UNIVERSITE PIERRE & MARIE CURIE
LA SCIENCE A PARIS

Institut Formation Doctorale

UNIVERSITA' DEGLI STUDI DI PADOVA

Sede Amministrativa:

Università degli Studi di Padova

Sede Consorziata:

Université Pierre et Marie Curie, Paris IV

Dipartimento di FISICA

DOTTORATO DI RICERCA IN FISICA

CICLO XIX

DATA CONSEGNA TESI: 31 Luglio 2008

THESE DE DOCTORAT DE
L'UNIVERSITE PIERRE ET MARIE CURIE

Spécialité

LOGIQUE DU VIVANT

Présentée par

M. Francesco Mosconi

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse/Titolo Tesi :

Fluctuations in biological molecules: tools to probe mechanical and structural properties of DNA and proteins

soutenue le ……

Coordinatore : Ch.mo Prof. Attilio Stella, U.Padova

devant le jury composé de :

Supervisore: Ch.mo Prof. Amos Maritan, U.Padova	Co-Directeur de thèse
Supervisore: M. David Bensimon, DR-CNRS UMR 8550	Co-Directeur de thèse
M. Jean-François Joanny, Professeur Université Paris IV	Rapporteur
M. Richard Lavery, DR-CNRS UMR 5086	Rapporteur
Ch.mo Prof. Fabio Mammano, U.Padova	Examinateur
Ch.mo Prof. Giampaolo Mistura, U.Padova	Examinateur

Dottorando : Francesco Mosconi

Université Pierre & Marie Curie - Paris 6
Bureau d'accueil, inscription des doctorants et base de données
Esc G, 2ème étage
15 rue de l'école de médecine
75270-PARIS CEDEX 06

Tél. Secrétariat : 01 42 34 68 35
Fax : 01 42 34 68 40
Tél. pour les étudiants de A à EM : 01 42 34 69 54
Tél. pour les étudiants de EN à ME : 01 42 34 68 41
Tél. pour les étudiants de MF à Z : 01 42 34 68 51
E-mail : scolarite.doctorat@upmc.fr

Alla mia famiglia, che ha sempre avuto fiducia in me e ad Anna, che mi ha insegnato la pazienza

Contents

Abstract

The traditional boundary between hard science (physics and mathematics) and soft sciences (chemistry and biology) is progressively fading away as the complexity inherent in the biological world is understood and mapped out thanks to a joint attack on two fronts. On the one side more quantitative experiments allow to investigate the details of the atomic structures of biological molecules and to measure with greater precision the laws of interaction among different molecules; on the other side, the massive introduction of information technology in the management and catalogation of the multitude of molecular components found inside a cell is allowing to gain deep insights in the complex dynamic equilibrium that regulates the network of interactions amog different molecules.

The work described in this thesis concerns the first side of the battlefield: the development of new techniques to allow quantitative measurements of biologically relevant quantities.

The work consisted in the design, construction and validation of three different experiments dealing with proteins and DNA mechanics. Key components of the cellular microcosmos, DNA and proteins are large macromolecules that constantly interact and accomplish most of the tasks needed by the cell to survive. The first part of the thesis summarises the known properties of these molecules and introduces the motivations driving the designed experiments.

Proteins catalyse chemical reactions in the cell and their threedimensional configuration gives each of them its specific function. The connections between structural and chemical properties of a protein are a suject largely unexplored.The second part of the thesis describes an experiment based on single molecule fluorescence microscopy designed to explore the dynamics of fluctuations of catalytic activity of a single enzyme.

The experiments described in this part have not yet given the hoped results. However part of the preliminary considerations done when building these setup were used to write the article F. Mosconi et al. "Some nonlinear challenges in biology", Nonlinearity 21 (2008) T131-T147.

DNA stores the genetic information needed by the cell to accomplish its tasks,

and such information must be physically stored, read, written and restored in different times during the cell cycle. The importance of a proper knowledge of its mechanical properties is fundamental if its interaction with proteins is to be understood. The third part of this thesis describes two different experiments based on the magnetic tweezers micro-manipulation technique, that attempt to measure some not yet entirely characterised mechanical properties of DNA.

The two experiments presented in this part gave interesting results. A new determination of the biologically relevant parameter C, the twist modulus of DNA was obtained developing a novel type of analysis to data collected using the standard magnetic tweezers apparatus. Also, a new type of "soft" magnetic tweezers that allows the simultaneous application of an external force and an external torque has been developed and validated to measure the torque response of a DNA molecule. The results described in this part of the thesis are summarised in two papers that are ready to be submitted.

Résumé

La division traditionnelle entre sciences dures (physique et mathématique) et sciences molles (chimie et biologie) devient de plus en plus faible en vue du fait que la complexité propre du monde biologique est comprise grâce à un attaque sur deux fronts. D'un coté, des expériences plus quantitatives permettent d'étudier les détails de la structure atomique des molécules biologiques et de mesurer avec une grande précision les lois d'interaction entre différentes molécules; d'autre part, l'introduction massive de l'informatique dans la gestion et la catalogation de la multitude de composants moléculaires qui se trouvent dans la cellule, permets d'obtenir une compréhension profonde du complexe équilibre dynamique qui régule le réseau des interactions entre différent molécules.

Le travail décrit dans cette thèse regarde le premier coté du champ de bataille: le développement de nouvelles techniques qui permettent des mesures quantitatives de quantités d'intérêt biologique.

Le travail a compris le projet, la construction et la validation de trois différents expériences concertantes la mécanique des protéines ou de l'ADN. Composants clefs du microcosme cellulaire, l'ADN et les protéines sont des grandes macromolecules qui interagissent continuellement et qui accomplissent la plus part des taches nécessaires pour la vie de la cellule. La première partie de cette thèse est un résume des propriétés connues de ces molécules et introduises les motivations qui sont a la base des expériences effectuées.

Les protéines catalysent les réactions chimiques dans la cellule et leur configuration trois-dimensionnel donne a chaque protéine sa fonction spécifique. Les connections entre les propriétés structurels et chimiques d'une protéine sont un domaine qui reste encore a explorer en détail. La deuxième partie de la thèse décrives un expérience basée sur les techniques de microscopie de fluorescence et de molécule unique pour explorer la dynamique des fluctuations d'activité d'un enzyme isole.

Les expériences décrites dans cette partie n'ont pas encore donne les résultats espérés. Cependant, une partie des considérations préliminaires a la construction de ces expériences a été utilisée pour l'écriture de l'article F. Mosconi et al. "Some

nonlinear challenges in biology", Nonlinearity 21 (2008) T131-T147.

L'ADN contiens l'information génétique nécessaire à la cellule pour accomplir toutes ses fonctions. Cette information doit être gardé, lue, copiée t restaurée a différents moments du cycle cellulaire. L'importance d'une connaissance détaillée des propriétés mécaniques de l'ADN est fondamental si l'on veut comprendre ses interactions avec les protéines. La troisième partie de cette thèse décrives deux différents expériences basées sur la technique des "pinces magnétiques", qui essayent de mesurer des propriétés mécaniques de l'ADN qui n'ont pas encore étés complètement caractérisées.

Les deux expériences présentées dans cette dernière partie ont donné des résultats intéressants. Une nouvelle détermination du paramètre d'intérêt biologique C, le module de torsion de l'ADN, a été obtenue en appliquant un nouveau type d'analyse a des données mesurées avec les "pinces magnétiques" traditionnelles. Aussi, une nouvelle type de pinces magnétiques "douces" qui permettent d'appliquer une force externe ainsi qu'un couple a été développé et validée pour mesurer la réponse de l'ADN a un couple externe. Les résultats décrits dans cette partie de la thèse sont résumées dans deux papiers qui sont prêts pour la soumission.

Riassunto

La tradizionale divisione tra scienze esatte (fisica e matematica) e scienze naturali (chimica e biologia) sta poco a poco scomparendo grazie al fatto che la complessita propria del mondo biologico e compresa e mappata attraverso un duplice attacco. Da un lato, degli esperimenti sempre piu' quantitativi permettono di studiare i dettagli della struttura atomica di molecole biologiche e di misurare con grande precisione le leggi d'interazione tra diverse molecole; dall'altro, l'introduzione massiva dell'informatica nella gestione e catalogazione della moltitudine di componenti molecolari che si trovano all'interno della cellula, permette d'ottenere una comprensione piu' profonda del complesso equilibrio dinamico che regola la rete delle interazioni tra diverse molecole.

Il lavoro descritto in questa tesi riguarda il primo fronte del campo di battaglia: lo sviluppo di nouve tecniche che permettono di compiere misure quantitative di parametri di interesse biologico.

Il lavoro ha compreso la progettazione, la costruzione e la validazione di tre diversi esperimenti volti allo studio della meccanica delle proteine o del DNA. Componenti chiave del microcosmo cellulare, DNA e proteine sono delle grandi macro-molecole che interagiscono continuamente, svolgendo la maggior parte dei compiti necessari alla sopravvivenza della cellula. La prima parte di questa tesi riassume le proprietà note di queste molecole e introduce le motivazioni alla base degli esperimenti svolti.

Le proteine svolgono il ruolo di catalizzatori delle reazioni chimiche all'interno della cellula e la loro conformazione tridimensionale determina la funzione specifica di ciascuna di esse. I legami tra proprietà strutturali e chimiche di una proteina sono un argomento ancora largamente inesplorato. La seconda parte della tesi descrive un esperimento basato sulle tecniche di microscopia di fluorescenza e di "molecola unica" volto a esplorare la dinamica delle fluttuazioni d'attività di un enzyma isolato.

Gli esperimenti descritti in questa parte non anno ancora dato i risultati sperati. Ciononostante, una parte delle considerazioni preliminari alla costruzione di questi esperimenti é stata utilizzata per scrivere l'articolo F. Mosconi et al.

"Some nonlinear challenges in biology", Nonlinearity 21 (2008) T131-T147.

Il DNA contiene l'informazione genetica necessaria alla cellula per svolgere tutte le sue funzioni. Quest'informazione dev'essere custodita, letta, copiata e riparata in diversi momenti del ciclo cellulare. L'importanza di una conoscenza dettagliata delle proprieta meccaniche del DNA è fondamentale se si vuol comprendere la sua interazione con le proteine. La terza parte di questa tesi descrive due diversi esperimenti basati sulla tecnica delle "pinzette magnetiche" volti a misurare alcune proprieta meccaniche del DNA che non sono ancora state completamente caratterizzate.

I due esperimenti presentati in quest'ultima parte han portato dei risultati interessanti. Si e' infatti ottenuta una nuova misura del parametro di interesse biologico C, il modulo di torsione del DNA, applicando un nuovo tipo di analisi a dei dati misurati con le "pinzette magnetiche" tradizionali. Inoltre, è stato sviluppato e validato un nuovo tipo di pinze magnetiche "dolci", che permettono di applicare una forza esterna contemporaneamente all'applicazione di un momento torcente. Con questo strumento si è stato possibile misurare la risposta del DNA a un momento torcente imposto dall'esterno. I risultati descritti in questa parte della tesi sono riassunti in due articoli che saranno presto inviati.

Acknowledgements

A PhD in experimental biophysics is a very intricate path, mine was, at least, and wish to express all my gratitude to all those who walked all or a part of this route with me.

First of all I would like to thank y two advisors David Bensimon and Amos Maritan, who share a common quality in the enthusiasm they put in their research work. I want to thank David for all the encouragement, the patience and the support he provided during these 3 years. I learnt a lot from his curiosity and his creativity in posing new questions to overcome obstacles. I want to thank him for welcoming me in the lab and for accepting a collaboration on a challenging project wih a university of a foreign country.

I want to thank Amos, for accepting to be my supervisor initially and for being enthusiast to let a student leave and go to study in a foreign laboratory. With him I began being fascinated by the biological world and I always admired his ability to gain simple insights into complex domains.

Vincent Croquette had no formal role in my education, yet his contribution is present in each accomplishment of this work. His extremely depp knowledge of physics, electronics, informatics and biology and his critical eye were a constat reference point all trhough this work.

Jean-Franǫis Allemand also contributed with stimulating discussion and most importantly with great culinary tips.

Nicolas Desprat taught me the fascination for microbiology, Thomas Julou, we shared frustrations and joys of our experimental days and I really appreciated that, Elise Praly, you brought the charm of a lady in an otherwise quite dull and dry room full of microscopes and Maria Manosa and Fang Yuan Ding you came to help her with an otherwise overwelming task.

I want expecially thank Maria for allowing me to use her picotwist extensively and not only at nights and weekends. Without your patience and kindness the main results of this thesis wouldn't have been possible.

I wish to thank Adrien Meglio for the many times you helped me out with french and Etienne Cavatore for all your jokes and music, Eric Perez for directing the laboratory, Nora Sadaui, Annie Ribaudeau and Marie Gefflot for the great professionality in dealing with the subtleties of french bureaucracy, Jose Silva da Quintas and Olivier Hombert for your magic craftsman's hands, Zaire and Fred for answering my geeky questions in informatics.

I wish to thank Keir Neuman for the incredible kindness and disponibility in front of the naive questions of a young student and for the attention and you put in formal presentations of colleagues at different conferences. I whish to thank Hiroaki Yokota and Omar Saleh for sharing the beginning of this journey and Timothee Lionnet and Giuseppe Lia for paving the way for younger generations to follow your path. Also, together with Gilles Charvin, you were excellent guides in another foreing land, providing help and assistence when needed.

I want to thank the equipe of Ludovic Jullien, for the close collaboration in the first project, Ludovic for the precious advise in photochemistry and wine tasting, Andrè for sharing the darkness in S13, Thomas LS for keeping up the idea of protein surface functionalisation, Thomas B for the vast amount of jazz music you made me discover, Sandrine for teaching me the rudiments of FCS, Isabelle fo her great kindness in all chemical matters.

I want to thank Marc Baaden and Sophie Sacquin-Mora for the precious time spent in discussing the details of a collaborative project that has yet to give the best fruits.

Richard Lavery and Jean-Franǫis Joanny for accepting to be rapporteurs, a very special thank in advance for accepting to read the work in full and judge it on the firm ground of science. I want to thank the equipe of statistical an biological physics in Padua for the inspiring discussions and for sharing the common hope of bringing Single molecule biophysics in the department. In particular I want to thank Attilio Stella for coordination of the doctoral school in Padua and for taking care of the subtle and sometims boring aspects of lgal mattersFlavio Seno and Antonio Trovato for teaching me a lot about proteins in the first year of my PhD, Gianluca Lattanzi and Luca Marsella for the interesting anecdotes about life in Trieste and finally Boris Marcone, Fulvio Baldovin and Francesco Zonta for keeping the atmosphere lively in the group. Sandro Azaele for the endless discussions about the foundations of QM and GR and for sharing a wonderful time in Padova, Marco Zanetti and Roberto Valandro for growing up physicists together.

I want to express all my thanks to those who helped me learning the details and tricks of fluorescence Single Molecule microscopy, in particular Shimon Weiss, claus Seidel, Oleg kritchewsky, Roman Shusterman,

Eyal Nir. A special thank for Terence Strick for the inspiring course he gave in Les Houches and for suggesting that contacting David was not out of reach for a young student: without such a simple advice this adventure would have never begun. Thanks to Simona Cocco, Didier Chatenay, Remy Monasson, Jean Dalibard for organizing the summer school in Les Houches in 2004, and all the people who attended the school for the invaluable experience we shared.

I want to thank Ala Trusina for being a wonderful guide in SF, Martin for the discussions of networks and ski, Franǫis Taddei and Ariel Lindner for the inspirng camp, Paolo Pierobon for attempting to carefully read the manuscript.

Last but not least I want to thank all those who shared good and bad times all though these four years, and without whom this would not have been possible.

Grazie ai miei genitori, per come mi avete cresciuto e per il vostro appoggio costante in questo lunghissimo progetto in terra straniera. Grazie ad Anna, sorella incredibile, geniale e profonda, le tue acute osservazioni mi stupiscono ogni volta. Grazie ad Anna, che hai condiviso con me gioe e dolori della prima parte di questo progetto. Grazie a Mariachiara, che ti conosco da sempre, grazie a Fra, per lo stupendo esempio di forza, graze a Marco per le chiacchiere e le canzoni e le storie di vita cosi' simili e cosi' diverse, Pigna per la colocation e per il verbois, il vertbois per Lotti, Dito, Daniela e Valentina, Alessio e Giovanna per le immancabili cene e pranzi domnicali, Ba per le cene e le telefonate, Zac per l'indispensabile computer portatile, Paolo e Mario per la Normandia e il sostegno nei momenti piu' duri, Jaki e Andre per le sciate organizzate in confcall, Angi per il sostegno quando e' servito, Lucia per il record di visite ormai imbattibile, Sil perche sei tu, Giullo amico profondo e sincero, Marco e Tati per le belle email, per il festival del cinema e per l'ADSL all'ultimo minuto, Cecilia, Matte, Chiara, Nano, Erica, Dave, Max, Cex, e tutti quelli che mi fan sentire a casa ogni volta che torno a Pd.

Je tiens aussi a remercier tous les amis qui m'ont accompagne en cette aventure francaise, si difficile au debut et si passionante apres. Merci Karine, Julien, Marie, Olga, Daniela, Francesca, Stefania pour la perseverance dans la maraude du mercredi. Merci a Beber, Joses, Alan, Marc, Fernand, Mitch, Momo, JeanLuc, Yacine, Riton, herve, Aude, Francis, Franǫis, Michel, Pierre, Andre, Gerhart, Alex et tous les autres S.A. du 5eme et 4eme que j'ai pu rencontrer et de qui j'ai pu apprendre beaucoup plus que donner.

Merci a Fab pour la coloc, la aison, la zyk et les discous du quotidien exceptionnel, merci a Ivan pour la recommendation VIP, a Essi pour la tendresse et le grand coeur, Tibault pour les chansons et le pont des arts pour l'ambience. Merci a Paname pour son atmosphere si misterieuse, froide au debut et charmante apres...

Part I

Introduction

Chapter 1

Fluctuations

Ma poichè quel che è distrutto, patisce; e quel che distrugge, non gode, e a poco andare è distrutto medesimamente; dimmi quello che nessun filosofo mi sa dire: a chi piace o a chi giova cotesta vita infelicissima dell'universo, conservata con danno e con morte di tutte le cose che lo compongono?

— G. Leopardi
Dialogo della Natura e di un Islandese (1824)

1.1 Introduction

All living organisms are made of cells. From simple unicellular organisms like bacteria, to complex organisms like ourselves, the cell is the basic subunit and understanding its functioning is key to understanding life. From a physical and chemical point of view, cells are nothing more than a droplet of water full of chemicals, bounded by a thin layer of fat. Despite the apparent simplicity, cells are capable of an incredible variety of complex tasks. They can move, divide, reproduce, differentiate, evolve and they can interact with the surrounding environment, processing information and synthesising chemical responses or mechanical reactions. How this is possible is one of the most fascinating subjects of biology.

Biology is essentially an experimental science, whose goal is to understand the engineering principles selected by evolution in the design of the great variety of organisms that are presently known. At all levels (from molecules to cells, from organisms to populations) the interactions between the various players are

strongly non-linear and usually saturate at sufficiently high concentrations (of molecules or cells). Molecular interactions occur at a microscopic level where Brownian fluctuations and thermally activated processes are often a key factor in the function of the system studied. Biological machines differ from man-made machines in that they are optimized to work in a very noisy environment. Because of this, bio-molecular processes are (extrinsically) stochastic and should be described within the framework of stochastic differential equations or via Fokker-Planck equations for the evolution of the probability density of given observables.

1.1.1 Extrinsic and intrinsic noise

Another fundamental element that characterises biological processes at all levels is the intrinsic discreteness of their players: the number of bio-molecules of a given type in a cell is often very small. For micron size bacteria such as E.coli, pH=7 may mean less than 100 H^+ ions per bacteria. These ions however diffuse very rapidly in and out of the cell, so that on the time scale of enzymatic reactions (ms) their average is well defined. It is not so for larger molecules which do not diffuse in and out of the cell, in particular DNA, mRNA and proteins. For example in bacteria the number of mRNA coding for a given protein is often less than 1 per generation [13] and less than 20 DNA-polymerase molecules are responsible for replication [19]! Besides this, large macromolecules can assume different conformations, and thus change biochemical properties. These two facts introduce an extra (intrinsic) level of stochasticity, and they may also qualitatively alter the behaviour of the system: as pointed out by S.Solomon [84], the outcome of differential equations can be radically different if the underlying variables are discrete or continuous.

1.1.2 Fluctuation dissipation theorem

Intuition that fluctuations and dissipation are intimately linked can be traced back to Maxwell, Clausius and Boltzmann [17]. They began understanding that relaxation phenomena, essentially described by diffusion coefficients and linked to dissipation, and transport coefficients such as viscosity, describing steady state non equilibrium flow, could be connected to the same molecular-scale properties (mean free path) in the framework of the kinetic theory of gases. The famous paper by Einstein in 1905 [34] and later on the theory of linear response and the mathematical formulation of the fluctuation-dissipation theorem (see [40] and references therein) showed that the near-equilibrium behaviour of a system could be predicted from the study of its reversible fluctuations in thermal equilibrium. Conversely, if both dissipation and equilibrium fluctuations can be experimentally measured, violation of the fluctuation-dissipation theorem can be observed if the linear-response hypothesis is not valid, as is often the case in biology. Microrhe-

ology on active gels [98] clearly showed that the fluctuation-dissipation theorem was violated when ATP, and thus chemical energy, was consumed. Biology thus provides a very interesting playground for the physicist interested in non-linear stochastic systems.

1.2 Observation of fluctuations: single molecule experiments

Biology and Biochemistry often use optical microscopes and ensemble measurements to characterize the systems studied. An optical microscope cannot resolve individual atoms and molecules, since they are smaller than the diffraction limit of visible light. X-ray crystallography or electron microscopy provide the necessary spatial resolution, at the price of observing the biological sample in state very different from the physiological state. For example, electron microscopy requires freezing or coating with metals, and it does not allow the observation of biologically active samples. On the other hand, fluorescence allows to reveal the presence of isolated molecules and follow their movement.

If we are interested in studying the stochastic fluctuations in a biological system and their connection to non-equilibrium response, we need to track fluctuations and be able to perturb the system in a controlled way, to observe its response. The opportunity to do this is offered by Single Molecule (SM) techniques [21]. Developed in the last twenty years, they allow the direct and indirect observation of biologically active single molecules in real time in their natural environment. Besides this, the miniaturisation of manipulation tools allows to interact directly with single biological molecules, through the application of a mechanical stress (force or torque).

1.2.1 Fluctuations, time average and ensemble average

Thermodynamics describes the average behaviour of a large ensemble of molecules by simple laws involving few macroscopic parameters. The behaviour of individual molecules is averaged over the population, which erases the presence of static disorder, i.e. the possibility that not all the molecules in the population behave in the same way. This is particularly true for large macromolecules, such as enzymes, where the chemical behaviour is intimately linked to the three-dimensional structure. For example, consider an enzyme with more than one biologically active configuration, each with its own catalytic rate. A bulk measurement of catalytic activity on a mixed population will yield an average activity and will not contain any information about the presence of different configurations. The possibility to measure the activity of a single enzyme and then to repeat the experimental over a large number of identical molecules, provides more information than the

ensemble measurement because the whole distribution of states, and not just the average, can be experimentally reconstructed.

Bulk measurements also imply averaging over time and lack of synchronicity. In this case, dynamic changes of behaviour (dynamic disorder) are hidden to the observer. Coming back to the enzyme example, the activity of a single enzyme could change over time, as a consequence of conformational fluctuations and interaction with the surrounding environment. Changes with time would be averaged, unless a high enough time resolution was obtained for the experimental observation. Even in this case, synchronisation across the population would be required to observe dynamic disorder in a bulk experiment.

SM techniques give access to realtime observation of individual molecules and they provide more information than bulk experiments. They open up the possibility of completely new observations, because of their power to resolve fluctuations in individual molecules.

1.2.2 Pulling, twisting and tracking

SM techniques [99, 80] allow both the observation and the interaction in real-time of isolated molecules. This is usually achieved through the coupling of a probe to the molecule. The probe is visualised in a microscope, and it reports about the behaviour of the observed molecule. For example, quantum dots are fluorescent nano-crystals about 20 nm in size. These fluorescent probes can be coupled to single enzymes whose motion can then be followed with a standard fluorescence microscope [100, 28].

Manipulation experiments require a microscopic handle to apply a controlled stress to an individual molecule. For example, single DNA manipulation is achieved by attaching the extremities of the DNA filament to micron-size beads which will serve both as position probes and as handles. The interest of micro-manipulation of biological molecules is evident if one considers that many proteins are motors that produce mechanical work. The whole process of DNA replication, for example, is achieved by proteins that move directionally along the DNA chain. In this context, notions like speed, force, torque, mechanical work and efficiency become the relevant quantities to describe the behaviour of such motors. Also, the static mechanical properties of biological molecules, such as resistance to a pulling force, can be studied as a preliminary step in understanding their interaction with one another.

In this thesis both real time tracking and manipulation techniques have been implemented in new experiments targeted at answering the problems proposed in section 1.3.

1.3 Open problems

1.3.1 Protein structure and function

Protein structures are thought to be determined by the minimisation of free energy in the landscape of all the possible configurations of the poly-peptidic chain. In the classical picture [3, 2] the energy landscape is smooth and it has a unique minimum, the native state, which is both stable against thermal fluctuations and kinetically accessible to allow for a fast folding. Furthermore, this state is supposed to have a well defined kinetic rate, its interaction with a substrate described by Michaelis-Menten (MM) dynamics. Soon after its formulation, the classical picture was questioned by experiments on myoglobin conducted by Frauenfelder and colleagues [6]. They suggested [39] that the protein energy landscape could resemble that of glassy materials, presenting many local minima separated by energy barriers of different heights. The local minima would correspond to different three-dimensional conformations assumed by the protein. With the development of high resolution experimental techniques like X-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance, evidence accumulated to support the idea that the same protein could assume different conformations. Only quite recently, though, the dynamics of hopping between states started to be assessed (see [50] for a recent account of progress on the subject).

Recent single molecule studies, in fact, opened a window on a previously unexplored area allowing to follow the fluctuations in catalytic activity of an isolated protein for long time. In such experiments [35, 31, 37] the catalytic rate of a single enzyme was monitored with the use of appropriate substrates which turned into fluorescent products upon catalysis by the enzyme. Both static and dynamic disorder were observed when repeating the experiment on enzymes of identical sequence, and the cause for these has yet to be clearly identified. It is not clear whether activity fluctuations are related to structural fluctuation and how the different time scales that appear in the time correlation of activity are to be interpreted. One way to investigate these questions is to study the dynamics of the protein structural fluctuations (using techniques such as single molecule fluorescence energy transfer [28]), in order to distinguish between thermal fluctuations about a well defined state and fluctuations exhibiting long-term correlation effects. In a recent paper [47], the authors used this technique to asses the dynamics of conformational changes of the enzyme Adenylate Kinase and the influence of substrate interaction. Also, recent numerical simulations [51, 52] seem to support the picture of a large range of timescales in the dynamics of enzyme conformational changes.

1.3.2 Measurement of torque in biological samples

The SM techniques developed in the past 15 years focused on measuring mechanical forces in biological samples. Force spectroscopy of biological molecules has now become a standard technique and the mechanical response of DNA, proteins and even cells has been extensively investigated [20]. This approach provided an invaluable amount of information on how biological molecules react to external forces, and gave insights on molecular devices that had not been studied before. Inside cells, motor proteins exist that do mechanical work to perform their task and a lot of effort has been addressed to studying their properties in detail.

Despite the many successes of force spectroscopy, the response to mechanical torque has been less investigated due to the lack of techniques that could measure it. Nevertheless, even at the level of single macromolecules, like DNA and proteins, torque is thought to play an important role in different biologically relevant processes. For example, proteins apply torques on DNA during cellular processes such as replication and transcription [103], DNA compaction [70], chromosomal condensation [56]. SM micromanipulation techniques have been successfully employed to study the behaviour of torsionally constrained DNA in great detail [88], [89], [59] and to investigate its interaction with torque sensitive enzymes such as topoisomerases [90] and helicases [62].

1.4 Proteins

1.4.1 Function of Proteins

Proteins are large macromolecules formed by a long polymeric chain, whose subunits are called amino-acids. Proteins are synthesised inside the cell to perform a particular function, and are destroyed when they are unnecessary or misfunctioning. Many proteins are in fact enzymes, that catalyse a chemical reaction allowing it to happen in conditions where it would normally take exceedingly long times. Enzymes are essential for the cell metabolism as they allow for the possibility to tune catalysis to the needs of the cell. These nano-machines intervene at each step of a chemical reaction so that it does not happen in uncontrolled way.

Proteins are also used for structural purposes, to give particular mechanical properties to cells or tissues. Example of these proteins include keratin in hair cells, actin in muscle cells or micro-tubules which form the internal scaffolding of a cell. Proteins are also used for signalling and recognition both inside and outside the cell, and in this case they are often found to be part of a large network of interaction. An external stimulus can trigger the production of a protein which is a the top of a cascade of signals which will eventually trigger a broader response from the cell. Examples of these include antibodies in the immune response or signalling cascades that trigger electrical signals in neurons.

Finally there are proteins that can truly be considered as nano-motors converting chemical energy into useful mechanical work. These are often involved in complex tasks requiring mechanical motion, like active transport of vesicles or DNA replication.

1.4.2 Protein constituents

Proteins are built from 20 standard amino-acids which, like nucleotides, share common features. Their chemical structure includes a carbon atom, called α-carbon, an amino group and a carboxyl group. A peptide bond is formed between amino-acids: a double resonant covalent bond, and it is very stable. The peptide bond confers some rigidity to the amino acid, so that subsequent α-carbons are generally co-planar. The other dihedral angles formed by the α-carbons with the side chains and with the β-carbon (which originally belonged to the carboxyl group) can freely rotate, giving the chain the possibility to fold into complex structures.

Amino-acids differ in the chemical groups attached to the α-carbon, or side chains, which confer to each of them its specific chemical characteristics. Amino-acids can be classified in two groups, hydrophobic and hydrophilic, which differ in their preference for water. This broad division in two classes becomes important when the three dimensional structure of a protein is considered.

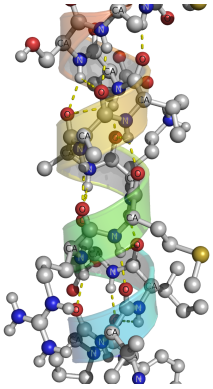
1.4.3 Protein structure

Each protein sequence assumes a particular three-dimensional shape, which determines its catalytic properties and thus its biological function. Short chains fold into well defined motifs, the most common of which are α-helices and β-sheets (see figure 1.1a). These motifs are ubiquitous and can be considered to be the building blocks with which proteins are built. These motifs are stabilised by the formation of hydrogen bonds between adjacent residues, but hydrogen bonding is not required for the formation of such patterns.

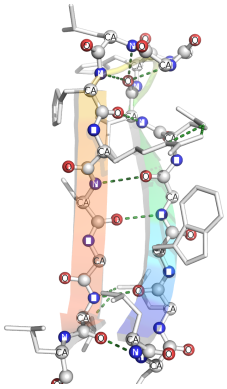
Secondary motifs, as they are called, are then ordered into a three-dimensional configuration that is called the tertiary structure (see figure 1.1c). This is often of globular shape, although more complicated forms exists. Hydrophobic amino-acids are found in the inner part of the globule, the core, while hydrophilic residues are prevalently found on the surface. The tertiary structure is stabilised by the formation of bonds between non-consecutive but adjacent residues. These can be ionic bonds but also covalent bonds, such as cysteine disulphide bonds. Also, proteins can be modified and enriched after folding with the addition of other molecules (sugars, metal ions, co-factors) or by the modification of some side chain. These are called post-translational modifications. The tertiary structure is also called the native state, because it is assumed to be the state in which

proteins are biologically active.

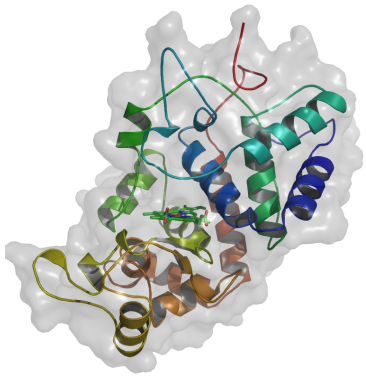
In some cases also a quaternary structure exists, obtained by the union, or docking, of different protein subunits, shaped to match one with the other, like Lego building blocks. The quaternary structure is stabilised through electrostatic interaction, facilitated by the presence of charged amino acids on the exterior of each subunit and by the complementarity of the surface shape.



(a) Alpha helix: the backbone forms a roughly helical shape and hydrogen bonds are established between adjacent turns



(b) Beta sheet: the backbone forms a zig-zag pattern, and hydrogen bonds are formed between a sheet and the next.



(c) Tertiary structure of a protein: the backbone of the polypeptidic chain is shown as a cartoon , highlighting the secondary structures (mainly helices in this case). The semi-transparent cloud represents the outer surface of the protein when side chains are taken into account.

Figure 1.1: Secondary and tertiary structures of a protein

As we stated in section 1.4.1, many proteins are enzymes, i.e. they catalyse a chemical reaction, lowering the energy barrier for activation. The three-dimensional structure of proteins is precisely shaped in order to be particularly adapted to this task. Catalysis is obtained by the interaction of the molecule with a particular region of the protein, called the active site (see figure 1.2. This region has often the shape of a pocket, into which the reactants fit nicely, like a key in a keyhole. Once bound in the active site, reactants find themselves very close in space. This reduces the activation energy for the chemical reaction to happen. The active site is thus the very important region that uniquely defines the chemical role of an enzyme.

1.4.4 Protein folding and native state

The process of protein folding[8] from the primary amino-acid sequence to the 3D native state is not fully understood. Already in 1968 C. Levinthal [60] pointed out that the number of degrees of freedom present in a protein chain is such

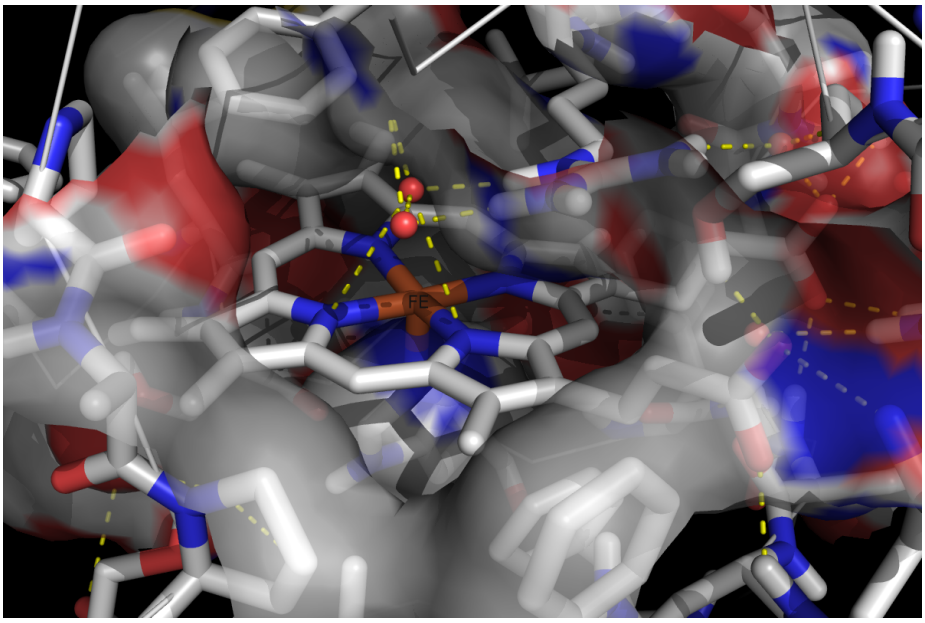


Figure 1.2: Close-up view of the heme co-factor contained in the active site pocket of HRP: the tertiary structure forms a channel that leads directly to the heme allowing subtrates to easily access it.

that, had the folding process been an extensive random search of all possible configurations, it would last longer than the age of the universe to fold a single protein. How proteins fold, and how they manage to find the correct shape in matters of micro-seconds is one of the most debated problems of the last 30 years. It has been suggested that secondary motifs are formed first and then tertiary structure is composed. Some studies suggest that the protein chain quickly forms a pre-fold which is topologically very similar to the folded structure, and then it adjusts the details. A particular proposal is found in [53] where the authors suggest that the energy landscape is pre-sculpted by geometrical considerations and all the protein sequence has to do is chose in a menu of pre-folds which is much smaller than the ensemble of all possible combinations given by a random chain.

Whatever the precise mechanism of folding is, the traditional assumption is that this process drives the protein to a unique stable native state. This conjecture, is traditionally taken as a postulate and goes under the name of Anfinsen's Dogma [3, 2]. It implies that the free energy minimum is unique, stable and kinetically accessible. Uniqueness implies that there are no other stable ground states in the free energy landscape explored by the protein during folding. Stability means that thermal fluctuations cannot unfold the protein, so that it can continue to perform its biological function even in the noisy cell environment. Accessibility is necessary to explain the rapidity of folding, in fact

the native configuration is assumed to be easily reachable from the unfolded configuration

This rather static picture is at odd with the immense number of degrees of freedom of the poly-peptide chain, and it is precisely this oddness that inspired some of the considerations that lead to the experiment detailed in chapter 3.

1.5 DNA

This section and the following one are intended to serve as reference for the biological concepts used in this manuscript. The main reference for all the information reported in the following sections is [1].

1.5.1 Nucleic Acids

The information needed by the cell to perform all its tasks is stored in nucleic acids. Nucleic acids are long polymers formed by a sequence of repeated subunits, whose pattern forms the basic code for information storage and transfer in the cellular world. The repeated subunits are called nucleotides: organic molecules composed of a sugar, a phosphate group and a nucleotide base. The sugar moleucule, ribose or deoxyribose (for RNA and DNA respectively) and the phosphate groups are united by a phosphodiester bond and form the backbone chain. Nucleotide bases exist in four varieties called, Adenine (A), Guanine (G), Cytosine (C), Thymine (T) (or Uracil (U) in RNA): four "chemical letters" that constitute the alphabet of our genetic code.

DNA is formed by two complementary strands wrapped around each other in helical fashion. Since its discovery by Watson and Crick in 1953, its chemical composition and its physical properties have been extensively studied and characterised. This macro-molecule is very stable to the perturbations of the environment, and it is not astonishing that it has been selected for the purpose of information storage.

RNA is commonly found as a single chain poly-nucleotide. Its backbone is very flexible and it can bend and wrap in very intricate manners. RNA is assumed to have appeared first in the history of life because it can both store information in its sequence of bases and it can perform catalysis (the RNA in ribosomes is the catalyser of proteins syntesis). This dual property provides a feedback mechanism to drive evolution, even in the absence of cells. In principle, an RNA world is possible in which RNA molecules fold to a form which catalyses the formation of new RNA molecules with the same sequence and so on.

