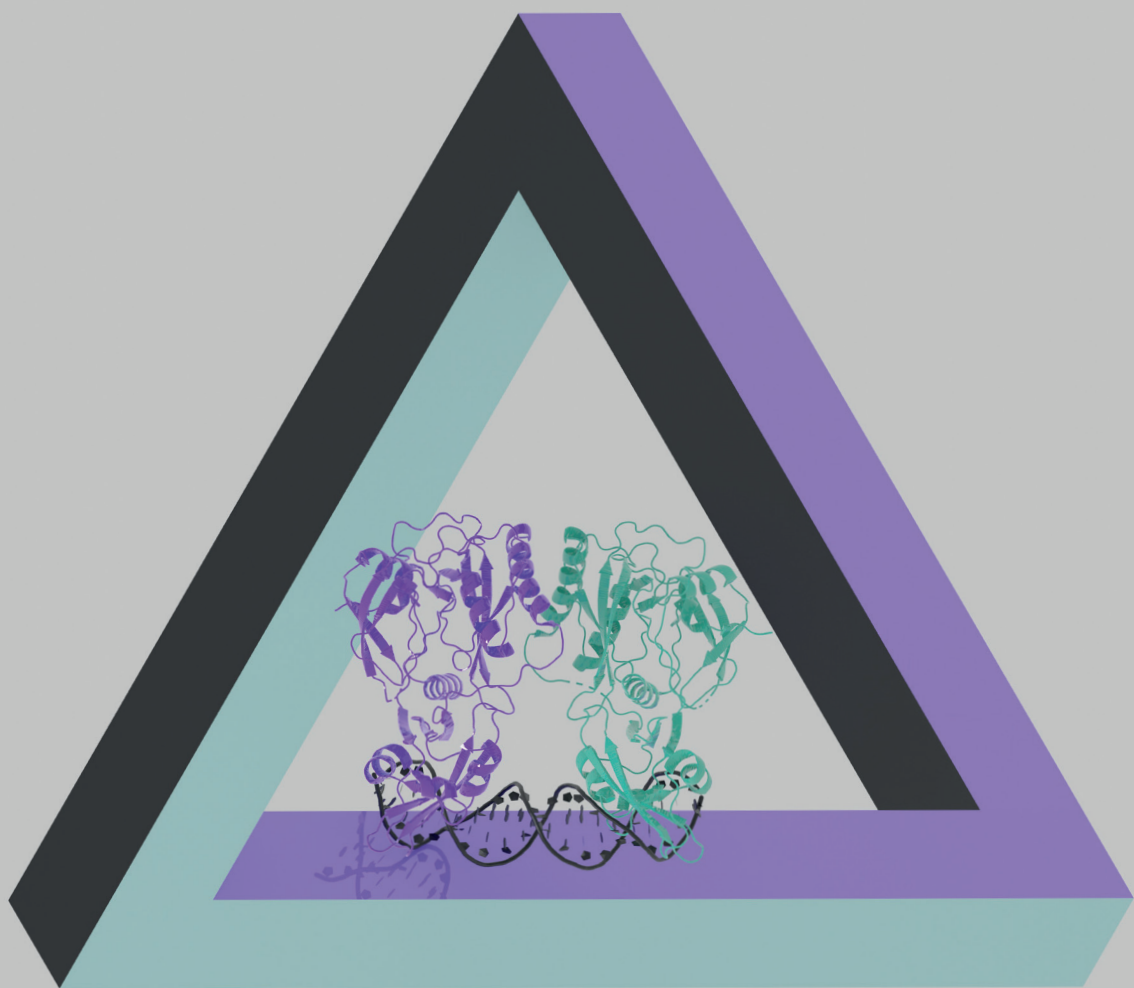# On the interaction between ARFs and AuxREs: from methods to models

Mattia Fontana

Propositions

1. Differences in the homotypic interactions of DBD and PB1 domains determine the affinity of ARFs towards AuxREs of different topology
   (this thesis)

2. The assumptions underlying the derivation of a model decides its domain of applicability and should be stated whenever a model is used
   (this thesis)

3. In selecting experimental methods one must remember that complexity comes at a price

4. Linear transformation of functions for fitting purposes is a relic from the past and should be treated as such

5. With the increasing popularity of workflows based on deep learning it is paramount to be aware of its limitations and pitfalls

6. Free software (free as in freedom) is the only viable escape from technological serfdom

7. Privacy is retreating faster than glaciers

Propositions belonging to the thesis entitled
"On the interaction between ARFs and AuxREs:
from methods to models"
Mattia Fontana
Wageningen, 7 April 2021

On the interaction between ARFs and AuxREs: from methods to models

Mattia Fontana

On the interaction between ARFs and AuxREs: from methods to models

Mattia Fontana

Contents

Chapter 1

Introduction

1

The molecular basis of inheritance became clear only after the discovery of the structure of DNA almost 70 years ago[1]. From this discovery, decades of scientific and technical advances in DNA sequencing allowed the complete genetic information (genome) of increasingly complex organisms to be determined. The first genome of a multicellular eukaryote to be sequenced was to one belonging to the worm *Caenorhabditis elegans* in 1998[2] followed two years later by two other illustrious model organisms, the insect *Drosophila melanogaster*[3] and the flowering plant *Arabidopsis thaliana*[4]. In the last twenty years, the genomes of thousands of eukaryotic species have been sequenced and the amount of sequencing is still increasing exponentially. Despite this incredible wealth of information, our understanding of life is progressing at a much lower pace. The genome contains all the information needed to produce every RNA and protein present in an organism but fails to capture the full complexity of life. In an analogy, we can imagine life as a symphony and the list of genes of a genome as list of musical instruments; here, the full beauty and complexity of the music is generated by the instruments playing together and interacting with each other in coordinated harmony. As the musical instruments of the analogy, also genes need to be turned on and off (expressed or not) in a precise manner to obtain the desired effect. This control of gene expression happens at every step down the *central dogma of molecular biology*[5], which states that genes are transcribed into messenger RNAs which in turn are translated into proteins. Transcriptional control being the first of such controls makes it arguably the most important, as regulation at this stage prevents the synthesis on superfluous intermediates.

The mediators of transcriptional control are called transcription factors (TF, or alternatively transcription regulators) and allow the cell to express only the genes needed to perform a particular function in the organism or to respond to stimuli from the environment (e.g. from other cells). TFs activate or repress transcription by binding specific DNA sequences (usually 6–8 bp, often called cis-regulatory sequences or response element) in the promoter region of the genes that they regulate (target genes). A common feature of eukaryotic TFs is their ability to form dimers, the reason behind this feature is twofold: (I) the short sequence recognized by a TF monomer appears by chance in the genome too often to be specific, so doubling the number of recognized bases exponentially increases specificity; (II) the increased number of stabilizing contacts with the DNA makes the interaction much stronger, showing cooperativity. Moreover, when the dimerization is not very strong, the interaction between the TF and its cis-regulatory sequence displays a

1

steeper response to the increase in concentration of the transcription factor so that the DNA target is bound in a more on/off like fashion.

Transcription factors are abundant in eukaryotic genomes ($\approx$3–10 % of protein-coding genes) and are divided into families based on the presence in their DNA binding domain (DBD) of characteristic motifs[6]. In the genome of *A. thaliana*, are present approximately 2000 TFs divided into $\approx$ 60 families[7]; of these families, about half are plant-specific as they are not related to families found in other eukaryotic lineages[8]. An important family of plant-specific TF is the auxin response factor (ARF) family which represent the main transcriptional effector to the presence of the plant hormone auxin.

Auxin signalling

Auxins are a class of structurally similar plant hormones characterized by an aromatic ring and a carboxyl group. Three auxins have been identified as active and endogenous in plants but indole-3-acetic acid (IAA) is the most abundant and most studied[9–12]; for this reasons the term auxin usually refers to IAA alone. Auxin induces two kind of responses: A fast one, which occurs in seconds to minutes after the application of the hormone and does not involve change in gene expression and a slow one ($>$10 minutes), which is a transcriptional response where genes are either activated or repressed. The mechanism of the fast response is still not completely established, as there is experimental evidence supporting different possible pathways[13]. On the other hand, the transcriptional response is well established and occurs via the nuclear auxin pathway (NAP); for this reason this is also dubbed the canonical pathway[14]. The NAP contains only three players: (I) TRANSPORT INHIBITOR RESISTANT1 / AUXIN SIGNALLING F-BOX (TIR1/AFB) (II) AUXIN / INDOLE-3-ACETIC-ACID (AUX/IAA) (III) AUXIN RESPONSE FACTOR (ARF)[15]. In this pathway, TIR1/AFB and Aux/IAA act as coreceptor for auxin as the hormone physically stabilizes the interaction between the two proteins[16,17]. Since TIR1/AFB is part of the SCF$^{\text{TIR/AFB}}$ ubiquitin ligase complex, the auxin-induced stabilization of the interaction with Aux/IAA leads to the ubiquitinization and proteasome-mediated degradation of the latter. When present, Aux/IAA binds ARF and the corepressor TOPLESS (TPL) preventing the activation of ARF's target genes; then, the effect of auxin is to free ARF from Aux/IAA repression leading to the trascription of ARFs target genes.

1

Fluorescence

Fluorescence is defined as the spontaneous emission of radiation from an electronically excited species which occurs during the irradiation of a substance by electromagnetic radiation[18]. Fluorescence emission ceases within nano/milliseconds after the irradiation stops since the electron in the excited orbital is still paired by opposite spin to the one in the ground state which makes the return to the ground state spin allowed[19].

A molecule that can display fluorescence is called a fluorophore (or fluorescent molecule). From a single-molecule perspective, fluorescence is caused by a fluorophore cycling between a ground electronic state (S_0) and an excited electronic state (S_1). A fluorophore in the ground state can undergo stimulated adsorption of an incoming photon which takes it to an excited state. In the timescale of the electronic transitions, there cannot be rearrangement of the nuclei (Franck-Condon principle); then, the absorption of the photon takes the molecule to an excited vibrational state of the excited electronic state. The molecule can then relax through a radiationless decay to lower vibrational states losing energy to surrounding molecules. The molecule can then return to the ground electronic state emitting a photon. Again the transition being vertical will lead to the molecule in an excited vibrational state of the ground electronic state followed by relaxation. The Franck-Condon principle applies to the molecules of the solvent as well, so that the adsorption of the incoming photon occurs in the solvent arrangement characteristic of the electronic ground state of the fluorophore while the photon emission happens in the solvent arrangement characteristic of the electronic excited state. Then, the relaxation of the solvent plays a role in defining the energy levels involved in the transitions and exert an effect on the adsorption and emission spectra of a dye. The presence of the two relaxation steps (of both molecule and solvent) cause the emitted photons to have lower energy (longer wavelength) than the excitation photon, an effect quantified by the *Stokes shift*.

Single-molecule fluorescence detection and TIRF microscopy

Single-molecule fluorescence detection (SMFD) allows to study biomolecular interactions and conformational dynamics while enabling the identification of subpopulations and transient intermediates[20–23]. In order to achieve a signal-to-noise ratio (SNR) compatible with SMFD, the excitation is confined to a small volume of the sample using high NA (numerical aperture) objectives in confocal or total

1

internal reflection configurations[24]. Total internal reflection fluorescent (TIRF) microscopy is based on the behaviour of light at the interface between two isotropic media that are characterized by different refractive index. The angle of incidence of the light that reaches the interface (θ_1) is linked to the angle of the refracted (or reflected) light (θ_2) by the Snell's law

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} \tag{1.1}$$

where n_1 and n_2 are the refractive indices of the two media. If light is travelling from a medium with higher reflective index to one with lower reflective index ($n_1 > n_2$), the refracted ray enters the second medium with wider angle than the incoming ray. As θ_1 increases, it reaches a value for which $\theta_2 = 90°$, called the critical angle. For values of θ_1 bigger than the critical angle, the ray is reflected back into the first medium but it creates an evanescent field at the interface that decreases exponentially into the second medium. This effect generates a very shallow profile of illumination ($\approx 200\,\mathrm{nm}$) that can excite fluorophores that are very close or tethered to the surface[25,26].

Förster resonance energy transfer

Förster resonance energy transfer (FRET) describes the non-radiative energy transfer from a donor fluorophore, present in an excited electronic state, to an acceptor chromophore (often a fluorophore), present in the ground electronic state. The efficiency E of this energy transfer is dependent on the distance between the donor and the acceptor fluorophores R according to the following function[27,28]

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6}, \tag{1.2}$$

where R_0 is the Förster radius, a constant that can be calculated as

$$R_0 = \sqrt[6]{\frac{9 \ln 10}{128 \pi^5 N_A} \frac{\kappa^2 \Phi_{D,0}}{n^4} J}, \tag{1.3}$$

where n is the refractive index of the medium, $\Phi_{D,0}$ is the fluorescence quantum yield of the donor in absence of FRET, κ^2 is the orientation factor, N_A is the Avogadro constant and J is the spectral overlap integral calculated as

$$J(\lambda) = \int_0^\infty F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda, \tag{1.4}$$

where $F_D(\lambda)$ is the normalized fluorescence of the donor and $\epsilon_A(\lambda)$ is the extinction coefficient of the acceptor. Each of these three equations is connected to one of the three requisites to observe FRET between two fluorophores: (I) there must be spectral overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor, (II) the orientation factor κ^2 must be bigger the 0, (III) the dyes needs to be at distance close to the value of their R_0, typical values of R_0 for organic dyes like the ones used throughout this thesis is $\approx$5–7 nm, which places the range in which FRET can occur to 3–10 nm. κ^2 can be calculated from the relative orientation of the transition dipole moments of the donor and acceptor but in the vast majority of the application this is not possible as the fluorophores are usually able to rotate freely and its value changes accordingly over time. A common solution to this problem is to assume the dyes to be in an isotropic dynamic averaging regime; in this regime, the dyes are expected to take all possible relative orientations during the energy transfer which leads to a mean value of $\kappa^2 = 2/3$.

There are two main ways to calculate the FRET efficiency E; the first is based on measurements of the lifetime of the donor in presence ($\tau_{D,A}$) or in absence of the acceptor ($\tau_{D,0}$)

$$E = 1 - \frac{(\tau_{D,A})}{(\tau_{D,0})};$$ (1.5)

the second method, used throughout this thesis, is based on the measurements of the emission intensities (number of photons) collected from the donor (DD) and from the acceptor (DA) during a donor excitation

$$E^* = \frac{DA}{DD + DA}.$$ (1.6)

The second method returns a value of FRET efficiency (E^*) that needs to be corrected for several factors if the aim of the experiment is to obtain accurate distances between the dyes[29,30].

Microfluidics

Microfluidics is the study or design of devices that have features with a characteristic length in the order of microns to millimeters. The physical behaviour of fluids at the microscale can be understood through the definition of dimensionless numbers that express the relative importance of different physical phenomena[31,32]. In the field of microfluidics, the most important and ultimately least relevant number is

1

the Reynolds number, Re, defined as

$$\text{Re} \equiv \frac{\rho U_0 L_0}{\eta} \tag{1.7}$$

where U_0 is the flow velocity inside the microchannel, L_0 is the characteristic length of the microchannel, ρ is the density of the fluid and η is the dynamic viscosity of the fluid. Re represent the ratio between inertial and viscous forces; systems where $\text{Re} \gg 1$ are dominated by inertia and display turbulent flow whereas systems characterized by $\text{Re} \ll 1$, like microfluidic devices, are dominated by viscosity and show laminar flow. The importance of Re relies in the fact that it shows how the microfluidics world is governed by viscosity whereas inertial effects are irrelevant; on the other hand, inside the microfluidic world Re itself becomes irrelevant as it tells nothing about the characteristic of a particular microfluidic device.

When microfluidic devices are designed to work in combination with camera-based single-molecule fluorescence detection, they allow the tracking of single fluorescent molecules as they pass through the field of view. The motion of the molecules is governed by advection and diffusion. Advection is the movement of the fluorescent molecules caused by the flow of the solution while diffusion is a random movement of the fluorescent molecules caused by the collisions with the molecules of the medium. Then, the information about the fluorescent molecule movement can be used to calculate its diffusion coefficient as well as the characteristics of the flow inside the microfluidic device[33]. In particular, the determination of the diffusion coefficient D is very informative as it is related to the radius of the fluorescent molecule r according to the Einstein-Stokes equation

$$D = \frac{k_B T}{6 \pi \eta r} \tag{1.8}$$

where k_B is the Boltzmann's constant, T is the temperature and η is the dynamic viscosity of the fluid.

Aim and outline of this thesis

In the past 20 years, the interaction between ARF and its response element has been studied extensively with a number of different experimental methods leading to a good understanding of ARF binding preferences. Despite this, little quantitative information was available about the strength and the kinetics of this interaction. The aim of this thesis is to develop methods that allow for a quantitative take

on the determination of ARF-DNA affinity and to derive appropriate models for it.

Chapter 2 reports the development of a quantitative method for the determination of the affinity and kinetics of ARF-DNA interaction using smFRET and smPIFE-FRET.

Chapter 3 describes the use of the method based on smPIFE-FRET to determine the effect of dimerization and oligomerization on the affinity and kinetics of ARF toward its DNA response element. Moreover it contains the derivation of an association model that describes the experimental data.

Chapter 4 presents a method for single-molecule fluorescence detection of labelled biomolecules using microfluidic devices featuring nanochannels.

Chapter 5 describes the application of nanochannels to study ARF-ARF interaction and ARF-DNA interaction without DNA immobilization.

Chapter 6 summarizes and discusses the main findings presented in this thesis and highlights possible future directions for the continuation of the research.

Chapter 7 contains a summary of this thesis.

References

1. Watson, J. & Crick, F. Molecular structure of nucleic acids. *Nature* **171,** 737–738 (1953).
2. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science* **282,** 2012–2018 (1998).
3. Adams, M. D. *et al.* The Genome Sequence of Drosophila melanogaster. *Science* **287,** 2185–2195 (2000).
4. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408,** 796–815 (2000).
5. CRICK, F. Central Dogma of Molecular Biology. *Nature* **227,** 561–563 (1970).
6. Gonzalez, D. H. in *Plant Transcription Factors* (ed Gonzalez, D. H.) 3–11 (Academic Press, Boston, 2016).
7. Hong, J. C. in *Plant Transcription Factors* (ed Gonzalez, D. H.) 35–56 (Academic Press, Boston, 2016).
8. Yamasaki, K. in *Plant Transcription Factors* (ed Gonzalez, D. H.) 57–72 (Academic Press, Boston, 2016).
9. Simon, S. & Petrášek, J. Why plants need more than one type of auxin. *Plant Science* **180,** 454–460 (2011).
10. Korasick, D. A., Enders, T. A. & Strader, L. C. Auxin biosynthesis and storage forms. *J Exp Bot* **64,** 2541–2555 (2013).
11. Lavy, M. & Estelle, M. Mechanisms of auxin signaling. *Development* **143,** 3226–3229 (2016).
12. Cook, S. D. An Historical Review of Phenylacetic Acid. *Plant Cell Physiol* **60,** 243–254 (2019).
13. Kubeš, M. & Napier, R. Non-canonical auxin signalling: fast and curious. *J Exp Bot* **70,** 2609–2614 (2019).

1

14. Leyser, O. Auxin Signaling. *Plant Physiology* **176,** 465–479 (2018).

15. Weijers, D. & Wagner, D. Transcriptional Responses to the Auxin Hormone. *Annual Review of Plant Biology* **67,** 539–574 (2016).

16. Tan, X., Calderon-Villalobos, L. I. A., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M. & Zheng, N. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446,** 640–645 (2007).

17. Calderón Villalobos, L. I. A., Lee, S., De Oliveira, C., Ivetac, A., Brandt, W., Armitage, L., Sheard, L. B., Tan, X., Parry, G., Mao, H., Zheng, N., Napier, R., Kepinski, S. & Estelle, M. A combinatorial TIR1/AFB-Aux/IAA co-receptor system for differential sensing of auxin. *Nature Chemical Biology* **8,** 477–485 (2012).

18. McNaught, A., Wilkinson, A., Pure, I. U. o. & Chemistry, A. *Compendium of Chemical Terminology: IUPAC Recommendations* (Blackwell Science, 1997).

19. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* (Springer, 2006).

20. Smiley, R. D. & Hammes, G. G. Single Molecule Studies of Enzyme Mechanisms. *Chem. Rev.* **106,** 3080–3094 (2006).

21. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in Single-Molecule Fluorescence Methods for Molecular Biology. *Annu. Rev. Biochem.* **77,** 51–76 (2008).

22. Claessen, V. I., Engelkamp, H., Christianen, P. C., Maan, J. C., Nolte, R. J., Blank, K. & Rowan, A. E. Single-Biomolecule Kinetics: The Art of Studying a Single Enzyme. *Annual Rev. Anal. Chem.* **3,** 319–340 (2010).

23. Küchler, A., Yoshimoto, M., Luginbühl, S., Mavelli, F. & Walde, P. Enzymatic reactions in confined environments. *Nature Nanotech* **11,** 409–420 (2016).

24. Walter, N. G., Huang, C.-Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat Methods* **5,** 475–489 (2008).

25. Axelrod, D. in *Methods in Enzymology* 1–33 (Academic Press, 2003).

26. Martin-Fernandez, M. L., Tynan, C. J. & Webb, S. E. D. A 'pocket guide' to total internal reflection fluorescence. *Journal of Microscopy* **252,** 16–22 (2013).

27. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **437,** 55–75 (1948).

28. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. in *Methods in Enzymology* (ed Walter, N. G.) 455–514 (Academic Press, 2010).

29. Lee, N. K., Kapanidis, A. N., Wang, Y., Michalet, X., Mukhopadhyay, J., Ebright, R. H. & Weiss, S. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophysical Journal* **88,** 2939–2953 (2005).

30. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43,** 1156–1171 (2014).

31. Squires, T. M. & Quake, S. R. Microfluidics: Fluid physics at the nanoliter scale. *Rev. Mod. Phys.* **77,** 977–1026 (2005).

32. Bruus, H. Acoustofluidics 1: Governing equations in microfluidics. *Lab on a Chip* **11,** 3742 (2011).

33. Fontana, M., Fijen, C., Lemay, S. G., Mathwig, K. & Hohlbein, J. High-throughput, non-equilibrium studies of single biomolecules using glass-made nanofluidic devices. *Lab Chip* **19,** 79–86 (2019).

1

Chapter 2

Studying affinity and kinetics of ARF-DNA interaction using smFRET and smPIFE-FRET

Mattia Fontana, Elmar van der Wijk, Simon Lindhoud, Willy van den Berg, Dolf Weijers and Johannes Hohlbein

2

Abstract

Auxin is a plant hormone that has been linked to almost all developmental processes in higher plants. The main effectors to Auxin at the transcriptional level are the transcription factors of the Auxin Response Factor (ARF) family. Despite their importance, little quantitative data is available on the binding affinity and kinetics of these transcription factors towards their response element. In this chapter, we developed single-molecule DNA-binding assays that use Förster resonance energy transfer alone or in combination with protein-induced fluorescence enhancement to quantitatively determine the binding affinity and kinetics of different ARF DNA-binding domains (ARF-DBDs) toward their response element. First, we validated these assays on previously reported *Arabidopsis thaliana* ARF-DBDs interactions with DNA. Then, we studied the binding preference of different *Marchantia polymorpha* ARF-DBDs. We found that ARF preference in binding is maintained in this early-divergent land plant, with the sequence TGTCGG being bound as much or even tighter than the canonical TGTCTC. Moreover, TGTCAA is bound with low affinity. Our finding shows that doubly labelled DNA assays offer a reliable, simple and yet powerful way of studying ARF-DNA interaction at the single-molecule level. In fact, we believe that the design principles of our assays can be applied to characterise any transcription factor/response element interaction.

Introduction

To better understand the function of transcription factors (TFs), an accurate knowledge of their binding affinity to their response elements (REs) on genomic DNA is crucial. The importance of this scientific quest led to the development of assays such as SPR (surface plasmon resonance[1]), EMSA (electrophoretic mobility shift assay[2]), Chip-Seq (chromatin immunoprecipitation sequencing[3]), PBM (protein binding microarray[4,5]), SELEX (systematic evolution of ligands by exponential enrichment[6]), and more recently DAP-Seq (DNA affinity purification sequencing[7]) and SIPDI (SYBR gold-based in vitro protein-DNA interaction[8]). All of these methods evaluate the binding of TFs, but vary greatly in the kind of information they can provide. Some methods provide genome-wide insight on protein-bound motifs (e.g., Chip-Seq and DAP-Seq) but suffer from a poor correlation between the obtained occupancy values and transcriptional activity[9]. Other methods (e.g., PBM, SPR, EMSA) report values of affinity and sometimes kinetics of binding but are limited in their accuracy (PBM), repeatability (SPR) and experimental feasibility (EMSA).

Many of these methods have been applied to study the interaction between the members of the Auxin Response Factor (ARF) transcription factor family. ARF were initially identified as being able to bind the auxin-responsive element (AuxRE) TGTCTC in EMSA experiments[10,11], later experiments using PBM returned TGTCGG as an even higher affinity binding site for the DNA Binding Domain (DBD) of *A. thaliana* ARF1, 5[12] and 3[5] while TGTCAA was found to be weakly bound by AtARF1-DBD and AtARF5-DBD[12]. Further studies based on DAP-Seq confirmed the preference of AtARF2 and 5[7,13] and several maize ARFs towards[14] TGTCGG; the same preference was found using other techniques like SIPDI[8] and luciferase assay[15] while testing several ARFs from various species. The wealth of information about ARFs binding preferences coming from these methods, although extremely valuable, fails to be quantitative about the affinities and do no contain any information about the kinetics of the system. We sought a method that would be simple, reliable, accurate and able to quantitatively asses both affinity and kinetics of binding. To this end, we based our approach on single-molecule Förster Resonance Energy Transfer (smFRET) alone or in combination with Protein-Induced Fluorescence Enhancement (smPIFE-FRET).

Since its introduction[16] in 1996, smFRET has been extensively used in the life sciences to probe intra- and inter-molecular conformations and dynamics of biomolecules[17]. SmFRET is a distance-dependent, non-radiative energy transfer between a donor fluorophore and an acceptor chromophore that can precisely and accurately report on distances between 3 and 10 nm[18], on a timescale that ranges from µs to minutes (depending on the microscope configuration and the type of sample[19]). These length and time scales are well suited to investigate the interactions and conformations of biomolecules; for this reason, smFRET is often referred to as a spectroscopic or molecular ruler. The lifetime of TFs on the DNA *in vitro* is typically in the range from 10^{-1} s to 10^5 s (reviewed in [20]). For analysing such slow dynamics, smFRET experiments are usually performed using surface-immobilized molecules and total internal reflection fluorescence (TIRF) microscopy[21] in combination with camera based detection to record hundreds of molecules in parallel[22]. The immobilization allows following individual molecules up to tens of minutes; these time series (called single-molecule time traces) can then be analysed using, for example, Hidden Markov Modelling (HMM) to obtain information about the kinetics of the system under study[23–28].

PIFE is an additional photophysical effect that can be used as a spectroscopic ruler to indicate changes in distances smaller than 3 nm[29–31]. PIFE relies on the modulation of the quantum yield of a suitable dye by the presence of a protein in its proximity. The molecular mechanism behind this modulation has been described in detail for the carbocyanine dye Cy3[32]. Cy3 molecules are excited from the thermodynamically favored *trans* ground-state isomer to the *trans* excited state isomer. The electronic configuration of the excited state allows for easier rotation around any of the central C-C bonds[33,34]. This rotation passes through a 90° twisted excited-state isomer that represents the starting point for an efficient deactivation route towards both *cis* and *trans* ground-state isomers[35]. Hence, the rotation (i.e. photoisomerization) decreases the quantum yield of the dye; conversely, any steric interaction that would limit the probability of photoisomerization results in an increased quantum yield i.e. PIFE.

As the PIFE mechanism of Cy3 is well understood and Cy3 can be combined with a red-shifted acceptor fluorophore to form a donor-acceptor FRET pair, the combination of smPIFE and smFRET in a single experiment is intriguing[36]. Several different implementations of smPIFE-FRET have been reported in literature; either aimed to combine the sensitive range of both methods resulting in a ruler that

spans from angstroms to nanometers[37] or to simultaneously monitor two distances without the need of complex multi-color approaches[38,39]. Cy3B is a fluorophore that is structurally similar to Cy3 but in which photoisomerization is sterically forbidden thanks to the addition of a rigid backbone around the central trimethine bond[40]. This structural property makes Cy3B largely insensitive to environmental effects affecting the dye's conformational dynamics. As a result, the brightness of Cy3B is maximised and tailing the one of Cy3 at maximum PIFE. Due to its superior photostablility and brightness, Cy3B is an excellent choice as a donor fluorophore in smFRET experiments; moreover, results obtained with Cy3B can be used to disentangle PIFE from FRET in smPIFE-FRET experiment in which Cy3 was used as donor[38].

In this chapter, we developed and applied a reliable DNA binding assay to measure the binding affinities and kinetics of ARF towards its response element in a quantitative manner. The method, based on smFRET alone or in combination with smPIFE, was tested on known interactions between *A. thaliana* ARFs and the DNA and was subsequently used to investigate the binding preference of two *M. Polymorpha* ARFs.

Materials and methods

Protein expression and purification

Protein expression and purification was carried out following the protocol used in reference [12] and [41]. Briefly, the genomic regions corresponding to the DNA binding domain (DBD) of *Arabidopsis thalania* ARF1, ARF2, ARF5 and *Marchantia polymorpha* ARF1 and ARF2 were amplified and cloned in an modified expression vector pTWIN1 (New England Biolabs) to generate fusions with the Chitin Binding Domain (CBD) and Intein. ARF-CBD fusion proteins were expressed in *E. coli* strain Rosetta DE3 (Novagen). Cells were inoculated in Difco Terrific Broth (BD), supplemented with ampicillin and grown to an OD_{600} of 0.5–0.7, protein expression was induced by adding IPTG and the temperature was switched from $37\,°C$ to $20\,°C$; the growth was continued for $20\,h$. Cells were harvested by centrifugation and resuspended in extraction buffer ($20\,mM$ Tris, $500\,mM$ NaCl, $1\,mM$ EDTA, $0.1\,\%$ NP-40 and $2\,mM$ MgCl2, pH 7.8, $10\,mg$ of DNase and $0.2\,mM$ PMSF). Cells were then lysed by passing the suspension twice through a French Pressure cell and cell-free extract was generated by centrifugation. The supernatant was loaded onto a

chitin column (New England Biolabs) and washed with 10 column volumes washing buffer (20 mM Tris, 500 mM, pH 7.8) using an AKTA explorer 100 (GE Healthcare). ARF-DBD proteins were eluted by 1 h incubation with 40 mM DTT in washing buffer. Proteins were concentrated into a final volume of approximately 1 mL using Amicon ultra-15 10K spin filters, and next passed over a Superdex 200PG size-exclusion chromatography column. ARF-DBD proteins were eluted using washing buffer with 1 mM DTT, concentrated using Amicon ultra-15 10K spin filters and stored until use at $-80\,^{\circ}$C.

Accessible volume simulations

The web server 3D-DART[42] was used to model a standard B-DNA structure for the 37 bp oligos used in the smPIFE-FRET experiments ("free" DNA structure) and to extract the geometrical information (roll, tilt and twist between consecutive base pairs) from the short 21 bp DNA present in the crystal structure of ARF1-DBD (PDB ID:4LDX). The smPIFE-FRET oligo was then remodelled using 3D-DART this time replacing, for the overlapping region, the standard B-DNA geometry with the one extracted from the crystal structure; this 37 bp bent oligo was then docked (rigid body docking) in place of the original oligo in the crystal structure of ARF1 using PyMOL ("bound" DNA structure).

The standard B-DNA oligo and the bent oligo bound to ARF1-DBD were used as starting points to model the accessible volumes (AVs) of the FRET pair using the FRET-restrained positioning and screening (FPS) software[43]. For the geometrical parameterization of the dyes, the dimension reported in table 2.1 were used. FPS was then used to calculate the distance between average dyes position (R_{mp}), the distribution of distances between dyes positions (R_{DA}) and the expected mean FRET efficiency (E). An in-house/custom-written MATLAB function was used to calculate the volume of the AVs.

DNA constructs

Single strand DNA oligos were ordered from Eurogentec. Each strand contained a 5-C6-amino-dT modification at the desired position for subsequent labelling. Some of the strands were purchased biotinylated at their 5'- end to allow for surface im-mobilization using a Neutravidin bridge. Strands were labelled with the desired dye (NHS-ester) following a modified version of the protocol provided by the dye man-

Table 2.1: AV simulations: Geometrical parameters of the dyes[43,44]

Dye	Linker	Linker length [Å]	Linker width [Å]	R_1 [Å]	R_2 [Å]	R_3 [Å]
Cy3	C6-NHS	20.0	4.5	6.8	3.0	1.5
Cy3B	C6-NHS	14.2	4.5	8.2	3.3	2.2
ATTO647N	C6-NHS	17.8	4.5	7.4	4.8	2.6

ufacturer and purified using polyacrylamide gel electrophoresis (20 % Acrylamide). DNA constructs were annealed by heating complementary single strands to 95 °C in annealing buffer (250 mM NaCl, 10 mM Tris HCl pH 8, 1 mM EDTA) followed by cooling down to room temperature performed overnight. The sequence of the oligo used throughout this chapter is 5'-AGTCTTTTTGGTCTCTGGTCGGTCGACAATCCGTGTG-3' with the exception of the titration to determine the binding preferences of MpARFs where the sequence 5'-ACTCTTTTTTGTCNNGGAAAGGNNGACAATCCGTGTG-3' was used (NN being either TC, GG or AA).

Single-molecule FRET and PIFE-FRET

Imaging was carried out on a home-built TIRF microscope, described previously[45]. The measurements were performed using alternating-laser excitation (ALEX)[36]; in this excitation scheme, an acquisition frame during which the donor is excited is followed by a frame in which the acceptor is directly excited and so on. The emission of the fluorophores is spectrally divided into two different detection channels on the emCCD camera sensor (Andor iXon 897 Ultra). This approach creates four photon streams, three of which are relevant; (1) donor emission after donor excitation (DD), (2) acceptor emission after donor excitation (DA, arising from FRET) and (3) acceptor emission after acceptor excitation (AA). The three photon streams can be used to calculate the raw FRET efficiency ($E^* = DA/(DD + DA)$) and stoichiometry ($S = (DD + DA)/(DD + DA + AA)$)[46]. E^* contains the information about the relative distance of the two fluorophores whereas S contains information about the photophysical state of a given molecule (allowing to filter out molecules missing an active donor or an active acceptor) and the information about the spatial confinement in case of environmental-sensitive dyes (PIFE). The camera acquisition time and the excitation time were set to 250 ms; laser powers were set to 1.5 mW

(Cy3) or $0.5\,\text{mW}$ (Cy3B) for green ($\lambda = 561\,\text{nm}$) and $0.3\,\text{mW}$ (Cy3) and $0.5\,\text{mW}$ (Cy3B) for red ($\lambda = 638\,\text{nm}$) lasers. The data reported in figure 2.2 is an initial characterization of the system recorded at $50\,\text{ms}$ and laser powers set to $8\,\text{mW}$ (Cy3) or $1.5\,\text{mW}$ (Cy3B) for green ($\lambda = 561\,\text{nm}$) and $0.75\,\text{mW}$ for red ($\lambda = 638\,\text{nm}$) lasers. The imaging buffer contained $137\,\text{mM}$ NaCl, $2.7\,\text{mM}$ KCl, $10\,\text{mM}$ phosphate, $1\,\text{mM}$ Trolox, $1\,\%$ gloxy and $1\,\%$ glucose.

Single-molecule titration experiments

Labelled dsDNA oligos were immobilized on a PEGylated glass coverslip as described previously[47]. In particular, the PEGylation was carried out inside the wells of silicone gaskets placed on the coverslip (Grace Bio-labs). Each titration was performed using a single well, washing it between data points with $600\,\mu\text{L}$ of 1x PBS buffer. Typically, each data point of a titration consists of three movies (1,000 frames each). In the experiments with MpARFs, a waiting step of 5 min was added before starting the acquisition of the first movie to allow the system to equilibrate.

Time traces analysis

First, the time traces from individual DNA molecules were filtered to remove the sections in which either the donor or the acceptor were inactive (due to fluorophore bleaching or blinking). A molecule was allowed to have values of E^* and S outside the thresholds (typically 0 to 0.8 for E^* and 0.3 to 0.7 for S) for a maximum of three consecutive data points; longer stays outside the thresholding range resulted in the trace being interrupted. In case the molecule reentered the allowed range for E^* and S, the data points were saved as a new trace. A minimum length of the traces was set to 50 data points. The filtered time traces were then loaded in the software package ebFRET to perform and empirical Bayesian Hidden Markov Modelling[25]. The analysis was performed assuming two states, with two restarts and a convergence threshold of 10^{-6}. The results of the analysis were exported as '.csv' and the transition matrix was used to calculate k_{on}, k_{off} and $K_{\text{d}} = k_{\text{off}}/k_{\text{on}}$.

Binding isotherms analysis

For each datapoint i in a titration, the fit of the FRET efficiency distribution with the two Gaussian pertaining to the free and bound populations returns an uncorrected bound fraction: $F_{\text{B}}^u(i)$. Even in absence of ARF, the Gaussian fit centered on the

E^* of the bound DNA returns an apparent bound fraction $F_{\mathrm{B}}^u(0) > 0$ (typically ≈ 0.1). This value is an indication of the error connected to the two population fit and can be later used to correct the entire titration under the assumption that in case the DNA would be completely bound, the expected apparent free fraction would have the same value ($F_{\mathrm{B}}^u(0) = F_{\mathrm{F}}^u(\infty)$). Then the corrected bound fraction for each datapoint can be calculated as

$$F_{\mathrm{B}}^c(i) = \frac{F_{\mathrm{B}}^u(i) - F_{\mathrm{B}}^u(0)}{1 - 2F_{\mathrm{B}}^u(0)}. \tag{2.1}$$

The corrected bound fraction (henceforth F_{B}) can be fit with the appropriate mathematical model for the interaction, in this case a simple binding isotherm

$$F_{\mathrm{B}} = \frac{[\mathrm{ARF}]}{K_{\mathrm{d}} + [\mathrm{ARF}]}. \tag{2.2}$$

In order to obtain a reliable confidence interval for the K_{d}, a set of 1000 random bootstrap samples, each containing n titrations, were generated from the n experimental titrations by stratified resampling[48,49] (a value of F_{B} for each concentration i is drawn n times from the original n values, with replacement) and fitted with the binding isotherm. The distribution of the obtained 1000 K_{d}s can be fitted with a log-normal distribution to obtain mean value and 95 % confidence interval.

