University of South Bohemia in České Budějovice
Faculty of Science

Single-molecule fluorescence detection in molecular biology

Ph.D. Thesis

Mgr. Tomáš Fessl

Supervisors: Prof. RNDr. František Vácha, Ph.D.
                RNDr. František Adamec, CSc.
University of South Bohemia, Faculty of Science and Biology Centre of the Academy of Sciences of the Czech Republic, v.v.i., Institute of Plant Molecular Biology

České Budějovice 2012

▨ Annotation

SMFD techniques offer genuine detection possibilities which are often inaccessible using ensemble methods. This was demonstrated in three projects investigating translocation activity of CHD4 protein, analysis of MS2 phage capsid assembly and in-cell characterization of DNA structure. In other projects, binding interactions between two fluorescent probes and a short oligonucleotide were characterized and all optical depth of focus extended microscope configuration for imaging of individual molecules inside bacterial cells was developed and tested.

▨ Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 11. 9. 2012 ....

jméno a příjmení

This thesis originated from a partnership of Faculty of Science, University of South Bohemia, and Institute of Plant Molecular Biology, Biology Centre of the ASCR, supporting doctoral studies in the Biophysics study programme.

▦ List of papers and author's contribution

The thesis is based on the following materials (listed chronologically):

I. **Fessl, T.**, Ben-Yaish, S., Vacha, F., Adamec, F., Zalevsky, Z., (2009). "Depth of focus extended microscope configuration for imaging of incorporated groups of molecules, DNA constructs and clusters inside bacterial cells." <u>Optics Communications</u> **282** (13): 2495-2501. (IF = 1.316, 4 citations)
Tomáš Fessl was responsible for planning of the experimental design, sample preparation, optical set-up assembly, imaging of individual fluorescent DNA molecules inside live E. coli cells, and participated in manuscript writing.

II. Kabelac, M., Zimandl, F., **Fessl, T.**, Chval, Z., Lankas, F., (2010). "A comparative study of the binding of QSY 21 and Rhodamine 6G fluorescence probes to DNA: structure and dynamics." <u>Physical Chemistry Chemical Physics</u> **12** (33): 9677-9684. (IF = 3.453, 3 citations)
Tomáš Fessl participated in planning of research design and manuscript writing.

III. Translocation activity of CHD4 protein at single-molecule level. First half of final report from FEBS sponsored stay at the Astbury Centre for Structural Molecular Biology, the University of Leeds in 2010.
Tomáš Fessl was responsible for DNA and protein labeling, DNA isolation and purification, single-pair FRET measurement and data analysis, interpretation and report writing.

IV. Study of the MS2 phage assembly mechanisms using fluorescence correlation spectroscopy. Second half of final report from FEBS sponsored stay at the Astbury Centre for Structural Molecular Biology, the University of Leeds in 2010.
Tomáš Fessl was responsible for statistical data analysis, interpretation, development of MATLAB software package computing regularized inverse Laplace transform for distribution fitting of fluorescence correlation spectroscopy data and report writing.

V. **Fessl, T.**, Adamec, F., Polivka, T., Foldynova Trantirkova, S., Vacha, F., Trantirek, L., (2012) "Towards characterization of DNA structure under physiological conditions *in vivo* at the single-molecule level using single-pair FRET." <u>Nucleic Acid Research</u>. (IF = 8.026)
Tomáš Fessl was responsible for planning of the experimental design, sample preparation, data measurement, data analysis, interpretation and participated in the manuscript writing.

Content

1. Introduction

The main task of this thesis is to introduce methods of Single-Molecule Fluorescence Detection (SMFD) to the extensive community of molecular biologists. SMFD methods often provide unprecedented straightforward views on various biological issues, which are directly inaccessible using ensemble methods. Despite all its benefits, SMFD is not widely known or used.

Perhaps the most important question is: why should the single-molecule concept be attractive for researchers in biological sciences? To address this question and demonstrate the full potential of SMFD techniques, here I report a broad introduction to SMFD applied to molecular biology, including its motivation, history, literature overview of current biological applications, basic principles, list of SMFD techniques and sample treatment.

Finally, the results section comprises actual work proving broad usefulness of SMFD approach to biology and two projects associated with precise realization of SMFD. Results namely include a proof of concept work introducing investigation of DNA structure under physiological conditions *in vivo*, characterization of CHD4 protein translocation activity and a study of MS2 bacteriophage capsid assembly. Another two projects involve *in silico* comparative study of the binding interaction of fluorescent probes Rhodamine 6G and QSY 21 to DNA and implementation of an optical element extending the depth of focus of SMFD apparatus.

1.1. Aims

a) To introduce capabilities of SMFD to the local community of molecular biologists and establish fluorescence lifetime based spFRET and time-resolved fluorescence polarization.

b) To extend the existing SMFD apparatus for fluorescence lifetime, lifetime-based single-pair FRET and time-resolved fluorescence anisotropy detection.

c) To demonstrate SMFD experimental powers in biologically relevant studies.

1.2. Biology in singulo: *motivation*

SMFD is a collection of techniques investigating the spectroscopic properties of a fluorophore with ultra high spatial resolution. Over the last two decades the growth of findings concerning behavior of biological systems on single-molecule level has been spectacular. As usual, the story started with technique development, immediately followed by applications to biological sciences.

A lot of reactions in living organisms are far too complex to be fully elucidated via standard ensemble experiments. Using conventional reductional perspective, all reactions in biology might be seen as a feat of individual enzyme, DNA or RNA molecules. Analyzing individual biomolecules one at a time will often yield us

phenomenally clear and astonishing information about their nature and function.

There are generally three main and invaluable advantages of SMFD compared with standard ensemble experiments. First, SMFD allows the direct detection of true statistical distribution that is otherwise obscured under the ensemble average [1], e.g. we can determine even very rare and/or short lived states of biomolecules [2]. Second exclusive feature of SMFD tools is that synchronization of individual processes, e.g. enzyme activity, diffusion in membrane and action of molecular machine is not needed. This is achieved by monitoring of individual molecules in real time and synchronization of their action afterwards (see Figure 1) [1]. Third, SMFD principle is well suited to explore fluctuating systems in equilibrium and to detect time trajectories and reaction pathways of individual participants in a non-equilibrated system. Concretely, SMFD allows us to monitor intermediates and follow time dependent pathways of biological reactions [2].

Nonetheless, what kind of information does SMFD methods provide about the host biomolecule? SMFD provides high spatial and temporal localization of individual objects. The localization of individual molecules can be achieved even with ~1 nm precision by fitting the record of a individual fluorescent probe (or two spectrally separated probes) to a two-dimensional Gaussian function [3].

Another SMFD technique, single-molecule fluorescence polarization, might be employed to monitor the spatial orientation of the host biomolecule, its domain or molecular binding [4][5]. Moreover, the distance within one biomolecule or between different biomolecules can be measured by detecting the efficiency of Förster Resonance Energy Transfer (FRET) between two suitable fluorescent probes in the range between $1 - 9$ nm [6]. Also the diffusion coefficient and subsequently Stokes radius (~hydrated size) of biomolecules can be monitored by Fluorescence Correlation Spectroscopy (FCS) [9]. FCS does not strictly belong to SMFD as the signal is usually averaged over thousands events, nevertheless it is measured at very low concentrations i.e. at any given time there are only one or a few biomolecules in the probed volume. All mentioned techniques will be described, in more detail, in subchapters of section 1.6.
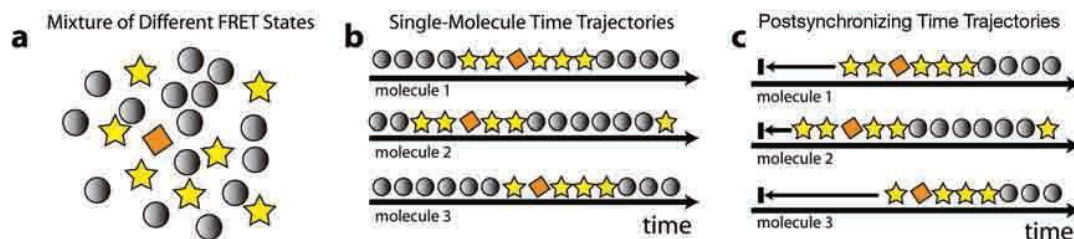

Figure 1: Single-molecule FRET trajectories post-synchronization

(a) Ensemble FRET represents the average value from the whole ensemble of molecules. (b) FRET time traces of individual molecules show the real states of molecules in real time. (c) Studied molecular events were post-synchronized within subsequent data analysis. Adopted from Selvin et al. *[1].*

1.3. Very brief history of SMFD techniques

At the beginning, SMFD was used as a complementary method to, at that time, well established line narrowing techniques, such as photon echo [7] and hole burning [8] to explore the fundamental dynamics of amorphous solid matrices in ultralow temperatures. Both photon echo and hole burning remove the chromophore inhomogeneous broadening and probe the interaction between chromophore and matrix via detection of homogeneous broadening. Nevertheless, photon echo and hole burning still detect averaged information.

First in the row of single-molecule spectroscopic experiments was the detection of absorption spectra of single pentacene molecules in a *p*-terphenyl matrix. This pioneering experiment was performed by Kador and Moerner in 1989 [9]. One year later, Orrit and Bernard measured the first fluorescence excitation spectrum of the same system [10]. The breakthrough step towards physiological conditions was done by Betzig and Chichester in 1993, when they introduced their SMFD experiments into room temperature [11]. Nevertheless their near-field scanning optical microscopy experiment was perturbed by the metal-coated tip [12]. Next milestone of SMFD employed in biological research was reported by Yanagida in 1995, as he monitored individual molecules in liquid solution (standard procedure of recent SMFD) [13]. Selvin *et al.* introduced single-pair FRET (spFRET) in 1996 [14]. Only four years later came the first *in vivo* applications of SMFD. It was independently realized almost at the same time by Schindler [18] and Yanagida [15]. Up until today, SMFD is a growing and flourishing sphere of research.

1.4. SMFD applied in biology

The interest of biologists in fluorescence detection of individual biomolecules can be followed back to 1977, when Hirschfeld detected an individual antibody tagged with less than 100 fluorophores [16]. In the consecutive year, Hirschfeld *et al.* presented an instrument capable of determining size and classification of individual viruses [21] [22]. Despite the fact that none of Hirschfeld's experiments was performed at single fluorophore level, his group laid the foundations of several SMFD essentials in background reduction. These were composed reduction of excitation volume, time-gated detection, and sample prephotobleaching to avoid impurities [21, 23]. Moreover Hirschfeld defined photobleaching as a limiting factor for the number of emitted photons per individual fluorophore [17].

1.4.1. Structural analysis

Structural information is based on precise knowledge of position, angle or distance between atoms, molecular domains or molecules as part of molecular complexes. Compared to other structural methods, using principles of Nuclear Magnetic Resonance (NMR), Electron Paramagnetic Resonance (EPR) and X-ray diffraction, SMFD methods have inherently lower structural resolution. However, the main power of SMFD methods, applied in structural studies, is in detection of short living species, determination of structural intermediates and in direct investigation of statistical distribution of conformational states. SpFRET and FCS proved to be the most valuable SMFD techniques to obtain static information about biomolecular

structure.

SpFRET measures interprobe distances, hence with strategic probe placement, several suggested structural models can be investigated. SpFRET may also serve as an excellent complementary method to high resolution methods as X-ray and NMR to probe the position of flexible domains within biomolecules. Muschielok *et al.* simplified this type of analysis and developed a nano-positioning system (similar to the well known geo-positioning system) MATLAB (MathWorks) software package that uses probabilistic data analysis to combine spFRET experimental data with a known structural model, previously determined by X-ray or NMR [25][26]. SpFRET is also suitable to study conformational distribution of intrinsically disordered proteins. Their intrinsic lack of stable structure allows them to bind to several different targets. The flexibility also provides precise control over the binding process inducible by phosphorylation or through interaction with other components of the cellular machinery [27]. The heterogeneous conformational distribution and unsynchronized action makes them ideal candidates for single-molecule study [18].

By contrast, the FCS analysis yields the Stokes radius (an approximation of hydrated size) derived from diffusion time. Thus it might be used to approximate the size of fluorescent objects and follow their assembly, composition or aggregation [29].

DNA and RNA structure involves many short living states and complicated structural dynamics with several intermediates even under equilibrium conditions. Therefore, it is well suited for SMFD studies. The majority of SMFD structural experiments on biomolecules deal with DNA and RNA structure, thus also this thesis focuses mainly on nucleic acids structure. Literature about protein structure and dynamics can be found elsewhere [19].

1.4.1.1. DNA structure

DNA structure and dynamics is primarily defined by its sequence. Nevertheless, these molecules are highly sensitive to their environmental conditions, such as polarity of the solvent [20], presence of particular ions and their concentration [32], molecular crowding [21] and hydration [22]. All mentioned factors influence stiffness and helical parameters of double-stranded DNA (dsDNA) molecules. Radical change in solvent polarity or salt concentration may even induce conformational transition between A, B and Z dsDNA form. Nevertheless, there is plethora of more complicated DNA structures, often involved in enzymatic actions tightly associated with basic cellular mechanisms, such as DNA replication and repair, whose structural mechanism still remains unclear. SMFD plays an important role in their investigation.

DNA duplexes: In 2008, Wozniak *et al.* precisely measured dsDNA subtle bends and kinks [35] using spFRET. Thus they showed the distance-derived limiting feature of DNA structure resolvable by this technique.

The same year, Iqbal *et al.* employed time-resolved spFRET to investigate the helical modulation of dsDNA [23]. They were able to distinguish between A and B form dsDNA using phase shift in the helical modulation of FRET efficiency with helix length. Thus they extended the conventional use of FRET, measuring of distance, adding information about mutual orientation of fluorophores from a set of measurements.

Hicks *et al.* examined the conformational diversity and dynamics of hybridization of 25 base pair long complementary DNA (cDNA) molecules. Individual single-stranded DNA (ssDNA) molecules were optically trapped within water nanodroplets and probed via spFRET at various temperatures and NaCl concentrations [24]. The studied system exhibited frequent conformational fluctuations, proposing that the freely diffusing short DNA oligonucleotides do not remain steadily in a B-form conformation. Moreover, atypical conformers constituting to an unfolding pathway appeared under low NaCl concentration.

DNA loops: DNA molecules form even more complicated structures than just double helixes. Moreover, they often fluctuate between alternating conformational states, which is behavior that can be followed using SMFD methods. Hairpin loop (stem loop) is a relatively simple DNA motif, important in protein binding, which can fluctuate between an open and closed form.

The kinetics of DNA hairpin loop fluctuations was investigated using a FRET fluctuation spectroscopy (for details see section 1.6.3) [25]. Bonnet *et al.* measured the activation energies associated with the opening and closing of a hairpin loop of different sizes, sequences and under various NaCl concentrations. The rate of opening of the hairpin stem was found to be independent of the characteristics of the loop, while the closing rate differed with the loop sequence and length. The closing rate was faster at higher NaCl concentrations. Regarding hairpin closing, the loop of adenosine repeats induced a decrease in rate and higher activation energies than a loop with thymine repeats.

Conformational fluctuations of the DNA hairpin were also studied as a function of temperature and viscosity using FRET fluctuation spectroscopy. The closing of the loop kinetics was interpreted with a two-state model. Resulting kinetics revealed a non-Arrhenius behavior. Wallace *et al.* also suggested that stacking interactions in the loop might increase the roughness of the free energy surface of the DNA hairpin [26].

Later in 2006, FRET fluctuation spectroscopy and Photon Counting Histograms (PCH) analysis were coupled to monitor thermodynamic parameters of a hairpin opening and closing in various NaCl concentrations [27]. Jung *et al.* found a three state mechanism of hairpin folding that involves a stable intermediate form of the DNA hairpin. The intermediate and open states were suggested to be in a rapid equilibrium between conformations.

In the same year, Kim *et al.* investigated the first step of hairpin folding. They detected hairpin formation via FCS monitoring a guanine-mediated quenching of suitably located fluorophore MR121. Resulting data suggested that the rate limiting step is not defined by stem sequential mismatches or loop dynamics, but by the interactions between loop and stem [28].

Recently, Probability Distribution Analysis (PDA), a statistical method, which simplifies analysis of static molecular species within a spFRET histogram, was employed to study dynamics of hairpin formation. [42] PDA revealed the timescale of hairpin conformational fluctuations. Reaction times were in good agreement with data obtained, one year before, using correlation-based analysis [29].

Branched DNA structures: Branched DNA structures are chief intermediates in DNA duplication, enzymatic repair of damaged replication forks, repair of DNA damage, and homologous recombination [30]. The structures of these molecules have a crucial impact on recognition by a multitude of DNA-processing proteins. Albeit crystal structures are known for a few of such molecules, e.g. the Holliday Junction (HJ) [31] but the detailed structures of majority of branched nucleic acids remains unknown. [32] SMFD techniques offer an ideal set of tools to investigate the structure of individual branched DNA intermediates without ensemble averaging [47].

The HJ is a central intermediate state of DNA for genetic processes. Branch migration is the key process by which the exchange of homologous DNA regions occurs. It might be spontaneous or regulated by proteins. Unfolding of the HJ is necessary for branch migration, which is a stepwise process comprising consequential migration and folding steps. Folding of the HJ is crucial, since one of the folded conformations terminates the branch migration phase. Moreover, in the unfolded state HJ quickly migrates over entire homology region in one hop [33].

Detecting spFRET on immobilized Holliday junctions, McKinney *et al.* found that the branch migration process has an intermediate, in the form of an open unstacked structure [34]. One year later the same group observed that the HJ fluctuation rate drops with increasing concentration of Mg^{2+}, whereas the ratio of populations of the states remains constant [2]. Dependence of stepwise branch migration dynamics of HJ on Mg^{2+} concentration was proved and theoretically explained one year later [50]. SMFD techniques were also employed to define the free-energy landscape of spontaneous branch migration. It was found to be highly rugged with sequence-dependent barriers and local energy minima [51].

Karymov *et al.* investigated the HJ hopping, analyzing spFRET time trajectories. Their results suggest that the transition length, referred as hop, is not necessarily 1 base pair, as stated before. By contrast, the junction is able to migrate over several DNA base pairs, detected as one hop. Derived distances confirm a partially opened, side-by-side model with adjacent double-helical arms slightly kinked at the four-way junction [52].

Recently, Palets *et al.* studied the effect of DNA defects - nicks - on the efficiency of HJ branch migration and the dynamics of the HJ. Combining results from AFM and spFRET analysis, their study revealed a dramatic nick moderated change of the HJ dynamics. The nick changed the structure and conformational dynamics of the junctions, leading to induction of new conformations that were not observed for the intact HJ [33].

Human intramolecular DNA quadruplexes: Structural distributions and dynamics of these vital DNA structures have also been studied using SMFD methods. SpFRET results revealed two subpopulations of human telomeric intramolecular quadruplex. These were interpreted as coexisting parallel and anti-parallel quadruplex conformers. The conformers were found to be interconvertible on a time-scale of minutes [53]. The free energy barrier in the folding energy landscape was estimated to be between 3 and 15 k_BT. Another study employed spFRET to find surprisingly extreme diversity of conformational states involved during the folding of immobilized human telomeric DNA [54].

Shirude *et al.* compared structure and dynamics of human genomic DNA G-quadruplex sequences that occur in the telomeres (htelo) and in the promoter of a proto-oncogene (c-kit-promoter G-quadruplex) [35]. Their spFRET-based study suggests that both c-kit and htelo intramolecular DNA G-quadruplexes are conformationally heterogeneous. The dynamic interconversions of htelo between distinct FRET states, on a time scale of seconds to minutes under physiological conditions, suggest that several of the conformational states can be relevant for the telomeres in live cells. The authors proposed that other components of the telosome might control which conformational state predominates in cells and at which phase of the cell's cycle it appears. Shirude *et al.* also found folded structures of the c-kit promoter and proved that such promoter G-quadruplex conformations are an alternative to the duplex structure. Moreover, they challenged the doctrine that G, C-rich DNA exclusively forms a stable duplex, and they supported hypothesis that basic molecular mechanisms involve changes in the topology of the DNA.

The effect of single-base mutations on the structural dynamics of human telomeric G-quadruplexes was explored by Lee and Kim [36]. Structural polymorphism was invoked by single guanine-thymine replacement in a G-track. In addition, analysis of the dynamics revealed that this small mutation induces fast unfolding and refolding events under physiological conditions [36].

1.4.1.2. RNA structure

Although synthesized as linear polymers, RNA molecules fold into complicated structures vital for cellular processes. Messenger, transfer and ribosomal RNA molecules play a key role in RNA editing and splicing, protein secretion, telomere maintenance, small-molecule sensing and reaction catalysis. How RNA molecules achieve their battery of functions with an ultimately limited amount of building blocks is a question of great interest. Nevertheless, the determination of rules defining RNA structure from primary structure is far from trivial. Many RNAs are poorly conserved, and their function cannot be deduced by simple homology searches based solely on primary structure [37]. Below is a review of the most recent SMFD-based publications about small tRNA molecules and a nearly complete overview of SMFD work on large RNA molecule with enzymatic function - RNase P RNA.

Structure and folding of tRNA: During translation, the ribosome accurately selects aminoacyl-tRNA (aa-tRNA) to fit the particular codon of mRNA, the exact mechanism of the selection is still unknown. Mishra *et al.* partly elucidated this process by employing spFRET coupled with fluorescence anisotropy. They provided evidence of random thermal motion (fluctuations) of tRNAs within the ribosome in nanosecond timescale. They showed that cognate aa-tRNA fluctuates less frequently than near-cognate. This result was counterintuitive since cognate aa-tRNA was expected to fluctuate more often to reach the ribosomal A-site faster than near-cognate. Furthermore, the authors suggested that both cognate and near-cognate aa-tRNA occupies the same position within the ribosome. Mishra *et al.* also proposed that fluctuating cognate aa-tRNA is guided by the ribosome toward the A-site by the conformational changes in the 30S subunit [38].

The energy landscape of RNA folding was found to be significantly affected by different counterions that neutralize the phosphate group negative charge on the RNA backbone and also play a key role in tertiary structure. Using spFRET, Dammertz *et al.* investigated the folding properties of the RNA transcript of human mitochondrial tRNA(Lys) in thermodynamic equilibrium. This transcript showed two different folded states and one unfolded. The authors quantified the stabilization effect on RNA tertiary structure for different types of cations. They found that small monovalent ions stabilize the probed structure better than larger ones. Divalent ions showed the most pronounced conformation-dependent binding specificity. The radii of these divalent ions and coordination properties were found to be responsible for shaping the folding free energy landscape [39].

Folding and structure of RNase P RNA: The ribonucleoprotein enzyme ribonuclease P (RNase P) catalyzes the maturation reaction of tRNAs by cleavage of the 5' leader sequence of precursor tRNAs in bacteria and eukaryotes. RNase P is a well-characterized large ribozyme, the RNA component catalyzes tRNA maturation *in vitro* without proteins [60]. Preceding ensemble studies of this large RNA suggested a two-step folding pathway. Within this pathway, the catalytic domain was found in three conformational states (unfolded, intermediate and folded native state). [40] The number of known Mg^{2+} dependent folding intermediates was extended in a study of RNase P RNA folding. SpFRET here revealed at least two previously unidentified folding intermediates [62].

Large RNAs with more complex intermediate structures usually have non-exponential kinetics of folding. Smith *et al.* employed spFRET to analyze the dynamics under low concentrations of Mg^{2+}. Surprisingly, they detected simple exponential kinetics. This proposes a small quantity of discrete conformational states [63].

Qu *et al.* extended spFRET investigation of this enzyme by controlled Mg^{2+} concentration oscillations [41]. The system exhibited complex kinetics, dependent on the periodically alternating concentration of Mg^{2+} in solution. The detected spFRET time series exhibited a quasi-periodic response to a square-wave pattern of Mg^{2+} concentration changes. Nonetheless, these measurements probed only one of the many degrees of freedom in the macromolecule. Thus, one year later, the same group identified the origin of these quasi-periodic responses in the back and forth switching in the positions of potential minima. The authors proposed a precise model and concluded that SMFD experiments with periodic perturbations provide qualitatively new information beyond the conventional results obtained at equilibrium. Using their novel approach they identified four folding substates [65].

1.4.2. Conformational dynamics and function

The precision of parameters, valuable for structural studies provided by SMFD methods, is usually low; however these methods are insuperable in detection of statistical distributions and temporal changes of structure i.e. in investigation of structural dynamics. Importantly, dynamics is inherently related to a biological function. SpFRET, FCS, PCH and fluorescence anisotropy are the most frequently

used techniques for investigation of conformational distributions in time, dynamics and function of biomolecules.

SpFRET reveals interprobe distance and might be captured in a time range usually between 1 ms and 100 s [42], whereas FCS detects signal fluctuations, which can originate from structural changes probed via several biophysical phenomena in the whole range of biologically interesting timescales[67]. PCH allows exploration of changes in probe brightness, such as fluorescence quenching or intensity increase. These changes might be caused by molecular reorganization, assembly or aggregation [68]. PCH shares the same time resolution with FCS [43]. Another technique, single-molecule fluorescence anisotropy monitors the velocity and rotational freedom of biomolecules or their domains. This feature is particularly useful for investigation of molecular machine motion, as these machines often rotate while translating along a linear rail (such as DNA or microtubule). For example, Forkey *et al.* employed 3D fluorescence anisotropy to examine tilting of the myosin light chain domain during structural dynamics of myosin V walk [44].

1.4.2.1. Biological assemblies

SMFD techniques are well suited to investigate the assembly of molecular complexes and their function. Important question of this field is: How is conformational change coupled with molecular binding in biology? Recent SMFD studies cover most of the important fields in biological function. A few of these studies, related to chromatin remodeling, chaperonine function and disease related protein misfolding, are overviewed below.

Protein folding: Nowadays, there are nearly 40 diseases known to be associated with protein misfolding. Identification of folding missteps might remarkably contribute to the development of new pharmaceuticals. More than twenty of these pathologies are associated with amyloids. Amyloids are insoluble fibrous protein aggregates, arising from at least 18 different misfolded proteins and polypeptides [71].

Liang *et al.* investigated the earliest steps of the folding of the Alzheimer's disease Aβ peptide employing real-time fluorescence imaging with photobleaching and FCS. They revealed the instantaneous formation of micrometer sized clusters. Moreover, these early intermediates kept properties of intermolecular molten globules, which served as the nucleating sites for amyloid growth [72].

Another protein whose misassembly causes severe disease is α-synuclein. It is an intrinsically disordered protein whose conformational change and subsequent aggregation into fibrous structures leads to Parkinson's disease. Trexler *et al.* utilized FCS and spFRET to probe the structure of α-synuclein under several aggregation prone conditions. Low pH and charged small molecules (spermine and heparin) induced conformational changes in α-synuclein and accelerated the aggregation. Observed conformational changes were dramatically different for both groups of environmental conditions [45].

Correct protein folding can be mediated via the action of molecular chaperons, although the mechanism of how they affect of protein folding is still unclear. Hofmann *et al.* employed spFRET coupled with microfluidics (see Figure 2) to investigate the protein folding inside the GroEL/GroES chaperonin cavity over a wide range of time

and ionic conditions. Investigation of protein rhodanese showed that confinement in the chaperonin decreases the rate of the C-terminal domain folding, but does not affect the folding rate of the N-terminal domain. The authors suggested that the chaperonin mechanism is not unique for all proteins and involves competition between intramolecular and intermolecular interactions, which define the folding rates and particular mechanisms of this reaction [74].
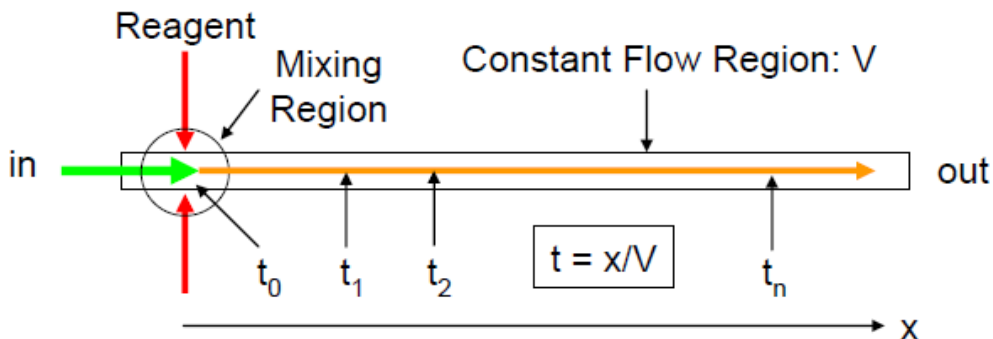

Figure 2: Microfluidics scheme for single-molecule stopped flow experiments.

Probed biomolecules and reagents are rapidly mixed in the mixing region on one side of the microfluidics chamber and flowed at a constant flow velocity toward the other end. Fluorescence is detected at different distances from the mixing region, corresponding to different time points t_i in the reaction. Adopted from [46].

Chromatin remodeling: Nucleosomes are multi-component macromolecular assemblies, functioning as the basic units of chromatin compaction in eukaryotic cells. They play a key role in regulation of basic biological processes, such as DNA transcription, repair and replication. They affect the accessibility of DNA through conformational dynamics like the transient unwrapping of DNA from the nucleosome, nucleosome repositioning, or partial dissociation of the DNA-nucleosome assembly. SMFD techniques have significantly helped to investigate such conformational dynamics, revealing a very dynamic organization of the nucleosome, modulated by post-translational modifications of the histone proteins (such as H3 K56 acetylation) and DNA sequence [47].

One such modification is lysine acetylation of histones. Neumann *et al.* reported that acetylation of H3 K56 does not directly affect the compaction of chromatin and has only modest effects on remodeling by SWI/SNF and RSC. Besides this, their spFRET experiments revealed that H3 K56 acetylation induces a seven fold increase in transient and spontaneous unwrapping and rewrapping of outer stretches of wrapped DNA from the nucleosome (DNA breathing) [77].

DNA breathing is suspected to be the main process governing access of proteins to nucleosomal DNA target sites. Recently, Tims *et al.* employed stopped-flow spFRET coupled with FCS [48] to measure the rates of unwrapping and rewrapping. They presented that the spontaneous DNA unwrapping rate decreases dramatically with location of the protein target site inside the nucleosome. In contrast, the rewrapping rate was found to decrease only slightly. These observations clarified the

well-known, strong position dependence of the accessibility of nucleosomal DNA [79].

A different mechanism revealing DNA from the wrapped state is nucleosome disassembly. Recently, Bohm at al. published an spFRET & FCS study presenting a new intermediate state before H2A-H2B dimer release. Their results suggest that the first step in nucleosome disassembly is the opening of the $(H3-H4)_2$ tetramer/(H2A-H2B) dimer interface, the second step is then H2A-H2B dimer release from the DNA and the last step involves $(H3-H4)_2$ tetramer removal. Nucleosome assembly follows the reverse pathway with the same intermediates [80].

Nucleosome assembly is also remodeled by ATP-driven machines. Chromatin assembly and remodeling factor (ACF) is a molecular motor, which generates regularly spaced nucleosomes. Blosser *et al.* utilized spFRET and fluorescence localization to investigate its mechanism. In the presence of ACF and ATP, the nucleosomes were translocated along DNA. This movement had two different kinetic pauses, after three or seven DNA base pairs of translocation. The binding of ACF, DNA translocation and exiting of translocation pauses were all proved to be ATP-dependent [81]. We studied CHD4, an ATP-driven chromatin remodeler with a different mechanism of action than ACF, for details see section 2.3.

1.4.2.2. Single-molecule enzymology

SMFD techniques brought several novel insights into the enzymatic function and kinetics of several proteins and RNAs, which were previously hidden in conventional ensemble measurements [49][50]. The most startling discovery in this field is an environment-independent diversity in the reaction activities for different individual enzyme molecules. This was first explained by diverse conformations of enzyme molecules; where two different conformational sub-states were ascribed different catalytic activities.

Lu *et al.* published a breakthrough work investigating the enzymatic activity of individual cholesterol oxidases that catalyse the oxidation of cholesterol by oxygen. They utilized fluorescence polarization, FCS and imaging of individual molecules of fluorescent cofactor, flavin adenine dinucleotide (FAD) [51]. Fluorescent FAD oxidizes cholesterol during the catalytic cycle and turns to a dark reduced form, $FADH_2$, which is then oxidized back to FAD by molecular oxygen. The authors monitored this cycling between bright and dark states to detect individual turnovers in real time. They analyzed fluorescence time trajectories resulting in an interesting observation that the enzymatic activity of individual molecules fluctuated over time. The analysis showed that enzyme turnover waiting times often depended on the previous ones (termed molecular memory effect). This was explained by slow conformational dynamics of the enzyme. There are also other enzymes which exhibit memory effects, validating the enzyme conformational fluctuations as a major reason of the variations in enzymatic activity [52][49][50].

In the year 2000, Polakowski *et al.* proposed that post-translational modifications may play an important role in the detected variations [53]. Nonetheless, the question whether the wide distribution of kinetic rates is determined by the inherent structural properties of enzyme proteins or whether biological systems play an active role and govern these rates via structural and environmental modifications remains open.

Recently, the group of Peter Lu analyzed fluctuations in T4 lysozyme nonequilibrium conformational kinetics (measured via spFRET) and found a significant bunching effect [54][55]. The authors explain the bunching effect as a convoluted outcome of multiple consecutive Poisson rate processes that are defined by protein functional motions under substrate−enzyme interactions.

The same group analyzed spFRET time trajectories of the enzyme 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase in a catalytic reaction. Their data exhibited intermittent coherence, suggesting that the catalytic reaction involves a multistep conformational motion along the coordinates of substrate−enzyme complex formation and product releasing. Analyzed coherence frequency determined that the correlation function analysis from spFRET time-trajectories increased with the increasing concentration of substrate. The authors suggested that intermittent coherence in conformational state changes (at the enzymatic reaction site) exists also in other conformation regulated enzymatic reactions [89]. For both described phenomena, the authors suggest strong consequences associated with biological oscillations, temporal and spatial functionality.

1.4.2.3. Molecular motors

Another rich area of SMFD-based exploration has been that of molecular motors. Molecular motors are biological molecular machines that perform a variety of key functions in the cell, including transport, synthesis and assembly. With respect to energetic efficiency, the majority of molecular motors are superior to recent man-made motors.

Molecular machines are usually enzymes, which are able to convert energy taken from a thermodynamically spontaneous chemical reaction (typically nucleotide hydrolysis) to a motion or a mechanical work, such as generation of torque or displacement. Due to their tiny size, many parts of these molecular machines operate at energies only slightly higher than the energy of their thermal bath. Therefore the operation of such parts is subjected to large fluctuations. These fluctuations were approved by single-molecule analyses [90]. Behaving as genuine thermodynamically open systems, molecular machines exchange energy and matter with their thermal baths and utilize the fluctuations to operate as energy rectifiers. This is exactly as the 'honest' Maxwell Demon that sits astride the line that separates stochastic from deterministic phenomena [91]. Molecular motors are often divided into functional groups such as rotary motors, nucleic acid motors, cytoskeletal motors and polymerization motors [56]. Molecular motors are investigated mostly via fluorescence polarization detection, FRET and super-resolution imaging techniques. Examples of SMFD studies on motor proteins are overviewed below.

Rotary motors: Nowadays, five natural rotary molecular motors are known, namely two molecular machines in each of ATP-synthase and V-ATPase and the bacterial flagellar motor. Rotation of the flagellar motor facilitates a helical propeller that drives the cell motility. The further rotary motors couple proton (or sodium cation) gradients to synthesize or hydrolyze ATP. The rotation within molecular complexes is here only part of the coupling mechanism, not the main function of the machines. The rotation mechanism of the F1 complex within ATP-synthase and the flagellar motor was

investigated under a wide range of environmental conditions [57].

Single-molecule investigation of F_1-ATPase started very recently, in 1997, when Noji and colleagues reported the first observation of its rotary motion in action [58]. However it was not strictly speaking a single-molecule experiment since they monitored rotation of an intensively labeled actin filament attached to the rotor unit. Using video microscopy of the rotating filament, the authors estimated almost ideal (approaching 100%) efficiency for this rotation molecular motor. Later, several SMFD experiments confirmed rotary motion of F_1-ATPase e.g. the first real single-molecule experiment, polarization–based study made by Adachi [95].

The first experiments were performed on motors immobilized to a glass cover slip. Later, Kaim *et al.* using fluorescence polarization [96] and Zimmermann *et al.* employing spFRET [97] transferred the studies into a biologically relevant environment and investigated the activity of F_0F_1-ATP synthase incorporated in membrane.

Although many structural details were known from crystallographical studies, the complete mechanism of the motor motion still remains unclear. In 2009, Düser *et al.* investigated stepwise c-ring rotation relative to the stator A subunits in E. coli F_0F_1–ATPase during ATP synthesis [59]. Employing spFRET techniques, they estimated the rotation angle step of the c-ring as approximately 36°. This angle corresponds to the suggested ten subunits of the c-ring. They also proved that every single proton directly drives the rotation of the F_O motor.

A complete mechanistic model can elucidate the unprecedented efficiency of molecular motors. Elastic torque transfer between F_O and F_1 is fundamental for enhancement of the kinetic efficiency of the complex. One known elastic domain is on the central rotor, between the two units, which are generating and depleting torque. Wachter *et al.* investigated whether the peripheral stalk of stator also serves as an elastic buffer. Particularly, they examined the homodimer b_2 (the most extended part of the stalk) from E. coli F-ATP synthase. By thermal rotary fluctuation analysis of two Q-dots attached to the stator, the authors determined the spring constant of the stator in response to bending and twisting. In wild type and also in all examined mutants the stator was at least 10 fold stiffer than the rotor and the molecular motor kept its proton-coupled activity [99].

Nucleic acid motors

Nucleic acids are similar to molecular railways on which molecular motors move as vehicles. They have a well-defined polarity with distinct chemistry, which attracts specific motor proteins to board and travel. Nucleic acid motors are often directional with variable speed and step size. Their processivity is defined by how long the motors stay on the rails before trailing off to an exit. The motion of nucleic acid motors is driven by ATP hydrolysis, which alternates their conformation, thus converting chemical energy into mechanical movement in a stepwise manner. Nucleic acid molecular motors include the genome processing motors, such as ribosomes, helicases, reverse transcriptases and polymerases which participate in natural gene transfer. A second group of motors involves viral packaging motors and other enzymes that process nucleic acids in a non-genomic manner. Motion of these motors cannot be effectively synchronized and detected using ensemble methods. Therefore SMFD

methods were often called to elucidate the mechanistic action underlying their function [60].

Bacteriophage DNA-packaging motor: During the replication cycle viruses compact their DNA in preparation for infection of host cells. The majority of bacteriophages can package their genomes extremely tightly. Packaging molecular motors must therefore overcome large forces to compress the whole viral genome into a small capsid. Despite the fact that SMFD methods alone do not allow force detection, they have a great potential to follow the kinetics and functional details of this mechanistic issue.

In a set of exemplary single-molecule experiments [101] Hugel *et al.* combined polarization-sensitive single-molecule fluorescence imaging and magnetic trapping to investigate the DNA packaging into the Bacteriophage φ29 capsid. They proved that subunits of the packaging motor act in a coordinated and processive manner. They also refuted that rotation of the motor connector complex plays a role in DNA packaging, even though structural models over the past 30 years coupled DNA movement to rotation of the connector relative to the capsid.

Sabanayagam *et al.* employed FCS to investigate the ATP-dependent kinetics of DNA packaging into the bacteriophage T4. Their results suggest that several 100 bp long DNA molecules can be sequentially translocated into a single bacteriophage prohead. Concurrently, FRET between GFP labeled prohead and the small organic dye Texas Red attached to DNA was utilized to verify ATP-dependent DNA packaging [61].

Recently, Ray *et al.* utilized coupled FCS and spFRET analysis to colocalize the ends of DNA molecules within the T4 phage capsid. Uniform FRET efficiencies between DNA ends in procapsids suggested that packaging motors translocate a DNA loop rather than a DNA end [103].

Helicases: Motor proteins referred to as helicases disentangle dsDNA into ssDNA. They play a crucial role in many molecular processes, such as DNA repair, recombination, replication and replication of viral genomes. A variety of SMFD methods has been used to investigate several types of helicases, producing detailed mechanistic insights into the function of this group of molecular motors [62]. For example, Balci *et al.* recently investigated the conformational states of E. coli Rep helicase undergoing ATP hydrolysis while bound to a partial-duplex DNA via spFRET [105]. Another recent SMFD study investigated the kinetics and thermodynamics of the initiation of RNA unwinding mediated by the viral RNA helicase NPH II utilizing Total Internal Reflection Fluorescence (TIRF) spFRET [63].

Polymerases: RNA polymerase (RNAP) translocates along the DNA template and transcribes it into messenger RNA, thus it initiates the gene expression process. During initiation of transcription, RNAP undergoes a process referred to as "abortive initiation". This process involves the synthesis of short RNA molecules, their subsequent release and the return of the RNAP to the initial promoter site [107][108]. A combined study employed spFRET in a confocal regime with magnetic trapping to examine initiation of abortive transcription by RNA polymerase [109]. Measured data

supported a model where initial transcription occurs via a DNA-scrunching mechanism. This theory involves a stressed intermediate as a driving force for promoter escape via release of the polymerase and binding factors.

The reverse transcriptase of the HIV virus (an RNA-dependent DNA polymerase) converts viral ssRNA into dsDNA. Liu *et al.* employed spFRET to investigate reverse transcriptase sliding on nucleic acid double strands and shuttling between opposite termini of the double strand. When this motor reaches the DNA 3′ end it flips into an orientation suitable for polymerization. Anti-HIV drugs affected the measured sliding kinetics [110].

Cytoskeletal motors

Cells employ molecular motors which translocate along the cytoskeletal network to transport various cargos including mRNAs, organelles and protein complexes. The cargo is transported along microtubule fibers by kinesins, which are moving towards the plus ends of microtubules and dyneins that move towards the minus ends of microtubules. Unlike kinesins and dyneins, myosins transport cargos along actin filaments [64].

Kinesins: Kinesin is an ATP-dependent molecular motor with two heads that transports organelles along the microtubular network in live cells. In very early work on this molecular motor, Vale *et al.* employed a TIRF approach to investigate the motility of individual kinesin motors translocating along microtubules without a cargo [65]. They found that kinesin is a very processive motor, with only a 1% probability of detachment from a microtubule per mechanical cycle. Besides this, they supported the model of a hand-over-hand movement mechanism, as they proved that single-headed kinesins do not exhibit processive motion. Further SMFD study employed single-molecule fluorescence polarization anisotropy and found a mobile ADP-bound state of kinesin, a key intermediate in the molecular walk [113]. Three years later, Yildiz *et al.* employed high-resolution single particle tracking to investigate the load-free translocation of individual kinesin motors with high accuracy [66]. They detected two alternating step sizes, and thus strongly supported the hand-over-hand model against the inchworm model of translocation movement. These experiments were subsequently transferred into a cellular environment. One such experiment detected 8 nm long steps of kinesin-mediated movement of peroxisomes labeled with GFP [67].

