spectroscopy (FCS), enabling the detection of dynamic processes on the sub-millisecond timescale. Pioneering applications of this approach have studied the PET quenching of TMR by guanines on dsDNA^{155,156} and of R6G by dGTP in solution¹⁵⁷. Technical advances in time-correlated single photon counting electronics have since enabled the measurement of photon-by-photon correlation functions over twelve orders of magnitude from picoseconds to minutes with picosecond timing resolution^{158,159}. PET-FCS has been used to study the initial steps of DNA hairpin folding¹⁶⁰, the structural dynamics of

mRNA¹⁶¹, the chain dynamics of unstructured single-strand overhangs¹⁶² and the kinetics of base flipping in dsDNA due to mismatches¹⁶³.

In proteins, fluorescence quenching is most notably made use of in Tryptophan-Induced Quenching (TriQ), a technique which exploits the PET quenching effect of tryptophan on the emission of certain fluorophores¹⁶⁴⁻¹⁶⁷. TriQ requires van der Waals contact for efficient quenching¹⁰⁴ and therefore can be used as a reporter for monitoring conformational dynamics in proteins, for interactions in the range ~5–15 Å. A commonly used probe for the TriQ approach is bimane, but dyes like BODIPY (boron-dipyrromethene), lucifer yellow, and ATTO-655 can also be utilized. The TriQ-bimane approach has been widely used to study¹⁶⁷ conformational changes in GPCRs like β -adrenergic receptors^{168,169}, the interaction of rhodopsin with transducin¹⁷⁰, etc. It has recently been shown that bimane fluorescence is also quenched by tyrosine in a distance-dependent manner and the phenomenon named tyrosine-induced quenching (TyrIQ)¹⁷¹.

Another similar application based on the distance-dependent modulation of fluorescence is called “Distance-encoding photoinduced electron transfer” (DEPET)¹⁷². DEPET is particularly suited for the precise determination of protein structure and conformational changes since it directly provides the distance between protein backbone and side-chain atoms. DEPET has been proven to be effective in determining how membrane depolarization changes intramolecular distances and side-chain orientations in the human large-conductance potassium (BK) channel, where cellular excitability is regulated. DEPET is also accurate enough to measure the length of rigid polyproline peptides in solution.

The reversibility of PET quenching also finds applications in antibody recognition. The reversible fluorescence quenching of oxazine dyes by tryptophan in short peptides has been used to identify the presence of p53-autoantibodies¹⁷³. In the free peptide, a ground-state complex is formed between the oxazine dye and the tryptophan residue, leading to static quenching. On binding to the antibody, a spontaneous conformational change is initiated in the peptide, resulting in a strong increase in fluorescence intensity¹¹⁶.

b) Interpreting the environment from fluorescence enhancement

The first instance of AIFE in Cy5 DNA-labels was reported for *E. coli* SSB binding to Cy3-labeled ssDNA¹³⁹. In a FRET experiment with Cy3 (donor) and Cy5 (acceptor) on the 3' and 5' ends, respectively, of a ssDNA wrapped on SSB tetramer, it was seen that although a significant FRET signal occurred, a corresponding decrease in Cy3 fluorescence was not observed because the fluorescence decrease by FRET is counteracted by the AIFE effect¹³⁹. This paved the way to study the diffusion and chemomechanical pushing of SSB proteins along ssDNA^{174,175}. In addition to SSB protein, binding of other motor proteins to ssDNA^{176,177} were also studied. It was also possible to follow the dynamics of DNA and RNA polymerases^{178,179} and detect the motion of helicases on nucleic acids at the single-molecule level¹⁸⁰.

The phenomenon of AIFE allowed monitoring of protein binding to a fluorescently labeled nucleic acid without having to label the protein. It was thus possible to quantify the protein binding constant, examine substrate specificity and explore binding kinetics¹³⁵. For example, RNA-binding of the protein RIG-I (Retinoic acid inducible–gene I) was studied by Myong et al.¹⁸⁰ using a dsRNA labeled with the fluorophore DY547. Markiewicz et al.¹⁸¹ used the fluorescence enhancement of Cy3-labeled DNA to characterize the association and dissociation kinetics of the binary complex formed between the Klenow fragment and DNA. The effect of monovalent ions and the effect of incorrect NTPs (nucleotide triphosphates) on the KF binding dynamics were also examined.

The main advantage of AIFE, when compared to FRET, arises from its insensitivity to acceptor bleaching or blinking. Stacking induced fluorescence enhancement (SIFI) is an alternative approach to FRET for studying DNA dynamics, provided there is dye-stacking interaction^{140,182}, and has been used to probe hybridization dynamics of a DNA hairpin at the single-molecule level.

There are other areas of protein study where AIFE has important contributions. The filament formation of the Rad51 gene, involved in DNA repair, was probed by AIFE using Cy3 attached to either ends of a ssDNA. The directionality of filament formation in this case was revealed to be in the 5' to 3' direction¹⁸³. A special case of interaction in a dye-NA system is the class of RNA aptamer –fluorophore systems like the RNA

mango/spinach aptamer-thiazole orange systems^{184,185}. These are RNA oligonucleotides that enhance the fluorescence of a dye like Thiazole orange up to approx. 1000 fold when non covalently bound to it. These aptamers are designed to bind to a specific target. Hence such systems are used to simultaneously track and visualize biologically important RNAs.

c) Dye labels in other notable fluorescence-based techniques

There are many popular biological techniques where fluorescent labels are at the core of experiment design. These techniques report on the structure and dynamics of the biomolecules, quantification of cellular processes or assist in drug development and clinical medicine.

Förster Resonance Energy Transfer (FRET) has been used in a wide spectrum of biological studies¹⁸⁶ not only because the accessible range of distances in FRET fall within the order of bio-macromolecular distances, but also because FRET allows the study of molecules in dynamic motion.

FRET has been mainly used to study protein structure and conformational changes¹⁸⁷. For example, FRET studies showed that binding of the 15.5K protein to U4 snRNA kink-turn induces kink-turn folding that results in a tightly kinked conformation in the RNA¹⁸⁸. In the bacteriophage T4 lysozyme, in addition to the open and closed states that exchange at 4 μ s in solution, a new minor state was disclosed by FRET studies³. FRET studies also discovered that free syntaxin exists in a dynamic equilibrium between a closed (inactive) state and an open (active) configuration with a relaxation time of 0.8 ms¹⁸⁹. Another example where FRET monitored the conformational dynamics was that of the extracellular ligand-binding domain of the metabotropic glutamate receptor (mGluR), which exists in a preformed equilibrium of conformational states. The presence of ligands tunes the conformational dynamics by modulating the transition rate between the states, and not by stabilizing the individual conformations¹⁹⁰.

Some of the other common applications of FRET are: (1) *Visualizing protein-protein interactions*, for example, between SNARE protein syntaxin-1A (Syn-1A) and sulfonyleurea receptors¹⁹¹ or the association between actin-binding protein cortactin and potassium channel Kv1.2¹⁹². (2) *Decoding the stoichiometry of protein complexes* like the

rod CNG channels, which was found to be composed of three CNGB1 and one CNGA1 subunit¹⁹³ (3) *As biosensors*¹⁹⁴. For example, the Src biosensor - when the substrate in this biosensor is phosphorylated by endogenous Src kinase after chemical or mechanical stimulation, the FRET efficiency changes such that we can quantify Src kinase activity in cells¹⁹⁵. Such FRET biosensors measure either a change in distance¹⁹⁶ or in fluorescence properties¹⁹⁷ of the labels.

Fluorescence in-situ hybridization (FISH) is a highly sensitive and specific macromolecule recognition tool that uses DNA fragments incorporated with fluorophore-coupled nucleotides to examine the presence of complementary sequences in fixed cells or tissues which are then visualized through a fluorescence microscope or an imaging system¹⁹⁸. This technique finds applications in the genetic diagnosis of constitutional common aneuploidies¹⁹⁹, microdeletion/microduplication syndromes²⁰⁰, and subtelomeric rearrangements²⁰¹ and to detect infectious microbes and parasites in human blood cells²⁰².

DNA points accumulation in nanoscale topography (DNA-PAINT) is a super-resolution technique where biological processes below the classical diffraction limit of light are visualized. In DNA-PAINT, the transient binding of short, dye-labeled oligonucleotide (imager) to its target-bound complementary (docking) strand is utilized to create molecular 'blinking' to enable stochastic super-resolution microscopy²⁰³⁻²⁰⁵.

Stochastic optical reconstruction microscopy (STORM) is another super-resolution imaging technique that utilizes single-molecule localization microscopy (SMLM) method. Here the continuous photo-switching of single fluorophores and their position determination are utilized to achieve super-resolved images²⁰⁶. This technique helped reveal information about the structures and functions of sub-cellular organelles with a spatial resolution at the molecular level²⁰⁷.

1.2 Motivation and objectives of present study

The entirety of the thesis deals with the position-dependent photophysical behavior of fluorescent dyes under different biomolecular environments. The main objectives of the thesis are (1) to investigate the environment-dependent dye behavior on biomolecules and thus employ the fluorophores to probe the local bio-environment and (2) to apply this knowledge to improve FRET experiments.

To reveal the distinct photophysical signatures of different micro-environments, the commonly used organic fluorophore, Alexa488 (a green fluorescent Rhodamine dye) and the red-emitting cyanine dyes, Alexa647 and Cy5, are employed. These fluorophores are commercially available, widely-used in FRET studies, have unique redox properties (Table 1.1) and exhibit distinctive photophysical behavior due to their differing propensities to isomerism and fluorescence quenching.

In the following passages, I describe how the main objectives of the thesis were explored. Consistent with the organization of the chapters in this thesis, I discuss the research problem in 3 parts. In **chapter 3**, the position-dependent fluorescent properties of dyes on RNA environments are explored; in **chapter 4**, the environment-sensitive dye properties in protein environments are discussed; and, in **chapter 5**, dye models for FRET label description are reviewed and further developed. Below, I define the questions raised, the experimental techniques used and offer a brief synopsis of the results achieved.

(1) Sensing the local environment and global conformation of RNA by flexibly coupled fluorescent labels

In this work (**chapter 3**) we try to answer the following questions:

- How do the dyes behave in different nucleic acid environments, namely, RNA duplex, three-way and four-way junction molecules?
- Do dyes undergo π -stacking interactions with the nucleobases at the helix ends or junctions?
- What are the contributions of dynamic and static PET quenching for the Alexa488 labels at different environments?

- What is the role of labeling chemistry and the corresponding distinct local environments on the photophysical behavior of Cy5 labels?
- Can we predict the photophysical properties based on dye label positions?
- How are the dyes distributed and oriented at distinctive positions on the RNA?
- What are the major RNA-dye interaction modes?
- What is the effect of RNA primary sequence on the dye photophysical behavior?
- How can we use the fluorophores as local and global environment sensors?
- How does the environmental dependence of dye properties affect FRET studies?

In order to answer these questions, we employ a range of experimental and simulation techniques. (i) Ensemble time resolved lifetime studies are used to study the quenching behavior, (ii) Ensemble time resolved anisotropy studies are employed to identify sterical constraints (iii) Spectral shifts from absorption and steady state emission measurements help indicate the nature and extent of dye-biomolecular interactions (iv) FCS studies are utilized to quantitatively evaluate quenching and cis-trans isomerism processes (v) Fluorescence quantum yield measurements are adopted to explore the mechanisms of PET quenching and reveal the extents of static/dynamic quenching. (vi) Accessible volume (AV) simulations assess the sterically accessible volume of the fluorophore and provide accurate insights into potential interactions and steric clashes (vii) Molecular dynamics (MD) simulations are employed to obtain atomistic insights into potential interaction modes, dye position distribution and orientations.

In summary, the fluorescence properties of Alexa488 and Cy5 in 104 different environments on RNA duplexes, three-way and four-way junctions are studied and their distinct photophysical signatures for the different RNA micro-environments unveiled. The results are corroborated by AV and MD simulations to provide structural insights on potential interaction modes. By comparing the spectroscopic signatures of fluorescent labels in the different RNA molecules and measuring the rotational diffusion of RNA using end stacked Cy5 dyes, I show that fluorescence spectroscopy offers the required sensitivity to elucidate the structure and dynamic properties of RNA. Lastly, I discuss the implications of these results for FRET studies. The insights obtained from the structural

and photophysical interactions of flexibly coupled fluorescent dyes with nucleic acids will improve the design and interpretation of future studies.

(2) Fluorescent dyes probe the local environment of protein, T4 Lysozyme: Exploring with experiments and simulations

In this section (**chapter 4**), I studied the protein, T4 lysozyme (T4L) as a model system to address the following questions :

- How do Alexa488 and Alexa647 dyes behave in different protein environments?
- Can we explain the source of dye quenching and steric restriction?
- Can we predict probable dye-protein interactions from the changes in fluorescence parameters?
- What are the main dye-protein interaction modes on an atomistic scale?
- Can we use the fluorophores as environmental probes and detect biomolecular processes?

With the help of experimental techniques like time resolved lifetime and anisotropy measurements, spectral studies, FCS and fluorescence quantum yield measurements, combined with AV and MD simulations, the aforementioned questions are tackled (similar to those used in chapter 3).

In summary, I use the bacteriophage T4 Lysozyme (T4L) as a prototype system for exploring the influence of the local dye environment in different protein environments using the commonly used fluorophores, Alexa488 and Alexa647. The experimental output from fluorescence spectroscopy is combined with computational simulations to give a detailed insight into the origin, type and extent of dye-protein interactions under various environments. The amino acid residues responsible for dye trapping and quenching, the effect of dye trapping on fluorescence quenching, the extent of static quenching and H-bonding interactions and the interference to dye motion from protein dynamics are presented. Finally, specific examples are shown where the coupled dyes act as sensors and report on changes in the local and global protein environment.

(3) Developing dye models for accurate FRET label description:

Förster resonance energy transfer (FRET) is a powerful technique used to study the conformational dynamics of biomolecules by employing covalently attached donor and

acceptor dyes. In order to arrive at accurate structural information from interdye distances, spatial distribution of dye positions must be included in the dye model description. In this chapter, we explore the following:

- How to describe dye spatial distribution in dye models used for FRET studies?
- How prior experimental knowledge can help in improving dye model description?

In summary, **chapter 5** reviews the different dye models available to describe the spatial distribution of dyes covalently attached to biomolecules. Prior knowledge of the environmental sensitivity of fluorophores on different RNA environments, especially those interactions which significantly alter the dye distribution, are incorporated into the label description for FRET dyes. A proof of concept study is carried out by comparing FRET donor-acceptor (DA) distances from experiments and simulations on a double stranded RNA (dsRNA) as a model system. When there are significant deviations between measured and calculated distances, the prior knowledge of position-dependent dye behavior is applied to modify the description of the dye distribution. For end-stacked dyes, the dye model is improved by introducing the concept of a hybrid accessible volume.

Chapter 2. Theoretical background

This chapter describes some of the basic concepts of fluorescence relevant to this thesis and the theory behind frequently employed fluorescence techniques.

2.1 Basics of fluorescence

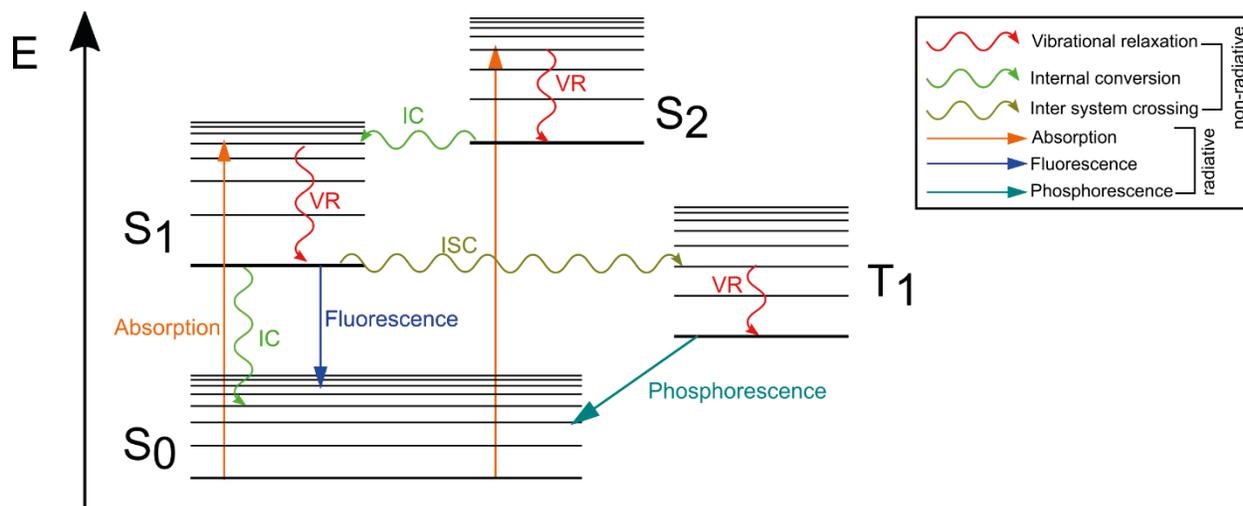


Figure 2.1. A simple Jablonski diagram depicting the processes that occur to a molecule on absorption of light. Wavy arrows indicate non-radiative transitions, while straight arrows indicate radiative transitions. S_0 = singlet ground state, S_1 = first singlet excited state, S_2 = second singlet excited state, T_1 = first triplet excited state, VR=vibrational relaxation, IC=internal conversion, ISC=intersystem crossing

A Jablonski diagram²⁰⁸ (Fig. 2.1) describes the processes that occur after irradiation of a molecule by light. Most molecules, at room temperature, are in the lowest vibrational level of the singlet electronic ground state (S_0). When a molecule absorbs a photon, it absorbs energy and gets excited from the vibrational ground state of S_0 to a higher vibrational state of S_1 . After excitation, the molecule relaxes through one of several competing pathways. These relaxation processes are classified into two categories, radiative and non-radiative processes.

a) Radiative and non-radiative processes

When the relaxation process is accompanied by the emission of a photon, it is called a radiative relaxation process. There are two types of radiative transition depending on the nature of the excited state: (1) Fluorescence: The spontaneous emission of a photon from an excited singlet state to a lower electronic state is termed fluorescence. The wavelength

of the emitted photon corresponds to the energy gap between the levels. Since the transition occurs from the lowest vibrational level of the excited state (Kasha's rule¹¹⁴), the wavelength of emitted light is independent of the excitation wavelength. Additionally, the Franck-Condon principle states that all electronic transitions occur without change in the position of the nuclei. As a result, if a transition between particular states has the highest probability in absorption, the reciprocal transition has the highest probability in emission (mirror image rule). (2) Phosphorescence: The radiative transition from the first triplet state, T_1 , to the singlet ground state, S_0 is called phosphorescence. Since this transition requires a change in multiplicity, it is spin-forbidden, but occurs due to spin-orbital coupling.

When a molecule in the excited state undergoes deactivation via a pathway where no photon is emitted, it is called a non-radiative relaxation. There are several non-radiative processes possible. (1) Internal Conversion (IC) which is the isoenergetic transition between two electronic states of the same multiplicity. (2) Intersystem Crossing (ISC), the isoenergetic transition between two electronic states of different multiplicities. (3) Vibrational relaxation (VR), which occurs when the molecule is in the higher vibrational levels of a given electronic state. Since the molecule is not in the vibrational ground state, the excess energy is dissipated within picoseconds to the environment through collisions or redistributed inside the molecule and the molecule relaxes to the lowest vibrational level of the electronic state. (4) Other non-radiative processes include energy transfer processes like FRET or PET and photochemical reactions.

b) Fluorescence lifetime and quantum yield

Fluorescence lifetime (τ) is the average time the molecule spends in the excited state prior to return to the ground state¹¹⁴. This is an important characteristic of fluorescence because it determines the time available for the fluorophore to interact with its environment in the excited state and therefore reveals valuable information about the dye environment. The relaxation from the excited state may occur through radiative or non-radiative processes with rate constants k_{rad} and k_{nr} , respectively, such that $\tau = \frac{1}{k_{rad} + k_{nr}}$.

In the absence of non-radiative processes, $\tau_n = \frac{1}{k_{rad}}$, which is called the intrinsic or natural lifetime.

Another important characteristic of fluorescence is the fluorescence quantum yield (Φ_F), (QY) defined as the number of photons emitted during fluorescence relative to the number of photons absorbed. Molecules with high quantum yields approach unity. For example, Rhodamine6G in ethanol has an impressive quantum yield of 0.95. Assuming that fluorescence is the only radiative process and has a rate constant k_F , the fluorescence quantum yield may be described as

$$\Phi_F = \frac{k_F}{k_F + k_{nr}}. \quad (2.1)$$

As the rate constant of non-radiative processes increase, a decrease in fluorescence quantum yield ensues. This event is called quenching and may occur upon intermolecular collisions (dynamic quenching) or by the formation of a ground state complex (static quenching). In static quenching, the quenched state persists on timescales much longer than the fluorescence lifetime, such that no change in the lifetime is observed, although the average quantum yield decreases.

c) Fluorescence anisotropy

Before irradiation, fluorophores are randomly oriented in solution. However, when interacting with polarized light, those fluorophores with transition moments oriented parallel to the electric field vectors of the photons are preferentially absorbed (photoselection). The resulting emission also occurs with the light polarized along a fixed axis in the fluorophore. If I_V and I_H are the fluorescence intensities of the vertically (V) and horizontally (H) polarized emission with respect to the excitation light, the anisotropy, r is given as:

$$r = \frac{I_V - I_H}{I_V + 2I_H}. \quad (2.2)$$

For fluorophores where the transition dipole moments are displaced by an angle β between the dipoles of absorption and emission, the fundamental or intrinsic anisotropy (r_0) is given as:

$$r_0 = 0.4 \cdot \frac{(3 \cos^2 \beta) - 1}{2} \quad (2.3)$$

For fluorophores with colinear transition dipole moments of absorption and emission, the fundamental anisotropy, $r_0 = 0.4$. For quantitative fluorescence measurements where it

is preferable to eliminate polarization effects, one makes use of magic angle conditions, namely, vertically polarized excitation and an emission polarizer oriented 54.7° from the vertical. Under these conditions, where $\beta = 54.7$, the intrinsic anisotropy $r_0 = 0$.

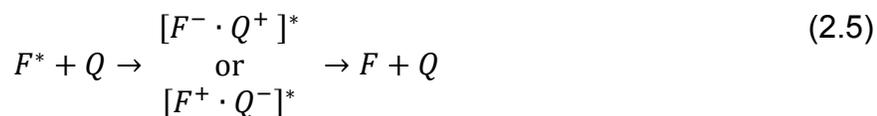
Rotational diffusion that occurs during the lifetime of the excited state displaces the emission dipole of the fluorophore such that the anisotropy values are lower than the fundamental anisotropy, r_0 . As a result, in non-viscous solutions, small organic fluorophores typically have anisotropies approaching zero. However, when the fluorophore is bound to a macromolecule, rotational diffusion is restricted, thereby increasing the anisotropy. For spherical rotors where τ is the fluorescence lifetime and ρ is the rotational correlation time, the Perrin equation gives:

$$\frac{r_0}{r} = 1 + \tau/\rho \quad (2.4)$$

Thus fluorescence anisotropy provides information about the microenvironment of the fluorophore, thereby helping the study of the biomolecular environment and interactions.

d) Fluorescence quenching by photoinduced electron transfer

For the systems described in this thesis, photoinduced electron transfer (PET) is the major non-radiative pathway for fluorescence quenching. PET is a distance-dependent fluorescence quenching process between a fluorophore and a quenching moiety requiring van der Waals contact. In PET, an electron is transferred between the quencher and the organic fluorophore. Depending on the electrochemical potentials, the excited fluorophore may hereby act as the electron donor or acceptor. PET involves the formation of a charge-transfer complex according to the reaction:



The energetics of PET quenching are described by the Rehm-Weller equation^{36,104,209}:

$$\Delta G_{CS} = E_{ox} - E_{red} - E_{0,0} + \Delta G^0(\epsilon), \quad (2.6)$$

where E_{ox} and E_{red} are the reduction potentials of the oxidized electron donor and the electron acceptor, $E_{0,0}$ is the energy of the singlet-singlet transition of the fluorophore and $\Delta G^0(\epsilon)$ describes the coulombic attraction of the ion pair that depends on the dielectric constant (ϵ) of the solvent²¹⁰.

e) Association-induced fluorescence enhancement

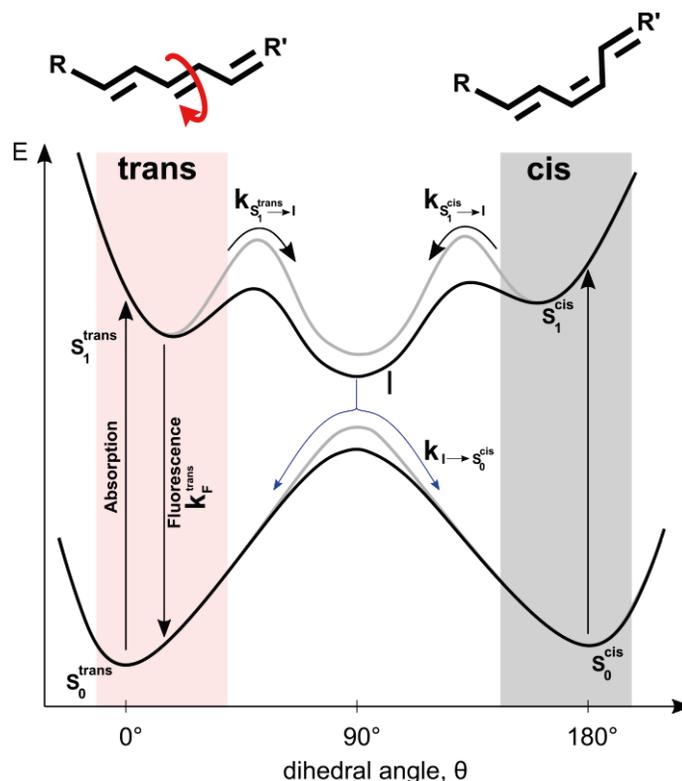


Figure 2.2. Potential energy diagram of photoisomerization as a function of the bond angle, θ , of the central conjugated double bond. The structures of the fluorescent *trans* and non-fluorescent *cis* isomers are shown above. The energy levels in the absence of steric hindrance are shown as black curves, while the modulated energy levels due to increased steric hindrance are shown in gray.

For fluorophores containing an extended π -system, like the cyanine dyes, the fluorescence intensity is seen to increase in a sterically constrained environment. This phenomenon is called association-induced fluorescence enhancement (AIFE). A simple energy landscape diagram to explain the photophysical mechanism of AIFE is illustrated in Fig. 2.2. The ground state of the π -bond-containing fluorophore is in an all-*trans* conformation and this electronic state is denoted as S_0^{trans} . Upon absorption of light, the molecule is excited to the S_1^{trans} state. Relaxation from this state can occur either (1) radiatively through fluorescence with a rate constant, k_F^{trans} or (2) via a twisted intermediate (I) formed by rotation ($\theta = 90^\circ$) around the central C-C double bond (red arrow) with rate constant $k_{S_1^{trans} \rightarrow I}$. Relaxation from the twisted intermediate may occur either back to the *trans* ground state, S_0^{trans} , or to the *cis* photoisomer S_0^{cis} . Since the *trans* configuration is energetically favorable, the population of the *cis* photoisomer will slowly

convert back into the trans state thermally. Back-conversion of the cis isomer S_0^{cis} can also occur after re-excitation into the S_1^{cis} and relaxation over the twisted intermediate via $S_0^{cis} \rightarrow S_1^{cis} \rightarrow I \rightarrow S_1^{trans}$. Steric hindrance increases the energy barrier to the twisted intermediate state, thereby blocking the non-radiative relaxation from the S_1^{trans} state by photoisomerization and enhancing the fluorescence. This is the reason for the fluorescence enhancement.

2.2 Time resolved fluorescence measurement

Fluorescence measurements are generally of two types: (1) Steady state measurements, performed with constant illumination and observation and (2) time-resolved measurements where the sample is exposed to pulsed light and the resulting intensity decay recorded with a high-speed detection system¹¹⁴. Steady state measurements can be considered as an average of the time-resolved phenomena over the intensity decay¹¹⁴. For a fluorophore with a single exponential intensity and anisotropy decay, the intensity at time t , $I(t)$ and the anisotropy at time t , $r(t)$ are given as:

$$I(t) = I_0 e^{-t/\tau} \quad (2.7)$$

$$r(t) = r_0 e^{-t/\rho} \quad (2.8)$$

where I_0 and r_0 are the intensities and anisotropies at $t = 0$, immediately following the excitation pulse, respectively, and τ and ρ are the fluorescence lifetime and rotational correlation time, respectively. Time-resolved measurements are important because a significant amount of molecular information is lost during the time averaging process in steady state measurements¹¹⁴.

Time-correlated single photon counting (TCSPC) is the general method used to measure time resolved lifetime and anisotropies. Here, a pulsed laser source is employed to excite the sample and a “time-stamping” principle¹⁵ is adopted where each detected photon is assigned 3 tags: channel number, microtime and macrotime (Fig. 2.3A). While the channel number of a given photon gives information on its spectral region and polarization, the microtime (t_i for the i^{th} photon) identifies the time elapsed between the excitation pulse and the photon detection, and the macrotime (N_i for the i^{th} photon) gives the number of excitation pulses that have occurred since the start of the measurement

until photon detection. From these tags and the time interval (Δt_L) between two consecutive laser pulses, the exact time of photon detection can be determined:

$$\text{Time of detection of } i^{\text{th}} \text{ photon} = t_i + N_i \Delta t_L \quad (2.9)$$

Coupling this information with that regarding the number of detected photons, the time traces of the fluorescent intensity for each channel can be generated ¹⁵.

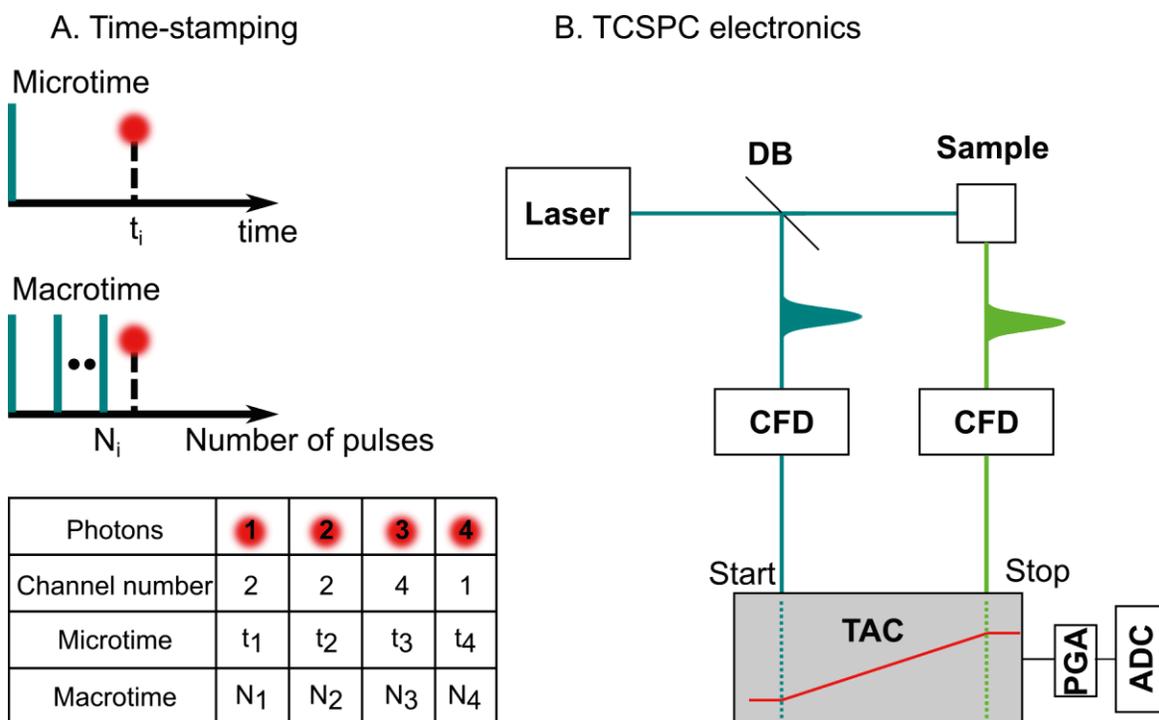


Figure 2.3. TCSPC basic concepts. (A) Time stamping by assigning 3 fluorescent tags to a detected photon - microtime, macrotime and channel number. (B) Basic TCSPC electronics where DB:dichroic beamsplitter, CFD: constant fraction discriminator, TAC: time-to-amplitude-converter, PGA: programmable gain amplifier and ADC: analog-to-digital-converter.

The basic electronics for measuring TCSPC¹¹⁴ (Fig. 2.3B) consists of a laser excitation source, a dichroic beam splitter (DB) to split the beam into two channels, a constant fraction discriminator (CFD) to measure the arrival time of the pulse, a time-to-amplitude-converter (TAC) to generate a voltage ramp that increases linearly with time, a second CFD to record the pulse from the detected photon, voltage-amplifying programmable gain amplifier (PGA) and an analog-to-digital-converter (ADC) to convert the voltage to a numerical value. In addition to the sample, an instrument response function (IRF) is recorded using a scattering solution (LUDOX) to correspond to a zero-lifetime sample.

This IRF is accounted for in the fit function via convolution and the model decay is given by:

$$F(t) = f_0 \cdot IRF(t - t_{\text{shift}}) \otimes f(t) + s \cdot IRF(t) + bg, \quad (2.10)$$

where f_0 is the amplitude scaling factor, $IRF(t - t_{\text{shift}})$ is the instrument response function shifted by the time t_{shift} , s is the scatter amplitude and bg is the constant background level. The “ \otimes ” sign designates circular convolution.

The fluorescence decay $f(t)$ is described by

$$f(t) = \sum_i x_i e^{-t/\tau_i}, \quad (2.11)$$

where τ_i and x_i are the fluorescence lifetimes and corresponding species fractions with $\sum_i x_i = 1$. The species weighted average lifetime may then be defined by:

$$\langle \tau \rangle_x = \sum_i x_i \tau_i \quad (2.12)$$

The anisotropy decay is described by:

$$r(t) = \sum_i r_i e^{-t/\rho_i}, \quad (2.13)$$

where ρ_i and r_i are the rotational correlation times and respective amplitudes.

2.3 Fluorescence correlation spectroscopy (FCS)

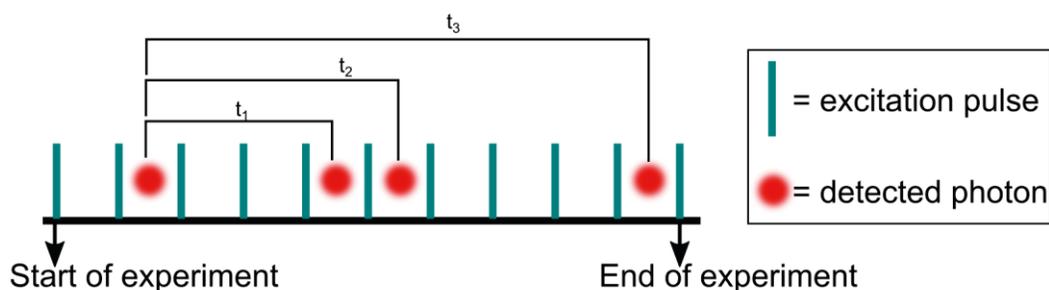
Fluorescence Correlation Spectroscopy (FCS) is a powerful statistical method that extracts information from fluctuations in fluorescence intensity of a freely diffusing molecule as it diffuses through a small detection volume^{19,21}. The core of this method is the calculation of the autocorrelation function of the fluorescence fluctuations. Essentially, the overlap between a fluorescence signal $F(t)$ with itself, after a time delay t_c , given by $F(t + t_c)$ is determined (Fig 2.4A). The autocorrelation function is thus given by¹⁵:

$$G(t_c) = \frac{\langle F(t) F(t + t_c) \rangle}{\langle F \rangle^2} \quad (2.14)$$

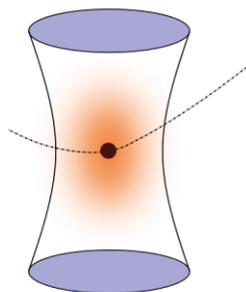
where F is the fluorescence intensity. To ensure that only the photons emitted by an individual molecule are correlated, a confocal setup where the detection volume is of the

order of a few femtoliters and a pinhole to reject signal from outside the desired detection volume are employed. The detection volume is approximated as a 3D Gaussian ellipsoid (Fig 2.4B). Since the autocorrelation amplitude is inversely proportional to the number of particles, we can determine the concentration and molecular brightness. A model function of the form given by eq. (2.15) is applied to fit the autocorrelation. The FCS model function can have separate terms for independent processes; each process distinguishable as long as they have sufficiently different characteristic correlation times. We can thus determine translational diffusion constant, triplet state lifetime and amplitude, the rotational diffusion constant and the excited state lifetime ²⁰.

A. Interphoton time for autocorrelation



B. Molecule diffusing through the confocal volume



C. Typical autocorrelation curve

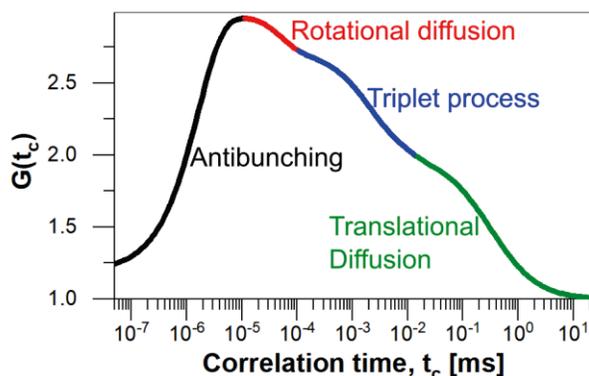


Figure 2.4. Basic concepts of FCS (A) Scheme describing interphoton time (t_1 , t_2 etc.) used in autocorrelation. Excitation pulses (vertical blue lines) occur in regular intervals and the detected photons are shown in red. (B) A freely diffusing molecule shown passing through the confocal detection volume. (C) Typical correlation curve showing a few of the processes detectable by FCS and their correlation time window.

For a sample with freely diffusing molecules (diffusion time t_{diff}), undergoing triplet kinetics (with time t_T and fraction of molecules in the triplet state, A_T) and quenching (with time t_Q and fraction of molecules in the quenched state, A_Q), with photon antibunching

described by time t_{Ab} and amplitude A_{Ab} , and rotational diffusion by depolarization time t_R and amplitude A_R , the fit model is given by the equation:

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{bright}}} \cdot \frac{1}{1 + \frac{t}{t_{\text{diff}}}} \cdot \frac{1}{\sqrt{1 + \frac{t}{\left(\frac{z_0}{\omega_0}\right)^2 \cdot t_{\text{diff}}}}} \quad (2.15)$$

$$\cdot \left(1 - A_T + A_T \cdot e^{-\frac{t}{t_T}} - A_Q + A_Q \cdot e^{-\frac{t}{t_Q}} - A_{Ab} \cdot e^{-\frac{t}{t_{Ab}}}\right)$$

$$\cdot \left(1 + A_R \cdot e^{-\frac{t}{t_R}}\right).$$

where ω_0 and z_0 are the $1/e^2$ radii of the 3D Gaussian shaped observation volume in the lateral and in axial dimension, respectively, and N_{bright} is the effective number of molecules in the bright state $N_{\text{bright}} = N_{\text{total}} \cdot (1 - T - Q)$, with N_{total} referring to the total number of molecules.

2.4 Förster resonance energy transfer (FRET)

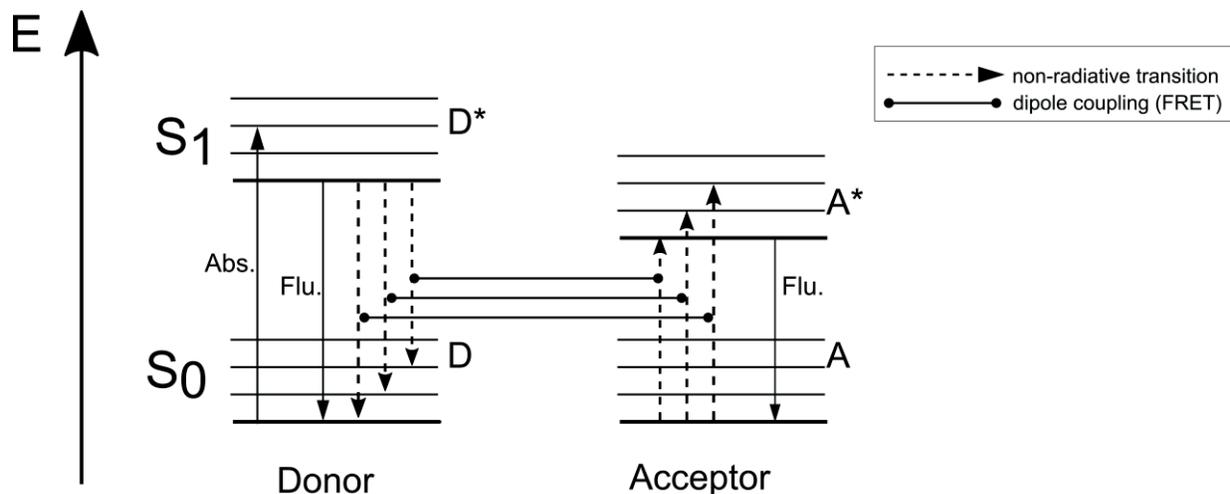


Figure 2.5. A Jablonski diagram depicting FRET between a donor and acceptor fluorophore. Abs.: absorption, Flu.: fluorescence, S₀, S₁: singlet ground and first excited state, D: donor ground state, D*: donor first excited state, A: acceptor ground state, A*: acceptor first excited state. The dipole-coupled transfer occur between non-radiative transitions of equal energies in the donor and acceptor systems and lead to FRET.

Förster resonance energy transfer (FRET) is a distance dependent phenomenon used to study structure and conformation changes in biomolecules^{17,18} employing two suitable

dyes, the donor and the acceptor. The rate constant of the energy transfer by FRET, k_{FRET} , is highly dependent on the distance (R_{DA}) between the two fluorophores as:

$$k_{FRET} = \frac{1}{\tau_D} \left(\frac{R_0}{R_{DA}} \right)^6 \quad (2.16)$$

where τ_D is the fluorescence lifetime of the donor, R_0 is the Förster radius and R_{DA} is the donor-acceptor distance. The efficiency of energy transfer between these dyes is related to the distance between them as:

$$E = \frac{1}{1 + R_{DA}^6/R_0^6} \quad (2.17)$$

where

$$R_0 = 0.2108 \sqrt[6]{\left(\frac{\Phi_{F,D} \kappa^2}{n^4} \right) J} \quad (2.18)$$

and

$$J = \int_0^\infty \bar{F}_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \quad (2.19)$$

Here, E corresponds to the FRET efficiency, $\Phi_{F,D}$ is the donor fluorescence quantum yield and n denotes the refractive index of the medium. The Förster radius, R_0 , denotes the distance for which the FRET efficiency is exactly 50%. This radius is characteristic to a given dye pair, the FRET between these two dyes being most sensitive to changes in distance in this region. For FRET to occur, there must be spectral overlap between the donor's fluorescence emission spectrum and the acceptor's absorption spectrum. J is the corresponding spectral overlap integral (generally in units $M^{-1}cm^{-1}nm^4$). \bar{F}_D is the normalized spectral radiant intensity of the excited donor [nm^{-1}] and $\varepsilon_A(\lambda)$ is the extinction coefficient of acceptor [$M^{-1} cm^{-1}$].

κ^2 is the dipole orientation factor which depends on the relative orientation of dipoles and is given by¹⁵:

$$\kappa^2 = (\sin \theta_D \sin \theta_A \cos \varphi - 2 \cos \theta_D \cos \theta_A)^2 \quad (2.20)$$

where the angles θ_D and θ_A define the dipole orientations with respect to the unitary interdye distance vector and φ corresponds to the angle between the two planes defined

by the transition dipole moment of each dye and the interdye distance vector. κ^2 can reach a maximum value of 4 when the dipole moments are colinear and a minimum value of 0 when the transition dipole moments of the donor and acceptor dyes are perpendicular. Due to rotational diffusion, the transition dipole moments change direction resulting in a continuously changing value of κ^2 ¹⁵. Therefore, for an isotropically diffusing pair of dyes, $\kappa^2 = 2/3$ ⁴⁹. In FRET experiments, this value is often adopted, under the assumption that the dyes are freely diffusing within the constraints of the linkers that tether them to the biomolecule.

Chapter 3. Sensing the local environment and global conformation of RNA by flexibly coupled fluorescent labels

Authors: Aiswaria Prakash¹, Christian A. Hanke^{1*}, Oleg Opanasyuk^{1*}, Olga Doroshenko¹, Ralf Kühnemuth¹, Noah Salama¹, Bekir Bulat¹, Anders Barth^{1,2†}, Claus A.M. Seidel^{1†}

Affiliations: ¹Chair for Molecular Physical Chemistry, Heinrich Heine University, Düsseldorf, Germany

²Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands.

* Equal contributions as second authors

†Corresponding authors with correspondence to: c.seidel@hhu.de, a.barth@hhu.de

Contributions: A.P., O.D., A.B. and C.A.M.S. designed the research. A.B. and C.A.M.S. supervised the project. A.P., O.D. and N.S. hybridized RNA molecules and performed time-resolved measurements. A.P. and B.B. performed steady state spectral measurements. A.P. performed FCS measurements. A.P., R.K. and N.S. performed pFCS measurements. A.P., R.K. and C.A.M.S. deciphered pFCS results with respect to cis-trans isomerism. C.A.H. performed and analyzed MD simulations and identified interaction modes. O.O. provided the models for analyzing time-resolved measurements, performed cluster analysis of MD simulations and computed theoretical rotational correlation functions. A.P. performed AV modeling of dye distributions and AV-related calculations. A.P. designed figures 3.1-3.4 and 3.6. A.P. and C.A.H. designed Fig. 3.5. A.P. and A.B. wrote the manuscript in consultation with all authors.

3.1 Abstract

Covalently attached fluorescent probes are excellent reporters of biomolecular structure and dynamics due to their high sensitivity to the biomolecular environment. However, to relate a variation of the photophysical parameters to structural changes of the host biomolecule, a detailed understanding of the fluorophore photophysics in different environments is required. Here, we studied the fluorescence properties of two commonly used organic fluorophores, the rhodamine derivative Alexa488 and the cyanine dye Cy5, in 104 different environments on RNA duplexes, three-way and four-way junctions. For both dyes, surface interactions are detected by an increase of the fluorescence anisotropy. In addition, Cy5 shows association-induced fluorescence enhancement, while Alexa488 is site-specifically quenched by guanine bases through photoinduced electron transfer. These effects lead to distinct photophysical signatures of different RNA micro-environments. Notably, end-stacking interactions are detected for labelling positions up to 7 bp away from the helix ends. The results are corroborated by accessible volume and molecular dynamics simulations to provide structural insights on potential interaction modes. By comparing the spectroscopic signatures of fluorescent labels in the different RNA molecules and measuring the rotational diffusion of RNA using end stacked Cy5 dyes, we show that fluorescence spectroscopy offers the required sensitivity to elucidate the dynamic tertiary structure of RNA. Lastly, we discuss the implications of our results for FRET studies. The presented results provide novel insights into structural and photophysical interactions of flexibly coupled fluorescent dyes with nucleic acids, which will improve the design and interpretation of future studies.

3.2 Introduction

Fluorescently labeled oligonucleotides are widely used in the life sciences in diagnostic assays¹⁰⁻¹², fluorescence imaging of live or fixed cells⁶⁻⁸, super-resolved fluorescence microscopy²¹¹ and in-vitro mechanistic studies of proteins^{1,2}, nucleic acids^{4,5} or DNA origami nanostructures^{212,213}. In these assays, fluorescent dyes are covalently linked to the nucleic acid backbone or bases at the termini or at internal positions along the oligonucleotide. Molecular interactions, ligand binding or conformational changes are then translated into changes of the fluorescence signal (i.e., the molecular brightness), or

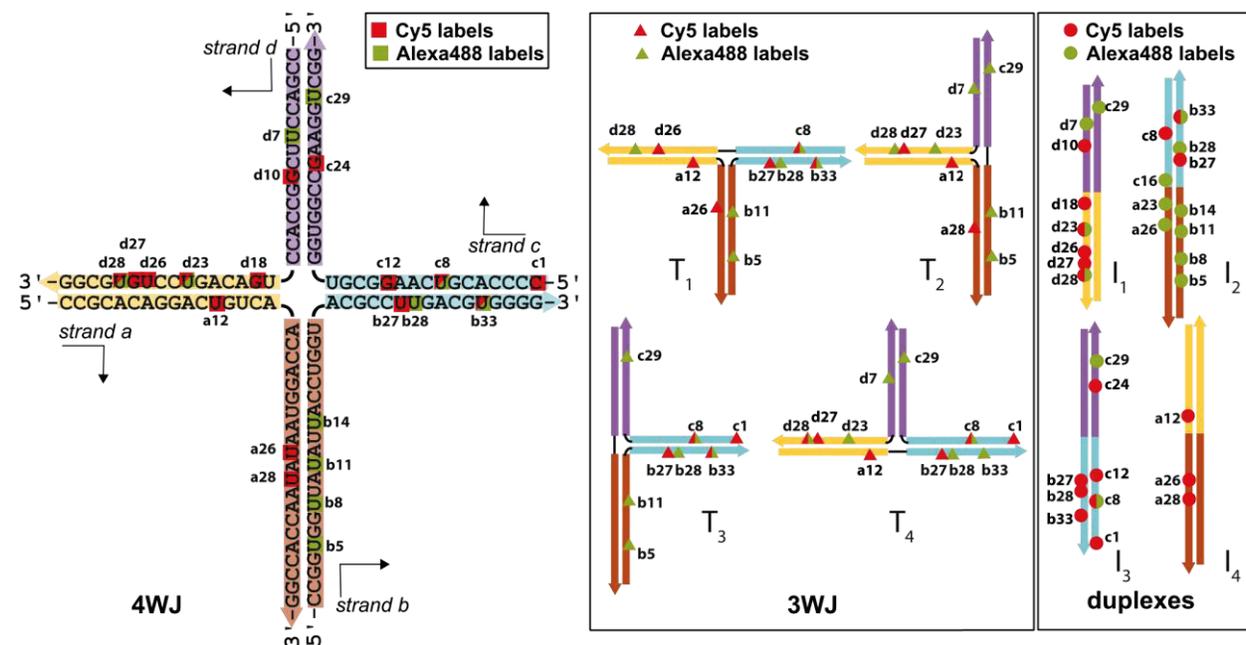
other photophysical parameters such as the fluorescence lifetime, anisotropy, or the absorption and emission spectrum. A prime example are so-called “molecular beacons”, short hairpin DNA structures with a quencher and fluorophore on the opposing ends that light up when binding to the complementary sequence by spatially separating the fluorophore-quencher pair^{8,12,151}. A detailed understanding of how the photophysical properties are affected by the molecular environment is crucial for the interpretation of experimental results and the design of improved fluorescent probes with higher sensitivity and specificity.

Fluorophores are commonly attached via long flexible linkers of ~2 nm length that enable them to interact with the biomolecular surface in the vicinity of the attachment site in a variety of ways. For nucleic acids, the propensity of the cyanine dyes Cy3/Cy5 and the dye rhodamine 6G to stack to the end of the double helix via π - π interactions has been studied in detail^{92,94-96}. Much less is known, however, about potential interactions of labels attached at internal positions with the backbone, the major and minor grooves, or with magnesium ions bound to the nucleic acid backbone^{90,214}. Upon interaction with the host molecule, the fluorescence signal is often reduced or enhanced. Rhodamine and oxazine dyes are quenched specifically via photon-induced electron transfer (PET) by the electronegative base guanine (or, in proteins, by the amino acid tryptophan)^{104,163}. If the guanine base is fully accessible, usually a π - π ground state complex is formed in which the excited state of the fluorophore is quenched on the picosecond time scale^{99,215}. However, PET may also occur in other geometries where the electron orbitals only partly overlap, or over longer lengths through a bridge of conjugated π -systems¹²⁷ or even a DNA helix^{124,216}. The high efficiency of the quenching process makes PET an ideal reporter for fast (nanosecond) conformational dynamics of structured²¹⁷ or intrinsically disordered^{218,219} proteins and nucleic acids^{161,162} when combined with fluorescence correlation spectroscopy (FCS). Cyanine dyes, on the other hand, show an increase in fluorescence when they are sterically restricted, an effect that was first utilized to detect the binding of proteins to nucleic acids and termed protein-induced fluorescence enhancement (PIFE)¹³⁵, even though it is not exclusive to protein binding. PIFE has since been widely applied to various systems^{141,147,182,220,221}. The correlation between steric restriction and molecular brightness is explained by a modulation of the

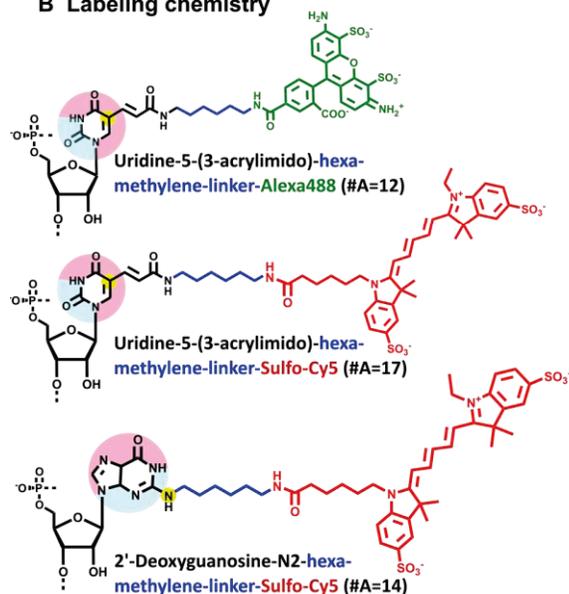
photoisomerization between the fluorescent trans and effectively non-fluorescent cis configurations^{136,222}. Blocking the rotation around the central conjugated polymethine chain increases the brightness of cyanine dyes by hindering the non-radiative relaxation from the first excited state through a 90° twisted intermediate, as well as by reducing the steady-state population of the cis state through photo-conversion. While the sensitivity of cyanine dyes to their environment can be used as a proximity reporter on the molecular scale, the synthesis of bridged cyanine dyes with optimal brightness and minimal environmental sensitivity has been reported for Cy3^{223,224} and recently also Cy5²²⁵⁻²²⁷.

The five intrinsic fluorescence properties are the spectral properties of absorption and fluorescence, $F(\lambda_A, \lambda_F)$, the molecular brightness, the quantum yield, Φ_F , and the fluorescence lifetime, τ , and anisotropy, r ¹⁶. These dimensions of the fluorescence signal offer a vast amount of information on the fluorophore, its immediate environment, and the host biomolecule, as they are sensitive to changes of the polarity of the environment, size of the solvent shell, presence of quenching moieties, steric restriction of the fluorophore, and the global rotation of the host biomolecule. Moreover, fluorescence correlation spectroscopy provides information on processes that modulate the fluorescence intensity, such as quenching, fluorescence enhancement, or molecular rotations when polarized detection is employed. With a detailed understanding of the behavior of fluorescent dyes on nucleic acids currently missing, it remains challenging to predict potential interactions of the fluorophore with the nucleic acids and their effect on the fluorescence properties. Decoupling the effect of local interactions on the fluorescence properties from the global dynamics of the biomolecule is of particular importance for understanding the structural dynamics of nucleic acids with complex folds, such as DNA Holliday junctions or catalytic RNA ribozymes, for which the biological function is intricately linked to their dynamic structure.

A Sequence and labeling positions in RNA molecules



B Labeling chemistry



C Labeling site in the grooves

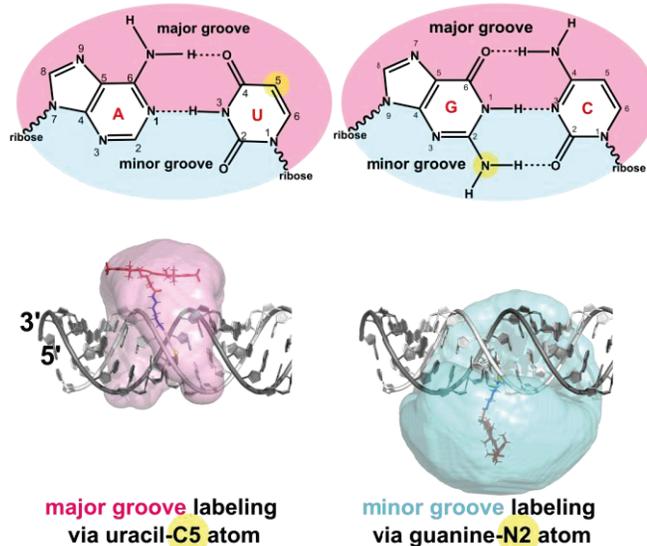


Figure 3.1: Design of fluorescently labeled RNA molecules. **A)** The four-way junction (4WJ) RNA molecule is comprised of the four strands a, b, c, and d. The red and green symbols represent the labeling sites for the dyes Cy5 and Alexa488, respectively. The derived three-way junction (3WJ) and duplex RNA molecules are schematically depicted on the right. The color of the fragments indicates the analogous helix in the 4WJ molecule. **B)** Chemical structures of the fluorophores Alexa488 (green) and Cy5 (red) covalently coupled to RNA with a flexible hexamethylene linker (blue) on modified uracil or guanine bases (black). The #A refers to the number of atoms in the flexible region between the ring systems of the nucleobase and dye. **C)** The different attachment points on the uracil and guanine bases place the dye linker in the major and minor grooves, respectively. This results in a preferential orientation of the linker towards the 3'-end for uracil attachment. Top: The attachment points (yellow) in the Watson-

Crick base pairs on the C5 atom in uracil and the N2 atom in guanine point towards the major (pink) or minor grooves (light blue), respectively. Bottom: Structural model of the RNA with the Cy5 dyes (red) attached to uracil (left) or guanine (right). The accessible volume of the Cy5 dye is shown in pink and light blue for attachment to the uracil and guanine bases, respectively.

Here we performed an extensive study of the photophysical properties of the commonly used organic fluorophores Alexa488 and Cy5 attached to double-stranded RNA duplexes, three-way and four-way junctions. We monitor the position-dependent fluorescence properties, i.e., the fluorescence lifetime, anisotropy, quantum yield, absorption and emission spectra, and photophysical dynamics at 22 different labeling sites on nine different RNA molecules, amounting to 104 combinations of fluorophore, labeling site and type of RNA molecule. Our extensive dataset provides a comprehensive view of the effect of the nucleic acid environment and the covalent attachment on the fluorescence properties, revealing unique signatures of distinct microenvironments on the RNA. The experimental results are corroborated by accessible volume simulations of the flexibly coupled fluorophores that predict the accessibility of interaction sites on the RNA surface. Further atomistic insights are gathered from all-atom molecular dynamics simulations of the coupled dyes that identify end-stacked conformations and the dominant interaction modes of internal labels within the major and minor grooves. The gathered insights drastically improve our understanding and the predictability of dye photophysics in nucleic acid systems and provide the necessary information for using fluorescent labels as environmental probes of the local and global structure of nucleic acids. We close by assessing the effect of the variations of the position-dependent dye behavior on FRET studies and propose guidelines to minimize undesirable effects to improve the accuracy of FRET studies.

3.3 Results

To study the influence of the biomolecular environment on the photophysical properties of the dyes Cy5 and Alexa488 flexibly coupled to RNA, we selected a variety of labeling sites on RNA molecules of increasing complexity (Fig. 3.1A). All molecules are derived from a common four-way junction (4WJ) whose junction region is identical to the one found in the hairpin ribozyme (PDB-ID: 1M5K, Supplementary Fig. 3.1). In addition to the complete 4WJ, we also consider possible fragments thereof, resulting in four possible three-way junctions (3WJ, T_{1-4}) and four duplexes (I_{1-4}) (Fig. 3.1A and Supplementary Fig.

3.2). The fluorescent labels are attached by long flexible linkers either to an uracil base at the C5 atom, or to a guanine base at the N2 atom (Fig. 3.1B). For uracil labels, the linker emerges from the major groove of the RNA, while the anchor point for guanine labels lies within the minor groove (Fig. 3.1C).

3.3.1 Photophysics of Cy5 are sensitive to the local environment.

We first investigated the position-dependent properties of sulfonated Cy5 tethered to RNA by recording the time-resolved fluorescence and anisotropy decays for 15 distinct labeling positions on RNA duplexes, three-way junctions (3WJ) and four-way junctions (4WJ), resulting in a total of 49 distinct environments (Fig. 3.1A, Supplementary Fig. 3.3). While free Cy5 in water shows a monoexponential fluorescence decay with lifetime of 0.95 ns, we observe multiexponential decays for all RNA-coupled Cy5 labels (Fig. 3.1B and Supplementary Table 3.1). The shortest lifetime component for all studied labels falls in the range from 0.9 to 1.1 ns (1.0 ± 0.1 ns). This component corresponds to the fraction of coupled dye that is not interacting with the RNA and freely moving within the constraints set by the linker. The longer lifetime component, on the other hand, covers a wider range between 1.5 to 2.1 ns (1.7 ± 0.2 ns). The larger spread of the lifetimes arises from interactions of the dye with the heterogeneous microenvironment on RNAs. No quenching mechanisms (such as the formation of long-lived dark states) were found to be relevant for Cy5 (Supplementary Fig. 3.4 and Supplementary Table 3.2).

The photophysics of cyanine dyes are governed by steric constraints that modulate the cis-trans photoisomerization^{43,90,131}. Steric restriction due to interactions with the local environment hinders the rotation around the central stretch of conjugated carbon-carbon double bonds and thus reduces the rate for cis-trans isomerization, which is the predominant non-radiative pathway in cyanine dyes. A detailed potential energy diagram of the Cy5 fluorophore is given in Supplementary Fig. 3.5. Consequently, steric hindrance increases the lifetime of the excited state, an effect that was originally termed protein-induced fluorescence enhancement (PIFE)¹³² and has since also been described for nucleic acids^{90,140}. As a more generally applicable acronym, we denote this effect here as association-induced fluorescence enhancement (AIFE). Steric restriction can be detected by the time-resolved anisotropy that measures the depolarization of the fluorescence signal due to rotational motion (Fig. 3.2B). The anisotropy decays for Cy5 are

characterized by two rotational correlation times. The shorter correlation time ($\rho_1 = 0.4$ - 0.8 ns) describes the fast rotation of the fluorophore (partially restricted by the environment), while the longer correlation time ($\rho_{\text{global}} > 15$ ns) reflects the overall tumbling motion of the RNA. The dye motion is interpreted based on the 'diffusion with traps' (DWT) dye model²²⁸, which assumes that the fluorophores either rotate freely or are trapped by sticking to the surface. The amplitude of the global tumbling motion of the RNA, given by the residual anisotropy, r_∞ , then directly reports on the fraction of fluorophores that stick to the host molecule by stable interactions. Meanwhile, non-interacting fluorophores can rotate freely due to the flexible C6-linker, resulting in an almost complete depolarization of the fluorescence signal.

A plot of the species-averaged fluorescence lifetimes $\langle \tau \rangle_x$ against the residual anisotropies r_∞ is given in Fig. 3.2A. A clear correlation is observed between the two variables over all positions with a Pearson correlation coefficient of 0.95. Labeling positions in similar RNA environments show similar photophysical parameters, resulting in three major clusters of labels whose attachment points are either (1) close to the 3'-end (magenta), (2) close to the junctions (purple) or (3) in the internal parts of the helices (blue). The cluster of internal labels can be further subcategorized into uracil-attached (dark blue) and guanine-attached labels (blue-gray). Labels close to the 3'-end show the highest average lifetimes (> 1.5 ns) and residual anisotropies ($r_\infty > 0.25$). Slightly lower values of $\langle \tau \rangle_x \sim 1.4$ - 1.5 ns and $r_\infty \sim 0.25$ are observed for labels near the junction, while the labels at internal positions along the RNA duplex exhibit the lowest lifetimes ($\langle \tau \rangle_x < 1.3$ ns) and residual anisotropies ($r_\infty < 0.2$). The sub-cluster of internal Cy5 labels with the dyes attached to guanine (blue-gray, Fig. 3.2A) exhibit a narrow range of lifetimes ($\langle \tau \rangle_x \sim 1.2 \pm 0.01$ ns) and residual anisotropies ($r_\infty \sim 0.18 \pm 0.01$).

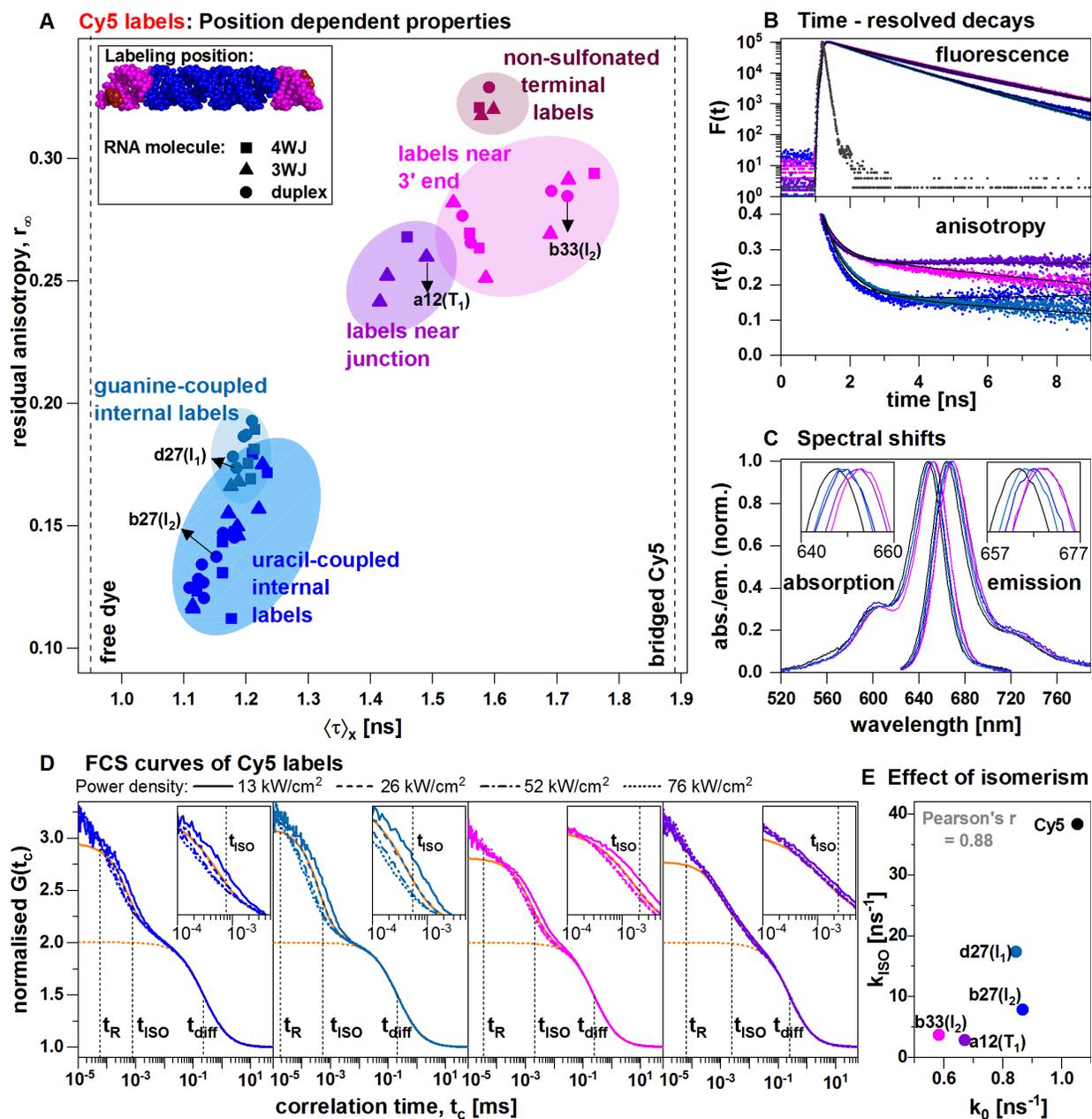


Figure 3.2: Photophysics of Cy5 in different RNA microenvironments. A) A plot of the species-averaged fluorescence lifetime $\langle \tau \rangle_x$ and residual anisotropy r_∞ reveals distinct clusters of Cy5 labels originating from the different RNA microenvironments (uracil-coupled internal labels: bright blue, guanine-coupled internal labels: blue-gray, near-3'-end labels: magenta, labels close to the junction: purple, non-sulfonated terminal labels: dark red). The inset illustrates the classification into the different RNA environments on the structural model of an RNA duplex. The dashed lines denote the lifetimes of the free Cy5 dye and a bridged Cy5 dye²²⁷ that does not undergo photo-isomerization of 1.9 ns. **B)** Fluorescence (top) and anisotropy decays (bottom) for representative Cy5 labels of the different clusters. The instrument response function is shown in dark grey. Fits are shown as solid lines. **C)** Comparison of absorption and emission spectra of representative Cy5 labels (colored) and free Cy5 dye (black). The peak

region is magnified in the insets. **D)** FCS curves of representative Cy5 labels of the different clusters reveal the environment-dependent modulation of the cis-trans photoisomerization. From left to right: uracil-coupled internal labels, guanine-coupled internal labels, labels near 3' end, labels near junction. Curves were acquired at power densities between 13-76 kW/cm². Fits to the model function given by eq. 3.25, accounting for translational diffusion (t_{diff}), two isomerization processes (t_{ISO}) and molecular rotation (t_{R}), are shown as black lines. The vertical dashed lines indicate the fitted relaxation times for the FCS curve at 26 kW/cm² excitation. For photo-isomerization, the weighted average of the two fitted relaxation times is shown. The orange curves indicate the partial model functions of the different contributions, assuming only diffusion (dashed line) or diffusion and photoisomerization (solid line). The displayed FCS curves represent the average of the parallel-perpendicular and perpendicular-parallel correlation functions, normalized to the total number of molecules. Insets show the curves around t_{ISO} . **E)** A plot of the rate constant of fluorescence decay (k_0) against the trans-to-cis isomerization rate constant (k_{ISO}) obtained from the FCS curves at 26 kW/cm² excitation power.

3.3.2 Cy5 labels close to the 3'-end stack to the terminal base pair

The longer lifetimes and larger residual anisotropies for the Cy5 labels close to the 3'-end are likely caused by π - π stacking interactions with the exposed terminal base pair, which hinders the cis-trans photoisomerization and enhances the fluorescence by AIFE. This is suggested by the fact that the high $\langle\tau\rangle_x$ values of 1.5-1.8 ns are mainly caused by an increased fraction of the long lifetime component (65-85%) which also shifts slightly to longer lifetimes (1.7-1.9 ns), indicating an increased population of a stacked conformation with high steric restriction. As a control for a fully-stacked conformation, we measured a non-sulfonated Cy5 label attached to the 5'-end phosphate of the RNA duplex with a three-carbon linker (Supplementary Fig. 3.6), which has previously been shown to stack to the terminal base pair^{91,92}. We observed long lifetimes and high residual anisotropies similar to the labels attached near the 3'-end (magenta in Fig. 3.2A), confirming that efficient stacking occurs for these labels despite a distance of 4-6 bp to the terminal base (Fig. 3.2A) due to the long flexible linker used to tether the dye to the base (Fig. 3.1B). Consistent results are obtained for the non-sulfonated Cy5 label on the three-way and four-way and junction molecules (Fig. 3.2A and Supplementary Table 3.1). The long lifetimes obtained for the near-3'-end labels are also in agreement with a lifetime of 1.9 ns obtained for a conformationally restrained bridged-Cy5²²⁷. Surprisingly, we observed low anisotropies and lifetimes for position d27 although it is situated at a similar distance to the end as the labels at position d26 and d28 (Fig. 3.2A). As label d27 is attached to a guanine residue, this indicates that the propensity for stacking not only depends on the distance but also the tether point on the duplex.

To further elucidate the π - π stacking interactions of the near-3'-end labels, we measured the absorption and emission spectra for representative Cy5 labels from different clusters (Fig. 3.2C). While the internal Cy5 labels showed a shift of 3 nm of the absorption maxima compared to the free dye, the near-3'-end label was red shifted by 5 nm. A similar shift was observed for the emission spectra (Fig. 3.2C). Although the absorption of the non-fluorescent cis state is red shifted²²⁹, it is unlikely that the observed red shift is caused by a change of the cis-trans equilibrium, as only a negligible population of the thermodynamically unfavored cis state is expected under the low excitation power applied in the ensemble experiments. Instead, because Cy5 exhibits a bathochromic shift with decreasing solvent polarity^{230,231}, the spectral shift is most likely caused by the less polar environment experienced by the interacting dye at the terminal base pair or within the grooves.

To obtain more detailed insights into the photoinduced isomerization, we performed full fluorescence correlation spectroscopy (FCS)¹⁵⁸ measurements on free Cy5 dye (Supplementary Fig. 3.7) and representative Cy5 labels from each cluster (Fig. 3.2D). The FCS curves show a characteristic power-dependent bunching term on the microsecond timescale, indicative of the cis-trans photoisomerization. We fitted the curves with a model composed of one translational diffusion term ($t_{\text{diff}} = 200\text{-}250 \mu\text{s}$), one rotational diffusion term ($t_{\text{R}} = 20\text{-}60 \text{ ns}$) and two additional bunching terms with relaxation times, $t_{\text{ISO},1}$ and $t_{\text{ISO},2}$ ($\sim 0.3\text{-}16 \mu\text{s}$) to describe the cis-trans photoisomerization (see Methods and Supplementary Table 3.3). We did not observe a significant contribution of triplet states under the applied excitation powers, in agreement with previous reports¹⁴¹. While the correlation amplitude of the photoisomerization process did not change with the excitation power, we observe a clear decrease of the relaxation times (Fig. 3.2D). This behavior is expected as the increased excitation rate accelerates both the forward and backward isomerization rates but has only a minor effect on the cis-trans equilibrium due to the similar absorption cross-sections of the cis and trans states¹⁴¹.

The power-dependent relaxation times, $t_{\text{ISO},1}$ and $t_{\text{ISO},2}$, are clearly influenced by the RNA environment as the steric restriction imposed by the interactions with the terminal bases hinders the conversion between the cis and trans isomers (Supplementary Fig. 3.8). Accordingly, the relaxation time of the cis-trans process is slower for the end label than

for the internal labels and the free dye (Fig. 3.2D, Table 3.1 and Supplementary Table 3.3). We also observed a decrease of the amplitude of the cis-trans process from the free dye over the internal labels to the end label (Fig. 3.2D and Table 3.1), suggesting that the RNA-dye interactions result in a shift of the equilibrium between the two photoisomers. From the measured amplitudes and relaxation times of the photoisomerization process, we computed the excitation-independent trans-to-cis isomerization rate k_{ISO} (Table 3.1, see Methods). We observed a good correlation (Pearson's correlation coefficient = 0.88) between k_{ISO} and k_0 (Fig. 3.2E), confirming that steric hindrance of the cis-trans isomerization is the cause for the increased fluorescence lifetimes of the Cy5 labels close to the junction and 3'-end.

Table 3.1. Photophysical properties of RNA-labeled Cy5 dyes: representative example and mean for different microenvironments.

| Dye description | | Free dye | Internal U-attached | | Internal G-attached | | Near-3'-end | | Near-junction | |
|-----------------------------------|------------------------------|----------|---------------------|------------|---------------------|------------|-------------|------------|---------------|------------|
| Label position | | - | b27 | Mean (SD) | d27 | Mean (SD) | b33 | Mean (SD) | a12 | Mean (SD) |
| Molecule | | - | I_2 | | I_1 | | I_2 | | T_1 | |
| Fluorescence decay analysis | x_1 | - | 0.22 | 0.24(0.04) | 0.37 | 0.36(0.01) | 0.80 | 0.75(0.07) | 0.4 | 0.37(0.03) |
| | τ_1, ns | - | 1.67 | 1.66(0.13) | 1.55 | 1.58(0.02) | 1.91 | 1.86(0.08) | 2.12 | 2.08(0.05) |
| | x_2 | 1 | 0.78 | 0.76(0.04) | 0.63 | 0.64(0.01) | 0.20 | 0.25(0.07) | 0.6 | 0.63(0.03) |
| | τ_2, ns | 0.95 | 1.01 | 1.00(0.03) | 0.97 | 0.98(0.01) | 0.94 | 0.95(0.04) | 1.06 | 1.07(0.01) |
| | $\langle \tau \rangle_x, ns$ | 0.95 | 1.15 | 1.16(0.04) | 1.18 | 1.20(0.01) | 1.72 | 1.63(0.08) | 1.49 | 1.45(0.03) |
| Time-resolved anisotropy analysis | r_1 | - | 0.24 | 0.24(0.02) | 0.21 | 0.20(0.01) | 0.09 | 0.10(0.01) | 0.12 | 0.12(0.01) |
| | ρ_1, ns | - | 0.74 | 0.67(0.06) | 0.64 | 0.65(0.05) | 0.52 | 0.54(0.08) | 0.51 | 0.54(0.05) |
| | r_∞ | - | 0.14 | 0.14(0.02) | 0.17 | 0.18(0.01) | 0.28 | 0.28(0.01) | 0.26 | 0.26(0.01) |
| | ρ_{global}, ns | - | >100 | >100 | 18.1 | 48.61(51) | 18.6 | 26.18(11) | >100 | >100 |
| FCS | A_{ISO} | 0.6 | 0.49 | - | 0.52 | - | 0.45 | - | 0.44 | - |
| | $t_{ISO}, \mu s$ | 0.3 | 0.77 | - | 0.54 | - | 2.07 | - | 2.4 | - |
| | k_{ISO}, ns^{-1} | 38.4 | 7.88 | - | 17.39 | - | 3.72 | - | 2.87 | - |
| Spectral maxima | $\lambda_{max}(abs), nm$ | 648 | 650 | - | 650 | - | 653 | - | 652 | - |
| | $\lambda_{max}(em), nm$ | 664 | 666 | - | 665 | - | 668 | - | 668 | - |

Fluorescence decays were fitted to a two-component model function with the fluorescence lifetimes τ_{1-2} and corresponding fractions x_{1-2} according to eq. 3.4. The species-weighted average fluorescence lifetime, $\langle \tau \rangle_x$ is calculated using eq. 3.5. Anisotropy decays were described by two components with rotational correlation times ρ_1 and ρ_{global} and corresponding fractions r_1 and r_∞ according to eq. 3.7. The amplitude of the slow component is the residual anisotropy r_∞ . A_{ISO} and t_{ISO} are the amplitudes and relaxation times for the cis-trans isomerization process obtained by fitting the FCS curves to the model function eq. 3.25. The rate constant of trans-to-cis isomerization, k_{ISO} is calculated using eq. 3.28.

$\lambda_{\max}(\text{abs})$ and $\lambda_{\max}(\text{em})$ are the peak wavelengths of the absorption spectra and emission spectra excited at 635 nm, respectively. For the time-resolved measurements, the mean and standard deviation (SD) of samples measured within a given cluster (microenvironment) are tabulated.

3.3.3 Cy5 labels interact with the exposed bases at the junction

The Cy5 labels near the junction region exhibit photophysical properties very similar to the labels near the 3'-end. These near-junction labels have high residual anisotropies (~ 0.26), lifetimes $\langle\tau\rangle_x$ ranging between 1.4-1.5 ns with a long lifetime component of ~ 2.1 ns (35-40%), and a ~ 5 nm red shifted absorption maximum compared to the free dye (Fig. 3.2C). The FCS curves (Fig. 3.2D) also show slower characteristic cis-trans relaxation times compared to the internal labels (Table 3.1).

For the label a12 on the 3WJ molecules (positioned four base pairs away from the junction), a clear restriction of photoisomerization is observed from the significantly longer lifetimes and higher residual anisotropies ($\langle\tau\rangle_x = 1.4\text{-}1.5$ ns, $r_\infty = 0.24\text{-}0.26$). This contrasts with the same label on the 4WJ molecule ($\langle\tau\rangle_x = 1.16$ ns, $r_\infty = 0.13$), which behaves like the corresponding label on the duplex ($\langle\tau\rangle_x = 1.12$ ns, $r_\infty = 0.13$). In the 4WJ, the helices most likely assume a coaxially stacked configuration similar to the native structure of the hairpin ribozyme (Supplementary Fig. 3.1)²³² or the Holliday junction in DNA²³³, resulting in a fully base-paired junction that is inaccessible to the dye. In the 3WJ, however, only two of the three helices can stack coaxially at a time. It is thus possible that the bases at the junction are exposed and accessible to the dye, enabling potential stacking interactions. The complex shape of the FCS curves for junction labels on the 3WJ molecules provides further evidence for such interactions (Fig. 3.2D and Supplementary Table 3.3). In certain cases, AIFE is also observed for labels near the junction on the 4WJ molecule. For example, the label d18 on the 4WJ, positioned one base pair away from the junction, exhibits a long lifetime ($\langle\tau\rangle_x = 1.46$ ns) and a high residual anisotropy ($r_\infty = 0.27$). In this case, steric restriction could occur because the dye is situated in the confined region between two coaxially stacked helices of the 4WJ.

In contrast to the labels near the ends or junction, dyes attached to internal positions undergo no base stacking but still experience some degree of steric restriction, as indicated by longer $\langle\tau\rangle_x$ values ranging between 1.1-1.3 ns compared to the free dye with a lifetime of 0.95 ns. This is also reflected in the FCS curves by the longer relaxation time

of the cis-trans process and a small red shift of the absorption and emission spectra of ~ 2 nm compared to the free dye. Among the internal labels, we observe distinct differences between uracil- and guanine-attached labels. Guanine-attached labels show slightly increased residual anisotropy and a narrower distribution for both lifetimes and residual anisotropies compared to uracil-attached labels. These features originate from differences in the attachment point and the local environment at labeling site, as discussed further in later sections.

3.3.4 Site-specific photophysics of Alexa488 are dominated by quenching

Next, we measured the position dependent fluorescence properties of the rhodamine-based dye Alexa488 for 14 different labeling positions on the different RNA molecules, resulting in a total of 51 distinct environments (Fig. 3.1A, Fig. 3.3A and Supplementary Fig. 3.9). The anisotropy decays were characterized by three rotational correlation times (Fig. 3.3B). The two faster correlation times describe the rotational motion of the fluorophore ($\rho_1 = 0.2$ ns) and the linker motion ($\rho_2 = 1-2$ ns) in the RNA environment, while the longest correlation time ($\rho_{\text{global}} = 7-40$ ns) measures the overall tumbling motion of the RNA. As mentioned for the Cy5 labels, the amplitude of this global motion is given the residual anisotropy r_∞ which reports on the fraction of fluorophores that interact with the host molecule. The fluorescence lifetime decays were characterized by two to three lifetime components (Supplementary Table 3.4). For the coupled dye, the species-averaged lifetime $\langle \tau \rangle_x$ is consistently found to be lower compared to the free dye ($\langle \tau \rangle_x = 4.1$ ns) due to quenching by photoinduced electron transfer (PET) from the nucleobase guanine^{96,104,125,234,235}. PET occurs upon orbital overlap either through (i) collisions or short-lived interactions (dynamic quenching) or (ii) formation of stable π -stacked complexes (static quenching). While both mechanisms reduce the fluorescence intensity, they have different effects on the observed lifetime. For dynamic quenching, the interaction occurs due to collisions with the quencher within the excited state lifetime of the dye, reducing the observed fluorescence lifetime. The reduction of the fluorescence lifetime relates to the frequency of collisions with accessible guanine bases. For static quenching, the long-lived quenched species is only weakly fluorescent with a short fluorescence lifetime in the picosecond range^{215,235,236}, which is generally too short to be detected in the lifetime analysis but results in a decrease of the steady-state fluorescence

intensity. The two longest lifetimes (around 1.7 ns and 4 ns) are assigned to environment dependent dynamic quenching. For some labels, a short lifetime component of ~ 0.3 ns was required to achieve a good fit which indicates a significant contribution of strongly quenched dyes.

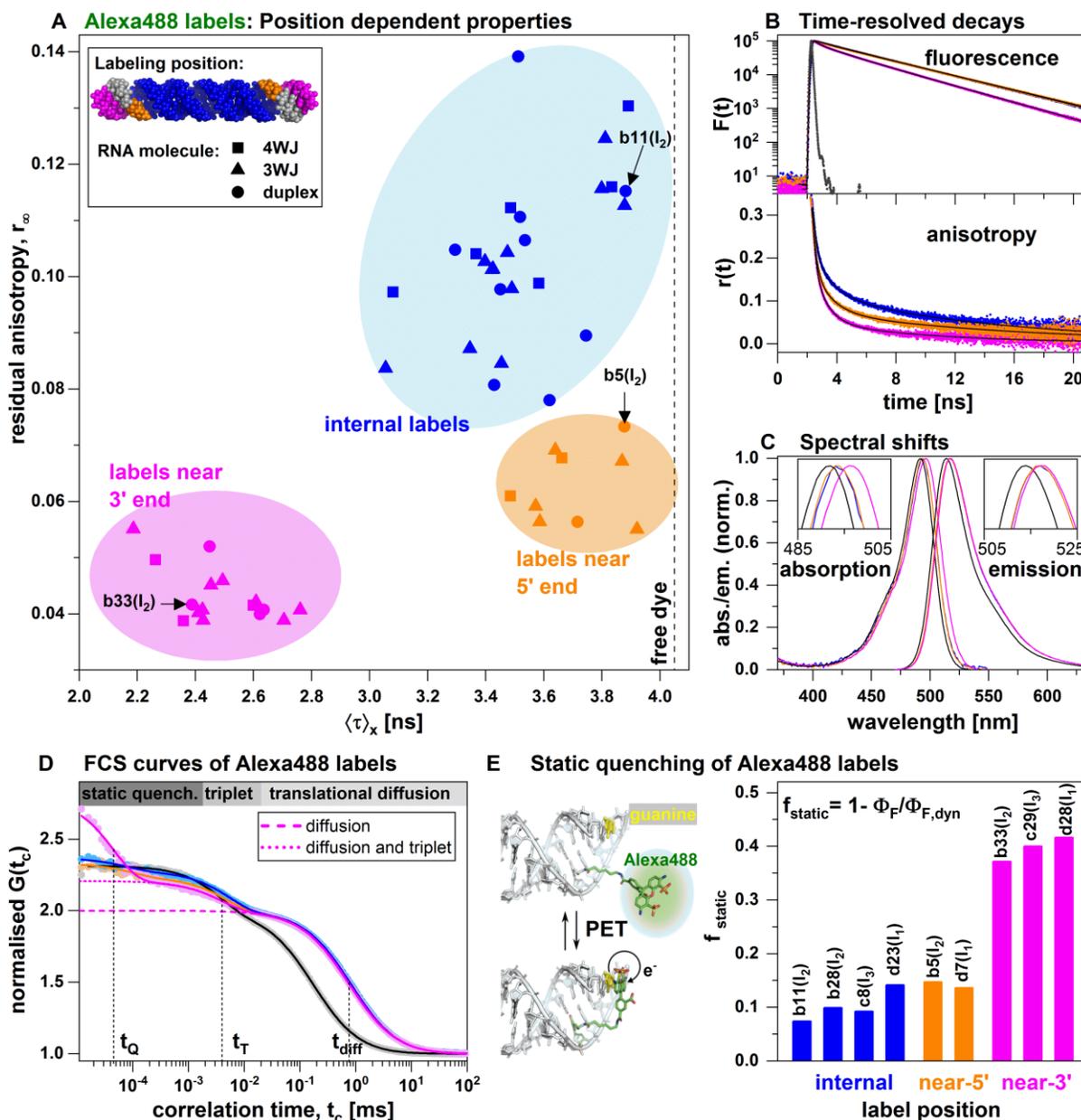


Figure 3.3: Photophysics of Alexa488 in different RNA microenvironments. A) A plot of the species-averaged fluorescence lifetime $\langle \tau \rangle_x$ and residual anisotropy r_∞ reveals distinct clusters of Alexa488 labels originating from the different RNA microenvironments (internal labels: bright blue, near-5'-end labels: orange, near-3'-end labels: magenta). The inset illustrates the classification into the different RNA environments on the structural model of an RNA duplex. **B)** Fluorescence (top) and anisotropy decays (bottom) for representative

Alexa488 labels of the different clusters. The instrument response function is shown in dark grey. Fits are shown as solid lines. **C**) Comparison of absorption and emission spectra of representative Alexa488 labels (colored as in A) and free Alexa488 dye (black). The peak region is magnified in the insets. **D**) FCS curves of representative Alexa488 labels of the different clusters (colored as in A) reveal position-dependent PET quenching. The FCS curve of the free Alexa488 dye is shown in black. Curves were fit to a model accounting for translational diffusion (t_{diff}), triplet photophysics (t_{T}) and PET quenching (t_{Q}) according to eq. 3.21. The near-3'-end label b33 on duplex I_2 shows a bunching term at ~ 45 ns due to PET quenching by the terminal guanine base. For this label, the model functions describing only diffusion and diffusion with triplet photophysics are shown as dashed lines. The displayed FCS curves represent the average of the parallel-perpendicular and perpendicular-parallel correlation functions, normalized to the total number of molecules. **E**) Estimated fraction of statically quenched Alexa488 labels, f_{static} , at different labeling positions. The statically quenched fraction is calculated according to eq. 3.18. Cartoon: Static quenching by PET occurs upon stacking of the fluorophore to the exposed guanine base (yellow) at the helix end.

Based on the species-averaged lifetimes and residual anisotropies (Fig. 3.3A), we identify three clusters that correspond to internal labels and labels close to the 3'- or 5'- ends. Close to the 3'-end, Alexa488 displays the lowest residual anisotropies ($r_{\infty} \sim 0.04$) and lifetimes ($\langle \tau \rangle_x < 2.8$ ns), while labels close to the 5'-end show higher residual anisotropies ($r_{\infty} \sim 0.07$) and significantly longer lifetimes ($\langle \tau \rangle_x > 3.5$ ns) compared to the near-3'-end labels. Internal Alexa488 labels show the highest residual anisotropies over a wide range from 0.08-0.14 and a larger spread of the lifetimes from 3-4 ns. Photophysical properties of representative examples from each cluster are given in Table 3.2. The existence of these clusters suggests position-specific interaction modes of Alexa488 in different microenvironments on RNA.

3.3.5 End-stacked Alexa488 fluorophores are efficiently quenched by PET

Alexa488 labels close to the 3'-end exhibit the shortest lifetimes ($\langle \tau \rangle_x = 2.2$ -2.8 ns) and their fluorescence decays contain a significant contribution of a short lifetime component (~ 300 ps), indicating highly efficient quenching by guanines through PET (Supplementary Table 3.4). At the helix ends, there is not only a high number of guanines in the studied sequence, but the π -electron system of the exposed terminal guanines is more accessible to the coupled dye than for the internal guanines, resulting in highly efficient PET quenching. Additionally, the linker is oriented towards the terminal nucleobases at the 3'-end. Neubauer *et al.* studied the dye Rhodamine 6G bound to the 5'-end of a DNA duplex, showing that the dye stacks onto the terminal base pair⁹⁶. Similar stacking interactions are thus expected for the rhodamine-based dye Alexa488 studied here.

Surprisingly, 3'-end labels exhibit significantly lower residual anisotropies of ($r_{\infty} = 0.03-0.05$) compared to the internal labels (Fig. 3.3A), while Cy5 labels close to the 3'-end showed the highest propensity for end stacking (Fig. 3.2A). This result is explained by the fact that the stacking of Alexa488 to the terminal guanine base leads to highly efficient static quenching by PET, making the stacked species virtually non-fluorescent and thus invisible in the ensemble experiments. The fluorescence signal hence originates exclusively from the uncomplexed fluorophores that are free to move and thus show low residual anisotropy. The reduced fluorescence lifetime for the 3'-end labels is therefore due to the dynamic quenching of the fraction of uncomplexed dyes by collisions and short-lived interactions with the high number of guanines present at helix ends. To confirm the formation of a non-fluorescent species for the 3'-end labels, we measured the steady-state quantum yield of representative Alexa488 labels from the different clusters using the low-intensity FCS approach²³⁷. By comparing the steady-state quantum yield Φ_F with the dynamic quantum yield $\Phi_{F,dyn}$ estimated from the fluorescence lifetime, we estimate the fraction of statically quenched species as $f_{static} = 1 - \Phi_F / \Phi_{F,dyn}$ (Fig. 3.3E). Indeed, the highest fraction of statically quenched species was observed for the near-3'-end labels (27-41%), whereas the internal and near 5' labels showed only a low amount of static quenching (6-18%).

While the distinction between static and dynamic quenching based on the effect of the quenching process on the fluorescence lifetime suggests the absence of dynamics for statically quenched species, the dark (statically quenched) species must transition to the bright (unquenched or dynamically quenched) state on timescales longer than the excited state lifetime. To probe this exchange, we performed FCS measurements of representative Alexa488 labels (Fig. 3.3D). For the near-3'-end label, a strong bunching term is observed at short correlation times (~40 ns) which is absent for the free dye and appears only as a small amplitude for the internal and near-5'-end labels. This term describes the dynamic exchange between the statically quenched and bright, unstacked species observed for the 3'-end labels (for FCS curves of additional Alexa488 labels see Supplementary Fig. 3.10 and Supplementary Table 3.5). The amplitude of this term is a measure of the amount of static quenching and correlates well with the fraction of the

statically quenched species determined from the quantum yield measurements (Supplementary Fig. 3.11C).

Table 3.2. Photophysical properties of RNA-labeled Alexa488 dyes: representative example and mean for different microenvironments.

| Dye description | | Free dye | Internal | | Near-5'-end | | Near-3'-end | |
|-----------------------------------|---|----------|----------|-------------|-------------|-------------|-------------|-------------|
| Label position | | | b11 | Mean (SD) | b5 | Mean (SD) | b33 | Mean (SD) |
| Molecule | | | I_2 | | I_2 | | I_2 | |
| Fluorescence decay analysis | x_1 | 1.00 | 0.94 | 0.82 (0.07) | 0.92 | 0.85 (0.06) | 0.59 | 0.60 (0.04) |
| | τ_1 , ns | 4.05 | 4.05 | 4.00 (0.09) | 4.11 | 4.10 (0.02) | 3.56 | 3.62 (0.08) |
| | x_2 | - | 0.06 | 0.13 (0.04) | 0.08 | 0.11 (0.03) | 0.17 | 0.16 (0.02) |
| | τ_2 , ns | - | 1.35 | 1.85 (0.40) | 1.31 | 1.68 (0.54) | 1.54 | 1.51 (0.05) |
| | x_3 | - | - | 0.08 (0.3) | - | 0.07 (0.01) | 0.24 | 0.24 (0.03) |
| | τ_3 , ns | - | - | 0.34 (0.11) | - | 0.50 (0.10) | 0.32 | 0.30 (0.02) |
| | $\langle\tau\rangle_x$, ns | 4.05 | 3.88 | 3.53 (0.23) | 3.88 | 3.70 (0.15) | 2.45 | 2.49 (0.16) |
| Time-resolved anisotropy analysis | r_1 | - | 0.16 | 0.17 (0.02) | 0.19 | 0.19 (0.01) | 0.18 | 0.20 (0.01) |
| | ρ_1 , ns | - | 0.24 | 0.24 (0.04) | 0.23 | 0.23 (0.00) | 0.2 | 0.22 (0.03) |
| | r_2 | - | 0.1 | 0.11 (0.01) | 0.11 | 0.13 (0.01) | 0.15 | 0.14 (0.01) |
| | ρ_2 , ns | - | 1.35 | 1.52 (0.26) | 1.24 | 1.36 (0.07) | 0.92 | 1.05 (0.09) |
| | r_∞ | - | 0.12 | 0.10 (0.01) | 0.07 | 0.06 (0.01) | 0.05 | 0.04 (0.00) |
| | ρ_{global} , ns | - | 12.99 | 17.01 (6.4) | 17.67 | 15.34 (5.0) | 8.07 | 10.72 (2.1) |
| FCS | A_Q | - | 0.07 | - | 0.07 | - | 0.33 | - |
| | t_Q , ns | - | 79.3 | - | 98.3 | - | 45.8 | - |
| Spectral maxima | $\lambda_{\text{max}}(\text{abs})$, nm | 492 | 493 | - | 493 | - | 496 | - |
| | $\lambda_{\text{max}}(\text{em})$, nm | 514 | 517 | - | 517 | - | 518 | - |
| f_{static} | | - | 0.08 | - | 0.14 | - | 0.4 | - |

Fluorescence decays were fitted to a three-component model with fluorescence lifetimes τ_{1-3} and corresponding fractions x_{1-3} according to eq. 3.4. The species-weighted average fluorescence lifetime, $\langle\tau\rangle_x$ is calculated using eq. 3.5. Anisotropy decays were described by three components with rotational correlation times ρ_1 , ρ_2 and ρ_{global} and corresponding fractions r_1 , r_2 and r_∞ according to eq. 3.7. The amplitude of the slow component is the residual anisotropy r_∞ . A_Q and t_Q are the amplitudes and relaxation times for the static quenching process obtained on fitting FCS curves to the model function eq. 3.21. $\lambda_{\text{max}}(\text{abs})$ and $\lambda_{\text{max}}(\text{em})$ are the peak wavelengths of the absorption spectra and emission spectra excited at 488 nm, respectively. f_{static} , the fraction of statically quenched species is determined from the fluorescence quantum yields using eq. 3.18. For the time-resolved measurements, the mean and standard deviation (SD) of samples measured within a given cluster (microenvironment) are tabulated.