In most organisms RNA performs intermediate tasks between the information storage task (which is commonly assigned to DNA) and the catalysis task (which is commonly assigned to proteins). RNA is used as messenger that carries the information to code for the generation of a protein. RNA is also used as

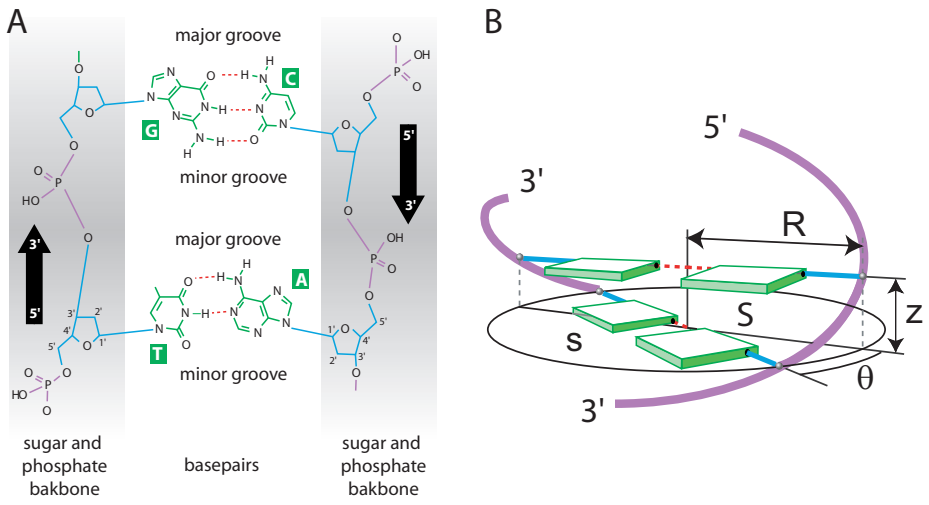


Figure 1.3: A: Flattened primary structure of DNA: DNA is composed by two antiparallel strands. Each strand is made of a chain of nucleotides linked by phosphodiester bonds. Each nucleotide is formed by a base (green), a sugar (blue), and a phosphate (purple). Bases are paired by hydrogen bonds (red dotted lines). The geometry of some bonds was modified for reasons of clarity. The strand sequence is conventionally read from the 5' to the 3' end. B: The two backbones made of sugars and phosphates wrap around each other forming a double helix. (Figure adapted from [61])

transporter for protein subunits (amino acids) during protein synthesis and it is a major constituents of ribosomes, which are the machines that carry out protein synthesis.

1.5.2 Function of DNA

The main function of DNA is to store the information necessary for all the tasks performed by a cell. Its enhanced stability with respect to RNA offers a selective advantage that most likely tilted the choice of evolution in its favour. In fact, the absence of a single hydroxyl group in the backbone sugars confers additional rigidity to the its chain. Furthermore, DNA is normally found in a double helical pair of complementary strands. This configuration is more reliable because random damages to one strand can be repaired using the complementary strand as a template. Also it provides a simple mean of duplicating the whole genetic repository, as required by cellular division (mitosis). During replication two new double helices are synthesised starting from each separated single strand.

Despite the huge effort of genome sequencing projects in the past 30 years, our comprehension of the genetic code is not complete. Coding sequences and regulatory sequences total 10% of the human genome. The rest is formed by repeated sequences, transposons, and sequences not yet understood. Table 1.5.2

summarises the information content of the human genome. The parallel with informatics can be made considering that four bases can be described with 2 bits of information, and thus a byte (= 8bits) can be used to uniquely identify a sequence of four bases.

Genome	3.2×10^9 bases	760 Mbytes
Unique sequences	10^9bases	240 Mbytes
Coding + regulatory	320×10^6bases	76 Mbytes
Coding sequences	160×10^6bases	38 Mbytes

Table 1.1: Human genome

1.5.3 Constituents of DNA

The two strands forming DNA are united by weak selective bonds, that form between conjugate pairs of bases: A can only form a pair with T and C only with G. These kind of pairings are called Watson-Crick pairings. They imply the formation of two (A-T) or three (C-G) hydrogen bonds between the outer edges of the base-pairs. When they discovered this mechanism of selective base pairing, Watson and Crick immediately understood that it provided a perfect way to preserve the genetic information from external damage and to copy it during cell division. In fact, although the two strands do not contain exactly the same sequence (they are complementary), they do contain exactly the same information. This enhances the robustness of the storage process: information transfer in a noisy environment cannot rely on coherence (contrary to analog communication) and thus content of a message has to be decoupled from the sources of noise, very much like digital signalling. DNA is perfectly apt to achieve this task. In fact, the relevant energy scale for biology is set by the thermal energy $E = k_B T = 4.1 \times 10^{-21} J$. The typical energy of a covalent bond ranges between $150 k_B T$ and $400 k_B T$. Phosphodiester bonds formed between phosphate and sugars are thus very stable compared with the thermal energy. On the other hand the energy of a single hydrogen bond is around $3 - 5 k_B T$ and thus Watson Crick base pairing is not energetically sufficient to guarantee the thermal stability of the double stranded DNA molecule. Stability is achieved through the cooperative action of other bonds: hydrogen bonds form between a base and the adjacent ones along the same strand (stacking interaction) conferring enhanced stability to each strand. Furthermore, the helical structure contributes to mechanically prevent the conjugated bases from separating and hydrophobic effect also conspire to keep the nucleotide bases hidden from solvent in the centre of the chain.

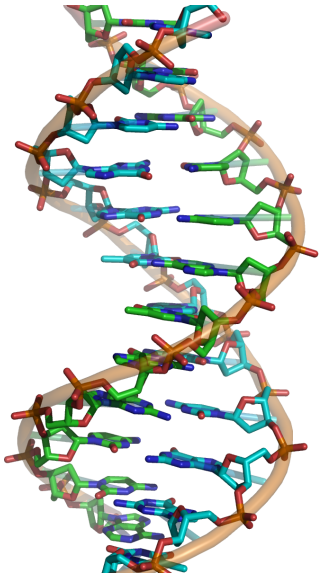


Figure 1.4: B DNA

1.5.4 DNA structure

The structure described in detail in the previous section is that of B-DNA. This is the form assumed by DNA in water solution under physiological conditions and it is also the form adopted by DNA inside cells in normal conditions (except during replication and division). In this form DNA is a long polymer of 2nm diameter, completing a helical turn every 10.4 base-pairs corresponding to a helical pitch of 3.6 nm and to 3.4$\mathring{A}$ per base-pair. In what follows we will often express the length of a DNA molecule in number of base pairs or more frequently in kbp. The conversion from kbp to metres is obtained multiplying the length of DNA in kbp times the height of a single base (0.34nm).

B-DNA is not the only form found in nature. Under particular conditions DNA can assume different forms, which have also been observed in X ray crystallography. These forms can depend on external conditions, such as hydration or salinity. They can be induced by the cell during particular phases of the cell cycle (for example during division DNA is compacted into chromosomes) or can be locally favoured by base pairs sequence. The last point, which still has to be investigated in greater detail, could provide a way to recognise particular sequence motifs from DNA structure, without actually reading the sequence, and could be a mechanism to help proteins to find their target sequence.

In condition of low humidity, DNA is found in A-DNA form. This is still a right double helix, but its diameter is larger and basepairs are tilted with respect to the B form. The Z form is found in high salt conditions or in presence of alcohol. It is a left handed helix in which phosphate backbones are turned in

the opposite sense. Finally P-DNA exists, in which the two strands are close to the centre of the helix, base interactions are disrupted and bases are exposed to solvent. This form has been found to be adopted in single molecule experiments when positive torsion is added together with a strong pulling force [63].

1.5.5 Polymer picture

All of the experiments detailed in this thesis will consider a coarse grained description of DNA. The mechanical stability of DNA in its normal configuration confer it a particular rigidity, which allows to coarse grain the base-pair structure to a large extent. In fact, due to its rigidity, the relevant length scale to describe the behaviour of DNA in solution at room temperature is set by the bending persistence length $A = 50nm$ or $150bp$. This is the relevant length scale to describe bending fluctuations of DNA. DNA in solution behaves like a semi-flexible polymer: thermal fluctuations of its structure will be negligible at length-scales lower than the persistence length, while they will become important to describe DNA conformation at larger length-scale.

Polymer physics predicts an order of magnitude for the radius of gyration occupied by a random polymer: $R_G = \sqrt{2LA}$, where L is the contour length of the polymer and A the persistence length. This means, for example, that the genome of the E. Coli bacteria ($5Mb = 1.7mm$) would occupy a sphere of radius $R_G \approx 13\mu m$, while the typical size of E. Coli is $1\mu m$!

Part II

Fluorescence Experiments on Proteins

This part is dedicated to the experiments conducted in the years 2005 and 2006, which focused on the relationship between protein structure and function. As introduced in section 1.3.1, the static picture of proteins has been increasingly challenged both a theoretical and an experimental point of view. Following the first pioneristic experiments conducted by Lu et al. [64] and by Edman et al. [31] on HRP, the idea of repeating the experimence in the presence of an external force, became increasingly tempting.

Adding an external tunable force to the experimental observation of protein activity in real time seemed an important step in the exploration of the links between structure deformations and changes in catalitic activity. The project was conducted in collaboration with two computational groups (IBPC, Paris and University of Padova, Italy) who were interested in formulating models for the description of protein structural changes in the presence of force.

This ambitious project did not go to a good end because of many reasons, probably the most importan being the little previous expermental experience with the subtleties of fluorescence microscopy of our group. Nevertheless, the amount of knowledge accumulated is worth being written for the future students not to repeat the same mistakes.

Chapter 2 is dedicated to the description of the problem, previous results and framing of the proposed experiments. Chapter 3 will review standard techniques in flurescence single molecule microscopy and chapter 4 will be dedicated to the description of the setups developed to address the problem.

IOP PUBLISHING NONLINEARITY

Nonlinearity **21** (2008) T131–T147 doi:10.1088/0951-7715/21/8/T03

OPEN PROBLEM

Some nonlinear challenges in biology

Francesco Mosconi, Thomas Julou, Nicolas Desprat, Deepak Kumar Sinha, Jean-François Allemand, Vincent Croquette and David Bensimon

LPS, ENS, UMR 8550 CNRS, 24 rue Lhomond, 75231 Paris Cedex 05, France

E-mail: David.Bensimon@lps.ens.fr

Received 21 April 2008
Published 10 July 2008
Online at stacks.iop.org/Non/21/T131

Abstract

Driven by a deluge of data, biology is undergoing a transition to a more quantitative science. Making sense of the data, building new models, asking the right questions and designing smart experiments to answer them are becoming ever more relevant. In this endeavour, nonlinear approaches can play a fundamental role. The biochemical reactions that underlie life are very often nonlinear. The functional features exhibited by biological systems at all levels (from the activity of an enzyme to the organization of a colony of ants, via the development of an organism or a functional module like the one responsible for chemotaxis in bacteria) are dynamically robust. They are often unaffected by order of magnitude variations in the dynamical parameters, in the number or concentrations of actors (molecules, cells, organisms) or external inputs (food, temperature, pH, etc). This type of structural robustness is also a common feature of nonlinear systems, exemplified by the fundamental role played by dynamical fixed points and attractors and by the use of generic equations (logistic map, Fisher–Kolmogorov equation, the Stefan problem, etc.) in the study of a plethora of nonlinear phenomena. However, biological systems differ from these examples in two important ways: the intrinsic stochasticity arising from the often very small number of actors and the role played by evolution. On an evolutionary time scale, nothing in biology is frozen. The systems observed today have evolved from solutions adopted in the past and they will have to adapt in response to future conditions. The evolvability of biological system uniquely characterizes them and is central to biology. As the great biologist T Dobzhansky once wrote: 'nothing in biology makes sense except in the light of evolution'.

Mathematics Subject Classification: 01-02, 92-02, 92C15, 92C45, 92D15, 92D25, 92D40

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Biology is essentially an experimental science, whose goal is to understand the engineering principles selected by evolution in the design of the great variety of organisms that are presently known. At all levels (from molecules to cells, from organisms to populations) the interactions between the various players are strongly nonlinear and usually saturate at sufficiently high concentrations (of molecules or cells). Molecular interactions occur at a microscopic level where Brownian motion and thermally activated processes are often a key factor in the function of the system studied (e.g. molecular motors). As a result all bio-molecular processes are (extrinsically) stochastic and should ideally be described within the framework of stochastic differential equations or via Fokker–Planck equations for the evolution of the probability density of given observables. Another fundamental element that characterizes biological processes at all levels is the intrinsic discreteness of its players: the number of bio-molecules of a given type in a cell is often very small. For micrometre size bacteria such as *E. coli*, pH = 7 may mean less than 100 H^+ ions per bacteria. These ions, however, diffuse very rapidly in and out of the cell, so that on the time scale of enzymatic reactions (ms) their average is well-defined. It is not so for larger molecules which do not diffuse in and out of the cell, in particular DNA, mRNA and enzymes. For example in bacteria the number of mRNA coding for a given protein is often less than 1 per generation [1] and less than 20 DNA-polymerase molecules are responsible for replication [2]! Besides the fact that this small number of molecules introduces an extra (intrinsic) level of stochasticity it may also qualitatively alter the behaviour of the system: as pointed out by Solomon [3], the outcome of differential equations can be radically different if the underlying variables are discrete or continuous.

A fundamental and still largely unsolved problem is how biological systems can perform their function effectively under such constraints of extrinsic and intrinsic stochasticity. Here the contribution of nonlinear approaches could be fundamental since bifurcation theory can be used to predict the robust (topological) behaviour of complex dynamical systems for which we often do not know many important parameters (rate and affinity constants, cooperativities, saturation concentrations, etc). This weakness, however, can be turned into a strength. Indeed since these rates can be altered—by environmental changes (temperature, nutrients, pH, salt, etc), by different levels of protein expressions within the cell (a phenomenon known as phenotypic variability) and by genetic mutations—often without affecting the behaviour of the organism, a strong requirement of any modelling of biological systems is that it should be robust under major alterations of many of the parameters entering the model. When the actors of some regulation and metabolic networks and their topology are known (or reasonably guessed), a quite common trend among theoretical biologists is to write the differential equations describing their interactions and simulate them by making educated guesses of the many dozens of parameters that often enter such models (see for example [4]). Due to their large number of parameters and the often semi-quantitative predictions they make, it may, however, be difficult to falsify them (the usually accepted criterion for a good theory). Moreover, even if such models are robust to changes in their parameters and can be falsified, it is not clear that one can learn much from such detailed modelling. Beyond the 'forest' of equations are there some simple underlying principles? Extracting the 'slow' or relevant modes of such models might be more illuminating. How can the system be reduced to the dynamics of these modes only?

An answer to these questions would be particularly welcome if we are to understand the evolution of biological systems. Indeed the trade-off between robustness and evolvability (the possibility to adapt to new conditions) is a central theme of research in biology at all levels of inquiry: how is evolution constrained by the choices already made? What are its

rules, namely, can we predict how a biological/ecological system is going to evolve under a given selection pressure? The interesting case here is presented when different responses are possible: which one is selected? With which probability? Can we predict these probabilities? Finally, a third fundamental issue has to do with the mechanisms of development. These are remarkably robust to many variations (temperature, nutrients, fluctuations in gene expression and in rate constants, etc) yet are capable of evolving. How this robustness of the developmental pathways is achieved and how that robustness is restricting the evolutionary options presented to the organism are fascinating and unresolved issues where innovative nonlinear modelling and experimental approaches are required.

In the following we will sketch how the various issues of modelling, stochasticity (intrinsic and extrinsic), robustness and evolvability have been addressed in a variety of contexts: protein structure/function (section 2), gene expression networks (section 3), morphogenetic pathways (section 4) and ecological systems (section 5). The interested reader is referred to the appendices, where a detailed example for each topic is illustrated (appendix A is an exception as it contains a very short summary of relevant concepts in molecular biology). Our purpose is of course not to review these subjects (for that see the excellent book by Gerhart and Kirshner [5]), but rather to 'whet the appetite' of our readers and convince them that there are many unsolved fundamental problems in biology where tools from nonlinear system dynamics could be applied with great success.

2. Proteins, biochemistry and evolution

Protein structures and functions have been extensively studied from a variety of perspectives. Although physics and chemistry became quantitative sciences earlier than biology, at the level of small biological molecules, such as proteins, there are still many unsolved issues, involving stochasticity and nonlinearity, which demand further investigation both theoretically and experimentally. For example, the definition of the ground (or native) state of a protein is becoming more and more puzzling as new experiments explore it in detail. Not only is the prediction of the protein structure from its primary (i.e., amino-acid) sequence still unsolved (for a very recent account of structure prediction methods see [6]) but the mere question of the existence of a well-defined ground state is not clear.

Protein structures are thought to be determined by the minimization of free energy in the landscape of all possible configurations of the poly-peptidic chain. In the classical picture [7,8] the energy landscape is smooth and it has a unique minimum, the native state, which is both stable against thermal fluctuations and kinetically accessible to allow for a fast folding. Furthermore, this state is supposed to have a well-defined kinetic rate, its interaction with a substrate described by Michaelis–Menten (MM) dynamics (see appendix B). Soon after its formulation, the classical picture was questioned by experiments on myoglobin conducted by Frauenfelder and colleagues [9]. They suggested [10] that the protein energy landscape could resemble that of glassy materials, presenting many local minima separated by energy barriers of different heights. The local minima would correspond to different three-dimensional conformations assumed by the protein. With the development of high resolution experimental techniques such as x-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance, evidence accumulated to support the idea that the same protein could assume different conformations. Only quite recently, though, the dynamics of hopping between states started to be assessed (see [11] for a recent account of progress on the subject).

Recent single molecule studies, in fact, opened a window on a previously unexplored area allowing us to follow the fluctuations in catalytic activity of an isolated protein for long time.

In such experiments [12–14] the catalytic rate of a single enzyme was monitored with the use of appropriate substrates which turned into fluorescent products upon catalysis by the enzyme. This method allowed the measurement of the time interval between successive reactions at high substrate concentrations (i.e., when diffusion is not the limiting factor). It appeared that such a time was not exponentially distributed (as one would expect if the enzyme had a well-defined conformation and kinetic rate). Rather its distribution could be fitted by a multi-exponential law, supporting the existence of a continuum of conformers, each characterized by different reaction kinetics (see appendix B). The time correlation of the activity of a single enzyme suggested that it spent up to seconds in each of these states, giving rise to memory effects in the enzyme conformational trajectory.

How many and how different are these conformers? Do they reflect thermal hopping between enzymatic structures of similar energy but different functionality or do they arise from the metabolic activity of the enzyme, kicked off from its native state by the reaction it catalyses and then moving chaotically on this manifold of excited states? One way to answer these questions is to study the dynamics of the protein structural fluctuations (using techniques such as single molecule fluorescence energy transfer [15]), in order to distinguish between thermal fluctuations about a well-defined state and fluctuations exhibiting long-term correlation effects. In a recent paper [16], the authors used this technique to assess the dynamics of conformational changes of the enzyme Adenylate Kinase and the influence of substrate interaction. Also, recent numerical simulations [17, 18] seem to support the picture of a large range of time scales in the dynamics of enzyme conformational changes.

The traditional picture of a homogeneous population of enzymes is no longer consistent with the amount of experimental and numerical data being gathered, and there is need for a theory taking into account the presence of a great number of different conformers. This effect introduces stochasticity (also called 'static disorder') across a population of enzymes. Similarly the possibility for a single enzyme to hop between different conformations introduces stochasticity (dubbed 'dynamic disorder') at the level of the activity of a single enzyme. These effects have to be taken into account if a satisfactory picture of enzyme catalysis has to be formulated. Yet, for the vast majority of known enzymes, and surely for those studied in the works mentioned above, the Michaelis–Menten kinetics seems to hold unperturbed, even for fluctuating single enzymes. A possible explanation was proposed in [19,20], yet one might expect departures from Michaelis–Menten-like kinetics in more general cases.

In the broader context of protein evolution, the previous results might also be related to the independent observation that proteins may evolve different functionalities by so called 'promiscuous activity' [21]. Since mutations are usually deleterious to the function of a given protein, it is important to understand how an enzyme with a specific function can mutate and evolve a new one. One known possibility for the gene coding is for the enzyme to be duplicated with evolution acting on one copy only, the other ensuring continuity of the vital activity; another is for the enzyme to have its main catalytic activity at one site and a weaker unrelated ('promiscuous') activity at another more distant one. This 'promiscuous' activity could then be acted upon by evolution without much interference with the main function of the protein (the 'promiscuous' site plays in real space the role of the duplicated gene). How is that picture altered by the existence of not one but many active states of the enzyme? Is the plasticity suggested in the previous paragraph useful for the evolution of protein function? Is it evolutionary easier to tailor these states to the needs of the organism or do they in contrast increase the constraints on the evolvability of the protein? These are fascinating issues that will require both sophisticated single molecule experiments on evolving enzymes and new theoretical approaches to these systems.

3. Gene expression and regulation, system biology

To understand the way cells operate it is not sufficient to study the structure and function of their molecular components. One has to grasp how the various molecules interact together in regulation and metabolic pathways to define the functional state of the cell (known as its phenotype, see appendix A). Although these pathways are interconnected, it has been proposed that just like man-made machines, they are composed of functional modules [22]. A functional module is a network of proteins (and possibly other molecules such as RNA, lipids and sugars) that are capable of performing a certain task: metabolizing glucose, making a flagellum, directing a bacterium to a source of food (chemotaxis), etc. A major goal of systems biology [23] is to understand the design principles of these modules. To that purpose one can adopt two approaches: analyse native modules or design modules that mimic natural ones.

For example an abrupt transition between a low and a high gene expression state can be ascribed to a bistable switch. Native switches are involved during the development of a multicellular organism [24] where the fate of cells results from complex inter-cellular interactions (see below). They are also observed in bacteria and single cell organisms (e.g. yeast) where the different phenotypes may result from multi-stable fixed points of the functional module dynamics (for a review see [25, 26]). For instance, bistability in gene expression was shown to be responsible for transitions in cell state such as cell competence [27, 28] (see appendix C) or sporulation [29] in *Bacillus subtilis* and sexual identity in yeast [30, 31].

Such a genetic switch was mimicked by introducing an artificial gene network decoupled from endogenous signalling pathways in bacteria [32, 33] and in yeast [34]. An important ingredient for these experimental models to function as a bistable switch seems to be the degree of cooperativity of the transcription factors regulating the genes of the network, namely, the requirement of strong nonlinearity in gene repression (or activation, see appendix C). One essential feature of a fate transition driven by a bistable switch is that at the transition point, the stochastic variation in the number of proteins in each bacterium allows the coexistence of two populations (phenotypes) each characterized by one of the fixed points. Consequently, while cells may be genetically identical (share the same genotype) they may exhibit a different phenotype.

The possibility of engineering desired biochemical pathways and networks in cells has given rise to a sub-domain of systems biology, known as synthetic biology. Inspired by the success of electrical solid state devices (transistors, micro-processors, memories, etc) researchers in that field are trying to develop functional modules (biochemical networks such as the just mentioned bistable switch) that perform logical functions [35], mimic cell–cell communication system [36] or exhibit oscillations [37]. Besides designing artificial networks in cells, there is also great effort to introduce metabolic pathways from one organism into another [38] (for example to graft the cellulose metabolism of the bacteria from termite guts into yeast to produce bio-ethanol). The experimental implementation of such networks is often not straightforward and a great deal is learnt on their regulation and their coupling to the general metabolism of the cell when attempting these grafts. The modelling of these networks is therefore both instructive and useful in helping to identify the key parameters that must be adjusted to get the sought after behaviour [39–41]. In particular, for oscillatory modules, the robustness of the desired response to fluctuations in the network's components and variation in its kinetic parameters is a key to its efficient implementation.

Even if a given function is successfully introduced into a cell, it has to be maintained over many generations. The evolutionary stability and evolvability of networks is a fundamental aspect of cell engineering which is still largely unexplored. In that respect, there has recently

been interesting experimental and theoretical investigations looking at the type of networks selected by evolutionary strategies to fulfil a desired function either in order to implement some artificial network *in vivo* [42] or to investigate the type of solutions arising *in silico* [43, 44]. Unravelling the rules behind the selection of certain solutions and being able to predict their probability of appearance in natural systems would represent a major advance in our understanding of the design principles of cellular networks.

The present approaches in systems biology consider the cell as a homogeneous medium where the components of a network diffuse and interact by the law of mass action. However, the spatial distribution of molecules inside a eukaryotic cell is not homogeneous: protein are targeted to specific structures (the nucleus, the membrane, etc) and they do not always diffuse freely: molecular motors actively transport molecules along the tracks of the cytoskeleton. Hence the structure of the cell appears to regulate the interaction between its constituents. How can these effects be incorporated in the theoretical framework of systems biology? What role do they play in cell differentiation where mechanical stresses on the cell membrane have been shown to influence the specialization of the cell into bone, muscle or neuronal tissue [45]?

4. Development

In multicellular organisms, cells are grouped to form tissues and organs. The foundations of morphogenesis, i.e., the developmental mechanisms by which organisms shape their form and patterns emerge can be traced back to influential works by D'Arcy Thompson in 1917 [47] and by Turing in 1952 [48]. In Turing's proposal, reaction–diffusion equations among appropriate signalling molecules (morphogens) provide a general framework for explaining the patterns observed in early development, the identity of a cell being set by the local value of the morphogen concentration [49]. However, while the existence of gradients and patterns of morphogens have now been firmly established, how these molecules interact both in space and in time to define a robust developmental programme is not fully understood. What type of nonlinear interactions among the various components of these networks (some of which may still be unknown) can explain the observed patterns and their precise spatial extent and temporal dynamics are basic questions that one has to elucidate. In this investigation of the possible mechanisms underlying morphogenesis one is helped (and constrained) by the fact that the developmental programme is surprisingly robust with respect to fluctuations in the environment (temperature, nutrients), in the concentration of morphogens and in the reaction rates (as a result of environmental fluctuations, genetic polymorphism or non-lethal mutations). For example, how do sharp boundaries of gene expression domains cope with the fluctuations of morphogens involved in the patterning of tissues [50, 51]? These observations suggest that a useful criterion to wean out theoretical models is to test their robustness to variations in concentrations, reaction rates, temperature, etc.

Appendix D illustrates some recent work where these issues were tackled and it exemplifies how the nonlinear degradation of morphogens could generate both robust and long range morphogen gradients. The mechanisms by which the degradation of morphogens might effectively be nonlinear are many. For instance the degradation of morphogens could be carried out by enzymes acting on morphogen dimers. In other cases (see for example [52]), an enzyme involved in the degradation pathway might be under control of the morphogen (the higher its concentration, the more enzyme is produced). Finally complex interactions among some components of a given pathway may give rise to non-exponential, i.e., robust, morphogen gradients. A very nice example of the use of nonlinear modelling in morphogenesis is presented

by dorso-ventral patterning in the fly. In this case Eldar *et al* [53] studied the robustness of the morphogen gradient to changes in its production rate (controlled by the number of copies of its gene). Since the topology of the regulation network was known (but not its nine kinetic parameters) they simulated the response of the network for values of the parameters varying by four orders of magnitude. Of the 66 000 tested network configurations, only 198 proved to be robust to changes in the morphogen production rate. These robust configurations were also immune to stochastic variation in the (small) number of interacting proteins. The robust networks were characterized by having a constraint on two of the network parameters, but not on the other seven. This helped identify the nonlinear mechanism responsible for the non-exponential and robust variation of the morphogen gradient. It also suggested a crucial test of the model: alter the constrained parameters and check the robustness of the response.

In these examples robustness to fluctuations in the morphogen concentration at the source was probed. However, as previously mentioned, the developmental programme is remarkably robust to many other perturbations. How these select the possible regulation networks and what constraints evolution imposes on the developmental programmes is not yet known. This brings us to an issue that is still largely unexplored: how does an organism achieve its correct size and proportion?

If a morphogen is produced at the head and degraded at the tail of an embryo, it has been proposed [54] that while its local concentration sets the cell's identity along the antero-posterior (AP) axis, its gradient determines the growth rate. When the gradient falls below a certain value growth stops. One then expects growth to stop far from the source much earlier than closer to it. This, however, is not the case. Growth stops almost instantly throughout the growing tissue. To explain what sets the mean size of organs, Shraiman [55] proposed that the local growth rate is set both by the morphogen concentration (growth increases monotonically when the morphogen concentration is above a certain threshold) and by the local stress P due to the differential growth rates (growth is maximal when the tissue is slightly under tension: $P_{\max} < 0$). The feedback provided by the local stress results in a more uniform growth rate and causes the tissue to stop growing when the cells further away from the morphogen source stop dividing. This mechanism may be used to set the size of the embryo and any other tissue, limb or in the case of the fly, its wing. The idea is appealing and merits further investigation. It is known that stresses can influence the developmental pathway of the fly [56] and are essential for proper bone development. However, little is know about how generic the use of tensile signalling in development is [56] and about what molecular mechanisms are used to couple these signals to gene expression [45]. With the existence of means to apply tensions locally in a tissue (for example optical and magnetic tweezers) these issues could now be tested and compared with robust theoretical models (see for example [57]).

5. Populations biology and evolution

Population biology has emerged as one of the first domains at the interface between mathematics and biology. From the beginning of the twentieth century, it gave rise to several fields ranging from population dynamics, to game theory and the modelling of adaptive evolution. These studies were seminal in many areas of nonlinear phenomena. The papers of Fisher in the 1930s on population dynamics [59] set the stage for nonlinear wave propagation through what became known as the Fisher–Kolmogorov equation. The seminal paper by May on the logistic map (inspired by a model of population dynamics [60]) introduced a generation of physicists to period doubling, chaos and nonlinear phenomena. Finally the rephrasing of game theory in the context of evolutionary biology has been exposed in great detail in the

excellent book by Maynard Smith [61]. In some sense, of all the contributions of nonlinear science to biology this is the most mature.

Over ecological time scales, the dynamics of a population of n individuals is a balance between its growth rate $g(n)$ and its death rate, $b(n)$: $\mathrm{d}n/\mathrm{d}t = g(n) - b(n)$. While the death rate of a population is usually assumed to grow linearly with n, $b(n) = bn$, its growth rate, while increasing with n, saturates at large n for the simple reason that the environmental resources are finite. For bacteria it has been shown by Monod [62] to be well fitted by $g(n) = an/(n_0 + n)$. However for mating populations, the growth rate at small population size may increase not linearly with n but for instance quadratically. This so called Allee effect [63] is crucial when considering the conservation of endangered species: it implies that there is a minimal size below which the population goes extinct. Spatial effects were also considered by Fisher in the case of diffusion of a bacterial population. If the death rate is sufficiently high so that the growth rate never saturates ($n \ll n_0$), it is easy to show that the Fisher–Kolmogorov equation for $x = an/n_0$ is obtained: $\mathrm{d}x/\mathrm{d}t = D\nabla^2 x + rx - x^2$ where $r = (a - bn_0)/n_0$.

Over evolutionary time scales, organisms in populations are selected on the basis of their reproductive efficiency with respect to environmental conditions. In order to assess this efficiency, Darwin introduced the concept of *fitness*. Two hundreds years later no general formulation of fitness has emerged yet, and evolutionary biologists still adapt its definition depending on the context. For instance, within clonal populations (of asexually reproducing individuals), the relative fitness w_{AB} of two sub-population (A and B, see appendix E) is defined with respect to their relative growth rates, $r_{A,B} = \mathrm{d}\ln n_{A,B}/\mathrm{d}t$, as $w_{AB} = r_A/r_B$. If A grown in competition with B outcompetes B then it is fitter than B (and $w_{AB} > 1$). Note that unlike energy, fitness is not a transitive property (see appendix E). While evolution occurs on large time scales, only instantaneous value of fitness can be deduced from the ecological dynamics of the population. This stresses the need to distinguish between ecological and evolutionary time scales in the formulation of the models.

Although most formulations of population dynamics are deterministic, several stochastic components may affect the dynamics. First, the intrinsic stochasticity caused by small numbers of individuals may lead to the extinction of the population (see appendix E). Second, genetic stochasticity may result in phenotypic variability within populations possibly affecting their dynamics. Finally, environmental fluctuations may affect differentially sub-populations. These empirical facts emphasize the need to incorporate stochasticity in the modelling of ecological and evolutionary processes. This was introduced for instance in evolutionary game theory where mixed and conditional strategies correspond to playing several pure strategies with different probabilities (see the book by Hofbauer and Sigmund [64] for a remarkable overview of game theory and population dynamics). Later on, it has been extensively done using agent-based simulations (as in cellular automata [65]), in particular to highlight the role of the interaction range (or spatial structure) on the outcome (see appendix E). The birth–death processes represent also a noteworthy attempt to rephrase population dynamics in a more stochastic manner that are, however, still confined to a few particular cases such as Lotka–Volterra dynamics [66]. Could this be extended to the study of other challenging issues like cooperation which may involve higher order nonlinearities? Which types of models (deterministic versus stochastic) are more relevant when confronted to experimental data?

On evolutionary time scales, the major challenge is to define a framework allowing the derivation of the evolutionary dynamics of a phenotype (e.g. a given trait such as height) from the ecological dynamics of the considered population. For the last two decades, adaptive dynamics (AD) has filled this gap, but their ecological relevance is still questioned (see the review by Waxman and Gavrilets [67] and the following comments in the same issue of *Journal of Evolutionary Biology*). Due to the dependence of this kind of evolutionary model

upon ecological ones, the nonlinear and stochastic features of the latter will necessarily be included in the former. More generally, is adaptation a continuous phenomena or are there abrupt first order transitions? Moreover how can one interlink the different levels of description of evolutionary models (from molecular to phenotypic descriptions)?

In fact, in these investigations, one is not hampered by the lack of theoretical approaches but by the dearth of experimentally controlled and reproducible data. Much ecological and evolutionary data is gathered from field studies in which conditions are tough to control, almost impossible to reproduce and with an outcome that is difficult to predict. Thus, even if various theoretical models fit the data, the uncertainty of the many uncontrolled parameters does not allow us to identify the most relevant one. However, while the evolution of most ecological systems is impossible to study on a human time scale, the ecology of microbial systems is still widely open for experimental investigations. One challenge is to devise long-term self-contained systems replicated many times and probing some well-defined question in evolution. Only then will we be able to compare quantitatively experimental observations and theoretical predictions, measure the probability of certain evolutionary scenarios and test issues of intrinsic stochasticity and robustness.

6. Conclusion

From proteins to genes and from organisms to populations, biological systems display recurrent themes: the nonlinearity of the interactions; the effects of fluctuations arising from environmental variations and the stochasticity inherent in the discreteness and small numbers of players (molecules, cells, individuals); the robustness of the response to these variations and to variation in many of the systems parameters (arising for example from mutations); the role played by evolution both in selecting the present systems and in their possibility to adapt in response to their environment and evolve (their evolvability). The last two themes, robustness and evolvability present contradictory requirements. While robustness implies stability with respect to external and internal perturbations, evolvability requires a measure of adaptation to these variations. How is the compromise between these constraints reached? A possible way out of this conundrum is suggested by the fact that robust processes coupled with the stochasticity arising from the small number of players can give rise to a minority of different phenotypes upon which selection can act.

Although the issues mentioned above can be addressed theoretically, they should also be tackled experimentally. Biology is witnessing a profound transition: from the essentially observational science of the last centuries to a quantitative discipline based on reverse engineering of complex systems: it aims to understand life as it is, not as it could be. The concepts of symmetry and universality which underlie the comprehension of physical systems from magnetism to chaos are of little use in biology where the devil is in the details. As a consequence generic models (e.g. a Landau–Ginzburg type of approach) elicit usually limited interest among biologists. While presenting a real challenge for a theorist, it is only by a thorough understanding of the experimental details that one may expect to build useful models from which hopefully some general evolutionary design principles will be learnt.

Acknowledgments

We acknowledge the financial support from CNRS, from ANR through the ANR-06-PCVI-0005 grant, from HFSP through the RGP003/2007 grant and from EU through a Bio-Nanoswitch grant.

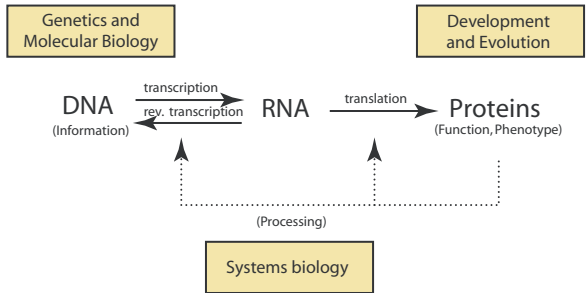


Figure A1. Central dogma of molecular biology.

Appendix A. Some biological notions

Because of their common ancestry all life forms share common motifs (enzymes, metabolic and regulation pathways, etc). The most fundamental one common to most, and dubbed 'the central dogma of molecular biology' (figure A1), is that information is encoded in the famed double helix of the DNA molecule (in the form of genes, i.e., DNA sequences encoding for proteins, regulatory sequences controlling the expression of the genes and an often large portion of DNA of unknown role). To give rise to proteins this information is first transcribed into messenger RNA, mRNA (by enzymes known as RNA-polymerases) and then translated into proteins (by a complex of tens of proteins and ribosomal RNA (rRNA), known as the ribosome). Life forms are mainly divided into two types: prokaryotes, which are single cell organisms lacking a nucleus (essentially a bag of DNA and enzymes enclosed by a lipidic membrane) that appeared about 3.5 billion years ago, and eukaryotes, which are cells with a structured environment possessing a number of so-called organelles (nucleus, mitochondria, etc) that probably evolved from prokaryotes about 1 billion years ago. Within a few hundred million years multicellular organisms appeared first as an assembly of eukaryotic cells and then as more complex organisms with cells specializing into tissues and organs. To proliferate cells divide after having replicated their DNA (with enzymes known as DNA-polymerases). In multicellular organisms, even though all the cells originate from the same initial one (a so-called monoclonal population sharing the same genotype), they specialize ('differentiate') to fill different function (i.e., adopt different phenotypes: skin, muscle, neuron, bone, etc). While the DNA molecule in these differentiated cells is the same, the pattern of expression of the encoded genes (i.e., the concentrations of the various proteins) is not. This pattern is under control of proteins known as transcription factors that bind to the regulatory sequences on the DNA and modulate gene expression. These transcription factors are themselves part of more general signalling networks that play a central role in development by coordinating the differentiation of cells within a given organism. The field in biology that studies gene networks and uses nonlinear modelling to describe them has recently acquired the name of systems biology [23].