Fluorescence Interference Contrast microscopy (FLIC), a technique developed to achieve z-resolution on the nanometer scale [116] was recently used to investigate the microtubule rotations during kinesin-1 mediated cargo transportation [68]. To investigate the microtubule rotation, Nitzsche *et al.* sparsely attached individual quantum dots to the gliding fibers and monitored their lateral and vertical motion in 3D with high accuracy. Their results showed that picking up a large cargo inhibited the microtubule rotation, which was otherwise observed. Nevertheless the large cargo did not affect the velocity of microtubule gliding.

Myosins: Myosins are involved in muscle contraction and also in a variety of other eukaryotic motility processes and cargo transportation.

Force and motility in the myosin-actin complex is thought to be mainly

generated by tilting of the molecular lever arm - the light chain domain of myosin. Forkey *et al.* employed single-molecule fluorescence polarization detection to investigate the extent and timing of tilting motions of brain myosin V in 3D with 20 and 40 ms time resolution. Their data supported the model of lever arm rotation (calmodulin-binding domain rotation) in myosin V. They also supported a 'hand-over-hand' model of mechanism for the translocation of myosin V motors along actin filaments [44].

Foldes-Papp *et al.* recently reported a revolutionary SMFD study of intramuscular actin-myosin complex motion during contraction of skeletal muscle myofibrils [118]. The authors employed single-molecule fluorescence anisotropy to observe the fluctuation in orientation of fluorescently labeled actin within contracted muscle fibril. During contraction, the fluctuating polarized photons arrived with bimodal normal distribution and the fluorescence anisotropy exhibited a slow decay of 9 s. Reported slow relaxation dynamics suggests molecular memory.

Dynein: Dynein molecular motors govern the axonemal motion of flagella and cilia and processively transport cargo along microtubules towards its minus end (i.e. towards the nucleus).

Reck-Peterson *et al.* investigated dynein processivity and stepping behavior. Here the authors employed an objective-type TIRF microscopy coupled with the FIONA method [69] and analyzed data with a step-finding program developed by Kerssemakers *et al.* [120]. The dynein motor walked mostly in 8 nm steps, albeit with different step lengths and backward and sideward steps were also detected. Single dynein domains exhibited a different stepping pattern, this suggested cooperation of two alternating domains which are shuffling between forward and backward positions [70].

The catalytic head domain of the dynein molecular motor is composed of six AAA+ domains arranged in a regular ring. To determine how the two catalytic heads interact and move relative to each other, deWitt *et al.* probed these heads with two spectrally resolved quantum dots and concurrently localized both dynein heads during processive motion. The authors applied the single-molecule high-resolution colocalization technique (SHREC) [122] to achieve 3 nm precise localization. Their results suggested that the catalytic heads remained separated, and that their stepping characteristics varied as a function of distance between heads. Additionally they also suggested that one active dynein head was sufficient for efficient translocation. Moreover an active dynein head was able to pull its inactive partner forward. Hence the authors conclude that the dynein molecular motor does not demand interhead coordination and that during its translocation utilizes a mechanism dramatically different from the hand-over-hand walk of myosin and kinesin [71].

1.4.3. SMFD inside living cells

Current knowledge of structure, dynamics and interactions of biomolecules arises largely from *in vitro* experiments. Generally, experimentalists in the *in vitro* studies try to reconstitute important biological processes using the smallest relevant number of purified (and often modified) components. Consequently, *in vitro* experiments often suffer from losing the complex cellular context. This complex

environment is formed by a multitude of different ions, metabolites, proteins, nucleic acids in addition to compartmentalization and molecular crowding. Thereupon, to validate the biological relevance of *in vitro* detected structures, statistical heterogeneity, stochasticity and kinetics, it is important to verify results via complementary *in vivo* SMFD experiments. SMFD is arguably the best candidate for examining the intracellular events amongst single-molecule techniques. Unlike other single-molecule techniques, such as AFM, electron microscopy or magnetic traps, SMFD can be introduced into the cellular interior without significant perturbation. Moreover, current development in the burgeoning field of genetically encoded fluorescent proteins allows studying of complicated biological processes (such as transcription) without the need of external labeling.

Many questions can be addressed using *in vivo* SMFD, [72] such as: what is the intracellular structure of biomolecules? What are the kinetics and rate-limiting steps of biomolecular processes? How are rate limiting steps influenced by DNA sequence and environmental factors? Are there statistical distributions of biomolecular behavior in cells - sub-cellular stochasticity? How important is stochasticity for components with low abundance, such as genomic DNA or low-copy number RNAs and proteins? Comprehensive information about *in vivo* SMFD can be found in these publications [125-127][128]. We employed spFRET as an intracellular reporter on DNA structure in a proof of concept study (see section 2.5). Further representative publications addressing above mentioned questions are sorted and overviewed below in this subchapter.

Membrane Proteins: The experiments within cell membranes are simplified into a 2D plane and thus are ideal for total internal reflection excitation [73]. This approach greatly improves the signal-to-background ratio. Two main phenomena, which were investigated via SMFD in plasmatic membrane, were lateral diffusion and stoichiometry. For example, Iino and colleagues [74] labeled a transmembrane protein, E-cadherin, with GFP and analyzed diffusing protein clusters, they detected the cluster diffusion coefficient and determined the cluster stoichiometry. Ulbrich *et al.* used the same approach to resolve the composition of NMDA receptors composed of NR1 and NR3 subunits tagged with GFP [75]. Later, Wieser *et al.* published a comprehensive review on the tracking of single molecules in the live cell plasma membrane [76]. Halemani *et al.* employed the high resolution technique, Stimulated Emission Depletion microscopy (STED) and comparative Fluorescence Recovery After Photobleaching (FRAP) to study structure and dynamics of the SNARE complex forming clusters in plasma membrane of PC12 cells [133].

Cytoplasmic Proteins: Cytoplasmatic protein translation occurs in the 3D volume, therefore it is much more difficult to probe compared to a 2D membrane system. This difficulty can be circumvented by a clever simplification, as Yu *et al.* did. They examined the gene expression rate of the membrane protein Tsr (tagged with fast maturing yellow fluorescent protein Venus), revealing that it appears in bursts with Poissonian distribution, i.e. more than one protein was translated from one mRNA [134]. The protein was produced in cytoplasm. Nevertheless after few minutes, it was transferred and incorporated into membrane, making it immobile and thus easily

detectable.

Molecules might also become immobile after binding to chromatin. Recently, Vukojevic *et al.* studied binding of synthetic Sex combs reduced (Scr) peptides to DNA using FCS and fluorescence imaging. They detected the kinetics of Scr-DNA interactions in live salivary gland cells, suggesting that the Scr find their specific target sites by multiple association/dissociation events [135].

However, the majority of proteins are not, anytime within their lifespan, tagged to a cytoplasmatic membrane or chromatin. Many groups have recently attempted to resolve the technically difficult task of tracking molecules in the 3D volume. For example, Sun *et al.* developed a method named Parallax to study the 3D trafficking of glucose-transporter 4 containing vesicles in living adipocytes [77]. One year later, Lew and colleagues characterized the localization precision of a novel method for 3D super-resolution imaging. This method involves examination of a double-helix point spread function [137]. Subsequently, they verified the applicability of the method in biological systems by tracking the movement of quantum dots in the cytoplasm of live COLO205 cells.

Molecular transport: Transport between the nucleus and cytoplasm passes through nuclear pores, with a depth of approximately 120 nm [78]. Thus, the transport happens over distances smaller than the conventional optical resolution. Grunwald *et al.* recently developed a fluorescence-based method to monitor the bidirectional transport. Applying this method the authors achieved 20-ms temporal and 26-nm spatial precision, enabling the capture these highly transient interactions in live cells. The authors then detected the kinetics of mRNA transport in mammalian cells and presented a three-step model consisting of docking (80 ms), transport (5–20 ms) and release (80 ms) [79].

Recently, Schuster *et al.* employed single particle tracking to investigate the ATP-driven bidirectional transport of early endosomes along microtubules in live fungal (Ustilago maydis) cells. They fused the endogenous dynein heavy chain, kinesin-3 and small endosomal GTPase Rab5a with fluorescent proteins and monitored both motor proteins carrying their cargo in this fungal cells. Analysis of tracking revealed that binding of retrograde moving dynein to anterograde moving endosomes alternates the transport direction and when dynein releases the early endosome, the organelles switch back to anterograde kinesin-3-based motion. Using quantitative photobleaching the authors proved that one dynein motor and four to five kinesin-3 motors bind to an early endosome [140].

RNA transcript: Bertrand and colleagues were pathfinders in the detection of RNA molecules in live cells [80]. RNA cannot yet be easily modified with endogenous fluorescent probes (such as GFP), therefore their fluorescent labeling must be circumvented. Bertrand *et al.* first came with the idea to utilize the bacteriophage MS2 protein's high specificity and affinity to a small RNA hairpin. MS2 proteins, fused with GFP, can be expressed in the presence of an RNA transcript with small MS2-binding hairpin repeats added. Thus, the immediate and strong binding of the hairpin and fluorescent construct allows localization and tracking of the RNA transcript and dynamics in the eukaryotic cell in real time.

Yunger *et al.* coupled the MS2-GFP system, FRAP and single particle tracking to monitor mRNA transcriptional kinetics of single-gene integrations in real time with 200 ms resolution [81]. One year later, Nevo-Dinur *et al.* utilized the MS2-GFP system to investigate the localization of mRNA in E. coli. They suggested that mRNAs are targeted to the future destination of their encoded proteins (cytoplasm, poles or inner membrane) in a translation-independent manner [82].This means that mRNAs can travel to a location where their encoded proteins are needed. As an example, they proved that *cis*-acting sequences of the membrane proteins are an indispensable and sufficient factor for mRNA membrane localization. Further, Treutlein *et al.* also utilized the MS2-GFP system and directly monitored the dynamics of mRNA transcription initiation, elongation and termination from an endogenous gene in live yeast cell [83].

Virus tracking: Viruses were tracked as single particles in multitude of live cells [84]. Recently, Van der Schaar *et al.* tracked single Dengue viruses (DENV) in live HeLa cells. They examined the cell entry pathway, endocytic trafficking, and fusion of DENV and proved e.g. that DENV enters cells solely via clathrin-mediated endocytosis. They also located the fusion of the viral and endosomal membrane to late endosomal compartments [146].

1.5. Principles of SMFD

Main task of SMFD in biology is to optically isolate and then examine individual molecules. These two conditions are accomplished by two ostensibly simple steps. The first step is to assure maximally one molecule can fluoresce in the probed volume at a time. The second condition can be fulfilled by providing sufficient signal-to-noise ratio (SNR) and signal-to-background ratio (SBR) for the single-molecule fluorescence signal in a reasonable averaging time.

To assure that only one molecule is excited and subsequently emit fluorescence at a time, SMFD usually combines use of an ultralow concentration of probed molecules in ultraclean solvents in minimized probing volume. The minimization of probing volume is usually achieved by the confocal principle, by the total internal reflection phenomenon or using the zero-mode waveguide (ZMW) approach (see below). The concentration required depends upon the volume probed by the pumping light to achieve the single-molecule limit. A scheme describing the photoselection process is shown in Figure 3.

Required SNR and SBR can be achieved by maximizing the fluorescence signal while minimizing background sources and the relative size of laser shot noise [85]. To maximize the fluorescence signal, a combination of several photophysical probe properties are needed. It is highly advantageous to have large absorption cross section, high photostability, high quantum yield, weak bottlenecks into dark states (as triplet states or *cis-trans* state switching), and excitation power below saturation level of the molecular absorption. Individual molecules can be investigated either freely diffusing or immobilized to a surface or cellular component.

Detection of freely diffusing molecules: Detection of freely diffusing molecules is arguably the best SMFD approach for fast monitoring of statistical distributions of the properties of individual biomolecules [148], as it probes large sets of molecules in a short time. Nonetheless, in this type of experiment long-range dynamics of one molecule cannot be directly followed. Generally, experiments on freely diffusing molecules minimize the structural and functional perturbation of host molecules, because they do not need any modification of the biomolecule with an anchoring system. In this approach, individual molecules are probed whilst freely diffusing through a small well defined excitation volume.

The excitation volume is most often realized via a tightly focused pumping beam in a highly diluted sample. A confocal optical set-up is usually employed to limit the observation volume in a sample and to define its size and shape. The confocal volume is typically on the order of femtoliters. This is attained by adding a "pinhole" in the image plane of the optical set-up which then removes out-of-focus photons [86]. Thus in roughly picomolar concentration, when an individual molecule of interest travels through the observation volume a burst of photons is emitted. This burst is then detected with a single-photon counter, such as a single-photon-counting avalanche photodiode or a single-photon-counting photomultiplier [87]. A comprehensive analysis of the statistics of fluorescence bursts was published by Enderlein *et al.* [88] The disadvantage of this set-up comes with the need for low concentration of labeled sample. Using higher concentrations is burdened with a too high background.

This flaw can be circumvented by the use of ZMW. This approach has a promising future, since the majority of biological systems associated with enzymatic catalysis or molecular binding require very high concentrations of active molecules, usually between micro- and millimolar concentration [89]. Thus, to perform single-molecule experiments under physiological conditions the observation volume must be decreased by 3–6 orders of magnitude compared with a diffraction limited set-up. ZMWs provide this reduction, as they consist of sub-wavelength holes with typical diameters in the 30–300 nm range, manufactured in a thin metallic coating on a glass substrate (Figure 3). ZMWs make use of the evanescent field generated at the aperture's entrance. The tiny cross-section of the aperture and even smaller depth of the evanescent field ensure that the observation volume is in atto- to zeptoliters. The presence of thousands of apertures on a single substrate enables parallel use of multiple ZMWs during a SMFD experiment and thus parallel detection of individual molecules [153]. This and the open configuration, allowing for an exchange of buffer, make the ZMW a suitable method for high throughput measurements, such as single-molecule real-time DNA sequencing [90].
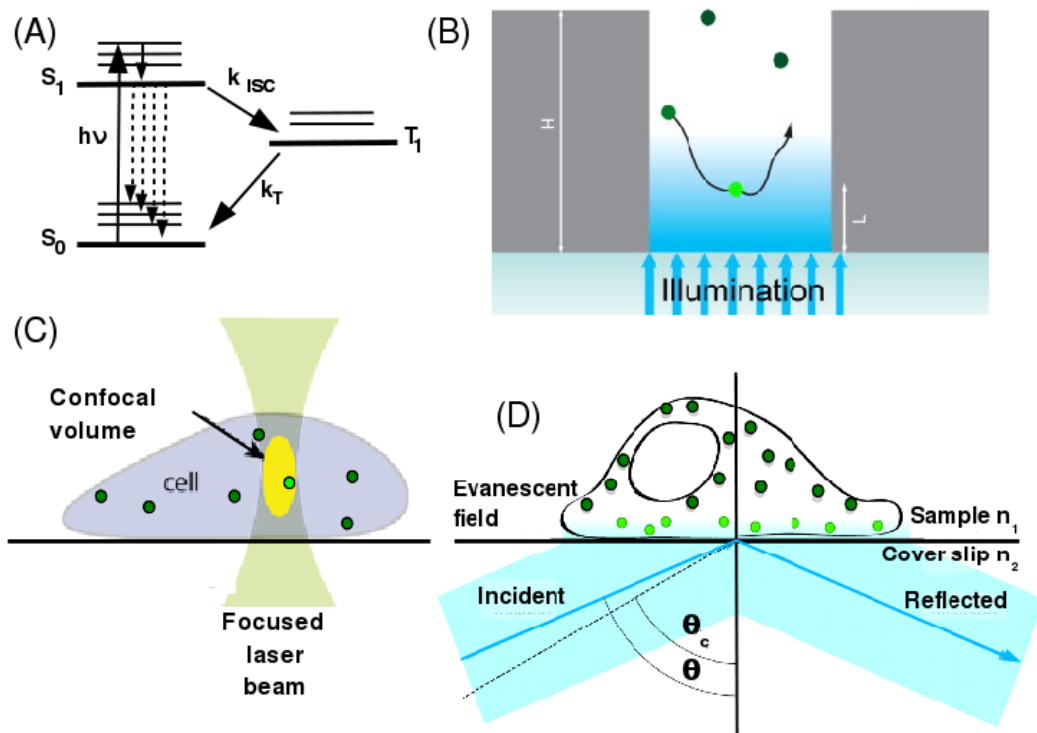
Figure 3: Scheme of photoselection process.

(A) Energy level scheme of fluorescence. S_0 depicts ground singlet level; S_1 depicts first excited level and T_1 is the lowest triplet level. For each of the three electron levels, several vibrational levels are shown. Pump light at energy hv excites the electron at ground level to the first excited level via dipole-allowed singlet–singlet transition. The forbidden intersystem crossing rate is depicted as k_{ISC}, and the triplet decay rate as k_T. Fluorescence emission is shown as dotted lines. Adopted from Moerner et al. [85]. (B) Implementation of zero-mode waveguides. A laser generates an evanescent field at the aperture's entrance and defines the atto- to zeptoliters observation volume. Only fluorophores diffusing into the evanescent field emit fluorescence, as indicated by the bright green color. Adopted from Moran-Mirabal et al. [153]. (C) Confocal observation volume. The light beam is focused in a small volume of the specimen. When a fluorophor enters this volume it can be excited and emit fluorescence (bright green circle). (D) TIRF illumination scheme. The excitation light beam comes at incidence angle (θ), which is bigger than the critical angle (θc), given by Snell's law. All angles are measured from the normal. The excitation beam is reflected off the interface between the coverslip and sample, generating an evanescent field in the sample side. Only fluorophores in the small volume defined by the evanescent field are excited and emit fluorescence (depicted as bright green circles). The refractive index at the sample side (n_1) must be smaller than the refractive index of the coverslip (n_2) to achieve total internal reflection. Adopted from Mattheyses et al. [155].

Detection of immobilized molecules: Many of the important biological reactions occur in the time range from milliseconds to seconds. Immobilization of biomolecules or molecular complexes allows acquisition of long signal time traces, providing a record of several repetitions of the probed reaction with precious information about its kinetics and dwell times. The most common immobilization techniques are described in section 1.7.3.

There are two basic approaches used to study individual tethered biomolecules, confocal scanning microscopy and TIRF microscopy (see Figure 3(C) and (D)). Confocal scanning microscopy is a variant of confocal microscopy where molecules are immobilized and the surface is scanned to monitor fluorescent biomolecules [87].

In TIRF microscopy [91] an excitation light beam is totally reflected at the interface between the material of the optical element and the water buffer, subsequently a very thin evanescent field is induced in the buffer. The evanescent field is usually formed using one of two optical elements, a TIRF prism [92] or a microscope objective (with numerical aperture of 1.45 and higher) [93]. The evanescent field in TIRF microscopy pumps fluorescent probes either at the interface or within ~100 nm of the glass surface. The intensity of the evanescent field (I) at any distance from cover slip (z) is described by:

$$I(z) = I_0 e^{-zd}$$ (Eq.1)

Here I_0 is the intensity of the evanescent field at the surface (z=0). I_0 is related to the intensity of the incident beam by a complex function of its incident angle and polarization. The depth of the evanescent field (d) refers to the distance from the interface at which the excitation intensity decays to 37% of I_0 and depends on the wavelength of excitation light, excitation light beam incidence angle and refractive indices of the cover slip and sample [94][155].

There are two main benefits of TIRF illumination over confocal. Firstly, TIRF is better if the system under investigation involves the interaction between immobilized molecules and a high concentration of freely diffusing fluorescent molecules. In this situation a confocal set-up would introduce a large background signal whereas the same experiment illuminated by TIRF is almost background free [160]. A second advantage of TIRF microscopy is the possibility to excite fluorescent probes in a large area on the surface at once. The fluorescence from all these molecules can be detected on a CCD camera, and thus a large number of time-trajectories can be quickly obtained in parallel. However, the time resolution of a CCD camera is significantly lower than that of avalanche photodiodes or photomultipliers. This limits the experiment to events slower than milliseconds.

1.6. SMFD techniques

Ever since their introduction, SMFD techniques have evolved and branched out to address various biological questions, which are otherwise inaccessible using ensemble methods. The current most common SMFD methods have the capability of probing dynamic biomolecular interactions via spFRET and fluorescence lifetime; tracking single objects over micrometer distances with nanometer resolution; and monitoring the translational and rotational motion of several systems via FCS [161]. The basic features of the most common SMFD techniques are listed in the following subchapters.

1.6.1. Fluorescence lifetime

The fluorescence lifetime refers to an average period of time the fluorescent probe stays in its excited state before emitting a photon, i.e. undergoes relaxation to the singlet ground state (Figure 1(a)). Recently the most common lifetime measurement techniques have involved repetitive pumping of the probed molecule with short pulses

from a light source such as a pulsed laser diode or a mode-locked laser. After each excitation pulse, at most one photon can be emitted from a single molecule. Photon arrival times are then detected on the single-photon counters and statistically processed to give time histograms. These histograms show fluorescence intensity decay, which can be mathematically represented as a sum of exponentials:

$$F(t) = \sum_i a_i \exp\left(\frac{t}{\tau_i}\right). \qquad (Eq.2)$$

In the above, F(t) is fluorescence intensity at time t, τ_i is the lifetime of component i and a_i its intensity contribution to the fluorescence decay. The measured fluorescence intensity decay data is given by the convolution of the instrument response function IRF(t) with the excitation light pulse. Thus the total intensity measured I(t) is:

$$I(t) = \int_0^t IRF(t-\tau) F(\tau) d\tau. \qquad (Eq.3)$$

The fitting of (Eq.2 convoluted to an instrument response function to experimental data is usually done by iterative convolution. A guess of the coefficients a_i and τ_i, is used to compute F(t), which is convoluted with the experimentally determined instrument response function IRF(t) to obtain I_{model}. Minimization of the difference between the I_{model} and the I(t), allows to find the most probable lifetimes. [95] An alternative imaging based technique for lifetime detection is called fluorescence-lifetime imaging microscopy (FLIM) [163].

An important advantage of the fluorescence lifetime lies in its independence from the initial fluorescence intensity. Thus, compared with purely intensity based techniques it has better tolerance to signal noise. Therefore fluorescence lifetimes have often been employed for robust detection of FRET [96] and fluorescence anisotropy [97][166]. Moreover, detection of lifetime shortening was also employed in exploring tRNA dynamics, [98] DNA structural fluctuations, [99] and electron transfer at interfaces [100].

1.6.2. Fluorescence Correlation Spectroscopy

FCS monitors the spontaneous fluctuations of a system as detected by fluctuations in fluorescence intensity in a microscopic detection volume (see Figure 4(b)). FCS enables the study of several biophysical phenomena. The broadest applicability in biology has been found in the detection of translational diffusion motion. However other phenomena, such as fluorescence blinking, rotational orientation motion and photon antibunching might also be utilized in biomolecular research (see below).
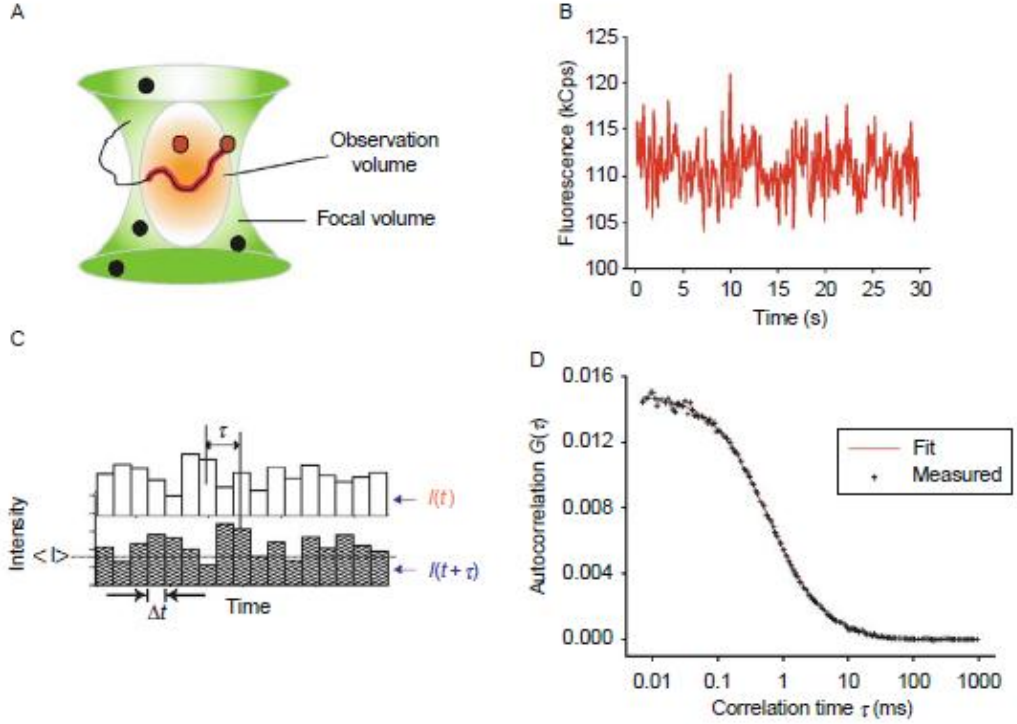
Figure 4: Fluorescence fluctuation spectroscopy

(A) Fluorescent molecules (black filled circles) diffuse through the focus (yellow) of a laser beam (green) where they are excited (red filled circles). Diffusion of molecules across the focal volume causes fluctuations in the measured fluorescence intensity signal, additional fluctuations in the same time scale might arise from molecule conformational kinetics. (B) A representative time trace of the detected fluorescence intensity fluctuation. (C) The fluorescence intensity time trace is binned into discrete time bins (Δt). The autocorrelation function (ACF) is computed from the signal fluctuation about its mean intensity <I> at time (t), compared with that at a time delay τ later (t + τ) (D) The ACF as a function of the logarithm of the lag (delay) time τ. Adopted from Chen et al. *[101].*

In FCS, the probing volume is most often realized via a confocal set-up. Molecules of interest diffuse across the probing volume and undergo excitation-emission turnovers. The temporal measure of fluorescence intensity fluctuations in confocal volume is examined by signal autocorrelation (see Eq. 4 and Figure 4(d)).

$$G(\tau) = \frac{\langle \delta I(t)\delta I(t+\tau)\rangle}{\langle I(t)\rangle^2} \qquad (Eq.4)$$

The similarity of the fluorescence intensity fluctuations $\delta I(t)$, given by $\delta I(t) = I(t) - \langle I(t)\rangle$, at time t and by $\delta I(t + \tau)$ at a time delay τ later is analyzed and normalized by the mean time averaged fluorescence intensity $\langle I(t)\rangle$ (Figure 4(c)). If the system is at equilibrium, the result of an autocorrelation analysis - autocorrelation function $G(\tau)$

(ACF) depends only on the time lag τ. Otherwise, for systems evolving in time we can easily collect a set of ACFs during the system progression and analyze the kinetics of its evolution. We utilized this approach, coupled with singular value decomposition and regularized distribution fitting to monitor the assembly of MS2 bacteriophage capsid. FCS revealed a gradual decrease in concentration of initial assembly components (capsid protein dimers and trRNA), a simultaneous increase in concentration of fully assembled capsids and a change in loading of assembly intermediates, for details see section 2.4.
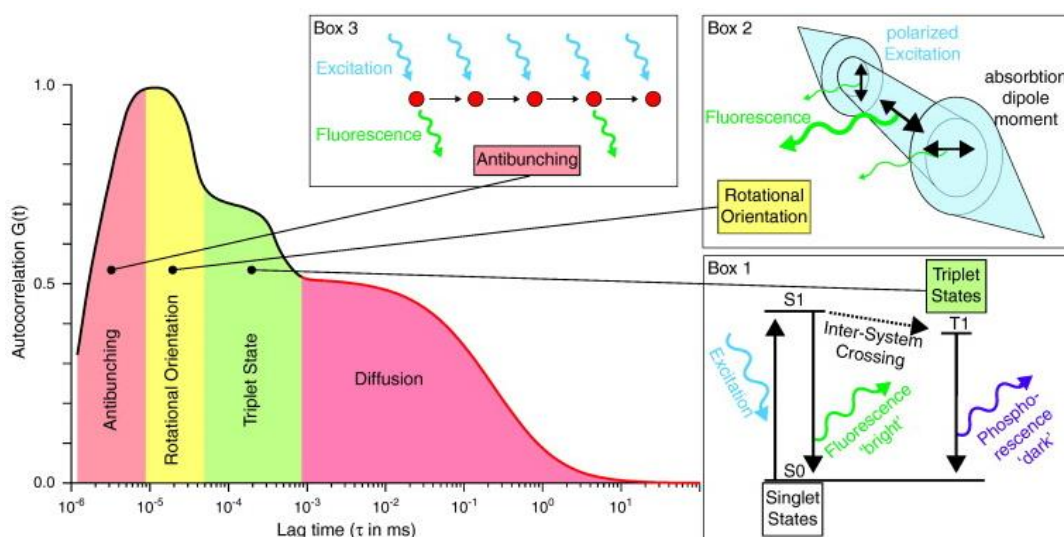

Figure 5: Scheme of complete autocorrelation curve.

Red region covers translational diffusion, green region shows triplet-state transitions (shown in Box 1), yellow region indicates rotational diffusion (shown in Box 2), yellow region covers fluorescence anti-bunching phenomenon (shown in Box 3). Adopted from Fitzpatrick et al. [171].

Information about involved phenomena can be extracted by fitting analytic models to the measured ACF in different time regions. An entire autocorrelation curve usually covers a nanosecond to second time scale. Fortunately, all basic phenomena extractable from ACF have well defined time regions (Figure 5), these are listed below.

Translational diffusion: Monitoring of translational motion and motion-derived parameters of biomolecules is the most common biological utilization of FCS. Due to its fine spatial resolution it is applicable to follow motion in live cells. Different types of motion and transportation were monitored using FCS, e.g. free diffusion [102], active transport [173], anomalous subdiffusion in crowded cellular cytoplasm [103] and systematic movements of glycoproteins in membrane [104].

In addition to direct characterization of biomolecular motion, FCS provides an estimate of the hydrated size of freely diffusing biomolecules. The Stokes radius (R_H) is, strictly speaking, the radius of an ideal hard sphere diffusing at identical speed as the probed molecules. Because biomolecules are not ideal spheres, the R_H is usually

smaller than the real hydrated radius of an examined molecule. In solutions with substantial friction, the Stokes' radius depends on viscosity and temperature via following equation.

$$R_H = \frac{k_B T}{6\pi\eta D}$$
(Eq.5)

In the above, η is viscosity, k_B is the Boltzmann constant, D is the diffusion coefficient and T is the temperature. However, to follow the changes of biomolecule, it is not necessary to precisely determine the actual size. For example, Sababayagam *et al.* employed detection of a change in the diffusion coefficient to monitor the ATP dependent packaging of labeled DNA oligonucleotides into a bacteriophage T4 prohead [105]. The diffusion coefficient derived from ACF analysis decreased from the original 11.7 $\mu m^2\ s^{-1}$ for the freely diffusing DNA to 0.66 $\mu m^2\ s^{-1}$ after packaging. This drop in diffusion velocity was caused by change in the hydrated radius of the labeled system and allowed the direct observation of the packaging process.

Detected diffusion might be also be related to molecular interactions. For example, a variant of FCS, namely Total Internal Reflection (TIR)-FCS was used to examine the binding of IgG antibody to the Fcg receptor-2 incorporated in a lipid bilayer [106]. Lieto *et al.* quantified bound verses unbound ratios as well as the binding constant of the interaction (2.4 - 0.4 mM). Another group employed a different variant of FCS, dual-color fluorescence cross-correlation spectroscopy, to monitor the binding of different pairs of MAP kinase signaling proteins [178]. Observed cross-correlation helped the authors to map the dynamic binding interactions of the MAP kinase signaling in living yeast cells.

Fluorescence blinking (fast flickering or intermittence) covers a time range of 10^{-7} - 10^{-6} s. Fluorescence blinking is characterized by fast reversible fluctuations in fluorescence quantum yield (on and off states). Known mechanisms behind fluorescence blinking involve intramolecular electron transfer [107], *cis-trans* photoisomerization of fluorophores e.g. cyanine dyes [108] and light-induced protonation of GFP [109] or dynamic spFRET [182]. However, the most common mechanism of fast blinking is the intersystem crossing (triplet state formation) which is common in most dyes under high intensity of excitation light. The triplet state lifetime is usually much longer than the fluorescence lifetime, resulting in off states lasting up to microseconds. The intersystem crossing parameters depend not only on the intensity of the excitation source, but also on the local environment of the fluorescent probes.

Environmental influence, such as the presence and concentration of molecular oxygen and heavy metal ions, effectively quenches triplet states and alters its kinetics [183]. Another interesting feature of blinking is the sensitivity of several fluorophores to local pH. These features make fast flickering a useful reporter for an intracellular environment. Well known single-molecule spectroscopic pH reporters are fluorescent proteins. The most pronounced pH sensitivity of fast flickering was reported in enhanced GFP protein (EGFP) and yellow fluorescent protein (YFP). Both these proteins exhibit pH-induced spectral shifts in microsecond to milliseconds time scale, causing temporary dark states, easily detectable by FCS [109]. Haupts *et al.* identified a reversible protonation of the chromophore of EGFP as a reason for the spectral shift

of fluorescence emission [110].

Since the protonation kinetics depends only on the local pH, EGFP and YFP can be used as sensitive, intrinsically calibrated pH reporters. Fast flickering of EGFP can also serve as molecular thermometers because its reversible protonation is strongly temperature-dependent at low pH [111].

Rotational orientation (aka rotational Brownian diffusion) appears typically in the time region from 10^{-8} to 10^{-7} s. As the fluorescent probe is excited preferentially with light of parallel polarization to its absorption dipole moment correctly oriented molecules have a larger probability to fluoresce. Emitted fluorescence photons are polarized in the direction of the probe's emission dipole. Hence, if the molecules rotate within the fluorescence lifetime of the probe, the orientation of emitted photons depends on the velocity of this rotation. Therefore, fluctuation of the fluorescence signal caused by molecular rotation appears only when at least one of the two following conditions is fulfilled. Either the molecules of interest are pumped with light of linear polarization at low intensity and/or the fluorescence is monitored in a polarization-dependent fashion [112]. For example this phenomenon might be used to estimate the overall shape of examined molecules and to follow its conformational change. However its reliable interpretation might be extremely complicated, depending on the type of excitation and detection and the shape of probed molecule [113].

Photon antibunching covers a range of very short times, approximately from 10^{-9} to 10^{-8} s. As mentioned above, one molecule can only emit at the most one fluorescence photon per one short exciting pulse. This phenomenon is called photon antibunching. The quantity of fluorophores in the probing volume can thus be determined due to proper statistical analysis. A typical application in biology is monitoring of biomolecular oligomerization [188].

1.6.3. Förster Resonance Transfer

FRET is strictly not a technique but a photophysical phenomenon, which has been widely employed to measure distances on a nanometer scale in many applications [114]. FRET is a distance-dependent nonradiative energy transfer from an energy donor fluorophore to an energy acceptor fluorophore. An energy donor fluorophore, in an excited state, may transfer energy to an energy acceptor fluorophore in the ground state. (Figure 6(a)) The transfer is limited within an interprobe distance between 2 and 10 nm via a non-radiative dipole–dipole coupling mechanism. Experimental FRET was first reported in the 1920s by Cairo, [115] but it was not fully elucidated up to the 1940s, when Förster published the first comprehensive theory [116, 117]. A substantial increase in interest came at the end of the 20th century with the merger of FRET and optical microscopy, facilitating the visualization of the *in vivo* macromolecular interactions. The rate of nonradiative energy transfer k_T is given by following equation.

(Eq.6)

$$k_T = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6$$

Here τ_D refers to the donor fluorescence lifetime, R_0 is the Förster radius - the distance

between energy donor and acceptor, at which the FRET efficiency is 50% (Figure 6(b)) - and r is the measured interprobe distance. The efficiency of the FRET (E) is defined by the ratio of the energy transferred from donor to acceptor to the sum of energy absorbed by the donor.

$$E = \frac{k_T}{k_T + k_D + \sum_{i \neq T,D} k_i} = \frac{R_0^6}{R_0^6 + r^6}$$

(Eq.7)

Here $k_D = 1/\tau_D$ refers to the rate at which donor emits photons and k_i refers to the rates of all other de-excitation pathways. The most important parameter for accurate distance determination in FRET analysis is R_0. It depends on several spectral parameters of probes used, such as the overlap of normalized emission spectrum of donor and the absorption spectrum of acceptor, the mutual orientation of the donor and acceptor transient dipole moments and the quantum yield of donor. Also, refractive index as an environmental factor affects the distance range of the interaction. A typical value of R_0 for organic dyes and fluorescent proteins is around 5 nm. A reasonable range of detectable distances is usually defined within $0.6–1.6 \cdot R_0$, (Figure 6(b)) but due to experimental limitations such as noise the useful range is usually smaller (e.g. 4–7 nm). Extensive lists of R_0 values for *D–A* FRET pairs are available [118].

An accurate determination of distance is often a complicated and technically demanding task, [194] nevertheless the occurrence of FRET always provides proof that the donor probe and acceptor probe associate closely. Thus, FRET is without doubt a suitable tool for detection of events such as the macromolecular binding, conformational change or stepping behavior of molecular motors.
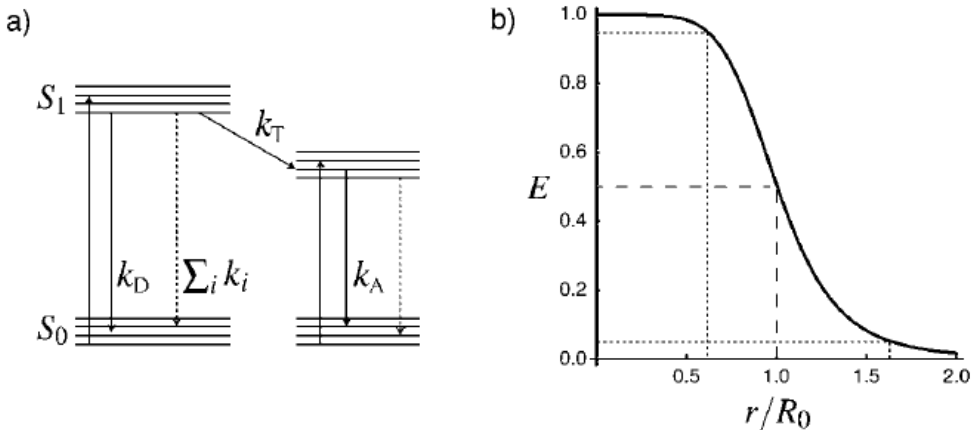

Figure 6: Basic principles of fluorescence resonance energy transfer.

a) Jablonski diagram of the most common processes in a FRET pair: An excited donor may relax to its ground state by fluorescence with a rate k_D, or by several other de-excitation processes with rates k_i. A donor can also non-radiativelly transfer its energy to an acceptor with a rate referred as k_T. An acceptor can fluoresce that energy with rate k_A. b) FRET efficiency (E) as a function of r/R_0, (normalized interprobe distance). Förster radius R_0 is defined as the distance where E=0.5. The employable distance region is in the range $r/R_0 \approx 0.6$ to $r/R_0 \approx 1.6$, corresponding to efficiency E=0.95 and E=0.05 respectively. Inside these borders, FRET efficiency has a steep distance dependency and FRET changes can be detected with high reliability. Adopted from Grecco et al. [195].

Methods to measure FRET: FRET rate k_T, cannot be directly determined at the single-molecular level. However, the efficiency can be monitored for example via measurement of quantum yield or fluorescence lifetime of participating fluorophores [119][120][198][6][121].

As implied from (Eq.7), the leaking of energy of excited donor through FRET means that the quantum yield and fluorescence lifetime of the donor decreases since a substantial amount of energy is transferred non-radiativelly. Within the same process the acceptor fluorophore is excited via FRET and subsequently releases part of that energy as fluorescence. Hence, FRET can be monitored via decrease in donor emission spectrum and subsequent increase in acceptor emission spectrum of individual molecules. This might be achieved through direct detection of full emission spectra or more simply via detection of fluorescence intensity at chosen wavelengths. The simpler variant can be easily detected using a CCD camera with two spectrally separated channels which are reflected to different regions of CCD chip [122].

The limitation of this technique appears when the relative concentrations of donor and acceptor vary. This might be surpassed via careful correction of spectral cross-talk [201, 202]. Moreover FRET analysis based on detection of both spectral channels frequently suffers from noise in the acceptor sensitized emission. To overcome this flaw, it is possible to detect donor fluorescence in presence of acceptor, subsequently photobleach the acceptor probe and thus obtain an unquenched reference for final analysis [123][124].

Since FRET decreases the lifetime of the excited state of the donor and modifies the excited state lifetime of the acceptor, fluorescence lifetime is often used as a robust, intensity independent FRET reporter. A lifetime approach in the form of FLIM has recently seen a steady increase in use, fostered by its reliability and quantitative possibilities [125][126].

FRET can be also detected by monitoring the depolarization of the energy of the acceptor emission via fluorescence anisotropy. This approach is based on the fact that the depolarization might occur only within the time window when the energy transfer occurs [127]. Polarization anisotropy makes possible monitoring of FRET between fluorescent probes of the same type [128]. This allows for example monitoring of protein clusters formed by proteins of the same type [209].

FRET fluctuation spectroscopy is an excellent tool for detecting FRET kinetics [67]. It couples FRET and FCS, utilizing simultaneous analysis of the ACF and crosscorrelation function of FRET probe pair within investigated biomolecules. This technique provides a robust detection approach, independent of the diffusion contributions and without artifacts related to dye photophysics and changes in the observation volume between measurements [129].
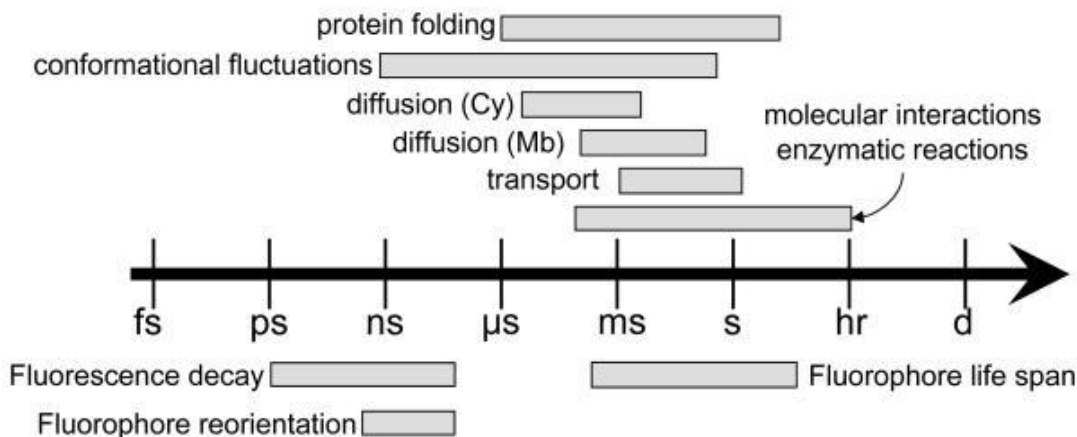
Figure 7: Approximate time scales of biological phenomena studied via SMFD.