Results

DNA substrate and experimental design

To test the sequence-dependent binding of ARF to DNA, we chose a 37 bp sequence from the promoter of the gene TMO5 of *Arabidopsis thaliana* known to be bound by ARF5 *in vivo*[50]. The sequence consists of two AuxREs (Fig. 2.1a, highlighted in grey), in an inverted configuration and separated by 7 bp (called IR7).

smFRET and smPIFE-FRET necessitates care to be taken in the choice of the positions where the donor and acceptor fluorophore will be attached (labeling positions); a good pair of labeling positions should lead to a noticeable change in FRET efficiency (E) and increase of the fluorescent quantum yield of Cy3 upon protein binding. In case of ARFs, the available crystal structure of the DNA binding domain (DBD) of *A. thaliana* ARF1 bound to an IR7[12] provides valuable prior structural information. Modeling different potential labelling positions using the FRET-restrained positioning and screening (FPS[43]) software (see Materials and Methods)

allowed us to optimize the expected change in FRET efficiency and PIFE. In particular, the software allows to calculate the region of space accessible to the dye (i.e. accessible volume, AV) for every potential labelling position. We calculated the AVs of a donor/acceptor pair in absence and in presence of the protein (Fig. 2.1b-c) allowing us to retrieve the expected change of the distance between the mean position of the dyes (R_{mp}) and the expected relative change in FRET Efficiency (E) upon protein binding. Furthermore, as the presence of the protein reduces the AV of Cy3 for some labelling positions, a first estimation on positions suitable for PIFE can be given. For the labelling positions used in this study (Fig. 2.1), the oligo in absence of protein ("free") has an R_{mp} of 6.70 nm (6.60 nm for Cy3B), an expected E of 0.17 ($R_{0,\text{Cy3}} = 5.1$ nm[38]; 0.40 for Cy3B [$R_{0,\text{Cy3B}} = 6.2$ nm[38]]) and an AV for Cy3 of 13.9 nm^3; upon ARF binding, the R_{mp} increases to 7.8 nm (7.6 nm for Cy3B), E decreases to 0.08 (0.22 for Cy3B) and the AV for Cy3 decreases to 10.5 nm^3. The large decrease of the AV for Cy3 ($> 20\,\%$) and the large decrease of E upon binding ($\Delta E \approx 0.15$) show that the chosen labelling position is a good candidate for both smFRET and smPIFE-FRET experiments.

We note that of the calculated 1.11 nm increase in the distance between the mean position of the dyes only 0.49 nm are caused by the binding-induced bending of the DNA; the remaining 0.62 nm are caused by the steric effect of the proteins reducing the accessible volume of the fluorophores thereby shifting the mean dye position.

Interactions between DNA and unlabelled ARF-DBDs can be detected using smFRET and smPIFE-FRET

We tested the ability of smPIFE-FRET and smFRET to report on ARF binding to the DNA oligo by performing measurements in absence and in presence of saturating concentration of *A. thaliana* ARF5-DBD. In brief, the construct was immobilized on the surface of a coverslip through a neutravidin bridge and imaged on a TIRF microscope using a buffer containing either 0 μM or 2 μM of protein; the resulting movies were analyzed to obtain apparent FRET efficiencies (E^*) and stoichiometries (S) (for details see Materials and Methods). Values of E^* and S are visualized in a 2D histogram (E^*S histogram); in this representation different population of molecules will cluster together according to their labelling and biochemical state (Fig. 2.2). The FRET efficiency contains information about the relative distance of the fluorophores whereas the stoichiometry encodes the information about their
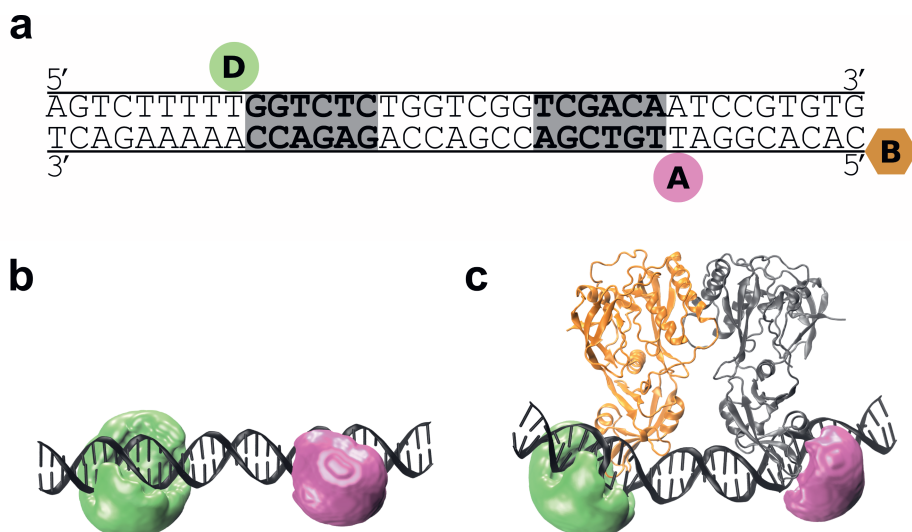
Figure 2.1: dsDNA oligo and accessible volumes (AVs). (a) Schematic representation of the dsDNA oligo; donor (either Cy3 or Cy3B, green) and acceptor (ATTO647N, magenta) indicate the positions for internal labelling; similarly, the biotin (brown) used to immobilize the dsDNA is conjugated to a modified 5'- end. The binding sites for ARF (AuxREs) are highlighted in gray. (b-c) Accessible volumes (AVs) for Cy3 and ATTO647N in absence (b) and presence (c) of AtARF1-DBD. The AVs show the volume that a fluorophore can explore thanks to its flexible linker. The parts of the AV facing each other become partially inaccessible to the dyes upon protein binding; this effect accounts for ($\approx 56\%$) of the expected distance change between the mean position of the dyes, with DNA bending having a smaller effect ($\approx 44\%$ of the total change). The fluorophores being on average further apart when the protein is present will lead to a decrease in FRET efficiency. In addition, as the AV of Cy3 decreases by $\approx 25\%$ upon ARF binding, PIFE is expected.

relative brightness; molecules that are missing either the donor or the acceptor will approach extreme values of stoichiometry and can therefore be filtered out easily (see black boxes in Fig. 2.2a and b). Moreover, the stoichiometry responds to changes in the quantum yield of the fluorophores making it the main observable for monitoring PIFE. The filtered data points can be projected on the 1D side histograms and fitted with single or multiple Gaussians to obtain quantitative information about the relative prevalence of different populations (or states) in the sample; moreover, these 1D histograms (typically the ones representing the distribution of E^*) can be stacked vertically to obtain a stackplot (Fig. 2.2c).

In the smPIFE-FRET experiments, the peak corresponding to the doubly labelled DNA construct in absence of proteins is centered at $E^* = 0.64$ and $S = 0.67$ (Fig. 2.2a, top); upon addition of $2\,\mu$M AtARF5-DBD the peak shifts to a lower FRET efficiency (0.45) in agreement with the expectations from the AV simulations. Moreover, the peak shifts towards an higher stoichiometry value (0.79) due to PIFE (Fig. 2.2a, bottom). SmFRET experiments return similar results (Fig. 2.2b), with a clear change in FRET efficiency (from 0.47 to 0.33) upon binding, but given the characteristics of the FRET pair used (Cy3B/ATTO647N), the change in stoichiometry is small ($\Delta S < 0.02$).

The affinity between AtARF5-DBD and IR7 can be strongly reduced by performing point mutations of key amino acids that mediate the interaction; in particular, the arginine in position 215 was found to be crucial for an high-affinity binding of a canonical IR7[12]. In presence of $2\,\mu$M AtAR5-DBD, almost all DNA is occupied by ARF. In contrast, using $2\,\mu$M AtARF5-DBD R215A only little binding is seen (Fig. 2.2c). These results confirm that the change in FRET efficiency seen upon AtARF5-DBD addition is caused by specific binding of the transcription factor to its response element rather then unspecific interaction of the protein with the DNA.

smPIFE-FRET can reliably quantify differences in ARFs/IR7 affinity

We asked if smPIFE-FRET is able to distinguish the affinity of different ARFs towards an IR7 response element by conducting titrations with increasing concentration of ARF. Assuming a simple binding model for bi-molecular association, the value of the dissociation constant (K_d) coincides with the concentration of ARF at which half of the DNA oligos are bound. Therefore, we are able to obtain the K_d of a given ARF-IR7 combination from the analysis of the distributions of FRET efficiencies.

For AtARF5-DBD, we saw binding at a concentration of $8\,$nM; at $16\,$nM the bound fraction is $\approx 30\,\%$ whereas at $32\,$nM most of the DNA is bound (Fig. 2.3). This finding indicates that the K_d of AtARF5-DBD to the tested IR7 has a value of $\approx 20\,$nM. This tight binding is consistent with the tested IR7 being bound *in vivo* by ARF5[50] . ARF5-DBD carrying the single mutation R215A shows a decrease in affinity of several orders of magnitude with the DNA being mostly free even at a
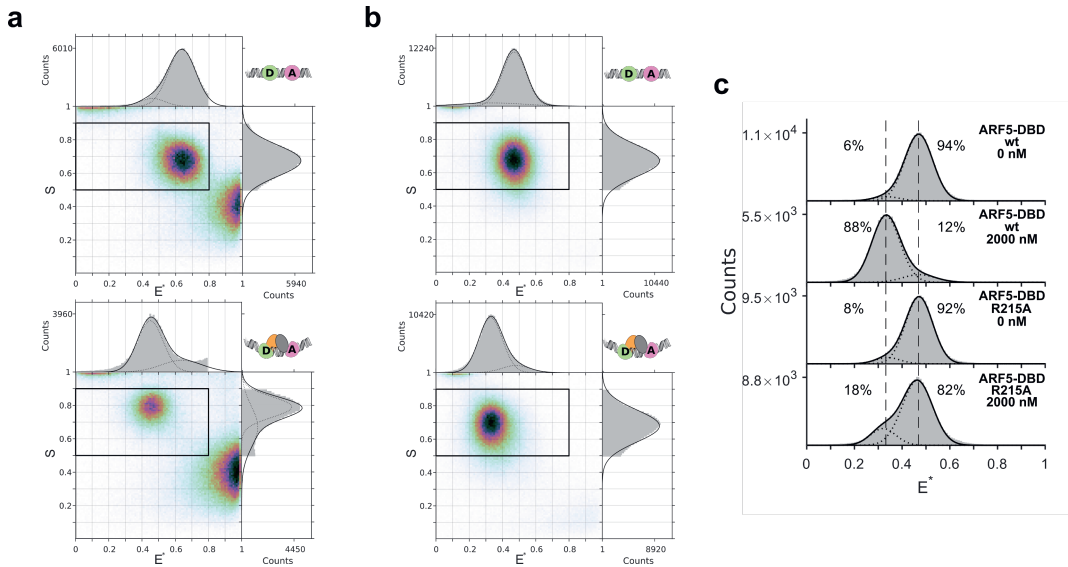
Figure 2.2: ARF-DNA binding detected with smPIFE-FRET and smFRET; E^*S histograms show the FRET efficiency and stoichiometry of fluorescently labelled molecules present in the sample. (a) smPIFE-FRET signature of a DNA when free (top) and when bound to AtARF5-DBD (bottom). A clear shift in E^* and S is visible upon binding as a result of the the increased distance between the average position of the dyes and the increased quantum yield of Cy3. (b) smFRET signature of a dsDNA when free (top) and when bound to AtARF5-DBD (bottom). Because the quantum yield of Cy3B is not affected by the presence of the protein, only the shift in E^* is visible. (c) smFRET experiments carried out in absence or presence of $2\,\mu M$ of AtARF5-DBD or AtARF5-DBD R215A; the small bound population seen upon AtARF5-DBD R215A addition is consistent with the expected lower affinity of this mutant and shows that the shift in E^* seen for the wild-type protein is caused by specific interaction with the response element rather than others unspecific effects.

protein concentration of 2048 nM, placing the K_d for this mutant in the µM range. We then tested ARF1-DBD and ARF2-DBD which showed a K_d of approximately 300 nM.

Taken together these results show that smPIFE-FRET can visualize ARF binding to an IR7 over a large range of ARF concentrations while providing a first estimate for the K_d of the interaction.

2
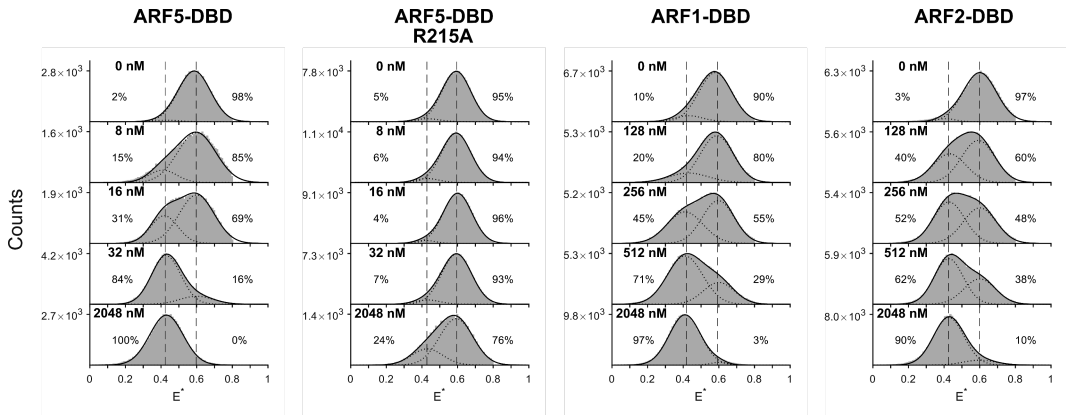
Figure 2.3: Stackplots of smPIFE-FRET titrations of various AtARF-DBDs. AtARF5-DBD shows tight binding with a K_d of $\approx 20\,\text{nM}$. When an amino acid that plays a key role in the interaction between ARF5-DBD and its response element is mutated (R215A), the binding is weakened and the K_d increases to the µM range. AtARF1-DBD and AtARF2-DBD show a similar affinity of approximately 300 nM.

smPIFE-FRET allows to study the kinetics of binding

The analysis of FRET efficiency distributions as shown previously can return values for the binding affinity (K_d) but leave the kinetics of the interactions unknown. Since the DNA construct is bound to the surface of the cover slip, however, we can obtain minute long single-molecule time traces that can be analysed using statistical packages for Hidden Markov Model (HMM) analysis (in this case ebFRET[25], see Materials and Methods) to retrieve the kinetics of binding and dissociation.

In presence of AtARF1-DBD or AtARF2-DBD at concentrations around the K_d, the time traces show clear transitions between the free and bound states (Fig. 2.4a and b respectively); in case of smPIFE-FRET experiments, the transitions can be seen both in the trace reporting the FRET efficiency of the molecule, which decreases upon binding (black trace) and in the trace that reports on the photons coming from the donor, which sees a noticeable increase due to PIFE (green trace). The HMM analysis of the time traces returns a set of parameters that best explain the observed signal (E^*) and the most likely sequence of hidden states (Fig. 2.4a-b, in blue). From these parameters, we calculated the kinetics constants of the system (k_{on} and k_{off}) and, from their ratio, the dissociation constant (K_d). Applying HMM analysis to every data point of a titration returned values for the kinetic

and equilibrium constants for every protein concentration tested (Fig. 2.4c, left); the values coming from these independent analysis were used to calculate the expected values for the constants and estimate their error (Fig. 2.4c, right). AtARF1-DBD and AtARF2-DBD show a similar value of k_{on} $(1.1(4) \times 10^{-4}\,\text{nM}^{-1}\text{s}^{-1}$ and $0.9(3) \times 10^{-4}\,\text{nM}^{-1}\text{s}^{-1})$ with their k_{off} showing a bigger difference $(0.08(2)\,\text{s}^{-1}$ and $0.05(2)\,\text{s}^{-1}$ respectively). The resulting K_{d} show a tighter binding for AtARF2-DBD $(500(50)\,\text{nM})$ compared with AtARF1-DBD $(800(100)\,\text{nM})$. This difference in affinity is driven by the change in k_{off}, while the value of k_{on} stays constant at $\approx 10^5\,\text{M}^{-1}\text{s}^{-1}$; a value within the typical range for association constants between transcription factors and their target $(10^4 - 10^9\,\text{M}^{-1}\text{s}^{-1})$[20].

smFRET allows to quantify the preference of ARF-DBDs towards AuxREs of different sequence

We tested if our method is able to capture preferences of sequence bound by different ARFs by measuring the affinity of two *Marchantia polymorpha* ARFs towards a DNA constructs bearing three variations of the last dinucleotide of the AuxRE.

MpARF1-DBD showed nearly the same affinity towards IR7-TGTCTC and IR7-TGTCGG (16 nM and 13 nM; 95 % CI [13:20] and [10:16] respectively) whereas the affinity towards IR7-TGTCAA was one order of magnitude lower at 140 nM, 95 % CI [120:170] (Fig. 2.5a). MpARF2-DBD showed a similar pattern (Fig. 2.5b), where the binding to IR7-TGTCTC and IR7-TGTCGG was comparable in magnitude (86 nM and 41 nM, respectively) but the difference in binding is statistically significant (95 % CI [72:102] and [32:51] respectively). The low binding affinity towards IR7-TGTCAA is even more pronounced with the value for the K_{d} entering the μM range (1.8 μM, 95 % CI [1.7:2.0]).

These results show, how the TGTCGG element is bound at least as tight as TGTCTC by ARF-DBDs of *M. polymorpha* while TGTCAA presents far weaker interactions. This finding confirms the trends seen in other plant species and suggests that the relative affinity towards different TGTCNN elements might be conserved across ARFs evolution.

2
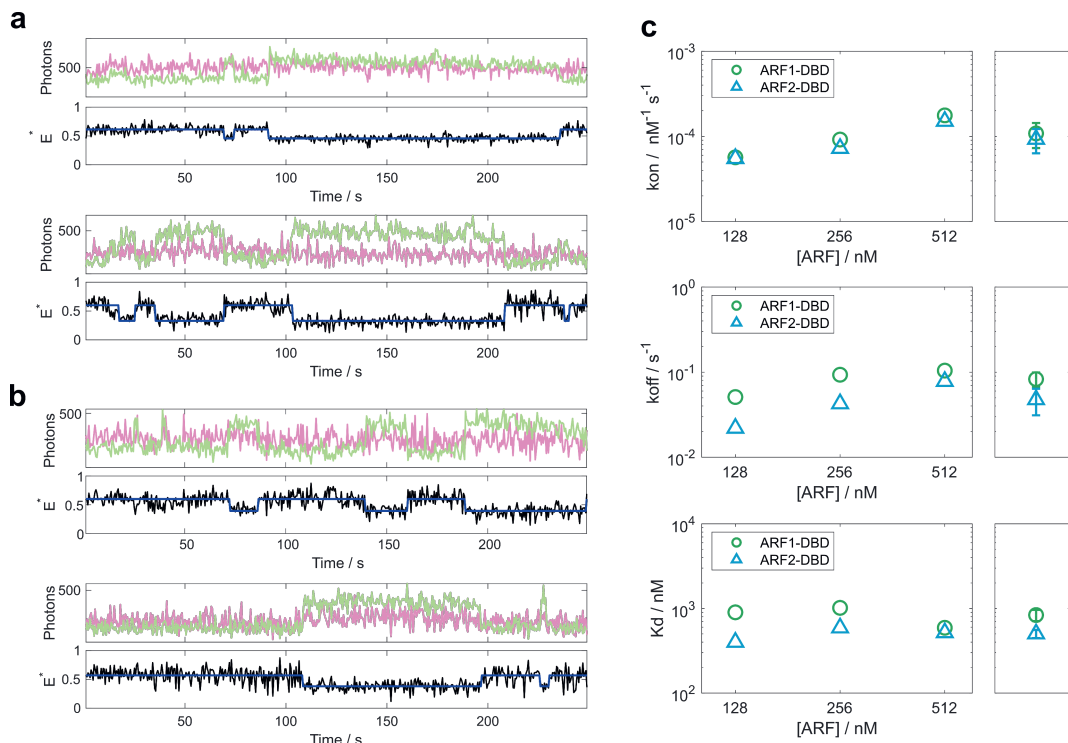
Figure 2.4: HMM analysis of smPIFE-FRET titrations. (a and b) Single-molecule time traces and HMM analysis. Labelled DNA molecules are observed for several minutes and the photons coming from the donor (green trace), the acceptor (magenta trace) and the corresponding FRET efficiency (black trace) are plotted as function of time. The FRET time traces are analyzed using HMM to return the most probable sequence of states (blue trace). Addition of 256 nM of AtARF1-DBD (a) or AtARF2-DBD (b) leads to modulations of the donor intensity (PIFE) and, together with the change in average distance between the donor and the acceptor, generates the observed changes in the FRET efficiency. (c) For each concentration of ARF tested, HMM analysis returns a set of kinetic constants (k_{on} and k_{off}) and, from their ratio, the dissociation constant (K_d); these values can be used to calculate the mean values (and relative standard errors) of the kinetic constants of the system (right pane). The k_{on} of AtARF1-DBD and AtARF2-DBD are similar ($1.1(4) \times 10^{-4}$ nM^{-1}s^{-1} and $0.9(3) \times 10^{-4}$ nM^{-1}s^{-1}). Larger difference are seen in the k_{off} ($0.08(2)$ s^{-1} and $0.05(2)$ s^{-1} respectively) which drives the difference seen in the K_d where AtARF2-DBD appears to bind more tightly (500(50) nM, AtARF1-DBD: 800(100) nM).
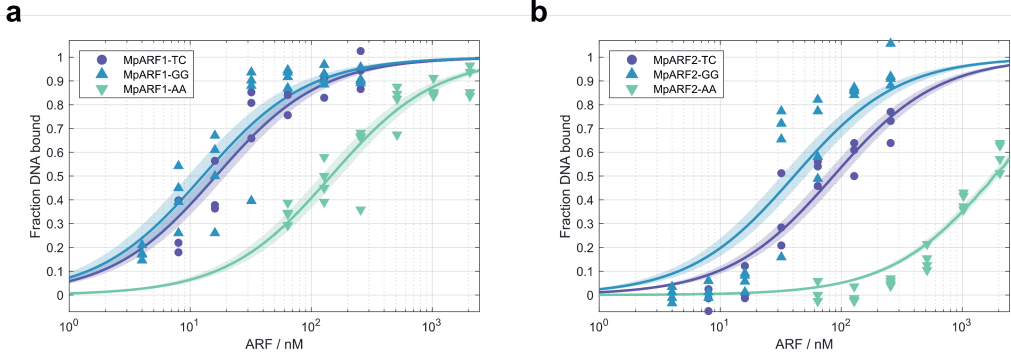
Figure 2.5: smFRET titration of (a) MpARF1-DBD and (b) MpARF2-DBD versus three different IR7 variants. MpARF1-DBD shows strong binding towards both IR7-TGTCTC and IR7-TGTCGG (with a K_d of 16 nM and 13 nM, respectively) but weaker affinity towards IR7-TGTCAA ($K_d = 140$ nM). MpARF2-DBD has slightly higher affinity towards IR7-TGTCGG ($K_d = 41$ nM) compared to IR7-TGTCTC ($K_d = 86$ nM). The binding to IR7-TGTCAA is only noticeable when MpARF2-DBD is added to the solution at µM concentration indicative of a dissociation constant of 1.8 µM. Either three or four independent titrations for each ARF-IR7 combination were acquired. The mean K_ds and relative 95 % confidence intervals are represented as solid lines and shaded areas and were calculated via bootstrapping of the original datasets (see Materials and Methods and Fig. S2.1)

Discussion

In this chapter we presented a DNA binding assay based on single-molecule Förster Resonance Energy Transfer (FRET) alone or in combination with Protein Induced Fluorescence Enhancement (PIFE).

We tested several *A. thaliana* ARF-DBDs and showed that the differences in binding stability and kinetics are captured by the method. The smFRET and smPIFE-FRET experiments in which the affinity of the single amino acid mutation R215A of AtARF5-DBD was probed, showed the expected weaker interaction with the tested IR7. Moreover, the experiments proved that the shift in E^* seen when wt AtARF5-DBD is added in concentration of tens on nM is due to the specific protein interaction with the tested IR7 DNA motif.

The HMM analysis of the single-molecule time traces of AtARF1-DBD and AtARF2-DBD returned a value of k_{on} in the order of $10^5 \, \mathrm{M^{-1}s^{-1}}$ which is consistent with a diffusion limited model of the association[20]. This finding implies that there are no rate-limiting conformational rearrangements on the association path between

the two macromolecules[51,52], which is confirmed by the high similarity of the crystal structures of ARF1-DBD alone and co-crystallized with an IR7[12].

ARFs are divided in three phylogenetically conserved classes called A, B and C[53]; A-ARFs are considered activators while B and C-ARFs are considered repressors[54–56]. Recently, a mechanism for B-ARFs repression based on competition for the same AuxREs bound by A-ARFs has been proposed as part of a minimal auxin response system in *Marchantia polymorpha*[41]. This model requires the affinity of class A-ARF and class B-ARF to be within one order of magnitude in order to allow for competition at near-stoichiometric A-ARF and B-ARF concentrations. We previously validated this model on IR7-TGTCTC, since the ratio between the affinity of MpARF2-DBD (a class B-ARF) and MpARF1-DBD (a class A-ARF) was found to be ≈ 5 ([41], Fig. 2.5); the results shown in this chapter further support the model for IR7-TGTCGG as the ratio between affinities is even smaller at ≈ 3. Interestingly, the ratio of the affinities towards IR7-TGTCAA is ≈ 13 which hints to a regulation for this element that does not involve competition between A and B-ARFs.

The method presented here can be generally applied to study protein-DNA interactions as long as the specific sequence targeted by the protein is known. If further structural data is available (e.g. crystal structures), the design of the DNA construct is simplified further by evaluating possible labelling positions *in silico*. This allows to reject quickly the ones that do not display significant changes in FRET efficiency and/or PIFE. Most importantly, the presented method does not rely on protein labelling but on custom labelled DNA oligos that are commercially available.

Conclusions

We studied the affinity and kinetics of ARF-DNA interaction using smFRET and smPIFE-FRET. We showed that our approach can reliably quantify the affinity and kinetics of ARF binding to its response element providing the kind of quantitative information that was not available in literature. The method presented in this chapter offers a convenient way to study protein-DNA interactions at the single-molecule level, further helping smFRET and smPIFE-FRET to become a standard and widespread tool in structural molecular biology.

References

1. Wegner, G. J., Lee, H. J., Marriott, G. & Corn, R. M. Fabrication of Histidine-Tagged Fusion Protein Arrays for Surface Plasmon Resonance Imaging Studies of ProteinProtein and ProteinDNA Interactions. *Anal. Chem.* **75,** 4740–4746 (2003).

2. Hellman, L. M. & Fried, M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat Protoc* **2,** 1849–1861 (2007).

3. Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O. L., He, A., Marra, M., Snyder, M. & Jones, S. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4,** 651–657 (2007).

4. Berger, M. F. & Bulyk, M. L. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat Protoc* **4,** 393–411 (2009).

5. Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P. & Solano, R. DNA-binding specificities of plant transcription factors and their potential to define target genes. *PNAS* **111,** 2367–2372 (2014).

6. Moyroud, E., Minguet, E. G., Ott, F., Yant, L., Posé, D., Monniaux, M., Blanchet, S., Bastien, O., Thévenon, E., Weigel, D., Schmid, M. & Parcy, F. Prediction of Regulatory Interactions from Genome Sequences Using a Biophysical Model for the Arabidopsis LEAFY Transcription Factor. *The Plant Cell* **23,** 1293–1306 (2011).

7. O'Malley, R. C., Huang, S.-s. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A. & Ecker, J. R. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165,** 1280–1292 (2016).

8. Schreier, S., Petla, B. P., Lin, T., Chakravarty, S. & Subramanian, S. A simple and sensitive SYBR Gold-based assay to quantify DNA–protein interactions. *Plant Mol Biol* **101,** 499–506 (2019).

9. Lickwar, C. R., Mueller, F., Hanlon, S. E., McNally, J. G. & Lieb, J. D. Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* **484,** 251–255 (2012).

10. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. ARF1, a Transcription Factor That Binds to Auxin Response Elements. *Science* **276,** 1865–1868 (1997).

11. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. Dimerization and DNA binding of auxin response factors. *The Plant Journal* **19,** 309–319 (1999).

12. Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., López-Vidrieo, I., Franco-Zorrilla, J. M., de Vries, S. C., Solano, R., Weijers, D. & Coll, M. Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell* **156,** 577–589 (2014).

13. Stigliani, A., Martin-Arevalillo, R., Lucas, J., Bessy, A., Vinos-Poyo, T., Mironova, V., Vernoux, T., Dumas, R. & Parcy, F. Capturing Auxin Response Factors Syntax Using DNA Binding Models. *Molecular Plant* **12,** 822–832 (2019).

14. Galli, M., Khakhar, A., Lu, Z., Chen, Z., Sen, S., Joshi, T., Nemhauser, J. L., Schmitz, R. J. & Gallavotti, A. The DNA binding landscape of the maize AUXIN RESPONSE FACTOR family. *Nature Communications* **9,** 4526 (2018).

15. Lanctot, A., Taylor-Teeples, M., Oki, E. A. & Nemhauser, J. L. Specificity in Auxin Responses Is Not Explained by the Promoter Preferences of Activator ARFs. *Plant Physiology* **182,** 1533–1536 (2020).

16. Ha, T., Enderle, T., Ogletree, D. F., Chemla, D. S., Selvin, P. R. & Weiss, S. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *PNAS* **93,** 6264–6268 (1996).

17. Lerner, E., Cordes, T., Ingargiola, A., Alhadid, Y., Chung, S., Michalet, X. & Weiss, S. Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science* **359** (2018).

18. Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kühnemuth, R., Rezaei Adariani, S., Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V., Bowen, M. E., Chen, H., Cordes, T., Eilert, T., Fijen, C., Gebhardt, C., Götz, M., Gouridis, G., Gratton, E., Ha, T., Hao, P., Hanke, C. A., Hartmann, A., Hendrix, J., Hildebrandt, L. L., Hirschfeld, V., Hohlbein, J., Hua, B., Hübner, C. G., Kallis, E., Kapanidis, A. N., Kim, J.-Y., Krainer, G., Lamb, D. C., Lee, N. K., Lemke, E. A., Levesque, B., Levitus, M., McCann, J. J., Naredi-Rainer, N., Nettels, D., Ngo, T., Qiu, R., Robb, N. C., Röcker, C., Sanabria, H., Schlierf, M., Schröder, T., Schuler, B., Seidel, H., Streit, L., Thurn, J., Tinnefeld, P., Tyagi, S., Vandenberk, N., Vera, A. M., Weninger, K. R., Wünsch, B., Yanez-Orozco, I. S., Michaelis, J., Seidel, C. A. M., Craggs, T. D. & Hugel, T. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nature Methods* **15,** 669–676 (2018).

19. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. in *Methods in Enzymology* (ed Walter, N. G.) 455–514 (Academic Press, 2010).

20. Kamagata, K., Murata, A., Itoh, Y. & Takahashi, S. Characterization of facilitated diffusion of tumor suppressor p53 along DNA using single-molecule fluorescence imaging. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews. Special issue on Bioimaging* **30,** 36–50 (2017).

21. Martin-Fernandez, M. L., Tynan, C. J. & Webb, S. E. D. A 'pocket guide' to total internal reflection fluorescence. *Journal of Microscopy* **252,** 16–22 (2013).

22. Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J. & Kapanidis, A. N. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal* **99,** 3102–3111 (2010).

23. McKinney, S. A., Joo, C. & Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophysical Journal* **91,** 1941–1951 (2006).

24. Okamoto, K. & Sako, Y. Variational Bayes Analysis of a Photon-Based Hidden Markov Model for Single-Molecule FRET Trajectories. *Biophysical Journal* **103,** 1315–1324 (2012).

25. Van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzalez, R. L. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys J* **106,** 1327–1337 (2014).

26. Schmid, S., Götz, M. & Hugel, T. Single-Molecule Analysis beyond Dwell Times: Demonstration and Assessment in and out of Equilibrium. *Biophysical Journal* **111,** 1375–1384 (2016).

27. Hadzic, M. C. A. S., Börner, R., König, S. L. B., Kowerko, D. & Sigel, R. K. O. Reliable State Identification and State Transition Detection in Fluorescence Intensity-Based Single-Molecule Förster Resonance Energy-Transfer Data. *The Journal of Physical Chemistry B* (2018).

28. Du, C. & Kou, S. C. Statistical Methodology in Single-Molecule Experiments. *Statist. Sci.* **35,** 75–91 (2020).

29. Hwang, H., Kim, H. & Myong, S. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences* **108,** 7414–7418 (2011).

30. Hwang, H. & Myong, S. Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chem. Soc. Rev.* **43,** 1221–1229 (2014).

31. Qiu, Y. & Myong, S. in *Single-Molecule Enzymology: Fluorescence-Based and High-Throughput Methods* (eds Spies, M. & Chemla, Y. R.) 33–51 (Academic Press, 2016).

32. Stennett, E. M. S., Ciuba, M. A., Lin, S. & Levitus, M. Demystifying PIFE: The Photophysics Behind the Protein-Induced Fluorescence Enhancement Phenomenon in Cy3. *J. Phys. Chem. Lett.* **6,** 1819–1823 (2015).

33. Murphy, S. & Schuster, G. B. Electronic Relaxation in a Series of Cyanine Dyes: Evidence for Electronic and Steric Control of the Rotational Rate. *J. Phys. Chem.* **99,** 8516–8518 (1995).

34. Cao, J., Wu, T., Hu, C., Liu, T., Sun, W., Fan, J. & Peng, X. The nature of the different environmental sensitivity of symmetrical and unsymmetrical cyanine dyes: an experimental and theoretical study. *Phys. Chem. Chem. Phys.* **14,** 13702–13708 (2012).

35. Sanborn, M. E., Connolly, B. K., Gurunathan, K. & Levitus, M. Fluorescence Properties and Photophysics of the Sulfoindocyanine Cy3 Linked Covalently to DNA. *J. Phys. Chem. B* **111,** 11064–11074 (2007).

36. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43,** 1156–1171 (2014).

37. Gidi, Y., Götte, M. & Cosa, G. Conformational Changes Spanning Angstroms to Nanometers via a Combined Protein-Induced Fluorescence Enhancement–Förster Resonance Energy Transfer Method. *J. Phys. Chem. B* **121,** 2039–2048 (2017).

38. Ploetz, E., Lerner, E., Husada, F., Roelfs, M., Chung, S., Hohlbein, J., Weiss, S. & Cordes, T. Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. *Scientific Reports* **6,** 33257 (2016).

39. Lerner, E., Ploetz, E., Hohlbein, J., Cordes, T. & Weiss, S. A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement–Förster-Type Resonance Energy Transfer (PIFE-FRET). *J. Phys. Chem. B* **120,** 6401–6410 (2016).

40. Cooper, M., Ebner, A., Briggs, M., Burrows, M., Gardner, N., Richardson, R. & West, R. Cy3B™: Improving the Performance of Cyanine Dyes. *Journal of Fluorescence* **14,** 145–150 (2004).

41. Kato, H., Mutte, S. K., Suzuki, H., Crespo, I., Das, S., Radoeva, T., Fontana, M., Yoshitake, Y., Hainiwa, E., van den Berg, W., Lindhoud, S., Ishizaki, K., Hohlbein, J., Borst, J. W., Boer, D. R., Nishihama, R., Kohchi, T. & Weijers, D. Design principles of a minimal auxin response system. *Nature Plants* **6,** 473–482 (2020).

42. Van Dijk, M. & Bonvin, A. M. J. J. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res* **37,** W235–W239 (suppl_2 2009).

43. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., Goody, R. S., Gohlke, H. & Seidel, C. A. M. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Meth* **9,** 1218–1225 (2012).

44. Craggs, T. D., Sustarsic, M., Plochowietz, A., Mosayebi, M., Kaju, H., Cuthbert, A., Hohlbein, J., Domicevica, L., Biggin, P. C., Doye, J. P. K. & Kapanidis, A. N. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res* **47,** 10788–10800 (2019).

45. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17,** 27862–27872 (2015).

46. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. & Weiss, S. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *PNAS* **101,** 8936–8941 (2004).

47. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Research* **43,** 5998–6008 (2015).

48. Efron, B. & Tibshirani, R. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statist. Sci.* **1,** 54–75 (1986).

49. Jones, G. Fitting and handling dose response data. *J Comput Aided Mol Des* **29,** 1–11 (2015).

50. Freire-Rios, A., Tanaka, K., Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., Fontana, M., Hohlbein, J., Boer, D. R., Mironova, V. & Weijers, D. Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in *Arabidopsis. Proc Natl Acad Sci USA,* 202009554 (2020).

51. Qin, S., Pang, X. & Zhou, H.-X. Automated Prediction of Protein Association Rate Constants. *Structure* **19,** 1744–1751 (2011).

52. Zhou, H.-X. & Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **118,** 1691–1741 (2018).

53. Mutte, S. K., Kato, H., Rothfels, C., Melkonian, M., Wong, G. K.-S. & Weijers, D. Origin and evolution of the nuclear auxin response system. *eLife* **7** (ed Yu, H.) e33399 (2018).