Appendix B. Fluctuating enzymatic activity

The catalytic activity of many enzymes is well described by the Michaelis–Menten (MM) kinetics: the enzyme E binds its substrate S forming a complex ES to yield the product P while regenerating the enzyme for the next catalytic cycle (see figure B1(a)). The model assumes that the enzyme has a unique catalytic state, with

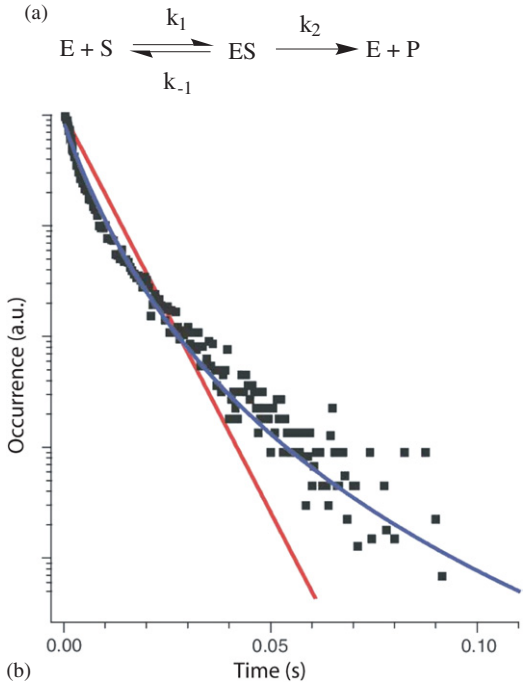


Figure B1. (*a*) Traditional interpretation of Michaelis–Menten kinetics for enzymatic reactions. (*b*) Histogram of enzymatic cycle times for a single β-galactosidase enzyme. Adapted by permission from Macmillan Publishers Ltd from [12], copyright (2006).

corresponding unique reaction rates (k_1, k_{-1}, k_2). The catalytic rate $(v = \mathrm{d}[P]/\mathrm{d}t)$ is given by

$$v = \frac{v_{\max}[S]}{[S] + K_M},$$

where the brackets [..] denote the concentration and $K_M = (k_{-1} + k_2)/k_1$ is known as the MM constant. MM dynamics is normally used to describe biochemical assays, where billions of copies of the same enzyme are involved, but a single molecule version of the Michaelis–Menten dynamics can be formulated in terms of the average time between two successive turnovers. Because the enzymatic cycle is a stochastic variable, for a MM enzyme one expects it to be exponentially distributed. However, this was not observed [12]. As shown in figure B1(*b*) the cycle time histogram for the enzyme β-galactosidase is better fit by a multi-exponential curve (curved line) than by a single exponential one (straight line). These results suggests that the catalytic (functional) conformation of an enzyme is not unique. The enzyme fluctuates between a large number of active configurations each with its own reaction rates. How general is this result? Under certain assumptions concerning the transition rates between the various states it was shown to be consistent with MM behaviour, but whether this can lead to non-MM kinetics in more general cases and how biologically, i.e., evolutionary, significant this process is, are still open questions.

Appendix C. Bistability in gene expression

Switching behaviour in biological systems is observed in many circumstances ranging from developmental biology to microbiology. The interpretation of morphogen gradients (see

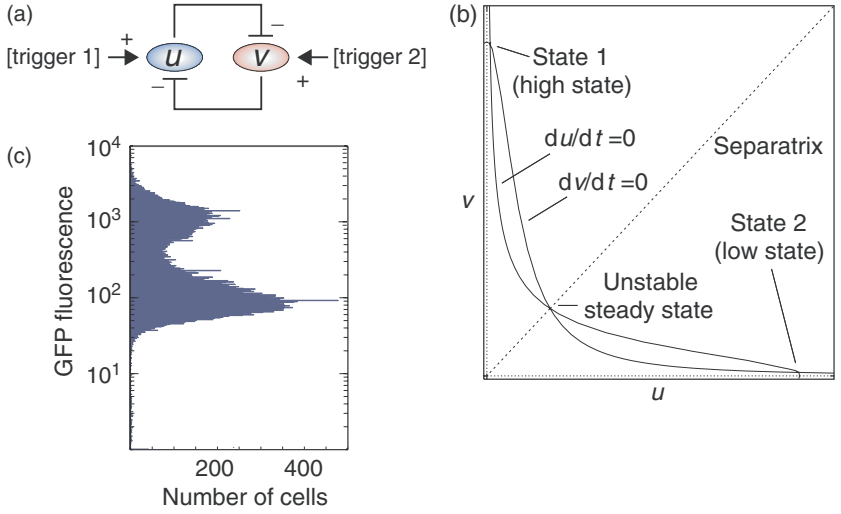


Figure C1. (*a*) Double negative feedback loop. In this circuit, two proteins A and B mutually repress each other. Reprinted from [25], Copyright (2002), with permission from Elsevier. (*b*) Phase space of the steady state solutions of the system described in (*a*). Reprinted by permission from Macmillan Publishers Ltd from [32], Copyright (2000). (*c*) Histogram of a population of cells near the bifurcation point (protein expression is monitored with fluorescent reporter). Adapted by permission from Macmillan Publishers Ltd from [32], Copyright (2000).

appendix D), which control well-defined stripe-like domains in a developing fly embryo, requires the conversion of a graded signal into an all-or-none differentiation response [46]. In the bacterium *Bacillus subtilis*, the ability to withstand stress conditions results from a switch between two states: a default non-competent state and a competent one. In the latter, the bacterium is able to incorporate DNA from the surrounding medium in the 'hope' of acquiring genetic features that will allow it to survive in the new conditions [27,28]. To model this behaviour, synthetic gene networks were engineered. Depending on the environmental conditions, which act on regulatory proteins, elementary circuits such as double-negative feedback loop (figure C1(*a*)) or positive feedback loop were shown to display bistable behaviour. In the bacterium *Escherichia coli*, a synthetic double-negative feedback circuit was engineered with two mutually repressing genes, whose products concentrations (u,v) are described by the equations [32]:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\alpha_1}{1+v^\beta} - u,$$
$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{\alpha_2}{1+u^\gamma} - v.$$

The first term on the RH describes the repression of one gene by the product(s) (e.g. protein) of the other gene. The second term accounts for the degradation of these proteins. α_i is the effective rate of synthesis of the repressor protein i. β and γ are the degrees of cooperativity (i.e., the number of proteins required for effective gene repression). If $\beta, \gamma > 1$ and $\alpha_1 \approx \alpha_2$, this system has two stable equilibrium points (figure C1(*b*)). The capacity of achieving multiple internal states in response to a single set of external inputs is noticeable in a population of clonal (i.e., genetically identical) cells. Indeed, the noise in gene expression may often be enough to project the cell into one of the two stable states (also known as phenotypes). This is most easily demonstrated by putting a fluorescent reporter protein under the control of one of the

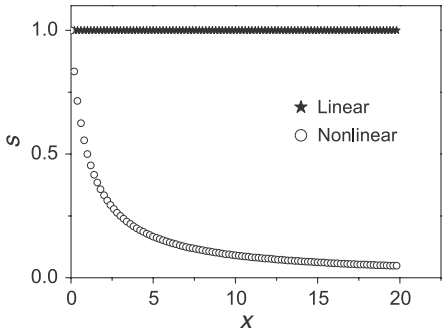


Figure D1. Sensitivity of the morphogen concentration with respect to fluctuations at the source when its degradation is linear or nonlinear.

gene products studied. As shown in figure C1(*c*) the distribution of fluorescence is bimodal. Thus, elementary gene network can convert graded input into a toggle switch.

Appendix D. Morphogens and development

Morphogens are molecules whose spatio-temporal concentration gradients provide positional information to the embryonic cells. The model of reference is the so-called 'french-flag' model [49] where the colours of the flag are determined by various thresholds of a putative morphogen gradient. This gradient is established via localized synthesis and subsequent diffusion as well as degradation within the embryo. As it is used to determine cell fate the gradient must be robust (with respect to fluctuation in production rate) and it must extend far enough (many cells) from its source to be useful. A simple model of localized synthesis and diffusion with linear decay rate generates an exponential gradient for the morphogen. With this type of gradient, robustness and long range cannot be achieved simultaneously (a long range gradient is not robust). If, however, the decay rate of the morphogen is nonlinear (for example if it is enzymatically degraded only in dimeric form) the gradient obtained can be both long range and robust [58]. Let us see how these conclusion can be easily derived. The diffusion equations for a morphogen whose degradation depends linearly or nonlinearly on the concentration can be written as

$$\frac{\partial L}{\partial t} = D\nabla^2 L - \alpha L \qquad \text{or} \qquad \frac{\partial L}{\partial t} = D\nabla^2 L - \beta L^n,$$

where D is the diffusion constant and α (β) is the linear (nonlinear) degradation rate. The corresponding steady-state solutions are

$$L(x) = L_0 \mathrm{e}^{\frac{-x}{\Delta_d}} \qquad \text{or} \qquad L(x) = \frac{A}{(x+\epsilon)^m},$$

where $\Delta_\mathrm{d} = \sqrt{D/\alpha}$, $m = 2/(n-1)$, $A = (Dm(m-1)/\beta)^{2/m}$, $\epsilon = (A/L_0)(1/m)$ and L_0 is the morphogen concentration at the source. A change in L_0 will modify the morphogen level everywhere, causing a shift in the cell fate boundary. One defines the sensitivity s to fluctuations in production rate as the relative change in the morphogen concentration $L(x)$ for a relative change in the concentration at the source: $s = \partial \ln L / \partial \ln L_0$. For linear decay rate the sensitivity is constant ($s = 1$). For nonlinear degradation, $s = \epsilon/(x+\epsilon)$, in other words the further away a cell is from the morphogen source the less sensitive it is to changes in its production rate and the more robust its fate is (see figure D1). Nonlinear decay of morphogen

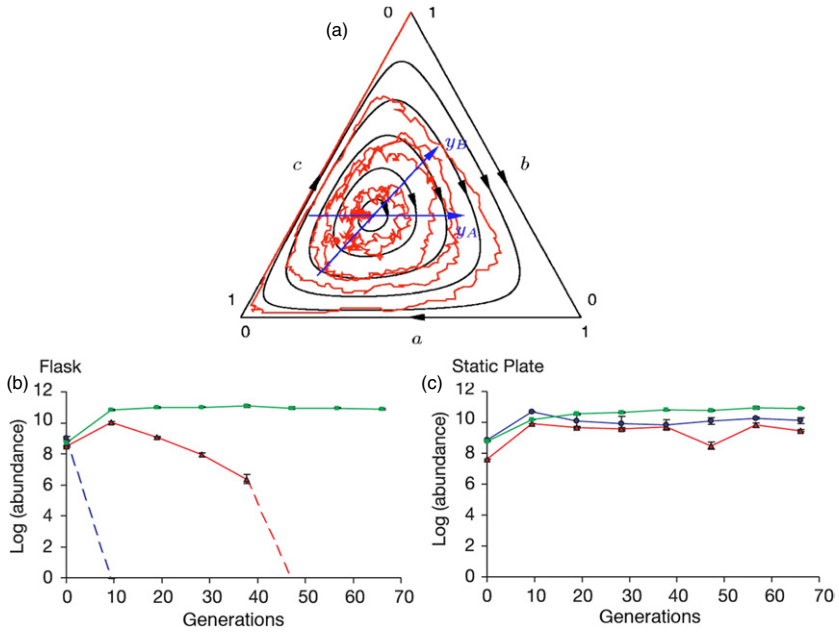


Figure E1. (*a*) Phase space diagram for the three species in a rock–paper–scissors game, showing the predicted cycles around an elliptic fixed point. The erratic flow denotes the trajectory obtained from a stochastic simulation with $N = 200$ agents: it spirals out to reach one of the three trivial fixed points (with only one population surviving). Reprinted with permission from [68]. Copyright (2006) by the American Physical Society. (*b*), (*c*) Results of an experimental realization of the game in bacteria: (*b*) when the experiment is conducted in a well-mixed flask; (*c*) when the experiment is conducted on an agar plate where the interaction range is smaller. Reprinted by permission from Macmillan Publishers Ltd from [71], Copyright (2002).

gradient has been reported in the dorso-ventral polarity of the fly and the patterning of its wing [58]. It could be crucial for establishing a long range and robust gradient of the relevant morphogens.

Appendix E. *In vivo* rock–paper–scissors game in bacteria

Consider a system of three bacterial strains where strain A outcompetes B, strain B dominates over C, and strain C in turn outperforms A, where γ, α and β are the respective competition coefficients. The dynamics of these strains densities (respectively, a, b, c) can be written as

$$\begin{aligned} \dot{a} &= a\,(\gamma b - \beta c) \qquad & A + B &\xrightarrow{\gamma} A + A, \\ \dot{b} &= b\,(\alpha c - \gamma a) \qquad & B + C &\xrightarrow{\alpha} B + B, \\ \dot{c} &= c\,(\beta a - \alpha b) \qquad & C + A &\xrightarrow{\beta} C + C. \end{aligned}$$

These equations possess three trivial stable fixed points $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ (only one of the three strains survives), and one elliptic fixed point (a^*, b^*, c^*) corresponding to coexistence of the three strains. Solutions for this equation can be computed for a variety of initial conditions, which yield stable cycles around (a^*, b^*, c^*) as shown in the phase space diagram (figure E1(*a*)). Using stochastic simulations, fluctuations in finite populations are shown to destabilize these orbits: single trajectories spiral out from (a^*, b^*, c^*), eventually reaching one of the three trivial fixed points. To experimentally study this type of dynamical

problems toxin secreting bacteria [69, 70] were used. Consider three types of bacteria: type A secretes a toxin (colicin) and is resistant to it; type C possesses only the resistance, while type B is not burdened by the production and secretion of the toxin nor by the production of the resistance to it. As a result, B outgrows C, which overcomes A which kills B. That ecological system has been studied when the system is well mixed (corresponding to the case where every bacterium can interact with every other) and when it grows on solid agar plates (where the bacteria interact with their neighbours). In the former case, one strain outcompetes the others in accordance with the stochastic simulations (figure E1(*b*)); in the later, dynamic coexistence can be observed over many generations (figure E1(*c*). These results emphasize the importance of the interaction range in the dynamical outcome: stochastic simulations have been widely used to show that local interactions favour cooperation but more general analytical studies are still rare (see [72]) and long-term experimental investigations almost non-existent.

References

[1] Bon M, McGowan S J and Cook P R 2006 Many expressed genes in bacteria and yeast are transcribed only once per cell cycle *FASEB J.* **20** 1721–3 (PMID: 16818468)

[2] Burgers P M and Kornberg A 1983 The cycling of *Escherichia coli* DNA polymerase iii holoenzyme in replication *J. Biol. Chem.* **258** 7669–75 (PMID: 6345527)

[3] Shnerb N M, Louzoun Y, Bettelheim E and Solomon S 2000 The importance of being discrete: life always wins on the surface *Proc. Natl Acad. Sci. USA* **97** 10322–4 (PMID: 10962027)

[4] Amonlirdviman K, Khare N A, Tree D R P, Chen W-S, Axelrod J D and Tomlin C J 2005 Mathematical modeling of planar cell polarity to understand domineering nonautonomy *Science* **307** 423–6 (PMID: 15662015)

[5] Gerhart J and Kirschner M 1997 *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability* 1st edn (Oxford: Blackwell)

[6] Helles G 2008 A comparative study of the reported performance of ab initio protein structure prediction algorithms *J. R. Soc. Interface/R. Soc.* **5** 387–96 (PMID: 18077243)

[7] Anfinsen C B, Haber E, Sela M and White F H 1961 The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain *Proc. Natl Acad. Sci. USA* **47** 1309–14 (PMID: 13683522)

[8] Anfinsen C B 1973 Principles that govern the folding of protein chains *Science* **181** 223–30 (PMID: 4124164)

[9] Austin R H, Beeson K W, Eisenstein L, Frauenfelder H and Gunsalus I C 1975 Dynamics of ligand binding to myoglobin *Biochemistry* **14** 5355–73 (PMID: 1191643)

[10] Frauenfelder H, Sligar S G and Wolynes P G 1991 The energy landscapes and motions of proteins *Science* **254** 1598–603 (PMID: 1749933)

[11] Henzler-Wildman K and Kern D 2007 Dynamic personalities of proteins *Nature* **450** 964–72 (PMID: 18075575)

[12] English B P, Min W, van Oijen A M, Lee K T, Luo G, Sun H, Cherayil B J, Kou S C and Xie X S 2006 Ever-fluctuating single enzyme molecules: Michaelis–Menten equation revisited *Nat. Chem. Biol.* **2** 87–94

[13] Edman L, Foldes-Papp Z, Wennmalm S and Rigler R 1999 The fluctuating enzyme: a single molecule approach *Chem. Phys.* **247** 11–22

[14] Flomenbom O *et al* 2005 Stretched exponential decay and correlations in the catalytic activity of fluctuating single lipase molecules *Proc. Natl Acad. Sci. USA* **102** 2368–72 (PMID: 15695587)

[15] Deniz A A, Dahan M, Grunwell J R, Ha T, Faulhaber A E, Chemla D S, Weiss S and Schultz P G 1999 Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of förster distance dependence and subpopulations *Proc. Natl Acad. Sci. USA* **96** 3670–5 (PMID: 10097095)

[16] Hanson J A, Duderstadt K, Watkins L P, Bhattacharyya S, Brokaw J, Chu J-W and Yang H 2007 Illuminating the mechanistic roles of enzyme conformational dynamics *Proc. Natl Acad. Sci. USA* **104** 18055–60 (PMID: 17989222)

[17] Henzler-Wildman K A, Lei M, Thai V, Kerns S J, Karplus M and Kern D 2007 A hierarchy of timescales in protein dynamics is linked to enzyme catalysis *Nature* **450** 913–6 (PMID: 18026087)

[18] Henzler-Wildman K A *et al* 2007 Intrinsic motions along an enzymatic reaction trajectory *Nature* **450** 838–44 (PMID: 18026086)

[19] Xue X, Liu F and Ou-Yang Z-C 2006 Single molecule michaelis–menten equation beyond quasistatic disorder *Phys. Rev. E, Stat. Nonlinear Soft Matter Phys.* **74** 030902 (PMID: 17025584)

[20] Min W, Gopich I V, English B P, Kou S C, Xie X S and Szabo A 2006 When does the michaelis–menten equation hold for fluctuating enzymes? *J. Phys. Chem.* B **110** 20093–7 (PMID: 17034179)

[21] Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C and Tawfik D S 2005 The 'evolvability' of promiscuous protein functions *Nature Genetics* **37** 73–6 (PMID: 15568024)
[22] Shen-Orr S S, Milo R, Mangan S and Alon U 2002 Network motifs in the transcriptional regulation network of *Escherichia coli Nature Genetics* **31** 64–8 (PMID: 11967538)
[23] Alon U 2006 *An Introduction to Systems Biology: Design Principles of Biological Circuits* 1st edn (London/Boca Raton, FL: Chapman and Hall/CRC)
[24] Ferrell J E and Machleder E M 1998 The biochemical basis of an all-or-none cell fate switch in xenopus oocytes *Science* **280** 895–8 (PMID: 9572732)
[25] Ferrell J E 2002 Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability *Curr. Opin. Cell Biol.* **14** 140–8 (PMID: 11891111)
[26] Dubnau D and Losick R 2006 Bistability in bacteria *Mol. Microbiol.* **61** 564–72 (PMID: 16879639)
[27] Süel G M, Kulkarni R P, Dworkin J, Garcia-Ojalvo J and Elowitz M B 2007 Tunability and noise dependence in differentiation dynamics *Science* **315** 1716–9 (PMID: 17379809)
[28] Maamar H, Raj A and Dubnau D 2007 Noise in gene expression determines cell fate in bacillus subtilis *Science* **317** 526–9 (PMID: 17569828)
[29] González-Pastor J E, Hobbs E C and Losick R 2003 Cannibalism by sporulating bacteria *Science* **301** 510–3 (PMID: 12817086)
[30] Andrews S S and Arkin A P 2007 Systems biology: a switch for sex *Curr. Biol.: CB* **17** R410–2 (PMID: 17550765)
[31] Ingolia N T and Murray A W 2007 Positive-feedback loops as a flexible biological module *Curr. Biol.: CB* **17** 668–77 (PMID: 17398098)
[32] Gardner T S, Cantor C R and Collins J J 2000 Construction of a genetic toggle switch in *Escherichia coli Nature* **403** 339–42 (PMID: 10659857)
[33] Atkinson M R, Savageau M A, Myers J T and Ninfa A J 2003 Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli Cell* **113** 597–607 (PMID: 12787501)
[34] Becskei A, Séraphin B and Serrano L 2001 Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion *EMBO J.* **20** 2528–35 (PMID: 11350942)
[35] Guet C C, Elowitz M B, Hsing W and Leibler S 2002 Combinatorial synthesis of genetic networks *Science* **296** 1466–70 (PMID: 12029133)
[36] Basu S, Gerchman Y, Collins C H, Arnold F H and Weiss R 2005 A synthetic multicellular system for programmed pattern formation *Nature* **434** 1130–4 (PMID: 15858574)
[37] Elowitz M B and Leibler S 2000 A synthetic oscillatory network of transcriptional regulators *Nature* **403** 335–8 (PMID: 10659856)
[38] Martin V J J, Pitera D J, Withers S T, Newman J D and Keasling J D 2003 Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids *Nature Biotechnol.* **21** 796–802 (PMID: 12778056)
[39] Thattai M and van Oudenaarden A 2001 Intrinsic noise in gene regulatory networks *Proc. Natl Acad. Sci. USA* **98** 8614–9 (PMID: 11438714)
[40] Swain P S, Elowitz M B and Siggia E D 2002 Intrinsic and extrinsic contributions to stochasticity in gene expression *Proc. Natl Acad. Sci. USA* **99** 12795–800 (PMID: 12237400)
[41] Paulsson J 2004 Summing up the noise in gene networks *Nature* **427** 415–8 (PMID: 14749823)
[42] Yokobayashi Y, Weiss R and Arnold F H 2002 Directed evolution of a genetic circuit *Proc. Natl Acad. Sci. USA* **99** 16587–91 (PMID: 12451174)
[43] François P and Hakim V 2004 Design of genetic networks with specified functions by evolution in silico *Proc. Natl Acad. Sci. USA* **101** 580–5 (PMID: 14704282)
[44] François P, Hakim V and Siggia E D 2007 Deriving structure from evolution: metazoan segmentation *Mol. Syst. Biol.* **3** 154 (PMID: 18091725)
[45] Engler A J, Sen S, Sweeney H L and Discher D E 2006 Matrix elasticity directs stem cell lineage specification *Cell* **126** 677–89 (PMID: 16923388)
[46] Clyde D E, Corado M S G, Wu X, Paré A, Papatsenko D and Small S 2003 A self-organizing system of repressor gradients establishes segmental complexity in drosophila *Nature* **426** 849–53 (PMID: 14685241)
[47] Thompson D W 1992 *On Growth and Form: The Complete Revised Edition*. (New York: Dover)
[48] Turing A M 1952 The chemical basis of morphogenesis *Phil. Trans. R. Soc. Lond.* B **237** 37–72
[49] Wolpert L, Smith J, Jessell T, Lawrence P, Robertson E and Meyerowitz E 2006 *Principles of Development* 3rd edn (New York: Oxford University Press)
[50] Gregor T, Tank D W, Wieschaus E F and Bialek W 2007 Probing the limits to positional information *Cell* **130** 153–64 (PMID: 17632062)
[51] Gregor T, Wieschaus E F, McGregor A P, Bialek W and Tank D W 2007 Stability and nuclear dynamics of the bicoid morphogen gradient *Cell* **130** 141–52 (PMID: 17632061)

[52] Eldar A, Shilo B-Z and Barkai N 2004 Elucidating mechanisms underlying robustness of morphogen gradients *Curr. Opin. Genetics Dev.* **14** 435–9 (PMID: 15261661)
[53] Eldar A, Dorfman R, Weiss D, Ashe H, Shilo B-Z and Barkai N 2002 Robustness of the bmp morphogen gradient in drosophila embryonic patterning *Nature* **419** 304–8 (PMID: 12239569)
[54] Day S J and Lawrence P A 2000 Measuring dimensions: the regulation of size and shape *Development* **127** 2977–87 (PMID: 10862736)
[55] Shraiman B I 2005 Mechanical feedback as a possible regulator of tissue growth *Proc. Natl Acad. Sci. USA* **102** 3318–23 (PMID: 15728365)
[56] Farge E 2003 Mechanical induction of twist in the drosophila foregut/stomodeal primordium *Curr. Biol.: CB* **13** 1365–77 (PMID: 12932320)
[57] Morante J, Desplan C and Celik A 2007 Generating patterned arrays of photoreceptors *Curr. Opin. Genetics Dev.* **17** 314–9 (PMID: 17616388)
[58] Eldar A, Rosin D, Shilo B-Z and Barkai N 2003 Self-enhanced ligand degradation underlies robustness of morphogen gradients *Dev. Cell* **5** 635–46 (PMID: 14536064)
[59] Fisher R A 1937 The wave of advance of advantageous genes *Ann. Eugenics* **7** 355–369
[60] May R M 1976 Simple mathematical models with very complicated dynamics *Nature* **261** 459–67 (PMID: 934280)
[61] Smith J M 1998 *Evolutionary Genetics* 2nd edn (Oxford: Oxford University Press)
[62] Monod J 1972 *Recherche sur la Croissance des Cultures Bacteriennes* (Paris: Hermann & Cie)
[63] Stephens P A, Sutherland W J and Freckleton R P 1999 What is the allee effect? *OIKOS* **87** 185–190
[64] Hofbauer J and Sigmund K 1998 *Evolutionary Games and Population Dynamics* (Cambridge: Cambridge University Press)
[65] Doebeli M and Hauert C 2005 Models of cooperation based on the prisoner's dilemma and the snowdrift game *Ecol. Lett.* **8** 748–766
[66] Gutiérrez R and Rosales M J 1998 Diffusion approximations for lotka? volterra type models *Stochastic Models* **14** 809
[67] Waxman D and Gavrilets S 2005 20 questions on adaptive dynamics *J. Evol. Biol.* **18** 1139–54 (PMID: 16135102)
[68] Reichenbach T, Mobilia M and Frey E 2006 Coexistence versus extinction in the stochastic cyclic lotka-volterra model *Phys. Rev. E, Stat. Nonlinear Soft Matter Phys.* **74** 051907 (PMID: 17279939)
[69] James R, Kleanthous C and Moore G R 1996 The biology of e colicins: paradigms and paradoxes *Microbiology* **142** 1569–80 (PMID: 8757721)
[70] Riley M A and Gordon D M 1999 The ecological role of bacteriocins in bacterial competition *Trends Microbiol.* **7** 129–33 (PMID: 10203843)
[71] Kerr B, Riley M A, Feldman M W and Bohannan B J M 2002 Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors *Nature* **418** 171–4 (PMID: 12110887)
[72] Le Galliard J-F, Ferrière R and Dieckmann U 2003 The adaptive dynamics of altruism in spatially heterogeneous populations *Evolution* **57** 1–17 (PMID: 12643563)

Chapter 2

Fluctuations in protein activity

2.1 Fluctuations in protein activity

2.1.1 Fluctuating enzymes

Section 1.3.1 introduced the concepts of static and dynamic disorder in proteins, suggesting that enzymes with identical amino-acid sequences could have different catalytic activities depending on the particular structural conformation they adopt.

One of the first observations of static disorder protein activity was claimed by Rotman in 1961 [78]. In this experience, the author sprayed a diluted solution containing β-galactosidase (see section 2.2.1 below) mixed with a fluorogenic substrate on a glass slide covered with a layer of silicone oil. He then covered the sample with a similar oil coated slide, so that the small microdroplets of enzyme+substrate solution would drop to the bottom slide and be trapped and isolated one from the other. Upon incubation for a certain time the author recorded the amount of processed substrate by measuring the fluorescence of each droplet with a fluorescence microscope for a population of droplets of equal size. The authors used the fraction of empty droplets (no fluorescence) and poisson statistics to verify that the chosen enzyme concentration made it statistically very unlikely for a droplet to contain more than one protein, and indeed they found a quantized distribution of intensities (the lowest peak corresponding to droplets containing only one enzyme). In the first experience the author found that the protein activity was distributed in an "all or none" fashion, the turnover rate being comparable with bulk measurements, and that thermal inactivation resulted would disrupt catalytic activity completely in a similar binary fashion. The experience was repeated with batches of proteins that had been stored under 40% saturated solution of ammonium sulfate at $4°C$ for at least 3 years [79] and

with proteins coming from an 8 year old crystal and contrary to the first experience a broader distribution of intensities was observed, suggesting the possibility that the enzyme activity was similarly distributed. While not completely ruling out that such a broad distribution was due to individual amino-acid modifications or structural damage in some proteins of the old batches, the author speculated that prolonged storaged had allowed the enzyme to reach a three-dimensional configuration of lower free energy with respect to freshly prepared enzymes.

Observation of single molecule enzymatic activity in real-time was then reported in [64], where Lu et al. monitored the fluorescence fluctuations of the active site of the flavo enzyme Cholesterole Oxydase (see section 2.2.1 below) trapped in agarose gel. With respect to the seminal experiences of Rotman, this experiment did not look only at static disorder but also at dynamic disorder. The single enzymes were trapped in agarose gel and identified by the appearance of bright fluorescent spots when raster scanning a planar portion of the gel with a confocal miscroscope. The authors could distinguish between "on" and "off" state of the enzyme and recorded "telegraphic" signals of the type displayed in figure (figure **??**).

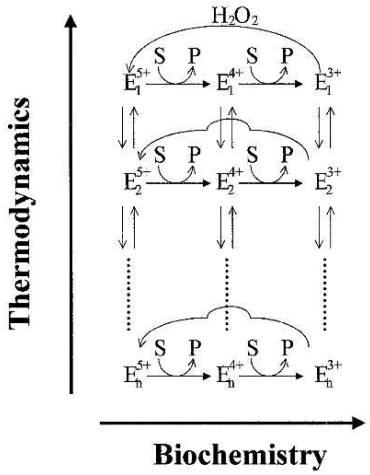


Figure 2.1: Kinetics proposed by Rigler et al. (figure from [31])

The study of the distribution of waiting times in such signals seemed to suggest at least two different active states were necessary to explain the observed distribution. However, good statistics could not be accumulated due to a limitation inherent in the activity detection scheme used. In fact, the flavin fluorophore in the active site of the enzyme would bleach after a certain time of observation, thus limiting the observation window to few minutes.

This limitation was partially addressed by Edman et al. [31], who reported the observation of activity of single Horseradish Peroxydases (see section 2.2.1 below) immobilized on a surface. Their experience proceeded by immobilisa-

tion of biotin conjugated proteins on a streptavidin coated glass coverslip and subsequent incubation in the presence of a non fluorescent protein substrate. Interaction with the protein would turn the substrate in a fluorescent product that was then detected by fluorescence confocal microscopy (3.2.1).

In this experiment as well immobilized proteins were identified by raster scanning a planar region near the surface with a focalised laser beam. The appearance of bright spots, not present in controls without enzymes, was interpreted as reaction product accumulated near an active protein. The experience proceeded by recording the fluorescence fluctuations at each bright spot for a certain amount of time. The sensitivity of this experience did not allow to resolve individual turnovers of the protein, so fluorescence correlation spectroscopy (FCS) (see section 3.3.1) was used to gain an insight on the dynamics of activity at various timescales.

The FCS curves observed could be interpreted with a kinetic model in which the enzme hopped between various states with differet catalytic activity (see figure 2.1), but the quality of their data did not allow precise estimation of the variability of the kinetic parameters. They also observed static disorder across the population of enzymes, but could not draw definitive conclusions on the origins of such a distribution. Also, long data acquisition was impeded by the accumulation of fluorescent product molecules in solution, which contributed to an increase in background noise.

Later on, another similar experience was conducted by Shi et al. [83] on dihydroorotate dehydrogenase (DHOD). Their approach was closer to that of [64] both for the technique used for immobilising the proteins, and the activity detection scheme. In fact the chosen enzyme is a monomeric membrane-associated flavoprotein whose active site becomes fluorescent when it catalyzes the oxidation of dihydroorotate to orotate. This offered a more stable identification for the enzymes, ruling out false positives provided by immobilized fluorophores. Immobilisation was achieved by sterical trapping of the proteins in an agarose gel, in order to avoid the introduction of static disorder due to contact with the surface or by different orientation of the bound proteins. In this .

The experiences mentioned were mainly affected by the same problem: the short window of observation time, the low sensitivity and the possible presence of false positives, had prevented the experimenters from directly observing single turnovers for a long time, in order to build up sound statistics for model building. Indirect evidence (see also section 2.1.2) was however convincing enough to challenge our curiosity push us to start a similar experiment.

In the years 2004 to 2006 important experimental works [94, 37, 35] completed the previous scenario by providing direct observation of individual turnovers in isolated proteins with much higher quality data.

In particular, the results of English et al. [35] are summarized in section 2 of

the article "Some nonlinear challenges in biology" that preceded this chapter.

2.1.2 Memory effect

Both the initial and the more recent observations pointed to the presence of a memory effect [64, 32, 104] in the time correlation of protein activity. In experiments where the enzyme can be labeled as active or inactive, memory effects can be investigated by studying the 2D histogram of "on" times separated by an an certain lapse of time. With cholesterol oxydase in mind, suppose to describe the active/inactive behaviour observed with a binary stochastic process in which the "on" state corresponds to the fluorescent flavin and the "off" state corresponds to non fluorescencent flavin. A 2D histogram can be built of the durations of two "on" times separated by a fixed time *tau*. If the stochastic process is Markovian, i.e. it has no memory effect, the duration of the second "on" time must be independent from the duration of the first. Conversely, the appearance of a diagonal feature in the 2D histograms indicates a correlation in the duration of on times, suggesting that the duration of the second depends on the duration of the first. This correlation can also be studied as a function of the separating interval of time and it was observed to be present on short time scales while absent on longer time scales [64]. The presence of such memory effects were interpreted as a hint that activity fluctuations could be linked to conformational fluctuations in the enzyme structure.

2.1.3 Open questions

These seminal results opened many questions. First of all, the investigation of the origin of activity fluctuations required the definition of a model that could describe these fluctuations in terms of structural deformations. The quantity observed in the experiments is substrate productuion, which contains information on protein activity fluctuations when substrate diffusion is not the rate limiting reaction step. However, no direct observation of structure deformations has been conducted together with activity monitoring. In this regard, it would be nice to be able to either monitor protein fluctuations with an imaging technique such as FRET or to be able to deform the structure by an external force while monitorning the changes in activity.

On the modelling side, there is currently no way of including catalytic activity in a Molecular Dynamics (MD) simulations. The definition of a good descriptor for the activity is subject of active research. Besides this, the total simulation time has not yet reached the microsecond to second scale observed in experiments. Thus it seems challenging and interesting to observe the real-time activity of a protein in the presence of an external pulling force and complete that observation computational investigation that explores the changes in the protein structure

and its conformational fluctuations about the ground state under the action of a stretching force.

2.2 Systems proposed

The design of our experiment was guided by initial considerations and was corrected along the way when new results were encountered and experience accumulated.

The initial choice was simply to repeat the experiment of [31]. We started with a setup that was a replica of theirs. This proved not to be a good approach, because the enzyme chosen, the search technique and the detection scheme were not optimal for the goals that we wanted to implement. This first version of the experimental setup will be called "0.0" in what follows, and it was ready when I joined the group in 2005. The following improvements will be called "1.0" and "2.0", each of which consisted in a different technique that tried to overcome the limits of the previous one.

The detailed technical description of each of these experiments is left to chapter 4, while here we will focus on the considerations that guided us in the choice of the enzyme and fluorophore. The constraints on the systems were many and of different origin, and compromise was required.

2.2.1 The choice of the enzyme

The choice of the enzyme was constrained by many considerations. Initially we looked at Horseradish Peroxidase (HRP), β-galactosidase and flavoenzymes which had been used by our predecessors. Flavo enzymes were fhe first to be ruled out. The fact that flurescence was intrinsic to the enzyme and that the observation window was limited by the dissociation of the flavin in their active site were considered limiting factors for the experience proposed (A possibility to overcome this limitation is suggested in [101] where the authors used FRET between the flavin and a donor molecule to visualize enzyme activity). β-galactosidase and HRP both had positive and negative qualities and we decided to keep both to explore different technical challenges of the setup developed.

β-galactosidase is a hydrolase enzyme that catalyzes the hydrolysis of β-galactosides into monosaccharides that is found in E.Coli. Recombinant production of the enzyme allows site specific attachment of DNA handles for pulling, which is clearly an interesting point for our goal. Also, a wide variety of commercial substrates that yield a fluorescent product exists, making it a good candidate for initial tests. The two main drawbacks are the size of the enzyme, which prevents any simulation and its tetrameric form, which complicates the analysis of activity state. Tetramer dissociation would also be a limiting factor, because the

enzyme would cease to be active upon dissociation, thus limiting the observation window.

Horseradish peroxidase (HRP), found in horseradish, is a small (44 kDa) enzyme that catalyses the decomposition of hydrogen peroxyde in the presence of hydrogen donors. It is used extensively in molecular biology applications primarily for its very fast turnover rate, which allow to quickly develop fluorescence in the presence of a suitable substrate. For our application, it presents complementary advantages and limitations with respect to β-galactosidase. Is is small enough to allow MD simulations and fast enough to allow quick accumulation of fluorescence product (which in the end turned out to be a limitation as well). The main limiting factor is the absence of recombinant HRP, its production being possible only in baculovirus systems. Furthermore. the inclusion of the heme in the apo-protein is to be done in a separate step which further complicates protein production. On the other hand, its routine use for Elisa tests makes it commercially available both in free and biotinilated form (non site specific). This suggested that HRP was also a good candidate for debugging the experimental setup.

Other systems considered where the EcoRV exonuclease and Guanylate Kinase (GK). EcoRV is a dimer protein that cuts in the middle of the sequence 5'–GATATC–3' on double stranded DNA. We proposed to use a modified 10-mer with a FRET [46, 45, 80] couple attached on the extremities. The enzyme cuts the double strand, thus revealing the fluorescence of the donor (D) fluorophore, otherwise quenched by the acceptor. This seemed particularly interesting because we were hoping to be able to detect both arrival time and activity by monitoring the two colours simultaneously. On the technical side this scheme involved the possibility to excite and detect two fluorophores of different colours, which would slightly complicate the setup.

Studied by Choi et al [23] for the investigation of the mechanism of allosteric control, GK is a small enzyme that mediates the transfer of a phosphate group from ATP to GMP, and is often part of large signalling cascades. For our purposes, it has the property of being small and possessing mutants available. We thought of developing a FRET based activity assay which used modified ATP and GMP moelcules. This scheme proved to be too challenging because the modified substrates reduced the affinity of the enzyme and no activity was detected in bulk experiments.

Other possible candidates that were not considered during the course of my experience but that could indeed prove interesting are lipases [37], proteases, ribozymes [108], or hydrolases such as amylases or xylanases.

2.2.2 The choice of the substrate

The choice of the substrate was considered together with the choice of enzyme. The substrate had to be non fluorescent and turn into a fluorescent product upon catalisys. The produce had to have a high fluorescence yield in order to allow single molecule detection (SMD).

Dihydrorhodamine-6G used in [31] was a good choice for HRP, with the only limitation that it spontaneously degraded with time due to interaction with the oxygen present in buffer. This contributed to narrowing the time window and pushed us to look for alternatives. A more stable susbtrate with similar photophysics is (10-acetyl-3 7-dihydroxyphenoxazine), or Amplex Red (Invitrogen, USA). This substrate also had the practical advantage of being commercially available and cheap. A plus for the choice of AR was given by te fact that a substrate for β-galactosidase exists (resorufin-β-D-galactopyranoside) that yields the same reaction product (resorufin).