Dynamic quenching resulting from collisional interactions only affects the excited states of the fluorophores. Consequently, no change in the absorption spectra compared to the

free dye is expected. In contrast, static quenching involves the formation of a stable ground-state complex between the dye and the terminal guanine, often resulting in a characteristic shift of the absorption spectrum²³⁵. As the complexed fluorophores absorb but do not emit, the formation of ground-state complexes is manifested as a significant red shift of the absorption spectra of 5 nm for the representative label from the near-3'-end cluster (Fig. 3.3C) and a characteristic shift between the absorption and excitation spectra (Supplementary Fig. 3.12). A much smaller shift of the absorption spectrum of ~2 nm compared to the free dye is observed for the internal and near-5'-end labels (Table 2, Fig. 3.3C and Supplementary Fig. 3.13A). On the contrary, the emission spectra show no significant environment-dependent shifts (Fig. 3.3C and Supplementary Fig. 3.13B).

Internal labels show a large spread of lifetimes and anisotropies (Fig. 3.3A). As will be discussed in detail below, this is also seen for the simple duplex molecules and hence most likely a sequence-specific effect. The presence of dynamic quenching for most internal labeling positions, as detected from the reduced lifetimes, suggests that the π -electron system of internal guanine bases is accessible to the fluorophore in suitable interaction geometries. Interestingly, we observe higher residual anisotropies for the internal labels compared to the near-5'-end labels, which could originate from a higher structural flexibility of the RNA at the helix ends as seen in molecular dynamic simulations (Supplementary Fig. 3.14). There are also distinct differences between the photophysical properties of Alexa488 labels close to the 3' and 5'-ends (Fig. 3.3A). Despite similar base pair distances from the helix ends, the near 5' labels show considerably longer lifetimes ($\langle\tau\rangle_x = 3.6\text{-}3.9$ ns) and moderately higher residual anisotropies ($r_\infty = 0.06\text{-}0.07$) compared to the near-3'-end labels ($\langle\tau\rangle_x = 2.2\text{-}2.8$ ns, $r_\infty = 0.04 - 0.06$). These observations suggest the absence of end-stacking for the Alexa488 dyes labeled close to the 5'-end.

3.3.6 Predicting photophysical properties from accessible volume simulations

To gain insights into the molecular interactions underlying the observed photophysical effects, we assess the sterically accessible volume (AV) of the fluorophores on RNA. The AV simulation is based on geometric modeling of the dye, which has been widely applied for the prediction of FRET-derived distances^{103,238-241}. The AV approach is simpler, faster,

and more convenient compared to alternative approaches such as coarse-grained or all-atom molecular dynamics simulations, while still providing accurate insights into potential interactions and steric clashes. The dye is modeled as a sphere with an empirical radius of R_{dye} , connected to the nucleobase by a flexible linker of length, L_{link} and width, w_{link} (Supplementary Fig. 3.15 and Supplementary Table 3.7). The set of sterically allowed dye positions then defines the AV. Several modifications to the simple AV model have been proposed that model the dye as an ellipsoid^{239,242} or apply weights to the sampled positions to account for the thermodynamic cost of linker extension^{243,244} or potential interactions with the molecular surface^{90,242}. Here, we apply the AV approach to visualize the accessible space of dyes at specific positions and quantify the accessibility of the dye to potential interaction sites to understand the photophysical observations.

3.3.7 Terminal base pair accessibility predicts Cy5 stacking interactions

In the experiments, we detected interactions with the terminal bases for labels that are situated up to six base pairs away from the helix end, facilitated by the long flexible dye linkers. To evaluate the accessibility of the terminal base pair for potential sticking interactions with the Cy5 fluorophore, we performed AV simulations on a model A-form RNA duplex (Fig. 3.4A). We define the terminal base pair accessibility as the fraction of the AV that is in reach of the terminal base pair, using a cutoff distance of 7 Å. The dependence of the terminal base pair accessibility on the 3' or 5'-ends as a function of the base pair distance is shown in Fig. 3.4A for Cy5 labels attached either to uracil (top, linker length 22 Å) or guanine (bottom, linker length 19.5 Å). The AV calculations indeed show that fluorescent labels can reach the ends of the RNA helix even if the dye is positioned at a distance of up to 8 base pairs. Uracil-attached labels show a higher terminal accessibility on the 3'-end than on the 5'-end, with maximum distances of 8 and 4 bp, respectively. Such a preference is not seen for the guanine-attached labels which can only reach the terminus up to a distance of 3 base pairs. Shortening the linker for uracil-attached labels to 19.5 Å did not significantly affect the results (Supplementary Fig. 3.16). To correlate the AV predictions with our experimental results, we plot the scaled residual anisotropies on the right axis (Fig. 3.4A), which show good qualitative agreement with the AV predictions for both uracil and guanine attached labels. In addition to the base pair distance, the effect of the rotation of the labeling position around the helical axis is

uracil show a drastically reduced accessibility of the terminal guanine compared to near-3'-end labels. This is exemplified by the Alexa488 labels b33 on the 3'-end (magenta) and b5 on the 5'-end (orange), both placed at a distance of 4 bp. The larger interaction volume (red) with the terminal guanine base (yellow) for the near-3'-end label is confirmed by the higher fraction of statically quenched species.

3.3.8 The behavior of Cy5 labels depends on the labeling strategy

To understand the differences between uracil and guanine Cy5 labels, we visualize the AVs of the guanine label d27 and the uracil label d26 in Fig. 3.4B. The AVs show clear differences in their shape and orientation due to the different attachment points on the bases (Fig. 3.1B). The uracil label emerges from the deep major groove which results in a restricted AV whose shape follows the groove. In contrast, the guanine label emerges from the shallow minor groove on the opposite side of the RNA, resulting in a large spherical AV. While the shapes of the AVs would suggest a higher rotational freedom for the guanine label, our data shows slightly higher lifetimes and residual anisotropies for internal guanine labels (average $\langle \tau \rangle_x = 1.2 \pm 0.01$, $r_\infty = 0.18 \pm 0.01$) compared to uracil labels (average $\langle \tau \rangle_x = 1.17 \pm 0.04$, $r_\infty = 0.14 \pm 0.02$, Fig. 3.2A), indicating that more interactions with the biomolecular surface occur for guanine labels. This observation is supported by a quantification of the accessible contact volume, defined as the AV within 7 Å distance to the RNA surface, which is found to be larger for internal Cy5 dyes coupled to guanine compared to uracil (Supplementary Fig. 3.17). The visual inspection of the AVs of the labels d26 and d27 in Fig. 3.4B also clearly show that the guanine label does not reach the helix end, while there is significant overlap of the AV with the terminal base pair for the uracil label, confirmed by the higher lifetime and residual anisotropy of the uracil label.

The largest difference between the two labeling strategies stems from the attachment point on the RNA. For uracil, the fluorophore is attached via the C5 atom situated in the major groove, while the dye is attached to guanine via the N2 atom located in the minor groove (Fig. 3.1C). The major and minor grooves differ in their ligand accessibility, polarity, hydrophobicity, and groove structure^{25,245,246}. The deep and narrow major groove is lined with negatively charged phosphate groups, while the wide and shallow minor groove exposes the hydroxyl groups of the sugar backbone (Supplementary Fig. 3.18)²⁴⁷. As the linker exits the groove, steric constraints render it unlikely to fold back into the

groove that it emerged from. Instead, the linker points the dye towards the neighboring grooves. Uracil-attached dyes thus access the neighboring minor groove, while guanine-attached dyes can reach the neighboring major groove. This potentially also facilitates sticking interactions with the sugar-phosphate backbone through hydrogen bonding and electrostatic interactions. For uracil labels, the linker exiting the narrow major groove points towards the 3'-end due to the steric clashes within the groove, while for guanine labels no such preference is dictated by the shallow and wide minor groove (Fig. 3.1C). In addition to the reduced linker length for guanine labels (19.5 Å for guanine compared to 22.0 Å for uracil coupled labels, Fig. 3.1B), this “pointing effect” of the major groove is the main reason for the different terminal base pair accessibilities of uracil labels for the 3' and 5'-ends (Fig. 3.4A). These results exemplify how different labeling chemistries can lead to drastically different local environments with varying accessibilities to trapping sites on the biomolecular surface.

3.3.9 PET quenching of Alexa488 correlates with terminal guanine accessibility

We detected a high propensity for static quenching by PET for Alexa488 labels close to the helix ends, whereby higher quenching was observed for labels close to the 3'-end. To understand this observation, we quantify the accessibility of the terminal guanine base for Alexa488 labels (linker length 20 Å, Fig. 3.4C). Mimicking the experimental construct, we place a cytosine on the 5'-end and a guanine on the 3'-end of the fluorophore-labeled strand (5'C/3'G model). A slightly higher accessibility of the terminal guanine is observed for 3'-end labels due to the 3'-end pointing effect of uracil-attached dyes, which reach up to 5 bp compared to 4 bp for 5'-end labels (Fig. 3.4A). This effect is absent for guanine-attached labels (Supplementary Fig. 3.19). To compare the AV predictions with the experiment, we scale the fraction of statically quenched dyes (f_{static}) to the terminal guanine accessibility (Fig. 3.4A, bottom panel). As predicted by the model, higher static quenching is observed at short distances of 3-4 base pairs for 3'-end labels compared to 5'-end labels. To investigate the impact of the terminal sequence, we additionally calculated the terminal guanine accessibility for a duplex where the G is placed on the 5'-end of the labeled strand and C on the 3'-end, although this geometry has not been assessed experimentally in this study (5'G/3'C model, Supplementary Fig. 3.20). This

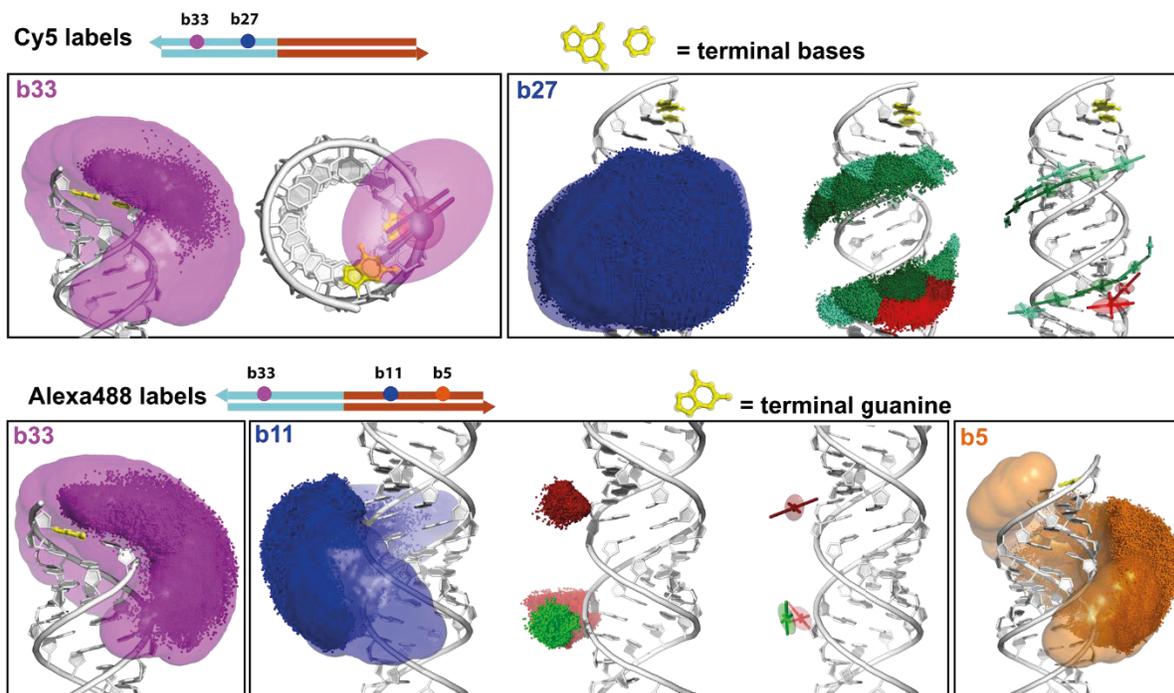
model predicts an overall higher terminal guanine accessibility compared to the 5'C/3'G model with a longer reach on the 3'-end compared to the 5'-end (up to 8 bp and 5 bp, respectively). The consistently higher overlap at the 3'-end for both model duplexes implies that the increased overlap due to the 3'-end pointing effect is more important than the position of the guanine within the terminal base pair.

An exemplary comparison of a near-3'-end and near-5'-end Alexa488 label, both situated at a distance of 4 bp from the helix end, is shown in Fig. 3.4C. Despite their similar relative position on the RNA helix, the AV of the 3'-end label b33 clearly shows a high degree of overlap with the terminal guanine ($f_{\text{static}} = 27\%$), whereas for the near-5'-end label b5 only a small fraction of the AV is in reach of the guanine base ($f_{\text{static}} = 13\%$). Two effects contribute to this difference. First, due to the attachment point at the C5 atom of uracil, the linker is oriented towards the 3'-end, increasing the overlap for the 3'-end label b33. Second, the terminal guanine is positioned on the same strand for the 3'-end label, while it is on the opposite strand for the 5'-end label, which significantly reduces the overlap for position b5 while still allowing sufficient overlap for position b33.

3.3.10 MD simulations explore RNA-dye interaction modes

To obtain atomistic insights into potential interaction modes, Dr. Christian Hanke performed all-atom molecular dynamics (MD) simulations of tethered dyes on RNA at internal positions and close to the 3' and 5'-ends (positions b5, b11 and b33 for Alexa488; positions b27 and b33 for Cy5, Fig. 3.5A). Five independent simulations of 8 μs length each were performed for each of the different labeling positions, starting from an extended linker configuration (see Methods). In the simulations, the internal labels at position b27 (Cy5) and b11 (Alexa488) as well as the near-5'-label b5 (Alexa488) explore a volume that corresponds well to the estimated AV (Fig. 3.5A). On the other hand, the spatial distribution for the near-3'-labels at position b33 are strongly biased towards the helix end for both fluorophores due to long-lived (>100 ns) sticking interactions with the terminal base pair (Fig. 3.5A and Supplementary Fig. 3.21). While the internal labels showed less interactions overall, we still observed frequent short-lived (10-100 ns) interactions with the RNA surface at these positions (Supplementary Fig. 3.21).

A Dye positional distribution and dipole orientations at different labeling sites



B Major interaction modes from MD simulations

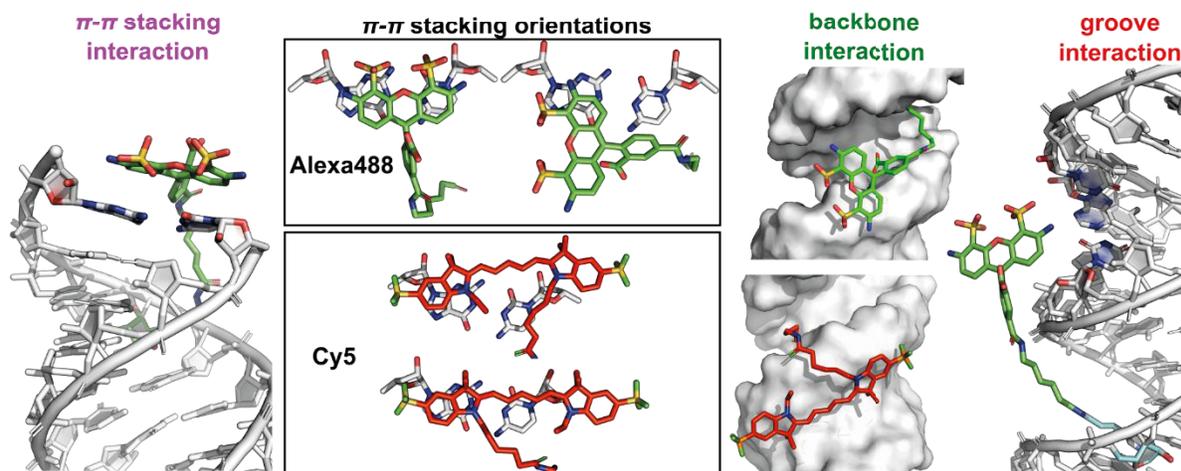


Figure 3.5. Molecular dynamics simulations provide atomistic insights into dye-RNA interaction modes. A) Distribution of dye positions and preferential dipole orientations. For representative Cy5 and Alexa488 labels, the spatial distributions sampled during the MD simulations are shown with the predicted AVs overlaid as transparent clouds. The scatter points represent the location of the C19 atom for Cy5 and the O3 atom for Alexa488 (Supplementary Fig. 3.22) at 20 ps intervals. Identified clusters are shown as ellipsoids and colored according to the type of dye-RNA interaction (green: backbone interaction, red: groove-interacting and magenta: π -stacking states). The ellipsoids indicate the width of the spatial distributions within the clusters. The rods show the preferential orientation of the transition dipole moment of the dye cluster (see Supplementary Fig. 3.22 for the definition of the dipole

moments). **B)** The major RNA-dye interaction modes observed in MD simulations include dye stacking interactions with the electron rich π systems of the exposed nucleobases at the termini, and interactions with the phosphate backbone or the neighboring grooves (right). For both dyes, stacking to the exposed terminal bases occurs in two distinct conformations.

To obtain more detailed insights into the potential interaction modes, Dr. Oleg Opanasyuk performed a cluster analysis of the dye positions to identify stable interactions (Fig. 3.5A). For each identified cluster, we compute the average position and variance of the spatial distribution (displayed as an ellipsoid) as well as the preferential orientation of the transition dipole moment (see Methods and Supplementary Fig. 3.22). For the labels at position b33, the main clusters were found in contact with the terminal base pairs which show narrow spatial distributions and a strong confinement of the dipole moment (Fig. 3.5A and Supplementary Fig. 3.23). A more complex picture arises for the internal labels. For Cy5 at position b27, many clusters were identified along the sugar-phosphate backbone (colored in green), with the transition dipole moment aligning along the backbone. Additionally, some minor states were found within the neighboring minor groove (red). In contrast, Alexa488 at position b11 predominantly formed stable interactions with the neighboring minor grooves or the backbone at defined positions. Corresponding results were obtained for the Alexa488 label at position b5 (Supplementary Fig. 3.23). These results suggest that the behavior of Cy5 at internal positions is dictated by unspecific interactions with the backbone, leading to a diffuse distribution of interacting states, while Alexa488 forms specific interactions with the bases in the neighboring groove or the backbone.

To further investigate the nature of the detected interactions, we show snapshots of the stable binding modes observed in the MD simulations in Fig. 3.5B. Due to the length of the linker, the Cy5 label at position b33 can only reach the terminal base pair from the proximal side where it stacks in two conformations that are rotated by 180° with respect to each other. The long conjugated π -system in the center of the chromophore prohibits that both indolenine groups bind the terminal base pair at the same time. In the two conformations, π - π stacking occurred predominantly with the terminal guanine with no preference for either of the indolenine groups of the Cy5 chromophore. A similar situation is observed for the Alexa488 label at the same position, which also reaches the terminal base pair from the proximal side. Stacking between the extended π -system of the

chromophore and the exposed terminal base pair occurred in two distinct conformations. Due to the smaller size of the chromophore, it can fully stack to the exposed terminal base pair akin to an additional base pair. Less frequently we also observed a partially stacked conformation where the chromophore only interacted with the terminal guanine. Both conformations should offer sufficient orbital overlap for highly efficient PET quenching.

For the internal label positions, interactions occurred either with the nearby RNA backbone or the neighboring minor grooves. Cy5 showed stable interactions of its two heterocycles with the sugar-phosphate backbone of the neighboring minor groove, either on one side or spanning across the minor groove (Fig. 3.5B). Such interactions are potentially mediated by hydrogen bonding between the sulfo groups of the fluorophore and the sugar backbone. On the other hand, stable interactions of Alexa488 at position b5 with the minor groove predominantly occurred in a heads-on conformation wherein the amino group of the chromophore potentially forms hydrogen bonds with the exposed bases within the groove (Fig. 3.5B). Different to the unspecific backbone interactions for Cy5, Alexa488 thus clearly formed stable interactions with the bases in the minor groove.

3.3.11 Origin of PET quenching for internal labels

The observation of dynamic quenching for internal Alexa488 raises the question of how the dye can access these guanines in a way that provides sufficient overlap of the electron orbitals for PET. Intercalation of the labeled dye with the nucleobases is unlikely due to the bulky and highly charged sulfo groups. Internal guanines could potentially unpair with their complementary base and rotate out of the nucleic acid double helix momentarily (base flipping)^{248,249}. To rule out this effect, we performed control measurements on an RNA duplex where the Alexa488 dye is placed in the vicinity of an internal guanine with a mismatch on the complementary strand (either A, U or G). We observed no reduction of the fluorescence lifetime at ambient temperature, indicating that base flipping is not significant (Supplementary Note 3.1). Instead, we speculate that the accessibility of internal guanines for PET is modulated by sequence-dependent local structural variation and flexibility of the RNA duplex, supported by the broad distributions of the fluorescence lifetimes and residual anisotropies for internal Alexa488 labels ($\langle\tau\rangle_x = 3.3\text{-}3.9$ ns, $r_\infty =$

0.08-0.14). Nucleic acid binding proteins recognize specific binding sites by local structural features of the duplex, such as twist, rise, tilt and roll²⁵⁰⁻²⁵³. It is thus likely that sequence-dependent bending and twisting of the RNA duplex modulates the accessibility of the nucleobases, as corroborated by the high flexibility of the RNA duplex observed in the MD simulations (Supplementary Fig. 3.14). Yet another effect that could enhance the structural flexibility of the RNA and promote PET quenching for labels close to the termini is fraying, i.e. the temporary unpairing of the terminal bases^{254,255}. Evidence for a higher structural flexibility of the helix ends is given by the correlation between the residual anisotropy and the distance to the nearest end for internal labels (Supplementary Fig. 3.24). Lastly, it has been postulated that PET quenching could be mediated over long distances by tunneling of electrons between the π -stacked bases along the helical axis¹²⁵. Indeed, we observed a moderate amount of PET quenching for labeling positions b14 with no immediate guanine in the surrounding (Supplementary Fig. 3.25), where tunneling-mediated PET quenching could play a role.

3.3.12 Effect of the primary sequence

To assess whether variation of the primary sequence could affect the dye properties, we selectively investigate the fluorescence lifetimes and residual anisotropies of internal dyes on duplex molecules (Fig. 3.6A). For Cy5, the narrow distribution of photophysical properties for guanine and uracil labels (with the exception of labeling position b27) suggests a low sensitivity to the primary sequence. On the other hand, internal Alexa488 labels show a large spread of residual anisotropies between 0.08-0.14 and fluorescence lifetimes between 3.3-3.9 ns, with no apparent correlation between the two quantities. Note that this spread is significantly larger than the measurement error obtained from repeated measurements (Supplementary Table 3.8). This indicates that PET quenching by internal guanines is sensitive to the primary sequence, potentially due to variations of the electron orbital overlap in different interaction geometries and local breathing dynamics of the RNA. While the strong PET quenching at the helix ends can be predicted from overlap of the dye accessible volume with nearby guanine bases (Supplementary Fig. 3.25), a more detailed model would be required to predict PET quenching of internal position that accounts for these effects. The broad distribution of the experimental residual anisotropies for internal Alexa488 labels could potentially originate from short-

lived sticking interactions mediated by hydrogen bonding, as supported by the MD simulations (Fig. 3.5B) and an apparent correlation between the residual anisotropy and the accessibility of hydroxyl group on the sugar (Supplementary Fig. 3.26).

3.3.13 Sensing the global structure

So far, we have investigated the effects of the primary and secondary structure of the RNA molecules, i.e., the linear base sequence and the helical structure of A-form RNA with the major and minor grooves. The tertiary structure additionally describes the arrangement of secondary structure elements in space and is especially relevant for the 3WJ and 4WJ molecules where the helical arms can change their orientation based on the coaxial stacking at the junction. Using fluorescent dyes, the tertiary structure of RNA could be probed by sensing the structural flexibility and dynamics at the labeling position or the spatial proximity to the junction or neighboring helices. To assess whether our measurements could provide such information, we compare the residual anisotropies of Cy5 labels on the duplex with the corresponding positions the 3WJ or 4WJ molecules (Fig. 3.6B). Indeed, some positions show significantly lower residual anisotropies on the duplex compared to the 3WJ or 4WJ, indicating that the dyes at these positions are sterically restricted (see label a12 on the 3WJ and d18 on the 4WJ, Fig. 3.6B). For these labels, AIFE occurs due to the significant overlap of the respective AVs with the junction region (Fig. 3.6A, inset), enabling stacking interactions with the exposed bases at the junction or promoting steric clashes with the neighboring helices. Interestingly, the Cy5 label at position a12 shows increased anisotropy for the 3WJ but not the 4WJ (Fig. 3.6A), indicating that the bases at the 3WJ are more exposed compared to the coaxially stacked 4WJ. We did not observe large differences for the residual anisotropies between the different 3WJ molecules studied here (Supplementary Fig. 3.27-28). Correspondingly, the FCS curves of the label a12 on the 3WJ, T1, is different from those on the 4WJ and the duplex (Supplementary Fig. 3.29 and Supplementary Table 3.9).

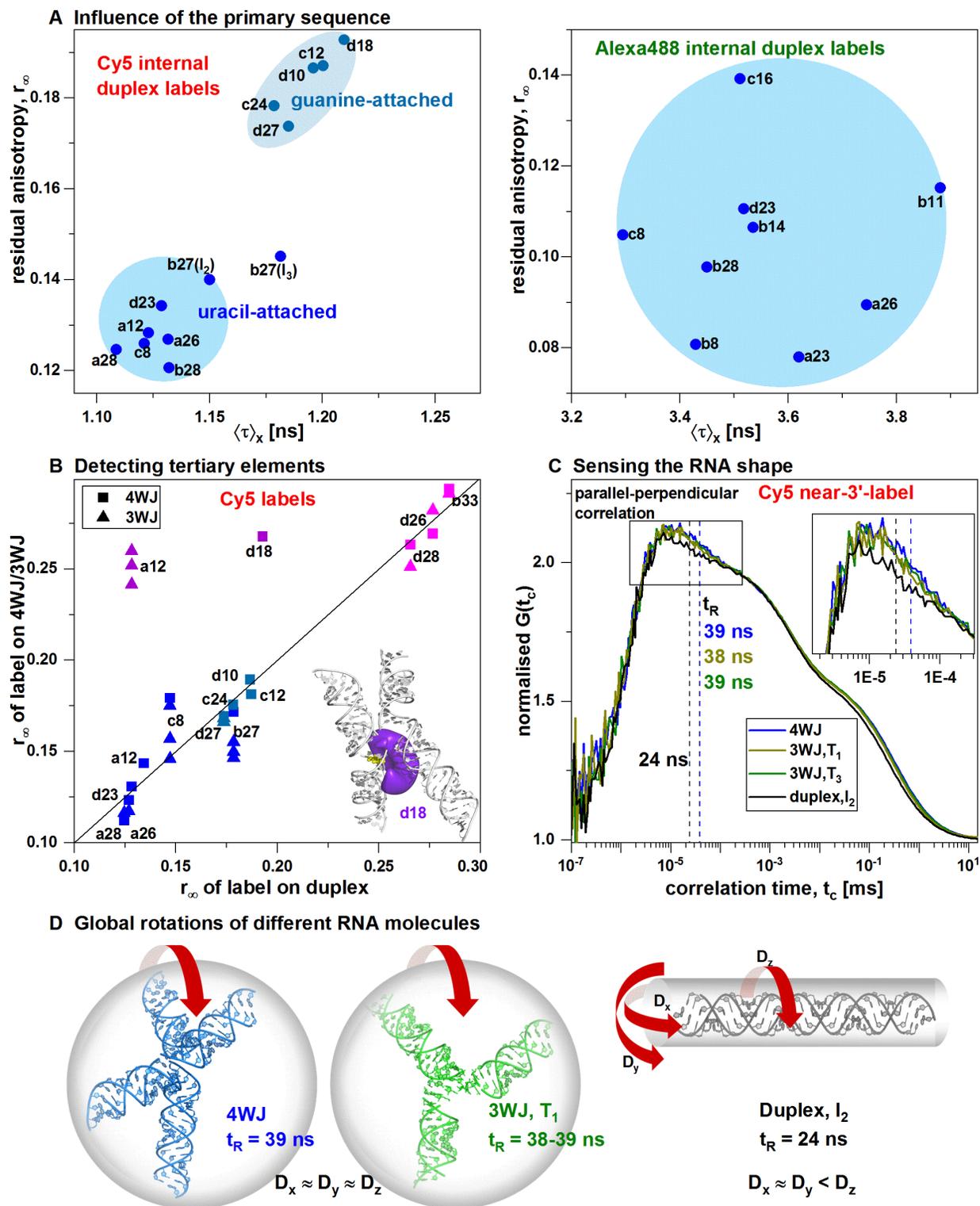


Figure 3.6: Sensing the RNA environment using fluorescent probes. A) The distribution of fluorescence lifetimes and residual anisotropies of internally labeled Cy5 and Alexa488 on RNA duplexes reveals the effect of the primary sequence on the photophysical properties. **B)** Comparison of the residual anisotropy of Cy5 labels on RNA duplexes with the same labeling

position on the 3WJ (triangle) or 4WJ (square) molecules. The diagonal line follows a 1:1 correlation between the axes. Significant differences of the residual anisotropies arise for labels in proximity of the junction as illustrated by the AV of the d18 label on the 4WJ (inset). The labeling site is highlighted in yellow. **C)** Polarization-resolved fluorescence correlation spectroscopy senses the size and shape of RNA molecules. The FCS curves of the end stacked Cy5 label b33 reveal different rotational correlation times for the different types of RNA molecules. The vertical dashed lines indicate the rotational correlation times (t_R) obtained from a fit to the model for a spherical rotor given by eq. 3.29. The inset shows a magnified view of the relevant region for rotational motion. Displayed curves are the parallel-perpendicular correlation functions measured at 11 kW/cm², normalized to the number of bright molecules. **D)** Rotational motions of the different RNA molecules. The 4WJ and 3WJ molecules show approximately isotropic rotation and are well described as spherical rotors. The cylindrical shape of the duplex results in anisotropic rotational diffusion with faster rotation around the helical axis. The red arrows show the directions of rotational diffusion. The length of the longest axis of rotation is approximately similar for all molecules. The structures for the 4WJ (PDBDEV_00000019), the 3WJ T₁²⁵⁶ and the duplex RNA²⁵⁷ have been adopted from the cited sources.

3.3.14 Detecting RNA shape and size by polarization-resolved FCS.

(The theoretical work in this section was performed by Dr. Oleg Opanasyuk)

The rotational tumbling of biomolecules is sensitive to their size and shape and can be predicted based on the atomic structure^{258,259} or coarse-grained representations²⁶⁰. For the molecules studied here, the rotational tumbling times are larger than 20 ns. These times are too long to be accurately resolved by time-resolved anisotropy experiments, which are limited by the excited-state lifetime of the fluorophore. However, the global tumbling motion can be resolved by polarization-resolved fluorescence correlation spectroscopy (pFCS) by investigating the correlation between the fluorescence emission in the parallel and perpendicular detection channels with respect to the polarization of the incident light^{261,262}.

For optimal sensitivity, the rotation of the fluorescent label should be maximally coupled to the rotation of the biomolecule, which can be achieved by using short, rigid linkers^{103,262,263}. Here, instead, we chose the Cy5 label at position b33 due to its high propensity to stack at the helix end. The pFCS curves reveal clear differences between the duplex and the 3WJ/4WJ molecules in the diffusional (0.1-1 ms) and rotational (10-100 ns) part of the correlation function (Fig. 3.6C, Supplementary Table 3.10). By fitting the curves to the model function for a spherical rotor, we obtained global rotational correlation times, ρ_{global} , of 24 ns for the duplex and 38-39 ns for the 3WJ/4WJ molecules. Interestingly, we observed no significant difference between the 3WJ and 4WJ molecules

despite the larger number of nucleotides of the latter, suggesting a similar hydrodynamic radius. Similar results were observed for Cy5 at position c1 and Alexa488 at position b11 (Supplementary Fig. 3.30-31 and Supplementary Table 3.11-12).

To assess the sensitivity of the pFCS curves to the shape and size of the molecule, we computed theoretical estimates of the diffusion tensors using hydrodynamic modeling²⁵⁸ based on the ideal rigid structure of the RNA duplex and FRET-derived structural models for the 3WJ and 4WJ molecules (see Methods)^{256,264}. Using this information, we model the theoretical pFCS correlation functions by accounting for the relative orientation of the fluorophore dipole with respect to the rotation axes of the RNA, as estimated from the MD simulations (Supplementary Note 3.3). The theoretically expected rotation times (defined as $1/6D$ where D is the rotational diffusion coefficient) for the 3WJ and 4WJ molecules are consistently longer (~50-70 ns) compared to the experimentally fitted values (~40 ns), while the theoretical prediction matches well for the RNA duplex (exp.: 20 ns, predicted: 24 ns) (Supplementary Fig. 3.32A-D and Supplementary Table 3.13-14). This indicates that additional processes must contribute to the depolarization of the fluorescence signal for the 3WJ and 4WJ molecules.

The predicted correlation functions of the various 3WJ and 4WJ molecules can be well approximated by the model function for the spherical rotor, while the duplex clearly deviates from the spherical model (Supplementary Fig. 3.32E-H). This is supported by an improved fit quality for the duplex when using the scaled predicted diffusion tensor compared to the spherical model (Supplementary Table 3.14). For the rod-shaped duplex, the rotation around the helical axis should cause less friction than for the other axes, resulting in a faster rotational diffusion coefficient around this axis ($D_z > D_{xy}$, Fig. 3.6D). To test this prediction, we fitted our experimental data to a generalized model for a prolate ellipsoid assuming equal rotational diffusion coefficients in the x-y plane ($D_x = D_y = D_{xy} < D_z$, Supplementary Note 3.3 and Supplementary Fig. 3.33). Indeed, the ratio of D_z to D_{xy} was largest for the duplex ($D_z/D_{xy} = 10.5$, predicted 5.7) and smaller for the 3WJ and 4WJ molecules ($D_z/D_{xy} = 2.5-3.0$, predicted 1.5-1.9, Supplementary Table 3.14). Interestingly, the measured ratio of D_z/D_{xy} exceeds the theoretical estimate approximately by a factor of two for all molecules, which is mostly due to a faster rotation around the helical z-axis compared to the prediction ($\Delta D_z = 3.3$ MHz for the 3WJ and

4WJ, and ~8 MHz for the duplex, Supplementary Table 3.14). This suggests that the faster rotational correlation times observed when the experimental data was fitted by a spherical model originate from additional depolarization processes along the helical axis due to wobbling or twisting motions. In summary, our results show that pFCS informs on the size and shape/sphericity of biomolecules if the rotation of the fluorescent label is sufficiently coupled to the global rotation, as is the case for end-stacked fluorophores. The faster rotational times compared to the theoretical predictions further indicate the presence of additional structural dynamics in the studied 3WJ and 4WJ molecules., which are not accounted for by the rigid model used for theoretical predictions.

3.4 Discussion

In this study, we characterized the dye behavior of Cy5 and Alexa488 fluorophores at various sites and local environments in different RNA molecules and identified distinct dye-RNA interaction modes. π - π stacking interaction between the dye and exposed bases occur predominantly at the helix termini or junctions. The wide distribution of photophysical properties for internal labels indicates interactions with the sugar-phosphate backbone, as well as with the bases within the neighboring major or minor grooves of the double helix. Most of the observed interactions occur over long distances (> 1 nm) and are hence facilitated by the long flexible dye linkers. In addition, the labeling chemistry was found to be a determining factor for the dye behavior. The different attachment points on the uracil and guanine bases determine whether the dye emerge from the major or minor groove for U and G, respectively, from where it can predominantly access the neighboring grooves of the opposite type. Moreover, for uracil labels, the deep major groove directs the dye linker towards the 3'-end, facilitating interactions with the helix ends.

These detailed insights into dye-RNA interactions provide a foundation for sensing the unknown nucleic acid environment using the spectroscopic methods presented here, allowing to, e.g., assess the proximity of the dye to the helix ends or secondary structure elements, estimate the local concentration of guanine, or even sense local structural dynamics and the global structure and shape of the nucleic acid. To this end, we summarize the information provided by the different photophysical observables and

experimental methods and how they report on the microenvironment of the fluorophore and the structure and dynamics of the biomolecule.

- (1) A bathochromic shift of the absorption spectrum generally correlates with dye sticking, most prominently at the ends of the double helix. Thus, a significantly red-shifted absorption spectrum indicates that a fluorescent label is in proximity to exposed bases either at the helix end or a junction where π - π stacking interactions could occur.
- (2) An increase of the residual anisotropy is an indicator for dye trapping or a sterically restricted environment. This effect is most pronounced for π - π -stacking interactions to the helix ends for cyanine labels. Special care must be taken to ensure that the trapped species is not statically quenched, as observed here for Alexa488.
- (3) For cyanine dyes, steric restriction also results in an increase of the fluorescence lifetime and brightness due to association-induced fluorescence enhancement which should correlate with an increase of the residual anisotropy.
- (4) For Alexa488 and other rhodamine or oxazine dyes, a reduction of the fluorescence lifetime indicates dynamic quenching due to a high concentration of partially accessible, buried guanines around the attachment site.
- (5) A reduction of the brightness and steady-state quantum yield of Alexa488 indicates long-lived interactions (>10 ns) with exposed guanine bases at the helix ends or junctions. The presence of static quenching is detected from a comparison of steady state and dynamic quantum yields and/or characteristic spectral shifts of the absorption spectrum.
- (6) FCS is sensitive to any interaction that affect the brightness of the fluorophore, either by quenching or enhancement, and provides quantitative information on the kinetics of the interaction. Possible interactions are revealed by a direct comparison with the FCS curve of the free dye or a dye-labeled reference molecule.

In general, the sensing of cyanine-based dyes by AIFE is independent of the nucleic acid sequence, while the PET-based sensing is specific to guanines. This allows for site-specific control through the placement of guanine bases at positions of interest, e.g., to

detect the accessibility of a specific base at the helix end or detect local melting of the double strand.

MD simulations provided further atomistic insights into potential interactions. Interactions with the helix ends were mediated by π - π stacking interactions which could occur in different geometries for both dyes. Alexa488 formed stable interactions in two conformations which are rotated by 180° with respect to each other, while the two end-stacked conformations of Cy5 involved only one of the two indolenine groups. These simulation results are in excellent agreement with structural studies of end-stacked dyes^{91,92,94-96}. Importantly, many of the interaction modes could be accurately predicted by AV simulations. This includes the lowered accessibility of the helix termini for 5'-end compared to 3'-end labels, the different accessibility of the major and minor grooves between uracil and guanine labels, the local quencher concentration, as well as the proximity to specific interaction sites. Particularly, AV simulations were well suited to predict the interactions with the terminal bases as seen from the good agreement between the predicted overlap from AV simulations and the residual anisotropy for Cy5 labels or the propensity of static quenching for Alexa488 labels (Fig. 3.4A). These results show that accurate predictions of dye-nucleic acid interactions can be obtained from the simple and efficient AV model, reducing the need for computationally expensive MD simulation.

Lastly, we explored the potential of polarization-resolved FCS to sense the size and shape of an RNA molecule. Using a Cy5 label with a high propensity for end-stacking ensured optimal coupling between the depolarization of the fluorescence signal and the rotational motion of the RNA. We detected a higher asymmetry and faster rotational motion for the duplex compared to the 3WJ/4WJ molecules, while the similar rotational correlation times for the 3WJ and 4WJ molecules indicate comparable hydrodynamic radii for these molecules. Rotational correlation times for the 3WJ and 4WJ molecules were also significantly shorter compared to theoretical predictions, suggesting the presence of additional conformational dynamics that lead to a faster depolarization of the fluorescence signal. These results highlight the potential of polarization-resolved FCS to sense the size, shape, and local flexibility of biomolecules.

3.4.1 Consequences of environmental dependence of photophysics for FRET studies.

Förster resonance energy transfer (FRET) is a widely used and powerful method to investigate the structural dynamics of nucleic acids over timescales ranging from picoseconds to minutes/hours on the ensemble or single-molecule level²⁶⁵. In FRET, the efficiency of the energy transfer from a donor to an acceptor fluorophore, coupled to the biomolecule at suitable positions, reports on the interdye distance. A clear understanding of the photophysics, linker dynamics, and interactions of the dye with the host biomolecule is essential to the design, proper interpretation, and quantitative evaluation of FRET experiments, especially if the distance information is to be used for integrative structural modeling and to avoid misinterpretation of photophysics as conformational dynamics^{242,266,267}. Therefore, we examine our results in the context of FRET studies and provide guidelines on how to minimize and correct for the effect of dye-biomolecule interactions.

In a typical FRET experiment, the measured fluorescence intensities in the donor and acceptor channel are corrected for background signal, spectral crosstalk of the donor signal into the acceptor channel, and direct excitation of the acceptor fluorophore by the donor excitation laser. From the corrected fluorescence intensities F_D and F_A , the accurate FRET efficiency is then calculated as:

$$E = \frac{F_A}{\gamma F_D + F_A} = \frac{1}{1 + \left(\frac{R_{DA}}{R_0}\right)^6}, \quad (3.1)$$

which relates the FRET efficiency E to the interdye distance R_{DA} . The estimation of accurate FRET efficiencies depends on several setup and dye dependent correction factors^{268,269}, of which the γ -factor is generally most important (and most challenging to determine). It accounts for the different detection efficiencies of the donor and acceptor fluorophore and is given by:

$$\gamma = \frac{g_{R|A} \Phi_{F,A}}{g_{G|D} \Phi_{F,D}}, \quad (3.2)$$

where $g_{R|A}$ and $g_{G|D}$ are the detection efficiencies of the donor and acceptor emission in the red and green detection channels, respectively. $\Phi_{F,A}$ and $\Phi_{F,D}$ are the quantum yields of the acceptor and donor fluorophores. Note that if the quantum yield of fluorophores that undergo transitions to dark states (e.g., triplet states or photoisomers) is estimated from the fluorescence lifetime, it needs to be additionally corrected for the fraction of molecules in the bright state a_b by $\Phi_F^{\text{eff}} = a_b \Phi_F$ (Supplementary Table 3.15).

The accuracy of the distance derived from a FRET experiment crucially depends on the Förster radius R_0 , the characteristic length scale of the energy transfer process. This quantity depends on the spectral overlap between the donor and acceptor fluorophores, J , the refractive index of the surrounding medium, n , the quantum yield of the donor fluorophore, $\Phi_{F,D}$, and the relative orientation of the donor emission and acceptor absorption transition dipole moments, expressed in the orientational factor κ^2 :

$$R_0 \sim \sqrt[6]{J \Phi_{F,D} \kappa^2 n^{-4}}. \quad (3.3)$$

In a previous benchmark study, the joint error of these quantities has been estimated to introduce a distance uncertainty of 5%²⁶⁸. Often, a single value for the Förster radius is estimated in FRET studies to convert the measured FRET efficiencies between dyes in different environments into distances.

Table 3.3. Estimation of Förster radius R_0 and correction factor γ for representative Alexa488-Cy5 donor-acceptor pairs with donor and acceptor at different labeling sites.

| Donor(D)- Acceptor(A) combination | Overlap integral, J [M ⁻¹ cm ⁻¹ nm ⁴] 10 ¹⁵ | Reduced Förster radius, R _{0J} [Å] | $\Phi_{F,D}$ | Förster radius, R ₀ [Å] | $\Phi_{F,A}$ | $\Phi_{F,D} /$ $\Phi_{F,A}$ | Detection efficiency ratio, g _{R A} /g _{G D} | γ -factor |
|---|---|--|--------------|--|--------------|--------------------------------|---|------------------|
| Free dyes | 1.27 | 55.5 | 0.96 | 51.5 | 0.30 | 3.17 | 0.74 | 0.23 |
| b11(D)-b27(A) | 1.36 | 56.1 | 0.90 | 51.5 | 0.37 | 2.43 | 0.74 | 0.30 |
| b33(D)-b27(A) | 1.35 | 56.0 | 0.35 | 44.0 | 0.37 | 0.95 | 0.74 | 0.78 |
| b11(D)-b33(A) | 1.27 | 55.5 | 0.90 | 50.9 | 0.48 | 1.88 | 0.74 | 0.39 |

The overlap integral J quantifies the spectral overlap between the emission spectrum of D and absorption spectrum of A, where D is the donor and A is the acceptor, computed according to eq. 3.35. R₀ is the Förster radius calculated

according to eq. 3.34 under the assumption of $\kappa^2=2/3$, and R_{0J} is the reduced Förster radius calculated according to eq. 3.37 by omitting the contributions of the donor quantum yield $\Phi_{F,D}$ and the orientational factor κ^2 . $\Phi_{F,D}$ and $\Phi_{F,A}$ are the fluorescence quantum yields of the donor and acceptor, respectively. $g_{R|A}$ and $g_{G|D}$ are the detection efficiencies of acceptor and donor dyes, respectively. The correction factor γ is applied for the normalization of fluorescence quantum yields and the detection efficiencies as given by eq. 3.2.

We first assess how much the observed variation of the photophysical properties contributes to the uncertainty of R_0 , which evaluates to 51.5 Å for the free dyes. We consider the ideal situation based on the parameters of the free dyes and compare it to representative dye pairs for two internal labels, and for two combinations of an internal and a near-3'-end label (Table 3.3). The overlap integral can deviate by up to ~15% compared to the free dyes due to the large spectral shift of the acceptor at position b33, however the propagated error on the Förster radius remains below 1 Å due to the sixth-root dependence in eq. 3.3. The donor fluorescence quantum yield $\Phi_{F,D}$, on the other hand, shows larger variation between 0.96 for the free dye and 0.35 for the near-3'-end label, resulting in deviations of the Förster radius of up to 8 Å. The variation of the photophysical properties also affect the γ -factor, and thus the inferred interdye distance R_{DA} . As the detection efficiency ratio is only affected by spectral shifts, it shows a variation of less than 5% (distance uncertainty < 1 Å). On the other hand, the large variations of the donor and acceptor quantum yields due to PET quenching and AIFE, respectively, result in a variation of the γ -factor by more than 2-fold between the exemplary dye pairs, corresponding to an approximate relative distance error of 10%. Taken together, while the observed spectral shifts affect both the Förster radius and setup-dependent correction factors, the error on the FRET-derived distances remains minor. Contrarily, the large variations of the donor fluorescence quantum yield propagate into a significant distance uncertainty if they remain unaccounted for. It should be noted that, technically, the donor fluorescence quantum yield is not required for the conversion of the measured fluorescence intensities into physical interdye distances (see Supplementary Note 3.2)^{270,271}. Thus, while any error related to $\Phi_{F,D}$ could in principle be avoided, care must be taken to account for the position-specific photophysical properties of the donor fluorophore if fluorescence intensities are first converted into accurate FRET efficiencies that are interpreted in terms of physical distances at a later step of the analysis.

In addition to the photophysical effects, we observed strong binding of the fluorophore to the helix termini and junctions in addition to weak interactions with the grooves and backbone. Such interactions can affect FRET experiments in three ways by 1) by biasing the spatial distribution of the fluorophore, 2) causing a deviation of the orientation factor κ^2 from its isotropic average, and 3) giving rise to a correlation between the interdyde distance and κ^2 , which will be discussed below.

First, the position of the fluorophore deviates from the average position estimated by AV calculations. Previously, dye-surface interactions have been treated in the AV framework by weighting the occupancy within the AV towards the surface by defining a contact volume²⁴². This unspecific treatment of surface interactions, however, does not account correctly for the specific interaction sites observed here, e.g., at the helix termini. In this case, weighting towards the surface layer would move the average dye position closer to the helical axis, while the interactions with the helix terminus results in an actual displacement *along* the helical axis towards the helix end. The accuracy of AV predictions could thus be further improved by defining specific interaction sites on the biomolecular surface to account for the spatial displacement of the fluorophore, which can be identified *in silico* by MD simulations⁹⁰ or experimentally using X-ray crystallography⁹² or NMR spectroscopy^{91,272} of fluorescently labeled biomolecules.

Second, the trapping interactions affect the orientation factor κ^2 which depends on the mutual orientation of the donor emission and acceptor excitation transition dipole moments. Generally, it is assumed that the dyes rotate freely on a timescale that is much faster than the excited state lifetime of the donor (<100 ps), leading to the isotropic average of $\langle \kappa^2 \rangle_{\text{iso}} = 2/3$. For the case of steric restriction, the uncertainty of κ^2 can be estimated from the residual anisotropy^{103,273,274}, with a typical value of 5%^{268,275}. In this model, axially symmetric distributions of the transition dipole orientation are assumed, and the residual anisotropy is interpreted in terms of the opening angle of a cone that describes the possible orientations of the transition dipole. The results presented here, however, reveal a drastically different behavior where the dye alternates between a freely rotating state and trapped conformations at distinct binding sites on the surface. This is especially true for the end-stacked conformations, where it has previously been shown

that a modulation of the FRET efficiency between two end-stacked dyes can arise due to the restricted dipole orientations^{94,95}.

Third, the trapping of dyes in preferential conformations gives rise to a correlation between the dye position and the mutual orientation of their dipole moments, resulting in a dependence between the donor-acceptor distance and the orientational factor κ^2 ²⁷⁶. An accurate treatment of this effect requires a characterization of the preferential dipole orientation of the trapped conformation and needs to account for the correlation between trapped states of the donor and acceptor fluorophores. The observed switching between free rotation and long-lived trapped states suggests a model where the residual anisotropy can be interpreted in terms of the fraction of trapped dyes. The variation of κ^2 (and hence R_0) may then be estimated by considering all combinations of free or trapped donor and acceptor dyes. While the precise assessment of the κ^2 error would require further knowledge of the dipole orientation in the trapped conformations, an upper-bound estimate can be obtained under the assumption of the least favorable dipole orientation for the trapped conformation ($\kappa^2 = 0$). This “diffusion-with-traps” model is expected to provide a more accurate and realistic estimation of the distance uncertainty compared to current approaches and is the subject of future work.

Finally, our results reveal a correlation between dye trapping and the brightness of the dye, which will likewise cause significant errors in the FRET-derived distances. Rhodamine-based labels may be quenched by guanine nucleobase through PET, resulting in a near depletion of the fluorescence signal, while for cyanine labels, the trapped state shows increased brightness due to the hindrance of the cis-trans isomerization. A refined AV model should thus not only account for the preferential occupancy of interaction sites and their predominant dipole orientations, but also consider position-dependent changes of the brightness, quantum yield and excited-state lifetime. Based on the presented insights, we summarize in Table 3.4 the effect of the local environment on the dye properties and the FRET process, outline how the different artifacts can be detected, and formulate guidelines on how to minimize related errors.

Table 3.4. Implications of local dye interactions for FRET studies.

| RNA-dye interaction | Affected dye families | How to detect | Magnitude of interaction | Implications for FRET | Magnitude of effect on FRET | How to minimize errors |
|--|-----------------------|---|---|--|--|---|
| Dye trapping at helix ends or junction | all | High residual anisotropy | Residual anisotropies differ from that of internal labels, $r_{\text{ex}} > 0.25$. | Error in distance calculations due to deviation in κ^2 from 2/3 and non-uniform dye distribution. | Distance error of up to 4 Å observed in FRET measurements ^a | Use end-weighted AV model for FRET distance calculations. Assess uncertainty due to κ^2 . |
| π - π stacking interactions | all | Red shift of absorption and emission spectra | Peak shifts up to 5 nm relative to free dye. | Error in spectral overlap integral. | Minor (<1 Å) | Measure spectra of fluorophores attached to biomolecule using single-labeled control. |
| Dynamic quenching of the donor by PET | Rhodamine, oxazine | Reduction of the fluorescence lifetime | Up to 40% reduced quantum yield | Deviation of Förster radius and correction factors | Large if undetected, minor if correctly identified | Use correct quantum yield for Förster radius calculation and γ correction factor determination |
| Static quenching of the donor by PET | Rhodamine, oxazine | Compare steady state and dynamic quantum yields | Up to 42% of statically quenched species detected | Lower detected signal, but no deviation of the FRET efficiency. | Lower signal to noise ratio can affect analysis | - |
| Modulation of cis-trans photoisomerization of the acceptor | cyanine | Photoisomerization is seen as a bunching term in FCS measurements | Population of the dark state varies between 40-60%, depending on excitation power. | Fluorescence quantum yield obtained from lifetime is overestimated | Large if undetected, minor if correctly identified | Correct $\Phi_{\text{F,A}}$ for the fraction of dark species involved in cis trans relaxation |

^a Estimated from FRET measurements on double stranded RNAs by comparison of the experimental distances to AV predictions (data not shown).

3.5 Conclusion and Outlook

In this study, we provided a detailed characterization of the surface interactions of covalently attached fluorescent labels on RNA, revealing distinct photophysical fingerprints of the different microenvironments on the RNA and highlighting the potential of single fluorescent labels to inform on the structural dynamics and hydrodynamic properties of biomacromolecules. Our results have direct implications for the modeling of interdyer distance in FRET studies. In addition to quantifying the sterically accessible volume, improved dye models should account for specific interaction sites on the biomolecular surface and the respective preferential orientation of the fluorophore's dipole moment. Improved modeling of the spatiotemporal dynamics of fluorescent labels will enable researchers to avoid photophysical artifacts during the design of experiments and significantly improve the accuracy of FRET experiments and FRET-derived structural models.

3.6 Materials and methods

General. The studied RNA molecules are derived from the RNA four-way junction found in the hairpin ribozyme (PDB-ID: 1M5K, Supplementary Fig. 3.1, 3-7 bp from the junction) which was extended to obtain the 4WJ molecule (Fig. 3.1A). RNA 3WJ and duplexes are derived from the 4WJ by omission of one or two of the helices, respectively. The naming of the labels corresponds to the analogous position in the 4WJ molecule. In our study, we employ the C5 position for uracil labeling and N2 position for Cy5 dye coupling at guanine (Fig.1B) where the amino-modified dU or dG is coupled with the NHS ester of the dye.

Unlabeled and fluorescently labeled RNA was obtained from Purimex (Grebenstein, Germany), IBA (Göttingen, Germany), Biomers (Ulm, Germany) or Dharmacon (Lafayette, CO, USA) (see Fig. 3.1 and Supplementary Fig. 3.2 for an overview of the RNA sequences, dye labeling positions and the structures of modified U and G nucleotides with linker and dye). All RNA was purified by polyacrylamide gel electrophoresis. In order to minimize the sticking of dyes to the negatively charged RNA helix, sulfonated Cy5 labels have been used ²⁷⁷, with the exception of the terminally-labeled duplex (see Supplementary Fig. 3.6). Free-dye reference measurements were performed with the 5-isomer of Alexa488 (Thermo Fischer Scientific) and Cy5-NHS-ester

(GE Healthcare Life Sciences), which was hydrolyzed before the experiment to obtain the carboxylic acid.

Hybridization of RNA. Hybridization was performed in phosphate buffer containing 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$, 100 mM KCl, and 20 mM MgCl_2 at pH 6.5. The same buffer was used later for all measurements. Concentrations of labeled strands for the hybridization were in the range of 0.5 to 1.5 μM . A three-fold excess of unlabeled strands was used. Samples were heated to 85°C inside a thermo-cycler (primus 96 advanced, peqLab, Erlangen, Germany) at 0.1 °C/s, immediately cooled down to 52 °C, kept for 2 h and quickly cooled down to and stored at 4°C until the measurement.

Fluorescence lifetime and anisotropy decay measurements and analysis.

Fluorescence decays were recorded using a FluoTime300 spectrometer (PicoQuant, Germany) equipped with a white light laser (SuperK Extreme EXW-12, NKT Photonics, Denmark) at a repetition rate of 20 MHz. Filters with cutoff wavelengths 495 nm and 645 nm (Farbglas FGL 495 and FGL 645, Thorlabs, New Jersey, USA) were used to reduce the contribution of the scattered light in Alexa488 and Cy5 samples respectively. All samples were measured in quartz ultra-micro cuvettes (Helma #105.252.85.40) with a total volume of 20 μl . A LUDOX scattering solution was used to record instrument response function (IRF) at 5% transmission. For Alexa488 samples, the excitation and emission wavelengths were set at 488 nm and 520 nm. For Cy5 samples, they were set at 635 nm and 665 nm, respectively. Temperature was maintained at 20°C with a temperature controller (TC 425, Quantum Northwest, USA).

To estimate the fluorescence lifetimes and anisotropies, the measured fluorescence decays (F_{exp}) with the polarizer-detector settings VV (parallel-parallel) and VH (parallel-perpendicular), were fitted jointly with the corresponding VM (parallel-magic angle) fluorescence decay.

The fluorescence decays $f(t)$ of single-labeled RNA molecules were described by

$$f(t) = \sum_i x_i e^{-t/\tau_i}, \quad (3.4)$$

where τ_i and x_i are the fluorescence lifetimes and corresponding species fractions with $\sum_i x_i = 1$. Up to three lifetime components were used in the analysis. The number of components chosen were always the minimum number required to achieve a reduced chi-squared of 1.15 and flat residuals. The species weighted average lifetime is then defined by:

$$\langle \tau \rangle_x = \sum_i x_i \tau_i. \quad (3.5)$$

For the fluorescence lifetime measurements, decays were measured under magic angle condition with the detection polarizer set at an angle of 54.7° (VM). In this case, the model decay is given by:

$$F_{exp,VM}(t) = f_0 \cdot IRF_{VM}(t - t_{shift}) * f(t) + s_{VM}(t) \cdot IRF_{VM}(t) + bg_{VM}, \quad (3.6)$$

where f_0 is the amplitude scaling factor, $IRF_{VM}(t - t_{shift})$ is the instrument response function shifted by the time t_{shift} , s_{VM} is the scatter amplitude and bg_{VM} is the constant background level. The * sign designates circular convolution.

The anisotropy decays were described by:

$$r(t) = \sum_i r_i e^{-t/\rho_i}, \quad (3.7)$$

where ρ_i and r_i are the rotational correlation times and respective amplitudes. The sum of the anisotropy amplitudes is equal to the fundamental anisotropy $\sum_i r_i = r_0$, where $r_0 = 0.38$. Up to three rotational components were used for the analysis. The residual anisotropy r_∞ is defined as the anisotropy amplitude of the component with the longest rotational correlation time (ρ_{global}) which reflects the overall tumbling motion of the molecule (global motion).

The model functions for the parallel (VV) and perpendicular (VH) fluorescence decays are then given by:

$$f_{VV}(t) = f(t)[1 + 2r(t)] = \left[\sum_i x^{(i)} e^{-t/\tau_i} \right] \left[1 + 2 \sum_j r_j e^{-t/\rho_j} \right], \quad (3.8)$$

$$f_{VH}(t) = f(t)[1 - r(t)] = \left[\sum_i x^{(i)} e^{-t/\tau_i} \right] \left[1 - \sum_j r_j e^{-t/\rho_j} \right]. \quad (3.9)$$

In analogy to eq. 3.6, the model decays are then given by:

$$F_{exp,VV}(t) = f_0 IRF_{VV}(t - t_{shift}) * f_{VV}(t) + s_{VV}(t) \cdot IRF_{VV}(t) + bg_{VV}, \quad (3.10)$$

$$F_{exp,VH}(t) = g_{VV/VH} f_0 IRF_{VH}(t - t_{shift}) * f_{VH}(t) + s_{VH}(t) \cdot IRF_{VH}(t) + bg_{VH}, \quad (3.11)$$

where $g_{VV/VH}$ is g-factor, f_0 is the amplitude scaling factor, $IRF_{VV}(t - t_{shift})$ and $IRF_{VH}(t - t_{shift})$ are the instrument response functions that are shifted by the time t_{shift} , s is the scatter amplitude and bg is the constant background level. The subscripts VV and VH indicate the excitation and detection polarizer settings (parallel-parallel and parallel-perpendicular, respectively).

Least-squares fitting was performed using the Levenberg-Marquardt method²⁷⁸ to estimate the relevant times and their fractions by minimizing the reduced chi-squared defined as:

$$\chi_{red.}^2 = \frac{1}{N_{data} - N_{param}} \sum_{k=1}^{N_{data}} \frac{(F_{model}^{(k)} - F_{exp}^{(k)})^2}{F_{exp}^{(k)}}, \quad (3.12)$$

where $F_{exp}^{(k)}$ and $F_{model}^{(k)}$ are the intensities of the measured and model decay in TCSPC bin k , and N_{data} and N_{param} are the number of bins in the TCSPC histogram and the number of fit parameters, respectively. The measurement uncertainty (i.e., the weights in the numerator) are approximated as $\sigma_k^2 = F_{exp}^{(k)}$ based on the Poisson statistics of photon counting. Details of further corrections applied are given in Chapter 7.

Absorption and emission spectral measurements: Absorption spectra were measured with a UV-Vis spectrometer (Cary 4000, Agilent Technologies) and fluorescence spectra with a fluorescence spectrometer (FluoroMax-4, Horiba Scientific) at room temperature in measurement buffer. For absorbance measurements, the optical density was kept below 0.01 to avoid inner filter effects. To exclude polarization effects, the fluorescence was collected at a 90° angle with the emission polarizer set to the magic angle (54.7°). For emission spectra of Alexa488 labels, the samples were excited at 460 nm and for the Cy5 labels at 614 nm to record the complete spectra. In a second step, the Cy5 dye was excited closer to the maximum at 635 nm to rule out effects due to excitation wavelength dependencies. Comparison of both spectra after normalization yielded exact overlap in all cases.

Dynamic fluorescence quantum yield ($\Phi_{F,dyn}$). In the absence of any static quenching process, the species-weighted average fluorescence lifetime $\langle\tau\rangle_x$ of a dye is proportional to its fluorescence quantum yield. Since this estimator of the quantum yield only accounts for dynamic quenching, we call it the dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) defined by:

$$\Phi_{F,dyn} = \Phi_{F,ref} \cdot \frac{\langle\tau\rangle_x}{\langle\tau\rangle_{x,ref}}, \quad (3.13)$$

where $\Phi_{F,ref}$ and $\langle\tau\rangle_{x,ref}$ are the fluorescence quantum yield and species-averaged fluorescence lifetimes of a reference sample with known quantum yield. For Alexa488, reference values of $\langle\tau\rangle_x = 4.05$ ns and $\Phi_{F,ref} = 0.964$ were used. For Cy5, $\langle\tau\rangle_x = 0.95$ ns and $\Phi_{F,ref} = 0.30$ were used. Reference quantum yields were determined as described below.

Determination of the steady-state fluorescence quantum yield (Φ_F). Two different approaches were used to determine the steady-state quantum yield.

(1) Conventional method based on the fluorescence intensity:

These measurements were performed according to reference ²⁷⁹ by comparing the emission of the sample with a spectrally similar standard reference dye. The steady-state quantum yield is determined from:

$$\Phi_F = \Phi_{F,ref} \cdot \frac{F_s}{F_{ref}} \cdot \frac{A_{ref}}{A_s} \cdot \left[\frac{n_s}{n_{ref}} \right]^2, \quad (3.14)$$

where Φ_F and $\Phi_{F,ref}$ are the steady state fluorescence quantum yields of the sample and reference, F is the integrated intensity of the fluorescence emission spectra; A is the measured absorbance at the excitation wavelength, n is the refractive index of the solution of measurement and the subscripts s and ref refer to the sample and reference, respectively. Practically, the emission and absorption spectra were measured thrice at three different analyte concentrations and the slope m of the plot between integrated fluorescence intensity and absorbance used to determine the fluorescence quantum yield as:

$$\Phi_F = \Phi_{F,ref} \cdot \frac{m_s}{m_{ref}} \cdot \left[\frac{n_s}{n_{ref}} \right]^2, \quad (3.15)$$

For the emission spectra of Alexa488 and Cy5 labels, the samples were excited at 460 nm and 590 nm, respectively. The reference dyes were either free Cy5 ($\Phi_F=0.30$) or Alexa488 ($\Phi_F=0.964$). The fluorescence quantum yields of free Cy5 was calibrated against Rhodamine 700 in ethanol with a $\Phi_{F,ref} = 0.38$ ²⁸⁰ and that of free Alexa488 was calibrated against Rhodamine 110 with a $\Phi_{F,ref} = 0.94$ (measured versus Rh6G in EtOH and Fluorescein in NaOH²⁷⁹). Details of corrections applied are given in Chapter 7.

(2) Fluorescence quantum yield measurements by low-intensity FCS (liFCS) method:

For the Alexa488 dyes, we additionally determined the quantum yield using the low-intensity FCS (liFCS) method²³⁷. In this approach, the molecular brightness is measured by FCS at low excitation intensities (in the range from 0.5-2 kW/cm²) and the slope m of the molecular brightness plotted against the excitation intensity is determined. The molecular brightness (MB) is defined by:

$$MB = \frac{\langle I \rangle}{N_{total}}, \quad (3.16)$$

where $\langle I \rangle$ is the average count rate and N_{total} is the particle number estimated from the FCS curve (see eq. 3.19). Analogous to eq. 3.14, the steady-state quantum yield is determined by comparison to a reference measured under identical conditions:

$$\Phi_F = \Phi_{F,\text{ref}} \cdot \frac{g_R}{g_S} \cdot \frac{\varepsilon_{\text{ref}}^{\lambda_{\text{ex}}}}{\varepsilon_S^{\lambda_{\text{ex}}}} \cdot \frac{m_S}{m_{\text{ref}}}, \quad (3.17)$$

where $\varepsilon^{\lambda_{\text{ex}}}$ is the extinction coefficient at the excitation wavelength λ_{ex} , g is the fraction of the signal that is detected (detection efficiency of the setup), and the subscripts s and ref refer to the sample and reference dyes, respectively. Free Alexa488 dye was used as a reference with $\Phi_{F,\text{ref}} = 0.964$.

Fraction of statically quenched species. From the steady state fluorescence quantum yield Φ_F , which accounts for both static and dynamic quenching, and the dynamic fluorescence quantum yield $\Phi_{F,\text{dyn}}$, which is only sensitive to dynamic quenching, we calculate the fraction of statically quenched species as:

$$f_{\text{static}} = 1 - \frac{\Phi_F}{\Phi_{F,\text{dyn}}} \quad (3.18)$$

FCS measurements of Alexa488 labels. FCS experiments of Alexa488 labels were performed on a custom-built confocal microscope. The fluorescent molecules were excited by a diode laser (Cobolt 06-MLD, Germany) at 488 nm in continuous wave mode. The laser was focused into the sample by a water-immersion objective lens (UPLAPO 60 NA = 1.2, Olympus, Hamburg, Germany). The fluorescence was collected by the same objective and separated from the excitation by a dichroic beam splitter (488/636 PC, AHF, Tübingen, Germany). In the emission path, a polarizing beam splitter (VISHTII, Gsänger) was used to split the beam into two polarizations (parallel and perpendicular to the polarization of the excitation laser light) to eliminate dead time and after-pulsing artefacts in the correlation functions. Fluorescence signals were passed through bandpass filters (HQ 533/467, Germany) and registered on avalanche photodiode detectors (APD, SPCM-AQR-14, Perkin-Elmer). Cross-correlation functions of the parallel and perpendicular signals were computed using a homebuilt hardware correlator²⁸¹. For the FCS measured at low excitation intensities, the average of parallel-perpendicular and perpendicular-parallel correlations were fitted to the 3D diffusion model^{21,282}.

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{total}}} \cdot G_{\text{diff}}(t_c), \quad (3.19)$$

where N_{total} refers to the total number of molecules. The diffusional part of the correlation function, $G_{\text{diff}}(t_c)$, is given by:

$$G_{\text{diff}}(t_c) = \frac{1}{1 + \frac{t}{t_{\text{diff}}}} \cdot \frac{1}{\sqrt{1 + \frac{t}{\left(\frac{z_0}{\omega_0}\right)^2 \cdot t_{\text{diff}}}}}, \quad (3.20)$$

where t_{diff} is the diffusion time and ω_0 and z_0 are the $1/e^2$ radii of the 3D Gaussian shaped observation volume in the lateral and in axial dimension, respectively (typically, $z_0/\omega_0 = 3-6$ for our setup). The curves were fitted over a range from 576 ns to 18 ms to exclude contributions from fast quenching dynamics on the nanosecond timescale and therefore no triplet states were detected.

The FCS measurements to investigate the static quenching of Alexa488 labels were performed at a higher power density of 15.3 kW/cm² to increase the signal-to-noise ratio of the FCS curves. The average of parallel-perpendicular and perpendicular-parallel correlation curves were fitted in the range from 4 ns to 100 ms with a 3D diffusion model and two bunching components at 3-5 μs and 45-130 ns that are assigned to triplet state formation and quenching dynamics, respectively:

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{bright}}} \cdot G_{\text{diff}}(t_c) \cdot \left(1 - A_T + A_T \cdot e^{-\frac{x}{t_T}} - A_Q + A_Q \cdot e^{-\frac{x}{t_Q}}\right), \quad (3.21)$$

where t_T and A_T are the time and fraction, respectively, of the molecules in the triplet state and t_Q and A_Q the corresponding quantities for statically quenched molecules, and the fraction of bright molecules is given by:

$$N_{\text{bright}} = N_{\text{total}} \cdot (1 - A_T - A_Q). \quad (3.22)$$

For the polarization-resolved FCS measurements of Alexa488 labels, the FCS curves were fit to the following model function:

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{bright}}} \cdot G_{\text{diff}}(t_c) \quad (3.23)$$

$$\cdot \left(1 - A_T + A_T \cdot e^{-\frac{x}{t_T}} - A_Q + A_Q \cdot e^{-\frac{x}{t_Q}} - A_{Ab} \cdot e^{-\frac{t}{t_{Ab}}} \right) \cdot G_{\text{rot}}^{\text{spher}}(t_c),$$

Photon antibunching is additionally accounted for by the photon anti bunching time t_{Ab} with amplitude A_{Ab} . The last term in eq. 3.23 is given by the rotational autocorrelation function of a spherical rotor²⁶²:

$$G_{\text{rot}}^{\text{spher}}(t_c) = 1 + A_R \left(\frac{1}{1+C} \cdot e^{-\frac{t}{t_{R1}}} + \frac{C}{1+C} \cdot e^{-\frac{t}{t_{R2}}} \right), \quad (3.24)$$

which contains two exponential terms with depolarization times t_{R1} and t_{R2} with $\frac{t_{R2}}{t_{R1}} = 0.3$, and amplitudes $A_{R1} = \frac{R}{1+C}$ and $A_{R2} = \frac{R \cdot C}{1+C}$, with $C = 0.0889$ for parallel-parallel correlation.

FCS measurements of Cy5 labels. For the Cy5 labels, polarization-resolved full FCS (pFCS) was performed on a confocal laser scanning microscope (FV1000, Olympus, Germany) equipped with a single photon counting module with picosecond time-resolution (counting electronics: HydraHarp400, PicoQuant, Berlin, Germany) at $20 \pm 1^\circ\text{C}$. Fluorescent molecules were excited by a diode laser (Olympus) operated in continuous wave mode with a wavelength of $\lambda = 635 \text{ nm}$. The detected signal passed the dichroic beam splitters (630 DCXR, Chroma, Vermont, USA) and was split into parallel and perpendicular polarized components by a polarizing beam splitter (PBS 101, Thorlabs, New Jersey, USA) and detected on MPD detectors (PDM-R by Micro Photon Devices, Bolzano, Italy) after passing a band pass filter (HQ 715/120, AHF).

To investigate the photoisomerization of Cy5 labels under different environments, pFCS measurements were performed at power densities of 13, 26, 52 and 78 kW/cm^2 . The average of parallel-perpendicular and perpendicular-parallel correlation curves were fitted to a model function $G(t_c)$ that accounts for translational diffusion in a 3D-Gaussian volume (diffusion time t_{diff}), rotational diffusion (depolarization time t_R and amplitude A_R), photon antibunching (time t_{Ab} and amplitude A_{Ab}) and two bunching terms that are assigned to the cis-trans isomerization process (relaxation times t_{ISO1} , t_{ISO2} and amplitudes A_{ISO1} , A_{ISO2}):

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{bright}}} \cdot G_{\text{diff}}(t_c) \cdot \left(1 - A_{ISO_1} + A_{ISO_1} \cdot e^{-\frac{t}{t_{ISO_1}}} - A_{ISO_2} + A_{ISO_2} \cdot e^{-\frac{t}{t_{ISO_2}}} - A_{Ab} \cdot e^{-\frac{t}{t_{Ab}}} \right) \cdot \left(1 + A_R \cdot e^{-\frac{t}{t_R}} \right). \quad (3.25)$$

Here, N_{bright} is the effective number of molecules in the bright state, given by:

$$N_{\text{bright}} = N_{\text{total}} \cdot (1 - A_{ISO_1} - A_{ISO_2}). \quad (3.26)$$

Since here we focus on the cis-trans photoisomerization, we use a simplified expression for the rotational contribution compared to eq. 3.23. Due to the large size of the pinhole used (200 μm), the detectors observe almost all molecules in the Gaussian shaped excitation volume. The fit uses two bunching terms A_{ISO_1} and A_{ISO_2} to describe the molecules in various excitation states in the cis-trans isomerization process. The sum of their amplitudes, $A_{ISO} = A_{ISO_1} + A_{ISO_2}$, yields the total fraction of molecules in the non-fluorescent cis state. Due to the distribution of excitation powers, the relaxation times are also distributed. The average cis-trans relaxation time t_{ISO} is obtained as the amplitude weighted average of the two fitted times, $t_{ISO} = A_{ISO_1} \cdot t_{ISO_1} + A_{ISO_2} \cdot t_{ISO_2}$. We do not assign any of the fitted bunching terms to the triplet state because 1) we do not observe a bunching term with a power-dependent amplitude in the microsecond range and 2) previous studies showed that a significant triplet state population of Cy5 is observed only at excitation intensities above 50 kW/cm^2 at 647 nm excitation wavelength¹⁴¹.

The excitation-dependent apparent trans-to-cis isomerization rate, k'_{ISO} , and the cis-to-trans back isomerization, k'_{BISO} , can be calculated from the amplitude, $A_{ISO} = k'_{ISO}/k'_{BISO}$, and the relaxation time, $1/t_{ISO} = k'_{ISO} + k'_{BISO}$, as¹⁴¹:

$$k'_{ISO} = \frac{A_{ISO}}{1 + A_{ISO}} \cdot \frac{1}{t_{ISO}} \quad \text{and} \quad k'_{BISO} = \frac{k'_{ISO}}{A_{ISO}}. \quad (3.27)$$

The excitation-independent trans-to-cis isomerization rate constant k_{ISO} is related to k'_{ISO} by:

$$k_{ISO} = \frac{k_{10} + k_{01}}{k_{01}} k'_{ISO}, \quad (3.28)$$

where k_{10} is the fluorescence decay rate constant estimated as $k_{10} = 1/\langle\tau\rangle_x$, and k_{01} is the excitation rate constant. The excitation rate constant is given by $k_{01} = \sigma_{01}\gamma_P I$ where σ_{01} is the absorption cross section, γ_P is the inverse photon energy and I is the excitation power.

Analysis of polarized-FCS measurements. For the measurement of rotational motions, pFCS measurements were performed at a power density of 11 kW/cm². The pFCS curves were fitted to the following model function:

$$G(t_c) = \text{offset} + \frac{1}{N_{\text{bright}}} G_{\text{diff}}(t_c) G_{\text{Ab,ISO}}(t_c) G_{\text{rot}}(t_c), \quad (3.29)$$

where $G_{\text{diff}}(t_c)$ is the translational diffusion correlation factor defined by eq. 3.20. $G_{\text{Ab,ISO}}(t_c)$ is the factor accounting antibunching and photoisomerization:

$$G_{\text{AB,ISO}}(t_c) = 1 - A_{\text{ISO}_1} + A_{\text{ISO}_1} \cdot e^{-\frac{t}{t_{\text{ISO}_1}}} - A_{\text{ISO}_2} + A_{\text{ISO}_2} \cdot e^{-\frac{t}{t_{\text{ISO}_2}}} - A_{\text{Ab}} \cdot e^{-\frac{t}{t_{\text{Ab}}}}, \quad (3.30)$$

and $G_{\text{rot}}(t_c)$ is the rotational contribution to the correlation function as given in eq. 3.24, with $C=-0.519$ for parallel-perpendicular correlation. This takes the general form:

$$G_{\text{rot}}(t_c) = 1 + f_R R_p \left(x_p g^{(2)}(t_c; f_\tau \mathbf{D}_{\text{rot}}, \boldsymbol{\mu}) + (1 - x_p) g^{(4)}(t_c; f_\tau \mathbf{D}_{\text{rot}}, \boldsymbol{\mu}) \right). \quad (3.31)$$

Here, the amplitudes R_p and x_p depend only on the choice of the excitation and emission polarizers, designated by the index p . The multiexponential functions $g^{(2)}$ and $g^{(4)}$ depend only on the diffusion tensor of the rigid rotating molecule, $\mathbf{D}_{\text{rot}} = (D_x, D_y, D_z)$, and the orientation of the transition dipole moment of the dye, $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$. To fit the experimental data using the theoretically predicted diffusion tensor, two scaling parameters were introduced to scale the common amplitude of time-dependent part of correlation (f_R), and the diffusion tensor (f_τ), which consequently scales the depolarization correlation times. These theoretical calculations were performed by Dr. Oleg Opanasyuk.

The rotational contribution as given in eq. 3.31 is derived using the solution of the Smoluchovski diffusion equation by the matrix method proposed by Kalmykov²⁸³. For the derivation, it is assumed that the excited state lifetime is significantly shorter than the rotational diffusion time (i.e., $6D_{\text{rot}}\tau \ll 1$). This allows the rotational contribution to the correlation function to be factored out and makes the correlation functions for the polarizer settings $p = (\parallel\parallel, \parallel\perp)$ and $(\parallel\perp, \parallel\parallel)$ indistinguishable. Here, $(\parallel\parallel, \parallel\perp)$ denotes that the signal detected in the parallel channel is correlated with the signal detected in the perpendicular channel, whereby both signals are obtained after excitation by parallel polarized light. A detailed derivation is presented in Supplementary Note 3.3. The theoretical values for the amplitudes R_p and x_p for common experimental polarizers setups are presented in the Supplementary Table 3.16 and 3.17.

The experimental polarization-resolved FCS curves for the two polarizer settings $p = (\parallel\parallel, \parallel\perp)$ and $(\parallel\perp, \parallel\parallel)$ were first averaged and then fitted to three different model functions for the rotational contribution in the form of eq. 3.31, assuming 1) isotropic rotation ($D_x = D_y = D_z = D$, spherical model), 2) rotation according to the expected diffusion tensor estimated from the molecular structure (scaled model), and 3) isotropic rotation described by prolate diffusion tensor ($D_x = D_y = D_{xy} < D_z$, prolate model), where the amplitude scaling factor f_R was fixed to the value obtained for the scaled model. The position of the transition moment μ was estimated based on the end-stacked conformation of Cy5 obtained from the MD simulations. The expected diffusion tensor was estimated using the HYDROPRO program²⁵⁹ using the molecular structures of duplex,3,4-way RNA junctions. Note that for the spherical model, the functions $g^{(2)}$ and $g^{(4)}$ in eq. 3.31 reduce to mono-exponential decays with rate constants $6f_\tau D_{\text{rot}}$ and $20f_\tau D_{\text{rot}}$ and hence do not depend on the orientation of transition dipole moment (see eq. 3.24).

For the fitting, the weights for data points in the diffusion range above 1 ms were suppressed by a factor of 20 to prioritize the fitting of the rotational contribution. For the prolate model, we additionally define the excess mobility around the helical z-axis compared to the HYDROPRO model, ΔD_z , by:

$$\Delta D_z = D_z(\text{fit}) - f_\tau D_z(\text{pred}), \quad (3.32)$$

where $D_z(\text{fit})$ and $D_z(\text{pred})$ are the fitted and predicted values for D_z . For the prolate model, the scaling factor f_τ is computed from the fitted and predicted average values of the x- and y-components of the diffusion tensor, D_{xy} , as:

$$f_\tau = \frac{D_{xy}(\text{fit})}{D_{xy}(\text{pred})}, \quad (3.33)$$

Computation of Förster radii. The Förster radius R_0 is the characteristic distance for a given dye pair at which the FRET efficiency is 50%:

$$R_0 = 0.2108 \sqrt[6]{\left(\frac{\Phi_{F,D}\kappa^2}{n^4}\right)J}. \quad (3.34)$$

It depends on the donor fluorescence quantum yield $\Phi_{F,D}$, the orientation factor κ^2 and the refractive index of the surrounding medium, n , and the spectral overlap between the donor emission spectrum and acceptor absorption spectrum, quantified by the spectral overlap integral J :

$$J = \int_0^\infty \bar{F}_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda, \quad (3.35)$$

where \bar{F}_D is the normalized emission spectrum donor [nm^{-1}] and $\varepsilon_A(\lambda)$ is the extinction coefficient of acceptor [$\text{M}^{-1} \text{cm}^{-1}$] at wavelength λ . The reduced Förster radius (R_{0r}) is defined analogous to eq. 3.34 by omitting the contribution of $\Phi_{F,D}$:

$$R_{0r} = 0.2108 \sqrt[6]{\left(\frac{\kappa^2}{n^4}\right)J} \quad (3.36)$$

Peulen et al. additionally introduced the reduced spectral Förster radius which combines only the contributions of n and J as:

$$R_{0J} = 0.2108 \sqrt[6]{\left(\frac{J}{n^4}\right)}, \quad (3.37)$$

which can usually be assumed to be constant over different labeling positions for a given dye pair.²⁷⁰

Computational methods

AV calculations. AV calculations were performed using the single-radius AV model which approximates the dye as a sphere with radius R_{dye} , tethered to the biomolecule by a flexible linker of length L_{link} with width w_{link} ¹⁰³. The linker length L_{link} was estimated from the fully extended conformation of the linker using the CambridgeSoft Chem3D software. The AV parameters were as follows: for Alexa488 coupled to uracil, $L_{\text{link}} = 20.0 \text{ \AA}$; for Cy5 coupled to uracil, $L_{\text{link}} = 22.0 \text{ \AA}$ and for Cy5 coupled to guanine, $L_{\text{link}} = 19.5 \text{ \AA}$. In all cases, $w_{\text{link}} = 2.5 \text{ \AA}$ and $R_{\text{dye}} = 4 \text{ \AA}$. The attachment point for the labels was the C5 atom of uracil or the N2 atom of guanine. Terminal base accessibilities was calculated as the fraction of the AV within 7 \AA of the terminal base pair for Cy5, or the terminal guanine for Alexa488. The geometric modeling was performed using the software *Olga*²⁶⁷ (<https://github.com/Fluorescence-Tools/Olga>).

The structure for the double-stranded RNA used for the AV simulations was generated using Nucleic Acid Builder from the AmberTools suite of programs²⁵⁷. The model duplexes used for predictions are 32 base pair long and consist of either all U-A or all G-C as internal base pairs, while the terminal bases are either G or C. For the extensively studied model duplex 5'C/3'G, the terminal bases at the 5'-ends are cytosines, and 3'-ends are guanines. For the model duplex 5'G/3'C (see inset in Supplementary Fig. 3.20), the terminal bases are guanines at the 5'-ends and cytosines at the 3'-ends.

Structural models of the RNA three-way and four-way junctions

Structural models for the RNA three-way^{256,264} and four-way²⁶⁴ junctions were established using the FRET-restrained structural modeling approach as described previously²⁴². Briefly, 50 donor-acceptor FRET pairs on the 4WJ and 42-45 FRET pairs on each of the four 3WJs were measured and their FRET-averaged distances and corresponding errors computed. Computational structural modeling, including MD and coarse-grained simulations, were employed to construct the full atom 3D structural models of RNA with precision of $\sim 2 \text{ \AA}$. To demonstrate that the inferred structure satisfies the experimental data, statistical analysis, employing the goodness of fit parameter and the RMSD to differentiate between the clusters of solutions, were conducted. For the 4WJ, additional coexisting minor populations were revealed²⁶⁴, however their distance uncertainties were

significantly larger than for the major conformer. In this study, the major conformer has thus been exclusively adopted as the 4WJ structure.

MD simulations. (Work performed by Dr. Christian Hanke) All simulation were performed using the ff99OL3 force field, which contains modifications to the ff99 force field²⁸⁴ by Perez et al.²⁸⁵ and Zgarbova et al.²⁸⁶. For simulations of the unlabeled RNA duplex, Mg²⁺ ions corresponding to a concentration of ~20 mM were placed within the first hydration shell of the RNA duplex following a protocol described earlier²⁸⁷ in order to ensure initial hexahydration of the Mg²⁺ ions. Subsequently, the system was placed in an octahedral box of TIP4PEW²⁸⁸ water molecules such that the minimal distance between the RNA and the edge of the system was at least 11 Å. In addition to the Mg²⁺ ions, K⁺ and Cl⁻ ions were added to achieve a concentration of ~80 mM and to neutralize the system. For Mg²⁺ ions, the parameters by Li & Merz²⁸⁹ were used, while for the monovalent ions, the parameters by Joung and Cheatham were used²⁹⁰.

For the simulations with the explicit dyes Alexa488 and Cy5 attached to the dsRNA, the respective base of the labeling position was replaced by a modified uracil base with the attached linker and dye molecule. Parameters for the modified base, linker, and dyes are from described in Spiegel et al.²⁹¹ The modified RNA was placed in a rectangular water box of TIP4PEW water molecules²⁸⁸ such that the minimal distance between RNA-dye system, and the edge of the system was at least 11 Å. The system was neutralized with Na⁺ ions using the parameters by Joung and Cheatham²⁹⁰. Positional restraints of 1 kcal mol⁻¹ Å⁻² were applied to the phosphorus atoms of the RNA throughout these simulations, as we were only interested in the movement of the dye with respect to the RNA.

Prior to the simulation, the system was thermalized adapting procedures described earlier²⁸⁷. In brief, each system was minimized by 500 steps of steepest decent minimization, followed by 2000 steps of conjugate gradient minimization. In the simulations, the particle mesh Ewald method²⁹² was used to treat long-range electrostatic interactions, and bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm²⁹³. During the next steps, harmonic force restraints with force constants of 5 kcal mol⁻¹ Å⁻² were applied to all solute atoms, and if present, to the Mg²⁺ ions and their first hydration shell water molecules. We carried out canonical ensemble

(NVT) MD simulations for 100 ps, during which the system was heated from 100 to 300 K. Subsequently, isothermal isobaric ensemble NPT simulations were run for 300 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints were gradually reduced to $1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ during 100 ps of NVT-MD simulations. This thermalization process was performed with different initial velocities for each repetition of the simulations.

In case of the unrestrained simulations of the RNA duplex only, all restraints were subsequently removed, while for the simulations with the dyes, positional restraints on the phosphorus atoms of the RNA were kept. Subsequently, NVT-MD simulations with a time step of 4 fs were performed using the GPU version of *pmemd*²⁹⁴⁻²⁹⁶ from the Amber18 suite of programs²⁵⁷.

For the unrestrained RNA duplex, we performed five independent simulations of 300 ns length each. We performed simulations for Alexa488 attached to position b5, b11, and b33, and for Cy5 attached to position b27 and b33. For each of these positions, five independent simulations of 8000 ns length each were performed. Conformations were extracted every 20 ps.

Cluster analysis of MD simulations (Performed by Dr. Oleg Opanasyuk)

The dye position and transition dipole moments (TDM) were defined based on the atom coordinates as shown in Supplementary Fig. 3.22. For Alexa488, the TDM is defined along the long axis of the central xanthylium unit using the middle points of the outermost pairs of C atoms. For Cy5, the two terminal C atoms of the polymethine chain are used. The central position of the TDM was taken as the middle point of these vectors.

Agglomerative hierarchical clustering using centroid linkage was performed based on the central coordinates of the dye using the *fastcluster* library (<https://github.com/cran/fastcluster>). The hierarchy of clusters was established using the *fcluster* method of the SciPy package²⁹⁷ using a maximum number of clusters of 70. Clusters were sorted according to their population weight. Any cluster with more than 1% contribution was classified as an interacting state, while all other clusters were assigned to diffusive states of the dye. Interacting states in close spatial proximity were manually combined. For the Cy5 label at position b27, the cluster assignment was performed

manually to improve the resolution of closely spaced conformations as described in Supplementary Fig. 3.34.

For each identified cluster, the mean of the position of the TDMs was calculated as the average of the coordinates, $\langle \mathbf{r} \rangle = (\langle x \rangle, \langle y \rangle, \langle z \rangle)$. The variance of the TDM position was calculated using the eigen-value decomposition. First, the direct product of the coordinates was calculated for each position:

$$\mathbf{r} \otimes \mathbf{r} = \begin{pmatrix} xx & yx & zx \\ xy & yy & zy \\ xz & yz & zz \end{pmatrix}. \quad (3.38)$$

Then, the variance-covariance matrix (Σ) was calculated as the average of the direct product of the coordinates:

$$\Sigma = \langle \mathbf{r} \otimes \mathbf{r} \rangle = \left\langle \begin{pmatrix} xx & yx & zx \\ xy & yy & zy \\ xz & yz & zz \end{pmatrix} \right\rangle, \quad (3.39)$$

The eigen-values λ_i of Σ are the principal variances of the positional distribution, $\sigma_i^2 = \lambda_i$ for $i = \{x, y, z\}$. The square roots of the variances, i.e., the standard deviations σ_i , define the width of the positional distribution along the eigen-vectors \mathbf{v}_i which define the principal axes of the variance ellipsoid.

For the calculation of average TDM orientations, the TDM vector \mathbf{m} was first normalized:

$$\boldsymbol{\mu} = \frac{\mathbf{m}}{|\mathbf{m}|}, \quad (3.40)$$

where $|\mathbf{m}|$ is the norm of the TDM vector. For the averaging of vector orientations, it is important to account for the bidirectionality of the TDM. This is done by including the inverse of each TDM to set of TDM $\boldsymbol{\mu}_i$ before averaging:

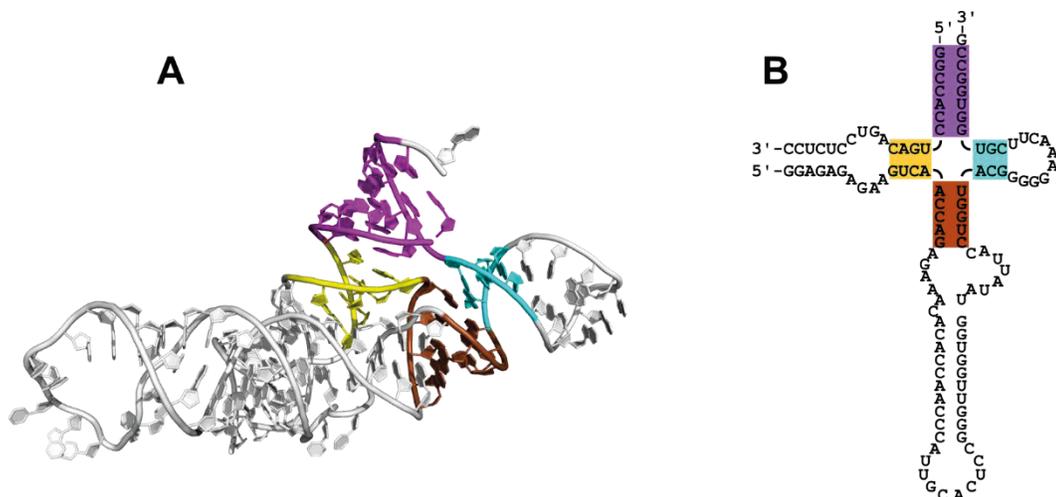
$$\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_1, \dots\} \rightarrow \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_1, \dots, -\boldsymbol{\mu}_1, -\boldsymbol{\mu}_2, -\boldsymbol{\mu}_1, \dots\}. \quad (3.41)$$

Then, the averaging of the normalized TDM is performed as described for the position vector \mathbf{r} above. As for the position vector, the eigen-values define the directional variances along the eigen-vectors. The normalized eigen vectors for TDM orientation represent the rotation matrix R_σ that transforms a TDM vector from the coordinate frame of the PDB atomic coordinates to the eigen-frame of the orientational Σ matrix.

