THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Single molecule methods for DNA-protein interaction studies

Carl Möller

Department of Life Sciences

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2025

Single molecule methods for DNA-protein interaction studies
Carl Möller
ISBN 978-91-8103-243-7

Acknowledgements, dedications, and similar personal statements in this Thesis, reflect the author's own views.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5701
ISSN 0346-718X

Department of Life Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
[Illustration of a microscope with a functionalized cover slide. The zoomed in view shows a typical sample of YOYO-1 dyed λ-DNA colocalizing with aptamer-QDOT labeled hMR. Further reading on the method and the experiments can be found under section 5.2 and 6.1.2.3]

[Chalmers Digitaltryck]
Gothenburg, Sweden 2025

Single molecule methods for DNA-protein interaction studies

Carl Möller
Department of Life Sciences
Chalmers University of Technology

ABSTRACT

In the volume of a typical cell, even a single molecule significantly affects the concentration. Therefore, at typical affinities, only a few proteins are required to saturate binding to, for example, DNA. Compartmentalization further lowers this threshold, underscoring the need for sensitivity and molecular recognition and highlighting the importance of precision in biological research.

Given the small number of molecules required to reach functional concentrations in a cell, it becomes clear that studying molecular interactions at the single-molecule level can provide important insights. However, studies of biochemical reactions, such as DNA-protein interactions, are typically done with bulk methods. This approach is robust, but the result is based on a time and population average of all the molecules present in the reaction. If multiple populations co-exist, the average can be misleading. By isolating molecules so that they can be studied one-by-one, as in single-molecule methods, it is possible to resolve separate populations that otherwise would have been hidden.

The original work presented in this Thesis explores the application of single-molecule methods for studying DNA-protein interactions. The work is focused on the use of quantitative fluorescence microscopy and involves the use of nanofluidics to study the DNA-repair protein complexes MRN (human) and MRX (yeast), and the role of their individual components in DNA end-joining. The potential of nanofluidics was further explored by establishing the influence of divalent metal ions and ATP on the binding of the fluorescent dye YOYO-1 to DNA and the impact this has on studies of active protein processes on DNA. The Thesis also encompasses work involving DNA-protein complexes immobilized on functionalized glass slides. It presents a novel quantitative method where colocalization with a fluorescent dCas9 is used to identify and size viral vectors isolated from cells. Furthermore, the same principle is used in a study showing that the oncogenic protein MYC is involved in alternating the activity of Topoisomerase 2A by co-condensation. Taken together, the work presented in Thesis provides valuable insights in how to isolate and immobilize DNA-protein complexes and combine this with quantitative fluorescence microscopy to extract meaningful biological insights from single-molecule data.

Keywords: DNA, nanofluidics, fluorescence microscopy, single molecule, DNA repair, biomolecular interactions

List of Publications

This Thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I. **Xrs2/NBS1 promote end-bridging activity of the MRE11-RAD50 complex**
Carl Möller, Rajhans Sharma, Robin Öz, Giordano Reginato, Elda Cannavo, Ilaria Ceppi, K.K. Sriram, Petr Cejka, Fredrik Westerlund
Biochem. Biophys. Res. Commun. **695**, 149464 (2023) doi:10.1016/j.bbrc.2023.149464.

II. **Effects of Mg2+ and ATP on YOYO-1 labelling of genomic DNA in single molecule nanofluidic experiments**
Carl Möller, Dennis Winters, Radhika Nambannor Kunnath, Sriram KK, Fredrik Westerlund
(Manuscript submitted to Biochemistry and Biophysics Reports, May 2025. Revised and resubmitted, July 2025)

III. **Cas9-leveraged single-molecule characterization of sparse plasmid vectors in heterogenous DNA samples**
Carl Möller, Luis Leal-Garza, Emanuele Celauro, Roberto Nitsch, Fredrik Westerlund
(Manuscript submitted to Applied Biochemistry and Biotechnology January 2025. Under revision, July 2025)

IV. **MYC modulates TOP2A diffusion to promote substrate detection and activity**
Donald Cameron, Kathryn Jackson, Alessia Loffreda, Carl Möller, Matteo Mazzocca, Evgeniya Pavlova, Bea Jagodic, Fredrik Westerlund, Davide Mazza, Laura Baranello
(Manuscript submitted to Nature Communications, July 2025)

Contribution Report

Below follows a description of the author's contribution to the papers appended in this Thesis:

I. C.M. designed and performed nanofluidic experiments together with R.S and R.Ö. C.M. hypothesized and designed experiments related to the use of aptamers for affinity based fluorescent labelling of proteins. C.M. designed and performed glass slide experiments. C.M. curated all data, developed necessary analysis tools and finalized the analysis, including visualization and interpretation of the data. C.M. wrote the paper in collaboration with F.W.

II. C.M. designed the nanofluidic experiments and performed the data analysis. C.M. performed electrophoresis and plate reader experiments as well as developed the necessary analysis tools (including visualization) for all data used in the study. C.M. wrote the manuscript in collaboration with D.W., S.KK. and F.W.

III. C.M. designed and performed all single molecule experiments as well as all adjacent experiments coupled to characterization and optimization. C.M. developed a full analysis pipeline for quantitation of single molecule data. C.M. designed and performed qPCR experiments and analysis, including isolation and preparation of genetic material. C.M. wrote the manuscript in collaboration with F.W.

IV. C.M. consulted on method development coupled to fluorescence microscopy experiments aiming to image single condensates immobilized on glass and the colocalization between fluorescently labeled DNA, TOP2A and MYC. C.M. performed the experiments and curated the data, developed the necessary analysis tools for quantitation of the single molecule data and contributed to the interpretation of the findings. C.M. provided feedback on the manuscript.

Preface

This dissertation was submitted for the partial fulfilment of the degree of Doctor of Philosophy. The original work presented in the dissertation was carried out between June 2020 and June 2025 at the Department of Life Sciences (previously Biology and Biological Engineering), Chalmers University of Technology, under the supervision of Professor Fredrik Westerlund. The research was funded by the European Research Council.

Carl Möller
July, 2025

Table of Contents

1 Introduction 1

2 DNA 5

2.1 Guiding principle of molecular biology 5

2.2 The chemical structure of DNA 6

2.3 Physical properties of DNA 7

2.3.1 DNA free in solution 7

2.3.2 DNA in confinement 8

3 Proteins 9

3.1 DNA-protein interactions 9

3.1.1 DNA-repair 9

3.1.2 CRISPR-Cas9 11

3.1.3 Genome maintenance 12

3.1.4 Liquid-Liquid Phase Separation 13

4 Fluorescence 15

4.1 Physical principle of fluorescence 15

4.2 Fluorophores 16

4.2.1 DNA-specific fluorophores 17

4.3 Fluorescence microscopy 18

5 Studying single biomolecules 21

5.1 Stretching DNA in nanochannels 21

5.2 Immobilizing biomolecules on glass 22

6 Original work 23

6.1 Nanofluidics for studying end-specific DNA-protein interactions 23

6.1.1 Information on DNA-protein interactions is inferred from fluorescence and molecule conformation 24

6.1.2 Analysis of cyclisation and concatemerization by Xrs2/NBS1 highlights their important role in DNA-tethering by the MRE11-RAD50 complex 26

6.2 Effects of Mg^{2+} and ATP on YOYO-1 labeling of genomic DNA in single molecule experiments 31

6.2.1 YOYO-1 binding constants reflects Mg^{2+} and ATP concentrations 32

6.2.2 Less photolytic damage of YOYO-1 labeled DNA when Mg^{2+} is present 33

6.3 Single-molecule characterization of plasmid vectors 35

6.3.1 Fluorescence intensity from YOYO-1 labeled DNA scales with size 35

6.3.2 GFP expression and qPCR of viral vectors from transduced HEK293 36

6.3.3 Intensity-based sizing coupled with dCas9 colocalization can be used to confidently resolve sparse DNA populations in heterogenous samples 38

6.4 Fluorescence microscopy of multi-component LLPS puncta 39

6.4.1 TOP2A forms puncta at physiological concentrations 39

6.4.2 MYC stimulates condensate formation by TOP2A 40

6.4.3 MYC accelerates TOP2A target search and activity 43

7 Concluding remarks 45

8 Acknowledgements 47

9 References 49

1 Introduction

A biochemical reaction is a process that takes place within a living organism. These reactions are typically catalyzed by proteins and are essential for sustaining life, as they underlie all biological functions. Cells operate at the nanoscale, where the number of molecules involved in key processes can be remarkably small.

A typical bacterium, such as *E. coli*, has a volume of ~1 μm^3. If we consider a single molecule or protein within this volume it corresponds to a concentration of ~1.6 nM. If we then consider a typical binding constant (K) in the range of 10^{-9} - 10^{-12} M one can see that a cell of this size does not require more than ~100 molecules or proteins to saturate the binding. Given that a cell is compartmentalized, this number becomes even smaller, highlighting how studying molecular interactions at the single-molecule level can provide a level of resolution that allows researchers to observe stochastic behaviors and transient interactions.

However, the most common way to study biochemical reactions, such as proteins interacting with deoxyribonucleic acid (DNA), is with bulk methods, where the readout is dependent on the behavior of a large number of molecules. This approach is robust, and the signal strength is typically depending on the number of molecules. However, the result is a time and population average of all the molecules present in the reaction. If multiple populations co-exist, the average can be misleading. By studying molecules one-by-one it is possible to discern different populations that otherwise would have been hidden in the ensemble average[1]. There is today an array of single molecule techniques that are developed to probe and characterize a reaction or interaction on the individual molecule level. In common for all single molecule techniques is the need to separate a molecule, or set of molecules, and the need for a readout that is sensitive enough at such a small scale.

One of the earliest reports of single molecule detection using fluorescence microscopy came in 1981, where the authors reported images of surface deposited DNA labeled with the fluorescent dye DAPI[2]. Since then, the field, and the available technologies, have developed immensely. The introduction of new DNA specific dyes, such as YOYO-1[3], provided an important improvement in signal specificity and contrast. YOYO-1 is a primary tool for single molecule DNA visualization[4] and while many basic aspects of the interaction with DNA is well characterized there remains open questions on its use in presence of multivalent ions.

The original work presented in this Thesis aims to demonstrate and explore the application of single molecule methodologies for studies of DNA-protein interactions. The work is revolving around the use of nanofluidics and surface immobilization combined with quantitative wide-field fluorescence microscopy to image single DNA molecules or DNA-protein complexes. While the work presented in this Thesis uses fluorescence as the readout when detecting and quantifying single DNA molecules, there are in many additional methods, where electron microscopy[5] (EM), atomic force microscopy[6,7] (AFM) and (magnetic and optical) tweezers[8,9] are common examples. Electron microscopy provides a high-resolution image of an immobilized molecule by detecting the transmission or scattering of an electron beam. AFM generates an image by tapping a very fine tip along a surface, detecting the difference in height. Magnetic and optical tweezers require both the ends of the DNA to be

attached to a surface or a bead. By pulling the bead, the DNA can be stretched and the force applied measured, thus probing events such as folding and unfolding of secondary structures or protein binding.

A common requirement when probing DNA on the single DNA molecule level is the need to extend the DNA from its coiled state into a linear conformation to visualize static and dynamic interactions with other molecules. Before the advent of nanofluidic confinement, a main topic of this Thesis, the main method to isolate and stretch individual DNA molecules (for studies with fluorescence microscopy) was stretching by flow or by anchoring the DNA to a surface so that the ends could be pulled apart. One of the earliest studies of DNA in nanochannels was published in 2004[10] and demonstrated how DNA could be stretched solely by confinement, without the need for anchoring to a surface. This development has brought on a vast array of applications across several research fields.

The insight into the role of DNA as a storage of information has changed the course of modern biology. Maintenance of DNA is a key and continuous process in the cell and roughly 500 protein-coding genes are inferred to be involved in upholding the integrity of the human genome. One single double-stranded break (DSB) may result in disruption of vital cellular processes and therefore requires immediate attention. While being well characterized in most prokaryotes and eukaryotes, details regarding important intermediate steps in DSB repair are still to be fully explored, for example how two broken ends are brought into proximity for repair to occur. **Paper I** explores the capabilities of nanofluidics as a tool for investigating end-specific reactions on DNA and reports an extensive study on the human and yeast homologs, MRN and MRX, and their role in DSB repair. By distinguishing between circularization and concatemerization events, the work demonstrates a framework for characterisation of the end-bridging activity of DNA repair proteins at the single-molecule level. In a systematic analysis of the MRN and MRX complexes and their subunits (NBS1 and Xrs2), the presented research reveals distinct mechanistic differences in how these complexes mediate DNA end-joining. These results contribute to further the understanding the balance between homologous recombination and non-homologous end joining and highlights the synergistic roles of subunits at a level of detail that is difficult to resolve with bulk assays.

Paper II investigates how to best setup studies of biochemical reactions in nanofluidic channels. The study investigates the effects of Mg^{2+} and ATP on YOYO-1 fluorescence and provides critical insights for experimental design in single-molecule studies. By quantifying how these cofactors influence dye binding and photostability, the work helps define suitable conditions for imaging active protein-DNA systems, which is essential for future studies involving ATP-dependent enzymes.

Papers III-IV revolve around the deposition of DNA-protein complexes on functionalized glass surfaces. **Paper III** introduces a method using Cas9 for identifying and sizing sparse DNA molecules in heterogeneous samples. The method offers a refinement free strategy for detecting structural features such as concatemers in miniaturized samples, providing an important concept and research tool with potential applications in gene therapy, synthetic biology, and diagnostics. **Paper IV** investigates the formation of biomolecular condensates by TOP2A and the protooncogene MYC. The study contributes to the use of fluorescence microscopy in the field of liquid-liquid phase separation (LLPS). By showing how MYC modulates condensate size,

number, and DNA content, the study adds to the understanding of how transcriptional regulators influence genome organization through phase-separated compartments.

Across all presented studies, the work demonstrates how single molecule methods and quantitative fluorescence microscopy can be integrated to study DNA-protein interactions and extract meaningful biological insights from single-molecule data.

2 DNA

The main function of DNA is as a medium for information storage. The information is stored as a four-letter code defined by the sequential appearance of adenosine (A), thymine (T), cytosine (C) and guanine (G). This basic sequence is divided in codons (three letter words) that encode specific amino acids, the building blocks of proteins. Taken together, a set of codons specifying an ordered sequence of amino acids is referred to as a gene. Each gene encodes instructions to assemble a protein. Proteins are the fundamental building blocks of all cells and are not only responsible for most of the elemental functions in a cell, but also interact with DNA in many different processes.

2.1 Guiding principle of molecular biology

The central dogma of molecular biology describes the generalized flow of genetic information from long term storage in DNA via ribonucleic acid (RNA) to protein[11] (**Figure 1)**. This pathway is a simplified description, but captures the essential transfer of the biological sequence information that is fundamental to life. The process of transcription, the first step in going from DNA to protein, copies the relevant sequence (gene) from DNA into RNA. RNA is a similar molecule to DNA but differs in its structure, where RNA is built by ribose instead of deoxyribose as well as uses uracil (U) instead of T. Proteins are then built in a process referred to as translation where an amino acid polymer is built according to the order of the codons in the RNA. The amino acid polymer will gain its function and 3D structure by folding itself, thus connecting distant parts of the chain[11,12].

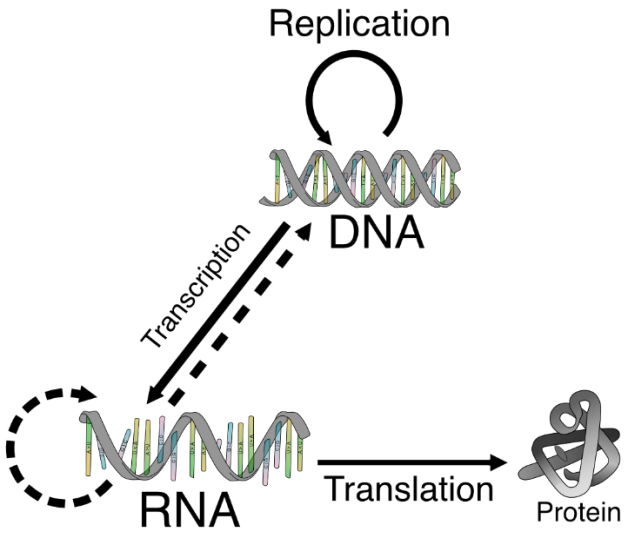

Figure 1. The central dogma of molecular biology describes the flow of information from DNA to proteins via RNA. This is fundamental to all biological processes. Adapted from "Central dogma of molecular biology"[13], licensed under CC BY-SA 3.0.

2.2 The chemical structure of DNA

DNA is composed of two single strands of DNA (ssDNA) that are held together by hydrogen bonding between opposing nucleobases, **Figure 2**. The determination of the double helical structure of DNA has famously been attributed to Watson and Crick[14], but has lately been revised to also acknowledge the contribution of Franklin[15]. Each polynucleotide chain consists of monomeric nucleotides, which are composed of a deoxyribose sugar and a phosphate group linked by a phosphodiester bond between the hydroxyl group of the 5′-carbon of one deoxyribose and the hydroxyl group of the 3′-carbon on the next. This gives rise to a directionality, since the chain extends by the addition of a nucleotide to the 3′-carbon hydroxyl group on the terminal deoxyribose. To each deoxyribose, one of the four nucleobases is attached: adenosine (A), thymine (T), guanine (G) or cytosine (C). In dsDNA, the nucleotides are organized in pairs, A with T and G with C (referred to as base pairs (bp)), that are held together by hydrogen bonding. This is what brings the two polynucleotide chains together to form the typical helix of double stranded DNA (dsDNA). At neutral pH, the bases are uncharged, and the sugar-phosphate backbone negatively charged. The stacking of the base pairs shields their hydrophobic structures inwards, exposing the polar backbone to the solvent.

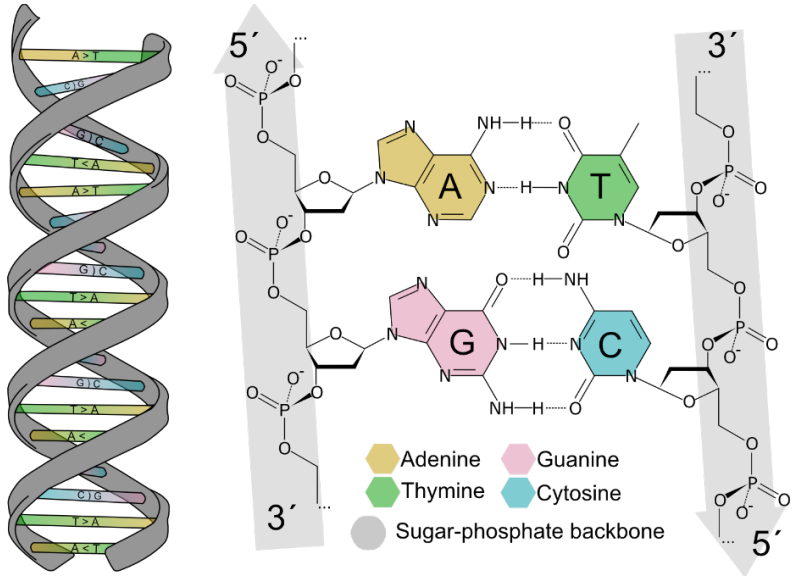

Figure 2. The helical structure of dsDNA with the four nucleobases and their respective interactions. dsDNA is composed of two polymeric strands of nucleotides. Each strand consists of a sugar-phosphate backbone that is built from alternating deoxyribose sugar and phosphate groups, connected by a phosphodiester bond. The nucleotides sit on the deoxyribose sugar and hybridize with a corresponding nucleotide on the opposing strand via hydrogen bonding. The arrows indicate the anti-parallel direction of each strand. Adapted from "DNA structure formula virgin"[16] and "DNA structure and bases"[17], licensed under CC BY-SA 3.0.

In cells, DNA will almost always be found in the B-form conformation[12], depicted in **Figure 2**. However, DNA can also take on other conformations such as A- and Z-form, which differ significantly from the B-form, but are still helical structures[12]. B-form DNA is a right-handed helical structure with a typical rise of 0.34 nm/bp at a 36-degree angle and a width of ~2 nm between the two opposing sugar-phosphate backbones[18], which gives a periodicity of 10 bp. The traces along the helical axis between the backbones are called grooves and because the strands are not completely paralell, the width and depth of the grooves differ, creating a major and minor groove. The groves often serve as binding sites for proteins and small molecules.

2.3 Physical properties of DNA

DNA is a polymer with structural and mechanical properties that influence its biological function. The physical characteristics, such contour length and response to confinement defines experimental techniques used to study DNA at the single-molecule level, where its conformation, extension, and interactions with proteins can be observed and quantified.