For EcoRV a taylor made substrate was designed and purchased from (MWG-Biotech, Germany). Considering structures and FRET distances we designed two different substrates (figure **??**), one with a shorter distance of . Unfortunately the short distance substrate was disliked by the protein while the long distance substrate's FRET efficiency was lower than 50%.

The substrates for GK were purchased from (Jena Bioscience, Germany) and

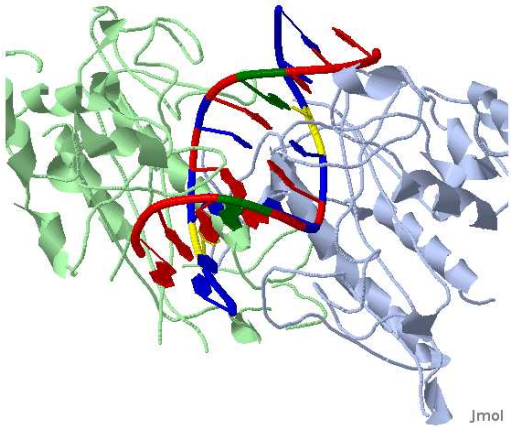
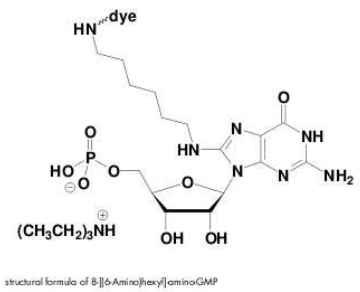
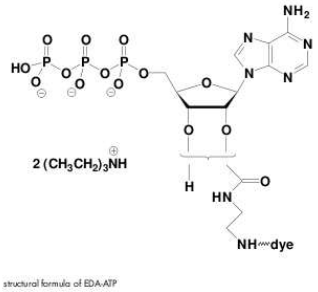


Figure 2.2: Crystal structure of ECORV with DNA substrate. Different bases are shown in different colors (A(red), T(blue), C(yellow), G(green)). The enzyme cuts in the middle of the sequence 5'-GATATC-3'. Two different modified substrates were used. Substrate 1 was: Cy5-GGATATCGGGGG-TMR, substrate 2 was: CY5-GGATATXCGGGGG, where X represents a modified T with an Alexa555 fluorophore attached. Substrate 1 had a low FRET efficiency, but catalytic activity was presered. Substrate 2 had a high FRET efficiency (due to the reduced distance) but no catalytic activity was detected.

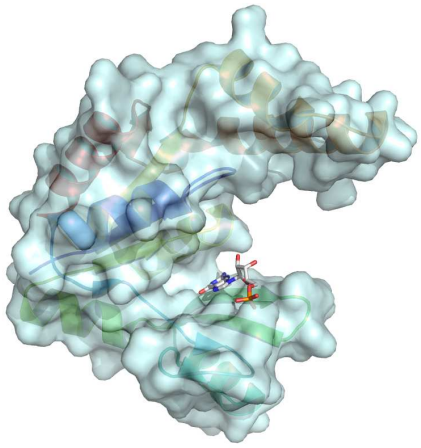
tested in bulk. Unfortunately also in this case, substrate modification resulted in total loss of activity.



(a) Modified ATP



(b) Modified GMP



(c) Guanylate kinase with GMP

Figure 2.3: Modified substrates for Guanylate kinase (GK) and crystal structure of the enzyme bound to GMP. The enzyme has roughly the form of a "C". GMP binds on the bottom hinge and ATP binds on the top hinge. The enzyme catalyzes the transfer of a phosphate group from ATP to GMP. Substrtes were modified by adding a fluorophore (ATTO-647N on ATP and ATTO-532 on GMP) and linker, which resulted in a loss of catalytic activity

2.2.3 The choice of the observation technique

The observation technique was chosen according to the type of biological system studied. Other considerations were: desing simplicity, detection efficiency, available material and cost. We sould probably have included development time as well, which we instead sistematically understimated.

The first setup we built (setup 0.0 section 4.1) was based on [31] and used two photon microscopy as a source of excitation. The main consideration of construction simplicity showed its limits when ectitation efficiency and fluorescence detection were measured and compared with the literature. As we will see in

section 3.1.3 the global efficiency of a fluorescence setup for single molecule detection can be measured by the molecular brightness *eta* which is defined as the number of photons counts per second detected from a single fluorescent molecule. This was proved in [54] to be a good measure of the efficiency also in the case of two photon excitation. In our experiments rhodamine-B and resorufin gave signal to noise ratios of about 1 (see chapter 4 for details) and suggested that the technique was impractical for SMD. An explication of the poor efficiency of this technique in SM excitation is commonly attributed to photobleaching and saturation [29, 10, 33]. Also, the practical need to share the two photon laser with other experiments obliged us to find a compromise on the excitation wavelength, which is a highly critical parameter for two photon excitation efficiency [49], further reducing the efficiency of our setup. Before abandoning the two photon excitation we also explored the possibilities offered by raster scanning of the surface. Section **??** describes the details.

Setup 1.0 (see section 4.2) substituted two photon excitation with the more conventional one photon excitation, but kept the stage raster scanning of the surface. The increase in fluorescence yield per molecule with this setup was of about 10 times. The excitation line (532nm) was chosen with resorufin in mind, and also because of the availability of a green laser in the lab. In this setup we added a high power bleaching beam like that of English et al. [35], but never managed to reproduce their bleaching efficiency. Personal discussion with the authors pointed out that the choice of a 560nm laser had been highly critical in their setup (see section 4.2.1 for details).

Finally, setup 2.0 aimed to improve the search time due to raster scanning of surface bound proteins. In this setup we added TIRF excitation and wide-field detection in order to be able to quickly look for proteins on a larger area. The possibilityto have confocal excitation and detection was kept because of the higher time resolution offered by this configuration.

Chapter 3

Fluorescence Single Molecule Microscopy

3.1 Fluorescence

3.1.1 Fluorescence photophysics

Fluorescence is based on spontaneous emission of light by small excited molecules. An electron of the molecule absorbs a photon an jumps from the ground state to an excited state. Successively a photon of longer wavelength is emitted back and the electron relaxes to the original state. The efficiency of this process is determined by the photon absorption and emission cross-sections which depend on wavelength and can be calculated from quantum mechanics.

All fluorescent organic molecules present delocalized electronic π wavefunctions , which arise when carbon hydrogens bound to carbon in a ring share their electrons. The presence of π orbitals confers rigidity to the organic molecule and conjugates electronic excitations to vibrational modes of the molecule. When an electron belonging to a π orbital absorbs an external photon, it jumps to an excited energy state. Due to conjugation the molecule can quickly relax some of the absorbed energy changing its vibrational mode before a photon is re-emitted and the electron relaxes to the ground state (or to a vibrational excitation of the ground state). The absorption and emission spectra of a fluorescent molecule present many bands that correspond to all possible energy jumps between any vibrational mode of the fundamental state and any vibrational mode of the first excited state.They are often symmetric mirror images one of the other.

A commonly used fluorophore is rhodamine6G (Rh6G), and I will continue the description taking it as an example candidate. To obtain the emission spectrum, Rh6G is excited at a fixed wavelength (at the maximum of absorption), and

the intensity of emitted light is recorded as a function of the emitted wavelength. Inversely, the absorption spectrum is obtained by changing the excitation wavelength and collecting ligth at a fixed wavelength (at the maximum of emission). In the case of Rh6G, the absorption peak is at 530nm, and the emission peak is at 566nm.

3.1.2 Photobleaching

Due to conjugation, fluorescent molecules are rather reactive and there exist many ways in which they can turn into a form which is not fluorescent. This irreversible switch off process is called photobleaching and many different mechanisms can cause it such as: charge transfer, proton transfer, cis-trans isomerization of double bonds, twisted intramolecular charge transfer, photooxidations, photoadditions, etc. Most of these mechanisms lead to non-fluorescent molecules, and depend both on the type of molecule and on external factors such as buffer conditions, temperature, excitation intensity.

3.1.3 Single Molecule Detection

Single molecule detection is possible when the signal emitted by a molecule is separable from the signal emitted by the background, i.e. when the signal to noise ratio is high enough to clearly distinguish the photons coming from a molecule from background photons.

3.2 Excitation techniques

3.2.1 Confocal Microscopy

Conventional one photon confocal microscopy uses laser beam coupled into a microscope objective. The beam forms a cone that converges to a diffraction limited spot on the focal plane. Fluorescent molecules present in the cone of light are excited and emit light. In order to achieve spatial selection and to isolate the light coming from a particular feature being observed at the focal spot, the light coming from out of focus molecules has to be discarded. To obtain this, the focal plane is imaged via a tube lens to an image plane, and a small aperture is placed and conjugated with the focal spot. This will create a spatial filter that selects light coming from a thin longitudinal slice near the focal plane [76]. In other terms, in one photon microscopy one excites all the molecules in the cone of light and then selects the light coming from the focal spot and discards all the other light.

Two-photons excitation (TPE) is a technique widely used in for in-vivo imaging, which presents several advantages with respect to traditional confocal microscopy. Contrary to standard one photon excitation (OPE), in TPE a second

order process is exploited, in which the fluorescent molecule absorbs two photons of higher wavelength nearly simultaneously, and then emits one photon of lower wavelength. Since this is a second order process, we expect the probability of this process to be much smaller than the one photon analogue.

In practice, TPE is obtained focusing a femtosecond pulsed laser beam through a high numerical aperture objective. This produces sufficient intensity only at the focal spot and thus only at the focal spot the fluorescent molecules will be excited. Thus, all the emitted light is coming from the desired spot and it carries signal information to be collected, without the need for pinhole spatial filtering. The setup simplicity of technique was one of the reasons that induced us to use it in the beginning.

Two photon absorption spectra are not just a multiplication by two of one photon absorption spectra. In fact, since two photons are involved, selection rules for the conservation of angular momentum are different in the one photon and two photon cases. Also, the shape and symmetry of the molecule come into play and this complicates the prediction of the two photon absorption efficiency from quantum mechanics calculation. Nevertheless, two photon absorption spectra can be measured and compared with one photon absorption spectra, and indeed a certain degree of correspondence can be found between the one photon and the two photon spectra [105, 11].

Another advantage of the two photon technique, that makes it preferable to one photon for in-vivo imaging, is the weaker interaction of infrared light with organic tissues. This implies that molecules located deeper inside a tissue can be excited. This property of the two photon technique appeared very interesting to observe the molecules emitted by a protein trapped in a gel. In fact we expected to have a lower background noise due to scattering of excitation light on the gel structure, and this consideration reinforced us in the choice of the two photon technique.

Finally, in TPE, the excitation range is well separated from the emission range. In the OPE, dichroic mirror and filters have to be carefully chosen in order to separate excitation and emission light. This subordinates the choice of mirrors and filters to the optical properties of the chosen fluorophore. In other words, a filter set adjusted to collect light from fluorescein cannot be used to excite or observe rhodamine. This exquisitely technical constraint is removed using the two photon technique, because a single dichroic separating infrared from visible can be used and then the infrared wavelength can be adjusted to meet the absorption peak of the chosen fluorophore. This was also one of our considerations when choosing the two photon technique.

3.2.2 TIRF

Total Internal Reflection Fluorescence (TIRF) [7] is a technique that combines the advantage of wide-field excitation with a very low background noise. When passing from a medium with high refractive index n_1 (like glass or oil) to a medium of lower refractive index n_2 (like water or air), light rays are deviated according to Snell's law (figure 3.2.2). If the incidence angle from the denser

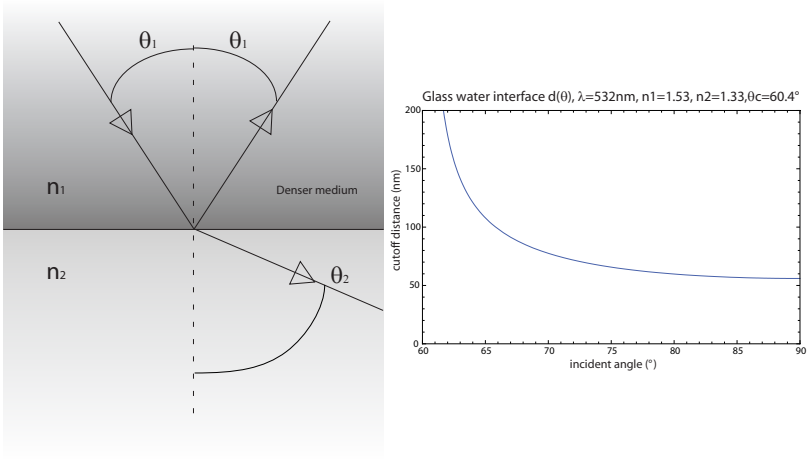


Figure 3.1: Snell's law and total internal reflection: A ray of light travels in a dense medium and it is partially reflected at the interface with a less dense medium. If the angle of incidence is large enough, all the light is reflected and the EM field intensity in the less dense medium decays exponentially with distance from the surface. The cutoff distance depends on the incidence angle.

medium is increased, a critical angle $\theta_c = \arcsin(n_1/n_2)$ is reached where light is completely reflected (internal reflection). The EM field intensity on the other side of the interface decays exponentially:

$$I(z) = I_0 e^{-z/d} \qquad \text{where} \qquad d = \frac{\lambda_0}{4\pi\sqrt{n_1^2 \sin^2\theta - n_2^2}} \tag{3.1}$$

In TIRF fluorescence microscopy, the excitation beam is arranged so that an evanescent wave forms at the glass/water interface of a sample. This means that only flurescent molecules in a volume slice of about $100nm$ thickness above the surface (see figure 3.2.2) will be excited and thus detected. This technique is perfect for both single molecule detection and tracking.

3.3 Detection techniques

3.3.1 FCS

Fluorescence Correlation Spectroscopy (FCS) is a technique aimed at extracting information from noisy environments, exploiting the time correlation functions of emitted light. This technique measures the intensity of light emitted at the focal spot as a function of time $I(t)$ and calculates the auto-correlation function of the signal $G(\tau) = \langle \delta I(t+\tau)\delta I(t)\rangle / \langle I(t)\rangle^2$. It was first used in confocal detection schemes (see [57] for a recent review) and then extended to TIRF configuration [48, 92]. It is useful to quantify the efficiency of a setup by measuring the number of photons extracted by a single fluorescent molecule.

Derivation of Correlation function for simple diffusion

The correlation function for simply diffusing molecules is:

$$G(\tau) = \frac{1}{N\sqrt{1+\left(\frac{\tau}{\tau_d}\right)}} \tag{3.2}$$

N is the average number of fluorescent molecules present in the focal spot and tau_d is the average diffusion time across the XY section of the focal spot. The quality of the setup depends on how many photons are collected when a diffusing molecule passes through the excitation volume $q =< I > /N$. This value can e measured as a function of external parameters such as input power, beam size at the back of the objective and laser wavelength..

3.3.2 Single photon counting and time tag (SPCTT)

SPCTT means that individual photons collected by the detector are not binned but tagged by their time of arrival. For lifetime applications and in general when high frequency sampling is required, the advantage of this technique consists in the reduction of space required for storage (if no photon arrives, no data is stored). This, and the high resolution on the time of arrival, allows more complex algorithms to be applied off-line in order to extract more information from the recorded signals. Traditional light detection, digitizes the intensity measured by the photon detector at a fixed sample rate. This implies one of the following: at high intensity/low sample rate the signal is averaged and information on dynamics at time scales shorter than the sampling frequency is lost; at low intensity/high sample rate many samples contain no photons and add just a trail of useless zero bins filling up data space. On the contrary, SPCTT uses a fast clock to digitize the time of arrival (with picosecond resolution). The size of data stored is then proportional to the number of photons received and not to the number of bins. Furthermore, all the information available is stored, and can be subsequently

treated off-line for calculation of intensity traces, correlation functions and other statistics on the collected light.

Chapter 4

Fluorescence setups

4.1 Setup 0.0

4.1.1 Description

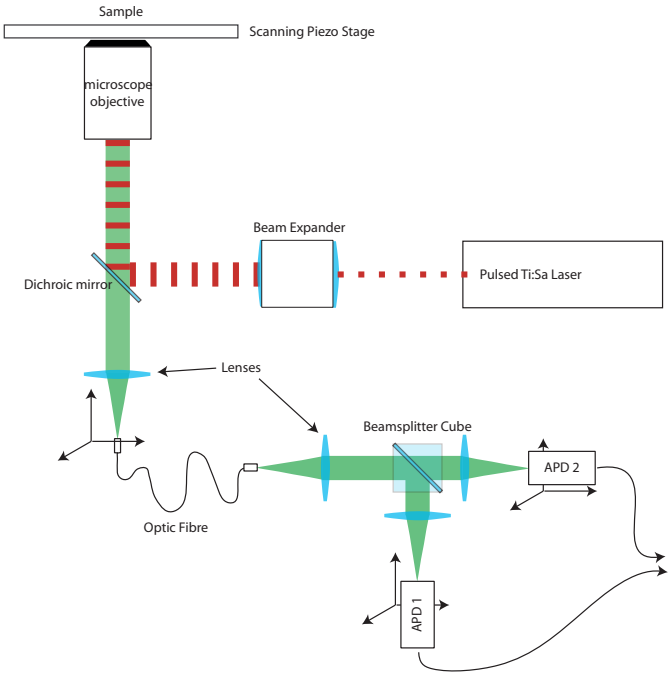


Figure 4.1: Scheme of setup 0.0

The first setup was quite simple in conception. A pulsed 820nm Ti:Sa laser beam (MIRA, Coherent, USA) was expanded to overfill the back aperture of an

inverted 60× microscope objective (UplanAPO, Olympus, Japan). This configuration focalized the beam to a diffraction limited spot on the focal plane to obtain two photon excitation (see section **??**. The small absorption cross section of this second order process implies that only the molecules found in a volume of approximately 1μm diameter and 4μm thickness are excited and emit light. This is the reason that first induced us to chose this technique for our application. The sample was attached to a scanning piezoelectric stage (P-527C, Physik Instrumente GmbH, Germany) that allows for a displacement of 200 μm in X and Y direction with a resolution of 20 nm. Coarser displacement was done manually releasing the screws that attach the sample holder to the stage and then fixing them again in the new position.

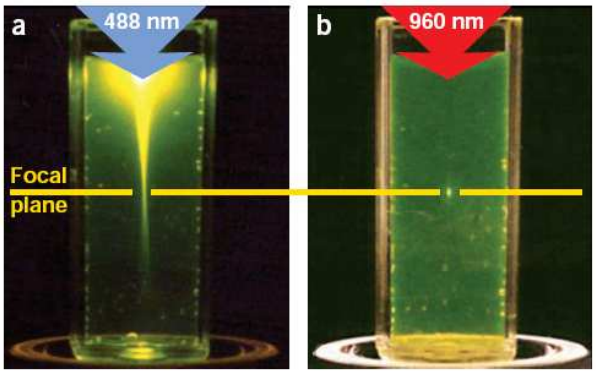


Figure 4.2: Two photon excitation (figure from [109])

Emitted light was collected through the objective and focused to a 200μm core multimode optical fibre (M25L02, Thorlabs GmbH, Germany). Detection was achieved by splitting the emergent beam in two halves and coupling each to an APD (SPCM-AQR-14, Perkin Elmer, USA). Each APD was mounted on an XY alignment and its position could be regulated with a nominal precision of 10μm. Unfortunately, mechanical hysteresis in the screws and the choice of support affected the overall alignment reproducibility. The setup was aligned following the procedure described in ([25]) which I briefly summarize in appendix A.3.

FCS was used to measure the efficiency of the setup. Figure 4.3 shows the typical shape of an FCS curve for fluorescein 100nm. The overall efficiency of the setup is defined as the average number of photon counts per second per molecule in the excitation volume. Figure 4.3 was obtained when the average count in each APD was about 100kHz, yielding a value of about 1kHz per molecule for common fluorophores like fluorescein and rhodamine-B.

4.1.2 Scanning stage

Stage scanning proved to be so painful that we finally abandoned it. Initially we thought of doing a pixel scan, stopping the stage at each step for acquisition. However, the typical response time for such a stage is $10ms$ at best. With a focal spot of $\tilde{3}$00nm an image of 256×256 pixels covering 85μm $\times 85\mu$m took about 11 minutes, an incredibly ong time for debugging purposes. Furthermore, the upper limit for this experimental configuration was set by the background noise increase due to product accumulation. Considering an average volume of $10\mu^3$m per protein (corresponding to atto-molar concentrations for the protein) and a a turnover rate of about 1kHz (for HRP) it is easy to realize that pico-molar concentrations were reached within a quarter of an hour.

4.1.3 Gel immobilisation

In [64] and [83] the authors used a gel to trap isolated proteins in a well defined three dimensional position. The approach seemed very interesting because of the use of the spin-coating technique, which confined the proteins to a 2D surface between two slides of gel. Agarose gel was mixed with proteins at aM concentration and with substrate and sandwiched between two coverslides. Scanning of the sample was performed looking for high intensity spots. When a bright spot was identified, the beam would be focalised and photons collected for inspection. This search procedure invariably yielded traces of decreasing intensity and we

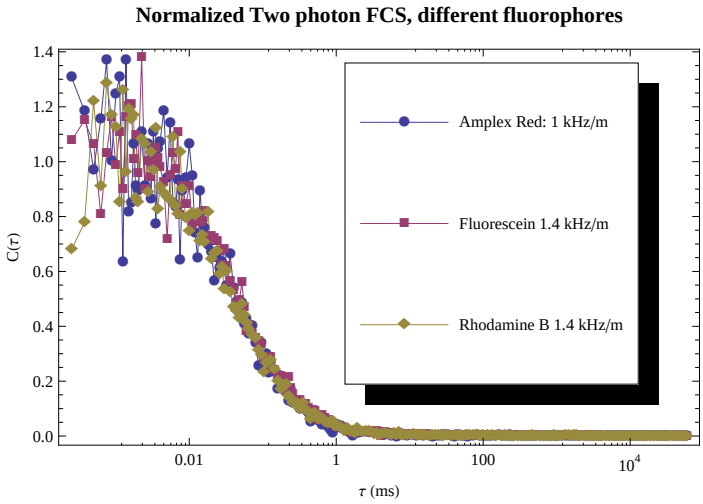
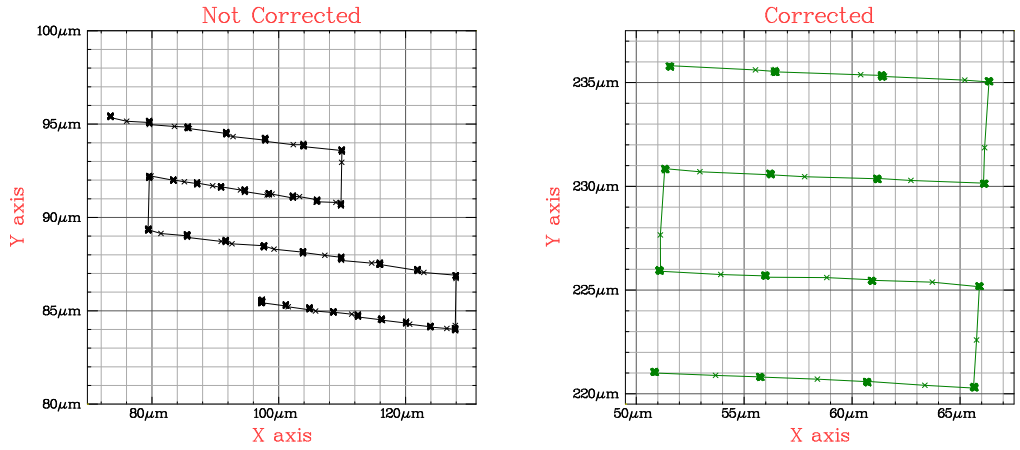
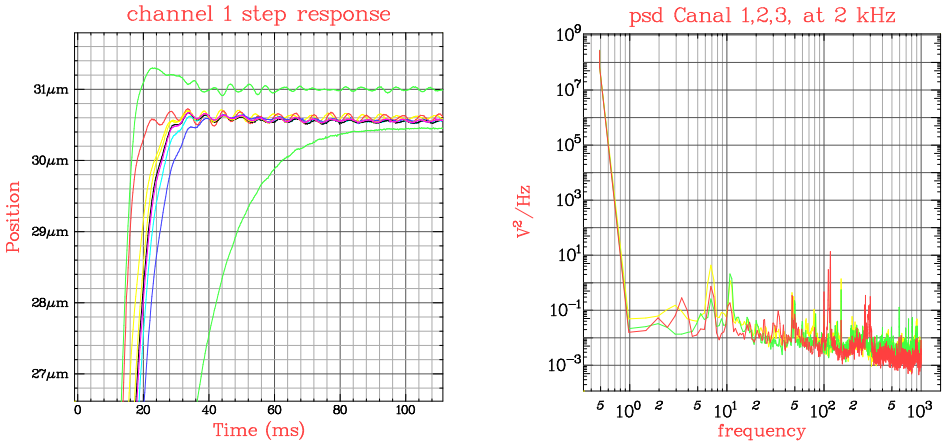


Figure 4.3: FCS curve obtained in the first setup. The curves were normalised for comparison. The average number of molecules in the focal spot was ≈ 100 for fluorescein and rhodamine-B and ≈ 10 for resorufin (Amplex Red after interaction with HRP).



(a) Not corrected (b) Corrected

Figure 4.4: Tracking the position of a bead anchored to the glass surface while it is displaced with the piezzo stage. Fig. 4.4a shows that the scanning stage does not respond with a step of equal size when displaced on the left or on the right. Fig. 4.4b shows the behaviour of the stage after the input signal has been corrected to take into account the asymmetry.



(a) (b)

Figure 4.5: Step response of the piezzo stage an power spectrum. The piezzo stage had low frequency resonances that were excited in scanning mode, leading to vibrations and deterioration of the quality of positioning.

concluded that the higher brightness was to be attributed to impurities in the gel rather than to active proteins.

4.1.4 Conclusion 0.0

[29] report an efficiency as high as 25 kHz per molecule for a two photon confocal setup comparable with ours. However, they observe a 3-fold reduction of efficiency switching the excitation from one photon to two photons. The authors of [33] studied the probability of bleaching for one and two photon excitation and reported that the total number of extracted photons is reduced 18 folds in two photon excitation with respect to one photon. In all our tests with comparable fluorescent dyes and laser power we never obtained more than 2-3 kHz per molecule with the fiber and 5-7 kHz per molecule with direct coupling of the microscope to the detectors. These values are one order of magnitude smaller that those reported in the literature, and

4.2 Setup 1.0

4.2.1 Description

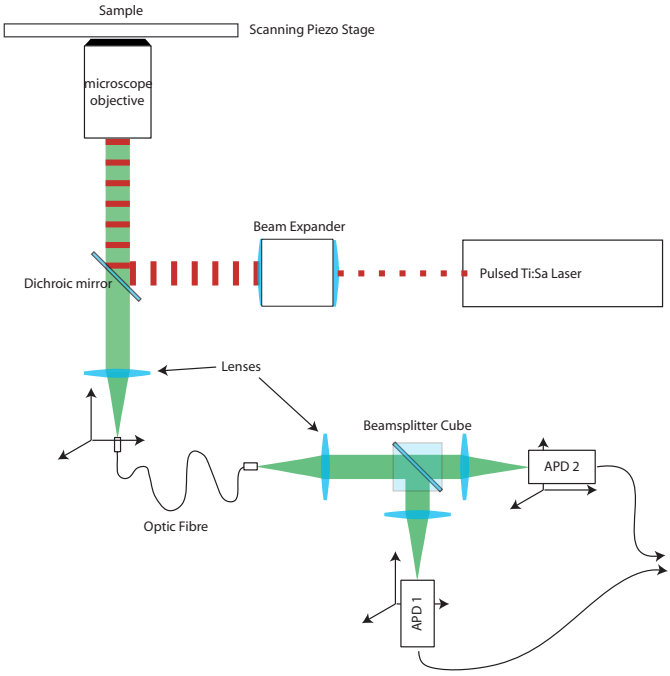


Figure 4.6: Scheme of setup 1.0

The second optical setup we built aimed to overcome the following limitations of setup 0.0:

- Low overall efficiency of the setup
- Background noise increase

We were inspired by te paper [35] and by discussion with experts at the OSBD conference.

Two photons excitation was substituted with 1 photon $532nm$ excitation. This change alone brought us from about to kHz per molecule for powers of .

Pinhole

The implementation of a one photon confocal setup implies not only the change of dichroic mirror but also a change on the detection pathway. In fact in this configurations, molecules are excited in the whole lightcone and not only at the focal spot. thus light coming from the focal volume has to be spatially selected in the detection pathway by optically conjugating a pinhole to the focal spot.

The role of the pinhole is not so much to filter light in the XY plane but in the longitudinal direction. Sometimes optical fibres are used for pinhole but we found that even a $200\mu m$ multimode fibre (Thorlabs GmbH, Germany) was degrading light transmission by at least 30% ad we decided to eliminate it from the collection pathway.

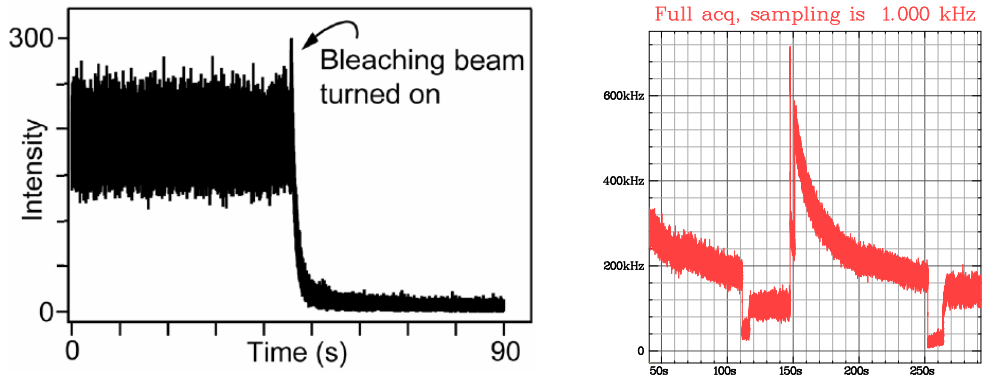
APDs

APDs were mounted on a more stable and precise alignment system (sensitivity 5μm). Also, we found that a proper criterion for good alignment was the observation of a plateau when moving the APD along the XY axes. This creterion is useful to get the correct alignment in the longitudinal direction. In fact, the optical image of the pinhole on the APD chip is about $100\mu m$ in diameter. The APD silicon detector is $170\mu m$ in diameter an dthus it must be possible to move the XY adjustment knowbs for some microns without noticing changes in the collected intensity. Over all, the improvements in detection and excitation contributed to increase the fluorescence per molecule from 1kHz per molecule to about 10kHz per molecule.

Background noise reduction

The other issue we tried to attack in this second version of the experiment was product accumulation leading to background noise increase. The following features were implemented:

- Gel trapping was abandoned in favour of surface binding. Slightly more complicated to obtain, this allowed us to reduce the total amount of protein present in the sample without increasing search times.
- Experiments were conducted in a gentle flow, in order to allow for a partial elimination of the accumulated fluorescent product.
- A powerful 532nm bleaching beam was added in order to bleach all the fluorescent molecules in a region of about $200\mu m$ of diameter around the focal spot.



(a) (b)

Figure 4.7: Bleaching beam comparison

Figure 4.7 compares the figure reported by English et al. in the supplementary material and the performance of our setup. The striking difference is that while they obtained a decrease in fluorescence when the bleaching beam was switched on, we obtained a sudden increase. It took us a lot of time to identify the cause of this. Leakage of laser light to the detection pathway was ruled out by adding filters and notch filters. Insufficient intensity for bleaching was also ruled out by turning off the bleaching beam after 10s and measuring the intensity, which was indeed lower than the initial one. Finally, discussion with the authors pointed out that the substrat might be excitable at 532nm and that, for the success of their experiment, the use of a 560nm laser had turned out to be a key issue. In the absence of such equipment the bleaching beam was abandoned.

4.2.2 Conclusion

Setup 1.0 brought moderate improvements on the excitation and detection efficiency but it was still palgued by a major fault: the use of a home built confocal scanning procedure made protein search so long and irreproducible that routine test of surface passivation protocols and quick debugging was impossible.

Furthermore, even if efficiency had been improved, it was still lower than that reported in the literature.

4.3 Setup 2.0

4.3.1 Description

It was only in Summer 2007, after a long period of insuccesses that the decision was taken to review once more the design of this setup. I was to expecially thank Prof. S. Weiss for the fruitful discussion and the critical approach that he brought in analyzing the failures of this experiment. Following is a list of the key flaws that were identified:

- The beam alignment algorithm was not consistent and we had no precise way to be sure of the alignment. this influenced the quality of our point spread function (PSF), which turned out to be very important when efficiency was considered.
- Single Molecule Detection had not been validated separately from FCS. In other words we had never checked that our setup was able to resolve single fluorescent molecules passing through the confocal volume.
- We didn't have debugging tools to quickly assess the quality of the alignment and the possible sources of noise.
- We had no precise idea of the amount of protein deposited on the surface, the only quantification being an apriori estimation and the surface scan. TIRF imaging was added to the setup for this reason.

In some respect the current setup resembles those described in [93, 48], with some changes that enhance its usability and sensitivity.

Excitation

The excitation patwhway was completely re-designed to improve the quality of the PSF. A 532 Nd:YAG laser (GCL-532-L, CrystaLaser, USA) provides the excitation light for fluorescence illumination. The light passes through a 1:1 telescope that corrects the divergence of the beam and is then separated into two parts by a 50% beamsplitter. Each beam is separately coupled to a single mode fibre (P1-460A-FC-2, Thorlabs GmbH, Germany) in order to eliminate spurious higher order spatial modes.

The first fiber is used for confocal illumination. It is coupled to a microscope objective (L-10X, Newport, USA) to obtain a parallel beam and then is sent to the dichroic mirror which reflects it into the microscope objective (UplanAPO,

Olympus, Japan) in order to obtain a diffraction limited spot. The fiber and the objective are fixed on an adjustable support (F-91-C1 with connector FPH-CA4, Newport, USA). The use of this piece instead of the alignment mirrors and lenses improved the overall stability of the alignment as well as the quality of the PSF.

The second fiber is used for TIRF imaging. A lens focalises the light coming out to the back focal plane of the microscope objective and a positioning stage is used to displace the incoming beam in order to obtain total internal reflection.

Switching from confocal to TIRF mode is possible by flipping a single mirror on the excitation pathway. It is important to observe that the most delicate confocal beam does not need this mirror, which is only inserted when TIRF is needed.

Sample

The sample is composed by a glass chamber, similar to the one described in section 6.2.2. A channel is cut into a thin parafilm, and two glass coverslips are used to provide the lower and upper cover of the channel. Two holes are pierced in the upper coverslide to allow for exchange of buffer solution. The sample is firmly attached to the piezo stage which is only used to center the confocal beam on a precise position in the sample, but not for scanning

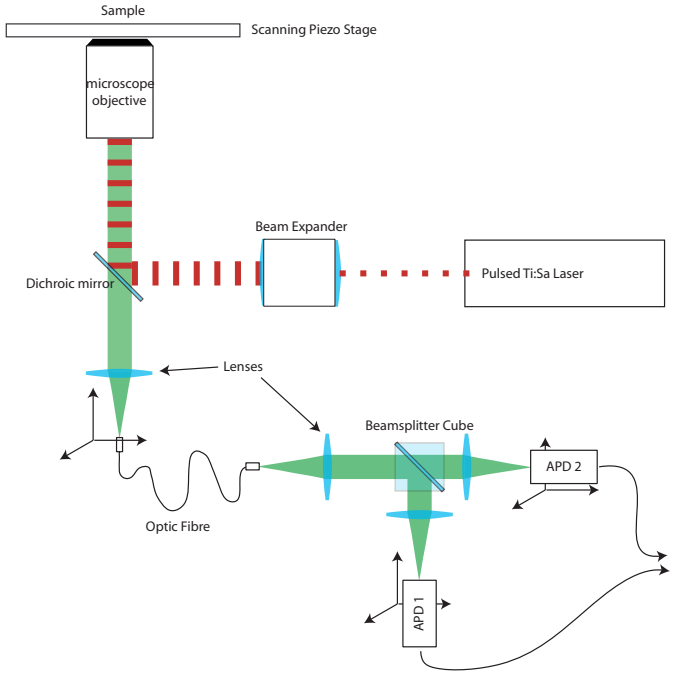


Figure 4.8: Scheme of setup 2.0

Emission

On the detection side another flippable mirror was added befor the pinhole, in order to send the collected light to an EMCCD (Ixon 885, Andor, UK). Another improvement was the insertion of a quick spectrometer to analize the spectrum of incoming light. This is obtained by sending the confocal beam through a diffractive Amici prism which separated the different wavelengths on a CCD camera (CV-M40, JAI, Denmark). Approximate calibration of the spectrometer was obtained by sending ligths of known wavelength through the microscope objective.

4.3.2 Single Molecule Detection

Single molecule detection in bulk was the first achievement of the new setup as is clearly shown in figure **??** which shows the intensity trace recorded by APDs for a sample containing free rhodamine-6G in solution at sub nanomolar concentration. Each burst corresponds to the passage of a single molecule in the confocal volume.

4.3.3 Surface preparation

The amount of protein deposited on the surface was qualitatively estimated by coupling HRP to labeled streptavidin. This showed that the estmated density of surface protein was actually much lower than the real one (see figure 4.9). We also observed sticking of product molecules on the surface, which could probably explain the origin of the disappearing of previously observed surface bright spots.

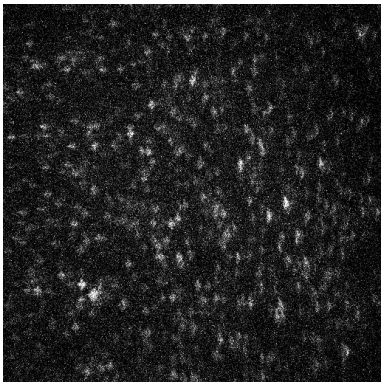


Figure 4.9: TMR labeled streptavidin molecules specifically bound to PEG-biotin on the sample glass surface, visualised by objective TIRF. The concentration of proteins at the surface can be partially controlled by modifying the ratio of PEG/PEG-biotin used to coat the surface.

4.4 Open questions and future development

Despite the many improvements and the long development of this setup, not even the first mileston of observing protein activity in real time has been obtained. What conclusions ca we draw from this experience? On the project management side the most important consideration is the experience factor. The project of an experiment in fluorescence spectroscopy without any previous experience in the field proved to be too audacious. In particular, untill summer of 2007, inexperience had prevented us from clarifying the preliminary goals to be achieved on the way to activity observation. Inexperience was a major factor for insuccess because we didnt' know what to benchmark the performance of our setup against. A clear example of this is the long time that passed before scanning two photon technique was abandoned because not suited to the experiment we envisioned. Since discussion with experts proved to be so fruitful, another lesson to be learned from this experinece is that it should probably have occurred earilier along the way. Scientific ideas are rooted in dialectic discussion and not in mere deduction and mind-power, but it is very easy to carve a hole of technical difficulties. This should have been considered during the months spent to obtain a properly functioning scanning stage.