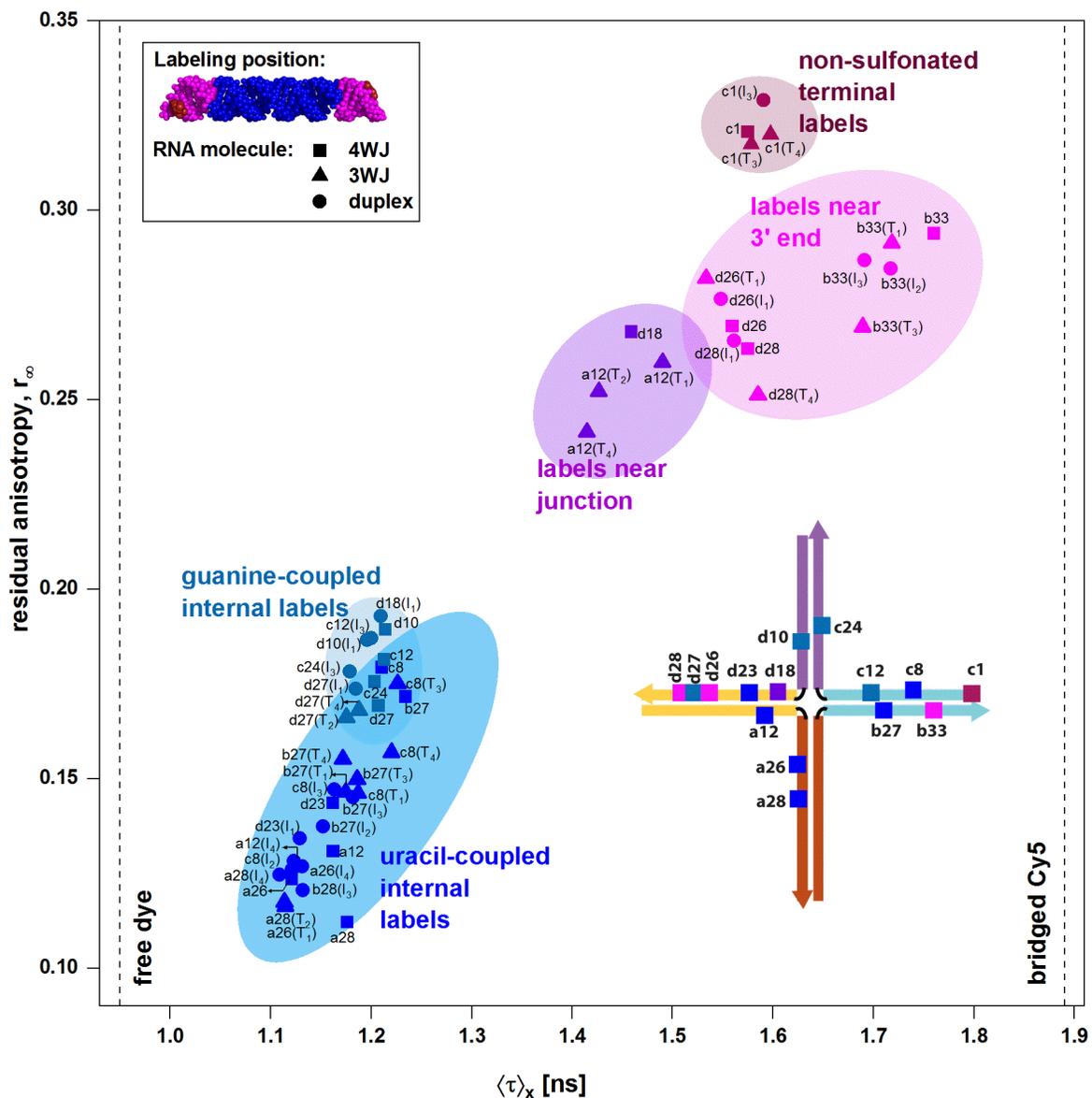
$$R_{\sigma} = \begin{pmatrix} \hat{v}_{xx} & \hat{v}_{yx} & \hat{v}_{zx} \\ \hat{v}_{xy} & \hat{v}_{yy} & \hat{v}_{zy} \\ \hat{v}_{xz} & \hat{v}_{yz} & \hat{v}_{zz} \end{pmatrix}. \quad (3.42)$$

3.7 Supporting information

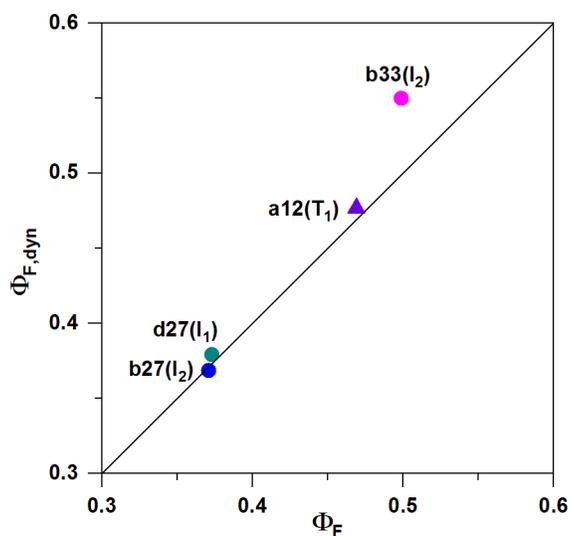
Supplementary Figures



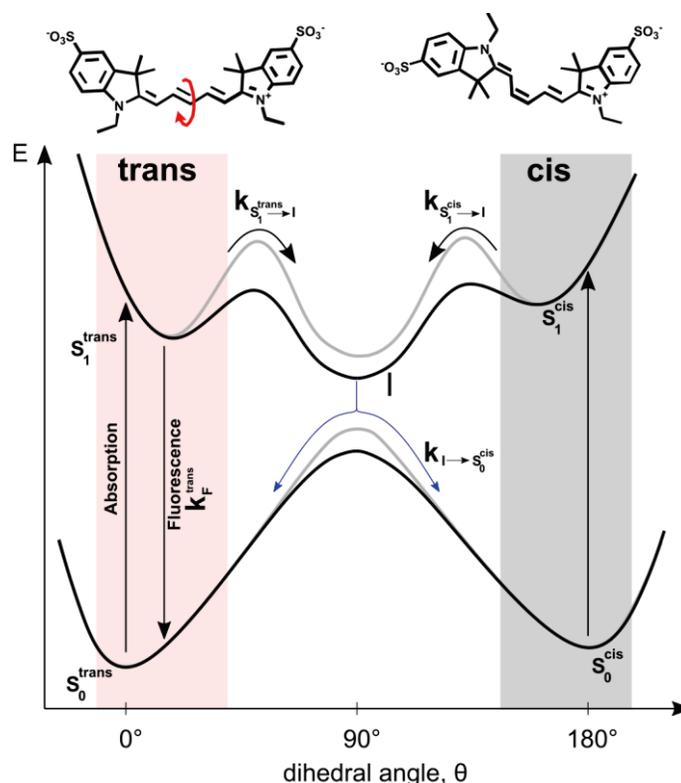
Supplementary Figure 3.1: The crystal structure (A) and sequence (B) of hairpin ribozyme (PDB: 1M5K)²³². The colored bases near the junction have been preserved in the RNA four-way junction molecule used in our study (see Fig. 3.1A in the main text).



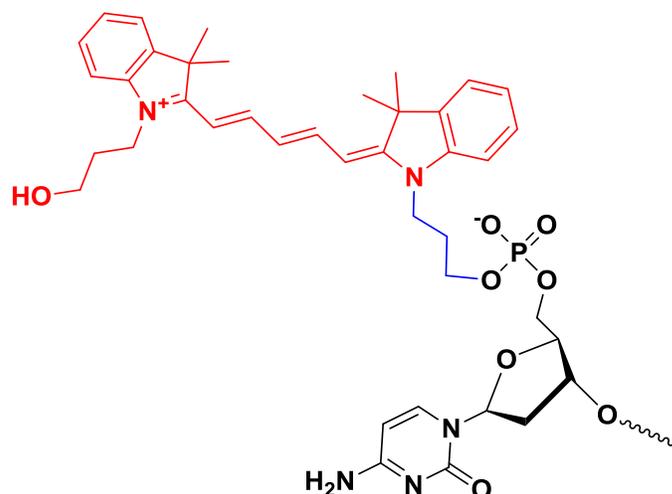
Supplementary Figure 3.3: Detailed plot of the species-weighted average fluorescence lifetime $\langle \tau \rangle_x$ and residual anisotropies r_∞ of Cy5 labels, including all label names. Positions of the studied labels within the RNA 4WJ parent structure are shown in the inset.



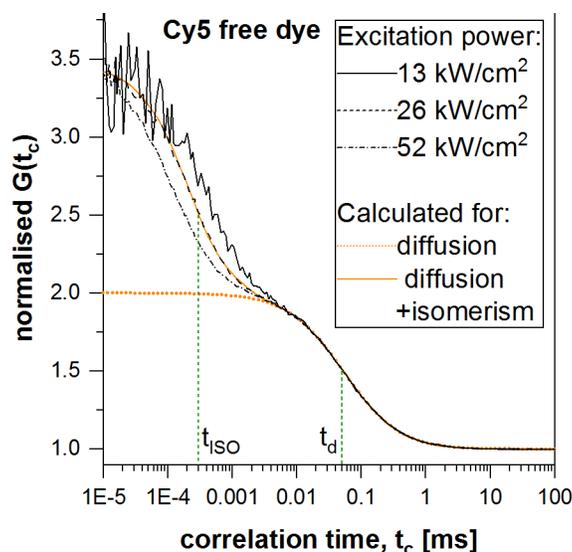
Supplementary Figure 3.4: Comparison of the dynamic ($\Phi_{F,dyn}$) and steady-state (Φ_F) quantum yields of representative Cy5 labels for the four clusters of Cy5 labels identified in Fig. 3.2A of the main text. No large differences are observed between the two estimates of the fluorescence quantum yield, indicating that there is no static quenching occurring for the Cy5 labels.



Supplementary Figure 3.5: Schematic potential energy diagram of Cy5 photoisomerization as a function of the dihedral angle, θ , of the central conjugated double bond. The structures of the fluorescent trans and non-fluorescent cis isomers of Cy5 are shown above. The energy levels in the absence of steric hindrance are shown as black lines, while the modulation of the energy levels with increased steric hindrance is shown as in gray. The ground state is in all-trans conformation and denoted as S_0^{trans} . Upon absorption of light, the molecule is excited to the S_1^{trans} state. Relaxation from this state can occur either (1) radiatively through fluorescence with a rate constant, k_F^{trans} or (2) via a twisted intermediate (I) formed by rotation ($\theta = 90^\circ$) around the C-C bond (red arrow) with rate constant $k_{S_1^{\text{trans}} \rightarrow I}$. Relaxation from the twisted intermediate occurs non-radiatively either back to the trans ground state, S_0^{trans} , or to the cis photoisomer S_0^{cis} . Since the trans configuration is energetically favorable, the population of the cis photoisomer will slowly convert back into the trans state thermally. Back-conversion of the cis isomer S_0^{cis} can also occur after re-excitation into the S_1^{cis} and relaxation over the twisted intermediate via $S_0^{\text{cis}} \rightarrow S_1^{\text{cis}} \rightarrow I \rightarrow S_1^{\text{trans}}$. Steric hindrance increases the energy barrier to the twisted intermediate state, thereby blocking the non-radiative relaxation from the S_1^{trans} state by photoisomerization and enhancing the fluorescence. This is the cause of association-induced fluorescence enhancement (AIFE).

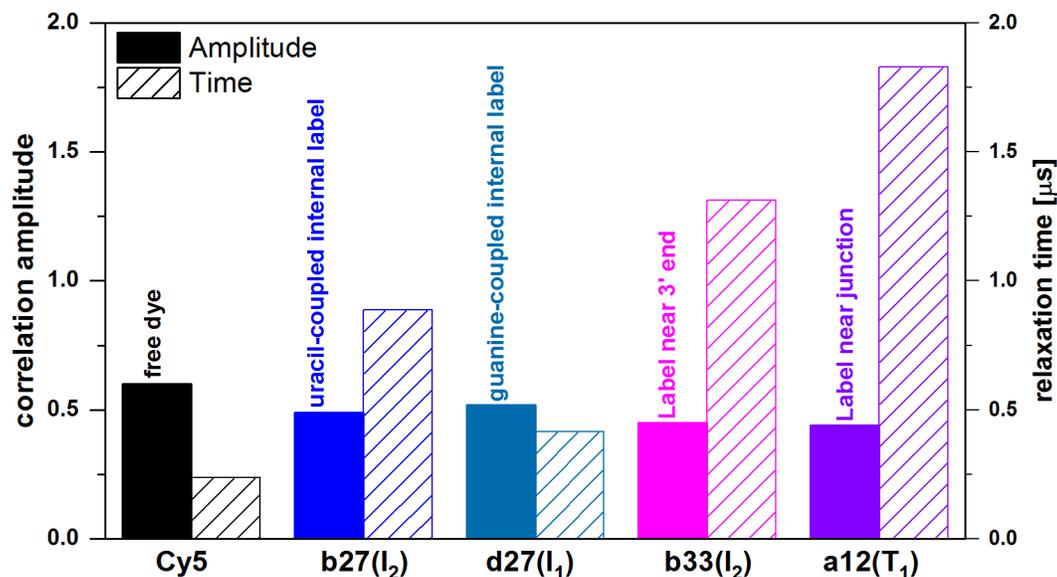


Supplementary Figure 3.6: Chemical structure of the terminally attached Cy5 fluorophore used as a control for end stacking. The Cy5 label is non-sulfonated and attached to the 5' terminal phosphate of the RNA duplex with a three-carbon linker. This design is similar to the labeling approach used in ^{91,92} where stable end stacking was shown by 2D NMR and X-ray crystallography studies.

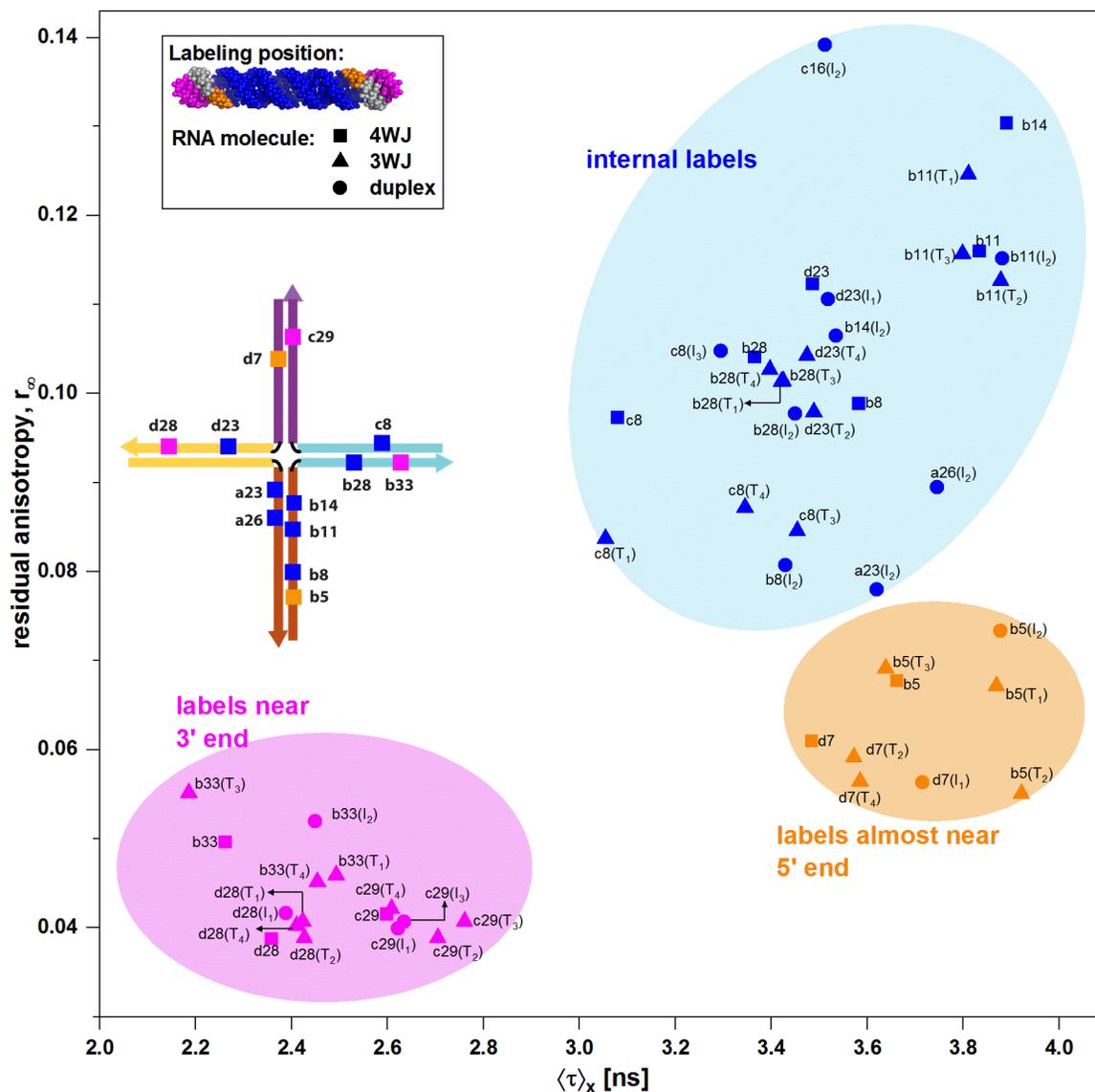


Supplementary Figure 3.7: Fluorescence correlation curves of free Cy5 dye measured at different power densities of 13, 26 and 52 kW/cm². The FCS curves were measured using polarized excitation and detection. Shown are the average curves of the parallel-perpendicular and perpendicular-parallel correlation functions, normalized to the total number of molecules N_{total} . The model correlation functions for translational diffusion only (dotted orange line) and translational diffusion with cis-trans photoisomerization (solid orange line) are calculated based

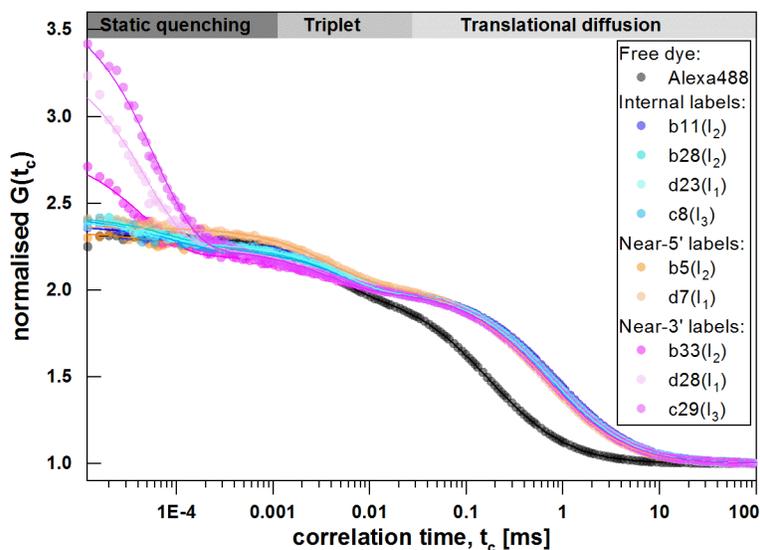
on the parameters estimated for the FCS curve obtained at a power density of 26 kW/cm² using the model given by eq. 3.25. The vertical green lines indicate the fitted relaxation times for translational diffusion (t_{diff}) and cis-trans photoisomerization (t_{iso}) for the FCS curve at 26 kW/cm². The corresponding FCS curves for the Cy5 labels on RNA at different excitation are shown in Fig. 3.2D of the main text. The fitted parameters are given in Supplementary Table 3.3.



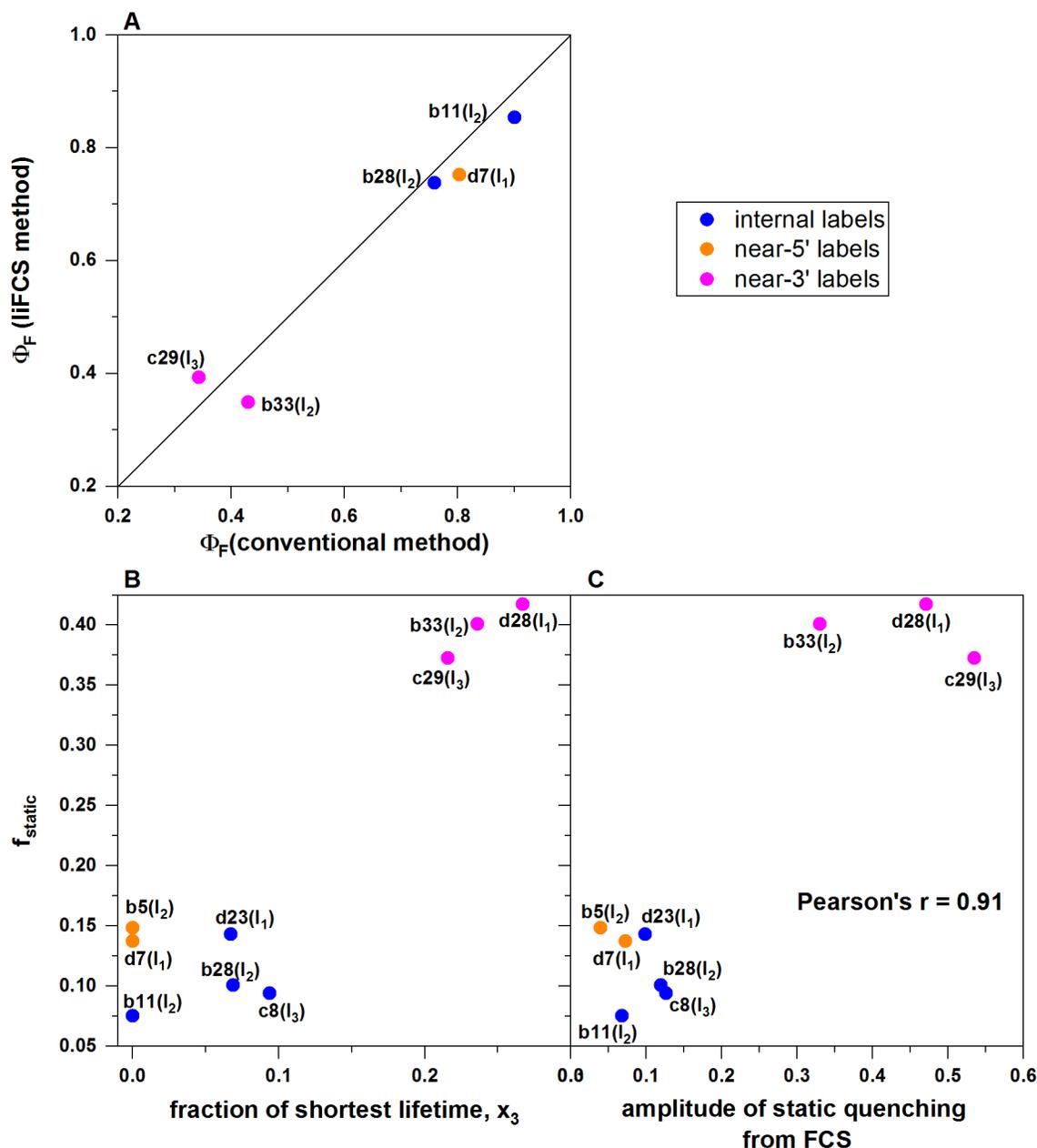
Supplementary Figure 3.8: Comparison of the cis-trans isomerization amplitudes and relaxation times for representative Cy5 labels. The given amplitude is the sum of the amplitudes of the two terms assigned to the cis-trans isomerization process, A_{ISO1} and A_{ISO2} , and the isomerization relaxation time is taken as the amplitude weighted average relaxation time of the two components. The relaxation times are more sensitive to the local environment than the corresponding amplitudes. The isomerization is found to be slower for the end label b33(I₂) and the near-junction label a12(T₁) compared to the free dye and the internal labels.



Supplementary Figure 3.9: Detailed plot of the species-weighted average fluorescence lifetime $\langle \tau \rangle_x$ and residual anisotropies r_∞ of Alexa488 labels, including all label names. Positions of the studied labels within the RNA 4WJ parent structure are shown in the inset.

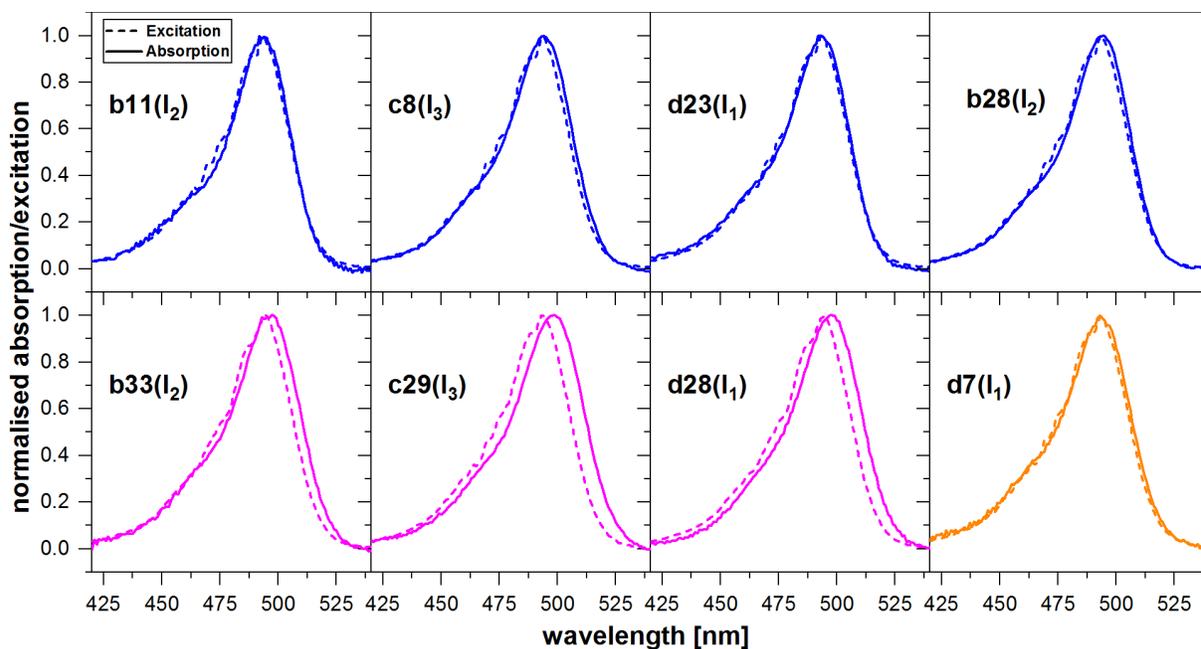


Supplementary Figure 3.10: FCS curves of Alexa488 labels measured at a power density of 15 kW/cm^2 with respective fits to the model given by eq. 3.21 (solid lines). The FCS curves were measured using polarized excitation and detection. Shown are average curves of the parallel-perpendicular and perpendicular-parallel correlation functions, normalized to the total number of molecules N_{total} . This figure includes additional curves to those in Fig. 3.3D of the main text. All near 3'-end labels (magenta) show a significant bunching term in the 40-60 ns range with a fraction of 30-50%, which is assigned to the static quenching process (see Supplementary Fig. 3.11C). This bunching term is reduced for the almost near 5'-labels (orange) and internal labels (blue) to a fraction of 7-13%. The FCS curve of the free Alexa488 dye is given for comparison. The fitted parameters are given in Supplementary Table 3.5.

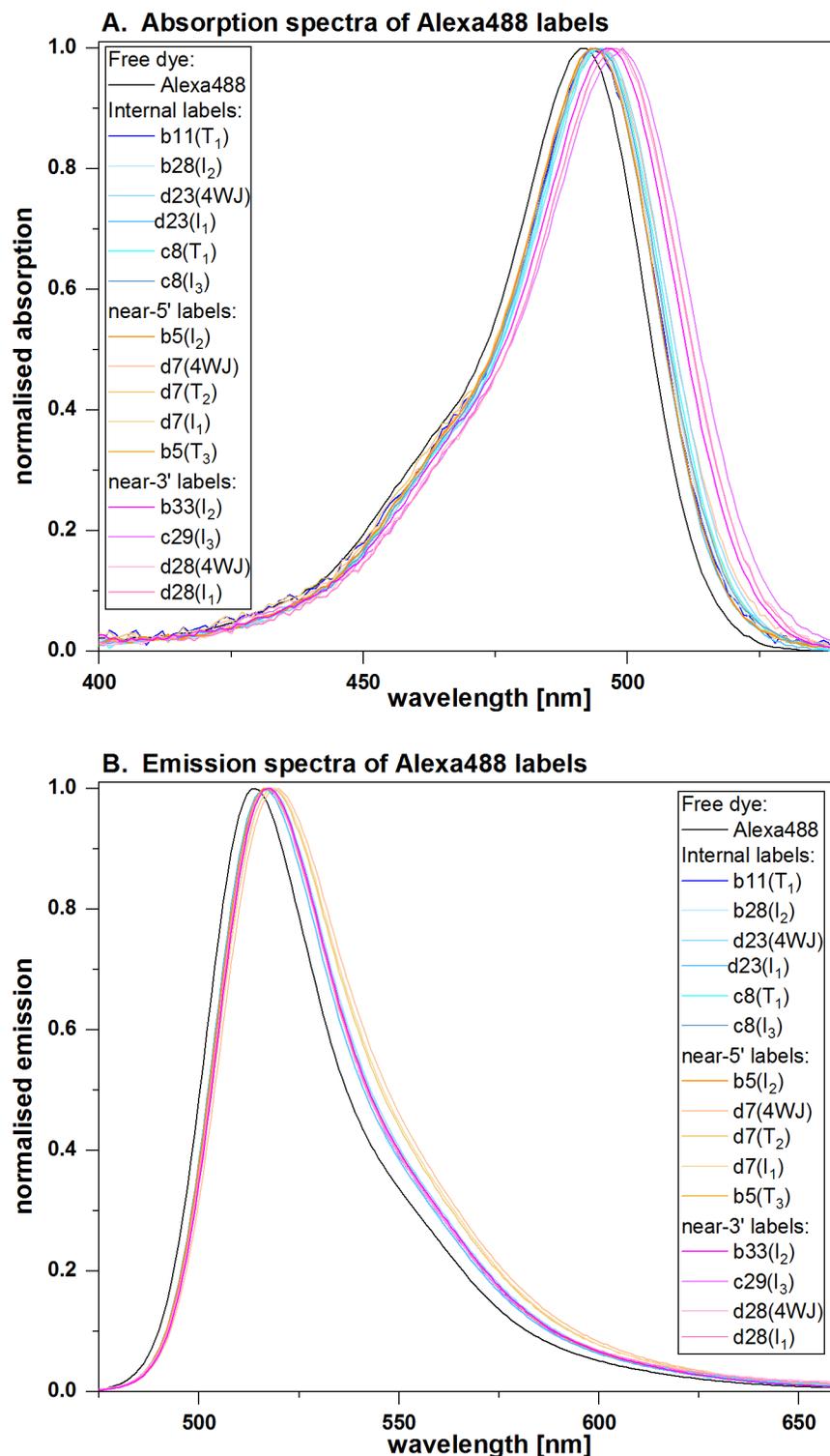


Supplementary Figure 3.11: A) A comparison of the steady-state quantum yield, Φ_F , measured by the low-intensity FCS (IIFCS) method and from the spectral emission shows that the IIFCS provides quantum yield estimates that are consistent with the conventional method (see Supplementary Table 3.6). **B)** A plot of the fraction of statically quenched species, f_{static} , calculated from the dynamic and steady-state fluorescence quantum yields using eq. 3.18, against the fraction of the shortest lifetime component, x_3 , in the ensemble lifetime measurements. The fraction of statically quenched species correlates well with the fraction of the shortest lifetime component (<300 ps) obtained from the ensemble fluorescent

lifetime measurements (Supplementary Table 3.4), indicating that the short lifetime component originates from static quenching. **C)** A plot of the fraction of statically quenched species f_{static} against the amplitude of the bunching term of the FCS curves in the range of 50 to 100 ns. The good correlation between the two quantities supports that the observed bunching originates from static quenching.

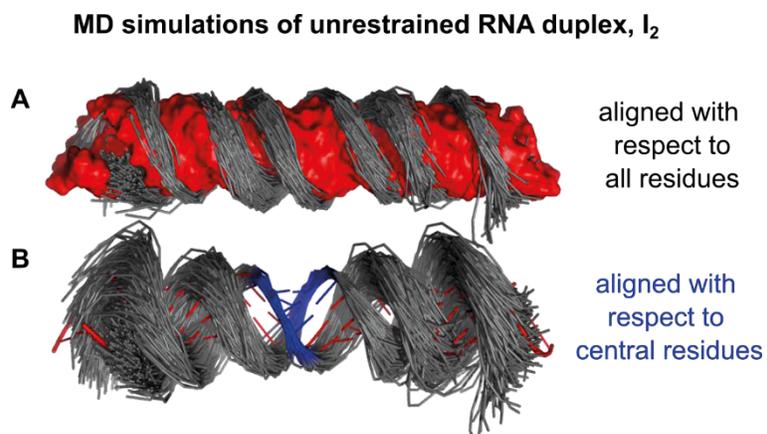


Supplementary Figure 3.12: Comparison of absorption and excitation spectra of Alexa488 labels (blue: internal labels, magenta: near 3'-end labels, orange: near 5'-end labels). The excitation spectra were recorded by scanning the excitation wavelength and monitoring the emission at 550 nm. For the labels near the 3'-end, a shift between the absorption and excitation spectra is detected that is absent for the internal, near 5'-end labels, and free dye control measurements. This is in support of the formation of stable π - π ground state complexes between the chromophore and terminal base pair.

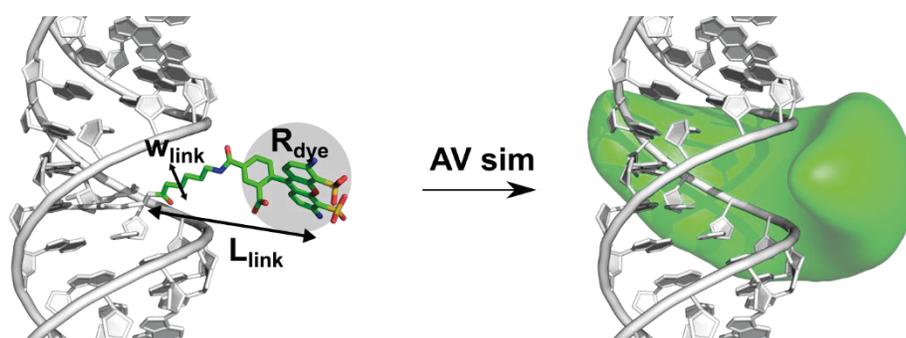


Supplementary Figure 3.13: A plot of the absorption and emission spectra of the free dye and select Alexa488 labels from each cluster. All labels on the RNA consistently show a shift of the absorption maximum of ~ 2 nm compared to the free dye, while the near 3'-end labels show a significantly larger shift of ~ 5 nm, indicating the formation of a stable ground state

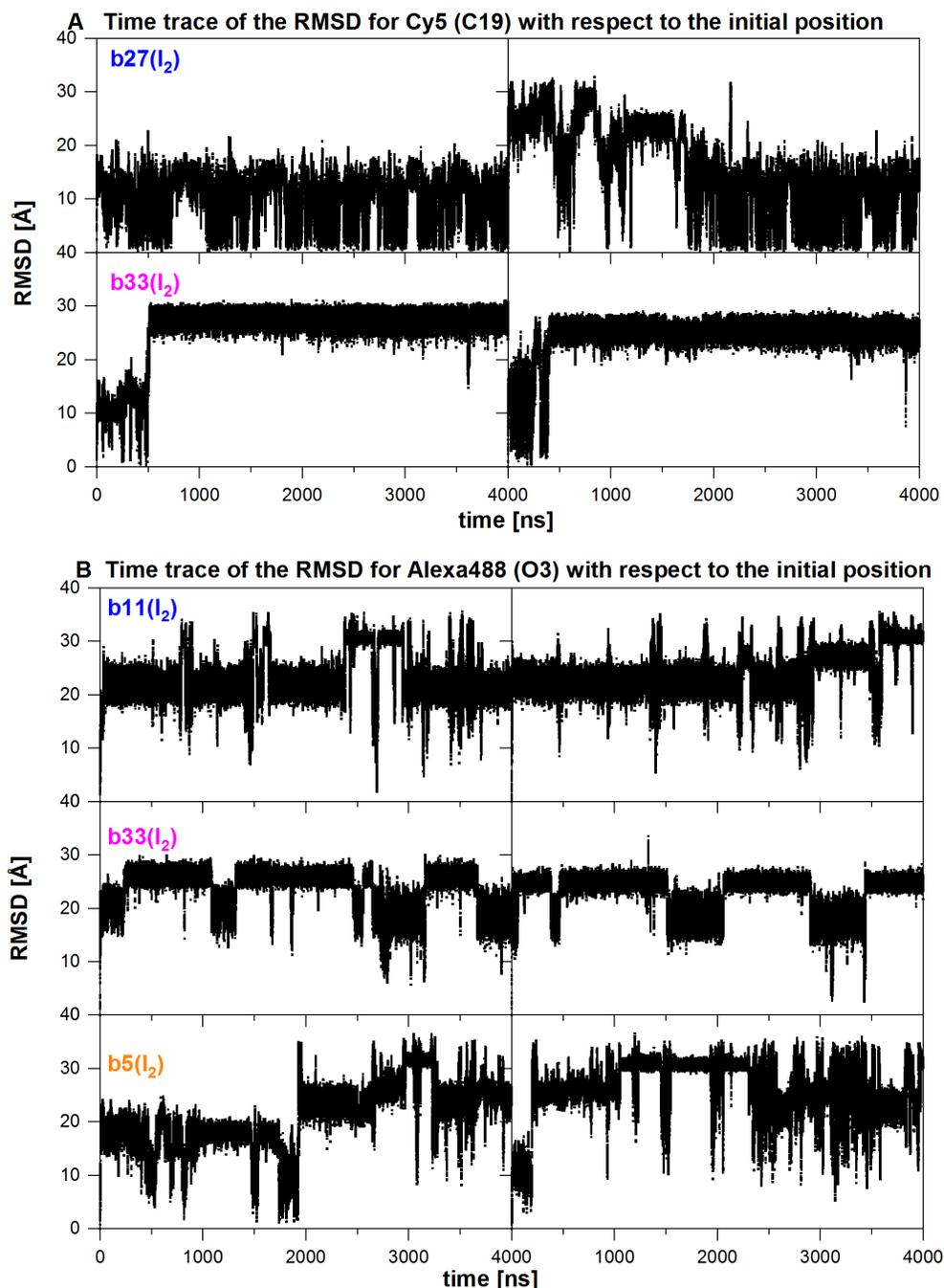
complex. The emission spectra of the labels on RNA are consistently shifted by 2 nm compared to the free dye, with the exception of the near 5'-end labels b5 and d7.



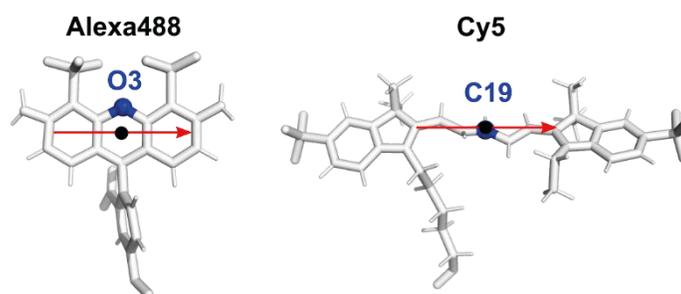
Supplementary Figure 3.14: Structural flexibility of the RNA duplex characterized by MD simulations. The starting structure (red) is overlaid with every 100th frame of the MD simulations (gray). **A)** Conformations are aligned to all atoms. The starting structure is shown in the surface representation. **B)** Extracted conformations aligned to the four central base pairs are highlighted in blue. The MD simulations show the flexibility of the RNA duplex which undergoes bending motions. The fluctuations are largest at the helix termini.



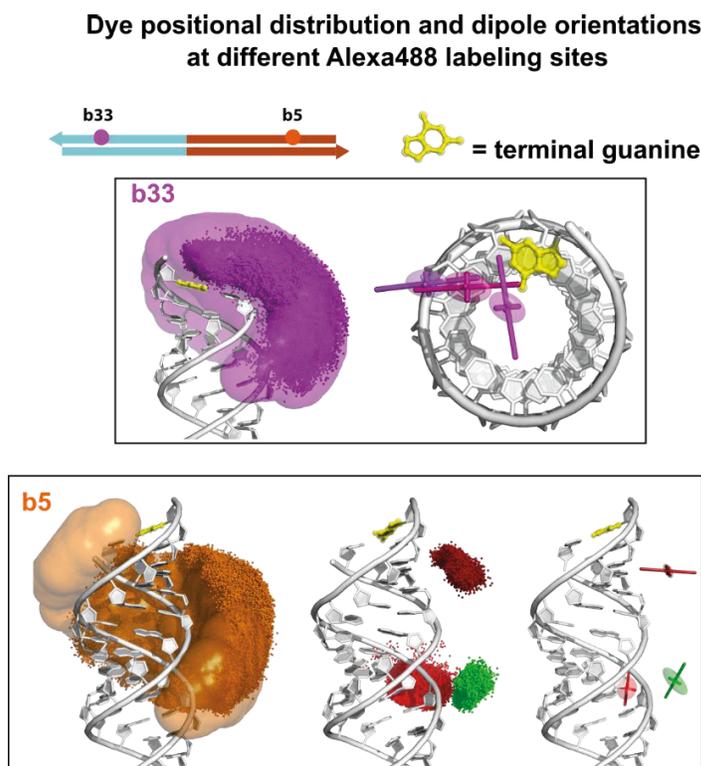
Supplementary Figure 3.15: Illustration of the parameters used in AV simulations. The dye is approximated by a sphere with radius R_{dye} and is attached to the RNA base by a flexible linker of length L_{link} and width w_{link} . The resulting spatial distribution of sterically allowed dye positions is shown in green.



Supplementary Figure 3.21: Exemplary time traces of the root-mean-square deviation (RMSD) with respect to the initial position of the dyes during the MD simulations. **A)** Cy5 label attached to uracil at the internal position b27 and the end position b33. **B)** Alexa488 label attached at internal position b11, the near 3'-end position b33 and almost near 5'-end at b5. The RMSD was calculated for the O3 atom in the case of Alexa488 and the C19 atom in the case of Cy5.

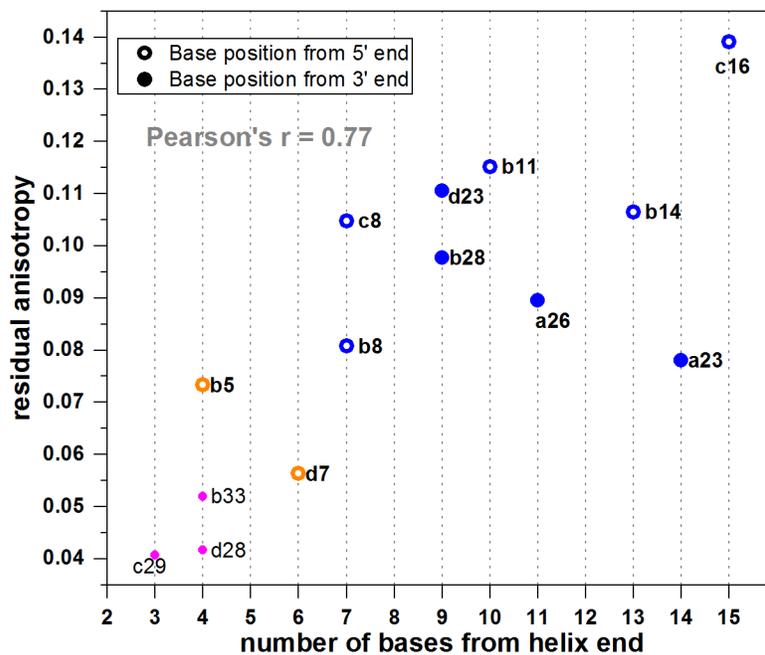


Supplementary Figure 3.22: Definition of transition dipole moment vectors used in the analysis of the MD simulations. Transition dipole moments (TDM) are indicated by red arrows. The center of the TDM, shown as black dots, is used to define the dye position in cluster analysis. The atoms, O3 and C19 (blue), serve to represent the location of Alexa488 and Cy5, respectively, in the depiction of complete (non-clustered) dye position distributions and for the calculation of time traces of the RMSD during the MD simulations, due to their positions being approximately at the center of the dye.

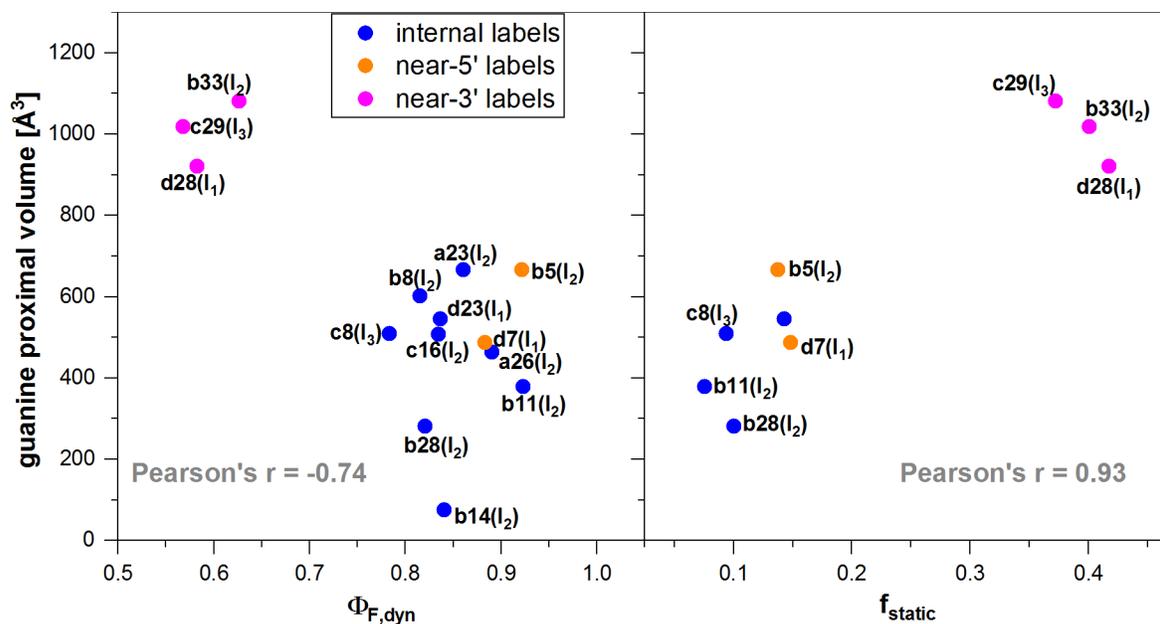


Supplementary Figure 3.23: Dye position distributions and dipole orientations of dye clusters obtained from MD for the Alexa488 labels at positions b33(I_2) and b5(I_2). Magenta clusters indicate π -stacking, green clusters have predominantly backbone interactions and

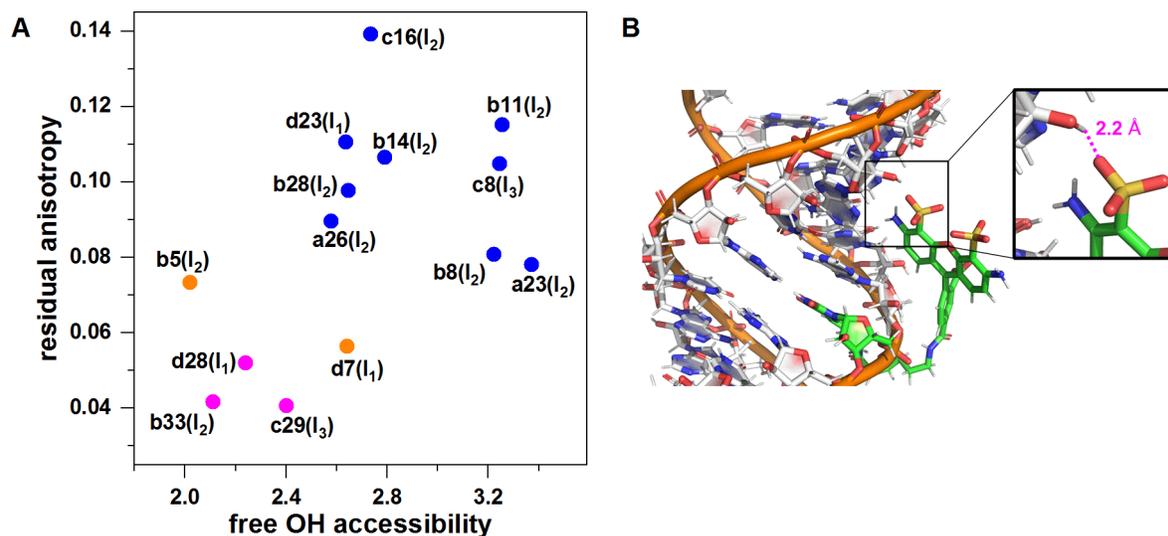
red clusters interact primarily in the groove region. The terminal guanine is colored yellow. The b33 label shows prominent end-stacking dyes in defined orientations. The near 5' label, b5, shows highly defined clusters originating from backbone or groove interactions.



Supplementary Figure 3.24: Residual anisotropies of Alexa488 labels on duplexes as a function of base count from the 5' (empty circles) and 3' (filled circles) helix ends. Internal labels are shown in blue and labels near 5'-end are shown in orange. The labels near the 3'-end undergo considerable static quenching and are therefore biased towards lower residual anisotropy because the interacting dyes contribute no photons to the fluorescence signal.

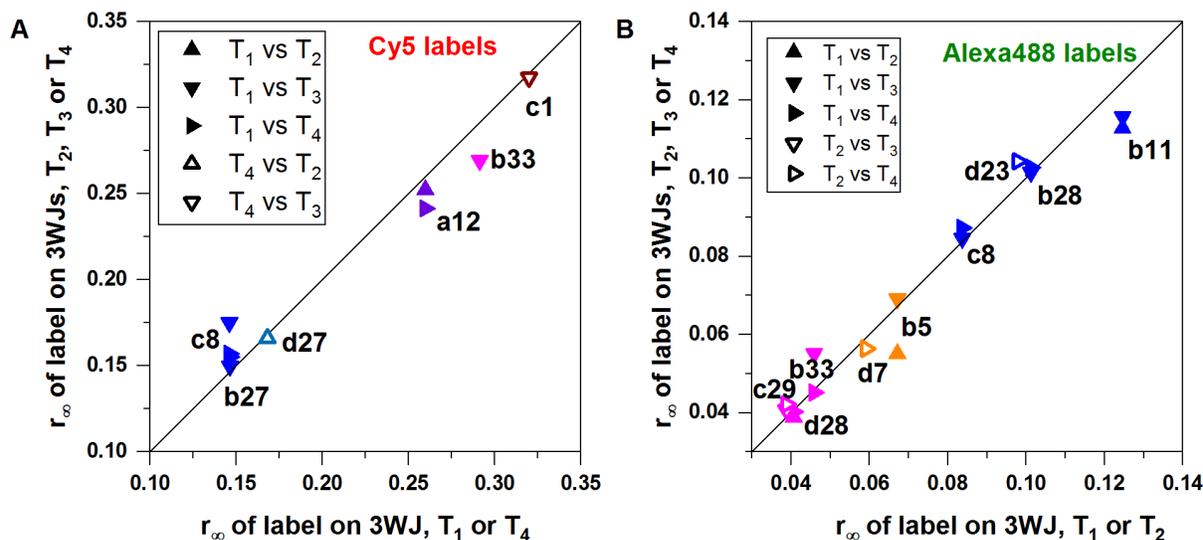


Supplementary Figure 3.25: A plot of the guanine proximal volume, estimated as the volume within the AV that is within 7 Å of any guanine, against the dynamic quantum yield, $\Phi_{F,dyn}$ (left), and the fraction of the statically quenched species, f_{static} (right), reveals a correlation (Pearson's correlation coefficient, r , given) for both quantities. A higher guanine proximal volume corresponds to a lower dynamic quantum yield and higher fraction of statically quenched dye.

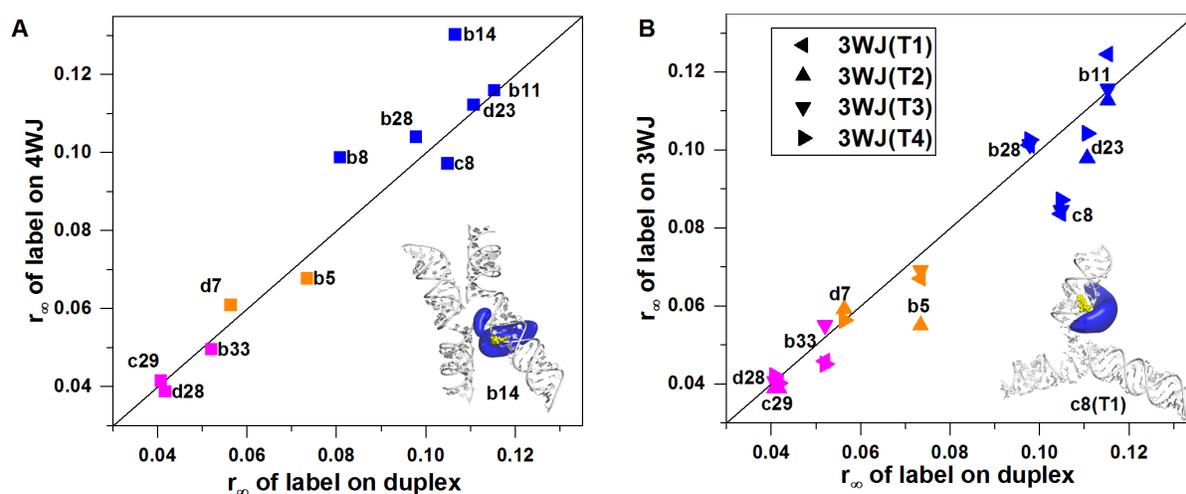


Supplementary Figure 3.26: Hydrogen bonds facilitate dye-RNA surface interactions. **A)** A plot of the free hydroxyl group surface accessibilities (defined here as the ratio of the volume of the AV cloud within 7Å of the free hydroxyl group to the duplex surface accessible within

7 Å) of Alexa488 labels on duplexes against their measured residual anisotropies. **B)** A snapshot from the MD simulations of Alexa488 at position b11, showing the formation of a hydrogen bond between a sulfo group of the dye and a hydroxyl group of a sugar residue.

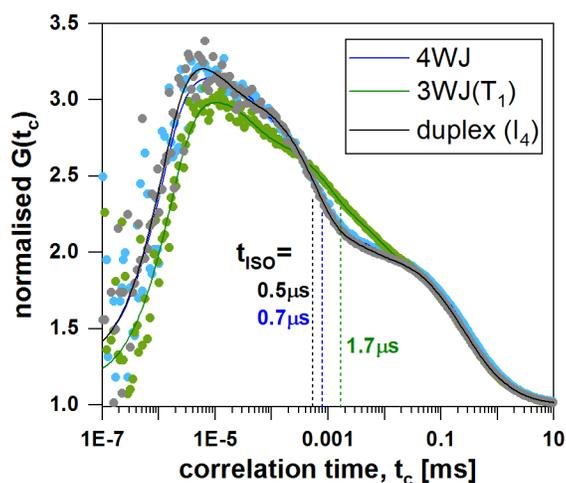


Supplementary Figure 3.27: Comparison of the residual anisotropies, r_∞ , obtained for Cy5 (A) and Alexa488 (B) labels on the different three-way junction. The 3WJ T₁ is compared against T₂, T₃ or T₄. The diagonal line indicates a 1:1 correlation. No significant differences are observed between the different three-way junction molecules. The observed variation of ± 0.01 is on the order of the experimental uncertainty.

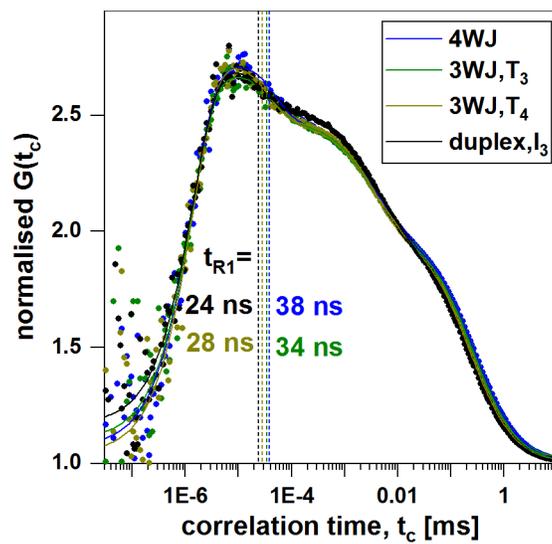


Supplementary Figure 3.28: A) Comparison of the residual anisotropies, r_∞ , of Alexa488 labels on the four-way junction and the corresponding duplex. The diagonal line indicates a

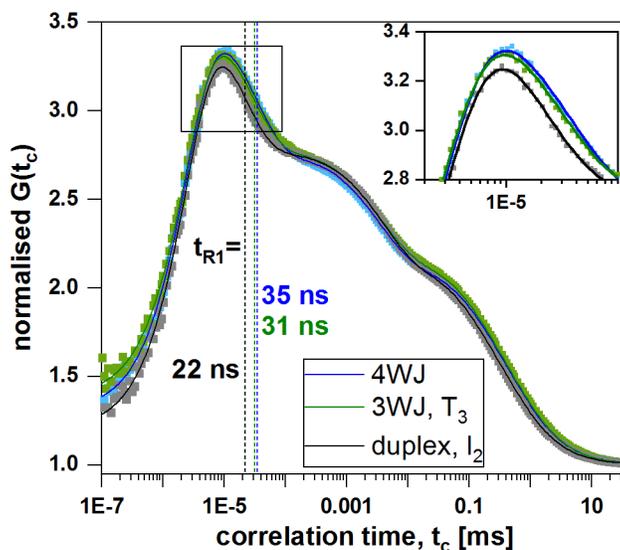
1:1 correlation. Label b14 shows the largest deviation, potentially due to proximity to the junction as illustrated by the AV of the b14 label on the 4WJ (inset). The structure of the RNA four-way junction was obtained using 50 FRET pairs as described in ²⁶⁴. **B)** Comparison of the residual anisotropies, r_{∞} , of Alexa488 labels on different three-way junction molecules and the corresponding duplex. The diagonal line indicates a 1:1 correlation. The inset shows the AV cloud of label c8 on the 3WJ T₁. The structure of the RNA three-way junction was obtained using 42-45 FRET pairs as described in ²⁶⁴ and ²⁵⁶. The labeling sites are highlighted in yellow in the insets.



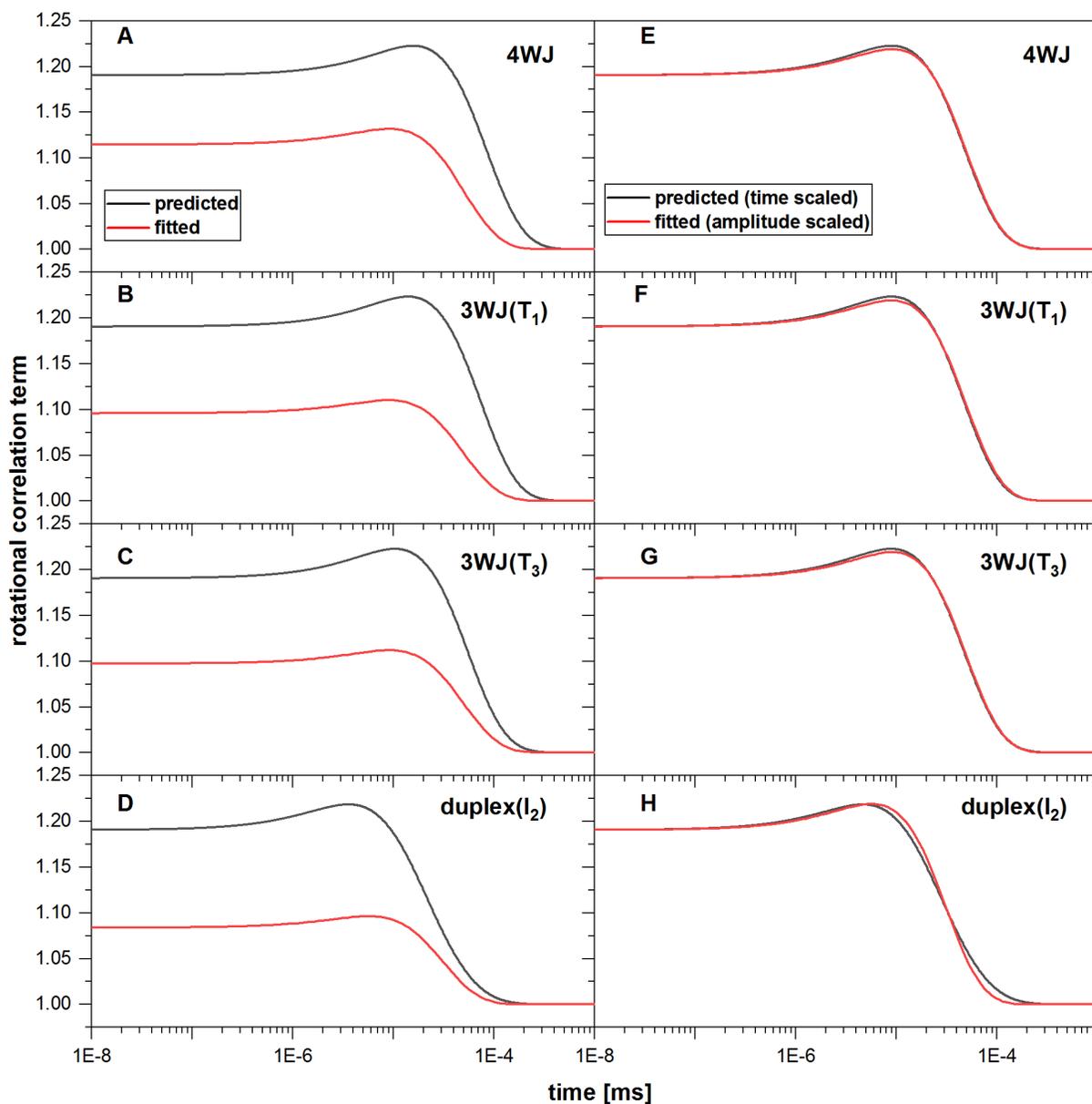
Supplementary Figure 3.29: FCS curves of Cy5 labels at position a12 in the four-way junction molecule (4WJ), the three-way junction T₁ and the corresponding duplex, I₄. The FCS curves were measured at a power density of 11 kW/cm² using polarized excitation and detection. Shown are parallel-perpendicular correlation functions, normalized to the total number of molecules N . The corresponding fits to the model given by eq. 3.25 are given as solid lines. The fitted parameters are listed in Supplementary Table 3.9. A clear difference is observed for the label close on the three-way junction T₁, while the FCS curves for the four-way junction and duplex overlap. The longer relaxation time observed for the cis-trans photoisomerization on the three-way junction indicates that the dye experiences steric restriction due to interaction with the nearby junction.



Supplementary Figure 3.30: Polarization-resolved FCS curves of the non-sulfonated Cy5 labels at the terminal position c1 on the four-way junction, the three-way junctions T_3 and T_4 and the duplex I_3 . The FCS curves were measured at a power density of 11 kW/cm^2 using polarized excitation and detection. Shown are the parallel-perpendicular correlation functions. FCS curves were fitted to the model given in eq. 3.29-30 with the generalized model function for a spherical rotor as given in eq. 3.29 and normalized to the total number of molecules N_{total} . Vertical lines indicate the fitted longest rotational diffusion times (t_{R1}). The rotational depolarization of this label monitors predominantly the global tumbling motion of the RNA because the non-sulfonated terminal Cy5 prominently stacks to the terminal bases (compare Supplementary Fig. 3.6). The fitted parameters are given in Supplementary Table 3.11.

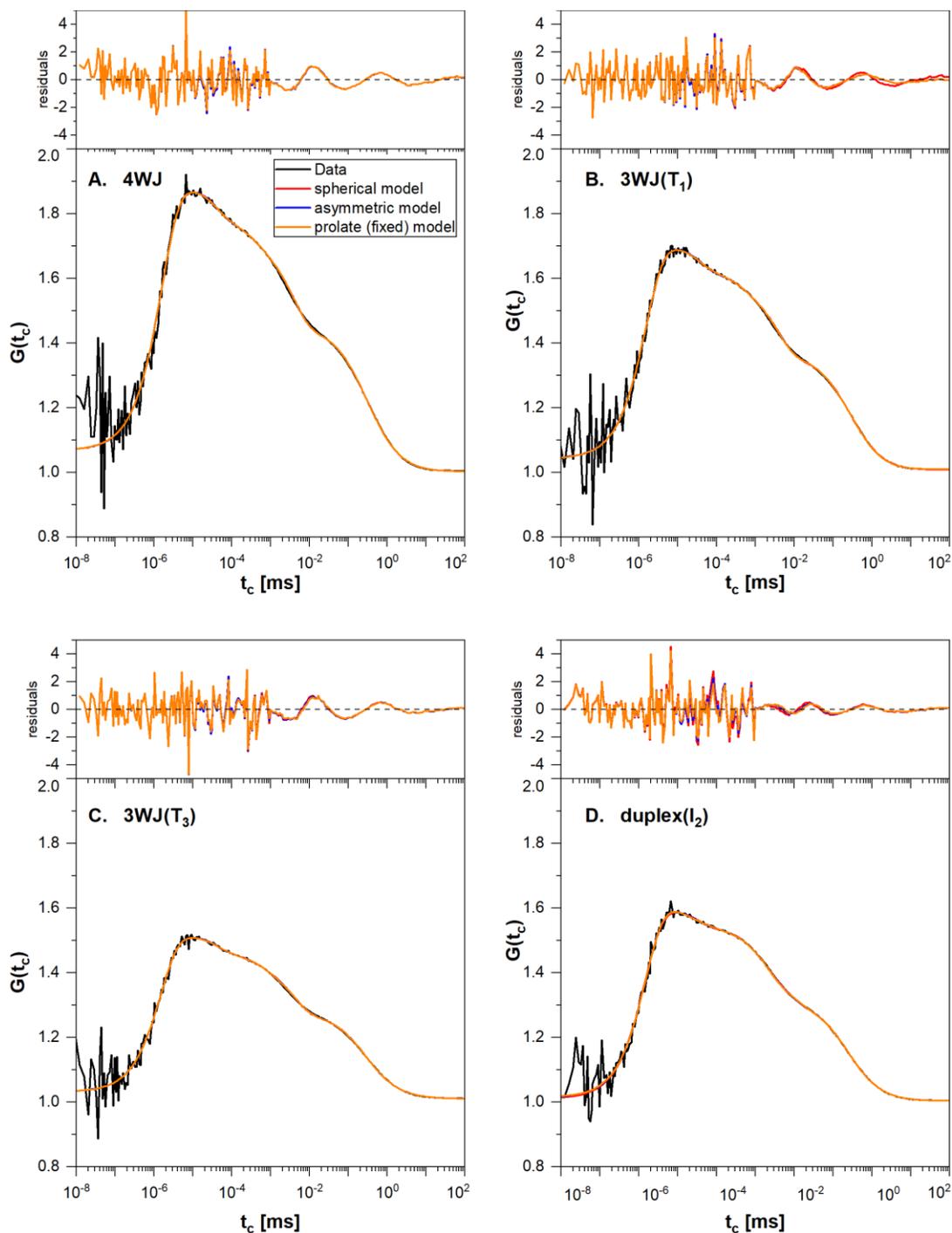


Supplementary Figure 3.31: Polarization-resolved FCS curves of Alexa488 labels at the internal position b11 on the four-way junction, the three-way junction T_3 and the duplex I_2 . The FCS curves were measured at a power density of 17 kW/cm^2 using polarized excitation and detection. Shown are the average of 2 parallel-parallel correlation functions. FCS curves were fitted to the generalized model function for a spherical rotor as given by eq. 3.23 and normalized to the total number of molecules N_{total} . Vertical lines indicate the fitted longest rotational diffusion times (t_{R1}). The inset magnifies the region of the curve that is relevant for the rotational contribution. The rotational times and their corresponding amplitudes are given in Supplementary Table 3.12.



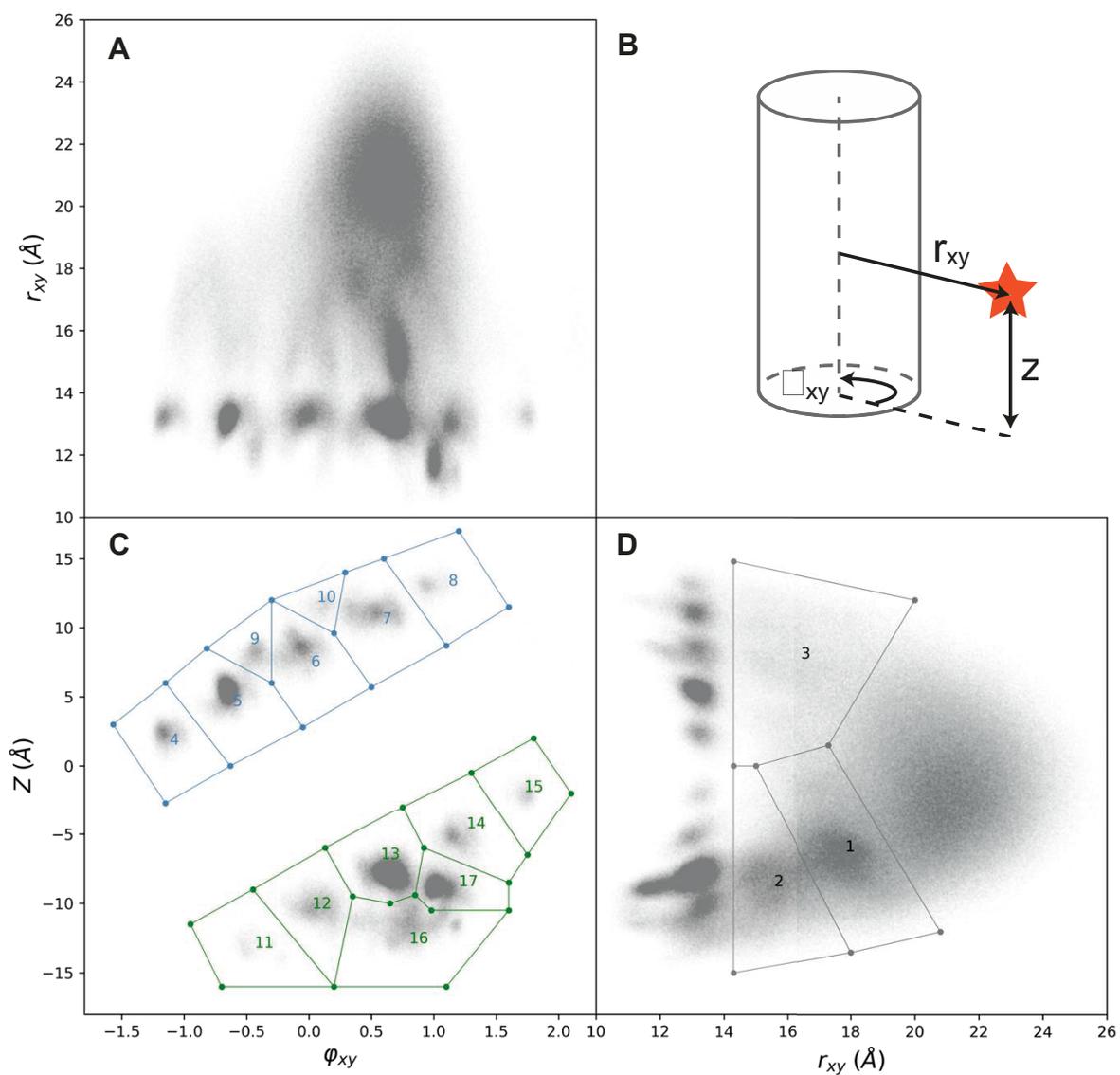
Supplementary Figure 3.32: Comparison of predicted and fitted rotational correlation terms for end-stacked Cy5 labels for the four-way junction (A), the three-way junctions T_1 (B) and T_3 (C), and the duplex I_2 (D). The parallel-perpendicular pFCS curves fitted to the model function for a spherical rotor as given in eq. 3.29. Shown is only the rotational correlation term, defined as $1 + A_R \left(\frac{1}{1+C} \cdot e^{-\frac{t}{t_{R1}}} + \frac{C}{1+C} \cdot e^{-\frac{t}{t_{R2}}} \right)$, where the two rotational terms are characterized by depolarization times t_{R1}, t_{R2} and amplitudes A_{R1}, A_{R2} . For the spherical rotor, $\frac{t_{R2}}{t_{R1}} = 0.3$, $A_{R1} = \frac{A_R}{1+C}$ and $A_{R2} = \frac{A_R \cdot C}{1+C}$, and $C = -0.519$ for parallel-perpendicular correlation functions²⁶². The

theoretical rotational correlation functions were calculated using the diffusion tensor obtained from the HydroPro software²⁵⁹ based on the FRET restrained structures of the three and four-way junctions and the average dipole moment of the fluorophore obtained from the MD simulations, as described in the methods. The amplitude of the rotational contribution to the correlation is underestimated by a factor of ~2 in the experiment compared to the prediction. The fitted parameters are given in Supplementary Table 3.14. The fitted curves were scaled with respect to their amplitude and time dependence to match the predicted curves (E-F). A good agreement is obtained for the three and four-way junction molecules, while the predicted duplex clearly deviates from the spherical rotor model. The scaling factors are given in Supplementary Table 3.13.



Supplementary Figure 3.33: Inferring the molecular shape of RNA from polarized-FCS curves. The experimental correlation curves (black curves) of the end-stacked Cy5 label b33 labeled on the different RNA molecules are compared with fits that assume a spherical model (red curve, eq. 3.29), the asymmetric model based on the predicted diffusion tensor (blue, eq. 3.31) and the prolate model with a fixed amplitude of the rotational contribution (orange, see

Methods). The quality of the fits show that the spherical model is a good approximation for the 4WJ and 3WJs. For the duplex, the fit quality is notably improved using the asymmetric model due to its rod shape. See Supplementary Table 3.14 for a summary of the fitted parameters.



Supplementary Figure 3.34: Manual clustering of the MD simulation for the Cy5 label at position b27. The spatial coordinates of the dye center position as shown in Supplementary Fig. 3.22 are defined in cylindrical coordinates as shown in **B**, with r_{xy} being the distance of the dye from the helical axis, ϕ_{xy} the angle within the lateral xy plane, and z the position along the helical axis. **A**) A plot of the distance from the helical axis, r_{xy} , against the angular position in the xy plane, ϕ_{xy} . **C**) A plot of the axial position, z , against the angular position in the xy plane,

φ_{xy} . **D)** A plot of the axial position, z , against the distance from the helical axis, r_{xy} . The manual clustering was performed as follows. First, diffusive states were identified in the z - r_{xy} projection shown in D based on a minimum distance from the helical axis of $r_{xy} > 14.3$ Å. After the diffusive states were removed, a total of 14 clusters were manually identified in the z - φ_{xy} parameter space as shown in C. The polygon edges were drawn through the regions of minimum density between the clusters. For each cluster, the mean dye position and transition dipole moment orientation was then calculated as described in the methods. Work performed by Dr. Oleg Opanasyuk.

Supplementary Tables

Supplementary Table 3.1. Fitted lifetime and anisotropy values for Cy5 labels.

| label name | molecule | x_1 | τ_1 (ns) | x_2 | τ_2 (ns) | $\langle\tau\rangle_x$ (ns) | r_1 | ρ_1 (ns) | r_2 (r_∞) | ρ_{global} (ns) |
|----------------------|----------|-------|---------------|-------|---------------|-----------------------------|-------|---------------|----------------------|-----------------------------|
| free dye | free dye | | | 1.00 | 0.95 | 0.95 | | | | |
| b27 | 4WJ | 0.25 | 1.82 | 0.75 | 1.04 | 1.23 | 0.21 | 0.64 | 0.17 | > 50 |
| b27(T ₃) | 3WJ | 0.25 | 1.75 | 0.75 | 1.00 | 1.19 | 0.23 | 0.55 | 0.15 | > 50 |
| b27(T ₁) | 3WJ | 0.20 | 1.78 | 0.80 | 1.03 | 1.17 | 0.23 | 0.64 | 0.15 | > 50 |
| b27(T ₄) | 3WJ | 0.22 | 1.70 | 0.78 | 1.02 | 1.17 | 0.22 | 0.75 | 0.16 | > 50 |
| b27(l ₂) | duplex | 0.22 | 1.67 | 0.78 | 1.01 | 1.15 | 0.24 | 0.74 | 0.14 | > 50 |
| b27(l ₃) | duplex | 0.19 | 1.89 | 0.81 | 1.02 | 1.18 | 0.23 | 0.63 | 0.15 | > 50 |
| d23 | 4WJ | 0.25 | 1.61 | 0.75 | 1.01 | 1.16 | 0.24 | 0.70 | 0.14 | > 50 |
| d23(l ₁) | duplex | 0.31 | 1.51 | 0.69 | 0.96 | 1.13 | 0.25 | 0.68 | 0.13 | 32.4 |
| c8 | 4WJ | 0.26 | 1.74 | 0.75 | 1.03 | 1.21 | 0.20 | 0.65 | 0.18 | > 50 |
| c8(T ₃) | 3WJ | 0.23 | 1.81 | 0.77 | 1.05 | 1.23 | 0.20 | 0.63 | 0.18 | > 50 |
| c8(T ₄) | 3WJ | 0.21 | 1.85 | 0.79 | 1.05 | 1.22 | 0.22 | 0.61 | 0.16 | > 50 |
| c8(T ₁) | 3WJ | 0.23 | 1.72 | 0.77 | 1.03 | 1.19 | 0.23 | 0.62 | 0.15 | > 50 |
| c8(l ₃) | duplex | 0.20 | 1.73 | 0.80 | 1.03 | 1.16 | 0.23 | 0.74 | 0.15 | > 50 |
| c8(l ₂) | duplex | 0.34 | 1.47 | 0.66 | 0.94 | 1.12 | 0.25 | 0.67 | 0.13 | 34.11 |
| a26 | 4WJ | 0.24 | 1.53 | 0.76 | 0.99 | 1.12 | 0.26 | 0.71 | 0.12 | > 50 |
| a26(T ₁) | 3WJ | 0.30 | 1.48 | 0.70 | 0.96 | 1.11 | 0.26 | 0.66 | 0.12 | > 50 |
| a26(l ₄) | duplex | 0.21 | 1.60 | 0.79 | 1.00 | 1.13 | 0.25 | 0.73 | 0.13 | > 50 |
| a28 | 4WJ | 0.30 | 1.58 | 0.70 | 1.00 | 1.18 | 0.27 | 0.69 | 0.11 | 20.90 |
| a28(T ₂) | 3WJ | 0.24 | 1.52 | 0.76 | 0.99 | 1.11 | 0.26 | 0.67 | 0.12 | > 50 |
| a28(l ₄) | duplex | 0.30 | 1.46 | 0.70 | 0.96 | 1.11 | 0.26 | 0.83 | 0.12 | > 50 |
| b28(l ₃) | duplex | 0.16 | 1.78 | 0.84 | 1.01 | 1.13 | 0.26 | 0.63 | 0.12 | > 50 |
| a12 | 4WJ | 0.22 | 1.71 | 0.78 | 1.00 | 1.16 | 0.25 | 0.55 | 0.13 | > 50 |
| a12(T ₁) | 3WJ | 0.40 | 2.12 | 0.60 | 1.06 | 1.49 | 0.12 | 0.51 | 0.26 | > 50 |
| a12(T ₂) | 3WJ | 0.34 | 2.11 | 0.66 | 1.07 | 1.43 | 0.13 | 0.52 | 0.25 | > 50 |
| a12(T ₄) | 3WJ | 0.35 | 2.07 | 0.65 | 1.06 | 1.41 | 0.14 | 0.53 | 0.24 | > 50 |
| a12(l ₄) | duplex | 0.24 | 1.55 | 0.76 | 0.99 | 1.12 | 0.25 | 0.73 | 0.13 | > 50 |
| d18 | 4WJ | 0.40 | 2.01 | 0.60 | 1.09 | 1.46 | 0.11 | 0.62 | 0.27 | > 50 |
| d18(l ₁) | duplex | 0.39 | 1.59 | 0.61 | 0.97 | 1.21 | 0.19 | 0.67 | 0.19 | 24.80 |
| d27 | 4WJ | 0.35 | 1.61 | 0.65 | 0.99 | 1.21 | 0.21 | 0.67 | 0.17 | > 50 |

| | | | | | | | | | | |
|----------------------|--------|------|------|------|------|------|------|------|------|-------|
| d27(T ₂) | 3WJ | 0.37 | 1.55 | 0.63 | 0.96 | 1.18 | 0.21 | 0.72 | 0.17 | 40.09 |
| d27(T ₄) | 3WJ | 0.35 | 1.58 | 0.65 | 0.97 | 1.19 | 0.21 | 0.66 | 0.17 | 39.76 |
| d27(I ₁) | duplex | 0.37 | 1.55 | 0.63 | 0.97 | 1.18 | 0.21 | 0.64 | 0.17 | 18.10 |
| c12 | 4WJ | 0.36 | 1.61 | 0.64 | 0.99 | 1.21 | 0.20 | 0.57 | 0.18 | > 50 |
| c12(I ₃) | duplex | 0.37 | 1.59 | 0.63 | 0.97 | 1.20 | 0.19 | 0.62 | 0.19 | 19.07 |
| c24 | 4WJ | 0.34 | 1.59 | 0.66 | 1.00 | 1.20 | 0.20 | 0.56 | 0.18 | 21.98 |
| c24(I ₃) | duplex | 0.37 | 1.54 | 0.63 | 0.96 | 1.18 | 0.20 | 0.66 | 0.18 | 18.33 |
| d10 | 4WJ | 0.36 | 1.60 | 0.64 | 0.99 | 1.21 | 0.19 | 0.74 | 0.19 | > 50 |
| d10(I ₁) | duplex | 0.37 | 1.57 | 0.63 | 0.98 | 1.20 | 0.19 | 0.64 | 0.19 | 18.52 |
| b33 | 4WJ | 0.85 | 1.91 | 0.15 | 0.91 | 1.76 | 0.09 | 0.54 | 0.29 | 32.94 |
| b33(T ₁) | 3WJ | 0.80 | 1.90 | 0.20 | 0.97 | 1.72 | 0.09 | 0.71 | 0.29 | 31.13 |
| b33(T ₃) | 3WJ | 0.72 | 1.95 | 0.28 | 1.02 | 1.69 | 0.11 | 0.60 | 0.27 | 26.68 |
| b33(I ₂) | duplex | 0.80 | 1.91 | 0.20 | 0.94 | 1.72 | 0.09 | 0.52 | 0.28 | 18.55 |
| b33(I ₃) | duplex | 0.79 | 1.89 | 0.21 | 0.93 | 1.69 | 0.09 | 0.44 | 0.29 | 15.36 |
| d26 | 4WJ | 0.78 | 1.74 | 0.22 | 0.92 | 1.56 | 0.11 | 0.56 | 0.27 | 26.58 |
| d26(T ₁) | 3WJ | 0.79 | 1.71 | 0.21 | 0.87 | 1.53 | 0.11 | 0.58 | 0.28 | 18.90 |
| d26(I ₁) | duplex | 0.75 | 1.74 | 0.25 | 0.96 | 1.55 | 0.10 | 0.53 | 0.28 | 15.03 |
| d28 | 4WJ | 0.65 | 1.90 | 0.35 | 0.99 | 1.57 | 0.12 | 0.53 | 0.26 | > 50 |
| d28(T ₄) | 3WJ | 0.67 | 1.89 | 0.33 | 0.96 | 1.59 | 0.13 | 0.42 | 0.25 | 29.63 |
| d28(I ₁) | duplex | 0.64 | 1.88 | 0.36 | 0.98 | 1.56 | 0.11 | 0.51 | 0.27 | 20.54 |
| c1 | 4WJ | 0.87 | 1.69 | 0.13 | 0.85 | 1.57 | 0.06 | 0.47 | 0.32 | 28.2 |
| c1(T ₃) | 3WJ | 0.86 | 1.69 | 0.14 | 0.87 | 1.58 | 0.06 | 0.41 | 0.32 | 20.66 |
| c1(T ₄) | 3WJ | 0.89 | 1.69 | 0.11 | 0.88 | 1.60 | 0.06 | 0.49 | 0.32 | 25.34 |
| c1(I ₃) | duplex | 0.90 | 1.68 | 0.10 | 0.81 | 1.59 | 0.05 | 0.40 | 0.33 | 13.58 |

The fitted lifetime components are τ_{1-2} and the corresponding fractions are x_{1-2} . The anisotropy decay is described by two depolarization times ρ_{1-2} with corresponding fractions, r_1 and r_∞ . The longest anisotropy component arises from the global motion and its amplitude is the residual anisotropy (r_∞). The colors given to the label names indicate the dye environment. Bright blue: internal labels attached to uracil, purple: labels near the junction, blue-gray: internal labels attached to guanine, magenta: labels near the 3'-end, non-sulfonated terminal labels: dark red.

Supplementary Table 3.2. Comparison of dynamic ($\Phi_{F,dyn}$) and steady state (Φ_F) quantum yields of Cy5 labels.

| Label name | Molecule | $\Phi_{F,dyn}$ | Φ_F |
|----------------------|----------|----------------|----------|
| b27(I ₂) | duplex | 0.37 | 0.37 |
| d27(I ₁) | duplex | 0.38 | 0.37 |
| b33(I ₂) | duplex | 0.55 | 0.50 |
| a12(T ₁) | duplex | 0.48 | 0.47 |

The dynamic quantum yield was determined from the species-averaged fluorescence lifetime given in Supplementary Table 3.1 using eq. 3.13. The steady-state quantum yields were determined by measuring the spectral emission with respect to a reference as described in the methods.

Supplementary Table 3.3. The fit results for the FCS curves of Cy5 labels.

| Label type | Label name | Power (kW/cm ²) | offset | N _{bright} | t _{diff} (ms) | z ₀ /w ₀ | A _{ISO1} | t _{ISO,1} (μs) | A _{ISO2} | t _{ISO,2} (μs) | A _{Ab} | t _{Ab} (ns) | A _R | t _R (ns) |
|---------------------|----------------------|-----------------------------|--------|---------------------|------------------------|--------------------------------|-------------------|-------------------------|-------------------|-------------------------|-----------------|----------------------|----------------|---------------------|
| Free dye | Cy5 | 13 | 1 | 0.93 | 0.05 | 6 | 0.25 | 1.26 | 0.35 | 0.23 | 0.95 | 0.27 | | |
| | | 26 | 1 | 0.85 | 0.05 | 6 | 0.16 | 0.98 | 0.44 | 0.19 | 0.96 | 0.67 | | |
| | | 52 | 1 | 0.74 | 0.05 | 6 | 0.16 | 0.69 | 0.44 | 0.1 | 0.95 | 0.6 | | |
| Internal U-attached | b27(l ₂) | 13 | 1 | 1.08 | 0.23 | 6 | 0.1 | 8.26 | 0.4 | 0.72 | 1.01 | 1.02 | 0.18 | 21.1 |
| | | 26 | 1 | 0.93 | 0.23 | 6 | 0.11 | 6.24 | 0.38 | 0.51 | 0.89 | 1.04 | 0.18 | 56.9 |
| | | 52 | 1 | 0.7 | 0.23 | 6 | 0.13 | 5.24 | 0.34 | 0.4 | 0.89 | 0.98 | 0.25 | 52.3 |
| | | 76 | 1 | 0.6 | 0.23 | 6 | 0.15 | 4.57 | 0.32 | 0.32 | 0.89 | 0.93 | 0.28 | 39.9 |
| Internal G-attached | d27(l ₁) | 13 | 1 | 1.31 | 0.21 | 6 | 0.09 | 3.97 | 0.44 | 0.7 | 0.84 | 1.36 | 0.21 | 18 |
| | | 26 | 1 | 1.14 | 0.21 | 6 | 0.11 | 2.42 | 0.42 | 0.38 | 0.92 | 1.06 | 0.2 | 17.8 |
| | | 52 | 1 | 0.85 | 0.21 | 6 | 0.1 | 2.41 | 0.41 | 0.27 | 0.89 | 0.98 | 0.23 | 21.1 |
| | | 76 | 1 | 0.74 | 0.21 | 6 | 0.11 | 2.47 | 0.38 | 0.24 | 0.91 | 0.87 | 0.23 | 34.6 |
| Near 3'-end | b33(l ₂) | 13 | 1 | 0.87 | 0.24 | 6 | 0.1 | 16.5 | 0.35 | 2.27 | 0.96 | 1.42 | 0.18 | 34.3 |
| | | 26 | 1 | 0.78 | 0.24 | 6 | 0.14 | 6.46 | 0.31 | 1.32 | 0.96 | 1.44 | 0.17 | 34.3 |
| | | 52 | 1 | 0.68 | 0.24 | 6 | 0.21 | 3.63 | 0.25 | 0.76 | 0.95 | 1.19 | 0.13 | 39.8 |
| | | 76 | 1 | 0.59 | 0.24 | 6 | 0.25 | 2.98 | 0.23 | 0.56 | 0.91 | 1.2 | 0.13 | 33.9 |
| Junction label | a12(T ₁) | 13 | 1 | 1.48 | 0.25 | 6 | 0.16 | 14.5 | 0.28 | 1.44 | 0.91 | 1.92 | 0.27 | 47.4 |
| | | 26 | 1 | 1.26 | 0.25 | 6 | 0.19 | 8.44 | 0.25 | 0.97 | 0.92 | 1.49 | 0.24 | 64.8 |
| | | 52 | 1 | 0.95 | 0.25 | 6 | 0.23 | 6.17 | 0.23 | 0.67 | 0.91 | 1.34 | 0.25 | 56.6 |
| | | 76 | 1 | 1.02 | 0.25 | 6 | 0.24 | 5.74 | 0.22 | 0.61 | 0.93 | 1.2 | 0.26 | 52.7 |

FCS curves were fitted to the model function described in eq. 3.25. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z₀/w₀ describes the shape of the observation volume, A_{ISO1-2} are the fitted amplitudes and t_{ISO1-2} the relaxation times of the cis-trans isomerization process, t_{Ab} is the photon anti bunching time with amplitude A_{Ab} and t_R is the rotational depolarization time with amplitude A_R. The FCS curves are displayed in Fig. 3.2D and Supplementary Fig. 3.7.

Supplementary Table 3.4. Fitted lifetime and anisotropy values for Alexa488 labels.

| label name | molecule | x ₁ | τ ₁ (ns) | x ₂ | τ ₂ (ns) | x ₃ | τ ₃ (ns) | ⟨τ⟩ _x (ns) | r ₁ | ρ1 (ns) | r ₂ | ρ2 (ns) | r ₃ (r _∞) | ρ3(global) (ns) |
|----------------------|----------|----------------|---------------------|----------------|---------------------|----------------|---------------------|-----------------------|----------------|---------|----------------|---------|----------------------------------|-----------------|
| free dye | free dye | 1.00 | 4.05 | | | | | | | | | | | |
| b11 | 4WJ | 0.93 | 4.02 | 0.07 | 1.45 | | | 3.83 | 0.16 | 0.22 | 0.11 | 1.37 | 0.12 | 17.81 |
| b11(T ₂) | 3WJ | 0.94 | 4.03 | 0.06 | 1.44 | | | 3.88 | 0.17 | 0.24 | 0.10 | 1.35 | 0.11 | 14.55 |
| b11(T ₃) | 3WJ | 0.92 | 4.00 | 0.08 | 1.43 | | | 3.80 | 0.16 | 0.24 | 0.11 | 1.40 | 0.12 | 16.80 |
| b11(T ₁) | 3WJ | 0.93 | 4.00 | 0.07 | 1.41 | | | 3.81 | 0.16 | 0.27 | 0.10 | 1.71 | 0.12 | 19.72 |
| b11(l ₂) | duplex | 0.94 | 4.05 | 0.06 | 1.35 | | | 3.88 | 0.16 | 0.24 | 0.10 | 1.35 | 0.12 | 12.99 |
| b28 | 4WJ | 0.77 | 3.91 | 0.16 | 2.07 | 0.07 | 0.32 | 3.37 | 0.16 | 0.25 | 0.12 | 1.63 | 0.10 | 18.76 |
| b28(T ₄) | 3WJ | 0.78 | 3.94 | 0.15 | 2.07 | 0.07 | 0.31 | 3.40 | 0.16 | 0.23 | 0.12 | 1.37 | 0.10 | 14.31 |
| b28(T ₃) | 3WJ | 0.78 | 3.95 | 0.15 | 2.14 | 0.07 | 0.28 | 3.42 | 0.16 | 0.23 | 0.12 | 1.39 | 0.10 | 13.60 |
| b28(T ₁) | 3WJ | 0.79 | 3.94 | 0.14 | 2.04 | 0.07 | 0.22 | 3.42 | 0.16 | 0.23 | 0.12 | 1.42 | 0.10 | 14.53 |
| b28(l ₂) | duplex | 0.80 | 3.96 | 0.14 | 2.07 | 0.07 | 0.25 | 3.45 | 0.16 | 0.23 | 0.12 | 1.43 | 0.10 | 11.78 |
| d23 | 4WJ | 0.80 | 3.97 | 0.13 | 2.28 | 0.07 | 0.35 | 3.49 | 0.16 | 0.25 | 0.11 | 1.52 | 0.11 | 16.29 |
| d23(T ₄) | 3WJ | 0.78 | 3.98 | 0.15 | 2.40 | 0.08 | 0.40 | 3.47 | 0.17 | 0.22 | 0.11 | 1.32 | 0.10 | 12.86 |
| d23(T ₂) | 3WJ | 0.78 | 3.99 | 0.14 | 2.39 | 0.08 | 0.36 | 3.49 | 0.18 | 0.30 | 0.10 | 1.82 | 0.10 | 16.02 |

| | | | | | | | | | | | | | | |
|----------------------|--------|------|------|------|------|------|------|------|------|------|------|------|------|-------|
| d23(I ₁) | duplex | 0.79 | 3.99 | 0.14 | 2.34 | 0.07 | 0.35 | 3.52 | 0.16 | 0.24 | 0.11 | 1.34 | 0.11 | 11.27 |
| c8 | 4WJ | 0.70 | 3.91 | 0.17 | 1.83 | 0.13 | 0.29 | 3.08 | 0.15 | 0.25 | 0.13 | 1.54 | 0.10 | 13.97 |
| c8(T ₄) | 3WJ | 0.80 | 3.85 | 0.20 | 1.25 | | | 3.35 | 0.19 | 0.33 | 0.10 | 2.09 | 0.09 | 19.98 |
| c8(T ₃) | 3WJ | 0.84 | 3.88 | 0.16 | 1.22 | | | 3.45 | 0.18 | 0.31 | 0.12 | 2.29 | 0.08 | 21.45 |
| c8(T ₁) | 3WJ | 0.71 | 3.93 | 0.14 | 1.78 | 0.16 | 0.26 | 3.05 | 0.17 | 0.23 | 0.13 | 1.32 | 0.08 | 13.27 |
| c8(I ₃) | duplex | 0.75 | 3.95 | 0.16 | 1.96 | 0.09 | 0.31 | 3.30 | 0.16 | 0.23 | 0.12 | 1.22 | 0.10 | 9.12 |
| b8 | 4WJ | 0.80 | 4.10 | 0.13 | 1.92 | 0.07 | 0.53 | 3.58 | 0.15 | 0.31 | 0.13 | 1.85 | 0.10 | 38.77 |
| b8(I ₂) | duplex | 0.75 | 4.06 | 0.16 | 2.07 | 0.09 | 0.57 | 3.43 | 0.17 | 0.22 | 0.13 | 1.34 | 0.08 | 13.35 |
| b14(I ₂) | duplex | 0.74 | 4.18 | 0.18 | 2.29 | 0.09 | 0.54 | 3.54 | 0.17 | 0.19 | 0.10 | 1.30 | 0.11 | 19.79 |
| b14 | 4WJ | 0.88 | 4.19 | 0.08 | 2.36 | 0.04 | 0.32 | 3.89 | 0.14 | 0.22 | 0.11 | 1.64 | 0.13 | 32.59 |
| a23(I ₂) | duplex | 0.85 | 4.06 | 0.16 | 1.24 | | | 3.62 | 0.18 | 0.25 | 0.12 | 1.55 | 0.08 | 14.33 |
| a26(I ₂) | duplex | 0.88 | 4.05 | 0.12 | 1.42 | | | 3.75 | 0.17 | 0.24 | 0.12 | 1.50 | 0.09 | 13.50 |
| c16(I ₂) | duplex | 0.80 | 4.19 | 0.07 | 1.81 | 0.13 | 0.19 | 3.51 | 0.24 | 0.14 | 0.10 | 1.46 | 0.14 | 20.94 |
| b5 | 4WJ | 0.78 | 4.13 | 0.15 | 2.60 | 0.07 | 0.63 | 3.66 | 0.19 | 0.23 | 0.12 | 1.40 | 0.07 | 21.03 |
| b5(T ₃) | 3WJ | 0.79 | 4.12 | 0.15 | 2.47 | 0.07 | 0.58 | 3.64 | 0.19 | 0.23 | 0.12 | 1.40 | 0.07 | 21.12 |
| b5(T ₂) | 3WJ | 0.93 | 4.11 | 0.07 | 1.26 | | | 3.92 | 0.22 | 0.24 | 0.11 | 1.36 | 0.06 | 14.39 |
| b5(T ₁) | 3WJ | 0.91 | 4.11 | 0.09 | 1.29 | | | 3.87 | 0.20 | 0.24 | 0.11 | 1.41 | 0.07 | 20.54 |
| b5(I ₂) | duplex | 0.92 | 4.11 | 0.08 | 1.31 | | | 3.88 | 0.19 | 0.23 | 0.11 | 1.24 | 0.07 | 17.67 |
| d7 | 4WJ | 0.80 | 4.07 | 0.11 | 1.65 | 0.09 | 0.41 | 3.48 | 0.17 | 0.23 | 0.15 | 1.44 | 0.06 | 13.58 |
| d7(T ₂) | 3WJ | 0.83 | 4.07 | 0.09 | 1.72 | 0.08 | 0.46 | 3.57 | 0.18 | 0.23 | 0.14 | 1.36 | 0.06 | 11.88 |
| d7(T ₄) | 3WJ | 0.83 | 4.07 | 0.10 | 1.71 | 0.07 | 0.41 | 3.59 | 0.18 | 0.23 | 0.14 | 1.31 | 0.06 | 10.04 |
| d7(I ₁) | duplex | 0.88 | 4.07 | 0.12 | 1.09 | | | 3.72 | 0.19 | 0.23 | 0.13 | 1.29 | 0.06 | 7.81 |
| b33 | 4WJ | 0.53 | 3.56 | 0.19 | 1.48 | 0.27 | 0.31 | 2.26 | 0.19 | 0.17 | 0.14 | 0.97 | 0.05 | 10.91 |
| b33(T ₄) | 3WJ | 0.60 | 3.54 | 0.18 | 1.53 | 0.23 | 0.32 | 2.45 | 0.20 | 0.20 | 0.14 | 0.96 | 0.05 | 8.68 |
| b33(T ₃) | 3WJ | 0.51 | 3.54 | 0.20 | 1.42 | 0.29 | 0.30 | 2.19 | 0.17 | 0.15 | 0.16 | 0.84 | 0.06 | 8.17 |
| b33(T ₁) | 3WJ | 0.60 | 3.58 | 0.17 | 1.62 | 0.23 | 0.34 | 2.49 | 0.20 | 0.23 | 0.14 | 1.07 | 0.05 | 10.80 |
| b33(I ₂) | duplex | 0.59 | 3.56 | 0.17 | 1.54 | 0.24 | 0.32 | 2.45 | 0.18 | 0.20 | 0.15 | 0.92 | 0.05 | 8.07 |
| d28 | 4WJ | 0.58 | 3.56 | 0.15 | 1.51 | 0.27 | 0.27 | 2.36 | 0.21 | 0.25 | 0.13 | 1.09 | 0.04 | 11.91 |
| d28(T ₂) | 3WJ | 0.60 | 3.57 | 0.14 | 1.53 | 0.26 | 0.28 | 2.43 | 0.21 | 0.25 | 0.13 | 1.10 | 0.04 | 15.08 |
| d28(T ₄) | 3WJ | 0.59 | 3.57 | 0.15 | 1.55 | 0.26 | 0.28 | 2.41 | 0.21 | 0.24 | 0.13 | 1.05 | 0.04 | 12.43 |
| d28(T ₁) | 3WJ | 0.60 | 3.56 | 0.14 | 1.51 | 0.26 | 0.27 | 2.42 | 0.21 | 0.25 | 0.13 | 1.05 | 0.04 | 12.51 |
| d28(I ₁) | duplex | 0.59 | 3.57 | 0.14 | 1.45 | 0.27 | 0.26 | 2.39 | 0.20 | 0.24 | 0.13 | 1.01 | 0.04 | 9.43 |
| c29 | 4WJ | 0.61 | 3.71 | 0.17 | 1.51 | 0.22 | 0.30 | 2.60 | 0.20 | 0.23 | 0.14 | 1.10 | 0.04 | 10.88 |
| c29(T ₄) | 3WJ | 0.62 | 3.72 | 0.16 | 1.48 | 0.22 | 0.29 | 2.61 | 0.20 | 0.22 | 0.14 | 1.03 | 0.04 | 8.76 |
| c29(T ₃) | 3WJ | 0.66 | 3.74 | 0.15 | 1.55 | 0.19 | 0.30 | 2.76 | 0.21 | 0.23 | 0.13 | 1.16 | 0.04 | 11.55 |
| c29(T ₂) | 3WJ | 0.65 | 3.74 | 0.16 | 1.51 | 0.20 | 0.29 | 2.70 | 0.21 | 0.24 | 0.13 | 1.20 | 0.04 | 13.78 |
| c29(I ₃) | duplex | 0.63 | 3.71 | 0.16 | 1.53 | 0.22 | 0.32 | 2.64 | 0.21 | 0.23 | 0.13 | 1.07 | 0.04 | 8.27 |
| c29(I ₁) | duplex | 0.63 | 3.66 | 0.16 | 1.48 | 0.20 | 0.30 | 2.62 | 0.21 | 0.24 | 0.13 | 1.17 | 0.04 | 10.35 |

The fitted lifetime components are τ_{1-3} and the corresponding fractions are x_{1-3} . The anisotropy decay is described by the three depolarization times ρ_{1-3} with corresponding fractions, r_{1-2} and r_{∞} . The longest anisotropy component arises from the global motion and its amplitude is the residual anisotropy (r_{∞}). The colors given to the label names indicate the dye environment. Bright blue: internal labels, orange: labels near the 5'-end, magenta: labels near the 3'-end.

