Single-molecule studies using junctured-DNA mounted on a magnetic tweezers setup : characterizing the dynamics of interaction between the FKBP12 protien and several of its ligands

François Stransky

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Ecole Normale Supérieure de Paris

Etude molécule-unique de la dynamique des interactions entre FKBP12 et plusieurs de ses ligands grâce à un échafaudage en ADN manipulé par des pinces magnétiques

Soutenue par
François STRANSKY
Le 25 Février 2021

Ecole doctorale n° 474
CRI-FIRE « Frontiers in Innoviation, Research and Education »

Spécialité
Biologie fondamentale

Composition du jury :

Jean-François, ALLEMAND Professeur, ENS de Paris	*Président*
Etienne, DAGUE Chargé de Recherche au CNRS , Université de Toulouse	*Rapporteur*
Laurent, LIMOZIN Directeur de Recherche au CNRS, Aix Marseille Université	*Rapporteur*
Jean-Louis, BANERES Directeur de Recherche au CNRS, Université de Montpelier	*Examinateur*
Charlie, GOSSE Chargé de Recherche au CNRS, ENS de Paris	*Co-Directeur de thèse*
Térence, STRICK Professeur, ENS de Paris	*Directeur de thèse*

IBENS Institut de Biologie de l'Ecole Normale Supérieure

PSL UNIVERSITÉ PARIS

Frontières Du Vivant

CRI Catalysing a fruitful future

Abstract

The ability of proteins to selectively bind to specific partners is at the core of information processing in biological systems. Understanding that ability is key in drug design, as drugs rely on the accurate and effective targeting of precise therapeutic targets. Building on the magnetic tweezer technique, our team has designed the "Jointed DNA" (J-DNA) technique, a molecular scaffold made of DNA with two interaction-capable molecular partners attached. Under an optical microscope with a magnetic tweezer, we are able to follow in real time the opening and closing of the scaffold, and using a competitor in solution we can deduce the equilibrium constants and kinetics of the interaction. We used a known ligand of the FKBP12 complex, JK313, as our reference partner on a scaffold with JK313 and FKBP12. We could then measure the affinity for their cellular target of other molecules proposed as drugs in the treatment of mTOR pathway pathologies. We then investigated the question of the force exerted on the scaffolds and developed an automated data-processing method using hidden markov modelling software. We conclude on the comparison of this method in therms of cost, speed, reliability and type of results to other current methods in biomedical and single-molecule research.

Acknowledgements

Throughout the writing of this dissertation I have received a great deal of support and assistance.

I would first like to thank Prof. Terence Strick and Dr. Charlie Gosse for their assistance and comments during my work and writing, their helpful comments on the results and this manuscript, and their passion for science.

I would like to thank my reviewers, Dr. Etienne Dague and Dr. Laurent Limozin for their insightful discussion on this manuscript as well as Prof. Jean-François Allemand and Dr. Jean-Louis Banères for their participation in the thesis jury. I would also like to thank my colleagues, Dr. Dorota Kostrz, Maryne Follenfant and Leah Friedman, for making the jDNA scaffold and their pioneering work on this topic. This thesis would not have been possible without you. I would also like to acknowlege Christian Meyners, Patrick Purder, Sebastian Pomplun, Christian Kozany, Jürgen Kolos and Felix Hausch for providing us with the competitors we used in this work.

I would like to thank professors Jean-François Allemand and Marc Nadal for their kind advice and stimulating discussions throughout the three and a half years I spent in the laboratory, and Dr. olivier Bensaude for his participation in my thesis advisory committee. I would also like to acknowledge Dr. Jean-François Léger and Dr. Laurent Bourdieu, my tutors during my studies at the ENS, for showing me that biology can mean building microscopes as much as mixing solutions.

I would also like to thank my most esteemed colleagues, (soon-to-be Dr.) James Portman and (recently) Dr. Katerina Nitsenko. The three years we spent together were interesting, sometimes for the best, sometimes for the worst, but I am very glad that you have been there. I do not believe I would have finished this without your support.

And finally, I would like to thank my parents Jan and Marie, my brother Stephen, my sister Jane and my friends, especially Luc Lehéricy for his help with Markov modelling, Marc De Visme for his stimulating conversations on game design and Sarah Krauss, for being there for me with her army of plush toys when I was down.

Contents

List of Figures

List of Tables

Foreword

My first encounter with single-molecule biology was in my first year at the university. We were studying the workings of mitochondrial metabolism and were about to discuss the transmembrane ATPase and proton pump. This large protein uses the difference in pH between the inner chamber of the mitochondrion and the outer compartment to create ATP - chemical energy usable by the cell. Since it uses a flow of H^+ ions to generate energy, I imagined it to be the "turbine" of the cellular power generator that is the mitochondrion, and remember thinking it "would be cool if it spun". Going home in the metro, I was flipping through the class documents for the next session when I saw images by Hiroyuki Noji and Kazuhiko Kinosita paper showing, using an actin filament affixed onto the protein under an electron microscope, that it was, indeed, spinning. Not only that, but the authors managed to also determine the direction, speed, and even stoechiometry of the reaction simply by observing the discrete steps the rotor took - they concluded the rotor had 12 steps for a full rotation, which was validated later by structural analysis.
I was awestruck in this metro. And from that point I very often wondered why we can't manipulate molecules individually like we handle machine parts.
Single-molecule tools answer that yes, we can. Nanopore sequencing allows to lo look directly at individual DNA strands and sequences by "feeling" the shape, the distribution of charges on each individual base. Optical or acoustic tweezers can grab and pull microscopic handles to stretch molecules, in all three directions of space while we observe in real-time. Magnetic tweezers can manipulate 50 different molecules in identical ways to acquire statistically significant data on identical molecules and show us if they really are identical. Today, those techniques are put to use not only for the study of life, but for the design of smart materials, micro-medical devices, and so much more.

Project overview

I have joined Terence and Charlie's team to work on a novel molecular scaffold : jointed DNA, or jDNA, which was at the time being developed by Dorota Kostrz and Jing-Long Wang. I joined the project with Maryne Follenfant and Leah Friedman.

My objective was to build upon preliminary results obtained by Mrs. Friedman on a scaffold where two molecular partners - FKBP12 and mTOR - interact with a third, rapamycin, in solution. That model proved complex to study. Since our aim was at first to show how jDNA could be used for analytical chemistry, we decided to move towards a simpler system : the FKBP12 ligands. We were in this greatly assisted by Dr. Christian Stephan Meyners, who collaborated with us and provide us with a series of ligands for FKBP12. We used those as a simple test-case for jDNA. First, we showed that we could determine the interaction rates between FKBP12 and a reference ligand, JK313, using competition assays. Then, we used those results to measure the interactions between FKBP12 and the other ligands. In addition, we also investigated the role of force in our measurements and showed that we can model the behavior of jDNA using common physical models.

Dissertation overview

The objective of this thesis is to showcase how the technology of jDNA can advance the fields of molecular biology, protein-ligand interaction analytical chemistry and drug discovery through the example of the study of FKBP12 competitive inhibitors. Therefore, the background section starts with the FKBP protein family, centering on FKBP12, and on their pharmaceutical importance. This shows us the need for precise measurements of drug-target interactions, and leads us to a general presentation of the different techniques that are used today to probe the interactions between proteins and their targets. From macroscopic measurements, we progressively reach single-molecule techniques and, from there, connecting the technique with the molecule, we present the DNA molecular scaffolds, their physical and chemical properties, and show why jDNA is an important step forwards in this field. We also show the different ways jDNA and similar constructions have already been used.
This will lead us to the presentation of the tools used in this work : the preparation of the jDNA scaffold and the attached proteins, the magnetic tweezers and the Hidden Markov Model analysis tools. Putting it all together, we present low-force molecular competition experiments on the jDNA scaffold. Finally, in the results section, we present the theoretical framework we developed for the interpretation of our results and the experimental results themselves. Moving to answer questions raised by those results, we analyse the effects of magnetic force on the behavior of the scaffold and provide additional perspectives on the mechanics of jDNA.
The dissertation is ended with a general discussion, which will allow us to consider the general implications of the technique and suggest how it can be positioned in the field of drug discovery and analytical chemistry.

Part I

Introduction

Chapter 1

FKBP12 : A model system

The physiological role of FKBP12 is the motivation for this work. A brief summary of the protein family and history is presented here. The practical and historic reasons for our choice of this model system will be detailed at the end of the chapter.

1.1 Rapamycin, FKBP12 and the FKBP family

1.1.1 Rapamycin

Rapamycin was firstly identified as an antifungal agent ; it was later discovered to have antitumoral properties in humans (*1*, *2*). Rapamycin strongly suppresses Interleukin 2 (IL-2) signalling (*3*) and thus leads to immunosuppression by reducing T cell proliferation. Therefore, it is a prime immunosuppressant, and used today in post-transplantation cases to prevent allograft rejection, sometimes combined with cyclosporin. It has the advantages of having a low renal toxicity and rejection rate - particularly important for renal transplantation. Nevertheless, cross-talk with other proteins, of the FKBP family in particular, can cause side-effects (*1*, *4*).

1.1.2 FKBP12 : receptor of Rapamycin

Rapamycin was shown to have a physiological interaction with immunosuppressant Tacrolimus (FK506) to suppress cell growth and proliferation. It was then revealed it was then revealed that rapamycin and FK506 are both interaction partners of a single protein, (*1*) FKBP12. FKBP12 (FK506-binding protein 12) was thus historically the first member of the family, identified as a receptor for the immunosuppressant drug Tacrolimus, or FK506 (*5*, *6*). FKBP12 is a protein with a molecular weight of 12 kDa. Like all FKBPs it contains the core peptidylprolyl cis/trans-isomerase (PPIase) activity domain required for signalling (*7*).
FKBP12 was first discovered during the investigation of the antifungal agent ra-

pamycin, which has antitumoral properties in humans. FKBP12 binds to rapamycin with an equilibrium dissociation constant of 0.2 nM ; a very strong affinity.

1.1.3 FKBPs in-vivo

The FKBPs were identified by structural and genetic analogy to FKBP12. They are a subclass of the immunophilin family and are present in all eukaryotes from yeast to Humans and ubiquitously expressed. They possess a cis-trans peptidyl-prolyl isomerase (PPIase) activity. Despite the name, all FKBP proteins do not bind FK506 or do so poorly, FKBP12 being the best binder of FK506. The family name comes from the homology of their binding sites to the binding site of FKBP12.
FKBPs have diverse but critical roles in cellular metabolic and immune regulations (FKBP13, for instance, is involved in controlling Immunoglobulin A secretion). Mice knockouts exhibit various defects - FKBP12 knockouts are deadly - but not all their effects are fully understood (*8*), owning specifically to the fact that the FKBP signalling networks are strongly interconnected.

1.1.4 Pharmaceutical significance of FKBPs

FKBP51 and 52 are involved in the regulation of steroid hormone receptors - specifically, FKBP51 inhibits their expression while FKBP52 enhances the activity of glucocortinoïd receptors. While FKBP52 mice knockouts have serious developmental defects, FKBP51 depletion improves stress hormone signalling in mouse models of depression. In humans, FKBP51 polymorphisms have been associated with stress-related psychiatric disorders (*9*). FKBPs are abundantly present in neuronal tissues and their expressions are elevated after nerve injury. FKBP ligands can cross the blood-brain barrier more readily than other neurotrophins. Artificial, non-immunosuppressive ligands to FKBP12, have shown immunoprotective and neuroregenerative effects on rodent models of Parkinson's disease (*7*).

1.2 The mTOR signalling pathways and their inhibition by FKBP12

1.2.1 mTOR, a metabolic control nodal point

mTOR is a large ($\approx$ 290 kDa) protein. It is a phosphoinositide-3-kinase (PI3K) that plays a vital role in energy and nutrient-level detection, and in responses to growth factors in particular (see Fig. 1.2). In metabolically significant organs such

as the liver, it interacts with insulin signalling. mTOR can bind to protein RAPTOR (Regulatory Associated Protein of TOR) and mL-ST8to create mTORC1 (*10*), or to RICTOR (rapamycin-insensitive companion of mammalian target of rapamycin) and SIN1 (also called Target of rapamycin complex 2 subunit) to form mTORC2. In both cases, it serves as the enzymatic core of the complex.

mTORC1

mTORC1 promotes adiposity and lipid accumulation [1]. mTORC1 promotes anabolic metabolism and leads to protein synthesis, ribosome biogenesis, nutrient transport, lipid synthesis and other processes in response to nutrients, growth factors and cellular energy(*1*, *11*). Cells with a high mTORC1 activity are significantly larger in size. Some studies even suggest that, as long-term inhibition of mTOR signalling mimicks dietary restriction effects on lifespan, inhibition of mTOR signalling could be a life-extending treatment for mammalians, especially humans ; mice studies seem to give evidence in this direction (*4*).
Overall, mTORC1 is a driver of the general metabolic activity of a cell, activated in situations or cellular pathways associated with high metabolic activity, promoting cell growth while its inhibition promotes autophagy and inhibits growth (*12*).

mTORC2

mTORC2 is activated by growth factors, in particular through the insulin/PI3K signaling pathway. The major substrates of mTORC2 are AGC-kinase family members. This phosphorylation activates Akt/PKB, where deregulation of Akt/PKB has been implicated in cancer and diabetes (*13*). RICTOR and mTORC2 have been shown to play an essential role in embryonic growth and development, perhaps due to the control that mTORC2 exerts on actin cytoskeleton organization (*14*).

1.2.2 mTORC1 inhibition by FKBP12 and rapamycin

The FKBP12-Rapamycin complex interacts with the FRB domain of mTOR. This interaction inhibits the functions of mTOR in the mTROC1 complex ; mTORC2, on the other hand, is rapamycin-insensitive [2]. The interaction is shown on Fig. 1.1. With physiological concentrations of FKBP12 and mTOR, rapamycin has a high selectivity for mTOR and is an effective inhibitor in the nanomolar range (*1*).

[1] This is linked to obesity and, if insulin responsiveness is lost, type 2 diabetes

[2] This is due to steric constraints that prevent the binding of FKBP12 to mTOR in mTORC2 (*13*)

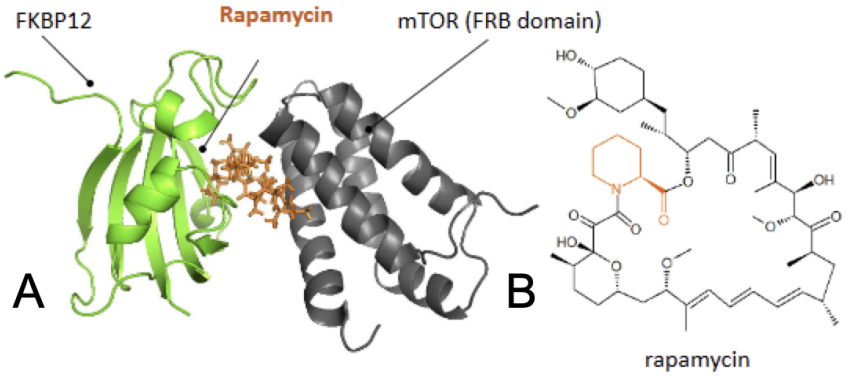


Figure 1.1: (A) crystal structures of of proteins FKBP12 and the FKBP12 - Rapamycin binding domain of mTOR, interacting with rapamycin. (B) Molecular structure of Rapamycin. The amine cycle in orange is essential for recognition by FKBP12. Adapted from Ref. (*18*)

Rapamycin and rapamycin analogues usage as anti-oncologues

The mTOR pathway is often activated in cancer. Rapamycin analogues (sometimes called "rapalogs") - temsirolimus or everolimus for example - have been developed and marketed as anti-oncogenes, but their benefits have been limited : it seems that only some pathways are inhibited effectively. In addition, negative feedback loops partially compensate the inhibitive effects, sometimes even promoting cancerous cell survival. For example, inhibition of mTORC1 promotes autophagy, increasing cell metabolism and increasing cell survival in hypoxic and poorly vascularized environments (frequent in cancer) (*15*).
Rapamycin is was mildly successful on patients suffering from renal cell carcinoma, even if they fail previous treatments (*16*). Current investigations aim at combining rapalgs with other pathway inhibitors to suppress all relevant pathways and increase their efficacy as anticancer drugs (*17*).

Because of their regulatory roles on the essential mTOR pathways, FKBPs, and FKBP12 in particular, are interesting targets for therapeutics.

1.2.3 Ligand development for FKBPs

Since the different FKBP proteins have similar FK binding domains, and because those domains are critical to their function, drugs targeting those metabolic and immune controllers must exhibit very strong specificity to be effective and avoid dangerous side-effects (*8*). With the increasing knowledge of the structure of the binding sites of the FKBP proteins, the development of synthetic ligands is today made easier.
For FKBP12 ligands, bridged bicyclic structures were shown to be effective partners

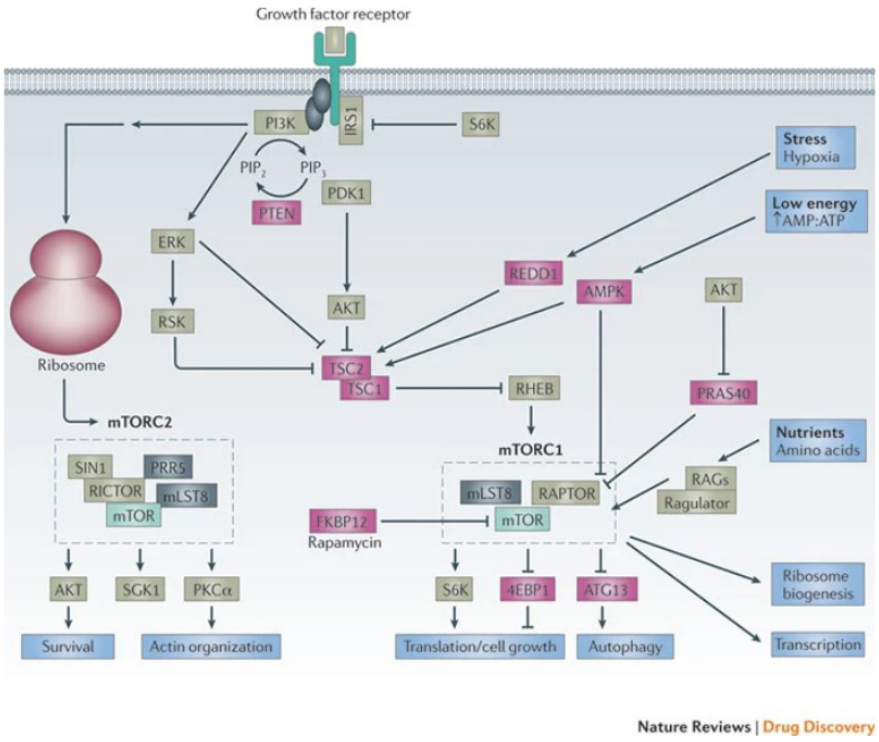


Figure 1.2: Simplified scheme of the mTOR signalling network (*1*)

since 1995 (*19*) and new compounds are improving on the core two-cycle model, always incorporating the motif shown in Fig. 1.3 (*20*, *21*). The recent work of the team of F. Hausch (*20*, *21*) has led to the synthesis of a series of artificial organic ligands to FKBP12, presented on Fig. 1.4.
In cooperation with these medicinal chemists, we have decided to use this system as a model binding reaction for the junctured DNA assay. Pomplun et al. measured the binding dynamics of the ligands with their targets using fluorescence polarization (cf. Section 3.1.1) ; this well-established method is a perfect point of comparison with our jDNA results.

Our team has already worked on the FKBP12-rapamycin-FRB complex before, and I have begun work on it in the first steps of my thesis. However, the three component-nature of the system made all calculations more difficult and made it harder to extract relevant information form the experiments. The collaboration with Hausch et al. allows us to move to a chemically simpler configuration.
Therefore, we will be using the FKBP12 model system to illustrate the use of junctured DNA in chemistry and drug design. To present further background in this field, the following section will cover basic principles in protein-ligand interactions to set the stage for the presentation of the devices and experimental methodologies that are used to measure those interactions.

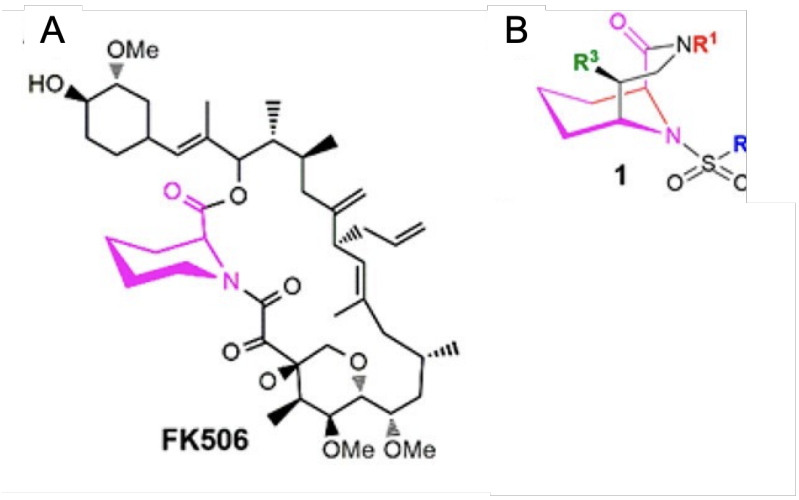


Figure 1.3: (A) Structure of FK506, molecular ligand of FKBP12. The pink motif indicates the core cycle, essential for binding. (B) Core motif of bicyclic compounds (mimicking the pink cycle in A) used in this work. Adapted from Ref. (*20*).

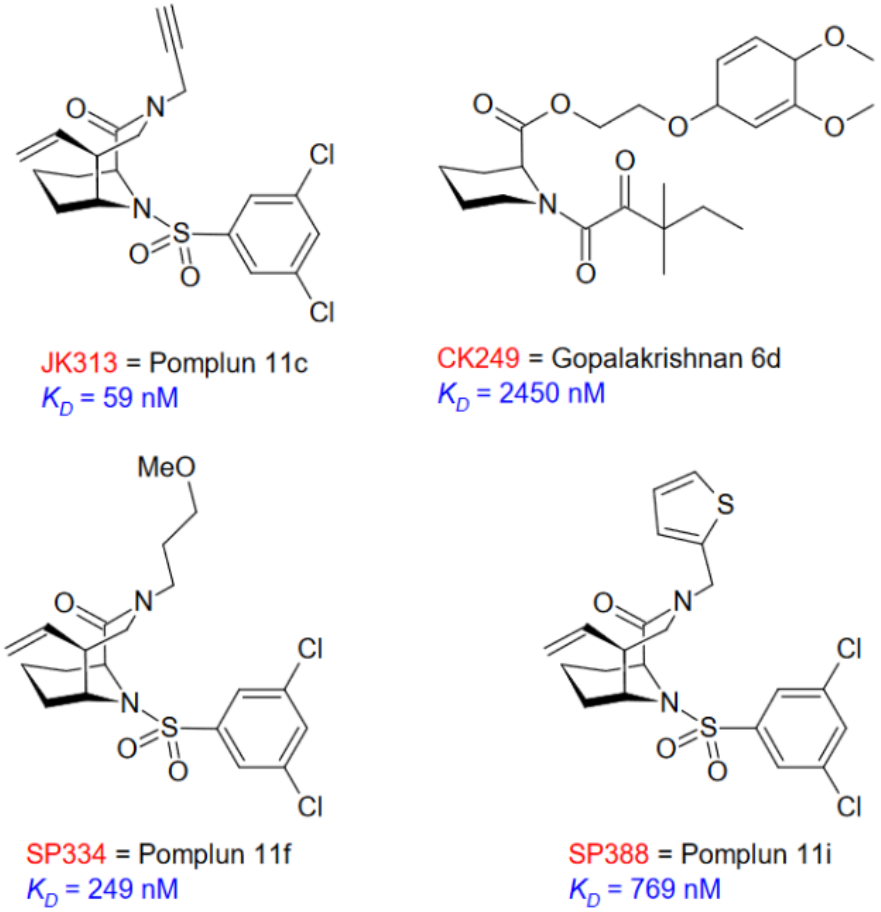


Figure 1.4: The four synthetic agonists of FKBP12 analyzed as model ligands in this work. Dissociation constants in blue were obtained using fluorescence polarization. Adapted from Ref. (*20*, *21*) and personal communication from the authors.

Chapter 2

Principles of protein-target interactions

2.1 General considerations

2.1.1 Macromolecular interactions

A macromolecule interacts with other molecules via *weak* - that is non-covalent - interactions. Those include ionic interactions, hydrogen bonds, and Van der Waals interactions. They exert at short distances (one nm and shorter) with weak energies comparatively to covalent bonds (on the order of tens of kJ/mol per interaction for ionic bonds, 10-40 kJ/mol for hydrogen bonds and 1 kJ/mol for Van der Waals bonds (*22*)). Those low energies allow for reversible binding between proteins and other macromolecules, which are required for all the processes that make up life.

Determining the interaction properties of macromoelcules is a major focus of biochemical and biophysical research, especially through the development of dedicated instruments. Titration experiments have been made during the 19 th century while fast mixing experiments have been performed in the first decade of the 20^{th} century. The temperature jump methods has been established in 1963 (*23–25*) and has contributed greatly to this field. Multiple other methodologies have since been devised to measure both equilibrium and transition states, based on different chemical and physical principles. The aim of this section is to provide an overview of those methods, to highlight their strengths and weaknesses and how single-molecule (SM) molecular interaction assays may play a role in the field.

2.1.2 Kinetic and thermodynamic descriptions

Chemical reactions can be described *kinetically* and *thermodynamically*. Kinetic descriptions (sometimes called rate descriptions) take into account the temporal evolution of a reaction, while the thermodynamic description only refers to systems at equilibrium, that is with no macroscopic evolution. Thermodynamic parameters are typically easier to measure, as they do not require measurements over time. They can also be deduced from kinetic parameters, as they correspond to the limit of the chemical reaction when time goes to infinity.
On the other hand, *kinetic* or *rate* constants describe the temporal evolution of a system, and its behavior out of equilibrium. Those parameters are richer, but harder to measure reliably.

2.1.3 Formalism of interactions

Let us examine the simplest molecular interaction : a protein P interacts reversibly with a ligand L - in our case, P would be FKBP12 and L a potential inhibitor. The properties of this interactions might shed additional light on the biological regulatory pathways controlled by this receptor ; in addition, if it is required, for a treatment for instance, to block this pathway, it is useful to both know the interaction physiologically and to test any putative drugs to ensure that, at the concentrations expected *in-vivo*, they will be able to block the interaction.

Reactions will be conventionally written as dissociations :

$$\mathrm{PL} \underset{\mathrm{k_+}}{\overset{\mathrm{k_-}}{\rightleftharpoons}} \mathrm{P} + \mathrm{L} \tag{2.1}$$

Where k_+[nM $^{-1}$s $^{-1}$] and k_-[s $^{-1}$] are respectively the association and dissociation rate constants. The equilibrium dissociation constant is $K = \frac{k_-}{k_+}$; this thermodynamic *dissociation* constant can also be expressed at equilibrium :

$$K = \frac{P_{eq}L_{eq}}{PL_{eq}} \tag{2.2}$$

P_{eq} is the concentration of P (L and PL respectively) at equilibrium. From a macroscopic point of view, the evolution of this simple system is also constrained by the conservation of matter :

$$P + PL = P_0$$

$$L + PL = L_0$$

With P_0 and L_0 being the initial concentrations of P and L. In our case, for simple association and dissociation kinetics, we have a *first-order* kinetic for both molecules. Thus, we can write :

$$\frac{dPL}{dt} = k_+ L \times P - k_- PL \tag{2.3}$$

Which can be solved as an exponential convergence. In experimental conditions, with receptors on a surface or a scaffold, the amount of ligand L is largely superior to the amount of protein P, so that $L \approx L_0$. The equation can then be solved as (with initial condition of no interactions at t = 0) :

$$PL(t) = \frac{k_+ L_0 P_0}{k_+ L_0 + k_-}(1 - e^{-t(k_+ L_0 + k_-)}) \tag{2.4}$$

2.1.4 Timescales of interaction

Macroscopic time-scales

From this resolution, we can see that the characteristic timescale here is $\tau = \frac{1}{k_+ L_0 + k_-}$. On a macroscopic level, τ is the order of time that the system will take to reach equilibrium if perturbed. When designing or choosing kinetic measurement methods, we will need to compare the temporal resolution of the method with this value. To prevent under-sampling effects, the temporal resolution needs to be at least one order of magnitude shorter than the characteristic timescale of the observed reaction.

Single-molecule timescales

If a molecule is observed in isolation, two time-constants emerge, one for the association (τ_P) and one for the dissociation ($\tau_P L$). Those time constants they no longer depend on the concentration of protein (since we consider there is only one molecule reacting) but k_+ still depends on the concentration of ligand in the medium. $k_+ L$ and k_- are homogeneous to the inverse of a time, and can be interpreted as the average duration of a molecular state (or dwell-time of P in that state). $\tau_{PL} = \frac{1}{k_-}$ is the average duration of the protein-ligand complex and $\tau_P = \frac{1}{[L_0]k_+}$ corresponds to the average duration of the non-interacting state of the protein.

Exponential distributions

Since the dwell-times in first-order kinetics are distributed according to an exponential distribution, the average dwell-time is also the inverse of the exponential decay

constant. If $f(t)$ is the probability distribution density of a state lasting t, we have :

$$f(t) = A.e^{\frac{-t}{\tau}} \tag{2.5}$$

With A a normalization constant.

Also notice that τ_{PL} is a constant - indeed, it only depends on the stability of the complex, not on any other factors such as the presence of ligand. Increasing the ligand concentration will only mean that once the association is broken (through thermal fluctuation), the complex will be reformed sooner.

Diffusion limit

The diffusion limit for association rate constants in water is estimated to be $10^9 s^{-1}.M^{-1}$. This represents the rate at which the molecules can diffuse in the solvant and collide with our protein, and is thus an upper limit to the value of k_+ (*26*). If we consider a complex with a typical dwell-time of 1s, we arrive at a dissociation equilibrium K = 1nM , so the nanomolar range would be a ballpark estimate for the lowest dissociation constants (ie. the molecules with the highest affinity) with interactions lasting for about a second. A complex with a lifetime of 1000 s. would be capped at a K of 1 pM . This is true in solution, but diffusion guides, such as surfaces can facilitate protein and ligand encounters and help circumvent this limit (*26*).

2.2 Objectives in drug discovery

With those considerations on how proteins bind to their targets, let us look back at the problem of drug development. Drugs are generally molecules that bind to an *in vivo* target, e.g. a receptor, an enzyme or any other physiological component in the target cell, to either activate it, or prevent its activity. Most are xenobiotics (not produced organically) of low molecular weight. To have an effect, the molecule must first get to its target, and then stay in its presence for long enough to have a chance to bind for a significant time. Understanding how they bind *in-vitro* is useful, but only if we understand how their *in-vitro* behavior affects their *in-vivo* efficacy. What type of measurement is needed for drug design ?

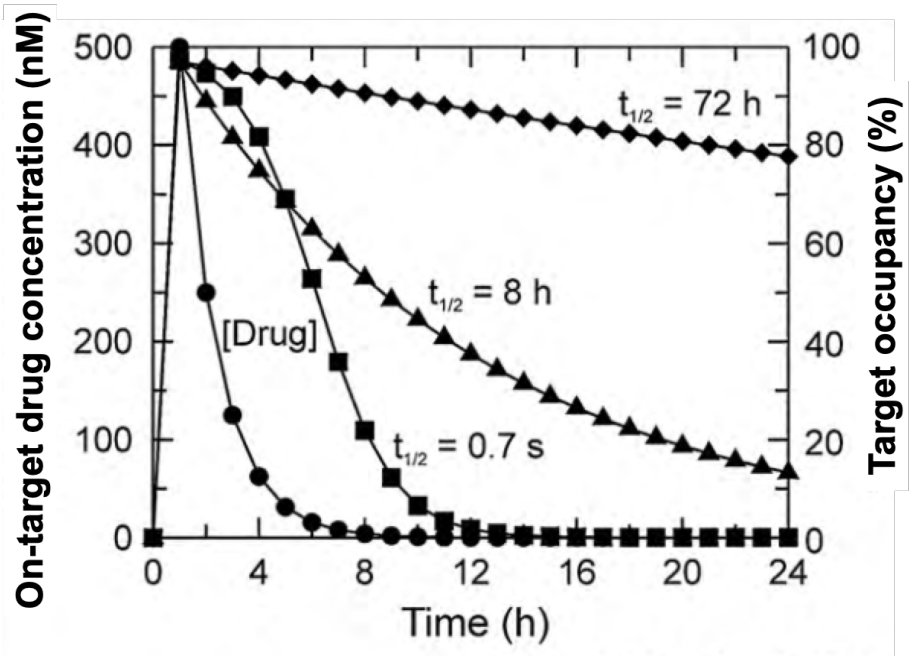


Figure 2.1: Residence time, pharmacokinetics and pharmacodynamics. An analysis that demonstrates how residence time affects the amount of drug–target complex (pharmacodynamics) as a function of time. The fractional occupancy of the target by the drug, given as a percentage, is shown for three drugs all of which have equilibrium dissociation constants of 14 nM for the final drug–target complex. $t_{1/2}$ is the average lifetime of the drug-target complex. Adapted from (*27*)

2.2.1 In-vitro measurements

In-vitro measurements are often made on a closed system : there is no inflow of outflow of components. Thus, the effects of a particular molecule can be effectively summarized by experiments at equilibrium. One can measure K, the equilibrium constant, or the concentration of drug required to get a 50 % decrease in activity of the target, IC_{50}.

2.2.2 In-vivo conditions

In-vivo however, the drug must not only bind to its molecular target, it must also bind quickly enough to have an effect before it is removed by blood circulation, destroyed by the renal system, oxidized or suffered other fates foreign organic compounds suffer in living bodies.
In addition, after a medication intake, the ambient drug concentrations will fall quickly. If the drug can remain bound for long periods of time it can outlast by far the drug molecules present in circulation.

Figure 2.1 shows the hypothetical case of multiple drugs binding to the same

molecular target with the same K, but different residence times. The figure illustrates how quickly the on-target concentration decreases. On that figure, the medical effectiveness of the drug can be seen as the integral over time of the target occupancy.

2.2.3 Goals of drug design

In multiple cases of clinical study, the lifetime of the drug-target complex correlates with the efficacy of the drug, especially when comparing different compounds from the same family(*27*). Therefore, it appears favorable to design drugs with low k_- values. However, side-effects follow the same pattern, and knowing the dissociation rate with undesirable targets can guide the dosage and frequency of uptake, for example, to minimize secondary effects.
For example, anti-inflammatory agents ibuprofen and naproxen are close analogs of acetylsalicylic acid, but their binding to target is rapid and quickly reversible. Aspirin, on the other hand, binds irreversibly with long-lasting effects, causing the known side-effect of increased risk of bleeding when taking aspirin (*28*).
Modern drug-design must thus aim to determine the kinetics of drug-target interactions, as they afford a better understanding of hwo the drug affects the organism.

2.3 Summary : Objectives in protein-ligand interaction measurement

Three elements are required to measure the interaction between a protein and a ligand :

Preparation of proteins This is usually done in transgenic bacteria. Then, the concentration of the proteins must be determined - various well-established methods are used, such as mass spectroscopy, chromatography and Bradford assays. This is not be the topic of this work.

An Observable - A way to determine what fraction of proteins are in protein-ligand complexes. In macroscopic terms, this means determining the concentration of protein-ligand complexes. In single-molecule terms, this means being able to distinguish between the two molecular conformations. This must be linked to an **observable** parameter, that can be measured during the experiment.

With this information alone, we can already determine the equilibrium constant K of the reaction, since we know the concentrations of both molecules introduced. Thus, this is enough for *thermodynamic* studies.

A Perturbation - A way to synchronize the association or dissociation reaction. In macroscopic studies, it means making certain that all molecules, diffusing freely in the medium, will begin interacting with their partners at the same time, under the same conditions. Some method must thus be found to perturb the equilibrium of the reaction, either to force the reaction to start or to simply change the position of the equilibrium consequently to an external impulse. The time-resolution of the observable relative to the interaction rate constants must also be considered.

Chapter 3

Classical ensemble techniques for equilibrium and rate constant measurements

This chapter will describe some of the most used methods currently available for the determination of kinetic and thermodynamic interaction constants of biological macromolecules.
We will treat ensemble and single-molecule assays separately, since they are fundamentally different. We will begin with the ensemble techniques. First, we will go over the different *observables* that are being used to measure interactions. The time constants that can be probed and reached will be an important consideration. Secondly, we will go over the different means of creating a *perturbation* in a chemical system for analysis - how to break the equilibrium to gain access to rate constants and kinetic data. This will lead us to the single-molecule experiments.
For every technique, we will also address the technical issues that arise from the methodology.

3.1 Macroscopic observables

Multiple very simple methods can be used to determine whether a protein and a ligand interact in a given environment, and to quantify the fraction of proteins in interaction. They need to be combined with a perturbation method to measure kinetic rates.

3.1.1 Molecular mass as an observable

Two molecules in complex together are more massive than each of the individual components. This gain in mass by interaction can be detected and measured with different techniques - some very simple, other more complex.

Gel Chromatography

A simple way to assay interactions between two proteins is chromatography : since migration speed is affected by molecular mass, the relative percentage of bound versus unbound protein can be estimated. This is a bulk method that relies on the assumption that migration conditions will not break nor affect the interaction between the partners : thus, it is only reliable for stable, long-term interactions. With shorter interactions, a phenomenon known as peak broadening occurs, whereupon some binding and unbinding occurs in the chromatography column itself and makes interpretation harder. Finding good conditions of flow, temperature and buffer is a difficult task in and of itself, making this technique simple in theory, but sometimes very hard to implement.
The quantification of the bound versus unbound fraction is another important factor of error. Various techniques may be used, such as scintillation spectrometry as in Ref. (*29*), which counts the radioactivity of a sample - this requires, of course, that one of the partners be radioactive, usually via $[H^3]$ labelling. Antibodies (fluorescent for instance) may also be used for detection. Also note that this technique only requires tools common in most biochemistry laboratories and may be used for quick estimations before using more complex methods.

Surface Plasmon Resonance

Surface plasmon resonance (SPR), is a widely-used technique based on the coupling of electromagnetic resonance frequencies of a metal dielectric surface, such as gold or silver, with adsorbed components on the surface such as antibodies able to capture soluble proteins. The amount of organic matter close to the surface - the change in mass of the bound complexes - slightly modifies the refractive index. This change can be detected by shining a laser beam on the surface and observing either the change in reflected angle or wavelength. The principle is illustrated on Fig. 3.1.
Typically, a standard SPR sensor can detect changes of $1.0 \cdot 10^{-5}$ to $1.0 \cdot 10^{-6}$ of units in refractive index, which can lead to changes of $\approx 1\ pg.mm^2$ (*30*), corresponding to typically hundreds of proteins being adsorbed, thus this technique requires a significant amount of materiel and is definitely not suitable for single-molecule analysis.

To bind one of the partners to the surface, various methods can be employed.

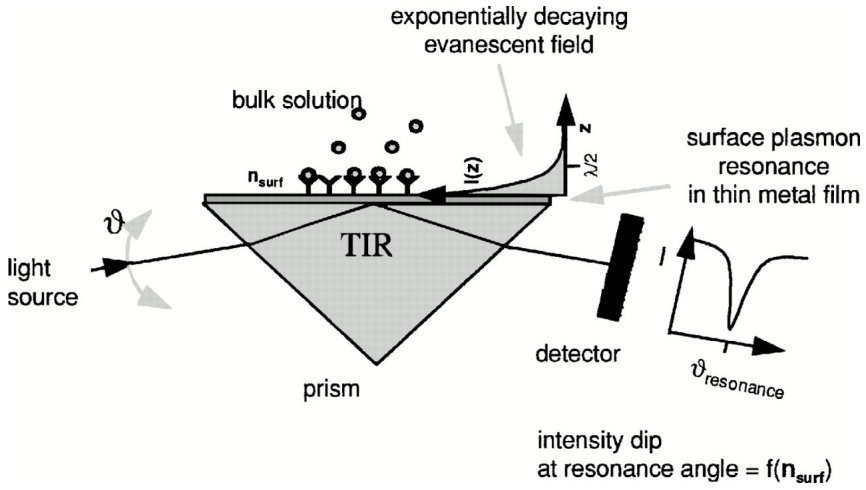


Figure 3.1: Principle of a surface plasmon resonance biosensor. Adapted from Ref. (*31*)

The first most popular is the coating of the metallic surface with alkyl thiols which renders the surface hydrophobic and suppresses non-specific interactions, then with a 200 μm layer of hydrophilic dextran matrix (*32*), and then using N-Hydroxysuccinamide to functionalize the dextran surface, leading to the formation of ester bonds with protein lysin groups (*33*). Thus method is simple, but can result in the denaturation of proteins and different proteins may be immobilized differently. Alternative methods, such as the use of biotinylated macromolecules and avidin- or streptavidin-coated surfaces, can be preferred and also used for the immobilization of DNA, RNA or vesicles. Antibodies may also be immobilized on the surface to bind specific proteins.

The response of an SPR biosensor being proportional to the mass of proteins on the surface (*30*, *31*), equilibrium constants of the interaction can be determined simply ; more complex interactions may be probed, such as three-partner interactions, by progressively saturating the buffer with different components, so that only one varies at a time (*34*). SPR measurements allow for the exploration of many equilibrium constants, but also suffer from limitations : in many cases, a single-step reaction as described above may seem to follow a two-step exponential temporal reaction, because of different sub-populations of bound proteins (*31*), which often requires doing experiments with different immobilisation techniques. The diffusion of soluble proteins to the surface may also be found to be limiting (association and dissociation kinetics seem single-exponential at low concentrations, but double-exponential at high concentrations) especially in immobilization matrices. Again, this often imposes the realization of multiple control experiments under different conditions to exclude the possibility of those artefacts.

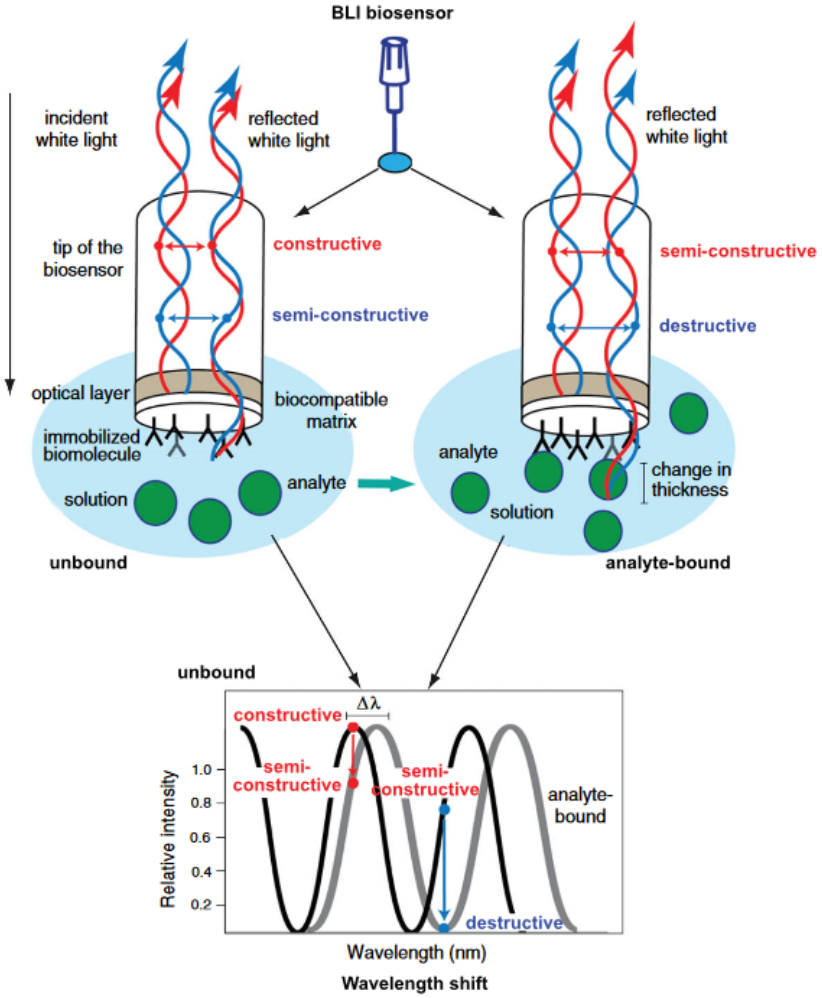


Figure 3.2: Principle of biolayer interference (BLI) interaction kinetic determination method. Adapted from Ref. (*35*)

The changes in electromagnetic resonance in the surface occur on a much faster timescale than the binding and unbinding of molecular partners (*31*), SPR can be used for kinetic experiments. For instance, the dissociation constant of a complex can be measured by binding one of the partners to an SPR-friendly surface, then washing this surface with buffer : the refractive index progressively comes back to empty, following an exponential decay which can be fitted to determine the dissociation constant (*34*). Unfortunately, this remains tied to the ability of the system to reach equilibrium (*31*) : especially when working with weak concentrations, the diffusion of the soluble molecule to the SPR surface may prove to be a limiting step, as we will discuss in Section 3.2.1.