2.3.1 DNA free in solution

DNA is a polymer and will, when free in solution, try to minimize its free energy by taking on a coiled structure, **Figure 3**. A common model used to represent the behavior of DNA is the worm-like chain (WLC) model, which assumes the DNA to behave like a continuous semi-flexible polymer[18,19]. The model uses two physical quantities of the DNA, the contour length (L) and the persistence length (P). The contour length describes the end-to-end distance of a fully stretched out polymer, or, in the case of B-DNA, the number of bp (N) multiplied with the average rise per bp (0.34 nm)[20]. The persistence length refers to the minimum length at which the polymer cannot be bent, or the distance over which the direction of the tangent to the chain is correlated[20]. This can be described by the average cosine of the angle θ between the tangents of the chain at a fixed position[18]:

$$\langle \cos\theta \rangle = e^{-L/P} \qquad (1)$$

For very short DNA molecules, where L<<P, the tangents will be highly correlated with an angle close to 0 giving a cosine of 1, equivalent to the behavior of a rigid rod. This is further supported by[18]:

$$\langle L^2 \rangle = 2p\left(L - P + Pe^{-L/P}\right) \qquad (2)$$

The persistence length (P) is proportional to the bending stiffness of the polymer chain and is for DNA around 50 nm (~147 bp) at physiological conditions[18]. Long DNA molecules in solution with L>>P will take on the conformation of a random coil to minimize the energy and maximize the entropy. The size of a coiled polymer can be described by the radius-of-gyration (R_G) which, for an ideal polymer, is given by random walk statistics[21]:

$$R_G^2 = \frac{Nb^2}{6} \qquad (3)$$

where N is the number of statistical segments and b the statistical segment length, from which $L = Nb$ follows. The electrostatic forces from the negatively charged sugar-phosphate backbone of a DNA molecule will stop it from fully collapsing. The distance at which the effect of electrostatic interactions is felt is the Debye length, which is roughly 1 nm (at physiological conditions) for the sugar-phosphate backbone[18]. This will change the effective width of the DNA, which can be roughly estimated as the intrinsic width (~2 nm) plus twice the Debye length. The effective width describes the effective interaction range of a backfolded DNA chain and essentially determines the closest distance between two DNA chains before they are repelled by electrostatic interactions. When this self-avoidance is considered, the expression for R_G is[21]:

$$R_G \cong L^{3/5} P^{1/5} w^{1/5} \qquad (4)$$

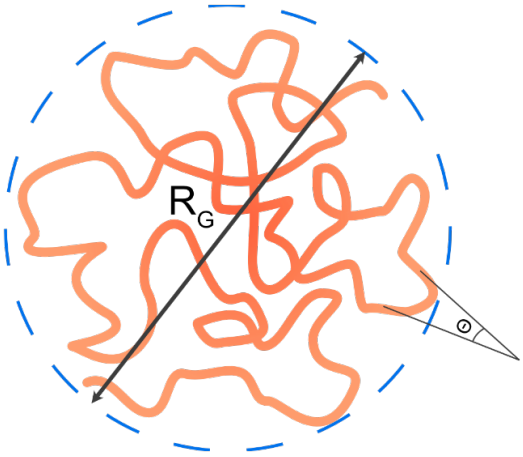

Figure 3. Depiction of a DNA polymer that has taken on a coiled structure to minimize its free energy. The radius of gyration (R_G) is the root-mean-square of the distance from each segment to the center of mass. The tangents from a representative segment give the angle (θ) from which the persistence length can be calculated according to Equation 1.

2.3.2 DNA in confinement

By confining a DNA molecule in a narrow channel, it is possible to disrupt the coiled structure and stretch the DNA in the channel direction. The channel dimensions need to satisfy $D_{AV} < R_G$ (where D_{AV} is the average cross-section area of a channel defined by $\sqrt{D_{width}D_{heigtht}}$) to force the molecule to stretch[20,21]. The self-avoidance of DNA causes it to distribute evenly in the channel direction. From this follows that the apparent extension of the molecule scales with its contour length, which in turn means that the position of an observed feature along DNA in a channel, directly corresponds to the position of the feature in the nucleotide sequence. The degree of confinement is usually divided into regimes in which the behavior of a polymer can be described by a statistical model (**Figure 4**). A molecule under weak confinement ($P \ll D_{AV} < R_G$) can be described by considering the polymer as a series of non-interacting spherical isometric units that have a radius equal to the channel dimensions. This model was first proposed by de Gennes[22] in 1977. When the confinement is increased ($P < D_{AV}$) the polymer follows a model presented by Odijk[23], giving the name to this regime. Under strong confinement the energy needed to loop the polymer increases and it takes on an elongated conformation, where it undulates between the channel walls.

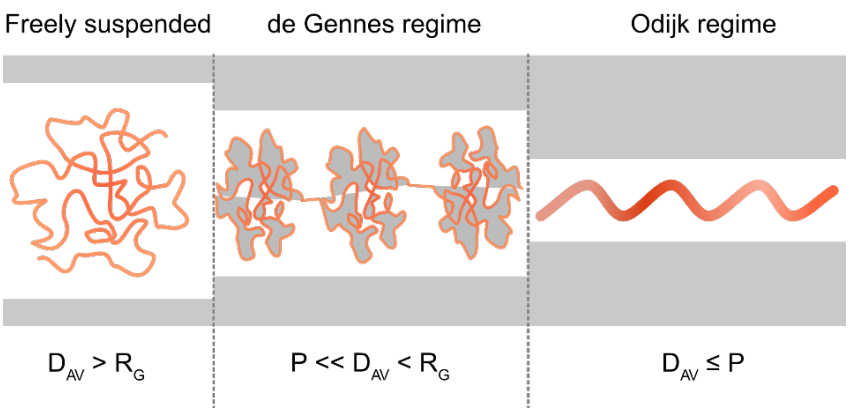

Figure 4. Schematic of a polymer under different degrees of confinement. When the channel dimensions are larger than R_G the polymer does not experience any confinement and will be freely suspended. As the dimensions are reduced below R_G (de Gennes regime) the polymer behavior can be described as a series of non-interacting units. When the channels are smaller than the persistence length, the polymer enters the Odijk regime.

3 Proteins

Proteins are the essential building blocks in a cell and make out a majority of its mass, catalyzing or executing almost all cellular functions. Just like DNA, proteins are linear polymers. The polymer chain is built by 20 different amino acids with differing chemical properties (such as acidity, polarity and charge). The three-dimensional shape (conformation) of a protein is achieved by folding of the polymer chain. The final shape is dependent on the amino acid sequence and their properties, and it is the conformation of a protein that determines its function. Each amino acid is comprised of an amino group ($-NH_2$), a carboxyl group (-COOH), and a variable side chain attached to an α-carbon. It is the side chain that determines the type of amino acid and the associated properties. The polymer backbone is formed by the peptide bonds between the amino group and the carboxyl group on two adjacent amino acids.[11]

3.1 DNA-protein interactions

DNA-protein interactions are central to virtually all aspects of genome function and involve several different basic interactions, such as unwinding of the double helix (e.g. helicases), synthesis (polymerases), forming (e.g. ligases) and breaking (e.g. nucleases) the phosphodiester bond, and sequence recognition (e.g. transcription factors). These interactions are governed by a combination of hydrogen bonding, hydrophobic interactions and electrostatic forces.

Proteins can recognize specific DNA sequences through direct contact between amino acids and nucleobases, or, since the local structure of DNA (helical twist, roll and tilt) is sequence dependent, the protein can be structurally compatible with a local structural variant. In addition to sequence-specific binding, many proteins interact with DNA in a non-specific manner, guided by the negatively charged phosphate backbone and the overall topology of the DNA.[18]

3.1.1 DNA-repair

Maintaining genomic integrity is of outmost importance to ensure cell viability. The DNA is under constant stress and suffers damage from environmental causes such as radiation and chemicals, failing endogenous processes or from reactive metabolites. It is estimated that the genome of a single cell is damaged around 70.000 times per day. Most of these damages are single-stranded breaks (SSB) and only a few are DSBs. DSBs are, nevertheless, the most dangerous type of damage and if they are not repaired correctly, they can lead to mutations or rearrangements in the DNA sequence, having potentially pathogenic consequences[11,24], or trigger cell-cycle arrest and cause apoptosis[25].

There are two main routes for repairing DSBs, non-homologous end-joining (NHEJ) and homologous recombination (HR)[26]. NHEJ is typically described as a “quick and dirty” mechanism since it is relatively error prone and repairs the DNA by simply ligating the broken ends without any regard to whether the sequence is correct, potentially causing mutations at the DSB site. HR is, in contrast, a very accurate process with high fidelity to the original sequence, but is only available in the late S and G2 phases of the cell cycle, due to the need of a sister

chromatid from which the correct sequence can be copied[27]. By using an intact copy of the genome as a template, the cell can ensure that there are no mutations introduced at the damaged site.

HR mediated repair of DSBs can be conceptualized as a three-step process: pre-synaptic, synaptic and post-synaptic[28] (**Figure 5**). The pre-synaptic step is initiated by end-resection of the broken DNA ends by the MRN (human) or MRX (yeast) complex, generating stretches of ssDNA. The ssDNA is subsequently bound by replication protein A (RPA) to protect the strands and mitigate secondary structures. The RPA is then replaced by Rad51, forming filaments along the ssDNA. This initiates the synaptic stage where the Rad51-ssDNA filament performs strand invasion and homology search of the repair template. When homology has been established, the process proceeds to the post-synaptic stage. In that last step the broken DNA is re-synthesized using the intact DNA as a template.

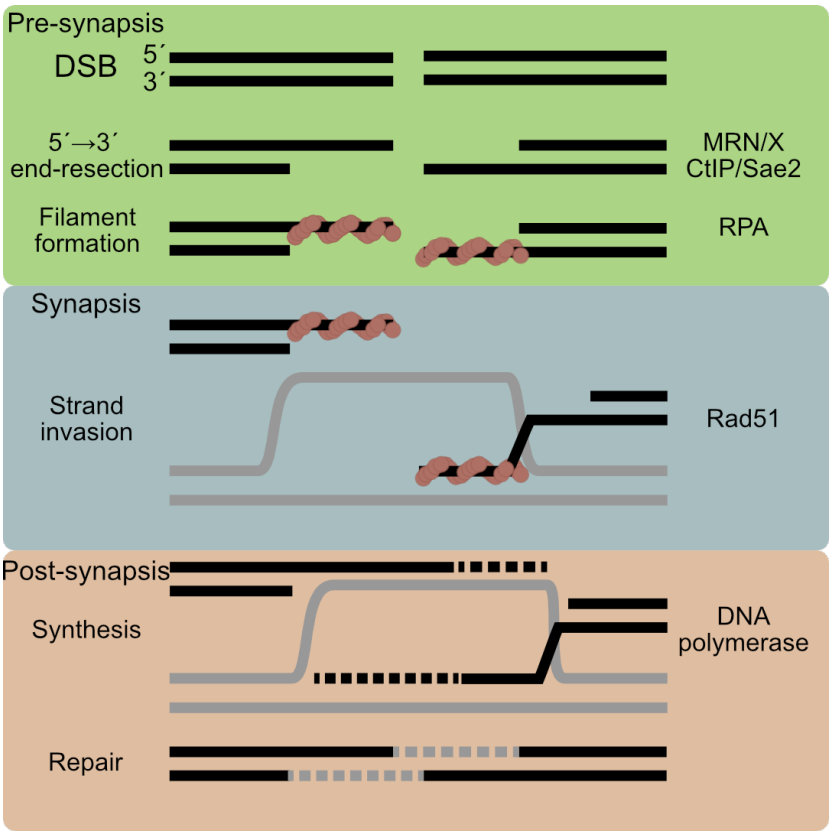

Figure 5. Conceptualized DNA repair by homologous recombination and the three stages together with the main responsible proteins within the human/yeast system. The pre-synaptic stage includes the 5´ to 3´ end-resection by MRN/X and filament formation on the ssDNA strands by RPA. Next RPA is replaced by Rad51 which then initiates the synaptic stage by strand invasion and homology search of the repair template. Once homology has been found the post-synaptic stage starts where the ssDNA strands are filled in by DNA polymerase using the intact DNA as a template.

In human cells, the key signalling and repair proteins MRE11, RAD50 and NBS1, that together form the MRN complex[29], are responsible for initiating DNA repair in the pre-synaptic step and are in turn dependent on CtIP as a co-factor[30]. The MRE11-RAD50 (MR) subcomplex is conserved across several domains of life, whereas NBS1 (with its homologue Xrs2 in *S. cerevisiae*) is conserved only in eukaryotes[31]. The co-factor corresponding to CtIP in yeast is Sae2[31,32]. MRN/X is one of the first protein complexes that is recruited to a DSB, where it initiates repair through ATM/Tel1 activation, DNA end resection, and holding the DNA ends in close proximity for repair[29,33]. Depending on the type of DNA lesion and the cell cycle stage, MRN/X channels the repair to either HR or NHEJ[34,35] in a way that is not completely understood.

The NBS1/Xrs2 protein is the least conserved subunit that markedly differs in structure and functionality between organisms and is also the least understood[36]. While NBS1 is essential for DNA end resection, Xrs2 is largely dispensable[37]. Studies have demonstrated that decreased DNA damage detection in *xrs2Δ* cells can be partially rescued by fusing Mre11 to a nuclear localization signal (NLS). This does not mitigate the loss in DNA repair by NHEJ, indicating the importance of Xrs2 in NHEJ[37,38]. Additionally, Mre11-NLS *xrs2Δ* cells showed significantly decreased DNA-tethering and replication fork stability[37]. The importance of Xrs2 in DNA-tethering is further highlighted by findings showing that MRX has a significant role in increasing the tethering and subsequent ligation of DNA ends by ligase complex Dnl4/Lif1, but yeast Mre11 on its own does not[39]. Interestingly, comparable experiments with human MRE11 and MRN together with Ligase IV/XRCC4 have demonstrated that neither MRE11 nor MRN are sufficient to promote the ligation of broken ends *in vitro*[40]. This, together with results showing that NBS1-deficient cells generate NHEJ events at the same level as wild type cells, but have a significant decrease in HR frequency[41], indicates the lesser involvement of NBS1 in NHEJ and highlights its importance in HR and ATM activation. The original work presented in this Thesis (**Paper I**) focuses on NBS1 and Xrs2 and their putative role in DNA tethering by the MRN/X complex in the pre-synaptic stage of DNA repair.

3.1.2 CRISPR-Cas9

The CRISPR-Cas system has had a revolutionizing impact in biological research, and especially in applications that require sequence specific interactions with DNA. While gene editing is the main example of this type of application, researchers are constantly finding new uses. The CRISPR- (Clustered Regularly Interspaced Palindromic Repeats) Cas (CRISPR associated) system is a programmable endonuclease that utilizes a short RNA molecule, with a part complementary to the target DNA sequence to introduce DSBs at specific sites. The Cas protein complexes with the RNA, making up a ribonucleoprotein (RNP). The complex scans the DNA for a complementary sequence and positions itself on the DNA by hybridizing the RNA to the DNA. The system was first identified in bacteria where it functions as an adaptive immune system. By saving a short sequence, specific for the infectious agent, any future infection can be thwarted by cutting the foreign DNA into pieces[42]. This is analogous to restriction enzymes, but with the added benefit of being adaptable to target any sequence, rather than a hardcoded motif.

Figure 6 depicts one of the most commonly used CRISPR-Cas systems, Cas9 from *S. pyogenes*, and how it is complexed with an RNA sequence, often referred to as the "guide RNA" (gRNA). The gRNA contains a secondary structure for the protein to bind to as well as a complement to the target sequence that allows it to bind to a specific piece of DNA. The target sequence has few limitations except that it is obligatory for it to include a small motif called the protospacer adjacent motif (PAM), which dictates the position of the DSB. By introducing two amino acid substitutions, researchers have developed a Cas9 variant that does not induce DSBs, referred to as "dead Cas9" due to its lack of nucleolytic activity[43] . By coupling this variant to fluorophores, it is possible to get a fluorescent signal correlated with binding to a specific DNA sequence, without inducing damage. This capability makes up a powerful tool for sequence-specific detection or single molecule tracking[44–46] .

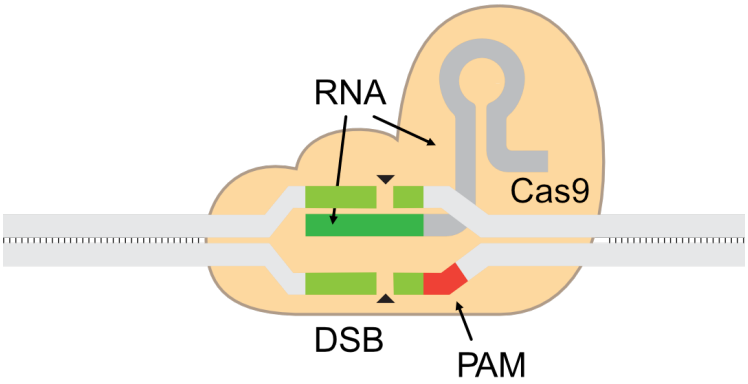

Figure 6. The CRISPR-Cas system from *S. pyogenes* with the Cas9 protein complexed to the guide RNA. The RNA includes the complement (dark green) to the target DNA sequence (light green) and a secondary structure. The obligatory PAM (red) defines where the DSB will be induced. Image retrieved and adapted from "DNA_Repair_after_CRISPR-Cas9_cut"[47] licensed under CC BY-SA 4.0

In **Paper III** of this Thesis, the dead Cas9 variant is utilized to develop a method which can detect single plasmids from DNA-samples without the need of purification steps or amplification. This work aims to address situations where the sample volume is very small or where purification is not possible due to lack of unique features on the target DNA molecule. This method is also well suited for applications where the target DNA contains repetitive sequences or contain a population of concatemers which cannot be readily resolved in methods relying on amplification, such as qPCR.

3.1.3 Genome maintenance

DNA is constantly exposed to internal metabolic processes that include the introduction of DSBs and supercoiling. To cope with this, cells have genomic maintenance systems that are tightly coordinated with transcription and replication. Disruptions in these processes can lead to genomic instability, a hallmark of many diseases, including cancer.

The MYC "protooncogene" is important in many human cancers where it regulates a multitude of cellular processes, either as a "universal amplifier" or as a "super-transcription factor". MYC-driven transcription also reorganizes chromatin and generates topological stress[91]. For MYC to work efficiently as a transcription amplifier, the elevated topological stress accompanied by increased transcription needs to be relieved. In a recent study, Das et al. found that MYC interacts with Topoisomerase 1 and Topoisomerase 2A/B (TOP2A/B) in formations they referred to as "topoisosomes"[92]. They hypothesized that MYC interacts with TOP2A to increase its activity and thus relieve the increased topological stress caused by the amplified transcription levels. DNA topoisomerases are important in genomic maintenance throughout the cell cycle, e.g. by removing supercoils to regulate conformations and removing knots or tangles on the DNA[92,93].

Topoisomerases are found throughout the nuclei and localize to the nucleolus and the nucleoplasm in a dynamic equilibrium[94]. This equilibrium can be shifted upon stimuli such as ATP depletion[95,96], which have brought forward questions regarding potential interaction partners within the nucleolus and how mobile proteins are retained within the compartment. Based on several earlier observations, a study from Jeong et al. established that TOP2A and TOP2B coalesce in biomolecular condensates through liquid-liquid phase separation (LLPS) [97]. This is a transition where a macromolecule is locally concentrated and separated from the bulk to create a membrane-less compartment with locally significantly higher concentration than in the bulk phase. Additionally, they found that condensate formation was stimulated by both DNA and RNA, which increased the fluidity of the condensates.

3.1.4 Liquid-Liquid Phase Separation

The concept of biomolecules organizing in densely packed structures through liquid-liquid phase separation (LLPS) is a way of explaining how cells can regulate processes temporally and spatially[48]. LLPS is a change in local concentration of one or several macromolecules such that they de-mix from the bulk solution and make a separate compartment. What is distinctive of a LLPS compartment is that it does not have a membrane. This allows it to fuse with neighboring condensates or split into smaller droplets. Additionally, biomolecular condensates can be transformed into materials with differing characteristics, such as viscous liquids, gels, and even solid aggregates[49].

Considering phase separation from a simplified thermodynamic perspective, the free energy of the solution needs to be considered. For molecules in solution that do not interact with each other, the free energy as a function of solute concentration (**Figure 7**) is unimodal (red curve). The molecules thus need to disperse homogeneously within the volume to minimize the free energy. If a small localized transient increase in concentration occurs, they will be redispersed by diffusive flux to regain the minimized state. On the other hand, if the solutes interact, the energy landscape changes and becomes multimodal (blue curve) allowing for several minima. The solute may then be separated into multiple concentrations within the volume to minimize the energy, which will create compartments of low and high solute density[48].

Multivalent molecules are molecules with multiple interaction sites to other molecules. Such molecules can be proteins, DNA or RNA, which have several places that can interact with other species or other types of ligands. This allows for larger assemblies to form and from this, phase separation can occur. Multivalency has become a governing explanation how phase separation is regulated as well as for other characteristics like physical properties, composition and biological function in biological condensates[48,50].

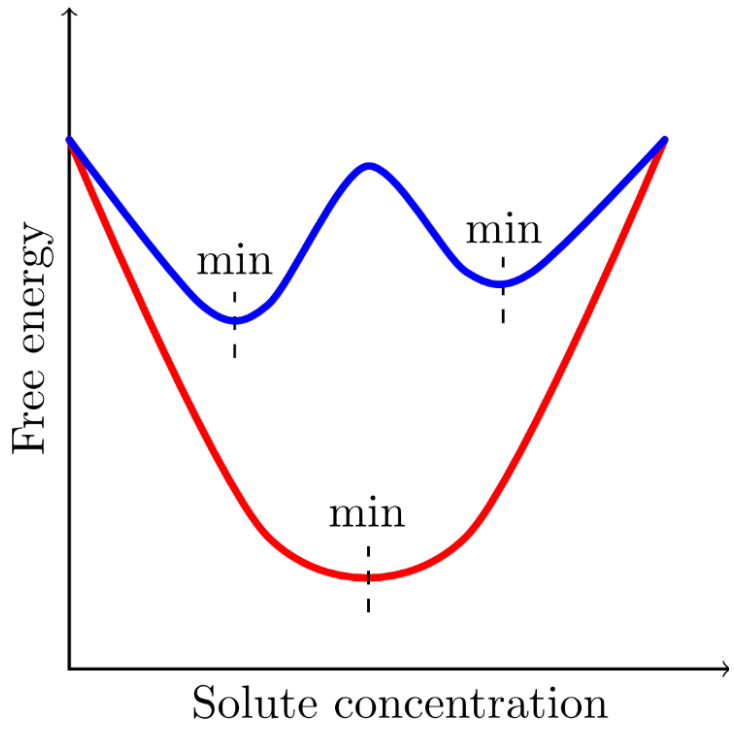

Figure 7. Free energy of a solute in a solvent. A given number of solute molecules that do not interact will be distributed homogeneously so that there is one system concentration. For interacting molecules, the solute will be distributed heterogeneously so that the system will contain multiple concentrations. The red line demonstrates the energy of macromolecules that do not interact. The blue line demonstrates the multimodality of macromolecules that interact.

