



**HAL**  
open science

# Les structures secondaires dans l'ARN : une étude par mesure de forces sur molécules uniques

Mathilde Bercy

► **To cite this version:**

Mathilde Bercy. Les structures secondaires dans l'ARN : une étude par mesure de forces sur molécules uniques. Physique [physics]. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066718 . tel-01409699

**HAL Id: tel-01409699**

**<https://theses.hal.science/tel-01409699>**

Submitted on 6 Dec 2016

**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.

**THESE DE DOCTORAT  
DE L'UNIVERSITE PIERRE ET MARIE CURIE**

**Spécialité : Physique**

**École doctorale : « Physique en Île-de-France »**

**réalisée**

**au Laboratoire de Nanobiophysique, ESPCI**

**présentée par**

**Mathilde BERCY**

**pour obtenir le grade de :**

**DOCTEUR DE L'UNIVERSITE PIERRE ET MARIE CURIE**

**Sujet de la thèse :**

**Les structures secondaires dans l'ARN :  
une étude par mesure de forces sur molécules uniques**

**soutenue le 1<sup>er</sup> décembre 2015**

**devant le jury composé de :**

<b>M.</b>	<b>Marc Dreyfus</b>	<b>Président du jury</b>
<b>M.</b>	<b>Emmanuel Margeat</b>	<b>Rapporteur</b>
<b>M.</b>	<b>Terence Strick</b>	<b>Rapporteur</b>
<b>M.</b>	<b>Thierry Bizebard</b>	<b>Examineur</b>
<b>M.</b>	<b>Erwin Peterman</b>	<b>Examineur</b>
<b>M.</b>	<b>Ulrich Bockelmann</b>	<b>Directeur de thèse</b>



# Remerciements

Je voudrais tout d'abord remercier les membres de mon jury pour leur participation à ma soutenance de thèse. Je tiens en particulier à souligner le travail de mes deux rapporteurs, Emmanuel Margeat et Terence Strick. Sans être particulièrement familiers de mon travail, ils ont immédiatement accepté ce rôle, et pris le temps d'analyser ma thèse en profondeur.

Un grand merci à Marc Dreyfus, qui avant de participer à mon jury avait déjà suivi l'évolution de ma thèse. Son expérience et sa bienveillance m'ont beaucoup appris, et permis d'avancer sur le difficile sujet des hélicases, entre autres.

Un autre grand merci à Thierry Bizebard, qui m'a appris presque tout ce que je sais de la biologie moléculaire. Merci pour la bonne humeur et l'entrain que tu apportes toujours avec toi, en plus de tes conseils.

Merci à Erwin Peterman, qui a fait le déplacement depuis Amsterdam. Je garde un excellent souvenir de la visite que j'ai eu la chance de faire de ton laboratoire, et j'espère que le travail qui se poursuivra après moi permettra de renforcer encore la collaboration entre nos deux équipes.

Je voudrais enfin remercier Ulrich Bockelmann, mon directeur de thèse, pour m'avoir proposé une place dans son laboratoire. Sans lui, rien de tout cela n'aurait évidemment été possible.

Mes pensées vont ensuite à Ismail : j'ai eu la chance de partager son bureau au début de ma thèse, et cela m'a permis de découvrir les meilleures pâtisseries parisiennes... Merci à ta gourmandise ! Plus sérieusement (en admettant qu'il y ait sujet plus sérieux que les gâteaux), merci pour ta disponibilité, ta patience, tes explications sur à peu près tous les équipements et le fonctionnement du laboratoire. Je ne peux absolument pas imaginer comment je m'en serais tirée sans toi. Merci encore, et j'espère bien qu'on continuera à partager encore longtemps nos meilleures pâtisseries !

Je veux aussi remercier Edouard et Laurent, mes co-thésards, qui plus que des collègues sont maintenant des amis. Merci de vous être battus pour intégrer le laboratoire, et y apporter un peu de vie ! Je n'oublierai pas les moments que nous avons partagés, et quel que soit le chemin que vous ayez choisi, bon vent (spécialement pour Edouard...) et bonne chance pour la suite.

Je n'oublie pas non plus Lena, dernière arrivée. Les deux petits mois que nous avons passés en commun au laboratoire sont passés bien vite ! Mais je pense que nous continuerons à échanger nos recettes bien après mon départ... En tout cas, je suis contente de savoir que mon travail sera poursuivi, et que je le laisse entre de bonnes mains.

Un petit mot aussi pour tous les stagiaires qui ont animé le labo ces dernières années !

Merci à Emilie et Emel avec qui j'ai adoré travailler. Et bien sûr à Hala, Samar, et Pierre. J'espère que nous resterons en contact encore longtemps.

Merci aussi à Kyle et Josette pour leurs conseils et les enrichissantes discussions que nous avons eues autour de la biologie (ou de la danse...).

Du côté de la physique, merci David pour les intéressantes discussions au sujet des hair-pins. Au delà de ça, c'était bien de savoir que quelqu'un d'extérieur à l'équipe s'intéressait à ce travail !

Encore un petit merci pour Pierre, qui m'a laissé son piège optique. Et qui a toujours été disponible et réactif lorsque j'avais des doutes sur son fonctionnement...

Un peu plus loin des manips mais toujours là au quotidien, je voudrais aussi remercier mes différents co-bureaux d'EC2M : Ziane, Maxime, Alex, Kunyun, et Guillaume. Ca a été très important pour moi d'être entourée d'autres thésards, et de partager leurs expériences. Merci d'avoir été là dans les bons et les moins bons moments (mention spéciale à Alex et Kunyun, merci !). J'en profite aussi pour remercier Michaël, Justine, et Antoine pour tous les thés et cafés partagés.

J'ai maintenant fait le tour du labo. Et je pense à tous les autres, ceux qui m'ont vraiment permis d'arriver au bout de ce doctorat. À François évidemment, qui a vécu toutes les galères de ma thèse alors qu'il était déjà bien occupé avec la sienne... Et qui m'a toujours soutenue, qui a toujours été là, et que je ne remercierai jamais assez.

Merci à mes amis qui ont eu la bonne idée de faire une thèse pas loin, et avec qui j'ai partagé une bonne partie de mes déjeuners ces dernières années. On a beaucoup râlé, certes, mais aussi pas mal rigolé, et surtout parlé de beaucoup de choses qui n'avaient rien de scientifique (!). Merci donc Binôme, Vincent, et Michel. Et là je n'espère pas, mais je suis sûre, que ni la France, ni l'Atlantique, ni je ne sais quoi d'autre, ne suffiront à nous éloigner.

Merci à mes parents qui ont toujours été (et sont toujours) sûrs que je réussirai n'importe quoi. Sans vous et votre confiance, c'est certain, je n'en serai pas là aujourd'hui. Merci de vous être intéressés, pour de vrai, à mon sujet de recherche, d'avoir même lu (et corrigé !) mon manuscrit. Et merci encore d'avoir compris qu'après tout ça, je choisisse une autre voie, et de continuer à me soutenir.

Et pour finir, merci à tous les autres ! À Yoan pour son code, à Robin pour beaucoup de bons moments, à Léontine pour ses plans toujours originaux, à Matthias pour sa capacité infinie à m'écouter râler... et à tous ceux que j'ai probablement oublié.

# Table des matières

<b>Remerciements</b>	<b>i</b>
<b>Introduction</b>	<b>1</b>
0.1 ARN/ADN : si proches, si différents . . . . .	1
0.2 Importance des ARNs structurés en biologie . . . . .	2
0.3 Outils : piège optique, molécules uniques . . . . .	5
<b>1 Dispositif expérimental</b>	<b>7</b>
1.1 Piège optique . . . . .	7
1.1.1 Principe du piégeage optique . . . . .	7
1.1.2 Dispositif utilisé : le double piège optique . . . . .	8
1.2 Échantillons biologiques . . . . .	10
1.2.1 Vision générale . . . . .	10
1.2.2 Préparation des échantillons . . . . .	12
1.2.3 Capture et étirement d'une construction . . . . .	12
1.3 Calibration . . . . .	13
<b>2 Comparaison des hairpins ARN / ADN</b>	<b>15</b>
2.1 Les hairpins d'ARN . . . . .	15
2.2 Étude de hairpins en molécule unique . . . . .	16
2.2.1 Les constructions . . . . .	16
2.2.2 Description des expériences réalisées . . . . .	17
2.2.3 Modèle hors d'équilibre pour le changement d'état des hairpins . . . . .	24
2.3 Hairpins under tension : RNA vs DNA . . . . .	29
2.3.1 Résumé des résultats obtenus . . . . .	29
2.3.2 Results . . . . .	29
2.3.3 Discussion . . . . .	35
2.3.4 Conclusion . . . . .	39
2.3.5 Supplementary Information . . . . .	41
<b>3 Interaction ARN / hélicases</b>	<b>47</b>
3.1 Les protéines hélicases . . . . .	47
3.1.1 Exemples de protéines à motif DEAD . . . . .	48
3.1.2 Étude des protéines hélicases en molécule unique . . . . .	50

3.2	Activité des protéines DEAD par mesures de forces . . . . .	50
3.2.1	Les constructions pour l'étude de DbpA . . . . .	50
3.2.2	Conditions expérimentales . . . . .	51
3.3	Les résultats obtenus . . . . .	53
3.3.1	Expériences à vitesse constante . . . . .	53
3.3.2	Expériences complémentaires . . . . .	60
3.4	Conclusion . . . . .	61
<b>4</b>	<b>Surétirement d'hybrides ARN/ADN</b>	<b>65</b>
4.1	L'ADN sous haute force : le surétirement . . . . .	65
4.1.1	Les débuts . . . . .	65
4.1.2	Les différents mécanismes de surétirement de l'ADN . . . . .	66
4.1.3	Modèle du « twistable worm » pour le surétirement de l'ADN. . . . .	69
4.1.4	Le cas de l'ARN . . . . .	70
4.2	Étude de molécules hybrides ARN/ADN en surétirement . . . . .	73
4.2.1	Motivation de l'étude . . . . .	73
4.2.2	Protocole pour la réalisation des constructions moléculaires . . . . .	76
4.2.3	Description d'une courbe de surétirement typique . . . . .	76
4.2.4	Analyse énergétique pour l'estimation de la force du plateau de peeling . . . . .	84
4.3	Caractérisation expérimentale . . . . .	89
4.3.1	Effet de la vitesse sur le mode de surétirement observé . . . . .	89
4.3.2	Effet du tampon expérimental sur le surétirement . . . . .	89
4.3.3	Discussion . . . . .	94
4.4	Perspectives . . . . .	99
	<b>Conclusion</b>	<b>101</b>
	<b>ANNEXE Calibration du piège optique : modèle de Langevin</b>	<b>103</b>
	<b>ANNEXE Coupling of proteins to microspheres</b>	<b>107</b>
	<b>ANNEXE Hairpins under tension : RNA versus DNA</b>	<b>111</b>
	<b>ANNEXE Test en FRET de l'activité de DbpA</b>	<b>121</b>
	<b>ANNEXE Analyse énergétique dans le cas des « bulles fondues »</b>	<b>125</b>
	<b>Bibliographie</b>	<b>134</b>

# Table des figures

0.1	Structure schématique des polyacides nucléiques . . . . .	2
0.2	Représentation de Lewis des pentoses . . . . .	3
0.3	Représentations schématiques d'une hélice d'ADN de type B . . . . .	4
0.4	Structures A et B des doubles hélices. . . . .	4
0.5	Faces d'interaction de Westhof. . . . .	5
0.6	Schéma d'un ARN de transfert . . . . .	6
1.1	Représentation schématique de l'action d'un laser sur une bille. . . . .	8
1.2	Schéma simplifié du montage du double piège optique. . . . .	9
1.3	Schéma simplifié du piégeage d'un échantillon . . . . .	11
2.1	Représentation schématique d'un hairpin . . . . .	16
2.2	Représentation schématique des quatre hairpins synthétisés. . . . .	17
2.3	Simulation en Dynamique Moléculaire des hairpins RNA10 et RNA18 . . . . .	18
2.4	Courbe force/position caractéristique d'un hairpins, et définition des grandeurs mesurée. . . . .	19
2.5	Exemples de distributions des forces d'ouverture et de fermeture. . . . .	21
2.6	Exemples de courbes force/position et ajustements WLC pour un ADN double brin et un hybride ADN/ARN . . . . .	23
2.7	Courbes théoriques d'étirement d'acides nucléiques simple brin de 36 ou 44 bases (modèle WLC) . . . . .	25
2.8	Schematic representation of the energy landscape used in our theoretical description . . . . .	26
2.9	Force versus extension curves at 50 nm/s for the four different hairpins . . . . .	31
2.10	Force versus extension curves for DNA10 at 50, 150, 300 and 450 nm/s . . . . .	31
2.11	Histograms of the hysteresis measured on RNA10 . . . . .	33
2.12	Histograms of the hysteresis measured at 150 nm/s on the four different hairpins . . . . .	34
2.13	Spontaneous transitions between folded and unfolded hairpin states . . . . .	36
2.14	Histograms of the hysteresis measured on the four hairpins at different velocities . . . . .	42
2.15	Unfolding energy landscapes calculated for the DNA and RNA hairpins . . . . .	44
3.1	Modèle 3D de la protéine Ded1. . . . .	48
3.2	Représentation de la fixation de DbpA au hairpin ARN H92 . . . . .	49
3.3	Séquence pour l'étude de l'activité de DbpA en molécule unique. . . . .	51



3.4	Séquence pour l'étude de l'activité de DbpA en molécule unique . . . . .	52
3.5	Courbe de réponse en force des constructions pour l'étude de DbpA en l'absence de la protéine . . . . .	54
3.6	Agrandissement de courbes de réponse en force des constructions pour l'étude de DbpA en l'absence de la protéine . . . . .	56
3.7	Courbe de réponse en force des constructions pour l'étude de DbpA en présence de la protéine . . . . .	57
3.8	Représentation schématique des différents profils d'ouverture en présence de DbpA . . . . .	58
3.9	Exemples des différents profils d'ouverture en présence de DbpA . . . . .	59
3.10	Probabilités correspondant à l'apparition des différents profils observés en absence et en présence de DbpA . . . . .	60
3.11	Représentation schématique du test de l'activité hélicase de DbpA sur un substrat fermé . . . . .	62
4.1	Représentation schématique des différents modes de surétirement . . . . .	67
4.2	Surétirement de l'ADN : bulles fondues <i>versus</i> peeling . . . . .	68
4.3	Exemple de surétirement d'un ADN double brin . . . . .	71
4.4	Exemple de surétirement d'un ARN double brin . . . . .	72
4.5	Schéma de la structure secondaire de l'ARN ribosomique 23S . . . . .	74
4.6	Schéma de la configuration de peeling pour l'étude de structures secondaires . . . . .	75
4.7	Schéma du protocole expérimental pour la fabrication des constructions hybrides à une extrémité libre . . . . .	77
4.8	Exemple de surétirement d'un double brin hybride ADN/ARN . . . . .	78
4.9	Exemple des différents modes de surétirement observés pour les hybrides ARN/ADN . . . . .	82
4.10	Exemple de réhybridation partielle de l'hybride ARN/ADN . . . . .	83
4.11	Estimation théorique de la force du plateau de peeling . . . . .	86
4.12	Effet de la vitesse sur le surétirement d'hybrides ARN/ADN . . . . .	90
4.13	Température de dénaturation d'hybrides ARN/ADN . . . . .	91
4.14	Exemples de courbes de réponse en force lors de l'étirement d'hybrides ARN/ADN dans différents tampons expérimentaux . . . . .	92
4.15	Force du plateau de peeling mesurée pour un étirement à 10 nm/s dans différents tampons expérimentaux . . . . .	93
4.16	Influence de la concentration en sels sur le plateau de surétirement . . . . .	95
4.17	Profil du plateau de peeling lors de plusieurs étirements successifs d'une même molécule . . . . .	96
4.18	Énergie des structures secondaires possibles dans le simple brin ARN généré par peeling . . . . .	97
4.19	Spectre de calibration du piège optique . . . . .	105
4.20	Oligomères utilisés pour le test en FRET de l'activité de DbpA en solution . . . . .	122
4.21	Spectres de fluorescence pour le test d'activité de DbpA en solution . . . . .	123

*TABLE DES FIGURES*

vii

4.22 Schema d'une « bulle fondue » . . . . . 125



# Introduction

*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.*

---

James. D. Watson  
and Francis. H. C. Crick

## 0.1 ARN/ADN : si proches, si différents

Les acides ribonucléique (ARN) et désoxyribonucléique (ADN) sont deux macromolécules de structure similaire, formées d'une chaîne plus ou moins longue de nucléotides. Les nucléotides sont des molécules comportant trois unités distinctes : une base azotée, un monosaccharide à cinq carbones (pentose), et un groupement phosphate (figure 0.1). Les polynucléotides d'ADN et d'ARN sont constitués en formant une chaîne de nucléotides, avec une liaison entre le pentose d'un nucléotide et le groupe phosphate du nucléotide suivant (figure 0.1). La chaîne ainsi obtenue est orientée. L'extrémité phosphate est dénommée 5', et l'extrémité sucre 3'. La seule différence entre les nucléotides composant l'ADN et ceux composant l'ARN vient de la nature du pentose : il s'agit d'un ribose pour l'ARN (figure 0.2), alors que l'ADN contient un désoxyribose (figure 0.2). Dans le désoxyribose, un groupement -OH est remplacé par un atome -H. Il existe différentes bases azotées, distinguées par leur nombre de cycles carbonés et par les groupements portés par ceux-ci. Dans la nature, trois sont retrouvées à la fois dans l'ADN et dans l'ARN. Il s'agit de la cytosine (C), de la guanine (G), et de l'adénine (A). La thymine (T) n'est présente que dans les molécules d'ADN, alors que l'uracile (U) se trouve dans l'ARN. Ces bases sont réparties en deux catégories : les pyrimidines, composées d'un seul cycle carboné (C, T, et U), et les purines, qui en contiennent deux (G et A).