On the scientific side, the pursuit of this experiment will have to pass thourh these intermediate goals:

1. Detection of protein activity in TIRF mode: This will probably require a change in excitation wavelength ($560nm$) and possibly the change in enzyme system. Two systems which have been successfully used in similar experiments are Alkaline Phosphatase (AP) [30, 26] and Lipase [37]. Unfortunately fluorescent substrates for AP do not exist in the wavelength used by our setup.

2. Modification of the enzyme for the application of a force. Recombinant production is essential for this step, which rules out HRP from the list of candidates. AP, lipase and β-galactosidase all exist in recombinant form, althought the tetrameric form of the latter, may render the whole modification step a little difficult. Attachment of handles is to be done via cystein modification. Our current protocol for DNA pulling uses Digoxygenin-Antidigoxygenin for binding DNA to the glass surface and biotin-streptavidin for binding the DNA to the magnetic beads. If TIRF illumination is kept, the protien must be within $100nm$ from the surface, which limits the length of the DNA linker to little more than one pesistence length or some hundredths of bases. A alternative route would be to directly modify the protein for Digoxygenin attachment, the feasibility of which is to be explored. The binding of beads will also imply higher background noise due to light scat-

tering from the bead. Thiss will be addressed by choosing a DNA linker of at least $1\mu m$ in order to keep beas out of the evanescent wave.

3. If the previous step is completed, the interpretation of the results will necessarily have to pass through a computational model capable of connecting structural deformations due to stresses on the protein surface with changes in the active site configuration and consequently in catalytic activity. This non trivial step has been partially attacked by our collaborators in [81] but will demand additional investigation for other protein systems.

Part III

Magnetic tweezers experimens on DNA

This part is dedicated to the experiments conducted in the years 2007 and 2008, which focused on the torque measurement on DNA. The molecular structure of DNA confers to the polymer chain both a bending rigidity. While the former is now quite well understood thanks to the experimental and theoretical effors of the past 15 years, the latter is less well characterised. This is due to the difficulty of conducting sensitive measurements as well as to the complications involved in modelling the problem.

The following four chapters are dedicated to this subject and to the description of the experimental work conducted to address this problem. Chapter 5 will summarize the current knowledge in coarse grained modelling of DNA, with a particular attention to the open questions regarding torque. Chapter 6 will introduce DNA micro-manipulation techniques and the magnetic tweezers apparatus in detail. Chapter 7 will present the results of precision measurements we conducted with the magnetic tweezers, and it will introduce a new method for estimating the applied torque in MT experiments. The aim of these experiments was both to validate a new paradigm for torque measurmeent on DNA and to address the fundamental question of the dependence of the twist rigidity on ionic conditions. Finally 8 will describe the development of a new apparatus that we developed to directly measure the torque applied to a DNA molecule.

Chapter 5

DNA elasticity

5.1 Topology of DNA: Linking number, Twist and Writhe

5.1.1 The linking number

The helical nature of DNA has profound implications for its mechanical properties and for its geometry. Nature is full of cases in which the DNA double helix forms a closed loop. For example, the genome of a bacteria is formed by a single closed loop of double stranded DNA; plasmids, small circular fragments, are used by bacteria to exchange genetic material, and are one of the basic tools of molecular biology. When DNA is closed in a loop, each of the two strand also forms a loop. It is not possible to separate the two strands completely, unless we break one of them. Mathematically this is formalised by defining the linking number Lk as the number of turns one strand makes around the other. This is a topological invariant, meaning that it does not change if DNA is deformed unbroken.

In the case of magnetic tweezers experiments, the DNA will not be a closed loop, but the equivalent of linking number can also be defined based on the observation that both ends of the DNA are firmly attached either to a glass slide or to a large bead so that rotation of one strand around the other is forbidden. In this case the relaxed state linking number is defined as $Lk_0 = N_{bp}/p$, where N_{bp} is the number of base pairs of the DNA fragment used in the experiment and $p = 3.6$nm is the helical pitch.

5.1.2 DNA supercoiling

Consider a linear piece of DNA in a sample containing a ligating enzyme. The chemical reaction catalysed by a ligase consists in sealing the two open ends of the DNA fragment, effectively forming a closed loop. Due to thermal agitation there is no reason to suppose that all the loops will be exactly closed so that

the two strands wrap around each other exactly Lk_0 times. In fact, experiments show that a symmetric distribution [106] peaked around Lk_0 will appear. We define the degree of supercoiling $\Delta Lk = Lk - Lk_0$ as the number of passages of one strand around the other that are needed to go from the relaxed state to the actual state. This quantity can be positive or negative, depending if turns were added in the sense of the helix (+) or in the opposite sense (-). We can also define the supercoiling density $\sigma = \Delta Lk/Lk_0$ in order to be able to compare the degree of supercoiling for molecules of different length.

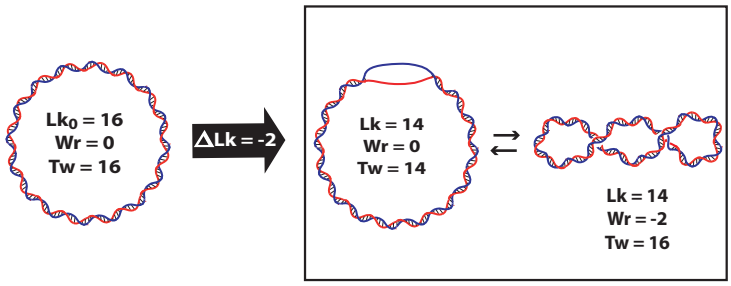


Figure 5.1: A circular DNA molecule in the relaxed state with $Lk_0 = 16$. If one strand is broken, turned two times around the other and then religated a deficit of $\Delta Lk = -2$ has been introduced in the molecule. This can be absorbed either as Twist or Writhe or a mixture of both(figure from [61])

5.1.3 Twist, Writhe and plectonemes

Changing the degree of supercoiling of a string remarkably changes its mechanical properties. The same observation is valid in the case of DNA, where changes in the degree of supercoiling affect both its local properties and its global conformation. In order to describe the effect of the change in linking number on DNA structure it is convenient to introduce two more variables related to it: Twist (Tw) and Writhe (Wr). Twist measures the local action of torsion and is defined as the number of times a strand wraps around the axis of the DNA filament. Writhe is related to the global properties of the DNA loop, and it is defined as the number of times the axis of the helix wraps around itself.

White's theorem [102] states that these two variables are not independent for a closed DNA loop, and that their sum is equal to the linking number ($Lk = Tw + Wr$). In the relaxed state the DNA helix forms a loop and writhe is null. A change in linking number can produce both a change in twist or in writhe depending on what is energetically favourable. This will depend on internal parameters like the bending persistence length or the twist persistence length, as

well as on external factors like buffer salinity or external force.

Fuller [41] obtained an exact formula for the writhe of two different configurations of DNA: the solenoid and the plectoneme. He showed that for fixed writhe, the configuration that minimises curvature energy is the plectoneme, which appears to be the state adopted by DNA loops with non zero supercoiling. The solenoid, energetically less favourable, is also found in nature in presence of binding proteins, histones, which are of paramount importance for the geometry of DNA in the nucleus of eukariotic cells.

Equipped with these definitions, we are now ready to begin the journey into the models that describe the elastic properties of DNA, to which section 5.2 is dedicated.

5.2 Standard models for DNA elasticity

The mechanical response of DNA to an external force can be characterized to a great detail borrowing the tools of mechanics of a flexible rod. This is quite remarkable, since it means that, in first approximation, we can forget about the molecular structure and in particular about the specific sequence of bases, because the elastic properties of this polymer are, in first approximation, independent of it. The bending properties of DNA can be characterised using a particular length-scale, named the persistence length (A). This length represents the scale of thermal bending fluctuations. It is about 50nm or 150bp. It is also interesting to notice that the persistence length is in general quite smaller than the typical length of a DNA molecule ($> 1000bp$), so that it will be possible to adopt it as the monomer unit of a coarse grained description of DNA.

DNA is a double helix and in section 5.1 we emphasised how this has non trivial consequences with respect to its properties under torsion. This chapter will show that it is possible to define also a twist persistence length C, which sets the scale of torsional fluctuations. A measurement of C is presented in chapter 7.

5.2.1 The Freely Jointed Chain Model (FJC)

The freely jointed chain model is nothing more than another formulation of a random walk in 3D space. A polymer is modelled as a chain of monomer units of length a, and no self-avoidance is taken into account. Every monomer is free to be oriented in any direction and its orientation does not depend on the orientation of the preceding or following monomer.

Call $\vec{t}_i$ the vector indicating the direction of a monomer. A chain of N

monomer will have average length:

$$\langle \vec{R}_0 \rangle = \langle \sum_{i=1}^{N} a\vec{t}_i \rangle = 0 \tag{5.1}$$

which is zero because of the random orientation of each monomer. In contrast to this, the mean squared length of the chain is not zero:

$$\langle \vec{R}_0^2 \rangle = \langle (\sum_{i=1}^{N} a\vec{t}_i)^2 \rangle = a^2(N + \sum_{i,j=1}^{N} \langle \vec{t}_i \vec{t}_j \rangle) = Na^2 \tag{5.2}$$

and it represents the average spread of the chain in 3D space. Here we recover the well known fact that the mean squared distance in a random walk is proportional to the number of steps in the walk. The entropy in the presence of an external force imposing an "end to end" distance $\langle \vec{R}^2 \rangle \ll \langle \vec{R}_0^2 \rangle$ on the extremities can easily be calculated:

$$S(L) = S_0 - 3\frac{k_B T}{2\langle \vec{R}_0^2 \rangle} \langle \vec{R}^2 \rangle \tag{5.3}$$

where S_0 is the entropy of the free chain in the absence of force. The entropy decreases as the square of the "end to end" distance: the number of possible configurations for the random polymer is decreased as it is obliged to stretch and align along the direction of the pulling force. This system does not have any potential energy, thus its free energy is obtained by $\mathcal{F} = -TS$. Supposing that the entropy is independent of force, which is the case for small "end to end" distances, it is possible to write:

$$\vec{F} = -\frac{\partial \mathcal{F}}{\partial \vec{R}} = -3\frac{k_B T}{\langle \vec{R}_0^2 \rangle} \langle \vec{R} \rangle = 3\frac{k_B T}{a} \frac{\langle \vec{R} \rangle}{L_O} \tag{5.4}$$

where $L_0 = Na$ is the maximal extension of the chain. Thus the random polymer opposes an "entropic force" to an external pulling force. This very simple model predicts that for small extension the force is proportional to the "end to end" distance, i.e. that the elastic constant of the "entropic spring" is constant.

The FJC model can also be resolved exactly. Its partition function can be written as:

$$Z = \langle e^{\sum_{i=1}^{N} \frac{Fa\cos\theta_i}{k_B T}} \rangle = \langle \prod_{i=1}^{N} e^{\frac{Fa\cos\theta_i}{k_B T}} \rangle \tag{5.5}$$

where θ_i is the angle formed by the i-th monomer with the direction imposed by the pulling force. The product factorises and each single term has value:

$$\langle e^{\frac{Fa\cos\theta}{k_B T}} \rangle = \frac{1}{2} \int_0^{\pi} \sin\theta e^{\frac{Fa\cos\theta}{k_B T}} = \frac{1}{2} \frac{\sinh(Fa/k_B T)}{Fa/k_B T} \tag{5.6}$$

The partition function is then easily calculated:

$$Z = (\frac{1}{2} \frac{\sinh(Fa/k_B T)}{Fa/k_B T})^N \tag{5.7}$$

and the extension as a function of force can be obtained as:

$$\langle z\rangle = -\frac{\partial \mathcal{F}}{\partial F} = \frac{\partial k_B T \ln(Z)}{\partial F} = L_0\left(\coth\left(\frac{Fa}{k_BT}\right) - \frac{Fa}{k_BT}\right) \overset{Fa/k_BT \ll 1}{\approx} -\frac{1}{3}L_0\frac{k_BT}{Fa} \tag{5.8}$$

The Kuhn length a, which is twice the persistence length, can be adjusted to fit the experimental data and we obtain $a \approx 100$nm, which is consistent with a persistence length of about 50nm. This model does not seem to fit the SM experimental data well. It is however a good approximation of polymers when stretched by a force $F \ll k_BT/a$, and it allows to have an intuitive understanding of the origin of the "entropic force".

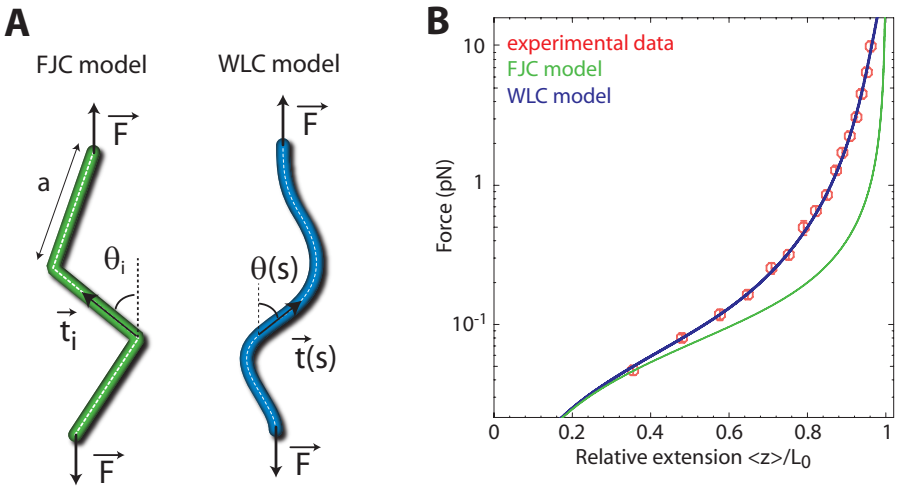


Figure 5.2: RLC and WLC fit of force extension curve (figure from [61])

5.2.2 Worm-like chain model (WLC)

The FJC model can be improved with the introduction of a cost for bending the chain, i.e. a curvature energy. This simple modification allows to describe very precisely the behaviour of DNA subject to an external pulling force. The WLC is defined for a continuum chain, so the relevant variable is the tangent vector $\vec{t}(s)$. The energy of the chain is:

$$\frac{\mathcal{E}}{k_BT} = \frac{1}{2}A\int_0^{L_0}(\frac{\partial \vec{t}}{\partial s})^2 ds - \frac{FA}{k_BT}\int_0^{L_0}\cos(\theta(s))ds \tag{5.9}$$

where the first term is associated to curvature and the second term is the dependence on force. The parameter A is the persistence length, which sets the characteristic length for the decay of correlations between the orientation of tangent vectors along the chain: $<\vec{t}(s)\vec{t}(s')> \approx exp(-|s-s'|/A)$. This means that

the persistence length sets the scale above which bending fluctuations become important.

The WLC model can be solved exactly [66], even if there isn't an analytical relation between force and extension. An approximate relation can however be proposed [14] which is correct to 0.1%:

$$F = \frac{k_B T}{A}\left(\frac{\langle z\rangle}{L_0} - \frac{1}{4} + \frac{1}{4\left(1-\frac{\langle z\rangle}{L_0}\right)^2} + \sum_{i=2}^{7} a_i \left(\frac{\langle z\rangle}{L_0}\right)^i\right) \quad \langle z\rangle/L_0 \overset{\approx}{\ll} 1 \; 3\frac{k_B T}{2A}\frac{\langle z\rangle}{L_0} \tag{5.10}$$

a_i are coefficients whose value is $a_2 = -0.5164228, a_3 = -2.737418, a_4 = -16.07497, a_5 = -38.87607, a_6 = -38.49944, a_7 = -14.17718$, and for small extensions the formula reproduces the result in 5.8 with $a = 2A$ as anticipated. This model is in excellent agreement with experimental data (see figure 5.2). The value of the persistence length A can be fitted and is found to be around 50nm for DNA at physiological conditions, and it is found to vary depending on salt conditions. This model allows to understand why the FJC is not correctly fitting the data. The best fit of the FJC model is obtained with a parameter $a = 2A$, i.e. the chain is considered as a rigid rod on this lengthscale. This underestimates the bending fluctuations, and the FJC sistematically overestimates the DNA length for a given force value.

The WLC model neglects the existence of a structural statistical curvature, due to the local sequence of DNA [27, 66, 71, 9].

5.2.3 High force correction

The WLC describes the entropic elasticity of a DNA molecule in the regime of small forces ($F \leq 10$pN). When the force approaches this value the molecule reaches its contour length and it responds to an increasing force by getting stretched. This effect can be taken into account adding a phenomenological term inspired by Hooke's law of elasticity [85]:

$$\frac{\mathcal{E}_{stretch}}{k_B T} = \int_0^{L_0} \frac{2\pi^2}{p^2} B\epsilon(s)^2 ds \tag{5.11}$$

where $p = 3.6$nm is the helical pitch and $B \approx 78$nm is the stretching modulus and $\epsilon(s) = \frac{\partial u}{\partial s} - 1$ represent the local stretching at the point identified by the coordinate s. We observe that as the molecule is stretched the coordinate s is to be viewed as a parameter, that does not correspond to the position along the molecule axis u any longer. To give an order of magnitude of this effect, the stretching of a base-pair by 10% of its usual extension costs $4k_B T$. This implies that the stretching observed at the scale of a DNA molecule is typically very small. The value of B is thus very difficult to measure directly. This regime describes the elongation of the molecule precisely up to a force of about 65pN.

At this force the molecule undergoes a dramatic phase transition and its length increases of about 70%[24].

5.2.4 Twist behaviour

The response of a molecule of DNA to a stretching force has been characterised in the previous section. In this section we will concentrate on what happens if the linking number of DNA is changed, adding negative or positive supercoiling. In fact, magnetic tweezers allow not only to apply an external force, but also an external fixed linking number constraint, which will incredibly affect the mechanical properties of DNA. The behaviour of DNA under a fixed torsional constraint has been characterised extensively in magnetic tweezers experiments, and this will be summarised below.

Consider a molecule of DNA firmly attached at both extremities. In magnetic tweezers experiments, the linking number of the molecule can be externally imposed through a rotation of the magnets (see section 6.2). This constrains the DNA molecule into a state of either positive or negative supercoiling and it modifies its "end to end" extension in a way that depends on force. The typical shape of the "extension versus σ" curve is the well known "hat curve".

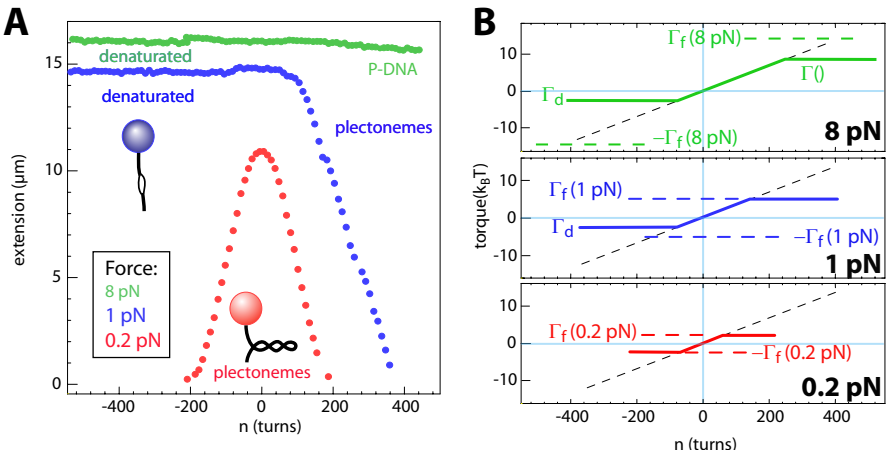


Figure 5.3: Effect of torsion on DNA. A: Extension-Torsion curves. At low force ($F = 0.2$pN, red curve), the molecule buckles and forms plectonemes. The curve is symmetric ($n \to -n$). At intermediate forces ($F = 1$pN, blue curve), negatively supercoiled DNA is denatured, while positively supercoiled DNA still buckles. At high force ($F = 8$pN, green curve), negatively supercoiled DNA is denatured and positively supercoiled is converted to P-DNA (figure adapted with permission from [22])

The shape of the hat curve depends on applied force, and three regimes can be identified. At low forces the hat curves are symmetric for both positive and negative supercoiling. Starting from the maximal extension for the given force, the

addition of supercoiling is absorbed by the molecule as an increase in Twist (see section 5.1). When a sufficient degree of supercoiling is inserted into the molecule, it buckles, supposedly forming tertiary structures known as plectonemes, and its length decreases abruptly (see section 7.1.6). This happens because it is energetically favourable for DNA to absorb the change in linking number as writhe rather than as twist. As more turns are added, the molecule length decreases linearly. A simple model can be used to estimate the energy associated with nucleation of the first plectoneme, modelised as a circle of radius R. The energy is formed by two terms: the first is associated with the cost of bending the DNA molecule to form a circle, the second associated with the cost of reducing its extension against the external force:

$$\mathcal{E}_{pl} = (2\pi R)\frac{1}{2}\frac{k_B T A}{R^2} + 2\pi R F \tag{5.12}$$

For fixed force the radius of the plectoneme is obtained by minimising this energy: $R = \sqrt{A k_B T/2F}$. For a typical force of 0.1pN this radius is about 30nm and the associated energy about 40pN nm = $10 k_B T$. Using this equation, a rough estimate for the buckling torque can be calculated, once the energy associated with torsion has been defined. This can be done as follows. Cosider turning of an angle θ one of the extremities of a molecule of length L, while holding the other firmly. This corresponds to a change in torsion $\Delta Tw = \theta/2\pi$. The energy associated with this torsional change is:

$$\mathcal{E}_{tw} = \frac{k_B T C \theta^2}{2L} \tag{5.13}$$

where $C \approx 100$nm is the torsional persistence length of the chain, which is a measure of the torsional rigidity of the moleucle. As we will shortly see, the torque exerted on a molecule can be defined as $\Gamma := -\partial E_{tw}/\partial\theta$. At the buckling transition the energy associated to the increase in twist must be equal to the energy associate with the formation of a loop :

$$2\pi\Gamma_b = 2\pi\sqrt{2FAk_BT} \tag{5.14}$$

This simple model has recently been challenged in [65], and section 5.3 will introduce the changes proposed. Here we observe that it gives a prediction against which torque measurements can be compared.

5.2.5 Behaviour of extension as a function of force and linking number

As force is increased above 0.5pN the behaviour of the hat curve changes dramatically when adding negative supercoil. The reduction in linking number is not absorbed in plectonemic supercoils but in local denaturation. It becomes energetically favourable to unwind the two strands rather than bending the double helix

to absorb the change in linking number as writhe. The critical torque necessary to open a denaturation bubble can be estimated considering that the opening of a bubble absorbing $\Delta LK = 1$ costs an energy of about $\Delta E_d = 20 k_B T$. This implies a torque $\Gamma_d = \Delta E_d / 2\pi \approx 13$pN.nm for each turn of negative supercoil added to the helix. The buckling torque for plectoneme formation is higher than this value, for forces of about 0.5pN. When the torque reaches this critcal value every additional turn opens a denaturation bubble of about 10.4bp.

For forces > 3pN, a transition is observed also with the addition of positive supercoiling, and almost no reduction of extension is observed for positive supercoling as well. This region corresponds to a change in structure of DNA to a form called P-DNA. The torque associated to this transition has been estimated to be $\Gamma_{P-DNA} \approx 34$pN.nm [18]. When the force is sufficiently high, the torque imposed on DNA reaches this value before the buckling instablity and it becomes energetically favourable for DNA to change its structure rather than to bend its backbone.

5.3 More elaborate models

A general framework for modelling the behaviour of a DNA molecule suject to an external force and a fixed Linking number constraint was proposed by J. Marko in [65]. Here we will summarize the proposal which will be then used in chapter 7 to estimate the torque acting on a DNA molecule.

Suppose that the free energy per length $\mathcal{F}(F, \sigma)$ of the DNA molecule as a function of stretching force (F) and degree of supercoiling (σ) is known. The average extension of the molecule and the average torque applied can be obtained by derivation of $\mathcal{F}$ with respect to the conjugate variables:

$$\frac{\langle z \rangle}{L} = -\frac{\partial \mathcal{F}}{\partial F} \tag{5.15}$$

$$\Gamma = \frac{1}{\omega_0} \frac{\partial \mathcal{F}}{\partial \sigma} \tag{5.16}$$

$$\tag{5.17}$$

where $\omega_0 = 2\pi/3.6$nm is a constant to obtain the value of torque (in fact $\omega_0 \sigma = (2\pi/p)(nL/p) = \theta/L$ i.e. the twist angle normalized by the length of the molecule). These equations are the starting point of tall the work done in chapter 7 so let us look at them in more detail. The first equation was encoutered before and states that the average extension of the molecule is equal to the derivative of the free energy with respect to force. Since force extension curves are well measured and acharacterized by the WLC model, equation 5.15 implies that the free energy $\mathcal{F}(F, \sigma)$ must reduce to the WLC free energy when $\sigma = 0$.The second equation is interesting for two reasons: first of all it formulates a predition on the

dependency of torque as a function of F and σ if $\mathcal{F}(F,\sigma)$ is known; secondly it can be used to estimate the value of torque if $\frac{\partial \mathcal{F}}{\partial \sigma}$ is experimentally measured.

5.3.1 Construction of $\mathcal{F}(F,\sigma)$

The exact form of $\mathcal{F}(F,\sigma)$ is not known. In [65] the author proposes a simple model that allows analytical treatment and thus allows to formulate predictions against which experiments can be compared. The author proposes to consider DNA under the action of force and fixed linking number constraint as a mixture of two distinct phases: one of purely stretched DNA and the other of purely supercoiled DNA. This formulation, which is borrowed from the description of mixed phases for gases, is a first order approximation, and the free energy is written as:

$$\mathcal{F}(\sigma) = x_s \mathcal{S}(\sigma_s) + (1 - x_s)\mathcal{P}(\sigma_p) \tag{5.18}$$

with the condition

$$\sigma = x_s \sigma_s + (1 - x_s)\sigma_p \tag{5.19}$$

x_s is the fraction of DNA that is a purely stretched with free energy $\mathcal{S}(\sigma)$ and $(1 - x_s)$ the fraction of purely plectonemic DNA with free energy $\mathcal{P}(\sigma)$. The construction makes sense when the two pure state free energies cross each other for a certain value of σ. This is the case for DNA, where at low σ the stretched molecule absorbs additional linking number by twisting, and at high σ the stretched molecule buckles to form plectonemes and stretching is absorbed by writhe fluctuations.

Due to the condition 5.19, the free energy of the mixed state in equation 5.18 depends linearly on the parameter σ. This is important because it implies that the extension versus force curve $z/L = -\frac{\partial \mathcal{F}}{\partial f}$ will depend linearly on σ in the mixed phase. This is indeed what we observe in the curves in section **??**. It also implies that torque is constant in the mixed phase.

5.3.2 Explicit forms for the free energies of the pure states

Stretched DNA can be described by the WLC model of section 5.2.2. An approximate formula for the twist contribution to the free energy of stretched DNA at large forces was prposed by Moroz and Nelson [68, 69]:

$$\mathcal{S}_{mn}(\sigma) = \frac{c_s}{2}\sigma^2 + o(\sigma^4) \qquad \text{with} \qquad c_s = c\left[1 - \frac{C}{4A}\sqrt{\frac{k_B T}{AF}}\right] \tag{5.20}$$

where $c = k_B T C \omega_0^2$ and $\omega_0 = 2\pi/p$. This formula gives an elastic form for the free energy dependence on twist, with a coefficient c_s that depends on force. The value of the coefficient c_s decreases with force because excess twist is partially

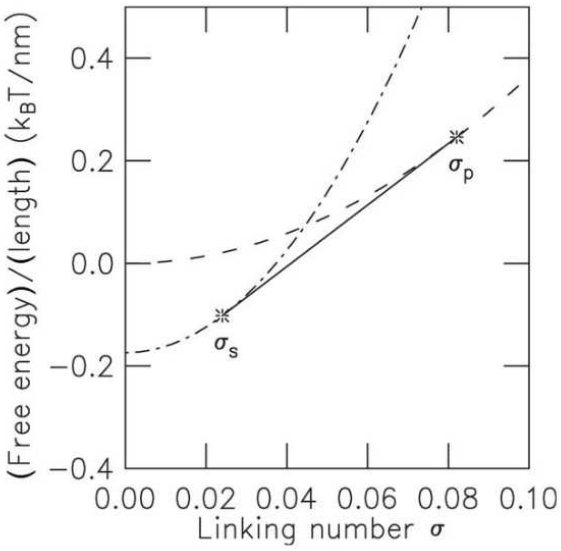


Figure 5.4: Double tangent construction (figure from [65])

absorbed into writhe fluctuations. To describe the stretched twisted DNA, J. Marko thus proposes to use a free energy of the form:

$$\mathcal{S}(F,\sigma) = S_{wlc}(F) + S_{mn}(F,\sigma) = -F + \sqrt{\frac{F}{\alpha}} + \frac{c}{2}\left[1 - \frac{C}{4A}\sqrt{\frac{\alpha}{F}}\right]\sigma^2 \qquad (5.21)$$

This equation contains also an approximate formula for the WLC energy is compatible with formula 5.10 for forces higher than 0.3 pN.

The dependence of the free energy of supercoiled DNA on linking number density is very difficult to calculate exactly. In fact self avoidance becomes relevant in the calculation of the entropy of a plectoneme and only approximate results exist. The form suggested in [65] is the following:

$$\mathcal{P}(\sigma) = \frac{p}{2}\sigma^2 \qquad (5.22)$$

which corresponds to the first term of an empirical series expansion obtained with montecarlo simulation [55, 96].

This formula depends on the empirical parameter $p = k_B T P \omega_0^2$ which corresponds to the twist persistence length of plectonemic DNA. Values for P in the range $20 - 30$nm have been proposed, but a detailed test of the validity of this approach is missing.

Equation 5.22 was obtained from an empirical expansion to second order in σ from Monte Carlo simulations on supercoiled DNA. The inclusion of only the quadratic term in the energy of the plectonemic state is justified by the desire to have an analytically solvable. The effect of this truncation is probably negligible

for small values of σ but it may become important as σ is increased. As we shall see this theory does not correctly predict the variations of DNA extension as a function of σ. The article presented in chapter 7 presents the results of a new set of measurements that cannot be correctly fitted unless terms of higher order in σ are included in formula 5.22. However, the introduction of these higher order terms implies the introduction of additional parameters, whose physical meaning is less clear.

With the caveats just discussed, the free energy of the mixed state can be written explicitly by eliminating the parameters σ_s and x_s

$$\mathcal{F}(F,\sigma) = \begin{cases} -\mathcal{S}_{wlc}(F) + \frac{1}{2}c_s(F)\sigma^2, & |\sigma| < |\sigma_s(F)|, \\ -\frac{\mathcal{S}_{wlc}(F)}{1-p/c_s(F)} + \sqrt{\frac{2p\mathcal{S}_{wlc}(F)}{1-p/c_s(F)}}|\sigma|, & |\sigma_s(F)| < |\sigma| < |\sigma_p(F)|, \\ \frac{1}{2}p\sigma^2, & |\sigma| > |\sigma_p(F)|, \end{cases} \tag{5.23}$$

and the torque dependence on force and degree of supercoiling can also be explicitly calculated as:

$$\tau(F,\sigma) = \begin{cases} c_s(F)/\omega_0, & |\sigma| < |\sigma_s(F)|, \\ 1/\omega_0\sqrt{\frac{2p\mathcal{S}_{wlc}(F)}{1-p/c_s(F)}}, & |\sigma_s(F)| < |\sigma| < |\sigma_p(F)|, \\ p/\omega_0, & |\sigma| > |\sigma_p(F)|, \end{cases} \tag{5.24}$$

It is important to observe that these formulas depend on three free parameters: A (the bending persistence length), C (the twist persistence length) and P (the plectonemic twist persistence length). A appears alone in the formula of the WLC model for the extension of torsionally relaxed DNA, and can be fitted separately. C appears in formula 5.20 in combination with A, in the region of σ less than the buckling transition. We can thus fit a single variable which is a combination of A and C in a fixed ratio. Finally, P appears only in the mixed phase and its value can be fitted once the values of A and C have been independently measured.

5.3.3 Complete model: RLC

Finally a model which includes thermal fluctuations has been proposed in [16]. The model, called Rod like chain (RLC) describes DNA as a semi-flexible polymer, integrating its curvature energy. This model gives a slightly different prediction for the twist persistence length C with respect to the model of Moroz and Nelson [68]. In fact the value of C=86nm [14], which is slightly lower than the C=110nm found in [68].

Finally, Monte Carlo simulations allow to model the behaviour of stretched and twisted DNA [97]. In simulations the effect of self avoidance can be taken into account, which becomes more and more important as writhe is accumulated in the polimer. These simulations gave an estimation for $C = 75 \pm 15$nm.

Chapter 6

Single molecule micro-manipulation techniques and the magnetic tweezer

6.1 Brief summary of Single Molecule pulling techniques

6.1.1 History

DNA micro-manipulation techniques(see [72, 43]) appeared at the beginning of the 1990s [86]. They are all based on the same principle of attaching a DNA molecule to a rigid surface on one side and to a movable element on the other. The movable element both enables the application of an external constraint as force or torque and it acts as a transducer for the measurement. The measured variable is the DNA's extension as deduced from the distance of the transducer to the attachment point at the other end of molecule.

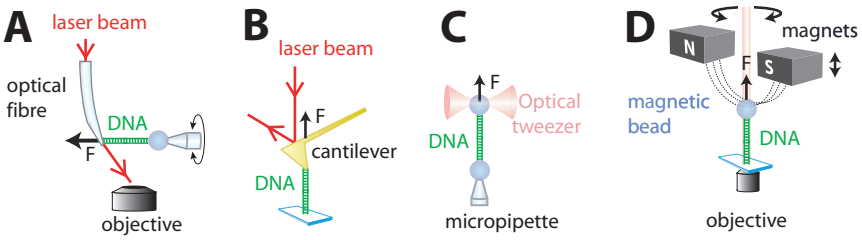


Figure 6.1: DNA micromanipulation methods. A: Optical microfibers, B: Atomic force microscopy, C: Optical Tweezers, D: Magnetic Tweezers. Figure adapted from [61]

6.1.2 Optical microfiber

This technique consists in holding the DNA between a glass micro-pipette at one end and an optical fibre at the other end (see fig 6.1A and [24, 58]). The pipette is then displaced with a micro-metric stage and the DNA molecule is stretched between the fibre and the pipette. The fibre itself will also move, subject to the pulling force, and we can monitor its position by illuminating it with a laser beam and collecting the light with a microscope objective. The applied force will be proportional to the bending rigidity times the displacement of the fibre and we can thus deduce the force. The bending rigidity can be calibrated using an external force of known value such as stokes drag resulting from a flow or using a fiber of known rigidity as a calibration tool. The length of the molecule can be deduced measuring the relative position of pipette and fibre. The spatial resolution of this technique is about 10nm and the force stiffness is between 10^{-2} and 10^{-6} N/m [24, 36].

6.1.3 Atomic Force Microscopy

Atomic force microscopy was first invented to probe surfaces with atomic resolution [12]. Its working principle is similar to that of music vinyl discs: a tiny cantilever is held close to the surface being scanned and a laser beam is shone onto it. The deflection of the cantilever deviates a reflected laser beam from its equilibrium position on a light detector. AFM tips can be coated with different kinds of proteins in order to be able to attach a biological molecule. With this functionalisation, a molecule of DNA or a protein can be suspended between the AFM tip and a rigid surface. As for the micro-pipette technique, displacing the surface with respect to the tip will stretch the molecule and deform the cantilever. Force measurements can be obtained calibrating the deflection against a known force and measuring the stage displacement. Typical AFM cantilevers are about 100 μm long, 5 μm large and about 100 nm thin [95], although even smaller cantilevers have been recently reported [107]. The typical stiffness of AFM varies between 1 N/m and 10^{-3} N/m.

6.1.4 Optical Tweezers

Optical tweezers use light to trap a small particle in three dimensions. They were first proposed in 1978 [4] and realised in 1986 [5]. In an optical tweezers setup, a powerful laser beam is focused to a diffraction limited spot through a high numerical aperture objective. Trapping is due to the spatial gradient in light intensity and acts in the direction of that gradient. A stable trap is obtained when that force is larger than the radiation pressure. A thorough review on optical tweezers is [91], more recent ones are [44, 67, 72, 43]. Optical tweezers have been used to trap bacteria and cells directly [5] or to trap small dielectric

particles which are then used as handles. Micro-spheres are now routinely used in optical tweezers setups. Their surface can be coated in a variety of ways to make them usable as handles to attach biological molecules. Optical tweezers can trap objects as small as 5 nm and can exert forces as large as 100pN with resolutions as fine as 10^{-4}pN (see [44, 67, 72]).

The position of a bead in optical tweezers is confined in all three dimensions, and can be measured in different ways. The first setups used video microscopy and single particle tracking to monitor the position of the bead. This method is affected by a low bandwidth, due to the limited acquisition frequency of the camera. A faster and commonly used solution is to conjugate the position of the trap with a quadrant photodiode. When the bead is in the centre of the trap the voltage is the same on all four quadrants of the photo-diode; for small displacements the voltage difference of two opposite quadrants is proportional to the displacement. Although this second method demands a precise calibration in order to relate the displacement to a voltage difference, it provides a higher bandwidth. The spatial resolution that can be obtained with this method is of about $1A$.

Optical tweezers impose a fixed position to a bead. Upon application of an external force on the trapped bead, it is displaced from its equilibrium. Knowing the trap stiffness and measuring the displacement allows to estimate the applied force.

Recent developments involve simultaneous trapping of many particles as well as the possibility to move the position of the trap in the longitudinal direction through the use of olographic filters (see [44] and references therein). Other possibilities involve the use of polarised light to transfer angular momentum to anisotropic particles, in order to apply a torsional stress to a biological sample. An example of this technique is reviewed in section **??**.

6.2 Magnetic Tweezers

Magnetic tweezers are based on a simple principle: a DNA molecule is attache to a superparamagnetic bead on one side and to a glass surface in an inverted microscope sample on the other side. A couple of permanent magnets placed above the sample generates a magnetic field in the horizontal direction. The bead is simultaneously subject a vertical pulling force due to the gradient of the magnetic field and to a torque that fixes the orientation of its magnetic moment in the horizontal direction of the magnetic field (see section and [42] for a quantitative description). The magnets poles are separated by a fraction of mm, which determines the scale of variation of the resulting magnetic field. This implies that the field is essentially uniform at the scale of the bead (μm) and the force is thus constant. This can be varied by varying the distance of the magnets

from the sample. A torsional constraint can be imposed changing thedirection of the magnetic field. The bead acts both as force transducer and as position sensor. In fact, its position is monitored in realtime through a microscope (typically at a frequency of 60Hz) and the force is calculated from the Brownian movement of the bead (see section 6.2.7 for a detailed explanation of the measurement principle principle). Typical resolution for the position measurement is 1-2 nm, while the resolution for the force depends on the acquisition time and on the bandwidth of the camera. The force range is between 0.01 pN and 100 pN, depending on the bead size.

