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

Single-Molecule Microscopy Study of Nano-Systems:

From synthetic photo-switchable nano-devices
to the dynamics of naturally occurring transcription factors

Gregor Johann Heiss

aus

Tamsweg, Österreich

2011

Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 (in der Fassung der sechsten Änderungssatzung vom 16. August 2010) von Herrn Prof. Don C. Lamb, betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde sebständig, ohne unerlaubte Hilfe erarbeitet.

München, am 15.11.2011

_____

(Unterschrift des Autors)

Dissertation eingereicht am: 17.11.2011
1. Gutachter: Prof. Dr. Don C. Lamb
2. Gutachter: Prof. Dr. Christoph Bräuchle
Mündliche Prüfung am: 21.12.2011

Abstract

In this work, techniques were developed and used to study the properties of molecules on a single-molecule level. Single-molecule techniques have the major advantage, that in contrast to ensemble measurements, they allow a detailed insight on the distribution and dynamics of single molecules without averaging over subpopulations. The use of Total Internal Reflection Fluorescence Microscopy (TIRFM) in combination with single-pair Förster Resonance Energy Transfer (spFRET) and Alternating Laser Excitation (ALEX) allows the identification of molecular-states by making quantitative measurements of distances in the Ångström range. The development of highly sensitive photon detectors and the use of versatile labeling techniques with photostable (synthetic or genetically-encoded) fluorophores, extended the application of TIRF microscopy to *in vitro* and *live-cell* experiments. Despite reducing the complexity of biological systems down to the single-molecule level, functions of individual molecules and interactions between them can be very sophisticated and challenging to analyze. Using information theory based methods, e.g. HMM, the dynamics extracted from single-molecule data was used to illuminate protein interactions and functions.

The highly regulated process of gene transcription plays a central role in living organisms. The TATA-box Binding Protein (TBP) is a Transcription Factor (TF) that mediates the formation of the Pre-Initiation Complex (PIC). The lifetime of TBP at the promoter site is controlled by the Modulator of transcription 1 (Mot1), an essential TBP-associated ATPase involved in repression and in activation of transcription. Based on ensemble measurements, various models for the mechanism of Mot1 have been proposed. However, little is known about how Mot1 liberates TBP from DNA. Using TIRF microscopy, the conformation and interaction of Mot1 with the TBP/DNA complex were monitored by spFRET. In contrast to the current understanding of how Mot1 works, Mot1 bound to the TBP/DNA complex is not able to directly disrupt the TBP/DNA complexes by ATP hydrolysis. Instead, Mot1's ATPase activity

induces a conformational change in the complex. The nature of this changed, "primed", conformation is the change of the bending dynamics of the DNA. The results presented in this work suggest a model in which this primed conformation is a destabilized TBP/DNA complex. The interaction with an additional Mot1 molecule is required in order to liberate TBP from DNA. The effect of Mot1 on the DNA dynamics is TBP binding orientation specific. Mot1 effects on the DNA bending dynamics are strongest for molecules where TBP is bound in the inverted binding orientation. The specificity of Mot1's regulation of DNA bending dynamics suggests that Mot1 preferably "primes" TBP bound in the inverted binding orientation. The mechanistic insight into the interaction of Mot1 with the TBP/DNA complex serves as a framework for understanding the role of Mot1 in gene up- and down-regulation.

In a second project, the same single-molecule techniques were used to fabricate and evaluate self-assembled optically controllable, nanodevices. Based on the specificity of Watson-Crick base pairing, DNA was used as a scaffold to position different fluorophores with nanometer accuracy. The functionality of these nanodevices was expanded by making them optically addressable by incorporation of the switchable fluorescent protein Dronpa. Two functions have been demonstrated: Signal enhancement using Optical Lock-In Detection (OLID) and pH sensing in a *live-cell* environment.

Contents

1. Fluorescence

1.1. Fundamentals

Whenever a molecule is exposed to light, there is a finite probability that this molecule will absorb the energy equivalent of the photon. This absorbed energy can drive molecular vibrations which in turn are converted into kinetic energy in form of heat. Alternatively, the absorbed energy can be radiatively released in the form of photons. A specific class of molecules, called *fluorophores*, are highly effective in radiative dissipation of the energy in the form of photons after they have been excited into their excited singlet states[1] by a process called *fluorescence*. This absorption and emission process is governed by the electromagnetic nature of light, the structure of the molecules involved and their environment 1.1. (*Fluorophores*) are able to absorb the energy of photons $E = h\nu$ by using the photon's energy to induce an electronic transition. The probability P of a transition between two state is given by the overlap integral:

$$P = \langle \psi_f \left| \mu \right| \psi_i \rangle, \tag{1.1}$$

where μ is the dipole operator, ψ_i and ψ_f are the overall wavefunctions for the initial and finale state, respectively. The overall wave functions are the product of the vibrational, electronic and spin wavefunctions. The electron configuration of the final state, compared to the initial state, can be shifted in respect to the equilibrium position of the nuclei. Since the electronic transition is fast

---

[1] If the exited state is a triplet state (parallel paired electron spins) the lifetime of the exited state is in the order of 1×10^0 s to 1×10^3 s due to violation of the $\delta S = 0$ selection rule. This process is called phosphorescence.

compared to the nuclei movement, the *Franck-Condon principle* approximates that an electronic transition occurs without changes in the nuclei position. As a result, the contribution from the vibrational wavefunction in the overlap integral favor electronic transitions from the ground state S_0 to one of the vibrational energy levels of the exited states $S_{1,2}$. The exited state readily relaxes to the lowest vibrational level of S_1 through *vibrational relaxation* [2] and in turn looses some of its excitation energy. The vibrational ground state of S_1 is the starting point for different deexciation pathways (e.g. FRET) of which fluorescence is the only radiative one. Since some of the excitation energy is lost by the vibrational relaxation, the energy liberated by the transition from S_1 to S_0 is lower than the energy provided by the absorbed photon. This phenomena is called *Stokes shift* [2] (Figure 1.1). This spectral shift allows separation of excitation from emission with optical spectral filters.
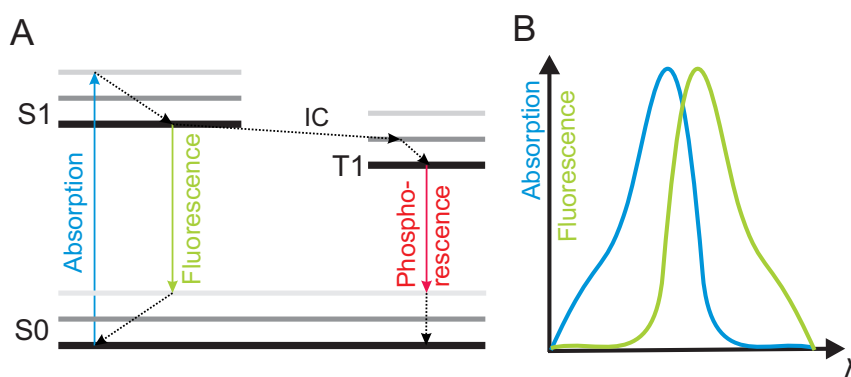

Figure 1.1.: Jablonski diagram. A: Absorption of an photon induces electronic transition into an vibrational excited state. After internal conversion the electron transfers to the vibrational ground state of S_1 from where it can return to S_0 through fluorescence or to the triplet state T_1 through intersystem crossing. The radiative transition from T_1 to S_0 is called *phosphorescence*. B: Absorption and fluorescence emission spectra illustrating the Stokes shift.

The transition from S_1 to the triplet state T_1 state does not conserve the spin. This transition requires a inversion of the electrons spin and is possible through

---

[2]The rate for vibrational relaxation is approximately one order of magnitude faster than the rate for fluorescence. Hence, fluorescence takes place from the vibrational ground sate of S_1 and the emission spectrum is independent from the excitation wavelength (Kasha Rule [1]).

spin-orbital interactions [3]. Due to the nature of those interactions the lifetime of triplet state is generally higher than the lifetime of singles states.

Besides absorption, excitation and emission of fluorophores, there are three other characterizing properties:

- The *absorption coefficient* ϵ is a measure of the probability that a photon will be absorbed.

- The *quantum yield* Φ is the efficiency of emitting a photon after absorbing.

- The *fluorescence lifetime* τ is the average time a molecule spends in the excited state.

The probability of absorbing and incident photon is fluorophore specific and can be quantified by the *Absorption* A with the dimensionless units OD. The strength ob Absorption A can be described by the ratio of incident light $I_{incident}$ to transmitted light $I_{transmitted}$ as given by

$$A = \log_{10} \frac{I_{incident}}{I_{transmitted}}. \tag{1.2}$$

The Absorbtion A [OD] is proportional to the extinction coefficient ϵ [mol$^-$1], the concentration c [mol] and the path length l [cm]. This relationship is known as the *Beer-Lambert law*,

$$A = \epsilon c l. \tag{1.3}$$

In the absence of any nonradiative decays pathways the lifetime is called the *natural lifetime* τ_n or *intrinsic lifetime* and is inversly proportional to the radiative decay rate k_{fl}

$$\tau_n = \frac{1}{k_{fl}}. \tag{1.4}$$

Due to the inevitable presence of non radiative pathways under normal measurement conditions, the measured fluorescence lifetime τ will be reduced [3]. Those other non radiative pathway k_{nr} directly competes with the radiative decay rate as given by

$$\tau = \frac{1}{k_{fl} + \sum k_{nr}}.$$ (1.5)

Any non radiative transition of the excited state energy will decrease the number of photons emitted through fluorescence and in turn the quantum yield as given by

$$\Phi = \frac{\text{Number of emitted photons}}{\text{Number of absorbed photons}} = \frac{k_{fl}}{k_{fl} + k_{nr}}.$$ (1.6)

Due to the quantum mechanical nature of electronic transitions, the transitions do not occur instantaneous, synchronized or at fixed determined times but stochastic. As a result, the fluorescence intensity of an ensemble of molecules as a function of time $F(t)$ after termination of excitation at time point $t = 0$ will decay continuously to 0. One general expression to quantify the mean time of an decay from any given function $F(t)$ is given by:

$$\tau = \frac{\int_0^\infty F(t)\, t\, dt}{\int_0^\infty I(t)\, dt}.$$ (1.7)

For an ensemble of fluorophores with a constant probability for photon emission, the decay of fluorescence intensity $I(t)$ with an initial fluorescence intensity of $I(0) = I_0$ follows a first order rate equation with the rate constant $k = \frac{1}{\tau}$:

---

[3]There are conditions where the lifetime can be prolonged such as low temperature or bimolecular interactions.

$$I(t) = I_0 \exp^{-kt} \tag{1.8}$$

1.2. Fluorescence Anisotropy

The absorption of an incident photon as well as fluorescence emission are governed by coupling between the dipole of the molecule with the electric field of the photon. Based on the molecular structure, fluorophores have an transition moment spatially aligned relative to their molecular axis. The strength and in turn the probability of electromagnetic coupling between the electric vector of the exciting photon and the transition moment of the fluorophore depend on their relative angle between them[4]. In solution experiments, the orientations of the fluorophores are randomly distributed, and an averaged relative orientation factor is used. The probability of photon absorption is the highest when the electronic dipole of the fluorophore is parallel to the electric field. The polarization of the emitted photon is determined by the orientation of the transition moment at the instant of emission and in turn the orientation of the molecule at that moment. The correlation of the polarizations between the incident and emitted photons can be quantified by the *fluorescence anisotropy* and defined as:

$$r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp}, \tag{1.9}$$

where $I_\parallel$ and $I_\perp$ are the intensities of the fluorescence emission detected with parallel and perpendicular polarization relative to the excitation light[5].

The time between photon absorption and emission, the fluorescence lifetime, determines the time during which the molecule rotates before fluorescence is emitted. Together with a measure of the rotational diffusion Θ, the *Fluorescence Anisotropy* r is given as:

---

[4]the laser used in this work for excitation provide linear polarized light

[5]r values range from -0.2 to 0.4 in ensemble measurements

$$r = \frac{r_0}{1 + \frac{\tau}{\Theta}}, \tag{1.10}$$

where r_0 is the fundamental anisotropy (without rotational diffusion) and τ the lifetime. The rotational correlation time Θ can be used to gain information about the hydrodynamic volume V of the molecule itself in an environment with given viscosity η and temperature T:

$$\Theta = \frac{\eta V}{RT} \tag{1.11}$$

In the case where proteins are labeled with synthetic fluorophores[6] consideration needs to be given to the influence of linkers. The time resolved anisotropy can display an initially high anisotropy with a fast decay due to the flexibility of the linker. This fast decay is due to the flexibility of the linker connecting the fluorophore to the macromolecule. The final/steady state anisotropy is then the value due to the anisotropy from the macromolecule to which the fluorophore is attached to.

1.3. Förster Resonance Energy Transfer

Förster Resonance Energy Transfer (FRET) is a resonant nonradiative energy transfer between two chromophores through near-field dipole-dipole coupling [4, 5, 6, 7]. The energy is transferred from the exited state of a *donor* molecule to the exited state of an *acceptor* molecule as illustrated in Figure 1.2.

Since all rates depopulating the exited state S_1 are in direct competition, the efficiency of energy transfer, E, can be expressed as ratio of the rate for FRET k_{FRET} to the sum of all radiate and non radiative rates:

$$E = \frac{k_{FRET}}{k_{fluorescence} + k_{FRET} + \sum k_{nonradiative}} \tag{1.12}$$

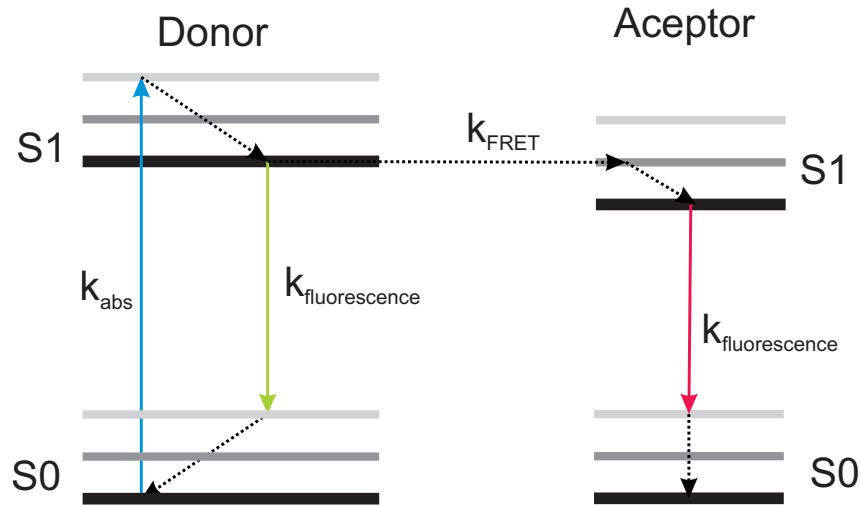[6]Proteins are commonly larger than synthetic fluorophore

Figure 1.2.: Jablonski diagram for FRET. The exited state energy from the vibrational ground state of S_1 can be dissipate by the transition to S_0 in form of fluorescence with the rate $k_{fluorescence}$ or it can be transfered to the exited state S_1 of an acceptor molecule through FRET with the rate k_{FRET}.

The rate for FRET k_{FRET} depends on following factors:

- The *quantum yield* of the donor in absence of the acceptor ϕ_d.

- The *intermolecular distance* R.

- The *spectral overlap* J between donor emission and acceptor absorption.

- The *orientation factor* κ^2 which depends on the angle between donor and acceptor dipole.

and can be calculated as[7]

$$k_{FRET} = \frac{9(ln10)\kappa^2\phi_d J}{128\pi^5 n^4 N \tau_d R^6} \qquad (1.13)$$

---

[7]this formulation is based on a dipol-approximation and valid if the intermolecular separation is much larger than the molecule dimension

where N is the the Avogadro number and n the refractive index of the medium between the dipoles. The normalized spectral overlap integral J is determined by:

$$J = \frac{\int_0^\infty F_d(\lambda)\epsilon_a(\lambda)\lambda^4 d\lambda}{\int_0^\infty F_d(\lambda)d\lambda}, \tag{1.14}$$

where $F_d(\lambda)$ is the donor fluorescence in the absence of the acceptor and $\epsilon_a(\lambda)$ the molar extinction coefficient of the acceptor.

For a given relative orientation of the acceptor with respect to the donor, the orientational factor κ^2 is given by:

$$\kappa^2 = \left(\cos\theta_{DA} - 3\cos\theta_D\cos\theta_A\right)^2, \tag{1.15}$$

where θ_{DA} is the angle between the transition dipole moments of the fluorophores ($\vec{M_A}$ and $\vec{M_D}$), and θ_D and θ_A are the angles between the transition dipole moments of donor and acceptor, respectively. The center of both dipoles are connected by the vector $\vec{r}$ as illustrated in Figure 1.3.

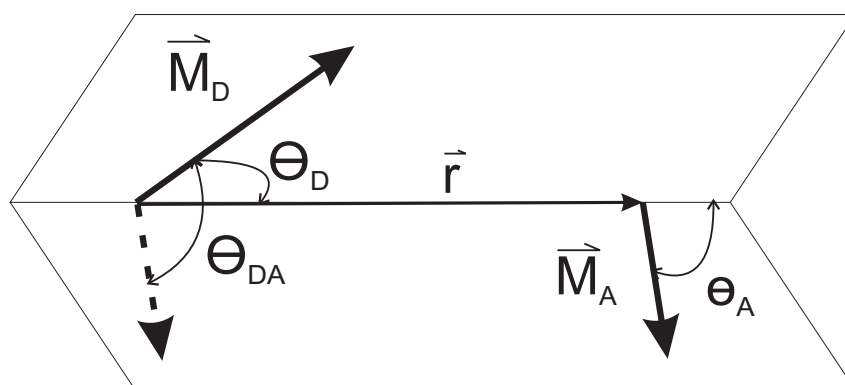**Figure 1.3.:** Definition of angles used to calculate κ^2.

In the case of non-correlated relative orientations between both dipoles, the average over all possible orientations can be calculated. This condition holds true if the linker connecting the fluorophores is flexible enough to allow rotational diffusion and in turn yields low Anisotropy values. In this case κ^2 is $\frac{2}{3}$ used [8].

The distance at which the rate for FRET equals the rate of the donor fluorescence ($k_{FRET} = k_{fluorescence}$) is defined as the *Förster distance* R_0. At this distance, the efficiency for FRET is $E = 50\%$. Based on the power law relation in Equation 1.13, the change of the rate for FRET k_{FRET} as function of the intermolecular distance R can be written as:

$$k_{FRET} = \frac{1}{\tau_d}(\frac{R_0}{R})^6.$$ (1.16)

The value of R_0 can be determined for a given FRET pair[8] and the FRET efficiency can be expressed simplified as function of the distance between donor and acceptor R.

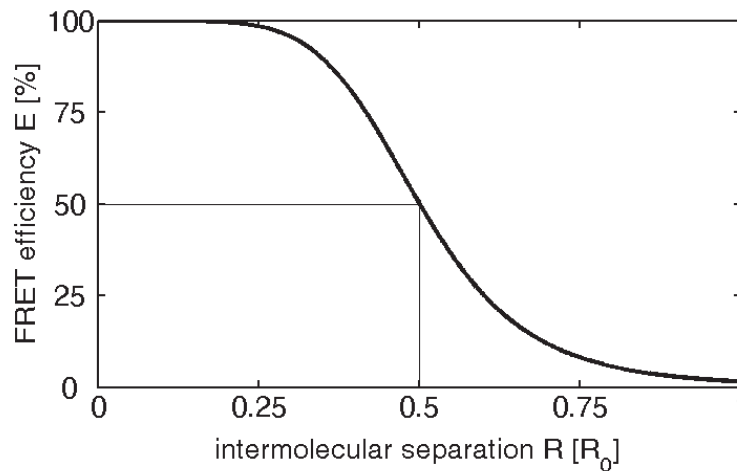**Figure 1.4.:** Power law relation between FRET efficiency E and intermolecular distance R

$$E(R) = \frac{1}{1 + (\frac{R_0}{R})^6}$$ (1.17)

---

[8]R_0 depends on κ^2 and n^4 and is sensitive to changes in the local environmental condition

2. Total Internal Reflection Microscopy

2.1. Fundamentals

One of the key factors for detecting the fluorescence signals from single molecules is the quality of the separation of the fluorescence of interest from the omnipresent fluorescent background. This quality is quantified by the Signal-to-Noise Ratio (SNR) and is defined as the ratio of the fluorescent signal to the fluorescent background. There are two apparent strategies to increase the SNR in single-molecule experiments - increase the photon flux of the signal or decrease the fluorescent background. *Total Internal Reflection Microscopy* (TIRFM) decreases the background signal by spatially limiting the excitation volume at the interface where the fluorophores of interest are immobilized. The excitation of the bulk solution, which potentially contributes to the fluorescent background, is minimized. Initially, TIRF was developed for cell imaging [9] but, with improved detection technologies, this technology is now also applicable to single-molecule experiments [10, 11].

2.2. Total Internal Reflection

In contrast to confocal microscopy, where the excitation volume is limited by the use of a pinhole, TIRFM makes use of the unique properties of an evanescent wave[12]. An evanescent wave is generated when light[1] strikes a dielectric

---

[1]evanescent waves are a solution of the general wave equation and not limited to photons from laser sources

interface (with the refractive indices n_1, n_2 and $n_2 < n_1$) at an incident angle of Θ_i exceeding the critical angle $\Theta_c = \sin^{-1} \frac{n_2}{n_1}$. The relation between the refractive indicies n and the bending angles Θ_i, Θ_t and Θ_r at an optical interface are given by Snell's law[2] (Equation 2.1) and is illustrated in Figure 2.1.
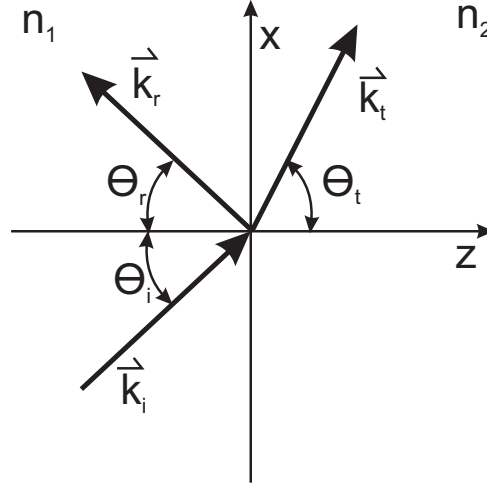

Figure 2.1.: Reflection and transmission of light at an interface in the case of non TIRF conditions ($\Theta_i < \Theta_c$)

$$\sin \Theta_i n_i = \sin \Theta_t n_t \tag{2.1}$$

The generation of the evanescent field can be derived based on the electromagnetic wave nature of light. The electric field of the transmitted wave E_t is given by:

$$E_t = E_0 e^{i(k_t \vec{r} - \omega t)}, \tag{2.2}$$

where E_0 is the amplitude, k_t is the wave vector, $\vec{r}$ the vector of propagation and ω the angular frequency.

For the geometry in Figure 2.1, the field can be written as:

$$E_t = E_0 e^{i(k_t \sin \Theta_t \hat{x} + k_t \cos \Theta_t \hat{y} - \omega t)}, \tag{2.3}$$

[2]The law of refraction was originally discovered by Ibn Sahl in 984 [13]

where $\hat{x}$ and $\hat{y}$ are the unit vectors.

Snell's law for TIRF conditions ($\Theta_i > \Theta_c$) is given by:

$$\cos \Theta_t = \sqrt[2]{1 - \sin^2 \Theta_i \left(\frac{n_1}{n_2}\right)^2}. \tag{2.4}$$

Using Equations 2.3 and 2.4, the electric field for the evanescent wave can be written as

$$E_t = E_0 e^{-\kappa z} e^{i(k_t x - \omega t)}, \tag{2.5}$$

with

$$\kappa = \frac{\omega}{c} \sqrt[2]{n_1^2 \sin^2 \Theta_i - n_2^2} \qquad \text{and} \qquad k = \frac{\omega n_1}{c} \sin \Theta_i. \tag{2.6}$$

The penetration depth d, where the energy density of the electric field $u = \frac{1}{2}\epsilon |E|^2$ decayed by the factor of e^{-1} is given by:

$$d = \frac{\lambda}{2\pi} \frac{1}{\sqrt{\left(n_1^2 \sin^2 \Theta_i - n_2^2\right)}}. \tag{2.7}$$

where ϵ is the permittivity of the medium. This penetration depth applies for s- and p-polarized excitation light [14].

The evanescent wave consists of an exponential decaying standing wave with the same frequency as the incident light and an exponential damping term. The decay of the evanescent wave amplitude for different incident angles is illustrated in Figure 2.2

The intensities of the s- and p-polarized components of the evanescent wave at the interface ($z = 0$) as a function of incident angle Θ_i are given by [15, 16]:
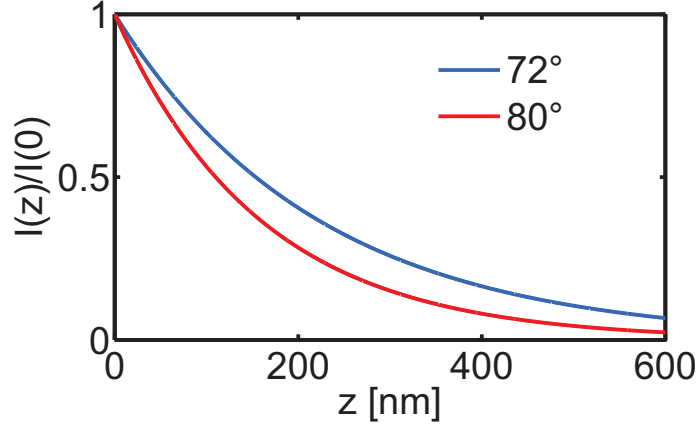
Figure 2.2.: Normalized exponential decay of evanescent field for different incident angles Θ_i for a index of refraction of $n_1 = 1.458$ and $n_2 = 1.333$ and an excitation with $\lambda = 532\,\mathrm{nm}$.

$$I_s(0) = |A_s|^2 4 \cos^2 \Theta_i \frac{n_1^2}{1 - n_2^2} \tag{2.8}$$

$$I_p(0) = |A_p|^2 \frac{4 cos^2 \Theta_i \left(2 \sin^2 \Theta_i - n_2^2/n_1^2\right)}{(n_2/n_1)^4 \cos^2 \Theta_i + \sin^2 \Theta_i - (n_2/n_1)^2} \tag{2.9}$$

and illustrated in Figure 2.3.

The polarization of the evanescent field depends on the polarization of the incident light polarization. For p-polarized light (polarized in the plane formed by the incident and reflective rays) the evanescent wave is elliptically polarized in the same plane as the incident light. Incident light with s-polarization gives rise to an s-polarized evanescent field. For both polarizations the wave front of the evanescent field travles parallel to the interface surface and is slightly phase shifted [17].
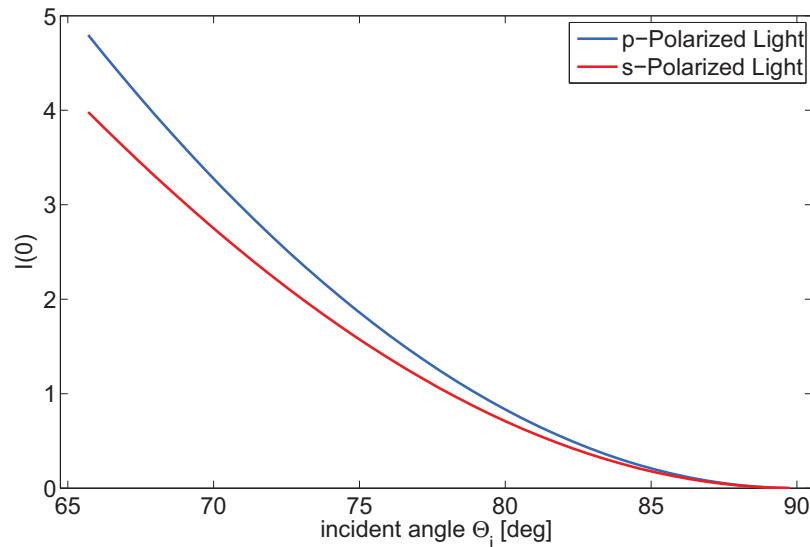
Figure 2.3.: Intensities of p- and s-polarized light of the evanescent field at the interface ($z = 0$). Intensities are normalized to incident intensities. Intensities of the evanescent field close to the interface can be several fold increased compared to the incident light. The influence of the interface to the emission field of the fluorophore limits the detectable signal [15]

2.3. Experimental Setup

There are two types of TIRF Microscopes available which differ in the way the evanescent wave is generated. In *objective-type* TIRF, the excitation is focused of-axis on the back-focal-plane and both the emission and excitation follow the same paths as in widefield microscopy. By parallel shifting the excitation in respect to the optical axis, the excitation will hit the focus plane of the objective under an angle. With sufficient off-axial positioning of the excitation TIRF conditions can be realized. There are optimized objectives for objective-type TIRF available (e.g. Apo TIRF 60x, Nikon). However, this scheme implies that most of the excitation light is reflected back to the objective. With the *prism-type* TIRF setup the excitation and emission pathways are separated within the setup. The excitation is guided through a prims to a water/quartz interface in the sample chamber which is opposite of the objective. Due to this geometry, the excitation light is reflected away from the objective. The advantage of the objective type TIRF is that it does not need a prism. This makes is suit-

able for life cell experiments, where cells are cultivated in open dishes. The prism-type TIRM has the advantage of lower background signal, making it a better choice for *in vitro* single-molecule experiments. The TIRF microscopy used in this study is based on the TIRF microscope installed by Peter Schlüsche [18] (TE 2000-U, Nikon) and was extended by two laser lines, broad-band fiber coupeling, an Acousto-Optic Tunable Filter (AOTF) with appropriate electronic controller and modified TIRF prism (illustrated in Figure 2.4). Using an AOTF in combination with a single-mode-fiber has the advantage of mechanical decoupling of the excitation source from the microscope, perfect beam profile, perfect spatial overlay of different laser lines and spectral filtering of individual laser lines[3].

As excitation sources a $405\,\text{nm}$ diode laser with $100\,\text{mW}$ (Cube 405-50E, Coherent), the $488\,\text{nm}$ line[4] of an AR:Kr Ion gas laser with $250\,\text{mW}$ (Stabilite 2018-RM, Coherent), a diode pumped Nd:YVO4 at $532\,\text{nm}$ with $100\,\text{mW}$ (GCL-100-L, CrystaLaser) and a He:Ne gas laser with $13\,\text{mW}$ at $633\,\text{nm}$ (Laser 2000) were used. The laser lines are combined by dichroic mirrors, guided though an AOTF (AOTFnC.4OO-650-PV-TN, Pegasus Optik) and coupled into a single mode polarization maintaining fiber (532/647 nm, OZ Optics) by an apochromatic[5] laser coupler (60SMS-1-4-RGBV11-47, Schäfter und Krichhoff). The laser light is decoupled by an apochromatic collimator (60FC-4-RGB11-47, Schäfter und Krichhoff) and guided to the prism surface under an angle of 72° relative to the surface used for sample immobilization. An achromatic lens ($f = 200\,\text{mm}$) is used to focus the excitation light onto the prism surface. The focusing is necessary to increase the laser energy density on the prism surface and to limited the area of excitation (and the area of photo bleaching) to the observation area. The fluorescence is collected by an 60x water immersion objective (CFI Planapochromat VC, Nikon, Na 1.2), spectrally separated by a dichroic mirror (630DCXR,AHF) and filtered through two bandpass filters (HQ585/80 and HQ715/150,AHF).

---

[3]this is of particular interest for diode lasers where beam profile and spectral bandwidth are not ideal

[4]multiple lines are available using a broadband mirror instead of a prism in the cavity
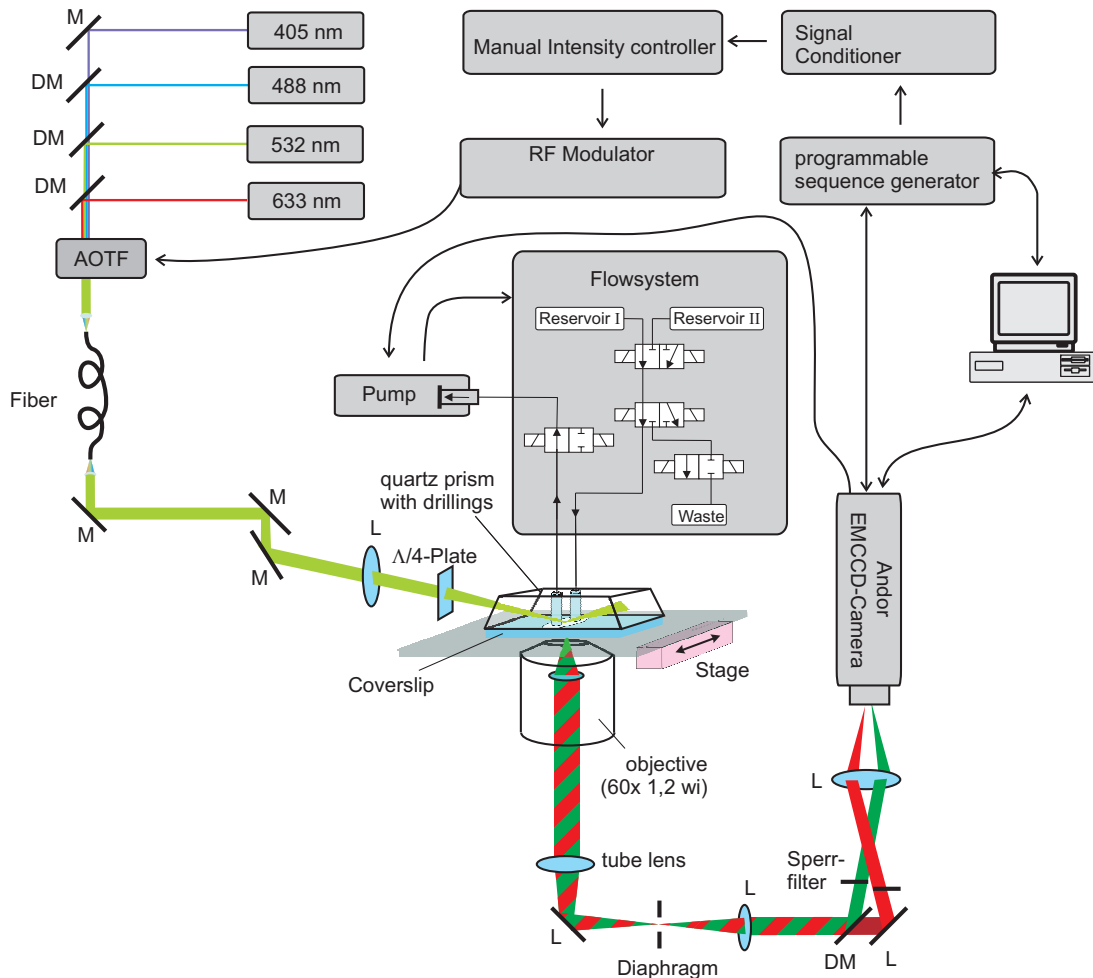
[5]corrected for three wavelengths

Figure 2.4.: Experimental Prism Type TIRF Setup (modified from [18]).

The selection of the laser wavelength used for excitation, the control over of its duration and the laser intensity is accomplished by the use of an AOTF. The AOTF is driven by an radio frequency modulator *RF-modulator* which in turn is controlled via analog input from a pulse- and amplitude generator. The interface at the prism surface is used to generate the evanescent field and to act as seal to form a chamber which can be connected to a flow system (see Section 2.3.1). The flow can either be controlled manually (connection to an syringe) or by a programmable syringe pump (PHD 2000, Harvard Apparatus), which is synchronized to the EMCCD camera (see Section 2.3.2).

Despite triggering the Andor Camera by the shutter control box with high precision, the acquisition time for a frame is determined internally by the camera. The pulse width of the triggering signal has a sub ms precision, but the acquisition time of the camera can vary several ms. This variations in exposure times can introduce additional noise to the measurements. One option to overcome this problems is to use a stroboscopic illumination. Using the AOTF the illumination time can be precisely controlled. Between exposures, the illumination can be switched off. This avoids unnecessary exposure which contributes to photo-bleaching and avoids exposure of the CCD chip during the frame transfer. Furthermore, processes much faster than the time resolution of the camera can be visualized. Instead of averaging the fluorescence signal during the whole exposure time, only a snapshot is taken. However, this does not increase the effective time resolution but can gain excess to processes not observable otherwise.

2.3.1. TIRF Prism

Commonly prism type TIRF setups attach a prism via an index matching fluid to a preformed flow chamber. With this method, the flow chamber can be formed by two cover slips separated by a spacer (e.g. Nesko film). In this work, the flow chamber was formed directly at the prism surface. This minimizes reflections by avoiding the introduction of an additional interface in the excitation path. The TIRF prisms are custom made (Zell Glass) from synthetic fused silica (Suprasil 2) with dimensions shown in Figure 2.5. The angle of the prisms are chosen to avoid dispersion on the incident side and to ensure overlapping foci for different excitation wavelengths. The prism angles are asymmetric to avoid that the beam exiting the prism will be reflected back to the area where the surface immobilized molecules are imaged.

Since sample immobilization takes place directly on the prism surface, the prism need to be cleaned and surface functionalized for each measurement. This tedious procedure is optimized by forming two flow chambers on each prism surface. This design allows two independent experiments with the same prism. Due to increased complexity of the spacer geometry, the nesko film spacer used

15 mm   D.: 1.45 mm +- 0.1 mm

4 mm

24

24 mm

4 mm
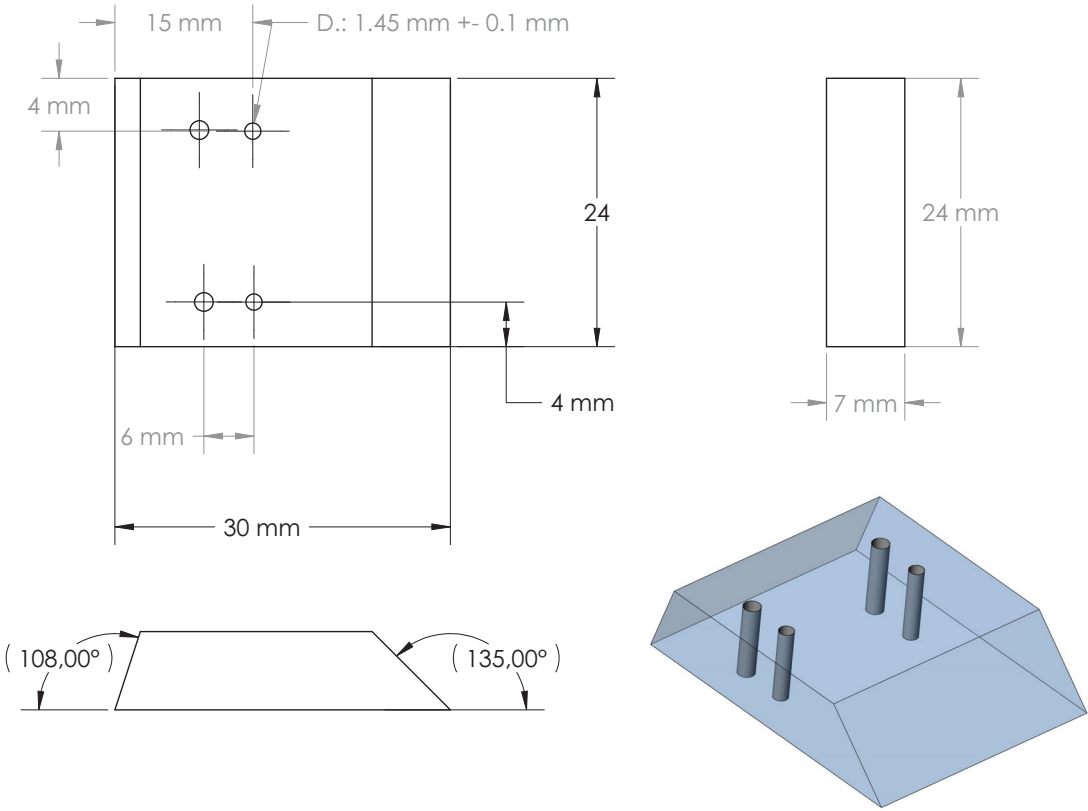
6 mm

7 mm

30 mm

(108,00°)   (135,00°)

Figure 2.5.: Flowchamber used for prism type TIRF

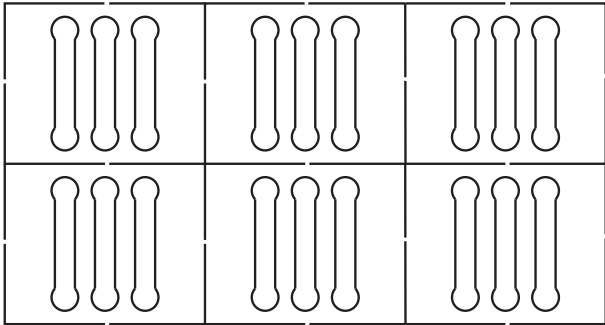to seal the flow chamber is cut by a laser-cutter. The template for the nesko seal is shown in Figure 2.6.

Figure 2.6.: TIRF Prism space template for laser the cutter

2.3.2. EMCCD

An *Electron-Multiplying Charged Coupled Device* (EMCCD) is based on a charge-coupled device (CCD) extended by a solid state Electron Multiplying (EM) register [19] [20]. The *frame transfer* mode enables faster acquisition rates by spatially separating image exposure and readout. This enable the simultaneous readout of the previous frame while the next frame is exposed. The gain register amplifies the charge based on the principle of impact ionization in the charge domain prior AD conversion [21]. Each shift in the EM register applies a minor amplification gain g which potentiates after N shifts to the Gain G as given by[6]

$$G = g^N. \tag{2.10}$$

The process of frame transfer and EM gain amplification is illustrated in Figure 2.7.

The output of an EMCCD is given as *greylevels* or *Analog Digital Units* (ADU).

In general, EMCCD cameras suffer from 4 types of noise:

- *Dark Current* - generated in the absence of light

- *Clock Induced Charge* - from reading the signal

- *Shot noise* - poison statistic of photon arrival times

- *excess noise factor* or *Multiplicative Noise* - distribution of amplification by impact ionization

- *Readout Noise* - noise from video chain electronics

Dark counts are randomly generated charges in the depletion area of the CCD Sensor. This electron-hole pairs can be generated thermally or by energy deposition of cosmic rays. Due to their sporadic occurrence they can be considered of shot noise nature. A typical values for Dark Current is $\sigma_{DC} =$

---

[6]voltages of $30 - 60\,\mathrm{V}$ for each shift yields single-step gain values of $1 - 1.6\,\%$
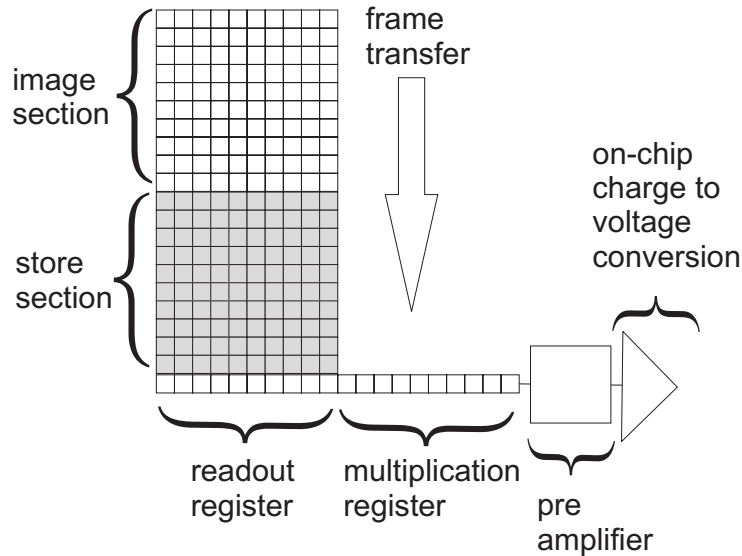
Figure 2.7.: Schematics of a frame transfer EMCCD readout [22]. The image section of the
CCD chip is exposed to light and is the area of photoelectron creation. In a frame transfer
step the electrons are transfered from the image-section to the store-section from where
the readout and gain multiplication is feed.