Supplementary Table 3.5: Fit results for the FCS curves of different Alexa488 labels.

| Label type | Label name | offset | N_{bright} | t_{diff} (ms) | z_0/w_0 | A_Q | t_Q (ns) | A_T | t_T (μs) |
|-------------|----------------------|--------|---------------------|------------------------|-----------|-------|------------|-------|-------------------------|
| Free dye | Alexa488 | 1 | 6.87 | 0.17 | 3.62 | | | 0.24 | 3.74 |
| Internal | b11(l ₂) | 1 | 4.09 | 0.84 | 3.34 | 0.07 | 79.34 | 0.2 | 5.28 |
| | b28(l ₂) | 1 | 4.32 | 0.76 | 3.72 | 0.12 | 86.06 | 0.18 | 4.9 |
| | d23(l ₁) | 1 | 9.6 | 0.77 | 4.06 | 0.1 | 66.26 | 0.19 | 5.01 |
| | c8(l ₃) | 1 | 7.02 | 0.72 | 4.92 | 0.13 | 73.17 | 0.17 | 5.21 |
| Near 5'-end | b5(l ₂) | 1 | 6.98 | 0.82 | 3.69 | 0.07 | 98.34 | 0.18 | 4.65 |
| | d7(l ₁) | 1 | 7.92 | 0.7 | 2.96 | 0.04 | 130.7 | 0.21 | 4.33 |
| Near 3'-end | b33(l ₂) | 1 | 7.98 | 0.76 | 4.73 | 0.33 | 45.82 | 0.12 | 3.92 |
| | d28(l ₁) | 1 | 4.79 | 0.71 | 4 | 0.47 | 46.85 | 0.11 | 3.59 |
| | c29(l ₃) | 1 | 5.03 | 0.7 | 4.17 | 0.53 | 60.49 | 0.09 | 3.71 |

FCS curves were fitted to the model function described in eq. 3.23. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z_0/w_0 describes the shape of the observation volume, t_T is the relaxation time of the triplet state with fraction A_T , and t_Q is the relaxation time of the static quenching process with the fraction of statically quenched molecules A_Q . The label d7 required an additional bunching term to achieve a good fit with an amplitude of 0.04 and a relaxation time of 80 μs . The FCS curves are displayed in Fig. 3.3D and Supplementary Fig. 3.10.

Supplementary Table 3.6: Comparison of dynamic ($\Phi_{F,\text{dyn}}$) and steady state (Φ_F) quantum yields of Alexa488 labels.

| Label name | Molecule | $\Phi_{F,\text{dyn}}$ | Φ_F (liFCS) | f_{static} | Φ_F (spectral) |
|----------------------|----------|-----------------------|------------------|---------------------|---------------------|
| b11(l ₂) | duplex | 0.92 | 0.85 | 0.08 | 0.9 |
| b28(l ₂) | duplex | 0.82 | 0.74 | 0.1 | 0.76 |
| d23(l ₁) | duplex | 0.84 | 0.72 | 0.14 | |
| c8(l ₃) | duplex | 0.78 | 0.71 | 0.09 | |
| b5(l ₂) | duplex | 0.92 | 0.8 | 0.14 | |
| d7(l ₁) | duplex | 0.88 | 0.75 | 0.15 | 0.8 |
| b33(l ₂) | duplex | 0.58 | 0.35 | 0.4 | 0.42 |
| d28(l ₁) | duplex | 0.57 | 0.33 | 0.42 | |
| c29(l ₃) | duplex | 0.63 | 0.39 | 0.37 | 0.34 |

The dynamic quantum yield was determined from the species-averaged fluorescence lifetime given in Supplementary Table 3.4 using eq. 3.13. The steady-state quantum yield was determined using the low-intensity FCS (liFCS) method. The fraction of statically quenched species (f_{static}) was calculated from the measured dynamic and steady-state quantum yields using eq. 3.18. As a control, the steady-state quantum yield was also measured from the spectral emission, which yielded consistent results (see Supplementary Fig. 3.11A for a comparison).

Supplementary Table 3.7: Parameters used in AV simulations.

| Parameter | Cy5 | Alexa488 |
|-----------------------|--|----------|
| L_{link} (Å) | 22.0 (U attached) 19.5 (G attached) | 20.0 |
| W_{link} (Å) | 2.5 | 2.5 |
| R_{dye} (Å) | 4.5 | 4.5 |

The dye is approximated by a sphere with radius R_{dye} and is attached to the RNA base by a flexible linker of length L_{link} and width W_{link} , as illustrated in Supplementary Fig. 3.15.

Supplementary Table 3.8: Standard deviations for the ensemble measurements of lifetime and residual anisotropy for Alexa488 and Cy5 labels.

| Alexa488 labels | Mean $\langle\tau\rangle_x \pm$ SD (ns) | Mean $r_{\infty} \pm$ SD | Cy5 labels | Mean $\langle\tau\rangle_x \pm$ SD (ns) | Mean $r_{\infty} \pm$ SD |
|----------------------|---|--------------------------|----------------------|---|--------------------------|
| b11(4WJ) | 3.81±0.04 | 0.12±0.01 | b27(4WJ) | 1.21±0.04 | 0.18±0.01 |
| b11(T ₂) | 3.85±0.04 | 0.12±0.01 | b27(l ₂) | 1.19±0.04 | 0.16±0.02 |
| b11(T ₃) | 3.78±0.03 | 0.12±0.01 | d23(4WJ) | 1.21±0.07 | 0.16±0.03 |
| b11(T ₁) | 3.80±0.01 | 0.13±0.01 | c8(4WJ) | 1.24±0.04 | 0.19±0.02 |
| b11(l ₂) | 3.86±0.03 | 0.12±0.01 | c8(l ₃) | 1.15±0.02 | 0.15±0.01 |
| d23(4WJ) | 3.55±0.09 | 0.11±0.01 | a26(4WJ) | 1.12±0.01 | 0.13±0.01 |
| c8(4WJ) | 3.10±0.03 | 0.09±0.01 | a12(T ₁) | 1.42±0.11 | 0.24±0.04 |
| c8(T ₃) | 3.38±0.10 | 0.09±0.01 | d27(4WJ) | 1.21±0.01 | 0.17±0.01 |
| c8(T ₁) | 3.20±0.16 | 0.08±0.02 | c12(4WJ) | 1.20±0.01 | - |
| c8(l ₃) | 3.27±0.04 | 0.10±0.01 | b33(4WJ) | 1.51±0.22 | 0.28±0.02 |
| b8(l ₂) | 3.47±0.06 | 0.08±0.01 | b33(l ₂) | 1.64±0.07 | 0.28±0.01 |
| b14(l ₂) | 3.49±0.07 | 0.11±0.01 | d28(4WJ) | 1.51±0.09 | 0.25±0.02 |
| b5(4WJ) | 3.52±0.20 | 0.07±0.01 | | | |
| b5(l ₂) | 3.83±0.06 | 0.07±0.01 | | | |
| d7(T ₂) | 3.57±0.01 | 0.06±0.01 | | | |
| b33(T ₃) | 2.13±0.08 | 0.05±0.01 | | | |

“Mean $\langle\tau\rangle_x \pm$ SD” and “Mean $r_{\infty} \pm$ SD” refer to the average of the species-weighted fluorescence lifetimes and residual anisotropies, respectively, over all repeated measurements for a given sample. The samples were measured between 2-4 times. SD is the standard deviation calculated for repeated measurements performed on different days. In many cases, the samples for the repeated measurements were newly hybridized from single strands.

Supplementary Table 3.9: Fit results for the FCS curves of Cy5 labels at position a12 in the different molecules.

| Molecule | offset | N_{bright} | t_{diff} (ms) | z_0/w_0 | A_{ISO1} | t_{ISO1} (μs) | A_{ISO2} | t_{ISO2} (μs) | A_{ISO} | t_{ISO} (μs) | A_{Ab} | t_{Ab} (ns) | A_{R} | t_{R1} (ns) |
|------------------------|--------|---------------------|------------------------|-----------|-------------------|-------------------------------------|-------------------|-------------------------------------|------------------|------------------------------------|-----------------|----------------------|----------------|----------------------|
| 4WJ | 1 | 2.5 | 0.29 | 6 | 0.08 | 5.57 | 0.41 | 0.61 | 0.49 | 0.72 | 0.88 | 1.13 | 0.08 | 35.76 |
| 3WJ, T ₁ | 1 | 2.5 | 0.24 | 6 | 0.2 | 7.28 | 0.24 | 0.87 | 0.44 | 1.65 | 0.93 | 1.58 | 0.11 | 33.58 |
| Duplex, I ₃ | 1 | 1.9 | 0.25 | 6 | 0.08 | 3.73 | 0.43 | 0.55 | 0.51 | 0.54 | 0.89 | 1.12 | 0.08 | 13.11 |

FCS curves were fitted to the model function described in eq. 3.25. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z_0/w_0 describes the shape of the observation volume, $A_{\text{ISO},1-2}$ are the fitted amplitudes and $t_{\text{ISO},1-2}$ the relaxation times of the cis-trans isomerization process, t_{Ab} is the photon antibunching time with amplitude A_{Ab} , t_{R1} is the longest rotational depolarization time and A_{R} is the amplitude of the rotational term. The total fraction of molecules in the cis state, $A_{\text{ISO}} = A_{\text{ISO1}} + A_{\text{ISO2}}$ and t_{ISO} is the amplitude weighted average of t_{ISO1} and t_{ISO2} . The FCS curves are displayed in Supplementary Fig. 3.29.

Supplementary Table 3.10: Fit results for FCS curves of Cy5 labels at position b33 on different molecules.

| Molecule | offset | N_{bright} | t_{diff} (ms) | z_0/w_0 | A_{ISO1} | t_{ISO1} (μs) | A_{ISO2} | t_{ISO2} (μs) | A_{Ab} | t_{Ab} (ns) | A_{R} | t_{R1} (ns) | t_{R2} (ns) |
|------------------------|--------|---------------------|------------------------|-----------|-------------------|-------------------------------------|-------------------|-------------------------------------|-----------------|----------------------|----------------|----------------------|----------------------|
| 4WJ | 1 | 1.3 | 0.31 | 6 | 0.16 | 7.36 | 0.26 | 1.64 | 0.9 | 1.62 | 0.11 | 38.9 | 11.7 |
| 3WJ, T ₁ | 1 | 1.6 | 0.27 | 6 | 0.17 | 6.62 | 0.25 | 1.62 | 0.93 | 1.51 | 0.10 | 38.3 | 11.5 |
| 3WJ, T ₃ | 1 | 2.0 | 0.29 | 6 | 0.14 | 7.88 | 0.28 | 1.8 | 0.94 | 1.5 | 0.10 | 38.7 | 11.6 |
| duplex, I ₂ | 1 | 1.9 | 0.24 | 6 | 0.15 | 8.18 | 0.28 | 1.81 | 0.96 | 1.53 | 0.08 | 24.2 | 7.26 |

FCS curves were fit to the model function specified in eq. 3.29, as shown in Fig. 3.6C. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z_0/w_0 is the ratio of the axial to the lateral width of the observation volume, $A_{\text{ISO},1-2}$ are the fitted amplitudes and $t_{\text{ISO},1-2}$ the relaxation times of the cis-trans isomerization process, t_{Ab} is the photon anti bunching time with amplitude A_{Ab} , $t_{\text{R1-2}}$ are the rotational depolarization times and A_{R} is the amplitude of the rotational term. The curves were measured at a power density of 11 kW/cm².

Supplementary Table 3.11: Fit results for the polarization-resolved FCS curves of Cy5 labels the terminal position c1 on the specified molecules.

| Molecule | A_{R} | t_{R1} (ns) | $t_{\text{R2}} = 0.3 \cdot t_{\text{R1}}$ (ns) |
|---------------------|----------------|----------------------|--|
| 4WJ | 0.13 | 38.03 | 11.41 |
| 3WJ, T ₃ | 0.11 | 34.13 | 10.24 |
| 3WJ, T ₄ | 0.13 | 28.36 | 8.51 |
| duplex | 0.09 | 23.91 | 7.17 |

FCS curves were fit to the model function for a spherical rotor as given by eq. 3.29. $t_{\text{R1-2}}$ are the rotational depolarization times with corresponding amplitudes $A_{\text{R1-2}}$. The FCS curves are shown in Supplementary Fig. 3.30.

Supplementary Table 3.12: Fit results for the polarization-resolved FCS curves of Alexa488 labels at the internal position b11 on the specified molecules.

| Molecule | A_R | t_{R1} (ns) | $t_{R2} = 0.3 \cdot t_{R1}$ (ns) |
|----------|-------|---------------|----------------------------------|
| 4WJ | 0.48 | 35.31 | 10.59 |
| 3WJ, T3 | 0.46 | 31.45 | 9.43 |
| duplex | 0.49 | 22.28 | 6.69 |

The FCS curves were fit to the model given in eq. 3.23 in the methods, using the model function of the spherical rotor as given in eq. 3.24 to describe the molecular rotation. FCS curves are shown in Supplementary Fig. 3.31.

Supplementary Table 3.13: Scaling factors used to overlay the fitted rotational correlation terms from the experimental pFCS curves using the model for the spherical rotor (eq. 3.29) with the predicted curves based on the HYDROPRO model as given in Supplementary Fig. 3.32.

| RNA molecule | t_R' (pred) [ns] | t_R' (fit) [ns] | t_R' (pred) / t_R' (fit) | $A_{R(\text{pred})}$ | $A_{R(\text{fit})}$ | $A_{R(\text{pred})} / A_{R(\text{fit})}$ |
|------------------------|--------------------|-------------------|------------------------------|----------------------|---------------------|--|
| 4WJ | 91.0 | 51.9 | 1.75 | 0.19 | 0.11 | 1.67 |
| 3WJ, T ₁ | 80.2 | 51.1 | 1.57 | 0.19 | 0.10 | 1.96 |
| 3WJ, T ₃ | 60.1 | 51.6 | 1.17 | 0.19 | 0.10 | 1.99 |
| Duplex, I ₂ | 25.3 | 32.8 | 0.77 | 0.19 | 0.08 | 2.27 |

The given relaxation times t_R' correspond to the time point where the amplitude-scaled correlation amplitude of the rotational term decayed to a value of 1.1. $A_{R(\text{fit})}$ is the amplitude of the rotational term obtained from the fitting of experimental FCS curve (Fig. 3.6C) to the model function given by eq. 3.29. $A_{R(\text{pred})} = 0.19$ is the theoretical amplitude of the rotation contribution for the spherical rotor obtained from eq. 3.46.

Supplementary Table 3.14: Fit results for polarization-resolved FCS curves of end-stacked Cy5 labels in the different RNA molecules.

| | duplex | | | 4WJ | | | 3WJ, T ₃ | | | 3WJ, T ₁ | | |
|---|--|---------------|--------------|---------------|---------------|--------------|---------------------|---------------|--------------|---------------------|---------------|--------------|
| | predicted rotational diffusion tensor | | | | | | | | | | | |
| D_x (MHz) | 3.30 | | | 1.97 | | | 2.33 | | | 2.09 | | |
| D_y (MHz) | 3.37 | | | 2.09 | | | 3.02 | | | 2.99 | | |
| D_z (MHz) | 19.06 | | | 3.33 | | | 4.97 | | | 3.97 | | |
| D (MHz) | 8.58 | | | 2.46 | | | 3.44 | | | 3.02 | | |
| 1/(6D) (ns) | 19.43 | | | 67.67 | | | 48.49 | | | 55.29 | | |
| 1/(20D) (ns) | 5.83 | | | 20.30 | | | 14.55 | | | 16.58 | | |
| D_z/D_{xy} | 5.72 | | | 1.64 | | | 1.86 | | | 1.56 | | |
| | fit result | | | | | | | | | | | |
| Model | spher. | scaled | prol. | spher. | scaled | prol. | spher. | scaled | prol. | spher. | scaled | prol. |
| $\chi^2_{\text{red.}}$ | 0.737 | 0.664 | 0.643 | 0.802 | 0.797 | 0.793 | 0.787 | 0.784 | 0.776 | 0.837 | 0.801 | 0.797 |
| A_R | 0.070 | 0.081 | 0.086* | 0.084 | 0.086 | 0.086* | 0.081 | 0.083 | 0.086* | 0.086 | 0.086 | 0.086* |
| f_r | - | 0.80 | 0.51 | - | 2.05 | 1.65 | - | 1.52 | 1.06 | - | 1.72 | 1.50 |
| D_x (MHz) | - | 2.64 | 1.70 | - | 4.03 | 3.34 | - | 3.54 | 2.84 | - | 3.60 | 3.81 |
| D_y (MHz) | - | 2.70 | 1.70 | - | 4.28 | 3.34 | - | 4.58 | 2.84 | - | 5.15 | 3.81 |
| D_z (MHz) | - | 15.27 | 17.82 | - | 6.82 | 8.83 | - | 7.54 | 8.62 | - | 6.84 | 9.40 |
| D (MHz) | 6.61 | 6.87 | 7.07 | 5.37 | 5.04 | 5.17 | 5.19 | 5.22 | 4.76 | 5.90 | 5.20 | 5.67 |
| 1/(6D) (ns) | 25.21 | 24.26 | 23.56 | 31.03 | 33.04 | 32.26 | 32.10 | 31.95 | 34.99 | 28.24 | 32.08 | 29.38 |
| 1/(20D) (ns) | 7.56 | 7.28 | 7.07 | 9.31 | 9.91 | 9.62 | 9.63 | 9.59 | 10.50 | 8.47 | 9.62 | 8.81 |
| Asym. (D_z/D_{xy}) | - | 5.72 | 10.48 | - | 1.64 | 2.64 | - | 1.86 | 3.04 | - | 1.56 | 2.47 |
| | excess D_z compared to prediction | | | | | | | | | | | |
| ΔD_z (MHz) | - | - | 8.1 | - | - | 3.35 | - | - | 3.34 | - | - | 3.45 |

Experimental curves were computed from the parallel and perpendicular signals and fit using the general model function given in eq. 3.29-30 using the spherical rotor model as given by eq. 3.29 or the asymmetric model function as described in the methods and given in eq. 3.31. The orientation of the transition dipole moment was fixed to the average orientation obtained from the MD simulations. For the asymmetric model function, the rotational diffusion coefficients, $D_{x/y/z}$, were fixed to the values obtained from the HYDROPRO software based on the FRET-restrained structural models of the RNA molecules and scaled by a common factor f_r . The theoretically predicted values for amplitude of the rotational correlation term, A_R , is $A_R = 0.19$. For the prolate fit model, the amplitude was fixed to $A_R = 0.086$ as obtained for the

asymmetric model (indicated by an asterisk, *). The asymmetry is defined as the ratio of D_z and D_{xy} , where D_{xy} is the average of the x- and y-components. The excess D_z compared to the prediction (ΔD_z), is computed as the difference between the fitted and scaled predicted values for D_z , $\Delta D_z = D_z(\text{fit}) - f_r D_z(\text{pred})$, where f_r for the prolate model is obtained as the ratio of the fitted and measured values for D_{xy} , $f_r = D_{xy}(\text{fit})/D_{xy}(\text{pred})$ (see eq. 3.32-33).

Supplementary Table 3.15: Fluorescence quantum yields for representative Cy5 labels.

| Cluster | Label position | Molecule | $\Phi_{F,\text{dyn}}$ | A_{ISO} | $\Phi_{F,\text{corr}}$ |
|---------------------------|----------------|------------------------|-----------------------|-----------|------------------------|
| Internal U-attached label | b27 | duplex, I ₂ | 0.40 | 0.49 | 0.20 |
| Internal G-attached label | d27 | duplex, I ₁ | 0.38 | 0.52 | 0.18 |
| Near-3'-end label | b33 | duplex, I ₂ | 0.55 | 0.45 | 0.30 |
| Near-junction label | a12 | 3WJ, T ₁ | 0.48 | 0.44 | 0.27 |

$\Phi_{F,\text{corr}}$ is the fluorescence quantum yield corrected for cis-trans dark state using the equation $\Phi_{F,\text{corr}} = \Phi_{F,\text{dyn}}(1 - A_{ISO})$, $\Phi_{F,\text{dyn}}$ is the fluorescence quantum yield calculated from species-weighted average lifetime, and A_{ISO} is the amplitude of cis-trans relaxation determined from FCS measurements.

Supplementary Table 3.16: The polarization coefficients $c_{(a_1 e_1), (a_2 e_2)}^{(L)}$ as given in eq. 3.10 for standard settings of the excitation and emission polarizers. The coefficients are scaled to be consistent with previous work by Kask et al.²⁹⁹ and must be multiplied by a factor $(5 \cdot 7 \cdot 9)^2$ for use in eq. 3.10.

| L | 0 | 2 | 4 |
|--|------|------|------|
| $c_{(III), (III)}^{(L)}$ | 3969 | 6480 | 576 |
| $c_{(II\perp), (II\perp)}^{(L)}$ | 441 | 180 | 324 |
| $c_{(III), (II\perp)}^{(L)}$ $c_{(II\perp), (III)}^{(L)}$ | 1323 | 540 | -288 |

Note that coefficients for both cross-correlations $c_{(III), (II\perp)}^{(L)}$ and $c_{(II\perp), (III)}^{(L)}$ are the same, and that an error in Kask et al. for the coefficient for $L = 4$ has been corrected. The subscripts denote the polarizations of the excitation-emission cycle, i.e., $c_{(III), (II\perp)}^{(L)}$ is the coefficient for excitation by parallelly polarized light, detection in the parallel channel, re-excitation by parallelly polarized light and detection in the perpendicular channel.

Supplementary Table 3.17: The normalized polarization coefficients R_p and x_p as used in eq. 3.31 of the main text and eq. 3.12 in Supplementary Note 3.3 for standard settings of the excitation and emission polarizers (p).

| p | R_p | x_p | C_p |
|------------------------------|------------------------------|-------------------------------|--------------------------------|
| (), () | $\frac{16}{9} \approx 1.778$ | $\frac{45}{49} \approx 0.918$ | $\frac{4}{45} \approx 0.089$ |
| (⊥), (⊥) | $\frac{8}{7} \approx 1.143$ | $\frac{5}{14} \approx 0.357$ | $\frac{9}{5} = 0.180$ |
| (), (⊥) (⊥), () | $\frac{4}{21} \approx 0.190$ | $\frac{15}{7} \approx 2.143$ | $-\frac{8}{15} \approx -0.533$ |

Supplementary Notes

Supplementary Note 3.1: Base flipping does not cause significant changes in photophysical properties at room temperature

To investigate whether internal guanines could spontaneously un-pair and flip out of the stacked helix to quench the neighboring Alexa488 labels by PET, control experiments were performed on RNA duplexes containing mismatches using the following sequence:



Here, a mismatch at position X (marked in red) may cause the guanine (yellow) adjacent to the dye attachment point on uracil (green) to flip out of the RNA helix. We studied different RNA duplexes where the complementary base to guanine (position X) is:

- (1) cytosine, resulting in a Watson Crick base pair GC,
- (2) adenine, resulting in a mismatched GA pair,
- (3) uracil, resulting in a mismatched GU pair,
- (4) or guanine, resulting in a mismatched GG pair.

The variants (2), (3) and (4) with mismatched base pairs could potentially show base flipping. The labeling chemistry and the length of duplex mimic duplexes I₁₋₄ of our primary study as closely as possible. We performed time-resolved fluorescence measurements of the different constructs and fitted the resulting decays to a three-component exponential decay and determined the dynamic quantum yield $\Phi_{F,dyn}$ and the steady-state quantum yield Φ_F (Supplementary Table SN1.1). We observed no significant dynamic or static quenching for the mismatches compared to the fully base-

paired duplex, indicating that no significant amount of base flipping occurs at ambient temperatures ($T = 20\text{ }^{\circ}\text{C}$). This excludes the possibility that PET quenching of Alexa488 labels could be caused by temporary unpairing of internal guanines.

Supplementary Table SN1.1: Investigation of base flipping of guanine bases. Photophysical properties of Alexa488 labels adjacent to base-paired or mismatched guanine bases, measured at 20°C . See Supplementary Note 3.1 for experimental details. The steady-state quantum yield Φ_F was estimated using the liFCS method.

| Mismatch | ensemble time resolved fluorescence | | | | | | | | ensemble time resolved anisotropy | | | | | | | |
|-------------------|-------------------------------------|------------------|-------|------------------|-------|------------------|--------------------------------|----------------|-----------------------------------|------------------|-------|------------------|---------------------------|----------------------|----------|--------------|
| | x_1 | τ_1 (ns) | x_2 | τ_2 (ns) | x_3 | τ_3 (ns) | $\langle\tau\rangle_x$ (ns) | $\Phi_{F,dyn}$ | r_1 | ρ_1 (ns) | r_2 | ρ_2 (ns) | r_3 (r_{∞}) | ρ_3 (global) | Φ_F | f_{static} |
| Fully base paired | 0.84 | 4.15 | 0.16 | 1.94 | | | 3.79 | 0.90 | 0.14 | 0.21 | 0.10 | 1.37 | 0.14 | 8.77 | 0.89 | 0.01 |
| With GA mismatch | 0.92 | 4.15 | 0.08 | 1.88 | | | 3.96 | 0.94 | 0.15 | 0.21 | 0.10 | 1.12 | 0.13 | 9.27 | 0.91 | 0.03 |
| With GU mismatch | 0.91 | 4.15 | 0.09 | 2.41 | | | 4.00 | 0.95 | 0.13 | 0.20 | 0.09 | 1.14 | 0.16 | 9.32 | 0.86 | 0.09 |
| With GG mismatch | 0.83 | 4.19 | 0.11 | 2.46 | 0.05 | 0.56 | 3.80 | 0.90 | 0.14 | 0.19 | 0.11 | 1.14 | 0.13 | 9.60 | 0.89 | 0.01 |

Supplementary Note 3.2: FRET-derived distances are independent of the donor quantum yield

In this section, we show that FRET-derived distances are independent of the donor quantum yield if the same donor quantum yield is used for the determination of the Förster radius R_0 and the calculation of accurate FRET efficiencies.

The Förster radius is defined as:

$$R_0 = 0.2108 \sqrt[6]{\Phi_{F,D} \kappa^2 J n^{-4}}, \quad (3.43)$$

where $\Phi_{F,D}$ is the donor fluorescence quantum yield, κ^2 is an orientational factor that depends on the mutual orientation of the donor emission and acceptor absorption transition dipole moments, J is the spectral overlap integral and n is the refractive index of the surrounding medium.

The FRET efficiency is calculated from the corrected signal in the donor and acceptor channel, F_D and F_A , as²⁶⁸:

$$E = \frac{F_A}{\gamma F_D + F_A}, \quad (3.44)$$

where the γ -factor corrects for the different detection efficiencies and quantum yields of the donor and acceptor fluorophores, defined as:

$$\gamma = \frac{g_{R|A} \Phi_{F,A}}{g_{G|D} \Phi_{F,D}}. \quad (3.45)$$

Note that F_D and F_A are corrected for background signal, crosstalk of the donor fluorophore into the acceptor detection channel, and direct excitation of the acceptor fluorophore by the donor excitation laser.

The FRET efficiency is related to the donor-acceptor separation distance R_{DA} by:

$$E = \frac{1}{1 + \left(\frac{R_{DA}}{R_0}\right)^6} \quad (3.46)$$

Combining the above equations, we obtain the direct relation between R_{DA} and the corrected signals:

$$R_{DA} = R_0 \left(\frac{1}{E} - 1\right)^{\frac{1}{6}} = R_0 \left(\frac{g_{R|A} \Phi_{F,A} F_D}{g_{G|D} \Phi_{F,D} F_A}\right)^{\frac{1}{6}} = R_{0,r} \left(\Phi_{F,A} \frac{g_{R|A} F_D}{g_{G|D} F_A}\right)^{\frac{1}{6}}, \quad (3.47)$$

where we define the reduced Förster radius $R_{0,r}$ without consideration of the donor quantum yield as:

$$R_{0,r} = 0.2108 \sqrt[6]{\kappa^2 J n^{-4}} = \Phi_{F,D}^{-1/6} R_0. \quad (3.48)$$

In eq. 3.47, the donor fluorescence quantum yield cancels. Thus, the FRET-derived distance R_{DA} is independent of the donor quantum yield. Identical considerations apply to time-resolved FRET measurements ²⁷⁰.

Supplementary Note 3.3: Derivation of rotational correlation functions for arbitrary diffusion tensors (performed by Dr. Oleg Opanasyuk)

Although the measured polarization-resolved FCS correlation functions are second-order fluorescence intensity correlation function, they represent 4th-order correlation functions $G^{(4)}(t_1, t'_1, t_2, t'_2)$ of the excitation and emission probabilities, meaning they depend on four time points in the following sequence of events. First, the fluorophore is excited by light of a given polarization at time t_1 , after which it can rotate during its excited state lifetime until it relaxes back to the ground state by emission of a photon at time t'_1 . After an additional time delay, the fluorophore is re-excited at time t_2 and the second photon is emitted at time t'_2 . This sequence of events is illustrated in the following scheme:

$$\begin{array}{c} \text{fl. delay } \Delta t_1 \qquad \qquad \qquad \text{fl. delay } \Delta t_2 \\ \overbrace{t_1 \xrightarrow{\Delta\Omega_1(\Delta t_1)} t'_1} \quad \xrightarrow{\Delta\Omega_{21}(t-\Delta t_2)} \quad \overbrace{t_2 \xrightarrow{\Delta\Omega_2(\Delta t_2)} t'_2} \\ \Omega_1 \qquad \qquad \qquad \Omega'_1 \qquad \qquad \qquad \Omega_2 \qquad \qquad \qquad \Omega'_2 \quad , \quad (3.49) \\ \underbrace{e_1 \xrightarrow{\text{correlation time } t_c} e_2} \end{array}$$

where Ω_i denotes the orientation of the fluorophore dipole moment at step i and e_i marks the emission of a photon at time t_i . The excitation probabilities at time t_1 and t_2 depend

on the polarization of the incident laser light, and the emission probabilities at times t'_1 and t'_2 depend on the polarization of the detection channel. The fourth-order correlation function for the rotational contribution is expressed as²⁹⁹:

$$G^{(4)}(t_1, t'_1, t_2, t'_2) = \langle a(\Omega_1)e(\Omega'_1)a(\Omega_2)e(\Omega'_2) \rangle, \quad (3.50)$$

where $a(\Omega)$ and $e(\Omega)$ denote the absorption and emission probabilities and the angular brackets denote the average over all orientations Ω_i . The second-order correlation function is then obtained by integration over all possible time delays $\Delta t_1 = t'_1 - t_1$ and $\Delta t_2 = t'_2 - t_2$ consistent with the correlation time t , i.e., the time between the two photon detection events at the times t'_1 and t'_2 . The problem can be simplified if the excited state lifetime is much shorter compared to the rotational diffusion time (i.e., $6D_{\text{rot}}\tau \ll 1$). In this case, the rotational contribution to the correlation function, $G_{\text{rot}}(t_c)$, is given by:

$$G_{\text{rot}}(t_c) = G^{(4)}(0,0,t_c,t_c) = \langle a(0)e(0)a(t_c)e(t_c) \rangle. \quad (3.51)$$

In this case the rotational contribution $G_{\text{rot}}(t)$ can be decoupled from all other contributions, and the total intensity correlation function can be factorized:

$$G(t) = \frac{1}{N_{\text{bright}}} \underbrace{G_{AB}(t)}_{\sim 1 \text{ ns}} \underbrace{G_{\text{rot}}(t)}_{\sim 10-100 \text{ ns}} \underbrace{G_{ISO}(t)}_{\sim 1 \mu\text{s}} \underbrace{G_{\text{diff}}(t)}_{\sim 1 \text{ ms}} + 1, \quad (3.52)$$

where the contribution of photon antibunching ($G_{AB}(t)$), photoisomerization ($G_{ISO}(t)$), and diffusion ($G_{\text{diff}}(t)$) are described as given above.

Explicit expressions for the rotational correlation function have previously been derived for the spherical and ellipsoidal rotors^{114,299}. Here, we derive expressions for the rotational correlation function for arbitrary diffusion tensors $\mathcal{D} = (D_x, D_y, D_z)$ based on the matrix method presented by Kalmykov²⁸³. According to this formalism, the rotational correlation functions for isotropic rotation can be expressed as the weighted sum of elements of Wigner D -matrices $D_{nm}^{(l)}(\Omega)$ averaged over the time-dependent angular distribution $p(\Omega, t)$:

$$G_{\text{rot}}(t_c) = \sum_l \mathbf{c}^{(l)} \cdot \mathcal{F}^{(l)}(t_c) = \sum_{l,p,q} C_{p,q}^{(l)} \mathcal{F}_{p,q}^{(l)}(t_c), \quad (3.53)$$

where the functions $\mathcal{F}_{p,q}^{(l)}(t_c)$ are given by:

$$\mathcal{F}_{p,q}^{(l)}(t) = \langle D_{p,q}^{(l)}(\Omega) \rangle_{p(\Omega,t)} = \int D_{p,q}^{(l)}(\Omega) p(\Omega, t) d\Omega, \quad (3.54)$$

and the angular probability distribution function $p(\Omega, t)$ follows the Smoluchovski diffusion equation. The diffusion averaged time-dependent functions $\mathcal{F}_{p,q}^{(l)}(t)$ can be expressed as:

$$\mathcal{F}^{(l)}(t) = e^{\mathbf{A}^{(l)}(\mathcal{D})t}, \quad (3.55)$$

where the matrix $\mathbf{A}^{(l)}$ is a function of the (diagonalized) rotational diffusion tensor $\mathcal{D} = (D_x, D_y, D_z)$:

$$\begin{aligned} [\mathbf{A}^{(l)}(\mathcal{D})]_{p,q} &= -2D(c_p^l \delta_{pq} + a_p^l \delta_{p,q-2} + a_{p-2}^l \delta_{p,q+2}); \\ c_p^l &= \frac{l(l+1)}{2} + \Delta p^2; \quad a_p^l = \frac{1}{4} \theta \sqrt{[l^2 - (p+1)^2][(l+1)^2 - (p+1)^2]}; \\ D &= \frac{D_x + D_y}{2}; \quad \Delta = \frac{D_z}{D_x + D_y} - \frac{1}{2}; \quad \theta = \frac{D_x - D_y}{D_x + D_y} \end{aligned} \quad (3.56)$$

Here, the parameter θ is a measure of asymmetry of the diffusion tensor \mathcal{D} in the xy plane, and the parameter Δ is the measure of asymmetry relative to the z-axis. In practice, the matrix exponential in eq. 3.55 can be calculated using the eigen-value decomposition of the matrix $\mathbf{A}^{(l)}$:

$$\mathcal{F}^{(l)}(t) = e^{\mathbf{A}^{(l)}t} = \mathbf{U}^{(l)} e^{\mathbf{\Lambda}^{(l)}t} \mathbf{U}^{(l)T}, \quad (3.57)$$

where the matrix \mathbf{U} consist of normalized eigen-vectors of $\mathbf{A}^{(l)}$ as rows and $\mathbf{\Lambda}$ is the diagonal matrix of eigen-values of $\mathbf{A}^{(l)}$.

For collinear absorption and emission transition dipole moments, the rotational correlation function $G_{\text{rot}}^{(a_1 e_1), (a_2 e_2)}(t_c)$ for a certain combination of excitation and emission polarizer settings $(a_1 e_1), (a_2 e_2)$, is given by:

$$G_{\text{rot}}^{(a_1 e_1), (a_2 e_2)}(t_c) = \sum_{L=0,2,4} c_{(a_1 e_1), (a_2 e_2)}^{(L)} g^{(L)}(t_c), \quad (3.58)$$

Here, the time-dependent functions $g^{(L)}(t)$ are given by:

$$g^{(L)}(t_c) = \sum_i c_i^{(L)} e^{-k_i^{(L)} t_c}; \quad c_i^{(L)} = \left| \sum_Q D_{Q,0}^{(L)}(\Omega_\mu) u_{Q,i}^L \right|^2; \quad \sum_i c_i^{(L)} = 1, \quad (3.59)$$

where $k_i^{(L)}$ are the eigenvalues and $u_{Q,i}^L$ the eigen-vectors of the matrix $\mathbf{A}^{(l)}$ defined in eq. 3.57, and Ω_μ are Euler angles defining the position of the transition dipole moment μ in the eigen-frame of the diffusion tensor, i.e., the frame where the diffusion tensor is diagonal. The respective amplitudes $c_{(a_1 e_1), (a_2 e_2)}^{(L)}$ of the functions $g^{(L)}(t)$ can be calculated for given combinations of excitation and emission polarizations²⁹⁹ and are given in Supplementary Table 3.16.

To account for the normalization of the experimental FCS curves, eq. 3.58 can be rewritten in the form:

$$\hat{G}_{\text{rot}}(t_c) = G_{\text{rot}}(t_c)/c_p^{(0)} = 1 + R_p \left(x_p g^{(2)}(t_c) + (1 - x_p) g^{(4)}(t_c) \right), \quad (3.60)$$

where:

$$R_p = \frac{c_p^{(2)} + c_p^{(4)}}{c_p^{(0)}}; \quad x_p = \frac{c_p^{(2)}}{c_p^{(2)} + c_p^{(4)}}; \quad p = (a_1 e_1), (a_2 e_2). \quad (3.61)$$

Alternatively, the parameter x_p can be replaced by the parameter C_p as used in eq. 3.31 in the main text defined as:

$$C_p = \frac{1 - x_p}{x_p} = \frac{c_p^{(4)}}{c_p^{(2)}}, \quad (3.62)$$

From which the normalized rotation correlation function, $\hat{G}_{\text{rot}}(t_c)$, is obtained as:

$$\hat{G}_{\text{rot}}(t_c) = 1 + R_p \left(\frac{1}{1 + C_p} g^{(2)}(t_c) + \frac{C_p}{1 + C_p} g^{(4)}(t_c) \right), \quad (3.63)$$

The values for the parameters R_p , x_p , and C_p are given in in Supplementary Table 3.17.

Chapter 4. Fluorescent dyes probe the local environment of protein T4 Lysozyme: Exploring with experiments and simulations

Authors: Aiswaria Prakash, Christian A. Hanke, Katherina Hemmen, Laura Stephan, Claus A.M. Seidel†

Affiliations:

Chair for Molecular Physical Chemistry, Heinrich Heine University, Düsseldorf, 40225 Düsseldorf, Germany

†Correspondence to: c.seidel@hhu.de

Contributions: A.P. and C.A.M.S. designed the research. K.H. and C.A.M.S. supervised the project. K.H., A.P. and L.S. performed the biochemistry methods. K.H., A.P. and L.S. performed the time-resolved measurements. A.P. and L.S. performed the steady state spectral measurements. L.S. performed the temperature-dependent measurements. A.P. performed the FCS measurements. K.H. performed the PIE-FRET measurements and the substrate binding studies. C.A.H. performed and analyzed MD simulations, identified and quantified potential interacting modes. A.P. performed AV modeling of dye distributions and AV-related calculations. A.P. designed all figures. A.P. wrote the manuscript in consultation with C.A.H. and C.A.M.S.

4.1 Introduction

Since the large diversity of functions in biomolecules is a direct manifestation of the chemical structure and dynamic behavior of the biomolecular system, their structure determination is crucial. There are several biomolecular research tools like X-ray crystallography³⁰⁰⁻³⁰², cryo-electron microscopy^{303,304} and nuclear magnetic resonance^{305,306} that are conventionally employed to probe the conformational structure and dynamics of biomolecules and report on their biomolecular interactions. Another important technique to study biomolecular structural dynamics is Förster resonance energy transfer (FRET), which is particularly useful for large and dynamic biomolecules and complexes. The FRET technique utilizes fluorescent dye labels, strategically placed on the biomolecule, and determines their inter-dye distance by measuring the relative changes in fluorescent properties^{15,159,187,265,307-309}. A thorough understanding of the fluorophore photophysics is essential for the choice and design of suitable fluorescent probes and the correct interpretation of the dye behavior. The local environment in the biomolecule has a profound effect on the fluorescence properties of the dye. A notable example is protein-induced fluorescence enhancement (PIFE). Here an increase in fluorescence intensity occurs when a protein binds to a nucleic acid in the proximity of a fluorescent probe like Cy3 or Cy5^{135,136,220}. In certain cases, fluorescence quenching may occur through interactions of amino acids like Tryptophan and organic fluorophores like rhodamine or oxazine due to photoinduced electron transfer (PET)^{104,105}. For example, when FRET is used to measure the distance between two regions of a protein, the presence of quenchers can cause systematic deviations in recovered distances²⁷⁰. Therefore, one must be cautious while interpreting changes in spectroscopic properties of the dye since they may originate from features of the local environment of the dye.

The influence of the local environment on the fluorophore properties can however be used to our advantage to report on the local biomolecular environment. Depending on the fluorophore, the protein and the solvent conditions, dye-protein interactions can occur in a variety of ways. The different modes of interactions or binding mechanisms include π -stacking³¹⁰⁻³¹², electrostatic interactions³¹³ and hydrophobic interactions³¹⁴. Interactions with a protein can have significant consequences on the dye behavior resulting in

fluorescence quenching¹⁰⁵, fluorescence enhancement^{135,136}, spectral shifts^{114,315}, and anisotropy changes^{101,102,316,317}.

In this study, we use the bacteriophage T4 Lysozyme (T4L) as a prototype system for exploring the influence of the local dye environment in proteins. It is a modular protein consisting of two subdomains, the N-terminal subdomain and the C-terminal subdomain linked by a long α -helix³¹⁸ (Helix C, Fig. 4.1A). Additionally, there is a highly flexible domain, which acts as a connector and integral to the structure of T4L (Helix A, grey Fig. 4.1A,C). T4L has been extensively used as a model protein in protein unfolding^{319,320} and ligand binding studies³²¹⁻³²³. It is very tolerant of mutations, even accommodating insertions and deletions³²⁴. Mutations introduced for the covalent labeling of organic fluorophores onto specific sites is therefore expected to cause little perturbations to the protein conformation. The function of T4L is the hydrolysis of the peptidoglycans of bacterial cell walls, this cleavage involving a hinge-bending motion between open and closed conformations. An even more compact minor state, referred to as the C₃ state, was discovered by Sanabria et.al.³ The equilibrium fractions of the 3 conformational states were found to be 54%(C₁), 30%(C₂) and 16%(C₃) of the wild type T4L. From the 500+ structural models of T4L available within the Protein Data Bank (PDB), we chose PDB 172L to represent the open (C₁) structure and 148L to represent the closed (C₂) structure of T4L^{3,267} (Fig. 4.1B).

The main objective of this study is to investigate the effect of fluorescent properties of coupled dyes in different protein environments and subsequently use the dye as a probe to sense the environment. The commonly used fluorophores, Alexa488 and Alexa647, were chosen and covalently coupled to chosen labeling sites in the protein via long flexible linkers. The experimental output from fluorescence spectroscopy is combined with computational simulations to give a detailed insight into the origin, type and extent of dye-protein interactions under various environments. The amino acid residues responsible for dye trapping and quenching, the effect of dye trapping on fluorescence quenching, the extent of static quenching and H-bonding interactions and the interference to dye motion from protein dynamics are presented. Finally, specific examples are shown where the coupled dyes act as sensors and report on changes in the local and global protein environment.

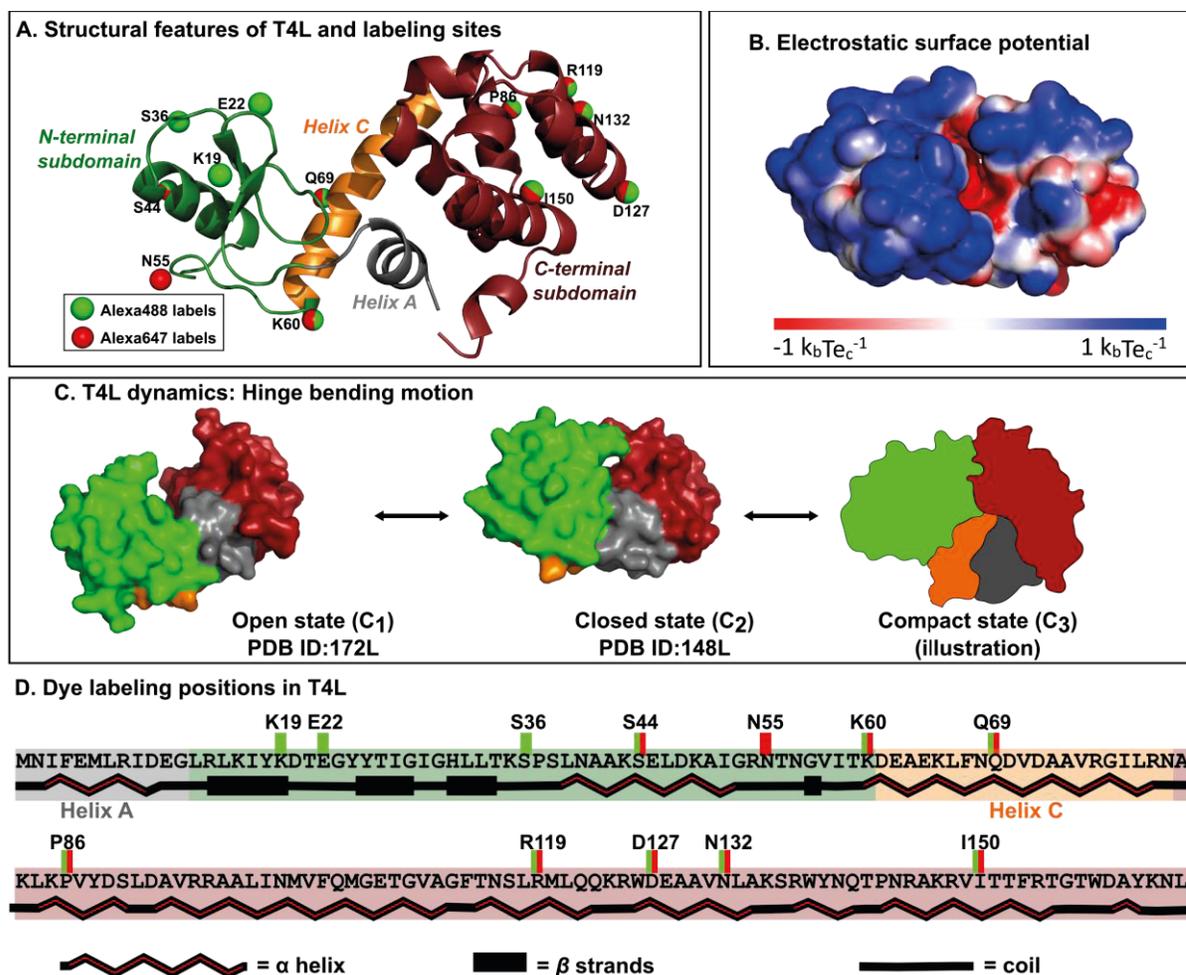


Figure 4.1. Structural details of protein and dye labels used in the study. (A) The T4 Lysozyme (T4L) protein comprises of a rigid C-terminal subdomain (dark red), an N-terminal subdomain (green) connected by a 21 amino acid long long α -helix (Helix C, orange) and a connector helix (Helix A, grey). The dye labeling sites are indicated as spheres (red: Alexa647 labels, green: Alexa488 labels). (B) The van der Waals surface of T4L (PDB: 1Lw9) color coded with the electrostatic surface potential (ESP, calculated by APBS³²⁵⁻³²⁷). The electrostatic potential is given here in dimensionless units of $k_b T e_c^{-1}$ where k_b , the Boltzmann's constant = $1.3806504 \times 10^{-23} \text{ J K}^{-1}$, T is the temperature in K and e_c is the charge of an electron, $e = 1.60217646 \times 10^{-19} \text{ C}$. (C) The function of T4L is realized by its hinge bending motion from open (C_1 , PDB: 172L) to closed (C_2 , PDB: 148L) to compact (C_3) conformations. (D) The amino acid sequence of T4L along with the labeling positions of Alexa488 (green) and Alexa647 (red) used in this study. The colored blocks behind the sequence indicate the corresponding region in (A). The secondary structural elements like helix, β -sheets and coils are also indicated as described by the legend.

4.2 Results and Discussion

In order to investigate the effects of the T4L protein environment on the photophysics and local motions of the coupled dyes Alexa488 and Alexa647, we use labeling sites span

over all subdomains of T4L (Fig. 4.1D). Subsequently, we combine fluorescence techniques and computational simulations to gain information on potential interactions between the protein and the dye. From the insights gathered from these studies, we could use the fluorophores as sensors and identify the characteristics of unknown local biomolecular environments or global events.

4.2.1 Alexa488 labels exhibit environment-sensitive photophysical properties

(The T4L variants P86C, R119C, D127C, N132C and I150C were prepared and measured by A.P., with help from L.S. The variants E22pAcF, S36pAcF, Q69pAcF, K19pAcF, S44pAcF, K60pAcF were prepared and measured by K.H., to the best of my knowledge)

In order to study the environment-dependent behavior of Alexa488 dye, ensemble time resolved fluorescence lifetime and anisotropy measurements of Alexa488 labels on 11 sites were studied (Fig. 4.1D, Supplementary Fig. 4.1). The original residue at each of these sites have been mutated to either cysteine (C) or pAcPhe (pAcF) to accommodate the dye labeling chemistry, thus leading to 11 T4L mutants. Hereafter, I refer to each labeled T4L variant by its mutation site which is also the labeling site (For example: P86C, S44pAcF etc.). Fig. 4.2A shows the spread of the species-weighted fluorescence lifetimes of Alexa488 labels over a time range $\langle\tau\rangle_x = 2.3\text{-}4.1$ ns and that of residual anisotropies over a wide range 0.15-0.24. For the free Alexa488 dye, the fluorescence is single exponential with a lifetime of 4.1 ns, while the fluorescence of the coupled labels is multiexponential, indicating interactions between the dye and the protein. Fluorescence decay curves of highly quenched (N132) and relatively less quenched (P86) Alexa488 labels are shown in Fig. 4.2B(top). The multiexponential fluorescence decays are characterized by two or three lifetime components (Table 4.1, Supplementary table 2). Among the labels studied, the longest lifetime was exhibited by S44, with fluorescence lifetime $\langle\tau\rangle_x = 4.1$ ns, similar to that of the unquenched dye; while E22 had the shortest lifetime $\langle\tau\rangle_x = 2.4$ ns. The anisotropy decay curves of the Alexa488 labels, N132 and P86, are shown in Fig. 4.2B(bottom). The decays are characterized by three rotational correlation times (Supplementary table 2), of which the longest correlation time describes the global rotation of the protein and the corresponding amplitude is a measure of the

rotational freedom of the molecule and is called the residual anisotropy (r_∞). The highest and lowest residual anisotropies were exhibited by K19 ($r_\infty = 0.25$) and R119 ($r_\infty = 0.15$), respectively. In presence of only transient interactions, we expect an inverse correlation between lifetime and residual anisotropy because when the lifetime of the dye is lowered due to quenching, the increased dye-quencher interactions is presumed to raise the anisotropy. However, this expected correlation is not observed for the studied Alexa488 labels. This indicates that there are specific environment-dependent interactions affecting the photophysical properties of these labels. The elements in the local protein environment affecting the fluorescence properties of the dye labels are investigated in the following sections.

Table 4.1. Photophysical properties of representative Alexa488 labels.

| Label position | x_1 | τ_1 , ns | x_2 | τ_2 , ns | x_3 | τ_3 , ns | $\langle\tau\rangle_x$, ns | r_∞ | Φ_F | $\Phi_{F,dy}$ n | f_{static} | λ_{max} (abs) nm | A_Q |
|--------------------------|-------|---------------|-------|---------------|-------|---------------|-----------------------------|------------|----------|--------------------|--------------|--------------------------------|-------|
| P86C | 0.81 | 4.22 | 0.13 | 2.45 | 0.06 | 0.31 | 3.76 | 0.18 | 0.84 | 0.89 | 0.06 | 496 | 0.06 |
| R119C | 0.65 | 4.04 | 0.24 | 1.90 | 0.11 | 0.45 | 3.14 | 0.15 | 0.68 | 0.75 | 0.08 | 495 | 0.13 |
| D127C | 0.71 | 4.13 | 0.20 | 2.35 | 0.08 | 0.38 | 3.46 | 0.19 | 0.72 | 0.82 | 0.12 | 496 | 0.13 |
| N132C | 0.62 | 3.99 | 0.29 | 2.07 | 0.09 | 0.43 | 3.10 | 0.22 | 0.68 | 0.74 | 0.09 | 496 | 0.11 |
| I150C | 0.70 | 4.04 | 0.23 | 2.40 | 0.07 | 0.40 | 3.40 | 0.20 | 0.75 | 0.81 | 0.08 | 496 | 0.15 |
| Alexa488 free dye | | | | | | | 4.05 | 0.02 | 0.96 | - | - | 491 | 0.00 |

The species-weighted average lifetime, $\langle\tau\rangle_x$, calculated using eq. 4.5 and residual anisotropy, r_∞ , are obtained from fluorescence and anisotropy decay analysis. The steady state fluorescence quantum yield (Φ_F), the dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) and the fraction of statically quenched species (f_{static}) are calculated using eq. 4.9-11. $\lambda_{max}(abs)$ is the wavelength at absorption maxima. A_Q is the amplitude of the static quenching process with a relaxation time 68 ns obtained on global fitting of FCS curves measured at 485 nm and 5.6 kW/cm² power to the model function given by eq. 4.15.

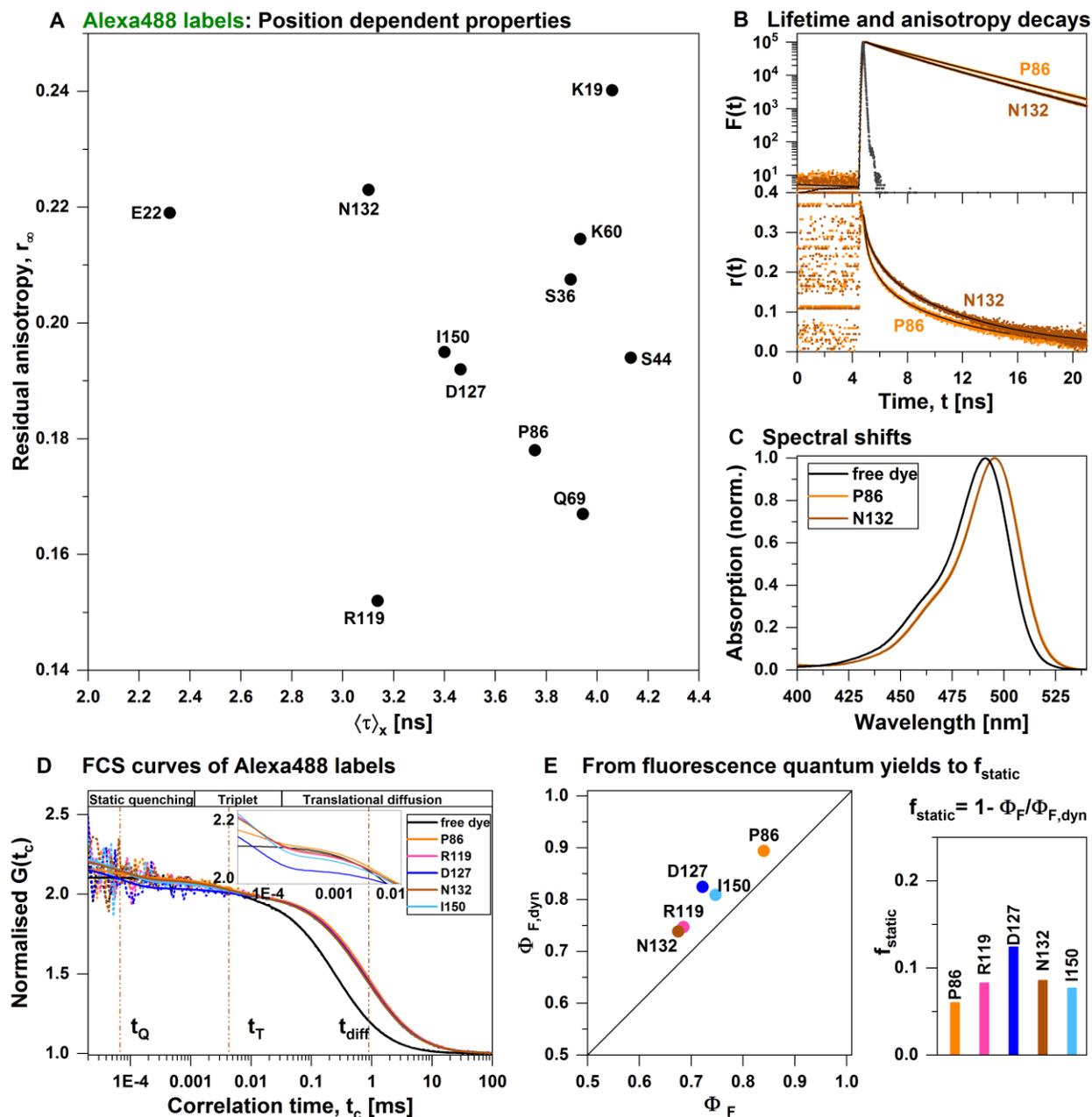


Figure 4.2. Photophysics of Alexa488 depends on local protein environment. (A) The species-weighted average fluorescence lifetime $\langle \tau \rangle_x$ and residual anisotropies exhibit a wide range of values due to difference in dye environments. An inverse correlation between these two observables can be seen only for 6 of the 11 labels, indicating specific interactions for each label. (B) Exemplary fluorescence and anisotropy decays of Alexa488 labels, P86 (orange) and N132 (brown), representative of less quenched and highly quenched labels, with the instrument response function (grey) and their corresponding fits (black lines). (C) Absorption spectra of Alexa488 free dye and representative labels, P86 and N132, show the red shift of labels compared to the free dye. (D) FCS curves (parallel-perpendicular correlations) of selected Alexa488 labels (dotted lines) and their fits (solid lines) measured at 485 nm and at an excitation power of 5.6 kW/cm². The vertical lines indicate the fitted relaxation times for the translational

diffusion (t_{diff}), triplet (t_r) and quenching (t_Q) for the N132 label. Inset: Zoom of static quenching regime showing fitted curves. (E) Steady state fluorescent quantum yield (Φ_F) and dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) are plotted for the selected Alexa488 labels. The black line shows a 1:1 relationship. The difference between the two values arises due to static quenching. The fraction of static quenching can be calculated using the displayed formula and is seen to be approx. 10% for the Alexa488 labels.

A combination of spectroscopic methods sheds light on Alexa488 fluorescence quenching. One of the major factors affecting the photophysical behavior of Alexa488 dye is PET quenching¹⁰⁸ where electron transfer occurs from an electron donor with suitable oxidation potential to the chromophore. The amino acids Trp, Tyr, Met and His can serve as effective electron donors for Alexa488¹⁰⁸ and other Rhodamine-based dyes¹⁰⁴. The mechanism of PET by amino acid residues involves a combination of both static and dynamic quenching pathways. While dynamic quenching, which occurs by collisional encounters, affects both fluorescence lifetime and intensity, static quenching reduces the intensity without influencing the lifetime^{104,105,108}.

Static quenching occurs by ground-state complex formation between the dye and a quencher. This results in the perturbation of the ground state and therefore in the red shift of the dye absorption¹¹⁴. The absorption spectra can thus act as indicators for dye-protein interactions, if such interactions lead to static quenching. In the present study, we observe a 5 nm shift of absorption spectra for the coupled dyes compared to the free dye, indicating potential stacking interactions (Fig. 4.2C). Since polarity differences and presence of H and J aggregates also result in absorption shifts, we cannot make conclusive statements about the quenching pathway from these spectral shifts alone. In addition, there is no significant difference in the absorption maxima between the measured coupled dyes, which further hampers our efforts to extract information about individual dye environments from the absorption spectra.

In order to quantify the magnitude of quenching, selected T4L labels and the free dye were probed by FCS measurements performed at a constant power density of 5.6 kW/cm² at 485 nm. A fluorescence intensity fluctuation at 68 ns was observed to varying degrees for the T4L labels, but was absent for the free dye. Since this is a protein environment-dependent observation, and amino acid quenching of Alexa488 is expected, this fluctuation is assigned to PET dye quenching. The fits of the FCS curves of R119, D127, N132 and I150 (Fig. 4.2D, Table 1, Supplementary table 3) show a fluorescence intensity

fluctuation with amplitude $Q > 11\%$ at a correlation time $t_Q = 68$ ns, while the amplitude of P86 is lower at $Q=6\%$. This lower quenching amplitude for P86 agrees with its longer lifetime from the time resolved measurements. It must be remembered that the FCS amplitude may be an average of more than 1 process occurring in the same time range, and does not necessarily reflect a single quenching mechanism.

To unravel the quenching pathway, we measured the fluorescence quantum yields for the free dye and selected T4L labels (Fig. 4.2E). We measure two types of fluorescence quantum yields : (i) The dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) measures the dynamic/collisional quenching and is proportional to the average lifetime $\langle \tau \rangle_x$ as given by eq. (4.10). (ii) The steady state fluorescence quantum yield (Φ_F) measures the reduction in fluorescence intensity relative to a standard reference and therefore accounts for both static and dynamic quenching processes.

From Fig. 4.2E, it is clear that $\Phi_{F,dyn}$ is consistently higher than Φ_F . This difference is due to static quenching which is not reflected in $\Phi_{F,dyn}$, but reduces Φ_F . From these quantities, the fraction of statically quenched species (f_{static}) can be calculated as given in eq. 4.14. For the measured Alexa488 labels, the f_{static} is between 6-12%, P86 being the least (6%) and D127 being the most (12%) statically quenched. From previous studies of Alexa488 labels on RNA (Chapter 3), f_{static} up to 41% were observed. Compared to that, the loss of fluorescence due to static quenching is considerably lower for the T4L protein labels. Therefore, we conclude that dynamic quenching is the major quenching pathway for the T4L-coupled Alexa488 labels. In proteins, the aromatic PET quenchers are usually found within hydrophobic pockets where dye accessibility is low; in contrast to the exposed guanine quencher at the nucleic acid helix ends (Chapter 3). Therefore, effective dye-quencher stacking, necessary for ground state complex formation in static quenching, is less probable in proteins compared to nucleic acids with dyes labeled close to the Guanine terminus.

4.2.2 Simulation studies predict probable dye-protein interactions

In order to investigate the behavior of Alexa488 dyes on the T4L protein two independent simulation approaches were applied: Molecular Dynamics (MD) simulations and accessible volume (AV) calculations. The MD simulations yield atomistic representations

of the dye attached to the protein while the AV simulations provide sterically allowed dye positions while assuming all positions to be equally probable^{103,242}. Since conformational sampling in MD simulations is time-consuming process, it was performed only on selected labels (P86, R119, D127, N132 and I150). For each labeling position, we performed 10 simulations of 3000 ns simulation time each. The more coarse-grained AV simulations, on the other hand, are based on simple geometric calculations, and were carried out for all measured labels.

a) MD simulations allow a closer look at interaction modes

(MD simulations and their analysis were carried out by Dr. Christian Hanke)

In order to investigate the sampled conformational space in the MD simulations, the positions of a selected atom in the dye (the xanthene oxygen atom, Supplementary Fig. 4.3B) over the course of the simulation is shown as spheres in Fig.4.3A and Supplementary Fig. 4.2. A high density of spheres at regions near the T4L surface suggest interactions and indicate potential dye trapping, whereas a more uniformly distribution of spheres would correspond to a less trapped label (e.g. P86). For a better understanding of dye interactions, we computed the dye-amino acid distance over the simulation time for several amino acids (Supplementary Fig. 4.3). From these traces, stable and long-lived interactions can be inferred as distinct plateaus with distances ~ 10 Å. Some of the frequently occurring interaction modes are presented in Fig. 4.3B and discussed below.

π -stacking interactions of dye with PET quenchers: In order to investigate potential π -stacking, we follow Vaiana et. al¹⁰⁶, who describe a combination of distances and angles between the dye and quencher by which π -interactions lead to effective PET quenching. The distance (r) is defined between the centers of dye and quencher ring systems and the angle (φ) is defined between the normal vectors of the planes of these ring systems (Supplementary Fig.4.4). These interaction parameters are calculated for the MD trajectories (Fig. 4.3B(1), right). In our simulations, the distance (r) was generally greater than 5.5 Å where stacking interactions were less probable. However, there were a few instances of suitable r and φ combinations that could result in quenching interactions due to effective π -overlap (Fig. 4.3B(1)). Since the quenching amino acid residues, Trp and

Tyr, are hydrophobic, they are usually buried within hydrophobic protein pockets and are thus inaccessible to the dyes. Therefore, strong π - π interactions are only occasionally observed in our simulations.

H-bonding interactions of dye with positively charged and other amino acids:

Among the amino acid residues on T4L, Lysine and Arginine are the most positively charged, due to protonation of their NH_2 and guanidine groups, respectively, at physiological pH used in the experiments. They have high aqueous pK_a 's: ~ 10.5 for Lys and ~ 13.8 for Arg³²⁸. These residues are responsible for the positively charged regions on the electrostatic potential map on the T4L surface (Supplementary Fig. 4.5). The positively charged Lys and Arg potentially participate in H-bonding with the negatively charged SO_3^- groups of the dye labels (Fig.4.1B). This type of electrostatic interaction is the most frequent of the interaction modes observed between the coupled dye and the protein in the simulations. The higher prevalence of Lys and Arg in T4L and their solvent-exposed natures contribute to the predominance of this interaction mode. For each of the 10 simulations performed for a given label, the percentage of simulation time where H-bonding was observed was calculated and displayed as box plots (Fig. 4.3B(2)); the criteria for H-bonding adopted from Raschka et al.³²⁹. It can be seen that N132 has consistently high H-bonding in all the simulations with the highest average among the labels measured; while the lowest average percentage of H bonding was exhibited by R119. This ranking is also reflected in their respective residual anisotropies. The positive charges in the amino acid side chains trap the negatively charged dye, thereby generating a "contact volume", which is an area of defined thickness close to the protein surface favored by the dye²⁴² where diffusion is lower. It is assumed that when trapped to Lys or Arg, the dye remains for a long time in the neighborhood of PET quenching amino acids, leading to quenching events which otherwise would not have been as probable. In addition to the positively charged Lys and Arg, neutral (non-charged) and polar amino acids like Asn and Gln are also seen to participate in H-bonding (Supplementary Fig. 4.6A).

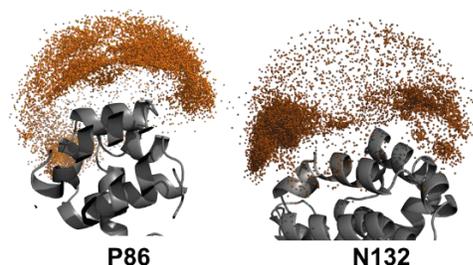
A less frequently observed interaction partner for the protein in the simulations is the dye-linker. The long flexible linker connecting the dye to the protein labeling site contains functional groups which can form H-bonding interactions with suitable amino acid

residues (Supplementary Fig. 4.6B). This particular mode of interaction is usually observed together with other dye-protein interaction modes.

Concerted interactions: Even though π -stacking interactions between the dye and Trp/Tyr are relatively rare in our simulations, H-bonding interactions between the oxygen of the SO^3 group in the dye label and the hydrogen from the NH and OH functional groups of Trp or Tyr are more frequently detected. Consistently, these dye interactions are reinforced due to the presence of another dye-protein bond in its proximity (Fig. 4.3B(3)). Such concerted interactions can be detected in the dye-quencher distance trace, where most instances of quencher-dye proximity are accompanied by dye proximity to a positively charged amino acid (Fig. 4.3B(3)).

Cation- π interactions: A common non-covalent interaction existing between the amino acids in a protein is the cation- π interaction³³⁰⁻³³². This is usually observed between the cationic side chain of Lys/Arg and the aromatic sidechain of Trp, Tyr and Phe within a distance of 6 Å. The fluorophore Alexa488 also possesses delocalized π electrons in its aromatic rings which is within the 6 Å distance from the cationic centres of Lys and Arg residues. Therefore, cation- π interactions are highly likely between Lys/Arg and the dye. This interaction mode has been detected in our MD simulations (Fig. 4.3B(4)). A screening of the MD trajectories to identify potential cation- π interactions was carried out using a geometric criterion ≤ 6 Å between the ammonium nitrogen (NZ) in Lys or the guanidinium carbon (CZ) in Arg and the midpoints of any of the 3 rings in the xanthene system of Alexa488. This screening provides us with an upper limit of the percentage of simulation time where cation- π interactions are possible. (Fig. 4.3B(4), box plot). Clearly, this percentage will be an overestimation of actual cation- π interactions since we consider only the geometric criteria, and not all cation- π contacts represent energetically favorable interactions. It can be seen that on average, N132 has the highest percentage of interactions, comparable to the % of H-bonding. The prevalence of H-bonding and cation- π interactions over the simulation period are remarkably high, emphasizing the role of Lysine and Arginine in affecting dye behavior.

A. Dye distributions from MD sim



B. Types of Alexa488-T4L interactions

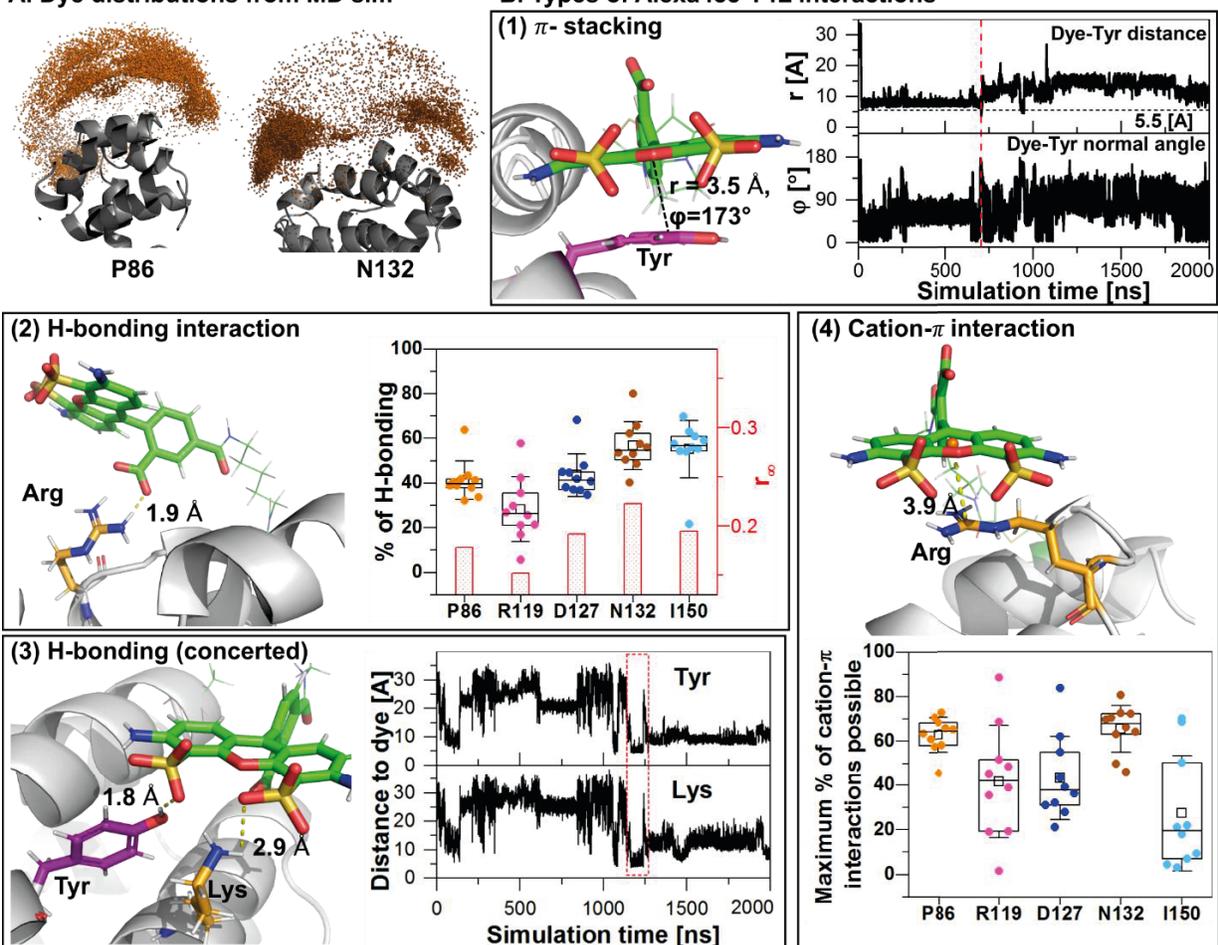


Figure 4.3. Molecular dynamics simulations probe Alexa488-T4L interactions. (A) Distributions of dye positions sampled during MD simulations for representative Alexa488 labels, P86 (orange) and N132 (brown). Spheres represent the location of the O3 atom for every 2 ns of the simulation time. (B) T4L-dye interaction modes observed during MD simulations. (1) π -stacking: Stacking interaction between the electron rich π -systems of the dye (green) and the quencher Tyr (magenta). Interaction parameters¹⁰⁶, r , the distance between the centres of the ring systems and ϕ , the angle between the normals to the plane of the ring systems for a part of the trajectory is displayed with a distance threshold of 5.5 Å (dashed black line). The part of the trajectory corresponding to the picture is indicated by the red dashed line. (2) H-bonding: The negatively charged groups on the dye and the NH groups of Arg undergo H-bonding. The box plot shows the percentage of simulation time where H-bonding is observed, each simulation displayed as a circle. (3) Concerted H-bonding: Multiple simultaneous H-bonding events can take place for the dye with different residues like Tyr and Lys (snapshot). This is detectable in the dye-amino acid distance plots (dashed red box). (4) Cation- π : The cationic side chain of Arg interacts with the aromatic π system of the dye. The box plot shows the percentage of simulation time where the interacting groups are within 6 Å, allowing potential cation- π interactions. For both box plots, each simulation is displayed as a circle, the hinges of the boxes show the 25th and 75th percentile, the horizontal lines inside boxes shows the median, the mean is displayed as a square inside a box and the whiskers show a standard deviation of 1.

b) AV simulations predict changes in fluorescence parameters

Because in all-atom MD simulation, sampling is limited, we use coarse grained AV simulations. In the current study, the accessible volume (AV) approach to calculate proximity/accessibility to various suspected interaction sites and correlate it to a fluorescence observable like the fluorescence lifetime or residual anisotropy. Since the protein T4L adopts both open (PDB: 172L) and closed (PDB: 148L) conformations, for accessibility calculations, a weighted average of the calculated quantity for the C₁ and C₂ structures were employed. Since the structure of C₃ is still under study, the equilibrium species fraction determined for the C₃ state, was instead assigned to the C₂ state for the purpose of this study. Thus AV-based calculations were weighted 54% for the C₁ state and 46% for the C₂ state. The following factors affecting Alexa488 behavior were studied:

Interaction with PET quenchers: PET quencher accessibility is calculated as the fraction of the interaction volume to the AV volume; where the interaction volume (Fig. 4.4A, inset) is defined as the volume of AV that is $\leq 4 \text{ \AA}$ from the quenchers, Trp, Tyr, Met and His. Since this accessibility as an estimate of quencher proximity and thereby, the concentration, we can study the effect of quenchers on the species-weighted average fluorescence lifetime (Fig.4A).

$$\text{PET quencher accessibility} = 0.54 \cdot \left[\frac{\text{Interaction volume}}{\text{Total AV volume}} \right]_{C_1} + 0.46 \cdot \left[\frac{\text{Interaction volume}}{\text{Total AV volume}} \right]_{C_2}$$

From Fig 4.4A, it is seen that unquenched labels have negligible quencher accessibility. A clear correlation between the calculated quencher accessibility and fluorescence lifetime is not observed, indicating that presence of PET quenchers is not the only factor determining the fluorescence behavior.

Interference due to protein dynamics: During the hinge-bending motion of T4L, it is possible for a dye to get trapped if the dye is labeled close to the cleft region. In order to investigate if such a trapping is detrimental to the fluorescence of a dye, we calculate the difference in AVs between the C₁ and C₂ states and correlate it to the residual anisotropy (Fig. 4.4B). Among the measured labels, E22, K60 and N132 exhibit the largest changes in AV, going from open to closed state, indicating that these labels are the most affected by the hinge bending motion. Interestingly, these labels also exhibit high residual

anisotropies. This correlation implies that during T4L hinge bending motion, some dyes may be trapped or “squeezed” by the protein depending on its labeling site and orientation.

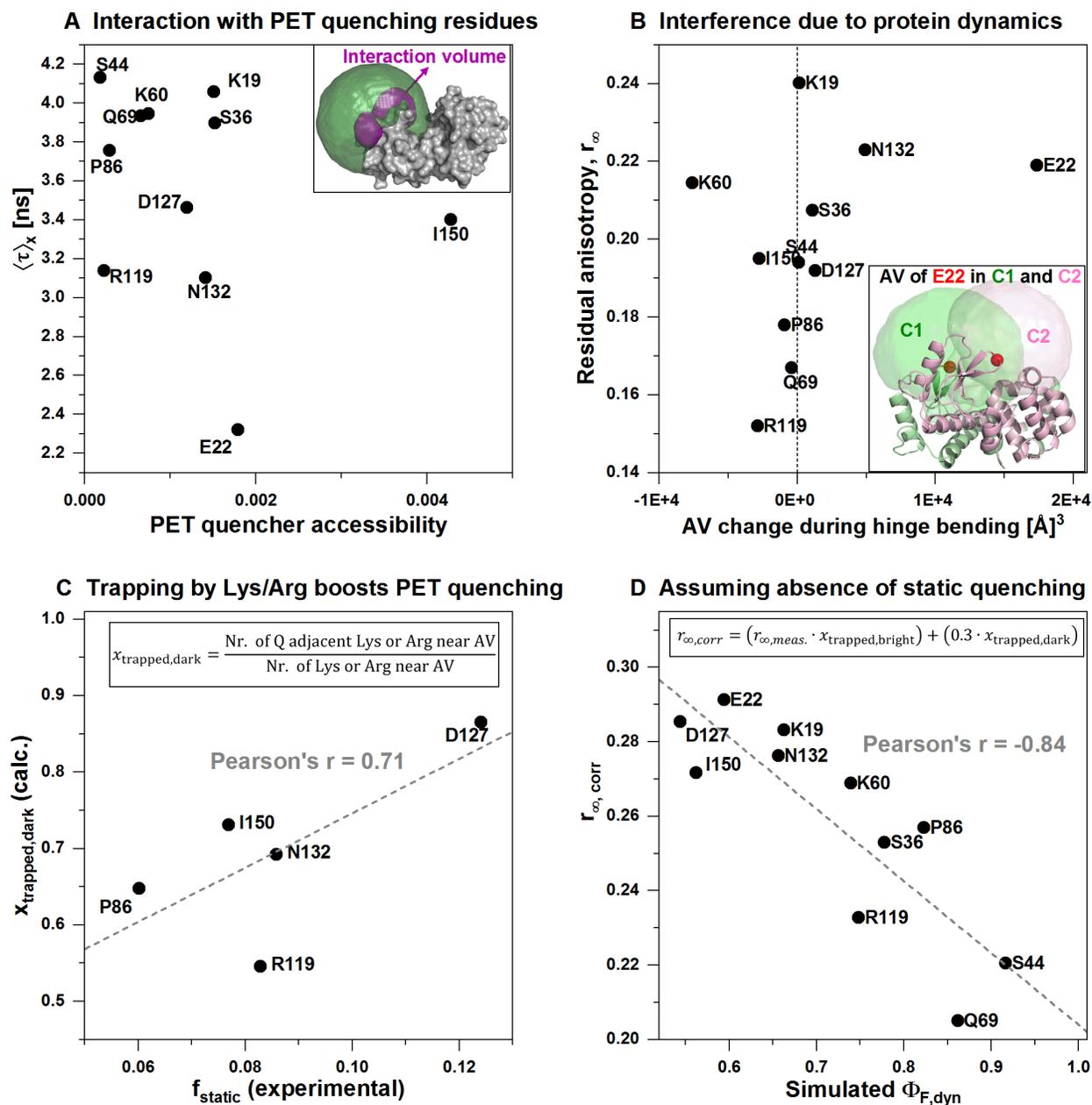


Figure 4.4. Predicting factors affecting Alexa488 behavior. (A) PET quenchers: The predicted accessibilities of the PET quenchers, Trp, Tyr, Met and His, have little correlation with the species-weighted average fluorescence lifetimes of the Alexa488 dyes. Inset shows the interaction volume of an exemplary AV. (B) Hinge bending motion of T4L: Change in AV between C₁ and C₂ states plotted against residual anisotropy show that the sites where dye labels are constrained due to protein dynamics have relatively high residual anisotropy. Inset: Example of an AV at labeling position E22 (red spheres) in C₁ (green) and C₂ (pink) states. (C)

Dye trapping: The fraction of statically quenched species (f_{static}) from experiment is well correlated to the fraction of trapped protein surface that is dark ($f_{\text{trapped,dark}}$) from calculations. This is due to the fluorescence of Lys/Arg-trapped dyes being quenched by adjacent PET quenchers. (D) Static PET: Assuming a hypothetical case of no static quenching, simulated dynamic fluorescence quantum yields is plotted as a function of the corrected residual anisotropies (equation displayed), yielding the inverse correlation expected in the absence of static quenching.

Interaction with Lys/Arg: From our MD studies, we could ascertain that in addition to the PET quenchers, positively charged amino acids, Lys/Arg, also interact with the dye for a considerable length of time. In the vicinity of these Lys/Arg residues, PET quenchers are also present. It is therefore reasonable to assume that there is a fraction of dye species which undergo PET quenching by Trp/Tyr while being trapped by Lys/Arg. To estimate the extent of trapped positions that are dark (statically quenched), we determine the number of trapped positions (that is, Lys or Arg residues) that are within 4 Å of the dye AV and find out how many of these residues have PET quenchers within 4 Å, since these would render the dye dark (more details in Supplementary note 4.1). Then, the species fraction of trapped dark surface is:

$$x_{\text{trapped,dark}} = \frac{\text{Number of quencher – adjacent Lys/Arg near AV}}{\text{Number of Lys/Arg near AV}} \quad (4.1)$$

This fraction of trapped surface that is dark from the AV calculations correlates reasonably well ($\rho = 0.71$) with the experimental fraction of statically quenched species (f_{static}), calculated from the fluorescence quantum yields (Fig. 4.4C); confirming that static quenching occurs when the dye is trapped near quencher-rich areas. The validity of this assumption was further tested by incorporating this fraction into fluorescence parameters obtained experimentally as described below.

It is known that residual anisotropy is a measure of dye restriction¹⁰³. However, the measured r_{∞} does not reflect the trapping at the dark surface, since static quenching removes the quenched molecules from observation and we observe only the bright molecules. We imagine a hypothetical situation where all quenching is dynamic and calculate the residual anisotropy in that situation. Due to the additive property of anisotropy, $r = \sum r_i \cdot f_i$. where r_i and f_i are the anisotropy and fluorescence fractions for the i^{th} species. In our case, there are 2 species: bright and dark. Therefore,

$$r = (r_{trapped,bright} \cdot f_{trapped,bright}) + (r_{trapped,dark} \cdot f_{trapped,dark}) \quad (4.2)$$

The Fluorescence fraction, $f_i = \frac{\langle \Phi_F \rangle_1 \cdot x_i}{\sum x_i \cdot \langle \Phi_F \rangle_i}$.

We approximate $\langle \Phi_F \rangle_1 \approx \langle \Phi_F \rangle_2$, since in this hypothetical case, all trapped fractions have the same brightness. Then we get $f_i \approx x_i$.

Thus, we can use the species fractions (x_i) instead of the fluorescence fractions (f_i) in eq (4.2)

$$r = (r_{trapped,bright} \cdot x_{trapped,bright}) + (r_{trapped,dark} \cdot x_{trapped,dark}) \quad (4.3)$$

Thus we can determine the corrected residual anisotropy, $r_{\infty,corr}$, as the sum of the measured r_{∞} (bright) and r_{∞} of dark molecules weighted by their respective fractions on the trapped surface. If we assume that the r_{∞} of dark molecules is 0.3 due to trapping, the corrected r_{∞} is:

$$r_{\infty,corr} = (r_{\infty,meas.} \cdot x_{trapped,bright}) + (0.3 \cdot x_{trapped,dark}) \quad (4.4)$$

$$\text{where } x_{trapped,bright} = 1 - x_{trapped,dark} \quad (4.5)$$

The corrected residual anisotropy accounts for the hypothetical case where all interactions are collisional (dynamic) and all static effects are absent. A previous study from Peulen et al.²⁷⁰ simulated the diffusion and dynamic quenching of Alexa488 dyes on different proteins by using a combination of AV and Brownian dynamics (BD) simulations. The resulting fluorescence quantum yield is purely dynamic in nature. We adopt this approach to obtain the simulated dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) for the T4L labels under study. The simulated and measured $\Phi_{F,dyn}$ are not equal and simulated $\Phi_{F,dyn} \leq$ measured $\Phi_{F,dyn}$ (Supplementary Fig.4.7). A good inverse correlation is observed when the simulated $\Phi_{F,dyn}$ is plotted as a function of $r_{\infty,corr}$ (Fig. 4.4D) with a correlation coefficient $r=-0.84$; thus validating our proposition of static quenching of dyes trapped at Lys/Arg residues with nearby quenchers.

In conclusion, we can summarize that the different environmental factors that alter the photophysical properties of Alexa488 dyes on T4L. The major quenching mechanism is PET which occur by a combination of dynamic and static quenching pathways. These quenching effects are boosted when the dye is trapped to a nearby positively charged

amino acid like Lys or Arg. The hinge bending motion of T4L also present spatial constraints to dyes labeled in its proximity. It may be worth noting that some of the observed correlations are expected to improve with the adoption of a C₃ structure in the future.

4.2.3 The photophysics of coupled-Alexa647 dyes are influenced by the T4L local environment

(The T4L variants S44pAcF and K60pAcF were prepared and measured by A.P. The variants S44C, N55C, Q69C, P86C, R119C, N132C, I150C and D127C were prepared and measured by K.H., to the best of my knowledge.)

In order to investigate the photophysical properties of a class of dyes quite different from Alexa488, we chose the cyanine dye, Alexa647. They are often used as a donor-acceptor dye pair in FRET experiments^{3,238,270}. Alexa647 is expected to exhibit association-induced fluorescence enhancement (AIFE) similar to the Cy5 dyes (Chapter 3). The fluorescence behavior of Alexa647 on 9 distinct labeling sites on T4L were investigated (Fig.4.5A, Supplementary Table 4.5) using ensemble time resolved fluorescence lifetime and anisotropy measurements. When Alexa647 dyes are coupled to different sites in T4L, they exhibit lifetimes ranging from $\langle\tau\rangle_x = 1.25$ ns to 1.5 ns, longer than that for the free Alexa647 dye ($\tau = 1.0$ ns³³³). The residual anisotropies, ranging between 0.16 - 0.31, are highly correlated to the corresponding lifetimes (Pearson's correlation coefficient, $r = 0.91$). The label K60 shows the longest lifetime and the highest residual anisotropy, while the P86 label demonstrates the shortest lifetime and lowest residual anisotropy. The fluorescence and anisotropy decay curves of these two labels are displayed in Fig.4.5B along with their fitted curves. Since the residual anisotropy is a measure of the environmental restriction to dye movement, it follows that, for the Alexa647 labels, the more sterically restricted dyes appear brighter with longer lifetimes, and vice versa. This phenomenon, called AIFE, is explained by the restriction to cis-trans isomerization of the Alexa647 dye (explained in detail in chapters 2 and 3)

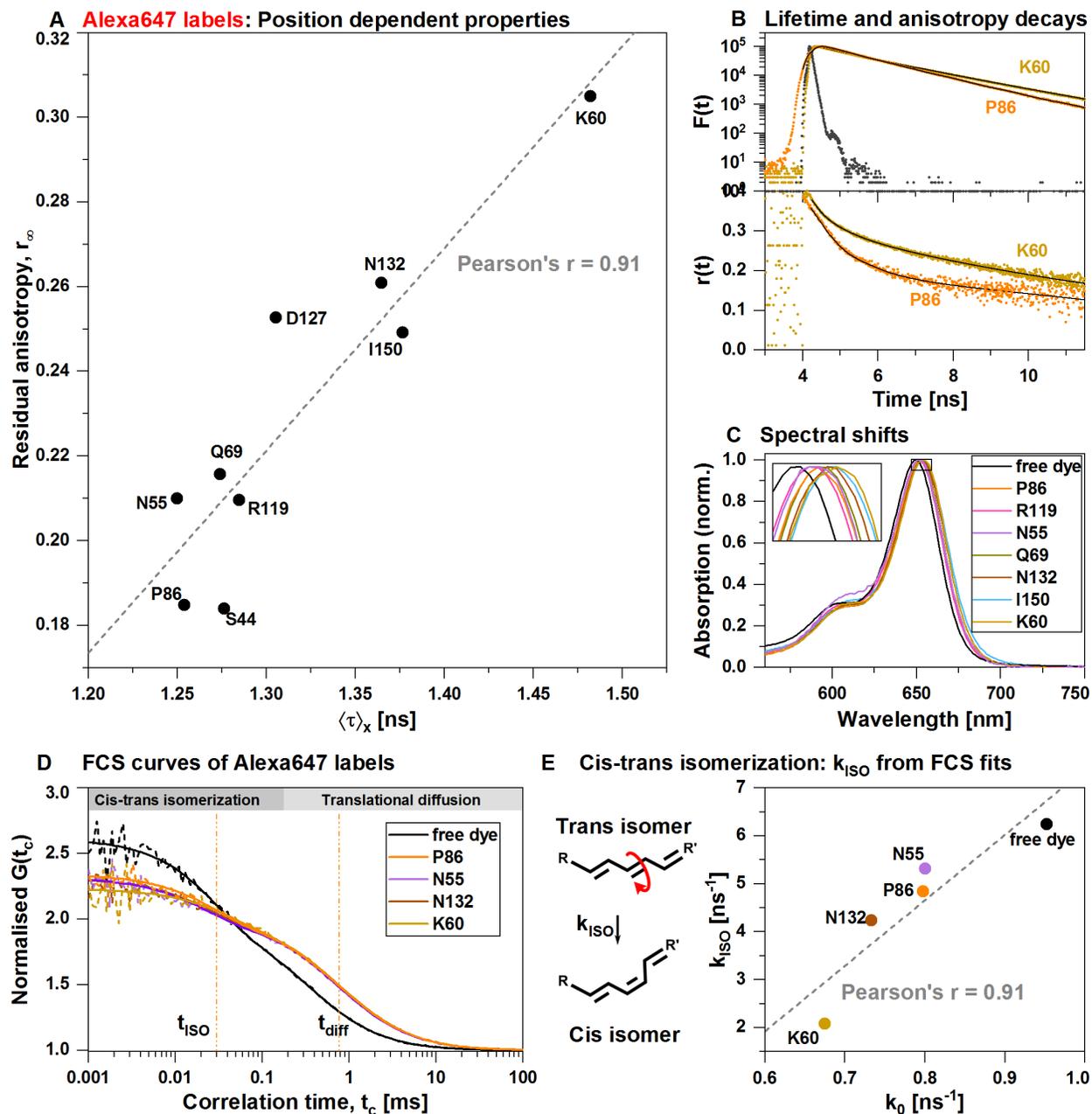


Figure 4.5. Photophysics of Alexa647 depends on the local protein environment. (A) The species-weighted average fluorescence lifetimes and residual anisotropies exhibit a wide range of values due to difference in dye environments. A positive correlation between these two observables can be seen throughout, indicating that the dye photophysics is influenced by its propensity to cis-trans isomerization. (B) Exemplary fluorescence and anisotropy decays of representative Alexa488 labels, P86 (orange) and K60 (yellow) with the instrument response function (grey) and their corresponding fits (black lines). (C) Absorption spectra of a few Alexa647 labels and the free dye show varying extents of red shift of labels compared to the free dye. The colors of the curves correspond to the label colors in panel A. (D) FCS curves (parallel-perpendicular correlations) of representative Alexa647 labels (dashed lines) and their fits (solid lines) measured under PIE signal at 635 nm and at an excitation power of 0.7 kW/cm².

The vertical lines indicate the fitted relaxation times for the translational diffusion (t_{diff}) and cis-trans isomerization (t_{ISO}) for the P86 label. (E) Sketch of cis-trans isomerization by bond rotation (red arrow) in Alexa647. Steric restriction to the dye movement affects the rate constant of this process (k_{ISO}) calculated from FCS fits. The decay rate constant and the residual anisotropy are highly correlated.

