



**HAL**  
open science

## Mécanismes de diffusion facilitée de l'enzyme de restriction EcoRV.

Isabelle Bonnet

► **To cite this version:**

Isabelle Bonnet. Mécanismes de diffusion facilitée de l'enzyme de restriction EcoRV.. Biophysique [physics.bio-ph]. Université Pierre et Marie Curie - Paris VI, 2007. Français. NNT : . tel-00197580

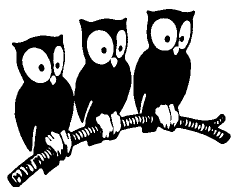
**HAL Id: tel-00197580**

**<https://theses.hal.science/tel-00197580>**

Submitted on 14 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Laboratoire Kastler Brossel  
Département de Physique  
École Normale Supérieure



## THÈSE de DOCTORAT de l'UNIVERSITÉ PARIS 6

Spécialité : Physique des Liquides

présentée par

**Isabelle BONNET**

---

# Mécanismes de diffusion facilitée de l'enzyme de restriction EcoRV

---

Soutenue le 30 octobre 2007

devant le jury composé de :

M. Denis Côte	.....	Président
M. Antoine Delon	.....	Rapporteur
M. Brahim Lounis	.....	Rapporteur
M. Olivier Bénichou	.....	Examineur
M. Wolfgang Wende	.....	Examineur
M. Pierre Desbiolles	.....	Directeur de thèse



# Table des matières

Notations	7
Introduction	9
<b>I Enjeux Biologiques</b>	<b>13</b>
I.1 Les interactions ADN/protéines . . . . .	13
I.1.1 Reconnaissance ADN-Protéine . . . . .	14
I.1.2 Diversité . . . . .	14
I.1.3 Modélisation . . . . .	15
I.1.4 Les techniques d'étude <i>in vitro</i> . . . . .	15
I.1.5 Importance . . . . .	16
I.2 Quand les protéines doivent trouver leur cible sur l'ADN.... . . . .	17
I.2.1 ... la localisation de la cible est très efficace . . . . .	17
I.2.2 ... l'ADN non spécifique joue un rôle . . . . .	18
I.2.3 ... les protéines usent de mécanismes de diffusion facilitée . . . . .	18
I.3 Notre protéine d'investigation : l'enzyme de restriction EcoRV . . . . .	22
I.3.1 Les enzymes de restriction . . . . .	22
I.3.2 EcoRV . . . . .	25
I.3.3 Comment EcoRV trouve-t-elle ses sites? . . . . .	27
<b>II Mise en place du dispositif expérimental</b>	<b>33</b>
II.1 Étudier les interactions ADN-protéine à l'échelle de la Molécule Unique . .	33
II.1.1 Motivations : s'affranchir du moyennage . . . . .	33
II.1.2 Dispositif : le montage de TIRFM . . . . .	34
II.1.3 Aperçu des autres approches moléculaires . . . . .	44
II.2 Rendre les enzymes fluorescentes . . . . .	48
II.2.1 Un lourd cahier des charges . . . . .	48
II.2.2 Des enzymes mutées pour un couplage facilité . . . . .	49
II.2.3 Les différents couplages . . . . .	49
II.2.4 Réussite du couplage et contrôle de l'activité . . . . .	51
II.3 Positionner correctement une molécule d'ADN . . . . .	58

II.3.1	Nécessité d'étirer les molécules d'ADN . . . . .	58
II.3.2	L'étirement spécifique de molécules d'ADN . . . . .	59
II.4	Visualiser l'interaction ADN - EcoRV . . . . .	67
II.4.1	Dispositif Expérimental . . . . .	67
<b>III</b>	<b>Étude de la diffusion facilitée d'EcoRV</b>	<b>71</b>
III.1	Démarche expérimentale . . . . .	71
III.1.1	Stratégie . . . . .	71
III.1.2	Protocole . . . . .	73
III.1.3	Acquisition . . . . .	75
III.2	EcoRV glisse le long de l'ADN . . . . .	77
III.2.1	La MSD : un outil d'analyse . . . . .	77
III.2.2	La diffusion 1D d'EcoRV-Cy3B . . . . .	82
III.2.3	Pooling des données . . . . .	83
III.2.4	Influence du tampon et du pH . . . . .	87
III.3	EcoRV saute aussi . . . . .	90
III.3.1	EcoRV saute . . . . .	90
III.3.2	La distribution des sauts . . . . .	94
III.3.3	Pourquoi EcoRV saute t-elle? . . . . .	94
III.3.4	Modélisation des sauts . . . . .	98
III.4	Conséquences de l'existence des petits sauts . . . . .	101
III.4.1	Temps d'interaction . . . . .	101
III.4.2	Longueur de sliding . . . . .	103
III.4.3	Coefficient de diffusion 1D . . . . .	104
III.4.4	Prévisions - Réalité <i>in vivo</i> . . . . .	104
III.5	Effet de la taille du fluorophore . . . . .	105
III.5.1	Théorie - modèle de Schurr : . . . . .	105
III.5.2	Résultats expérimentaux : . . . . .	105
<b>A</b>	<b>Fabrication d'un ADN de 9,2 kb aux extrémités biotinilées ne comportant aucun site spécifique pour EcoRV.</b>	<b>115</b>
A.1	Principe . . . . .	115
A.2	Choix d'un plasmide support : le T7 . . . . .	115
A.3	Fabrication des fragments biotinilés . . . . .	116
A.4	Ligation . . . . .	117
<b>B</b>	<b>Programme de tracking</b>	<b>119</b>
B.1	Répérer un spot fluorescent . . . . .	119
B.2	Description de l'interface . . . . .	120

<b>C</b>	<b>Fluctuations de l'ADN étiré</b>	<b>123</b>
C.1	Modèle billes-ressorts (modèle de Rouse)	123
C.2	Modes normaux d'oscillation	124
C.3	Application à l'ADN	124
C.3.1	Élasticité entropique	124
C.3.2	Linéarisation	125
C.3.3	Coefficient de friction	126
C.3.4	Modes normaux d'oscillation :	126
C.3.5	Fluctuations spatiales	126
C.4	Détermination expérimentale	127
C.4.1	Temps de corrélation	127
C.4.2	Amplitude des fluctuations	127
C.4.3	Conclusion	128
	<b>Bibliographie</b>	<b>129</b>



## Notations

ADN	Acide DésoxyriboNucléique
AFM	Atomic Force Microscope : Microscope à Force Atomique
APD	Avalanche Photodiode : Photodiode à avalanche
APTES	3-AminoPropyl TriEthoxy Silane
bp	Base Paire : Paire de Base
CCD	Charge Coupled Device
Da	Dalton - $1 \text{ Da} = 1 \text{ g.mol}^{-1}$
kDa	KiloDalton
DTT	Dithiothreitol
FCS	Fluorescent Correlation Spectroscopy : Microscopie de Corrélation de Fluorescence
FRAP	Fluorescence Recovery After Photobleaching : Retour de Fluorescence Après Photoblanchiment
$K_A, K_D$	Constantes thermodynamiques d'association et de dissociation
$k_B$	Constante de Boltzmann = $1,38.10^{-23} \text{ m}^2.\text{kg}.\text{s}^{-2}.\text{K}^{-1}$
$k_{on}, k_{off}$	Constantes cinétiques d'association et de dissociation d'un couple récepteur-ligand
$\lambda_A$	Longueur d'onde maximale d'absorption
$\lambda_E$	Longueur d'onde maximale d'émission
kp ou kbp	Kilo Paires de Base
$MW$	Molecular Weight : Masse Molaire
MET	Microscope Électronique à Transmission
MFPP	Modulated fringe pattern photobleaching
$NA$	Numerical Aperture : Ouverture Numérique
PBS	Phosphate Buffer Saline
PB	Phosphate Buffer
PCR	Polymerase Chain Reaction : amplification en chaîne par polymérase
PSF	Point Spread Function : réponse impulsionnelle du système
QD	Quantum Dot : nanocristal semi-conducteur
$R_G$	Rayon de Giration d'un polymère
RMN	Résonance Magnétique Nucléaire
Sav	Streptavidine
SPT	Single Particle Tracking : Suivi de Molécule Unique
TCEP	Tris[2-carboxyethyl] phosphine
TIRFM	Total Internal Reflection Fluorescence Microscopy : Microscopie de fluorescence par onde évanescente
U.V	Ultra - Violet





# Introduction

La biologie, « science du vivant », est une science jeune. Son histoire débute au milieu du XIX<sup>e</sup> siècle, avec la rationalisation des connaissances et l'abandon de la théorie vitaliste, qui faisait appel à une mystérieuse énergie vitale afin d'expliquer le vivant. L'application réussie des outils de la physique (microscopie) et de la chimie (analyse et identification) donnait naissance à ce qu'on appela plus tard la biologie moléculaire. La découverte en 1953 de la structure de l'ADN par diffraction des rayons X fut un bouleversement dans l'étude des phénomènes biologiques : la dimension moléculaire y est introduite.

L'organisme vivant est depuis décrit comme un assemblage très complexe de molécules interagissant entre elles selon les lois de la physique et de la chimie. L'ADN (support de l'information génétique), l'ARN (structure transitoire permettant sa transmission à une machine de traduction) et les protéines (constituants de bases qui font fonctionner l'organisme) sont les trois acteurs principaux de ces réactions biophysiques et biochimiques. La compréhension du fonctionnement de ces interactions est par conséquent essentielle d'un point de vue fondamental et appliqué puisqu'elle peut contribuer au développement de molécules thérapeutiques. Elle n'en reste pas moins un problème biologique majeur car l'activité cellulaire repose en fait sur de véritables cascades d'interactions extrêmement complexes faisant même intervenir des boucles de rétroaction.

Même si des études à l'interface physique-biologie sont menées depuis de nombreuses années, ce n'est que récemment que les physiciens travaillent sur des systèmes à l'échelle de la molécule unique grâce à l'apparition des techniques de micromanipulation de molécules d'ADN. Ces techniques ont initialement servi à sonder les propriétés mécaniques du polymère ADN. Elles permettent aujourd'hui, combinées à des techniques de mesure de forces ou à la microscopie de fluorescence, d'étudier la dynamique d'interaction ADN - protéine.

Au cours de cette thèse, nous nous sommes particulièrement intéressés à la dynamique de recherche de cibles par les protéines. Parmi les nombreuses protéines qui interagissent avec l'ADN, certaines sont site-spécifiques et sont capables de trouver très rapidement leur cible, constituée de quelques paires de bases, sur de l'ADN comportant plusieurs di-

zaines de milliers de paires de base. C'est le cas des enzymes de restriction de type II, protéines présentes dans les bactéries et destinées à lutter contre les attaques virales. Les biochimistes ont multiplié les expériences pour tenter d'expliquer ce surprenant constat. Ils sont parvenus à montrer que les protéines interagissent avec de l'ADN non-spécifique pour trouver leur cible sur l'ADN : on parle de « diffusion facilitée ». Le mécanisme sous-jacent de la diffusion facilitée n'est toujours pas élucidé, il est généralement admis que les protéines alternent entre des processus de « sliding » (diffusion unidimensionnelle le long du sillon de l'ADN) et des processus de « jumping » (succession de saut). Mais quelle est l'importance relative de ces deux processus ?

Pour répondre à cette question, nous avons décidé d'étudier l'interaction non-spécifique de l'enzyme de restriction EcoRV avec l'ADN. EcoRV a été étudiée en profondeur par les biochimistes (mesures cinétiques). Ils ont également étudié certains aspects de la diffusion facilitée d'EcoRV : mise en évidence du rôle de l'ADN non spécifique, mise en évidence du sliding, et, plus récemment, du jumping. Néanmoins, ces résultats reposent sur l'interprétation de données expérimentales impliquant des modèles controversés. En tant que physiciens, nous avons cherché à observer directement la diffusion facilitée d'EcoRV sur de l'ADN non spécifique. Pour ce faire, nous avons combiné les techniques de micromanipulation d'ADN et celles de microscopie de fluorescence par onde évanescente. Dans nos expériences *in vitro*, l'ADN est étiré au dessus d'une surface et les enzymes sont couplées à des fluorophores. Nous sommes ainsi parvenus à visualiser l'interaction d'EcoRV sur une unique molécule d'ADN. L'étude de cette dynamique d'interaction a permis de mettre en évidence la diffusion facilitée d'EcoRV.

Ce mémoire s'articule sur 3 chapitres.

★ Le premier chapitre intitulé « Enjeux Biologiques » présente la problématique biologique liée à la localisation rapide de séquences cibles par les protéines. Il débute par une courte description des interactions ADN-protéines, puis décrit l'état de l'art des différentes études faites par les biochimistes sur les enzymes de restriction. Ces travaux, basés sur des mesures d'ensemble, ont permis d'avancer différents mécanismes possibles de diffusion facilitée et apportent les premiers éléments de réponse. Néanmoins, ils ne suffisent pas à élucider clairement les mécanismes mis en jeu, légitimant notre approche à l'échelle de la molécule individuelle.

★ Le chapitre suivant détaille la mise en œuvre du dispositif expérimental. La première partie de ce chapitre fait la transition avec le premier puisqu'elle passe en revue les différentes approches pour étudier les interactions ADN-protéines à l'échelle de la molécule individuelle. Nous y justifierons le choix de la microscopie de fluorescence par onde éva-

nescente. Les deuxième et troisième parties sont consacrées à la réalisation expérimentale proprement dite : le couplage des enzymes aux fluorophores d'une part et l'étirement des molécules d'ADN d'autre part.

★ Le dernier chapitre est consacré aux résultats obtenus durant ce travail de thèse concernant la diffusion facilitée d'EcoRV. Le résultat principal est l'observation directe de la diffusion facilitée de l'enzyme, qui apparaît comme une combinaison de sliding et de jumping. Les paramètres physiques du sliding (coefficient de diffusion, durée du processus et longueur d'ADN explorée) sont mesurés et discutés. Le jumping observé est identifié expérimentalement à des excursions 3D. La distribution de sauts obtenue expérimentalement est ensuite confrontée à un modèle que nous avons développé en collaboration avec des théoriciens (C. Loverdo - O. Bénichou, LPTL, UPMC). Dans ce modèle, une enzyme qui effectue un saut correspond à une enzyme, qui initialement sur l'ADN, se dissocie de l'ADN, diffuse à 3D dans l'espace puis se réassocie plus loin sur l'ADN. Ce modèle simple permet de rendre compte de nos observations expérimentales, et de donner une vision quantitative synthétique de la diffusion facilitée d'EcoRV.

Nos résultats apportent une contribution importante à l'étude de la diffusion facilitée de l'enzyme de restriction EcoRV. Plus loin, les méthodes mises au point pour étirer l'ADN et observer la dynamique d'interaction d'enzymes fluorescentes avec cet ADN ainsi que notre analyse reposant sur un modèle original pourront s'appliquer à l'étude d'autres protéines interagissant avec l'ADN.



# Chapitre I

## Enjeux Biologiques

*Au cœur de la cellule, les interactions ADN-protéine sont incessantes et essentielles à son bon fonctionnement. Lors de ces interactions, il n'est pas rare que les protéines doivent localiser des courtes séquences de nucléotides particulières. Cette localisation a généralement lieu avec une efficacité surprenante, ce qui a poussé, il y a quelques années, les biochimistes à s'interroger sur les mécanismes à l'origine de cette localisation « accélérée ». De nombreux modèles susceptibles d'accélérer la localisation des séquences-cibles ont ainsi été proposés ; cependant, le mécanisme sous-jacent de localisation n'a toujours pas été identifié avec certitude.*

*Dans ce chapitre liminaire, nous exposerons les enjeux biologiques soulevés par cette question : comment les protéines localisent-elles leurs cibles sur l'ADN ? Pour cela, nous nous efforcerons de décrire brièvement les bases des interactions ADN-protéines. Nous présenterons ensuite des modèles proposés par les biochimistes pour rendre compte de l'efficacité des protéines qui doivent chercher une « cible ». Enfin, nous expliquerons pourquoi nous avons choisi l'enzyme de restriction EcoRV durant ce travail de thèse pour étudier les interactions ADN-protéines.*

### I.1 Les interactions ADN/protéines

La grande variété de protéines capables de se lier à l'ADN détermine en partie la grande diversité des interactions ADN-protéines. Les connaissances actuelles des mécanismes de ces interactions résultent d'approches pluridisciplinaires utilisant par exemple les outils de la chimie et de la physique. Parmi les techniques de biophysique, les techniques de micromanipulation d'ADN et d'imageries de molécules individuelles se sont particulièrement développées ces dernières années bousculant les approches plus traditionnelles des mesures d'ensemble des biochimistes et permettant d'apporter de nouveaux éléments de réponse (ces techniques seront présentées au Chapitre II).

## **1.1.1 Reconnaissance ADN-Protéine**

### **1.1.1.1 Reconnaissance**

Les protéines diffèrent entre elles par leur chaîne d'acides aminés qui leur confère des propriétés différentes (chargée positivement ou négativement, hydrophobe ou hydrophile) et une certaine structure tridimensionnelle. Les groupes restant accessibles dans la protéine repliée déterminent son affinité avec les autres entités biologiques : protéines, membrane, ADN.

Dans le cas des interactions ADN-protéines, différents types de liaison interviennent entre les acides aminés des protéines et les paires de bases de l'ADN : liaisons hydrogènes, liaisons de van der Waals, liaisons électrostatiques, ... Ces différentes liaisons définissent une affinité globale de la protéine pour l'ADN. Le type de liaison mis en jeu dans ces complexes ADN-protéines va jouer sur leur stabilité.

Pour résumer, l'interaction entre une protéine et l'ADN résulte d'une reconnaissance chimique (propriétés des liaisons) et structurale (accessibilité des sites).

### **1.1.1.2 Spécificité**

Les interactions ADN-protéine peuvent être classées en deux catégories : les interactions en une séquence spécifique (régulateurs transcriptionnels, histone-like bactériennes<sup>1</sup>, enzymes de restriction) ou non (topoisomérase, histones).

Lors des interactions spécifiques, les protéines reconnaissent une séquence bien précise, appelée séquence cible. Pour la majorité de ces protéines, cette séquence est reconnue du côté du grand sillon de la molécule d'ADN, plus riche en « information chimique » et là où les bases sont plus accessibles. La spécificité de reconnaissance est ainsi liée aux bases et acides aminés mis en jeu lors de l'interaction.

## **1.1.2 Diversité**

Les interactions ADN-protéine sont très diversifiées. Pour illustrer cette variété, nous pouvons citer quelques exemples, parmi les plus étudiés, de protéines interagissant avec l'ADN :

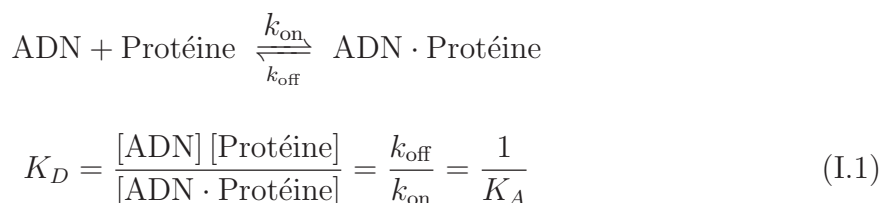
- les polymérases enchaînent des nucléotides en polymères d'ADN ou d'ARN.
- les histones structurent l'ADN.
- les facteurs de transcriptions régulent l'expression des gènes (activation ou inhibition).
- les enzymes de restriction reconnaissent et coupent une séquence particulière de l'ADN double brin.

---

<sup>1</sup>La bactérie utilise des protéines de type-histone pour replier l'ADN dans une structure condensée, appelée nucléoïde bactérien.

### I.1.3 Modélisation

L'interaction d'une protéine avec l'ADN est caractérisée par une constante d'équilibre thermodynamique d'association  $K_A$  ou de dissociation  $K_D = 1/K_A$ , reliées à des constantes cinétiques d'association  $k_{\text{on}}$  et de dissociation  $k_{\text{off}}$ . Le  $K_D$  est situé entre le nM et le  $\mu\text{M}$  en fonction de l'affinité de la protéine pour l'ADN. Dans le cas simple d'une protéine interagissant avec un seul site sur l'ADN,  $K_A$  est relié à la concentration en ADN, en protéine, et en complexe ADN·protéine selon :



$k_{\text{on}}$  est une constante de vitesse du second ordre (réaction bimoléculaire - s'exprime en  $\text{M}^{-1} \cdot \text{s}^{-1}$ ),  $k_{\text{off}}$  est une constante de vitesse du premier ordre (réaction monomoléculaire - s'exprime en  $\text{s}^{-1}$ ).

### I.1.4 Les techniques d'étude *in vitro*

D'une manière générale, l'étude des interactions protéine-acides nucléiques est assez délicate car elle nécessite le contrôle et l'observation (directe ou non) de l'association de deux molécules biologiques. Néanmoins, de nombreuses techniques de biochimie ont été utilisées pour caractériser *in vitro* les interactions ADN-protéine, et ce, bien avant le développement de ces dernières années des techniques de micromanipulation des molécules d'ADN. Faisons un rapide aperçu des techniques qui ont été utilisées pour étudier les interactions ADN-protéine.

#### I.1.4.1 Biochimie : mesures d'ensemble

Le lecteur désireux d'en savoir plus sur ces techniques pourra se référer aux chapitres 7 et 8 de [1].

**Les techniques de gels retard** furent les premiers outils utilisés pour étudier les interactions ADN-protéine. Elles se basent sur l'analyse des changements de mobilité électrophorétiques de l'ADN selon qu'il est ou non lié à la protéine d'intérêt. La stœchiométrie des complexes ADN-protéine et l'affinité relative d'une même protéine pour différents substrats ont ainsi été étudiées [2].

**Les techniques d'empreinte de digestion d'ADN** permettent d'analyser l'accessibilité d'une séquence par une protéine capable de couper l'ADN (enzyme) pour caractériser les nucléotides impliqués dans l'interaction ADN-protéine.



### I.1.4.2 Biologie structurale : détermination des structures

La cristallographie par rayon X ou la RMN permet de déterminer les structures des complexes ADN-protéine à l'échelle atomique. Ces méthodes, bien que lourdes à mettre en œuvre (nécessité d'avoir un cristal bien ordonné pour les rayons X, molécules purifiées pour la RMN) sont très informatives mais restent néanmoins inadaptées au suivi de la dynamique d'association par exemple.

### I.1.4.3 Biophysique : expériences à l'échelle de la molécule individuelle

Dans ces nouvelles approches en molécule unique, il faut distinguer les techniques de microscopie (fluorescence, MET, AFM) et les techniques de mesure de force (pinces optiques et magnétiques, AFM, micropipettes). Ces différentes techniques seront présentées et discutées au paragraphe II.1 du chapitre suivant.

## I.1.5 Importance

La diversité des interactions ADN-protéine est à elle seule une preuve de leur importance pour le fonctionnement biologique. De manière plus précise : i) elles régissent le stockage et la transmission de l'information génétique où les protéines participent à la condensation<sup>2</sup> (histones, histone-like bactériennes), à la réplication (polymérase) à la réparation (enzymes de réparation) ; ii) elles interviennent dans les mécanismes de l'expression génétique (transcription, décodage du code génétique) iii) elles sont impliquées dans le contrôle de l'expression génétique : régulation de la transcription, régulation de la traduction, inhibition de l'expression génétique et destruction du matériel génétique lors des infections virales. Il va de soi que la spécificité des reconnaissances est essentielle dans l'ensemble de ces mécanismes.

Élucider clairement le mécanisme de translocation des protéines le long de l'ADN est d'un intérêt majeur puisque ce mécanisme contrôle la cinétique des génétiques. Comprendre le mécanisme d'interaction ADN-protéine peut également être utile pour guider l'ingénierie d'enzymes modifiées, ou à plus long terme, l'ingénierie de nouvelles protéines régulatrices se fixant sur une séquence donnée.

---

<sup>2</sup>Il faut se souvenir qu'une molécule d'ADN a une longueur typique de l'ordre du  $\mu\text{m}$ . Une cellule ayant typiquement une taille de 1 à 10  $\mu\text{m}^3$ , la condensation est indispensable.

## 1.2 Quand les protéines doivent trouver leur cible sur l'ADN....

### 1.2.1 ... la localisation de la cible est très efficace

Les séquences cibles reconnues par les protéines sont généralement de courtes séquences constituées de quelques paires de bases sur des molécules d'ADN pouvant en compter quelques dizaines de milliers (ADN des bactériophages) voire des centaines de milliers (ADN chromosomique). C'est dire que le problème de la localisation du point de vue de la protéine est difficile. Il est néanmoins crucial que le temps de recherche des protéines soit optimisé afin d'assurer la survie des organismes. Des mesures cinétiques de réactivité ont d'ailleurs prouvé que certaines de ces protéines étaient très efficaces, parvenant à localiser leur cible avec des vitesses plus élevées que celles atteintes avec un simple processus de diffusion tridimensionnelle. Ainsi, dès 1970, Riggs *et al.* ont estimé la constante d'association du répresseur Lac sur l'ADN à :  $k_{\text{on}} = 10^{10} \text{ M}^{-1} \cdot \text{s}^{-1}$  [3], soit deux ordres de grandeur supérieure à celle estimée dans un modèle d'association régie par la diffusion 3D donnant :  $k_{\text{on}}^{3\text{D}} = 10^7 - 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$  [4]. Or, cette protéine ne consomme pas d'ATP, son mouvement ne peut donc être que diffusif. Cette observation a été reconduite sur de nombreuses autres protéines : comment rendre compte de ce constat paradoxal ?

#### Taux d'association maximal par diffusion tridimensionnelle :

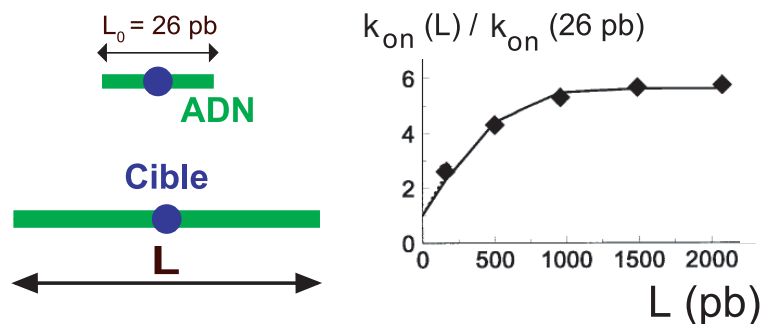
Considérons une protéine diffusive de masse moléculaire quelques dizaines de kDa. Cette protéine a une taille de l'ordre de  $\sim 5 \text{ nm}$ . Le mouvement de la protéine est une marche aléatoire à 3D de coefficient de diffusion  $D_3$ . Ainsi au bout d'un temps  $t$ , la protéine a parcouru une distance caractéristique  $r$  telle que :  $r^2 \sim D_3 t$ . Le coefficient  $D_3$  est donné par la loi d'Einstein :  $D_3 = \frac{k_B T}{6\pi\eta R}$  où  $R$  est le rayon de la protéine. Pour une protéine de diamètre 5 nm, le coefficient de diffusion tridimensionnel est par conséquent :  $D_3 = 100 \mu\text{m}^2 \cdot \text{s}^{-1}$ .

Berg a calculé la probabilité de collision entre une protéine diffusive et une cible, la protéine étant située à la distance  $r$  de sa cible de taille  $a$  :  $\mathcal{P} = \frac{a}{r}$  [5]. Il est possible de retrouver ce résultat en raisonnant en terme d'ordre de grandeur. L'espace de recherche est découpé en cube de côté  $a$ . La protéine explore une zone de taille  $a$  pendant le temps  $\frac{a^2}{D_3}$ , et une zone de taille  $r$  pendant le temps  $t_{\text{diff}} = \frac{r^2}{D_3}$ . On dénombre  $\left(\frac{r}{a}\right)^3$  zones de taille  $a$  dans une zone de taille  $r$ . Or, le nombre de sites de taille  $a$  explorés pendant  $t_{\text{diff}}$  est donné par  $\left(\frac{r}{a}\right)^2$ , on en déduit la probabilité de collision de la protéine avec sa cible :  $\mathcal{P} = \frac{a}{r}$ . En notant  $c$  la concentration en cibles dans la solution, chaque protéine est en moyenne à la distance  $r^3 = \frac{1}{c}$  d'une cible. Le temps nécessaire à la localisation d'une cible est :  $t = t_{\text{diff}} \times \frac{1}{\mathcal{P}} = \frac{1}{Dac}$ . On en déduit la constante d'association  $k_{\text{on}} = Da$ .

De manière plus rigoureuse, en utilisant l'équation de Debye-Smoluchowski, il a été montré que  $k_{\text{on}} = 4\pi Da$  [6, 7, 8]. En prenant  $a = 0,34 \text{ nm}$  la taille de la cible (taille d'une paire de base), sachant que 1 nM équivaut à 1 molécule/ $\mu\text{m}^3$ , il vient  $k_{\text{on}} = 3 \cdot 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$ . Il s'agit d'un taux d'association théorique, qui est une limite supérieure au taux d'association réel. En toute rigueur, il faudrait prendre en compte les contraintes liées aux conformations des biomolécules qui diminuent ce taux d'association maximal [9] ainsi que les effets électrostatiques qui modifient également ce taux d'association [10].

## I.2.2 ... l'ADN non spécifique joue un rôle

De nombreux biochimistes se sont alors intéressés à ce problème et ont expérimentalement montré que l'association des protéines à leur séquence cible est accélérée avec la quantité d'ADN non spécifique l'entourant (cf. Figure I.1). Sur la base de ces observations expérimentales, ils ont proposé le terme de « diffusion facilitée » signifiant que l'ADN non spécifique joue un rôle dans la localisation de cible spécifique. La plupart de ces études ont été menées sur le répresseur Lac [11] et sur les enzymes de restriction [12, 13, 14].



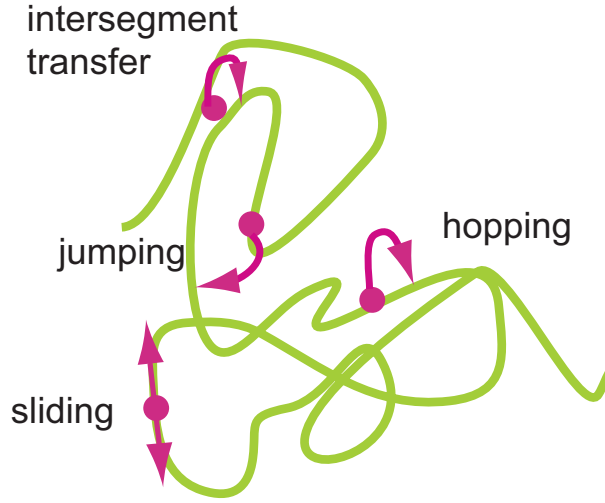
**Figure I.1:** Preuve expérimentale de l'existence d'un mécanisme de « diffusion facilitée ». Les enzymes de restriction de type II coupent l'ADN en un endroit spécifique appelé « séquence cible ». La mesure de la vitesse de coupure ( $k_{on}$ ) de molécules d'ADN comportant une unique séquence-cible entourée d'ADN non spécifique de taille variable  $L$  indique que la localisation de la cible est accélérée lorsque celle-ci est entourée d'ADN non-spécifique de plus en plus long. L'effet sature au delà de quelques centaines de paires de bases. Extrait de [15].

## I.2.3 ... les protéines utilisent des mécanismes de diffusion facilitée

Les biochimistes ont ensuite proposé différents mécanismes microscopiques de diffusion facilitée. Tous ces mécanismes incluent une première étape de diffusion de la protéine dans l'espace à 3 dimensions jusqu'à trouver l'ADN. La seconde étape correspond à la diffusion facilitée proprement dite et diffère selon les modèles. Nous allons décrire les mécanismes les plus couramment rencontrés dans la littérature. Ces mécanismes sont définis un peu différemment d'un article à l'autre : nous avons conservé la classification utilisée par Steven Halford [16, 17] et Alfred Pingoud [18]. Pour décrire ces mécanismes, illustrés sur la Figure I.2, nous avons gardé la terminologie anglo-saxonne.

### I.2.3.1 Le sliding

**Description :** le sliding (la traduction française serait glissement) correspond à une diffusion unidimensionnelle le long de l'ADN. Adam et Delbruck ont été les premiers à proposer une réduction de la dimension de l'espace de recherche (de 3 à 1) pour expliquer



**Figure I.2:** Les modèles de diffusion facilitée les plus souvent rencontrés dans la littérature.

l'efficacité de la protéine RecA [19]. La plupart des modèles de sliding présument que la protéine se déplace le long du grand sillon de l'ADN : à petite échelle, le mouvement est donc hélicoïdal. D'autres modèles ont récemment été proposés : un mouvement de diffusion 2D sur le cylindre d'ADN [20] ou encore une diffusion dans un petit volume autour de l'ADN [21]. Dans la suite de ce manuscrit, nous utiliserons le terme de sliding pour décrire une diffusion 1D de la protéine le long de l'ADN.

**Modélisation - efficacité :** le sliding au sens diffusion unidimensionnelle le long de l'ADN peut être modélisé par une marche aléatoire 1D symétrique de la protéine dont le pas est une paire de base, notée  $a = 0,34$  nm dans la suite. Plus précisément, après chaque pas, l'enzyme a une probabilité  $\mathcal{P}_{\text{off}}$  de se dissocier de l'ADN et une probabilité  $(1 - \mathcal{P}_{\text{off}})$  d'y rester et de faire un pas dans l'une ( $\mathcal{P}_-$ ) ou l'autre des directions ( $\mathcal{P}_+$ ) avec des probabilités identiques ( $\mathcal{P}_- = \mathcal{P}_+ = (1 - \mathcal{P}_{\text{off}})/2$ ).

$(1 - \mathcal{P}_{\text{off}})$  est déterminée par l'énergie de liaison protéine-ADN non spécifique et détermine le nombre moyen  $N$  de pas effectués avant la dissociation. En effet, la probabilité pour que la protéine reste accrochée à l'ADN à la suite de  $n$  pas de sliding est  $p(n) = A(1 - \mathcal{P}_{\text{off}})^n$  avec  $A$  une constante de normalisation

$$\int_{n=0}^{\infty} p(n) dn = 1 \implies A = -\ln(1 - \mathcal{P}_{\text{off}}),$$

Le nombre moyen  $N$  de pas est par conséquent égal à

$$N = \int_{n=0}^{\infty} np(n) dn = -\frac{1}{\ln(1 - \mathcal{P}_{\text{off}})}.$$

Dans l'hypothèse où  $\mathcal{P}_{\text{off}} \ll 1$ , il vient  $N = 1/\mathcal{P}_{\text{off}}$ . Le mouvement étant diffusif, le nombre de sites différents explorés par la protéine est de l'ordre de  $\sqrt{N}$ , on en déduit la

distance parcourue par sliding :  $\ell_{\text{sliding}} = \frac{a}{\sqrt{\mathcal{P}_{\text{off}}}}$ , à laquelle correspond un temps caractéristique de sliding  $\tau_{\text{sliding}} = \frac{\ell_{\text{sliding}}^2}{D_1} = \frac{1}{k_{\text{off}}}$  où  $D_1$  est le coefficient de diffusion [6, 22]. En 1998, Jeltsch et Pingoud ont proposé pour l'enzyme EcoRV la valeur de  $10^{-6}$  pour  $\mathcal{P}_{\text{off}}$  [15]. Erskine *et al* ont déterminé la constante de dissociation pour EcoRV en présence de 10 mM de  $\text{MgCl}_2$  :  $k_{\text{off}} = 1,2 \text{ s}^{-1}$  [23], nous pouvons alors estimer le coefficient de diffusion  $D_1 = \frac{a^2 k_{\text{off}}}{P_{\text{off}}} \sim 0,1 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ .

Il faut appréhender  $\ell_{\text{sliding}}$  comme un paramètre étant donné qu'il dépend énormément des conditions expérimentales qui influent sur l'affinité non-spécifique des protéines à l'ADN telle que la salinité du milieu [11]. Néanmoins, nous pouvons garder à l'esprit que  $N$  pas de sliding conduisent à l'exploration de  $\sqrt{N}$  sites différents pour une durée caractéristique de  $\frac{N}{D_1}$  : pour des petites valeurs de  $N$ , le sliding est un processus efficace, cependant au fur et à mesure que  $N$  augmente l'exploration des sites devient redondante.

### 1.2.3.2 Le hopping - jumping

**Description :** les molécules d'ADN sont repliées sous forme de pelote en solution ; après s'être dissociée de l'ADN, une protéine a donc une probabilité non négligeable de s'y réassocier rapidement. Les mécanismes de hopping et de jumping décrivent de telles situations qui correspondent à des dissociations-réassociations successives de la protéine sur l'ADN, autrement dits des sauts. Selon la taille du saut, on utilise le terme hopping (arrivée proche du site de départ : déplacement  $< 20$  pb) ou jumping (arrivée complètement décorrélée du site de départ, déplacement  $> 20$  pb). Cette limite de 20 pb est arbitrairement fixée par Halford [17], dans la suite, nous ne garderons que le terme jumping.

**Modélisation - efficacité :** dans le modèle de jumping,  $N$  sauts conduisent à l'exploration de  $N$  sites sur l'ADN. Le temps moyen pour faire un saut peut-être estimé avec le majorant suivant :  $t_{\text{saut}} = \frac{R_G^2}{D_3}$  où  $D_3$  est le coefficient tridimensionnel de la protéine et  $R_G$  le rayon de giration de l'ADN en pelote donné par  $R_G = aM^{\frac{3}{4}}$  où  $a$  est la taille d'une paire de bases,  $M$ , le nombre de monomères<sup>3</sup>. Pour un ADN de taille 10 kb utilisé lors de nos expériences, il vient :  $R_G \approx 350$  nm. En prenant comme valeur caractéristique  $D_3 = 50 \mu\text{m}^2 \cdot \text{s}^{-1}$ , il vient  $t_{\text{saut}} = 2,4$  ms, bien plus grand que le temps moyen pour faire un déplacement élémentaire dans le modèle du sliding  $t_{\text{sliding}} = \frac{a^2}{2D_1} \approx 6 \mu\text{s}$  pour un coefficient  $D_1$  de  $0,01 \mu\text{m}^2 \cdot \text{s}^{-1}$  (valeur inspirée de celle obtenue lors de nos expériences, cf. Chapitre III).

---

<sup>3</sup>Cf. Chapitre suivant, section II.3.1

Même si  $N$  sauts conduisent à l'exploration de  $N$  sites, le temps moyen requis pour faire un saut est trois ordres de grandeur supérieurs au temps nécessaire pour faire un pas en sliding : le jumping n'est donc pas forcément plus efficace que le sliding.

### I.2.3.3 L'intersegment transfert

Dans le cas de protéines dimériques ayant deux sites de fixation à l'ADN, l'intersegment transfert est un mécanisme qui décrit le passage de la protéine d'un endroit de l'ADN à un autre, relativement éloignés<sup>4</sup> mais temporairement proches : lorsque deux segments d'ADN sont suffisamment proches, la protéine déjà fixée au premier segment par l'un de ses sites bascule sur l'autre segment en s'y fixant par son second site tout en se détachant du premier [24]. Nous ne considérerons plus ce mécanisme par la suite puisque notre protéine d'étude (l'enzyme de restriction EcoRV) ne possède qu'un seul site de fixation à l'ADN.

### I.2.3.4 Modèle mixte : sliding - jumping

Récemment, les expériences de biochimistes ont été analysées en terme d'une combinaison de mécanismes 1D et 3D, combinaison qui présente l'intérêt de réduire le temps nécessaire à la localisation de séquences-cibles [16, 25, 26] : la succession d'associations et de dissociations (jumping) permet d'explorer successivement par sliding des petites zones de l'ADN probablement différentes, rendant la recherche moins redondante que dans un modèle de sliding pur.

Un modèle simple permet d'évaluer le temps nécessaire à l'exploration d'une molécule d'ADN de longueur  $L$ , supposée grande devant la longueur de persistance de l'ADN (qui vaut typiquement 50 nm en conditions classiques). Une protéine alterne entre des phases de sliding et des phases de diffusion 3D (jumping) à l'intérieur de la pelote d'ADN. On néglige toute dissociation amenant la protéine hors du domaine de l'ADN. Les phases de sliding sont caractérisées par un coefficient de diffusion  $D_1$  et une longueur moyenne  $\ell_1$  de la zone explorée avant dissociation. Les phases de jumping sont caractérisées par un coefficient de diffusion  $D_3$  et une distance moyenne  $\ell_3$  entre les sites de dissociation et de réassociation. Cette longueur  $\ell_3$  dépend bien évidemment de la longueur de la molécule d'ADN et de sa conformation.

La durée moyenne d'une phase de sliding est  $\ell_1^2/D_1$ , et la durée moyenne d'une phase de jumping est  $\ell_3^2/D_3$ . Faisons l'hypothèse simplificatrice que les zones explorées par sliding sont disjointes : l'ensemble de la molécule est donc explorée à l'issue de  $L/\ell_1$  événements d'association puisque nous avons vus que  $\tau_3 \ll \tau_1$ . Le temps nécessaire à la localisation

---

<sup>4</sup>À cause de la longueur de persistance de l'ADN

de la séquence-cible est une fonction de  $\ell_1$  et s’exprime :

$$\tau(\ell_1) = \frac{L}{\ell_1} \left( \frac{\ell_1^2}{D_1} + \frac{\ell_3^2}{D_3} \right).$$

Ce temps devient très élevé pour  $\ell_1 \rightarrow 0$  (pur jumping), car le nombre d’événements d’association/dissociation, de durée non nulle, est alors très grand. Il est également grand pour  $\ell_1 \rightarrow L$  (pur sliding) en raison de la durée proportionnelle à  $\ell_1^2$  des phases de sliding. Il passe par un unique minimum  $\ell_1$  lorsque les phases de diffusion 1D sur l’ADN et 3D dans la pelote ont même durée :

$$\ell_1 = \ell_3 \sqrt{\frac{D_1}{D_3}}. \tag{I.2}$$

Le rapport  $D_1/D_3$  peut être estimé à l’aide de considérations hydrodynamiques [4] : il a une valeur de 1/10 à 1/1000 pour une protéine de taille nanométrique.

Pour une molécule d’ADN du bactériophage T7 utilisé dans nos expériences, de longueur 9,2 kb soit 3,2  $\mu\text{m}$ ,  $\ell_3$  est de l’ordre de  $R_G$  soit 0,35  $\mu\text{m}$ . Finalement,  $\ell_1 \simeq 10$  à 100 nm  $\simeq 30$  à 300 paires de bases. Pour des ADN de bactériophages, de taille 40 kb, ce modèle simple permet de prédire une longueur de sliding optimale de l’ordre de quelques dizaines à quelques centaines de paires de bases sur les ADN de bactériophages. Comme nous le verrons dans la partie suivante, ces valeurs sont cohérentes avec celles qui ont été prévues par Jeltsch et Pingoud pour l’enzyme EcoRV [15].

### **1.2.3.5 Réalité *in vivo***

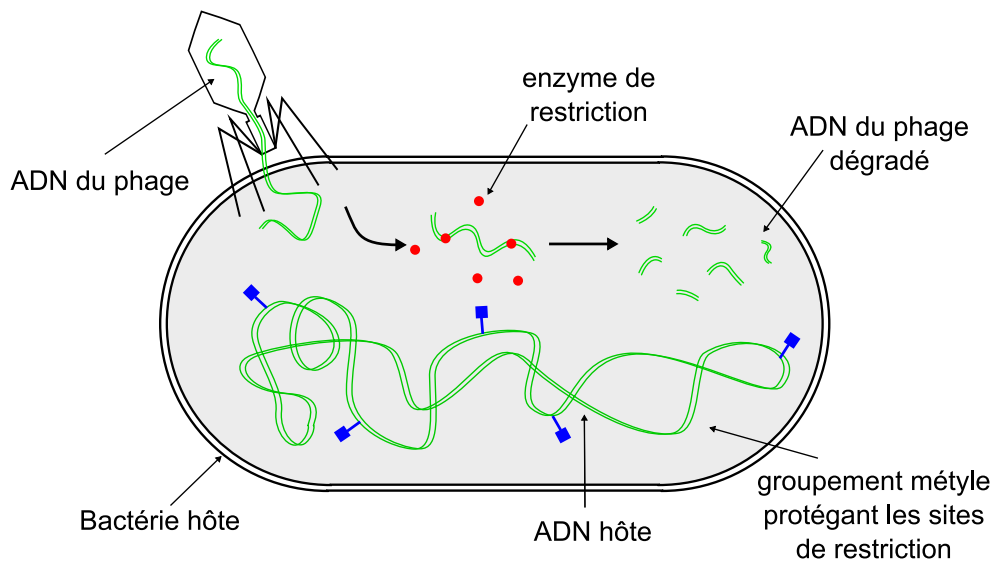
*In vivo*, l’ADN est recouvert de nombreuses protéines qui constituent des obstacles à la diffusion 1D [27] : le sliding ne peut donc avoir lieu *in vivo* que sur des distances inférieures à celles séparant ces obstacles. Néanmoins, il est raisonnable d’imaginer que l’ADN des phages est exempt de protéines juste après avoir été injecté dans la bactérie : dans cette situation, les enzymes de restriction pourraient user de sliding *in vivo* comme elles le font *in vitro* [28, 29].

## **1.3 Notre protéine d’investigation : l’enzyme de restriction EcoRV**

### **1.3.1 Les enzymes de restriction**

#### **1.3.1.1 Le mécanisme de défense des bactéries**

Les premières enzymes de restriction ont été découvertes en 1962 par les biologistes Aber, Nathans et Smith (prix Nobel de médecine 1978) qui avaient remarqué que certaines



**Figure I.3: Le système de défense des bactéries.** Les enzymes de restriction de type II (points rouges) reconnaissent et coupent de courtes séquences de nucléotides des ADNs des phages. Les méthylases protègent l'ADN bactérien en ajoutant des groupes méthyles (carrés bleus) sur l'ADN bactérien.

bactéries avaient la propriété de dégrader de l'ADN étranger introduit dans leur cytosol. La recherche de l'explication de ce phénomène les a amenés à la découverte d'une nouvelle classe d'enzymes présentes chez les bactéries hôtes : les enzymes de restriction. Lorsque les bactéries sont parasitées par des virus à ADN (nommés bactériophages), les enzymes de restriction coupent les ADNs étrangers pour les empêcher de s'insérer dans leur propre ADN et de se dupliquer. La destruction du génome du virus bloque ainsi l'infection. Étant donné que les séquences reconnues par les enzymes sont courtes (de l'ordre de quelques paires de bases), il est statistiquement probable qu'elles existent sur un ADN viral long de plusieurs milliers de paires de base. De plus, les bactéries produisent beaucoup d'enzymes de restriction (chez *E.Coli*, on connaît plus d'une centaine d'enzyme de restriction différentes), ce qui permet *a priori* de couper quasiment n'importe quel ADN viral.