2

54. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. Activation and repression of transcription by auxin-response factors. *PNAS* **96,** 5844–5849 (1999).
55. Tiwari, S. B., Hagen, G. & Guilfoyle, T. The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. *The Plant Cell* **15,** 533–543 (2003).
56. Roosjen, M., Paque, S. & Weijers, D. Auxin Response Factors: output control in auxin biology. *J Exp Bot* **69,** 179–188 (2018).

Supplementary material
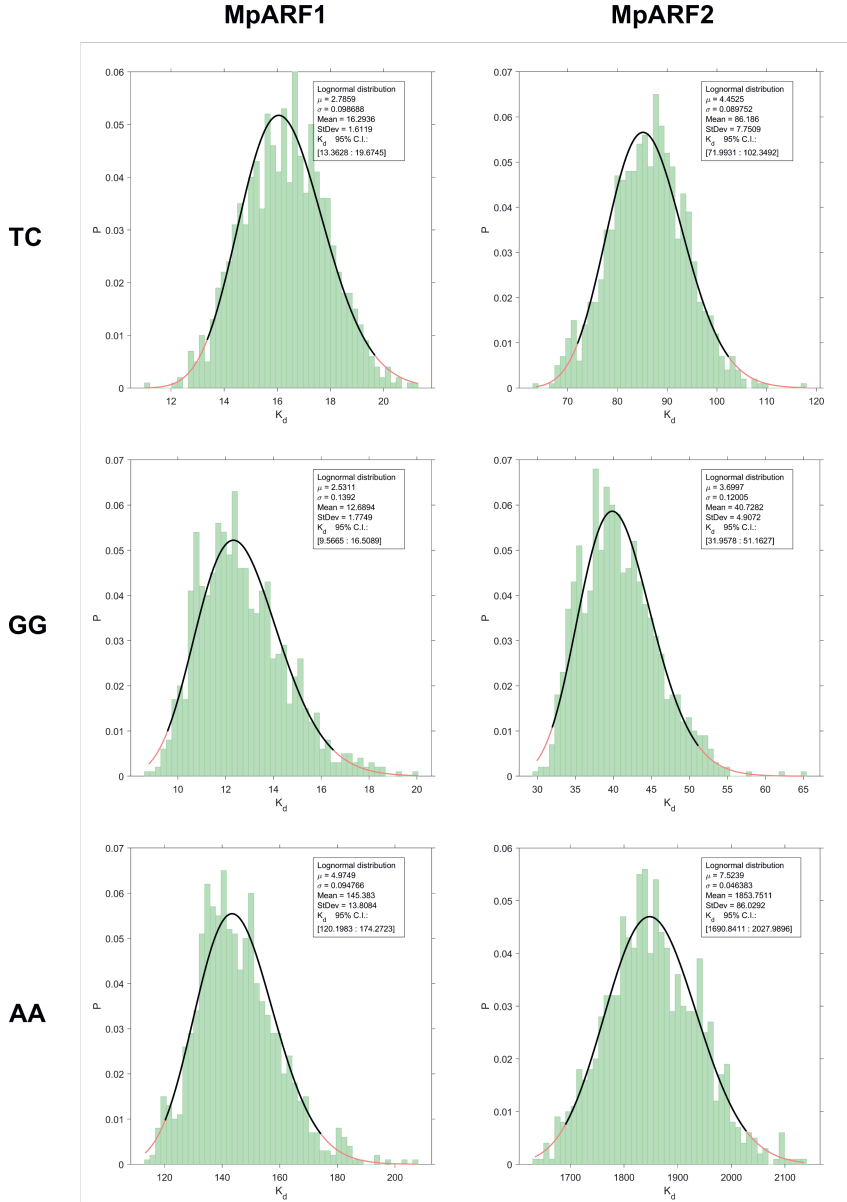
Figure S2.1: Log-normal fits of the probability distributions for the K_ds. The fits are drawn in black up to the limits of their 95 % CI for the K_d and in red for values that exceed them. The relevant information about the fits are annotated inside the plots.

Chapter 3

Affinity and kinetics of ARF-AuxRE interaction: a tale of monomers and dimers

Mattia Fontana, Mark Roosjen, Willy van den Berg, Dolf Weijers and Johannes Hohlbein

3

Abstract

The plant hormone auxin coordinates many growth and development processes mainly through the modulation of gene expression. The transcriptional response to auxin is mediated by the auxin response factor (ARF) transcription factor family. Monomers of this family recognize a DNA motif (TGTC[TC]/[GG]) called the auxin-responsive element (AuxRE). ARFs can homodimerize through their DNA binding domains (DBD) thereby enabling cooperative binding for a bipartite inverted AuxRE (IR7). In addition to the DBD, most ARFs contain a Phox and Bem1p (PB1) domain which is capable of homotypic interactions; while this domain is thought to promote protein dimerization and DNA binding on composite AuxREs there has been no quantitative observation of these effects, nor are the relative contributions of the two dimerization domains understood. Here, we utilize a DNA binding assay based on smPIFE-FRET to study the equilibrium and kinetics of the interaction of several *A. thaliana* ARFs with an IR7 AuxRE. We demonstrate that in case of AtARF2, the PB1 domain increases the affinity towards the DNA response element through stabilization of the protein dimer. Similarly, dimer stability can explain most of the differences in affinity and kinetics seen for the DBDs of different AtARFs. Lastly, we derived an analytical solution for a four-states model that explains both the kinetics and the affinity of the interaction between AtARF2 and IR7. The model further indicates that ARF dimerization both on the DNA and in solution are relevant for the interaction. Our work demonstrates that the affinity of ARFs towards composite DNA response elements can be tuned by small changes of their dimerization equilibrium and that this effect has probably major implications for ARF-mediated transcriptional activity.

Introduction

The plant hormone auxin plays a pivotal role in many cellular and developmental processes. Auxin triggers both non-transcriptional and transcriptional responses with the latter being controlled by the nuclear auxin pathway[1–5]. This pathway involves three main players: the transcription factor ARF, its repressor Aux/IAA and the ubiquitin ligase complex SCF$^{TIR1/AFB}$. Binding of auxin to TIR1/AFB enables the recognition and ubiquitination of Aux/IAA. Upon degradation of Aux/IAA, ARF is able to modulate the expression of its downstream target genes.

The interaction between ARFs and Aux/IAAs is mediated by the Phox and Bem1p (PB1) domain that is present in both proteins. The PB1 domain features two oppositely charged surfaces (type I/II or AB [acid basic] PB1 domain) that can oligomerize head to tail[6–9]. Remarkably, this structural characteristic opens scenarios of homo and hetero oligomerization among and between Aux/IAAs and ARFs. In addition to the PB1 domain, ARFs consist of two other domains, the Middle Region (MR) and the DNA Binding Domain (DBD). The MR domain is predicted to be intrinsically disordered[5] and its amino acid sequence differs between the three phylogenetically separated ARF clades (A,B and C)[10,11]. When tested for their effect on gene expression, some ARFs activate auxin-responsive genes while other repress them; in general, activator ARFs belong to class A (e.g. *A. thaliana* ARF5) while repressor ARFs belong to class B (e.g. *A. thaliana* ARF1 and 2). The DBD domain physically interacts with its DNA response element called AuxRE (auxin-responsive element). This cis-regulatory element was first identified in promoters of an auxin-responsive genes in pea[12] and soybean[13,14] and was found to be necessary for the inducibility of auxin. The canonical TGTCTC recognition sequence was later shown to be bound by different members of the ARF family[15,16]. More recently, the TGTCGG recognition sequence was found to have an even higher affinity for ARFs *in vitro*[17–19] and was used to create an enhanced artificial auxin response reporter[20]. Single AuxREs are bound by single ARF monomers but ARF DBDs can dimerize in solution and bind cooperatively to composite response elements bearing two AuxRE in inverted configuration (IR)[17]; moreover, ARF dimerization through its DBD is necessary for ARF function *in vivo*[17,21]. Interestingly, the PB1 domain seems to have different effects on different class A ARFs as its deletion in *M. polymorpha* ARF1 generates a loss-of-function mutant[22] whereas in AtARF5 the mutant maintains its function and is hyperactive[23]. The effect of the homotypic interaction of ARF PB1 domains of another class A ARF, AtARF19, has been studied using synthetic auxin

response circuits in yeast, showing that mutating either the positive or the negative side of the PB1 domain reduces its transcriptional output[24].

These studies show that much is known about the structures and the relevant interactions among the players of the auxin nuclear pathway but at the same time they highlight the lack of a quantitative view on this system. In particular, the effect of the dimer/monomer equilibrium on the interaction between ARFs and between ARFs and AuxREs or the effect of mutations on the DBD and PB1 domains on ARF dimerization have never been studied quantitatively. This lack of information about the affinity and kinetics leaves open questions about which interactions are relevant.

Here, we utilize the DNA-binding assay based on smPIFE-FRET presented in chapter 2 to quantitatively assess the binding affinities between different *A. thaliana* ARFs and a response element composed of two AuxREs in an inverted repeat configuration with a spacing of 7 bp. We found that the presence of the PB1 domain increases the affinity of AtARF2 towards the tested composite response element. In fact, this effect is caused by an increased stability of the dimer whereas AtARF2 oligomerization has no sizable effect. The analysis of the kinetics of the interaction between AtARF2 and the DNA response element revealed that the presence of the PB1 domain increases the k_{on} thanks to the stabilization of the dimer in solution while decreases the k_{off} thanks to the stabilization of the dimer bound to the DNA. Further analysis of variants of AtARF5-DBD showed that changes in dimer stability generated through mutations in the DBD domain displays the same pattern on the kinetics as the ones generated by changes in the PB1 domain, highlighting that stable protein dimers are proxy to stable DNA binding, no matter the source of their stability. Lastly, we introduce a general four-state cyclic model to quantify the mechanisms of ARF interaction with bipartite DNA response elements.

Materials and methods

Protein expression and purification

Protein expression and purification was carried out following the same protocol used in reference [17]. Briefly, the genomic regions corresponding to the DNA binding domain (DBD) of *Arabidopsis thalania* ARF1, ARF2 ARF5 and full-length ARF2 were amplified and cloned in an modified expression vector pTWIN1 (New England Biolabs) to generate fusions with the Chitin Binding Domain (CBD) and Intein.

ARF-CBD fusion proteins were expressed in *E. coli* strain Rosetta DE3 (Novagen). Cells were inoculated in Difco Terrific Broth (BD), supplemented with ampicillin and grown to an OD_{600} of 0.5–0.7, protein expression was induced by adding IPTG and the temperature was switched from $37\,°C$ to $20\,°C$; the growth was continued for 20 h. Cells were harvested by centrifugation and resuspended in extraction buffer (20 mM Tris, 500 mM NaCl, 1 mM EDTA, 0.1 % NP-40 and 2 mM MgCl2, pH 7.8, 10 mg of DNase and 0.2 mM PMSF). Cells were then lysed by passing the suspension twice through a French Pressure cell and cell-free extract was generated by centrifugation. The supernatant was loaded onto a chitin column (New England Biolabs) and washed with 10 column volumes washing buffer (20 mM Tris, 500 mM, pH 7.8) using an AKTA explorer 100 (GE Healthcare). ARF-DBD proteins were eluted by 1 h incubation with 40 mM DTT in washing buffer. Proteins were concentrated using Amicon ultra-15 10K spin filters, and next passed over a Superdex 200PG size-exclusion chromatography column. ARF-DBD proteins were eluted using washing buffer with 1 mM DTT, concentrated using Amicon ultra-15 10K spin filters and stored until use at $-80\,°C$.

DNA constructs

Single strand DNA oligos were ordered from Eurogentec. Each strand contained a 5-C6-amino-dT modification at the desired position for labelling. Some of the strands were purchased biotinylated at their 5'- end to allow for surface immobilization using a Neutravidin bridge. Strands were labelled with the desired dye (NHS-ester) following a modified version of the protocol provided by the dye manufacturer and purified using polyacrylamide gel electrophoresis (20 % Acrylamide). DNA constructs were annealed by heating complementary single strands to $95\,°C$ in annealing buffer (250 mM NaCl, 10 mM Tris HCl pH 8, 1 mM EDTA) followed by cooling down to room temperature performed overnight.

Single-molecule PIFE-FRET

Imaging was carried out on a home-built TIRF microscope, described previously[25]. The measurements were performed using alternating-laser excitation (ALEX)[26]; in this excitation scheme, each frame during which the donor is excited is followed by a frame in which the acceptor is directly excited. The emission of the fluorophores is spectrally divided into two different detection channels on the emCCD camera

sensor (Andor iXon 897 Ultra). This approach creates four photon streams, three of which are relevant; (1) donor emission after donor excitation (DD), (2) acceptor emission after donor excitation (DA, arising from FRET) and (3) acceptor emission after acceptor excitation (AA). The three photon streams can be used to calculate the raw FRET efficiency ($E^* = DA/(DD + DA)$) and stoichiometry ($S = (DD + DA)/(DD + DA + AA)$)[27]. E^* contains the information about the relative distance of the two fluorophores whereas S contains information about the photophysical state of a given molecule (allowing to filter out molecules missing an active donor or an active acceptor). The camera acquisition time and the excitation time were set to 250 ms; laser powers were set to 3 mW (Cy3) for green ($\lambda = 561$ nm) and 0.5 mW for red ($\lambda = 638$ nm) lasers. The imaging buffer contained 137 mM NaCl, 2.7 mM KCl, 10 mM phosphate, 1 mM Trolox, 1 % gloxy and 1 % glucose.

Single-molecule titration experiments

Labelled dsDNA oligos were immobilized on a PEGylated glass coverslip as described previously[28]. In particular, the PEGylation was carried out inside the wells of silicone gaskets placed on the coverslip (Grace Bio-labs). Each titration was performed using a single well, washing it between data points with 600 µL of 1x PBS buffer. The final washing step consisted of three washings separated by 15 minutes. Typically, each data point consisted of four movies (1000 frames each).

Binding isotherms analysis

The fit of the FRET efficiency distribution with the two Gaussian pertaining to the free and bound populations returns an uncorrected fraction bound for each datapoint i: $F_{\mathrm{B}}^u(i)$. When no protein is added, the Gaussian fit centered on the E^* of the bound DNA return an uncorrected fraction bound $F_{\mathrm{B}}^u(0) > 0$ (typically ≈ 0.1). This value is an indication of the error connected to the two population fit and can be used to renormalize the entire titration under the assumption that, in case the DNA would be completely bound, the expected uncorrected fraction free would have the same value ($F_{\mathrm{B}}^u(0) = F_{\mathrm{F}}^u(\infty)$). Then, the corrected fraction bound for each data point can be calculated as

$$F_{\mathrm{B}}^c(i) = \frac{F_{\mathrm{B}}^u(i) - F_{\mathrm{B}}^u(0)}{1 - 2F_{\mathrm{B}}^u(0)}. \tag{3.1}$$

The corrected fraction bound (henceforth F_{B}) can be fit with the appropriate mathematical model for the interaction.

Time traces analysis

First, the time traces from individual DNA molecules were filtered to remove sections in which either the donor or the acceptor were inactive (due to fluorophore bleaching or blinking). A molecule was allowed to take values of E^* and S outside the thresholds (typically 0 to 0.85 for E^* and 0.5 to 0.9 for S) for a maximum of three consecutive data points; longer stays outside the thresholding range resulted in the trace being interrupted. In case the molecule reentered the allowed range for E^* and S, the data points were saved as a new trace. A minimum length of the traces was set to 50 data points. The filtered time traces were then loaded in the software package ebFRET to perform and empirical Bayesian Hidden Markov Modelling[29]. The analysis was performed assuming two states, with two restarts and a convergence threshold of 10^{-6}. The results of the analysis were exported as '.csv' text files and the transition matrix was used to calculate k_{on}, k_{off} and $K_d = k_{off}/k_{on}$.

Results

The affinity of AtARF2 towards an IR7 composite AuxRE is increased by the PB1 domain

Following up on a ChIP experiment published in [24], we tested if the interactions between the PB1 domains increases the affinity of ARFs towards a composite AuxRE. We designed smPIFE-FRET experiments in which the binding of ARFs to a small doubly-labelled dsDNA oligo containing two AuxREs in an inverted configuration spaced by 7 basepairs (IR7) leads to a decrease of FRET efficiency (Fig. 3.1, see Chapter 2 for details on the assay). We then performed titrations with increasing concentrations of AtARF2-DBD, AtARF2-FL (full length), AtARF2-FL K2S, and AtARF2-FL OPCA (Fig. 3.2a). AtARF2-FL K2S (K737S) and AtARF2-FL OPCA (D797-8S) carry mutations of amino acids on the positive (K2S) and negative (OPCA) side of the PB1 domain respectively, both of which were shown to impair the interaction between PB1 domains[6]. These two mutants were tested individually or mixed in a 50/50 ratio (AtARF2-FL KpO).

The FRET efficiency distributions of the DNA sensor show a single population centered at $E^* = 0.59$ in absence of proteins. With increasing concentration of AtARF2, a second population representing the bound DNA fraction (centered at $E^* = 0.42$) appeared. For saturating protein concentrations (2048 nM) all DNA strands are bound. To demonstrate that the shift seen during the titration is gen-
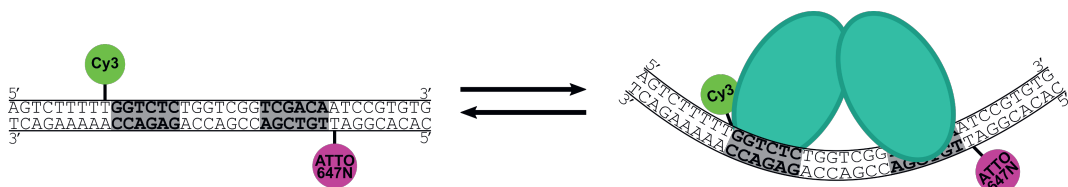
Figure 3.1: Schematic representation of the DNA-binding assay used to evaluate ARF binding; the dsDNA is labelled with Cy3 and Atto647N on the opposite sides of the response element. Upon protein binding, the increased distance between the dyes generates a decrease in FRET efficiency.

erated by specific binding of ARF to the DNA and that the binding is reversible, we performed a washing step at the end of each titration that reverted the FRET efficiency distributions to the ones seen in absence of any protein.

We then calculated the fraction of DNA bound by ARF (F_B) for every datapoint (see Materials and Methods) in the titration and fitted each dataset with a simple binding isotherm ($F_B = [ARF]_T/(K_d + [ARF]_T)$, Fig. 3.2b) to obtain the K_d of the interaction. AtARF2-DBD shows a K_d of 240 nM (95 % CI [120:360]) whereas AtARF2-FL shows a much tighter binding ($K_d = 39$ nM, CI [31:47]) suggesting that the stabilization of the dimer caused by the homotypic interaction of the PB1 domain does stabilize the association between ARF and the IR7 response element. This hypothesis is confirmed by the fact that when the PB1 domain is mutated on either its positive (K2S) of negative (OPCA) side the increase in affinity is strongly reduced (to 140 nM [100:180] and 220 nM [190:250] respectively). Furthermore, when the two mutants are mixed in a 50/50 ratio (KpO), the gain in affinity is restored to levels comparable with the wild type FL variant ($K_d = 57$ nM, CI [35:78]). This further demonstrates that on this RE topology, oligomerization does not play an important role in stabilizing the interaction and that the dimer is the relevant species for DNA binding.

The PB1 domain increases the probability of association and decreases the probability of dissociation

We then tested the kinetics of AtARF2/IR7 interaction by analyzing relevant datapoints of the titrations using ebFRET[29], a MATLAB suite for Hidden Markov Model (HMM) analysis of single-molecule time traces. As the DNA oligos are immobilized, their interaction with the proteins in solution can be monitored for several minutes
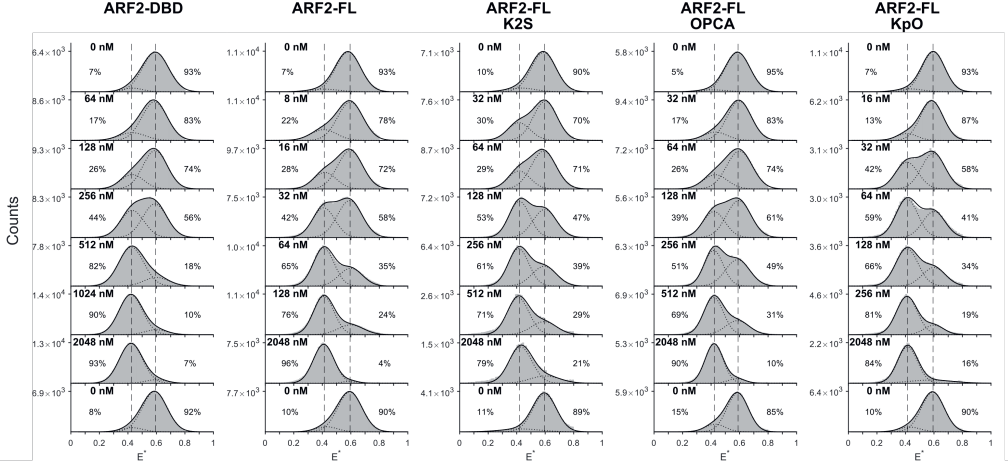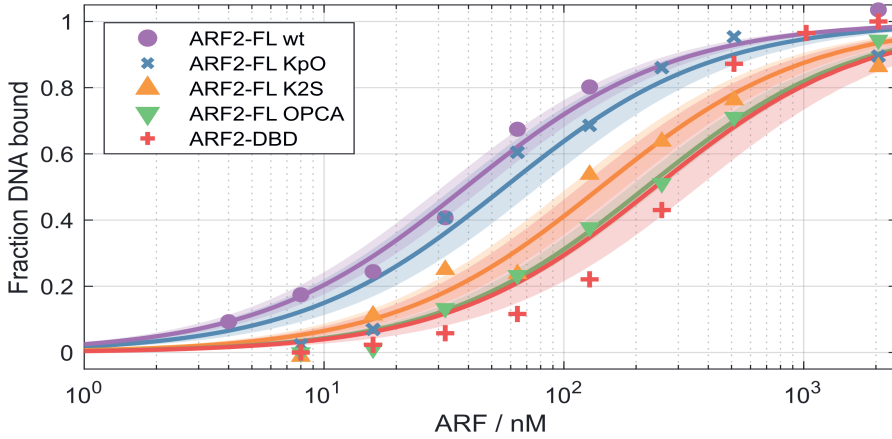
Figure 3.2: Effect of the PB1 domain on the affinity between AtARF2 and an IR7 AuxRE. (a) Titrations of the dsDNA with several ARF variants. The dsDNA alone has a FRET efficiency $E^* = 0.59$; as the protein concentration increases the population of bound DNA (centered at $E^* = 0.42$) increases until all the DNA is bound (saturating condition). A washing step suffice to reset the system proving that the bound population is generated by specific and reversible binding of ARF. (b) The fraction of DNA bound as function of the protein concentration can be fitted using simple binding isotherms obtaining values of K_d for the interactions. The 95 % confidence interval is reported as shaded areas, the only datasets that show significantly tighter binding compared to AtARF2-DBD are AtARF2-FL and AtARF2-FL KpO which indicates that the homotypic interaction of the PB1 domain is increasing the affinity of AtARF2 towards the response element through stabilization of the dimer.

(250 s in our experiments). The analysed FRET traces returned the most probable hidden state sequence for each trace (Fig. 3.3a) according to the set of kinetic parameter that best explain the transitions and states seen in the entire dataset.

To facilitate the identification of transitions by ebFRET, we selected the three concentrations that were closest to having equal population of bound and free DNA for each AtARF2 variant. Each of these three concentrations (represented as Δ, $\square$ and ∇ in Fig. 3.3b) were tested during at least three independent "short" titrations for each ARF variant; each of these datapoints was analyzed independently and returned a value for k_{on}, a value for k_{off} and, from their ratio, a value for K_d (Fig. 3.3b). The mean values of k_{on} show a clear trend in which ARF variants with higher affinity show faster association. On the other hand, the k_{off} show similar values for all the FL variants whilst AtARF2-DBD has a faster k_{off} accompanied by a large spread in the individual values. The resulting K_ds show the expected trend, with AtARF2-DBD having the lowest affinity (400 nM, 95 % CI [220:590]), AtARF2-FL KpO and wt showing the tightest binding (78 nM CI [47:109] and 53 nM CI [29:76], respectively) and AtARF2-FL K2S and OPCA having an affinity in between the two (140 nM CI [80:190] and 110 nM CI [20:190], respectively).

The trends seen in the k_{on} and k_{off} suggest that the analysis of the kinetics using HMM is capturing mostly the interaction between the dimer and the DNA and that the interaction between the monomer and the DNA occurs on a timescale shorter than the 500 ms acquisition time used in our experiments. Under this hypothesis, the apparent k_{on} can be approximated by the probability of the DNA encountering a dimer which is increased by the enhanced dimer stability in solution brought by the PB1 domain. The k_{off} is also a function of the dimer stability but when it is bound to the DNA, as both the dissociation of either a dimer from the DNA or one monomer from the other monomer and the DNA contribute to its value. The nearly identical dissociation kinetics of all FL variants suggests that the presence of the PB1 domain, even when the homotypic affinity is reduced by mutations, is sufficient to stabilize the ARF-FL dimer on the DNA and to make the probability of dissociation of a single ARF-FL monomer from the other monomer and the DNA negligible compared to the dissociation of the ARF-FL dimer from the DNA. On the other hand, AtARF2-DBD shows a faster dissociation kinetics where the consensus found by the HMM analysis converges on the sum of the k_{off}s for both dissociation processes. In fact, we explain the observed increased spread of k_{off} values as being
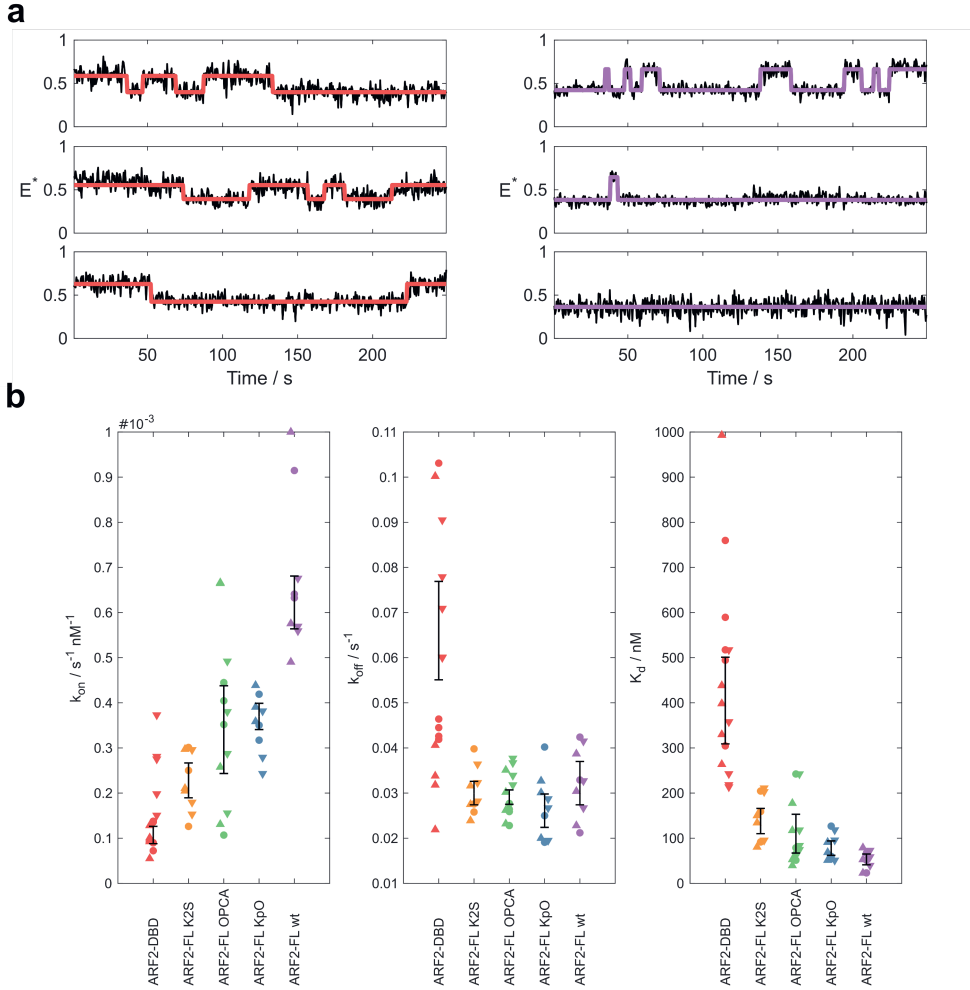
Figure 3.3: Kinetics of the interaction between AtARF2 variants and IR7 AuxRE. (a) Example of FRET efficiency time traces of single doubly-labelled dsDNA in presence of 128 nM of either AtARF2-DBD (left) or AtARF2-FL wt (right). The FRET efficiency is reported in black and the most probable sequence of hidden states returned by ebFRET is represented in red (AtARF2-DBD) and purple (AtARF2-FL wt). (b) Kinetics parameters obtained from ebFRET. For each ARF variant three concentrations closest to having equal bound and free population were chosen and repeated in at least three independent titrations. The concentrations were 128, 256, 512 nM for AtARF2-DBD, 64, 128 and 256 nM for AtARF2-FL K2S and OPCA and 32, 64, and 128 nM for AtARF2-FL and ARF2-FL KpO. Each repeat of each concentration was analyzed independently using ebFRET and the resulting k_{on} and k_{off} (and, from their ratio the K_d) are marked with Δ, $\square$ and ∇ in order of increasing ARF2 concentration. The black interval represent ± 1 standard deviation from the weighted mean values..

the result of fitting the convolution of two dissociation processes with a single kinetic parameter.

Stability of the dimer influences the binding kinetics of different ARF-DBDs

In the previous paragraph we showed that the added dimer stability induced by the presence of the PB1 domain influences the kinetics of the binding of ARF towards its DNA response element. Here, we asked if the ability of ARFs to dimerize through the DBD domain alone could have the same effect by testing the DBDs of different AtARFs.

Experiments with AtARF1-DBD showed similar values of k_{on} and k_{off} ($1.3 \times 10^{-4}\,\mathrm{nM^{-1}s^{-1}}$ 95 % CI [0.8:1.8], $0.080\,\mathrm{s^{-1}}$ 95 % CI [0.062:0.098], respectively) as AtARF2-DBD ($1.1 \times 10^{-4}\,\mathrm{nM^{-1}s^{-1}}$ 95 % CI [0.7:1.4], $0.066\,\mathrm{s^{-1}}$ 95 % CI [0.045:0.087] , respectively; Fig. 3.4). On the other hand, AtARF5-DBD showed a 5-fold increase in k_{on} ($5.9 \times 10^{-4}\,\mathrm{nM^{-1}s^{-1}}$ 95 % CI [2.3:9.4]) and an 8-fold reduction in k_{off} ($0.0085\,\mathrm{s^{-1}}$ 95 % CI [0.0051:0.0118]) compared to AtARF2-DBD; which lead to a K_d of 15 nM (95 % CI [12:18]). In analogy with the considerations made for AtARF2-FL, the increase in k_{on} and part of the decrease in k_{off} can be explained with AtARF5-DBDs forming a tighter dimer compared to AtARF1 and AtARF2 DBDs. To prove this hypothesis, we tested AtARF5-DBD G279N, a single amino acid mutation known to reduce AtARF5-DBD dimerization[17]. Strikingly, the kinetics of the interaction between AtARF5-DBD G279N and the IR7 became similar to the ones of AtARF1 and AtARF2 DBDs ($1.2 \times 10^{-4}\,\mathrm{nM^{-1}s^{-1}}$ 95 % CI [1.0:1.5], $0.10\,\mathrm{s^{-1}}$ 95 % CI [0.04:0.17]) validating hour hypothesis. Our results suggest that AtARF5-DBD forms tighter dimers compared to AtARF1-DBD and AtARF2-DBD and this explains part of the higher affinity towards the IR7. Finally, we tested AtARF5-DBD R215A, a mutant in which a key amino acid for the interaction with the DNA is mutated. This mutant showed a 13-fold reduction of k_{on} compared to the wild-type ($0.46 \times 10^{-4}\,\mathrm{nM^{-1}s^{-1}}$ 95 % CI [0.41:0.51] as well as a 39-fold increase of k_{off} ($0.33\,\mathrm{s^{-1}}$ 95 % CI [0.14:0.52]) which translates in a reduction of affinity of three orders of magnitude. We note that the reduction of k_{on} is consistent in magnitude with the effect of charge neutralization of DNA-contacting residues seen in other protein-DNA interactions[30] and is a reminder of the importance of charged residues in defining association kinetics[31,32].
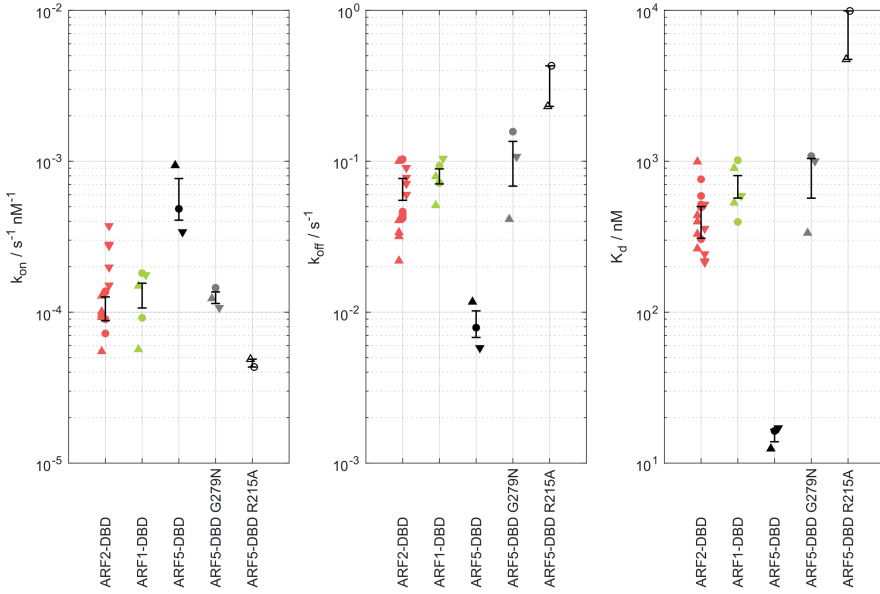
Figure 3.4: Kinetics of the interaction between AtARF-DBDs and IR7. Kinetic parameters obtained from HMM analysis using ebFRET. The datapoints are marked with Δ, $\square$ and ∇ in order of increasing ARF concentration. The concentrations were 128, 256, 512 nM for AtARF2-DBD and AtARF1-DBD, 8, 16 and 32 nM for AtARF5-DBD, 64, 128 and 256 nM for AtARF5-DBD G279N and 128 and 512 nM for AtARF5-DBD R215A. The error bars represent the standard deviations of the mean values. AtARF2-DBD and AtARF1-DBD behaved similarly while AtARF5-DBD showed increased k_{on} and decreased k_{off}. Consistent with a model in which part of the difference in kinetic can be explained by an increased stability of AtARF5-DBD dimer, a weakening of the dimerisation (G279N mutant) leads to kinetic parameters that resemble the ones of AtARF1-DBD and AtARF2-DBD. In addition, AtARF5-DBD R215A mutant in a key amino acid for the interaction with DNA showed a one order of magnitude reduction of k_{on} and an almost two order of magnitude increase of k_{off} compared to the wild-type.

3

The interaction between ARF and an IR7 follows a four-states cyclic reaction mechanism

So far we explained the trends seen in the observed k_{on} and k_{off} in a qualitative way, as being influenced by changes of the stability of ARF dimers in solution and on the DNA. Our reasoning implied the existence of ARF in both monomeric and dimeric forms and opens questions about the possible equilibrium landscapes of their interactions with the DNA. In this paragraph we seek to explain the equilibrium and kinetics of the binding seen for AtARF2/IR7 interaction using a unified, quantitative model.

In case of an IR7 the most general model consists of a four-states cyclic association mechanism (Fig. 3.5a). This model allows for monomers or dimers to bind the DNA, for monomers and dimers to exist in solution and for dimers to form or dissociate both in solution or on the DNA. The system is defined by four k_{on}s and four k_{off}s or alternatively by four equilibrium constants (Ks). The presence of the PB1 domain (either wild-type or mutant) should not change the contacting interface between the DBDs and the DNA; hence, it is safe to assume that the k_{off} of the dimer from the DNA ($k_{\mathrm{off,DF}}$) and of the monomer from the DNA ($k_{\mathrm{off,MF}}$) has the same value for all AtARF2-variants. Moreover, we assume that the PB1 domain has limited influence on the k_{on}s of the system which stay diffusion limited. The only constants that are expected to be influenced by changes in the stability of the dimer induced by the PB1 domain are the ones associated with the separation of two monomers: the equilibrium dissociation constant of the dimer in solution ($K_{\mathrm{I}} = k_{\mathrm{off,DS}}/k_{\mathrm{on,DS}}$, see fig. 3.5a) and $k_{\mathrm{off,DM}}$, which encompasses the stability of the dimer on the DNA. Following these assumption, the titrations and the kinetic analysis of time traces can be fitted with a system consisting of three parameters shared across all AtARF2 datasets (k_{on}, $k_{\mathrm{off,DF}}$ and $k_{\mathrm{off,MF}}$) and the dataset-specific parameter K_{I}. We note that the dataset-specific parameter $k_{\mathrm{off,DM}}$ is defined, in our derivation of the system, by the imposition of microscopic reversibility[33] and is not a free parameter of the fit. Moreover, the single k_{on} is defined as the microscopic k_{on} that a monomer displays when binding a single AuxRE and hence it is equal to $k_{\mathrm{on,MM}}$ but it is half the value of $k_{\mathrm{on,M}}$ and $k_{\mathrm{on,D}}$.