1×10^{-3} electron/pixel/sec. Readout Noise and Clock induced Charge (CIC)
are charges generated during each readout process by shifting the charge in
the register. The number of shifts a charge undergoes depends on the posi-
tion within the chip. For the sake of simplicity this noise factor is given as
an average [23]. Typical values are $\sigma_{RN} = < 1$ electron/pixelreadout and
$\sigma_{CIC} = 1$ electron/line respectively. This two noise factors are exposure time
independent and occur for every read frame once. This source of noise is also
the key factor when comparing single photon counting detectors. Despite the
high *Quantum Efficiency* (QE) of up to 85% only when the photon flux is high
enough an EMCCD provides better SNR than a GaAs detectors [23].

Due to the quantum mechanic nature of light, the photons energy is absorbed in
distinct packages of energy. The number of energy packages is finite and their
arrival time is poissonian distributed. The amount of shot-noise is a function of
the number of particles N detected[7] and can be calculated as $\sigma_{SN} = \sqrt{N}$.

---

[7]quantum efficiency impacts shot noise by decreasing the number of detected photons

Due to the probabilistic nature of impact ionization the amplification factor at impact ionization step varies. Based of theoretical considerations this excess noise factor, F, can be calculated as a function of the number of shift steps N and the EG gain G [22]. This noise factor is also called multiplicative noise and is defined as

$$F^2 = \frac{\sigma_{out}^2}{G^2 \sigma_{input}^2} \tag{2.11}$$

and can be calculated by

$$F^2 = 2(G-1)G^{\left(\frac{-N+1}{N}\right)} + \frac{1}{G}, \tag{2.12}$$

wherer G is the effective Gain and N the number of elements in the EM shift register. For common settings for single-molecule experiments ($G > 200$ and $N = 512$) $F^2 = 1.98$.

The Readout Noise[8], $\sigma_{readout}$, is the contribution of noise after the EM register. This type of noise is not amplified by the EM gain and is due to the analog electronic and AD conversion. The readout noise is only amplified by the analog *pre amplification gain* A. Taking all noise factors into account, the total Noise σ_{total} is given by

$$\sigma_{total} = \sqrt{A^2 \sigma_{readout}^2 + F^2 G^2 A^2 (\sigma_{dark}^2 + \sigma_{signal}^2 + \sigma_{CIC}^2)}. \tag{2.13}$$

The signal S can be written as function of the number of photons incident to the detector surface[9]

$$S = G \cdot n_{photons} \tag{2.14}$$

The *Signal to Noise Ratio* (SNR) is given by

---

[8]also called *video chain gain*

[9]for the sake of simplicity, the quantum efficiency QE is set to 1

$$SNR = \frac{S}{N} = \frac{S}{\sqrt{\frac{\sigma_{readout}^2}{G^2} + F^2(S + \sigma_{signal}^2 + \sigma_{CIC}^2)}}. \qquad (2.15)$$

In order to convert the ADU into the absolute number of photons it is essential to know the absolute gain value. Some camera manufacturer (including Andor) offer EM-CCD camera systems with *linear gain* and claim a 1:1 relation between the gain set in the acquisition software and the physical gain. The simplest way to confirm linearity of the gain is to measure the mean ADU counts as a function of the gain. The illumination is kept constant within the linear range of the EMCCD. This mean intensity values can be normalized to the mean intensity at Gain zero to obtain the relative amplification. One method to determine the absolute Gain directly without the need for normalization is the *mean variance test* [22]. As the name suggests, the mean variance test uses the functional elation between the output signal S_{out} and the noise factors σ_{signal} and σ_{dark} as given by

$$S_{out} = AG(S + S_{dark}) = AG(\sigma_{signal}^2 + \sigma_{dark}^2), \qquad (2.16)$$

where A is the total video chain gain, S the the mean signal from a photon source and S_{dark} the mean signal from the background[10]. The total variance on the camera output signal is given by:

$$\sigma_{tot}^2 = A^2 G^2 F^2(\sigma_{signal}^2 + \sigma_{dark}^2) + A^2 \sigma_{readout}^2. \qquad (2.17)$$

Combining Equation 2.16 with Equation 2.17 we get the relation between the total signal noise σ_{tot} and the mean output signal S_{out}:

$$\sigma_{tot}^2 = AGF^2 S_{out} + A^2 \sigma_{readout}^2 \qquad (2.18)$$

---

[10]commonly a constant bias is added by the acquisition software

The video chain gain, A, can be determined by setting the gain value to 1 and recording the mean ADU values for different exposure times[11]. The noise in the image is calculated by calculating the variation in the difference of two images. The background is determined for every shutter setting. Since the mean-variance test requires a significant amount of images to be taken, the process of image acquisition and analysis was automated with MATLAB using the Andor SDK package.

As seen in Equation 2.18, the video chain gain is the slope of the fit shown in Figure 2.8.
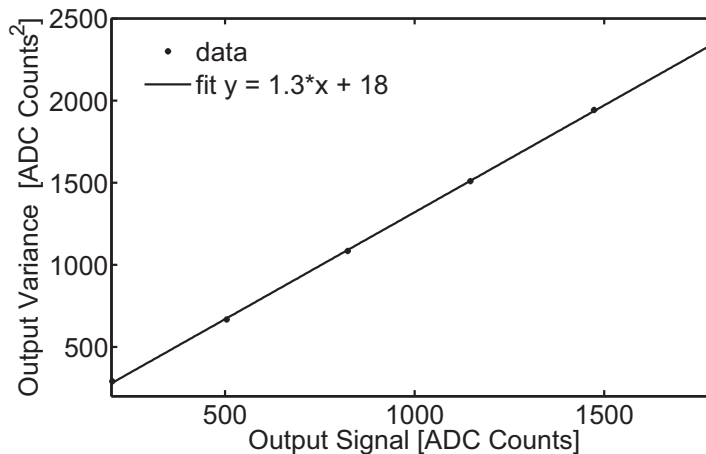

Figure 2.8.: The output variance from an EM-CCD camera vs. the output signal measured at
    EM-gain setting of 1. The slope of the curve (1.3) is the video chain gain A.

The relation between software gain and hardware gain can be obtained by repeating the analysis used to obtain the video chain gain, but now for different gain values, G. Since the video chain gain was determined in the previous step, the absolute hardware gain can be calculated directly[12] and is shown in Figure 2.9.

The physical gain is linear but the slope is less than one (reproducible for independent measurements). The meaning of a slope of .91 is that if the gain is

---

[11]The video chain gain, A, includes the on-chip charge to voltage conversion factor. A software preamp factor
    setting of 5.2 was used

[12]the excess noise factor F^2 is set to 2
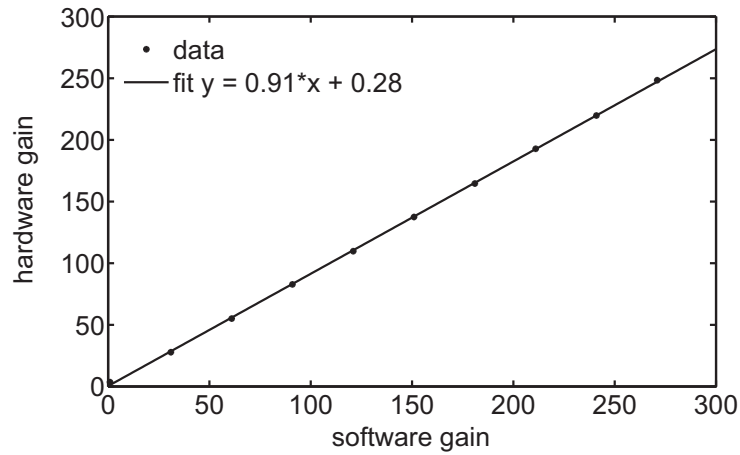
Figure 2.9.: The physical (hardware) gain as function of the software gain. The relation is linear with a slope of 0.91 and an intersect of 0.28.

set in the software to e.g. 100, the physical gain is only 91. The relation between software and hardware gain need to be taking into account, whenever the absolute number of photons are calculated from the ADC counts.

3. Data Analysis

To extract information from single molecules, a number of analysis routines needed to be programmed. One of those routines is a software toolbox for information extraction, conditioning and analysis of single-molecule TIRF (smTIRF) data. For details about the *Trace Intensity Analysis* program (TRACY), see Appendix A.5. The TIRF microscope in this work was used to record fluorescence intensities of individual surface immobilized fluorescent molecules. The size of the fluorescent emitters is significantly smaller than the *point spread function* of the setup, which results in the detection of diffraction limited spots. In contrast to confocal imaging techniques, where the sample is scanned and the image is reconstructed from a sequential stream of intensity data, widefield TIRF uses a camera to record intensities from an area as *image stacks*. In the case of a camera, the spectrally separated fluorescence signals from one area of the surface are imaged on different regions (=channels) on the camera chip. The workflow for extraction of these fluorescence intensities from different channels and their correction for detection efficiencies and optical abberations is illustrated in the flow chart shown in Figure 3.1.

The toolbox can also be used to analyze data obtained from other single-molecule experiments, e.g. confocal images of surface immobilized complexes.

3.1. Mapping

In many Microsopy applications the fluorescence intensity of the sample is recorded with a digital camera. EMCCD cameras have one of the highes detection efficiencies available. However, an EMCCD camera can not distinguish
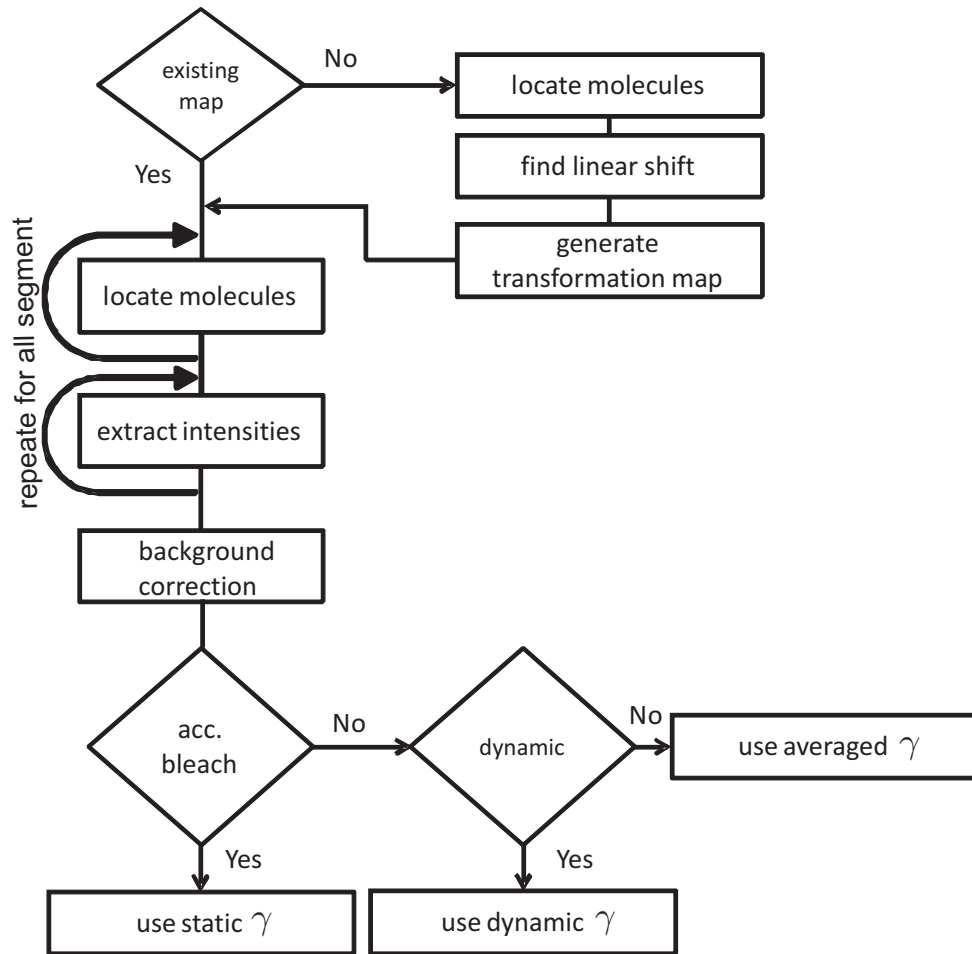
Figure 3.1.: Flow chart for smTIRF analysis. Optical abberations between spectrally seperated channels are corrected by a transformation map. This map is generated by approximating the linear shift between single moleucles within each channel and computing a spacial transofrmation (see Section 3.1). The intensity information from each molecule is extracted in a two step process. First, the position of the molecules within the whole image stack is determined. Secondly, the signal- and local background-intensities are extracted (see Section 3.2). After applying background corrections, the gamma factor is determined and used to correct the total intensity and the FRET values (see 3.3)

between photon energies and it operates in an intensity modus. For applications where the energies of the photons is of interest e.g. FRET, livel-cell imaging with multiple labels, the fluroescence need to be seperated spectrally. In order to be able to spectrally seperate the fluorescence and at the same time maintain the high detection efficiency of an EMCCD camera, dichroic mirrors are

used. With the help of dichroic mirrors, the spectrally seperated images can be imaged on different regions (=channels) of the camera chip. The individual channel can be overlayed in a post processing step. However, non linear distortions introduced through optical abberations are always present and will cause the channels to be spatially distorted in respect to each other. As a result, the images of the individual channels only overlap to a certain extend and a correction for the optical abberations within the channels in the form of a *transformation map* is neccesarry. In the case of TIRF microsocpy where the FRET efficiency between two chromophores is of interest, it is sufficient to only correct the location of the chromophores to identify pairs of fluorescent spots which belong to the same molecule.

The process of generating the image transformation between two channels is a two step process. In the first step, the locations of the spots in both channels are identified as shown in Figure 3.2. In the second step, the locations of the spots are matched to identify fluorescent spots that belong to the same molecule. Matching pairs of spots are used to generate the transformation map.

Individual spots within one channel are identified by a "clean" - algorithm.

The SN ratio is improved by using an intensity image which is averaged over a user-defined number of frames. The background of the intensity image is set to zero by applying a threshold. The threshold is calculated based on the intensity histogram of the image in such a way that the resulting image contains only 5% of the brightest pixels[1]. Within this thresholded image, the pixel with the maximum intensity is located. A *Region Of Interest* (ROI) around this pixel is selected and used for analysis as shown in Figure 3.3. The pixel values of the ROI are multiplied with a circular mask and the result is averaged to obtain the average signal[2]. A second ring shaped mask is used to calculate the local background around the spot[3] as shown in Figure 3.3.

---

[1]450 non overlaping molecules would cover less than 5% of the image pixels.

[2]The number of photons collected from a particular molecule can be calculated by integration of the background corrected signal under consideration of the photoelectron conversion factor and EM-gain.

[3]The calculation of the background can be refined. Background pixels with intensities above a certain standard deviation from the mean background can be ignored.
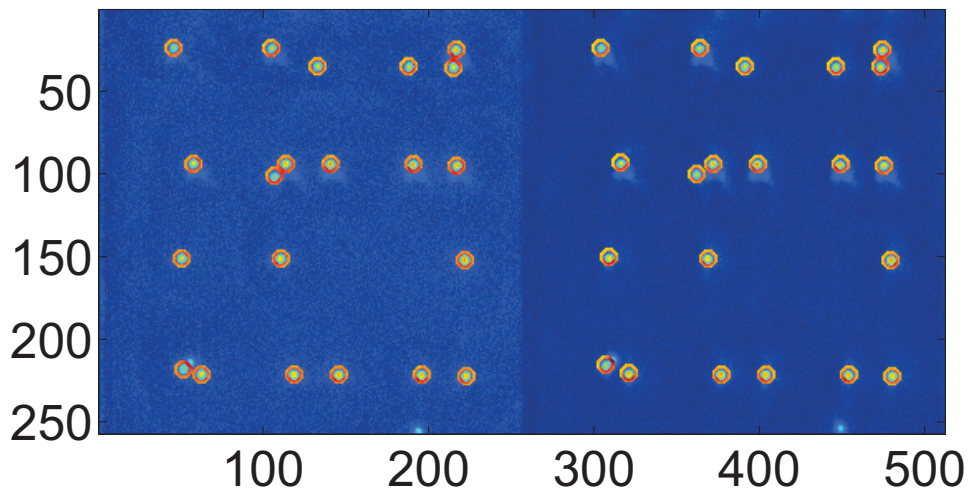
Figure 3.2.: Image of fluorescent beads immobilized at the prism surface. Individual beads are
localized as indicated by circles in both channels. The left half is the green channel and
the right half the red channel.

Figure 3.3.: Intensity image of a surface immobilized bead. Left: full ROI, middle: ROI with
applied mask before calculation of the mean signal, right: ROI with applied background
mask used to average the local background signal around the spot.

The ratio between the mean signal and the mean background intensity is used to judge if a fluorescent spot is circular and of diffraction limited size. If the signal originat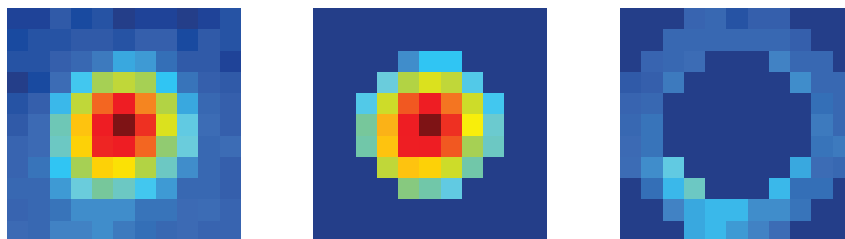es from a source which is bigger than the diffraction limit, the fluorescence will extend to the pixels used to calculate the background. As a result, the signal to background ratio will be smaller and this spots can be identified and discarded.

After analysis of the ROI the pixel values of the ROI are "cleared" by setting them to zero. The analysis is then repeated as long as there are non-zero pixels left.

If the distance between two spots is small enough to detect the fluorescent signal from one spot in the background of the other spot, both spots are discarded.

Once the fluorescent spots in both channels are localized, pairs of spots need to be identified which belong to the same molecule. In a first approximation the linear shift between both channels is determined. This linear shift can be approximated by stepwise linear shifting the coordinates of the spots found in one channel. The distance between the locations of the spots in one channel to the shifted spots in the other channel is calculated. The number of pairs whose distance is below a certain tolerance is determined. The tolerance is chosen big enough to allow matching of pairs despite nonlinear abberations but small enough to avoid formation of pairs which do not belong to the same molecule. The optimal linear shift is reached when the maximum number of matching pairs can be found as shown in Figure 3.4. The locations of the spots at the optimal shift position are used as "anchor" points for the generation of the transformation map.

The locations of the spots in one channel are used as *base points* and the locations of the corresponding spots in the other channel as *input points*. The image transformation is generated in a way that the *base points* are mapped to the *input points*. The image transformation was implemented using the build-in *cp2tform* and *imtransform* function in MATLAB. The *cp2tform* function can infer a spatial transformation between a set of vectors. Each position of a pixel within an image can be considered to be a vector. The *cp2tform* function was used to infer a spatial transformation between the pixels of two images based on
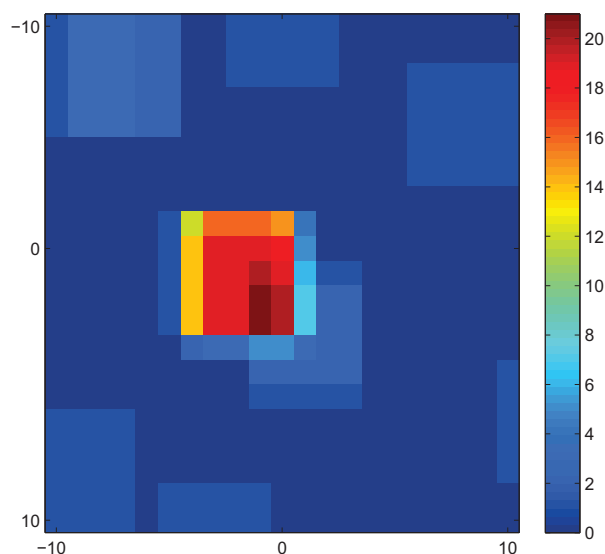
Figure 3.4.: The number of matching pairs found as function of a linear shift in x and y direction between the channels.

the input and base points. A third order polynomial was used as tranformation function. To infer the mapping function, a minimum of 10 mapping points is required. In the case of an image containing molecules which undergo FRET, the locations of the fluorescent donor and acceptor molecules can be used as mapping points in both channels. Alternatively, a movie from surface immobilized fluorescent beads can be recorded and their locations used. The use of beads with a broad fluorescent spectra (e.g. Nil Red, Kisker Biotech) have the advantage of a robust localization due to a stable and intense fluorescent signal of the same bead in both channels. Even when used at stock concentrations, the beads attach to the surface slowly. The quality of the transformation map depends on the density of the beads e.g. if many mapping points are available in one part of the image, the mapping in this part of the image will be better. A little trick was used to generate mapping points covering the full field of the image despite the slow attachment rate of the beads. Individual images with only a few immobilized beads are recorded several times with a slight shift between them. Those images are combined by averaging them to give an image with a sufficient amount of control points as shown in Figure 3.2.

Figure 3.5.: Images of the overlay of the green and red channel before and after applying the image transformation map to the red channel. The transformation properly corrects the any optical image distortion. The transformation map is only applyed to the image to visualize the performance of the mapping. The raw image data is always used for extraction of the intensity information and analysis.

When fluorescent beads are used to generate a map, each bead will be visible in both channels and no mismatch can occur. If a sample with FRET is used, the locations of non fretting molecules does not correlate. The number of random mismatches will be small in comparison to the number of correctly matched pairs and will not effect the approximation of the linear shift. However, wrongly matched pairs can distort the mapping function locally. The quality of the map can be checked by applying the map to the whole image. A zoom in on two beads before and after image transformation is shown in Figure 3.5.

The transformation map is saved in the data directory. Prior to image analysis, the presence of a map is checked and, if no map can be found, a new one is generated.

3.2. Extraction of single-molecule intensities

The size of the image stack obtained by the TIRF microscope depends not only on the resolution in x, y but also on the number of frames recorded. For the settings used in this work, an image stack has a size of about $500\,\mathrm{MB}$ for every

1000 frames. The size of the image stack which can be loaded at once into the memory of a PC is limited by hard- and software. As a consequence, the number of frames which can be analyzed is limited if the whole stack needs to be present in memory at the same time. To bypass this limitation, the image stack is divided into segments with user-defined length and each segment is analyzed sequentially. As a consequence, the image stack need to be loaded two times. The first time the image stack is loaded to localize the molecules in all frames within the whole image stack and the second time to extract the intensity information at the positions identified in the first step. For the localization procedure, a user defined number of frames out of the image segment is averaged and within this average the individual molecules are identified. This procedure is repeated for all segments with the exception that ROIs where molecules have already been found are excluded. For some applications, it is sufficient to localize individual molecules which are present at the beginning of the recording and extract their intensity traces. This work flow also allows the identification of molecules which are not present at the beginning of the recording but bind at any time later.

Based on the transformation map which corrects for the relative optical distortion between two channels (see Section 3.1), there are two approaches to extract the intensity information of a single-molecule from both channels. One is to transform one channel with the inverse transformation which caused the distortion and then localize the molecules in one channel to determine its intensity. The location of the molecules in the other channels are given by the linear shift between the channels. The other approach, which was implemented in this work, is to localize the molecules in both channels and use the image transformation to transform their location only. Fluorescent spots which transformed locations match can be considered to belong to the same molecule. The intensities are then extracted from the raw data. The transformation of the whole image is computational more expansive than to transform only the locations of a few molecules. Furthermore, the transformation of the raw intensity images modifies the size of the fluorescent spots and would change the fluorescence intensities recorded from them. The signal and local background intensities of pairs of spots (*traces*) which belong to the same molecule are extracted using the same circular masks as shown in Figure 3.3.

3.3. Signal correction

The fluorescence intensity traces from single molecules, obtained by methods described in the previous section, need to be corrected for the background signal (extracted as described in the previous Section), relative detection efficiency γ and spectral cross-talk β. The background can be either directly subtracted or smoothed by a moving average prior subtraction. Direct subtraction of a noisy signal does add additional noise. To minimize this additional noise introduced by background subtraction, the background can be smoothed with a moving average filter. Alternatively, a user defined number of frames at the end of each trace can be averaged, assuming that the fluorescent signal is photobleached by then. This single averaged intensity value is subtracted from the whole signal trace and does not add any noise factor to the signal. The γ-factor accounts for the relative detection efficiencies (η_A and η_D) and for differences in the quantum yields ϕ_A and ϕ_D in different channels:

$$\gamma = \frac{\eta_A \phi_A}{\eta_D \phi_D} \tag{3.1}$$

The γ-factor is different for every molecule and can be determined experimental by comparing the fluorescent intensities before and after acceptor bleaching as shown in Figure 3.6 and Equation 3.2.

$$\gamma - \beta = \frac{\langle I_A \rangle_{before} - \langle I_A \rangle_{after}}{\langle I_D \rangle_{after} - \langle I_D \rangle_{before}} = \frac{\Delta I_A}{\Delta I_D} \tag{3.2}$$

Alternatively, if the trace shows dynamic in FRET, the γ-factor can be adjusted to minimize the variations in the total intensity.

This approach can be implemented practically by minimizing the variance of the total intensity as a function of the γ-factor as described in Equation 3.3 and shown in Figure 3.7.

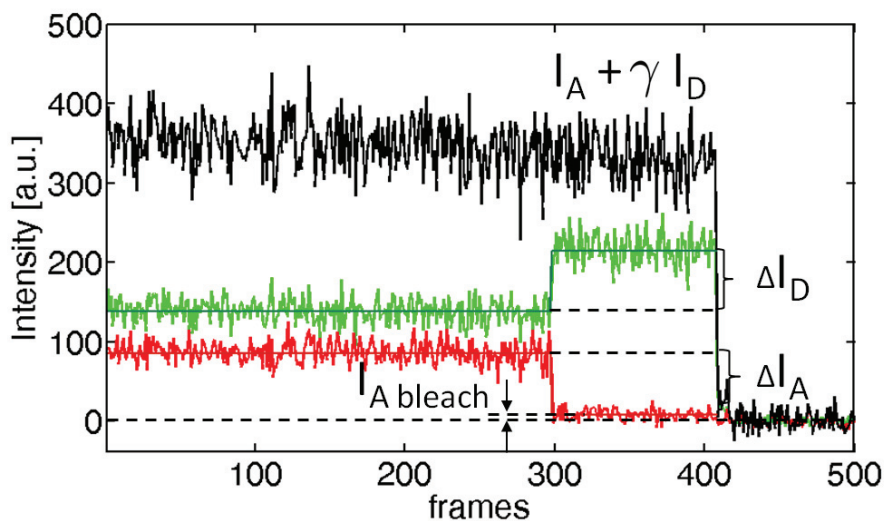$$\frac{d}{d\gamma}\mathrm{std}(I_A + \gamma I_D) = 0 \tag{3.3}$$

Figure 3.6.: γ- and β-factor determination from static FRET traces.

This method assumes that different states within the dynamic FRET trace have the same γ-factor. The γ-factor can change when the local chromophor environment changes, e.g. through protein binding. If the γ-factor for different FRET states differs significantly, the γ- corrected total intensity will correlate with the FRET trace.

The γ-factors obtained by minimizing the standard deviation of the total intensity does estimate the true γ-factors very precisely as shown by a simulation with 500 traces as shown in Figure 3.9.

The γ-factors obtained from static or dynamic traces can be averaged and applied to molecules where no gamma could be determined. A flow diagram illustrating the procedures from this Section is shown in Figure 3.1. The γ-factor can be adjusted at any point during or after the analysis without the need to reanalyze the data. A typical distribution of the γ-factor is shown in Figure 3.10.

The spectral cross-talk is accounted for by a second correction, the β-factor. The β-factor accounts for crosstalk from the donor to the acceptor channel and can be calculated by the ratio of red to the green fluorescence which is detected after acceptor photobleaching (see Figure 3.6). Without any spectral crosstalk,
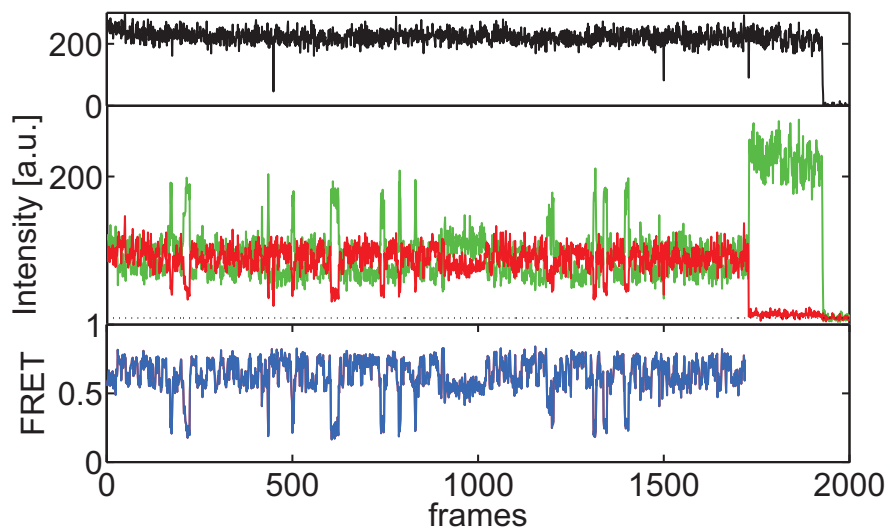
Figure 3.7.: Alternative γ-factor determination from dynamic FRET traces. Upper panel: total Intensity with dynamic gamma correction. Middle panel: Donor and acceptor fluorescence. Lower panel: FRET trace with transitions between three states.

Figure 3.8.: Alternative γ-factor determination from dynamic FRET traces. The plot shows the standard deviation of the total intensity from the trace shown in Figure 3.7 as a function of the γ-factor. At the minimum of the standard deviation, the γ-factor is 0.55.

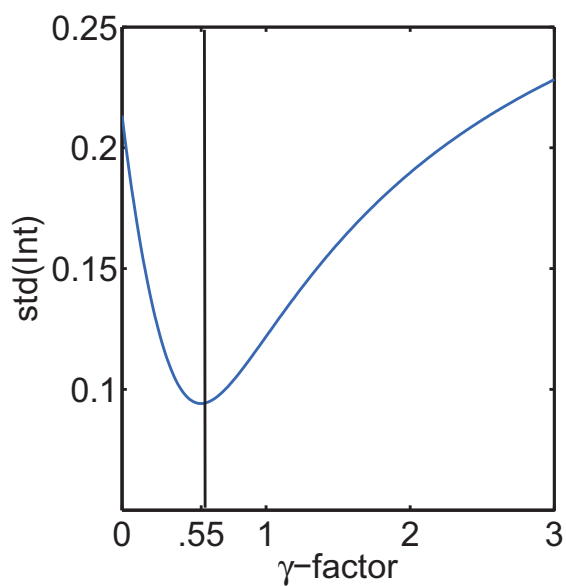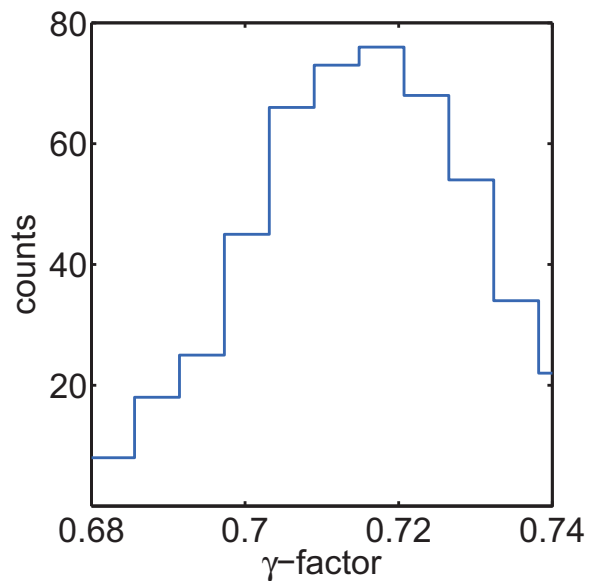**Figure 3.9.:** Distribution of the γ-factor determined from 500 simulated dynamic FRET traces with 500 frames each. The FRET values used for the simulation are 30% and 70% with a γ-factor of 0.7. When fitted to a Normal distribution, the mean μ and variance σ values are $0,713$ and $0,0153$. The slight shift to higher γ-factors is due to the asymmetry in the function of the standard deviation as seen in Figure 3.8
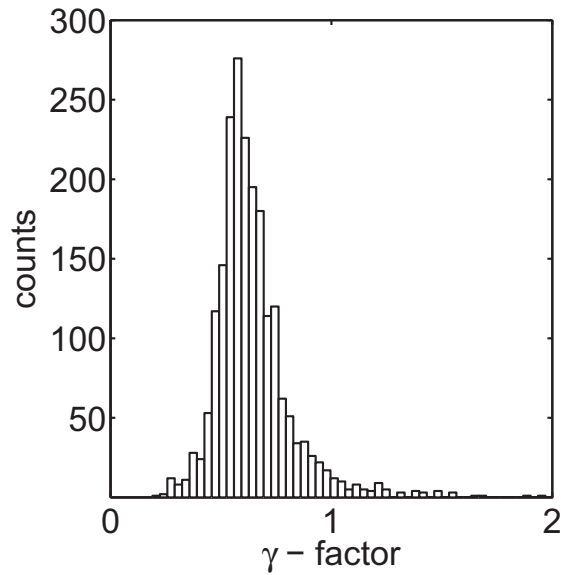
Figure 3.10.: Distribution of γ-factors for the Atto647N/Atto532 dye pair on DNA with a mean value of 0.77. The γ-factor is sample specific and can vary significantly e.g. if a chromophor is attached to a DNA or Protein

no fluorescence would be observed in the red channel after acceptor bleaching. However, due to the spectral properties of the dyes some of the donor emission spectra will always overlap with the emission of the acceptor.

$$\beta = \left| \frac{\langle I_A \rangle}{\langle I_D \rangle} \right|_{\text{after acc. bleach}} \tag{3.4}$$

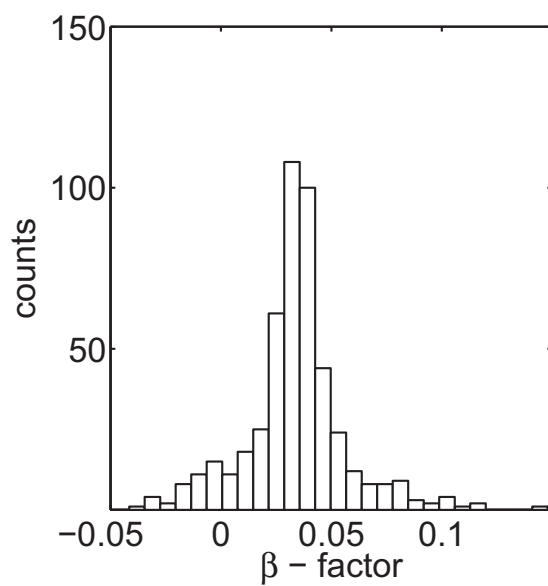A typical distribution of β-factor is shown in Figure 3.11.

Figure 3.11.: Distribution of β-factors for the Atto647N/Atto532 dye pair on DNA with a mean value of 3.5%

3.4. The Hidden Markov Model

In a *Markov Model* (MM), the discrete states of the system are directly observable and the sequence of states is governed by the transition probability between the states. Similar to a MM, a *Hidden Markov Model* HMM is a model that describes a system assuming a set of discrete states and a constant transition probability between them. When the states are "hidden" e.g. behind noise and only the (noisy) outputs are observed a HMM can be used. As the name suggests, the transitions between the states are governed by a *Markovian process*, a stochastic process without memory. For the case of a time-series of states, this means that the probability of transition to another state only depends upon the current state of the system. As a result, each transition between two states can be described by a probability or single rate constant. For HMM, a second stochastic process is involved in the generation of the observable output. The distribution of the output from each hidden state is called *emission function*.

The framework for HMM was developed in the 1960s [24, 25] and used for speech and pattern recognition [26, 27, 28]. It was soon widely used for e.g. pattern recognition within electrocardiography (ECG) curves [29], modeling the current through ion channels [30] or DNA sequence identification [31]. For a review about development and applications of HMM see [32]. Recently, the use of HMM was extended to single-molecule fluorescence techniques. After the potential use of HMM was demonstrated through simulations [33, 34], the first quantitative results were published in 2006, where HMM was used to quantify the number of fluorescent labeled domains by finding the number of bleaching steps in fluorescence intensity traces [35]. Another application was to monitor RecA filament formation on DNA by spFRET with single monomer resolution [36]. Recently, a new method was introduced to extend the HMM approach in single-molecule experiments to non-Markovian processes by using the HMM to find variations in intensity trajectories [37].

Model Parameters

A HMM model is governed by two stochastic processes. One process describes the transition between discrete states of a time-series, and the other process describes the distribution of the observations given the states. The time series $Q = \{q_t\}$ consists of a sequence of N discrete states S_i. The state q_t at time t depends only on the prior state q_{t-1} (see Equation 3.5). The probability for transitions between two states is given by the transition probability matrix $P_{ij} = p_{i \to j}$ and is constant over time (see Equation 3.6).

$$P(q_{t+1} = S_i | Q) = P(q_{t+1} = S_i | q_t) \tag{3.5}$$

$$P(q_{t+1} = S_i | q_t) = P(q_{t+2} = S_i | qt + 1) \tag{3.6}$$

The second stochastic process is the generation of the observations $X = \{x_t\}$ from the states S_i by the *emission function* $f_i(x)$. The distribution of the observed emission from the state S_i is given by f_i. This *emission function* f_i can be a multivariant Gaussian distribution. In this work, the FRET values of each state S_i were modeled as single Gaussian distributions with mean μ_i and variance σ_i as given in Equation 3.7.

$$f(x | \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{\left(x - \mu_i^2\right)}{2\sigma_i^2}} \tag{3.7}$$

The knowledge about the system is incorporated in the *prior* π. In the case of single-molecule FRET traces, the prior defines the probability distribution of the initial FRET state. Since, in the experiments reported here, no information about the initial FRET state is available, all FRET states are assumed to occur with the same probability. This assumption can be expressed as a *flat prior*, where each initial FRET state has the same probability to occur. For the application of HMM on single-molecule data, the HMM model M consists of the following three parameters:

1. The transition matrix P which contains the transition probabilities:

$$p_{ij} = p(q_t = S_i | q_{t+1} = S_j) \qquad (3.8)$$

2. The prior π which contains the prior knowledge about the system e.g. distribution of initial FRET values

$$\pi_i = p(q_{t=0} = S_i) \qquad (3.9)$$

3. The number of N states and the distribution of outputs. e.g. N Gaussian distributions with mean μ_i and variance σ_i

$$f_i(x) = p\left(x | q_t = S_i\right) = f\left(x | \mu_i, \sigma_i\right) \qquad (3.10)$$

Transitions between the different states in the *hidden* layer generate through the *emission function* a sequence of observations. A scheme of a HMM with three states is shown in Figure 3.12.
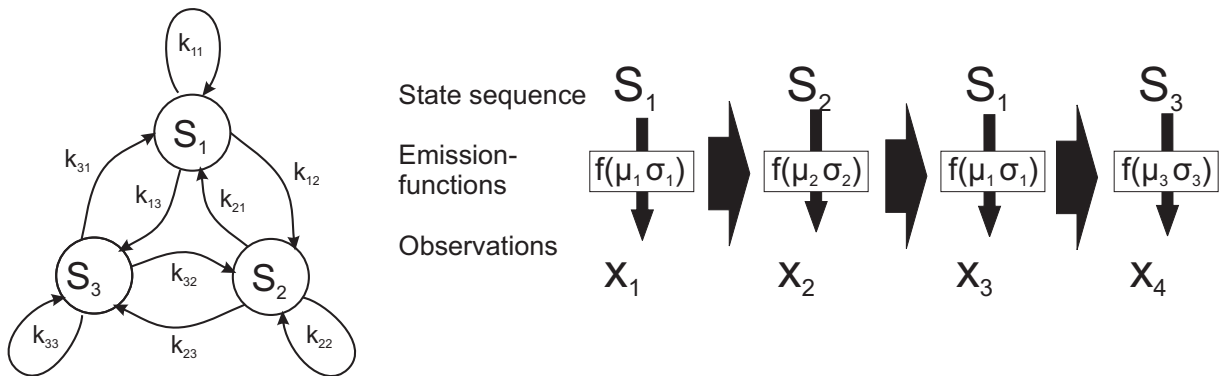
Figure 3.12.: Schematic of a HMM with 3 states and generation of the observations S_i. Left: Three states S_i are connected by single exponential rate constants k_{ij}. Each state S_i transfers with the rate k_{ii} to itself. The relation to the state dwell time is given by Equation 3.14. Right: An example for a hidden state sequence $Q = S_1, S_2, S_1, S_3$ which generates the observation sequence $X = x_1, x_2, x_3.x_4$ through the emission function $f(\mu_i, \sigma_i)$.

The transition rates K can be calculated from the transition probabilities P as shown in Equation 3.11.

$$K = \frac{(P - \mathbb{1})}{\Delta t} \tag{3.11}$$

where $\mathbb{1}$ is the identity matrix and Δt the duration of one step in the sequence e.g. the exposure time of the camera.