Cis-trans isomerization is modulated in restrictive environments. Cis-trans isomerism in Alexa647 labels exist due to the double-bonded group in the chromophore situated between two bulky heterocyclic rings (Fig. 4.1C). The potential energy barrier for cis-trans isomerization is increased in restrictive environments¹³⁶. As a result, the rate constant of photoisomerization decreases and the excited state deactivates mainly radiatively, thus enhancing the fluorescence. This gives rise to the phenomenon of association-induced fluorescence enhancement (AIFE) (Chapter 2), similar to PIFE^{135,136,220}, whereby the dye association with the local environment brings about an increase in fluorescence.

The observed photophysical responses of Alexa647 labels are thus governed by the rates of cis-trans isomerism. The wide distribution of lifetimes and residual anisotropies demonstrate the heterogeneous microenvironment of these labels. Alexa647 labels coupled in the vicinity of highly restrictive protein environments have larger residual anisotropies and increased brightness and longer lifetimes due to AIFE. Based on this, labels like K60, I150, N132 etc. are predicted to be in sterically confined states and likely trapped to the protein surface; while the environment of Alexa647 labels like P86 or S44 is expected to be mostly unrestrictive (Fig. 4.5A, Table 2).

The excellent correlation between the residual anisotropies and their corresponding lifetimes presents us with a novel information. A plot of fluorescence-averaged lifetime, $\langle\tau\rangle_f$ and the corresponding residual anisotropies (data as in Supplementary Table 4.5) yields a linear fit (Pearson's $r = 0.82$). When this linear function is extrapolated to the fluorescence lifetime of the free dye ($\tau=1$ ns), a residual anisotropy of 0.057 (denoted as $r_{\infty,extrapol.}$) is obtained. This may be considered the anisotropy of a covalently bonded Alexa647 dye that not trapped by the protein surface. Sindbert et.al. described the wobbling in a cone (WIC) model for a Cy5 dye and characterized the wobbling motion¹⁰³ using the equation:

$$\frac{1}{2} \cos \theta_{cone} (1 + \cos \theta_{cone}) = \sqrt{\frac{r_{\infty}}{r_0}} \quad (4.6)$$

where $\cos \theta_{cone}$ is the opening half angle.

The Alexa647 dye used here is structurally similar to the Cy5 dye and more importantly, its transition dipole moment is also parallel to the linker axis, allowing us to adopt the above equation. Replacing r_{∞} with $r_{\infty,extrapol.}$ in eq. (4.6), we calculate the opening half angle to be $\theta_{cone} = 59.3^{\circ}$.

Absorption spectra measured for Alexa647 labels show that the labels are red shifted compared to the free dye (Fig. 4.5C), indicating dye-T4L interactions. From the inset, it is seen that the most red-shifted labels have the highest fluorescence enhancement and vice versa. That is, these spectral shifts indicate the extent of dye-protein interactions.

To probe the cis-trans process in the Alexa647 labels, we tracked the relaxation processes for selected representative labels using fluorescent correlation spectroscopy (FCS) at 0.7 kW/cm^2 (Fig. 4.5D, Table 2, Supplementary Table 6). The free dye, being the least restricted, exhibited the highest amplitude at 0.38, for the cis-trans relaxation process; while the K60 label exhibited the lowest amplitude at 0.19 and the longest correlation time for the cis trans process ($38 \mu\text{s}$). The sketch in Fig. 4.5E shows the central double bond of Alexa647 undergoing geometric isomerism. The rate constant of isomerization (k_{ISO}) from the trans to cis isomer, can be determined from the amplitude and time of the cis-trans relaxation obtained from the FCS fits¹⁴¹ with equations (4.17) and (4.18). It shows good correlation (Pearson's $r = 0.91$) to the rate constant of fluorescence decay, k_0 , in accordance with the proposed AIFE mechanism. Highly trapped dyes like the label at K60 show lower rate constants ($k_{ISO} = 2.1 \text{ ns}^{-1}$), while relatively mobile dye labels have higher rate constants ($k_{ISO} = 4.2\text{-}5.3 \text{ ns}^{-1}$) and the free dye exhibits the highest $k_{ISO} = 6.3 \text{ ns}^{-1}$.

Table 4.2. Photophysical properties of representative Alexa647 labels.

| Label position | $\langle\tau\rangle_x$, ns | r_∞ | $\lambda_{\max}(\text{abs})$, nm | FCS @ 0.7kW/cm ² , 635 nm | | |
|-------------------|-----------------------------|------------|--------------------------------------|--------------------------------------|----------------------------------|-------------------------------------|
| | | | | A_{ISO} | t_{ISO} , μs | k_{ISO} , ns ⁻¹ |
| N55C | 1.25 | 0.21 | 652 | 0.24 | 19.89 | 5.31 |
| P86C | 1.25 | 0.18 | 652 | 0.24 | 23.17 | 4.84 |
| N132C | 1.36 | 0.26 | 653 | 0.24 | 24.39 | 4.23 |
| K60pAcF | 1.48 | 0.31 | 654 | 0.19 | 37.66 | 2.08 |
| Alexa647 free dye | 1.0 | - | 650 | 0.38 | 27.66 | 6.25 |

The species-weighted average lifetime, $\langle\tau\rangle_x$, calculated using eq. 4.5 and residual anisotropy, r_∞ , are obtained from fluorescence and anisotropy decay analysis. The free dye lifetime is from literature³³³. $\lambda_{\max}(\text{abs})$ is the wavelength at absorption maxima. A_{ISO} and t_{ISO} are the amplitudes and relaxation times for the cis-trans isomerization process obtained on fitting FCS curves measured at an excitation of 0.7 kW/cm² at 635 nm to the model function given by eq. 4.16. The rate constant of trans to cis isomerization, k_{ISO} , is calculated using equations, 4.17 and 4.18.

4.2.4 The source of steric restriction unraveled by AV simulations

Change in dye AV on protein bending and accessibility of Alexa647 to positively charged amino acids were calculated using the AV approach (for settings, refer to the Methods section) to reveal the source of steric restriction of Alexa647 labels. The protein dynamics, specifically, the hinge bending motion of T4L, is expected to interfere with the movement of the coupled Alexa647 dye. To investigate this possibility, the change in accessible volumes between C_1 and C_2 states was computed, analogous to that for the Alexa488 labels. The label K60 shows the highest deviation and correlates with high residual anisotropy.

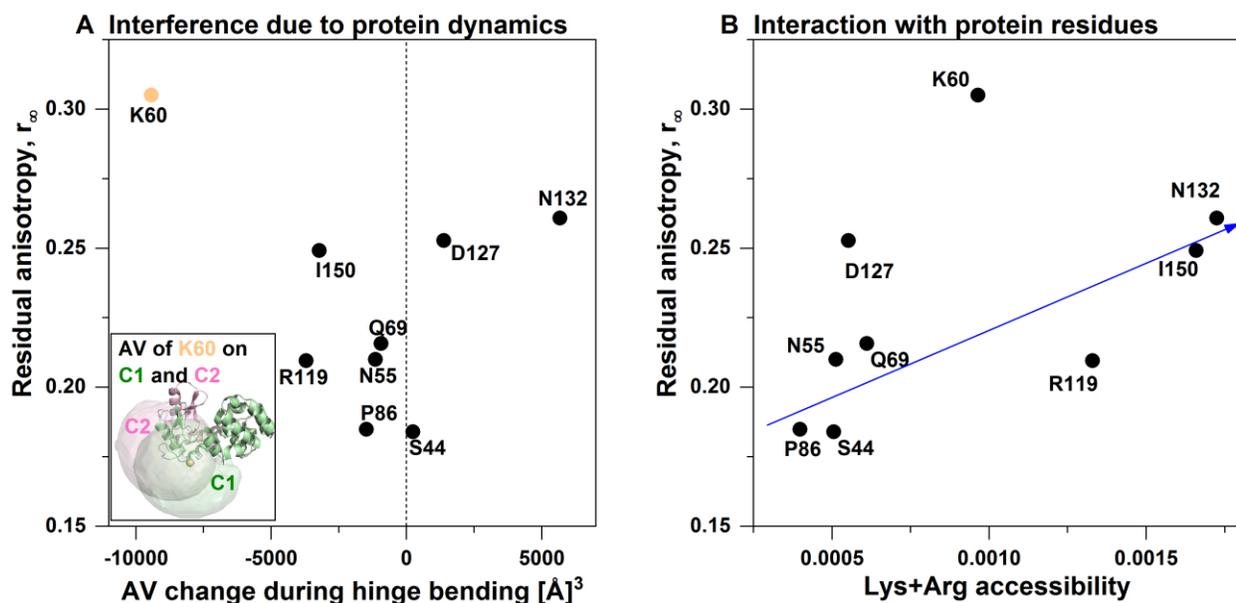


Figure 4.6. Alexa647 trapping sites in T4L. (A) Hinge bending motion of T4L: Change in AV between C_1 and C_2 states plotted against residual anisotropy show that the sites where dye labels are constrained due to protein dynamics have relatively high residual anisotropy. Inset: Example of an AV at labeling position K60 (spheres) in C_1 (green) and C_2 (pink) states. (B) Except for K60, where the residual anisotropy originates from interference from protein motion, the predicted accessibilities of the positively charged residues correlates moderately well with the measured residual anisotropy. The blue arrow is added as a guide for the eye (and is not a linear fit).

Since the Alexa647 dye is net negatively charged due to its sulfonated nature, it is expected to have affinity towards positive charges and can get trapped at these amino acid sites. The Lys+Arg accessibilities, defined as the fraction of the AV cloud within 4 \AA of any arginine or lysine, was calculated for PDB:172L (C_1 state) and PDB:142L (C_2 state) and were weighted by their prevalence in the cysteine-free pseudo-wild type of T4L³. When the accessibilities were plotted against the residual anisotropies, we did not get a good correlation, although a positive trend could be detected (Fig. 4.6B, blue arrow). For the label K60, the r_∞ is higher than expected for the corresponding predicted dye accessibility due to dye squeezing during protein dynamics. For the rest of the labels, the residual anisotropy is higher for labels with higher accessibilities to Lys/Arg. This implies that the molecular source of steric restriction is the dye trapping at the Lys and Arg residues. The absence of a good correlation could be due to the opposing effect of water quenching, whereby protonated solvents quench the fluorescence of red-emitting dyes

due to energy transfer from the excited state of the dye to the harmonics and combination bands of OH vibrational modes of water. The extent of this quenching effect depends on the solvent accessibility of the fluorophore and is therefore different at distinct protein environments.

4.2.5 Singly labeled dyes act as reporters of local environment and global events

In the previous sections, we investigated meticulously how the fluorescent properties of coupled Alexa488 and Alexa647 dyes are affected by the local environment. The insights from these studies equip us to identify the characteristics of an unknown local biomolecular environment, since their features will be reflected in the dye fluorescence properties. Thus, these fluorophores act as sensors by providing information on the local and global environments.

a) Information about local environment

Stern- Volmer plots for the Alexa488 labels can be used to assess the presence of PET quenchers in the neighborhood of the coupling site. For a cyanine-based dye like Alexa647, appreciable AIFE, characterized by high lifetimes and residual anisotropies, are indicative of a large number of Lys/Arg in the vicinity of the labeling site. Additionally, high residual anisotropies could indicate proximity to the active site of T4L. The T4L active site is located in the cleft region situated between the C and N terminal subdomains. The hinge bending motion of the protein would trap a dye in this region and therefore restrict its free motion. This would increase the residual anisotropy significantly, thereby indicating proximity to the active site.

b) Detecting protein unfolding and denaturation

(The temperature-dependent studies shown here were carried out by L.S.)

The study of unfolding and denaturation of proteins is an important area of research³³⁴ since it indicates the protein stability and the conditions which a protein can endure. Fluorescence is one among the many techniques used to investigate this phenomenon. Although FRET between donor and acceptor dyes are generally used to monitor the unfolding steps, singly labeled dyes can also provide valuable information. Fig. 4.7A displays the fluorescence quantum yields determined using eq. 4.12, for Alexa488 dye

labels coupled at positions P86, R119, D127, N132 and I150 in the rigid C-terminal sub domain, at temperatures ranging from 10-45°C. As the temperature is initially increased up to 35°C, the dyes show very little fluctuations in quantum yields, implying that the structure of the C-terminal domain of the protein is not perturbed. The slight increase in the quantum yield of some of the labels could be due to one of the following reasons:(1) An increasing distance from quenchers: The label P86, for example, is quenched by Tyrosine, Y88, and Tryptophan, W126, in the open T4L state (Fig.4.7A, inset). On increasing temperature, the quencher W126 on the neighboring helix may be moving further away, causing an increase in fluorescence quantum yield. (2) A reduction in static quenching at higher temperatures. Above 35°C, we begin to see marked changes in dye behavior for all labels, signifying increased perturbances in the protein environment. The fluorescence quantum yield of the free Alexa488 dye, however, does not vary over the measured temperature range. This confirms that the changes in the fluorescence properties of the dye labels occur due to environmental effects and is not due to degradation of the dye itself.

c) Detecting substrate binding

(The T4L variants used in this study were prepared and measured by K.H.)

Binding of the substrate, peptidoglycan, at the active site of protein T4L, can be detected by appropriately labeled Alexa647 labels. It is known that on substrate binding, the proportion of closed conformers is increased³, thereby creating a sterically restrictive environment and increasing accessibility to the interacting residues on the surface of T4L. Since for the Alexa647 dye, this steric restriction lowers the rate of cis-trans isomerization, measuring the rate constant of isomerization, helps us predict substrate binding.

To demonstrate this, I show the FCS curves of Alexa647 label I150 in the presence and absence of peptidoglycan (Fig. 4.7B, Supplementary Table 7). An additional functional mutation E11A, is introduced in the active site of the protein since this mutant is known to bind and unbind the substrate, but not cleave it³. The equilibrium ratios of the different states for the substrate-free functional mutant is 75%(C₁), 12%(C₂) and 13%(C₃); while that of the functional mutant in presence of substrate is 56%(C₁), 27%(C₂) and 17%(C₃)³ (Fig. 4.7B, upper inset). When excited at 635 nm with an excitation power of 0.7 kW/cm²,

the power-dependent rate constant of isomerization, $k'_{ISO}=7.8\text{ ms}^{-1}$ decreases to 4.9 ms^{-1} in the presence of the substrate (Fig. 4.7B).

The source of increased steric restriction for the closed C_2 conformer was investigated and revealed to be due to an increase in accessibility to positively charged amino acids. Using the AV approach, it was ascertained that when T4L changes from open to closed state, there is a 26% increase in dye accessibility to Lys/Arg for a label at I150 (Fig. 4.7B, lower inset). The increased dye trapping is reflected in the decreased rate constant of cis-trans isomerization. A suitably labeled Alexa647 dye can therefore report on the equilibrium shift between open and closed states of T4L and in turn, detect substrate binding.

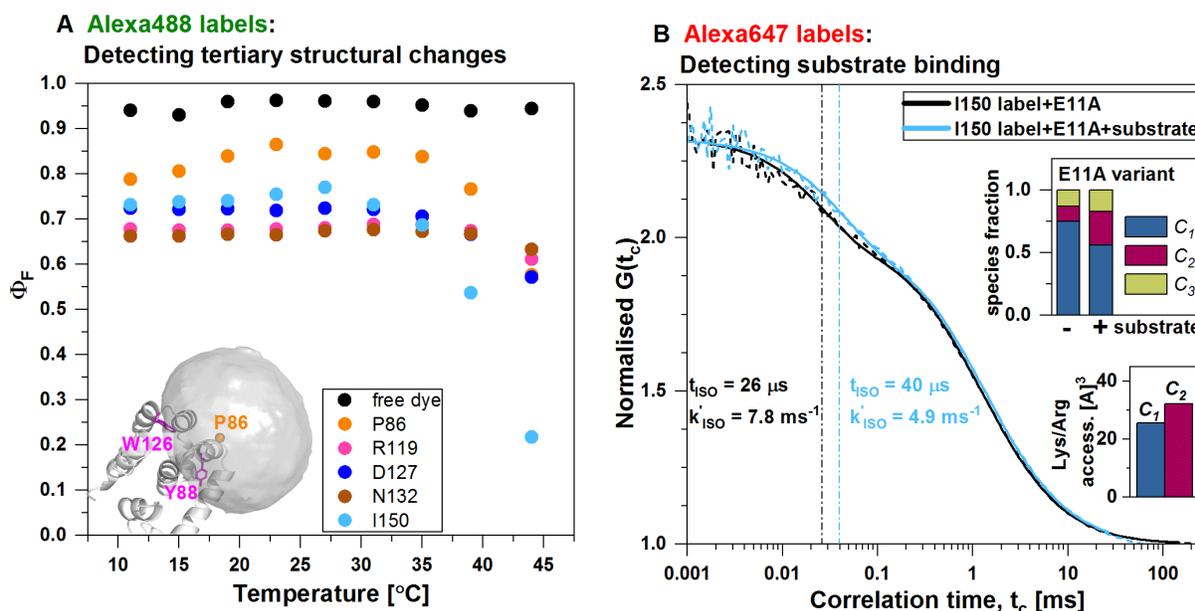


Figure 4.7. Dye as reporter of local and global environments. (A) The tertiary structural changes occurring in the T4L structure upon increasing the temperature can be sensed by the Alexa488 labels. The changes in the fluorescence quantum yields reflect the change in PET quencher distances as the structural elements begin to unfold. (B) Alexa647 labels detects substrate binding in E11A variant where the substrate is not cleaved on binding. The parallel-perpendicular FCS curve of the label I150 displays longer relaxation times for the cis-trans isomerization process on substrate binding with a corresponding decrease of the rate of isomerization. Substrate binding on T4L(E11A) corresponds to a larger proportion of closed states (upper inset³). The dye accessibility to Lys/Arg residues of the I150 label in these closed states are higher than in the open state. (lower inset); this increased steric restriction leading to decreased isomerization.

4.3 Conclusions and outlook

A combined investigation of the fluorescence properties of T4L-labeled Alexa488 and Alexa647 dyes using spectroscopy and simulations helped identify fluorescence quenching sites and sterically constraining environments. In addition to PET quenching amino acids, dye trapping was observed at positively charged residues like Lys and Arg. Multiple modes of dye-protein interactions like π - π stacking, H-bonding and cation- π interactions were recognized by MD simulations. AV simulations verified that protein dynamics could cause steric restrictions to dye motion, which are subsequently reflected in their photophysical properties. AV and BD simulations also successfully examined the source of quenching in Alexa488 labels.

From the insights gathered on dye-protein interactions, we are equipped to sense an unknown environment of a dye-labeled protein. With appropriate experiment design, dye labeled proteins can act as local or global environment sensors. In addition to locally assessing the proximity to PET quenchers or sterically restrictive environments, global events like substrate binding and protein denaturation can also be inferred. Thus, studying the local environment effect on the photophysical properties of covalently labeled dyes can potentially answer important structural questions in biology.

4.4 Materials and methods

Sample preparation. Site-directed mutagenesis was used to generate the cysteine and Amber (unnatural amino acid, *p*-acetylphenylalanine) mutants using the TAG codon in the pseudo-wild-type T4L. It contains the mutations C54T and C97A in comparison to wild-type T4L. This is done to replace all naturally-occurring cysteine residues. The generated mutants were cloned into the pET11a vector (Life Technologies Corp.)³³⁵. Transformation of the plasmid (containing the gene with the mutants at specific sites) into competent cells, either BL21(DE3) or Dh5 α E. coli cells (Life Technologies Corp.), protein production and purification were carried out according to the procedure outlined in Sanabria et.al³. Gel electrophoresis was performed to verify purity at different stages of protein purification and labeling (Supplementary fig.4.8).

Site-specific labeling of T4L: The sites with *p*-acetyl-L-phenylalanine (pAcPhe) were labeled by coupling its keto group to the hydroxylamine of the fluorophore under study.

The sites with cysteine residue were coupled by Michael addition of its thiol group with the maleimide of the appropriate fluorophore. For Alexa488 dyes, all measurements were done on singly labeled samples. However, the second mutation was present without being labeled (Supplementary Table 1). For the study of Alexa647 labels, doubly labeled samples were obtained by site-specific, orthogonal labeling; the second label being the Alexa488 dye (Supplementary Table 4). However, the employed measurement procedures ensure the acquisition of Alexa647 signal uninfluenced by the Alexa488 labeling. The maleimide and hydroxyl amine derivatives of Alexa488 and Alexa647 were obtained from Life Technologies Corp. For substrate binding studies of labeled T4L mutants with peptidoglycan (from *Micrococcus luteus*, Sigma-Aldrich), the procedure applied is described in Sanabria et.al³. Description of protocols followed for the synthesis of T4L and dye labeling are given in Chapter 7.

Fluorescence lifetime and anisotropy decay measurements and analysis:

Fluorescence decays were recorded using a FluoTime300 spectrometer (PicoQuant, Germany) equipped with a white light laser (SuperK Extreme EXW-12, NKT Photonics, Denmark) or an IBH-5000U (IBH-Scotland) with a diode laser LDH-P-C-470 (PicoQuant, Germany) for the green dye or diode laser LDH-8-1-126 (PicoQuant, Germany) for the red dye. To reduce the contribution of the scattered light, filters (Farbglas FGL 495 and FGL 645, Thorlabs, New Jersey, USA) were used to cutoff wavelengths at 495 nm and 645 nm for Alexa488 and Alexa647 samples respectively. Samples for measurement (20 μ l) were taken in quartz ultra-micro cuvettes (Helma #105.252.85.40). The IRF was recorded using the highly scattering LUDOX solution. For Alexa488 samples, the excitation and emission wavelengths were set at 488 nm and 520 nm. For Alexa647 data, the doubly labeled samples were used, but due to the excitation of sample at 635nm and emission recording at 665nm, we observe only the direct acceptor photophysics and not the FRET sensitized acceptor behavior.

To estimate the fluorescence lifetimes, the fluorescence decays with polarizer-detector setting, VM (parallel-magic angle), was fitted; while for anisotropies, the measured decays with the polarizer-detector settings, VV (parallel-parallel) and VH (parallel-perpendicular), were fitted. The fittings were performed using either inhouse softwares, PyFia (a python-

based fitting software) or `chisurf` (<https://github.com/Fluorescence-Tools/chisurf>) or using MATLAB scripts and the model functions are as described in Chapter 3.

The fluorescence decays $f(t)$ of single-labeled T4L molecules were described by

$$f(t) = \sum_i x^{(i)} e^{-t/\tau_i}, \quad (4.7)$$

where τ_i and $x^{(i)}$ are the fluorescence lifetimes and corresponding species fractions with $\sum_i x^{(i)} = 1$. Up to three lifetime components were used in the analysis. The number of components chosen were always the minimum number required to achieve a χ^2 of 1.15 and well-behaved residuals. The species-weighted average fluorescence lifetime ($\langle\tau\rangle_x$) and the fluorescence-weighted average fluorescence lifetime ($\langle\tau\rangle_f$) are then defined as:

$$\begin{aligned} \langle\tau\rangle_x &= \sum_i x^{(i)} \tau_i \\ \langle\tau\rangle_f &= \frac{\sum_i x^{(i)} \tau_i^2}{\sum_i x^{(i)} \tau_i} \end{aligned} \quad (4.8)$$

The anisotropy decays were described by:

$$r(t) = \sum_i r_i e^{-t/\rho_i}, \quad (4.9)$$

where ρ_i and r_i are the rotational correlation times and respective amplitudes. The sum of the anisotropy amplitudes is equal to the fundamental anisotropy $\sum_i r_i = r_0$, where $r_0 = 0.38$. For description of fits, refer to the Methods section in Chapter 3. Up to three rotational components were used for the analysis. The residual anisotropy r_∞ is defined as the anisotropy amplitude of the component with the longest rotational correlation time (ρ_{global}) which reflects the overall tumbling motion of the molecule (global motion).

Dynamic fluorescence quantum yield ($\Phi_{F,dyn}$). The dynamic fluorescence quantum yield ($\Phi_{F,dyn}$) is defined for those molecules undergoing dynamic/transient processes only. Since the species-weighted average fluorescence lifetime $\langle\tau\rangle_x$ of a dye is proportional to its fluorescence quantum yield,

$$\Phi_{F,dyn} = \Phi_{F,ref} \cdot \frac{\langle\tau\rangle_x}{\langle\tau\rangle_{x,ref}}, \quad (4.10)$$

where $\Phi_{F,ref}$ and $\langle\tau\rangle_{x,ref}$ are the fluorescence quantum yield and species-averaged fluorescence lifetimes of a reference sample with known quantum yield, respectively. Alexa488 free dye was used as the reference, and the measured values of $\langle\tau\rangle_x = 4.05$ ns and $\Phi_{F,ref} = 0.964$ were used.

Absorption and emission spectral measurements: Absorption spectra were measured with a UV-Vis spectrometer (Cary 4000, Agilent Technologies) at optical densities below 0.02 to avoid inner filter effects. The fluorescence spectra for fluorescence quantum yield measurements of Alexa488 labels were recorded with a fluorescence spectrometer (FluoroMax-4, Horiba Scientific) at an excitation of 460 nm. To exclude polarization effects, the fluorescence was collected at a 90° angle with the emission polarizer set to the magic angle (54.7°). All measurements were done at room temperature in the measurement buffer, unless otherwise specified.

Steady-state fluorescence quantum yield (Φ_F). For the Φ_F of Alexa488 labels at 20°C, measurements were performed according to³³⁶ by comparing the emission of the sample with Alexa488 free dye as the reference dye. The steady-state quantum yield is determined from:

$$\Phi_F = \Phi_{F,ref} \cdot \frac{F_s}{F_{ref}} \cdot \frac{A_{ref}}{A_s} \cdot \left[\frac{n_s}{n_{ref}} \right]^2 \quad (4.11)$$

where Φ_F and $\Phi_{F,ref}$ are the steady state fluorescence quantum yields of the sample and reference, F is the integrated intensity of the fluorescence emission spectra; A is the measured absorbance at the excitation wavelength, n is the refractive index of the solution of measurement and the subscripts s and ref refer to the sample and reference, respectively. The emission and absorption spectra were measured thrice at 3 different concentrations and the slope m of the plot between integrated fluorescence intensity and absorbance used to determine the fluorescence quantum yield as:

$$\Phi_F = \Phi_{F,ref} \cdot \frac{m_s}{m_{ref}} \cdot \left[\frac{n_s}{n_{ref}} \right]^2 \quad (4.12)$$

The $\Phi_{F,ref}$ of Alexa488 was measured to be 0.964 against a secondary standard reference dye, Rh110 whose $\Phi_F = 0.94$. This value was, in turn, calibrated vs the primary reference standards, Rh6G in EtOH and Fluorescein in NaOH³³⁶.

For determining the Φ_F at different temperatures in the range from 11°C to 44°C, the fluorescence intensities of the samples at 520 nm at different temperature conditions were measured and compared with the emission intensity at 19° C (explained further in Supplementary note 4.2). The steady state fluorescence quantum yields for all the labels at 20°C were measured as described above. Assuming that $\Phi_F^{20^\circ\text{C}} = \Phi_F^{19^\circ\text{C}}$, the quantum yields at temperatures from 11°C to 44°C were determined from their respective fluorescence intensities at 520 nm as:

$$\Phi_F^{T^\circ\text{C}} = \Phi_F^{19^\circ\text{C}} * \left(\frac{F_{520}^{T^\circ\text{C}}}{F_{520}^{19^\circ\text{C}}} \right) \quad (4.13)$$

where $\Phi_F^{T^\circ\text{C}}$ is the fluorescence quantum yield at temperature, T°C, and $F_{520}^{T^\circ\text{C}}$ is the fluorescence intensity at 520 nm at temperature, T°C.

Fraction of statically quenched species. From the steady state fluorescence quantum yield Φ_F , and the dynamic fluorescence quantum yield $\Phi_{F,dyn}$, we calculate the fraction of statically quenched species as:

$$f_{static} = 1 - \frac{\Phi_F}{\Phi_{F,dyn}} \quad (4.14)$$

FCS measurements. FCS experiments of Alexa488 labels were performed on a custom-built confocal microscope. The fluorescent molecules were excited by a diode laser (Cobolt 06-MLD, Germany) at 488 nm in continuous wave mode. The laser was focused into the sample by a water-immersion objective lens (UPLAPO 60 NA = 1.2, Olympus, Hamburg, Germany). The fluorescence was separated from the excitation by a dichroic beam splitter (488/636 PC, AHF, Tübingen, Germany). A polarizing beam splitter (VISHTII, Gsänger) was used to split the emission beam into two polarizations (parallel and perpendicular to the polarization of the excitation laser light) to eliminate dead time and after-pulsing artefacts in the correlation functions. Fluorescence signals were passed through bandpass filters (HQ 533/467, Germany) and registered on avalanche photodiode detectors (APD, SPCM-AQR-14, Perkin-Elmer). Cross-correlation functions of the parallel and perpendicular signals were computed using a homebuilt hardware correlator²³⁹. The FCS measurements to investigate the static quenching of Alexa488 labels were performed at an excitation power of 5.6 kW/cm² and wavelength of 485 nm.

The correlated curves were fitted in the range from 12 ns to 110 ms with a model comprising of two 3D diffusion terms and two bunching components at 4.3 μ s and 68 ns that are assigned to triplet state formation and quenching dynamics and fitted globally between the different dye labels, respectively:

$$G(t_c) = offset + \frac{1}{N_{bright}} \left[\frac{f_{dye}}{1 + \frac{t}{t_{diff,dye}}} \cdot \frac{1}{\sqrt{1 + \frac{t}{\left(\frac{z_0}{\omega_0}\right)^2 \cdot t_{diff,dye}}}} + (1 - f_{dye}) \cdot \frac{1}{1 + \frac{t}{t_{diff}}} \cdot \frac{1}{\sqrt{1 + \frac{t}{\left(\frac{z_0}{\omega_0}\right)^2 \cdot t_{diff}}} \right] \cdot (1 - A_T + A_T \cdot e^{-\frac{t}{t_T}} - A_Q + A_Q \cdot e^{-\frac{t}{t_Q}}) \quad (4.15)$$

where $N_{bright} = N_{total} \cdot (1 - A_T - A_Q)$

Here N_{total} refers to the total number of molecules and $t_{diff,dye}$ and t_{diff} are the diffusion times of the free dye and the labeled dye, respectively, f_{dye} is the fraction of free dye, ω_0 and z_0 are the $1/e^2$ radii of the 3D Gaussian shaped observation volume in the lateral and in axial dimension, respectively (typically, $z_0/\omega_0 = 3-6$ for our setup). t_T and A_T are the time and fraction, respectively, of the molecules in the triplet state and t_Q and A_Q are the relaxation time of static quenching and fraction of the statically quenched molecules.

FCS measurements of Alexa647 labels. Here, T4L samples doubly labeled with Alexa488 and Alexa647 dyes were subjected to pulsed interleaved excitation (PIE)³³⁷ which employs two pulsed lasers operating at identical frequencies, but with a delay corresponding to half the repetition period. The microtime information is used to distinguish photons based on their excitation source and detection channel. Alternating laser excitation (ALEX)³³⁸ allows us to extract the Alexa647 signals detected in parallel and perpendicular channels after direct excitation of this dye with the laser at 635 nm. The experimental set up is described in Hellenkamp et al.²³⁸. The excitation power of the laser applied was 0.7 kW/cm². Cross-correlation functions of the parallel and perpendicular signals were computed using a homebuilt hardware²³⁹. The correlated curves were fitted in the range from 94 ns to 150 ms with a model comprising of a diffusion term and a bunching term 20-40 μ s assigned to cis-trans isomerization as given below:

$$G(t_c) = offset + \frac{1}{N_{bright}} \cdot \frac{1}{1 + \frac{t}{t_{diff}}} \cdot \frac{1}{\sqrt{\left(1 + \frac{t}{\left(\frac{z_0}{\omega_0}\right)^2 \cdot t_{diff}}\right)}} \cdot \left(1 - A_{ISO} + A_{ISO} \cdot e^{-\frac{t}{t_{ISO}}}\right) \quad (4.16)$$

where $N_{bright} = N_{total} \cdot (1 - A_{ISO})$

Here N_{total} refers to the total number of molecules and t_{ISO} and A_{ISO} are the relaxation time and fraction, respectively, of molecules undergoing cis-trans isomerization. The rest of the parameters are the same as described in eq. 4.15.

The excitation power-dependent trans→cis isomerization rate constant, k'_{ISO} , and the cis→trans back isomerization, k'_{BISO} , can be calculated from the amplitude, $A_{ISO} = \frac{k'_{ISO}}{k'_{BISO}}$ and the relaxation time, $\frac{1}{t_{ISO}} = k'_{ISO} + k'_{BISO}$, as:

$$k'_{ISO} = \frac{A_{ISO}}{1 + A_{ISO}} \cdot \frac{1}{t_{ISO}} \quad \text{and} \quad k'_{BISO} = \frac{k'_{ISO}}{A_{ISO}} \quad (4.17)$$

From Widengren et al.¹⁴¹, we know that $k'_{ISO} = \frac{k_{01}}{k_{10} + k_{01}} \cdot k_{ISO}$. where k_{10} is the fluorescence decay rate constant, $k_{10} = \frac{1}{\langle \tau \rangle_x}$ and k_{01} is the rate constant of excitation. Since $k_{01} = \sigma_{01} \cdot \gamma_P \cdot I$, where σ_{01} is the absorption cross section, γ_P is the inverse photon energy and I is the excitation power, we can calculate the rate constant of trans→cis isomerization independent of the excitation power as:

$$k_{ISO} = \frac{k_{01}}{k_{10} + k_{01}} \cdot k'_{ISO} \quad (4.18)$$

Computational methods

AV calculations. AV calculations were performed using the 3-radii AV model, which calculated 3 single radius AV with different radii and overlays them to get the final AV. Each single radius AV approximates the dye as a sphere with radius R_{dye} , tethered to the biomolecule by a flexible linker of length L_{link} with width w_{link} ¹⁰³ (Supplementary Fig.4.9). The linker length L_{link} was estimated from the fully extended conformation of the linker

using the CambridgeSoft Chem3D software. The AV parameters were as follows: for Alexa488, $L_{\text{link}} = 20 \text{ \AA}$; $w_{\text{link}} = 4.5 \text{ \AA}$ and $R_{\text{dye}} = 5 \text{ \AA}, 4.5 \text{ \AA}, 1.5 \text{ \AA}$ and for Cy5, $L_{\text{link}} = 22 \text{ \AA}$, $w_{\text{link}} = 4.5 \text{ \AA}$ and $R_{\text{dye}} = 11 \text{ \AA}, 3 \text{ \AA}, 3.5 \text{ \AA}$. The attachment point for the labels was the β -carbon. The geometric modeling was performed using the software Olga (<https://github.com/Fluorescence-Tools/Olga>). The accessibilities were calculated as the volume of the AV within 4 \AA of the interested amino acid residue.

Simulating dynamic fluorescence quantum yield. A combination of AV and Brownian dynamics (BD) simulations were used to simulate dynamic quenching of Alexa488 using the procedure developed by Peulen et al.²⁷⁰. The dye is approximated as a sphere of radius $R_{\text{dye}} = 3.5 \text{ \AA}$ tethered to T4L with a linker of length, $L_{\text{link}} = 20 \text{ \AA}$ and width, $w_{\text{link}} = 4.5 \text{ \AA}$, diffusing with a diffusion coefficient, $D = 10 \text{ \AA}^2 \text{ ns}^{-1}$ within its accessible volume. Quenching with Trp, Tyr, Met and His is approximated by a step function where the dye is quenched with a rate constant, $k_{\text{Q}} = 2 \text{ ns}^{-1}$ if it is closer than distance $R_{\text{rad}} = 8.5 \text{ \AA}$ from the protein surface. A “sticky” region is defined closer than distance, $R_{\text{surface}} = 8 \text{ \AA}$ from the protein surface, where the diffusion coefficient of the dye is slowed down by a factor 0.1. Length of the BD simulation was 10 \mu s with a step length of 2 ps .

MD simulations. In order to investigate the dynamics of Alexa488 attached to the T4L protein, molecular dynamics simulations for five T4L variants (P86C, R119C, D127C, N132C, I150C) were performed using the GPU-Version^{294,295} of the Amber software suite (version 16)²⁵⁷. The ff14SB force field³³⁹ was used for the protein, while the force field parameters by Graen et al.³⁴⁰, which were adapted to Amber by Schepers et al.³⁴¹ were used for the dye. The starting structure of the protein was extracted from the crystal structure with PDB-ID 1LW9³⁴² by removing all molecules except for the protein. Subsequently, the dye was placed at the respective labeling position, and the original amino acid was replaced by the cysteine with attached maleimide linker. The protein-dye complex was then placed in a truncated octahedral water box of TIP4PEW^{288,343} water molecules. Na^+ and Cl^- ions were added as counter ions to reach a concentration of 0.1 M . Thermalization of the prepare system was done following previously described protocols³⁴⁴. For each variant, 10 independent simulations were performed by starting the thermalization process at slightly different initial temperature resulting in different initial

velocities of the atoms for each simulation. Each simulation was run for 3000 ns of simulation time, where coordinates were saved every 20 ps.

In addition to identifying potential modes of interaction, additional analyses were performed over all trajectories:

(a) Potential stacking interactions of the dye and quenching atoms: In order to investigate potential stacking interactions of Alexa488 with the quenching amino acids we used geometric criteria following Vaiana et al.¹⁰⁶, to calculate interaction parameters, r and φ , where r is the distance between the centers of the dye and the quenching amino acid and φ is the angle between the normal vectors of the rings (Supplementary Fig.4.4). Tryptophan and tyrosine were considered as potential quenching amino acids.

(b) Hydrogen bond analysis between Alexa488 and positively charged amino acids: In order to investigate potential hydrogen bonds between the dye Alexa488 and the positively charged amino acids, lysine and arginine, we used geometric criteria following the definitions for hydrogen bond from Raschka et al.³²⁹. Potential hydrogen bond acceptors on Alexa488 were the oxygens bound to the sulfur atoms, as well as the oxygens on the 2-carboxy group of the chromophore. The hydrogens on the NH₂ group of the lysine residue, as well as those on the two NH₂ groups of the guanidino group in the arginine residue were considered as potential hydrogen bond donors. A hydrogen bond is counted as present if three criteria are fulfilled at the same time (Supplementary Fig.4.10): (1) The distance between the hydrogen of the donor and the acceptor atom is smaller than 2.5 Å. (2) The angle between the donor nitrogen, the donor hydrogen, and the acceptor atom is between 120 and 180 degrees. (3) The angle between the donor hydrogen atom, the acceptor atom, and the carbon atom next to the acceptor is between 90 and 180 degrees. Distances and angles were calculated for every frame of the trajectory and then evaluated for the presence of potential hydrogen bonds.

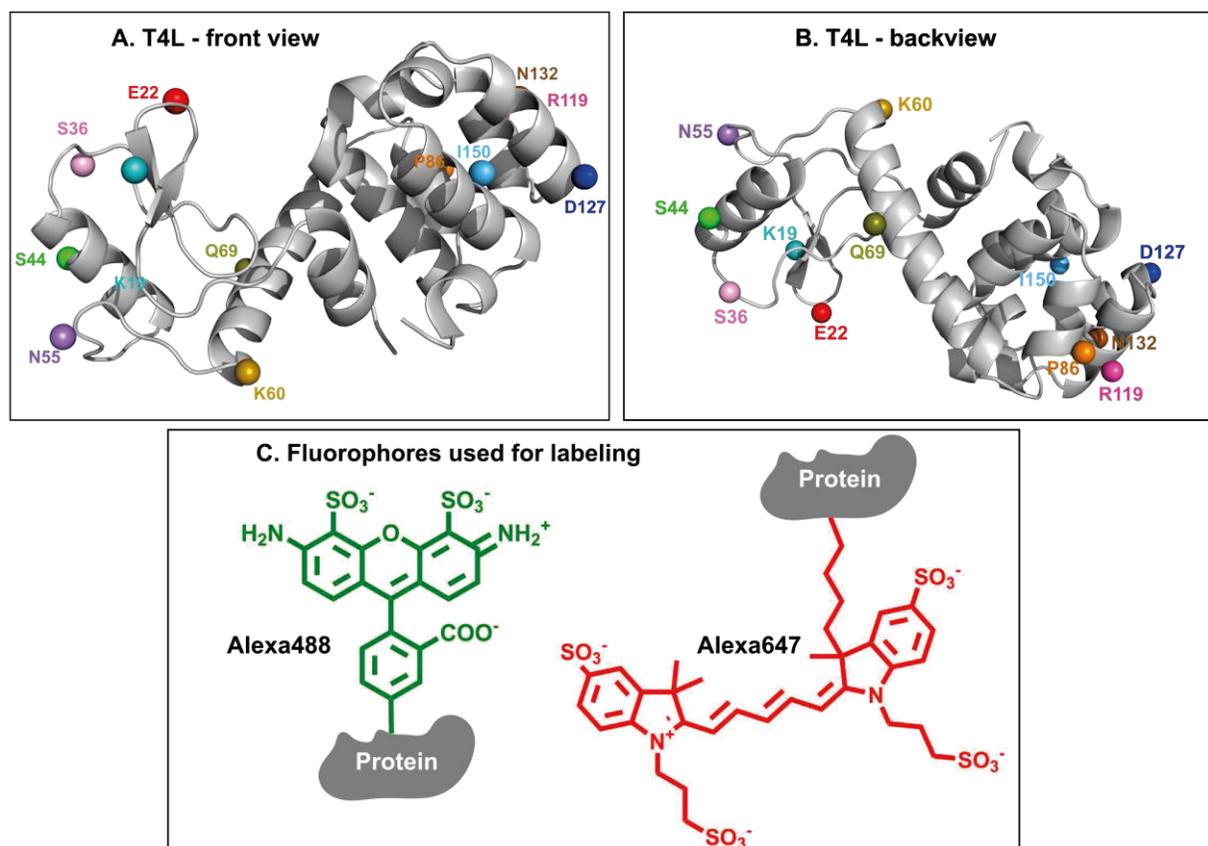
(c) Dye-amino acid distances between the oxygen atom of the xanthene ring of Alexa488 and the quenching amino acids Trp, Tyr, His and Met, and the positively charged amino acids, Lys and Arg were plotted for every 20 ps. The atoms chosen for distance calculations are (Supplementary Fig.4.3B) either carbon atoms approximately in the

middle of the aromatic rings or the sulfur atom, the case of Met. For Lys and Arg, the carbon atom closest to the positively charged nitro groups are chosen.

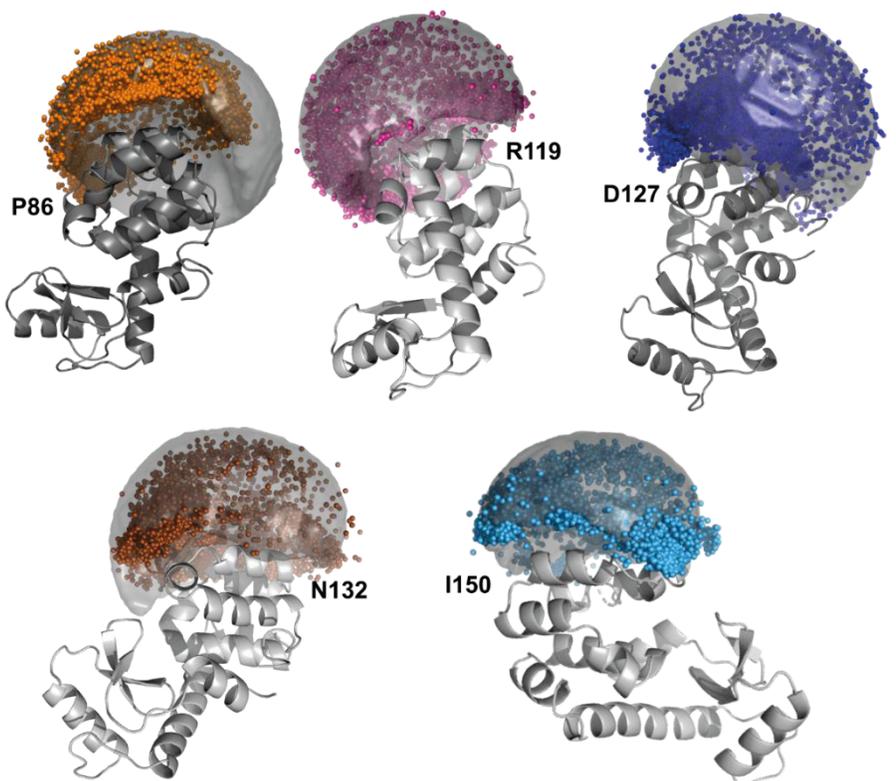
(d) Potential cation- π interactions were screened based on the geometric criteria of 6 \AA^{331} between the centres of any of the rings in the xanthene group of Alexa488 and the ammonium nitrogen (NZ) in Lys or the guanidinium carbon (CZ) in Arg. The screening was performed for every frame of the trajectory and then evaluated for the presence of potential cation- π interactions.

4.5 Supporting information

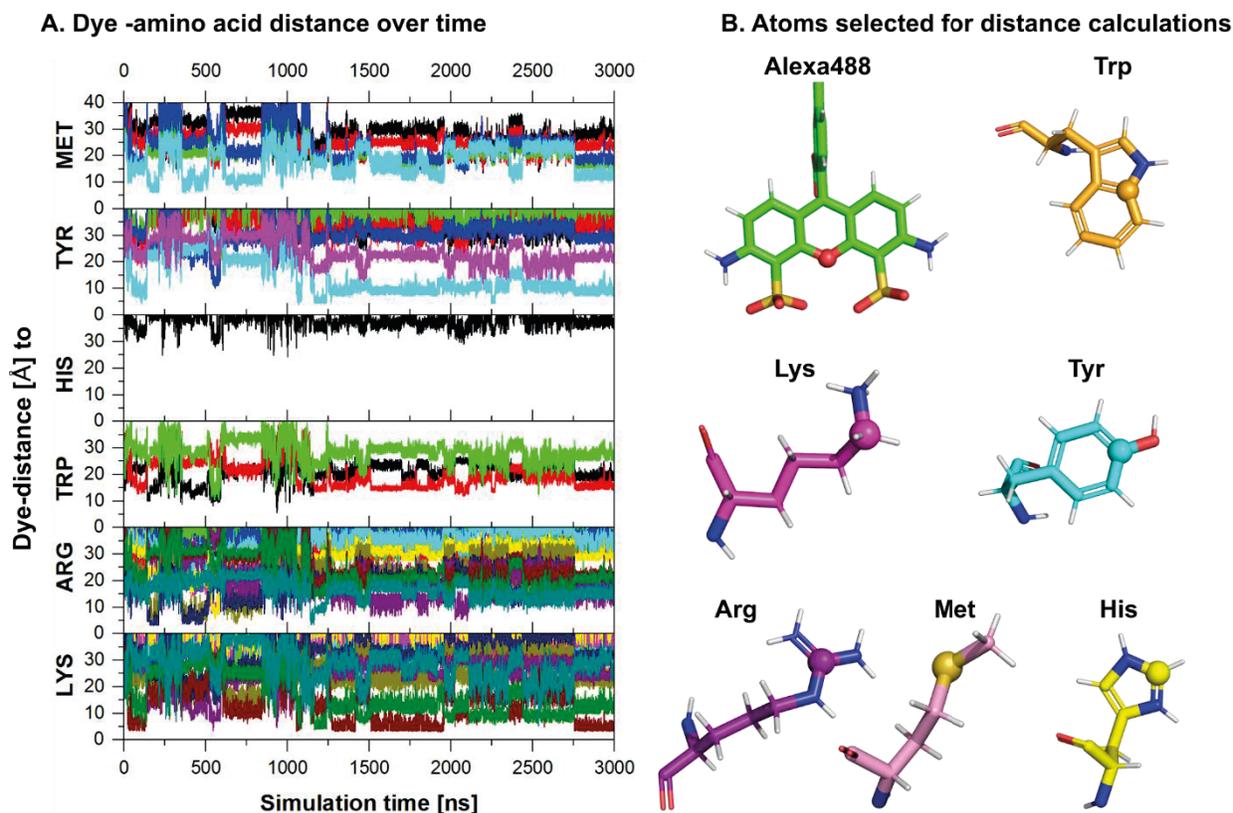
Supplementary Figures



Supplementary Figure 4.1: (A,B) The crystal structure of T4 lysozyme (C_1 state, PDB: 172L). The colored spheres represent the labeling sites of dyes used in the study. The mutations introduced for dye labeling and details of other mutations are given in Supplementary Tables 1 (Alexa488 labels) and 3 (Alexa647 labels). (B) Fluorophores used for labeling in this study.

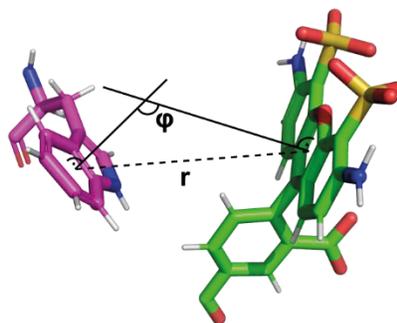


Supplementary Figure 4.2. Distribution of dye positions sampled during MD simulations for representative Alexa488 labels on protein T4L; the spheres representing the location of the O3 atom of Alexa488 every 2 ns of the simulation time. For each label, a total of 10 simulations, each of length 3 μ s are shown. These are overlaid with the accessible volumes (transparent clouds) determined from the AV simulations.

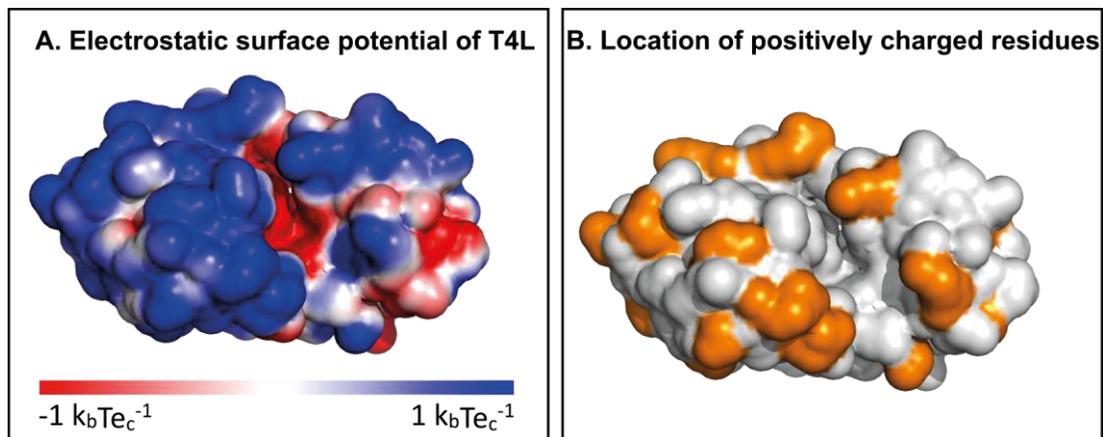


Supplementary Figure 4.3. (A) Dye- amino acid distance over time from MD simulations for every 20 ps is plotted. This representative example is from one of the 10 simulations of the Alexa488 label, D127. The different colors are due to the presence of multiple amino acid in T4L, namely, There are 5 Met, 6 Tyr, 1 His, 3 Trp, 13 Arg, 13 Lys. (B) The atoms chosen for distance calculations are indicated as spheres.

π - stacking interaction parameters

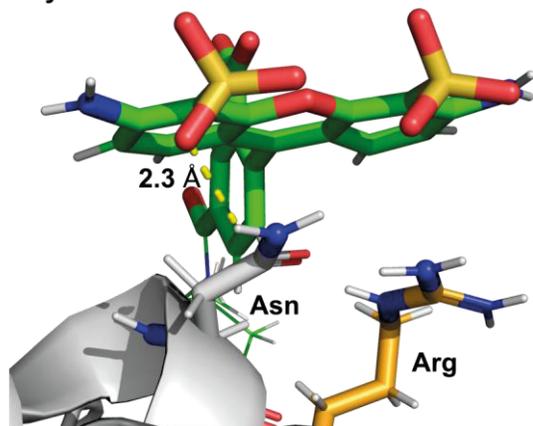


Supplementary Figure 4.4. Description of interaction parameters for π -stacking. The distance, r , is the distance between the centres of the ring systems and ϕ is the angle between the normals to the plane of the ring systems. The definitions are as described in Vaiana et al.¹⁰⁶

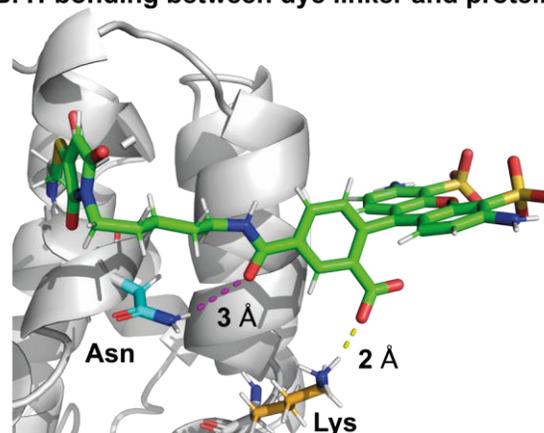


Supplementary Figure 4.5. Charge distribution on the T4L surface. The electrostatic surface potential (from APBS calculations with default parameters) show regions of positive charge (A) which coincide with the location of the positively charged amino acids, Lys and Arg, marked in orange (B). The electrostatic potential is given here in dimensionless units of $k_b T e_c^{-1}$ where k_b , the Boltzmann's constant = $1.3806504 \times 10^{-23} \text{ J K}^{-1}$, T is the temperature in K and e_c is the charge of an electron, $e = 1.60217646 \times 10^{-19} \text{ C}$.

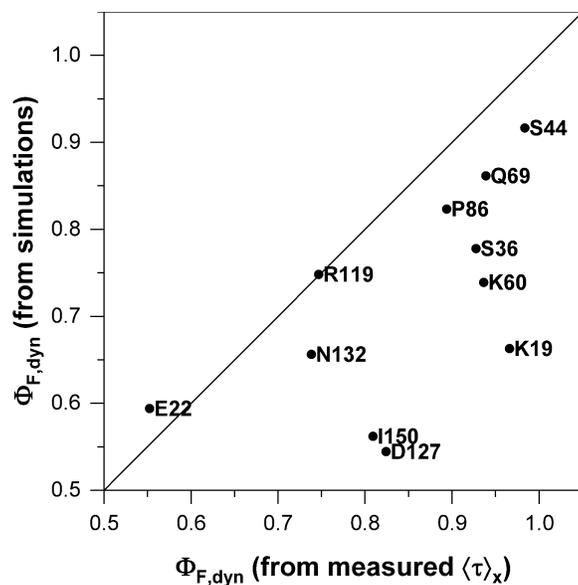
A. Dye interaction with basic amino acids



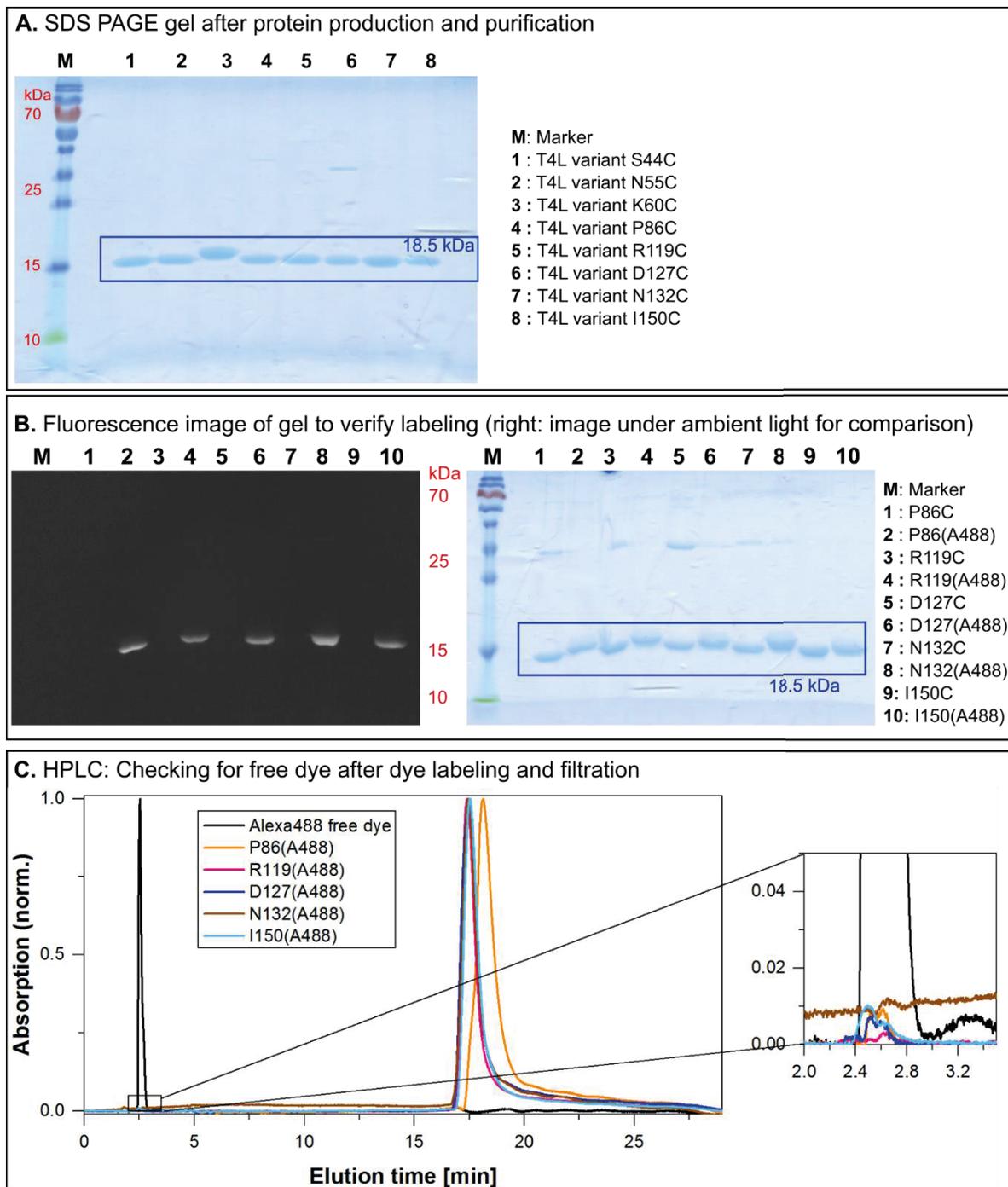
B. H-bonding between dye linker and protein



Supplementary Figure 4.6. Some interesting H-bonding interactions from MD simulations (A) H-bonding between dye and the amino acid, Asn (B) H-bonding interaction of dye-linker with amino acid (Asn) of T4L. Simultaneous H-bonding of dye with Lys is also observed.

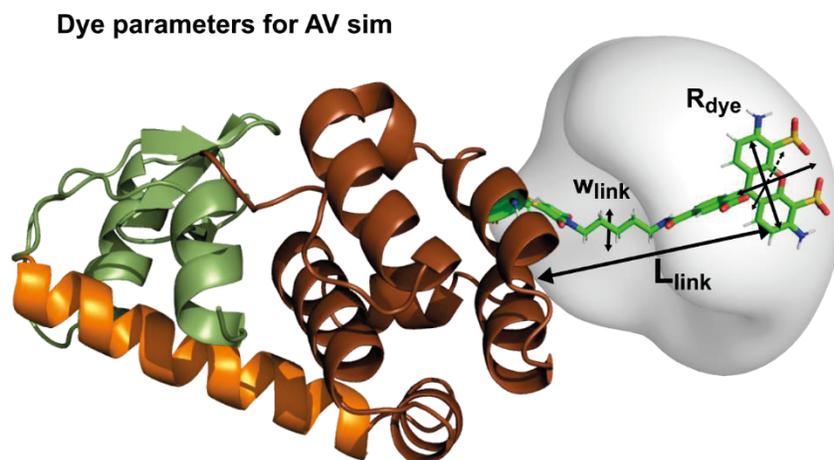


Supplementary Figure 4.7. Comparing measured and simulated dynamic fluorescence quantum yields. The measured quantities do not account for those interactions which result in static quenching. The simulations, on the other hand, assume that all interactions are dynamic. Therefore it includes all possible interactions, leading to a lower quantum yield than $\Phi_{F,dyn}$ calculated from experimentally determined $\langle\tau\rangle_x$. The black line shows a 1:1 relationship.

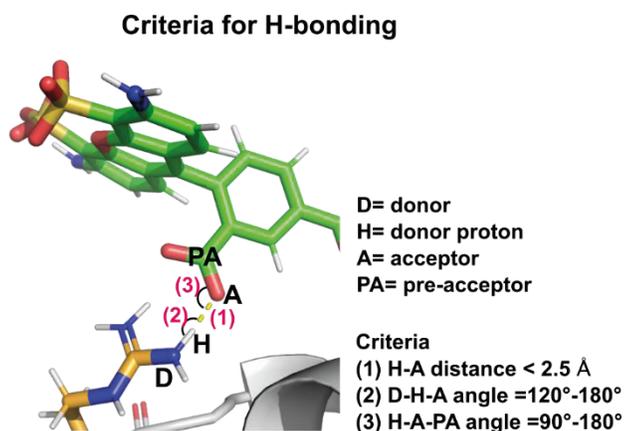


Supplementary Figure 4.8. (A) SDS polyacrylamide gelelectrophoresis (PAGE) was performed to evaluate the purity of samples after protein production and purification by ion exchange chromatography and stained with SimplyBlue™ SafeStain dye. The representative gel is shown here displays the expected band at 18.5k Da corresponding to the size of T4L. Marker (M) is GeneRuler™ 1kbp DNA ladder (B) After dye labeling the presence of the fluorescent label was verified with Fluorescence image of the gel captured at 460 nm excitation using a Cy2 emission filter (Amersham Imager 680, Cytiva, USA).The image of the same gel

under ambient light conditions is shown on the right to display the marker and the unlabeled protein variants which are non-fluorescent. (C) The labeled dyes checked for the presence of free dye using HPLC.



Supplementary Figure 4.9. Illustration of the parameters used in AV simulations. The dye is attached to the protein by a flexible linker of length L_{link} and width w_{link} . The 3 physical dimensions of the fluorophore are used to calculate the AV and independent calculations made by approximating the dye to a sphere of radius $R_{dye(i)}$ each time. The 3 AV simulations are then superimposed to obtain a spatial distribution of allowed dye positions (white cloud).



Supplementary Figure 4.10. The 3 criteria for assigning H bonding includes one distance and two angles³²⁹ as given in the figure above. These criteria were adopted in the determination of percentage of H-bonding in MD simulations.

Supplementary Tables

Supplementary Table 4.1. Mutations introduced to T4L wt* for investigating the photophysical properties of Alexa488 labels.

| Sample # | Alexa488 labels studied | Mutations introduced in T4L wt* for dye labeling | Derivation to T4L wt* | Presence of other mutations in the studied sample (unlabeled) |
|----------|-------------------------|--|-----------------------|---|
| 1 | K19 | K19Amber | pAcF at aa 19 | I150C |
| 2 | K19 | K19Amber | pAcF at aa 20 | N132C |
| 3 | E22 | E22Amber | pAcF at aa 22 | P86C |
| 4 | E22 | E22Amber | pAcF at aa 22 | D127C |
| 5 | S36 | S36Amber | pAcF at aa 36 | P86C |
| 6 | S36 | S36Amber | pAcF at aa 36 | N132C |
| 7 | S44 | S44Amber | pAcF at aa 36 | Q69C |
| 8 | S44 | S44Amber | pAcF at aa 44 | P86C |
| 9 | S44 | S44Amber | pAcF at aa 44 | R119C |
| 10 | S44 | S44Amber | pAcF at aa 44 | D127C |
| 11 | S44 | S44Amber | pAcF at aa 44 | N132C |
| 12 | S44 | S44Amber | pAcF at aa 44 | I150C |
| 13 | K60 | K60Amber | pAcF at aa 60 | P86C |
| 14 | K60 | K60Amber | pAcF at aa 60 | R119C |
| 15 | K60 | K60Amber | pAcF at aa 60 | N132C |
| 16 | K60 | K60Amber | pAcF at aa 60 | I150C |
| 17 | Q69 | Q69Amber | pAcF at aa 69 | P86C |
| 18 | Q69 | Q69Amber | pAcF at aa 69 | R119C |
| 19 | Q69 | Q69Amber | pAcF at aa 69 | N132C |
| 20 | Q69 | Q69Amber | pAcF at aa 69 | I150C |
| 21 | P86 | P86C | Cys at aa 86 | - |
| 22 | R119 | R119C | Cys at aa 119 | - |
| 23 | D127 | D127C | Cys at aa 127 | - |
| 24 | N132 | N132C | Cys at aa 132 | - |
| 25 | I150 | I150C | Cys at aa 150 | - |

Supplementary Table 4.2. Fitted lifetimes and anisotropy values for Alexa488 labels

| Label #. | Label | x_1 | $\tau_{1, ns}$ | x_2 | $\tau_{2, ns}$ | x_3 | $\tau_{3, ns}$ | $\langle\tau\rangle_x, ns$ | Av. $\langle\tau\rangle_x$ | r_1 | $\rho_{1, ns}$ | r_2 | $\rho_{2, ns}$ | $r_3 (r_\infty)$ | $\rho_{3, ns}$ |
|----------|----------|-------|----------------|-------|----------------|-------|----------------|----------------------------|----------------------------|-------|----------------|-------|----------------|------------------|----------------|
| | free dye | 1.00 | 4.05 | | | | | 4.05 | 4.05 | | | | | | |
| 1 | K19 | 0.02 | 7.56 | 0.91 | 4.28 | 0.07 | 0.78 | 4.09 | 4.06 | 0.08 | 0.22 | 0.05 | 1.84 | 0.25 | 12.77 |
| 2 | K19 | 0.28 | 5.07 | 0.64 | 3.99 | 0.08 | 0.70 | 4.03 | | 0.10 | 0.20 | 0.05 | 1.63 | 0.23 | 12.00 |
| 3 | E22 | 0.39 | 4.05 | 0.45 | 1.85 | 0.16 | 0.76 | 2.52 | 2.44 | 0.04 | 0.13 | 0.09 | 1.70 | 0.24 | 12.02 |
| 4 | E22 | 0.36 | 3.83 | 0.47 | 1.83 | 0.17 | 0.67 | 2.36 | | 0.05 | 0.14 | 0.12 | 1.78 | 0.20 | 16.33 |
| 5 | S36 | 0.81 | 4.40 | 0.14 | 2.28 | 0.05 | 0.49 | 3.91 | 3.92 | 0.10 | 0.17 | 0.09 | 1.44 | 0.19 | 12.00 |
| 6 | S36 | 0.83 | 4.39 | 0.13 | 2.23 | 0.05 | 0.46 | 3.93 | | 0.08 | 0.13 | 0.07 | 0.91 | 0.22 | 8.12 |
| 7 | S44 | 0.94 | 4.32 | 0.06 | 1.75 | | | 4.17 | 4.14 | 0.09 | 0.15 | 0.09 | 1.17 | 0.19 | 9.33 |
| 8 | S44 | 0.94 | 4.32 | 0.06 | 1.75 | | | 4.17 | | 0.11 | 0.19 | 0.07 | 1.27 | 0.20 | 8.60 |
| 9 | S44 | 0.96 | 4.32 | 0.04 | 1.44 | | | 4.20 | | | | | | | 8.60 |
| 10 | S44 | 0.85 | 4.28 | 0.10 | 2.25 | 0.05 | 0.39 | 3.87 | | | | | | | 9.33 |
| 11 | S44 | 0.96 | 4.32 | 0.04 | 1.44 | | | 4.20 | | | | | | | 8.60 |
| 12 | S44 | 0.96 | 4.32 | 0.04 | 1.44 | | | 4.20 | | | | | | | 8.60 |
| 13 | K60 | 0.94 | 4.12 | 0.06 | 2.07 | | | 4.00 | 3.93 | 0.10 | 0.18 | 0.08 | 1.24 | 0.20 | 8.24 |
| 14 | K60 | 0.91 | 4.26 | 0.09 | 1.81 | | | 4.04 | | 0.08 | 0.16 | 0.07 | 1.17 | 0.23 | 9.00 |
| 15 | K60 | 0.89 | 4.15 | 0.11 | 1.78 | | | 3.89 | | | | | | | 8.24 |
| 16 | K60 | 0.88 | 4.09 | 0.12 | 1.76 | | | 3.81 | | | | | | | 8.24 |
| 17 | Q69 | 0.94 | 4.20 | 0.06 | 1.52 | | | 4.04 | 3.94 | 0.12 | 0.18 | 0.09 | 1.07 | 0.17 | 8.26 |
| 18 | Q69 | 0.88 | 4.20 | 0.12 | 1.64 | | | 3.89 | | | | | | | 8.26 |
| 19 | Q69 | 0.89 | 4.20 | 0.11 | 1.47 | | | 3.90 | | | | | | | 8.26 |
| 20 | Q69 | 0.89 | 4.20 | 0.11 | 1.88 | | | 3.94 | | | | | | | 8.26 |
| 21 | P86 | 0.81 | 4.22 | 0.13 | 2.45 | 0.06 | 0.31 | 3.76 | 3.76 | 0.10 | 0.13 | 0.10 | 1.07 | 0.18 | 7.74 |
| 22 | R119 | 0.65 | 4.04 | 0.24 | 1.90 | 0.11 | 0.45 | 3.14 | 3.14 | 0.10 | 0.17 | 0.12 | 1.20 | 0.15 | 6.77 |
| 23 | D127 | 0.71 | 4.13 | 0.20 | 2.35 | 0.08 | 0.38 | 3.46 | 3.46 | 0.09 | 0.14 | 0.09 | 1.03 | 0.19 | 9.00 |
| 24 | N132 | 0.62 | 3.99 | 0.29 | 2.07 | 0.09 | 0.43 | 3.10 | 3.10 | 0.06 | 0.12 | 0.09 | 1.13 | 0.22 | 7.97 |
| 25 | I150 | 0.70 | 4.04 | 0.23 | 2.40 | 0.07 | 0.40 | 3.40 | 3.40 | 0.08 | 0.13 | 0.10 | 1.05 | 0.20 | 7.92 |

Label # corresponds to the sample number assigned in Supplementary table 1. The fitted lifetime components are τ_{1-3} and the corresponding fractions are x_{1-3} . Fitting for anisotropy yielded the components ρ_{1-3} and corresponding fractions, b_{1-3} . The longest anisotropy component arises from the global motion and its amplitude is the residual anisotropy (r_∞). $\langle\tau\rangle_x$ is the species-weighted average fluorescence lifetime. Av. $\langle\tau\rangle_x$ is the average $\langle\tau\rangle_x$ for a given label position.

Supplementary Table 4.3. Fit results for the FCS curves of Alexa488 labels.

| Label # | Alexa488 label | offset | N_{bright} | t_{diff} (ms) | z_0/w_0 | $f_{\text{free dye}}$ | $t_{\text{diff, free dye}}$ (ms) | A_T | t_T (ns) | A_Q | t_Q (μs) |
|---------|-----------------|--------|---------------------|------------------------|-----------|-----------------------|----------------------------------|-------|------------|-------------|-------------------------|
| 21 | P86 | 1.00 | 3.78 | 0.90 | 4.19 | 0.06 | 0.26 | 0.09 | 4.28 | 0.06 | 67.87 |
| 22 | R119 | 1.00 | 3.14 | 0.90 | 4.19 | 0.11 | 0.26 | 0.07 | 4.28 | 0.13 | 67.87 |
| 23 | D127 | 1.00 | 3.10 | 0.90 | 4.19 | 0.14 | 0.26 | 0.03 | 4.28 | 0.13 | 67.87 |
| 24 | N132 | 1.00 | 5.68 | 0.90 | 4.19 | 0.17 | 0.26 | 0.08 | 4.28 | 0.11 | 67.87 |
| 25 | I150 | 1.00 | 3.50 | 0.90 | 4.19 | 0.17 | 0.26 | 0.06 | 4.28 | 0.15 | 67.87 |
| | free dye | 1.00 | 3.01 | 0.90 | 4.19 | 1.00 | 0.26 | 0.09 | 4.28 | 0.00 | 67.87 |

Label # corresponds to the sample number assigned in Supplementary table 1. FCS was measured at 635 nm excitation and power of 0.7 kW/cm² and the average of parallel-perpendicular and perpendicular-parallel correlation curves fitted to the model function described in eq. 4.15. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z_0/w_0 describes the shape of the observation volume, t_T and A_T are the time and fraction, respectively, of the molecules in the triplet state and t_Q and A_Q , of the statically quenched molecules. Due to presence of free dye, an additional diffusion term is used for fitting which describes the diffusion of free dye. All relaxation times are globally fitted between the different molecules.

Supplementary Table 4.4. Mutations introduced to T4L wt* for investigating the photophysical studies of Alexa647 labels.

| Sample # | Alexa647 labels studied | Mutations introduced in T4L wt* for Alexa647 labeling | Derivation to wt* | Other mutations and labels in the studied sample | |
|----------|-------------------------|---|-------------------|--|--------------|
| | | | | Mutations | Labeled with |
| 26 | S44 | S44C | Cys at aa 44 | E5Amber(pAcF) | Alexa488 |
| 27 | S44 | S44Amber | pAcF at aa 44 | R119C | Alexa488 |
| 28 | N55 | N55C | Cys at aa 55 | R8Amber(pAcF) | Alexa488 |
| 29 | Q69 | Q69C | Cys at aa 69 | R8Amber(pAcF) | Alexa488 |
| 30 | Q69 | Q69C | Cys at aa 69 | K19Amber(pAcF) | Alexa488 |
| 31 | Q69 | Q69C | Cys at aa 69 | S44Amber(pAcF) | Alexa488 |
| 32 | Q69 | Q69C | Cys at aa 69 | N55Amber(pAcF) | Alexa488 |
| 33 | P86 | P86C | Cys at aa 86 | R8Amber(pAcF) | Alexa488 |
| 34 | P86 | P86C | Cys at aa 86 | K19Amber(pAcF) | Alexa488 |
| 35 | P86 | P86C | Cys at aa 86 | E22Amber(pAcF) | Alexa488 |
| 36 | P86 | P86C | Cys at aa 86 | S36Amber(pAcF) | Alexa488 |
| 37 | P86 | P86C | Cys at aa 86 | S44Amber(pAcF) | Alexa488 |
| 38 | P86 | P86C | Cys at aa 86 | K60Amber(pAcF) | Alexa488 |
| 39 | P86 | P86C | Cys at aa 86 | Q69Amber(pAcF) | Alexa488 |
| 40 | R119 | R119C | Cys at aa 119 | D70Amber(pAcF) | Alexa488 |
| 41 | R119 | R119C | Cys at aa 119 | Q69Amber(pAcF) | Alexa488 |
| 42 | R119 | R119C | Cys at aa 119 | R8Amber(pAcF) | Alexa488 |
| 43 | R119 | R119C | Cys at aa 119 | K19Amber(pAcF) | Alexa488 |
| 44 | R119 | R119C | Cys at aa 119 | N55Amber(pAcF) | Alexa488 |
| 45 | R119 | R119C | Cys at aa 119 | K60Amber(pAcF) | Alexa488 |
| 46 | R119 | R119C | Cys at aa 119 | S44Amber(pAcF) | Alexa488 |

| | | | | | |
|-----|--------------|----------|---------------|----------------|----------|
| 47 | N132 | N132C | Cys at aa 132 | S44Amber(pAcF) | Alexa488 |
| 48 | N132 | N132C | Cys at aa 132 | R8Amber(pAcF) | Alexa488 |
| 49 | N132 | N132C | Cys at aa 132 | K19Amber(pAcF) | Alexa488 |
| 50 | N132 | N132C | Cys at aa 132 | S36Amber(pAcF) | Alexa488 |
| 51 | N132 | N132C | Cys at aa 132 | N55Amber(pAcF) | Alexa488 |
| 52 | N132 | N132C | Cys at aa 132 | K60Amber(pAcF) | Alexa488 |
| 53 | N132 | N132C | Cys at aa 132 | D70Amber(pAcF) | Alexa488 |
| 54 | N132 | N132C | Cys at aa 132 | E5Amber(pAcF) | Alexa488 |
| 55 | N132 | N132C | Cys at aa 132 | Q69Amber(pAcF) | Alexa488 |
| 56 | I150 | I150C | Cys at aa 150 | N55Amber(pAcF) | Alexa488 |
| 57 | I150 | I150C | Cys at aa 150 | Q69Amber(pAcF) | Alexa488 |
| 58 | K60 | K60Amber | pAcF at aa 60 | N132C | Alexa488 |
| 59 | D127 | D127C | Cys at aa 127 | S44Amber(pAcF) | Alexa488 |
| 60 | D127 | D127C | Cys at aa 127 | E22Amber(pAcF) | Alexa488 |
| 61* | I150* | I150C | Cys at aa 150 | S44Amber(pAcF) | Alexa488 |

* This label contains additional E11A functional mutation

In all these samples, Alexa488 labels are also present, but do not affect the fluorescence signals from Alexa647 labels under study due to direct excitation at 635nm and emission monitoring at 665nm ensuring Alexa647-only properties, similar to a singly labeled sample (explained further in the methods section).

Supplementary Table 4.5. Fitted lifetimes and anisotropy values for Alexa647 labels

| #. | Label | x_1 | τ_1 , ns | x_2 | τ_2 , ns | x_3 | τ_3 , ns | $\langle\tau\rangle_x$, ns | $\langle\tau\rangle_f$, ns | Av. $\langle\tau\rangle_x$ | r_1 | ρ_1 , ns | r_2 | ρ_2 , ns | r_3 (r_∞) | ρ_3 , ns |
|----|-------|-------|---------------|-------|---------------|-------|---------------|-----------------------------|-----------------------------|----------------------------|-------|---------------|-------|---------------|----------------------|---------------|
| 26 | S44 | 0.30 | 1.68 | 0.65 | 1.12 | 0.05 | 0.32 | 1.24 | 1.33 | 1.28 | 0.08 | 0.03 | 0.14 | 0.70 | 0.17 | 10.68 |
| 27 | S44 | 0.40 | 1.77 | 0.53 | 1.09 | 0.06 | 0.26 | 1.31 | 1.45 | | 0.03 | 0.06 | 0.15 | 0.62 | 0.20 | 12.75 |
| 28 | N55 | 0.35 | 1.67 | 0.59 | 1.09 | 0.06 | 0.34 | 1.25 | 1.35 | 1.25 | 0.03 | 0.14 | 0.12 | 0.84 | 0.24 | 12.11 |
| 29 | Q69 | 0.09 | 2.13 | 0.62 | 1.37 | 0.30 | 0.79 | 1.26 | 1.37 | 1.27 | 0.08 | 0.17 | 0.12 | 0.86 | 0.20 | 11.84 |
| 30 | Q69 | 0.14 | 2.00 | 0.58 | 1.29 | 0.28 | 0.71 | 1.23 | 1.36 | | 0.06 | 0.08 | 0.12 | 0.57 | 0.21 | 13.96 |
| 31 | Q69 | 0.29 | 1.89 | 0.60 | 1.24 | 0.10 | 0.52 | 1.35 | 1.48 | | 0.05 | 0.05 | 0.11 | 0.73 | 0.23 | 12.53 |
| 32 | Q69 | 0.09 | 2.12 | 0.60 | 1.37 | 0.30 | 0.76 | 1.26 | 1.38 | | 0.04 | 0.08 | 0.13 | 0.66 | 0.22 | 13.12 |
| 33 | P86 | 0.28 | 1.75 | 0.72 | 1.05 | | | 1.25 | 1.33 | 1.25 | 0.06 | 0.10 | 0.15 | 0.79 | 0.18 | 11.26 |
| 34 | P86 | 0.02 | 2.37 | 0.35 | 1.59 | 0.64 | 1.00 | 1.22 | 1.31 | | 0.07 | 0.04 | 0.16 | 0.71 | 0.17 | 12.58 |
| 35 | P86 | 0.26 | 1.83 | 0.66 | 1.19 | 0.08 | 0.53 | 1.30 | 1.4 | | 0.04 | 0.17 | 0.14 | 0.87 | 0.21 | 11.38 |
| 36 | P86 | 0.33 | 1.79 | 0.62 | 1.14 | 0.05 | 0.36 | 1.32 | 1.42 | | 0.04 | 0.09 | 0.14 | 0.79 | 0.21 | 10.73 |
| 37 | P86 | 0.03 | 2.39 | 0.58 | 1.40 | 0.39 | 0.81 | 1.20 | 1.31 | | 0.06 | 0.09 | 0.17 | 0.78 | 0.16 | 19.93 |
| 38 | P86 | 0.06 | 2.16 | 0.61 | 1.37 | 0.34 | 0.86 | 1.25 | 1.33 | | 0.03 | 0.06 | 0.17 | 0.89 | 0.19 | 17.79 |
| 39 | P86 | 0.10 | 1.96 | 0.47 | 1.36 | 0.43 | 0.93 | 1.24 | 1.32 | | 0.06 | 0.05 | 0.15 | 0.91 | 0.18 | 14.09 |
| 40 | R119 | 0.02 | 2.57 | 0.55 | 1.65 | 0.43 | 0.97 | 1.37 | 1.47 | 1.28 | 0.05 | 0.06 | 0.07 | 0.73 | 0.27 | 10.13 |
| 41 | R119 | 0.31 | 1.77 | 0.69 | 1.05 | | | 1.28 | 1.37 | | 0.11 | 0.35 | 0.10 | 1.72 | 0.19 | 19.97 |
| 42 | R119 | 0.39 | 1.73 | 0.57 | 1.08 | 0.05 | 0.36 | 1.30 | 1.41 | | 0.04 | 0.12 | 0.13 | 0.80 | 0.22 | 11.18 |
| 43 | R119 | 0.13 | 1.96 | 0.63 | 1.30 | 0.24 | 0.73 | 1.25 | 1.36 | | 0.03 | 0.03 | 0.15 | 0.80 | 0.21 | 17.19 |
| 44 | R119 | 0.22 | 1.85 | 0.62 | 1.26 | 0.16 | 0.47 | 1.26 | 1.4 | | 0.04 | 0.16 | 0.14 | 0.90 | 0.21 | 14.11 |
| 45 | R119 | 0.15 | 1.99 | 0.71 | 1.26 | 0.14 | 0.60 | 1.27 | 1.38 | | 0.03 | 0.17 | 0.13 | 0.82 | 0.23 | 11.47 |
| 46 | R119 | 0.03 | 2.54 | 0.64 | 1.45 | 0.33 | 0.77 | 1.25 | 1.37 | | 0.09 | 0.02 | 0.15 | 0.99 | 0.15 | 14.80 |
| 47 | N132 | 0.02 | 2.68 | 0.64 | 1.56 | 0.34 | 0.84 | 1.34 | 1.45 | 1.36 | 0.07 | 0.12 | 0.06 | 0.67 | 0.27 | 10.45 |

| | | | | | | | | | | | | | | | | |
|----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|
| 48 | N132 | 0.18 | 1.99 | 0.62 | 1.34 | 0.20 | 0.72 | 1.33 | 1.44 | | 0.03 | 0.07 | 0.11 | 0.63 | 0.25 | 11.14 |
| 49 | N132 | 0.15 | 2.18 | 0.66 | 1.54 | 0.20 | 0.73 | 1.48 | 1.6 | | 0.04 | 0.16 | 0.08 | 0.96 | 0.28 | 10.37 |
| 50 | N132 | 0.21 | 1.98 | 0.61 | 1.40 | 0.18 | 0.72 | 1.40 | 1.51 | | 0.03 | 0.09 | 0.07 | 0.76 | 0.29 | 10.10 |
| 51 | N132 | 0.20 | 1.93 | 0.65 | 1.27 | 0.15 | 0.60 | 1.30 | 1.42 | | 0.10 | 0.12 | 0.08 | 0.78 | 0.22 | 11.38 |
| 52 | N132 | 0.28 | 1.87 | 0.57 | 1.26 | 0.15 | 0.70 | 1.35 | 1.45 | | 0.05 | 0.02 | 0.07 | 0.53 | 0.28 | 8.61 |
| 53 | N132 | 0.01 | 2.98 | 0.54 | 1.69 | 0.46 | 0.99 | 1.37 | 1.47 | | 0.05 | 0.07 | 0.07 | 0.75 | 0.27 | 9.80 |
| 54 | N132 | 0.44 | 1.78 | 0.51 | 1.10 | 0.05 | 0.22 | 1.35 | 1.48 | | 0.03 | 0.09 | 0.07 | 0.74 | 0.29 | 10.84 |
| 55 | N132 | 0.50 | 1.74 | 0.50 | 0.99 | | | 1.36 | 1.47 | | 0.11 | 0.23 | 0.06 | 2.86 | 0.22 | 11.06 |
| 56 | I150 | 0.60 | 0.93 | 0.40 | 1.88 | | | 1.31 | 1.47 | 1.38 | | | 0.15 | 0.87 | 0.23 | 29.46 |
| 57 | I150 | 0.35 | 0.84 | 0.05 | 2.65 | 0.59 | 1.70 | 1.44 | 1.61 | | | | 0.11 | 0.44 | 0.27 | 9.20 |
| 58 | K60 | 0.46 | 1.98 | 0.48 | 1.16 | 0.06 | 0.28 | 1.48 | 1.65 | 1.48 | 0.02 | 0.16 | 0.06 | 0.75 | 0.31 | 12.16 |
| 59 | D127 | 0.39 | 1.75 | 0.55 | 1.09 | 0.06 | 0.16 | 1.29 | 1.43 | | 0.04 | 0.13 | 0.09 | 0.83 | 0.25 | 11.78 |
| 60 | D127 | 0.38 | 1.79 | 0.56 | 1.11 | 0.05 | 0.19 | 1.32 | 1.46 | 1.31 | 0.05 | 0.32 | 0.08 | 1.17 | 0.25 | 11.27 |

corresponds to the sample number assigned in Supplementary table 4. The fitted lifetime components are τ_{1-3} and the corresponding fractions are x_{1-3} . Fitting for anisotropy yielded the components ρ_{1-3} and corresponding fractions, r_{1-3} . The longest anisotropy component arises from the global motion and its amplitude is the residual anisotropy (r_{∞}). $\langle\tau\rangle_x$ is the species-weighted average fluorescence lifetime. $\langle\tau\rangle_f$ is the fluorescence-average lifetime. Av. $\langle\tau\rangle_x$ is the average $\langle\tau\rangle_x$ for a given label position.

Supplementary Table 4.6. Fit results for the FCS curves of Alexa647 labels.

| Label # | Alexa647 label | Alexa488 position | offset | N _{bright} | t _{diff} (ms) | z ₀ /w ₀ | A _{iso} | t _{iso} (μs) |
|---------|----------------|-------------------|--------|---------------------|------------------------|--------------------------------|------------------|-----------------------|
| 54 | N132 | E5 | 1.00 | 0.18 | 0.77 | 6.27 | 0.24 | 22.48 |
| 50 | N132 | S36 | 1.00 | 0.15 | 0.81 | 6.58 | 0.25 | 26.31 |
| 28 | N55 | R8 | 1.00 | 0.22 | 0.74 | 6.00 | 0.24 | 19.89 |
| 35 | P86 | E22 | 1.00 | 0.18 | 0.76 | 5.45 | 0.23 | 20.12 |
| 36 | P86 | S36 | 1.00 | 0.15 | 0.77 | 5.98 | 0.25 | 26.21 |
| 58 | K60 | N132 | 1.00 | 0.17 | 0.79 | 5.19 | 0.19 | 37.66 |
| | Free dye | - | 1.00 | 19.15 | 0.34 | 6.00 | 0.38 | 27.66 |

Label # corresponds to the sample number assigned in Supplementary table 4. FCS was measured at 635 nm excitation and power of 5.6 kW/cm² and the average of parallel-perpendicular and perpendicular-parallel correlation curves fitted to the model function as described in eq. 4.16. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z₀/w₀ describes the shape of the observation volume, A_{iso} is the fitted amplitudes and t_{iso} the relaxation times of the cis-trans isomerization process. The experimental conditions are described in the methods section.

Supplementary Table 4.7. Fit results for the FCS curves of Alexa647 labels in the substrate binding study

| Alexa647 label | Other mutations | substrate | offset | N _{bright} | t _{diff} (ms) | z ₀ /w ₀ | A _{iso} | t _{iso} (μs) |
|----------------|-----------------|-----------|--------|---------------------|------------------------|--------------------------------|------------------|-----------------------|
| I150 | E11A, S44Amber | No | 1 | 0.15 | 1.26 | 6.1 | 0.25 | 25.51 |
| I150 | E11A, S44Amber | Yes | 1.14 | 0.25 | 1.37 | 6 | 0.24 | 39.64 |

Fit results for the FCS curves of Alexa647 labels at position I150 and functional mutation E11A in presence and absence of substrate peptidoglycan. E11A corresponds to Sample # 61 from Supplementary table 4. FCS was measured at 635

nm excitation and power of 0.7 kW/cm² and the average of parallel-perpendicular and perpendicular-parallel correlation curves fitted to the model function described in eq. 4.16. N_{bright} is the effective number of molecules in the bright state, t_{diff} is the diffusion time, z_0/w_0 describes the shape of the observation volume, A_{iso} is the fitted amplitudes and t_{iso} the relaxation times of the cis-trans isomerization process.

Supplementary Notes

Supplementary Note 4.1. Calculation of trapped dark surface

Using AV simulations, the Lysines and Arginines within 4 Å of the accessible volume of the dye was identified for both the C₁ and C₂ conformations. Lys and Arg are referred to as the “trapping positions” due to their high affinity for interaction.

Supplementary Table 4.8. Trapping positions within 4 Å of the dye AV in the C₁ conformation

| Alexa488 label | Lys/Arg within 4 Å in the C ₁ state | | | | | | | | | Total Nr. |
|----------------|--|------|------|------|------|------|------|------|------|-----------|
| E22 | K19 | K35 | R14 | R137 | R145 | | | | | 5 |
| S36 | K19 | K35 | K43 | K48 | R52 | | | | | 5 |
| K60 | K16 | K65 | K162 | R8 | R14 | R52 | | | | 6 |
| Q69 | K35 | K48 | R52 | K65 | R76 | R80 | R96 | | | 7 |
| K19 | K16 | K35 | K43 | R148 | R145 | | | | | 5 |
| S44 | K35 | K43 | K48 | R52 | | | | | | 4 |
| P86 | R76 | R80 | K83 | K85 | R95 | R96 | R119 | K124 | R125 | 9 |
| R119 | R80 | K83 | K85 | K124 | R125 | K135 | | | | 6 |
| D127 | R95 | R119 | K124 | R125 | K135 | K147 | R154 | | | 7 |
| N132 | K83 | R119 | K124 | R125 | K135 | R137 | R154 | | | 7 |
| I150 | K83 | R119 | K124 | R125 | K135 | R137 | R154 | | | 7 |

Supplementary Table 4.9. Trapping positions within 4 Å of the dye AV in the C₂ conformation

| Alexa488 label | Lys/Arg within 4 Å in the C ₂ state | | | | | | | | | Total Nr. |
|----------------|--|------|------|------|------|------|------|------|--|-----------|
| E22 | K19 | K35 | R137 | | | | | | | 3 |
| S36 | K19 | K35 | K43 | K48 | R137 | | | | | 5 |
| K60 | R8 | R14 | K16 | R52 | K65 | | | | | 5 |
| Q69 | R8 | K35 | K48 | R52 | K65 | R76 | R80 | | | 7 |
| K19 | R8 | R14 | K16 | K35 | K43 | R137 | R148 | | | 7 |
| S44 | K35 | K43 | K48 | R52 | | | | | | 4 |
| P86 | R76 | R80 | K83 | K85 | R96 | R119 | K124 | R125 | | 8 |
| R119 | K83 | K85 | K124 | R125 | K135 | | | | | 5 |
| D127 | R95 | R119 | K124 | R125 | K135 | K147 | R148 | R154 | | 8 |
| N132 | K83 | R119 | R125 | K135 | R137 | R154 | | | | 6 |

| | | | | | | | | | | |
|------|------|------|------|------|--|--|--|--|--|---|
| I150 | K135 | K147 | R148 | R154 | | | | | | 4 |
|------|------|------|------|------|--|--|--|--|--|---|

For each individual Lys and Arg, PET quenchers within 4 Å was identified. It was seen that the PET quenchers within 4 Å of a given Lys or Arg are the same in both C₁ and C₂ conformations.

Supplementary Table 4.10. PET quenchers near trapped positions in both C₁ and C₂ conformations

| Lysines/Arginines | Names of quenchers within 4[Å] |
|-------------------|--------------------------------|
| K16 | Y18 |
| K19 | Y18, Y24, Y25 |
| K35 | Y24 |
| K43 | None |
| K48 | None |
| K60 | None |
| K65 | None |
| K83 | None |
| K85 | Y88 |
| K124 | M120, W126 |
| K135 | None |
| K147 | Y139 |
| K162 | W158, Y161 |
| R8 | M6 |
| R14 | Y18 |
| R52 | None |
| R76 | None |
| R80 | None |
| R95 | W126 |
| R96 | Y88 |
| R119 | M120 |
| R125 | M120, W126 |
| R137 | W138, Y139 |
| R145 | None |
| R148 | Y161 |
| R154 | W126 |

From the above table, we can see which of the Lys/Arg positions have atleast one quencher within 4 Å (adjacent area). A dye stuck at these positions has a higher probability to be quenched. Then the fraction of trapped dark surface is calculated from

$$\text{eq. 4.1: } x_{\text{trapped,dark}} = \frac{\text{Nr. of Q adjacent Lys or Arg near AV}}{\text{Nr. of Lys or Arg near AV}}$$

Supplementary Table 4.11. Calculation of trapped dark surface for C₁ and C₂ conformations and their weighted average.

| Alexa488 label | Nr. of trapped positions near dye | | Nr. of quencher-adjacent trapped positions near dye | | $x_{\text{trapped,dark}}$ | | |
|----------------|-----------------------------------|----------------|---|----------------|---------------------------|----------------|------------------|
| | C ₁ | C ₂ | C ₁ | C ₂ | C ₁ | C ₂ | weighted average |
| E22 | 5 | 3 | 4 | 3 | 0.80 | 1.00 | 0.89 |
| S36 | 5 | 5 | 2 | 3 | 0.40 | 0.60 | 0.49 |
| K60 | 6 | 5 | 4 | 3 | 0.67 | 0.60 | 0.64 |
| Q69 | 7 | 7 | 2 | 2 | 0.29 | 0.29 | 0.29 |
| K19 | 5 | 7 | 3 | 6 | 0.60 | 0.86 | 0.72 |
| S44 | 4 | 4 | 1 | 1 | 0.25 | 0.25 | 0.25 |
| P86 | 9 | 8 | 6 | 5 | 0.67 | 0.63 | 0.65 |
| R119 | 6 | 5 | 3 | 3 | 0.50 | 0.60 | 0.55 |
| D127 | 7 | 8 | 6 | 7 | 0.86 | 0.88 | 0.87 |
| N132 | 7 | 6 | 5 | 4 | 0.71 | 0.67 | 0.69 |
| I150 | 7 | 4 | 5 | 3 | 0.71 | 0.75 | 0.73 |

The weights are the fraction of open and closed (including compact) states for the T4L wt* determined previously by Sanabria et al.³. Weighted average of $x_{\text{trapped,dark}} = (0.54 * x_{\text{trapped,dark}}(C1)) + (0.46 * x_{\text{trapped,dark}}(C2))$

Supplementary Note 4.2. Calculating Φ_F for a series of temperatures

For the determination of fluorescence quantum yield of Alexa488 labels at a range of temperatures from 11°C to 44°C, at first the fluorescence intensities at 520 nm emission was measured at an excitation of 485 nm.