Les chaînes de nucléotides décrites forment des acides nucléiques simple brin. C'est dans cet état que se trouve en général l'ARN contenu dans les cellules. L'ADN est par contre présent sous la forme d'une double hélice, formée par l'enroulement de deux chaînes de nucléotides (figure 0.3). Cela est rendu possible par l'établissement de liaisons hydrogène entre les bases azotées complémentaires de chaque chaîne. Une pyrimidine va établir des liaisons H avec une

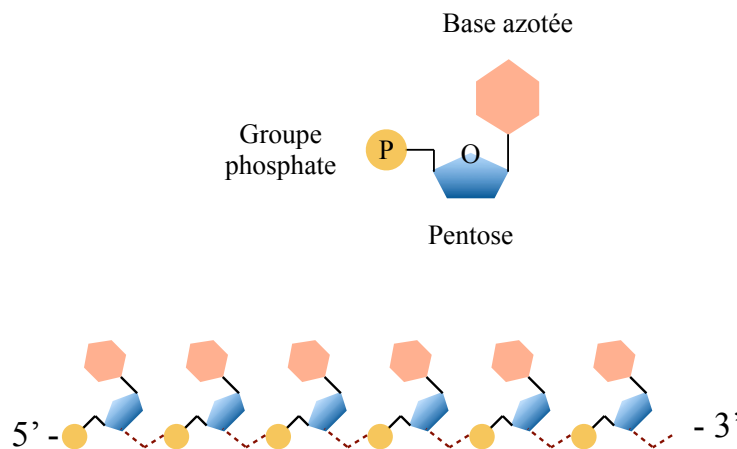


Figure 0.1 **Structure schématique des polyacides nucléiques.** Chaque nucléotide est composé d'un groupe phosphate, d'un pentose, et d'une base azotée (haut). Les chaînes sont obtenues en formant une liaison entre le pentose d'un nucléotide et le groupe phosphate du nucléotide suivant (bas).

purine, selon les couples canoniques suivants : A et T sont reliées par deux liaisons hydrogène, C et G par trois.

Lorsque deux simples brins d'ARN complémentaires sont mis en présence, une double hélice peut aussi se former spontanément. La thymine est alors remplacée par l'uracile dans la liaison avec l'adénine. De la même façon, il est possible de former une hélice hybride composée d'un brin d'ADN et d'un brin d'ARN. Malgré les structures chimiques très proches de l'ADN et de l'ARN, les hélices obtenues par appariement de deux brins d'ADN sont différentes de celles obtenues avec deux brins d'ARN, ou des hélices hybrides (figure 0.4). En conditions physiologiques, les doubles hélices d'ADN sont des hélices droites de type B, de rayon 1 nm et dont le pas hélical vaut 3.4 nm pour 10 paires de bases. Dans les mêmes conditions, l'ARN confère à la double hélice une structure droite de type A, plus compacte, avec 2.8 nm de pas hélical et 11 paires de bases par tour, et un rayon de 1.3 nm. Les hélices hybrides sont aussi de type A. Cette différenciation a une cause stéréochimique : la présence du groupement -OH en 2' dans les brins d'ARN est incompatible avec la formation d'une hélice B.

## 0.2 Importance des ARNs structurés en biologie

Les récents progrès des techniques de séquençage ont permis une exploration inédite du contenu en ARN des cellules biologiques. Bien que seuls 2% environ du génome humain correspondent à une séquence d'ARN messenger codant pour une protéine, la majorité de l'ADN est transcrit en ARN. Les ARNs non codants, dont la taille peut varier de quelques bases à plusieurs centaines, occupent des fonctions variées, allant de la catalyse à la régulation de l'expression génique. Ces fonctions sont rendues possibles par la capacité qu'a l'ARN, transcrit en

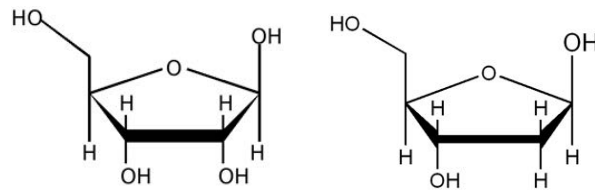


Figure 0.2 **Structure chimique des pentoses.** Gauche : représentation de Lewis du ribose, constituant des nucléotides d'ARN. Droite : représentation de Lewis du désoxyribose, constituant des nucléotides d'ADN.

simple brin, de se structurer dans l'espace. Cette structuration est en général décomposée en structure secondaire et tertiaire. Elle est déterminée en premier lieu par la séquence du polymère d'ARN, combinaison des quatre nucléotides standards A,U,G, et C. Ceux-ci pouvant former des liaisons hydrogène doubles ou triples dans les paires Watson-Crick classiques (A=U ; G≡C), un simple brin d'ARN peut appairer localement deux zones complémentaires de sa séquence. Cette structuration secondaire "canonique" induit l'apparition d'hélices double brins qui mettent en contact deux parties potentiellement très distantes dans la séquence linéaire. Ces liaisons "canoniques" n'impliquent cependant que 60% des bases dans les ARNs structurés. Les autres peuvent interagir par des liaisons hydrogène non canoniques mettant en jeu les autres faces des bases, à savoir la face "Hoogsteen" et la face "sucre" [2]. Chaque interaction pouvant se faire en conformation *cis* ou *trans*, 12 géométries sont possibles, et toutes ont été observées. Un appariement fréquent dans les structures secondaires d'ARN est le G-U (paire « wobble », ou bancale), qui ne fait pas partie des paires canoniques, mais implique toutefois les faces Watson Crick des deux nucléotides.

Un niveau d'interactions supplémentaire forme des structures tertiaires ou supra-hélicoïdales, où des régions non-appariées sont mises en contact, ou mises en contact avec des hélices. Ces motifs tertiaires reposent le plus souvent sur l'empilement des bases (stacking), favorisé d'un point de vue énergétique par la géométrie plane des bases azotées. Les ARNs de transfert sont un exemple classique de structuration secondaire et tertiaire (figure 0.6). Parmi les 76 bases de leur séquence totale, 41 sont impliquées dans une structure hélicoïdale, et 72 dans une interaction d'empilement. Le hairpin, structure secondaire minimale, fait l'objet du chapitre 3 de ce travail.

Les diversités structurelle et fonctionnelle présentées par les molécules d'ARN sont un argument fort dans le modèle du "monde ARN". Dans cette théorie, l'origine de la vie sur Terre remonterait à l'émergence de l'ARN, polymère capable d'assumer à la fois les fonctions de support de l'information génétique et les fonctions catalytiques vitales à une cellule.

Les fonctions catalytiques de structures d'ARN complexes sont bien illustrées par le cas du ribosome. Cette machinerie cellulaire sophistiquée est nécessaire à la vie de toute cellule : le ribosome assure la fonction de traduction de l'ARN messager en protéine. Il est lui-même composé d'un assemblage de plusieurs longs brins d'ARN fortement structurés et de nombreuses protéines. Les protéines assurent la structuration et le maintien du complexe, et la

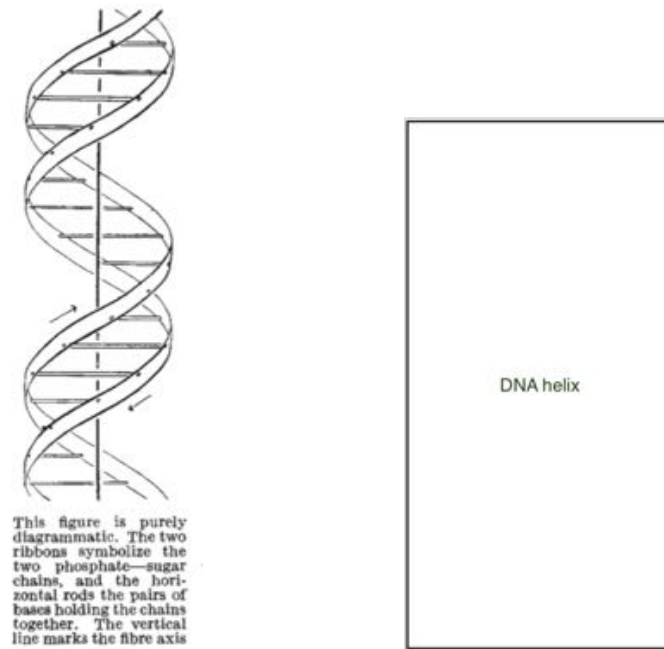


Figure 0.3 **Représentations schématiques d'une hélice d'ADN de type B.** Gauche : première représentation, proposée par J. D. Watson et F. H. C. Crick en 1953 [1]. Figure reproduite avec la permission de Macmillan Publishers Ltd : Nature, copyright 1953. Droite : Sur cette deuxième représentation figurent aussi les différents constituants des nucléotides (P : groupe phosphate ; D : désoxyribose ; A/T/G/C : base azotée).

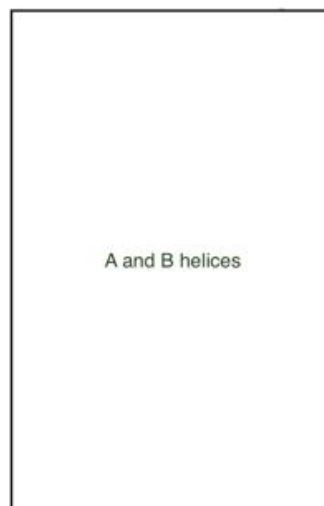


Figure 0.4 **Structures A et B des doubles hélices.** Gauche : structure A d'une double hélice adoptée par l'ARN et les hybrides en conditions physiologiques. Droite : structure B d'une double hélice adoptée par l'ADN en conditions physiologiques.

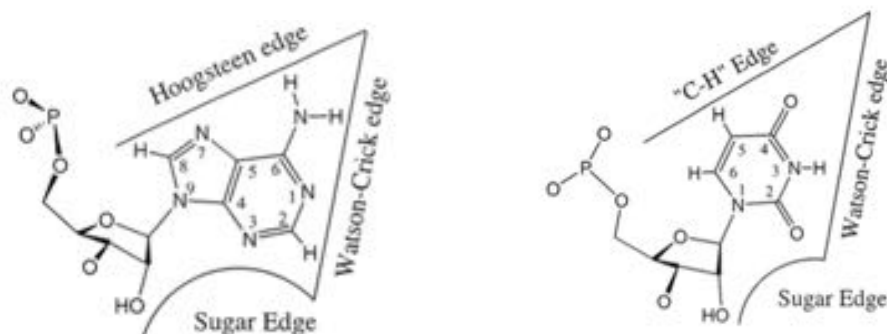


Figure 0.5 **Représentation schématique des différentes faces d'interaction de Westhof sur des bases ARN.** En plus de la face Watson-Crick impliquée dans les liaisons canoniques, les bases azotées peuvent interagir par leur face « sucre » ou leur face « Hoogsteen » (appelée face « C-H » dans les pyrimidines). Figure issue de Leontis *et al.*, *Nucleic Acids Research*, 2002, avec la permission de Oxford University Press [2].

fonction catalytique du ribosome (fonction peptidyl transferase) est entièrement assurée par des résidus ARN. Le ribosome sera mentionné à plusieurs reprises dans ce travail, en particulier dans les chapitres 3 et 4.

### 0.3 Outils : piège optique, molécules uniques

Les expériences classiques de biologie moléculaire ont permis d'apprendre beaucoup sur la structure des acides nucléiques, ainsi que sur leur interaction avec certaines protéines. Ces expériences ne permettent toutefois que l'étude de comportements moyens, évalués sur une population. Les techniques sur molécules uniques développées plus récemment permettent au contraire d'étudier le comportement individuel d'une molécule d'ADN ou d'ARN, seule ou en interaction avec une protéine. Il devient alors possible d'accéder à des données supplémentaires, telles que la variabilité entre molécules, le mécanisme d'action d'une protéine, les différentes étapes d'une réaction...

Nous avons ici mis à profit la technique de mesure de force par piégeage optique pour l'étude de structures secondaires d'ARN. Différents aspects ont été abordés. Dans un premier temps, nous nous sommes intéressés à une structure secondaire modèle, le hairpin : une étude comparative de la réponse dynamique des hairpins d'ARN et d'ADN a été réalisée. La question de l'interaction entre les structures secondaires d'ARN et les protéines a ensuite été abordée. Dans le cadre d'une recherche plus large sur l'assemblage du ribosome, l'action d'une hélicase sur un fragment d'ARN ribosomique a été étudiée. Enfin, une méthode d'étude des interactions ARN/protéines en configuration de surétirement est en cours de développement. Cette technique permettrait l'étude de plusieurs protéines ribosomiques sur leur ARN cible dans des conditions proches de celles présentes *in vivo*.



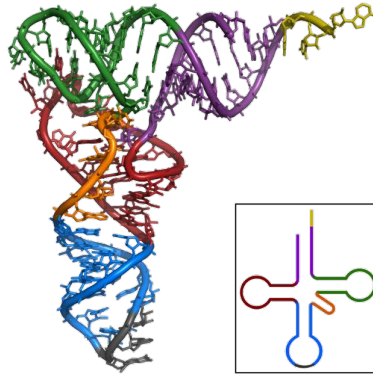


Figure 0.6 **Représentation de la structuration secondaire et tertiaire d'un ARN de transfert (ARNt).** L'encart présente la structure secondaire seule. Les boucles rouge (boucle D), verte (boucle TΨC), et orange (boucle variable) présentent des interactions tertiaires fortes qui donnent à l'ARNt sa forme caractéristique. La boucle bleu contient l'anti-codon (gris), qui vient interagir avec l'ARN messenger à traduire. La tige violette est liée à un acide aminé par son extrémité jaune. La tige violette et les tiges à la base de chaque boucle (rouge, verte, et bleue) forment des structures hélicoïdales [3].



# Chapitre 1

## Dispositif expérimental pour l'étude de molécules uniques par piégeage optique

*J'accorde une grande importance au matos.*

---

Michel, par Bruno Podalydès,  
dans « Comme un avion »

### 1.1 Piège optique

#### 1.1.1 Principe du piégeage optique

Le piégeage optique provient de l'interaction entre des photons et un objet. Ce phénomène se comprend bien en considérant le cas limite de l'optique géométrique, lorsque la longueur d'onde du laser est petite devant la taille des objets à piéger. Ce n'est pas rigoureusement le cas dans cette étude puisque les billes ont une dimension comparable à la longueur d'onde, mais cette vision permet une compréhension assez intuitive du piégeage optique. Lorsqu'un faisceau lumineux rencontre une interface, l'interaction est de deux types : d'une part les photons exercent une pression de radiation sur l'interface ; d'autre part, lorsqu'ils sont réfractés entre les deux milieux d'indice différent, ils transfèrent une certaine quantité de mouvement à l'interface. Les deux milieux sont ici une bille (silice ou polystyrène) et son milieu d'immersion (solution aqueuse). La force résultante peut être décomposée en deux contributions (Figure 1.1) : l'une dirigée dans le sens de propagation de la lumière (pression de radiation) et l'autre vers les fortes intensités lumineuses (force de gradient due à la réfraction). Dans le cas d'un faisceau homogène, la force de gradient est négligeable devant la pression de radiation et une bille est entraînée par le flux lumineux [4]. Si par contre le faisceau est fortement focalisé, la force de gradient devient dominante, et la bille est attirée puis piégée près du point focal [5]

(légèrement décalée par l'effet de la pression de radiation). Un faisceau laser d'une puissance de l'ordre du Watt fortement focalisé permet de créer un piège que l'on peut assimiler à un puits de potentiel équivalent à celui créé par un ressort, et dont la constante de raideur serait de l'ordre de 250 pN/ $\mu\text{m}$  pour une bille de polystyrène de taille micrométrique.

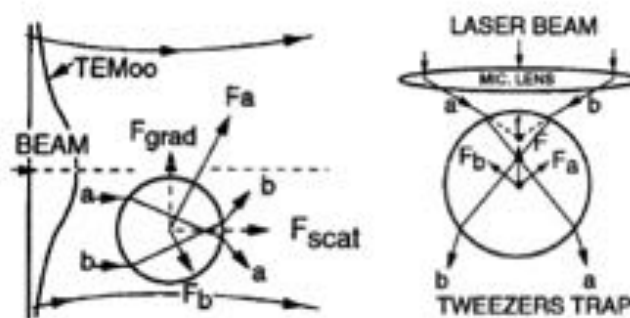


Figure 1.1 **Représentation schématique de l'action d'un laser sur une bille.** La réfraction des faisceaux  $a$  et  $b$  représentés génère des forces dénotées respectivement  $F_a$  et  $F_b$ . La force résultante  $F$  peut être décomposée en une composante colinéaire au faisceau lumineux ( $F_{scat}$ ) et une qui lui est orthogonale ( $F_{grad}$ ). Gauche : effet d'un laser non focalisé sur une bille. La bille est ramenée vers le centre du faisceau par la force de gradient  $F_{grad}$ , et déplacée dans le sens de propagation de la lumière par la pression de radiation  $F_{scat}$ . Droite : effet d'un laser fortement focalisé sur une bille. La force de gradient est très dominante, et la bille est piégée légèrement sous le point focal  $f$  du laser. Figure issue de Ashkin, *Biophysical Journal*, 1992.[5].

La théorie généralisée de Lorenz-Mie permet de décrire exactement les forces générées au niveau du piège optique. Elle repose sur l'étude de la diffusion de la lumière par un objet de forme quelconque [6]. Son développement ne sera pas exposé ici.

### 1.1.2 Dispositif utilisé : le double piège optique

Dans le cas d'expériences impliquant des objets biologiques, la longueur d'onde du laser de piégeage doit être choisie de façon à limiter les dégâts sur l'échantillon. La fenêtre optimale est située dans l'infra-rouge, entre 800 et 1200 nm. Le laser utilisé ici est de type Nd : YVO4 (cristal d'yttrium vanadate), de longueur d'onde 1064 nm. Celui-ci présente une stabilité suffisante à la fois en puissance (1% sur quelques minutes) et en pointage ( $5\mu\text{rad}/^\circ\text{C}$ ) pour obtenir une bonne précision sur la mesure de force. Afin de limiter les fluctuations du faisceau dues aux flux d'air, le système est isolé par une protections en PVC.

Les deux pièges sont générés par le même faisceau laser (Figure 1.2) : ceci est obtenu en décomposant le faisceau selon deux polarisations orthogonales. Chaque polarisation suit ensuite un chemin optique différent. La première forme un faisceau qui est réfléchi sur un ensemble de miroirs fixes, et constitue un piège immobile, à une position bien définie. Le second faisceau va former un piège mobile, dont la position doit être précisément contrôlée.