Biolayer Interference

Relying on similar principles as SPR, biolayer interference (BLI) is a recently-developed technique (*35*, *36*) in which one of two molecular partners is adsorbed on

the functionalised surface of a fiber-optic rod. The rod is then dipped in a solution containing the other partner. A laser beam is lit through the rod ; a fraction of the light will reflect at the end of the rod while another will reach the polymer matrix on the tip and be reflected upon it. Both return beams then interfere. When interacting molecules attach to the tip, they extend very slightly the tip's light-conductive area, thus increasing the equivalent length of the path of light of the beam that reflects on the matrix, causing a shift in the interference pattern - see Figure 3.2
As the binding reaction occurs at the tip of an optical fiber, the surface to coat is smaller and thus requires less materiel and is subject to shorter diffusion times, reducing the uncertainty in determination of diffusion times. The method also lends itself very well to parallel experiments, with commercial devices existing with 96 fiber optic rods that can sample standard 96 well-plates simultaneously, using the same buffers and incubation times and thus reducing variability ; controls can also be made alongside experiments. The optics required are less complex and less expensive than in SPR for example, facilitating deployment.

Fluorescence polarization

A fluorophore is a molecule that absorbs incoming light in a certain frequency range, moving into an excited state. Then, after a short duration(measured in nanoseconds) , it emits a less-energetic photon. The fluorophore it is still subject to Brownian motion during this time : it diffuses and rotates. This rotation matters when the polarization of fluorescence is detected : the correlation between excitation and emission increases with the molecular mass. The rotation correlation time is $\theta = \frac{\eta V}{k_B T}$, where η is the kinetic viscosity and V the molecular volume - a quick estimation leads to a correlation time on the order of the nanosecond, much shorter than any biological interaction times (most often superior to the ms) (*37*).
When a large population of fluorescent molecules is exposed to polarized stimulation, the polarization level of the output gives an indication of the mass of the fluorescent molecules (*38*). One can observe a solution of fluorescent proteins and gradually increase the concentration of partner molecules, and see the increase of the anisotropy in the emitted fluorescent light.

On Fig. 3.3 is depicted a typical titration curve for the interaction of the FK1 binding-domain of FKBP51 in a fluorescence polarization assay with soluble molecular partners (here synthetic partner 1d). With the increasing concentration of partner, the EC50 - the concentration of FKBP51 required so that half FKBP51 molecules be bound - increases, and a linear interpolation may be drawn (close-up) ; the slope of the line is the equilibrium constant. This exact method was used

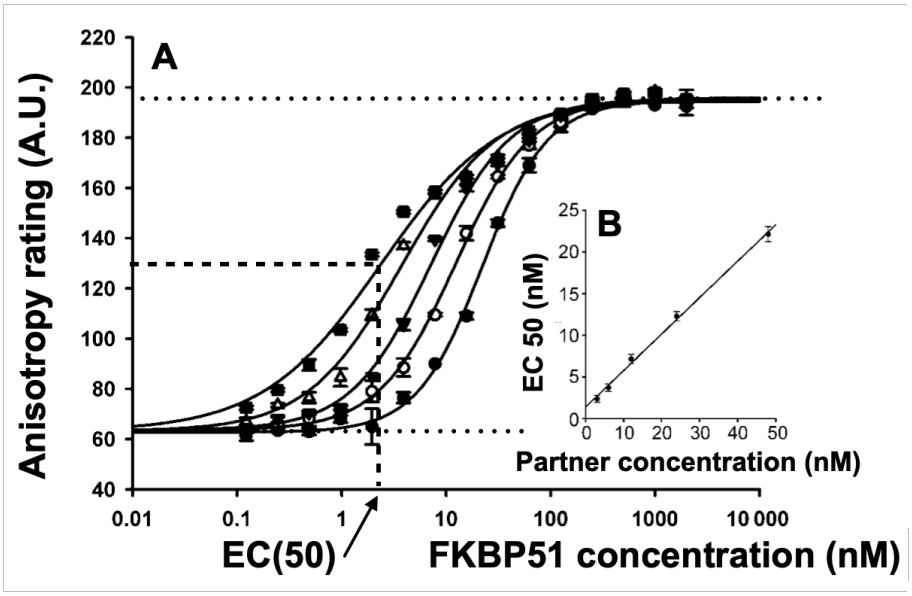


Figure 3.3: (A) Titration of partner 1d by FKBP51 observed with fluorescence polarization. The successive curves correspond to increasing partner concentrations (3 nM to 48 nM). For each partner concentration, the EC50 corresponds to the concentration of FKBP51 required for attaining half the maximum increase of anisotropy. (B) Linear increase of EC50 with 1d partner concentration. The slope of this line corresponds to the equilibrium constant.

by Pomplun et al. to determine the equilibrium constant of the FKBP12 ligands studied in this work (*20*, *21*).

With this method, it is possible to measure the interactions of many molecular partners, but with a significant number of limitations. The method only works if the fluorescent molecule is smaller that the partner, as its angular diffusion rate must be significantly altered for a change to be detected. It requires a large amount of fluorescently labelled molecules and is an *equilibrium* method as the decorrelation happens much faster than the interactions being probed.

3.1.2 Thermal energy as an observable

A widely used family of methods, especially in pharmaceutical industry are the calorimetric titrations, and especially isothermal titration calorimetry (ITC). This well-established method uses a calorimeter to measure the heat output of a reaction, in order to determine the enthalpic energy variation and stoechiometry.

An ITC instrument is made of two cells (coated with a strong heat-dispersing material, such as copper or gold) separated from each another and from the environment by an adiabatic shield ; thermal sensors and an electrical heating system, such as a Peltier module, is implemented in the sample cell - see Fig.3.4. The reaction happens in the sample cell, emitting (if the reaction is exothermic) or absorbing (if the reaction is endothermic) heat. A feed-back system maintains the temperature

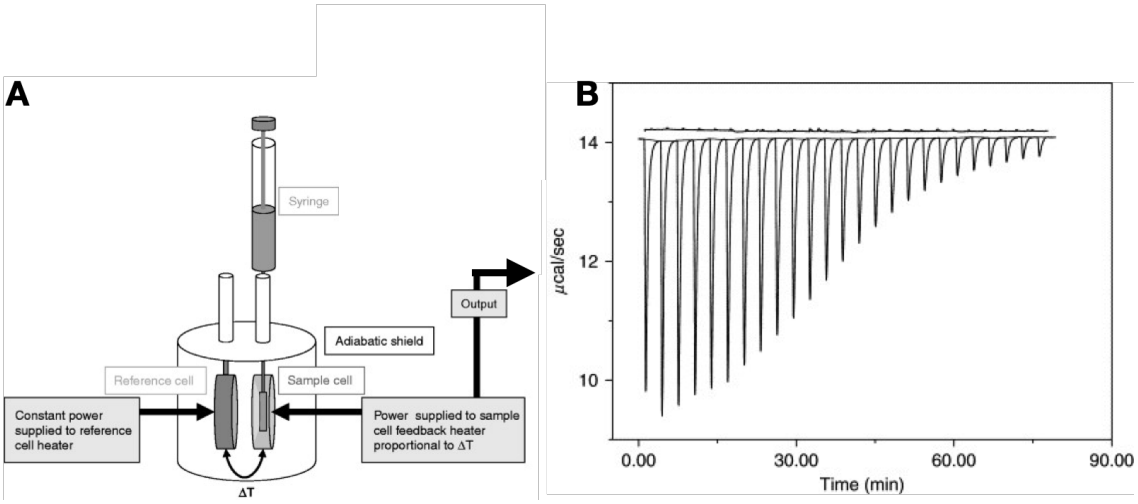


Figure 3.4: (A) Schematic representation of an ITC measurement device (B) Sample result trace from an ITC measurement. Each spike represents the energy required to keep the sample cell at thermal equilibrium with the reference cell after injection of a constant volume of titrant. Notice that the energy required is reduced as the titration progresses, showing the reduction in relative entropy gain. Adapted from (*39*)

constant in the cell. The power required to compensate the ΔH change of the cell, measured as the current delivered to the Peltier module, is monitored. If the concentrations of reagents present in the cell are known, the molar change in enthalpy of the reaction can be determined.

A typical experiment is run by injecting small, controlled volumes of reagent into the sample chamber, and integrating the changes over those injections until the reaction is complete. Once the injection of additional reagent leads to no thermal change (after correcting for artifacts such as the mechanical energy of the mixing) the stoechiometry of the reaction can be determined, and the value of K deduced ; one can then obtain the free Gibbs energy change $\Delta G^0 = -RTln(K^0)$ and deduce the entropy change via the equation $\Delta G = \Delta H - T\Delta S$.

This method is widely used as it is very robust and can give access to not only the thermodynamic properties of a reaction, but also give insights into kinetics using the temperature-dependence of rate constants as shown by the Arrhenius equation (Equation 4.1) (*40*). With measurements that can be accurate up to 0.4 μJ equilibrium constants up to $10^8 M^{-1}$ can be determined, as well as reaction kinetic times of 2 ms to 20 s. This comes however at the price of many sources of possible artifacts, as any heat generation in the system can cause errors in measurements, including mixing of different buffers, bubbles appearing during mixing (requiring reagents to be de-gassed first, which can also lead to evaporation of solvent and changes of con-

centration), enthalpy change due to uptake or release of protons by buffers during a reaction. This requires many control experiments to be made with blank mixes of buffers, so that their effects may be subtracted from the raw data (*41*).

3.1.3 Chemical properties : a reaction as an observable

If the complex formed by two molecules has chemical activity, such as an enzymatic catalytic property, it is possible to test for interaction by testing for those properties. *Reaction observables* can be used if one of the partners to be assayed is an enzyme, or has detectable properties that are modified by the other partner. For example, the protein FKBP12 has rotamase activity, that is it can isomerase certain peptides (*29*, *42*). This isomerization can be tracked using spectroscopy, as both conformeres have different absorption spectra. Therefore, it is possible to measure the activity of the same amount of enzyme in presence of different concentrations of ligand, and deduce from there the concentration of ligand required to remove half of the native catalytic activity : assuming bound enzymes are inactive, this allows us to determine the interaction constants of the two partners.
Of course, this assumption of inactivity is a very strong one, and it can be difficult to justify. This method, whilst simple, only works if a suitably quantitative reaction can be found.

3.1.4 Spectroscopic observables

Fluorescence intensity

The most intuitive, and indeed one of the most popular, means of following protein-ligand interactions are optical methods. Light interacts with molecules in solution on the femtosecond timescale, and thus can probe reactions with kinetics close to the diffusion limit. Light is very convenient to use, and optics are readily available on the market.
Close proximity between molecules does not only lead to chemical interaction, but also to optical interactions. Under a given excitation signal, the emission intensity of a fluorophore is affected by the lifetime of the excited fluorophore as well as by the quantum yield, which is the ratio of light energy emitted to absorbed. Those values can be affected by the environment of the fluorophore ; for instance, close proximity of aromatic groups can quench part of the energy, or close molecules can provide non-radiative energy dissipation pathways, reducing yield. Thus, the intensity of a fluorescent probe - especially located close to a binding site - can be quantitatively affected by the presence of a ligand in solution. This type of technique

are nevertheless very empirical as the effect of the ligand on fluorescence cannot be predicted (*43*).
Tryptophane, one of the 20 constitutive amino acids in proteins, is also fluorescent and is the highest contributor of to the intrinsic fluorescence of proteins. This intrinsic fluorescence can be used in this way to probe protein interactions with other molecules. Intrinsic tryptophane fluorescence has the great advantage of using the native protein and removing the need for molecular biology to engineer fluorescent tags. The main drawback is the necessity for specialized microscopes (UV is blocked by certain optics, for instance) and the relative complexity of the analysis (*43*).

Förster resonance energy Transfer (FRET)

Another well-known phenomenon is Förster resonance energy transfer, or FRET : it is the modification of the emission spectrum of a fluorophore A (Acceptor) when in close proximity of another fluorophore D (Donor) emitting at the excitation frequency of A.
Bulk studies can easily be conducted by mixing probe-linked ligands and proteins and following the fluorescence of the mix ; for example, in Ref. (*44*), the authors probe the interaction of proteins FRB, FKBP12 and rapamycin using FRET ; mixing both FRB and FKBP12, they progressively increase the concentration of rapamycin and observe an increase in the fluorescent acceptor, and a corresponding decrease in the fluorescent emission of the donor, often quantified as the ratio R of donor / acceptor fluorescence for maximal contrast and normalization, as seen on Fig. 3.5. The changes of R are proportional to the amount of closely interacting proteins. This protocol has the advantage of allowing, with a single experimental setup, the probe of multiple concentrations of rapamycin and thus the determination of multiple equilibrium constants, as it reacts quickly to changes in concentration. In addition, tracking non-FRET fluorescence creates integral controls and references, making it easier to compare replication experiments (Vertical dotted lines on Fig. 3.5 B).

In-situ studies can also be made, using the fluorescence to probe interactions inside a cell, for instance, and explore compartimentalized reactions and the propagation of a signal in a cell. (*43*). It is a flexible method that builds upon biological materials and instruments that often exist, as fluorescently tagged proteins and FRET microscopes are relatively common nowadays.

Nuclear magnetic resonance

Nuclear magnetic resonance spectroscopy (NMR) is a technique known for its applications in structural determination of molecules. Very succinctly, NMR uses the

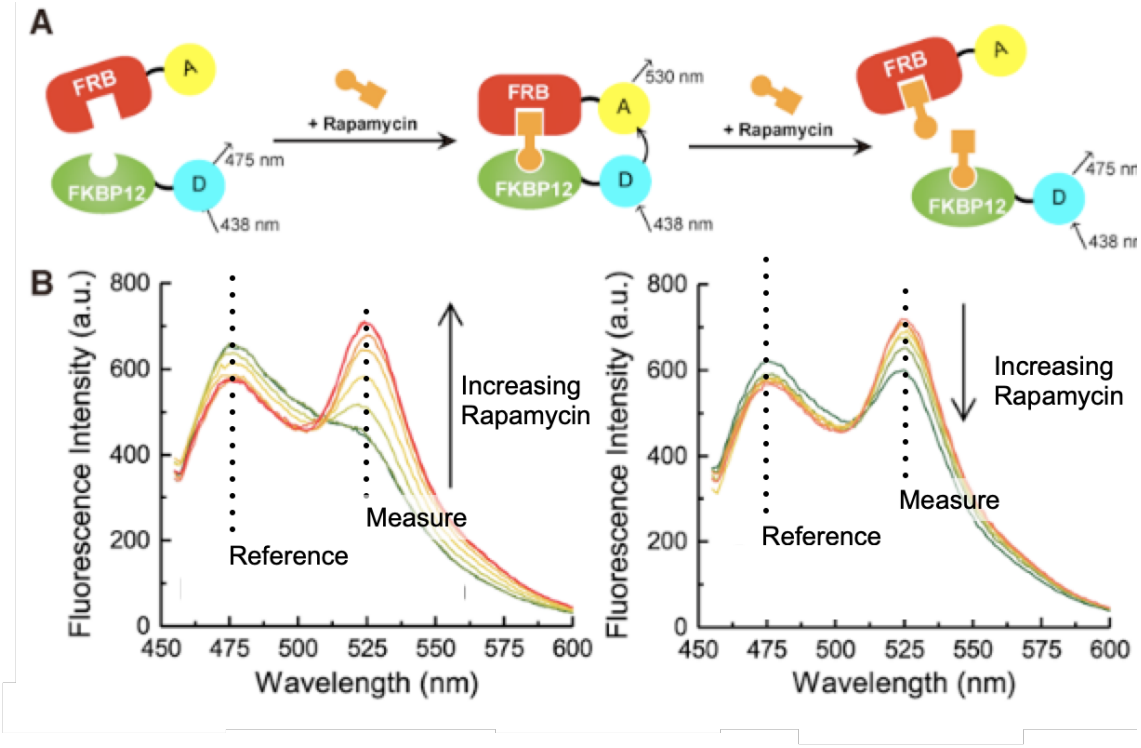


Figure 3.5: FRET probing of interactions between FRB and FKBP12. (A) Principle of the probe. (B) Increasing concentrations of rapamycin increase then decrease the FRET signal at 530 nm while the reference wavelength at 475 nm remains constant. Adapted from Ref. (*44*)

resonance frequencies of atomic nuclei to probe the bonds and chemical environments of the various atoms in a molecule, thus obtaining valuable insights into the molecular structure. It can also be used to probe the changes in structure accompanying protein-ligand interactions and thus to determine interaction rates (*45*, *46*). In this way, NMR can be used to detect changes in the chemical environment of a known chemical group during interaction. For instance, titration experiments can be made with varying concentrations of protein : a shift in the resonance spectra with increasing concentration can be interpreted as an increase in association.
Spectral changes can also use spin diffusion to probe proximity, using for instance a phenomenon known as saturation transfer difference (STD) between nuclei in close proximity (*47*). Very briefly, a change in the spin of certain atoms is induced - the atoms are pumped. That change in spin can be, to an extent depending on proximity, propagated to nearby other atoms ; this propagation can be detected. If the protein is being "pumped", comparing the spectra of liegand between pumped and non-pumped states leads to determination of an STD amplification factor, from which the K can be calculated. This method is also very useful to screen libraries of putative agonists (*47*). Finally, as the spin transfer effects are localized, the method can even determine on which chemical groups of the ligand the interaction takes

Table 3.1: Summary of observables used in ensemble protein-ligand interaction assays.

Observable	Typical methods	Advantages	Drawbacks
Complex mass	Surface Plasmon Resonnance	Relatively simple ; compatible with flow controls	Significant mass must be attached to both partners ; very sensitive to nonspecific interactions.
Temperature	Isothermal Calorimetric Titration	Inexpensive, wide range of use	Requires large sample volumes ; thermal and pH artifacts are possible, certain reaction cannot be measured because of thermodynamic properties.
Reaction	Activity assays	Simple to observe	A reaction must be found ; positive and negative controls are required.
Spectroscopic	FRET, Fluorescence intensity	Intrinsic properties ; non-damaging	May require probes ; bleaching limits lifetime

place.
One of the advantages of NMR studies in this application lies in the high (mM affinities can be determined) and the complimentary with structural studies that can be made. Unfortunately, this method requires very high concentrations of compound and thus sub- μ M affinities are difficult to determine. NMR it also requires expensive equipment, and the analysis of NMR results is very technical.

3.1.5 Summary of macroscopic observables

Many different parameters are representative of the state of a binding and unbinding reaction, with varying accuracy and precision. Not all can be used for every situation, but there are many different approaches to each. A summary is presented in Table 3.1.

3.2 Macroscopic perturbative methods

To measure the rates, it is necessary to break the equilibrium so that the rate appears in the temporal evolution of the observable. Different methods can be used to synchronize the reaction, to make it start at the same time across a volume of reagents. An overview will be presented here.
The perturbation must allow a precise measurement of the reaction rate constants. To this end, the perturbation characteristic times must be significantly shorter than the reaction characteristic time constants.

3.2.1 Concentration as a perturbation parameter

The reaction speed of multimolecular reactions depends on the concentration of the reactants. Thus, concentration is an important control parameter and controlling the concentration is the most obvious way placing it out of equilibrium - as the simplest example, one could start a reaction by putting in contact the reactants. However, on a macroscopic scale, making certain that all the reaction volume is simultaneously put at the same concentration is no mean feat. Mere diffusion is often insufficient since typical diffusion constants of solutes in water will be on the order of $10^{-5} cm^2/s$ for small molecules or $10^{-7} cm^2/s$ for globular proteins and less for long polymers such as DNA (*48*). This leads to equilibration times on the order of $10^5 - 10^7$ (days) for centimeter distances, the size of a typical spectrophotometric measurement cell. Other approaches might disturb the measurement process, such as turbulent mixing (which can cause cavitation and bubbles as well as make the solution unfit for acquisition while mixing). Thus, many different methods have been designed and used to perform fast changes of concentrations. A few examples will be presented here.

Ensemble flow methods

With a reactant attached to a surface - glass for instance, such as is the case with SPR experiment - it is possible to flow in and out of the measurement chamber a buffer containing the molecular partner. It is then important to know whether such a flow is sufficient - or not - to change the local concentration quickly enough that the measurement of the rate constants be accurate. Poiseuille's law of flow is an obstacle here : at the edges of a channel, the flow is zero, leading to a quadratic flow profile as seen on Fig. 3.6. This is a problem because the binding reaction itself happens on the walls of the channel, where one partner is immobilized. Thus, flow methods involve diffusion, but over a reduced distance.

All of those additional diffusion steps can cause the rate constants to appear lower than they actually are. This issue has, indeed, been reported, for instance, with SPR studies (*31*). Those effects can sometimes be detected if binding and unbinding appear to follow a double-exponential, rather than a single-exponential, law (the two rate constants of binding and diffusion can be detected) (*26*, *49*).

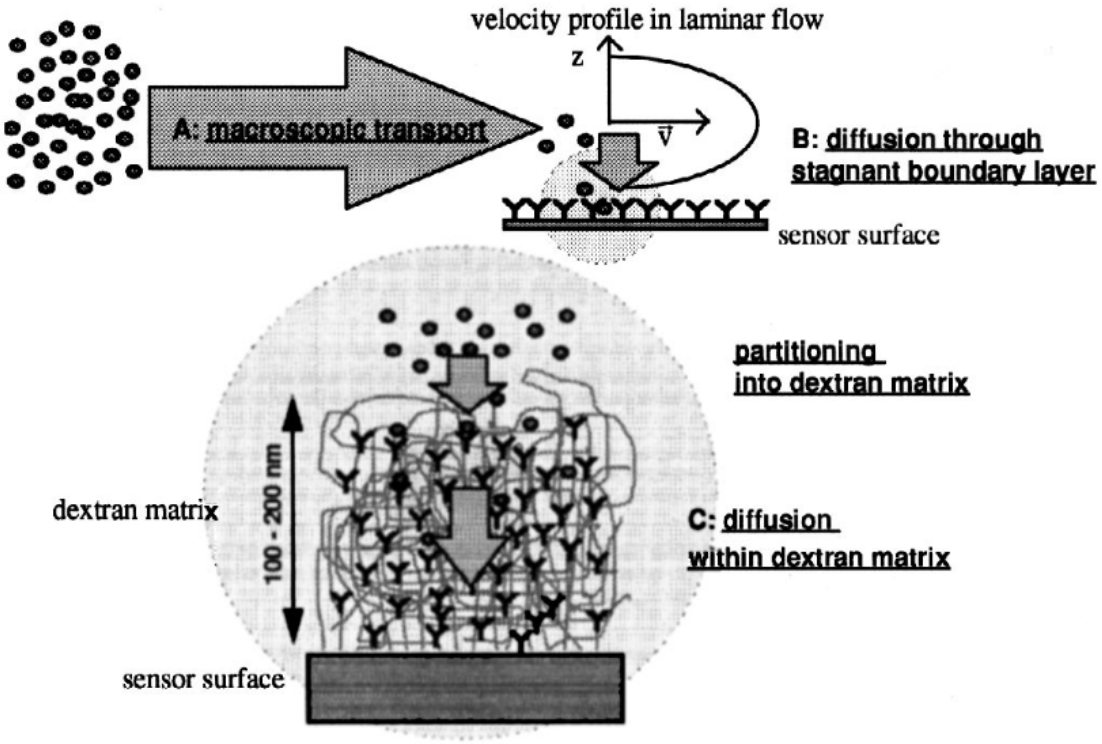


Figure 3.6: Steps of a buffer wash during an SPR experiment. (A) Flow leads to macroscopic transport of the ligand in solution (B) Boundnary layer does not allow flow, short-distance diffusion happens (C) Dextran matrix containing fixed molecular partner limits the diffusion through it, leading to slower diffusion and access to some binding site buried deeper. Adapted from Ref. (*31*)

Stopped-flow

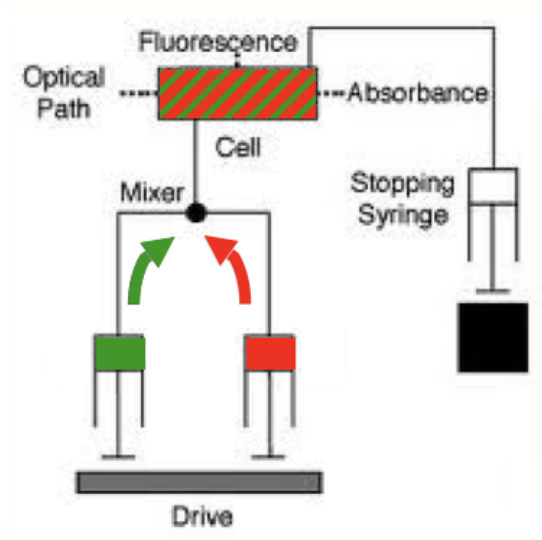


Figure 3.7: Principle of the stopped-flow technique. Various signals, such as fluorescence and/or absorbance, can be measured. (*50*)

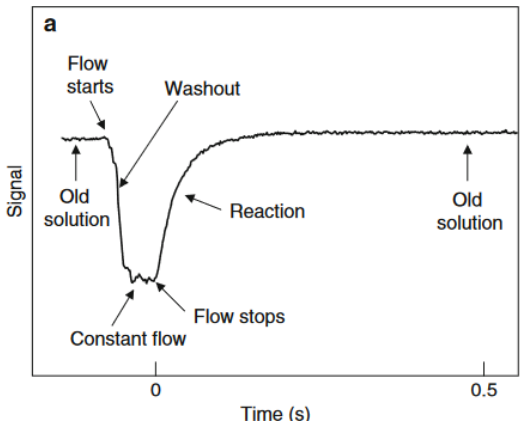


Figure 3.8: Sample trace obtained from a stopped-flow experiment. Recording usually starts at time 0, when the counter-flow is started. Notice that the upward jump in signal does not exactly align with time 0 : this difference corresponds to the instrument's *dead time* Adapted from Ref. (*50*)

Stopped-flow techniques are a vast family of experimental methods to probe the kinetics of fast reactions. The reaction being assayed happens when two solutions are mixed together. Using syringe-pumps, a steady state of flow is reached in a chamber. A third pump, working against the previous ones, stops the flow (hence the name) in the chamber, and the reaction is observed via various optical methods (spectrophotometry, fluorescence, circular dichroism, etc.) depending on the reaction (*50*). See Fig. 3.8 .

This method allows to pinpoint very precisely the time at which the reaction begins in the flow chamber : the main limitation is the *dead volume* of the instrument, that is the volume that must be filled before the flow stops. This volume depends on the instrument but is typically on the order of tens of microliters for commercial apparatuses - divided by the flow, it gives the dead time, typically 1 ms (*51*). Reactions with half-times shorter than this cannot be studied.

One major disadvantage of this method is that filling the instrument and syringes requires large amounts of sample volumes - from hundreds to tens of microliters depending on the instrument model.

Also note the variant named quenched-flow, where after rapid mixing the two solutions are mixed with a quenching agent, stopping the reaction (acid, or metal chelators like EDTA are most common). The sample of known age can then be analyzed via HPLC or other methods. This system can provide samples separated

by up to 5 millisecond intervals, but uses even more reagent and is limited by the quality of the mixing . The advantage of real-time observation is also lost (*51*).

Flash photolysis

The technique of flash photolysis uses light-sensitive compounds (usually UV-sensitive) to activate the reaction. The technique was first used for the study of light-dependent reactions, but the arrival of UV-sensitive molecular cages on the market has generalized the usage. Cages can be used to release either ligand directly or to modify the environment, for instance by releasing H^+ ions or Ca^{2+} ions when studying reactions that are sensitive to it. This can lead to very quick, localized modifications of the equilibrium and, combined with other optical methods, allows the study of interaction kinetics. Note, however, that the method introduces a second reaction (the photolysis itself) and it must thus be proven first that the photolysis reaction is fast in comparison to the reaction being studied (photolysis typically takes less than a ms). It is also necessary to ensure that the UV light does not damage the compounds (*51*).

3.2.2 Temperature as a perturbation

Temperature Jump Relaxation Spectroscopy

Using a pulsed laser, a very small volume of an experimental cell can be heated extremely quickly (current devices can achieve a change of 10-20 K in 5-10 ps) thus bringing the solution out of equilibrium. This method can offer a very high temporal resolution, allowing even to monitor certain reaction intermediates. Van't Hoff's law ($\frac{dln(K)}{dT} = \frac{\Delta_r H^0}{RT^2}$) limits the sensitivity of the method, as our capacity to spot changes in K is linked to the change in enthalpy. Biological binding and unbinding reactions typically show a high reaction enthalpy, in the 20-200 kJ/mol range and thus are well-suited to this method (*52*).
Temperature changes are also useful when dealing with more complex systems. Oscillating temperature changes have been used to investigate certain kinetics in a system and probe it using different temperature oscillation periods (*53*).

Temperature induced pH jump

An alternative method is the Temperature induced pH-jump (*25*) : the sample is complemented with a buffer that exhibits a high variation of pKa with temperature, and the sample is heated with a laser device. The change in pH is often enough to displace the equilibrium and create a synchronised departure from equilibrium, enabling reliable kinetic measurements. This has the advantage of circumventing

the ΔH requirement of a T-jump measurement, but with the added pH change as a potential source of noise. In addition, the buffers used for this type of experiment need to be chosen very carefully to obtain the desired effect.

3.2.3 Pressure as a perturbation

Reaction equilibria are pressure-sensitive. This sensitivity comes from the fact that a chemical reaction causes a change in volume. This change in volume is linked to pressure as the reaction must work against the pressure - hence, a change in volume ΔV requires a work $W = -P\Delta V$ to be achieved. Hence, changing the pressure on an at-equilibrium reaction displaces the equilibrium and causes a re-settling of it, which can be tracked and interpolated. Pressure jumps can be best generated with a pizeoelectric crystal that can create a strong and fast pressure jump. The pressure change can be low : for instance, a system pre-pressured at 4 MPa exposed to an increase of 7 mPa of pressure can cause a 1 % change in observable reaction rate as illustrated in Ref. (*54*). If the changes are low, they can be executed extremely quickly, typically in about 50 microseconds. The changes propagate at the speed of sound, precluding any diffusion or inertia effects. Thus, it is possible to collect multiple repeated experiments quickly and average them (*51*). In addition, pressure can be maintained constant over long periods of time with ease (contrary to local temperature increases which revert to baseline through diffusion) and the change affects the entire cell, making diffusion a non-issue. Everything considered, the method is robust and reliable but not widely used in the protein-protein interaction field of research - it is much more widespread in use for the study of protein folding, for instance (*55*).

3.2.4 Comparative summary : ensemble methods

We have presented an overview of the most popular macroscopic approaches to equilibrium and rate constant determination in protein-ligand interactions. The most important methods (combining observable and perturbation) are summarized in Table 3.2. One of the most important drawbacks of macroscopic measurements is the sample consumption, especially in flow methods in which mixing prevents the sample from being reused. Current technological development aims at reducing as much as possible the sample volumes, since not only will this reduce the cost but also decrease the time required to propagate a perturbation and thus increase temporal resolution. The ultimate reduction in volume, of course, are single-molecular methods.

Table 3.2: Methods for ensemble protein-ligand interaction assays. Adapted from Ref. (*35*)

Method	Temporal resolution (ms)	Sample size (μL)	Observables	Notes
ITC	Depends on reaction $\Delta_r H$	200	Heat production	Measures thermodynamic potential(G,H, and S) due to binding
Fluorescence polarization	1	25	Molecular mass change	Requires reagent labelling ; lifetime of label limiting.
SPR	100	100 - 1000	Absorbed mass	Flow methods limited by diffusion in boundary layers ; expensive
Stopped-flow	1	200 - 2000	Absorbance, fluorescence	Requires large amounts of sample
Temperature jump relaxation spectroscopy	1 ms	10	Absorbance, fluorescence	Complex heating and cooling apparatus ; requires high enthalpy change and corrections for pH and temperature changes.

Chapter 4

Single-molecule principles and methods

Contrary to what happens at the macroscopic scale, when dealing with individual molecules, there is no notion of equilibrium : a molecule is either in interaction, or not. Single-molecule studies of protein-ligand interactions thus always enable kinetic measurements. However, they have other weak points that need to be accounted for.

The conceptual gap between macroscopic and single-molecule assays is bridged by the ergodic hypothesis, which states that statistical data gathered on multiple molecules (macroscopic measurements) and on the same molecule, repeatedly (single-molecule) are generated by identical random distributions. Single-molecule studies are uniquely placed to investigate the validity of this hypothesis, as it is possible to compare different molecules under the same conditions. With sufficient resolution, single-molecule studies can also detect reaction intermediates between states and afford a deeper understanding of the reaction.

Single-molecule studies generally conform to two experimental modes, active and passive measurements. We will first discuss their general principles and objectives before presenting examples.

4.1 Generalities on single-molecule measurements in protein analytical chemistry

4.1.1 Active and passive approaches

We will make a distinction between *active* and *passive* studies. Active techniques exert force - via electrical or mechanical means - on the molecule for the observation, prompting a change in behavior, and measure this change (time, amplitude,

etc.). These approaches reveal very interesting energetic aspects of a protein-ligand interaction landscape. Other approaches are *passive* - purely observational - and consist of observing the spontaneous fluctuations of a single molecule under as little external influence as possible.

4.1.2 Single-molecule experimental protocols

We propose here a simple classification of single-molecule experimental protocols.

Constant-force methods can be active or passive. They require simple observation of molecular transitions over time, at a known force - or no force at all. This very simple setup makes it possible to detect the dwell-times : the time the protein spends in the open (non-interacting) and closed (interacting) states. From there, it is possible to infer the dissociation and equilibrium constants.

Force-step methods are active. They alternatively apply high and low forces on the protein-ligand complex. During the low-force (typically less than 1 fN) periods, the protein and ligand couples are left at rest, with the aim of favoring interactions. The force is then quickly raised to high force ; the value of this force ,depending on the instrument, ranges from piconewtons to hundreds of piconewtons. This pulling force can rupture the PL complex, and this rupture can be detected ; the time taken can be connected to the energy of the interaction via the Arrhenius-Bell equation (Equation 4.2 below).
This method can also be used to probe the equilibrium constant by measuring the percentage of cycles that lead to a dissociation event : this percentage is the ratio of association in the low-force state. Thus, with the knowledge of the dissociation rate constant, the association and equilibrium constants can be determined.

Force-ramp - also active - methods track the response of the system to an increasing force, aiming to detect breaking points. This is very useful for detecting reaction intermediates ; one of its core uses is the study DNA hairpin (*57*) or protein unfolding (*58*). Note that the force-ramp approach is in no means the only one to allow detection of intermediates.

Different protocols yield different results : passive observation can be used to measure the dwell-times at zero external force, but this constraint imposes the drawback of requiring more time no energy is provided to the system and thus experiments are generally longer. On the other hand, the active study of relaxation allows the probing of the energy landscape, but zero-force interaction constants can

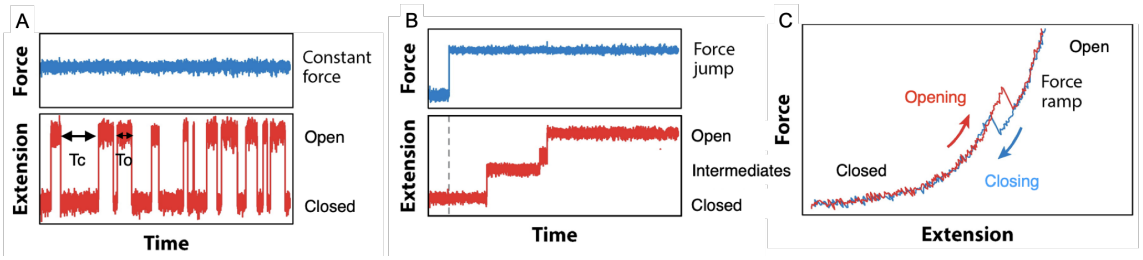


Figure 4.1: Example time-traces in single-molecular experiments. (A) constant-force measurements show spontaneous fluctuations between open and closed states (B) Force-jump experiment shows the stability of the closed complex and reveals intermediate dissociation steps (C) Force-ramp experiment offers insight into the binding and unbinding dynamics, and illustrates the different energetic pathways of both reactions. Adapted from Ref. (*56*)

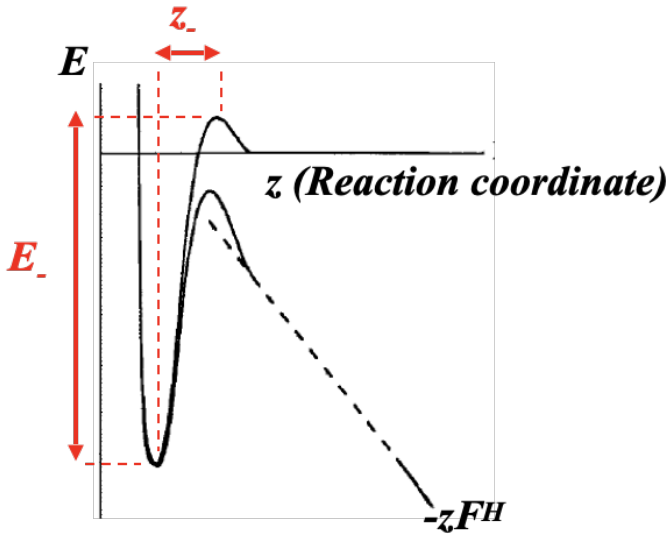


Figure 4.2: Schematic representation of an energy landscape of the dissociation of a protein-ligand complex, with and without a contributing force F^H. Adapted from Ref. (*59*)

only be extrapolated.

Force and energy on the single-molecule : the Arrhenius-Bell equation

Under zero-force, the dwell-times of the PL complex are distributed according to a single-exponential law of parameter k_-, thus the average dwell-times of the closed complex τ_{PL} and open complex τ_P can be expressed as :

$$\tau_{PL} = \frac{1}{k_-} \quad \tau_P = \frac{1}{k_+[L]}$$

Unbinding reactions can be seen as escaping from a potential well of height E_-, the activation energy barrier for dissociation. The Arrhenius-Bell equation provides

the relationship between the unbinding rate constant and the energy barrier (*59–61*) :

$$k_- = A_- exp(\frac{-E_-}{k_B T}) \tag{4.1}$$

With k_- the opening rate constant , A_- a dimensional constant, k_B the Boltzmann constant and T the absolute temperature. E_- would be expressed in energy units per molecule [1].

Exploring energy landscapes is the typical use case for the use of active single-molecule experiments. The force is directed along an axis, making it easy to define reaction coordinates, and to determine the displacement induced. Thus, the energy modification caused by the active force (force times displacement, or integral of force along the displacement) can be calculated. If a pulling force F^H is applied, the activation energy is modified by $F^H.z_-$ where z_- is the transition distance, modifying the energy landscape. As the pulling force is aligned with the opening reaction, the Arrhenius equation becomes :

$$k_-(F) = A_- exp(\frac{-E_- + F^H z_-}{k_B T}) \tag{4.2}$$

This equation represents the fact that the energy barrier is reduced by the external application of force. In logarithmic scale, this can be translated to a simple line :

$$ln(k_-(F)) = ln(A_-) + \frac{-E_-}{k_B T} + \frac{F^H z_-}{k_B T} \tag{4.3}$$

Using measurements at different forces, it is possible to use linear interpolation to deduce a zero-force dissociation constant. Then, using experiments at different temperatures, E_- can be calculated via a secondary regression. This process is illustrated on Fig. 4.3. Those experiments can thus be made with a wide variety of techniques, including AFM, Optical and Magnetic tweezing.

Probing structural dynamics

If the anchoring point on the molecule to study can be chosen (using protein engineering or organic chemistry) then the pulling direction can be precisely controlled. This opens multiple new avenues of investigation for structural dynamics as the conformational adjustments of proteins can be measured to the angstrom in chosen directions, for example in response a binding event (*18*, *62*). This, however, requires new scaffold - or at least protein - synthesis for every pulling direction required. Fortunately, as a scaffold harbors two different partners, all combinations

[1] Or in energy units per mole replacing k_B by R in Equation 4.1.

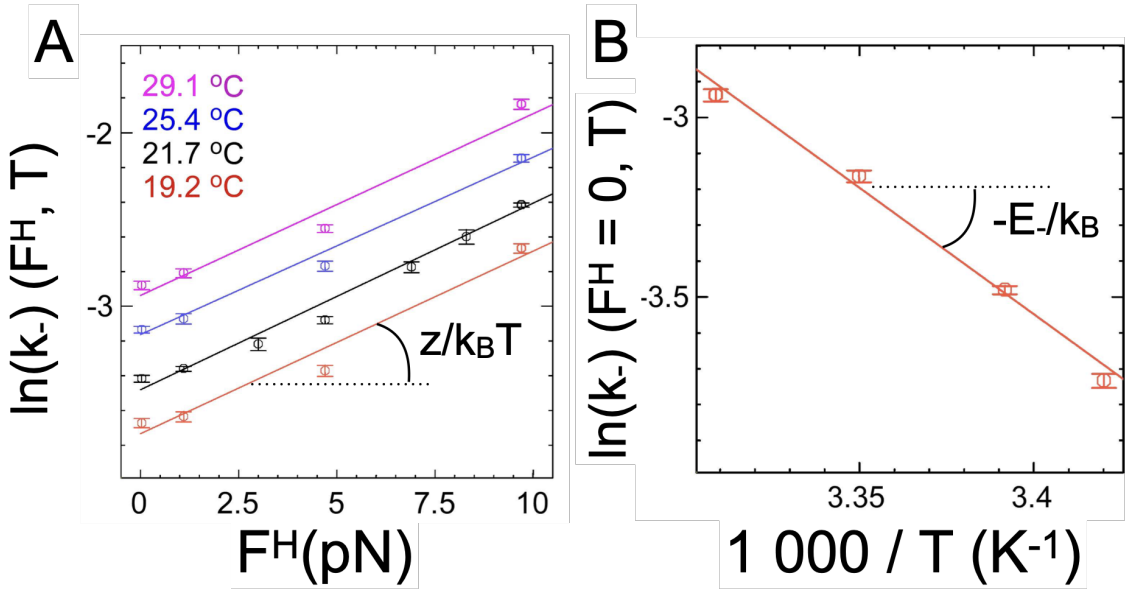


Figure 4.3: Single-molecule determination of E_{-} (A) Linear interpolation of the equation 4.3 on FKBP12-rapamycin-FRB dissociation at different temperatures. Note that the slopes are equal while the intercepts differ. z_{-} is here of 0.44 nm. (B) Linear interpolation of the intercepts of the previous panel in function of the inverse of the temperature yields the dissociation activation energy $E_{-} = 58.6 \pm 1.7$ kJ.mol^{-1}. Adapted from Ref (*18*)

can be investigated, yielding more different pulling axes per protein, as shown in Fig. 4.4.

4.2 Single-molecule techniques for molecular interactions

Most tools can only do either passive (as they are incapable of exerting force) or active (as they use force and resistance to observe activity). Some techniques can, however, exert variable forces that can be made small, so as to come close to zero-force behavior. To present an organized overview of the existing single-molecule techniques, we will begin with the historical electrical techniques (observing membrane ion channel dynamics) and show how they have been adapted to study non-membrane protein dynamics. Then, we will move on to mechanical techniques that control force to observe displacement. Finally, we will discuss optical observational techniques, passive-only tools that can complement the others.