Pour se protéger de leurs propres enzymes de restrictions, et éviter ainsi qu'elles ne coupent l'ADN de leur propre génome, les bactéries fabriquent d'autres enzymes, appelées méthylases. Ces méthylases protègent l'ADN bactérien en ajoutant un groupement méthyl sur les sites d'action des enzymes de restriction. L'association de ces deux enzymes, l'une de coupure (enzyme de restriction) et l'autre de protection (méthylases) constitue le mécanisme de défense des bactéries contre les virus et s'appelle système de restriction/modification.



### I.3.1.2 Propriétés

Les enzymes de restriction sont capables de cliver les deux brins d'ADN : elles appartiennent à la classe des endonucléases, c'est-à-dire des enzymes qui coupent à l'intérieur d'un acide nucléique, les liaisons phosphodiester entre deux nucléotides.

### I.3.1.3 Les différentes familles

Les systèmes de restriction et de modification ont été classés en trois principaux groupes [30].

Les enzymes des systèmes de type I et III possèdent à la fois l'activité endonucléase et méthylase. Ces enzymes reconnaissent une séquence d'ADN spécifique mais coupent en un endroit plus ou moins éloigné du site reconnu. Les enzymes de type I coupent 1000 à 5000 paires de bases plus loin du site de reconnaissance. Les enzymes de type III consomment de l'ATP et coupent une vingtaine de nucléotides après la séquence de reconnaissance. La suite de ce manuscrit portera entièrement sur le troisième type de système de restriction-modification : le type II, le plus abondant et le plus étudié.

Les enzymes de restriction de type II ne possèdent que l'activité endonucléase (une autre enzyme possède l'activité méthylase). Elles ne consomment pas d'ATP et leur site de coupure est le même que leur site de reconnaissance, le magnésium joue le rôle de cofacteur dans la réaction de clivage. Toutes les endonucléases de type II sont des protéines homodimériques et sont structurellement très proches les unes des autres.

Leur capacité à couper spécifiquement, comparable à des « ciseaux » à ADN fait des enzymes de restriction de type II un outil simple et fondamental pour les biologistes moléculaires puisqu'elles permettent de fragmenter l'ADN en segments de taille définie : leur découverte a ainsi constitué une avancée décisive dans la manipulation de l'ADN et ses gènes. On dispose aujourd'hui de plusieurs centaines de ces enzymes pouvant couper l'ADN selon différentes séquences. Pour une revue complète sur les enzymes de type II, le lecteur peut se reporter à la revue de Jeltsch et Pingoud [31].

### I.3.1.4 Pourquoi avoir choisi les enzymes de restriction de type II ?

Les enzymes de restriction sont l'objet de nombreuses études sur les processus de localisation des séquences cibles sur l'ADN. Plusieurs raisons motivent ces études et nous ont d'ailleurs conduit à étudier l'une d'entre elles : EcoRV.

- Les enzymes de restriction sont par nature un exemple de protéines qui doivent localiser une courte séquence cible sur de très longues molécules d'ADN. La plupart des enzymes de restriction de type II reconnaissent une séquence de 4 à 6 paires de

bases, les ADNs des bactériophages sont des molécules de plusieurs dizaines de milliers de paires de bases : l'ADN du phage T7 contient 39 kb.

- ▶ La propriété que les enzymes ont de couper l'ADN en petits fragments (on parle de digestion enzymatique) a été grandement exploitée par les biochimistes depuis le début des années 80 qui ont pu en tirer parti en utilisant les techniques de gels d'électrophorèse. Ces techniques permettent notamment d'accéder à la taille des fragments d'ADN présents dans un échantillon donnant ainsi accès aux caractéristiques cinétiques. Les principales études portaient sur les enzymes EcoRI et EcoRV et ont abouti à des conclusions parfois divergentes : certains groupes ont ainsi montré la possibilité d'un mouvement de sliding *in vitro*, d'autres ont mis en évidence l'implication de processus 3D également *in vitro*. Il est à présent établi que la localisation des séquences cibles utilise conjointement ces deux mécanismes, leur importance relative reste néanmoins inconnue. Nous ferons le point sur les différentes expériences qui ont été menées au cours de ces vingt dernières années sur ce sujet ainsi que sur leurs limitations à la section suivante (section I.3.3).
- ▶ Les gènes codant pour ces enzymes ayant été pour la plupart isolés, plusieurs groupes de biochimistes sont capables de les produire en quantité importante, ainsi que d'en synthétiser des mutants. Le nombre élevé de publications concernant la diffusion facilitée d'EcoRV nous a d'ailleurs conduit à contacter A. Pingoud (Justus-Liebig Universität, Giessen, Allemagne) pour lui proposer une collaboration : Wolfgang Wende a accepté de nous donner des échantillons de l'enzyme EcoRV modifiée pour mener nos expériences.

### I.3.2 EcoRV

#### I.3.2.1 Nomenclature

EcoRV est produite par la bactérie *Escherichia coli*. Le nom des enzymes de restriction est dérivé du nom de l'espèce ou de variété de la bactérie qui la produit. Il faut écrire en majuscules l'initiale du genre de la bactérie d'où l'enzyme a été extraite (E), suivi des deux initiales du nom de l'espèce (co), suivi d'une lettre ou d'un nombre pour désigner la variété ou souche bactérienne (R) et enfin un chiffre romain (V) pour désigner le numéro d'ordre.

#### I.3.2.2 Séquence cible et coupure

EcoRV est une enzyme de restriction de type II qui a largement été étudiée par les biochimistes. Cette enzyme reconnaît la séquence spécifique des 6 paires de base suivante : GATATC et la coupe en son milieu, il s'agit d'une coupure à bout franc comme l'illustre

la Figure I.4. EcoRV se présente sous la forme d'un homodimère de masse moléculaire 57

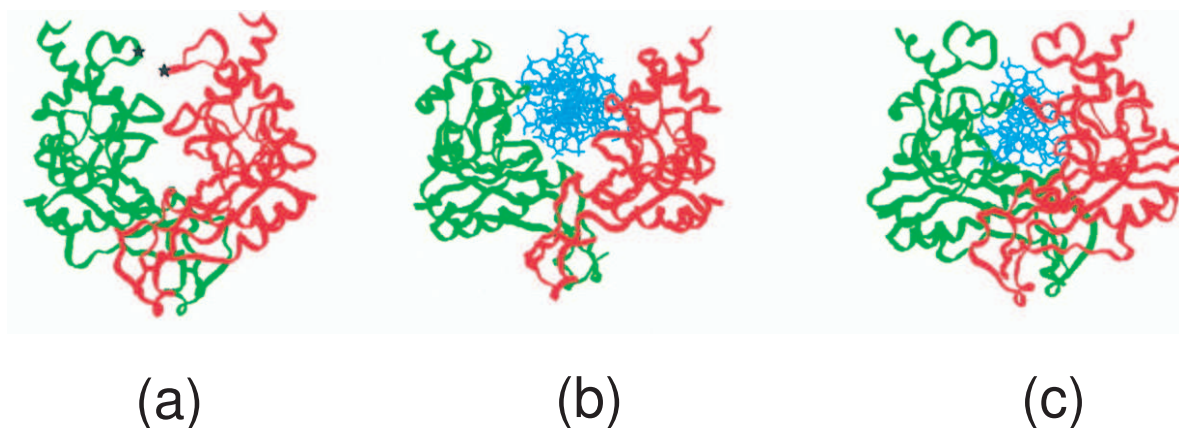


**Figure I.4:** Site de restriction de l'enzyme EcoRV. L'enzyme reconnaît la séquence palindromique GATATC et coupe en son milieu (trait vert).

kDa. Chaque sous-unité de la protéine coupe un brin de la molécule d'ADN. La capacité d'EcoRV à reconnaître spécifiquement sa séquence-cible et à la cliver dépend de la présence d'ions divalents. En l'absence d'ions magnésium, EcoRV manifeste peu de spécificité pour sa séquence cible, qu'elle est incapable de cliver [32]. En leur présence, EcoRV se lie plus fortement à sa séquence-cible qu'à l'ADN non-spécifique. Cependant, sa capacité à cliver l'ADN est variable selon les ions utilisés : le clivage est plus efficace avec des ions  $\text{Mg}^{2+}$ , alors qu'il n'est pas observé en présence d'ions  $\text{Ca}^{2+}$  [33, 34].

### I.3.2.3 Structure cristallographique

EcoRV a été cristallisée en 1993 par Winkler *et al.* dans trois configurations différentes : enzyme libre, enzyme liée non-spécifiquement à l'ADN et enzyme liée spécifiquement à l'ADN (cf. Figure I.5, [35]). Les structures cristallographiques du complexe enzyme-ADN spécifique montrent que EcoRV établit des contacts sur 3 ou 4 bases de part et d'autre de la séquence cible. La comparaison des complexes spécifiques et non spécifiques montre que l'ADN est courbé au niveau de la séquence cible.



**Figure I.5: Structure d'EcoRV :** (a) enzyme libre (b) enzyme liée à l'ADN non spécifique (c) enzyme liée à l'ADN spécifique. Les dimères de la protéine apparaissent en rouge et vert, l'ADN en bleu. D'après [35].

### 1.3.3 Comment EcoRV trouve-t-elle ses sites ?

Élucider le mécanisme avec lequel les enzymes de restriction trouvent leur cibles est une question d'actualité. Comme nous l'avons aperçu à la section I.2.2, Jeltsch *et al.* ont montré en 1995 que la présence d'ADN non-spécifique accélérerait le processus de recherche de cible d'EcoRV [29]. Il est ainsi admis qu'EcoRV use de mécanismes de « diffusion facilitée » pour trouver sa cible : l'enzyme trouve l'ADN par diffusion tridimensionnelle (processus lent) et recherche ensuite sa cible (processus rapide). En ce qui concerne le mécanisme sous-jacent de cette étape rapide, la question est ouverte [16]. Pendant longtemps, il a été proposé qu'un grand nombre d'enzymes de type II se déplaçaient sur l'ADN par sliding [29, 13, 36]. Cependant des études plus récentes tendent à montrer que des alternances de cycle d'association/dissociation seraient des mécanismes probables [37, 17]. Dans la partie qui suit, nous ferons le point sur les différentes études qui ont été menées sur EcoRV jusqu'à maintenant dans le but de saisir le mécanisme de diffusion facilitée. EcoRV n'est pas la seule enzyme de restriction qui a été étudiée dans ce sens : EcoRI [28, 13] ou encore BamHI [38, 13] sont également de bons sujets d'études. Par souci de simplification, nous ne présenterons que les principaux résultats ayant trait à EcoRV.

#### 1.3.3.1 Preuves expérimentales du sliding

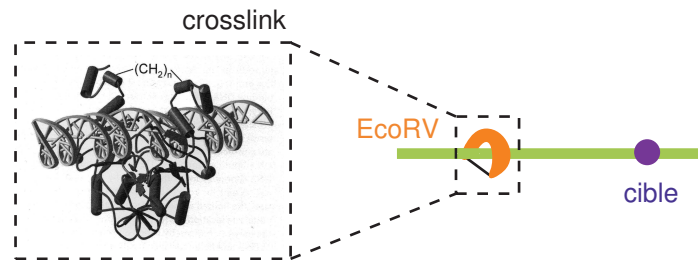
**Action sur des molécules d'ADN contenant deux séquences-cibles proches d'une extrémité.** En 1998, l'équipe d'A. Pingoud s'est intéressée à des molécules d'ADN contenant deux séquences cibles pour EcoRV proches d'une extrémité de l'ADN (cf. Figure I.6). Ils ont ainsi comparé les vitesses de coupure pour ces deux sites : ils ont observé une coupure préférentielle de la séquence cible intérieure [15]. Ce résultat peut être vu comme une preuve d'existence de sliding. En effet, les processus de jumping n'ont aucune raison de favoriser une des cibles par rapport à l'autre, au contraire, le sliding induit une nette préférence pour la cible intérieure puisque la protéine a plus de chance de s'associer initialement à l'ADN du côté intérieur.



**Figure I.6: Preuve expérimentale du sliding 1.** Deux séquences cibles proches sont insérées à l'extrémité d'une molécule d'ADN. La séquence cible intérieure est préférentiellement coupée.

**Enzyme liée autour de l'ADN de manière covalente.** En 1998 également, la même équipe a réussi à produire un mutant d'EcoRV dont les deux dimères sont reliés de ma-

nière covalente autour d'une molécule d'ADN contenant une séquence cible pour EcoRV (cf. Figure I.7, [39]). Cette construction est relativement contraignante puisqu'elle élimine toute possibilité de déplacement tridimensionnel : les enzymes parviennent néanmoins à trouver leur cible et à couper l'ADN. Ce résultat prouve que ces enzymes EcoRV peuvent localiser leur cible uniquement par sliding, sans toutefois démontrer que la diffusion unidimensionnelle est l'unique mécanisme dans la localisation des séquences-cibles par les enzymes sauvages.

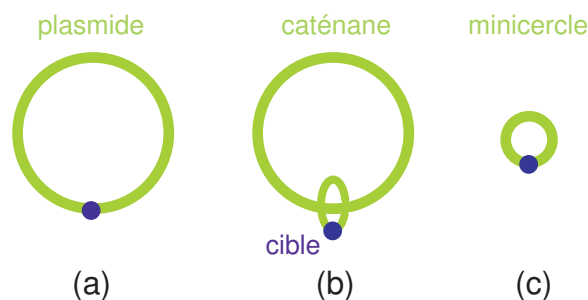


**Figure I.7: Preuve expérimentale du sliding 2.** Représentation schématique de l'expérience menée par Schulze *et al.* Les bras de l'enzyme EcoRV sont reliés de manière covalente autour de l'ADN.

### I.3.3.2 Preuves expérimentales du jumping

**Différentes topologies d'ADN** En 2003, Gowers et Halford ont comparé l'action d'EcoRV sur 3 substrats d'ADN différents contenant chacun une séquence-cible pour EcoRV comme l'illustre la figure I.8 [40] :

- un plasmide de taille 3400 pb contenant la séquence cible (construction n° (a)),
- une caténane constituée d'un minicercle d'ADN de taille 300 pb et comprenant la séquence cible, ce minicercle est accroché autour d'un plasmide plus long (3100 pb) (construction n° (b)),
- les mêmes minicercle et plasmide non accrochés entre-eux (construction n° (c)).



**Figure I.8: Preuve expérimentale du jumping 1.** Différents substrats d'ADN contenant une séquence-cible pour EcoRV. EcoRV trouve préférentiellement sa cible sur les constructions plasmide (a) et caténane (b).

Expérimentalement, il a été observé que la coupure des enzymes est plus efficace sur le plasmide (a) et sur le caténane (b) plutôt que sur le mini-cercle (c). Dans l'hypothèse d'un mécanisme de sliding, l'action sur le plasmide (a) devrait être plus rapide que dans le cas du caténane (b), puisque le minicercle et le plasmide sont disjoints dans la construction (b) : l'hypothèse n'est donc pas la bonne. En revanche, en supposant un mécanisme de jumping, une enzyme initialement située sur le plasmide a autant de chance de trouver sa cible pour les constructions (a) et (b). De plus, étant donné que le volume occupé par le minicercle (c) est plus petit que dans les autres cas, l'enzyme a plus de chance de s'en éloigner après dissociation, retardant ainsi la localisation. Les auteurs excluent ainsi un modèle de sliding pur, tranchant ainsi pour un modèle de diffusion facilitée à 3D.

**ADN surenroulé** Dans le même article, les auteurs ont également montré que les enzymes localisaient plus rapidement leurs cibles sur des molécules d'ADN surenroulées plutôt que sur des molécules d'ADN relâchées. *A priori*, un mécanisme de diffusion facilitée à 1D ne doit pas dépendre de la conformation de l'ADN. En revanche, dans le cas d'un mouvement de jumping, le surenroulement augmente fortement la « densité d'ADN » favorisant la probabilité de se réattacher à l'ADN après dissociation, c'est-à-dire favorisant les sauts.

### 1.3.3.3 Sliding - jumping : quelles contributions ?

Aujourd'hui, les différents groupes s'accordent à dire qu'il y a une implication conjointe des mécanismes 1D et 3D dans la localisation de séquences-cibles. La question intéressante qui se pose ensuite est la suivante : quelle est l'importance relative de ces deux mécanismes ? Pour y répondre, ces mêmes groupes ont réalisé des expériences de processivité qui théoriquement peuvent séparer les contributions des mécanismes 1D/3D.

**Les expériences de processivité** Dans les expériences de processivité, les substrats d'ADN possèdent deux séquences-cibles pour l'enzyme d'intérêt. La processivité d'une enzyme de restriction définit sa capacité à couper successivement les deux séquences cibles en fonction de la distance qui les sépare. D'un point de vue pratique, elles sont réalisées en large excès d'ADN afin de ne pouvoir imputer la coupure qu'à l'action processive d'une même enzyme.

Cette capacité à recouper après avoir coupé une première fois dépend naturellement du mouvement de la protéine le long de l'ADN : mesurer la processivité d'une enzyme en fonction de la distance inter-site doit permettre de remonter au mécanisme qui régit le mouvement de la protéine. Ainsi, Jack *et al.* ont montré en 1982 que l'enzyme de restriction EcoRI ne se dissocie pas de l'ADN après avoir trouvé une cible, mais continue à se mouvoir sur l'ADN par sliding [12]. Néanmoins, les expériences de processivité ont une

limitation fondamentale étant donné que leur interprétation nécessite de les confronter aux prédictions théoriques issues des modèles de mécanismes microscopiques. Toutes les étapes de ces mécanismes (passage non-spécifique au spécifique, coupure, dissociation) sont spéculatives, actuellement, aucune preuve expérimentale ne permet de trancher.

Le groupe de S. Halford a analysé en 2000 la processivité d'EcoRV sur des molécules d'ADN contenant deux séquences cibles séparées par des distances de 54 à 764 pb [17]. En supposant que le mécanisme microscopique de déplacement de l'enzyme est du sliding, on note  $\mathcal{P}_{\text{diff}}$  le rapport  $\mathcal{P}_{\text{diff}} = \frac{k_S}{k_{\text{off}}}$  où  $k_S$  et  $k_{\text{off}}$  sont respectivement les constantes cinétiques de translocation et de dissociation de l'enzyme sur/à l'ADN. La probabilité  $\mathcal{P}$  de faire un pas sur l'ADN plutôt que de s'en dissocier est alors donnée par :

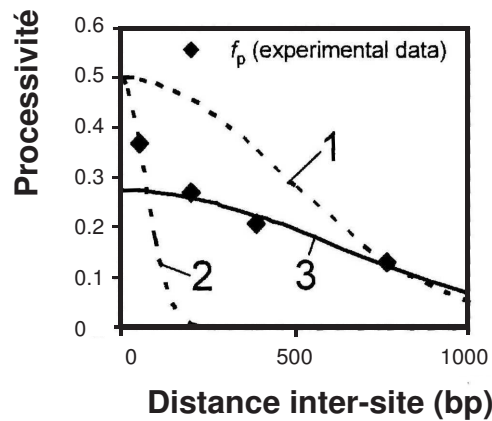
$$\mathcal{P} = \frac{\mathcal{P}_{\text{diff}}}{\mathcal{P}_{\text{diff}} + 1}. \quad (\text{I.3})$$

Si  $N$  paires de bases séparent les deux séquences-cibles, Stanford *et al.* définissent la processivité [17] :

$$f_P = f_{\text{max}} \times \left[ \frac{\mathcal{P}_{\text{diff}}}{\mathcal{P}_{\text{diff}} + 1} \right]^{N^2}, \quad (\text{I.4})$$

où  $f_{\text{max}}$  est la processivité maximale. À partir de là, les conclusions divergent entre Halford et Pingoud :

- Selon le groupe d'Halford, l'enzyme animée de sliding, se retrouve obligatoirement sur l'un des deux fragments d'ADN après la première coupure, elle a donc une probabilité 1/2 de se trouver sur le fragment contenant la seconde cible :  $f_{\text{max}} = 1/2$ . En prenant cette valeur de  $f_{\text{max}}$ , leurs ajustements (courbes 1 et 2 de la Figure I.9 ne coïncident pas avec leurs points expérimentaux (losanges). Ils en déduisent que la localisation des séquences cibles se fait essentiellement par jumping. Leurs analyses prédisent une longueur parcourue par sliding inférieure à 25 paires de bases avant la dissociation.
- Pour le groupe de Pingoud, EcoRV est une enzyme qui coupe à bout franc : par conséquent, rien ne garantit que l'enzyme reste sur l'ADN immédiatement après la coupure. Les deux fragments restant peuvent alors diffuser librement indépendamment de l'enzyme (ce qui correspond à une phase de jumping) : de fait, la processivité maximale doit être inférieure à 1/2. En laissant le paramètre  $f_{\text{max}}$  variable dans leurs ajustements, cette équipe trouve une interpolation correcte des points expérimentaux de Halford pour une processivité maximale de 0,27. Selon eux, le sliding reste le mécanisme principal et ils annoncent une exploration de 1000 pb sur l'ADN avant la dissociation [18].



**Figure I.9:** Les différentes analyses des expériences de processivité. Variation de la processivité d'EcoRV en fonction de la distance inter-site. Les losanges sont les points expérimentaux pris par l'équipe de S. Halford [17], les courbes 1 et 2 correspondent à leurs ajustements non pertinents basés sur un modèle de sliding dans lequel la processivité tend vers 1/2 lorsque la distance inter-site tend vers zéro (« processivité maximale »). La courbe 3 est l'ajustement du groupe de A. Pingoud, selon un modèle dans lequel la processivité maximale est laissée comme paramètre ajustable (déterminée à 0,27). Extrait de [18].

Cette difficulté à interpréter les résultats ne permet donc pas de déterminer la réelle contribution des mécanismes de sliding et de jumping. Dans le chapitre suivant, nous allons décrire l'approche en molécule unique que nous avons développée pour visualiser directement les interactions non spécifiques ADN-EcoRV. Le dernier chapitre de ce mémoire illustrera les éléments de réponse qui peuvent être obtenus étoffant ainsi les conclusions faites par les biochimistes.





# Chapitre II

## Mise en place du dispositif expérimental

*Ce deuxième chapitre est consacré à la présentation du dispositif expérimental que nous avons utilisé. Nos expériences visent à l'observation des interactions ADN-protéines par microscopie de fluorescence à l'échelle de la molécule individuelle. Pour y parvenir, deux étapes sont nécessaires :*

- En solution, les molécules d'ADN sont sous forme de pelote : dans une telle configuration, elles ne peuvent être résolues optiquement. Nous avons mis en place une technique permettant de dérouler une pelote d'ADN sous l'effet d'un flux : l'étirement d'ADN. Ainsi étirée, la molécule d'ADN est idéalement positionnée pour interagir avec les protéines.*
- Pour être repérées, les enzymes doivent être couplées à des fluorophores. Cette étape de marquage est délicate : les fluorophores ne doivent pas gêner l'action de l'enzyme et doivent pouvoir être détectés avec des résolutions spatiale et temporelle suffisantes.*

### II.1 Étudier les interactions ADN-protéine à l'échelle de la Molécule Unique

#### II.1.1 Motivations : s'affranchir du moyennage

##### II.1.1.1 Accéder aux inhomogénéités

Étudier des objets de taille nanométrique consistait, il y a encore une dizaine d'années, à travailler sur de grands ensembles de populations. Depuis, de nouvelles techniques dédiées à la micromanipulation d'objet sont apparues rendant les études à l'échelle d'une molécule individuelle possibles et offrant de nombreux intérêts. Alors que les mesures d'ensemble donnent accès à la moyenne d'une grandeur donnée sur un grand nombre de

particules, les observations à l'échelle de la molécule individuelle permettent de s'affranchir de tout effet de moyenne et d'obtenir la distribution associée à la grandeur en question, c'est-à-dire d'accéder à sa valeur moyenne et à ses fluctuations. L'analyse des dispersions des paramètres moléculaires permet alors de remonter aux inhomogénéités de population.

### II.1.1.2 Accéder aux processus stochastiques

La dynamique des objets que nous étudions est contrôlée par des processus stochastiques. Il n'y a donc aucune synchronisation à l'échelle d'un ensemble de molécules : ces processus ne peuvent donc pas être observés lors des mesures d'ensemble. En revanche, étudier l'état d'une molécule au cours du temps doit permettre d'accéder aux caractéristiques de ces processus stochastiques. Il est ainsi possible de visualiser les fluctuations temporelles des paramètres moléculaires, ceci, même si on est à l'équilibre thermodynamique.

### II.1.1.3 Les cas des mesures de diffusion

Nombreuses sont les techniques physiques ou chimiques qui permettent de suivre les paramètres d'une cinétique, pour en citer quelques-unes : spectrophotométrie, fluorescence, radioactivité, dosages immunologiques. Les biochimistes ont largement étudié les cinétiques d'interaction enzyme-protéine dans le cadre de mesures d'ensemble. Généralement, les expériences mesurent la variation de concentration du substrat ou du produit en fonction du temps et cela pour une concentration donnée d'enzyme et différentes valeurs de la concentration en substrat. La connaissance des valeurs de ces différents paramètres permet de déterminer les constantes cinétiques et de remonter à une valeur du coefficient de diffusion. Ces mesures d'ensemble ne permettent pas de remonter à la dynamique du processus de diffusion. Le suivi en temps réel de l'interaction d'une enzyme unique avec une molécule d'ADN unique permettra de mieux comprendre le processus de diffusion facilitée des enzymes de restriction.

## II.1.2 Dispositif : le montage de TIRFM

Pour visualiser l'interaction de l'enzyme de restriction EcoRV avec l'ADN non spécifique, nous avons choisi d'utiliser un montage de microscopie de fluorescence par excitation en onde évanescente (TIRFM). Cette technique est en pratique la mieux adaptée à nos objectifs comme nous le verrons à la section II.1.3 dans laquelle nous décrirons sommairement les autres approches moléculaires qui sont actuellement utilisées dans les études des interactions ADN-protéines.

### II.1.2.1 La microscopie de fluorescence

La microscopie de fluorescence permet d'étudier des processus sur des échelles spatiales et temporelles étendues (du nanomètre au micromètre et de la nanoseconde à la seconde). Elle constitue actuellement un outil de base dans de nombreux domaines de la recherche : de la biologie moléculaire à la physique en passant par la biochimie. Rappelons brièvement les caractéristiques générales du processus d'émission de fluorescence.

**Historique** Sir George Gabriel Stokes<sup>1</sup> fut le premier à observer le phénomène de fluorescence en 1852 : lorsqu'il éclaire le minéral de fluorite avec des radiations ultraviolettes, une vive lueur bleue apparaît. Il donne le nom de fluorescence à ce phénomène lumineux. Stokes montre ensuite que la longueur d'onde de la lumière de fluorescence est plus grande que celle de la lumière d'excitation (propriété connue sous le nom de déplacement de Stokes) et que l'émission de lumière s'arrête dès que l'excitation est coupée. De manière générale, on parle de photoluminescence pour les processus qui se traduisent par l'émission de photons par une molécule qui a été irradiée par un faisceau lumineux. La photoluminescence distingue la phosphorescence de la fluorescence, selon la nature des états fondamentaux et excités. La Figure II.1 représente le diagramme de Jablonski

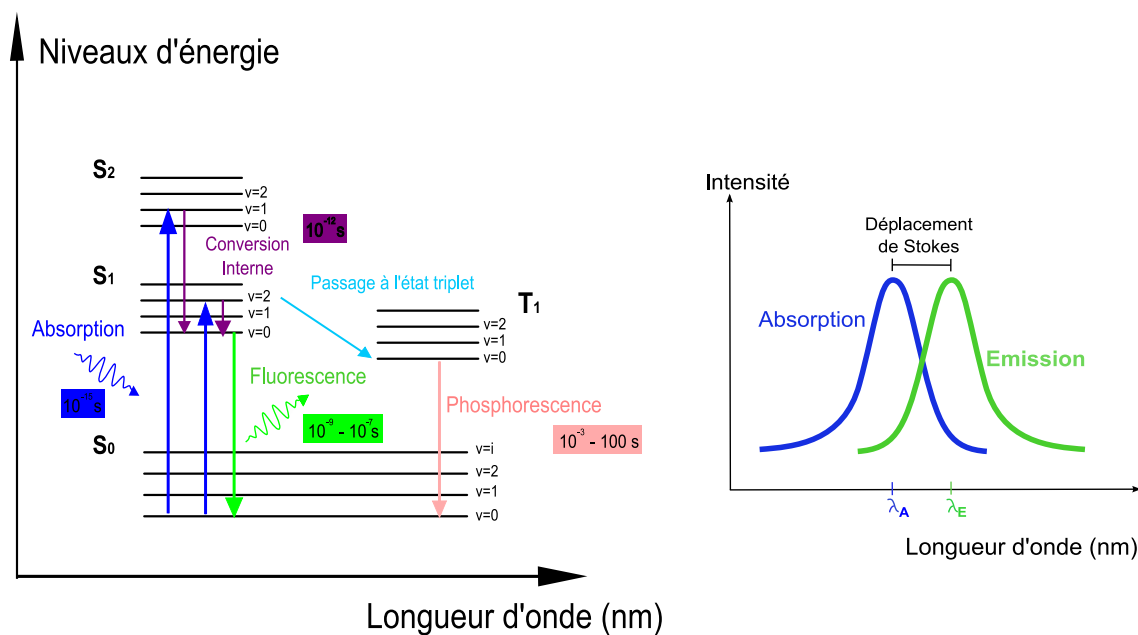
**Éléments de théorie** Le processus d'absorption d'un photon par une molécule initialement dans un état singulet  $S_0$  l'amène sur un niveau singulet  $S_1$ . Les règles de sélection de mécanique quantique imposent que l'état vibrationnel de l'état d'arrivée doit être plus haut que celui de l'état de départ (principe de Franck-Condon, cf. Figure II.1). À température ambiante, la molécule retombe rapidement vers l'état fondamental de  $S_1$  ou  $S_2$  : ce processus est appelé conversion interne. Plusieurs évolutions sont alors possibles :

- la molécule retombe vers l'état fondamental  $S_0$  et émet un photon de fluorescence,
- la molécule passe vers l'état triplet  $T_1$  par une transition non radiative. Le retour de  $T_1$  à  $S_0$  peut s'accompagner de l'émission d'un photon de phosphorescence.
- la molécule revient, de manière non-radiative dans son état fondamental sans émission de photons. Ces transitions non radiatives sont plus ou moins fréquentes en fonction de divers paramètres liés à l'environnement direct de la molécule.

Les conversions internes qui se produisent lors de la désexcitation d'une molécule sont à l'origine des pertes d'énergie entre les photons d'excitation et les photons d'émission. Ainsi, la longueur d'onde centrale d'émission  $\lambda_E$  est toujours supérieure à la longueur d'onde centrale d'absorption  $\lambda_A$ . Cette propriété porte le nom de déplacement de Stokes (cf. Figure II.1).

---

<sup>1</sup>Mathématicien et physicien irlandais (1819-1903), célèbre entre autres par sa théorie ondulatoire de la lumière (1852) et le concept d'éther qui l'accompagne.



**Figure II.1: Principe de l'émission de fluorescence.** **Droite :** Le diagramme de Jablonski donne un aperçu détaillé des différentes possibilités d'absorption et d'émission de photons par une molécule.  $S_0$ ,  $S_1$  et  $S_2$  désignent respectivement l'état fondamental singulet, le premier et le deuxième états excités singulets.  $T_1$  est le premier état excité triplet. Pour chacun de ces états, apparaissent les différents états vibrationnels notés  $v=0, v=1, v=2 \dots$  **Gauche :** Le déplacement de Stokes : profil caractéristique des spectres d'absorption et d'émission de fluorescence d'une molécule.

**Rendement quantique et temps de vie de fluorescence** Le temps de vie de fluorescence  $\tau_f$  correspond à la durée de vie de l'état excité. La désexcitation d'une molécule peut se faire de manière radiative (émission de fluorescence) ou non. Ces deux types de processus sont caractérisés par des probabilités de transition  $k_0 = \frac{1}{\tau_0}$  et  $k_{nr} = \frac{1}{\tau_{nr}}$ . Ainsi, le nombre  $N$  de fluorophores qui se trouvent dans l'état excité varie exponentiellement au cours du temps selon l'équation :  $\frac{dN}{dt} = -(k_0 + k_{nr}) N$ , On définit ainsi le temps de vie de fluorescence  $\tau_f = \frac{1}{k_0 + k_{nr}}$ . Typiquement, la durée de vie de l'état excité d'une molécule fluorescente autrement appelée fluorophore varie de quelques dizaines de picosecondes à quelques nanosecondes.

Le rendement quantique  $\eta$  est défini comme le rapport entre le nombre de photons absorbés par la molécule et le nombre de photons de fluorescence émis. On montre que :  $\eta = \frac{\tau_f}{\tau_0}$ . Les processus qui favorisent les désexcitations non-radiatives (pH, force ionique, ...) diminuent à la fois temps de vie de fluorescence et rendement quantique. Ces processus sont regroupés sous le terme de *quenching* de la fluorescence.

### Un outil pour la détection de biomolécules individuelles

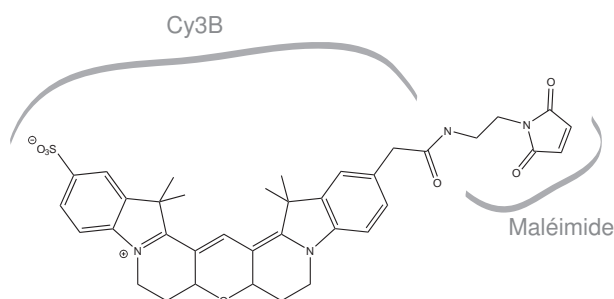
**Les marqueurs fluorescents** : très peu de biomolécules sont naturellement fluorescentes : certains cofacteurs d'enzyme tels que les NADH absorbent dans le bleu, les flavines absorbent dans l'ultra-violet. Ce sont néanmoins de très mauvais fluorophores qui rendent délicate la détection au niveau de la molécule individuelle. La plupart des molécules doivent donc être couplées à des fluorophores extrinsèques pour pouvoir être repérées en microscopie de fluorescence. Actuellement, plusieurs types de sondes sont utilisées dans les expériences de suivi de molécule unique :

- les fluorophores organiques (Cy3, Alexa488, Rhodamine, FITC, ...)
- les nanocristaux semi-conducteurs fluorescents (CdSe/ZnS,...)
- les billes non-fluorescentes (d'or, de latex)
- les protéines fluorescentes (GFP, YFP, ...)
- les micro/nano sphères fluorescentes (renfermant des fluorophores organiques)

En fonction des attentes et des contraintes liées au système biologique, quatre critères doivent être pris en compte pour choisir le fluorophore : la taille, la photostabilité, la capacité à être conjugué à la molécule d'intérêt dans un bon rapport stœchiométrique, et la facilité à être détecté. Au cours de cette thèse, nous avons principalement utilisé des fluorophores organiques et, dans une moindre mesure, des nanocristaux semi-conducteurs.

► **Les fluorophores organiques** sont actuellement les sondes fluorescentes les plus utilisées en biologie. Ce sont des petites molécules organiques qui présentent dans leur structure des groupements chimiques appelés chromophores (tels que C=O, C=C, C=N,

noyau aromatique...) qui possèdent des électrons délocalisés (cf. Figure II.2). En absorbant un photon, les chromophores passent dans un état excité et reviennent à leur état fondamental en émettant de la lumière. Les fluorophores sont synthétisés chimiquement et existent dans le commerce sur une large gamme de longueurs d'onde (de 350 nm à 800 nm). Il existe également une grande variété de dérivés de ces fluorophores possédant des fonctions chimiques réactives adaptées à de nombreux types de couplages biologiques : fluorophores conjugués à des anticorps, à de la streptavidine, à des nucléotides, ... Une grande partie des fluorophores que nous avons utilisés appartient à la famille des cyanines : Cy3B (*Amersham Biosciences*), Sav-Cy3 (*Jackson*), Alexa488 et Sav-Alexa546 (*Invitrogen*). Ces fluorophores sont fréquemment utilisés car ils sont de bons fluorophores (rendement quantique élevé, faible photoblanchiment).

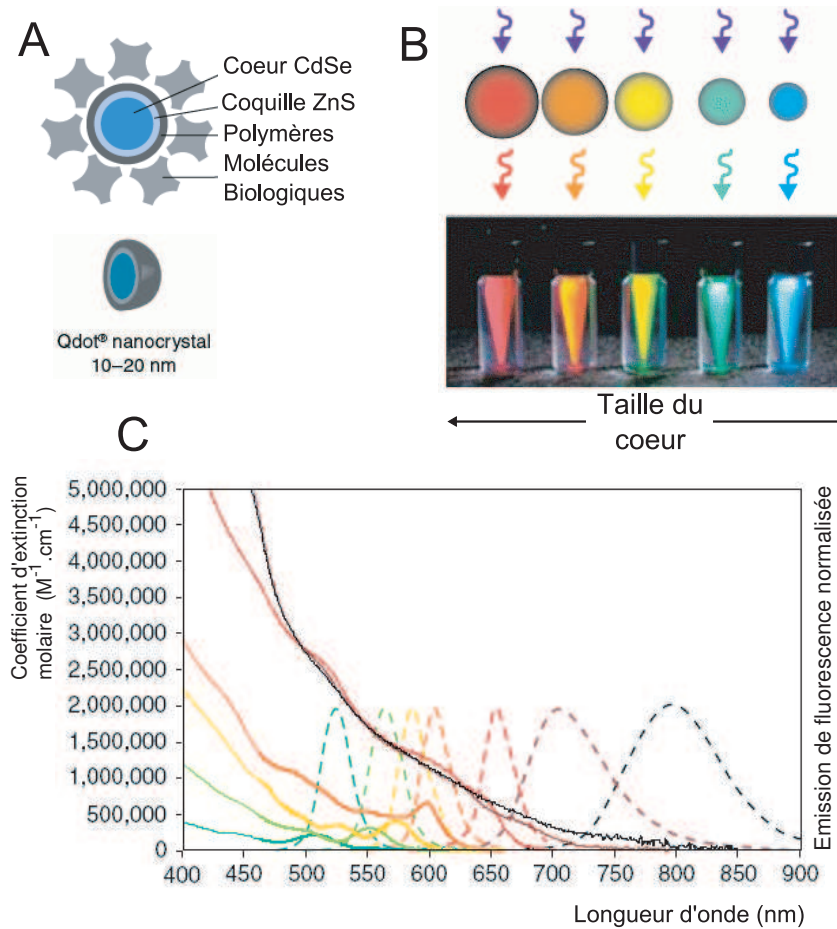


**Figure II.2:** Structure de l'un des fluorophores organiques utilisé durant ce travail de thèse : le Cy3B ester de maléimide.

★ La diversité des couleurs et des dérivés disponibles est l'un des avantages des fluorophores organiques. De plus, leur petite taille (typiquement 1 nm) permet de marquer les molécules sans perturber *a priori* leur comportement.

★ Le phénomène de photoblanchiment est un problème rencontré lorsque l'on utilise des fluorophores organiques. Après irradiation, les fluorophores ont tendance à être chimiquement plus réactifs [41] : des réactions photo-chimiques irréversibles (oxydations, polymérisations,...) peuvent alors créer de nouvelles espèces non fluorescentes. Cela se traduit par une baisse progressive de l'intensité de fluorescence.

► **Les Quantum Dots (QD)** sont des particules sphériques de taille nanométrique composées de semi-conducteurs fluorescents. Ceux que nous utilisons sont constitués d'un cœur de séléniure de cadmium (CdSe) recouvert d'une coquille de sulfure de zinc (ZnS). Depuis une dizaine d'années, ils peuvent être synthétisés par voie chimique avec un diamètre bien contrôlé, entre 5 et quelques dizaines de nanomètres. Les nanocristaux sont disponibles dans le commerce (*Invitrogen*, *Evident*) déjà conjugués à diverses biomolécules : anticorps, streptavidine, ...



**Figure II.3: Les nanocristaux semi-conducteurs.** **A** : schéma d'un nanocristal de CdSe (haut : vue en coupe, bas : vue de profil). **B** : fluorescence de solutions de nanocristaux de différentes tailles excités à la lampe U.V. **C** : Spectres d'absorption (traits pleins) et d'émission (traits pointillés) de différents nanocristaux

Les QD présentent des propriétés optiques remarquables :

- ★ **spectre d'excitation large, spectre d'émission étroit** : les nanocristaux absorbent toute radiation d'énergie supérieure à leur énergie de gap :  $E_{exc} \geq E_{gap}$ , en revanche si  $E_{exc} < E_{gap}$ , la lumière n'est pas absorbée. Lors de l'excitation, une paire électron-trou est créée : cette paire peut se désexciter par émission d'un photon, dont la longueur d'onde est définie par la taille du cœur du semi-conducteur.

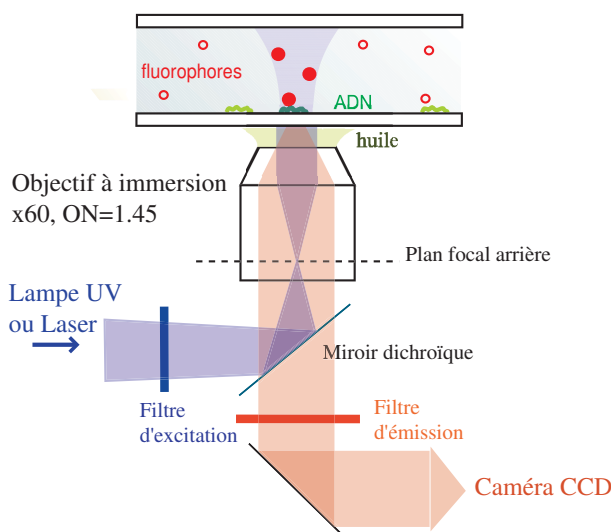
Le spectre d'absorption des nanocristaux est donc large puisqu'ils absorbent toutes les fréquences au dessus de leur gap. Leur spectre d'émission, lié à la désexcitation de la paire électron-trou est symétrique et étroit avec une largeur à mi-hauteur de 25 à 30 nm. Il est donc possible d'exciter simultanément, avec une même source, des nanocristaux de tailles différentes donc de couleurs différentes. On réalise ainsi une détection "multi-couleurs" en évitant le phénomène de *cross-talk*, ce qui n'est pas envisageable avec les fluorophores organiques (cf. Figure II.3)



- ★ **brillance et photostabilité** : la principale limitation des colorants organiques est sans doute le phénomène de photoblanchiment. Dans le cas des nanocristaux, la coquille protectrice en ZnS permet de protéger la paire électron-trou de l'environnement extérieur. Les nanocristaux présentent une grande photostabilité, permettant des observations sur des échelles de temps supérieures à la minute. En réduisant les processus de recombinaison non radiative, cette coquille permet de plus d'obtenir des rendements quantiques élevés ( $\eta > 0,5$ ), similaires, voire légèrement plus faibles, à ceux des meilleurs fluorophores organiques ( $\eta_{\text{Cy3B}} > 0,7$ ). La propriété qu'ont les nanocristaux d'émettre plus de photons de fluorescence que les fluorophores organiques est due à leur section efficace d'absorption, 10 à 100 fois plus élevée que celles de molécules organiques.
  
- ★ **un large choix de couleurs** : la taille du cœur du nanocristal fixe la longueur d'onde d'émission : la lumière émise évolue du bleu au rouge quand la taille du cœur augmente. Avec un même matériau, on obtient ainsi en changeant la taille du cœur des fluorophores des couleurs différentes. En changeant de semi-conducteur, toute la gamme spectrale peut être couverte, depuis le proche I.R. jusqu'au proche U.V.
  
- ★ **un clignotement caractéristique** : la fluorescence d'un nanocristal n'est pas constante au cours du temps, elle présente des alternances de périodes "brillantes" et "noires", phénomène caractéristique des nanocristaux, appelé clignotement, de l'anglais *blinking*. Ce phénomène de clignotement occasionne certes des contraintes sur le suivi de nanocristaux individuels, néanmoins il est la preuve directe de l'unicité du nanocristal observé.

Grâce à leur résistance au photoblanchiment et leur forte luminescence, les nanocristaux sont actuellement les fluorophores les plus performants d'un point de vue photophysique. Dans le but de faire des expériences de suivi à l'échelle de la molécule individuelle, il est très facile de détecter des nanocristaux uniques ; en revanche, ils présentent quelques inconvénients dans le cadre de nos expériences de suivi de l'interaction ADN-protéine :

- il n'est actuellement pas aisé de coupler spécifiquement les QDots aux protéines d'intérêt en maîtrisant la stœchiométrie.
- leur taille, supérieure à une quinzaine de nanomètres peut gêner dans certaines applications.
- leur scintillement rend délicate toute expérience de suivi temporel même s'il existe aujourd'hui des algorithmes de suivi de fluorophores clignotants au cours du temps [42].



**Figure II.4:** Schéma de principe du microscope à épifluorescence éclairé avec une lampe à mercure ou un laser.

C'est pour ces raisons que nous leur avons préféré les fluorophores organiques dans un premier temps.

### II.1.2.2 Les différentes techniques de microscopie de fluorescence :

Différentes techniques de microscopie de fluorescence ont été développées dans le but d'exploration du monde biologique. L'apparition ces vingt dernières années de nouvelles technologies optiques (sondes plus brillantes, plus stables) et électroniques (caméras plus sensibles, plus rapides) a permis d'améliorer les résolutions temporelles et spatiales faisant de la microscopie de fluorescence un outil très puissant. Faisons un bref aperçu des principales techniques existantes.

**Microscopie de fluorescence par transmission :** un faisceau excitateur illumine entièrement l'échantillon. La lumière de fluorescence émise est collectée par l'intermédiaire d'un objectif de microscope. Cette technique présente de nombreux inconvénients, en particulier la résolution axiale est mauvaise : elle se limite par conséquent à l'étude d'échantillons extrêmement fins.