SMFD techniques are perfectly suited to investigate diverse biophysical phenomena such as conformational fluctuation, protein folding, diffusion in the membrane (Mb) in cytoplasm (Cy), motions of molecular motors, enzymatic dynamics and protein interactions. The short fluorescence lifetimes of common fluorescent probes allows detection of fast fluctuations due to its sensitivity to non-radiative de-excitation pathways such as FRET or charge transfer, while slower processes can be monitored by imaging, correlation methods and intensity fluctuations. Adopted from Michalet et al. [213].

Model-free data analysis: FRET data usually does not demand any complicated model to extract useful information as compared with for example FCS [161]. Nevertheless, the FRET time trace consists of fluorescence signal from different FRET states as well as from noise of an instrumental, statistical, or spectroscopic nature. Thus, complex statistical analysis should be applied to extract reliable information. For example Kerssemakers *et al.* adopted an automated step-finding algorithm to analyze the FRET time trace of DNA unwinding by an NS3 helicase [130]. McKinney *et al.* used hidden Markov model analysis to study even more complex behavior, such as bidirectional transitions [131]. They used this algorithm to examine binding and dissociation of RecA protein monomers on DNA [62].

1.6.4. Brief list of further SMFD techniques

Fluorescence lifetime, FRET and FCS are the most common among SMFD techniques. However, there are several other useful methods, which are not so widespread, mainly because of high technical and interpretation demands or expensive instrumentation. This group includes the diverse field of super-resolution light microscopy methods, photon-counting histograms, polarization imaging etc.

Super-resolution light microscopy is a variant of light microscopy with enhanced spatial resolution. The spatial resolution of conventional light microscopy is defined by the diffraction of light as stated by Ernst Abbe in 1873 [132]. In general, the full width at half-maximum of detected point spread function is used as an approximation of the spatial resolution. In another, in terms of SMFD a more suitable definition, resolution is stated as the shortest distance between two points (individual molecules) on a

specimen that can still be distinguished by the detection system as separate entities. High numerical aperture microscope objectives in a wide-field optical microscope configuration usually provide the resolution of approximately 250 nm. Super-resolution techniques offer substantially higher resolution than that, defined by a diffraction limit, i.e. they localize individual molecules with much higher precision.

There are two main groups of super-resolution light microscopy methods. The first group contains deterministic techniques; these are usually expensive and technically demanding, therefore used only by a limited number of experimentalists. Several of these techniques are based on the nonlinear response of fluorophores to excitation, e.g. stimulated emission depletion [133] and ground state depletion microscopy [134].

The second group covers stochastic techniques. Reversible fluctuation between dark and bright states is an inherent property of individual fluorophores. Moreover, it can be chemically and optically enhanced. When the fluorophores at close vicinity (below diffraction limit) undergo unsynchronized blinking, they fluoresce at separate times, and thus become resolvable in time. Due to the fit of a fluorescence image of time-separated individual molecules to 2D Gaussian function; spatial resolution up to one nanometer can be achieved. Another, significantly quicker and almost equally precise (~5 nm) algorithms are based on algebraic centroid estimations. These algorithms are moreover less sensitive to low SNR [135][136].

The stochastic methods include super-resolution optical fluctuation imaging (SOFI) [219] and purely single-molecule localization methods such as photoactivated localization microscopy (PALM) [137], stochastic optical reconstruction microscopy (STORM) [138], single-molecule high-resolution imaging with photobleaching (SHRIMP) [139] and fluorescence imaging with one nanometer accuracy (FIONA) [223].
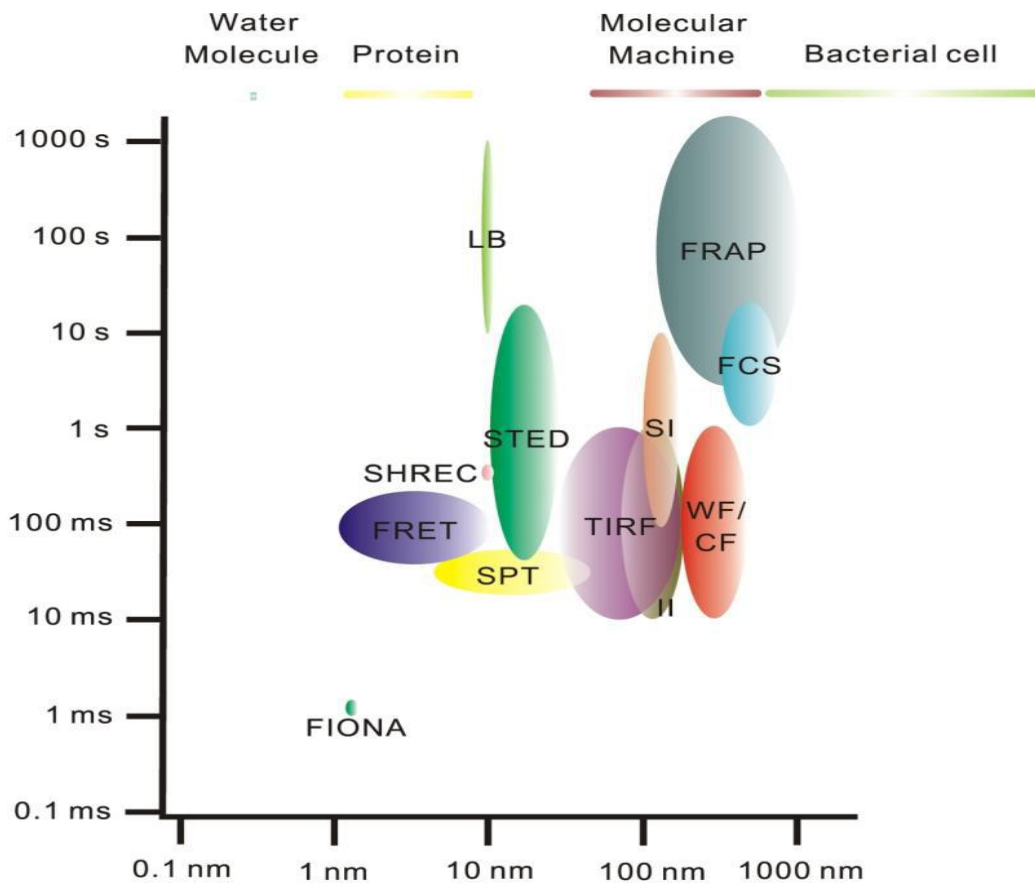
Figure 8: Comparison of the spatial and temporal resolutions of SMFD techniques.

Both axis have logarithmic scales. The typical size of a water molecule, protein, molecular machine and bacterial cell are shown as the top panel. The horizontal bright side of each oval shape, representing one SMFD technique, approximates the z-localization accuracy and the dark side approximates the x-y-localization accuracy. For example, WF/CF and TIRF have very alike x-y-resolution; however TIRF has a significantly better z-resolution. For FRAP and FCS the horizontal scale shows the limiting size of the probed volume. The vertical (time) axis refers to the acquisition time of one measurement; the reciprocal value of this time defines the maximal sampling rate of detection. Used abbreviations refers to: fluorescence correlation spectroscopy (FCS), fluorescence imaging with one nanometer accuracy (FIONA), fluorescence recovery after photobleaching (FRAP), Förster resonance energy transfer (FRET), interference illumination (II), localization-based (LB); single-molecule high-resolution co-localization (SHREC), structured illumination (SI), single particle tracking (SPT), stimulated emission depletion (STED), total internal reflection fluorescence (TIRF), wide-field/confocal (WF/CF). Adopted from Chiu et al. *[224].*

Photon Counting Histograms (PCH): PCH captures the statistical distribution of fluorescence intensity fluctuations in an observation volume [140][141]. The detected statistical distribution can be uniquely described with two parameters for each examined species. These parameters are the molecular brightness of the probe and the average number of molecules within the probing volume. Due to its sensitivity to the

molecular brightness, PCH allows separation of a mixture of fluorescent species into initial compounds, based on a different brightness. This ability proved to be useful to the study molecular aggregation [142]. Moreover PCH is a complementary technique to FCS [143].

Alternating Laser Excitation (ALEX): Direct excitation of both fluorophores of the FRET pair allows the detection of stoichiometry and the dissociation constant with high accuracy. Importantly, ALEX also separates a low FRET population from incomplete FRET pairs (donor-only molecules) [229]. To study more complex systems, Lee introduced in 2007 three-color alternating excitation [230]. The described configuration enables precise measurement of FRET between three fluorescent probes, providing the possibility of investigating more complex biomolecular assemblies and structures than conventional two-color FRET.

Polarization imaging: This method monitors the 3D orientation of the fluorophore via analysis of the unique fluorescence patterns of defocused individual molecules upon polarized excitation and detection [144]. This technique has much lower temporal resolution than non-imaging techniques, nevertheless it requires a simpler experimental setup and provides more accurate information on the single-molecular orientation than widely used time-resolved fluorescence anisotropy. Typical SMFD set-up can be easily changed to monitor the "defocused imaging". Toprak *et al.* employed this approach in biology and monitored an emission pattern dependent on the 3D orientation of a myosin construct [232].

Multiparameter fluorescence detection: Several SMFD methods can be combined to harvest useful information about individual molecules. Widengren *et al.* took this idea to a new level and measured fluorescence anisotropy, fluorescence lifetime, fluorescence intensity, excitation spectrum, fluorescence spectrum, fluorescence quantum yield at once on individual molecules freely diffusing in solution [233].

1.7. Sample treatment

Individual fluorophores are known for their various inherent sensitivity to the local environment, including pH, viscosity and polarity of solvent, vicinity of fluorescence quenchers, such as guanine [145] and tryptophan [146]. Hence, the correct choice of environmental conditions and fluorescent probe can prevent several interpretation artifacts. Sensitive detection of individual molecules demands specific spectroscopic parameters of the fluorescent probe. Moreover, the labeling of probed biomolecules should not perturb its structure and function, while preserving close and well-defined probe attachment.

Individual molecules also inherently exhibit photobleaching and blinking. These phenomena complicate data interpretation, thus it is beneficial to avoid them, using anti-photobleaching and anti-blinking cocktails. Moreover, monitoring of single-molecules over long time ranges often requires non-perturbative immobilization, which usually demands passivation of the cover slip surface. All these tasks and issues are discussed below in separated subchapters.

1.7.1. Labeling

SMFD of labeled biomolecules is possible only due to thousands to millions of light induced excitation – de-excitation cycles between the ground state and the first excited electronic state (see Figure 3(a)) [236][147]. Therefore probes used for SMFD should have high brightness, large Stokes shift, high photostability, as well as additional properties specific to SMFD [42]. Important parameters of fluorescent probes and several site-specific labeling strategies are listed below.

Probe parameters
Singlet-state saturation: One of the most important bottlenecks for repeated excitation of a fluorescent probe is the relaxation to the ground state [236]. Hence, probes with shorter fluorescence lifetimes (~ 0.5 -5 ns) are preferred. This condition has an exception in lifetime-based experiments, which by contrast, demand long fluorescence lifetime to increase the observation time window.

Brightness: Brightness is an important parameter describing the rate of photons emitted by the fluorescent probe. The parameter is defined as the product of the molar absorptivity (ε_{exc}) and the fluorescence quantum yield of the probe (Q_Y). Limiting parameters for a SMFD probe are: ε_{exc}>20 000 cm^{-1} M^{-1} and Qy>0.1 [148]. Brightness is affected by the local environment, for example proximity of quenching agents and buffer pH, viscosity and solvent polarity.

Triplet-state saturation: Singlet-triplet transition (intersystem crossing) usually leads to a spin forbidden return to the ground state. This transition results in micro- to millisecond dark states. For some probes, effective triplet-state saturation leads to a significant decrease in the emission intensity. Such probes then are not suitable for SMFD without proper sample treatment. The lifetime of the triplet state can be significantly decreased by the ground triplet state of molecular oxygen ($^3O_2$) [149]. However oxygen often causes photodestruction of the probe, thus samples are usually deoxygenated. The dilemma can be solved by a combination of oxygen scavengers and triplet-state quenchers, specific for a particular fluorescent probe.

Signal stability and fluorescence intermittency: Fluorescent probes transiently interact with their local environment resulting in spectral jumps and/or jumps in quantum yield (blinking). Quantum yield can quickly jump to zero, creating several dark states. These dark states can last a long time, in order of seconds, resulting in fluorescence intermittency that reduces the number of emitted photons and alters time trajectories of fluorescence signal.

Resistance to photodestruction: Organic probes incline to irreversible photodestruction, mainly due to photo-oxidation or photo-reduction [150]. The photosurvival time determines the total number of fluorescence photons and limit the time of observation. Common SMFD probes emit 10^5–10^6 photons before their photodestruction [147]. The quantum yields of photodestruction and photosurvival times significantly vary amongst fluorescent probes [151]. Photodestruction is a key issue in experiments on immobilized biomolecules that report on long time trajectories

of complex reactions. The ideal photosurvival time should allow probing of several time scales of reaction completion. The photosurvival time might be extended with various anti-photobleaching and triplet quenching cocktails (see section 1.7.2).

Probe hydrophilicity and linker size: Probe attachment via short rigid linkers is preferred as it causes smaller perturbations to the structure of probed biomolecules. Moreover short rigid linkers produce less fluorophore noise. This noise is generated by alternations in the spectral properties of probe caused by random changes in its local environment. Shorter linkers provide smaller accessible volume and therefore decreases the probability of probe quenching [152]. Water solubility of fluorescent probes is also important, hydrophilic probes reduce hydrophobic interactions with biomolecules and surfaces, thus reducing quenching and uncertainty in the location of the probe. We studied this phenomenon in a comparative study of noncovalent binding interactions between a short B-DNA oligonucleotide and terminally attached nonfluorescent quencher QSY 21 and Rhodamine 6G, for details see section 2.2. Both analyzed dyes show substantial binding with the terminal base pair including several binding motifs.

Site-specific labeling

Site-specific attachment of fluorescent probes to proteins requires precise knowledge of the protein structure, strategic choice of conjugation chemistry, careful optimization of the labeling reaction and rigorous verification of the efficiency and function of the labeled proteins. Labeling of DNA and RNA is usually much easier, and several fluorescent probes or reactive groups might be implemented using automated solid-phase synthesis. Moreover labeled nucleic acids are commonly provided by commercial manufacturers.

In vitro **labeling of proteins:** The high site-specificity of protein labeling is not usually required for imaging applications, as far as the probe does not affect the function of the biomolecule. Nevertheless, when researchers seek precise orientation or distance information, site-specificity becomes an important issue.

  The most common technique is the cysteine-specific reaction with thiol-reactive reagents. Proteins with exposed cysteine residues are here covalently labeled by iodoacetamide or maleimide conjugates of fluorescent probes [153][154]. Cystein is the preferred labeling target, since it is rare, and can be easily substituted with another amino acid via site-directed mutagenesis [155]. If the protein of interest has no surface-exposed cystein residues, a carefully chosen amino acid at the chosen locus can be substituted by a cystein [155]. Vice versa, unfitting or redundant cystein can be replaced with serine, an amino acid with similar properties [246].

  To monitor biomolecular orientation using fluorescence polarization experiments, the orientation of a transition dipole moment of fluorescent probe can be fixed. Proper orientation of the transition dipole moment with respect to the structure of host biomolecule can be set by intramolecular cross-linking of two suitably spaced cystein residues with a bisfunctional cystein-reactive fluorescent probe [113, 247].

In vitro **labeling of nucleic acids:** Fluorescent labeling of DNA and RNA is usually achieved by enzymatic reactions. Organic fluorescent probes are covalently attached to

short primers or nucleoside triphosphates and then transferred using PCR amplification or using nucleic acid polymerases or terminal polynucleotide transferases [156][157][158]. The direct chemical incorporation of probes into DNA or RNA [159] is rare. Labeling of primers and/or nucleoside triphosphates is described elsewhere [160].

***In vivo* labeling of proteins:** There exist three main strategies of site-specific labeling of proteins in live cells which are suitable for the purposes of SMFD. The first and easiest strategy is labeling proteins *in vitro*, and subsequent incorporation into the live cell through a whole battery of incorporation techniques, such as chemo- and electro-permeabilization, endocytosis and microinjection. This strategy has allowed for example FRET monitoring of dimerization of the epidermal growth factor [161] and detection of individual β-galactosidase molecules inside the cell nucleus [162].

The second strategy involves a high-affinity protein–ligand interaction used for *in situ* attachment of fluorescent probes to specific amino acid sequences or protein structures inside live cells. This approach is more complicated than the previous one, since the site-specificity of the probe attachment chemistry must be highly selective to prevent nonspecific binding to different biomolecules. Moreover the probes with all the conjugation chemistry must be nontoxic and permeable through the cell membrane. For example this *in situ* labeling strategy was successfully used for monitoring of individual ion channels on the membrane of Jurkat cells [163]. Recently, there have been several amino acid motifs and labeling strategies established using an *in situ* approach. One variant of this strategy targets a four cystein α-helical motif incorporated to a protein [164]. Typical reagents representing this strategy are fluorescein arsenical helix binder (FLASH), and resorufine arsenical helix binder (REASH). FLASH and REASH based labeling can be combined with several genetically-encoded fluorescent proteins for FRET applications. Nevertheless, this strategy is not applicable in some cell lines due to known nonspecific interactions with proteins containing cystein rich segments [257]. A similar approach, called epitope tagging, utilizes site-specific single-chain antibodies. These antibodies are then co-expressed with investigated proteins in the cell [165]. The well known hexahistidine tag can also be used as a probe targeting sequence, utilizing the high affinity binding of this sequence to transition-metal complexes such as $(Ni^{2+}$-nitrilotriacetic acid)$_2$[166].

The third strategy employs genetically encoded fluorescent proteins, such as green fluorescent protein (GFP) [167]. Here, the DNA sequence coding for the fluorescent probe is inserted immediately alongside to the protein coding sequence resulting in expression of a hybrid GFP-fusion protein. Recently, mainly thanks to the effort of the Tsien group, there are variants of GFP and other genetically encoded fluorescent proteins which cover the whole visible spectrum [168]. Howbeit, fluorescent proteins have a few limitations. Their incorporation is limited to the protein termini and can cause perturbations due to their large size [164]. Nonetheless, there is a plethora of successful observations of single GFP-fusion proteins in living cells [74].

1.7.2. Anti-photobleaching and anti-blinking cocktails

Monitoring the complex multistep behavior of biomolecules demands a prolonged and stable fluorescence signal, ideally on the minute timescale. Nevertheless, as discussed above, fluorescent probes show fast photobleaching and/or fluctuations in quantum yield. Both these photophysical phenomena obstruct the interpretation of fluorescence time traces (such as spFRET time traces) and determine the observation time [169][170]. Several photophysical and chemical events have been proved to evoke signal instability, including triplet states [171], radical ions [172][173] and conformational isomers [267].

Molecular oxygen (O_2) is a key player in fluorescent probe stability. The advantage of triplet oxygen [$^3O_2$] molecules in SMFD exists in quenching triplet states of the fluorescent probes, which are responsible for blinking. Nevertheless, this reaction leads to the formation of highly reactive species such as singlet oxygen, hydroxyl radicals and hydrogen peroxides[149][174][175]. These reactive molecules can then oxidize exposed chemical groups in fluorescent probes, the amino acids cysteine, histidine, tyrosine and tryptophan and also the DNA nulceobase guanosine [176][177]. Oxidative damage drops the probe's quantum yield to zero and impairs the structure and function of biomolecules. In short, aqueous buffer with saturated dissolved O_2 will result in minimal blinking of fluorescent probes but oxidative damage to biomolecules and probes is extensive.

Recent methods for fluorescence signal stabilization and prolongation typically involve deoxygenation, [149] using an enzymatic O_2-scavenging system and chemical anti-blinking additives. The prevalent system utilizes a combination of glucose oxidase and catalase (GODCAT) in a glucose buffer [272][273] as deoxygenation agents. The GODCAT system extends the photosurvival to the tens of second. Fluorescence blinking is often minimized by addition of a chemical reducing agent - millimolar β-mercaptoethanol [178][147]. However, there are still several flaws limiting the use of this system. Obviously, the enzymatic nature of GODCAT limits its application strictly to *in vitro* experiments. Besides, mismatched GODCAT enzymatic activity produces hydrogen peroxide, enhancing the oxidative damage of biomolecules and fluorescent probes [179]. The main flaws of β-mercaptoethanol implementation inhere in induced long dark states (in order of seconds), or slow blinking [276]. Rasnik *et al.* described fluorescence stabilizing effects of β-mercaptoethanol and characterized its alternative - Trolox, a water-soluble analog of vitamin E. Trolox favorably minimizes slow blinking and photobleaching, as compared to β-mercaptoethanol [180].

Several other chemical additives have been employed as molecular oxygen scavengers, such as *n*-propyl gallate [181], *p*-phenylenediamine [182], or protocatechuic acid / protocatechuate-3,4-dioxygenase [276]. Carotenoids have been used as an alternative to singlet-oxygen scavengers [86] and ascorbic acid as oxygen-metabolite scavenger [280].

1.7.3. Immobilization and surface passivation

Biomolecules are often immobilized to prolong observation times, which are than limited only by photosurvival of fluorescent probes [183]. Immobilization systems in SMFD experiments employ "brick lying" systems. The most common system uses biotinylated bovine serum albumin (BSA) protein bound nonspecifically to an aminosilan treated glass surface as the ground layer. The second layer consists of multivalent avidin proteins (such as neutravidin) which bind to the ground layer biotin. The last layer consists of the biomolecule of interest, modified with a biotin anchor, this biotin again binds to the multivalent avidin protein (Figure 9(a)).

Other brick lying system uses aminosilaned glass surface treated with a polyethylene glycol (PEG) coating, carrying Ni^{2+} or Cu^{2+} ions chelated to iminodiacetic acid groups. These groups efficiently bind to the hexahistidine tag (Figure 9(c)).

Recently, immobilization click chemistry has pervaded [6][183][184]. All brick lying systems occasionally perturb the structure and function due to possible direct structural interactions between the examined biomolecule and anchoring system.

Alternatively, biomolecules of interest might be entrapped inside lipid vesicles attached to a PEG-functionalized surface [185][284] (Figure 9(d)). This combination localizes biomolecules to a sufficiently small volume to avoid signal loss and does not require any complicated anchoring system, hence generally does not interfere with biomolecules.

A general issue of biomolecular immobilization is nonspecific interaction with the microscope slide, which has substantial negative charge at neutral pH. Situation is generally easier for nucleic acids as their negative charge prevents their "sticking" to the glass surface [186][286]. Complications come with the studies involving protein molecules. Their study is more challenging. Due to partial positive charges, proteins can non-specifically interact with the negatively charged glass surface [187]. A common solution of this issue, employs passivation of the glass surface by a thin neutral layer consisting of, for example, PEG (Figure 9(a) versus (b)) [6][188]. Two of the less common immobilization methods, the click chemistry and lipid vesicle encapsulation, do not suffer with nonspecific interactions [183][185], hence do not require any passivation.
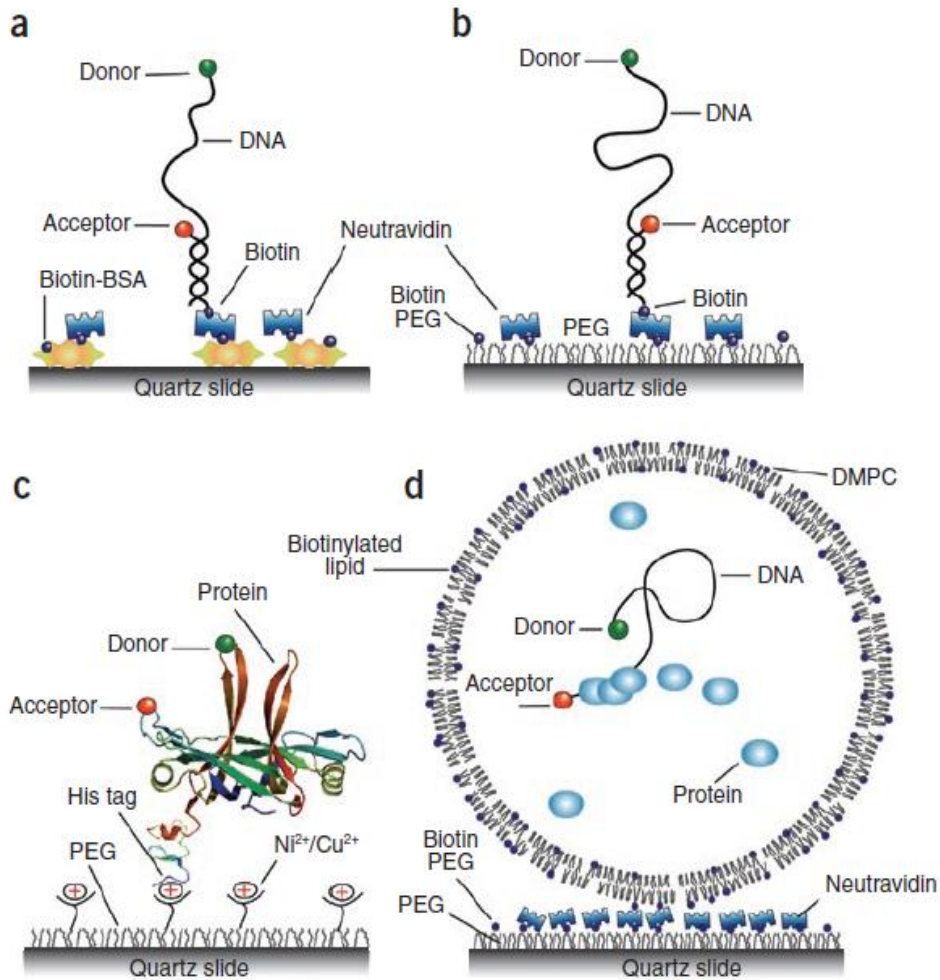
Figure 9: Immobilization strategies in SMFD experiments.

*(**a**) Biotinylated bovine serum albumin proteins bound nonspecifically to the aminosilanized glass surface anchor biotinylated double stranded nucleic acid with the help of neutravidin. (**b**) A mixture of biotin-PEG and PEG is covalently connected to aminosilanized glass surface. Biotins on the PEG can bind biomolecules with attached biotin tether with the aid of a sandwiched avidin protein. The PEG layer avoids nonspecific binding interactions. (**c**) PEG-coated surfaces can be modified to carry Ni^{2+} or Cu^{2+} chelated with iminodiacetic acid groups to bind 6 His-tagged biomolecules. (**d**) An example of immobilization using vesicle encapsulation using dimyristoyl phospatidylcholine (DMPC). At room temperature, DMPC vesicles are permeable to small molecules (as ATP) and can be used to entrap larger biomolecules. Biotinylated lipids allow specific binding of the vesicles to the biotin-avidin-PEG aminosilanized surface. Adopted from Roy* et al. *[6].*

1.7.4. Autofluorescence and diffraction issues

The reduction of noise and background is a key issue of all SMFD techniques [2]. The SNR and SBR assign the capability of the optical instrumentation to accurately examine individual molecules. High SBR and SNR allow reliable determination of fluorescence parameters. The main factors contributing to noise and background are fluorescent impurities in buffers, Raman scattering of water, Rayleigh scattering and autofluorescence of cellular components. Increases in instrumentation-related SNR and SBR can be achieved by proper choice of fluorophores, excitation source and optics. Rayleigh diffraction in a sample might be minimized by use of ultrapure solvents and exceptionally clean, highly polished quartz microscope slides. Ultrapure chemicals should be used to rigorously exclude undesirable fluorescent impurities. Favorable fluorophores are shifted to the red part of visible spectrum and with substantial Stokes shift, as they allow simple discrimination from background fluorescence (usually blue-shifted) and from Rayleigh scattering.

Autofluorescence: A variety of molecules in living cells and tissues have their own intrinsic fluorescence. Well documented examples of autofluorescence sources are nicotinamide adenine dinucleotide (NADH), flavoproteins [289] and elastin [290]. It is also induced by second harmonic generation in collagen structures [189]. Also amino acids with aromatic rings produce autofluorescence. The intensity of intrinsic fluorescence depends on the excitation wavelength used [190] and also on the localization inside living cells. In HeLa cells, the cytoplasm produces double the intensity of intrinsic fluorescence than the nucleus with the same excitation [191]. To decrease autofluorescence, it is advisable to use cell lines with minimal background and to eventually prephotobleach them. The red shifted excitation source and fluorophores might also decrease intrinsic fluorescence, as the majority of autofluorescent molecular species has absorption spectra in blue spectral region.

References

1. Selvin, P.R., ed. *Single-Molecule Techniques: A Laboratory Manual*. ed. T. Ha. 2008: Urbana-Champaign.
2. Weiss, S., *Fluorescence spectroscopy of single biomolecules.* Science, 1999. **283**(5408): p. 1676-1683.
3. Gordon, M. and P.R. Selvin, *Single-molecule high resolution imaging microscopy with photobleaching (SHRIMP): A new approach to high resolution fluoresence imaging.* Biophysical Journal, 2004. **86**(1): p. 180a-180a.
4. Gradinaru, C.C., et al., *Fluorescence anisotropy: from single molecules to live cells.* Analyst, 2010. **135**(3): p. 452-459.
5. Burghardt, T.P., *Single Molecule Fluorescence Image Patterns Linked to Dipole Orientation and Axial Position: Application to Myosin Cross-Bridges in Muscle Fibers.* Plos One, 2011. **6**(2).
6. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET.* Nature Methods, 2008. **5**(6): p. 507-516.
7. Molenkamp, L.W. and D.A. Wiersma, *Optical Dephasing in Organic Amorphous Systems - a Photon-Echo and Hole-Burning Study of Pentacene in Polymethylmethacrylate.* Journal of Chemical Physics, 1985. **83**(1): p. 1-9.
8. Kozankiewicz, B., J. Bernard, and M. Orrit, *Single-Molecule Lines and Spectral Hole-Burning of Terrylene in Different Matrices.* Journal of Chemical Physics, 1994. **101**(11): p. 9377-9383.
9. Moerner, W.E. and L. Kador, *Optical-Detection and Spectroscopy of Single Molecules in a Solid.* Physical Review Letters, 1989. **62**(21): p. 2535-2538.
10. Orrit, M. and J. Bernard, *Single Pentacene Molecules Detected by Fluorescence Excitation in a Para-Terphenyl Crystal.* Physical Review Letters, 1990. **65**(21): p. 2716-2719.
11. Betzig, E. and R.J. Chichester, *Single Molecules Observed by near-Field Scanning Optical Microscopy.* Science, 1993. **262**(5138): p. 1422-1425.
12. Ambrose, W.P., et al., *Alterations of Single-Molecule Fluorescence Lifetimes in near-Field Optical Microscopy.* Science, 1994. **265**(5170): p. 364-367.
13. Funatsu, T., et al., *Imaging of Single Fluorescent Molecules and Individual Atp Turnovers by Single Myosin Molecules in Aqueous-Solution.* Nature, 1995. **374**(6522): p. 555-559.
14. Selvin, P.R., et al., *Fluorescence resonance energy transfer between a single donor and a single acceptor molecule.* Biophysical Journal, 1996. **70**(2): p. Wp302-Wp302.
15. Yanagida, T., et al., *Single-molecule imaging of EGFR signalling on the surface of living cells.* Nature Cell Biology, 2000. **2**(3): p. 168-172.
16. Hirschfeld, T., *Optical Microscopic Observation of Single Small Molecules.* Applied Optics, 1976. **15**(12): p. 2965-2966.

17. Hirschfeld, T., *Quantum Efficiency Independence of Time Integrated Emission from a Fluorescent Molecule.* Applied Optics, 1976. **15**(12): p. 3135-3139.

18. Milles, S. and E.A. Lemke, *Single Molecule Study of the Intrinsically Disordered FG-Repeat Nucleoporin 153.* Biophysical Journal, 2011. **101**(7): p. 1710-1719.

19. Wang, D. and E. Geva, *Protein structure and dynamics from single-molecule fluorescence resonance energy transfer.* Journal of Physical Chemistry B, 2005. **109**(4): p. 1626-1634.

20. Narainui, H., H. Akutsu, and Y. Kyogoku, *Alcohol Induced B-a Transition of Dnas with Different Base Compositions Studied by Circular-Dichroism.* Journal of Biochemistry, 1985. **98**(3): p. 629-636.

21. Li, G.W., O.G. Berg, and J. Elf, *Effects of macromolecular crowding and DNA looping on gene regulation kinetics.* Nature Physics, 2009. **5**(4): p. 294-297.

22. Preisler, R.S., et al., *The B-Form to Z-Form Transition of Poly(Dg-M(5)Dc) Is Sensitive to Neutral Solutes through an Osmotic-Stress.* Biochemistry, 1995. **34**(44): p. 14400-14407.

23. Iqbal, A., et al., *Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids.* Proc Natl Acad Sci U S A, 2008. **105**(32): p. 11176-81.

24. Hicks, S., J. Case, and A. Jofre, *Conformational Diversity of Short DNA Duplex.* Journal of Physical Chemistry B, 2010. **114**(46): p. 15134-15140.

25. Bonnet, G., O. Krichevsky, and A. Libchaber, *Kinetics of conformational fluctuations in DNA hairpin-loops.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(15): p. 8602-8606.

26. Wallace, M.I., et al., *Non-Arrhenius kinetics for the loop closure of a DNA hairpin.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(10): p. 5584-5589.

27. Jung, J.Y. and A. Van Orden, *A three-state mechanism for DNA hairpin folding characterized by multiparameter fluorescence fluctuation spectroscopy.* Journal of the American Chemical Society, 2006. **128**(4): p. 1240-1249.

28. Kim, J., et al., *The initial step of DNA hairpin folding: a kinetic analysis using fluorescence correlation spectroscopy.* Nucleic Acids Research, 2006. **34**(9): p. 2516-2527.

29. Santoso, Y. and A.N. Kapanidis, *Probing Biomolecular Structures and Dynamics of Single Molecules Using In-gel Alternating-Laser Excitation.* Analytical Chemistry, 2009. **81**(23): p. 9561-9570.

30. Heller, R.C. and K.J. Marians, *Replisome assembly and the direct restart of stalled replication forks.* Nature Reviews Molecular Cell Biology, 2006. **7**(12): p. 932-943.

31. Ortiz-Lombardia, M., et al., *Crystal structure of a DNA Holliday junction.* Nature Structural Biology, 1999. **6**(10): p. 913-917.

32. Lilley, D.M.J., *Structures of helical junctions in nucleic acids.* Quarterly Reviews of Biophysics, 2000. **33**(2): p. 109-159.

33. Palets, D., et al., *Effect of Single-Strand Break on Branch Migration and Folding Dynamics of Holliday Junctions.* Biophysical Journal, 2010. **99**(6): p. 1916-1924.

34. McKinney, S.A., et al., *Structural dynamics of individual Holliday junctions.* Nature Structural Biology, 2003. **10**(2): p. 93-97.

35. Shirude, P.S. and S. Balasubramanian, *Single molecule conformational analysis of DNA G-quadruplexes.* Biochimie, 2008. **90**(8): p. 1197-1206.

36. Lee, J.Y. and D.S. Kim, *Dramatic effect of single-base mutation on the conformational dynamics of human telomeric G-quadruplex.* Nucleic Acids Research, 2009. **37**(11): p. 3625-3634.

37. Krasteva, V., *RNA structures.* Nature Methods, 2012. **9**(38).

38. Mishra, P.P., et al., *Codon-dependent tRNA fluctuations monitored with fluorescence polarization.* Biophysical Journal, 2010. **99**(11): p. 3849-58.

39. Dammertz, K., et al., *Single-molecule FRET studies of counterion effects on the free energy landscape of human mitochondrial lysine tRNA.* Biochemistry, 2011. **50**(15): p. 3107-15.

40. Fang, X.W., T. Pan, and T.R. Sosnick, *Mg2+-dependent folding of a large ribozyme without kinetic traps.* Nature Structural Biology, 1999. **6**(12): p. 1091-1095.

41. Qu, X., et al., *Single-molecule nonequilibrium periodic Mg2+-concentration jump experiments reveal details of the early folding pathways of a large RNA.* Proc Natl Acad Sci U S A, 2008. **105**(18): p. 6602-7.

42. Ha, T., *Single-molecule fluorescence resonance energy transfer.* Methods, 2001. **25**(1): p. 78-86.

43. Palo, K., et al., *Fluorescence intensity multiple distributions analysis: Concurrent determination of diffusion times and molecular brightness.* Biophysical Journal, 2000. **79**(6): p. 2858-2866.

44. Forkey, J.N., et al., *Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization.* Nature, 2003. **422**(6930): p. 399-404.

45. Trexler, A.J. and E. Rhoades, *Single molecule characterization of alpha-synuclein in aggregation-prone states.* Biophysical Journal, 2010. **99**(9): p. 3048-55.

46. Michalet, X., et al., *New photon-counting detectors for single-molecule fluorescence spectroscopy and imaging.* Advanced Photon Counting Techniques V, 2011. **8033**.

47. Buning, R. and J. van Noort, *Single-pair FRET experiments on nucleosome conformational dynamics.* Biochimie, 2010. **92**(12): p. 1729-1740.

48. Tims, H.S. and J. Widom, *Stopped-flow fluorescence resonance energy transfer for analysis of nucleosome dynamics.* Methods, 2007. **41**(3): p. 296-303.

49. Min, W., et al., *Fluctuating enzymes: Lessons from single-molecule studies.*

Accounts of Chemical Research, 2005. **38**(12): p. 923-931.

50. Smiley, R.D. and G.G. Hammes, *Single molecule studies of enzyme mechanisms.* Chemical Reviews, 2006. **106**(8): p. 3080-3094.

51. Lu, H.P., L.Y. Xun, and X.S. Xie, *Single-molecule enzymatic dynamics.* Science, 1998. **282**(5395): p. 1877-1882.

52. Edman, L. and R. Rigler, *Memory landscapes of single-enzyme molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(15): p. 8266-8271.

53. Polakowski, R., et al., *Single molecules of highly purified bacterial alkaline phosphatase have identical activity.* Journal of the American Chemical Society, 2000. **122**(20): p. 4853-4855.

54. Wang, Y. and H.P. Lu, *Bunching effect in single-molecule T4 lysozyme nonequilibrium conformational dynamics under enzymatic reactions.* Journal of Physical Chemistry B, 2010. **114**(19): p. 6669-74.

55. Lu, H.P., *Revealing time bunching effect in single-molecule enzyme conformational dynamics.* Physical Chemistry Chemical Physics, 2011. **13**(15): p. 6734-6749.

56. Nelson, P., M. Radosavljevic, and S. Bromberg, eds. *Biological physics.* 2004.

57. Pilizota, T., Y. Sowa, and R.M. Berry, *Single-Molecule Studies of Rotary Molecular Motors.* Handbook of Single-Molecule Biophysics, 2009: p. 183-216.

58. Noji, H., et al., *Direct observation of the rotation of F-1-ATPase.* Nature, 1997. **386**(6622): p. 299-302.

59. Duser, M.G., et al., *36 degrees step size of proton-driven c-ring rotation in FoF1-ATP synthase.* EMBO J, 2009. **28**(18): p. 2689-96.

60. Myong, S. and T. Ha, *Stepwise translocation of nucleic acid motors.* Current Opinion in Structural Biology, 2010. **20**(1): p. 121-127.

61. Sabanayagam, C.R., et al., *Viral DNA packaging studied by fluorescence correlation spectroscopy.* Biophysical Journal, 2007. **93**(4): p. L17-9.

62. Joo, C., et al., *Real-time observation of RecA filament dynamics with single monomer resolution.* Cell, 2006. **126**(3): p. 515-527.

63. Fairman-Williams, M.E. and E. Jankowsky, *Unwinding Initiation by the Viral RNA Helicase NPH-II.* Journal of Molecular Biology, 2011.

64. Vale, R.D., *The molecular motor toolbox for intracellular transport.* Cell, 2003. **112**(4): p. 467-480.

65. Vale, R.D., et al., *Direct observation of single kinesin molecules moving along microtubules.* Nature, 1996. **380**(6573): p. 451-453.

66. Yildiz, A., et al., *Kinesin walks hand-over-hand.* Science, 2004. **303**(5658): p. 676-678.

67. Kural, C., et al., *Kinesin and dynein move a peroxisome* in vivo*: A tug-of-war or coordinated movement?* Science, 2005. **308**(5727): p. 1469-1472.

68. Nitzsche, B., F. Ruhnow, and S. Diez, *Quantum-dot-assisted characterization of microtubule rotations during cargo transport.* Nature

Nanotechnology, 2008. **3**(9): p. 552-556.

69. Yildiz, A., et al., *Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization.* Science, 2003. **300**(5628): p. 2061-2065.

70. Reck-Peterson, S.L., et al., *Single-molecule analysis of dynein processivity and stepping behavior.* Cell, 2006. **126**(2): p. 335-348.

71. DeWitt, M.A., et al., *Cytoplasmic dynein moves through uncoordinated stepping of the AAA+ ring domains.* Science, 2012. **335**(6065): p. 221-5.

72. Hohlbein, J., et al., *Surfing on a new wave of single-molecule fluorescence methods.* Physical Biology, 2010. **7**(3).

73. Sako, Y. and T. Yanagida, *Single-molecule visualization in cell biology.* Nature Reviews Molecular Cell Biology, 2003: p. Ss1-Ss5.

74. Iino, R., I. Koyama, and A. Kusumi, *Single molecule imaging of green fluorescent proteins in living cells: E-cadherin forms oligomers on the free cell surface.* Biophysical Journal, 2001. **80**(6): p. 2667-2677.

75. Ulbrich, M.H. and E.Y. Isacoff, *Subunit counting in membrane-bound proteins.* Nature Methods, 2007. **4**(4): p. 319-321.

76. Wieser, S. and G.J. Schutz, *Tracking single molecules in the live cell plasma membrane-Do's and Don't's.* Methods, 2008. **46**(2): p. 131-140.

77. Sun, Y., et al., *Parallax: High Accuracy Three-Dimensional Single Molecule Tracking Using Split Images.* Nano Letters, 2009. **9**(7): p. 2676-2682.

78. Rout, M.P. and J.D. Aitchison, *The nuclear pore complex as a transport machine.* Journal of Biological Chemistry, 2001. **276**(20): p. 16593-16596.

79. Grunwald, D. and R.H. Singer, In vivo *imaging of labelled endogenous beta-actin mRNA during nucleocytoplasmic transport.* Nature, 2010. **467**(7315): p. 604-7.

80. Bertrand, E., et al., *Localization of ASH1 mRNA particles in living yeast.* Molecular Cell, 1998. **2**(4): p. 437-445.

81. Yunger, S., et al., *Single-allele analysis of transcription kinetics in living mammalian cells.* Nature Methods, 2010. **7**(8): p. 631-U41.

82. Nevo-Dinur, K., et al., *Translation-Independent Localization of mRNA in E. coli.* Science, 2011. **331**(6020): p. 1081-1084.

83. Treutlein, B. and J. Michaelis, *Direct Observation of Single RNA Polymerase Processing through a Single Endogenous Gene in a Living Yeast Cell.* Angew Chem Int Ed Engl, 2011.

84. Zhuang, X.W. and B. Brandenburg, *Virus trafficking - learning from single-virus tracking.* Nature Reviews Microbiology, 2007. **5**(3): p. 197-208.

85. Moerner, W.E. and D.P. Fromm, *Methods of single-molecule fluorescence spectroscopy and microscopy.* Review of Scientific Instruments, 2003. **74**(8): p. 3597-3619.