The fraction of DNA bound as a function of the total concentration of ARF was fitted using the equation derived for the four-state association mechanism (see Supporting note 1) while the observed k_{on} and k_{off} were fitted according to the

Table 3.1: Global fit: values and uncertainty of the fitting parameters reported as mean [95 % CI]

Protein	k_{on} [nM^{-1}s^{-1}]	$k_{off,MF}$ [s^{-1}]	$k_{off,DF}$ [s^{-1}]	K_I [nM]
AtARF2-DBD	6.6 [4.5:8.8]$\times10^{-4}$	1.1 [0.2:2.0]	2.8 [2.4:3.2]$\times10^{-2}$	1.9 [0.8:3.0]$\times10^{3}$
AtARF2-FL OPCA	6.6 [4.5:8.8]$\times10^{-4}$	1.1 [0.2:2.0]	2.8 [2.4:3.2]$\times10^{-2}$	3.6 [0.3:6.8]$\times10^{2}$
AtARF2-FL K2S	6.6 [4.5:8.8]$\times10^{-4}$	1.1 [0.2:2.0]	2.8 [2.4:3.2]$\times10^{-2}$	3.1 [0.2:6.0]$\times10^{2}$
AtARF2-FL KpO	6.6 [4.5:8.8]$\times10^{-4}$	1.1 [0.2:2.0]	2.8 [2.4:3.2]$\times10^{-2}$	2.6 [-1.6:6.8]$\times10^{1}$
AtARF2-FL K2S	6.6 [4.5:8.8]$\times10^{-4}$	1.1 [0.2:2.0]	2.8 [2.4:3.2]$\times10^{-2}$	0.0 [-0.4:0.4]$\times10^{-3}$

expected theoretical transition probabilities (see Supporting note 2); the global fit returned values for k_{on}, $k_{off,DF}$, $k_{off,MF}$ and K_Is that best explained the titrations and the kinetic analysis (see Table 3.1); the results of the fit are reported as coloured lines for the titrations and black markers for the observed k_{on} and k_{off} (Fig. 3.5b and Fig. 3.5c, respectively). The global fit of the bound fraction converged to a K_I of 0 nM for AtARF2-FL wt; in this situation the equation of the bound fraction for the four-states system simplifies to a simple binding isotherm for the dimer (see Supporting note 1). The global fit of the other datasets captured the shift of the binding to higher [ARF]$_T$ thanks to increasing values of K_I which corresponds to a decrease in dimer stability. The global fit for AtARF2FL KpO is still very close to the one of the wt variant but because of the decrease in dimer stability ($K_I \approx 0.03\,\mu$M) its steepness is increased. The two AtARF2-FL mutants, K2S and OPCA, show a similar value of K_I ($\approx 0.3\,\mu$M) but here the increased steepness of the function returned a worse fit. Lastly, the titration of AtARF2-DBD, which was poorly fitted by the simple binding isotherm, returned a value of K_I of $\approx 2\,\mu$M.

The global fit off of the kinetics with the four-states model captured both the trends of observed k_{on} and k_{off} (Fig. 3.5c). The fit returned similar values of observed k_{on} for the datasets of AtARF2-DBD, AtARF2-FL OPCA and AtARF2-FL K2S, at around $2 \times 10^{-4}\,$nM^{-1}s^{-1}. When the two mutants were mixed together (AtARF2-FL KpO), the value increases to $\approx 4 \times 10^{-4}\,$nM^{-1}s^{-1}. AtARF2-FL wt represents a limiting case since only dimers are present ($K_I = 0$) and the value of observed k_{on} converged to half the true k_{on} of the dimer at $6.6 \times 10^{-4}\,$nM^{-1}s^{-1} (half because the observed rate is still divide by [ARF]$_T$). The fit for the observed k_{on} at the three different protein concentrations of AtARF2-FL wt shows a slight downward trend which is explained by the decrease survival probability of the DNA

3

in the free state as the dimer concentration increases (see Supporting note 2). The fits of the other AtARF2 variants show an opposite trend in which the observed k_{on} increases with increasing concentration of AtARF2. The main reason is that at higher protein concentration the equilibrium of ARF dimerization is shifted towards the dimer, thereby increasing its effective concentration. The fits for the observed k_{off} of AtARF2-FL wt and KpO converge to the value characteristic of the dissociation kinetic of the dimer from the DNA ($k_{off,DF} = 0.028\,\mathrm{s}^{-1}$). For the other datasets (AtARF2-FL K2S, AtARF2-FL OPCA and AtARF2-DBD) the dissociation of the dimer from the DNA is caused also by the loss of a monomer which becomes more likely than the dissociation of the dimer from the DNA in the case of AtARF2-DBD ($k_{off,DM} = 0.036\,\mathrm{s}^{-1}$).

The solution of the four-states system generated by the global fit shows that the monomer is the predominant species in solution for most of AtARF2 variants during most of the titrations (Fig. 3.5d top, solid lines). Strikingly, the fraction of DNA bound by a monomer never exceeds 11 % (Fig. 3.5d top, dashed lines). Looking at the association mechanism shown in figure 3.5a, there are two routes that lead from free DNA to dimer-bound DNA. In the first route, the dimer forms (and dissociates) in solution and then binds the DNA. In the second route, the dimer forms (and dissociates) on the DNA. While the first route is the only one used by AtARF2-FL and AtARF2-FL KpO, dimerization on the DNA becomes relevant for AtARF2-FL K2S and OPCA while becoming the preferred way to form a dimer-bound DNA for AtARF2-DBD (Fig. 3.5d bottom).

Discussion

In this chapter we studied the equilibrium and kinetics of ARF-AuxRE interaction focusing on the importance of dimer/monomer equilibrium of ARFs. We showed that the presence of the PB1 domain increases the affinity of AtARF2 towards an IR7 composite AuxRE and that this effect is caused by the increased stability of the protein dimer. Moreover, our results show that oligomerization does not further enhance the affinity towards bipartite response elements. This behaviour is consistent with the fact that additional monomers (beside the initial two) do not have any AuxRE left to further stabilize the binding to the DNA. This said, the effect of the PB1 domain seen in our experiments predicts that oligomerization should be relevant on response elements comprising of more than two AuxREs as the PB1 domain would enable cooperative binding beyond the dimer (under the
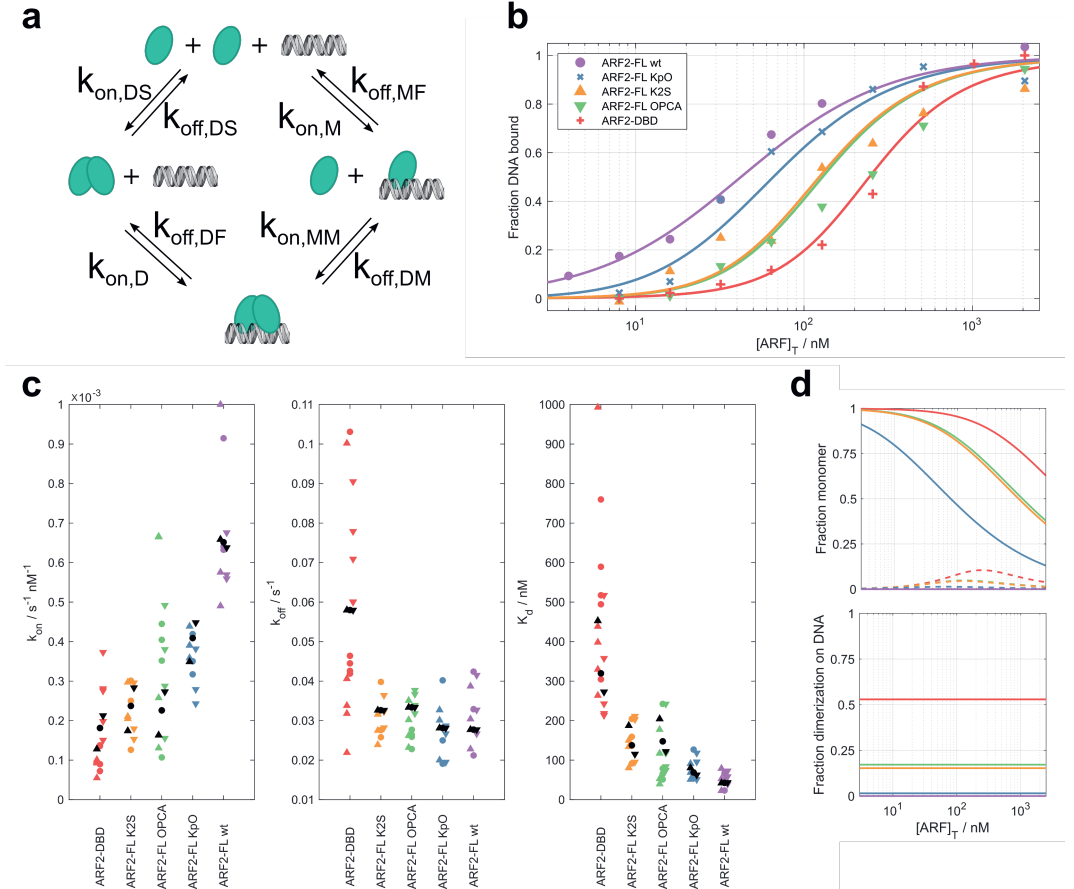
Figure 3.5: AtARF2-IR7 interaction follows a four-states cyclic association mechanism. (a) Four-states association mechanism for ARF-DNA interaction. The binding of a dimer to a bipartite response element can happen either through two successive binding events of a monomer or through direct binding of a dimer that formed in solution. (b-c) Global fit of the fraction of DNA bound (b, coloured lines) and kinetics of binding (c, black markers). (d) Features of the four states system as solved by the global fit. Top: In solution, the monomer is the most abundant species (solid lines). On the DNA, the monomer accounts for less then 11 % of the bound DNA (dashed lines). Note that for AtARF2-FL the monomer is not present since $K_I = 0$. Bottom: The binding of AtARF2 occurs by direct binding of the dimer in case of AtARF2-FL wt and AtARF2-FL KpO; this is still the preferred way for AtARF2-FL K2S and OPCA. For AtARF2-DBD the preferred way of binding is through dimerization on the DNA.

3

assumption of a flexible middle region, MR). Binding of ARF to bipartite AuxREs in the other two possible orientations (directed repeat DR and everted repeat ER), should resemble the one seen for the IR with the difference that the dimerization through the DBD domain should not be possible. Strikingly, titrations presented in a recent publication by Freire-Rios et al.[34] confirmed this hypothesis; the binding of AtARF1-DBD and AtARF5-DBD to a bipartite DR5 was compatible with a simple binding isotherm, whereas binding to an IR8 showed steeper response, similar to the one seen here for AtARF2-DBD. The equations of the 4-states model presented in this chapter can be applied to a bipartite DR5 without any modification other than the meaning of K_I which will no longer coincide with the one of the dimerization K_d in solution. Our model predicts that repeating the experiments presented by Freire-Rios et al.[34] with full length ARFs will return a tighter binding for both DR5 and IR8 but the steepness of the curve should swap; the DR5 should show a steep response while the IR8 should display a simple binding isotherm.

Looking at the kinetics of AtARF2/DNA interaction, the observed k_{on} and k_{off} can be explained as being function of the stability of the protein dimer in solution and bound to the DNA, respectively. The same was found true for AtARF5-DBD in which a mutation in the dimerization domain reduced the k_{on} and increased the k_{off}. This finding suggests that the tighter binding of AtARF5-DBD compared to AtARF1 and AtARF2 DBDs is in part due to the higher stability of its dimer. Then, the DBD alone should be able to achieve stable binding on an IR even in absence of the PB1 domain; this hypothesis explains why AtARF5ΔPB1 is a gain-of-function mutant that can activates auxin-responsive genes even in absence of the hormone[23]. Full-length ARF5 is known to bind many differently spaced IR, DR and ER[19,35] but following our hypothesis, we expect AtARF5ΔPB1 to modulate only genes whose promoter contains IR AuxREs.

Conclusions

In this work we presented a quantitative analysis of the interaction between members of the ARF transcription factor family and their DNA response element. Our results highlight the importance of protein dimerization in defining the affinity of the interaction between ARF and an IR7 composite AuxRE. Since stable DNA binding can be achieved only by stable dimerization/oligomerization, the topology of composite AuxREs could in theory dictate the affinity towards different ARF members differentiating them based on the relative strength of the homotypic interaction through

their DBD and PB1 domains. Tweaking the affinities of different ARFs towards the same DNA sequence can be achieved by affecting their dimerization properties and opens the possibility for easy evolution of complex interactions between members of the family. The quantitative approach we employed in this chapter allowed us to verify previous hypotheses about the affinities at play in ARF-AuxRE interaction and we envision that combining quantitative single-molecule studies with biochemical bulk-studies will help to develop a complete understanding of the nuclear auxin pathway.

3

References

1. Grones, P. & Friml, J. Auxin transporters and binding proteins at a glance. *J Cell Sci* **128,** 1–7 (2015).
2. Lavy, M. & Estelle, M. Mechanisms of auxin signaling. *Development* **143,** 3226–3229 (2016).
3. Weijers, D. & Wagner, D. Transcriptional Responses to the Auxin Hormone. *Annual Review of Plant Biology* **67,** 539–574 (2016).
4. Parcy, F., Vernoux, T. & Dumas, R. A Glimpse beyond Structures in Auxin-Dependent Transcription. *Trends in Plant Science* **21,** 574–583 (2016).
5. Roosjen, M., Paque, S. & Weijers, D. Auxin Response Factors: output control in auxin biology. *J Exp Bot* **69,** 179–188 (2018).
6. Nanao, M. H., Vinos-Poyo, T., Brunoud, G., Thévenon, E., Mazzoleni, M., Mast, D., Lainé, S., Wang, S., Hagen, G., Li, H., Guilfoyle, T. J., Parcy, F., Vernoux, T. & Dumas, R. Structural basis for oligomerization of auxin transcriptional regulators. *Nature Communications* **5,** 3617 (2014).
7. Korasick, D. A., Westfall, C. S., Lee, S. G., Nanao, M. H., Dumas, R., Hagen, G., Guilfoyle, T. J., Jez, J. M. & Strader, L. C. Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression. *PNAS* **111,** 5427–5432 (2014).
8. Han, M., Park, Y., Kim, I., Kim, E.-H., Yu, T.-K., Rhee, S. & Suh, J.-Y. Structural basis for the auxin-induced transcriptional regulation by Aux/IAA17. *PNAS* **111,** 18613–18618 (2014).
9. Dinesh, D. C., Kovermann, M., Gopalswamy, M., Hellmuth, A., Villalobos, L. I. A. C., Lilie, H., Balbach, J. & Abel, S. Solution structure of the PsIAA4 oligomerization domain reveals interaction modes for transcription factors in early auxin response. *PNAS* **112,** 6230–6235 (2015).
10. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. Activation and repression of transcription by auxin-response factors. *PNAS* **96,** 5844–5849 (1999).
11. Tiwari, S. B., Hagen, G. & Guilfoyle, T. The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. *The Plant Cell* **15,** 533–543 (2003).
12. Ballas, N., Wong, L.-M. & Theologis, A. Identification of the Auxin-responsive Element, AuxRE, in the Primary indoleacetic Acid-inducible Gene, PS-IAA4/5, of Pea (Pisum sativum). *Journal of Molecular Biology* **233,** 580–596 (1993).
13. Li, Y., Liu, Z. B., Shi, X., Hagen, G. & Guilfoyle, T. J. An Auxin-Inducible Element in Soybean SAUR Promoters. *Plant Physiology* **106,** 37–43 (1994).
14. Ulmasov, T., Liu, Z. B., Hagen, G. & Guilfoyle, T. J. Composite structure of auxin response elements. *The Plant Cell* **7,** 1611–1623 (1995).
15. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. ARF1, a Transcription Factor That Binds to Auxin Response Elements. *Science* **276,** 1865–1868 (1997).
16. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. Dimerization and DNA binding of auxin response factors. *The Plant Journal* **19,** 309–319 (1999).
17. Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., López-Vidrieo, I., Franco-Zorrilla, J. M., de Vries, S. C., Solano, R., Weijers, D. & Coll, M. Structural Basis

for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell* **156,** 577–589 (2014).

18. Franco-Zorrilla, J. M., López-Vidriero, I., Carrasco, J. L., Godoy, M., Vera, P. & Solano, R. DNA-binding specificities of plant transcription factors and their potential to define target genes. *PNAS* **111,** 2367–2372 (2014).

19. O'Malley, R. C., Huang, S.-s. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A. & Ecker, J. R. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165,** 1280–1292 (2016).

20. Liao, C.-Y., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T. & Weijers, D. Reporters for sensitive and quantitative measurement of auxin response. *Nature Methods* **12,** 207–210 (2015).

21. Liu, J., Hua, W., Hu, Z., Yang, H., Zhang, L., Li, R., Deng, L., Sun, X., Wang, X. & Wang, H. Natural variation in *ARF18* gene simultaneously affects seed weight and silique length in polyploid rapeseed. *Proc Natl Acad Sci USA* **112,** E5123–E5132 (2015).

22. Kato, H., Mutte, S. K., Suzuki, H., Crespo, I., Das, S., Radoeva, T., Fontana, M., Yoshitake, Y., Hainiwa, E., van den Berg, W., Lindhoud, S., Ishizaki, K., Hohlbein, J., Borst, J. W., Boer, D. R., Nishihama, R., Kohchi, T. & Weijers, D. Design principles of a minimal auxin response system. *Nature Plants* **6,** 473–482 (2020).

23. Krogan, N. T., Ckurshumova, W., Marcos, D., Caragea, A. E. & Berleth, T. Deletion of MP/ARF5 domains III and IV reveals a requirement for Aux/IAA regulation in Arabidopsis leaf vascular patterning. *New Phytologist* **194,** 391–401 (2012).

24. Pierre-Jerome, E., Moss, B. L., Lanctot, A., Hageman, A. & Nemhauser, J. L. Functional analysis of molecular interactions in synthetic auxin response circuits. *PNAS* **113,** 11354–11359 (2016).

25. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17,** 27862–27872 (2015).

26. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43,** 1156–1171 (2014).

27. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. & Weiss, S. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *PNAS* **101,** 8936–8941 (2004).

28. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Research* **43,** 5998–6008 (2015).

29. Van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzalez, R. L. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys J* **106,** 1327–1337 (2014).

30. Hancock, S. P., Hiller, D. A., Perona, J. J. & Jen-Jacobson, L. The Energetic Contribution of Induced Electrostatic Asymmetry to DNA Bending by a Site-Specific Protein. *Journal of Molecular Biology* **406,** 285–312 (2011).

31. Schreiber, G., Haran, G. & Zhou, H.-X. Fundamental Aspects of ProteinProtein Association Kinetics. *Chem. Rev.* **109,** 839–860 (2009).

32. Zhou, H.-X. & Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **118,** 1691–1741 (2018).

33. Colquhoun, D., Dowsland, K. A., Beato, M. & Plested, A. J. R. How to impose microscopic reversibility in complex reaction mechanisms. *Biophysical journal* **86,** 3510–3518 (2004).

34. Freire-Rios, A., Tanaka, K., Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., Fontana, M., Hohlbein, J., Boer, D. R., Mironova, V. & Weijers, D. Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in *Arabidopsis*. *Proc Natl Acad Sci USA,* 202009554 (2020).

35. Stigliani, A., Martin-Arevalillo, R., Lucas, J., Bessy, A., Vinos-Poyo, T., Mironova, V., Vernoux, T., Dumas, R. & Parcy, F. Capturing Auxin Response Factors Syntax Using DNA Binding Models. *Molecular Plant* **12,** 822–832 (2019).

3

Supplementary material

Supporting Note 1: Derivation binding isotherm four states system

The dsDNA containing the DNA response element can be found in three states: free (F), bound to a monomer (M) and bound to a dimer (D). The DNA oligo has a certain probability P to be in each state and can transition to a different state by either binding or dissociating from a monomer or a dimer. Applying the law of mass action allows to write the following three ordinary differential equations, which can further be simplified assuming the system is in equilibrium

$$\frac{dP_F}{dt}^{\!\!\!0} = k_{\text{off,DF}} * P_D + k_{\text{off,MF}} * P_M - k_{\text{on,M}} * [\text{ARF}] * P_F - k_{\text{on,D}} * [\text{ARF}_2] * P_F \tag{S3.1}$$

$$\frac{dP_M}{dt}^{\!\!\!0} = k_{\text{on,M}} * [\text{ARF}] * P_F + k_{\text{off,DM}} * P_D - k_{\text{off,MF}} * P_M - k_{\text{on,MM}} * [\text{ARF}] * P_M \tag{S3.2}$$

$$\frac{dP_D}{dt}^{\!\!\!0} = k_{\text{on,D}} * [\text{ARF}_2] * P_F + k_{\text{on,MM}} * [\text{ARF}] * P_M - k_{\text{off,DM}} * P_D - k_{\text{off,DF}} * P_D. \tag{S3.3}$$

Moreover, the sum of the probabilities of the DNA being in any of the three states must be unity

$$P_F + P_M + P_D = 1. \tag{S3.4}$$

Assuming that the fraction of ARF bound to the DNA is negligible compared to the one of ARF in solution, it is possible to derive the following equation linking the concentration of dimers in solution with the total concentration of ARF added to the solution

$$[\text{ARF}_2] = \frac{4[\text{ARF}]_T + K_I - \sqrt{K_I * (8[\text{ARF}]_T + K_I)}}{8} \tag{S3.5}$$

where K_I is the dissociation constant of the protein dimer into two monomers. Since the four states system forms a closed circle, one of the kinetic constant must be obtain from the equation imposing microscopic reversibility[1]

$$k_{\text{off,DM}} = K_I \frac{k_{\text{on,M}} * k_{\text{on,MM}} * k_{\text{off,DF}}}{k_{\text{off,MF}} * k_{\text{on,D}}}. \tag{S3.6}$$

The equations S3.1 to S3.6 form a system that can be simplified by substitution to obtain the equations for the expected P_M and P_D as function of $[\text{ARF}]_T$. Moreover, we can define the system as being function of a single microscopic association kinetic constant $k_{\text{on,mic}}$, which represents the kinetics of the association of a monomer to a single AuxRE. Then, the three k_{on}s of the four states model can be defined in relation to this single constant ($k_{\text{on,MM}} = k_{\text{on,mic}}$, $k_{\text{on,M}} = 2k_{\text{on,mic}}$ and $k_{\text{on,D}} = 2k_{\text{on,mic}}$).

$$P_M = \frac{8k_{\text{on,mic}}k_{\text{off,DF}}K_I\left(\left(K_I k_{\text{on,mic}}+k_{\text{off,MF}}\right)A-\frac{k_{\text{on,mic}}B}{2}+\left(-\frac{K_I}{2}+2[\text{ARF}]_T\right)k_{\text{on,mic}}-k_{\text{off,MF}}\right)}{\substack{8k_{\text{on,mic}}\left(-\frac{k_{\text{off,MF}}k_{\text{on,mic}}B}{8}+\left(\left(k_{\text{off,DF}}-\frac{3k_{\text{off,MF}}}{8}\right)K_I+\frac{[\text{ARF}]_T k_{\text{off,MF}}}{2}\right)k_{\text{on,mic}}+\frac{3k_{\text{off,DF}}k_{\text{off,MF}}}{2}\right)K_I A- \\ -4K_I\left(\left(k_{\text{off,DF}}-\frac{3k_{\text{off,MF}}}{4}\right)K_I-4[\text{ARF}]_T\left(k_{\text{off,DF}}+\frac{3k_{\text{off,MF}}}{4}\right)\right)k_{\text{on,mic}}^2+16k_{\text{off,DF}}k_{\text{off,MF}}^2- \\ -4k_{\text{on,mic}}\left(K_I\left(k_{\text{off,DF}}-\frac{k_{\text{off,MF}}}{4}\right)k_{\text{on,mic}}+k_{\text{off,MF}}^2\right)B+ \\ +4\left(\left(k_{\text{off,DF}}+k_{\text{off,MF}}\right)K_I+4[\text{ARF}]_T k_{\text{off,MF}}\right)k_{\text{off,MF}}k_{\text{on,mic}}}},$$

$$\text{(S3.7)}$$

$$P_D = \frac{-3k_{\text{on,mic}}\left(K_I k_{\text{on,mic}}\left(K_I-\frac{4[\text{ARF}]_T}{3}+\frac{B}{3}\right)A+\left(\frac{4k_{\text{off,MF}}}{3}-\frac{K_I k_{\text{on,mic}}}{3}\right)B-\left(K_I+4[\text{ARF}]_T\right)\left(K_I k_{\text{on,mic}}+\frac{4k_{\text{off,MF}}}{3}\right)\right)k_{\text{off,MF}}}{\substack{8k_{\text{on,mic}}\left(-\frac{k_{\text{off,MF}}k_{\text{on,mic}}B}{8}+\left(\left(k_{\text{off,DF}}-\frac{3k_{\text{off,MF}}}{8}\right)K_I+\frac{[\text{ARF}]_T k_{\text{off,MF}}}{2}\right)k_{\text{on,mic}}+\frac{3k_{\text{off,DF}}k_{\text{off,MF}}}{2}\right)K_I A- \\ -4K_I\left(\left(k_{\text{off,DF}}-\frac{3k_{\text{off,MF}}}{4}\right)K_I-4[\text{ARF}]_T\left(k_{\text{off,DF}}+\frac{3k_{\text{off,MF}}}{4}\right)\right)k_{\text{on,mic}}^2+16k_{\text{off,DF}}k_{\text{off,MF}}^2- \\ -4k_{\text{on,mic}}\left(K_I\left(k_{\text{off,DF}}-\frac{k_{\text{off,MF}}}{4}\right)k_{\text{on,mic}}+k_{\text{off,MF}}^2\right)B+ \\ +4\left(\left(k_{\text{off,DF}}+k_{\text{off,MF}}\right)K_I+4[\text{ARF}]_T k_{\text{off,MF}}\right)k_{\text{off,MF}}k_{\text{on,mic}}}},$$

$$\text{(S3.8)}$$

where $A = \sqrt{(8[\text{ARF}]_T + K_I)/K_I}$ and $B = \sqrt{K_I(8[\text{ARF}]_T + K_I)}$. Then, P_M and P_D are defined by 4 parameters: $k_{\text{on,mic}}$, $k_{\text{off,MF}}$, $k_{\text{off,DF}}$ and K_I. The experimental determination of the fraction bound is based on fitting two Gaussians at the E^*s of the free and dimer-bound DNA; the monomer is expected to reside halfway between these two populations and it is reasonable to assume that half of the monomer bound population will be erroneously accounted as free and half as bound. Following this assumption we can define the fraction bound as

$$F_B = P_D + \frac{1}{2}P_M. \tag{S3.9}$$

In the limit of an infinitely stable protein dimer (i.e. $K_I = 0$), the function for F_B (S3.9) simplifies to the simple binding isotherm determined only by the kinetic constants of the dimer binding

$$F_B = \frac{[\text{ARF}]_T}{[\text{ARF}]_T + 2\frac{k_{\text{off,DF}}}{k_{\text{on,D}}}}. \tag{S3.10}$$

On the other extreme, as the dimer becomes less stable in solution, the binding event of a protein dimer to the DNA becomes rare and, consequently, the binding events of monomers becomes more independent; in this scenario the function for F_B (S3.9) numerically tends to the Klotz equation[2]

$$F_{B,Kl} = \frac{1}{2}\frac{K_1[ARF] + 2K_1K_2[ARF]^2}{1 + K_1[ARF] + K_1K_2[ARF]^2}$$
(S3.11)

and the DNA can be seen as a receptor featuring two binding sites for the ligand monomer ARF. Here, K_1 and K_2 are stoichiometric association constants. For independent binding, the Klotz equation simplifies to a simple binding isotherm in which the dissociation constant is defined by the microscopic kinetic constants related to the interaction of one monomer to one binding site ($K_{d,mic} = k_{off,mic}/k_{on,mic}$). Strikingly, using the four-states model allows calculating the expected K_I in case of independent binding ($K_I = 1.4 \times 10^5$ nM for the global fit presented in Fig 3.5). This number represent the limit of instability of the dimer in solution over which the dimerization has no effect on the binding capabilities of the protein.

Summarizing the behaviour of the binding isotherm of the four-states system (Eq. S3.9), at values of K_I close to zero it approximates the simple binding isotherm of the dimer, for intermediate values of K_I the function is steeper capturing the expected cooperativity, while at high values of K_I the function can be approximated with the Klotz equation for independent binding and its steepness decreases back to a value consistent with a simple binding isotherm.

Supporting Note 2: Derivation of observed k_{on} and k_{off} four states system

Given the four state model we can write the following notable transition probabilities involving changes in the binding state of the DNA

$$P_{F \mapsto M}(t) = \frac{k_{on,M}}{FR} * [ARF] * e^{-\left(\frac{k_{on,D}}{FR} * [ARF_2] + \frac{k_{on,M}}{FR} * [ARF]\right)t}$$

$$P_{M \mapsto D}(t) = \frac{k_{on,MM}}{FR} * [ARF] * e^{-\left(\frac{k_{off,MF}}{FR} + \frac{k_{on,MM}}{FR} * [ARF]\right)t}$$

$$P_{F \mapsto D}(t) = \frac{k_{on,D}}{FR} * [ARF_2] * e^{-\left(\frac{k_{on,D}}{FR} * [ARF_2] + \frac{k_{on,M}}{FR} * [ARF]\right)t} \qquad (S3.12)$$

$$P_{D \mapsto M}(t) = \frac{k_{off,DM}}{FR} * e^{-\left(\frac{k_{off,DF}}{FR} + \frac{k_{off,DM}}{FR}\right)t}$$

$$P_{D \mapsto F}(t) = \frac{k_{off,DF}}{FR} * e^{-\left(\frac{k_{off,DF}}{FR} + \frac{k_{off,DM}}{FR}\right)t},$$

where FR is the frame rate in fps (f s^{-1}) and the time t is expressed in frames (f). Since the monomer is expected to have short dwell time on the DNA (compared to the frame time of 500 ms), the only transitions that can be observed are the ones that lead to a protein dimer bound to the DNA. Then, the observed transition probability associated with "binding" and "dissociation" can be derived as

$$P_{F,M \mapsto D} = \frac{P_F}{P_F + P_M}(P_{F \mapsto D} + P_{F \mapsto M} * P_{M \mapsto D} * \tau_M) + \frac{P_M}{P_F + P_M}P_{M \mapsto D}, \qquad (S3.13)$$

$$P_{D \mapsto F,M} = P_{D \mapsto F} + P_{D \mapsto M}, \qquad (S3.14)$$

where τ_M is the dwell time of the monomer on the DNA ($\tau_M = FR/(k_{off,MF} + k_{on,MM} * [ARF])$). Then, these probabilities can be converted to the observed k_{on} and k_{off}

$$k_{on,Obs} = \frac{P_{F,M \mapsto D}}{[ARF]_T}FR, \qquad (S3.15)$$

$$k_{off,Obs} = P_{D \mapsto F,M}FR. \qquad (S3.16)$$

We note that for the global fit the k_{on}s in the transition probabilities were substituted with a single k_{on} ($k_{on,mic}$, see Supporting note 1).

References

1. Colquhoun, D., Dowsland, K. A., Beato, M. & Plested, A. J. R. How to impose microscopic reversibility in complex reaction mechanisms. *Biophysical journal* **86,** 3510–3518 (2004).
2. Klotz, I. *Introduction to Biomolecular Energetics: Including Ligand–Receptor Interactions* (Elsevier Science, 2012).

3

Chapter 4

High-throughput, non-equilibrium studies of single biomolecules using glass-made nanofluidic devices

4

Abstract

Single-molecule detection schemes offer powerful means to overcome static and dynamic heterogeneity inherent to complex samples. However, probing biomolecular interactions and reactions with high throughput and time resolution remains challenging, often requiring surface-immobilized entities. Here, we introduce glass-made nanofluidic devices for the high-throughput detection of freely-diffusing single biomolecules by camera-based fluorescence microscopy. Nanochannels of 200 nm height and a width of several micrometers confine the movement of biomolecules. Using pressure-driven flow through an array of parallel nanochannels and by tracking the movement of fluorescently labelled DNA oligonucleotides, we observe conformational changes with high throughput. In a device geometry featuring a T-shaped junction of nanochannels, we drive steady-state non-equilibrium conditions by continuously mixing reactants and triggering chemical reactions. We use the device to probe the conformational equilibrium of a DNA hairpin as well as to continuously observe DNA synthesis in real time. Our platform offers a straightforward and robust method for studying reaction kinetics at the single-molecule level.

Introduction

Single-molecule techniques are playing an increasingly important role in identifying and analyzing static or dynamic interactions of (bio-)molecules with high spatiotemporal resolution[1–3]. Among the large number of available frameworks[4], single-molecule fluorescence detection (SMFD) schemes, such as confocal or camera-based total-internal reflection fluorescence (TIRF) microscopy, are frequently chosen due to their simplicity, robustness and ease of use[5]. However, users often have to choose between high throughput or high time resolution[6–8], many experiments require surface immobilization techniques to extend the observation time, and it is difficult to access non-equilibrium conditions or follow biochemical reactions.

Many creative solutions have been proposed to overcome these hurdles using fluidic platforms[9,10]. These include mixers for studying single-molecule kinetics[11,12], titration devices[13], devices to confine molecules in polymer nanochannels[14], electroosmotic molecular traps[15], or microfluidic droplets containing individual enzyme analytes[16]. These fluidic platforms are, despite their clear benefits, not yet widely used due to complex fabrication procedures, limited reusability and configurability or restrictions in integration into microscopy platforms.

Here, we introduce a fluidic platform for SMFD consisting of nano-/microchannel devices fabricated entirely in glass thereby overcoming limitations imposed by previous designs utilizing polydimethylsiloxane (PDMS) such as the susceptibility to swelling by some organic solvents and the limited modifiability of the surface to prevent non-specific adsorption. The design goals of our platform were (1) to achieve single-molecule detection at high time resolution by geometrically confining the flow through nanochannels (200 nm channel height and micro-scale width and length), (2) to use camera-based detection to monitor many molecules in parallel, (3) to enable straightforward device integration, and (4) to observe chemical reactions at the single-molecule level and in real time using mixing in a T-shaped nanochannel.

We utilize the new nanofluidic devices for high-throughput sensing of short DNA oligonucleotides in an array of parallel nanochannels by detecting single-molecule Förster resonance energy transfer (smFRET) between a donor and an acceptor fluorophore attached to the DNA. Using an updated excitation scheme, several hundred thousand events are combined to gather reliable single-molecule data within minutes.

We probe the conformational equilibrium of a DNA hairpin in nanochannel devices with a mixing geometry. By imaging the T-junction, we probe conformations before, during and after mixing with a buffer containing a high salt concentration that stabilizes the closed hairpin conformation.

Finally, we continuously monitor reaction kinetics at the single-molecule level by observing DNA synthesis upon mixing DNA oligonucleotides, DNA polymerases and nucleotides. Reactions are observed by a change in FRET efficiency downstream in the outlet nanochannel allowing us to translate a *spatial position* into a corresponding moment in the *time evolution* of a reaction. Thus, our platform demonstrates the ability for a wide range of applications in probing complex kinetics.

Materials and methods

Device fabrication and surface passivation

Nano-/microfluidic devices were fabricated by Micronit with a 45 mm × 15 mm footprint to be compatible with a commercially available chip holder Fluidic Connect PRO (Micronit Microtechnologies B.V., The Netherlands). In brief, the devices consist of two thermally bonded borosilicate glass layers with photolithographically defined and wet-etched nano- and microchannels. Microchannels and ports to connect to tubing are structured in the upper 1 mm thick glass, while nanochannels are etched into the bottom layer. Here, coverslip type D263 borosilicate glass (Schott) with a thickness of 175 μm was used enabling high-resolution fluorescence imaging with inverted oil-immersion microscope objectives with shortest working distances whilst minimizing undesirable autofluorescence.

To prevent non-specific adsorption in experiments with proteins, channels were passivated with PEG using a variation of a method described previously[17]. The fluidic devices were first flushed and incubated (3 × 5 min) with a 1 : 50 (vol/vol) Vectabond : acetone solution. All subsequent washing and passivation steps were performed by flushing the channels with ∼100 μL of the respective solutions. PEGylated channels were filled with PBS and stored in a humid chamber at 4 °C.

DNA and buffers

Fluorescently labelled oligonucleotides were ordered from IBA, Germany. To construct the DNA hairpin, a 30-mer primer sequence (biotin-5'-CCT CAT TCT TCG

TCC CAT TAC CAT ACA TCC-3') was annealed to a 75-mer hairpin sequence (5'-TGG ATT AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA TCC ATT GGA TGT ATG GTA ATG GGA CGA AGA ATG AGG-3'). The primer was internally labelled with ATTO647N at the -12 position; the hairpin was labelled with Cy3B at the 5' end. Gapped DNA, used to study the binding of DNA polymerase I (Klenow fragment, KF), was constructed using the same primer sequence, annealed to a template strand (5'- CCA CGA AGC AGG CTC TAC TCT CTA AGG ATG TAT GGT AAT GGG ACG AAG AAT GAG G-3') and a downstream complementary strand (5'-TAG AGA GTA GAG CCT GCT TCG TGG-3'). The template strand was labelled with Cy3B at the +12 position. The DNA sensor used for the polymerization experiment consists of the same primer and template sequences, but with ATTO647N on the -7 position and Cy3B on the +25 position, respectively. dNTPs were ordered from Sigma-Aldrich/Merck, Germany.

DNA constructs (as well as KF, if specified) were diluted in an imaging buffer containing 50 mM Tris HCl (pH 7.5), 100 µg/mL BSA, 10 mM $MgCl_2$, 5 % glycerol, 1 mM DTT, 1 mM Trolox, 1 % glucose oxidase/catalase and 1 % glucose. Trolox is a triplet-state quencher and prevents fluorophore blinking. Glucose, glucose oxidase and catalase was used as an oxygen scavenger system to prevent premature photobleaching of fluorophores[18,19]. The concentration of gapped DNA was 1 nM and, if used, the concentration of KF was 10 nM. DNA hairpin concentrations were 500 pM in parallel channels and 1 nM in the mixing channel (DNA hairpins were diluted in a similar imaging buffer without magnesium, but with additional NaCl as specified). Prior to mixing, the concentration of DNA polymerization sensors was 1 nM; the concentration of KF was 5 nM, and the concentration of dNTPs was 200 µM each. For this polymerization experiment, we added neutravidin directly to the imaging buffer in a concentration of 0.6 µg/mL to block 5' end of the biotinylated DNA primer. We found this prevents the formation of a low E* state, the cause of which is probably binding of multiple KF polymerases to the same DNA molecule. The imaging buffer was applied to a 100 µL syringe (ILS, Germany), which was then connected to the nanofluidic device using ethylene tetrafluoroethylene (ETFE) tubing (1/16″ outer diameter, 0.010″ inner diameter; Micronit Microtechnologies B.V., The Netherlands).