**Microscopie à épifluorescence :** l'objectif du microscope sert à la fois à focaliser la lumière excitatrice sur l'échantillon et à collecter la lumière de fluorescence émise (cf. Figure II.4).

**Microscopie confocale :** Minsky dépose un brevet en 1957 sur le principe de la microscopie confocale de fluorescence, basé sur le filtrage spatial [43]. L'idée consiste à utiliser

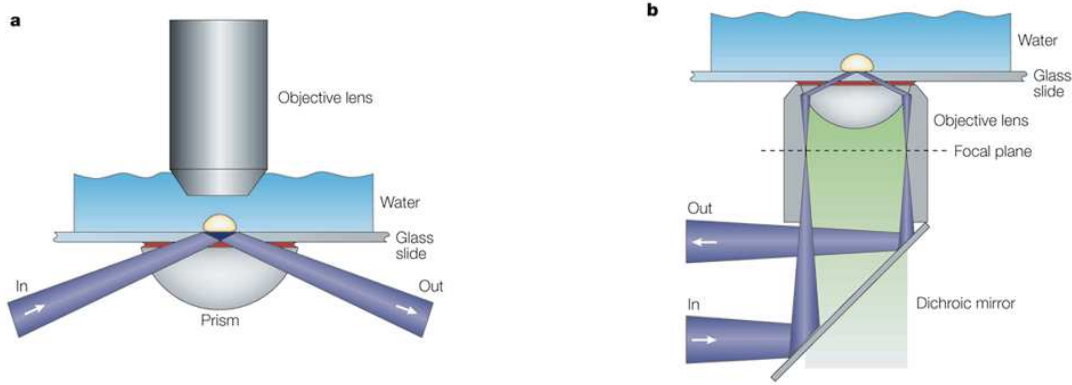
un montage à épifluorescence, sur lequel on ajoute un filtre (typiquement, un trou de filtrage) qui permet de sélectionner uniquement les photons émis depuis le plan focal. Cette technique présente l'avantage d'avoir des résolutions axiale et latérale réduites de 30% par rapport à la microscopie plein champ.

**Microscopie biphotonique :** le processus d'absorption à 2 photons se traduit par l'absorption simultanée de deux photons d'énergie  $h\nu_1$  et  $h\nu_2$ , la somme des absorptions  $h\nu_1 + h\nu_2$  étant de l'ordre de l'énergie de la transition  $h\nu_{em}$ . La possibilité de telles transitions a été théoriquement prédite par Göpper-Mayer en 1931 et expérimentalement vérifiée en 1961 par Kaiser et Garrett [44]. La microscopie biphotonique présente de nombreux avantages. Le premier réside dans son volume d'excitation, intrinsèquement confiné, ce qui permet d'atteindre des résolutions axiales tridimensionnelles comparables à celles obtenues en microscopie confocale. Ensuite, la longueur d'onde d'émission de fluorescence est environ le double de la longueur d'onde d'excitation : il est donc très facile de se débarrasser du faisceau excitateur à la détection. La microscopie biphotonique nécessite en revanche des sources de puissances suffisantes car les sections efficaces d'absorption à deux photons sont relativement faibles : les sources utilisées sont souvent des sources impulsionnelles.

### II.1.2.3 L'excitation par onde évanescente : TIRFM

À première vue, la microscopie de fluorescence classique (par épifluorescence) semble être un outil de choix pour l'étude du suivi de la dynamique des protéines, dans la mesure où la protéine est marquée. Cependant les images obtenues avec ce type de microscopie montrent qu'il est difficile d'atteindre la résolution nécessaire à l'observation d'une unique protéine. En effet, l'excitation avec une lampe de mercure a lieu sur toute l'épaisseur de l'échantillon. Ainsi, à l'image nette des protéines présentes dans le plan focal se superposent la lumière parasite, issue des protéines situées hors de ce plan. Ce « bruit de fond », superposition d'un fond continu et de fluctuations, réduit considérablement le rapport signal à bruit, rendant les observations à l'échelle de la molécule unique difficiles.

Deux approches permettent de se débarrasser du fond et d'augmenter le rapport signal à bruit. La première consiste à collecter les photons provenant uniquement du plan de focalisation en utilisant un diaphragme : il s'agit de la microscopie confocale [43]. L'autre approche repose sur un champ d'excitation limité, c'est le principe de la microscopie par onde évanescente (TIRFM pour *Total Internal Reflection Fluorescence Microscopy*). Dans ce cas, la technique pour limiter le volume d'excitation consiste à utiliser le phénomène de réflexion totale au niveau d'une interface verre/milieu. Une onde évanescente, de quelques centaines de nanomètres de portée apparaît de l'autre côté de la surface. Seuls les fluorophores proches de cette surface sont alors excités.



**Figure II.5:** Les deux catégories de montage utilisées en microscopie TIRFM. **(a)** : Montage avec prisme. **(b)** : Montage pour lequel l'excitation de la fluorescence et sa détection se font à travers l'objectif. D'après [45].

**Principe** L'onde évanescente apparait lorsqu'un rayon lumineux se réfléchit totalement sur un dioptre séparant un milieu d'indice  $n_1$  d'un milieu d'indice plus petit  $n_2$ . L'angle d'incidence  $\theta$  est supérieur à un certain angle limite :  $\theta_c = \arcsin(n_2/n_1)$ . Le champ électrique évanescent crée une intensité qui décroît exponentiellement lorsqu'on s'éloigne de l'interface :

$$I(z) = I_0 \exp\left(-\frac{z}{\delta}\right), \quad (\text{II.1})$$

où  $\delta$  est la portée de l'onde évanescente, qui varie en fonction de la longueur d'onde  $\lambda$  du rayon incident :

$$\delta = \frac{\lambda}{4\pi\sqrt{n_1^2 \sin^2(\theta) - n_2^2}}. \quad (\text{II.2})$$

**Le montage de TIRFM** Il existe deux catégories de montage permettant une observation en microscopie de TIRFM. Dans la première catégorie, le faisceau laser vient frapper l'interface verre/milieu après avoir traversé un prisme de verre, et la détection de la fluorescence a lieu du côté opposé. Nous avons opté pour l'autre catégorie en utilisant un microscope inversé : l'excitation et l'émission se font à travers l'objectif. Cette configuration nécessite l'utilisation d'objectifs à grande ouverture numérique  $NA$  puisque c'est elle qui limite l'angle d'incidence du faisceau laser au niveau de l'interface verre/milieu :  $NA = n \sin(\theta)$

En pratique, un laser est réfléchi au niveau de l'interface verre/eau ( $n_{\text{verre}} = 1,5$  et  $n_{\text{eau}} = 1,33$ ), pour laquelle  $\theta_C = 63^\circ$  par l'intermédiaire d'un objectif à immersion à huile ( $n_{\text{huile}} = 1,55$ ) de grande ouverture numérique :  $NA = 1,45$ . Le rayon le plus incliné pouvant entrer dans l'objectif et en ressortir fait un angle  $\theta = 69,3^\circ$ . Nous pouvons alors estimer la portée  $\delta$  de l'onde évanescente pour les deux lasers que nous avons utilisés :

$$\delta_{532\text{nm}} \approx 95 \text{ nm} \text{ et } \delta_{640\text{nm}} \approx 114 \text{ nm} \quad (\text{II.3})$$

### II.1.3 Aperçu des autres approches moléculaires

Les expériences de micromanipulation de molécules uniques, destinées à élargir le champ d'investigation de la biologie moléculaire, se sont développées il y a quelques années. Parmi les avancées marquantes, nous pouvons citer celle du groupe de Kinoshita (Yokohama, Japon) qui a réussi à visualiser par microscopie, des moteurs moléculaires rotatifs de 10 nm de diamètre, tournant à un rythme effréné (8000 tr/min) comportant un stator et un rotor en le marquant avec un filament d'actine fluorescent [46].

Les techniques de micromanipulation de molécules biologiques sont variées : pinces optiques, pinces magnétiques, microscopie AFM, les manipulations d'objet par contact (micro-pipette)... Dans la suite, nous nous attacherons à ne décrire que quelques uns des nombreux travaux appliqués aux études des interactions ADN-protéine à l'échelle de la molécule unique.

De nombreuses expériences de micromanipulation de molécules d'ADN utilisent des billes fonctionnalisées attachées à une (plus rarement deux) extrémités de la molécule. La méthode la plus usuelle consiste à attacher l'une des extrémités de l'ADN à une lamelle de verre et l'autre à une microbille. La position des billes est repérée par microscopie et mesurée simultanément. Selon les techniques, la bille peut être libre de fluctuer sous l'effet de l'agitation thermique ou manipulée par l'expérimentateur. Dans ce cas, la longueur et/ou la torsion de l'ADN peuvent être contrôlés par action sur la bille. L'interaction d'une protéine exerçant une force sur l'ADN est susceptible de modifier la topologie de l'ADN : les variations de taille et/ou de torsion de la molécule d'ADN reflètent ainsi l'action des protéines.

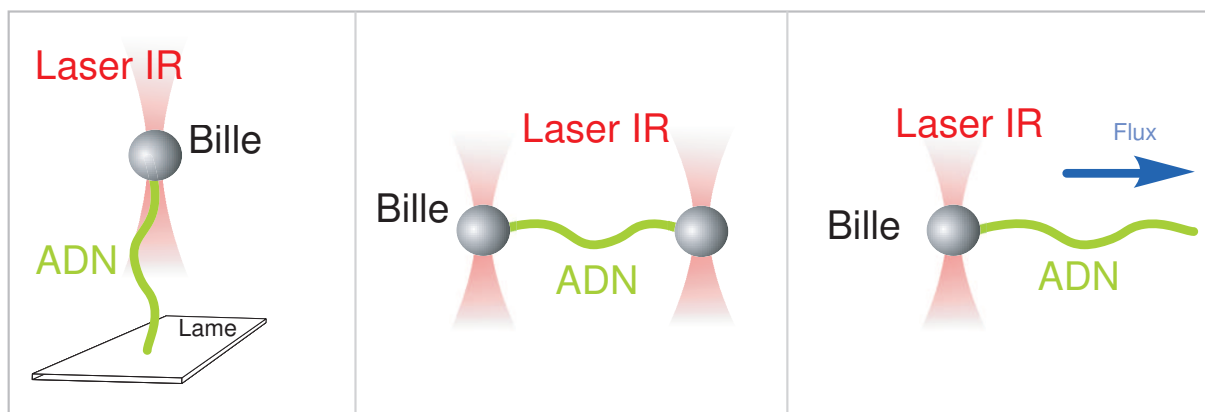
#### II.1.3.1 Bille non piégée

Une première méthode consiste à attacher une extrémité de l'ADN à une surface et à coupler l'autre extrémité à une bille de taille micrométrique. Les actions des enzymes modifiant la longueur de l'ADN se répercutent sur les positions de la bille et peuvent ainsi être caractérisées [47, 48, 49].

#### II.1.3.2 Les pinces optiques

Les pinces optiques utilisent des billes micrométriques diélectriques que l'on piège au point de focalisation d'un faisceau laser. La force créée est proportionnelle au gradient de l'intensité du champ électrique et peut-être déterminée par mesure interférométrique de la position de la bille par rapport au centre du piège. Les forces ainsi créées (de quelques à quelques centaines de pN) permettent de contrôler l'extension de l'ADN [50].

Différentes configurations utilisant des pièges optiques existent (cf. Figure II.6). Lorsque la molécule d'ADN est fixée par l'une de ses extrémités à la surface et à la bille par l'autre



**Figure II.6: Pincés Optiques** : les différentes configurations des expériences utilisant des pincés optiques pour étudier les interactions ADN-protéine. **Gauche** : l'ADN est fixé à la surface par une de ses extrémités et à une bille par l'autre. **Milieu** : double piège optique pour lequel chaque extrémité de l'ADN est fixée à une bille. **Droite** : l'ADN est attaché à une bille par une de ses extrémités et étiré par un flux hydrodynamique.

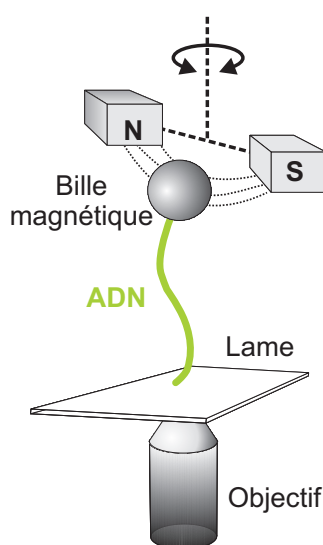
bout, la force qu'exerce une enzyme sur l'ADN peut être déterminée [51]. Avec deux pincés optiques, l'ADN peut être étiré, et l'action d'une enzyme de restriction [52] ou d'une polymérase [53] peut ainsi être étudiée. En appliquant un flux hydrodynamique à une molécule d'ADN tenue par une pince d'un côté, celle-ci se déplie : l'ADN étiré peut alors servir de substrat à des protéines. Dans la plupart de ce type d'expérience, la protéine étudiée n'était pas fluorescente ; l'ADN en revanche était teinté à l'aide d'une molécule fluorescente. L'assemblage de la chromatine [54], la digestion par une exonucléase [55] ou par une endonucléase [56], la dénaturation par une hélicase [57] ont été ainsi récemment étudiés.

L'étude en microscopie de fluorescence en présence de pincés optiques n'est pas simple. La taille micrométrique des billes utilisées ne rend possible l'utilisation d'une imagerie en onde évanescente qu'au prix d'un dispositif expérimental complexe, dans lequel la surface de verre utilisée est structurée [53]. De fait, les utilisateurs de pincés optiques voulant utiliser la microscopie de fluorescence ont souvent recours à l'illumination de l'ensemble de l'échantillon (il s'agit de l'excitation en épifluorescence), configuration qui favorise le photoblanchiment. En outre, les billes diélectriques diffusent la lumière IR utilisée pour les piéger, diminuant le rapport signal à bruit de détection en épifluorescence..

### II.1.3.3 Les pincés magnétiques

En accrochant les extrémités d'une molécule d'ADN à une surface de verre et une microbille paramagnétique, il est possible de manipuler mécaniquement l'ADN en agissant sur la bille au moyen d'un champ magnétique produit par des aimants (cf. Figure II.7).

L'ADN n'a pas besoin d'être marqué, seule la bille est observée. Ces aimants sont réglables en hauteur et peuvent tourner sur eux-mêmes permettant de contrôler ainsi la contrainte de traction et de torsion appliquée sur la molécule d'ADN. L'analyse du mouvement brownien de la bille permet de déterminer la force de traction exercée sur l'ADN et la mesure de la hauteur de la bille par rapport à la surface de verre, traduit la valeur de l'extension de la molécule d'ADN. Dans ces expériences de pinces magnétiques, l'ADN est toujours perpendiculaire à la surface à laquelle il est attaché. Si un processus biologique a pour effet de faire varier l'extension de la molécule d'ADN, il est alors possible d'étudier ce processus en temps réel. Les systèmes de pièges magnétiques ont fourni de nombreuses études sur les topoisomérases [58, 59].



**Figure II.7: Pinces Magnétiques :** la bille, de taille quelques microns, est placée dans un gradient de champ magnétique généré par des aimants permanents, ce qui permet de contrôler la force appliquée sur l'ADN. De plus, la rotation du champ magnétique entraîne la bille et induit une contrainte de torsion sur l'ADN (paramètre important pour sa fonction biologique). Les forces sont mesurées, sans calibration, par simple mesure du mouvement Brownien de la bille. Les forces appliquées vont de la dizaine de fN à la centaine de pN.

Une éventuelle utilisation de la microscopie de fluorescence par onde évanescente (excitation sur une centaine de nm) serait restreinte à l'étude de l'ADN sur une zone très proche de la surface. À notre connaissance, aucun travail n'a été publié en microscopie de fluorescence dans ces configurations pour étudier des interactions entre ADN et protéines.

### II.1.3.4 ADN fixé par une de ses extrémités à la surface

Une autre approche consiste à attacher directement une des extrémités de l'ADN à une surface. L'ADN est ensuite étiré à l'aide d'un champ électrique [60, 61] ou d'un écoulement [62, 63, 64]. La préparation de surfaces permettant l'utilisation de techniques électrophorétiques est délicate, ce qui représente un inconvénient majeur : de nombreux

tests sont en effet nécessaires à la conduite de ces expériences, par exemple pour déterminer les concentrations optimales en protéines ou réduire les interactions non-spécifiques des protéines avec la surface. Il est donc indispensable de pouvoir bénéficier facilement et rapidement d'un nombre important de surfaces. De plus, les champs électriques élevés utilisés (supérieurs à  $10^6$  V/m dans [60]) peuvent perturber les solutions.

### II.1.3.5 Le peignage moléculaire

Le peignage est une autre technique permettant de déplier les molécules d'ADN [65, 66]. Il permet simplement d'étirer parallèlement un grand nombre de molécules d'ADN sur une surface dans des conditions où elles peuvent s'y accrocher par leurs extrémités. En pratique, il faut placer l'ADN à un pH acide (typiquement pH=5,5) pour que les extrémités de l'ADN se désapparent et puissent s'ancrer à une surface hydrophobe. La passage d'un ménisque air-eau déroule alors la molécule d'ADN, et la colle à la surface. Expérimentalement, on peut, au choix, plonger une lame de verre rendue hydrophobe dans une solution d'ADN à un pH=5.5 durant quelques minutes puis la retirer, ou déposer une microgoutte d'une solution d'ADN sur une lamelle hydrophobe et attendre son évaporation qui provoque la rétraction de la ligne de contact de la goutte. Nous suggérons au lecteur désireux de connaître les détails qui ne sont pas présentés ici de consulter les références [67, 68].

Les molécules ainsi peignées peuvent servir de substrat à des protéines. Deux études ont ainsi permis de mettre en évidence l'incorporation de nucléotides fluorescents par une polymérase à ADN [69] et à ARN [70].

C'est une technique simple à mettre en œuvre puisqu'elle ne requiert que l'utilisation de surfaces hydrophobes et ne nécessite aucune modification des molécules d'ADN. Néanmoins, le surétirement des molécules d'ADN (150% en moyenne) et leurs nombreux points d'ancrage à la surface limitent l'utilisation du peignage pour l'étude des interactions ADN-protéines.

### II.1.3.6 La microscopie de force atomique (AFM)

Le principe de base de la microscopie de force atomique (AFM pour *Atomic Force Microscopy*) repose sur les travaux de Tabor et Israelachvili [71]. Il consiste à balayer la surface d'un objet par une pointe de taille nanométrique suffisamment fine pour détecter, en terme de force, l'effet de l'interaction de cette pointe avec les atomes de la surface. L'AFM permet alors de reconstruire la topologie de la surface. Utilisé en biologie, l'AFM permet d'observer à l'air ambiant ou même en milieu liquide les molécules individuelles adsorbées sur une surface, avec une résolution nanométrique ! Les surfaces utilisées doivent être planes et propres à l'échelle atomique, c'est pour cette raison que l'on utilise le mica pour y déposer les molécules d'intérêt. La résolution temporelle d'un AFM n'est



actuellement pas performante : quelques dizaines de secondes séparent deux images.

En utilisant un AFM, une équipe américaine a pu observer qualitativement en 1994 le sliding de l'ARN polymérase sur l'ADN, dans des conditions où les interactions ADN-protéines sont fortement perturbées par les interactions des protéines avec les surfaces de mica [72]. Plus tard, en 1998, une équipe de Twente a observé la dynamique d'interaction de la photolyase à l'ADN [73].

Dans le but d'étudier la dynamique de l'interaction ADN-EcoRV, nous ne pouvons utiliser l'AFM puisqu'il n'est pas adapté aux études sur des échelles de temps caractéristiques inférieures à la seconde.

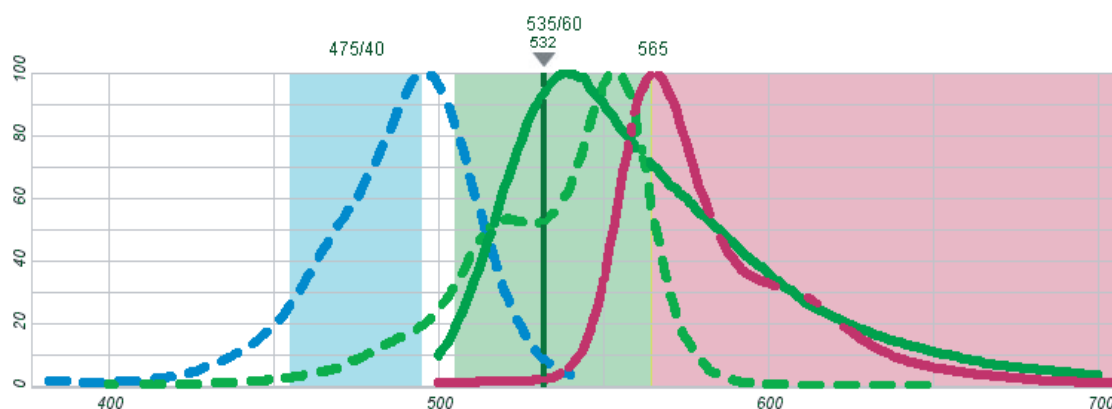
## II.2 Rendre les enzymes fluorescentes

### II.2.1 Un lourd cahier des charges

**Contrôle de la stœchiométrie :** le but d'un marquage fluorescent pour des études à l'échelle de la molécule individuelle est de pouvoir suivre une unique particule, qui reflète le mouvement d'une seule molécule d'intérêt biologique, en l'occurrence EcoRV. Jusqu'à maintenant, les expériences de visualisation d'interactions ADN-enzymes faites dans l'équipe avaient été effectuées avec un excès d'enzymes : en moyenne, une dizaine d'enzymes par nanocristal [68]. Avec une telle stœchiométrie, on ne peut pas rejeter l'idée d'un mouvement coordonné des enzymes le long de la molécule d'ADN. En effet, pourquoi ne pas imaginer un processus de « rolling » qui consisterait en des diffusions successives de plusieurs enzymes d'un même nanocristal accompagné d'un mouvement de rotation du nanocristal sur lui même ? Pour lever cette ambiguïté, il faut absolument choisir la stœchiométrie pour ne pas avoir plusieurs protéines par fluorophores. En outre, pour éviter les signaux factices, il ne faut pas avoir de fluorophores libres (i.e. non couplés). L'idéal est donc de travailler avec au plus un fluorophore par protéine.

Pour s'affranchir du signal de fluorescence résiduel émis par le fluorophore qui sera utilisé pour repérer l'ADN (*SyBR Gold* émettant dans le vert, *Molecular Probes*, cf. Section II.3), nous avons choisi d'utiliser des fluorophores qui émettent dans le rouge pour marquer les protéines : Cy3B, sav-Cy3, QDot605 streptavidine (cf. Figure II.8). Ces 3 fluorophores présentent l'avantage d'avoir des tailles différentes (cf. Section II.2.4.3), ils permettront ainsi d'étudier l'influence de la taille du fluorophore sur les caractéristiques physique de la diffusion facilitée d'EcoRV.

**Des enzymes encore actives après couplage :** rendre les enzymes fluorescentes ne doit pas altérer leur activité biologique. À la fin des différents couplages, nous faisons systématiquement plusieurs tests (détaillés après) qui permettent de vérifier que les enzymes



**Figure II.8:** Spectres d'émission (trait pointillé) et d'absorption (trait plein) du SyBR (bleu et vert) et du Cy3B (vert et rouge). Les filtres utilisés pour la détection de l'ADN sont : excitation (475AF40), dichroïque (505DRLP) et détection (535AF60). Les filtres utilisés pour la détection des protéines sont : dichroïque (560DRLP) et détection (565ALP).

sont toujours efficaces (vis-à-vis de leur site de coupure).

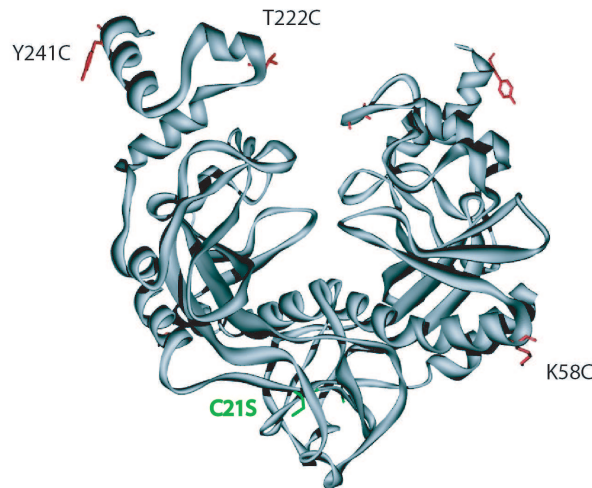
### II.2.2 Des enzymes mutées pour un couplage facilité

Nos collaborateurs allemands (A. Pingoud et W. Wende, Giessen, Allemagne) nous fournissent en enzymes EcoRV modifiées. EcoRV possède naturellement une cystéine relativement proche du site de fixation de l'enzyme à l'ADN (c21). Cette cystéine a été remplacée par une sérine pour tous les mutants dont nous disposons. Ces mutants C21S présentent les mêmes propriétés cinétiques et d'attachement à l'ADN. Les différentes mutations consistent ensuite à placer un nouveau résidu cystéine à des positions bien définies : K58C, T222C, Y241C (cf. Figure II.9). Grâce à la chimie des maléimides, ces résidus sont modifiés pour y ajouter une biotine ou le fluorophore Alexa488.

Pour résumer, nos collaborateurs nous ont donné des enzymes EcoRV biotinilées et des enzymes marquées au colorant Alexa488. Ces enzymes sont biologiquement actives.

### II.2.3 Les différents couplages

Tous les couplages ont lieu sur la glace pour ne pas réchauffer les enzymes, dans des tampons dont la concentration en sel est supérieure à 100 mM afin d'éviter l'agrégation des protéines. À l'issue des couplages, quelques microlitres sont mis de côté pour les différents tests d'activité et on réalise avec le reste un stock glycérol (50 % v/v) qui permet de conserver les enzymes pendant plusieurs mois au congélateur à - 20°C.



**Figure II.9:** Les deux mutations effectuées sur EcoRV par W. Wende (groupe d'A. Pingoud) : C21S et selon l'échantillon K58C, T222C et Y241C.

### II.2.3.1 EcoRV - Cy3B

Nous utilisons des colorants Cy3B fonctionnalisés à des esters de maléimide (*Amersham*, cf. Figure II.2) qui réagissent sur les groupements sulfhydryles -SH libres des protéines. Le couplage est effectué en suivant le protocole commercial. Nous travaillons avec le mutant K58C/C21S de concentration 423,1  $\mu\text{M}$ , qui possède une cystéine accessible.

Nous travaillons avec des concentrations initiales en enzyme de 40  $\mu\text{M}$ , et à un pH proche de 7,5 pour exalter le caractère nucléophile des groupements thiols, et favoriser ainsi leur réaction avec l'ester de maléimide.

**Réduction d'EcoRV** La première étape consiste à réduire les ponts disulfures d'EcoRV par ajout d'un large excès (100 molaire) de TCEP (*Amersham*) frais sous diazote. Après réduction par le TCEP, il faut éliminer tous les groupements -SH résiduels<sup>2</sup>, cela se fait par passage de la solution d'enzymes sur des colonnes d'exclusion (ZEBA Micro Desalt Spin Columns, *Pierce*, réf : 89877). Le rendement de cette purification est estimé à 85 % par mesure de l'absorption à 280 nm.

**Élimination des colorants libres** Après couplage, les enzymes sont purifiées des colorants libres en utilisant ces mêmes colonnes, après en avoir vérifié l'efficacité sur des colorants seuls. À cette étape, il faut ajouter 100 mM de sel à la solution d'enzymes purifiées.

**Couplage** Deux couplages différents, à des ratio  $n_{\text{dye}}/n_{\text{prot}}$  différents ont été faits. Le premier pour un ratio  $n_{\text{dye}}/n_{\text{prot}} = 1$  et le second pour un ratio  $n_{\text{dye}}/n_{\text{prot}} = 5$ . Dans les deux cas, la proportion d'enzymes effectivement marquées a été estimée à 0,4 (par FCS

<sup>2</sup>Il y initialement 0,1 mM de DTT dans le tampon de conservation des enzymes.

et mesure des absorptions). À l'issue du couplage, nous disposons de solution d'enzymes d'EcoRV de concentration 15  $\mu\text{M}$ , dont un peu plus d'un tiers est marquée au Cy3B.

### II.2.3.2 EcoRV - streptavidine-Cy3

**Couplage** Nous avons utilisé le mutant K58C WT biotinilé de concentration  $c = 0,8 \mu\text{M}$ . La streptavidine-Cy3 (*Jackson*, réf : 016-160-084) utilisée est à  $0,9 \text{ mg/mL} = 15 \mu\text{M}$ . Le couplage est réalisé pour avoir 2 fois plus de moles de streptavidine-Cy3 que d'enzymes dans le Pipes à  $\text{pH}=6,8$  contenant de la caséine à  $1 \text{ mg/mL}$  et assurant une concentration totale en sel de  $100 \text{ mM}$ . Le couplage a lieu dans la glace pendant 1h, avec agitation régulière douce. Dans ces conditions, aucune purification n'est nécessaire. On dispose finalement d'enzymes marquées de concentration  $0,4 \mu\text{M}$ .

### II.2.3.3 EcoRV - QD

**Couplage** Nous procédons de la même manière que pour le couplage à la streptavidine-Cy3. En revanche, les enzymes couplées aux nanocristaux ne sont pas conservées au congélateur, mais seulement quelques jours au réfrigérateur.

**Purification** Nous pouvons éventuellement purifier les nanocristaux couplés des enzymes libres en utilisant des colonnes de chromatographie d'exclusion de taille (colonnes de séphadex, *Superdex 200 prep* chez *Amersham*). Les enzymes libres, de taille  $60 \text{ kDa}$ , sont retardées en se piégeant dans les pores de la colonne et éluent beaucoup plus lentement que les nanocristaux de taille  $500 \text{ kDa}$  ou que les complexes (nanocristaux + enzymes). Après un étalonnage de la colonne en faisant migrer un mélange de nanocristaux et d'enzymes fluorescentes marquées à l'Alexa488 pour estimer les vitesses de migration, les gouttes ne contenant que les nanocristaux couplés (premières gouttes fluorescentes à la lampe UV) sont récupérées.

## II.2.4 Réussite du couplage et contrôle de l'activité

### II.2.4.1 Vérification de l'activité : digestion sur gel

Après couplage, l'activité biologique des enzymes (reconnaissance et coupure de la séquence cible) est systématiquement vérifiée par électrophorèse sur gel d'agarose. Pour cela, nous incubons de l'ADN  $\lambda$  qui contient 21 séquences cibles pour l'enzyme EcoRV. La digestion a lieu dans le tampon recommandé par *New England Biolabs* : le NEB3. Deux témoins (ADN  $\lambda$  sans enzyme et ADN  $\lambda$  avec enzyme non couplées) sont systématiquement réalisés. La figure II.10 est la photo d'un gel d'agarose réalisé pour tester l'activité des enzymes couplées au Cy3B : les enzymes couplées coupent encore l'ADN  $\lambda$ , preuve que le couplage n'altère pas (du moins, pas trop) l'activité biologique des enzymes.



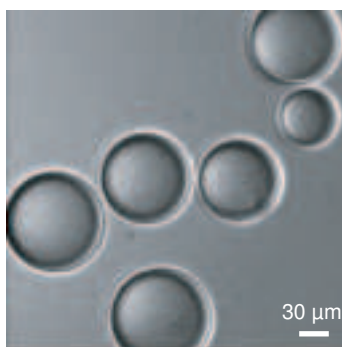
**Figure II.10:** Gel d'agarose à 0,7 % après digestion de 100 ng d'ADN  $\lambda$ . **Puits n° 1 et 5 :** marqueurs de taille 1kb (*Fermentas*). **Puits n° 2 :** ADN  $\lambda$  seul (témoin). **Puits n° 3 :** ADN  $\lambda$  + EcoRV-Y241C WT avant couplage. **Puits n° 4 :** ADN  $\lambda$  + EcoRV-K58C WT après couplage au Cy3B

### II.2.4.2 Contrôle du marquage : test à l'héparine

L'héparine est un polymère chargé dont la structure ressemble à celle de l'ADN. À force ionique physiologique, les endonucléases présentent une forte affinité pour l'héparine. En revanche, pour des concentrations en sel élevées (typiquement supérieure à 700 mM), les enzymes se détachent de l'héparine.

L'idée est de tester l'affinité entre l'héparine et les enzymes éventuellement marquées par analyse de la fluorescence de billes d'héparine pour déterminer s'il y a ou non des protéines réellement marquées. Pratiquement, nous utilisons des billes d'agarose recouvertes d'héparine (*Sigma-Aldrich*, réf : H3392). Macroscopiquement, une solution d'héparine est une suspension trouble qui sédimente rapidement (les billes tombent au fond du tube sous l'effet de leur masse en 1 à 2 minutes). Microscopiquement, les billes d'agarose présentent des tailles dispersées (cf. Figure II.11).

**Protocole** Les billes d'héparine sont d'abord lavées plusieurs fois avec une solution de tensioactifs (typiquement, le Triton à 0,1 %) pour éliminer l'éthanol présent dans le tampon de conservation. Elles sont alors resuspendues dans un tampon approprié pour EcoRV : nous avons choisi, par simplicité, le NEB Buffer 3 de chez *New England Biolabs*, auquel nous avons ajouté 0,1% de caséine.



**Figure II.11: Suspension de billes d'héparine :** la taille des billes varie entre 40 et 200  $\mu\text{m}$ .

La première étape consiste à mettre en présence un large excès d'enzymes couplées avec les billes d'héparine lavées pendant 30 min sous agitation permanente dans la glace dans un tampon contenant 1 mg/mL de caséine. Deux témoins sont réalisés : l'un avec les billes seules, le second avec le fluorophore utilisé pour les différents couplages. Après ce temps d'incubation, les billes sont rincées avec le NEB3 afin d'éliminer les enzymes non fixées à l'héparine. Quelques  $\mu\text{L}$  de la solution sont déposés sur une lamelle de microscope et une image de fluorescence est prise. Le reste de la solution est ensuite élué au sel 1M puis resuspendu dans le NEB3 : une seconde image de fluorescence est prise. Toutes les images de fluorescence apparaissent dans le tableau II.1 avec la même échelle de gris. Trois observations importantes peuvent être faites :

- en présence d'enzymes fluorescentes, la tache fluorescente des billes est annulaire
- l'élué au sel 1M fait disparaître la fluorescence des billes
- l'adsorption non-spécifique des fluorophores (cas d'Alexa488) sur les billes génère une fluorescence uniforme qui ne disparaît pas après élué au sel.

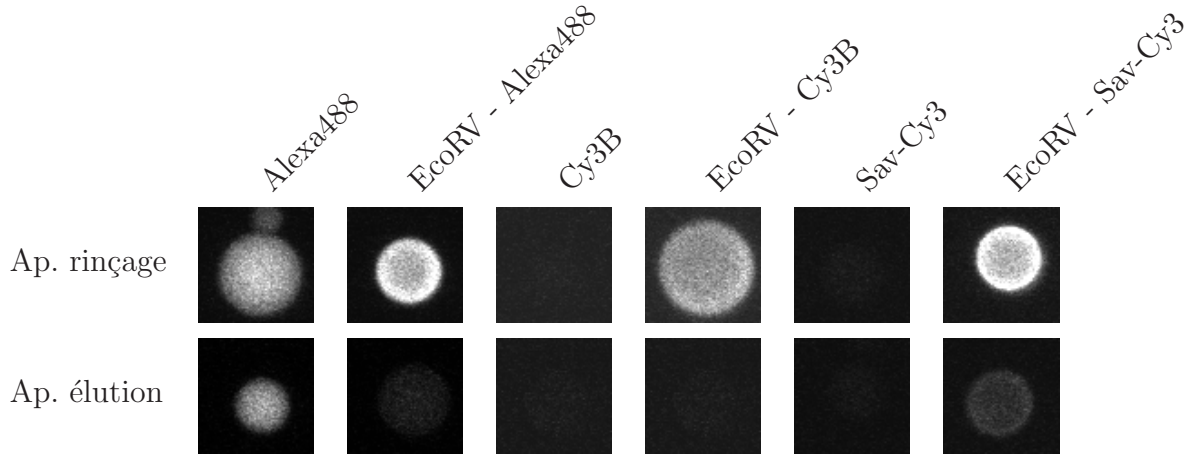
L'apparence annulaire de la tache de fluorescence dénote une répartition uniforme des fluorophores sur la surface des billes : l'interaction spécifique enzymes-héparine est, sans surprise, homogène.

Après un couplage, l'observation d'une tache annulaire sur les billes d'héparine est une preuve de la présence d'enzymes fluorescentes dans notre échantillon.

### II.2.4.3 Détermination de la taille du complexe enzyme-fluorophore : FCS à 2 photons

*Acquisitions faites par P. Neveu & J.-F. Allemand, LPS, ENS.*

Nous utilisons la formule de Stokes-Einstein qui relie le coefficient de diffusion tridimensionnel  $D_3$  au rayon hydrodynamique  $R_H$  en fonction de  $\eta = 10^{-3}$  Pa.s, la viscosité



**Tableau II.1:** Images de fluorescence de billes d'héparine individuelles. Témoins : colorants seuls (Alexa488, Cy3B et streptavidine-Cy3) et enzymes couplées. Test positif pour les couplages : EcoRV - Cy3B et EcoRV - streptavidine-Cy3

dynamique du milieu :

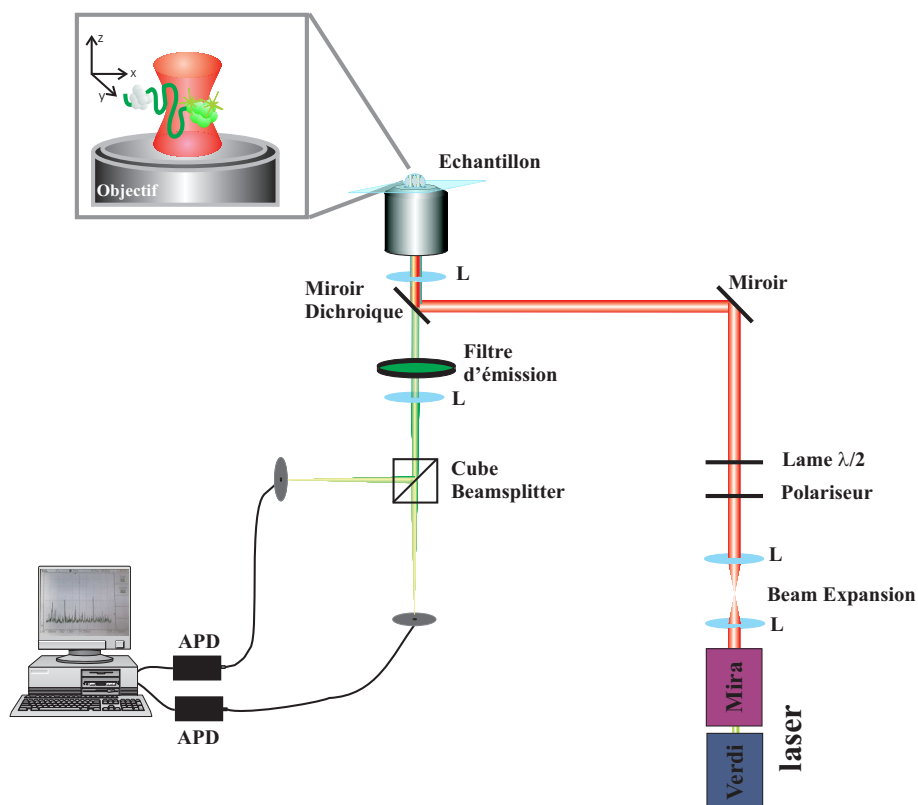
$$D_3 = \frac{k_B T}{6\pi\eta R_H}. \quad (\text{II.4})$$

**Présentation de la FCS** La microscopie de corrélation des fluctuations de l'intensité de fluorescence (FCS pour *Fluctuation Correlation Spectroscopy*) est une technique de haute résolution temporelle et spatiale d'analyse de milieux très dilués en molécules fluorescentes basée sur les fluctuations temporelles de l'intensité de fluorescence. L'analyse de ces corrélations permet d'obtenir des informations très précises sur les phénomènes responsables de ces fluctuations d'intensité : diffusion moléculaire, diffusion membranaire, réactions photochimiques, ... Les principales applications de cette technique sont les mesures de coefficients de diffusion et la détermination de la concentration en fluorophore dans le milieu d'investigation.

**Aspects théoriques** Dans notre situation, les mesures de FCS reposent sur l'analyse des fluctuations d'intensité liées à la diffusion de biomolécules colorées en solution aqueuse qui traversent un faisceau laser (cf. Figure II.12). Le signal de fluorescence analysé provient d'un volume bien défini de la solution, volume caractéristique du montage et qui doit être préalablement déterminé. Les fluctuations d'intensité sont détectées par des photorécepteurs.

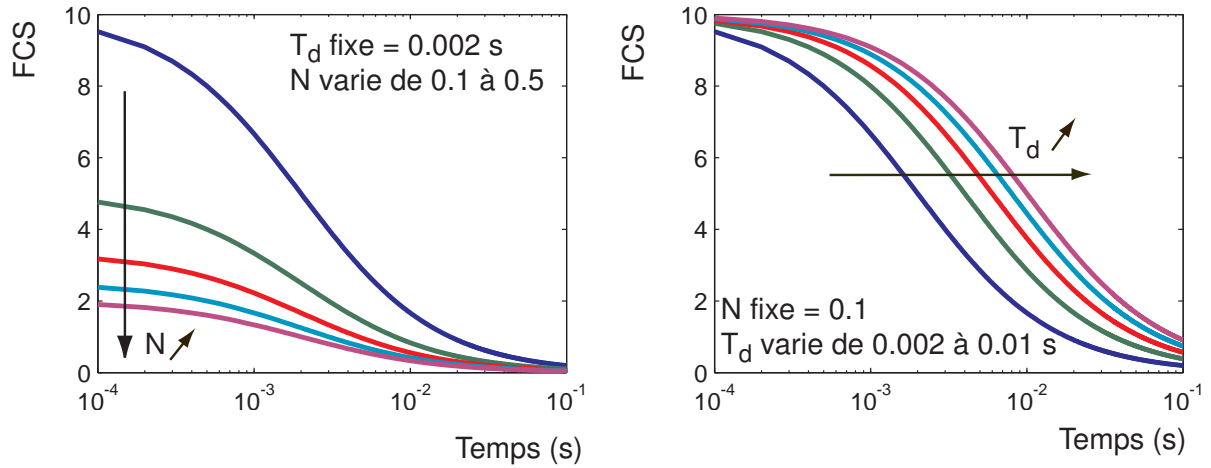
La fonction d'autocorrélation temporelle des fluctuations du signal  $G(\tau)$  de fluorescence  $F$  permet d'accéder au temps caractéristique de cette dynamique de fluorophores et la variance fournit le nombre moyen de molécules présentes dans le volume d'excitation :

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t + \tau) \rangle}{\langle F(t) \rangle^2}. \quad (\text{II.5})$$



**Figure II.12: Schéma d'une expérience de FCS.** Le faisceau laser excitateur (732 nm) apparaît en rouge. En vert, les fluorophores excités lorsqu'ils traversent le volume focal. Sur l'écran d'ordinateur, on peut voir un oscillographe obtenu avec une solution de rhodamine à une concentration  $c = 2 \text{ nM}$ . Dans ces conditions de concentration, on voit apparaître régulièrement des pics qui se détachent du bruit, correspondant au passage d'une molécule fluorescente dans le volume de référence. Les temps d'arrivée sont rangés dans les registres d'une carte informatique comportant une horloge interne, et analysés par l'intermédiaire d'un logiciel « Multi Tau Correlation » qui génère une fonction d'autocorrélation expérimentale. Cette fonction est ajustée par une fonction inspirée de celle de l'équation (II.6), donnant ainsi accès au temps de diffusion de l'espèce fluorescente et au nombre moyen de molécules dans le volume de référence.





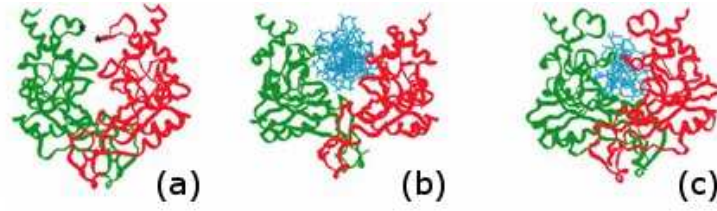
**Figure II.13:** Explication de la lecture d'une courbe de FCS. Simulation de courbes de FCS à une seule espèce fluorescente. **Gauche** : le temps de diffusion de l'espèce a été fixé à 2 ms, le nombre moyen de particule  $N$  variant de 0,1 à 0,5 : moins il y a d'espèces fluorescentes, plus l'ordonnée à l'origine de la courbe de FCS est grand. **Droite** : le nombre de particules diffusantes est fixé à 0,1, le temps de diffusion variant de 0,002 à 0,01 s.

**Dispositif expérimental** L'excitation des molécules se fait par l'intermédiaire d'un laser femtoseconde 732 nm focalisé dans l'échantillon (excitation biphotonique). La lumière de fluorescence est récoltée sur des photodiodes à avalanche. Pour une solution contenant une unique espèce fluorescente, l'expression finale de la fonction d'autocorrélation normalisée des fluctuations Browniennes de l'intensité s'écrit :

$$G(\tau) = \frac{1}{N} \frac{1}{1 + \frac{\tau}{\tau_D}}, \quad (\text{II.6})$$

où  $N$  est le nombre moyen de fluorophores dans le volume d'excitation et  $\tau_D$  le temps de diffusion de l'espèce considérée dans le volume d'excitation. Nous utilisons l'équation II.6 pour ajuster les courbes expérimentales et remonter aux valeurs de  $N$  et  $\tau_D$  (cf. Figure II.13). Sous excitation biphotonique,  $\tau_D$  relie le coefficient de diffusion tri-dimensionnel  $D_3$  au *waist*  $w_{xy}$  selon :  $\tau_D = \frac{w_{xy}^2}{8D_3}$  [74].