4.2.1 Electrical single-molecule techniques

The first single-molecule observations were made on transmembrane ion channels, using patch-clamp technique : using a glass pipette to isolate a small patch of neuron

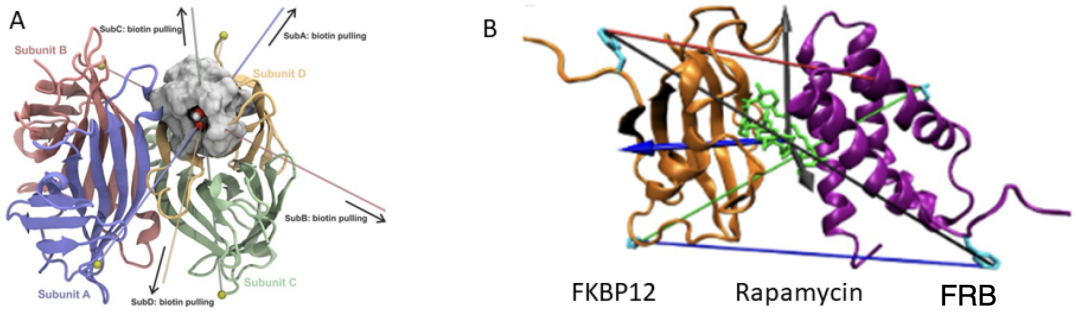


Figure 4.4: (A) AFM pulling direction experiment probes different interaction strengths between Biotin and the four domains of streptavidin. Adapted from Ref. (*62*). (B) Structure of the ternary complex of FKBP12-Rapamycin-FRB with 2 possible different modified amino acids for attachment to a magnetic tweezer. This makes it possible to investigate complex rupture along 4 different axes (blue, orange, green and black lines). Adapted from Ref. (*18*).

membrane that, on average, contains only one channel. This way, it is possible to observe the opening and closing of the channel in real-time by measuring the current flowing through the pipette , and to build dynamic models that reflect the reality of a single protein and of its states - and gating mechanisms (*63*).

Nanopore force spectroscopy

Nanopores are molecular structures inserted into a biological membrane - an artificial lipid bilayer for instance, or a cell membrane - and allowing the passage of certain macromolecules through. By placing the nanopore between two otherwise electrically separated compartments, it is possible to observe the flow of ions through the pore by measuring the resulting current. Alternatively, by imposing a voltage difference using an outside generator or different ion concentrations in both chambers, a current of charged particles can be forced. This current, however will not only depend on the potential difference but also on the width of the pore. If a large molecule passes through, the current will be reduced, and this reduction can be connected to the size of the molecule, its conformation, size of domains, etc. To quote Varongchayakul et al. (Ref (*64*)) : *"pushing an elephant through a door significantly reduces the flow of air through the door while an ant quickly and easily crawls through the door barely, if at all, disturbing the airflow"*.
This tool has been adapted for various DNA-based applications such as nanopore sequencing. Recently, however, it has also started being used as an interaction sensor between proteins and DNA sites in a technique called nanopore force spectroscpy (*65*). A protein and its target DNA site are put in one chamber, and the potential difference is adjusted so that the complex - which is negatively charged - is pulled through the pore. The protein is however too big to pass the pore, so it is detached

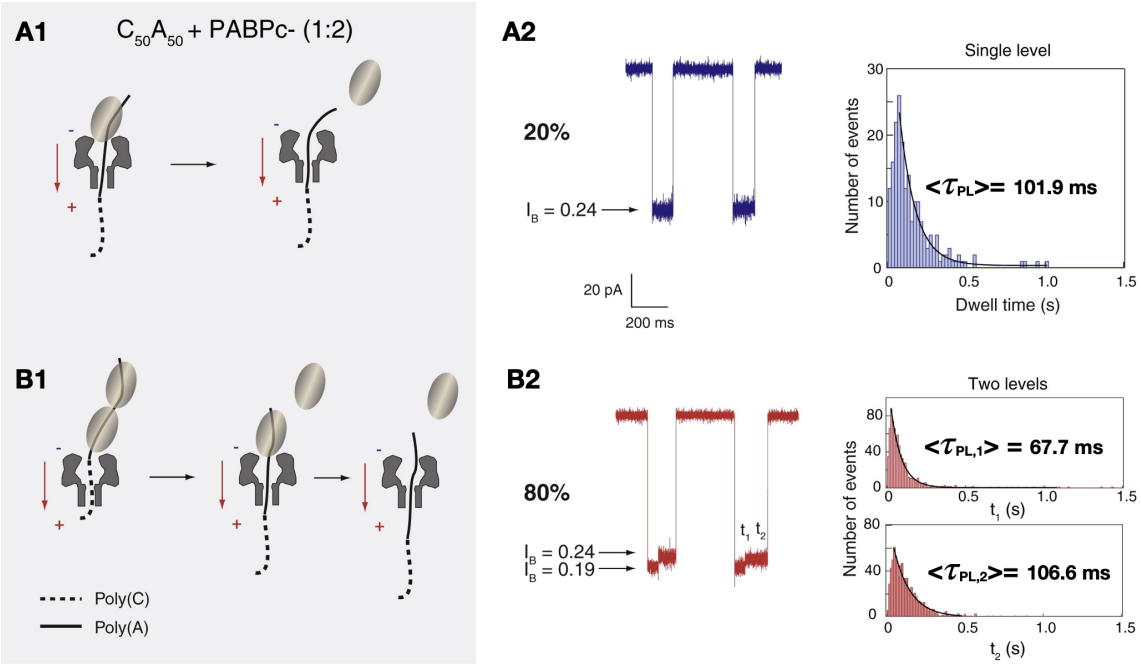


Figure 4.5: Two types of detachment between a DNA construct serving as a target for protein PABPc translocating through an α-hemolysin nanopore. (A1) a single protein is bound on the DNA. (A2) The translocation can be detected and timed, measuring the dwell-times leads to the average on-target dwell-time, inverse k_{-} . (A2) Two proteins are bound (B2) The current with 2 bound proteins is lower than with only one bound protein, since the obstruction is larger. Two dissociation events can clearly be seen, and the rate constants of both events can be interpolated. Adapted from Ref. (*66*).

from the DNA as the later is pulled through. The dwell time - time required for translocation and dissociation - can be easily seen on the recording of the electric current, as while the protein is inside the pore, it markedly obstructs the flow of ions - see figure 4.5. The pulling force can be evaluated as the chargeof the molecule times the voltage gradient, so this experiment is homologous to a force-clamp experiment (*66*). What is very interesting with this experiment is that it does not require any scaffolds to be built, only to have pores of the right size to separate one partner from the next. They do require the pull-through partner to be strongly charged however, so they are perfect for the study of Protein-DNA binding (as a quick estimate, pulling DNA - 3 base pairs - through a 1 nm long pore with a transmembrane potential of 100 mV exerts a force on the order of the piconewton) (*65*). Finally, this technique can detect even small changes in the binding conformation, making it practical for the analysis of multi-step dissociation mechanics. Force-ramps can also be created by ramping the voltage. It is not suited to all cases, requiring charged partners to function.

4.2.2 Mechanical force single-molecule methods

Using mechanical rather than electric force removes the need for charged partners. Very different means can be used to exert force on a molecular structure, and how that force is controlled is an important issue. There are two different ways in general to control the force exerted on a sample : either by controlling the force (force-clamp) or by controlling the force potential (position-clamp). Magnetic, acoustic and electrostatic force-based approaches have a fixed force as the device creating the force is much larger than the molecules being manipulated. Mechanical (AFM) or optical devices, on the other hand, are exerted by a potential and so depend on the position, and require a precise feedback loop to maintain a constant force. Constant-force and force-step approaches are typically easier with force-control while force-ramp protocols make the best use of position-control devices.

Atomic Force Microscopy

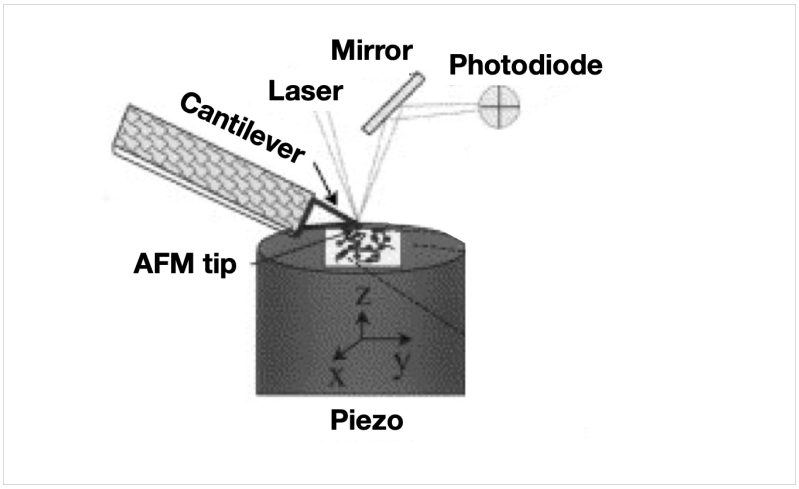


Figure 4.6: Standard set-up of an AFM device. Adapted from Ref. (*67*)

Atomic force microscopy, or AFM, relies on a sharp tip mounted on a flexible cantilever, held by computer-controlled piezoactuators. When the cantilever contacts or experiences forces from the surface, it is deflected from its position, and the defection can be measured using the reflection of a laser beam off the back of the cantilever head (Fig. 4.6). The stiffness of the cantilever is key in obtaining a high force resolution, and modern cantilevers can apply forces between tens and thousands of pN (*68*). The resolution is limited only by cantilever materials and laser power, which limits the accuracy when determining deflection. To obtain multiple interaction events and remove bias, AFM is often operated by "tapping" the surface repeatedly with the cantilever head - this way, the interactions happen vertically in a reproducible manner (*67*).

To measure interaction strengths, a surface is coated with one partner while the head is coated with another. The cantilever reaches oscillation, then is approached to the surface. As the cantilever contacts the surface, it is deflected from it and the deflection is detected. Then, once it reaches the end of the oscillation, it retracts again, leading to decreasing deflection again, and finally return to baseline. A typical trace is shown in Fig. 4.7 To extract biological data from such experiments, it must be performed such as to minimize any undesirable forces - thus, the experiment must be performed in liquid buffer, and the solutions must be made with highest purity as any type of contaminant might affect the measurements with nonspecific interactions. The attachment of proteins to the surface must be strong, as a weakly bound molecule could be displaced by the strong forces exerted by the cantilever ; unfortunately, a too strong attachment might reduce the capacity of a molecule to interact and, thus, diminish biological activity (*61*).

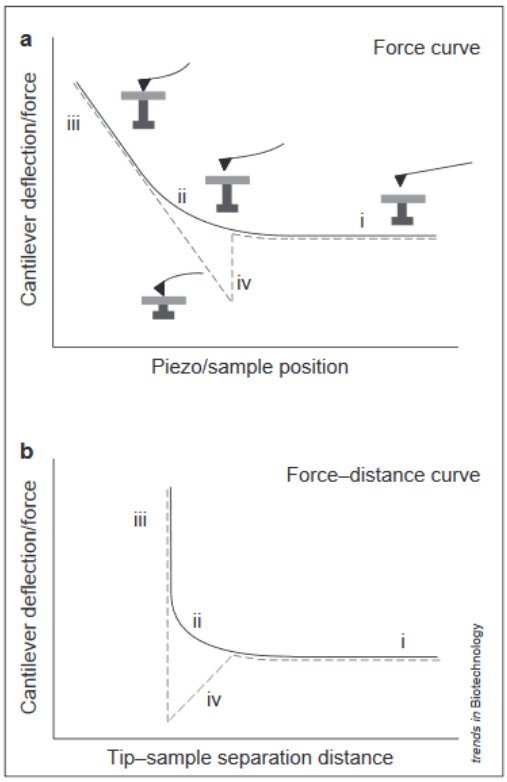


Figure 4.7: A. Steps of a typical AFM-surface contact trace : 1. Approach 2. Tip approaching contact point ; Van der Waals forces start being detected 3. Contact with surface and increasing force and deflection 4. Brief adhesion of tip to surface via weak interactions. Adapted from Ref. (*69*).

Interaction strength can be determined with the force at which the cantilever jumps off the surface : indeed, this strength corresponds to the rupture force that must be exerted to break interaction, and can be deduced from the deflection with a simple spring calculation.

It should be noted that the method is extremely flexible and has been successfully applied to not only protein/protein interactions, but also DNA/protein or DNA/DNA for example, or even cell-surface interaction forces. (*69*)

Many variations of the method exist, for example for affixing the proteins to the cantilever heads, the methods of approach, etc. While the method is very well-known and controlled, it suffers from the necessity of long experiments, requiring a complex set-up. only one molecule may be investigated at a time, limiting data generation. Finally, the attachment of the protein to the cantilever is an important topic, and controls must be performed to make sure that is does not affect the measurements.

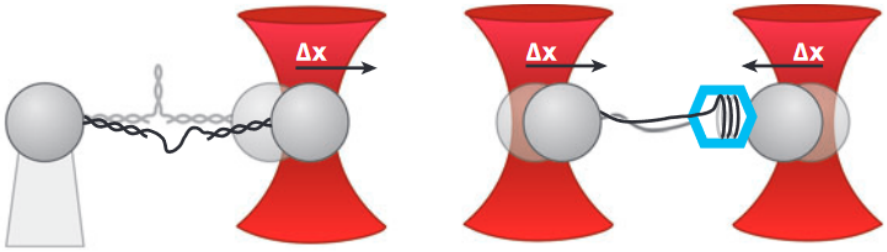


Figure 4.8: (Left) Sample optical trap set-ups for study of secondary DNA structures between an optical trap and a pipette tip (Right) A protein between two optical traps. Adapted from Ref. (*70*).

Optical tweezing

Optical tweezers are a nanomanipulation tool originally developed by Ashkin et al. (*71*) that has since been used in many biological studies (*70*). The core principle of this technique is to use the fact that light carries momentum, and when its propagation is changed, for example by a transparent glass bead, it exerts a force in return. This is harnessed by focusing a near-infrared laser beam onto such a transparent bead : in a well-chosen configuration, the forces exerted by the laser create a potential well that not only confines the bead at the chosen focal point but also behaves like a linear spring, with the recall force proportional to the distance to trap center. Thus, it is quite simple to estimate the force being exerted upon the bead, as it changes linearly with bead displacement. The laser beam can be driven with optical control tools, such as piezo-mounted mirror or accousto-optical deflectors ; feedback loops can be established with the determination of bead position (with an optical camera or a 4-quadrant photodiode) to maintain a constant force, or the lasers can stay immobile for constant position experiments. This makes this method very flexible, as it can be used for both active and passive studies.

Those set-ups were initially developed for nucleic acid research. Tools for the attachment of nucleic acids to beads, such as biotin-streptavidin linkage, are common. Usually, this kind of studies involve trapping the sample held on DNA or RNA tethers between two beads, held in two different optical traps or in an optical trap and a micropipette suction trap (see fig. 4.8).

However, optical tweezers can and have been applied to the measurement of protein-protein interactions. For example (*72*) studying the interaction between the von Willebrand factor and platelet surface glycoproteins (see Fig. 4.9).

Optical tweezers have many advantages : they allow for a wide range of forces (from hundreds of femtonewtons to hundreds of piconewtons) and can manipulate objects with two handles in every direction. The technique is real-time, and force feedback loops can be set up to maintain a constant force or to change it automatically when

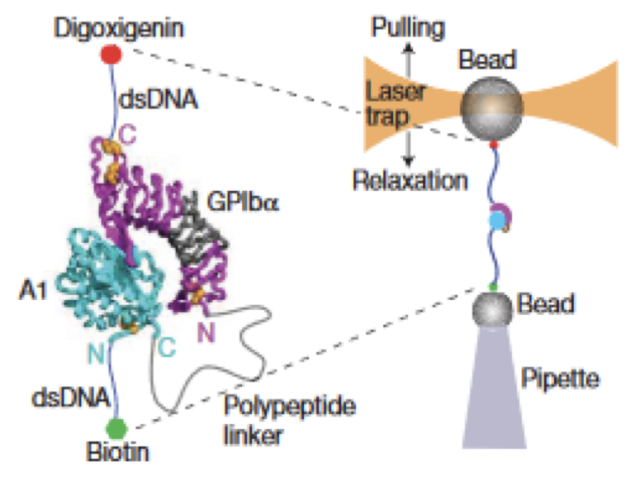


Figure 4.9: Receptor and ligand in Single-Molecule construction using optical tweezer setup. Adapted from Ref. (*72*).

certain conditions are met - unfortunately, a constant feedback loop also takes time, which means that reaction times under 1 ms are difficult to probe as the feedback loop cannot be established properly in those short intervals (*73*, *74*).

Magnetic tweezers

A magnetic tweezers is a device coupled with an optical microscope to manipulate microscale objects using magnetic force. The sample is attached on one side to a glass surface and, on the other, to a paramagnetic micro-bead (1 μm in diameter usually). A pair of magnets is placed above the surface and pull the magnetic beads away. The magnets can be rotated, which can provide torque. Because of the size of the bead, multiple beads (up to 50-80) can be observed simultaneously, under identical conditions. The force exerted is uniform and typically in a range from 20 fN to 10 pN.

Magnetic tweezers have multiple advantages in the field. The technique can provide torque. The ranges of force are wide, which makes it possible to do both active and passive measurements. Many experimental points can be acquired at once, and it can be coupled with other methods such as fluorescence. It is also relatively simple to set-up and run. Multiple drawbacks are however to note : changing the force requires a macroscopic movement, creating experimental dead-times between forces. The evolution of the magnetic field in function of magnet position requires prior calibration as the force cannot be simply calculated. Finally, the metallic beads reflect light, meaning that direct illumination from lasers can be problematic, and fluorescence is restricted to indirect excitation such as TIRF (*75*, *76*).

Acoustic force spectroscopy

Acoustic tweezers are a tool initially developed for the bulk sorting of particles, but have recently gained interest in the molecular biochemistry (*77*, *78*). One or multiple piezoelectric actuators generate sound waves in the experimental medium (water) creating wells of potential that can attract micro-particles, such as plastic beads of very small diameter (nm and above). If the bead is attached to a molecular scaffold, it can be pulled apart with a chosen pulling force, which can be set low (from 0.3 fN to 200 pN) (*79*). Alternatively, a uniform potential gradient can be created to exert identical forces upon a large number of particles. This technique can generate precise forces in the piconewton to tens of piconewton range (*77*, *79*) on one or multiple beads. Acoustic tweezers cannot reach the spatial positioning precision of optical tweezers due the lower frequencies of sound waves (*77*). Heat dissipation can be an issue in some set-ups. The technique is extremely flexible and still improving, being able to affect many particles with an identical force that can be modulated quickly as needed.

Laminar flow chamber

Laminar flow chamber-based techniques are a broad family of techniques, largely used for the study of cell-cell adhesion. They can be applied to protein-protein interactions as well. One molecular partner is attached to the glass surface of a flow chamber under an optical microscope while the other partner is coated on micrometer-sized beads. The beads are then flown through the chamber, and tracked over time using the microscope. When both molecular partners interact, the beads stop their movement. The duration of those arrest events can be linked to the force exerted by the flow on the beads (which can be measured or calculated) and to the energy of the binding interaction, making it even possible to probe the energy landscape of dissociation.
Used in this capacity, the technique suffers from multiple drawbacks : it is necessary to prove that dissociation events only involve a single molecular binding event ; the force can be difficult to establish precisely and non-specific interactions between the bead and surface must be carefully taken into account in control experiments (*80*, *81*).

4.2.3 Observation-based methods

In addition to force-based methods, electrical or mechanical, single-molecule assays can be done using only optical means. Mostly using fluorescence, those methods oftentimes struggle to generate a strong enough signal-to-noise ratio but have the advantage of minimally perturbing the sample. They also require less mechanical

access to the sample, making them usable in conjunction with other force-dependent strategies.

Direct observation

The direct observation of the rotation of the F1-ATPase by Noji et al. (Ref. (*82*)) is a prime example of passive single-molecule observational experiment. They attached a short length of actin on top of the mitochondrial membrane complex to observe the rotation of one of its subunits. Even better, they were able to quantify the rate and stoechiometry of proton gradient to ATP conversion by observing the most represented angular motions of said actin filament and deducing the number of individual steps required for a full turn.
Direct observation of molecular interaction is not always possible, but can be done, either by electron microscopy or by using fluorescent proteins and ligands. Surface proteins or other configurations where the proteins condense in a one or two dimensional configuration are often well suited to direct study (*73*).

Correlative fluorescence microscopy

To visualize the association and dissociation timing of partners in solution, fluorescent microscopy is a popular choice, because of its relative ease of implementation and flexibility. TIRF - total internal refraction fluorescence is very well-suited for SM experiments. the protein is immobilized on a glass surface and an evanescent field is generated (depth of about 100 nm). Only the proteins and ligand molecules very close to the surface are excited, which considerably decreases background noise and aids in focusing. The residence times of colocalization (times during which two different fluorophores appear to be on the same position) are recorded, and a distribution can be interpolated from them, leading to the determination of interaction rate constants (*83*). Multiple binding events can be recorded at once, so more data can be collected per experiment.
This approach does have many limitations. As it relies on optical imaging, it cannot co-localize proteins more accurately than the diffraction limitation (typically 100 nm) : it is therefore impossible to distinguish effectively between random close proximity due to diffusion and genuine interaction, and control experiments are required to remove the effect. Photobleaching of fluorophores can be an issue, so for long experiments, the field of view needs to be refreshed. Finally, the concentration needs to be low enough that only one fluorescent ligand will be present in a diffraction-limited area around the target for a single interaction to be reliably detected (typically lower than 10 nM (*84*, *85*)).
Förster resonnance energy transfer (FRET) techniques can be used to circumvent

those limitations. FRET interactions are typically established at distances of the order of 10 nm, and are a much better indicator of interaction than simple colocalization. (*85*, *86*).
Background fluorescent noise remains a problem at high concentrations, even in FRET. Practically speaking, concentrations higher than 10 nM can become problematic (*85*) to work with. Labelling only a fraction of the molecules circumvents the issue, but requires computation to reach rate constants. Control experiments are be required to prove that labelled and unlabelled molecules behave comparably. Novel optical techniques can be helpful here, such as STED (stimulated emission depletion) where fluorescence around target samples is depleted using another emission bead, which further increases resolution (*87*). Zero-mode waveguides allow for the construction of very small spaces (50 nm on the side and 30 nm in depth), reducing the amount of background noise while still maintaining high concentrations (*88*). The use of oxygen scavenging compounds can increase significantly the average lifetime of fluorescent compounds (*89*) while imaging at lower frequencies (such as once-per-minute imaging pulses) allows to experiment on very slow reactions, so experiments lasting days are possible. (*90*)

Fluorescence correlation spectroscopy

Fluorescence correlation spectroscopy (FCS) uses the autocorrelation of a fluorescent signal issued from a small volume containing reactive fluorescent species. By illuminating a femtoliter volume with a laser, it is possible to acquire only on average one molecule at a time - if fluorophores are concentrated at 10^{-9} M, a volume of 1 fL means about one molecule will be imaged at a time (*24*). The objective is to detect so-called "quantum bursts" - single emissions of fluorescence photons. The fluorescence signal is acquired and correlated with itself over time. This autocorrelation time primarily depends on the fluorophore diffusion speed in the medium. If a chemical reaction occurs during the observation and modifies quantum yield of the fluorescence (such as association of two molecules with FRET probes), that quantum rate change can be quantitatively connected to the rate constants of the reaction (*91*). In addition, much like in Sec. 3.1.1, the diffusion properties of molecules depend on their size, which changes with binding, so it is possible to determine the equilibrium constant in this manner. Variations also exist, such as confocal fluorescence coincidence analysis, which measures coincident emissions in the same femto-volume of two different fluorophores - this can establish interactions between the two labelled proteins (*92*).
This method can be seen as a slight outsider among single-molecule methods, as while it is capable of detecting single-molecular events, it does not track a singular individual molecule.

4.3 Overview : strengths and weaknesses of single-molecule methods

We have shown a panel of currently used single-molecule techniques, both force-based and zero-force. A summary is provided on Table 4.1. For drug-design and discovery requirements, a temporal resolution of the order of the second is often enough, and forces should be low to allow the easiest interpolation of zero-force conditions, or to measure precisely the variation of interaction dynamics with force. Repeatability of the experiments is essential for comparison across compounds. Thus, fluorescence, optical trapping, magnetic trapping and acoustic force spectroscopy seem most useful for the investigation of dissociation rate constants, while AFM is more suited to the investigation of higher-energy structures such as protein (un)folding and DNA secondary structures.

Table 4.1: Table summarising main single-molecule techniques used to probe protein-protein interactions. Adapted from Ref. (*73*).

Technique	Temporal resolution (ms)	Force (pN)	Spatial resolution (nm)	Advantages	Limitations
AFM	1	$1 - 10^4$	0.5	Can probe covalent bonds ;	Attachment geometries ; high stiffness ; high force ; position-control
Optical trap	10	$0.1-100$	0.1	Precise geometry, fast change	Photo-damage ; handle processing ; position-control
Fluorescence methods	1	Passive	1	Ensemble measurements possible. Detects internal motions.	Dye bleaching ; requires fluorophore attachment
Direct observation	1000 and longer	Passive	Varies	Can be combined with flow ; parallel measurements	Only specific cases
Magnetic tweezers	100	$10^{-2} - 20$	10	Torque is possible. Parallel measurements	Large particles ; requires scaffold attachment
Nanopore force spectroscopy	10	$0.1 - 20$	varies	Native proteins ; can probe inner protein movements	Charged molecules only
Accoustic force spectroscopy	1	$10^{-4} - 200$	100	Three-dimension control of position, fast adjustment	Heat evacuation issues ; position- ; parallel measurements

4.4 Hidden Markov models in analytical chemistry

The result of a single-molecule observation is usually a signal trajectory of current, position, or fluorescence intensity, with changes marking a change in protein conformation or interaction. This type of trace is the ideal case for the usage of hidden Markov models in data analysis, which is the next data-analysis tool we will present.

4.4.1 Generalities on hidden Markov models

A hidden Markov model (HMM) is a method of understanding and analyzing data. Very briefly, signals that can be measured - observables - are considered to be produced by multiple different distributions. The system itself moves from one state to the other, and each state generates observables based on a different distribution. However, the states are hidden as they cannot be seen directly ; it is only possible to detect the observables they generate. But it is possible to infer from the observables the distributions generated by the different states and the state of the system at any given time. See figure 4.10 for an illustration.

Under an HMM model, the system is in only one state at a time and can change state with a given constant probability at every time interval. The transition probabilities from every state to another are constant, which means that the time spent in every state is distributed exponentially. The probabilities differ for every transition (and some transitions might be impossible) but the probability of any given transition does not change over time.

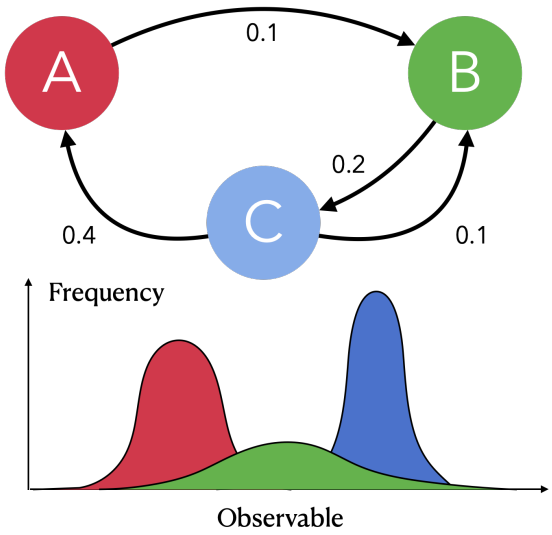


Figure 4.10: Example representation of a Gaussian HMM model. From the 3 states above come the values of the observables at every data point. The arrows represent the transition probabilities at every data point.

HMMs have been used in the field of molecular biology since the 80s, especially for the analysis of ion channels and other single-molecule experiments (*93*). Indeed, ion channels are an almost textbook application example ; the system being the channel, which can be in two different states (open and closed) and the observable is the current. The current at each time-frame is generated by a Gaussian distribution, convolution of the average current and the experimental noise.

4.4.2 Computer inference of hidden Markov models

Many computer libraries and algorithms exist to infer a state description of a time-trace of observables. The most popular are based on the Viterbi algorithm (*94*) which must be told the number of expected different states, and initial guesses for the observable values of those states. Like machine-learning processes, the Viterbi algorithm first learns how to translate observables to states on a segment of the data by determining what is the observable output function of every hidden state, then uses these values to determine, for every time-point, in which hidden state the system is. The result of this second step is the time-series of states called maximum likelihood.

4.4.3 Hidden Markov Models in protein-ligand interactions

The probability of transition from one state to another being constant, HMMs do not exhibit latency. Instead, in order to simulate reactions of higher order in an HMM, states are added. To continue with the ion channel example, some ion channels have a refractory period after closing, during which they cannot open again. This would be represented in an HMM model by creating a 3rd state, that has the same distribution of currents as the regular closed state, with another transition to a regular, non-refractory closed state with a distinct probability.
Thus, HMMs are a natural fit for the modelling of molecular dynamics, as they give tools to accommodate first-order kinetics and, with the introduction of enough transition states, can model any molecular dynamics with sufficient data. The ongoing development of single-molecule studies has led to corresponding improvements in analytical tools. Multiple different hidden Markov model approaches have been proposed to deduce the states that can be occupied by a molecule from a history of transitions (*93*, *95*), either by using the Viterbi algorithm (*94*) to deduce it from the interpolted transition probabilities or by using thresholding approaches, essentially filtering the signal. Research is ongoing, especially to reduce the need for *a priori* assumptions for the fitting and, for example, determining the number of different states required to simulate a system (*74*, *96*).

Chapter 5

DNA scaffolds as a new class of soft tools in single-molecule instrumentation

We have shown how proteins interact with ligands, and what type of information on those interactions we require. We have shown how single molecules can be manipulated with force. The final piece is a structure that would allow us to conveniently manipulate them. Those structures are single-molecule scaffolds. Our team has developed a molecular DNA construct to study non-homologous end joining, a reparation system that ligates DNA blunt ends at the single-molecule level (*97*). It appeared that a similar strategy could be pushed forwards to study association between any arbitrary molecular partners. The first demonstration was made on FKBP12-rapamycin-FRB system by Kostrz et al. (Ref. (*18*)). Since my work demonstrates further applications, will follow a presentation of this scaffold, motivated by the properties of DNA and the requirements for single-molecule force spectroscopy scaffolds.

5.1 Generalities on molecular scaffolds

5.1.1 Principles and requirements

Mechanical single-molecule manipulation requires handles so that force may be applied and measured. If only a single molecule is investigated, it is simple, only an attachment to the nanomanipulation tool is needed. For multi-component reactions, a naive protocol would simply pull both components apart, and in the best case only one dissociation event could be seen. A leash must thus be added to bridge both partners together and prevent them from separating by force traction or diffusion. Thus, molecular scaffolds are artificial chemical constructs that maintain both part-

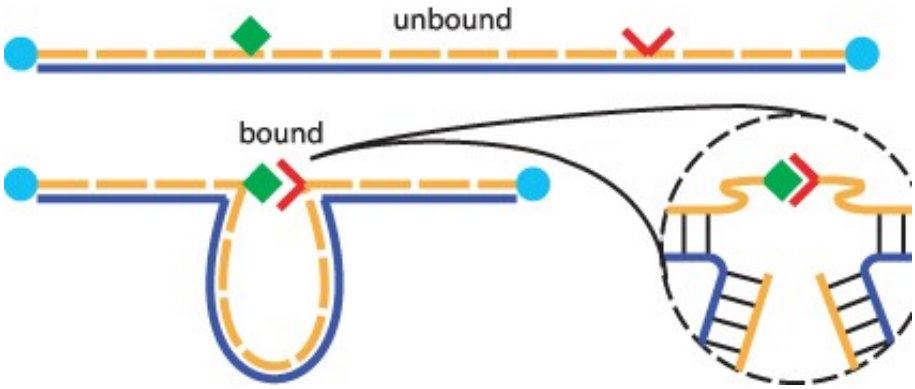


Figure 5.1: Schematic representation of a DNA origami scaffold used to investigate protein-ligand interactions, built using a single-strand DNA as a template to assemble the origami with the protein (red) and ligand (green) on it. The blue spheres can be manipulated with optical tweezers or an AFM, for example. Adapted from Ref (*98*).

ners at a chosen distance even at high force. By structurally associating a protein with a ligand, they enable association reactions to happen under low, but nonzero forces and thus repetitive experiment cycles. They maintain pairings, always associating a single protein with an identical ligand and thus can explore intermolecular variability.

Finally, experimental convenience must be a focus of scaffold design as the same scaffold can be used for different proteins by changing the attached sample. This reduces the number of different scaffolds required and affords the modularity desired for drug discovery. For a scaffold to be useful, it must be able to withstand the forces to which it will be exposed, be easy to manipulate and construct. Rigidity is a key property - low rigidity will increase experimental noise while too high rigidity can prevent the scaffold from closing and the molecules from interacting.

5.1.2 Diversity of molecular scaffolds

With the rise of single-molecule biochemistry, very different scaffolds have been tried and tested under different conditions. Polymer scaffolds, using for example PEG constructs, have been proposed (*99*). Protein-based scaffolds have been successfully utilized (*72*, *100*, *101*) for single-molecule investigations. Unfortunately, the unnatural polypeptidic sequences are complex to express and purify, and they must be re-engineered whenever the tested protein or ligand is to be changed. DNA origamis, noncovalently bound single-strand DNA oligomers that self-assemble to form the scaffold, have also been proposed (see Fig. 5.1). DNA origamis offer a much simpler assembly protocol and more modularity. However, they may result in partially mismatched pairings, and can even change conformation on repeated

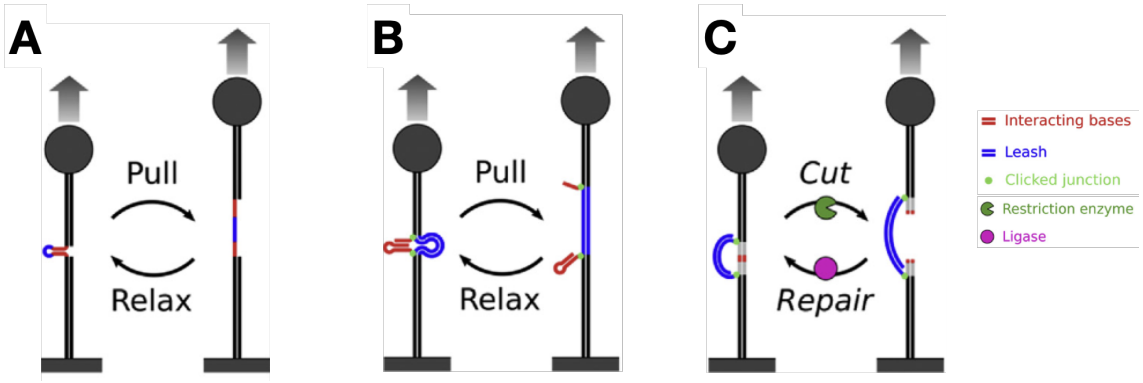


Figure 5.2: Overview of different experimental configurations using single molecular scaffolds for the study of protein action or interaction. For the sake of clarity, all selected examples correspond to measurements conducted with magnetic tweezers. **(A)** Folding/unfolding of a DNA hairpin **(B)** Association/dissociation of a DNA triplex **(C)** Digestion/Religation of a DNA restriction site. Figure adapted from Ref. (*109*).

force-loading (*98*, *102*, *103*). Under high forces, the separation of the origamis needs to be accounted for (*102*). Harnessing the ease-of-handling of DNA, covalent double-strand DNA scaffolds have been proposed.

5.1.3 Double-stranded DNA scaffolds

Double-stranded DNA (dsDNA) has been widely investigated under tension (*73*, *104*, *105*) and has been used as a handle for single-molecule experiments (*73*, *106*) but to qualify as a scaffold, the construction must allow repeated measurements. Constructs were designed to study DNA hairpins (Fig. 5.2 A) (*107*) under force, or DNA other DNA secondary structures like triplexes (Fig. 5.2 B) (*108*). In those cases, it is the structure under investigation that keeps the scaffold together. A recent new development is the addition of dsDNA leashes to maintain a scaffold cohesive, such as in (*97*) (Fig. 5.2 C) for the investigation of DNA break repairs.

Two sets of properties make DNA an ideal material for molecular scaffolds. The first is the chemical properties : resistance to chemical damage and modularity of DNA that allows an easy construction and modification of DNA scaffolds. The second are the physical and mechanical properties - rigidity means decreased noise and compatibility with a wide range of forces.

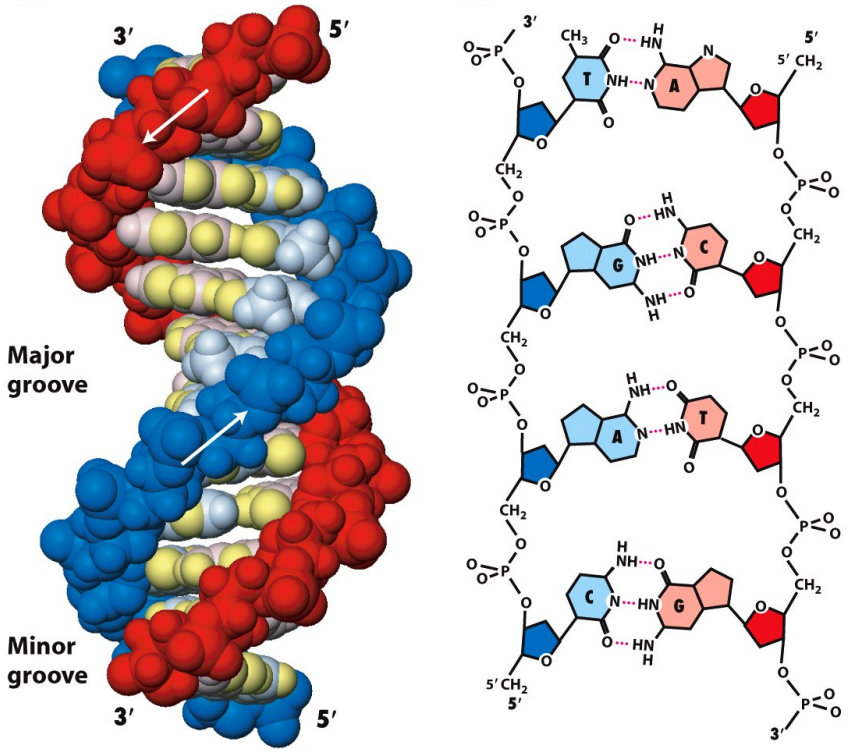


Figure 5.3: Schematic structure of dsDNA (Left) Cartoon representation of the B-helix conformation of dsDNA (Right). Schematic representation of the base pairings underpinning the structure of dsDNA. Adapted from (*112*).

5.2 Chemical properties and modularity of the DNA polymer

5.2.1 Base pairing

Desoxyribonucleic acid (DNA) is the vector of genetic information in living cells. It is a macromolecule composed of two single-strands, each a polymer assembled from desoxyribonucleotides. The complementary nature of the bases (*110*) allows the storage copy and edition of genetic information, both *in-vivo* and *in-vitro*. Its chemical, physical and biological stability has been extensively studied (*111*), as it is critical to its function of preserving genetic information and allowing its use and propagation. The numerous ways in which this information can be altered are key in understanding genetic instability, mutation, and evolution. Those properties are brought forth by unique mechanical properties that make it a perfect tool for the construction of molecular scaffolds. The aim of this section will be to show why DNA has such properties, how they can be harnessed and how DNA scaffolds are built, and why.

The interaction between each strand and the opposite is made via ionic and hy-

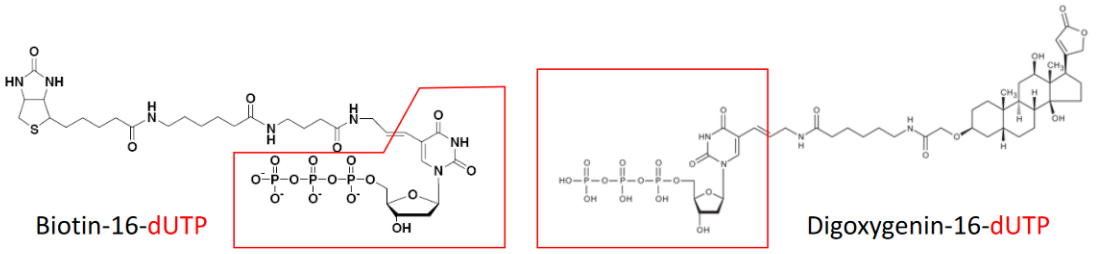


Figure 5.4: Chemical formulae of Biotin-16-dUTP and Digoxygenin-11-dUTP

drogen bonds - see Figure 5.3. The unique property of base-pairing (Adenine pairs with thymine ; cytosine pairs with guanine) is at the heart of the transmission and expression of genetic information. From a chemical point of view, DNA is remarkably stable when compared to proteins, resistant to hydrolysis and denaturation (*111*).

Single-stranded DNA polymers in solution will bind depending on sequence complementarity, and it is thus possible to assemble double-stranded DNA, or even larger complexes, from single-stranded short DNA sequences (sometimes called oligonucleotides). This is the logic behind the design of DNA origamis (*113*) or the ligation of DNA sequences using cohesive overhangs. The ease of the artificial synthesis of DNA means the design of custom DNA structures is inexpensive and fast.

5.2.2 Bioorganic modification and conjugate synthesis

To interface DNA with other types of molecules, modified DNA bases can be incorporated into strands, either during PCR or oligonucleotide synthesis. Bases modified with digoxygenin can be bound with antibodies. Bases modified with biotin can be bound to avidin or streptavidin (*114*, *115*), which can itself be bound to another protein or surface. Bases modified with DBCO [1] or azides can be used for click-chemistry to bind any kind of protein or xenobiotic (Section 7.2.1). Fig. 5.4 shows example chemical formulae of modified bases that can be used to attach a scaffold to other structural elements.

5.3 Physical and mechanical properties of the DNA polymer

5.3.1 DNA is a rigid polymer

The interaction between the two strands of DNA is achieved via ionic and hydrogen bonds - see Figure 5.3. dsDNA is further stabilised by interactions between the

[1] Dibenzocyclooctyne-amine

aromatic cycles in the successive bases (pi-stacking). Hydrophobic compaction imposed by the solvent water also increases stability, as the bases themselves are quite hydrophobic. Thus, the assembly itself is very rigid, and while it can be bent under thermal fluctuations, the effects of a bend are propagated along the main axis at a distance. This distance is the *persistence length* (ξ) of dsDNA, measured at 53 nm, or 150 bp, in our conditions(*116–120*).

The mechanical properties of DNA are strongly dependent on ionic conditions (salt concentration in the medium). We will be focusing here on concentrations used in this study and similar ones (about 100 mM chloride salts - see Section 7.3.1) which correspond to saturating ionic conditions. Ion concentrations dropping under 20 nM lead to increased persistence length and rigidity due to fewer screening of electrostatic interactions (*121*).

It can be shown that $\xi = \frac{A}{k_B T}$ with A the flexural rigidity of the material (a constant) and $k_B T$ the thermal energy (*122*, *123*). As temperature increases, thermal agitation decorrelates the orientation of successive monomers and the persistence length of DNA decreases. At zero Kelvin, the polymer would be a straight rod.

DNA remains flexible under forces or torsion ; indeed, in physiological settings it is compacted in the (prokaryotic) cells or (eukaryotic) nuclei by HU or histone protein complexes. The persistence length of 53 nm is indeed critical, as it is longer than protein complexes (histone octamers, which are relatively large protein complexes, are 6.4 nm in radius) (*112*) but shorter than cellular lengths (prokaryotic cells being on the order of 1 000 nm). This interplay between rigidity and flexibility is key in the design and usage of DNA-based molecular scaffolds, as its rigidity means lower fluctuations and less deformation under tension allowing easier transmission of forces.