Supplementary Table 4.12. Measured fluorescence intensity at 520nm emission for Alexa488 labels at temperatures from 11°C to 44°C when excited at 485 nm.

| T | Fluorescence intensity (a.u.) at 520 nm, $F_{520}^{T^\circ\text{C}}$ | | | | | |
|-----------|--|--------------|--------------|--------------|--------------|-----------------|
| [°C] | P86C | R119C | D127C | N132C | I150C | Alexa488 |
| 11 | 16041317 | 15367226 | 13957341 | 18131650 | 13806152 | 24900509 |
| 15 | 16410097 | 15298145 | 13905215 | 18151396 | 13924999 | 24631524 |
| 19 | 17079680 | 15316162 | 13911038 | 18261736 | 13968596 | 25419476 |
| 23 | 17593382 | 15370934 | 13851696 | 18225571 | 14241184 | 25501367 |
| 27 | 17183657 | 15437148 | 13947645 | 18464957 | 14519915 | 25454423 |
| 31 | 17259223 | 15594141 | 13891722 | 18513190 | 13786136 | 25426918 |
| 35 | 17056977 | 15587857 | 13602229 | 18429354 | 12957114 | 25201431 |

| | | | | | | |
|-----------|----------|----------|----------|----------|----------|----------|
| 39 | 15584192 | 15266411 | 12839434 | 18286451 | 10111314 | 24885994 |
| 44 | 11729957 | 13840126 | 11017298 | 17321462 | 4096617 | 25026799 |

The steady state fluorescence quantum yields for all the labels at 20°C were measured as described in literature³³⁶ and calculated according to equations (4.11) and (4.12). It was then assumed that $\Phi_F^{20^\circ\text{C}} = \Phi_F^{19^\circ\text{C}}$. Now that we have $\Phi_F^{19^\circ\text{C}}$, the quantum yields at the other temperatures were determined from their respective fluorescence intensities at 520 nm as:

$$\Phi_F^{T^\circ\text{C}} = \Phi_F^{19^\circ\text{C}} * \left(\frac{F_{520}^{T^\circ\text{C}}}{F_{520}^{19^\circ\text{C}}} \right) \quad (4.19)$$

where $\Phi_F^{T^\circ\text{C}}$ is the fluorescence quantum yield at temperature, T°C and $F_{520}^{T^\circ\text{C}}$ is the fluorescence intensity at 520 nm at temperature, T°C.

Supplementary Table 4.13. Calculated fluorescence quantum yield of Alexa488 labels.

| T | Steady state fluorescence quantum yield, Φ_F | | | | | |
|-----------|---|-------------|-------------|-------------|-------------|-------------|
| [°C] | P86C | R119C | D127C | N132C | I150C | Alexa488 |
| 11 | 0.79 | 0.69 | 0.72 | 0.67 | 0.74 | 0.94 |
| 15 | 0.81 | 0.68 | 0.72 | 0.67 | 0.74 | 0.93 |
| 19 | 0.84 | 0.68 | 0.72 | 0.68 | 0.75 | 0.96 |
| 23 | 0.87 | 0.69 | 0.72 | 0.67 | 0.76 | 0.96 |
| 27 | 0.85 | 0.69 | 0.72 | 0.68 | 0.78 | 0.96 |
| 31 | 0.85 | 0.70 | 0.72 | 0.68 | 0.74 | 0.96 |
| 35 | 0.84 | 0.70 | 0.71 | 0.68 | 0.69 | 0.95 |
| 39 | 0.77 | 0.68 | 0.67 | 0.68 | 0.54 | 0.94 |
| 44 | 0.58 | 0.62 | 0.57 | 0.64 | 0.22 | 0.95 |

Chapter 5. Developing dye models for accurate FRET label description

Authors: Aiswaria Prakash¹, Christian A. Hanke¹, Olga Doroshenko¹, Mykola Dimura^{1,2}, Thomas-Otavio Peulen¹, Holger Gohlke², Claus A.M. Seidel¹

Affiliations:

¹Chair for Molecular Physical Chemistry, Heinrich Heine University, Düsseldorf, 40225 Düsseldorf, Germany

²Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University, Düsseldorf, 40225 Düsseldorf, Germany

Contributions: M.D., T.O.P initiated and designed the project. C.A.M.S and H.G supervised the project. M.D., T.O.P, C.A.H., C.A.M.S, H.G and A.P. were involved in the review of dye models, including Fig. 5.1. O.D. and T.O.P measured and analyzed FRET samples for the RNA calibration study. C.A.H. performed ACV and hybrid AV modeling and related calculations. A.P. measured the residual anisotropy used in dye modeling. A.P. designed the figures 5.2 and 5.3 and composed the chapter.

Section 5.1.1 of this chapter is an updated extract from the published paper:

Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M., Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Current Opinion in Structural Biology* **2016**, 40, 163-185

5.1 Introduction

Förster resonance energy transfer (FRET)^{18,345} is a powerful approach to study the conformational dynamics of biomolecules since it can measure structural dynamics over 12 orders of magnitude^{159,265}. This ratiometric approach involves non-radiative energy transfer from a high-energy fluorophore (donor, D) to a lower-energy fluorophore (acceptor, A). When the fluorophores are present within the FRET range, dipole-dipole coupling can occur, leading to energy transfer with efficiency, $E \propto (R_{DA})^{-6}$, where R_{DA} is the interdye distance. Since the fluorophores are usually attached to the biomolecule using long flexible linkers to allow free dye rotation, there are significant uncertainties in the dye position. Therefore, to arrive at accurate protein structural information from interdye distances, spatial distribution of dye positions must be included^{103,239,242} in the dye model description.

5.1.1 A review on the dye description of organic fluorescence labels

(This section is an updated excerpt from a co-authored paper, “Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems” published in the journal, *Current Opinion in Structural Biology* by Dimura et al.²⁴²).

Both EPR spectroscopy with the double electron-electron resonance (DEER) method³⁴⁶ and FRET spectroscopy are frequently used to obtain structural information on proteins by selectively coupling labels via flexible linkers (i.e., nitroxide spin-labels or donor and acceptor dyes, respectively). Thus, the labels can explore a large conformational space, such that a distribution of inter-label distances rather than a single inter-label distance is observed experimentally. In structural analysis, both techniques face the same problem of describing the localization of their respective labels and connecting the measured distances with structural information on the protein. This is a typical inverse problem, where both EPR³⁴⁶⁻³⁴⁹ and FRET^{239 241 350} spectroscopies have developed similar solution strategies of assuming a structural model to compute a specific distance distribution between the labels and comparing the simulated and experimental values. Analogous to super-resolution microscopy, where the knowledge of the optical point spread function allows for localizing single emitters far beyond the optical resolution limit,

a detailed knowledge of the label distribution for a given biomolecular structure is required to maximize the structural resolution of the respective technique.

All-atom molecular dynamics (MD) simulations^{4,150,340,351-355} seem to be the most appropriate solution to describe the label localization in integrative modeling (Figure 1A-C). However, while the relevant time-scale of fluorescence and biomolecular dynamics covers picoseconds to hours, current MD simulations reach only the micro- to millisecond timescale. Furthermore, MD simulations are time-consuming and limited in their capability to sample the conformational space, in particular since the dye distribution changes with conformational changes of the biomolecule. MD simulations^{351,353,356} as well as experimental fluorescence^{103,357,358} and EPR³⁵⁹ studies find long-lived conformational states, where the labels are trapped on the biomolecular surface. Trapping of cyanine dyes is known to change their fluorescence quantum yields³⁵⁸ such that this process is utilized by others for probing protein-nucleic acid interactions by protein-induced fluorescence enhancement (PIFE)^{135,220}. Now, a more general term, association-induced fluorescence enhancement (AIFE) has been introduced, where the fluorescence enhancement is due to dye association with any biomolecule. To avoid quenching and mobility reduction by the surface effects, most commercially available dyes used for FRET experiments have long flexible linkers (~ 20 Å), and the fluorophore carries sulfonic acid groups for improved solubility. The large flexibility of the linker also ensures a high orientational mobility of the dyes such that the assumption of the isotropic average for the dipole orientation factor (i.e., $\kappa^2 = 2/3$) is feasible, which is an important factor determining the accuracy of distance calculations by FRET measurements. In conclusion, the atomistic treatment of dye labels, together with the biomolecular system by MD simulations, is helpful for analyzing protein label interactions. However, it is currently not feasible to use MD simulations as a rapid routine method to predict the conformational sampling of the labels, which is needed for finding the structural model within a large ensemble that is most consistent with FRET data, as shown below.

Computationally faster algorithms use a coarse-grained representation, e.g. the label is approximated by a flexibly linked sphere^{348,360} (see Figure 1a-c), to compute the sterically allowed accessible volume (AV), which reproduced MD simulations well¹⁰³. The use of a geometric search algorithm reduces the computational time for the AV of a single FRET-

label to less than 50 milliseconds on a conventional desktop computer²³⁹. The computed AV represents uniform population densities ρ_{Dye} of the spatial dye localizations (Figure 1D). The mesh represents the surface of the dye positions most distant from the attachment point and includes all sterically allowed dye positions. To compute fluorescence observables, the spatial population densities $\rho_{\text{Dye}}(x,y,z)$ of the dye localizations is needed.

Two convenient assumptions for limiting cases were initially used for the models describing the spatial and orientational dye density in various hybrid-FRET modeling approaches: (1) The Nano Positioning System (NPS) assumes a label adopting a single position with a defined conformation of the linker and the fluorophore wobbling in a cone (Localized antenna/Single cone model)³⁶⁰. This means that all fluorophores of an ensemble of molecules are located at the same position (defined by a delta function δ) with respect to the macromolecule: $\rho_{\text{Dye}}(x,y,z) = \delta(x_{\text{Dye}},y_{\text{Dye}},z_{\text{Dye}})$. Since the position $(x_{\text{Dye}},y_{\text{Dye}},z_{\text{Dye}})$ of the dye is initially unknown multiple measurements are used to localize it within a credible volume for a given confidence level (Figure 1d, left). (2) As an approximation of a freely diffusing fast rotating dye all orientations and all positions within the AV are equally populated; $\rho_{\text{Dye}}(x,y,z)$ is constant and the fluorophore is isotropic oriented within the AV (Free diffusion/Iso model)^{239 103}. The effect of different dye models on the spatial population densities $\rho_{\text{Dye}}(x,y,z)$ along the vector R , pointing from the linker attachment site to the mean dye position in the AV, is visualized in the lower panel of Figure 1D.

More recently, other dye models were suggested^{90,243,361}. To consider small fractions of fluorescent fluorophores trapped on the biomolecular surface⁹⁰, the classical AV description (using the Free diffusion/Iso model) was extended by defining an additional interacting surface layer (the contact volume of defined thickness is highlighted in violet in Figure 1D) with a preferential surface residence of the dye, which may be approximated by a step function. This surface preference is specific for the dye and the labeling site. This combination yields the Free diffusion + Contact/Iso model for an accessible and contact volume (ACV), which predicts the compaction of the spatial dye distribution due to trapped dye fractions without significant increase of the computational cost. Höfig et al.²⁴³ introduced a weighted AV description, where the densities ρ_{Dye} are enriched by

empirical weighting with a Gaussian function leading to a higher local dye surface preference but only close to the attachment point (Gaussian-chain (GC) diffusion/Iso model) (Figure 1D).

The comparison of the normalized densities ρ_{Dye} of the four dye models in Figure 1D shows that the mean dye positions of an ACV or a weighted AV are closer to the protein interface as compared the AV with a uniform density. This dye model specific shift of the mean positions becomes important for the predicted interdye distances if the other dye of the FRET pair is attached on the opposite side of the protein. To avoid an arbitrary choice of dye models, it is important to define rational criteria based on experimental evidence through observable fluorescence parameters. For example, time-resolved fluorescence data map the translational and orientational dye diffusion and the resulting DA distance distributions by fluorescence quenching (studied in Chapters 2+3), anisotropy (studied in Chapters 2+3) and FRET.

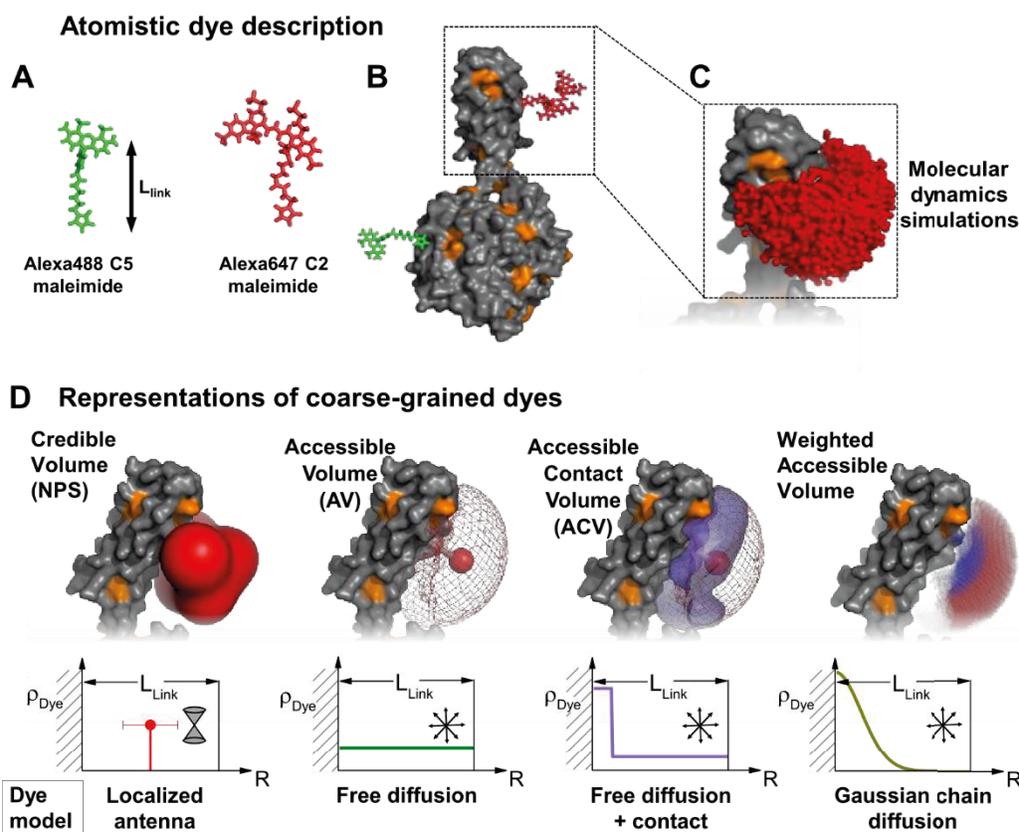


Figure 5.1. Description of fluorescent labels in simulation. Treatment of dye labels in FRET-based modeling. (a) The donor and acceptor dyes, for example, Alexa488 C5 maleimide (green) and Alexa647 C2 maleimide (red), respectively, are typically attached to the

biomolecule via long (~20 Å) flexible linkers. (b) Dye labels attached to the protein Atlastin-1 (grey surface) to give an impression of the sizes. Quenching amino acids on the protein surface are highlighted in orange. (c) Molecular dynamics simulations provide the spatial distribution of dye molecules. (d) Representation of coarse-grained dye labels. On the top different dye representations are shown. On the bottom the spatial density ρ_{Dye} along a vector R starting at the attachment point in the direction of the dye mean position is shown for the corresponding dye models. The original Nano Positioning System (NPS)³⁶⁰ assumes the accessible volume as prior information (uncertainty), which is reduced by a set of FRET measurements with fixed satellites resulting in an uncertainty distribution (red) of a putatively fixed localized dye (antenna). The orientation of the dyes follows a diffusion in a cone model highlighted by a pictogram. Following R, the dye is located at a specific position (vertical red line) with an uncertainty. Accessible volume (AV, mesh) models provide the sterically allowed space of the dye molecule attached to the protein as calculated by the FPS program²³⁹. Here, the linkers of Alexa488 and Alexa647 are approximated as flexible tubes with width of $L_{width} = 4.5$ Å and length of $L_{link} = 20.5$ Å and $L_{link} = 21.0$ Å, respectively. Three radii were used to describe the dyes (5.0 Å, 4.5 Å, 1.5 Å) and (11.0 Å, 4.7 Å, 1.5 Å) for Alexa488 and Alexa647, respectively. The large sphere indicates the mean dye position. For a dye freely diffusing inside of its AV a uniform spatial distribution is assumed. The accessible and contact volume (ACV,⁹⁰) provides a similar description as the AV, but defines an area close to the surface as contact volume (violet). In this figure the density ρ_{Dye} in the contact volume is weighted six times stronger and defined as part of the AV that is closer than 3 Å from the macromolecular surface. For the corresponding model, where a dye freely diffuses within the AV and its diffusion is hindered close to the surface, the spatial density ρ_{Dye} along R is approximated by a step function: $\rho_{Dye}(R < 3 \text{ Å}) = 6 \cdot \rho_{Dye}(R \geq 3 \text{ Å})$. The weighted accessible volume is a modification of AV where ρ_{Dye} along R is approximated by a Gaussian chain-inspired empirical weighting function²⁴³. To illustrate the effect of the proposed weighting of ρ_{Dye} the weighted AV is colored from blue (high density) to red (low density). (Adapted from Fig. 1 of Dimura et al.²⁴²)

5.1.2 Prior experimental input and its role in FRET experiments

Prior experimental information (such as the 3D structure of a protein) are helpful in designing FRET experiments and guiding labeling strategies so that informative sites can be recognized, those labeling sites can be avoided where mutations for dye coupling can affect the protein function or stability³⁶²⁻³⁶⁵ or where the dye is likely to be highly quenched¹⁰⁴ due to the characteristics of the local environment (Chapter 3+4). Additionally, insights from previous structural studies are valuable for mapping FRET label positions so that new experiments can be designed to either confirm prior structural knowledge (from X-ray crystallography, NMR etc.) or invalidate a previous assumption by generating a three dimensional structural model (posterior hypothesis)²⁴².

Prior information can not only help in designing informative labels, but also in better label description. For example, residual anisotropy from time resolved decay experiments is used to define the increased dye occupancy within the contact volume in the Free

diffusion + Contact/Iso model²⁴². In nucleic acids, it is known that dye labeling at the terminal base^{92,94-96} or even 4-6 base pairs from the terminus (Chapter 3) can cause the dye to stack on the terminal base similar to base stacking. Knowledge of such specific interactions which significantly alter the dye distribution can be incorporated into the label description for FRET studies. To this end, we compare FRET donor-acceptor (DA) distances from experiments and simulations on a double stranded RNA (dsRNA) as a model system. When there are significant deviations between measured and calculated distances, we apply our prior knowledge of position-dependent dye behavior to better describe the dye distribution by introducing the concept of a hybrid AV.

5.2 Results and Discussion

In this study, the suitability of the Free diffusion + Contact/Iso model²⁴² to describe the distribution of donor and acceptor dyes is investigated and a new dye model, called the **Hybrid AV model**, is developed for labels with highly specific localized interactions to improve their dye description. FRET distance calibration studies are carried out on dsRNA molecules with Alexa488 as the donor dye and Cy5 as the acceptor dye. FRET distance calibration with single FRET pairs was carried out for 24 donor-acceptor (DA) combinations (Fig. 5.2A, Table 1). The chosen label positions belong to distinct RNA environments (as outlined in Chapter 3) and the FRET pairs studied here fall under 4 categories: (1) the acceptor is close to the 3' end and the donor is in the internal regions of the helix, away from the ends (Fig. 5.2B, filled magenta circles) (2) the donor is close to the 3' end and the acceptor is internal (Fig. 5.2B, magenta circle outlines) (3) both the donor and acceptor dyes are in the internal regions (Fig. 5.2B, filled blue circles) (4) both the donor and acceptor dyes are close to the 3' ends on either ends of the helix (Fig. 5.2B, magenta circle with disconnected outline). The dye linker attachment on dsRNA occurs at the C5 atom for uracil and cytosine and the N2 atom for guanine.

5.2.1 dsRNA rulers: Experimental vs predicted distance $\langle R_{DA} \rangle_E$

The double stranded RNA is used for distance calibration since it is a fairly rigid system whose structure is well-defined such that there is least uncertainty arising from structural considerations. Five dsRNA molecules are studied here (Fig. 5.2A). Of these, the molecules *extended I₁* and *extended I₂* are similar in sequence to *I₁* and *I₂*, respectively;

but with 4 extra base pairs added at one of the helix termini, thus “extending” the molecule. The extended sequences allow comparison of dye labeling sites (e.g. c29 and b33) with the same immediate surrounding primary sequence, but different end proximities. They are used as FRET rulers in single-molecule FRET (smFRET) spectroscopy studies where the distances between the dyes are varied and energy transfer efficiencies obtained. The FRET-averaged donor-acceptor distance $\langle R_{DA} \rangle_E$ corresponding to the FRET efficiency is calculated from experiments and compared with theoretical predictions.

Table 5.1. Experimental and calculated FRET distance, $\langle R_{DA} \rangle_E$. For measured residual anisotropies, see Chapter 3.

| FRET pair | Molecule | Position from end | $\langle R_{DA} \rangle_{E,exp}$ Å | $\langle R_{DA} \rangle_{E,cal}$ Å |
|---------------|----------------|----------------------|---------------------------------------|---------------------------------------|
| c29(D)-d15(A) | l_1 | (D) near 3'-end | 49.2 | 51.8 |
| c29(D)-d18(A) | l_1 | (D) near 3'-end | 50.4 | 51.2 |
| c29(D)-d20(A) | l_1 | (D) near 3'-end | 57.5 | 59.6 |
| c29(D)-d23(A) | l_1 | (D) near 3'-end | 63.3 | 68.1 |
| b14(D)-c8(A) | l_2 | Internal label | 42.9 | 41.4 |
| b11(D)-c8(A) | l_2 | Internal label | 50.4 | 49.6 |
| b8(D)-c8(A) | l_2 | Internal label | 58.0 | 59.1 |
| b5(D)-c8(A) | l_2 | Internal label | 61.2 | 64.2 |
| a26(D)-b27(A) | l_2 | Internal label | 57.4 | 54.0 |
| a26(D)-b28(A) | l_2 | Internal label | 60.1 | 55.4 |
| a26(D)-b33(A) | l_2 | (A) near 3'-end | 73.7 | 71.4 |
| c16(D)-b33(A) | l_2 | (A) near 3'-end | 54.2 | 49.6 |
| a18(D)-b33(A) | l_2 | (A) near 3'-end | 61.8 | 53.3 |
| a21(D)-b33(A) | l_2 | (A) near 3'-end | 57.8 | 54.4 |
| a23(D)-b33(A) | l_2 | (A) near 3'-end | 70.2 | 62.1 |
| c29(D)-b27(A) | l_3 | (D) near 3'-end | 61.9 | 64.3 |
| c29(D)-b28(A) | l_3 | (D) near 3'-end | 65.4 | 67.7 |
| c29(D)-b33(A) | l_3 | (D), (A) near 3'-end | 84.7 | 82.7 |
| c29(D)-d15(A) | Extended l_1 | Internal label | 51.0 | 50.8 |
| c29(D)-d18(A) | Extended l_1 | Internal label | 50.1 | 48.4 |
| c29(D)-d23(A) | Extended l_1 | Internal label | 66.8 | 65.8 |

| | | | | |
|---------------|-------------------------|----------------|------|------|
| c16(D)-b33(A) | Extended I ₂ | Internal label | 53.7 | 51.0 |
| a18(D)-b33(A) | Extended I ₂ | Internal label | 57.1 | 53.1 |
| a23(D)-b33(A) | Extended I ₂ | Internal label | 62.3 | 61.7 |

Experimental and calculated FRET distance, $\langle R_{DA} \rangle_E$ for all measured dsRNA samples in the distance calibration study. (D) and (A) denotes the donor and acceptor labels, respectively. The corresponding sequences and labeling positions are given in Figure 5.2A.

From experiment: Single-molecule confocal experiments are performed on the FRET labeled dsRNA molecules, and the single molecules are detected as photon bursts (of a certain intensity and duration) during their free diffusion through the confocal volume. From the collected photon bursts, a variety of fluorescence information can be extracted due to single-molecule multiparameter fluorescence detection (smMFD)^{15,16,366,367}. Photon Distribution Analysis (PDA) is performed to eliminate the photon shot-noise contribution and examine the heterogeneity of the sample^{337,357,368}. A quantitative analysis of the FRET efficiency histogram is carried out using an appropriate model that accounts for shot noise, background contributions and additional broadening due to complex acceptor photophysics and fits one or two states described by Gaussian distributions (assigned to a FRET population and a donor only population). The distance corresponding to the FRET populations is referred to as the experimental/measured donor-acceptor distance, $\langle R_{DA} \rangle_{E,exp}$.

From predictions: Since the RNA structure is known, the DA distances can be predicted from modeling the dye distributions. We use the previously described Free diffusion+Contact/Iso model²⁴² to describe the accessible and contact volume (ACV) of the dyes from the dye radii and the lengths and widths of the linkers used for labeling. Due to interactions between the fluorophore and the biomolecule, the dye is less mobile near the surface of the biomolecule. Ensemble fluorescence anisotropy measurements²³⁹ were employed to calculate the residual anisotropy r_∞ . From this, the fraction of the fluorophore trapped on the surface of the protein (s) is determined¹⁰³:

$$s = \sqrt{\frac{r_\infty}{r_0}} \quad (5.1)$$

with the fundamental anisotropy $r_0=0.38$. The measurements were performed on single labeled dsRNA molecules and analyzed as described in Chapter 3. From the trapped

fraction and the contact volume (defined as the volume of AV within 2Å from the RNA surface), the spatial dye densities are calculated³.

Distances between random points in the donor and acceptor ACVs are calculated and converted to efficiencies using eq. 5.2. These FRET efficiencies are averaged over the distribution and converted back into to the DA distance, referred to as the predicted or calculated distance, $\langle R_{DA} \rangle_{E,cal}$.

The interdye distance is related to the FRET efficiency as:

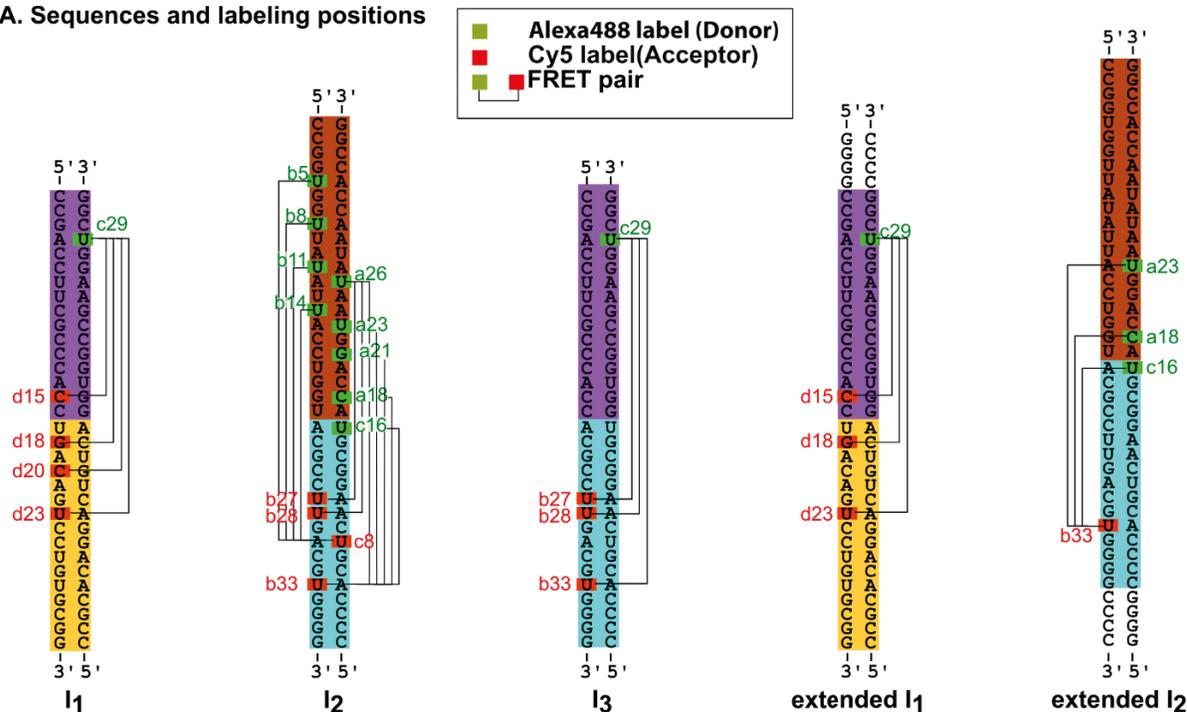
$$E = \frac{1}{1 + \left(\frac{R_{DA}}{R_0}\right)^6} \quad (5.2)$$

where R_0 is the Förster radius and R_{DA} is the donor-acceptor distance.

It may be noted that $\langle R_{DA} \rangle_E$ calculated here, is different from R_{mp} (distance between mean dye positions) and $\langle R_{DA} \rangle$ (measured by ensemble TCSPC)¹⁰³.

Comparison: When both donor and acceptor label are in the internal regions of the helix (Fig. 5.2B, blue circles), there is good agreement between experimental and calculated distances. However, for the acceptor (Cy5) labels close to the 3'-end, there is a marked deviation in $\langle R_{DA} \rangle_E$ (Fig. 5.2B, filled magenta circles). This originates from their ability to reach the exposed terminal bases of the helix and therefore form strong π -stacking interactions there (Chapter 3 and others^{92,94,95}). π -interactions have been studied not only for dyes labeled on the terminal bases^{94,95}, but also for dyes labeled a few base pairs away from the helix terminus, where effective stacking interactions depend on the length of linker, attachment atom in the RNA base and the kind of helix end (5' or 3'). These π -stacking interactions significantly alter the dye distribution and the dye density is the highest in the region close to the terminal base (Chapter 3). To describe these 3'-end labels, the coarse-grained Free diffusion+Contact/Iso label description is insufficient because the dye densities are more concentrated near the terminal bases than over the accessible nucleic acid surface.

A. Sequences and labeling positions



B. Correlating experimental and calculated FRET distances

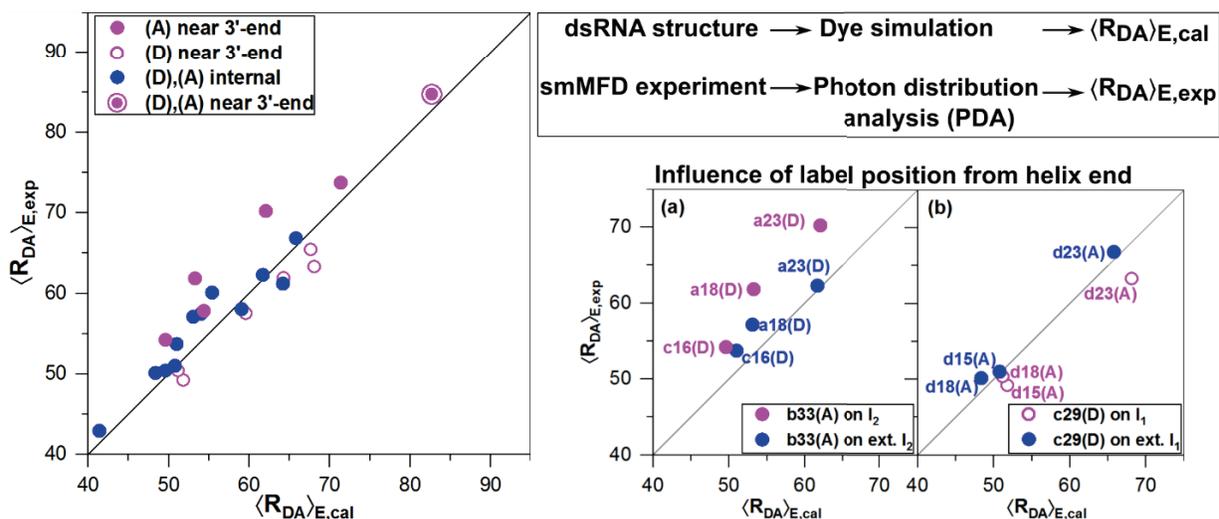


Figure 5.2. FRET distance calibration study. (A) dsRNA sequences and donor(D)-acceptor(A) labeling positions, with donor label in green and acceptor label in red. The 24 single FRET pairs measured in the study are displayed with connecting lines between D and A. The names of the labeling positions are derived from their corresponding label names in the four-way junction molecule studied in Chapter 3. (B) FRET-averaged distances from the Free diffusion+Contact/Iso (ACV) dye model calculations, $\langle R_{DA} \rangle_{E,cal}$, are plotted as a function of FRET-averaged distances from the single molecule MFD experiments, $\langle R_{DA} \rangle_{E,exp}$ showing all measured FRET pairs from panel (A). From this plot, a set of selected points are reexamined in the bottom right plot to highlight the effect of acceptor and donor label proximity to the helix end. (a) The acceptor b33(A) has the same neighboring base pairs sequence in molecules I₂ (magenta filled circle) and extended I₂ (blue), but differs in its proximity to the 3'-end. (b) The

donor c29(D) on molecules I_1 (magenta circle outline) and extended I_1 (blue), differs in the proximity to the 3'-end. A brief flow chart depicting the main steps for obtaining experimental and calculated $\langle R_{DA} \rangle_E$ are also given. For detailed explanation of the steps, see sections 2.1 and 4.

Comparison of same labeling positions in different molecules: The molecule I_2 has the same primary sequence as the molecule *extended* I_2 , and differs only by 4 additional base pairs at the dye labeled end on *extended* I_2 . This alters the proximity of the acceptor b33(A) to the 3'-end, bringing about drastic changes in its photophysical properties. Comparing FRET samples with the same donor labels and b33(A) acceptor label on the molecules I_2 and *extended* I_2 (Fig. 5.2B(a)), we observe deviations between the experimental and calculated distances when b33(A) is on molecule I_2 , where π -stacking with the terminal bases is possible. Meanwhile, the FRET samples with b33(A) on *extended* I_2 show very good agreement between experimental and calculated $\langle R_{DA} \rangle_E$, since the acceptor dye is further away from the helix end and cannot stack with the terminal bases.

A similar comparison of the Alexa488 donor at position c29(D) on molecules I_1 and *extended* I_1 (Fig. 5.2B(b)) yields rather similar agreement between experimental and calculated $\langle R_{DA} \rangle_E$ whether the dye was in the proximity of the 3'-end (as in I_1) or further away from it (as in *extended* I_1). Although donor (Alexa488) dyes labeled near the 3'-end are known to undergo π -stacking at the terminal base (Chapter 3 and Neubauer et al.)⁹⁶ like an additional base pair, the altered dye distribution due to this interaction is not expected to affect the agreement between $\langle R_{DA} \rangle_{E,cal}$ and $\langle R_{DA} \rangle_{E,exp}$ in the same way as for the Cy5 dye. This is because the stacked species are non-fluorescent due to static PET quenching with the terminal guanines and do not contribute to the observed donor signal in MFD experiments. Nevertheless, dynamic quenching interactions occurring between the Alexa488 dye and the exposed terminal guanine is expected to bias the dye distribution towards the terminus.

Improving the estimation of trapped dye fraction: From the correlation between the residual anisotropies and lifetimes of Cy5 labels studied in Chapter 3 (Fig.3.2), we can arrive at a special case of a covalently attached non-interacting dye. That is, a dye which is covalently attached to the RNA, but otherwise free, not being trapped at the RNA surface. The residual anisotropy of such a species is obtained from the linear fit of the

correlation extrapolated to the free dye lifetime (denoted as $r_{\infty,extrapol.}$). Using this to modify eq. 5.1 gives a more accurate estimation for trapped dye fraction since it accounts for the cone angle as described in a wobbling in a cone (WIC) model¹⁰³. The improved trapped dye fraction (s') is then:

$$s' = \sqrt{\frac{r_{\infty,meas.} - r_{\infty,extrapol.}}{r_0}} = \sqrt{\frac{\Delta r_{\infty}}{r_0}} \quad (5.3)$$

where Δr_{∞} is the difference in the residual anisotropy of the measured sample ($r_{\infty,meas.}$) and the residual anisotropy of the covalently attached non-interacting dye ($r_{\infty,extrapol.}$) determined by extrapolation (*vide supra*).

5.2.2 Improving dye model based on prior knowledge

It is well documented that rhodamine and cyanine dyes labeled close to or at the terminal base undergo strong π -stacking interactions with the exposed terminal bases^{92,94-96}. For the cyanine labels, these interactions result in association-induced fluorescent enhancement (AIFE) and a corresponding increase in residual anisotropy (r_{∞}). For a typical Cy5 label coupled at a nucleobase in the internal regions of the RNA double helix, the residual anisotropy was measured to be 0.18 (Chapter 3). However, when Cy5 is attached to a position close to the helix end, the time resolved anisotropy measurements yield $r_{\infty}=0.27$. Thus there is a 50% increase in residual anisotropy due to π -stacking. This information from previous experiments with single labels can be used to improve the dye description for DA distance calculations.

The dye distribution as described by the Free diffusion+Contact/Iso model²⁴² can be modified to describe a label that is π -stacked to the terminal base, by the addition of an accessible volume (AV) close to the terminal base. The additional AV generated is assigned a weight of 50%, since this is the percentage of increase in residual anisotropy due to π -stacking (see eq. 5.4 in Methods section). It is simulated exactly like an AV, but with a shorter linker length and assuming the terminal base as the attachment point. This hybrid AV model can better describe the end label distribution.

In this concept of hybrid accessible volume (hybrid AV), the dye densities are not uniform over the RNA surface (as in Free diffusion+Contact/Iso model) nor enriched near the

attachment point (as in Gaussian chain diffusion model). Instead, the dye density (ρ_{Dye}) varies along the long helical axis (z-axis) as shown in Fig. 5.3B. Close to the helix terminus, there is an enrichment of ρ_{Dye} due to the “localized trap” furnished by the π -stacked dyes. As we move along the z-axis, away from the terminus outwards, the contact volume of the dye maintains a moderately high ρ_{Dye} , albeit slightly lower than the end-trapped region. This region close to the RNA where surface interactions are more probable (“distributed traps” in contrast to the “localized trap” for the end-stacking) is best described by the contact volume part of an ACV. Beyond this region, the dye density sharply decreases as it extends into the region of unhindered dye movement, and then falls to zero outside the accessible volume. The resulting dye model is called a hybrid AV model because it is described using a combination of the Free diffusion+Contact/Iso model (ACV for the “distributed traps”) and the Free diffusion/Iso model (additional AV for the “localized trap”).

In this study, the hybrid AV is determined by carrying out the following steps: (1) The ACV is simulated with $L_{link} = 22 \text{ \AA}$, $w_{link} = 2.5 \text{ \AA}$ and $R_{dye} = 3.5 \text{ \AA}$, contact volume thickness of 2 \AA (empirical value) and a trapped dye fraction calculated from the residual anisotropy according to eq. 5.1 (2) To account for the “localized trap” at the helix terminus (Chapter 3, Fig 3.5A), an additional accessible volume is simulated; the attachment point being the C6 atom of the terminal residue of the labeled strand and a short $L_{link} = 7 \text{ \AA}$, $w_{link} = 2.5 \text{ \AA}$ and $R_{dye} = 3.5 \text{ \AA}$ (3) In order to combine the ACV with the end-stacked AV for calculating inter-dye distances, individual $\langle R_{DA} \rangle_E$ values for both the ACV and the AV to a common donor-AV are calculated ($\langle R_{DA} \rangle_{E,ACV}$ and $\langle R_{DA} \rangle_{E,AV}$, respectively), and subsequently converted to FRET efficiencies (E_{ACV} and E_{AV} , respectively) for weighted averaging. From spectroscopic measurements it was estimated that the increase in residual anisotropy due to π -stacking at the helix end is approx. 50% (see eq. 5.4 in Methods section) and, therefore, the additional AV was assigned a weight of 50%. Based on the assigned weights, the FRET efficiencies from ACV and AV are combined as $E_{hybrid} = 0.5 * E_{ACV} + 0.5 * E_{AV}$, and then converted back to the distance, $\langle R_{DA} \rangle_{E,hybrid}$.

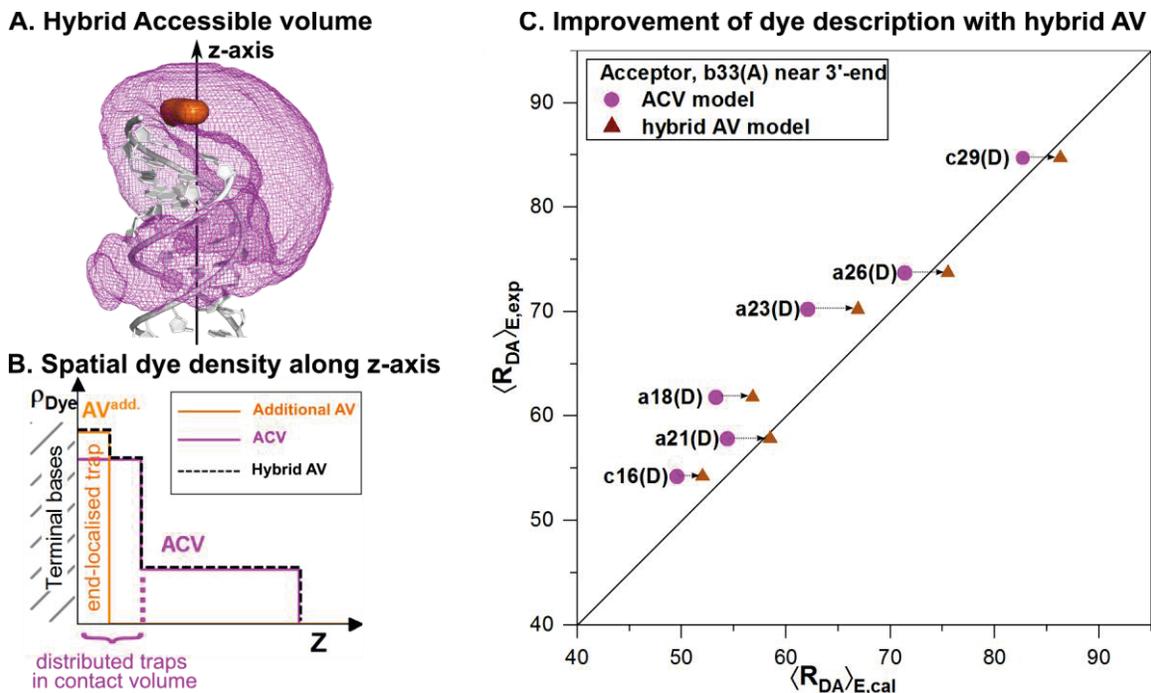


Figure 5.3. Hybrid accessible volume (hybrid AV) model improves dye description in FRET distance calculations. (A) Hybrid AV: The ACV (mesh representation) and terminal AV (orange surface) together make up the hybrid AV. The helix axis of the RNA is assigned as the z-axis (black line with arrowhead pointing outwards) (B) The spatial dye density along the z-axis is displayed originating from the terminal bases and extending outwards. The dye density (ρ_{Dye}) distribution is described by a combination of the additional AV ($AV^{add.}$, orange), due to end dye stacking, and the ACV (magenta). The $AV^{add.}$ represents the end-localised trap region while the contact volume of the ACV may be described as a region of distributed traps. The dye density along z-axis for the hybrid AV is shown in black dashed lines. (C) Hybrid AV improves agreement between experimental and calculated $\langle R_{DA} \rangle_E$ values for acceptor labels near the 3'-end. The thin dashed arrows are drawn to connect points belonging to the same sample.

The $\langle R_{DA} \rangle_{E,hybrid}$ values calculated using the hybrid AV model for the acceptor end labels (Fig. 5.3) is found to be in better agreement with the experimental $\langle R_{DA} \rangle_{E,exp}$ than the $\langle R_{DA} \rangle_{E,ACV}$ from the Free diffusion+contact/Iso model. In comparison to the $\langle R_{DA} \rangle_{E,ACV}$, the $\langle R_{DA} \rangle_{E,hybrid}$ is generally longer due to the higher weighting of dye position distribution at the extreme end of the RNA helix, away from the attachment position which is closer to the donor label.

5.3 Conclusions

In this chapter, the different dye distribution models used to describe fluorescence labels for FRET distance calculations were reviewed. A calibration study was performed with 24

single FRET pairs (Alexa488 donor-Cy5 acceptor) on five double stranded RNA molecules where FRET-averaged donor-acceptor (DA) distances, $\langle R_{DA} \rangle_E$, from MFD experiments and ACV simulations were compared. The Free diffusion + Contact/Iso model (ACV model) was found to be generally suitable to describe the distribution of dyes coupled in the internal regions of the dsRNA. However, where π -stacking is prominent, as for Cy5 labels attached near the helix ends, this model was insufficient. It was demonstrated that inclusion of prior experimental inputs like anisotropy can improve the dye model description when appropriately incorporated into the dye model. A hybrid AV model, a modification of the established ACV model, is proposed. The suitability of the hybrid AV model for π -stacked end labels is demonstrated by a better agreement between experimental and calculated $\langle R_{DA} \rangle_E$. Thus, researchers undertaking FRET studies should attempt to inform themselves of potential interactions that may significantly alter the dye distribution around the biomolecule, and incorporate that into the label description for more accurate FRET-derived structures.

5.4 Materials and Methods

smMFD experiment. Single molecule MFD experiments were carried out and analyzed by Simon Sindbert, Thomas Peulen and Olga Doroshenko according to the procedure described elsewhere²⁵⁶.

Dye simulations. In the simple Free diffusion+Iso model^{103,242}, all allowed positions of the dye are considered to be equally probable, which allows us to define the accessible volume (AV)¹⁰³. The dye is approximated as a sphere with radius R_{dye} , tethered to the biomolecule by a flexible linker of length L_{link} with width w_{link} ¹⁰³. The AV parameters were as follows: for Alexa488, $L_{link} = 20$ Å; for Cy5, $L_{link} = 22$ Å and in both cases, $w_{link} = 2.5$ Å and $R_{dye} = 3.5$ Å. The attachment point for the labels was the C5 atom when attached to uracil or cytosine, and the N2 atom when attached to guanine as described in Chapter 2. The geometric modeling was performed using the software Olga (<https://github.com/Fluorescence-Tools/Olga>). The structures for the double-stranded RNAs used for the AV simulations were generated using the Nucleic Acid Builder from the AmberTools suite of programs²⁵⁷.

In the free diffusion + Contact/Iso model, an additional volume near the RNA surface with thickness 2\AA and dye density calculated from the trapped dye fraction is simulated³. The experimental input of residual anisotropy was measured by ensemble time resolved anisotropy measurements according to the procedure described in Chapter 3 and used to calculate the trapped dye fraction (eq. 4.1)).

The hybrid accessible volume model is a combination of the ACV model enriched by an additional AV near the helix termini whose weighting is determined from the measured residual anisotropies. This is applied only for the labels coupled near the helix termini that stack on the helix end. The weighting for the additional AV is given by the increase of the residual anisotropy due to stacking:

$$\Delta r_{\infty,stack} = \frac{r_{\infty,end} - r_{\infty,internal}}{r_{\infty,end}} \quad (5.4)$$

where the subscripts *end* and *internal* refers to end-labeled and internally labeled dyes.

From chapter 3, $r_{\infty,end} = 0.29$ and $r_{\infty,internal} = 0.14$ Thus $\Delta r_{\infty,stack} = 0.49 \approx 0.5$.

The dye parameters employed for AV and ACV simulations are detailed in the main text.

Chapter 6. Summary

In this thesis, I investigated the environment-dependent behavior of commonly used organic fluorescent dyes, Alexa488, Cy5 and Alexa647 in distinct biomolecular environments when they are covalently labeled to various distinct sites in RNA or the protein, T4 Lysozyme. A total of 104 distinct labels were studied in RNA systems and 21 labels in the protein T4L.

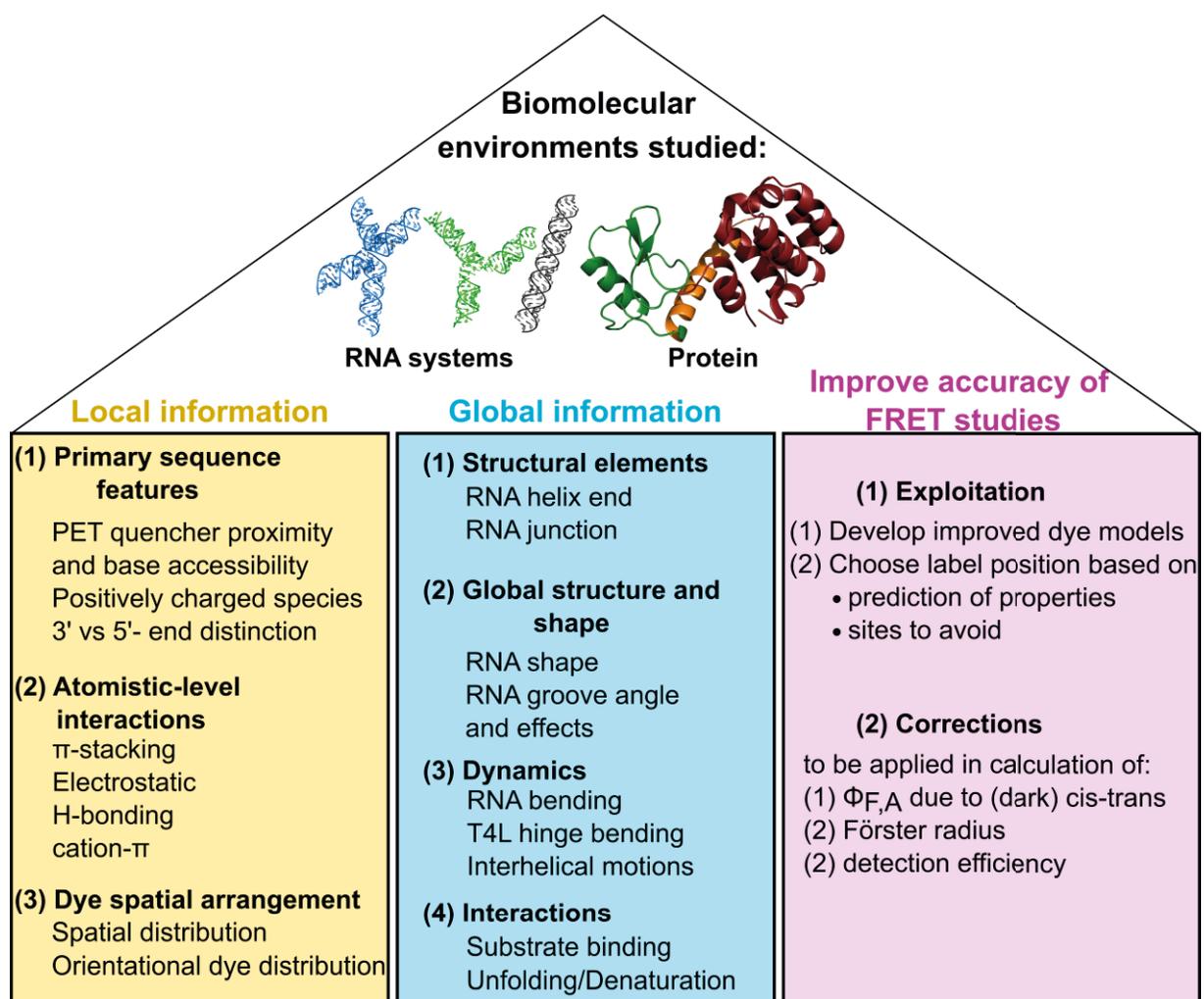


Figure 6.1. Summary of achievements.

Investigation of the local and global environment-dependent dye behavior on biomolecules, their application as biomolecular probes and adoption of this knowledge to improve FRET experiment design and analysis were carried out in 3 parts in my research. The first part focuses on relatively rigid RNA systems, the second, on the model protein

T4L, and the third focuses exclusively on developing FRET models for distance analysis. The experimental techniques included an array of spectroscopic methods and substantiation of experimental results with simulation studies. (i) Ensemble and subensemble lifetime studies, to give a glimpse into quenching environments, (ii) Residual anisotropies to identify sterically constrained environments, (iii) Spectral shifts from absorption measurements, to indicate the extent of dye-biomolecular interactions, (iv) FCS studies, to quantitatively evaluate quenching, cis-trans isomerism processes and global rotation, (v) fluorescence quantum yield measurements to explore the mechanisms of PET quenching and reveal the extents of static/dynamic quenching, (vi) AV simulations and (vii) MD simulations.

Achievements of my research studies (summarized in Fig. 6.1)

Chapter 3: Sensing the local environment and global conformation of RNA by flexibly coupled fluorescent labels

(a) Insights obtained about behavior of Alexa488 and Cy5 fluorophores in distinct RNA environments in RNA duplexes, three-way and four-way junction molecules.

- Dyes undergo π -stacking with the nucleobases at the helix ends or junctions (Fig. 3.2, 3.3, 3.4, 3.5 and 3.6).
- Alexa488 labels at distinct environments are quenched by dynamic and static PET to different degrees depending on the sequence (presence of PET quencher, Guanine) and the π -stacking potential (higher at exposed bases at the helix ends and junction) (Fig. 3.3, 3.4 and 3.5).
- Dye behavior depends on the labeling chemistry employed. A labeling site in the deep major groove directs the dye linker towards the 3' end, while the shallow minor groove provides no directional specificity (Fig. 3.1 and 3.2) so that dye distributions and orientations are distinctive for different positions on the RNA (Fig. 3.4).
- AV simulations allowed the examination of the difference between the 3' and 5' ends, steric constraints at the junction, difference in major- and minor-groove attached labels and proximity to suspected interaction sites. It could also predict dye behavior in relation to the helix end accessibilities (Fig. 3.4).

- Major RNA-dye interactions revealed by MD simulations: π -stacking, groove and backbone interactions (Fig. 3.5).

(b) Employment of fluorophores as local and global environment probes

- Fluorophores act as local sensors. From the insights gathered on dye-biomolecular interactions, we were equipped to sense the unknown environment of a dye-labeled nucleic acid and potentially assess the dye proximity to the helix ends/junction or gauge quencher concentration (Fig. 3.6).
- Appropriately chosen fluorophores can also sense the global environment of the RNA, like the global structure and shape (from FCS, Fig. 3.6), and recognize inter-helical motions (from anisotropy and FCS studies, Fig. 3.2) and RNA bending (Supplementary Fig. 3.14 and 3.24).

(c) Analyzed the implications for FRET studies

Understanding dye behavior at different positions provide knowledge that can be applied to improve the accuracy of FRET results by the careful selection of labeling sites, modification of experiment design, interpretation of results and choice of dye models. We therefore examined our results in the context of FRET studies and

- brought attention to problematic labeling sites like the nucleic acid helix ends, junctions or labeling sites in the vicinity of PET quenchers like guanine, Trp and Tyr.
- Guidelines are provided to recognize potential dye-environmental interactions that could be detrimental for FRET studies
- Proposals to minimize these effects include avoiding sites with strong dye-RNA interactions, corrections for Förster radius, detection efficiency and acceptor fluorescence quantum yield.

Chapter 4: Fluorescent dyes probe the local environment of protein, T4 Lysozyme: Exploring with experiments and simulations

(a) Discerning the behavior of Alexa488 and Alexa647 in distinct T4L environments.

- AV simulations, together with Brownian dynamics simulations, were shown to be successful in explaining the experimentally observed photophysical properties of T4L-

coupled dye labels. The sources of quenching in Alexa488 labels and steric restriction in Alexa647 labels were revealed.

- Alexa488 labels at distinct environments are quenched by dynamic and static PET to different degrees depending on the sequence (presence of PET quenchers, Trp, Tyr, His and Met) and the π -stacking potential. Compared to RNA-coupled Alexa488 dyes, the T4L-coupled dyes exhibit static PET to a much lesser degree and the quenching mechanism is predominantly dynamic.
- Dye-quenching by PET quenchers were boosted in presence of positively charged amino acids which trap the negatively charged dyes (by electrostatic forces) near to the quenchers.
- There are mainly 2 sources of steric restrictions: (1) Constraints due to dye trapping at positively charged amino acids like Lys or Arg and (2) due to squeezing within the cleft volume between subdomains during hinge bending.
- Dye-protein interactions were studied on an atomistic scale and shown to exhibit π - π stacking, multiple H-bonding and cation- π interactions.

(b) As local and global environment probes

- Local sensor: From fluorescence properties of the dyes labeled at a particular position, the proximity to quenchers, positively charged amino acids, cleft region can be inferred.
- The fluorophores can be used as environmental probes to detect biomolecular processes like substrate binding and protein denaturation.

(3) Developing dye models for accurate FRET label description:

- The different dye distribution models (NPS, AV, ACV and weighted AV) used to describe fluorescence labels for FRET distance calculations were reviewed.
- FRET-averaged donor-acceptor (DA) distances, $\langle RDA \rangle_E$, from experiments and model simulations compared.
- Hybrid AV model: When there is significant π -stacking, it was demonstrated that inclusion of prior experimental inputs like anisotropy can improve the dye model description. A hybrid AV model was developed which showed improvement in dye description by better correlating experimental and calculated $\langle RDA \rangle_E$ values.

Conclusion: The investigations carried out in this thesis consolidate the factors affecting fluorophore behavior as shown in figure 6.2. The work described in this thesis helps researchers using fluorescence-label based techniques, especially FRET, to avoid pitfalls during the design of experiments and interpretation of observations. The results emphasize the importance of accounting for specific interaction sites on the biomolecule while choosing dye models, in order to significantly improve the accuracy of FRET experiments. Ultimately, this work emphasizes that the environment sensitivity of dye labels is not a nuisance; instead they present a wealth of information that highlights the potential of these labels to inform on the structure, dynamics and hydrodynamic properties of biomolecules. In this way, this work advocates for the use of fluorophores as probes to report on local and global biomolecular environments (Figure 3.6 and 4.7).

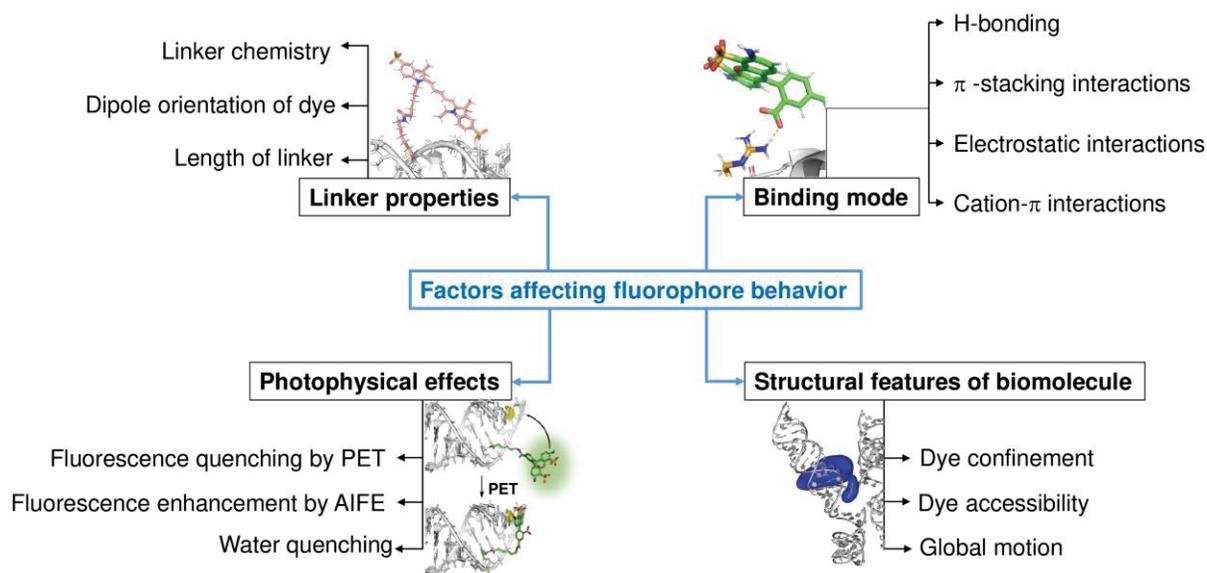


Figure 6.2. Overview of factors affecting fluorophore behavior.

Chapter 7. Additional methods

This chapter contains details on the corrections applied during the analysis of spectroscopic techniques in chapter 3, and detailed protocols of the biochemical methods employed in the course of the work for chapter 4 of this thesis.

7.1 Corrections applied during spectroscopic analysis (employed in chapter 3)

a) Corrections applied during fluorescence quantum yield measurements:

Relative steady state fluorescence quantum yield was measured by comparing to a standard reference dye using one of the following approaches: Conventional method based on integrated fluorescence³³⁶ or liFCS method based on brightness measurement from low intensity FCS²³⁷.

The measured spectra in both cases were corrected during analysis as described below:

(1) Absorption spectrum – Corrections for baseline & Scattered light.

For baseline correction (constant or linear offset):

To determine a constant offset, select a base point at a wavelength above the S1 absorption range of the dye. The correction with a linear offset requires a second interpolation point. This $\lambda(\min)$ is set in a range of minimum absorbance before the main band. A generally non-zero absorption there can be taken into account with a dye-specific value relative to the maximum

$$f = A_{\lambda(\text{offset})}/A_{\lambda(\text{max})} \quad (7.4)$$

$$A_{\text{cor}} = A_{\text{exp}} - b\lambda - c \quad (7.5)$$

For scatter correction:

A third interpolation point is necessary $\lambda(UV)$ with a dye specific absorbance value relative to the maximum.

$$f_u = A_{\lambda(UV)}/A_{\lambda(\text{max})} \quad (7.6)$$

$$A_{\text{cor}} = A_{\text{exp}} - a(1/\lambda^4) - b\lambda - c \quad (7.7)$$

(2) Excitation at $\lambda(ex)$ - Inner Filter Effect (attenuation of excitation light).

Model adapted from ^{369,370}

$$F(corr) = \frac{F(experiment) * 2.303 * E(\lambda) * \Delta I_x}{10^{E(\lambda) * I_x} * 10^{E(\lambda) * \frac{\Delta I_x}{2}} - 10^{-E(\lambda) * \frac{\Delta I_x}{2}}} \quad (7.8)$$

where $E(\lambda)$ is the absorbance at $\lambda(ex)$; I_x is the relative beam position (cuvette centre is fixed at 0.5); ΔI_x is the relative beam width, ($0 < \Delta I_x \leq 1$ default 0.5).

(3) Wavelength calibration: Absorption vs. excitation

An erroneous wavelength calibration of the spectrometer can be taken into account once the shift is experimentally determined. The absorbance value relevant for excitation is then extracted from the absorption spectrum by interpolation.

(4) Fluorescence spectrum - solvent background

The solvent fluorescence spectrum (background) is subtracted from the fluorescence spectrum of the sample.

(5) Fluorescence spectrum - dark current

Dark current varying over the measurement day (especially with uncooled detectors) is only partially compensated for, in the subtraction of the solvent spectrum. A dark current correction must be performed where a fluorescence-free point in the fluorescence spectrum is used to scale the correction spectrum of the spectrometer. This is then subtracted from the measured values.

(6) Fluorescence spectrum - Inner Filter Effect (reabsorption of fluorescence)

Reabsorption based on the respective absorption spectrum can be considered to correct the fluorescence spectrum based on the relative beam position in cuvette.

(7) Fluorescence spectrum - extrapolation (not for reference dye)

For incompletely recorded fluorescence spectra (due to the excitation being performed at long wavelengths), the integral signal can be scaled with a constant factor.

b) Corrections applied for polarizer settings in eTCSPC measurements

It was noticed that the emission polarizer was not perfectly calibrated for some of the measurements. A shift of 1 degree was determined. For the data measured during this period, the polarizer settings were retroactively “corrected” by reanalyzing the data accounting for the 1 degree shift of emission polarizer. For this, rotational correlation functions were derived for any arbitrary angle of polarizer. The following section is the work from Dr. Oleg Opanasyuk from the Seidel lab.

Deriving fluorescence intensities for isotopically diffusing molecules and arbitrary orientation of emission polarizer:

The intensity of fluorescence for a given polarizer setup (a, e) is proportional to the correlation function:

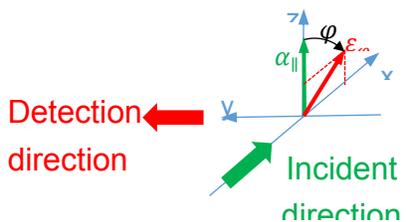
$$I_{a,e}(t) \propto g_{a,e}(t) = \langle p_a(0)p_e(t) \rangle_{\Omega_a(0), \Omega_{ea}(t)}, \quad (7.9)$$

where probabilities $p_a(0), p_e(t)$ are the probability of the absorption of photon at the time 0 and the probability of the absorption at the time t corresponding. These probabilities in turn are proportional to squared scalar products of unit incident (α) and detection (ϵ) polarizer vectors with corresponding the unit absorption transition moment (TM) at time 0 ($\mu_a(0)$) and the emission transition moment at time t ($\mu_e(t)$).

$$\begin{aligned} p_a(0) &= p_{a,0} |\alpha^* \cdot \mu_a(0)|^2 \\ p_e(t) &= p_{e,0} |\epsilon^* \cdot \mu_e(t)|^2 \end{aligned} \quad (7.10)$$

In eq. 7.9 the averaging is done over positions of absorption TMs at the time 0 ($\Omega_a(0)$) and mutual orientations of the absorption TM at the time 0 and the emission TMs at the time t ($\Omega_{ea}(t)$).

Let's define the laboratory coordinate frame as shown:



In this frame the incident and detection polarizer vectors are given by:

$$\alpha_{\parallel} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \quad \varepsilon_{\varphi} = \begin{pmatrix} \sin \varphi \\ 0 \\ \cos \varphi \end{pmatrix} \quad (7.11)$$

Let us also assume that orientations of molecules are distributed isotropically at time 0 and that they diffuse isotropically. It can be shown that for such setup the orientational part of the correlation function (with factor $p_{a,0}p_{e,0}$ omitted) take form:

$$\boxed{g_{\parallel,\varphi}(t) = \frac{1}{9} + \frac{1}{9} (3 \cos^2 \varphi - 1) r(t)}, \quad (7.12)$$

with factor $r(t)$ given by:

$$r(t) = \frac{3}{5} \sum_{ij} M_{ai}^{(2)} M_{ej}^{(2)*} \left\langle D_{ji}^{(2)}(\Omega_{ea}(t)) \right\rangle_{\Omega_{ea}(t)}, \quad (7.13)$$

where $M_{ai}^{(2)}$, $M_{ej}^{(2)}$ are second order components of spherical tensors of the absorption and emission transition moments in the coordinate frames aligned with eigen vectors of molecular diffusion tensor; $D_{ji}^{(2)}(\Omega_{ea}(t))$ are elements of Wigner D-matrix of the angle ($\Omega_{ea}(t)$) between absorption and emission vectors.

The substitution of standard emission polarizer angles gives next expressions for fluorescence intensities:

$$\begin{aligned} \varphi = 0^\circ: & \quad I_{\parallel,\parallel}(t) = I_0(1 + 2 r(t)); \\ \varphi = 90^\circ: & \quad I_{\parallel,\perp}(t) = I_0(1 - r(t)); \\ \varphi = \text{magic angle}: & \quad I_{\parallel,M}(t) = I_0 \end{aligned} \quad (7.14)$$

The time dependent factor $r(t)$ is equal to the fluorescence anisotropy, which becomes clear from substitution of intensities from the eq.7.14 to the standard definition of anisotropy:

$$\frac{I_{\parallel,\parallel}(t) - I_{\parallel,\perp}(t)}{I_{\parallel,\parallel}(t) + 2 I_{\parallel,\perp}(t)} = \frac{g_{\parallel,\parallel}(t) - g_{\parallel,\perp}(t)}{g_{\parallel,\parallel}(t) + 2 g_{\parallel,\perp}(t)} = r(t) \quad (7.15)$$

Thus, the eq.7.13 is the expression of a fluorescence anisotropy in terms of orientations of transition moments relative to molecular coordinate frames and some time dependent factors $\langle D_{ji}^{(2)}(\Omega_{ea}(t)) \rangle_{\Omega_{ea}(t)}$ which can be easily calculated as functions of the diffusion tensor of molecules.

Errors introduced by the emission polarizer errors to the total fluorescence intensity

Using eq.7.12 and 7.14, we can calculate the relative errors to the total intensity measured by the magic angle and combination methods.

$$\delta I_0(\Delta\varphi) = \frac{I_0(\Delta\varphi)}{I_0} - 1 \quad (7.16)$$

magic angle: $\delta I_0^M(\Delta\varphi) = (3 \cos^2(\varphi_M + \Delta\varphi) - 1) r$; $\varphi_M \approx 54.7^\circ$

combination: $\delta I_0^C(\Delta\varphi) = \sin^2 \Delta\varphi r$

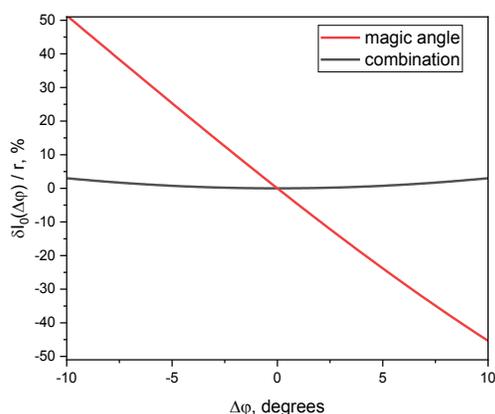


Figure S7. Comparison of relative errors in fluorescence intensity introduced by the emission polarizer by different measurement methods.

So, the magic angle method is much more prone to errors caused by inaccuracy of the emission polarizer. However, it should be noted that the combination method is prone to errors caused by erroneous g-factor and IRF measurements.

7.2 Biochemical methods (employed in chapter 4)

The biochemical protocols detailed below³³⁵ were followed to achieve dye-labeled T4L samples for fluorescence studies in Chapter 4. The T4 lysozyme mutants were cloned into the pET11a vector (Novagen®, Merck KGaA, Darmstadt, Germany). For the expression of proteins with the unnatural amino acid pAcPhe, the pEvolv-AcPhe plasmid was used.

a) Site-directed mutagenesis

In order to insert a specific mutation in a protein, the respective codon in the gene of interest is modified using site-directed mutagenesis. This is accomplished with one temperature-gradient PCR and one complementary pair of primer carrying the modified base triplet. For each mutation, five reaction mixtures for scanning five different annealing temperatures were prepared, with each mixture containing 5 μ l 10x Pfu buffer +MgSO₄, 1 μ l dNTP mix (10 mM each), 10 ng DNA template, 0.5 μ l forward primer, 0.5 μ l reverse primer, filled upto 49 μ l with ddH₂O and 1 μ l of polymerase added at last step. The five different annealing temperatures used are $(t_m - x)$ °C, where t_m is the primer melting temperature and $x = 5-10$. In the PCR program, the temperature gradient (T_{grad}) across the PCR sample heating block is chosen such that the slots where the 5 different samples are kept, receive the chosen annealing temperature. Then the PCR program is: 95 °C (3 min) \rightarrow 95 °C (30 s) \rightarrow T_{grad} (30 s) \rightarrow 72 °C (12 min) \rightarrow Repeat previous 3 steps 25 times \rightarrow 72 °C (15 min) \rightarrow Store at 4 °C. In order to digest any remaining template DNA that was not mutated in the PCR, 0.5 μ L of DpnI enzyme was added at 37 °C and kept for 1 hr. Subsequently, 10 μ L of each reaction mixture was mixed with 5x DNA loading dye (10 mM Tris-HCl, pH 7.5, 0.2 % (w/v) orange G, 1 mM EDTA, 50 % glycerol), loaded onto a 1 % agarose gel and a gel electrophoresis carried out. 2 μ L of each positive reaction mixture was transformed into E.coli DH5 α cells.

b) Agarose gel electrophoresis

In order to check for the successful amplification of plasmids during PCR, we run an agarose gel electrophoresis. A 1% agarose gel, prepared from broad range Agarose and 1xTAE buffer (40 mM Tris-HCl pH 8.2, 20 mM sodium acetate, 1 mM EDTA), was casted in a mold and 2 μ L of Midori Green distributed in the gel and let to harden for 30 min. The gel was kept in the electrophoresis chamber and flooded with 1x TAE-buffer. The samples containing 2 μ L 5x DNA loading dye were pipetted into the pockets of the gel along with 5 μ L of the GeneRuler 1kb ladder as a size marker. The electrophoresis was carried out at a constant voltage of 100 V for 30 min. The gels were then viewed under UV-light.

c) Preparation of chemically competent cells

Under sterile conditions, 5 μ L of the required bacterial strain (DH5 α , BL21 etc.) was added to 5 mL of LB-medium and incubated at 37 °C overnight while shaking. Then 100 mL of LB were inoculated with 1/100 of its volume with the overnight culture and 1/50 of its volume with a mix of 500 mM MgCl₂, 500 mM MgSO₄. The cells were grown at 37°C while shaking until an OD₆₀₀ of 0.5- 0.6 was reached, and then harvested in 2 ml aliquots by centrifugation at 8000*g for 3 min. The cell pellet obtained was resuspended in 1 mL of TMF-buffer (100 mM CaCl₂, 50 mM RbCl₂, 40 mM MnCl₂) and incubated for 30 min on ice. The cells were again harvested at 8000*g for 3 min, resuspended in 250 μ L of TMFG-buffer (TMF-buffer, 20 % (v/v) glycerol) and stored at -80 °C until use.

d) Preparation of plasmid DNA

As the first step for preparation of plasmid DNA, a culture of E. coli DH5 α cells in LB medium containing the respective antibiotic was incubated overnight at 37 °C while shaking. Then, the sonification of the bacterial culture is carried out to obtain the lysate. This is then further purified with the QIAprep Hispeed Midi Kit. The concentration was determined using UV spectrometry at 260 nm and dispatched for sequencing to GATC Biotech (Konstanz, Germany).

e) Plasmid transformation

The plasmid DNA was transformed to the competent bacterial strain, BL21, via heat shock transformation. The cells were thawed on ice for 30 min and the plasmid was added. For co-transformation, 30 ng of pEvolv (which carries a Chloramphenicol resistance), and 15

ng of the Amber-mutant carrying pET11a (which carries an Ampicillin resistance) plasmids, were added simultaneously. This was incubated on ice for 15 min and a heat shock applied by placing the cells were for 45 sec at 42 °C. This was followed by an incubation of 2 min on ice. The cells were spun down for 1 min at an acceleration of 13000 *g, redissolved in 250 µl of LB-medium and grown while shaking for 30 min at 37 °C prior to plating on LB-agar plates containing both Chloramphenicol and Ampicillin antibiotics.

f) Protein production

From the plates containing freshly transformed cells, a colony was picked and incubated overnight at 37 °C in 100 ml of LB medium containing the respective antibiotics while shaking at 150 rpm. This is the preculture. Then a large volume (1 or 2 L) of LB media is taken and inoculated with 1/50 of its volume of preculture and the respective antibiotics (main culture). The cells were grown while shaking at 37 °C until an OD₆₀₀ of ~ 0.6 was reached. Then, the temperature is lowered to 30 °C and the protein production induced by the addition of 1M IPTG to bring the final concentration in the culture to 1mM IPTG. Prior to induction and every two hours after induction, a 1 ml sample was taken and stored to monitor the protein production. 6 hours after induction, the cells were harvested by centrifugation for 15 min at 10000*g and the pellets stored at -20°C until purification. To check, whether the protein had been produced, the collected samples were applied onto SDS PAGE. For a better comparison, samples were normalized to have the same OD₆₀₀ by resuspension in buffer.

For the production of proteins containing the unnatural amino acid pAcPhe, 0.4 g/L of this chemical was added additionally to the main culture during the inoculation stage and the cells grown as before. At the stage of IPTG induction, 4 g/L of arabinose was also added due to the pEvolv-AcPhe plasmid.

g) SDS polyacrylamide gel electrophoresis (SDS-PAGE)

In order to evaluate the purity and yield during the protein production, SDS-PAGE was carried out. The gels were composed of the upper stacking gel (5 % acrylamide) and the lower resolving gel (15 % acrylamide). 12 µL of the resuspended protein solutions, together with 3 µL of SDS loading buffer were applied on the gel. Samples taken from the growth culture were heated for 5 min at 95 °C prior to application to the gel to release the

protein from the bacterial cells. After heating, centrifugation is carried out to separate the proteins from other insoluble bacterial cell fragments. The electrophoresis was carried out at a constant voltage of 200 V until the dye front reached the end of the gel. After electrophoresis, the gel was washed thrice with water to remove SDS and then stained with SimplyBlue™ SafeStain dye (Invitrogen) while gently shaking for 1 hr and viewed under ambient light.

h) Protein purification

Lysis: The harvested and stored cells after induction were thawed on ice in CatIEX buffer A (25 mM HEPES pH 7.5, 0.1 mM EDTA, 5 mM DTT). Subsequently, cell lysis was performed on ice via sonification for three cycles of 30 s ON – 60 s OFF with an ultrasonicator at 40 % power. Centrifugation for 45 min at 30'000 * g at 4 °C separates insoluble cell fragments.

Ion exchange chromatography (IEX): In an ion exchange chromatography (IEX) the proteins are separated on the basis of their charge, using a solution of varying ionic strength. In a cationic IEX (used here), the column material is negatively charged and the protein is in a buffer with a low salt concentration and a pH value below its pI and is therefore positively charged. The protein to be purified was loaded on a HiTrap SP Sepharose FF 5/50 column (GE Healthcare), which was preequilibrated with CatIEX buffer A. For the whole process, the flow rate was maintained constant at 1.5 mL/min. After the protein was loaded, the column was washed with buffer A upto 5 times the column volume (CV) to remove the unbound protein. The different steps (washing, elution etc.) can be followed in real time by detection with UV/visible light absorbance and conductivity measurements. For the protein elution, an increasing salt gradient from 0 % to 100 % CatIEX buffer B (25 mM HEPES pH 7.5, 0.1 mM EDTA, 1 M NaCl, 5 mM DTT) was applied upto 10 CV. The gradient was achieved through mixing of CatIEX buffers A and B. Fractions of 2 ml were collected during elution and analyzed on an SDS gel.

i) Protein concentration and buffer exchange

The pooled protein fractions from IEX were first filtered through an Amicon concentrator (Millipore) with a molecular weight cutoff (MWCO) of 50 kDa to remove impurities of high molecular weight and then concentrated using a 10 kDa filter. A swinging bucket

centrifuge was used and the samples periodically stirred to remove any concentration gradient. For T4L mutants containing cysteine residue, 5mM DTT was added to reduce the cysteines and incubated for 1 hour before the filtration step.

The concentrated protein was then diluted 100 fold with the desired buffer and concentrated again and the procedure repeated three times. The concentration of the unlabeled protein was determined photometrically using the Beer Lambert law at a wavelength of 260 nm and 280 nm.

j) Site-specific labeling

T4L variants were labeled using orthogonal chemistry. For the sites containing a cysteine residue, Michael addition between the thiol-group of the cysteine and maleimide group of the dye derivative was carried out. For the sites with the unnatural amino acid pAcPhe, a ketoxime is formed with the hydroxyl amine dye derivative.

Michael addition: A fivefold molar excess of dye-maleimide was added to the selected variant and the reaction mixture was incubated in the dark for 2 h at room temperature in PBS. Subsequently, the excess of dye was removed. For single fluorophore labeling, the buffer was exchanged to PBS and the degree of labeling (DOL) determined.

Orthogonal labeling with pAcPhe: After the removal of excess dye, the buffer is exchanged to Keto-labeling buffer (50 mM Sodium acetate, 150 mM NaCl, adjusted with acetic acid to pH 4) using a desalting column (Nap5 or Nap10, GE Healthcare). Immediately after elution, a fivefold molar excess of the dye-hydroxylamine was added to the protein solution. The reaction mixture is incubated for 72 h in the dark at 37 °C, and then centrifuged for 30 min at 13'000 * g to remove any precipitated protein. The clear supernatant was collected and stored. The pelleted part was dissolved slowly in 100 – 300 µL of PBS containing 8 M urea. For efficient refolding of T4L while decreasing the urea concentration, the clear supernatant solution (containing no urea) was added in several small fractions to the dissolved pellet, each addition was followed by severe mixing. To remove the non-reacted dye, the buffer was exchanged to PBS using a desalting column (Nap10 or PD10, GE Healthcare). The labeled protein was concentrated using concentrators with a MWCO of 10 kDa. To ensure further removal of free dye, the

concentrated protein was diluted 15 fold with PBS and concentrated again. This process was repeated twice, and then the concentration and degree of labeling determined.

k) Determination of labeled protein concentration and the degree of labeling (DOL)

The concentration of labeled protein c_P and its degree of labeling were determined via absorption spectrometry using the following equations:

$$c_P = \left(Abs_{280} - (Abs_{496} \cdot cf_{D,280}) \right) \cdot \frac{dF}{\varepsilon_P \cdot d} \quad (7.1)$$

$$DOL_{A488} = \frac{c_{A488}}{c_P} \quad (7.2)$$

$$DOL_{A647} = \frac{c_{A647}}{c_P}$$

where Abs_{280} and Abs_{496} are the absorptions measured at 280 nm and 496 nm respectively, cf is the correction factor with the subscripts denoting the dye (D= donor, Alexa488 and A=acceptor, Alexa647) and the absorption wavelength. The dilution factor is given by dF , ε_P is the extinction coefficient of the protein, and d is the pathlength.

For double-labeled samples, the absorption of the acceptor dye in the donor absorption region must also be accounted for. In this case, the concentration is given by:

$$c_P = \left(Abs_{280} - \left[\left((Abs_{496} - Abs_{650} \cdot cf_{A,496}) \cdot cf_{D,280} \right) + Abs_{650} \cdot cf_{A,280} \right] \right) \cdot \frac{dF}{\varepsilon_P \cdot d} \quad (7.3)$$

and the labeling degree (DOL) calculated using the eq. (7.2) above.

Frequently used abbreviations

| | Meaning | | Meaning |
|------------------|--|--------------------|---|
| 3WJ | Three-way junction | MFD | multi parameter fluorescence detection |
| 4WJ | Four-way junction | MWCO | molecular weight cutoff |
| ACV | Accessible contact volume | NA | nucleic acid |
| AIFE | Association-induced fluorescence enhancement | NAIFE | Nucleic acid-induced fluorescence enhancement |
| ALEX | Alternating laser excitation | NHE | Normal hydrogen electrode |
| APBS | Adaptive Poisson-Boltzmann Solver | NHS | N-hydroxy succinimide |
| APD | avalanche photodiode | NMR | Nuclear magnetic resonance |
| AV | Accessible volume | NPS | Nano Positioning System |
| BD | Brownian dynamics | PCR | polymerase chain reaction |
| DA | donor-acceptor | PDB | Protein database |
| DEPET | distance-encoding PET | PET | Photoinduced electron transfer |
| DNA | Deoxyribonucleic acid | pFCS | polarization-resolved fluorescence correlation spectroscopy |
| DNA-PAINT | DNA-Point Accumulation in Nanoscale Topography | PIE-FRET | Pulsed interleaved excitation-FRET |
| DOL | Degree of labeling | PIFE | Protein-induced fluorescence enhancement |
| ds | double strand | QY | quantum yield |
| DTT | Dithiothreitol (a redox reagent) | RIFE | RNA-induced fluorescence enhancement |
| DWT | Diffusion with traps | RNA | Ribonucleic acid |
| EDTA | Ethylenediaminetetraacetic acid | SCE | Saturated calomel electrode |
| EPR | Electron paramagnetic resonance | SDS-PAGE | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| ESP | electrostatic surface potential | SIFI | Stacking-induced fluorescence increase |
| eTCSPC | ensemble TCSPC | ss | single strand |
| FAM | Fluorescein amidite | SSB protein | Single-strand DNA-binding protein |
| FCS | Fluorescence correlation spectroscopy | STORM | Stochastic Optical Reconstruction Microscopy |
| fFCS | filtered FCS | T4L | T4 lysozyme (Lysozyme of bacteriophage, T4) |
| FISH | Fluorescence in situ hybridization | TCSPC | Time-correlated single photon counting |
| FRET | Förster resonance energy transfer | TM | transition moment |
| IC | Internal conversion | TMR, TAMRA | Tetramethyl Rhodamine |
| IEX | Ion exchange chromatography | UAA | unnatural amino acid |
| IPTG | Isopropyl β - d-1-thiogalactopyranoside | UV | Ultraviolet |
| IRF | instrument response function | UV-Vis | Ultraviolet-Visible |
| ISC | Inter system crossing | VR | Vibrational relaxation |
| LB | Lysogeny broth/Luria-Bertani | WIC | Wobbling-in-cone |
| MD | Molecular dynamics | WT | Wild type |

Acknowledgements

This thesis work was accomplished with the help, support and guidance of many people.

First and foremost, I would like to thank my supervisor, Prof. Dr. Claus Seidel, for his support and guidance throughout and for giving me the opportunity to learn advanced techniques and collaborate with leading minds in this field.

I also wish to thank my mentor, Prof Lutz Schmitt, for his support during my PhD period.

I am extremely grateful to Dr. Anders Barth for his supervision, the lively discussions and collaborations for paper writing.

My sincere gratitude to Dr. Christian Hanke, for the constant support, prompt help, encouraging words and reviewing most of my thesis.

I sincerely thank Dr. Olga Doroshenko for her patience during my training period, her scientific collaborations and her friendship.

I deeply appreciate Dr. Mykola Dimura, Dr. Thomas Peulen, and Dr. Oleg Opanasyuk for their enthusiastic discussions and dedication to science.

My sincere thanks to:

Dr. Katherina Hemmen and Dr. Jakub Kubiak for training me in biochemical protocols and the very helpful scientific discussions.

Dr. Suren Felekyan and Dr. Ralf Kühnemuth, for being accessible and their tremendous help with measurements, analysis and interpretations.

Milana Popara, for a friendly office atmosphere, highly practical scientific discussions and all-round support.

Alexander Larbig, for his patient explanations and for reviewing a large part of my thesis.

Noah Salama, for his willingness to help and his infectious positivity.

Veronika Mendorf, Bekir Bulat, Matthias Fröscher and Martin Schramm for technical and bureaucratic support.

I thank all the past and present members of Prof. Seidel's lab, including Dr. Deborah Sandrin, Dr. Qijun Ma, Dr. Hayk Vardanyan, Dr. Anne-Marie Griefe, Peter Zentis, Julian Koch, Dragana Sretenovic, Bianca Reschke, Julian Folz, Jan-Hendrik Budde, Nicolaas van der Voort and Laura Vogel for their invaluable support.

I would like to thank my family for their constant encouragement and understanding. I express my deep gratitude to my husband, Dr. Biju Francis, for his insightful comments, valuable friendship and supporting my dreams.

Finally, I gratefully acknowledge the German Academic Exchange Service (DAAD) for awarding me their research scholarship to pursue my doctoral programme in Germany.

References

1. Schluesche, P.; Stelzer, G.; Piaia, E.; Lamb, D. C.; Meisterernst, M., NC2 mobilizes TBP on core promoter TATA boxes. *Nat Struct Mol Biol* **2007**, *14* (12), 1196-1201.
2. Kilic, S.; Felekyan, S.; Doroshenko, O.; Boichenko, I.; Dimura, M.; Vardanyan, H.; Bryan, L. C.; Arya, G.; Seidel, C. A. M.; Fierz, B., Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1 α . *Nat Commun* **2018**, *9* (1), 235.
3. Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T. O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kuhnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M., Resolving dynamics and function of transient states in single enzyme molecules. *Nat Commun* **2020**, *11* (1).
4. Wozniak, A. K.; Schröder, G. F.; Grubmüller, H.; Seidel, C. A. M.; Oesterhelt, F., Single-molecule FRET measures bends and kinks in DNA. *Proc Natl Acad Sci USA* **2008**, *105* (47), 18337-18342.
5. McKinney, S. A.; Déclais, A.-C.; Lilley, D. M.; Ha, T., Structural dynamics of individual Holliday junctions. *Nat Struct Biol* **2003**, *10* (2), 93-97.
6. Cawte, A. D.; Unrau, P. J.; Rueda, D. S., Live cell imaging of single RNA molecules with fluorogenic Mango II arrays. *Nat Commun* **2020**, *11* (1), 1-11.
7. Chen, K. H.; Boettiger, A. N.; Moffitt, J. R.; Wang, S.; Zhuang, X., Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **2015**, *348* (6233), 412.
8. Kuang, T.; Chang, L.; Peng, X.; Hu, X.; Gallego-Perez, D., Molecular Beacon Nano-Sensors for Probing Living Cancer Cells. *Trends Biotechnol* **2017**, *35* (4), 347-359.
9. Teng, K. W.; Ishitsuka, Y.; Ren, P.; Youn, Y. A.; Deng, X.; Ge, P. H.; Belmont, A. S.; Selvin, P. R., Labeling proteins inside living cells using external fluorophores for microscopy. *Elife* **2016**, *5*.
10. Marras, S. A. E., Interactive fluorophore and quencher pairs for labeling fluorescent nucleic acid hybridization probes. *Mol Biotechnol* **2008**, *38* (3), 247-255.
11. Nazarenko, I.; Pires, R.; Lowe, B.; Obaidy, M.; Rashtchian, A., Effect of primary and secondary structure of oligodeoxyribonucleotides on the fluorescent properties of conjugated dyes. *Nucleic Acids Res* **2002**, *30* (9), 2089-2095.
12. Ochmann, S. E.; Vietz, C.; Trofymchuk, K.; Acuna, G. P.; Lalkens, B.; Tinnefeld, P., Optical Nanoantenna for Single Molecule-Based Detection of Zika Virus Nucleic Acids without Molecular Multiplication. *Anal Chem* **2017**, *89* (23), 13000-13007.
13. Brodsky, R. A.; Mukhina, G. L.; Li, S. Y.; Nelson, K. L.; Chiurazzi, P. L.; Buckley, J. T.; Borowitz, M. J., Improved detection and characterization of paroxysmal nocturnal hemoglobinuria using fluorescent aerolysin. *American Journal of Clinical Pathology* **2000**, *114* (3), 459-466.
14. Rissin, D. M.; Kan, C. W.; Campbell, T. G.; Howes, S. C.; Fournier, D. R.; Song, L.; Piech, T.; Patel, P. P.; Chang, L.; Rivnak, A. J.; Ferrell, E. P.; Randall, J. D.; Provuncher, G. K.; Walt, D. R.; Duffy, D. C., Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nat Biotechnol* **2010**, *28* (6), 595-599.
15. Sisamakris, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M., Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Meth Enzymol* **2010**, *475*, 455-514.

16. Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A. M., Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Anal Chem* **2006**, *78* (6), 2039-2050.
17. Stryer, L.; Haugland, R. P., Energy transfer - A spectroscopic ruler. *Proc Natl Acad Sci USA* **1967**, *58* (2), 719-726.
18. Forster, T., Zwischenmolekulare energiewanderung und fluoreszenz. *Annalen Der Physik* **1948**, *2* (1-2), 55-75.
19. Rigler, R.; Elson, E. S., *Fluorescence Correlation Spectroscopy*. Springer Berlin, Heidelberg: 2001; Vol. 65, p 1-487.
20. Bacia, K.; Haustein, E.; Schwille, P., Fluorescence correlation spectroscopy: principles and applications. *Cold Spring Harb Protoc* **2014**, *2014* (7), 709-25.
21. Ries, J.; Schwille, P., Fluorescence correlation spectroscopy. *Bioessays* **2012**, *34* (5), 361-368.
22. Toseland, C. P., Fluorescent labeling and modification of proteins. *J Chem Biol* **2013**, *6* (3), 85-95.
23. Dziuba, D.; Didier, P.; Ciaco, S.; Barth, A.; Seidel, C. A. M.; Mely, Y., Fundamental photophysics of isomorphous and expanded fluorescent nucleoside analogues. *Chem Soc Rev* **2021**, *50* (12), 7062-7107.
24. Korlach, J.; Bibillo, A.; Wegener, J.; Peluso, P.; Pham, T. T.; Park, I.; Clark, S.; Otto, G. A.; Turner, S. W., Long, processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides. *Nucleosides Nucleotides & Nucleic Acids* **2008**, *27* (9), 1072-1083.
25. Hottin, A.; Marx, A., Structural Insights into the Processing of Nucleobase-Modified Nucleotides by DNA Polymerases. *Acc Chem Res* **2016**, *49* (3), 418-427.
26. Sinkeldam, R. W.; Greco, N. J.; Tor, Y., Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem Rev* **2010**, *110* (5), 2579-2619.
27. Verma, S.; Eckstein, F., Modified oligonucleotides: Synthesis and strategy for users. *Annu Rev Biochem* **1998**, *67*, 99-134.
28. Kricka, L. J.; Fortina, P., Analytical Ancestry: "Firsts" in Fluorescent Labeling of Nucleosides, Nucleotides, and Nucleic Acids. *Clinical Chemistry* **2009**, *55* (4), 670-683.
29. Zohar, H.; Muller, S. J., Labeling DNA for single-molecule experiments: methods of labeling internal specific sequences on double-stranded DNA. *Nanoscale* **2011**, *3* (8), 3027-3039.
30. Goodchild, J., Conjugates of Oligonucleotides and Modified Oligonucleotides: A Review of Their Synthesis and Properties. *Bioconjug Chem* **1990**, *1* (3), 165-187.
31. Xu, W.; Chan, K. M.; Kool, E. T., Fluorescent nucleobases as tools for studying DNA and RNA. *Nature Chemistry* **2017**, *9* (11), 1043-1055.
32. Bell, N. M.; Micklefield, J., Chemical Modification of Oligonucleotides for Therapeutic, Bioanalytical and other Applications. *ChemBioChem* **2009**, *10* (17), 2691-2703.
33. Gottfried, A.; Weinhold, E., Sequence-specific covalent labelling of DNA. *Biochemical Society Transactions* **2011**, *39*, 623-628.
34. Dong, Y. C.; Liu, D. S.; Yang, Z. Q., A brief review of methods for terminal functionalization of DNA. *Methods* **2014**, *67* (2), 116-122.
35. Thuong, N. T.; Asseline, U., Modification of the 5' Terminus of Oligonucleotides for Attachment of Reporter and Conjugate Groups. *Curr Protoc in Nucleic Acid Chem* **2000**, *00* (1), 4.2.1-4.2.33.

36. Seidel, C. A. M.; Schulz, A.; Sauer, M. H. M., Nucleobase-specific quenching of fluorescent dyes .1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *J Phys Chem* **1996**, *100* (13), 5541-5553.
37. Eckstein, F., Nucleoside phosphorothioates. *Annu Rev Biochem* **1985**, *54*, 367-402.
38. Olsen, D. B.; Sayers, J. R.; Eckstein, F., Site-directed mutagenesis of single-stranded and double-stranded DNA by phosphorothioate approach. *Meth Enzymol* **1993**, *217*, 189-217.
39. Fidanza, J. A.; McLaughlin, L. W., Introduction of reporter groups at specific sites in DNA containing phosphorothioate diesters. *J Am Chem Soc* **1989**, *111* (25), 9117-9119.
40. Hodges, R. R.; Conway, N. E.; McLaughlin, L. W., Post-assay covalent labeling of phosphorothioate-containing nucleic-acids with multiple fluorescent markers. *Biochem* **1989**, *28* (1), 261-267.
41. Hartmann, A.; Krainer, G.; Schlierf, M., Different Fluorophore Labeling Strategies and Designs Affect Millisecond Kinetics of DNA Hairpins. *Molecules* **2014**, *19* (9), 13735-13754.
42. Sanborn, M. E.; Connolly, B. K.; Gurunathan, K.; Levitus, M., Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA. *J Phys Chem B* **2007**, *111* (37), 11064-11074.
43. Nguyen, B.; Ciuba, M. A.; Kozlov, A. G.; Levitus, M.; Lohman, T. M., Protein Environment and DNA Orientation Affect Protein-Induced Cy3 Fluorescence Enhancement. *Biophys J* **2019**, *117* (1), 66-73.
44. Walter, N. G.; Burke, J. M., Real-time monitoring of hairpin ribozyme kinetics through base-specific quenching of fluorescein-labeled substrates. *RNA* **1997**, *3* (4), 392-404.
45. Markiewicz, W. T.; Wyrzykiewicz, T. K., Universal solid supports for the synthesis of oligonucleotides with terminal 3'-phosphates. *Nucleic Acids Res* **1989**, *17* (18), 7149-7158.
46. Ozaki, H.; McLaughlin, L. W., The estimation of distances between specific backbone-labeled sites in DNA using fluorescence resonance energy-transfer. *Nucleic Acids Res* **1992**, *20* (19), 5205-5214.
47. Fidanza, J. A.; Ozaki, H.; McLaughlin, L. W., Site-specific labeling of DNA-sequences containing phosphorothioate diesters. *J Am Chem Soc* **1992**, *114* (14), 5509-5517.
48. Ranjit, S.; Gurunathan, K.; Levitus, M., Photophysics of Backbone Fluorescent DNA Modifications: Reducing Uncertainties in FRET. *J Phys Chem B* **2009**, *113* (22), 7861-7866.
49. van der Meer, B. W., Kappa-squared: from nuisance to new sense. *Rev Mol Biotechnol* **2002**, *82* (3), 181-196.
50. Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J., Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc Natl Acad Sci USA* **2008**, *105* (32), 11176-11181.
51. Giller, G.; Tasara, T.; Angerer, B.; Muhlegger, K.; Amacker, M.; Winter, H., Incorporation of reporter molecule-labeled nucleotides by DNA polymerases. I. Chemical synthesis of various reporter group-labeled 2'-deoxyribonucleoside-5'-triphosphates. *Nucleic Acids Res* **2003**, *31* (10), 2630-2635.
52. Weeks, K. M.; Crothers, D. M., Major groove accessibility of RNA. *Science* **1993**, *261* (5128), 1574-1577.
53. Frugier, M.; Schimmel, P., Subtle atomic group discrimination in the RNA minor groove. *Proc Natl Acad Sci USA* **1997**, *94* (21), 11291-11294.