**EcoRV -Cy3B** d'après la structure cristallographique (protéine globulaire de 57 kDa), nous pouvons évaluer le rayon de l'homodimère à 3,5 nm (cf. Figure II.14 qui permet de comparer la taille de l'ADN (2 nm) à celle de l'enzyme). Ce qui est raisonnable si on le compare aux rayons hydrodynamiques qui ont été mesurés pour des protéines par MFPP ( $R_H = 3,4 \pm 0,4$  nm pour la phosphoglucomutase de masse  $MW = 60$  kDa, d'après [75]). Les expériences de FCS que nous avons réalisées donnent un rayon  $R_H = 4$  nm lorsque l'enzyme est marquée avec Cy3B, quelque soit le tampon. Nous en déduisons le coefficient



**Figure II.14:** (a) EcoRV en solution. (b) Complexe EcoRV/ADN non spécifique. (c) Complexe EcoRV/ADN spécifique. Les deux sous-unités de l'enzyme apparaissent en vert et rouge, l'ADN en bleu. Extrait de [35]

de diffusion 3D :

$$R_H^{\text{EcoRV}} = 4 \pm 0,5 \text{ nm} \implies D_3^{\text{EcoRV}} = 54 \pm 7 \mu\text{m}^2.\text{s}^{-1} \quad (\text{II.7})$$

**EcoRV -Sav-Cy3 :** les masses moléculaires de ces deux protéines globulaires étant proches ( $MW_{\text{EcoRV}} = 57 \text{ kDa}$  et  $MW_{\text{Sav}} = 60 \text{ kDa}$ ), leurs rayons hydrodynamiques devraient être similaires. En revanche le rayon hydrodynamique de l'association des deux protéines n'est pas simplement la somme des rayons. Pour deux sphères identiques de rayon  $R_H$ , nous pouvons présumer que le rayon hydrodynamique associé aux deux sphères collées est [76, 77] :  $R = 4R_H/3$ . Dans notre cas, on attendrait donc un rayon de l'ordre de 5,3 nm. Les expériences de FCS donnent  $R_H = 5,3 \text{ nm}$ , en bon accord avec le résultat du modèle considéré. Le coefficient de diffusion 3D est alors :

$$R_H^{\text{EcoRV-Sav-Cy3}} = 5,3 \pm 0,5 \text{ nm} \implies D_3^{\text{EcoRV-Sav-Cy3}} = 40 \pm 4 \mu\text{m}^2.\text{s}^{-1} \quad (\text{II.8})$$

**EcoRV - QD 605 Sav :** nous avons également passé un échantillon de QDot605 - Sav pour caractériser leur rayon hydrodynamique. Nous avons trouvé :

$$R_H^{\text{EcoRV-QD605Sav}} = 16 \pm 0,5 \text{ nm} \implies D_3^{\text{EcoRV-QD605Sav}} = 14 \pm 0,5 \mu\text{m}^2.\text{s}^{-1}$$

Ce résultat est satisfaisant puisque dans la littérature, la taille de QDots recouverts de streptavidine est estimée à 15-20 nm [78, 79].

Complexe	$D_3 (\mu\text{m}^2.\text{s}^{-1})$	$R_H (\text{nm})$
EcoRV - Cy3B	$54 \pm 7$	$4 \pm 0,5$
EcoRV - Sav-Cy3	$40 \pm 4$	$5,3 \pm 0,5$
QD605-Sav	$14 \pm 0,5$	$16 \pm 0,5$

**Tableau II.2:** Détermination des rayons hydrodynamiques par FCS.

## II.3 Positionner correctement une molécule d'ADN

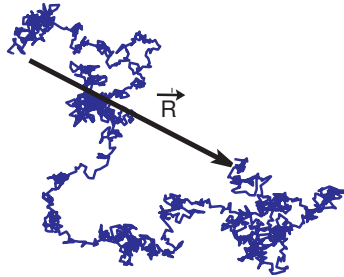
### II.3.1 Nécessité d'étirer les molécules d'ADN

Les molécules d'ADN sur lesquelles nous avons travaillé sont des ADN linéaires de taille 9,2 kb. Ces ADN, sous forme de pelotes en solution ont une taille apparente  $R_g$  donnée par :

$$R_g \approx aN^\nu, \quad (\text{II.9})$$

où  $a$  est la taille d'un maillon élémentaire (segment),  $N$  le nombre de segment et  $\nu$ , l'exposant de Flory, proche de 3/5 à 3D. En prenant comme taille de segment la longueur de persistance  $L_P$  de l'ADN ( $L_P \approx 50$  nm), nous obtenons  $R_g$  (10 kb)  $\approx 600$  nm pour un ADN de taille 10 kb.

Le modèle de chaîne idéale (ou gaussienne) est le plus simple que l'on puisse imaginer pour décrire un polymère. Dans ce modèle, on néglige toutes les interactions entre monomères, de taille constante  $a$ . Les monomères peuvent alors indépendamment les uns des autres prendre toutes les directions possibles. La structure spatiale d'une telle chaîne est équivalente à celle d'une marche aléatoire à 3D, d'où le terme de chaîne gaussienne.



Pelote Gaussienne -  $N = 2000$

Pour un nombre  $N$  de monomères (ou segments) suffisamment grand, la distance  $R$  entre les deux extrémités de la chaîne suit une loi normale.  $R$  est le rayon de la pelote occupée par le polymère,  $\vec{r}_i$  les segments de taille  $a$  :

$$\langle \vec{R}^2 \rangle = \left\langle \left( \sum_i^N \vec{r}_i \right)^2 \right\rangle$$

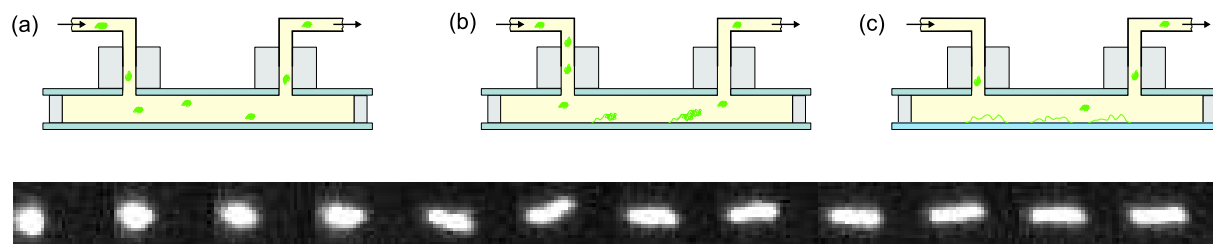
$$R = N \times a$$

Le comportement plus général des polymères a initialement été étudié par P. Flory [80]. Il montre une dépendance du rayon similaire à celle prédite par le modèle gaussien, mais avec un exposant de Flory  $\nu$  qui dépend de la dimension de l'espace.

Or la résolution d'un dispositif optique est donnée par la limite de Rayleigh :  $\delta = \frac{0,61\lambda}{NA}$  où  $NA$  est l'ouverture numérique et vaut 1,45 pour les objectifs à immersion à huile que nous utilisons, ce qui donne une résolution d'environ  $\delta \approx 223$  nm pour  $\lambda = 530$  nm. Il est donc difficile de discerner les contours d'une molécule d'ADN en solution avec notre microscope.

En 2001, l'équipe a développé une technique permettant d'étirer une molécule d'ADN sous l'action d'un flux hydrodynamique [81]. L'étirement de molécules d'ADN sur lamelle hydrophobe est une technique bien maîtrisée : ainsi étirées, les molécules d'ADN ne sont accrochées à la surface que par leurs extrémités et sont libres de fluctuer sous l'effet de

l'agitation thermique [81].



**Figure II.15: Haut :** Principe de l'étirement d'ADN sous flux (a) les molécules d'ADN sous forme de pelote sont injectées dans une cavité (b) certaines molécules peuvent s'accrocher par l'une de leurs extrémités à la surface (c) si la seconde extrémité s'accroche, la molécule est étirée. **Bas :** Extraits d'un film montrant une molécule d'ADN accrochée par une de ses extrémités s'étirant sous le cisaillement du flux hydrodynamique appliqué dans la direction gauche-droite. Le temps s'écoule de gauche à droite.

Cependant, un tel protocole n'assure pas un accrochage spécifique de l'ADN à la surface. De plus, il n'est pas rare que les molécules d'ADN se cassent au cours des manipulations (lors du pipetage par exemple) conduisant à une dispersion dans la taille des molécules étirées. C'est pour ces raisons que nous avons décidé d'améliorer la technique d'étirement afin d'assurer un accrochage spécifique des extrémités des molécules d'ADN à la surface.

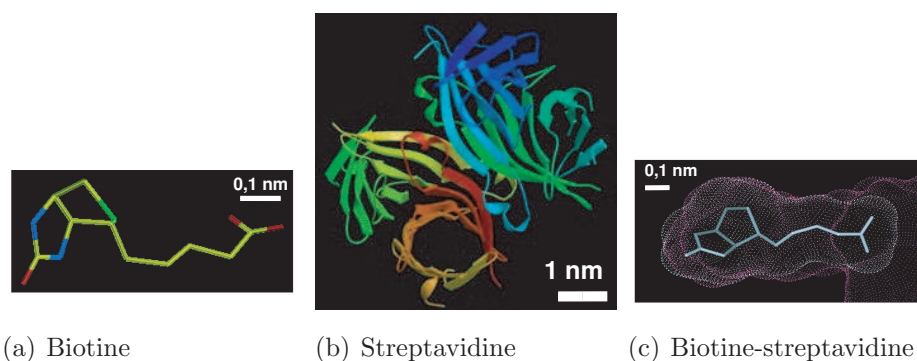
## II.3.2 L'étirement spécifique de molécules d'ADN

### II.3.2.1 Principe

Nous avons choisi de travailler sur des surfaces recouvertes de streptavidine pour étirer des molécules d'ADN aux extrémités biotinilées. Le couple biotine-streptavidine est un couple récepteur-ligand bien décrit dans la littérature qui présente l'avantage d'avoir la constante d'association la plus élevée parmi les molécules biologiques :  $K_A = 10^{15} \text{ M}^{-1}$ . Chaque molécule de streptavidine possède 4 sites de fixation à la biotine.

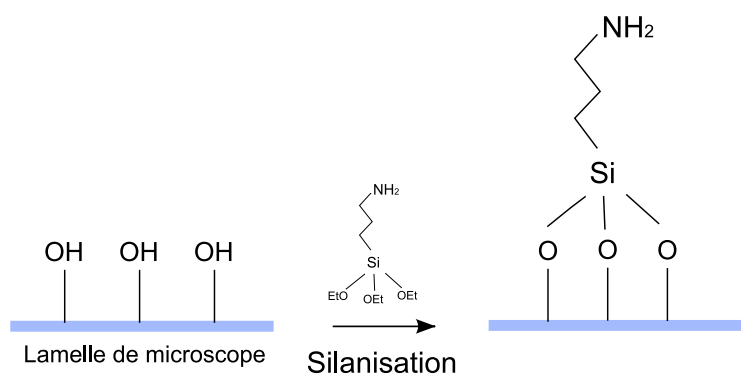
### II.3.2.2 Fabriquer une surface de streptavidine

**Nettoyer les surfaces** Avant toute modification et utilisation, les différentes surfaces que nous utilisons (lamelles, bloc de PDMS, lames porte-objet) sont nettoyées par application d'un plasma d'oxygène pendant plusieurs minutes (Plasma Cleaner, *Harrick*) : pratiquement, une pompe à huile permet de faire le vide ( $10^{-2} - 10^{-3} \text{ mbar}$ ) dans une enceinte dans laquelle est introduit de l'oxygène gazeux qui sous l'effet d'une décharge haute fréquence (micro-onde) va s'ioniser. Ce plasma permet d'enlever toutes les impuretés organiques présentes à la surface du verre et rend les surfaces hydrophiles.



**Figure II.16:** Représentation des molécules de biotine (a) et de streptavidine (b). La figure (c) est une représentation schématique d'une molécule de biotine insérée dans une des quatre poches hydrophobes de la streptavidine.

**Silanisation des surfaces** La première étape pour construire une surface de streptavidine consiste à greffer des groupements  $\text{-NH}_3^+$  à la surface du verre en utilisant des aminosilanes<sup>3</sup> (APTES, *Sigma*, réf : A3648). La silanisation des surfaces inorganiques est possible lorsqu'il y a des groupements hydroxyles  $\text{-OH}$  (comme c'est le cas pour le verre), ceux-ci réagissent avec les silanes pour former des liaisons oxanes stables  $\text{Si-O-}$  (cf. Figure II.17). Pour que la silanisation soit possible il faut d'abord se débarrasser des ions et des impuretés de la surface, nous le faisons en passant les lamelles au Plasma Cleaner.

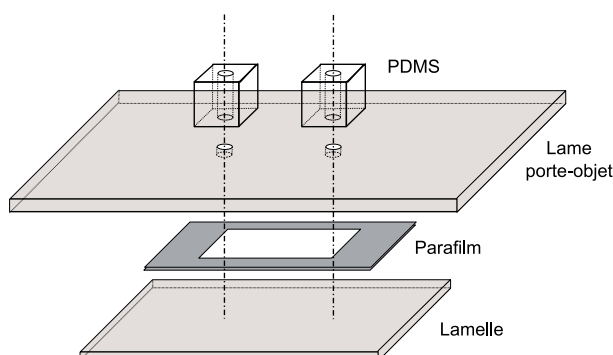


**Figure II.17:** Principe de la silanisation par APTES

En pratique, nous utilisons des lamelles couvre-objet de microscope ( $24 \times 36$  mm, *Amilabo*) que nous nettoyons d'abord à l'éthanol, puis au Plasma Cleaner. La silanisation de ces lamelles consiste à les faire tremper dans une solution contenant les APTES durant une nuit à  $4^\circ\text{C}$ . À l'issue de ce bain, les lamelles sont rincées à l'éthanol, puis à l'eau milliQ, et enfin séchées à l'air comprimé. Elles peuvent être conservées sous vide d'air minimal (cloche à vide) pendant quelques semaines. Les lamelles de verre ainsi silanisées sont donc recouvertes d'une couche d'amine  $\text{NH}_2$ , ou d'ions  $\text{NH}_3^+$  en milieu aqueux.

<sup>3</sup>Les silanes sont des composés du silicium, analogues des alcanes

**Fabrication des cavités** Les cavités que nous utilisons sont faites « maison », et réutilisables en partie. En pratique, nous perçons deux trous espacés de 15 mm dans une lame porte-objet. Sur ces deux trous sont montés des blocs de PDMS d'environ 1 cm<sup>3</sup> percés. L'adhésion du PDMS au verre se fait par légère pression après avoir passé au Plasma Cleaner PDMS et lame porte-objet : l'action du plasma oxyde la surface du PDMS et produit des groupements silanols Si–OH analogues à ceux présents sur du verre propre. Les groupements silanols du verre et du PDMS interagissent fortement entre eux, formant des liaisons hydrogènes. Cette partie de la cavité est réutilisée après nettoyage à l'éthanol et passage au Plasma Cleaner.



**Figure II.18:** Schéma des cavités à étirement

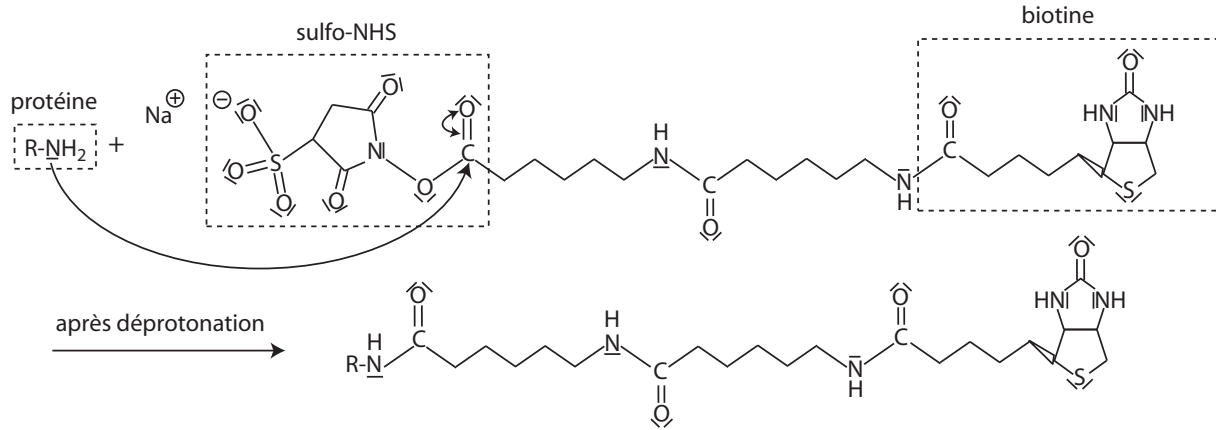
Pour fabriquer une cavité à écoulement, nous découpons dans une feuille de parafilm un canal rectangulaire de dimension 2×20 mm (cf. Figure II.18). Ce canal est placé entre une lamelle silanisée et une lame « porte-objet cavité », le tout est posé sur un bloc chauffant durant quelques secondes pour faire fondre le parafilm (50 °C) qui soude ainsi la lame à la lamelle (cf. Figure II.18).

### Préparer une surface de streptavidine

1) **Couche de caséine biotiniée** : La caséine est une protéine contenant de nombreux résidus apolaires : elle est donc amphiphile. Il en existe de différentes sortes :  $\alpha$ ,  $\beta$ ,  $\kappa$ . Ce sont des polyélectrolytes légèrement chargés qui comportent une partie hydrophile et une partie hydrophobe, elles s'apparentent donc à un tensio-actif. Elles forment des micelles (la concentration micellaire critique varie entre 0,5 et 2 mg.mL<sup>-1</sup>). Elles peuvent s'adsorber sur des surfaces hydrophiles ou hydrophobes, en exhibant l'une ou l'autre de leur région. Sur des surfaces hydrophobes, l'arrangement est assez compact, mais sur des surfaces hydrophiles, des multicouches sont facilement formées.

► **Solution de caséine** : nous utilisons du Blocking Reagent (*Roche*) en guise de caséine, principalement constitué de caséine  $\kappa$ . Le Blocking Reagent

est préalablement dilué à 10 mg/mL dans le tampon Pipes 20 mM, NaCl 150 mM, pH=6,8. La dissolution étant lente, la solution est laissée sur rotateur pendant 8 heures. Cette dilution peut-être conservée une semaine au réfrigérateur.



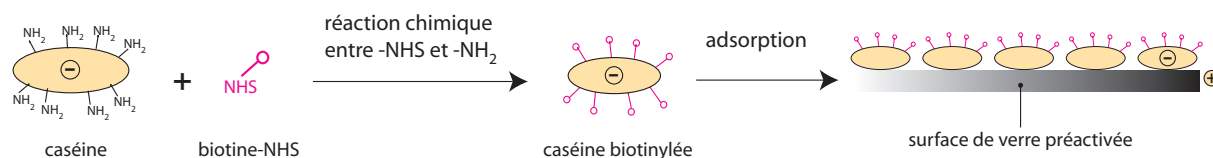
**Figure II.19:** Formule semi-développée de la sulfo -NHS-LC-LC-biotine et réaction avec une protéine.

► **La solution de caséine biotinilée** est préparée à l'aide de biotine fonctionnalisée N - hydroxysuccinimide (-NHS) qui réagissent avec les amines primaires des protéines (résidus lysines). L'emploi du groupement -NHS pour former des liaisons amides avec les amines primaires est très répandu (marquage en fluorescence de protéines, fonctionnalisation de nanoparticules), notamment du fait de sa grande solubilité dans de nombreux solvants, dont l'eau. Pour que le couplage soit efficace, le groupement amine primaire ne doit pas être ionisé, il faut donc travailler dans des tampons dont le pH est compris entre 7,5 et 9. Mais à ce pH, l'hydrolyse du groupement -NHS augmente également.

Il y a 10 lysines sur la  $\kappa$ -caséine bovine. L'idée est de mettre une biotine par lysine. Nous préparons ainsi une solution de sulfo-NHS biotine (*Pierce*, réf : 21420) à 100 mg/mL dans du DMSO. On ajoute 2  $\mu$ L de cette solution avec 40  $\mu$ L de caséine dans 160  $\mu$ L de PBS. Nous obtenons ainsi une solution pour laquelle la concentration en biotine est une vingtaine de fois supérieure à celle de la caséine (MW(biotine-NHS) = 669 Da et MW(caséine) = 24 kDa). La réaction s'effectue en 30 mn sous agitation douce à température ambiante.

► **L'adsorption de la caséine biotinilée sur les surfaces silanisées** constitue la première étape pour fabriquer une surface de streptavidine. L'adsorption de la caséine biotinilée sur la lamelle silanisée se fait car la caséine, ayant un pI de 4,5, apparaît fortement négative à un pH de 6,8 et se fixe donc électrostatiquement

aux ions ammonium. Nous faisons passer 30  $\mu\text{L}$  de caséine biotinylée dans la cavité. Après un temps d'incubation de 10 minutes, la cavité est soigneusement rincée avec le tampon de travail.



**Figure II.20:** Adsorption de la caséine biotinylée sur le verre : la réaction entre la biotine fonctionnalisée par le groupement N-hydroxysuccinimide (NHS) et la caséine permet l'obtention de caséine biotinylée que l'on adsorbe ensuite sur du verre silanisé.

**2) Couche de streptavidine :** on peut ensuite déposer la streptavidine (*Sigma*) à 0,1 mg/mL dans le Pipes 20 mM, NaCl 100 mM, pH=6,8, qui s'attache à la biotine de manière quasi-covalente. Après un deuxième temps d'incubation de 10 minutes également, la cavité est de nouveau rincée avec le Pipes 20 mM, NaCl 50 mM. La streptavidine possède 4 sites de fixation pour la biotine, en travaillant dans de telles conditions, on réussit à recouvrir tout le tapis de biotine et à laisser des biotines libres pour permettre par la suite l'accrochage de l'ADN.

Dans ces différentes étapes, il est nécessaire de travailler dans des conditions telles qu'une couche dense de molécules soit déposée à chaque fois : il faut que les ions  $\text{NH}_3^+$  recouvrent l'ensemble de la surface. De même, les caséines doivent être suffisamment nombreuses pour pouvoir entièrement tapisser les  $\text{NH}_3^+$ . Ainsi, la streptavidine ne « voit » qu'une surface homogène de biotine (cf. Figure II.21).

### II.3.2.3 Biotiniler les extrémités des molécules d'ADN

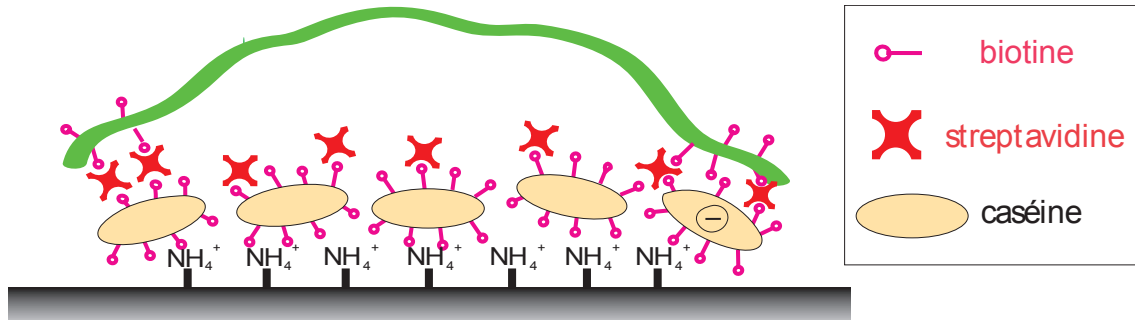
En pratique, nous synthétisons des fragments d'ADN biotinylés par PCR que nous liguons aux extrémités d'une molécule d'ADN d'intérêt (cf. Annexe A). Pour les études des interactions ADN/enzyme de restriction, cette méthode présente l'avantage de pouvoir choisir la séquence d'ADN à étirer, c'est-à-dire la présence ou non de séquences cibles. En l'occurrence, pour étudier la diffusion facilitée d'EcoRV, nous avons choisi de travailler avec des ADNs ne comportant pas de séquence cible pour cette enzyme.

### II.3.2.4 Étirer l'ADN biotinilé sur une surface de streptavidine

Une fois la surface de streptavidine réalisée, on peut injecter dans la cavité les molécules d'ADN biotinylées à une concentration proche du ng/ $\mu\text{L}$ . Une grande partie des molécules d'ADN, initialement en pelote s'attache à la surface par une de leurs extrémités. En appliquant un flux hydrodynamique, les pelotes se déroulent et les extrémités libres se



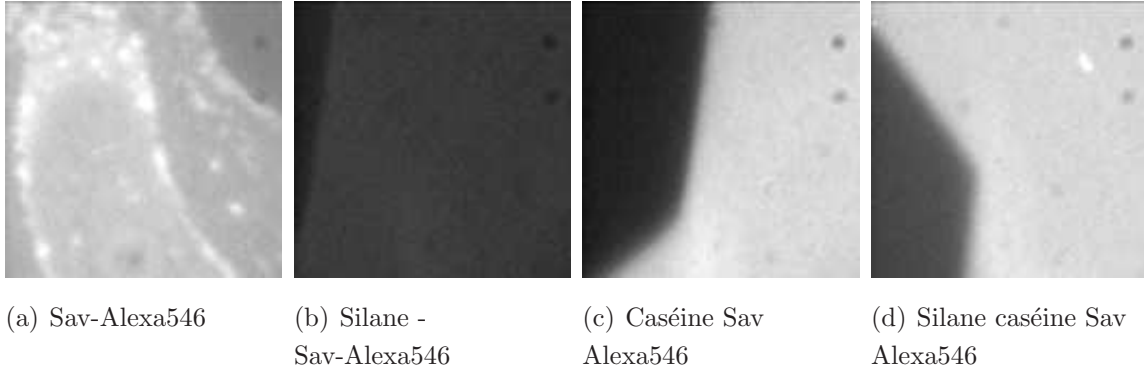
collent à leur tour à la surface : les molécules d'ADN sont ainsi étirées. L'ADN peut ensuite être visualisé en ajoutant un fluorophore tel que le *SyBR Gold* (*Molecular Probes*).



**Figure II.21:** Étirement des molécules d'ADN : schéma des différentes couches pour construire une surface de streptavidine - les échelles ne sont volontairement pas respectées.

En pratique, nous avons testé différents protocoles pour tenter de fabriquer une couche dense et homogène de streptavidine. Pour ce faire, nous avons utilisé des lamelles de microscope nettoyées au Plasma Cleaner sur lesquelles nous avons déposé plus ou moins directement de la streptavidine marquée à l'Alexa546 (*Invitrogen*). Les différents protocoles que nous avons tentés ainsi que les observations correspondantes sont décrits ci-dessous. Les images de fluorescence obtenues pour les différents protocoles apparaissent sur la figure II.22 avec la même échelle de gris. Les concentrations utilisées sont les mêmes dans tous les cas.

- (a) **Streptavidine-Alexa546** : la première idée consistait à déposer directement la streptavidine sur les lamelles (Figure II.22(a)). Il y a alors peu de streptavidine fluorescente et la couche obtenue n'est pas homogène, on distingue même des zones d'agrégation.
- (b) **Silane - Streptavidine-Alexa546** : nous avons ensuite essayé de déposer la streptavidine sur des lamelles silanisées (Figure II.22(b)). Dans ce cas, la répartition semble uniforme mais il y a très peu de streptavidine.
- (c) **Caséine - Streptavidine-Alexa546** : nous avons alors déposé la streptavidine sur une première couche de caséine biotinilée (Figure II.22(c)). La couche de streptavidine obtenue est cette fois dense et homogène : la caséine se dépose non-spécifiquement sur le verre et permet ensuite l'adsorption de la streptavidine.
- (d) **Silane - Caséine - Streptavidine-Alexa546** : la streptavidine est déposée sur des lamelles silanisées recouvertes de caséine biotinilée (Figure II.22(d)). L'allure de la couche de streptavidine obtenue est la même que pour le cas précédent, à la différence près que nous n'utilisons plus les interactions non-spécifiques caséine-verre. Le protocole pourrait être peu reproductible si on ne comptait que sur les interactions non-spécifiques.



**Figure II.22:** Les différentes tentatives pour obtenir une surface de streptavidine. Les images (c) et (d) ont été prises juste après avoir déplacé la lamelle, on peut ainsi voir la marque du diaphragme qui indique la limite entre la zone précédemment observée (zone sombre - photoblanchie) et la zone actuellement éclairée (zone claire)

L'analyse, certes qualitative, de ces différentes images justifie cependant le choix de notre protocole (situation (d)) et nous confirme que la couche de streptavidine obtenue dans ce cas est dense et homogène.

### II.3.2.5 Estimation du taux d'étirement de l'ADN par un flux de cisaillement

Dans nos expériences, des fragments de T7 biotinilé de taille  $3,1 \mu\text{m}$  (9,2 kb) sont étirés avec le flux maximal de notre tire-seringue : 90 mL/heure.

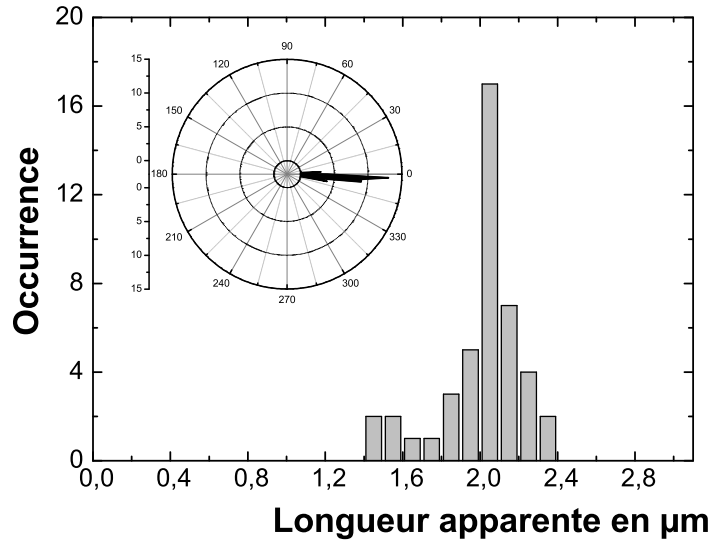
Il est possible d'estimer le taux d'étirement des molécules d'ADN, en utilisant le nombre de Weissenberg  $W_i^4$ , qui dans le cas d'un polymère soumis à un écoulement s'exprime simplement :  $W_i = \dot{\gamma}\tau$  où  $\dot{\gamma}$  est le gradient transverse de vitesse et  $\tau$  le temps de relaxation du polymère seul en solution [63, 82, 64, 83]. L'écoulement entre la lame et la lamelle qui forment la cavité est un Poiseuille : en prenant le vecteur  $\mathbf{x}$  l'axe du gradient de pression, et le vecteur  $\mathbf{z}$  la normale aux plaques (les plaques sont situées en  $z = \pm h/2$ ), la vitesse de l'écoulement est parallèle aux plaques :  $\mathbf{v} = v(z) \mathbf{u}_x$  :

$$v(z) = v_{\max} \left( 1 - \frac{4z^2}{h^2} \right), \quad (\text{II.10})$$

où  $v_{\max}$  est la vitesse maximale de l'écoulement, liée au gradient de pression et au débit volumique par unité de longueur  $Q_v$  :

$$v_{\max} = \frac{h^2}{4\eta} \frac{dp}{dx} = \frac{3Q_v}{4h}.$$

<sup>4</sup>Nombre sans dimension qui représente le rapport entre le temps de relaxation du fluide et un autre temps caractéristique d'un processus d'intérêt, ici le temps de relaxation du polymère.



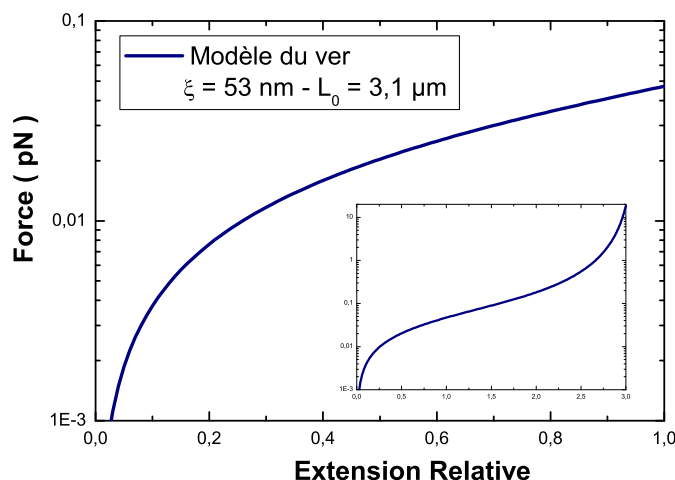
**Figure II.23:** Statistiques faites sur 44 molécules d'ADN étirées en cavité dans les mêmes conditions. La longueur apparente moyenne est de  $2,1 \pm 0,2 \mu\text{m}$ . Insert : Statistiques des angles que font les molécules d'ADN étirées avec la direction du flux appliqué.

Nous en déduisons le gradient de vitesse transverse au niveau de la lamelle, là où est étiré l'ADN.

$$\dot{\gamma}(z) = \frac{6Q_v}{h^3}z \implies \dot{\gamma}(-h/2) \approx 4.10^3 \text{ s}^{-1}.$$

Dans l'article [64], Ladoux et Doyle ont mesuré le temps de relaxation de molécules d'ADN de différentes tailles dont l'une des extrémités est attachée à une surface. En extrapolant leurs résultats expérimentaux, nous trouvons  $\tau = 0,025 \text{ s}$  pour une molécule d'ADN de taille  $3,1 \mu\text{m}$ .

Finalement, le nombre de Weissenberg peut-être estimé à 100. D'après la revue [83], le taux d'étirement correspondant à cette valeur du nombre de Weissenberg est 0,7. En mesurant à la caméra la taille des molécules étirées avec un débit de 90 mL/heure, nous trouvons une taille apparente de  $2,1 \mu\text{m}$ , soit un taux d'étirement moyen de 70 % (Figure II.23). D'après la courbe force-extension calculée avec le modèle du ver de l'ADN, une telle extension est obtenue en appliquant une force de 0,03 pN (Figure II.24).



**Figure II.24:** Extension d'une molécule d'ADN en fonction de la force appliquée. L'élasticité de l'ADN étant décrit ici par le modèle du ver (Worm Like Chain Model, [84]).

## II.4 Visualiser l'interaction ADN - EcoRV

### II.4.1 Dispositif Expérimental

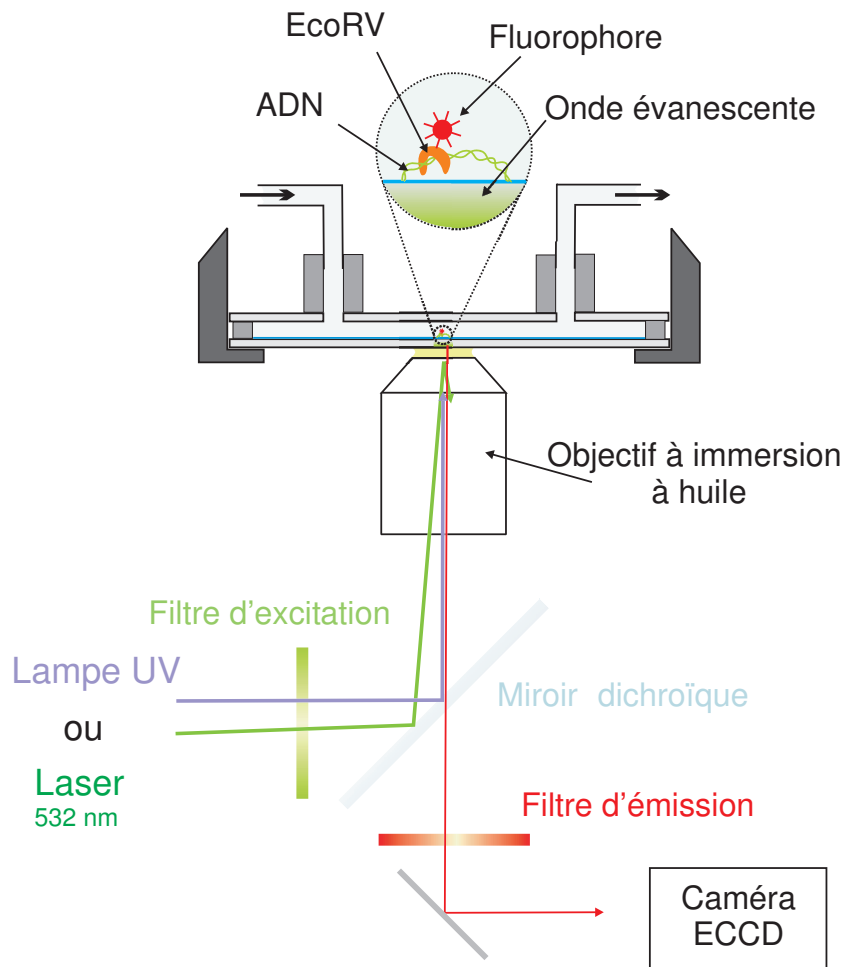
#### II.4.1.1 Montage de TIRFM

Les expériences sont réalisées sur le montage de TIRFM schématisé sur la figure II.25. L'élément principal est un microscope inversé (IX 70, *Olympus*), disposant d'un objectif à immersion à huile de grossissement  $\times 60$  d'ouverture numérique 1,45. En plus du laser, nous disposons d'une lampe U.V. à spectre large, conjuguée avec le plan focal arrière de l'objectif, permettant de travailler en configuration plus classique d'épifluorescence. Les caméras de détection sont conjuguées avec la lame qui est elle-même dans le plan focal de l'objectif. Les jeux de filtres (excitation, émission et miroir dichroïque) sont adaptés en fonction des lampes et des sondes fluorescentes utilisées. Une caméra Ixon DX885 de  $128 \times 128$  pixels de taille  $24 \mu\text{m}$  (*Andor technology*) est adaptée sur une sortie latérale du microscope permettant d'acquérir les signaux avec une grande sensibilité. Une caméra CCD classique, placée au niveau de l'oculaire permet une observation en champ large sur un moniteur. Les cavités à étirement sont placées au dessus de l'objectif du microscope.

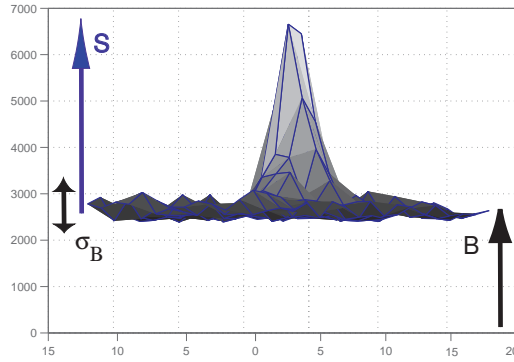
#### II.4.1.2 Une caméra adaptée au suivi de molécule unique

Nous avons utilisé une caméra EMCCD<sup>5</sup> (Ixon DV860), dont le principe de base est celui d'un capteur CCD : les photons incidents créent une paire électrons-trous, les électrons libres s'accumulent alors dans des puits de potentiel agencés en matrice CCD (effet photo-

<sup>5</sup>CCD à multiplications d'électrons



**Figure II.25:** Schéma du dispositif expérimental utilisé pour observer les interactions ADN-EcoRV à l'échelle de la molécule unique et en microscopie de fluorescence par onde évanescente.



**Figure II.26:** Définition du signal à bruit

conducteur). En pratique, le signal de fluorescence  $S$  obtenu est inévitablement entaché par des bruits, d'origines différentes. On distingue généralement 3 principales sources de bruit :

**Le bruit de photons** est une propriété fondamentale de la nature quantique de la lumière. Aussi appelé bruit quantique, il est relié aux fluctuations statistiques du nombre de photons incidents. L'intervalle entre chaque photon incident est régi par une statistique de Poisson, c'est pourquoi le bruit quantique est en  $\sqrt{I}$  où  $I$  est le flux incident de photons. Le bruit de photons est indépendant de la caméra.

**Le bruit d'obscurité**  $B_{\text{obs}}$  lié aux électrons générés par effet thermique au sein de la structure en silicium du CCD. Ce phénomène est indépendant du flux de photons incidents. Il est généralement réduit par refroidissement ; dans notre cas, des modules Peltier permettent de descendre le capteur à  $-70^{\circ}\text{C}$ .

**Le bruit de lecture**  $B_{\text{lect}}$  généré par l'électronique de la caméra qui convertit les charges en valeur numérique. La technologie EMCCD est une technologie d'amplification du signal avant la lecture, ce qui réduit le bruit de lecture, particulièrement intéressant dans le cas de signaux faibles.

Le rapport signal à bruit  $SNR$  correspond à l'amplitude d'un signal comparé à l'incertitude sur la provenance de ce signal comme l'illustre la Figure II.26. En notant  $\eta_Q$  le rendement quantique de la caméra et  $B_f$  le bruit de fond provenant de l'échantillon même, le  $SNR$  est donné par la formule suivante :

$$SNR = \frac{\eta_Q I T}{\sqrt{(I + B_f) \eta_Q T + B_{\text{obs}} T + B_{\text{lect}}^2}} \quad (\text{II.11})$$

Le capteur de la caméra est de type *back-illuminated*, ce qui signifie que la CCD est éclairée par l'arrière afin de minimiser les pertes en traversant moins de silicium. L'efficacité quantique atteint alors plus de 90 %. De plus, la caméra dispose d'un système de *frame transfert* qui consiste à diviser le registre parallèle en deux parties identiques. La

première partie est utilisée pour réaliser l'acquisition alors que la deuxième partie servira de zone de stockage temporaire. Ainsi, un transfert rapide de la totalité des charges de la zone d'exposition vers la zone de stockage permet de libérer la zone d'exposition pour une nouvelle exposition alors que la zone de stockage organise le transfert des charges vers l'amplificateur. L'acquisition a lieu en continu et le temps d'acquisition peut être descendu à 2 ms pour une zone d'observation de  $32 \times 32$  pixels. Toutes ces caractéristiques ont permis la visualisation de fluorophores uniques.

L'ensemble des mises au point réalisées (couplage stœchiométrique des enzymes, préparation des surfaces, étirement spécifique de l'ADN), nous sommes parvenus à observer l'interaction d'une unique enzyme sur une seule molécule d'ADN comme le témoigne le chapitre suivant.

# Chapitre III

## Étude de la diffusion facilitée d'EcoRV

*Afin de caractériser la diffusion facilitée d'EcoRV, nous avons étudié son interaction avec des molécules d'ADN étirées ne comportant aucun site de restriction pour cette enzyme.*

*Dans ce dernier chapitre, nous présentons d'abord le protocole expérimental que nous avons élaboré pour observer l'interaction d'une unique enzyme EcoRV fluorescente avec une seule molécule d'ADN étirée. Ensuite, nous décrirons les techniques d'analyse que nous avons mises au point pour traiter les données expérimentales et caractériser ainsi la diffusion facilitée de l'enzyme EcoRV. Enfin, nous proposons un modèle analytique qui rend compte de nos observations.*

### III.1 Démarche expérimentale

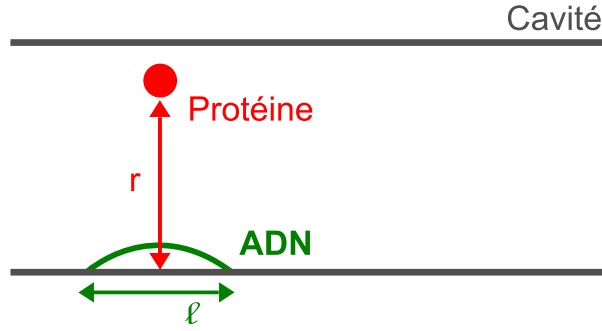
#### III.1.1 Stratégie

##### III.1.1.1 Choix de l'ADN

**ADN non spécifique :** l'ADN du bactériophage T7 est un ADN linéaire de 39,9 kb qui ne comporte aucun site de restriction pour EcoRV. C'est le plasmide support de notre construction d'ADN non-spécifique pour l'enzyme EcoRV : nous en avons récupéré un fragment de 8,2 kb par digestion enzymatique. En parallèle, deux fragments de 500 pb sont synthétisés par PCR en présence de nucléotides-biotine de manière à pouvoir être ligués aux extrémités du fragment de T7 digéré (cf. Annexe A).

**Taille idéale des molécules d'ADN pour l'étirement :** la taille des molécules d'ADN que nous étirons doit être judicieusement choisie : trop courtes, les molécules d'ADN ne seraient pas résolues ; trop longues, l'amplitude des fluctuations thermiques et leur temps de corrélation gêneraient l'étude du suivi d'une protéine interagissant avec l'ADN (cf. Annexe C). La résolution optique  $\delta$  de notre microscope est d'environ 220 nm (cf. section





**Figure III.1:** Estimation de la concentration optimale en enzymes.

II.3.1). Étirée à 70%, une molécule d'ADN doit avoir une longueur de contour minimale  $L_{min}$  pour que ses deux extrémités soient résolues telle que :

$$L_{min} = \frac{1}{0,70} \times \delta \approx 920 \text{ pb.}$$

Nous avons choisi de travailler avec des molécules d'ADN de taille  $L \approx 10 L_{min}$  : pour les détails techniques de la construction de molécules d'ADN de cette taille aux extrémités biotinilées à partir de l'ADN de T7, le lecteur peut consulter l'Annexe A.