5.3.2 DNA is an entropic spring

Polymers can be modelled as successions of hinged straight rods. If we suppose the rods are freely jointed, the length of a rod that best describes the polymer is the Kuhn length and, in the case of room-temperature dsDNA, it is equal to twice the persistence length, 105 nm (*124*). When under no force, the Kuhn polymer follows a random walk like a self-avoiding coil. This form maximizes entropy by maximizing the number of different configurations the entire polymer can adopt (*118*). For a polymer with a fixed number of elements, as the extension increases, the number of possible conformations decreases and more energy has to be spent to fight the thermal fluctuations. Stretching DNA requires force, and one can calculate a stiffness (a force per unit of distance) that generates an opposing force resisting

extension. This stiffness is reminiscent of a spring, with one key caveat : the stiffness *increases* with temperature, in opposition to regular springs which behave inversely. Hence the notion that polymers, and DNA in particular, are entropic springs.

5.3.3 DNA extension under force is controlled by the Worm-Like Chain model

On short spatial scales, DNA behaves like a rigid rod ; on long spatial scales, it behaves like a random coil. Unifying those behaviors under a single coherent model requires taking into account the energy associated with the bending of the polymer. The freely-rotating chain builds on the Kuhn model by associating a bending energy to the angles between segments. The Worm-Like Chain model (WLC) takes this model further by assuming the elements are infinitely small. Thus, instead of placing constraints on angles between two adjacent segments, constraints are put on the curvature of the polymer itself. This way, the polymer maintains correlations between its elements at short ranges, with a harmonic bending potential.

The force-extension relationship of the Worm-Like Chain is a complex topic that has been explored by multiple works(*116*, *125*, *126*). We have chosen to use a model that has been well-tested in the context of magnetic tweezers and published in (*117*). This model starts with a well-established approximation of the worm-like chain establishing that, for a fixed (ie. fully extended) contour length L_0 and a persistence length ξ, to stretch the polymer to an end-to-end length L, the stretching force required is :

$$F = \frac{k_B T}{\xi} \left(\frac{1}{4(1 - L/L_0)^2} - \frac{1}{4} + \frac{L}{L_0} \right) \tag{5.1}$$

This equation, established by (*127*) has the desired properties and fits very well with experiments at L small (where the force converges linerarly to 0) or $L \to L_0$ (where the Force becomes infinite) but was noted to have error up to 10 % for $L \approx 0.5 L_0$ (*125*). To increase it's accuracy, a corrective polynomial (7th degree) has been calculated by Bouchiat et al. (Ref. (*117*)), allowing for theory-to-measurement error of less than 0.01 %. This relationship is illustrated in Fig. 5.5 by a typical "force-extension" curve, which is easily readable in log space.

This model is only true if the fundamental hypothesis of the worm-like chain is accurate, that is that the number of elements in the chain is large. The equation is typically used in the context of magnetic tweezers for DNA molecules of a few thousand base pairs and more ; with a persistence length of 150 base pairs, this means that the molecules usually studied have at least more than 10 persistence

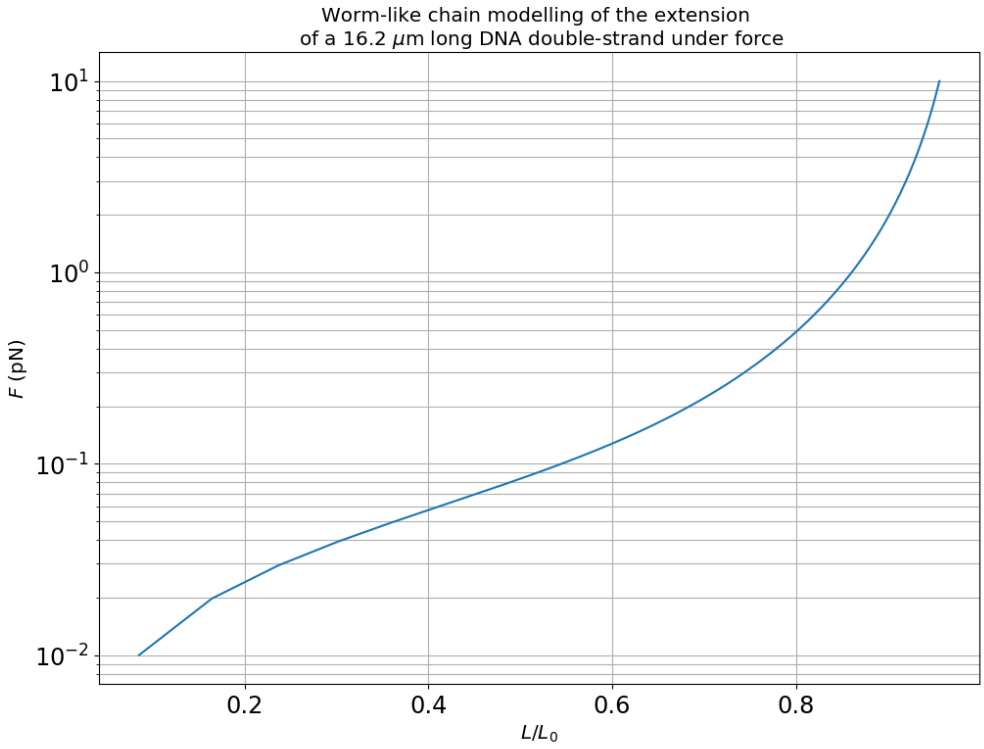


Figure 5.5: Worm-like chain modelling of the extension of DNA under force as described in Equation 5.1

lengths.

At forces higher than 10 pN, DNA experiences phase transitions which take us out of the present scope (*123*).

5.3.4 DNA is a supercoilable polymer

DNA is a double-stranded polymer : it cannot rotate freely around its own axis. This means that it can accommodate torsion when its ends are rotated one relative to the other ; this modifies the natural coiling of 10.3 bp per turn via a process called *supercoiling*. Formally, two parameters can be defined : *twist* is the number of times the double strand wraps around itself while *writhe* is the number of times the axis of a molecule crosses itself. For a molecule whose ends cannot freely rotate - either because they are attached to a support or because the molecule is circular - the sum Twist + Writhe is constant, and called the linking number. When the twist number is different from the native twist (the energy minimum), the molecule is called supercoiled, either positively (if twist is higher than native) or negatively (if twist is lower). When the supercoiling reaches a certain point, it causes the DNA to buckle and creates plectonemes on it, which transfer torsional energy into rotational energy, and cause the stretched polymer to visibly shorten, while maintaining torque

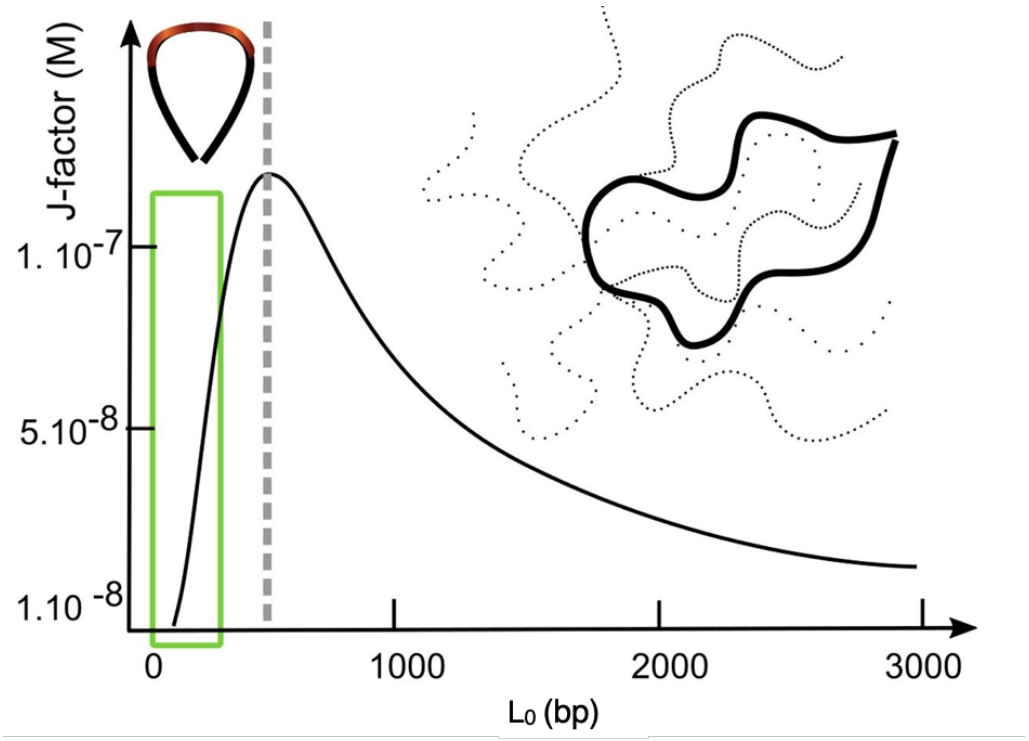


Figure 5.6: J-factor as a function of contour length L_0 given by Yamakawa-Stockmayer's model (*128*). The corresponding cartoons of DNA show the mechanical deformation of short DNA beam (red gradient) and the entropy-driven constraints on long DNA polymers with dotted lines representing the non-circularized conformations. The green highlight shows the high-torsion regime. Adapted from (*129*).

constant.

In this work, we will be working on DNA scaffolds which contain single-covalent "pivots" meaning that the DNA pieces we will be observing will not be torsionally constrained. This means we do not have access to the supercoiling tools we would normally use, and in particular we cannot use it to measure the length of a DNA molecule under observation. This will be key in Sections 6.3.3 and 9.1. This also negates any torsional effects on the interacting engrafted molecules.

5.3.5 Looping description of DNA

Another approach to DNA flexibility in solution is to consider the dynamics of DNA "loops". For experimental convenience the looping probability of a single strand is commonly expressed in terms of the local molar concentration of one extremity with respect to the other, and is called the Stockmayer j-factor j_M [M]. j_M is also related to the free energy of loop formation. If both ends of DNA carry a protein or ligand, the loop stability depends on the comparison of j_M with the interaction equilibrium constant K (*130*). To create a scaffold that favors interactions, one needs to maximize loop formation.

The loop probability as a function of loop length is a trade-off between two competing effects. Loop probability decreases when the contour length is reduced as the radius of curvature is reduced and dsDNA must be bent more strongly. On the other hand, the loop probability also decreases when the contour length increases : the entropy goes up and the extremities are diluted. For salt concentrations of 150 mM, a maximum at around 500 base pairs - 3.3 persistence lengths - has been experimentally measured in (*128*, *131*, *132*), simulated in (*132*), or calculated from WLC modelling in (*128*).j_M is evaluated at 120 nM at the maximum point. Experimental approaches on end-to-end binding probability confirm this observation, with a maximum j_M at 550 base pairs (*129*, *133*).

5.4 Junctured - DNA (jDNA) scaffolds

5.4.1 Structure of jDNA

The scaffold used in this work has been made and used previously by the laboratory and team, and in particular J-L. Wang and subsequently D. Kostrz (respectively Ref. (*97*) and Ref.(*18*)). It consists of three DNA strands, two "Shanks" and one leash (see figure 5.7). The leash is attached to the shanks with click-chemistry (Section 7.2.1) with a single covalent bond. The shanks incorporate, at their extremities, biotin and digoxygenin-bearing DNA fragments, thus allowing it to bond to streptavidin-coated magnetic beads and anti-digoxygenin antibodies on a glass surface, respectively. At juncture points in the middle of the scaffold, short 9-bp single-strand DNA tips are found. The design of those tips is voluntarily strongly modular, so that any molecule (such as a protein) can be attached to the scaffold. The molecule only has to be conjugated to a 9-base pair DNA oligonucleotide ; by complementarity between the oligonucleotide and the aforementioned tips, both spontaneously associate. An enzymatic ligation reaction reestablishes covalent bonding between the molecules and the scaffold.

This scaffold has all the properties that can be expected : it allows us to probe the behavior of a single molecular interaction, but the leash prevents the two sides from dissociating completely, thus allowing the interaction to reform under weak force. Forces can be applied on the system via magnetic tweezing : when the complex is closed, the leash is not stretched and therefore does not prevent opening. When the complex is open under low force, the leash is relaxed and enables encounter and association of proteins. When the leash is open under strong force (1 pN) the proteins are distant and cannot encounter or re-associate. All those operations are fully reversible.

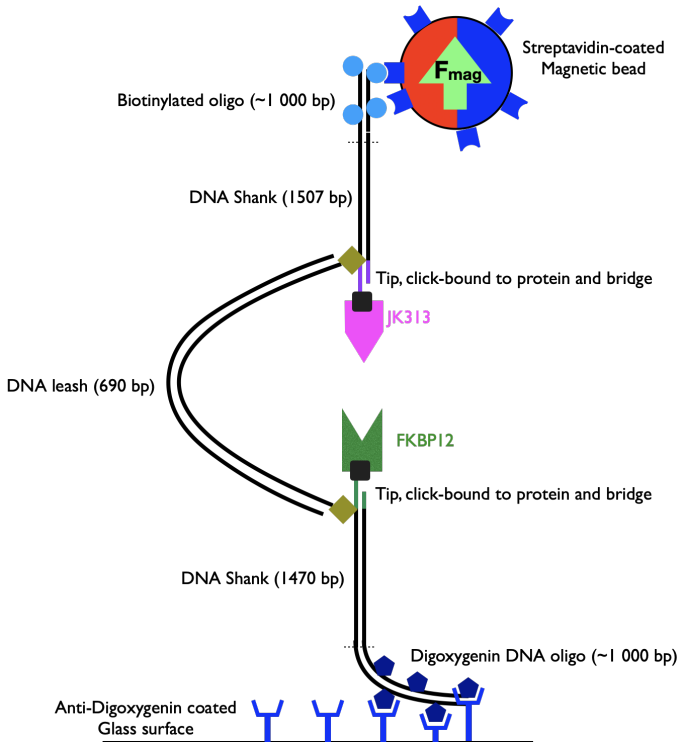


Figure 5.7: Structure of the J-DNA scaffold used in this work. See text for description of elements.

Because of the length of the shanks relative to the length of the bridge, any non-specific interaction between a loaded molecule and either the surface or the bead will display a much higher change in extension than for scaffold closing. This allows us to separate specific from non-specific binding events. Thus, this assembly allows us to manipulate individually tens of molecules in parallel using the magnetic tweezer. The scaffold in position can be washed with different buffers as needed, and can stay under observation for a week if not left to dry.

5.4.2 Assembly of jDNA scaffolds

This construction has been made very recently (*18*, *97*) based on three DNA strands : two *shanks* of 1500 bp and a *bridge* of 690 bp. The construction strategy uses a mixture of classical restriction-ligation DNA assembly as well as azide click-chemistry, and is summarized on Fig. 7.1 .
The length of the shank is chosen so as to maximize the chances of contact between the tips. As mentioned in Section 5.3.5, the j-factor is maximal at contour lengths between 500 and 600 bp. Our tips are 30 bp from the click branching points, to which must be added the 9 bp length of the oligomers and the size of the protein or ligand itself, on the order of the nm (5 bp or less). Thus, the net distance between the protein and ligand is equivalent to circa 600 bp. Of course, this is quick estimation since the pulling force, even if low, will have an effect on the j-factor, as will

the extra flexibility afforded by the click hinges. The effect of leash length on the dynamics of the scaffold would be an interesting physical question in and by itself.

Part II

Materials and Methods

Chapter 6

Magnetic tweezers instrumentation

6.1 Overview

Magnetic tweezers are device coupled with an optical microscope that can manipulate microscale objects using magnetic force. They are made of four primary components : a sample chamber, an inverted optical microscope, a magnet head with position control systems, and a temperature control system.

All of those components are controlled from a single computer, that integrates all signals and controls - in our case on the software PicoJai (commercially available under the company name PicoTwist, Paris, France).

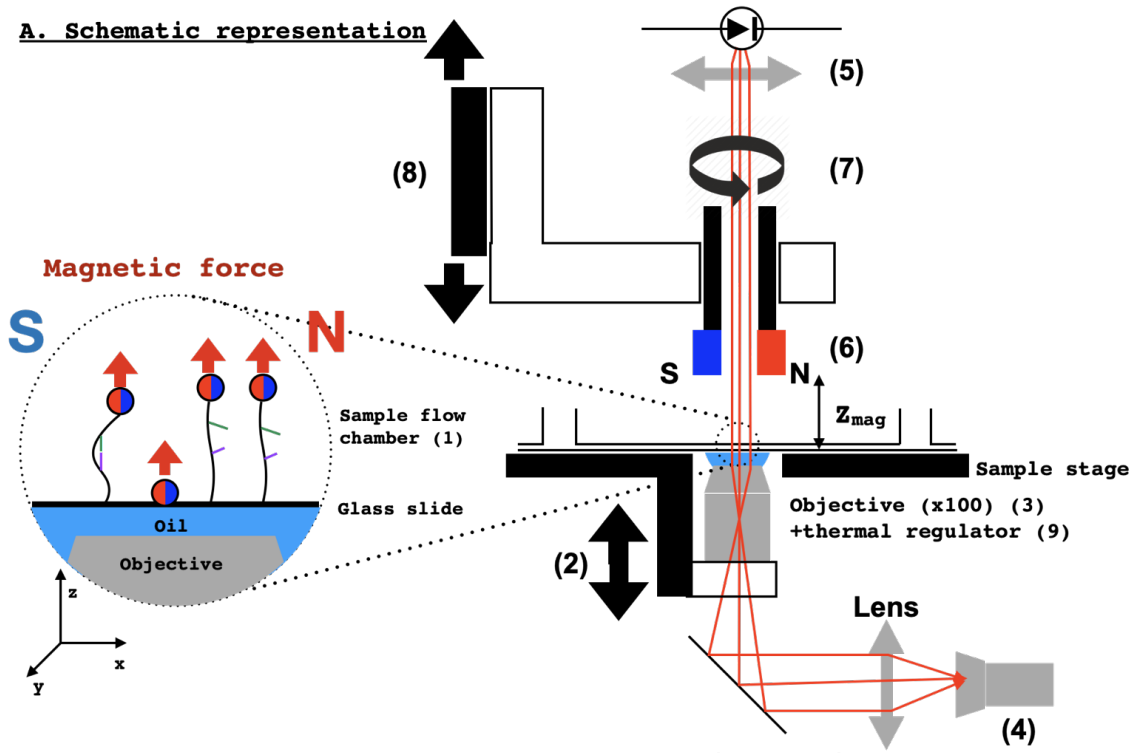
There are multiple similar microscopes in our laboratory. They differ in terms of maximum force, motor speeds and magnification, all changes that do not have a major impact on this work. For the sake of clarity, I will focus on the microscope that was used for the competition assays.

A schematic overview and photograph can be found on Fig. 6.1.

6.2 Technical background

6.2.1 Sample flow chamber

Our samples are contained in a chamber made of two glass surfaces delimiting a flow channel so that the buffer may be changed. The surfaces are coated polystyrene and functionalized with anti-digoxygenin antibodies - see Section 7.1. The chamber is delimited on the sides with a melted parafilm spacer so as to make a channel 3 mm wide and 35 mm long. Two holes on the top slide providing access to reservoirs,



Figures legend

1. Sample-holding plate with the sample
2. Piezoelectric micro-positioner holding the objective
3. Oil immersion microscope objective
4. CCD camera with convergent imaging lens
5. Blue LED light with a convergent lens providing illumination
6. Rotating head with two magnets positioned around the illumination shaft
7. Rotational servomotor for rotation of the magnet turret
8. Linear servomotor controlling altitude of the magnet turret
9. Thermal blocks for heat regulation

B. Photography

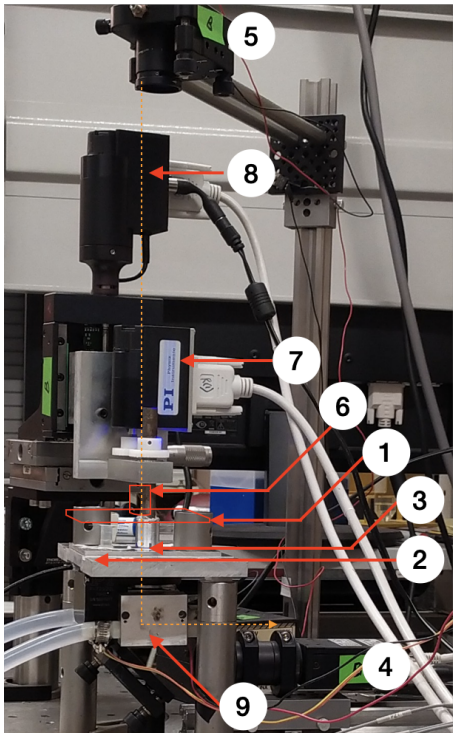


Figure 6.1: Magnetic tweezer setup

creating a chamber of volume circa. 100 μL . DNA scaffolds and captured by the digoxygenin endings.

6.2.2 Inverted optical microscope

The set-up amounts to an ordinary inverted microscope with a x100 Olympus (PlanC, 1.25 NA 100x, Olympus Rungis, France) oil-immersion objective. With an additional 180 mm focal distance tube lens, we achieve a standard x100 magnification, which allows us to see plainly the 1-micrometer magnetic beads we use.
To detect objects under the microscope, illumination is provided by a blue (in this case) or (on some models) near-infrared LED colimated through a simple lens. The light passes through the turret and between the magnets, through the sample and immersion oil onto the objective. A CCD camera (jAi CM-140 GE, jAi Ltd Europe) with a tube lens detects this light and images the field of view at a sampling frequency of 31.08 Hz. Recording is done on PC running the PicoJai software suite (PicoTwist).
Focus is controlled by a piezoelectric crystal (Physik Instrumente P-725.CLQ, Physik Instrumente, Karlsruhe, Germany). The piezoelectric micro-positioning stage has a range of motion 100 μm and a precision of 0.1 μm. The objective is mounted on top of this piezo. As the objective moves up or down, the focus can be adjusted. On some microscopes, an additional manual focus exists that can move the sample plate up or down for coarse focus adjustment.

6.2.3 Magnet heads

The magnetic tweezers *per se* are made of two neodymium-iron-boron (NdFeBr) magnets (10 mm long, 3 mm high and 1 mm thick) of opposite polar orientations affixed on a rotating turret above the field of view - the illuminating light passes between them. They create the magnetic field that exerts force upon the beads tethered to the scaffold. The magnets can be rotated using servo-motor M-136 (Physik Instrumente, Karlsruhe, Germany) and moved up or down using servo-motor M-126.DG (Physik Instrumente, Karlsruhe, Germany), thus controlling rotation of the bead and vertical applied force.

On the 1 μm diameter paramagnetic beads we use (Section 7.3.2 for details) the forces can apply forces ranging from zero (for magnets 5 mm away from the top glass slide) up to 10 pN (for magnets in contact with the top glass slide). A calibration routine allows one to know the typical force applied to the beads (see Section 9.1 for details). Typically, the inhomogeneity of mass of the bead, as well as the inhomogeneity of the magnetic material adds 10 % variation to the mean force.
The distance between the magnets and the sample, important for consistency of

force measurements, is calibrated as follows : a single glass coverslip is placed under the microscope, on the same holder. The magnet turret is then approached (by steps of 50 μm) until physical contact with the coverslip (seen on the microscope ; the objective is spring-loaded and not damaged by this contact). The position of the turret at contact is considered as the reference point for $Z_{mag} = 0$. With this protocol, Z_{mag} has a precision of 50 μm.

6.2.4 Temperature control system

To control precisely the temperature of the sample and minimize aberrations from thermal fluctuations (caused by the heat dilatation of materials in particular) we have a heat-control apparatus installed on the objective. An aluminium bloc encloses the brass objective ; thermal communication is ensured with vacuum grease. The heating bloc itself is made of two thermoelectric Peltier modules that warm up or cool down the objective bloc based on the reading of a platinum probe (F3141 Thinfilm RTD element, Omega, Manchester, U.K.) installed on it. Two heat-sinks are installed and can be water-cooled if necessary. This set-up ensures the temperature on the objective is precise to 0.01 °C.
The heating of the sample is achieved by heat-transfer through the oil and the thin glass slide to the sample. Calibration experiments have been made in which a thermal probe (same as above) is placed on a glass slide on the objective and the temperatures on the objective and surface are compared. This allowed us to establish precisely the relationship between the temperature on-objective and on-sample, so any temperature can be set on the sample with a precision of 0.5 °C.

6.3 Bead position measurement

This section describes the automated routines that executes by the PicoJai software used by our team and written by D. Bensimon, V. Croquette, T. Strick and J-F. Allemand (*104*, *118*, *122*, *123*).

6.3.1 Horizontal position measurement

The horizontal movements of the beads can be easily followed by the camera. An algorithm tracks the centroid of the bead by using the symmetry of the pattern : the centoid of the bead is assumed to be located at a point so that the correlation between both sides of that point is highest. This gives us access to the x and y positions of the bead with a 10 nm accuracy (*123*).

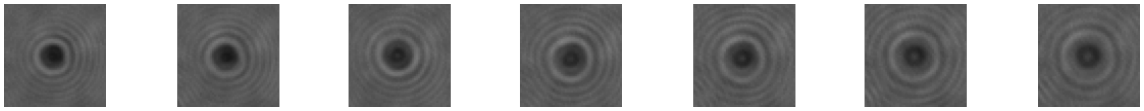


Figure 6.2: Appearance under the microscope of a single bead, at varying z altitude (or position of the objective). From left to right, the bead moves closer to the objective by 500 nm per frame.

6.3.2 Vertical position measurement

The bead's image on the camera shows a diffraction pattern which is minimal when the bead is in the focal plane of the objective and which grows the farther the bead moves from the focal plane (see Fig. 6.2). Thus, to determine the vertical position of a given bead, we can calibrate and then use this pattern as a look-up table. The objective is placed atop a piezo micropositioning stage. With an appropriate feedback loop, we have sub-nanometer precision in objective positioning.

Calibration In order to determine the vertical z movement of the bead, a calibration is first required. To calibrate, the force is increased to maximum - about 10 pN - and the piezo is simply moved across a wide position range (typically 6 μm) by 100 nm increments. At each position, the appearance of the bead is recorded, and an average diffraction pattern is established and memorized. Once the calibration is done, to determine the vertical position of the bead we observe the diffraction pattern and compare it with the look-up table constructed from the stack of calibration images.

Figure 6.2 shows typical calibration images obtained for a bead. Notice that, as the focus moves, the rings grow wider and wider - thus, the software can interpolate bead altitude as a continuous function of the diffraction pattern and obtain altitude measurements, and reach a z-determination precision up to the nm (*134*).

Drift correction To account for low-frequency mechanical drifts in the system (thermal dilatation of microscope components, for instance) all measurements are made in reference to a bead glued to the glass surface, which provides us with a zero position reference inside the sample chamber.

Absolute vertical position For situations where the absolute vertical position with respect to the glass surface are required, it is possible to obtain it by rotating the magnets until the DNA is completely supercoiled and the bead touches the glass surface, which gives the z = 0 position. This can only be done with DNA constructs

that are topologically constrained (entirely double-stranded) so this cannot be done with jDNA, which exhibits single covalent bond attachments at the shank junctions.

6.3.3 Measuring Force in magnetic tweezing experiments

The pendulum analogy

Let us consider a dsDNA molecule attached to the glass coverslip on the bottom and to a magnetic bead at the top. Three forces are exerted on the bead : the vertical pull from the magnets, the tension from the DNA molecule, and the Brownian fluctuations from the water solution, a force of average intensity 0 but with non-zero standard deviation resulting from the energy drawn from the thermal batch via the equipartition theorem [1]. The force generated by the magnets is vertical because that is the orientation of the magnetic field gradient. The Brownian fluctuations will displace the bead from the equilibrium center position. As can be seen on figure 6.3, we can assimilate the situation to an inverted pendulum and project the forces onto an axis tangent to the direction of bead oscillation. The friction caused by the solvent should be taken into account ; this will be addressed later.
By analogy with a harmonic pendulum, we pose that the DNA of contour length L_C has extension L (which depends on the force, see Section 5.3.3) and that the bead is displaced from the central position by a distance δx causing the DNA to tilt from equilibrium by an angle θ.

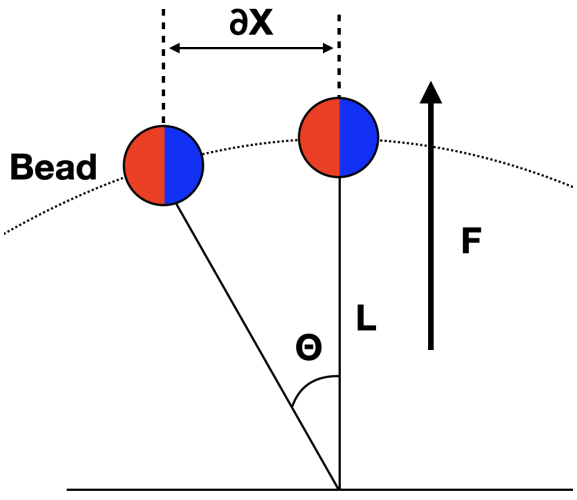


Figure 6.3: Geometry of the pendulum approximation

The vertical force generated by the magnetic tweezer projects onto two orthogonal axes, one parallel to the DNA and one perpendicular to the DNA and which serves as a restoring force to return the bead to the center equilibrium position:

$$F.\sin\theta \approx F.\theta = F.\frac{\delta x}{L} \qquad (6.1)$$

This is a simple spring equation for a spring of rigidity F/L. Thus, the energy of the pendulum when deviated δx from its resting position is $\frac{1}{2}\frac{F}{L}.\delta x^2$. This energy is provided by the random collisions with molecules in the environment, and

[1] The weight of the bead and the buoyancy, which are on the order of 10 fN in total, are constant and can be directly integrated into the resulting vertical magnetic force.

according to the equipartition theorem this energy is equal to $\frac{1}{2}k_BT$ per degree of freedom, so 1 when projecting onto a single axis (*117*). We thus reach the following equation :

$$F = \frac{k_BTL}{\delta x^2} \tag{6.2}$$

With k_B the Boltzmann constant, T the absolute temperature, L the length of the DNA molecule and δx^2 the variance of the bead position in x. If we consider a sample DNA molecule of 3 000 base pairs in length (crystallographic length of about 1 μm) at high forces, we will see a position variance of about 20 nm, corresponding to forces on the order of a 10 pN. On the other hand, at low forces, the oscillations will be on the order of 100 nm, and the forces will be in the hundreds of femtonewton range.
Thus, the measurement of the force exerted upon the molecule by the magnetic tweezer can be reduced to the measuring of the horizontal displacements of the bead and the extension of the molecule connecting the bead to the glass surface. This is true, of course, of the oscillations parallel to the direction of the magnetic field - that is, the projection of the bead on an equipotential line in the magnetic field.

Fourier analysis

This simple explanation is however not sufficient. As the microscope is subject to drift - thermal in particular - and the camera's sampling rate can cause unwanted filtering effects. At lower oscillation frequencies, the dampening effect of friction by the solvent also has a significant effect. At low forces in particular, Equation 6.2 diverges strongly from observation. Thus, it is useful to analyze the Brownian motion not in real space, but in the frequency space. A very brief overview is presented here - readers interested in more details can consult (*118*, *123*).
A cutoff frequency ω_c can be measured in the experimental frequency spectrum of the bead. This frequency is the highest frequency that is sampled correctly by our measurement system. Indeed, Brownian motion, by definition, has constant components in Fourier space - all frequencies are equally represented. In reality however, because of camera sampling and hydrodynamic dampening, this is not true. Taking into account the viscous drag on the bead, one finds the theoretical value of ω_c as :

$$\omega_c = \frac{F}{\gamma\eta R} \tag{6.3}$$

With F the magnetic force, γ the viscous drag coefficient and η the dynamic viscosity of water (*123*). We can use this to evaluate the time that is required for

two measurements to be independent : the characteristic time of an oscillation is $T_c = 1/\omega_c$: this is the time motion takes to be dampened, and we thus get one independent measurement every T_c time interval. This is essential to determine the precision of our force determination : we need to quantify the uncertainty on δx^2 to have the standard error on the mean of F.

Coming back to the impact of the camera sampling rate, our camera has an acquisition rate of 31.08 Hz. Therefore, we cannot use any frequencies higher than this to calculate $< \delta x^2 >$. Moreover, frequencies that are close to the camera sampling rate also lose in relevance, as sampling bias can strongly affect them. (*123*) shows that, for a camera sampling frequency equal to 10% of the cutoff frequency, the observed spectrum intensity is only equal to 0.97 of the theoretical value. We usually take this convention that we can only measure a force reliably if the cutoff frequency is lower than one-tenth of the imaging frequency, so 3.1 Hz in our case.

The cutoff frequency can not only be expressed analytically, but can be measured as well. Calculating the power spectrum of an oscillating bead at a given force (Fig. 6.4) one can see it remains constant at lower frequencies but decreases in the higher frequencies, because of hydrodynamic dampening (*123*) and camera imaging biases (*118*, *123*). The frequency after which the amplitude decreases exponentially is the cutoff frequency ; it is the highest frequency we can analyse reliably.

The uncertainty of a single measurement of the position, since x follows a Gaussian distribution, is equal to the measurement itself ; thus, the S.E.M. on the measurement of $< \delta x^2 >$ is equal to $< \delta x^2 > /\sqrt{n}$ where n is the number of independent measurements we can do, which requires nT_c observation time to obtain. As a general rule, we select observation times so that $n > 100$.

One major issue with this approach is the time required to measure a force. As our cutoff frequencies typically range in the 0.1 Hz - 1 Hz range, an observation requires between 100 and 1000 seconds to complete.

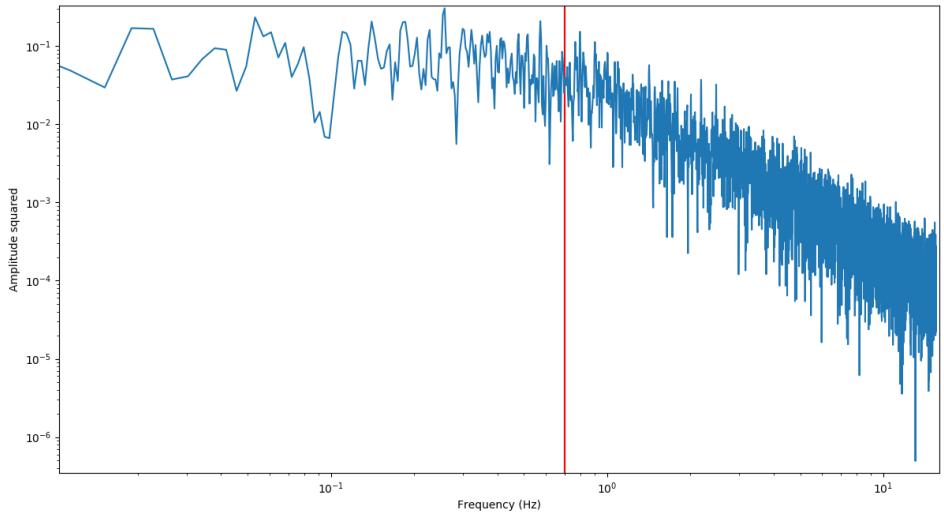


Figure 6.4: FFT power spectrum of the horizontal fluctuations of a sample time-trace at 50 fN. Notice that at lower frequencies, the power spectrum is constant but diminishes after 0.7 Hz (red line) as the sampling and hydrodynamic dampening move it away from idealized brownian fluctuations. The observed cutoff frequency here is thus 0.7 Hz.

Chapter 7

Single-molecule junctured DNA observation protocols

This section provides a technical description of jDNA magnetic tweezer experiments, from surface preparation, jDNA construction and protein attachment to injection, observation data acquisition and analysis.

Overview An experiment is done by flushing the channel of a polystyrene/Anti-Dig coated surface with a suspension of magnetic beads with DNA scaffolds ("washing"), waiting for them to settle (and the antibodies on the surface to bind the digoxygenin in the DNA) and then removing the unbound beads and scaffolds. The bead movement can then be observed and analysed.

All chemicals were from Sigma-Aldrich, Merck, Darmstadt, Germanypowders, 99.5 % grade (BioUltra line when applicable) and were used without further purification.

7.1 Surface preparation

We use glass-activated, polystyrene-coated surfaces with anti-digoxygenin antibodies on them. Glass surfaces are from Menzel-Gläzer, Fischer Scientific, dimensions 24 × 60 mm. The bottom surfaces are kept as is, while the top surfaces are drilled using a sand-blaster with two 2 mm holes to create openings in the channel.
The glass surfaces are first cleaned with ethanol and water to rid them from the largest impurities.This is followed with a sonication step, 10 min. in absolute Ethanol and then two 15 min. intervals in potassium hydroxide, 1M. Between each bath, the surfaces are rinsed in MiliQ water. The surface are then left one hour in MiliQ water and the water is changed every 15 minutes.
The surfaces are then dried under a flow of pressurized nitrogen and placed for 90s

in a plasma oven (PDC-002 from Harrick Plasma, Ithaca NY, United States) in air. They are then coated with primer (Adhesion Promoter AR 300-80 new from AllResist, Strausberg, Garmany) on a spin-coater and baked for 5 minutes at 110°C on a hot-plate. They are then coated again with 0.5 % m polystyrene dissolved in toluene. The experimental chamber is then assembled with one bottom surface, one perforated "top" surface and parafilm walls melted at 60°C.
To coat the surface with anti-digoxygenin antibodies (Unconjugated polyclonal sheep purified immunoglobulin IgG, Roche), the chamber is then filled with a 10 μg/mL anti-dig solution in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) and incubated overnight at 37°C. Finally, the chamber is blocked with a solution of 100 mM KCl, 20 mM HEPES pH 7.8, 5 mM $MgCl_2$, 0,1 %m Tween 20, 5.0 mg/mL BSA and 2 mM DTT for 90 minutes at 37°C. The buffer is then replaced with mTOR (cf. 7.3.1); the surfaces can then be stored at 4°C.

7.2 Junctured DNA preparation

7.2.1 Click-chemistry

"Click" chemistry is a class of organic reactions creating heteroatom links (C-X-C) between organic molecules. They are called "click" as their core principle is to attach two "building blocks" of a molecule. Click reactions share the common principles of being modular, high-yield, stereo-specific and usually not requiring solvents other than water and purification other than crystallization. They are also very strongly thermodynamically driven - the reactions are sometimes called "spring-loaded" because of the very energetic nature of the reagents. (*135*, *136*)
The reaction used in our case is the Strain Promoted Alkyne-Azide Cycloaddition (SPAAC) conjugation of dibenzocyclooctyne (DBCO) with an azide linker - 4-azido-l-phenylalanine, or AzF - to form a stable triazole. This reaction is fast, irreversible, total and specific. It can be done in water with no specific catalysts (notably, no cytotoxic copper unlike other click reactions). Compounds can be bought from suppliers in ready-to-use kits (*18*, *137*).

7.2.2 Scaffold synthesis

The scaffold used in my experiments was prepared by D. Kostrz and M. Follenfant according to protocols given in Ref. (*18*). The general workflow is depicted in Fig. 7.1. Relative to the reference, the primers used for the construction were altered, resulting in a symmetrical construct rather than the published asymmetric one. All tips and primer sequences are given in (*18*), with the exception of J1 and J2 which are given in Fig. 7.2.

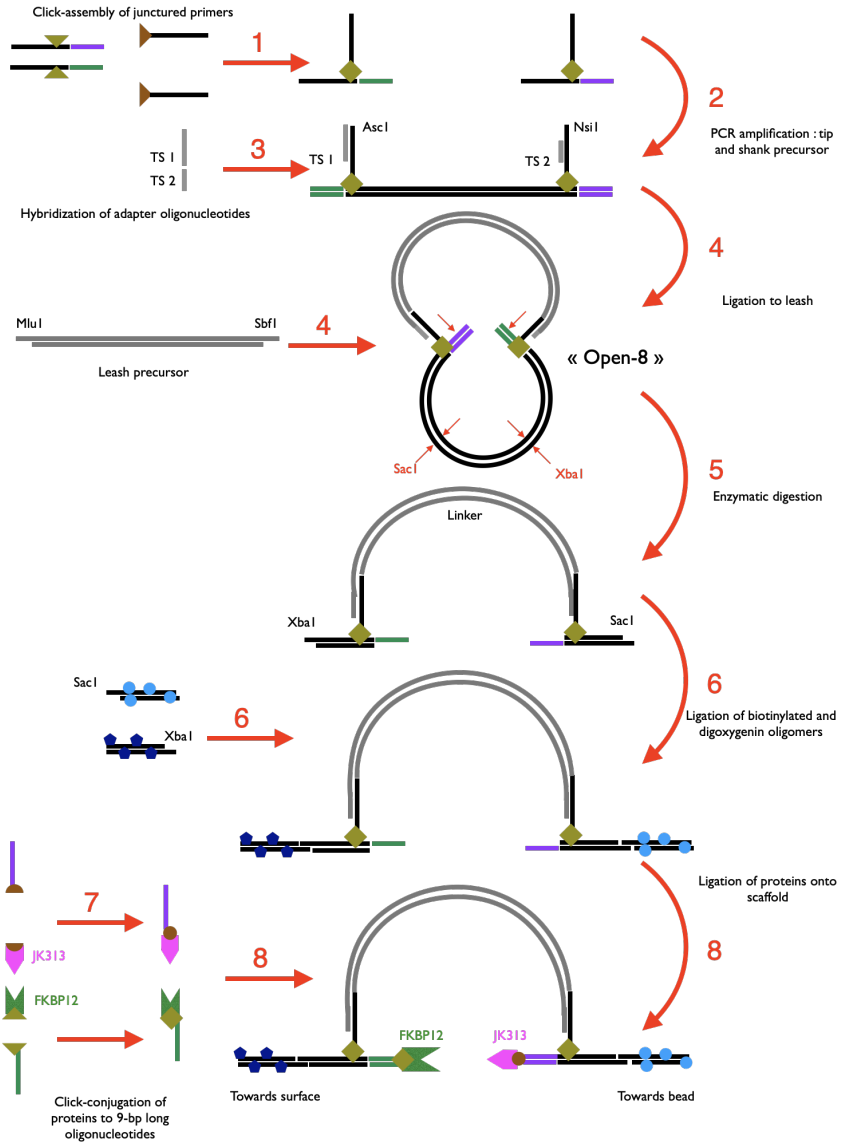


Figure 7.1: Assembly of the J-DNA JK313/FKBP12 scaffold used in this work. 1.The two Primer-J junctions are obtained by click chemistry between the Oligo-TS and the Oligo-L strands. The resulting sequences are in Fig. 7.2. 2.The precursor of the "tips" and "shanks" is synthesised via PCR, using the oligos TS1 and TS2 as primers and the charomid 9-5 Δ SbfI plasmid as template. 3.Complementary oligonucleotides are hybridized onto the leash precursors. 4.A PCR amplification step is run on the charomid 9-5 Δ SbfI template (*18*) to create the leash. Digestion by MluI and SbfI restriction enzymes (New England Biolabs) creates the ligation sites. The leash is ligated onto the precursors using the created open restriction sites. This creates the "Open 8" configuration. 5.Enzymatic restriction digestion by enzymes SacI and XbaI creates two separate shanks as well as linearizes the scaffold. Also creates overhangs for ligation of biotin and digoxygenin oligomers. 6. Sticky shank ends containing dUTP-Biotin and dUTP-Digoxygenin (Roche) were prepared according to (*18*) and digested with SacI and XbaI restriction enzymes (New England BioLabs) and ligated on the scaffold, creates an empty scaffold. 7.Proteins or ligands are conjugated to 9-bp long oligonucleotides using click chemistry. 8.Using complimentarity between those oligonucleotides and the available sites on the scaffold, the protein-oligomer conjugates are ligated in position onto the scaffold.

Figure 7.2: Final sequences of the junctured primers used to build the jDNA scaffold, step 1 of Fig. 7.1.

7.2.3 Conjugation of ligand and protein to the scaffold

The conjugation of the JK313 drug to the AzP2 oligonucleotide was performed using the CuAAC Biomolecule Reaction Buffer Kit from Jena Bioscience (Jena, Germany), following the guidelines provided by the manufacturer. This protocol was implemented by M. Follenfant and C. Gosse.
A premix containing copper sulfate at 33.33 mM and the tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine ligand (THPA) at 166.66 mM was first freshly prepared by adding 1 volume of $CuSO_4$ at 100 mM in water to 2 volume of THPA at 250 mM in water, resulting in a blue solution. In a second tube were mixed 66.4 μL of sodium phosphate buffer at 100 mM (pH 7) with 16.6 μL of AzP2 at 0.3 mM in water (5 nmol, 1 equiv) and 1 μL of JK313 at 20 mM in DMSO (20 nmol, 4 equiv). Then, 6 μL of the $CuSO_4$ / THPA premix were added (0.2 μmol, 40 equiv. / 1 μmol, 200 equiv. , respectively) ; the tube was vortexed and spun down briefly. Finally, the reaction mixture was completed with 10 μL of sodium ascorbate at 1 M in water (10 μ mol) and, after vortexing, incubation took place at 37 °C for 1 h.