While the transition rates might be intuitive to calculate, the dwell times (diagonal elements of K) need a closer look. The dwell time of a state is defined by the average time the HMM model will stay in the same state. A state sequence Q where the state S_i is conserved for d time steps is shown in Equation 3.12.

$$Q_\tau = \left\{ q_1 = S_i, q_2 = S_i, \cdots, q_{d+1} = S_{i \neq j)} \right\} \tag{3.12}$$

The probability that a HMM M will stay in a state i for d time steps and transfers to any other state j at the time $d + 1$ is given by:

$$P(Q_\tau | M) = p_{ii}^{(d-1)}(1 - p_{ii}) = p_i(d). \tag{3.13}$$

With this probability density function of the duration d of state S_i, the average time in state S_i can be calculated by using a Taylor series expansion [38] as:

$$\tau_i = \sum_{d=1}^{\inf} = dp_i(d) = \frac{1}{1 - p_{ij}} \Delta t \tag{3.14}$$

Training

The aim of training a HMM is to modify its parameters in such a way that the likelihood that the HMM model M can describe the observations X is maximized. Based on the optimized HMM, the underlying hidden state sequence Q can be retrieved. In other words, the HMM parameters M are trained with the observation X to maximize the probability of the observation given the Model $P(X|M)$. The training of an HMM involves two steps:

1. The likelihood that the observation X can be described by the model M must be calculated $P(X|M)$

2. The model parameters M must be modified to maximize $P(X|M)$

The probability to make an observation X given a sequence of hidden states Q is given by:

$$P(X|Q, M) = p(x_1|q_1) \cdot p(x_2|q_2) \cdots p(x_T|q_T)$$
$$= \prod_{t=1}^{T} p(x_t|q_t) = \prod_{t=1}^{T} f(q_t). \tag{3.15}$$

The probability of a certain hidden state sequence given the HMM model parameters M is given by:

$$P(Q|M) = p(q_1, q_2, \cdots, q_T|M) = p(q_{t-1}|q_t) \cdots p(q_{T-1}|q_T)$$
$$= p(q_1) \prod_{t=2}^{T} p(q_t|q_{t-1}) = \pi_{q_1} \prod_{t=2}^{T} p_{q_{t-1}q_t}. \tag{3.16}$$

The probability for a certain observation of a sequence X is given by the joint probability of Equation 3.15 and 3.16. By summing over all possible N^T state sequence Q_i, the probability of the observation of the sequence X given the HMM model $P(X|M)$ can be written as:

$$P(X|M) = \sum_{Q_i} P(X|Q, M) \cdot P(Q|M) \tag{3.17}$$

For each possible state sequence, $2T$ calculations are needed in the sum of Equation 3.17 which yields a total of $2TN^T$ calculations. Since it is clearly not feasible to use Equation 3.17 to calculate $P(X|M)$ a more efficient approach is needed. A commonly used training algorithm was developed by Baum et. al. which makes use of an *forward-backward* algorithm to calculate the probabilities at each time step independently [39, 38, 40]. This is possible because the

Markov conditions state that each state only depends on its previous one. At each time step t, each of the N states can undergo a transition into N states, making TN^2 possible paths for one trace. The probability α to be in state j at the time point t can be calculated recursive by the product of the transition probability p_{ij} and the probability that the observation x_t is due to emission from the state S_j. This emission probability can be calculated by evaluating the emission function f_j for the observation x_t. The *forward estimate* $\alpha_t(j)$ is given by:

$$
\begin{aligned}
\alpha_1(j) &= \pi_j f_j(x_t) \\
\alpha_t(j) &= \sum_{i=1}^{N} \alpha_{t-1}(i) p_{ij} f_j(x_t) \qquad \text{with} \qquad 2 \le t \le T.
\end{aligned}
\tag{3.18}
$$

This probability contains all transitions until time t. The probability β_i of transitions after t can be calculated similar by summing over the transitions recursive starting at the end $t = T$ of the path. This *backward estimate* $\beta_t(i)$ is given by:

$$
\begin{aligned}
\beta_1(i) &= 1 \\
\beta_t(i) &= \sum_{i=1}^{N} \beta_{t+1}(j) p_{ij} f_j(x_{t+1}) \qquad \text{with} \qquad t = \{T-1, T-2, \cdots, 1\}
\end{aligned}
\tag{3.19}
$$

The probability $\gamma_i(t)$ of being in state i at the time step t can be calculated by the product of Equation 3.18 and 3.19 normalized by the time invariant likelihood $P(X|M)$.

$$
\gamma_i(t) = P\left(q(t) = S_i | X, M\right) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{i=1}^{N} \alpha_i(t)\beta_i(t)} = \frac{\alpha_i(t)\beta_i(t)}{P(X|M)}
\tag{3.20}
$$

Additionally, the probability $\epsilon_t(i, j)$ for a transition from state i at the time t and to the state j at the time $t + 1$ can be written as:

$$\epsilon_t(i,j) = P\left(q_t = S_i, q_{t+1} = S_j, X | M\right) = \frac{\alpha_i(t)p_{ij}f_j\left(x_{t+1}\right)\beta_j(t+1)}{P(X|M)} \quad (3.21)$$

The optimization of the HMM parameters is an iterative, two-step process. In the first step, the likelihood is estimated (E-Step) and in the second step, this estimate is used to modify the parameters to maximize the likelihood (M-Step). These steps are then iteratively repeated until convergence is reached. There are two approaches to derive the EM-algorithm, a classic one based on estimation of the number of transitions for discrete observations x [38] or a general one based on the derivative of the likelihood function to find the optimized parameters [41, 42].

In the classical derivation, during the E-Step the probability of transitions between two states ($S_i \rightarrow S_j$) are estimated by summing $\epsilon_t(i,j)$ over all time steps. This probability can be interpreted as the number of transitions expected between two states and is given by:

$$\sum_{t=1}^{T-1} \epsilon_t(i,j). \quad (3.22)$$

The overall number of transitions from state i can be estimated by summing the probability of being in state i:

$$\sum_{t=1}^{T-1} \gamma_i(t). \quad (3.23)$$

Using Equations 3.22 and 3.23, optimized model parameters can be calculated in the M-Step. The new transition probability p_{ij}^{new} is the expected number of transitions between $S_i \rightarrow S_j$ divided by the expected number of transitions from S_i.

$$p_{ij}^{new} = \frac{\sum_{t=1}^{T} \epsilon_t(i,j)}{\sum_{t=1}^{T} \gamma_i(t)} \tag{3.24}$$

The new prior π_i^{new} is the probability of being in state i at time $t = 1$:

$$\pi_i^{new} = \gamma_i(1). \tag{3.25}$$

The probability for making the observation $x_t = f_i(x_t = S_i)$ at time t in the state i is the expected time in the state S_i while observing a certain $x(t)$ divided by the total amount of time spent in state S_i:

$$f_i(x)^{new} = \frac{\sum_{t=1}^{T} \gamma_i(t) * \delta_{x,S_i}}{\sum_{t=1}^{T} \gamma_i(t)} \tag{3.26}$$

For continuous values of x (as in the case of FRET), we have continuous emission functions $f_i(x)$ and the equations need to be adjusted to obtain the mean and variance variables. The updated values for the mean μ_i and sigma σ_i for a HMM with single Gaussian distributed emission functions is given by:

$$\mu_i^{new} = \frac{\sum_{t=1}^{T} \gamma_i(t) x_t}{\sum_{t=1}^{T} \gamma_i(t)} \tag{3.27}$$

$$\sigma_i^{2\,new} = \frac{\sum_{t=1}^{T} \gamma_i(t)(x_t - \mu_i)^2}{\sum_{t=1}^{T} \gamma_i(t)}. \tag{3.28}$$

Once the optimized HMM parameters are found, the next step is to find the single most likely sequence of hidden states Q given the model M and the observation X. Since the HMM provides not only the distribution of the observation values but also the probabilities for the transitions between the states, a less likely transition to a certain state will also result in a lower population of this state. Without this knowledge about the transition probabilities, one could only use the knowledge about the emission function to determine the most likely

state sequence. This would mean that there is always a finite probability to visit any state even if this state actually is never populated. Similar to the *forward* algorithm the probability $\delta_t(j)$ of the most likely sequence terminating at the state j at time T can be calculated recursively by:

$$\delta_t(j) = \max_{1 \leq i \leq N} \left[\delta_{t-1}(j) p_{ij} \right] f_j(x_t) \quad 2 \leq t \leq T \text{ and } 1 \leq j \leq N. \quad (3.29)$$

By choosing the state j with the highest final probability at time T and back-tracking, the single best state sequence can be found. This method is named after its inventor Viterbi who introduced it in 1967 [43].

Application of HMM to spFRET

In single-molecule applications, FRET is used as a molecular ruler to measure intra- and intermolecular distances. The information from FRET can be used to monitor protein and DNA conformations e.g. during protein folding, protein-protein interactions or protein-DNA interactions like TBP to DNA binding. In these applications, a shot-noise limited FRET distribution can be interpreted as one conformational state. Commonly, states are identified as subpopulations in FRET histograms with Gaussian distribution. However, the distribution of the measured FRET values can also be used to obtain information about the distribution of the underlying FRET states by employing a *Photon Distribution Analysis* (PDA) [44].

For a static Gaussian-distributed FRET signal, the "hidden" FRET state can be determined directly without a HMM by computing the average FRET value. The different FRET states within a dynamic FRET signal can theoretically be separated manually or by means of a HMM. A manual selection of the different states suffers from a subjective bias of what is to be considered a FRET state and practical feasibility[4]. If the separation of the FRET states is significantly larger than the noise within the FRET signal, identification of the FRET states based on a threshold can be performed. In this case, an automated FRET state

---

[4]The experimental data used in this work contains hundreds of thousands of transitions

segmentation based on e.g. FRET histograms, could be implemented. In contrast to the HMM approach, threshold methods do not consider transition rates or FRET distributions to identify FRET states and, therefore, will always have an inferior performance.

Several aspects need to be considered when applying a HMM to single-molecule FRET data: How many states are present, which FRET distribution to use, which parameters to train and how to train them. There are different approaches to determine the number of states. The simplest is to infer the number of states directly from the data e.g. by estimating the number of states within each FRET trace or from a frame-wise FRET histogram. If the number of HMM states was overestimated, the HMM tends to "cluster" several hidden states to very similar FRET values and not all HMM states need to be occupied in the Viterbi-path. An effective way to visualize this "clustering" of states is the *Transition Density Plot* which is described later. A more objective approach is the use of the *Bayesian Information Criterion* (BIC) [45]. Similar to a least squares fit, the likelihood of an HMM will increase with a increased number of states. With the BIC, it is possible to compare the likelihoods of different models by adding a "penalty" term for additional states. For experimental data, its application is debatable because finding a defined global maximum in the BIC as function of the number of states can be challenging [46].

One advantage of using a HMM is, that it does take the distribution of the individual states (the *emission function*) into account. The parameters of the emission function add degrees of freedom to the model and need to be properly initialized and trained. The fluorescence intensity distributions of donor and acceptor signal are broadened by background fluorescence, shot-noise and the additional noise factor of the EMCCD camera. Their intensities are Poisson distributed and can be approximated by a Gaussian distribution. Since FRET is calculated as a ratio between donor and acceptor fluorescence, the distribution of the resulting FRET signal is non trivial (see Appendix A [46]). The FRET signal is best approximated by a *beta* distribution (see Chapter 4), but for the sake of simplicity, a Gaussian approximation is commonly used [36]. Recently,

a HMM model for the analysis of single-molecule FRET traces was developed which adjusts the distribution of the emission function at each time point based on the total intensities and the FRET value [46]. When training a HMM, not all parameters need to be optimized, some parameters can be fixed to the initial value e.g. if the FRET distributions are known, it is sufficient to train only the transition matrix. Constraining the training decreases the computational cost and makes it possible to e.g. prevent artificially broadening of σ due to fluorescent blinking of the fluorophores. In principle, there are two possibilities to train a HMM on FRET data: either, an individual HMM is trained on the data from each molecule, or one HMM is trained on the data from all molecules. When training a HMM on a set of traces the FRET distribution obtained from the HMM model does not only contain the broadening of the FRET signal but also contains heterogeneities between different molecules. Special care needs to be taken in the case where subpopulations with different properties are present e.g. states with similar FRET values but with different transittiton rates might be assigned to the same state by a broader distribution.

The visualization of the HMM model can be done either by the previously introduced Viterbi-path or by means of the *Transition Density Plot* (TDP). The TDP is a two dimensional histogram where the transitions from all m molecules from an initial state q_i^m to the final state q_f^m are superimposed as two dimensional Gaussian distributions given by:

$$\text{TDP}(\text{FRET}_\text{i}, \text{FRET}_\text{f}) = \sum_{\text{transitions,traces}} e^{\frac{1}{2}\left(\frac{\text{FRET}_\text{i}-q_i^m}{2\sigma_i^m}+\frac{\text{FRET}_\text{f}-q_f^m}{2\sigma_f^m}\right)^2} \tag{3.30}$$

and shown in Figure 3.13. The diagonal elements would correspond to a lack of transition for a given time step and is not included in the TDP.

The FRET distributions- and rates can be directly obtained from the trained HMM or by the generation of a *Transition Density Plot* (TDP).

If the HMM is trained without optimizing σ, the σ values can be set to a user defined default value for all transitions. The TDP has is symmetric if the system

is in equilibrium. Each Transition between two states appears two times, first time for the forward reaction and the second time for the backward reaction. A system with transitions between two states will display two clusters in the TDP plot. In the case that all possible transitions between i states are possible, $i^2 - i$ clusters of transition will appear in the TDP plot. The number of clusters will be reduced to $2i - 2$, if only transitions between neighbors are possible.
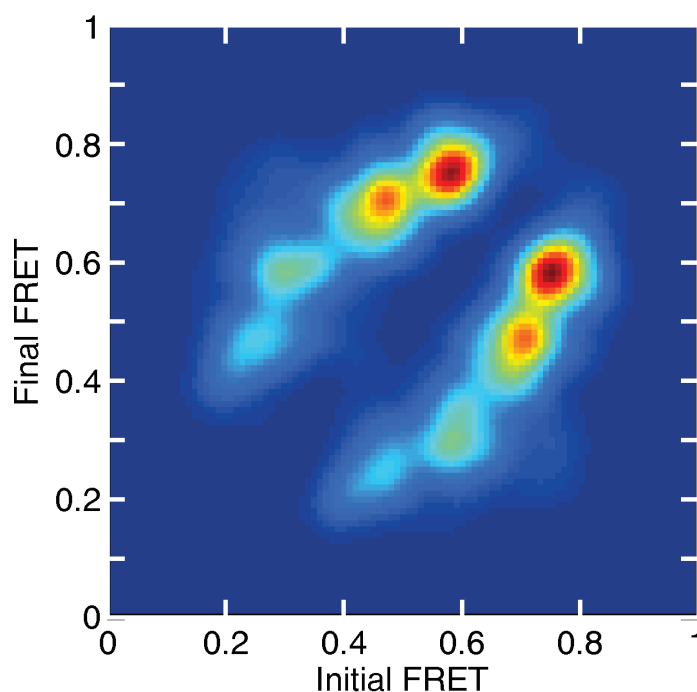

Figure 3.13.: Transition Density plot of the FRET values from a double-labeled DNA bound to TBP. The TDP is generated from a HMM trained with 243 traces containing 10153 transitions and trained with a fixed $\sigma = 0.03$. Eight clusters of transitions from 6 states are visible.

Within the TDP shown in Figure 3.13, clusters of transitions can be selected by an elliptical selection tool. A dwell time histogram of all the transitions within the selection can be composed. This dwell time histogram can be used to obtain the transition rate by fitting it to e.g. a monoexponential decay shown in Figure 3.14.
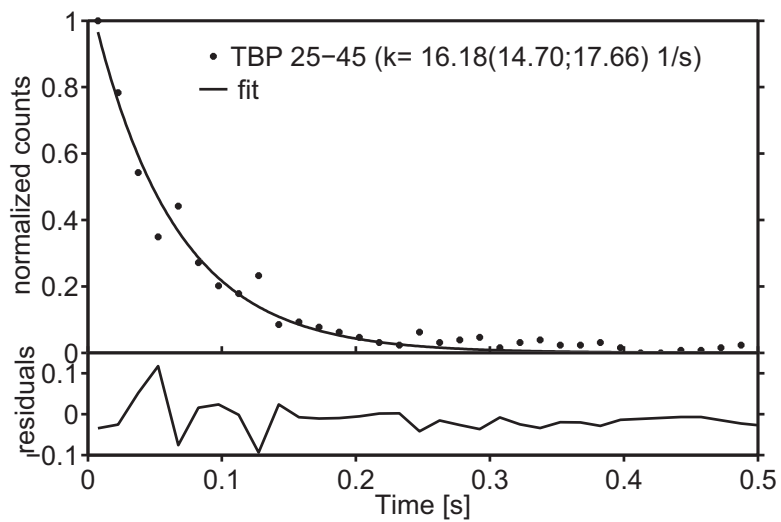
Figure 3.14.: Dwell time histogram of a cluster of transitions selected from Figure 3.13 and fitted to a monoexponential decay. The dwell time is based on the cluster of transitions from 25% to 45% FRET.

4. Single-Molecule Photoswitchable Nanodevice

Miniaturization of structures and manipulation of matter on an atomic scale have an immense potential. With the advancements in technology it is not only possible to further increase integration for e.g. computer circuitry, data and energy storage [47], but also new "enabling" technologies are emerging. Such new developments in the field of nano-bio-technology range from diagnosis to treatment of diseases e.g. bio-marker screening or drug delivery [48, 49]. In the field of material science, nanotechnology can enable synthesis of novel materials with unique properties e.g. graphene with applications yet to be discover [50, 51].

4.1. Introduction

Two general approaches for the manipulation and fabrication of objects in the nanometer scale have been proposed [52]. The *Top-down approach* where a smaller object is formed from a larger one by means of micro fabrication such as micro patterning, photo lithography, STM or AFM [53, 54]. In contrast, the *Bottom-up fabrication* uses chemical properties of single molecules for molecular self-assembly such as carbon-nanotubes, quantum-dots or DNA-origami [55, 56, 57]. The advantage of the self assembled Bottom-up fabrication method is the possibility of massive parallelization. However, adding complexity and functionality are currently a challenge.

Based on the specificity of Watson-Crick base pairing, nucleic acids are an excellent platform to direct self assembly of structures in the nanometer scale. Fabrication of complex shapes as well as implementation of mechanic functions have been demonstrated in the past [57, 58, 59]. The use of fluorescent labels increases functionality and enables the readout of the conformation of such devices for example by spFRET [60, 61]. In this project, a self-assembled DNA-based nanodevice was investigated that can be optically manipulated based on the optical switching properties of the fluorescent protein Dronpa. Using Dronpa as a fluorescent donor and the synthetic dye Atto647N as an acceptor, the fluorescent state of individual self-assembled nanodevices was read out and manipulated.

4.2. Fluorescent Protein Dronpa

The Fluorescent Protein Dronpa[1] is derived from the faintly fluorescent coral *Pectiniidae*. Its fluorescent clone (22G) was isolated from a cDNA library and expressed in *E. coli*. The monomeric version (22Gm3) of Dronpa with a weight of $28.8\,\mathrm{kD}$ was obtained by rational and random mutations [62]. It was found to have two chromophore conformations with two corresponding absorption bands at $390\,\mathrm{nm}$ and $503\,\mathrm{nm}$ [63, 64]. Due to the low fluorescence quantum yield of the conformation with the absorption band at $390\,\mathrm{nm}$, absorption at this wavelength does not yield significant fluorescence. However, this so called *dark state* can be photoconverted into the conformation with an absorption band at $503\,\mathrm{nm}$ through photoisomerization within ms time scale [65, 66, 67]. In contrast to the dark state, this *bright state* has a high fluorescence quantum yield of $\theta_{fl} = 0.85$. Apart from the fluorescence emission upon absorption at $503\,\mathrm{nm}$ the conformation of the chromophore is photoconverted into the non fluorescent dark state (see Figure 4.1) [68, 69, 70]. This conversion process is reversible. In this project, Dronpa was used in combination with a synthetic acceptor fluorophore as a switchable FRET donor.

---

[1]The name Dronpa was derived from the ninja term for vanishing "dron" and the abbreviation for photoactivation "pa"
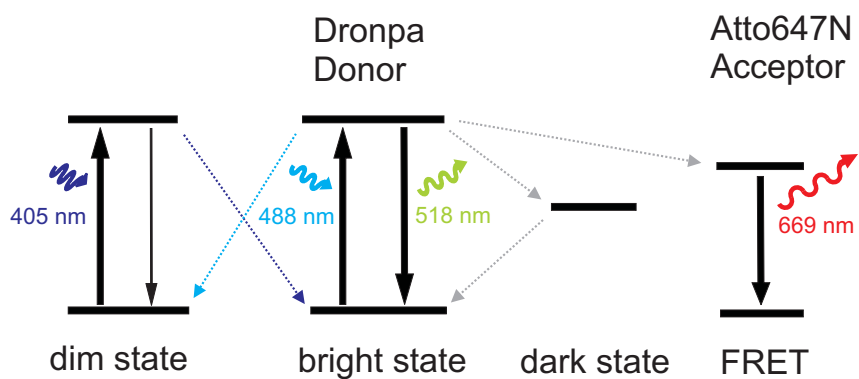
Figure 4.1.: Schematics of the photodynamics of Dronpa in the presence of an acceptor. The dim state can be excited with 405 nm light and is converted into the fluorescent bright state without fluorescence emission. Illumination of Dronpa in the bright state with 488 nm can either yield fluorescence with an emission maximum at 518 nm or induce switching to a dark state (triplet state or another non-fluorescence intermediate state) or reversible photoconversion to the dim state or FRET with the Atto647N acceptor.

4.2.1. Structure

In the fluorescent bright state, Dronpa emits green light with a maximum at 518 nm upon excitation at $488\,\mathrm{nm}$. The crystal structure revealed a GFP-like fold consisting of eleven-stranded ß-barrel enclosing an α-helix containing the chromophore, a moiety formed from the Cys62-Tyr63-Gly64 (CYG) tripeptide [64, 71, 72, 73]. In the fluorescencent bright(cis-state), the conformation of the CYG chain is slightly non-coplanar. Whereas in the dark(trans-state), the CYG ring is tilted by $\tilde{3}0\%$.
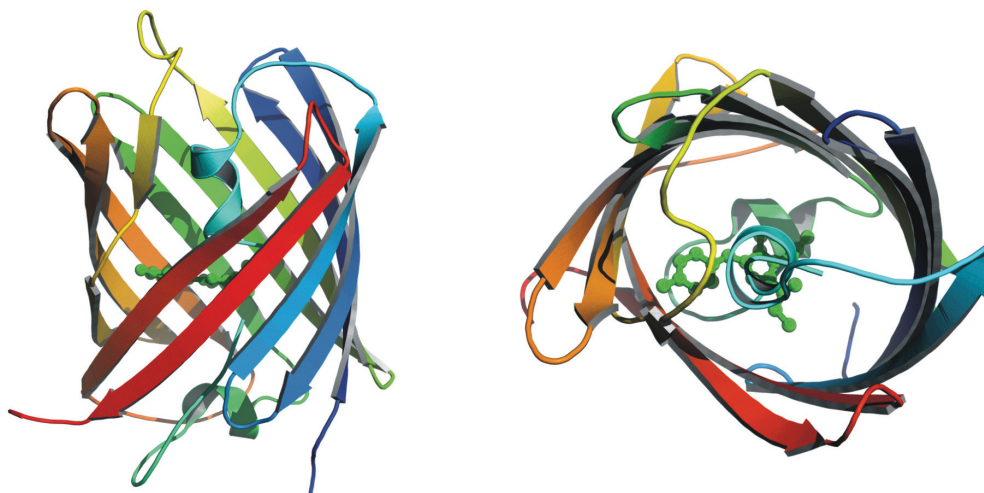

Figure 4.2.: Crystal structure of Dronpa, front and top view rendered from PDB structure 2GXO. The CYG tripeptide chromophore is shown in green as ball-and-stick model.

4.3. Sample preparation

Using bioconjugation methods, Dronpa was covalently coupled to a 20 base DNA (DNA1) strand at position 10 through a C2 amino linker. Alkylamino-modified DNA was activated with the heterobispecific crosslinker sulfosuccinimidyl-4-(N-maleimidomethyl) cyclohexane-1-carboxalate (sSMCC), which contains a thiol-reactive maleimide functionality in addition to an amino-reactive NHS

group. The resulting maleimide-modified DNA was then allowed to react with cystein residues of the Dronpa. The Dronpa-DNA conjugate was purified by Anion Exchange Fast Protein Liquid Chromatography (FPLC) [74, 75]. A second 20 base DNA (DNA2) was labeled with the synthetic dye Atto647N at positions 3 or 5 (IBA GmbH). To combine both DNA strands, a third 40 base DNA strand (DNA3), complementary to DNA1 and DNA2, was used as shown in Figure 4.3. DNA1 and DNA2 were phosphorylated and ligated using T4 Kinase and Ligase (New England Biolabs Inc). The conjugates were stored in TETBS buffer (20mM tris-HCL, 150mM NaCl, 5mM EDTA, 0.005% Tween20, pH 7.5).
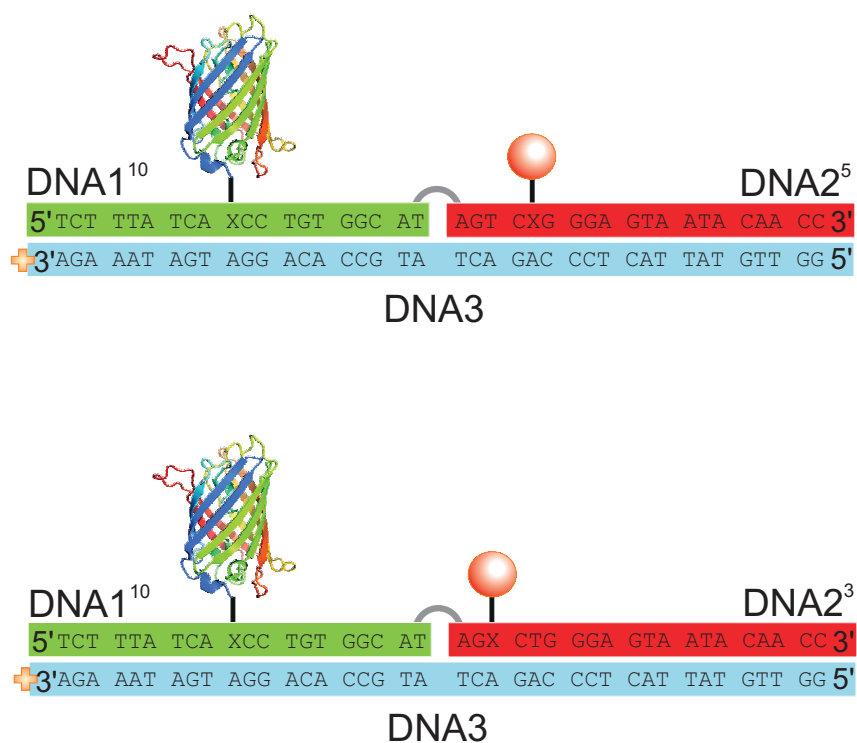

Figure 4.3.: Schematics of the two DNA nanodevices used. The complex consist of a 20 base DNA fragment linked with Dronpa at position 10 from the 5' end (DNA1), a 20 base DNA strand with Atto647N attached to position 5 (upper complex) or 3 (lower complex) from the 5' end (DNA2), and a 40 base DNA strand complementary to both DNA1 and DNA2 with a biotin attached to the 3' end. The grey link between DNA1 and DNA2 signifies the ligation of the two DNAs

Using a surface functionalized TIRF Prism the biotinylated 3'-end of DNA3 was tethered to the surface as shown in Figure 4.4.
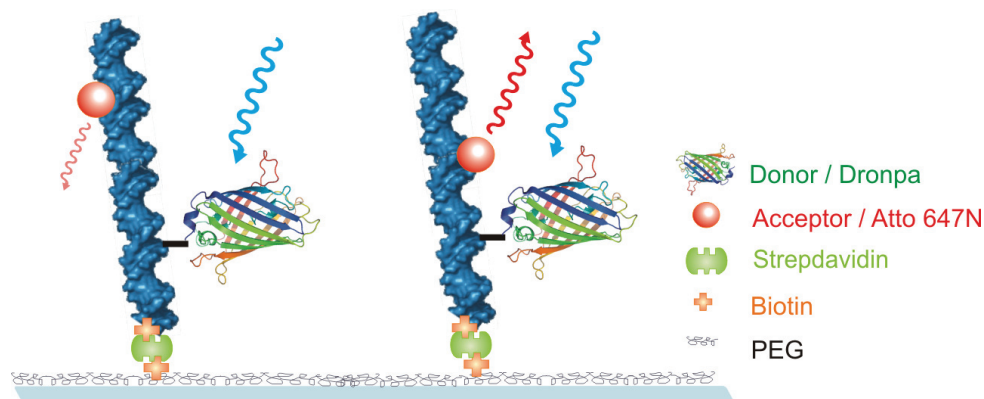

Figure 4.4.: A schematic representation of the immobilization method used for the experiments. Two construct with 15 and 13 base pairs separation between donor and acceptor molecule are shown on the left and right, respectively. Both constructs were tethered to the surface through the biotinylated 3' end of DNA3 via a PEG-biotin-streptavidin linkage.

With the Experimental Setup described in 2.3, fluorescence from individual immobilized constructs was imaged as shown in Figure 4.5.
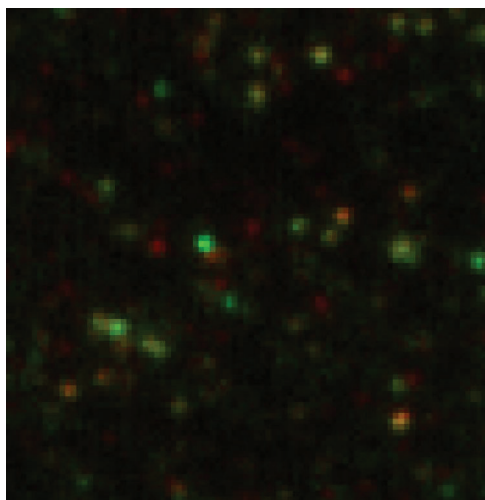

Figure 4.5.: The fluorescence detected from individual nanodevices tethered to the surface. The fluorescence from Dronpa and Atto647 are color-coded in green and red respectively.

For live-cell measurements, the construct was incubated with $300\,\mathrm{nm}$ polystyrol beads (Kisker Biotech) coated with streptavidin. Unbound DNA was removed by centrifuging beads and replacing the supernant with buffer.

4.4. Dronpa as a FRET switch

4.4.1. Single-Molecule experiments

To investigate the photoswitching properties of the nanodevice, the fluorescence intensities of individual surface-immobilizes nanodevices were recorded. The response of the fluorophore to changes in the illumination was measured by using an alternating excitation scheme. With the experimental setup described in Chapter 2.3, the fluorescence under different combinations of illumination from both laser sources ($405\,\mathrm{nm}$ and $488\,\mathrm{nm}$) was recorded. The following combination of excitation wavelengths were used for the single-molecule experiments, as illustrated in Figure 4.6, lower panel:

Table 4.1.: Summary of the excitation scheme used for the single-molecule experiments.

frames	excitation	emission	Dronpa conformation
$1-5$	$405\,\mathrm{nm}$	no fluorescence	conversion to bright-state
$6-15$	$488\,\mathrm{nm}$	exponential decay	conversion to dark-state
$16-25$	$488\,\mathrm{nm}$ + $405\,\mathrm{nm}$	constant fluorescence	continuous reactivation
$26-35$	$488\,\mathrm{nm}$	exponential decay	conversion to dark-state

The response of the donor (Dronpa) and acceptor (Atto647N) fluorescence averaged from 4092 molecules is shown in Figure 4.6, upper panel. The observed fluorescence during illumination with $405\,\mathrm{nm}$ is less than 2% of the maximum fluorescence. This is due to the low fluorescence quantum yield of the dark state.

In contrast to ensemble measurements or averaged single-molecule measurements (see Figure 4.6), fluorescence of single Dronpa molecules doesn't change
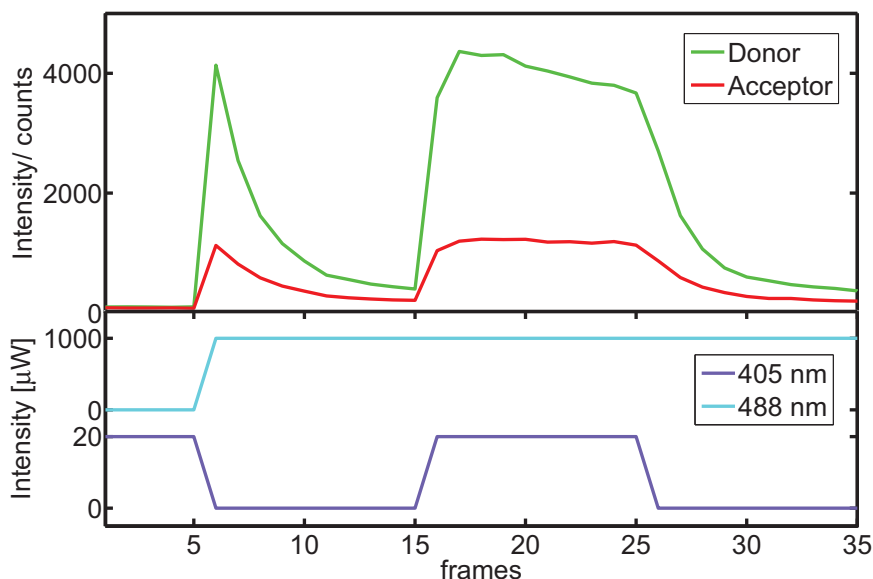
Figure 4.6.: The photoswitching of Dronpa averaged over 4092 single-molecule traces (upper panel). The excitation sequence is shown in the lower panel. The Dronpa intensity is shown in green and the Atto647N intensity is shown in red.

gradually during photo-conversion. Photoisomerization is completed very fast compared to the time resolution of the experimental setup [76] and in turn only discrete transitions between the bright and dark states are observed as shown in Figure 4.7.

During illumination with $405\,\mathrm{nm} + 488\,\mathrm{nm}$ light, the donor and acceptor fluorescence is approximately constant. This shows the stability of the construct indicated by static FRET. To evaluate how sensitive the FRET measurements are to changes in the position where the Dronpa protein is linked to the DNA, both constructs shown in Figure 4.3 were investigated. Since the Atto647N dye used is much more photostable than Dronpa, no photobleaching of Atto647N could be observed, which would be needed for determination of the γ-factor. To evaluate changes in FRET and in turn distance, the *Proximity Ratio* (PR) was used. The PR, given in Eq. 4.1, was used to quantify the amount of energy transfered between Dronpa and Atto647N molecule:
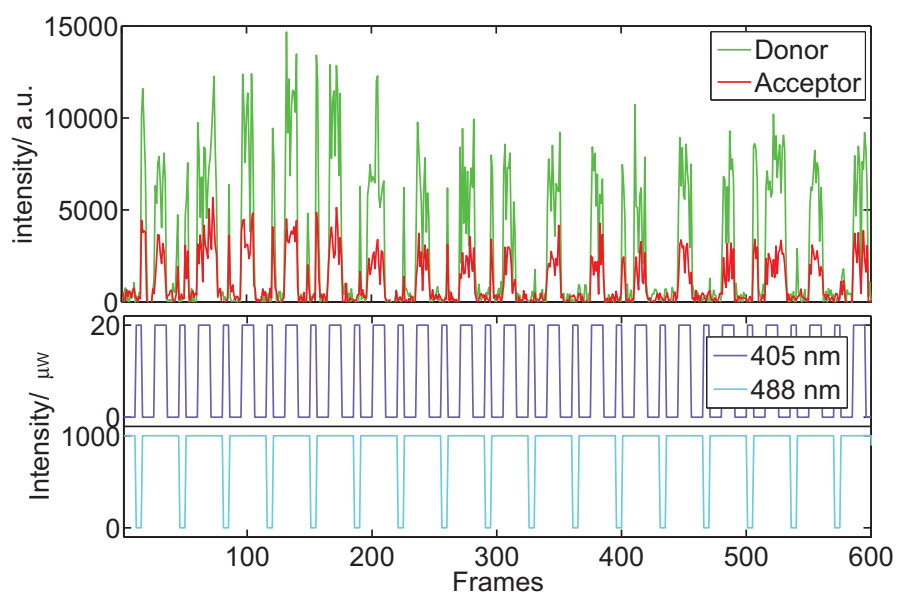
Figure 4.7.: The photoswitching of a single Dronpa-DNA-Atto647N complexe with a time resolution of $100\,\mathrm{ms}$ per frame. The Dronpa intensity is shown in green and the Atto647N intensity is shown in red (upper panel). The excitation sequence is shown in the lower panel.

$$PR = \frac{I_A}{I_D + I_A} \tag{4.1}$$

The distribution of PR values can be approximated by a Gaussian distribution. However, for PR values close to 0 or 1, a Beta distribution is more accurate [77] and is given by

$$P\left(x|a,b\right) = \frac{\Gamma\left(a+b\right)}{\Gamma\left(a\right)+\Gamma\left(b\right)}x^{a-1}\left(1-x\right)^{b-1}, \tag{4.2}$$

with mean μ and variance σ^2

$$\mu = \frac{a}{a+b} \qquad \text{and} \qquad \sigma^2 = \frac{a*b}{\left(a+b\right)^2\left(a+b+1\right)} \tag{4.3}$$

The PR was used to compare the constructs and test if three bases difference in donor-aceptor distance can be resolved by FRET measurements. The mean PR for each construct was obtained by fitting the histogram of PR values with a Beta distribution as shown in Figure 4.8. Only PR values obtained during illumination with $405\,\text{nm}$ + $488\,\text{nm}$, where the donor and acceptor intensities are constant, were used.

Differences in FRET values should be reflected in the average time before photoconversion. Higher FRET values mean higher rate for energy transfer between the excited state of the Dronpa molecule and the acceptor. The higher this rate is, the less time Dronpa spends in the excited state. Since photoconversion takes place from the excited state, the average time before photoconversion should be longer for the construct with higher FRET. This relation between FRET and photoconversion time was confirmed by comparing a construct with and without acceptor. The time for photo-conversion from the bright to the dark state upon excitation with $488\,\text{nm}$ light was determined by fitting the fluorescence decay which is averaged from single molecules as shown in Figure 4.9.

Thus, the presence of the acceptor not only acts as protection from photobleaching (see A.4) but also protects from photoconversion.
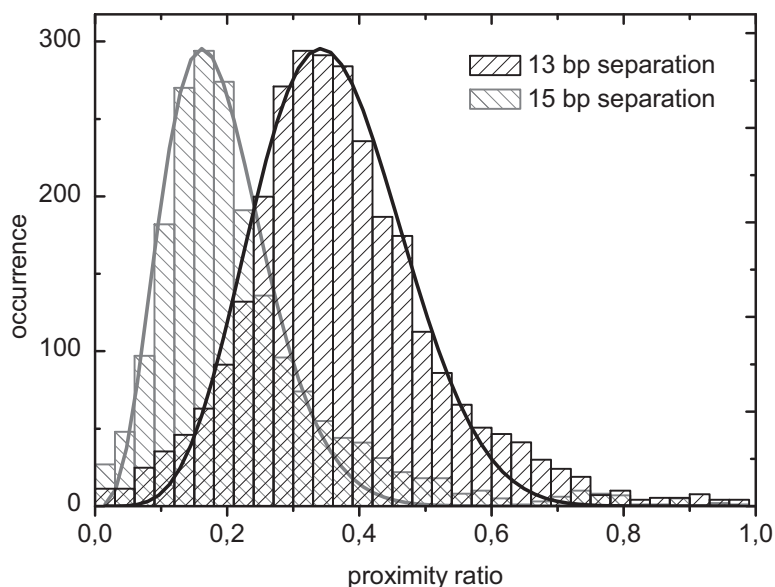
Figure 4.8.: Histogram of proximity ratios for individual switching events of single molecules for the two constructs shown in Fig 1A with 13 bp (black) and 15 bp (grey) separation between attachments points of Dronpa and Atto647N. The histograms were fit with a Beta distribution, yielding a mean proximity value and variance of $\mu = 0.36$, $\sigma^2 = 0.012$ and $\mu = 0.19$, $\sigma^2 = 0.006$ for the constructs with 13 bp and 15 bp separation between donor and acceptor respectively.
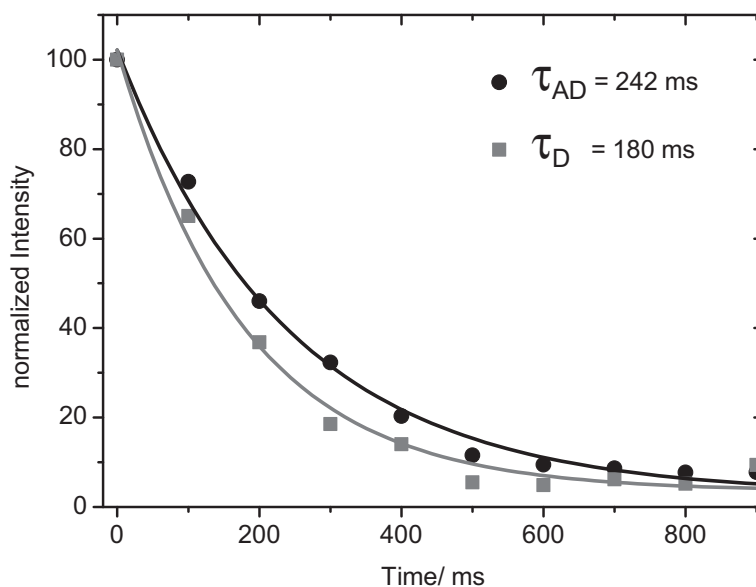
Figure 4.9.: The average fluorescence decay of Dronpa with 488 nm excitation switching to the dark state in the presence (black) and absence (grey) of an acceptor. In constructs with a photoactive acceptor, the donor-acceptor separation was 13 bp. Lines represent a fit to an exponential decay with lifetimes of 242 ms and 180 ms for constructs with and without a FRET acceptor respectively.

4.4.2. Live-Cell Imaging Measurements

With the development of genetically-encoded fluorescent-proteins e.g. GFP, live-cell imaging has become a versatile tool to investigate the functions of proteins in their native environment [78]. One of the biggest technical challenges for live-cell experiments is to keep the cells healthy while probing their internal functions. The probes used to gather information from inside the cells have to fulfill several requirements. Beside being small enough in size to prevent interference with the cells native function, they need to be addressable remotely without mechanical interactions. Optical controllability together with nano meter size of the construct used in the project described in the previous section, makes it predestined for use in the live-cell environment. In combination with lock-in amplification, the nanodevice can be used to actively enhance the SN ratio and to sense the pH value in the local environment.