The high concentrations in the interior of a biomolecular condensate will give obvious advantages for functions that depend on proximity and interaction of individual molecules. This can be enhancement or suppression of reaction rates[50], for example the catalytic activity of a reaction due to co-concentration of substrates and enzymes. An increase in the effective concentration of substrate will drive the reaction towards the product, while the exclusion of substrate can limit a reaction. If a condensate is organized so that inhibitors are excluded, it can temporarily increase the output of a certain reaction. Additionally, if a reaction can utilize several types of substrates, but only one gives the required product, the reaction can be directed to a specific product by exclusion of unwanted substrate.

LLPS plays a central role in the interaction and function of many DNA related processes. A condensate can concentrate DNA-binding proteins, such as transcription factors[51] and DNA-repair proteins[52], acting as a regulatory mechanism or enhancing the probability of productive interactions. Furthermore the, aforementioned, selective inclusion or exclusion of proteins, enables precise spatial and temporal regulation. This dynamic compartmentalization also acts as a buffer towards fluctuations in molecular concentrations, contributing to the robustness of cellular processes[53]. Disruptions in these interactions, whether through mutations or environmental stress, can lead to misregulation and are increasingly associated with disease states[54].

Paper IV investigates the formation of TOP2A condensates and their dynamic interaction with the oncogenic MYC protein, focused on imaging and quantifying DNA-protein complexes immobilized on functionalized glass surfaces. These experiments characterize how MYC influences key properties of TOP2A condensates, such as their size and density. Section 6.4 outlines the scope and main findings of the study, with emphasis on the methods, data, and results generated through my work.

4 Fluorescence

Fluorescence is the emission of light when a molecule relaxes from a higher to a lower energy state and is routinely used as a tool for research within biology. Few biological macromolecules exhibit fluorescence at practically useful levels and thus need to be coupled to fluorescent molecules or fluorescens proteins. The development of bright and specific fluorophores has enabled fluorescence microscopy to become a fundamental tool in many research contexts.

4.1 Physical principle of fluorescence

Light can be described as electromagnetic radiation in the form of a wave with a specific wavelength (λ) and frequency (ν) that travels at the speed of light (c). These characteristics are related according to $c = \lambda\upsilon$. The energy of light is quantized in photons, meaning that it is possible to treat light as particles, rather than as a continuous wave, where each photon carries a set amount of energy. This was proposed by Max Planck in 1900 and the energy of a photon relates to the frequency of the wave through the Planck constant (h) as $E = h\upsilon$, the energy of light is thus dependent on the wavelength[55,56].

The atomic model describes an atom as a positively charged nucleus surrounded by negatively charged electrons. The electrons are not randomly distributed but occupy orbits, which are regions of space where an electron is most likely to be found. The orbits are arranged in shells, and each shell corresponds to a discrete energy level. To jump between shells, an electron needs to lose or gain a specific amount of energy. Atoms and molecules normally exist in a stable ground state. When an atom or molecule absorbs energy (i.e. from light) the electrons jump to a higher energy shell, thus entering an excited state. As the molecule or atom loses the extra energy, the molecule relaxes back to the ground state[55,56].

The process of excitation and relaxation between the singlet ground state (S_0) and an excited state (S_n) is usually visualized through a Jablonski diagram, **Figure 8**. At each electronic state (S_n) the molecule can exist on several vibrational levels. Transition from higher to lower vibrational levels is referred to as internal conversion (IC) and results in the loss of energy as heat. As the molecule reaches the lowest vibrational level of S_1, relaxation to the ground state can occur through either IC, radiative decay (fluorescence) or intersystem crossing (ISC) to the triplet state. Which process that occurs depends on the molecular structure, the temperature and the surrounding environment[57]. When relaxation occurs through fluorescence, the energy is emitted as photons. When a molecule is excited, it will change its dipole moment, causing rearrangement of surrounding solvent molecules. This is referred to as solvent relaxation and reduces the energy of the excited state. This loss in energy causes a shift in the emitted wavelength compared to the wavelength of the absorbed light. This shift is called the *Stoke's shift* and is typically expressed as the difference between the wavelengths where the maximum absorption and emission is observed[58].

Electrons exist in pairs, each one spinning in the opposite direction to the other. ISC occurs when one of the electrons changes spin, making both electrons spin in the same direction. The state formed is called the triplet excited state (T_n) and the relaxation to S_0 can either go via non-radiative ISC and IC or a photon emitting process called phosphorescence. Fluorescence

and phosphorescence differ in two key aspects. The typical time a molecule spends in an excited state before fluorescence (fluorescence lifetime (τ_f)) is in the 10^{-8} s range while the lifetime of phosphorescence is in the milliseconds to seconds range. Since relaxation through phosphorescence entails lower energy levels, the Stoke's shift is also typically larger. Phosphorescence is generally less probable than fluorescence due to the spin-forbidden nature of transitions from the triplet to the singlet ground state.

Absorptivity (ε) and fluorescence quantum yield (Φ_f) are important characteristics for a fluorescent molecule. The absorptivity describes a molecule's ability to absorb light at a certain wavelength and the quantum yield describes the number of photons emitted per absorbed photon, i.e. how much of the absorbed energy that is emitted as fluorescence. The quantum yield can never reach unity due to Stokes' losses[58]. From an experimental perspective, the absorptivity and quantum yield of a fluorophore should ideally be high in combination with a large Stoke's shift. This ensures a bright signal where it is possible to distinguish between excitation and emission light. A short fluorescence lifetime is also beneficial when using fluorescent probes since the shorter lifetime limits the amount of diffusion or interaction and thus increases the spatial information from the signal.

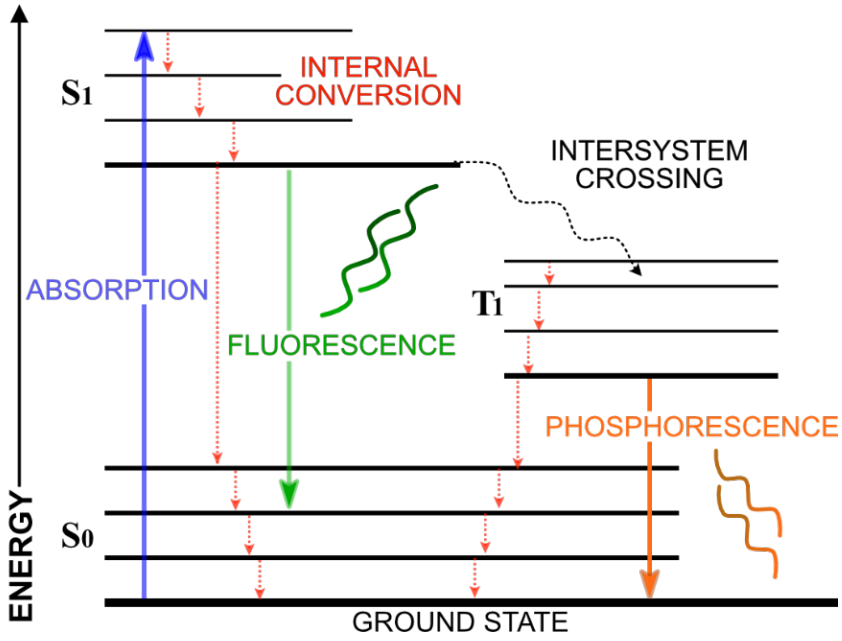

Figure 8. Electronic transitions between ground and excited states visualized as a Jablonski diagram. When an atom or molecule absorbs energy, an electron jumps from the ground state (S_0) to an excited state (here S_1), shown by the blue arrow. The excited electron will mainly lower its energy through non-radiative transitions, either through IC (dashed red arrows) as heat or via ISC (dashed black arrow) to the triplet state (T_1). The electron can also relax back to S_0 through a radiative process, such as fluorescence or phosphorescence, shown by the green and orange arrow, respectively. When relaxation happens through a radiative process, the energy will be emitted as photons at a wavelength corresponding to the energy difference between the excited state (S_n) and the singlet ground sate (S_0).

4.2 Fluorophores

Fluorophores are molecules that have an inherent ability to emit photons following absorption of light. They are typically divided into two categories: intrinsic and extrinsic fluorophores. Intrinsic fluorophores occur naturally and are inherently incorporated into the biomolecule under study. A common example is the amino acid tryptophan (Trp (W)). While intrinsic fluorescence is convenient, many biomolecules (including DNA) do not exhibit fluorescence or have intrinsic fluorophores. Instead, a non-fluorescent molecule can be fluorescently labeled

by coupling it to a fluorophore. Extrinsic fluorophores are thus molecules that can be used as fluorescent labels by binding them, covalently or non-covalently, to the target of interest. There are countless options of fluorophores, covering a large part of the visible light spectrum. **Figure 9** shows the excitation and emission spectra for three common extrinsic fluorophores and how they relate to each other. The three fluorophores showcased here were used in the work presented within this Thesis.

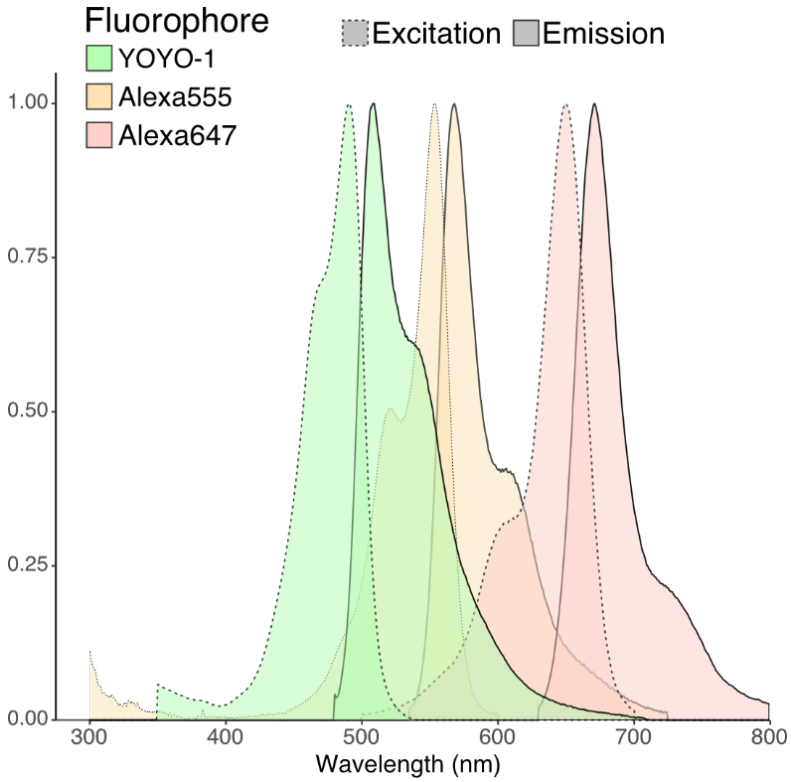

Figure 9. Excitation (dotted lines) and emission (solid lines) spectra for YOYO-1, Alexa555 and Alexa647. YOYO-1 is a common DNA dye that binds specifically to dsDNA with very high affinity. Alexa555 and Alexa647 are common fluorophores used as extrinsic labels on proteins or to covalently label DNA. The combination of these three fluorophores is suitable when multiple labels are needed on the same target or when looking at heterogenous samples where multiple targets need to be detected, since they are spectrally separated.

4.2.1 DNA-specific fluorophores

Since DNA does not possess any intrinsic fluorescence, an extrinsic fluorophore is needed for fluorescence-based studies. There are multiple types of DNA-specific fluorescent dyes available with differing binding modes, affinities and spectral properties. The work within this Thesis utilizes the homo-dimeric derivative of oxazole yellow (YO), YOYO-1. **Figure 10A** shows the YOYO-1 structure. YOYO-1 is practically non-fluorescent in solution, whereas it shows a ~400–1000-fold increase in emission quantum yield upon intercalation with dsDNA[59]. Intercalation is the insertion of a ligand into the double-helix in the space between the stacked nucleobase-pairs. YOYO-1 primarily binds to DNA by bis-intercalation, inserting both YO units (**Figure 10B)**, at bp:dye ratios above 6:1, with an increasing contribution from external binding at higher dye loading[60]. The significant difference in quantum yield between bound and unbound dye is attributed to decreased rotational freedom in the bridge between the benzoxazole and quinoline moieties of the YO unit[61] upon intercalation. One YOYO-1 molecule occupies ~3-4 bp[9,62] and the intercalation of YOYO-1 into DNA affects both the mechanical and structural properties of DNA[62–66]. The intercalation of YOYO-1 elongates the DNA by 0.51 nm and unwinds the helix 24 degrees per bound molecule[64].

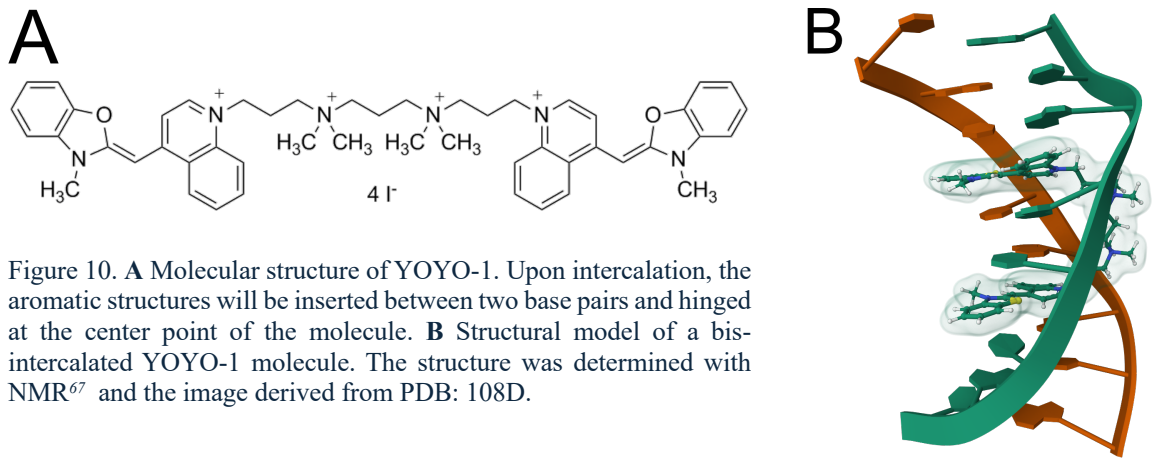

Figure 10. **A** Molecular structure of YOYO-1. Upon intercalation, the aromatic structures will be inserted between two base pairs and hinged at the center point of the molecule. **B** Structural model of a bis-intercalated YOYO-1 molecule. The structure was determined with NMR[67] and the image derived from PDB: 108D.

4.3 Fluorescence microscopy

An optical microscope, or light microscope, uses a set of lenses to generate magnified images of an object with visible light. The lenses are contained within an objective and collect the transmitted, reflected or emitted light from an illuminated sample.

A fluorescence microscope is specifically built to excite and collect the emission from fluorescent samples, **Figure 11A**. The main components of a fluorescence microscope comprise a light source (poly- or monochromatic), excitation filter, dichroic mirror, objective, emission filter and detector. If the light source is monochromatic, e.g. generates light of a specific wavelength (such as LEDs or lasers), the excitation filter can be removed, else unwanted wavelengths need to be filtered out. The dichroic mirror is a color selective mirror that reflects light below and transmits light above a certain wavelength. Typically, this threshold is set so that the excitation light is reflected, and the emission light is transmitted. Before the emitted light reaches the detector, it is passed through an emission filter that removes any backscattered light from the sample and light outside of the filters' bandwidth. This ensures that only light originating from fluorescent components of the sample reaches the detector.

As light travels through the apertures of the lenses, diffraction will occur. When a point emitter is imaged, such as a fluorophore, it will have an intensity distribution with a central peak and rings at increasing distances, due to the diffraction, called a *point spread function* (PSF). The spatial resolution (d), e.g. the minimum distance needed between the centers of two PSFs to differentiate between them, of a light microscope is limited by the wavelength of light (λ) and the numerical aperture (NA) of the objective according to the *Rayleigh criterion*: $d = 0.61\frac{\lambda}{NA}$, **Figure 11B.** NA is a dimensionless characteristic number that describes the range of angles at which an objective can transmit or accept light, defined as $NA = n \sin\theta$, where n is the refractive index and θ the half-angle of the light cone entering or exiting the objective. A similar criterion was developed by E. Abbe, referred to as the *Abbe diffraction limit*, which sets the minimal distance to $d = 0.5\frac{\lambda}{NA}$. Considering these restrictions, an objective with NA=1.46 and the typical emission from YOYO-1 (509 nm), gives a theoretical spatial resolution of ~175-212 nm. This means that it is not possible to discern two separate YOYO-1 point emitters with centers at a distance below ~175-212 nm.

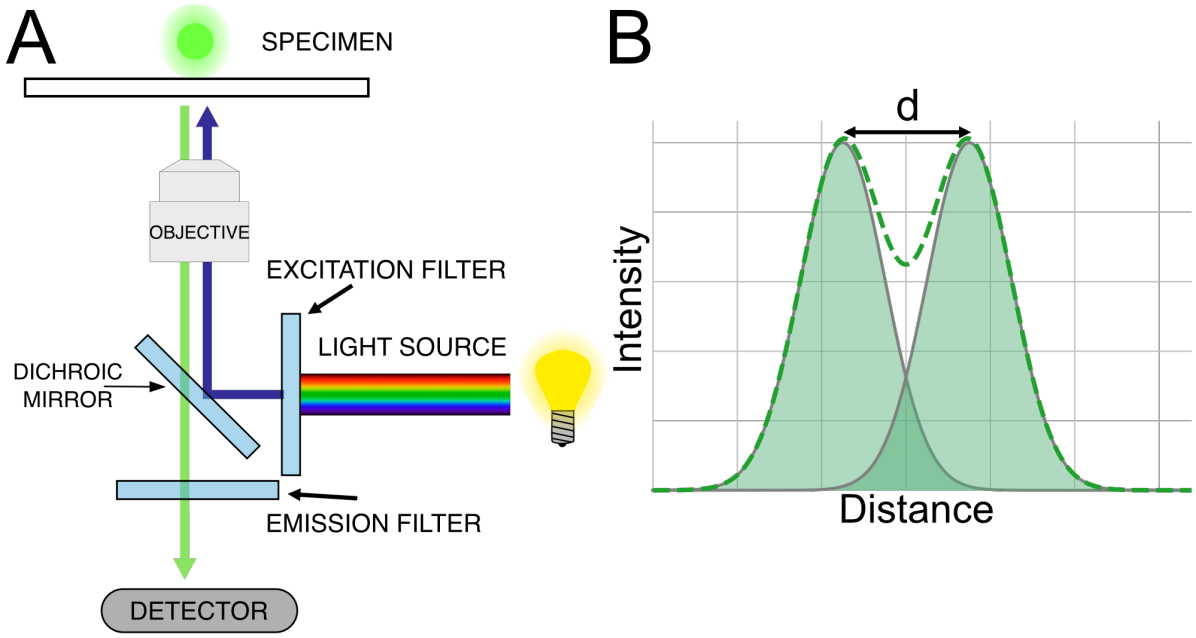

Figure 11. **A** Schematic of the main components of a fluorescence microscope and the light path from light source to detector. The light source generates light at a desired intensity and wavelength. In setups where the generated light is non-monochromatic the light passes through the excitation filter to remove unwanted wavelengths. The dichroic mirror reflects only specific wavelengths to the specimen through the objective to excite the fluorophores. The emitted fluorescence then travels back through the objective, the dichroic mirror and the emission filter, which removes any back scattered light from the specimen, and on to the detector. **B** Visualization of two neighboring PSFs and the typical distance defined by the Rayleigh-Abbe limit that describes the diffraction limited spatial resolution of a fluorescence microscope.

5 Studying single biomolecules

Single molecule analysis has the potential to uncover the stochastic and heterogenous behavior of individual molecules within a sample. This heterogeneity is typically lost in bulk measurements, where the result is an average of the entire population. Single molecule analysis is a wide field that encompasses many different methods and techniques to isolate and observe the behavior and interaction of single molecules. Beside the approaches used for the work presented within this Thesis, which is based on fluorescence microscopy, there are multiple methods and strategies to detect and manipulate single DNA molecules. What is common for most single molecule methods involving DNA is the need to extend the DNA from its coiled state into a linear conformation to visualize static and dynamic interactions with other molecules. This can be done by anchoring the ends of the molecule and actively applying force[8] , by immobilizing the molecules while applying a hydrodynamic flow[68] or confining them in nanofluidic channels[69] , as described in this Thesis.