To separate the oligonucleotide from the click reagents that have been introduced in excess we precipitated it by adding 10 volumes of lithium perchlorate at 3 % in acetone, this solution having been precooled at - 20 °C. The tube was placed in the freezer for 30 min before being centrifuged at 4 °C, 10 000 rpm for 5 min. The supernatant was removed, replaced by 1 mL of acetone, also precooled at - 20 °C, and the tube was again centrifuged at 4 °C, 10 000 rpm for 5 min. The supernatant was discarded and the pellet dissolved in 100 μL of water. A whitish suspension was obtained, which was clarified thanks to an extra centrifugation step. The oligonucleotide contained in the aqueous supernatant was further purified by repeating the above precipitation and washing procedure. After solvent evaporation the oligonucleotide contained in the pellet was dissolved in 4 mL of water.

This solution was injected on a Sep-Pak Classic C18 Cartridge (Waters, Milford,

MA) that had been first wetted with 10 mL of acetonitrile and then equilibrated with 10 mL of water. Contaminants were washed away using 10 mL of water and the oligonucleotide was recovered using 2 mL of acetonitrile. Solvent was finally evaporated using a SpeedVac vacuum concentrator and the conjugate was dissolved in water.

Conjugation of FKBP12-M0AzF DBCO-P1 was performed as in (*18*). Likewise, the binding of the JK313-P1 and FKBP12-P2 oligomers to the jDNA was performed by M. Follenfant as described in (*18*) and Section 7.2.2.

7.3 Sample preparation

7.3.1 Experimental buffers

All experiments are conducted in a buffer we named "mTOR" (100 mM KCl, 20 mM HEPES pH 7.8, 5 mM $MgCl_2$, 0.1 % vol. Tween 20, 0.5 mg/mL BSA, 2 mM DTT).
FKBP12 is stored at -20°C in protein buffer (50 mM Tris equilibrated to pH 7.5 with HCl, 0.3 M NaCl, 10 % vol. glycerol) to minimize freezing damage. Dilutions are conducted in the same buffer.
For experiments in which FKBP12 serves as the competitor, final injections in chamber are prepared by mixing 40 μ L of protein solution in protein buffer with 1 960 μL of mTOR, so that the glycerol concentration in the experimental chamber is constant throughout the experiment at 0.2 % v/v glycerol (even in the absence of competitor).
Competitor drugs shown on Fig. 1.4, kindly provided by Hausch et al. (*20*), were kept at -20°C in 100 % DMSO.
When those drugs are used as competitors, injection solutions are made in 0.05 % v/v DMSO in mTOR, likewise constant throughout the experiment, even in absence of competitor.
All experiments are conducted at a temperature of 28°C in-sample, unless otherwise noted.

7.3.2 Injection

The experimental chambers have two wells, one on each side of the 2 mm channel, of approximate capacity 500 μL . A flow can be created by filling one of the wells and letting the liquid through by gravity.
5 μL of 10 mg.mL^{-1} beads (Dynal MyOne Dynabeads Streptavidin C1 from Sigma-Aldrich, Merck, Darmstadt, Germany) are mixed with 200 μL of mTor buffer, then

agglomerated by a magnet to remove the buffer and resuspended in 35 μL of mTor buffer. Those beads are then added to 1 μL of jDNA scaffolds as described in 5.2 carrying the protein FKBP12 and the drug JK313, in solution in 50 mM NaCl at 50 pM. The mixture is incubated for 2 min at room temperature and injected onto the surface and left for 10 minutes to bind. Unattached beads are washed out with 2 mL of mTor buffer under zero magnetic force.
To avoid contaminating the field of view with free beads, all washing steps are made in a direction opposite to the direction of the bead injection.

7.4 Force pattern application

There are two primary modes of experiments used with jDNA. We described them as passive - ie. constant-force, and active, or force-step.

Passive experiments correspond to a passive regime where the force is left constant, at approximately 100 fN (always less than 300 fN). The movements of the bead are recorded with no intervention.
As we will show in Section 8, we aim to interpolate the zero-force parameters. Thus, the exact force at which the scaffolds are observed will not be a central factor in the analysis. The force we use is thus a compromise between the need to keep force high enough to distinguish between open and closed conformations, and the need to have as many association reactions as possible to increase the number of collected data points.

Active experiments, corresponding to the force-cycle regime, allow experimentation at higher forces. However, at those forces, scaffolds cannot close spontaneously (at least not at reasonable frequencies). Thus, the experiment is conducted by alternating low-force (typically under 10 fN) and high-forces (on the order of the pN). At low forces, scaffolds can close, tough we cannot observe that change. When the force is increased, the interaction can be broken after a while, and one can observe a two-step increase in bead height, one as the closed scaffold is extended and a second one, later, once the binding is broken and the scaffold deploys at full length.
The duration of every cycle is chosen so that 99 % of interactions are finished before the cycle is ended, determined in prior experiments .

7.5 Titration experiments

7.5.1 Scaffold selection

On the entire surface, a single field of view is chosen. About 60 individual beads are present on a well-chosen single field of view. However, not all scaffolds are capable of interacting since the yield of the ligation of oligonucleotides to the scaffold is not 100 % (Section 7.2.2).
The scaffolds are therefore first assayed for their capacity to interact. This is done by a force-cycling experiment between low (less than 10 fN) and high (1 pN) forces for 20 s each. Only functional scaffolds will display a two-step increase in extension on the force increase ; on average, only about 25 % of scaffolds (circa 15 per experiment) display an interaction. Only those will be analysed.

7.5.2 Measurement protocol

The chosen scaffolds are monitored at a constant average force of about 100 fN (magnet head 2.5 mm away from the sample, see Section 6.2.3) for half a day. Depending on the experimental situation (daytime or nighttime experiments, weekend recording, etc.) a recording lasts for 6 to 24 hours. Nighttime experiments are typically longer, and while we strive to make our acquisitions last for 8h, sometimes this is not practical in our working shifts.
Using the protocol described in Section 6.3, the vertical position of each bead is acquired over time at the imaging frequency of the camera (31.08 Hz). The resulting $z(t)$ traces for each scaffold are our data, and allows us to observe transitions between the open and closed conformations of the scaffold. The 100 fN force has been selected for it offers a good compromise between low force (minimal effect on the interaction, see Section 9.1.1) and being high enough to keep experimental noise low enough.

7.5.3 Selection of competitor concentration

The same sample, with the same field of view, is kept for a series of measurements - typically 10 - at different competitor concentrations. Thus, it is used for a whole week, with two different measurements per workday (one in the day and one overnight).
The competitor concentrations chosen are 0 nM and a range of values both above and below the estimated K for the assayed reaction between the scaffold and the competitor (either based on preliminary experiments or on data provided by Pomplun et al. (*20*)). Typical points will be at regular increments of 0.25 K, starting at

0. There is a trade-off between the need to cover a large competitor concentration range - to have a clear reading of the induced changes - and the fact that higher competitor concentrations reduce the number of interactions and thus increases the experimental time.

7.5.4 Competitor concentration change

The first experiment is conducted without any competitor. As we cannot guarantee that a wash step will be perfect, we need to keep increasing our competitor concentration as to minimize the effects of an incomplete wash. For every condition change, the chamber (approximate volume 100 μL) is washed through with 2 mL of appropriate solution. The solution that is pushed out is discarded. During the wash, the force is increased to a maximum (approximately 10 pN) to minimize the chances of beads being ripped away or stuck to the surface with flow.
A control experiment was performed to assay the thoroughness of this change. After a titration experiment on [FKBP12:JK313] jDNA scaffold using FKBP12 as a competitor, the experimental solution at 60 nM FKBP12 was exchanged with a 0 nM solution. The scaffold dynamics were observed as usual, and used to estimate the true concentration of FKBP12 in the chamber by comparing them to the results from the titration. From those observations, it appeared that the experimental chamber contained FKBP12 at circa. 5 nM. The concentration change protocol is not very effective, with about 10 % of the original solution remaining after a change. This comforts us in our decision to only work with increasing concentrations of competitor. The concentration change could bear improvement, such as with a microfluidic flow system (*138*).

7.6 Data analysis

7.6.1 Principles of the conformation detection algorithm

A single measurement requires about 8 hours to be completed for every concentration point. Manual curation of the data is possible but strongly time-consuming and would not allow us to have as many data points as we now have. As we will see from the results, this would have been problematic.
The objective was to write code that would allow us to automate the process as much as possible. After consideration, we have chosen to use HMMs for the reasons outlined earlier, and not transition detection, a second best choice. We used, for the bulk of the work, a library based on the popular sckit-learn Numpy package, HMMLEARN (*139*).

Terminology

A point of terminology must be addressed first. In our experiments, as will be described in detail in Section 8, the scaffold can be in two observable conformations : open and closed. However, the open conformation can correspond to two different chemical states : either the molecule on the scaffold is unbound or the scaffold is bound to a competitor molecule in solution. We attempted at first to use an HMM detection algorithm to distinguish the hidden states, but that approach never gave us satisfying results, possibly because the amount of data collected was insufficient to properly establish the durations of the two hidden states.
Therefore, the analysis will be split in three major steps. In the first we analyse the raw data : the vertical position of the bead over time $z(t)$. In the second step, we use the HMM algorithm to digitize the trace and determine the conformation of the scaffold (Closed, C or Open, O) over time [1] . Finally, from the conformation trace we extract the durations of open and closed events : those durations are analyzed and used to determine interaction properties according to Section 8.

7.6.2 Raw data processing

The raw data - $z(t)$ for each scaffold - is imported from xvin ; the following steps are implemented in python :

1. Subtraction of z-position of a reference bead

To account for experimental noise (thermal dilatation of the microscope parts, vibration caused by the water cooling pump, slight light fluctuations, etc.) we subtract the z-position of a reference bead from the z-position of every bead of interest. The reference bead is chosen by the experimenter from those stuck on the surface and that display the least movement.

2. Removal of artefacts

Experimental conditions and tracking software errors sometimes make certain data points unusable. They are removed at this point. Two types of artifacts are generally encountered : points outside of the vertical range of measurement are removed automatically (their time positions are kept empty) and time periods with tracking errors (for example because the tracking algorithm lost tracking of the bead) are recognized visually and removed by hand. When patches of points are removed by hand, a buffer time of 500s is removed on either side so as to not contaminate relevant

[1] Notice, therefore, that the output of our HMM analysis are not states, but conformations.

data (this duration was chosen arbitrarily as largely superior to the characteristic open and closed durations). Depending on experimental conditions, this can affect between 0 and 50 % of a measurement. Measurements that have less than 6 hours of usable data are discarded.

3. Window filter application

The semi-processed data we obtain is relatively noisy in the sense that a typical histogram of z-positions for a single time trace will not yield two separate Gaussians. This explains why we cannot use a simple thresholding algorithm and must use the more powerful HMM. In addition, in a given conformation, the fluctuations of the bead around its mean altitude do not appear to be Gaussian ; rather, their distribution is skewed downwards. This is a well-known phenomenon in our technology, which we believe is due to non-specific electrostatic interactions between the antibodies adsorbed on the surface and the streptavidin coating the bead [2]. Unfortunately, this does interfere with the HMM algorithm which can sometimes detect those downwards bursts as conformation transitions.
This problem is solved by averaging the data first using a 50-point wide (1.6 s) rectangular window filter. This reduces the standard deviation of our data in either state from about (depending on the exact trace) 67 nm to 22 nm.

4. Removal of long-term drift

While traces are stable on the scale of tens of minutes, changes in ambient temperature can cause slight dilatation of the instruments and induce thermal drift, even if the sample itself is kept at a constant temperature. This is normally solved by subtracting the altitude of a reference bead stuck on the surface from $z(t)$ (Step 1). However, even with this precaution some very slight long-term drift remains, possibly as a result of ambient light changes and dilatation of the glass surface. Those changes are minimal in amplitude - on the order of 10 nm - however, they can still perturb the HMM algorithm. To solve this, a 10^th order polynomial is fitted to the entire trace and subtracted from $z(t)$. This step removes the very long wavelengths and stabilizes the trace so that both open and closed states remain at the same altitudes over time. Since all we are interested in is relative, not absolute vertical displacement, this has no consequences on the results. We obtain the blue signal on Fig. 7.3.

[2] This is further corroborated by the fact that these interactions are reduced in number and magnitude if the bead's tether is longer, and they manifest in short, discrete "bursts" towards the surface (see Fig. 7.3, orange boxes and arrows)

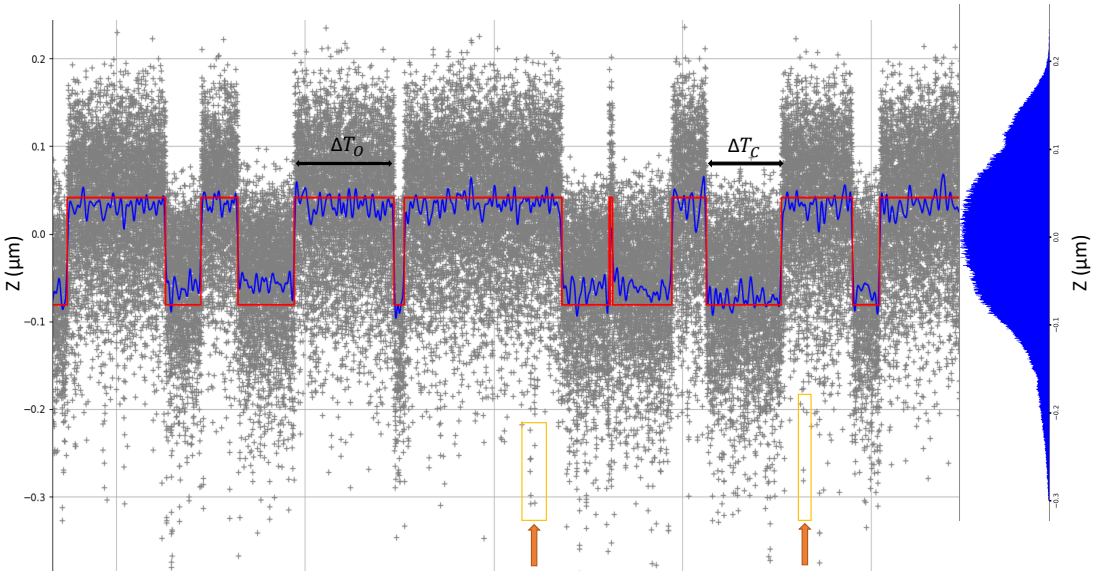


Figure 7.3: An example time trace of a [FKBP12:JK313] jDNA scaffold under 100 fN force, no competition. Gray crosses are the raw data points ; the blue curve is the filtered data trace used for state assignation and red curve is state assignation. Notice the downwards "spikes" (orange arrows). On the right : histogram of positions over the entire data trace (8 hours)

7.6.3 Conformation determination

The automated HMM learning routines that determine the conformations do not work well if one of the conformations is over-represented in a sample, in practice by a factor of more than 10:1. This typically happens if the number of interactions in a given sample is low.

5. Determination of the learning region

To get around this problem, we restrict the section of the signal trajectory used for learning by focusing on a region where that the number of open and closed states is relatively balanced. At first, we did this manually. This was later improved by running a preliminary Viterbi algorithm on the entire time-trace, which determines the region of time where the states are the most balanced. This area is then used for the second HMM learning step.

6. HMM learning

We run the HMM Viterbi algorithm on the time region found previously. We expect two distinct conformations so we used a two-state algorithm.

7. Assigning conformations

Once the learning is complete, we assign a conformation (open or closed) to every time point. We visually verify every HMM interpolation for major defects. The data can then be separated by conformation.

8. Event extraction

For the remainder of the analysis, we extract the events from our time-trace. An event is simply a length of time during which the scaffold is in one conformation. For each event, the average z position $< z >$ and standard deviation are recorded as well as conformation (open or closed) and its duration Δt.

7.6.4 Event analysis

What follows here is a brief discussion of technical observations made on the results provided by the algorithm. All figures are taken from a single scaffold in a measurement with no competitor.

Event $< z >$ position

To verify the quality of the conformation assignation, we traced a histogram of the average height $< z >$ of each event. The results are plotted on Fig. 7.4. We may see that while $< z >$ correlates strongly with the conformation (open or closed) of the scaffold, there are some events with intermediate $< z >$. We then plotted $< z >$ in function of the event duration Δt (see Fig. 7.6). We see that the events of intermediate altitude are in fact very short events. Since the movement of the bead is subject to inertia, transitions between the two conformations are not instantaneous, hence the average position during short events is different from the average position for longer ones. We see that the $< z >$ of open events only becomes distinct from the $< z >$ of closed events for events lasting 3.5s or more.
To avoid mis-classification of shorter events, we decide to not use the events shorter than 3.5 s when determining characteristic times. This threshold of 3.5s is indicated by the dotted line in Fig. 7.6. The height distribution of events longer than 3.5 s is shown on Fig. 7.5, and we can indeed see that those events have much more coherent heights. Excluding short events also has the secondary benefit of removing any artifacts caused by the limited sampling frequency (31.08 Hz).

Event durations and characteristic times

One of the main foci of this work is the time a protein-ligand complex spends in interaction. As noted in 2.1.4, the duration of first-order reaction is distributed as a

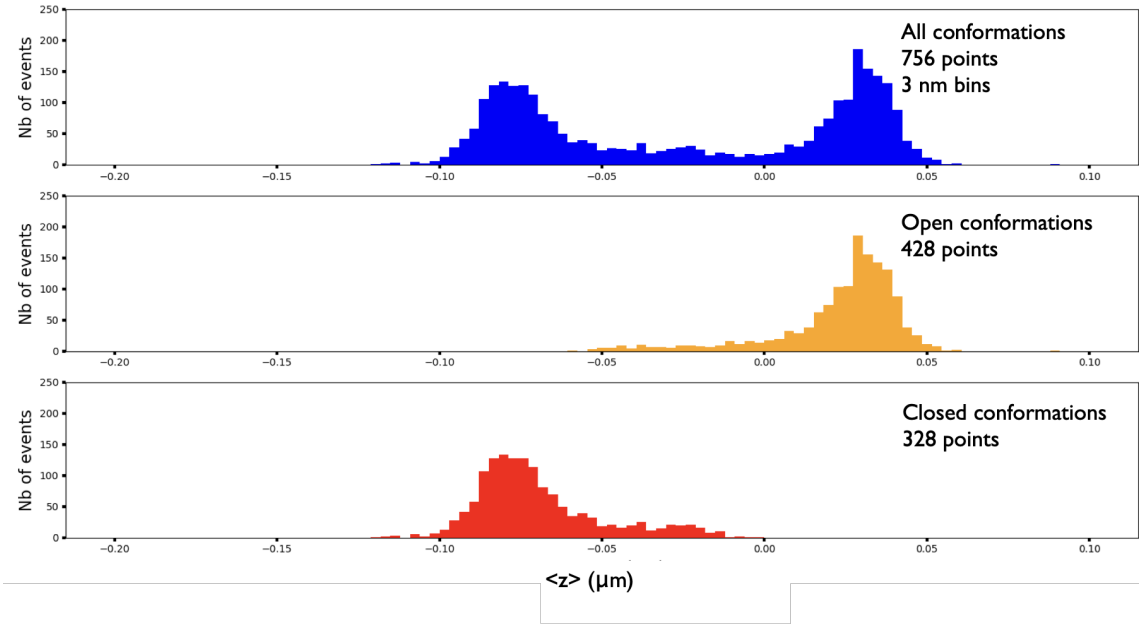


Figure 7.4: Distribution of event $< z >$ for a typical 8 hour measurement on a [FKBP12:JK313] jDNA scaffold with no competitor in solution. Data acquired in mTOR buffer at 28°C under applied force 100 fN.

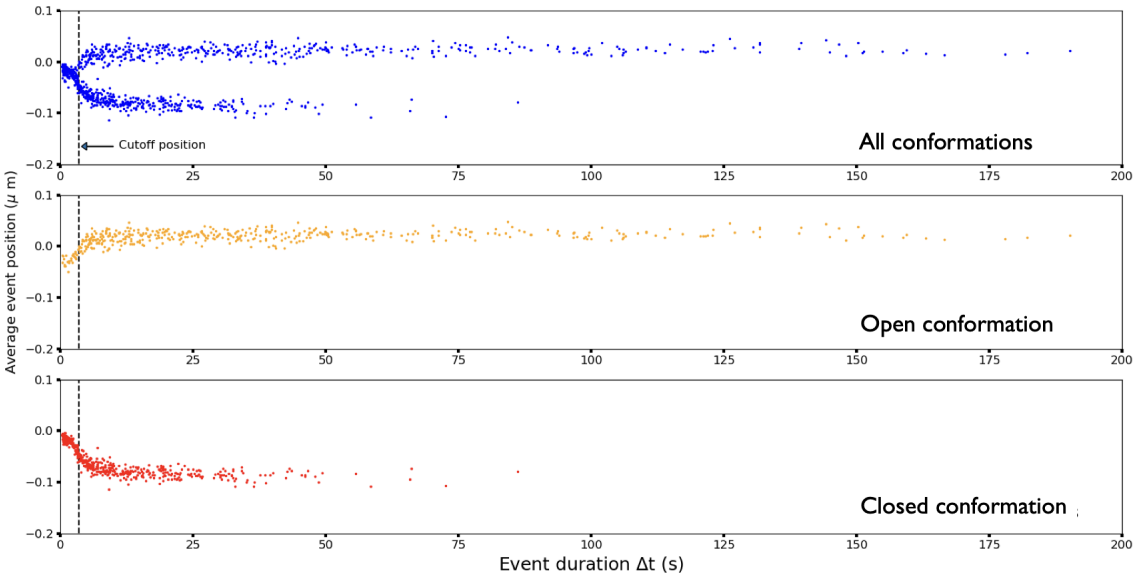


Figure 7.5: Correlation between event $< z >$ and event duration for a typical 8-hour experiment on a [FKBP12:JK313] jDNA scaffold with no competitor

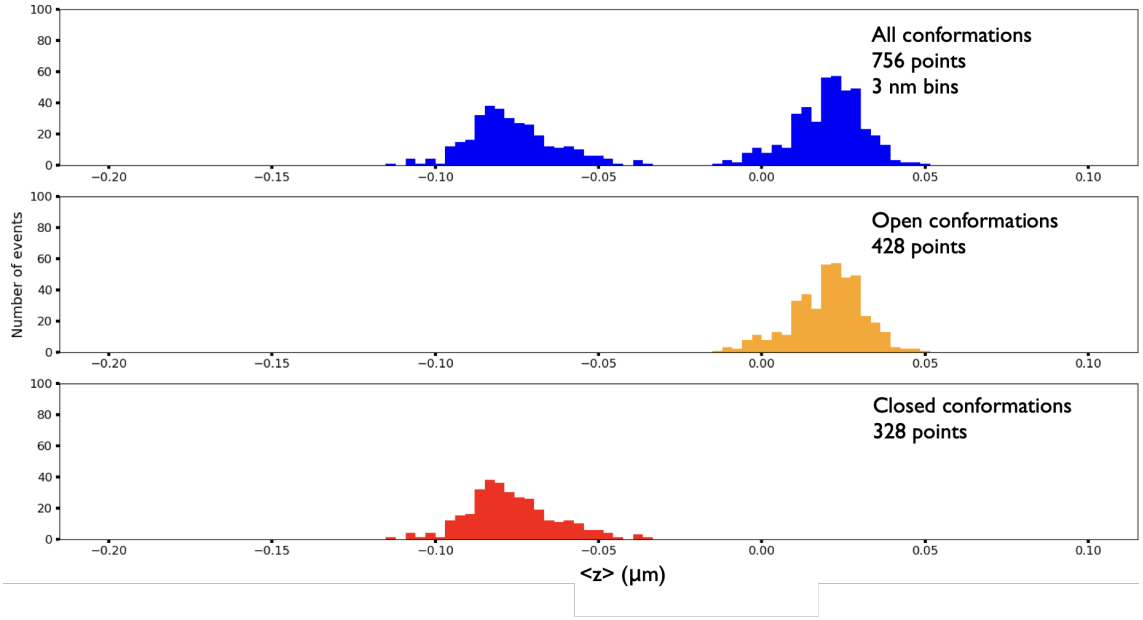


Figure 7.6: Distribution of event $< z >$ (events longer than 3.5s) for a typical 8-hour experiment on a [FKBP12:JK313] jDNA scaffold with no competitor weighted by event length. Data acquired in mTOR buffer at 28°C under applied force 100 fN.

decaying exponential, and the characteristic time is the inverse of the rate constant. Therefore, it makes sense to plot the histograms of Δt, both of open and closed events, and attempt to interpolate them with an exponential. For this interpolation, we only take events that are longer than the 3.5s threshold established earlier. We use the `curve_fit` function of the python `scipy.optimize` library (*140*), which uses a non-linear least-squares optimization method.
We obtain, on most experiments, histograms that are well-interpolated by a single decaying exponential (Fig. 7.7) [3]. Events shorter than 3.5s would break this exponential distribution . This shows that the algorithm does not detect all the short events, which would skew the results if they were included. This comforts us in our decision to exclude those events, which is done in every characteristic time determination made in this work. As may be seen from this figure, events shorter than 3.5s typically amount to 5 % (open) or 10 % (closed) of all detected events.

Population ratio

The population ratio $\Delta t_O/\Delta t_C$ is the ratio of the total time a scaffolds spends in the open conformation, divided by the time spent closed. We will show in Section 8 that theoretically this ratio verifies $\Delta t_O/\Delta t_C = I + S \times P_{trans}$ where I an P can be linked to the equilibrium constants of the binding reaction while P_{trans} is the

[3]Equation $f = Ae^{-t/\tau}$, with A a normalization prefactor (so that the integral of the curve is equal to 1) and τ the characteristic time, inverse of the rate constant

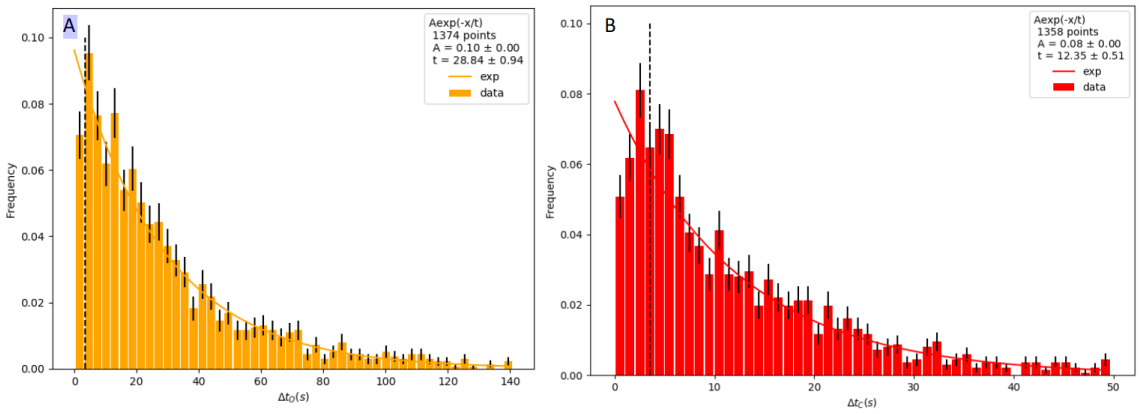


Figure 7.7: Duration of open (A) and closed (B) events on an 8 hour observation of [FKBP12:JK313] jDNA with no competitor. Bins are 3 s (A) or 1 s (B) intervals. The continuous line corresponds to the single-exponential interpolation. The dotted line corresponds to the 3.5-second threshold chosen for exclusion. The black error bars are s.e.m.

concentration of the competitor. It is therefore essential for the interpretation of our results.

7.6.5 Simulation of Markov dynamics of protein-ligand interactions

We are going to observe random events on single molecules. It is thus interesting to evaluate how much differences we can expect from one experiment to the next because of chance. In other words, if we observed identical scaffolds, how much variation would we observe ?

To answer this question we ran simulations of Markov dynamics, using estimated association and dissociation parameters from preliminary JK313-FKBP12 competition assays, and ran the analysis on every single trace.

Simulated trajectory generation

We simulated 200 8-hour measurements for concentrations of FKBP12 between 0 and 60 nM, assuming an on-scaffold concentration of $P_{cis} = 14$ nM, an association rate constant $k_+ = 0.0024\text{s}^{-1}.\text{nM}^{-1}$ and dissociation rate constant $k_- = 0.11\text{s}^{-1}$. Those simulations took into account the existence of 3 different states : C_{cis} (The scaffold is closed, with interactions between the two molecules on the scaffold), O_{free} (no interaction) and O_{trans} (interaction between one molecule on the scaffold with a competitor in solution).

Figure 7.8 shows example traces obtained from simulations. As expected, the frequency of trans-interactions increases with the concentration of competitor ; also note that transitions between open and closed only happen between closed and open

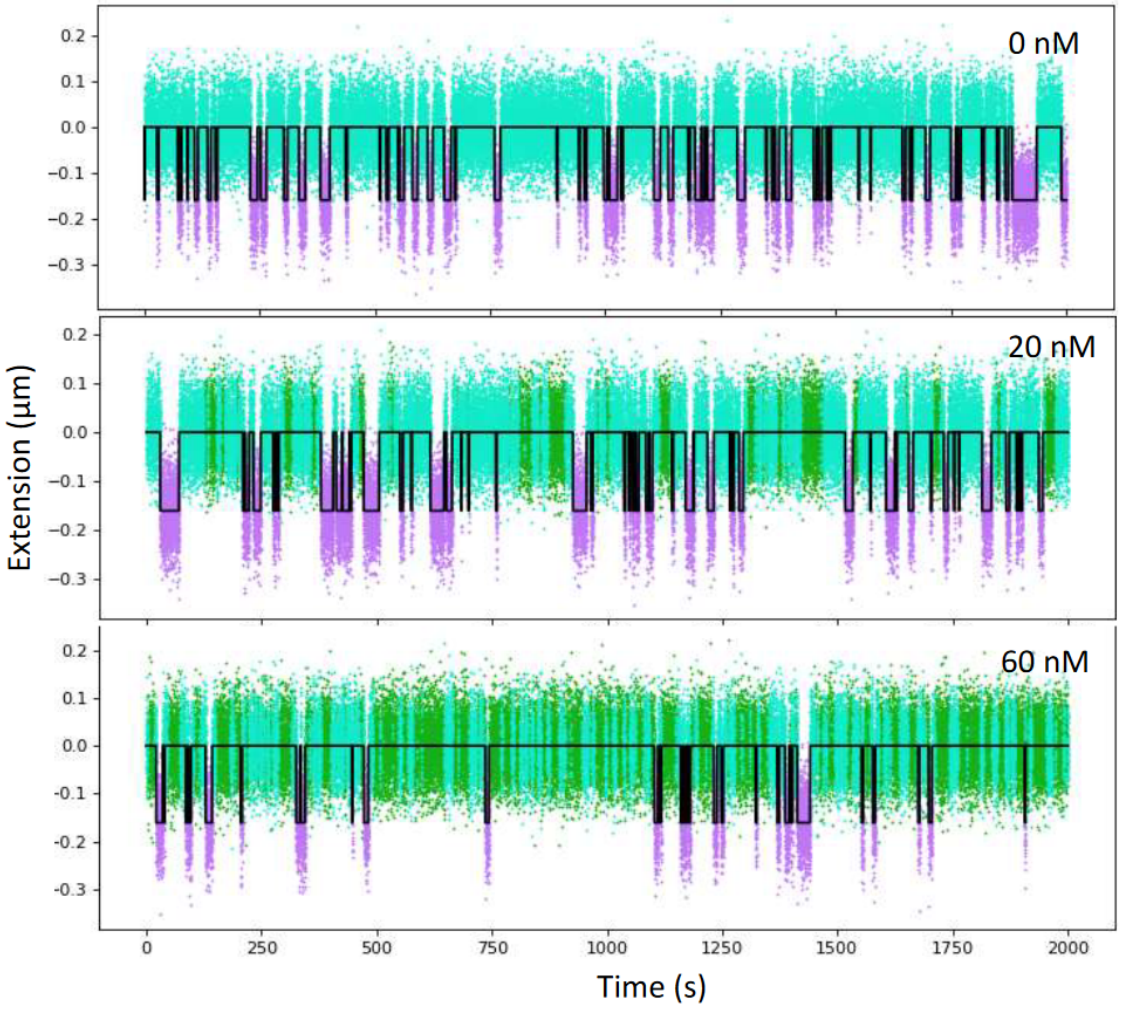


Figure 7.8: Example 2000 s extracts from simulated traces using Markov dynamics. Purple points indicate a closed scaffold (C_{cis} state) ; light green points are open scaffold with no interaction (O_{free}) ; dark green is scaffold with an interaction with an in-solution compound (O_{trans}). Black line is the conformation assignment using the analysis algorithm.

with no interaction. The trajectories were generated in python using the HMMLEARN (*139*) package by creating a 3-state Markov model with transition probabilities given by the above constants. The observable signal output from all states follows a Gaussian distribution ; both open states have an identical distribution of $< z >= 0$ and $\sigma = 0.05\mu m$. The closed state has $< z >= -0.16\mu m$ and $\sigma = 0.05\mu m$. The time step used in the simulation was 1/31.08 s so that the simulations have the exactly same temporal resolution as the experiments.

Detection protocol validation

This simulation is an opportunity to validate our detection protocol, tough the simulation only simulates a Gaussian distribution. In that configuration, the detection algorithm has over 99.9 % correct event recognition, and perfect recognition for

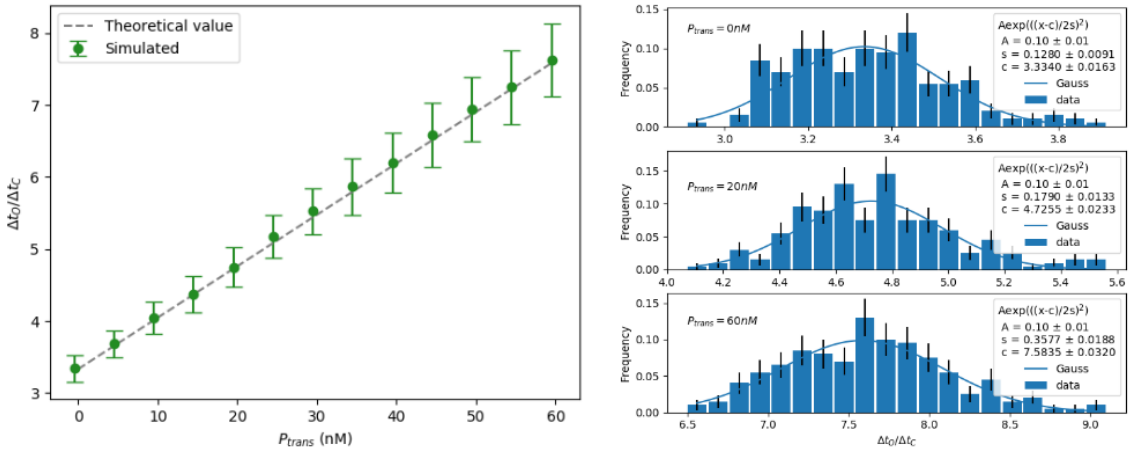


Figure 7.9: Simulated total open to closed conformations dwell-times. Averages and standard deviations were compiled on 200 8-hour simulated experiments.

events longer than 1 s.

Error estimation for population ratio

The primary question we wanted to answer with the simulations is how much dispersion we can expect from the experimental results. Indeed, since the experiment is limited in time, some variability can be expected from purely statistical sources. We computed, for the 200 simulated trajectories, the value of the $\Delta t_O/\Delta t_C$ ratio for each trace, and then computed the average and standard deviation. The results are indicated on Figure 7.9. In addition, using linear regression we found that the s.e.m. on the ratio is equal to 6.3 ± 0.14 % of the value of the ratio itself.
Rounding up, we will use the estimation that the ratios have an expected s.e.m. of 7 % of their value. We assume that this approximation will hold for all the measurements three reasons. Firstly, the total acquisition time is always close to 8 h . Secondly, we always use the same scaffold and finally, we always adjusted the competitor concentration so that the $\Delta t_O/\Delta t_C$ ratio is kept between 3 and 10, values explored by those simulations.

Using the linear fit of the $\Delta t_O/\Delta t_C$ ratio presented in 8.1.3 - Equation 8.14 we also determined, for every simulation, the values of P_{cis} and K. The resulting values are distributed according to Gaussian distributions [4].
For a simulated value of $K_0 = 47.6nM$, we obtain a value for $K_0 = 45.7$ nM and a standard deviation of 2.7 nM - a 6 % error on a single experiment. For P_{cis} simulated at $14nM$ we obtain a value of $13.98nM$ with a standard deviation of $0.73nM$ - a 5 % error on average on an individual experiment. The results are plotted on Fig. 7.10.

[4]Note that our simulations were made using preliminary data for the values of the affinity constants, so the values differ from the final results on the proteins

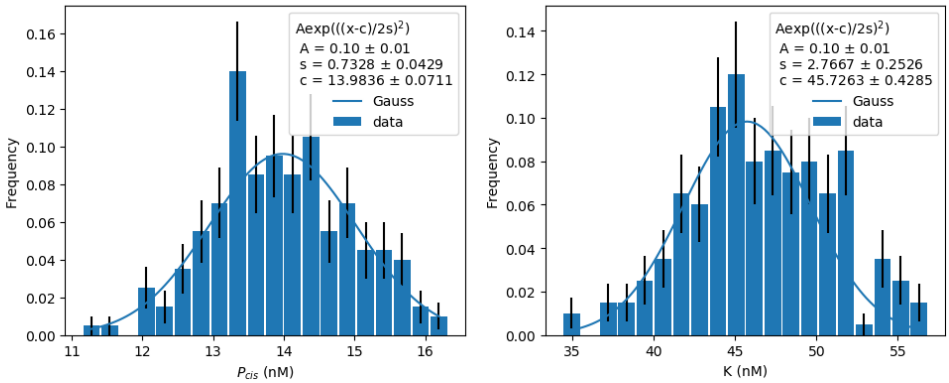


Figure 7.10: Simulated K and P_{cis} values obtained from 200 simulations. Line corresponds to Gaussian interpolation ; black error bars are s.e.m. on the bar.

The takeaway of those simulations is that even tough the experiments are limited in time, if we assume that all the scaffolds behave identically, we would expect an inter-scaffold variability on the order of 5 % in the extracted K and P_{cis}.

7.6.6 Analysis of force-cycling experiments

The same algorithm can be used for the force cycle experiments, the main difference being that the data traces are first divided into segments based on the position of the magnets. Three states are used : low force, closed at high force and open at high force. Each state corresponds to a different conformation. The transitions between the states can then be determined. No filtering (step 3) is needed, as the higher force reduces noise.

Part III

Results

Chapter 8

Titration experiments on jDNA

In this section, we will cover the passive measurement protocol shown in the materials section.

8.1 Theoretical framework

Recall from Section 2.1.3 that the unbinding reaction of a protein-ligand complex is :

$$\mathrm{PL} \underset{\mathrm{k_+}}{\overset{\mathrm{k_-}}{\rightleftharpoons}} \mathrm{P} + \mathrm{L} \tag{8.1}$$

As in 2.1.3, we will note P, L and PL the concentration of the species ; the rate constants are k_- $[s^{-1}]$ and k_+ $[M^{-1}.s^{-1}]$. The equilibrium constant $K = \frac{k_-}{k_+}$ is homogeneous to a concentration. While in our work P mainly refers to the protein and L to the (non-proteic) ligand, the theoretical framework holds for any two molecules on the scaffold.

These constants are what would be measured in a bulk assay, that is independent of Force. On the scaffold, the notion of concentration of molecules is ill-defined. Tough one could try to estimate that concentration by determining the available space for the free diffusion of the molecules on the scaffold, it would not be possible to account for the restrictions in diffusion and rotation imposed on the molecules by the steric constraints imposed by the scaffold, for instance. It is thus much more convenient to define the apparent first-rate reaction constants k_{+cis} and k_{-cis} , which take all the various, poorly accessible phenomena into account. Those constants can be readily measured by observing the opening and closing of a scaffold.

8.1.1 Kinetic model

With no partner in solution

The objective then is to relate those apparent constants to the in-solution constants. Under the hypothesis that the force is negligible on the scaffold, we have $k_{-cis} = k_-$ as the dissociation reaction depends only on PL. For the association reaction, we define an apparent concentration of P, P_{cis} so that $k_{+cis} = P_{cis} \times k_+$. In summary :

$$k_{-cis} = k_- \quad k_{+cis} = k_+ P_{cis}$$

In other words, the dynamic between the open O_{free} and closed C_{trans} conformations of the isolated scaffold is represented by the reaction :

$$\mathrm{C_{cis}} \underset{\mathrm{k_{+cis}}}{\overset{\mathrm{k_-}}{\rightleftharpoons}} \mathrm{O_{free}} \tag{8.2}$$

We have thus two main parameters : K and P_{cis}. While the equilibrium constant K is self-explanatory, P_{cis} is less clear: is the *apparent* concentration of P experienced by the protein L engrafted on the scaffold. It is, indeed, the concentration of one molecule of P inside the volume delimited by the scaffold, as seen by partner L. This parameter integrates the behavior of the scaffold (shape, size, tension, etc.), effect of force on the reaction, and all other unaccounted effects.

With a partner in solution

If soluble molecule P of concentration P_{trans} is added in the medium with the scaffold, an extra-scaffold binding reaction is possible :

$$\mathrm{P_{trans}} + \mathrm{L} \underset{\mathrm{k_{-trans}}}{\overset{\mathrm{k_{+trans}}}{\rightleftharpoons}} \mathrm{P_{trans}\,L} \tag{8.3}$$

Under the hypothesis that the complexes behave identically on the scaffold and in solution, we can propose $k_{-trans} = k_-$. This hypothesis will have to be reexamined later, in the discussion. Also note that competition with soluble L works identically.

States and conformations

The scaffold can be in two *observable* conformations, open (high extension) and closed (short extension). If the scaffold is isolated in the solution, they correspond to the two states of the molecular partners : open with no interaction O_{free} and closed in intra-scaffold (cis) binding, C_{cis}. However, if we add a molecular partner

in the solution, a third state, O_{trans}, becomes possible. This state however corresponds to the same observable conformation, open, and so we have two observable conformations corresponding to three hidden states. In summary, we have two reactions :

$$\mathrm{C_{cis}} \underset{\mathrm{k_{+cis}}}{\overset{\mathrm{k_-}}{\rightleftharpoons}} \mathrm{O_{free}} \quad \mathrm{O_{trans}} \underset{\mathrm{k_{+trans}}}{\overset{\mathrm{k_-}}{\rightleftharpoons}} \mathrm{O_{free}} \tag{8.4}$$

Constant concentration of P_{trans}

In all the rest of this work, we will assume that the concentration of P_{trans} is constant in time and across the scaffolds under observation. We work at competitor concentrations of 1 nM and higher ; at those concentrations, the typical distance between molecules in the solution is 1 μm (*84*). At the same time, the distance between the scaffolds is at the barest minimum of 10 μm (and there are typically only 50 scaffolds on a 150 by 150 μm surface). The scaffolds are only on two dimensions, so that the bulk of the solution acts as a reserve. We do not, therefore, expect any localized depletion of P_{trans} by the scaffolds.

Accounting for force

We have talked about the effects of forces on chemical reactions in the first chapter. Our goal is to measure *in-solution* values for our constants, and thus to remove the effects of force from the experiments. Previous observations in our team by Follenfant et al. (Ref (*141*)) have shown on force-cycling experiments that the dissociation rate constant k_- verifies the following relationship with force F :

$$k_-(F) = k_-^0 e^{\frac{z_D}{k_B T}F}$$

With k_- the value of the dissociation rate constant at zero force, z_D the distance to the transition state towards dissociation (measured at 0.38 nm). We thus have $\left(\frac{\partial k_-}{\partial F}\right)_{F\approx 0} = \frac{z_D}{k_B T}$, and for $F \approx 100fN$, we get a 1 % variation for the value of the dissociation constant. Thus, at the forces used for passive measurements, the dissociation does not appear to be affected significantly by force. For the association rate, the effects of the force can be included in the value of the intra-scaffold concentration P_{cis} .