86. Pawley, J.B., ed. *Handbook of Biological Confocal Microscopy*. 1995, Plenum: New York.

87. Haran, G., *Single-molecule fluorescence spectroscopy of biomolecular folding.* Journal of Physics-Condensed Matter, 2003. **15**(32): p. R1291-

R1317.

88. Enderlein, J., et al., *The statistics of single molecule detection: an overview.* Bioimaging 1997. **5**: p. 88–98.

89. Samiee, K.T., et al., *lambda-repressor oligomerization kinetics at high concentrations using fluorescence correlation spectroscopy in zero-mode waveguides.* Biophysical Journal, 2005. **88**(3): p. 2145-2153.

90. Braslavsky, I., et al., *Sequence information can be obtained from single DNA molecules.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(7): p. 3960-3964.

91. Axelrod, D., *Total Internal-Reflection Fluorescence Microscopy.* Methods in Cell Biology, 1989. **30**: p. 245-270.

92. Conibear, P.B. and C.R. Bagshaw, *A comparison of optical geometries for combined flash photolysis and total internal reflection fluorescence microscopy.* Journal of Microscopy-Oxford, 2000. **200**: p. 218-229.

93. Tokunaga, M., et al., *Single molecule imaging of fluorophores and enzymatic reactions achieved by objective-type total internal reflection fluorescence microscopy.* Biochemical and Biophysical Research Communications, 1997. **235**(1): p. 47-53.

94. Axelrod, D., *Total internal reflection fluorescence microscopy in cell biology.* Traffic, 2001. **2**(11): p. 764-774.

95. Pelet, S., et al., *A fast global fitting algorithm for fluorescence lifetime imaging microscopy based on image segmentation.* Biophysical Journal, 2004. **87**(4): p. 2807-17.

96. Wallrabe, H. and A. Periasamy, *Imaging protein molecules using FRET and FLIM microscopy.* Current Opinion in Biotechnology, 2005. **16**(1): p. 19-27.

97. Clayton, A.H.A., et al., *Dynamic fluorescence anisotropy imaging microscopy in the frequency domain (rFLIM).* Biophysical Journal, 2002. **83**(3): p. 1631-1649.

98. Jia, Y.W., et al., *Nonexponential kinetics of a single tRNA(Phe) molecule under physiological conditions.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(15): p. 7932-7936.

99. Wennmalm, S., L. Edman, and R. Rigler, *Non-ergodic behaviour in conformational transitions of single DNA molecules.* Chemical Physics, 1999. **247**(1): p. 61-67.

100. Lu, H.P. and X.S. Xie, *Single-molecule kinetics of interfacial electron transfer.* Journal of Physical Chemistry B, 1997. **101**(15): p. 2753-2757.

101. Chen, H.M., E.R. Farkas, and W.W. Webb, In Vivo *Applications of Fluorescence Correlation Spectroscopy.* Biophysical Tools for Biologists, Vol 2: In Vivo Techniques, 2008. **89**: p. 3-+.

102. Schwille, P., F. Oehlenschlager, and N.G. Walter, *Quantitative hybridization kinetics of DNA probes to RNA in solution followed by diffusional fluorescence correlation analysis.* Biochemistry, 1996. **35**(31): p. 10182-10193.

103.   Weiss, M., et al., *Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells.* Biophysical Journal, 2004. **87**(5): p. 3518-3524.

104.   Sheetz, M.P., et al., *Nanometer-Level Analysis Demonstrates That Lipid Flow Does Not Drive Membrane Glycoprotein Movements.* Nature, 1989. **340**(6231): p. 284-288.

105.   Sabanayagam, C.R., et al., *Viral DNA packaging studied by fluorescence correlation spectroscopy.* Biophysical Journal, 2007. **93**(4): p. L17-L19.

106.   Lieto, A.M., R.C. Cush, and N.L. Thompson, *Ligand-receptor kinetics measured by total internal reflection with fluorescence correlation spectroscopy.* Biophysical Journal, 2003. **85**(5): p. 3294-3302.

107.   Edman, L., U. Mets, and R. Rigler, *Conformational transitions monitored for single molecules in solution.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(13): p. 6710-6715.

108.   Harvey, B.J., C. Perez, and M. Levitus, *DNA sequence-dependent enhancement of Cy3 fluorescence.* Photochemical & Photobiological Sciences, 2009. **8**(8): p. 1105-1110.

109.   Schwille, P., et al., *Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(1): p. 151-156.

110.   Haupts, U., et al., *Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(23): p. 13573-13578.

111.   Wong, F.H.C., et al., *A molecular thermometer based on fluorescent protein blinking.* Journal of the American Chemical Society, 2007. **129**(34): p. 10302-+.

112.   Kask, P., et al., *Fluorescence Correlation Spectroscopy in the Nanosecond Time Range - Rotational Diffusion of Bovine Carbonic Anhydrase-B.* European Biophysics Journal with Biophysics Letters, 1987. **14**(4): p. 257-261.

113.   Ehrenberg, M. and R. Rigler, *Fluorescence Correlation Spectroscopy Applied to Rotational Diffusion of Macromolecules.* Quarterly Reviews of Biophysics, 1976. **9**(1): p. 69-81.

114.   Clegg, R.M., *Fluorescence resonance energy transfer.* Curr Opin Biotechnol, 1995. **6**(1): p. 103-10.

115.   Cario, G. and J. Franck, *Concerning the break down of hydrogen molecules by excited quicksilver atoms.* Zeitschrift Fur Physik, 1922. **11**: p. 161-166.

116.   Forster, T., *Energiewanderung Und Fluoreszenz.* Naturwissenschaften, 1946. **33**(6): p. 166-175.

117.   Forster, T., *Zwischenmolekulare Energiewanderung Und Fluoreszenz.* Annalen Der Physik, 1948. **2**(1-2): p. 55-75.

118.   Wu, P. and L. Brand, *Resonance energy transfer: methods and applications.* Anal Biochem, 1994. **218**(1): p. 1-13.

119.	Wouters, F.S., P.J. Verveer, and P.I.H. Bastiaens, *Imaging biochemistry inside cells.* Trends in Cell Biology, 2001. **11**(5): p. 203-211.

120.	Jares-Erijman, E.A. and T.M. Jovin, *FRET imaging.* Nature Biotechnology, 2003. **21**(11): p. 1387-1395.

121.	Becker, W., et al., *FRET measurements by TCSPC laser scanning microscopy.* Photon Migration, Optical Coherence Tomography, and Microscopy, 2001. **2**(31): p. 94-98.

122.	Adams, S.R., et al., *Fluorescence Ratio Imaging of Cyclic-Amp in Single Cells.* Nature, 1991. **349**(6311): p. 694-697.

123.	Bastiaens, P.I.H., et al., *Imaging the intracellular trafficking and state of the AB(5) quaternary structure of cholera toxin.* Embo Journal, 1996. **15**(16): p. 4246-4253.

124.	Demarco, I.A., et al., *Monitoring dynamic protein interactions with photoquenching FRET.* Nature Methods, 2006. **3**(7): p. 519-524.

125.	Gadella, T.W.J., T.M. Jovin, and R.M. Clegg, *Fluorescence Lifetime Imaging Microscopy (Flim) - Spatial-Resolution of Microstructures on the Nanosecond Time-Scale.* Biophysical Chemistry, 1993. **48**(2): p. 221-239.

126.	Bastiaens, P.I.H. and A. Squire, *Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell.* Trends in Cell Biology, 1999. **9**(2): p. 48-52.

127.	Rizzo, M.A. and D.W. Piston, *High-contrast imaging of fluorescent protein FRET by fluorescence polarization microscopy.* Biophysical Journal, 2005. **88**(2): p. L14-L16.

128.	Tramier, M. and M. Coppey-Moisan, *Fluorescence anisotropy imaging microscopy for homo-FRET in living cells.* Fluorescent Proteins, Second Edition, 2008. **85**: p. 395-+.

129.	Torres, T. and M. Levitus, *Measuring conformational dynamics: A new FCS-FRET approach.* Journal of Physical Chemistry B, 2007. **111**(25): p. 7392-7400.

130.	Myong, S., et al., *Spring-loaded mechanism of DNA unwinding by hepatitis C virus NS3 helicase.* Science, 2007. **317**(5837): p. 513-516.

131.	McKinney, S.A., C. Joo, and T. Ha, *Analysis of single-molecule FRET trajectories using hidden Markov modeling.* Biophysical Journal, 2006. **91**(5): p. 1941-1951.

132.	Abbe, E., *Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrmehmung.* Archiv für Mikroskopische Anatomie 1873. **9**: p. 413–420.

133.	Hell, S.W. and J. Wichmann, *Breaking the Diffraction Resolution Limit by Stimulated-Emission - Stimulated-Emission-Depletion Fluorescence Microscopy.* Optics Letters, 1994. **19**(11): p. 780-782.

134.	Bretschneider, S., C. Eggeling, and S.W. Hell, *Breaking the diffraction barrier in fluorescence microscopy by optical shelving.* Physical Review Letters, 2007. **98**(21).

135.	Berglund, A.J., et al., *Fast, bias-free algorithm for tracking single particles with variable size and shape.* Optics Express, 2008. **16**(18): p. 14064-

14075.

136. Andersson, S.B., *Localization of a fluorescent source without numerical fitting.* Optics Express, 2008. **16**(23): p. 18714-18724.

137. Hess, S.T., T.P.K. Girirajan, and M.D. Mason, *Ultra-high resolution imaging by fluorescence photoactivation localization microscopy.* Biophysical Journal, 2006. **91**(11): p. 4258-4272.

138. Rust, M.J., M. Bates, and X.W. Zhuang, *Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM).* Nature Methods, 2006. **3**(10): p. 793-795.

139. Gordon, M.P., T. Ha, and P.R. Selvin, *Single-molecule high-resolution imaging with photobleaching.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(17): p. 6462-6465.

140. Hillesheim, L.N. and J.D. Muller, *The photon counting histogram in fluorescence fluctuation spectroscopy with non-ideal photodetectors.* Biophysical Journal, 2003. **85**(3): p. 1948-1958.

141. Chen, Y., et al., *The photon counting histogram in fluorescence fluctuation spectroscopy.* Biophysical Journal, 1999. **77**(1): p. 553-567.

142. Muller, J.D., Y. Chen, and E. Gratton, *Resolving heterogeneity on the single molecular level with the photon-counting histogram.* Biophysical Journal, 2000. **78**(1): p. 474-486.

143. Skakun, V.V., et al., *Global analysis of autocorrelation functions and photon counting distributions.* Front Biosci (Elite Ed), 2011. **3**: p. 489-505.

144. Sick, B., B. Hecht, and L. Novotny, *Orientational imaging of single molecules by annular illumination.* Physical Review Letters, 2000. **85**(21): p. 4482-4485.

145. Heinlein, T., et al., *Photoinduced electron transfer between fluorescent dyes and guanosine residues in DNA-hairpins.* Journal of Physical Chemistry B, 2003. **107**(31): p. 7957-7964.

146. Doose, S., H. Neuweiler, and M. Sauer, *A close look at fluorescence quenching of organic dyes by tryptophan.* Chemphyschem, 2005. **6**(11): p. 2277-2285.

147. Nie, S.M. and R.N. Zare, *Optical detection of single molecules.* Annual Review of Biophysics and Biomolecular Structure, 1997. **26**: p. 567-596.

148. Kapanidis, A.N. and S. Weiss, *Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules.* Journal of Chemical Physics, 2002. **117**(24): p. 10953-10964.

149. Hubner, C.G., et al., *Direct observation of the triplet lifetime quenching of single dye molecules by molecular oxygen.* Journal of Chemical Physics, 2001. **115**(21): p. 9619-9622.

150. Wilkinson, F., D.J. Mcgarvey, and A.F. Olea, *Excited Triplet-State Interactions with Molecular-Oxygen - Influence of Charge-Transfer on the Bimolecular Quenching Rate Constants and the Yields of Singlet Oxygen (O-2(Asterisk),(1)Delta(G)) for Substituted Naphthalenes in Various Solvents.* Journal of Physical Chemistry, 1994. **98**(14): p. 3762-3769.

151. Eggeling, C., et al., *Photobleaching of fluorescent dyes under conditions used for single-molecule detection: Evidence of two-step photolysis.* Analytical Chemistry, 1998. **70**(13): p. 2651-2659.

152. Sindbert, S., et al., *Accurate distance determination of nucleic acids via Forster resonance energy transfer: implications of dye linker length and rigidity.* Journal of the American Chemical Society, 2011. **133**(8): p. 2463-80.

153. Waggoner, A., *Covalent labeling of proteins and nucleic acids with fluorophores.* Methods Enzymol, 1995. **246**: p. 362-73.

154. Selvin, P.R., *Fluorescence resonance energy transfer.* Methods Enzymol, 1995. **246**: p. 300-34.

155. Kunkel, T.A., K. Bebenek, and J. McClary, *Efficient site-directed mutagenesis using uracil-containing DNA.* Methods Enzymol, 1991. **204**: p. 125-39.

156. Ansorge, W., et al., *Non-Radioactive Automated Sequencing of Oligonucleotides by Chemical Degradation.* Nucleic Acids Research, 1988. **16**(5): p. 2203-2206.

157. Prober, J.M., et al., *A System for Rapid DNA Sequencing with Fluorescent Chain-Terminating Dideoxynucleotides.* Science, 1987. **238**(4825): p. 336-341.

158. Smith, L.M., et al., *The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis.* Nucleic Acids Research, 1985. **13**(7): p. 2399-412.

159. Roget, A., H. Bazin, and R. Teoule, *Synthesis and Use of Labeled Nucleoside Phosphoramidite Building-Blocks Bearing a Reporter Group - Biotinyl, Dinitrophenyl, Pyrenyl and Dansyl.* Nucleic Acids Research, 1989. **17**(19): p. 7643-7651.

160. Proudnikov, D. and A. Mirzabekov, *Chemical methods of DNA and RNA fluorescent labeling.* Nucleic Acids Research, 1996. **24**(22): p. 4535-4542.

161. Sako, Y., S. Minoghchi, and T. Yanagida, *Single molecule imaging of signal-transduction of epidermal growth factor receptor (EGFR) stimulated by EGF on the surface of A431 cells.* Biophysical Journal, 2000. **78**(1): p. 68a-68a.

162. Kues, T., R. Peters, and U. Kubitscheck, *Visualization and tracking of single protein molecules in the cell nucleus.* Biophysical Journal, 2001. **80**(6): p. 2954-2967.

163. Schütz, G.J., Pastushenko, V. P., Gruber, H. J., Knaus, H.-G., Pragl, B., Schindler, H., *3D Imaging of Individual Ion Channels in Live Cells at 40nm Resolution.* Single Molecules, 2000. **1**(1): p. 25-31.

164. Griffin, B.A., S.R. Adams, and R.Y. Tsien, *Specific covalent labeling of recombinant protein molecules inside live cells.* Science, 1998. **281**(5374): p. 269-272.

165. Farinas, J. and A.S. Verkman, *Receptor-mediated targeting of fluorescent*

probes in living cells. Journal of Biological Chemistry, 1999. **274**(12): p. 7603-7606.

166. Kapanidis, A.N., Y.W. Ebright, and R.H. Ebright, *Site-specific incorporation of fluorescent probes into protein: Hexahistidine-tag-mediated fluorescent labeling with (Ni2+: Nitrilotriacetic acid)(n)-fluorochrome conjugates.* Journal of the American Chemical Society, 2001. **123**(48): p. 12123-12125.

167. Tsien, R.Y., *The green fluorescent protein.* Annual Review of Biochemistry, 1998. **67**: p. 509-544.

168. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins.* Nature Methods, 2005. **2**(12): p. 905-909.

169. Bagshaw, C.R. and D. Cherny, *Blinking fluorophores: what do they tell us about protein dynamics?* Biochemical Society Transactions, 2006. **34**: p. 979-982.

170. Ha, T.J., et al., *Temporal fluctuations of fluorescence resonance energy transfer between two dyes conjugated to a single protein.* Chemical Physics, 1999. **247**(1): p. 107-118.

171. Ha, T., et al., *Quantum jumps of single molecules at room temperature.* Chemical Physics Letters, 1997. **271**(1-3): p. 1-5.

172. Zondervan, R., et al., *Photoblinking of rhodamine 6G in poly(vinyl alcohol): Radical dark state formed through the triplet.* Journal of Physical Chemistry A, 2003. **107**(35): p. 6770-6776.

173. Sauer, M., et al., *Dynamics of the electron transfer reaction between an oxazine dye and DNA oligonucleotides monitored on the single-molecule level.* Chemical Physics Letters, 1998. **284**(3-4): p. 153-163.

174. Renn, A., J. Seelig, and V. Sandoghdar, *Oxygen-dependent photochemistry of fluorescent dyes studied at the single molecule level.* Molecular Physics, 2006. **104**(3): p. 409-414.

175. Piwonski, H., et al., *Optimal oxygen concentration for the detection of single indocarbocyanine molecules in a polymeric matrix.* Chemical Physics Letters, 2005. **405**(4-6): p. 352-356.

176. Sies, H. and C.F.M. Menck, *Singlet Oxygen Induced DNA Damage.* Mutation Research, 1992. **275**(3-6): p. 367-375.

177. Davies, M.J., *Reactive species formed on proteins exposed to singlet oxygen.* Photochemical & Photobiological Sciences, 2004. **3**(1): p. 17-25.

178. Kishino, A. and T. Yanagida, *Force Measurements by Micromanipulation of a Single Actin Filament by Glass Needles.* Nature, 1988. **334**(6177): p. 74-76.

179. Englander, S.W., D.B. Calhoun, and J.J. Englander, *Biochemistry without Oxygen.* Analytical Biochemistry, 1987. **161**(2): p. 300-306.

180. Rasnik, I., S.A. McKinney, and T. Ha, *Nonblinking and longlasting single-molecule fluorescence imaging.* Nature Methods, 2006. **3**(11): p. 891-893.

181. Gaigalas, A.K., et al., *Photodegradation of fluorescein in solutions containing n-propyl gallate.* Journal of Physical Chemistry A, 2004.

108(20): p. 4378-4384.

182. Florijn, R.J., et al., *Analysis of Antifading Reagents for Fluorescence Microscopy.* Cytometry, 1995. **19**(2): p. 177-182.

183. Aleman, E.A., H.S. Pedini, and D. Rueda, *Covalent-Bond-Based Immobilization Approaches for Single-Molecule Fluorescence.* Chembiochem, 2009. **10**(18): p. 2862-2866.

184. del Rio, A., et al., *Stretching Single Talin Rod Molecules Activates Vinculin Binding.* Science, 2009. **323**(5914): p. 638-641.

185. Okumus, B., et al., *Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic.* Biophysical Journal, 2004. **87**(4): p. 2798-2806.

186. Zhuang, X., *Single-molecule RNA science.* Annu Rev Biophys Biomol Struct, 2005. **34**: p. 399-414.

187. Rasnik, I., S.A. Mckinney, and T. Ha, *Surfaces and orientations: Much to FRET about?* Accounts of Chemical Research, 2005. **38**(7): p. 542-548.

188. Ha, T., et al., *Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase.* Nature, 2002. **419**(6907): p. 638-641.

189. Williams, R.M., W.R. Zipfel, and W.W. Webb, *Interpreting second-harmonic generation images of collagen I fibrils.* Biophysical Journal, 2005. **88**(2): p. 1377-1386.

190. Schwille, P., et al., *Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation.* Biophysical Journal, 1999. **77**(4): p. 2251-2265.

191. Chen, Y., et al., *Molecular brightness characterization of EGFP* in vivo *by fluorescence fluctuation spectroscopy.* Biophysical Journal, 2002. **82**(1): p. 133-144.

2. Results

2.1. Depth of focus extended microscope configuration

Fessl, T., Ben-Yaish, S., Vacha, F., Adamec, F., Zalevsky, Z., (2009). "Depth of focus extended microscope configuration for imaging of incorporated groups of molecules, DNA constructs and clusters inside bacterial cells." <u>Optics Communications</u> **282** (13): 2495-2501.

Depth of focus extended microscope configuration for imaging of incorporated groups of molecules, DNA constructs and clusters inside bacterial cells

Tomas Fessl[a,b], Shai Ben Yaish[c], Frantisek Vacha[a,b,d], Frantisek Adamec[b,d], Zeev Zalevsky[c]*

[a] Faculty of Science, University of South Bohemia, Branisovska 31, 37005 Ceske Budejovice, Czech Republic
[b] Biology Centre of Academy of Sciences, Branisovska 31, 37005 Ceske Budejovice, Czech Republic
[c] School of Engineering, Bar Ilan University, Bar Ilan, Ramat Gan 52900, Israel
[e] Institute of Physical Biology, University of South Bohemia, Zamek 136, 37333 Nove Hrady, Czech Republic

Abstract

Imaging of small objects such as single molecules, DNA clusters and single bacterial cells is problematic not only due to the lateral resolution that is obtainable in currently existing microscopy but also, and as much fundamentally limiting, due to the lack of sufficient axial depth of focus to have the full object focused simultaneously. Extension in depth of focus is helpful also for single molecule steady state FRET measurements. In this technique it is crucial to obtain data from many well focused molecules, which are often located in different axial depths.

In this paper we present the implementation of an all-optical and a real time technique of extension in the depth of focus that may be incorporated in any high NA microscope system and to be used for the above mentioned applications. We demonstrate experimentally how after the integration of special optical element in high NA 100× objective lens of a single molecule imaging microscope system, the depth of focus is significantly improved while maintaining the same lateral resolution in imaging applications of incorporated groups of molecules, DNA constructs and clusters inside bacterial cells.

Keywords

Fourier transform optics; Optical sensing

1. Introduction

In recent years the topic of lateral super resolution has become an important direction of research while large variety of approaches were developed and deployed in different imaging systems [1]. One of the aspects of improving the resolution capabilities of an imaging system is related to the axial rather than the lateral dimension. Having the object larger than the depth of focus extent provided by the imager, will generate lateral blurring and loss of spatial frequencies and features.

Different approaches were developed during the years to extend the depth of focus of imaging systems. Some are a combination of special optical element that

codes the aperture plane of the imaging lens and a digital post processing algorithm [2], [3], [4] and [5], some are related to aperture apodization by absorptive mask [6], [7], [8], [9] and [10] and others include the usage of a diffraction optical phase elements such as multi focal lenses or elements with spatially dense distributions [11], [12] and [13]. Other interesting approaches included tailoring the modulation transfer functions with high focal depth [14] and usage of logarithmic asphere lenses [15].

One interesting approach of extended depth of focus (EDOF) was presented in Refs. [16] and [17]. There an all-optical way for realization of the extension was demonstrated by attaching a phase-affecting element. The attached element is constructed from a binary phase pattern with spatially low frequency transitions that codes the entrance pupil of the lens. The presented approach had several important features: since this optical element contains low spatial frequencies, it is not sensitive to chromatic aberrations and dispersion (as other diffractive optical elements do) and it has high energetic efficiency not only in the element plane but rather in the object plane. In addition its fabrication is simple and cheap and thus its possible integration into an objective lens. The optical element has high energetic efficiency of close to 100% since it is a phase only element and thus it does not cause apodization by absorptive mask and in addition the phase element has no spatial high frequencies and thus there is no energy loss due to diffraction orders directing energy outside the region of interest.

In this EDOF technology the extension in the depth of focus was obtained by interference of the energy sent by the various parts of the lens aperture to the region of interest.

This all-optical technology was demonstrated also for ophthalmic applications [18] and proved to show improved performance in extending the depth of focus.

In this paper we design and realize the concept of Refs. [16], [17] for microscopy applications related to imaging of small structures such single molecules [19], DNA clusters and biological cells. In those applications the spatial features are very small and thus the natural depth of focus of the imaging objective is very short. This causes two undesired effects: several objects positioned one near the other cannot be focused simultaneously and single object also has blurred lateral regions because its 3D structure that sometimes extends beyond the depth of focus provided by the imaging system.

Since the proposed all-optical EDOF technology can allow real time extension of focus without energetic losses nor damage to the color fidelity we used the proposed concept for generating an optical binary phase only element that is added to the objective lens of special microscope configuration that is used for imaging applications of single molecules and small biological structures.

The novelty and the goals of this paper are as follows:

- To design and to construct an integrated 100× objective lens with high NA (of 0.95) that provides not only high transversal resolution but also increased depth of focus.
- The optical element that we aim to add to the lens should not increase its aberrations (such as chromatic aberrations) to avoid reduction in the transversal resolution.
- To incorporate this new lens into special microscope system allowing the imaging of group of single molecules, incorporated DNA constructs and clusters inside bacterial cells.
- Due to the application we are aiming for, the energetic efficiency is very important. Therefore, the extension in depth of focus and the improvement of the resolution should not reduce the efficiency of the energetic transmission of the constructed microscope system.
- To demonstrate experimentally that indeed the new lens design provides improved imaging performance for the above mentioned applications, in real time (e.g. in vivo) and in an all-optical way (without digital processing or without usage of a computer or a screen in the loop) such that human observer could see the outcome through the ocular lens of the microscope.

To the best of our knowledge the construction and the experimental demonstration of this special microscope configuration aiming to image very small objects such as group of single molecules, incorporated DNA constructs and clusters inside bacterial cells and having integrated high resolution objective lens with significantly increased depth of focus that improves the imaging performance for such small objects, has not been realized before.

Note that the new optical design of the objective lens used the basic concept of Ref. [17] as the starting working point but then an iterative simulated annealing optimization algorithm was applied in which the constrain of preserving all the spatial frequencies in the MTF plane obtained stronger emphasis, i.e. we took the working point coming from the analytical solution of Ref. [17] and applied simulated annealing optimization on it to optimize our trade-offs which are related to the spectral contrast of the various frequencies versus the focus extension. The parameters that were varied in the simulated annealing process were the diameter of the annular phase mask disc as well as the value of its phase.

In Section 2 we perform the numerical design of the element adapted to the parameters of our microscope configuration. In Section 3 we present the constructed experimental system and the obtained experimental results. The paper is concluded in Section 4.

2. Extension in depth of focus

The mathematical formulation used in Ref. [17] includes defining mathematically a binary phase only element positioned on the exit pupil of an imaging lens. In this case the Optical Transfer Function (OTF) which is the auto correlation of the Coherence Transfer Function (CTF) equals to:

$$H(\mu; Z_i) = \frac{\int_{-\infty}^{\infty} P\left(x + \frac{iZ_i\mu}{2}\right) \sum_{n=1}^{N} \exp\left(ia_n rect\left(\frac{x + \frac{iZ_i\mu}{2} - n\Delta x}{\Delta x}\right)\right) P^*\left(x - \frac{iZ_i\mu}{2}\right) \sum_{n=1}^{N} \exp\left(-ia_n rect\left(\frac{x - \frac{iZ_i\mu}{2} - n\Delta x}{\Delta x}\right)\right) dx}{\int_{-\infty}^{\infty} |P(x)|^2 dx}$$

(Eq.1)

where a_n are binary coefficients equal either to zero or to a certain phase modulation depth: $a_n = (0, \Delta_\phi)$ of the phase only element that we design. Δ_ϕ is the phase depth of modulation. Δx represents the spatial segments of the element. λ is the wavelength and μ is the coordinate of the OTF plane. P is the aperture of the lens having coordinates of x (the plane of the CTF) and Z_i is the distance between the imaging lens and the sensor. Since we do not want to create a diffractive optical element, i.e. spatially high frequencies (such that there will be no wavelength dependence and no chromatic aberrations) we force $\Delta x \gg \lambda$.

We impose mathematical constrain that the expression of Eq. (1) will have maximal value for the minimal contrasts defined by the OTF within a predetermined range of spatial frequencies. The result obtained in Ref. [17] was of an annular like disc structure with phase Δ_ϕ of close to $\pi/2$ (for the green wavelength of 532 nm). The annular like binary phase disc had external dimension of 4.1 mm.

The obtained result is understandable since the annular shape phase only element can cancel the sign inversions (by addition of proper phase to the spatial frequencies where the inversion was) of the quadratic phase generated while defocusing. Although in order to cancel the phase inversion one has to add a phase of π, the best solution in this case was the phase of $\pi/2$. The phase of $\pi/2$ was essential since due to the requirement for continues focused region, one had to cancel the inversions while the quadratic phase appeared but also when one was in focus and there was no quadratic phase. Thus $\pi/2$ is the phase that equally contributes when the object is defocused and when it is in focus.

Thus, the mathematical analytical solution after imposing the previously mentioned constrains over Eq. (1) yields an annular like phase shape element with binary phase of approximately $\pi/2$.

We applied this mathematical solution into Zemax software which is often used for lenses designs (we followed the concepts that were extensively described in Refs. [17] and [18]). In the software we simulated a full microscope system having 100× objective lens of Olympus with NA of 0.95. We assumed that in the microscope imaging system the distance between the lens and the image plane is Z_i = 250 mm and between the imaging lens and the object is Z_o = 2.5 mm. The focal length of the objective is F = 2.475247 mm and its aperture D = 4.70296 mm. Thus, the F number is $F_\# = F/D = 0.5263$.

In the simulations presented in Fig. 1 we compute the through focus Modulation Transfer Function (MTF) at the imaging plane (after magnification of 100×) for spatial frequency of 8 cycles per mm (i.e. in the object plane this spatial frequency is equivalent to frequency of 800 cycles per mm). From the obtained results

one may see improvement of at least a factor of 2 in the resulted depth of focus. In Fig. 1a and b we show the through focus MTF without the EDOF element at focus in Fig. 1a and after a shift of +1 μm from the focal position in Fig. 1b. One may see that shift of +1 μm destroyed completely the imaging contrast and resulted with a low pass for the spatial features. Therefore, the overall depth of focus of this objective without the usage of EDOF element is about 1.5 μm (±0.75 μm). When the EDOF element is added one may see that shift of +1 μm (Fig. 1c) or −1 μm (Fig. 1d), from the focal position, preserves the same contrast as obtained while in focus. Therefore, in this case the measurable overall depth of focus is approximately 3 μm (±1.5 μm).

The increase of the depth of focus from 1.5 μm into 3 μm in this high resolution microscope can significantly improve the imaging performance for small structures such as a group of single molecules, DNA constructs and clusters inside bacterial cells, since for all of those examples the axial dimension of those objects requires higher depth of focus in order to capture at once their entire structure or in order to see more such structures in-focus in the field of view of the microscope.
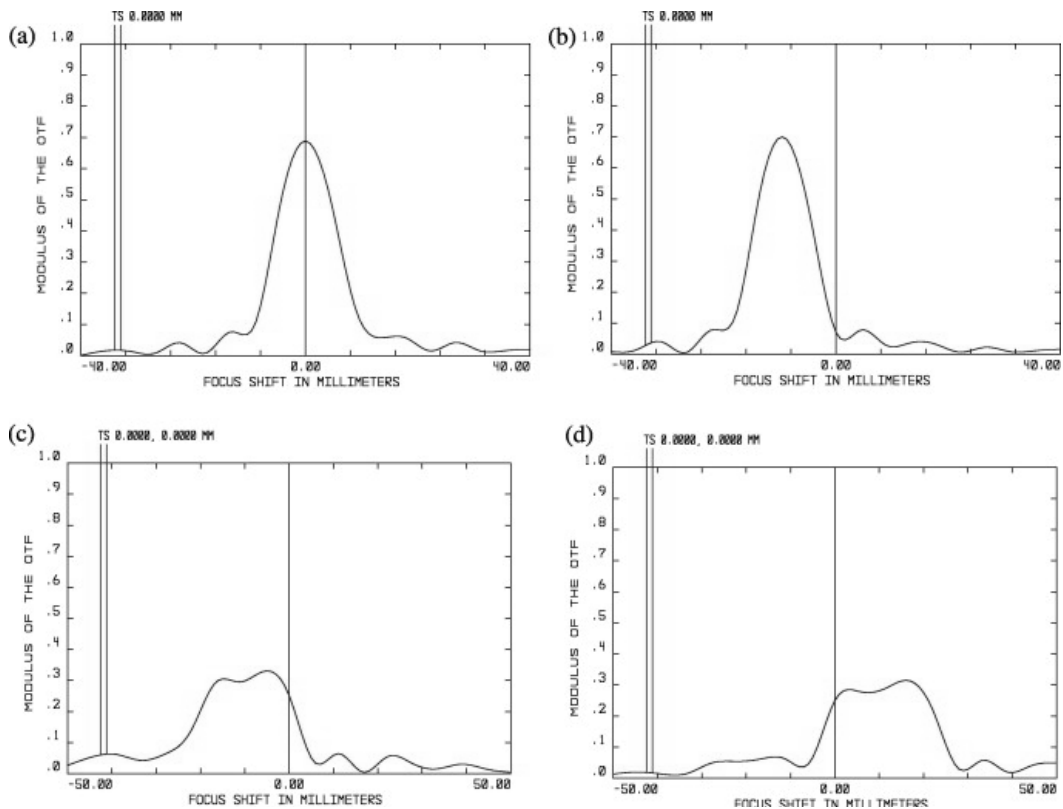
Fig. 1. Numerical simulations of through focus MTF for spatial frequency of 800 cycles/mm (in the object plane). (a) Without EDOF at focus. (b) Without EDOF with shift of +1 μm from focus. (c) With EDOF at position of +1 μm from focus. (d) With EDOF at position of −1 μm from focus.

From the results of the simulation designs of Fig. 1 one may see that a reduction of about 50% in contrast was obtained. However, please note that this was obtained for a very high spatial frequency of 800 cycles/mm. For a bit lower frequencies lower contrast reduction is obtained and the average reduction in contrast is not that significant (it is less than about 10%). Indeed reduction in signal to noise ratio (SNR) may be compensated by longer integration time and thus one may ask the question of maybe it is simpler just to leave the integration time as is and not to use the EDOF approach and instead increase the axial resolution by mechanical scanning. Regarding this point please note that the reduced SNR is not exactly equivalent to axial scanning. Specifically in the explored application, fast and not static objects are being imaged. Axial scanning will lead to capturing different molecules or other microscopic objects each time because they are constantly moving. There is an advantage of seeing all the axial information at one snap shot while the reduction in the contrast may be compensated by increasing the illumination power and by using a sensor with higher dynamic range or better SNR. This hardware that can compensate the reduction in contrast cannot compensate losing axial resolution.

We have fabricated the designed element using photolithography on SU-8 photo-resist on top of a plastic substrate and added the element, after proper optical alignment, into the experimental setup of the single molecule imaging microscope. Since the phase of this interference element is approximately $\pi/2$, it is very thin (less than 0.3 μm) and can easily be integrated/incorporated into the objective lens of the microscope system.

Please note that if extension in the depth of focus comes on extent of lateral resolution one may as well reduce the NA. However, the point of this paper is exactly to demonstrate the opposite. The proposed approach extends depth of focus WITHOUT loosing lateral resolution and this is why this new improved objective is so suitable for microscopy applications as those explored by this paper. The extension in the depth of focus comes on the extent of some reduction in contrast but not in lose of lateral resolution. This means that the Fourier transform of the point spread function contains all spatial frequencies while some have slightly reduced contrast.

3. Experimental results

The schematic sketch as well as an image (upper right corner) of the microscope system appears in Fig. 2.

The fabricated EDOF element that was designed for this system was positioned attached to the objective lens, aligned and tested with several biological samples. The microscope configuration consists of: inverted Olympus IX70 microscope, Triax 320 imaging spectrograph with back illuminated liquid nitrogen cooled CCD camera (Spectrum One, Jobin Yvon, 2048 × 512 pixels, pixel size 13.5 × 13.5 μm) and picosecond-pulse laser diode module (PicoQuant LDH-D-C-640, 641 nm, lin. polar) as an excitation source.
We enlarged the objects with objective Olympus 100×, NA of 0.95, UMPLANFL infinity/0, and excited samples with 641 nm by total internal reflection (TIRF).
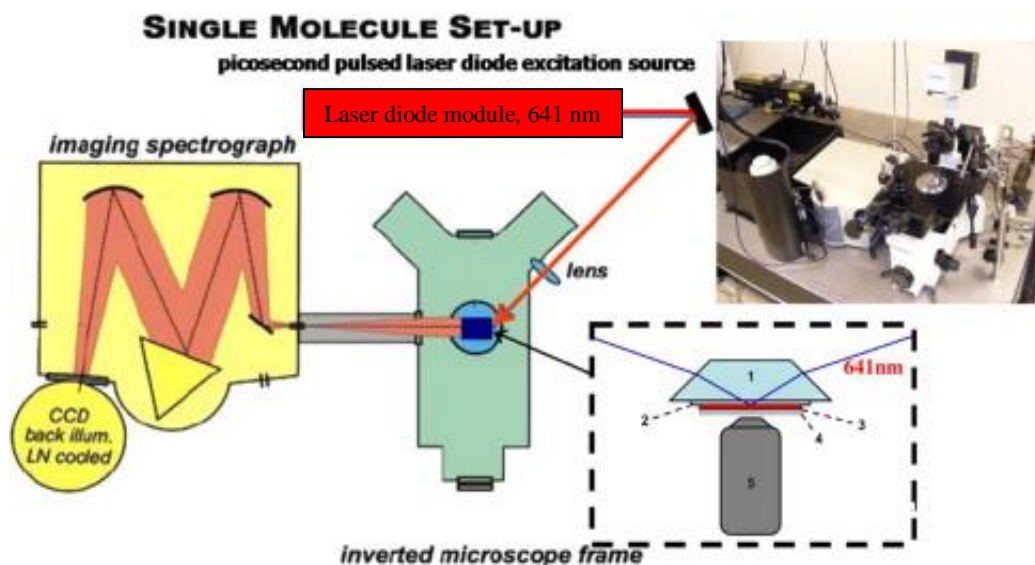
Fig. 2. The experimental setup for single molecule imaging. Item (1) is a quartz TIRF prism, item (2) is quartz cover slip with spin coated layer, item (3) is thin polymer layer with sample, item (4) is the EDOF mask and item (5) is the objective lens.

The microscope is equipped with an Olympus filter cube (Olympus, Japan) containing Raman emitter RS 664 LP (679.3–1497.7 nm), HC-Laser Clean-up MaxDiode 640/8 (Semrock, Germany).

In Fig. 3 we used samples that include a few *E. coli* cells. They are immobilized in a thick layer of polyvinyl alcohol (PVA) (MW 145 000, 98% hydrolyzed, Merck). This places them into different axial depths. There are two transmission images, one with (Fig. 3a) and one without (Fig. 3b) the EDOF element. One may clearly see that when the EDOF element was added more cells are seen in focus.

Inside the *E. coli* cells, there are incorporated double strand labeled oligonucleotides which have 22 base pairs, they are modified by fluorescent label Alexa Fluor 647. Those are incorporated DNA constructs. In Fig. 4 we imaged those samples. Here in contrast to the previous experiments we did not try to position several objects in different axial depths. Nevertheless, even when the objects are in the same plane due to their 3D structure, the usage of the EDOF element significantly improved the axial imaging resolution. The single almost round spots are single molecules. Other shapes on fluorescent images are clusters of constructs, which normally fill the whole bacteria (as could be seen when focusing through the cell). One may clearly see that in Fig. 4a where the EDOF element was used all the clusters are seen due to the extended depth of focus and thus we have an improved axial resolution. In Fig. 4b we did not use the EDOF element and thus the clusters are visible only in the focused part of the *E. coli* cells.
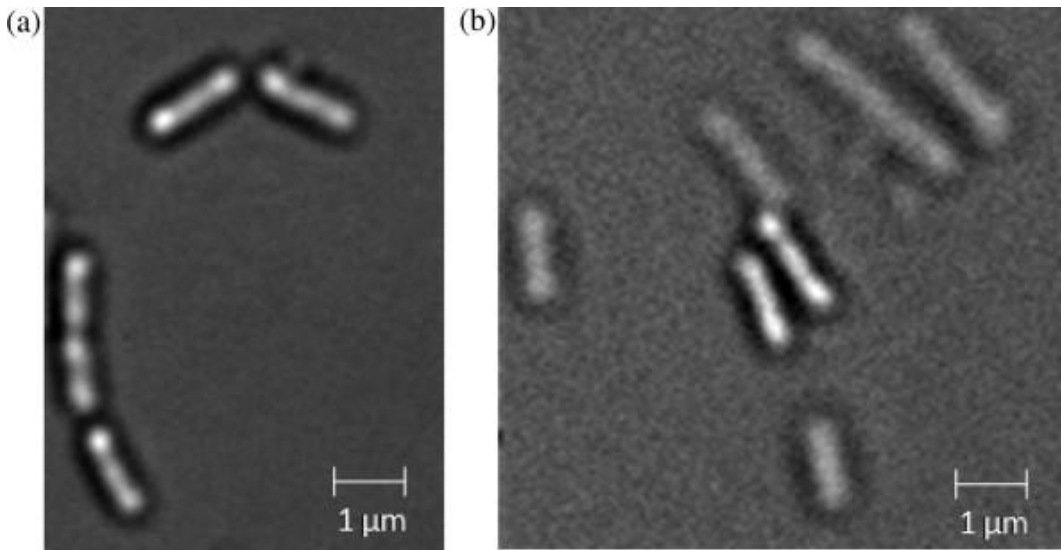
Fig. 3. Samples including a few E. coli *cells: (a) Transmission image with the EDOF element. (b) Transmission image without the EDOF element.*
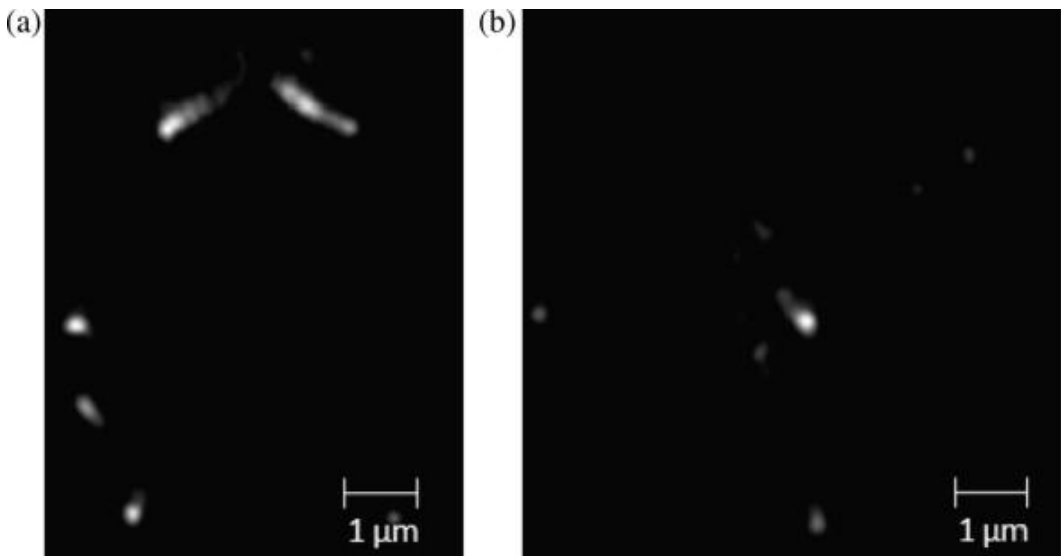

Fig. 4. Imaging of "in-cell" incorporated DNA clusters. (a) Fluorescence image of two clusters inside bacterial cell with the EDOF element. (b) Fluorescence image of cluster without the EDOF element.