5.1 Stretching DNA in nanochannels

In many single-molecule methods the DNA ends typically need to be anchored so that the molecule can be stretched out and imaged. For example, is it possible to immobilize and stretch a DNA molecule by anchoring one end on a ledge and apply a hydrodynamic flow, stretching the molecule in the direction of the flow, typically referred to as “DNA curtains”. Similarly, in force spectroscopy, where the molecule needs to be anchored to surfaces that are pulled apart to exert force on the molecule. The downside is that the DNA cannot be considered to be free in solution and that they require the ends of the DNA to be modified and essentially blocked from interactions. By confining the DNA in a geometry smaller than its hydrodynamic radius (R_G), the molecule will spontaneously adopt an extended conformation. Thus, the molecule is free to move, and the ends can be kept free and unmodified, allowing the molecule to interact with itself and neighboring DNA molecules, as well as moving freely within the limits of the confinement. This is particularly advantageous when studying end-specific DNA processes, such as DSB repair. **Figure 12** shows a schematic of a nanofluidic chip used for the work presented in **Papers I-II**. The design is comprised of two pairs of loading wells that are connected by microchannels. The microchannels are in turn bridged by an array of nanochannels. The DNA is manipulated and transported within the system by flow. By applying pressure, via nitrogen gas, on one or more of the wells the liquid can be driven in a desired direction. Since confinement is energetically unfavorable for a DNA molecule, it will experience an entropic barrier at the nanochannel entrance, i.e. the DNA will not spontaneously enter nanochannels. To overcome this barrier a small amount of force needs to be applied by creating a flow across the nanofluidic array. This will essentially push the DNA molecules into nanochannels forcing them to extend as they enter.

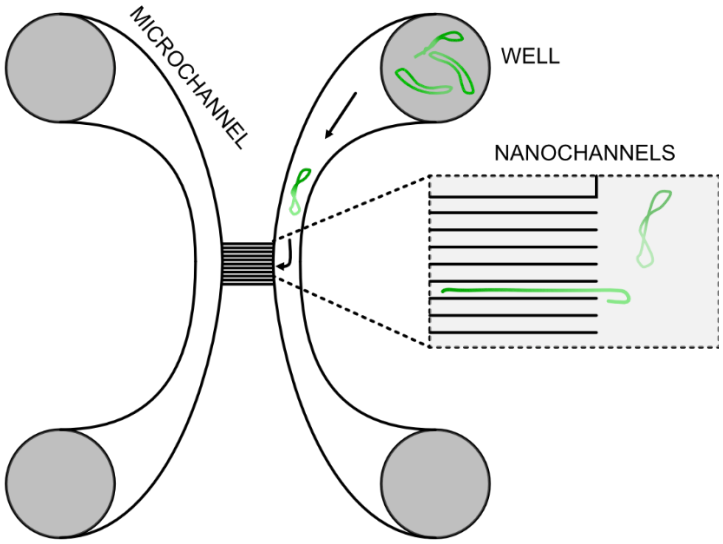

Figure 12. Schematic (not to scale) of a nanofluidic chip used for the work presented in this Thesis. The chip consists of four loading wells connected in pairs by microchannels. The two microchannels are bridged by an array of nanochannels, creating a system in which single DNA molecules can be manipulated, isolated and observed. The chip is operated by applying pressurized nitrogen gas on one or several of the wells to create a flow in the desired direction. The DNA is maneuvered into the nanochannels by applying pressure on both wells on one side of the nanochannel array to direct flow through the nanochannels. This forces the DNA to extend from its coiled conformation as it is pushed into the channel. Once the DNA has entered the nanochannel, the flow is stopped, trapping the DNA in an extended conformation, allowing for observation and data collection.

5.2 Immobilizing biomolecules on glass

A common and simple way to stretch a DNA molecule is use the negative charge of the DNA phosphate backbone to immobilize DNA on a positively charged surface. To stretch the DNA, it needs to be deposited while it is pulled across the surface. This can be done by sandwiching two surfaces and introducing a small amount of sample at one of the edges. The liquid will be pulled across the surface by capillary forces and the DNA will be deposited at the air-liquid interface in the travel direction. **Figure 13** depicts the two main steps included in the protocol used for the work presented in this Thesis (**Papers I, III-IV**), where the glass surface is functionalized by a monolayer of silanes ((3-Aminopropyl)triethoxysilane (APTES) and allyltrimethoxysilane (ATMS)). The monolayer provides a hydrophobic surface with a net positive charge, **Figure 13A**. As the liquid is introduced between the functionalized coverslip and a microscope slide, **Figure 13B**, the DNA is deposited and stretched as the liquid is pulled across the surface by capillary forces.

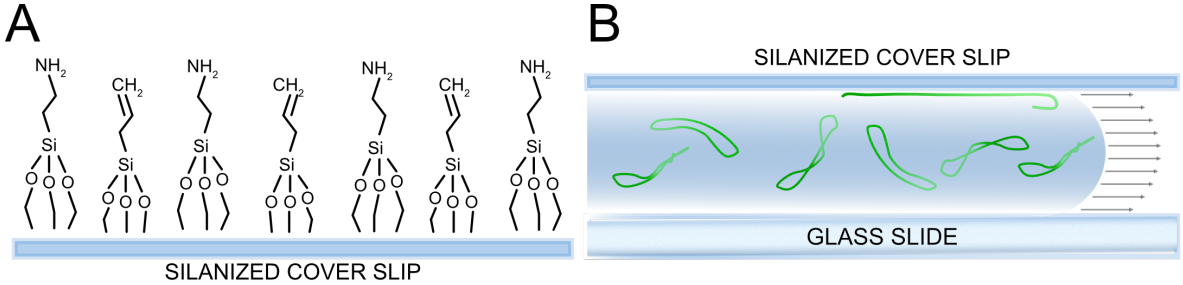

Figure 13. **A** Glass cover slip functionalized with a mixture of silanes with amine and vinyl terminal groups. The amine groups are positively charged. **B** DNA molecules are stretched by capillary force between the silanized cover slip and a glass slide.

6 Original work

In this chapter the original work in the appended papers (**Papers I-IV**) is summarized and discussed. The work has focused on single molecule analysis of DNA-protein interactions using fluorescence microscopy and is centered around two main methodologies. The first method uses nanofluidics to study DNA-protein interactions, and in particular DNA repair proteins. The second method involves immobilization of DNA-protein complexes on a glass surface. **Paper I** investigates the DNA-repair protein complex MRN/X, from human and yeast, and the role of the individual components in DNA end-joining. **Paper II** is focused on the influence of Mg^{2+} and ATP on the binding of YOYO-1 to DNA and the potential impact this can have on studies of active protein processes in nanofluidics. **Paper III** demonstrates a quantitative method where Cas9 is used to identify sparse DNA molecules in heterogenous samples and size them based on their fluorescence intensity. **Paper IV** is a study on how the oncogenic protein MYC is involved in the formation and organization of TOP2A biomolecular condensates where my contribution encompasses quantitative fluorescence microscopy data showing the colocalization of the proteins within the biomolecular condensates and the change in characteristics of the condensates upon co-condensation with MYC.

6.1 Nanofluidics for studying end-specific DNA-protein interactions

In order to study interactions occurring along a DNA molecule, the molecule needs to be extended from its coiled state. For a DNA molecule to spontaneously stretch to an extension close to its contour length (L), the dimensions of the nanochannel must satisfy $D_{AV} < R_G$. The channels in the nanofluidic chip (schematic in **Figure 12**) that was primarily used for the experiments in **Papers I-II**, had a width of 100 nm and a depth of 150 nm, equating to D_{AV} = 122 nm. In **Paper II** an additional chip design was used, with the same channel dimensions and basic principle, but containing ten individual nanofluidic arrays. This chip design is referred to as the "multiplex chip" since it is possible to image ten different samples in parallel[70]. For single molecule studies, DNA from the lambda-phage (λ-DNA) is typically used as a model DNA. The lambda-phage genome is 48 502 bp long (L ~ 16.5 μm) and can be produced in large quantities, making it readily available. After being extracted from its host cell it has 12 bp complementary single stranded ends, making the ends easy to modify as well as enabling production of concatemers and circles. Since the R_G of λ-DNA is ~500 nm at 0.1 M salt, the DNA molecules will extend upon confinement in the nanochannels used. In the work presented in this Thesis, YOYO-1, introduced in section **4.2.1** and **4.3,** has been used to visualize the DNA. Other DNA specific alternatives such as 4′,6-diamidino-2-phenylindole (DAPI), Thiazole Orange Homodimer (TOTO-1) or SYTOX Orange, each covering different wavelengths, are available but do not provide the same contrast as YOYO-1.

6.1.1 Information on DNA-protein interactions is inferred from fluorescence and molecule conformation

The main benefit of the nanofluidic system is the possibility to study DNA substrates with free ends. This allows proteins to interact with the DNA without potential perturbations related to DNA tethers and, most importantly, it allows ends of DNA molecules to meet, as they would be able to do in the cell. In the work presented within this Thesis, all DNA-protein complexes have been studied at equilibrium, meaning that the protein has been allowed to bind to the DNA prior to the introduction of YOYO-1 to the DNA. This approach minimizes the potential perturbation and bias introduced by the intercalation of YOYO-1 to the DNA.

When studying DNA under confinement, the DNA is usually imaged over time and the recorded images are visualized as kymographs. A kymograph is essentially a representation of a DNA molecule for each time-point stacked to form a two-dimensional image with time on one axis and distance on the other. The typical appearance of a fluorescently stained λ-DNA molecule (**Figure 14A**) shows an even emission intensity along the entire DNA molecule. As mentioned, the ends of λ-DNA can be joined to form circles (**Figure 14B**) or concatemers (**Figure 14C**), where multiple DNA molecules have been connected by the ends to generate longer complexes.

In the work presented in **Paper I** the studied proteins are not fluorescently labeled. This means that there is no direct detection of the proteins acting on the DNA. Instead, any evidence of DNA-protein interactions must be deduced through the fluorescence emission intensity along the DNA molecule (from bound YOYO-1) and the conformation of the DNA molecule. The presence of proteins may result in the formation of more complex conformations such as knots, interlocked circles or other structures involving multiple molecules, **Figure 14D**. By comparing data from λ-DNA with and without proteins present, it is possible to get information on the interaction between the proteins of interest and the DNA. This information can consist of changes in extension, changes in the relative occurrence of circles or concatemers, introduction of complex structures (**Figure 14D)** or, in cases of sequence specific binding[71], variations in the fluorescence intensity along the DNA.

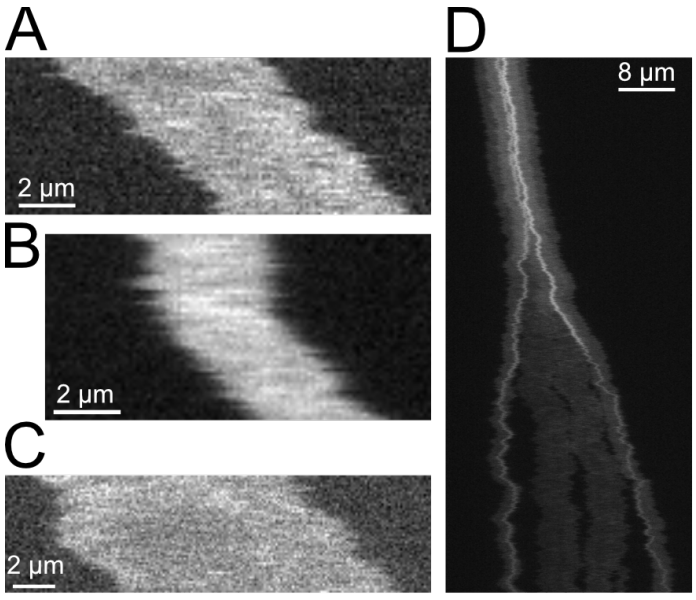

Figure 14. **A** Representative kymograph of a YOYO-1 labeled monomeric λ-DNA molecule under confinement. The molecule is imaged over time, each row in the kymograph represents one time-point and shows the extension and intensity of the molecule. **B** Representative kymograph of a circularized λ-DNA molecule, with shorter extension and higher intensity compared to the linear molecule in **A**. **C** Representative kymograph of a concatemeric λ-DNA molecule, with a longer extension compared to the monomeric molecule in **A**. **D** Kymograph of a YOYO-1 labeled λ-DNA molecule in the presence of human DNA-repair complex MRN with higher order structures induced by protein interaction.

6.1.1.1 Extension and thermal energy guides molecule categorization

The extension of naked DNA in nanofluidic channels is mainly determined by the channel dimensions and the ionic strength of the surrounding buffer[21]. Under confinement, thermal energy will give rise to fluctuations of the molecule extension. This can be quantified by the standard deviation of the extension. A circular molecule will experience a higher degree of confinement due to its double-folded nature and the self-avoidance between the chains, which will lower the flexibility of the polymer, resulting in lower thermal fluctuations. This leads to that circular molecules will have a lower standard deviation of the extension[72]. This characteristic can be used to deduce the conformation of each molecule; by plotting the standard deviation against the extension, it is possible to find clusters of molecules with similar characteristics and get a quantitative measure on the abundance of different conformational states[71].

To visualize and compare the properties between individual molecules in a large data set it is useful to generate a scatterplot with a variable of interest on each axis. **Figure 15A** reports a typical scatterplot constructed from the extension and standard deviation of bare λ-DNA. Each point represents an individual molecule, and by the way molecules cluster it is possible to determine their conformation and categorize them accordingly. The large central cluster (yellow) is the linear λ-DNA with an average extension of ~6.5 μm, whereas the smaller (red) cluster down to the left consists of the few circles that were found in the bare λ-DNA sample. To find the conformational state of each molecule the scatter plot is crudely divided according to each major peak in the distribution along each axis. The full width at half-max (FWHM) of each peak on each axis was used to define a bounding box within which the majority of the molecules within a category were expected to be found. Since the extension of a circular molecule is expected to be half of a linear molecule it is easy to assign each cluster to a conformational category. Concatemers are defined as molecules with an extension higher than max_{EXT} of the linear cluster. To assign a category to each molecule Euclidean distance hierarchical clustering was used to generate 100 small clusters. If > 2/3 of the cluster members were situated within a bounding box, the molecules were assigned accordingly.

Figure 15B reports the corresponding relative abundance of each category. From the data in **Figure 15B** it is possible to calculate the "end-joining"-ratio (EJ-ratio). The EJ-ratio is a composite measurement of the formation of both circles and concatemers. The EJ-ratio indicates the overall probability for λ-DNA ends to come in sufficient proximity to hybridize. The individual numbers of P_{conc} (teal bar) and P_{circ} (orange bar) report the likelihood for an end-joining event to result in a circle or concatemer. The results in **Figure 15C** show that there is a 5% chance of spontaneously joining two ends in a bare λ-DNA sample. For each end-joining event, there is 44% chance of it being an intramolecular event (circle formation) and a 56% chance for it to be an intermolecular event (concatemer formation).

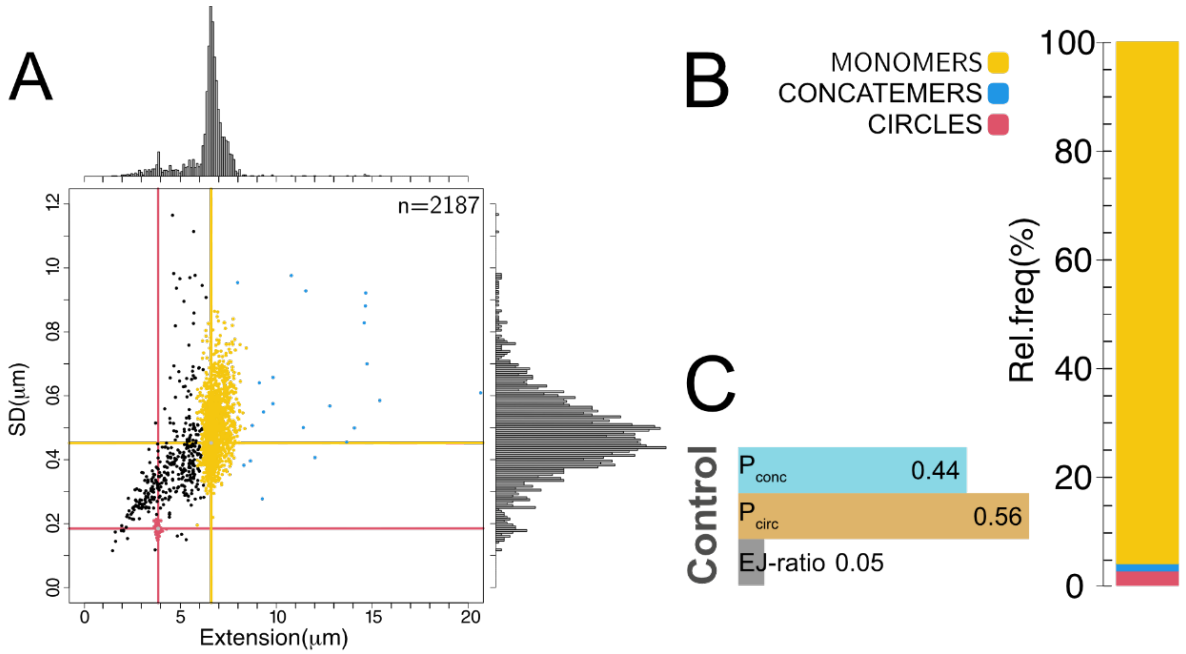

Figure 15. **A** Scatterplot from a control sample of λ-DNA showing the standard deviation of the extension versus the extension. Each point represents a molecule, and the colors represent molecules that have been categorized as belonging to one of the conformational populations. Black points are molecules that fall outside the categories due to being e.g. fragmented. **B** Relative abundance of the different conformational populations found in a λ-DNA control sample. **C** "End-joining" ratio for a control sample of λ-DNA. The EJ ratio is the relation between the number of observed monomers and the number of end-joining events, where the number of end-joining events is based on the number of circles plus the number of monomers found in concatemers. A low ratio indicates a low number of "end-joining events" per observed monomer, i.e. the likelihood of a monomer to be involved in an "end-joining event", either circularization or concatemerization.

6.1.2 Analysis of cyclisation and concatemerization by Xrs2/NBS1 highlights their important role in DNA-tethering by the MRE11-RAD50 complex

Spontaneous cyclization or concatemerization of bare λ-DNA is typically rare, as discussed in the previous section. A spontaneous end-joining event relies on the ends coming in close proximity to hybridize and is thus limited by the diffusion of the DNA ends. Since DNA is a self-avoiding polymer, the ends do not diffuse freely. Data have shown that a typical cyclisation rate for λ-DNA at room temperature is 4/day[73], but is increased greatly by introduction of multivalent condensing agents that compact the molecule, bringing the ends in close proximity[74]. This makes it possible to characterize the end-bridging interactions of a protein by comparing the formation of circular or concatemeric λ-DNA between samples with and without a protein.

In **Paper I** the main goal of the study was to characterize how the MRN/X complex tethers and binds to DNA and to determine the contribution of the individual subunits. Additionally, the study aimed to further explore and demonstrate the use of nanofluidics and conformational analysis within the context of DNA-protein interactions. The MRN/X complex consists of three separate subunits (MRE11, RAD50 and NBS1/Xrs2) and is involved in DNA-repair in both humans and yeast. The complex has an important role in the early processing of broken ends and is involved in the channeling of DNA-repair into the different pathways, NHEJ or HR. The main experimental results from imaging of DNA-protein complexes in nanochannels is the conformational fractions describing the relative abundance of circles and

concatemers for each condition and the end-joining ratio (described in section 6.1.1) derived from this data (see appended **Paper I**, **Figures 3-4**). The study demonstrated that, while MRN primarily bridges DNA in an intramolecular manner, leading to mainly formation of circles, MRX primarily promotes the hybridization of ends from two different DNA molecules, leading to a larger fraction of concatemers. The tendency for MRN to form circles is in line with existing research showing that MRN is essential for HR. The bridging of intramolecularly distant parts is analogous to the tethering of parallel strands during the pre-synaptic and synaptic stages of HR (**Figure 5**). In comparison, the tendency of MRX to promote concatemerization rather than circularization could be considered as evidence for a more end-specific binding by MRX where it does not primarily bridge intramolecularly distant parts but rather form synapses between broken ends, analogous to the role of MRX in NHEJ. Interestingly, the difference between MRN and MRX can be partially explained by the ways that NBS1 and Xrs2 subunits bridge DNA, as will be further discussed in the next section.

While it is evident that the observed change in conformational states is in direct correlation with the presence of proteins, the described methods need to be interpreted and used as an analogue to cellular processes, rather than a direct representation, to better understand the DNA bridging and end-joining activities of the proteins of interest.

6.1.2.1 MRN and MRX exhibit distinct DNA tethering capacities attributed to NBS1 and Xrs2

The study included the three samples for each species. The human proteins MRN, human MR (hMR) and NBS1, and the respective homologues MRX, yeast MR (yMR) and Xrs2 from yeast (*S. cerevisiae*). Each protein (or protein complex) was imaged at three different concentrations, with data demonstrating that all included conditions show end-joining activity. **Figure 16** reports the EJ-ratio for each protein and concentration. MRN and MRX have comparable EJ-ratios, i.e. they bring ends together at a similar efficiency, but they do it differently. MRN has a very high likelihood for circle formation while MRX primarily promotes concatemer formation. Both hMR and yMR showed very low EJ-ratios at comparable concentration, but experiments (data not shown) at increased concentrations of hMR and yMR, revealed that both human and yeast protein subcomplexes have end-joining activity. This indicates that NBS1 and Xrs2 are dispensable for end-joining to occur, but contribute greatly to the end-joining by MRN/X. Looking at the effects of NBS1 and Xrs2 alone, two important insights emerge. The EJ-ratios and the types of end-joining event differ greatly. Compared to NBS1, Xrs2 shows a near double EJ-ratio at comparable concentrations and a near inverse relation in P_{conc}/P_{circ}. The doubled EJ-ratio would be an indication of Xrs2 having relatively higher intrinsic DNA binding compared to NBS1. Considering the strong tendency for Xrs2 to promote concatemer formation it could be attributed to the specific binding to duplex-single strand DNA junctions[75] which would direct Xrs2 to the ends of the λ-DNA. This could in turn alter the diffusion of the ends, making it more likely for them to encounter an end from a neighboring molecule, thus favoring the formation of concatemers. This contrasts with the end-joining by NBS1 which suggests intramolecular bridging. The results for NBS1 and Xrs2 demonstrate that they both play an integral part in the DNA bridging by the full MRN and MRX complexes, respectively, as well as have distinct DNA-binding capacity on their own. It is interesting to notice that neither NBS1, Xrs2 nor h/yMR show the same degree of bridging as the full MRN and MRX complexes. This suggests that there is a synergistic effect when NBS1 and Xrs2 bind to their respective MR partners.