8.1.2 Characteristic time analysis

Our primary experimental observable we have is determining whether the complex is, at a given time, open or closed.

Off-rate

The off-rate, k_- is the inverse of the average time required to dissociate the complex. On an exponential, this average is equal to the characteristic time. To happen, the dissociation reaction takes the time spent in a closed complex - thus :

$$\tau_C = \frac{1}{k_-}$$

k_- is independent from P_{trans}, and can be measured in every condition. The average value for τ_C in an experiment will be written as M, a function of (P_{trans}).

On-rate

With no competitor In conditions without competition we have :

$$\tau_O = \frac{1}{k_+ \times P_{cis}}$$

In that simplified case, the event durations are distributed exponentially. Thus, to evaluate τ_O and τ_C we can simply plot the distribution of event durations and interpolate them with a decreasing exponential function.

With a competitor in solution The association reaction is bimolecular and thus cannot be evaluated without the knowledge of P_{cis} . We can only directly observe the dwell times of open events, as we are unable to distinguish between the O_{free} and O_{trans} states. An open conformation event can correspond to a single instance of O_{free}, or it can correspond to a sequence of O_{free} and O_{trans} events, and we cannot distinguish them.
Thus, in theory, the distribution of the open events durations should follow a more complex pattern. According to Shelly et al. (Ref. (*142*)), the open states are distributed according to a bi-exponential distribution of characteristic times τ_{O-} and τ_{O+}. If we consider the association constant outside of the scaffold to be k_{+trans} and its dissociation k_{-trans} then the two time constants can be written as follows :

$$\tau_{O\pm} = \frac{2}{\alpha \pm \sqrt{\alpha^2 - 4k_{+cis}k_-}} \tag{8.5}$$

with

$$\alpha = k_- + k_{+trans} + k_{+cis}$$

The weights of both lifetimes are expressed as :

$$w_{O\pm} = \pm k_{+cis} \frac{(1/\tau_{O\pm}) - k_{-}}{(1/\tau_{O+}) - (1/\tau_{O-})} \tag{8.6}$$

This set of equations is complex, and interpolation based on experimental data appears difficult. Indeed, this would require us to be able to interpolate a double-exponential decay in lifetimes from our experimental findings, that we have been unable to achieve reliably, because we cannot effectively acquire events shorter than 3.5s. This could be an issue as we cannot directly extract the lifetime of a protein-competitor interaction event. Thus we need to aim to not interpret directly the observed lifetime of the open states in competition assays.
Those distributions can still be approximated by an exponential, as will be seen later.

8.1.3 Population analysis

As noted above, we cannot distinguish the O_{free} state from the O_{trans} state, as they both correspond to the open observable conformation. As understanding the dwell-times of conformations is complex, we will instead use a measurement strategy based on the ratio of occupation of both states.

We consider the fraction of time spent by the scaffold in both conformations. It stands that those fractions verify $C + O = 1$ and the states verify $C_{cis} + O_{free} + O_{trans} = 1$. If we consider an equilibrium situation, we can write the equilibrium constants for both reactions of binding and unbinding, inside the scaffold ("cis") and outside of the scaffold ("trans").

Heterogeneous competition : competition by one of the engrafted partners

We add in the solution a competitor of concentration P_{trans}. While P_{cis} is not directly accessible to measurement, the amount of competitor protein put in solution is known. Thus, we will be using competition assays to determine indirectly P_{cis} relative to P_{trans} . We define two equilibrium constants :

$$K_{cis} = \frac{P_{cis} O_{free}}{C_{cis}} \tag{8.7}$$

$$K_{trans} = \frac{P_{trans} O_{free}}{O_{trans}} \tag{8.8}$$

From state occupation to rate constants

We calculate the ratio of time spent in the open state relative to the time spent in closed state :

$$\frac{O_{free} + O_{trans}}{C_{cis}} = \frac{O_{free}}{C_{cis}} + \frac{O_{trans} \times O_{free}}{C_{cis} \times O_{free}} = \frac{O_{free}}{C_{cis}} \left(1 + \frac{O_{trans}}{O_{free}}\right)$$

From there, one can arrive to the following theoretical equation using the mass action laws 8.7 and 8.8 :

$$\frac{O_{free} + O_{trans}}{c_{cis}} = \frac{K_{cis}}{P_{cis}} \left(1 + \frac{P_{trans}}{K_{trans}}\right) \tag{8.9}$$

Now, we are left with finding a means to evaluate the leftmost ratio, which links the occupation probabilities of each state. We can either do so by computing an *average* distribution of state occupations (eg. what is the average percentage of time the scaffold spends in each state ?) or we could try to pinpoint more precisely the duration of every state, then calculate the average duration of each state and obtain their ratios. The first option, much more practical, leads us to the following equation, that links the average time spent in both states to the thermodynamic parameters :

$$\frac{\Sigma \Delta t_O}{\Sigma \Delta t_C} = \frac{K_{cis}}{P_{cis}} \left(1 + \frac{P_{trans}}{K_{trans}}\right) \tag{8.10}$$

The dependence of the $\frac{\Sigma \Delta t_O}{\Sigma \Delta t_C}$ ratio on the competitor concentration is linear, which is the reason this particular formula will be used. To simplify further analysis, we will call the intercept of the interpolation line I and the slope S : this finally leads us to the interpolation equation 8.11 :

$$\frac{\Sigma \Delta t_O}{\Sigma \Delta t_C} = I + S \times P_{trans} \tag{8.11}$$

With $I = \frac{K_{cis}}{P_{cis}}$ and $S = \frac{K_{cis}}{K_{trans} P_{cis}}$, so that we may very simply obtain $K_{trans} = \frac{I}{S}$. Note that we cannot obtain K_{cis} and P_{cis} from heterogeneous competition alone.

Homogeneous competition : competition by one of the engrafted partners

In a more specific case, if we add in the solution one of the molecules attached to the scaffold, the equilibrium constants of the *cis* and *trans* reactions are the same K and verify :

$$K = \frac{P_{cis} O_{free}}{C_{cis}} \tag{8.12}$$

$$K = \frac{P_{trans} O_{free}}{O_{trans}} \tag{8.13}$$

Using the same reasoning as previously, we reach this equation :

$$\frac{\Sigma \Delta t_O}{\Sigma \Delta t_C} = \frac{K}{P_{cis}} \left(1 + \frac{P_{trans}}{K} \right) \tag{8.14}$$

We can likewise define $I = \frac{K}{P_{cis}}$ and $S = \frac{1}{P_{cis}}$, so that we may very simply obtain $K = \frac{I}{S}$ using equation 8.11. We can also deduce $P_{cis} = 1/I$.

8.2 Results : Competition by one of the engrafted partners

The first experiment we will present is a homogeneous competition experiment in which we use the [FKBP12:JK313] jDNA scaffold shown in Section 5.4.1 in competition against soluble FKBP12. Our aim is twofold : first, to verify that the theory we established in the previous Section holds on a relatively simple experimental setup by calculating the equilibrium constant K. Secondly, we want to characterize the behavior of jDNA by calculating P_{cis} and observing variability between scaffolds.
The experiments follows the spontaneous fluctuation protocols and data analysis outlined in Materials and Methods (Sections 7.3 and 7.6).

8.2.1 Event durations and rate constants

For every scaffold at every concentration of competitor, the duration of open and closed events are collected and histograms are plotted. We choose to not use events shorter than 3.5 s for our analysis, since we have observed that from that threshold we are no longer capable of detecting reliably all events, and separating short events from artifacts (Section 7.6.4). A single-exponential best-fit is calculated for every case. A typical figure (for one scaffold across different concentrations of FKBP12) is shown on Figures 8.1 and 8.2 for open and closed conformations respectively. Figure 8.4 shows that the dwell-time of the scaffold in closed conformation is constant with P_{trans} . Figure 8.3, on the other hand, shows the increase of the dwell-time in the open conformation (obtained via *single-exponential* interpolation, see below) with P_{trans} , an increase that appears linear.
Notice that experimental noise increases with P_{trans} , as the events are longer, so less events are recorded per experiments.

τ_C analysis and k_- determination

The closed complex lifetime is independent of P_{trans} in accordance with the theory. We define M to be the average value for each scaffold. Across 27 scaffolds, we obtain an average value of $M = 10.3 \pm 2.8$ s, which leads us to the value of $k_- = 0.097 \pm 0.026\text{s}^{-1}$.

τ_O analysis and k_+ determination

With no competitor, $\tau_O = 27.4 \pm 8.6$s across all 27 scaffolds, which leads us to the value of $k_+ = 3.710^{-4} \pm 9,310^{-5}\text{s}{-}1\ \text{nM}^{-1}$ (After determining P_{cis} in Section 8.2.3). However, the analysis of the open dwell-times with a competitor is more complex

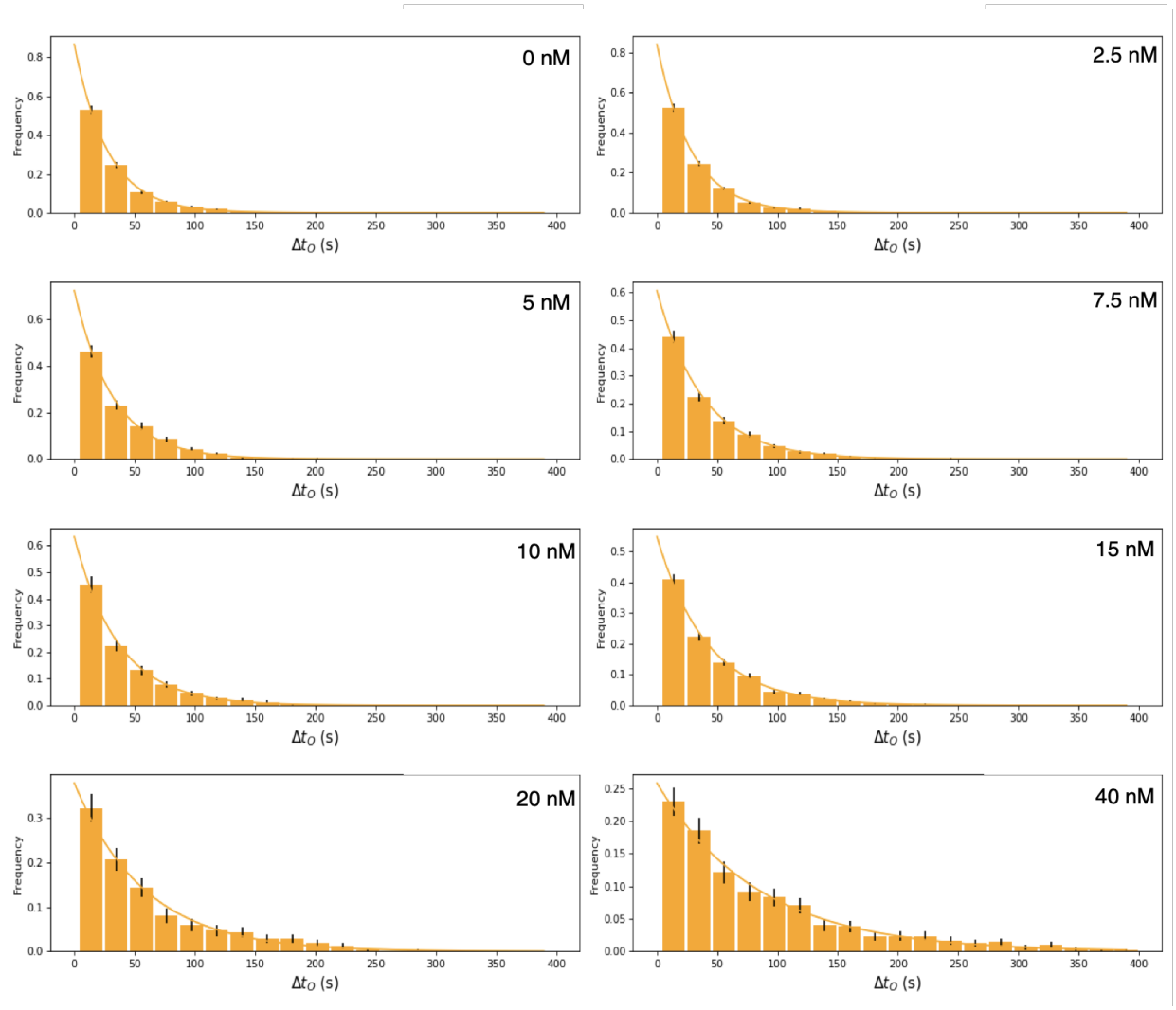


Figure 8.1: Typical histograms of open events durations acquired on [FKBP12:JK313] jDNA scaffolds with increasing concentrations of FKBP12 at 28°C in mTOR at 100 fN; vertical error bars are s.e.m. Events shorter than 3.5s were not included in the histograms. Bin size is 25s.

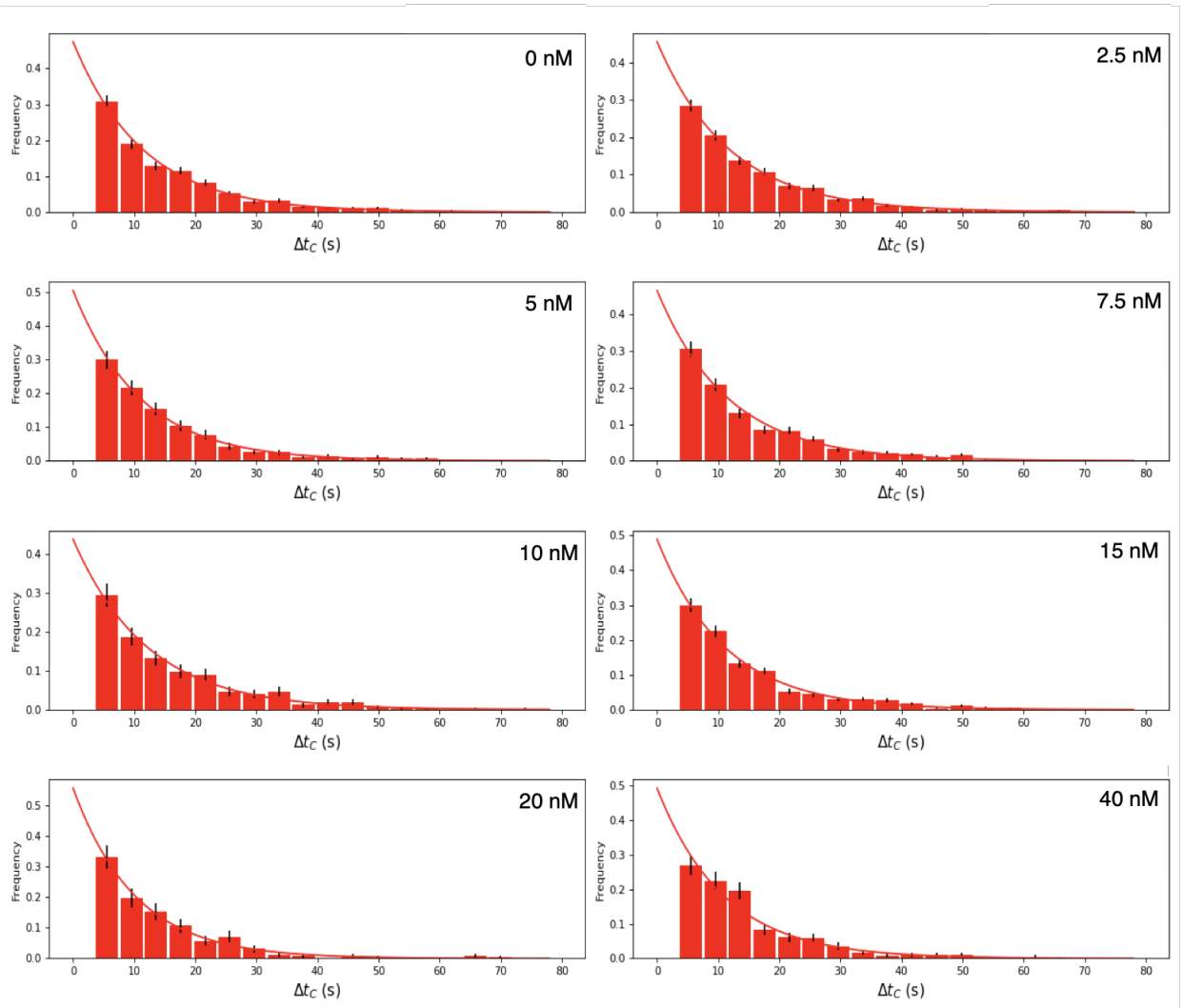


Figure 8.2: Typical histograms of closed events durations acquired on [FKBP12:JK313] jDNA scaffolds with increasing concentrations of FKBP12 at 28°C in mTOR at 100 fN; vertical error bars are s.e.m. Events shorter than 3.5s were not included in the histograms. . Bin size is 6s.

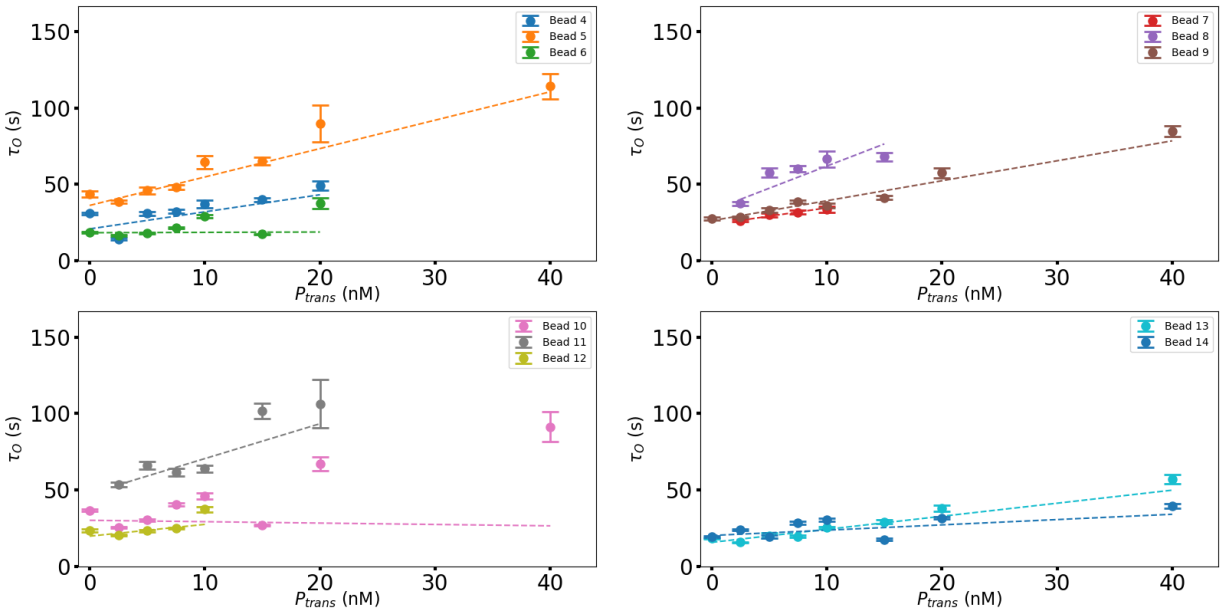


Figure 8.3: Characteristic open time of events from experimental series 1, observed at 28°C in mTOR at 100 fN on [FKBP12:JK313] jDNA scaffolds with increasing concentrations of FKBP12 (Ptrans). Error bars are s.e.m. deduced from exponential fit.

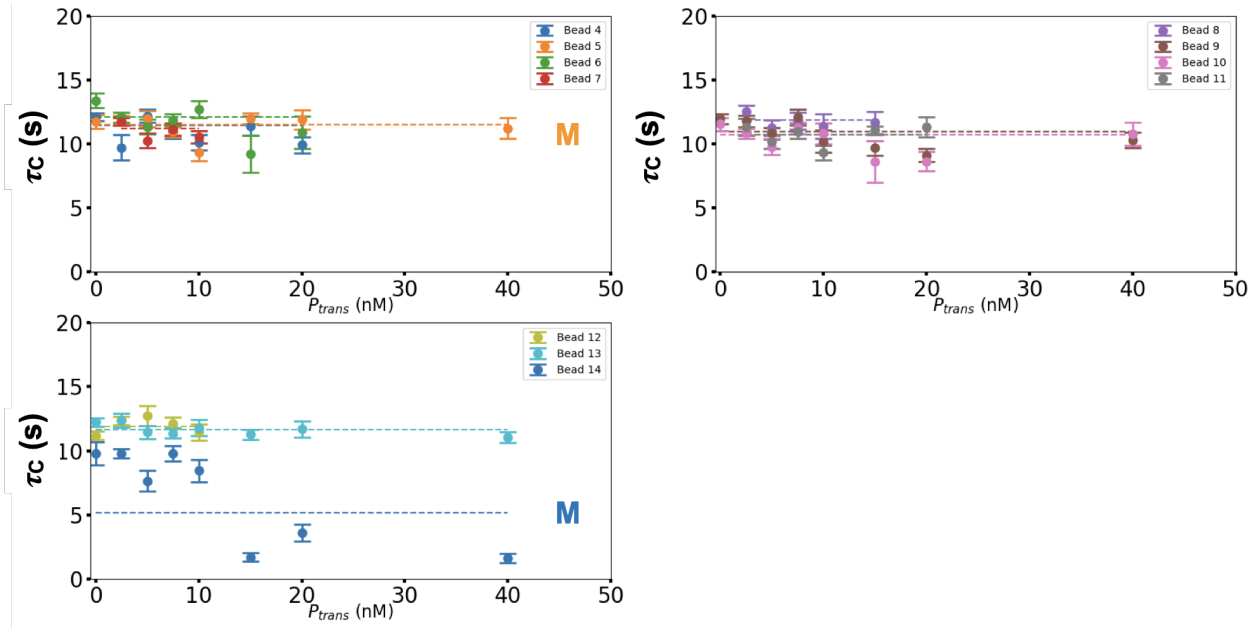


Figure 8.4: Characteristic closed time of events from experimental series 1, observed at 28°C in mTOR at 100 fN on [FKBP12:JK313] jDNA scaffolds with increasing concentrations of FKBP12 (Ptrans). Dotted lines are constant value best-fit for M. Error bars are s.e.m. deduced from exponential fit.

as mentioned in Section 8.1.2, the underlying issue being the double-exponential distribution of the open dwell-times.

Comparison with theoretical distributions

Figure 8.5 shows the expected values of $\tau_{O,+}$ and $\tau_{O,-}$ using the aforementioned values of k_+ and k_- under Equation 8.5 (A2). As can be seen, we expect them to be on the same order of magnitude in most of the competitor range (At $P_{trans} = 30$ nM, $\tau_{O,-} = 11\ \tau_{O,+}$). on the other hand, their weights are very different, with the exponential of distribution $\tau_{O,+}$ being in the minority. (B1-3) on the same figure shows expected open event duration distributions , and we can see that, especially for long events at high competitor concentrations, the short lifetime-exponential is eclipsed by the long lifetimes. This explains why a single exponential is a correct approximation on Fig. 8.1. The number of events that can be observed and attributed to every state is therefore lower (depending on experiment and scaffold, between 200 and 500). Since the two exponential constants are close, we are unable to distinguish them. We are, however, able to interpolate the open dwell-times with a single-exponential distribution, and it is this approximate solution that we have used here.

A very recent work by Rieu et al. (*143*) used similar methods and analyses on DNA hairpins. Contrary to the present work, they were able to distinguish both exponential constants. This is made possible by the fact that they compare two different physical phenomena - DNA hairpin (un-)folding and protein-protein associations, that have very different time constants, with values at circa. 10 ms and 1s respectively (while we expect them to be at 10s and 20-80s). The quick interaction also allowed for much faster data acquisition, with over 7000 points per experiment, whereas here we collect 200-500 points per condition and bead. Those are advantages that we do not have here, and so we must use population dynamics to probe reliably the dynamics of our scaffold.

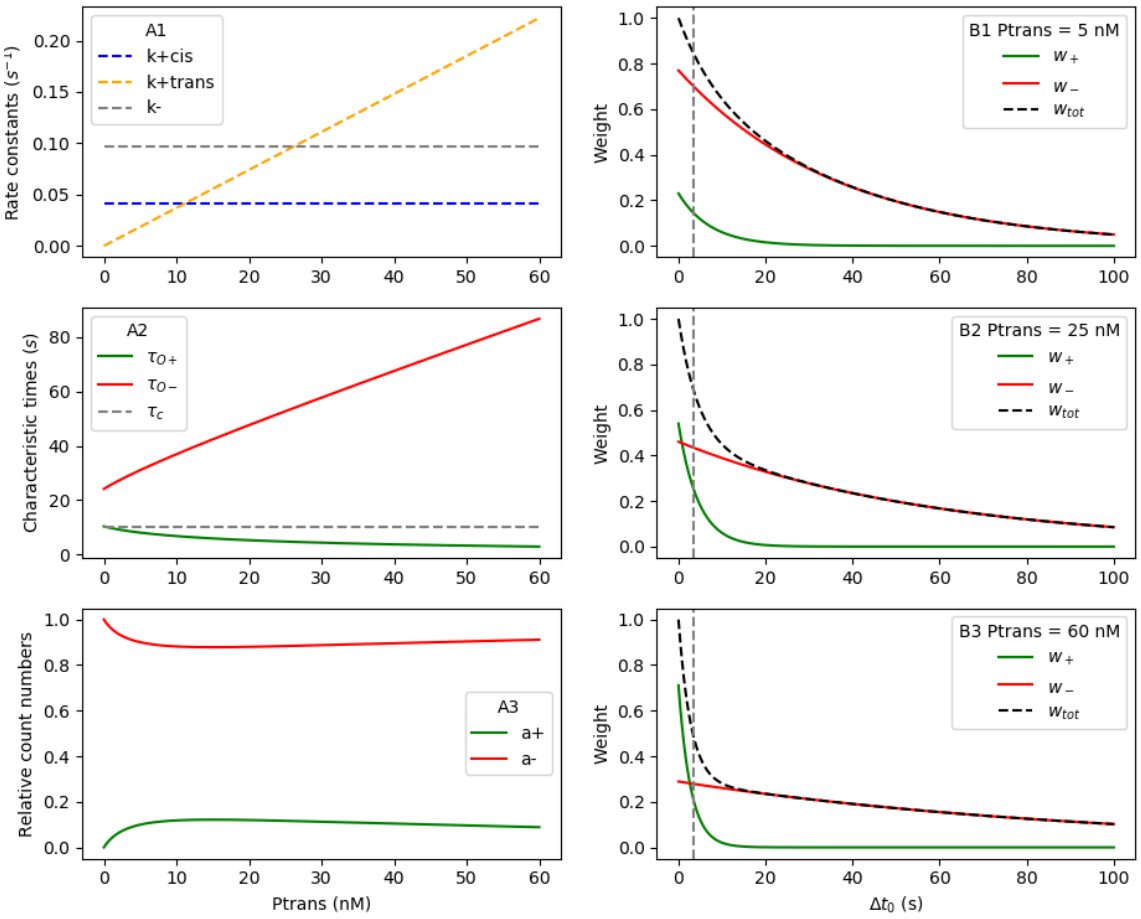


Figure 8.5: (A1) Theoretical values for the rate constants as calculated from the simulated $k_+ = 0.0024 s^{-1} nM^{-1}$ and $k_- = 0.11 s^{-1}$ (A2) Theoretical values for $\tau_{O,+}$ and $\tau_{O,-}$ calculated from experimental results using Equation 8.5. (A3) Theoretical relative importance of the two exponential distributions studied in Equation 8.6 in function of P_{trans} (B1-3) theoretical distributions of the open state events for P_{trans} of 5, 25 and 60 nM respectively, contributed by the two exponential distributions. The black dotted curve corresponds to the resulting combined distribution ; the gray dotted vertical line is the 3.5s threshold under which events are not taken into account for the analysis.

8.2.2 Population analysis

For every scaffold, a ratio of $\frac{\Sigma\Delta t_O}{\Sigma\Delta t_C}$ (the ratio of time spent in the open state versus the time spent in the closed state) is determined at every concentration of FKBP12. According to Equation 8.11 , there is a linear correlation between this ratio and competitor concentration. We interpolate a line through our points (using least-square optimization) : the slope S of this line is equal to the inverse of P_{cis} while the y-axis intercept I is equal to K_{cis} . The equilibrium constant can be determined by the relationship $K = K_{cis}P_{cis}$, for each individual scaffold, or by fitting the equation $K = I/S$.

Evaluating the error on the measurements

The analytical approach we have taken here only produces one data point per concentration of competitor and scaffold. Thus, there is no straightforward way of determining a standard error on the mean for each measurement point.
As we only acquire for a limited amount of time, under-sampling could be an issue. To estimate the error caused by sampling, we have used Markov dynamics simulations (Section 7.6.5).
The takeaway of those simulations is that even tough the experiments are limited in time, if we assume that all the scaffolds behave identically, we would expect an inter-scaffold variability on the order of 5 % in the parameters. We used those simulations to determine the s.e.m. on the values of the $\frac{\Sigma\Delta to_{free/trans}}{\Sigma\Delta tc_{cis}}$ ratio for each point of competitor concentration and were used to draw the error bars shown on the figures.

Individual scaffold analysis

Over 3 replicates of the experiment, we obtain multiple measurements of open-to-closed ratios for different scaffolds and concentrations of soluble FKBP12. We can observe the linear correlation between FKBP12 concentration and ratio, verifying equation 8.11.
The results show standard deviations of 7.43 nM for K and 6.65 for P_{cis} , which translates to percentage dispersion of 38 % and 60 % respectively. When compared with the expected statistical errors mentioned above, we see that statistical errors do not (solely) account for the differences in parameter values observed. We must postulate that the behavior of different scaffolds is physically different. This will be covered in more detail in Section 8.2.3.

Global analysis

We summarize all the scaffolds onto a single scatter plot showing I in function of S. A line is fitted onto this cloud of points. The slope of the best-fit line is the value of the K equilibrium constant. This method has the advantage of nullifying internally the effect of P_{cis} on K, under the assumption that P_{cis} is affected by the force to which a given scaffold is exposed. Resulting figures can be found on Fig. 8.7.
Regrouping the results from the tree replicates, we find $K = 18.2 \pm 1.3nM$ from the fit, while the average value of K from the individual linear fits is 19.6 nM with a standard deviation of 7.44 nM.

Summary

Figure 8.6 shows an example of the plots that we obtain (for the first experimental repeat). Table 8.1 summarizes the three replicates.

Table 8.1: Results of 3 replicates of affinity measurements at 28°C in mTOR at 100 fN on [FKBP12:JK313] jDNA

Replicate	K value from scatter plot (Fig. 8.7)	Average individual K ($< I/S >$)	Average individual P_{cis} value (nM) ($1/ < S >$)	Nb. scaffolds
1	14.6 ± 0.6	14.8 ± 2.0	7.4 ± 3.0	11
2	19.9 ± 1.7	21.6 ± 4.8	9.7 ± 2.6	9
3	21.4 ± 4.0	24.7 ± 8.6	19.0 ± 8.6	7

The scatter plot interpolations yield a lower s.e.m. thanks to the linear fitting algorithm, and will be used for the rest of the work. The error obtained, on the order of 20 % on individual replicates and 5 % on the general interpolation are on-par with classical methods, including ensemble methods.

8.2.3 Measuring the equivalent cis-concentration

The equivalent cis-concentration can be deduced from the variation of the open-to-closed ratio as described in equation 8.11. From the experiments, we deduce its value at 11.2 ± 6.7 nM. Notice that a strong variability between the different scaffolds exists. Bearing in mind that P_{cis} depends upon all the hidden variables for the individual scaffolds, we must try to identify the origins of this heterogeneity.
Different hypotheses could be made to explain the observed differences between scaffolds : they could be due to experimental errors, to differences in local conditions

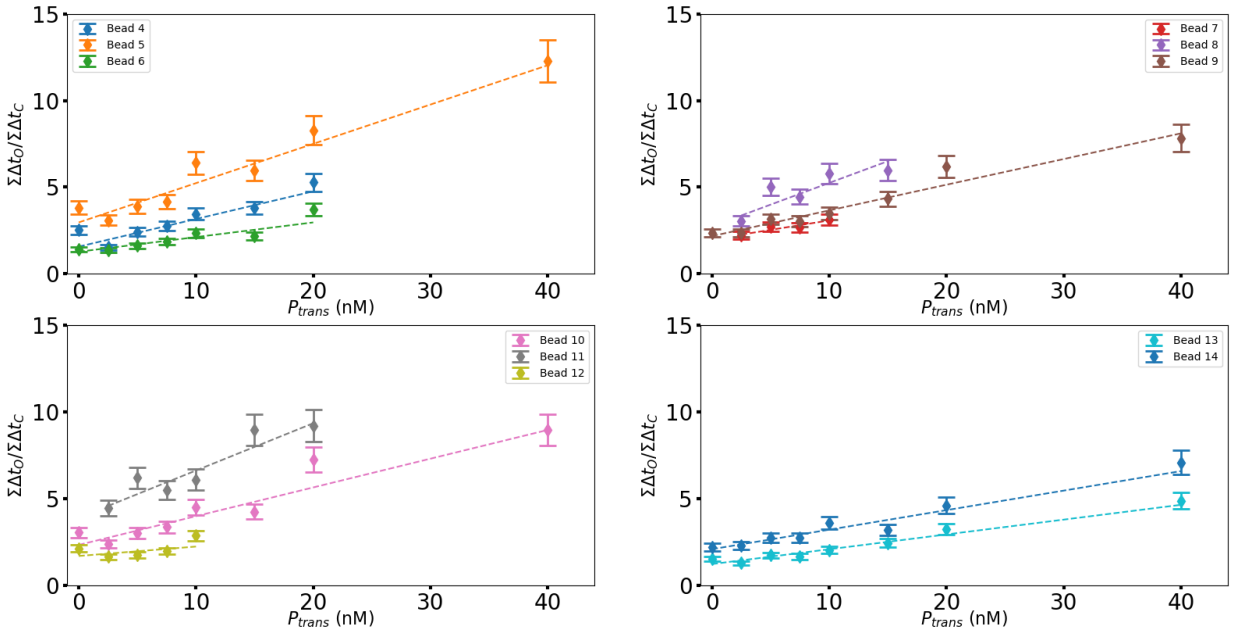


Figure 8.6: Time spent in open and closed states of [FKBP12:JK313] jDNA scaffolds at 28°C under increasing concentrations of FKBP12 (P_{trans}), determined at 28°C in mTOR at 100 fN. Lines are fit by equation $\Sigma\Delta t_O/\Sigma\Delta t_C = I + P_{trans} \times S$

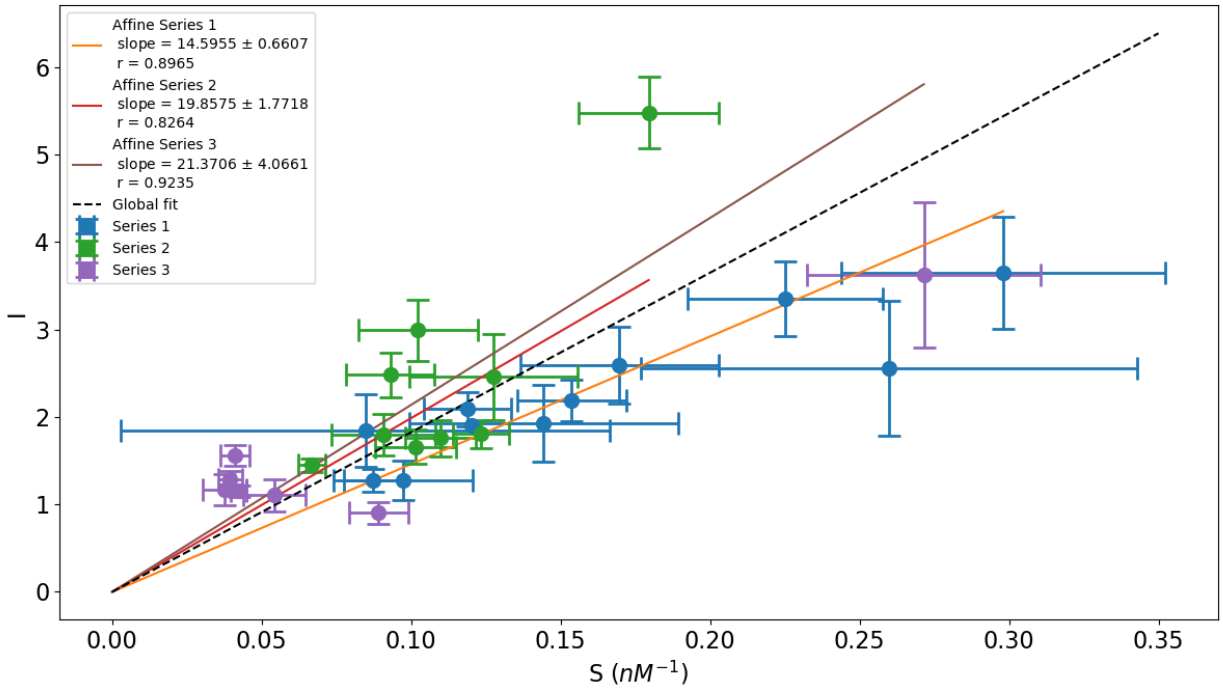


Figure 8.7: Interpolation plots of $I = K \times S$ on all beads from 3 experimental replicates of the experiment. Black line corresponds to the overall fit. Error bars are s.e.m. from previous linear interpolations. All experiments were made on [FKBP12:JK313] jDNA scaffolds at 28°C in mTOR, at 100 fN.

or differences in the individual protein molecules affixed to the scaffolds. We tried different approaches to resolve those issues, covered in sections 8.2.2 and 8.2.3.

Physical significance of P_{cis}

In the course of this work, we have shown how jDNA competition assays can be used to measure protein-ligand interactions. By expressing the in-solution affinity constant K and defining an on-scaffold concentration P_{cis} , we have been able to remove the effects of the force on the scaffolds from our measurements and determine the zero-force value of K.

Defining an equivalent on-scaffold concentration of one ligand relative to the other has allowed us to conceptually equate on-scaffold measurements with in-solution ones. This value, homogeneous to a concentration, could be explained as the concentration of one molecule in the volume of the scaffold. Very summarily, we know that the difference in scaffold height between open and closed state is about 100 nm. If we consider that the molecules move in a sphere of this diameter, the volume is circa 10^{-18} L for one molecule ; this volume containing one molecule we reach a molar concentration of $\frac{1}{N_A V} \approx 1\mu M$. We have determined P_{cis} to be on the order of 10 nM, simple volume considerations are not enough to explain our observations. Multiple other factors certainly influence the closing of the scaffold, *reducing* the probability of interaction.

Applied force Assuming a pulling force of 100 fN, the energy required for contact is $E = F \times d$ with d the distance, which is on the order of $7.8 \times 10^{-21} J$ for one molecule here. This must be compared with the energy barrier for binding and unbinding which, for FKBP12-JK313, has been measured at 120 kJ.mol−1 or 1.6×10^{-19} J for a single molecule (*141*). The effect of the force is significant, but at least one degree of magnitude under the activation energy.

Orientation Solution dynamics assume that molecules can freely rotate. With molecules affixed on the scaffold, their rotation capacity is reduced. While we have not investigate the effects of the change in orientation here, pulling experiments on FKBP12-JK313 complexes show that changing the anchor point of FKBP12 by an angle of 60° changes the energy required for separation by circa $k_B T \approx 4 \times 10^{-21}$ J per molecule, which is low compared to the energy of the interaction itself (*141*).

Shank bending The closing of the scaffold requires the closing of the DNA loop of the shank ; as mentioned in Section 5.3.5, this energy can be expressed as

an equivalent concentration in the form of the j-factor. For dsDNA of 700 bp, we expect a j-factor of about 100 nM ; finding a value on the order of 10 nM for our scaffold is thus not surprising, as the bending is just a prerequisite for association.

One central question in this case is whether a fixed value of P_{cis} can be established for a given scaffold. If such is the case, it would drastically simplify the analysis competition experiments by allowing us to get a reading of affinity with a single concentration measurement. Due to the dispersion of the results we observe, however, it is possible that such a determination could only be done on a per-scaffold basis, or would at the very least involve knowing the force to which a scaffold is exposed.

Variability of P_{cis}

A strong variability has been observed in P_{cis} ; we attempt to propose explanations. We know from experience with the magnetic tweezers that the force exerted on each individual scaffold is not identical - the mass of the beads and the inhomogeneity of the magnetic field account for this. Thus, the question becomes how strong this inhomogeneity is and what is its effect on the interaction. We will assume that K, being the *in-solution* equilibrium coefficient, is not affected by force. However, P_{cis} can easily be affected : indeed, as it represents the "concentration" of the molecules on the scaffold, a stronger pulling force could pull them apart, increasing the volume in which the tips can diffuse. This could explain the strong variability we have observed between scaffolds. We attempt to evaluate the forces pulling on the scaffold in 9.1.
A final hypothesis that we cannot discard out of hand is that the proteins engrafted on the scaffold, while being theoretically identical, are not : they could be denatured - at least partly - or folded slightly differently. Differences in the scaffold because of PCR irregularities are also possible. To see those differences, we would need to first evaluate the two other sources of errors and determine if they explain the differences we see.

8.2.4 Limitations of the method and proposed future developments

Using characteristic time ratios to determine K

We have used the $\Sigma\Delta t_O/\Sigma\Delta t_C$ ratio as the ratio of state occupancy on the scaffold ; however, we could also use the ratio of characteristic times τ_O and τ_C determined by exponential interpolation on the event duration in the same manner. The fact that the open-state events Δt_O are theoretically not single-exponentially distributed has

led us to not use that approach, but as we observe on Fig. 8.5, we can approximate the Δt_O distributions by a single exponential, which would allow the use of this approach.

Solution change

Changing the solution containing the competitors in the experimental channel is not a simple task. We have seen in Section 7.5.4 that in our protocol, we have a contamination of the final solution by the previous one of about 10 % v/v. While this is not a very important issue when the solution is changed with increasing concentrations of competitor, it will be an issue if different competitors are to be used on the same scaffold. The fact that the change is done manually also reduces reproductibility. An improved design to reduce dead volume and possible back-flows from the downstream well is required.
A logical improvement would be to implement microfluidic devices to change to solutions : the channel we currently have is larger than needed (a field of view is at most 150 μm on the side, our channel currently is 3 mm on a side). Solutions such as proposed by (*138*) could conceivably allow for a wash volume of 100 μL or less. There must be more than a single field of view per experiment, so that we may pick and choose the best one, but a smaller - and thinner - channel would allow the use of less buffer volume.
The protocol as described is a fairly good candidate for automation. The changes of solution, repeated twice per day, only requires injection and removal of liquids, which can be automated easily. The set-up of the experiment consists of injections of solutions and movement of magnets, which could also be automated, although the choice of a field of view still requires human action. The storage of the competitor solutions could be an issue if the competition is made by proteins, as they normally require storage at -80 °C. Storage of inorganic competitors is less stringent and the competitor solutions could be pre-prepared, which would allow the experiment to function for a week with no human intervention.

Improving the ligation efficiency of molecules to the scaffold

Our current protocols allow us to observe 15 scaffolds at a time under the microscope, which represents roughly between 20 and 50 % of the total number of molecules in the field of view. Unfortunately, we must also contend with a strong attrition rate, as only 8-12 scaffolds remain capable of interaction through the 1-week experiment. The two main causes of scaffold loss are the detachment of the scaffold from the surface or the bead (rupture of the biotin/streptavidin or digoxygenin/antibody bonds) or the bead getting stuck on the surface through non-specific interaction. To im-

prove the number of scaffolds that can be tracked throughout the experiment, the most important step seems to be the attachment of the protein or ligand oligomers to the scaffold, to reduce the number of scaffolds that show no interaction. This will require the improvement of the ligation of those oligomers to the scaffold (Section 7.2.3) to attain better yields, as this appears to be the limiting step in the protocol.