Lock-in amplification is based on the principle of modulating the signal of interest with a specific carrier wave. To recover the signal of interest in the presence of a strong noise floor, the measured signal is multiplied with the carrier wave and intergated or filtered by a low-pass filter. Only those components of the measured signal which match the frequency and phase of the carrier wave contribute to the output. Noise which is not correlated with the carrier wave does not contribute to the output and thus the Signal to Noise ratio is improved. There are several technical implementations of this method, mostly used in electronic signal amplification. Recently, lock-in amplification was extended to optical signals and applied to amplify FRET and this enabled detection of FRET signals below 1% [79]. The method of using Lock-in amplification in combination with fluorescent probes was named *Optical Lock-In Detection* (OLID). In the experiment described in this section, lock-in amplification was used for signal enhancement and localization of spatially confined regions containing an optically switchable construct. To spatially confine and increase the local concentration of the signal generating label, $300\,\mathrm{nm}$ polystyrol beads were linked with the construct used in the previous experiments. To test the usability of this method in the live-cell environment, stably eGFP-tubulin tranfected HeLa cells were incubated with those beads. Imaging was performed on a spinning-disk

confocal microscope (Revolution System, Andor). The switching properties of Dronpa were used to modulate the fluorescence intensity with a carrier wave, in this case the frequency of the alternating illumination at $405\,\mathrm{nm}$. The cells were continuously illuminated at $488\,\mathrm{nm}$ and additionally, in two out of 10 frames, at $405\,\mathrm{nm}$. This additional excitation was synchronized to the detection to ensure reproducible modulation. The average fluorescence intensity image is shown in Figure 4.10. This image contains both channels, the green channel (left), with the Dronpa and eGFP fluorescence signal and the red channel (right), with the fluorescence signal originating from FRET.
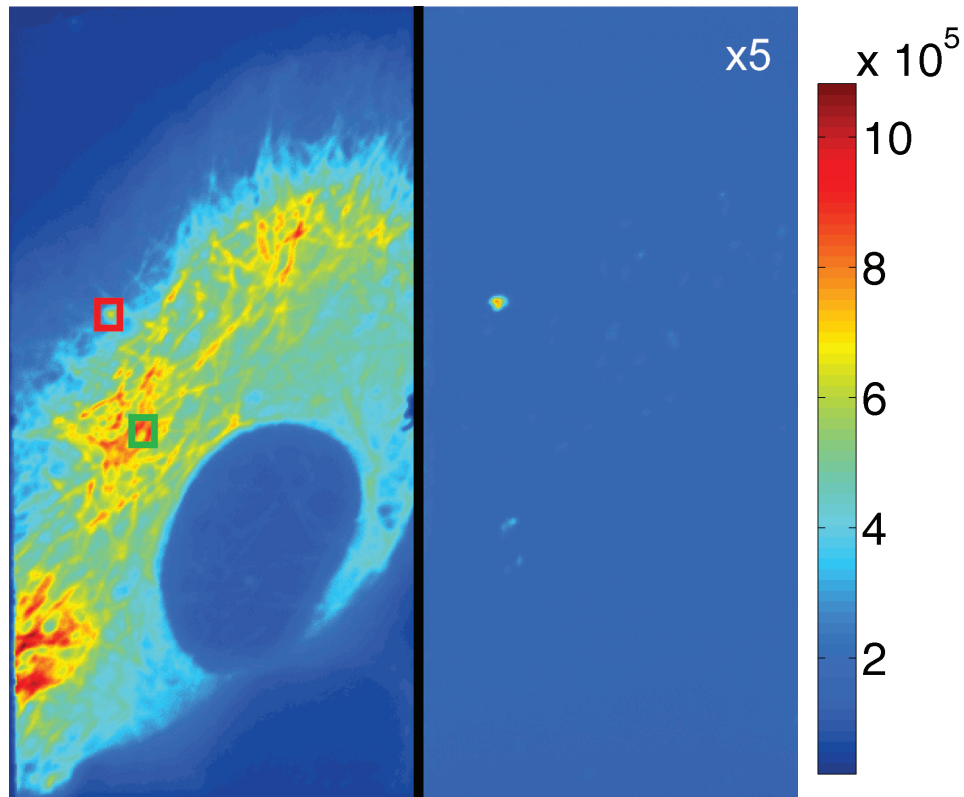
Figure 4.10.: Live-cell imaging experiments with a photoswitchable nanodevice. HeLa cells
stably transfected with eGFP-tubulin were incubated for 2 hours with $300\,\mathrm{nm}$ polystyrol
beads coated with streptavidin and the Dronpa-Atto647N-DNA constructs. The excitation
is alternated with 2 frames of 405 nm excitation followed by 8 frames of 488 nm excitation
and data was collected at a frame rate of 10 Hz or one switching-cycle per second. The
fluorescence intensity averaged over two cycles (without 405 nm excitation) is shown
with the eGFP and Dronpa fluorescence in the left channel and the FRET signal in the
right channel (multiplied by a factor of 5). The red and green squares indicate the regions
used for the FFT-Anylsis shown in Figures 4.12

Since the majority of the green fluorescence originates from the eGFP signal,
the Dronpa coated beads can not be located in the mean intensity image. In
the intensity image, the fluorescence signal is filtered based on emission spec-
tra. The left channel contains green fluorescence and the right channel the red
fluorescence. Since both Dronpa and eGFP have similar emission spectra, both
fluorophores are detected in the green channel and a filtering based on spec-
tral properties is difficult to achieve. In the red channel, fluorescence origi-
nating from the FRET acceptor was detected. Since the fluorescence from the

Atto647N, used as the acceptor in the construct, is the only source of red fluorescence, the bead is clearly visible in the red channel. In contrast to spectral filtering of the fluorescence signal, OLID separates the signals based on switching dynamics. One possible implementation of OLID is the use of a *Fast Fourier Transformation* (FFT) analysis to calculate the power spectrum of the fluorescence intensities at each pixel [24] and to apply a frequency filter to the output. The fluorescence intensities and power spectrum from the pixels from two selected regions (highlighted in Figure 4.10) are shown in Figure 4.12.

By modulating the signal with the carrier wave, the signal of interest is shifted to higher frequencies, away from the DC contribution of the constant eGFP fluorescence. A FFT analysis was performed at each pixel of the intensity image. A bandpass frequency filter summing up all power contributions at the switching frequency was applied to the power spectrum. The output of the frequency filter was used to compose a filtered image as shown in Figure 4.13.

From their physical characteristics, optically controlled nanodevices are ideal probes for investigating processes in living cells. The possibility of constructing a nanodevice which can reversibly switch FRET was demonstrated *in vitro*. In the next step, the functionality of this nanodevice was tested in the context of live-cell measurements. The photo-conversion properties were used to localize the nanodevice in the presence of other fluorescence labels in live-cells using the recently introduced method of optical lock-in detection (OLID) [79]. The principle of OLID is to switch the fluorescence intensity of a particular construct in a controlled fashion and then detect signals that change at the same frequency. Hence, the signal from the nanodevice is distinguished from that of cellular fluorescence through the switching behavior of the nanodevice rather than spectral separation. The method presented in this Section could be used to simultaneously track and probe individual constructs in live cells (e.g. uptake of virus particles). The time resolution needed for tracking could be achieved by the development of fast switching proteins and cameras with shorter exposure times. The methods described in this Section could be extended to combine different Fluorescent Proteins (FP) within FRET distance. Resonance energy transfer over multiple FP's could be realized. Using photo-switchable FP's, the direction of the energy transfer could be controlled. The nanodevice used in this
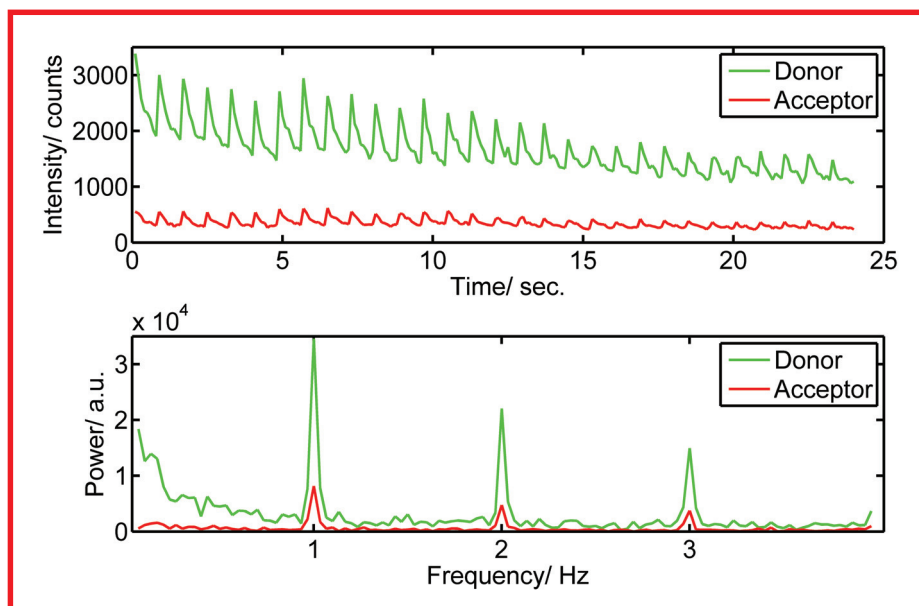
Figure 4.11.: Intensity and power spectrum of a region indicated by a red square in Figure 4.10.
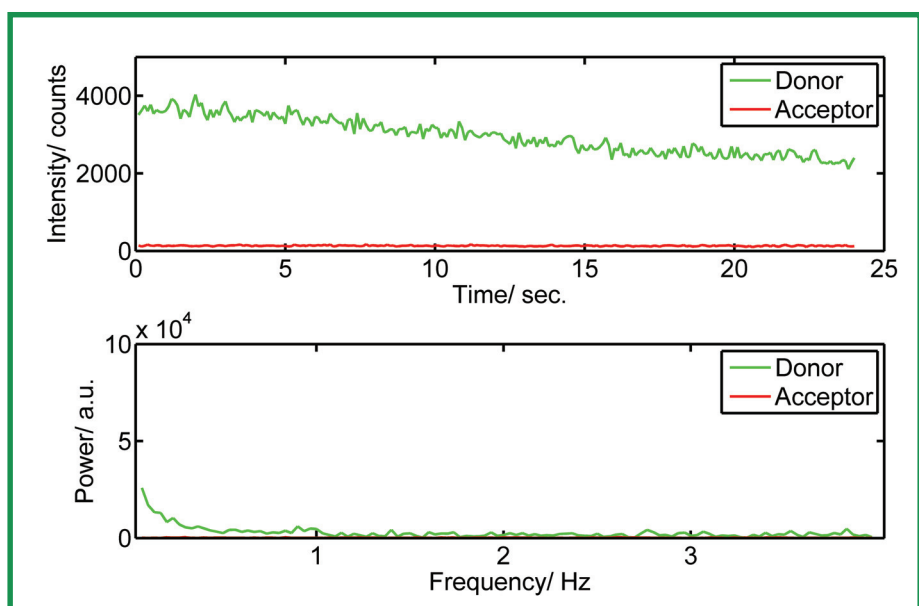

Figure 4.12.: Intensity and power spectrum of the two regions indicated by a red and green square in Figure 4.10. **Red**: The fluorescence intensity is strongly modulated by the $1\,\mathrm{Hz}$ carrier wave. **Green:** Slow photobleaching of fluorescence and no modulation of the fluorescence with the carrier wave.
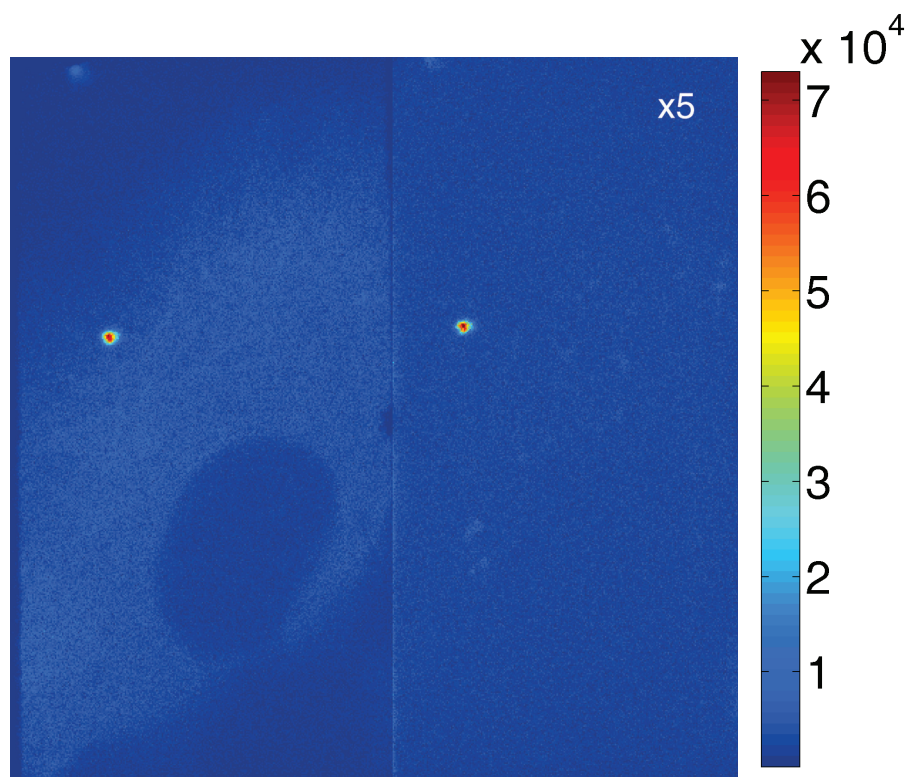
Figure 4.13.: Image of the frequency filtered FFT computed for the intensities at each pixel. The regions with strongly modulated fluorescence intensities are amplified. The Dronpa coated bead is clearly visible in this image while the constant eGFP fluorescence of the HeLa cell is filtered out.

project is basically a logical *And*-gate. Based on the combination of different photo-switchable gates, an optically addressable "Calculator" could be realized.

4.4.3. Dronpa as pH sensor

Fluorescent non-invasive indicators of pH are of great interest to measure the proton concentration in the cytosol of cells. Typical methods to optically sense pH are fluorescence intensity based and measure the pH dependence of the absolute fluorescence intensity [80] or the pH dependence of the fluorescence spectrum [81, 82, 83, 84, 85, 86, 87, 88]. Advanced methods have been developed which are based on the pH dependence of the fluorescence lifetime [89, 90]. Fluorescence lifetime based methods require extensive experimental equipment and intensity-based pH sensing methods have various other shortcomings such as uncertainty of the local label concentration due to lack of control in label expression, dilution or photo-bleaching and difficulties in discriminating between different sources of fluorescence (e.g. autofluorescence or presence of other fluorescent markers). The protonation state of Dronpa and in turn its absorption spectra strongly depend on the pH of the local aqueous environment (see Figure 4.14). The efficiency of photo-conversion between the two fluorescence states depends on the absorption and in term on the pH for a constant illumination intensity. Since Dronpa can be photo-converted repetitively, as shown in the previous section, the dynamic of the photo-conversion can be probed continuously. Based on the photo-conversion dynamic of Dronpa, a method for pH sensing in the live-cell environment was developed using a spinning-disc confocal microscope. The combination of standard time domain microscopy to obtain the raw intensity images with the advantages of the frequency domain analysis enables this Lock-in amplified pH sensing (LAPS) method to measure pH independently of the absolute fluorescence intensity, photo-bleaching or any static fluorescence, making this method well suited for investigating processes in a live cell context.

A Dronpa construct, based on the construct used in the experiment described in Section 4.4.2 but without an acceptor, was linked to polysterole beads as shown in Figure 4.15.

The beads were imaged with a spinning-disc confocal microscope[2] using constant $488\,\mathrm{nm}$ illumination and 2 out of 10 frames were illuminated with $405\,\mathrm{nm}$

---

[2]The $405\,\mathrm{nm}$ laser was not part of the commercial setup and was guided directly to the objective externally
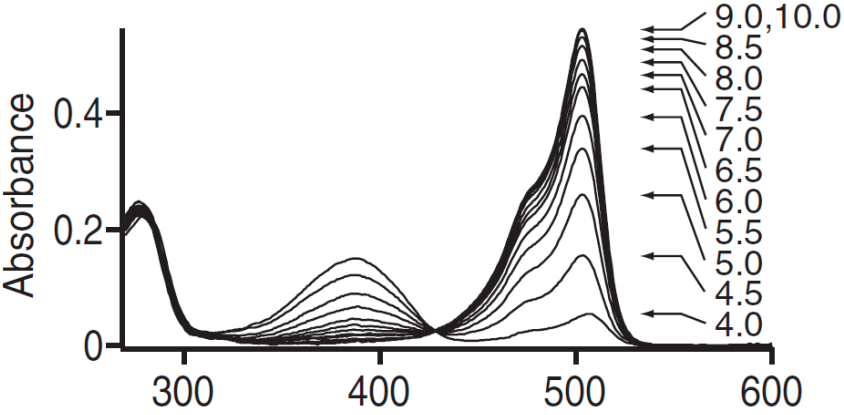
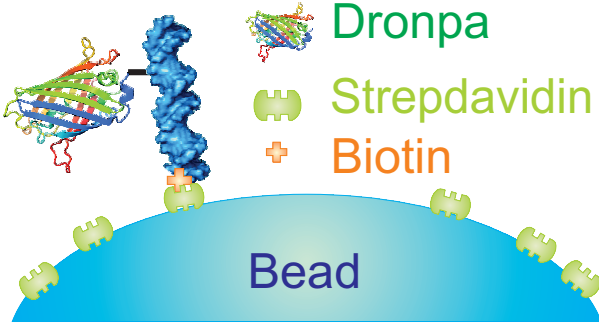**Figure 4.14.:** Absorption spectrum of Dronpa under different pH conditions (from [63]).

Figure 4.15.: A schematic of construct linked to the surface of beads.

and illumination after every 10 frames, as described in Section 4.4.2 (see Figure 4.16). The excitation cycle was repeated 30 times and a TIFF image stack was recorded.
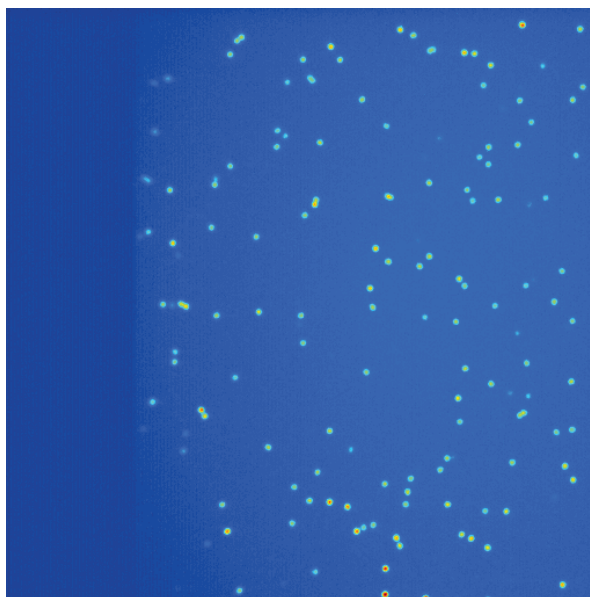
Figure 4.16.: Average intensity image of beads at the surface obtained by spinning-disc confocal microscopy.

The intensity information in each pixel was used to compute a FFT in time. To determine the strength of the modulation at each of these pixels, a frequency band-pass filter was used. The band-pass filter extracts the amplitude of the frequency component at the frequency used for modulation ($1\,\mathrm{Hz}$) and their higher harmonics. The amplitudes are obtained for each pixel and are used to compose the filtered OLID image shown in Figure 4.17.

The intensities of individual beads did stretch out over several pixels and the intensity information of those pixels were averaged to obtain the average intensity of each bead at every frame. Each individual *Regions Of Interest* (ROI) covers the pixels containing the fluorescence of individual beads. To automate this process, a threshold was applied to the OLID image and the individual ROIs

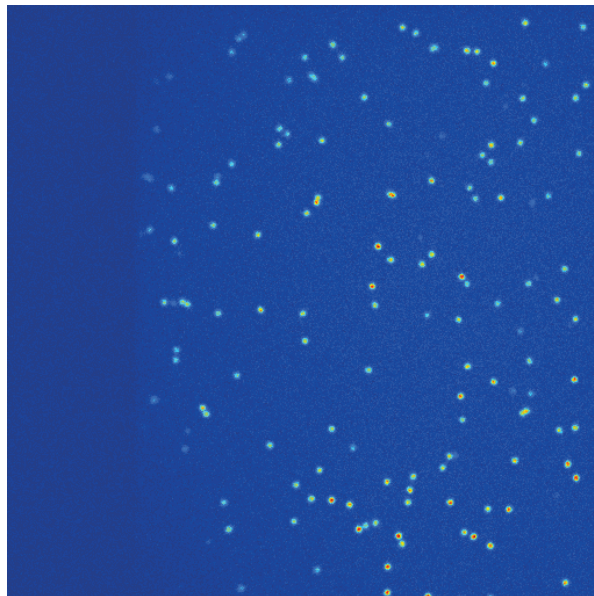**Figure 4.17.:** Frequency filtered intensity image. A FFT analysis was applied to each individual pixel and the result filtered with a band pass filter. The contrast is increased compared to Figure 4.16

were identified. The OLID image has higher contrast than the average intensity image and in turn is best suited to perform the isolation of pixels containing Dronpa fluorescence. The average modulation of the fluorescence intensities for a ROI is shown in Figure 4.18.
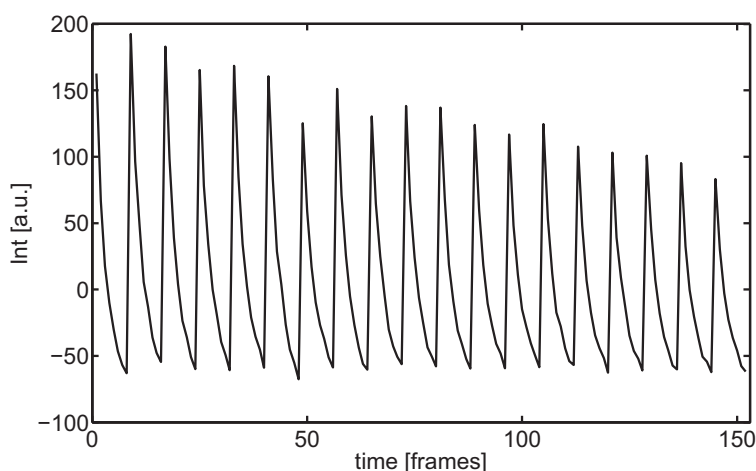

Figure 4.18.: Modulation of the fluorescence intensity averaged over the pixels of one ROI. The intensity was corrected for photobleaching by subtraction of the mean intensity of every switching cycle.

The average of the exponential decays for each switching cycle should stay constant and can be used to correct for photobleaching by setting it to zero. This can be done by subtracting the average intensity from every switching cycle. The two frames with additional $405\,\mathrm{nm}$ fluorescence were cut out and only the frames containing the decays were considered. This averaged intensity trace is fitted to a mono-exponential decay as shown in Figure 4.19.

Since the photo-conversion from the bright to the dark state occurs in the excited state of the bright-state (see Figure 4.1), the photo-conversion dynamic strongly depends on the intensity of the $488\,\mathrm{nm}$ illumination. Care needs to be taken to account for inhomogeneities in the illumination profile. To correct for an inhomogeneous illumination, the fluorescence intensity of a homogeneous fluorescent sample (Fluorescent slides, Chroma Technology) was imaged[3] (see Figure 4.20). This image was normalized to 1 by dividing every pixel by the

---

[3]The Andor Ixon+ EM-CCD camera has almost perfect linearity in the intensity range used
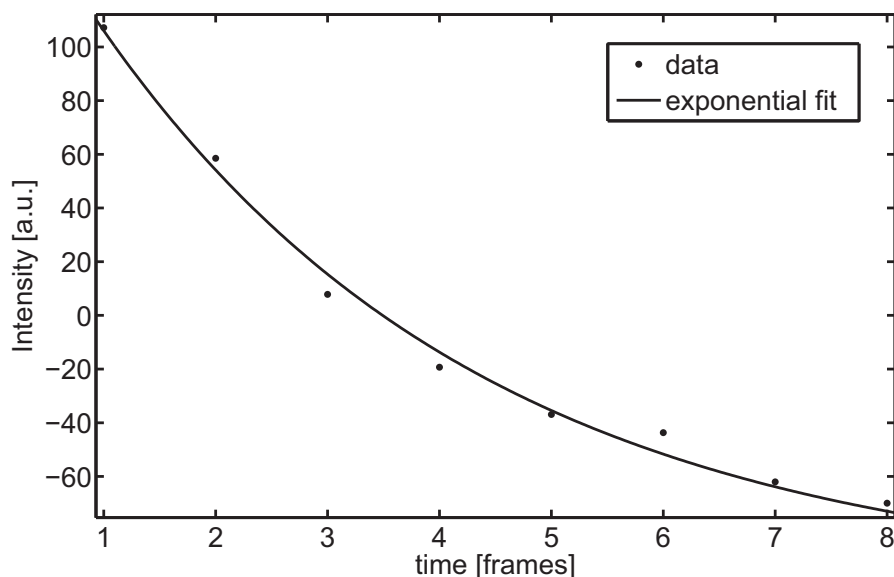
Figure 4.19.: Modulation of the fluorescence intensity for one ROI. The Intensity was corrected for photobleaching by subtraction of the mean intensity. The intensity is fitted to an mono-exponential decay.

maximum intensity. Using the normalized illumination profile, the switching times for different beads was corrected based on their locations on the surface.

The farther away a bead is from the center with maximum $488\,\mathrm{nm}$ intensity (see Figure 4.20), the lower the local illumination intensity is. The lower the illumination at $488\,\mathrm{nm}$, the longer the fluorescence on-time is. By multiplying the fluorescence on-time with the normalized intensity map, the influence of the illumination profile was accounted for. The effect of this normalization procedure on the fluorescence on-time is shown in Figure 4.21.

The analysis was repeated for all beads localized within the field of view. For each ROI, the mean fluorescence on-time was obtained. The decay time of every bead was used to compose an image where the fluorescence on-time is color-coded as shown in Figure 4.22.

In order to calibrate this fluorescence on-time with a pH value, measurements under different pH conditions were performed. The lifetime of the bright state, the average fluorescence on-time, decreases with increasing buffer pH. The life-
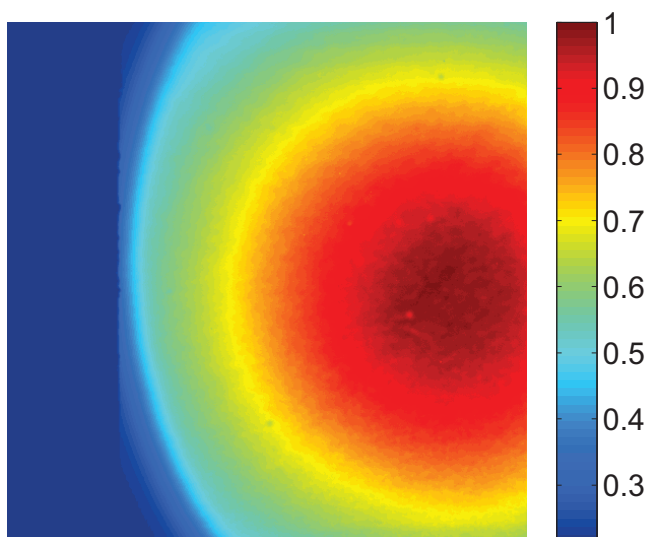
Figure 4.20.: Profile of fluorescence excitation normalized to 1
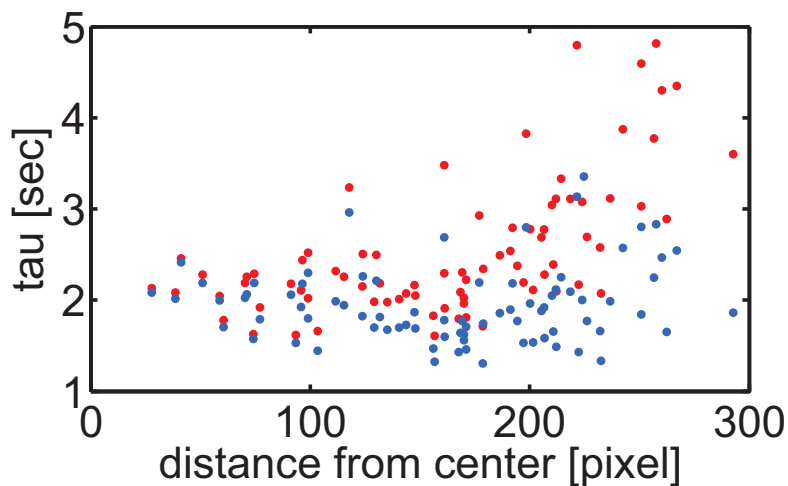
Figure 4.21.: Fluorescence on-time as function of distance from the center with maximum intensity. Red: without correction of the inhomogeneous illumination profile, Blue: with correction

time averaged over all ROIs as function of the buffer pH is shown in Figure 4.23. The standard deviation between different ROIs is shown as error bars in the plot.
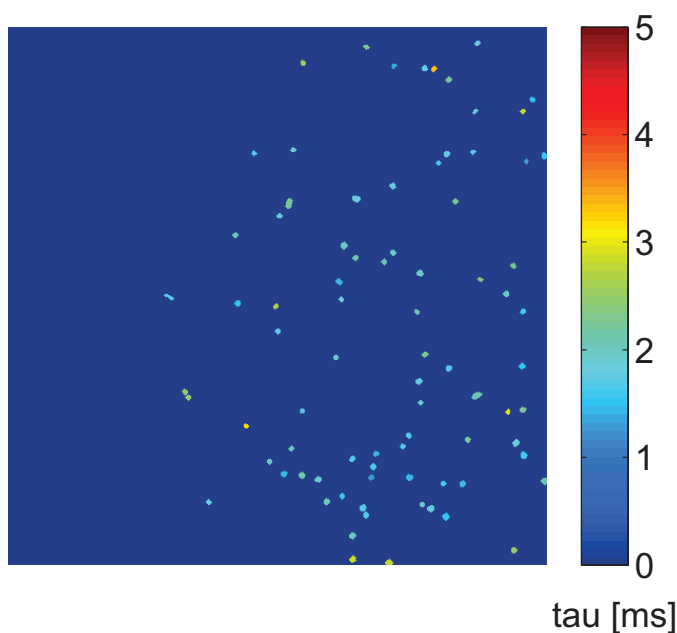
Figure 4.22.: Image of ROIs color-coded with fluorescence on-time corrected for the illumination profile.

The sensitivity of the pH measurement depends on the slope of the curve and is biggest for small pH values. The relation between the lifetime $\tau(pH)$ and the pH value was compared with the absorption spectrum of Dronpa. The lifetime $\tau(pH)$ is in good agreement with the change of the absorption spectra of Dronpa as function of the pH value (see Figure 4.14), as shown in Figure 4.23 as red triangles.

The slope of $\tau(pH)$ as a function of the switching life time decreases with increasing pH value. As a result, the accuracy of pH determination is higher for low pH values. For pH values above 6.5, the accuracy of the lifetime determination is not sufficient to give a upper bound of the pH values. The accuracy of the pH measurements is given in Table 4.2.
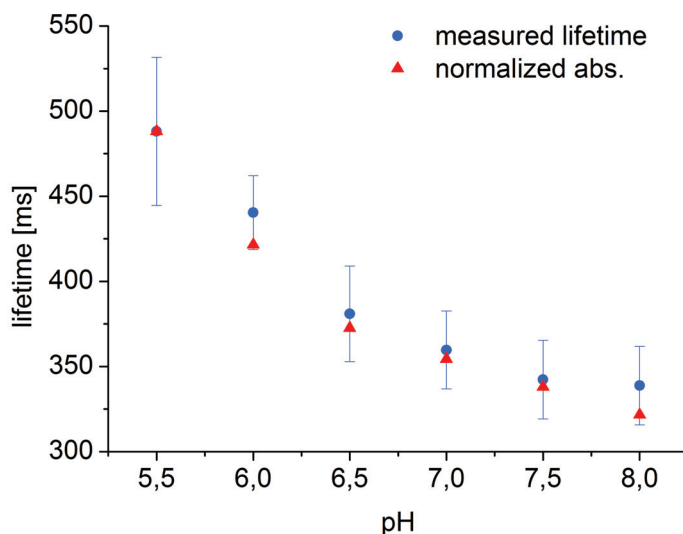
Figure 4.23.: Average fluorescence on-time as function of buffer pH. circles: on-times from fitting fluorescence decay; triangles: absorption at $488\,nm$ from 4.23, normalized to the measured lifetime at pH 5.5. Error bars are the standard deviation of ROIs.

Table 4.2.: Accuracy of pH measurements based on the switching time of Dronpa. The function $\tau(pH)$ was fitted with a second-order polynomial and used to solve the pH values given the measured life times τ +- the standard deviations of the lifetime measurements.

buffer pH	-	+
5.5	5.2	5.9
6.0	5.7	6.1
6.5	6.3	7.1
7.0	6.6	-
7.5	6.9	-
8.0	6.9	-

For each set of experiments, the $488\,nm$ excitation intensity was kept constant. The relationship between excitation intensity and fluorescence on-time is linear as reported previously [66]. To determine the slope, the experiment was repeated under various excitation conditions.
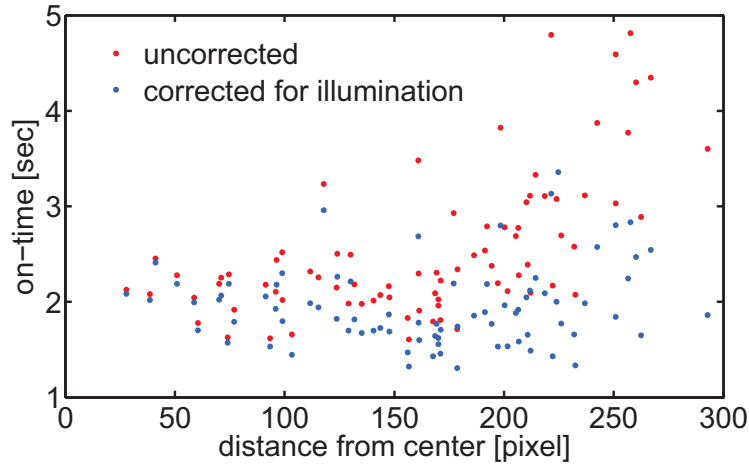
Figure 4.24.: Fluorescence on-time as function of illumination intensity at 488 nm. The slope obtained from the linear fit of -8.4 is in agreement with the slope of -7.6 calculated from the data published in [66]. For a theoretical derivation, see supplimental information in [66]

The application of the nanodevice, containing Dronpa as active element, for sensing the pH value in the live-cell environment, was demonstrated. The switching dynamics need to be calibrated for the excitation conditions to obtain absolute pH values. There are two factors which can be improved: The accuracy of pH determination and the time resolution. The accuracy of the pH determination depends on the quality of the switching-time measurement. The accuracy of the switching-time measurement can be improved by increasing the fluorescent signal obtained from the nanodevice. The fluorescence intensity of the nanodevice can be increased by increasing excitation intensity. Increasing the illumination at 488 nm will also make the switching time faster. One approach to improve both factors, time resolution and accuracy, is the use of fluorescent proteins with fast switching times and cameras with short exposure times.

5. Modulation of TBP binding to the TATA-box by Mot1

5.1. Introduction

Regulation of gene expression is essential for an organism as it allows a cell to adapt to changes in the environment [91] and it drives the process of cellular differentiation [92, 93, 94]. Although this regulation is essential, misguided regulation can be fatal by leading to uncontrolled cell growth and tumor formation [95]. This regulation can take place at various steps during protein biosynthesis starting from transcriptional control till translational regulation[96, 97]. For transcriptional regulation, TBP occupancy of the promoter site is of particular interest (see Section 5.1.3). The amount of TBP associated with the promoter can be up-regulated by active recruitment of TBP to the promoter site by interactions with other proteins. One of the proteins which is able to recruit TBP to the promoter site is the Spt3p subunit of the multi subunit SAGA (Spt-Ada-Gcn5-acetyltransferase) complex. The SAGA complex is required for transcription for some *RNA polymerase II* (PolII) dependent genes [98, 99]. Negative regulation of transcription can be accomplished by reducing the number of TBP molecules available for promoter binding through dimerization [100, 101], mobilizing of TBP at the core promoter site by the Negative Cofactor 2 (NC2) [102, 103] or by dissociation of bound TBP through the ATPase activity of *Modifier of transcription 1* (Mot1). In this work the role of the Mot1 on TBP binding to the TATA-box is investigated. In Section 5.2.1 and 5.2.4 a model of how Mot1 dissociates TBP is developed and in Section 5.2.2 a model for the role of Mot1 for gene activation is proposed.

5.1.1. Protein biosynthesis

Proteins are build from a chain of amino acids. The sequence of the amino acids is determined by the sequence of the DNA. The region of the DNA which contains the sequence for a particular protein is called *gene*. Twenty different amino acids can be encoded by the DNA. Amino acids consist of an amino group, a carboxylic acid group and a side-chain. Individual amino acids are polymerized in a linear chain by the formation on an amide linkage. This condensation reaction does not occur directly but instead individual amino acids are attached to specific transfer RNA (tRNA) through an ester bond. Each tRNA contains an *anticodon*, a base triplet used to match an amino acid to the sequence encoded in the mRNA. With the help of ribosomes those tRNAs are joined together to form the *primaty structure*, a polypeptide chain with a defined sequence of amino acids. The *secondary structure* is the three-dimensional folding of local segments and can be classified by geometric terms in α-helix or β-sheet. The *secondary structure* is formed by hydrogen bonds between the backbone of the chain. The three-dimensional structure of the protein is defined by the positions of the atoms and is called the *tertiary structure*. Proteins can form a compound of several individual subunits. This complex can contain several proteins of the same type or can be a mixture of different proteins. The spacial structure of this complex is called the *quaternary structure*. Despite the complexity of the interactions involved in formation of the *quaternary structure*, structure and function of proteins are determined by the sequence of the peptide chain. Great effort has been taken to optimize computational methods to predict the final structure given only the sequence information of the amino acids. However, due to the enormous degrees of freedom in the folding process, predicting the *tertiary structure* or *quaternary structure* is challenging. The computational methods available today are at best of limited accuracy [104]. Determination of protein structures depend largely on well established methods such as X-ray crystallography, NMR-spectroscopy and increasingly on cryo-electron microscopy [105]. Recently, single-molecule methods were used to infer structural information from FRET data. Pairs of positions within a protein were labeled with donor and acceptor fluorophores. The FRET values recorded were used to determine the relative positions of the labels and to monitor conformational changes [106, 107].

5.1.2. The TATA-box

The so called TATA-box is composed of a locally confined and evolutionary highly conserved A-T-rich DNA sequence. It is located about 30 bp upstream of about 10% of genes in humans [108]. Initially, it was called the "Goldberg-Hogness" box, who first observed this sequence [109]. The importance of the TATA-box is in its role as primary binding site of the TATA-box binding protein (TBP), which itself is the platform for assembly of the multi-subunit Preinitiation Complex (PIC) [110, 111]. However, only about 10% of mammalian genes are TATA-box associated but all require TBP binding for PIC formation [112, 113]. The TATA-box belongs to the group of core promoters such as BRE, MTE, and XCPE1. The core promoters are defined as sequences located in the proximity of the gene *Transcription Start Siteable* (TSS) and which are able to initiate transcription [114]. The TATA-box and subsequently TBP binding nucleates the PIC formation and assembly in all three polymerases (POLI, II and III) in archaeal and eukaryotic organsisms. Binding of TBP to the TATA-box is the first step in the transcription of ribosomal-, messanger- and transport-RNA. It was reported from *in vivo* experiments that for some genes the orientation of the transcription is not determined by the orientation of the TATA-box and it was suggested that interaction of TBP and the PIC with other TFs ultimatiely determines the transcirptional orientation [115, 116]. Modulation of the interaction of TBP with promoter sites is therefore an efficient strategy for regulating gene expression *in vivo*.

5.1.3. The TATA-box binding Protein

The *TATA-box binding protein* (TBP) is a 27kD evolutionary highly conserved monomeric protein which consists of two symmetrical subdomains, each with five beta-strands and two alpha-helices [117, 118, 119, 120]. It is speculated that TBP has evolved from a homodimer by gene duplication and fusion [121]. TBP binds the TATA-box with a modest specificity of $10^3 - 10^5$ fold preference for binding the TATA-box over a random DNA sequence [122, 100, 123, 124, 125] with an high affinity of about 10 nM [126, 127, 128]. Binding of TBP to pro-

moters that do not contain a TATA sequence require additional TBP-associated factors (TAFs) [129]. Binding of TBP to DNA observed *in vitro* is highly stable [130] and induces an approximately 90° bend of the DNA towards the major groove as shown in Figure 5.1.

A   B
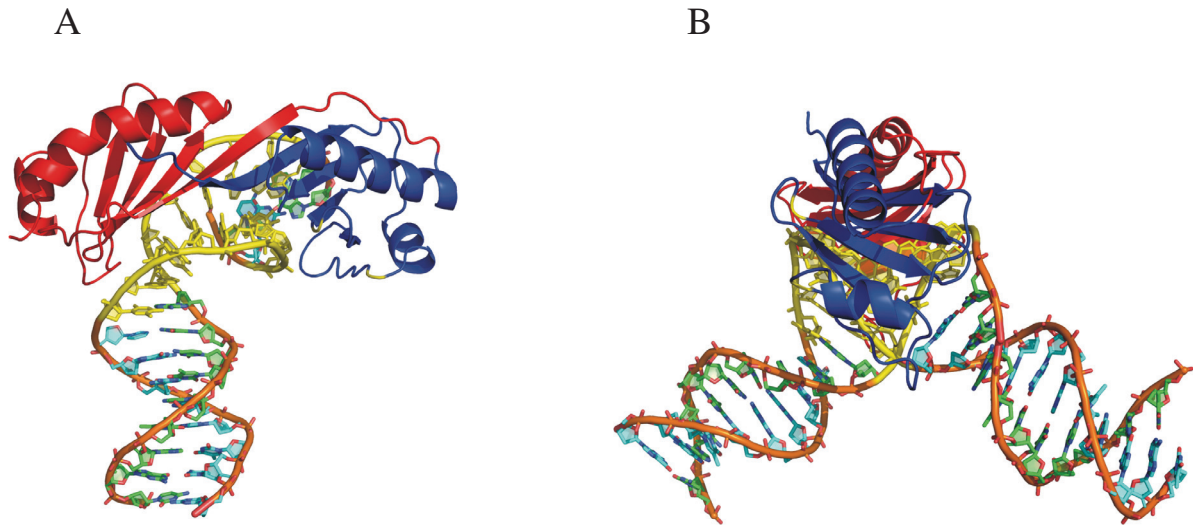
Figure 5.1.: Crystal Structure of TBP bound to the TATA-box (based on PDB structure 1RM1). N- and C-Terminal half of TBP are color coded in red and blue, respectively. The TATA-box is color coded yellow. A: front view, B: side view

It was reported that the DNA bending plays a role in nucleosome sliding [131]. In contrast to *in vitro* studies, *in vivo* experiments revealed a highly dynamic pool of TBP molecules [132, 133]. TBP binds the TATA-box through its curved eight-stranded antiparallel beta sheet which recognizes the minor groove of the TATA-box via an induced-fit mechanism. Upon binding of TBP to DNA, TBP inserts two pairs of phenylalanins into the first T-A and between the last 2 bp of the TATA-box. Additionally, formation of hydrogen bonds between two asparagines (N69 and N159) and the center of the TATA-box also contributes to the deformation of the DNA [134, 111]. The binding affinity of TBP to the TATA-box is lowered by interaction of the N-terminal of TBP with the TATA-box containing DNA. It was reported that removal of the N-terminal domain increases the binding affinity of TBP to the TATA-box [135]. The first crystal structures suggested uni-directional binding of the highly symmetrical TBP to

the TATA-box [122, 134, 136, 137]. However, in the following years it became clear that TBP binds the TATA-box in both directions [138] and almost without preferred orientation [116, 139, 121]. The preference for the proper binding orientation is enhanced by TFIIA or TFIIB to approximately 80% [140, 141]. It has been suggested that binding and bending of the TATA-box by TBP takes place in a single step [128]. Based on FRET ensemble experiments a three step model of TATA-box binding of TBP was proposed with the formation of a stable final binary complex as seen in the crystal structure [142, 128, 143, 144]. This model was supported by the findings of Tolic-Nørrelykke who observed that TBP induces a three step dynamic in the bending of the TATA-box containing DNA [145]. TBPs function as PIC nucleation site is essential for transcription by all three RNA polymerases [146, 147]. It was shown that TBP occupany at the core promoter site is correlated with transcriptional activity [148]. Since TBP binding to the TATA-box is the initial and rate limiting step of PIC formation and in turn gene transcription, modulation of the interaction of TBP and DNA is a prime target for gene regulation [149, 150]. This concept is supported by the finding that the TBP occupancy of the TATA-box correlates with the amount of transcription product of that gene [148]. Since only a minority of human genes are associated with the TATA-box as promoter, TBP binding to non TATA sequences is required [151, 125, 113]. The influence of TBP dimerization on the binding of TBP to DNA and in turn gene expression is still under discussion [152, 118].