### III.1.1.2 Concentration d'enzyme adaptée aux expériences en molécule unique

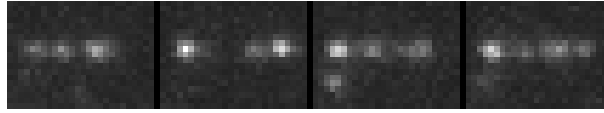
**Estimation :** nous pouvons, par un calcul d'ordre de grandeur, estimer la concentration optimale d'enzymes pour observer un événement d'interaction ADN-protéines toutes les 30 s avec notre dispositif expérimental (cf. Figure III.1). En s'inspirant de [16], nous pouvons exprimer la probabilité  $\mathcal{P}$  par unité de temps pour une protéine de trouver par diffusion une molécule d'ADN fixe de longueur  $\ell$  située à la distance  $r \gg \ell$  :  $\mathcal{P} = \ell/r$  (capture par diffusion). Le temps caractéristique nécessaire à la protéine pour se déplacer sur cette distance  $r$  est  $t_{diff} = \frac{r^2}{D_3}$  où  $D_3$  est le coefficient de diffusion tridimensionnel de la protéine. Le temps total d'association  $t_{ass}$  vaut alors :

$$t_{ass} = t_{diff} \times \frac{1}{\mathcal{P}} = \frac{r^3}{D_3 \ell}. \quad (\text{III.1})$$

En posant  $c$  la concentration moyenne d'enzymes dans le canal, il vient  $r^3 = \frac{1}{c}$  soit :

$$c = \frac{1}{D_3 \ell t_{ass}}. \quad (\text{III.2})$$

Dans nos expériences, les molécules d'ADN, de longueur 9200 pb, sont étirées de 70% en moyenne, ce qui donne  $\ell = 2,2 \mu\text{m}$ . En prenant une taille caractéristique de 5 nm pour les enzymes marquées, on obtient  $D_3 = 4.10^{-11} \text{ m}^2.\text{s}^{-1}$ . Nous souhaitons typiquement avoir une interaction toutes les 30 s, ce qui donne  $c \approx 10^{12} \text{ enzymes.L}^{-1}$ , c'est-à-dire des concentrations de l'ordre du pM. Ce calcul suppose qu'à chaque collision avec l'ADN,



**Figure III.2:** Séquences extraites d'un film d'interaction ADN - EcoRV-Cy3B. La concentration en enzyme ( $c = 0,7$  nM) est trop élevée pour une étude à l'échelle de la molécule individuelle puisque les protéines recouvrent entièrement l'ADN.

la protéine s'associe, ce qui n'est pas le cas : cette concentration est donc une limite inférieure. Expérimentalement, nous avons effectué différentes dilutions pour déterminer la concentration optimale et nous avons trouvé une concentration plus grande. En effet, pour une concentration en enzymes marquées de 70 pM, les événements d'interaction sont raisonnablement espacés dans le temps : à cette concentration, il est rare de voir deux protéines interagissant avec l'ADN en même temps. En revanche, lorsque la concentration en enzyme est plus élevée, l'ADN est entièrement recouvert d'enzymes comme l'illustre la Figure III.2. Finalement, les expériences ont été réalisées à une concentration en enzyme marquées de 70 pM.

**Comparaison :** dans la littérature [85, 86], nous trouvons les valeurs de la vitesse de réaction d'association d'EcoRV à l'ADN. Pour des conditions similaires aux nôtres, ces vitesses sont de l'ordre de la dizaine de  $\text{pM}\cdot\text{min}^{-1}$ . Nous pouvons retrouver ce résultat à condition de connaître la concentration en ADN puisque :  $v = k_{\text{on}} [\text{ADN}] [\text{EcoRV}]$  avec

$$k_{\text{on}} = \frac{1}{t_{\text{ass}} [\text{EcoRV}]}$$

Pour estimer la concentration d'ADN en cavité, il suffit de connaître la densité de molécules étirées. Nous avons typiquement une molécule d'ADN étirée par zone de  $128 \times 128$  pixels. La taille d'un pixel est de 126 nm, la hauteur de la cavité est de  $140 \mu\text{m}$ , nous travaillons à une concentration :  $[\text{ADN}] \sim 4,5 \cdot 10^{-15}$  M. En prenant  $c = 10^{-12}$  M, nous déduisons une vitesse de réaction  $v = 10^{-13} \text{ M}\cdot\text{s}^{-1} = 6 \text{ pM}\cdot\text{min}^{-1}$ , en accord avec la littérature.

## III.1.2 Protocole

### III.1.2.1 Réduire les interactions non-spécifiques

Après avoir étiré l'ADN, la surface doit être passivée en y déposant de la caséine à  $0,2 \text{ mg/mL}$  pendant un temps d'incubation de 10 min. Cette couche de caséine permet de réduire considérablement les interactions non-spécifiques enzymes-surface [68], qui, dans notre configuration expérimentale de TIRF contribueraient largement au bruit de fond. En parallèle, il faut repérer une dizaine de molécules d'ADN correctement positionnées : bien orientées dans la direction du flux et suffisamment étirées (pour réduire au maximum l'effet des fluctuations de l'ADN). Les positions de ces molécules d'ADN sont enregistrées

par l'intermédiaire d'un logiciel qui pilote la platine motorisée du microscope : il est donc possible d'observer séquentiellement plusieurs molécules d'ADN étirées.

### III.1.2.2 Choisir un tampon d'interaction adapté

L'étape de passivation réalisée, il faut changer le milieu pour travailler dans un tampon adapté à l'enzyme EcoRV. Inspirés de la littérature, nous avons adopté la composition suivante :

- Tampon biologique (Pipes, Hepes, Tris, ...) 20 mM
- NaCl 10 mM
- MgCl<sub>2</sub> 10 mM
- DTT 1 mM
- Caséine 0,2 mg/mL

Les ions magnésium Mg<sup>2+</sup> sont cofacteurs de l'enzyme EcoRV, la caséine limite l'agrégation des protéines, le DTT est un agent réducteur, qui sert à stabiliser les protéines (sous l'action de l'oxygène de l'air et en présence de cations métalliques, les résidus de cystéine s'oxydent).

**Salinité :** lorsque la concentration en sel augmente, la portée des interactions électrostatiques ADN-protéine diminue, réduisant ainsi l'affinité ADN-protéine. Cela a d'ailleurs été vérifié expérimentalement en montrant que la durée d'interaction des protéines avec l'ADN diminue lorsque la salinité augmente [68, 87]. Pour observer des événements d'interaction ADN-protéine suffisamment longs, nous avons choisi de travailler à faible concentration en NaCl (10 mM).

**Force ionique :** la force ionique  $F_i$  d'une telle solution s'exprime en mol.L<sup>-1</sup> et se calcule de la manière suivante :  $F_i = \frac{1}{2} \sum_i c_i z_i^2$ . où  $c_i$  est la concentration de l'ion  $i$ , et  $z_i$  sa charge. Ainsi :

$$F_i(\text{NaCl}, c) = \frac{1}{2} (c(\text{Na}^+) + c(\text{Cl}^-)) = c. \quad (\text{III.3})$$

$$F_i(\text{MgCl}_2, c) = \frac{1}{2} (4c(\text{Mg}^{2+}) + c(\text{Cl}^-)) = 3c. \quad (\text{III.4})$$

Le Pipes est un diacide symétrique de pK<sub>A</sub> = 6,8 ; le Hepes est un monoacide de pK<sub>A</sub> = 7,5. Nous les avons utilisés à des pH tels que pH=pK<sub>A</sub> ou relativement proche ( $|\text{pH} - \text{pK}_A| \leq 1$ ). Le pH a été ajusté par ajout de soude : il faut également tenir compte de ces contre-ions ajoutés. En faisant l'hypothèse que pour ces pHs, [acide] = [base], nous pouvons calculer la force ionique des tampons utilisés :

$$F_i(\text{Pipes}, c, \text{pH} = 6, 8) = F_i(\text{A}^{2-}) + F_i(\text{Na}^+) = \frac{1}{2} \left( 4\frac{c}{2} + c \right) = \frac{3c}{2}. \quad (\text{III.5})$$

$$F_i(\text{Hepes}, c, \text{pH} = 7, 5) = F_i(\text{A}^-) + F_i(\text{Na}^+) = \frac{1}{2} \left( \frac{c}{2} + \frac{c}{2} \right) = \frac{c}{2}. \quad (\text{III.6})$$

La force ionique totale des tampons utilisés est par conséquent :

$$F_i (\text{Tampon d'interaction Pipes}) \approx 70 \text{ mM.} \quad (\text{III.7})$$

$$F_i (\text{Tampon d'interaction Hepes}) \approx 50 \text{ mM.} \quad (\text{III.8})$$

La force ionique de la solution joue un rôle important dans les propriétés mécaniques du polymère chargé ADN puisque l'écrantage des charges affecte sa longueur de persistance [88]. Chaque couple ADN-protéine doit être vu comme un système à part entière, dans lequel les interactions spécifiques ou non dépendent beaucoup des conditions ioniques du milieu (force ionique, présence d'ions divalents). La force ionique physiologique est de 150 mM. Travailler à force ionique pas trop élevée permet d'augmenter le temps d'interaction ADN-protéine, même si celui-ci peut aussi dépendre de la nature du tampon (anionique, cationique, amphotère) [85].

### III.1.3 Acquisition

#### III.1.3.1 Démarche :

Une fois les positions des molécules d'ADN repérées, nous rinçons abondamment la cavité avec le tampon d'interaction. Ce rinçage permet accessoirement de se débarrasser du *SyBR Gold* car les ions magnésium se fixent sur l'ADN, délogeant ainsi le marqueur fluorescent. Il est important de se débarrasser du *SyBR Gold* pour s'affranchir d'éventuels artéfacts liés aux fluorophores lors des événements d'interaction. La cavité est ensuite déplacée pour placer une molécule d'ADN précédemment repérée dans le champ de l'objectif. Les protéines fluorescentes sont alors injectées à la concentration de 70 pM.

Nous utilisons le montage de TIRF décrit précédemment. La fluorescence est collectée par l'intermédiaire d'un objectif  $\times 60$  à immersion à huile et d'ouverture numérique  $NA = 1,45$  (*Olympus*). Les filtres que nous avons utilisés pour la détection du *SyBR Gold* sont les suivants : filtre d'excitation 475AF40, filtre d'émission 535AF45 et dichroïque 505DRLP (*Omega Filters*). Les fluorophores (Cy3B, Sav-Cy3, QD605) couplés aux enzymes sont excités par un laser à 532 nm. Pour les détecter, nous avons utilisé : filtre d'émission 595AF60 et miroir dichroïque 560DRLP. Nous réalisons des séries de films pour enregistrer le mouvement des protéines fluorescentes qui interagissent avec les molécules d'ADN. Il s'agit typiquement de films de 6000 images en continu avec un temps d'exposition  $t_{\text{acq}}$  de 20 ms (sauf mention contraire).

#### III.1.3.2 Type de données :

Tous les films ont été pris avec le logiciel Ixon (*Andor Technology*). Pour chaque zone d'observation de taille  $4 \mu\text{m} \times 4 \mu\text{m}$ , nous disposons de plusieurs séquences d'image. Notre

Paramètres	Valeur
Largeur d'un pixel	$h \approx 126 \text{ nm}$
Ouverture numérique	$NA = 1,45$
Temps d'acquisition	$t_{\text{acq}} = 20 \text{ ms}$

**Tableau III.1:** Paramètres du système d'acquisition (TIRFM et caméra EMCCD).

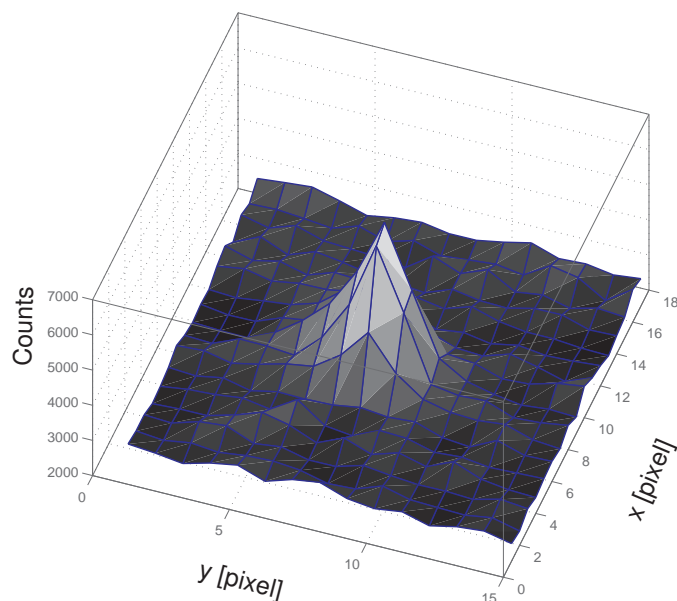
dispositif permet d'obtenir des séquences d'images 2D, monochromes et quantifiées sur 14 bits, les paramètres étant répertoriés dans le tableau III.1. En superposant l'ensemble des images d'un film donné, nous localisons la molécule d'ADN correspondante puisque les signaux de fluorescence de toutes les protéines ayant interagi avec l'ADN se superposent, ainsi que la fluorescence résiduelle des dernières molécules de *SyBR Gold* dont le spectre d'émission déborde légèrement dans le rouge.

### III.1.3.3 Image idéale des sondes fluorescentes à travers le microscope :

Les fluorophores que nous avons utilisés peuvent être considérés comme des sources ponctuelles dans la mesure où leur taille (de l'ordre du nanomètre) sont largement inférieures à la longueur d'onde de la lumière qu'ils émettent. Le microscope pouvant être considéré comme un système linéaire, l'image de ces sondes à travers le microscope correspond à la réponse impulsionnelle du système optique. D'après la théorie de la diffraction, le profil normalisé à 1 de cette réponse impulsionnelle est donné par la fonction de Bessel de première espèce, d'ordre 1 et dont le premier zéro définit le rayon d'Airy  $R_A = \frac{0,61\lambda}{NA}$ . Ce profil théorique peut-être approximé par un profil gaussien, d'écart-type  $\sigma_G$  plus facile à manipuler d'un point de vue numérique et souvent noté *PSF* (acronyme de Point Spread Function) :

$$PSF^{\text{norm}}(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma_G^2}\right). \quad (\text{III.9})$$

L'image d'une sonde fluorescente est donc une gaussienne dont l'amplitude dépend du nombre de photons émis et dont l'étalement spatial ne dépend, à travers le rayon d'Airy, que de paramètres physiques connus tels que la longueur d'onde d'émission  $\lambda$  et l'ouverture numérique  $NA$ . Deux méthodes peuvent être envisagées pour estimer l'écart-type  $\sigma_G$  : soit appliquer une régression gaussienne sur la réponse impulsionnelle théorique du système optique (ce qui donne  $\sigma_G = R_A/3$  [89]), soit appliquer une régression gaussienne sur des données expérimentales. Nous avons choisi la seconde option car elle permet de prendre en compte les aberrations géométriques et les effets de pixelisation propres au montage optique utilisé. Pour le montage optique dont les paramètres sont répertoriés dans le tableau III.1, l'écart-type mesuré expérimentalement pour  $\lambda = 600 \text{ nm}$  est :  $\sigma_G \approx 200 \text{ nm} \approx 1,6 \text{ pixel}$ . Ainsi, chaque spot de fluorescence a la forme d'une gaussienne dont l'essentiel



**Figure III.3:** Image de fluorescence 3D d'une molécule de Cy3B adsorbée à la surface d'une lamelle; taille d'un pixel = 126 nm, temps d'exposition = 20 ms. Ce spot de fluorescence est une gaussienne 2D s'étalant sur un carré d'environ 5 pixels de côté.

de l'énergie est concentré sur un carré d'environ 5 pixels de côté (cf. Figure III.3).

Dans la suite de ce manuscrit, nous appelons « spot » toute image ponctuelle de fluorescence. Un programme de *tracking* a été écrit en MatLab (*Mathworks*) pour reconstruire les trajectoires des protéines sur l'ADN. Pour plus de détails sur ce programme, j'invite le lecteur à lire l'Annexe B.

## III.2 EcoRV glisse le long de l'ADN

### III.2.1 La MSD : un outil d'analyse

#### III.2.1.1 Mouvement Brownien et coefficient de diffusion

Dans le cas d'une diffusion pure le long de l'ADN, le mouvement d'une protéine est modélisé par un mouvement Brownien unidimensionnel. Plusieurs formalismes mathématiques existent pour décrire le mouvement Brownien : la loi de la diffusion d'Einstein (théorème de fluctuation/dissipation : équipartition de l'énergie), le modèle de Langevin (calcul stochastique) et le modèle de marche au hasard. En adoptant ce dernier formalisme, la position  $x(t + \tau)$  d'une particule à l'instant  $t + \tau$  est reliée à sa position  $x(t)$  à l'instant  $t$  comme suit :

$$x(t + \tau) = x(t) + \xi \sqrt{2D_1\tau}. \quad (\text{III.10})$$

où  $D_1$  est le coefficient de diffusion et  $\xi$  une variable aléatoire gaussienne de moyenne nulle  $\langle \xi \rangle = 0$  et de variance  $\langle \xi^2 \rangle = 1$ . Cette formulation du mouvement Brownien est équivalente au modèle de Langevin dans la limite où les effets inertiels sont négligeables devant les effets liés à la viscosité<sup>1</sup>. C'est aussi la solution exacte de l'équation de Smoluchowski<sup>2</sup> en dimension 1 et en l'absence de potentiel extérieur.

### III.2.1.2 Le déplacement quadratique moyen : MSD

**Définition et calcul :** le déplacement quadratique moyen ou MSD pour *Mean Square Displacement* est une fonction bien adaptée à l'analyse de trajectoires. La MSD  $\rho(t)$  à l'instant  $t$  est initialement définie pour une population infinie et homogène de particules se déplaçant sous l'effet des mêmes lois physiques.  $\rho(t)$  est alors obtenu sur la moyenne sur l'ensemble des trajectoires des différentes particules. En vertu du principe d'ergodicité, cette moyenne d'ensemble peut-être calculée à partir d'une moyenne temporelle effectuée sur une seule trajectoire de durée infinie :

$$\rho(t) = \langle (x(t) - x(0))^2 \rangle = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-T/2}^{T/2} (x(t + \tau) - x(\tau))^2 d\tau. \quad (\text{III.11})$$

où  $\langle \rangle$  désigne une moyenne d'ensemble sur toutes les trajectoires  $\{x^{(m)}(t)\}_{m \in \mathbb{N}}$ .

En pratique, nous disposons de trajectoires de taille finie, contenant  $N$  éléments de coordonnées temporellement séparés de  $\Delta t = t_{\text{acc}}$ , l'intervalle de temps entre deux images consécutives. Dans ce cas, une estimation de la MSD peut être obtenue par le calcul de la fonction discrète suivante [90, 91] :

$$\rho(n\Delta t) = \frac{1}{N-n} \sum_{i=1}^{N-n} (x_{i+n} - x_i)^2. \quad (\text{III.12})$$

**La MSD : signature de la nature des mouvements** ci-dessous, nous calculons les MSDs des différents mouvements de spots que nous sommes susceptibles de rencontrer dans notre configuration expérimentale :

► **Spot diffusant librement 1D** : il s'agit d'une diffusion Brownienne, également appelée diffusion pure ou libre. La MSD se calcule aisément à partir de l'équation III.10 :

$$\rho_x(t) = \langle 2D_1 t \xi^2 \rangle = 2D_1 t. \quad (\text{III.13})$$

---

<sup>1</sup>Ce qui est vrai pour des temps  $\tau$  suffisamment longs :  $\tau \gg \frac{m}{\gamma}$  où  $\gamma$  est le coefficient de friction et  $m$  la masse de la particule considérée.

<sup>2</sup>Marian von Smoluchowski introduit la notion de marche aléatoire et fait le lien avec la théorie des jeux en 1906.

- **Spot fixe sur l'ADN** : le spot est animé du mouvement de fluctuation de l'ADN. À ces mouvements de fluctuations correspondent des fonctions de corrélations de position longitudinales et transverses de la forme  $\langle u(t)u(t + \tau) \rangle = U^2 e^{-\frac{t}{\tau_u}}$  où  $u$  désigne la direction ( $x$  pour la direction de l'ADN,  $y$  pour la direction transverse),  $U$  est l'amplitude maximale des fluctuations longitudinales ( $X$ ) ou transverses ( $Y$ ),  $\tau_u$  le temps caractéristique de ces fluctuations (cf. Annexe C). Dans nos conditions expérimentales,  $t \gg \tau_u$ , la MSD est constante :

$$\rho_u(t) = 2U^2 \left(1 - e^{-\frac{t}{\tau_u}}\right) \xrightarrow{t \gg \tau_u} 2U^2. \quad (\text{III.14})$$

- **Spot diffusant le long de l'ADN** : en faisant l'hypothèse légitime que le mouvement de diffusion et le mouvement de fluctuation ne sont pas corrélés. En posant  $x$  la direction de l'ADN, il vient :

$$\rho_x(t) = 2X^2 + 2D_1 t. \quad (\text{III.15})$$

Dans ce cas, la MSD est une fonction affine et sa pente donne accès au coefficient de diffusion.

**Les sources d'erreurs de la MSD expérimentale** lorsqu'on considère une trajectoire réelle, le calcul de la MSD donné à l'équation III.12 est entaché par l'incertitude liée à la mesure expérimentale. Deux principales sources d'erreur affectent cette estimation :

- **L'erreur statistique** liée au nombre fini de points dans la trajectoire étudiée. Cette erreur dépend du nombre  $N$  de points dans la trajectoire mais aussi du point  $t = n\Delta t$  auquel nous nous plaçons pour calculer la MSD. Cette erreur est d'autant plus petite que le nombre  $N$  de points est grand et que le point  $n$  auquel on se place est petit puisque le nombre de points intervenant dans le calcul de la MSD décroît avec  $n$ . D'après [90], nous pouvons considérer que seul le premier quart des points de la MSD est statistiquement significatif pour une trajectoire individuelle : les derniers points de la MSD n'ont alors aucune signification physique.

★ Dans le cas d'une seule trajectoire de taille  $N$ , purement Brownienne 1D, il est possible de calculer l'erreur statistique en chaque point  $n$  tels que  $n \leq N/2$  [91] :

$$\Delta\rho_1(n\Delta t) = 2D_{\text{exp}} n\Delta t \frac{4n^2(N-n) + 2N - n - n^3}{6n(N-n)^2}. \quad (\text{III.16})$$

★ À cette erreur s'ajoute celle liée aux fluctuations de l'ADN. J'invite le lecteur à lire la thèse d'A. Crut pour les détails des calculs qui l'ont conduit à exprimer la variance de la MSD pour un mouvement de fluctuations dans un modèle billes-ressorts [68] :

$$\Delta\rho_{2,u}(n\Delta t) = \frac{8\zeta_u^4}{N-n}. \quad (\text{III.17})$$



où  $\zeta_u$  est une constante relative à la raideur selon la direction  $u$ . Au final, en supposant que le mouvement diffusif et les fluctuations sont indépendants, A. Crut montre que l'erreur sur la mesure de la MSD est [68] :

$$\Delta\rho_u = \Delta\rho_1 + \Delta\rho_{2,u} + 4\frac{\langle\rho_1\rangle\langle\rho_{2,u}\rangle}{N-n}. \quad (\text{III.18})$$

Expérimentalement,  $D_1$  est mesuré en calculant la pente du déplacement quadratique moyen. Par conséquent,  $D_1$  est déterminé à cette incertitude près.

- **L'erreur systématique liée à la précision de pointé** : la position d'une particule à chaque instant  $t$  est déterminée à une erreur près : la précision de pointé du dispositif optique. Si  $x_r$  est la position réelle de la particule et  $x_m$  la position mesurée, nous avons  $x_m = x_r + \epsilon$  avec  $\epsilon$  une variable aléatoire suivant une loi normale de moyenne nulle et d'écart-type  $\langle\epsilon^2\rangle = \sigma^2$  où  $\sigma$  est la précision de pointé.

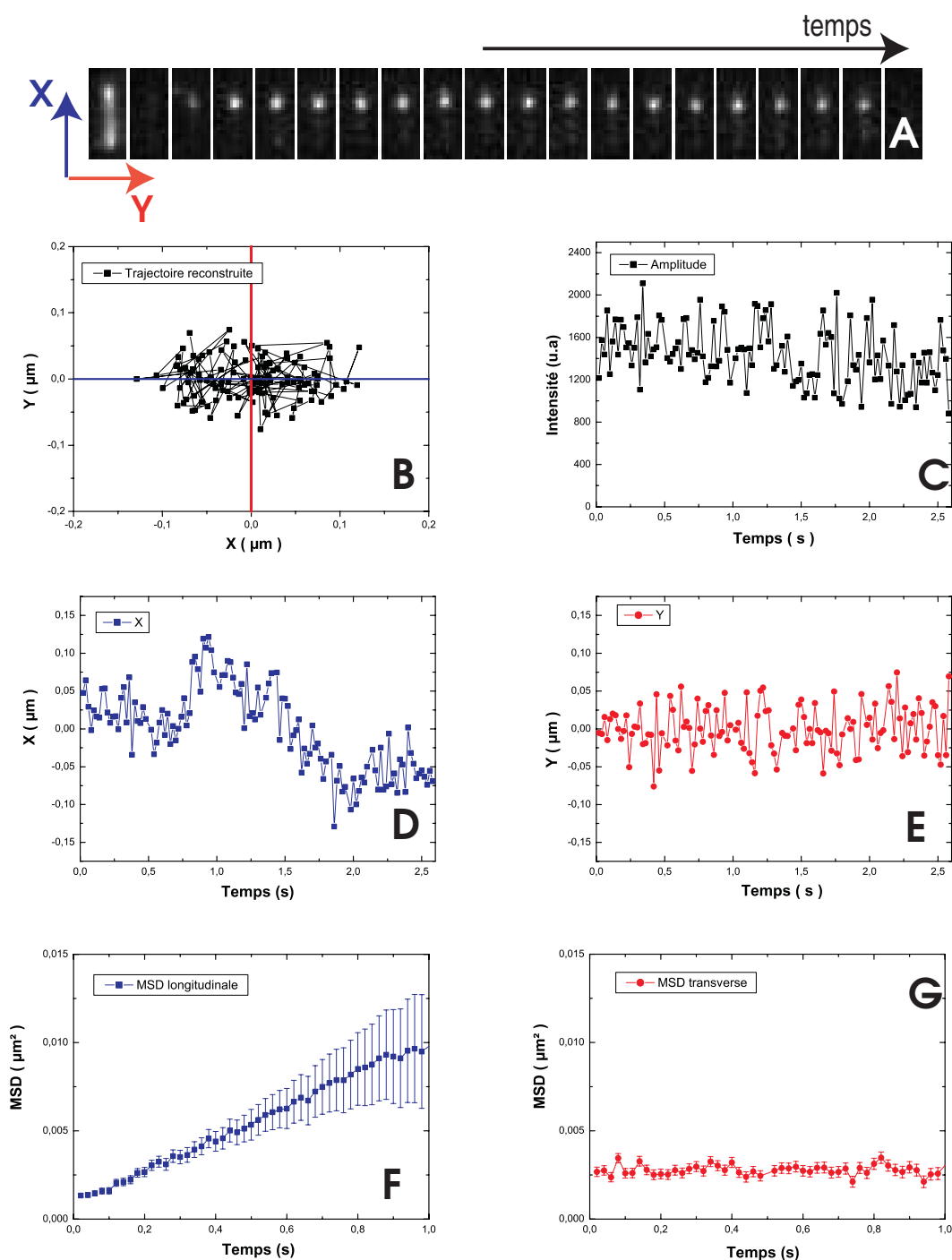
La précision de pointé ajoute donc une composante continue dans la MSD. Elle intervient également dans la détermination du  $D_1$  nécessaire aux calculs des barres d'erreur (Eq.III.16).

**Expression du coefficient de diffusion  $D_1$**  : aux erreurs expérimentales près, le coefficient de diffusion caractéristique d'un mouvement diffusif à 1D correspond donc à la moitié de la pente de la MSD. Nous avons ainsi déterminé expérimentalement des coefficients par ajustements linéaires des MSDs. Sauf indication contraire, ces ajustements ont été faits sur 5 points, à partir du deuxième (c'est-à-dire  $n = 1$ ). Toutefois, le coefficient de diffusion  $D_{\text{exp}}$  ainsi mesuré est un coefficient de diffusion apparent qui ne tient pas compte du taux d'étirement  $\tau$  de l'ADN  $< 100\%$ . Nous devons en tenir compte dans le calcul du coefficient de diffusion  $D_1$  :

$$D_1 = \frac{D_{\text{exp}}}{\tau^2}. \quad (\text{III.19})$$

$\tau$  est déterminé en comparant la taille apparente de chaque molécule d'ADN à la taille de contour de l'ADN  $L = 9200 \text{ pb} = 3,1 \mu\text{m}$ . Dans nos expériences,  $\tau$  vérifie :  $0,6 \leq \tau \leq 0,75$ . Par la suite, toutes les distances ou coefficients de diffusion annoncés sont systématiquement corrigés du taux d'étirement.

Le programme de *tracking* que nous avons écrit inclut le calcul des MSDs longitudinales (selon l'axe de la molécule d'ADN) et transverses (perpendiculaires à l'ADN) des différentes trajectoires. L'analyse des MSDS au cours du temps nous renseigne sur la nature des trajectoires : une MSD longitudinale linéaire témoigne d'un mouvement diffusif, une MSD transverse constante est signe d'un mouvement de fluctuations.



**Figure III.4: Interaction d'une enzyme EcoRV-Cy3B avec l'ADN. A-gauche :** image de l'ADN marqué au *SyBR*. **A-autres images :** extrait d'un film sur la même zone montrant l'enzyme marquée à Cy3B interagissant avec l'ADN pendant 2,6 s. **B :** Trajectoire de la protéine reconstruite. Axe bleu : axe longitudinal (X), axe rouge : axe transverse (Y). **C :** Amplitude du spot au cours du temps. **D et E :** Composantes longitudinale (fluctuation + mouvement diffusif) et transverse (fluctuation) de la trajectoire. **F :** MSDs longitudinale (variation linéaire  $\leftrightarrow$  mouvement diffusif : on mesure un coefficient de diffusion  $D_1 \approx 0,9 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ ). **G :** MSD transverse (constante  $\leftrightarrow$  mouvement confiné).

### III.2.2 La diffusion 1D d'EcoRV-Cy3B

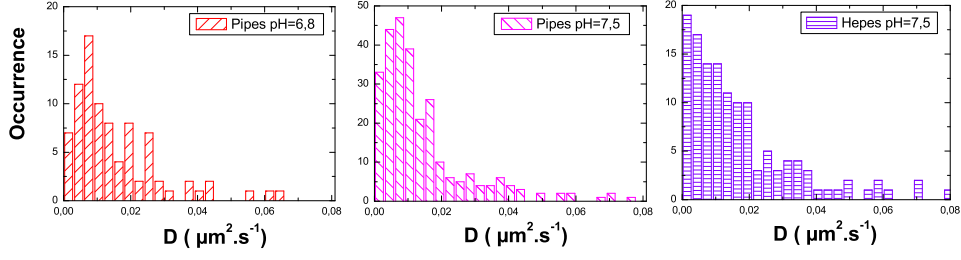
Dans l'hypothèse où la diffusion de l'enzyme le long de l'ADN serait en partie régie par l'hydrodynamique, son coefficient de diffusion 1D serait inversement proportionnel à un coefficient de friction hydrodynamique, lui-même dépendant de la taille de l'enzyme. Dans une certaine mesure, il est donc raisonnable de penser que la taille du fluorophore utilisé pour marquer les enzymes influe sur le coefficient de diffusion de l'enzyme libre : [4]. Pour caractériser précisément la diffusion facilitée d'EcoRV, nous avons choisi de débiter les expériences d'interaction ADN-EcoRV en utilisant les enzymes couplées au Cy3B, puisque le Cy3B est le fluorophore de plus petite taille ( $MW \# 1 \text{ kDa} \sim 1 \text{ nm}$ ) dont nous disposons. Les expériences ont été réalisées dans différents tampons d'interaction : Pipes à  $\text{pH} = 6,8$  et Hepes à  $\text{pH} = 6,8$  et  $7,5$ .

**Coefficient de diffusion :** La figure III.4 est un exemple d'interaction EcoRV-Cy3B - ADN. Dans cet exemple, la protéine reste sur l'ADN pendant 2,6 s. Le programme de *tracking* a tout d'abord permis de reconstruire la trajectoire de la protéine durant cette phase d'interaction (Figure III.4 B, D et E). Les MSDs sont ensuite calculées (Figure III.4 F et G) : la MSD longitudinale peut être ajustée par une droite, alors que la MSD transverse reste constante au cours du temps, preuves que l'enzyme diffuse à 1D le long de l'ADN avec, dans cet exemple, un coefficient  $D_1 = 0,9 \pm 0,05 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ .

Les événements d'interaction pour lesquels la trajectoire comporte plus de 30 points peuvent ainsi être traités individuellement puisque les MSDs peuvent être calculées de manière significatives sur les 7 premiers points de ces trajectoires [90]. Dans plus de la moitié des interactions, le mouvement de la protéine est une diffusion longitudinale et le coefficient de diffusion  $D_1$  est déterminé en ajustant les 5 premiers points des MSDs longitudinales par une droite. Dans les autres cas, les MSDs longitudinale et transverse sont constantes : le spot « tracké » est fixe sur l'ADN. Cette absence de diffusion peut avoir plusieurs origines : enzyme collée à la surface, enzyme non active, agrégats d'enzymes.

Les histogrammes des coefficients de diffusion  $D_1$  déterminés dans les différentes conditions expérimentales apparaissent sur la figure du tableau III.2 qui donne les valeurs moyennes  $\bar{D}$  ainsi que les écart-type  $\sigma_D$  de ces distributions : ces distributions sont similaires, tous les histogrammes présentent un maximum autour de  $D_1 \sim 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ . Cette valeur est comparable aux coefficients de diffusion récemment reportés dans le cadre d'études menées à l'échelle de la molécule individuelle et portant sur la diffusion de protéine [87, 92, 93, 94].

$D_1$  est environ 3 ordres de grandeurs plus petit que le coefficient tridimensionnel déterminé par FCS au chapitre précédent  $D_3 \simeq 50 \mu\text{m}^2 \cdot \text{s}^{-1}$ . Cette réduction drastique



	Pipes pH=6,8	Pipes pH=7,5	Hepes pH=7,5
$N$	104	275	134
$\bar{D}$ ( $10^{-2}\mu\text{m}^2.\text{s}^{-1}$ )	1,6	1,5	1,8
$\sigma_D$ ( $10^{-2}\mu\text{m}^2.\text{s}^{-1}$ )	1,3	2	2,2

**Tableau III.2: Haut :** Distribution des coefficients de diffusion 1D d'EcoRV-Cy3B dans les différents tampons d'interaction. Ces coefficients sont issus de l'extrapolation linéaire des MSDs longitudinales des trajectoires comportant plus de 30 points. **Bas :** Moyenne et écart-type de ces coefficients de diffusion 1D.  $N$  est le nombre de trajectoire supérieure à 30 points pour lesquelles les MSDs individuelles ont été calculées.

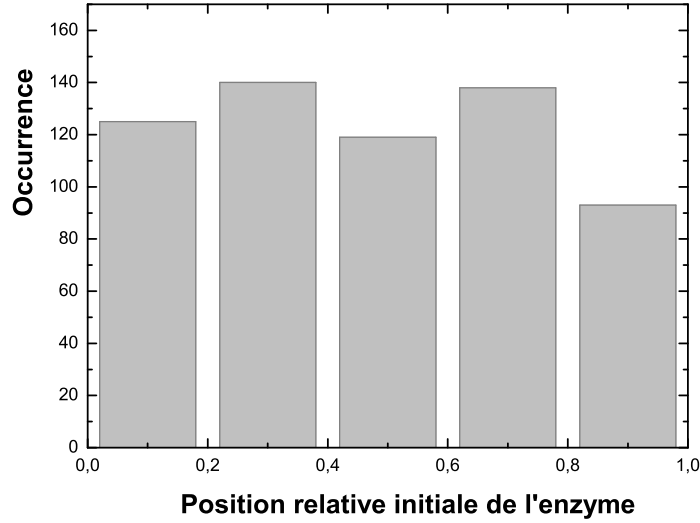
peut être attribuée au mouvement hydrodynamique de l'enzyme glissant sur l'ADN [4] et/ou à des modulations du potentiel d'interaction ADN-protéine (lorsque le sliding est considéré comme une diffusion dans un profil énergétique dépendant de la séquence [25]).

**Positions initiales** Nous avons également extrait les positions initiales des trajectoires des enzymes. Les interactions qui ont eu lieu aux extrémités des molécules d'ADN, là où la molécule d'ADN est accrochée à la surface, ont systématiquement été enlevées. L'histogramme des positions initiales relatives au segment ADN apparaît sur la Figure III.5 : ces positions sont également réparties le long de l'ADN. Ce résultat montre que la configuration des molécules d'ADN dans notre dispositif expérimental ne favorise pas certaines approches de l'enzyme.

### III.2.3 Pooling des données

#### III.2.3.1 Présentation

**Intérêt :** tous les événements d'interaction ne sont pas suffisamment longs pour pouvoir en calculer des MSDs significatives. En particulier, il est difficile d'apprécier une MSD pour les trajectoires qui présentent moins de 20 points. Pour prendre en compte ces événements courts, nous avons choisi de rassembler les données. C'est ce que nous avons choisi d'appeler le « pooling » des trajectoires. Nous garderons la terminologie anglaise



**Figure III.5: Histogramme des positions d'association d'EcoRV sur l'ADN étiré.** Toutes conditions confondues. Toutes les positions initiales sont équivalentes : la configuration « ADN étiré » est bien adaptée aux études des interactions ADN-protéine.

et ses déclinaisons françaises (pooler, poolé) dans la suite de ce manuscrit.

**Principe :** pooler les données d'un ensemble de trajectoires consiste à regrouper non pas les trajectoires (ensemble des coordonnées) mais les distances des trajectoires. Plus précisément, pour  $m = 1..M$  trajectoires de durée  $\{N_m\}_{m=1..M}$ , nous extrayons les distances :  $\left\{ (x_{i+n}^{(m)}(t) - x_i^{(m)}(t)) \right\}_{n=1..N_m; m=1..M}$ . Les distances d'un même intervalle temporel  $n \times dt$  ( $n$  fixé) sont concaténées : nous obtenons des vecteurs de distances  $\mathbf{dist}_n$  avec  $n \in \llbracket 1, \max(N_m) \rrbracket$ . La MSD poolée au point  $n \times \Delta t$  est alors calculée en prenant le carré de la moyenne du vecteur  $\mathbf{dist}_n$ .

Pour  $M$  trajectoires de durée  $\{N_m\}$ , exprimons la MSD poolée à un instant  $t = n\Delta t$  tel que  $\forall m, n \leq N_m$ . L'ensemble des distances entre  $n$  points, toutes trajectoires confondues, a une taille  $\mathcal{N}$  :

$$\mathcal{N} = \sum_{m=1}^M (N_m - n) = \sum_{m=1}^M N_m - nM$$

La MSD poolée s'exprime alors :

$$\rho^{\text{Pool}}(n\Delta t) = \frac{1}{\sum_{m=1}^M N_m - nM} \sum_{m=1}^M \sum_{i=1}^{N_m-n} \left( x_{i+n}^{(m)} - x_i^{(m)} \right)^2. \quad (\text{III.20})$$

Pour la raison invoquée en III.2.1.2, les distances considérées dans l'expression III.20 sont les distances réelles, c'est-à-dire les distances apparentes corrigées du taux d'étirement de

Jeu de paramètres	M=100 $T_{\text{moy}} = 1,5 \text{ s}$ $\sigma = 30 \text{ nm}$ K=100	M=100 $T_{\text{moy}} = 1,5 \text{ s}$ $\sigma = 50 \text{ nm}$ K=100	M=100 $T_{\text{moy}} = 15 \text{ s}$ $\sigma = 30 \text{ nm}$ K=100	M=1000 $T_{\text{moy}} = 1,5 \text{ s}$ $\sigma = 30 \text{ nm}$ K=100	M=100 $T_{\text{moy}} = 1,5 \text{ s}$ $\sigma = 30 \text{ nm}$ K=1000
$\Delta\rho \text{ (}\mu\text{m}^2.\text{s}^{-1}\text{)}$	$5.10^{-3}$	$6.10^{-3}$	$2.10^{-3}$	$2.10^{-3}$	$5.10^{-3}$

**Tableau III.3:** Erreur commise sur les coefficients de diffusion  $D_1$  par pooling selon les différents paramètres des simulations.

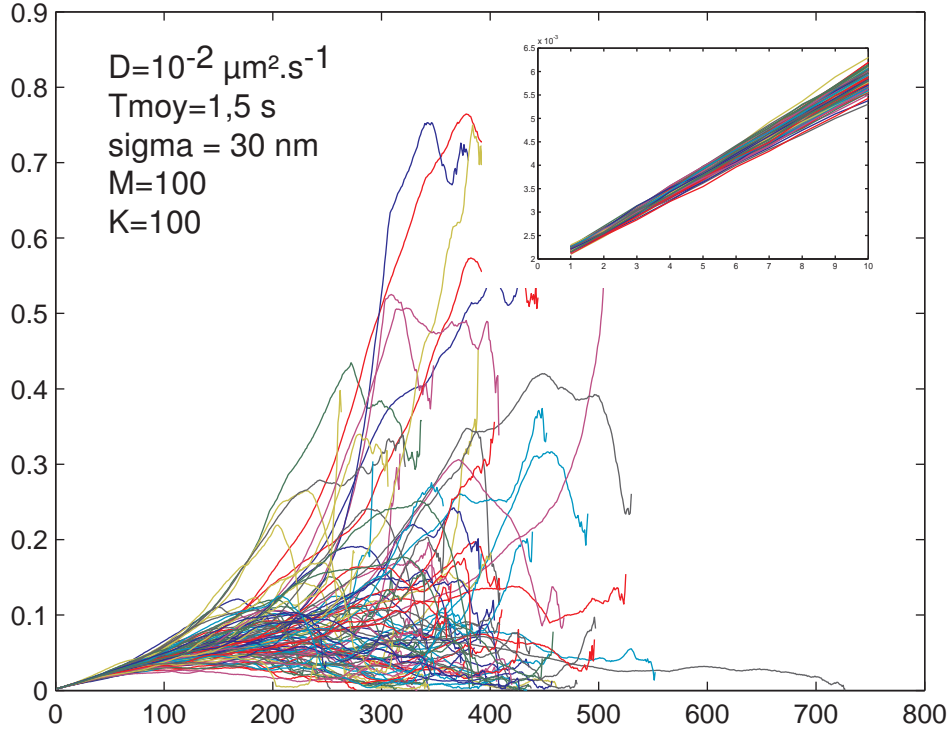
la molécule d'ADN. Les MSDs poolées ont été calculées en regroupant plus d'une centaine d'événements d'interaction pour chaque jeu de conditions.

**Précision du pooling - Simulation** Estimer les barres d'erreurs relatives à la MSD poolée n'est pas aisé. Or, il est toujours délicat de donner une courbe expérimentale sans y faire figurer les barres d'erreurs. Pour estimer les incertitudes sur la MSD après pooling, nous avons fait des simulations sous MatLab. Ces simulations tirent  $M$  marches aléatoires à 1D (pendant des  $M$  trajectoires expérimentales) de coefficient de diffusion constant  $D_1$  et de durée  $T_m$ .  $T_m$  suit une distribution exponentielle décroissante, de temps caractéristique  $T_{\text{moy}}$ . Pour prendre en compte la précision de pointé expérimentale, nous avons ajouté aux coordonnées un bruit gaussien de moyenne nulle et de largeur  $\sigma$ . Les distances des trajectoires sont regroupées pour calculer une MSD poolée. Ce tirage est reconduit  $K$  fois : nous obtenons ainsi  $K$  MSDs poolées. Les  $K$  coefficients de diffusion issus des MSDs simulées sont comparés au  $D_1$  pour estimer l'erreur commise.

Durant ces simulations, les paramètres  $\sigma$ ,  $M$ ,  $T_{\text{moy}}$  et  $K$  ont été changé ( $\sigma = 30$  ou  $50$  nm ;  $M = 100$  ou  $200$  trajectoires ;  $T_{\text{moy}} = 1,5$  ou  $15$  s ;  $K = 100$  ou  $1000$  répétitions).  $D_1$  est resté constant, égal à  $10^{-2} \mu\text{m}^2.\text{s}^{-1}$ . Seul  $T_{\text{moy}}$  a entraîné des variations appréciables comme le montre le Tableau III.3 qui donne l'erreur commise  $\Delta\rho$  après 1 pooling :  $\Delta\rho = \sqrt{M} \times \text{sd}$  où sd est l'écart-type des coefficient de diffusion trouvés lors de ces simulations après  $M$  pooling. Logiquement, le calcul d'une MSD poolée sur un ensemble de trajectoires moyennement plus longues ( $T_{\text{moy}} = 15$  s) est plus précis.

Sur la figure III.6 apparaissent les  $K=100$  MSDs poolées correspondant au jeu de paramètre suivant :  $M = 100$  trajectoires,  $\sigma = 30$  nm et  $T_{\text{moy}} = 1,5$  s. Ces simulations donnent deux informations importantes :

- seuls les premiers points de la MSD poolée doivent être pris en compte pour le calcul de la pente et donc du coefficient de diffusion. Dans la suite, nous calculerons les coefficients de diffusion par ajustement linéaire des MSDs poolées sur 5 points à partir du deuxième.



**Figure III.6: Simulation du pooling.** 100 « poolings » réalisés dans des conditions identiques à celles de notre situation expérimentale ont été simulés : les 100 MSDs obtenues ne se superposent que sur les tous premiers points, ensuite, leurs allures divergent. L'insert est un zoom du coin inférieur gauche du graphe (l'axe des abscisses est gradué en unité de  $t_{acq}$ ). Il y a une bonne correspondance pour les 5 premiers points des différentes MSDs poolées, l'erreur sur cette plage est évaluée à  $\Delta\rho = \pm 3.10^{-3} \mu\text{m}^2.\text{s}^{-1}$ .

- la différence entre la pente la plus petite et la plus grande donne une estimation de l'erreur  $\Delta\rho$  commise après 1 pooling. Pour cette simulation, nous trouvons une différence de  $5.10^{-3} \mu\text{m}^2.\text{s}^{-1}$  entre les deux pentes extrêmes. En prenant l'arrondi supérieur, nous estimons l'erreur à  $0,3.10^{-2} \mu\text{m}^2.\text{s}^{-1}$ , ce qui permet de faire figurer des barres d'erreur sur les graphes des MSD poolées qui suivront.