8.3 Results : comparing the affinity of other drugs for FKBP12

The scaffold can also be used as a tool to probe molecular interactions between FKBP12 and other agonists using the competition protocol. By observing the effects of concentration increases of various agonists on the scaffold dynamics, one can infer the affinity constants of the trans interaction.
The objective here is above all methodological : we wish to establish how much data can be gathered in a week of experiments with a given agonist, whether this gives satisfying results and whether those results are close to already published results.

8.3.1 Data acquisition and analysis

The experiments are conducted identically to the experiments on FKBP12 competition, with one main difference : as the drugs are stored in DMSO and not glycerol buffer, the final experimental buffer is 0.05 % v/v DMSO instead of glycerol. Every different competitor is run in a different sample chamber to avoid cross-contamination, and concentrations are always increasing across a given series. Results are analysed in the same way.
The experiments are conducted in spontaneous-fluctuation, at a fixed force of approximately 100 fN. They are analysed using the algorithm described in section 7.6.

8.3.2 Results and comparison across drugs

Five different drugs have been tested using our protocol. For every one, we have received a preliminary estimate of their K from Pomplun et al. (References (*20*, *21*)) , which has allowed us to select the range of competitor concentration we will be using. Those ranges are different as we need to acquire as many opening and closing events as possible ; thus, we want to have relatively short open state events (recall that closed state events are invariant in duration) and thus keep the open-to-closed duration ratios to a minimum while still exploring larger ranges of concentrations.

The analysis was conducted using the procedure described in Section 8.1.3. The ratio of time spent in the open state relative to the closed state is plotted in function of the competitor concentration. A line is fitted to those points and the intercept I and slope S are calculated. Figure 8.8 illustrates a few different open-to-closed ratios for the different drugs we tested. Since we have shown with Equation 8.10 that $K_{trans} = I/S$, we can deduce the value of K_{trans} for each different scaffold. This value can be averaged to obtain $< K_{trans} >$. Alternatively, we draw a scatter

Table 8.2: Table comparing affinity constants of different studied rapamycin analogues on FKBP12/JK313 jDNA at 28°C with measurements from Hausch et al. (Ref. (*20*, *21*))

Drug	K_{trans} from Ref. (nM)	Mesured K_{trans} from scatter plot (Fig 8.9) (nM)	Average K_{trans} = $< I/S >$ (nM)	Nb. of scaffolds
PPU 278	1.8 ± 0.1	1.3 ± 0.2	1.56 ± 0.59	9
JK 313	59 ± 7	71 ± 11	92 ± 20	8
SP 344	249 ± 2.4	256 ± 15	261 ± 96	13
SP 388	769 ± 24	367 ± 30	366 ± 79	9

plot I in function of S on figure 8.9 ; a linear interpolation going through zero of the intercept to the slope allows us to determine the K equilibrium constant. This yields the measured value for K_{trans}.
Table 8.2 summarizes our findings along with the estimated values of the constants from the literature.
Another visualization tool is the plotting of $Log(I)$ versus $Log(S)$, since we have $Log(K_{trans})+Log(S) = Log(I)$: on such a plot, one can read the value of $-Log(K_{trans}) = pK_D$ as the intercept at $Log(S) = 0$ - see Fig. 8.11.

On- and off-rates in heterogeneous competition

Unfortunately, we cannot determine the rate constants of the trans-association as we are unable to detect trans-associations directly. The precise study of the distribution of the open events could allow us to determine the k_{+cis} but as of yet, we were not able to collect enough data to provide a reliable determination. As shown on Fig. 8.5, there is a possibility of interpreting long event distributions as rate constants, but no conclusive results have been found yet.

8.3.3 Heterogeneous competition : discussion

Efficacy of the method

For the 3 selected competitors we were able to determine an equilibrium constant with a reasonable 10 % margin of error. Our method worked across a wide range of equilibrium concentrations (from nanomolar to micromolar). Attempts at measuring the equilibrium constant of CK249 (estimated coefficient value of 2.5 μM - not shown on figures) did not yield satisfying results (we did not see a significant increase in open times relative to closed from 0 to 5 μM of competitor) so this could be an upper limit to the concentration we can explore. While our results fall relatively close to

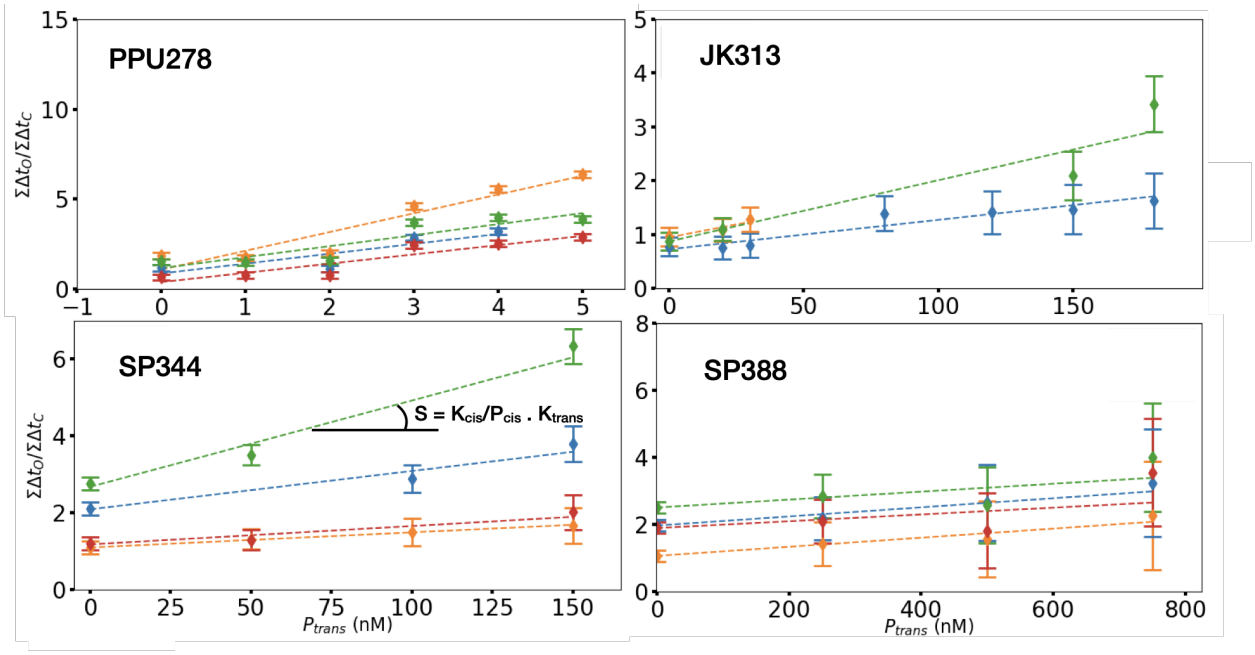


Figure 8.8: Time spent in open and closed states of FKBP12/JK313 for different jDNA scaffolds at 28°C under increasing concentrations of competitor drugs PPU278, SP388, JK313 and SP344. Notice that the x-axis (concentration of competitor) differs for each different competitor. Fits are to the equation $\Sigma\Delta t_O/\Sigma\Delta t_C = I + P_{trans} \times S$.

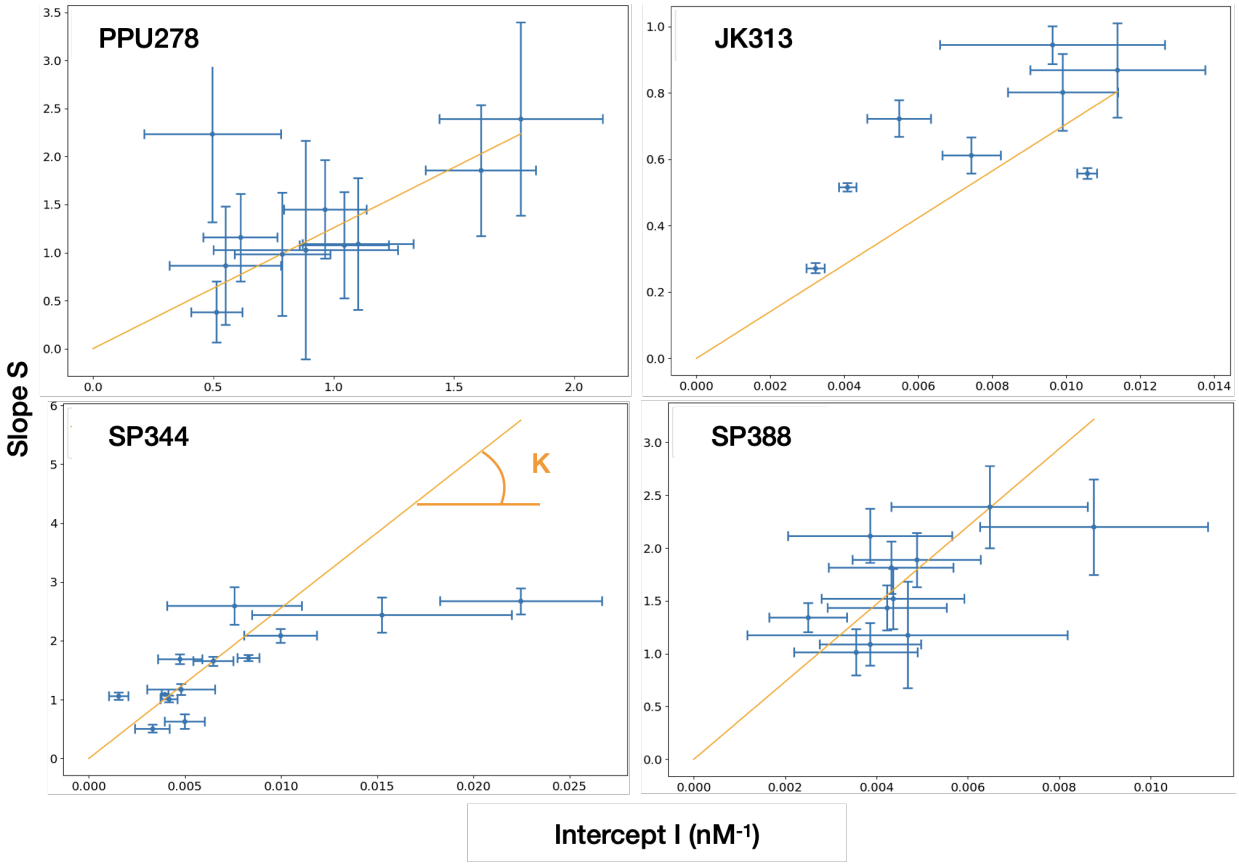


Figure 8.9: Time spent in open and closed states of FKBP12/JK313 jDNA scaffolds at 28°C under increasing concentrations of competitor drugs PPU278, SP388, JK313 and SP344. Orange line fit is to the equation $I = K_{trans} \times S$

the results of Pomplun et al. , some differences must be noted, especially for SP 388, which has the highest K of all our molecules. For this latter case, it must be noted that, for technical reasons, competition could not be achieved at concentrations higher than 750 nM ($\approx K$) - therefore, the range of competition was narrower than for the other drugs, which could lead to reduced accuracy. Experimental repeats on those molecules would be most welcome.
The concentrations of our proteins were assayed by a Bradford test, which typically has an incertitude of 10%, which should also be compounded to our results.

Reverse competition by JK313

JK313 is one of the compounds engrafted on the scaffold ; it has been analysed with all the other drugs here in order to perform a reverse competition test. A strong difference must be noted between the results measured with JK313 in solution ($K = 71 \pm 11$ nM) and FKBP12 in solution ($K = 18.2 \pm 1.9$ nM). The measurement with JK313 in solution is much closer to the result obtained by Ref (*20*) ($K = 59 \pm 7$ nM) than with FKBP12 in solution.
While no definite answer can be given, one can be proposed. FKBP12 is a protein, while JK313 is a small drug molecule. Therefore, the molecular dynamics of FKBP12 could be less affected by its engraftment on the scaffold than JK313, explaining why measuring the interaction between FKBP12 on the scaffold and JK313 in-solution would lead to more accurate measurements than between JK313 on the scaffold and FKBP12 in-solution. However, this observation would require more detailed experiments to be confirmed.

Concentration of the competitor solution

The measurement of protein concentrations is a very difficult topic, universally encountered in titration assays, that we have not attempted to breach. Bradford assays have a typical error rate of 10 %, which is similar to the error rate we have observed in our measurements ; as such, we do not believe that it has a major impact on the validity of our results.

Interest of the method

It is interesting to ask the question of how this method compares with the other methods shown in the introductory chapters of this thesis.
jDNA force-spectroscopy greatly facilitates parallel measurements of single molecules under identical conditions, a property that is shared with acoustic force spectroscopy and, when no force is applied, certain fluorescence-based methods such as TIRF.

Magnetic tweezers are very convenient for long measurements as the force is maintained passively , reducing the number of possible perturbations. Relative to other single-molecule techniques, requires less complex set-ups (an inverted microscope and the magnetic turret) than, for example, AFM or optical tweezers. However, the technique is also less present on the market and most commercial set-ups still require considerable user adjustments. The ability to observe large numbers of molecules at once is rare among SM techniques, shared only by fluorescence and acoustic tweezers. Acoustic tweezers, with their capacity to exert low forces on multiple beads at once, are a good alternative to magnetic tweezers for jDNA competition experiments.

Heterogeneous competition determination of rate constants In heterogeneous competition (the molecule in solution is not on the scaffold), there is no simple way to determine rate constants, as we cannot observe directly any trans-reactions. A path forwards would be to notice that, as depicted on Fig. 8.5 B, the long exponential time constant dominates for long open events. Thus, neglecting τ_{O+} in the distribution would allow us to determine the dominant characteristic time τ_{O-} and use Equation 8.6 to determine the closing rate constant for the trans interaction, and deduce the opening rate from K.
We have not yet managed to obtain satisfying results in that direction however, mostly as this approach necessitates reliable data at high competitor concentrations (higher than P_{cis}), while our experiments were focused on obtaining readings at competitor concentrations close to P_{cis} .

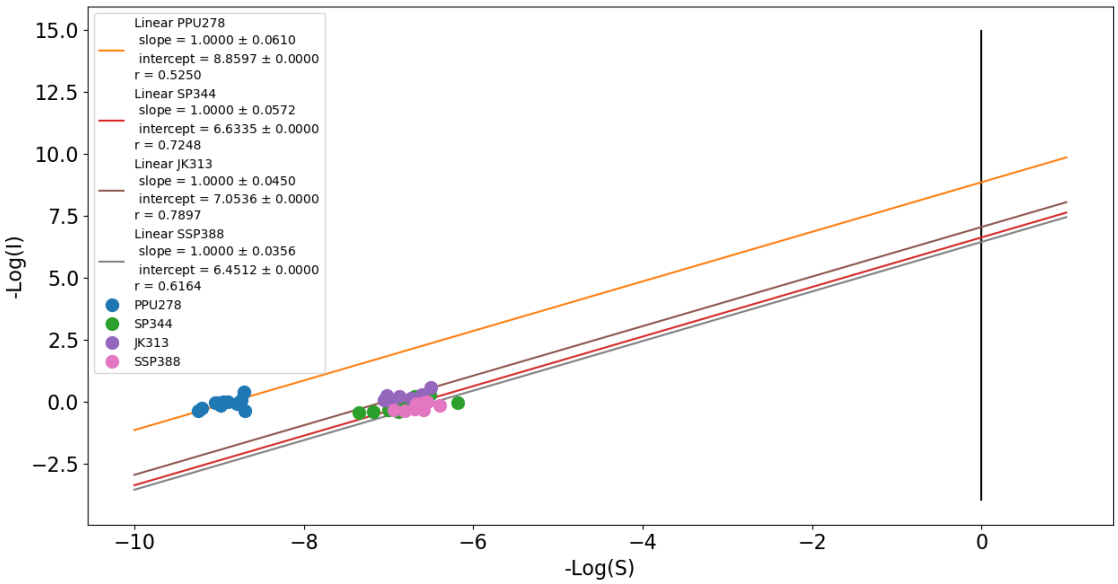


Figure 8.10: Double negative logarithmic plot of I versus S for the 4 competitors studied in this work. Intercept is the pK value. Lines are fits to $-Log(I) = pK - Log(S)$

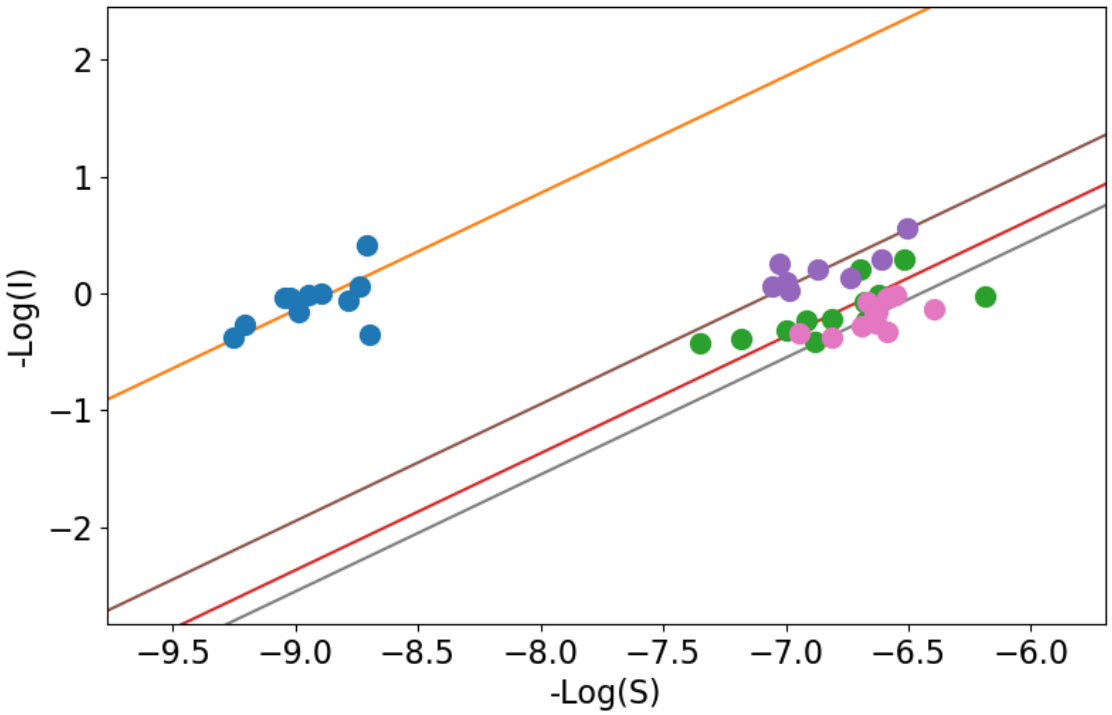


Figure 8.11: Close-up of previous figure

Chapter 9

Mechanical measurements

9.1 Force as a key parameter

jDNA is a construct made of double-stranded DNA and, as we have already mentioned, DNA acts as an entropic spring, its extension under constraint controlled by the worm-like chain equation (*144*). jDNA scaffolds have thus the interesting property of being able to report the force they are under if their extension can be measured. This will be the aim of this Section.

9.1.1 Using the Arrhenius-Bell equation

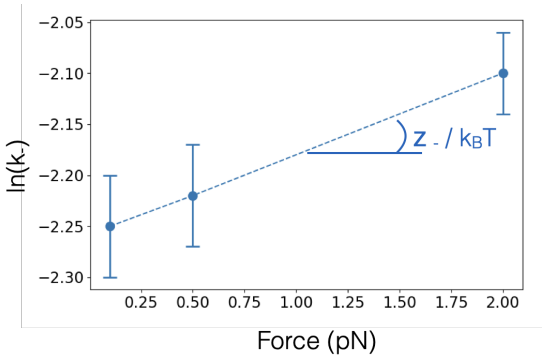


Figure 9.1: Measured k_- on [FKBP12:JK313] jDNA using force-cycling techniques depending on dissociation force. Adapted from Ref. (*141*).

As we have mentioned earlier, it is possible to use jDNA to probe the effects of force on molecular interactions. Follenfant et al. (Ref (*18*, *141*)) have probed the effects of force on the k_- of [FKBP12:JK313] jDNA , as summarized on Fig. 9.1. High forces do indeed increase the observed dissociation rate of the scaffold as can be expected by lowering the energy barrier for dissociation, as shown by the Arrhenius equation (4.1). To get the in-solution rate constants, we use interpolation ; on Figure 9.1, this means drawing the interpolation line and reading the value of k_- on the y-axis intercept. Doing so, however, requires a precise knowledge of the force. Follenfant et al. determines force using the calibration routine described in Section 6.3.3 : the force is assumed to be identical

for all scaffolds under the microscope. However, this assumption is not necessarily true as we will discuss in the following section.
As we have seen in the previous chapter, we observe a strong inter-scaffold variability on K and P_{cis} in our experiments. Since we know force can have an effect on the rate values, we need to find a way to measure the force *individually* on each scaffold to account for it and compensate for it to find zero-force rate values.

9.1.2 Sources of force inhomogeneity

Two factors mainly contribute to the inhomogeneity of the force across a field of view, and prevents us from reliably determining force for every position of the magnets.

Magnetic field inhomogeneity The magnetic field is not exactly homogeneous. That can be seen by moving a single bead across it and tracking its lateral deviations : as those fluctuations are linked to force, one can thus plot a map of force relative to position in the field of view.
This has been investigated in a short experiment. A regular 11-kbp dsDNA molecule has been assembled under the microscope. The force was measured using the method explained in Section 6.3.3 and Equation 6.2 for different positions of the magnets. The extension of the DNA was also assayed : this has allowed us to determine the force-extension relationship for the dsDNA, which follows the worm-like chain equation described in Section 5.3.3. Then the surface was moved under the microscope, with the magnets immobile. The changes in extension of the molecule were recorded at different arrest points, and the extension was used to determine (using the measured WLC equation) precisely the force detected by that molecule at that position under the microscope. The results are represented on Figure 9.2, and show a range of forces between 140 and 170 fN on a single field of view for a single molecule. A reverse experiment, where different scaffolds would be prepared on a surface and then moved to stand at the same coordinate for comparison, has yet to be done.

Bead magnetic moment difference The mass of the beads, their size and magnetic moment are not identical. Those variations can be seen to reach up to 10 % in force, and so contribute greatly to the differences in force observed in a group of scaffolds in a same experiment.

9.1.3 Measuring force on jDNA

The simplest way forwards would be to determine a force-extension relationship using the methods described in Section 6.3.3 : supercoiling a dsDNA molecule

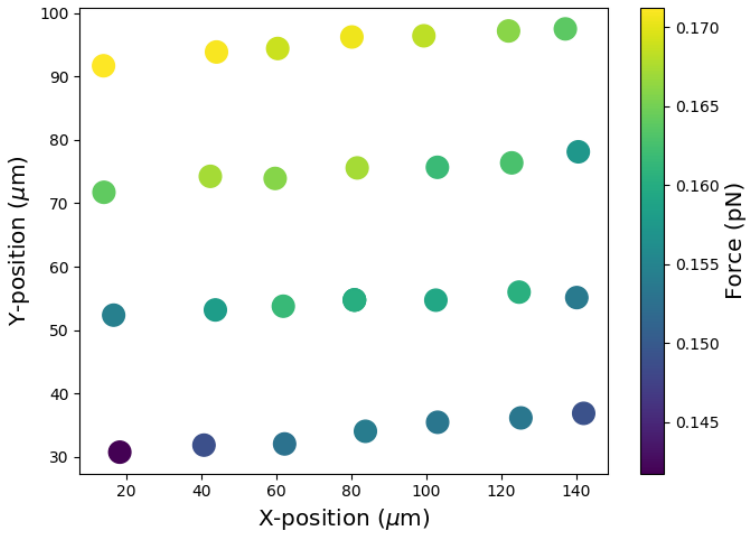


Figure 9.2: Force variations in the field of view of our microscope, measured via force-extension relationship on an 11-kbp dsDNA at 24°C in mTOR buffer. Each point was acquired for 300s for force measurement.

to read contour length, then measure lateral oscillation amplitude across different magnet positions and use the pendulum equation (6.2) to calculate the force. Unfortunately, this proves difficult for technical reasons : as the scaffold is not supercoilable, we cannot get a good measurement of the scaffolds length. True, one could argue that this length is known to us, but two factors make it intrinsically variable : first, the attachment of the scaffold to the surface and the bead is not perfect, an a certain amount of the extreme oligos can be free, increasing contour length. More critically, we cannot guarantee that the bead will be attached to the scaffold at a pole ; the attachment point could be anywhere between a pole and the equator, which drastically reduces the observed contour length of the scaffold, as shown on figure 9.3. The beads we use having 1 μm in diameter, the largest error we can observe is 0.5 μm . Our scaffold being $\approx$ 1.1 μm in length, this is a large

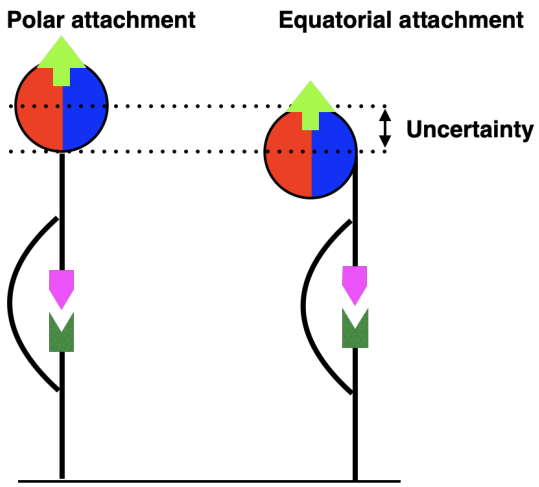


Figure 9.3: Two DNA fragments of the same length can appear to have different contour lengths if they are attached differently to the bead. Pictured is the maximum difference that can be observed between polar and equatorial bead attachment.

error.
One proposed solution to this problem is to use longer molecules, since the error in contour length is fixed (on the order of the bead radius) and using longer scaffolds would decrease its relative importance. This is unfortunately not practical : increasing the length of the scaffold bridge would result in a wider gap between molecules and an exponentially reduced interaction change by reducing the J-factor (see Section 5.3.5) while increasing the length of the arms without lengthening the bridge would reduce the relative change in extension between open and closed conformations, reducing the accuracy of detection.
One parameter that is readily measurable is the difference in extension between the closed scaffold and the open scaffold, Δz. Finding a relation between Δz and force would enable us to know the force on every scaffold in any condition, and will be out aim here.

9.2 Experimental determination of an average Force-extension jump relationship using a magnetic mean field

We can determine the average force-field for a certain configuration of the microscope and magnets. Then, by measuring enough scaffolds under this average field, we could determine an average value of Δz and assume that the average Δz corresponds to the average Force. Repeating this process in different force-fields, an average force-Δz relationship can be established

9.2.1 Measuring a mean force field

To measure the mean force in a field of view, at a certain magnetic position, we use the pendulum equation (cf. 6.2). Calibration uses 11 kb DNA molecules, which can be supercoiled by magnet rotation. We rotate the magnets to positively supercoil the molecules, causing the beads to descend until they hit the surface. Their position is then acquired before restoring the rotation to zero. Thus, the height difference between the fully extended and fully coiled bead z positions is a measurement of contour length. The magnets are then moved and positioned at various positions and a lateral fluctuation measurement is conduced. Using the pendulum equation, the force is determined for that position of the magnets.
We have collated those measurements for 27 identical molecules to compare variations at magnet positions ranging from 0.5 mm above the surface to 2.5 mm above the sample (see Section 6.2.3 for details). The Figure 9.4 shows the logarithm of the

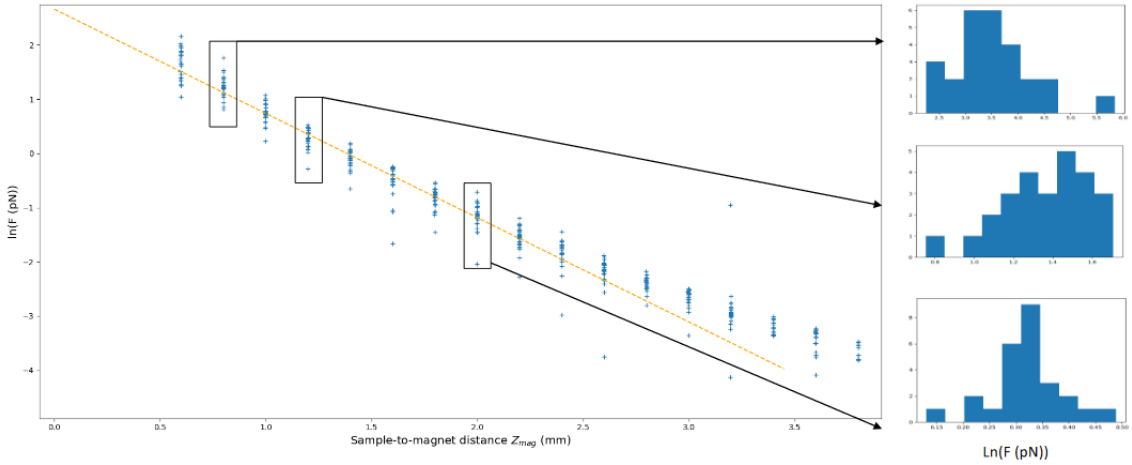


Figure 9.4: Effects of magnet distance on Force measured on 11-kb DNA in mTOR at 24°C. Blue crosses are experimental points collected from 27 different beads under 5 different experiments, in identical conditions. Orange line is a linear fit onto points between 0.5 and 2.5 mm Zmag (magnet-to-sample distance). On the right, histograms of point distributions for indicated Z_{mag} points

measured forces at those positions.

We observe an exponential decay of Force with magnet distance, at least for magnets close to the surface. As our experiments typically happen in the 100 fN range, this exponential dependence is true where we are concerned. At lower forces, the force appears to decrease slower than expected - the reasons for this are beyond the scope of this work.

At higher forces, the line we see in a logarithmic representation leads us to a single-exponential law. Using the linear interpolation, we obtain $F = F_0 e^{-\frac{Z_{mag}}{Z_0}}$ with $F_0 = 14.3 \pm 1.4\ pN$ and $Z_0 = 0.52 \pm 0.03\ mm$ in the force domain that interests us.

9.2.2 Mean force-extension curve for the J-DNA scaffold

On the same microscope, under the same conditions, we then proceeded with a Force-cycle experiment. Scaffolds with no competitor were alternately exposed to a low force (under 10 fN) for 10s and a known force between 6 and 0.1 pN, for varying durations (increased when the force is decreased). When the scaffold was exposed to a low force, it could close easily ; when the force increased, a closed scaffold could be opened. The time required for this opening increases exponentially with force (cf. (*18*)), which is why more time has to be spent at higher lower forces in the higher-force state. An overview of the experiment is depicted in 9.5

In this force-cycling assay, at a given force, we can detect three different extensions of the molecule : one at low force, and two at higher force, one preceding the other : the higher one corresponds to the "open" configuration and the lower one to

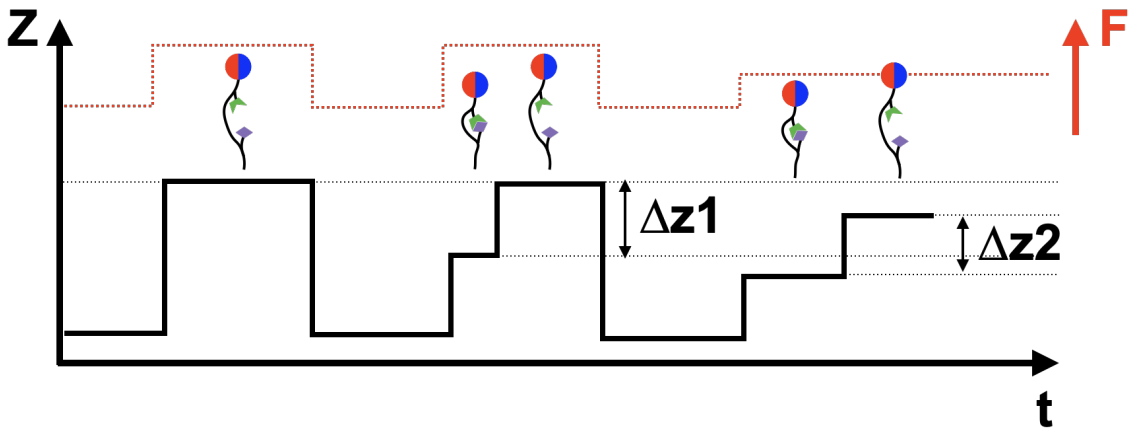


Figure 9.5: Schematic depiction of the force-dz measurement using cycling force. In the first cycle, we see the bead going upwards in response to a force increase, showing DNA scaffold extension. In the second cycle, we observe a transition from the closed to open states while the force is high, and can measure the associated jump amplitude at that force. In the third cycle, at a lower high force, both open and closed states are lower, but the difference between them also decreases ($\Delta z1 > \Delta z2$)

the "closed". Furthermore, as the force decreases, the open and closed states both become lower, but the distance between them also diminishes. That distance is the difference between the closed and open states, Δz.

This relationship between force and the amplitude of the extension jump Δz is analogous to a force-extension curve for a regular piece of DNA. Thus, we at first expected it to follow the worm-like chain equation. We plotted the measured change in extension Δz against the mean force as determined for a given position of the magnets in the previous section and obtained the following curve in 9.6.

We attempted to interpolate this curve using the WLC model using the equation from Section 5.3.3, but that was met with mitigated results. True, we did manage to find a best-fit that explains the observations, but that curve gives us a persistence length of 27 nm, not the 53 nm we would expect for DNA. The contour length was determined to be 0.22 μm, a good fit for an expected value of 0.236 μm.

9.2.3 Comparison with known behavior of short WLC models

Those results are consistent with recent works on short DNA molecules. Indeed, recall that the Worm-Like Chain model supposes that the whole chain is "long" relative to the persistence length, which we could contest here, our bridge only spanning about 4 persistence lengths. The modelling of short polymers (under 10 persistence lengths) has advanced with the use of the Finite Worm-Like Chain model (FWLC) that takes into account the fact that the chain is, well, finite, as well as boundary conditions at the edges (*145*). Very recently (*146*), this model was used to

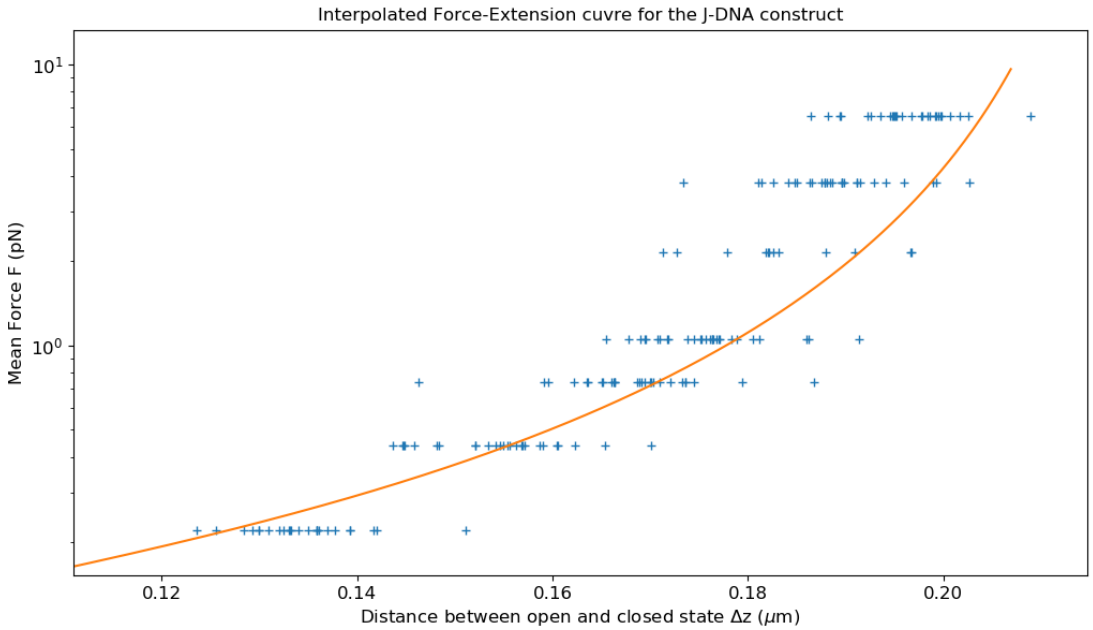


Figure 9.6: Semi-logarithmic plot of jump amplitudes for different mean-force values, obtained for 25 beads among 5 independent series. The orange line corresponds to a WLC model (Equation 5.1) best fit with contour length estimated at $0.22\mu m$ and persistence length estimated at $27 \pm 0.4\ nm$

experimentally fit sub-micrometer DNA molecules. An empirical formula emerged, relating the "observed" persistence length (in the WLC model) of a short DNA strand to its contour length L_C :

$$\xi_{WLC} = \frac{\xi_\infty}{1 + \frac{\alpha \xi_\infty}{L}} \tag{9.1}$$

Where ξ_{WLC} is the equivalent persistence length as fitted by the Worm-Like Chain, ξ_∞ is the persistence length of a sufficiently long polymer (53 nm for DNA) and α is a fitting constant determined by (*145*) to be 2.78. In our case, we have $\xi_\infty = 53$ nm, and the contour length of our DNA is 230 nm, leading to an expected ξ_{WLC} value of 32 nm - not exactly our result, but much closer.

It seems therefore that the "bridge" of jDNA behaves similarly to a short strand of DNA. Differences can be explained by experimental uncertainty, but also by the boundary conditions (freely rotating here versus stiffly attached in the previously mentioned experiments) and the fact that our bridge does have different structural constraints than a regular DNA strand.

9.3 Direct determination of force under a scaffold

We have attempted an alternate strategy to determine directly the force on a scaffold from data obtained in-situ.

9.3.1 Pendulum equation on scaffolds

Recall that the force is related to the lateral Brownian motions of the bead via the pendulum equation (cf. 6.2) :

$$F = \frac{k_B T L}{< \delta x^2 >} \tag{9.2}$$

For a given scaffold in a constant magnetic field, the force is constant, as neither the magnets, nor the bead changes, and its movement is negligible on the relevant scales (largely under 1 micrometer in every direction, in comparison to the surface-to-magnet distance measured in millimeters). However, the contour length changes between scaffold opening and closing, by a distance dz. Thus, using the pendulum equation in both closed and open states, we can write that the force is equal between the closed conformation (contour length L_0) and the open conformation ($L_0 + \Delta z$) :

$$F = \frac{L_0 + \Delta z}{< \delta x_0^2 >} = \frac{L_O}{< \delta x_C^2 >}$$

This allows us to deduce the contour length of the molecule :

$$L_0 = \Delta z \left(\frac{< \delta x_O^2 >}{< \delta x_C^2 >} - 1 \right)^{-1} \tag{9.3}$$

And thus the Force:

$$F = \frac{k_B T \Delta z}{< \delta x_O^2 > - < \delta x_C^2 >} \tag{9.4}$$

With x_O (resp. x_C) the open and closed x-positions. This equation can theoretically allow us to determine the force of any scaffold under observation but requires very accurate determinations of both lateral and vertical fluctuations of bead position to reach an acceptable error margin.

9.3.2 Determining force using the pendulum equation directly on scaffolds

The approach we propose here is simple : on a given trace, we know the time at which the scaffold is open or closed. We extract the x and y positions of the bead at every time and classify them accordingly. Then, we evaluate the lateral fluctuations in both conformations, and deduce the force using Equation 9.4. To date, we have not been able to obtain accurate enough force measurements using this technique. Our main issue is that we need to remove certain high frequencies from the lateral fluctuations as shown on Fig. 6.4, but since we need to work with data recomposed

from different segments of varying lengths, this process is complex. To evaluate the quality of this process, we apply the analysis to subsequent acquisitions on the same scaffolds, as they should exhibit identical forces.
To-date, we have been able to obtain estimations of the force of a correct order of magnitude (100 fN), but not precise enough to reliably distinguish between beads, with on-scaffold variations on the order of 50 fN.

9.4 Physical interpretation of the Cis-concentration parameter

We have suggested previously that the parameter we call P_{cis} , the equivalent concentration of both partners on the scaffold, could be used to take into account the force with our analysis.
Using our results on the titration of the [FKBP12:JK313] jDNA scaffold with FKBP12 in solution, we plotted the correlation between I (K/P_{cis}) and S $(1/P_{cis})$ and the force, calculated using the force-curve extension determined in 9.2. The results are depicted on Fig 9.7 (A and B). The values of K and P_{cis} are determined for each scaffold according to Equation 8.14 ; their dependence with force is plotted on the same figure (C and D).

The first surprising observation is the very large force dispersion that we observe, stronger than would be expected. This could be due to the fact that our determination of Δz is less precise than expected (recall from Section 7.6 that the z noise we observe is not exactly Gaussian, which diminishes our ability to determine $< z >$ precisely).
The resulting scatter plots do not yield a clear linear interpolation ($r < 0.5$). We can nevertheless observe a trend of force-dependence of P_{cis} . This is consistent with our hypotheses, and supports our assumption that the value we obtain for K is the zero-force. The same trend is less visible for K, and might be purely artefactual - however, it also cannot be said that K is independent of force.

9.5 Discussion : effects of force on jDNA interactions

The results presented in this sections are not conclusive, but strongly suggest a relationship between force and jDNA dynamics. However, as has been pointed out in 9.1.1, we do not believe that the force differences observed in our experiments are enough to cause significant changes in the dissociation dynamics of FKBP12-JK313.

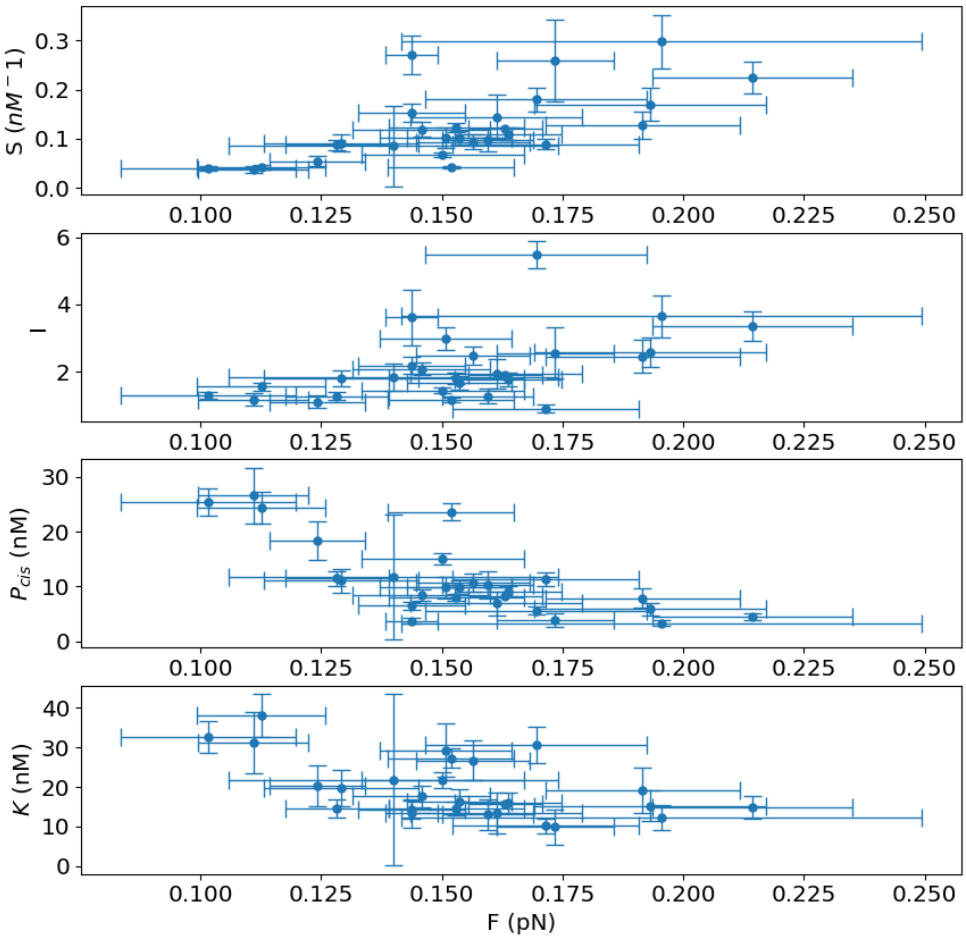


Figure 9.7: Force-dependence of (A) I, (B) S, (C) K and (D) P_{cis} . Error bars are s.e.m. Data is from homogeneous competition experiments presented in Section 8.2, on [FKBP12:JK313] jDNA scaffolds at 28 °C in mTOR.