The first use of magnetic beads to manipulate DNA dates back to 1992 [86], when the group of C. Bustamante used them to obtain the first measurement of the elasticity of a single DNA molecule. Their technique used a couple of magnets to hold the bead and a viscous flow to stretch the molecule. Since then, the technique has been pursued by different experimental groups, leading to the development of an apparatus that goes under the name of magnetic tweezers [88].

6.2.1 Description

The working principle of the static magnetic tweezers is extensively described in [87, 42, 61]. A part of my PhD thesis consisted in the improvement of this device to allow for torque measurement. I will detail here some technical information to introduce the reader to the standard magnetic tweezers apparatus. This will be helpful in the chapters 7 and 8, when the torque measurements will be presented.

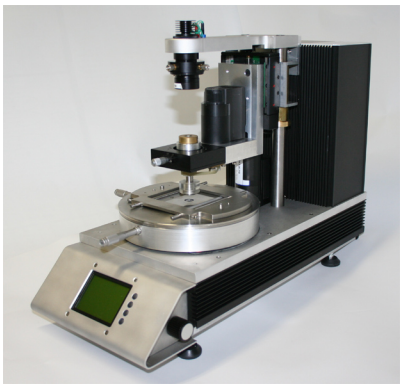


Figure 6.2: Ensemble view of the Magnetic Tweezers apparatus developed by PicoTwist (picture courtesy of www.picotwist.com)

The first magnetic tweezers were built adapting commercial inverted microscopes with the addition of magnets suspended above the sample. Experience showed that the main problem with commercial devices were thermal drifts of the microscope. These pose a serious problem, since the measurement of force in magnetic tweezers relies on the precise measurement of the molecule length.

This lead to the design custom built device (currently commercialized by Picotwist (www.picotwist.com)). The main difference of this device with respect to a commercial microscope, consists in a careful choice of the materials for the support of the objective and sample that allows for the compensation of thermal drifts.

6.2.2 Sample

The chamber where the DNA will be manipulated is made using a glass microscope cover-slides. One of them has two holes, 1mm in diameter, that are pierced using a sand blower. These two slides are held together to form a chamber by a piece of double sided adhesive tape or Parafilm (Alcan Packaging, Canada), in which a channel 4 mm wide and 40 mm long is cut. The union of the three parts yields a sealed chamber with an entry and an exit. Toroidal rubber joints are used to connect the chamber thus prepared to a plexiglass reservoir and a rubber tube that allows to exchange buffers and to infuse enzyme and other reagents into the micro-channel (see figure 6.3). The device is easily prepared and can be used for a month. Since the strength of the pulling force depends on the distance between the magnets and the beads, it is of the greatest importance to reduce the height of the chamber to the minimum. For high force applications, a different chamber can be prepared, in which the top part is obtained using a stretched mylar film. Its thickness of $50\mu m$ is much smaller than the 170 microns cover-slides, allowing for even higher forces to be reached without increasing the magnetic bead size. In fact, since the force is measured using Brownian fluctuations of the bead, increasing the size of the bead increases the time needed to average out Brownian fluctuations. This chamber imposes a minimal bead/magnets separation of $100\mu m$ allowing to exert forces up to $15pN$ for a $1\mu m$ bead (MyOne, Dynal).

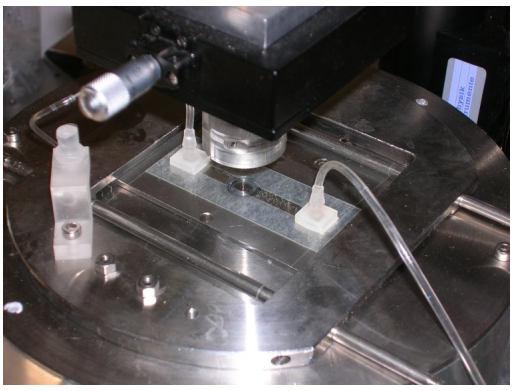


Figure 6.3: The magnetic tweezers slide (picture courtesy of www.picotwist.com)

6.2.3 Thermalised stage

The sample chamber is held above the microscope objective by a thermally stabilised stage which encloses the objective and which allows for horizontal micrometric displacement of the sample. Thermal stabilisation is achieved with two Peltier cells that connect this piece with the main body of the device, allowing for a stabilisation within 0.01 degree. However, this is not enough to completely eliminate the drifts due to thermal dilatation, which are of the order of 1 μm per $^{\circ}C$. Thus, to further increase the position stability, a compensation system is used: the objective is mounted on a ring that compensates for its dilatation, so that the distance between the focal distance is kept constant at the nanometre scale. The couple objective-compensation ring is placed on a piezoelectric holder with capacitive closed-loop position control. This allows to set the position of the microscope objective with respect to the sample plane with nanometre precision. This is used to calibrate the images for position tracking along the axial direction.

6.2.4 Illumination

The illumination of our device must be optimized to form diffraction rings around the beads that are used to determine the longitudinal (z) distance of the bead from the focal plane. The light must be sufficiently coherent to produce nice rings around the beads but not too coherent to produce diffraction patterns around distant features like dust particles on the upper surface. A convenient choice consists in focalising the light coming from a red Light Emission Diode (LED) and in imaging the focal plane on a CCD camera (JAI CV-M30) with 60 Hz frame rate.

6.2.5 Realtime tracking

The images acquired on the camera are used to track the position of the bead in 3D in realtime, and exploit the diffraction rings formed around the bead [73]. The centre of the bead on the XY plane is first obtained. Then the rings are used to obtain the Z coordinate. In fact, the further the beads are from the focal plane, the larger are the diffraction rings. It is then possible to calibrate the changes in the size of the diffraction rings against known distances and then use an inverse algorithm to deduce the Z position from the rings size. It is necessary to repeat a calibration procedure when a new bead is found, which consists of stepping with the piezoelectric holder in regular steps and acquiring some images at each step. The images are then averaged and filtered and finally the radial profile of the rings is measured. This gives us a function that correlates Z positions with a particular radial profile. It is then sufficient to calculate the profile of the rings at each detected frame, and use an inverse procedure to deduce the relative distance between the bead and the focal plane of the objective. Since the objective and

the sample stage are held at rest, this gives a measurement of the position of the bead in the longitudinal direction.

Finally, the process can be further improved tracking simultaneously two beads, one of which is firmly attached to the glass surface. The position of the two beads can thus be simultaneously measured allowing us to essentially eliminate any drift of the mechanical parts of the microscope.

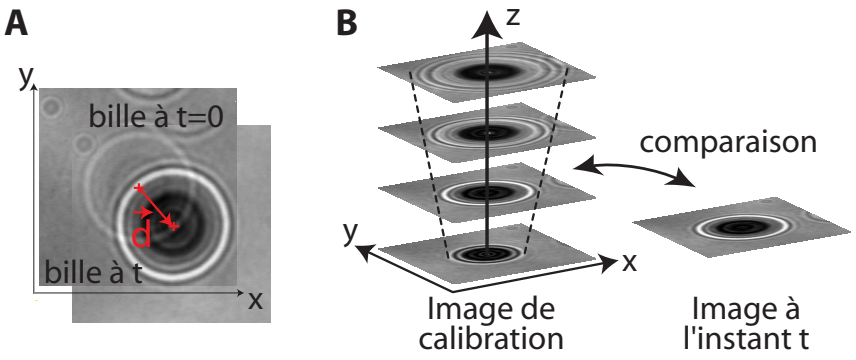


Figure 6.4: Position tracking principle: A: Measure in the XY plane. the position of the centroid of the diffraction rings is recorded for every video frame. By correlating the intensity radial profile at time t with that at time t=0 we obtain the displacement of the bead in the XY plane. B: Measurement along the optical axis. Before each experiment the relative distance of the bead from the focal plane is varied while keeping the bead at a fixed position and changing the position of the microscope objective. A set of calibration images is obtained, by recording an image at each objective position. During the experiment, the bead image is compared in real-time to the calibration images allowing to reconstruct the longitudinal position of the bead

6.2.6 Force application

A couple of permanent magnets (NeFeB) is positioned above the sample with the opposite poles facing each other and a gap smaller than a mm in between. The generated magnetic field is horizontally oriented and its value is around 1 Tesla at the centre of the gap. Its strength decreases almost exponentially along the optical axis z. The size of the gap sets the typical lengthscale for the exponential decrease.

The magnetic beads used are made of a super-paramagnetic material, with very high magnetic susceptibility and almost zero permanent magnetisation. In practice, the beads are composed of a mix of ferromagnetic nanocrystals, suspended in a non-magnetic matrix [74]. These nano-domains are small enough to lose their magnetisation under the effect of thermal fluctuations. When a bead is immersed in a magnetic field, it acquires a magnetic moment com-

posed of two contributions: the first, most important (of order $10^{-11} e.m.u. = 10^{-14} A.m^2$) is aligned along the magnetic field. The second, much weaker (of order $10^{-13} e.m.u. = 10^{-16} A.m^2$) , is aligned along a direction of preferential magnetisation, which is fixed with respect to the bead and is due to a slight anisotropy in the distribution and orientation of the magnetic nano-particles. The anisotropic contribution creates a torque, of order $10^4 pN.nm$ that tents to align the direction of preferential magnetisation with the external magnetic field. In the classical magnetic tweezers here described, this torque exceeds by at least two orders of magnitude all biologically generated torques. This means that beads will always align with the external magnetic field.

The magnetisation curve of the magnetic beads (MyOne, Dynal), follows a classical Langevin law. The interaction between the magnetic dipole formed $\vec{\mu}$ and the magnetic field $\vec{B}$ generates a force $\vec{f} = (\vec{\mu}\vec{\nabla})\vec{B}$ and a torque $\vec{\Gamma} = \vec{\mu} \times \vec{B}$. As we have mentioned, the magnetic torque tends to align the direction of the preferential axis of the bead with the magnetic field, while the force is vertical and tends to attract the bead towards the magnets. We can thus estimate quite simply the force applied to the bead by measuring the distance of the magnets from the beads and using the magnetisation curve given by the producer. However, this estimation is not sufficient to determine the force on the bead accurately, due to small dispersion in bead size and magnetization, of the order of 15%. We are thus obliged to measure directly the force. To that purpose we use the Brownian fluctuations of the bead and the theorem of equipartition of energy.

6.2.7 Force Measurement: Time Analysis

Let's first have a look at the movement of the bead. The tethered bead is subject to the external magnetic force, to the tension of the DNA molecule and to thermal shocks from Langevin force F_L. The viscous drag, in the approximation that the bead is far from the surface is given by the Stokes formula: $F_v = 6\pi\eta r v$ ([91, 82]). Under the action of the stretching force, the DNA molecule behaves like a spring that tends to recall the bead to its equilibrium position. Let's look at fluctuations in the XY plane. Consider a small displacement δx of the bead from the vertical of the attachment point on the surface: the stretched DNA molecule will react with a force $\delta F = F\delta x/\langle z \rangle$ along the x direction, because the tension T of the DNA molecule will tend to equilibrate the pulling force. We can thus write the following equation for the movement of the bead along the x axis:

$$k_x x + 6\pi\eta r \frac{\partial x}{\partial t} = F_{L,x} \tag{6.1}$$

Where $k_x = F/\langle z \rangle$ is the spring constant of our system and the mass term can be neglected because it's negligible with respect to the other terms. In other words we are considering a Low Reynolds number regime.

The theorem of the equipartition of energy establishes that the mean value of the energy for each degree of freedom is equal to $\frac{1}{2}k_BT$. If we apply to the position x of the bead we obtain:

$$\frac{1}{2}k_BT = \frac{1}{2}k_x\langle\delta x^2\rangle = \frac{1}{2}\frac{F\langle\delta x^2\rangle}{\langle z\rangle} \qquad \rightarrow \qquad F = \frac{k_BT\langle z\rangle}{\langle\delta x^2\rangle} \tag{6.2}$$

This tells us that we will be able to measure forces by measuring Brownian fluctuations in the transverse plane and the average length of the molecule. This also tells us that the higher the force, the smaller the fluctuations, and vice-versa. To estimate the statistical averages $\langle z\rangle$ and $\langle\delta x^2\rangle$ and eliminate drifts and other systematic problems it is advantageous to work in Fourier space.

6.2.8 Force Measurement: Fourier analysis

Equation 6.1 becomes:

$$(k_x + i12\pi^2\eta r f)\hat{x} = \hat{F}_{L,x} \tag{6.3}$$

where f is the frequency in Hz and " $\hat{.}$ " indicates the Fourier transform. The fluctuations power spectrum is:

$$|\hat{x}^2(f)| = \frac{12\pi^2 k_BT\eta r}{k_x^2}\frac{1}{1+(f/f_0)} \tag{6.4}$$

where we used the fact that the Fourier transform of a convolution is a product, and since the correlation function of the Langevin force is a delta function, its Fourier spectrum is a constant white noise of amplitude: $|F_{L,x}^2| = 12\pi\eta r k_BT$. The spectrum of fluctuations is a Lorentzian with a cut-off frequency $f_0 = k_x/12\pi^2\eta r = F/(12\pi^2\eta r z)$. Above this frequency, fluctuations are damped by the viscous drag of water, which implies that it is not necessary to have an infinite bandwidth in order to measure cluctuations correctly. In fact, it is sufficient to measure the position of the bead with a frequency f_{acq} large with respect to f_0 in order to correctly sample the signal, because the contribution of frequencies higher than f_0 is negligible. Nevertheless, when the cut-off frequencyis of the order of the Nyquist frequency $f_{\text{Nyquist}} = f_{acq}/2$ or larger it is impossible to measure correctly the force exerted on the bead [77].

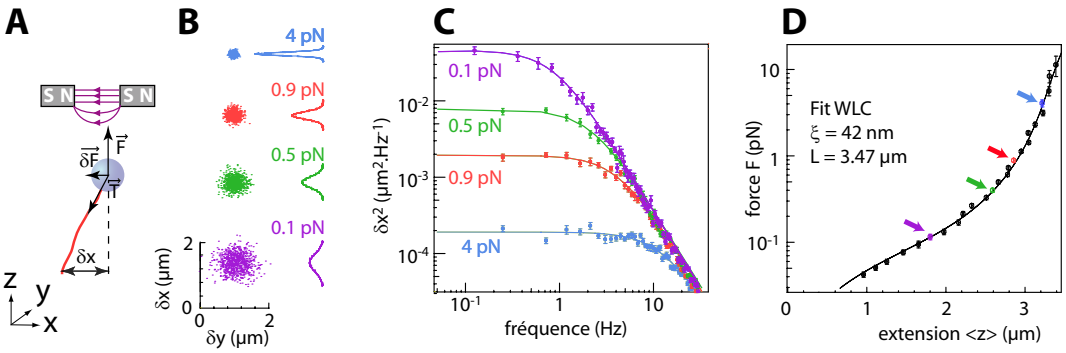


Figure 6.5: Measurement of force using transverse fluctuations. A: the DNA-bead system. B,C,D: experimental data. B Transverse fluctuations of the bead. C: Spectrum of fluctuations of transverse position of the bead. D: Force extensio curve for a double stranded DNA molecule. (figure adapted from [61])

Chapter 7

Torque measurement with standard magnetic tweezers

Measurement of torque with magnetic tweezers

Francesco Mosconi,* Jean Franqis Allemand, David Bensimon, and Vincent Croquette†
LPS, ENS, UMR 8550 CNRS, 24 rue Lhomond, 75231 Paris Cedex 05, France
(Dated: July 31, 2008)

We measured the extension change of a single DNA molecule trapped in a magnetic tweezer for different values of supercoiling density and in different ionic conditions. Using the method proposed in **??** we estimated the value of torque applied on the molecule in previously unprobed small force regime.

PACS numbers:
Keywords:

In all living organisms DNA is twisted. This is the result of the activity of enzymes, named topoisomerases, that regulate the torsional state of DNA. A torque in DNA is then required for DNA's biological activity. Torque does not only affect DNA elasticity and structure [1**?** –3] but also many cellular processes are influenced by DNA supercoiling, or a torque in the molecule: DNA replication, DNA transcription[4] and repair [5], DNA compaction [6] and chromosomal condensation[7]

Experimentally torque in a DNA molecule can be controled by modifying its linking number, i.e. the number times of the two DNA strands are intertwined [8, 9]. Nevertheless, because of DNA elastic properties, the linking number is absorbed bt DNA partly as pure twist and partly as writhe. Since only the twist contributes to torque, for a given linking number, torque will depend on the possiblity for the molecule to bend on itself to accumulate writhe, and thus it will depend on the force applied on the molecule. For this reason, and because they allow a simpler way to modify the linking number than classical biochemical experiments, micromanipulation experiments provide a more direct approach to investigate the role torque in the structure of a DNA molecule [1–3, 10] and its role in DNA/protein interactions [11–23]. In particular the magnetic tweezers, where a DNA molecule is attached between a glass surface and a magnetic bead, has proven to be a simple way to induce a torque in a DNA molecule. In these experiments, a magnetic field control the pulling force on the molecule but also the variations of the linking number. As mentioned above, the torque is not a simple combination of these parameters and thus its measurement is not direct. This limitation can be overcome by other techniques. Bryant et al.[2] have used the rotational drag of a small bead bound to a nicked DNA molecule (a molecule with a break in one of its two strands allowing the relaxation of the torque) to deduce the torque applied to the segment of DNA that could be twisted. Deufel et al.[24] used polarized light to apply a control torque on anisotropic particles trapped in an optical tweezers. The quite complex experimental approaches required for these measurements, as well as their difficulty to explore low (< 1pN) force ranges, call for new torque measurement strategies that are both simple to implement and suitable to low force regimes. The method exposed in this paper meets both requirements, since it uses a standard magnetic trap set-up that is technically simpler and better suited to measure low forces. In the following we will show that this set-up can be used to determine torque on DNA (albeit not via the angular fluctuations of the trapped bead), allowing one to obtain an independent estimation of biologically relevant mechanical parameters, such as the critical torque for denaturation and the DNA twist modulus C.

This torque measurement method is suggested in a recent pubblication by Zhang et al.[25]. The authors noticed that the free energy $\mathcal{F}$ of a twisted and stretched DNA molecule can generally be written as a function of both controlled variables: the force F and the number of turns n (or total rotational angle: $\theta = 2\pi n$) as $\mathcal{F}(F, \theta)$. Since the mean extension of the molecule at a given force is: $l = -\partial\mathcal{F}/\partial F$ and the mean torque is $\tau = \partial\mathcal{F}/\partial\theta$, one readily derives an expression for the mean torque at a given force $\tau(F, \theta)$ from a measurement of the decrease in extension with increased coiling $\partial l/\partial\theta$:

$$\tau(F,\theta) = \tau(F_0,\theta) - \int_{F_0}^{F} dF' \left(\frac{\partial l}{\partial \theta}\right)_{F'} \qquad (1)$$

Since the angular rotation is known and the force and change in extension with rotation are easily measurable, the determination of the torque difference is reduced to a problem of sampling these variables finely enough to estimate the above integral with sufficient precision. The integration constant $\tau(F_0, \theta)$ is a priory unkonwn, but can be determined or measured in different ways. As we will see below, our procedure allows to estimate this constant by using the observation that $\tau(F_0, \theta)$ varies linearly with θ (before the buckling transition) and the requirement that $\tau(F_0, 0) = 0$.

We measured the extension of a DNA molecule (of total length $l_0 = 17.8kb \approx 6\mu$m) in various salt conditions and for different values of force F and degree of supercoiling σ ($= nh/l_0$, where $h = 3.6$nm is the DNA pitch), see (Fig.1). Such measurements have been described before[1, 26]. Briefly at low forces (F less than about

0.4pN) the curves are symmetric. The extension is maximal at $\sigma = 0$ and decreases non-linearly at small values of σ due to twist fluctuations[27, 28]. Past the buckling threshold (at $\sigma = \sigma_s$), the molecule coils on itself with a constant slope ($\partial l/\partial\sigma$) to form tertiary structures known as plectonemes or supercoils. While below buckling (i.e. when $\sigma < \sigma_s$) the torque increases with increased rotation, it is constant in the plectonemic regime[29] (when $\sigma > \sigma_s$). For larger forces the curve becomes asymmetric as DNA denatures for negative supercoilings at a critical torque [2**?**](i.e. when $\sigma < -\sigma_d$). For this reason we have estimated the torque only for positive degrees of supercoiling, although this estimate should also be valid at low forces for negative supercoilings.

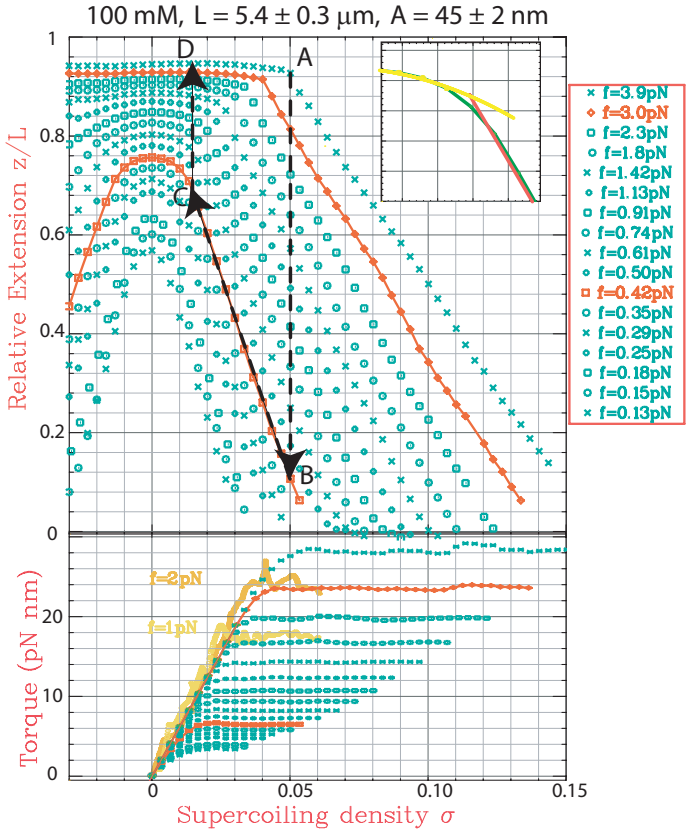


FIG. 1: (upper) Variation of the mean relative extension of a DNA molecule l/l_0 as a function of the degree of supercoiling $\sigma = nh/l_0$ (with $h = 3.6$nm the DNA pitch) for the 100mM NaCl series. The torque at point D is obtained by integrating Eq.1 from A to B and from C to D and taking into account that the torque in the plectonemic regime at fixed force (i.e. between B and C) is constant. Lower part shows the reconstructed torque curves.

One difficulty in the estimation of torque from Eq.1 is that derivating experimental data results in a very noisy signal. To overcome this problem we tested two different strategies (see below) to obtain a smooth determination of the derivative $\frac{\partial l}{\partial\theta}$ at each measured value of force and sigma. The first strategy consisted in fitting each hat curve to a quadratic polynomial in σ when $-\sigma_d < \sigma < \sigma_s$ and to a linear polynomial when $\sigma > \sigma_s$.

$$l(\sigma) = \begin{cases} b_0 + b_2\sigma^2 & -\sigma_d < \sigma < \sigma_s \\ a_0 + a_1\sigma & \sigma > \sigma_s \end{cases} \tag{2}$$

The value of σ_s is then defined as the intersection point of the two curves. To compute the torque at various forces and degrees of supercoiling, we then start from the highest stretching force ($F_{max} = 3.9$ pN in the series shown in fig 1) and the highest torque state investigated, namely point A at the buckling transition ($\sigma = \sigma_{s,max}$) in Fig.1. The absolute value of torque in A is unknown and it will play the role of the constant $\tau(F_0, \theta)$ in Eq. 1. We then compute the change in torque at point D (for which $\sigma < \sigma_{s,max}$), by integrating Eq.1 along path AB (a path of constant σ but varying force, see Fig.1) from F_{max} to F_B. Taking into account the fact that the torques at points C and B are equal, we can then calculate the torque difference along the path CD from force $F_C = F_B$ to force F_D along the path of costant sigma $\sigma_C = \sigma_D$, and substract it from the previous quantity to obtain the torque difference $\Delta\tau(F_D, \sigma_D) = \tau_A(F_{max}, \sigma_{max}) - \tau_D(F_D, \sigma_D)$. Having recorded a large set of curves we can repeat the procedure extending the integration to lower and lower values of force and σ. Notice that due to the choice of fitting procedure, the first integral involves integrating only on the slopes of the linear parts, while the second integral only concerns integrating on the derivative of the curved parts:

$$\int_{ABCD} \frac{\partial l}{\partial\theta} df = \int_{F_A}^{F_B=F_C} a_1(F)dF + \int_{F_B=F_C}^{F_D} 2b_2(F)\sigma_s(F_B)dF \tag{3}$$

With this procedure it is important to chose the value of F_B so that the buckling point $\sigma_s(F_B)$ corresponds to the desired point σ_D. This method allows the estimation of τVSσ curves at different forces and the results of this procedure are shown in Fig.2.

This strategy obliges a quadratic fit on the top part of the hat curve, and expecially at low salt the fitted hat curve and the real data do not perfectly match at the buckling transition (see inset in Fig 1). Furthermore, at low force where the curvature of the hat is large, it was not always possible to determine an intersection of the fitted second order polynomial and the fitted straight line. We thus devised a second strategy to obtain the value of the local derivative at each data point, without assumptions on the global shape of the hat curve. We fitted a polynomial of order 2 using the 2 nearest neighobours on each side and taking the derivative of the fitted curve as a good estimation of the local derivative. This procedure is called Savitzky-Golay derivative smoothing **??** and it

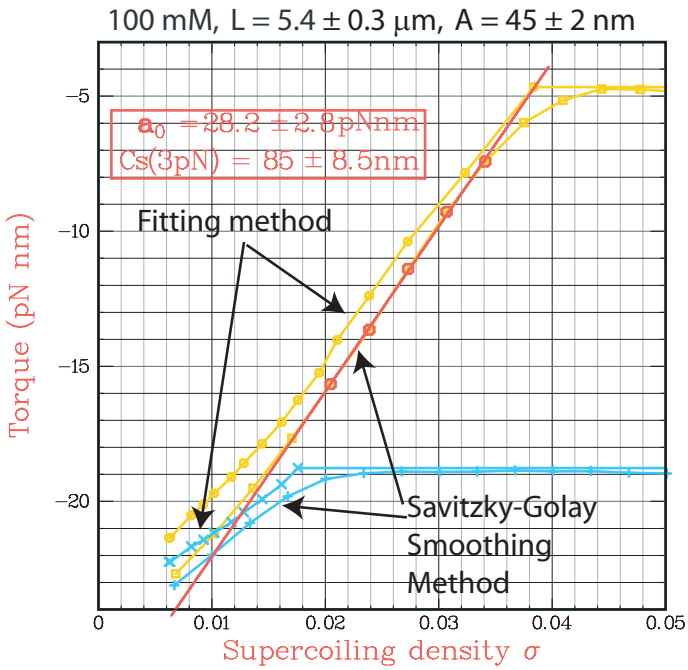


FIG. 2: Comparison of the two methods used to obtain a smooth derivative of the extension versus supercoiling curves ($F = 3.0$pN and $F = 0.74$pN are shown) . The two methods yield very similar results, and are both affected by an overestimation of torque in the low σ region. The red line shows a linear fit to the upper part of the reconstructed curve. This linear fit will be used to determine the unknown offset and to correct the low σ region.

yields a numerically calculated local derivative which is more well behaved and less affected by noise due to discrete sampling. This can be used directly to calculate the integral in equation 1. Figure 2 shows some curves obtained with this method and the two methods agree withing ±1pN.

The methods described yield torque differences and the only unknown, the torque at point A, can be determined by the requirement that $\tau(F, 0) = 0$. We decided to estimate this by fitting a straight line to the top part of the angular variation of torque (its dependence on σ) for the largest values of force calculated. We observe that the low σ part of the obtained curves, seems to depart from the expected linear behaviour [29], tending to saturate towards a higher value. We thus decided to proceed as follows: we reconstructed the angular dependence of torque for the highest available force before the transition to P-DNA and took it as reference. The low sigma part of the curve was rejected, and replaced with the line fitted on the high sigma part of the angular dependence before the buckling transition. The obtained patched curve was taken as reference curve and all the other curves of torque versus σ were obtained in the usual way by integrating on smoothed derivatives as previously described. This procedure yields similar curves and does not alter the value of critical torque (see below), but it corrects what we think is a sistematic overestimation of torque at low values of sigma. The curves displayed in the lower part of figure 1 are obtained with this method and they qualitatively agree very well with previously measured data [30]. There seem to be an difference in offset (e.g. our torque curve at 3pN agrees well with their data at 2pN), which could indicate that our estimation of the offset is not correct (more on this below), or a problem in the measure of force.

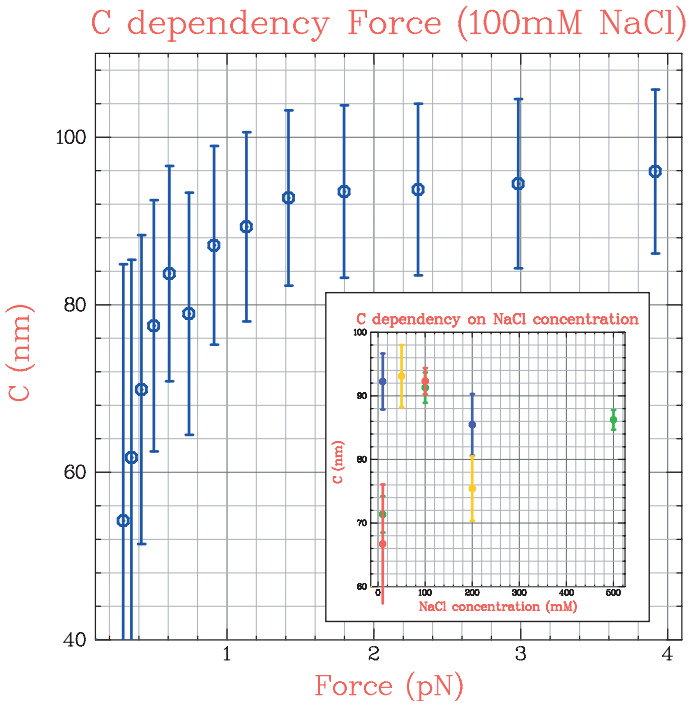


FIG. 3: Variation of the torsional persistence length of DNA with force (data shown for the 100mM series. Inset: dependence of average C (1pN< F < 3pN) on ionic conditions.

The slope of these curves yields the value of the effective DNA twist persistence length as a function of force, after appropriate constants are taken into account. Nelson and Moroz [31] predicted the behavious of C_s as a function of force to be (see [29] for the particular form used here):

$$C_s = C\left[1 - \frac{C}{4A}\sqrt{\frac{k_B T}{AF}}\right] \qquad (4)$$

We invert this formua and plot the value of C obtained from C_s as a function of force in Fig.3. While C remains roughly constat for forces higher than 1pN it appears to dramatically decresase at lower forces, even if Moroz-Nelson correction is taken into account. We cannot tell if this is due to the procedure used to estimate C or if it points to the interesting discovery that the twist persistence length at very low forces becomes shorter than

expected since no other measures exist at low values of force. However we observe that error bars at low force are larger because of the smaller number of points used to determine C as well as the larger impact of numerical errors in the integration procedure. We also notice that the values of C obtained by a quadratic fit of the top part of the hat curves are systematically higher than those estimated with our method and this discrepacy remains to be explored in greater detail. Finally, since different data series have been measured, we decided to average the values of C obtained at forcew between 1pN and 3pN and plotted each value obtained as a function of $NaCl$ concentration (Inset in Fig. 3).

The present approach allows us to compute the buckling torque of a DNA molecule as a function of force in various salt conditions. The results are shown in Fig.4. These results do not seem agree with recent direct measurement using optical traps[30], in particular, the critical torques obtained with our estimation method seem to be sistematically lower than those directly measured with optical traps. This contrasts with the fact that the angular decrease of extension agrees quite well (see below, and Fig. 5). We observe that there appear to be an offset difference of $8-10$pNnm between the data of [30] and our series recorded at similar salt conditions. Since we only have direct access to torque differences, it is possible that our procedure to estimate the offset is not correct. However, we observe that arbitrarily introducing this offset in our data would yield unacceptably high values of critical torques at very low force.

A simple model of DNA buckling[**?**] suggests that the buckling torque scales as $\sqrt{AF/k_BT}$ (where A is the DNA's persistence length). The results shown in Fig.4 are systematically lower than the values obtained with this simple model. For the same value of force, the buckling torque decreases with increasing salt as expected from the observed decrease in the DNA persistence length from $56 \pm 2nm$ at low salt to $40 \pm 2nm$ at high salt.

J.Marko has recently suggested a heuristic theory to describe the behavior of a stretched DNA molecule under twist. In his model, DNA molecules in the plectonemic regime partition between an unstretched plectonemic supercoil phase with torsional stiffness P and a stretched and twisted DNA molecule with persistence length A and an effective torsional rigidity C_s (that varies slightly with force[31]). The model of J.Marko has only three parameters: the DNA persistence length A and torsional rigidity C and the plectonemic torsional stiffness P. It makes a number of predictions on the variation of the extension with the degree of supercoiling and the variation of the buckling torque with force that can be compared with experiments. The predictions of Marko's model are in qualitative agreement with our measurements of critical torque, as shown in the inset of Fig.4. However, Fig.5 shows that they cannot fit with only three parameters all of our results. Furthermore, the value of P tends to become very low (P=10nm at 200nM $NaCl$) when high

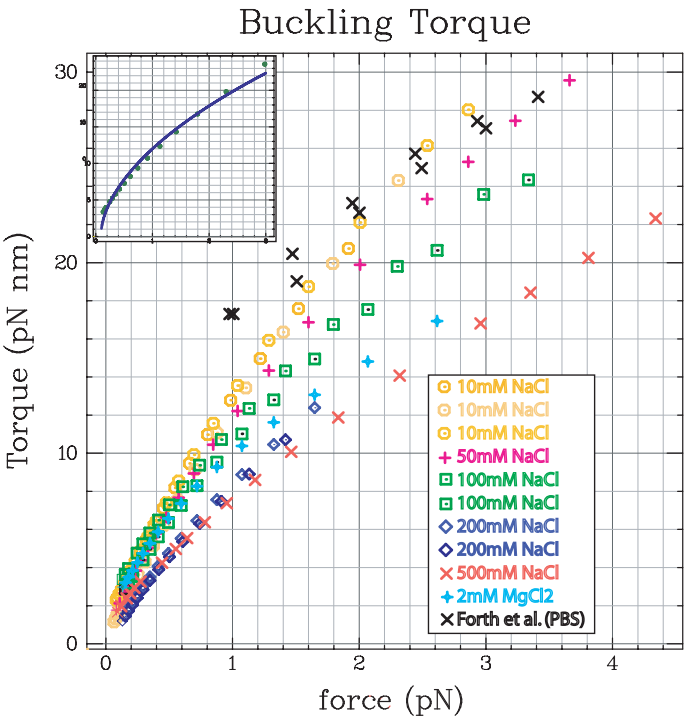


FIG. 4: Buckling torque for different buffer conditions (10mM, 50mM, 100mM, 200mM, 500mM $NaCl$, 2mM $MgCl_2$). Also shown are the results of ref.[30] using optical tweezers.

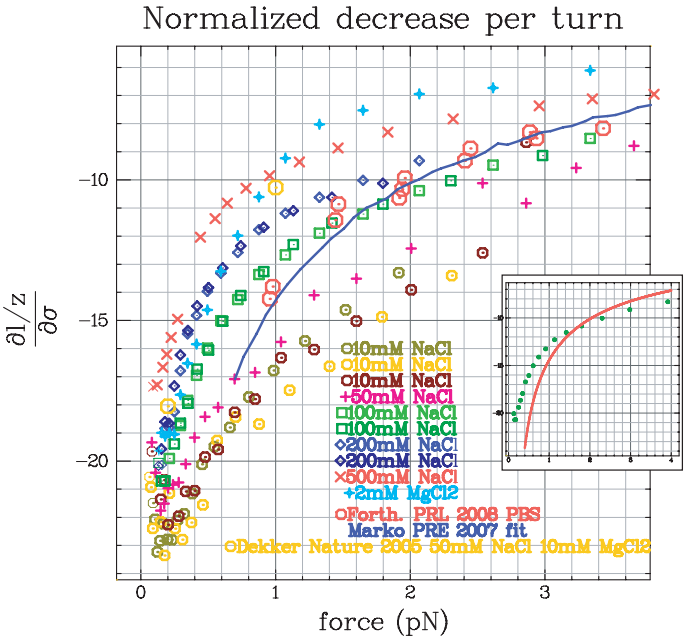


FIG. 5: Angular variation of extension as a function of force. Data agree well with previous publication, but seems to depart from current models

salt series are fitted, and the curve departs from our data at forces lower than 1pN. Also it has proved impossible to fit the angular variation of the extension in the plectonemic regime with a single value of P. The known variation of the plectonemic diameter with force possibly affects the estimate of the torsional energy of the plectonemic supercoiled phase.

* mosconi@lps.ens.fr
† www.lps.ens.fr/biophys

[1] T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, Science (New York, N.Y.) **271**, 1835 (1996), ISSN 00368075, pMID: 8596951.

[2] Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante, Nature **424**, 338 (2003), ISSN 0028-0836, URL http://dx.doi.org/10.1038/nature01810.

[3] J. F. Leger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko, Physical Review Letters **83**, 1066 (1999), copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org, URL http://link.aps.org/abstract/PRL/v83/p1066.

[4] H.-Y. Wu, S. Shyy, J. C. Wang, and L. F. Liu, Cell **53**, 433 (1988), URL http://www.sciencedirect.com/science?_ob=ArticleURL\&_udi=B6WSN-4GB6N6G-B6\&_user=10\&_coverDate=05%2F06%2F1988

[5] A. Stasiak and E. D. Capua, Nature **299**, 185 (1982), ISSN 00280836, pMID: 7050731.

[6] R. H. Morse and R. T. Simpson, Cell **54**, 285 (1988).

[7] R. D. Kornberg and Y. Lorch, Cell **98**, 285 (1999), ISSN 0092-8674, pMID: 10458604, URL http://www.ncbi.nlm.nih.gov/pubmed/10458604.

[8] J. Vinograd, J. Lebowitz, R. Radloff, R. Watson, and P. Laipis, Proceedings of the National Academy of Sciences of the United States of America **53** (1965).

[9] J. H. White, American Journal of Mathematics **91**, 693 (1969).

[10] A. Dawid, F. Guillemot, C. Breme, V. Croquette, and F. Heslot, Physical Review Letters **96**, 188102 (2006), URL http://link.aps.org/abstract/PRL/v96/e188102.