4

Single-molecule detection

We used a home built TIRF microscope and a fiber-coupled laser engine (Omicron, Germany) equipped with lasers of four different wavelengths (405 nm, 473 nm, 561 nm, and 638 nm). Laser intensities were independently controlled by a home-written LabVIEW program. Divergent light from the fiber output was collimated ($f = 30$ mm, Thorlabs, Germany) and focused by a second lens ($f = 200$ mm, Thorlabs, Germany) into the backfocal plane of a $100\times$ NA 1.49 objective (Nikon, Japan). A polychroic filter and a multi-bandpass filter (Chroma, USA) prevented laser light from entering the emission path. A tube lens focused the emission on an aperture, which reduced the effective field of view to a rectangle. Next, the light was spectrally split into a blue, a green and a red emission channel. The three beams were focused on an Ixon Ultra 897 emCCD (Andor, UK) with 512 by 512 pixels (maximum acquisition rate: 56 Hz at full frame and 100 Hz after cropping the frame to 343 by 256 pixels). In our configuration, one pixel on the camera corresponded to a length and width of 112 nm in the sample plane. We used a rapid automated modular microscope (RAMM) system as a stage holder (ASI, USA), combined with motorized x, y-scanning stage and a z-piezo for focusing.

Molecules were excited with a laser power of 130 mW (measured after the fiber output, 561 nm and 638 nm laser) for the fluidic experiments and with 30 mW (561 nm) and 15 mW (638 nm) for the surface immobilized experiments. Excitation in total-internal reflection mode ensured the highest possible signal-to-noise ratio. A stroboscopic alternating-laser excitation (sALEX[20]) scheme was used to reduce motion blur of diffusing molecules. Laser pulse widths were 1.5 ms (fluidic devices) and 3 ms (surface immobilized experiments) in a frame time of 10 ms. Green and red pulses were aligned back-to-back, so that particle movement between a green and a red frame was minimal.

Extracting intensities and positions from single molecules

Particles were localized and tracked with a home-written variation on GaussStorm[21] (Matlab). Time traces, histograms and binned maps were generated with custom-built software packages. We applied a bandpass filter to find local intensity maxima before fitting the local maxima with elliptical 2D Gaussian functions from which we obtained the photon count as well as the position with sub-pixel accuracy[21]. After filtering of localizations based on their intensities, our algorithm links corresponding

particles in the green and red channels and calculates FRET efficiency E^* and stoichiometry ratio S for each molecule.

Tracking of single molecules and determination of flow speeds and diffusion coefficients

A tracking algorithm was applied to track particles from frame to frame[22]. The distribution of the displacements of single particles was used to calculate mean flow speed and mean diffusion coefficient in the entire field of view or in selected regions. In the case of one-dimensional diffusion along the direction of the flow, the probability distribution for the position of the particle (ρ) is Gaussian with variance proportional to the diffusivity D of the particle ($\sigma^2 = 2Dt$) and mean proportional to the flow speed v ($\mu = vt$). For an example of displacement distribution see Fig. S4.5.

FRET and alternating laser excitation (ALEX)

The apparent FRET efficiency E^* is calculated from the emission intensities of donor and acceptor after donor excitation (denoted DD and DA, respectively) for each molecule in each camera frame according to

$$E^* = \frac{DA}{DD + DA}.$$
(4.1)

To verify the presence of an acceptor fluorophore on the DNA we applied alternating laser excitation (ALEX) in which every frame of donor excitation is followed by a direct excitation of the acceptor fluorophore using a second laser resulting in a third photon stream (AA) for each molecule[23–25]. The detection of AA in addition to DD and DA allows for calculating the stoichiometry ratio S, defined as

$$S = \frac{DD + DA}{DD + DA + AA}.$$
(4.2)

S can be used to filter molecules: molecules with a stoichiometry close to 0 have no photoactive donor (e.g. because of donor bleaching), and molecules with a stoichiometry close to 1 have no photoactive acceptor. Depending on the ratio of the laser intensities for direct donor and acceptor excitation, a stoichiometry around 0.5 represents molecules having both a photoactive donor and acceptor that can undergo FRET. In this work, we show FRET data based on molecules that have

stoichiometry values $0.3 \leq S \leq 0.8$ (nanochannel data) or $0.5 \leq S \leq 0.9$ (immobilized hairpins).

Results

Single-molecule fluorescence detection of nanometer sized, non-immobilized entities is challenging due to the limited photon budget of the fluorophores. In the case of surfaceimmobilized molecules, the integration of photon emission from a diffraction limited spot quickly leads to an acceptable signal-to-noise ratio, but moving particles distribute the emitted photons over a larger area on the camera, compromising the signal-to-noise ratio. For this reason, fluorophore excitation must be limited to an interval as short as possible. In the following, we describe the fluidic design and flow control enabling conditions in which the induced flow is slow enough as not to cause additional broadening of the point spread function, and enables the observation of DNA–protein interactions, salt-dependent conformational equilibria of DNA, and enzymatic driven DNA synthesis.

Achieving low nanochannel flow speeds *via* parallel flow control

The first device consists of parallel nanochannels wet-etched into a first glass wafer and connected via wafer-bonding with feeding microchannel etched into a second glass wafer (Fig. 4.1, see also Materials and methods). We use the principle of parallel flow control (PFC[26,27]) in which liquid driven by a syringe pump is divided into nanochannels and a parallel bypassing microchannel. Due to the high ratio of hydraulic resistance between the nanochannels and the microchannel, the flow speed is reduced by orders of magnitude in the nanochannel, and the dead volume in feeding channels and tubing is replaced quickly (see also Fig. S4.2 and Supporting Note 1). The resulting typical flow rates in the nanochannels are in the range of 0.2–20 pL/min per channel, corresponding to average flow velocities of 5–500 nm/ms. We validated our implementation of PFC by measuring the flow speeds (see Materials and methods) and computing velocity profiles at different pump rates (Fig. S4.3). The shape and magnitude of the velocity profile show an excellent agreement with the theoretical model for Poiseuille flow in rectangular nanochannels at the pump rate tested experimentally (Fig. S4.4, Supporting Note 2).

The second device applies, for the first time, the PFC approach to a mixing geometry. Here, two syringe pumps deliver flow to two microchannels and two respective parallel feeding nanochannel inlets which merge to a single nanochannel
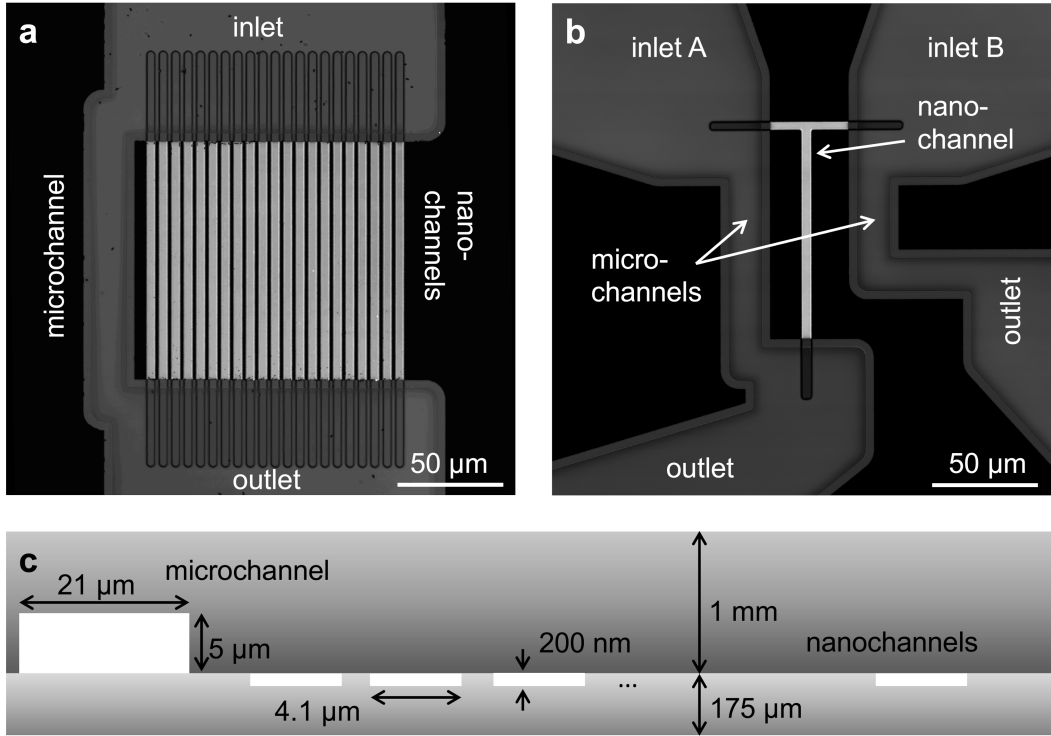
Figure 4.1: Design of glass nanofluidic chips. (a and b) High-resolution confocal scans based on reflection of light. The parallel channel design (a) contains 21 straight nanochannels (l × w × h: 120 by 4.1 by 0.2 μm) and a microchannel (w × h: 21 by 5 μm). The design of the mixing device (b) contains a single T-shaped nanochannel (horizontal part: l × w × h: 40 by 3.8 by 0.2 μm, vertical part l × w × h: 100 by 4.7 by 0.2 μm) in between two microchannels (l × w × h: 95 by 21 by 5 μm). Fluidic chips fit in a Micronit chip holder for easy connection to tubes and pumps (see also Fig. S4.1). (c) Schematic cross-section of the parallel channel array, showing the dimensions of the microchannel and the nanochannels (not to scale).

at a T-junction (Fig. S4.2). The mass transfer in such device is summarized by the Péclet number

$$\text{Pe} = \frac{v_y w}{D} \tag{4.3}$$

in which v_y is the velocity along the channel's axis, w is the channel's width and D is the diffusion coefficient of the transported species. This dimensionless number evaluates the relative magnitude of convection (flow along the channel) to diffusion (Brownian motion), and in the T-sensor geometry it estimates how many channel widths downstream of the junction mixing will be completed[28]. Given the low veloc-

ities required to achieve single molecule detection and the high diffusion coefficients of the molecules and ions used in our experiments, the mixing nanochannels operates at Péclet number ranging from 0.01–100. In the limit of low Péclet numbers (Pe < 10), axial diffusion (i.e. diffusion of a species from one inlet upstream into the other inlet) has been predicted for this geometry but had never been experimentally observed[29,30]. The presence of axial diffusion is confirmed in our experiments.

Detecting fluorescence of moving particles inside nanochannels

As probes, we opted for doubly labelled DNA molecules providing a FRET efficiency E^* that is calculated using photon numbers after donor excitation emitted by the donor (DD) and the acceptor (DA), respectively (see Materials and methods). To verify the presence of the acceptor, we applied alternating-laser excitation (ALEX) in which every camera frame of donor excitation is followed by a frame of direct acceptor excitation creating photon numbers AA (Materials and methods)[23,24]. Since motion blur severely affects the ability to determine emission intensities of individual molecules inside the nanochannels we opted for stroboscopic ALEX (sALEX[20]) in which the molecules are only excited for a fraction of the camera exposure time. We further aligned the green and red laser pulses back-to-back to facilitate successive localizations in the acceptor channel allowing to link donor and acceptor signals (Fig. 4.2a, top). The stroboscopic excitation pulses (1.5 ms) lead to spots which are barely affected by motion blur and have an excellent signal-to-noise ratio allowing us to fit 2D Gaussians to obtain photon numbers and positions with sub-pixel localization accuracies for individual DNA molecules (Fig. 4.2a, bottom).

Parallel nanochannels: visualizing DNA–protein interactions

To prove that, after passivating the glass surface using PEG (see Materials and methods), our parallel device can be used to study protein binding without immobilization, we measured a sample containing 1 nM gapped dsDNA in the absence and presence of 10 nM of DNA polymerase I (Klenow fragment, KF). The nanochannels were imaged for five minutes obtaining $\sim 10^5$ FRET efficiency readouts from single molecules flowing through the field of view; this allowed us to compute high quality FRET efficiency histograms (Fig. 4.2b).

The mean FRET efficiency for the DNA molecules in the absence of KF shows a single species centered around $E^* = 0.4$. The addition of KF
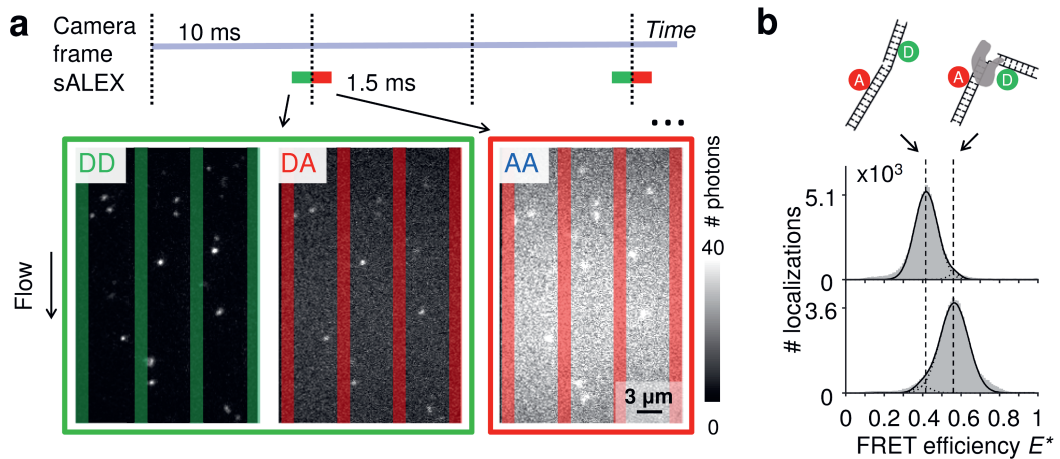
Figure 4.2: Detection of single molecules and DNA–protein interactions inside parallel nanochannels. (a) Top: Schematic of stroboscopic alternating laser excitation (sALEX); the excitation time is considerably shorter than the required acquisition time of a camera frame, reducing motion blur. The laser pulses are placed back to back to facilitate the linking of the AA signal to the correct DD/DA couple of signals. Bottom: Example of raw movie frames showing the gapped DNA construct flowing through the parallel nanochannels at a flow speed of 218(1) nm/ms (mean value and 95 % interval of confidence). Excitation colors are indicated by the surrounding boxes. For each molecule, photon counts of DD and DA after donor excitation are determined simultaneously. The photon counts of AA after acceptor excitation are collected during the next camera frame. Nanochannel boundaries are indicated with green bars (green detection channel) or red bars (red detection channel). (b) Top: Schematic of a gapped DNA construct in its free and bound conformation. The dsDNA is labelled with a donor and acceptor dyes located on opposite sides of the one-nucleotide gap. Bottom: E^* histograms (100 bins) of the gapped DNA construct, measured in the parallel channels for 5 minutes. Top: 1 nM DNA. Bottom: 1 nM DNA in presence of 10 nM DNA polymerase I (KF), a DNA polymerase known to change the conformation of gapped DNA upon binding. Dashed lines are added for visual guidance. For full E^*/S histograms see Fig. S4.7a.

causes an increase of FRET efficiency to 0.6 which reflects the expected shortened donor–acceptor distance upon binding and DNA bending[31]. We note that using freely diffusing molecules, we obtained a second independent readout for the binding in the form of the decreased apparent diffusion coefficient along the channel upon KF addition and complex formation from $D_{\mathrm{DNA}} = 35.3(3)\,\mu\mathrm{m}^2/\mathrm{s}$ to $D_{\mathrm{DNA/KF}} = 30.6(3)\,\mu\mathrm{m}^2/\mathrm{s}$.

Parallel nanochannels: resolving salt-dependent conformational equilibria of DNA hairpins

To compare the accessible dynamic range of FRET measurements in nanochannels with those of immobilized entities, we utilized DNA hairpins that can interconvert between an open (low FRET) and closed (high FRET) conformations. The chosen DNA hairpin is mostly open at $0\,M$ NaCl and mostly closed at $1\,M$ NaCl, while showing fast opening and closing at an intermediate concentration of $0.5\,M$ NaCl[20]. Upon addition of salt, we further note a small shift of the FRET peak of the open conformation due to a salt-depended compaction of the single stranded DNA overhang[32]. The FRET histograms of the freely flowing DNA hairpins are similar to the immobilized ones (Fig. 4.3a), indicating that we access a broad FRET range and resolve the individual conformational states of "open" and "closed". Combining consecutive localizations from individual molecules allowed us to obtain time traces that show opening and closing of individual DNA hairpins in real time (Fig. 4.3b).

Nanofluidic mixing: spatially resolving shifts in the conformational equilibria of DNA hairpins

Next, we validated the mixing devices by shifting the conformational equilibrium of the DNA hairpin from mostly open to mostly closed. To this aim, we mixed a buffer containing $1\,M$ NaCl with a buffer containing DNA hairpins at $0\,M$ NaCl. The viscosity of the solution containing NaCl is expected to be approximately $10\,\%$ higher than the solution with zero salt[33]; the pump rates were then adjusted by visual inspection.

We constructed a map of spatially resolved, binned FRET efficiencies for various flow rates (Fig. 4.3c–e, center). For each bin, the color indicates the median FRET efficiency, but 1D FRET efficiency histograms can be reconstructed from any part of the field of view. We selected two regions, one from the right inlet and one below the mixing intersection (Fig. 4.3c–e, top and bottom histograms).

At the three pump settings tested, the mixing device operates at Pe < 20 for DNA ($D_{DNA} \approx 35\,\mu m^2/s$) and Pe < 0.4 ($D_{NaCl} \approx 2000\,\mu m^2/s$[28]) for salt ions. Under these conditions, identification of a well-localized onset of mixing is not possible since axial diffusion becomes relevant: at the lowest flow speeds at which DNA molecules are transported with Pe ≈ 5 (Fig. 4.3c, center), many molecules diffuse back in the salt inlet whereas maintaining a convective displacement toward the
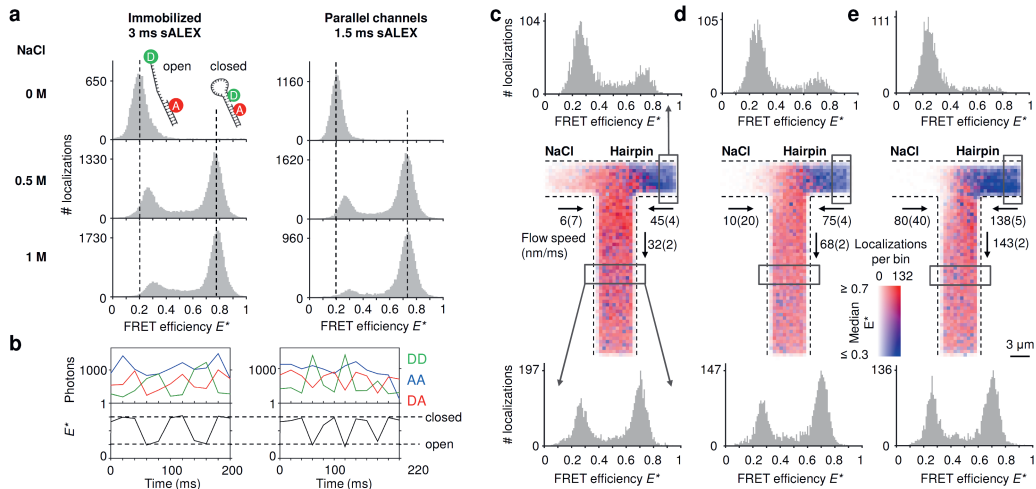
Figure 4.3: Detecting conformational equilibria of DNA hairpins in nanochannels. (a) Single-molecule detection of DNA molecules immobilized or non-immobilized within parallel nanochannels. The conformational equilibrium of DNA hairpins is salt dependent; DNA hairpins are mostly open (low FRET) at [NaCl] = 0 M; the conformational equilibrium shifts towards the closed state (high FRET) with increasing salt concentrations. The DNA hairpin shows similar behavior when imaged in parallel nanochannels compared to a standard immobilized sample. For full E^*/S histograms see Fig. S4.7b. (b) Single-molecule time traces obtained from tracking individual molecules flowing through the channels at [NaCl] = 0.5 M. The AA signal (blue trace) remains largely constant. The clear anti-correlation of the DD (green trace) and the DA (red trace) signal and the resulting FRET time traces (black trace) indicate conformational changes in the millisecond timescale of the DNA hairpin during its passage. Dashed vertical lines are added for visual guidance indicating the expected FRET efficiencies of the open and closed DNA hairpin conformation, respectively. (c–e) Mixing of a high salt solution [NaCl] = 1 M with DNA hairpins [NaCl] = 0 M at increasing flow speeds. Center: binned maps (4×4 camera pixels per bin), in which color represents the median FRET efficiency and opacity indicates the number of localizations in a bin over the time span of the measurement. Given the low Péclet number, axial diffusion of the DNA hairpin in the salt inlet is visible at lower flow speeds (given with 95 % interval of confidence) (c and d) but becomes negligible at the highest one (e). Top: FRET histograms reconstructed from the inlet (boxed region); the salt diffuses into the inlet channel of the hairpin stabilizing the closed conformation of the hairpin at the lowest flow rate; this effect decreases at higher flow rates. Bottom: FRET histograms reconstructed below the T-junction (boxed region); at higher flow speed, the concentration of the salt in the outlet decreases because the effect of axial diffusion becomes less prevalent originating a rise in the open conformation compared to the lower flow speeds.

junction. The effect of axial diffusion on DNA transport diminishes at higher flow rates in agreement with finite element Comsol simulations for these transport conditions (Fig. S4.6a, Supporting Note 3). As the diffusion coefficient of salt is almost two orders of magnitude higher than the one of the DNA hairpin the simulations show a considerable amount of axial diffusion for this species; in particular, the concentration of the salt in the outlet is predicted to be between 400 mM for the highest flow rate up to 500 mM at the lowest flow rate. Experimentally, the salt concentration in a given region of a nanochannel can be estimated from the FRET efficiency profile reconstructed from that region. The histograms generated from the inlets (Fig. 4.3c and d, top) confirm that the salt diffuses to the other inlet indicated by a small "closed" population and its relative decrease with increasing flow speed. The histograms taken below the intersection show a relative increase of the "open" population for increasing flow speeds, consistent with the expected lower salt concentration.

Nanofluidic mixing: triggering enzymatic catalysis of DNA synthesis

To establish our mixing device as a tool for non-equilibrium studies, we utilized a FRET-based DNA polymerization sensor. The sensor consists of a primer DNA, labelled with an acceptor fluorophore, that is annealed to a long template DNA in which the single stranded overhang (25 bases) is labelled with a donor fluorophore (see Materials and methods). Since the single strand overhang is randomly coiled, the FRET efficiency is high ($E^* \sim 0.6$) before decreasing substantially upon addition of DNA polymerases and nucleotides (dNTPs: deoxyribonucleotide triphosphates) to $E^* \sim 0.2$. When the DNA construct is immobilized on a surface, we previously found that full polymerization adding 25 bases takes on average 1.6 s[34]. Due to this long reaction time, we acquired movies using three different field of views and stacked them together (Fig. 4.4). Two different pump settings, which resulted in a flow speed of approximately 20 µm/s and 10 µm/s in the common nanochannel and a residence time in a given field of view of 1.5 s and 3 s respectively.

In the control without dNTPs, the median FRET efficiencies binned and traced back to the field of view stays constant (Fig. 4.4b); upon addition of nucleotides in the left inlet (200 µM), the mapped median FRET efficiencies decrease from $E^* \sim 0.6$ to $E^* \sim 0.2$ (Fig. 4.4c; T-maps) as the underlying population of the fully polymerized sensor increases (Fig. 4.4b, 1D FRET histograms). Halving the
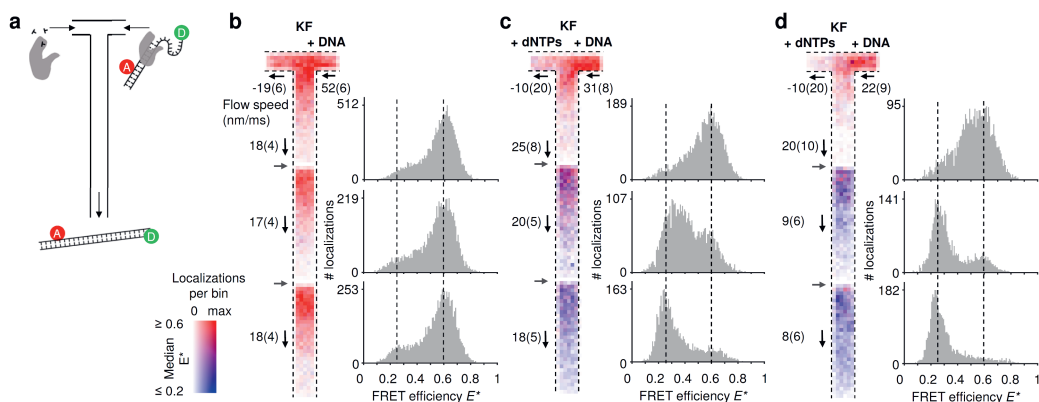
Figure 4.4: Monitoring DNA polymerization in mixing nanochannels. (a) Using the T-configuration, nucleotides and DNA polymerases (KF) are flowing from the left inlet encountering the doubly labelled DNA polymerization sensor and DNA polymerases from the right inlet. Successful DNA polymerization increases the distance between donor and acceptor fluorophore leading to a decrease in FRET efficiency. (b–d) Binned maps (8 × 8 camera pixels per bin) and corresponding FRET efficiency histograms containing data from three adjacent field of views (FOV); the borders of new FOVs are indicated with grey arrows. Color maps represent the median FRET efficiency; opacity indicates the number of localizations per bin, normalized for each map separately. Fewer molecules are localized in the lower parts of each FOV due to premature fluorophore photobleaching. Experimentally derived flow speeds are given with their mean value and 95 % interval of confidence. (b) Control experiment in absence of nucleotides. The DNA sensors show no changes in the mapped FRET efficiencies (map max = 126 localizations per bin) and the FRET histograms (100 bins) establish $E^* = 0.6$ as the center point of the high FRET species. (c) After addition of nucleotides via the left inlet, the mapped median FRET efficiencies of the FRET sensor decreases from ~ 0.6 to 0.25 towards the outlet. Almost complete polymerization is achieved at the third FOV (map max = 66). (d) Decreasing the flow speed by half, shows that the polymerization is be almost completed in the second FOV, consistent with the doubling the residence time in the first FOV (map max = 58). For full E^*/S histograms see Fig. S4.7c.

flow speed doubles the residence time in a given field of view and the polymerization is almost complete already at the second field of view (Fig. 4.4d) confirming the expected polymerization time of 1–2 s. The single molecule displacement analysis revealed the presence of backflow in the dNTPs inlet; nevertheless, simulations of such conditions revealed a concentration of dNTP after the junction higher than 20 μM for the fast flow speed and higher than 40 μM for the slow flow speed (Fig. S4.6b). The

effect of the axial diffusion of the dNTPs in the T-sensor allowed the polymerization to take place even in presence of backflow.

Discussion

We introduced two nanofluidic device designs for single-molecule detection under equilibrium and non-equilibrium conditions: the first one with a series of parallel nanochannels for equilibrium studies, and the second one with a T-shaped nanochannel for non-equilibrium studies. Both designs combine camera-based single-molecule detection, which allows monitoring many fluorescent molecules in parallel, with constant replenishment of freely flowing molecules. An updated stroboscopic excitation scheme utilizing back-to-back illumination resulted in near-circular emission spots on the camera sensor. Furthermore, in both designs we employed the concept of parallel flow control allowing us to work with simple syringe pumps whilst obtaining extremely low flow speeds that are necessary for single-molecule detection using image-based data acquisition.

Our design featuring the parallel array of nanochannels is especially suited for high-throughput measurements. Using 1 nM of fluorescently labelled DNA hairpins, we obtained a throughput of $\sim 10^4$ FRET data points per minute of measurement, even though the current field of view on our setup (29 µm by 19 µm) covers only 3 out of 21 channels.

In single-focus, diffusion based confocal microscopy, fluorescence bursts of molecules passing the focus are collected with the time between individual bursts being kept long enough to avoid doubly occupancy in the focal volume[7]. Assuming 500 ms required between obtaining two single FRET data points, a 1 minute measurements yields only around 100 FRET data points; more than two orders of magnitude less than in our nanochannels. The throughput of camerabased smFRET detection with immobilized samples can in principle be higher than in our devices[35], but requires elaborate protocols for surface immobilization restricting the range of potential applications and potentially introducing surface-induced artefacts. Furthermore, as experiments at high time resolution require high laser powers leading to premature photo-bleaching, the continuous replenishment of molecules in our method is advantageous.

Using our nanochannel mixing devices, we first accessed non-equilibrium conditions by mixing primarily open DNA hairpins from one inlet with a high-salt

solution from a second inlet triggering hairpin closing. Additionally, we observed polymerization of 25 bases on a DNA template by a DNA polymerase, illustrating that complex biological reactions can be followed in real time and in a continuous fashion.

In our current mixing design, the distance from junction to outlet is 100 µm, corresponding to a maximum residence time in the nanochannel of around 10 s. To gain access to further time points after mixing, designs using meandering channels could be implemented as demonstrated for confocal microscopy[11,36] or widefield microscopy[37]. Furthermore, our current field of view is cropped by a factor of two do ensure data acquisition at 100 Hz. With the use of faster cameras (sCMOS) and by reducing the overall magnification of the optical system (*e.g.*, by replacing the $100\times$ TIRF objective with a $60\times$ objective), more molecules could be simultaneously observed for a longer time over a larger area and with greater time resolution[35].

Conclusions

The introduced devices enable single-molecule studies with high throughput and millisecond time resolution under both equilibrium and non-equilibrium conditions. We used our platform to determine conformational changes and kinetic DNA synthesis by mapping the timescale of single reaction as a position in the nanochannel. We believe that the flexibility of the underlying design principles and the glass-only fabrication will allow for many applications in the (bio-) chemical sciences.

References

1. Smiley, R. D. & Hammes, G. G. Single Molecule Studies of Enzyme Mechanisms. *Chem. Rev.* **106,** 3080–3094 (2006).
2. Claessen, V. I., Engelkamp, H., Christianen, P. C., Maan, J. C., Nolte, R. J., Blank, K. & Rowan, A. E. Single-Biomolecule Kinetics: The Art of Studying a Single Enzyme. *Annual Rev. Anal. Chem.* **3,** 319–340 (2010).
3. Küchler, A., Yoshimoto, M., Luginbühl, S., Mavelli, F. & Walde, P. Enzymatic reactions in confined environments. *Nature Nanotech* **11,** 409–420 (2016).
4. Walter, N. G., Huang, C.-Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat Methods* **5,** 475–489 (2008).
5. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in Single-Molecule Fluorescence Methods for Molecular Biology. *Annu. Rev. Biochem.* **77,** 51–76 (2008).
6. Hohlbein, J., Gryte, K., Heilemann, M. & Kapanidis, A. N. Surfing on a new wave of single-molecule fluorescence methods. *Phys. Biol.* **7,** 031001 (2010).
7. Santoso, Y., Joyce, C. M., Potapova, O., Reste, L. L., Hohlbein, J., Torella, J. P., Grindley, N. D. F. & Kapanidis, A. N. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *PNAS* **107,** 715–720 (2010).

8. Hohlbein, J., Aigrain, L., Craggs, T. D., Bermek, O., Potapova, O., Shoolizadeh, P., Grindley, N. D. F., Joyce, C. M. & Kapanidis, A. N. Conformational landscapes of DNA polymerase I and mutator derivatives establish fidelity checkpoints for nucleotide insertion. *Nature Communications* **4,** 2131 (2013).

9. Mathwig, K., Chi, Q., Lemay, S. G. & Rassaei, L. Handling and Sensing of Single Enzyme Molecules: From Fluorescence Detection towards Nanoscale Electrical Measurements. *ChemPhysChem* **17,** 452–457 (2016).

10. Banterle, N. & Lemke, E. A. Nanoscale devices for linkerless long-term single-molecule observation. *Current Opinion in Biotechnology. Systems biology • Nanobiotechnology* **39,** 105–112 (2016).

11. Wunderlich, B., Nettels, D., Benke, S., Clark, J., Weidner, S., Hofmann, H., Pfeil, S. H. & Schuler, B. Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat Protoc* **8,** 1459–1474 (2013).

12. Gambin, Y., VanDelinder, V., Ferreon, A. C. M., Lemke, E. A., Groisman, A. & Deniz, A. A. Visualizing a one-way protein encounter complex by ultrafast single-molecule mixing. *Nat Methods* **8,** 239–241 (2011).

13. Kim, S., Streets, A. M., Lin, R. R., Quake, S. R., Weiss, S. & Majumdar, D. S. High-throughput single-molecule optofluidic analysis. *Nat Meth* **8,** 242–245 (2011).

14. Tyagi, S., VanDelinder, V., Banterle, N., Fuertes, G., Milles, S., Agez, M. & Lemke, E. A. Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat Meth* **11,** 297–300 (2014).

15. Lesoine, J. F., Venkataraman, P. A., Maloney, P. C., Dumont, M. E. & Novotny, L. Nanochannel-Based Single Molecule Recycling. *Nano Lett.* **12,** 3273–3278 (2012).

16. Arayanarakool, R., Shui, L., M. Kengen, S. W., Berg, A. v. d. & T. Eijkel, J. C. Single-enzyme analysis in a droplet-based micro- and nanofluidic system. *Lab on a Chip* **13,** 1955–1962 (2013).

17. Evans, G. W., Hohlbein, J., Craggs, T., Aigrain, L. & Kapanidis, A. N. Real-time single-molecule studies of the motions of DNA polymerase fingers illuminate DNA synthesis mechanisms. *Nucleic Acids Research* **43,** 5998–6008 (2015).

18. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* **3,** 891–893 (2006).

19. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131,** 5018–5019 (2009).

20. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17,** 27862–27872 (2015).

21. Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J. & Kapanidis, A. N. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal* **99,** 3102–3111 (2010).

22. Crocker, J. C. & Grier, D. G. Methods of Digital Video Microscopy for Colloidal Studies. *Journal of Colloid and Interface Science* **179,** 298–310 (1996).

23. Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E. & Weiss, S. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *PNAS* **101,** 8936–8941 (2004).

24. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43,** 1156–1171 (2014).

25. Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kühnemuth, R., Rezaei Adariani, S., Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V., Bowen, M. E., Chen, H., Cordes, T., Eilert, T., Fijen, C., Gebhardt, C., Götz, M., Gouridis, G., Gratton, E., Ha, T., Hao, P., Hanke, C. A., Hartmann, A., Hendrix, J., Hildebrandt, L. L., Hirschfeld, V., Hohlbein, J., Hua, B., Hübner, C. G., Kallis, E., Kapanidis, A. N., Kim, J.-Y., Krainer, G., Lamb, D. C., Lee, N. K., Lemke, E. A., Levesque, B., Levitus, M., McCann, J. J., Naredi-Rainer, N., Nettels, D., Ngo, T., Qiu, R., Robb, N. C., Röcker, C., Sanabria, H., Schlierf, M., Schröder, T., Schuler, B., Seidel, H., Streit, L., Thurn, J., Tinnefeld, P.,

Tyagi, S., Vandenberk, N., Vera, A. M., Weninger, K. R., Wünsch, B., Yanez-Orozco, I. S., Michaelis, J., Seidel, C. A. M., Craggs, T. D. & Hugel, T. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nature Methods* **15,** 669–676 (2018).

26. Liang, H., Nam, W. J. & Fonash, S. J. A Novel Parallel Flow Control (PFC) System for Syringe-Driven Nanofluidics. **3,** 3 (2008).

27. Mathwig, K., Mampallil, D., Kang, S. & Lemay, S. G. Electrical Cross-Correlation Spectroscopy: Measuring Picoliter-per-Minute Flows in Nanochannels. *Phys. Rev. Lett.* **109,** 118302 (2012).

28. Squires, T. M. & Quake, S. R. Microfluidics: Fluid physics at the nanoliter scale. *Rev. Mod. Phys.* **77,** 977–1026 (2005).

29. Song, H., Wang, Y. & Pant, K. Cross-stream diffusion under pressure-driven flow in microchannels with arbitrary aspect ratios: a phase diagram study using a three-dimensional analytical model. *Microfluid Nanofluid* **12,** 265–277 (2012).

30. Sadeghi, A. Analytical solutions for species transport in a T-sensor at low peclet numbers. *AIChE Journal* **62,** 4119–4130 (2016).

31. Craggs, T. D., Sustarsic, M., Plochowietz, A., Mosayebi, M., Kaju, H., Cuthbert, A., Hohlbein, J., Domicevica, L., Biggin, P. C., Doye, J. P. K. & Kapanidis, A. N. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res* **47,** 10788–10800 (2019).

32. Murphy, M. C., Rasnik, I., Cheng, W., Lohman, T. M. & Ha, T. Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. *Biophysical Journal* **86,** 2530–2537 (2004).

33. *CRC Handbook of Chemistry and Physics, 86th Edition* (ed Lide, D.) 2544 (CRC Press, 2005).

34. Fijen, C., Silva, A. M., Hochkoeppler, A. & Hohlbein, J. A single-molecule FRET sensor for monitoring DNA synthesis in real time. *Physical Chemistry Chemical Physics* **19,** 4222–4230 (2017).

35. Juette, M. F., Terry, D. S., Wasserman, M. R., Altman, R. B., Zhou, Z., Zhao, H. & Blanchard, S. C. Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat Methods* **13,** 341–344 (2016).