Therefore, the effects of force would be mainly on the dynamics of the association, which not only includes the molecular association itself, but also the closing step of the scaffold. P_{cis} representing the virtual concentration of one partner on the scaffold relative to the other, it seems natural to postulate that it would be most affected by force : the extension of the scaffold bridge, controlled *via* the WLC equation by force, defines the space in which the partners can diffuse and encounter one another. By the definition of P_{cis} and K, on the other hand, K represents the in-solution equilibrium constant, and is thus expected to be force-independent.
An experimentally observed dependence of K on force would imply that one of the hypotheses made in Section 8 are not verified. Either the dissociation rate constant k_- is force-dependent (which, as we have discussed in Section 9.1.1, seems to not be the case) or it is not true that the association rates are governed by the ratios of cis and trans concentrations alone (recall from Section 8.1.1 that we assume that $K_{+cis}/k_{+}trans = P_{cis}/P_{trans}$). More experimental repeats will be required before conclusions can be drawn on this topic.
Finally, it will also be necessary to determine whether the force explains all the heterogeneity we observe on P_{cis}. If this is not the case, other possible sources of dispersion will need to be explored, such as the possibility of different interaction properties of individual proteins, maybe because of folding differences, or chemical damage.

Part IV

Discussion

Chapter 10

Conclusion and perspectives

10.1 jDNA titration assays for rate and equilibrium measurements

In this work, we have shown that single-molecule scaffolds can be used to obtain accurate readings of the association and dissociation rate constants, as well as equilibrium constants. Using a titration against a known concentration of competitor in the solution, it is possible to obtain the values that would be measured in bulk assays, removing the effects of force and scaffold dynamics on the measurement. Equilibrium constants can be measured with a simple population analysis, comparing state occupancy of scaffold conformation, while rate constants can be accessed using the distribution of conformation durations.
Using a competitor interacting with the scaffold, but not present on the scaffold, it is possible to obtain the equilibrium constant for interactions between that molecule in solution and its partner on the scaffold. We have shown that this method can explore a wide range of competitors, across two orders of magnitude of equilibrium constants.

10.2 jDNA and competition in the context of single-molecule assays

Versatility of competition methods Competition appears as yet another tool in the design of single-molecule assays, one that bridges the gap between single-molecule and in-solution measurements. Competition assays offer a view of interactions that cannot be directly detected on the scaffold, which allows a single construct to serve in a wide range of experiments. This approach can be extended to many

different scaffolds and single-molecule approaches, including optical tweezers and fluorescence measurements (*147*). Competition on scaffolds has even been used in bulk measurements on molecular scaffolds undergoing electrophoretic migration (*148*).

Improvement of jDNA scaffolds DNA is a key component of scaffolds since it allows both structural rigidity (through physical properties) and experimental flexibility (thanks to its chemical properties). The capacity to construct synthetic DNA efficiently allows a scaffold to incorporate handles, structuring elements, engraft samples (proteins, drugs, DNA hairpins, etc.) and even accessory sensor elements. The development of specific molecular "rulers" that report on force is a recent advancement. For instance, Ref. (*143*) uses a DNA hairpin incorporated in the scaffold to measure the force exerted on each individual scaffold. Other approaches, based on the DNA origami, can report on the force detected in an interaction (*149*), or even exert a specific force (*108*, *150*, *151*). FRET fluorescent probes can provide additional information on the conformation of a scaffold, or on the position of proteins on it (*152*, *153*) .

Building on this work, many new improvements to the jDNA scaffold could be explored. The structure of the jDNA scaffold has an influence on the dynamics of the interaction. While its design was guided by work on the closing of dsDNA loops, the fact that the leash is bound to the shanks with a single-strand molecular bond can change the dynamics of the closing. Experiments with shorter or longer leashes could offer additional insight into the relationship between P_{cis} and the length of the leash. It has been shown that, for short dsDNA, the j-factor increases exponentially with length (*129*) ; this would have to be confirmed here. But since the current length of the leash has been chosen to be close to a length that maximizes the j-factor, it is unlikely that massive gains in association rates can be achieved through this approach. The use of a single-strand leash, with a lower persistence length, could also increase the closing probability, but at the cost of a higher experimental noise.

The length of the tips bearing the protein and ligands could also be a factor to consider ; in this work, they have been chosen short so that they behave rigidly, but an argument could be made to choose them longer to reduce the average distance between the molecular partners and reduce P_{cis} , increasing association rate. Faster reactions would allow proportionately shorter acquisition times. While we cannot change the dissociation time, decreasing P_{cis} would reduce the τ_O, increasing the amount of data collected per experiment.

Finally, the length of the shanks is also a factor to consider. Longer scaffolds exhibit more noise, but they also reduce the number of interactions with the surface, which

could alleviate certain data-analysis issues encountered in Sections 7.6 and 8.2.3.

Improving the temporal resolution Collecting shorter events could also a possible way towards reducing the duration of experiments. The use of a faster camera is a possibility, but the viscous friction exerted on the bead by the buffer during a conformation change is also an issue, as we have seen. To reduce its impact, smaller beads could be utilized, with a corresponding increase of the magnetic force field to compensate. Smaller beads offer a shorter relaxation time, but their increased sensitivity to Brownian motion will increase experimental noise.

The use of single-molecule scaffolds, and DNA scaffolds in particular is an ongoing development in the field of molecular biology and biological chemistry. This work takes place in a global effort to streamline the use of DNA as a scaffolding element, with the aim to create tools that are robust and flexible for multiple types of analyses. Competition and jDNA scaffolds are both tools that can be integrated in other approaches for a wide variety of purposes.

10.3 jDNA competition assays and drug discovery

This work began as our team was developing the jDNA construct to study nonhomologous end-joining DNA repair (*97*). The focus then shifted to using the same construct to study protein-ligand interactions under high force. The natural next step was to focus on interactions at low forces and explore potential applications to drug discovery.
The jDNA competition methodology we demonstrated offers a few distinctive advantages to the established ensemble methods.

Reagent consumption The consumption of reagents is an important factor to consider for proteins, especially human or viral ones that are difficult to produce. Stopped-flow, the most conservative ensemble method, requires 100 μL of reagent at a concentration close to the equilibrium constant for each test ; this bar could be attained by using a microfluidic set-up in our experiments. Furthermore, the consumption of engrafted molecules is extremely low, using only 1 μL of 50 pM scaffold solution, which is appreciable for the study of rare proteins, such as viral proteins or mammalian large proteic complexes.

Temporal resolution For the type of interactions studied here, temporal resolution is not limiting : with an imaging frequency of 30 Hz, events of characteristic

times in the tens of seconds can be fully measured. Stopped-flow and surface plasmon resonnance have temporal resolutions that can reach the milisecond, but the limiting factor is the timescale of the perturbation that is used to measure a kinetic. Single-molecule experiments bypass this constraint by being equilibrium measurements and requiring no perturbation, eliminating an important possible source of bias. Shorter reactions can be measured by increasing the imaging frequency.

Duration of the experiment The jDNA competition method still needs to be improved in terms of throughput and capabilities to be used in a more systematic way. Our experiments were made on a week, for 8-16 hours per point of concentration. With $\tau_C \approx 10$ s and $\tau_O \approx 30 - 100$ s, this means that we acquire between 1400 and 250 open and closed cycles per condition. It appears from our observations that 200 cycles are enough to see clear distribution histograms, so that the experiment could be shortened. The number of different concentrations to be explored could also be reduced, potentially allowing the testing of different competitors in the same experimental run. Combined with microfluidic approaches, this could allow us to run multiple different drugs per week of experiments and get a comparable level of precision.

jDNA for drug discovery ? The jDNA competition assay is a self-contained assay, with a single scaffold enabling the exploration of interactions with multiple different partners, force-dependence and energy landscape. This allows us to propose the jDNA as a valid tool for single-molecule analysis of drug-target interactions, offering parallel measurements, low consumption of protein and direct analysis. jDNA-based techniques are thus well-suited to drug refinement steps of the discovery process, when a reduced number of candidate molecules are compared to find the best candidates for in-vivo assays.

The usage of this technique, however, must be compared to the techniques already existing in the industry and academia. While bulk techniques do have drawbacks - consumption of reagent and necessity of using probes on the proteins for most of them - they are also widely accepted and used, which makes them very attractive for industrial applications. The most established techniques, such as stopped-flow, have been optimized for tens of years to give the best signal and consume minimal amounts of sample : it will be difficult for a new technique to challenge their dominance of the market.

The development of single-molecule techniques, therefore, will be easier for specific applications, most critically the measurement of force and the determination of the energy landscape of a given reaction. Understanding the energy landscape of a drug-target interaction is essential for effective pharmacodynamics, but of particular note

for the study of complex structural interactions, such as the von Wildebrandt factor (*72*) or the T-cell antigen surface receptor (*154*), where tension plays a role in the physiological functions, or the study of molecular motors, such as the RNA polymerase, for instance. Moving forwards, as the role of force in physiological processes is better understood, drug molecules taking force into account can be expected to emerge.

Bibliography

(*1*) Benjamin, D., Colombi, M., Moroni, C., and Hall, M. N. (2011). Rapamycin passes the torch: a new generation of mTOR inhibitors. *Nature Reviews Drug Discovery 10*, 868–880.

(*2*) Vézina, C., Kudelski, A., and Sehgal, S. N. (1975). Rapamycin, a new antifungal antibiotic. *The Journal of Antibiotics 28*, 721–726.

(*3*) Martel, R. R., Klicius, J., and Galet, S. (2011). Inhibition of the immune response by rapamycin, a new antifungal antibiotic. *Canadian Journal of Physiology and Pharmacology*, DOI: 10.1139/y77-007.

(*4*) Harrison, D. E., Strong, R., Sharp, Z. D., Nelson, J. F., Astle, C. M., Flurkey, K., Nadon, N. L., Wilkinson, J. E., Frenkel, K., Carter, C. S., Pahor, M., Javors, M. A., Fernandez, E., and Miller, R. A. (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature 460*, 392–395.

(*5*) Schreiber, S. L., and Crabtree, G. R. (1992). The mechanism of action of cyclosporin A and FK506. *Immunology Today 13*, 136–142.

(*6*) Kino, T., Hatanaka, H., Hashimoto, M., Nishiyama, M., Goto, T., Okuhara, M., Kohsaka, M., Aoki, H., and Imanaka, H. (1987). FK-506, A novel immunosuppressant isolated from streptomyces. *The Journal of Antibiotics 40*, 1249–1255.

(*7*) Kang, C. B., Hong, Y., Dhe-Paganon, S., and Yoon, H. S. (2008). FKBP Family Proteins: Immunophilins with Versatile Biological Functions. *Neurosignals 16*, 318–325.

(*8*) Kolos, J. M., Voll, A. M., Bauder, M., and Hausch, F. (2018). FKBP Ligands—Where We Are and Where to Go? *Frontiers in Pharmacology 9*, DOI: 10.3389/fphar.2018.01425.

(*9*) Gaali, S. et al. (2015). Selective inhibitors of the FK506-binding protein 51 by induced fit. *Nature Chemical Biology 11*, 33–37.

(*10*) Yang, H., Jiang, X., Li, B., Yang, H. J., Miller, M., Yang, A., Dhar, A., and Pavletich, N. P. (2017). Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40. *Nature 552*, 368–373.

(*11*) Huang, K., and Fingar, D. C. (2014). Growing knowledge of the mTOR signaling network. *Seminars in Cell & Developmental Biology 36*, 79–90.

(*12*) Tokunaga, C., Yoshino, K.-i., and Yonezawa, K. (2004). mTOR integrates amino acid- and energy-sensing pathways. *Biochemical and Biophysical Research Communications 313*, 443–446.

(*13*) Sarbassov, D. D., Guertin, D. A., Ali, S. M., and Sabatini, D. M. (2005). Phosphorylation and Regulation of Akt/PKB by the Rictor-mTOR Complex. *Science 307*, 1098–1101.

(*14*) Shiota, C., Woo, J.-T., Lindner, J., Shelton, K. D., and Magnuson, M. A. (2006). Multiallelic Disruption of the rictor Gene in Mice Reveals that mTOR Complex 2 Is Essential for Fetal Growth and Viability. *Developmental Cell 11*, 583–589.

(*15*) Kim, J., and Guan, K.-L. (2019). mTOR as a central hub of nutrient signalling and cell growth. *Nature Cell Biology 21*, 63–71.

(*16*) Motzer, R. J., Escudier, B., Oudard, S., Hutson, T. E., Porta, C., Bracarda, S., Grünwald, V., Thompson, J. A., Figlin, R. A., Hollaender, N., Urbanowitz, G., Berg, W. J., Kay, A., Lebwohl, D., and Ravaud, A. (2008). Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. *The Lancet 372*, 449–456.

(*17*) Rao, R. D., Mladek, A. C., Lamont, J. D., Goble, J. M., Erlichman, C., James, C. D., and Sarkaria, J. N. (2005). Disruption of Parallel and Converging Signaling Pathways Contributes to the Synergistic Antitumor Effects of Simultaneous mTOR and EGFR Inhibition in GBM Cells. *Neoplasia (New York, N.Y.) 7*, 921–929.

(*18*) Kostrz, D., Wayment-Steele, H. K., Wang, J. L., Follenfant, M., Pande, V. S., Strick, T. R., and Gosse, C. (2019). A modular DNA scaffold to study protein–protein interactions at single-molecule resolution. *Nature Nanotechnology 14*, 988–993.

(*19*) Babine, R. (1995). Design, synthesis and X-ray crystallographic studies of novel FKBP-12 ligands. *Bioorganic & Medicinal Chemistry Letters 5*, 1719–1724.

(*20*) Pomplun, S. et al. (2018). Chemogenomic Profiling of Human and Microbial FK506-Binding Proteins. *Journal of Medicinal Chemistry 61*, 3660–3673.

(*21*) Ranganath Gopalakrishnan, Kozany, C., Gaali, S., Kress, C., Hoogeland, B., Bracher, A., and Hausch, F. (2012). Evaluation of Synthetic FK506 Analogues as Ligands for the FK506-Binding Proteins 51 and 52. *Journal of Medicinal Chemistry 55*, 4114–4122.

(*22*) Israelachvili, J. N. In *Intermolecular and Surface Forces (Third Edition)*, Israelachvili, J. N., Ed.; Academic Press: Boston, 2011, pp 151–167.

(*23*) Eigen, M., and Hammes, G. G. In *Advances in Enzymology and Related Areas of Molecular Biology*; John Wiley & Sons, Ltd: 1963, pp 1–38.

(*24*) Eigen, M., and Rigler, R. (1994). Sorting single molecules: application to diagnostics and evolutionary biotechnology. *Proceedings of the National Academy of Sciences 91*, 5740–5747.

(*25*) Callender, R., and Dyer, R. B. (2006). Advances in Time-Resolved Approaches To Characterize the Dynamical Nature of Enzymatic Catalysis. *Chemical Reviews 106*, 3031–3042.

(*26*) Berg, O. G., and von Hippel, P. H. (1985). Diffusion-controlled macromolecular interactions. *Annual Review of Biophysics and Biophysical Chemistry 14*, 131–160.

(*27*) Lu, H., and Tonge, P. J. (2010). Drug–target residence time: critical information for lead optimization. *Current Opinion in Chemical Biology 14*, 467–474.

(*28*) Núñez, S., Venhorst, J., and Kruse, C. G. (2012). Target–drug interactions: first principles and their application to drug discovery. *Drug Discovery Today 17*, 10–22.

(*29*) Bierer, B. E., Mattila, P. S., Standaert, R. F., Herzenberg, L. A., Burakoff, S. J., Crabtree, G., and Schreiber, S. L. (1990). Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. *Proceedings of the National Academy of Sciences 87*, 9231–9235.

(*30*) Hoa, X. D., Kirk, A. G., and Tabrizian, M. (2007). Towards integrated and sensitive surface plasmon resonance biosensors: A review of recent progress. *Biosensors and Bioelectronics 23*, 151–160.

(*31*) Schuck, P. (1997). Reliable determination of binding affinity and kinetics using surface plasmon resonance biosensors. *Current Opinion in Biotechnology 8*, 498–502.

(*32*) Löfås, S., and Johnsson, B. (1990). A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *Journal of the Chemical Society, Chemical Communications 0*, 1526–1528.

(*33*) Johnsson, B., Löfås, S., and Lindquist, G. (1991). Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. *Analytical Biochemistry 198*, 268–277.

(*34*) Banaszynski, L. A., Liu, C. W., and Wandless, T. J. (2005). Characterization of the FKBP·Rapamycin·FRB Ternary Complex. *Journal of the American Chemical Society 127*, 4715–4721.

(*35*) Sultana, A., and Lee, J. E. (2015). Measuring Protein-Protein and Protein-Nucleic Acid Interactions by Biolayer Interferometry. *Current Protocols in Protein Science 79*, 19.25.1–19.25.26.

(*36*) Wilson, J. L., Scott, I. M., and McMurry, J. L. (2010). Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. *Biochemistry and Molecular Biology Education 38*, 400–407.

(*37*) Hall, M. D., Yasgar, A., Peryea, T., Braisted, J. C., Jadhav, A., Simeonov, A., and Coussens, N. P. (2016). Fluorescence polarization assays in high-throughput screening and drug discovery: a review. *Methods and applications in fluorescence 4*, 022001.

(*38*) Moerke, N. J. (2009). Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding. *Current Protocols in Chemical Biology 1*, 1–15.

(*39*) Freyer, M. W., and Lewis, E. A. In *Methods in Cell Biology*; Biophysical Tools for Biologists, Volume One: In Vitro Techniques, Vol. 84; Academic Press: 2008, pp 79–113.

(*40*) Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., Disdier, F., Baltzinger, M., and Dumas, P. (2012). kinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *Journal of the American Chemical Society 134*, 559–565.

(*41*) Zogg, A., Stoessel, F., Fischer, U., and Hungerbühler, K. (2004). Isothermal reaction calorimetry as a tool for kinetic analysis. *Thermochimica Acta 419*, 1–17.

(*42*) Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J., and Speicher, D. W. (1984). Cyclophilin: a specific cytosolic binding protein for cyclosporin A. *Science 226*, 544–547.

(*43*) Tamura, T., Kioi, Y., Miki, T., Tsukiji, S., and Hamachi, I. (2013). Fluorophore Labeling of Native FKBP12 by Ligand-Directed Tosyl Chemistry Allows Detection of Its Molecular Interactions in Vitro and in Living Cells. *Journal of the American Chemical Society 135*, 6782–6785.

(*44*) Lu, C., and Wang, Z.-X. (2017). Quantitative Analysis of Ligand Induced Heterodimerization of Two Distinct Receptors. *Analytical Chemistry 89*, 6926–6930.

(*45*) Pellecchia, M. (2005). Solution Nuclear Magnetic Resonance Spectroscopy Techniques for Probing Intermolecular Interactions. *Chemistry & Biology 12*, 961–971.

(*46*) Leone, M., Crowell, K. J., Chen, J., Jung, D., Chiang, G. G., Sareth, S., Abraham, R. T., and Pellecchia, M. (2006). The FRB Domain of mTOR: NMR Solution Structure and Inhibitor Design, *Biochemistry 45*, 10294–10302.

(*47*) Mayer, M., and Meyer, B. (1999). Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition 38*, 1784–1788.

(*48*) Estévez-Torres, A., Gosse, C., Le Saux, T., Allemand, J.-F., Croquette, V., Berthoumieux, H., Lemarchand, A., and Jullien, L. (2007). Fourier Analysis To Measure Diffusion Coefficients and Resolve Mixtures on a Continuous Electrophoresis Chip. *Analytical Chemistry 79*, 8222–8231.

(*49*) Piehler, J. (2005). New methodologies for measuring protein interactions in vivo and in vitro. *Current Opinion in Structural Biology 15*, 4–14.

(*50*) Roberts, G., and Watts, A., *Encyclopedia of biophysics*; Springer: 2019.

(*51*) Eccleston, J. F., Martin, S. R., and Schilstra, M. J. In *Methods in Cell Biology*; Biophysical Tools for Biologists, Volume One: In Vitro Techniques, Vol. 84; Academic Press: 2008, pp 445–477.

(*52*) Zrelli, K., Barilero, T., Cavatore, E., Berthoumieux, H., Le Saux, T., Croquette, V., Lemarchand, A., Gosse, C., and Jullien, L. (2011). Temperature Modulation and Quadrature Detection for Selective Titration of Two-State Exchanging Reactants. *Analytical Chemistry 83*, 2476–2484.

(*53*) Quérard, J., Le Saux, T., Gautier, A., Alcor, D., Croquette, V., Lemarchand, A., Gosse, C., and Jullien, L. (2016). Kinetics of Reactive Modules Adds Discriminative Dimensions for Selective Cell Imaging. *ChemPhysChem 17*, 1396–1413.

(*54*) Pearson, D. S., Holtermann, G., Ellison, P., Cremo, C., and Geeves, M. A. (2002). A novel pressure-jump apparatus for the microvolume analysis of protein–ligand and protein–protein interactions: its application to nucleotide binding to skeletal-muscle and smooth-muscle myosin subfragment-1. *Biochemical Journal 366*, 643–651.

(*55*) Silva, J. L., Foguel, D., and Royer, C. A. (2001). Pressure provides new insights into protein folding, dynamics and structure. *Trends in Biochemical Sciences 26*, 612–618.

(*56*) Woodside, M. T., and Block, S. M. (2014). Reconstructing Folding Energy Landscapes by Single-Molecule Force Spectroscopy. *Annual Review of Biophysics 43*, 19–39.

(*57*) Gupta, A. N., Vincent, A., Neupane, K., Yu, H., Wang, F., and Woodside, M. T. (2011). Experimental validation of free-energy-landscape reconstruction from non-equilibrium single-molecule force spectroscopy measurements. *Nature Physics 7*, 631–634.

(*58*) Dietz, H., and Rief, M. (2004). Exploring the energy landscape of GFP by single-molecule mechanical experiments. *Proceedings of the National Academy of Sciences 101*, 16192–16197.

(*59*) Evans, E. (2001). Probing the relation between force–lifetime–and chemistry in single molecular bonds. *Annual Review of Biophysics and Biomolecular Structure 30*, 105–128.

(*60*) Bell, G. I. (1978). Models for the specific adhesion of cells to cells. *Science 200*, Publisher: American Association for the Advancement of Science Section: Articles, 618–627.

(*61*) Merkel, R., Nassoy, P., Leung, A., Ritchie, K., and Evans, E. (1999). Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature 397*, 50–53.

(*62*) Sedlak, S. M., Schendel, L. C., Gaub, H. E., and Bernardi, R. C. (2020). Streptavidin/biotin: Tethering geometry defines unbinding mechanics. *Science Advances 6*, DOI: `10.1126/sciadv.aay5999`.

(*63*) Sakmann, B., and Neher, E. (1984). Patch clamp techniques for studying ionic channels in excitable membranes. *Annual Review of Physiology 46*, 455–472.

(*64*) Varongchayakul, N., Song, J., Meller, A., and W. Grinstaff, M. (2018). Single-molecule protein sensing in a nanopore: a tutorial. *Chemical Society Reviews 47*, 8512–8524.

(*65*) Dudko, O. K., Mathé, J., and Meller, A. In *Methods in Enzymology*, Walter, N. G., Ed.; Single Molecule Tools, Part B:Super-Resolution, Particle Tracking, Multiparameter, and Force Based Methods, Vol. 475; Academic Press: 2010, pp 565–589.

(*66*) Lin, J., Fabian, M., Sonenberg, N., and Meller, A. (2012). Nanopore Detachment Kinetics of Poly(A) Binding Proteins from RNA Molecules Reveals the Critical Role of C-Terminus Interactions. *Biophysical Journal 102*, 1427–1434.

(*67*) Zlatanova, J., Lindsay, S. M., and Leuba, S. H. (2000). Single molecule force spectroscopy in biology using the atomic force microscope. *Progress in Biophysics and Molecular Biology 74*, 37–61.

(*68*) Weisel, J. W., Shuman, H., and Litvinov, R. I. (2003). Protein–protein unbinding induced by force: single-molecule studies. *Current Opinion in Structural Biology 13*, 227–235.

(*69*) Heinz, W. F., and Hoh, J. H. (1999). Spatially resolved force spectroscopy of biological surfaces using the atomic force microscope. *Trends in Biotechnology 17*, 143–150.

(*70*) Moffitt, J. R., Chemla, Y. R., Smith, S. B., and Bustamante, C. (2008). Recent Advances in Optical Tweezers. *Annual Review of Biochemistry 77*, 205–228.

(*71*) Ashkin, A. (1970). Acceleration and Trapping of Particles by Radiation Pressure. *Physical Review Letters 24*, 156–159.

(*72*) Kim, J., Zhang, C.-Z., Zhang, X., and Springer, T. A. (2010). A mechanically stabilized receptor–ligand flex-bond important in the vasculature. *Nature 466*, 992–995.

(*73*) Greenleaf, W. J., Frieda, K. L., Foster, D. A. N., Woodside, M. T., and Block, S. M. (2008). Direct Observation of Hierarchical Folding in Single Riboswitch Aptamers. *Science 319*, 630–633.

(*74*) Greenfeld, M., Pavlichin, D. S., Mabuchi, H., and Herschlag, D. (2012). Single Molecule Analysis Research Tool (SMART): An Integrated Approach for Analyzing Single Molecule Data. *PLOS ONE 7*, e30024.

(*75*) Graves, E. T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., and Strick, T. R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. *Nature Structural & Molecular Biology 22*, 452–457.

(*76*) Gosse, C., and Croquette, V. (2002). Magnetic Tweezers: Micromanipulation and Force Measurement at the Molecular Level. *Biophysical Journal 82*, 3314–3329.

(*77*) Ozcelik, A., Rufo, J., Guo, F., Gu, Y., Li, P., Lata, J., and Huang, T. J. (2018). Acoustic tweezers for the life sciences. *Nature Methods 15*, 1021–1028.

(*78*) Neuman, K. C., and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods 5*, 491–505.

(*79*) Kamsma, D., Creyghton, R., Sitters, G., Wuite, G. J. L., and Peterman, E. J. G. (2016). Tuning the Music: Acoustic Force Spectroscopy (AFS) 2.0. *Methods 105*, 26–33.

(*80*) Pierres, A., Touchard, D., Benoliel, A.-M., and Bongrand, P. (2002). Dissecting Streptavidin-Biotin Interaction with a Laminar Flow Chamber. *Biophysical Journal 82*, 3214–3223.

(*81*) Pierres, A., Vitte, J., Benoliel, A.-M., and Bongrand, P. (2006). Dissecting individual ligand–receptor bonds with a laminar flow chamber. *Biophysical Reviews and Letters 01*, Publisher: World Scientific Publishing Co., 231–257.

(*82*) Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K. (1997). Direct observation of the rotation of F1-ATPase. *Nature 386*, 299–302.

(*83*) Wang, Y., Guo, L., Golding, I., Cox, E. C., and Ong, N. P. (2009). Quantitative Transcription Factor Binding Kinetics at the Single-Molecule Level. *Biophysical Journal 96*, 609–620.

(*84*) Erickson, H. P. (2009). Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online 11*, 32–51.

(*85*) van Oijen, A. M. (2011). Single-molecule approaches to characterizing kinetics of biomolecular interactions. *Current Opinion in Biotechnology 22*, 75–80.

(*86*) Karymov, M. A., Krasnoslobodtsev, A. V., and Lyubchenko, Y. L. (2007). Dynamics of Synaptic SfiI-DNA Complex: Single-Molecule Fluorescence Analysis. *Biophysical Journal 92*, 3241–3250.

(*87*) Kasper, R., Harke, B., Forthmann, C., Tinnefeld, P., Hell, S. W., and Sauer, M. (2010). Single-Molecule STED Microscopy with Photostable Organic Fluorophores. *small 6*, 1379–1384.

(*88*) Miyake, T., Tanii, T., Sonobe, H., Akahori, R., Shimamoto, N., Ueno, T., Funatsu, T., and Ohdomari, I. (2008). Real-Time Imaging of Single-Molecule Fluorescence with a Zero-Mode Waveguide for the Analysis of Protein−Protein Interaction. *Analytical Chemistry 80*, 6018–6022.

(*89*) Yokota, H. (2020). Fluorescence microscopy for visualizing single-molecule protein dynamics. *Biochimica et Biophysica Acta (BBA) - General Subjects 1864*, DOI: 10.1016/j.bbagen.2019.05.005.

(*90*) Elenko, M. P., Szostak, J. W., and van Oijen, A. M. (2009). Single-Molecule Imaging of an in Vitro-Evolved RNA Aptamer Reveals Homogeneous Ligand Binding Kinetics. *Journal of the American Chemical Society 131*, 9866–9867.

(*91*) García-Sáez, A. J., and Schwille, P. (2008). Fluorescence correlation spectroscopy for the study of membrane dynamics and protein/lipid interactions. *Methods 46*, 116–122.

(*92*) Winkler, T., Kettling, U., Koltermann, A., and Eigen, M. (1999). Confocal fluorescence coincidence analysis: An approach to ultra high-throughput screening. *Proceedings of the National Academy of Sciences 96*, 1375–1378.

(*93*) Kinz-Thompson, C. D., Bailey, N. A., and Gonzalez, R. L. In *Methods in Enzymology*, Spies, M., and Chemla, Y. R., Eds.; Single-Molecule Enzymology: Fluorescence-Based and High-Throughput Methods, Vol. 581; Academic Press: 2016, pp 187–225.

(*94*) Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory 13*, Conference Name: IEEE Transactions on Information Theory, 260–269.

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

(*96*) Lehéricy, L. (2019). Consistent order estimation for nonparametric hidden Markov models. *Bernoulli 25*, 464–498.

(*97*) Wang, J. L., Duboc, C., Wu, Q., Ochi, T., Liang, S., Tsutakawa, S. E., Lees-Miller, S. P., Nadal, M., Tainer, J. A., Blundell, T. L., and Strick, T. R. (2018). Dissection of DNA double-strand-break repair using novel single-molecule forceps. *Nature Structural & Molecular Biology 25*, 482–487.

(*98*) Halvorsen, K., Schaak, D., and Wong, W. P. (2011). Nanoengineering a single-molecule mechanical switch using DNA self-assembly. *Nanotechnology 22*, 494005.

(*99*) Krasnoslobodtsev, A. (2015). A Flexible Nanoarray Approach for the Assembly and Probing of Molecular Complexes. *Biophysical Journal 108*, 2333–2339.

(*100*) Gao, Y., Zorman, S., Gundersen, G., Xi, Z., Ma, L., Sirinakis, G., Rothman, J. E., and Zhang, Y. (2012). Single Reconstituted Neuronal SNARE Complexes Zipper in Three Distinct Stages. *Science 337*, 1340–1343.

(*101*) Rognoni, L., Stigler, J., Pelz, B., Ylänne, J., and Rief, M. (2012). Dynamic force sensing of filamin revealed in single-molecule experiments. *Proceedings of the National Academy of Sciences 109*, Publisher: National Academy of Sciences Section: Biological Sciences, 19679–19684.

(*102*) Koirala, D., Shrestha, P., Emura, T., Hidaka, K., Mandal, S., Endo, M., Sugiyama, H., and Mao, H. (2014). Single-Molecule Mechanochemical Sensing Using DNA Origami Nanostructures. *Angewandte Chemie 126*, 8275–8279.

(*103*) Kilchherr, F., Wachauf, C., Pelz, B., Rief, M., Zacharias, M., and Dietz, H. (2016). Single-molecule dissection of stacking forces in DNA. *Science 353*, DOI: `10.1126/science.aaf5508`.

(*104*) Strick, T. R., Croquette, V., and Bensimon, D. (2000). Single-molecule analysis of DNA uncoiling by a type II topoisomerase. *Nature 404*, 901–904.

(*105*) Kostiuk, G., Dikić, J., Schwarz, F. W., Sasnauskas, G., Seidel, R., and Siksnys, V. (2017). The dynamics of the monomeric restriction endonuclease BcnI during its interaction with DNA. *Nucleic Acids Research 45*, 5968–5979.

(*106*) van den Broek, B., Noom, M. C., and Wuite, G. J. L. (2005). DNA-tension dependence of restriction enzyme activity reveals mechanochemical properties of the reaction pathway. *Nucleic Acids Research 33*, 2676–2684.

(*107*) Manosas, M., Camunas-Soler, J., Croquette, V., and Ritort, F. (2017). Single molecule high-throughput footprinting of small and large DNA ligands. *Nature Communications 8*, 304.

(*108*) Li, N., Wang, J., Ma, K., Liang, L., Mi, L., Huang, W., Ma, X., Wang, Z., Zheng, W., Xu, L., Chen, J.-H., and Yu, Z. (2019). The dynamics of forming a triplex in an artificial telomere inferred by DNA mechanics. *Nucleic Acids Research 47*, e86–e86.

(*109*) Gosse, C., Strick, T. R., and Kostrz, D. (2019). Molecular scaffolds: when DNA becomes the hardware for single-molecule investigations. *Current Opinion in Chemical Biology 53*, 192–203.

(*110*) Watson, J. D., and Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature 171*, 737–738.

(*111*) Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature 362*, 709–715.

(*112*) Harvey, L., and Arnold, B. In *Molecular Cell Biology*, 7th ed.; macmillian higher education, pp 115–164.

(*113*) Castro, C. E., Kilchherr, F., Kim, D.-N., Shiao, E. L., Wauer, T., Wortmann, P., Bathe, M., and Dietz, H. (2011). A primer to scaffolded DNA origami. *Nature Methods 8*, 221–229.

(*114*) Vale, G., and Dell'Orto, P. (1992). Non-readioactive nucleiec acid probes: labelling and detection procedures. *Liver 12*, _eprint: https://onlinelibrary.wiley.com/doi/pdf/0676.1992.tb01056.x, 243–251.

(*115*) Höltke, H. J., Ankenbauer, W., Mühlegger, K., Rein, R., Sagner, G., Seibl, R., and Walter, T. (1995). The digoxigenin (DIG) system for non-radioactive labelling and detection of nucleic acids–an overview. *Cellular and Molecular Biology 41*, 883–905.

(*116*) Vologodskii, A. (1994). DNA Extension under the Action of an External Force. *Macromolecules 27*, 5623–5625.

(*117*) Bouchiat, C., Wang, M. D., Allemand, J. .-.-F., Strick, T., Block, S. M., and Croquette, V. (1999). Estimating the Persistence Length of a Worm-Like Chain Molecule from Force-Extension Measurements. *Biophysical Journal 76*, 409–413.

(*118*) Strick, T. R., Allemand, J.-F., Bensimon, D., Bensimon, A., and Croquette, V. (1996). The Elasticity of a Single Supercoiled DNA Molecule. *Science 271*, 1835–1837.

(*119*) Shimada, J., and Yamakawa, H. (1985). Statistical mechanics of DNA topoisomers: The helical worm-like chain. *Journal of Molecular Biology 184*, 319–329.

(*120*) Mosconi, F., Allemand, J. F., Bensimon, D., and Croquette, V. (2009). Measurement of the Torque on a Single Stretched and Twisted DNA Using Magnetic Tweezers. *Physical Review Letters 102*, 078301.

(*121*) Baumann, C. G., Smith, S. B., Bloomfield, V. A., and Bustamante, C. (1997). Ionic effects on the elasticity of single DNA molecules. *Proceedings of the National Academy of Sciences 94*, 6185–6190.

(*122*) Strick, T. R., Allemand, J. .-.-F., Bensimon, D., and Croquette, V. (1998). Behavior of Supercoiled DNA. *Biophysical Journal 74*, 2016–2028.

(*123*) Allemand, J.-F. Micro-Manipulation de molécules d'ADN isolées, thesis, Paris: Paris 6, 1997, 120 pp.

(*124*) Smith, S. B., and Bendich, A. J. (1990). Electrophoretic charge density and persistence length of DNA as measured by fluorescence microscopy. *Biopolymers 29*, 1167–1173.

(*125*) Wang, M. D., Yin, H., Landick, R., Gelles, J., and Block, S. M. (1997). Stretching DNA with optical tweezers. *Biophysical Journal 72*, 1335–1346.

(*126*) Marko, J. F., and Siggia, E. D. (1995). Stretching DNA. *Macromolecules 28*, 8759–8770.

(*127*) Marko, J., Siggia, E., Bustamante, C., and Smith, S. (1994). Entropic elasticity of lambda-phage DNA. *265*, 1599–1600.

(*128*) Shimada, J., and Yamakawa, H. (1984). Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules 17*, Publisher: American Chemical Society, 689–698.

(*129*) Tardin, C. (2017). The mechanics of DNA loops bridged by proteins unveiled by single-molecule experiments. *Biochimie 142*, 80–92.

(*130*) Allemand, J.-F., Cocco, S., Douarche, N., and Lia, G. (2006). Loops in DNA: An overview of experimental and theoretical approaches. *The European Physical Journal E 19*, 293–302.

(*131*) Levene, S. D., Giovan, S. M., Hanke, A., and Shoura, M. J. (2013). The thermodynamics of DNA loop formation, from J to Z. *Biochemical Society Transactions 41*, 513–518.

(*132*) Douarche, N., and Cocco, S. (2005). Protein-mediated DNA loops: Effects of protein bridge size and kinks. *Physical Review E 72*, Publisher: American Physical Society, 061902.

(*133*) Ringrose, L., Chabanis, S., Angrand, P.-O., Woodroofe, C., and Stewart, A. (1999). Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *The EMBO Journal 18*, 6630–6641.

(*134*) Strick, T. Enroulement mecanique de l'adn et relaxation par les topoisomerases, Publication Title: http://www.theses.fr, thesis, Paris 6, 1999.

(*135*) Kolb, H. C., Finn, M. G., and Sharpless, K. B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition 40*, 2004–2021.

(*136*) Hong, V., Presolski, S. I., Ma, C., and Finn, M. G. (2009). Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angewandte Chemie 121*, 10063–10067.

(*137*) Click Chemistry tools - Copper free click chemistry (SPAAC), Click Chemistry Tools `https://clickchemistrytools.com/cu-free-click-chemistry-spaac/` (accessed Nov. 26, 2020).

(*138*) Gosse, C., Stanescu, S., Frederick, J., Lefrançois, S., Vecchiola, A., Moskura, M., Swaraj, S., Belkhou, R., Watts, B., Haltebourg, P., Blot, C., Daillant, J., Guenoun, P., and Chevallard, C. (2020). A pressure-actuated flow cell for soft X-ray spectromicroscopy in liquid media. *Lab on a Chip 20*, 3213–3229.

(*139*) Hmmlearn hmmlearn `https://hmmlearn.readthedocs.io/en/latest/index.html` (accessed Dec. 11, 2020).

(*140*) Scipy scipy.optimize.curve_fit — SciPy v1.5.4 Reference Guide `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html` (accessed Dec. 14, 2020).

(*141*) Follenfant, M., Stransky, F., Kostrz, D., Wang, J. L., J.Kolos, P.Purder, C.Meyners, F.Hausch, Strick, T., and Gosse, C. Characterization of drug-target binding lifetimes using a junctured-DNA (jDNA) scaffold combined to magnetic tweezers, Poster presentation, Poster presentation, 64th annual meeting of the Biophysical Society, San Diego, California, 2020.

(*142*) Shelley, C., and Magleby, K. L. (2008). Linking Exponential Components to Kinetic States in Markov Models for Single-Channel Gating. *Journal of General Physiology 132*, 295–312.

(*143*) Rieu, M., Valle-Orero, J., Ducos, B., Allemand, J.-F., and Croquette, V. (2020). Fluorescence-free quantification of protein/nucleic-acid binding through single-molecule kinetic locking. *bioRxiv*, 2020.09.30.321232.

(*144*) Smith, S. B., Finzi, L., and Bustamante, C. (1992). Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science 258*, 1122–1126.

(*145*) Seol, Y., Li, J., Nelson, P. C., Perkins, T. T., and Betterton, M. D. (2007). Elasticity of Short DNA Molecules: Theory and Experiment for Contour Lengths of 0.6–7m. *Biophysical Journal 93*, 4360–4373.

(*146*) Shon, M. J., Rah, S.-H., and Yoon, T.-Y. (2019). Submicrometer elasticity of double-stranded DNA revealed by precision force-extension measurements with magnetic tweezers. *Science Advances 5*, eaav1697.

(*147*) Schickinger, M., Zacharias, M., and Dietz, H. (2018). Tethered multifluorophore motion reveals equilibrium transition kinetics of single DNA double helices. *Proceedings of the National Academy of Sciences 115*, E7512–E7521.

(*148*) Koussa, M. A., Halvorsen, K., Ward, A., and Wong, W. P. (2015). DNA nanoswitches: a quantitative platform for gel-based biomolecular interaction analysis. *Nature Methods 12*, Number: 2 Publisher: Nature Publishing Group, 123–126.

(*149*) Hudoba, M. W., Luo, Y., Zacharias, A., Poirier, M. G., and Castro, C. E. (2017). Dynamic DNA Origami Device for Measuring Compressive Depletion Forces. *ACS Nano 11*, 6566–6573.

(*150*) Nickels, P. C., Wünsch, B., Holzmeister, P., Bae, W., Kneer, L. M., Grohmann, D., Tinnefeld, P., and Liedl, T. (2016). Molecular force spectroscopy with a DNA origami–based nanoscopic force clamp. *Science 354*, 305–307.

(*151*) Blanchard, A. T., and Salaita, K. (2019). Emerging uses of DNA mechanical devices. *Science 365*, 1080–1081.

(*152*) Long, X., Parks, J. W., Bagshaw, C. R., and Stone, M. D. (2013). Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy. *Nucleic Acids Research 41*, 2746–2755.

(*153*) Mitra, J., Makurath, M. A., Ngo, T. T. M., Troitskaia, A., Chemla, Y. R., and Ha, T. (2019). Extreme mechanical diversity of human telomeric DNA revealed by fluorescence-force spectroscopy. *Proceedings of the National Academy of Sciences 116*, 8350–8359.

(*154*) Liu, Y., Blanchfield, L., Ma, V. P.-Y., Andargachew, R., Galior, K., Liu, Z., Evavold, B., and Salaita, K. (2016). DNA-based nanoparticle tension sensors reveal that T-cell receptors transmit defined pN forces to their antigens for enhanced fidelity. *Proceedings of the National Academy of Sciences 113*, 5610–5615.

RÉSUMÉ

Nous présentons ici une nouvelle méthode de mesure des interactions entre molécules biologiques. Ceci est utilisé dans la recherche médicale, pour savoir si un médicament va fonctionner sur une cible par exemple. Pour cela, nous accrochons deux molécules que nous voulons étudier à un "échafaudage" moléculaire observable au microscope. Ainsi, nous pouvons voir si les molécules accrochées se lient ou se détachent, et à quelle vitesse. De plus, l'échafaudage étant aimanté, nous pouvons "tirer" sur ces molécules avec une pince magnétique et donc appliquer une force, ou encore ajouter d'autres molécules (par exemple, un autre médicament) dans le milieu et comparer l'efficacité de différentes molécules. Le travail présenté ici se concentre sur le développent de cet outil et sur les outils d'analyse informatique associés. Nous appliquons ces techniques à l'étude du système FKBP12-Rapamycine, central dans la régulation du métabolisme et de la réponse immunitaire.

MOTS CLÉS

Biochimie, Interaction, Molécule unique, Physique statistique, Biophysique, Cinétique

ABSTRACT

We present here a novel method measuring interactions between biological molecules. This is useful in medical research, to determine if a drug will work on a target, for example. To do that, we stick two molecules we want to study to a molecular "scaffold" we can observe under the microscope. Thus, we can see if the molecules bind or separate, and at which speeds. In addition, the scaffold being magnet-sensitive, we can "pull" on those molecules with a magnetic tweezer and apply force, or add other molecules (for example, other drugs) in the medium and compare the efficiency of different molecules. This thesis focuses on the design and improvement of this tool and of the associated data analysis tools. We apply our technique to the study of the FKBP12-Rapamycin system, essential for the regulation of metabolism and immune responses.

KEYWORDS

Biochemistry, Interaction, Single-Molecule, Statistical physics, Biophysics, Kinetics