[11] T. R. Strick, V. Croquette, and D. Bensimon, Nature **404**, 901 (2000), URL http://dx.doi.org/10.1038/35009144.

[12] M. D. Stone, Z. Bryant, N. J. Crisona, S. B. Smith, A. Vologodskii, C. Bustamante, and N. R. Cozzarelli, Proceedings of the National Academy of Sciences of the United States of America **100**, 8654 (2003), ISSN 0027-8424, pMID: 12857958, URL http://www.ncbi.nlm.nih.gov/pubmed/12857958.

[13] A. Revyakin, C. Liu, R. H. Ebright, and T. R. Strick, Science (New York, N.Y.) **314**, 1139 (2006), ISSN 1095-9203, pMID: 17110577, URL http://www.ncbi.nlm.nih.gov/pubmed/17110577.

[14] A. Revyakin, R. H. Ebright, and T. R. Strick, Proceedings of the National Academy of Sciences of the United States of America **101**, 4776 (2004), ISSN 0027-8424, pMID: 15037753, URL http://www.ncbi.nlm.nih.gov/pubmed/15037753.

[15] R. Seidel, J. van Noort, C. van der Scheer, J. G. P. Bloom, N. H. Dekker, C. F. Dutta, A. Blundell, T. Robinson, K. Firman, and C. Dekker, Nature Structural & Molecular Biology **11**, 838 (2004), ISSN 1545-9993, pMID: 15300241, URL http://www.ncbi.nlm.nih.gov/pubmed/15300241.

[16] R. Seidel, J. G. P. Bloom, C. Dekker, and M. D. Szczelkun, The EMBO Journal **27**, 1388 (2008), ISSN 1460-2075, pMID: 18388857, URL http://www.ncbi.nlm.nih.gov/pubmed/18388857.

[17] S. Bigot, O. A. Saleh, F. Cornet, J.-F. Allemand, and F.-X. Barre, Nat Struct Mol Biol **13**, 1026 (2006), ISSN 1545-9993, URL http://dx.doi.org/10.1038/nsmb1159.

[18] O. A. Saleh, S. Bigot, F.-X. Barre, and J.-F. Allemand, Nat Struct Mol Biol **12**, 436 (2005), ISSN 1545-9993, URL http://dx.doi.org/10.1038/nsmb926.

[19] G. Lia, E. Praly, H. Ferreira, C. Stockdale, Y. C. Tse-Dinh, D. Dunlap, V. Croquette, D. Bensimon, and T. Owen-Hughes, Mol Cell **21**, 417 (2006), URL http://dx.doi.org/10.1016/j.molcel.2005.12.013.

[20] G. Lia, D. Bensimon, V. Croquette, J.-F. Allemand, D. Dunlap, D. E. A. Lewis, S. Adhya, and L. Finzi, Proceedings of the National Academy of Sciences of the United States of America **100**, 11373 (2003), ISSN 0027-8424, pMID: 14500788, URL http://www.ncbi.nlm.nih.gov/pubmed/14500788.

[21] J. Gore, Z. Bryant, M. D. Stone, M. Nöllmann, N. R. Cozzarelli, and C. Bustamante, Nature **439**, 100 (2006), ISSN 1476-4687, pMID: 16397501, URL http://www.ncbi.nlm.nih.gov/pubmed/16397501.

[22] D. A. Koster, V. Croquette, C. Dekker, S. Shuman, and N. H. Dekker, Nature **434**, 671 (2005), ISSN 14764687, pMID: 15800630.

[23] A. Bancaud, N. C. e Silva, M. Barbi, G. Wagner, J.-F. Allemand, J. Mozziconacci, C. Lavelle, V. Croquette, J.-M. Victor, A. Prunell, et al., Nature Structural & Molecular Biology **13**, 444 (2006), ISSN 1545-9993, pMID: 16622406, URL http://www.ncbi.nlm.nih.gov/pubmed/16622406.

[24] C. Deufel, S. Forth, C. R. Simmons, S. Dejgosha, and M. D. Wang, Nat Meth **4**, 223 (2007), ISSN 1548-7091, URL http://dx.doi.org/10.1038/nmeth1013.

[25] H. Zhang and J. F. Marko, Physical review. E, Statistical, nonlinear, and soft matter physics **77**, 031916 (2008), ISSN 1539-3755, pMID: 18517431.

[26] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette, Biophysical journal **74**, 2016 (1998), ISSN 0006-3495.

[27] J. Moroz and P. Nelson, Macromolecules **31**, 6333 (1998), ISSN 0024-9297, URL http://pubs3.acs.org/acs/journals/doilookup?in_doi=10.1

[28] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette, BIOPHYSICAL JOURNAL **76**, 409 (1999), ISSN 0006-3495.

[29] J. F. Marko, Physical review. E, Statistical, nonlinear, and soft matter physics **76**, 021926 (2007), ISSN 15393755, pMID: 17930084.

[30] S. Forth, C. Deufel, M. Y. Sheinin, B. Daniels, J. P. Sethna, and M. D. Wang, Physical Review Letters **100**, 148301 (2008), URL http://link.aps.org/abstract/PRL/v100/e148301.

[31] J. D. Moroz and P. Nelson, Proceedings of the National Academy of Sciences of the United States of America **94**,

14418 (1997), ISSN 00278424, pMID: 9405627.

7.1 Supplementary material

7.1.1 DNA substrate

The measurements described in this section were performed on a standard magnetic tweezers apparatus with a DNA substrate of 17.8kb. This gives a nominal extension of $L_0 \approx 6\mu$m. DNA was functionalised on both ends

7.1.2 Beads preparation

DNA substrate is incubated with 10μL of 1μm diameter superparamagnetic (My-One, Dynal) beads. Before incubation the beads are rinsed in PBS three times. Rinsing is achieved with a magnetic separator that allows to eliminate the liquid buffer while keeping the beads at the bottom of the test tube. After rinsing, the beads are re-suspended in 10μL of passivation buffer A.2 and mixed with 0.5μL of DNA substrate.

The mixture DNA beads is incubated for some minutes to allow the binding of functionalised DNA molecules to the surface of magnetic beads. Then, the mixture is diluted 5 times and 5-10μL of the diluted mix are injected in the passivated chamber. The beads will cover the surface and some of them will remain attached specifically. Incubation of the beads in the chamber is allowed for some minutes to increase the probability that beads are bound to the surface.

Then, a gentle flow is injected in the chamber in order to remove all the beads that are not bound to the surface. This step requires about one hour at least, because a very slow flow has to be used in order not to eliminate all the beads. The concentration of beads with respect to DNA is chosen in such a way that quite few beads remain attached to the surface. This choice is necessary to reduce the possibility that a bead is attached by more than one DNA molecule.

After all unbound beads are removed the chamber is sealed on both ends to reduce evaporation and the experiment is started.

7.1.3 Preliminary checks

A bead attached to a single DNA molecules is searched by approaching the magnets to the surface in order pull on beads. A good candidate bead detaches from the surface and its distance from the surface depends on the rotation of the magnets.

Once a candidate bead is identified, a curve of extension versus rotation is recorded at low force. In this regime we expect to see a symmetric behaviour of the extension of the molecule when both negative and positive supercoiling are introduced. If this is the case, the position of maximal extension is recorded and fixed to indicate the relaxed state of the molecule. As it was explained in

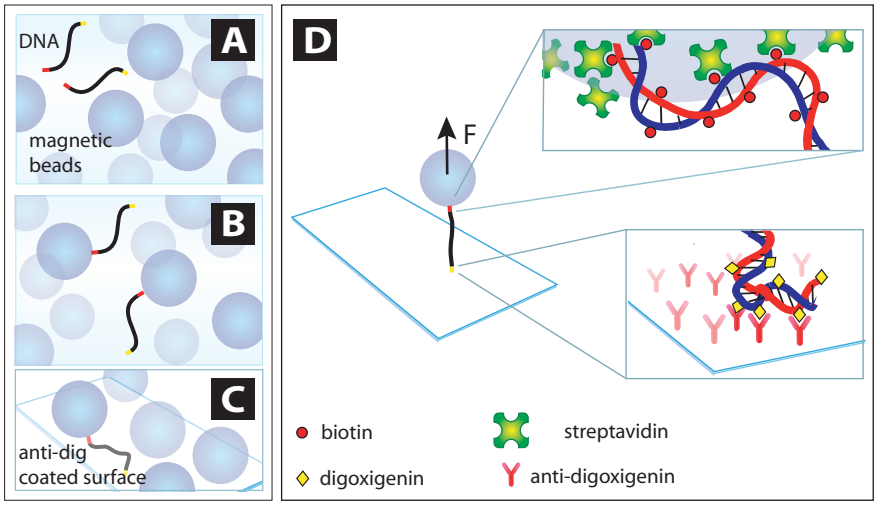


Figure 7.1: Incubation of beads

section 5.1, this is the state in which the double helix is torsionally relaxed and its linking number corresponds to $Lk_0 = L/p$.

In this position, a force-extension curve is recorded. As we have seen in section 5.2, the behaviour of torsionally relaxed DNA subject to an external force is well described by the WLC model, which depends only on two external parameters: the bending persistence length and the contour length of the molecule. Fitting the force extension curve to the WLC model serves as a consistency check, because the fitted values for A and L can be compared with the known values.

7.1.4 Limitations of the measurement technique

Various factors intervene and limit the domain of measurable forces. Upper bounds on the values of applicable forces are set by two factors. First of all the geometry of the magnetic tweezers setup only allow to approach the magnets up to the contact with the Mylar surface at the top of the chamber. With one micron beads this allows to apply a maximal force of about 15pN. There is also another limiting factor, which is the acquisition frequency of the camera. In fact, in order to correctly estimate the pulling force, the spectrum of fluctuations of the bead has to be measured as explained in section 6.2.8. Fitting of this spectrum to a Lorentzian curve yields the cut off frequency, which is related to the stiffness of the magnetic trap and to the size of the bead. If the force is to high, small beads fluctuate too fast for their fluctuations to be correctly sampled at video resolution. This fact limits the highest measurable force with MyOne beads to roughly 5pN.

7.1.5 Limitations in the method for torque estimation

The method for torque estimation from hat curves proposed in our paper relies on the assumption that torque is constant along the straigth linear part of the lowest force hat curve. This assumption seems to be justified in view of currently available experimental data from direct torque measurements [38, 18] and the theoretical framework proposed in [65]. However, other models [15] and preliminary data from T. Strick (personal communication) seem to suggest that this is not completerly true.

An upper bound on the error committed with this assumption is given by the value of buckling torque for the lowest force value measured in our experiences, which can be estimated from the simple model proposed in section 5.2.4:

$$\Gamma_b(0.1pN) = \sqrt{\frac{2fA}{k_BT}} \approx 1.6pNnm \tag{7.1}$$

This is to be considered as a possible sistematic error in our measurements, but it is however quite small when compared to the values of torques estimated for the high force curves.

Another source of uncertainty in our meethod could be the presence of the glass surface, which tends to flatten the hat curves and thus to bias the value of the slope of the linear part towards lower absolute values. Given the length of our molecule, for the lowest force measured the bucklin occurs at more than 2.5μm from the surcafe and we estimate that any effect due to surface proximity is negligible at this distance.

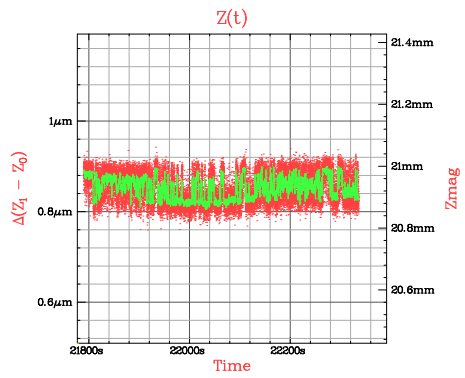
7.1.6 Buckling instability

The results proposed in our treat the buckling instability as if it were a smooth transition. This is not correct, and indeed the behaviour of the DNA molecule in the vicinity of the buckling transition appears to be richer than expected.

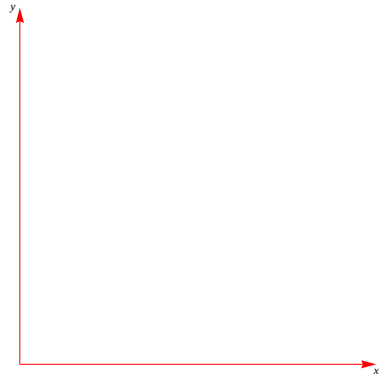
Jumps

Forth et al. [38] reported the observation of abrupt jumps in the extension of the molecule near the buckling transition. We investigted this previously unreported behaviour recording the extension of the molecule near the buckling transition at higher resolution in force and sigma.

Figure 7.2a displays a behaviour we observed near the buckling transition. It looks like a telegraphic signal typical of a two states systems hopping back and forth from one state to the other. Such signal could be observed by fixing the value of supercoiling density σ and letting the force vary in steps of very small size (see figure 7.2. Jumps are observed over a very narrow range of forces and can rarely be observed if the force is kept constat while varying σ. In fact, since



(a) Telegraphic signal of jumps



(b) Tecnique used to observe the jumps jumps

small misalignments of the magnets are sufficient to let the force vary slightly from one angular position to another, hat curves are usually recorded in steps of one complete turn or more integer turns. This could explain whi such jumps had not been observed before: 1 turn is in fact sufficient to pass from the high state to the low state for a molecule of small size. For longer molecules, acquisition of a hat curve usually procedes in steps of 5 or 10 turns, and the same type of inconvenient is present.

Our current data on jumps are not enough to draw definitive conclusions. We could onlyobserve them at high force, where brownian fluctuations of the bead in the longitudinal direction are smaller than the size of the jump. At lower forces the jumps should still be detectable as an additional widening in the brownian fluctuations, but the measurement conducted so far did not allow to confirm this.

Jump size and salt dependency

We observed jumps on a very short (0.851μm) molecule and on the long 6μm molecule. The size of such jumps seems to change in the two cases, and this demans more investigation.

Appearance of jumps seemed to be favoured in high salt conditions, althogh a spread of fluctuations could be observed in low salt buffers as well.

Interpretation

The observed telegraphic behaviour was interpreted by the authors of [38] as the appearance of the first plectoneme loop. If this is correct the size of the jumps should depend on salt conditions but not on the molecule length. In fact, the typical size of a loop was estimated in section 5.2.4 to be of the order of 40nm for a force of 0.1pN and it depend on the persistence length A which has been observed to vary with salt condition. On the other hand, the size of the loop should be independent of the length of the molecule, which did not seem to be

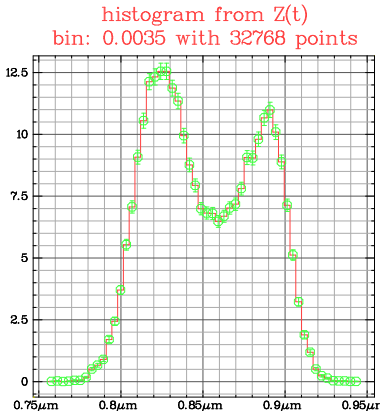


Figure 7.2: Histogram of the size of jumps

the case in our preliminary measurements. If this were to be confirmed another explanation should be looked for. The model proposed in [75], where buckling is attributed to the formation of global helical structures could be a good candidate to explain this.

Chapter 8

Soft magnetic tweezer

Soft magnetic tweezer

Francesco Mosconi,* Jean Franǫis Allemand, and Vincent Croquette
LPS, ENS, UMR 8550 CNRS, 24 rue Lhomond, 75231 Paris Cedex 05, France
(Dated: July 31, 2008)

PACS numbers:
Keywords:

I. INTRODUCTION

Force spectroscopy was introduced in biology in the early nineties, through the development of Single Molecule micromanipulation techniques such as optical tweezers (OT) [1, 2] and magnetic tweezers (MT) [3, 4]. Since then, the force response of many biological systems has been studied and the mechanical properties of DNA and proteins [5] have been extensively characterised. While most man made polymers are free to swivel about their links, many biological molecules (DNA, actin, microtubules, collagen etc.) can bear a torsional load. It is thus very interesting to introduce torque spectroscopy in biology to study the mechanical response of bio-macromolecules to twist deformations. In particular, DNA is a long cylindrical object, which forms a spiral staircase and it is natural to assume that DNA interacting motors may apply and/or respond to torsional stress on the molecule.

Although the conception and develpment of torque-sensitive measurements seems natural, the difficulty in the construction of such devices is due to the high sensitivity necessary to measure torques on single molecules. The DNA denaturation torque, for example, has been reported to be as low as $10pNnm$ or $2.5k_BT$[6]. A technique previously developed [7] to measure and apply torque to single biomolecules relies on a modification of the OT that exploits angular momentum transfer between a laser beam and a polarizable material. This technique has allowed to measure the torque response of DNA for stretching forces larger than $1pN$. Un fortunately, lower force regimes become difficult to explore with this apparatus. Another technique [6], based on the video tracking of twist relaxation, allowed to perform a non-equilibrium measurement of the response to torque of a stretched DNA molecule.

MT are a micromanipulation technique that can apply forces in the $[0.1 - 10]pN$ range, while imposing a torsional constraint on the trapped molecule. It is thus natural to develop an extension of the MT in which a torque of desired intensity can be applied. The traditional implementation of the MT [4] uses static magnets to create an horizontally oriented magnetic field with a strong vertical gradient. The interaction between the magnetic dipole of the superparamagnetic bead $\vec{\mu}$ and the magnetic field $\vec{B}$ generates a force $\vec{f} = (\vec{\mu}\vec{\nabla})\vec{B}$ and a torque $\vec{\Gamma} = \vec{\mu} \times \vec{B}$. The magnetic torque tends to align the direction of the preferential axis of the bead with the magnetic field, while the force is vertical and tends to attract the bead towards the magnets.

This imposes a stiff constraint on the angular orientation of the trapped magnetic bead. The device presented in this article relaxes that constraint by using a mangetic field rotating faster than the bead can. Torque measurement can then be obtained in a way similar to force measurement through tha application of the fluctuation-dissipation theorem. In fact, standard MT use the transverse fluctuations $\langle \delta x^2 \rangle$ of the bead to determine the force $F = k_BTl/\langle \delta x^2 \rangle$ pulling on a DNA of mean extension l. Similarly, torque can be estimated from angular fluctuations (see below). The advantages of this proposal are in its simplicity and its applicability to low force regimes.

II. DESCRIPTION OF THE SYSTEM

A. The tosional stiffness of a magnetic tweezer

Our device is an improvement of the magnetic tweezers that allows for the simultaneous application and measurement of a force and a controlled torque to a DNA molecule. The main difference with respect to the magnetic tweezers is that it is capable to soften the torsional constraint given by the static magnetic field used in the magnetic tweezers. To explain this point let us give some order of magnitudes: biological forces at the scale of the single molecule are in the range of picoNewtons (pN) while torques are in the range of picoNewtons.nanometers (pN.nm). The magnetic forces exerted by magnetic tweezers corresponds to the vertical field gradient times the magnetization. If we assume the magnetic field to decrease exponentially with the distance to the sample, for the sake of simplicity:

$$B = B_0 e^{\frac{-z}{b}} \tag{1}$$

where B_0 is the maximum strength of the field, z is the distance between the magnet and the bead and b is the decay length of the field (b is typically the gap size of the polar pieces), the force is immediately obtained through derivation with respect to the distance variable z, and its

*mosconi@lps.ens.fr

value is:

$$F = \frac{\mu B_0}{b} e^{\frac{-z}{b}} = \frac{\mu B_0}{b} \tag{2}$$

If the bead has an anisotropy α in magnetic susceptibility, the torque applied by the bead will be:

$$\tau = \alpha \mu B_0 \sin\theta \tag{3}$$

where θ is the angle between the field direction and that of the bead magnetization. We can easily relate the torque to the applied force:

$$\tau = \alpha F b \sin\theta \tag{4}$$

Since one needs to apply a force of the order of 1pN, in the case of small angle the torque will be this force multiplied by $(\alpha b\theta)$. Typically α is of the order of a few percent but b is the millimetre range. This leads to an estimated torque value of 10000 pN.nm for a significant angular deviation θ! Clearly this torque is considerably too strong compared with biological values.

Our device allows to considerably reduce the torsional stiffness of the magnetic tweezers in such a manner that the angular fluctuations of position of the paramagnetic bead related to Brownian motion become now clearly visible. Depending upon the protocol used, this device can be used to study the DNA under constant torque or constant torsion.

B. Softening the torsional constraint

The soft magnetic tweezers exploits a non-linear coupling regime which is encountered in a viscous rotator system composed by a magnetized bead in a viscous fluid under the effect of a rotating magnetic field. This system exhibits two characteristic behaviours. The first one is the synchronous mode where the magnetization follows the direction of the magnetic field in perfectly synchronous manner (this behaviour corresponds to the normal use of a compass). The second type of behaviour appears when we increase the angular velocity of the magnetic field. Above a well defined velocity, the bead will ultimately be unable to follow the magnetic field and will exhibits a new behaviour with a complex motion made of fast oscillations and slow rotation: this phenomenon is well known in microrheology applications [8, 9] and we will call it "slippage". By "slippage" we mean the fact that the orientation of magnetization vector of the paramagnetic bead is no longer coupled to the orientation of the bead itself. This happens because the bead is immersed in a viscous fluid, which exerts a drag that depends on the angular velocity of the bead. In the synchronous mode, to compensate for the viscous torque, the bead magnetization and the magnetic field form an angle and thus the magnitude of the torque applied by the magnetic field increases with increasing angle. Since the torque depends on the vector product of the magnetic field vector and the magnetization direction, its value will be maximal when the two vectors are perpendicular. When the external magnetic field is rotated sufficiently fast, the viscous drag will be so high that it exceeds this maximum value and the bead will not be able to follow the rotation of the magnetic field any longer. In the slippage regime, the magnetization of the bead rotates with the magnetic field but the bead does not follow.

Mathematically, a magnetic field $\vec{B}$ is applied to a paramagnetic bead immersed in a viscous fluid. The magnetic moment developed by the bead, $\vec{\mu}$, tends to align the principal anysotropy axis of the bead with the direction of the external magnetic field. If the external field is rotated, the stationary change is balanced by the magnetic torque:

$$\gamma\dot{\theta}(t) = \alpha\mu B \sin\left(2(\Omega t - \theta(t))\right) \tag{5}$$

where μ is the magnetic moment developed by the bead, $\gamma = 8\pi\eta r^3$ is the rotational viscous drag coefficient, α is the anysotropy and Ω is the angular velocity of the magnetic field.

For a slowly rotating field, the bead responds by rotating with same rate with a phase lag: $\theta(t) = \Omega t - \phi$, where the phase lag ϕ between the angular position of the magnetic field and that of the bead is set by the condition:

$$\sin(2\phi) = \frac{\gamma\Omega}{\alpha\mu B} \tag{6}$$

Increasing the angular velocity of the magnetic field results in an increased torsional viscous drag on the bead, which produces an increased phase lag between $\vec{B}$ and $\vec{\mu}$. When Ω reaches the maximum value $\Omega_c = \frac{\alpha\mu B}{\gamma}$ the increase in viscous drag cannot be compensated by the magnetic torque any longer. For values of $\Omega > \Omega_c$ the system enters a non-linear regime where the motion of the θ variable is a sum of a fast oscillation and a slow rotation. A perturbative solution of the e.o.m. for large Ω can be obtained by performing the change of variable $t \to \Omega t$ and then considering a series expansion in $\epsilon = \alpha\mu B/\gamma\Omega$ of equation 5:

$$\begin{aligned}\theta(t) &= \theta_0 - \frac{\epsilon}{2\Omega}\cos(\Omega t - \theta_0) + \frac{\epsilon^2}{2}\left(\Omega t - \theta_0 + \frac{1}{4}\sin(4(\Omega t - \theta_0)\right) \quad (7)\\ &= \theta_0 + \frac{A}{2\Omega}\left(A(t - \frac{\theta_0}{\Omega}) - \cos(\Omega t - \theta_0)\right) + o(\frac{A}{\Omega}) \quad (8)\end{aligned}$$

where $A = \frac{\alpha\mu B}{\gamma}$ (The interested reader is referred to appendix A for the detailed calculation). This approximate solution is composed by a linear term and an oscillating term. The average angular velocity of the bead is obtained by considering the derivative with respect to time of the linear term: $\bar{\theta} = A^2/(2\Omega)$. The first order oscillating term indicate that the motion of the bead will be a superposition of a slow rotation (the previous term) and a fast oscillation at frequency Ω. This implies that for a magnetic field rotating at high frequency, a slow rotation can be obtained.

C. Implementation

The soft magnetic tweezers apparatus, is implemented substituting the permanent magnets of magnetic tweezers with electromagnets. Other characteristics such as parallel illumination and sample preparation are the same as in [4]. Briefly, a glass chamber is prepared, fixed on the stage of an inverted microscope and illuminated with parallel light. Magnets are suspended just above the chamber in order to provide a vertical pulling force and an orizontally aligned magnetic field. Light is collected through a 100X il immersion microscope objective (Olympus, France) and the image is formed on 25Hz CCD camera (Sony, France). DNA is attached to the bottom of the glass chamber on one end and on a paramagnetic bead on the other end. The image of the bead is used to track is position in the three dimension.

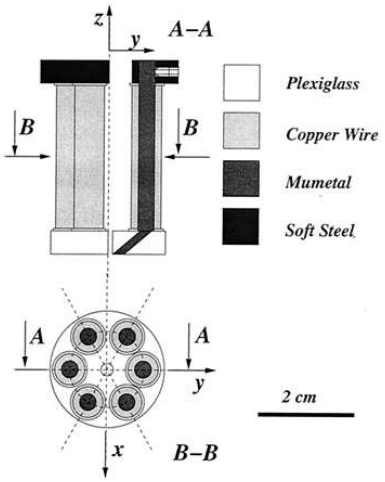


FIG. 1: Setup

This apparatus can work in the exact same mode as the traditional one, if a static current is applied to the electromagnets. Nevertheless, since there are no moving parts, it is possible to rotate the orientation of the magnetic field at higher frequencies. The current necessary to generate a magnetic field of desired intensity is provided to the 6 coils by a 6 power amplifiers that take an input signal generated by a computer controlled DAC card.

1. Coils

The six coils (Lima 600880, Vicenza, Italy) are disposed on the six corners of a hexagon. They are made of copper wire and have a resistance of $10 \pm 2\Omega$. Each of them encloses a cylinder of mumetal (Goodfellow, Cambridge, U.K.) slightly longer than the coil. These six polar pieces are inserted in a soft steel cap on the top, and bent towards the centre at the bottom. The XC15 soft steel cap (Tonnetot Metaux, Fontenay-sous-Bois, France) is designed to close the field lines in the system, thus increasing the magnitude of the magnetic field. Finally the bent tips are also inserted in a Plexiglass ring which confers overall stability to the ensemble.

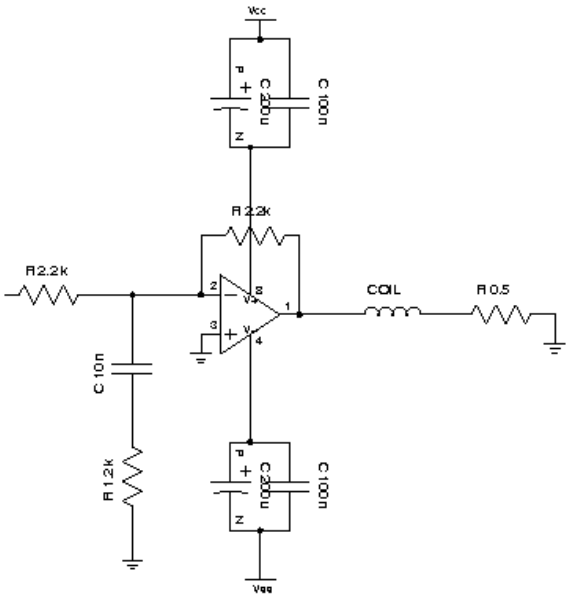


FIG. 2: Scheme of the power amplifier

2. Amplifiers

The current that drives the coils is obtained by a home built six channel voltage-voltage amplifier with feedback (see figure 2 for a schematic diagram of one channel). The current passing through the coils depends on the impedance of the coil and thus on the frequency of the signal. We chose this amplifier configuration because it has a fast response and thus allows to apply fast modulations to the current intensity. The price we pay in this configuration is that the field is not simply proportional to the applied voltage, and so care has to be taken when asking the value of the magnetic field.

D. Signal Generation

The voltage signal is generated by an analog output computer board (NI-PCI6733, National Instruments, USA), which can generate analog samples between -10V and 10V on eight independent channels, at maximal frequency of 1MHz per channel. The board also includes an memory of 4096 samples and can access larger buffers using DMA channels to the computer main memory. In order to generate a signal with a frequency in the range

1-10kHz, modulated in amplitude and shape in realtime without interruption, we actively regenerate the buffer of samples that are sent to the board. Since the new samples are calculated in realtime, the signal can be modified in realtime with an update frequency given by $f_u = \frac{2f_s}{N}$.

If this update frequency is too high, the computer will not be able to calculate all the samples before the following update and thus the board will come to a halt due to the lack of input. On the other hand, a too low frequency would lead to unacceptable waiting times between the change of the driving signal and its actuation on the card output. We empirically found that an update frequency between $10Hz$ and $50Hz$, satisfies both criteria. All routines are written in C using the NIDAQmx API. The interested reader is referred to A for the details of the generation of the rotating magnetic field signal.

III. SYSTEM CHARACTERISATION

A. Preliminary checks

Figure 3 displays the decoupling of bead motion with increasing frequency.

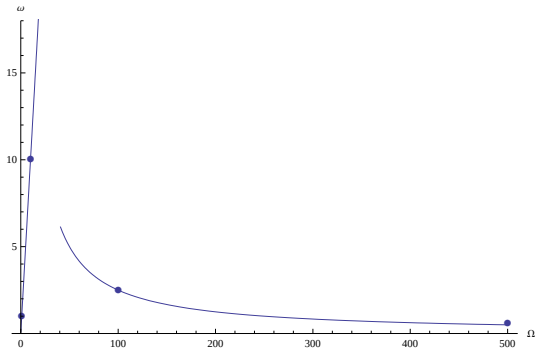


FIG. 3: The rotation frequency of a bead doublet as a function of applied rotation frequency

The critical decoupling frequency appears to be $\Omega_c \approx 16Hz$, although a precise determination is difficult due to the low number of data points. We also observe that the height of the beads with respect to the surface decreases with increasing frequency of rotation, which means that also the applied force depends on Ω. We will discuss this limitation in section IV. Operating the device in this regime yields a fixed torque trap in which a small average torque of size $\gamma\eta r^3\dot{\bar{\theta}}$ is constantly applied to the bead. In this configuration, the bead rotates with constant angular velocity and oscillates very fast about its mean position.

In certain applications, expecially when dealing with torsionally constrainable DNA, it may not be suitable to have a constantly rotating bead. In fact, if the DNA molecule is coilable, i.e. if it can accumulate torsion, rotation of the bead will change its extension, ultimately approaching the bead to the surface or even leading to breaking the DNA molecule. For these reasons, it is desirable to have a second configuration of the device in which the average torque imposed is zero and a fixed angular position is imposed on the bead with a quadratic angular potential of stiffness k_θ. In this configuration the bead is subject to brownian fluctuations whose size is determined by the stiffness via the fluctuation dissipation theorem:

$$\langle \Delta\theta^2 \rangle = \frac{k_B T}{k_\theta} \tag{9}$$

A configuration of this kind can be obtained by inverting the sense of rotation of the magnetic field after a certain number of turns. This inverts the direction of the applied torque and gives an average zero torque when the limit of infinite time is considered. We implemented this configuration in our device by performing N turns in one direction and N in the opposite direction. Fig. 4 displays a comparison of the amplitude of angular fluctuations for the static MT and this configuration and it can be seen that the size of fluctuations is increased, corresponding to a softening of the angular stiffness.

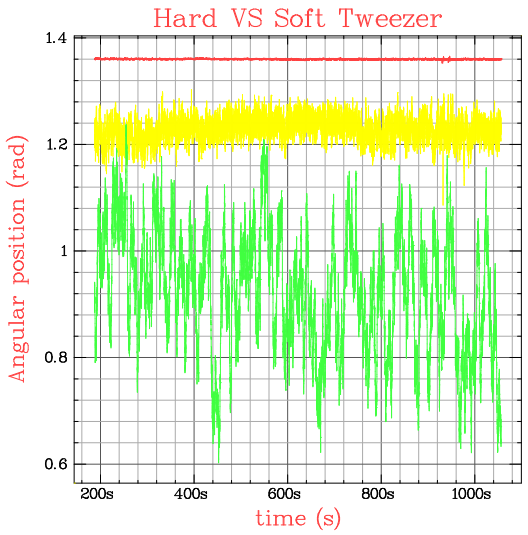


FIG. 4: Brownian fluctuations of the angular variable for the traditional MT, the soft MT and the soft MT with second order corrections

Unfortunately, the abrupt inversion of the rotation at a fixed angle θ_0 breaks the rotational symmetry of the trap and the bead's main axis of anysotropy tends to align along the preferential directionθ_0. The precise origin of this alignent is not completely understood, by it is possibly linked with the histeresis of the mumetal used in the cores of the electromagnets, preventing the inversion from being instantaneous and complete. Furthermore, the position of the six polar pieces appears to introduce yet another static contribution to a preferential orientaton of the bead's axis, the two effects conspiring to give

an effective angular potential with a minimum at an uncontrollable angular position.

In order to overcome these limitations, we devised and implemented two additional corrections to the field generating signal. In order to soften the impact of the sharp inversion we modulated the amplitude of the signal with a $(1-\cos(2\Omega t/N))$ wave so that the field reached zero amplitued at the inversion points. This modulation implies a modulation of force (which is linked to the amplitude of the field) and requires the inversion to be done sufficiently fast, so that the average force experience by the bead is constant and no large oscillation is observed. Inverting the rotation every N=5 turns at 1ms per turn (1kHz) implies an inversion frequency of 200Hz which is still much higher than the viscous cut-off frequency for brownian fluctuations in the longitudinal direction. In other words, this 200Hz force modulation is damped by the viscous fluid and the bead experiences a constat average force.

This modulation is not sufficient to completely eliminate the anisotropy of the field, so we are obliged to intruce a second correction: we impose that the curve described bt the magnetic field vector is not a circle but an ellipse. By chosing the orientation of the main axis of the ellipse and by varying the ellipticity parameter, we can effectively cancel the remaining anysotropy, thus lowering the stiffness of the trap to values that allow measurement of torques of biological size via equation 9. The additional bonus of this elliptic correction is that, by playing with the orientation and the ellipticity parameters, an arbitrary preferential angular direction with desired stiffness can be imposed, which is very suitable for experiments on coilable DNA mocules. The interested reader is referred to the appenix B for the details of the signal generation. Figure 4 displays the angular fluctuations of a bead doublet for the cases of static magnetic field, fast rotating magnetic field with inversions and fast roating field with inversions, modulation and second order elliptic correction. In particular, the ellipticity parameter allows us to finely control the stiffness of the trap as is shown in figure 5

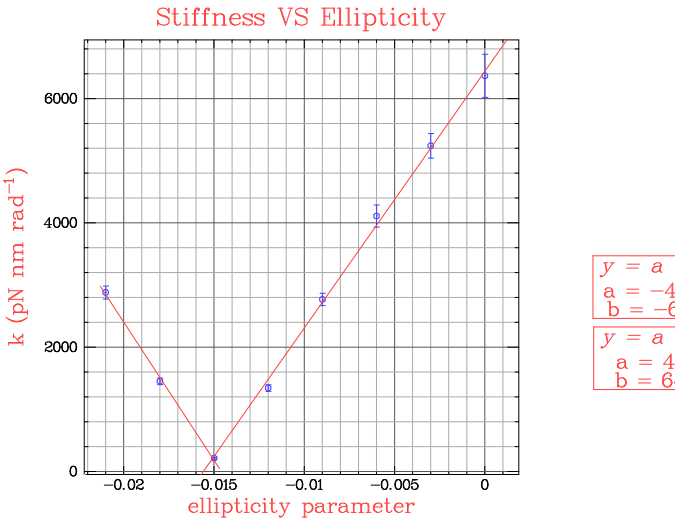


FIG. 5: Stiffness of the trap versus the ellipticity parameter

B. Validation on DNA

A soft angular potential can be used to determine an applied torque by tracking angular deviations through the formula:

$$\langle \tau \rangle = k_\theta \langle \theta - \theta_0 \rangle \tag{10}$$

where θ_0 is the minimum of the angular potential. We validated our device by measuring the response to torque of a DNA molecule.

DNA has a chiral geometry and its twist rigidity has been characterized in a number of experiments [10]. In particular, a torsionally constrained DNA molecule, subject to an external pulling force, is expected to respond with a torque that increases linearly as a function of the number of turns added to the molecule. When the number of added tuns overcomes the critical value for buckling, the torque stabilises to a fixed value and it remains constant.

We injected paramagnetic beads attached to DNA in the flow cell and measured angular deviations from the equilibrium position as a function of the number of turns applied. In order to facilitate angular tracking we searched for DNA molecules attached to bead doublets. The experience proceeded as follows: when a coilable molecule attached to a bead doublet was identified, its extension versus rotation at fixed force was measured in order to identify the position of relaxation. In this position the molecule reaches its maximum extension for a given force. Once this is identified we proceeded recording the angular fluctuations of the beads when $-75, -45, -15, 15, 45, 75$ turns were added to the molecule. Each position was maintained for about 7 minutes and then 30 turns were added to switch to the next position. The whole sequence was repeated 100 times in order to accumulate enough statistics to refine the sampling of the average position. The use of bead doublets enormously increased the relaxation times of the system (viscous drag for a 2 bead system is not $8\pi\eta r^3$ but $28\pi\eta r^3$. The choice of integration times of 7 minutes is a compromise between the need to average the angular fluctuations long enough and the risk of introducing low frequency noise in the measurment due to thermal drifts. Figure 6 shows a curve acquired at $f = 0.5pN$ which seems to indicate that indeed our apparatus is able to detect torque variations and to perform absolute torque measurement once calibrated using the fluctuation-dissipation theorem.

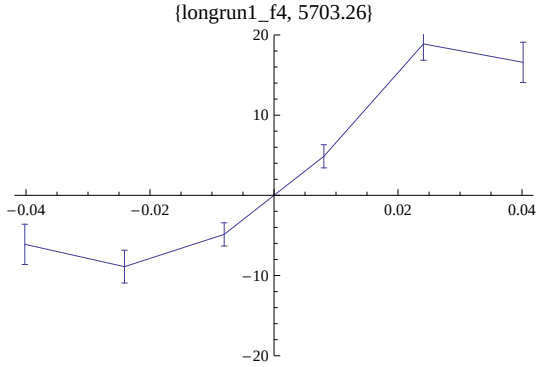


FIG. 6:

IV. DISCUSSION

The current limitation of our apparatus mainly consists in the long time required for measurments. Figure 6 required 3 days of continuous acquisition on a single DNA molecule. This limitation is severe if the action of an enzyme is to be studied dynamically but fortunately it is mainly due to technical factors: the use of bead doublets for ease of tracking but most of all the high viscous drag of the large $4.5\mu m$ beads that were used to apply the pulling force. As mentioned above, we noticed that the pulling force decreases when the rotation frequency of the magnetic field is increased to values higher than $1kHz$. This is due to the inductance of the coils that depends on the frequency of the alternating current and also to Foucault dissipative currents developed in the metallic cores. With the current implementation of the device, it was not possible to apply a pulling force to beads of smaller size, and since relaxation times are proportional to the cube of the radius of the bead, great improvements can be achieved if smaller size beads are used. This will hopefully be possible with the use of lamellar materials together with coils of lower inductance.