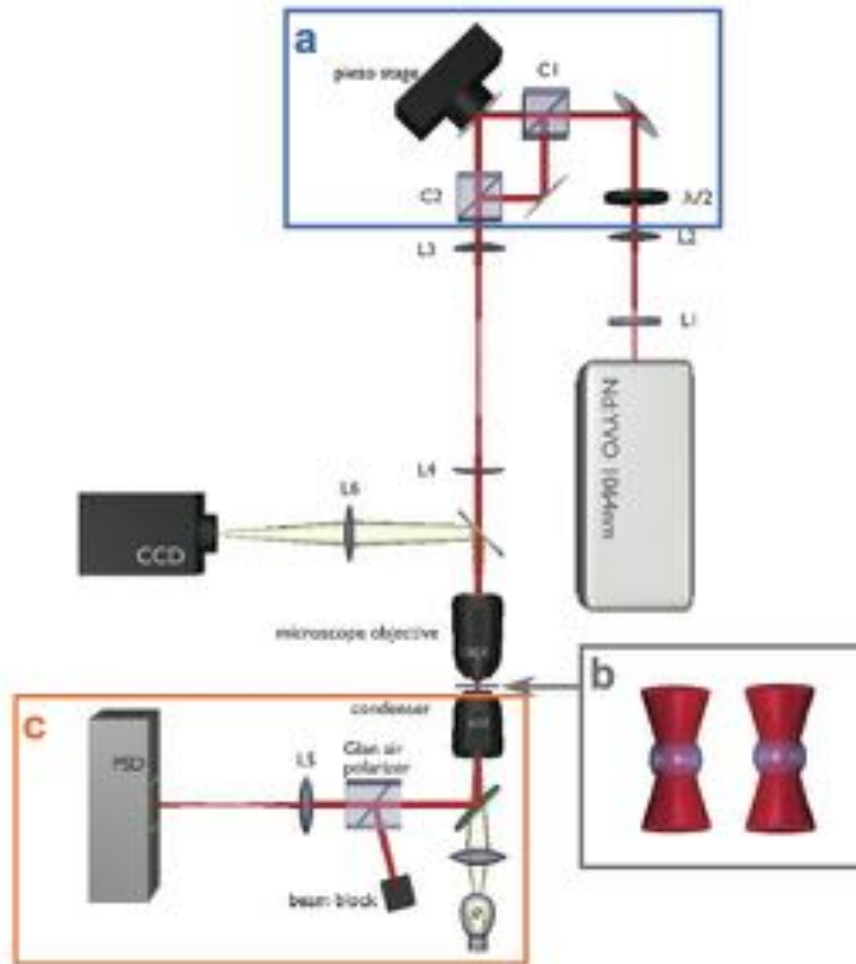


Figure 1.2 Schéma simplifié du montage du double piège optique. a. Création de deux faisceaux de polarisations orthogonales (C1), l'un mobile (piezo stage) et l'autre fixe. Le module acousto-optique est inséré entre le cube C1 et le miroir mobile. Recombinaison des faisceaux (C2). b. Création de deux pièges optiques par focalisation à travers un objectif de microscope de forte ouverture numérique (NA=1.4). c. Mesure de force par détection de la position de la bille piégée dans le potentiel fixe (détecteur PSD). Le faisceau de polarisation correspondant au piège mobile est éjecté (Glan air polarizer). Figure reproduite avec la permission de Mangeol et al., *Review of Scientific Instruments*, 2008., Copyright 2008, AIP Publishing LLC [7].

Pour cela, un miroir monté sur une cale piézo-électrique est placé sur le chemin optique de ce faisceau. Grâce à ce dispositif, le piège mobile peut être éloigné jusqu'à 15  $\mu\text{m}$  du piège fixe, avec une précision de l'ordre du nanomètre. La génération des deux pièges par le même faisceau laser permet d'augmenter significativement la précision du montage : comme les deux pièges se déplacent ensemble lors de perturbations, leur écartement relatif reste stable.

La focalisation qui permet la création des pièges optiques est réalisée par le premier objectif du microscope. Son ouverture numérique doit donc être la plus grande possible. L'objectif installé est un Nikon 100X d'ouverture numérique 1.4, optimisé dans l'infrarouge. La focalisation est optimale près de la lamelle, c'est donc là que les billes seront piégées. Le second objectif collecte la lumière émergente afin de permettre la mesure de force. L'ouverture numérique doit être assez grande pour collecter le maximum du faisceau émergent, et ainsi améliorer la précision sur la mesure. Il s'agit ici d'un objectif Olympus 60X à ouverture numérique 1.2.

La mesure de force est effectuée sur la polarisation correspondant au piège fixe, par détection en pupille arrière. Il s'agit d'évaluer les variations de position du centre de gravité de l'intensité lumineuse sur la pupille arrière de l'objectif. Ces variations sont induites par les déplacements de la bille située dans le piège. Comme le piège est équivalent à un puits de potentiel harmonique, la relation  $\vec{F} = -k_p \vec{x}$  permet d'obtenir la force correspondante. Si le piège a été correctement calibré, la force peut ainsi être mesurée avec une précision inférieure au piconewton. Dans le montage, une combinaison de lentilles permet de faire l'image de la pupille arrière du microscope sur le détecteur PSD (Position Sensitive Detector).

En pratique, la polarisation n'est pas exactement conservée à la traversée des objectifs du microscope. Par conséquent les faisceaux peuvent interférer, ce qui perturbe la mesure de force. Pour éviter ce problème, un modulateur acousto-optique est inséré sur le chemin du laser correspondant au piège mobile, afin de décaler sa fréquence [7]. Les interférences ne sont pas réellement supprimées, mais elles varient trop rapidement pour être détectées.

L'imagerie de l'échantillon sur une caméra CCD, par le même chemin optique, permet de visualiser les billes sur un écran.

## 1.2 Échantillons biologiques

### 1.2.1 Vision générale

Les échantillons biologiques étudiés ici sont des polymères d'acides nucléiques, ADN ou ARN, contenant plusieurs milliers de bases. Il en résulte des molécules longues de quelques micromètres et de section nanométrique. Dans les solutions aqueuses utilisées, ces molécules se présentent sous forme de pelotes sub-micrométriques. Elles ne sont alors ni directement observables en microscopie optique, ni manipulables par piégeage optique. Afin de permettre leur étude, les extrémités de chaque molécule vont être fixées à des billes micrométriques. Lorsque les billes sont placées dans les pièges optiques, il est possible de les séparer d'une distance contrôlée, et de mesurer la force qui est appliquée à la molécule (Figure 1.3).

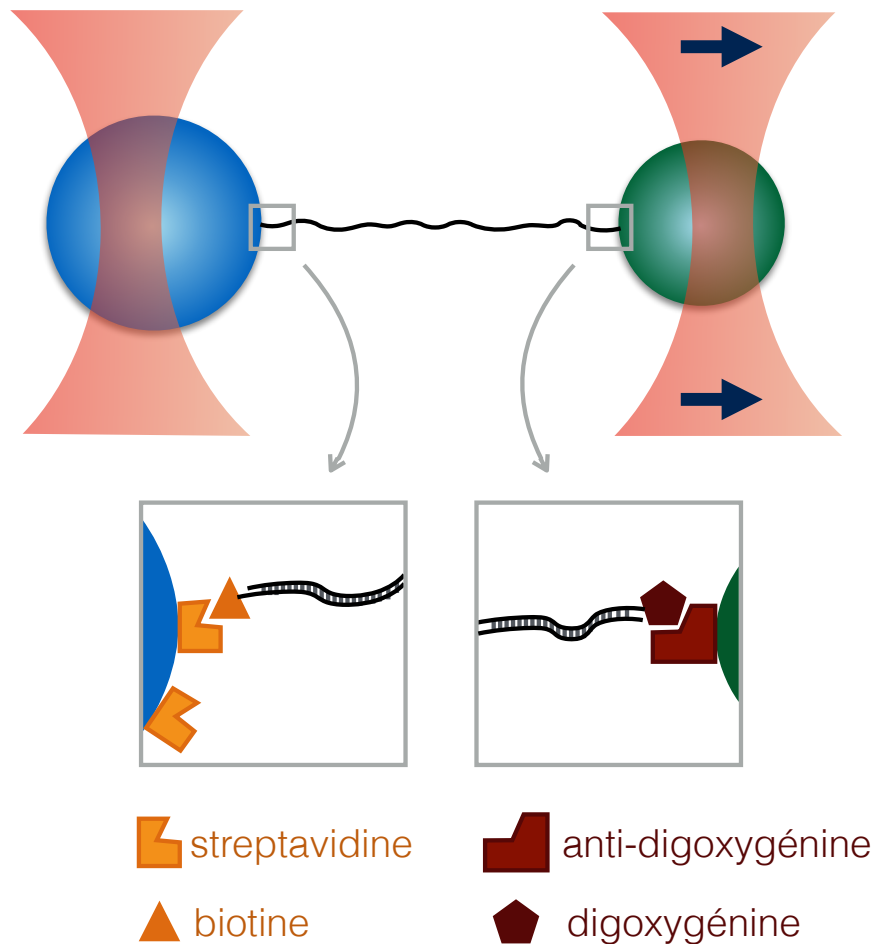


Figure 1.3 **Schéma simplifié du piégeage d'un échantillon.** Les constructions d'acides nucléiques sont fonctionnalisées à leurs extrémités par une biotine ou une digoxygénine (panneau inférieur). Ces fonctions peuvent former des liaisons avec des molécules de biotine ou d'anti-digoxygénine, respectivement. Par ce biais, les constructions d'acides nucléiques peuvent être fixées à des billes micrométriques fonctionnalisées par de la biotine ou de l'anti-digoxygénine. Les billes micrométriques peuvent être attrapées avec les pièges optiques. Il est alors possible d'exercer une force sur les constructions moléculaires en déplaçant l'un des pièges (panneau supérieur).

### 1.2.2 Préparation des échantillons

L'attachement des polymères d'acides nucléiques aux billes est réalisé par des couplages ligand/récepteur. Certains nucléotides proches des extrémités sont fonctionnalisés avec l'un des éléments du couple, l'autre élément étant fixé chimiquement aux billes. Selon les expériences, les attachements peuvent être simples ou multiples. Les couples utilisés ici sont biotine/streptavidine et digoxygénine/anti-digoxygénine (Figure 1.3). Le protocole de fonctionnalisation des billes (par de la streptavidine ou de l'anti-digoxygénine) est dérivé du protocole d'un kit commercial, et est donné en annexe (Covalent coupling of proteins to carboxylated microspheres).

Les billes fonctionnalisées et les constructions d'acides nucléiques sont stockées séparément. Chaque expérience commence donc par une étape où les billes et les constructions sont mises en présence. Afin de faciliter la rencontre des différents éléments, l'échantillon est centrifugé (6 minutes à 30g). Les billes sont ainsi concentrées au fond du tube, ce qui augmente la probabilité qu'une construction se fixe à deux billes. Après centrifugation, le mélange est laissé à incuber pendant une heure à température ambiante. La concentration de constructions est choisie assez faible pour minimiser les attachements multiples, mais assez grande pour que le temps de recherche d'un « duplexe » (deux billes rattachées par une construction) ne soit pas trop long. Typiquement, 0.5  $\mu\text{L}$  de solution de constructions à 5 ng/ $\mu\text{L}$  sont mis à incuber avec 1  $\mu\text{L}$  de solution de billes à  $10^5$  billes/ $\mu\text{L}$ . Le temps de recherche d'un duplexe est alors en général inférieur à 5 minutes. Après incubation, la préparation est diluée dans une solution adaptée à la manipulation, jusqu'à un volume final de 20  $\mu\text{L}$ . Afin d'éviter les interactions non spécifiques entre billes, il est important que le pH soit proche de 7.4. Les tampons utilisés pour les différentes expériences réalisées seront précisés dans les chapitres correspondants. L'échantillon est ensuite déposé entre deux lamelles de microscope (épaisseur #1,  $\sim 160 \mu\text{m}$ ) séparées par des bandes de parafilm. La cellule ainsi créée est scellée à la cire (1/3 vaseline, 1/3 paraffine, 1/3 lanoline en masse), afin d'éviter l'évaporation de la solution [8].

### 1.2.3 Capture et étirement d'une construction

La cellule contenant l'échantillon est ensuite placée entre les deux objectifs. La première étape consiste à repérer un duplexe. Afin de le capturer, le piège mobile est initialement éloigné du piège fixe d'une distance légèrement supérieure à la somme des rayons des deux billes. Une fois piégé, le duplexe est amené à une distance de 5  $\mu\text{m}$  au dessus de la lamelle inférieure de la cellule, distance qui offre à la fois une bonne qualité de piégeage et de faibles interactions hydrodynamiques entre les billes et la surface. Le déplacement et la vitesse du piège mobile sont ensuite choisis en fonction de la mesure à réaliser. La fréquence d'acquisition de la mesure doit être ajustée. Les constructions ont été conçues de manière à pouvoir être utilisées plusieurs fois : lorsque la construction a été étirée, il est possible de rapprocher les pièges pour qu'elle reprenne sa conformation initiale. Une même construction peut donc être étirée et relaxée jusqu'à ce qu'elle (ou ses attachements) se rompe et que les billes se séparent.

## 1.3 Calibration

Après chaque mesure, une calibration du piège fixe (utilisé pour la mesure de force) est effectuée, afin de déterminer sa raideur. Celle-ci dépend notamment du diamètre, de la forme exacte et de l'indice diélectrique des billes. Toutes ces grandeurs étant sujettes à fluctuations, la raideur peut varier d'une mesure à l'autre. Une calibration est donc effectuée après chaque expérience. La calibration est basée sur la mesure du mouvement brownien d'une bille dans le piège. Ce mouvement est décrit par l'équation de Langevin qui, passée dans le domaine de Fourier, permet de relier l'amplitude du mouvement à la raideur du piège et au coefficient de conversion entre la tension mesurée et la force. Afin de prendre en compte l'interaction hydrodynamique entre la bille et la lamelle (séparées de  $5 \mu\text{m}$ ), le coefficient de frottement de la bille est corrigé. Le spectre en fréquence du mouvement brownien est mesuré grâce à un analyseur de spectre, et son ajustement par la fonction théorique permet d'obtenir la raideur  $k_p$  (en  $\text{pN}/\mu\text{m}$ ) et le coefficient de conversion tension/force ( $\text{pN}/\text{V}$ ). Plus de détails concernant la calibration sont donnés en annexe (Calibration).





# Chapitre 2

## Comparaison des hairpins d'ARN avec leur équivalent ADN

*Le génie n'est plus qu'une grande aptitude à la patience.*

---

Georges-Louis Buffon

### 2.1 Les hairpins d'ARN : une structure secondaire fondamentale en biologie

Le hairpin (ou « tige-boucle ») est la forme minimale de structure secondaire que peut former un polymère d'acides nucléiques simple brin (Figure 2.1). Lorsque la séquence présente deux fragments antiparallèles complémentaires, le brin peut se replier pour permettre l'appariement de ces fragments. Une double hélice est alors localement formée, dont l'une des extrémités est fermée par une boucle simple brin. Cette structure peut se former indifféremment sur les brins d'ADN ou d'ARN. Dans les structures secondaires d'ARN plus complexes, le hairpin est le motif le plus fréquemment retrouvé, d'où l'intérêt de son étude approfondie. Un exemple de structure complexe est l'ARN ribosomique 16S, où 70% des nucléotides sont impliqués dans des structures en hairpin [9, 10].

*In vivo*, les interactions secondaires dans l'ARN sont assez stables pour que la plupart des ARNs simple brin en forment naturellement. Dans le cas du hairpin, la tige hélicoïdale et la boucle peuvent être de taille variable. A cause des répulsions stériques, la taille minimale d'une boucle d'ARN est de 3 nucléotides. Les boucles à 4 nucléotides sont particulièrement fréquentes, et peuvent être fortement stabilisées par des liaisons internes non Watson-Crick. De manière générale, les hairpins d'ARN sont très importants dans la structuration des ARN, et dans les différentes fonctions qu'ils assurent *in vivo*, ce qui recouvre notamment le contrôle de l'expression génique. La bivalence entre un état « ouvert » (ARN simple brin, absence de hairpin) et un état « fermé » (hairpin formé) leur confère la possibilité de contrôler un



Figure 2.1 **Représentation schématique d'un hairpin.** Les bases appariées constituent la tige, fermée d'un côté par une boucle de nucléotides non appariés.

processus : une des conformations l'autorise, l'autre le bloque. Ainsi, la présence d'un hairpin sur un brin d'ARN messager (ARNm) en cours de transcription peut décider de la poursuite ou de l'interruption du processus, et donc moduler l'expression du gène concerné [11, 12, 13, 14]. Cette modulation peut aussi avoir lieu au niveau des ARNm totalement transcrits. La présence d'un hairpin peut alors favoriser, ou au contraire empêcher la reconnaissance de l'ARNm par certaines protéines, avec pour effet de moduler la durée de vie de l'ARNm et donc sa probabilité de traduction [15, 16].

Les hairpins d'ADN sont beaucoup moins fréquents *in vivo*, l'ADN étant majoritairement présent sous forme double brins. Ils existent cependant, par exemple dans certains virus à ADN [17]. *In vitro*, des structures secondaires se forment spontanément dans l'ADN simple brin, et des formes présentant une activité catalytique ont pu être obtenues (déoxyribozymes, [18]).

## 2.2 Étude de hairpins en molécule unique

Afin de comparer les propriétés dynamiques plutôt que structurales des hairpins d'ARN par rapport à leurs homologues d'ADN, nous avons préparé des hairpins identiques, réalisés sous les deux formes ADN et ARN. L'étude comparative de ces quatre hairpins a fait l'objet d'une publication [19]. Je vais donc dans un premier temps présenter une description assez générale de ces hairpins, et des expériences réalisées. Les résultats particuliers obtenus de ces expériences seront présentés dans la forme sous laquelle ils ont été publiés, en anglais donc, avec de légers remaniements. Ils seront bien évidemment accompagnés d'un résumé en français. L'article complet, sous sa forme originale, est inclus aux annexes.