5.1.4. The Mot1 protein

Mot1 (Modulator of transcription 1) is a $210 \, \text{kDa}$ protein encoded in the BUR3-gene[1] and a member of the Snf2/Swi2-family of ATPases [154, 155, 156]. Despite sharing conserved helicase domains Mot1 does not track along DNA [157] and has no helicase activity [158, 159, 160]. Mot1 binds to TBP and to the binary TBP/DNA complex but does not bind to DNA alone [161, 160].

Mot1 requires a 17bp upstream DNA handle to dissociate TBP from DNA [162]. The HEAT repeats contained in the N-terminal of Mot1 are required to medi-

---

[1]the human homolog is BTAF1 [153]
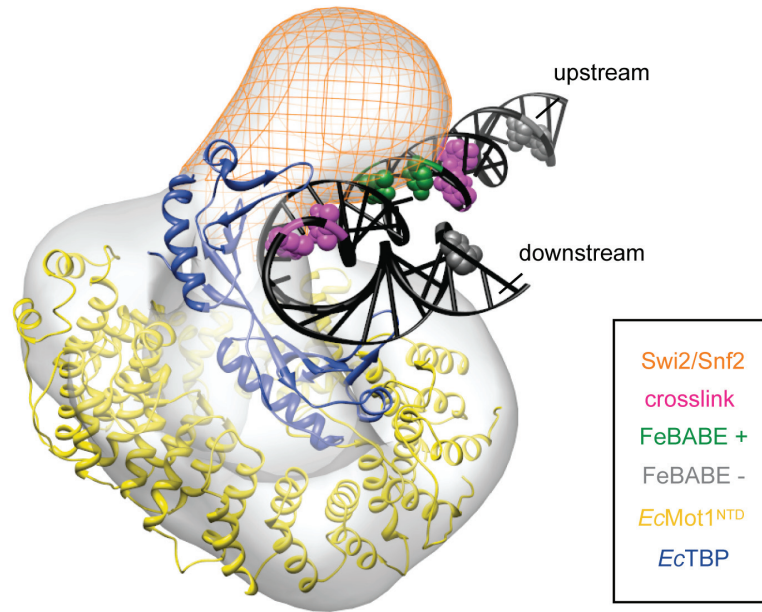
Figure 5.2.: Crystal Structure of Mot1 bound to the TBP (kindly provided by Petra Wollmann). TBP is shown in blue and Mot1 in yellow.

ate association with TBP [163] and potentially serves as a scaffold for protein-protein interactions [164]. The C-terminus contains the ATPase domain essential for TBP dissociation. Mot1's function as a repressor and activator is essential for regulation of transcription by Polymerase II *in vivo* [161]. Gene repression can be explained by Mot1's ability *in vitro* to remove TBP from DNA in an ATP dependent manner [161]. Several models have been discussed regarding how Mot1 utilizes the energy provided by ATP to dissociate TBP from DNA including induced conformational changes within the structure of TBP [165], conformational changes induced in the DNA [165], mechanical displacement of TBP through translocation along the DNA [158] and ATP driven conformational change of Mot1 which modulates TBP/DNA binding [166]. However, how Mot1's ATPase activity contributes to gene activation is not fully understood. Several models for the role of Mot1 in gene activation have been discussed, in-

cluding recycling and redistribution of TBP from non-active transcription start sites [167, 157], modulation of TBP binding affinity to DNA [161, 165, 168] or interactions with nucleosome remodeling [169, 170].

5.2. Interaction of Mot1 with the TBP/DNA binary complex

To investigate whether the presence of Mot1 modulates the conformation of the binary TBP/DNA complex, two different approaches were used. One approach is to monitor the changes in conformation between TBP and DNA. In this case a specifically donor labeled TBP mutant together with an acceptor labeled ds-DNA containing the TATA-box serves as a FRET pair (see Section 5.2.1). With this assay, the relative position of the label in the N-terminal part of TBP with respect to the downstream part of the TATA-box can be observed. Since TBP and DNA are fluorescently labeled the formation of the binary complex can be monitored directly by FRET. The second approach is to monitor the changes in the DNA conformation by FRET between a FRET pair attached to the dsDNA (see Section 5.2.2).

5.2.1. Conformational changes of the TBP/DNA complex

TBP binds tightly to the minor groove of the TATA-box and interacts with the DNA several base pairs up-and downstream from it [158]. It was reported that Mot1 binds to the upstream part of the TATA-box and that Mot1 binding to this binary complex does change the interaction of TBP with the TATA-box flanking regions [158]. It was shown that Mot1 dissociates TBP from DNA in an ATP-dependent manner [171]. In a recently published crystal structure of the binary Mot1/TBP complex, it was shown that Mot1 binds TBP directly in the saddle shaped TBP/TATA-box binding surface via a loop like structure [166]. In this

section, the influence of Mot1 binding on the conformation of the TBP/DNA complex is investigated.

Experiments and Results

The assay described in this Section is used to monitor the influences of Mot1 on the conformation of the TBP/DNA complex. FRET between the labeled C61 mutant of TBP and labeled DNA containing the TATA-box from the H2B promoter was recorded as illustrated in Figure 5.3A. After preincubating TBP with DNA, Mot1 was added and allowed to bind to the binary TBP/DNA complex. Individual ternary complexes were surface immobilized. To remove any excess of unbound complexes and free TBP, Mot1 or DNA the flow chamber was flushed with working buffer prior data acquisition. In contrast to ensemble measurements where unbound Mot1 or TBP remains in solution, here only the immobilized binary/ternary complexes are available for the reaction.

Molecules with two types of FRET behavior could be observed: molecules that displayed either static or dynamic FRET during the time window of observation. The majority ($\tilde{6}0\%$) of DNA/TBP/Mot1 complexes before ATP addition display static FRET[2]. Interestingly, the fraction of static complexes increased after ATP addition to over 90% at the cost of a decreasing amount of dynamic complexes. The absolute number of complexes did not significantly change after ATP addition.

A sample trace for static FRET is shown in Figure 5.3 B. The transitions observed from complexes showing dynamic FRET includes fast stepwise transitions and gradual changes in FRET (see Supplemental Figures in Appendix A.3 for sample trace). Due to this heterogeneity in FRET dynamics, only complexes with static FRET were considered for analysis and for composing the histograms in Figure 5.3.

---

[2]The number of molecules for each type was estimated by averaging the number of molecules found within the
field of view for different recordings.

The molecule-wise FRET histogram of the preincubation binary TBP/DNA complexes, compared with the FRET histogram of the ternary complex, before and after $1\,\text{mM}$ ATP addition is shown in Figure 5.3C.

The parameters from fitting the FRET histogram shown in Figure 5.3 with two or three Gaussian is given in Table 5.1. Upon preincubation with Mot1, a third high FRET conformation appears and is accounted for by extending the double-Gaussian fit model used to fit the TBP/DNA and TBP/DNA/Mot1 histograms to a triple-Gaussian distribution.

Table 5.1.: Parameters of the normal-distributions obtained by fitting the molecule-wise FRET histogram shown in Figure 5.3. μ and σ values are given in % of FRET. The relative amplitude of the Gaussian distribution A is given in %.

	S_1			S_2			S_3		
	A	μ	σ	A	μ	σ	A	μ	σ
TBP/DNA	85	25	16	15	34	34	-	-	-
TBP/DNA/Mot1	15	25	15	13	41	32	72	78	21
+ATP	75	24	18	25	32	36	-	-	-

For the binary TBP/DNA complex, a low FRET efficiency of about 25% with a shoulder at 35% is observed. If the complexes are preincubated with Mot1, a drastic increase in FRET between the label on the DNA upstream of the TATA-box and TBP is observed. This induced increase of FRET is reversed by ATP addition as shown in Figure 5.3C. To test if ATP binding or ATP hydrolysis restores the low FRET population, the same experiment was performed with ATPγS. Addition of $1\,\text{mM}$ ATPγS almost completely restored the low-FRET population. The minor fraction of molecules in the high-FRET population (not shown in the fit for table 5.2) disappeared after additional sequential addition of $1\,\text{mM}$ ATP. The values from fitting the FRET histograms with a double-Gaussian distribution is given in Table 5.2.

Since only TBP and the DNA are fluorescently labeled in these experiments, Mot1 binding cannot be observed directly. The appearance of the high FRET population after preincubation with Mot1 strongly suggests that the molecules with high FRET values are the ones which have Mot1 bound. However, the
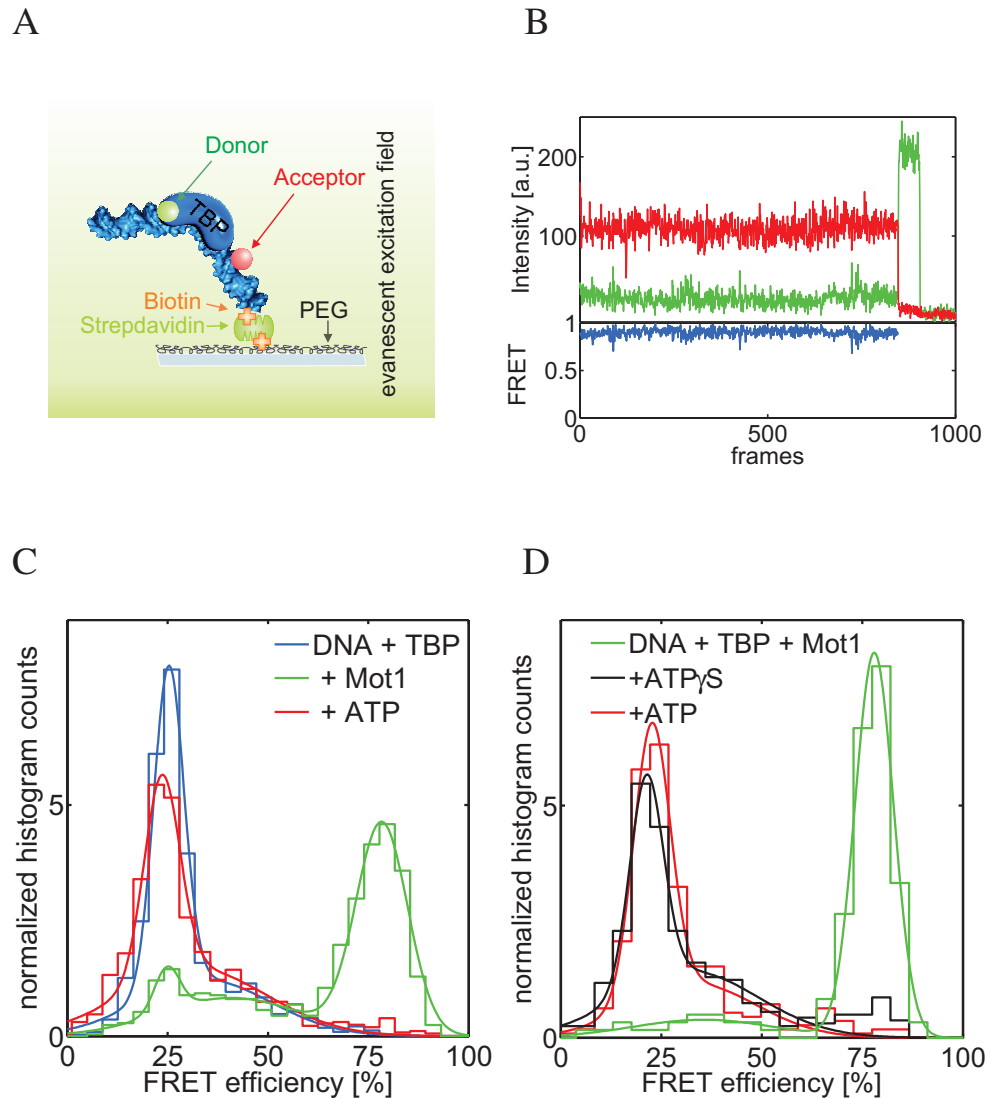
Figure 5.3.: Conformational changes of the TBP/DNA complex induced by preincubation with Mot1. Initial FRET distribution was restored after sequential addition of ATP/ATPγS. **A:** Scheme of surface immobilization of sample. The TBP-containing complexes are bound to a quartz surface, passivated by PEGylation, via a biotin-streptavidin bond. **B:** Single-molecule trace of a static DNA/TBP/Mot1 complex. Upper panel: The donor and acceptor intensities are shown in green and red, respectively. The donor intensity increases after single step acceptor photo-bleaching. Lower panel: The calculated FRET efficiency is shown in blue. **C:** Molecule-wise FRET histogram of complexes on the H2B promoter: the TBP/DNA binary complex (blue), of the preincubated TBP/DNA/Mot1 ternary complex before ATP addition (green) and after addition of 1mM ATP to the ternary complexes (red). The ATP addition was performed in the same flow-chamber under the same buffer conditions as the measurements of the ternary TBP/DNA/Mot1 complex. **D:** Molecule-wise FRET histogram of complexes on the H2B promoter: the ternary preincubated TBP/DNA/Mot1 complex (green), after addition of 1mM ATPγS (black) and after sequential addition of 1mM ATP (red).

Table 5.2.: Parameters of the normal-distributions obtained by fitting the molecule-wise FRET histogram shown in Figure 5.3. μ and σ values are given in % of FRET. The relative amplitude of the Gaussian distribution A is given in %

	S_1			S_2		
	A	μ	σ	A	μ	σ
TBP/DNA/Mot1	4	35	35	96	78	18
+ATPγS	76	21	17	24	32	35
+ATP	83	23	18	17	33	33

molecules that remain at a low FRET efficiency before ATP addition or return to a low FRET efficiency after ATP addition are not necessarily molecules without Mot1 bound. To determine whether bound Mot1 is associated with a high FRET state, the following three-color experiment was performed. For this experiment, in addition to labeled TBP and DNA, Mot1 was labeled with Alexa488. Since the presence of Mot1 was determined by the fluorescence intensity and not FRET, is was sufficient to label Mot1 nonspecifically. The concentration of labeled Mot1 used for these single-molecule experiments was lower than the concentration of unlabeled Mot1 in the previous Section. The concentration of labeled proteins which can be added to the flow chamber is limited because any fluorescent label present in the solution contributes to the background fluorescence despite the exponential decay of the evanescent excitation field. Additionally, the amount of unspecific binding of labeled proteins to the surface increases with increased protein concentration. The fluorescence was recorded using the same two-color TIRF setup described in Chapter 2 but with an altered excitation scheme. The laser excitation switched every frame between $488\,\text{nm}$ and $532\,\text{nm}$. In the first frame, the fluorescence during $488\,\text{nm}$ excitation from the Alexa488 dyes was recorded in the green channel. In the next frame the excitation is switched to $532\,\text{nm}$ and the fluorescence from the Atto532 dye is recorded in the same green channel. The fluorescence of the Atto647N dye due to FRET from the Atto532 dye is recorded in the red channel. This excitation scheme is then repeated for the whole acquisition. The separation of the recorded fluorescence intensities is based on the excitation scheme and is performed after acquisition in a MATLAB program (see Appendix A.5).

Complexes containing Mot1 were identified by the presence of fluorescence in the green channel after $488\,\mathrm{nm}$ excitation. Direct excitation of the Atto532 dye by $488\,\mathrm{nm}$ excitation was corrected for by subtracting a fraction of the Atto532 fluorescence after $532\,\mathrm{nm}$ excitation from the fluorescence recorded from illumination at $488\,\mathrm{nm}$. The histogram of the FRET values between TBP and DNA is shown in Figure 5.4.
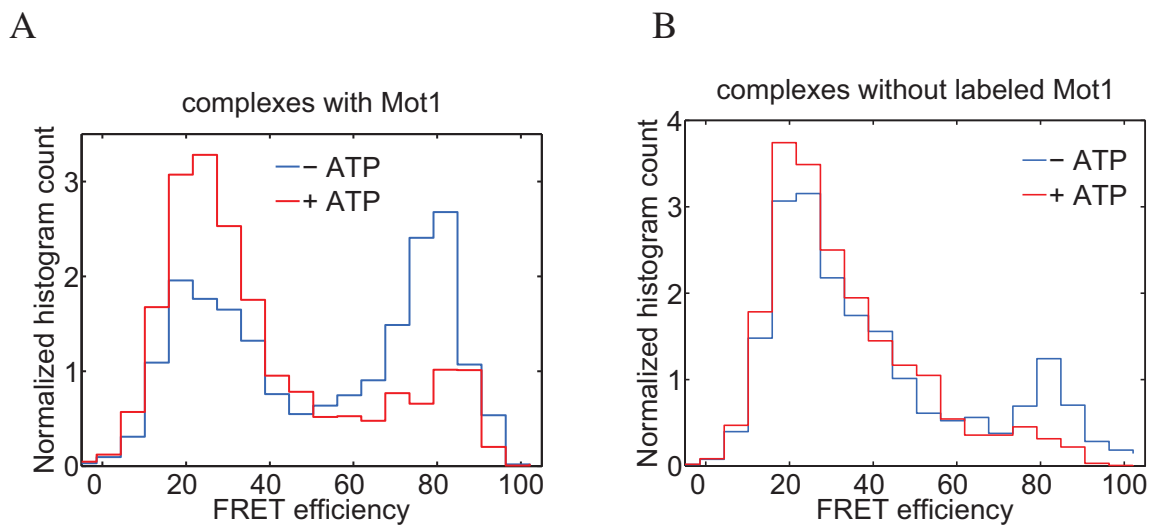

Figure 5.4.: Three-color experiment with Alexa488 labeled Mot1, and Atto532/Atto647N labeled TBP and DNA. A: Normalized histogram counts of frame-wise FRET between DNA and TBP of complexes containing Alexa488 labeled Mot1. Blue: before addition of ATP. Red: after addition of 1mM ATP. B: Normalized histogram counts of frame-wise FRET between DNA and TBP of complexes without detection of Alexa488 labeled Mot1. Blue: before addition of ATP. Red: after addition of 1mM ATP

There are two reasons for complexes not to fluoresce after $488\,\mathrm{nm}$ excitation. Either they don't have Mot1 bound or the Mot1 bound is not labeled due to limited labeling efficiency. In contrast to Figure 5.3B & C, the concentration of labeled Mot1 used in this experiment is not sufficient to form primarily ternary complex. As a result, a fraction of the binary TBP/DNA complexes are without bound Mot1 and a small fraction of ternary complexes with unlabeled Mot1 were present in this experiment.

The complexes without a detectable fluorescence signal from the labeled Mot1 mainly populate the low FRET state as shown in Figure 5.4B. As expected, the

minor fraction of molecules in the high FRET state disappears upon addition of $1\,\mathrm{mM}$ ATP. This fraction of molecules in the high FRET state is likely complexes containing unlabeled Mot1 due to the limited labeling efficiency.

Complexes containing labeled Mot1 show two subpopulations in the FRET distribution as shown in 5.4A. After ATP addition, the relative amount of molecules in the high FRET conformation is reduced, as expected from the two color experiments. However, after ATP addition molecules containing Mot1 with a low FRET efficiency are still observed.

To rule out that influences of the surface alter the interactions of the TBP with Mot1 and DNA, single-molecule solution experiments were performed with a *Multiparameter Fluorescence Detection* (MFD) setup. Samples were prepared as for the TIRF experiments and measured for sequential 2 hours periods. The same increase in FRET efficiency upon preincubation with Mot1 was observed as shown in Figure 5.5. Upon ATP/ATPγS addition, the higher FRET components vanished.

Discussion

Preincubated surface-immobilized binary TBP/DNA and ternary TBP/DNA/Mot1 complexes are highly stable and can be observed for hours. In contrast to ensemble measurements, where unbound DNA, TBP or Mot1 are present in solution, in this assay the function of individual isolated complexes is observed without interaction with additional free proteins. The FRET values obtained here are sensitive to changes in distance between the N-terminal half (position C61) of TBP with DNA downstream of the TATA-box which makes this assay suitable to monitor the interaction between TBP and DNA. The majority of molecules show static FRET values. This indicates that, in the majority of the complexes, TBP binds stably and tightly to the TATA-box. The heterogeneity of the dynamic complexes suggests a diverse spectrum of interactions between the TBP/DNA molecules. Experiments that are sensitive to the DNA conformation alone (see Section 5.2.2) lack such heterogeneities, suggesting that TBP/DNA interactions and not conformational changes of the DNA are the cause. The

histogram of static FRET values of the binary TBP/DNA complex can be fit-
ted with a distribution of two Gaussians. An explanation for the shoulder of
the FRET distribution (State S_2) in Figure 5.3 is that both TBP binding orien-
tations have slightly shifted FRET values as suggested previously [103] or a
conformation where TBP binds at a slightly shifted position with respect to the
TATA-box. The molecules with this medium-FRET state S_2 ($30 - 40\%$) are
significantly less affected by Mot1 or ATP addition than the molecules in the
high- or low-FRET state S_1, S_3. This suggests that the Mot1 induced conforma-
tional change of TBP mainly affects TBP bound in a certain state (in State S_1 or
S_3). Although FRET was measured between TBP and DNA, possible explana-
tions for the increase in the FRET value should not be limited to protein/DNA
interactions. Possible explanations include that Mot1 binding rearranges TBP
with respect to the TATA-box or induces a conformational change within TBP.
Another possibility is that Mot1 binding to the upstream part of the TATA-box
induces a conformation within the DNA which propagates through the TATA-
box and alters the TBP/DNA interaction. Rearrangement of TBP with respect
to the DNA imply that the TBP/DNA-interaction is altered and TBP would bind
to a shifted TATA-box, which can change the bending of the DNA [172]. The
influence of Mot1 on the DNA conformation is investigated in experiments de-
scribed in 5.2.2.

In contrast to ensemble measurements where Mot1 ATPase activity dissociates
TBP from DNA, in this assay without free Mot1 in solution, no dissociation of
TBP was observed but rather a significant decrease in FRET efficiency. Both,
ATP or ATPγS addition did decrease the FRET efficiency. This clearly shows
a functional ATPase activity of Mot1 upon ATP binding which does not cause
TBP displacement. The question that arises is what effect Mot1's ATPase ac-
tivity has on the TBP/DNA complex. The FRET values measured between
two labeling positions, (position C61 of TBP and the DNA downstream of the
TATA-box) before Mot1 and after ATP addition are very similar. However, this
does not imply that those complexes overall have the same structure. Although
the three-color experiments described here indicate that molecules adopting the
low-FRET conformation can have Mot1 bound and that ATP addition does not
necessarily dissociate Mot1 from the binary complex. To illuminate the effects
of Mot1's ATPase activity on bound TBP, the conformations of the DNA were

investigated by measuring the FRET efficiency of a double-labeled DNA, as described in the next section.

5.2.2. Modulation of the TATA-box conformation

Binding of TBP to the TATA-box induces a $\tilde{9}0°$ bent of the DNA as reported by X-ray crystallography and *in vitro* studies [122, 134, 136, 137, 143, 145] as shown in Figure 5.1. It has been shown that the transcription activity correlates with the bending angle [172]. A dsDNA which contains two fluorescence labels separated by 18 base pairs was used as a FRET sensor. The FRET labels are located up- and downstream of the TATA-box (see Appendix A.3 for sequence) as illustrated in Figure 5.6.

The fluorescent labels were placed in the flanking regions of the TATA-box in such a way that they were sensitive to bending within the TATA-box. The TBP used in these experiments was not fluorescently labeled and hence the presence or binding state of TBP to the DNA could not be determined directly. In contrast to the experiments described in section 5.2.1 where only complexes containing TBP are visible and show FRET, the experiments described here do not contain this "intrinsic" selection for complexes containing TBP. As a result, every immobilized DNA molecule, also molecules without bound TBP, are analyzed.

Experiments and Results

The complex were allowed to be formed prior surface immobilization during a preincubation step. After immobilization of the molecules to the surface of the flow chamber via a streptavidin-biotin bond, any unbound molecules were removed by flushing the flow chamber with working buffer. The majority of complexes (90%) analyzed showed static FRET values which are indistinguishable from DNA alone as shown in Figure 5.7.

The dynamic of the remaining non-static double-labeled DNA molecules is homogeneous and exhibit clear distinct states with stepwise transitions between

them as shown in Figure 5.8. Previously it was reported that binding and bending of TBP to the TATA-box takes place in a single step [143, 173]. The results from the Group of Parkhurst support the model in which molecules with the same FRET values as DNA alone are complexes without TBP bound.

Donor and acceptor intensities are anticorrelated while the total intensity (corrected for differences in the detection efficiencies) is constant. This indicates that the different FRET values are due to changes in the distance between donor and acceptor and not due to uncorrelated noise or changes in the properties of the dye. It is obvious that the FRET values of this molecule show transitions between 3 distinct FRET states. Each of these FRET states is a result of a distinct FRET pair separation and in turn conformation of the DNA. The FRET states are "hidden" behind noise and the FRET values measured from each of the FRET states have a distribution described by the *emission function*. Using a HMM analysis, the underlying FRET states and their sequence were determined from the measured FRET values. The frame-wise histogram of the FRET values for all molecules with dynamic behavior are shown in Figure 5.9A. Using the information obtained from the HMM analysis, a 2D-histogram of the FRET states before and after transition can be composed as shown as the Transition Density Plot (TDP) in Figure 5.9B (see Chapter 3.4). Each transition is superimposed as a 2D-Gaussian with a fixed width into the TDP. Distinct clusters emerged when all transitions from dynamic molecules are plotted. It is obvious that the TDP shown in Figure 5.9B contains transitions from two populations of molecules (P_1 and P_2). Starting from any arbitrary FRET state, the molecule can only propagate through half of the transition clusters and no transitions between the two populations can be observed. Using a selection tool, clusters of transitions can be combined and the molecules which generated those transitions can be identified. Based on the transition clusters within the TDP and this selection tool, those two populations of molecules can be separated as shown by color codes in Figure 5.9D. The molecules in either of those two populations shows transitions between three states (S_1, S_2 and S_3). The three FRET states for each population are slightly shifted with respect to each other.
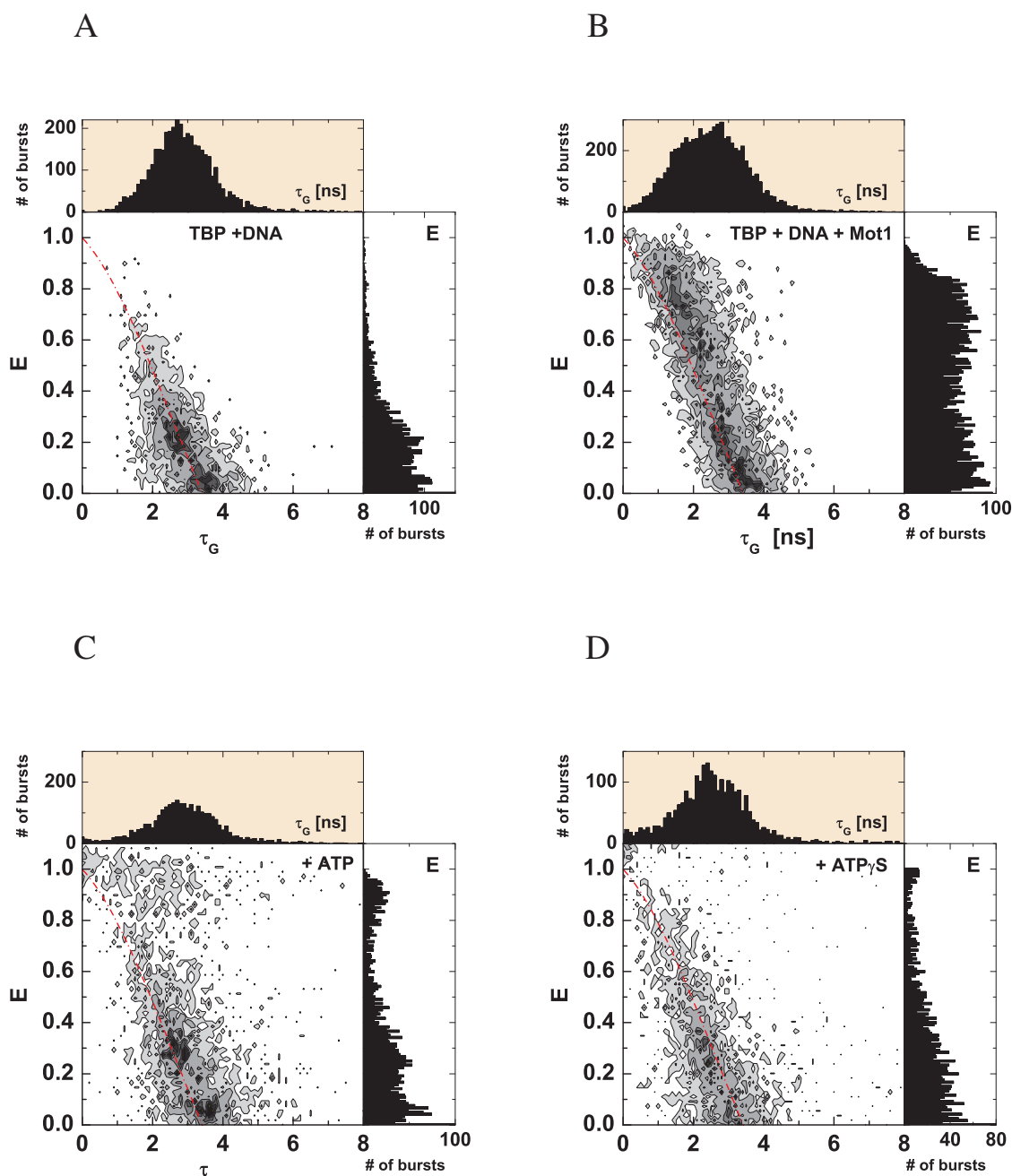
Figure 5.5.: Burst Analysis of preincubated complexes. 2D histograms of FRET efficiency vs. lifetime τ for Atto532 fluorescently labeled TBP. The relative amount of double-labeled complexes in comparison to detected donor containing complexes were estimated by selecting the bursts based on stoichiometry and are given in parenthesis. A: binary TBP/DNA complex with low FRET distribution (17%). B: TBP/DNA/Mot1 ternary complex with increased FRET (15%). C: after addition of $1\,\text{mM}$ ATP (6%). D: after addition of 1mM ATPγS (4%). The deviation from the idealized lifetime vs. E curve (shown in red) indicates that the low FRET state is quenched in the case of ATPγS addition

Figure 5.6.: Illustration of TBP bound to a double-labeled DNA containing a TATA-box.
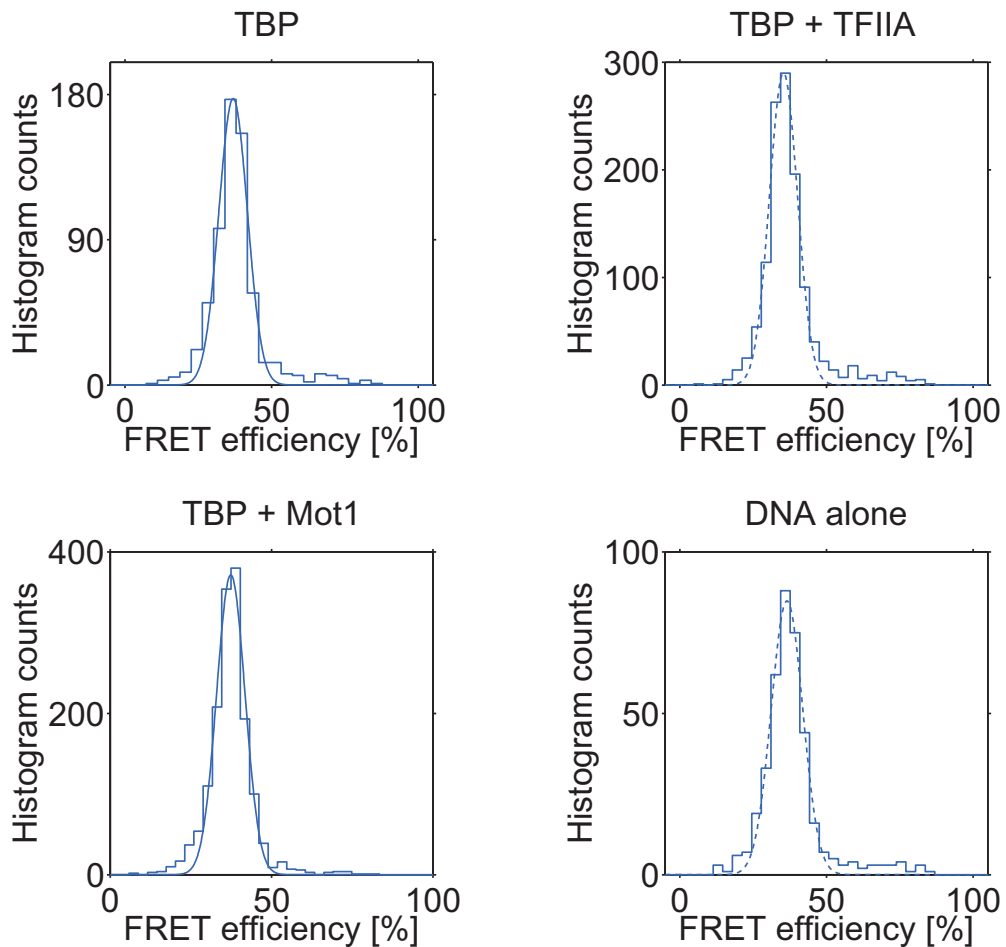
Figure 5.7.: Molecules wise FRET histograms of static double-labeled DNA molecules for the different complexes. The histograms are fitted with Gaussian distributions. The mean and the 95% confidence interval of the FRET values are obtained from the fit. The confidence interval is derived from the *standard error of the sample mean* (SEM). The values for the DNA, DNA/TBP, DNA/TBP/Mot1, DNA/TBP/TFIIA are $37,74 \pm 0,49$ %, $38,07 \pm 0,39$ %, $36,77 \pm 0,89$ %, $34,81 \pm 0,49$ %, respecively.
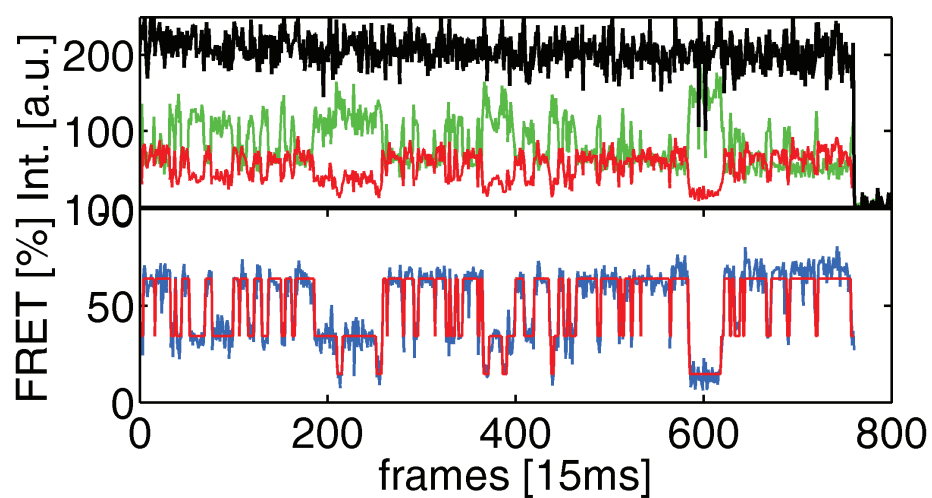
Figure 5.8.: Intensity and FRET trace of a single surface immobilized DNA/TBP complex. Donor and acceptor intensities are displayed in red and green, respectively. The total intensity is scaled by a factor of $1,6$ to best fit in the axis window and is displayed in black. The FRET values are plotted in blue and the fitted HMM states sequence in red (see Chapter 3.4 for details).
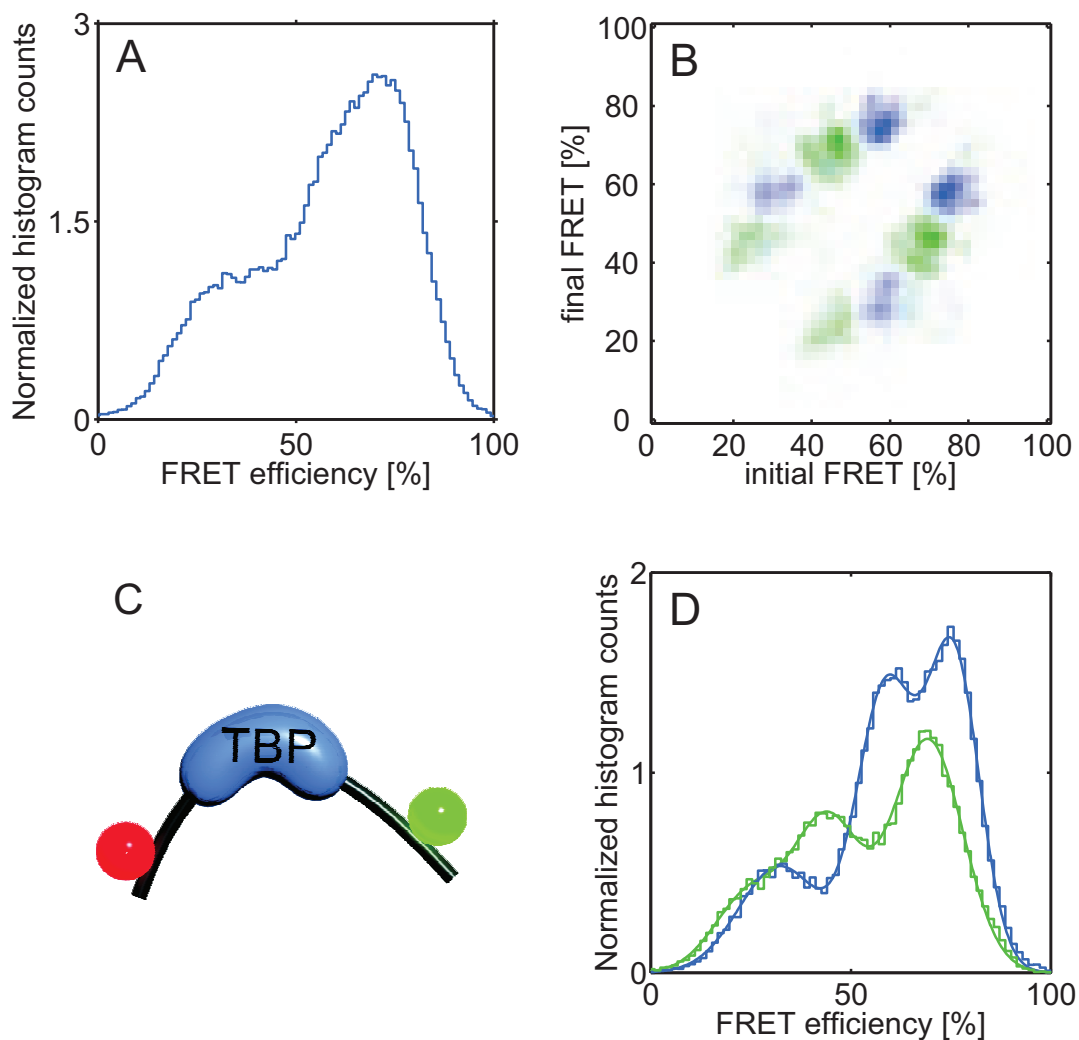
Figure 5.9.: Conformational dynamics of the TBP/DNA complex observed by measuring changes in FRET of a double-labeled dsDNA. **A**: A normalized frame-wise FRET histogram of the whole dynamic population ($P_1 + P_2$). **B**: A transition Density Plot (TDP) of the whole dynamic population. Both populations P_1 and P_2 are colored in green and blue, respectively. **C**: A schematic of the TBP/DNA complex. **D**: The frame-wise FRET histogram of the two populations. The total areas under both histograms are normalized to 1.

The results of fitting the frame-wise FRET of the binary TBP/DNA complex is given in Table 5.3.

Table 5.3.: Parameters of the normal-distributions obtained by fitting the frame-wise FRET histogram of the TBP/DNA complexes. μ and σ values are given in % of FRET. The relative area under the Gaussian distribution A is given in %.

TBP/DNA	S_1			S_2			S_3		
	A	μ	σ	A	μ	σ	A	μ	σ
P_1	18	25	9	51	45	8	31	69	9
P_2	21	32	10	39	59	7	40	76	7

Similar results are obtained when this experiment is repeated with TFIIA present during preincubation as shown in Figure 5.10. TFIIA is known to promote complex formation in one of the two binding orientations of TBP with respect to the TATA-box [140, 141]. In contrast to the TBP/DNA complex where only two populations of molecules were observed, in this case a third population P_3 appears. Molecules within this third population also exhibit three FRET states as shown in red in Figure 5.10B.
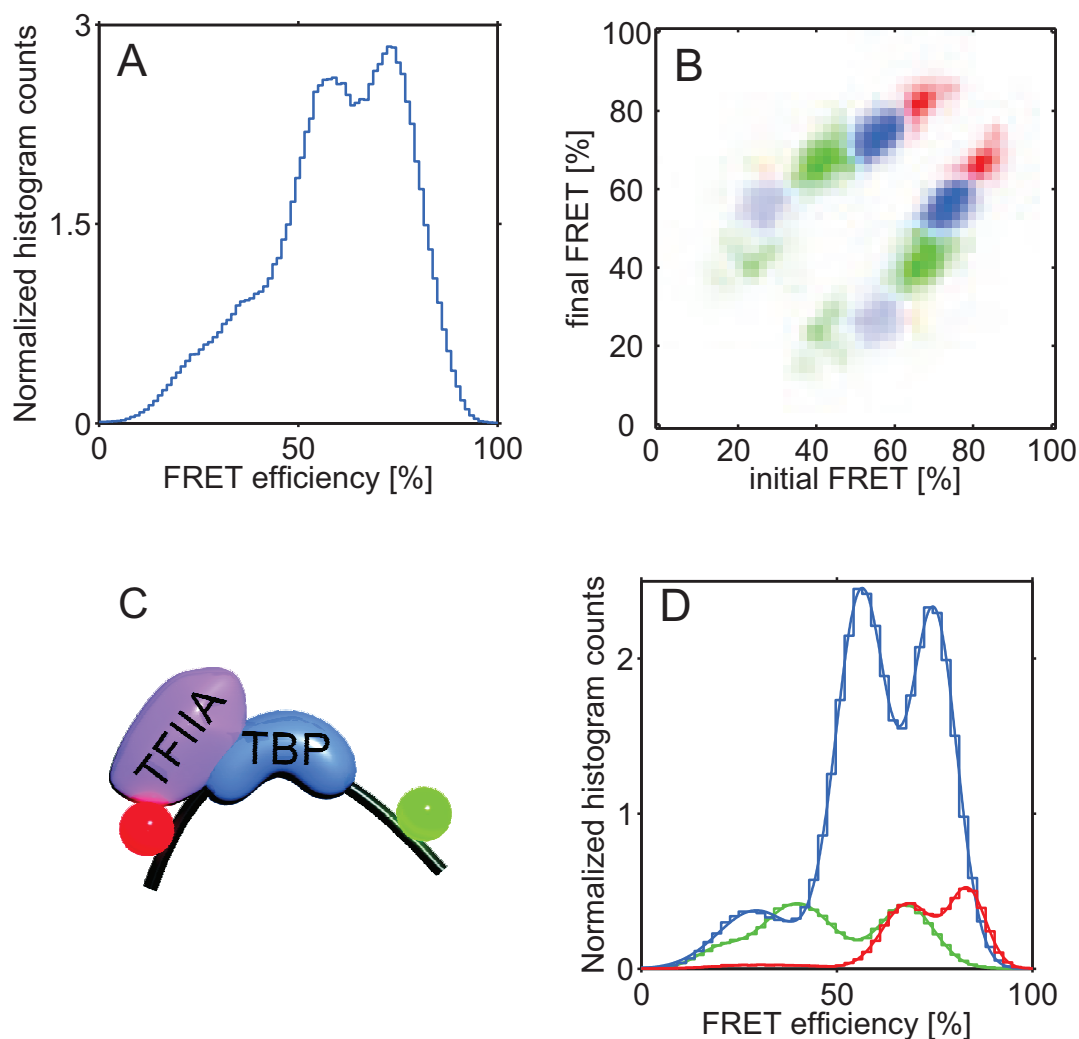
Figure 5.10.: Conformational dynamics of the TBP/DNA/TFIIA complex observed by measuring changes in FRET of a double-labeled dsDNA. **A**: A normalized frame-wise FRET histogram of the whole dynamic population ($P_1 + P_2 + P_3$). **B**: A transition Density Plot (TDP) of the whole dynamic population. Both populations P_1, P_2 and P_3 are colored in green, blue and red, respectively. **C**: A schematic of the TBP/DNA/TFIIA complex. **D**: The frame-wise FRET histogram of the three populations. The total areas under all three histograms are normalized to 1.