**Comparaison Pooling - Moyenne des MSDs :** la question suivante peut venir à l'esprit du lecteur : pourquoi ne pas avoir pris la moyenne des MSDs individuelles ? Y a-t-il un intérêt à travailler par pooling ? Pour  $m = 1..M$  trajectoires, la MSD moyenne  $\bar{\rho}$  à l'instant  $n \Delta t$  est donnée par :

$$\bar{\rho}(n\Delta t) = \frac{1}{M} \sum_{m=1}^M \left( \frac{1}{N_m - n} \sum_{i=1}^{N_m - n} \left( x_{i+n}^{(m)} - x_i^{(m)} \right)^2 \right). \quad (\text{III.21})$$

en ayant encore supposé que  $\forall m, n \leq N_m$ . Le coefficient de diffusion  $\bar{D}$  issu de  $\bar{\rho}$  :

$\bar{\rho}(t) = 2\bar{D}t$  est ainsi égal à la moyenne des coefficients de diffusion  $D_m$  obtenus séparément pour chaque trajectoire :

$$\bar{D} = \frac{1}{M} \sum_{m=1}^M D_m. \quad (\text{III.22})$$

Nous pouvons comparer le coefficient de diffusion  $\bar{D}$  avec le coefficient  $D^{\text{Pool}}$  obtenu par pooling. La loi de diffusion  $\left| x_{i+n}^{(m)} - x_i^{(m)} \right|^2 \simeq D_m \times n\Delta t$  permet d'estimer la MSD poolée donnée à l'équation III.20 :

$$\rho^{\text{Pool}} \simeq \frac{\sum_{m=1}^M D_m (N_m - n)}{\sum_{m=1}^M (N_m - n)} \times n\Delta t. \quad (\text{III.23})$$

D'où l'expression approchée du coefficient  $D^{\text{Pool}}$  :

$$D^{\text{Pool}} \simeq \frac{\sum_{m=1}^M D_m (N_m - n)}{\sum_{m=1}^M (N_m - n)}. \quad (\text{III.24})$$

$D^{\text{Pool}}$  apparaît comme une moyenne des coefficients  $D_m$  pondérée par le nombre de points de chaque trajectoire  $m$ , limitant ainsi les erreurs liées au faible nombre de points de certaines trajectoires lorsque l'on fait simplement la moyenne. C'est parce qu'il permet d'inclure les événements courts dans les analyses que nous avons choisi le pooling.

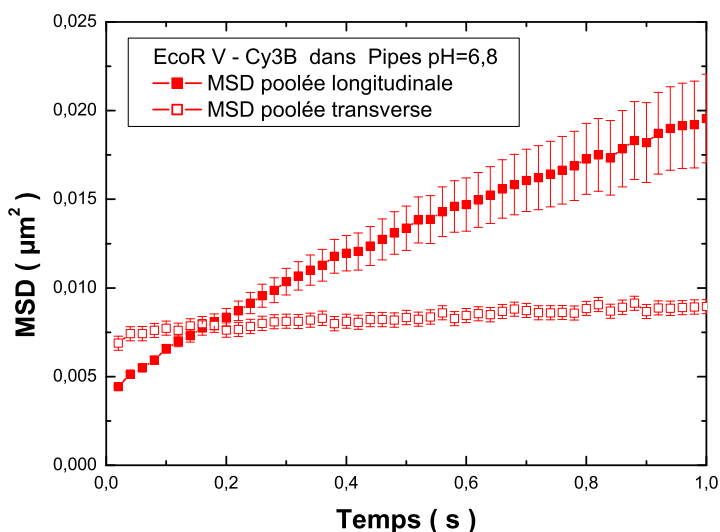
**Pooling des données EcoRV-Cy3B dans le Pipes pH = 6,8 :** la Figure III.7 montre les MSDs longitudinales et transverses obtenues après pooling de 120 événements d'interaction de l'enzyme EcoRV marquée au Cy3B interagissant avec de l'ADN étiré dans le tampon Pipes à pH=6,8. La MSD poolée longitudinale varie linéairement, l'ajustement sur les 5 premiers points permet de déduire un coefficient de diffusion :  $D_1^{\text{Pipes } 6,8} = 1,2 \pm 0,3 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ , comparable aux coefficients de diffusion des trajectoires prises individuellement (cf. Section précédente). La MSD poolée transverse est constante, témoignant d'un mouvement confiné.

### III.2.4 Influence du tampon et du pH

#### Influence du pH :

En travaillant dans un même tampon, nous avons voulu voir quelle était l'influence du pH. Pour cela, nous avons travaillé avec le Pipes à deux pHs différents : 6,8 et 7,5. Les MSDs poolées apparaissent sur la figure III.8.





**Figure III.7:** Pooling de 120 événements d'interaction EcoRV-Cy3B - ADN. La MSD poolée longitudinale (transverse) apparaît en carré plein (vide).

L'ajustement sur les premiers points permet de déterminer le coefficient de diffusion de l'enzyme EcoRV marquée à Cy3B dans le Pipes à pH=6,8 :  $D_1^{\text{Pipes } 6,8} \approx 1,2 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ . Il n'y a donc pas de changement par rapport à  $D_1^{\text{Pipes } 7,5}$  : le sliding d'EcoRV ne dépend donc pas a priori du pH dans la gamme 6,8-7,5, qui, il est vrai, est assez restreinte, mais 0,5 unité pH avait suffi aux auteurs de [85] à observer des différences notables.

### Influence du tampon :

En mesurant les constantes de Michaelis<sup>3</sup> et les constantes de dissociation non-spécifiques  $k_{off}^{ns}$ , Wenner *et al.* ont montré que la concentration ainsi que la nature du tampon pouvaient significativement jouer sur les interactions entre EcoRV et l'ADN spécifique ou non [85]. Inspirés de cette étude, nous avons voulu voir quelle était l'influence du tampon d'interaction dans notre configuration expérimentale. Nous avons travaillé avec deux tampons ajustables à pH = 7,5 : le Pipes et le Hepes. Pour ces deux tampons, nous avons comparé la MSD poolée (cf. Figure III.9). Nous avons trouvé un coefficient de diffusion dans le tampon Hepes :  $D_1^{\text{Hepes } 7,5} \approx 1,5 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ , légèrement plus grand que celui dans le Pipes :  $D_1^{\text{Pipes } 7,5} \approx 1,2 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ . La différence entre ces deux valeurs est incluse dans les barreaux d'erreur : l'éventuel effet du tampon n'est pas décelable sur la diffusion 1D d'EcoRV.

<sup>3</sup>Constante thermodynamique qui traduit l'affinité d'une enzyme pour son substrat.

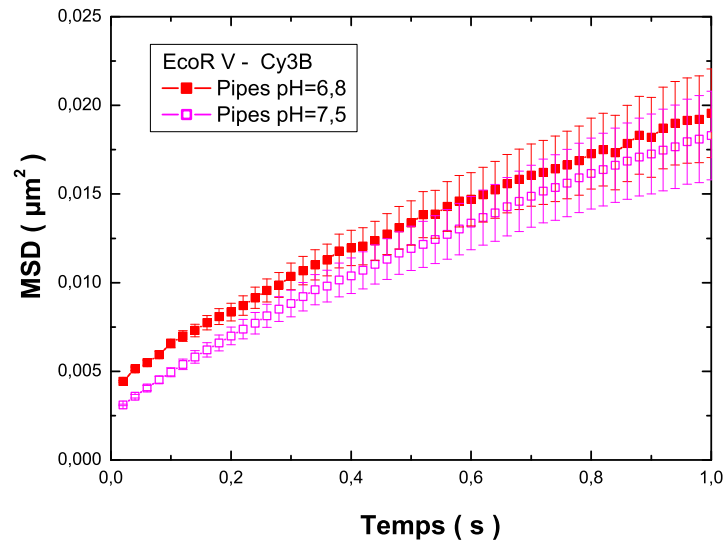


Figure III.8: Le sliding d'EcoRV ne dépend pas du pH dans la gamme 6,8 - 7,5.

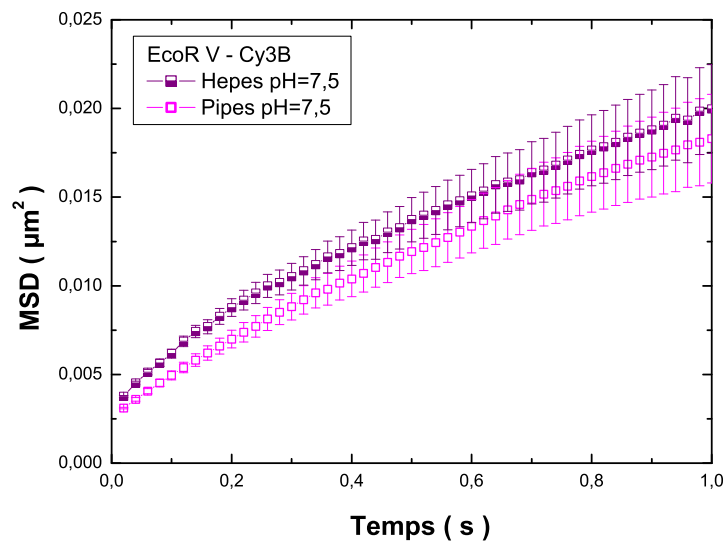


Figure III.9: Le sliding d'EcoRV ne dépend pas du tampon d'interaction.

## III.3 EcoRV saute aussi

### III.3.1 EcoRV saute

#### III.3.1.1 Observations expérimentales

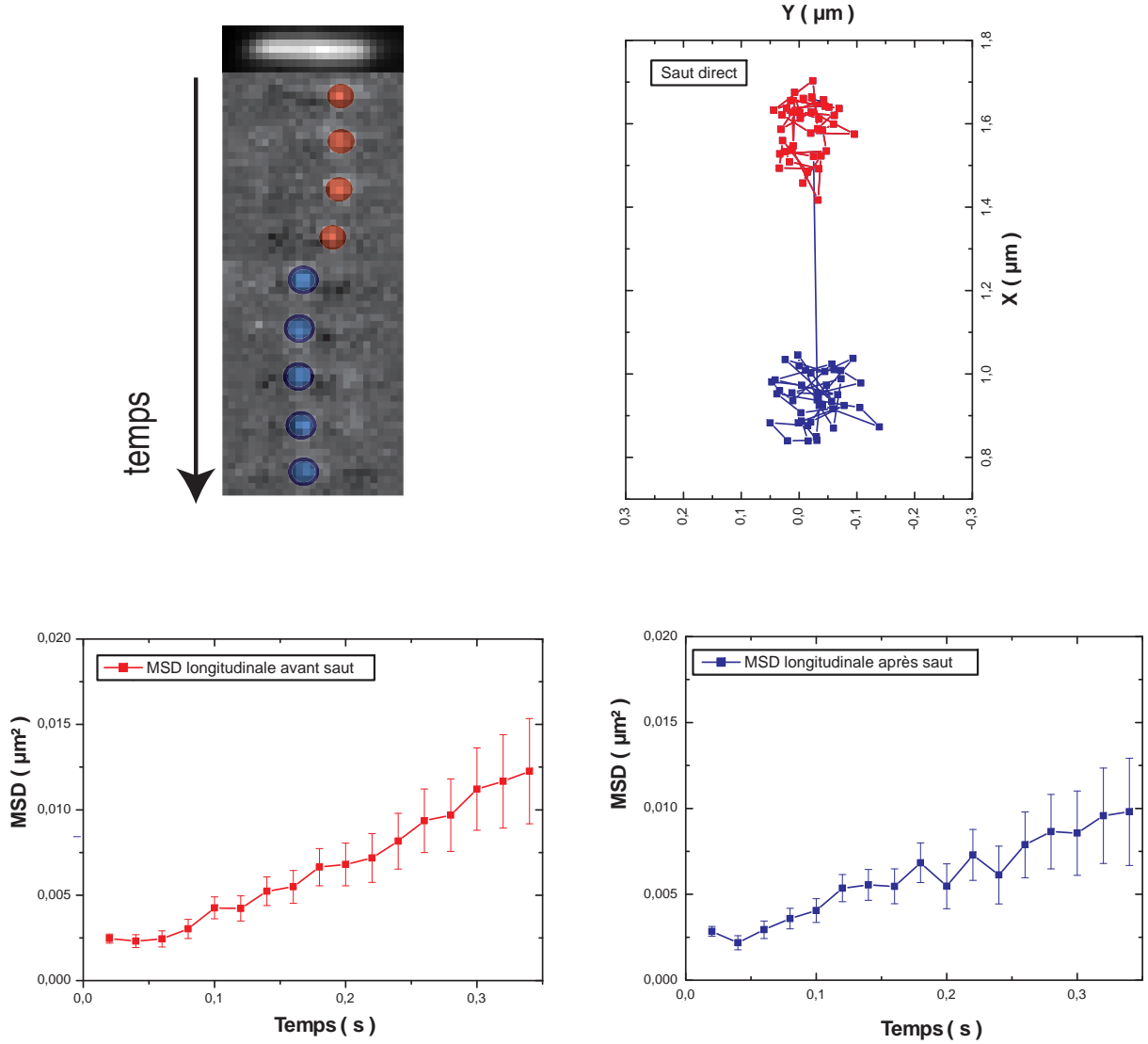
**Description :** sur plusieurs films expérimentaux d'interaction ADN-protéine, nous avons observé de longs et brefs déplacements de la protéine le long de l'ADN. Par la suite nous désignerons « sauts » ce type de déplacement le long de l'ADN, supérieur à  $x_C = 200$  nm, en une ou deux images successives. Nous justifions ci-dessous le choix de la valeur de  $x_C$ . La Figure III.10 montre un saut de 550 nm le long de l'ADN d'une enzyme marquée au Cy3B qui a eu lieu en moins de  $t_{\text{acq}}$ , la Figure III.11 est un exemple de saut de 300 nm qui a eu lieu en l'espace de 2 images successives (soit  $2t_{\text{acq}}$ ).

**Conséquences :** les sauts supérieurs à 200 nm modifient la MSD comme le prouvent les graphes de la Figure III.12 relatifs au saut de la Figure III.10 :

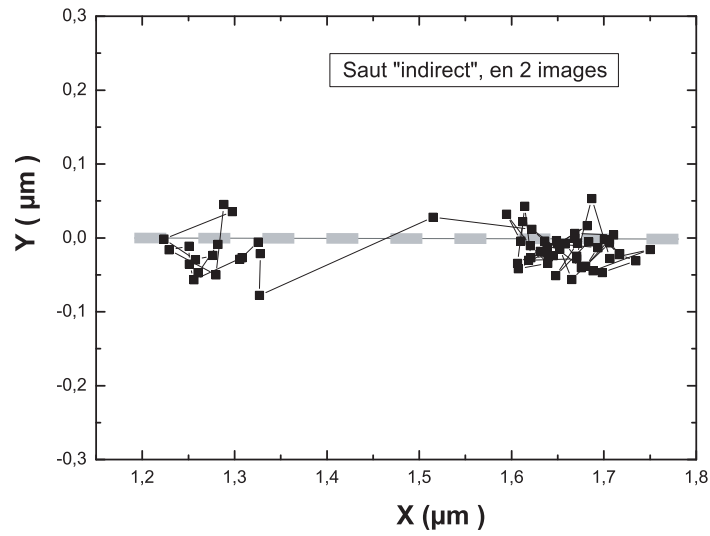
- les carrés noirs représentent la MSD calculée sur l'ensemble des points de la trajectoire en tenant compte du saut. L'ajustement linéaire de cette MSD donne un coefficient de diffusion de  $0,18 \mu\text{m}^2 \cdot \text{s}^{-1}$ .
- les ronds rouges et bleus correspondent aux MSD calculées sur les deux trajectoires secondaires obtenues en enlevant le saut (rouge : avant saut - bleu : après saut). En considérant ces portions de trajectoires, les coefficients de diffusion obtenus sont  $D_1^{\text{av saut}} = 3,1 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$  et  $D_1^{\text{ap saut}} = 2,6 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ , identiques aux incertitudes près et similaires au  $D_1^{\text{Pool}}$ .

Lors des analyses, ce sont les coefficients de diffusion calculés sur les trajectoires secondaires qu'il faut considérer. Par conséquent, les grands sauts ( $\geq 200$  nm) doivent être exclus des analyses. Nous verrons par la suite que les petits sauts ( $\leq 200$  nm) n'interviennent pas dans la valeur estimée du coefficient de diffusion (cf. Section III.4). Dans l'étude précédente portant sur le sliding, n'ont été considérés que les événements d'interaction sans sauts apparents : lorsqu'une trajectoire présentait  $N$  sauts, elle a été scindée en  $N+1$  trajectoires secondaires.

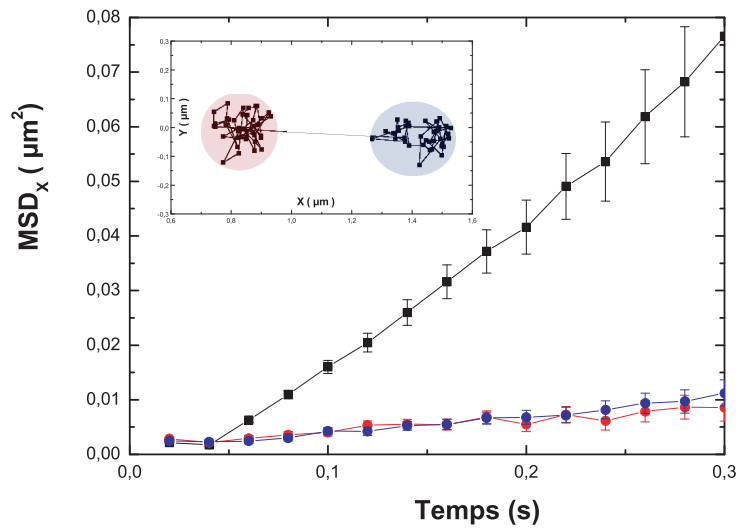
**Critères :** dans nos conditions expérimentales, l'amplitude caractéristique des fluctuations longitudinales est de l'ordre de 80 nm pour un temps caractéristique de 1 ms (cf. Annexe C). La distance moyenne parcourue par diffusion par la protéine pendant le temps  $t_{\text{acq}} = 20$  ms est de 30 nm. Il faut aussi tenir compte de la précision de pointé  $\sigma \approx 30$  nm. Pour s'affranchir de manière certaine de la composante Brownienne du mouvement et des fluctuations longitudinales de l'ADN, nous avons fixé la limite inférieure de la taille des sauts à  $x_C = 200$  nm. Pour des déplacements plus petits, principalement à cause des



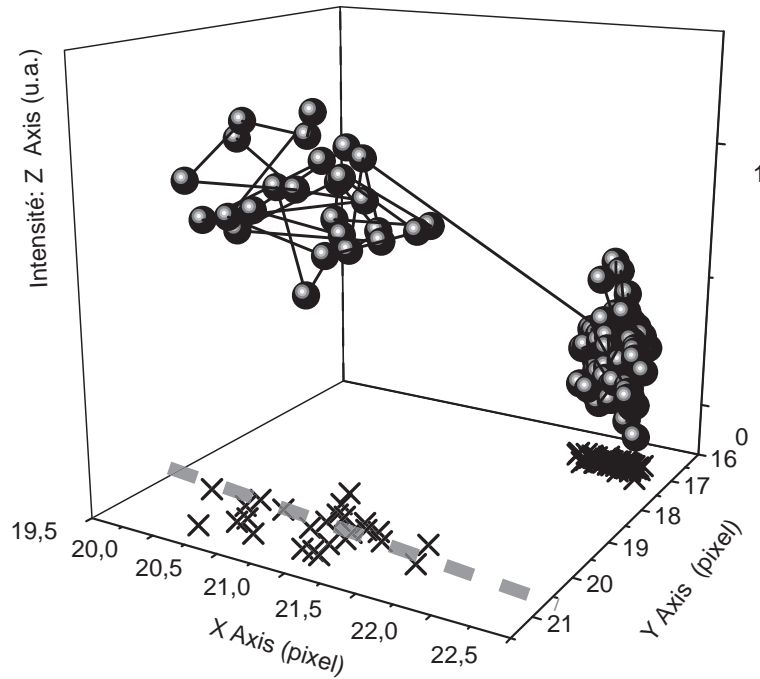
**Figure III.10: Un saut de 550 nm le long de l'ADN en 1 frame. Haut gauche :** Haut : image de fluorescence de l'ADN marqué au *SyBR* - Dessous : extrait du film sur la même zone de l'enzyme marquée à Cy3B interagissant avec l'ADN (cercles rouges : 4 images juste avant le saut, cercles bleus : 5 images juste après le saut). **Haut droite :** trajectoire reconstruite ( $X$  désigne l'axe longitudinal,  $Y$  l'axe transverse de l'ADN). Rouge : avant le saut - Bleu : après le saut. **Bas gauche :** MSD calculée sur la première partie de la trajectoire (avant saut),  $D_1^{\text{av saut}} = 3,1 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ . **Bas droite :** MSD calculée sur la seconde partie de la trajectoire (après saut),  $D_1^{\text{ap saut}} = 2,6 \cdot 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ .



**Figure III.11: Un saut de 300 nm le long de l'ADN en 2 frames.** La direction de l'ADN est repérée par le segment en pointillé.



**Figure III.12: Effet des sauts sur le calcul de la MSD.** Si l'on inclut le saut de 550 nm de la figure III.10 dans le calcul de la MSD, nous obtenons la courbe de plus haute pente (noir) : le coefficient de diffusion calculé sur cette MSD est beaucoup plus élevé et ne correspond plus à celui de la diffusion 1D des enzymes. Les MSDs obtenues lorsque le saut est enlevé apparaissent en rouge (avant saut) et bleu (après saut).



**Figure III.13: Exemple de « faux » saut :** dans cet exemple, l'enzyme saute depuis l'ADN vers la surface. L'ADN est repéré par sa projection dans le plan de la lamelle (trait pointillé). Les boules noires représentent les différentes positions de la protéine au cours du temps, les croix sont les projections de ces positions dans le plan de la lamelle. Au début (points situés sur la gauche de la figure), l'enzyme diffuse sur l'ADN, on devine le mouvement de diffusion et le mouvement de fluctuation ; à la fin (points les plus à droite), l'enzyme est collée à la surface. Ce genre de « saut » n'est pas pris en compte lors de nos analyses. Cet événement permet accessoirement de mesurer la portée  $d$  de l'onde évanescente puisque l'intensité  $I$  est reliée à la côte  $z$  selon :  $\ln \left( \frac{I(z_1)}{I(z_2)} \right) = \frac{z_2 - z_1}{d}$ . Pour cet exemple,  $d$  a été estimée à 105 nm, en accord avec la théorie.

fluctuations de l'ADN, il est impossible de savoir si le mouvement est du sliding ou du jumping.

**Occurrence :** ces sauts ont été observés sur environ 10% des interactions. Cependant, il n'est pas rare (1 cas sur 5) de ne pas disposer de suffisamment d'images avant et/ou après ce qui apparaît comme un saut : dans ce cas, nous ne pouvons savoir s'il s'agit d'un réel saut le long de l'ADN ou d'un saut vers et/ou depuis la surface comme illustré sur la Figure III.13. Par précaution, ces sauts potentiels n'ont pas été pris en compte dans les analyses.

## III.3.2 La distribution des sauts

### III.3.2.1 La cumulative inversée : un outil approprié

**Description :** pour analyser la distribution des sauts, nous avons choisi de travailler avec la cumulative inversée de la taille des sauts observés. À cause de notre résolution optique, nous ne pouvons pas voir les sauts  $\leq 200$  nm, nous ne pouvons donc pas travailler avec la cumulative habituelle.

La cumulative inversée  $\mathcal{C}(x)$  est en quelque sorte la primitive de la distribution de probabilité  $\mathcal{P}(x)$  de faire un saut de taille  $x$  puisqu'elle donne le nombre de saut observés supérieurs à  $x$ . En fait, nous travaillons avec la cumulative inversée *normalisée* par rapport au nombre total d'événements d'interaction  $N_{\text{int}}$  déterminé pour chaque condition expérimentale en ajustant la cumulative des durées d'interaction par une exponentielle décroissante selon  $\mathcal{C}(t) = N_{\text{int}}e^{-\frac{t}{T}}$  (voir la Figure III.19).

### III.3.2.2 Influence du tampon et du pH

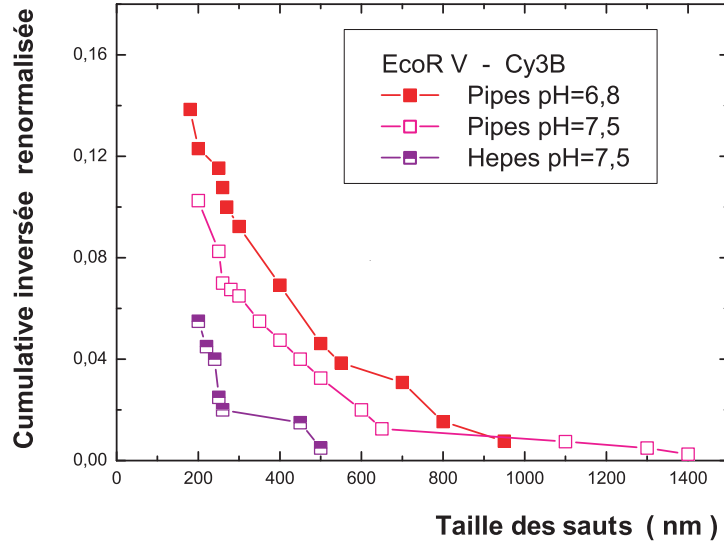
Alors que le coefficient de diffusion 1D semble indépendant des conditions expérimentales, le nombre de saut observés en dépend quantitativement. Sur la Figure III.14 sont représentées les cumulatives des sauts observés pour l'enzyme EcoRV-Cy3B dans les différentes conditions précédemment testées (tampon et pH). Plusieurs remarques peuvent être faites quant aux différences entre ces cumulatives :

- le nombre moyen de sauts visibles par interaction (donné par l'ordonnée en  $x_C = 200$  nm) change selon les conditions. Il y a en effet plus de saut dans le tampon Pipes que dans le tampon Hepes. En outre, dans, le tampon Pipes, il y a plus de saut à bas pH.
- l'allure générale des cumulatives change également en fonction des conditions : à nombre d'interaction sensiblement égal, nous n'avons pas observé de saut supérieur à 500 nm dans le tampon Hepes, ce qui n'est pas le cas dans le tampon Pipes.

## III.3.3 Pourquoi EcoRV saute t-elle ?

### III.3.3.1 Une enzyme ne peut pas en cacher une autre

La première idée qui peut venir à l'esprit lorsque nous observons les sauts en visionnant les films est qu'il s'agit en fait de deux enzymes différentes : alors que la première enzyme se détache de l'ADN, la seconde s'accroche à l'ADN immédiatement après, en moins de  $2t_{\text{acq}} = 40$  ms. Nous avons exclu ce cas de figure en considérant la fréquence d'association expérimentale des enzymes à l'ADN à la concentration à laquelle nous travaillons. En effet, le temps moyen entre 2 interactions sur une même molécule d'ADN est de  $T = 30$  s. Par conséquent, la probabilité pour avoir deux événements d'interaction qui se suivent à moins



**Figure III.14:** Cumulative inversée et renormalisée des sauts observés pour l'enzyme marquée au Cy3B. **Carrés pleins** : Pipes, pH=6,8 - **Carrés vides** : Pipes, pH=7,5 - **Carrés barrés** : Hepes, pH=7,5

de 2 images est donnée par la réalisation d'une loi de Poisson de paramètre  $\lambda = \frac{2t_{\text{accq}}}{T}$  pour  $k = 2$  :

$$\mathcal{P}(k) = e^{-\lambda} \frac{\lambda^k}{k!} \implies \mathcal{P}(2) \approx 9.10^{-7} \ll 1\%, \quad (\text{III.25})$$

Il s'agit bien de la même enzyme : les sauts que nous observons sont par conséquent dus à des déplacements brefs d'une même enzyme.

### III.3.3.2 La diffusion à 1D n'explique pas les sauts.

Dans l'hypothèse où le mouvement d'EcoRV serait purement diffusif à 1D de coefficient de diffusion  $D_1$ , les sauts correspondraient à des déplacements diffusifs grands devant la moyenne  $\bar{x}$  des déplacements observés dont un majorant est :  $\bar{x} + \sigma = \sqrt{2D_1\Delta t} + 0,03 \approx \sqrt{2 \times 0,01 \times 0,04} + 0,03 = 60 \text{ nm}$ , avec  $\sigma$  la précision de pointé. Pour estimer la fréquence de tels déplacements, il faut connaître la probabilité des événements rares d'une distribution gaussienne.

Dans la limite visqueuse où l'inertie de la particule Brownienne est négligée, la densité de probabilité  $p(x, t)$  dans le formalisme de Langevin obéit à l'équation de diffusion :

$$\frac{\partial p}{\partial t} = D_1 \frac{\partial^2 p}{\partial x^2}, \quad (\text{III.26})$$



La solution fondamentale de l'équation III.26 est une gaussienne de variance  $\omega^2 = 2D_1\Delta t$  :

$$p(x, t) = \frac{1}{\sqrt{4\pi D_1 t}} \exp\left(-\frac{x^2}{4D_1 t}\right) = \frac{1}{\sqrt{2\pi\omega}} \exp\left(-\frac{x^2}{2\omega^2}\right). \quad (\text{III.27})$$

La probabilité d'avoir un déplacement de taille  $x$  supérieur à  $n\omega$  est donnée par :

$$\bar{\mathcal{P}}(x \geq n\omega) = 2 \times \int_{n\omega}^{\infty} \frac{1}{\sqrt{2\pi\omega}} \exp\left(-\frac{x^2}{2\omega^2}\right) dx = \text{erfc}\left(\frac{n}{\sqrt{2}}\right). \quad (\text{III.28})$$

où erfc est la fonction erreur complémentaire définie par :

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-u^2} du.$$

En prenant  $D_1 = 10^{-2} \mu\text{m}^2.\text{s}^{-1}$  et  $\Delta t = 40$  ms, c'est-à-dire  $\omega = 28$  nm, nous pouvons calculer la probabilité d'avoir un déplacement de  $7\omega \sim 200$  nm pendant l'intervalle  $\Delta t$  :

$$\bar{\mathcal{P}}_{7\omega} = \text{erfc}\left(\frac{7}{\sqrt{2}}\right) \approx 2,6 \cdot 10^{-12}. \quad (\text{III.29})$$

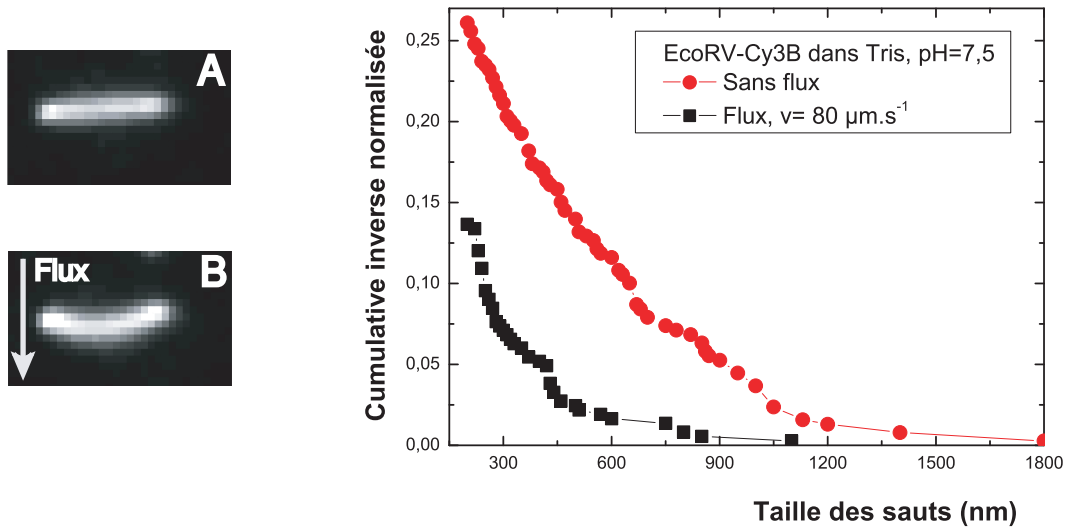
Une interaction dure en moyenne  $T = 4,4$  s,  $\bar{\mathcal{P}}_{7\omega}$  doit donc être multiplié par  $\frac{T}{\Delta t}$ , il vient  $\mathcal{P}_{7\omega} = 2,9 \cdot 10^{-10}$ . Cette probabilité reste extrêmement faible et signifie qu'il faudrait attendre plus de 400 ans pour observer un tel événement ! Elle ne rend donc pas compte de la fréquence des sauts observés : les sauts ne sont pas dus au sliding.

### III.3.3.3 Existe-t-il une diffusion 3D ou 1D rapide ?

Nous pourrions imaginer que lorsque l'enzyme se décroche de l'ADN, elle reste à proximité de celui-ci, comme guidée par un tunnel et diffuse toujours à 1D mais sans frottement avec l'ADN. Cela pourrait être la conséquence d'une interaction électrostatique entre l'enzyme et l'ADN qui confinerait l'enzyme ou plus simplement, un effet probabiliste.

Une autre hypothèse consiste à dire que les sauts observés correspondent à des excursions 3D de la protéine sans aucune interaction avec l'ADN. Avec un coefficient tridimensionnel  $D_3 \sim 50 \mu\text{m}^2.\text{s}^{-1}$ , la distance couverte par la protéine pendant 40 ms est de l'ordre du  $\mu\text{m}$ . L'hypothèse est plausible.

À ce niveau, il n'est pas possible de savoir si les sauts observés sont dus à des mécanismes 1D ou 3D. Pour identifier le mécanisme sous-jacent, des expériences d'interaction ADN-protéine *sous flux* ont été réalisées par Andréas Biebricher, post-doctorant dans l'équipe. En pratique, nous appliquons en permanence un flux hydrodynamique perpendiculairement à la direction de l'ADN étiré : l'ADN est courbé (cf. Figure III.15, A et B). Ces expériences consistent à regarder quel est l'effet du flux sur la distribution des sauts pour élucider le mécanisme. *A priori*, le flux ne doit pas affecter les mécanismes de diffusion 1D, alors que lors de sauts 3D, les protéines seront entraînées par le flux : la distribution des sauts devraient changer.



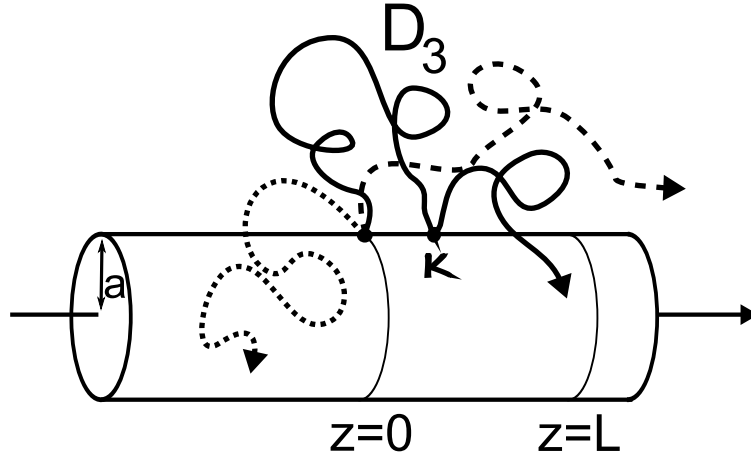
**Figure III.15: Expériences réalisées sous flux.** En rouge (ronds), la cumulative des sauts observés pour l'enzyme marquée à Cy3B dans le Tris, pH=7,5. En noir (carrés), la cumulative des sauts observés en présence d'un flux transverse (B) de vitesse  $v = 80 \mu\text{m.s}^{-1}$ .

### III.3.3.4 Il existe une diffusion 3D.

La Figure III.15 montre les cumulatives des sauts d'EcoRV-Cy3B avec (carrés noirs) ou sans (ronds rouges) flux. La vitesse de l'écoulement appliquée est dans ce cas  $v = 70 \mu\text{m.s}^{-1}$ . En présence de flux, non seulement, le nombre moyen de sauts par interaction est réduit de moitié, mais la cumulative est déplacée vers les petits sauts : seuls 5% des sauts sont supérieurs à 700 nm en présence de flux, contre 30% sinon. Cette diminution du nombre de grand sauts s'explique aisément dans le cas d'un mécanisme de diffusion 3D : les sauts de taille supérieure à  $\frac{D_3}{v} = 700 \text{ nm}$  sont les plus sensibles au flux. Ces observations sont des preuves considérables en faveur d'un mécanisme 3D pour expliquer les sauts d'EcoRV.

Les expériences ont été réalisées dans plusieurs tampons, à différents pHs. Les conclusions sont les mêmes mais nous avons choisi de montrer les résultats pour le tampon Tris, pH=7,5, pour lequel l'effet est le plus important. Pour plus de détails sur les résultats de ces expériences, je renvoie le lecteur à [95].

Nous avons gardé l'hypothèse de saut 3D pour développer une approche numérique dont les détails sont décrits à la section suivante.



**Figure III.16: Modèle utilisé dans les calculs.** L'ADN est un cylindre d'axe  $z$ , de longueur  $L$ , de rayon  $a = a^* + R_H$  où  $a^*$  est le rayon de la molécule d'ADN et  $R_H$  la taille de l'enzyme.

### III.3.4 Modélisation des sauts

Pour vérifier l'hypothèse d'un processus de diffusion 3D complémentaire au sliding, et afin de comprendre l'influence des différents paramètres, nous avons calculé les cumulatives analytiques associées à un modèle de diffusion 3D. Cette étude fait l'objet d'une collaboration avec Claude Loverdo, Raphaël Voituriez et Olivier Bénichou, du Laboratoire de Physique Théorique de la Matière Condensée de l'UPMC.

#### III.3.4.1 Description du modèle

Nous utilisons les coordonnées cylindriques. L'ADN est modélisé par un cylindre de longueur  $L$  selon l'axe  $z$  et de rayon  $a^*$  ( $L = 2,2 \mu\text{m}$  et  $a^* = 1\text{nm}$ ). Pour prendre en compte la taille de l'enzyme, nous prenons un rayon hydrodynamique effectif  $a$  tel que  $a = a^* + R_H$ , où  $R_H$  est le rayon de l'enzyme. L'enzyme quitte l'ADN à  $t = 0$  et y revient à l'instant  $t$ . Elle diffuse à 3 dimensions avec un coefficient  $D_3$ . Selon les observations expérimentales, nous considérons que l'enzyme est perdue si l'excursion à 3D dure plus longtemps que  $t_{obs} = 2\Delta t = 40 \text{ ms}$  ou si elle se rattache en dehors du segment  $[0, L]$  (cf. Figure III.16).

Soit  $P(z|t_{obs})$  la probabilité de faire un saut de taille  $z$  pendant le temps d'observation  $t_{obs}$ .  $P(z|t_{obs})$  peut être reliée à  $P(z|t)$ , la probabilité de s'être déplacé d'une distance  $z$  pendant le temps  $t$  :

$$P(z|t_{obs}) = \int_0^{t_{obs}} P(z|t)f(t)dt, \quad (\text{III.30})$$

où  $f(t)$  est la probabilité de revenir sur le cylindre à l'instant  $t$ .

Le problème est invariant par translation selon  $z$  :  $f(t)$  peut être ramenée à la probabilité en deux dimensions de revenir sur un disque de rayon  $a$  à l'instant  $t$  en ayant

quitté  $r_0$  à l'instant  $t = 0$ . À chaque rencontre avec le cylindre de l'ADN, l'enzyme a une probabilité finie de s'adsorber. Cette hypothèse nécessite d'introduire un paramètre  $\kappa$ , caractéristique de la longueur de ce processus d'adsorption. La transformée de Laplace  $F(s)$  de  $f(t)$  peut alors être exprimée en fonction de  $K$ , la fonction de Bessel modifiée de seconde espèce comme suit [96] :

$$F(s) = \frac{K_0\left(a\sqrt{\frac{s}{D_3}}\right)}{K_0\left(a\sqrt{\frac{s}{D_3}}\right) + \kappa^{-1}\sqrt{\frac{s}{D_3}} \times K_1\left(a\sqrt{\frac{s}{D_3}}\right)}. \quad (\text{III.31})$$

Inverser cette fonction n'est analytiquement pas facile, cela est fait numériquement [97]. Nous pouvons ensuite calculer l'intégrale de la densité de sauts :

$$c(z, t_{obs}) = \int_z^L P(u, t_{obs}) \times \frac{L-u}{L} du. \quad (\text{III.32})$$

où le facteur  $\frac{L-u}{L}$  rend compte de la taille finie de l'ADN. La probabilité qu'une enzyme s'échappe de l'ADN pendant le temps  $t_{obs}$  est donnée par  $1 - c(0, t_{obs})$ . Lorsque l'enzyme s'échappe, l'interaction est considérée comme terminée. D'où l'expression du nombre moyen  $N_{tot}$  de sauts par interaction :

$$N_{tot} = \frac{1}{1 - c(0, t_{obs})}.$$

La probabilité d'observer un saut de taille supérieure à  $z$  correspond au produit de la densité de probabilité de saut supérieur à  $z$  par le nombre moyen de sauts :

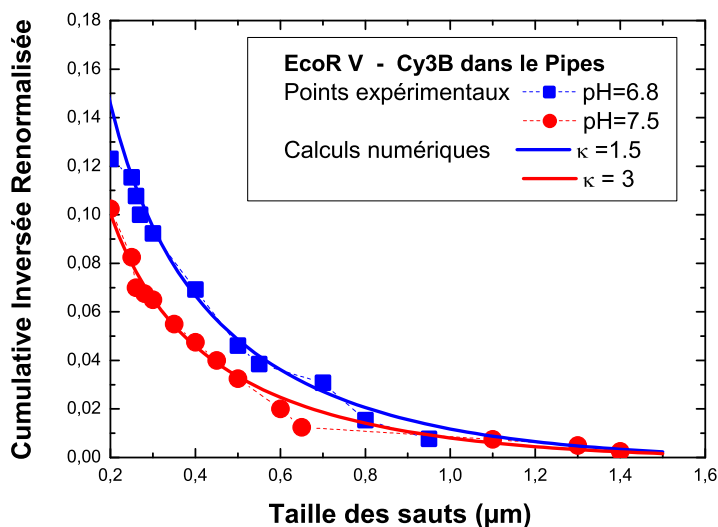
$$\mathcal{C}(z, t_{obs}) = N_{tot} c(z, t_{obs}).$$

$\mathcal{C}(z, t_{obs})$  équivaut à la cumulative inversée normalisée expérimentale.  $\kappa$  est le seul paramètre ajustable dans ces calculs.

### III.3.4.2 Résultats pour EcoRV - Cy3B

En prenant les valeurs :  $a = 5$  nm et  $D_3 = 54 \mu\text{m}^2.\text{s}^{-1}$ , justifiées par les mesures de FCS décrites au chapitre précédent (partie II.2.4.3) et une taille de l'ADN  $L = 2,2 \mu\text{m}$ , ce modèle a permis d'obtenir des courbes analytiques des cumulatives qui se superposent très bien aux points expérimentaux pour l'enzyme marquée à Cy3B dans le tampon Pipes (cf. Figure III.17). Les valeurs du paramètre d'adsorption  $\kappa$  pour le Pipes à pH=6,8 et à pH=7,5 sont respectivement  $\kappa_{6,8} = 1,5 \text{ nm}^{-1}$  et  $\kappa_{7,5} = 3 \text{ nm}^{-1}$ .

Dans les autres tampons (Hepes et Tris), la concordance expérience-calcul n'est que qualitative. Un modèle plus complet qui prendrait en compte les effets électrostatiques devrait permettre de réduire ces écarts.



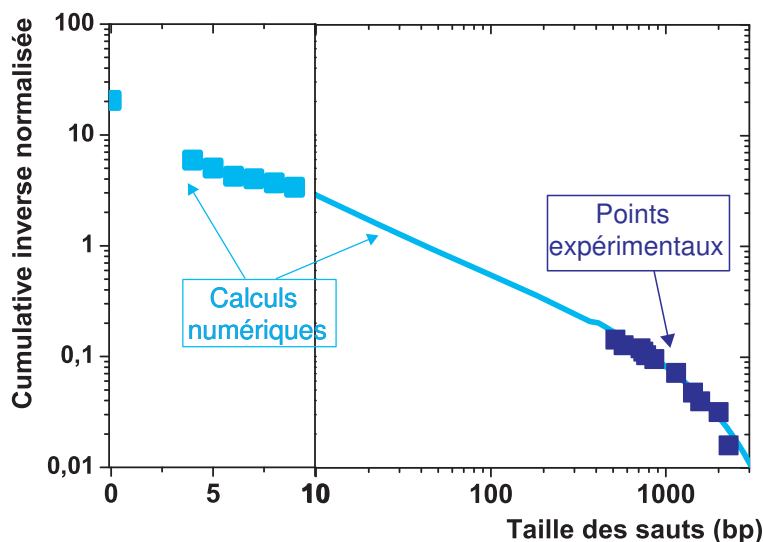
**Figure III.17:** Superposition des cumulatives renormalisées expérimentales (ronds et carrés) et analytiques (traits continus) pour les sauts observés avec l'enzyme EcoRV-Cy3B dans le tampon Pipes. Haut : pH= 6,8 - Bas : pH=7,5.

### III.3.4.3 Conclusions

À partir d'un modèle simple, nous parvenons à expliquer les sauts d'EcoRV observés expérimentalement. Ce modèle consiste à assimiler les sauts à des excursions 3D : la protéine quitte l'ADN, diffuse à 3D dans l'espace puis se réassocie sur l'ADN, cela en un temps limité pour rendre compte des conditions expérimentales : si l'enzyme n'est pas revenue sur l'ADN, nous considérons qu'elle est perdue. Ce modèle apporte deux informations importantes :

★ **L'enzyme saute par diffusion 3D** : nous obtenons le bon ordre de grandeur du nombre de sauts observés expérimentalement. En outre, les cumulatives expérimentales et calculées ont la même allure pour  $z \geq 200$  nm. Notre modèle, à un seul paramètre ajustable, rend parfaitement compte de l'expérience. Même si cela peut paraître surprenant, les sauts correspondent à de la diffusion à 3D.