54. Hiratsuka, T., New ribose-modified fluorescent analogs of adenine and guanine-nucleotides available as substrates for various enzymes. *Biochimica Et Biophysica Acta* **1983**, *742* (3), 496-508.
55. Hiratsuka, T.; Uchida, K., Preparation and properties of 2'(or 3')-o-(2,4,6-trinitrophenyl) adenosine 5'-triphosphate, an analog of adenosine-triphosphate. *Biochimica Et Biophysica Acta* **1973**, *320* (3), 635-647.
56. Divita, G.; Goody, R. S.; Gautheron, D. C.; Dipietro, A., Structural mapping of catalytic site with respect to alpha-subunit and noncatalytic site in yeast mitochondrial F1-ATPase using fluorescence resonance energy-transfer. *J Biol Chem* **1993**, *268* (18), 13178-13186.
57. Grubmeyer, C.; Cross, R. L.; Penefsky, H. S., Mechanism Of ATP hydrolysis by beef-heart mitochondrial ATPase - rate constants for elementary steps in catalysis at a single site. *J Biol Chem* **1982**, *257* (20), 2092-2100.
58. Shapiro, A. B.; McCarty, R. E., Alteration of the nucleotide-binding site asymmetry of chloroplast coupling factor-i by catalysis. *J Biol Chem* **1988**, *263* (28), 14160-14165.
59. Scheidig, A. J.; Franken, S. M.; Corrie, J. E. T.; Reid, G. P.; Wittinghofer, A.; Pai, E. F.; Goody, R. S., X-Ray Crystal-Structure Analysis Of The Catalytic Domain Of The Oncogene Product P21(H-RAS) complexed with caged GTP and MANT DGPPNHP. *J Mol Biol* **1995**, *253* (1), 132-150.
60. Lee, L. G.; Connell, C. R.; Woo, S. L.; Cheng, R. D.; McArdle, B. F.; Fuller, C. W.; Halloran, N. D.; Wilson, R. K., DNA Sequencing With Dye-Labeled Terminators And T7 DNA-polymerase - effect of dyes and DNTPS on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res* **1992**, *20* (10), 2471-2483.
61. Johnson, A. E., Fluorescence approaches for determining protein conformations, interactions and mechanisms at membranes. *Traffic* **2005**, *6* (12), 1078-1092.
62. Shuster, S. O.; Lee, J. C., Tryptophan Probes of TDP-43 C-Terminal Domain Amyloid Formation. *J Phys Chem B* **2021**, *125* (15), 3781-3789.
63. Hellmann, N.; Schneider, D., Hands On: Using Tryptophan Fluorescence Spectroscopy to Study Protein Structure. In *Protein Supersecondary Structures: Methods and Protocols*, Kister, A. E., Ed. Springer New York: New York, NY, 2019; pp 379-401.
64. Ghisaidoobe, A. B. T.; Chung, S. J., Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Forster Resonance Energy Transfer Techniques. *Int. J. Mol Sci* **2014**, *15* (12), 22518-22538.
65. Gerdes, H. H.; Kaether, C., Green fluorescent protein: Applications in cell biology. *Febs Lett* **1996**, *389* (1), 44-47.
66. Kremers, G. J.; Gilbert, S. G.; Cranfill, P. J.; Davidson, M. W.; Piston, D. W., Fluorescent proteins at a glance. *Journal of Cell Science* **2011**, *124* (2), 157-160.
67. Sun, X. L.; Zhang, A. H.; Baker, B.; Sun, L.; Howard, A.; Buswell, J.; Maurel, D.; Masharina, A.; Johnsson, K.; Noren, C. J.; Xu, M. Q.; Correa, I. R., Development of SNAP-Tag Fluorogenic Probes for Wash-Free Fluorescence Imaging. *ChemBioChem* **2011**, *12* (14), 2217-2226.
68. Kapanidis, A. N.; Ebright, Y. W.; Ebright, R. H., Site-Specific Incorporation of Fluorescent Probes into Protein: Hexahistidine-Tag-Mediated Fluorescent Labeling with (Ni²⁺:Nitrilotriacetic Acid)_n-Fluorochrome Conjugates. *J Am Chem Soc* **2001**, *123* (48), 12123-12125.
69. Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohane, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman,

- K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V., HatoTag: A novel protein labeling technology for cell imaging and protein analysis. *Acs Chemical Biology* **2008**, *3* (6), 373-382.
70. Shu, X. K.; Lev-Ram, V.; Deerinck, T. J.; Qi, Y. C.; Ramko, E. B.; Davidson, M. W.; Jin, Y. S.; Ellisman, M. H.; Tsien, R. Y., A Genetically Encoded Tag for Correlated Light and Electron Microscopy of Intact Cells, Tissues, and Organisms. *Plos Biology* **2011**, *9* (4).
71. Walling, M. A.; Novak, J. A.; Shepard, J. R. E., Quantum Dots for Live Cell and In Vivo Imaging. *Int. J. Mol Sci* **2009**, *10* (2), 441-491.
72. Alivisatos, A. P.; Gu, W. W.; Larabell, C., Quantum dots as cellular probes. *Annual Review of Biomedical Engineering* **2005**, *7*, 55-76.
73. Kim, Y. G.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S., Efficient site-specific Labeling of proteins via cysteines. *Bioconjug Chem* **2008**, *19* (3), 786-791.
74. Lemke, E. A., Site-specific labeling of proteins for single-molecule FRET measurements using genetically encoded ketone functionalities. *Methods Mol Biol* **2011**, *751*, 3-15.
75. Brustad, E. M.; Lemke, E. A.; Schultz, P. G.; Deniz, A. A., A General and Efficient Method for the Site-Specific Dual-Labeling of Proteins for Single Molecule Fluorescence Resonance Energy Transfer. *J Am Chem Soc* **2008**, *130* (52), 17664.
76. Beatty, K. E.; Xie, F.; Wang, Q.; Tirrell, D. A., Selective dye-labeling of newly synthesized proteins in bacterial cells. *J Am Chem Soc* **2005**, *127* (41), 14150-14151.
77. Tyagi, S.; Lemke, E. A., Genetically Encoded Click Chemistry for Single-Molecule FRET of Proteins. In *Laboratory Methods in Cell Biology: Imaging*, Conn, P. M., Ed. 2012; Vol. 113, pp 169-187.
78. Nikic, I.; Kang, J. H.; Girona, G. E.; Aramburu, I. V.; Lemke, E. A., Labeling proteins on live mammalian cells using click chemistry. *Nat Protoc* **2015**, *10* (5), 780-791.
79. Zangoli, M.; Di Maria, F.; Zucchetti, E.; Bossio, C.; Antognazza, M. R.; Lanzani, G.; Mazzaro, R.; Corticelli, F.; Baroncini, M.; Barbarella, G., Engineering thiophene-based nanoparticles to induce phototransduction in live cells under illumination. *Nanoscale* **2017**, *9* (26), 9202-9209.
80. Rozkiewicz, D. I.; Gierlich, J.; Burley, G. A.; Gutschmiedl, K.; Carell, T.; Ravoo, B. J.; Reinhoudt, D. N., Transfer printing of DNA by "Click" chemistry. *ChemBioChem* **2007**, *8* (16), 1997-2002.
81. Gramlich, P. M. E.; Warncke, S.; Gierlich, J.; Carell, T., Click-click-click: Single to triple modification of DNA. *Angew Chem Int. Ed.* **2008**, *47* (18), 3442-3444.
82. Seo, T. S.; Li, Z. M.; Ruparel, H.; Ju, J. Y., Click chemistry to construct fluorescent oligonucleotides for DNA sequencing. *J Org Chem* **2003**, *68* (2), 609-612.
83. Braun, T.; Drescher, M.; Summerer, D., Expanding the Genetic Code for Site-Directed Spin-Labeling. *Int. J. Mol Sci* **2019**, *20* (2).
84. Zhu, Z. R.; Chao, J.; Yu, H.; Waggoner, A. S., Directly labeled DNA probes using fluorescent nucleotides with different length linkers. *Nucleic Acids Res* **1994**, *22* (16), 3418-3422.
85. Vamosi, G.; Gohlke, C.; Clegg, R. M., Fluorescence characteristics of 5-carboxytetramethylrhodamine linked covalently to the 5' end of oligonucleotides: Multiple conformers of single-stranded and double-stranded dye-DNA complexes. *Biophys J* **1996**, *71* (2), 972-994.
86. Chow, C. S.; Bogdan, F. M., A Structural Basis for RNA-Ligand Interactions. *Chem Rev* **1997**, *97* (5), 1489-1514.

87. Neville, D. M.; Davies, D. R., Interaction Of Acridine Dyes With DNA - an X-ray diffraction and optical investigation. *J Mol Biol* **1966**, *17* (1), 57-&.
88. Sayed, M.; Krishnamurthy, B.; Pal, H., Unraveling multiple binding modes of acridine orange to DNA using a multispectroscopic approach. *Phys Chem Chem Phys* **2016**, *18* (35), 24642-24653.
89. Kataev, E. A.; Shumilova, T. A.; Fiedler, B.; Anacker, T.; Friedrich, J., Understanding Stacking Interactions between an Aromatic Ring and Nucleobases in Aqueous Solution: Experimental and Theoretical Study. *J Org Chem* **2016**, *81* (15), 6505-6514.
90. Steffen, F. D.; Sigel, R. K. O.; Borner, R., An atomistic view on carbocyanine photophysics in the realm of RNA. *Phys Chem Chem Phys* **2016**, *18* (42), 29045-29055.
91. Iqbal, A.; Wang, L.; Thompson, K. C.; Lilley, D. M. J.; Norman, D. G., The structure of cyanine 5 terminally attached to double-stranded DNA: Implications for FRET studies. *Biochem* **2008**, *47* (30), 7857-7862.
92. Liu, Y. J.; Lilley, D. M. J., Crystal Structures of Cyanine Fluorophores Stacked onto the End of Double-Stranded RNA. *Biophys J* **2017**, *113* (11), 2336-2343.
93. Norman, D. G.; Grainger, R. J.; Uhrin, D.; Lilley, D. M. J., Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies. *Biochem* **2000**, *39* (21), 6317-6324.
94. Ouellet, J.; Schorr, S.; Iqbal, A.; Wilson, T. J.; Lilley, D. M. J., Orientation of Cyanine Fluorophores Terminally Attached to DNA via Long, Flexible Tethers. *Biophys J* **2011**, *101* (5), 1148-1154.
95. Urnavicius, L.; McPhee, S. A.; Lilley, D. M. J.; Norman, D. G., The Structure of Sulfoindocarbocyanine 3 Terminally Attached to dsDNA via a Long, Flexible Tether. *Biophys J* **2012**, *102* (3), 561-568.
96. Neubauer, H.; Gaiko, N.; Berger, S.; Schaffer, J.; Eggeling, C.; Tuma, J.; Verdier, L.; Seidel, C. A. M.; Griesinger, C.; Volkmer, A., Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy. *J Am Chem Soc* **2007**, *129* (42), 12746-12755.
97. Kupstat, A.; Ritschel, T.; Kumke, M. U., Oxazine Dye-Conjugated DNA Oligonucleotides: Forster Resonance Energy Transfer in View of Molecular Dye-DNA Interactions. *Bioconjug Chem* **2011**, *22* (12), 2546-2557.
98. Heinlein, T.; Knemeyer, J.-P.; Piester, O.; Sauer, M., Photoinduced electron transfer between fluorescent dyes and guanosine residues in DNA-hairpins. *J Phys Chem B* **2003**, *107* (31), 7957-7964.
99. Li, X.; Zhu, R.; Yu, A.; Zhao, X. S., Ultrafast Photoinduced Electron Transfer between Tetramethylrhodamine and Guanosine in Aqueous Solution. *J Phys Chem B* **2011**, *115* (19), 6265-6271.
100. Ranjit, S.; Levitus, M., Probing the Interaction Between Fluorophores and DNA Nucleotides by Fluorescence Correlation Spectroscopy and Fluorescence Quenching. *J Photochem Photobiol* **2012**, *88* (4), 782-791.
101. Rusinova, E.; Tretyachenko-Ladokhina, V.; Vele, O. E.; Senear, D. F.; Ross, J. B. A., Alexa and Oregon Green dyes as fluorescence anisotropy probes for measuring protein-protein and protein-nucleic acid interactions. *Anal Biochem* **2002**, *308* (1), 18-25.
102. Gijsbers, A.; Nishigaki, T.; Sanchez-Puig, N., Fluorescence Anisotropy as a Tool to Study Protein-protein Interactions. *Jove-Journal of Visualized Experiments* **2016**, (116).

103. Sindbert, S.; Kalinin, S.; Hien, N.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Muller, S.; Seidel, C. A. M., Accurate Distance Determination of Nucleic Acids via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J Am Chem Soc* **2011**, *133* (8), 2463-2480.
104. Doose, S.; Neuweiler, H.; Sauer, M., Fluorescence Quenching by Photoinduced Electron Transfer: A Reporter for Conformational Dynamics of Macromolecules. *ChemPhysChem* **2009**, *10* (9-10), 1389-1398.
105. Doose, S.; Neuweiler, H.; Sauer, M., A close look at fluorescence quenching of organic dyes by tryptophan. *ChemPhysChem* **2005**, *6* (11), 2277-2285.
106. Vaiana, A. C.; Neuweiler, H.; Schulz, A.; Wolfrum, J.; Sauer, M.; Smith, J. C., Fluorescence quenching of dyes by tryptophan: Interactions at atomic detail from combination of experiment and computer simulation. *J Am Chem Soc* **2003**, *125* (47), 14564-14572.
107. Chen, T.; Fu, L.; Zu, L., Steady-state and time-resolved fluorescence of tetramethylrhodamine attached to DNA: correlation with DNA sequences. *Luminescence* **2013**, *28* (6), 860-864.
108. Chen, H. M.; Ahsan, S. S.; Santiago-Berrios, M. B.; Abruna, H. D.; Webb, W. W., Mechanisms of Quenching of Alexa Fluorophores by Natural Amino Acids. *J Am Chem Soc* **2010**, *132* (21), 7244-7245.
109. Noble, J. E.; Wang, L.; Cole, K. D.; Gaigalas, A. K., The effect of overhanging nucleotides on fluorescence properties of hybridising oligonucleotides labelled with Alexa-488 and FAM fluorophores. *Biophys Chem* **2005**, *113* (3), 255-263.
110. Kurata, S.; Kanagawa, T.; Yamada, K.; Torimura, M.; Yokomaku, T.; Kamagata, Y.; Kurane, R., Fluorescent quenching-based quantitative detection of specific DNA/RNA using a BODIPY (R) FL-labeled probe or primer. *Nucleic Acids Res* **2001**, *29* (6).
111. Torimura, M.; Kurata, S.; Yamada, K.; Yokomaku, T.; Kamagata, Y.; Kanagawa, T.; Kurane, R., Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Analytical Sciences* **2001**, *17* (1), 155-160.
112. Fries, J. R. Charakterisierung einzelner Moleküle in Lösung mit Rhodamin-Farbstoffen. Universität-Gesamthochschule Siegen 1998.
113. Lewis, F.; Letsinger, R.; Wasielewski, M., Dynamics of photoinduced charge transfer and hole transport in synthetic DNA hairpins. *Acc Chem Res* **2001**, *34* (2), 159-70.
114. Lakowicz, J. R., *Principles of Fluorescence Spectroscopy*. Third ed.; Springer: New York, 2006.
115. Seidel, C.; Orth, A.; Greulich, K. O., Electronic effects on the fluorescence of tyrosine in small peptides. *Photochem Photobiol* **1993**, *58* (2), 178-184.
116. Marme, N.; Knemeyer, J. P.; Sauer, M.; Wolfrum, J., Inter- and intramolecular fluorescence quenching of organic dyes by tryptophan. *Bioconjug Chem* **2003**, *14* (6), 1133-1139.
117. Neuweiler, H.; Doose, S.; Sauer, M., A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. *Proc Natl Acad Sci USA* **2005**, *102* (46), 16650-16655.
118. Ohya, Y.; Yabuki, K.; Ouchi, T., Sequence dependence of fluorescent quenching of chromophores covalently bonded to oligo-DNAs before and after duplex formation. *Supramolecular Chemistry* **2003**, *15* (2), 149-154.
119. Berger, S. Untersuchung der Einzelmoleküldynamik mit multidimensionaler Fluoreszenzspektroskopie. Georg-August-University Goettingen, 2001.

120. Johansson, M. K., Choosing reporter-quencher pairs for efficient quenching through formation of intramolecular dimers. *Methods Mol Biol* **2006**, *335*, 17-29.
121. Marras, S. A. E.; Kramer, F. R.; Tyagi, S., Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res* **2002**, *30* (21).
122. Giese, B., Long-distance electron transfer through DNA. *Annu Rev Biochem* **2002**, *71*, 51-70.
123. Vogelsang, J.; Cordes, T.; Tinnefeld, P., Single-molecule photophysics of oxazines on DNA and its application in a FRET switch. *Photochem. Photobiol. Sci* **2009**, *8* (4), 486-496.
124. Murphy, C. J.; Arkin, M. R.; Jenkins, Y.; Ghatlia, N. D.; Bossmann, S. H.; Turro, N. J.; Barton, J. K., Long-range photoinduced electron-transfer through a DNA helix. *Science* **1993**, *262* (5136), 1025-1029.
125. Mao, H. H.; Luo, G. H.; Zhan, Y. X.; Zhang, J.; Yao, S.; Yu, Y., The mechanism and regularity of quenching the effect of bases on fluorophores: the base-quenched probe method. *Analyst* **2018**, *143* (14), 3292-3301.
126. Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L., Nature of biological electron-transfer. *Nature* **1992**, *355* (6363), 796-802.
127. Pourtois, G.; Beljonne, D.; Cornil, J.; Ratner, M. A.; Bredas, J. L., Photoinduced electron-transfer processes along molecular wires based on phenylenevinylene oligomers: A quantum-chemical insight. *J Am Chem Soc* **2002**, *124* (16), 4436-4447.
128. Gilbert, M.; Albinsson, B., Photoinduced charge and energy transfer in molecular wires. *Chem Soc Rev* **2015**, *44* (4), 845-862.
129. Porath, D.; Bezryadin, A.; de Vries, S.; Dekker, C., Direct measurement of electrical transport through DNA molecules. *Nature* **2000**, *403* (6770), 635-638.
130. Shin, Y. G. K.; Newton, M. D.; Isied, S. S., Distance dependence of electron transfer across peptides with different secondary structures: The role of peptide energetics and electronic coupling. *J Am Chem Soc* **2003**, *125* (13), 3722-3732.
131. Stennett, E. M. S.; Ciuba, M. A.; Levitus, M., Photophysical processes in single molecule organic fluorescent probes. *Chem Soc Rev* **2014**, *43* (4), 1057-1075.
132. Hwang, H.; Kim, H.; Myong, S., Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proc Natl Acad Sci USA* **2011**, *108* (18), 7414-7418.
133. Nygren, J.; Svanvik, N.; Kubista, M., The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **1998**, *46* (1), 39-51.
134. Robertson, K. L.; Yu, L. P.; Armitage, B. A.; Lopez, A. J.; Peteanu, L. A., Fluorescent PNA probes as hybridization labels for biological RNA. *Biochem* **2006**, *45* (19), 6066-6074.
135. Hwang, H.; Myong, S., Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions. *Chem Soc Rev* **2014**, *43* (4), 1221-1229.
136. Stennett, E. M. S.; Ciuba, M. A.; Lin, S.; Levitus, M., Demystifying PIFE: The Photophysics Behind the Protein-Induced Fluorescence Enhancement Phenomenon in Cy3. *J Phys Chem Lett* **2015**, *6* (10), 1819-1823.
137. Maillard, J.; Rumble, C. A.; Furstenberg, A., Red-Emitting Fluorophores as Local Water-Sensing Probes. *Journal of Physical Chemistry B* **2021**, *125* (34), 9727-9737.
138. Maillard, J.; Klehs, K.; Rumble, C.; Vauthey, E.; Heilemann, M.; Furstenberg, A., Universal quenching of common fluorescent probes by water and alcohols. *Chemical Science* **2021**, *12* (4), 1352-1362.

139. Kozlov, A. G.; Lohman, T. M., Stopped-flow studies of the kinetics of single-stranded DNA binding and wrapping around the Escherichia coli SSB tetramer. *Biochem* **2002**, *41* (19), 6032-6044.
140. Morten, M. J.; Lopez, S. G.; Steinmark, I. E.; Rafferty, A.; Magennis, S. W., Stacking-induced fluorescence increase reveals allosteric interactions through DNA. *Nucleic Acids Res* **2018**, *46* (21), 11618-11626.
141. Widengren, J.; Schwille, P., Characterization of photoinduced isomerization and back-isomerization of the cyanine dye Cy5 by fluorescence correlation spectroscopy. *J Phys Chem A* **2000**, *104* (27), 6416-6428.
142. Harvey, B. J.; Perez, C.; Levitus, M., DNA sequence-dependent enhancement of Cy3 fluorescence. *Photochem. Photobiol. Sci* **2009**, *8* (8), 1105-1110.
143. Aramendia, P. F.; Negri, R. M.; Sanroman, E., Temperature-dependence of fluorescence and photoisomerization in symmetrical carbocyanines - Influence of medium viscosity and molecular-structure. *J Phys Chem* **1994**, *98* (12), 3165-3173.
144. Chibisov, A. K.; Zakharova, G. V.; Gorner, H.; Sogulyaev, Y. A.; Mushkalo, I. L.; Tolmachev, A. I., Photorelaxation processes in covalently-linked indocarbocyanine and thiocarbocyanine dyes. *J Phys Chem* **1995**, *99* (3), 886-893.
145. Harvey, B. J.; Levitus, M., Nucleobase-Specific Enhancement of Cy3 Fluorescence. *Journal of Fluorescence* **2009**, *19* (3), 443-448.
146. Massey, M.; Algar, W. R.; Krull, U. J., Fluorescence resonance energy transfer (FRET) for DNA biosensors: FRET pairs and Forster distances for various dye-DNA conjugates. *Analytica Chimica Acta* **2006**, *568* (1-2), 181-189.
147. Rashid, F.; Raducanu, V. S.; Zaher, M. S.; Tehseen, M.; Habuchi, S.; Hamdan, S. M., Initial state of DNA-Dye complex sets the stage for protein induced fluorescence modulation. *Nat Commun* **2019**, *10*.
148. Randolph, J. B.; Waggoner, A. S., Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. *Nucleic Acids Res* **1997**, *25* (14), 2923-2929.
149. Spiriti, J.; Binder, J. K.; Levitus, M.; van der Vaart, A., Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair. *Biophys J* **2011**, *100* (4), 1049-1057.
150. Milas, P.; Gamari, B. D.; Parrott, L.; Krueger, B. P.; Rahmanseresht, S.; Moore, J.; Goldner, L. S., Indocyanine Dyes Approach Free Rotation at the 3' Terminus of A-RNA: A Comparison with the 5' Terminus and Consequences for Fluorescence Resonance Energy Transfer. *J Phys Chem B* **2013**, *117* (29), 8649-8658.
151. Tyagi, S.; Kramer, F. R., Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **1996**, *14* (3), 303-308.
152. Knemeyer, J. P.; Marme, N.; Sauer, M., Probes for detection of specific DNA sequences at the single-molecule level. *Anal Chem* **2000**, *72* (16), 3717-3724.
153. Crockett, A. O.; Wittwer, C. T., Fluorescein-labeled oligonucleotides for real-time PCR: Using the inherent quenching of deoxyguanosine nucleotides. *Anal Biochem* **2001**, *290* (1), 89-97.
154. Nazarenko, I.; Lowe, B.; Darfler, M.; Ikonomi, P.; Schuster, D.; Rashtchian, A., Multiplex quantitative PCR using self-quenched primers labeled with a single fluorophore. *Nucleic Acids Res* **2002**, *30* (9), e37-e37.
155. Edman, L.; Mets, U.; Rigler, R., Conformational transitions monitored for single molecules in solution. *Proc Natl Acad Sci USA* **1996**, *93* (13), 6710-6715.

156. Eggeling, C.; Fries, J. R.; Brand, L.; Gunther, R.; Seidel, C. A. M., Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proc Natl Acad Sci USA* **1998**, *95* (4), 1556-1561.
157. Widengren, J.; Dapprich, J.; Rigler, R., Fast interactions between Rh6G and dGTP in water studied by fluorescence correlation spectroscopy. *J Chem Phys* **1997**, *216* (3), 417-426.
158. Felekyan, S.; Kühnemuth, R.; Kudryavtsev, V.; Sandhagen, C.; Becker, W.; Seidel, C. A. M., Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection. *Rev Sci Instrum* **2005**, *76* (8), 083104.
159. Nettels, D.; Hoffmann, A.; Schuler, B., Unfolded Protein and Peptide Dynamics Investigated with Single-Molecule FRET and Correlation Spectroscopy from Picoseconds to Seconds†. *J Phys Chem B* **2008**, *112* (19), 6137-6146.
160. Kim, J.; Doose, S.; Neuweiler, H.; Sauer, M., The initial step of DNA hairpin folding: a kinetic analysis using fluorescence correlation spectroscopy. *Nucleic Acids Res* **2006**, *34* (9), 2516-2527.
161. Schüttpelz, M.; Schöning, J. C.; Doose, S. r.; Neuweiler, H.; Peters, E.; Staiger, D.; Sauer, M., Changes in Conformational Dynamics of mRNA upon At GRP7 Binding Studied by Fluorescence Correlation Spectroscopy. *J Am Chem Soc* **2008**, *130* (29), 9507-9513.
162. Qu, P.; Yang, X.; Li, X.; Zhou, X.; Zhao, X., Direct measurement of the rates and barriers on forward and reverse diffusions of intramolecular collision in overhang oligonucleotides. *J Phys Chem B* **2010**, *114* (24), 8235-43.
163. Yin, Y. D.; Yang, L. J.; Zheng, G. Q.; Gu, C.; Yi, C. Q.; He, C.; Gao, Y. Q.; Zhao, X. S., Dynamics of spontaneous flipping of a mismatched base in DNA duplex. *Proc Natl Acad Sci USA* **2014**, *111* (22), 8043-8048.
164. Mansoor, S. E.; McHaourab, H. S.; Farrens, D. L., Mapping proximity within proteins using fluorescence spectroscopy. A study of T4 lysozyme showing that tryptophan residues quench bimane fluorescence. *Biochem* **2002**, *41* (8), 2475-2484.
165. Mansoor, S. E.; DeWitt, M. A.; Farrens, D. L., Distance Mapping in Proteins Using Fluorescence Spectroscopy: The Tryptophan-Induced Quenching (TrIQ) Method. *Biochem* **2010**, *49* (45), 9722-9731.
166. Callis, P. R., Binding phenomena and fluorescence quenching. II: Photophysics of aromatic residues and dependence of fluorescence spectra on protein conformation. *J Mol Struct* **2014**, *1077*, 22-29.
167. Raghuraman, H.; Chatterjee, S.; Das, A., Site-Directed Fluorescence Approaches for Dynamic Structural Biology of Membrane Peptides and Proteins. *Frontiers in Molecular Biosciences* **2019**, *6*.
168. Yao, X. J.; Parnot, C.; Deupi, X.; Ratnala, V. R. P.; Swaminath, G.; Farrens, D.; Kobilka, B., Coupling ligand structure to specific conformational switches in the beta(2)-adrenoceptor. *Nat Chem Biol* **2006**, *2* (8), 417-422.
169. Tsukamoto, H.; Farrens, D. L.; Koyanagi, M.; Terakita, A., The Magnitude of the Light-induced Conformational Change in Different Rhodopsins Correlates with Their Ability to Activate G Proteins. *J Biol Chem* **2009**, *284* (31), 20676-20683.
170. Janz, J. M.; Farrens, D. L., Rhodopsin activation exposes a key hydrophobic binding site for the transducin alpha-subunit C terminus. *J Biol Chem* **2004**, *279* (28), 29767-29773.
171. Brunette, A. M. J.; Farrens, D. L., Distance Mapping in Proteins Using Fluorescence Spectroscopy: Tyrosine, like Tryptophan, Quenches Bimane Fluorescence in a Distance-Dependent Manner. *Biochem* **2014**, *53* (40), 6290-6301.

172. Pantazis, A.; Westerberg, K.; Althoff, T.; Abramson, J.; Olcese, R., Harnessing photoinduced electron transfer to optically determine protein sub-nanoscale atomic distances. *Nat Commun* **2018**, *9*.
173. Neuweiler, H.; Schulz, A.; Vaiana, A.; Smith, J.; Kaul, S.; Wolfrum, J.; Sauer, M., Peptide-based molecular beacons for the detection of p53-autoantibodies in human sera. *Angew. Chemie* **2002**.
174. Nguyen, B.; Sokoloski, J.; Galletto, R.; Elson, E. L.; Wold, M. S.; Lohman, T. M., Diffusion of Human Replication Protein A along Single-Stranded DNA. *J Mol Biol* **2014**, *426* (19), 3246-3261.
175. Sokoloski, J. E.; Kozlov, A. G.; Galletto, R.; Lohman, T. M., Chemo-mechanical pushing of proteins along single-stranded DNA. *Proc Natl Acad Sci USA* **2016**, *113* (22), 6194-6199.
176. Fischer, C. J.; Maluf, N. K.; Lohman, T. M., Mechanism of ATP-dependent Translocation of E.coli UvrD Monomers Along Single-stranded DNA. *J Mol Biol* **2004**, *344* (5), 1287-1309.
177. Wu, C. G.; Bradford, C.; Lohman, T. M., Escherichia coli RecBC helicase has two translocase activities controlled by a single ATPase motor. *Nat Struct Mol Biol* **2010**, *17* (10), 1210.
178. Luo, G.; Wang, M.; Konigsberg, W. H.; Xie, X. S., Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proc Natl Acad Sci USA* **2007**, *104* (31), 12610-12615.
179. Sorokina, M.; Koh, H. R.; Patel, S. S.; Ha, T., Fluorescent Lifetime Trajectories of a Single Fluorophore Reveal Reaction Intermediates During Transcription Initiation. *J Am Chem Soc* **2009**, *131* (28), 9630-9631.
180. Myong, S.; Cui, S.; Cornish, P. V.; Kirchhofer, A.; Gack, M. U.; Jung, J. U.; Hopfner, K. P.; Ha, T., Cytosolic Viral Sensor RIG-I Is a 5'-Triphosphate-Dependent Translocase on Double-Stranded RNA. *Science* **2009**, *323* (5917), 1070-1074.
181. Markiewicz, R. P.; Vrtis, K. B.; Rueda, D.; Romano, L. J., Single-molecule microscopy reveals new insights into nucleotide selection by DNA polymerase I. *Nucleic Acids Res* **2012**, *40* (16), 7975-7984.
182. Morten, M. J.; Steinmark, I. E.; Magennis, S. W., Probing DNA Dynamics: Stacking-Induced Fluorescence Increase (SIFI) versus FRET. *ChemPhotoChem* **2020**.
183. Antony, E.; Tomko, E. J.; Xiao, Q.; Krejci, L.; Lohman, T. M.; Ellenberger, T., Srs2 Disassembles Rad51 Filaments by a Protein-Protein Interaction Triggering ATP Turnover and Dissociation of Rad51 from DNA. *Molecular Cell* **2009**, *35* (1), 105-115.
184. Autour, A.; Jeng, S. C. Y.; Cawte, A. D.; Abdolazadeh, A.; Galli, A.; Panchapakesan, S. S. S.; Rueda, D.; Ryckelynck, M.; Unrau, P. J., Fluorogenic RNA Mango aptamers for imaging small non-coding RNAs in mammalian cells. *Nat Commun* **2018**, *9*.
185. Dolgosheina, E. V.; Jeng, S. C. Y.; Panchapakesan, S. S. S.; Cojocar, R.; Chen, P. S. K.; Wilson, P. D.; Hawkins, N.; Wiggins, P. A.; Unrau, P. J., RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *Acs Chemical Biology* **2014**, *9* (10), 2412-2420.
186. Ma, L. L.; Yang, F.; Zheng, J., Application of fluorescence resonance energy transfer in protein studies. *J Mol Struct* **2014**, *1077*, 87-100.
187. Mazal, H.; Haran, G., Single-molecule FRET methods to study the dynamics of proteins at work. *Curr Opin Biomed Eng* **2019**, *12*, 8-17.
188. Wozniak, A. K.; Nottrott, S.; Kuhn-Holsken, E.; Schroder, G. F.; Grubmuller, H.; Luhrmann, R.; Seidel, C. A. M.; Oesterhelt, F., Detecting protein-induced folding of the U4

- snRNA kink-turn by single-molecule multiparameter FRET measurements. *RNA* **2005**, *11* (10), 1545-1554.
189. Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Hausteiner, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M., Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc Natl Acad Sci USA* **2003**, *100* (26), 15516-15521.
190. Olofsson, L.; Felekyan, S.; Doumazane, E.; Scholler, P.; Fabre, L.; Zwier, J. M.; Rondard, P.; Seidel, C. A. M.; Pin, J. P.; Margeat, E., Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy. *Nat Commun* **2014**, *5*.
191. Kang, Y. H.; Zhang, Y.; Liang, T.; Leung, Y. M.; Ng, B.; Xie, H. L.; Chang, N.; Chan, J.; Shyng, S. L.; Tsushima, R. G.; Gaisano, H. Y., ATP Modulates Interaction of Syntaxin-1A with Sulfonylurea Receptor 1 to Regulate Pancreatic beta-Cell K-ATP Channels. *J Biol Chem* **2011**, *286* (7), 5876-5883.
192. Williams, M. R.; Markey, J. C.; Doczi, M. A.; Morielli, A. D., An essential role for cortactin in the modulation of the potassium channel Kv1. 2. *Proc Natl Acad Sci USA* **2007**, *104* (44), 17412-17417.
193. Zheng, J.; Trudeau, M. C.; Zagotta, W. N., Rod cyclic nucleotide-gated channels have a stoichiometry of three CNGA1 subunits and one CNGB1 subunit. *Neuron* **2002**, *36* (5), 891-896.
194. Liu, L. W.; He, F. C.; Yu, Y. Y.; Wang, Y. X., Application of FRET Biosensors in Mechanobiology and Mechanopharmacological Screening. *Frontiers in Bioengineering and Biotechnology* **2020**, *8*.
195. Komatsu, N.; Aoki, K.; Yamada, M.; Yukinaga, H.; Fujita, Y.; Kamioka, Y.; Matsuda, M., Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Molecular Biology of the Cell* **2011**, *22* (23), 4647-4656.
196. Wang, Y. X.; Botvinick, E. L.; Zhao, Y. H.; Berns, M. W.; Usami, S.; Tsien, R. Y.; Chien, S., Visualizing the mechanical activation of Src. *Nature* **2005**, *434* (7036), 1040-1045.
197. Rupprecht, C.; Wingen, M.; Potzke, J.; Gensch, T.; Jaeger, K. E.; Drepper, T., A novel FbFP-based biosensor toolbox for sensitive in vivo determination of intracellular pH. *J Biotechnol* **2017**, *258*, 25-32.
198. Cui, C. H.; Shu, W.; Li, P. N., Fluorescence In situ Hybridization: Cell-Based Genetic Diagnostic and Research Applications. *Frontiers in Cell and Developmental Biology* **2016**, *4*.
199. Ried, T.; Landes, G.; Dackowski, W.; Klinger, K.; Ward, D. C., Multicolor fluorescence in situ hybridization for the simultaneous detection of probe sets for chromosomes 13, 18, 21, X and Y in uncultured amniotic fluid cells. *Human Molecular Genetics* **1992**, *1* (5), 307-313.
200. Wei, Y.; Xu, F.; Li, P. N., Technology-Driven and Evidence-Based Genomic Analysis for Integrated Pediatric and Prenatal Genetics Evaluation. *Journal of Genetics and Genomics* **2013**, *40* (1), 1-14.
201. Ning, Y.; Roschke, A.; Smith, A. C. M.; Macha, M.; Precht, K.; Riethman, H.; Ledbetter, D. H.; Flint, J.; Horsley, S.; Regan, R.; Kearney, L.; Knight, S.; Kvaloy, K.; Brown, W. R. A., A complete set of human telomeric probes and their clinical application. *Nature Genetics* **1996**, *14* (1), 86-89.
202. Shah, J.; Mark, O.; Weltman, H.; Barcelo, N.; Lo, W.; Wronska, D.; Kakkilaya, S.; Rao, A.; Bhat, S. T.; Sinha, R.; Omar, S.; O'Bare, P.; Moro, M.; Gilman, R. H.; Harris, N., Fluorescence In Situ Hybridization (FISH) Assays for Diagnosing Malaria in Endemic Areas. *Plos One* **2015**, *10* (9).

203. Schueder, F.; Stein, J.; Stehr, F.; Auer, A.; Sperl, B.; Strauss, M. T.; Schwille, P.; Jungmann, R., An order of magnitude faster DNA-PAINT imaging by optimized sequence design and buffer conditions. *Nat Meth* **2019**, *16* (11), 1101-+.
204. Schnitzbauer, J.; Strauss, M. T.; Schlichthaerle, T.; Schueder, F.; Jungmann, R., Super-resolution microscopy with DNA-PAINT. *Nat Protoc* **2017**, *12* (6), 1198-1228.
205. Inuma, R.; Ke, Y. G.; Jungmann, R.; Schlichthaerle, T.; Woehrstein, J. B.; Yin, P., Polyhedra Self-Assembled from DNA Tripods and Characterized with 3D DNA-PAINT. *Science* **2014**, *344* (6179), 65-69.
206. Samanta, S.; Gong, W. J.; Li, W.; Sharma, A.; Shim, I.; Zhang, W.; Das, P.; Pan, W. H.; Liu, L. W.; Yang, Z. G.; Qu, J. L.; Kim, J. S., Organic fluorescent probes for stochastic optical reconstruction microscopy (STORM): Recent highlights. *Coordination Chemistry Reviews* **2019**, *380*, 17-34.
207. Danial, J. S. H.; Aguib, Y.; Yacoub, M. H., Advanced fluorescence microscopy techniques for the life sciences. *Glob Cardiol Sci Pract* **2016**, *2016* (2), e201616.
208. Jabłoński, A., Über den Mechanismus der Photolumineszenz von Farbstoffphosphoren. *Zeitschrift für Physik* **1935**, *94* (1), 38-46.
209. Rehm, D.; Weller, A., Kinetics of Fluorescence Quenching by Electron and H-Atom Transfer. *Israel Journal of Chemistry* **1970**, *8* (2), 259-271.
210. Weller, A., Exciplex and radical pairs in photochemical electron-transfer. *Pure Appl. Chem* **1982**, *54* (10), 1885-1888.
211. Jungmann, R.; Avendaño, M.; Woehrstein, J.; Dai, M.; Shih, W.; Yin, P., Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat Methods* **2014**, *11* (3), 313-8.
212. Kopperger, E.; List, J.; Madhira, S.; Rothfischer, F.; Lamb, D.; Simmel, F., A self-assembled nanoscale robotic arm controlled by electric fields. *Science* **2018**, *359* (6373), 296-301.
213. Bartnik, K.; Barth, A.; Pilo-Pais, M.; Crevenna, A.; Liedl, T.; Lamb, D., A DNA Origami Platform for Single-Pair Förster Resonance Energy Transfer Investigation of DNA-DNA Interactions and Ligation. *J Am Chem Soc* **2020**, *142* (2), 815-825.
214. Grilley, D.; Soto, A. M.; Draper, D. E., Direct quantitation of Mg(2+)-RNA interactions by use of a fluorescent dye. In *Meth Enzymol*, Johnson, M. L.; Holt, J. M.; Ackers, G. K., Eds. 2009; Vol. 455, pp 71-94.
215. Zhu, R. X.; Li, X.; Zhao, X. S.; Yu, A. C., Photophysical Properties of Atto655 Dye in the Presence of Guanosine and Tryptophan in Aqueous Solution. *J Phys Chem B* **2011**, *115* (17), 5001-5007.
216. Kelley, S. O.; Barton, J. K., Electron Transfer Between Bases in Double Helical DNA. *Science* **1999**, *283* (5400), 375-381.
217. Schulze, A.; Beliu, G.; Helmerich, D. A.; Schubert, J.; Pearl, L. H.; Prodromou, C.; Neuweiler, H., Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism. *Nat Chem Biol* **2016**, *12* (8), 628-+.
218. Luitz, M. P.; Barth, A.; Crevenna, A. H.; Bomblies, R.; Lamb, D. C.; Zacharias, M., Covalent dye attachment influences the dynamics and conformational properties of flexible peptides. *Plos One* **2017**, *12* (5).
219. Neuweiler, H.; Johnson, C. M.; Fersht, A. R., Direct observation of ultrafast folding and denatured state dynamics in single protein molecules. *Proc Natl Acad Sci USA* **2009**, *106* (44), 18569-18574.

220. Lerner, E.; Ploetz, E.; Hohlbein, J.; Cordes, T.; Weiss, S., A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement-Forster-Type Resonance Energy Transfer (PIFE-FRET). *J Phys Chem B* **2016**, *120* (26), 6401-6410.
221. Ploetz, E.; Lerner, E.; Husada, F.; Roelfs, M.; Chung, S.; Hohlbein, J.; Weiss, S.; Cordes, T., Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. *Sci Rep* **2016**, *6*, 33257.
222. Levitus, M.; Ranjit, S., Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments. *Q Rev Biophys* **2011**, *44* (1), 123-151.
223. Cooper, M.; Ebner, A.; Briggs, M.; Burrows, M.; Gardner, N.; Richardson, R.; West, R., Cy3B (TM): Improving the performance of cyanine dyes. *Journal of Fluorescence* **2004**, *14* (2), 145-150.
224. Waggoner, A. S.; Mujumdar, R. B. Rigidized trimethine cyanine dyes. 6,133,445, Oct. 17, 2000.
225. Michie, M. S.; Gotz, R.; Franke, C.; Bowler, M.; Kumari, N.; Magidson, V.; Levitus, M.; Loncarek, J.; Sauer, M.; Schnermann, M. J., Cyanine Conformational Restraint in the Far-Red Range. *J Am Chem Soc* **2017**, *139* (36), 12406-12409.
226. Matikonda, S. S.; Hammersley, G.; Kumari, N.; Grabenhorst, L.; Glembockyte, V.; Tinnefeld, P.; Ivanic, J.; Levitus, M.; Schnermann, M. J., Impact of Cyanine Conformational Restraint in the Near-Infrared Range. *J Org Chem* **2020**, *85* (9), 5907-5915.
227. Eiring, P.; McLaughlin, R.; Matikonda, S. S.; Han, Z. Y.; Grabenhorst, L.; Helmerich, D. A.; Meub, M.; Beliu, G.; Luciano, M.; Bandi, V.; Zijlstra, N.; Shi, Z. D.; Tarasov, S. G.; Swenson, R.; Tinnefeld, P.; Glembockyte, V.; Cordes, T.; Sauer, M.; Schnermann, M. J., Targetable Conformationally Restricted Cyanines Enable Photon-Count-Limited Applications. *Angew Chem Int. Ed.* **2021**, *60* (51), 26685-26693.
228. Agam, G.; Gebhardt, C.; Popara, M.; Mächtel, R.; Folz, J.; Ambrose, B.; Chamachi, N.; Chung, S. Y.; Craggs, T. D.; de Boer, M.; Grohmann, D.; Ha, T.; Hartmann, A.; Hendrix, J.; Hirschfeld, V.; Hübner, C. G.; Hugel, T.; Kammerer, D.; Kang, H.-S.; Kapanidis, A. N.; Krainer, G.; Kramm, K.; Lemke, E.; Lerner, E.; Margeat, E.; Martens, K.; Michaelis, J.; Mitra, J.; Moya Muñoz, G. G.; Quast, R.; Robb, N. B.; Sattler, M.; Schlierf, M.; Schneider, J.; Schröder, T.; Sefer, A.; Tan, P. S.; Thurn, J.; Tinnefeld, P.; van Noort, J.; Weiss, S.; Wendler, N.; Zijlstra, N.; Barth, A.; Seidel, C. A. M.; Lamb, D. C.; Cordes, T., Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *bioRxiv* **2022**, 2022.08.03.502619.
229. Huang, Z. X.; Ji, D. M.; Wang, S. F.; Xia, A. D.; Koberling, F.; Patting, M.; Erdmann, R., Spectral identification of specific photophysics of Cy5 by means of ensemble and single molecule measurements. *J Phys Chem A* **2006**, *110* (1), 45-50.
230. Mishra, A.; Behera, R.; Behera, P.; Mishra, B.; Behera, G., Cyanines during the 1990s: A Review. *Chem Rev* **2000**, *100* (6), 1973-2012.
231. Wang, L.; Fan, J.; Qiao, X.; Peng, X.; Dai, B.; Wang, B.; Sun, S.; Zhang, L.; Zhang, Y., Novel asymmetric Cy5 dyes: Synthesis, photostabilities and high sensitivity in protein fluorescence labeling. *J Photochem Photobiol A* **2010**, *210* (2-3), 168-172.
232. Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferre-D'Amare, A. R., Transition state stabilization by a catalytic RNA. *Science* **2002**, *298* (5597), 1421-1424.
233. D M J Lilley, a.; Clegg, R. M., The Structure of the Four-Way Junction in DNA. *Annual Review of Biophysics and Biomolecular Structure* **1993**, *22* (1), 299-328.

234. Heinlein, T.; Knemeyer, J.-P.; Piestert, O.; Sauer, M., Photoinduced electron transfer between fluorescent dyes and guanosine residues in DNA-hairpins. *J Phys Chem B* **2003**, *107* (31), 7957-7964.
235. Li, X.; Zhu, R.; Yu, A.; Zhao, X. S., Ultrafast Photoinduced Electron Transfer between Tetramethylrhodamine and Guanosine in Aqueous Solution. *J Phys Chem B* **2011**, *115* (19), 6265-6271.
236. Zhong, D.; Zewail, A. H., Femtosecond dynamics of flavoproteins: charge separation and recombination in riboflavine (vitamin B2)-binding protein and in glucose oxidase enzyme. *Proc Natl Acad Sci U S A* **2001**, *98* (21), 11867-11872.
237. Kempe, D.; Schone, A.; Fitter, J.; Gabba, M., Accurate Fluorescence Quantum Yield Determination by Fluorescence Correlation Spectroscopy. *J Phys Chem B* **2015**, *119* (13), 4668-4672.
238. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Adariani, S. R.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H. T.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P. Y.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B. Y.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R. Y.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nat Meth* **2018**, *15* (9), 669+.
239. Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M., A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Meth* **2012**, *9* (12), 1218-U129.
240. Steffen, F. D.; Sigel, R. K. O.; Börner, R., FRETraj: integrating single-molecule spectroscopy with molecular dynamics. *Bioinformatics* **2021**, *37* (21), 3953-3955.
241. Muschielok, A., Michaelis, J. *FRETnps Tools*, Version 2.6.1; Munich, Germany, 2008.
242. Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M., Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr Opin Struct Biol* **2016**, *40*, 163-185.
243. Höfig, H.; Gabba, M.; Poblete, S.; Kempe, D.; Fitter, J., Inter-dye distance distributions studied by a combination of single-molecule FRET-filtered lifetime measurements and a weighted accessible volume (wAV) algorithm. *Molecules* **2014**, *19* (12), 19269-19291.
244. Greife, A.; Felekyan, S.; Ma, Q. J.; Gertzen, C. G. W.; Spomer, L.; Dimura, M.; Peulen, T. O.; Wöhler, C.; Häussinger, D.; Gohlke, H.; Keitel, V.; Seidel, C. A. M., Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: An MFIS-FRET and integrative modelling study. *Sci Rep* **2016**, Advance Article.
245. Barawkar, D. A.; Ganesh, K. N., Fluorescent d(CGCGAATTCGCG) - characterization of major groove polarity and study of minor-groove interactions through a major groove semantophore conjugate. *Nucleic Acids Res* **1995**, *23* (1), 159-164.
246. Frugier, M.; Schimmel, P., Subtle atomic group discrimination in the RNA minor groove. *Proc Natl Acad Sci U S A* **1997**, *94* (21), 11291-11294.
247. Vukovic, L.; Koh, H. R.; Myong, S.; Schulten, K., Substrate Recognition and Specificity of Double-Stranded RNA Binding Proteins. *Biochem* **2014**, *53* (21), 3457-3466.

248. Cheng, X. D.; Blumenthal, R. M., Finding a basis for flipping bases. *Structure* **1996**, *4* (6), 639-645.
249. Banavali, N. K.; MacKerell, A. D., Free energy and structural pathways of base flipping in a DNA GCGC containing sequence. *J Mol Biol* **2002**, *319* (1), 141-160.
250. Wozniak, A. K.; Schroder, G. F.; Grubmuller, H.; Seidel, C. A. M.; Oesterhelt, F., Single-molecule FRET measures bends and kinks in DNA. *Proc Natl Acad Sci U S A* **2008**, *105* (47), 18337-18342.
251. Gabrielian, A.; Pongor, S., Correlation of intrinsic DNA curvature with DNA property periodicity. *Febs Lett* **1996**, *393* (1), 65-68.
252. Goodsell, D. S.; Dickerson, R. E., Bending and curvature calculations in B-DNA. *Nucleic Acids Res* **1994**, *22* (24), 5497-5503.
253. Ulyanov, N. B.; James, T. L., [4] Statistical analysis of DNA duplex structural features. *Methods Enzymol.* **1995**, *261*, 90-120.
254. Zgarbova, M.; Otyepka, M.; Sponer, J.; Lankas, F.; Jurecka, P., Base Pair Fraying in Molecular Dynamics Simulations of DNA and RNA. *J Chem Theory Comput* **2014**, *10* (8), 3177-3189.
255. Andreatta, D.; Sen, S.; Lustres, J. L. P.; Kovalenko, S. A.; Ernsting, N. P.; Murphy, C. J.; Coleman, R. S.; Berg, M. A., Ultrafast dynamics in DNA: "fraying" at the end of the helix. *J Am Chem Soc* **2006**, *128* (21), 6885-6892.
256. Doroshenko, O. Accurate structural determination of biomolecules via single-molecule high-precision FRET measurements. Heinrich Heine Universität Düsseldorf, Düsseldorf, 2020.
257. Case, D. A.; Betz, R. M.; Botello-Smith, W.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A., AMBER 2016. University of California, San Francisco, 2016.
258. Amoros, D.; Ortega, A.; de la Torre, J. G., Prediction of Hydrodynamic and Other Solution Properties of Partially Disordered Proteins with a Simple, Coarse-Grained Model. *J Chem Theory Comput* **2013**, *9* (3), 1678-1685.
259. Ortega, A.; Amorós, D.; García de la Torre, J., Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophys J* **2011**, *101* (4), 892-898.
260. Bae, S. H.; Dyson, H. J.; Wright, P. E., Prediction of the Rotational Tumbling Time for Proteins with Disordered Segments. *J Am Chem Soc* **2009**, *131* (19), 6814-6821.
261. Loman, A.; Gregor, I.; Stutz, C.; Mund, M.; Enderlein, J., Measuring rotational diffusion of macromolecules by fluorescence correlation spectroscopy. *Photochem Photobiol Sci* **2010**, *9* (5), 627-636.
262. Mockel, C.; Kubiak, J.; Schillinger, O.; Kuhnemuth, R.; Della Corte, D.; Schroder, G. F.; Willbold, D.; Strodel, B.; Seidel, C. A. M.; Neudecker, P., Integrated NMR, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics. *J Phys Chem B* **2019**, *123* (7), 1453-1480.
263. Viegas, A.; Dollinger, P.; Verma, N.; Kubiak, J.; Viennet, T.; Seidel, C. A. M.; Gohlke, H.; Etkorn, M.; Kovacic, F.; Jaeger, K. E., Structural and dynamic insights revealing how lipase

- binding domain MD1 of *Pseudomonas aeruginosa* foldase affects lipase activation. *Sci Rep* **2020**, *10* (1).
264. Vardanyan, H. FRET restrained high-precision structural modeling of RNA junctions. Heinrich Heine Universität Düsseldorf, Düsseldorf, 2014.
265. Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Borner, R.; Chung, H. S.; Cordes, T.; Craggs, T. D.; Deniz, A. A.; Diao, J. J.; Fei, J. Y.; Gonzalez, R. L.; Gopich, I. V.; Ha, T.; Hanke, C. A.; Haran, G.; Hatzakis, N. S.; Hohng, S.; Hong, S. C.; Hugel, T.; Ingargiola, A.; Joo, C.; Kapanidis, A. N.; Kim, H. D.; Laurence, T.; Lee, N. K.; Lee, T. H.; Lemke, E. A.; Margeat, E.; Michaelis, J.; Michalet, X.; Myong, S.; Nettels, D.; Peulen, T. O.; Ploetz, E.; Razvag, Y.; Robb, N. C.; Schuler, B.; Soleimanejad, H.; Tang, C.; Vafabakhsh, R.; Lamb, D. C.; Seidel, C. A. M.; Weiss, S., FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *Elife* **2021**, *10*.
266. Chung, H. S.; Cellmer, T.; Louis, J. M.; Eaton, W. A., Measuring ultrafast protein folding rates from photon-by-photon analysis of single molecule fluorescence trajectories. *Chem Phys* **2013**, *422*, 229-237.
267. Dimura, M.; Peulen, T. O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H., Automated and optimally FRET-assisted structural modeling. *Nat Commun* **2020**, *11* (1).
268. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Adariani, S. R.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H. T.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P. Y.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B. Y.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R. Y.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat Meth* **2018**, *15* (9), 669-676.
269. Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S., Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* **2005**, *88* (4), 2939-2953.
270. Peulen, T. O.; Opanasyuk, O.; Seidel, C. A. M., Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J Phys Chem B* **2017**, *121* (35), 8211-8241.
271. Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wohrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M., Multiparameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 reverse transcriptase: primer/template complexes. *Proc Natl Acad Sci U S A* **2003**, *100* (4), 1655-1660.
272. Sanchez-Rico, C.; von Voithenberg, L. V.; Warner, L.; Lamb, D. C.; Sattler, M., Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. *Chem Eur J* **2017**, *23* (57), 14267-14277.
273. Dale, R. E.; Eisinger, J.; Blumberg, W. E., Orientational freedom of molecular probes - orientation factor in intra-molecular energy-transfer. *Biophys J* **1979**, *26* (2), 161-193.
274. Ivanov, V.; Li, M.; Mizuuchi, K., Impact of Emission Anisotropy on Fluorescence Spectroscopy and FRET Distance Measurements. *Biophys J* **2009**, *97* (3), 922-929.

275. Hellenkamp, B.; Wortmann, P.; Kandzia, F.; Zacharias, M.; Hugel, T., Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nat Meth* **2017**, *14* (2), 174-180.
276. Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmuller, H., Structural Heterogeneity and Quantitative FRET Efficiency Distributions of Polyprolines through a Hybrid Atomistic Simulation and Monte Carlo Approach. *Plos One* **2011**, *6* (5).
277. Harvey, B. J.; Levitus, M., Nucleobase-Specific Enhancement of Cy3 Fluorescence. *J Fluoresc* **2009**, *19* (3), 443-448.
278. Moré, J. J. In *The Levenberg-Marquardt algorithm: Implementation and theory*, Berlin, Heidelberg, Springer Berlin Heidelberg: Berlin, Heidelberg, 1978; pp 105-116.
279. Brouwer, A. M., Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure Appl Chem* **2011**, *83* (12), 2213-2228.
280. Arden-Jacob, J. Neue langwellige Xanthen-Farbstoffe für Fluoreszenzsonden und Farbstofflaser. Universität-Gesamthochschule Siegen, Shaker Verlag GmbH (10 Mar. 1993), 1992.
281. Kalinin, S.; Kuhnemuth, R.; Vardanyan, H.; Seidel, C. A. M., Note: A 4 ns hardware photon correlator based on a general-purpose field-programmable gate array development board implemented in a compact setup for fluorescence correlation spectroscopy. *Rev Sci Instrum* **2012**, *83* (9).
282. Rigler, R.; Mets, U., *Diffusion of single molecules through a Gaussian laser beam*. SPIE: 1993; Vol. 1921.
283. Kalmykov, Y. P., Calculation of orientational correlation functions for free anisotropic rotational diffusion revisited. *J Chem Phys* **2009**, *130* (13).
284. Wang, J. M.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* **2000**, *21* (12), 1049-1074.
285. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M., Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers. *Biophys J* **2007**, *92* (11), 3817-3829.
286. Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P., Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J Chem Theory Comput* **2011**, *7* (9), 2886-2902.
287. Hanke, C. A.; Gohlke, H., Force Field Dependence of Riboswitch Dynamics. *Computational Methods for Understanding Riboswitches* **2015**, *553*, 163-191.
288. Horn, H. W.; Swope, W. C.; Pitara, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T., Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* **2004**, *120* (20), 9665-9678.
289. Li, P. F.; Merz, K. M., Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J Chem Theory Comput* **2014**, *10* (1), 289-297.
290. Joung, I. S.; Cheatham, T. E., Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* **2008**, *112* (30), 9020-9041.
291. Spiegel, J. D.; Fulle, S.; Kleinschmidt, M.; Gohlke, H.; Marian, C. M., Failure of the IDA in FRET Systems at Close Inter-Dye Distances Is Moderated by Frequent Low $k(2)$ Values. *J Phys Chem B* **2016**, *120* (34), 8845-8862.

292. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. *J Chem Phys* **1993**, *98* (12), 10089-10092.
293. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *J Comput Phys* **1977**, *23* (3), 327-341.
294. Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput* **2012**, *8* (5), 1542-1555.
295. Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J Chem Theory Comput* **2013**, *9* (9), 3878-3888.
296. Le Grand, S.; Gotz, A. W.; Walker, R. C., SPFP: Speed without compromise-A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* **2013**, *184* (2), 374-380.
297. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, I.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. N. H.; Pedregosa, F.; van Mulbregt, P.; SciPy, C., SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Meth* **2020**, *17* (3), 261-272.
298. Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J. H.; Liles, K.; Chun, M. J.; Li, P.; Gohara, D. W.; Dolinsky, T.; Konecny, R.; Koes, D. R.; Nielsen, J. E.; Head-Gordon, T.; Geng, W. H.; Krasny, R.; Wei, G. W.; Holst, M. J.; McCammon, J. A.; Baker, N. A., Improvements to the APBS biomolecular solvation software suite. *Protein Science* **2018**, *27* (1), 112-128.
299. Kask, P.; Piksarv, P.; Pooga, M.; Mets, U.; Lippmaa, E., Separation of the rotational contribution in fluorescence correlation experiments. *Biophys. J.* **1989**, *55* (2), 213-220.
300. Ilari, A.; Savino, C., Protein Structure Determination by X-Ray Crystallography. In *Bioinformatics: Data, Sequence Analysis and Evolution*, Keith, J. M., Ed. Humana Press: Totowa, NJ, 2008; pp 63-87.
301. McPherson, A.; Gavira, J. A., Introduction to protein crystallization. *Acta Crystallographica Section F-Structural Biology Communications* **2014**, *70*, 2-20.
302. Ronda, L.; Bruno, S.; Bettati, S.; Storici, P.; Mozzarelli, A., From protein structure to function via single crystal optical spectroscopy. *Frontiers in Molecular Biosciences* **2015**, *2*.
303. Bai, X. C.; McMullan, G.; Scheres, S. H. W., How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences* **2015**, *40* (1), 49-57.
304. Frank, J., Time-resolved cryo-electron microscopy: Recent progress. *J Struc Biol* **2017**, *200* (3), 303-306.
305. Palmer, A. G., NMR characterization of the dynamics of biomacromolecules. *Chem Rev* **2004**, *104* (8), 3623-3640.
306. Markwick, P. R. L.; Malliavin, T.; Nilges, M., Structural Biology by NMR: Structure, Dynamics, and Interactions. *PLOS Computational Biology* **2008**, *4* (9), e1000168.
307. Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S., Toward dynamic structural biology: Two decades of single-molecule Forster resonance energy transfer. *Science* **2018**, *359* (6373), 288-+.

308. Sasmal, D. K.; Pulido, L. E.; Kasal, S.; Huang, J., Single-molecule fluorescence resonance energy transfer in molecular biology. *Nanoscale* **2016**, *8* (48), 19928-19944.
309. Schuler, B.; Lipman, E. A.; Eaton, W. A., Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **2002**, *419* (6908), 743-747.
310. Feng, Y. T.; Liu, W. J.; Mercade-Prieto, R.; Chen, X. D., Dye-protein interactions between Rhodamine B and whey proteins that affect the photoproperties of the dye. *J Photochem Photobiol A* **2021**, *408*.
311. Sharma, A.; Enderlein, J.; Kumbhakar, M., Photon Antibunching Reveals Static and Dynamic Quenching Interaction of Tryptophan with Atto-655. *J Phys Chem Lett* **2017**, *8* (23), 5821-5826.
312. Zosel, F.; Haenni, D.; Soranno, A.; Nettels, D.; Schuler, B., Combining short- and long-range fluorescence reporters with simulations to explore the intramolecular dynamics of an intrinsically disordered protein. *J Chem Phys* **2017**, *147* (15).
313. Nardecchia, I.; Lechelon, M.; Gori, M.; Donato, I.; Preto, J.; Floriani, E.; Jaeger, S.; Mailfert, S.; Marguet, D.; Ferrier, P.; Pettini, M., Detection of long-range electrostatic interactions between charged molecules by means of fluorescence correlation spectroscopy. *Physical Review E* **2017**, *96* (2).
314. Grabolle, M.; Pauli, J.; Brehm, R.; Resch-Genger, U., Structural control of dye-protein binding, aggregation and hydrophilicity in a series of asymmetric cyanines. *Dyes and Pigments* **2014**, *103*, 118-126.
315. Klymchenko, A. S.; Mely, Y., Fluorescent Environment-Sensitive Dyes as Reporters of Biomolecular Interactions. In *Fluorescence-Based Biosensors: From Concepts to Applications*, Morris, M. C., Ed. 2013; Vol. 113, pp 35-58.
316. Trabulo, S.; Cardoso, A. L.; Cardoso, A. M. S.; Duzgunes, N.; Jurado, A. S.; de Lima, M. C. P., Cell-penetrating peptide-based systems for nucleic acid delivery: A biological and biophysical approach. In *Nanomedicine: Infectious Diseases, Immunotherapy, Diagnostics, Antifibrotics, Toxicology and Gene Medicine*, Duzgunes, N., Ed. 2012; Vol. 509, pp 277-300.
317. Majumdar, A.; Mukhopadhyay, S., Fluorescence Depolarization Kinetics to Study the Conformational Preference, Structural Plasticity, Binding, and Assembly of Intrinsically Disordered Proteins. In *Intrinsically Disordered Proteins*, Rhoades, E., Ed. 2018; Vol. 611, pp 347-381.
318. Rosemond, S. N.; Hamadani, K. M.; Cate, J. H. D.; Marqusee, S., Modulating long-range energetics via helix stabilization: A case study using T4 lysozyme. *Protein Science* **2018**, *27* (12), 2084-2093.
319. Zheng, W. J.; Glenn, P., Probing the folded state and mechanical unfolding pathways of T4 lysozyme using all-atom and coarse-grained molecular simulation. *J Chem Phys* **2015**, *142* (3).
320. Elwell, M.; Schellman, J., Phage-T4 lysozyme - physical-properties and reversible unfolding. *Biochimica Et Biophysica Acta* **1975**, *386* (1), 309-323.
321. Mondal, J.; Ahalawat, N.; Pandit, S.; Kay, L. E.; Vallurupalli, P., Atomic resolution mechanism of ligand binding to a solvent inaccessible cavity in T4 lysozyme. *Plos Computational Biology* **2018**, *14* (5).
322. Lopez, C. J.; Yang, Z. Y.; Altenbach, C.; Hubbell, W. L., Conformational selection and adaptation to ligand binding in T4 lysozyme cavity mutants. *Proc Natl Acad Sci USA* **2013**, *110* (46), E4306-E4315.

323. Bradford, S. Y. C.; El Khoury, L.; Ge, Y. H.; Osato, M.; Mobley, D. L.; Fischer, M., Temperature artifacts in protein structures bias ligand-binding predictions. *Chemical Science* **2021**, *12* (34), 11275-11293.
324. Baase, W. A.; Liu, L. J.; Tronrud, D. E.; Matthews, B. W., Lessons from the lysozyme of phage T4. *Protein Science* **2010**, *19* (4), 631-641.
325. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **2001**, *98* (18), 10037-10041.
326. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **2004**, *32*, W665-W667.
327. Lerner, M.; Carlson, H., APBS plugin for PyMOL. *Ann Arbor: University of Michigan* **2006**, 522.
328. Armstrong, C. T.; Mason, P. E.; Anderson, J. L. R.; Dempsey, C. E., Arginine side chain interactions and the role of arginine as a gating charge carrier in voltage sensitive ion channels. *Sci Rep* **2016**, *6*.
329. Raschka, S.; Wolf, A. J.; Bemister-Buffington, J.; Kuhn, L. A., Protein-ligand interfaces are polarized: discovery of a strong trend for intermolecular hydrogen bonds to favor donors on the protein side with implications for predicting and designing ligand complexes. *Journal of Computer-Aided Molecular Design* **2018**, *32* (4), 511-528.
330. Dougherty, D. A., The Cation-pi Interaction. *Acc Chem Res* **2013**, *46* (4), 885-893.
331. Gallivan, J. P.; Dougherty, D. A., Cation-pi interactions in structural biology. *Proc Natl Acad Sci USA* **1999**, *96* (17), 9459-9464.
332. Ma, J. C.; Dougherty, D. A., The cation-pi interaction. *Chem Rev* **1997**, *97* (5), 1303-1324.
333. Wiederschain, G. Y., The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. *Biochem (Moscow)* **2011**, *76* (11), 1276-1276.
334. Matouschek, A., Protein unfolding - an important process in vivo? *Curr Opin Struct Biol* **2003**, *13* (1), 98-109.
335. Hemmen, K. Structure and conformational fluctuations of Phage T4 Lysozyme under native and denaturing conditions. Heinrich-Heine University, Düsseldorf, Düsseldorf, 2017.
336. Brouwer, A. M., Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure Appl. Chem* **2011**, *83* (12), 2213-2228.
337. Kudryavtsev, V.; Sikor, M.; Kalinin, S.; Mokranjac, D.; Seidel, C. A. M.; Lamb, D. C., Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements. *ChemPhysChem* **2012**, *13* (4), 1060-1078.
338. Kapanidis, A. N.; Lee, 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. *Proc Natl Acad Sci USA* **2004**, *101* (24), 8936-8941.
339. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015**, *11* (8), 3696-3713.
340. Graen, T.; Hoefling, M.; Grubmuller, H., AMBER-DYES: Characterization of Charge Fluctuations and Force Field Parameterization of Fluorescent Dyes for Molecular Dynamics Simulations. *J Chem Theory Comput* **2014**, *10* (12), 5505-5512.
341. Schepers, B.; Gohlke, H., AMBER-DYES in AMBER: Implementation of fluorophore and linker parameters into AmberTools. *J Chem Phys* **2020**, *152* (22).

342. Gassner, N. C.; Baase, W. A.; Mooers, B. H. M.; Busam, R. D.; Weaver, L. H.; Lindstrom, J. D.; Quillin, M. L.; Matthews, B. W., Multiple methionine substitutions are tolerated in T4 lysozyme and have coupled effects on folding and stability. *Biophys Chem* **2003**, *100* (1-3), 325-340.
343. Horn, H. W.; Swope, W. C.; Pitera, J. W., Characterization of the TIP4P-Ew water model: Vapor pressure and boiling point. *J Chem Phys* **2005**, *123* (19).
344. Gohlke, H.; Kiel, C.; Case, D. A., Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes. *J Mol Biol* **2003**, *330* (4), 891-913.
345. Förster, T., Experimentelle und theoretische Untersuchung des zwischenmolekularen Übergangs von Elektronenanregungsenergie. *Zeitschrift für Naturforschung A* **1949**, *4* (5), 321-327.
346. Jeschke, G., DEER distance measurements on proteins. In *Annual Review of Physical Chemistry, Vol 63*, Johnson, M. A.; Martinez, T. J., Eds. 2012; Vol. 63, pp 419-446.
347. Columbus, L.; Hubbell, W. L., A new spin on protein dynamics. *Trends in Biochemical Sciences* **2002**, *27* (6), 288-295.
348. Cai, Q.; Kusnetzow, A. K.; Hideg, K.; Price, E. A.; Haworth, I. S.; Qin, P. Z., Nanometer distance measurements in RNA using site-directed spin labeling. *Biophys J* **2007**, *93* (6), 2110-2117.
349. Hagelueken, G.; Abdullin, D.; Schiemann, O., mtsslSuite: Probing biomolecular conformation by spin-labeling studies. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt A*, Qin, P. Z.; Warncke, K., Eds. 2015; Vol. 563, pp 595-622.
350. Brunger, A. T.; Strop, P.; Vrljic, M.; Chu, S.; Weninger, K. R., Three-dimensional molecular modeling with single molecule FRET. *J Struct Biol* **2011**, *173* (3), 497-505.
351. Klose, D.; Klare, J. P.; Grohmann, D.; Kay, C. W. M.; Werner, F.; Steinhoff, H. J., Simulation vs. reality: A comparison of in silico distance predictions with DEER and FRET measurements. *Plos One* **2012**, *7* (6), e39492.
352. Choi, U. B.; Strop, P.; Vrljic, M.; Chu, S.; Brunger, A. T.; Weninger, K. R., Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. *Nat Struct Mol Biol* **2010**, *17* (3), 318-U84.
353. Hoefling, M.; Lima, N.; Haenni, D.; Seidel, C. A. M.; Schuler, B.; Grubmüller, H., Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and Monte Carlo approach. *Plos One* **2011**, *6* (5), e19791.
354. Best, R. B.; Hofmann, H.; Nettels, D.; Schuler, B., Quantitative interpretation of FRET experiments via molecular simulation: Force field and validation. *Biophys J* **2015**, *108* (11), 2721-2731.
355. Liao, J.-M.; Wang, Y.-T.; Chen, C.-L., Computer simulation to investigate the FRET application in DNA hybridization systems. *Phys Chem Chem Phys* **2011**, *13* (21), 10364-71.
356. Walczewska-Szewc, K.; Corry, B., Accounting for dye diffusion and orientation when relating FRET measurements to distances: three simple computational methods. *Phys Chem Chem Phys* **2014**, *16* (24), 12317-12326.
357. Kalinin, S.; Sisamakos, E.; Magennis, S. W.; Felekyan, S.; Seidel, C. A. M., On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *J Phys Chem B* **2010**, *114*, 6197-6206.

358. Levitus, M.; Ranjit, S., Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments. *Q Rev Biophys* **2010**, 1-29.
359. Abdullin, D.; Hagelueken, G.; Schiemann, O., Determination of nitroxide spin label conformations via dPELDOR and X-ray crystallography. *Phys Chem Chem Phys* **2016**, 18 (15), 10428-10437.
360. Muschielok, A.; Andrecka, J.; Jawhari, A.; Bruckner, F.; Cramer, P.; Michaelis, J., A nano-positioning system for macromolecular structural analysis. *Nat Meth* **2008**, 5 (11), 965-971.
361. Beckers, M.; Drechsler, F.; Eilert, T.; Nagy, J.; Michaelis, J., Quantitative structural information from single-molecule FRET. *Faraday Discussions* **2015**, 184, 117-129.
362. Brieger, A.; Plotz, G.; Hinrichsen, I.; Passmann, S.; Adam, R.; Zeuzem, S., C-Terminal Fluorescent Labeling Impairs Functionality of DNA Mismatch Repair Proteins. *Plos One* **2012**, 7 (2).
363. Matthews, B. W., Studies on protein stability with t4 lysozyme. *Advances in Protein Chemistry, Vol 46* **1995**, 46, 249-278.
364. Rennell, D.; Bouvier, S. E.; Hardy, L. W.; Poteete, A. R., Systematic mutation of bacteriophage-T4 lysozyme. *J Mol Biol* **1991**, 222 (1), 67-87.
365. Veenstra, D. L.; Kollman, P. A., Modeling protein stability: a theoretical analysis of the stability of T4 lysozyme mutants. *Protein Engineering* **1997**, 10 (7), 789-807.
366. Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A., Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J Biotechnol* **2001**, 86 (3), 163-80.
367. Kühnemuth, R.; Seidel, C. A. M., Principles of single molecule multiparameter fluorescence spectroscopy. *Single Molecules* **2001**, 2 (4), 251-254.
368. Kalinin, S.; Felekyan, S.; Antonik, M.; Seidel, C. A. M., Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer. *J Phys Chem B* **2007**, 111 (34), 10253-10262.
369. Lutz, H.-P.; Luisi, P. L., Correction for Inner Filter Effects in Fluorescence Spectroscopy. *Helvetica Chimica Acta* **1983**, 66 (7), 1929-1935.
370. Parker, C. A.; Barnes, W. J., Some experiments with spectrofluorimeters and filter fluorimeters. *Analyst* **1957**, 82 (978), 606-618.

Published paper

Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems.

Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M.

Current Opinion in Structural Biology **2016**, 40, 163-185.

Contributions: All authors wrote the manuscript. C.A.M.S and H.G supervised the project. M.D., T.O.P, C.A.H. and A.P. prepared figures. T.P., C.H. and A.P. simulated and analysed fluorescence data. M.D. performed and implemented FRET guided structural modelling. M.D. and T.P. analysed the modelling results.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Structural Biology

Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems

Mykola Dimura^{1,2,3}, Thomas O Peulen^{1,3}, Christian A Hanke¹,
Aiswaria Prakash¹, Holger Gohlke² and Claus AM Seidel¹



Förster Resonance Energy Transfer (FRET) combined with single-molecule spectroscopy probes macromolecular structure and dynamics and identifies coexisting conformational states. We review recent methodological developments in integrative structural modeling by satisfying spatial restraints on networks of FRET pairs (hybrid-FRET). We discuss procedures to incorporate prior structural knowledge and to obtain optimal distance networks. Finally, a workflow for hybrid-FRET is presented that automates integrative structural modeling and experiment planning to put hybrid-FRET on rails. To test this workflow, we simulate realistic single-molecule experiments and resolve three protein conformers, exchanging at 30 μ s and 10 ms, with accuracies of 1–3 Å RMSD versus the target structure. Incorporation of data from other spectroscopies and imaging is also discussed.

Addresses

¹Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

²Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Corresponding author: Seidel, Claus AM (cseidel@hhu.de)

³Contributed equally.

Current Opinion in Structural Biology 2016, 40:163–185

This review comes from a themed issue on **Biophysical and molecular biological methods**

Edited by **Petra Fromme** and **Andrej Sali**

For a complete overview see the [Issue](#) and the [Editorial](#)

<http://dx.doi.org/10.1016/j.sbi.2016.11.012>

0959-440/© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

The measurement of Förster resonance energy transfer (FRET) [1^{**},2^{**},3] from a donor (D) to an acceptor (A) fluorophore has become a popular biophysical method that can yield unique insights into the structure and the structural exchange dynamics of labeled biomolecular systems. FRET has applications in two major research areas. The first application makes use of the high time resolution and single-molecule capabilities to study the

kinetics and to detect intermediates of exchanging systems with a limited number of DA pairs [4]. The second application utilizes FRET to study structures of biomolecules in solution using a larger number of DA pairs, to obtain detailed structural insights on biomolecules in solution [5]. The first approach is successful in correlating structural dynamics and function by directly mapping the timescales of exchange and the pathways between conformational states with biomolecules on the single-molecule (sm) level [4,6]. Various equilibrium processes were studied such as folding of proteins [7,8] or nucleic acids [9–11], assembly and disassembly of complexes [12–14], and enzymatic cycles [15,16]. A versatile set of single-molecule based measurements and analysis techniques allows one to cover a wide range of time scales ranging from picoseconds [17], tens of nanoseconds [18,19], microseconds [13,20^{*},21^{*},22,23^{*},24^{*}] to milliseconds and seconds [25^{*},26,27^{**}]. Moreover, perturbation techniques such as temperature modulation [28] and microfluidic mixers for non-equilibrium experiments [29^{*}] widen the use of smFRET experiments even further.

In the second approach, FRET is applied as a spectroscopic ruler for quantitative structural studies by measuring a larger set of single-pair DA pairs, which is necessary for interrogating the most important degrees of freedom of the system (i.e., the data sparseness must be sufficiently low [30^{*}]). In the absence of prior structural knowledge, such a set should ideally be a connected network. Pioneering work on the ensemble and single-molecule level determined quantitative FRET-based structural models of essentially static systems such as polyprolines [5,31], various nucleic acids [32–34,35^{*}], and large biomolecular complexes [36,37]. This approach benefits from single-molecule methods, which increase the effective resolution, to discriminate among coexisting states in solution. Combining methodologies of both major application areas harbors the potential to resolve structural models of dynamically exchanging coexisting states.

FRET measurements can only report inter-dye distances, and the number of FRET pairs is limited. Integrated methods can leverage experimental data that by itself would be insufficient to determine structures with satisfactory accuracy [38,39]. Therefore, integrative structural modeling relying on molecular simulations as well as FRET data is essential for achieving FRET-based

atomistic structural models. In the following article, we refer to this hybrid approach as hybrid-FRET modeling. Due to the long-range FRET information (typically 20–100 Å), the hybrid-FRET modeling approach is especially suitable to characterize the tertiary, super-tertiary [40], and quaternary structures of biomolecular systems. Benchmark studies for FRET-restrained structural hybrid modeling of biomolecules and their complexes versus crystal structures have demonstrated the accuracy of the integrative approach [41,42]. A dramatic improvement in the precision and accuracy (root mean square deviation (RMSD) of 0.5 Å) is achieved by explicitly considering spatial distributions of dye positions, which greatly reduces uncertainties due to flexible dye linkers [42].

There are many reasons for further improving FRET-based structural modeling. In view of the huge dynamic time range from picoseconds to hours that can be covered by fluorescence, smFRET experiments and hybrid-FRET modeling are an especially promising combination to determine structure and dynamics simultaneously. Many other biophysical techniques have shown that a dynamic view on protein structures can give many important insights because the dynamic properties control their functions, as shown for various research areas such as signaling [43], enzymatic catalysis [44], (mis) folding [45,46], and modulation of allostery [47]. Several structured [48–50] and partially unstructured [51,52,53] proteins were characterized by the FRET approach. Even dynamic and large biomolecular systems were studied such as ligand-induced conformational changes of the membrane-bound SNARE complex [54] and the conformational dynamics of single HIV-1 envelope trimers on the surface of native virions [55].

The dynamic smFRET experiments combined with integrative modeling should be especially useful to detect and directly characterize the conformational heterogeneity of biomolecules in the presence of thermally excited conformational states with a lifetime in the microsecond time range, which complements recently improved NMR spectroscopy methods [56]. Moreover, there are many reasons (e.g., limited solubility, unsuitable molecular size, limited stability, inadequate purity, or large conformational heterogeneity) why the structures of biological macromolecules often prove intractable to mainstream structural biology methods (i.e., crystallography, NMR spectroscopy, and cryo-electron microscopy) [38]. Exploiting the time resolution, sensitivity, and selectivity of fluorescence, integrative structural modeling, combining FRET experiments with computer simulations and comparative protein structure modeling [57] as prior information is developing into a promising complementary method for structural biology of dynamic biomolecules.

This perspective article focuses on advances in FRET-based structural modeling and the application to dynamic

biomolecular systems. To this end, we first review FRET techniques that can yield information for structural modeling of biomolecules. We then summarize the current state of the art techniques for modeling the FRET dye label behavior, which is important for FRET-based structural modeling. Subsequently, we introduce a typical workflow for hybrid-FRET modeling, which is then applied in a test case. We demonstrate an application of fluorescence methods to a three-state model system to show the potential of FRET-based structural modeling and to assess the currently achievable precision and accuracy. For this test case, we introduce an automated procedure for determining the most informative single-pair FRET labeling sites and simulate experimental data under realistic conditions, which are then used for structural modeling. In the outlook, we discuss the potential directions for future improvements with respect to experimental techniques and computational modeling to improve the precision and accuracy of hybrid models. Finally, we discuss applications of hybrid-FRET methods in imaging to realize an integrated molecular fluorescence microscope combining optical and computational microscopy (i.e., coarse-grained and all-atom molecular (dynamics) simulations) [58] to describe suitably labeled biomolecular systems by atomistic structural models.

FRET-based structural modeling

Experimental techniques

Considering the conformational species (i), FRET between the tethered fluorophores D and A with the DA-distance $R_{DA}^{(i)}$ is characterized by the yield of energy transfer from D to A, usually called transfer efficiency $E^{(i)}$ or FRET efficiency (Eq. (1a)). This yield is the ratio of the rate constant $k_{RET}^{(i)}$ relative to all other processes depopulating the excited donor with the total rate constant $k_{D,0}$. Eq. (1b), derived by Theodor Förster [1,2], allows one to compute $E^{(i)}$ in units of the coupling constant R_0 (Förster radius). R_0 is specific for the used dye pair. It depends on the refractive index of the medium n , the spectral overlap integral of the D fluorescence and the A absorption spectrum J , the factor κ^2 for the relative orientation of the D and A dipole, and the D fluorescence quantum yield in absence of an acceptor, $\Phi_{F,D(0)}$ [3].

$$E^{(i)} = \frac{k_{RET}^{(i)}}{k_{D,0} + k_{RET}^{(i)}} = \frac{1}{1 + \left(R_{DA}^{(i)}/R_0\right)^6} \quad (1)$$

$$= \frac{F_A^{(i)}}{\gamma F_D^{(i)} + F_A^{(i)}} = 1 - \frac{\tau_{D(A)}^{(i)}}{\tau_{D(0)}^{(i)}}$$

The FRET-efficiency E can be determined by fluorescence intensities of the sensitized acceptor emission F_A and the quenched donor fluorescence F_D (Eq. (1c)). The

accuracy of E relies critically on several factors: (1) the purity of the samples, that is, the Donor-only reference for a correct calibration and the degree of labeling of the FRET sample; and (2) the properties of the labels and the setup. The setup and the samples should be carefully characterized for computing the experimental calibration factor $\gamma = (x_{D,A} g_A \Phi_{F,A}) / (g_D \Phi_{F,D(0)})$ given by the spectrum-dependent detection efficiencies of the D and A fluorescence (g_D and g_A), and the fluorescence quantum yields $\Phi_{F,D(0)}$ and $\Phi_{F,A}$ of D and A respectively. Often $\Phi_{F,D(0)}$ and $\Phi_{F,A}$ are assumed to be identical for all conformational species (for details and a more complex analysis, see [59**,60**,61*]). In single-molecule experiments, the excitation irradiances are usually high. Thus, saturation effects of the fluorophores due to formation of long lived dark-states must be considered in the fluorescence quantum yield. For example, the cyanine-based acceptor dyes such as Cy5 and Alexa647, undergo photo-induced cis-trans isomerization [62], such that the fraction of bright trans species $x_{b,A}$ drops significantly below one. $x_{b,A}$ can be obtained by analyzing the acceptor autocorrelation function of the FRET experiment [63,64*].

Time-resolved measurements of donor fluorescence decays $f(t)$ by, for example, time-correlated single photon counting (TCSPC) are useful for precise FRET measurements, because they offer three main advantages. (1) The FRET efficiency can be determined without instrumental calibrations via the excited state lifetimes τ of the D-only reference sample $\tau_{D(0)} = 1/k_{D,0}$ and the DA sample $\tau^{(i)}_{D(A)} = (1/k_{D,0} + 1/k^{(i)}_{RET})$ (Eq. (1d)). These correspond to the slopes of the decay curves (see Box 1, panel 1). (2) The joint decay analysis of the D-only reference $f_{D(0)}(t)$ and the FRET sample $f_{D(A)}(t)$ allows relating the FRET rate constant $k^{(i)}_{RET}$ directly to $R^{(i)}_{DA}$ (Eq. (2), for details see [65]). (3) The curvature of $f_{D(A)}(t)$ contains information on the heterogeneity of the FRET sample such that a distance distribution $x(k^{(i)}_{RET})$ with species fractions $x^{(i)}$ can directly be resolved (Eq. (2)).

$$f_{D(0)}(t) = e^{-(k_{D,0})t},$$

$$f_{D(A)}(t) = f_{D(0)}(t) \cdot \int_0^\infty x(k^{(i)}_{RET}) \cdot e^{-k^{(i)}_{RET}t} dk^{(i)}_{RET} \quad \text{with} \quad (2)$$

$$k^{(i)}_{RET} = k_{D,0}(\Phi_{F,D(0)}) \cdot \left(\frac{R_0(n, \kappa^2, \Phi_{F,D(0)})}{R_{DA}^{(i)}} \right)^6$$

Depending on the complexity of the sample, FRET experiments can be performed on the single-molecule, sub-ensemble (selectively averaged single-molecule events), and ensemble levels, where each technique has its advantages. Intensity-based ensemble FRET measurements are relatively easy to perform; however, these

experiments yield only average observables due to ensemble averaging over the mixture of the molecules in the sample (Eq. (3)), and the sample must be well-characterized (purity, degree of labeling, homogeneity, fluorescence quantum yields of D and A) [59**]. Ideally, for such a mixture of N species with fractions $x^{(i)}$ and transfer efficiencies $E^{(i)}$, the average transfer efficiency is given by:

$$E = \sum_{i=1}^N x^{(i)} E^{(i)} \quad (3)$$

Single-molecule measurements are currently the most common type of FRET experiments, because they can resolve distributions of FRET observables and provide kinetic information over 15 orders of magnitude in time. The main advantage of sm fluorescence spectroscopy is the ability to resolve static (multiple distinguishable static states) and dynamic (interconverting states) heterogeneities. There are two main measurement setup types for smFRET which differ significantly in their time resolution: (1) total internal reflection fluorescence (TIRF) microscopy [66*] with emCCD camera detection and (2) confocal detection with fast single-photon counting point detectors [67]. TIRF is widely used for immobilized molecules. The experimental time resolution is usually limited by the frame rate (~ 60 Hz) of the camera. Recently Farooq and Hohlbein [68] presented a generalized excitation scheme (sALEX) that combines the concepts of alternating laser excitation (ALEX) [69,70] and stroboscopic illumination [71,72] to improve the time resolution at least 10 fold. Confocal setups can be used for studying immobilized and freely diffusing molecules. Felekyan et al. [73] presented a dead-time free configuration with four detectors and two data acquisition boards, which enables time-resolved measurements over 15 orders of magnitude in time (Box 1). The immobilization of single molecules on a coated glass surface can be either directly accomplished via tags [66*] or indirectly via encapsulation in a liposome, which is immobilized afterwards [74].

Sm experiments with confocal detection of freely diffusing molecules are characterized by reproducible and stable detection efficiencies. In combination with pulsed excitation or pulsed interleaved excitation (PIE) [64*,75*] time-resolved fluorescence spectroscopy with multi-parameter fluorescence detection (MFD) [60**,76,77] can be conveniently realized for all FRET relevant fluorescence observables in a single measurement with the same sample. Most importantly, time-resolved anisotropy decay curves are obtained to estimate the uncertainty of the FRET orientation factor κ^2 [61*]. Additionally, a series of combined 2D fluorescence parameter histograms can be generated, which allows determining all necessary FRET calibration parameters [60**,64*]. Finally, MFD increases the species selectivity, which is essential for

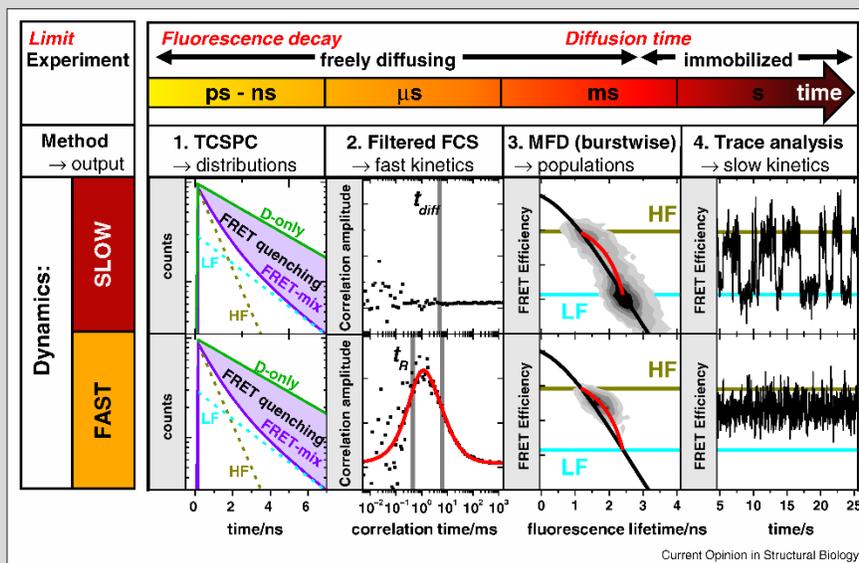
166 Biophysical and molecular biological methods

Box 1 A combination of fluorescence spectroscopy methodologies records dynamics over 15 orders of magnitude in time and captures a large variety of biologically relevant processes such as conformational transitions or folding/unfolding events [22,29*,60**,150].

The fluorescence lifetime of typical fluorophores decays in the pico-second to nanosecond range and sets the lower limit to interrogate conformational dynamics. The upper limit is set by the time a single molecule can be observed. Immobilized molecules can be studied for seconds while fluorescence bursts due to freely diffusing molecules are observed only for milliseconds in confocal setups.

Experimental information can be obtained by (1) time-correlated single photon counting (TCSPC) histograms, (2) fluorescence correlation spectroscopy (filtered FCS), (3) burst-wise single-molecule analysis, and (4) the analysis of FRET efficiency traces. The influence of the exchange dynamics on the observed data is outlined for a system with two FRET states, Low-FRET (LF) and High-FRET (HF), which are either in fast (sub milliseconds, lower row) or in slow (seconds, upper row) exchange.

- (1) In the analysis, fluorescence decays of FRET rate constants of HF and LF are averaged over the fluorescence lifetime of the dye (nanoseconds). Therefore, a joint analysis of a FRET sample and a D-only sample (shown in green) provides a static picture of a molecular ensemble and decomposes into HF and LF components shown as dashed dark yellow and cyan lines, respectively.
- (2) In FCS, fluorescent photons are correlated to provide kinetic information in the submicro-second to milliseconds regime. Filtered FCS (fFCS) additionally utilizes the photon distribution with respect to an excitation pulse [23*] to highlight the kinetic exchange between the HF and LF species by the species cross-correlation functions (SCC, black dots). If the kinetic exchange is significantly slower than the observation time (t_{diff} – diffusion time), no correlation is observed. Otherwise, a pronounced anti-correlation in the SCC is visible, which is analyzed (red curve) to yield the relaxation time (t_R) of the HF/LF exchange.
- (3) In burst-wise analysis, fluorescence is integrated over milliseconds to provide the FRET efficiency and fluorescence lifetime of freely diffusing single molecules. A 2D-histogram allows interrogating the kinetics between states and reveals dynamics by changes in the shapes. Theoretical static (black) and dynamic FRET lines (red) serve as references for the location of the FRET populations in the 2D parameter histograms. A shift of the observed population towards longer lifetimes with respect to the static FRET line indicates conformational dynamics.
- (4) By recording fluorescence intensities of immobilized or liposomal encapsulated single molecules over a longer time with a lower time resolution, processes significantly slower than the typical burst duration ($>ms$) are interrogated. FRET efficiency time traces of single molecules reveal stochastic jumps between HF and LF if the integration time of each frame is shorter than the dwell time of the state (upper plot); these jumps are averaged if the dwell time of the states are short (lower plot).

Box 1

Fluorescence timescales to study molecular kinetics using FRET.

sub-ensemble TCSPC of sub-populations, that is, selective averaging by integrating all single-molecule events of a resolved FRET species to reduce the data noise dramatically (for details, see Box 1).

To perform quantitative FRET measurements, all procedures for determining the calibration factor γ in Eq. (1) (TIRF [66*,78,79*,80], confocal [60**,64*]) rely on well characterized standards, which are usually DNA rulers

labeled with a FRET pair using the dyes of interest. A worldwide consortium of research groups on FRET is currently working to establish a set of common recommendations for measurement procedures, data analysis, and joint reference molecules to maximize the accuracy of the FRET observables measured in different laboratories.

Description of the labels

Both EPR spectroscopy with the double electron–electron resonance (DEER) method [81] and FRET spectroscopy are frequently used to obtain structural information on proteins by selectively coupling labels via flexible linkers (i.e., nitroxide spin-labels or donor and acceptor dyes, respectively). Thus, the labels can explore a large conformational space, such that a distribution of inter-label distances rather than a single inter-label distance is observed experimentally. In structural analysis, both techniques face the same problem of describing the localization of their respective labels and connecting the measured distances with structural information on the protein. This is a typical inverse problem, where both EPR [81–84] and FRET [42**,85**,86**] spectroscopies have developed similar solution strategies of assuming a structural model to compute a specific distance distribution between the labels and comparing the simulated and experimental values. Analogous to super-resolution microscopy, where the knowledge of the optical point spread function allows for localizing single emitters far beyond the optical resolution limit, a detailed knowledge of the label distribution for a given biomolecular structure is required to maximize the structural resolution of the respective technique. The so far used dye models are reviewed in this section, and their implementation in various toolkits for integrative FRET-based structural modeling is described in Section ‘Procedures for hybrid-FRET modeling’.

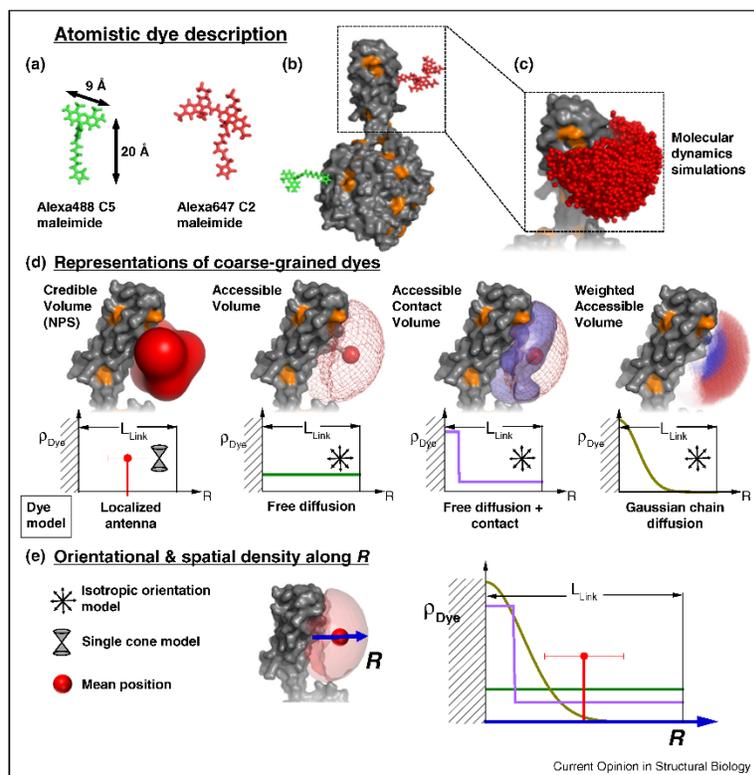
All-atom molecular dynamics (MD) simulations [33,87–93] seem to be the most appropriate solution to describe the label localization in integrative modeling (Figure 1a–c). However, while the relevant time-scale of fluorescence and biomolecular dynamics covers picoseconds to hours (Box 1), current MD simulations reach only the micro- to millisecond timescale. Furthermore, MD simulations are time-consuming and limited in their capability to sample the conformational space, in particular since the dye distribution changes with conformational changes of the biomolecule. MD simulations [87,89,94] as well as experimental fluorescence [61*,95,96*] and EPR [97] studies find long-lived conformational states, where the labels are trapped on the biomolecular surface. Trapping of cyanine dyes is known to change their fluorescence quantum yields [96*] such that this process is utilized by others for probing protein–nucleic acid interactions by protein-induced fluorescence enhancement (PIFE) [98,99*]. To avoid quenching and mobility reduction

by the surface effects, most commercially available dyes used for FRET experiments have long flexible linkers (~20 Å), and the fluorophore carries sulfonic acid groups for improved solubility. The large flexibility of the linker also ensures a high orientational mobility of the dyes such that the assumption of the isotropic average for the dipole orientation factor (i.e., $\kappa^2 = 2/3$) is feasible, which is an important factor determining the accuracy of distance calculations by FRET measurements. In conclusion, the atomistic treatment of dye labels, together with the biomolecular system, by MD simulations is helpful for analyzing protein label interactions. However, it is currently not feasible to use MD simulations as a rapid routine method to predict the conformational sampling of the labels, which is needed for finding the structural model within a large ensemble that is most consistent with FRET data, as shown below.