Another example for imaging of small objects is seen in Fig. 5 where a few molecules are incorporated inside the bacterial cell and some others are outside the focal plane. In Fig. 5a where the EDOF element was used much more molecules are seen due to the extended depth of focus and thus improved axial resolution is obtained. In Fig. 5b we did not use the EDOF element and thus much less molecules are visible.
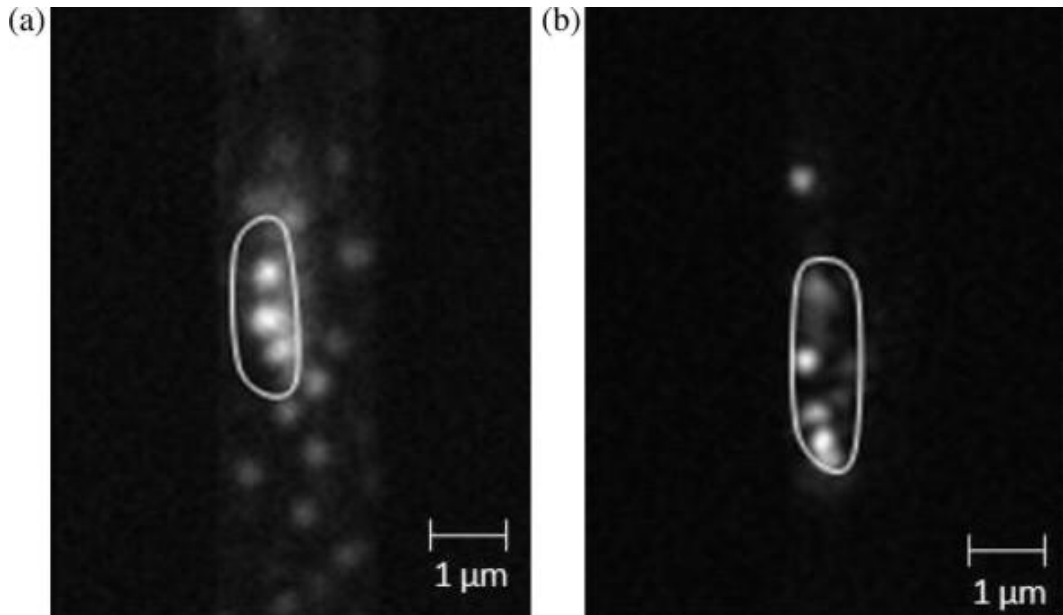
Fig. 5. "In-cell" imaging of few molecules that were incorporated inside the bacterial cell. (a) Fluorescence image with the EDOF element. (b) Fluorescence image without the EDOF element. White line borders the cell shape.

Note that the depth of focus enhancement factor in the experiments validated the numerical designs. The depth of focus extension factor that was obtained in the experiments was also close to two as obtained in the numerical simulations. Larger extension factor can allow many interesting biological studies such as observing dynamics throughout whole cell body. In our future work we intend to improve the optimization of the EDOF element such that larger extension factors are to be obtained while not as increasing the contrast related compromise.

In further investigation we modified a bit the experimental configuration as it appears in Fig. 6. The microscope configuration consists of: inverted Olympus IX70 microscope, Triax 320 imaging spectrograph with back illuminated liquid nitrogen cooled CCD camera (Spectrum One, Jobin Yvon, 2048×512 pixels, pixel size 13.5×13.5 μm) and HeCd laser (lin. polarized cw. laser at 442 nm – Kimmon, Japan) as an excitation source. We enlarged the objects with objective Olympus 100×, NA of 0.95, UMPLANFL infinity/0. The microscope is equipped with an Olympus filter cube (Olympus, Japan) containing Razor edge long pass filter 442 nm (Semrock, Germany).
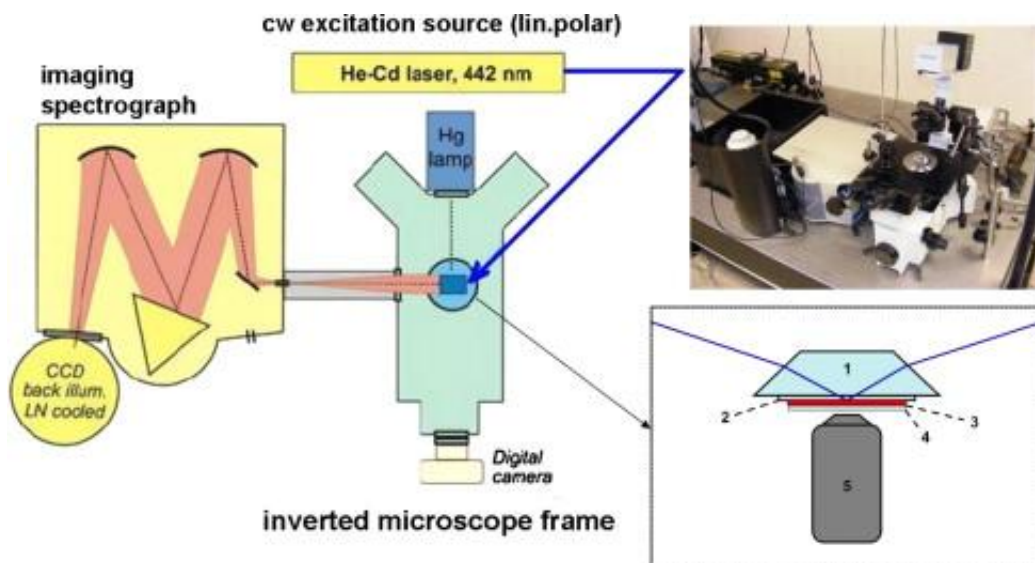
Fig. 6. The schematic sketch and image of the modified experimental setup. Item (1) is a quartz TIRF prism, item (2) is quartz cover slip with spin coated layer, item (3) is thin polymer layer with sample, item (4) is the EDOF mask and item (5) is the objective lens.

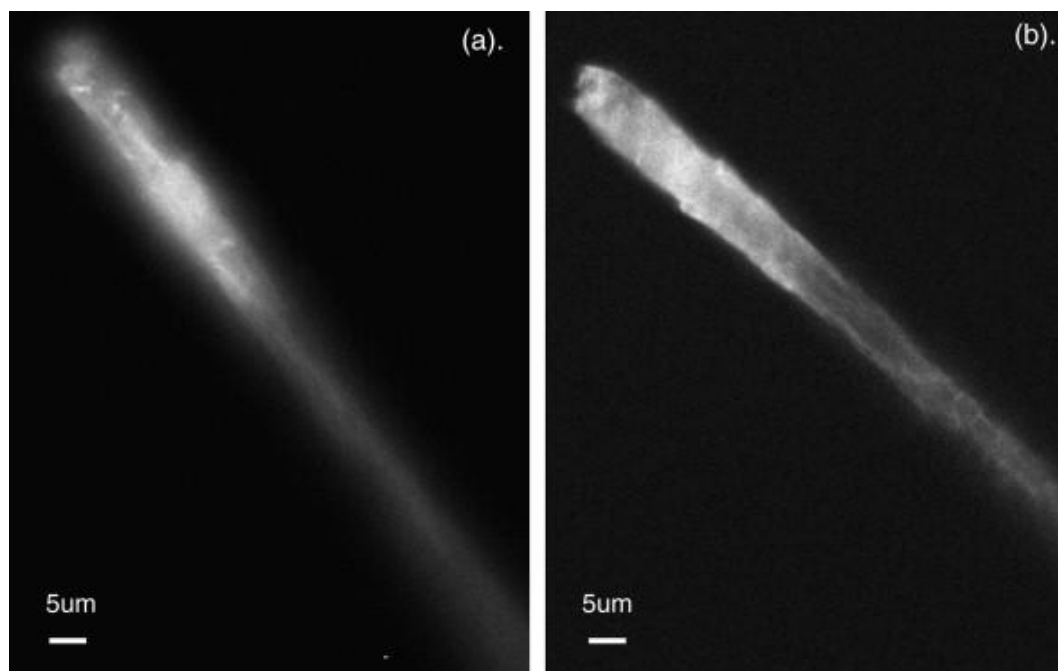The results from this configuration appear in Fig. 7 and Fig. 8.

Fig. 7. Chlorophyll fluorescence image of moss (Ceratodon purpureus) fragment. (a) Without EDOF mask. (b) With the EDOF mask.
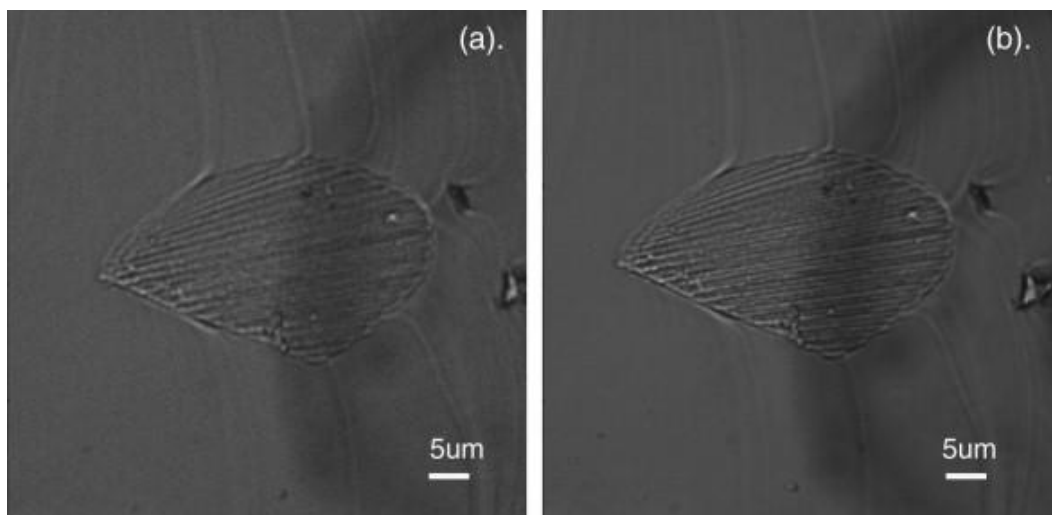
Fig. 8. Transmission image of flake-like crystal structure of PVA. (a) Without EDOF mask. (b) With the EDOF mask.

In Fig. 7 the object was a chlorophyll fluorescence image of moss (Ceratodon purpureus) fragment. Leaves of Ceratodon purpureus were cut into oligocellular linear fragments. Linear chlorophyll containing fragment was fixed in a firm position in the polyvinyl alcohol (PVA) matrix. Due to the tilted position of fragment in the PVA matrix the ends are in different axial depths (the upper-left part of fragment is in focal plane, the down-left part is the most distant visible part). Therefore, we can compare the depth of focus, using the EDOF mask. The object did not move. One may see in Fig. 7a the image obtained without the EDOF mask and in Fig. 7b with the mask. It is easy to see the significant improvement in the depth of focus.

In order to show that the extension in the depth of focus did not reduce the lateral resolution we have captured the images of Fig. 8. The images of Fig. 8 are the transmission image of flake-like crystal structure of PVA. The dark circle in the background is a disc shaped light-shield applied at the condenser side of the cover slide. Due to its fine structure this object can be used as a lateral resolution microscopic target. In Fig. 8a we present the captured image without the EDOF mask and in Fig. 8b it is with the mask. One may see that the lateral resolution is not reduced due to the addition of the EDOF mask.

4. Conclusions

In this paper we have presented the implementation of an all optical approach for extended depth of focus in single molecule imaging microscopy system.

The designed phase only optical element for the extension in depth of focus was fabricated and positioned attaching the objective lens of the microscope. Then, it was experimentally tested in imaging of sub-micron biological samples as incorporated molecules, DNA constructs and clusters inside bacterial cells.

Improved performance of imaging were experimentally demonstrated for plurality of objects positioned in different depths as well as for single small object

having 3D structure larger than the existing depth of focus range of the given imaging lens.

Acknowledgments

References

[1] Z. Zalevsky, D. Mendlovic, A. W. Lohmann, Progress in Optics, vol. XL, Chapter 4: ''Optical system with improved resolving power'', 1999.
[2] W. T. Cathy, E. R Dowski, Apparatus and method for extending depth of field in image projection system, US patent 6069738, May 2000.
[3] W. T. Cathy, Extended depth field optics for human vision, PCT Publication WO 03/052492, June2003.
[4] E. R. Dowski, W. T. Cathey, Appl. Opt. 34 (1995) 1859.
[5] J. van der Gracht, E. Dowski, M. Taylor, D. Deaver, Opt. Lett. 21 (1996) 919.
[6] C. M. Hammond, Apparatus and method for reducing imaging errors in imaging systems having an extended depth of field, US patent 6097856, August 2000.
[7] D. Miller, E.Blanko, System and method for increasing the depth of focus of the human eye, US patent 6554424, April 2003.
[8] N. Atebara, D. Miller, Masked intraocular lens and method for treating a patient with cataracts, US patent 4955904, September 1990.
[9] J. O. Castaneda, E. Tepichin, A. Diaz, Appl. Opt. 28 (1989) 2666.
[10] J. O. Castaneda, L. R. Berriel-Valdos, Appl. Opt. 29 (1990) 994.
[11] E. Ben-Eliezer, Z. Zalevsky, E. Marom, N. Konforti, D. Mendlovic, All optical extended depth of field imaging system, PCT Publication WO 03/076984, September 2003.
[12] E. Ben-Eliezer, Z. Zalevsky, E. Marom, N. Konforti, J. Opt. A: Pure Appl. Opt. 5 (2003) S164.
[13] E. Ben-Eliezer, E. Marom, N. Konforti, Z. Zalevsky, Appl. Opt. 44 (2005) 2792.
[15] W. Chi, N. George, Opt. Lett. 26 (2001) 875.
[16] Z. Zalevsky, Optical method and system for extended depth of focus, US patent No. 7061693 and US patent No. 7365917.
[17] Z. Zalevsky, A. Shemer, A. Zlotnik, E. Ben-Eliezer, E. Marom, Opt. Express 14 (2006) 2631.
[18] Z. Zalevsky, S. Ben Yaish, O. Yehezkel, M. Belkin, Opt. Express 15 (2007) 10790.
[19] G. Chirico, F. Cannone, S. Beretta, G. Baldini, A. Diaspro, Microsc. Res. Tech. 55 (5) (2001) 359.

2.2. Binding of QSY 21 and Rhodamine 6G probes to DNA

Kabelac, M., Zimandl, F., **Fessl, T.**, Chval, Z., Lankas, F., (2010). "A comparative study of the binding of QSY 21 and Rhodamine 6G fluorescence probes to DNA: structure and dynamics." <u>Physical Chemistry Chemical Physics</u> **12** (33): 9677-9684.

A comparative study of the binding of QSY 21 and Rhodamine 6G fluorescence probes to DNA: structure and dynamics†

Martin Kabeláč,*[ab] Filip Zimandl,[b] Tomáš Fessl,[c] Zdeněk Chval[d] and Filip Lankaš*[a]

[a] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic. E-mail: martin.kabelac@uochb.cas.cz, filip.lankas@uochb.cas.cz; Fax: +420 220 410 320; Tel: +420 220 410 319

[b] Faculty of Natural Sciences, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic

[c] Institute of Physical Biology, University of South Bohemia, Zámek 136, 373 33 Nové Hrady, Czech Republic

[d] Faculty of Health and Social Studies, University of South Bohemia, J. Boreckého 27, 37 011 České Budějovice, Czech Republic

† Electronic supplementary information (ESI) available: The videos and corresponding xyz trajectories of the MD simulations of DNA–QSY and DNA–RHO complexes. See DOI: 10.1039/c004020g

Abstract

Molecular dynamics (MD) simulations and *ab initio* quantum chemical calculations were employed to investigate the structure, dynamics and interactions of the QSY 21 nonfluorescent quencher and the fluorescence dye Rhodamine 6G bound to a B-DNA decamer. For QSY 21, two binding motifs were observed. In the first motif, the central xanthene ring is stacked on one base of the adjacent cytosine–guanine DNA base pair, whereas one of the 2,3-dihydro-1-indolyl aromatic side rings is stacked on the other base. In the second motif, the QSY 21 stacking interaction with the DNA base pair is mediated only by one of the side rings. Several transitions between the motifs are observed during a MD simulation. The *ab initio* calculations show that none of these motifs is energetically preferred. Two binding motifs were found also for Rhodamine 6G, with the xanthene ring stacked predominantly either on the cytosine or on the guanine. These results suggest that the side rings of QSY 21 play a crucial role in its stacking on the DNA and indicate novel binding mode absent in the case of Rhodamine 6G, which lacks aromatic side rings.

Introduction

Many processes in living organisms can be experimentally studied by fluorescence spectroscopy, most commonly in the far red and near-infrared (NIR) spectral region. There are several advantages to working in this region. Longer wave- length NIR fluorescence assays such as Rhodamine 6G, Cy3 and Cy5 can eliminate the tissue background signal because of the extremely low intrinsic fluorescence in the NIR as compared with traditional assays using shorter wavelength donor quencher pairs. NIR

assays also benefit from the enhanced tissue penetration of light near 650 to 900 nm, inexpensive laser diode excitation and reduced scattering. For these reasons, many new far-red and near infrared fluorescent probes have appeared on the market and with a lot of them, the structure, dynamics and interactions with labeled macromolecules like DNA or protein have been elucidated.[1]

One of the methods of fluorescence spectroscopy is Forster Resonance Energy Transfer (FRET). In FRET studies, a fluorescence or absorption characteristics are used to determine the rate of resonant energy transfer (k_{FRET}) from an energy donor (D) to an energy acceptor (A). D and A can be fluorophores of different origin. The FRET rate is often measured in time-resolved experiments that provide k_{FRET} more-or-less directly, or in steady-state experiments that yield k_{FRET} from the efficiency of D to A energy transfer relative to the fluorescence intensity of D. The rate of resonant energy transfer can be related to the distance between D and A, R_{DA}, using a theory developed by Forster in the 1940s. Recently, a number of groups have used molecular dynamics (MD) simulations to aid in the understanding of FRET experiments. Nonfluorescent quenching probes, also known as dark quenchers, are commonly used in FRET based assays including molecular beacons and nucleic acid hybridization real-time PCR, as well as in DNA charge transfer studies. A very promising tool seem to be FRET-based microarrays in which the immobilized strand is labeled with a donor fluorophore and the analyte target is labeled with a fluorescence quencher.[2] Unlike the traditional DNA microarrays this method allows visualization of both hybridized and unhybridized strands, eliminates artifacts[3] and enables real-time detection.[4]

The dark quencher QSY 21 is an efficient energy transfer acceptor of the far red and NIR fluorescent probes. It works in the wavelength range of 540–750 nm, frequently used in FRET applications. It has suitable spectral characteristics—a molar extinction coefficient of 90 000 $cm^{-1}M^{-1}$, and zero quantum yield (QY), meaning that in normal conditions it does not emit fluorescence. These properties make QSY 21 an excellent FRET acceptor. Its structure is similar to a popular laser probe and the fluorescent probe Rhodamine 6G.[5] The QSY 21 chromophore consists of two 2,3-dihydro-1-indolyl rings (denoted as DIHI 1 and DIHI 2 in this work) and one benzene ring attached to the central xanthene ring (see Fig. 1). QSY 21 is covalently bonded to the phosphate group of an adjacent DNA nucleotide via a phosphoramidite linker. An important factor affecting the spectral properties of FRET assays is the stacking interaction between the probe and the terminal DNA base pair. During the stacking interaction, a charge transfer between the base-pair and the probe takes place.[6] This leads to changes in the QY, molecular absorption cross-section and excited state lifetime. It is therefore important to resolve the probe position, dynamics, strength of binding and the interaction mechanism. To the best of our knowledge, unlike in the case of Rhodamine 6G, no X-ray or 2D-NMR study of QSY 21 bound to DNA has been published. Here we address the problem of the

structure, dynamics and interactions of QSY 21 bound to a B-DNA oligomer using theoretical methods, namely molecular dynamics (MD) and *ab initio* quantum chemical computations. We also report an analogous study involving Rhodamine 6G and compare the results for these two probes.
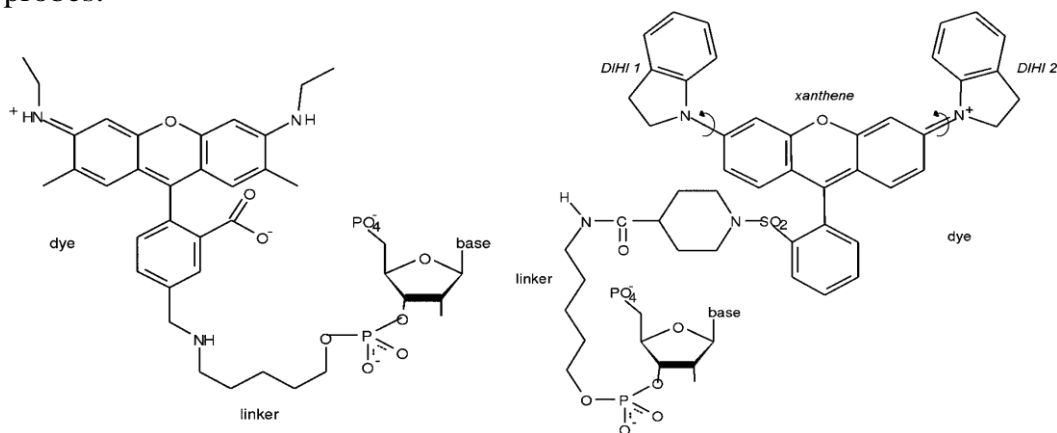

Fig. 1 *2-D representations of Rhodamine 6G (left) and QSY 21 (right) probes bonded to DNA via an aliphatic linker.*

Methods

MD simulations

The atomic resolution, explicit solvent molecular dynamics simulations of the 5'–CCACTGCAGG–3' B-DNA duplex decamer with covalently bonded QSY 21 probe (DNA–QSY), the decamer with a covalently bonded Rhodamine 6G (DNA–RHO), and the decamer alone (bare DNA) were performed using the AMBER 9 suite of programs. The parmbsc0 modification of the parm99 force field[7] for DNA was used, together with the TIP3P water model and Na^+ counterions. The standard parm99 parameterisation for the ions was also adopted.

The GAFF[8] force field parameters were employed for the Rhodamine 6G (RHO) and QSY 21 probes. The RESP charges for the probes were derived from the molecular electrostatic potential using the HF/6-31G* level of computation performed by the Gaussian 03 program,[9] in connection with the Antechamber module of AMBER. The starting structure of the DNA decamer in all of the simulations was a canonical B-DNA built using the program nucgen (part of Amber Tools). The probe was attached to the 5' end of the cytosine nucleotide. For the DNA–RHO complex, the dominant conformation found by NMR (PDB code 2V3L, model 1)[10] was adopted. The starting structure of the DNA–QSY complex was built manually in analogy to the structure of DNA–RHO.

An equilibration protocol consisting of a series of energy minimizations and short constrained MD runs was followed by a MD production using an NPT ensemble at 298 K and a pressure of 1 atm.

Periodic boundary conditions, the Particle Mesh Ewald method to treat long-range electrostatic interactions, and SHAKE on hydrogen atoms were used. Each trajectory was prolonged to 100 ns. The structures were saved every 1 ps. The conformation of the probes in terms of the DNA was monitored using two sets of internal coordinates. First, the central xanthene ring was mapped onto a virtual base pair of two cytosines, and its relative position and orientation with respect to the adjacent base pair of the DNA was described using the basepair step parameters commonly employed in DNA structural analysis, i.e. tilt, roll, twist, shift, slide and rise.[11] The 3DNA program[12] was used to compute the parameters. The most informative parameters turned out to be slide and twist for the DNA–QSY complex, and twist and shift for DNA–RHO. Besides that, the orientation of each of the 2,3-dihydro-1-indolyl side rings of QSY 21 (denoted as DIHI 1 and DIHI 2 in this work) in terms of the xanthene ring was monitored using a dihedral angle associated with the connecting bond between the xanthene ring and a side ring (see Fig. 1). The trajectories were also inspected in the VMD molecular graphics program. In order to investigate the influence of the probe on the structure of the DNA, we compared the average structural parameters of the simulated bare DNA with the DNA in the complexes. The 3DNA program[12] was used for the calculation. MD snapshots with at least one intra-basepair hydrogen bond broken or with at least one backbone torsion angle gamma in a non-canonical conformation were excluded from the analysis. Both phenomena represent rare events which can perturb the simulated DNA structure.[13,14]

Ab initio calculations of probe interactions with DNA

To estimate the strength of binding of the probes to DNA during MD simulation, two models were chosen. The small one includes the nearest pair of DNA bases (guanine and cytosine; the sugar–phosphate backbone was removed and replaced by hydrogen atoms), and the probe is mimicked just by the xanthene ring with the two side groups (i.e. the DIHI rings for QSY 21 and two aliphatic chains for RHO, see Fig. 2). The large model is composed from the nearest pair of DNA nucleosides (guanosine and cytidine) and a complete molecule of the probe with linker (the connecting phosphate group was removed and replaced by hydrogen atoms, see Fig. 2).

The computations were performed on every 100th frame saved from the MD simulations (1000 structures in total). To avoid problems with the SCF convergence in *ab initio* calculations, 100 steps of approximate self-consistent-charge, density functional tight-binding method (SCC-DF-TB-D)[15] minimization were applied. This optimisation does not lead to any significant change in geometry of the complexes. The single-point interaction energies of the complexes were calculated by the Resolution of Identity DFT method (the TPSS functional)[16] with the TZVP basis set using the Turbomole 6.0 program package.[17] The

empirical dispersion correction term[18] for the functional was added to describe the stacking interactions properly.
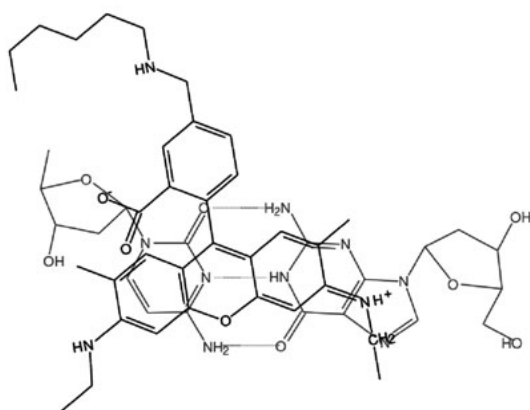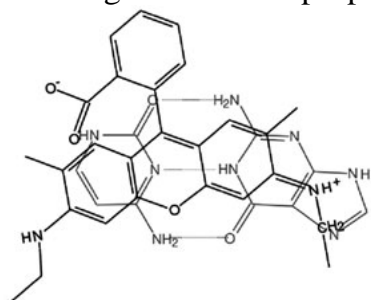

Fig. 2 The two models used for determination of the interaction energy between RHO (black) and the adjacent base pair (grey). The small model is shown above, the large one below.

Results and discussion

The MD simulations of DNA–QSY and DNA–RHO complexes

The initial structure of the DNA–QSY complex changes within ca. 6 ns of production MD into another conformation, which we call State Ia. In this conformation, the planes of the xanthene ring and of the adjacent base pair remain parallel, but the xanthene ring is stacked only on the cytosine. Its relative displacement with respect to the base pair is characterized by a slide of roughly − 2 Å and a twist of around 10° (see the upper part of Fig. 3 and 4). The phenyl ring remains on the minor groove side. One of the side rings (DIHI 1) is stacked on the guanine. The stacking is rather loose, as the side ring is able to switch its orientation by 180° (at ca. 10 ns). The other side ring (DIHI 2) is free to rotate and passes through multiple states between the dihedral angles of -40° and -160°.

At 27 ns, the system switches to a new conformational substate, State Ib. The xanthene ring slides along its long axis during the transition, without changing its orientation. This is reflected in the change of slide from - 2 Å to 3 Å, whereas the twist remains around - 10°. The change of the slide is clearly visible in Fig. 4. In the new state, the xanthene ring is stacked on the guanine whereas the DIHI 2 side ring is stacked on the cytosine, its dihedral angle remaining close to -20°. The DIHI 1 ring, in contrast, is now free to rotate and visits two well-defined states characterized by dihedral angle values of 40° and 140° respectively. In both State Ia and State Ib, the vertical separation between the xanthene ring and the base pair, as characterized by the value of rise, is roughly 3.5 Å and thus close to the vertical separation between base pairs in DNA.
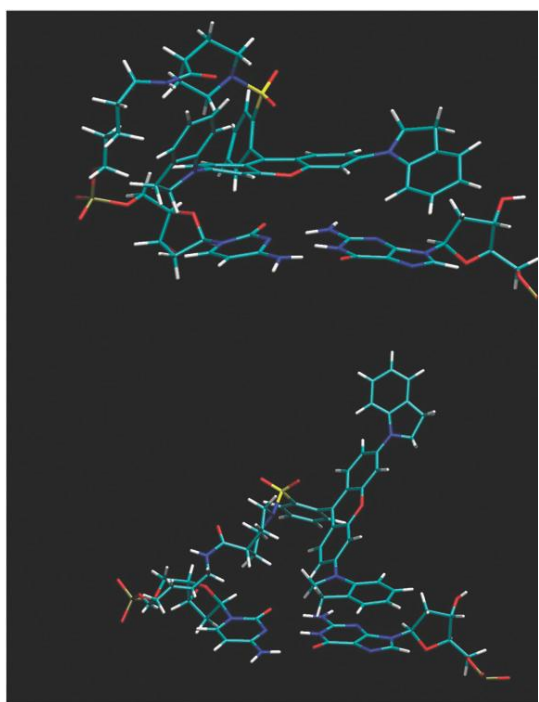
Fig. 3 The two main binding motifs obtained from the MD simulation of the DNA–QSY complex. A covalently bonded dye and the adjacent base pair are depicted. The structure above corresponds to the State Ia. The State II is represented by the structure below.
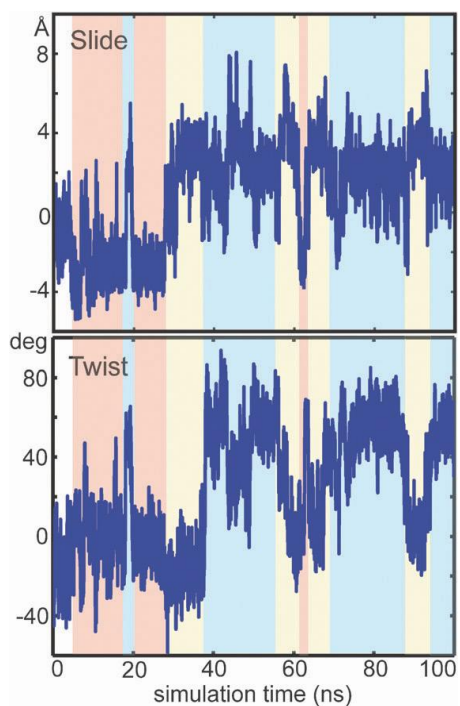
Fig. 4 The development of the slide and twist parameters characterising the orientation of the QSY 21 probe vs. the adjacent DNA base pair, during the 100 ns MD simulation. Different conformational substates (see text) are depicted by colour bands (State Ia in red, State Ib in yellow, State II in blue).

The second main conformational state (State II, see Fig. 3, the bottom part) is reached at 37 ns. This state fundamentally differs from the previous ones in that the xanthene ring is now completely unstacked and turned, pointing roughly towards the minor groove side of the DNA base pair. This transition is accompanied by a dramatic change in the twist (see bottom panel of Fig. 4) which now attains values around 60°, whereas the slide keeps oscillating around its State Ib value of 3 Å. The side ring DIHI 2 is stacked on the guanine, its dihedral angle staying close to 150°. The other side ring, DIHI 1, remains free to rotate and behaves as it did in State Ia.

Occasional interconversions among all of the states are observed. First, a short Ia to II switch is seen at roughly 20 ns. Around 60 ns, a change from II to Ib to Ia and back again takes place, and there is another change from II to Ib and back again around 90 ns. These transitions can be identified in Fig. 4. In summary, the MD simulation reveals two main conformational states, differing in the interaction pattern between the QSY 21 probe and the adjacent DNA base pair. One state (substates Ia and Ib) is characterised by a parallel arrangement of the xanthene ring and the base pair. The xanthene ring is stacked on one of the bases, whereas a side ring is stacked on the other one. State II, in contrast, includes the xanthene ring unstacked and turned away from the base pair, with the only stacking interaction being the one between a base and a side ring. All of the states seem to be easily accessible, because several transitions between them are observed.

Two main conformational states are observed also for the DNA−RHO complex (see Fig. 5). Its conformation shifts within a few ns of production MD from the initial structure into a similar state where the phenyl ring remains on the minor groove side but the xanthene ring is stacked predominantly on the cytosine (State A). The twist and shift fluctuate around 10° and 5 Å respectively. At 28 ns, this state irreversibly changes into a conformation in which the xanthene ring is loosely stacked on the guanine and the phenyl ring is located on the major groove side (State B). The transition is reflected by an abrupt change in the twist which now fluctuates around 100°, whereas the shift spans a range of several Å (see Fig. 6).
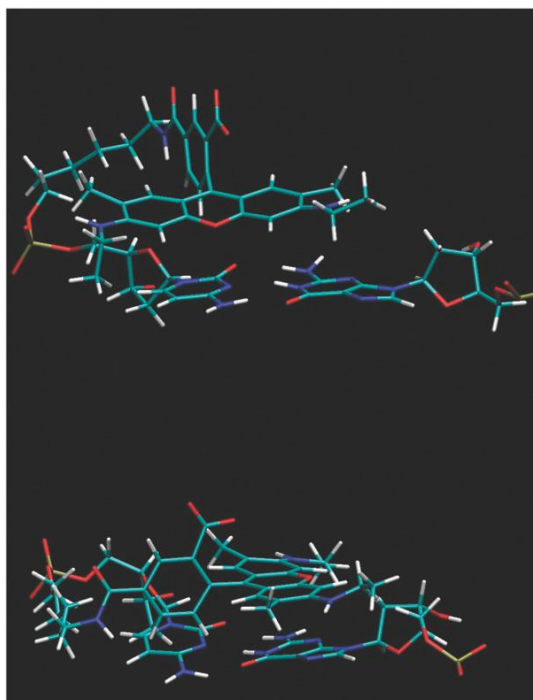
Fig. 5 *The two main binding motifs obtained from the MD simulation of the DNA–RHO complex. The covalently bonded dye and the adjacent base pair are depicted. The structure above corresponds to State A. State B is represented by the structure below.*
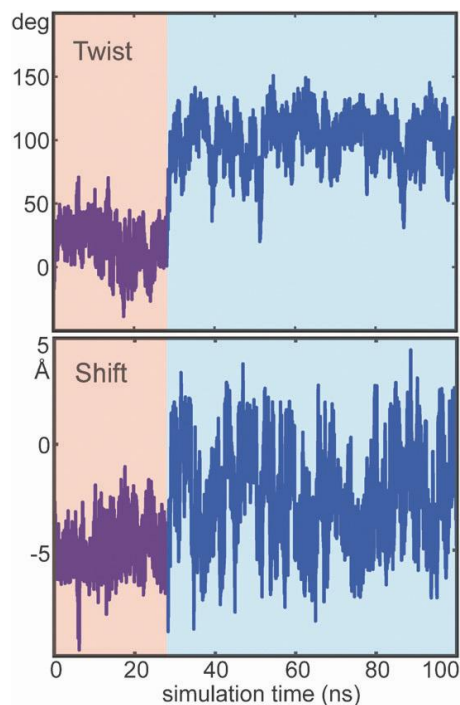
Fig. 6 *The development of the twist and slide parameters characterizing the orientation of the RHO probe vs. the adjacent DNA base pair, during the 100 ns MD simulation. Different conformational substates (see text) are depicted by colour bands (State A in red, State B in blue)*

These findings can be compared with the published structural analysis of the DNA complex with RHO. The NMR data of Neubauer *et al.*[10] indicate one dominant substate (80% population) in which the xanthene ring occupies an intermediate position between guanine and cytosine and the phenyl ring is located on the minor groove side of the adjacent base pair (PDB Code 2V3L, Model 1). In addition they identified a minor substate (PDB Code 2V3L, Model 2) in which the xanthene ring has turned around, so that it is stacked primarily on the guanine and the phenyl ring is located on the major groove side. The MD simulations of Ivanova *et al.*[19] are roughly consistent with these data.

The major NMR conformation was used in the starting structure of our simulated DNA–RHO complex, as described in the Methods section. The new substate reached early in the simulation still resembles the major NMR conformation in that the phenyl ring remains at the minor groove side, but the xanthene ring is stacked on the cytosine instead of being located in between the two bases. After the conformational switch at 28 ns, the new substate closely resembles the minor NMR structure of Neubauer

74

et al. The transition is accompanied by a rotation of the xanthene ring about the DNA helical axis as reflected by the change in twist of 90°. This can be compared to the analogous rotation of 125° between the two bound states in the experiments of Neubauer *et al.*

Our system remains in the new substate until the end of the simulation. The absence of any further conformational transition could be expected, since the experiments by Neubauer *et al.* suggest the time scale of the exchange between the substates in the millisecond range. We emphasize that the xanthene ring fluctuates vigorously, spanning a range of positions and becoming even transiently detached from the base pair. A small population of a detached state was also identified by Neubauer *et al.*, although its exchange rate is even smaller than that the one between the bound states. It is interesting that the major NMR substate, which is supposed to interact strongly with the guanine, actually has a much smaller sterical overlap with the guanine than the minor NMR state. Notice, however, that the electronic interaction is a complex quantum mechanical effect and is not correlated in any simple way with the sterical overlap of the two compounds.

In the case of the DNA–QSY complex, our States Ia and Ib resemble the major substate found by NMR for the DNA–RHO complex, but in our case the xanthene ring prefers to stack over just one of the bases whereas a side ring stacks on the other base. Our State II in which the stacking is mediated by a side ring only is of course absent in the case of Rhodamine 6G, which lacks aromatic side rings.

The molecular mechanics force fields used for our MD simulations were parmbsc0 for DNA and GAFF/RESP/ HF6-31G* for the probe (see Methods section). These are standard choices which have already been tested on different systems. Whereas the use of a polarizable force field might potentially improve the accuracy of the calculations, it seems that the existing polarizable force fields are not yet mature enough to be routinely used. Moreover, they are computationally more expensive, which would make it more difficult to achieve the relatively long simulation times crucial to identify the substates of our   systems and transitions among them.

The influence of the probe on the DNA structure

It is important to verify to what extent the presence of the probe influences the structure of the DNA oligomer to which it is attached. To this end, we compared the structures of the DNA molecule in the three systems, DNA–RHO, DNA–QSY and bare DNA. Since the DNA base sequence and the simulation conditions are exactly the same in the three cases, any difference among the structures can only stem from the influence of the probe or from an insufficient convergence of the simulations.

The DNA structure was described by the mean values of the intra-basepair parameters (buckle, propeller, opening, shear, stretch and stagger)

and the inter-basepair, or the step parameters (tilt, roll, twist, shift, slide and rise), computed over various portions of the trajectory. The parameters for each base pair or step were considered separately. For the DNA–QSY complex, the means were calculated for the following intervals: 1–37 ns (which correspond to States Ia and Ib) and 37–100 ns (where State II prevails). For the DNA–RHO complex, the intervals were 1–28 and 28–100 ns, corresponding to the two observed substates. Thus, the intervals capture the individual conformational substates of each complex. For the bare DNA, we considered the interval of 1–100 ns. MD snapshots in each interval were filtered for broken intra-basepair hydrogen bonds and non-canonical gamma torsion angles as described in the Methods section. In each case at least 60% of the snapshots passed through the filtering. The parameters' means were also computed separately for the two halves of each interval to estimate the error of the mean due to incomplete convergence. The error estimate for the difference of means is then the sum of the estimated errors of the two means. We call a difference statistically significant if it exceeds its estimated error.

The means computed for each interval (or, equivalently, for each substate) of each system were mutually compared. All of the DNA structures were in general very close to each other. Statistically significant differences in the mean parameters exceeding 11 or 0.1 Å only involve the two base pairs and two steps closest to the probe. Their values are as follows: less than 8° for the buckle, less than 6° for the propeller, at most 2° for the twist and roll, less than 0.3 Å for the shift and less than 0.2 Å for the slide. The differences in the other parameters for these pairs and steps did not exceed 1° or 0.1 Å.

Neubauer *et al.* reported an experimental indication of the reduced terminal base pair fraying in the DNA–RHO complex as compared to bare DNA. We have observed an analogous phenomenon in our simulations: in the complexes, the base pair adjacent to the probe is frayed for 1% of the simulation time, but it is frayed for roughly 10% of the time in the case of bare DNA.

We conclude that, apart from the reduced fraying of the base pair nearest to the probe, the influence of the QSY 21 or Rhodamine 6G probes on the DNA structure is very small and is limited to the two base pairs and steps closest to the probe.

The rotation barriers of the QSY 21 aromatic side ring

In contrast to the aliphatic side chains of RHO, the DIHI side rings of QSY 21 significantly contribute to the stabilization of the complex, as seen in the MD simulation. Thus, the knowledge of the energy barrier of the rotation of the DIHI ring in terms of the main xanthene ring would be beneficial. To estimate the barrier height we used a model structure, 2,3-dihydro-1-indolyl benzene (see Fig. 7).
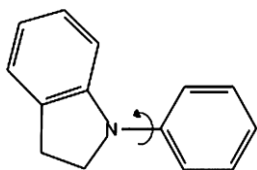
Fig. 7 *The structure of 2,3-dihydro-1-indolyl benzene with a rotatable bond depicted.*

Our model is similar to the biphenyl molecule, which has been extensively studied both experimentally and theoretically. Recent experimental studies indicate that the energy barriers separating the minimum from the planar and perpendicular arrangements of biphenyl are nearly the same, corresponding to the values of 1.4 and 1.6 kcal mol^{-1}, respectively, with the equilibrium torsion angle of 44.4°.[20] Even the highest level *ab initio* methods have a tendency to overestimate slightly the barriers, namely by about 0.5 kcal mol^{-1}.[21,22]

Our results for the rotation of the DIHI ring in QSY 21 show that the structure of the minimum is characterized by the torsion angle of 41.4° calculated at the MP2/cc-pVTZ level of theory. The barrier is slightly larger than for the biphenyl molecule: 3.63 kcal mol^{-1} (the corresponding equilibrium angle is -11.2°) for the ''planar'' form and 2.09 kcal mol^{-1} (equilibrium angle 80.6°) for the ''perpendicular'' form. The analogous data for biphenyl correspond to the values of 2.39 and 2.04 kcal mol^{-1}, respectively. These values suggest a hindered rotation of the DIHI ring in the QSY 21 molecule. The role of the ZPVE energy is marginal, leading to a correction in the relative barrier heights of less than 0.1 kcal mol^{-1}. Not even the effect of the environment is crucial. The rotation barrier is systematically smaller by 0.2–0.3 kcal mol^{-1} in implicit water than in vacuum despite the high-level method and basis set used.