36. Dingfelder, F., Wunderlich, B., Benke, S., Zosel, F., Zijlstra, N., Nettels, D. & Schuler, B. Rapid Microfluidic Double-Jump Mixing Device for Single-Molecule Spectroscopy. *J. Am. Chem. Soc.* **139,** 6062–6065 (2017).

37. Freitag, C., Noble, C., Fritzsche, J., Persson, F., Reiter-Schad, M., Nilsson, A. N., Granéli, A., Ambjörnsson, T., Mir, K. U. & Tegenfeldt, J. O. Visualizing the entire DNA from a chromosome in a single frame. *Biomicrofluidics* **9,** 044114 (2015).

4

Supplementary material
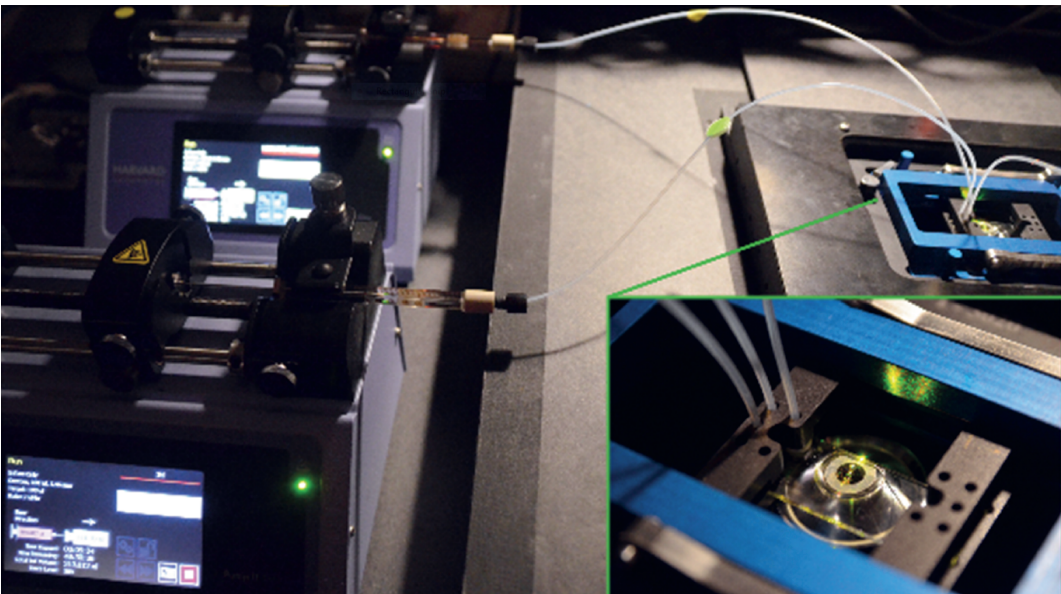
Figure S4.1: Overview of the fluidic setup on a home-built microscope.
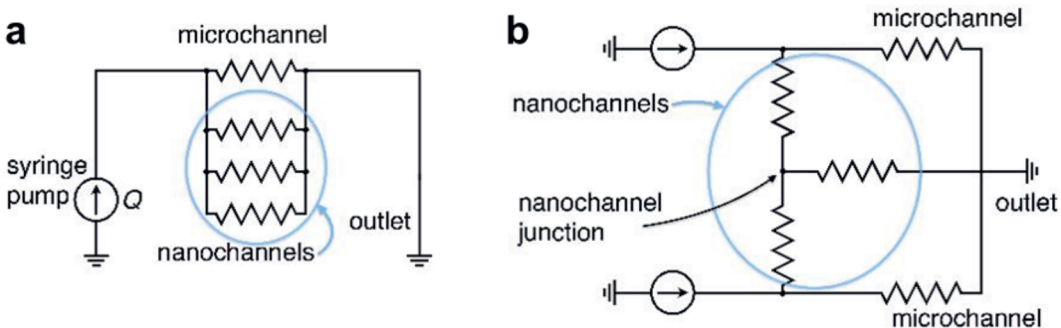
Figure S4.2: Equivalent circuit diagram of the nano-/microfluidic channel geometries. Syringe pumps act as current sources delivering constant flow rates Q and channels oppose flow by a hydraulic resistance R_{hyd}. (a) Parallel flow configuration of a bypassing microchannel and a nanochannel array. (b) Diagram of a micro-/nanofluidic mixing circuit. Here, reaction products are imaged at and downstream the nanochannel junction.
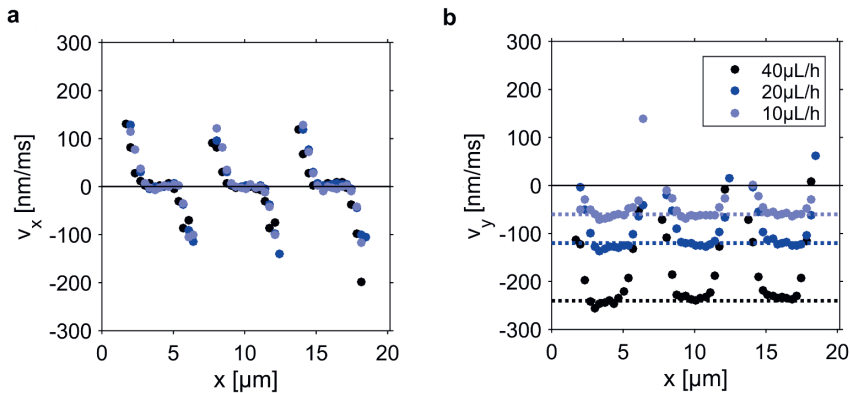
Figure S4.3: Experimental flow velocity profiles for different pump rates in three parallel nanochannels (40, 20, and 10 μL/h). (a) Velocity profiles along x: this direction is orthogonal to the applied flow thus the profiles are not influenced by the pump rate and are averaged around zero. The negative and positive velocities at the edges of each nanochannels arise from the border effect, in which a molecule very close to the nanochannel wall cannot diffuse further hence on average will diffuse back toward the central axis of the channel (b) Velocity profiles along y: Reference lines are drawn at $(-60, -120$ and -240 nm/ms. The flow speed registered inside the nanochannels scales linearly with the pump rate as expected for our system. All the profiles show a plateau for the velocity in the center of the channels (see also Fig. S4.4).
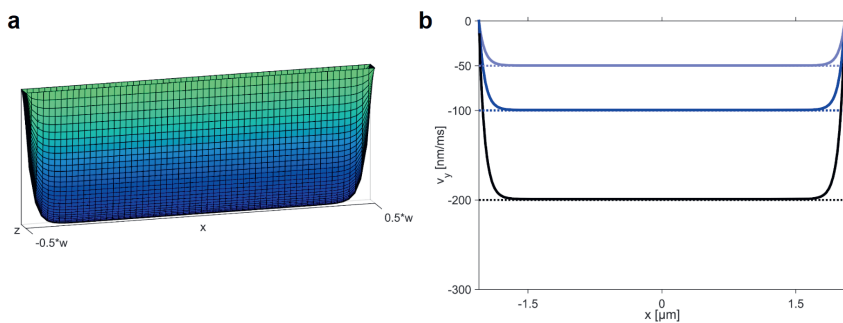
Figure S4.4: Theoretical 3D and 2D velocity profiles (see also Supporting Note 2). (a) Theoretical 3D velocity surface for Poiseuille flow in a nanochannel. The velocity surface was calculated using the dimensions of one parallel nanochannel. Along the height of the channels (z, 0.2 μm) the velocity profile is parabolic. Along most of the width (x, here 4.1 μm) the flow speed is constant (parallel plate-like scenario). (b) Theoretical 2D velocity profiles for Poiseuille flow in a nanochannel. Theoretical velocity profiles calculated for different pump rates: 10, 20, and 40 μL/h; the value of the plateau velocity (v_{max}) scales linearly with the pump rate.

4
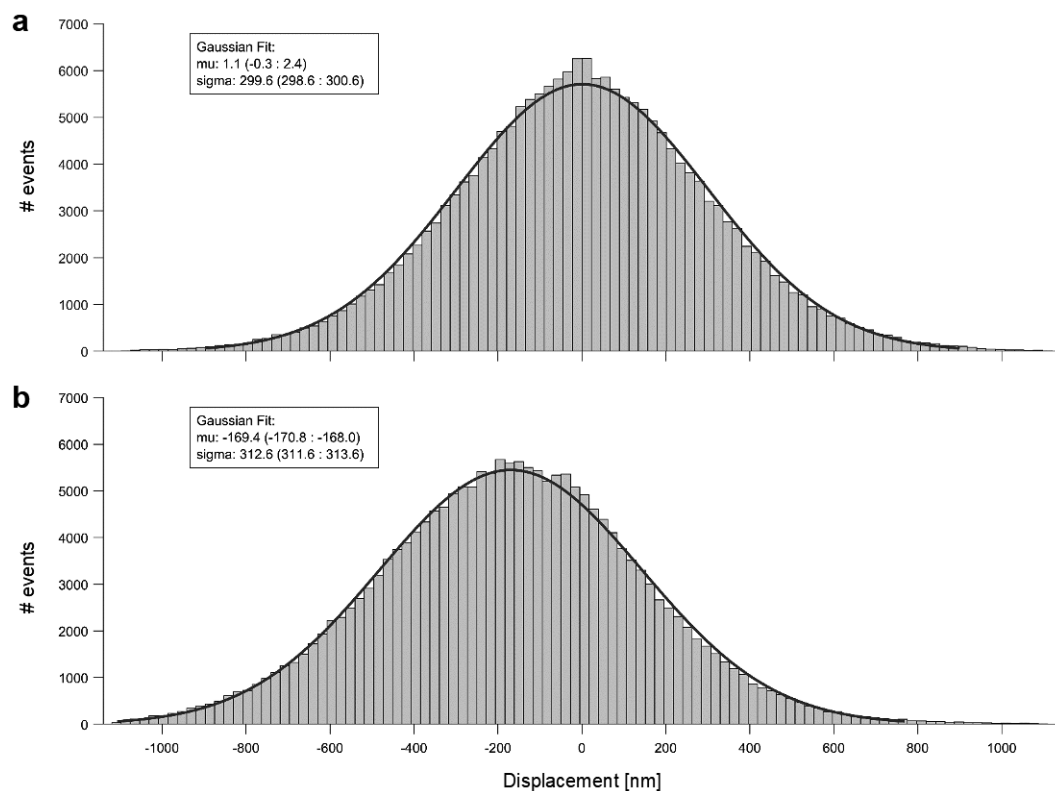
Figure S4.5: Example of frame-wide displacement distributions. Bending sensor in absence of KF in parallel nanochannels at pump rate of $20\,\mu\text{L/h}$. The displacement distribution is a normal distribution with mean proportional to the flow applied on that direction and variance proportional to the diffusivity of the tracked species (see also Material and Methods). (a) The distribution along x is centered on zero since there is no flow along this direction. Note that the 95 % interval of confidence for the mean (μ) contains the value 0. (b) The mean of the displacement distribution along y is shifted to $-169.4\,\text{nm}$; this, given the $1.5\,\text{ms}$ between the green and red laser excitations results in a mean velocity across the field of view of $112.9(9)\,\text{nm/ms}$ (95 % interval of confidence on the mean).
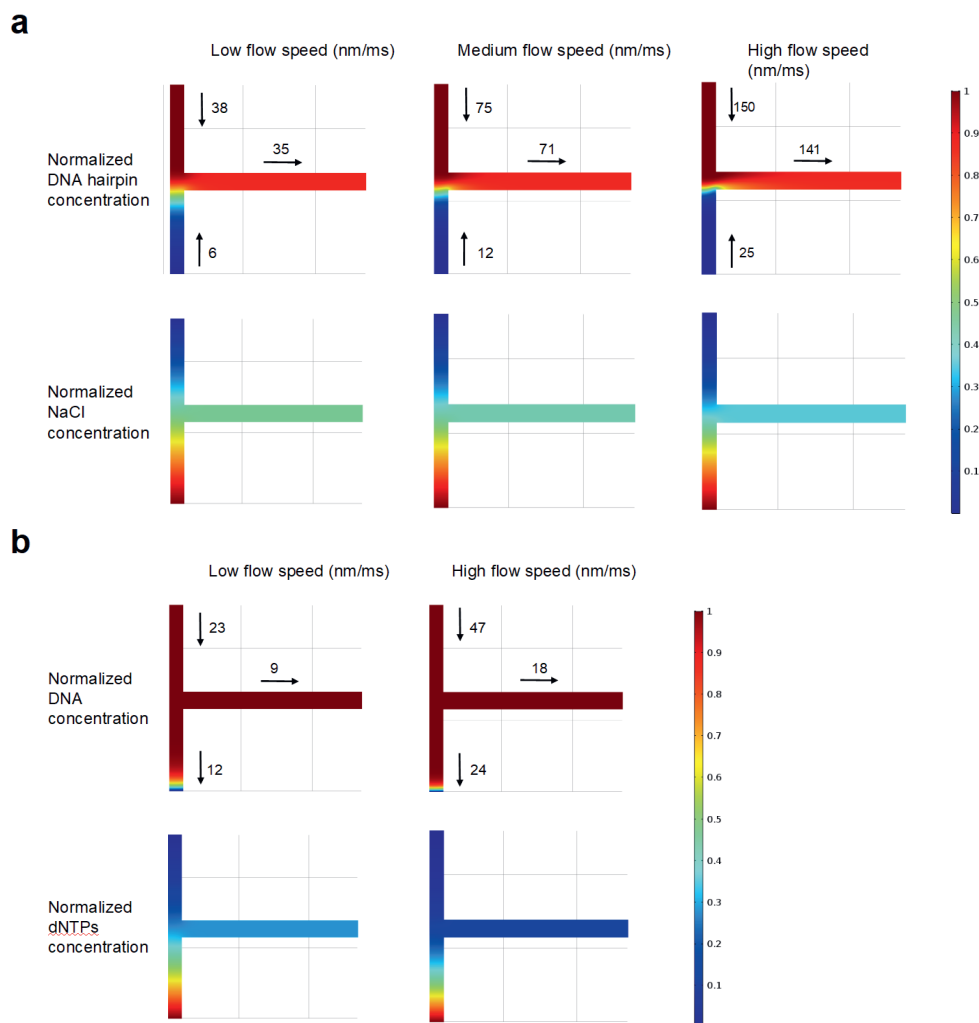
Figure S4.6: Finite element simulations (COMSOL, see also Supporting Note 3). A three-dimensional geometry of inlet channels of combined $52\,\mu\mathrm{m} \times 3.8\,\mu\mathrm{m} \times 0.2\,\mu\mathrm{m}$ and a downstream outlet/observation channel of $50\,\mu\mathrm{m} \times 4.7\,\mu\mathrm{m} \times 0.2\,\mu\mathrm{m}$ was used. Concentration profiles in the center plane at $z = 100\,\mathrm{nm}$ are shown. (a) Simulations of the DNA hairpin experiment in the mixing nanochannels. The simulated flow velocities for the three tested pump settings are reported. The experimentally obtained value for the diffusion coefficient of the hairpin was used $(35\,\mu\mathrm{m}^2/\mathrm{s})$, together with an approximate value for the diffusion coefficient of salt $(2000\,\mu\mathrm{m}^2/\mathrm{s})$. (b) Simulation of the DNA polimerization experiment. The simulated flow velocities for the two tested flow rates are reported. The experimentally obtained value for the diffusion coefficient of the DNA was used $(25\,\mu\mathrm{m}^2/\mathrm{s})$, together with an approximate value for the diffusion coefficient of the dNTPs $(400\,\mu\mathrm{m}^2/\mathrm{s})$.

4
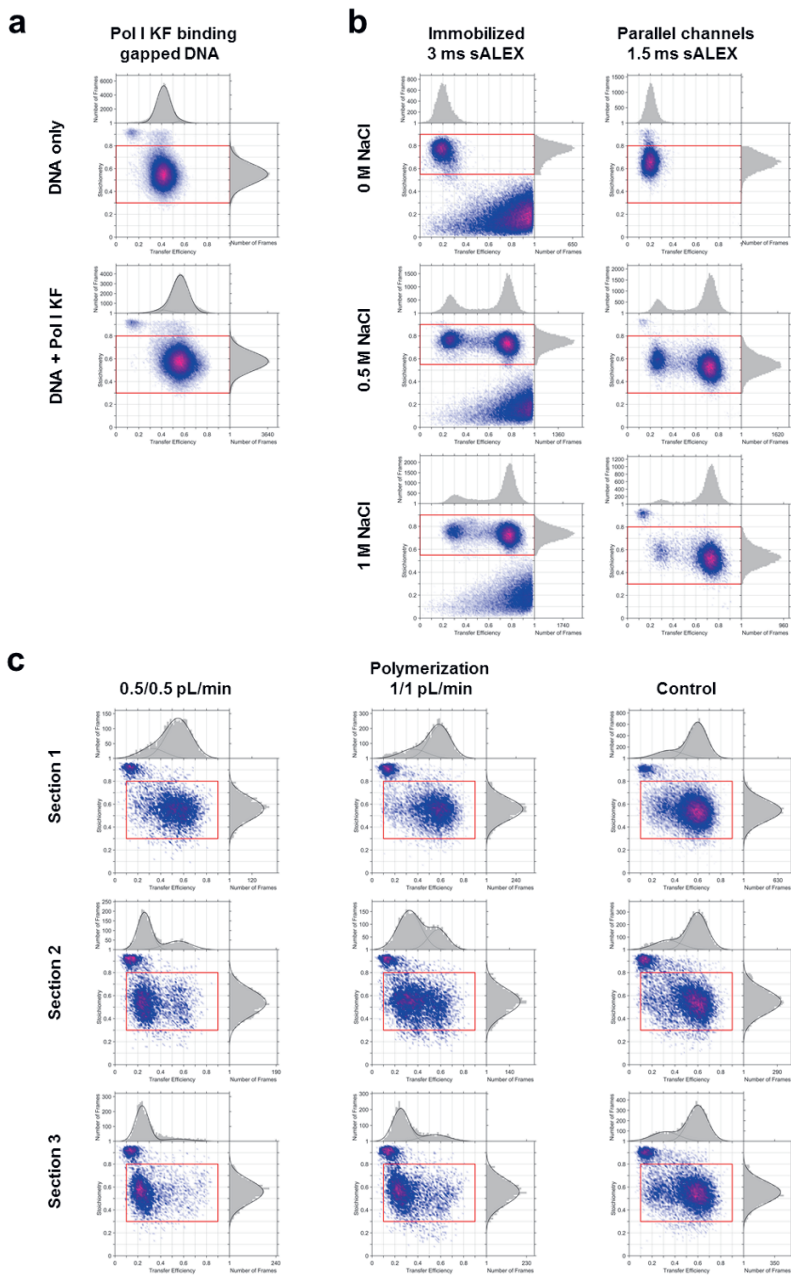
Figure S4.7: E^*/S histograms of data shown in the main figures. (a) Fig. 4.2b. (b) Fig. 4.3a. (c) Fig. 4.4b-d. Only data inside the red box is projected on the axes. FRET efficiency E^* is on the x-axis, stoichiometry S is on the y-axis.

Supporting Note 1: Achieving parallel control

Flow was driven by syringe pumps (Pump 11 Pico Plus Elite; Harvard Apparatus, USA). In Parallel Flow Control (PFC), low flow rates are achieved by dividing the syringe flow rate Q into the microchannel and nanochannels according to the channels' hydraulic resistance, which is calculated as[1,2]

$$R_{hyd} \approx \frac{12\eta L}{1 - 0.63\frac{h}{w}} \cdot \frac{1}{h^3 w} \quad \text{for} \quad h < w \tag{S4.1}$$

Here, η is the dynamic viscosity and h, w and L are the height, width and length of a nanochannel, respectively (A pressure Δp across a channel drops according to the Hagen-Poiseuille Law $Q = \Delta p / R_{hyd}$, and for parallel resistances $R_{array} = 1/\sum R_{hyd,i}^{-1}$ applies.). Due to the cubic dependence h^3, the flow rate in each nanochannel is strongly reduced compared to the microchannel. For high flow rate of 20 pL/min inside one nanochannel of the parallel design, the pump generates a pressure of about 15 kPa while operating at 85 µL/h. This large ratio of flows allows the dead volume in the 1 mm wide and up to 2 cm long feeding microchannels (0.1 µL) to be replaced in seconds. For the mixing device, two syringe pumps deliver flow to both feeding nanochannel inlets (each 20 µm long, 3.8 µm wide and 200 nm high). A single nanochannel (100 µm long, 4.7 µm wide and 200 nm high) is positioned downstream of the junction. Two bypassing microchannels lead to an overall reduction factor of 140.000 of the nanochannel flow compared to the combined microchannel flows. We note that a geometry similar to the one of our parallel devices, where narrow channels are flanked with wide channels has been previously used in literature to achieve a different goal: maintain hydrostatic equilibrium between reservoirs connected by the channels[3,4]. Any small difference in hydrostatic pressure between the two reservoirs would be rapidly equilibrated through the wide channels; at the same time, thanks to the high reduction factor of this geometry the flow induced in the narrow channels would stay negligible, allowing to perform experiment where this condition is required.

4

Supporting Note 2: Poiseuille flow in rectangular nanochannels

The velocity for Poiseuille flow in our nanochannel geometry can be calculated using (reviewed in refererence[1])

$$v_y(x, z) = \frac{4h^2 \Delta p}{\pi^3 \eta L} \sum_{n.odd}^{\infty} \frac{1}{n^3} \left[1 - \frac{cosh\left(n\pi \frac{x}{h}\right)}{cosh\left(n\pi \frac{x}{2h}\right)} \right] sin\left(n\pi \frac{z}{h}\right) \qquad \text{(S4.2)}$$

Here, Δp is the pressure difference between the two ends of the channels, η is the dynamic viscosity of the fluid, and h, w and L are the height, width and length of a nanochannel, respectively. The resulting flow surface (Fig. S4.4a) shows a central velocity plateau along most of the width (x) of the nanochannel whereas along the height (z) a parabolic profile is present. While taking a movie of molecules flowing through the nanochannels we are acquiring the projection of the position of the particles in the x,y-plane; still, particle speed will also be a function of height. By calculating the displacements on the plane, we are neglecting the position of the molecule along z and the velocity profiles obtained in this way will therefore represent the velocity averaged along z:

$$v_y(x) = \frac{1}{h} \int_0^h v_y(x, z) dz \qquad \text{(S4.3)}$$

Using this equation, it is possible to calculate the expected velocity profile as seen in our experiments (see Fig. S4.4b).

Supporting Note 3: Finite element simulations

Three-dimensional finite element modeling (COMSOL Multiphysics 5.2a) was employed to determine steady-state concentration profiles of the different species. The geometry consists a T-shaped channel with a top segment of $52\,\mu m \times 3.8\,\mu m \times 0.2\,\mu m$ connected to downstream channel of segment $50\,\mu m \times 4.7\,\mu m \times 0.2\,\mu m$ μm as shown in Figure S4.6. A laminar flow profile was determined by solving the Stokes equations for an incompressible fluid:

$$\vec{\nabla} p = \eta \nabla^2 \vec{v} \quad , \quad \vec{\nabla} \vec{v} = 0 \tag{S4.4}$$

As boundary conditions, inflow velocities as indicated in Figure S4.6, an outlet exit pressure $0\,Pa$, and no-slip conditions for all other boundaries were chosen. The calculated flow profile $\vec{v}$ was used to evaluate drift diffusion equations

$$\vec{v}\,\vec{\nabla} c_i = D_i \nabla^2 c_i \tag{S4.5}$$

where c_i and D_i are the concentration and diffusivities of DNA hairpin, salt, DNA and dNTP as defined in the caption of Figure S4.6. We chose boundary conditions of c(DNA hairpin) = 1 mM and c(salt) = 0 mM at the right/upper inlet and of c(DNA hairpin) = 0 mM and c(salt) = 1 mM for the left/lower inlet. For mixing in the polymerization experiment we chose c(DNA) = 1 mM and c(dNTP) = 0 mM at the right/upper inlet and of c(DNA) = 0 mM and c(dNTP) = 1 mM at the left/lower inlet, respectively.

References

1. Bruus, H. Acoustofluidics 1: Governing equations in microfluidics. *Lab on a Chip* **11,** 3742 (2011).
2. Mathwig, K. & Lemay, S. G. Mass transport in electrochemical nanogap sensors. *Electrochimica Acta* **112,** 943–949 (2013).
3. Pagliara, S., Dettmer, S. L., Misiunas, K., Lea, L., Tan, Y. & Keyser, U. F. Diffusion coefficients and particle transport in synthetic membrane channels. *The European Physical Journal Special Topics* **223,** 3145–3163 (2014).
4. Misiunas, K., Pagliara, S., Lauga, E., Lister, J. R. & Keyser, U. F. Nondecaying Hydrodynamic Interactions along Narrow Channels. *Phys. Rev. Lett.* **115,** 038301 (2015).

Chapter 5

Characterization of ARF interactions using SMFD in nanofluidic devices

Mattia Fontana, Šarūnė Ivanovaitė, Simon Lindhoud, Willy van den Berg, Dolf Weijers and Johannes Hohlbein

5

Abstract

Single-molecule fluorescence detection (SMFD) schemes offer powerful ways to study biomolecules and their interactions. In particular, camera-based SMFD of freely diffusing molecules minimizes potential surface-induced artifacts while enabling high-throughput measurements. In this chapter, we utilize a nanofluidic device to perform SMFD experiments on a single-molecule TIRF microscope. We studied the interaction of unlabelled, transcription factors (ARF-DBDs) with a doubly-labelled DNA oligo containing their response element. We show that the binding of ARF-DBD to the DNA can be detected by both changes in the FRET efficiency between an donor and an acceptor fluorophore on the DNA and the diffusion coefficient of the DNA. In addition, our data on fluorescently labelled ARF-DBDs showed that, at nanomolar concentrations, they are present exclusively as monomers. Lastly, the ability of a DNA sporting two binding sites for ARF monomers to stabilize the protein dimer was directly observed. In general, the additional information coming from SMFD of freely diffusing molecules proved to be instrumental in shading more light on the interactions between ARF-DBDs monomers and between ARF-DBDs and their response element.

Introduction

Single-molecule techniques are playing an increasingly important role in the investigation of properties and dynamics of proteins and biomolecular complexes thanks to their ability to resolve inter- and intramolecular heterogeneity. In particular, fluorescence-based techniques became widely employed thanks to their ease of use[1]. Traditionally, single-molecule fluorescence detection (SMFD) can be carried out either on a confocal microscope, which use one or more avalanche photodiodes as point detectors, or on a wide-field microscope used in total internal reflection (TIR) mode, which use emCCD or sCMOS cameras to monitor hundreds of molecules in parallel. SMFD of free-diffusing molecules on a confocal microscope allows for high time resolution (typically μs) at the expense of throughput and short observation times while SMFD of surface-immobilized molecules on a TIRF microscope displays a somehow complementary behaviour with lower time resolution (typically ms[2]) compensated by high throughput and long observation times. During the past decade, different frameworks were proposed to overcome the limitations imposed by these traditional implementations of SMFD. For confocal microscopy, the main focus has been prolonging the observation times[3–6] while in TIRF-based applications, the aim was to eliminate the need of sample immobilization[7–9]. Performing SMFD experiments on a TIRF microscope without immobilization allows to minimize surface-induced artifacts whilst maintaining the high throughput inherent to camera-based detection schemes.

Here, we utilize a glass-made fluidic device[8] on a TIRF microscope to study some members of the transcription factor family of Auxin Response Factors (ARFs) and their interaction with their response element via SMFD. We were able to detect the interaction between ARF DNA Binding Domain (ARF-DBD) with a doubly labelled DNA construct containing the response element by monitoring the change in its FRET signature and the diffusion coefficient upon DNA binding. We then studied fluorescently labelled ARF-DBDs to show that they are present in their monomeric form at the concentration used in our experiments ($\approx$1 nM), which contrast with the dimeric form portrayed in multiple crystal structures[10,11]. Finally, we show that the presence of the DNA construct is likely to support ARF dimerization.

Materials and methods

Device fabrication, surface passivation and cleaning

The fluidic devices were fabricated by Micronit (Micronit Microtechnologies B.V., The Netherlands) following the same methodology and design described previously[8]. To prevent non-specific adsorption of the analytes, we opted for passivating cleaned nanochannels using polyethylene glycol (mPEG, MW = 5000 Da, Laysan Bio Inc., USA) as follows. First, the channels were washed with acetone and then incubated for 5 minutes with a 50:1 acetone:Vectabond solution; then, channels were washed with acetone (1X) followed up by MilliQ (3X) and incubated with a solution of mPEG in MOPS buffer for 3 hours. The pegylated channels were then washed with PBS buffer and stored in a wet chamber at 4 °C. After the experiments, the devices were washed by flowing and incubating a solution of 50:1 MilliQ:Hellmanex®III (Hellma, Germany). After rinsing with MilliQ, the devices were burned in a furnace at 500 °C. The detailed protocols for passivation and cleaning can be found in the Supplementary Materials.

Protein expression and purification

The DBDs of MpARF1, MpARF2, AtARF1 and AtARF5 were amplified and cloned in a modified pTWIN1 vector. The expression and purification of recombinant ARF-DBDs were performed as previously described[10,11].

Protein labelling

The labelling buffer is composed of 20 parts of solution 1 (274 mM NaCl, 5.4 mM KCl, 10 mM phosphate) and one part of solution 2 (0.2 M NaHCO$_3$, pH adjusted to 9 using NaOH); the pH of the final labelling buffer was adjusted to 8.3 using solution 1 or 2. Buffer exchange was carried out by two consecutive runs on Zeba™Spin Desalting Columns 7MWCO (Thermo Fisher Scientific, USA) that was equilibrated in labelling buffer.

For labelling ARF5-DBD, 60 µL of a 3.5 mg/mL ARF5-DBD solution was supplemented with 12 µL of 800 µM DNA duplex and 1 mM TCEP (final protein concentration: 65 µM, final DNA concentration: 133 µM). 5 µL of 10 mM label was added. The labelling reaction was carried out overnight at 4 °C; then, the labeled

Table 5.1: AV simulations: Geometrical parameters of the dyes taken from reference [14].

Dye	Linker	Linker length [Å]	Linker width [Å]	R_1 [Å]	R_2 [Å]	R_3 [Å]
Cy3B	C6-NHS	14.2	4.5	8.2	3.3	2.2
ATTO647N	C6-NHS	17.8	4.5	7.4	4.8	2.6

protein was separated from free dye on a PD-10 desalting column (Cytiva, USA) that was equilibrated in PBS supplemented with 5 mM DTT.

For ARF1-DBD, the labelling was performed as described for ARF5-DBD but in absence of the DNA duplex.

Accessible Volume simulations

The web server 3D-DART[12] was used to first model a straight, standard B-DNA structure for the oligo used in the single-molecule FRET (smFRET) experiments. A second, bent oligo was modelled using the geometrical information extracted from the short DNA present in the crystal structure of MpARF2[11]. The bent oligo was then rigid-docked in place of the short oligo that was present in the crystal structure. Both oligos were then used as starting points to model the accessible volumes of the FRET pair using the FRET-restrained positioning and screening software[13]. The accessible volume is defined by the region of space that each dye can explore given its own geometry, the geometry of the linker and the attachment point on the biomolecule (see table 5.1).

Single-molecule detection and tracking

Single-molecule detection was performed using a TIRF microscope (previously described[8]) equipped with a fiber-coupled laser engine (Omicron, Germany). The triggering of the lasers and the camera as well as the laser intensities were controlled by a home-written LabVIEW program. The power of the lasers was set to 100 % which results in a measured laser power at the fiber output of approximately 280 mW for the green excitation (561 nm) and 140 mW for the red excitation (638 nm). The illumination was based on the stroboscopic alternating-laser excitation (sALEX[2]) scheme with laser pulses of 1.5 ms in a frame time of 10 ms. Green and red pulses of

two subsequent frames were placed back-to-back to facilitate tracking between these coupled frames. Particles were localized using a modified version of GaussStorm[15] and tracked using Maria Kilfoil's MATLAB porting of a tracking algorithm developed in IDL[16]. For samples containing only one fluorophore (e.g. singly-labelled proteins), a scheme based on stroboscopic back-to-back illumination was applied to the single excitation laser. Per condition, we acquired 4 movies of 4000 frames for a total acquisition time of 160 s.

Particle displacements analysis

The particle displacements between two back-to-back illuminated frames were obtained using 2D Gaussian fitting of the imaged point spread functions as described previously[8]. From the variance of the 1D Gaussian distributions of all one-dimensional displacements in either x, or y-direction (σ^2, also called mean-square-displacement *MSD*), we calculated the mean particle diffusion coefficient D using

$$D = \frac{\sigma^2}{2t}, \tag{5.1}$$

here, t is the time between the mean time of two back-to-back illuminations (1.5 ms). In case the displacements are calculated along the direction of the flow, the mean flow speed ($\langle v \rangle$) is determined by the mean (μ) of their distribution

$$\langle v \rangle = \frac{\mu}{t}. \tag{5.2}$$

We note that equations 5.1 and 5.2 do not take into consideration the uncertainty associated with the localization. While equation 5.2 holds true, calculating the diffusion coefficient from 5.1 requires knowledge on the variance that was generated by the localization uncertainty (of both localizations) and its subtraction from the total variance to isolate the component that was generated by the Brownian motion. Consequently, diffusion coefficients calculated without accounting for the localization uncertainty will overestimate the real value.

Calculation of theoretical diffusion coefficients

The predicted diffusion coefficients presented in this chapter were calculated using HYDROPRO[17]. This software uses the crystal structure of a macromolecule to estimate its diffusion coefficient. The calculations were performed loading the pdb

file of the protein or complex of interest, setting the temperature to 20 °C and the viscosity of the solution to 1 cP; the calculations were performed in *Mode 1* (shell-mode from atomic-level) using a radius of the atomic elements of 2.84 Å.

Results

SMFD in parallel nanochannels

Detecting non-immobilised, nanometer-sized molecules in fluidic devices using camera-based microscopy is challenging as the movement of the molecule due to diffusion and advection within the time of a single camera frame results in the spreading of emitted photons over many pixels. Given the already limited photon budget of single emitters, the movement further decreases the signal-to-noise ratio (SNR).

To mitigate the motion blur caused by advection, the fluidic design should provide sufficiently slow flow speeds while still allowing for fast exchange of sample solutions ensuring low dead times. A convenient way of implementing such characteristics is to use parallel flow control (PFC[18,19]). In our implementation of PFC[8], the flow coming from a syringe pump is divided between the imaging channels (i.e. nanochannels) and a bypassing microchannel (Fig. 5.1a-b) such that most of the liquid passes through the microchannel due to its lower hydraulic resistance thereby reducing the fluid velocity inside the nanochannels by several orders of magnitude (to values in the tens or hundreds of nanometres per millisecond). This arrangement further assures that the dead volume in the tubing and feeding channel is replaced in minutes (i.e. reduced dead time). The nanochannels have an height of 200 nm thereby limiting the movement of the fluorescent molecule to the evanescent field of the TIRF microscope and enabling reliable single-particle tracking (Fig. 5.1b). Another source of spreading of the emitted photons over additional pixels is the movement of the fluorescent molecule due to diffusion; the effect can be strongly reduced by applying a stroboscopic illumination scheme in which the molecule is illuminated only for a fraction of the acquisition time of the frame[2]. Moreover, the illumination of two neighbouring frames can be arranged in a back-to-back configuration (Fig. 5.1c) to further minimize the displacements between those frames thus allowing for higher concentrations of molecule in the channels before successful tracking is compromised by overlapping pathways.
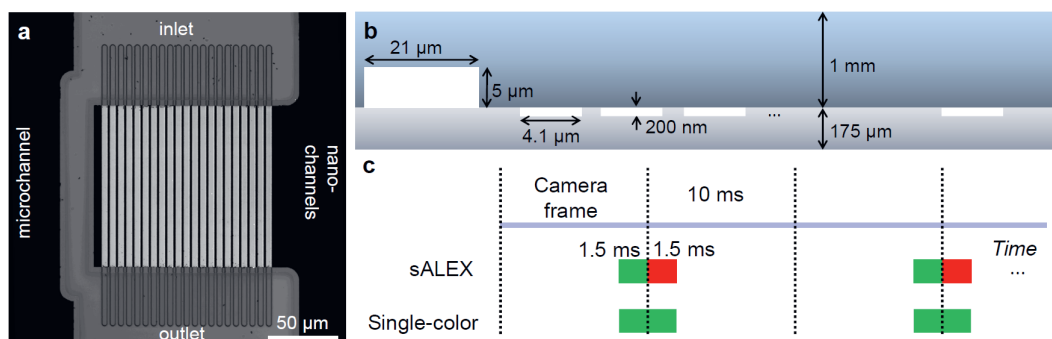
Figure 5.1: Design of the fluidic device and excitation (Adapted from reference [8]). (a) High-resolution confocal reflection scan of the parallel nanochannels device; both the array of imaging nanochannels (l × w × h: 120 by 4.1 by 0.2 μm) and the bypassing microchannel (l x w × h: 120 by 21 by 5 μm) are visible. (b) Schematic cross-section of the parallel channel array, showing the dimensions of the microchannel and the nanochannels. The microchannel has a larger cross-section thereby causing lower resistance to the fluid. As a result, most of the sample liquid passes through the microchannel minimising both the speed of the analyte inside the nanochannels and the dead time of the fluidic system. (c) Schematic representation of the laser excitation schemes. The concept of stroboscopic illumination is used to reduce the motion blur of single-particle localization while the excitation of two successive frames are placed back-to-back to facilitate tracking and allowing for higher concentrations of fluorescent molecules to be analysed.