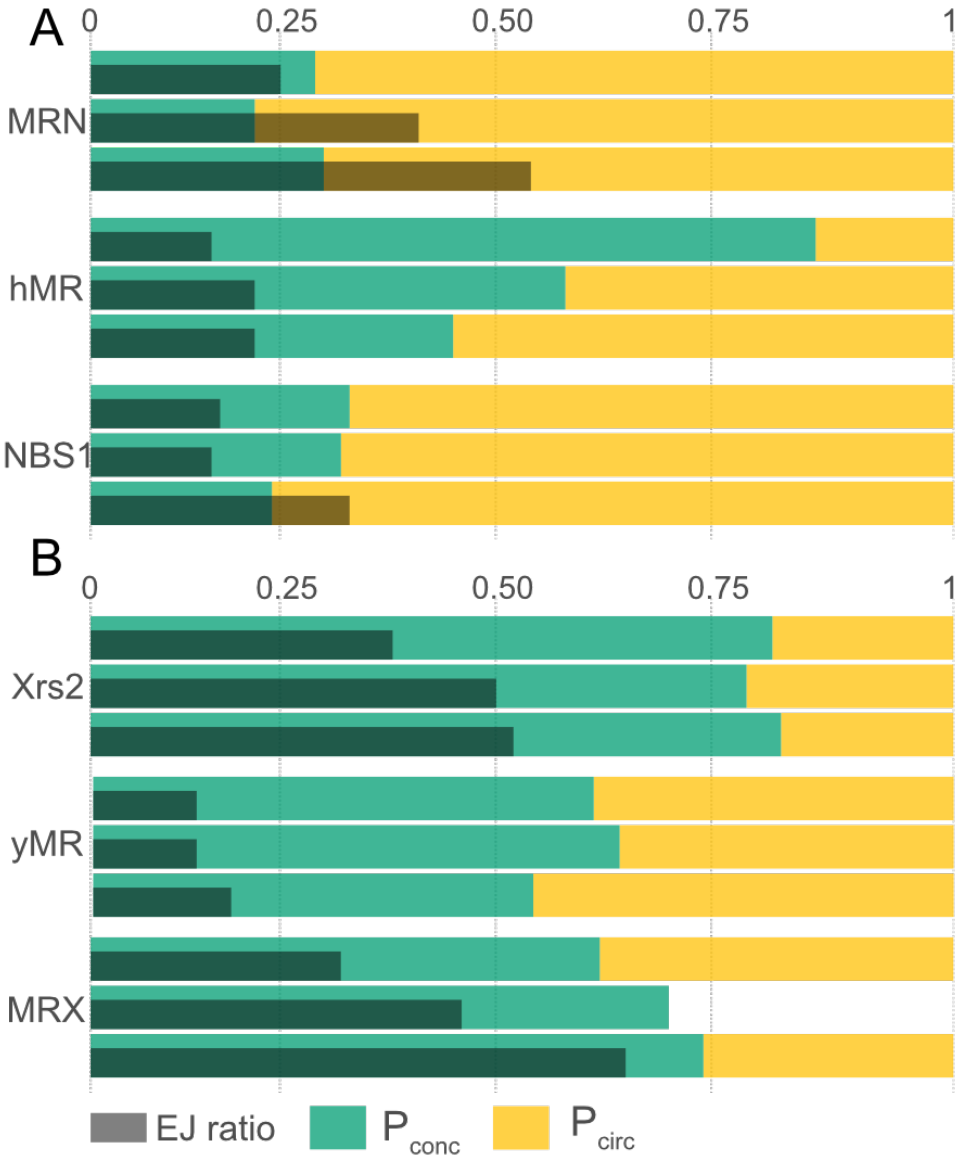

Figure 16. EJ-ratio and the likelihood for an end-joining event to generate a circle (Pcirc) or a concatemer (Pconc) for **A** MRN (80,60,40 nM), hMR (80,60,40 nM) and NBS1 (80,40,20 nM) and **B** MRX (80,60,40 nM), yMR (80,60,40 nM), and Xrs2 (80,60,40 nM). All data is based on three or more replicates.

6.1.2.2 Aptamer-based protein labelling underscores nanofluidic observations

To further validate the observations in the nanofluidic setup the DNA-protein complexes were immobilized on functionalized glass slides. This provided two benefits: First and foremost, it allowed the use of aptamer-based protein labelling to confirm the binding of the proteins to the DNA. Additionally, it provided a way to investigate the DNA aggregates that were too large to enter the nanochannels and could not be included in the conformational analysis.

Aptamer-based labelling is a way to add an extrinsic fluorescent label to proteins that are either difficult to label with other methods or when the yield of purified protein is limiting. The aptamer used was a 6xHis-tag specific aptamer with a biotin-tag. This aptamer was bound to streptavidin-conjugated QDOTS and added to the sample after the proteins had bound to the DNA. **Figure 17** shows the results of such experiments for all the included proteins. For MRN many of the DNA-protein complexes had a brighter, slightly bulbous end or ends with a

stretched-out part, like a tail. The protein in these formations were primarily found at the bulbous ends indicating that the protein was bridging distant parts of the molecule that was unraveled upon deposition. In addition, small, compacted molecules, which can be interpreted as single λ-DNA molecules held together by one or several proteins, was observed (**Figure 17A:** top row). In contrast, MRX generated very large complexes when bound to λ-DNA, with the QDOTs colocalizing in the center of the complex, indicating a high concentration of proteins in the center (**Figure 17B:** top row). In experiments with h/yMR, a majority of the molecules were single monomeric λ-DNA or small concatemers with indications of MR-proteins bound at the ends (**Figure 17A-B:** second row, left column). The h/yMR samples differed in that yMR also showed an appreciable number of examples where the protein, in addition to the localization on the DNA ends, could be found along the length of the DNA, whereas hMR almost exclusively was found at the ends of the DNA. When imaging NBS1 bound to DNA, many of the observed DNA-protein complexes were comparably small and condensed (**Figure 17A:** second row, right column). Experiments with Xrs2, on the other hand, showed very large DNA complexes that consisted of a large number of λ-DNA molecules with a large center of proteins (**Figure 17B:** second row, right column), similar to the observations for MRX.

6.1.2.3 Data from aptamer-based experiments corroborates observations from nanofluidics

The results from the aptamer-based imaging experiments align well with the observations made in the nanofluidic setup. While the EJ-ratio measurements demonstrated that MRN and MRX have comparable end-joining efficiencies, the imaging highlights further that they do so through distinct outcomes. The smaller condensed structures observed for MRN is consistent with the interpretation from nanofluidic experiments. In contrast, are the protein-dense aggregates formed MRX more indicative of concatemer formation. The low EJ-ratios observed for hMR and yMR were reflected in the imaging data, where most DNA molecules appeared as monomers or small concatemers with protein localized at the ends. The formation of small, condensed complexes, by NBS1 is consistent with intermolecular bridging and the observations mirror the results from MRN. Similarly to MRX did Xrs2 generated large, multi-molecular assemblies with central protein localization. These observations reinforce the conclusion that NBS1 and Xrs2 contribute significantly to the bridging activity of the full MRN and MRX complexes.

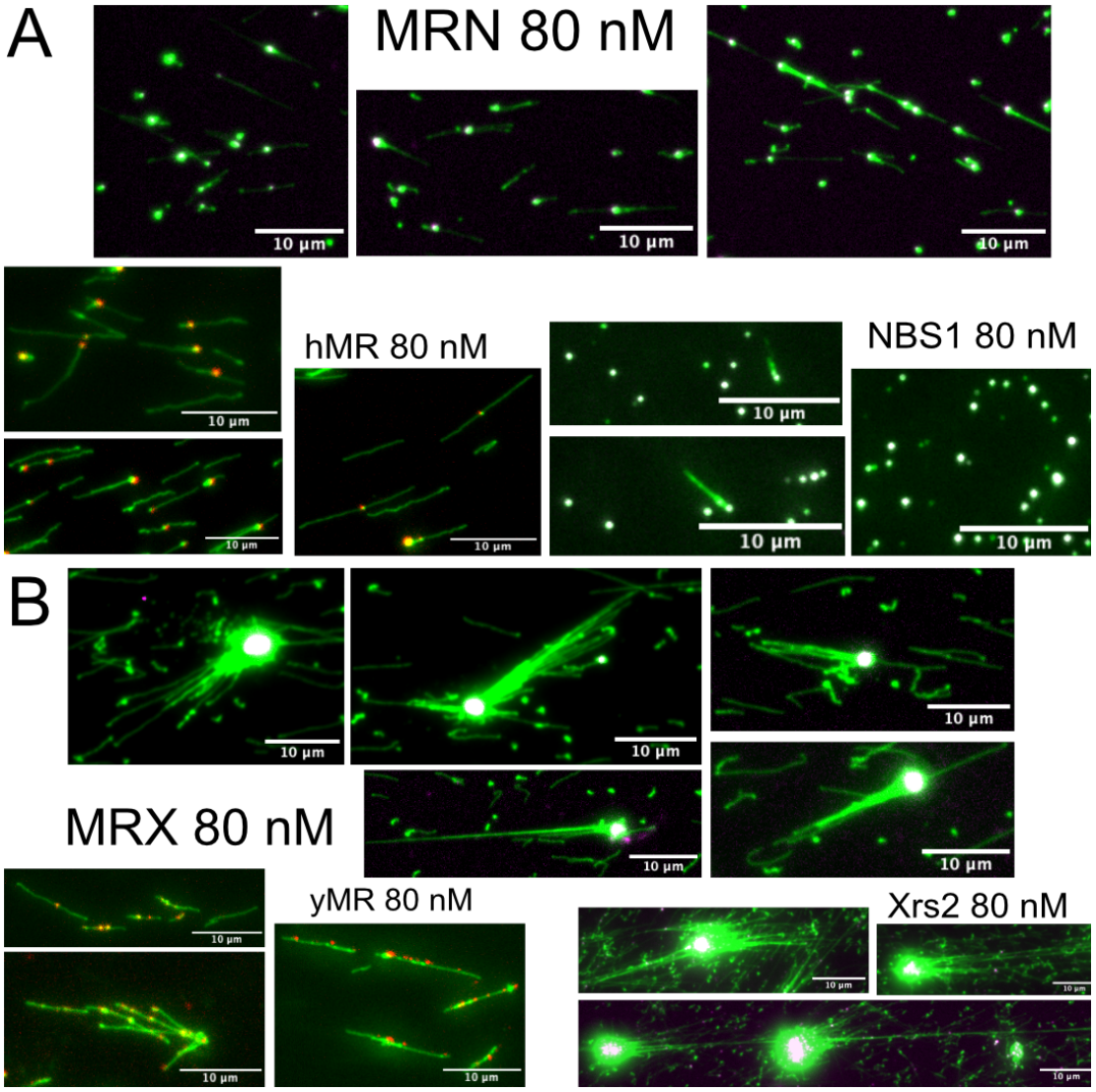

Figure 16. **A** MRN, hMR and NBS1 bound to λ-DNA and labeled with aptamer coated QDOTS targeting the 6xHIS-tag used for protein purification purposes. The DNA-protein complex was deposited on a functionalized glass slide and imaged by fluorescence microscopy. **B** Same as **A** for MRX, yMR and Xrs2.

Paper I highlights the power of nanofluidics when dissecting DNA-protein interactions at the single-molecule level by distinguishing between circularization (intramolecular) and concatemerization (intermolecular) of λ-DNA. The main result of the study concluded that MRN promotes intramolecular bridging, leading to formation of circular DNA, while MRX promotes the formation of long DNA concatemers. These findings are reflected in their roles within DNA repair: MRN is mainly active in HR (requiring sister chromatid proximity) and MRX in NHEJ (joining broken ends). From analysis of the individual constituents of the MRN/X complexes it was found that NBS1 contributes to the ability of MRN to bridge DNA ends and that Xrs2 significantly boosts the end-joining by MRX. Both NBS1 and Xrs2 bind DNA on their own, but their effects are amplified in the full MRN/X complexes. Similar observations were made for y/hMR, alone they show limited DNA bridging, even at high concentrations, but bridging efficiency increase significantly when NBS1 or Xrs2 are present, indicating a synergistic effect. Together this suggests that NBS1 and Xrs2 are critical for the DNA-bridging function of MRN and MRX, respectively, and that their roles align with the distinct repair pathways: NBS1 with HR and Xrs2 with NHEJ.

6.2 Effects of Mg^{2+} and ATP on YOYO-1 labeling of genomic DNA in single molecule experiments

Many proteins that interact with DNA have an ATP-driven activity that is dependent on a cofactor, such as Mg^{2+}, to function properly. The study presented in **Paper I** was done without the presence of ATP or Mg^{2+}, although both MRN and MRX do have ATP dependent nucleolytic activity that processes the broken ends. However, the study was limited to comparing the DNA-binding activity of the proteins prior to processing of the ends and ATP is not involved in this binding step. As previously mentioned, YOYO-1 is a typical dye used for visualization of DNA in nanochannels by fluorescence microscopy. The affinity of YOYO-1 to DNA is influenced by the ionic strength[60,76,77] and an ionic strength below 100 mM is preferable for stable YOYO-1 binding[64]. Of note is that the mentioned studies do not include multivalent ions. The ionic strength also influences the conformation of nanoconfined DNA, leading to a decreasing extension with increasing ionic strength[78]. These effects make the choice of buffer an integral part of experiments that involve YOYO-1-labeled DNA in nanoconfinement.

When performing pilot experiments for, amongst others, **Paper I**, it was found that the addition of ATP and/or Mg^{2+} to YOYO-1-labeled DNA interferes with the imaging in the nanofluidic channels. Since the ability to include ATP-dependent proteins in future nanofluidic studies is of great interest. These observations, in addition to previous nanofluidic studies that demonstrated that YOYO-1 is not a suitable dye at a Mg^{2+} concentrations of 5 mM or higher[79] and that YOYO-1 labeled DNA is affected by ATP[80], motivated further investigation.

It is well known that Mg^{2+} interacts with, and has a stabilizing effect on, DNA[18,81,82] and that Mg^{2+} acts as a cofactor for many ATP driven processes[12], where Mg^{2+} and ATP often exist as a $[MgATP]^{2-}$ complex[83]. With this in mind, it is expected that the relation between YOYO-1, Mg^{2+} and ATP would be best described by a competitive binding model where Mg^{2+} alters the affinity of YOYO-1 to DNA, mainly by electrostatic screening[81,82].

Paper II reports a study on the influence of Mg^{2+} and ATP on the fluorescence emission intensity of YOYO-1-labeled and confined DNA. In addition to the nanofluidic device presented in **Paper I**, **Paper II** utilized a recently developed parallelized nanofluidic device,[70,84] capable of handling ten different samples at once. By coupling this device with a programmable pneumatic pump, simultaneous automated data collection of a large set of conditions was possible. By exploring a range of Mg^{2+} and ATP concentrations at two different dye loadings, it was found that Mg^{2+} and ATP greatly impact the fluorescence intensity from YOYO-1 labeled DNA, **Figure 18A-B**. The data shows a decrease in fluorescence intensity independent of the dye loading, with an almost abolished fluorescent signal at 5 mM Mg^{2+} both with and without ATP. When ATP was added in excess to Mg^{2+} the loss of fluorescence was countered, as expected due to complexation between ATP and Mg^{2+}. Since ATP contributes to the ionic strength of the solution one would expect a loss of fluorescence intensity when ATP is present in the buffer (in addition to Mg^{2+}), if the system was only influenced by the ionic strength. Since the addition of ATP instead mitigates the fluorescence intensity drop, it points towards the suggested competitive binding between Mg^{2+} and YOYO-1 rather than a simple dependence on ionic strength.

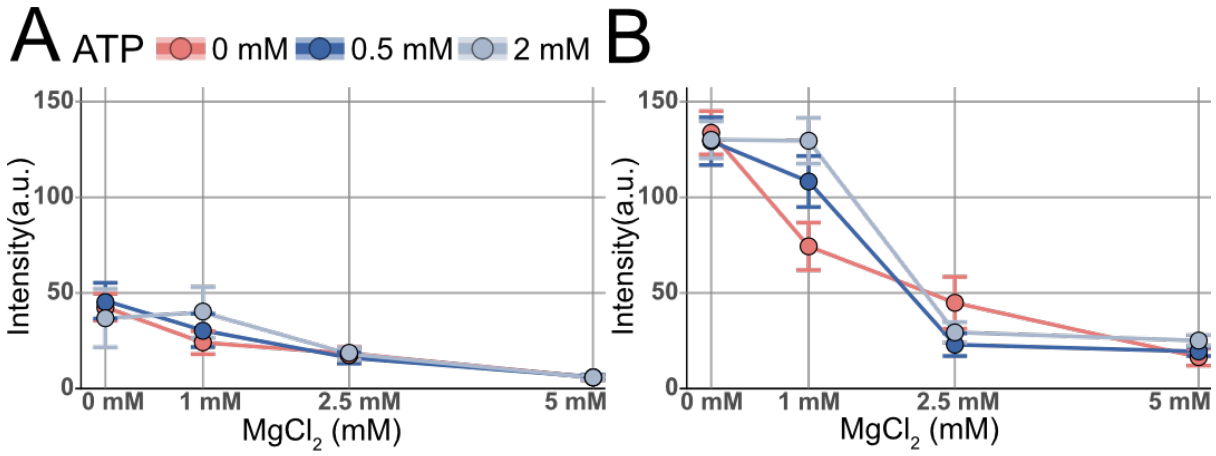

Figure 18. Mean, background adjusted, fluorescence intensity of monomeric λ-DNA at varying ATP and Mg^{2+} concentrations at a 1:10 (**A**) and 1:3 (**B**) dye:bp ratio, respectively.

6.2.1 Mg^{2+} and ATP concentrations reflected by YOYO-1 binding constants

To further characterize the influence of Mg^{2+} and ATP on YOYO-1, binding constants were determined for all considered conditions by measuring the fluorescence at increasing levels of YOYO-1 with a plate reader. The binding constant was determined by fitting a binding model to the emission. The applied model:

$$[L_T] = [L_f] + \frac{K[L_f]}{1+K[L_f]}\left(\frac{[S_T]}{n}\right) \quad (5)$$

was derived from the equilibrium constant:

$$K = \frac{[LS]}{[L_f][S_f]} \quad (6)$$

and conservation of mass:

$$[L_T] = [L_f] + [LS] \quad (7)$$

$$[S_T] = [S_f] + [LS] \quad (8)$$

Where L is ligand, S is substrate and subscript *f* and *T* is *free* and *Total*. The term including *n* was added to adjust for YOYO-1 binding more than 1 bp.

Figure 19 reports the normalised fluorescence from an increasing concentration of YOYO-1 bound to DNA under a large set of buffer conditions. **Figure 19A** shows the difference with increasing concentrations of Mg^{2+}. The maximum fluorescence signal was reached at a dye:bp ratio of 1:1. At higher dye loadings (data not shown) the intensity dropped as YOYO-1 self-quenches[85]. **Figure 19B** shows the corresponding data from conditions with increasing concentration of ATP and **Figure 19C** shows experiments with 2 mM ATP and increasing concentration of Mg^{2+}. **Figure 19D** reports the calculated binding constants for all experiments combining ATP and Mg^{2+}. From the data in **Figure 19** it is apparent that the affinity of YOYO-1 for DNA is reduced in the presence of Mg^{2+} and ATP alone as well as in combination, but to a smaller extent when both were present. This is in line with the assumption that it is Mg^{2+} that affects the affinity of YOYO-1 and that complexation between Mg^{2+} and ATP reduces Mg^{2+} interacting with DNA.