### 2.2.1 Les constructions

D'un point de vue thermodynamique, la longueur et la composition de la tige et de la boucle d'un hairpin jouent un rôle fondamental [20]. En modifiant la stabilité de la structure, ces éléments ont une influence directe sur son énergie libre. Nous avons réalisé deux hairpins de tige identique (13 paires de bases, séquence fixée avec  $\sim 50\%$  de GC) et de boucle de

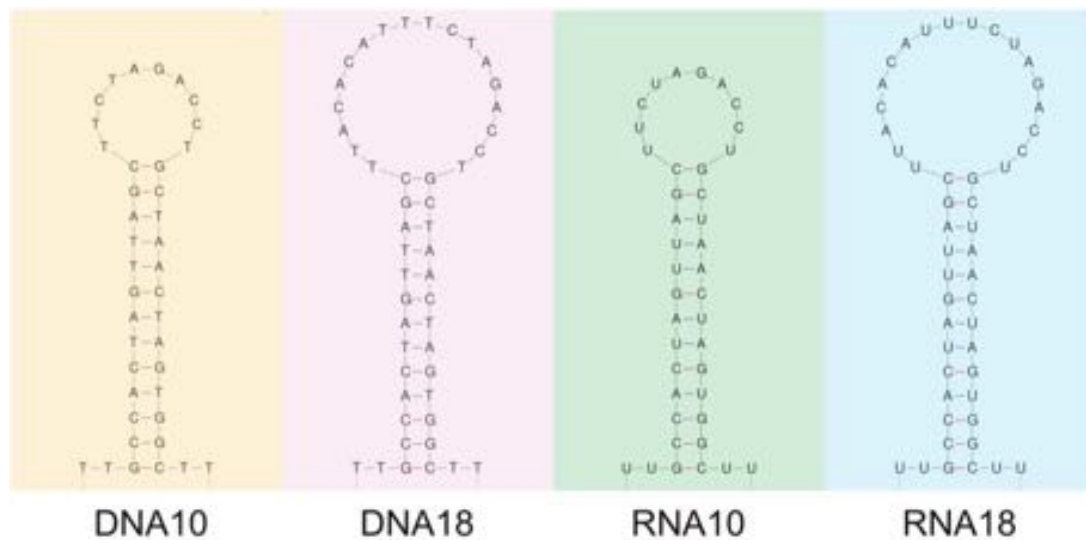


Figure 2.2 **Représentation schématique des quatre hairpins synthétisés.** Les quatre hairpins présentent une tige de séquence identique. Les boucles ont une séquence identique deux à deux, avec une longueur de 10 nucléotides pour DNA10 et RNA10, et de 18 nucléotides pour DNA18 et RNA18. Pour les deux hairpins ARN, les bases Thymines sont remplacées par des Uraciles. Figure issue de [19]

longueur différente (10 ou 18 bases). Ces hairpins ont été synthétisés en ARN et en ADN, les bases Uraciles étant remplacées par des Thymines (Figure 2.2). 4 hairpins sont donc obtenus.

Chacun est inséré entre deux poignées d'environ 2000 bases, ADN ou ARN selon la nature chimique du hairpin. Afin de rendre ces constructions manipulables avec les pinces optiques, la molécule simple brin contenant la séquence du hairpin est hybridée, de part et d'autre de cette séquence, à deux molécules d'ADN simple brin complémentaires. Ces ADNs ont été fonctionnalisés au préalable, afin de permettre leur attachement aux billes utilisées pour le piégeage.

Une simulation en dynamique moléculaire des deux hairpins d'ARN (réalisée par Tristan Cragolini à l'IBPC) est présentée figure 2.3. Cette représentation met en évidence les interactions non Watson-Crick susceptibles d'être établies dans les boucles d'ARN. De plus cette représentation tient compte de la rotation véritable au sein de l'hélice, et illustre ainsi le fait que la tige choisie pour les hairpins, longue de 13 paires de bases, correspond à légèrement plus d'un pas hélicial (11 paires de bases). Ceci est aussi vrai dans le cas de l'ADN, où un pas de l'hélice B implique 10 paires de bases.

## 2.2.2 Description des expériences réalisées

### Réponse caractéristique d'un hairpin sous extension

La plupart des expériences avec les hairpins ont été réalisées à vitesse constante. Dans cette configuration, le piège mobile est éloigné, puis rapproché du piège fixe à une vitesse

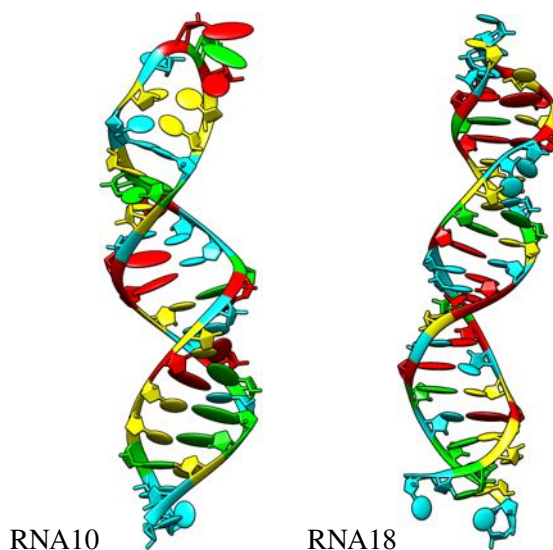


Figure 2.3 **Simulation en Dynamique Moléculaire des hairpins RNA10 et RNA18.** *Les interactions non Watson-Crick dans la boucle sont prises en compte. Simulations réalisées par Tristan Cragolini.*

donnée, choisie au préalable. Des mesures ont été réalisées à 50, 150, 300 et 450 nm/s. Dans tous les cas, la réponse en force mesurée présente une forme similaire. Un exemple de courbe enregistrée à 50 nm/s sur le hairpin RNA10 est présentée figure 2.4. La courbe foncée correspond à l'éloignement des pièges, et donc à l'extension de la construction. Dans un premier temps, pour une extension inférieure à  $\sim 1 \mu\text{m}$ , la construction s'étire à très basse force. Il s'agit de la réponse entropique des poignées (hybride ADN/ARN ici) : la pelote formée par le polymère se déroule, la molécule prend une forme allongée. Pour des extensions plus élevées, la force augmente avec une pente qui augmente elle-aussi. Il s'agit de la réponse élastique des poignées, qui sont progressivement étirées au-delà de leur longueur cristallographique ( $L_c = 1.13 \mu\text{m}$ ). Cette réponse en force est bien décrite par le modèle du ver extensible (extensible WLC), qui sera détaillé dans le paragraphe suivant. Une chute nette en force est ensuite observée, ici aux alentours de 13.5 pN. Elle correspond à l'ouverture du hairpin, c'est-à-dire à la rupture des liaisons Watson-Crick constituant la tige. Ce phénomène est stochastique, la force à laquelle il est observé est donc distribuée autour d'une valeur moyenne. Les extensions supérieures correspondent à la réponse élastique des constructions après déstabilisation du hairpin. Elles comportent alors un fragment simple brin (hairpin ouvert) entre les poignées double brin.

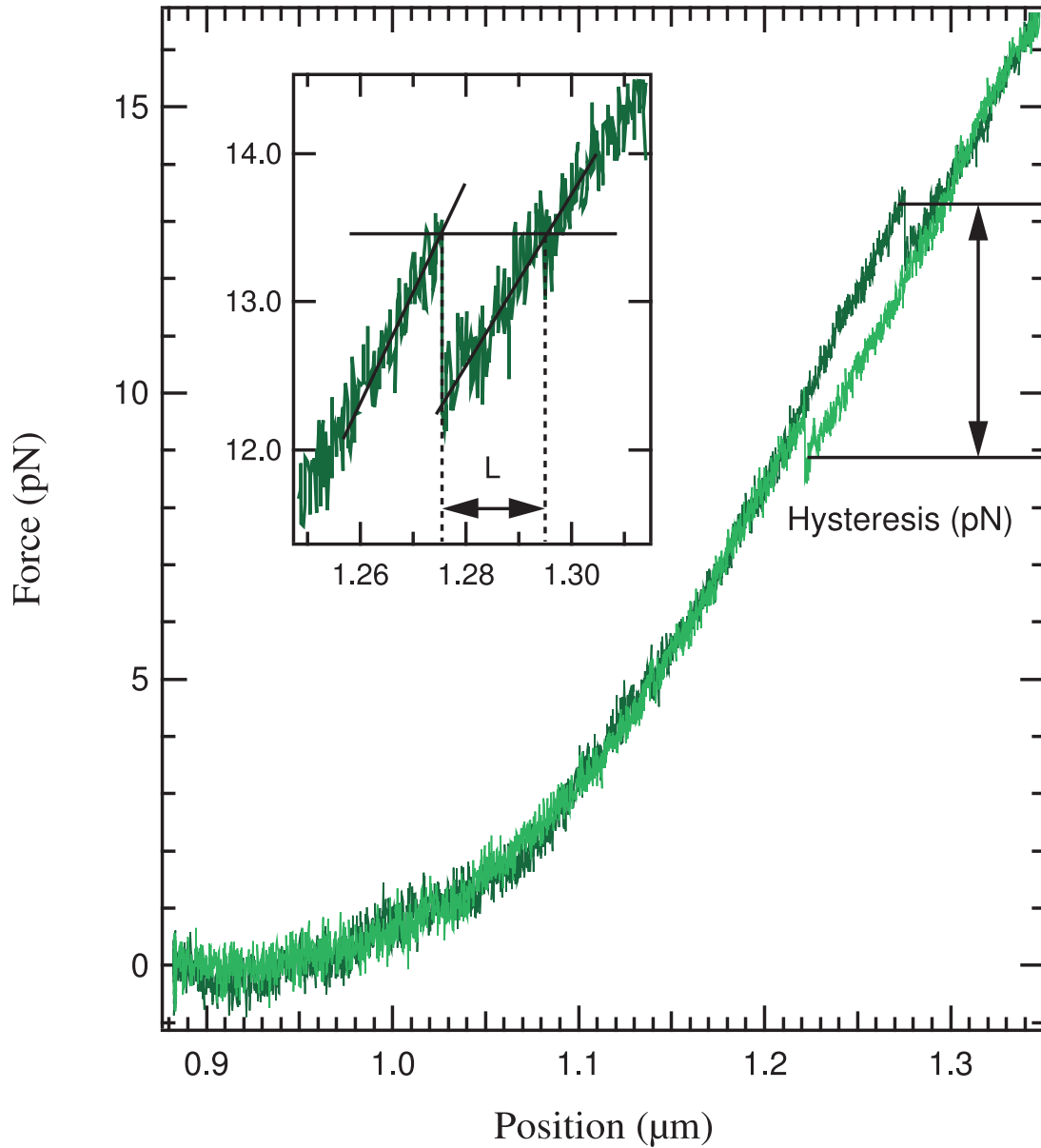


Figure 2.4 **Réponse en force du hairpin RNA10 à un cycle extension/relaxation à 50 nm/s.** La courbe vert foncé correspond à l'extension de la construction, et la courbe vert clair à sa relaxation. Les sauts en forces correspondant à l'ouverture ( $\sim 13.5$  pN) et à la fermeture ( $\sim 9$  pN) du hairpin sont repérés par des lignes noires. L'hystérésis est définie comme la différence entre les forces d'ouverture et de fermeture. L'encart présente un zoom sur l'ouverture du hairpin, et la mesure de la longueur  $L$  correspondant à la transition entre les états fermé et ouvert.

La déstabilisation du hairpin est un phénomène réversible : lorsque les extrémités de la construction sont rapprochées (courbe claire sur la figure 2.4), le hairpin peut se reformer. Cela se traduit par une augmentation brutale de la force, observée ici à  $\sim 9$  pN. Sur un cycle, les courbes d'extension et de retour sont similaires, à la fois dans la zone où le hairpin est présent et dans celle où il est déstabilisé. L'ouverture et la fermeture du hairpin se produisent par contre à des forces différentes, ce qui donne lieu à une hystérésis en force. Celle-ci est définie comme la différence entre les forces d'ouverture et de fermeture (figure 2.4). Les forces d'ouverture et de fermeture ont été mesurées pour les quatre hairpins, aux quatre vitesses étudiées. Pour chaque construction mesurée, plusieurs cycles ont été effectués. Un exemple des distributions des forces d'ouverture et de fermeture (mesurées sur le hairpin RNA10) est présenté figure 2.5. Sur un cycle d'extension/relaxation d'une construction, les forces d'ouverture et de fermeture correspondent à des événements indépendants. L'hystérésis peut donc être calculée à partir d'une ouverture et d'une fermeture non successives. Les forces d'ouverture et de fermeture mesurées sur  $N$  cycles d'une même molécule ont été appariées pour générer les  $N^2$  valeurs d'hystérésis possibles. Les probabilités associées aux différentes valeurs d'hystérésis sont pondérées afin de respecter le nombre de cycles réellement mesurés sur la construction. Pour une vitesse d'extension donnée, les valeurs ainsi obtenues sur différentes constructions correspondant au même hairpin sont regroupées, et un histogramme est réalisé. Suivant cette méthode, les histogrammes d'hystérésis ont été générés pour les quatre hairpins, aux quatre vitesses étudiées. Ces résultats seront présentés et discutés dans la suite de ce chapitre.

Les forces d'ouverture et de fermeture les plus probables à 50 nm/s sont présentées dans le tableau 2.3 (section 2.3). Elles sont mesurées comme le maximum d'un fit gaussien des distributions d'ouvertures et de fermetures, comme présenté figure 2.5. La force moyenne de transition ( $F_T$ ) est définie comme la moyenne des forces d'ouverture et de fermeture les plus probables, et est mesurée pour chaque hairpin (Tableau 2.2). L'analyse des courbes force/extension permet aussi d'obtenir la longueur gagnée à l'ouverture du hairpin (« transition length »  $L$ ). Celle-ci correspond à la distance entre les courbes d'extension de la construction avant et après ouverture du hairpin, mesurée à la force d'ouverture la plus probable du hairpin considéré à 50 nm/s (encart Figure 2.4, Tableau 2.2). Cette vitesse a été choisie comme celle la plus proche d'une situation d'ouverture à l'équilibre.

### **Le modèle du ver pour l'extension des polyacides nucléiques**

Le modèle WLC (ou modèle du ver) est utilisé pour décrire le comportement des polymères semi-flexibles. Dans ce modèle, la chaîne de polymère est considérée comme un cylindre isotrope continûment flexible (contrairement au modèle de la chaîne librement jointe). Ce modèle simple doit être adapté lorsque l'on considère le cas de polymères extensibles soumis à une tension. Lorsque cette tension augmente, la chaîne passe d'un régime dominé par l'entropie à un régime d'extension élastique. Dans le cas des faibles élongations et déformations, et pour des chaînes grandes devant leur longueur de persistance, la longueur bout à bout  $x$  d'une chaîne soumise à une force  $F$  peut s'exprimer sous la forme suivante (solution d'Odijk [21]) :

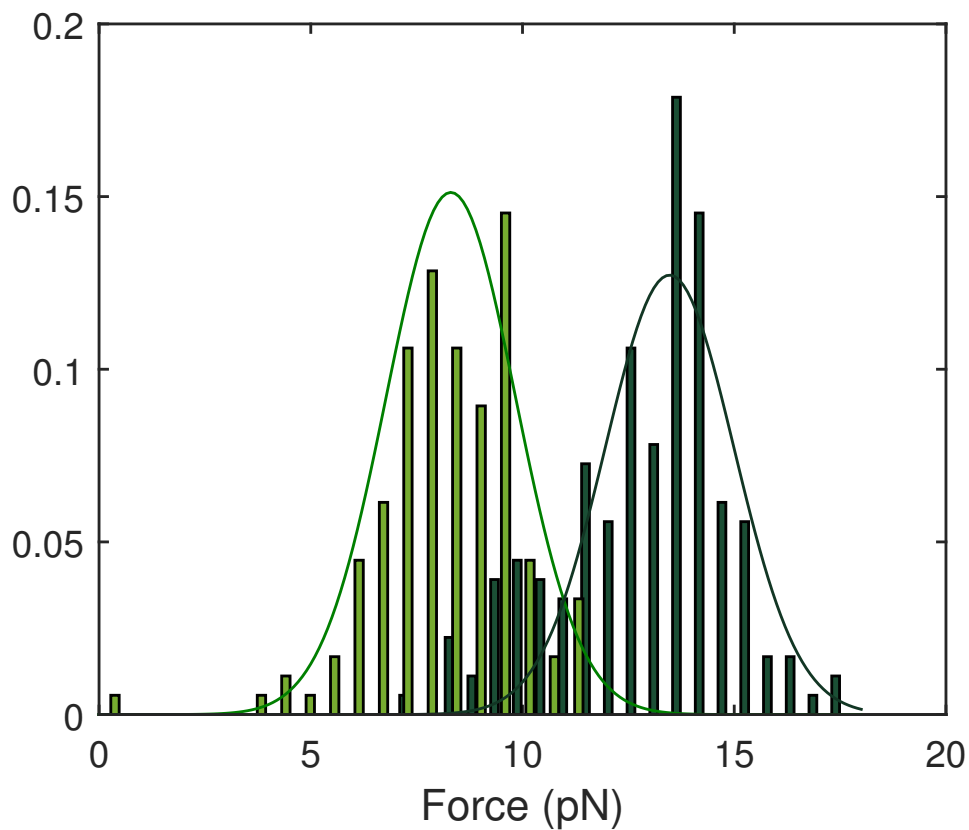


Figure 2.5 **Distributions des forces d'ouverture (vert foncé) et de fermeture (vert clair) de RNA10.** Les mesures ont été effectuées à 50 nm/s. Les histogrammes sont réalisés à partir de 179 événements d'ouverture et 148 de fermeture. Les courbes correspondent à des fits gaussiens de ces distributions. Figure issue de "Hairpins under tension : RNA versus DNA", Bercy et al., *Nucl. Acids Res.*, 2015 [19].



$$\langle x \rangle = L_c \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p} + \frac{F}{K}} \right)$$

Le terme entropique fait intervenir l'énergie d'agitation thermique  $k_B T$  (avec  $k_B$  la constante de Boltzmann, et  $T$  la température) et la longueur de persistance  $L_p$ , et domine à très faibles forces. Le terme enthalpique prend en compte le module élastique  $K$  de la chaîne, mais néglige tout effet de torsion. C'est ce modèle qui est le mieux adapté à la description de l'étirement des chaînes d'acides nucléiques dans la limite des basses forces ( $F \leq 35$  pN) [22]. Dans le cas des expériences de piégeage optique, un terme doit être ajouté à la solution pour décrire le déplacement des billes dans les pièges. Ce terme est de type loi de Hooke, avec pour paramètre la raideur équivalente aux deux pièges  $k_p$ . La solution complète est donc de la forme :

$$\langle x \rangle = L_c \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p} + \frac{F}{K}} \right) + \frac{F}{k_p}$$

### Application du modèle WLC à nos constructions

Pour des forces allant jusqu'à quelques dizaines de piconewtons, les courbes d'extension de chaînes d'ADN double brin ou d'hybrides ADN/ARN peuvent être ajustées par la solution d'Odijk du modèle WLC extensible, présentée ci-dessus. La figure 2.6 présente un exemple de courbes force extension enregistrées sur les hairpins RNA18 et DNA18, et de leur ajustement par le modèle WLC. La différence de longueur cristallographique entre la construction à poignées hybrides (RNA18) et celle à poignées ADN (DNA18) est clairement visible : la zone de réponse élastique apparaît à plus basse extension pour RNA18 que pour DNA18. L'ajustement par le modèle WLC donne pour les longueurs cristallographique et de persistance les valeurs suivantes :

	$L_c$	$L_p$
ARN/ADN	1.17 $\mu\text{m}$	47 nm
ADN/ADN	1.3 $\mu\text{m}$	43 nm

Les caractéristiques attendues pour l'hélice hybride ADN/ARN sont bien retrouvées : cette hélice de type A apparaît plus courte et plus rigide que l'hélice de type B, à nombre de nucléotides et séquence identiques.