Computationally faster algorithms use a coarse-grained representation, e.g. the label is approximated by a flexibly linked sphere ([83,85**], see Figure 1a–c), to compute the sterically allowed accessible volume (AV), which reproduced MD simulations well [61*]. The use of a geometric search algorithm reduces the computational time for the AV of a single FRET-label to less than 50 milliseconds on a conventional desktop computer [42**]. The computed AV represents uniform population densities ρ_{Dye} of the spatial dye localizations (Figure 1d). The mesh represents the surface of the dye positions most distant from the attachment point and includes all sterically allowed dye positions. To compute fluorescence observables the spatial population densities $\rho_{Dye}(x,y,z)$ of the dye localizations is needed.

Two convenient assumptions for limiting cases were initially used for the models describing the spatial and orientational dye density in various hybrid-FRET modeling approaches: (1) The original dye model of the Nano Positioning System (NPS) assumed a label adopting a single position with a defined conformation of the linker and the fluorophore wobbling in a cone (*Localized antennal Single cone model*) [85**]. This means that all fluorophores of an ensemble of molecules are located at the same position (defined by a delta function δ) with respect to the macromolecule: $\rho_{Dye}(x,y,z) = \delta(x_{Dye},y_{Dye},z_{Dye})$. Since the position $(x_{Dye},y_{Dye},z_{Dye})$ of the dye is initially unknown, multiple measurements are used to localize it within a credible volume for a given confidence level (Figure 1d, left). (2) As an approximation of a freely diffusing fast rotating dye all orientations and all positions within the AV are equally populated; $\rho_{Dye}(x,y,z)$ is constant and the fluorophore is isotropically oriented within the AV (*Free diffusion/Iso model*) [42**,61*]. The effect of different dye models on the spatial population densities $\rho_{Dye}(x,y,z)$ along the vector R , pointing from the linker attachment site to the mean dye position in the AV, is visualized in the lower panel of Figure 1d.

Figure 1



Description of fluorescent labels in simulation. Treatment of dye labels in FRET-based modeling. **(a)** The donor and acceptor dyes, for example, Alexa488 C5 maleimide (green) and Alexa647 C2 maleimide (red), respectively, are typically attached to the biomolecule via long (~20 Å) flexible linkers. **(b)** Dye labels attached to the protein Atlastin-1 (grey surface) to give an impression of the sizes. Quenching amino acids on the protein surface are highlighted in orange. **(c)** Molecular dynamics simulations provide the spatial distribution of dye molecules. **(d)** Representation of coarse-grained dye labels. On the top different dye representations are shown. On the bottom the spatial density ρ_{Dye} along a vector R starting at the attachment point in the direction of the dye mean position is shown for the corresponding dye models. The original Nano Positioning System (NPS) [85**] assumes the accessible volume as *prior* information (uncertainty), which is reduced by a set of FRET measurements with fixed satellites resulting in an uncertainty distribution (red) of a putatively *fixed* localized dye (antenna). The orientation of the dyes follows a diffusion in a cone model highlighted by a pictogram. Following R , the dye is located at a specific position (vertical red line) with an uncertainty. Accessible volume (AV, mesh) models provide the sterically allowed space of the dye molecule attached to the protein as calculated by the FPS program [42**]. Here, the linkers of Alexa488 and Alexa647 are approximated as flexible tubes with width of $L_{width} = 4.5$ Å and length of $L_{link} = 20.5$ Å and $L_{link} = 21.0$ Å, respectively. Three radii were used to describe the dyes (5.0 Å, 4.5 Å, 1.5 Å) and (11.0 Å, 4.7 Å, 1.5 Å) for Alexa488 and Alexa647, respectively. The large sphere indicates the mean dye position. For a dye freely diffusing inside of its AV a uniform spatial distribution is assumed. The accessible and contact volume (ACV, [102]) provides a similar description as the AV, but defines an area close to the surface as contact volume (violet). In this figure the density ρ_{Dye} in the contact volume is weighted six times stronger and defined as part of the AV that is closer than 3 Å from the macromolecular surface. For the corresponding model, where a dye freely diffuses within the AV and its diffusion is hindered close to the surface, the spatial density ρ_{Dye} along R is approximated by a step function: $\rho_{Dye} (R < 3 \text{ Å}) = 6 \rho_{Dye} (R \geq 3 \text{ Å})$. The weighted accessible volume is a modification of AV where ρ_{Dye} along R is approximated by a Gaussian chain-inspired empirical weighting function [101]. To illustrate the effect of the proposed weighting of ρ_{Dye} the weighted AV is colored from blue (high density) to red (low density). **(e)** Comparison of the normalized spatial population densities ρ_{Dye} of the above dye models.

More recently, other dye models were suggested [100*,101,102]. To consider small fractions of fluorescent fluorophores trapped on the biomolecular surface [102], the classical AV description was extended (using the *Free diffusion/Iso model*) by defining an additional interacting surface layer (the contact volume of defined thickness is highlighted in violet in Figure 1d) with a preferential surface residence of the dye, which may be approximated by a step function. This surface preference is specific for the dye and the labeling site. This combination yields the *Free diffusion + Contact/Iso model* for an accessible and contact volume (ACV), which predicts the compaction of the spatial dye distribution due to trapped dye fractions without significant increase of the computational cost. Höfing et al. [101] introduced a weighted AV description, where the densities ρ_{Dye} are enriched by empirical weighting with a Gaussian function leading to a higher local dye surface preference but only close to the attachment point (*Gaussian-chain (GC) diffusion/Iso model*) (Figure 1d,e).

The comparison of the normalized densities ρ_{Dye} of the four dye models in Figure 1e shows that the mean dye positions of an ACV or a weighted AV are closer to the protein interface as compared to the AV with a uniform density. This dye model specific shift of the mean positions becomes important for the predicted inter-dye distances if the other dye of the FRET pair is attached on the opposite side of the protein. To avoid an arbitrary choice of dye models, it is important to define rational criteria based on experimental evidence through observable fluorescence parameters. For example, time-resolved fluorescence data map the translational and orientational dye diffusion and the resulting DA distance distributions by fluorescence quenching, anisotropy and FRET.

So far we have discussed four dye models that have been developed for synthetic organic fluorescent labels, which are usually not too large (Figure 1a,b) so that their steric requirements are small. However, if the more bulky fluorescent proteins (FPs), such as GFP and mCherry, are used as labels in fusion proteins, we have to account for the significant steric hindrance of the FPs and their linkers. For this case, self-avoiding walk (SAW) weighting of the linker configuration (*SAW-Diffusion/Iso model*) was applied in our group. The scaled weighting factors were determined by computationally tractable MD simulations for sampling configurations of an explicit linker/label combination alone, followed by calculations of conformational free energies to weight each linker/label configuration in the presence of the biomolecule [103]. We applied this dye model in live cell hybrid-FRET studies for analyzing the architecture of biomolecular complexes labeled with FPs ([103,104*], further details see outlook). Notably, this *SAW-Diffusion/Iso model* has similar effects on the conformational space in the AV as the rotamer

library model that is widely used for the analysis of DEER experiments in EPR spectroscopy [81]. Unlike the weighted AV description, the steric effects in the SAW-weighting lead to a strong reduction of ρ_{Dye} in proximity to the attachment site so that ρ_{Dye} has a maximum far away from the surface.

Error estimation

For quantitative FRET studies, a number of control tasks should be performed for an appropriate analysis and interpretation of the experimental observables [60**]: (1) analysis of the fluorescence properties of D or A by analyzing multiple fluorescence parameters with respect to quenching and trapping [37,60**], which could affect R_0 [105] and the calibration of the intensity-based experiments; (2) interpretation of discrete distributions of FRET efficiency E and anisotropy in confocal single-molecule experiments by photon distribution analysis, which allows separating structural heterogeneities from stochastic variations [106*,107]; (3) validation of the chosen dye model by analyzing (i) the linker-induced distance distributions of single FRET states by TCSPC or sub-ensemble TCSPC, (Eq. (2)), and (ii) the time-resolved anisotropy to scrutinize the dye mobility [61*] for verifying the assumption of $\kappa^2 = 2/3$; (4) verification by sub-ensemble TCSPC and FCS (Box 1) that a population with a single FRET efficiency is indeed a single species and that it is not narrowed by a dynamic exchange (i.e. that it is actually a mixture of fast inter-converting states); and (5) verification by functional assays (e.g. binding constants of ligands, catalytic activity, structural stability sensed by thermal or chemical denaturation) that the label does not disturb the biomolecule.

The analysis of FRET measurements provides the inter-dye distances R_{DA} as sparse input data. To take advantage of this information, rigorous error estimation is necessary for integrative modeling such that all uncertainties that have been considered correctly propagate to the final structural model [30*]. According to the error propagation rules, the overall uncertainties of the experimental inputs (the inter-dye distances R_{DA}) for the search of structural model contain three major contributions.

- (1) In single photon counting, the statistical relative error is inversely proportional to the square root of the registered photons that translate to the experimental distance uncertainty, $\Delta R_{DA}^2(E)$, via the experimentally determined FRET efficiencies E or FRET rate constants k_{RET} . Systematic error in $\Delta R_{DA}^2(E)$ is introduced primarily by two factors: (i) Preprocessing of experimental data by approximate fluorescence models (e.g. Eq. (2)) and (ii) Inaccurate instrumental calibrations and reference samples. Furthermore, if dynamic multi-component systems are studied, $\Delta R_{DA}^2(E)$ usually also contains the uncertainty due

- to data ambiguity and incoherence, resulting in FRET observables being correlated [30^{*}]. To obtain a distance set for a specific FRET species (i.e. conformer) from analyzing a large set of studied FRET pairs, species fractions and specific kinetic properties [27^{**}], which are global (joint) for the whole data set, are the best choices to assign FRET observables to the corresponding specific FRET species. The reason is that a global analysis of multiple independent datasets by a joint model minimizes the ambiguity by making use of correlations between otherwise independent model parameters. A global analysis procedure will be demonstrated below using a sample test case.
- (2) The uncertainty $\Delta R_{DA}^2(\kappa^2)$ of using the correct orientation factor κ^2 directly influences R_0 , so that the accuracy of the computed distances, R_{DA} , is affected. The error range of κ^2 can be estimated by anisotropy measurements [61^{*}].
 - (3) The uncertainty $\Delta R_{DA}^2(R_D, R_A)$ of using an appropriate dye model describing the spatial positions R_D, R_A of D and A, respectively, influences the accuracy of predicting appropriate distances from structural models.

In the end, all the errors above propagate to structural models obtained via sampling of the conformational space and define the accuracy of the structural models with respect to the true structure and the precision of the model ensemble.

Procedures for hybrid-FRET modeling

Due to the sparse information provided by experimental techniques such as FRET or EPR measurements, given the large number of degrees of freedom in atomistic models of biomolecules, a complete determination of biomolecular models with atomic resolution using only these experimental data is impossible. Fortunately, computational methods can provide complementary information, such as detailed structural features of biomolecules. At the same time, many computational methods also face challenges such as sufficient sampling and correct identification of near-native structures in a structural ensemble. The combination of computational methods with experimental input within a hybrid approach can facilitate generation and verification of the detailed biomolecular structures [108^{*}], because experimental FRET data can provide external information on the architecture of proteins and biomolecular complexes to guide the modeling procedure.

Although many FRET-based structural studies were performed since Förster's publication in 1948 [1^{**}], the obtained structural models were shown primarily as cartoons. In recent years, a number of integrative computational modeling methods [33,51,85^{**},86^{**},89,109] were developed to derive and present structural models in a

more quantitative manner. The high accuracy and performance of hybrid-FRET modeling was demonstrated in two benchmark studies with static structures [41^{*},42^{**}]. In the last decade, many applications for small and large complex biomolecular systems were published, albeit the models remained on a rather descriptive level. Considering flexibly linked dyes, currently three software toolkits, Nano Positioning System (NPS) [85^{**}], Crystallography & NMR System (CNS) [88,110] and FRET-restrained positioning and screening (FPS) [42^{**}], are publicly available; they implement different methodologies, use partially different assumptions and dye models. Therefore, common standards for FRET analysis, generally available joint toolkits with well-defined workflows and widely accepted procedures should be established to facilitate quantitative hybrid-FRET modeling.

The key to hybrid modeling lies in the fine interplay between the computer simulations and the experiments to accomplish the most effective synergies between the strengths of both sides. On the experimental side, accuracy is achieved by appropriate consideration of spatial dye distributions, while precision estimation stems from rigorous error analysis. Moreover, the establishment of a statistically appropriate quantitative scoring function for judging agreement between the structural models and FRET observables is far from trivial. Three major routes currently exist to find a structural model satisfying FRET observables best. First, in the screening approach used by FPS, computational methods are used to initially generate a conformational ensemble that is subsequently quantitatively evaluated in terms of its agreement with the experimental FRET data. Second, in the multibody docking approach, used by CNS and FPS, the integration of FRET data into the computational modeling is achieved by addition of experiment-based distances with corresponding uncertainties as restraints that define a harmonic potential for the fluorophore center positions treated as pseudoatoms rigidly connected to the biomolecule [42^{**},86^{**}]. The defined "FRET-forces" induce the docking process of the labeled domains and molecules. Many iterations (typically > 10 000) with distinct starting conditions (i.e. different randomly assigned orientations of domains and molecules, different relative conformations of flexible domains) are usually performed for each set of FRET distance restraints to cover the configurational space. As a third option, the Markov Chain Monte Carlo (MCMC) [111^{**}] approach, among others, allows circumventing the difficulties with geometric restraints. Examples for the use of MCMC guiding in conjunction with data from FRET experiments include the generation of an open conformation of Syntaxin 1 [48] and the determination of structures of large biomolecular complexes, such as a RNAP complex [36].

Appropriate dye models, in combination with quantitative FRET studies, are crucial for the accuracy of

integrative modeling. Previously [37,85**], studies utilizing the toolkit for the Nano Positioning System (NPS) have assumed the *Localized antenna/Single cone model* and have used Bayesian parameter estimation for three dimensional dye localization to extract structural information from a network of FRET measurements. The AV has been used as a *prior*, which is refined by experiments to resolve potential locations of the dyes as credible dye volume (*posterior*). The inclusion of more complex dye models into NPS has been introduced recently [100*,112]. In its current stage of development, NPS tests the consistency of a given structural model with DA distance sets by comparing the overlap of these credible dye volumes (*posterior*, depicted in Figure 1d) with the AVs (*prior*) of the considered structural model. Initially, CNS and FPS both used the *Free Diffusion/Iso model* to consider the distribution of flexibly linked dyes. In a benchmark study with FPS, docking a DNA primer-template to HIV-1 reverse transcriptase, Kalinin et al. have shown that it is crucial for the accuracy of the docked complex to explicitly consider the spatial dye distribution [42**]. CNS uses fluorophore center positions as pseudoatoms rigidly connected to the biomolecule. The position of the pseudoatom is taken as the average position of the fluorophore relative to the molecule as obtained from a molecular dynamics simulation [86**,88]. Owing to the explicit treatment of the dye distance distributions, FPS can serve as an appropriate interface for this experimental input. Thus additional dye models have been included in FPS meanwhile for improving the accuracy of hybrid-FRET modeling for specific combinations of systems and labels: (1) *Free diffusion/Iso model* for organic fluorophores tethered to nucleic acids [42**,61*], (2) *Free diffusion + Contact/Iso model* for organic fluorophores tethered to proteins (this work), and (3) *SAW-Diffusion/Iso model* for fusion proteins [103,104*]. Finally, Preus et al. [35*] developed the toolkit FRET-matrix for modeling FRET between probes possessing limited diffusional and rotational freedom. This toolkit is especially useful for localized fluorophores in nucleic acids as a replacement for one of the natural bases [113] giving the advantage of reporting from internal sites of interest.

Currently, structural models from hybrid-FRET modeling cannot be deposited in the Worldwide Protein data Bank (wwPDB) [114,115]. At present, the authors have two alternatives to archive the coordinates of suggested structural models: (1) provide them as supplementary information with the publication or (2) deposit them in the Model Archive (MA, www.modelarchive.org), which assigns a unique stable accession code (DOI) to each deposited model. The Model Archive is being developed following a community recommendation during a workshop on applications of protein models in biomedical research [57] as part of the Protein Model Portal (<http://www.proteinmodelportal.org>; [116]). Advantageously, it offers

the opportunity to apply consistent assessment and validation criteria to the complete set of structural models available for proteins. Moreover, it allows providing in depth information about the simulations performed and the parameters and constraints used. Therefore we decided to deposit all data of the presented test case study in the Model Archive with the DOI: 10.5452/ma-a2hbq.

Finally, it is important to note that the wwPDB is aware of the need for a large and general repository for structural models of biomolecular systems that have been obtained by integrative modeling using varied types of experimental data and theoretical information [38]. Establishing community-wide accepted standards for measuring, analyzing, and describing FRET data is an additional step necessary for organized data deposition.

Hybrid-FRET modeling on rails: a case study

Which information can be obtained by a hybrid-FRET study?

The aim of hybrid-FRET modeling is to describe macromolecules that potentially adopt multiple dynamically exchanging conformations in thermal equilibrium by three-dimensional structural models derived from prior structural knowledge and multiple quantitative FRET measurements [42**,88]. To provide an informative answer, our hybrid-FRET modeling procedure is organized such that a spectrum of possible solutions (structural arrangements) is collected from all structural background information (e.g. X-ray structures, homology/comparative models and/or at least educated guesses) whose suitability is under question. We use this initial information (*prior*) in step I of the workflow considering two perspectives. From the first perspective, experiments are designed to *confirm* our *prior* structural knowledge. If the experiments prove to be inconsistent with this *prior*, we consider the design also from the second perspective that the experimental information can be used to generate a three-dimensional structural model as *posterior* hypothesis. To test our *prior* (i.e. the structural background information), we seek DA pairs with the largest power in proving our null hypothesis that the initial structural information is incorrect (step I of the workflow). We identify such DA pairs by exploring a range of possible conformational motions of the system, given the *prior* knowledge. At this stage, unbiased and extensive coverage of the structural space is more important than the accuracy of structures or the density of the generated ensemble. To extract the most useful information from the system of interest, we established an iterative workflow for hybrid-FRET modeling consisting of five steps:

Step I: Defining specific questions about the biomolecular system of interest and initializing modeling by collecting *prior* knowledge to generate an initial structural ensemble and to determine the most informative DA pairs.

Step II: Performing calibrated FRET measurements and quantitative analysis.

Step III: Computational modeling to improve the sampling density of the initial ensemble guided by the obtained experimental FRET data. Identifying the best structures by screening and computing the precision of the obtained ensemble.

Step IV: Judging the agreement between the modeled structure and experimental data to decide whether the steps I to III should be repeated.

Step V: Assessing the quality of the obtained conformational ensemble with respect to the FRET data (precision and accuracy) and to the structural modeling (judgement by short-range stereochemical criteria).

Choosing informative DA pairs — an experimental design problem. The choice of DA pairs for FRET experiments is an essential initial step. It represents a classical experimental design problem [117] which influences the obtained structural model in terms of the precision and accuracy. More explicitly, the question is, which DA pairs should be measured such that the obtained information is maximized with respect to its power to discriminate structures of the initial ensemble. Two cases need to be distinguished for the selection of DA pairs. If structural pre-knowledge is unavailable or not fully considered, the localization of the fluorophores themselves is of interest. Thus, connected networks of DA pairs are needed and the dyes are localized by trilateration approaches [37,100*].

In the second case, if structural pre-knowledge is available, usually not all DA pairs are equally informative. Therefore it is advantageous to select a small set of most informative DA pairs because in real experiments, acquisition of reliable data is challenging due to the involvement of multiple steps, that is, mutagenesis, protein expression, purification, biochemical activity assays and measurements.

Here, we define distance sets to be the most informative if they lead to the highest expected precision of a structural model. To achieve high precision with a minimum set of DA pairs $\{DA^{(i)}\}$ for a well-planned experiment, we formalize the design of $\{DA^{(i)}\}$ to automate generation of FRET-based structural models. Thus, we establish a quality parameter, $\langle\Delta_{set}\rangle_{ref}$, for a given $\{DA^{(i)}\}$ by defining its expected precision (uncertainty). First, each conformer is used as a reference to compute the reference-specific precision Δ_{set} for $\{DA^{(i)}\}$. Here the precision corresponds to the weighted average C_{α} atom RMSD over all conformers with respect to the chosen reference conformer. The weight of the conformer is determined by its confidence in being a worse model compared to the reference

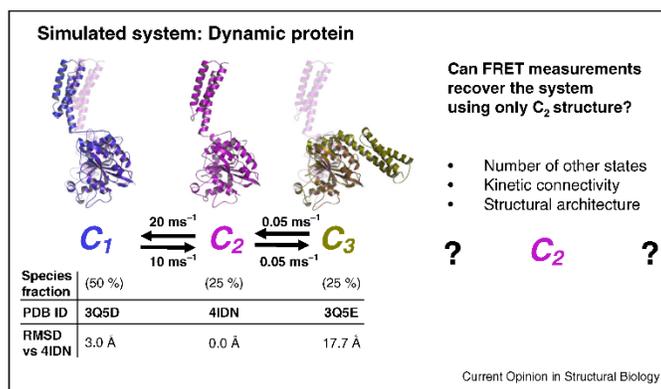
model as judged by squared, error weighted deviations. Second, the average over the reference-specific precisions yields $\langle\Delta_{set}\rangle_{ref}$. To find an informative distance set $\{DA^{(i)}\}$, we apply a greedy backward elimination feature selection algorithm [118] to minimize $\langle\Delta_{set}\rangle_{ref}$. In this algorithm, DA pairs that lead to the smallest decline of $\langle\Delta_{set}\rangle_{ref}$ (i.e. smallest increase in RMSD) were iteratively excluded from the set. As the set of possible DA pairs is limited to several thousand pairs, the computational burden of this algorithm is low enough to complete within ~ 24 h on a desktop PC.

For the first iteration of the hybrid-FRET modeling workflow (Figure 3), ten distances were selected (Figure 4a). In real experiments, suggested DA pairs might be discarded to minimize errors (see Section 'Error estimation'). Structural aspects relevant for biochemical activity (binding pockets, stabilizing salt bridges, and so on) and fluorescence properties (quenching amino acids) should be considered in the selection procedure to minimize the effects of the labels on the biomolecular function and to optimize their fluorescence properties.

Generating hybrid-FRET models for a heterogeneous system in the presence of fast dynamics for the test case *Atlastin-1*. To study the accuracy and precision of hybrid-FRET modeling, we designed a test case and simulated a protein with three exchanging conformers. The used structures of the conformers are structurally well-characterized by X-ray crystallography. We simulated typical data traces of single-molecule FRET experiments with multi-parameter fluorescence detection for typical conditions [60**,119]. In this way, the known X-ray structures can serve as unquestioned references for the assessment of the structural models obtained by hybrid-FRET modeling using our FPS toolkit.

We chose to study the cytoplasmic part (residues 1–446) of the large GTPase human Atlastin-1, which belongs to the dynamin superfamily. This part consists of two flexibly connected domains, the N-terminal GTPase domain (G domain) and the C-terminal middle domain, for which three distinct conformations in the functional cycle were found by X-ray crystallography. For the simulations of smFRET experiments, we chose the monomeric subunits of two crystallographic dimers (PDB-IDs 3Q5D [120] and 4IDN [121]) and a tetramer (PDB-ID 3Q5E [120]), depicted in Figure 2, named here C_1 , C_2 , and C_3 . To the best of our knowledge, nothing is known about conformational dynamics of Atlastin-1 in solution. Thus, in a Gedankenexperiment, we assumed the existence of a putative dynamic exchange equilibrium with the linear reaction scheme $C_1 \rightleftharpoons C_2 \rightleftharpoons C_3$ (Figure 2). The conformer C_3 differs significantly from the conformers C_1 and C_2 . In contrast, the structures of conformers C_1 and C_2 are similar to each other (RMSD 3.0 Å). Therefore, we simulated a fast exchange between C_1 and C_2 (kinetic

Figure 2



Presentation of the test case Atlastin. Overview of the benchmark test system. For the benchmark test, we used three crystal structures of Atlastin-1, which are termed C_1 , C_2 , and C_3 (PDB-ID: 3Q5D (chain A) [120], 4IDN (chain A) [121], and 3Q5E (chain A) [120], respectively). We simulate the kinetic exchange between C_1 and C_2 to be fast, and the exchange between C_2 and C_3 to be slow. Our task was to recover the number of other states, their kinetic connectivity, and their structures, given only the structure (PDB-ID: 4IDN) and information from smFRET spectroscopy.

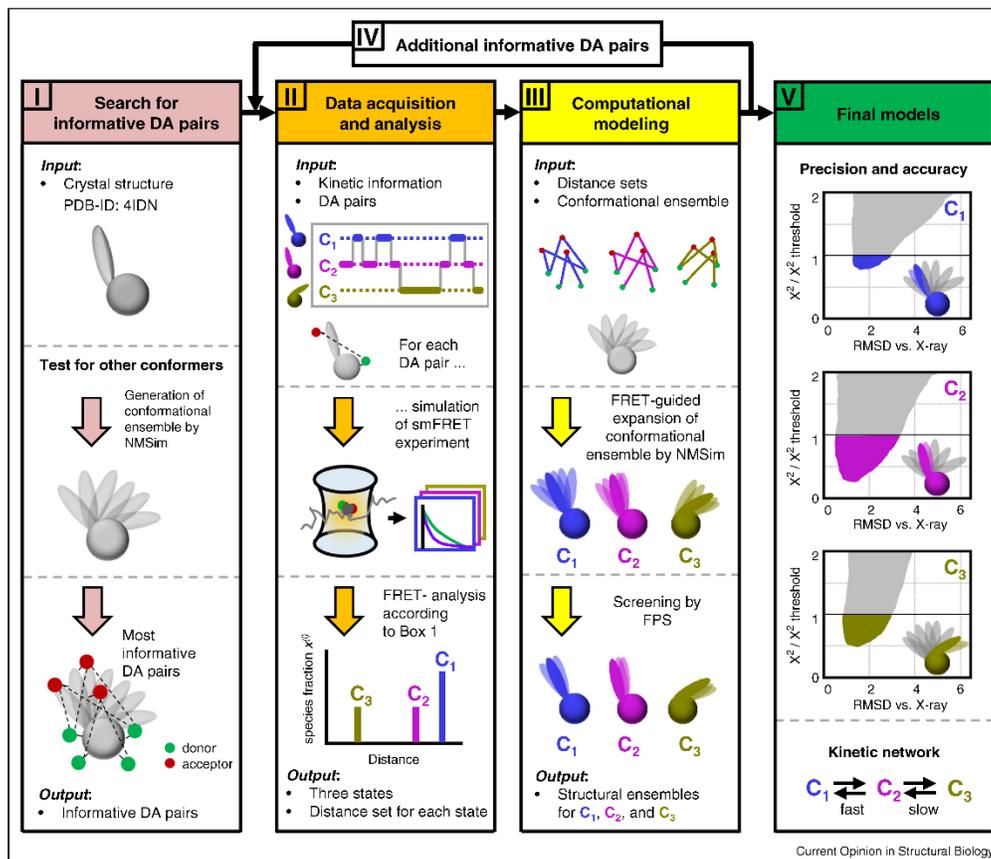
relaxation time $t_{R,1-2} = 30 \mu\text{s}$) and a slow exchange between C_2 and C_3 ($t_{R,2-3} = 10 \text{ ms}$). In the test case, we utilized the crystal structure (PDB-ID: 4IDN) as a *prior* for structural simulations. Note, however, that all experimental FRET observables were simulated from the three “true” crystallographic structures to test whether they could be recovered by hybrid-FRET modeling. Thus, we aimed at finding suitable structural models for C_1 , C_2 , and C_3 and recovering the kinetic exchange equilibrium by treating the simulated photon traces identical to single-molecule FRET experiments using a data generator (for details see [119]) that generated photon streams of typical single-molecule experiments by Brownian dynamics (BD) simulations.

In real experiments, fluorescence decays are complex due to DA distance distributions, brightness distributions due to the confocal excitation profile and experimental nuisances such as instrumental response function and detector dark counts. All these aspects are reproduced by the simulations of freely diffusing molecules, to generate realistic photon traces. Technical details are given in detail in [119] and are outlined below. As in real single photon counting experiments, the Poissonian statistics determines the experimental noise and thus the statistical errors of the subsequent analysis. The simulations correspond to ~ 3 measurement hours with current instrumentation in the Seidel group. In this test case, we want to study whether the simulated data of typical experimental quality allows us to recover accurate hybrid-FRET models and explore experimental limits on their precision.

FRET on rails: Step I. We started our hybrid-FRET modeling (Figure 3) by testing the null hypothesis that our *prior* will *not* be described by experiments. To design an optimal set of DA pairs for this test, we use the toolkit NMSim to generate an ensemble unbiased by experimental FRET data with 400 000 conformers (RMSD from PDB-ID: 4IDN up to 26 Å) as alternative candidates to the *prior*. NMSim is a normal mode-based geometric simulation approach for multiscale modeling of protein conformational changes that incorporates information about preferred directions of protein motions into a geometric simulation algorithm [122,123]. The obtained ensemble was clustered and used to select an optimal set of DA pairs according to three criteria. At first, all residues that are positioned on the protein surface and are located far enough from quenching amino acid residues (Trp, Tyr, His, Met) [124] were selected. Secondly, from all pairwise combinations of these labeling positions, those that result in average inter-dye distances $>30 \text{ Å}$ were selected. Thirdly, we determined the most informative distances from a matrix of inter-dye distances as described above.

FRET on rails: Steps II + III. Next, we have to test the null hypothesis using this initial distance set. In real situations, we have to prepare a set of samples and perform a series of measurements of freely diffusing molecules. However, in this test study, we replace the experimental data acquisition by simulations for a set of 10 hypothetical FRET samples with the mixtures of appropriately labeled Atlastin-1 conformers.

Figure 3

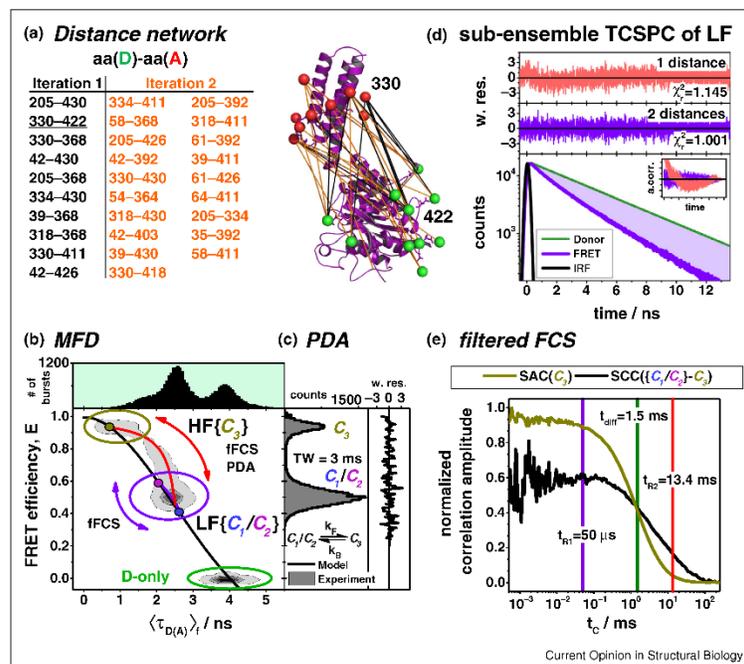


Complete workflow of FRET-restrained structural modeling. Workflow for the benchmark test using three distinct crystal structures of AtIastin-1. **Step I:** Starting from the crystal structure (PDB-ID: 4IDN), we generated a conformational ensemble by NMSim, which was subsequently clustered. From this, we derived the most informative donor-acceptor (DA) pairs, which can distinguish between these clusters. These DA pairs serve as one of the inputs for the next step. **Step II:** The previously determined DA pairs, crystal structures, and the kinetic scheme (Figure 2) were used to simulate smFRET experiments. In the simulations, the molecules randomly exchange their conformation, diffuse freely, and emit fluorescence if they diffuse through the confocal detection volume. The simulated photon trajectories were analyzed according to Box 1 to yield for each DA pair a conformer-specific distance and a species fraction x^0 . This results in a distance set for each of the three conformers, which is used in the next step. **Step III:** Structural modeling was based on the distance sets determined in step II and the conformational ensembles generated in step I. Using fitted distance sets, we applied FRET-guided NMSim to expand the conformational ensemble of each conformer. Subsequently, we screened the resulting conformational ensemble with the FPS software to identify conformers that are in agreement with the simulated experiment. **Step IV:** A second round repeating steps II and III, which considers additional DA pairs, improved the discrimination power and increased the resolution. **Step V:** The final models were chosen by selecting those that are within the 68% confidence level.

In *in silico* MFD-experiments, we simulate: (1) diffusion of individual molecules; (2) dynamic exchange of the states; (3) the confocal excitation profile and observation volume; (4) the photon emission as Poissonian process, which follows the conformer-specific fluorescence decay $f_{DA}(t)$

(Eq. (2)) described by the excited state depopulation with $k_{D,0}$ and FRET-induced quenching with a distributed k_{RET}^0 due to flexibly linked dyes; and (5) experimental nuisances such as the instrumental response function (IRF), background fluorescence and experimental calibration

Figure 4



Analysis of simulated FRET data. Fluorescence analysis workflow. **(a)** The distance network used to simulate the fluorescence experiments was constructed in two iterations. After a first analysis round (iteration 1, black), the network was extended by additional distances (iteration 2, orange). The analysis workflow is shown for the highlighted DA pair. **(b)** Confocal sm experiments with pulsed excitation (73.56 MHz) and time-resolved multi-parameter fluorescence detection (i.e. spectral and polarization resolved detection) [60**] were simulated for every DA pair using a data generator as described in detail in [119], that generated photon streams of typical single-molecule experiments by Brownian dynamics (BD) simulations. The fluorescence decay of D and A in absence of FRET was modeled by single exponential decays using a fluorescence lifetime of D of $\tau_{D(0)} = 4$ ns and A of $\tau_A = 1$ ns, respectively. Their time-resolved anisotropies were modeled using mean rotational correlation times of 1 ns. The fluorescence signal was modeled to yield $\gamma = 1$ as calibration parameter. The DA distance distributions were simulated by ACV simulations, which resulted in an average DA distance half-width of 10.4 Å. The optical detection profile of the setup was simulated by a 3D-Gaussian, which decayed at a radius of $w_0 = 0.5$ μ m (x, y direction) and $z_0 = 2.25$ μ m (z direction) to a value of $1/e^2$. To assure single-molecule conditions we simulate experiments with a mean number of 0.004 molecules in the focus which diffuse with a diffusion time $t_{diff} = 1.5$ ms as defined by FCS. First, single molecule bursts of the interconverting conformers (C_1 , C_2 , and C_3) were classified in 2D frequency histograms by their FRET-efficiency and intensity-weighted average donor fluorescence lifetime $\langle \tau_{D(A)} \rangle_f$. The number of molecules (bursts) is scaled from light grey (lowest) to black (highest). The 1D histogram of $\langle \tau_{D(A)} \rangle_f$ is given as a projection on the top. In the 2D-histogram, three peaks a high FRET (HF, yellow), low FRET (LF, violet) and a peak corresponding to FRET-inactive molecules (D-only, green) are visible. These peaks can be described by static and dynamic FRET lines. The static FRET line (black) describes the theoretical relationships between the FRET efficiency and $\langle \tau_{D(A)} \rangle_f$ for all molecules in the absence of protein dynamics. Deviation from the static FRET line towards longer fluorescence lifetimes indicates conformational dynamics. Dynamic FRET lines are defined by the limiting states of the dynamic processes. The FRET-efficiencies and $\langle \tau_{D(A)} \rangle_f$ of the conformers C_1 , C_2 , and C_3 are shown as yellow, magenta, and blue filled circles, respectively. Dynamic FRET lines of the C_1 - C_2 transition and a C_3 - C_1/C_2 transition are shown in violet and red, respectively. C_1 and C_2 are in fast exchange (violet arrow) while C_3 exchanges slowly with C_2 (red arrow). Therefore, only an average of C_1 and C_2 LF(C_1/C_2) is resolved. **(c)** Dynamic PDA (time-window, TW = 3 ms) characterizes the slow C_3 - C_1/C_2 exchange by the analysis of FRET-efficiency histograms (data, grey) by a kinetic two-state model (fit, black line) and recovers fluorescence-averaged distances of C_3 (33.6 Å) and C_1/C_2 (52.5 Å). Weighted residuals are shown to the right. **(d)** Sub-ensemble donor fluorescence decay analysis of the LF population resolves C_1 and C_2 as individual components. The donor fluorescence decay in the absence of FRET serves as reference. The instrument response function (IRF) is shown as black line. The magenta region shows the number of photons of the donor quenched due to FRET. On top of the fluorescence decay, the weighted residuals (w.res.) of a one-component (1 distance, red) and a two-component model (2 distances, violet) are shown. As visualized by the auto-correlation of the weighted residuals, the one-component model is significantly worse and is therefore discarded (p value > 0.99) **(e)** Filtered fluorescence correlation spectroscopy (FCS) computes the species-specific cross-correlation (SCC) between HF(C_1) and LF(C_1/C_2). The species auto correlation (SAC) of C_3 recovers the diffusion time t_{diff} . The presence of two anticorrelation terms in the SCC indicates three kinetic states. The analysis of the SCC recovers characteristic relaxation times t_{R1} and t_{R2} of the C_1 , C_2 , and C_3 kinetics.

factors. Finally, to generate brightness Q and fluorescence decays $f_{D(A)}(t)$ for each state in the simulated single-molecule experiment, we used the crystal structures as input. This involved two steps. First, DA distance distributions $\kappa(R_{DA})$ of the selected FRET pair were calculated using two inputs, the crystal structures of C_1 , C_2 , and C_3 , respectively, and the *Free diffusion+ContactIso model* defining the corresponding accessible contact volumes. Second, the obtained respective $\kappa(R_{DA})$ for C_1 , C_2 , and C_3 were used to calculate the FRET observables (fluorescence intensity decays $f(t)$ and transfer efficiencies F) according to Eqs. (1)–(3) (the assumed fluorescence parameters of the dyes are compiled in the caption of Figure 4). The transfer efficiencies E were used to calculate the brightness Q as described previously [60**]. All simulation parameters are compiled in the caption of Figure 4b.

The analysis of the fluorescence signals of the sm experiments simulated in step II of the workflow follows the procedures described in Box 1. First, the bright fluorescence bursts due to single-molecule events with durations of a few milliseconds were identified and selected in the time trace by separating them from the low background of 1–2 kHz. Second, for each single-molecule burst, all fluorescence parameters were determined as described in [60**]. To check for the presence of exchange dynamics, the FRET indicators, FRET efficiency E and the donor fluorescence lifetime $\tau_{D(A)}$ were plotted in two dimensional frequency MFD histograms (Figure 4b) to analyze the location of the bursts using the static (black) and dynamic (red) FRET lines [60**].

The number of conformers and their species fractions were identified as follows. A time window-based analysis by dynamic photon distribution analysis (dynPDA) [24*] resolved the FRET averaged distances $\langle R_{DA} \rangle_E$ of a small population (~25%) with high FRET efficiencies, referred to as HF population, which is in slow exchange with a second population with a lower FRET efficiency (~75%), referred to as LF population (Figure 4c). To test whether these populations are homogeneous, we performed a sub-ensemble TCSPC analysis of both populations. While the fluorescence decay of the HF population could be fitted with a single distance distribution (i.e. a single FRET species) (Eq. (2)), two distance distributions (i.e. two FRET species) were needed to describe the decay of the LF population (Figure 4d). The necessity of two FRET species in the LF population was judged by comparing the goodness of the fits for one and two distance distributions using weighted residuals (w.res.), the autocorrelation function (a.corr.) of the residuals, and χ^2 . Overall, sub-ensemble TCSPC and PDA analysis allowed us to resolve three FRET species. PDA identifies two dynamically exchanging populations. One of these populations is resolved into two distinct populations by sub-ensemble TCSPC (seTCSPC). We formally assign the HF population to the conformer (C_3) and the LF

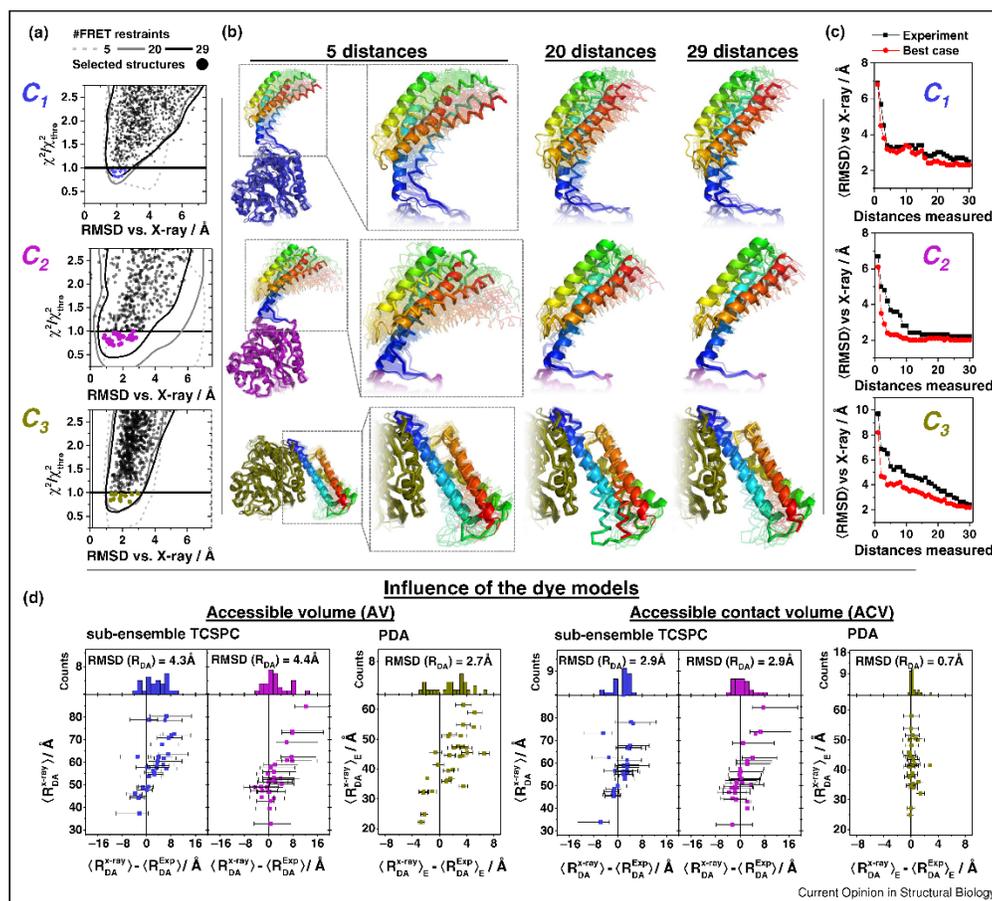
population to a dynamic mixture of C_1 and C_2 , which is separated from C_3 by a large kinetic barrier with a relaxation time of ~10 ms.

To resolve the heterogeneity of the averaged $\{C_1/C_2\}$ population, the fluorescence decays of the respective sub-ensembles for all 10 simulated single-pair FRET experiments were analyzed by a joint model, which assumed global species fractions of the $\{C_1/C_2\}$ population and Gaussian-distributed distances. Only this joint analysis is capable of recovering the experimental global species fractions $x_1 = 50\%$ and $x_2 = 25\%$, which agree well with input values used in the simulation (Figure 2). The presence of exchange kinetics is independently detected by filtered fluorescence correlation spectroscopy (fFCS) [23*,119] (Figure 4c) by computing the species cross-correlation function (SCC($\{C_1/C_2\} - C_3$)) that analyzes the exchange between the fast interchanging populations $\{C_1/C_2\}$ and C_3 . The SCC exhibits two relaxation times, which is additional evidence for the existence of three conformational states. The fast relaxation time $\tau_{R1} = 50 \mu\text{s}$ exactly corresponds to the one used in the simulations to describe the exchange between C_1 and C_2 . Even though the lifetimes of the conformers C_1 and C_2 are short (100 and 50 μs , respectively), MFD, seTCSPC and fFCS analysis unequivocally identified three FRET species together with their species fractions, respective distances and corresponding exchange kinetics. The errors of these distances $\Delta R_{DA}^2(E)$ were determined by propagating the experimental noise to the model parameters. As this analysis revealed three states, we accept the null hypothesis and reject the initial assumption that the crystal structure describes the experimental sample.

Therefore, in step III of the workflow, we explore the information contained in the experiments by screening the entire ensemble using FPS [42**]. This screen selects an ensemble of iteration 1 for the structural models of C_1 , C_2 and C_3 as *posterior* hypothesis.

FRET on rails: Step IV. To improve the precision of our *posterior* model, we performed the iteration step IV by repeating the DA pair selection procedure now using the reduced ensemble of iteration 1. Based on this analysis, we chose 19 additional informative DA pairs (Figure 4a). For these pairs, we again simulated smFRET experiments following step II as described above. After the analysis, our distance set for each conformer contained 29 distances. We then used these expanded distance sets of iterations 1 and 2 to model the respective structures applying FRET-guided NMSim simulations and subsequent screening by FPS (step III). We monitored the improvement of accuracy and precision for each conformer (Figure 5c). Since the RMSD value from the X-ray structure levels off at 25–29 distances for the best-case prediction and the recovered experimental accuracy for

Figure 5



Assessment of the hybrid-FRET structural models. Outcome of hybrid-FRET structural modeling of the three Atlasin-1 conformers. **(a)** The sum of squared deviations weighted by the estimated experimental error χ^2 relative to a threshold value $\chi^2_{\text{threshold}}$ estimated for the confidence level of 68% with respect to the best structure is plotted against C_{α} root-mean-square deviations (RMSD) from the corresponding crystal structure for 5 (dashed contour), 20 (grey contour), and 29 (black contour) measured distances. Black dots represent conformers corresponding to the set of 29 distances. Contours are drawn using a kernel density estimate. Colored dots represent selected conformations for the set of 29 distances. RMSD versus X-ray is calculated excluding flexible loops using the residues 35–99, 122–147, 157–189, 209–237, 257–277, 292–332, and 349–437. **(b)** Overlay of the crystal structures (cartoon representation), selected ensembles (transparent ribbon, 68% confidence), and best (χ^2_{min}) structures (solid ribbon) for the sets of 5, 20, and 29 distances. **(c)** Improvement of the corresponding accuracy with respect to the number of distances measured. Accuracy is calculated as χ^2 -weighted average C_{α} RMSD from the corresponding crystal structure. Black line represents the improvement using experimentally measured distances, the red line represents the best-case scenario where all the distances measured agree perfectly with the ones predicted for the crystal structure. **(d)** Comparison for the deviations between measured distances and distances predicted from the crystal structure by AV and ACV dye models.

all conformers, further measurements would not improve the accuracy of the obtained structures for Atlasin-1. However, in the case of remaining ambiguity for the solutions, additional iterations through steps IV and II

(determining DA pairs, simulating and analyzing smFRET experiments), and step III (structural modeling and subsequent screening) could improve the models even further.

FRET on rails: Step V (Figure 3). We screened all generated structures using FPS, which yielded FRET-selected ensembles of iteration 2 for each conformer C_1 , C_2 , and C_3 , using the confidence level threshold of 68%.

Accuracy and precision of FRET-derived structural models for the test case Atlastin-1. The accuracy and precision of the hybrid-FRET structural models obtained by the iterative workflow (Figure 3) is summarized in Figure 5a. We judge the accuracy of the method by plotting the conformer specific f -value ($\chi^2/\chi_{\text{threshold}}^2$ ratio) versus the C_α RMSD of the modeled structures from the corresponding target crystal structure. The $\chi_{\text{threshold}}^2$ value is estimated for the confidence level of 68%. For all states, the accuracy of the selected structures at a confidence level of 68% using 29 distances ranges between 1 and 3 Å. As we are interested in large-scale rearrangements of the super-tertiary structures, we exclude minor rearrangements of the sub-domains for the estimation of our accuracy (Figure 5a). To minimize the effect of non-uniform sampling, we use cluster representatives with a RMSD threshold of 1 Å (complete linkage) [125]. Cluster representatives below a confidence level of 68% for 29 distances are shown as colored dots. An overlay of the selected cluster representatives and the best model structure (highlighted by solid ribbons) visualizes the precision, which is given by the structural diversity within the selection (Figure 5b). To emphasize the differences, we aligned the representatives to the rigid G domain of the conformers. Within our precision (confidence level 68%) we can distinguish all three conformers, even C_1 and C_2 (i.e. the conformational ensembles do not overlap).

To capture the general dependence of the accuracy on the number of measured distances, we calculate an average RMSD that is weighted by the probability of agreement with the experimental distances as judged by χ^2 of the corresponding structural model. In Figure 5c, this average RMSD is used as a measure of accuracy, and shown as a function of the number of measured distances sorted by their information content. Interestingly, this RMSD decays fast with the number of measured distances and levels off in the studied system at an RMSD value of 2 Å. The precision and accuracy of the structural model is expected to depend on three major factors: (1) the noise (uncertainty) of the experimental data, (2) its sparseness and (3) outliers. Indeed, by increasing the number of distances, the effect of noise and data sparseness decrease for all conformers as highlighted in Figure 5c. The precision of the FRET-ensemble (4 to 5 Å) selected by 5 distances improves to 2 Å when 29 distances are measured. The influence of experimental noise and sparseness is seen by comparing the RMSD dependence on the number of measured distances in Figure 5c for the best case with accurate distances (black) and the real case with experimentally determined DA distances (red).

Notably, already small optimally informative set of DA pairs may provide a high accuracy (Figure 5c, red), because in the structural models obtained by NMSim (or any other computer simulation procedure) C_α atoms are interdependent. Therefore, only 29 FRET measurements with realistic errors (Figure 5c, black) localize the C_α positions to an RMSD of ~ 2 Å; the accuracy in our test does not improve significantly even if more than 15 measurements are performed. Notably, even for a large optimal experimental dataset, a residual RMSD error will be observed, as intrinsic degrees of backbone freedom within the structural model (relatively small fluctuations of C_α atoms) are not sensed by FRET. More FRET measurements would not alleviate this problem; that is, this test study reached the maximum possible precision. Finally, the third effect considering the possibility that the measurements might contain a low number of outliers, may be minimized by increasing the number of measurements, reducing the risk of systematic errors.

In crystallography, short range (up to 3 Å) stereo-chemical criteria, that is, deviations of atomic distances, angles, and dihedrals from their mean values, are used as quality criteria for obtained structures. Using such criteria (Mol-Probity [126]), the FRET-selected models generated by NMSim generally score better than corresponding X-ray structures. Therefore, we assess the overall quality of the structural model by experimental quality measures. As such, deviations between experimentally determined distances (R_{DA}^{Exp}) and corresponding model distances (R_{DA}^{Model}) are utilized as a quality indicator. Knowing the target structures in our test case, we used this indicator to assess the accuracy of different dye models (Figure 5d). We simulated our experimental information using the ACV dye model (Figure 3, step II) and, accordingly, the ACV model clearly outperformed the alternative AV model with an $\text{RMSD}(R_{DA})$ that is lower by ~ 1.5 – 2 Å when comparing R_{DA}^{Exp} with $R_{DA}^{\text{Model}} = R_{DA}^{\text{X-ray}}$ (Figure 5d). Thus, parameters of coarse grained dye models such as the preferential surface residence of the dye (ACV model) may be refined experimentally by calibration studies.

Outlook

Hybrid/integrative modeling: Combining FRET spectroscopy with other fluorescence and biophysical methods. To interrogate different molecular length scales, to address distinct sample properties (e.g. backbone, side chain, shape) and to cover different time scales, hybrid modeling can utilize the complementary information from a multitude of experimental techniques such as small-angle neutron scattering (SANS), small-angle X-ray scattering (SAXS), NMR-spectroscopy, EPR-spectroscopy and FRET-spectroscopy in one study [53*,109,127–129]. Integrative modeling is a thriving direction in methodological development directed towards

a multi-dimensional structural and dynamical description of complex biomacromolecular systems [38]. However, the error assessment of the data can often be treated only semi-quantitatively, because it is difficult, perhaps impossible, to determine specific errors of each method and their relative weights in a joint analysis.

In single photon counting it is straightforward to estimate error of FRET and other fluorescence measurements due to the Poisson statistics for photon noise. Besides FRET, other fluorescence methods are usually convenient and can also provide restraints for structural modeling as described in the following examples. Photo-induced electron-transfer (PET) probes the close proximity of the fluorophore to certain electron-rich quenching amino acid residues [130,131] and hence senses conformational dynamics [132]. Fluorescence polarization senses dye mobility that can be influenced by local and global rotation [133] and by complex formation where the label can be trapped [134]. The fluorescence spectra of polarity-sensitive fluorophores, such as intrinsic fluorescent amino acid tryptophan, sense their exposure to water [135]. FRET between the fluorophores and PET with electron rich amino acid residues are often competing processes in quenching the donor fluorescence, so that this scenario was either judged as valuable information [136] or as experimental nuisance [17]. Both methods can be combined to obtain distance (PET: short range, FCS: long range) [137] as well as kinetic information [138]. Information similar to that from PET can be obtained from protein-induced fluorescence enhancement (PIFE) experiments with cyanine dyes [98,99]. Finally, SAXS and polarization resolved FCS measurements [139,140] contain information on the anisotropic rotational diffusion of macromolecules, which reports on the size and the global shape of the macromolecules.

Maximizing data utilization by structure-based forward modeling. In forward modeling [111**], the chosen model parameters are verified by a direct comparison to the experimental data. In fluorescence experiments with single photon counting detection, forward modeling of the registered fluorescence decays and FRET efficiency histograms has a long tradition (e.g. convolution of instrument response function instead of deconvolution), because in this way the Poisson statistics of the counted photons is preserved and the experimental noise (shot-noise) is exactly known. However, since structural modeling is typically performed using distance restraints (in backward modeling; i.e. the fit is decoupled from structural model), experimental data is preprocessed by applying fluorescence models (for instance Eqs. (1)–(3)). This can lead to overestimation of the errors, if the model parameters (distances) are correlated. However, to obtain independent errors for the distance restraints, the multi-dimensional uncertainties of the distances are projected

to one dimension (i.e. marginalized). In this process, the valuable information on correlations is lost. Similar to peak assignment problems in NMR spectroscopy [111**], correlations can be preserved by forward modeling which utilizes a structural model as a global model for the experimental data, maximizing the recovered information and minimizing ambiguities. Dye models must be applied to directly link the structural model to the fluorescence decay. For this purpose AV, AGV or more complex spatial dye distributions can be used. In this forward fitting approach the structural model is varied for optimal agreement with experiment.

Hybrid-FRET studies in live cells. The FRET technique offers unique advantages for characterizing large biomolecules with high specificity and sensitivity in living cells or *in vitro*. Time-resolved FRET studies can deal with heterogeneous samples and make use of the single-molecule advantage so that a quantitative FRET analysis can be combined with normal [141,142] and super-resolution [143] imaging to gain insight into the biological function of biomolecules in their cellular context (e.g. localizing complexes inside of cells [141], quantifying binding to interaction partners [104*], and validating structural models [103,144,145]). For example, we could show by detailed hybrid-FRET studies of murine immune defense Guanylate binding proteins with GFP and mCherry fusions that they undergo reversible structural transitions between monomeric, dimeric and oligomeric states as revealed by a quantitative analysis of homo- and hetero-FRET [104*]. Notably, the degree of oligomerization is specific to the localization in live cells. Detailed hybrid-FRET studies allowed us to characterize the structures of the dimer and oligomer.

Comparative studies of cell-like environment and dilute solutions were used to extrapolate excluded volume effects from *in vitro* experiments to live-cells [146*], and to identify biomolecular stabilization mechanism caused by macromolecular crowding [147]. Microinjection allows one to control the concentration of biomolecules labeled by photostable fluorophores and therefore enables the study of crowding effects in live cells [28]. In this way, single-molecule conditions can be achieved by tracking microinjected *in vitro* purified FRET-labeled proteins to study complex formation and conformational changes of individual proteins [148]. This approach, combined with fast confocal detection, was used to probe protein dynamics from millisecond down to the nanosecond regime [149].

In conclusion, the presented hybrid-FRET methods actually allow realizing an integrated molecular fluorescence microscope combining optical and computational microscopy [58] at a huge spatial and temporal range to display suitably labeled biomolecular systems at unprecedented resolution by atomistic structural models.

Further information

For the presented test case study on Atlastin-1, all structural models presented in Figure 5, additional detailed information on the applied procedures and the simulated sm FRET data are deposited in the Model Archive with the DOI: 10.5452/ma-a2hbq.

Conflict of interest

Nothing declared.

Acknowledgments

We thank Stanislav Kalinin, Jens Michaelis and Oleg Opanasyuk for fruitful discussions on integrative FRET-modeling. We are grateful to Suren Felekyan for supporting the fluorescence simulations and Katherina Hemmen for her support in preparing figures. This work was supported by the European Research Council through the Advanced Grant 2014 hybridFRET (671208). The authors acknowledge the stimulating networking support by the COST Action CM1306 – Understanding Movement and Mechanism in Molecular Machines.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Förster T: **Zwischenmolekulare Energiewanderung und Fluoreszenz.** *Ann Phys* 1948, **437**:55-75.
In this seminal paper Förster derived the correct distance and orientation dependence of the FRET-rate constant for the first time. Additionally, energy migration within regular crystallographic lattices (the second part of this paper) is studied, which is still of interest for crystal like structures for instance, formed by membrane associated proteins such as GPCRs.
 2. Förster T: **Experimentelle und theoretische Untersuchung des zwischenmolekularen Übergangs von Elektronenanregungsenergie.** *Z Naturforsch A* 1949, **4**:321-327.
In this second seminal paper Förster derived experimental observables starting from simple schemes of rate constants. Equations for fluorescence decays of donors randomly surrounded by acceptor fluorophores are presented. In this sense this fundamental paper is the basis for understanding most FRET-experiments with flexibly coupled dyes.
 3. van der Meer BW, Cooker G, Chen SY: **Resonance Energy Transfer: Theory and Data.** VCH Publishers; 1994.
 4. Weiss S: **Fluorescence spectroscopy of single biomolecules.** *Science* 1999, **283**:1676-1683.
 5. Stryer L: **Fluorescence energy transfer as a spectroscopic ruler.** *Annu Rev Biochem* 1978, **47**:819-846.
 6. Ha T, Enderle 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, **93**:6264-6268.
 7. Pirchi M, Ziv G, Riven I, Cohen SS, Zohar N, Barak Y, Haran G: **Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein.** *Nat Commun* 2011, **2**:493-499.
 8. Chung HS, Eaton WA: **Single-molecule fluorescence probes dynamics of barrier crossing.** *Nature* 2013, **502**:685-688.
 9. Zhuang XW, Kim H, Pereira MJB, Babcock HP, Walter NG, Chu S: **Correlating structural dynamics and function in single ribozyme molecules.** *Science* 2002, **296**:1473-1476.
 10. Kim HD, Nienhaus GU, Ha T, Orr JW, Williamson JR, Chu S: **Mg²⁺-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules.** *Proc Natl Acad Sci U S A* 2002, **99**:4284-4289.
 11. Al-Hashimi HM, Walter NG: **RNA dynamics: it is about time.** *Curr Opin Struct Biol* 2008, **18**:321-329.
 12. Roy R, Kozlov AG, Lohman TM, Ha T: **SSB protein diffusion on single-stranded DNA stimulates RecA filament formation.** *Nature* 2009, **461**:1092-1097.
 13. Olofsson L, Felekyan S, Doumazane E, Scholler P, Fabre L, Zwier JM, Rondard P, Seidel CAM, Pin JP, Margeat E: **Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy.** *Nat Commun* 2014, **5**:e5206.
 14. Robb NC, Te Velthuis AJ, Wieneke R, Tampe R, Cordes T, Fodor E, Kapanidis AN: **Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA.** *Nucleic Acids Res* 2016, **44**.
 15. Diez M, Zimmermann B, Börsch M, König M, Schweinberger E, Steigmiller S, Reuter R, Felekyan S, Kudryavtsev V, Seidel CAM, Gräber P: **Proton-powered subunit rotation in single membrane-bound F₁F₀-ATP synthase.** *Nat Struct Mol Biol* 2004, **11**:135-141.
 16. Lu HP: **Sizing up single-molecule enzymatic conformational dynamics.** *Chem Soc Rev* 2014, **43**:1118-1143.
 17. Winkler JR: **FRETting over the spectroscopic ruler.** *Science* 2013, **339**:1530-1531.
 18. Nettels D, Gopich IV, Hoffmann A, Schuler B: **Ultrafast dynamics of protein collapse from single-molecule photon statistics.** *Proc Natl Acad Sci U S A* 2007, **104**:2655-2660.
 19. Soranno A, Buchli B, Nettels D, Cheng RR, Müller-Spahn S, Pfeil SH, Hoffmann A, Lipman EA, Makarov DE, Schuler B: **Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy.** *Proc Natl Acad Sci U S A* 2012, **109**:17800-17806.
 20. Gopich I, Szabo A: **Theory of photon statistics in single-molecule Förster resonance energy transfer.** *J Chem Phys* 2005, **122**:014707.
Gopich and Szabo describe FRET-efficiency histograms (FEHs) of single-molecule FRET experiments by the photon statistics. As they consider the effect of conformational dynamics on FEH, they set the basis for quantitative time-window analysis of FEHs of dynamically exchanging species.
 21. Chung HS, Gopich IV: **Fast single-molecule FRET spectroscopy: theory and experiment.** *Phys Chem Chem Phys* 2014, **16**:18644-18657.
In this perspective, the authors demonstrate the improvement of the dynamic range and accuracy of sm fluorescence spectroscopy at a given photon count rate by considering each and every photon and introducing a maximum likelihood (ML) method. By analyzing photon trajectories with recorded photon colors and inter-photon times of smFRET studies (protein folding), their ML method recovers parameters of a kinetic exchange model.
 22. Chung HS, McHale K, Louis JM, Eaton WA: **Single-molecule fluorescence experiments determine protein folding transition path times.** *Science* 2012, **335**:981-984.
 23. Felekyan S, Sanabria H, Kalinin S, Kühnemuth R, Seidel CAM: **Analyzing Förster Resonance Energy Transfer with fluctuation algorithms.** *Methods Enzymol* 2013, **519**:39-85.
This review describes the theory and application of fluorescence correlation spectroscopy (FCS) with different analysis procedures (classical FRET-FCS and FRET-specific species correlation by filtered FCS) to analyze FRET fluctuations in order to distinguish between different species. Filtered FCS is especially useful for systems with complex dynamics.
 24. Kalinin S, Valeri A, Antonik M, Felekyan S, Seidel CAM: **Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states.** *J Phys Chem B* 2010, **114**:7983-7995.
The authors present a method to analyze FRET efficiency histograms, taking into account shot noise and background contributions. The method can be applied to resolve the dynamically exchanging states and their kinetics. Furthermore, the authors present analysis tools to identify conformational dynamics by comparing the intensity-based and time-resolved FRET indicators F_D , F_A and τ_D (Eq. (1)) by calculating FRET-lines for the static and dynamic case, respectively (see Box 1). The provided theory is valuable to investigate dynamic systems.

25. McKinney SA, Joo C, Ha T: **Analysis of single-molecule FRET trajectories using hidden Markov modeling.** *Biophys J* 2006, **91**:1941-1951.
- The authors present a robust and unbiased algorithm to recover kinetic schemes from time-binned sm FRET trajectories. A number of different states and state-to-state transition probabilities can be recovered, together with the corresponding FRET efficiencies for transitions slower than the bin width. This is an effective analysis method for widely applied TIRF measurements.
26. König SLB, Hadzic M, Fiorini E, Börner R, Kowanko D, Blanckenhorn WU, Sigel RKO: **BOBA FRET: Bootstrap-based analysis of single-molecule FRET data.** *PLoS ONE* 2013, **8**:e84157.
27. Keller BG, Kobitski A, Jaschke A, Nienhaus GU, Noe F: **Complex RNA folding kinetics revealed by single-molecule FRET and Hidden Markov Models.** *J Am Chem Soc* 2014, **136**:4534-4543.
- A rigorous hidden Markov Model (HMM) analysis procedure is shown to recover states based on kinetics rather than just FRET efficiencies. For Diels-Alderase ribozyme, a seven state model with fastest interconversion timescale of 3 ms is recovered. The approach enables reconstruction of complex kinetic schemes at high temporal resolution.
28. Gao M, Gnutz D, Orban A, Appel B, Righetti F, Winter R, Narberhaus F, Muller S, Ebbinghaus S: **RNA hairpin folding in the crowded cell.** *Angew Chem Int Ed* 2016, **55**:3224-3228.
29. Schuler B, Hofmann H: **Single-molecule spectroscopy of protein folding dynamics-expanding scope and timescales.** *Curr Opin Struct Biol* 2013, **23**:36-47.
- The authors review a broad spectrum of sm fluorescence methods based on sm FRET and photo-induced electron transfer (PET) that have been developed for probing protein structure and dynamics, especially in structurally heterogeneous systems. They give an excellent overview on the applications of the kinetic experiments by discussing many processes coupled with protein folding such as misfolding, binding, transition path times and interactions with chaperones.
30. Schneidman-Duhovny D, Pellarin R, Sali A: **Uncertainty in integrative structural modeling.** *Curr Opin Struct Biol* 2014, **28**:96-104.
- This review defines the need for standards and tools to assess the input data and resulting models in integrative structural modeling. To tackle problems in integrative structural modeling they define standards for a common terminology. Furthermore, the authors point out possible sources of model uncertainty.
31. Stryer L, Haugland RP: **Energy transfer: a spectroscopic ruler.** *Proc Natl Acad Sci U S A* 1967, **58**:719-726.
32. Murchie AI, Clegg RM, von Kitzing E, Duckett DR, Diekmann S, Lilley DMJ: **Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules.** *Nature* 1989, **341**:763-766.
33. Wozniak AK, Schröder GF, Grubmüller H, Seidel CAM, Oesterhelt F: **Single-molecule FRET measures bends and kinks in DNA.** *Proc Natl Acad Sci U S A* 2008, **105**:18337-18342.
34. Preus S, Wilhelmsson LM: **Advances in quantitative FRET-based methods for studying nucleic acids.** *ChemBioChem* 2012, **13**:1990-2001.
35. Preus S, Kilsa K, Miannay FA, Albinsson B, Wilhelmsson LM: **FRETmatrix: a general methodology for the simulation and analysis of FRET in nucleic acids.** *Nucleic Acids Res* 2013, **41**:e18.
- The presented methodology addresses the problem of deriving structures using fluorophores with fixed positions and fixed dipoles using information from fluorescence decays combined with fits directed by the structural model. The resulting toolkit FRETmatrix (<http://www.fluortools.com/software/fretmatrix>) is applied to study nucleic acid structures in 3D.
36. Mekler V, Kortkhonija E, Mukhopadhyay J, Knight J, Revyakin A, Kapanidis AN, Niu W, Ebricht YW, Levy R, Ebricht RH: **Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex.** *Cell* 2002, **108**:599-614.
37. Nagy J, Gröhmman D, Cheung ACM, Schulz S, Smollett K, Werner F, Michaelis J: **Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS.** *Nat Commun* 2015, **6**:e6161.
38. Sali A, Berman HM, Schwede T, Trehella J, Kleywegt G, Burley SK, Markley J, Nakamura H, Adams P, Bonvin A, Chiu W et al.: **Outcome of the first wwPDB Hybrid/Integrative methods task force workshop.** *Structure* 2015, **23**:1156-1167.
39. Schneider M, Belson A, Rappsilber J, Brock O: **Blind testing of cross-linking/mass spectrometry hybrid methods in CASP11.** *Proteins* 2016, **84**:152-163.
40. Tompa P: **On the supertertiary structure of proteins.** *Nat Chem Biol* 2012, **8**:597-600.
41. Knight JL, Mekler V, Mukhopadhyay J, Ebricht RH, Levy RM: **Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability.** *Biophys J* 2005, **88**:925-938.
- This benchmark study demonstrates positioning accuracies between 7 and 10 Å for small molecules relative to the crystallographically defined binding sites without explicitly considering dye distributions.
42. Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, Goody RS, Gohlke H, Seidel CAM: **A toolkit and benchmark study for FRET-restrained high-precision structural modeling.** *Nat Meth* 2012, **9**:1218-1227.
- The authors present a comprehensive toolkit (FRET Positioning and Screening, FPS, <http://www.mpc.hhu.de/software/fps.html>) for quantitative FRET-restrained modeling for applications in structural biology, which takes the spatial distributions of flexible linked dyes into account. In a benchmark study (docking a DNA primer-template to HIV-1 reverse transcriptase) the RMSD of the structural model to the known X-ray structure was only 0.5 Å. The FPS toolkit has two main functionalities: (1) Planning FRET experiments by calculating fluorescence observables for the given structural models; (2) FRET-based screening and multibody docking determine quantitative structural models.
43. Smock RG, Gierasch LM: **Sending signals dynamically.** *Science* 2009, **324**:198-203.
44. Henzler-Wildman K, Kern D: **Dynamic personalities of proteins.** *Nature* 2007, **450**:964-972.
45. Neudecker P, Robustelli P, Cavalli A, Walsh P, Lundstrom P, Zarrine-Afsar A, Sharpe S, Vendruscolo M, Kay LE: **Structure of an intermediate state in protein folding and aggregation.** *Science* 2012, **336**:362-366.
46. Banerjee PR, Deniz AA: **Shedding light on protein folding landscapes by single-molecule fluorescence.** *Chem Soc Rev* 2014, **43**:1172-1188.
47. Ferreon ACM, Ferreon JC, Wright PE, Deniz AA: **Modulation of allostery by protein intrinsic disorder.** *Nature* 2013, **498**:390-394.
48. Margittai M, Widengren J, Schweinberger E, Schröder GF, Felekyan S, Hausteiner E, König M, Fasshauer D, Grubmüller H, Jahn R, Seidel CAM: **Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1.** *Proc Natl Acad Sci U S A* 2003, **100**:15516-15521.
49. McCann JJ, Zheng LQ, Chiantia S, Bowen ME: **Domain orientation in the N-terminal PDZ tandem from PSD-95 is maintained in the full-length protein.** *Structure* 2011, **19**:810-820.
50. McCann JJ, Zheng L, Rohrbeck D, Felekyan S, Kühnemuth R, Sutton RB, Seidel CAM, Bowen ME: **Supertertiary structure of the synaptic MAGuK scaffold proteins is conserved.** *Proc Natl Acad Sci U S A* 2012, **109**:15775-15780.
51. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA: **Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations.** *Proc Natl Acad Sci U S A* 2007, **104**:1528-1533.
52. Kellner R, Hofmann H, Barducci A, Wunderlich B, Nettels D, Schuler B: **Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein.** *Proc Natl Acad Sci U S A* 2014, **111**:13355-13360.
53. Aznauryan M, Delgado L, Soranno A, Nettels D, Huang JR, Labhardt AM, Grzesiek S, Schuler B: **Comprehensive structural and dynamical view of an unfolded protein from the**

182 Biophysical and molecular biological methods

- combination of single-molecule FRET, NMR, and SAXS.** *Proc Natl Acad Sci U S A* 2016, **113**:E5389-E5398.
- In this integrative structural study, the authors combine three of the most powerful biophysical methods (NMR spectroscopy, small angle X-ray scattering and smFRET measurements) available to obtain a comprehensive view on the structure and dynamics of the unfolded protein, ubiquitin, that would not be available from any of the individual methods.
54. Choi UB, Zhao ML, Zhang YX, Lai Y, Brunger AT: **Complexin induces a conformational change at the membrane-proximal C-terminal end of the SNARE complex.** *Life* 2016, **5**:e16886.
 55. Munro JB, Gorman J, Ma XC, Zhou Z, Arthos J, Burton DR, Koff WC, Courter JR, Smith AB, Kwong PD, Blanchard SC *et al.*: **Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions.** *Science* 2014, **346**:759-763.
 56. Kay LE: **New views of functionally dynamic proteins by solution NMR spectroscopy.** *J Mol Biol* 2016, **428**:323-331.
 57. Schwede T, Sali A, Honig B, Levitt M, Berman HM, Jones D, Brenner SE, Burley SK, Das R, Dokholyan NV, Dunbrack RL Jr *et al.*: **Outcome of a workshop on applications of protein models in biomedical research.** *Structure* 2009, **17**:151-159.
 58. Lee EH, Hsin J, Sotomayor M, Comellas G, Schulten K: **Discovery through the computational microscope.** *Structure* 2009, **17**:1295-1306.
 59. Clegg RM: **Fluorescence resonance energy transfer and nucleic acids.** *Methods Enzymol* 1992, **211**:353-388.
 This review provides an introduction to FRET and presents elegant and simple methods to determine absolute FRET efficiencies by steady state fluorometers. This allows one to obtain structural insights on the ensemble level using standard instrumentation.
 60. Sisamakris E, Valeri A, Kalinin S, Rothwell PJ, Seidel CAM: **Accurate single-molecule FRET studies using multiparameter fluorescence detection.** *Methods Enzymol* 2010, **475**:455-514.
 This review gives an introduction to single-molecule FRET experiments and analyses. It focuses on confocal detection of freely diffusing molecules by multi-parameter fluorescence detection. The authors describe qualitative and quantitative analysis methods for single-molecule FRET measurements and their prerequisites. Quantitative analysis of FRET measurements is the basis for FRET-based structural modeling.
 61. Sindbert S, Kalinin S, Hien N, Kienzler A, Clima L, Bannwarth W, Appel B, Müller S, Seidel CAM: **Accurate distance determination of nucleic acids via Förster Resonance Energy Transfer: implications of dye linker length and rigidity.** *J Am Chem Soc* 2011, **133**:2463-2480.
 This study relates dye linker length flexibilities to experimental FRET observables by measuring the fluorescence lifetimes and anisotropies of donor (Alexa 488) and acceptor (Cy5) dyes, respectively, tethered to dsRNA and dsDNA. For dyes with short linkers and unknown local environments, positional-related and orientational-related uncertainties outweigh better defined dye positions. For FRET-based structure determination of very high accuracy using dyes with long linkers, a procedure (accessible volume simulations) to model the spatial dye distributions is introduced.
 62. Widengren J, Schweinberger E, Berger S, Seidel CAM: **Two new concepts to measure fluorescence resonance energy transfer via fluorescence correlation spectroscopy: theory and experimental realizations.** *J Phys Chem A* 2001, **105**:6851-6866.
 63. Rothwell PJ, Berger S, Kensch O, Felekyan S, Antonik M, Wöhrl BM, Restle T, Goody RS, Seidel CAM: **Multi-parameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 Reverse Transcriptase:primer/template complexes.** *Proc Natl Acad Sci U S A* 2003, **100**:1655-1660.
 64. Kudryavtsev V, Sikor M, Kalinin S, Mokranjac D, Seidel CAM, Lamb DC: **Combining MFD and PIE for accurate single-pair Förster Resonance Energy Transfer measurements.** *ChemPhysChem* 2012, **13**:1060-1078.
 A combination of multiparameter fluorescence detection (MFD), pulsed interleaved excitation (PIE) and smFRET experiments on freely diffusing molecules is presented. MFD monitors a variety of fluorescence parameters simultaneously while PIE directly excites the acceptor to probe its presence and photoactivity. A single MFD-PIE measurement provides all necessary calibration factors for accurate analysis of smFRET histograms and excludes artifacts due to acceptor blinking or photobleaching. The performance of MFD-PIE measurements is demonstrated studying the interdomain separation in DnaK, the major bacterial heat shock protein 70 (Hsp70).
 65. Vöpel T, Hengstenberg CS, Peulen TO, Ajaj Y, Seidel CAM, Herrmann C, Klare JP: **Triphosphate induced dimerization of human guanylate binding protein 1 involves association of the C-terminal helices: a joint double electron-electron resonance and FRET study.** *Biochemistry* 2014, **53**:4590-4600.
 66. Roy R, Hohng S, Ha T: **A practical guide to single-molecule FRET.** *Nat Meth* 2008, **5**:507-516.
 The majority of the single-molecule FRET community performs total internal reflection (TIRF) experiments. This guide gives an introduction to FRET, TIRF, and discusses potential artifacts relevant for quantitative FRET-based structural modeling.
 67. Rigler R, Mets Ü, Widengren J, Kask P: **Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion.** *Eur Biophys J* 1993, **22**:169-175.
 68. Farooq S, Hohlbein J: **Camera-based single-molecule FRET detection with improved time resolution.** *Phys Chem Chem Phys* 2015, **17**:27862-27872.
 69. Lee NK, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebright RH, Weiss S: **Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation.** *Biophys J* 2005, **88**:2939-2953.
 70. Hohlbein J, Craggs TD, Cordes T: **Alternating-laser excitation: single-molecule FRET and beyond.** *Chem Soc Rev* 2014, **43**:1156-1171.
 71. Elf J, Li GW, Xie XS: **Probing transcription factor dynamics at the single-molecule level in a living cell.** *Science* 2007, **316**:1191-1194.
 72. Flors C, Hotta J, Uji IH, Dedecker P, Ando R, Mizuno H, Miyawaki A, Hofkens J: **A stroboscopic approach for fast photoactivation-localization microscopy with Dronpa mutants.** *J Am Chem Soc* 2007, **129**:13970-13977.
 73. Felekyan S, Kühnemuth R, Kudryavtsev V, Sandhagen C, Becker W, Seidel CAM: **Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection.** *Rev Sci Instrum* 2005, **76**:083104.
 74. Boukobza E, Sonnenfeld A, Haran G: **Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy.** *J Phys Chem B* 2001, **105**:12165-12170.
 75. Müller BK, Zaychikov E, Bräuchle C, Lamb DC: **Pulsed interleaved excitation.** *Biophys J* 2005, **89**:3508-3522.
 The authors present the method and analysis procedures of pulsed interleaved excitation (PIE) for the first time. PIE measurements significantly extend the capabilities of multiple-color fluorescence imaging, fluorescence cross-correlation spectroscopy (FCCS), and smFRET measurements.
 76. Kühnemuth R, Seidel CAM: **Principles of single molecule multiparameter fluorescence spectroscopy.** *Single Mol* 2001, **2**:251-254.
 77. Widengren J, Kudryavtsev V, Antonik M, Berger S, Gerken M, Seidel CAM: **Single-molecule detection and identification of multiple species by multiparameter fluorescence detection.** *Anal Chem* 2006, **78**:2039-2050.
 78. McCann JJ, Choi UB, Zheng L, Weninger K, Bowen ME: **Optimizing methods to recover absolute FRET efficiency from immobilized single molecules.** *Biophys J* 2010, **99**:961-970.
 79. Hildebrandt LL, Preus S, Birkedal V: **Quantitative single molecule FRET efficiencies using TIRF microscopy.** *Faraday Discussions* 2015, **184**:131-142.
 The authors introduced new software (<http://www.isms.au.dk>) to determine quantitative FRET efficiencies directly from individual fluorescence time traces of surface immobilized DNA molecules without the need for external calibrants.
 80. Preus S, Hildebrandt LL, Birkedal V: **Optimal background estimators in single-molecule FRET microscopy.** *Biophys J* 2016, **111**:1278-1286.

81. Jeschke G: **DEER distance measurements on proteins**. In *Annual Review of Physical Chemistry*, Vol. 63. Edited by Johnson MA, Martinez TJ. 2012:419-446.
82. Columbus L, Hubbell WL: **A new spin on protein dynamics**. *Trends Biochem Sci* 2002, **27**:288-295.
83. Cai Q, Kusnetzow AK, Hideg K, Price EA, Haworth IS, Qin PZ: **Nanometer distance measurements in RNA using site-directed spin labeling**. *Biophys J* 2007, **93**:2110-2117.
84. Hagelueken G, Abdullin D, Schiemann O: **mtssiSuite: Probing biomolecular conformation by spin-labeling studies**. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions*, Pt A. 563. Edited by Qin PZ, Warncke K. 2015:595-622.
85. Muschielok A, Andrecka J, Jawhari A, Bruckner F, Cramer P, Michaelis J: **A nano-positioning system for macromolecular structural analysis**. *Nat Meth* 2008, **5**:965-971.
- The nano positioning system (NPS) methodology (<https://www.uni-ulm.de/nawi/nawi-biophys/forschung/forschung-michaelis/method-development/software/nps/>) determines the spatially distributed uncertainties from a set of FRET experiments, assuming the dyes adopt fixed positions in space. A Bayesian approach to address such a problem is presented which localizes dyes with fixed locations.
86. Brunger AT, Strop P, Vrljic M, Chu S, Weninger KR: **Three-dimensional molecular modeling with single molecule FRET**. *J Struct Biol* 2011, **173**:497-505.
- Technical details on implementing FRET-restraints into modelling techniques originally developed for NMR to FRET are given. The presented implementation is publicly available in the established software CNS (Crystallography and NMR system; <http://cns-online.org/v1.3/>).
87. Klose D, Klare JP, Grohmann D, Kay CWM, Werner F, Steinhoff HJ: **Simulation vs. reality: a comparison of in silico distance predictions with DEER and FRET measurements**. *PLoS ONE* 2012, **7**:e39492.
88. Choi UB, Strop P, Vrljic M, Chu S, Brunger AT, Weninger KR: **Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex**. *Nat Struct Mol Biol* 2010, **17**:318-324.
89. Hoeffling M, Lima N, Haenni D, Seidel CAM, Schuler B, Grubmüller H: **Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and Monte Carlo approach**. *PLoS ONE* 2011, **6**:e19791.
90. Graen T, Hoeffling M, Grubmüller H: **AMBER-DYES: characterization of charge fluctuations and force field parameterization of fluorescent dyes for molecular dynamics simulations**. *J Chem Theory Comput* 2014, **10**:5505-5512.
91. Best RB, Hofmann H, Nettels D, Schuler B: **Quantitative interpretation of FRET experiments via molecular simulation: force field and validation**. *Biophys J* 2015, **108**:2721-2731.
92. Milas P, Gamari BD, Parrot L, Krueger BP, Rahmanseresht S, Moore J, Goldner LS: **Indocyanine dyes approach free rotation at the 3' terminus of A-RNA: a comparison with the 5' terminus and consequences for Fluorescence Resonance Energy Transfer**. *J Phys Chem B* 2013, **117**:8649-8658.
93. Liao J-M, Wang Y-T, Chen C-L: **Computer simulation to investigate the FRET application in DNA hybridization systems**. *Phys Chem Chem Phys* 2011, **13**:10364-10371.
94. Walczewska-Szewc K, Corry B: **Accounting for dye diffusion and orientation when relating FRET measurements to distances: three simple computational methods**. *Phys Chem Chem Phys* 2014, **16**:12317-12326.
95. Kalinin S, Sisamakris E, Magennis SW, Felekyan S, Seidel CAM: **On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits**. *J Phys Chem B* 2010, **114**:6197-6206.
96. Levitus M, Ranjit S: **Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments**. *Q Rev Biophys* 2010:1-29.
- This review focuses on the photophysical and photochemical properties of the most frequently used cyanine dyes in single-molecule FRET studies. Non-fluorescent dark states and methods to enhance dye fluorescence are discussed. Effects of dark states in the context of single-molecule FRET and localization studies are described.
97. Abdullin D, Hagelueken G, Schiemann O: **Determination of nitroxide spin label conformations via dPELDOR and X-ray crystallography**. *Phys Chem Chem Phys* 2016, **18**:10428-10437.
98. Hwang H, Myong S: **Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions**. *Chem Soc Rev* 2014, **43**:1221-1229.
99. 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* 2016, **120**:6401-6410.
- Protein-induced fluorescence enhancement (PIFE) serves as a complementary molecular ruler at molecular distances inaccessible to other spectroscopic rulers such as FRET or photoinduced electron transfer (PET). The authors provide experimental evidence that PIFE is proportional to the contact volume of the dye's accessible volume with protein interface.
100. Beckers M, Drechsler F, Eilert T, Nagy J, Michaelis J: **Quantitative structural information from single-molecule FRET**. *Faraday Discuss* 2015, **184**:117-129.
- The classic NPS methodology is extended by assumptions (e.g. dynamic dipole averaging), which drastically increases the precision in locating the position of the dyes by multiple FRET measurements. This study demonstrates that the dye description is essential to achieve a high precision to localize dyes.
101. Höfig H, Gabba M, Poblete S, Kempe D, Fitter J: **Inter-dye distance distributions studied by a combination of single-molecule FRET-filtered lifetime measurements and a weighted accessible volume (wAV) algorithm**. *Molecules* 2014, **19**:19269-19291.
102. Steffen FD, Sigel RKO, Börner R: **An atomistic view on carbocyanine photophysics in the realm of RNA**. *PCCP* 2016, **18** Advance Article.
103. Greife A, Felekyan S, Ma QJ, Gertzen CGW, Spomer L, Dimura M, Peulen TO, Wöhler C, Häussinger D, Gohlke H, Keitel V et al.: **Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study**. *Sci Rep* 2016. Advance Article.
104. Kravets E, Degrandi D, Ma QJ, Peulen TO, Klumpers V, Felekyan S, Kühnemuth R, Weidtkamp-Peters S, Seidel CAM, Pfeffer K: **Guanylate binding proteins directly attack *Toxoplasma gondii* via supramolecular complexes**. *eLife* 2016, **5**:e11479.
- The authors use multi-parameter fluorescence image spectroscopy in live cells to validate structural models and determine equilibrium constants to study the role of immuno-proteins in intracellular pathogen defense. This study shows the application of forward modeling using live-cell data.
105. Iqbal A, Arslan S, Okumus B, Wilson TJ, Giraud G, Norman DG, Ha T, Lilley DMJ: **Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids**. *Proc Natl Acad Sci U S A* 2008, **105**:11176-11181.
106. Antonik M, Felekyan S, Gaiduk A, Seidel CAM: **Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis**. *J Phys Chem B* 2006, **110**:6970-6978.
- The authors established the probability distribution analysis (PDA) method for the quantitative analysis of FRET efficiency (E) histograms to determine with high precision (<1% of the Förster radius) the originating E value of a shot-noise-limited distribution. Any deviations can be assigned to static and/or dynamic heterogeneities of the molecular system.
107. Kalinin S, Felekyan S, Antonik M, Seidel CAM: **Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer**. *J Phys Chem B* 2007, **111**:10253-10262.
108. Raval A, Piana S, Eastwood MP, Shaw DE: **Assessment of the utility of contact-based restraints in accelerating the prediction of protein structure using molecular dynamics simulations**. *Protein Sci* 2016, **25**:19-29.
- An excellent demonstration of how sparse restraints similar to those provided by FRET, enable MD-based protein structure determination.

184 Biophysical and molecular biological methods

109. Boura E, Rozycki B, Chung HS, Herrick DZ, Canagarajah B, Cafiso DS, Eaton WA, Hummer G, Hurley JH: **Solution structure of the ESCRT-I and -II supercomplex: Implications for membrane budding and scission.** *Structure* 2012, **20**:874-886.
110. Brunger AT: **Version 1.2 of the crystallography and NMR system.** *Nat Protocols* 2007, **2**:2728-2733.
111. Rieping W, Habeck M, Nilges M: **Inferential structure determination.** *Science* 2005, **309**:303-306.
The authors describe a forward modeling Bayesian approach to improve the structural quality of a model by taking into account indirect structural data, such as NMR data. Similar to NMR, ambiguous experimental data poses a challenge to structural modeling by FRET
112. Muschielok A, Michaelis J: **Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks.** *J Phys Chem B* 2011, **115**:11927-11937.
113. Dumat B, Larsen AF, Wilhelmsson LM: **Studying Z-DNA and B- to Z-DNA transitions using a cytosine analogue FRET-pair.** *Nucleic Acids Res* 2016, **44**:e101.
114. Berman H, Henrick K, Nakamura H: **Announcing the worldwide Protein Data Bank.** *Nat Struct Biol* 2003, **10**:980-980.
115. Berman HM, Burley SK, Kleywegt GJ, Markley JL, Nakamura H, Velankar S: **The archiving and dissemination of biological structure data.** *Curr Opin Struct Biol* 2016, **40**:17-22.
116. Haas J, Roth S, Arnold K, Kiefer F, Schmidt T, Bordoli L, Schwede T: **The Protein Model Portal — a comprehensive resource for protein structure and model information.** *Database* 2013:bat031.
117. NIST/SEMATECH: **e-Handbook of Statistical Methods.** 2016: <http://www.itl.nist.gov/div898/handbook/>.
118. Guyon I, Elisseeff A: **An introduction to variable and feature selection.** *J Mach Learn Res* 2003, **3**:1157-1182.
119. Felekyan S, Kalinin S, Sanabria H, Valeri A, Seidel CAM: **Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules.** *ChemPhysChem* 2012, **13**:1036-1053.
120. Byrnes LJ, Sondermann H: **Structural basis for the nucleotide-dependent dimerization of the large G protein atlastin-1/SPG3A.** *Proc Natl Acad Sci U S A* 2011, **108**:2216-2221.
121. Byrnes LJ, Singh A, Szeto K, Benven NM, O'Donnell JP, Zipfel WR, Sondermann H: **Structural basis for conformational switching and GTP loading of the large G protein atlastin.** *EMBO J* 2013, **32**:369-384.
122. Ahmed A, Gohlke H: **Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory.** *Proteins* 2006, **63**:1038-1051.
123. Ahmed A, Rippmann F, Barnickel G, Gohlke H: **A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins.** *J Chem Inf Model* 2011, **51**:1604-1622.
124. Chen HM, Ahsan SS, Santiago-Berrios MB, Abruna HD, Webb WW: **Mechanisms of quenching of Alexa fluorophores by natural amino acids.** *J Am Chem Soc* 2010, **132**:7244-7245.
125. Mullner D: **fastcluster: fast hierarchical, agglomerative clustering routines for R and Python.** *J Stat Softw* 2013, **53**:1-18.
126. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC: **MolProbity: all-atom structure validation for macromolecular crystallography.** *Acta Crystallogr Sect D: Biol Crystallogr* 2010, **66**:12-21.
127. Delaforge E, Milles S, Bouvignies G, Bouvier D, Boivin S, Salvi N, Maurin D, Martel A, Round A, Lemke EA, Jensen MR *et al.*: **Large-scale conformational dynamics control H5N1 influenza polymerase PB2 binding to Importin alpha.** *J Am Chem Soc* 2015, **137**:15122-15134.
128. Choudhury AR, Sikorska E, van den Boom J, Bayer P, Popenza L, Sztukowski K, Jurga S, Bonomi M, Sali A, Zhukov I, Passamonti S *et al.*: **Structural model of the biltranslocase transmembrane domain supported by NMR and FRET data.** *PLoS ONE* 2015:10.
129. Boura E, Rozycki B, Herrick DZ, Chung HS, Vecer J, Eaton WA, Cafiso DS, Hummer G, Hurley JH: **Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy.** *Proc Natl Acad Sci U S A* 2011, **108**:9437-9442.
130. Doose S, Neuweiler H, Sauer M: **Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules.** *ChemPhysChem* 2009, **10**:1389-1398.
131. Noe F, Doose S, Daidone I, Lollmann M, Sauer M, Chodera JD, Smith JC: **Dynamical fingerprints for probing individual relaxation processes in biomolecular dynamics with simulations and kinetic experiments.** *Proc Natl Acad Sci U S A* 2011, **108**:4822-4827.
132. Schulze A, Beliu G, Helmerich DA, Schubert J, Pearl LH, Prodromou C, Neuweiler H: **Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism.** *Nat Chem Biol* 2016, **12**:628-635.
133. Schröder GF, Alexiev U, Grubmüller H: **Simulation of fluorescence anisotropy experiments: Probing protein dynamics.** *Biophys J* 2005, **89**:3757-3770.
134. Cristovao M, Sisamakos E, Hingorani MM, Marx AD, Jung CP, Rothwell PJ, Seidel CAM, Friedhoff P: **Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA.** *Nucleic Acids Res* 2012, **40**:5448-5464.
135. Lakowicz JR: **Principles of Fluorescence Spectroscopy.** Springer Science & Business Media; 2013.
136. Yang H, Luo GB, Karnchanaphanurach P, Louie TM, Rech I, Cova S, Xun LY, Xie XS: **Protein conformational dynamics probed by single-molecule electron transfer.** *Science* 2003, **302**:262-266.
137. Mansoor SE, DeWitt MA, Farness DL: **Distance mapping in proteins using fluorescence spectroscopy: the tryptophan-induced quenching (TriQ) method.** *Biochemistry* 2010, **49**:9722-9731.
138. Haenni D, Zosel F, Raymond L, Nettels D, Schuler B: **Intramolecular distances and dynamics from the combined photon statistics of single-molecule FRET and Photoinduced Electron Transfer.** *J Phys Chem B* 2013, **117**:13015-13028.
139. Kask P, Piksarv P, Pooga M, Mets Ü, Lippmaa E: **Separation of the rotational contribution in fluorescence correlation experiments.** *Biophys J* 1989, **55**:213-220.
140. Pieper CM, Enderlein J: **Fluorescence correlation spectroscopy as a tool for measuring the rotational diffusion of macromolecules.** *Chem Phys Lett* 2011, **516**:1-11.
141. Stahl Y, Grabowski S, Bleckmann A, Kuhnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hulsewede A, Felekyan S *et al.*: **Moderation of Arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes.** *Curr Biol* 2013, **23**:362-371.
142. Somssich M, Ma QJ, Weidtkamp-Peters S, Stahl Y, Felekyan S, Bleckmann A, Seidel CAM, Simon R: **Real-time dynamics of peptide ligand-dependent receptor complex formation in planta.** *Sci Signal* 2015:8.
143. Blom H, Widengren J: **STED microscopy — towards broadened use and scope of applications.** *Curr Opin Chem Biol* 2014, **20**:127-133.
144. Bonomi M, Pellarin R, Kim SJ, Russel D, Sundin BA, Riffle M, Jaschob D, Ramsden R, Davis TN, Muller EGD, Sali A: **Determining protein complex structures based on a Bayesian model of in vivo Förster Resonance Energy Transfer (FRET) data.** *Mol Cell Proteomics* 2014, **13**:2812-2823.
145. Vogel SS, van der Meer BW, Blank PS: **Estimating the distance separating fluorescent protein FRET pairs.** *Methods* 2014, **66**:131-138.
146. Gnutt D, Gao M, Brylski O, Heyden M, Ebbinghaus S: **Excluded-volume effects in living cells.** *Angew Chem Int Ed* 2015, **54**:2548-2551.

The effect of macromolecular crowding inside living cells is studied by a FRET based sensor. This study is of relevance to relate *in vitro* FRET-studies to live-cell measurements.

147. Senske M, Tork L, Born B, Havenith M, Herrmann C, Ebbinghaus S: **Protein stabilization by macromolecular crowding through enthalpy rather than entropy.** *J Am Chem Soc* 2014, **136**:9036-9041.
148. Sakon JJ, Weninger KR: **Detecting the conformation of individual proteins in live cells.** *Nat Meth* 2010, **7**:203-205.
149. König I, Zarrine-Afsar A, Aznauryan M, Soranno A, Wunderlich B, Dingfelder F, Stuber JC, Pluckthun A, Nettels D, Schuler B: **Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells.** *Nat Meth* 2015, **12**:773-779.
150. Ratzke C, Mickler M, Hellenkamp B, Buchner J, Hugel T: **Dynamics of heat shock protein 90 C-terminal dimerization is an important part of its conformational cycle.** *Proc Natl Acad Sci U S A* 2010, **107**:16101-16106.