These conclusions are fully confirmed by time evolution of the torsion angles in the course of the MD simulation. In the conformational states where the side ring is not stacked, it rotates around the linker bond. The rotation is not entirely free but the barriers are rather low. This is indicated by the fast transitions among the rotation substates, the lifetime of each substate being of the order of 100 ps to 1 ns. Moreover, the most populated rotation substates correspond to the energy minima as revealed by the QM calculations. The low rotation barriers ensure a good adaptation of the torsion angle in those states where the side ring is stacked, so that the stacking is close to optimal and not compromised by the torsion angle barriers.

Ab initio interaction energies between the probes and DNA

The interaction energy between the probe and adjacent DNA base pair was calculated using two models (see Fig. 2) on geometries obtained at each 100 ps of the MD trajectory, as described in the Methods section. Whereas the small model covers only the interaction between the probe and the adjacent base pair, in the large model the interaction of the linker and the sugar rings is also included. The results are shown in Fig. 8. For the DNA−QSY complex the transition between State Ib and State II in MD simulation is accompanied by a significant decrease of interaction energy by about 8 kcal mol^{-1} in the small model, whereas the States Ia and Ib are isoenergetical and thus undistinguishable in this graph. In the large model the average interaction energy is roughly 11 kcal mol^{-1} greater than in the small one (-33 vs. -22 kcal mol^{-1}) and the energy difference between States Ia/Ib and II is significantly decreased. The less favorable stacking interaction between the probe and the DNA bases in State II is probably compensated by advantageous interactions of the probe and the linker with the DNA backbone, thus making all three states roughly isoenergetic.

The interaction between RHO and DNA is significantly weaker than in the case of QSY 21, due to the absence of the aromatic side rings in RHO. The difference between interaction energies is about 5 kcal mol^{-1} for the small model and 8 kcal mol^{-1} for the large one. No significant differences in stability between the binding motifs observed in the RHO–DNA complexes were found.

Besides the calculations in the gas phase, we also performed computations of the interaction energy in an implicit solvent with dielectric permitivity corresponding to water. The presence of the solvent leads to a systematic destabilization of the complexes, i.e. their interaction energies become less negative. The energies, however, still retain roughly 50% of their original values, and the order of interaction energies of the substates remains unchanged.

This suggests that the order of interaction energies is to a significant extent governed by the solvent-independent dispersion forces. Thus, the use of a more sophisticated solvent model should not have a decisive effect on the relative stability of the substates, as already suggested by the recent work of Vladimirov *et al.*[23]

The proper choice of the QM model used to compute the interaction energies is a delicate problem. The model should be comprehensive enough to capture all the important interactions, but an excessively complex model would not be easy to characterize unambiguously.[24] In our case, the large model describes both the interactions between the probe and the DNA bases and those between the probe and the DNA backbone. Our small model, then, describes the probe–base interactions separately, thus enabling us to separate the influences of the two contributions.
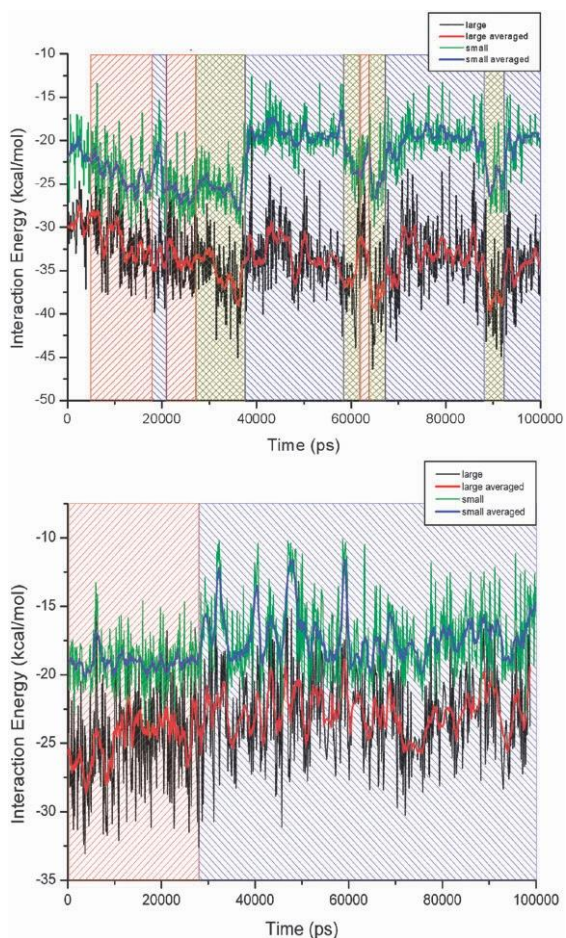
Fig. 8 *The interaction energy between QSY 21 (above) and RHO (bottom) probes with the adjacent base pair obtained from snapshots with a step of 100 ps from the MD simulation. The black curve corresponds to the interaction in the large model, whereas green one to the small model, see Fig. 2. Thick red and blue lines represent sliding averages of interaction energies using a 1 ns window size. Different conformational substates (see text) are depicted by colour bands (for QSY State Ia in red, State Ib in yellow, State II in blue; for RHO State A in red, State B in blue).*

Conclusions

In this work we investigated the structure, dynamics and energetics of molecular complexes involving the fluorescence probes Rhodamine 6G and QSY 21 bound to a DNA oligomer. We used explicit solvent, atomic resolution molecular dynamics (MD) as well as *ab initio* quantum chemical calculations. Two fundamentally different binding motifs have been identified for the DNA–QSY complex. One motif involves the xanthene ring of QSY 21 parallel with the adjacent DNA base pair and stacked on one of the bases whereas a side ring is stacked on the other base. The second motif is characterised by the xanthene ring completely unstacked

79

and pointing towards the minor groove side of the DNA base pair, whereas the stacking with the DNA is mediated only by one of the side rings. The two motifs are almost isoenergetic and multiple transitions between them are observed during the MD simulation. Whereas the first motif is somehow analogous to the main experimental binding motif of Rhodamine 6G, the second motif is new and made possible only by the presence of the aromatic side rings attached to the central xanthene ring in QSY 21. The flexibility of the torsion angles between the xanthene ring and the side rings enables the system to adopt optimal stacking geometries.

The DNA complex with Rhodamine 6G, previously studied experimentally and by short MD simulations, was investigated here for comparison. Our MD simulations, an order of magnitude longer that the published ones, demonstrated a transition to the conformational substate closely resembling the minor substate found by the NMR experiments. Since the exchange between the minor and the major experimental substates takes place at the millisecond scale, no further transition is expected in MD.

The detailed knowledge of the binding motifs of a probe bound to DNA and of the stacking interactions between the probe and the DNA is very important because of the expected changes in the probe's photophysical properties upon binding. The key role of the interaction between the probe and guanine should be emphasized. If the possibility of the probe stacking onto guanine is not considered, these changes could be misinterpreted as a change in donor–acceptor distance.

For a precise interpretation of the FRET experiment, theoretical calculations of the magnitude and orientation of the transient dipole moment of the probe in the vicinity of DNA, as well as a possible role of charge transfer between the probe and guanine will have to be evaluated. This is going to be an important topic in our future work.

Acknowledgements

Notes and references

1   B. Huang, M. Bates and X. W. Zhuang, *Annu. Rev. Biochem.*,2009, **78**, 993–1016.
2   D. E. Schwartz, P. Gong and K. L. Shepard, Biosens. *Bioelectron.*,2008, **24**, 383–390.
3   H. Kim, M. D. Kane, S. Kim, W. Dominguez, B. M. Applegate and S. Savikhin, *Biosens. Bioelectron.*, 2007, **22**, 1041–1047.
4   A. Hassibi, H. Vikalo, J. L. Riechmann and B. Hassibi, *Nucleic Acids Res.*, 2009, **37**, e132.
5   D. C. Swan, R. A. Tucker, B. P. Holloway and J. P. Icenogle, *J. Clin. Microbiol.*, 1997, **35**, 886–891.
6   M. A. O'Neill and J. K. Barton, *Proc. Natl. Acad. Sci*. U. S. A., 2002, **99**, 16543–16550.
7   A. Perez, I. Marchan, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton and M. Orozco, Biophys. J., 2007, 92, 3817–3829.
8   J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
9   M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople*, GAUSSIAN 03* (Revision C.02), Gaussian, Inc., Wallingford, CT, 2004.
10   H. Neubauer, N. Gaiko, S. Berger, J. Schaffer, C. Eggeling, J. Tuma, L. Verdier, C. A. Seidel, C. Griesinger and A. Volkmer, **J. Am. Chem. Soc.**, 2007, **129**, 12746–12755.
11   R. E. Dickerson*, Nucleic Acids Res.,* 1989, **17**, 1797–1803.
12   X. J. Lu and W. K. Olson, *Nucleic Acids Res.*, 2003, **17**, 5108–5121.
13   F. Lankas, N. Spackova, M. Moakher, P. Enkhbayar and J. Sponer, *Nucleic Acids Res.*, 2010, DOI: 10.1093/nar/gkq001, published online ahead of print.
14   F. Lankas, O. Gonzalez, L. M. Heffer, G. Stoll, M. Moakher and J. H. Maddocks, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10565–10588.
15   M. Elstner, *Theor. Chem. Acc.*, 2006, **116**, 316–325.
16   J. M. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
17   R. Ahlrichs, M. Bar, M. Haser, H. Horn and C. Kolmel, *Chem Phys. Lett.*, 1989, **162**, 165–169.
18   S. Grimme, *J. Comput. Chem.*, 2004, **25**, 1463–1473.
19   A. Ivanova, G. Jezierski, E. Vladimirov and N. Rosch, *Biomacromolecules*, 2007, **8**, 3429–3438.
20   O. Bastiansen and S. Samdal, *J. Mol. Struct.*, 1985, **128**, 115–125.
21   J. C. Sancho-Garcia and J. Cornil, *J. Chem. Theory Comput.*, 2005, **1**, 581–589.
22   M. P. Johansson and J. Olsen, *J. Chem. Theory Comput*., 2008, **4**, 1460–1471.
23   E. Vladimirov, A. Ivanova and N. Rosch, *J. Phys. Chem. B*, 2009, **113**, 4425–4434.
24   D. Svozil, J. E. Sponer, I. Marchan, A. Perez, T. E. Cheatham, F. Forti, F. J. Luque, M. Orozco and J. Sponer, *J. Phys. Chem. B*, 2008, **112**, 8188–8197.9684

2.3. Translocation activity of CHD4 protein at single-molecule level

First half of final report from FEBS sponsored stay at the Astbury Centre for Structural Molecular Biology, the University of Leeds in 2010.

Translocation activity of CHD4 protein at single-molecule level

Abstract

The Chromodomain Helicase DNA-binding protein 4 (CHD4) is the main component of the nucleosome remodelling and histone deacetylase complex. CHD4 also plays important role in signalling and repair after double-strand DNA breaks. The mechanisms by which CHD4 interacts with nucleosomes and rewind the DNA remain poorly understood. Here we report a single-molecule FRET characterization of the translocation activity of individual CHD4 protein in real time. We conclude that CHD4 acts as an ATP-driven bidirectional translocase with translocation step size of 2-3 DNA base pairs. We also suggest that DNA recognition and subsequent binding is not ATP-dependent.

Introduction

In eukaryotic cells essential nucleic acid transactions, such as transcription, replication, repair and recombination are performed in the context of chromatin structure. In turn, chromatin structure is maintained and modified by remodelling complexes, which couple the energy of ATP hydrolysis to the assembly and mobilization of nucleosomes.

ATP-dependent chromatin remodelling complexes can be divided into several subfamilies (e.g. ISWI, SWI/SNF, CHD/Mi2, INO80 and SNF2/RAD54) depending on their composition and function [1]. Different subfamilies exhibit diverse remodelling activities. CHD4 (also called Mi2-beta) is the main component of the NUcleosome Remodelling and histone Deacetylase complex (NURD). Protein has importance also in medicine; the exact role is not yet fully understood, but patients with dermatomyositis develop antibodies against CHD4 [2].

CHD4 contains two Plant HomeoDomain (PHD) motifs, followed by a tandem of chromodomains (DNA and a SNF2 type helicase domain. Recently, it was demonstrated that the PHD domains bind to histone H3 N-terminal tail while the chromodomains interact with DNA [11]. However, intermediates and pathways through which CHD4 remodels nucleosome structure remain mostly unknown.

Moreover previous year, the role of phosphorylated CHD4 in signalling and repair after double-stranded DNA breaks was published. Urquhart *et al.* worked with CHD4 as a target of ATM kinase, as ATM kinase activity is a primary driving force for chromatin alterations emanating from double-stranded break induction. They demonstrated that phosphorylation of CHD4 Ser-1349 resulted in an increased chromatin retention and suggested that ATM-dependent phosphorylation might result in CHD4 becoming more tightly associated with the chromatin and this may allow for opening of the chromatin structure and subsequent repair [3]. Larsen *et al.* reported CHD4 as a factor that becomes transiently immobilized on chromatin after ionizing radiation, which is known to cause double strand breaks. Thus, CHD4 emerges as a novel genome caretaker and a factor that facilitates both checkpoint signalling and repair events after DNA damage [4].

The basic facts about CHD4 as ATP-dependent translocase were not known. Since we had CHD4 construct and nucleosome binding DNA sequence, we could solve subsequent questions.

1) Is the translocase directional towards 3', 5' or bidirectional?
2) How long are the dwell and translocation times, and how are they dependent on ATP concentration?
3) What is the size of the translocation step?
4) Is the binding and translocation of CHD4 protein to nucleosomes and DNA ATP-dependent?

Materials and methods

Protein purification and labeling

CHD4 construct prepared by Mora [11] was stored as frozen stock solution. When defrosted it was re-purified by size exclusion chromatography to remove possible aggregates
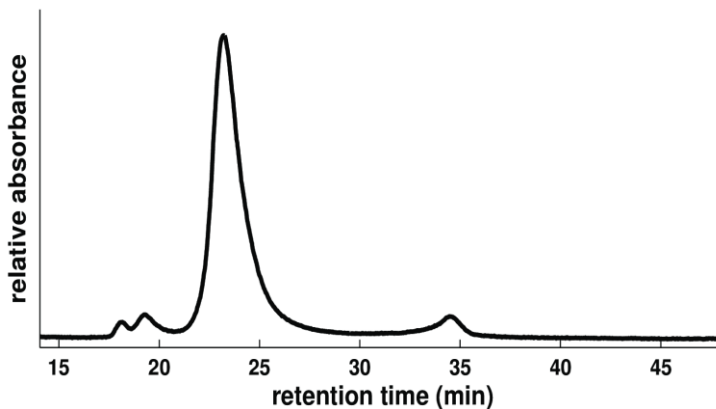

Figure 1: Liquid chromatography elution profile (A_{280}) of CHD4 before labeling. The stock of CHD4 was purified by size exclusion chromatography on Superdex 200 column, flow 0.2 mL/min in 20 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT at room temperature. CHD4 was recollected between 21st and 27th minute.

Afterwards, CHD4 construct was labeled with an acceptor dye. First we tried ATTO 590 maleimide, but probe's emission was quenched by the protein itself, as a second attempt, we successfully used Alexa Fluor 590 maleimide. Maleimide reacts with cysteines on CHD4 surface. Cysteines were chosen, since they are least abundant amino acids in our construct, the proportion is 1.4%, that means 22 cysteines per protein molecule and only few of them are exposed on the surface without any steric hindrance handicapping the labeling reaction.

Labeling procedure followed manufacturer's protocol. Consequently labeled constructs were purified by dialysis overnight and stored in darkness at 4°C. We chose

dialysis instead of LC to prevent the dilution of sample. Labeling efficiency was quantified by absorbance at protein abs. maximum 280 nm and dye abs. maximum 602 nm, as dye spectrum was red shifted, after the conjugation to protein. The molar ratio of label to protein was 1.9.

Preparation of labeled 601 sequence

The 601 sequence template was prepared by digesting pGEM-3z/601 plasmid with BamHI and HindIII.

<div align="center">(a)   (b)</div>
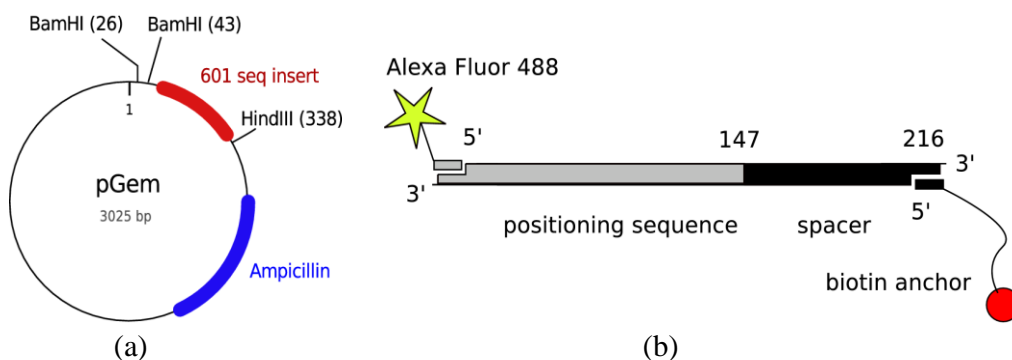
Primers: *Forward: 5'- AF 488 -GCCCTGGAGAATCCCGGTGC -3'*
               Reverse: 5'- Biotin -CAGGTCGGGAGCTCGGAACACTATC -3'

Figure 2: (a) Plasmid pGEM-3z/601 (Addgene plasmid 26656). This plasmid contains an 147 bp nucleosome positioning sequence and is referred to as clone "601" [5]. (b) Preparation of labeled and biotinilated construct. Isolated 601 sequence was modified with Alexa Fluor 488 and biotin via PCR.

The sequence was amplified by PCR using Alexa Fluor 488 modified and biotinylated primers. The labeling efficiency and emission brightness of donor dye attached to DNA was checked through absorption and emission spectrum.

Ensemble FRET

Prior to single molecular experiments, ensemble FRET was examined. We determined the efficiency of Förster resonance energy transfer using serial dilutions of FRET partners and ATP. FRET was calculated as:

$$E_{FRET} = 1 - \frac{F_{D(A)}}{F_D},$$
(Eq.1)

where E_{FRET} is FRET efficacy, $F_{D(A)}$ is fluorescence of donor in the presence of acceptor and F_D is fluorescence of donor without acceptor. The emission spectra of donor were corrected for dilution, quenching and direct excitation of acceptor. With increasing molar excess of protein to DNA (1:1, 6:1 and 11:1), FRET efficiency asymptotically approached ~25%. Ensemble FRET was not sensitive to ATP concentration and stood constant over time.

Ensemble measurements were performed using a Shimadzu fluorometer model RF/5301PC. The sample was excited at 488 nm and emission spectra were measured between 495 and 700 nm at room temperature.

Single-molecule experiments

DNA immobilization

Glass cover slip surfaces were cleaned and prepared using a method of Ha and co-workers [6]. Cover slips were incubated with Vectabond reagent (Vector Laboratories, Burlingame, USA) according to the manufacturer's instructions after which they were coated with 25% (w/v) polyethyleneglycol succinimidyl ester (PEG-NHS, Rapp Polymere, Tubingen, Germany) and 0.25% (w/v) biotinylated-PEG-NHS (Rapp Polymere) in 0.1M sodium bicarbonate (pH 8.3) for 3 h in order to eliminate non-specific adsorption of proteins. The functionalized surface was then rinsed with 10 mM Tris, pH 7.4 and incubated with 0.2 mg/ml Immunopure-streptavidin (Pierce Biotechnology, Rockford, USA) in the same buffer for 1 h and rinsed with 3 x 400 μl of buffer A (25 mM Tris acetate, pH 8.0 with 8 mM magnesium acetate, 100 mM KCl, 1 mM dithiothreitol, and 3.5% (w/v) poly-ethylene glycol 6000). Excess buffer was blotted away and Alexa Fluor 488 labeled, biotinylated DNA in buffer A was applied to the surface and allowed to bind for 30 min. Excess unbound DNA was then removed by blotting and rinsing with three times 400 μl of buffer A. Density of immobilized DNA molecules on cover slip was optimized (Figure 3) using final concentration of 158 nM. To avoid photo-destruction, we held samples in anti-photobleaching cocktail; 1.25 mM propyl gallate, 5 mM DTT, 5 mM cysteamine, 1.5 mM β-mercaptoethanol [7] [8].
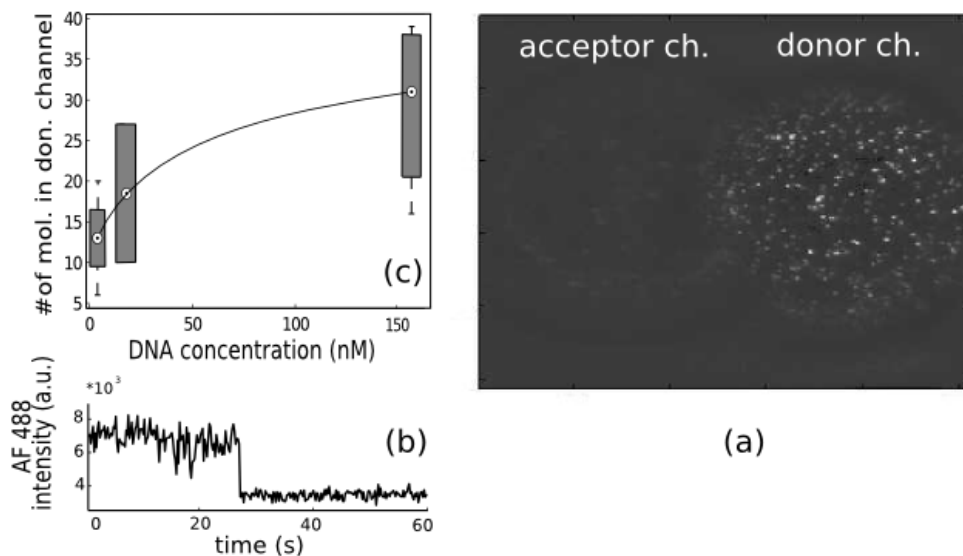

Figure 3: DNA immobilization procedure optimization. DNA concentration of 158 nmol gave satisfactory results with distinguishable molecules in sufficient quantity (other examined concentration: 15.8 and 3.2 nmol). (c) Box and whisker plot (median, 25th, 75th percentiles, extreme data points) optimized DNA concentration. (a) Representative TIRFM image of sample prepared with opt. DNA concentration and laser power 1.6 mW. Single spots representing individual molecules appeared in donor channel with minimal leak to acceptor channel (b) Analyzed molecules showed single step photobleaching behaviour typical for single molecule fluorescence.

SpFRET total internal reflection fluorescence microscopy

TIRF enables simultaneous detection of immobilized single molecules since it eliminates background from bulk solution. The collected images were processed in MATLAB (MathWorks) to yield FRET efficiencies as a function of time for each molecule. The efficiency of energy transfer was calculated as FRET proximity ratio:

$$P = \frac{I_A}{I_A + I_D},$$

(Eq.2)

where P is proximity ratio, I_A and I_D are the corrected acceptor and donor signals, respectively, in each integration time [9]. Single molecular experiments were performed on custom build TIRFM FRET instrument. The set-up is equipped with a high numerical aperture (NA) microscope objective (100x, 1.45 NA infinity corrected oil immersion Plan-Fluar, Zeiss, Germany), Coherent Sapphire 488 nm excitation laser and electron-multiplying charge-coupled-device camera (EM-CCD, iXon, Andor Technology) with suitable filter set spectrally splitting the photon stream into donor and acceptor channels. More detailed description of the instrument can be found in Gell's book [10]. The repetition rate of image collection was 5 frames per second. For single molecule FRET experiments, we optimized the protein concentration to 333 nmol, at this value, acceptor signal was optimal and background from the bulk solution was still minimal.

Results and discussion

When examining immobilized labeled DNA molecules alone, significant signal appeared only in donor channel, after addition of labeled protein, spatially correlated spots appeared also in acceptor channel (Figure 4). The proximity ratios calculated from the time trajectories corresponding to the correlated spots showed stable FRET values and single, rarely double step photobleaching.

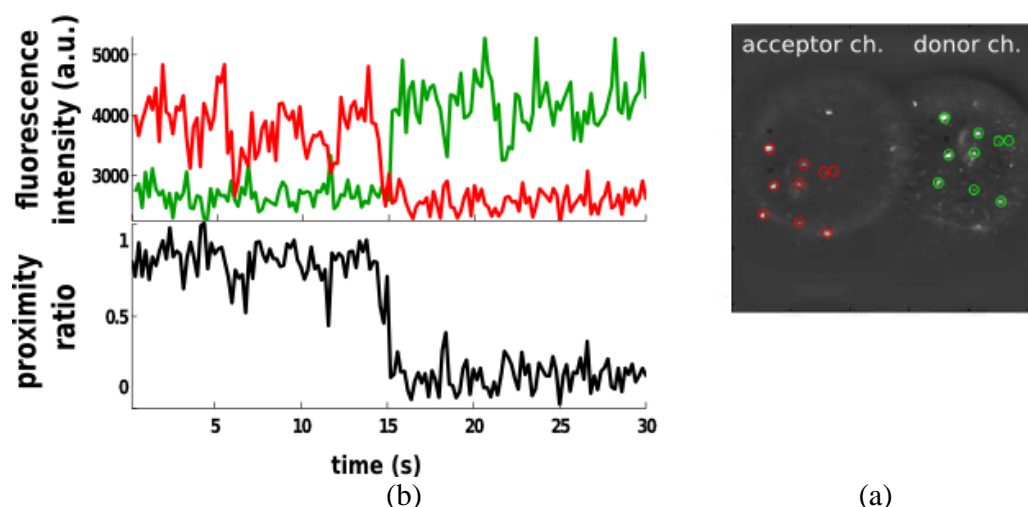(b)   (a)

Figure 4: Total internal reflection fluorescence microscopy of immobilized DNA molecules labeled with donor dye and bound CHD4 labeled with acceptor. (a) The two spectral channels show the same area of a sample recorded simultaneously for green (Alexa Fluor 488) and red (Alexa Fluor 594) fluorescence. Spatially correlated features were observed, they are highlighted with coloured circles. (b) In the upper part, there are intensity trajectories

processed from a TIRFM recorded image sequence. The anticorrelation of donor and acceptor emission is clearly visible. Single step acceptor photobleaching was captured at the 14th second. At this point red (acceptor) fluorescence intensity falls down (photobleaches) and green (donor) signal increases, as the energy transfer is interrupted. This is also recorded by proximity ratio time trajectory (bottom part), as a transition from values fluctuating about 0.85 to effectively zero.

The above results indicated that no stepping or translocation events occurred without ATP. This was expected, as CHD4 is ATP driven translocase, therefore we titrated labeled and immobilized DNA protein complexes with increasing concentration of ATP and searched for stepping pattern in proximity ratio resulting from the protein moving along DNA. Stepping trajectories such as that shown in Figure 5 can be explored in terms of translocation direction, dwell time, translocation time and the step size (Figure 5).
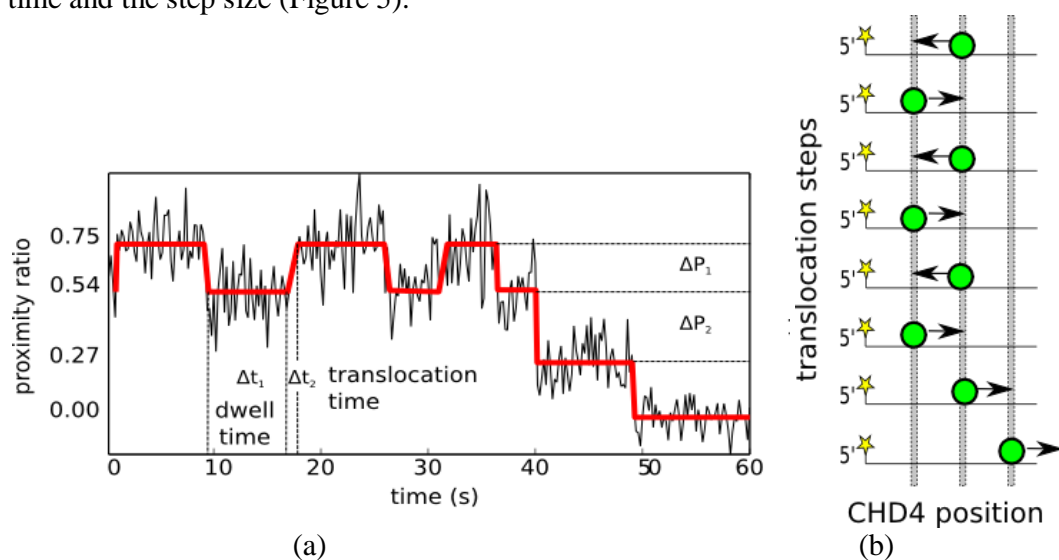

(a)   (b)

Figure 5: (a) Manual detection of dwell and translocation times. Dwell time (Δt_1) represents waiting time till next step occurs. Translocation time (Δt_2) represents time necessary to translocate during one step without waiting period. This particular time trajectory captured one immobilized DNA molecule with bound CHD4 after addition of 1 mM ATP. Three proximity ratio levels were occupied (0.75, 0.54 and 0.27), CHD4 construct translocated to proximity ratio 0.75, then moves to 0.54, then back to 0.75, back to 0.54, again to 0.75, back to 0.54, than to 0.27 and then the protein moved out of range of this FRET pair coverage. ΔP_1 and ΔP_1 represent difference in proximity ratio within closest occupied proximity levels. (b) Scheme of bidirectional translocation consistent with the time trajectory in (a). Labeled DNA is represented by a black solid L-shaped line with a yellow star (Alexa Fluor 488) and labeled CHD4 construct by a green circle, three detected positions are highlighted with grey zones, arrows represent the direction of translocation.
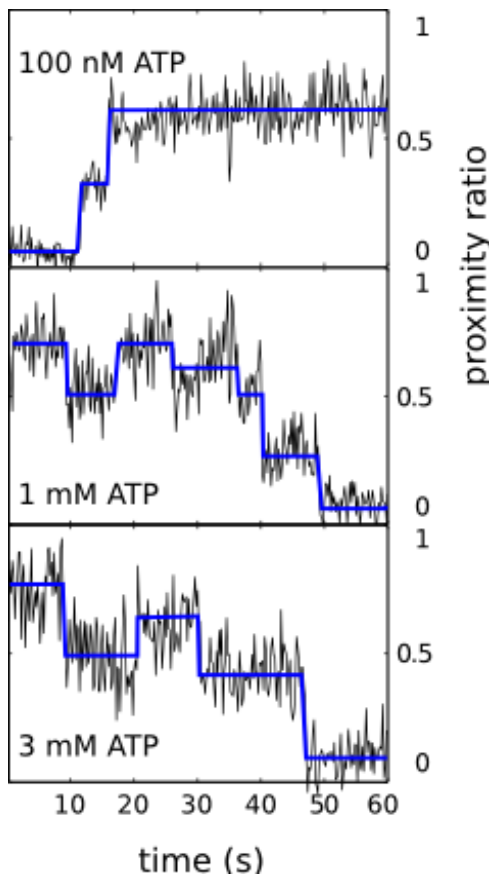
Figure 6: Time trajectories of FRET proximity ratio show step behaviour in the presence of ATP. Without ATP, trajectories have simple flat profiles. Multiple steps were observed only when ATP was added. Representative trajectories for particular ATP concentrations (100 nm, 1 mM and 3 mM) are shown. Major part of trajectories showed bidirectional steps.

We processed all trajectories, collecting dwell times, translocation times and step sizes (Figure 5). Counter intuitively, with increasing ATP concentration we observed an increase in the dwell time and also in the translocation time (Figure 7(a) and 7(b)). From the FRET trajectories, it is clear that CHD4 behaves as a bidirectional translocase (Figure 5 and 6). Using Förster radius of 60 Å, the translocation step size was estimated as 8.4 ± 0.9 Å (mean ± standard deviation). The axial rise per residue for the b-form DNA is 3.34 Å thus we estimated the step size to 2-3 base pairs.
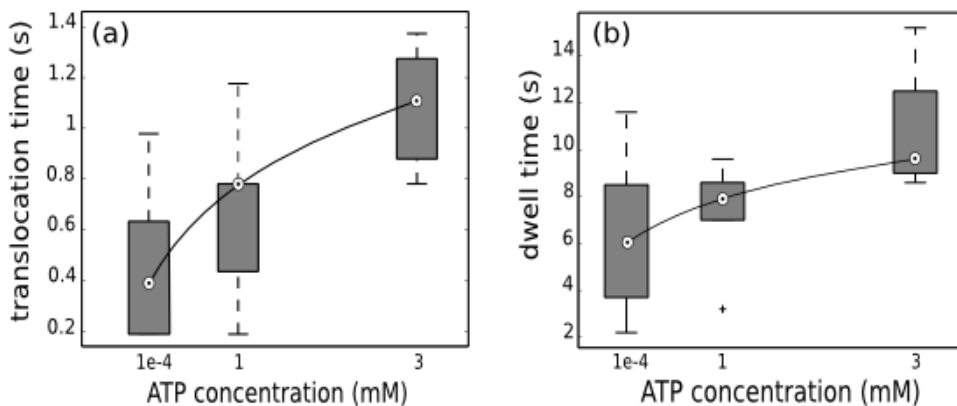

Figure 7: Kinetic characteristics of CHD4 DNA interaction when titrated with ATP. a) Dwell time increased with increasing ATP concentration. b) Translocation time increased with increasing ATP concentration.

From the presented experimental results, we can conclude that binding of CHD4 to DNA is not ATP-dependent, since molecules of labeled protein were visible

after addition of CHD4 only without ATP.

On the other hand translocation of CHD4 along DNA molecule is ATP-driven. Without ATP, no steps in FRET trajectories were detected. After addition of ATP, trajectories showed stepping behaviour, typical for movement of molecular machine.

From the proximity ratio steps, the directionality of CHD4 as translocase can be detected. As these steps were detected in both directions i.e. towards 5' end and also 3' end of DNA, CHD4 seems to be bidirectional translocase. Translocation step size was estimated to be 2-3 base pairs per step.

Hence we can summarize that CHD4 protein without phosphorylation moves along DNA in both directions as molecular motor driven by symmetric chemical potential. Its localization/immobilization at particular foci as in case of double-stranded break signalization and repair is probably controlled by posttranslational modifications as phosphorylation or interaction with other proteins or modified histones, creating a particular functional complex.

Acknowledgements

References

1. Blosser, T.R., et al., *Dynamics of nucleosome remodelling by individual ACF complexes.* Nature, 2009. **462**(7276): p. 1022-U79.
2. Zhang, Y., et al., *The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities.* Cell, 1998. **95**(2): p. 279-289.
3. Urquhart, A.J., et al., *ATM mediated phosphorylation of CHD4 contributes to genome maintenance.* Genome Integr, 2011. **2**(1): p. 1.
4. Larsen, D.H., et al., *The chromatin-remodeling factor CHD4 coordinates signaling and repair after DNA damage.* J Cell Biol, 2010. **190**(5): p. 731-40.
5. Lowary, P.T. and J. Widom, *New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning.* Journal of Molecular Biology, 1998. **276**(1): p. 19-42.
6. Ha, T., et al., *Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase.* Nature, 2002. **419**(6907): p. 638-641.
7. Sheetz, M.P. and D.E. Koppel, *Membrane Damage Caused by Irradiation of Fluorescent Concanavalin-A.* Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(7): p. 3314-3317.
8. Giloh, H. and J.W. Sedat, *Fluorescence Microscopy - Reduced Photobleaching of Rhodamine and Fluorescein Protein Conjugates by Normal-Propyl Gallate.* Science, 1982. **217**(4566): p. 1252-1255.
9. Tezuka-Kawakami, T., et al., *Urea-induced unfolding of the immunity*

protein Im9 monitored by spFRET. Biophysical Journal, 2006. **91**(5): p. L42-L44.

10. Christopher Gell, D.B.a.A.S., *Handbook of Single Molecule Fluorescence Spectroscopy.* 2006.

11. Morra R, Lee BM, Shaw H, Tuma R, Mancini EJ., *Concerted action of the PHD, chromo and motor domains regulates the human chromatin remodelling ATPase CHD4.* FEBS Lett. 2012 Jul 30;586(16):2513-21. Epub 2012 Jun 27.

2.4. Study of the MS2 phage assembly mechanisms using FCS.

Second half of final report from FEBS sponsored stay at the Astbury Centre for Structural Molecular Biology, the University of Leeds in 2010.

Output:

Poster:

"Exploring the mechanism of single-stranded RNA virus assembly using single molecule spectroscopy and ion mobility spectrometry-mass-spectrometry"; 3rd Mathematical Virology Workshop, 16th - 20th August 2010 Ambleside, UK

Study of the MS2 phage assembly mechanisms using fluorescence correlation spectroscopy

Abstract

Bacteriophage MS2 serves as a model for the assembly of other single-stranded RNA viruses. Capsid assembly is non-equilibrium heterogeneous process involving intermediated states. Most of the methods previously used to follow this process (e.g. electron microscopy and dynamic light scattering) have biased sensitivity towards differently sized components of the assembly reaction and they demand relatively high concentration of capsid protein. To overcome these limitations, we employed Fluorescence Correlation Spectroscopy (FCS) as a highly sensitive real-time method to resolve the assembly of viral capsids. We further tested two statistical methods: probability distribution fitting and singular value decomposition to evaluate the progress and outcome of capsid assembly, as detected via FCS. FCS was sensitive enough to observe a temporary kinetic trap. Use of advanced data processing is necessary to reveal the time dependent changes of stoichiometry of individual assembly components.

Introduction

Single-stranded RNA (ssRNA) viruses are complex RNA processing systems that perform packaging, replication and transcription of their genomes. Many of these viruses (e.g. human rhinovirus causing common cold, also poliovirus) pose a menace to mankind having a great social and economic impact (e.g. foot and mouth disease virus). Understanding mechanisms underlying ssRNA virus assembly can help to define new antiviral drug targets for different viral diseases and provide new tools for design and testing. Moreover, the assembly models might be used for creation of a new nanostructure and viral capsids are currently under investigations as vehicles for targeted drug delivery and vaccine development.

Due to the extensive biochemical and structural information available, MS2 bacteriophage is one of the best characterized virus assembly models. MS2 is a member of the Leviviridae family of viruses that infect male Escherichia coli cells via an initial attachment to the bacterial F-pilus.

MS2 has a single-stranded, positive-sense RNA genome of 3569 nucleotides that encodes four gene products: coat protein (CP), replicase, lysis protein and maturation protein. CP is the most highly expressed of the gene products and 180 copies assemble to form an icosahedral protein shell that encapsidates the genome in the mature virion. A single copy of the maturation protein is also incorporated into the virion and during infection binds to the F-pilus of target cells. The main chain of the coat protein (CP) has a ten-stranded antiparallel beta sandwich at the N-terminus followed by a kinked alpha -helix at the C-terminus. The C-terminal alpha-helices of two CP monomers interdigitate to form a non-covalent stable dimer (CP2) (Figure 1). The dimer can be isolated from the virion by acid dissociation and constitute the building block for self-assembly *in vitro*. In the T=3 capsid, the CP is found in three

distinct conformations, termed A, B and C, consistent with the quasi-equivalent symmetry required to construct a T=3 structure. The capsids contain two types of dimer: an asymmetric A/B dimer, and a symmetric C/C dimer (Figure 1(b)).

It was shown that capsid reassembly can be triggered *in vitro* by a sequence-specific RNA-protein interaction [1]. Acid disassembled CP protein which forms dimers can bind a stem-loop RNA called TR of 19 nucleotides from the MS2 phage genomic sequence initiating virus assembly process. RNA-protein binding thus achieves two functions via a single molecular recognition event: translational repression of the replicase and creation of an assembly competent A/B dimer complex on the viral RNA [1] [2].

Using non-covalent mass spectrometry it was demonstrated that the CP2 TR-RNA complex is kinetically trapped, i.e. that it is unable to assemble beyond this point in the absence of other molecular species [2]. The effect of the trap can be relieved by subsequent addition of RNA-free CP2 to this complex, resulting in the rapid appearance of higher-order assembly intermediates [3]. Equivalent behaviour can be observed when recombinant T=3 shells are analyzed [1]. This result suggested that the RNA-free and bound forms of CP2 were different conformers, in particular that they mimicked the symmetric (C/C) and the asymmetric (A/B) conformers needed to define the T=3 shell [1].

Overall, the simplified assembly model for the MS2 phage can be shown as on Figure 1(a). A pair of C/C-subunits (top left, shaded pink) binds a TR-RNA resulting in a conformational switch into an A/B-like asymmetric dimer (shaded green and blue). This complex binds an additional C/C-like dimmer creating the first higher order intermediate on the pathway to capsid formation. A variety of sub-pathways involving addition of CP dimers (arrows) can then generate the major assigned intermediate comprising the first 3-fold axis of the virus particle.
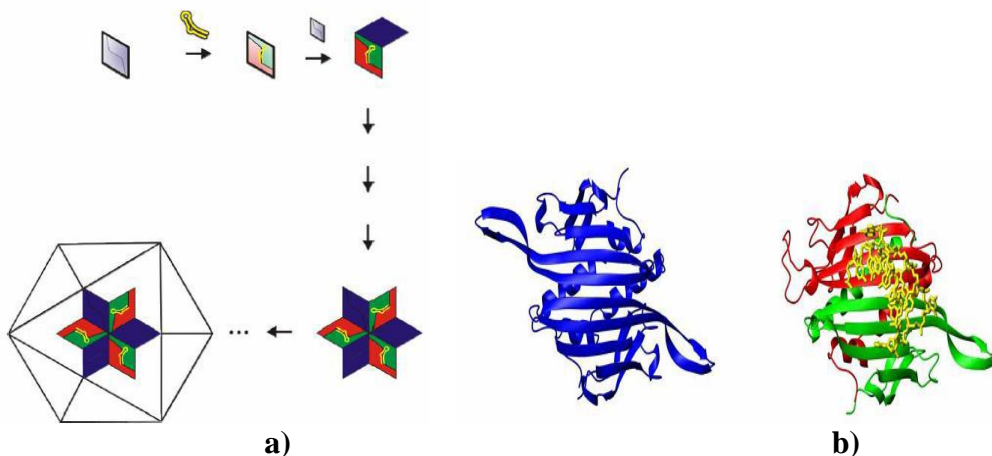

a)   **b)**

Figure 1: Model for the initiation of T=3 capsid assembly. In the right panel: Structures of the C/C (blue) and A/B (red/green) MS2 quasi-equivalent CP dimers. The TR-RNA (yellow stick model) is shown bound to the A/B dimer. Both images adapted from Rolfsson et al *[2].*

From previous work, it appears that no substantial assembly beyond the initial RNA-coat protein dimer complex occurs up to five minutes [4] [5]. Several techniques have been applied to follow virus assembly in real time. Among them are Ion-Mobility ElectroSpray-Ionisation Mass Spectrometry (IM-ESI-MS) [6] [7], light scattering [2] [8] and Dynamic Light Scattering (DLS) [9] [10]. The main disadvantage of these techniques is their biased sensitivity towards differently sized components of the assembly reaction and the need for relatively high protein concentrations. Under those conditions the protein-protein interactions dominate over RNA-protein recognition. Therefore, these techniques do not give relevant stoichiometric data about different intermediates during the assembly process and obscure the role of RNA in the assembly mechanism. Other, indispensable, drawback of these methods comprises the inherent impossibility to employ them in *in vivo* environment.