Pour modéliser la réponse aux forces supérieures à  $\sim 35$  pN, il faut prendre en compte la structure hélicoïdale des polymères d'acides nucléiques, qui induit un couplage entre l'élongation et la torsion des chaînes [23]. Cette description sera développée au chapitre 4.

Le modèle WLC peut aussi être utilisé pour décrire l'extension d'acides nucléiques simple brin [22]. Dans nos expériences, une telle extension se produit lors de l'ouverture du hairpin : le simple brin est généré sous tension, ce qui induit un coût énergétique supplémentaire à la déstabilisation des bases de la tige. Cette énergie d'étirement peut être calculée par intégration

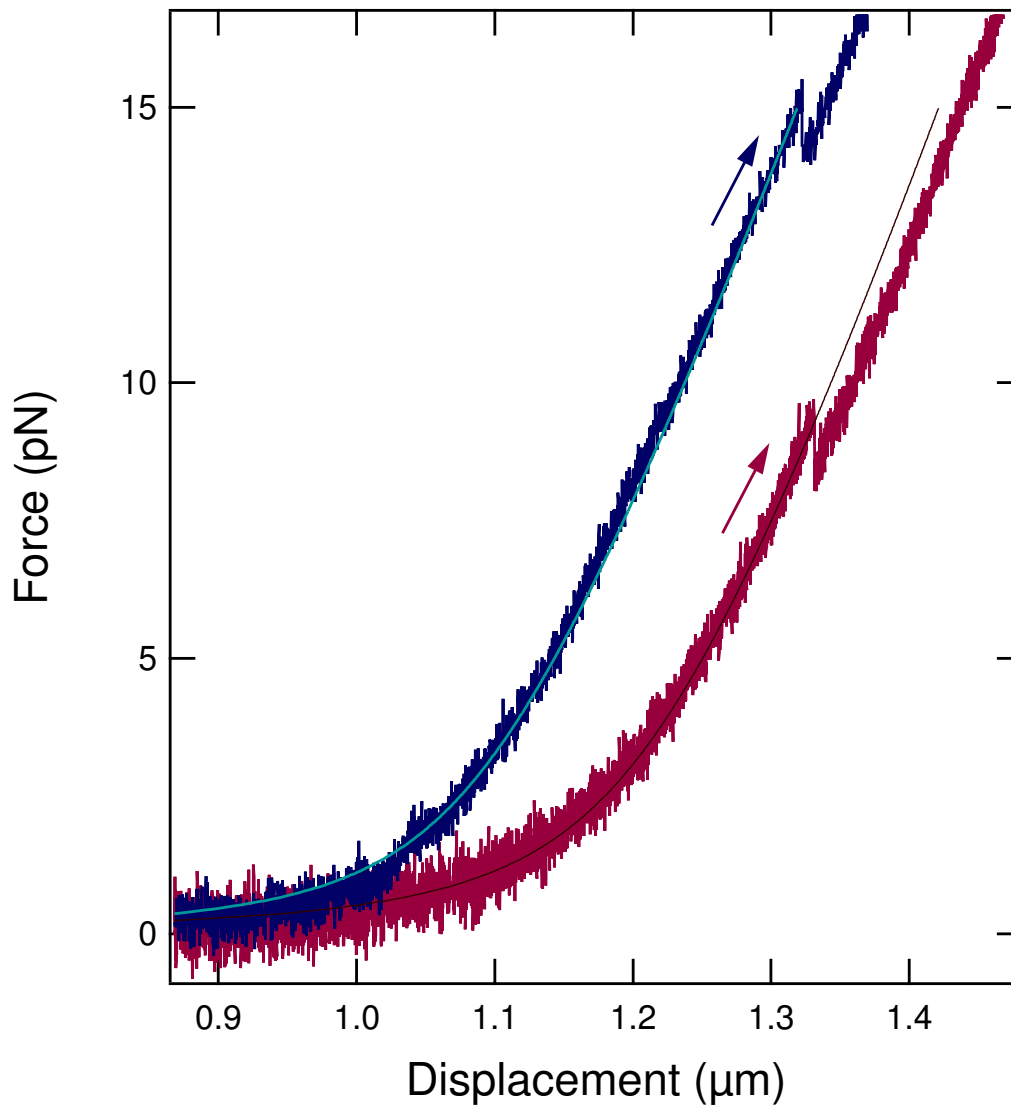


Figure 2.6 Exemples de courbes force/position pour l'extension d'un ADN double brin (rouge) et d'un hybride ADN/ARN (bleu). L'ajustement par le modèle WLC est représenté par les lignes continues. Les paramètres ajustés sont la longueur cristallographique  $L_c$  et la longueur de persistence  $L_p$ . Les valeurs obtenues sont les suivantes. Hélice hybride ARN/ADN :  $L_{c,ARN}=1.17 \mu\text{m}$  ;  $L_{p,ARN}=47 \text{ nm}$ . Hélice ADN/ADN :  $L_{c,ADN}=1.3 \mu\text{m}$  ;  $L_{p,ADN}=43 \text{ nm}$ . Figure issue de "Hairpins under tension : RNA versus DNA", Bercy et al., Nucl. Acids Res., 2015 [19].

de la courbe d'étirement du simple brin correspondant jusqu'à la force à laquelle l'ouverture se produit. Les courbes d'étirement de ce simple brin ne sont pas directement accessibles par notre mesure, mais peuvent être générées par le modèle WLC à condition de disposer des valeurs de longueur cristallographique  $l_0$  et de persistance  $l_p$  correspondantes. Pour les acides nucléiques simple brin, ces valeurs sont très dépendantes du tampon expérimental. Afin d'obtenir les valeurs correspondant au mieux à nos conditions, nous avons déterminé celles qui présentaient le meilleur accord avec les valeurs de force moyenne de transition  $F_t$  et de longueur à l'ouverture  $L$  mesurées. Pour cela, nous avons généré les courbes force/extension correspondant à des simples brins de 36 ou 44 nucléotides (longueur totale de nos hairpins dépliés) pour différents couples  $(l_0; l_p)$  inspirés de la littérature [24, 25, 20, 26] (figure 2.7). Le couple présentant la meilleure coïncidence avec le point  $(F_t, L)$  pour les deux longueurs de hairpins a été choisi. Ce travail a été fait en parallèle pour les deux hairpins ADN (présentés figure 2.7), et pour les deux ARN. Les valeurs retenues sont les suivantes :

	$l_0$	$l_p$
ARN	0.65 nm/nucléotide	1.37 nm
ADN	0.7 nm/nucléotide	1.2 nm

### 2.2.3 Modèle hors d'équilibre pour le changement d'état des hairpins

Cette partie a été initialement rédigée pour les sections « Material and Methods » et « Supporting Information » de l'article "*Hairpins under tension : RNA versus DNA*", Bercy et al., *Nucl. Acids Res.*, 2015 [19]. Elle présente les détails du modèle utilisé pour l'analyse thermodynamique et cinétique des hairpins, et a été conservée en anglais. Dans le cadre de cette thèse, il m'a paru plus naturel de réunir ces descriptions et de les faire figurer préalablement à l'analyse des résultats.

#### Theoretical description

Hairpin folding and unfolding are complex dynamical processes. They depend on several characteristics of the molecular construct. The base sequence of the hairpin stem induces a complex energy landscape and the size of the loop and the elasticity of the DNA/RNA handles also are important. Moreover, the force measurement by itself influences the experimental result, via the trap stiffness and the externally imposed distance versus time protocol. The imposed distance between the centers of the optical traps is subject to position noise arising from piezo-control electronics, mechanical vibrations and acoustic perturbations.

We use an out-of-equilibrium statistical physics description that is closely related to earlier publications [27, 28, 20, 29, 30, 31] (usually referred to as Bell-Evans model), and including a minimal set of parameters. We replace the complex molecular energy landscape by the single-barrier landscape presented in figure 2.8 and the elasticities of the molecular construct and optical traps are lumped together in one effective stiffness  $k_{eff}$ . A free energy barrier of height  $E$  separates the folded ( $x = 0$ ) and unfolded ( $x = L$ ) states. The free energies of the two states

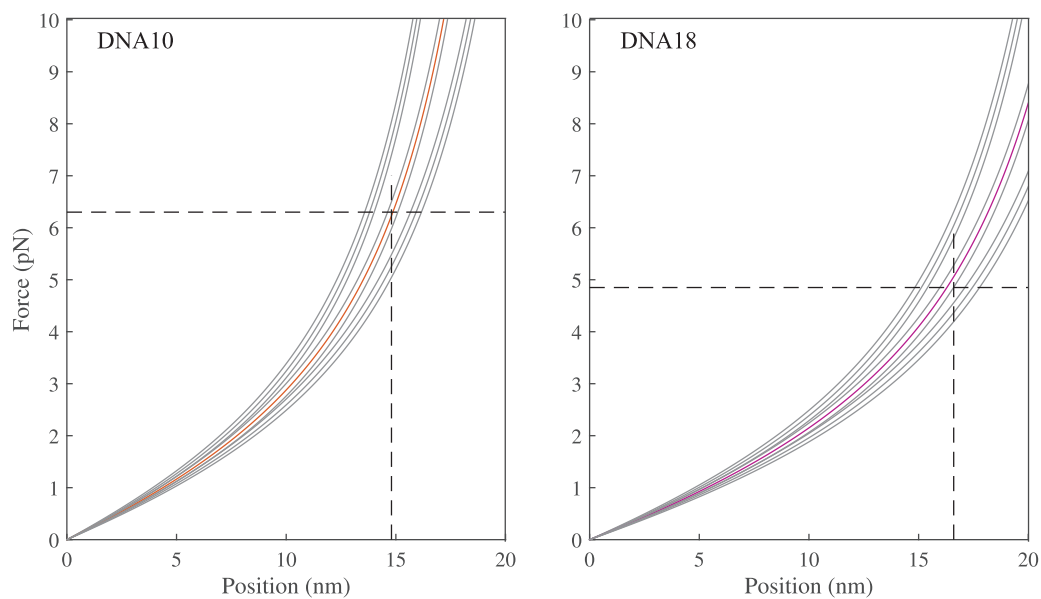


Figure 2.7 **Courbes théoriques d'étirement d'acides nucléiques simples brins (modèle WLC).** Les courbes ont été générées pour différents couples de paramètres  $(l_0 ; l_p)$  inspirés de la littérature [24, 25, 20, 26]. Les valeurs  $(F_t, L)$  mesurées pour nos hairpins ont été reportées (droites pointillées), sur le panneau de gauche pour l'ADN à boucle courte (DNA10) et celui de droite pour l'ADN à boucle longue (DNA18). Le couple  $(l_0 ; l_p)$  présentant la meilleure coïncidence avec ces valeurs a été choisi, et la courbe correspondante est ici représentée en couleur.

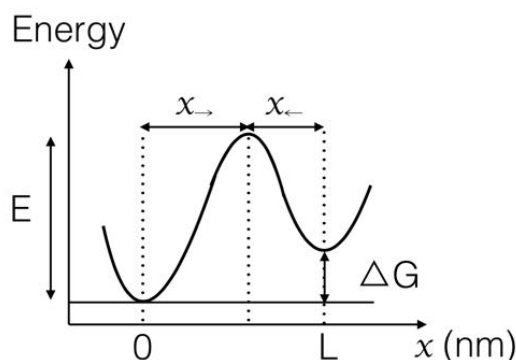


Figure 2.8 **Schematic representation of the simplified energy landscape used in our theoretical description.** The state of the folded hairpin occurs at  $x = 0$  and the state of the unfolded hairpin at  $x = L$ . The distances between the transition state and the folded and unfolded states are denoted by  $x_{\rightarrow}$  and  $x_{\leftarrow}$ , respectively. By definition, we have  $L = x_{\rightarrow} + x_{\leftarrow}$ .

differ by  $\Delta G$ . In the absence of force, the folded to unfolded transition rate is  $k_0 = k(F = 0) = v_0 \exp(-E/k_B T)$  (it includes the barrier height  $E$  and the attempt frequency  $v_0$ ). The reverse process is associated to a transition rate  $k_1 = k_0 \exp(\Delta G/k_B T) = v_0 \exp(-(E - \Delta G)/k_B T)$ . Applying an external force  $F$  tilts this landscape. In the limit of high barrier ( $E \gg k_B T$ ) and small applied force ( $F \ll E/x_{\rightarrow}$ ), the position of the transition state  $x_{\rightarrow}$  does not vary significantly and the barrier for unfolding is reduced by an amount  $F x_{\rightarrow}$ . Increasing (resp. decreasing) the external force linearly with time within this model, the unfolding probability  $p_{\rightarrow}(F)$  (resp. folding probability  $p_{\leftarrow}(F)$ ) can be derived analytically. Both probabilities explicitly depend on the pulling speed.

### Derivation of the unfolding and folding probabilities, and force hysteresis expression

Let us start considering the two-state free energy landscape presented in figure 2.8. At  $t = 0$ , the molecule is in the folded state ( $x = 0$ ). Force increases with a constant rate  $r$  according to

$$F = rt = k_{eff} vt \quad (2.2.1)$$

with an effective stiffness  $k_{eff}$  and a displacement velocity  $v$ . We define  $S(F)$  as the probability that the molecule is still folded at time  $t$  when the force has increased to a value  $F$ . Assuming first order kinetics, we have

$$\frac{dS}{dt} = -k(F(t)) S(t) \quad (2.2.2)$$

where  $k(F)$  is a force-dependent transition rate from the folded to the unfolded state. Assuming  $k(F) = v_0 \exp(-(E - F x_{\rightarrow})/k_B T)$  with an attempt frequency  $v_0$ , Eq. 2.2.2 is easily solved for

the linear force ramp of Eq. 2.2.1. :

$$S(F) = \exp\left(-\frac{1}{r} \int_0^F k(F') dF'\right)$$

Introducing the distribution of rupture forces  $p(F)$  that is related to  $S(F)$  by

$$\frac{dS}{dt} dt = -p(F) dF$$

we obtain the unfolding probability

$$p_{\rightarrow}(F) = \frac{k(F)}{r} \exp\left(-\frac{1}{r} \int_0^F k(F') dF'\right) \quad (2.2.3)$$

and thus

$$p_{\rightarrow}(F) = C_{\rightarrow} \exp\left(\frac{F x_{\rightarrow}}{k_B T} - \frac{k_0 k_B T}{r x_{\rightarrow}} \exp\left(\frac{F x_{\rightarrow}}{k_B T}\right)\right) \quad (2.2.4)$$

where  $k_0 = k(F=0) = v_0 \exp(-E/k_B T)$ .

The same procedure can be used for the folding probability  $p_{\leftarrow}(F)$ , which leads to

$$p_{\leftarrow}(F) = C_{\leftarrow} \exp\left(-\frac{F x_{\leftarrow}}{k_B T} - \frac{k_1 k_B T}{r x_{\leftarrow}} \exp\left(-\frac{F x_{\leftarrow}}{k_B T}\right)\right) \quad (2.2.5)$$

with  $k_1 = k_0 \exp(\Delta G/k_B T) = v_0 \exp(-(E - \Delta G)/k_B T)$ .

$C_{\rightarrow}$  and  $C_{\leftarrow}$  in Eqs. 2.2.4-5 are normalisation factors which come out when Eq. 2.2.3 (and the corresponding one for  $p_{\leftarrow}$ ) is calculated. They ensure the normalisation  $\int_0^{+\infty} p_{\rightarrow}(F') dF' = 1$  (resp.  $p_{\leftarrow}$ ).

$$C_{\rightarrow} = \frac{k_0}{r} \exp\left(\frac{k_0 k_B T}{r x_{\rightarrow}}\right)$$

and

$$C_{\leftarrow} = \frac{k_1}{r}$$

provided that the starting force of the decreasing force ramp is large enough to prevent folding events. In that case, the exact initial force doesn't influence the folding probability. In our experiments, 20 pN ensures this condition for all the hairpins studied.