The results of fitting the frame-wise FRET of the TBP/TFIIA/DNA complex is given in Table 5.4.

Table 5.4.: Parameters of the normal-distributions obtained by fitting the frame-wise FRET histogram shown in Figure 5.10 from the TBP/DNA/TFIIA. μ and σ values are given in % of FRET. The relative area under the Gaussian distribution A is given in %.

| TBP/DNA | S_1 | | | S_2 | | | S_3 | | |
	A	μ	σ	A	μ	σ	A	μ	σ
P_1	9	20	12	53	40	9	39	68	7
P_2	11	29	10	48	56	7	41	75	6
P_3	6	32	12	51	68	7	44	83	5

In the case of complexes formed with an additional sequentially second preincubation step with Mot1, again three populations were observed with each three states (see Figure 5.11).

The results of fitting the frame-wise FRET of the TBP/Mot1 complex after ATP addition is given in Table 5.5.

Table 5.5.: Parameters of the normal-distributions of TBP/DNA/Mot1 obtained by fitting the frame-wise FRET histogram. μ and σ values are given in % of FRET. The relative area under the Gaussian distribution A is given in %.

| TBP/DNA | S_1 | | | S_2 | | | S_3 | | |
	A	μ	σ	A	μ	σ	A	μ	σ
P_1	15	27	9	39	45	9	46	69	7
P_2	9	30	9	43	57	8	47	75	6
P_3	51	34	11	44	60	6	5	80	6

The effect of adding $1\,\mathrm{mM}$ ATP to the ternary TBP/DNA/Mot1 complexes (without free Mot1 in solution) are shown in Figure 5.12. The FRET states observed under those conditions are the same as observed after TFIIA preincubation. Again, three populations (P_1, P_2, P_3) with each three FRET states were observed.

The results of fitting the frame-wise FRET of the TBP/Mot1/DNA complex after ATP addition is given in Table 5.6.
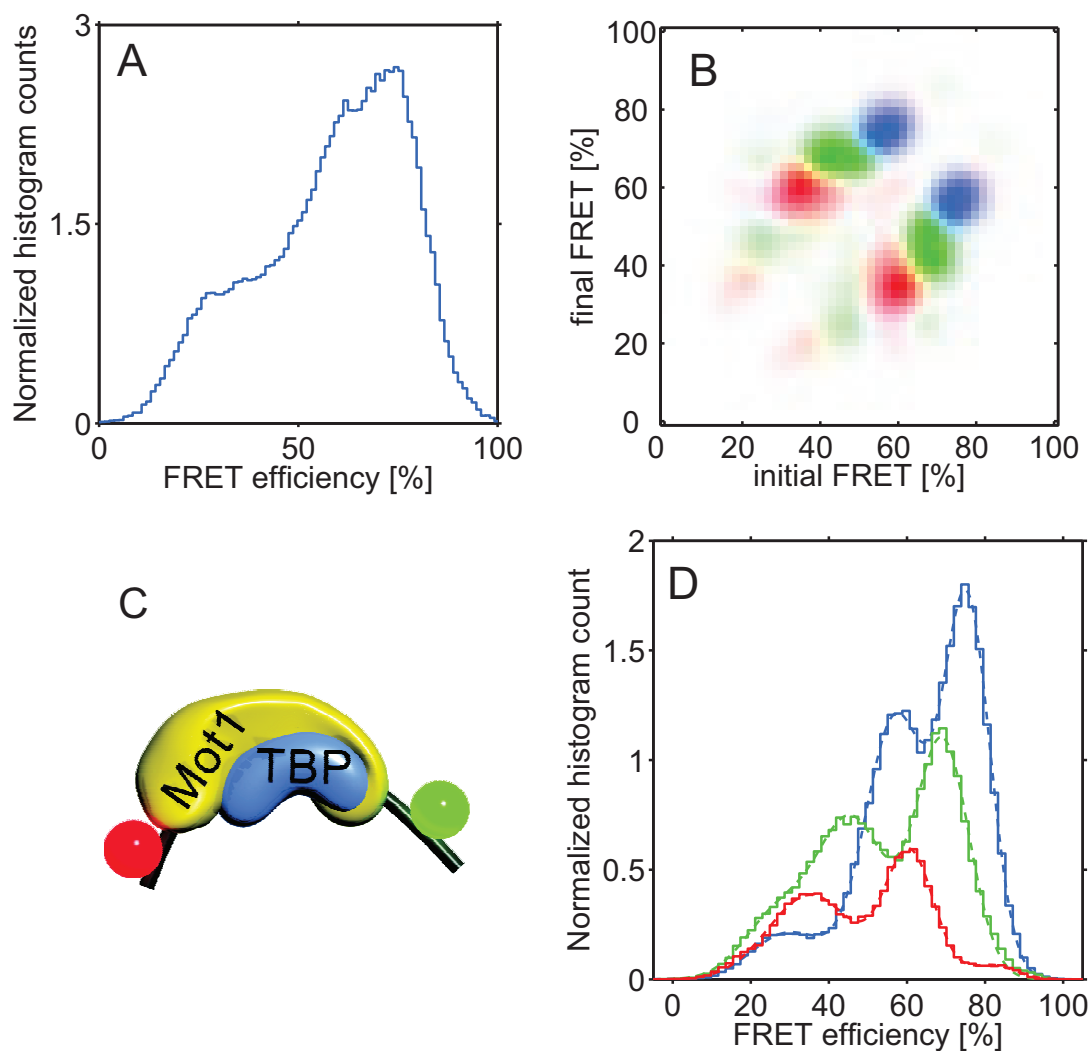
Figure 5.11.: Conformational dynamics of the TBP/DNA/Mot1 complex observed by measuring changes in FRET of a double-labeled dsDNA. **A**: A normalized frame-wise FRET histogram of the whole dynamic population. **B**: A transition Density Plot (TDP) of the whole dynamic population. The populations P_1, P_2 and P_3 are colored in green, blue and red, respectively. **C**: A schematic of the TBP/DNA/Mot1 complex. **D**: The frame-wise FRET histogram of the populations. The total areas under all histograms are normalized to 1.

The relative amount of molecules in either population P_1 or P_2 cannot be judged from the area under the states in the frame-wise histograms. Since the relative
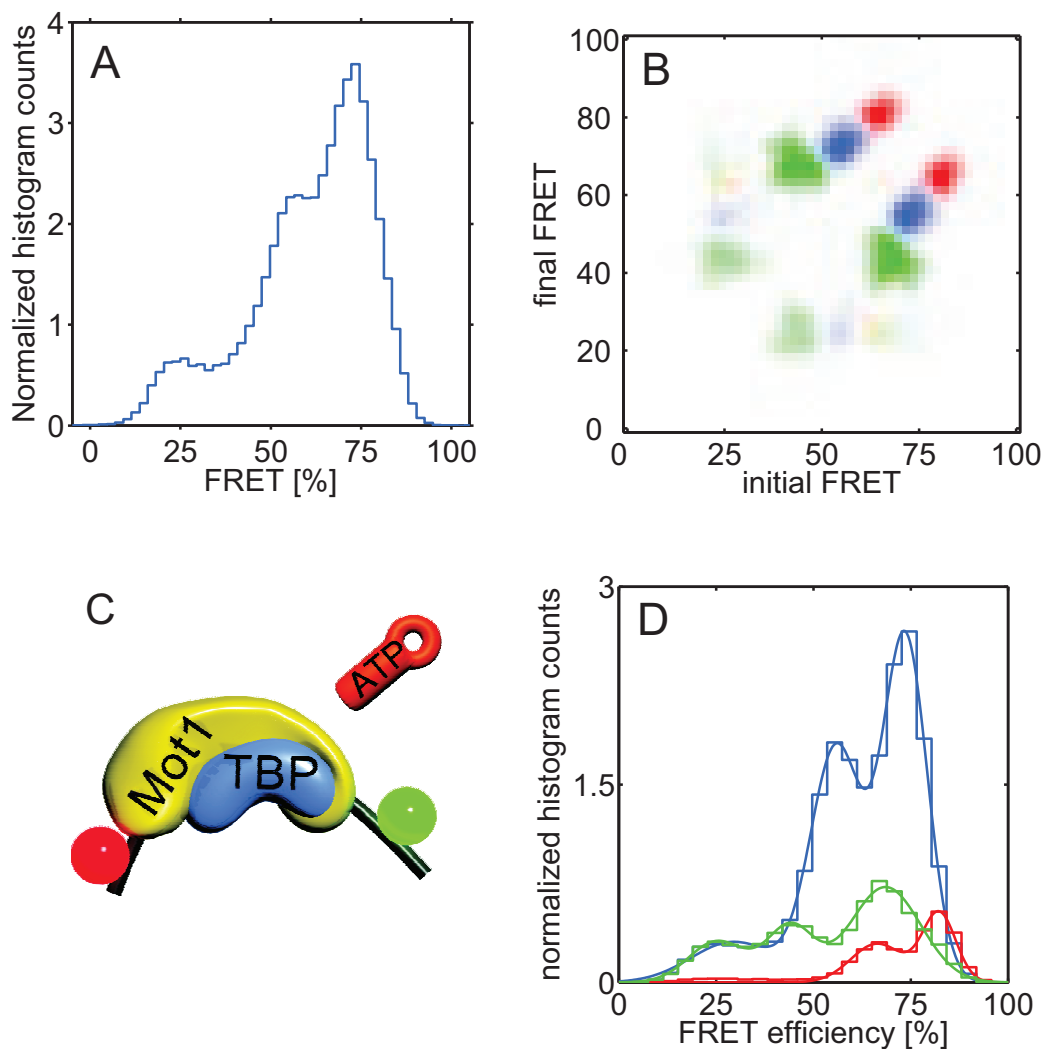
Figure 5.12.: Conformational dynamics of the TBP/DNA complex after $1\,\text{mM}$ ATP addition observed by measuring changes in FRET of a double-labeled dsDNA. **A**: A normalized frame-wise FRET histogram of the whole dynamic population ($P_1 + P_2 + P_3$). **B**: A transition Density Plot (TDP) of the whole dynamic population. Populations P_1, P_2 and P_3 are colored in green, blue and red, respectively. **C**: A schematic of the TBP/DNA/Mot1 complex. **D**: The frame-wise FRET histogram of the three populations. The total areas under all histograms are normalized to 1.

number of frames does not only depend on the number of molecules but mainly on the transitions between the different FRET states and to some extend on

Table 5.6.: Parameters of the normal-distributions obtained by fitting the frame-wise FRET histogram of the TBP/DNA/Mot1 complex after ATP addition as shown in Figure 5.12. μ and σ values are given in % of FRET. The relative area under the Gaussian distribution A is given in %.

	S_1			S_2			S_3		
TBP/DNA	A	μ	σ	A	μ	σ	A	μ	σ
P_1	22	25	7	31	45	7	47	68	7
P_2	11	30	11	38	56	6	50	74	6
P_3	15	32	14	43	67	7	42	82	4

the rate of photobleaching. Interestingly, the presence of TFIIA significantly increases the relative amount of molecules within the population P_2.

Table 5.7.: Relative amount of molecules within each population for different measurement conditions

	TBP/DNA	TBP/DNA/TFIIA	TBP/DNA/Mot1	TBP/DNA/Mot1 + ATP
P_1	49%	19%	40%	44%
P_2	51%	81%	60%	56%

From the molecules with dynamic FRET behavior, a frame-wise histogram of the FRET values can be composed (see Figure 5.9A). This FRET histogram is broadly distributed and the number of underlying FRET states not visible. The knowledge about the population to which each molecule belongs allows a separation of the molecules in both populations (e.g. Figure 5.9B). For each of the populations an individual frame-wise FRET histogram can be composed and the three individual FRET states become visible (e.g. Figure 5.9D). To minimize the broadening of the histogram due to shot noise, the histogram can be composed from the FRET states obtained by the HMM. The HMM is trained for every molecule locally and returns the optimized underlying FRET states of the individual FRET trace. For each population a separate frame-wise histogram of the HMM FRET states can be composed as shown in Figure 5.13.

The distributions of the states are narrower due to the fact that the mean FRET value found by the HMM is more precise than the individual frame-wise FRET
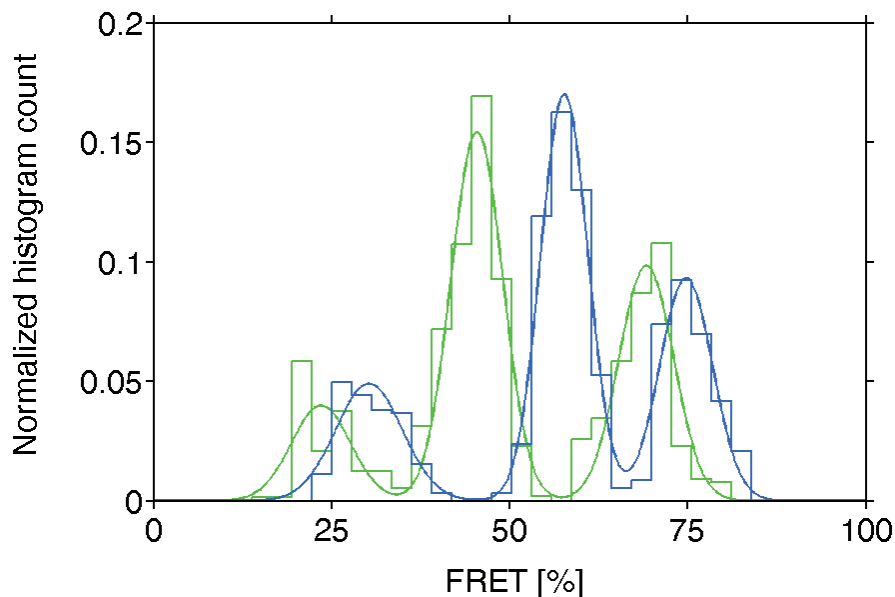
Figure 5.13.: Histogram of the FRET states of TBP/DNA binary complexes found by a HMM analysis. The two subpopulations are color-coded in green and blue.

values. The FRET distributions of the states found by the HMM are fitted to a normal distribution (see Table 5.8).

Table 5.8.: Fit results of the state-wise FRET histogram for both populations of the TBP/DNA complex using a Gaussian distribution. The error of the mean, $\pm$, is given as 95% confidence limits. The relative area under the Gaussian distribution A is given in %. The values, μ and σ are given in % of FRET.

	S_1				S_2				S_3			
TBP/DNA	A	μ	σ	$\pm$	A	μ	σ	$\pm$	A	μ	σ	$\pm$
P_1	12	30.2	3.2	0.12	59	57.6	3.3	0.11	29	74.7	3.7	0.14
P_2	12	23.5	4.2	0.16	55	45.4	3.6	0.09	33	69.2	3.8	0.15

Discussion

In contrast to ensemble measurements where Mot1 liberates TBP from DNA, no dissociation could be detected in the single-molecule experiments. Double-labeled DNA was used to investigate the effect of Mot1 on the conformation of the TBP/DNA complex. The majority of the double-labeled DNA molecules incubated with TBP, TBP/Mot1 or TBP/TFIIA exhibit static FRET indistinguishable from the FRET values observed from DNA alone (see Figure 5.7). In the contrast to experiments in Section 5.2.1, where FRET between TBP and DNA was measured, in this section only the DNA is labeled and the unlabeled TBP is not directly visible. Since TBP binding induced conformational change of the TATA-box is widely acknowledged, it is most likely that molecules showing the same FRET values as DNA alone do not have TBP bound. After excluding the molecules without bound TBP, as determined from the FRET value, only highly dynamic TBP/DNA complexes remain. In contrast to results published previously [144] where the formation of a stable highly bent TBP/DNA conformation was reported, no complex with static high-FRET values could be observed. This is not surprising considering that the previous results were obtained through ensemble measurements, which are not able to resolve dynamic FRET. The dynamics of the TBP induced DNA bending was reported from two single-molecule experiments [145, 174]. In one study, the TBP binding induced DNA bending was reported to involve three dynamic conformations of DNA. The results were obtained by measuring the *root mean square displacement* (RMSD) of a bead tethered to the surface by a TATA-box containing DNA [145]. From 10 molecules observed in this study, only one was shown to spend an extended time in a stably bent conformation. The findings were interpreted in the framework of a static final bound TBP/DNA complex as proposed by Parkhurst [144]. In the second study, TATA-box bending was monitored by smFRET between the fluorescently labeled TBP and a label on the DNA upstream of the TATA-box. Six different bending conformations were observed upon TBP binding.

For the TBP/DNA complex, two different populations of molecules could be identified based on transitions between three FRET states. It was reported that TBP binds DNA in both binding orientations with only modest affinity for the

orientation shown in the crystal structures and that the preference of TBP toward the preferred binding orientation is greatly increased by addition of TAIIF (from 64% to 84%) [140]. The interpretation of the two populations as the two binding orientations of TBP is supported by the relative amount of molecules in either population before and after TFIIA preincubaiton. The increase in the binding specificity is perfeclty reflected in the shift of the relative amount of molecules observed in this single-molecule experiment (see Table 5.7). This finding strongly suggests that the molecules belonging to population P_1 are complexes where TBP is bound to the TATA-box in the inverted orientation and that population II P_2 are complexes where TBP is bound to the DNA in the preferred orientation, as shown in the crystal structure [122].

An additional high FRET population, P_3, was observed whenever an additional protein interacted with the binary TBP/DNA complex. The population P_3 has overlapping FRET states with the other two populations P_1 or P_2. All the transition rates of the population P_3 are comparable to the transition rates of population P_2 for Mot1 preincubation and to P_1 in the case of TFIIA preincubation and ATP addition to the TBP/Mot1/DNA complex (see Section 5.2.3 and Table 5.2.3). The γ-factors for population P_3 differ from the average γ-factor of the other two populations ($\langle \gamma_{Mot1} \rangle = 0, 65$, $\gamma_{\mathrm{Mot1}\ P_3} = 0, 74$ and $\langle \gamma_{\mathrm{TFIIA}} \rangle = 0, 59$, $\gamma_{\mathrm{TFIIA}\ P_3} = 0, 47$ with a 95% confidence interval < 0.03). The chromophores were located 3 base pairs up- and 5 bp down-stream of the TATA-box with a total separation of 18 base pairs. The presence of an additional bound protein can interfere with the dyes when the complex is in a specific conformation. One possible explanation is a protein/dye interaction which alters the ϕ_A- and in turn the γ-factor of one of the FRET state. The dynamic γ-correction used for dynamic FRET traces assumes a constant γ-factor throughout the trace. A slightly shifted γ-factor for one of the three states would not be detectable but could cause a slight shift in the FRET value of one of the states. The molecules with one slightly shifted FRET state would be identified as an additional population within the TDP plot.

The observed FRET states S_i of the DNA within each population (P_1, P_2 and P_3) do not change upon addition of Mot1 or TFIIA. However, the relative number of molecules within each population and the transition rates between the states do change. As a result, the observed ensemble FRET value can change despite the fact that the individual FRET states stay the same (see Chapter 5.3).

During the time window of observation[3], the orientation of TBP on the DNA is conserved. Neither TFIIA nor Mot1 were able to directly invert the binding orientation of TBP. The increase in the number of TBP/DNA complexes with the preferred binding orientation in the presence of TFIIA can be explained by two mechanisms. Either the binding affinity or the stability of the binary TBP/DNA complex is modulated by TFIIA (see Section 5.2.3). The results from this section support a model in which the inversion of TBP binding orientation requires TBP dissociation and rebinding.

The three FRET states within each of the subpopulations reflect different DNA conformations. Binding of TBP to the DNA and bending of the DNA involves local strand unwinding, twisting and tilting of the DNA in three dimensions, as observed in the crystal structure. The specification of a quantitative model for the conformations involved would require more information than can be obtained from a single spFRET measurement alone. Despite the complexity of the bending process, the different FRET efficiencies observed from a double-labeled DNA can be interpreted by a simple DNA model [143]. The three FRET states, S_1-S_3, observed for either binding orientation can be interpreted in the framework of a simple one dimensional model of DNA. The DNA conformations are labeled with increasing FRET efficiency from S_1 for the lowest efficiency over S_2 to the highest efficiency state S_3. The changes in the FRET values can be interpreted as kinks in the DNA (based on the two-kink model [172, 143]). The transitions between the different FRET values can be interpreted as the dynamics of bending (for details see Appendix A.2). Using the FRET values from Table 5.8, the conformations of the TATA-box containing DNA are illustrated in Figure 5.14 based a the two-kink model for the DNA.

---

[3]The average observation time for the TBP/DNA complex was $5.8\,\mathrm{s}$

From the TDP plots (see Figures 5.9-5.12B), it is obvious that the three FRET states within each population are connected linearly. In other words, no transition between the highest and the lowest FRET value are observed, only "nearest-neighbor-transitions" occur.
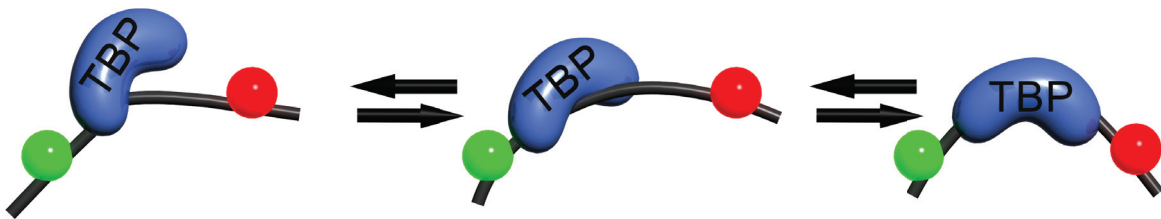

Figure 5.14.: Illustration of the different angles of TBP induced DNA bending. For the population $P2$, the bending angles calculated with a two-kink model are: $68°$ and $96°$ for the intermediate and the maximum bent state, respectively. The FRET value of the lowest bent state is lower than the FRET value of the DNA alone and can't be interpreted with this simplified model. This is attributed to DNA unwinding.

The FRET values measured in ensemble experiments for e.g. the TBP/DNA binary complex are the FRET values, averaged over both populations and depend not only on the eight rates between the six different states but also on the relative number of molecules found in each of those populations. A change in the ensemble measured FRET value of a double-labeled DNA containing TBP does not necessarily mean that the FRET value and in turn the bending of DNA changed. The results in this Section show that interaction of TFIIA and Mot1 with the TBP/DNA complex does not change the bending of the DNA. The influence of Mot1/TFIIA on the dynamics of the transitions between the different conformational states is described in the next section.

5.2.3. Kinetic of TATA-box bending

To gain insight into the dynamics of the transitions between the different DNA conformations, the rates of the transitions need to be analyzed. Applying a HMM analysis to the FRET traces not only provides information about the FRET values of the individual states, but also about their dwell times. Kinetic

information about the transitions were extracted based on the TDP plot as shown in Figure 5.15. Within the TDP plot, a cluster of transition was selected with an elliptical selection tool.
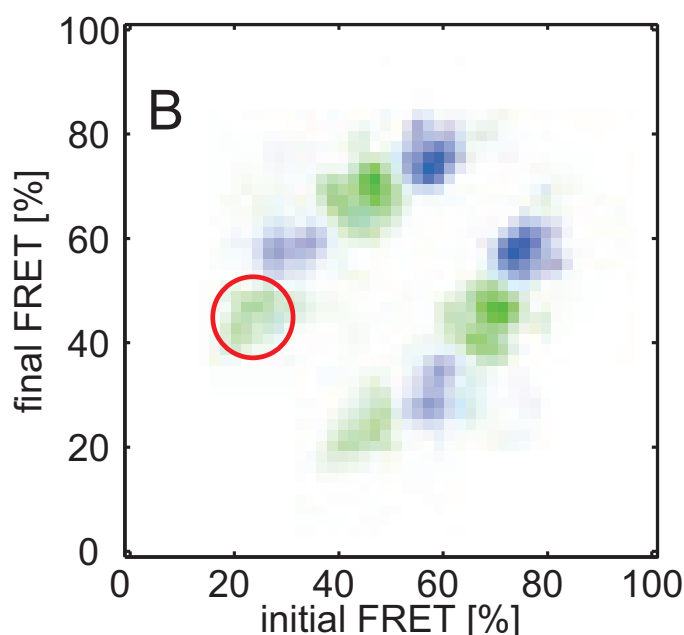

Figure 5.15.: Transition Density Plot (TDP) of TBP/DNA binary complexes. Subpopulations are color coded in green and blue. A cluster of transitions is selected with an elliptical selection tool shown in red

All transitions within the region of the selection tool were used to compose a dwell time histogram. The dwell time histogram of each cluster of transitions was fitted to a mono-exponential decay to obtain the rate for a particular transition as shown in Figure 5.16.

Applying this analysis to all transitions for both subpopulations, all transition rates were obtained as given in Table 5.10 and 5.9 (see Appendix A.6 for individual fits).

Using a three state linear model each of the three FRET states is interpreted as a distinct DNA conformation, the transition rates are visualized in Figure 5.17 and 5.18.
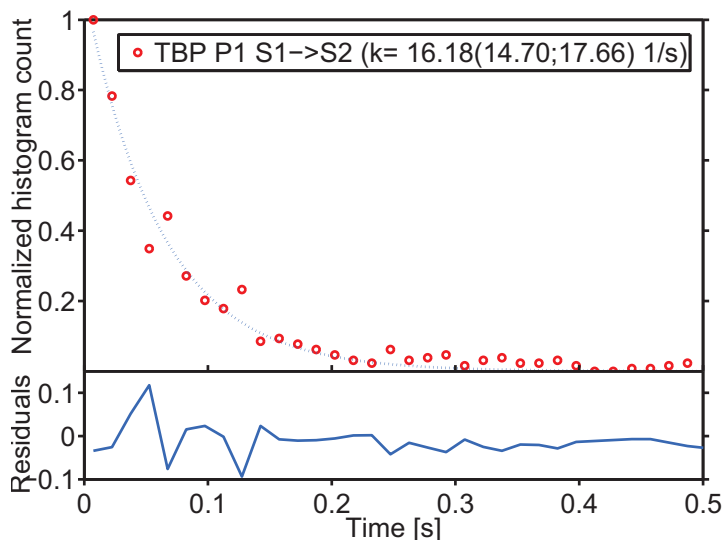
Figure 5.16.: Dwell time of the FRET state of the DNA/TBP complexes in population I. Upper panel: Histogram and monoexponential fit of the dwell times from the selection shown in Figure 5.15. Lower panel: Residuals

Table 5.9.: Transition rates of molecules in the population P_1. Transition rates are given in s^{-1}. Rates are obtained from a monoexponential fit of the dwell time histograms (see Appendix A.6 for fits and confidence intervals). The error is given as standard deviation of the rates from randomly chosen subsets of molecules in the same state.

$S_{start} \to S_{stop}$	TBP	$\pm$	TFIIA	$\pm$	Mot1	$\pm$	+ATP	$\pm$
$S_1 \to S_2$	16	5.9	7	0.9	11	5.2	10	1.3
$S_2 \to S_1$	41	1.1	53	6.7	36	7.2	62	4.6
$S_2 \to S_3$	25	1.1	22	0.8	16	0.9	29	0.6
$S_3 \to S_2$	12	1.9	7	0.8	9	1.9	10	2.9

Table 5.10.: Transition rates of molecules in the population P_2. Transition rates are given in s^{-1}. Rates are obtained from a monoexponential fit of the dwell time histograms. The error is given as standard deviation of the rates from randomly chosen subsets of molecules in the same state.

$S_{start} \to S_{stop}$	TBP	$\pm$	TFIIA	$\pm$	Mot1	$\pm$	+ATP	$\pm$
$S_1 \to S_2$	13	2,3	10.92	0,9	12	1,9	14	3,0
$S_2 \to S_1$	15	0,6	15.74	12	29	9	32	4,5
$S_2 \to S_3$	13	0,3	7.91	3,0	10	8	11	0,8
$S_3 \to S_2$	7	1,2	4.68	0,3	6	1,2	7	0,6

Table 5.11.: Transition rates of molecules in the population P_3. Transition rates are given in s^{-1}. Rates are obtained from a monoexponential fit of the dwell time histogram of the transition start state S_{start}. The error is given as standard deviation of the rates from randomly chosen subsets of molecules in the same state.

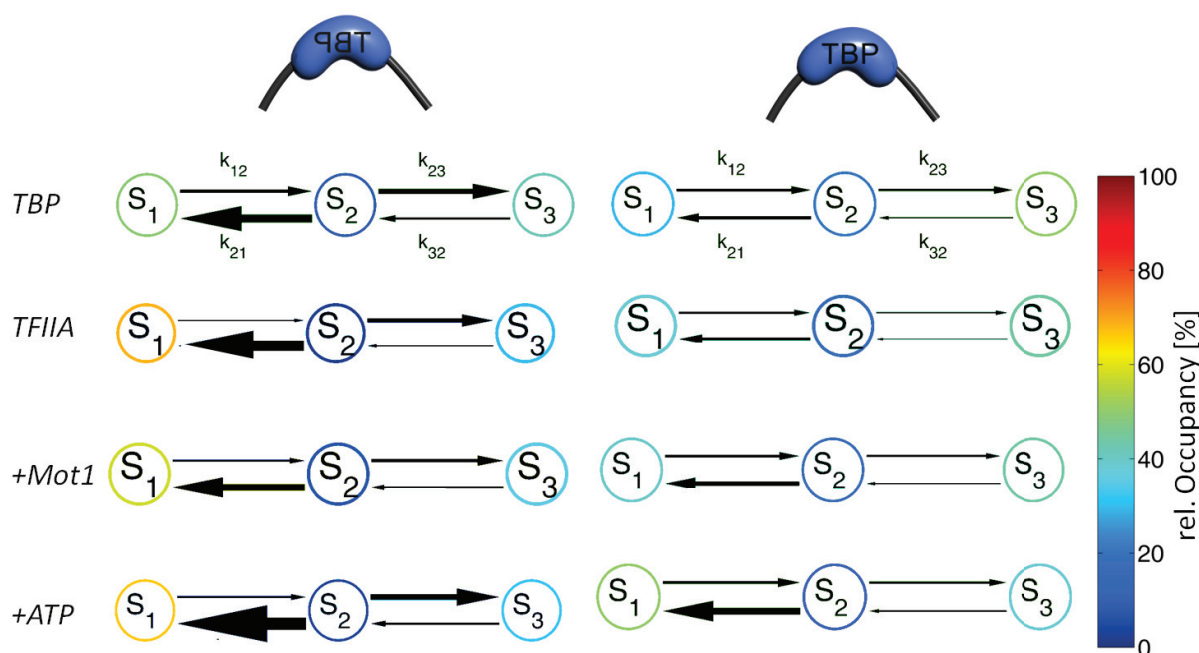| $S_{start} \rightarrow S_{stop}$	+ATP	$\pm$	TFIIA	$\pm$	Mot1	$\pm$
$S_1 \rightarrow S_2$	19	5,5	11	0,1	10	1,7
$S_2 \rightarrow S_1$	38	8,4	11	0,2	35	7
$S_2 \rightarrow S_3$	12	1,9	6	0,1	16	2,6
$S_3 \rightarrow S_2$	6	1,5	4	0,6	12	2,2

Figure 5.17.: Schematic representations of the kinetic rates for both binding orientations (population I and II) under different experimental conditions. Individual rates are coded by line thickness. The relative occupancy of each state is color coded and calculated based on the rates assuming equilibrium conditions. Within each binding orientation (P_1 and P_2) three FRET states (S_1,S_2,S_3) are observed. The transition rates k_{ij} from the state S_i to the state S_j are obtained by fitting the state dwell times of state S_i as a single-exponential decay.

The rates for the additional high-FRET states in the case of ATP addition and TFIIA/Mot1 preincubation are shown in Figure 5.18
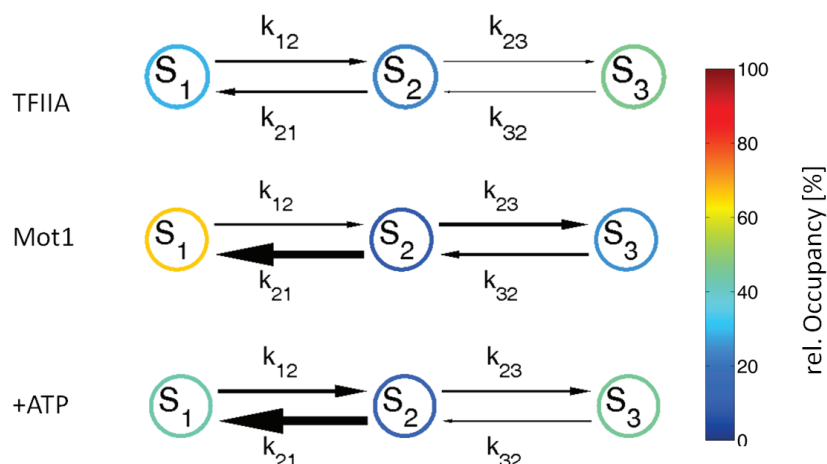

Figure 5.18.: Additional populations $P3$ in the case of TFIIA, Mot1 or ATP addition to the TBP/DNA/Mot1 complex. Individual rates are coded by line thickness. The relative occupancy of each state is color coded and calculated based on the rates assuming equilibrium conditions. Three FRET states (S_1,S_2,S_3) are observed. The transition rates k_{ij} from the state S_i to the state S_j are obtained by fitting the state dwell times of state S_i as a single-exponential decay.

The relative occupancy, as indicated by color coding in Figure 5.17 & 5.18 is based on the kinetic rates obtained from the fits. If the forward and backward rate are similar in magnitude, the relative occupancy is less sensitive to the errors in the determination of individual rates. However, if the magnitude of the rates differ significantly, errors will have a stronger effect on the calculated occupancy.

Discussion

TBP binding to the TATA-box is a highly dynamic process involving three DNA conformations. The transitions between the three conformations are next-neighboring transitions as visible from the TDP plot. The three FRET efficiencies of each population can be interpreted as three DNA conformations. Based

on a linear three step model, the transition rates obtained from the HMM analysis can be used to compare the different experimental conditions measured for the different TBP binding orientations.

From the transition rates it becomes clear that the effect of Mot1, TFIIA and ATP binding to Mot1 is TBP orientation specific. The rate between the intermediate S_2 and lowest bent S_1 conformation is strongly affected for the inverted TBP binding orientation. Preincubation with TFIIA or addition of ATP to the Mot1/TBP/DNA complex drastically increases the transition rate from state $S2$ to state $S1$ while preincubation with Mot1 does the opposite.

Previously it was reported that TFIIA promotes the formation of TBP/DNA complexes in the orientation observed in the crystal structure [140]. In principle, TFIIA can use two mechanisms to modulate binding specificity: orientation specific modulation of either the on- or off-rate for binding. The change in the relative number of molecules in either population supports a model in which the minimal bent state S_1 is also the state with the lowest TBP/DNA complex stability and that TFIIA increases TBP binding specificity by destabilizing TBP bound in the wrong orientation.

The strongest increase of rates is observed for the S_2 to S_1 transition after ATP addition to the TBP/DNA/Mot1 complex. As observed in the previous sections, TBP does not dissociate from the TATA-box. The rate from state S_2 to the S_1 state for inverted bound TBP/Mot1 complex after ATP addition is the fastest rate observed in this assay. This observation supports the model in which Mot1's ATPase activity is a TBP binding orientation specific destabilization of the TBP/DNA complex.

So far, no dissociation of TBP from the TATA-box could be observed. From the transition rates, it can be concluded that the ATPase activity of Mot1 promotes the formation of a linearized conformation of the DNA. The interaction of this primed TBP/Mot1/DNA complex with additional Mot1 molecules is investigated in the following section.

5.2.4. Dissociation of TBP from the TATA-box

From *in vitro* experiments, it was reported that Mot1's ATPase activity directly dissociates TBP from DNA [161, 171, 165]. However, in the experiments described in the previous sections, no dissociation of TBP from the TATA-box could be observed. The influence of Mot1/TFIIA is reflected in the change of the transitions rates for the highly dynamic TBP/DNA complex as reported in Section 5.2.2. To verify the dissociation activity of Mot1 on a single-molecule level, the binding of single TBP molecules to the TATA-box under the influence of Mot1 was observed. The binding of TBP to DNA was monitored by measuring the fluorescence intensities of the fluorescently labeled TBP molecules. As long as fluorescence could be observed, the TBP was assumed to be bound to the DNA. Disappearance of the fluorescence signal after ATP addition can be due to two factors: photobleaching (see Appendix A.4) or dissociation of TBP from the DNA. The rate for photobleaching of a donor molecule depends on the time that a donor molecule spends in the excited state and in turn on the FRET value. From the experiments in Section 5.2.1, we know that ATP addition induces a significant change in the FRET efficiency and in turn will alter the rate of donor bleaching. To avoid any influence of FRET on the rate of donor photobleaching, TBP was labeled as acceptor in the experiments described in this section.

Experiments and Results

The preincubated ternary complex, TBP/Mot1/DNA, were immobilized at the surface and unbound complexes were washed away by flushing with buffer. This washing step was performed with working buffer and removed free Mot1 from the sample chamber. In contrast to ensemble measurements, where unbound TBP and Mot1 are omnipresent, in this experiment two different initial conditions can be established. The experiment can be performed without any residual protein in solution by removing them with a flushing step or by adding Mot1 to the buffer that is used to add ATP. This way a reproducible amount of unbound Mot1 is added to the reaction chamber. The different experimental conditions

are illustrated in Figure 5.19.

Prior to each dissociation experiment, the characteristic photobleaching time was determined[4]. This *Time Before Disappearance* (TBD) was used to account for photobleaching and to identify any additional contributions to the disappearance of the fluorescence.

The experiment was prepared in such a way that the recording was started before addition of ATP to the flow chamber and care was taken that ATP addition did not cause focus drift. The effect of ATP addition to a sample preparation in the flow chamber can only be recorded once per ATP addition. This limits the number of molecules which are available for the analysis. The number of molecules available for analysis was increased by doubling the field of view and recording the fluorescence from the red channel only. The first 10 frames were recorded during excitation of the Atto532 labeled DNA with $532\,\mathrm{nm}$ light. The fluorescence recorded in the red channel during this first 10 frames was due to FRET between the donor labeled DNA and the acceptor labeled TBP. This first 10 frames were used to identify the complexes where TBP was bound to the TATA-box. Only molecules where fluorescence in the red channel due to FRET could be observed, where selected for analysis. After the first 10 frames and before ATP addition the excitation was switched to $647\,\mathrm{nm}$.

The decay in the number of TBP/Mot1/DNA complexes over time for both addition schemes (ATP $\pm$ Mot1) are shown in Figure 5.20. The frames before ATP addition are removed and the decay normalized.

The previous experiment was repeated but without Mot1 preincubation. The binary TBP/DNA complexes were immobilized and ATP together with Mot1 added. No significant dissociation could be observed as shown in Figure 5.21.

The results were fitted to a mono-exponential decay. The rates obtained from the experiments with ATP and Mot1 addition were corrected by subtracting the rate of photobleaching. The resulting rate was measured as a function of ATP concentration as shown in Figure 5.22.

---

[4]Significant variations of the photobleaching time between different flow chambers were observed, which makes an individual calibration necessary.
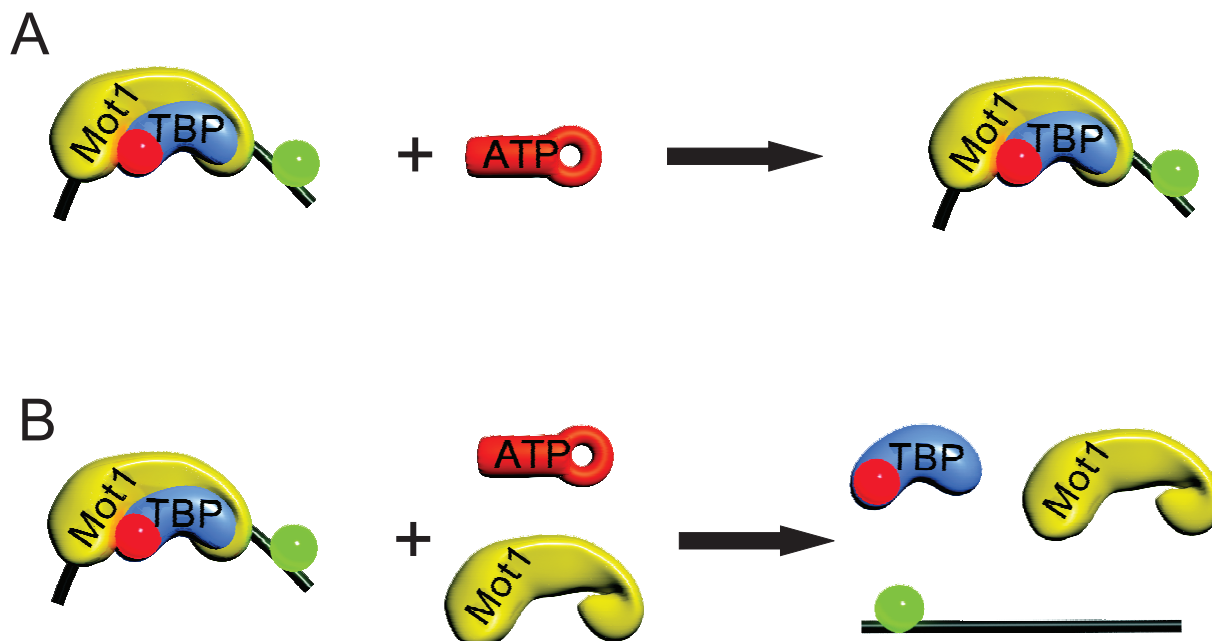
Figure 5.19.: Illustration of the Mot1-catalyzed TBP-DNA dissociation. A: Reaction scheme
without free Mot1 in solution by removing any unbound proteins by flushing with buffer
prior to ATP addition. B: Reaction scheme with Mot1 by flushing the reaction chamber
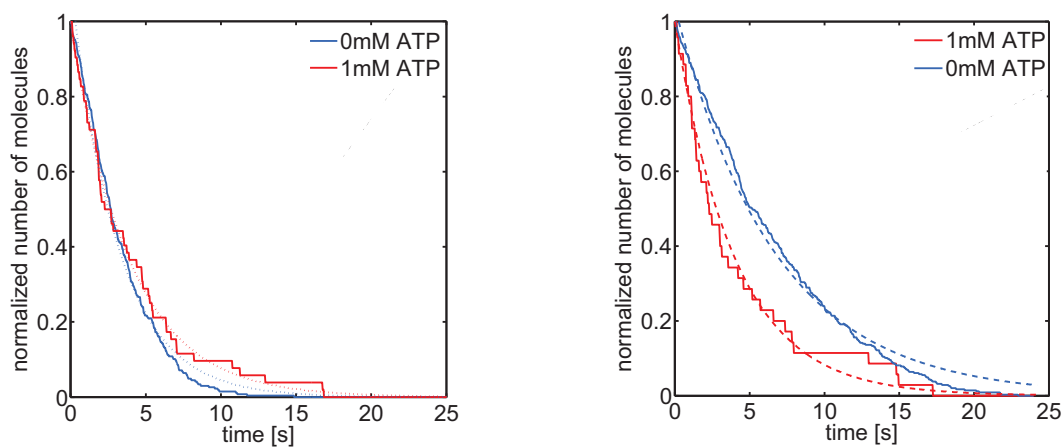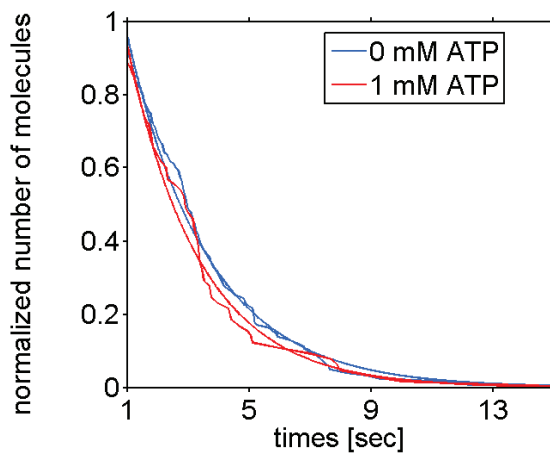with a Mot1 and ATP containing

Figure 5.20.: Kinetics of Mot1-catalyzed TBP-DNA dissociation measured by the decay of
number of fluorescent molecules. Left: addition of 1mM ATP to DNA/TBP/Mot1 ternary
complexes. The red curve shows the decay of number of fluorescent molecules due to
photobleaching (without ATP), the blue curve shows the decay in the number of fluores-
cent molecules after ATP addition. Right: same experiment as before but with the addition
of $1\,\text{mM}$ ATP together with $3.4\,\text{nm}$ Mot1.