To overcome this disadvantage we utilized Fluorescence Correlation Spectroscopy (FCS) [11] to follow the MS2 assembly in real time. Its high sensitivity, low required concentrations, and the very small observed volume offer an effective tool to obtain assembly kinetics. FCS relies on measurement of diffusion time (~hydrodynamic radius, R_h) of fluorescently labelled particles passing through the confocal volume. Moreover the technique is not seriously affected by the presence of various unlabeled molecules or molecular complexes. This feature has facilitated FCS applications within living cells [12].

The result of a typical FCS experiment is an AutoCorrelation Function (ACF) which reflects fluctuations due to diffusion and certain photophysical processes such as triplet state formation. In order to extract meaningful information such as the hydrodynamic radius it is necessary to fit ACF to a model parametric function accounting for all these processes. This usually allows for one or two species with distinct Rh to be identified. Nevertheless this standard approach to FCS data processing does not provide full information about highly heterogeneous systems such as assembly reactions. Therefore a novel approach was needed.

In this work we have collected time resolved FCS data of MS2 assembly and explored three approaches to ACF decomposition:

1) Regularized Inverse Laplace Transform (RILT) procedure that computes a probability distribution of diffusion times consistent with the ACF data

2) Two-state system model. Simple model testing hypothesis, that no detectable intermediates are present in the capsid assembly time progression.

3) Singular Value Decomposition (SVD), matrix factorisation based on linear algebra, which can reveal the information about the minimum number of components explaining the system at given noise level and concentration loadings of these components.

Material and methods

My contribution to this project consisted of development of a mathematical software package for data processing of FCS record of the time assembly series, software testing and data analysis. Biochemical part of work, EM and FCS measurements were done by Msc. Alexander Borodavka from Dr. Tuma's lab.

Distribution fitting by regularized inverse Laplace transform

FCS gives us the molecular diffusion times and with known confocal volume, we can calculate the hydrodynamic radius, thus derive the size of the diffusing particles (e.g. assembly intermediates). Because the assembly reaction is not synchronized, a complex mixture of different sizes ranging from the coat dimer up to the complete capsids is expected. In such a mixture of particles with different diffusion coefficients, the ACF is a sum of the contributions of the individual species.

Mathematical approach we utilized for obtaining a size distribution from a given ACF was regularized distribution fitting. This method relies on a customized inverse Laplace transform which is computed using a regularized least squares procedure. The FCS autocorrelation function can be approximated by an integral over the diffusion times in this form:

$$G(\tau) = \left(1 + \beta \cdot e^{-\lambda}\right) \cdot \left(\frac{1}{N} \int_{\tau_{min}}^{\tau_{max}} \rho(\tau_{diff}) \cdot g(\tau, \tau_{diff}) \cdot d\tau_{diff}\right) + 1, \qquad (Eq.1)$$

where $G(\tau)$ is autocorrelation function of fluorescence intensity time fluctuation, β is amplitude of the triplet term, λ is the triplet relaxation rate, N is mean number of particles in the excitation volume element, The $\rho(\tau_{diff})$ is desired normalized distribution, τ_{diff} is diffusion time, τ is macrotime and $g(\tau, \tau_{diff})$ is the core term expressing diffusion part of ACF defined by equation:

$$g(\tau, \tau_{diff}) = \left(1 + \frac{\tau}{\tau_{diff}}\right)^{-1} \cdot \left(1 + \frac{\tau}{\tau_{diff} \cdot \kappa^2}\right)^2, \qquad (Eq.2)$$

where K^2 is the axial ratio of the confocal volume, z_0/w_0, where w_0 and z_0 are the half-widths of the focus in the x–y plane, and in the z-direction, respectively. K^2 was determined experimentally from diffusion of the free dye Rhodamine G.

The task is to find the inversion of the integral transformation in Eq. (1), obtaining the distribution $\rho(\tau_{diff})$ from a measured, often noisy ACF. Inversion problems of this kind are known to be highly ill-posed, i.e. an infinite number of solutions compatible with experimental data exists, which may deviate considerably from each other. However, the distribution $\rho(\tau_{diff})$ in Eq. (1) is positive (i.e. negative values are not allowed) and additive. For finding such a distribution function that is physically reasonable and compatible with the data, a variety of regularization methods exist [13]. Regularization in our routine is based on minimization of the second derivative of distribution vector with weight, which is square of chosen regularization parameter (adopted from Provencher [14]). Minimized term (I) is:

$$I = L + \text{Re} \, g, \qquad (Eq.3)$$

where L is the likelihood of the fit, given by the sum of the squared deviations of the model points from the data (M points):

$$L = \sum_{j=1}^{M} \left(G_{theo}(\tau_j) - G_{\exp}(\tau_j) \right)^2 \qquad (Eq.4)$$

and Reg is so called regularizor:

$$\operatorname{Re} g = \alpha^2 \sum \left(w \cdot (r - R \cdot \rho(\tau_{diff}))^2 \right), \qquad (Eq.5)$$

where α is the regularisation parameter, w are weights, r is second derivative of $\rho(\tau_{diff})$ and R is a matrix determining the form of the regularizor. With higher regularization parameter, the resulting distribution is smoother and regular. Weights can be also sensitive to signal to noise ratio (SNR) in different parts of the ACF curve.

Similar method was previously successfully used for DLS data processing [15]. The lack of this approach for FCS data analysis has been chiefly due to the complications arising from the non-diffusional variables of ACF such as triplet relaxation time (β) , triplet amplitude (λ) and number of molecules in focal volume (N) (see Eq. (1)). This flaw was solved by built in subroutines utilizing discrete nonlinear parametric fit of particular ACF, which gives discrete values for non-diffusional variables [13]. These subroutines can call fits with single or multiple diffusion time components.

Two-state system model

In the absence of detectable intermediates the composition of a dynamic system at any time can be described as a superposition of the starting and the final state.

In our particular case this means that each of the ACF shall be a linear combination of the starting (t=0) ACF (CP2:TR RNA + addition of excess of CP2) and the end point (t=∞), representing capsids. In matrix formulation:

$$I = C \cdot I^{0\infty}, \qquad (Eq.6)$$

where matrix $I^{0\infty}$ contains the two row vectors corresponding to I(0) and I(∞), C is the unknown matrix containing the respective concentrations and I contains the measured FCS data. The unknown concentrations can be obtained by standard linear least-squares procedure (^{T} indicates matrix transposition):

$$C = I \cdot I^{0\infty T} \cdot \left(I^{0\infty} \cdot I^{0\infty T} \right)^{-1}. \qquad (Eq.7)$$

And the reconstructed data I' could be obtained by substituting the concentration matrix obtained from least-square procedure into Eq. 6. From the difference I-I' compared with noise, we evaluated, whether the two states are sufficient to explain the system.

Singular value decomposition

The SVD theorem states that for a real m by n matrix, here I (m×n) representing set of collected ACFs (Figure 5), there exist orthogonal matrices U (m×m) and V (n×n) such as:

$$U^T \cdot I \cdot V = S, \qquad (Eq.8)$$

where T indicates matrix transposition, S is a diagonal matrix with decreasing non-negative diagonal values. Selecting only a limited number (k) of s values and setting the rest to zero suffices to approximate I with a matrix I(1:k) of rank k. The effective rank is equal to the number of linearly independent components that are sufficient to

approximate the whole data set within the set error, i.e. noise level [17].

Results and discussion

RILT distribution fitting

The RILT procedure was tested on free dye (Alexa Fluor 488) and purified reaction components (CP2, CP2:TR complex and fully assembled capsids). The distribution maxima (upper panel of Figure 2) did not significantly differ from single component discrete fits. After successful testing, we employed the routine to process capsid assembly time series. During the reassembly process, the distribution broadened and additional peaks occurred (lower panel of Figure 2).
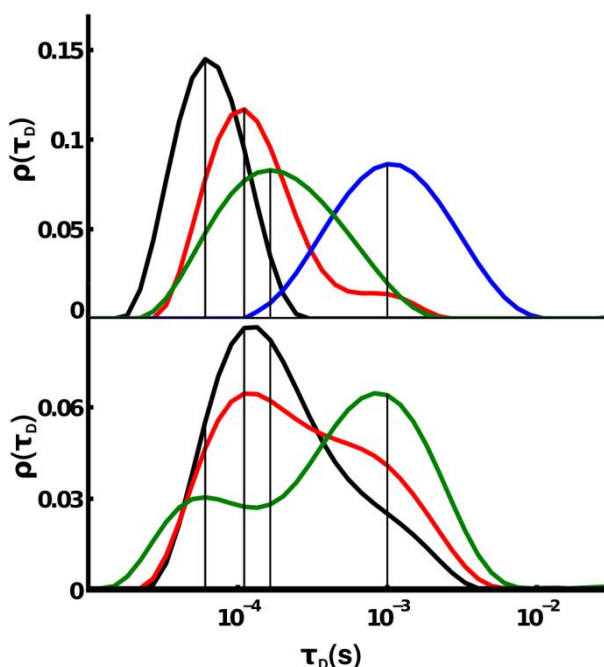

Figure 2: Upper panel. Diffusion time distributions of free Alexa Fluor 488 (black curve), CP2 (red curve), CP2:TR-RNA complex (green curve) and purified capsids (blue curve). The distributions of purified components of assembly reaction are fully distinguishable. The diffusion times represented by their peak maxima correspond to the discrete fit results. Lower panel. Diffusion time distributions of representative assembly times after initiation of reassembly by adding of excess of CP to the kinetically trapped CP2:TR-RNA complex, 0 min – black curve, 55 min – red curve, 111 min – green curve. Diffusion times corresponding to individual components are highlighted by vertical lines. Experimental data were noisy, therefore higher regularisation parameter ($\alpha=0.7$) was required to obtain stable solution.
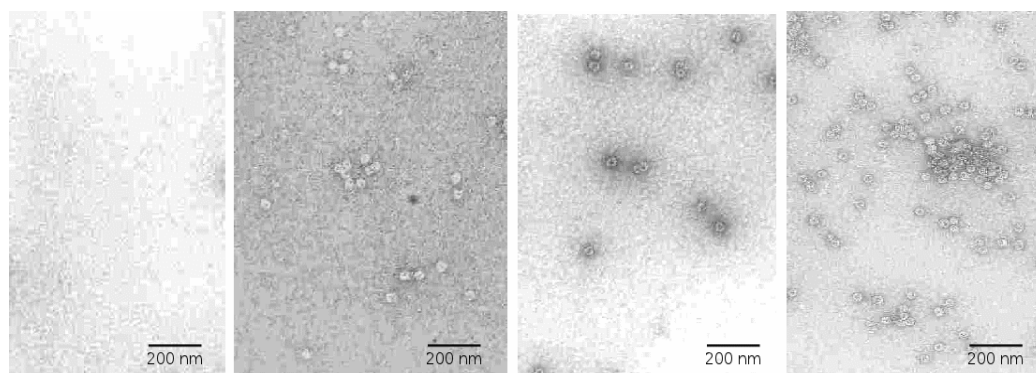
| Dissociated coat protein | 45 min | 120 min | 240 min |

Figure 3: Electron Microscopy examination of labeled MS2 assembly. Micrographs of dissociated coat protein (CP) and assembly mixture; 45, 120 and 240 minutes after assembly initiation.

As seen in the Figure 2, the distributions in the assembly series can be explained by a combination of contributions from dimers and fully assembled capsids with a small contribution from intermediates. Assembly kinetics can be derived by simple extraction of the peak intensity for each distinct component (see Figure 4(a)).

Initial CP2 TR-RNA concentration affects capsid formation kinetics

Current results (Figure 4(a)) confirm the sensitivity of the assembly time-trace to the initial assembly component stoichiometry. Two initial concentrations of the coat protein were employed and resulted in dataset 1 and 2, respectively (Figure 4). System with the higher initial concentration resulted in a reversible temporary kinetic trap (transient lag phase). After leaving the kinetic trap, the system progressed to completion. This illustrates that the method is sensitive to delineate complex kinetic behaviour.

Resulting assembly time progression was also investigated by Electron Microscopy (EM) (see Figure 3). EM roughly confirms the process estimated by RILT. Nevertheless, the presence and nature of intermediates cannot be directly derived from the RILT procedure. This is due to the noise affecting the data, adherent to usage of higher regularisation parameter; hence we tested different data processing methods based on linear algebra. Such methods should first of all answer the question whether the time resolved data set contains enough information to delineate any assembly intermediates.

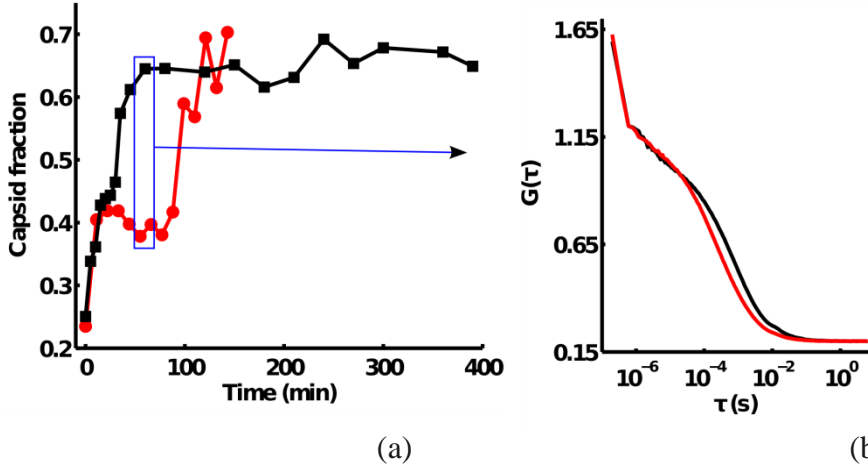(a)   (b)

Figure 4: Calculated changes of capsid fraction over the time. (a) Two assembly time series with different initial concentration of CP2 and TR-RNA computed from RILT-derived distributions. (b) Right panel shows ACFs from the middle part of the assembly time series (in blue box) as an illustration of the significant difference in the diffusion time. System with higher initial concentration of CP2 TR-RNA (red curve), on average 1.5 molecules in confocal volume; second shown dataset (black curve) had 0.9 molecules per confocal volume. Their time-traces split about the 20th minute from the assembly start, where the sample with higher initial concentration falls into the kinetics trap.

Two-state system model

To prove existence of detectable intermediates in the capsid assembly, we tested simple two-state system model. In the absence of intermediates, system can be described as a superposition of the starting and the final state, i.e. initial assembly mixture and assembled capsids (see Eq. 6 and 7). From the difference between fitted model and raw experimental data (Figure 5) compared with noise level, we evaluated, whether the two-state model is sufficient to explain the system progression (see Figure 6).

As seen in Figure 6 two state model did not satisfactorily explain the data, thus other components, representing species with intermediate particle sizes should be added. The total number of detectable species is equal to the effective rank of the data matrix I which can be determined by SVD [16].
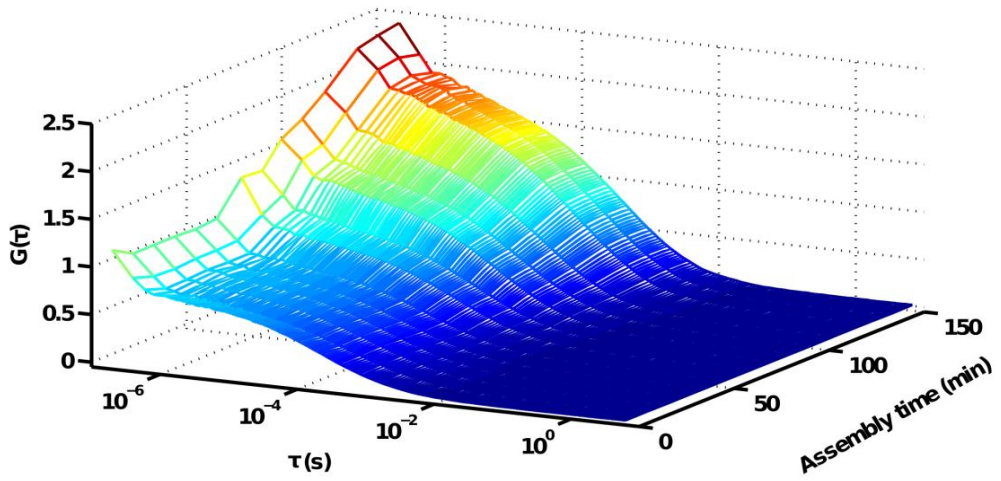
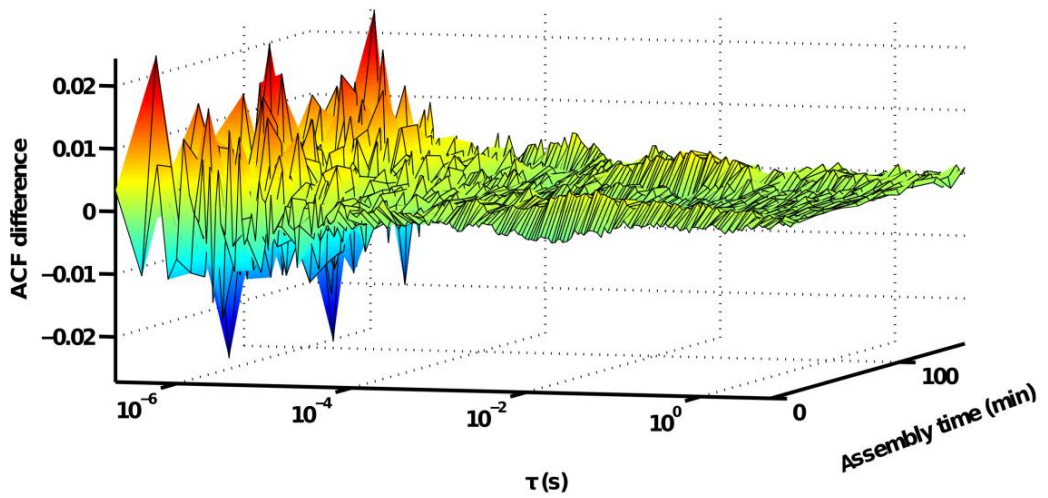*Figure 5: Representative set of experimentally collected time resolved ACFs during assembly.*

Figure 6: Difference of measured and least-square two-state fitted data. The main discrepancy is visible in the triplet region; nevertheless differences higher than noise are located also in the diffusion region 0.1-1 ms.

Singular value decomposition

In order to examine whether the FCS data contain contributions from intermediates and to estimate their concentration loadings we used singular value decomposition analysis. In our experiments it was sufficient to include three to four components to describe the set of time resolved autocorrelation functions (Figure 7).
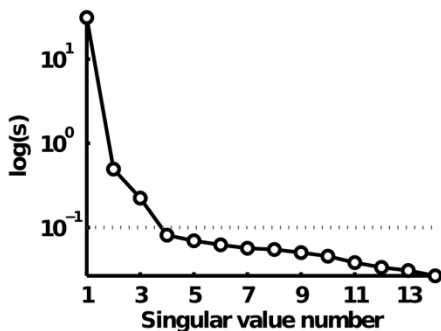

Figure 7: Plot of singular values. The dotted line represents the experimental noise level estimated from repeated collections of the purified fully assembled capsid. Only singular values higher than noise were considered in further analysis.

From SVD vectors to physically meaningful ACFs and concentration loadings

SVD yields the linearly independent components which in their original orthogonal form are devoid of physical meaning. For example, their values alternate between positive and negative and do not directly represent physically meaningful ACF curves or concentration loadings (Figure 8).
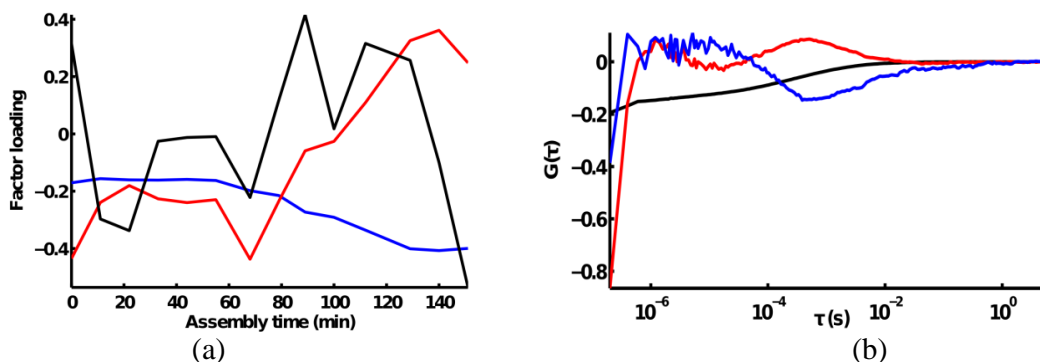

Figure 8: Raw SVD vectors with three factors included for the dataset with higher initial concentration. (A) Concentration (factor) loadings. (B) ACF factors.

However, the SVD scattering vectors represent a convenient orthogonal basis for obtaining a plausible and physically meaningful set of ACF curves as their linear combination. Constrained least squares method and 3D rotation were used to obtain ACF factors and concentration loadings which were non-negative to within the noise level (Figure 9). Rotation is a form of linear combination, frequently used with factor analysis and related statistical methods. Obtained results were in good agreement with output of RILT analysis (compare Figure 4(a) and 9(c)).
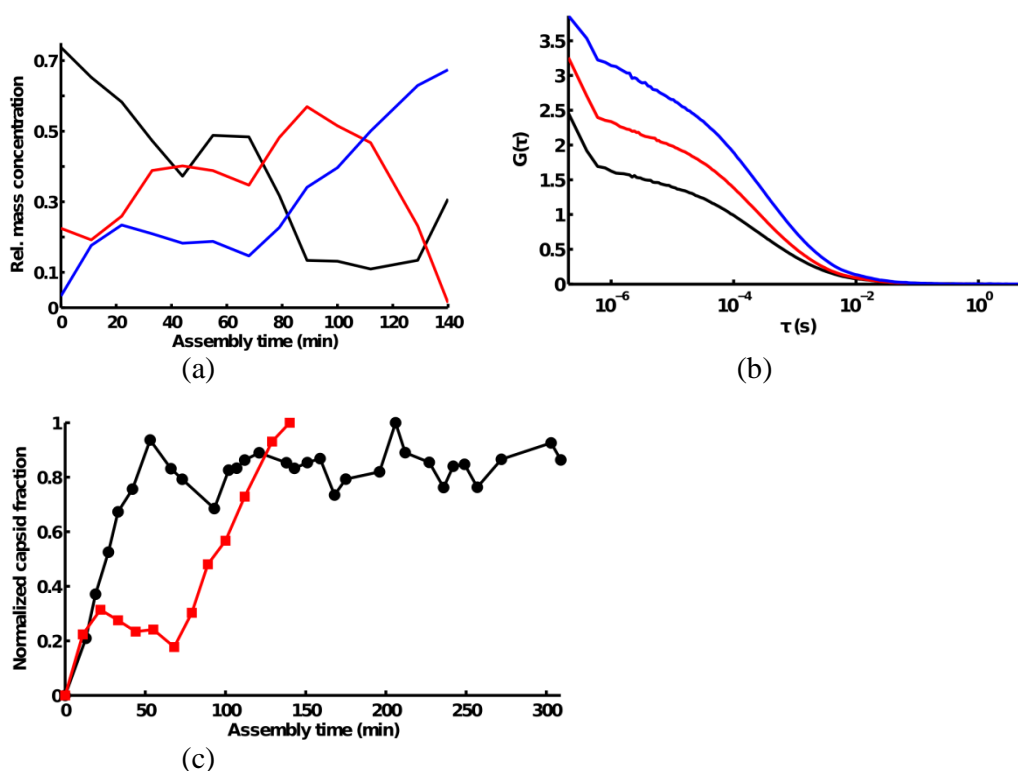

Figure 9: *SVD analysis of dataset with different initial concentration of assembly mixture. SVD was calculated with three factors, first representing CP2 (black), second for intermediates (blue) and third for fully assembled capsids (red). SVD output was rotated, until non-negative, physically meaningful result was found. (a) Loading of rotated factors of more concentrated sample. (b) Rotated ACFs corresponding to (a). (c) Time-resolved formation of capsids. The same datasets as in previous RILT analysis were tested. Sample with higher concentration was again kinetically trapped (red line). The similarity in trend and shape of curves is obvious (compare with Figure 4(a)).*

SVD computed and rotated ACFs can be further analysed with RILT procedure to provide diffusion time distributions (Figure 10). RILT procedure of linearly independent components (ACFs) calculated by SVD shows three distinct distributions. First and third distribution, corresponding to CP2 (blue) and capsids (green), respectively, has two separated peaks (Figure 10). CP2 distribution is shifted to shorter diffusion times and second peak of capsid component is broader. Second distribution,

corresponding to intermediates shows three clearly separated peaks of which the middle likely represents the average size of the intermediates.
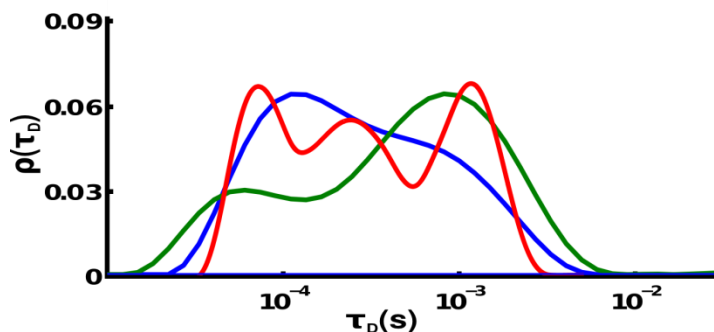

Figure 10: Probability distributions of diffusion times calculated from rotated SVD generated factors. The second factor (red line) is assigned to intermediates, first factor (blue line) to initial mixture and third (green line) to fully assembled capsids. Compared with distribution derived directly from measured ACFs, new peak at 240 μs appeared in the probability distribution derived from the second factor generated by rotated SVD.

Conclusions

FCS has been successfully applied in investigation of virus capsid assembly. FCS was sensitive enough to observe a temporary kinetic trap. Use of advanced data processing is necessary to reveal the time dependent changes of stoichiometry of individual assembly components.

The distribution calculated with the RILT method was stable and the average size of the particles calculated from the effective diffusion time was in good agreement with the data determined using the discrete-component fit for the test cases. Main advantage of RILT procedure lies in fact that it gives diffusion time distribution of each measured assembly step or component. Nevertheless RILT does not necessarily directly reveal the true concentration loading of the intermediates.

SVD, on the other hand, quickly reveals concentration loading of each detected factor. The SVD concentration factor loading agrees with assembly time traces obtained from RILT procedure.

Combination of both methods seems to be most promising, it can benefit from the capability of SVD to follow the loading of individual assembly factors and the probability distribution of SVD derived ACF's diffusion times. Both procedures utilizing FCS experimental data provide comparable results in accord with results obtained by EM or MS.

Acknowledgements

References

1. Stockley, P.G., et al., *A simple, RNA-mediated allosteric switch controls the pathway to formation of a T=3 viral capsid.* Journal of Molecular Biology, 2007. **369**(2): p. 541-552.
2. Rolfsson, O., et al., *RNA packing specificity and folding during assembly of the bacteriophage MS2.* Computational and Mathematical Methods in Medicine, 2008. **9**(3-4): p. 339-349.
3. Zlotnick, A., *Theoretical aspects of virus capsid assembly.* Journal of Molecular Recognition, 2005. **18**(6): p. 479-490.
4. Ni, C.Z., et al., *Crystal-Structure of the Ms2 Coat Protein Dimer - Implications for Rna-Binding and Virus Assembly.* Structure, 1995. **3**(3): p. 255-263.
5. Lago, H., et al., *Probing the kinetics of formation of the bacteriophage MS2 translational operator complex: Identification of a protein conformer unable to bind RNA.* Journal of Molecular Biology, 2001. **305**(5): p. 1131-1144.
6. Tito, M.A., et al., *Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid.* Journal of the American Chemical Society, 2000. **122**(14): p. 3550-3551.
7. Morton, V.L., et al., *Insights into Virus Capsid Assembly from Non-Covalent Mass Spectrometry.* Mass Spectrometry Reviews, 2008. **27**(6): p. 575-595.
8. Beckett, D. and O.C. Uhlenbeck, *Ribonucleoprotein Complexes of R17 Coat Protein and a Translational Operator Analog.* Journal of Molecular Biology, 1988. **204**(4): p. 927-938.
9. Kainov, D.E., et al., *Conserved intermediates on the assembly pathway of double-stranded RNA bacteriophages.* Journal of Molecular Biology, 2003. **328**(4): p. 791-804.
10. Epand, R.M. and P. Russo, *Changes in the morphology of influenza virus as measured by dynamic light scattering.* Biophysical Journal, 2001. **80**(1): p. 418a-418a.
11. Elson, E.L. and D. Magde, *Fluorescence Correlation Spectroscopy .1. Conceptual Basis and Theory.* Biopolymers, 1974. **13**(1): p. 1-27.
12. Schwille, P., *Fluorescence correlation spectroscopy and its potential for intracellular applications.* Cell Biochemistry and Biophysics, 2001. **34**(3): p. 383-408.
13. Modos, K., et al., *Maximum-entropy decomposition of fluorescence correlation spectroscopy data: application to liposome-human serum albumin association.* European Biophysics Journal with Biophysics Letters, 2004. **33**(1): p. 59-67.
14. Provencher, S.W., *Contin - a General-Purpose Constrained Regularization Program for Inverting Noisy Linear Algebraic and Integral-Equations.* Computer Physics Communications, 1982. **27**(3): p. 229-242.
15. Livesey, A.K., P. Licinio, and M. Delaye, *Maximum-Entropy Analysis of*

Quasi-Elastic Light-Scattering from Colloidal Dispersions. Journal of Chemical Physics, 1986. **84**(9): p. 5102-5107.

16.   Golub, G.H., *Matrix computation and the theory of moments.* Proceedings of the International Congress of Mathematicians, Vols 1 and 2, 1995: p. 1440-1448 1605.

17.   Tuma, R., et al., *Detection of intermediates and kinetic control during assembly of bacteriophage P22 procapsid.* Journal of Molecular Biology, 2008. **381**(5): p. 1395-1406.

2.5. Towards characterization of DNA structure under physiological conditions *in vivo* at the single-molecule level using single-pair FRET.

Fessl, T., Adamec, F., Polivka, T., Foldynova Trantirkova, S., Vacha, F., Trantirek, L., (2012) "Towards characterization of DNA structure under physiological conditions *in vivo* at the single-molecule level using single-pair FRET." <u>Nucleic Acids Research</u>10.1093/nar/gks333.

Towards characterization of DNA structure under physiological conditions *in vivo* at the single-molecule level using single-pair FRET

Tomáš Fessl[1,2], František Adamec[1,2], Tomáš Polívka[1,2], Silvie Foldynová-Trantírková[2], František Vácha[1,2,*] and Lukáš Trantírek[3,*]

[1]*Faculty of Sciences, University of South Bohemia,* [2]*Biology Centre ASCR, v.v.i., Branišovská 31, 370 05 České Budějovice, Czech Republic and* [3]*Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

*To whom correspondence should be addressed. Tel: +31 00420387772244; Fax: +31 00420385310366; Email: vacha@jcu.cz. Correspondence may also be addressed to Lukáš Trantírek. Tel: +31 0031302534111; Fax: +31 0031302540980; Email: l.trantirek@uu.nl

Abstract

Fluorescence resonance energy transfer (FRET) under *in vivo* conditions is a well-established technique for the evaluation of populations of protein bound/unbound nucleic acid (NA) molecules or NA hybridization kinetics. However, *in vivo* FRET has not been applied to *in vivo* quantitative conformational analysis of NA thus far. Here we explored parameters critical for characterization of NA structure using single-pair (sp)FRET in the complex cellular environment of a living *Escherichia coli* cell. Our measurements showed that the fluorophore properties in the cellular environment differed from those acquired under *in vitro* conditions. The precision for the interprobe distance determination from FRET efficiency values acquired *in vivo* was found lower (~31%) compared to that acquired in diluted buffers (13%). Our numerical simulations suggest that despite its low precision, the in-cell FRET measurements can be successfully applied to discriminate among various structural models. The main advantage of the in-cell spFRET setup presented here over other established techniques allowing conformational analysis *in vivo* is that it allows investigation of NA structure in various cell types and in a native cellular environment, which is not disturbed by either introduced bulk NA or by the use of chemical transfectants.

Introduction

Inside a living cell, nucleic acids (NA) are exposed to a very complex environment and their structures are, depending on their sequence, modulated by non-specific factors such as viscosity, molecular crowding or by specific interactions with ions and small molecular weight compounds (1). While the complexity of the intracellular environment as a general property of every living organism is generally appreciated, there is a lack of appropriate tools to analyze NA structures in a cellular context.

The first attempt to characterize nucleic acid structures at high resolution under native conditions was recently performed by Hansel *et al.* using nuclear magnetic resonance (NMR) spectroscopy inside living *Xenopus laevis* oocytes (2). Recently,

108

Krstic *et al.* extended the technique of the in-cell NMR to pulsed electron–electron double-resonance (PELDOR) spectroscopy (3). In both techniques, exogenous DNA/RNA constructs were mechanically introduced into the *X. laevis* oocytes/eggs, followed by NMR or PELDOR investigations. Although both methods provide unique information on NA structure 'in living cells', this phrase does not necessarily mean 'in a native cellular environment'. Due to the low inherent sensitivity of NMR and PELDOR detection and high cellular background, the methods require unnaturally

large quantities ($\sim$250–500 µM) of exogenous NA to be deposited into the cells (2,3). As illustrated in Figure 1, the number of NA base-pairs deposited into the cell for the in-cell NMR/PELDOR experiments was more than 15 000 times higher than the number of base-pairs in the genomic DNA (2–4). Although the injected cells for in-cell NMR/PELDOR experiments are viable, the addition of such a high concentration of NA might modulate the intracellular environment and bias the structural readout. In practical terms, applications of in-cell NMR and PELDOR spectroscopy are presently limited to *X. laevis* oocytes and eggs, which are large enough to allow for the introduction of high concentrations of exogenous NA via injection (2–4).
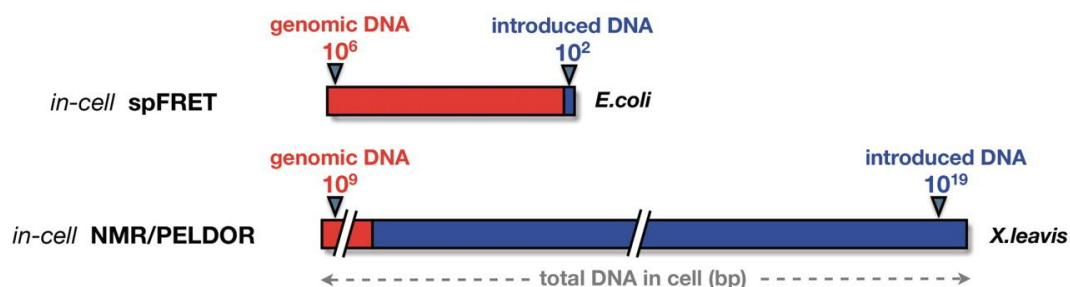

Figure 1.
 Schematic representation of requirements of in-cell NMR/PELDOR and in-cell spFRET on concentration of the exogenous NA introduced into cellular environment. The concentrations are displayed relative to average size of genomic DNA of E. coli *and* X. laevis.

Undoubtedly, the development of new techniques that would allow characterization of NA structure under physiological conditions *in vivo* is of general interest. Ideally, such techniques would avoid the deposition of large quantities of NA in the cellular environment, allow measurements in a variety of cell types and provide complementary information to the existing techniques of in-cell NMR and PELDOR.

One of the most obvious candidates among the available methods that in principle allow quantitative characterization of NA structure under native conditions is fluorescence resonance energy transfer (FRET) (5–7). FRET has been successfully used to obtain quantitative long-range information on NA structure under *in vitro* conditions (8–11). A number of studies also demonstrated that FRET measurements could be performed inside living cells (12–14). Although these state-of-the-art *in vivo* FRET applications are currently limited to the evaluation of populations of protein bound/unbound NA molecules or NA–NA hybridization kinetics and usually require the introduction of a bulk amount of fluorescently labeled NA into cells with the use of chemical transfectants, they clearly reveal the potential for quantitative characterization of nucleic acid structure using *in vivo* FRET (12–16). Rapid advances

in live-cell imaging, demonstrated e.g. by direct observation of the kinetics of transcription of single nascent mRNA molecules (17), have now enabled acquiring quantitative FRET data in living cells at single molecule level.

Here, we explored the possibility of quantitative characterization of nucleic acid structure inside living cells at the single-molecule level using single-pair FRET (spFRET). The proposed approach avoids the need to introduce bulk amounts of DNA into the cells and the use of chemical transfectants, which both disturb composition of the intracellular environment. The parameters critical for characterization of NA structure using spFRET were evaluated for the ATTO-680/ATTO-740 donor–acceptor system based on the experimentally determined variation of FRET efficiencies for a series of terminally labeled DNA duplexes with a length of 8–18 bp, both *in vitro* and *in vivo*, in *Escherichia coli* cells. The application potential and limitations of in-cell spFRET for structural analysis of NA are discussed.

Materials and methods
DNA constructs
Unmodified DNA oligonucleotides were purchased from Generi-Biotech (Hradec Kralove, Czech Republic). Fluorescently labeled DNA oligonucleotides coupled with fluorescent dyes ATTO680 and ATTO740 (ATTO-TEC GmbH) via flexible C_6 linkers at the 5′-termini were purchased from Sigma-Aldrich. Table 1 lists the DNA constructs used in this study. For both *in vitro* and *in vivo* spFRET measurements, both single-stranded DNA conjugated either with ATTO 680 or ATTO 740 was dissolved in *buffer A* (0.01% Tween 20, 200 µm sodium L-ascorbate, 0.3 mM EDTA, 10 mM TRIS, pH 7.4). Subsequently, the ATTO680–DNA was mixed with ATTO740–DNA of corresponding size at a ratio of 1:1.15. To allow annealing of complementary strands, the mixture was heated to 95°C for 5 min and left to slowly cool to room temperature.

Table 1.
DNA constructs employed for spFRET measurements

DNA length [bp]	ATTO680-5′ -> 3′	ATTO740-5′ -> 3
8	CCTGCAGG	CCTGCAGG
10	CCTGCAGTGG	CCACTGCAGG
12	CCTGCAGTACGG	CCGTACTGCAGG
14	CCTGCACGACCTGG	CCAGGTCGTGCAGG
16	CCTGCACGACCTGTGG	CCACAGGTCGTGCAGG

Absorption and fluorescence spectroscopy
Absorption and emission spectra were acquired at room temperature in a 10-mm silica glass cuvette in *buffer A* at DNA concentrations of 0.1 and 10 µM for measurements of fluorescence and absorption, respectively. Emission spectra were recorded on a Fluorolog spectrofluorometer (*SPEX*, USA). A Raman emitter RS 664 LP (679–1497 nm) was used to filter emission spectra from the excitation light. The absorption spectra were recorded on a UV300 spectrophotometer (Spectronic Unicam, UK).

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were acquired on a JASCO J-715 spectropolarimeter (JASCO Corporation, Tokyo, Japan) in 10-mm silica glass cuvettes at room temperature. The CD spectra were acquired in *buffer A* at a DNA concentration of 10 µM. The melting temperatures of the DNA constructs were determined by the temperature dependence of the molar ellipticity in the global minimum of the CD spectra. A thermostatic water bath combined with a thermoelectric Peltier device was used to control the temperature. The individual points of the melting curve were collected after 15 min of temperature equilibration at each temperature step.

Time-resolved fluorescence anisotropy

Time-resolved fluorescence anisotropy decay measurements were performed to evaluate the rotational mobility of the donor and acceptor fluorophores. The anisotropy decay *r(t)* was constructed from polarized intensity decays as

$$r(t) = \frac{I_{\parallel}(t) - GI_{\perp}(t)}{I_{\parallel}(t) + 2GI_{\perp}(t)},$$

(Eq.1)

where $I_{\parallel}$ and $I_{\perp}$ are the intensity decays recorded with the emission polarizer oriented parallel or perpendicular, respectively, relative to the vertically polarized excitations and *G* is the detector correction factor.

Fluorescence anisotropy measurements were performed in *buffer A* at a DNA concentration of 0.1 µM. The fluorescence anisotropy measurements were performed with a modified instrumental setup as for lifetime measurements (*vide infra*) with the addition of a Glan–Taylor prism to separate the two mutually perpendicular polarization channels and an additional avalanche photo diode (SPCM-AQR-16) used to detect the second polarization channel.

The instrument response function was deconvoluted from the lifetime decays of the fluorophores, and the decays were fitted using a Fluofit module (18) of Matlab (Mathworks, USA). In most cases, one short decay component, not related to donor excited-state lifetime, was used to eliminate the direct acceptor excitation and scattered excitation light in the cellular environment. The collected lifetime probability distribution functions were analyzed through the maximum likelihood estimation with the Matlab function '*gmdistribution.fit*', with the multi-component distribution analysis providing the means, covariance and logarithmic likelihood of the fit.

Time-resolved fluorescence anisotropy decays were fitted to a model consisting of the sum of two exponentials, one accounting for the fast local motion of the fluorophore and one accounting for the slow global tumbling of the entire molecule:

$$r(t) = \beta_F \exp\left(-\frac{t}{-\alpha_F}\right) + \beta_S \exp\left(-\frac{t}{-\alpha_S}\right)$$

(Eq.2)

The parameters β_F, β_S, α_F and α_S characterize the amplitudes and rotational correlation times of fast (F) and slow (S) motion, respectively.

Time-resolved *in vitro* FRET measurements and Förster radius determination

The *in vitro* FRET efficiencies were measured using the time-resolved method for the ensemble of molecules (19). Acquired FRET efficiencies were translated into interprobe distances using Supplementary Equation S1. The Förster radius in Supplementary Equation S1 was calculated from the experimentally acquired medium refractive index, the extinction coefficient of the acceptor, the fluorescence quantum yield of the donor, the orientation factor κ^2 and the overlap integral of FRET pair spectra in *buffer A* (Supplementary Table S3). Quantum yield of ATTO680 was determined relative to Rhodamine B according to the procedure of Williams *et al.* (20). The emission spectra used to determine the quantum yield were corrected for detection efficiency in different spectral regions. The extinction coefficient of the acceptor in *buffer A/lysate* was estimated from absorbancies acquired for various dilutions of the acceptor in *buffer A/lysate* relative to the absorbance of the ATTO740 dissolved in water using Lambert–Beer equation (all the measurements were conducted in the linear region of the detector) (Supplementary Table S3). Extinction coefficient of the ATTO740 in water was provided by manufacturer (ATTO-TEC, Germany).

The orientation factor, κ^2, was estimated from the steady-state fluorescence anisotropies following the procedure of Dale *et al.* (21) (Supplementary Figure S1) using a single nano-positioning system according to the approach described by Muschielok *et al*. (22). Estimation of κ^2 from the steady-state anisotropies was preferred to that one based on time-resolved anisotropy although resulting estimate is usually less precise (23). The reasons for this choice were anomalous characteristics of time-resolved anisotropies in crude bacterial lysate lacking exponential decays (Supplementary Figure S3).