The sum of two distributions (which is also a distribution) corresponds to their convolution. The distribution of hysteresis thus corresponds to the convolution of the distributions  $p_{\rightarrow}(F)$  and  $p_{\leftarrow}(-F)$  :

$$p_{hys}(\Delta F) = (p_{\rightarrow} * p_{\leftarrow})(\Delta F) = \int_{-\infty}^{+\infty} p_{\rightarrow}(F) p_{\leftarrow}(F - \Delta F) dF \quad (2.2.6)$$

$p_{\rightarrow}(F)$  and  $p_{\leftarrow}(-F)$  are calculated numerically using Eqs. 2.2.4-5, and their convolution is computed by the corresponding Matlab function.

---

Strictly speaking,  $p_{\rightarrow}(F)$  and  $p_{\leftarrow}(F)$  are probability densities, of dimension  $N^{-1}$ . For the sake of simplicity, we have retained the simple notation "probability". In the figures, we present real probabilities, i.e. the probability density is integrated over the force interval of the bin.

### Parameters and relation to theoretical descriptions of the literature

This model corresponds to the one described in the literature first by Evans [29] (relying on former work by Kramers [28] and Bell [27]) and more recently enriched by Dudko and coworkers among others [32, 33], particularly by taking into account various geometries for the landscape shape. In the aim of using a small number of parameters for our comparison, we choose to consider the case of a linear force ramp and use Bell's formula to describe  $k(F)$  [27]. This doesn't introduce new parameters for the shape of the landscape. This description involves a set of 3 parameters for the unfolding process, namely  $x_{\rightarrow}$ ,  $k_0 = v_0 \exp(-E/k_B T)$ , and  $r = k_{eff}v$ . Two additional ones are needed for the folding process :  $x_{\leftarrow}$  and  $k_1$ . Note that the attempt frequency  $v_0$  and the energy barrier  $E$  are not considered separately, but merge in the effective attempt frequency  $k_0$ .

Since  $L$  can be measured (Figure 2.4), we express  $x_{\leftarrow}$  as  $L - x_{\rightarrow}$ . We assume the same attempt frequency  $v_0$  for unfolding and folding, yielding to  $k_1 = k_0 \exp(\Delta G/k_B T)$ . With  $\Delta G$  obtained by calculation (mfold and stretching correction, see previous section and Table 2.2),  $k_1$  is defined by the knowledge of  $k_0$ . In addition, the loading rate  $r$  is determined since the effective stiffness  $k_{eff}$  is measured, and the velocity  $v$  experimentally imposed. This leaves us with only two fitting parameters, namely  $x_{\rightarrow}$  and  $k_0$ .

### Fitting procedure

The theoretical force hysteresis distribution is used as fitting function for the measured force hysteresis distributions, with  $x_{\rightarrow}$  and  $k_0$  as fitting parameters. The fits are done with a Matlab least-square fit algorithm. For each hairpin, the hysteresis distributions of the four velocities are fitted together as follows. In a first step, independent fits are performed for each velocity. For each of the two parameters an average value is calculated from these fits. These averages are used as starting values in an iterative procedure to obtain a global best fit, with  $x_{\rightarrow}$  constraint to a single value (within 0.2 nm tolerance) and  $k_0$  allowed to vary within a factor of 10.

The quality of the fits is characterised by the root mean squared error (*rmse*) :

$$rmse = \frac{1}{n_b} \sqrt{\sum_i (p_i - p_{hys,i})^2} \quad (2.2.7)$$

with  $p_i$  representing the experimental probability associated to the bin  $i$ ,  $p_{hys,i}$  the calculated probability for the bin  $i$ , and  $n_b$  the number of bins. For all histograms of this work, the *rmse* are in the range of  $(1-5) \times 10^{-3}$ .

## 2.3 Hairpins under tension : RNA vs DNA

### 2.3.1 Résumé des résultats obtenus

La comparaison de deux hairpins d'ARN de même tige et de boucles de tailles différentes avec leurs équivalents réalisés en ADN a donné les résultats suivants :

- dans des conditions expérimentales identiques, la force nécessaire à la déstabilisation du hairpin est toujours inférieure pour le hairpin d'ADN que pour son équivalent d'ARN.
- pour tous les hairpins, l'hystérésis entre les forces d'ouverture et de fermeture augmente avec la vitesse d'étirement.
- lors d'expériences à position constante, des changements d'état spontanés sont observés sur le hairpin d'ADN à grande boucle (DNA18), mais absents pour la structure d'ARN équivalente (RNA18). Pour les hairpins à boucle courte, de telles transitions sont observées aussi bien sur la structure ARN (RNA10) que la structure ADN (DNA10).
- en accord avec l'observation précédente, le taux de transition  $k_0$  obtenu pour les hairpins d'ADN est supérieur à celui des structures d'ARN équivalentes. Cela traduit une dynamique plus rapide des structures d'ADN par rapport à leur équivalent ARN.
- le paysage énergétique des hairpins étudiés apparaît fortement asymétrique, avec dans tous les cas un état de transition  $x_{\rightarrow}$  plus proche de l'état fermé que de l'état ouvert. Cette caractéristique est plus prononcée pour les hairpins d'ARN que pour leur structure d'ADN équivalente. Cela traduit un processus coopératif lors de l'ouverture des hairpins, avec une déstabilisation complète de la structure lorsque les premières paires de bases de la tige sont ouvertes ( $\sim 3$  premières paires pour RNA10 et 18, et  $\sim 5-6$  pour DNA10 et 18).

L'obtention de ces résultats est détaillée dans la section suivante (extraite de [19]), de même que leur discussion approfondie.

### 2.3.2 Results

Cette partie est extraite de l'article "*Hairpins under tension : RNA versus DNA*", Bercy et al., *Nucl. Acids Res.*, 2015 [19].

In the beginning of a measurement with varying extension, the mobile trap is moved apart from the fixed trap with a constant velocity. This movement is stopped at a relative extension of the construct (ratio of total length under tension to crystallographic length) of around 1.4, where the hairpin is always in the unfolded state. The movement is then inverted, so that the traps reapproach at the same speed and the hairpin can refold. This cycle is performed several times on the same molecule.

In Figure 2.9, we present typical force versus extension curves for the four hairpins at 50 nm/s. The left part of each curve (small displacement range not entirely shown) corresponds to the entropic response of the double-stranded handles where the force remains small. Force increases with a rising slope when the displacement approaches the crystallographic length of the molecular construct. This regime corresponds to the elastic response of the double-stranded



handles, well described by the Worm Like Chain model [34] (Figure 2.6). For the DNA hairpin constructs, the handles are double-stranded DNA, whereas the RNA hairpin constructs have hybrid DNA/RNA handles. This explains the length difference at the starting point of the elastic response. At a certain extension, a sudden drop in force by about 1 pN is observed, which corresponds to the unfolding of the hairpin. After this event, the force increases again, now stretching the whole construct with the double-stranded handles and the unfolded hairpin in series. No intermediate state is observed during hairpin unfolding.

When releasing the strain by inverting the direction of displacement and bringing the traps closer to each other, the hairpin can fold back into its initial structure. As no intermediate state is observed during the corresponding sudden upward rise in force, this process also appears to be cooperative.

The stretch and release curves are indistinguishable, except for the region where the unfolding (resp. folding) occurs. For a velocity of 50 nm/s and for all constructs except DNA10, a force hysteresis between unfolding and folding is observed. The hysteresis indicates that transitions occur out of thermal equilibrium. On the other hand, DNA10 shows several force flips between the folded and unfolded states during dynamic stretching and releasing of the construct (Figure 2.9.a, inset). These force flips are a signature of a close-to-equilibrium dynamics [35]. They are no longer observed and hysteresis becomes notable when the displacement velocity is increased to 150 nm/s and above (Figure 2.10). No force flips are observed for the other three hairpins.

Experimental data presented in table 1 show that, both for the DNA and RNA hairpins, the most probable folding force decreases with increasing loop size, while the unfolding force is less sensitive to this parameter. Both the unfolding and the folding forces of the RNA hairpins are higher than the corresponding forces of the DNA hairpins.

Measuring the successive unfolding (resp. folding) forces during stretch/release cycles on different molecules allowed us to build the hysteresis histograms for the four hairpins at the different displacement velocities. For DNA10 at 50 nm/s, when flips occur, only the first unfolding and first refolding forces are measured, according to the model used. Figure 2.12 shows such histograms for the four hairpins at 150 nm/s. The histograms corresponding to the other velocities are presented in figure 2.14. Our data show that for a given hairpin structure and under identical experimental conditions (same experimental buffer, same temperature, same pulling speed...), the hysteresis is notably higher for RNA than for DNA. We find that the hysteresis increases with loop length for both DNA and RNA.

For all four hairpins, the force hysteresis increases with displacement velocity. Figure 2.11 presents histograms of the hysteresis recorded for RNA10 at 50, 150, 300 and 450 nm/s as an example. A clear increase of the mean hysteresis is visible, with its value going from 5 pN at 50 nm/s to 7.5 pN at 450 nm/s. The hysteresis values for the four hairpins are given in SI, table 2.5. For each hairpin and pulling speed, variability in the unfolding and folding forces is seen, inducing a variability in the hysteresis. This behaviour is expected from the stochastic character of the unfolding and folding processes.

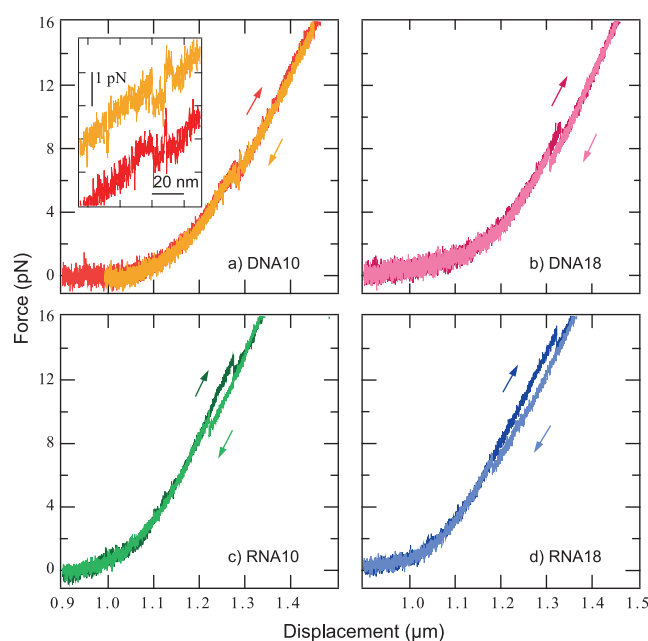


Figure 2.9 **Force versus extension curves at 50 nm/s for the four different hairpins.** The darker curves correspond to the stretching of the construct, whereas the lighter ones represent the release process. For both the DNA and RNA hairpins, the hysteresis between unfolding and folding increases with loop length. For the same hairpin sequence, the hysteresis is larger for the RNA construct than for the DNA one. The inset provides a closer look to the force flips observed with DNA10. In this inset, the unfolding and folding curves are shifted vertically for clarity.

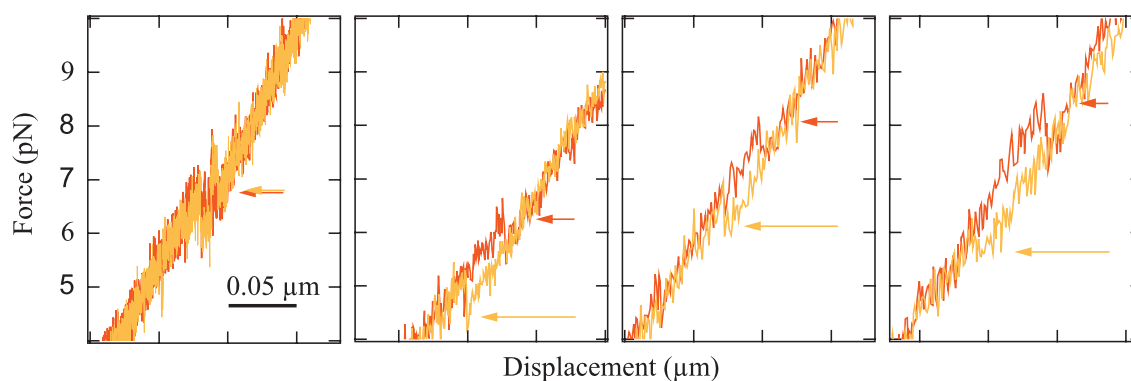


Figure 2.10 **Force versus extension curves for DNA10 at 50, 150, 300 and 450 nm/s (from left to right).** The darker curves correspond to the stretching of the construct, whereas the lighter ones represent the release process. The arrows indicate the force of the unfolding and folding events. The force flips observed with DNA10 at 50 nm/s vanishes at 150 nm/s and above. The hysteresis between unfolding and folding increases with the pulling velocity.

Tableau 2.1 **Most probable folding and unfolding forces, determined from measurements at a constant velocity of 50 nm/s. See figure 2.5 for an exemple of unfolding and folding forces distribution.**

hairpin	unfolding force (pN)	folding force (pN)
DNA10	6.4	6.2
DNA18	6.1	3.6
RNA10	13.6	8.8
RNA18	14.1	5.8

*The presented force values correspond to the maxima of a gaussian fit of the measured distributions of the forces where unfolding or folding occurs. The unfolding force distributions are build from 77 measured unfolding events for DNA10, 93 events for DNA18, 179 events for RNA10 and 146 events for RNA18. The folding force distributions are based on 79 measured folding events for DNA10, 70 events for DNA18, 148 events for RNA10 and 114 events for RNA18. The widths of these distributions are about 2.5 pN (FWHM).*

Theoretical analysis of the hysteresis histograms enables us to extract thermodynamic and kinetic information about the hairpin structures. The experimental distributions were fitted to calculated probability distributions, as defined in the previous section. For each hairpin, the fitting procedure is applied over the global data set, combining the four different speeds. The calculated force hysteresis distribution is presented by the solid line in figure 2.12.

Measured mean transition forces  $F_t$  (average of the mean opening force and the mean folding force at 50 nm/s) and transition lengths  $L$ , as well as theoretical free energy values that were obtained by simulation are presented in Table 2.2. Two parameters, the distance  $x_{\rightarrow}$  between folded state and transition state and the characteristic transition rate  $k_0$  were fitted to the experimental data and are presented in Table 2.3. The quality of the fits is good, with a root mean squared error in the range  $(1.0 - 5.0) \times 10^{-3}$  for each fitted histogram (*rmse*, see previous section for definition) and allows us to univocally determine the two parameters.

The results for  $x_{\rightarrow}$  indicate a clear asymmetry in the energy landscape for both DNA and RNA hairpins (ratio  $x_{\rightarrow}/L < 0.5$ ). The DNA values of  $x_{\rightarrow}/L \sim 0.3$  are higher than the RNA values of  $x_{\rightarrow}/L \sim 0.17$ . The transition rates  $k_0$  differ by a factor of about 10 between DNA18 and RNA18. The ratio is also about 10 between DNA10 and RNA10. These results confirm and quantify the faster dynamics of DNA as compared to RNA, already suggested by the appearance of force flips in the DNA10 curves at 50 nm/s.

The observation that the transition rates and hence the dynamics differ between the DNA and RNA hairpins is further confirmed by additional measurements performed with the four

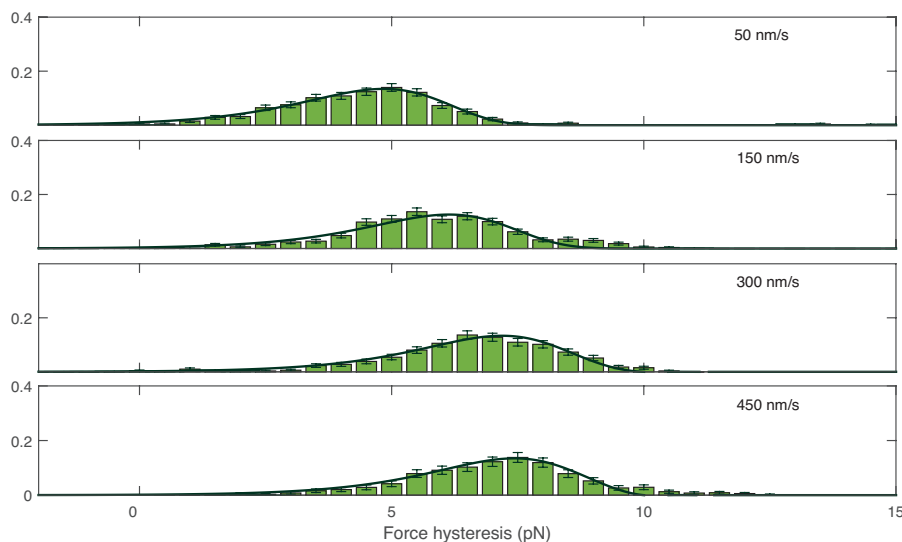


Figure 2.11 **Histograms of the hysteresis measured on RNA10 at 50, 150, 300, and 450 nm/s.** The solid lines correspond to the fit with the calculated convolutions of the unfolding and folding probabilities (Eqs. 4-6). Experimental data for 50 nm/s come from 148 stretch/release cycles on 48 molecules. 150 nm/s : 152 cycles on 44 molecules. 300 nm/s : 125 cycles on 16 molecules. 450 nm/s : 91 cycles on 7 molecules. There is good agreement between the experimental data and the calculated probability distributions. For this figure, the *rmse* is in the range  $(1.4-2.7) \times 10^{-3}$  for each panel.

Tableau 2.2 **Data obtained from measurements at 50 nm/s (a) or theoretical prediction (b).**

hairpin	(a)	(a)	(b)	(b)	(b)
	transition force $F_t$ (pN)	transition length $L$ (nm)	$\Delta G$ mfold (kJ/mol)	$\Delta G$ stretch (kJ/mol)	$\Delta G$ total (kJ/mol)
DNA10	6.3	$14.8 \pm 2$	$63 \pm 3$	$20 \pm 1$	$83 \pm 4$
DNA18	4.85	$16.6 \pm 2$	$60 \pm 3$	$18 \pm 1$	$78 \pm 4$
RNA10	11.2	$17.7 \pm 2$	$89 \pm 4$	$32 \pm 2$	$121 \pm 6$
RNA18	9.95	$18.4 \pm 2$	$87 \pm 4$	$35 \pm 2$	$122 \pm 6$

*mfold* : <http://mfold.rna.albany.edu>, see materials and methods and supplementary information for parameter determination.