Figure 5.21.: Kinetics of Mot1-catalyzed TBP-DNA dissociation measured by the decay of number of fluorescent molecules. Addition of $1\,\text{mM}$ ATP together with $3.4\,\text{nm}$ Mot1 to preincubated binary TBP/DNA complexes. In contrast to Figure 5.20 no significant dissociation is observed if TBP is not preincubated with Mot1.
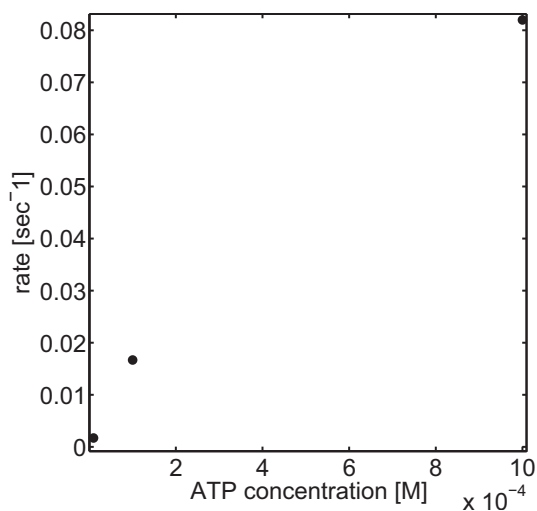

Figure 5.22.: Dependency of the TBP dissociation rate as function of ATP concentration.

Discussion

The Mot1 ATPase activity does not directly dissociate TBP from DNA when no free Mot1 is available in solution. TBP is liberated only in the presence of free Mot1 in solution. The results of these experiments are in agreement with *in vitro* experiments where TBP displacement by Mot1 ATPase activity was demonstrated [161]. The important difference between previous ensemble measurements and the single-molecule data is the presence or absence of free Mot1 in solution. The rate of the TBP dissociation is ATP and Mot1 concentration dependent. The lack of TBP dissociation could be a result of insufficient Mot1 binding during the preincubation. This can be ruled out since Mot1 readily binds the binary TBP/DNA complex as shown in Figure 5.3, where the majority of molecules undergo a conformational change during Mot1 preincubation.

Whether or not the complexes after ATP addition (without free Mot1 in solution) still have Mot1 bound cannot be judged from these experiments alone. Two lines of evidence support a model in which Mot1 is still bound to the TBP/DNA complex after ATP addition: In the experiments with double-labeled DNA (see Section 5.2.2), an altered DNA dynamic after ATP addition was observed, suggesting that Mot1 is still present after ATP addition and, in the three-color experiments (see Figure 5.4), Mot1 could still be detected after ATP addition.

The differences in Mot1 ATPase activity for experiments with and without free Mot1 in solution can be explained by two models: Mot1 dissociates TBP from the TBP/binary complex during ATPase activity with a low efficiency. In this framework, with only ATP in the solution, Mot1 has only one round of ATPase hydrolysis to liberate TBP. Rebinding or binding of another Mot1 molecule is very unlikely due to the low local concentration of Mot1 molecules available at the prism surface. If Mot1 is supplied in the solution together with ATP, repeated rebinding is possible and, despite the low dissociation efficiency, TBP liberation becomes possible. An other possibility is that the ATPase activity of Mot1 does not dissociate TBP from DNA but rather modulates and weakens the binding. In this model, Mot1 does not necessary need to be displaced from TBP to explain the lack of TBP dissociation activity. If additional Mot1 is avail-

able in solution, it dissociates TBP by interactions with the "primed" TBP/DNA complex. In this model, Mot1 has two distinct functions, modulation of the TBP/DNA conformation and dissociation of the primed TBP/DNA complex. It is important to note that, in this model, more than one Mot1 molecule is required for TBP liberation.

5.3. Discussion

Several models have been proposed for how Mot1 dissociates TBP from the TATA-box. The proposed models include direct displacement of TBP by translocation of Mot1 along the DNA [171], indirect displacement by modulations of the interactions of TBP with DNA [160] or a combination of both [166]. These models explain how Mot1 liberates TBP from DNA and in turn gene repression. However, these models have difficulties in explaining the role of Mot1 in gene up-regulation [165]. As a general framework for how Mot1 is able to up-regulate gene expression, it was proposed that Mot1's ATPase activity liberates transcriptional inactive TBP and in turn redistributes TBP along DNA [157, 159, 167]. However, little is known about the details of this mechanism - how and if this mechanism is TBP binding orientation specific.

Based on the dissociation experiments in Section 5.2.3, two models of Mot1's ATPase activity can be proposed: Either, Mot1 is inefficient in TBP removal and requires several rounds of ATP hydrolysis [165] or Mot1 displaces TBP in a two step mechanism where the first step is the formation of a "primed" conformation and in the second step, this primed TBP is liberated upon interaction with an additional Mot1 molecule, as illustrated in Figure 5.23.

Several lines of evidence support the model shown in Figure 5.23: The results from the experiments in Section 5.2.1 indicate that Mot1 is still bound to the TBP/DNA complex after ATP binding. Since ATP is available in excess in the solution, Mot1 bound to the TBP/DNA complex can undergo several rounds of ATP hydrolysis without TBP displacement. The addition of Mot1 together with ATP to TBP/DNA complexes did not yield significant dissociation (see Figure 5.21), supporting the model in which the presence of a primed TBP/DNA complex is required for TBP dissociation.

The nature of the primed conformation was investigated with the experiments described in Section 5.2.2 where the TBP induced dynamics of the TATA-box was analyzed. The two binding orientations of TBP to the TATA-box were identified based on the transitions between the DNA conformations. Interestingly,
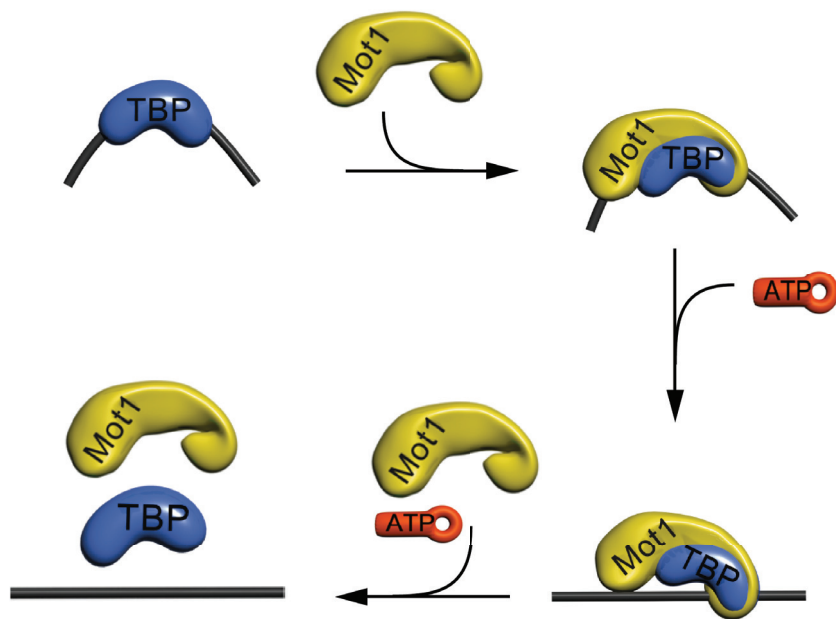
Figure 5.23.: Schematics of Mot1 induced TBP dissociation. Mot1 binds to the binary TBP complex and forms a stable ternary TBP/DNA/Mot1 complex. This binding step induces a conformational change in the TBP/DNA interactions. Addition of ATP "primes" the complex in a linearized DNA conformation. Complexes where TBP is bound in the inverted binding orientation are the primary target for priming. Addition of Mot1 together with ATP liberates TBP from the DNA in an ATP and Mot1 dependent manner

the binding orientation of TBP is conserved during the time window of observation ($15\,\text{s}$). This suggests that inversion of TBP's binding orientation (by TFIIA) requires the dissociation and rebinding of individual TBP molecules.

The DNA shows transitions between three different conformations for each binding orientation. Using a HMM model, the FRET states and the transition rates between them could be determined. The influence of ATP addition to the TBP/DNA/Mot1 complex is strongly TBP orientation specific. These changes are not changes in FRET values but changes in the rates between the different FRET states. Mot1's ATPase activity, similar to the effect of preincubation with TFIIA, promotes straightening of the DNA. In the case of TFIIA, the formation of complexes where TBP is bound in the inverted binding orientation is suppressed.

This findings can be explained with a model in which the linearized conformation of the DNA state S_1 is a destabilized TBP/DNA conformation prone to dissociation. Since the effects of ATP addition is the strongest for the inverted binding orientation of TBP, these molecules are expected to dissociate faster from the TATA-box. This model is supported by an *in vivo* study of the interaction of the URA1 gene with Mot1. It was suggested that Mot1 up-reagulates transcription by liberating TBP bound in the transcriptionally inactive reversed-binding orientation [175].

To investigate how well the rates from the model can reproduce the measured ensemble FRET values, an ensemble FRET value of the model was calculated based on the relative occupancy in each of the three states, assuming equilibrated states and constant rates. The average FRET value for the individual populations was calculated by the weighting the measured FRET values of the states with the relative occupancies and is given in Table 5.12.

The model correctly reproduces the observation that the population P_1 has a lower FRET efficiency than the state P_2 under all measuring conditions. The observation that the FRET efficiency of the DNA in the P_1 population is lowered and in P_2 population is increased after ATP addition to the TBP/DNA/Mot1 complex compared to the TBP/DNA complex is consistent with our model.

Table 5.12.: Measured and calculated averaged FRET values of population P_1, P_2, P_3. The calculated FRET values are based on the occupancy obtained from the rates in the 3 state model. The values are given in % of FRET.

	calculated			measured		
	P_1	P_2	$P3$	P_1	P_2	P_3
TBP/DNA	45	61		53	59	
TBP/DNA/TFIIA	35	63	59	49	60	72
TBP/DNA/Mot1	41	59	47	54	63	48
+ATP	38	63	58	51	61	68

However, the model overestimates the FRET values of P_2 and underestimates those of P_1. This is mainly due to the fact that P_1 is populated less often as predicted by the model, which can have several explanations. One possible reason is the limited accuracy in determining the relative occupancy of the FRET states due to uncertainties in determination of the rates. The divergence between modeled and measured FRET is larger when the difference of the back and forward rates are larger. The precision in determine the rates is limited by the ability of the HMM to find the correct transitions. The bigger the difference in the rates $S_i \rightarrow S_j$ and $S_j \rightarrow S_i$, the larger will be the influence of precision of the rates on the calculation of the relative occupancy. Another possible reason for the differences between the prediction of the model and the measured data could be that there are more states involved in the system than accounted for in the model. These additional states could have the same or similar FRET values which are not distinguishable. The transitions from state S_3 to state S_2 (see Appendix A.6) for all populations deviate from a single exponential decay which could suggest that state S_3 consists of two substates, which are not distinguishable by FRET alone. The complexity of the model was limited to 6 FRET states with 8 rates (only counting populations P_1 and P_2) because of the number of unknown parameters involved in the TBP binding process. Four color FRET experiments with labeled DNA, TBP and Mot1 will provide additional information about the conformation and composition of the complexes and can be used to identify

subpopulations and to refine the model.

A comparison of the single-molecule FRET values to FRET values obtained from ensemble measurements is difficult because of the different measuring conditions and the lack of γ-factor correction for individual molecules. However, the average FRET values of all dynamic molecules can be calculated and are given in Table 5.13.

Table 5.13.: Mean FRET values for all dynamic molecules measured from different complexes. The average is calculated once with γ correction and once with setting the γ-factor to 1. The values are given in % of FRET.

	γ-corrected	$\gamma = 1$
TBP/DNA	51	52
TBP/DNA/TFIIA	47	59
TBP/DNA/Mot1	55	47
+ATP	58	47

A relative decrease in the mean FRET efficiency in the case of Mot1 preincubation compared to TBP/DNA alone is seen when the γ-factor is not corrected. This suggests that the decrease is due to changes in the physical properties of the dye and not due to an actual change in the FRET efficiency. The average FRET efficiency after ATP addition to the TBP/DNA/Mot1 complex increases mainly due to the appearance of the population P_3 (the mean FRET efficiency of the population P_1 is decreased relative to TBP/DNA and TBP/Mot1 after ATP addition).

It is important to note that the model is based on a minimal, three-step process and a complete complex formation is assumed. Even in the case of partial complex formation, this model provides a basis for determining changes in rates and serves as a framework to understand the function of Mot1. The particular measuring conditions provided by single-molecule experiments, the absence of free Mot1 in solution and the possibility to separate individual populations, gives specific and novel insight into the mechanism of Mot1 ATPase activity.

6. Summary and Outlook

To understand the complexity of biological processes, single-molecule methods and nano-biotechnological probes are required to characterize interactions on a molecular scale. In this work, Total Internal Reflection (TIRF) Microscopy methods were implemented, optimized and used to study protein-DNA interactions and to characterize nanodevices.

In the first project, the interactions of TBP, DNA and Mot1 were monitored by single-pair FRET (spFRET). The dynamics of the FRET signal observed by single-molecule experiments contains information about different conformations and the transitions between them. Using a HMM approach to analyze the dynamics of TBP binding that induces DNA bending, populations with different binding orientations of TBP to the TATA-box could be identified. Each TBP binding orientation is associated with three highly dynamic DNA conformations. The effect of Mot1's ATPase activity on the TBP/DNA complex is strongly specific to the TBP binding orientation. In the inverted binding orientation, Mot1 promotes a linearization of the TATA-box containing DNA by increasing the rate of transition to the unbent conformation. Interestingly, the conformations (as identified by their FRET values) do not change, only the rates of the transitions between them changes. The observation of an ATP dependent change in the dynamics of the TBP/DNA complex without liberation of TBP suggests the formation TBP/DNA complex "primed" for dissociation. In contrast to the observation in ensemble measurements, where Mot1 completely liberates TBP from DNA, no dissociation was observed in the single-molecule experiments. Only when free Mot1 in nM concentrations was available in solution with the ATP during aditition, could TBP dissociation be observed. These findings are in perfect agreement with the model proposed in this work in which Mot1's ATP activity induces a primed TBP/DNA conformation which requires the interaction with an additional Mot1 molecule for TBP liberation. This observation serves as a basis to study the nature of the interaction of the primed TBP/DNA complex with additional molecules such as NC2 [176, 177, 178]. While the role of Mot1 in gene repression is readily explained by the liberation

of TBP from the TATA-box, the specificity of Mot1's activity to the binding orientation of TBP provides a model to understand the role of Mot1 in gene up-regulation. The findings in this work support the hypothesis that displacement of inverted bound TBP is one function of Mot1 in gene activation *in vivo* by promoting the PIC formation through liberation of transcriptionally inactive bound TBP [175]. To further expand the methods described in this work and to gather knowledge about the molecular basis of life, future developments are required. One possibility for increasing the information obtained from single-molecule experiments is to apply a combination of multi-parameter fluorescence techniques (e.g. 4-color TIRF, MFD) and increased positioning/analysis accuracy (e.g. NPS) with information theory based analysis (HMM).

In the second project, DNA was used as a scaffold to position different fluorophores with nanometer accuracy. A photo-switchable nanodevice was created that could be tracked in live cells. Dronpa, a photo-switchable fluorescent protein, was positioned within FRET-distance of a synthetic dye. By optically switching Dronpa between a fluorescent darkstate and a brightstate, the energy transfer (FRET) between the protein and the dye can be switched on and off. The physical characteristics of DNA were used to build optically addressable nanodevices which have dimensions ideal for probing cellular processes. The implementation of OLID could be used to simultaneously track (with enhanced contrast) and probe (e.g. pH) labeled proteins in live cells. With the development of new genetically encoded fluorescent proteins, this track-and-probe mechanism could be expanded to measure additional parameters in live-cells. Additionally, the use of OLID in combination with self assembled nanodevices could lead the way to optical read-out and active control of next-generation nano-machinery.

Appendix A.

Appendices

A.1. Protein labeling

For the experiments presented here, we used TBP from yeast with Cys - Ala mutations at positions 78 and 164 and a Ser - Cys mutation at position 61, kindly provided by Tony Weil and Professor Jens Michaelis. The labeling was performed as described in [18] and following the instructions of the dye manufacturer.

A.2. Bending angle calculation

Using a symmetric two-kink model shown in Figure A.1, the FRET values can be used to calculate the bending angles [172, 143].
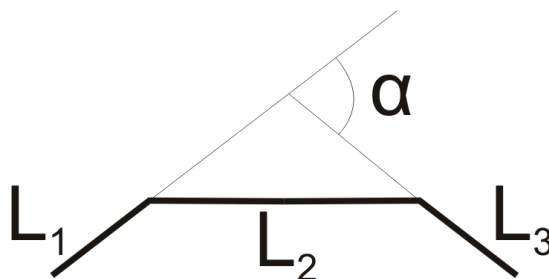
Figure A.1.: Schematics of two-kink model for DNA bending

The length of $L_2 = 20.4\,\text{nm}$ is given by the number of base pairs of the TATA-box. The distances between the donor and acceptor molecule obtained by FRET are R_{free} for the free unbent DNA and R_{bound} for the DNA bound and bent by TBP.

$$\alpha = 2 * cos^{-}1\frac{R_{bound} - L_2}{R_{free} - L_2} \tag{A.1}$$

A.3. DNA Sequences

The sequence of the DNA used to measure the conformation of the binary TBP/DNA complex is:

```
Biotin - 5' CTTCACCTTA TTTGCATAAG CGATTCTATA TAAAAGCGCC
XTGTCATACC CTGCTCACGC TGTTTTTCCT TTTCGTTGGC
```

where X labels the position of the fluorophore.

The sequence of the double-labeled DNA used to measure the conformation of the DNA is:

```
Biotin - 5' CTTCACCTT ATTTGCATAA GCGATTCTAT ATAAAAGCGC
CXCGCTCTA CCCTGCTCAC GCTGTTTTTC CTTTTCGTTGGC -3'
```

with the complimentary strand:

```
3' - GAAGTGGAAT AAACGTATTC GCYAAGATAT ATTTTCGCGG
AGCGAGATGG GACGAGTGC GACAAAAAGG AAAAGCAACCG-5'
```

where X and Y labels the position of the fluorophores.

A.4. Influence of FRET on photo-bleaching

The influence of FRET on the Time Before Bleaching (TBB) is shown in Figure A.2 and can be corrected by accounting for the FRET value of every molecule. Two populations of double-labeled molecules, with different donor-acceptor separations, were mixed and surface-immobilized.
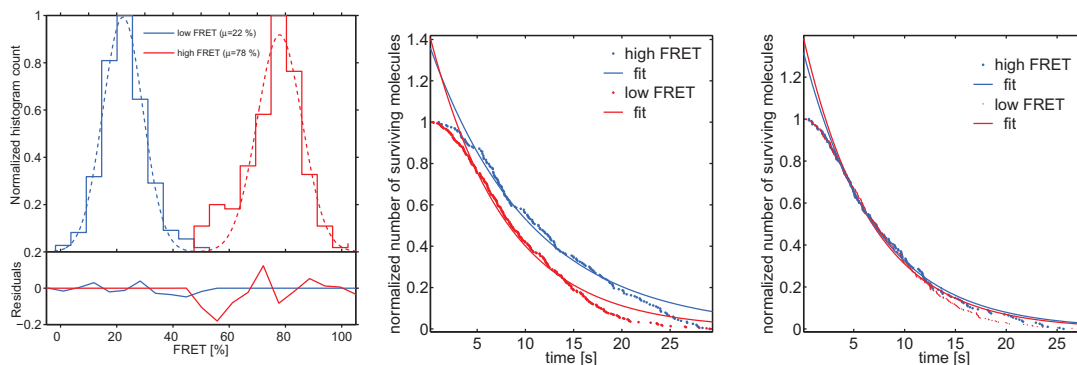
Figure A.2.: The influence of FRET on the time before bleaching. Left: Molecular wise FRET histogram. Individual molecules are separated, based on their average FRET value, into high- and low-FRET population shown in blue and red, respectively. Center: Cumulative dwell-time histogram for both populations with rates of $k = 0.095s^{-}1$ and $k = 0.127s^{-}1$ for the high- and low-FRET population. Right: Cumulative dwell-time histogram for both populations after correction for FRET with rates of $k = 0.141s^{-}1$ and $k = 0.139s^{-}1$ for the high- and low-FRET populations.

The time before bleaching τ_0 in the absence of a FRET acceptor can be calculated from the average FRET value E_{avg} and the measured time before bleaching τ_E in the presence of the acceptor is given by:

$$\tau_0 = \tau_E \left(1 - E_{avg})\right) \tag{A.2}$$
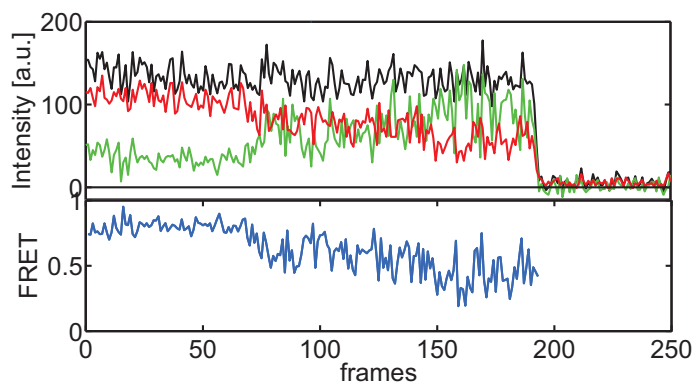
Figure A.3.: Intensity and FRET trace for a complex with labeled TBP and labeled DNA. The FRET values shown in blue change continuously to lower values.

A.5. Trace Intensity Analysis

The *Trace Intensity Analysis* (TRACY) toolbox was programmed in MATLAB and used to perform following tasks:

1. Extract single-molecule intensity trajectories (ALEX, TIRF, single color,...) from image stacks

2. Selection of time window with meaningful fluorescence intensities for each molecule

3. Fluorescence intensity corrections

4. Identification of molecular subpopulation and statistical analysis

All functions described in this section can be accessed through the graphical user interface GUI in the *TRACY* program (see screenshot shown in Figure A.4). Each molecule can be "tagged" with a user defined property. The list of "tags" can be user modified by the "ADD" and "DELETE" button. In the main window the trace intensities and the calculated FRET values are displayed in two separat windows. The control over which information to display (e.g. background, total intensity, HMM) is located below the tag list. The FRET histogram of the current molecule is displayed in the frame-wise histogram in the lower left corner.

The analysis of the intensity and FRET traces is performed in the *Trace Histogram Analysis* (THA) GUI. Population of molecules can be selected based on their tags and imported into THA. Display options include frame-wise or molecule-wise FRET histograms, cumulative or regular dwell-time histograms and intensity based histograms as shown in Figure A.6. Each histogram can be fitted with either mono-, bi- or trippel- exponential or Gaussian distribution. The fit can be applied to each data set individually or user-defined models can be fitted globally across datasets as shown in Figure A.5.
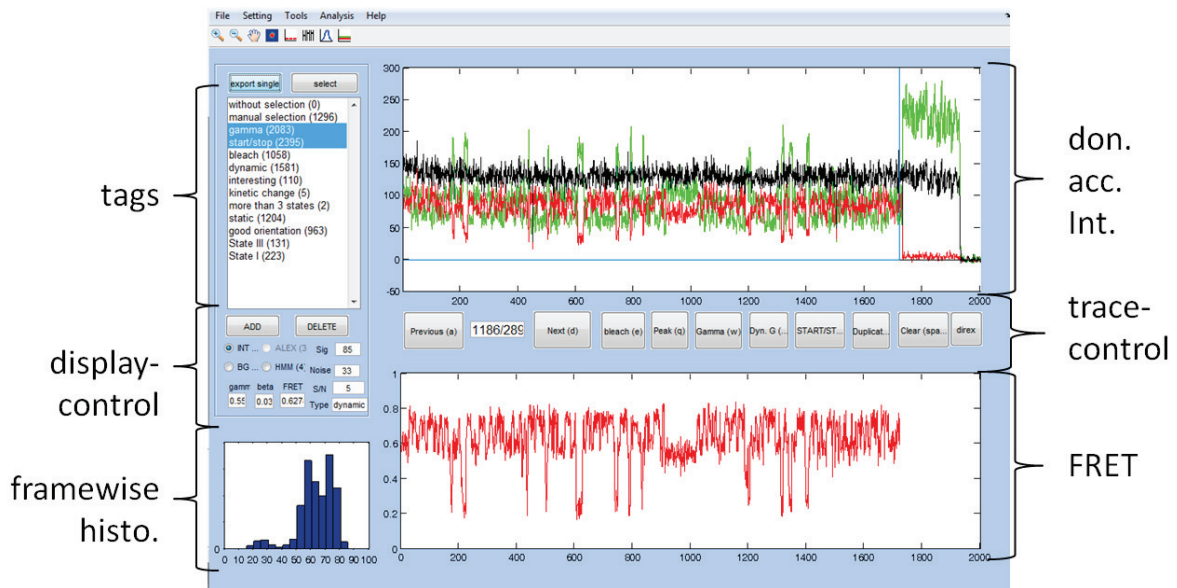
Figure A.4.: Screenshot of the main GUI of TRACY.

Figure A.5.: Screenshot of the global fitting function within the *Trace Histogram Analysis* GUI in TRACY for a 4 well model for two data sets. The parameters of individual datasets are listed as rows and the parameters as columns. Each parameter can be fixed to an initial value of set to be optimized globally.
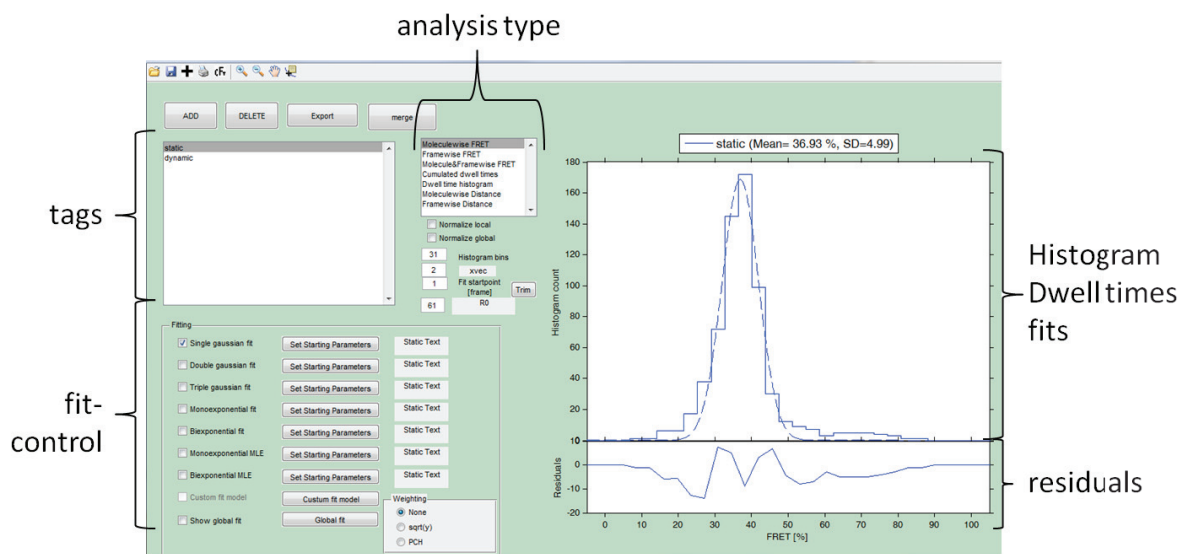
Figure A.6.: Screenshot of the *Trace Histogram Analysis* GUI in TRACY.

A.6. TDP kinetic rates fits

The dwell time histograms of the individual states Si before the transitions into another state Sj are shown in this Section. The forward and backward transitions between two states are show in one Figure. The normalized decays are fitted with the mono-exponential function $f(x) = e^{(-kt)}$.
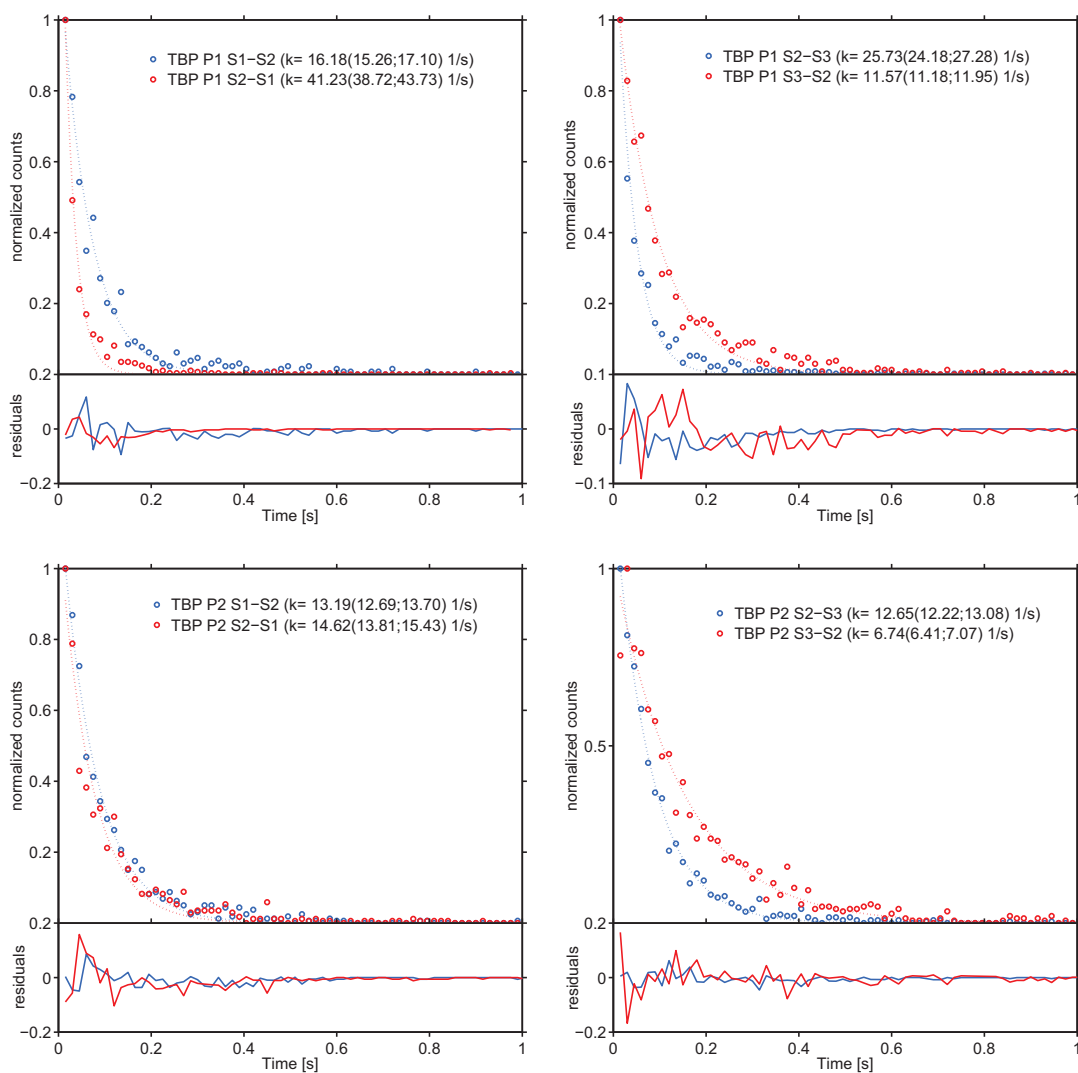
Figure A.7.: Dwell times and mono-exponential fits for the transitions of the TBP/DNA complex. The fitting parameters and the 95% confidence intervals are given in the figure legend

Figure A.8.: Dwell times and mono-exponential fits for the transitions of the TBP/DNA/TFIIA complex. The fitting parameters and the 95% confidence intervals are given in the figure legend

Figure A.9.: Dwell times and mono-exponential fits for the transitions of the TBP/DNA/Mot1 complex. The fitting parameters and the 95% confidence intervals are given in the figure legend
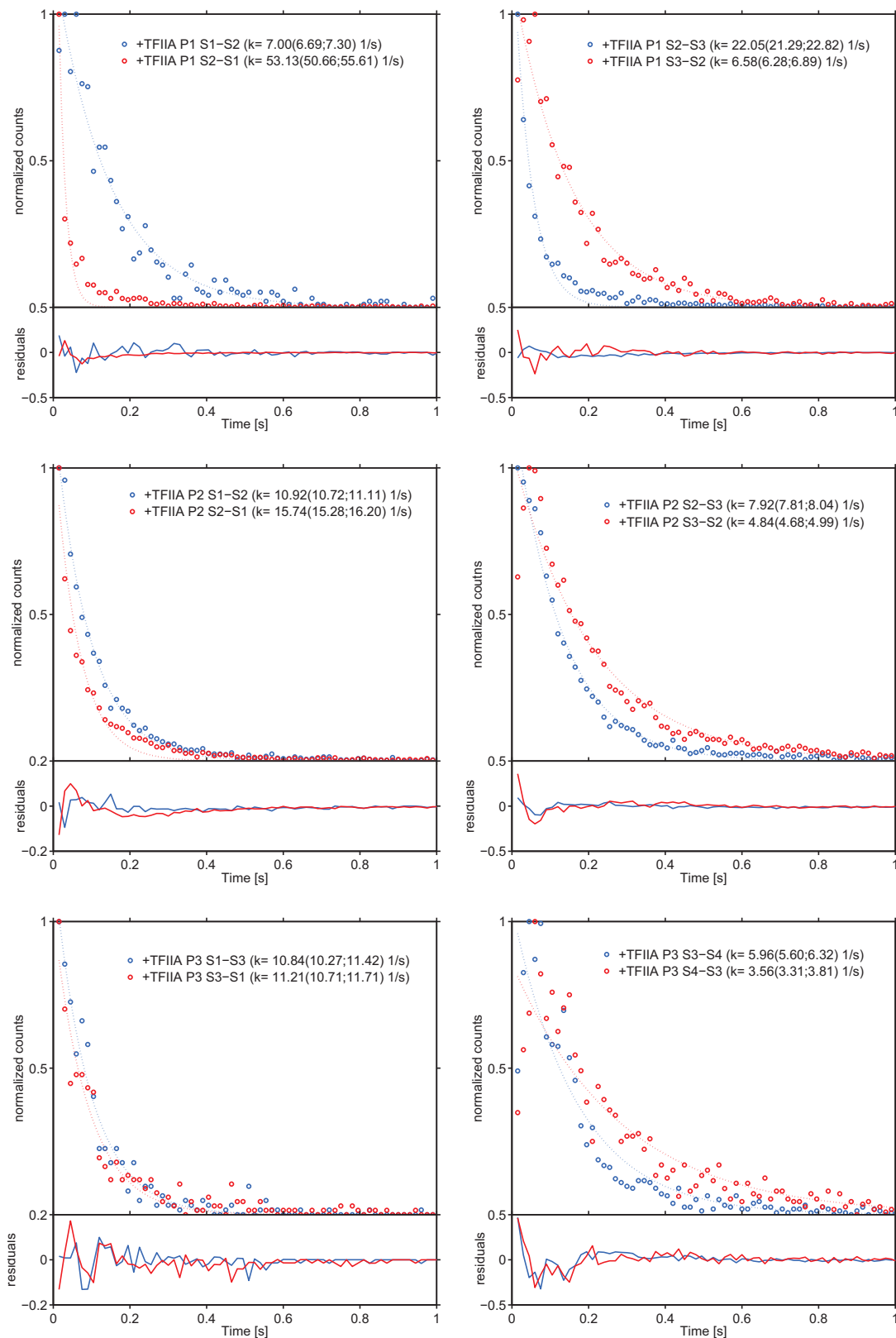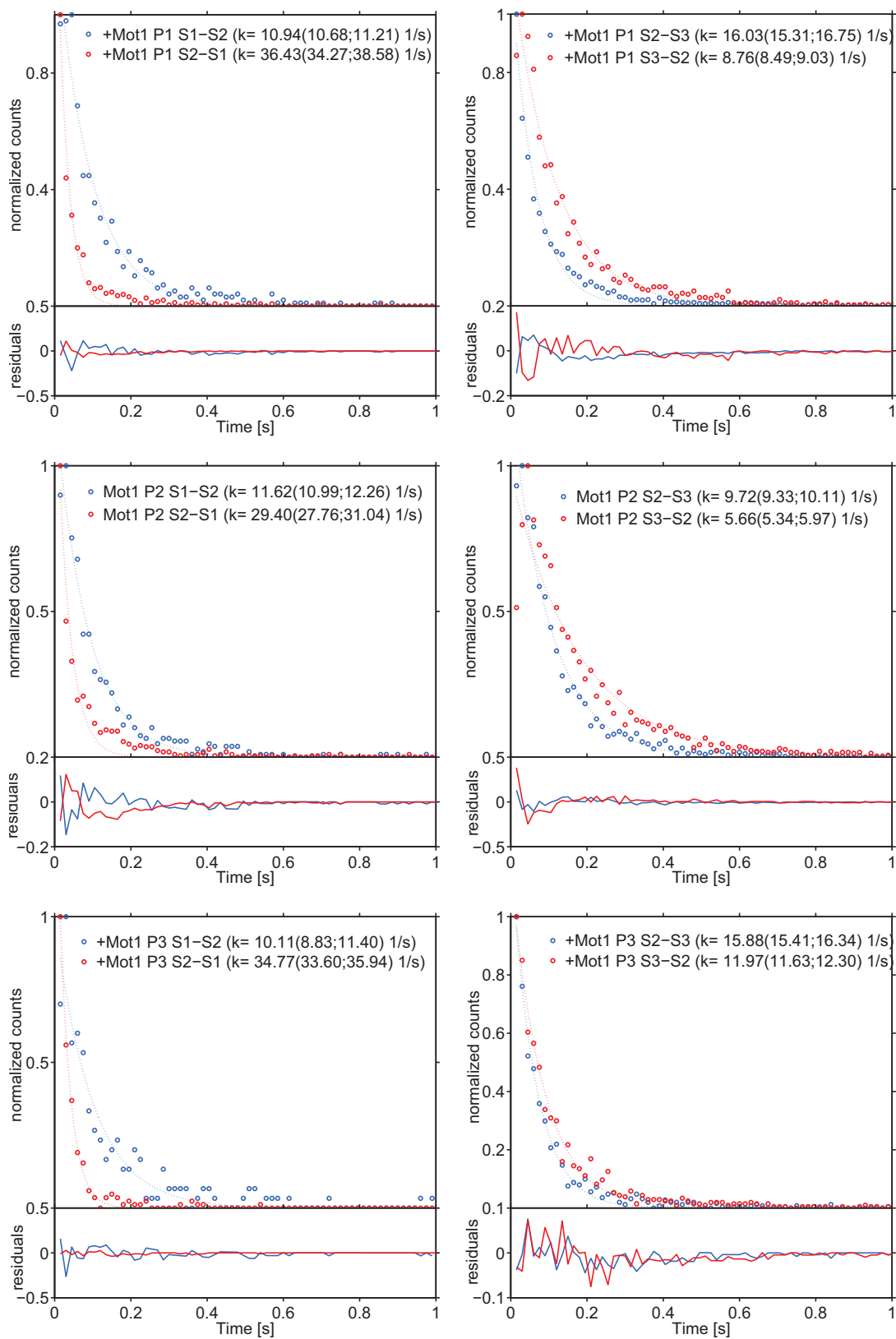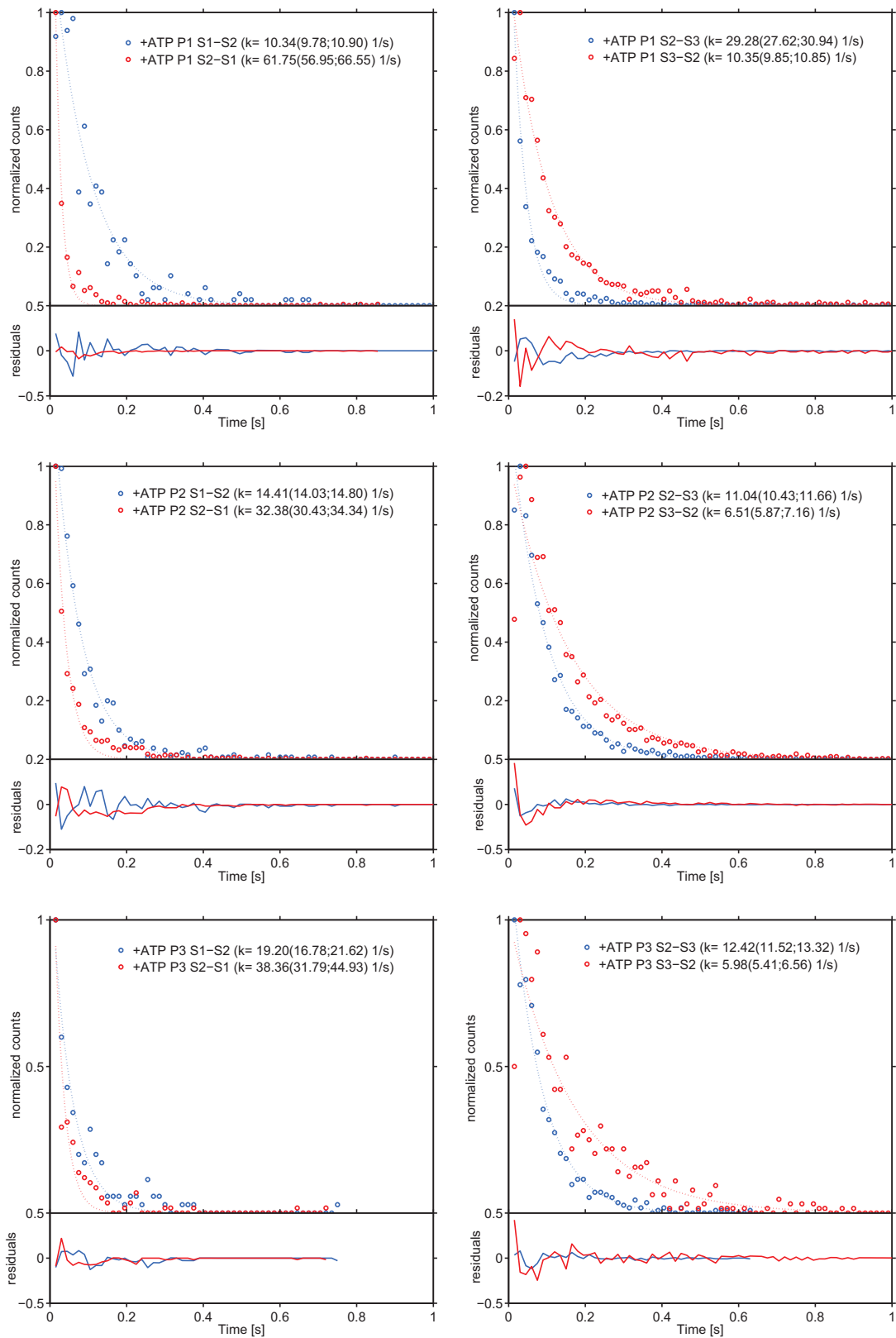
Figure A.10.: Dwell times and mono-exponential fits for the transitions of the TBP/DNA/Mot1 complex after ATP addition. The fitting parameters and the 95% confidence intervals are given in the figure legend

Bibliography

[1] M. Kasha. Characterization of electronic transitions in complex molecules. *Discussions of the Faraday Society*, 9:14–19, 1950.