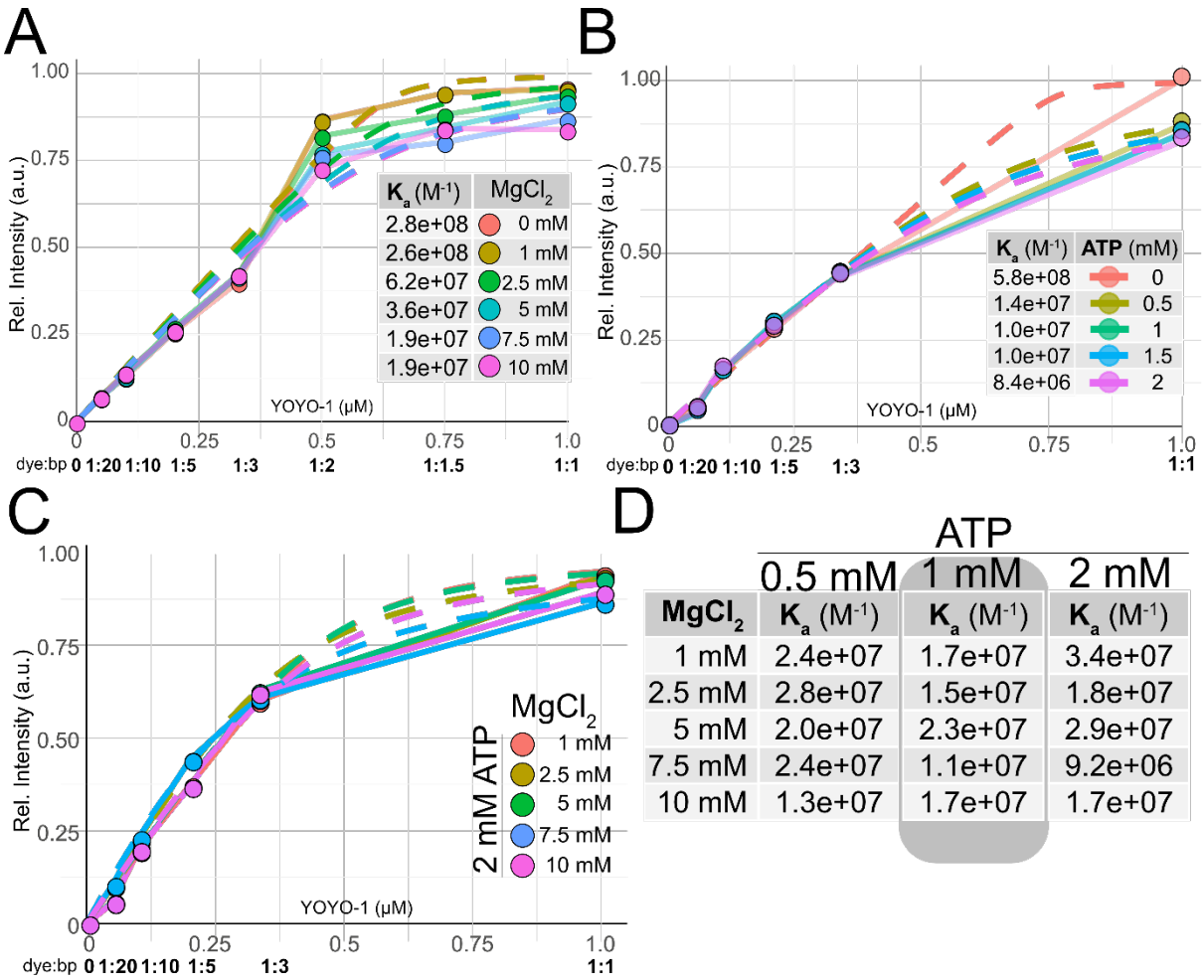

Figure 19. **A** Normalized fluorescence at increasing YOYO-1 concentrations at a constant DNA concentration of 1 μM (bp) and increasing concentrations of Mg^{2+}. Dashed lines represent the fitted model, and the table inset reports the K_a values. **B** Normalized fluorescence at increasing YOYO-1 concentrations at a constant DNA concentration of 1 μM (bp) and increasing concentrations of ATP. Dashed lines represent the fitted model, and the table inset reports the K_a values. **C** Normalized fluorescence at increasing YOYO-1 concentrations at a constant DNA concentration of 1 μM (bp), constant ATP concentration of 2 mM and increasing concentrations of Mg^{2+}. Dashed lines represent the fitted model. **D** Table with the associated constants from experiments at three fixed ATP concentrations and varying Mg^{2+} concentrations.

6.2.2 Less photolytic damage of YOYO-1 labeled DNA when Mg^{2+} is present

When inspecting individual kymographs, **Figure 20A**, there was a noticeable difference in photoinduced fragmentation damage at different conditions, where the presence of Mg^{2+} seemingly reduced the rate of fragmentation. To verify that the presence of Mg^{2+} had an effect on the photolytic fragmentation DNA molecules, an assay was performed where samples were exposed to light for a set amount of time, loaded on a gel and analyzed by electrophoresis. On the gel, the fragmented DNA appeared as a smear below the undamaged linear DNA. The bands from undamaged and damaged DNA were quantified and plotted against time, **Figure 20B**.

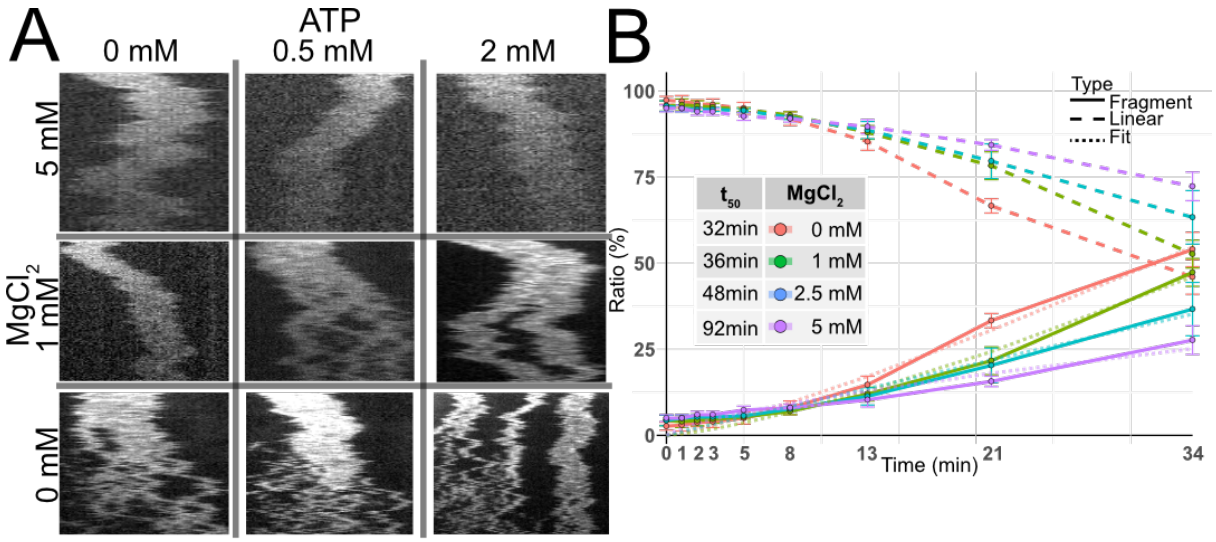

Figure 20. **A** Representative kymographs from bleaching/fragmentation experiments with λ-DNA stained with YOYO-1 at a 1:10 bp:dye ratio and varying concentrations of Mg^{2+} and ATP. **B** Densiometric quantification of results from gel electrophoresis. The fragmented DNA was defined as all signal below the linear band. The dashed line is the normalized intensity of the intact linear fraction, and the solid line is the fragmented fraction. The dotted line is the fitted power-law ($y = rt^n$) describing the fragmentation rate for each condition. The table reports the extrapolated time at which 50% of the DNA is fragmented (t_{50}).

Together the findings from this study show that Mg^{2+} has a great influence on the affinity of YOYO-1 to DNA. This can be explained by Mg^{2+} interacting with the phosphate backbone of the DNA[86], resulting in a charge screening effect that decreases the otherwise strong electrostatic interaction between YOYO-1 and DNA. In conclusion, these findings highlight the critical influence of Mg^{2+} and ATP on the fluorescence from YOYO-1-labeled. The observed loss in fluorescence intensity at higher Mg^{2+} concentrations, and its partial recovery upon addition of ATP, underscores the importance of careful buffer optimization for single-molecule imaging. These insights are essential for designing future nanofluidic experiments involving active, ATP-dependent protein processes, where maintaining both fluorescence signal integrity and biological activity is paramount.

6.3 Single-molecule characterization of plasmid vectors

Detection and characterization of small circular DNA molecules in heterogenous samples is a reoccurring practice when investigating, for example, extrachromosomal circular DNA[87], non-viral episomal vectors[88], viruses[89] and plasmids. In **Paper III,** a method was developed that can identify sparse DNA molecules in a heterogenous sample using picogram of input material without the need for amplification or removal of background DNA. The method leverages the high sequence specificity and affinity of a fluorescent and nucleolytically dead Cas9 (dCas9) to identify target DNA populations, coupled with the linear dependence of YOYO-1 fluorescence on DNA size. The method was validated with DNA samples extracted from rAAV transduced HEK293 cells and it was demonstrated that it is possible to determine the size distribution of rAAV vectors in cells.

6.3.1 Fluorescence intensity from YOYO-1 labeled DNA scales with size

The developed method is based on two core capabilities, sizing of circular DNA and sequence specific identification of plasmids. By immobilizing the DNA sample on functionalized glass slides in a sufficiently low density to allow single molecule detection, it was possible to detect single plasmids and quantify their YOYO-1 signal. The sizing is based on the correlation between fluorescence intensity from YOYO-1 and the size of the DNA molecule[90]. Since YOYO-1 is non-fluorescent when unbound, all fluorescence signal from a detected punctum can be assumed to be related to the presence of DNA. To verify that the sized plasmid is the one of interest, dCas9 with an ATTO550-labeled guide RNA (gRNA), targeting a unique sequence in the plasmid of interest, is added to the sample. By only considering plasmids colocalized with the fluorescence from ATTO550, it is possible to separate the target plasmids from background DNA.

Figure 21A shows a representative image of a YOYO-1 labeled plasmid (green) colocalized with fluorescent dCas9 (magenta). To determine the size of an unknown DNA sample, a calibration curve was constructed from three known plasmid samples. The plasmids were chosen based on availability and size. **Figure 21B** shows an image with representative examples of the three plasmids together with a 2D intensity profile. The plasmids were imaged as a mixed sample and the image, together with the 2D intensity profile, demonstrates clearly how the different sizes appear as puncta with increasing size and intensity. By outlining the boundaries of each punctum, the average intensity of the different plasmids could be calculated. By taking the area weighted and background corrected intensity of each punctum and plot it as a distribution, **Figure 21C**, it is possible to discern three separate populations and determine the average intensity for each included DNA size. The population mean for each plasmid size was then plotted against its known size. By doing a linear regression of the three data points, a simple model of how the intensity scales with DNA size (bp) could be determined, **Figure 21D.** This model was later used to determine the size of unknown DNA samples by extrapolating the expected size for a given intensity.

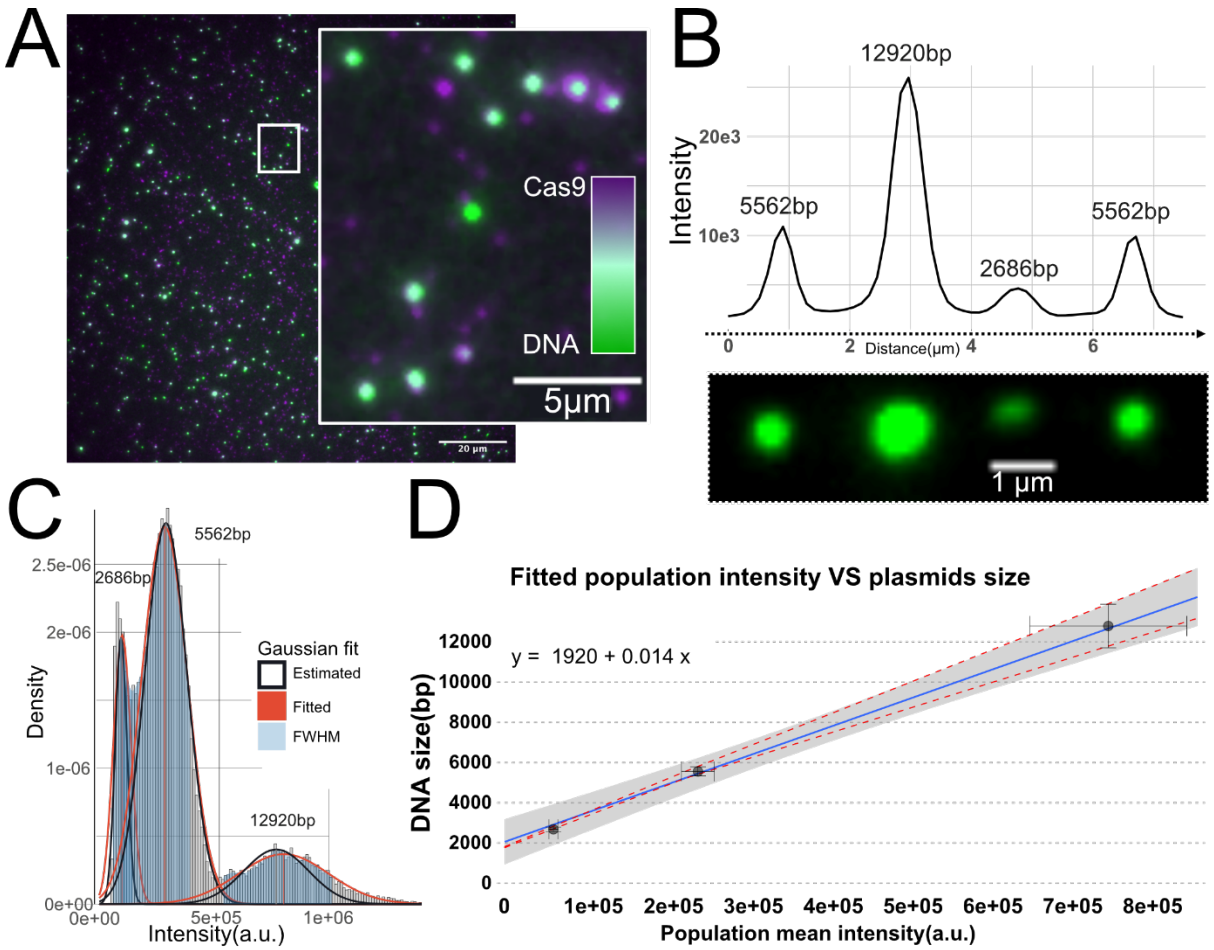

Figure 21. **A** Representative image of colocalization between plasmid DNA (pAAV-U6-sgRNA-CMV-GFP) and dCas9 complexed with ATTO550-labeled RNA targeting the AAV2-ITR. **B** Intensity profile along four plasmids of three different sizes. **C** Histogram of molecule intensities with fitted gaussians from a representative experiment with a mixed plasmid sample. Each molecule intensity is background adjusted. **D** Regression line fitted to known DNA size vs molecule intensity mapping based on the average population intensity extracted from fitted gaussians with the 99% CI outlined in gray. Horizontal error bars represent the SD of the population means from four replicates, thus representing the deviation in average population intensity between experiments, while the vertical error bars represent the average estimation error with the red dashed line outlining the upper and lower bounds of the estimation error along the regression.

6.3.2 GFP expression and qPCR of viral vectors from transduced HEK293

To generate a set of samples to be used for validation of the method, HEK293 cells were transduced with rAAV packed with plasmids designed for the expression of GFP. The GFP expression was used to confirm the success and efficiency of the transduction. The transduction was done as a step gradient of multiplicity of infection (MOI), ranging from 1 to 10k. Phase contrast and fluorescence (FITC) images were recorded for each well and all conditions, **Figure 22A**. The transduction was done in a 96-well plate and since not all wells were fully confluent, the GFP signal needed to be normalized against the actual area covered by cells rather than the entire well. The phase images were segmented to mask any areas not containing cells. By comparing the integrated GFP signal it was found that all conditions had a detectable level of GFP, confirming the presence of virus vectors and viable cells, **Figure 22B**. The cells were subsequently lysed, and the extracted DNA was analyzed using electrophoresis. **Figure 22C** show the resulting image from the electrophoresis and shows that the extracted DNA was

heterogenous in size and that the expected viral vector did not appear in a detectable quantity. qPCR was used to further confirm the presence of viral vectors in the transduced cells and to determine the number of vectors that could be expected in an average cell. To determine the viral vector copy number a standard curve, based on serial dilution of the same vector as used for transduction, was constructed. By extrapolation, the relative amount of viral DNA per ng of extracted DNA could be determined, **Figure 22D**. From this data it was concluded that only the highest MOI contained enough viral DNA to be used for further experiments.

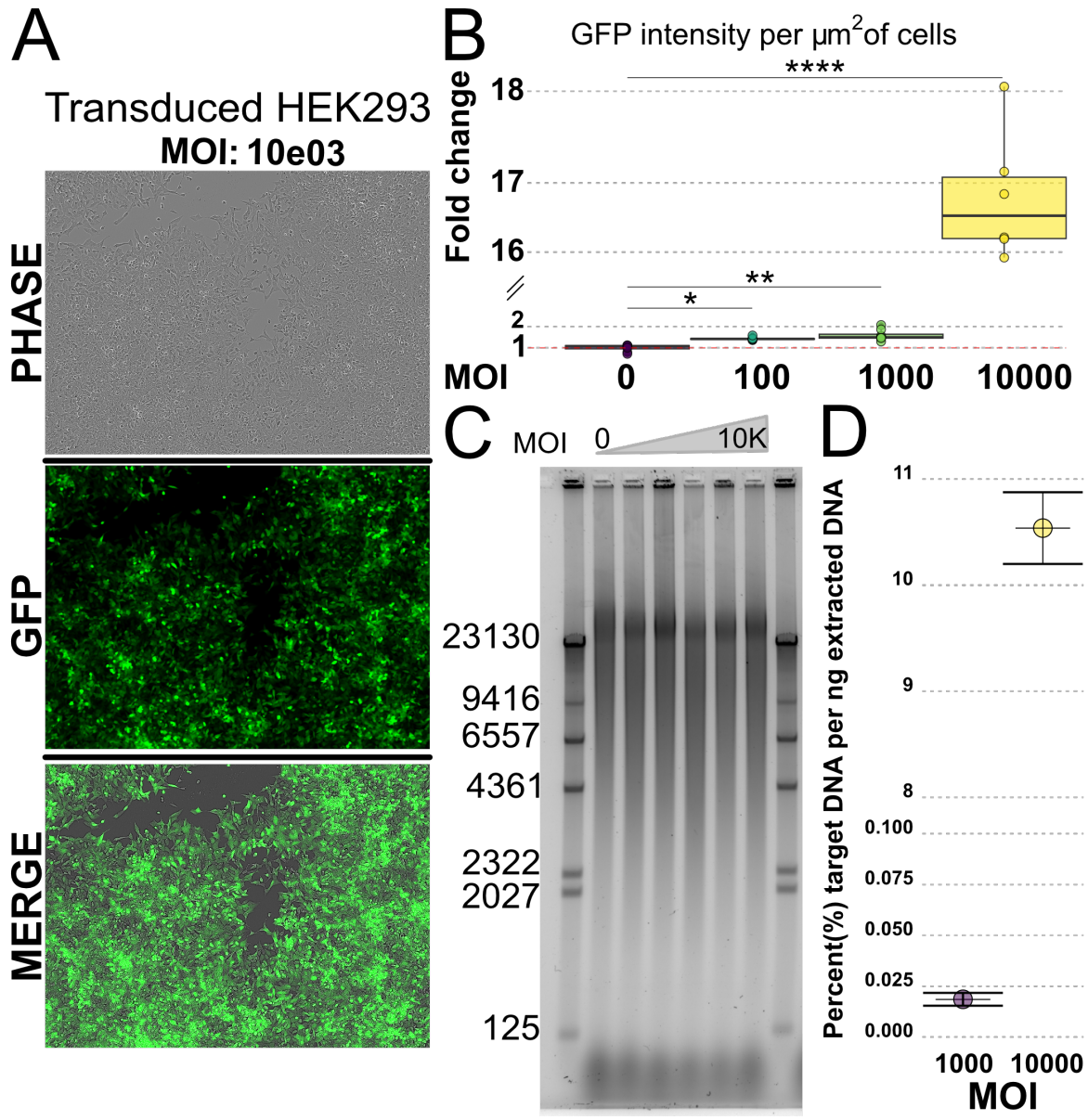

Figure 22. **A** Representative image from HEK293 cells transduced with MOI 10e03. From the top: phase contrast image, fluorescence image, merged image. **B** Fold change of GFP signal relative to negative control. MOI 1 and MOI 10 did not show significant increase in GFP intensity and were therefore omitted. Significance was determine using Dunnett's test. **C** Representative gel from electrophoresis of DNA extracted from rAAV transduced HEK293 cells. **D** Results from qPCR analysis represented as a relative amount of target vector per nanogram of extracted DNA. Only MOI 1k and 10k are shown.

6.3.3 Intensity-based sizing coupled with dCas9 colocalization can be used to confidently resolve sparse DNA populations in heterogenous samples

To investigate if the method could detect and size the viral DNA extracted from transduced cells, fluorescently tagged dCas9-RNA complexes were added to YOYO-1-stained cell-derived DNA and deposited on functionalized glass slides. DNA from the MOI 0 condition was used as control and for background subtraction of unspecific colocalization. **Figure 23A** provides a representative example from an experiment with DNA (green) from MOI 10k colocalizing with fluorescently tagged dCas9-RNA complexes (magenta). The image shows how the circular plasmid DNA is deposited as puncta whereas genomic DNA fragments stretch out in their linear form. The zoom-in shows that, despite adjacent molecules having an apparent similar intensity, some are not colocalized with a dCas9 and are thus most likely not rAAV vectors. The average colocalization degree in the MOI 0 (6.0%) samples **Figure 23C**, revealed that there was a non-negligible amount of unspecific colocalization. Interestingly, the proportional colocalization between positive (MOI 10k: 12.8%) and negative (MOI 0k: 6.0%) samples was the same for cell-derived DNA (**Figure 23C**) as for plasmid DNA (appended **Paper III** Figure 2E), i.e. the degree of colocalization in control samples is roughly 50% of the colocalization in positive samples. To adjust for the non-negligible unspecific colocalization, the histogram of the control sample was subtracted from the histogram of the positive sample. **Figure 23B** shows a background subtracted histogram of molecule intensities from dCas9 colocalized DNA. The histogram reveals two major subpopulations. A main population at 2512 bp that represents a monomeric rAAV vector (expected size 2684 bp) and a subpopulation at 5827 bp, around the size of a dimeric concatemer (~5368bp). When comparing the relative abundance between the monomeric and dimeric subpopulations the monomeric subpopulation is in clear majority (88%).