★ **Il y a beaucoup de petits sauts indétectables** : les calculs ont montré qu'il existait aussi des sauts inférieurs à 200 nm, coupure que nous avons fixée à cause des contraintes expérimentales. Ces « petits » sauts ne sont pas détectables dans notre configuration expérimentale mais sont en fait nombreux comme le montre la Figure III.18 : il y a en moyenne par interaction 20,5 sauts pour  $\kappa = 1,5 \text{ nm}^{-1}$  et 27,1 sauts pour  $\kappa = 3 \text{ nm}^{-1}$ . Nous discutons de leur conséquences dans la section suivante. De plus, la Figure III.18 nous indique que les sauts sont pour la plupart inférieurs à 10 nm.



**Figure III.18: Existence de petits sauts non détectables. Cas du Pipes pH=6,8.** Cumulatives inverses normalisées obtenues numériquement (cyan) avec le jeu de paramètres en accord avec les points expérimentaux (bleu) pour EcoRV-Cy3B. L'unité de l'axe des abscisses (taille des sauts) est la paire de base : échelle linéaire entre 0 et 10 pb, puis échelle logarithmique. L'ordonnée à l'origine donne le nombre total  $N=20$  de sauts invisibles.

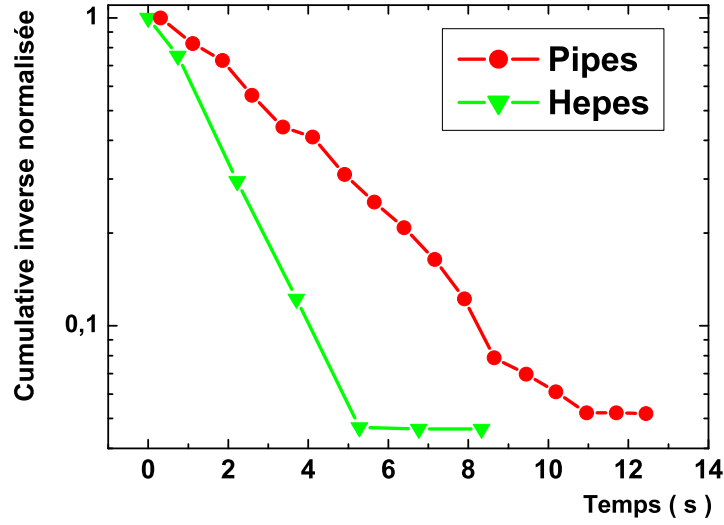
## III.4 Conséquences de l'existence des petits sauts

D'après le modèle, dans le Pipes à pH=6,8 ( $\kappa^{-1} = 1,5$  nm), il y a en moyenne  $N = 20$  sauts inférieurs à 200 nm par interaction. Ce que nous avons appelé événement de sliding n'est donc pas constitué d'une seule association-dissociation, mais de  $N+1$  association-dissociation-réassociation. Ces sauts indétectables influent sur les paramètres que nous avons estimés : temps d'interaction, longueur de sliding, et coefficient de diffusion.

### III.4.1 Temps d'interaction

**Estimation expérimentale du temps d'interaction :** nous avons travaillé à très faible puissance lumineuse d'excitation ( $\mathcal{P}_{\text{Laser}} = 5-10 \text{ W.cm}^{-2}$  au lieu de 100) afin de s'affranchir du phénomène de photoblanchiment. En contrepartie, nous avons dû augmenter le temps d'acquisition (300 ms au lieu de 20) pour pouvoir détecter les enzymes. Pour les différents tampons, nous avons enregistré plus d'une centaine d'événements d'interaction et nous avons tracé les cumulatives des durées des trajectoires en échelle semi-log (cf. Figure III.19).

Les cumulatives des durées en échelle log sont ajustées par des droites. Les temps issus de ces ajustements sont :  $\tau = 4,4 \pm 0,5$  s pour le Pipes et  $\tau = 2,1 \pm 0,4$  s pour le Hepes.



**Figure III.19: Durée de l'interaction EcoRV-Cy3B - ADN non spécifique.** Cumulatives normalisées des durées d'interaction en échelle semi-log.

La durée d'interaction est donc de l'ordre de grandeur de la seconde, du même ordre de grandeur que les durées déduites des expériences de biochimie [23]. Malgré les petits sauts, la distribution des temps reste exponentielle : nous allons montrer que cela s'explique par un calcul simple.

**Temps d'interaction prédit par les calculs :** un mouvement de sliding pur est caractérisé par un temps d'interaction exponentiel. Regardons quel est l'effet des petits sauts sur le temps d'interaction ADN-protéine. Dans la suite, puisque  $D_3 \gg D_1$ , nous considérons le temps passé en 3D négligeable.

Considérons une interaction comprenant en fait  $N$  excursions 3D. Soit  $T_N$  la durée de cette interaction,  $\tau_i$  la durée du  $i$ -ième passage en diffusion 1D :  $T_N = \sum_{i=1}^N \tau_i$ . Soit  $P_N(t)$  la probabilité d'une durée d'interaction  $t$  sachant qu'il y a eu  $N$  passages par la phase 1D. Soit  $\phi_{T_N}(u)$  la transformée de Fourier de  $P_N(t)$ , c'est-à-dire que  $\phi_{T_N}(u) = \int_0^{+\infty} e^{iut} P_N(t) dt$ , et  $\phi_\tau(u)$  la transformée de Fourier de la probabilité de la durée d'une phase 1D. En posant  $P(N)$  la probabilité que l'enzyme passe  $N$  fois par la phase 1D avant de se perdre, il vient :

$$\phi_{T_N}(u) = P(N) (\phi_\tau(u))^N. \quad (\text{III.33})$$

Nous supposons la distribution des temps passés dans la phase 1D est exponentielle de paramètre  $\lambda$  :  $\phi_\tau(u) = \frac{1}{1 - \frac{ju}{\lambda}}$ . Nous supposons que l'enzyme a une probabilité  $p$  de

revenir et  $q = 1 - p$  de se perdre pour un passage dans la phase 3D. Dans le calcul du paragraphe III.3.4.1,  $p = \mathcal{C}(0, t_{obs})$  et  $q = \frac{1}{N_{tot}}$ . Nous avons :  $P(N) = qp^{N-1}$  et :

$$\phi_{T_N}(u) = qp^{N-1} \left( \frac{1}{1 - \frac{ju}{\lambda}} \right)^N. \quad (\text{III.34})$$

Soit  $\phi_T(u)$  la transformée de Fourier de la probabilité de la durée d'interaction  $T$  :

$$\phi_T(u) = \sum_{i=1}^{\infty} \phi_{T_i}(u) = \sum_{i=1}^{\infty} qp^{i-1} \left( \frac{1}{1 - \frac{ju}{\lambda}} \right)^i = \frac{q}{q - \frac{ju}{\lambda}}. \quad (\text{III.35})$$

En prenant la transformée de Fourier inverse, nous trouvons :

$$P(t) = \lambda q e^{-\lambda t} \quad \text{soit} \quad P = \frac{\lambda}{N_{tot}} e^{-\frac{\lambda}{N_{tot}} t}. \quad (\text{III.36})$$

Cela correspond à une distribution exponentielle des temps d'interaction, de paramètre  $\lambda/N_{tot}$ , ce qui a été observé expérimentalement. Nous avons mesuré  $T = 4,4$  s (cf. Figure III.19), donc une phase de sliding pure dure en moyenne :  $\tau = \frac{1}{\lambda} = \frac{T}{N} = 220$  ms.

### III.4.2 Longueur de sliding

La présence des  $N$  sauts augmente *a priori* la longueur apparente de sliding. Sans les considérer, nous avons mesuré un temps d'interaction de 4,4 s et un coefficient de diffusion  $D_1 \sim 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ , c'est-à-dire une longueur apparente de sliding  $\ell_1 \sim 900$  pb, bien plus grande que les prédictions des mesures d'ensemble, inférieure à 100 bp [17]. Cependant, si l'on prend en compte ces  $N = 20$  sauts, la durée réelle du sliding devient  $4,4/20 \sim 220$  ms, et la distance parcourue par sliding devient  $\ell_1 \sim 190$  pb.

**Comparaison avec les études précédentes** Cette valeur de 200 pb est similaire à celle proposée par Coppey *et al* pour EcoRV dans leur approche stochastique (260 pb). Cette approche consiste à modéliser le mécanisme de recherche de cible par les protéines par des phases d'interaction non spécifique ADN-protéine entrecoupées d'excursions 3D. Ces excursions 3D sont modélisées par des sauts aléatoires non corrélés [26].

En analysant la processivité d'EcoRV en fonction de la distance inter-site, Stanford *et al* propose une distance plus petite pour la longueur de sliding : 50 pb [17]. L'écart avec notre estimation peut s'expliquer par le fait que nous travaillons à plus basse force ionique.



### III.4.3 Coefficient de diffusion 1D

Evaluons l'effet des petits sauts sur le coefficient de diffusion 1D apparent noté  $D_a$ . Si il y a  $N$  petits sauts durant un événement d'interaction de durée  $t_{\text{int}}$ , il y a  $(N + 1)$  événements de sliding de distance moyenne  $\ell_1$  et de durée moyenne  $\tau_1$ , et  $n$  sauts de durée moyenne  $\tau_3$  et de longueur moyenne  $\ell_3$ . La distance  $\mathcal{L}$  parcourue par l'enzyme durant le temps  $t$  est par conséquent la somme des distances parcourues durant toute les phases de sliding et de jumping :

$$\mathcal{L} = \sum_{i=1}^{N+1} \ell_{1,i} + \sum_{i=1}^N \ell_{3,i}. \quad (\text{III.37})$$

En supposant que toutes ces phases sont indépendantes, il vient :

$$\langle \mathcal{L}^2(t_{\text{int}}) \rangle = (N + 1) \langle \ell_1^2 \rangle + N \langle \ell_3^2 \rangle \quad (\text{III.38})$$

$$\approx N (\langle \ell_1^2 \rangle + \langle \ell_3^2 \rangle) \quad \text{pour } N \gg 1 \quad (\text{III.39})$$

$$\approx \frac{t_{\text{int}}}{\tau_1} (\langle \ell_1^2 \rangle + \langle \ell_3^2 \rangle) \quad \text{puisque } \tau_3 \ll \tau_1 \text{ donc } N \approx \frac{t_{\text{int}}}{\tau_1} \quad (\text{III.40})$$

$$\approx 2D_1 t_{\text{int}} \left( 1 + \frac{\langle \ell_3^2 \rangle}{\langle \ell_1^2 \rangle} \right). \quad (\text{III.41})$$

Or  $\langle \mathcal{L}^2(t_{\text{int}}) \rangle = 2D_a t_{\text{int}}$  donc :  $D_a \approx D_1 \left( 1 + \frac{\langle \ell_3^2 \rangle}{\langle \ell_1^2 \rangle} \right)$ .

Traisons le cas du Pipes pH=6,8. D'après les cumulatives analytiques, il y a environ  $N = 20$  sauts inférieurs à 200 nm par interaction. En outre :

- La distance moyenne parcourue durant ces 20 sauts est donnée par :

$$\sqrt{\langle \ell_3^2 \rangle} = \sqrt{\int_0^{200 \text{ nm}} z^2 \mathcal{P}(z) dz} \approx 11,4 \text{ nm}.$$

La durée d'un de ces sauts est d'environ  $\tau_3 = \frac{\langle \ell_3^2 \rangle}{D_3} \approx 3 \mu\text{s}$ , beaucoup plus petit que la durée apparente de sliding.

- La durée moyenne d'interaction est  $t_{\text{int}} = 4,4 \text{ s}$ , le temps moyen de sliding est donc de  $\tau_1 = t_{\text{int}}/N = 220 \text{ ms}$ , soit  $\ell_1 \sim 190 \text{ pb} = 60 \text{ nm}$ .

Ce qui donne  $\frac{D_a - D_1}{D_1} \approx 0,04$  : l'effet des petits sauts est négligeable dans notre étude.

### III.4.4 Prévisions - Réalité in vivo

La conformation en pelote de l'ADN dans les bactéries est une configuration qui favorise certainement le jumping. En effet, après s'être dissociée de l'ADN, une enzyme s'y réassocie plus rapidement quand l'ADN est compacté. *In vivo*, le jumping joue sûrement

un rôle plus important que celui que nous avons observé sur l'ADN étiré, du moins pour les grands sauts. En revanche, à cause de la longueur de persistance de l'ADN (50 nm en conditions standards) les conformations ADN étiré - ADN en pelote ne doivent pas influencer sur le nombre de petits sauts.

## III.5 Effet de la taille du fluorophore

### III.5.1 Théorie - modèle de Schurr :

D'après Schurr, le coefficient de diffusion d'une protéine sur l'ADN est vraisemblablement régi, au moins en partie, par des paramètres hydrodynamiques (les éventuels frottements protéines-ADN sont négligés) [4]. Plus précisément, Schurr suppose tout d'abord que le sliding des protéines est une diffusion unidimensionnelle dans un mouvement hélicoïdal selon le grand sillon de l'ADN. De plus, il admet que les protéines conservent la même orientation par rapport à la double-hélice d'ADN tout au long du mouvement. Le coefficient de diffusion de la protéine le long de l'ADN est alors inversement proportionnel à un coefficient de friction  $f_{\text{eff}}$  :  $D_1 = k_B T / f_{\text{eff}}$  avec :

$$f_{\text{eff}} = 6\pi\eta R_H \left[ 1 + \frac{4}{3} (2\pi)^2 \left( \frac{R_H(\text{nm})}{h} \right)^2 \right] \quad (\text{III.42})$$

où  $\eta$  est la viscosité de l'eau ( $10^{-3}$  Pa.s $^{-1}$ ) et où  $h = 3,4$  nm est la longueur correspondant à un tour dans une hélice d'ADN double-brin. Le premier terme correspond à la friction associée à la composante translationnelle du mouvement, le second traduit la friction causée par la composante rotationnelle.

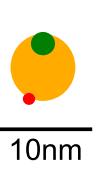
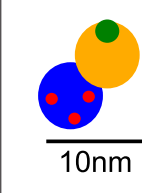
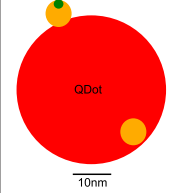
Pour des objets de rayon hydrodynamique supérieur au nanomètre, le terme de translation est négligeable devant le terme rotationnel, il vient :

$$f_{\text{eff}} \simeq \frac{32\pi^3\eta R_H^3}{(3,4 \text{ nm})^2}. \quad (\text{III.43})$$

Si  $f_{\text{eff}}$  varie avec le rayon hydrodynamique  $R_H$  du complexe protéine-fluorophore, le coefficient de diffusion du complexe doit également dépendre de la taille du fluorophore et varier selon  $R_H^{-3}$ .

### III.5.2 Résultats expérimentaux :

Nous avons répété les expériences avec les autres couplages dont nous disposons (EcoRV - Sav-Cy3 et EcoRV-QD605) pour voir quelle était l'influence de la taille du fluorophore sur la diffusion facilitée d'EcoRV. Le Tableau III.4 reprend les caractéristiques des différents complexes enzyme-fluorophore à notre disposition. Nous avons choisi de faire ces expériences dans le Pipes pH=6,8, dans lequel les interactions non-spécifiques entre la streptavidine et la surface sont réduites.

	EcoRV - Cy3B	EcoRV - sav - Cy3	EcoRV - QD605
$R_H$ (nm)	4	5,3	16
$D_3$ ( $\mu\text{m}^2.\text{s}^{-1}$ )	54	40	14
Représentation			

**Tableau III.4:** Rayons et coefficients de diffusions des complexes enzyme-fluorophore déterminés par FCS.

### III.5.2.1 Étude de EcoRV - Sav-Cy3

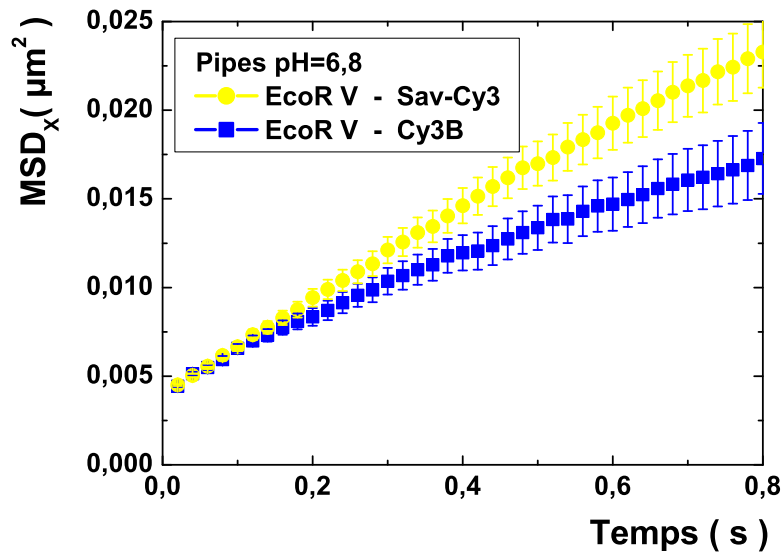
► La Figure III.20 représente les MSD poolées des composantes longitudinales d'EcoRV interagissant avec l'ADN : les carrés bleus correspondent à l'enzyme marquée au petit fluorophore Cy3B ( $MW \# 1 \text{ kDa}$ ) et les ronds jaunes sont les points obtenus avec l'enzyme marquée à Sav-Cy3 ( $MW \# 60 \text{ kDa}$ ). Les 10 premiers points de ces MSDs sont superposés. L'ajustement sur les 5 premiers points donne  $D_1^{\text{Sav-Cy3}} \approx 1,3.10^{-2} \mu\text{m}^2.\text{s}^{-1}$  contre  $D_1^{\text{Cy3B}} \approx 1,2.10^{-2} \mu\text{m}^2.\text{s}^{-1}$  : il n'y a donc pas *a priori* de changement des caractéristiques de la diffusion 1D de la protéine le long de l'ADN pour ces tailles de fluorophores.

En revanche, aux temps longs, les MSDs divergent : la MSD pour l'enzyme EcoRV marquée à Sav-Cy3 croît plus rapidement. Le mouvement d'EcoRV-Cy3B est confiné aux temps longs. L'explication pourrait venir du fait que ces enzymes ne sont marquées que dans un ratio 1 :3, des enzymes libres s'accumulent certainement sur l'ADN, limitant ainsi le mouvement des enzymes marquées.

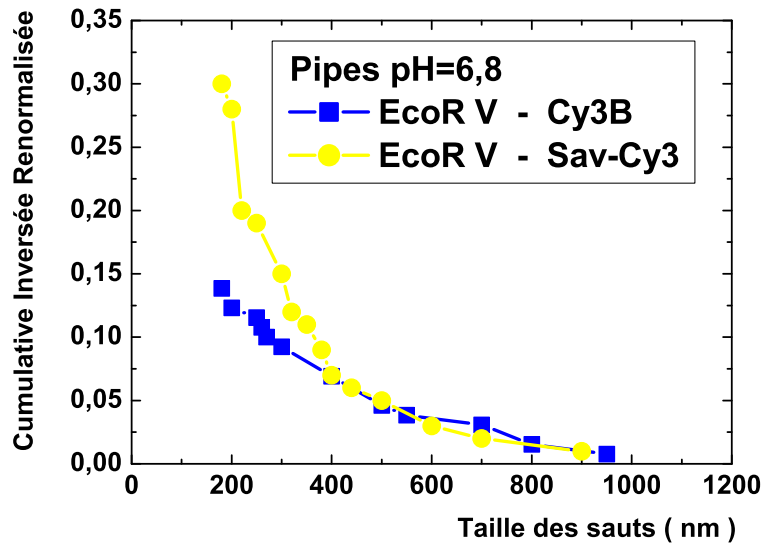
Quelques expériences avec les enzymes marquées à Sav-Cy3 ont été faites dans le Hepes pH=7,5 : les conclusions ont été les mêmes.

► Nous avons ensuite comparé les distributions des sauts : la Figure III.21 montre les cumulatives des sauts pour EcoRV-Cy3B et pour EcoRV-Sav-Cy3, dans le Pipes à pH=6,8. La différence entre ces deux cumulatives réside dans le nombre moyen de sauts par interaction, puisque la fréquence des sauts  $\geq 200 \text{ nm}$  est 2 fois plus élevée pour l'enzyme marquée à Sav-Cy3 par rapport à l'enzyme marquée à Cy3B.

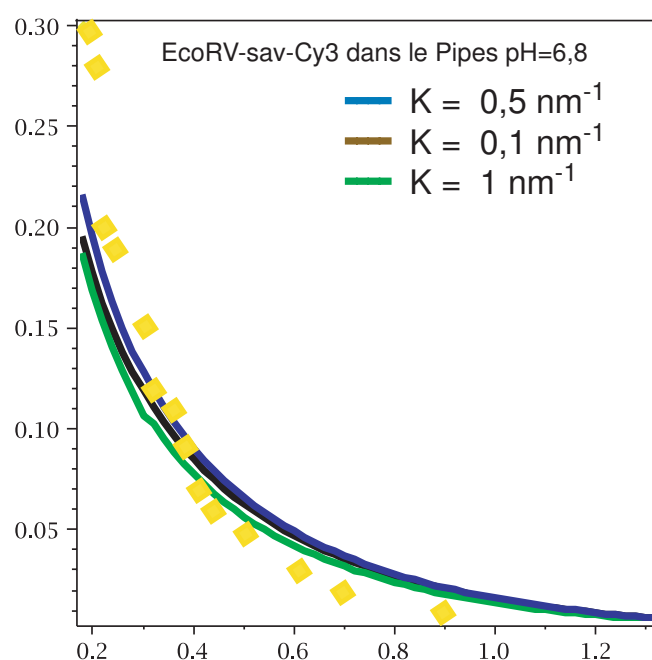
Nous avons tenté d'ajuster la cumulative expérimentale des sauts d'EcoRV - Sav-Cy3 avec les calculs en prenant les valeurs :  $a = 6,3 \text{ nm}$  et  $D_3 = 40 \mu\text{m}^2.\text{s}^{-1}$ . L'accord n'est pas bon comme le montre la Figure III.22.



**Figure III.20:** Comparaison des MSDs d'ensemble pour EcoRV - Sav-Cy3 (ronds jaunes) et EcoRV - Cy3B (carrés bleus). Les 10 premiers points se confondent aux incertitudes expérimentales près : la taille du fluorophore ne change pas le coefficient 1D mesuré.



**Figure III.21:** Cumulative inversée et renormalisée des sauts observés pour l'enzyme marquée au Cy3B (carrés bleus) et l'enzyme marquée à Sav-Cy3 (ronds jaunes) dans des Pipes, pH=6,8.



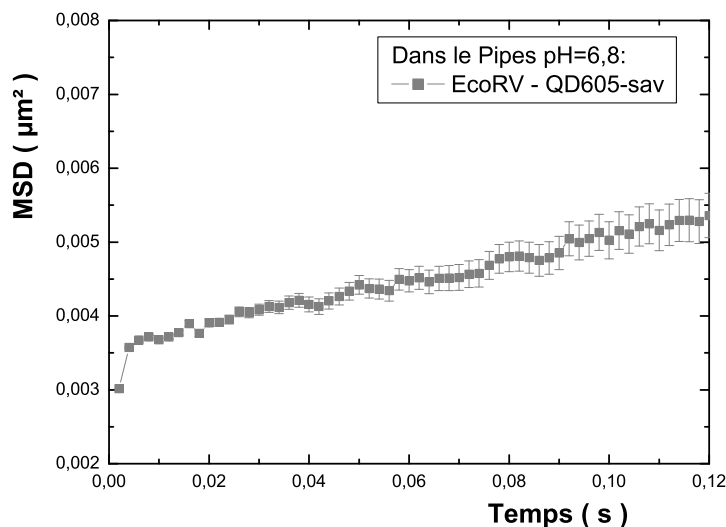
**Figure III.22: Cumulative des sauts pour EcoRV- Sav-Cy3 dans le Pipes, pH=6,8.** Points expérimentaux (jaune) et calculs ( $\kappa^{-1} = 1$  nm (vert),  $\kappa^{-1} = 2$  nm (bleu),  $\kappa^{-1} = 10$  nm (noir)).

### III.5.2.2 Étude de EcoRV - QD605-sav

Quelques expériences d'interaction EcoRV - ADN ont également été menées avec des enzymes couplées à des nanocristaux semi-conducteurs QD605-Sav. Les expériences ont aussi été faites dans le Pipes à pH=6,8, dans les mêmes conditions expérimentales que précédemment, à ceci près que le temps d'acquisition a été descendu à 2 ms, étant donné que les QDs sont des sondes dix fois plus brillantes que les fluorophores organiques.

Globalement, ces expériences se sont révélées plus délicates que les autres : à la concentration de travail permettant des observations en molécule unique, les événements sont moins fréquents qu'avec les autres fluorophores. Nous pouvons proposer une explication : le fait qu'il n'y ait statistiquement qu'une seule enzyme par nanocristal et que le nanocristal soit environ 4 fois plus gros que l'enzyme limite les chances pour que la protéine soit bien située par rapport à l'ADN lors des rencontres ADN-{nanocristal+enzyme}. L'autre difficulté que nous avons rencontrée lors de ces expériences est qu'il n'est pas rare que les nanocristaux se fixent à l'ADN sans pour autant y diffuser à 1D.

Le nombre d'événement de sliding enregistrés (une vingtaine) n'est actuellement pas suffisant pour faire des statistiques significatives : ces expériences doivent être poursui-



**Figure III.23:** MSD poolée pour l'enzyme marquée aux nanocristaux. La taille des QDs modifie le coefficient de diffusion de l'enzyme puisque  $D_1 \sim 0,7.10^{-2} \mu\text{m}^2.\text{s}^{-1}$ .

vies. Néanmoins, nous disposons de données suffisantes pour pooler les trajectoires. La Figure III.23 représente la MSD poolée : la taille des nanocristaux modifie le coefficient de diffusion puisque la pente de la MSD poolée en prenant en compte le taux d'étirement de l'ADN est de  $\sim 1,3.10^{-2} \mu\text{m}^2.\text{s}^{-1}$ , ce qui donne un coefficient de diffusion  $D_1 \sim 0,7.10^{-2} \mu\text{m}^2.\text{s}^{-1}$ . Cette valeur est en accord avec les résultats obtenus par A. Crut durant sa thèse [68].

Nous n'avons pas observé suffisamment de saut pour tracer une cumulative.

### III.5.2.3 Conclusion

Dans l'hypothèse où la diffusion est en grande partie régie par l'hydrodynamique, le coefficient de diffusion devrait changer en fonction de la taille du complexe enzyme-fluorophore. Or ce n'est pas notre observation : il reste à comprendre pourquoi nous observons plus de sauts que prévus par la diffusion 3D. Une explication pourrait venir de l'existence des petits sauts : comme le laisse présager la cumulative de la Figure III.22, nous pouvons supposer qu'il y aura encore plus de petits sauts pour EcoRV-Sav-Cy3 que pour EcoRV-Cy3B. Dans ce cas, l'effet des petits sauts sur le coefficient de diffusion ne sera plus négligeable : le  $D_1$  que nous avons estimé pour EcoRV - Sav-Cy3 sera moins précis. Si cette explication est la bonne, le fait que nous mesurons le même coefficient de diffusion pour les deux complexes est une joyeuse coïncidence !

Afin de comprendre l'effet de la taille, des expériences complémentaires sont en cours.



# Conclusion et Perspectives

Le travail présenté dans ce mémoire avait pour objectif d'apporter des éléments de réponse à une importante question en biologie : comment les protéines trouvent-elles leur cible sur l'ADN ? Pour cela, nous avons choisi d'utiliser les techniques de microscopie de fluorescence pour visualiser directement l'interaction ADN-protéine à l'échelle de la molécule individuelle. Nous avons mené cette étude avec l'enzyme de restriction EcoRV, qui a fait l'objet de nombreuses études biochimiques, mais pas à l'échelle de la molécule unique. En fait, nous nous sommes limités aux interactions non spécifiques ADN - protéines, interactions décisives dans la recherche de cible spécifique et qualifiées de « diffusion facilitée ».

Nous avons réussi à visualiser l'interaction d'une unique enzyme EcoRV avec de l'ADN non spécifique : ce mémoire décrit sa mise en œuvre expérimentale. Même si les études des interactions ADN-protéine à l'échelle de la molécule unique connaissent un intérêt croissant depuis plusieurs années, une telle visualisation reste délicate, et n'est offerte qu'à peu d'équipes dans le monde puisqu'elle nécessite d'allier les techniques de micromanipulation de molécules d'ADN aux sensibilités de détection de molécules fluorescentes individuelles.

Après avoir détaillé les deux grandes étapes techniques indispensables : le couplage des enzymes à des fluorophores d'une part et le positionnement des molécules d'ADN d'autre part, nous avons présenté et analysé nos résultats expérimentaux. Le résultat majeur de ce travail est l'observation de la « diffusion facilitée » proprement dite d'EcoRV : durant l'interaction avec l'ADN, l'enzyme alterne entre des phases de sliding (diffusion 1D le long de l'ADN) et du jumping (sauts 3D). Nous sommes parvenus à mesurer les paramètres physiques du sliding : coefficient de diffusion 1D ( $10^{-2} \mu\text{m}^2.\text{s}^{-1}$ ), longueur ( $\sim 200$  pb). De plus, nous avons relié les sauts observés à des phénomènes de diffusion 3D et un modèle simple a permis de retrouver numériquement la fréquence ( $\sim 15$  %) et la taille des sauts observés. Ces résultats sont originaux puisque nous sommes les premiers à reporter l'observation, l'analyse et la discussion de l'existence de tels sauts dans les mécanismes de diffusion facilitée.

Pour analyser les données, nous nous sommes efforcés de trouver les meilleurs outils, cela nous a conduit à développer des approches nouvelles adaptées à notre situation. Dans



l'étude du sliding, la méthode du pooling permet de regrouper les données de plusieurs trajectoires et de les traiter globalement. Les trajectoires « courtes » peuvent être désormais incluses dans les calculs des MSDs pour déterminer un coefficient de diffusion. Pour analyser les sauts, nous avons choisi d'étudier les cumulatives inverses, directement comparables aux probabilités de longueur par saut, grandeurs accessibles numériquement.

Pendant la rédaction de ce manuscrit, l'équipe a obtenu des résultats complémentaires confirmant les premières observations. Des expériences ont été faites dans d'autres tampons à pH=7,5 (Tris et PB). En outre, pour écarter l'éventuelle influence de la charge des différents marqueurs utilisés (tous négativement chargés aux pHs étudiés), des expériences ont été menées en ayant couplé l'enzyme à un fluorophore positivement chargé : Atto 647N ( $\lambda_{abs} = 635$  nm,  $\lambda_{em} = 659$  nm, *Atto Technology*, Allemagne). Les conclusions sont restées inchangées :

- le coefficient de diffusion  $D_1$  de l'enzyme le long de l'ADN ne change pas.
- en revanche, l'occurrence ainsi que la taille des sauts 3D dépend du milieu d'interaction.

Notre dispositif est opérationnel, fiable et précis (résolution spatiale :  $\sim 30$  nm, résolution temporelle : 20 ms) et les résultats obtenus (expérimentaux et numériques) encouragent fortement à poursuivre les expériences.

★ Une suite immédiate consisterait à finaliser l'étude de EcoRV-Cy3B en caractérisant totalement l'influence des conditions expérimentales sur les paramètres de la diffusion facilitée : force ionique, concentration en enzyme, absence de  $Mg^{2+}$ , taux d'étirement de la molécule, température. Des paramètres comme le taux d'association pourraient ainsi être mesurés. Ces expériences nécessitent de collecter une quantité de données mais méritent d'être faites. Plus tard, l'insertion d'une séquence cible sur l'ADN permettra d'analyser l'interaction spécifique.

★ Ensuite, il serait très intéressant d'identifier clairement le rôle que joue l'hydrodynamique dans la diffusion de l'enzyme. Réitérer les expériences avec les autres couplages (EcoRV-Sav-Cy3 et EcoRV-QD605) apporterait les premiers éléments de réponse. Dans l'hypothèse où l'hydrodynamique régit la diffusion 1D de la protéine,  $D_1$  dépend non seulement du rayon hydrodynamique mais aussi de la viscosité du milieu : changer la viscosité en travaillant dans une solution de sucre devrait changer le coefficient de diffusion.

★ À plus long terme, il apparaît essentiel de savoir si les résultats établis sont propres à l'enzyme EcoRV ou non. Notre système expérimental autorise l'utilisation d'autres protéines : la distribution des sauts est-elle spécifique à l'enzyme ? Les outils d'analyses dont nous disposons sont suffisamment solides, le dispositif expérimental relativement flexible, il ne devrait pas être difficile de caractériser la dynamique d'interaction d'autres

protéines avec l'ADN, et d'envisager l'étude des protéines processives. En ce moment, Pierre-Louis Porté, doctorant de l'équipe, mène des expériences similaires mais avec une autre enzyme : la DNA méthylTransférase MTaqI. Cette enzyme appartient au système de restriction-méthylation de la bactérie *Thermus aquaticus*, et protège l'ADN de l'action de ses propres enzymes de restriction en méthylant les adénines des sites de coupures. Les methyl-transférases jouent un rôle important dans la régulation de l'expression génique. Nous bénéficions de MTaqI spécifiquement marquées en C-terminal, loin du site spécifique (Elmar Weinhold, Aachen, Allemagne). Le mécanisme qui permet à une méthyltransférase de trouver sa cible n'a pas fait l'objet d'études particulières, contrairement au cas des enzymes de restriction. Nous abordons donc un sujet très intéressant et complètement ouvert.

★ À encore plus long terme, adapter l'expérience *in vivo* (directement dans la bactérie) serait certainement très instructif.

Cette liste est loin d'être exhaustive : les possibilités sont nombreuses et ouvertes.



# Annexe A

## Fabrication d'un ADN de 9,2 kb aux extrémités biotinilées ne comportant aucun site spécifique pour EcoRV.

### A.1 Principe

L'idée consiste à générer de courts fragments biotinilés que l'on ligue aux extrémités d'une molécule d'ADN d'intérêt. En l'occurrence, nous avons choisi un ADN linéaire ne comportant pas le site de restriction pour l'enzyme EcoRV et dont la taille est comprise entre 7 et 10 kb. Pour générer les extrémités biotinilées, nous avons amplifié par PCR des fragments de 0,5 kbp du plasmide pBluescriptSK+ en présence de nucléotides biotinilés. Ces produits de PCR sont ensuite digérés de manière à produire des extrémités à bout collant compatibles avec celles de l'ADN linéarisé. Pour avoir des extrémités compatibles, il suffit de travailler avec des enzymes qui laissent des nucléotides sortants et qui ne reconnaissent pas des séquences palindromiques.

### A.2 Choix d'un plasmide support : le T7

Nous avons utilisé l'ADN du bactériophage T7 (*Biocentric*), long de 39,9 kb et qui ne comporte aucun site de reconnaissance pour l'enzyme EcoRV. Pour obtenir un ADN d'une dizaine de kb, auquel nous pourrions liguer des extrémités biotinilées, le T7 est digéré par une enzyme de restriction non palindromique générant des bouts collants : nous avons utilisé l'enzyme BsmBI. La digestion du T7 par BsmBI produit 17 fragments, dont le plus long fait 8,2 kb. Ce fragment est récupéré après électrophorèse sur gel d'agarose et extraction (Gel Extraction Kit, *Qiagen*) puis précipitation à l'éthanol et extraction au phénol/chloroforme. Au final, on dispose de fragments de T7 long de 8,2 kb, ne comportant pas de séquence-cible pour EcoRV, et de concentration de quelques centaines de ng/ $\mu$ L.

### A.3 Fabrication des fragments biotinilés

Des amorces de PCR ont été conçues (*Eurogentec*, cf. tableau A.1) pour amplifier une séquence de 0,5 kpb du plasmide pBluescriptSK+ (*Stratagene*). Elles ont été fabriquées avec un site de reconnaissance pour l'enzyme de restriction BsaI. BsaI est une enzyme de classe IIS qui possède un site de reconnaissance (en gras) et un site de coupure (en gras italique) distincts : nous avons ainsi pu choisir les séquences des bouts collants laissés par BsaI pour qu'elles puissent être compatibles à celles du fragment de 9,2 kb de T7 digéré par BsmBI.

Amorces	Séquences
FwC6-T7	5'-CGCTTGGTCTCT <b>CGGAC</b> GGTATCAGCTCACTCAAAG -3'
FwC7-T7	5'-CGCTTGGTCTCT <b>GTTG</b> CGGTATCAGCTCACTCAAAG -3'
Rv	5'-GGCGATAAGTCGTGTCTTAC-3'

**Tableau A.1:** Amorces de PCR utilisées pour produire des ADNs de 0,5 kpb biotinilés.

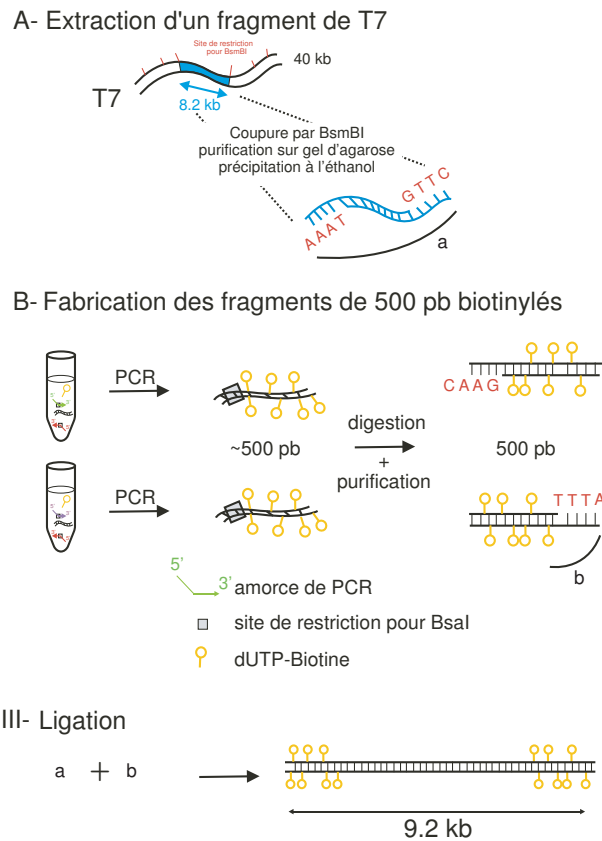
La PCR est réalisée dans le tampon Taq DNA polymérase (*New England's Biolabs*) avec 2 mM de magnésium, en présence de 8 pg/ $\mu$ L de pBluescriptSK+. Le rapport entre la concentration en dUTP-modifié biotine (*Roche*) et la concentration totale en dNTP (*Fermentas*) est de 2/3 (cf. tableau A.2) pour obtenir en moyenne 70 biotines incorporées par fragment. La réaction de PCR est effectuée dans un thermocycleur (*Eppendorf*). Le cycle suivant a été répété 30 fois :

- 30 secondes à 94°C (dénaturation)
- 30 secondes à 61°C (hybridation)
- 60 secondes à 72°C puis 10s/cycle (élongation)

Tampon Taq	H <sub>2</sub> O	primer fw	primer rw	dNTP mix	dUTP-biot	pBS	Taq
10x		10 $\mu$ M	10 $\mu$ M	2 mM	1 mM	0,2 ng/ $\mu$ L	5U/ $\mu$ L
5	23	8	8	1,25	0,84	2	1

**Tableau A.2:** Conditions expérimentales de la PCR pour la fabrication des extrémités biotinilées. Les volumes sont exprimés en  $\mu$ L.

Les produits de PCR sont purifiés (PCR Purification Kit, *Qiagen*) et digérés par Bsa I (*New England's Biolabs*). Après digestion, les produits de digestion des fragments sont éliminés en utilisant ce même kit de purification, reprécipités à l'éthanol, et enfin resuspendus dans du Tris-EDTA. La concentration finale des produits de PCR est typiquement



**Figure A.1:** Principe de la fabrication du fragment de T7 biotinylé aux extrémités

25 ng/ $\mu$ L (mesure de l'OD 260 nm au spectrophotomètre (Nanodrop Nd-1000, *Nanodrop*) et estimation par comparaison avec des quantités étalons déposées sur un gel d'agarose).

## A.4 Ligation

Les fragments biotinylés sont ligués (dans un ratio 2 :1 molaire puisqu'il s'agit de bouts collants) au fragment de T7 en utilisant la T4 DNA ligase (*New England's Biolabs*) en large excès, la ligation a lieu durant une nuit à 4°C. Les molécules d'ADN liguées sont finalement purifiées en utilisant le même kit que celui utilisé après la PCR. Généralement, à la fin de ce protocole, on récupère des ADNs aux extrémités biotinylées dont la concentration est de l'ordre de 10-15 ng/ $\mu$ L. La mesure du rapport des absorbances à 260 nm et 280 nm est proche de 1,8, ce qui est très satisfaisant.



# Annexe B

## Programme de tracking

Ce programme a été écrit en MatLab (*The Mathworks*) par Pierre-Louis Porté. Il a été conçu pour suivre un fluorophore unique susceptible de diffuser à une dimension. Il permet de déterminer la trajectoire du fluorophore puis d'analyser son mouvement.

### B.1 Répérer un spot fluorescent

La taille des fluorophores utilisés (de l'ordre du nm) est largement inférieure à la longueur d'onde qu'ils émettent (de l'ordre d'une centaine de nm), par conséquent, ils peuvent être considérés comme des sources ponctuelles. La réponse impulsionnelle théorique ( $PSF_B$ ) du système optique est une fonction de Bessel dont le premier zéro définit le rayon d'Airy  $R_A$ <sup>1</sup> :

$$R_A \approx \frac{1,22 \lambda}{2 NA}, \quad (\text{B.1})$$

où  $NA$  est l'ouverture numérique du système ( $NA = 1,45$  pour l'objectif  $60\times$  à immersion à huile que nous utilisons) et  $\lambda$  désigne la longueur d'onde d'émission du fluorophore. Pour simplifier les traitements numériques, cette  $PSF_B$  peut-être raisonnablement modélisée (Figure B.1) par une fonction gaussienne 2D isotrope de largeur  $\sigma_G = R_A/3$  [89] :

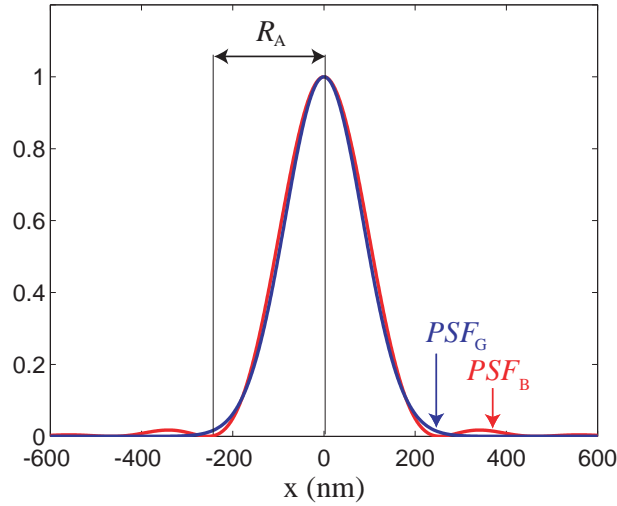
$$PSF_G(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma_G^2}\right) \quad \text{avec} \quad \sigma_G = \frac{R_A}{3}. \quad (\text{B.2})$$

L'image idéale en sortie du microscope est numérisée par la caméra CCD dont l'ensemble des cellules photosensibles constitue une grille sur laquelle l'image est échantillonnée spatialement. L'intensité en chaque pixel étant codée sur 14 bits, les effets dus à la discrétisation de ces valeurs après numérisation sont considérés comme négligeables. A

---

<sup>1</sup>Si l'on prend le critère de Rayleigh pour définir le pouvoir séparateur de notre système optique, alors celui-ci est égal à  $R_A$





**Figure B.1:** Profil normalisé de la réponse impulsionnelle du système optique : modèle théorique  $PSF_B$  (rouge pour  $NA = 1,45$  et  $\lambda = 600$  nm) et approximation gaussienne  $PSF_G$  (bleu) avec  $\sigma_G = R_A/3$

ce signal vient s'ajouter un offset constant  $L_B$  appelé ligne de base, et systématiquement ajouté au signal de fluorescence par la caméra CCD. Dans toute la suite, on appellera spot le signal d'un fluorophore sur une image. Notre méthode pour détecter les spots consiste à sélectionner une zone de recherche dans laquelle une régression au sens des moindres carrés est effectuée en utilisant la fonction bidimensionnelle suivante :

$$I(x, y) = L_B + A \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma_G^2}\right). \quad (\text{B.3})$$

Suivre un spot fluorescent au cours du temps à partir d'un film revient ainsi à déterminer les valeurs de  $I(x, y)$  pour chaque image du film, ce qui revient à déterminer les variables de la gaussienne 2D donnée à l'équation B.3. Il s'agit d'un ajustement à 5 paramètres : la largeur de la gaussienne  $\sigma_G$ , l'amplitude  $A$  et aussi l'abscisse  $x_0$ , l'ordonnée  $y_0$  et la ligne de base  $L_B$ . C'est à l'utilisateur de choisir le paramètre initial  $\sigma_G$ .

## B.2 Description de l'interface

Le programme écrit par P.-L. Porté est une interface graphique qui calcule les paramètres physiques liés à la diffusion 1D d'un fluorophore le long d'un axe de référence que l'on choisit : temps d'interaction, MSD transverse et longitudinale, coefficient de diffusion transverse et longitudinal.