FRET data processing

The most likely donor and acceptor positions and their uncertainties in all the terminally labeled DNA constructs were calculated using the Nano Positioning System (NPS) incorporated in FRETnps Tools (22). The probability distribution function (PDF) of acceptor position was derived from FRET measured efficiencies and from the accessible volume (AV) of donor dye simulated according the procedure of Sindbert *et al.* (23). In agreement with *in vitro* CD data, the canonical B-DNA model with defined attachment points for acceptor and donor along with the knowledge of the length (20 Å) and width (4.5 Å) of the flexible C_6 linker was used in the simulation of the AV for both dyes (23). As the manufacturer has not yet published the structures of the ATTO680 and ATTO740 dyes, each dye was approximated by sphere with a diameter of 5 Å in the AV simulations.

Incorporation of DNA into *E. coli* cells

The BL21 StarTM(DE3)plysS One Shot (Invitrogen, US) strain of *E. coli* was used for all *in vivo* FRET measurements. The DNA constructs were introduced into *E. coli* cells using heat shock with an optimized protocol allowing the introduction of one to five DNA molecules per cell (for details see Supplementary Data). After DNA incorporation, the cells were washed three times in PBS buffer (137 mM NaCl, 2.7 mM

KCl, 10 mM $Na_2 HPO_4$, 1.8 mM KH_2PO_4, pH 7.4) to remove the extracellular, unincorporated DNA constructs. Subsequently, the transfected cells were immobilized in a thin layer of polyvinyl alcohol (MW 145 000, 98% hydrolyzed, Merck) using the spin-coating method (24).

The information on the number of molecules per cell was obtained via visual counting. This approach was made possible by exploiting the element that extends the depth of focus of the microscope objective (EDOF), as described in Fessl *et al.* (25). When using this element, the setup can also detect molecules that are out of focus using the plane microscope objective. Such defocused molecules increase the fluorescence background and might contaminate measured fluorescence lifetime decays. The EDOF element was therefore used to optimize the procedure for the incorporation of single DNA molecules into the cells.

Single-pair time-resolved in-cell FRET measurements

Escherichia coli cells with incorporated DNA constructs were immobilized in a thin layer of PVA. Intracellular fluorophores were localized and individual molecules were spatially separated by slits in an imaging spectrograph. The lifetimes of single constructs were examined. Lifetime histograms were created from approximately 60 molecules for each oligonucleotide length.

The values of E_{FRET} were measured using the time-resolved method as

$$E_{FRET} = 1 - \frac{\tau'_D}{\tau_D},$$

(Eq.3)

where τ'_D and τ_D are the donor fluorescence lifetimes in the presence and absence of an acceptor, respectively.

The single-pair time-resolved FRET measurements were performed using a setup composed of an inverted Olympus IX70 microscope, Triax 320 imaging spectrograph with back-illuminated liquid nitrogen-cooled CCD camera (Spectrum One, Jobin Yvon, 2048 × 512 pixels, pixel size 13.5 × 13.5 μm) and a picosecond-pulse laser diode module (PicoQuant LDH-D-C-640) that generated pulses at 640 nm with linear polarization as an excitation source. The microscope was equipped with an objective providing 100× magnification (Olympus, NA of 1.35, UplanApo), an Olympus filter cube (Olympus, Japan) containing a Raman emitter RS 664 LP (679.3–1497.7 nm), and a donor bandpass BrightLine HC 720/13 and HC-Laser Clean-up MaxDiode 640/8 (Semrock, Germany). For measurements of the fluorescence lifetimes, a time-correlated single photon counting avalanche photodiode (Perkin Elmer, SPCM-AQR-16) was attached to the side exit of the Triax monochromator (Jobin Yvon Inc., USA). The samples were excited at 640 nm by a total internal reflection prism (25). Acquired in-cell FRET efficiencies were evaluated in terms of the most likely donor and acceptor positions and their uncertainties using the same procedure as described for *in vitro* data. However, in contrast to *in vitro* FRET evaluation, the Förster radius constructed from the medium refractive index, the extinction coefficient of the acceptor, the fluorescence quantum yield of the donor, the orientation factor κ^2 and the overlap integral of FRET pair spectra all acquired in crude bacterial lysate was employed for E_{FRET}-distance conversion.

Preparation of the bacterial lysate

Bacteria were cultivated following the same procedure as for single molecule experiments in a total volume of 1 l. To wash the cells from the cultivation medium, the cell suspension was pelleted via centrifugation at 4°C for 10 min at $4500g$. The supernatant was removed and the cell pellet was resuspended in 200 ml of ice-cold PBS buffer. This procedure was repeated three times with fresh PBS. Following the last step of centrifugation, the PBS buffer was removed and a semi-dry bacterial pellet was obtained. To obtain the bacterial lysate, the bacterial pellet was sonicated on ice. Cell debris from the crude cell homogenate was removed by centrifugation at 4°C for 1 h at $23\,500g$. The supernatant, herein referred as bacterial lysate, was used for FRET measurements.

Results

ATTO680 and ATTO740 do not affect helical conformation of DNA

In general, the principal disadvantage of the FRET technique is the required presence of donor and acceptor fluorophore tags on a probed NA fragment. In principle, these tags might interact with NA fragments and influence both the structure of the NA fragment under study and affect the mobility of fluorophores, which biases the interpretation of E_{FRET} via modulation of the orientation factor, κ^2 (Supplementary Equation S5). To evaluate the potential interference of ATTO680 and ATTO740 fluorophores with conformation of the DNA, the influence of the fluorophores on the conformation of the DNA was investigated using CD spectroscopy, thermo-melting analysis and time-resolved fluorescence anisotropy measurements.

The comparison of CD spectra of unmodified DNA and DNA modified with ATTO680, ATTO740, and both ATTO680 and ATTO740 fluorophores suggested that ATTO680 and ATTO740 did not disturb the global helical geometry of the DNA construct (Supplementary Figure S2). For all the constructs, the CD spectra exhibited characteristic features corresponding to the B-form of DNA. Minor differences among the CD spectra, melting temperatures, and distinct kinetics of anisotropy decays between DNA–ATTO680 and DNA–ATTO740 suggested that while ATTO680 has free mobility and moderately promotes terminal base-pair opening, the mobility of ATTO740 is partially hindered due to stabilizing interaction with the DNA (Supplementary Figure S2–S3 and Supplementary Table S1).

Modulation of E_FRET in a DNA duplex series *in vitro* versus *in vivo*

To evaluate both the potential of E_{FRET} measurements for characterization of NA structure within a complex cellular environment and impact of the intracellular *E. coli* environment on helical geometry of our DNA constructs, the variations of E_{FRET} for a series of five DNA duplexes of a length of 8–16 bp terminally labeled with ATTO680/ATTO740 fluorophores in free solution and in living *E. coli* cells were investigated (Supplementary Figure S4). The resulting E_{FRET}, both *in vitro* and *in vivo*, were plotted as a function of the helix length in Figure 2. Overall, *in vitro* E_{FRET} values decreased with helix length, yet there was an apparent local increase in E_{FRET} at 10 bp.
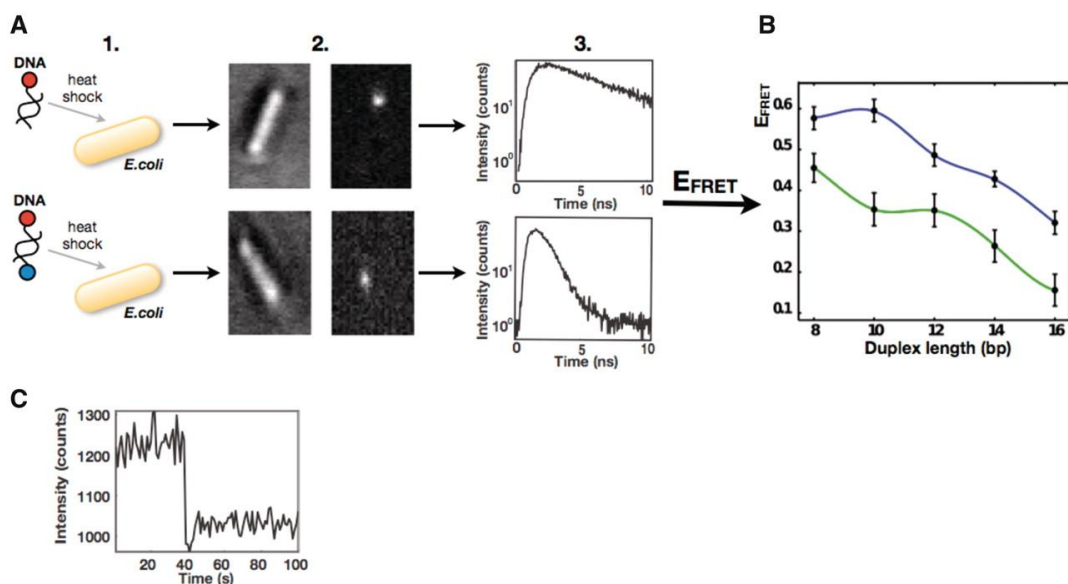
Figure 2.

(A) Simplified workflow of the in-cell spFRET experiment. Step 1—Bacterial cells were separately transformed with DNA constructs labeled with ATTO680 or with the complete FRET pair (ATTO680 and ATTO740). Step 2—Fluorescent molecules inside E. coli *cells were localized. Left sub-figure shows transmission image of a bacterial cell with the introduced DNA. The sub-figure on the right is fluorescent image of the same cell. Step 3—Fluorescence lifetimes from the constructs labeled with the donor (ATTO680) and from the constructs labeled with the complete FRET pair (ATTO680 and ATTO740) were acquired and used for calculation of the FRET efficiencies (E_{FRET}). (B) Modulation of E_{FRET} in a duplex DNA series of variable length inside living bacterial cells (green line) versus corresponding modulation* in vitro *(blue line). The experimental data were fitted with spline. (C) A representative photon-trace from DNA construct labeled with the donor (ATTO680) inside* E. coli *cell showing single step photobleaching. The trace demonstrates that the fluorescence signal comes from single molecule.*

Based on the characteristic pattern of the CD spectrum (Supplementary Figure S2), the observed profile for E_{FRET} dependence on helix length can be associated with helical parameters corresponding to the B-form of DNA.

Analogously to the situation *in vitro*, the *in vivo* E_{FRET} values decreased with helix length. However, in contrast to *in vitro* data, the E_{FRET} values displayed a local maximum in E_{FRET} at 12 bp (Figure 2). Comparison of the phase modulation between the *in vivo* and *in vitro* E_{FRET} profiles suggested differences in either DNA helical geometry and/or effective relative orientations of the donor and acceptor transition dipole moments. In addition to the differences in phase modulation between *in vivo* and *in vitro* E_{FRET} profiles, the *in vivo* E_{FRET} values appeared systematically reduced in amplitude compared to the *in vitro* measurements. The reduced E_{FRET} values indicated altered optical properties of environment and fluorophores inside the cells compared to those in the phosphate buffer used for *in vitro* measurements (Supplementary Equation S5). Indeed, the experimentally determined refractive index of the phosphate buffer, $n_{in\ vitro} = 1.33$, was lower than the refractive index of the bacterial lysate, $n_{in\ vivo} = 1.40$.

However, the difference in refractive indices accounted only for ~40–50% of reduction in measured E_{FRET} values based on the estimation from Supplementary Equation S2. The fact that the experimentally observed attenuation of E_{FRET} was almost twice as much as it was expected to be based on the corrected refractive index suggests that the cellular environment influences optical properties of fluorophores.

Altogether, the differences in E_{FRET} amplitude and phase modulation between *in vitro* and *in vivo* data strongly indicate that for quantitative analysis of an NA structure inside living cells, alterations in both spectral properties of fluorophores and their linker lengths and flexibilities due to the cellular environment need to be considered in the process of the in-cell spFRET data analysis. As for the spectral properties of fluorophores, it is, unfortunately, experimentally impossible to acquire Q_0 and J (Supplementary Equation S5) in the context of an intact living cell due to technical difficulties with recording the absorbance spectra and quantum yield of individual molecules. Therefore, to account for altered spectral properties of fluorophores in a cellular environment, we acquired the extinction coefficient of the acceptor, the fluorescence quantum yield of the donor, the orientation factor and the overlap integral of the FRET pair using a bulk amount of DNA mixed with the crude lysate from bacterial cells. Quantitative agreement between the fluorophore emission spectra and lifetimes recorded both *in vivo* and in the crude bacterial lysate indicated that the fluorophores' spectral properties in the crude bacterial lysate were similar to those in intact living cells (Supplementary Figure S5 and Supplementary Table S4).

The fluorophores' spectral properties and orientational factor from fluorescence anisotropies, both acquired in the bacterial lysate along with experimentally acquired refractive index of the bacterial lysate (see 'Materials and Methods' section), were subsequently used to calculate Förster distance (R_0), whose *a priori* knowledge is essential for interpretation of in-cell E_{FRET} values in terms of interprobe distances (Supplementary Equations S1 and S5). Supplementary Figure S6 shows the R_0 distributions derived from measurements in crude bacterial lysate and diluted buffer. While the distribution of R_0 *in vitro* was characterized by a median distance of $R_0 = 64.2$ Å, and first and third quantiles of 60.1 and 69.5 Å, respectively, the distribution of R_0 in the crude bacterial lysate was marked by a median distance of $R_0 = 47.1$ Å, and first and third quantiles of 39.6 and 53.2 Å, respectively. Significantly broader distribution of R_0 distances in the bacterial lysate is a consequence of higher fluorescence anisotropies of both dyes in this inherently heterogeneous environment that translates into larger variability of the orientational factor in bacterial lysate compared to that in the diluted solution. The large variability of the orientational factor is reflected in lower precision of distance determination from *in vivo* E_{FRET} measurements compared to measurements under *in vitro* conditions. From R_0 distributions (Supplementary Figure S6), the precision for the interprobe distance determination from E_{FRET} values was expected to be ~13.9% and 31.4% for *in vitro* and in-cell data, respectively (Supplementary Equation S4).

Structural interpretation of *in vivo* spFRET data

The double-helical DNA fragments of randomized sequences, analogous to those employed in our study, adopt B-DNA conformation in diluted aqueous solutions of monovalent ions (26,27). However, the same DNAs can undergo a B−A transition that is induced by a number of natively occurring agents, such as multivalent ions (28–30), osmolytes mimicking lower water content inside the cells (30–33) and proteins (34–36).

To assess which conformation the DNA adopts inside living prokaryotic cells, and simultaneously assess resolution power of the in-cell spFRET measurements, the derived R_0 values from the crude bacterial homogenate were used to express experimentally acquired in-cell FRET efficiencies for a series of five DNA duplexes with a length of 8–16 bp in terms of the donor–acceptor PDF. Resulting interprobe distance PDFs for individual DNA constructs are shown in Figure 3. Experimentally acquired in-cell interprobe distance PDFs for individual constructs were then compared to corresponding PDFs that were simulated based on geometric AVs of fluorophores positioned on the canonical A- and B-form DNA models according to a recently proposed technique by Sindbert *et al.* (23). Comparison of the experimental interprobe distance PDFs derived from in-cell spFRET efficiencies along with simulated PDFs for A- and B-form of DNA are displayed in Figure 3. Statistical comparison between the experimental and model PDFs clearly indicated better agreement of in-cell spFRET experimental data with B-DNA compared to the A-DNA model (Table 2).
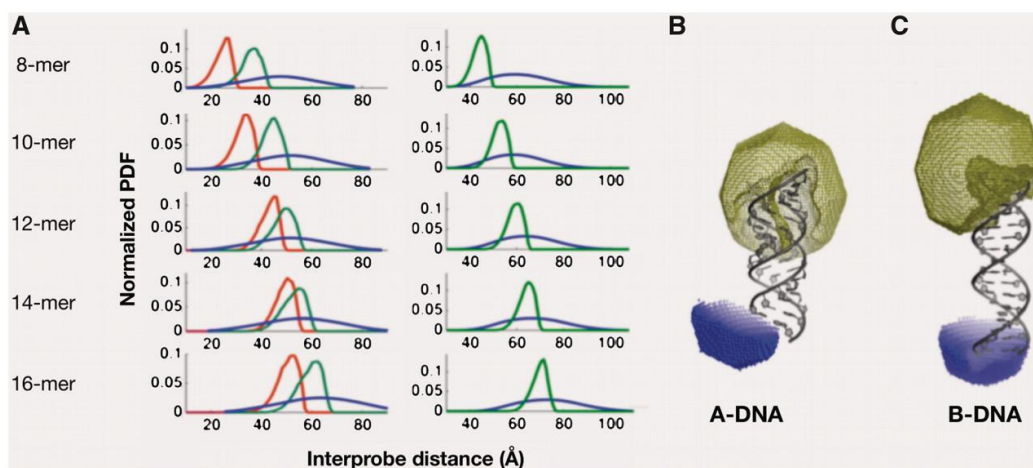

Figure 3.

(A) Left panel: Overlaps between interprobe distance PDFs derived from in-cell FRET measurements (blue line) and 'in-cell' PDFs simulated using FRET NPS tools for A- (red line) and B-form (green line) of DNA. Right panel: Overlaps between PDFs derived from in vitro *FRET measurements (blue line) and* 'in vitro' *PDFs simulated for B-form of DNA. (B) and (C) AVs of ATTO680 (olive mesh) simulated for A-form and B-form DNA, respectively, and FRET derived and constrained positions of ATTO740 (blue spheres). Both simulations accounted for altered fluorophores properties in cellular environment. For details on the simulation approach, see 'Materials and Methods' section.*

Table 2.
Probabilistic interpretation of the in-cell FRET data

bp	A-DNA	B-DNA
8	0.87	0.60
10	0.92	0.81
12	0.80	0.67
14	0.51	0.39
16	0.65	0.43

Quantification of a distance between the 'in-cell' PDFs simulated using FRET NPS tools (22) for A- and B-form of DNA and 'reference' interprobe distance PDFs derived from in-cell FRET measurements using the Kolmogorov–Smirnov (KS) statistic. Identical distributions are marked by 0 value in the KS statistics. Note that B-DNA model has lower KS values compared to A-DNA model for all the DNA constructs indicating the B-form of DNA to be a more likely conformation in intact bacterial cells compared to A-DNA.

Discussion

In the present study, we investigated the variation of E_{FRET} for a series of five DNA duplexes with a length of 8–16 bp terminally labeled with ATTO680 and ATTO740 donor–acceptor pair in a solution and in living *E. coli* cells. Our measurements showed that for the ATTO680/740 donor–acceptor system, the fluorophore properties in the cellular environment differed from those acquired under *in vitro* conditions. While it is experimentally impossible to acquire all spectral properties of the fluorophores for individual molecules under heterogeneous conditions of a living cell, our measurements indicated that these could be approximated by their measurements using a bulk amount of DNA mixed in a crude cellular homogenate.

Comparison between *in vitro* R_0 value distribution and that in crude bacterial lysate indicated that the precision for the interprobe distance determination from E_{FRET} values acquired *in vivo* was significantly lower (~31%) compared to that acquired in diluted buffers (~13%).

The low level of precision of *in vivo*-derived distances limits the *ab initio* structure analysis based on acquired in-cell E_{FRET} data and makes NA structure characterization from in-cell FRET data dependent on probabilistic evaluation. Still, the in-cell E_{FRET} data can be useful to assess the consistency of structural data acquired from E_{FRET} values measured in a complex cellular environment with various structural models. Comparison of interprobe distance PDF derived from experimentally acquired E_{FRET} values in intact bacterial cells and R_0 values corresponding to crude bacterial homogenate with the simulated PDFs corresponding to A- and B-forms of DNA indicated the B-form of DNA to be a more likely conformation in intact bacterial cells compared to A-DNA. This finding is in agreement with recent data on conformation of double-helical DNA oligonucleotides inside living eukaryotic cells derived using pulsed electron–electron double-resonance spectroscopy (3,4).

Considering the inherently low precision of in-cell spFRET measurements, we attempted to assess the application potential of the in-cell spFRET for structural

analysis of NAs. We simulated several typical situations encountered in DNA structure analysis such as B–A and B–Z transitions, or structural deformations in DNA due to oxidative DNA damage, protein–DNA binding or DNA–drug interactions (for details see Supplementary Figure S7). Using nano-positional models for fluorophore locations derived based on structural information from the PDB database and the derived precision of interprobe distance determination under in-cell-like conditions, we constructed the model interprobe distance PDFs. Comparison of the model interprobe distance PDFs for B- and Z-DNA, as well as for B-DNA and DNA containing α-anomeric adenosine, which are representative of the damaged right-handed double-helical DNA marked by a minor groove bend, revealed that the structural changes corresponding to the B–Z transition or the structural DNA damage couldn't be readily resolved using in-cell spFRET measurements (Supplementary Figure S7 and Supplementary Table S5). However, comparison between the model PDFs for DNA bound to protein, namely *E. coli* endonuclease IV, or to base-binding drugs, namely 2-amino-1-methyl-6-phenylimidazo[4,5-2 B]pyridine and cisplatin, and PDF for a naked DNA revealed the same or larger differences than observed for analyzed B–A transition (Supplementary Table S5), which is suggestive that in-cell spFRET could find its applications in structural studies involving DNA–drug and/or protein binding. Nevertheless, the low resolving power of in-cell E_{FRET} implies that for unbiased *in vivo* structural analysis of NAs, the in-cell FRET should be combined with other *in vivo* high-resolution techniques such as in-cell NMR or PELDOR (2–4).

Table 3 compares limitations and application potential of the in-cell spFRET with other established techniques for *in vivo* structural characterization of DNA, namely in-cell NMR and in-cell PELDOR. While both in-cell NMR and PELDOR have higher resolution and precision, respectively, compared to in-cell spFRET, their main drawback lies in the requirement of ultra-high concentrations of exogenous NA that needs to be deposited into the living cells (Figure 1). At the moment, the requirement of ultra-high concentration of exogenous NA for in-cell NMR and PELDOR represents the main source of controversy in the field of cellular structural biology of NAs, as the introduced bulk amount of NA might alter the environmental conditions in the cell and, therefore, affect the structure of the studied molecule. One of the known artifacts of the ultra-high concentration of introduced DNA is the dimerization of double-stranded DNA fragments introduced in *X. laevis* cells (3). This artifact would be naturally avoided in the in-cell spFRET setup.

Another consequence of the required ultra-high concentrations of exogenous NA for in-cell NMR and PELDOR is that these techniques are currently limited to *X. laevis* oocytes and eggs, which are large cells that can be mechanically injected with bulk amount of DNA. In contrast to in-cell NMR and PELDOR, the in-cell spFRET measurements can in principle be performed in both prokaryotic as well as eukaryotic cells. To demonstrate that the in-cell spFRET setup is compatible not only with *E. coli* cells but also with mammalian cells, we performed in-cell spFRET measurements in human epitheloid carcinoma (HeLa) cells (Supplementary Figure S9). One of the foreseen applications of the in-cell FRET thus could be its use for studies aimed at evaluating the influence of cell-type specific environments on DNA structure. In this respect, however, it needs to be emphasized that not all cell types are amenable to in-cell spFRET measurement. The cell-type choice for in-cell spFRET is generally

limited by three factors: the cell's autofluorescence, light scattering, and size. For example, large cells, such as the *X. laevis* oocytes/eggs used for in-cell NMR/PELDOR, cannot readily be used for in-cell spFRET measurements because they have high autofluorescence and their size allows focusing only on single molecules close to the cell surface.

Table 3.
A comparison of in-cell spFRET with other techniques for the in vivo *structural characterization of DNA*

	In-cell NMR	**In-cell PELDOR**	**In-cell spFRET**
Disturbance of native environment	Yes	Yes	No
Cell type	*X. laevis* egg/oocyte	*X. laevis* egg/oocyte	*E. coli,* mammalian cells[a]
Toxicity	Sequence dependent[b]	Sequence dependent[b]	No
Subcellular localization	Nucleus/cytosol[c]	Nucleus/cytosol[c]	Nucleus[d]
Tag requirement	No	Yes	Yes
Measurement time span	Hours	< 70 min[e]	Hours
Structural information[f]	Short-range	Long-range	Long-range

[a]*Illustrated here for human epitheloid carcinoma (HeLa) cells (Supplementary Figure S9).*
[b]*See the text for details.*
[c]*For both in-cell NMR and PELDOR, the issue of the intracellular localization of introduced DNA has not been properly addressed in the literature. In X. laevis, localization experiments using fluorescently-labeled DNA mini-haiprin suggest that ∼90% of introduced DNA is localized in the nucleus and ∼10% in the cytosol (41).*
[d]*Applies for mammalian cells (37–40).*
[e]*Ref (3,4,42).*
[f]*Short-range—typically < 5 Å; long range—typically < 50 Å.*

One of the important advantages of the in-cell spFRET (in eukaryotic cells) over both in-cell NMR/PELDOR is physiologically relevant localization of the introduced DNA. Under native conditions, i.e. inside living cell, the DNA is localized in the nucleus. Although low concentrations of exogenous DNA introduced into cell's cytoplasm tend to spontaneously localize and concentrate in the nucleus (37–40), the nuclei of *X. laevis* oocytes seems unable to accommodate the large quantities of introduced DNA that are required for both in-cell NMR and PELDOR measurements (41). Consequently, the structural information acquired from in-cell NMR/PELDOR is a superposition of information from both the physiologically relevant nucleus and the non-physiologically relevant cytosol (for details see Table 3).

Abstracting from the requirement that an ultra-high concentration of DNA be deposited into the interior of the cells and from the restriction to the *X. laevis* system, the in-cell PELDOR represents an analogy to the proposed in-cell FRET technique. Both in-cell spFRET and PELDOR provide long-range structural information. Both

techniques also require the attachment of reporter tags, namely, fluorophores for FRET and radicals for PELDOR. However, compared with in-cell PELDOR, the in-cell spFRET setup offers one particularly important advantage, namely extended measurement time. Whereas the in-cell spFRET measurement time is primarily limited by the life span of the cell (typically several hours for the ATTO680/ATTO740 donor–acceptor system used in this study), the in-cell PELDOR measurement time is typically limited to ~70 min due to the rapid reduction of the spin labels by the cellular environment (3,4,42).

Last but not least, the introduction of exogenous NA into cells at the single-molecule level is non-toxic to the cells. While the introduction of single molecules into the cell does not impose any toxicity to the cell regardless of the nature of the nucleic acid fragment, it was observed by our group (unpublished data) and others (43) that the introduction of large quantities of specific DNA motifs might be toxic to the cells. Although reduction of the amount of DNA can diminish its toxicity, the DNA concentrations required to avoid toxicity are far below the detection limits of both in-cell NMR and in-cell PELDOR. In this respect, the in-cell spFRET might be regarded as the only alternative to study DNA molecules that display toxicity at high concentrations.

Taken together, the inherent lack of precision appears to be the main disadvantage of in-cell spFRET. Future development of the in-cell spFRET technique thus should include surveying the effect of the cellular environment on other donor–acceptor systems with particular attention paid to anchoring fluorophores to the NAs to increase precision of interprobe distance determination based on in-cell spFRET data.

Nonetheless, the fact that the FRET measurements *in vivo* can be achieved on single molecules makes the in-cell spFRET the only technique allowing *in vivo* quantitative characterization of NA structure without significant disturbance of the native cellular environment. Performing quantitative spFRET measurements compared to bulk in-cell FRET measurements avoids the need to introduce bulk amounts of DNA into the cells and the use of chemical transfectants, which both disturb composition of the intracellular environment. Given the unique features of the in-cell spFRET setup, this technique has the potential to become an important tool to investigate the role of intracellular environment composition variations on the structure of NA through the cell-cycle progression.

Supplementary data

Supplementary Data are available at NAR Online: Supplementary Tables 1–5, Supplementary Figures 1–9, Supplementary Theory and Methods and Supplementary References [44–48].

Funding

Acknowledgments

References

1. Ito Y, Selenko P. *Cellular structural biology. Curr. Opin. Struct. Biol. 2010;20:640-648.*
2. Hansel R, Foldynova-Trantirkova S, Lohr F, Buck J, Bongartz E, Bamberg E, Schwalbe H, Dotsch V, Trantirek L. *Evaluation of parameters critical for observing nucleic acids inside living* Xenopus laevis *oocytes by in-cell NMR spectroscopy. J. Am. Chem. Soc. 2009;131:15761-15768.*
3. Krstic I, Hansel R, Romainczyk O, Engels JW, Dotsch V, Prisner TF. *Long-range distance measurements on nucleic acids in cells by pulsed EPR spectroscopy. Angew. Chem. Int. Edit. 2011;50:5070-5074.*
4. Azarkh M, Okle O, Singh V, Seemann IT, Hartig JS, Dietrich DR, Drescher M. *Long-range distance determination in a DNA model system inside* Xenopus laevis oocytes *by in-cell spin-label EPR. Chembiochem 2011;12:1192-1195.*
5. Selvin PR. *The renaissance of fluorescence resonance energy transfer. Nat. Struct. Biol. 2000;7:730-734.*
6. Stryer L. *Fluorescence energy-transfer as a spectroscopic ruler. Annu. Rev. Biochem. 1978;47:819-846.*
7. Stryer L, Haugland RP. *Energy transfer - a spectroscopic ruler. Proc. Natl Acad. Sci. USA 1967;58:719-726.*
8. Hurley DJ, Tor Y. *Donor/acceptor interactions in systematically modified Ru-II-Os-II oligonucleotides. J. Am. Chem. Soc. 2002;124:13231-13241.*
9. Lilley DMJ, Wilson TJ. *Fluorescence resonance energy transfer as a structural tool for nucleic acids. Curr. Opin. Chem. Biol. 2000;4:507-517.*
10. Woodside MT, Anthony PC, Behnke-Parks WM, Larizadeh K, Herschlag D, Block SM. *Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. Science 2006;314:1001-1004.*
11. Wozniak AK, Schroder GF, Grubmuller H, Seidel CAM, Oesterhelt F. *Single-molecule FRET measures bends and kinks in DNA. Proc. Natl Acad. Sci. USA 2008;105:18337-18342.*
12. Ishii M, Ikushima M, Kurachi Y. In vivo *interaction between RGS4 and calmodulin visualized with FRET techniques: possible involvement of lipid raft. Biochem. Biophys. Res. Commun. 2005;338:839-846.*
13. Gaibelet G, Planchenault T, Mazeres S, Dumas F, Arenzana-Seisdedos F, Lopez A, Lagane B, Bachelerie F. *CD4 and CCR5 constitutively interact at the plasma membrane of living cells: a confocal fluorescence resonance energy transfer-based approach. J. Biol. Chem. 2006;281:37921-37929.*

14. Chilibeck KA, Wu T, Liang C, Schellenberg MJ, Gesner EM, Lynch JM, MacMillan AM. *FRET analysis of* in vivo *dimerization by RNA-editing enzymes. J. Biol. Chem. 2006;281:16530-16535.*

15. Schoen I, Krammer H, Braun D. *Hybridization kinetics is different inside cells. Proc. Natl Acad. Sci. USA 2009;106:21649-21654.*

16. Uchiyama H, Hirano K, Kashiwasake-Jibu M, Taira K. *Detection of undegraded oligonucleotides* in vivo *by fluorescence resonance energy transfer. J. Biol. Chem. 1996;271:380-384.*

17. Treutlein B, Michaelis J. *Direct observation of single RNA polymerase processing through a single endogenous gene in a living yeast cell. Angew. Chem. Int. Ed. Engl. 2011;50:9788-9790.*

18. Enderlein J, Erdmann R. *Fast fitting of multi-exponential decay curves. Opt. Commun. 1997;134:371-378.*

19. Huang F, Lerner E, Sato S, Amir D, Haas E, Fersht AR. *Time-resolved fluorescence resonance energy transfer study shows a compact denatured state of the B domain of protein A. Biochemistry 2009;48:3468-3476.*

20. Williams ATR, Winfield SA, Miller JN. *Relative fluorescence quantum yields using a computer-controlled luminescence spectrometer. Analyst 1983;108:1067-1071.*

21. Dale RE, Eisinger J, Blumberg WE. *Orientational freedom of molecular probes - orientation factor in intra-molecular energy-transfer. Biophys. J. 1979;26:161-193.*

22. Muschielok A, Andrecka J, Jawhari A, Bruckner F, Cramer P, Michaelis J. *A nano-positioning system for macromolecular structural analysis. Nat. Methods 2008;5:965-971.*

23. Sindbert S, Kalinin S, Nguyen H, Kienzler A, Clima L, Bannwarth W, Appel B, Muller S, Seidel CA. *Accurate distance determination of nucleic acids via Forster resonance energy transfer: implications of dye linker length and rigidity. J. Am. Chem. Soc. 2011;133:2463-2480.*

24. Clifford JN, Bell TDM, Tinnefeld P, Heilemann M, Melnikov SM, Hotta J, Sliwa M, Dedecker P, Sauer M, Hofkens J, et al. *Fluorescence of single molecules in polymer films: Sensitivity of blinking to local environment. J. Phys. Chem. B 2007;111:6987-6991.*

25. Fessl T, Ben-Yaish S, Vacha F, Adamec F, Zalevsky Z. *Depth of focus extended microscope configuration for imaging of incorporated groups of molecules, DNA constructs and clusters inside bacterial cells. Opt. Commun. 2009;282:2495-2501.*

26. Wang JC. *Helical repeat of DNA in solution. Proc. Natl Acad. Sci. USA 1979;76:200-203.*

27. Kypr J, Chladkova J, Zimulova M, Vorlickova M. *Aqueous trifluorethanol solutions simulate the environment of DNA in the crystalline state. Nucleic Acids Res. 1999;27:3466-3473.*

28. Minyat EE, Ivanov VI, Kritzyn AM, Minchenkova LE, Schyolkina AK. *Spermine and spermidine-induced B to A transition of DNA in solution. J. Mol. Biol. 1979;128:397-409.*

29. Xu Q, Shoemaker RK, Braunlin WH. *Induction of B-A transitions of deoxyoligonucleotides by multivalent cations in dilute aqueous solution. Biophys. J. 1993;65:1039-1049.*

30. Robinson H, Wang AH. *Neomycin, spermine and hexaamminecobalt (III) share common structural motifs in converting B- to A-DNA. Nucleic Acids Res. 1996;24:676-682.*

31. Brahms J, Mommaerts WF. *A study of conformation of nucleic acids in solution by means of circular dichroism. J. Mol. Biol. 1964;10:73-88.*

32. Ivanov VI, Minchenkova LE, Minyat EE, Frank-Kamenetskii MD, Schyolkina AK. *The B to A transition of DNA in solution. J. Mol. Biol. 1974;87:817-833.*

33. Vorlickova M, Minyat EE, Kypr J. *Cooperative changes in the chiroptical properties of DNA induced by methanol. Biopolymers 1984;23:1-4.*

34. Florentiev VL, Ivanov VI. *RNA polymerase: two-step mechanism with overlapping steps. Nature 1970;228:519-522.*

35. Beabealashvily RS, Ivanov VI, Minchenkova LE, Savotchkina LP. *RNA polymerase-DNA complexes. I. The study of the conformation of nucleic acids at the growing point of RNA in an RNA polymerase-DNA system. Biochim. Biophys. Acta 1972;259:35-40.*

36. Setlow P. *DNA in dormant spores of Bacillus species is in an A-like conformation. Mol. Microbiol. 1992;6:563-567.*

37. Chin DJ, Green GA, Zon G, Szoka FC Jr., Straubinger RM. *Rapid nuclear accumulation of injected oligodeoxyribonucleotides. New Biologist 1990;2:1091-1100.*

38. Leonetti JP, Mechti N, Degols G, Gagnor C, Lebleu B. *Intracellular distribution of microinjected antisense oligonucleotides. Proc. Natl Acad. Sci. USA 1991;88:2702-2706.*

39. Clarenc JP, Lebleu B, Leonetti JP. *Characterization of the nuclear binding sites of oligodeoxyribonucleotides and their analogs. J. Biol. Chem. 1993;268:5600-5604.*

40. Fisher TL, Terhorst T, Cao X, Wagner RW. *Intracellular disposition and metabolism of fluorescently-labeled unmodified and modified oligonucleotides microinjected into mammalian cells. Nucleic Acids Res. 1993;21:3857-3865.*

41. Hansel R. *2009. Evaluation of parameters critical for observing nucleic acids inside living Xenopus laevis oocytes by in-cell NMR spectroscopy. PhD thesis - Johan Wolfgang Goethe University, Frankfurt am Main, Germany.*

42. Azarkh M, Okle O, Eyring P, Dietrich DR, Drescher M. *Evaluation of spin labels for in-cell EPR by analysis of nitroxide reduction in cell extract of Xenopus laevis oocytes. J. Magn. Reson. 2011;212:450-454.*

43. Sedoris KC, Thomas SD, Clarkson CR, Muench D, Islam A, Singh R, Miller DM. *Genomic c-Myc quadruplex DNA selectively kills leukemia. Mol. Cancer Ther. 2012;11:66-76.*

3. Conclusions

SMFD techniques offer genuine detection possibilities that are often inaccessible using ensemble methods. This was demonstrated in three projects investigating translocation activity of CHD4 protein, analysis of MS2 phage capsid assembly and in-cell characterization of DNA structure. In another project, binding interactions between two fluorescent probes and a short oligonucleotide were characterized. Knowledge of such information is valuable for FRET applications

SMFD instrument upgrade

Imaging of small objects such as individual molecules in bacterial cells is difficult due to the lateral resolution, but also due to the lack of sufficient axial depth of focus. The second condition allows having the full cell volume focused simultaneously. We implemented an optical real-time technique of extension in the depth of focus that can be incorporated in any high numerical aperture microscope set-up. This technique can be used for imaging of individual molecules, especially inside live cells. The resulting optical element was integrated into a single-molecule optical system. We subsequently demonstrated an improvement in the depth of focus while maintaining the same lateral resolution.

Later, we used this element to optimize the incorporation of labeled oligonucleotides into *E. coli* cells. Due to the extension in focal depth, we were able to detect virtually all fluorescent molecules within the volume of an E. coli cell, and hence find the proper incorporation conditions and minimize the portion of background which is otherwise formed by molecules out of focus.

During my Ph.D. study I also participated in an upgrade of the existing steady-state SMFD instrument to time-resolved apparatus with the capability to detect two-channel time-resolved photon streams. Time-resolved apparatus can be utilized to detect robust lifetime-based FRET and time-resolved fluorescence anisotropy.

DNA-probe binding interaction

For quantitative FRET applications, it is crucial to precisely characterize the interaction of fluorescent probes with the probed biomolecule. We investigated noncovalent binding interactions between a short B-DNA oligonucleotide and terminally attached nonfluorescent quencher QSY 21 and Rhodamine 6G. Using molecular dynamics and *ab initio* quantum chemical calculations, we have found that QSY 21 has two binding motifs. In the first motif, the central xanthene ring (with transition absorption dipole) is stacked on one base of the adjacent cytosine−guanine DNA base pair, whereas one of the 2, 3-dihydro-1-indolyl aromatic side rings is stacked on the other base. In the second motif, the QSY 21 stacking interaction with the DNA base pair is mediated only by one of the side rings. Several transitions between the motifs were observed during a molecular dynamics simulation. The *ab initio* calculations showed that none of these motifs is energetically preferred.

Two binding motifs were also found for Rhodamine 6G, with the xanthene ring stacked predominantly on either the cytosine or on the guanine. Our results suggest that the side rings of QSY 21 play a crucial role in its stacking on the DNA and indicate a novel binding mode absent in the case of Rhodamine 6G, which lacks

aromatic side rings. Besides this we characterized the position of the probe with respect to the DNA molecule. These DNA – probe interactions must be considered within quantitative FRET analysis to increase the distance determination accuracy and to minimize possible misinterpretations.

Translocation activity of CHD4 protein at single-molecule level
An exclusive feature of SMFD tools is that synchronization of individual enzymatic processes and action of molecular motors is not needed. We utilized this feature on the characterization of translocation activity of CHD4 chromatin remodeler. SpFRET was employed to investigate CHD4 translocation, revealing step size, directionality of translocation and the ATP dependence of the whole action.

From the presented spFRET data, we suggest that binding of CHD4 to DNA is not ATP dependent. This was supported by an ensemble FRET experiment where after addition of ATP no measurable change in FRET efficiency appeared. By contrast, translocation of CHD4 along a DNA molecule seems to be ATP driven, since after addition of ATP, spFRET time trajectories exhibited stepping behavior, typical for the movement of a molecular machine. From the proximity ratio steps, the translocation directionality of CHD4 can be qualified. As these steps were detected in both directions i.e. towards the 5' end and also the 3' end of DNA, we propose that CHD4 is a bidirectional translocase. The size of the translocation step was estimated to 2-3 DNA base pairs.

We propose that the CHD4 protein without modification moves along DNA in both directions as a molecular motor driven by symmetric chemical potential. Its localization at particular loci, such as during double strand break signalization and repair, is probably controlled by post-translational modifications, or by interaction with other proteins or modified histones, creating a functional complex.

FCS analysis of MS2 phage capsid assembly
Molecular biology benefits from the ability of SMFD to investigate statistical distributions, which can be followed in a long time scale. We used FCS to monitor the assembly of MS2 bacteriophage capsids. Here, we employed FCS to follow changes in the diffusion coefficient of intermediates during the assembly pathway from initial components (capsid proteins and trRNA) to fully assembled MS2 capsids.

However, since FCS is not strictly single-molecule technique, it does not give direct statistical distributions. Use of advanced data processing is necessary to reveal the time dependent changes of stoichiometry of individual assembly components. Hence, we applied singular value decomposition and regularized inverse Laplace distribution fitting and compared their outputs.

Both methods were sensitive enough to observe a temporary kinetic trap and provided comparable results according to data obtained by electron microscopy or mass spectrometry. The main advantage of this distribution fitting procedure is in characterization of diffusion times gathered for every single measured assembly step and purified component. Thus during the progression of the assembly reaction we can analyze changes in diffusion times (~size) of growing capsid intermediates and observe decreasing concentration of initial components, while the concentration of capsid is increasing. Nevertheless this method is time demanding and does not

necessarily reveal the true and unique concentration loading of the intermediates.

By contrast, singular value decomposition quickly reveals concentration loading of each detected factor (component). A combination of both methods seems to be most promising. It benefits from the capability of singular value decomposition to follow loading of individual assembly factors, which were beforehand determined by the probability distribution fitting of measured data.

In-cell characterization of DNA structure

SMFD techniques offer unique possibilities to non-invasively probe biomolecular structure and function in their native environment. Furthermore biomolecules can be probed with respect to their localization, for example proteins can be tracked and probed during their passage through different cellular compartments and DNA properties can be investigated inside prokaryotic cells or localized in the nucleus of eukaryotic cells. Moreover, SMFD techniques allow timing of biomolecular studies; thus biomolecules can be studied in a particular phase of cell cycle.

Here we explored parameters critical for characterization of nucleic acid structure using spFRET inside living E. coli cells. These parameters involve mainly spectroscopic properties of both fluorophores forming the FRET pair. Our measurements showed that the fluorophore properties in the cellular environment differed from those acquired under *in vitro* conditions. The precision for the interprobe distance determination from FRET efficiency values acquired *in vivo* was found to be lower (~ 31%) compared to that acquired in diluted buffer (13%). Our data suggest that despite its low precision, the in-cell FRET measurements can be successfully applied to discriminate among various structural models.

Single-molecule fluorescence detection in molecular biology.
Ph.D. Thesis Series, 2012, No. 14