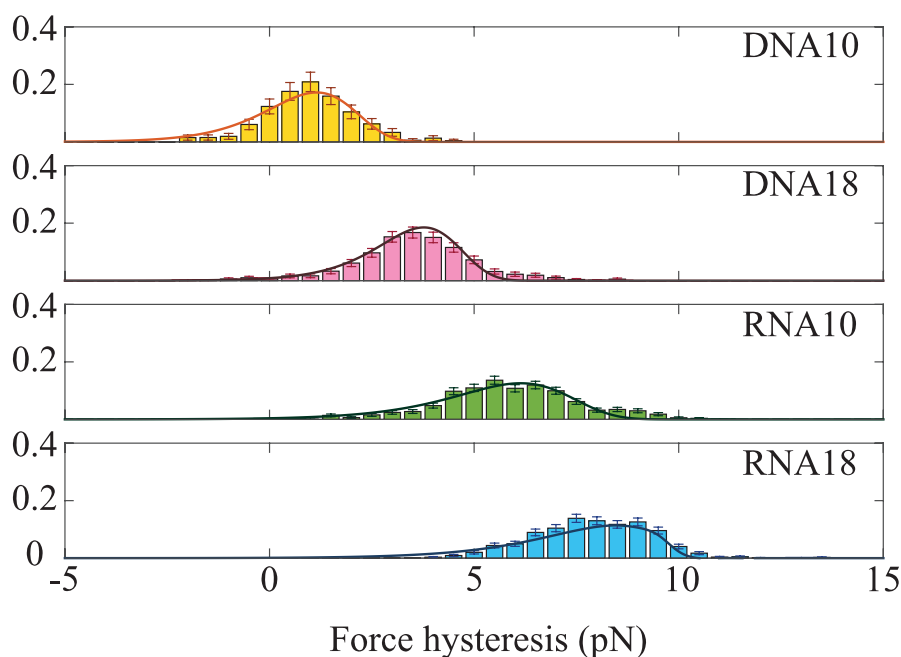


Figure 2.12 **Histograms of the hysteresis measured at 150 nm/s on the four different hairpins.** The experimental distributions are fitted to calculated convolution of the unfolding and folding probability distributions (solid line), as explained in the Materials and Methods section and in SI (section S1, Eqs 4-6). For each fitted histogram, the root mean squared error (*rmse*) is in the range  $(2.0 - 3.0) \times 10^{-3}$ . For DNA10 the data come from 46 stretch/release cycles on 12 molecules. DNA18 : 95 cycles on 13 molecules. RNA10 : 152 cycles on 44 molecules. RNA18 : 150 cycles on 21 molecules.

Tableau 2.3 **Parameters deduced from hysteresis measurements for the four hairpins.**

hairpin	transition state	transition state	transition rate
	$x_{\rightarrow}$ (nm)	$x_{\rightarrow}/L$	$k_0$ ( $s^{-1}$ )
DNA10	$4.4 \pm 0.2$	0.30	$(1.0 - 3.0) \times 10^{-4}$
DNA18	$4.7 \pm 0.2$	0.28	$(3.8 - 8.5) \times 10^{-5}$
RNA10	$3.0 \pm 0.2$	0.17	$(2.9 - 5.5) \times 10^{-5}$
RNA18	$3.1 \pm 0.2$	0.17	$(3.0 - 10) \times 10^{-6}$

By definition, we have  $x_{\leftarrow} = L - x_{\rightarrow}$  and  $x_{\leftarrow}/L = 1 - x_{\rightarrow}/L$ .

$$k_0 = k(F = 0) = v_0 \exp(-E/k_B T)$$

hairpins at constant extension. Here, the molecular construct is brought to a defined extension, which is then maintained constant during 10 seconds to 2 minutes. If the energy barrier is not too high, there are extension values where the hairpin can spontaneously flip between its folded and unfolded states [36, 35]; the measured force flips are induced by thermal fluctuations. If the corresponding dwell times do not significantly exceed the time-scale of the measurement, these spontaneous transitions can be observed experimentally.

Figure 2.13 (top) shows the force recorded on RNA10 as a function of time, while extension is held constant apart from stepwise increases. At the smallest extension, the hairpin remains folded most of the time, but already displays some brief passages to the unfolded state. The mean lifetime in the unfolded state progressively increases with extension, and at a given point equals the one in the folded state (this situation is presented in the middle panel). At even larger extensions, the unfolded state becomes energetically favorable, and the hairpin only folds sporadically. The same global behavior is observed for DNA10 (two bottom panels), but the transitions occur more frequently and the dwell-times of the folded and unfolded states are shorter than for RNA10. This observation is in line with the different  $k_0$  values of RNA10 and DNA10, presented in Table 3.

DNA10, DNA18 and RNA10 present spontaneous force flips at constant extension. For RNA18 however, no extension exists where both spontaneous unfolding and folding occurs, although the two RNA hairpins are close in terms of the theoretically predicted  $\Delta G$  values. This difference between RNA10 and RNA18 is consistent with the observation that the  $k_0$  value of RNA18 obtained from the varying extension data is an order of magnitude smaller than the one of RNA10.

### 2.3.3 Discussion

Cette partie est extraite de l'article "*Hairpins under tension : RNA versus DNA*", Bercy *et al.*, *Nucl. Acids Res.*, 2015 [19].

We observe that the RNA hairpins unfold and fold at higher force than the DNA hairpins. This may be expected from the known fact that in terms of free energy difference  $\Delta G$  an RNA duplex is more stable than the DNA duplex of equivalent sequence. Note however that there are three problems with this simplified view. First, the denaturation mechanisms under force and by thermal melting are different [37]. Whereas thermal melting involves independent and stochastic fluctuations of the stem bonds until denaturation occurs, force-induced unfolding is a directed process. Even if the bonds still fluctuate, the stem is sequentially opened from the bottom of the hairpin to the top. Similarly, the initial conditions for hairpin thermal folding are different from the ones for force reduction. Whereas folding of a thermally denatured molecule starts from a relatively compact conformation, the force measurement imposes a fully stretched initial construct. Second, the unfolding force can be estimated by a relation of the type  $F = \Delta G/L$  only if the transition occurs sufficiently close to thermal equilibrium. In the non-equilibrium case the actual barrier height, the attempt frequency and the force loading rate are the main parameters. There is no general relation between the equilibrium  $\Delta G$  and the barrier height. Third, the elastic energy usually plays a significant role in force-induced unfolding.

---

Extrait de l'article "*Hairpins under tension : RNA versus DNA*", Bercy *et al.*, *Nucl. Acids Res.*, 2015 [19].

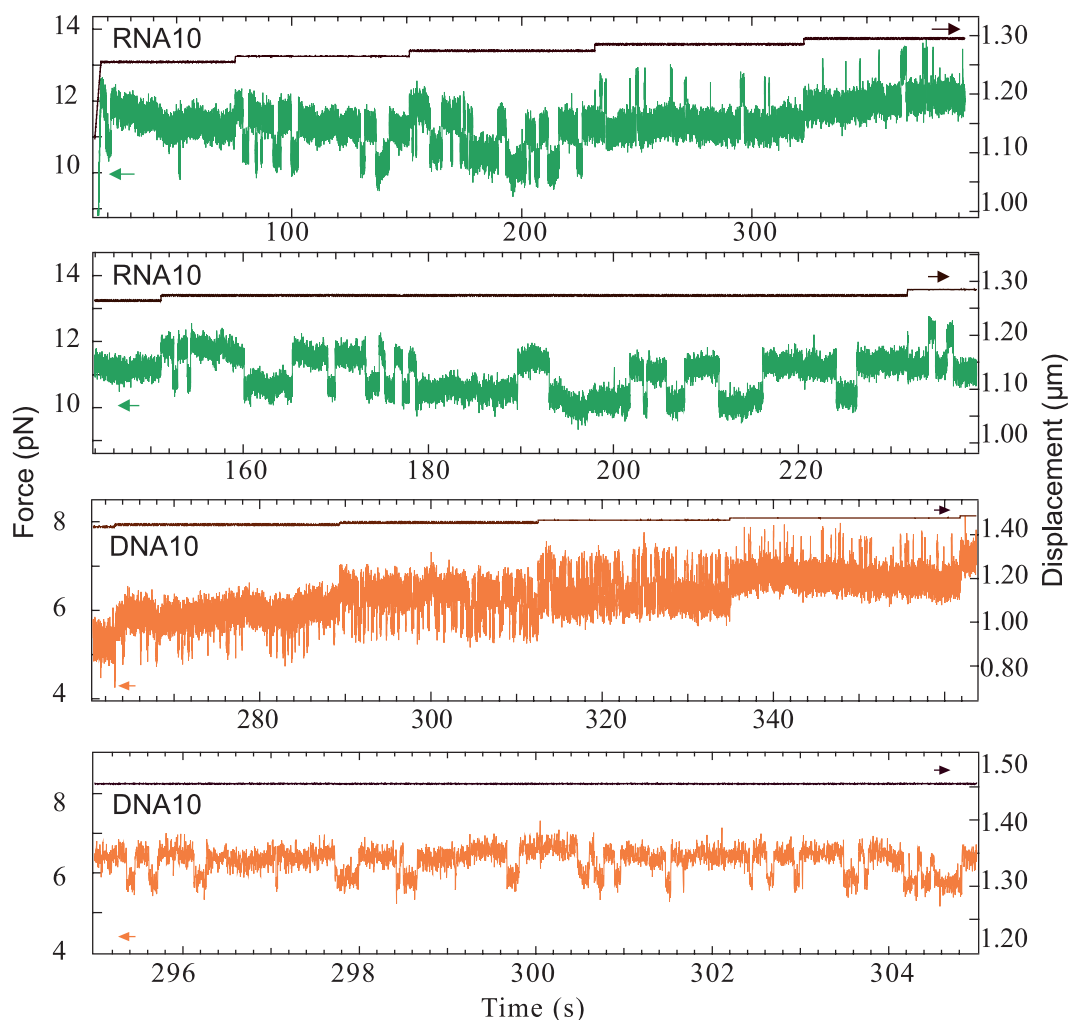


Figure 2.13 **Spontaneous transitions between folded and unfolded hairpin states at constant extension.** (Top panel) Force recorded on RNA10 (green) as a function of time, while the extension is constant apart from stepwise increases (black). Force flipping is observed for each extension value. (Second panel) Zoom on an extension where both states have a similar probability of occupation : the average dwell times are 1.77 s in the unfolded state and 3.1 s in the folded state. Presented time range : 80 s. (Third panel) Force recorded on DNA10 (orange) as a function of time, while the extension is constant apart from stepwise increases (black). Force flipping is observed for each extension value. (Bottom panel) Zoom on an extension where both states have a similar probability of occupation : the average dwell times are 0.11 s in the unfolded state and 0.25 s in the folded state. Presented time range : 10 s. The transition rate is much higher for DNA10 than for RNA10.

If for the sake of simplicity we consider the equilibrium case, the estimation of the opening force is of the form  $F = \Delta G_{tot}/L$ , but the molecular elasticity influences both the characteristic energy ( $\Delta G_{tot} = \Delta G_{mfold} + \Delta G_{stretch}$ , in the notations of Table 2) and the characteristic length  $L$  ( $L$  is the sum of the molecule length at zero-force and the force-induced extension). As can be seen from Table 2, the stretching energy contributions are sizeable for both DNA and RNA.

The hysteresis is larger for RNA18 than for RNA10 and the same holds for DNA18 compared to DNA10. This mainly arises from a loop-length dependence of the folding force. As shown in Table 1, the folding forces significantly decrease with increasing loop size, while the unfolding forces change weakly. To initiate hairpin folding a loop must form to approach the bases of the stem for base pairing. As folding occurs against an external force  $F$  in the present case, this initiation requires a mechanical work of the order  $W = FL$ , where  $L$  denotes the length of the unfolded loop. For similar  $W$  (energy fluctuations of similar amplitude), the different  $L$  thus implies that folding of the small-loop hairpin occurs at higher  $F$  than folding of the large-loop one. Of note, it was found by fluorescence techniques that even without external force the rate of hairpin closing decreases with increasing loop size [38].

We find that the transition rate  $k_0$  is smaller for the RNA hairpins than for their DNA counterparts. As shown in Figure 2.13, the RNA hairpin that exhibits smaller  $k_0$  also shows longer dwell times in both hairpin states. Smaller  $k_0$  leads to larger hysteresis in the constant velocity measurement, since both unfolding and folding happen statistically later on the force ramp. This relation between  $k_0$  on the one hand and magnitude of the force hysteresis on the other hand is observed in our measurements. RNA10 (resp. 18) exhibits smaller  $k_0$  and larger hysteresis than DNA 10 (resp. 18) at given velocity. For example, the transition length  $L$  corresponding to RNA10 unfolding (resp. folding) is  $\sim 20$  nm. Considering a typical equilibrium dwell time of 1 s for this hairpin (estimated from Figure 2.13, second panel), a pulling velocity of 50 nm/s induces a displacement of 50 nm during this time. This displacement is larger than  $L$ , and the folding (resp. unfolding) transition is thus highly unlikely. The unfolding transition (resp. folding) thus occurs out of equilibrium at this speed. On the contrary, the shorter dwell time of DNA10 ( $\sim 0.1$  s) induces a displacement ( $\sim 5$  nm) that is smaller than  $L$  ( $\sim 15$  nm) and allows for flipping between the two states. In this case, the measurement can be considered as close to equilibrium (Figure 2.9.a, inset).

As shown in table 2.3, the transition state is closer to the folded state than to the unfolded state. We convert the values of  $x_{\rightarrow}$  to the number of unfolded nucleotides, using a procedure described earlier [39]. For DNA10 (resp. DNA18) the transition state corresponds to 9 (resp. 11) unfolded nucleotides, or base pair 5-6 of the hairpin stem counted from the bottom. For RNA10 and RNA18 we obtain 5 (resp. 6) nucleotides, corresponding to base pair 3 from the bottom. We tentatively attribute these positions to the stem sequence : the first base pairs from the bottom are three G·C (Figure 2.2), conferring a greater stability to the bottom compared to the top of the stem. This interpretation is consistent with the fact that no intermediate state is observed during hairpin unfolding. Simply speaking, when the mechanical work becomes suf-



ficient to unfold the initial G·C rich regions, the whole stem bursts open cooperatively. Within our model that assumes a single transition state, the folding of the hairpin is correspondingly expected when the G·C rich part of the stem reforms. The fact that no intermediate states were observed in our measurements indicates that the folding and unfolding transitions occur with a high degree of cooperativity.

For comparison, we have calculated unfolding free energy landscapes of the hairpins. This study is described in section S8 of the supplementary information. The calculations are based on mfold free energies, include the sequential energy contributions of the different base pairs of the stem and the single-stranded loop, but do not include cooperativity. Interestingly the calculated landscapes exhibit their maxima at the top of the hairpin stem, in contrast to the measurements indicating transition states closer to the bottom of the stem. The differences are significant. First, they provide additional support to the idea that cooperativity is important for all investigated hairpins. Moreover, the comparison suggests a higher degree of cooperativity in RNA than in DNA, since we experimentally observe that the RNA hairpins exhibit smaller  $x_{\rightarrow}$  than the DNA hairpins.

There are structural arguments both for the stem and the loop that support a stronger cooperativity of the RNA hairpin as compared to the DNA hairpin. The stem of the RNA hairpin exhibiting a A-form helical structure is shorter and therefore more compact than the stem of the DNA hairpin, which is a B-form double helix. In the RNA helix, the constraint caused by the opening fork may thus more easily extend over several base pairs than in the DNA helix. It is expected that interactions between nucleotides in the hairpin loop contribute in a non-negligible manner to the overall stability. Indeed RNA is known (and DNA suspected) to build non Watson-Crick basepairs and we note that mfold does not fully take the corresponding free-energy contributions into account [2, 40]. The 2'-OH group of the RNA nucleotide forms hydrogen bonds, while the 2'-H of DNA does not. In this regard, NMR studies of chimeric RNA stem/DNA loop hairpins revealed that 2'-H in the hairpin loop reduces the overall stability and also suggests that the structure of the stem, A-form or B-form, strongly influences the loop structure [41].

We tentatively attribute the slower dynamics of the RNA hairpins to a stronger cooperativity. For strong cooperativity, barriers against folding and unfolding may be localized and high, since they can cumulate energy contributions of many structural elements. With weak cooperativity, on the other hand, the energy contributions of the structural elements distribute more equally over the reaction coordinate, high local barriers are less probable to occur and therefore passage of this landscape by a random walk that is biased by the external force is faster.

Measurements on more complex RNA structures showed pronounced structures in the force versus extension curves (or length versus force curves for force-clamp measurements) and hysteresis occurred even at low pulling speed [42, 43, 39], in qualitative agreement with

the present findings. Recently, an investigation of the mechanical properties of double-stranded RNA under force and torque has been published [44]. The authors applied torque to a torsionally constraint double-stranded RNA construct using magnetic tweezers. They investigated in particular a plectonemic buckling transition. They find that the characteristic transition rate of this buckling transition is two orders of magnitude smaller for double-stranded RNA than for double-stranded DNA. It is interesting to mention this study in the present context because it shows a slower dynamics for RNA than for DNA, although the buckling transition is different from hairpin folding.

A typical RNA molecule has a complex three-dimensional shape stabilized by several duplexes. The individual duplexes are connected by single strands of varying length, which allows for propagation of conformational changes within the RNA structure. This propagation of strain and displacement is influenced by the elasticity of the connecting single strands. Forces of the order  $k/\Delta x$  develop, where  $k$  is the stiffness of the molecular linker and  $\Delta x$  the relative displacement. This global picture suggests that the configuration studied in the present paper represents an elementary building block of complex three-dimensional structures formed by a single-stranded nucleic acid chain. Specifically, our result that the dynamics of folding and unfolding under mechanical load is slower in RNA hairpins than in the corresponding DNA hairpins translates itself into predicting that RNA structures are dynamically more stable. In other words, structures formed by single-stranded DNA rather than RNA exhibit significantly faster inter-conversion between alternative configurations. There are of course good reasons of evolutionary origin to explain why the domain of biologically-active three-dimensional nucleic acid structures is traditionally occupied by RNA. Our suggestion that RNA structures are dynamically more stable than DNA structures may be an interesting additional argument to understand why single-stranded DNA was not selected to become a molecule that assures its biological role by a complex internal structure.