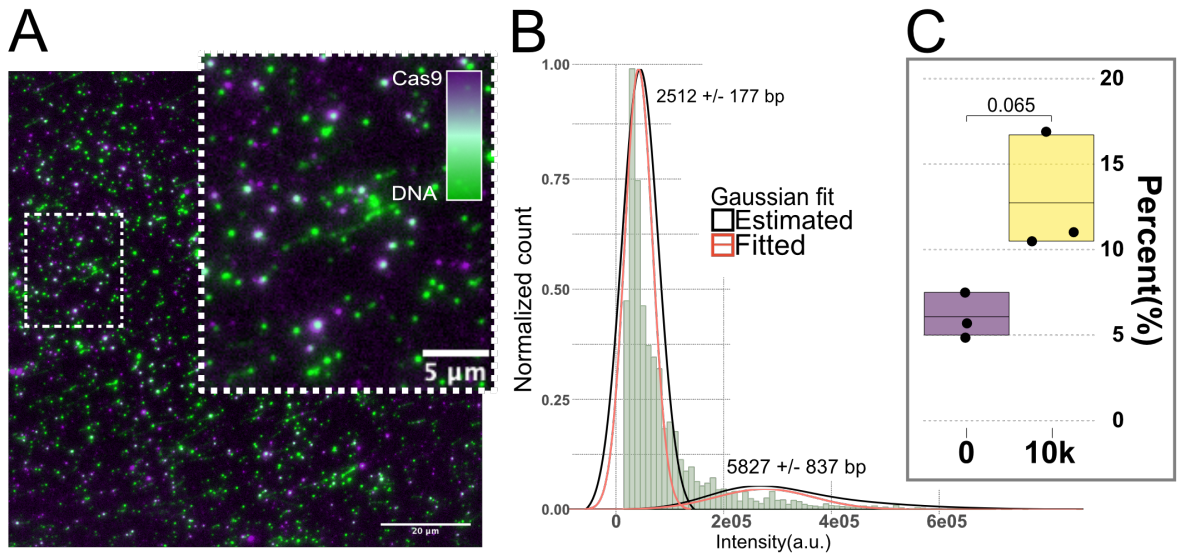

Figure 23. **A** Representative image of colocalization between cell derived DNA (MOI 10K) and dCas9 complexed with ATTO550 labeled RNA targeting the AAV2-ITR. **B** Histogram of the background adjusted colocalized fraction of detected DNA in MOI 10K based on pooled data from multiple experiments. 1D K-means clustering was used to identify individual populations and the bp size of each population was given by the equation from Figure 20D together with the associated estimation error. **C** Relative colocalization degree from multiple experiments using cell derived DNA from either MOI 0 or MOI 10K.

This study presents a novel, amplification-free method for detecting and sizing sparse populations of circular DNA molecules in heterogeneous samples using single-molecule fluorescence microscopy. The method provides detection and quantification of structural features, such as concatemers, and was successfully validated by identifying monomeric and dimeric episomal rAAV vectors in DNA samples derived from transduced cells.

Within rAAV-based therapy, this method could be applied to track the fate of episomal vectors *in vivo* over time to assess stability, integration, and expression efficiency. Furthermore, it is possible to detect and characterize extrachromosomal circular DNA (eccDNA) in tumor samples, which are often associated with oncogene amplification. With further development and optimization, it is also possible to find new applications, such as adaptation of the method for single-cell analysis to study heterogeneity in vector uptake and expression or monitoring of synthetic plasmid behavior in engineered cells, especially in low-copy or transient expression systems.

6.4 Fluorescence microscopy of multi-component LLPS puncta

The study presented in **Paper IV** includes a wide set of methods and techniques used to investigate the formation of TOP2A condensates and the dynamic relation with the oncogenic MYC protein, both *in vivo* and *in vitro*. My contribution to the study includes imaging and quantification of DNA-protein complexes immobilized on functionalized glass substrates to characterize how MYC influences properties such as size and density of TOP2A condensates. The following section will present the scope of the study and its key findings, but will be otherwise limited to the methods, data and results produced by me.

The study confirms the organization of TOP2A in LLPS compartments and demonstrates that TOP2A maintains a dynamic equilibrium between nucleolar and transcription condensates in human cells. This suggests a model where a fraction of the TOP2A accumulated in the nucleolus is part of a buffer system that changes the equilibrium between the nucleolus and the nucleoplasm upon stimuli. The equilibrium is connected to a stress response that can be triggered by the build-up of topological events in the nucleoplasm. The study found that MYC accelerates TOP2A diffusion in cells, and reduces the size of the TOP2A condensates *in vitro*, as well as the size of the protein complexes formed by TOP2A in cells. By increasing TOP2A diffusion, MYC favors substrate detection and increased TOP2A engagement on the chromatin genome-wide, revealing the mechanism for MYC stimulation of TOP2A activity.

6.4.1 TOP2A forms puncta at physiological concentrations

In previous publications, TOP2A condensation has been studied with confocal microscopy, in either well plates, or on untreated microscope slides[97]. In this study, glass coverslips were functionalized with silane to generate a surface with net positive charge that attracts the negatively charged DNA, thus keeping the condensates stationary. To visualize the condensates, the proteins were covalently tagged with fluorophores, and the DNA was labeled with YOYO-1. The use of spectrally separated fluorophores allowed for simultaneous imaging of all components (DNA, TOP2A and MYC) and the possibility to study colocalization of the components in the condensates. The experimental setup could detect puncta at protein

concentrations as low as 5 nM. **Figure 24A** shows representative images of TOP2A condensates formed with DNA at TOP2A concentrations ranging from 5-25 nM. The apparent size and density of the condensates scaled with increasing protein concentration, as expected. Interestingly, not all observed condensates colocalized with DNA although there is a strong correlation between high TOP2A emission intensity and a high signal from the DNA (data not shown), indicating that DNA is an integral part of the formation of large TOP2A puncta. The inset in **Figure 24B** highlights two condensates (from experiments with 10 nM TOP2A) where there is a clear colocalization between a large TOP2A condensate and the DNA substrate. In the intensity profile these two puncta show up as spatially overlapping peaks in each channel, confirming that the formation of TOP2A condensates is stimulated by DNA.

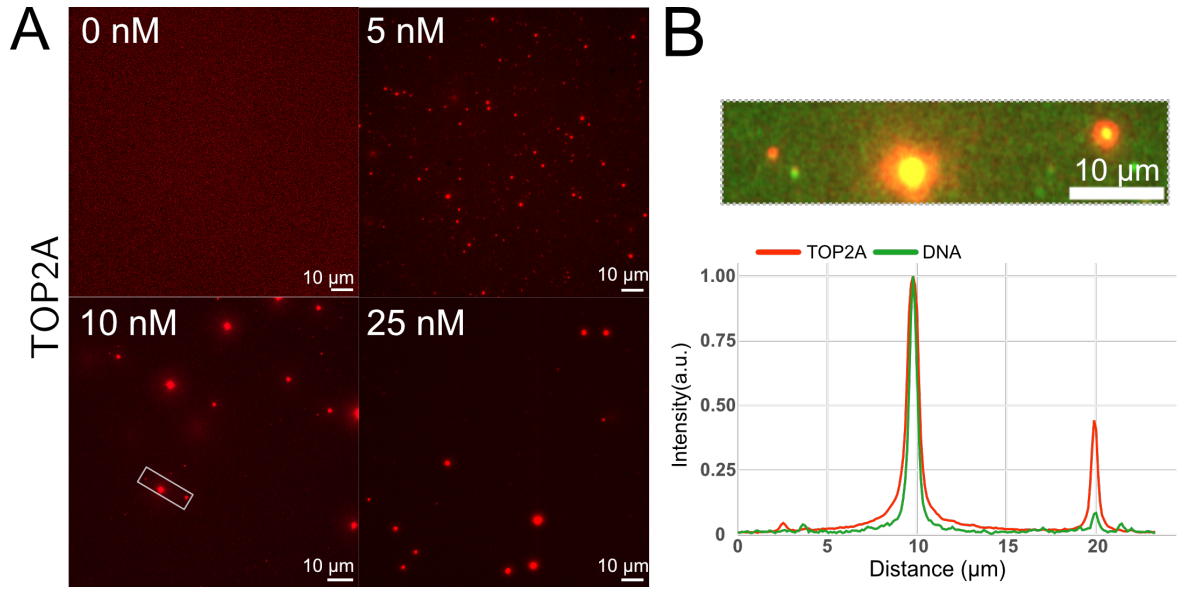

Figure 24. Representative images of TOP2A puncta at varying concentrations. The DNA was labeled with YOYO-1 and TOP2A with ATTO647.

6.4.2 MYC stimulates condensate formation by TOP2A

Next, MYC was added to investigate if the interaction between MYC and TOP2A reported by Das et al.[92] was affecting the condensate formation. The experimental setup was at this stage altered so that DNA was labeled with SYTOX Orange, TOP2A with Alexa488 and MYC with Alexa647. This change did not entail any changes in the detection or analysis. Experiments combining both proteins confirm the colocalization between TOP2A, DNA and MYC. Representative images from such experiments are reported in **Figure 25A.** The inset in **Figure 25B** highlights two puncta. In the intensity profile the puncta show up as spatially overlapping peaks in each channel, confirming that the puncta contain both TOP2A, DNA and MYC. Interestingly, the colocalization of TOP2A and MYC is not obligatory despite both being present in the sample. This could either indicate that the involvement of MYC in the condensate formation does not necessarily need to be a matter of co-condensation, but could potentially be explained by MYC having more of a catalytic role. Or, that MYC leaves the TOP2A condensate once the DNA has been processed.

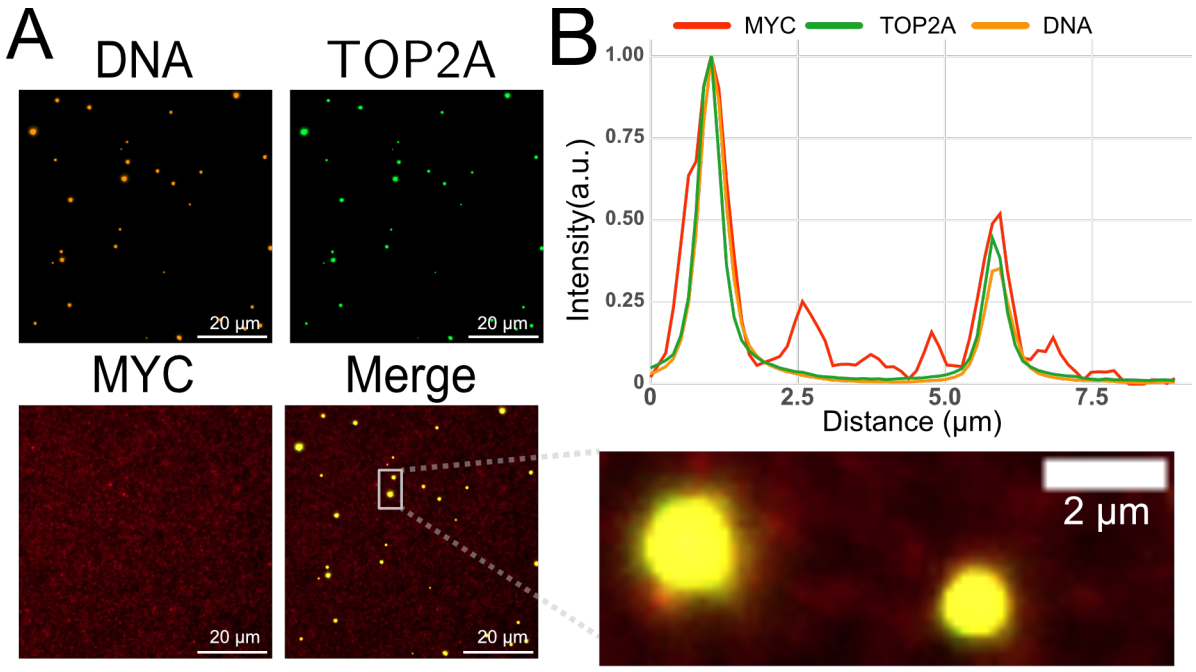

Figure 25. **A** Representative images of TOP2A puncta colocalizing with DNA and MYC. **B** Inset from merged image in **A** with intensity profile for all three channels indicating the spatial overlap between DNA, TOP2A and MYC puncta.

Having established that TOP2A puncta can be detected at very low TOP2A concentrations and that MYC indeed colocalizes with TOP2A, the TOP2A and MYC concentrations were fixed to 25 nM and 100 nM, respectively. Additional to the wild type MYC, a non-interacting delta variant of MYC consisting only of residues 200-320 (ΔMYC) was used as a control to further map MYCs involvement in condensate formation. The ΔMYC variant has previously been shown to elicit no effect on TOP2A activity[92]. Quantification of multiple replicates revealed that the presence of MYC decreased the average area of a TOP2A punctum, **Figure 26A,** while the ΔMYC variant did not. This indicates that MYC is involved in the organization of the TOP2A condensates in such a way that the condensates either contain fewer TOP2A proteins per condensate or that MYC alters the condensate properties so that they interact more with the functionalized surface, spreading out and increasing the area. When categorizing the detected TOP2A puncta based on their DNA content it is clear that DNA plays an important role in the regulation of their size. **Figure 26B** shows the results from a t-test comparing the average size of TOP2A puncta with and without DNA. All conditions show a significant increase in size when DNA is present. Interestingly, the condition including MYC shows a more statistically significant difference.

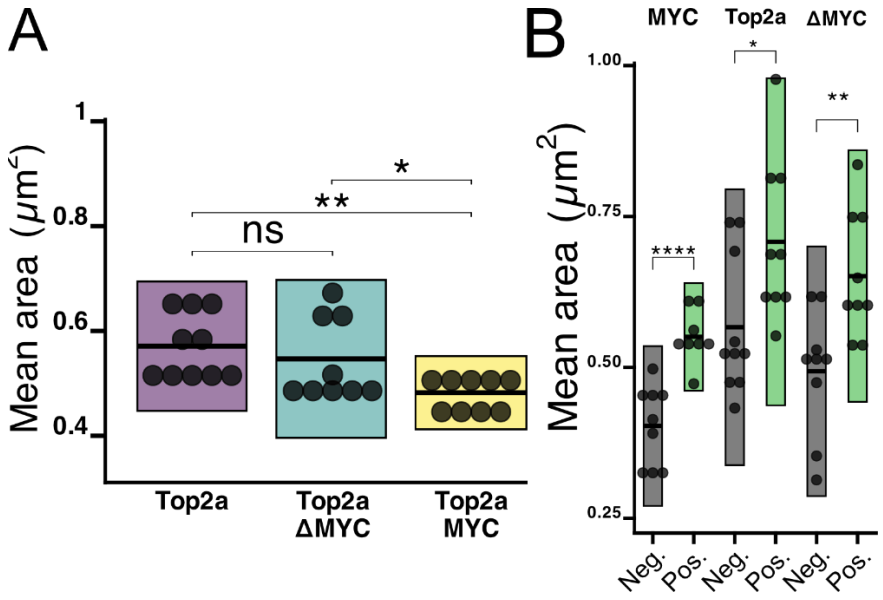

Figure 26. **A** Mean area of TOP2A puncta across at least 9 replicates. A t-test shows a significant decrease in area when MYC is present, but no effect by the delta variant MYC$^{200-320}$. **B** Mean area of TOP2A puncta with (pos.) and without (neg.) colocalization with DNA. t-test shows that the presence of DNA significantly increases the area for all conditions. The condition with MYC shows much higher confidence in the statistical difference compared to other conditions.

Looking to the intensity of each TOP2A puncta, there is an overall decrease in intensity in the presence of MYC, although not significant, **Figure 27A**. The ΔMYC variant also shows a non-significant decrease in intensity, but not as noticeable as for MYC. Since the intensity is correlated with the number of TOP2A molecules present in a condensate, the decrease in intensity suggests that the smaller surface area is less likely explained by surface interaction, but rather due to MYC influencing the growth of the condensates. Comparing the intensity between TOP2A puncta with and without DNA (**Figure 27B**), there is a significantly higher intensity for DNA containing puncta in the MYC condition, but not for the other conditions, indicating that the suggested effect by MYC is DNA dependent. From the imaged area and the volume of the applied sample it was possible to extrapolate how many condensates were present in each sample and condition. **Figure 27C** shows the results from these extrapolations and there is a significant increase in number of puncta when MYC is present.

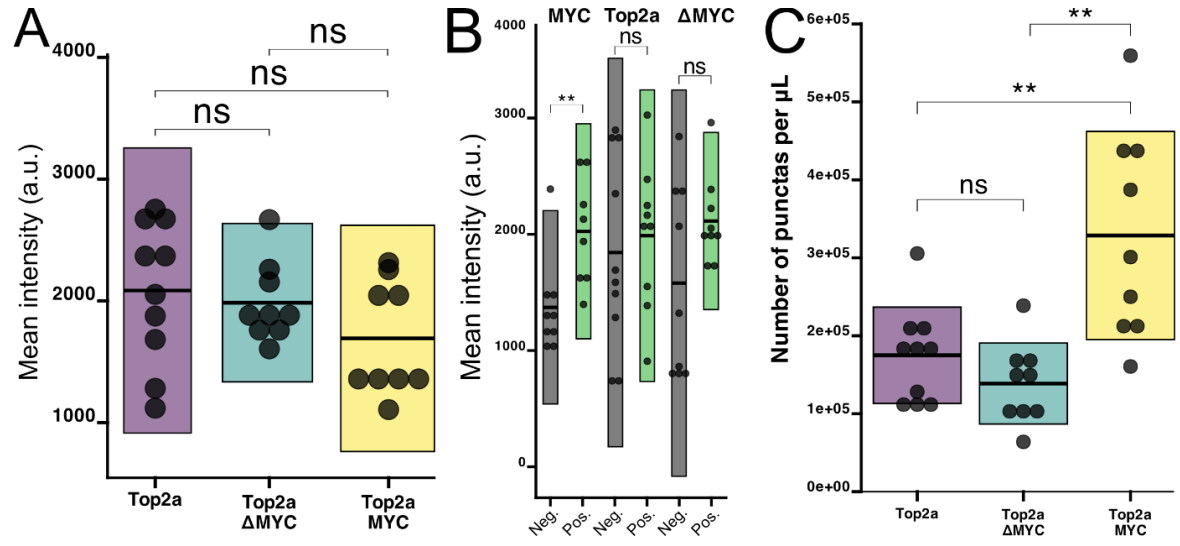

Figure 27. **A** Mean intensity of TOP2A puncta across at least 9 replicates. A t-test shows a non-significant difference between either of the conditions. **B** Mean intensity of TOP2A puncta with and without colocalization with DNA. t-test shows that the presence of DNA significantly increases the intensity for the condition with MYC, but not for the others. **C** Number of puncta found per µL of sample. A t-test shows that the presence of MYC significantly increases the number of puncta.

6.4.3 MYC accelerates TOP2A target search and activity

Taking the presented data together, we observed a behavior where MYC increases the number of TOP2A puncta by limiting their overall size. Results also showed that MYC favors DNA-bound TOP2A. This could be explained by MYC having a kinetic effect in the sense that an accelerated formation of condensates would decrease the chance of larger condensates to form. Alternatively, MYC changes the internal organization of the condensates by interfering with TOP2A protein-protein interactions to change the viscoelasticity. A more loosely kept structure would allow for increased mobility, within, and from the condensates, increasing the chances of a TOP2A finding target DNA. This would decrease the overall size due to a greater exchange between the condensates and the surroundings. Since condensate characteristics showed to be dependent on the binding of DNA it can be assumed that the effect of MYC is connected to the activity of TOP2A.

7 Concluding remarks

For a cell to stay viable it is of utmost importance that the integrity of DNA is maintained. The genome is constantly damaged, either by, for example, cellular processes during translation and transcription or by external sources, like radiation and toxins. The unsuccessful or unfaithful repair of these damages can have pathological consequences; it is therefore important to learn more about how these processes work and what the interaction between DNA and the proteins looks like.

The study presented in **Paper I** on the MRN/X complexes and their involvement in DNA repair revealed distinct mechanisms of DNA end-bridging, highlighting the nuanced roles of NBS1 and Xrs2 in homologous recombination and non-homologous end joining, respectively. These findings underscore the importance of dissecting protein complex subunit contributions to understand the orchestration of DNA repair pathways. Further studies should include efforts to study the DNA bridging activity of MRN/X and its subunits in an active state, in the presence of ATP. Furthermore, the inclusion of alternate DNA substrates would increase the breadth of the applied analysis where modifications to the ends of the DNA could provide additional information.

The investigation into the effects of Mg^{2+} and ATP on YOYO-1 labeling, presented in **Paper II**, emphasize the critical interplay between biochemical conditions and fluorescence-based readouts. This insight is essential for designing robust single-molecule experiments, particularly when studying active enzymatic processes.

By leveraging nanofluidic confinement and fluorescence microscopy, the work presented in this Thesis demonstrates how individual molecular events, often obscured in ensemble measurements, can be visualized, quantified, and interpreted in biologically meaningful ways. To further broaden the analytical framework that have been applied in the nanofluidic studies it would be of interest to increase the understanding on the kinetics behind circularization and concatemerization. While the formation of circles can be explained in a sufficient manner, the joining of two neighboring molecules has less apparent theoretical grounds and gaining insight into this process would enable a more detailed interpretation of end-joining events.