Dans notre cas, l'axe de référence d'intérêt est celui d'une molécule d'ADN étirée. L'utilisateur doit charger une image en fluorescence de la molécule d'ADN étirée, qui apparaît dans une première fenêtre. L'axe de référence est défini en marquant les deux

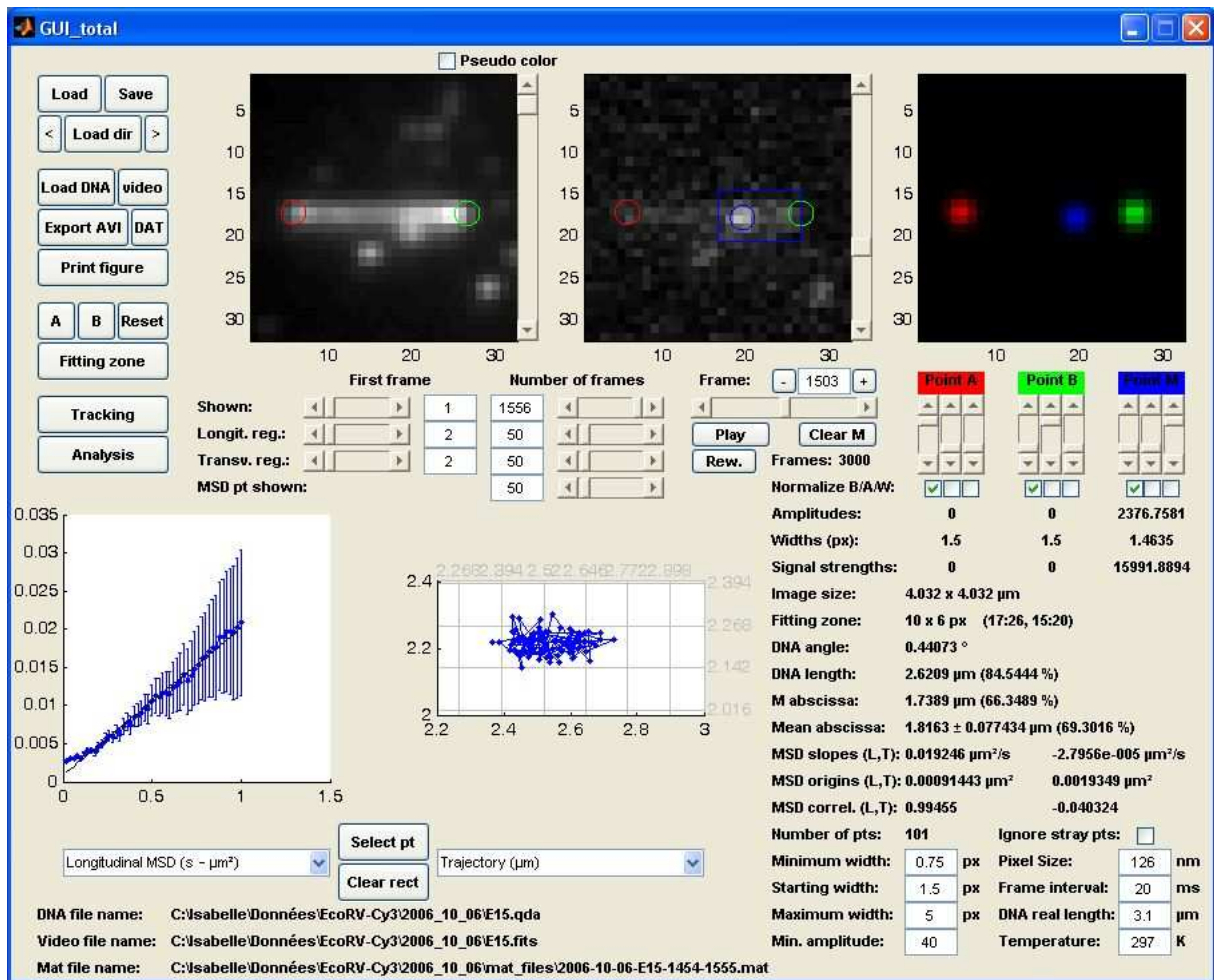


Figure B.2: Interface graphique utilisée pour analyser les vidéos.

extrémités de la molécule qui apparaissent en vert et rouge sur la figure B.2. La droite passant par ces deux points définit l'axe longitudinal de la molécule d'ADN. L'axe transverse lui est orthogonal. Si l'axe n'est pas défini, le programme prend l'axe  $x$  comme référence longitudinale, et l'axe  $y$  comme référence transverse.

Le film à analyser est ensuite chargé dans une deuxième fenêtre. Quand l'utilisateur veut suivre un spot fluorescent, il délimite d'abord une zone de recherche (« fitting zone », repérée par un rectangle bleu) : le programme détermine alors dans la zone choisie la  $PSF_G$  gaussienne à 2D la plus adéquate pour chaque image. Le centre théorique  $(x_0, y_0)$  est ainsi déterminé, donnant les coordonnées de chaque point au cours du temps, la précision atteinte est de  $\sim 35$  nm. L'utilisateur décide d'arrêter le tracking quand il le souhaite. Pendant le tracking, apparaît en temps-réel, dans une troisième fenêtre, la trajectoire reconstruite dont chaque point correspond au centre de la gaussienne 2D déterminée. Lorsque l'utilisateur décide d'arrêter le tracking, il peut au choix décider d'afficher :

- la trajectoire  $(x, y)$  du spot
- l'abscisse  $x(t)$  du spot
- l'ordonnée  $y(t)$  du spot
- l'amplitude  $A(t)$  de la gaussienne, i.e. l'intensité au centre du spot
- la largeur de la gaussienne  $\sigma_G$
- la base  $L_B(t)$

En fixant des critères tels qu'un seuil de détection minimal ou une corrélation spatiale, l'utilisateur peut enlever les spots détectés qui pourraient être des artéfacts. Une fois la trajectoire reconstruite, l'analyse peut alors être lancée : calcul des MSDs 1D transverse et longitudinale et régression linéaire de ces MSDs en choisissant le nombre de points sur lequel l'ajustement doit être fait. S'affichent alors la MSD calculée, avec les barres d'erreur calculées selon l'expression donnée à l'équation III.18 et le fit linéaire dont on déduit les coefficients de diffusion.

# Annexe C

## Fluctuations de l'ADN étiré

Nous considérons une molécule d'ADN de longueur de contour  $L$  et de longueur étirée  $R$ . Les extrémités de la molécule sont supposées fixes, par exemple attachées à une surface. La longueur de persistance de l'ADN est notée  $L_p$ . Nous allons estimer l'amplitude et le temps de corrélation associés aux fluctuations Browniennes transverses et longitudinales de la molécule.

### C.1 Modèle billes-ressorts (modèle de Rouse)

La molécule d'ADN est divisée en  $N - 1$  segments de longueur  $L_0$  telle que  $L \gg L_0 \gg L_p$ . Aux extrémités de ces segments sont placées des billes, repérées par un indice  $n$  et leur position  $\mathbf{r}_n$ . On a donc  $N$  billes ( $1 \leq n \leq N$ ) disposées régulièrement le long de l'ADN, les billes  $n = 1$  et  $n = N$  étant immobiles. Les segments d'ADN reliant les billes sont assimilés à des ressorts de raideur  $k$ , qui exercent sur les billes une force de tension  $\mathbf{T}$  de module  $T(r) = kr$ . Le mouvement d'une bille d'indice  $n$  est régi par l'équation du mouvement [98] :

$$\zeta \frac{\partial \mathbf{r}_n}{\partial t} = -k(\mathbf{r}_n - \mathbf{r}_{n-1}) + k(\mathbf{r}_{n+1} - \mathbf{r}_n) + \mathbf{L}_n,$$

où  $\zeta$  est un coefficient de frottement et  $\mathbf{L}_n$  la force de Langevin qui s'exerce sur la bille dont les coordonnées  $L_{n,\alpha}$  et  $L_{m,\beta}$  dans le plan vérifient :

$$\langle L_{n,\alpha}(t)L_{m,\beta}(t') \rangle = 2\zeta k_B T \delta(t - t') \delta_{\alpha\beta} \delta_{nm}, \quad (\text{C.1})$$

où  $\alpha$  et  $\beta$  désignent les deux composantes de  $\mathbf{L}_n$  et  $\langle \rangle$  est une moyenne statistique. A la limite continue, l'équation du mouvement s'écrit :

$$\zeta \frac{\partial \mathbf{r}_n}{\partial t} = k \frac{\partial^2 \mathbf{r}_n}{\partial n^2} + \mathbf{L}_n. \quad (\text{C.2})$$

En toute rigueur, les équations du mouvement écrites ci-dessus sont valables pour  $2 \leq n \leq N - 1$ . Pour qu'elles le soient également pour les billes  $n = 1$  et  $n = N$ , on ajoute

deux billes immobiles supplémentaires  $n = 0$  et  $n = N + 1$  en  $\mathbf{r}_0 = \mathbf{r}_1$  et  $\mathbf{r}_{N+1} = \mathbf{r}_N$ . Les conditions aux limites s'écrivent alors :

$$\left. \frac{\partial \mathbf{r}_n}{\partial n} \right|_0 = 0 \quad \text{et} \quad \left. \frac{\partial \mathbf{r}_n}{\partial n} \right|_N = 0. \quad (\text{C.3})$$

## C.2 Modes normaux d'oscillation

L'équation du mouvement (C.2) peut être résolue en introduisant les modes normaux d'oscillation :

$$\mathbf{X}_p = \frac{1}{N} \int_0^N dn \mathbf{r}_n \sin\left(\frac{pn\pi}{N}\right).$$

En partant de C.2, après deux intégrations par partie et en utilisant les conditions aux limites (C.3), il vient :

$$\zeta \frac{\partial \mathbf{X}_p}{\partial t} = -k \left(\frac{p\pi}{N}\right)^2 \mathbf{X}_p + \frac{1}{N} \int_0^N dn \mathbf{L}_n \sin\left(\frac{pn\pi}{N}\right).$$

A partir du dernier terme, on veut construire une fonction de Langevin  $\mathbf{L}_p$  qui vérifie une relation analogue à la relation (C.1) et qui permette que l'équation du mouvement vérifiée par  $\mathbf{X}_p$  prenne une forme simple. En posant :

$$\mathbf{L}_p = 2 \int_0^N dn \mathbf{L}_n \sin\left(\frac{pn\pi}{N}\right),$$

on vérifie facilement que :

$$\langle L_{p,\alpha}(t) L_{q,\beta}(t') \rangle = 2\zeta_p k_B T \delta(t - t') \delta_{\alpha\beta} \delta_{pq}, \quad \text{avec} \quad \zeta_0 = N\zeta \quad \text{et} \quad \zeta_p = 2N\zeta \quad \text{pour} \quad p \geq 1. \quad (\text{C.4})$$

L'équation du mouvement devient :

$$\zeta_p \frac{\partial \mathbf{X}_p}{\partial t} = -k_p \mathbf{X}_p + \mathbf{L}_p, \quad \text{avec} \quad k_p = \frac{2\pi^2 p^2}{N} k. \quad (\text{C.5})$$

La fonction de corrélation spatiale s'obtient facilement à partir de cette équation :

$$\langle X_{p\alpha}(t) X_{q\beta}(0) \rangle = \delta_{pq} \delta_{\alpha\beta} \frac{k_B T}{k_p} \exp(-t/\tau_p) \quad \text{avec} \quad \tau_p = \frac{1}{p^2} \frac{\zeta_1}{k_1} = \frac{1}{p^2} \left(\frac{N}{\pi}\right)^2 \frac{\zeta}{k}. \quad (\text{C.6})$$

## C.3 Application à l'ADN

### C.3.1 Élasticité entropique

Le modèle précédent peut être appliqué à l'ADN en prenant quelques précautions, car la force exercée par les ressorts sur les billes n'est pas élastique. Si  $r$  est la distance entre deux billes successives, la force liée à la raideur entropique de l'ADN s'écrit :

$$T(r) = k(r)r \quad \text{avec} \quad k(r) = k_0 f\left(\frac{r}{L_0}\right), \quad k_0 = \frac{3k_B T}{2L_0 L_p} \quad \text{et} \quad f\left(\frac{r}{L_0}\right) = \frac{2}{3} + \frac{L_0}{6r} \left( \frac{1}{(1 - r/L_0)^2} - 1 \right). \quad (\text{C.7})$$

La fonction  $f(r/L_0)$  traduit la non-linéarité de la force de rappel, cette expression n'est pas tout à fait exacte, il manque des termes numériques qui introduisent une correction de l'ordre du % [99]. Elle est construite de sorte que  $k \simeq k_0$  lorsque  $r/L_0$  tend vers 0.

### C.3.2 Linéarisation

On retrouve une force élastique en supposant que l'amplitude des fluctuations de l'ADN est faible devant  $L_0$  et donc que  $r(t)/L_0 \simeq R/L$ , ce qui conduit à une constante de raideur constante  $k = k_0 f(R/L)$ , différente selon que sont considérées les fluctuations transverses ou longitudinales de l'ADN.

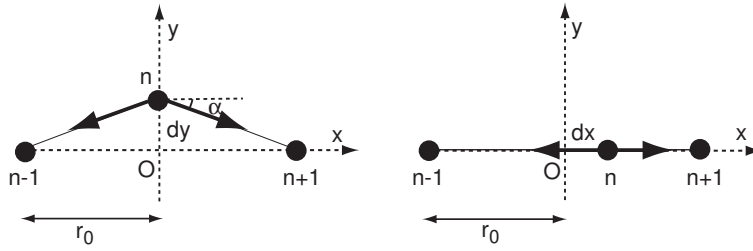


Figure C.1: Linéarisation de la tension exercée par l'ADN sur une bille.

**Constante de raideur associée aux fluctuations transverses :** considérons trois billes successives de l'ADN, initialement alignées suivant la direction  $(Ox)$ . La bille du centre est déplacée d'une quantité  $dy$  le long de l'axe  $(Oy)$ . La force  $\mathbf{F}(dy)$  qui s'exerce sur la bille s'écrit :

$$\mathbf{F}(dy) = -2k \left( \sqrt{r_0^2 + dy^2} \right) \sqrt{r_0^2 + dy^2} \sin \alpha \mathbf{j} \quad (\text{C.8})$$

$$= -2k \left( \sqrt{r_0^2 + dy^2} \right) \sqrt{r_0^2 + dy^2} \frac{dy}{\sqrt{r_0^2 + dy^2}} \mathbf{j} \quad (\text{C.9})$$

$$= -2k(r_0)dy \mathbf{j}, \quad (\text{C.10})$$

force de rappel dont la constante de raideur  $k_{\perp}$  est :

$$k_{\perp} = k(R/L) = k_0 \left( \frac{2}{3} + \frac{L}{6R} \left( \frac{1}{(1 - R/L)^2} - 1 \right) \right) = k_0 f_{\perp}(R/L),$$

où  $f_{\perp} = f$ . Dans notre situation, l'ADN est étiré à  $\sim 2/3$ , on trouve  $f(2/3) = 8/3$ .

**Constante de raideur associée aux fluctuations longitudinales :** la bille du centre est cette fois déplacée d'une quantité  $dx$  le long de l'ADN. A l'ordre 1, la force  $\mathbf{F}(dx)$  qui s'exerce sur la bille déplacée de  $dx$  s'écrit :

$$\mathbf{F}(dx) = k(r_0 - dx)(r_0 - dx) \mathbf{i} - k(r_0 + dx)(r_0 + dx) \mathbf{i} \quad (\text{C.11})$$

$$= -2 \left. \frac{d(k(r)r)}{dr} \right|_{r_0} dx \mathbf{i}, \quad (\text{C.12})$$

force de rappel dont la constante de raideur  $k_{\parallel}$  est :

$$k_{\parallel} = \left. \frac{d(k(r)r)}{dr} \right|_{R/L} = k_0 \left( \frac{2}{3} + \frac{1}{3} \frac{1}{(1-R/L)^3} \right) = k_0 f_{\parallel}(R/L).$$

On trouve ainsi  $f_{\parallel}(2/3) = 29/3$ . La constante de raideur longitudinale est donc supérieure à la constante de raideur transverse dans la gamme d'étirement qui nous intéresse.

### C.3.3 Coefficient de friction

Le coefficient  $\zeta$  introduit dans la partie précédente a une expression qui peut être plus ou moins compliquée, en particulier si nous tenons compte des couplages hydrodynamiques à longues distances entre billes (modèle de Zimm). Pour simplifier, nous écrivons :  $\zeta = \eta L_0 = \eta L/N$ , où  $\eta = 10^{-3}$  Pa.s est la viscosité dynamique de l'eau qui donne le bon ordre de grandeur de  $\zeta$ . Nous cherchons à obtenir simplement un ordre de grandeur, il est donc inutile de vouloir raffiner le modèle en prenant compte le fait que l'enzyme change  $\zeta$ .

### C.3.4 Modes normaux d'oscillation :

Les constantes  $k_p$  et  $\zeta_p$  introduites dans la partie précédente s'écrivent pour l'ADN :

$$k_p = 3\pi^2 p^2 \frac{k_B T}{L L_p} f \quad \text{et} \quad \zeta_p = 2\eta L,$$

où  $f = f_{\parallel}(R/L)$  ou  $f_{\perp}(R/L)$  suivant que l'on considère les modes longitudinaux ou transverses. Ces constantes ne dépendent plus de  $N$ , et permettent d'exprimer la fonction de corrélation :

$$\langle X_{p\alpha}(t) X_{q\beta}(0) \rangle = \delta_{pq} \delta_{\alpha\beta} \frac{1}{p^2} \frac{L L_p}{3\pi^2} \frac{1}{f} \exp(-t/\tau_p) \quad \text{avec} \quad \tau_p = \frac{1}{p^2} \frac{2\eta L^2 L_p}{3\pi^2 k_B T} \frac{1}{f}.$$

Amplitude et temps de corrélation décroissent en  $1/p^2$ , donc rapidement. Nous simplifions le problème en ne conservant que le premier terme  $p = 1$  pour la suite des calculs.

### C.3.5 Fluctuations spatiales

Considérons un point de l'ADN de coordonnées moyennes ( $\langle x_n \rangle = nR/N$ ,  $\langle y_n \rangle = 0$ ). Une des coordonnées, par exemple  $y_n(t)$ , est liée à  $Y_p$  par la relation :

$$y_n(t) = 2 \sum_{p=1}^N Y_p(t) \sin\left(\frac{pn\pi}{N}\right) \simeq 2Y_1(t) \sin\left(\frac{n\pi}{N}\right),$$

et la fonction de corrélation est donnée par :

$$\langle y_n(t) y_n(0) \rangle = 4 \langle Y_1(t) Y_1(0) \rangle = \frac{4L L_p}{3\pi^2} \frac{1}{f_{\perp}} \sin^2\left(\frac{\langle x_n \rangle}{R} \pi\right) \exp(-t/\tau_1).$$

La moyenne sur le brin d'ADN est donc :

$$\langle\langle y_n(t)y_n(0) \rangle\rangle = \frac{2LL_p}{3\pi^2} \frac{1}{f_\perp} \exp(-t/\tau_1).$$

Pour une molécule d'ADN de longueur 8.5 kb ( $L \simeq 3 \mu\text{m}$ ), avec un taux d'étirement de 2/3, il vient :

- Fluctuations longitudinales :

$$\langle\langle x^2(0) \rangle\rangle = \frac{2LL_p}{3\pi^2} \frac{1}{f_\parallel} = (30 \text{ nm})^2 \quad \text{et} \quad \tau_\ell = \frac{2\eta L^2 L_p}{3\pi^2 k_B T} \frac{1}{f_\parallel} = 0,8 \text{ ms}$$

- Fluctuations transverses :

$$\langle\langle y^2(0) \rangle\rangle = \frac{2LL_p}{3\pi^2} \frac{1}{f_\perp} = (60 \text{ nm})^2 \quad \text{et} \quad \tau_{tr} = \frac{2\eta L^2 L_p}{3\pi^2 k_B T} \frac{1}{f_\perp} = 3 \text{ ms}.$$

## C.4 Détermination expérimentale

Les expériences d'interaction avec l'enzyme EcoR V marquée aux QDot-605 ont pu être réalisées avec un temps d'acquisition  $\Delta t = 2 \text{ ms}$  grâce à la brillance des quantum-dots. Pour remonter aux valeurs des temps de corrélations longitudinal  $\tau_\ell$  et transverse  $\tau_{tr}$  ainsi qu'aux amplitudes correspondantes  $\sigma_\ell$  et  $\sigma_{tr}$ , il suffit d'ajuster les premiers points des courbes expérimentales des MSD poolées longitudinales et transverses par les fonctions données en C.13 :

$$\rho_\ell(t) = 2\sigma_{\text{tot},\ell}^2 \times \left(1 - e^{-\frac{t}{\tau_\ell}}\right) + B \cdot t \quad \text{et} \quad \rho_{tr}(t) = 2\sigma_{\text{tot},tr}^2 \times \left(1 - e^{-\frac{t}{\tau_{tr}}}\right). \quad (\text{C.13})$$

Il est légitime d'utiliser ces expressions puisque les événements pris en compte dans le pooling correspondent à des enzymes aléatoirement réparties sur l'ADN. Les courbes des MSDs poolées apparaissent sur la Figure C.2 et les paramètres de l'ajustement des 25 premiers points des MSD poolées apparaissent dans le tableau C.1.

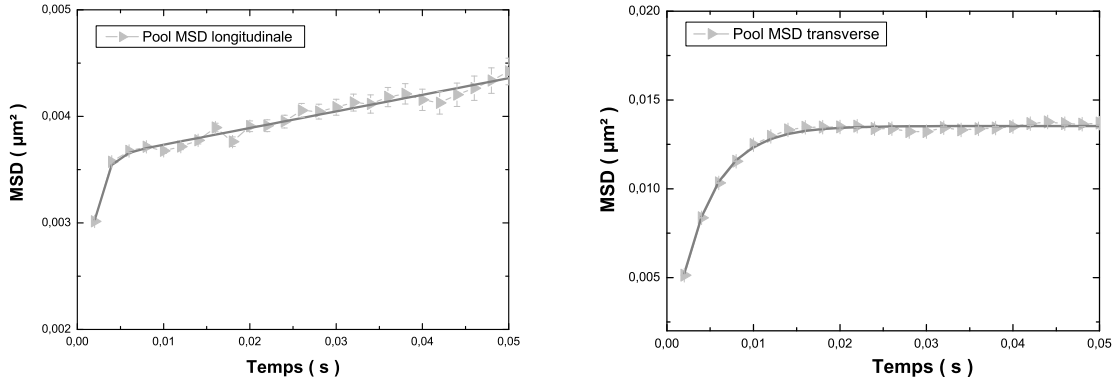
### C.4.1 Temps de corrélation

Les temps déterminés par l'ajustement des courbes expérimentales selon les équations C.13 sont comparables aux temps théoriquement prédits puisque nous avons trouvé  $\tau_\ell \approx 1,1 \text{ ms}$  et  $\tau_{tr} \approx 4 \text{ ms}$ .

### C.4.2 Amplitude des fluctuations

L'ajustement expérimental des MSDs poolées donne :  $\sigma_{\text{tot},\ell} \approx 42 \text{ nm}$  et  $\sigma_{\text{tot},tr} \approx 82 \text{ nm}$ . Pour pouvoir les comparer aux prédictions théoriques, il faut tenir compte de la précision de pointé  $\sigma_{P,\ell/tr}$  :  $\sigma_{\text{tot},\ell/tr}^2 = \sigma_{\ell/tr}^2 + \sigma_{P,\ell/tr}^2$ . En prenant une précision de pointé identique dans les 2 directions et égale à  $\sigma_P = 30 \text{ nm}$ , il vient :  $\sigma_\ell \approx 30 \text{ nm}$  et  $\sigma_{tr} \approx 76 \text{ nm}$ , valeurs cohérentes avec celles déterminées avec le modèle précédemment décrit.





**Figure C.2:** Détermination expérimentale des amplitudes et temps de corrélation des molécules d'ADN étirées à partir des premiers points de la MSD poolée. **Gauche** : MSD longitudinale poolée (triangle) ajustée par  $\rho_\ell$  (trait plein). **Droite** : MSD transverse poolée (triangle) ajustée par  $\rho_{tr}$  (trait plein). Les paramètres des ajustements apparaissent dans le Tableau C.1

### C.4.3 Conclusion

Le tableau C.1 montre que les valeurs expérimentales sont cohérentes avec les prédictions théoriques : le modèle est bien adapté.

Les temps de corrélation déterminés expérimentalement sont légèrement supérieurs aux valeurs attendues, on peut supposer que la présence des nanocristaux (de rayon hydrodynamique  $\sim 15$  nm) modifie le coefficient  $\zeta$ , qui aurait été sous-estimé dans notre modèle.

	Paramètres	Valeurs Expérimentales	Valeurs Théoriques
Longitudinal	$\sigma_\ell$	30 nm	30 nm
	$\tau_\ell$	1,1 ms	0,8 ms
Transverse	$\sigma_{tr}$	76 nm	60 nm
	$\tau_{tr}$	4 ms	3ms

**Tableau C.1:** Amplitudes et temps de corrélation caractéristiques des fluctuations des molécules d'ADN étirées.

# Bibliographie

- [1] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell - Fourth Edition*. Garland Science Publishing, 2002.
- [2] Lane D., Prentki P., and Chandler M. Use of gel retardation to analyse protein-nucleic acid interactions. *Microbiol. Rev.*, 56 :509–528, 1992.
- [3] Riggs A. D., Bourgeois S., and Cohn M. The lac repressor-operator interaction. III. Kinetics studies. *Biophys. Chem.*, 9 :413–414, 1970.
- [4] Schurr J. M. The one-dimensional diffusion coefficient of proteins absorbed on DNA. Hydrodynamic considerations. *Biophys. Chem.*, 9(4) :413–414, 1975.
- [5] Howard C. Berg. *Random Walks in Biology*. Princeton University Press, 1993.
- [6] Berg O. G., Winter R. B., and von Hippel P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. I : models and theory. *Biochemistry*, 20 :6929–6948, 1981.
- [7] Richter P. H. and Eigen M. Diffusion controlled reaction rates in spheroidal geometry. Application to repressor-operator association and membrane bound enzymes. *Biophys. Chem.*, 2 :255–263, 1974.
- [8] von Hippel P. H. and Berg O. G. Facilitated target location in biological systems. *J. Biol. Chem.*, 264(2) :675–678, 1989.
- [9] Schreiber G. Kinetic studies of protein-protein interactions. *Curr. Opin. Struct. Biol.*, 12(1) :41–47, 2002.
- [10] Record M. T., Zhang W., and Anderson C. F. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes : a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem.*, 51 :281–353, 1998.
- [11] Winter R. B., Berg O. G., and von Hippel P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. III, the *Escherichia coli* lac repressor-operator interaction : kinetic measurements and conclusions. *Biochemistry*, 20 :6961–6977, 1981.

## Bibliographie

---

- [12] Jack W. E., Terry B. J., and Modrich P. Involvement of outside DNA, sequences in the major kinetic path by which EcoR I, endonuclease locates and cleaves its recognition sequence. *Proc. Natl. Acad. Sci. USA.*, 79(13) :4010–4014, 1982.
- [13] Ehbrecht H.-J., Pingoud A., Urbanke C., Mass G., and Gualerzi C. Linear diffusion of restriction endonucleases on DNA. *J. Biol. Chem.*, 260 :6160–6166, 1985.
- [14] Terry B. J., Jack W. E., and Modrich P. Facilitated diffusion during catalysis by EcoRI, endonuclease. nonspecific interactions in EcoR I, catalysis. *J. Biol. Chem.*, 260 :13130–13137, 1985.
- [15] Jeltsch A. and Pingoud A. Kinetic characterization of linear diffusion of the restriction endonuclease EcoRV, on DNA. *Biochemistry*, 37(8) :2160–2169, 1998.
- [16] Halford S. E. and Marko J. F. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.*, 32(10) :3040–3052, 2004.
- [17] Stanford N. P., Szczelkun M. D., Marko J. F., and Halford S. E. One- and three-dimensional pathways for proteins to reach specific DNA sites. *EMBO J.*, 19(23) :6546–6557, 2000.
- [18] Jeltsch A. and Urbanke C. *Restriction endonucleases*. A. Pingoud, New York, 2004.
- [19] Adam G. and Delbruck M. Reduction of dimensionality in biological diffusion processes. *Structural Chemistry and Molecular Biology*, pages 198–215, 1968.
- [20] Kampmann M. Obstacle bypass in protein motion along DNA, by two-dimensional rather than one-dimensional sliding. *J. Biol. Chem.*, pages 38715–38820, 2004.
- [21] Zhou H. X. A model for the mediation of processivity of DNA-targeting proteins by nonspecific binding : dependence on DNA, length and presence of obstacles. 88(3) :1608–1615, 2005.
- [22] Belotserkovskii B. P. and Zarling D. A. Analysis of a one-dimensional random walk with irreversible losses at each step : applications for protein movement on DNA. *J. Theor. Biol.*, 226(2) :195–203, 2004.
- [23] Erskine S. G., Baldwin G. S., and Halford S. E. Rapid-reaction analysis of plasmid DNA, cleavage by the EcoRV, restriction endonuclease. *Biochemistry*, 36(24) :7567–7576, 1997.
- [24] Milsom S. E., Halford S. E., Embleton M. L., and Szczelkun. Analysis of DNA, looping interactions by type II, restriction enzymes that requires two copies of their recognition sites. *J. Mol. Biol.*, 31 :517–528, 2001.
- [25] Slutsky M. and Mirny L. A. Kinetics of protein-DNA, interaction : facilitated target location in sequence-dependent potential. *Biophys. J.*, 87(6) :4021–4035, 2004.
- [26] Coppey M., Bénichou O., Voituriez R., and Moreau M. Kinetics of target site localization of a protein on DNA : a stochastic approach. *Biophys. J.*, 87(3) :1640–1649, 2004.

## Bibliographie

---

- [27] Hildebrandt E. R. and Cozzarelli N. R. Comparison of recombination in vitro and in *E. coli* cells : measure of the effective concentration of DNA, *in vivo*. *Cell.*, 1995(81) :331–340, 3.
- [28] Jeltsch A., Alves J., Wolfes H., Maass G., and Pingoud A. Pausing of the restriction endonuclease EcoRI,during linear diffusion on DNA. *Biochemistry.*, 33(34) :10215–10219, 1994.
- [29] Jeltsch A., Wenz C., Stahl F., and Pingoud A. Linear diffusion of the restriction endonuclease EcoRV on DNA is essential for the *in vivo* function of the enzyme. *EMBO J.*, 15(18) :5104–5111, 1996.
- [30] Roberts R. J. and al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, 31(7) :1805–1812, 2003.
- [31] Pingoud A. and Jeltsch A. Structure and function of type II restriction endonuclease. *Nucleic Acids Res.*, 29(18) :3705–3727, 2001.
- [32] Taylor J. D., Badcoe I. G., Clarke A. R., and Halford S. E. EcoRV, restriction endonuclease binds all DNA, sequences with equal affinity. *Biochemistry.*, 30(36) :8743–8753, 1991.
- [33] Vipond I. B., Baldwin G. S., and Halford S. E. Divalent metal ions at the active sites of the EcoRV, and EcoRI, restriction endonucleases. *Biochemistry*, 34(697) :697–704, 1995.
- [34] Baldwin G. S., Sessions R. B., Erskine S. G., and Halford S. E. DNA, cleavage by the EcoRV, restriction endonuclease : roles of divalent metal ions in specificity and catalysis. *J. Mol. Biol.*, 288(1) :87–103, 1999.
- [35] Winkler F.K., Banner D.W., Oefner C., Tsernoglou D., Brown R. S., Heathman S. P., Bryan R. K., Martin P. D., Petratos K., and Wilson K. S. The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.*, 2 :1781–1795, 1993.
- [36] Shimamoto N. One-dimensional diffusion of proteins along DNA. *The Journal of Biological Chemistry*, 274(22) :15293–15296, may 1999.
- [37] Halford S. E. Hopping, jumping and looping by restriction enzymes. *Novartis Medal Lecture*, 29 :363–374, april 2001.
- [38] Viadiu H. and Aggarwal A. K. Structure of BamHI, bound to non-specific DNA : a model for DNA, sliding. *Mol Cell.*, 5(5) :889–895, 2000.
- [39] Schulze C., Jeltsch A., Franke I., Urbanke C., and Pingoud A. Crosslinking the EcoRV restriction endonuclease across the DNA-binding site reveals transient intermediates and conformational changes of the enzyme during DNA binding and catalytic turnover. *EMBO J.*, 17(22) :6757–6766, 1998.

- [40] Gowers D. M. and Halford S. E. Protein motion from non-specific to specific DNA, by three-dimensional routes aided by supercoiling. *EMBO J.*, 22(6) :1410–1418, 2003.
- [41] Rost F. W. D. *Fluorescence : physics and chemistry.*, volume 1. Cambridge University Press, 1992.
- [42] Stéphane Bonneau. *Chemins minimaux en analyse d'images : nouvelles contributions et applications à l'imagerie biologique.* Thèse de doctorat, Université Paris Dauphine, déc 2006.
- [43] Marvin Minsky. Confocal scanning microscope. Technical report, 3 novembre 1955.
- [44] Kaiser W. and Garrett C. G. B. Two-photon excitation in  $\text{CaF}_2:\text{Eu}^{2+}$ . *Phys. Rev. Lett.*, 7 :229–231, 1961.
- [45] Steyer J. A. and Almers W. A real-time view of life within 100 nm of the plasma membrane. *Nat. Rev. Mol. Cell. Biol.*, 2(4) :268–275, 2001.
- [46] Noji H., Yasuda R., Yoshida M., and Kinoshita K. Direct observation of the rotation of F1-ATPase. *Nature*, 386 :299–302, 1997.
- [47] van den Broek B., Vanzi F., Normanno D, Pavone F. S., and Wuite G. J. Real-time observation of DNA, looping dynamics of type IIE, restriction enzymes NaeI and NarI. *Nucleic Acids Res.*, 34 :167–174, 2006.
- [48] Tolic-Norrelykke S. F., Rasmussen M. B., Pavone F. S., Berg-Sorensen K., and Odershede L. B. Stepwise bending of DNA, by a single TATA,-box binding protein. *Biophys. J.*, 90 :3697–3703, 2006.
- [49] Vanzi F., Broggio C., Sacconi L., and Pavone F. S. Lac repressor hinge flexibility and DNA, looping : single molecule kinetics by tethered particle motion. *Nucleic Acids Res.*, 34 :3409–3420, 2006.
- [50] Wang M. D., Schnitzer M. J., Yin H., Landick R., Gelles J., and Block S. M. Force and velocity measured for single molecules of RNA, polymerase. *Science*, 282 :902–907, 1998.
- [51] Thomen P., Lopez P. J., and Heslot F. Unravelling the mechanism of RNA,-polymerase forward motion by using mechanical force. *Phys. Rev. Lett.*, 94 :128102, 2005.
- [52] van den Broek B., Noom M. C., and Wuite G.J. DNA-tension dependence of restriction enzyme activity reveals mechanochemical properties of the reaction pathway. *Nucleic Acids Res.*, 33 :276, 2005.
- [53] Harada Y., Funatsu T., Murakami K., Nonoyama Y., Ishihama A., and Yanagida T. Single-molecule imaging of RNA polymerase-DNA interactions in real time. *Biophys. J.*, 76 :709–715, 1998.

- [54] Ladoux B., Quivy J. P., Doyle P., du Roure O., Almouzni G., and Viovy J. L. Force and velocity measured for single molecules of RNA, polymerase. *Proc. Natl. Acad. Sci. USA.*, 97 :14251–14257, 2000.
- [55] Matsuura S., Komatsu J., Hirano K., Yasuda H., Takashima K., Katsura S., and Mizuno A. Real-time observation of a single DNA, digestion by lambda exonuclease under a fluorescence microscope field. *Nucleic Acids Res.*, 29 :16, 2001.
- [56] Schafer B., Nasanshargal B., Monajembashi S., Gemeinhardt H., Greulich K. O., and Uhl V. Study of single-molecule dynamics and reactions with classic light microscopy. *Cytometry*, 36 :209–216, 1999.
- [57] Spies M., Bianco P. R., Dillingham M. S., Handa N., Baskin R. J., and Kowalczykowski S. C. A molecular throttle : the recombination hotspot chi controls DNA, translocation by the RecBCD, helicase. *Cytometry*, 114 :647–654, 2003.
- [58] Koster D. A., Croquette V., Dekker C., Shuman S., and Dekker N. H. Friction and torque govern the relaxation of DNA, supercoils by eukaryotic topoisomerase IB. *Nature*, 434 :671–674, 2005.
- [59] Charvin G., Bensimon D., and Croquette V. Single-molecule study of DNA, unlinking by eukaryotic and prokaryotic type-II, topoisomerases. *Proc. Natl. Acad. Sci. USA.*, 100 :17, 2003.
- [60] Washizu M., Kurosawa O., Arai I., Suzuki S., and Shimamoto N. Applications of electrostatic stretch-and-positioning of DNA. volume 31, pages 447–456, 1995.
- [61] Dewarrat F. Calame M. and Schönenberger C. Orientation and positioning of DNA, molecules with an electric field technique. *Single Mol.*, 3 :189–193, 2002.
- [62] Perkins T. T., Quake S. R., Smith D. E., and Chu S. Relaxation of a single DNA,molecule observed by optical microscopy. *Science*, 264 :822–826, 1994.
- [63] Doyle P. S., Ladoux B., and Viovy J.L. Dynamics of a tethered polymer in shear flow. *Phys. Rev. Lett.*, 84 :4769–4772, 2000.
- [64] Ladoux B. and Doyle P. S. Stretching tethered DNA,chains in shear flow. *Europhys. Lett.*, 52 :511–517, 2000.
- [65] Bensimon A., Simon A., Chiffaudel A., Croquette V., Heslot F., and Bensimon D. Alignment and sensitive detection of DNA,by a moving interface. *Science*, 265 :2096–2098, 1994.
- [66] Bensimon D., Simon A. J., Croquette V. V., and Bensimon A. Stretching DNA,with a receding meniscus : Experiments and models. *Phys. Rev. Lett.*, 74 :4754–4757, 1995.
- [67] Jean-François Allemand. *Micromanipulations de molécules d'ADN isolées*. Thèse de doctorat, Université Pierre et Marie Curie, 1997.

## Bibliographie

---

- [68] Aurélien Crut. *Diffusion de l'enzyme de restriction EcoRV sur des molécules d'ADN étirées : une étude par microscopie de fluorescence à l'échelle de la molécule individuelle*. Thèse de doctorat, Université de Pierre et Marie Curie, nov 2005.
- [69] Meng X., Cai W., and Schwartz D. C. Inhibition of restriction endonuclease activity by DNA-binding fluorochromes. *J. Biomol. Struct. Dyn.*, 13 :945–951, 1996.
- [70] Guéroui Z., Place C., Freyssingas E., and Berge B. Observation by fluorescence microscopy of transcription on single combed DNA. *Proc. Natl. Acad. Sci. US A.*, 99 :6005–6010, 2002.
- [71] Israelachvili J. N. and Tabor D. The measurement of Van der Waals dispersion forces in the range 1.5 to 130 nm. *Proc. Roy. Soc. London A*, 331(1584) :19–38, 1972.
- [72] Guthold M., Zhu X., Rivetti C., Yang G., Thomson N. H., Kasas S., Hansma H. G., Smith B., Hansma P. K., and Bustamante C. Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase. *Biophys. J.*, 77 :2284–2294, 1999.
- [73] van Noort S. J., van der Werf K. O., Eker A. P., Wyman C., de Groot B. G., van Hulst N. F., and Greve J. Direct visualization of dynamic protein-DNA interactions with a dedicated atomic force microscope. *Biophys. J.*, 74 :2840–2849, 1998.
- [74] Schwille P., Haupts U., Maiti S., and Webb W. W. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophys. J.*, 77 :2251–2265, 1999.
- [75] Arrio-Dupont M., Foucault G., Vacher M., Devaux P. F., and Cribier S. Translational diffusion of globular proteins in the cytoplasm of cultured muscle cells. *Biophys. J.*, (78) :901–907, 2000.
- [76] Roberts C.H. Estimating friction coefficients of mixed globular/chain molecules, such as protein/DNA complexes. *Biophys. J.*, 69 :840–848, 1995.
- [77] Bloomfield V., Dalton W. O., and Van Holde K. E. Frictional coefficients of multi-subunit structures. I : theory. *Biopolymers*, 5 :135–148, 1967.
- [78] Michalet X., Pinaud F. F., Bentolila L. A., Tsay J. M., Doose S. Li J. J., Sundaresan G., Wu A. M., Gambhir S. S., and Weiss S. Quantum dots for live cells, in vivo imaging, and diagnostics. *Science*, 307(5709) :538–544, 2005.
- [79] Doose S., Tsay J. M., Pinaud F., and Weiss S. Comparison of photophysical and colloidal properties of biocompatible semiconductor nanocrystals using fluorescence correlation spectroscopy. *Anal. Chem.*, 77(7) :2235–2242, 2005.
- [80] Flory P. *Statistics of Chain Molecules*. Interscience Publisher, New York, 1969.
- [81] Crut A., Lasne D., Allemand J.-F., Dahan M., and Desbiolles P. Transverse fluctuations of single DNA molecules attached at both extremities to a surface. *Phys. Rev. E*, 67, may 2003.

- [82] Marko J. F. and Siggia E. D. Stretching DNA. *Macromolecules*, 28 :8759–8770, 1995.
- [83] Ladoux B., Quivy J.-P., Doyle P. S., Almouzni G., and Viovy J.-L. Direct imaging of single-molecules : from dynamics of a single DNA,chain to the study of complex DNA-protein interactions. *Science Progress*, 84 :267–290, 2001.
- [84] Bustamante C., Marko J. F., Siggia E. D., and Smith S. Entropic elasticity of  $\lambda$ -phage DNA. *Science*, 265 :1599–1600, september 1994.
- [85] Wenner J. R. and Bloomfield V.A. Buffer effects on EcoRV, kinetics as measured by fluorescent staining and digital imaging of plasmid cleavage. *Anal. Biochem.*, 268(2) :201–212, 1999.
- [86] Wenner J. R. and Bloomfield V. A. Crowding effects on EcoRV kinetics and binding. *Biophys J.* 77, 77(6) :3234–3241, 1999.
- [87] Blainey P. C., van Oijen A. M., Banerjee A., Verdine G. L., and Xie X. S. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl. Acad. Sci. USA.*, 103(15) :5752–5757, 2006.
- [88] Manghi M. and Netz R. R. Variational theory for a single polyelectrolyte revisited. *Eur. Phys. J. E*, 14(1) :67–77, 2004.
- [89] Bonneau S., Cohen L., and Dahan M. A multiple target approach for single quantum dot tracking. volume 1, pages 664–667, 2004.
- [90] Saxton M. J. Single-particle tracking : the distribution of diffusion coefficients. *Biophys. J.*, 72 :1744–53, 1997.
- [91] Qian H. Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophys. J.*, 60 :1910–1921, 1991.
- [92] Wang Y. M., Austin R. H., and Cox E. C. Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys. Rev. Lett.*, 97(048302), 2006.
- [93] Granéli A., Yeykal C. C., Robertson R. B., and Greene E. C. Long-distance lateral diffusion of human Rad51 on double-stranded DNA. *Proc Natl Acad Sci U S A.*, 103(5) :1221–1226, 2006.
- [94] Kim J. H. and Larson R. G. Single-molecule analysis of 1D diffusion and transcription elongation of T7 RNA polymerase along individual stretched DNA molecules. *Nucleic Acids Res.*, 35(11), 2007.
- [95] Bonnet I.\*, Biebricher A.\*, Loverdo C., Voituriez R., Bénichou O., Escudé C., Wende W., Pingoud A., and Desbailles P. Sliding and jumping of single ecorv enzymes along stretched dna. *En cours de soumission*.
- [96] Redner S. *A guide to first passage time processe*. Cambridge University Press, 2001.
- [97] A. Vogt. [http ://www.mapleprimes.com/blog/alec/](http://www.mapleprimes.com/blog/alec/), 2006.



## Bibliographie

---

- [98] Masao Doi and Sam F. Edwards. *The theory of polymer dynamics*. Oxford Science, 1986.
- [99] Bouchiat C., Wang M. D., Allemand J.-F., Strick T., Block S. M., and Croquette V. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.*, 76 :409–413, 1999.

**Résumé :** Certaines protéines qui interagissent avec l'ADN sont capables de trouver très rapidement une séquence spécifique de quelques paires de bases sur des ADNs pouvant contenir plusieurs milliers de paires de bases. Pour expliquer ce phénomène connu des biologistes depuis plus de trente ans, plusieurs mécanismes de localisation, regroupés sous le terme de « diffusion facilitée », ont été proposés. Tous supposent une association initiale à de l'ADN non spécifique, suivie d'une translocation le long de l'ADN jusqu'au site de reconnaissance. Le mécanisme qui sous-tend cette translocation est toujours discuté, et pourrait impliquer une diffusion linéaire le long de l'ADN, dénommée « sliding », et/ou une série de sauts (« jumping »).

Pour élucider le mécanisme de diffusion facilitée, nous avons utilisé la microscopie de fluorescence par onde évanescente pour visualiser l'interaction d'une unique protéine -l'enzyme de restriction EcoRV- avec de l'ADN non-spécifique. Dans nos expériences *in vitro*, une molécule d'ADN est étirée, attachée par ses extrémités à une surface, et les enzymes sont couplées à des fluorophores (Cy3B). Nous sommes parvenus ainsi à visualiser directement les processus de « sliding » et de « jumping ». Nous avons mesuré le coefficient de diffusion 1D des enzymes le long de l'ADN ( $D_1 \sim 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ ) ainsi que le nombre de saut par interaction ( $\sim 20$ ). Un modèle simple a permis à partir des résultats expérimentaux, de caractériser la diffusion facilitée d'EcoRV.

**Mots clés :** ADN — EcoRV — molécule unique — microscopie de fluorescence — onde évanescente — diffusion facilitée.

**Abstract :** Some proteins can rapidly locate a short recognition site within large DNA using "facilitated diffusion", a process in which protein first bind to the DNA via non-specific interactions and then translocate along DNA to the specific sequence. "Sliding" (1D diffusion) and "jumping" (3D translocations) on DNA are two mechanisms that are believed to accelerate target location, but their respective roles in facilitated diffusion remains to be clarified.

Using total internal fluorescence microscopy, we studied the "facilitated diffusion" of the restriction endonuclease EcoRV along stretched DNA on the single molecule level. In addition to sliding, we observed jumps which were demonstrated to stem from 3D excursions of the enzyme between distant DNA sites. We developed an analytical model based on 3D diffusion to account for both the jump frequency and the jump length distribution. We deduced from our study key parameters regarding the facilitated diffusion of EcoRV, such as the 1D diffusion coefficient ( $D_1 \sim 10^{-2} \mu\text{m}^2 \cdot \text{s}^{-1}$ ) and the number of jumps ( $\sim 20$ ) occurring during an interaction event with non-specific DNA.

**Keywords :** DNA — EcoRV — single molecule — fluorescence microscopy — evanescent wave — facilitated diffusion.