[2] G.G Stokes. On the change of refrangibility of light. *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences*, II:463–562, 1852.

[3] B. R. Henry and M. Kasha. Radiationless molecular electronic transitions. *Annual Review of Physical Chemistry*, 19:161, 1968.

[4] Perrin J. Fluorescence et induction moléculaire par résonance. *Comptes rendus*, 184:1097–1100, 1927.

[5] T. Förster. Energiewanderung und fluoreszenz. *Naturwissenschaften*, 33(6):166–175, 1946.

[6] T. Förster. Zwischenmolekulare energiewanderung und fluoreszenz. *Annalen Der Physik*, 2(1-2):55–75, 1948.

[7] R. M. Clegg. Fluorescence resonance energy-transfer and nucleic-acids. *Methods in Enzymology*, 211:353–388, 1992.

[8] R. E. Dale, J. Eisinger, and W. E. Blumberg. The orientational freedom of molecular probes. the orientation factor in intramolecular energy transfer. *Biophysical Journal*, 26(2):161–93, 1979.

[9] D. Axelrod. Cell-substrate contacts illuminated by total internal-reflection fluorescence. *Journal of Cell Biology*, 89(1):141–145, 1981.

[10] T. Funatsu, Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida. Imaging of single fluorescent molecules and individual atp turnovers by single myosin molecules in aqueous-solution. *Nature*, 374(6522):555–559, 1995.

[11] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss. Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences of the United States of America*, 93(13):6264–6268, 1996.

[12] R. H. Webb. Confocal optical microscopy. *Reports on Progress in Physics*, 59(3):427–471, 1996.

[13] Roshdi Rashed. A pioneer in anaclastics: Ibn sahl on burning mirrors and lenses. *The University of Chicago Press*, 81:464–491, 1990.

[14] D. Axelrod, S. E. Sund, and J. A. Swanson. Cell membrane orientation visualized by polarized total internal reflection fluorescence. *Biophysical Journal*, 77(4):2266–2283, 1999.

[15] J. Mertz. Radiative absorption, fluorescence, and scattering of a classical dipole near a lossless interface: a unified description. *Journal of the Optical Society of America B-Optical Physics*, 17(11):1906–1913, 2000.

[16] M. Oheim. Imaging transmitter release. ii. a practical guide to evanescent-wave imaging. *Lasers in Medical Science*, 16(3):159–170, 2001.

[17] John Correia. Biophysical tools for biologists. *Methods in Cell Biology*, 89:170, 2008.

[18] Peter Schlüsche. *Untersuchung der Funktion und Dynamik von DNA-Transkriptionsfaktoren mittels Einzelmolekül-Fluoreszenzmikroskopie*. PhD thesis, 2007.

[19] M. F. Tompsett, G. F. Amelio, W. J. Bertram, R. R. Buckley, W. J. Mcnamara, Mikkelse.Jc, and D. A. Sealer. Charge-coupled imaging devices - experimental results. *Ieee Transactions on Electron Devices*, Ed18(11):992, 1971.

[20] W. S. Boyle and G. E. Smith. Charge coupled semiconductor devices. *Bell System Technical Journal*, 49(4):587, 1970.

[21] S. K. Madan, B. Bhaumik, and J. M. Vasi. Experimental-observation of avalanche multiplication in charge-coupled-devices. *Ieee Transactions on Electron Devices*, 30(6):694–699, 1983.

[22] M. S. Robbins and B. J. Hadwen. The noise performance of electron multiplying charge-coupled devices. *Ieee Transactions on Electron Devices*, 50(5):1227–1232, 2003.

[23] O. Daigle, J. L. Gach, C. Guillaume, C. Carignan, P. Balard, and O. Boissin. L3ccd results in pure photon counting mode. *Optical and Infrared Detectors for Astronomy*, 5499:219–227 556, 2004.

[24] James W. Cooley; John W. Tukey. An algorithm for the machine calculation of complex fourier series. *Math. Comp.*, 19:297–301, 1965.

[25] L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state markov chains. *Annals of Mathematical Statistics*, 37(6):1554–&, 1966.

[26] J. Raviv. Decision making in markov chains applied to problem of pattern recognition. *Ieee Transactions on Information Theory*, 13(4):536, 1967.

[27] F. Jelinek. Continuous speech recognition by statistical-methods. *Proceedings of the Ieee*, 64(4):532–556, 1976.

[28] H. Bourlard and N. Morgan. Continuous speech recognition by connectionist statistical-methods. *IEEE Transactions on Neural Networks*, 4(6):893–909, 1993.

[29] D. A. Coast, G. G. Cano, and S. A. Briller. Use of hidden markov-models for electrocardiographic signal analysis. *Journal of Electrocardiology*, 23:184–191, 1991.

[30] F. G. Ball and J. A. Rice. Stochastic-models for ion channels - introduction and bibliography. *Mathematical Biosciences*, 112(2):189–206, 1992.

[31] G. A. Churchill. Stochastic-models for heterogeneous dna-sequences. *Bulletin of Mathematical Biology*, 51(1):79–94, 1989.

[32] Y. Ephraim and N. Merhav. Hidden markov processes. *Ieee Transactions on Information Theory*, 48(6):1518–1569, 2002.

[33] M. Andrec, R. M. Levy, and D. S. Talaga. Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden markov models. *Journal of Physical Chemistry A*, 107(38):7454–7464, 2003.

[34] G. F. Schroder and H. Grubmuller. Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. *Journal of Chemical Physics*, 119(18):9920–9924, 2003.

[35] D. S. Talaga, T. C. Messina, H. Y. Kim, and J. T. Giurleo. Hidden markov model analysis of multichromophore photobleaching. *Journal of Physical Chemistry B*, 110(33):16366–16376, 2006.

[36] T. Ha, C. Joo, S. A. McKinney, M. Nakamura, I. Rasnik, and S. Myong. Real-time observation of reca filament dynamics with single monomer resolution. *Cell*, 126(3):515–527, 2006.

[37] D. L. Ensign and V. S. Pande. Bayesian detection of intensity changes in single molecule and molecular dynamics trajectories. *The journal of physical chemistry. B*, 114(1):280–92, 2010.

[38] L. R. Rabiner. A tutorial on hidden markov-models and selected applications in speech recognition. *Proceedings of the Ieee*, 77(2):257–286, 1989.

[39] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in statistical analysis of probabilistic functions of markov chains. *Annals of Mathematical Statistics*, 41(1):164–&, 1970.

[40] L. E. Baum and J. A. Eagon. An inequality with applications to statistical estimation for probabilistic functions of markov processes and to a model

for ecology. *Bulletin of the American Mathematical Society*, 73(3):360–&, 1967.

[41] Jeff A. Bilmes. A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *International Computer Science Institute*, Berkeley CA, 1998.

[42] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via em algorithm. *Journal of the Royal Statistical Society Series B-Methodological*, 39(1):1–38, 1977.

[43] Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, 13(2):260, 1967.

[44] S. Kalinin, S. Felekyan, M. Antonik, and C. A. Seidel. Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer. *The journal of physical chemistry. B*, 111(34):10253–62, 2007.

[45] Schwarz G. Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464, 1987.

[46] Nawid Zarrabi. *Hidden Markov Modelle fuer Einzelmoleküldaten*. PhD thesis, Fakultät Mathematik und Physik der Universität Stuttgart, 2010.

[47] P. G. Bruce, B. Scrosati, and J. M. Tarascon. Nanomaterials for rechargeable lithium batteries. *Angewandte Chemie*, 47(16):2930–46, 2008.

[48] O. Salata. Applications of nanoparticles in biology and medicine. *Journal of nanobiotechnology*, 2(1):3, 2004.

[49] N. O'Farrell, A. Houlton, and B. R. Horrocks. Silicon nanoparticles: applications in cell biology and medicine. *International journal of nanomedicine*, 1(4):451–72, 2006.

[50] A. K. Geim and K. S. Novoselov. The rise of graphene. *Nature materials*, 6(3):183–91, 2007.

[51] B. C. Edwards. The space elevator: a new tool for space studies. *Gravitational and space biology bulletin : publication of the American Society for Gravitational and Space Biology*, 16(2):101–5, 2003.

[52] R. P Feynman. There's plenty of room at the bottom. *Engineering and Science, Caltech*, 23(5):22–36, 1960.

[53] D. M. Eigler and E. K. Schweizer. Positioning single atoms with a scanning tunneling microscope. *Nature*, 344(6266):524–526, 1990.

[54] T. Junno, K. Deppert, L. Montelius, and L. Samuelson. Controlled manipulation of nanoparticles with an atomic-force microscope. *Applied Physics Letters*, 66(26):3627–3629, 1995.

[55] M. Pumera. Carbon nanotube biosensors based on electrochemical detection. *Methods in molecular biology*, 625:205–12, 2010.

[56] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, and S. Weiss. Quantum dots for live cells, in vivo imaging, and diagnostics. *Science*, 307(5709):538–44, 2005.

[57] P. W. K. Rothemund. Folding dna to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, 2006.

[58] A. J. Turberfield and J. Bath. Dna nanomachines. *Nature Nanotechnology*, 2(5):275–284, 2007.

[59] D. S. Liu, Y. Yang, and H. J. Liu. Dna based nanomachines. *Progress in Chemistry*, 20(2-3):197–207, 2008.

[60] Y. Krishnan, S. Modi, M. G. Swetha, D. Goswami, G. D. Gupta, and S. Mayor. A dna nanomachine that maps spatial and temporal ph changes inside living cells. *Nature Nanotechnology*, 4(5):325–330, 2009.

[61] P. Tinnefeld, I. H. Stein, V. Schuller, P. Bohm, and T. Liedl. Single-molecule fret ruler based on rigid dna origami blocks. *Chemphyschem*, 12(3):689–695, 2011.

[62] R. Y. Tsien, R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, and D. A. Zacharias. A monomeric red fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7877–7882, 2002.

[63] R. Ando, H. Mizuno, and A. Miyawaki. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science*, 306(5700):1370–3, 2004.

[64] M. Andresen, A. C. Stiel, S. Trowitzsch, G. Weber, C. Eggeling, M. C. Wahl, S. W. Hell, and S. Jakobs. Structural basis for reversible photoswitching in dronpa. *Proceedings of the National Academy of Sciences of the United States of America*, 104(32):13005–9, 2007.

[65] P. Dedecker, J. Hotta, R. Ando, A. Miyawaki, Y. Engelborghs, and J. Hofkens. Fast and reversible photoswitching of the fluorescent protein dronpa as evidenced by fluorescence correlation spectroscopy. *Biophysical Journal*, 91(5):L45–L47, 2006.

[66] S. Habuchi, R. Ando, P. Dedecker, W. Verheijen, H. Mizuno, A. Miyawaki, and J. Hofkens. Reversible single-molecule photoswitching in the gfp-like fluorescent protein dronpa. *Proc Natl Acad Sci U S A*, 102(27):9511–6, 2005.

[67] V. Voliani, R. Bizzarri, R. Nifosi, S. Abbruzzetti, E. Grandi, C. Viappiani, and F. Beltram. Cis-trans photoisomerization of fluorescent-protein chromophores. *Journal of Physical Chemistry B*, 112(34):10714–10722, 2008.

[68] M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, and S. J. Remington. Crystal structure of the aequorea victoria green fluorescent protein. *Science*, 273(5280):1392–5, 1996.

[69] P. G. Wilmann, K. Turcic, J. M. Battad, M. C. Wilce, R. J. Devenish, M. Prescott, and J. Rossjohn. The 1.7 a crystal structure of dronpa: a photoswitchable green fluorescent protein. *J Mol Biol*, 364(2):213–24, 2006.

[70] M. Zimmer. Green fluorescent protein (gfp): applications, structure, and related photophysical behavior. *Chem Rev*, 102(3):759–81, 2002.

[71] R. Heim, A. B. Cubitt, and R. Y. Tsien. Improved green fluorescence. *Nature*, 373(6516):663–4, 1995.

[72] R. Heim and R. Y. Tsien. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current biology : CB*, 6(2):178–82, 1996.

[73] A. C. Stiel, S. Trowitzsch, G. Weber, M. Andresen, C. Eggeling, S. W. Hell, S. Jakobs, and M. C. Wahl. 1.8 a bright-state structure of the reversibly switchable fluorescent protein dronpa guides the generation of fast switching variants. *The Biochemical journal*, 402(1):35–42, 2007.

[74] C. M. Niemeyer, F. Kukolka, B. K. Muller, S. Paternoster, A. Arndt, C. Brauchle, and D. C. Lamb. A single-molecule forster resonance energy transfer analysis of fluorescent dna-protein conjugates for nanobiotechnology. *Small*, 2(8-9):1083–1089, 2006.

[75] C. M. Niemeyer, V. Lapienc, F. Kukolka, K. Kiko, and A. Arndt. Conjugation of fluorescent proteins with dna oligonucleotides. *Bioconjugate Chemistry*, 21(5):921–927, 2010.

[76] C. Flors, E. Fron, G. Schweitzer, S. Habuchi, H. Mizuno, R. Ando, F. C. De Schryver, A. Miyawaki, and J. Hofkens. Ultrafast excited-state dynamics of the photoswitchable protein dronpa. *Journal of the American Chemical Society*, 129(16):4870–+, 2007.

[77] I. Gopich and A. Szabo. Theory of photon statistics in single-molecule forster resonance energy transfer. *Journal of Chemical Physics*, 122(1):–, 2005.

[78] R. Y. Tsien. The green fluorescent protein. *Annual Review of Biochemistry*, 67:509–44, 1998.

[79] G. Marriott, S. Mao, R. K. P. Benninger, Y. L. Yan, C. Petchprayoon, D. Jackson, C. J. Easley, and D. W. Piston. Optical lock-in detection of fret using synthetic and genetically encoded optical switches. *Biophysical Journal*, 94(11):4515–4524, 2008.

[80] R. Ongjoco, K. Szkutnicka, and V. P. Cirillo. Glucose-transport in vesicles reconstituted from saccharomyces-cerevisiae membranes and liposomes. *Journal of Bacteriology*, 169(7):2926–2931, 1987.

[81] J. E. Whitaker, R. P. Haugland, and F. G. Prendergast. Spectral and photophysical studies of benzo[c]xanthene dyes - dual emission ph sensors. *Analytical Biochemistry*, 194(2):330–344, 1991.

[82] M. Kneen, J. Farinas, Y. X. Li, and A. S. Verkman. Green fluorescent protein as a non-invasive intracellular ph indicator. *Biophysical Journal*, 74(2):A181–A181, 1998.

[83] G. Miesenbock, D. A. De Angelis, and J. E. Rothman. Visualizing secretion and synaptic transmission with ph-sensitive green fluorescent proteins. *Nature*, 394(6689):192–5, 1998.

[84] X. H. Shi, T. B. McAnaney, P. Abbyad, H. Jung, S. J. Remington, and S. G. Boxer. Green fluorescent protein variants as ratiometric dual emission ph sensors: Temperature dependence of proton transfer. *Biophysical Journal*, 88(1):161A–161A, 2005.

[85] Mark Bradley, Lois Alexander, Karen Duncan, Mourad Chennaoui, and Anita C. Jones. ph sensing in living cells using fluorescent microspheres. *Bioorganic & Medicinal Chemistry Letters*, 18(1):313–317, 2008.

[86] S. M. Borisov, D. L. Herrod, and I. Klimant. Fluorescent poly(styrene-block-vinylpyrrolidone) nanobeads for optical sensing of ph. *Sensors and Actuators B-Chemical*, 139(1):52–58, 2009.

[87] Y. A. Povrozin, L. I. Markova, A. L. Tatarets, V. I. Sidorov, E. A. Terpetschnig, and L. D. Patsenker. Near-infrared, dual-ratiometric fluores-

cent label for measurement of ph. *Analytical Biochemistry*, 390(2):136–140, 2009.

[88] Anja Schulz, Stephanie Hornig, Tim Liebert, Eckhard Birckner, Thomas Heinze, and Gerhard J. Mohr. Evaluation of fluorescent polysaccharide nanoparticles for ph-sensing. *Org Biomol Chem*, 7(9):1884–9, 2009.

[89] C. Hille, M. Berg, L. Bressel, D. Munzke, P. Primus, H. G. Lohmannsroben, and C. Dosche. Time-domain fluorescence lifetime imaging for intracellular ph sensing in living tissues. *Analytical and Bioanalytical Chemistry*, 391(5):1871–1879, 2008.

[90] T. Nakabayashi and N. Ohta. Studies on microenvironment in a single cell using fluorescence lifetime imaging microscopy. *Bunseki Kagaku*, 58(6):473–485, 2009.

[91] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3:318–56, 1961.

[92] M. E. Greenberg, A. Brunet, L. B. Sweeney, J. F. Sturgill, K. F. Chua, P. L. Greer, Y. X. Lin, H. Tran, S. E. Ross, R. Mostoslavsky, H. Y. Cohen, L. S. Hu, H. L. Cheng, M. P. Jedrychowski, S. P. Gygi, D. A. Sinclair, and F. W. Alt. Stress-dependent regulation of foxo transcription factors by the sirt1 deacetylase. *Science*, 303(5666):2011–2015, 2004.

[93] R. Jaenisch and R. Young. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell*, 132(4):567–582, 2008.

[94] D. L. Johnson, S. A. S. Johnson, L. Dubeau, M. Kawalek, A. Dervan, A. H. Schonthal, and C. V. Dang. Increased expression of tata-binding protein, the central transcription factor, can contribute to oncogenesis. *Molecular and Cellular Biology*, 23(9):3043–3051, 2003.

[95] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander. Molecular classification of cancer: Class

discovery and class prediction by gene expression monitoring. *Science*, 286(5439):531–537, 1999.

[96] K. Chen and N. Rajewsky. The evolution of gene regulation by transcription factors and micrornas. *Nature reviews. Genetics*, 8(2):93–103, 2007.

[97] S. L. Berger. The complex language of chromatin regulation during transcription. *Nature*, 447(7143):407–12, 2007.

[98] S. P. Baker and P. A. Grant. The saga continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene*, 26(37):5329–40, 2007.

[99] S. R. Bhaumik and M. R. Green. Differential requirement of saga components for recruitment of tata-box-binding protein to promoters in vivo. *Molecular and Cellular Biology*, 22(21):7365–7371, 2002.

[100] R. A. Coleman and B. F. Pugh. Evidence for functional binding and stable sliding of the tata-binding protein on nonspecific dna. *Journal of Biological Chemistry*, 270(23):13850–13859, 1995.

[101] R. A. Coleman, A. K. P. Taggart, S. Burma, J. J. Chicca, and B. F. Pugh. Tfiia regulates tbp and tfiid dimers. *Molecular Cell*, 4(3):451–457, 1999.

[102] G. Prelich. Saccharomyces cerevisiae bur6 encodes a drap1/nc2 alpha homolog that has both positive and negative roles in transcription in vivo. *Molecular and Cellular Biology*, 17(4):2057–2065, 1997.

[103] P. Schlüsche, G. Stelzer, E. Piaia, D. C. Lamb, and M. Meisterernst. Nc2 mobilizes tbp on core promoter tata boxes. *Nature Structural and Molecular Biology*, 14(12):1196–1201, 2007.

[104] O. Dor and Y. Zhou. Achieving 80structure prediction by large-scale training. *Proteins*, 66(4):838–45, 2007.

[105] M. van Heel, B. Gowen, R. Matadeen, E. V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan. Single-particle electron cryo-microscopy: towards atomic resolution. *Quarterly reviews of biophysics*, 33(4):307–69, 2000.

[106] A. Muschielok, J. Andrecka, A. Jawhari, F. Bruckner, P. Cramer, and J. Michaelis. A nano-positioning system for macromolecular structural analysis. *Nature methods*, 5(11):965–71, 2008.

[107] M. Margittai, J. Widengren, E. Schweinberger, G. F. Schroder, S. Felekyan, E. Haustein, M. Konig, D. Fasshauer, H. Grubmuller, R. Jahn, and C. A. Seidel. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15516–21, 2003.

[108] V. B. Bajic, S. L. Tan, A. Christoffels, C. Schonbach, L. Lipovich, L. Yang, O. Hofmann, A. Kruger, W. Hide, C. Kai, J. Kawai, D. A. Hume, P. Carninci, and Y. Hayashizaki. Mice and men: Their promoter properties. *Plos Genetics*, 2(4):614–626, 2006.

[109] R. P. Lifton, M. L. Goldberg, R. W. Karp, and D. S. Hogness. Organization of histone genes in drosophila-melanogaster - functional and evolutionary implications. *Cold Spring Harbor Symposia on Quantitative Biology*, 42:1047–1051, 1977.

[110] M. C. Thomas and C. M. Chiang. The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology*, 41(3):105–178, 2006.

[111] S. K. Burley and R. G. Roeder. Biochemistry and structural biology of transcription factor iid (tfiid). *Annual Review of Biochemistry*, 65:769–799, 1996.

[112] E. Martinez, Q. A. Zhou, N. D. Letoile, T. Oelgeschlager, A. J. Berk, and R. G. Roeder. Core promoter-specific function of a mutant transcription

factor tfiid defective in tata-box binding. *Proceedings of the National Academy of Sciences of the United States of America*, 92(25):11864–11868, 1995.

[113] B. F. Pugh and R. Tjian. Transcription from a tata-less promoter requires a multisubunit tfiid complex. *Genes & Development*, 5(11):1935–1945, 1991.

[114] T. Juven-Gershon, J. Y. Hsu, J. W. M. Theisen, and J. T. Kadonaga. The rna polymerase ii core promoter - the gateway to transcription. *Current Opinion in Cell Biology*, 20(3):253–259, 2008.

[115] J. J. Li, R. H. Kim, and J. Sodek. An inverted tata box directs downstream transcription of the bone sialoprotein gene. *Biochemical Journal*, 310:33–40, 1995.

[116] L. C. Xu, M. Thali, and W. Schaffner. Upstream box tata box order is the major determinant of the direction of transcription. *Nucleic Acids Research*, 19(24):6699–6704, 1991.

[117] M. Horikoshi, C. K. Wang, H. Fujii, J. A. Cromlish, P. A. Weil, and R. G. Roeder. Purification of a yeast tata box-binding protein that exhibits human transcription factor-iid activity. *Proceedings of the National Academy of Sciences of the United States of America*, 86(13):4843–4847, 1989.

[118] K. M. Campbell, R. T. Ranallo, L. A. Stargell, and K. J. Lumb. Reevaluation of transcriptional regulation by tata-binding protein oligomerization: Predominance of monomers. *Biochemistry*, 39(10):2633–2638, 2000.

[119] G. A. Patikoglou, J. L. Kim, L. P. Sun, S. H. Yang, T. Kodadek, and S. K. Burley. Tata element recognition by the tata box-binding protein has been conserved throughout evolution. *Genes & Development*, 13(24):3217–3230, 1999.

[120] T. Miyata, D. Hoshiyama, and K. Kuma. Extremely reduced evolutionary rate of tata-box binding protein in higher vertebrates and its evolutionary implications. *Gene*, 280(1-2):169–173, 2001.

[121] M. Ouhammouch, W. Hausner, and E. P. Geiduschek. Tbp domain symmetry in basal and activated archaeal transcription. *Molecular Microbiology*, 71(1):123–131, 2009.

[122] Y. C. Kim, J. H. Geiger, S. Hahn, and P. B. Sigler. Crystal-structure of a yeast tbp tata-box complex. *Nature*, 365(6446):512–520, 1993.

[123] A. J. Bonham, T. Neumann, M. Tirrell, and N. O. Reich. Tracking transcription factor complexes on dna using total internal reflectance fluorescence protein binding microarrays. *Nucleic Acids Research*, 37(13):–, 2009.

[124] S. Hahn, S. Buratowski, P. A. Sharp, and L. Guarente. Yeast tata-binding protein tfiid binds to tata elements with both consensus and nonconsensus dna-sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 86(15):5718–5722, 1989.

[125] P. Carninci, A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. M. Semple, M. S. Taylor, P. G. Engstrom, M. C. Frith, A. R. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume, and Y. Hayashizaki. Genome-wide analysis of mammalian promoter architecture and evolution (vol 38, pg 626, 2006). *Nature Genetics*, 39(9):1174–1174, 2007.

[126] A. J. Bonham, G. Braun, I. Pavel, M. Moskovits, and N. O. Reich. Detection of sequence-specific protein-dna interactions via surface enhanced

resonance raman scattering. *Journal of the American Chemical Society*, 129(47):14572–+, 2007.

[127] G. M. Perezhoward, P. A. Weil, and J. M. Beechem. Yeast tata-binding protein-interaction with dna - fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry*, 34(25):8005–8017, 1995.

[128] K. M. Masters, K. M. Parkhurst, M. A. Daugherty, and L. J. Parkhurst. Native human tata-binding protein simultaneously binds and bends promoter dna without a slow isomerization step or tfiib requirement. *Journal of Biological Chemistry*, 278(34):31685–31690, 2003.

[129] M. R. Green. Tbp-associated factors (taf(ii)s): multiple, selective transcriptional mediators in common complexes. *Trends in Biochemical Sciences*, 25(2):59–63, 2000.

[130] B. S. Wolner and J. D. Gralla. Tata-flanking sequences influence the rate and stability of tata-binding protein and tfiib binding. *Journal of Biological Chemistry*, 276(9):6260–6266, 2001.

[131] S. Lomvardas and D. Thanos. Nucleosome sliding via tbp dna binding in vivo. *Cell*, 106(6):685–696, 2001.

[132] R. O. Sprouse, T. S. Karpova, F. Mueller, A. Dasgupta, J. G. McNally, and D. T. Auble. Regulation of tata-binding protein dynamics in living yeast cells. *Proceedings of the National Academy of Sciences of the United States of America*, 105(36):13304–13308, 2008.

[133] Melissa N. Wells Kunal Poorey, Rebekka O. Sprouse. Rna synthesis precision is regulated by preinitiation complex turnover. *Genome Research*, 20:1679–1688, 2010.

[134] J. L. Kim and S. K. Burley. 1.9-angstrom resolution refined structure of tbp recognizing the minor-groove of tataaaag. *Nature Structural Biology*, 1(9):638–653, 1994.

[135] R. Kuddus and M. C. Schmidt. Effect of the non-conserved n-terminus on the dna binding activity of the yeast tata binding protein. *Nucleic acids research*, 21(8):1789–96, 1993.

[136] D. B. Nikolov, H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley. Crystal-structure of a tfiib-tbp-tata-element ternary complex. *Nature*, 377(6545):119–128, 1995.

[137] J. H. Geiger, S. Hahn, S. Lee, and P. B. Sigler. Crystal structure of the yeast tfiia/tbp/dna complex. *Science*, 272(5263):830–836, 1996.

[138] P. F. Kosa, G. Ghosh, B. S. DeDecker, and P. B. Sigler. The 2.1-angstrom crystal structure of an archaeal preinitiation complex: Tata-box-binding protein/transcription factor (ii)b core/tata-box. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12):6042–6047, 1997.

[139] Y. C. Liu and A. Schepartz. Kinetic preference for oriented dna binding by the yeast tata-binding protein tbp. *Biochemistry*, 40(21):6257–6266, 2001.

[140] J. M. Cox, M. M. Hayward, J. F. Sanchez, L. D. Gegnas, S. vanderZee, J. H. Dennis, P. B. Sigler, and A. Schepartz. Bidirectional binding of the tata box binding protein to the tata box. *Proceedings of the National Academy of Sciences of the United States of America*, 94(25):13475–13480, 1997.

[141] A. R. Kays and A. Schepartz. Virtually unidirectional binding of tbp to the admlp tata box within the quaternary complex with tfiia and tfiib. *Chemistry & Biology*, 7(8):601–610, 2000.

[142] R. F. Delgadillo, J. E. Whittington, L. K. Parkhurst, and L. J. Parkhurst. The tata-binding protein core domain in solution variably bends tata sequences via a three-step binding mechanism. *Biochemistry*, 48(8):1801–1809, 2009.

[143] K. M. Parkhurst, M. Brenowitz, and L. J. Parkhurst. Simultaneous binding and bending of promoter dna by the tata binding protein: Real time kinetic measurements. *Biochemistry*, 35(23):7459–7465, 1996.

[144] K. M. Parkhurst, R. M. Richards, M. Brenowitz, and L. J. Parkhurst. Intermediate species possessing bent dna are present along the pathway to formation of a final tbp-tata complex. *Journal of Molecular Biology*, 289(5):1327–1341, 1999.

[145] S. F. Tolic-Norrelykke, M. B. Rasmussen, F. S. Pavone, K. Berg-Sorensen, and L. B. Oddershede. Stepwise bending of dna by a single tata-box binding protein. *Biophysical Journal*, 90(10):3694–3703, 2006.

[146] B. P. Cormack and K. Struhl. The tata-binding protein is required for transcription by all 3 nuclear-rna polymerases in yeast-cells. *Cell*, 69(4):685–696, 1992.

[147] R. J. White and S. P. Jackson. The tata-binding protein: a central role in transcription by rna polymerases i, ii and iii. *Trends in genetics : TIG*, 8(8):284–8, 1992.

[148] K. Struhl, L. Kuras, P. Kosa, and M. Mencia. Taf-containing and taf-independent forms of transcriptionally active tbp in vivo. *Science*, 288(5469):1244–1248, 2000.

[149] B. F. Pugh. Control of gene expression through regulation of the tata-binding protein. *Gene*, 255(1):1–14, 2000.

[150] T. I. Lee and R. A. Young. Regulation of gene expression by tbp-associated proteins. *Genes & Development*, 12(10):1398–408, 1998.

[151] P. C. FitzGerald, A. Shlyakhtenko, A. A. Mir, and C. Vinson. Clustering of dna sequences in human promoters. *Genome Research*, 14(8):1562–1574, 2004.

[152] A. J. Jackson-Fisher, C. Chitikila, M. Mitra, and B. F. Pugh. A role for tbp dimerization in preventing unregulated gene expression. *Molecular Cell*, 3(6):717–727, 1999.

[153] Pugh Chicca, Auble. Cloning and biochemical characterization of taf-172, a human homolog of yeast mot1. *Molecular and Cellular Biology*, 18(3):1701–10, 1998.

[154] A. E. Gorbalenya and E. V. Koonin. Helicases - amino-acid-sequence comparisons and structure-function-relationships. *Current Opinion in Structural Biology*, 3(3):419–429, 1993.

[155] J. A. Eisen, K. S. Sweder, and P. C. Hanawalt. Evolution of the snf2 family of proteins - subfamilies with distinct sequences and functions. *Nucleic Acids Research*, 23(14):2715–2723, 1995.

[156] J. L. Davis, R. Kunisawa, and J. Thorner. A presumptive helicase (mot1 gene-product) affects gene-expression and is required for viability in the yeast saccharomyces-cerevisiae. *Molecular and Cellular Biology*, 12(4):1879–1892, 1992.

[157] T. A. Muldrow, A. M. Campbell, P. A. Weil, and D. T. Auble. Mot1 can activate basal transcription in vitro by regulating the distribution of tata binding protein between promoter and nonpromoter sites. *Molecular and Cellular Biology*, 19(4):2835–2845, 1999.

[158] R. P. Darst, D. Y. Wang, and D. T. Auble. Mot1-catalyzed tbp-dna disruption: uncoupling dna conformational change and role of upstream dna. *Embo Journal*, 20(8):2028–2040, 2001.

[159] D. Poon, A. M. Campbell, Y. Bai, and P. A. Weil. Yeast taf170 is encoded by mot1 and exists in a tata box-binding protein (tbp)-tbp-associated factor complex distinct from transcription factor-iid. *Journal of Biological Chemistry*, 269(37):23135–23140, 1994.

[160] J. Thorner, J. I. Adamkewicz, C. G. F. Mueller, K. E. Hansen, and W. A. Prud'homme. Purification and enzymic properties of mot1 atpase, a regulator of basal transcription in the yeast saccharomyces cerevisiae. *Journal of Biological Chemistry*, 275(28):21158–21168, 2000.

[161] D. T. Auble, K. E. Hansen, C. G. F. Mueller, W. S. Lane, J. Thorner, and S. Hahn. Mot1, a global repressor of rna-polymerase-ii transcription, inhibits tbp binding to dna by an atp-dependent mechanism. *Genes & Development*, 8(16):1920–1934, 1994.

[162] D. T. Auble, R. O. Sprouse, and M. Brenowitz. Snf2/swi2-related atpase mot1 drives displacement of tata-binding protein by gripping dna. *Embo Journal*, 25(7):1492–1504, 2006.

[163] J. Thorner, J. I. Adamkewicz, K. E. Hansen, W. A. Prud'homme, and J. L. Davis. High affinity interaction of yeast transcriptional regulator, mot1, with tata box-binding protein (tbp). *Journal of Biological Chemistry*, 276(15):11883–11894, 2001.

[164] B. Kobe, T. Gleichmann, J. Horne, I. G. Jennings, P. D. Scotney, and T. Teh. Turn up the heat. *Structure*, 7(5):R91–7, 1999.

[165] R. P. Darst, A. Dasgupta, C. M. Zhu, J. Y. Hsu, A. Vroom, T. Muldrow, and D. T. Auble. Mot1 regulates the dna binding activity of free tata-binding protein in an atp-dependent manner. *Journal of Biological Chemistry*, 278(15):13216–13226, 2003.

[166] K. P. Hopfner, P. Wollmann, S. Cui, R. Viswanathan, O. Berninghausen, M. N. Wells, M. Moldt, G. Witte, A. Butryn, P. Wendler, R. Beckmann, and D. T. Auble. Structure and mechanism of the swi2/snf2 remodeller mot1 in complex with its substrate tbp. *Nature*, 475(7356):403–U168, 2011.

[167] A. Dasgupta, S. A. Juedes, R. O. Sprouse, and D. T. Auble. Mot1-mediated control of transcription complex assembly and activity. *The EMBO journal*, 24(9):1717–29, 2005.

[168] O. H. Gumbs, A. M. Campbell, and P. A. Weil. High-affinity dna binding by a mot1p-tbp complex: implications for taf-independent transcription. *The EMBO journal*, 22(12):3131–41, 2003.

[169] I. Topalidou, M. Papamichos-Chronakis, G. Thireos, and D. Tzamarias. Spt3 and mot1 cooperate in nucleosome remodeling independently of tbp recruitment. *Embo Journal*, 23(9):1943–1948, 2004.

[170] C. J. van Oevelen, H. A. van Teeffelen, and H. T. Timmers. Differential requirement of saga subunits for mot1p and taf1p recruitment in gene activation. *Molecular and Cellular Biology*, 25(12):4863–72, 2005.

[171] D. T. Auble, R. P. Darst, and D. Y. Wang. Mot1-catalyzed tbp-dna disruption: uncoupling dna conformational change and role of upstream dna. *Embo Journal*, 20(8):2028–2040, 2001.

[172] J. O. Wu, K. M. Parkhurst, R. M. Powell, and L. J. Parkhurst. Dna sequence-dependent differences in tata-binding protein-induced dna bending in solution are highly sensitive to osmolytes. *Journal of Biological Chemistry*, 276(18):14623–14627, 2001.

[173] L. J. Parkhurst, K. M. Parkhurst, R. Powell, J. Wu, and S. Williams. Time-resolved fluorescence resonance energy transfer studies of dna bending in double-stranded oligonucleotides and in dna-protein complexes. *Biopolymers*, 61(3):180–200, 2002.

[174] P. Schlüsche, G. Heiss, M. Meisteremst, and D. C. Lamb. Dynamics of tbp binding to the tata box - art. no. 68620e. *Single Molecule Spectroscopy and Imaging*, 6862:E8620–E8620 120, 2008.

[175] R. O. Sprouse, I. Shcherbakova, H. Y. Cheng, E. Jamison, M. Brenowitz, and D. T. Auble. Function and structural organization of mot1 bound to a natural target promoter. *Journal of Biological Chemistry*, 283(36):24935–24948, 2008.

[176] P. de Graaf, F. Mousson, B. Geverts, E. Scheer, L. Tora, A. B. Houtsmuller, and H. T. M. Timmers. Chromatin interaction of tata-binding protein is dynamically regulated in human cells. *Journal of Cell Science*, 123(15):2663–2671, 2010.

[177] L. Tora and H. T. M. Timmers. The tata box regulates tata-binding protein (tbp) dynamics in vivo. *Trends in Biochemical Sciences*, 35(6):309–314, 2010.

[178] M. P. Klejman, L. A. Pereira, H. J. van Zeeburg, S. Gilfillan, M. Meister-ernst, and H. T. Timmers. Nc2alpha interacts with btaf1 and stimulates its atp-dependent association with tata-binding protein. *Molecular and Cellular Biology*, 24(22):10072–82, 2004.

List of abbreviations

AFM	Atomic Force Microscopy
ADP	Adenosine Diphosphate
ADU	Analog Digital Units
ATP	Adenosine Triphosphate
ALEX	Alternating Laser Excitation
AOTF	Acousto-Optic Tunable Filter
BIC	Bayesian Information Criterion
CIC	Clock Induced Charge
eGFP	Enhanced Green Fluorescent Protein
EMCCD	Electron-Multiplying Charged Coupled Device
FCS	Fluorescence Correlation Spectroscopy
FFT	Fast Fourier Transformaiton
FP	Fluorescent Protein
FRET	Förster Resonance Energy Transfer
GFP	Green Fluorescent Protein
HMM	Hidden Markov Model
HKM	HEPES with Potassium and Magnesium
Mot1	Modulator of transcription 1
MFD	Multiparameter Fluorescence Detection
NC2	Negative Cofactor 2
NPS	Nano Positioning System
OLID	Optical Lock-In Detection
PIC	Preinitiation Complex
PDA	Photon Distribution Analysis
PDB	Protein Data Base
PR	Proximity Ratio
PSF	Point Spread Function

QE	Quantum Efficiency
ROI	Region Of Interest
SNR	Signal-to-Noise Ratio
STM	Scanning Tunneling Microscope
spFRET	single-pair Förster Resonance Energy Transfer
TF	Transcription Factor
TBP	TATA-box Binding Protein
TDP	Transition Density Plot
TIRFM	Total Internal Reflection Fluorescence Microscopy

Acknowledgements

I thank Prof. Don Lamb, for his support and the opportunity to do my PhD research in his group. His insight, passion and endurance for science did continuously inspire my view on research. The freedom to follow up on ideas and the trust he put in me made this endeavor a precious professional and personal experience.

I thank my second advisor, Prof. Christoph Bräuchle, for bringing me on board on his excellent one-of-a-kind research institute. His support within the IDK-NBT gave me access to networks and meetings such as the "Lindau Meeting of Nobel Laureates". I very much enjoyed to work in this knowledgeable and resourceful environment.

I thank Prof. David Auble, my collaborator for the Mot1 project, for his insight, endurance and his biologist perspective on physical problems. Without his support, this project would not have been possible. He always found the right words at the right time and kept his optimism high despite loosing multiple protein shipments to FedEx mishandling.

I thank Prof. Promod Pratap, who visited our group during his sabbatical, for introducing me to membrane proteins and on how to extract them from ducklings.

I thank Prof. Christof M. Niemeyer, my collaborator for the Dronpa project, for providing me with valuable samples.

I thank Petra Wollmann for providing me with renderings of the Mot1 crystal-structure and discussions about the same.

I gratefully acknowledge the Elitenetzwerk Bayern (IDK-NBT), the Center for Nanoscience (CENS) and the SFB646 for funding, networking and support on many occasions.

I am happy and proud to have had such great colleagues from AK Lamb, AK Bräuchle and AK Michaelis. My impressions of my very first day at AK still

Curriculum Vitae

Gregor Johann Heiss, born 9.6.1980 in Tamsweg, Austria

Education
2007 - 2011 PhD student in the group of Prof. Don Lamb, LMU Munich
2004 - 2006 MSc. in physics, University of Illinois, Urbana-Champaign, USA

Publications
Parts of this work have been published or are in preparation for publication.

- M. Singleton, G. Heiss, A. Hübler: "Optimization of ramified absorber networks doing desalination" Phys. Rev. E 83, 2011

- M. Singleton, G. Heiss, A. Hubler: "The Optimal Shape of Roots" Complexity, Vol. 15, 9-11, 2010

- D.M. Togashi, A. G. Ryder, G. Heiss: "Quantifying adsorbed protein on surfaces using confocal fluorescence microscopy" Colloids and Surfaces B: Biointerfaces, 2009

- G. Heiss, V. Lapiene, F. Kukolka, C.M. Niemeyer, C. Bräuchle, D.C. Lamb: "Single-Molecule Investigations of a Photoswitchable Nanodevice" Small 5 (10) 1169-75, 2009

- P. Schluesche, G. Heiss, M. Meisterernst, D. C. Lamb: "Dynamics of TBP binding to the TATA-box" Proceedings of the SPIE, 2008

- O. Holub, M. J. Seufferheld, C. Gohlke, Govindjee, G. J. Heiss and R. M. Clegg: "Fluorescence Lifetime Imaging Microscopy of Chlamydomonas reinhardtii: non-photochemical quenching mutants and the effect of photosynthetic inhibitors on the slow chlorophyll fluorescence transient" Journal of Microscopy, Vol. 226, 90-120, 2007

- D. E. Chandler, Z. K. Majumdar, G. J. Heiss: "Ruby Crystal for Demonstrating Time- and Frequency-Domain Methods of Fluorescence Lifetime Measurements" Journal of Fluorescence, Vol. 16, 793-807, 2006