Surface immobilization of DNA offers a simple, robust and versatile platform for single-molecule studies, complementary to nanofluidics. This is especially beneficial when dealing with small sample quantities or expensive reagents. **Paper III** presents a Cas9-based detection method for sparse plasmids in heterogeneous samples. The method is a scalable and amplification-free approach for DNA sizing and identification of DNA in sample volumes comparable to a single cell. This technique holds promise for applications in gene therapy, synthetic biology, and diagnostics. To ensure robust results in future applications it would be of interest to include engineered Cas9 variants with increased specificity. Furthermore, the use of multiple gRNAs would allow multiplexed detection, which could either be used to increase the robustness of the signal or to widen the number of species detectable in a sample. It would also be possible to miniaturize the method to be used on a compartmentalized glass slide to enable automated detection of multiple samples.

By immobilizing molecules on a surface, it is possible to get a snapshot of the sample. This is essential when studying dynamic processes such as biomolecular condensate formation where the sample characteristics is constantly changing. The contribution to **Paper IV** includes an assay that quantifies the influence of MYC on TOP2A biomolecular phase separation and its role in genome organization and transcriptional regulation. The observed MYC-induced changes in condensate dynamics suggest a regulatory mechanism that may be exploited in cancer biology. The use of fluorescence microscopy to study biomolecular condensates is robust and simple. To add a layer of information, it would be suitable to use nanofluidic devices where the condensates are free to diffuse. This would provide precise mass estimations as well as geometric information that could lead to additional insights.

In all, the work presented in this Thesis demonstrates the use of fluorescence wide-field microscopy within quantitative single-molecule methods. The ease of use and relative accessibility of fluorescence wide-field microscopy makes it a perfect tool for quantitative single molecule assays with a wide range of applications.

8 Acknowledgements

I would like to express my sincere gratitude to all the people for their direct or indirect support during all the hard work involved in this Thesis. I would like to extend a special thank you to the following people:

First and foremost, my supervisor **Fredrik**. For your incomparable trust and belief in me to explore and solve problems on my own terms. It has been testing at times but ultimately allowed me to develop into an independent scientist with a broad skillset and great confidence.

My examiner **Elin** and my co-supervisor **Marcus**, for your valuable support and advice and for always being there whenever needed.

All current and former members of the Westerlund group. A special thanks to **Robin** who helped me when I started and set the bar for great science. **Gaurav**, for always being available when I had questions, needed support or just inspiration. **Sriram**, for all the great discussions and for listening whenever I had a "new" idea. **Karolin**, for your support and thorough feedback of this Thesis. **My**, for your unfaltering optimism and reassuring support. **Evgeniya**, for your great friendship, all the days in the lab and your willingness to listen every time my confidence was wavering. **Anusha**, for your relentless work and interesting discussions. **Luis** and **Vinoth**, for some great ideas and pointers. **Florian** for all the days in the lab, working side-by-side and discussing new ideas. **Radhika** for your great support and invaluable feedback. **Dennis** for all the days in the lab and your contributions to the Mg^{2+}/YOYO project.

Furthermore, I would like to extend a big thank you and express my gratitude to all colleagues at the divisions of Chemical biology and Physical chemistry for creating such a nice and friendly work environment. A special thank you to **Ann-Britt** for being my buddy and providing a lot of support and help ever since.

Petr and his team members Elda, Ilaria and Giordano at IRB in Bellinzona, Switzerland for the great collaboration on the human and yeast HR project. Laura, Donald and Kathryn at Karolinska Institutet for bringing me in on a great project in a very interesting field. Special thanks to Donald for your great commitment and critical eye. I hope we can collaborate again. I would also like to thank all co-authors from AstraZeneca for their resourceful and valuable contributions to the Cas9 project. Tobias Ambjörnsson and Albertas Dvirnas at Lund University for all the support.

To **Linnea**. Thank you for your unwavering support during all the struggles and sleepless nights. For the late-night discussions and for pushing me to just try it one more time.
I would not have wanted to do it without you.

9 References

1. Bustamante, C. Direct Observation and Manipulation of Single DNA Molecules using Fluorescence Microscopy. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 415–446 (1991).

2. Matsumoto, S., Morikawa, K. & Yanagida, M. Light microscopic structure of DNA in solution studied by the 4′,6-diamidino-2-phenylindole staining method. *J. Mol. Biol.* **152**, 501–516 (1981).

3. Rye, H. S. *et al.* Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications. *Nucleic Acids Res.* **20**, 2803–2812 (1992).

4. Jin, X. & Jo, K. Handbook of Chemical Biology of Nucleic Acids. 1–30 (2022) doi:10.1007/978-981-16-1313-5_53-1.

5. Kissling, V. M. *et al.* Mre11-Rad50 oligomerization promotes DNA double-strand break repair. *Nature Communications 2022 13:1* **13**, 1–16 (2022).

6. Kalle, W. & Strappe, P. Atomic force microscopy on chromosomes, chromatin and DNA: A review. *Micron* **43**, 1224–1231 (2012).

7. Li, X., Gucht, J. van der, Erni, P. & Vries, R. de. Active microrheology of protein condensates using colloidal probe-AFM. *J Colloid Interf Sci* **632**, 357–366 (2023).

8. Bustamante, C. J., Chemla, Y. R., Liu, S. & Wang, M. D. Optical tweezers in single-molecule biophysics. *Nat Rev Methods Primers* **1**, 25 (2021).

9. Wang, Y., Sischka, A., Walhorn, V., Tönsing, K. & Anselmetti, D. Nanomechanics of Fluorescent DNA Dyes on DNA Investigated by Magnetic Tweezers. *Biophys. J.* **111**, 1604–1611 (2016).

10. Tegenfeldt, J. O. *et al.* The dynamics of genomic-length DNA molecules in 100-nm channels. *Proc. Natl. Acad. Sci.* **101**, 10979–10983 (2004).

11. Alberts, B. *et al.* Molecular Biology of the Cell. (2007) doi:10.1201/9780203833445.

12. Appling, C. K. M. K. van H. D., Holde, K. van, Appling, D. & Anthony-Cahill, S. J. *Biochemistry*. (Pearson, 2013).

13. Commons, W. *File:Central Dogma of Molecular Biology.Svg --- Wikimedia Commons, the Free Media Repository*.

14. WATSON, J. D. & CRICK, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).

15. Cobb, M. & Comfort, N. What Rosalind Franklin truly contributed to the discovery of DNA's structure. *Nature* **616**, 657–660 (2023).

16. Commons, W. *File:DNA Structure Formula Virgin.Svg --- Wikimedia Commons, the Free Media Repository*.

17. Commons, W. *File:DNA Structure and Bases.Svg --- Wikimedia Commons, the Free Media Repository*.

18. Bloomfield, V. A. *Nucleic Acids : Structures, Properties, and Functions*. (2000).

19. Rivetti, C., Walker, C. & Bustamante, C. Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility11Edited by D. Draper. *J. Mol. Biol.* **280**, 41–59 (1998).

20. Reisner, W., Pedersen, J. N. & Austin, R. H. DNA confinement in nanochannels: Physics and biological applications. *Reports on Progress in Physics* **75**, (2012).

21. Frykholm, K., Müller, V., Dorfman, K. D. & Westerlund, F. DNA in nanochannels: theory and applications. *Quarterly Reviews of Biophysics* (2022) doi:10.1017/s0033583522000117.

22. Gennes, P. G. de. One long chain among shorter chains. *J. Polym. Sci.: Polym. Symp.* **61**, 313–315 (1977).

23. Odijk, T. On the limiting law solution of the cylindrical Poisson-Boltzmann equation for polyelectrolytes. *Chem. Phys. Lett.* **100**, 145–150 (1983).

24. Huang, R. & Zhou, P.-K. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct. Target. Ther.* **6**, 254 (2021).

25. Rich, T., Allen, R. L. & Wyllie, A. H. Defying death after DNA damage. *Nature* **407**, 777–783 (2000).

26. Ranjha, L., Howard, S. M. & Cejka, P. Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma* **127**, 187–214 (2018).

27. Mao, Z., Bozzella, M., Seluanov, A. & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle* **7**, 2902–2906 (2008).

28. Heyer, W.-D., Ehmsen, K. T. & Liu, J. Regulation of Homologous Recombination in Eukaryotes. *Annu. Rev. Genet.* **44**, 113–139 (2010).

29. Syed, A. & Tainer, J. A. The MRE11-RAD50-NBS1 Complex Conducts the Orchestration of Damage Signaling and Outcomes to Stress in DNA Replication and Repair. *Annual Review of Biochemistry* **87**, 263–294 (2018).

30. Anand, R., Ranjha, L., Cannavo, E. & Cejka, P. Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Mol Cell* **64**, 940–950 (2016).

31. Gnügge, R. & Symington, L. S. DNA end resection during homologous recombination. *Current Opinion in Genetics & Development* **71**, 99–105 (2021).

32. Cannavo, E. & Cejka, P. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature* **514**, 122–125 (2014).

33. Lafrance-Vanasse, J., Williams, G. J. & Tainer, J. A. Envisioning the dynamics and flexibility of Mre11-Rad50-Nbs1 complex to decipher its roles in DNA replication and repair. *Progress in Biophysics and Molecular Biology* **117**, 182–193 (2015).

34. Schiller, C. B. *et al.* Structure of Mre11–Nbs1 complex yields insights into ataxia-telangiectasia–like disease mutations and DNA damage signaling. *Nature Structural & Molecular Biology 2012 19:7* **19**, 693–700 (2012).

35. Nakada, D., Matsumoto, K. & Sugimoto, K. ATM-related Tel1 associates with double-strand breaks through an Xrs2-dependent mechanism. *Genes & Development* **17**, 1957–1962 (2003).

36. Reginato, G., Cannavo, E. & Cejka, P. Physiological protein blocks direct the Mre11–Rad50–Xrs2 and Sae2 nuclease complex to initiate DNA end resection. *Genes Dev.* **31**, 2325–2330 (2017).

37. Oh, J., Al-Zain, A., Cannavo, E., Cejka, P. & Symington, L. S. Xrs2 Dependent and Independent Functions of the Mre11-Rad50 Complex. *Molecular Cell* **64**, 405–415 (2016).

38. Oh, J., Lee, S. J., Rothstein, R. & Symington, L. S. Xrs2 and Tel1 Independently Contribute to MR-Mediated DNA Tethering and Replisome Stability. *Cell Reports* **25**, 1681-1692.e4 (2018).

39. Chen, L., Trujillo, K., Ramos, W., Sung, P. & Tomkinson, A. E. Promotion of Dnl4-Catalyzed DNA End-Joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 Complexes. *Mol Cell* **8**, 1105–1115 (2001).

40. Paull, T. T. & Gellert, M. A mechanistic basis for Mre11-directed DNA joining at microhomologies. *Proc National Acad Sci* **97**, 6409–6414 (2000).

41. Tauchi, H. *et al.* Nbs1 is essential for DNA repair by homologous recombination in higher vertebrate cells. *Nature* **420**, 93–98 (2002).

42. Koonin, E. V. & Makarova, K. S. Origins and evolution of CRISPR-Cas systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**, 20180087 (2019).

43. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

44. Deng, W., Shi, X., Tjian, R., Lionnet, T. & Singer, R. H. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc. Natl. Acad. Sci.* **112**, 11870–11875 (2015).

45. Ishii, T. *et al.* RNA-guided endonuclease – in situ labelling (RGEN-ISL): a fast CRISPR/Cas9-based method to label genomic sequences in various species. *N. Phytol.* **222**, 1652–1661 (2019).

46. Tian, M., Zhang, R. & Li, J. Emergence of CRISPR/Cas9-mediated bioimaging: A new dawn of in-situ detection. *Biosens. Bioelectron.* **232**, 115302 (2023).

47. Commons, W. File:DNA Repair after CRISPR-Cas9 cut.svg --- Wikimedia Commons, the free media repository. Preprint at https://commons.wikimedia.org/w/index.php?title=File:DNA_Repair_after_CRISPR-Cas9_cut.svg&oldid=873648677.

48. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Bio* **18**, 285–298 (2017).

49. Tong, X. *et al.* Liquid–liquid phase separation in tumor biology. *Signal Transduction and Targeted Therapy 2022 7:1* **7**, 1–22 (2022).

50. Lyon, A. S., Peeples, W. B. & Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat Rev Mol Cell Bio* **22**, 215–235 (2021).

51. Renger, R. *et al.* Co-condensation of proteins with single- and double-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America* **119**, e2107871119 (2022).

52. Spegg, V. & Altmeyer, M. Biomolecular condensates at sites of DNA damage: More than just a phase. *DNA Repair* **106**, 103179 (2021).

53. Dai, Y., You, L. & Chilkoti, A. Engineering synthetic biomolecular condensates. *Nat. Rev. Bioeng.* **1**, 466–480 (2023).

54. Sabari, B. R. Biomolecular Condensates and Gene Activation in Development and Disease. *Developmental Cell* **55**, 84–96 (2020).

55. Burrows, A., Holman, J., Parsons, A., Pilling, G. & Price, G. *Chemistry3: Introducing Inorganic, Organic and Physical Chemistry*. (Oxford University Press, UK United Kingdom).

56. Jones, Loretta. & Atkins, P. W. *Chemical Principles : The Quest for Insight*. (W.H. Freeman, New York :).

57. Atkins, P., Paula, J. de & Friedman, R. *Physical Chemistry: Quanta, Matter, and Change*. (Oxford University Press).

58. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. (2006). doi:10.1007/978-0-387-46312-4.

59. Rye, H. S. *et al.* Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: properties and applications. *Nucleic Acids Res.* **20**, 2803–2812 (1992).

60. Larsson, A., Carlsson, C. & Jonsson, M. Characterization of the binding of YO to [poly(dA-dT)]2 and [poly(dG-dC)]2, and of the fluorescent properties of YO and YOYO complexed with the polynucleotides and double-stranded DNA. *Biopolymers* **36**, 153–167 (1995).

61. Carlsson, C., Larsson, A., Jonsson, M., Albinsson, B. & Norden, B. Optical and Photophysical Properties of the Oxazole Yellow DNA Probes YO and YOYO. *J. Phys. Chem.* **98**, 10313–10321 (1994).

62. Kundukad, B., Yan, J. & Doyle, P. S. Effect of YOYO-1 on the mechanical properties of DNA. *Soft Matter* **10**, 9721–9728 (2014).

63. Murade, C. U., Subramaniam, V., Otto, C. & Bennink, M. L. Interaction of Oxazole Yellow Dyes with DNA Studied with Hybrid Optical Tweezers and Fluorescence Microscopy. *Biophys. J.* **97**, 835–843 (2009).

64. Günther, K., Mertig, M. & Seidel, R. Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Research* **38**, 6526–6532 (2010).

65. Tegenfeldt, J. O. *et al.* The dynamics of genomic-length DNA molecules in 100-nm channels. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 10979–10983 (2004).

66. Iarko, V. *et al.* Extension of nanoconfined DNA: Quantitative comparison between experiment and theory. *PHYSICAL REVIEW E* **92**, 62701 (2015).

67. Spielmann, H. P., Wemmer, D. E. & Jacobsen, J. P. Solution Structure of a DNA Complex with the Fluorescent Bis-Intercalator TOTO Determined by NMR Spectroscopy. *Biochemistry* **34**, 8542–8553 (1995).

68. Ganji, M. *et al.* Real-time imaging of DNA loop extrusion by condensin. *Science* **360**, 102–105 (2018).

69. Frykholm, K., Nyberg, L. K. & Westerlund, F. Exploring DNA-protein interactions on the single DNA molecule level using nanofluidic tools. *Integrative Biology (United Kingdom)* **9**, 650–661 (2017).

70. KK, S. *et al.* A Parallelized Nanofluidic Device for High-Throughput Optical DNA Mapping of Bacterial Plasmids. *Micromachines* **12**, 1234 (2021).

71. Öz, R. *et al.* Phosphorylated CtIP bridges DNA to promote annealing of broken ends. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 21403–21412 (2020).

72. Alizadehheidari, M. *et al.* Nanoconfined Circular and Linear DNA: Equilibrium Conformations and Unfolding Kinetics. *Macromolecules* **48**, 871–878 (2015).

73. Wang, J. C. & Davidson, N. Thermodynamic and kinetic studies on the interconversion between the linear and circular forms of phage lambda DNA. *J. Mol. Biol.* **15**, 111–123 (1966).

74. Jary, D., Lal, J. & Sikorav, J.-L. Accelerated cyclization of λ DNA. *C. R. l'Académie des Sci. - Ser. III - Sci. Vie* **321**, 1–4 (1998).

75. Trujillo, K. M. *et al.* Yeast xrs2 binds DNA and helps target rad50 and mre11 to DNA ends. *The Journal of biological chemistry* **278**, 48957–48964 (2003).

76. Nyberg, L., Persson, F., Åkerman, B. & Westerlund, F. Heterogeneous staining: a tool for studies of how fluorescent dyes affect the physical properties of DNA. *Nucleic Acids Res.* **41**, e184–e184 (2013).

77. Gautam, D., Pandey, S. & Chen, J. Effect of Flow Rate and Ionic Strength on the Stabilities of YOYO-1 and YO-PRO-1 Intercalated in DNA Molecules. *J. Phys. Chem. B* **127**, 2450–2456 (2023).

78. Reisner, W. *et al.* Nanoconfinement-Enhanced Conformational Response of Single DNA Molecules to Changes in Ionic Environment. *Phys. Rev. Lett.* **99**, 058302 (2007).

79. Roushan, M. *et al.* Motor-like DNA motion due to an ATP-hydrolyzing protein under nanoconfinement. *Scientific Reports 2018 8:1* **8**, 1–8 (2018).

80. Roushan, M., Azad, Z., Lim, S. F., Wang, H. & Riehn, R. Interference of ATP with the fluorescent probes YOYO-1 andYOYO-3 modifies the mechanical properties of intercalator-stained DNA confined in nanochannels. *Mikrochimica acta* **182**, 1561–1565 (2015).

81. Yoo, J. & Aksimentiev, A. Competitive Binding of Cations to Duplex DNA Revealed through Molecular Dynamics Simulations. *J. Phys. Chem. B* **116**, 12946–12954 (2012).

82. Serec, K., Babić, S. D. & Tomić, S. Magnesium ions reversibly bind to DNA double stranded helix in thin films. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **268**, 120663 (2022).

83. Storer, A. C. & Cornish-Bowden, A. Concentration of MgATP2- and other ions in solution. Calculation of the true concentrations of species present in mixtures of associating ions. *Biochem. J.* **159**, 1–5 (1976).

84. Pavlova, E. *et al.* High-throughput single-molecule nanofluidic studies on B. subtilis Rok protein interaction with DNA. *QRB Discov.* **6**, (2025).

85. Wang, L., Pyle, J. R., Cimatu, K. L. A. & Chen, J. Ultrafast transient absorption spectra of photoexcited YOYO-1 molecules call for additional investigations of their fluorescence quenching mechanism. *J. Photochem. Photobiol. A: Chem.* **367**, 411–419 (2018).

86. Guéroult, M., Boittin, O., Mauffret, O., Etchebest, C. & Hartmann, B. Mg2+ in the Major Groove Modulates B-DNA Structure and Dynamics. *PLoS ONE* **7**, e41704 (2012).

87. Møller, H. D. *et al.* Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun.* **9**, 1069 (2018).

88. Mulia, G. E., Picanço-Castro, V., Stavrou, E. F., Athanassiadou, A. & Figueiredo, M. L. Advances in the Development and the Applications of Nonviral, Episomal Vectors for Gene Therapy. *Hum. Gene Ther.* **32**, 1076–1095 (2021).

89. Sihn, C.-R. *et al.* Molecular analysis of AAV5-hFVIII-SQ vector-genome-processing kinetics in transduced mouse and nonhuman primate livers. *Mol. Ther. - Methods Clin. Dev.* **24**, 142–153 (2022).

90. Torchinsky, D. & Ebenstein, Y. Sizing femtogram amounts of dsDNA by single-molecule counting. *Nucleic Acids Research* **44**, e17–e17 (2016).

91. Baranello, L., Levens, D., Gupta, A. & Kouzine, F. The importance of being supercoiled: How DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.* **1819**, 632–638 (2012).

92. Das, S. K. *et al.* MYC assembles and stimulates topoisomerases 1 and 2 in a "topoisome." *Molecular Cell* **82**, 140-158.e12 (2022).

93. Pommier, Y., Nussenzweig, A., Takeda, S. & Austin, C. Human topoisomerases and their roles in genome stability and organization. *Nature Reviews Molecular Cell Biology 2022* 1–21 (2022) doi:10.1038/s41580-022-00452-3.

94. Christensen, M. O. *et al.* Dynamics of human DNA topoisomerases IIα and IIβ in living cells. *J. Cell Biol.* **157**, 31–44 (2002).

95. Morotomi-Yano, K. & Yano, K. Nucleolar translocation of human DNA topoisomerase II by ATP depletion and its disruption by the RNA polymerase I inhibitor BMH-21. *Sci. Rep.* **11**, 21533 (2021).

96. Morotomi-Yano, K., Saito, S., Adachi, N. & Yano, K. Dynamic behavior of DNA topoisomerase IIβ in response to DNA double-strand breaks. *Sci. Rep.* **8**, 10344 (2018).

97. Jeong, J., Lee, J. H., Carcamo, C. C., Parker, M. W. & Berger, J. M. DNA-Stimulated Liquid-Liquid phase separation by eukaryotic topoisomerase ii modulates catalytic function. *Elife* **11**, e81786 (2022).