Another limitation of the current setup is the heat developed by the coils when high frequency fields are used. This heat produces thermal drifts of the sample, and complicates the longitudinal tracking required for force measurement (although differential tracking by comparing the position of the bead with a bead fixed on the glass surface practically resolves the issue). This problem will also be partly solved with lamellar material that will not produce as much heat as the current cores.

In conclusion, we built a new apparatus that allows the application and measurement of torques of biological size. Our device is an extension of the magnetic tweezers apparatus and was validated with the measurement of torque on a single DNA. Technical improvements are on the way in order to overcome the current limitations of dynamic range.

Appendix A: Approximate solution of equation A4

We can rescale the time variable from t to $t' = \Omega t$. Equation A4 becomes:

$$\dot{\xi}(t) = \frac{A}{\Omega}\sin 2(t - \xi(t)) \tag{A1}$$

with $\xi(t) = \theta(t/\Omega)$. Increasing the value of Ω can be viewed as effectively decreasing the value of the coupling constant before the sine. For $\Omega \to \infty$ the angular velocity tends to zero.

We can treat equation A1 considering an expansion of the solution in a series of powers of $\epsilon = A/\Omega$. We write:

$$\xi(t) = \xi_0 + \epsilon\xi_1(t) + \epsilon^2\xi_2(t) + \dots \tag{A2}$$
$$\dot{\xi}(t) = \epsilon\dot{\xi}_1(t) + \epsilon^2\dot{\xi}_2(t) + \dots \tag{A3}$$

and then:

$$\epsilon\sin(2(t-\xi(t))) = \epsilon\sin(2(t-\xi_0) - 2(\epsilon\xi_1(t) - \epsilon^2\xi_2(t) - \dots)) \tag{A4}$$
$$= \epsilon\sin(2(t-\xi_0))\cos(2(\epsilon\xi_1(t) + \epsilon^2\xi_2(t) + \dots)) \tag{A5}$$
$$-\epsilon\cos(2(t-\xi_0))\sin(2(\epsilon\xi_1(t) + \epsilon^2\xi_2(t) + \dots)) \tag{A6}$$
$$= \epsilon\sin(2(t-\xi_0)) - 2\epsilon^2\xi_1(t)\cos(2(t-\xi_0) + \dots) \tag{A7}$$

where we first expanded $\sin 2(t-\xi(t))$ using the identity $\sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b)$, and then approximated the sine and cosine for small values of ϵ. The approximate solution is calculated at the desired order in ϵ by grouping the terms of the same order. Here we consider only terms up to order ϵ^2, obtaining:

$$\xi(t) = \xi_0 - \frac{A}{\Omega}\cos 2(t-\xi_0) + \frac{A^2}{2\Omega^2}(t-\xi_0) + \frac{A^2}{8\Omega^2}\sin(4(t-\xi_0)) \tag{A8}$$

the value of ξ_0 is calculated using the initial condition. Translating back this equation to the equation for the variable θ we obtain:

$$\theta(t) = \theta_0 - \frac{A}{\Omega}\cos 2(\Omega t-\theta_0) + \frac{A^2}{2\Omega^2}(\Omega t-\theta_0) + \frac{A^2}{8\Omega^2}\sin(4(\Omega t-\theta_0)) \tag{A9}$$

Considering terms up to order ϵ^2, the solution is a superposition of a slow linear motion of angular speed $\frac{A^2}{2\Omega}$ and an oscillatory motion whose main contribution comes from the cosine term.

Appendix B: Rotating field construction

In order to apply a determined magnetic field, a relationship between the voltage applied to each channel and the total magnetic field must be established. Consider a couple of opposed electromagnets. A static magnetic field can be generated providing opposite continuous currents to the two coils. The generated magnetic field intensity will be proportional to the electric current passing through the copper coils, which is determined by the voltage applied and the impedance of the solenoid. The total

magnetic field generated by the three couples of magnets can be viewed as a linear combination of three magnetic fields generated by coupled opposite pairs of solenoids. The six elecromagnets are grouped in three couples of inverse polarities. This allows us to drive the six electromagnets with three signals, reducing the number of output channels required.

1. Rotating Phase

In order for the slippage phenomenon to appear, the magnetic vector field B has to be rotated sufficiently fast for the viscous drag to prevent the magnetic bead from following. The rotating phase $\theta(n,t)$ is defined as a function of time and of the number of turns n. The rotating phase is the junction of two pieces: a uniform anticlockwise motion for n turns, followed by a uniform anticlockwise motion.

$$\theta(n,t) = \begin{cases} 4n\pi(t - \lfloor t \rfloor) & t - \lfloor t \rfloor < 0.5 \\ 4n\pi(2 - t - \lfloor 2 - t \rfloor) & t - \lfloor t \rfloor > 0.5 \end{cases} \quad \text{(B1)}$$

where $\lfloor t \rfloor$ means Floor[t], i.e. the integer part of t (for positive t).

This form of rotating field is not sufficient to reduce the angular stiffness of the trap to the desired value.

2. Amplitude Modulation

The cosine amplitude modulation is defined as:

$$B(t,b) = 1 - b\cos\left(\frac{2\pi(t - \lfloor t \rfloor)}{0.5}\right) \quad \text{(B2)}$$

This depends on a parameter $b \in [0,1]$, which allows to switch continuously from no modulation to full modulation. b=0 means no modulation

3. Ellipse correction

In order to further reduce the angular stiffness of the trap, we impose that our magnetic field rotates on an ellipse, rather than on a circle. This allows us to correct the imperfect alignment of the six coils that tends to impose a preferential direction to the trap. Figure 7 shows the curve described by the magnetic field vector in XY plane.

An elliptical distortion of this figure is defined as:

$$E(e) = \begin{pmatrix} 1+e & 0 \\ 0 & 1-e \end{pmatrix} \quad \text{(B3)}$$

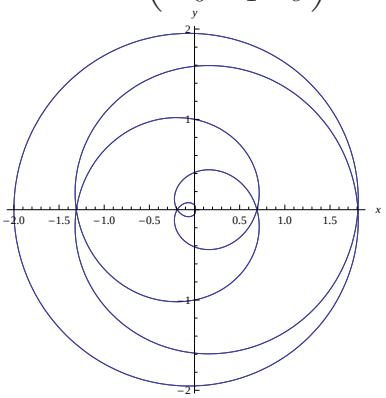
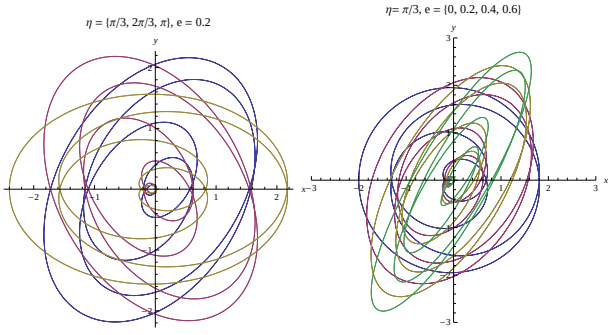


FIG. 7: The curve described by the magnetic field vector in the XY plane

And to allow for arbitrary rotation of the axes of the ellipse, the matrix:

$$T(e,\eta) = R(\eta)ER(-\eta) = \begin{pmatrix} 1 + e\cos[2\eta] & e\sin[2\eta] \\ e\sin[2\eta] & 1 - e\cos[2\eta] \end{pmatrix} \quad \text{(B4)}$$

can be applied to the rotating field vector. The relevant parameters are η controlling the angle formed by the ellipse major axis with the X axis, and e controlling the ellipticity.



(a) η (b) e

FIG. 8: Deformation of the ellipse depending on the choice of η and e

[1] K. Svoboda and S. M. Block, Annual review of biophysics and biomolecular structure **23**, 247 (1994), ISSN 10568700, pMID: 7919782.

[2] K. C. Neuman and A. Nagy, Nature Methods **5**, 491 (2008), ISSN 1548-7105, pMID: 18511917, URL `http://www.ncbi.nlm.nih.gov/pubmed/18511917`.

[3] S. B. Smith, L. Finzi, and C. Bustamante, Science (New York, N.Y.) **258**, 1122 (1992), ISSN 00368075, pMID: 1439819.

[4] C. Gosse and V. Croquette, Biophysical journal **82**, 3314 (2002), ISSN 00063495, pMID: 12023254.

[5] C. Bustamante, S. B. Smith, J. Liphardt, and D. Smith, Current Opinion in Structural Biology **10**, 279 (2000), URL http://www.sciencedirect.com/science/article/B6VS6-40GY8KC-4/2/b780ab2bb5f5193a7cae05f8692291d4/.

[6] Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante, Nature **424**, 338 (2003), ISSN 0028-0836, URL http://dx.doi.org/10.1038/nature01810.

[7] C. Deufel, S. Forth, C. R. Simmons, S. Dejgosha, and M. D. Wang, Nat Meth **4**, 223 (2007), ISSN 1548-7091, URL http://dx.doi.org/10.1038/nmeth1013.

[8] F. Amblard, B. Yurke, A. Pargellis, and S. Leibler, Review of Scientific Instruments **67**, 818 (1996), URL http://link.aip.org/link/?RSI/67/818/1.

[9] B. H. McNaughton, R. R. Agayan, R. Clarke, R. G. Smith, and R. Kopelman, Applied Physics Letters **91**, 224105 (2007), URL http://link.aip.org/link/?APL/91/224105/1.

[10] T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette, Science (New York, N.Y.) **271**, 1835 (1996), ISSN 00368075, pMID: 8596951.

Conclusions

This thesis presented the conception and development of different experiments in SM biophysics and I will summarize here the main results obtained.

In chapter 7 a new method to estimate the torque applied on DNA molecule in magnetic tseezers experiments was developed and succesfully applied to measure the critical torque of DNA in various sal tconditions. The results obtained agree with previous measurements and allow the determination of torque in previously unexplored regions of the parameter space, namely at low force. The measured data does not seem to agree with current models that describe torque and the amount of data collected opens the way for new theoreticl models to be tested. An independent determinatio of the biologically relevant twist modulust C has been determined and the dependence of this quantity on ionic conditions was studied. Observation of newly reported jumping behviour at the buckling transition was observed and more work is necessary to investigate its nature.

Chapter 8 presented the development of a system to measure torque on DNA directly, by lowering the angular stiffness of a magnetic trap and measuring angular fluctuations of the probing bead. A proof of principle validation was obtained by measuring the torque of a DNA molecule at different supercoiling densities, but use of the apparatus in dynamic situations involving DNA-enzyme interaction is not yet possible due to the long acquisition times required by the measurement technique.

Finally, chapter 4 presented the development of a setup to measure fluctuations of activity of a single enzyme. The results achieved and the improvements obtained in the three different setups tested bring the hope that it will be possible to measure this phenomenon soon.

Bibliography

[1] Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, and James D. Watson. *Molecular Biology of the Cell*. Garland, 3 edition, March 1994. [cited at p. 14]

[2] C B Anfinsen. Principles that govern the folding of protein chains. *Science (New York, N.Y.)*, 181:223–30, July 1973. PMID: 4124164. [cited at p. 9, 13]

[3] C B Anfinsen, E Haber, M Sela, and F H White. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America*, 47:1309–14, September 1961. PMID: 13683522. [cited at p. 9, 13]

[4] A. Ashkin. Trapping of atoms by resonance radiation pressure. *Physical Review Letters*, 40:729, March 1978. Copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org. [cited at p. 84]

[5] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, 11:288–290, 1986. [cited at p. 84]

[6] R H Austin, K W Beeson, L Eisenstein, H Frauenfelder, and I C Gunsalus. Dynamics of ligand binding to myoglobin. *Biochemistry*, 14:5355–73, December 1975. PMID: 1191643. [cited at p. 9]

[7] D Axelrod, T P Burghardt, and N L Thompson. Total internal reflection fluorescence. *Annual Review of Biophysics and Bioengineering*, 13:247–68, 1984. PMID: 6378070. [cited at p. 52]

[8] M A Basharov. Protein folding. *Journal of Cellular and Molecular Medicine*, 7:223–37, 2003. PMID: 14594547. [cited at p. 12]

[9] D. Bensimon, D. Dohmi, and M. M?zard. Stretching a heteropolymer. *Europhysics Letters*, 42:97–102, April 1998. [cited at p. 76]

[10] Keith Berland and Guoqing Shen. Excitation saturation in two-photon fluorescence correlation spectroscopy. *Applied Optics*, 42:5566–76, September 2003. PMID: 14526848. [cited at p. 47]

[11] F Bestvater, E Spiess, G Stobrawa, M Hacker, T Feurer, T Porwol, U Berchner-Pfannschmidt, C Wotzlaw, and H Acker. Two-photon fluorescence absorption and emission spectra of dyes relevant for cell imaging. *Journal of microscopy*, 208:108–15, November 2002. PMID: 12423261. [cited at p. 51]

[12] Binnig, Quate, and Gerber. Atomic force microscope. *Phys Rev Lett*, 56:930–933, March 1986. PMID: 10033323. [cited at p. 84]

[13] Michaël Bon, Simon J McGowan, and Peter R Cook. Many expressed genes in bacteria and yeast are transcribed only once per cell cycle. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 20:1721–3, August 2006. PMID: 16818468. [cited at p. 6]

[14] C. Bouchiat and M. Mézard. Elasticity model of a supercoiled dna molecule. *Physical Review Letters*, 80:1556, February 1998. Copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org. [cited at p. 76, 82]

[15] C. Bouchiat and M. Mézard. Elastic rod model of a supercoiled dna molecule. *The European Physical Journal E - Soft Matter*, 2:377–402, 2000. [cited at p. 102]

[16] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *BIOPHYSICAL JOURNAL*, 76:409–413, 1999. [cited at p. 82]

[17] Stephen G. Brush. *The Kind of Motion We Call Heat : Physics and the Atomists.* Elsevier Science Pub Co, September 1986. [cited at p. 6]

[18] Zev Bryant, Michael D. Stone, Jeff Gore, Steven B. Smith, Nicholas R. Cozzarelli, and Carlos Bustamante. Structural transitions and elasticity from torque measurements on dna. *Nature*, 424:338–341, 2003. [cited at p. 79, 102]

[19] P M Burgers and A Kornberg. The cycling of escherichia coli dna polymerase iii holoenzyme in replication. *The Journal of biological chemistry*, 258:7669–75, June 1983. PMID: 6345527. [cited at p. 6]

[20] Carlos Bustamante, Zev Bryant, and Steven B Smith. Ten years of tension: single-molecule dna mechanics. *Nature*, 421:423–7, 2003. PMID: 12540915. [cited at p. 10]

[21] Carlos Bustamante, Steven B Smith, Jan Liphardt, and Doug Smith. Single-molecule studies of dna mechanics. *Current Opinion in Structural Biology*, 10:279–285, 2000. [cited at p. 7]

[22] G. Charvin, J F Allemand, T. R. Strick, D. Bensimon, and V. Croquette. Twisting dna: single molecule studies. *Contemporary Physics*, 45:383–403, October 2004. [cited at p. 77, 135]

[23] Brian Choi and Giovanni Zocchi. Guanylate kinase, induced fit, and the allosteric spring probe. *Biophysical Journal*, 92:1651–8, March 2007. PMID: 17142284. [cited at p. 44]

[24] P Cluzel, A Lebrun, C Heller, R Lavery, J L Viovy, D Chatenay, and F Caron. Dna: an extensible molecule. *Science (New York, N.Y.)*, 271:792–4, February 1996. PMID: 8628993. [cited at p. 77, 84]

[25] Emmanuelle Cogné-Laage. *Contributions de chimiste à l'imagerie microscopique par excitation biphotonique.* PhD thesis, Université Paris 6, October 2003. [cited at p. 56]

[26] D.B. Craig, E.A. Arriaga, J.C.Y. Wong, H. Lu, and N.J. Dovichi. Studies on single alkaline phosphatase molecules: Reaction rate and activation energy of a reaction catalyzed by a single molecule and the effect of thermal denaturation-the death of an enzyme. *Journal of the American Chemical Society*, 118:5245–5253, June 1996. [cited at p. 65]

[27] D M Crothers, T E Haran, and J G Nadeau. Intrinsically bent dna. *The Journal of Biological Chemistry*, 265:7093–6, May 1990. PMID: 2185240. [cited at p. 76]

[28] A A Deniz, M Dahan, J R Grunwell, T Ha, A E Faulhaber, D S Chemla, S Weiss, and P G Schultz. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of förster distance dependence and subpopulations. *Proceedings of the National Academy of Sciences of the United States of America*, 96:3670–5, March 1999. PMID: 10097095. [cited at p. 8, 9]

[29] Dittrich and Schwille. Photobleaching and stabilization of. fluorophores used for single-molecule analysis. with one- and two-photon excitation. *Applied Physics B: Lasers and Optics*, 73:829–837, 2001. [cited at p. 47, 59]

[30] Ashley C. Dyck and Douglas B. Craig. Individual molecules of thermostable alkaline phosphatase support different catalytic rates at room temperature. *Luminescence*, 17:15–18, 2002. [cited at p. 65]

[31] Lars Edman, Zeno Foldes-Papp, Stefan Wennmalm, and Rudolf Rigler. The fluctuating enzyme: a single molecule approach. *Chemical Physics*, 247:11–22, 1999. [cited at p. 9, 21, 40, 43, 45, 46, 133]

[32] Lars Edman and Rudolf Rigler. Memory landscapes of single-enzyme molecules. *Proceedings of the National Academy of Sciences*, page 130589397, 2000. [cited at p. 42]

[33] Christian Eggeling, Andreas Volkmer, and Claus A M Seidel. Molecular photobleaching kinetics of rhodamine 6g by one- and two-photon induced confocal fluorescence microscopy. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 6:791–804, May 2005. PMID: 15884061. [cited at p. 47, 59]

[34] Albert Einstein. On the motion of small particles suspended in liquids at rest required by the molecular-kinetik theory of heat. *Annalen*, 17:549–560, 1905. [cited at p. 6]

[35] Brian P English, Wei Min, Antoine M van Oijen, Kang Taek Lee, Guobin Luo, Hongye Sun, Binny J Cherayil, S C Kou, and X Sunney Xie. Ever-fluctuating single enzyme molecules: Michaelis-menten equation revisited. *Nat. Chem. Biol.*, 2:87–94, February 2006. [cited at p. 9, 41, 47, 60]

[36] B Essevaz-Roulet, U Bockelmann, and F Heslot. Mechanical separation of the complementary strands of dna. *Proceedings of the National Academy of Sciences of the United States of America*, 94:11935–40, October 1997. PMID: 9342340. [cited at p. 84]

[37] Ophir Flomenbom, Kelly Velonia, Davey Loos, Sadahiro Masuo, Mircea Cotlet, Yves Engelborghs, Johan Hofkens, Alan E Rowan, Roeland J M Nolte, Mark Van der Auweraer, Frans C de Schryver, and Joseph Klafter. Stretched exponential decay and correlations in the catalytic activity of fluctuating single lipase molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 102:2368–72, February 2005. PMID: 15695587. [cited at p. 9, 41, 44, 65]

[38] Scott Forth, Christopher Deufel, Maxim Y. Sheinin, Bryan Daniels, James P. Sethna, and Michelle D. Wang. Abrupt buckling transition observed during the plectoneme formation of individual dna molecules. *Physical Review Letters*, 100:148301–4, April 2008. [cited at p. 102, 103]

[39] H Frauenfelder, S G Sligar, and P G Wolynes. The energy landscapes and motions of proteins. *Science (New York, N.Y.)*, 254:1598–603, December 1991. PMID: 1749933. [cited at p. 9]

[40] E. Frey and K. Kroy. Brownian motion: a paradigm of soft matter and biological physics. *Annalen der Physik*, 14:20–50, 2005. [cited at p. 6]

[41] F B Fuller. The writhing number of a space curve. *Proceedings of the National Academy of Sciences of the United States of America*, 68:815–9, April 1971. PMID: 5279522. [cited at p. 73]

[42] Charlie Gosse and Vincent Croquette. Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophysical journal*, 82:3314–29, June 2002. PMID: 12023254. [cited at p. 85, 86]

[43] William J Greenleaf, Michael T Woodside, and Steven M Block. High-resolution, single-molecule measurements of biomolecular motion. *Annual Review of Biophysics and Biomolecular Structure*, 36:171–90, 2008. PMID: 17328679. [cited at p. 83, 84]

[44] David G. Grier. A revolution in optical manipulation. *Nat Photon*, 424:810–816, 2003. [cited at p. 84, 85]

[45] T Ha. Single-molecule fluorescence resonance energy transfer. *Methods (San Diego, Calif.)*, 25:78–86, September 2001. PMID: 11558999. [cited at p. 44]

[46] T. Ha, T. Enderle, D F Ogletree, D S Chemla, P R Selvin, and S. Weiss. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences of the United States of America*, 93:6264?6268, June 1996. PMC39010. [cited at p. 44]

[47] Jeffrey A Hanson, Karl Duderstadt, Lucas P Watkins, Sucharita Bhattacharyya, Jason Brokaw, Jhih-Wei Chu, and Haw Yang. Illuminating the mechanistic roles of enzyme conformational dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 104:18055–60, November 2007. PMID: 17989222. [cited at p. 9]

[48] Kai Hassler, Marcel Leutenegger, Per Rigler, Ramachandra Rao, Rudolf Rigler, Michael G?sch, and Theo Lasser. Total internal reflection fluorescence correlation spectroscopy (tir-fcs) with low background and high count-rate per molecule. *Optics Express*, 13:7415–7423, 2005. [cited at p. 53, 62]

[49] Katrin G. Heinze, Andre Koltermann, and Petra Schwille. Simultaneous two-photon excitation of distinct labels for dual-color fluorescence crosscorrelation analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 97:10377?10382, September 2000. PMC27032. [cited at p. 47]

[50] Katherine Henzler-Wildman and Dorothee Kern. Dynamic personalities of proteins. *Nature*, 450:964–72, December 2007. PMID: 18075575. [cited at p. 9]

[51] Katherine A Henzler-Wildman, Ming Lei, Vu Thai, S Jordan Kerns, Martin Karplus, and Dorothee Kern. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, 450:913–6, December 2007. PMID: 18026087. [cited at p. 9]

[52] Katherine A Henzler-Wildman, Vu Thai, Ming Lei, Maria Ott, Magnus Wolf-Watz, Tim Fenn, Ed Pozharski, Mark A Wilson, Gregory A Petsko, Martin Karplus, Christian G Hübner, and Dorothee Kern. Intrinsic motions along an enzymatic reaction trajectory. *Nature*, 450:838–44, December 2007. PMID: 18026086. [cited at p. 9]

[53] X Hoang, A Trovato, F Seno, R Banavar, and A Maritan. Geometry and symmetry presculpt the free-energy landscape of proteins. *Proc Natl Acad Sci U S A*, 2004. [cited at p. 13]

[54] Vijay Iyer, Molly J. Rossow, and M. Neal Waxham. Peak two-photon molecular brightness of fluorophores is a robust measure of quantum efficiency and photostability. *Journal of the Optical Society of America B*, 23:1420–1433, July 2006. [cited at p. 47]

[55] K V Klenin, A V Vologodskii, V V Anshelevich, A M Dykhne, and M D Frank-Kamenetskii. Computer simulation of dna supercoiling. *Journal of molecular biology*, 217:413–9, February 1991. PMID: 1994032. [cited at p. 81]

[56] R D Kornberg and Y Lorch. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, 98:285–94, August 1999. PMID: 10458604. [cited at p. 10]

[57] Oleg Krichevsky and Gr?goire Bonnet. Fluorescence correlation spectroscopy: the technique and its applications. *Reports of Progress in Physics*, 65:251–297, February 2002. [cited at p. 53]

[58] J F Leger, J Robert, L Bourdieu, D Chatenay, and J F Marko. Reca binding to a single double-stranded dna molecule: a possible role of dna conformational fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, 95:12295–9, October 1998. PMID: 9770480. [cited at p. 84]

[59] J. F. Leger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, and J. F. Marko. Structural transitions of a twisted and stretched dna molecule. *Physical*

Review Letters, 83:1066, 1999. Copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org. [cited at p. 10]

[60] Cyrus Levinthal. Are there pathways to protein folding ? *Journal de Chimie Physique*, 65:44, 1968. [cited at p. 12]

[61] Timothée Lionnet. *Mécanique de l'ADN et étude des hélicases en molécule unique.* PhD thesis, Université Paris 6, 2006. [cited at p. 15, 72, 75, 83, 86, 92, 133, 134, 135]

[62] Timothée Lionnet, Alexandre Dawid, Sarah Bigot, François-Xavier Barre, Omar A Saleh, François Heslot, Jean-François Allemand, David Bensimon, and Vincent Croquette. Dna mechanics as a tool to probe helicase and translocase activity. *Nucleic Acids Res*, 34:4232–4244, 2006. [cited at p. 10]

[63] Timothée Lionnet, Sylvain Joubaud, Richard Lavery, David Bensimon, and Vincent Croquette. Wringing out dna. *Physical Review Letters*, 96:178102, May 2006. PMID: 16712339. [cited at p. 18]

[64] H P Lu, L Xun, and X S Xie. Single-molecule enzymatic dynamics. *Science (New York, N.Y.)*, 282:1877–82, December 1998. PMID: 9836635. [cited at p. 21, 40, 41, 42, 57]

[65] John F Marko. Torque and dynamics of linking number relaxation in stretched supercoiled dna. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 76:021926, August 2007. PMID: 17930084. [cited at p. 78, 79, 80, 81, 102, 135]

[66] John F Marko and Eric D Siggia. Stretching dna. *Macromolecules*, 28:8759–8770, 1995. [cited at p. 76]

[67] Jeffrey R Moffitt, Yann R Chemla, Steven B Smith, and Carlos Bustamante. Recent advances in optical tweezers. *Annual Review of Biochemistry*, 77:205–228, July 2008. PMID: 18307407. [cited at p. 84, 85]

[68] J D Moroz and P Nelson. Torsional directed walks, entropic elasticity, and dna twist stiffness. *Proceedings of the National Academy of Sciences of the United States of America*, 94:14418–22, December 1997. PMID: 9405627. [cited at p. 80, 82]

[69] J.D. Moroz and P. Nelson. Entropic elasticity of twist-storing polymers. *Macromolecules*, 31:6333–6347, 1998. [cited at p. 80]

[70] R. H. Morse and R. T. Simpson. Dna in the nucleosome. *Cell*, 54:285–287, 1988. [cited at p. 10]

[71] Philip Nelson. Sequence-disorder effects on dna entropic elasticity. *Physical Review Letters*, 80:5810, June 1998. Copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org. [cited at p. 76]

[72] Keir C Neuman and Attila Nagy. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods*, 5:491–505, June 2008. PMID: 18511917. [cited at p. 83, 84, 85]

[73] Ovryn and Izen. Imaging of transparent spheres through a planar interface using a high-numerical-aperture optical microscope. *J Opt Soc Am A Opt Image Sci Vis*, 17:1202–1213, 2000. [cited at p. 88]

[74] Q. A. Pankhurst, J. Connolly, S. K. Jones, and J. Dobson. Applications of magnetic nanoparticles in biomedicine. *Journal of Physics D: Applied Physics*, 36:R167–R181, 2003. [cited at p. 89]

[75] S. Panyukov and Y. Rabin. Thermal fluctuations of elastic filaments with spontaneous curvature and torsion. *Physical Review Letters*, 85:2404, 2000. Copyright (C) 2008 The American Physical Society; Please report any problems to prola@aps.org. [cited at p. 104]

[76] James Pawley. *Handbook of Biological Confocal Microscopy*. Springer, 2nd ed. edition, March 1995. [cited at p. 50]

[77] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2 edition, October 1992. [cited at p. 91]

[78] Boris Rotman. Measurement of activity of single molecules of ? -d-galactosidase. *Proceedings of the National Academy of Sciences of the United States of America*, 47:1981–1991, December 1961. ArticleType: primary_article / Full publication date: Dec. 15, 1961 / Copyright ? 1961 National Academy of Sciences. [cited at p. 39]

[79] Boris Rotman. *Partial loss of activity of individual molecules of aged ?-galactosidase*, page 279?289. Beckwith J R, Zipser D. , editors. The lactose operon. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory, 1970. [cited at p. 39]

[80] Rahul Roy, Sungchul Hohng, and Taekjip Ha. A practical guide to single-molecule fret. *Nat Meth*, 5:507–516, June 2008. [cited at p. 8, 44]

[81] Sophie Sacquin-Mora, Emilie Laforet, and Richard Lavery. Locating the active sites of enzymes using mechanical properties. *Proteins*, 67:350–9, May 2007. PMID: 17311346. [cited at p. 66]

[82] P. G. Saffman. Low reynolds number hydrodynamics. by j. happel & howard brenner. prentice-hall, 1965. 553 pp. ?6. *Journal of Fluid Mechanics Digital Archive*, 28:826–828, 2006. [cited at p. 90]

[83] Jue Shi, Bruce A Palfey, Joe Dertouzos, Kaj Frank Jensen, Ari Gafni, and Duncan Steel. Multiple states of the tyr318leu mutant of dihydroorotate dehydrogenase revealed by single-molecule kinetics. *Journal of the American Chemical Society*, 126:6914–22, June 2004. PMID: 15174861. [cited at p. 41, 57]

[84] N M Shnerb, Y Louzoun, E Bettelheim, and S Solomon. The importance of being discrete: life always wins on the surface. *Proceedings of the National Academy of Sciences of the United States of America*, 97:10322–4, September 2000. PMID: 10962027. [cited at p. 6]

[85] S B Smith, Y Cui, and C Bustamante. Overstretching b-dna: the elastic response of individual double-stranded and single-stranded dna molecules. *Science (New York, N.Y.)*, 271:795–9, February 1996. PMID: 8628994. [cited at p. 76]

[86] S B Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science (New York, N.Y.)*, 258:1122–6, November 1992. PMID: 1439819. [cited at p. 83, 86]

[87] T R Strick. *Enroulement mécnique de l'ADN et relaxation par les topoisomérases.* PhD thesis, Université Paris 6, 1996. [cited at p. 86]

[88] T R Strick, J F Allemand, D Bensimon, A Bensimon, and V Croquette. The elasticity of a single supercoiled dna molecule. *Science (New York, N.Y.)*, 271:1835–7, March 1996. PMID: 8596951. [cited at p. 10, 86]

[89] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette. Behavior of supercoiled dna. *Biophysical journal*, 74:2016–2028, April 1998. [cited at p. 10]

[90] T. R. Strick, V. Croquette, and D. Bensimon. Single-molecule analysis of dna uncoiling by a type ii topoisomerase. *Nature*, 404:901–904, April 2000. [cited at p. 10]

[91] K Svoboda and S M Block. Biological applications of optical forces. *Annual review of biophysics and biomolecular structure*, 23:247–85, 1994. PMID: 7919782. [cited at p. 84, 90]

[92] Nancy L Thompson and Bridgett L Steele. Total internal reflection with fluorescence correlation spectroscopy. *Nat. Protocols*, 2:878–890, April 2007. [cited at p. 53]

[93] Makio Tokunaga, Kazuo Kitamura, Kiwamu Saito, Atsuko Hikikoshi Iwane, and Toshio Yanagida. Single molecule imaging of fluorophores and enzymatic reactions achieved by objective-type total internal reflection fluorescence microscopy. *Biochemical and Biophysical Research Communications*, 235:47–53, 1997. [cited at p. 62]

[94] Kelly Velonia, Ophir Flomenbom, Davey Loos, Sadahiro Masuo, Mircea Cotlet, Yves Engelborghs, Johan Hofkens, Alan E Rowan, Joseph Klafter, Roeland J M Nolte, and Frans C de Schryver. Single-enzyme kinetics of calb-catalyzed hydrolysis. *Angewandte Chemie (International ed. in English)*, 44:560–4, 2005. PMID: 15619259. [cited at p. 41]

[95] Mario B. Viani, Tilman E. Schaffer, Ami Chand, Matthias Rief, Hermann E. Gaub, and Paul K. Hansma. Small cantilevers for force spectroscopy of single molecules. *Journal of Applied Physics*, 86:2258–2262, 1999. [cited at p. 84]

[96] A V Vologodskii, S D Levene, K V Klenin, M Frank-Kamenetskii, and N R Cozzarelli. Conformational and thermodynamic properties of supercoiled dna. *Journal of molecular biology*, 227:1224–43, October 1992. PMID: 1433295. [cited at p. 81]

[97] A V Vologodskii and J F Marko. Extension of torsionally stressed dna by external force. *Biophysical Journal*, 73:123–32, July 1997. PMID: 9199777. [cited at p. 82]

[98] T. A. Waigh. Microrheology of complex fluids. *Reports on Progress in Physics*, 68:685–742, 2005. [cited at p. 7]

[99] Nils G Walter, Cheng-Yen Huang, Anthony J Manzo, and Mohamed A Sobhy. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature Methods*, 5:475–89, June 2008. PMID: 18511916. [cited at p. 8]

[100] S. Weiss. Fluorescence spectroscopy of single biomolecules. *SCIENCE*, 283:1676–1683, March 1999. [cited at p. 8]

[101] Adrie H Westphal, Andrey Matorin, Mark A Hink, Jan Willem Borst, Willem J H van Berkel, and Antonie J W G Visser. Real-time enzyme dynamics illustrated with fluorescence spectroscopy of p-hydroxybenzoate hydroxylase. *The Journal of biological chemistry*, 281:11074–81, April 2006. PMID: 16492664. [cited at p. 43]

[102] James H. White. Self-linking and the gauss integral in higher dimensions. *American Journal of Mathematics*, 91:693–728, July 1969. [cited at p. 72]

[103] Hai-Young Wu, Shihua Shyy, James C. Wang, and Leroy F. Liu. Transcription generates positively and negatively supercoiled domains in the template. *Cell*, 53:433–440, May 1988. [cited at p. 10]

[104] X S Xie and H P Lu. Single-molecule enzymology. *The Journal of biological chemistry*, 274:15967–70, June 1999. PMID: 10347141. [cited at p. 42]

[105] Chris Xu and Watt W. Webb. Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm. *Journal of the Optical Society of America B Optical Physics*, 13:481–491, March 1996. [cited at p. 51]

[106] J Yan, M O Magnasco, and J F Marko. A kinetic proofreading mechanism for disentanglement of dna by topoisomerases. *Nature*, 401:932–5, October 1999. PMID: 10553912. [cited at p. 72]

[107] J. L. Yang, M. Despont, U. Drechsler, B. W. Hoogenboom, P. L. T. M. Frederix, S. Martin, A. Engel, P. Vettiger, and H. J. Hug. Miniaturized single-crystal silicon cantilevers for scanning force microscopy. *Applied Physics Letters*, 86:134101–3, March 2005. [cited at p. 84]

[108] Xiaowei Zhuang, Harold Kim, Miguel J. B. Pereira, Hazen P. Babcock, Nils G. Walter, and Steven Chu. Correlating structural dynamics and function in single ribozyme molecules. *Science*, 296:1473–1476, May 2002. [cited at p. 44]

[109] Warren R Zipfel, Rebecca M Williams, and Watt W Webb. Nonlinear magic: multiphoton microscopy in the biosciences. *Nature biotechnology*, 21:1369–77, November 2003. PMID: 14595365. [cited at p. 56, 134]

Appendices

Appendix A

Technical references

A.1 Magnetic field decomposition

A.1.1 Basis

$$\begin{pmatrix} \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{1}{\sqrt{3}} \\ -\frac{1}{3} & \frac{1}{\sqrt{3}} \end{pmatrix} \tag{A.1}$$

$$R = \text{PseudoInverse}[A];\text{MatrixForm}[\%] \tag{A.2}$$

$$\begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \tag{A.3}$$

$$R.A;\text{MatrixForm}[\%] \tag{A.4}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{A.5}$$

$$\text{Transpose}[R].\text{Transpose}[A];\text{MatrixForm}[\%] \tag{A.6}$$

$$\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix} \tag{A.7}$$

A.2 Buffers and protocols

HRP buffer

- 50mM Sodium Phosphate Buffer
- 0.2 % BSA (W384A, Promega, France)
- 0.2 % F-127 Puronic surfactant (P2443-250G, Sigma-Aldrich, Germany)
- H_20_2 (needed concentration)

Bead Passivation buffer

- PBS
- 0.2 % BSA (W384A, Promega, France)
- 0.2 % F-127 Puronic surfactant (P2443-250G, Sigma-Aldrich, Germany)
- 5 mM EDTA
- 10mM Sodium Azide (S-8032, Sigma-Aldrich, Germany)

Standard buffer

- 10mM Sodium Phosphate buffer pH 8.0
- 0.1 % BSA (W384A, Promega, France)
- 0.1 % F-127 Puronic surfactant (P2443-250G, Sigma-Aldrich, Germany)
- 1mM EDTA
- 10mM Sodium Azide (S-8032, Sigma-Aldrich, Germany)

TE NaCl buffer

- 10mM Tris HCl pH 8.0
- 1mM EDTA
- 0.1 % BSA (W384A, Promega, France)
- 0.1 % F-127 Puronic surfactant (P2443-250G, Sigma-Aldrich, Germany)
- NaCl (needed concentration)

MgCl2 buffer

- 10mM Tris HCl pH 8.0
- 0.1 % BSA (W384A, Promega, France)
- 0.1 % F-127 Puronic surfactant (P2443-250G, Sigma-Aldrich, Germany)
- MgCl2 (needed concentration)

A.2.1 DNA functionalization protocol

A.3 Setup alignment

A.3.1 0.0 alignment

Excitation alignment

Objective focus

A drop of water (40 μL) is deposed on the objective and the sample is placed on the scanning stage. The objective is brought at 5mm distance from the lower surface of the bottom glass coverslide of the sample. It is then slosly displaced towards the sample while monitoring the fluorescence counts detected by the APDs. Dark counts are of about 500Hz for each APD. When the focal spot reaches the glass coverslide the signal increases to a value of more than 10 kHz due to the reflections of IR light from the glass surface. Inside the glass coverslide the counts are again low (1-2kHz). Beyond the interface glass/water the counts depend on the solution contained in the sample. For typical solutions of 50-100nM of Fluorescein or Rhodamine-B used for calibration, a countrate of 60-80Hz per detector was reached.

Detection alignment

A.3.2 1.0 alignment

A.3.3 2.0 alignment

List of Figures

List of Tables