The interaction between ARF and labelled DNA can be visualized using smFRET in nanochannels

We tested the feasibility of visualizing the interaction between ARF and its DNA response element in nanofluidic devices by flowing a DNA construct through the channels in presence or absence of ARF transcription factors. We utilized a doubly-labelled DNA construct similar to the one we previously used as immobilized entity in single-molecule FRET experiments to determine the binding affinity of the DNA Binding Domains (DBDs) of *Marchantia polymorpha* ARF 1 and 2[11]. The construct contains two high affinity AuxREs (TGTCGG) in an inverted topology with a spacer of 7 bp (IR7) and is labelled with a FRET pair featuring Cy3B as the donor and ATTO647N as the acceptor fluorophore (Fig. 5.2a). The accessible volumes of the dyes, as well as the average distances between each other, were calculated using the FPS software[13] (see also Materials and Methods). Upon binding to DNA, ARF is expected to bend the DNA and sterically confine the movement of the dyes (Fig. 5.2b); thus, ARF binding causes a decrease in the FRET efficiency of the DNA

construct[11]. Measuring the DNA construct alone (2.5 nM in PBS buffer containing oxygen scavenger system[20,21]: 1 mM Trolox, 1 % gloxy and 1 % glucose) in the fluidic devices let to a single population centered at $E^* = 0.39$ (Fig. 5.2c); a second sample containing saturating concentration (256 nM) of MpARF2-DBD showed the characteristic shift toward a lower FRET efficiency ($E^* = 0.31$, Fig 5.2c, bottom). The fraction of bound DNA in the sample containing 256 nM of MpARF2-DBD was 86 %, a value close to the ones obtained in experiments with immobilized DNA samples 94(8) % (see Chapter 2, Fig. 2.5b).

We asked to which extend the use of parallel nanochannels allows performing titration experiments without the supervision of an operator. To this end, we opted for sequentially loading the samples containing DNA and increasing concentration of ARF inside the syringe or, alternatively, inside a long connecting tubing. To prevent mixing of neighbouring samples, we separated them by an air bubble. This arrangement abolishes dispersing effects created by convection (i.e. Taylor dispersion) while the air bubbles provide a visual cue for the position of the sample. We tested the general feasibility by loading two samples separated by an air bubble into a syringe before feeding the nanochannels. The first sample contained the DNA construct in PBS 1X buffer (2.5 nM, Fig. 5.2d,e top) while the second contained both DNA and 256 nM of either MpARF2-DBD (Fig. 5.2d, bottom) or MpARF1-DBD (Fig. 5.2e, bottom). The ARF containing samples showed the characteristic shift towards lower FRET efficiencies when compared to the one without ARF; however, the noise present on the E^*S histogram makes the evaluation of the DNA bound fraction inaccurate. The time duration of an entire titration combined with the presence of air bubbles between the samples makes the use of enzymatic oxygen scavenging system (OSS) problematic; given the presence of molecular oxygen in the sample, we explain the lower quality of the data to be caused by a relevant portion of the counts in the E^*S histogram originating from doubly-labelled DNA molecules in which either the donor or the acceptor fluorophore photobleached during the acquisition of a single frame.

Taken together, these results show that the characteristic shift in FRET efficiency seen in experiments with the surface-immobilised DNA construct can be also seen using freely diffusing DNA. This finding further confirms that the shift in E^* is caused by specific ARF to DNA binding rather than surface-induced artefacts (e.g. protein absorption on the surface).
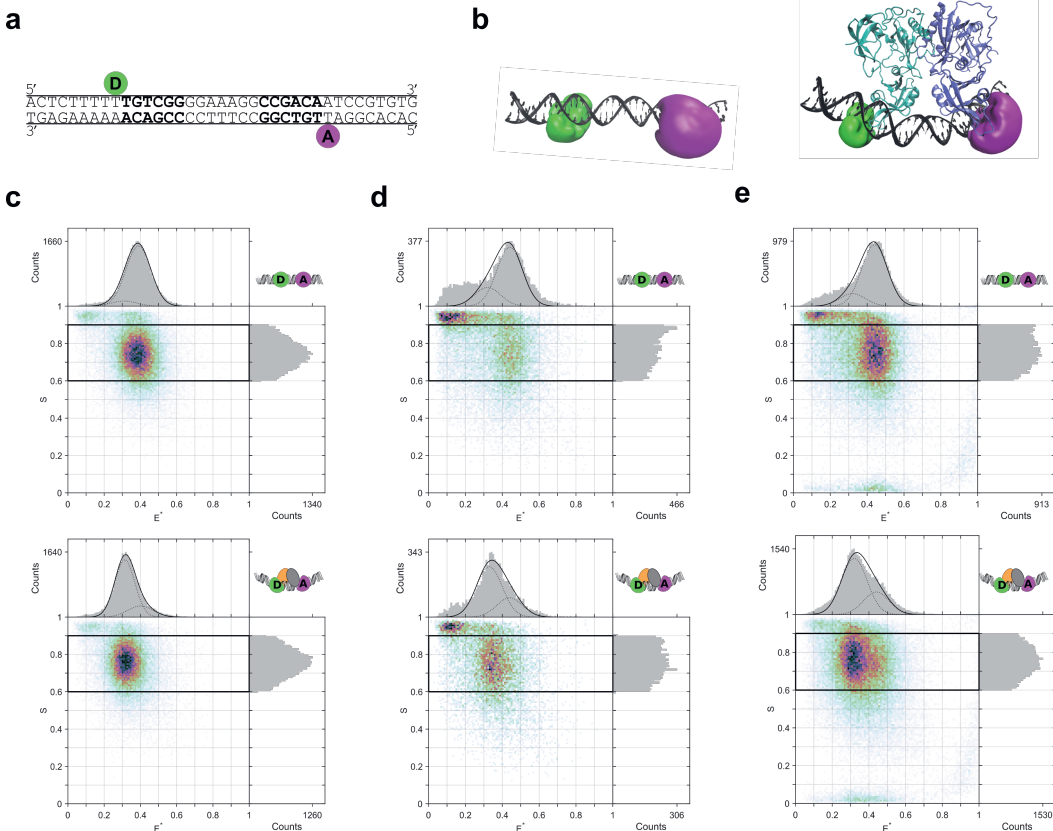
Figure 5.2: Visualizing ARF-DNA interactions in nanofluidic devices. (a) Schematic representation of the doubly-labelled DNA construct. The two AuxREs are highlighted in bold. Attachment points for the donor (D, Cy3B) and acceptor (A, ATTO647N) fluorophores are shown in green and magenta respectively. (b) Accessible Volume (AV) simulation for the dyes for free (left) and ARF-bound (right) DNA construct. The accessible volume is the region of space that can be visited by the dye thanks to the flexible linker that attaches it to the DNA. When the protein is present, the mean distance between the dyes increases resulting in the decrease of the FRET efficiency E^*. (c) E^*S histograms of a sample containing the DNA construct (top) and a sample additionally containing a saturating concentration of MpARF2-DBD (256 nM, bottom). The characteristic change in FRET efficiency is noticeable and the fraction of bound DNA can be calculated from the relative areas of the two Gaussians centered at the E^* of the free and bound DNA. (d) The same samples of (c) were loaded consecutively on the syringe while being separated by a air bubble, and without the standard enzymatic oxygen scavenger system being present in the buffer. The expected shift in FRET efficiency upon ARF binding can be seen, but the reduced data quality due to higher photobleaching does not allow to reliably identify the relative amount of bound and free DNA. (e) Same experiment as in (d) but with MpARF1-DBD instead of MpARF2-DBD.
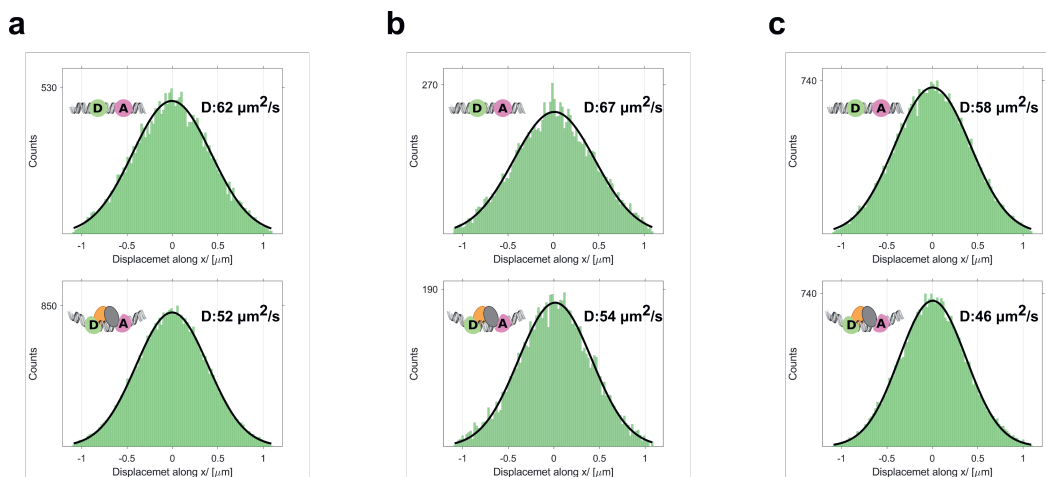
Changes in the diffusion coefficient of labelled DNA can be used to monitor ARF-DNA interaction

In the previous section we described experiments in which doubly-labelled single DNA constructs were imaged and tracked to determined their FRET efficiency (E^*) in absence or in presence of ARF; the resulting change in E^* was then used as readout for ARF binding. In this section, we will analyse the change in diffusion coefficient (D) of the DNA constructs to obtain an independent second readout for binding.

During two back-to-back illuminations, the DNA construct in the nanochannels will move due to flow and diffusion. In the direction perpendicular to the flow, the movement is due exclusively to Brownian motion and the distribution of the displacements can be used to calculate the diffusion coefficient ($D = \sigma^2/2t$, see also Materials and methods). The diffusion coefficient of the DNA construct in 1x PBS buffer (without OSS) was found to be 62 and 67 μm^2/s in two independent experiments (Fig. 5.3a,b top) and thereby close to the theoretical one of 69.8 μm^2/s (see Materials and Methods). Upon addition of either MpARF1-DBD or MpARF2-DBD, the diffusion coefficient of the DNA construct decreases to 52 μm^2/s and 54 μm^2/s respectively (Fig. 5.3a,b bottom); these values are again consistent with the theoretical value for the diffusion coefficient of the complex formed by a dimer of MpARF2-DBD bound to the DNA construct which is 51.0 μm^2/s. The analysis of the diffusion coefficients in presence of an OSS shows decreased diffusion coefficients for both free DNA and bound DNA. This effect is likely to be caused by (I) the reduced photobleaching which in turn reduces the average localization error and (II) the increased viscosity of the sample when OSS are present in the solution. (e.g. 1 %glucose = +2 %viscosity[22]). Similar to the experiments without the OSS, addition of ARF decreased the diffusion coefficient of the DNA by around 12 μm^2/s. We note that even though the change in viscosity upon OSS addition has an effect on the measured diffusion coefficients, the same does not hold for the addition of ARF as at the tested ARF concentrations the effect is too small to be detectable (see Supporting Note 1).

ARF-DBDs at nanomolar concentration are monomeric

A common feature of ARF-DBDs as portrayed in crystal structures is that they form dimers caused by interactions between their dimerization domains (DD). Unfortu-

Figure 5.3: Displacement analysis of the labelled DNA construct in absence or presence of MpARF1-DBD and MpARF2-DBD. (a) The diffusion coefficient of the DNA construct decreases by $10\,\mu m^2/s$ when 256 nM of MpARF1-DBD is added to the solution, consistent with the formation of a complex of the DNA construct with a protein dimer. (b) Same experiment as (a) but using MpARF2-DBD instead of MpARF1-DBD; the starting and final diffusion coefficients are in good agreement with the results in (a). (c) Same experiment as in (b) but with the addition of an enzymatic oxygen scavenger system (Glucose Oxidase/Catalase, glucose and Trolox); the decrease in diffusion coefficient upon MpARF2-DBD addition is consistent with (a) and (b) but the absolute values are lower due to the decrease in localization error and an increase of the viscosity of the solution.

nately, the structures do not provide information on the dimerization equilibrium in solution. To the best of our knowledge, the only data available (acquired using SAXS, Small Angle X-ray Scattering) reported $\approx 60\,\%$ dimer when $78\,\mu M$ AtARF1-DBD was added in solution[10]. This places a dissociation constant of the dimer to be $\sim 10\,\mu M$. Given the scarcity of information about ARF dimerization in solution we decided to perform experiments in parallel nanochannels. We imaged stochastically labelled AtARF1-DBD and AtARF5-DBD ($\approx 1\,nM$) whilst applying the back-to-back stroboscopic illumination scheme[8]; the proteins were tracked and the distributions of displacements were determined (Fig. 5.4).

Despite the passivisation of the glass surfaces with PEG, a small but noticeable fraction ($< 20\,\%$) of single-molecule localization came from ARFs that were adsorbed on the surface; for this reason a two Gaussian fit was employed, in which one Gaussian fits the displacements of the free flowing fraction and the second one

the apparent displacements of the adsorbed population. Fitting the distribution of displacements perpendicularly to the direction of flow (along x, Fig. 5.4,top) resulted in two Gaussian distributions centred at 0. The wider Gaussian belongs to the free-flowing molecules and its width is proportional to the diffusion coefficient of the free-flowing species while the narrow one belongs to the adsorbed population and its width is equal to double the localization error for an immobile particle (double since it comes from the sum of the errors of the two localizations used to calculate the displacement, see Materials and methods). An apparent diffusion coefficient for the adsorbed population was calculated and subtracted from the one of the free-moving particles to account for the localization error; we note that although moving particles are expected to have a higher localization error due to motion blur, we used stroboscopic illumination which minimizes this contribution. We obtained diffusion coefficients of $78\,\mu m^2/s$ for AtARF1-DBD labelled with Cy5, $65\,\mu m^2/s$ for AtARF5-DBD labelled with ATTO647N and $69\,\mu m^2/s$ for AtARF5-DBD labelled with Cy3B. These values can be compared with the theoretical ones for monomeric and dimeric AtARF5-DBD (extracted from the crystal structures PDBID:4LDU, see Materials and methods) of $73\,\mu m^2/s$ and $55\,\mu m^2/s$ respectively (given the high similarity in structure and size, the values obtained for AtARF1-DBD using the crystal structure PDBID:4lDV showed no difference when rounded to two significant digits). The comparison of the simulated and experimental values for the diffusion coefficient suggests that, at the concentration tested ($\approx 1\,nM$), both AtARF1-DBD and AtARF5-DBD are present as monomers in solutions. Fitting the distribution of displacements along the direction of the flow (along y, Fig. 5.4,bottom) proved that the fraction with a low apparent diffusion coefficient corresponded to absorbed molecules and not proteins aggregates. Here the narrow distribution is still centred around zero while the wide population of the freely-moving particles is centred around a value which is proportional to the mean flow speed in the field of view ($\langle v \rangle$, see Materials and methods). In general, the lack of a narrower population centred around the mean flow speed proves that protein aggregation, if present, is negligible. We note that the diffusion coefficient of the moving particle along y is affected by additional broadening due to the inhomogeneous velocity field inside the channels and cannot be used to determine the diffusion coefficient directly.
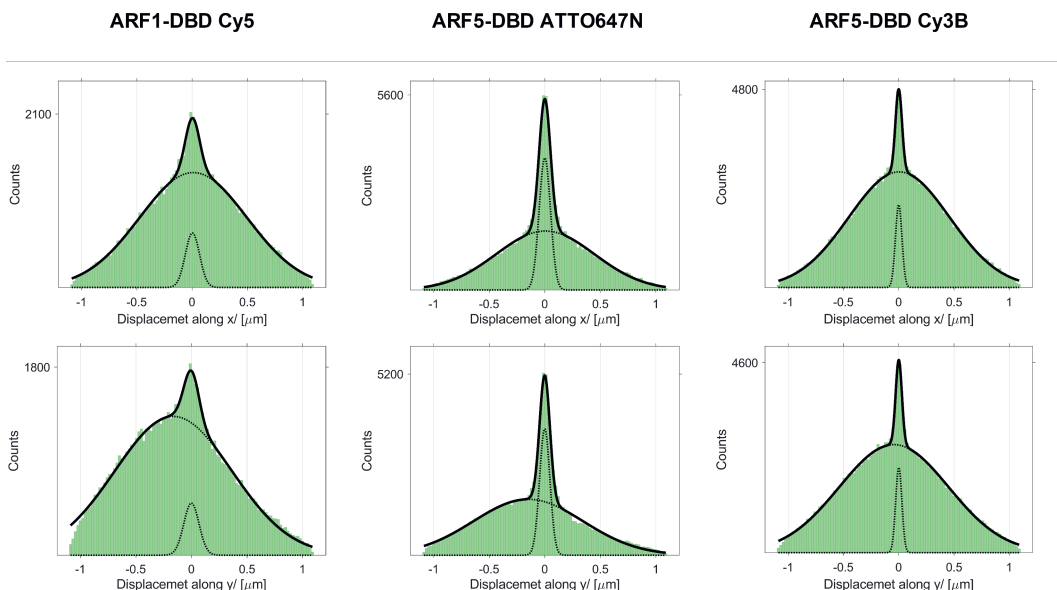
Figure 5.4: Displacements analysis for AtARF1-DBD (Cy5), AtARF5-DBD (ATTO647N) and AtARF5-DBD (Cy3B). The displacements perpendicular to the flow (along x, top) and parallel to the flow (along y, bottom) are best fitted with two Gaussians. The slow population (narrow Gaussian) belongs to the proteins absorbed to the surface of the channel as it appears centred around zero even along the direction of flow. The width of the Gaussian fit for this population equals double the localization error and was used to correct the diffusion coefficient D of the freely moving population (wide Gaussian). The resulting diffusion coefficient along x was $78\,\mu\mathrm{m}^2/\mathrm{s}$ for AtARF1-DBD labelled with Cy5, $65\,\mu\mathrm{m}^2/\mathrm{s}$ for AtARF5-DBD labelled with ATTO647N and $69\,\mu\mathrm{m}^2/\mathrm{s}$ for AtARF5-DBD labelled with Cy3B.

DNA promotes dimerization of AtARF5-DBD in solution

The standard DNA response element containing two AuxREs sequences arranged in an inverted manner and spaced by $7\,\mathrm{bp}$ (IR7) is known to be bound in a cooperative fashion by ARFs[10]; in chapter 3 we showed that this topology stabilizes the dimeric form of *A. thaliana* ARF2 with it being 40 times more stable on the DNA compared to in solution. We asked if this stabilization could be directly seen in solution using labelled ARFs and an unlabelled dsDNA oligo containing an IR7 response element.

We performed smFRET experiments using sALEX with samples containing $\approx 1\,\mathrm{nM}$ AtARF5-DBD (Cy3B) and $\approx 1\,\mathrm{nM}$ AtARF5-DBD (ATTO647N) in presence or in absence of $1\,\mathrm{nM}$ unlabeled DNA bearing the IR7 from the TMO5 promoter

known to be bound by AtARF5 *in vivo*[23] (see also Chapter 3). As the labelling positions of the protein were stochastic, multiple dyes can be located on the surface of each protein. Then, mixing two samples containing AtARF5-DBD labelled with either donor or acceptor dyes does not generate a well determined FRET signature for the dimer. However, the degree of dimerization can be inferred by the appearance of a signal at E^* larger than the one of the donor-only population $E_d^* \approx 0.1$. Without DNA being added in solution, the donor-only protein ($E^* < 0.3$) represented 89 % of the FRET distribution (Fig.5.5a). The remaining 11 % ($E^* > 0.3$) is characterized by a rather noisy estimation for the diffusion coefficient with values closer to the one of the monomer ($D_{th} = 73\,\mu m^2/s$). Therefore, AtARF5-DBD in absence of DNA is mostly found in its monomeric form, which further confirms the results presented in the previous section. When unlabelled IR7 was added at a concentration similar to the one of the protein ([DNA] $= 1\,nM$), approximatively 20 % of the FRET distribution is found at $E^* > 0.3$ and the diffusion coefficient of such fraction was consistent with the one of a dimer bound to the DNA ($D_{th} = 50\,\mu m^2/s$, Fig. 5.5b).

These results are consistent with a model in which a DNA bearing an IR7 DNA response element is capable of stabilizing the protein dimer by interacting simultaneously with two monomers and underlines once again the importance of cooperativity in this protein-DNA interaction.

Discussion

In this work, we used single-molecule fluorescence detection performed in fluidic devices to study the DNA binding domain of several members of the family of Auxin Response Factor (ARF) transcription factors and the interaction between them and their DNA response element.

We found that the interaction between doubly-labelled dsDNA bearing the response element and unlabelled ARF led to changes in both the FRET efficiency and the diffusion coefficient. Notably, the readout via a change in diffusion can be compared to the theoretical values for the DNA in the bound and unbound states. Furthermore, monitoring the diffusivity serves as an internal quality control check as potential surface-induced artifacts can easily by identified. We further investigated the possibility of performing titrations without the intervention of an operator. The samples were loaded sequentially inside the syringe and were separated using a bubble of air. This approach impeded the use of standard enzymatic oxygen scavenging
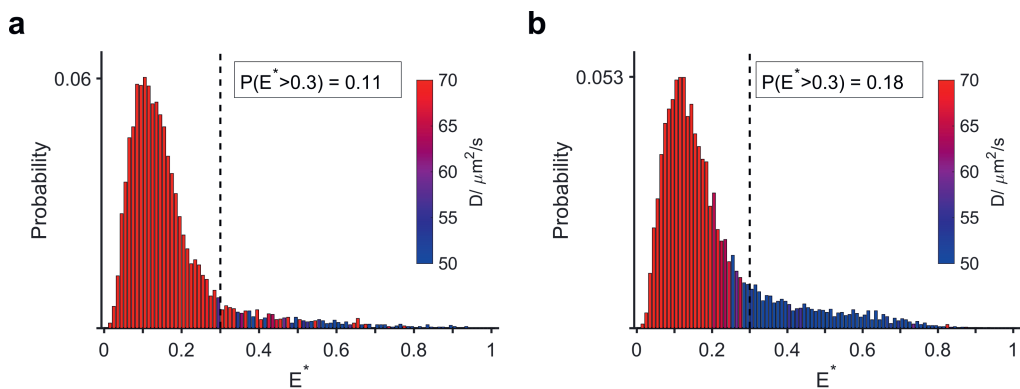
Figure 5.5: FRET efficiency distributions of samples containing donor and acceptor labelled AtARF5-DBD ($\approx 1\,$nM each) and different concentrations of unlabeled DNA. (a) In absence of DNA, the fraction of the probability distribution above $E^* = 0.3$ is characterized by a diffusion coefficient close to the one of the monomer ARFs, indicating that this fraction represents noise from the measurement. (b) When DNA is present at approximately the same concentration of the protein ($1\,$nM), the fraction of ARFs showing $E^* > 0.3$ increases. The associated diffusion coefficients decrease, consistent with the formation of a bulkier complex of a ARF dimer bound to DNA.

systems which in turn reduced the overall data quality; nevertheless, the characteristic reduction in FRET efficiency and diffusion coefficient in the samples containing ARF-DBDs was seen. In future experiments, the samples could be loaded in an inert environment (e.g. inert gas atmosphere), to suppress the presence of oxygen between the samples and enable the usage of effective oxygen scavenger systems.

The diffusion coefficients D of stochastically labeled AtARF5-DBD and AtARF1-DBD were then used to determine their state of oligomerization in free solution. The values found were compatible with a situation in which the DBDs are present mainly as monomers whereas in crystal structures ARF-DBDs are always present as dimers. This apparent discrepancy serves as a stark reminder that static images of crystal structures, although extremely useful, fail to capture the dynamic nature of biological processes. We further note that the analysis of the displacements along the flow inside the nanochannels allows detecting protein aggregation, if present, which would lead to a slowly diffusing population characterized by advection. As for the individuation of possible surface-induced artifacts, this is a quality control of the sample that is embedded into the methodology.

Lastly, we tested if the addition of unlabbeled DNA to a solution containing a mix of AtARF5-DBDs labelled with either donor or acceptor would help the formation of a dimer. Our result confirmed that, in absence of DNA and at a total concentration of $1\,\mathrm{nM}$, AtARF5-DBD is present primarily in its monomeric form. Moreover, the addition of DNA supported dimerization as seen by the increase on the population of molecules having $E^* > 0.3$ and $D \approx 50\,\mathrm{\mu m^2/s}$.

Conclusions

In this work we used SMFD inside parallel nanochannels to study the interaction between ARF-DBDs and between ARF-DBDs and their genomic response element. We showed that our approach is able to report on ARF binding by detecting changes in the FRET efficiency similar to experiments with immobilized DNA samples, while obtaining additional information in the form of the displacements distributions. The analysis of this distribution is able to identify the population of absorbed molecules and to separate it from the population of potential aggregates. Moreover, the obtained diffusion coefficient can be used to infer the nature of the tracked molecule/-complex. Taken together, our approach allowed us (I) to check for surface induced artifacts and sample quality *during* the experiment, (II) to confirm the results coming from changes in E^* with a second readout, (III) to infer the oligomerization state of ARF-DBDs in solution and (IV) to confirm the effect of the presence of the response element in solution on the dimerization of ARF-DBDs. In general, the combination of information from smFRET measurements together with displacement analysis makes SMFD inside nanochannels a powerful method to study protein-protein and protein-DNA association and dissociation.

References

1. Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. Advances in Single-Molecule Fluorescence Methods for Molecular Biology. *Annu. Rev. Biochem.* **77,** 51–76 (2008).
2. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17,** 27862–27872 (2015).
3. Kim, S., Streets, A. M., Lin, R. R., Quake, S. R., Weiss, S. & Majumdar, D. S. High-throughput single-molecule optofluidic analysis. *Nat Meth* **8,** 242–245 (2011).
4. Wunderlich, B., Nettels, D., Benke, S., Clark, J., Weidner, S., Hofmann, H., Pfeil, S. H. & Schuler, B. Microfluidic mixer designed for performing single-molecule kinetics with confocal detection on timescales from milliseconds to minutes. *Nat Protoc* **8,** 1459–1474 (2013).
5. Tyagi, S., VanDelinder, V., Banterle, N., Fuertes, G., Milles, S., Agez, M. & Lemke, E. A. Continuous throughput and long-term observation of single-molecule FRET without immobilization. *Nat Meth* **11,** 297–300 (2014).
6. Kim, J.-Y., Kim, C. & Lee, N. K. Real-time submillisecond single-molecule FRET dynamics of freely diffusing molecules with liposome tethering. *Nature Communications* **6,** 6992 (2015).

7. Leslie, S. R., Fields, A. P. & Cohen, A. E. Convex Lens-Induced Confinement for Imaging Single Molecules. *Anal. Chem.* **82,** 6224–6229 (2010).

8. Fontana, M., Fijen, C., Lemay, S. G., Mathwig, K. & Hohlbein, J. High-throughput, non-equilibrium studies of single biomolecules using glass-made nanofluidic devices. *Lab Chip* **19,** 79–86 (2019).

9. Gilboa, B., Jing, B., Cui, T. J., Sow, M., Plochowietz, A., Mazumder, A. & Kapanidis, A. N. Confinement-Free Wide-Field Ratiometric Tracking of Single Fluorescent Molecules. *Biophysical Journal* **117,** 2141–2153 (2019).

10. Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., López-Vidrieo, I., Franco-Zorrilla, J. M., de Vries, S. C., Solano, R., Weijers, D. & Coll, M. Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell* **156,** 577–589 (2014).

11. Kato, H., Mutte, S. K., Suzuki, H., Crespo, I., Das, S., Radoeva, T., Fontana, M., Yoshitake, Y., Hainiwa, E., van den Berg, W., Lindhoud, S., Ishizaki, K., Hohlbein, J., Borst, J. W., Boer, D. R., Nishihama, R., Kohchi, T. & Weijers, D. Design principles of a minimal auxin response system. *Nature Plants* **6,** 473–482 (2020).

12. Van Dijk, M. & Bonvin, A. M. J. J. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res* **37,** W235–W239 (suppl_2 2009).

13. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., Goody, R. S., Gohlke, H. & Seidel, C. A. M. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Meth* **9,** 1218–1225 (2012).

14. Craggs, T. D., Sustarsic, M., Plochowietz, A., Mosayebi, M., Kaju, H., Cuthbert, A., Hohlbein, J., Domicevica, L., Biggin, P. C., Doye, J. P. K. & Kapanidis, A. N. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res* **47,** 10788–10800 (2019).

15. Holden, S. J., Uphoff, S., Hohlbein, J., Yadin, D., Le Reste, L., Britton, O. J. & Kapanidis, A. N. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal* **99,** 3102–3111 (2010).

16. Crocker, J. C. & Grier, D. G. Methods of Digital Video Microscopy for Colloidal Studies. *Journal of Colloid and Interface Science* **179,** 298–310 (1996).

17. Ortega, A., Amorós, D. & García de la Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal* **101,** 892–898 (2011).

18. Liang, H., Nam, W. J. & Fonash, S. J. A Novel Parallel Flow Control (PFC) System for Syringe-Driven Nanofluidics. **3,** 3 (2008).

19. Mathwig, K., Mampallil, D., Kang, S. & Lemay, S. G. Electrical Cross-Correlation Spectroscopy: Measuring Picoliter-per-Minute Flows in Nanochannels. *Phys. Rev. Lett.* **109,** 118302 (2012).

20. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* **3,** 891–893 (2006).

21. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *J. Am. Chem. Soc.* **131,** 5018–5019 (2009).

22. *CRC Handbook of Chemistry and Physics, 97th Edition* (ed Haynes, W. M.) 2643 (CRC Press, 2016).

23. Freire-Rios, A., Tanaka, K., Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., Fontana, M., Hohlbein, J., Boer, D. R., Mironova, V. & Weijers, D. Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in *Arabidopsis. Proc Natl Acad Sci USA,* 202009554 (2020).

Supplementary material

Supporting Note 1: Effect of protein content on the viscosity of a solution

Particles dissolved in a liquid act as obstacles that hinder its movement; this results in an increase of the viscosity of the liquid. A mathematical description of this effect was first derived by Einstein in his doctoral thesis[1,2] under the assumption that the liquid contains rigid and spherical particles at very low concentration (very low solid fraction φ)

$$\eta = \eta_0 \left(1 + \frac{5}{2}\varphi\right) \tag{S5.1}$$

and later extended by Batchelor[3,4] for monodisperse suspension with a maximum solid fraction $\varphi = 0.15 - 0.2$

$$\eta = \eta_0 \left(1 + 2.5\varphi + 6.2\varphi^2\right) \tag{S5.2}$$

where η_0 is the viscosity of the liquid in absence of particles and η is the viscosity of the suspension. This theoretical framework can be used to predict the effect of the presence of the protein on the viscosity of the solution[5,6]. In our case, we have a solution containing 256 nM of MpARF2 monomers (0.0108 mg/mL) and a monomer volume of $45 \, \text{nm}^3$ (average of the two monomers found in PDBID:6SDG, calculated using VADAR 1.5[7]), which results in a solid fraction φ of 6.75×10^{-6}. This value is three orders of magnitude lower than the solid fraction necessary to increase the viscosity by 1% ($\varphi = 4 \times 10^{-3}$).

Supporting Note 2: On the use of Cy3 in nanochannels

The use of Cy3 instead of Cy3B in nanochannels is challenging since its lower photostability (compared to Cy3B) is exacerbated at the high laser powers needed for stroboscopic ALEX. To obtain a comparable photon count from the single-molecule localizations, a ~ 6 fold increase in laser power would be required but cannot be achieved since the excitation power currently used is already close to the limit imposed by the objective. A way around this technical limitation could be prolonging the green excitation; however, this comes at the price of lowering the effectiveness of stroboscopic excitation, with the photons being spread on a bigger area and lowering the obtainable signal to noise ratio. See image below.
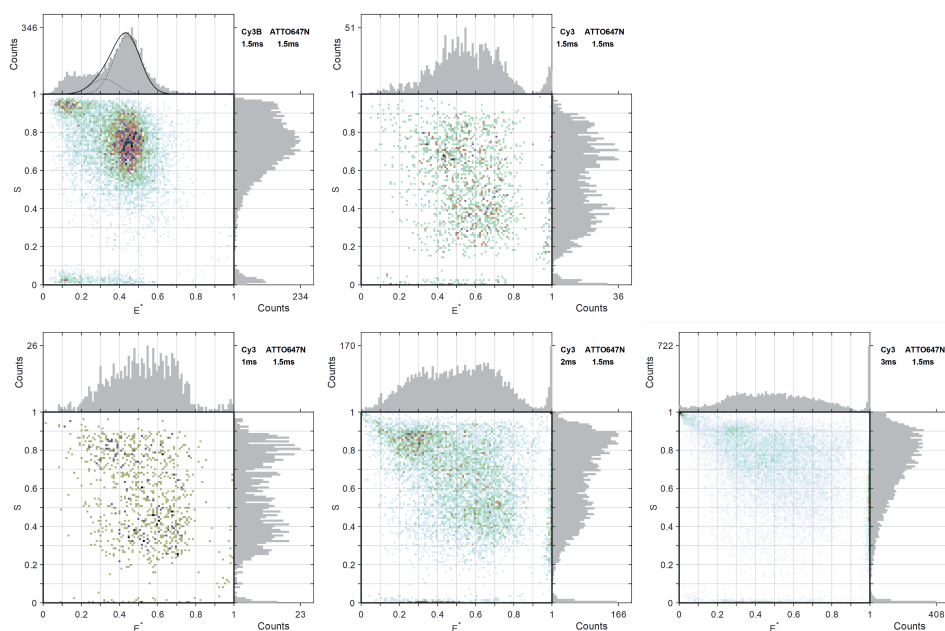
Figure S5.1: E^*S histograms comparing Cy3B/ATTO647N-labelled DNA with with Cy3/ATTO647N-labelled DNA imaged using different donor excitation times (both without OSS). Cy3 being several times dimmer than Cy3B makes single molecule detection challenging with the standard $1.5\,\mathrm{ms}$ illumination used for Cy3B. The laser being already set to maximum power leaves as only way of increasing the excitation to increase the excitation time. This comes at the cost of reducing the effectiveness of the stroboscopic illumination in reducing motion blur. As result in none of the tried excitation schemes distinguishable populations are present.

Passivation protocol for fluidic devices

1. Fill the channel with spectroscopy grade acetone. Use a glass Pasteur pipette.

2. Leave the acetone incubating for 5 min

3. Empty the channel. Use pressurized N_2, connect the gas outlet to a P1000 pipette tip and use it to push out the acetone

4. Prepare Vectabond solution (aminosilane solution). Add 0.1 mL of Vectabond reagent to 5 mL of acetone and mix thoroughly

5. Fill the channel with Vectabond solution. Put a droplet ($\approx 20\,\mu L$) of Vectabond solution on one end of the channel and move it into the channel by pulling from the other end. This can be done using a P1000 pipette (in case the pipette does not manage to pull enough use a syringe pump)

6. Leave the channel filled with Vectabond solution for 5 min

7. Empty the channel

8. Rinse the channel with acetone:

    (1) Fill the channel as in step 1

    (2) Leave the acetone incubating for 1 min

    (3) Empty the channel

9. Rinse the channel three times with milliQ:

    (1) Fill the channel as in step 5

    (2) Leave the MilliQ incubating for 1 min

    (3) Empty the channel

10. In a 1.5 mL Eppendorf tube, dissolve 40 mg of mPEG in 200 μL of MOPS buffer (50 mM MOPS, PH 7.5), mix the solution carefully

11. Spin down any remaining mPEG aggregates by 1 min centrifugation at 16000g

12. Fill the channel with mPEG solution as described in step 5. Do not take the solution from the bottom of the tube since you might collect the mPEG aggregates. Once the channel is filled, put two droplets of mPEG solution on the channel ends to prevent the channel to dry

13. Leave the device in a wet chamber at room temperature for 3 h. During the incubation check the droplets at the channel ends and add MOPS buffer if they are drying out

14. Empty the channel

15. Rinse the channel 3 times with PBS:

    (1) Fill the channel as in step 5

    (2) Leave the PBS incubating for 1 min

    (3) Empty the channel

16. Fill the channel with PBS leaving two droplets at its ends. Store the passivated channel in a wet chamber at 4 °C

5

Cleaning protocol for fluidic devices

1. Wash the exterior of the fluidic device soap and hot water to remove visible dirt and immersion oil

2. Rinse the chip with deionized water and dry it using a tissue paper

3. Prepare cleaning solution for the channels. Add 1 mL of Hellmanex®III (Hellma, Germany) detergent to 50 mL MilliQ

4. Fill the channels with the cleaning solution using a syringe pump

5. Place the device into a suitable container (*e.g.* a Petri dish or a beaker) and pour the remaining detergent solution into it. Cover the container and store it in the fume hood overnight

6. Rinse the chip with deionized water and dry it using a tissue paper

7. Empty the channels. Use pressurized N_2, connect the gas outlet to a P1000 pipette tip and use it to push out the cleaning solution

8. Wash channels three times with MilliQ times

    (1) Fill the channels. Put a droplet ($\approx 20\,\mu$L) of PBS on one end of the channel and move it into the channel by pulling from the other end. This can be done using a P1000 pipette (in case the pipette does not manage to pull enough, use a syringe pump)

    (2) Leave the MilliQ incubating for 1 min

    (3) Empty the channels

9. Burn the devices in a furnace

    (1) Heating rate 90 °C/h

    (2) Plateau at 500 °C for 1 h

    (3) Cooling rate 150 °C/h

10. Store cleaned devices in clean tinfoil at room temperature.

References

1. Einstein, A. Eine neue Bestimmung der Moleküldimensionen. *Annalen der Physik* **324,** 289–306 (1906).
2. Einstein, A. Berichtigung zu meiner Arbeit: "Eine neue Bestimmung der Moleküldimensionen". *Annalen der Physik* **339,** 591–592 (1911).
3. Batchelor, G. K. The stress system in a suspension of force-free particles. *Journal of Fluid Mechanics* **41,** 545–570 (1970).
4. Batchelor, G. K. The effect of Brownian motion on the bulk stress in a suspension of spherical particles. *Journal of Fluid Mechanics* **83,** 97–117 (1977).
5. Tu, R. S. & Breedveld, V. Microrheological detection of protein unfolding. *Phys. Rev. E* **72,** 041914 (2005).
6. Gonçalves, A. D., Alexander, C., Roberts, C. J., Spain, S. G., Uddin, S. & Allen, S. The effect of protein concentration on the viscosity of a recombinant albumin solution formulation. *RSC Adv.* **6,** 15143–15154 (2016).
7. Willard, L. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Research* **31,** 3316–3319 (2003).

5

Chapter 6

General discussion

6

Throughout this thesis we developed and applied methods to study protein-DNA and protein-protein interactions one molecule at a time. We focused our study on an important transcription factor of plants called Auxin Response Factor (ARF) and its interaction with its DNA response element. We set to obtain a quantitative understanding of this system that was not present in literature. To this aim, we based our DNA-binding assay on two well-characterized photophysical effects, namely Förster Resonance Energy Transfer (FRET) and Protein-Induced Fluorescence Enhancement (PIFE).

ARF-DNA interaction: a quantitative view

As discussed in **chapter 2**, the interaction between ARF and its response element (AuxRE) has been studied extensively and with a wide range of techniques. Despite this, limited quantitative information about the interaction was available. In **chapter 2** and **3** we applied our methods to the known interaction between an IR7 (two AuxREs in an inverted repeat and spaced by 7 bp) present in the promoter of the gene TMO5 in *Arabidopsis thaliana* and ARF5[1]; we found that the K_d of the interaction is 15(2) nM. When we tested two class B ARF (AtARF2-DBD and AtARF1-DBD) on the same DNA construct, the K_d was ≈ 500 nM. Interestingly, when a single mutation in a key amino acid for the dimerization of the DBD was introduced, AtARF5-DBD showed a behaviour similar to AtARF2-DBD and AtARF1-DBD both in the kinetics and the affinity (see **chapter 3**). This result showed once again the importance of protein dimerization to obtain stable binding on the DNA and highlighted the capacity of our method to obtain quantitative information.

We tested the DBDs of AtARF1 and AtARF5 on different AuxRE topologies (published in [1]) showing that they bind directed repeats (DR5) with an ≈ 5 fold decrease in affinity compared to an IR8. Strikingly, titrations using the DR5 construct followed a simple binding isotherm whereas titration using IR8 showed a steeper binding response (the latter being a characteristic seen for all DBDs when titrated on inverted topologies constructs throughout this thesis). This behaviour is consistent with the fact that the DBDs cannot dimerize through their dimerization domain when binding to repeated response elements which are in a direct orientation. This hypothesis can be tested in future experiments by titrating ARF DBDs mutated in their dimerization domain; these variants should bind the DR5 with affinity comparable to the wild type while showing a strong reduction of affinity

towards the IR8. Regarding the PB1 domain, thanks to the predicted flexibility of the middle region[2], it should be able to stabilize the binding on any topology; to confirm this conclusion, future experiments with full-length ARF variants would be necessary.

PB1: different functions, one domain

The PB1 domain is present in both ARFs and their repressors Aux/IAAs and mediates their association through an homotypic interaction. Since this interaction is head-tail, it should allow hetero and homo oligomerization of ARFs and AuxIAAs. The role of the PB1 domain to mediate the binding of Aux/IAA to ARFs allowing for repression to take place is well established but to which extent the PB1 domain helps to stabilize DNA binding has been an open question. In **chapter 3** we tested the binding affinity of different full-length variants of AtARF2 towards an IR7 response element. We found that the addition of the PB1 domain caused a 5-fold increased in affinity and that this effect was due to the enhanced stability of the dimer. Strikingly, through the analysis of both kinetics and affinity data based on a four-states model, where ARF can bind the DNA as monomer or dimer and can dimerize in solution or on the DNA, we were able to quantify the protein dimerization affinity. For example, AtARF2-DBD showed a dimerization K_d of $\approx 2\,\mu$M. Through the comparison of the kinetics of AtARF2-DBD-IR7 and AtARF2-FL-IR7 with AtARF5-DBD-IR7, it is possible to place the dimerization K_d of AtARF5-DBD in the hundreds of nM range; this higher dimer stability in conjugation with an intrinsic higher affinity of single monomers to AuxREs allows AtARF5-DBD to display stable binding of the DNA even in absence of the PB1 domain. This evidence is in agreement with *in vivo* experiments in which AtARF5ΔPB1 was seen to be a gain-of-function mutant that can activates auxin-responsive genes even in absence of the hormone[3]. Then, the role of the PB1 domain of AtARF5 appears to be only to bind the PB1 domain of Aux/IAAs coupling the transcriptional output with the presence of auxin. This hypothesis is confirmed by the fact that the PB1 domain of AtARF5 has a homodimerization K_d of 870 nM but an heterodimerization K_d with the PB1 domain of Aux/IAA17 of 73 nM[4]. Strengthening this hypothesis, AtARF5 and many others activator ARFs (also called class A ARFs) have been found to interact with many different Aux/IAA in a series of protein-protein interaction assays[5–16]. A different scenario is seen in case of AtARF2 (a class B ARF), where our data suggest that the interaction between PB1 of different AtARF2 monomers might be necessary to

achieve the stable binding to the DNA necessary for the protein function. This behaviour of the PB1 domain might be a common feature of other repressor (class B) ARFs and explains why this class of ARFs has been seen to interact with fewer members of the Aux/IAA family[5,13–15,17].

Summarizing, a model where the PB1 domain has two different function in the two mayor ARF classes can be drawn. In class A ARFs, the PB1 domain serves as mediator of auxin responsiveness whereas in class B ARF it stabilizes the DNA binding. This model of action for the PB1 domain is similar to the one found in *M. polymorpha* as part of the minimal auxin response system that we recently published ([18]) with one difference: MpARF1 (the only class A ARF in this species) cannot function without its PB1 domain. Therefore, the PB1 domain of class A ARFs in *M. polymorpha* probably has the double function of stabilizing the binding to the DNA and interacting with MpAux/IAA. This double function opens the possibility of a double repression by Aux/IAA, where to the standard recruitment of the corepressor TOPLESS[19] (TPL) an additional destabilization of ARF-DNA interaction might play a role. We note that the relevance and entity of this second type of repression can be easily tested using the DNA-binding assays presented in this thesis.

SMFD in microfluidic devices

In **chapter 4** we presented two microfluidic devices designed to provide the through-put typical of single-molecule fluorescence detection (SMFD) on a total internal reflection (TIRF) microscope while imaging non-immobilized molecules. The first device using an arrangement of parallel nanochannels allows to study systems that are in equilibrium. The second device features mixing nanochannels, allowing to mix two different samples and triggers reactions directly in the field of view; this allows to keep monitoring the reacting species before, during and after the reaction has started since it is kept in a steady state. In both designs, the molecules are kept inside the evanescent field thanks to the shallow depth of the nanochannels of only 200 nm. When imaging flowing and free-diffusing molecules it is important to reduce the motion blur induced by the movement. To this end we employed parallel flow control (PFC[20,21]) to obtain the required slow flow velocities (5–500 nm/ms) while the illumination was carried out with an updated stroboscopic illumination pattern that allowed to reduce motion blur while facilitating the tracking of the single molecules between coupled frames[22].

We tested the parallel nanochannels design studying (I) the well characterized interaction between DNA polymerase I (KF) and a doubly-labelled gapped DNA[23] and (II) the conformational dynamics of a DNA hairpin[24]. In both cases, the results were similar to the ones obtained with immobilized molecules with the added benefit that bleached molecules were promptly replaced by new ones thanks to the convective motion of the fluid in the channels. This is particularly important for applications in which camera-based detection is pushed to its limits (studying sub-millisecond dynamics) since the high laser intensities that are required will photobleach all fluorescent molecules immobilized in a given field of view in seconds even in presence of oxygen scavenging systems.

The parallel nanochannels proved useful in studying the interaction between ARF-DBD and its response element. In **chapter 5** we demonstrated that quantitative data similar to the one coming from immobilized samples can be obtained from smFRET experiments using free-flowing doubly-labelled DNA constructs. The similarities of these results served as an additional proof that surface immobilization does not interfere with the assays introduced in **chapter 2**. Moreover, inside the nanochannels binding was detected with a second, independent, readout which is the decrease in diffusion coefficient upon binding. The values of the diffusion coefficients of fluorescently labelled AtARF5-DBD and AtARF1-DBD was then used to asses that at the low concentrations tested in our channels ($\approx 1\,\text{nM}$) ARF-DBDs are present as monomers which is consistent with the data available in literature[25] as well as the one presented in **chapter 3**. As an additional information, the analysis of the diffusion coefficient of labelled ARF in solution provides a fast way to check for the presence of protein aggregates.

Finally, flowing a mixture of AtARF5-DBDs labelled with a donor or an acceptor further proved that no dimer is formed in solution at these concentrations since no FRET was detected. When an unlabelled DNA containing an IR7 response element was added, a small population displaying FRET and characterized by slow diffusion coefficient emerged. This result is consistent with the formation of dimer on the DNA which is possible at such low ARF-DBD concentration thanks to the IR7-induced stabilization of the dimer. This effect is predicted by the four-state model presented in **chapter 3** and highlights how all interactions between the members of a complex contributes cooperatively to its stability.

Future Outlook

In the past years, our understanding of the nuclear auxin pathway has increased considerably but the are still many outstanding questions that can (and probably will) be answered in the next years. As discussed earlier in this chapter, the effect of the presence of the PB1 domain when AuxREs are in an everted or directed repeat can be easily analyzed using the methods presented **chapter 2** and **3**; moreover, if Aux/IAA can act also via destabilization of ARF-DNA interaction can be checked using our methods. These experiments are performed adding increasing amounts of Aux/IAA while the concentration of ARF is kept close to the one that gives 50 % DNA bound, expecting to see a decrease in bound population as the concentration of repressor increases. When we tested AtARF2-FL with AtAux/IAA 12 and AtAux/IAA 33, no effect on the binding was observed even when the repressor was present in μM concentration. These negative results obtained using a class B ARF are consistent with the lack of interaction seen using other methods[15] and can be repeated in the future as negative control for experiments in which full-length class A ARF will be tested.

Since ARFs (and Aux/IAA) can potentially hetero oligomerize, understanding the composition and stoichiometry of ARF complexes is challenging, but necessary to understand the function of these proteins. We deem single-molecule pull-down[26] (SiMPull) as the method that can best answer this question. SiMPull has been already used in planta to reveal the stoichiometry of the HD-ZIPIII:LITTLE ZIPPER complex[27] and can easily be applied to the study of ARF and ARF:Aux/IAA complexes.

In this thesis we applied single-molecules fluorescence techniques to study ARF-DNA and ARF-ARF interactions. We see our work as being part of a bigger change of paradigm, where method development is coupled to the problem it has to solve. We foresee a future in which there will be a tighter integration between scientist that develop methods and scientist who utilize them.

References

1. Freire-Rios, A., Tanaka, K., Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., Fontana, M., Hohlbein, J., Boer, D. R., Mironova, V. & Weijers, D. Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in *Arabidopsis. Proc Natl Acad Sci USA,* 202009554 (2020).
2. Roosjen, M., Paque, S. & Weijers, D. Auxin Response Factors: output control in auxin biology. *J Exp Bot* **69,** 179–188 (2018).

3. Krogan, N. T., Ckurshumova, W., Marcos, D., Caragea, A. E. & Berleth, T. Deletion of MP/ARF5 domains III and IV reveals a requirement for Aux/IAA regulation in Arabidopsis leaf vascular patterning. *New Phytologist* **194,** 391–401 (2012).

4. Han, M., Park, Y., Kim, I., Kim, E.-H., Yu, T.-K., Rhee, S. & Suh, J.-Y. Structural basis for the auxin-induced transcriptional regulation by Aux/IAA17. *PNAS* **111,** 18613–18618 (2014).

5. Ouellet, F., Overvoorde, P. J. & Theologis, A. IAA17/AXR3: Biochemical Insight into an Auxin Mutant Phenotype. *The Plant Cell* **13,** 829–841 (2001).

6. Hamann, T., Benkova, E., Bäurle, I., Kientz, M. & Jürgens, G. The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes & Development* **16,** 1610–1615 (2002).

7. Tatematsu, K., Kumagai, S., Muto, H., Sato, A., Watahiki, M. K., Harper, R. M., Liscum, E. & Yamamoto, K. T. MASSUGU2 Encodes Aux/IAA19, an Auxin-Regulated Protein That Functions Together with the Transcriptional Activator NPH4/ARF7 to Regulate Differential Growth Responses of Hypocotyl and Formation of Lateral Roots in Arabidopsis thaliana. *The Plant Cell* **16,** 379–393 (2004).

8. Hardtke, C. S., Ckurshumova, W., Vidaurre, D. P., Singh, S. A., Stamatiou, G., Tiwari, S. B., Hagen, G., Guilfoyle, T. J. & Berleth, T. Overlapping and non-redundant functions of the Arabidopsis auxin response factors MONOPTEROS and NONPHOTOTROPIC HYPOCOTYL 4. *Development* **131,** 1089–1100 (2004).

9. Weijers, D., Benkova, E., Jäger, K. E., Schlereth, A., Hamann, T., Kientz, M., Wilmoth, J. C., Reed, J. W. & Jürgens, G. Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *The EMBO Journal* **24,** 1874–1885 (2005).

10. Fukaki, H., Nakao, Y., Okushima, Y., Theologis, A. & Tasaka, M. Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in Arabidopsis. *The Plant Journal* **44,** 382–395 (2005).

11. Muto, H., Nagao, I., Demura, T., Fukuda, H., Kinjo, M. & Yamamoto, K. T. Fluorescence Cross-Correlation Analyses of the Molecular Interaction between an Aux/IAA Protein, MSG2/IAA19, and Protein–Protein Interaction Domains of Auxin Response Factors of Arabidopsis Expressed in HeLa Cells. *Plant and Cell Physiology* **47,** 1095–1101 (2006).

12. Uehara, T., Okushima, Y., Mimura, T., Tasaka, M. & Fukaki, H. Domain II Mutations in CRANE/IAA18 Suppress Lateral Root Formation and Affect Shoot Development in Arabidopsis thaliana. *Plant and Cell Physiology* **49,** 1025–1038 (2008).

13. Shen, C., Wang, S., Bai, Y., Wu, Y., Zhang, S., Chen, M., Guilfoyle, T. J., Wu, P. & Qi, Y. Functional analysis of the structural domain of ARF proteins in rice (Oryza sativa L.) *Journal of Experimental Botany* **61,** 3971–3981 (2010).

14. Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Daele, H. V. d., Legrand, J., Oliva, M., Das, P., Larrieu, A., Wells, D., Guédon, Y., Armitage, L., Picard, F., Guyomarc'h, S., Cellier, C., Parry, G., Koumproglou, R., Doonan, J. H., Estelle, M., Godin, C., Kepinski, S., Bennett, M., Veylder, L. D. & Traas, J. The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Molecular Systems Biology* **7,** 508 (2011).

15. Piya, S., Shrestha, S. K., Binder, B., Stewart, C. N. J. & Hewezi, T. Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis. *Front. Plant Sci.* **5** (2014).

16. Pierre-Jerome, E., Moss, B. L., Lanctot, A., Hageman, A. & Nemhauser, J. L. Functional analysis of molecular interactions in synthetic auxin response circuits. *PNAS* **113,** 11354–11359 (2016).

17. Ulmasov, T., Murfett, J., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *The Plant Cell* **9,** 1963–1971 (1997).

18. Kato, H., Mutte, S. K., Suzuki, H., Crespo, I., Das, S., Radoeva, T., Fontana, M., Yoshitake, Y., Hainiwa, E., van den Berg, W., Lindhoud, S., Ishizaki, K., Hohlbein, J., Borst, J. W., Boer, D. R.,

Nishihama, R., Kohchi, T. & Weijers, D. Design principles of a minimal auxin response system. *Nature Plants* **6,** 473–482 (2020).

19. Szemenyei, H., Hannon, M. & Long, J. A. TOPLESS Mediates Auxin-Dependent Transcriptional Repression During Arabidopsis Embryogenesis. *Science* **319,** 1384–1386 (2008).

20. Liang, H., Nam, W. J. & Fonash, S. J. A Novel Parallel Flow Control (PFC) System for Syringe-Driven Nanofluidics. **3,** 3 (2008).

21. Mathwig, K., Mampallil, D., Kang, S. & Lemay, S. G. Electrical Cross-Correlation Spectroscopy: Measuring Picoliter-per-Minute Flows in Nanochannels. *Phys. Rev. Lett.* **109,** 118302 (2012).

22. Fontana, M., Fijen, C., Lemay, S. G., Mathwig, K. & Hohlbein, J. High-throughput, non-equilibrium studies of single biomolecules using glass-made nanofluidic devices. *Lab Chip* **19,** 79–86 (2019).

23. Craggs, T. D., Sustarsic, M., Plochowietz, A., Mosayebi, M., Kaju, H., Cuthbert, A., Hohlbein, J., Domicevica, L., Biggin, P. C., Doye, J. P. K. & Kapanidis, A. N. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res* **47,** 10788–10800 (2019).

24. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17,** 27862–27872 (2015).

25. Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., López-Vidrieo, I., Franco-Zorrilla, J. M., de Vries, S. C., Solano, R., Weijers, D. & Coll, M. Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell* **156,** 577–589 (2014).

26. Fareh, M., Loeff, L., Szczepaniak, M., Haagsma, A. C., Yeom, K.-H. & Joo, C. Single-molecule pull-down for investigating protein–nucleic acid interactions. *Methods. Single molecule probing by fluorescence and force detection* **105,** 99–108 (2016).

27. Husbands, A. Y., Aggarwal, V., Ha, T. & Timmermans, M. C. In Planta Single-Molecule Pull-Down Reveals Tetrameric Stoichiometry of HD-ZIPIII:LITTLE ZIPPER Complexes. *The Plant Cell* **28,** 1783–1794 (2016).

6

Chapter 7

Summary

7

The plant hormone auxin is a key regulator of many growth and developmental processes. The presence of auxin changes the gene expression of the cell via a short pathway called the nuclear auxin pathway (NAP). The NAP contains three players: TIR1/AFB, Aux/IAA and the transcription factor ARF. Under conditions of low auxin, Aux/IAA binds ARF repressing it. At higher concentrations of auxin, TIR1/AFB binds Aux/IAA and marks it for degradation, thereby freeing ARF. Members of the ARF family bind a DNA motif called AuxRE (TGTCTC) and their DNA binding domains (DBDs) are known to dimerize and bind cooperatively on response elements composed by two AuxREs present in an inverted orientation. Although the affinities at play inside the NAP and between ARFs and AuxREs had been studied extensively with a range of techniques, a quantitative view on this system was still lacking. Throughout this thesis, we therefore developed and applied methods with the goal of obtaining a quantitative understanding of ARF-ARF and ARF-AuxRE interactions.

In **chapter 2** we developed and utilized a method based on smFRET and smPIFE-FRET to study the interaction between ARF and composite DNA response elements. We tested several *A. thaliana* ARF-DBDs and showed that the differences in binding stability and kinetics are quantitatively captured by the method. The smFRET and smPIFE-FRET experiments in which the affinity of a DNA-binding deficient mutant of AtARF5-DBD (R215A) was probed, proved that the shift in FRET efficiency seen when wt AtARF5-DBD is added in concentration of tens on nM is due to the specific protein interaction with the DNA motif. We then tested the binding preferences of *Marchantia polymorpha* ARF1 and ARF2, which are, respectively, the only member of the phylogenetically conserved class A-ARF and B-ARF in the species. Whereas A-ARFs are considered activators, B-ARFs are considered repressors. In *M. polymorpha*, a mechanism in which B-ARFs repression is based on competition for the same AuxREs bound by A-ARFs had been proposed as part of a minimal auxin response system. We validated this model on a composite response element carrying two TGTCTC in an inverted repeat spaced by 7 basepair (IR7-TGTCTC) showing that the ratio between the affinity of MpARF2-DBD and MpARF1-DBD is ≈ 5, allowing for competition; similarly, IR7-TGTCGG showed a ratio of ≈ 3. Interestingly, IR7-TGTCAA showed a ratio of ≈ 13 which hints to a regulation for this element that does not involve competition between A and B-ARFs.

7

We then applied the method based on smPIFE-FRET to determine the effect of dimerization and oligomerization on the affinity and kinetics of ARF-AuxRE interaction. In **chapter 3** we tested the binding affinity of different variants of AtARF2 towards an IR7 response element. We showed that the presence of the PB1 domain increases the affinity of this protein-DNA interaction and that this effect is caused by the increased stability of the protein dimer brought by the homotypic interaction of the PB1 domain; the effect of further oligomerization on the affinity towards this composite AuxRE was found to be negligible. The kinetics of AtARF2-IR7 interaction, showed and increase of the observed k_{on} and a decrease in the observed k_{off} in the full-length protein compared to the DBD, with full-length variants carrying mutations that reduce the interaction between the PB1 domains (K2S and OPCA) showing intermediate values. Remarkably, both the trends in the kinetics and the differences in affinity towards the DNA shown by the different AtARF2 variants were captured by a global fit based on a four-states association model. The model explained the differences between variants as being caused by changes in the dimerization equilibrium of the proteins and returned a dimerization K_d for AtARF2-DBD of $\approx 2\,\mu$M.

In **chapter 4** we introduced two microfluidic devices aimed for single-molecule fluorescence detection on a TIRF microscope without requiring the immobilisation of DNA constructs. Both designs feature a (nano-)channel height of 200 nm which confines the fluorescent molecules into the evanescent field of the TIRF microscope allowing for single-molecule tracking to take place. The first design featuring parallel nanochannels was tested studying the well characterized interaction between DNA polymerase I (KF) and a gapped DNA construct and the conformational dynamics of a DNA hairpin; these experiments returned results comparable with the ones obtained in smFRET experiments with immobilized DNA molecules. Using the second design featuring a T-shaped mixing (nano-)channel we accessed non-equilibrium conditions by mixing primarily open DNA hairpins with a high-salt solution triggering the closing of DNA hairpins. Moreover, we observed polymerization of 25 bases on a DNA template by a DNA polymerase, illustrating that complex biological reactions can be followed in real time and in a continuous fashion.

The parallel nanochannels proved useful in studying the interaction between ARF-DBDs and between ARF-DBD and its DNA response element. In **chapter 5** we demonstrated that quantitative data similar to the one coming from immobilized samples can be obtained from smFRET experiments using free-flowing

doubly-labelled DNA constructs; moreover, a decrease in diffusion coefficient can be detected and used as an independent readout for the DNA binding.

The values of the diffusion coefficients of fluorescently labelled AtARF5-DBD and AtARF1-DBD showed that they are present as monomers at the $\approx 1\,\text{nM}$ concentrations tested. Finally, flowing a mixture of AtARF5-DBDs labelled with a donor or an acceptor showed that the addition of unlabelled DNA containing an IR7 response element is likely to stabilize the dimer as a population displaying FRET and characterized by slow diffusion coefficient emerged. This stabilization once again underlines the importance of cooperativity for the stability of biomolecular complexes.

7

Acknowledgments

Here we are, at the end of my thesis, which coincides with the end of my PhD. As Elcaro said "Everything that has an end has a beginning" and, as the concepts of end and beginning are deeply connected, it makes sense at the end of a journey to look back to its start. Six years ago I decided that I wanted to do my MSc thesis at the Laboratory of Biophysics. I bought a one-way ticket to the Netherlands and I landed on a sunny day of April. The sunny and warm weather continued for several days, long enough that I started doubting my model of reality concerning the climate of Northern Europe. Then came May and with it the clouds. I haven't seen the sun since.

During my master thesis I came to appreciate Johannes technology-driven approach to science, where new methods and microscopes are conceived. I also liked his style of supervision, always eager to help and to provide guidance. During the 4 years of my PhD my opinion on Johannes has not changed. **Johannes**, I want to thank you for your support and your patience during the times of my PhD where I was overwhelmed by my own negativity.

Next I would like to thank **Dolf** for being available for discussion on research direction and for offering insightful perspectives on the results (good and bad) that I had during my PhD. I appreciate your effort to create a team out of the Plant Development group and Biochemistry at large. I will never forget the Christmas Dinners, the lab weekends, and that time when I beat you at beer pong. Talking about BIC activities, I want to thank the people that participated at the PhD-Trip in Barcelona (**Heidi**, **Thijs**, **Tom**, **Maritza**, **Berdien**, **Nicole**, **Wouter**, **Margo** and **Dasha**). From that trip I got some of the best memories of my PhD, but what is more important is that it allowed me to get to know you and bond with you. **Sacco**, thank you for organizing it and for working as our personal tour guide for a

week. Moreover, thank you for all the discussions we had in the coffee room, seldom serious but always great.

Doing Science is a team effort so I would like to thank all the people that helped me in my endeavor. **Willy**, **Simon** and **Mark**, thank you for your support regarding protein purification. **John**, **Arjen**, **Rob**, **Cor**, **Brenda** and **Laura**, thank you for the technical support you gave me. **Carel**, I want to thank you for all you taught me during my MSc, knowledge that helped me a lot during my PhD; moreover, I'd like to thank you for all your bad jokes (they were not that bad after all). A big thank goes to all my students (**Meike**, **Ernesto**, **Sven**, **Šarūnė** and **Elmar**), you did a wonderful job.

I would like to thank **all** my collegues at BIP (past and present) which contributed to the warm and friendly environment of the Lab. A big shout-out goes to **Shazia**, once the social backbone of BIP, still a great friend. **Koen**, I've always enjoyed discussing with you, about science, technology, anything really. It was great to have you in the same office for 4 years. **Klausja**, thank you for the bike rides and the yoga lessons; fortunately, we had to do the pose of the "fallen rider" only during the latter. **Morwarid**, thank you for making me feel at home in your container and for giving me Dutch lessons. **Suyeon**, thank you for the great time we had together, at Onder de Linden, at Su's café or fixing your bike. I'm looking forward to visit Jeju with you.

Seleen, I would like to thank you and your family. You included me in your inner circle like no one else did during my time in The Netherlands and for this I am thankful. That night in Den Haag when we were surrounded by fire and explosions I was not afraid because you were there with me. **Max**, we are the last survivors of the initial group from Plantsoen. I enjoyed the time we spent together in the past years and I'm looking forward to have more dinners and parties together.

Cambiando lingua, vorrei ringraziare **Alejandro Hochkoeppler**. Il vero inizio della mia avventura nei Paesi Bassi é stato quando, durante una pausa tra due ore di lezione, decisi di chiedere al Professore di un corso interessantissimo (Ingegneria Proteica) se avesse contatti all'estero per una tesi. Il resto é storia. Grazie Alejandro anche per essere stato il mio External Advisor durante il mio PhD.

Un ringraziamento speciale va alle persone che sono risucite a starmi vicine nonostanze la distanza. Grazie agli (ex) membri di **Casa Grinton**. Nonostante i

142

molti anni passati da quando vivevamo sotto lo stesso tetto ed il fatto di essere oramai sparpagliati in mezzo mondo, vi sento ancora vicini. Anche nei momenti piú bui sapevo che eravate sempre ad un tap di distanza (e per questo devo ringraziare anche **Signal**). So che posso contare su di voi.

Parlando di persone contemporaneamente lontane e vicine, un grazie va ad un gruppo di soliti noti: gli **Amici Scopa**. In quanto a scopa il 2020 é andato com'è andato, iniziato con una disfatta clamorosa (ti chiedo ancora scusa **Norman**) seguita da un virus che ha impedito ogni possibilitá di rivincita. Scopa a parte le mie vacanze a Tonezza non sarebbero state le stesse senza di voi. Un ringraziamento va anche a **Danilo**, **Giorgia**, **Daniel** e **Daniela**, sempre presenti quando una sciabolata diventa necessaria. Chiudendo il ciclo tonezzano, ringrazio la pizzeria **La Pizza**, sia i colletti bianchi che i colletti blu. Siete sempre un punto di ritrovo importante. Un grazie a **Deza**, nonostante le poche occasioni di incontro durante gli ultimi anni ti sento vicino come quando vivevamo a 500 metri di distanza.

Chi invece mi é sempre vicino per forza di cose (specialmente in caso di lockdown) é **Nicola**. Grazie per avermi introdotto al fantastico mondo dell'Ingegneria Navale e ai suoi abitanti. Spero di averti convinto del fatto che il Veneto (ed in particolare Tonezza) é il posto migliore del mondo.

Infine vorrei ringraziare la mia famiglia. Un grazie a mia zia **Linda** e alle cugine **Cinzia** (Kugi) e **Monica**. Un grazie speciale va a mia **mamma**, a mio **papá** e a mia **sorella** che mi accolgono a braccia aperte ogni volta che torno a casa. Grazie per avermi sempre sostenuto.

List of publications

This thesis

Fontana, M.[*] , Fijen, C.[*], Lemay, S.G., Mathwig, K. and Hohlbein, J. "High-throughput, non-equilibrium studies of single biomolecules using glass-made nanofluidic devices" *Lab on a Chip*

Kato, H., Mutte, S. K[*]., Suzuki, H.[*], Crespo, I.[*], Das, S.[*], Radoeva, T.[*], **Fontana, M.**[*], Yoshitake, Y.,Hainiwa, E., van den Berg, W., Lindhoud, S., Ishizaki, K., Hohlbein, J., Borst, J.W., Boer, D.R.,Nishihama, R., Kohchi, T. and Weijers, D. "Design principles of a minimal auxin response system" *Nature Plants* (2020)

Freire-Rios, A., Tanaka, K, Crespo, I., van der Wijk, E., Sizentsova, Y., Levitsky, V., Lindhoud, S., **Fontana, M.**, Hohlbein, J., Boer, DR., Mironova, V. and Weijers, D. "Architecture of DNA elements mediating ARF transcription factor binding and auxin-responsive gene expression in *Arabidopsis*" *PNAS* (2020)

Other publications

Fijen, C, Mahmoud, M. M., Kronenberg, M., Kaup, R., **Fontana, M.**, Towle-Weicksel, J. B., Sweasy, J. B., and Hohlbein, J. "Probing the nucleotide dependent conformational landscape of Pol β-DNA complexes using smFRET" *The Journal of Biological Chemistry* (2020)

[*] indicates equal contribution

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: **Mattia Fontana**
Date: **7 April 2021**
Group: **Biophysics & Biochemistry**
University: **Wageningen University & Research**

1) Start-Up Phase	*date*	*cp*
► **First presentation of your project**		
Illuminating plant hormone responses at the single-molecule level	30 Sep 2016	1.5
► **Writing or rewriting a project proposal**		
► **MSc courses**		
Subtotal Start-Up Phase		1.5

2) Scientific Exposure	*date*	*cp*
► **EPS PhD student days**		
EPS PhD student days, Soest	9-10 Feb 2017	0.6
EPS PhD student days, Soest	15-16 Feb 2018	0.6
EPS PhD student days, Soest	10-11 Feb 2020	0.6
► **EPS theme symposia**		
EPS theme 1 'Developmental Biology of Plants', Leiden	28 Feb 2017	0.3
EPS theme 1 'Developmental Biology of Plants', Leiden	31 Jan 2019	0.3
► **Lunteren Days and other national platforms**		
Annual Meeting 'Experimental Plant Sciences', Lunteren	10-11 Apr 2017	0.6
Annual Meeting 'Experimental Plant Sciences', Lunteren	9-10 Apr 2018	0.6
DutchBiophysics, Veldhoven	3-4 Oct 2016	0.6
DutchBiophysics, Veldhoven	2-3 Oct 2017	0.6
DutchBiophysics, Veldhoven	1-2 Oct 2018	0.6
► **Seminars (series), workshops and symposia**		
Symposium: 1st WURomics Symposium: Technology-Driven Innovation for Plant Breeding	15 Dec 2016	0.3
Symposium: Molecules @ WURk	8 Apr 2017	0.2
WEES Seminar: Prof. Richard Lenski, Dynamics of Adaptation and Genome Evolution in a Long-Term Experiment	31 Aug 2017	0.1
WMLS Seminar Series: Prof. Claus Seidel, Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems in vitro and in live cells	19 Oct 2017	0.1
Symposium: Future Biochemistry	26 Oct 2017	0.2
WCS Seminar: Prof. Thomas Ward, Artificial Metalloenzymes: Challenges and Opportunities	19 Dec 2017	0.1
Symposium: What is Life?	12-14 Mar 2018	0.6
WMLS Seminar Series: Prof. Hans Aertsl, Glycolipids in health and disease	19 Apr 2018	0.1
WMLS Seminar Series: Prof. Marileen Dogterom, Reconstituting minimal spindles in artificial confinement	21 Jun 2018	0.1
Seminar: Dr. Victoria Mironova, What we learnt about auxin transcriptional regulation from meta-analysis of whole genome data	27 Jun 2018	0.1
Seminar: Prof. Johan R.C. van der Maarel, Shaping the bacterial genome by nucleoid associated proteins Hfq, HU, and H-NS and some other stories	4 Jul 2018	0.1
Seminar: Prof. Jiri Friml, Evolution of auxin transport and efficient root gravitropism	18 Dec 2018	0.1
WMLS Seminar Series: Prof. Wolf Frommer, Logistics: Allocation of carbon and energy for yield and pathogen resistance	17 Jan 2019	0.1
WMLS Seminar Series: Prof. Timothy Noël, Innovation in synthetic methodology through use of flow	21 Feb 2019	0.1
WMLS Seminar Series: Prof. Enrico Mastrobattista, Lipid and protein-based drug delivery systems at the nano-scale	19 Sep 2019	0.1
WMLS Seminar Series: Dr. Remus Dame, Understanding chromosome organization in bacteria and archaea	17 Oct 2019	0.1
Symposium: NMR/MRI, plants and foods: does size matter?	25 Oct 2019	0.2
► **Seminar plus**		
► **International symposia and congresses**		
Structure and Dynamics of Biomolecules, Hünfeld, Germany	14-16 Feb 2019	0.9
PicoQuant 25th Anniversary Single Molecule Workshop, Berlin, Germany	3-6 Sep 2019	1.2
► **Presentations**		
Talk: ICFO's Training and Development Program, Barcelona, Spain	18 May 2017	1.0
Poster: DutchBiophysics, Veldhoven	2-3 Oct 2017	1.0
Poster: DutchBiophysics, Veldhoven	1-2 Oct 2018	1.0
Poster: PicoQuant 25th Anniversary Single Molecule Workshop, Berlin, Germany	3-6 Sep 2019	1.0
► **3rd year interview**		
► **Excursions**		
PhD Trip Biochemistry, Barcelona, Spain	15-19 May 2017	1.0
EPS PhD council company visit Dümmen Orange, De Lier	15 Jun 2018	0.3
EPS PhD council company visit Koppert Biological Systems, Berkel en Rodenrijs	26 Oct 2018	0.2
Subtotal Scientific Exposure		15.7

3) In-Depth Studies	*date*	*cp*
► **Advanced scientific courses & workshops**		
Postgraduate Course: Transcription Factors and Transcriptional Regulation, EPS	12-14 Dec 2016	1.0
Workshop: Advanced Fluoroscence Spectroscopy and Imaging, Düsseldorf, Germany	19-20 Sep 2018	0.6
► **Journal club**		
Journal club, Laboratory of Biophysics and Biochemistry	2016-2020	3.0
► **Individual research training**		
Subtotal In-Depth Studies		4.6

4) Personal Development	*date*	*cp*
► **General skill training courses**		
Symposium Publish for Impact, Wageningen University & Research Library	7 Feb 2017	0.2
PhD Introduction Course, EPS	16 Feb 2017	0.3
The Essentials of Scientific Writing and Presenting, Wageningen Graduate Schools	10-21 Mar 2017	1.2
Reviewing a Scientific Paper, Wageningen Graduate Schools	23 Mar 2017	0.1
PhD Workshop Carousel, Wageningen Graduate Schools	7 Apr 2017	0.3
Communication with the Media and the General Public, Wageningen Graduate Schools	9 Nov - 11 Dec 2017	1.0
Transferable Skills, EPS	28 Jun 2018	0.2
Basic Dutch 2, Wageningen in'to Languages	Sep-Dec 2019	1.5
Career Perspectives, Wageningen Graduate Schools	Oct-Nov 2019	1.6
► **Organisation of meetings, PhD courses or outreach activities**		
► **Membership of EPS PhD Council**		
Subtotal Personal Development		6.4

5) Teaching & Supervision Duties	*date*	*cp*
► **Courses**		
BIP-32306 Biophotonics and Photosynthesis	2016-2019	
BIP-10303 Physics MLS	2016-2019	3.0
BIP-32803 Biophysical Imaging	2017-2018	
BIC-20806 Enzymology	2018-2019	
► **Supervision of BSc/MSc students**		
BSc thesis: Characterization of DNA polymerase β binding to DNA. Meike Kronenberg	2018	
Internship: Artificial Virus-like Coat Proteins:	2018	
BSc thesis: Illuminating the AUX/IAA-ARF interaction. Sven Koens	2019	
Internship: Study of plant transcription factors and their interaction with DNA using single-molecule Förster resonance energy transfer (smFRET) in nanofluidic devices. Šarūnė Ivanovaitė	2019/2020	3.0
MSc thesis: Regulation of ARF-DNA binding affinity by AuxRE sequence and configuration. Elmar van der Wijk	2019/2020	
Subtotal Teaching & Supervision Duties		6.0

TOTAL NUMBER OF CREDIT POINTS*	**34.2**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

* A credit represents a normative study load of 28 hours of study.

About the author

Mattia Fontana was born on August 19, 1990 in Thiene, Italy. During his high school he studied mining engineering (ITIS *U. Follador*, Agordo, Italy). Being fascinated by the rapid development of biotechnology he decided to change field and enrolled at the University of Trento. He obtained his bachelor degree in Biotechnology (*cum laude*) with the thesis entitled "Post-transcriptional regulation mediated by the RNA-binding protein SCP160". He then completed *cum laude* the Master of Science in Industrial Biotechnology at the University of Bologna with the thesis entitled "Characterization of nanofluidics devices for high-throughput single-molecule fluorescence detection". The research for the MSc thesis was carried out at the Laboratory of Biophysics at Wageningen University & Research under the supervision of Johannes Hohlbein. In 2016, he started as a PhD candidate in a collaborative project between the Laboratory of Biophysics and Biochemistry at Wageningen University & research, the results of which are presented in this thesis.

Colophon

Cover: "Model of reality and reality of models", Mattia Fontana

Layout and cover design by the author

Mattia Fontana, 2021