#### 2.3.4 Conclusion

Cette partie est extraite de l'article "*Hairpins under tension : RNA versus DNA*", Bercy *et al.*, *Nucl. Acids Res.*, 2015 [19].

Two RNA hairpin structures of different loop size were compared with their DNA equivalents. Under the same experimental conditions, the forces needed to unfold the DNA hairpins are systematically lower. Hysteresis between unfolding and folding is more pronounced for RNA than for DNA. For all studied hairpin types, the hysteresis increases with the pulling speed. Imposing a constant distance, spontaneous force flips are observed for the large loop DNA hairpin, but are absent for the equivalent RNA structure. For small loop hairpins, flipping between the folded and unfolded states occurs with both DNA and RNA. Comparison of the experimental data with a theoretical description allowed us to estimate the position of the transition state and a characteristic transition rate. The latter is found to be smaller for RNA than DNA. From all these results emerges the picture that RNA exhibits a more pronounced

out-of-equilibrium character than DNA. The force-induced conformational dynamics of hairpins is slower in RNA than in DNA. The results suggest that structures made of RNA are more stable against internal and external forces than structures made of single-stranded DNA and therefore the former are more apt to form complex structures that are required for an active biological function.

### 2.3.5 Supplementary Information

Cette partie est extraite de l'article "*Hairpins under tension : RNA versus DNA*", Bercy *et al.*, *Nucl. Acids Res.*, 2015 [19].

#### **Precision of the measured force hysteresis compared to the precision of separately measured folding and unfolding forces.**

Calibration uncertainty limits the precision of the measured force to about 5 % in our experimental case. Major causes are variability in sample viscosity, uncertainty in the distance between the coverslip surface and the beads and variations in diameter, refractive index and deviations from spherical shape of the beads. These quantities influence the force calibration derived from the power spectral density of a captured bead. Experimental noise also affects the precision of the measured force and amounts to  $\Delta F \simeq 0.25$  pN for the unfolding and folding forces and acquisition rates of this study. The precision is slightly lower for the determination of the zero-force baseline, possibly due to the close proximity of the beads that can lead to cross-talk and weak rearrangement of the beads in the traps. We estimate the precision of our zero-force baseline to  $\Delta F_0 \simeq 0.5$  pN. Given these contributions, we can estimate the precision of the measured unfolding force  $\Delta F_{un}$ , folding force  $\Delta F_{fo}$  and force hysteresis  $\Delta F_{hys}$  as follows.

$$\begin{aligned}\Delta F_{un} &= \Delta F + \Delta F_0 + 0.05(F_{un}) \\ &\simeq 0.75\text{pN} + 0.05(F_{un})\end{aligned}$$

$$\begin{aligned}\Delta F_{fo} &= \Delta F + \Delta F_0 + 0.05(F_{fo}) \\ &\simeq 0.75\text{pN} + 0.05(F_{fo})\end{aligned}$$

$$\begin{aligned}\Delta F_{hys} &= 2 \Delta F + 0.05(F_{un} - F_{fo}) \\ &\simeq 0.5\text{pN} + 0.05(F_{un} - F_{fo})\end{aligned}$$

This simple analysis shows that we have a higher precision on the hysteresis measurement than on the opening and closing forces. Calculated  $\Delta F_{un}$ ,  $\Delta F_{fo}$  and  $\Delta F_{hys}$  for the four hairpins at 50 nm/s are provided in Table 2.4.

Tableau 2.4 **Estimation of the error in the measured unfolding force, folding force and force hysteresis for the four hairpins at 50 nm/s.**

hairpin	$\Delta F_{un}$ (pN)	$\Delta F_{fo}$ (pN)	$\Delta F_{hys}$ (pN)
DNA10	1.05	1.05	0.50
DNA18	1.05	0.90	0.65
RNA10	1.40	1.20	0.75
RNA18	1.50	1.05	0.90

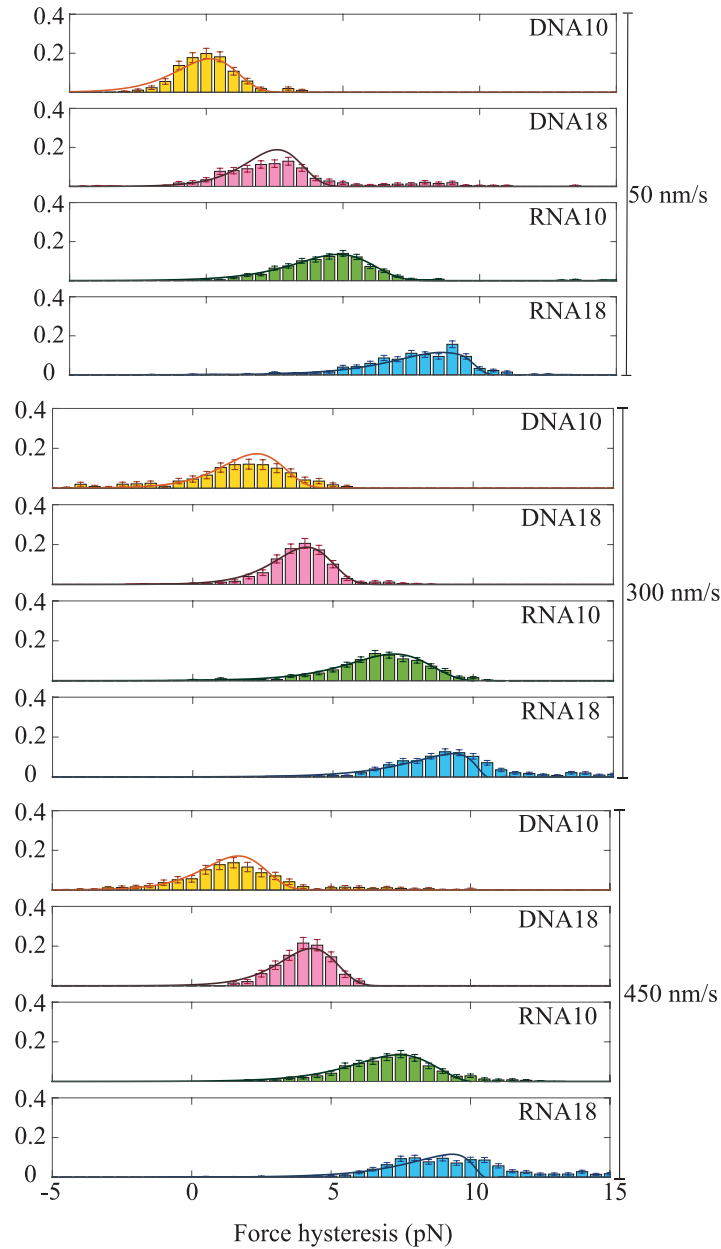
**Hysteresis distributions for the four hairpins at 50, 300, and 450 nm/s.**

Figure 2.14 **Histograms of the hysteresis measured on the four hairpins at 50 (top), 300 (middle), and 450 (bottom) nm/s.** The solid lines correspond to the fit with the calculated convolutions of  $p_{\rightarrow}$  and  $p_{\leftarrow}$  (Eqs. 4-6). The number of measured molecules and cycles are given in Table S3. The *rmse* is in the range  $(1.0-5.0) \times 10^{-3}$  for each fitted histogram.

### Calculation of hysteresis histograms and associated error bars.

On a stretch and release cycle, the unfolding and the folding forces are the result of independent events, and are thus uncorrelated. Consequently, for a given molecule and at a given speed, the hysteresis  $F_{um} - F_{fo}$  can be calculated combining values from different cycles. The histogram (for a given hairpin, at one speed) was obtained in three steps : first, all the unfolding forces measured were paired with all the folding forces measured on the same molecule in order to generate all possible hysteresis values. Second, the obtained values were weighted depending on the actual number of stretch and release cycles performed on this molecule. Third, a histogram was built from the values obtained with all the molecules probed. The number of event  $n_i$  in each histogram bin was normalised to the probability  $p_i = n_i/N$ , with  $N$  the total number of cycles performed. The error bars have been determined using the variance of the binomial law [45], leading to the following uncertainty in the probability :  $p_i \pm (p_i(1-p_i)/N)^{1/2}$ . They are shown for every histogram of this publication.

### Theoretical unfolding free-energy landscapes

In this section, we present energy landscapes obtained by sequential calculation of mfold free-energy contributions. We consider the expression

$$E_{tot}(j) = \sum_{i=1}^j E(i) - F l(j,F), \quad (2.3.1)$$

with

$$E(i) = \begin{cases} E_{bp}(i) + E_{stretch} & : i \leq n_{stem} \\ E_{loop}/n_{loop} + E_{stretch} & : n_{stem} < i \leq n_{stem} + n_{loop}, \end{cases}$$

where  $n_{stem} = 13$ ,  $n_{loop} = 5$  for DNA10 and RNA10 and  $n_{loop} = 9$  for DNA18 and RNA18. The binding energies of the base pairs  $E_{bp}$  and the loop energies  $E_{loop}$  are obtained by mfold.  $E_{stretch}$  denotes the energy required to extend a single-strand of two nucleotides from zero-force to  $F$ . The term  $-F l(j,F)$  on the right of Eq.8 accounts for the mechanical work of the generated single-strand of  $j$  base pairs exhibiting a length  $l(j,F)$  at force  $F$ .  $E_{stretch}$  and  $l(j,F)$  are determined as described in section S1. This one-dimensional theoretical description of the unfolding energy landscape has been used in earlier studies, for instance by the Block group [20] and our group [35, 46].

In figure 2.15, we present the landscapes calculated for the hairpin structures at different forces. In all cases of relevance to our measurements the maximum of the landscape occurs at the last base pair of the stem, i.e. at the top of the stem in the representation of main text figure 2.2. The experimental results, however, indicate transition states that are closer to the bottom of the stem. A way to account for this discrepancy consists in assuming that the base pairs in the vicinity of the opening fork are destabilized. This is a way to introduce cooperativity. More pronounced destabilisation corresponds to stronger cooperativity and would lead to smaller  $x_{\rightarrow}$ .

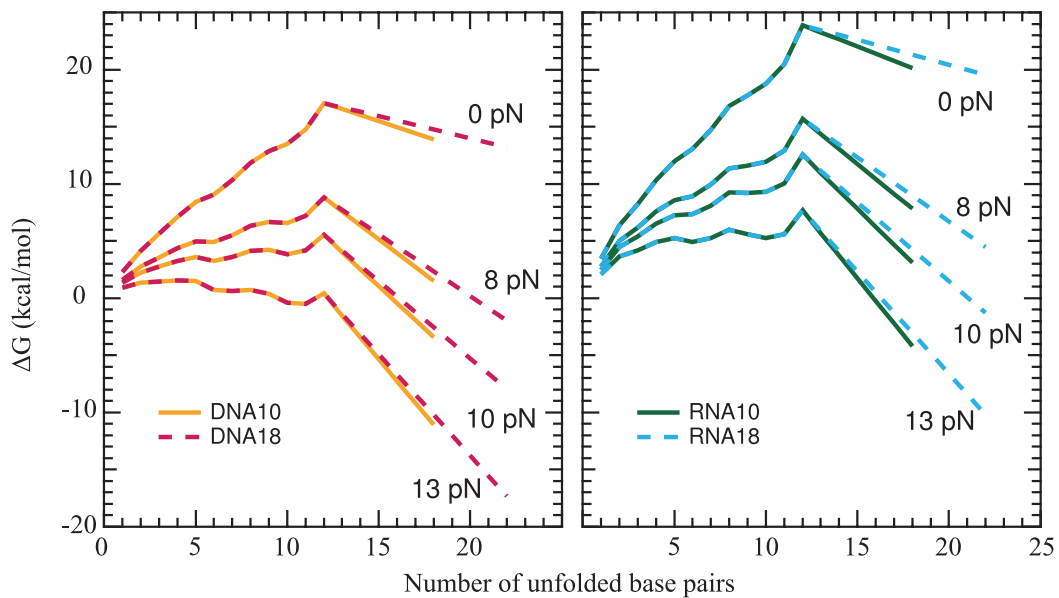


Figure 2.15 **Unfolding energy landscapes** calculated for the DNA (left panel) and RNA (right panel) hairpin structures at 0, 8, 10, and 13 pN externally imposed force. The dashed lines represent the large loop hairpins (DNA18 and RNA18), and the solid lines the smaller loop ones (DNA10 and RNA10). The stem of the two DNA (resp. RNA) hairpins being identical, at a given force the curves are superimposed from 1 to 12 unfolded base pairs.

As a main result this supplementary section indicates that cooperativity is required to account for the experimentally observed positions of the transition state. In agreement with the discussion section of the main text, our observation of significantly smaller  $x_{\rightarrow}$  for RNA than for DNA suggests that cooperativity is stronger in RNA.

### **Influence of $\Delta G$ on the fitted parameters**

At constant trap distance force flipping between the folded and unfolded states can be observed for hairpins DNA10, RNA10 and DNA18. We measure the force  $F_{1/2}$ , for which the average dwell times in the folded and unfolded states are equal. This allows us to experimentally estimate  $\Delta G$ . We have  $\Delta G = F_{1/2}L$ , where  $L$  is the change in the hairpin length induced by a transition at the force  $F_{1/2}$ . Measured  $F_{1/2}$  and derived  $\Delta G$  values are presented in table 2.7. The table contains no line for RNA18, since flipping is not observed for this hairpin. The energies obtained from  $F_{1/2}$  and the  $\Delta G_{tot}$  values of Table 2 agree within 30%. Using the energies obtained from  $F_{1/2}$  rather than the theoretically predicted  $\Delta G_{tot}$  values in the fitting procedure changes neither  $x_{\rightarrow}$  nor the ratios between the transition rates  $k_0$ , as can be seen by comparing table 2.7 and table 2.3.



Tableau 2.5 **Increase of hysteresis with displacement velocity.** The values correspond to the maximum of the fit of the measured hysteresis distribution to our out-of-equilibrium model.

hairpin	hysteresis at 50 nm/s (pN)	hysteresis at 450 nm/s (pN)
DNA10	0	2
DNA18	2.5	4.5
RNA10	5	7.5
RNA18	8.5	9.5

Tableau 2.6 **Number of molecules and stretch-release cycles that were measured and analyzed.**

hairpin	50 nm/s (mol./cyc.)	150 nm/s (mol./cyc.)	300 nm/s (mol./cyc.)	450 nm/s (mol./cyc.)
DNA10	15 / 70	12 / 46	8 / 49	5 / 52
DNA18	15 / 70	13 / 95	7 / 68	6 / 51
RNA10	48 / 148	44 / 152	16 / 125	7 / 91
RNA18	29 / 114	21 / 150	17 / 122	9 / 124

Tableau 2.7 **Force  $F_{1/2}$  measured at constant trap distance, energy  $\Delta G$  derived from  $F_{1/2}$ , and the parameters  $x_{\rightarrow}$  and  $k_0$  obtained by fitting our hysteresis measurements using  $\Delta G$  derived from  $F_{1/2}$ .**

hairpin	$F_{1/2}$ (pN)	$\Delta G$ (kJ/mol)	$x_{\rightarrow}$ (nm)	$k_0$ (s <sup>-1</sup> )
DNA10	6.3	56	4.4	$(2.4 - 7.1) \times 10^{-3}$
DNA18	7.0	70	4.6	$(1.0 - 2.5) \times 10^{-4}$
RNA10	11.0	117	3.1	$(1.8 - 3.0) \times 10^{-5}$

# Chapitre 3

## Action des protéines hélicases sur les structures secondaires d'ARN

*Le biologiste passe, la grenouille reste...*

---

Jean Rostand

### 3.1 Les protéines hélicases

Les protéines hélicases sont associées à l'ouverture des hélices double brin d'acides nucléiques. Certaines sont spécifiques des doubles brins d'ADN, d'autres des doubles brins d'ARN. D'autres encore sont polyvalentes, au moins *in vitro*. La spécialité des hélicases acide nucléique spécifiques peut provenir (entre autres) de la différence de structure des doubles hélices selon l'acide nucléique qui les compose : l'ADN leur confère une structure de type B, alors que l'ARN provoque une structure de type A (Introduction). L'accessibilité des petit et grand sillons dépend du type d'hélice.

Selon la présence ou l'absence de certains motifs d'acides aminés, les hélicases (ADN et ARN) sont classées selon 5 « superfamilles ». Les protéines à motif DEAD, étudiées dans ce projet, font par exemple partie de la superfamille SF2. Les protéines de cette famille sont retrouvées dans quasiment tous les processus impliquant des polymères d'acides nucléiques : transcription, traduction, formation du ribosome [47, 48]... Leur structure est proche de celle de nombreuses hélicases à ADN processives (capables de séparer activement de longs fragments d'ADN double brin), bien que les hélicases à ARN le soient rarement. Elles comportent deux domaines de type RecA (protéine bactérienne de recombinaison) permettant la liaison au substrat (Figure 3.1). Une molécule d'ATP peut se fixer entre ces deux domaines. Le « coeur hélicase » (constitué des deux domaines RecA) contient 12 motifs à séquences caractéristiques, qui permettent de classer ces protéines en différents sous-groupes. Le motif DEAD est un exemple de motif II.

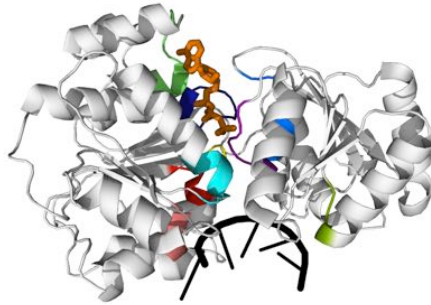


Figure 3.1 **Modèle 3D de la protéine Ded1.** La structure tertiaire présentée est basée sur la structure cristallographique résolue d'une protéine de séquence proche. Les deux domaines de type RecA sont séparés, et l'ATP vient se placer au centre. Le substrat ARN est aussi représenté.

### Les protéines à motif DEAD

Les protéines DEAD contiennent la séquence Asp-Glu-Ala-Asp (D-E-A-D) en motif II. Ce sont des protéines très conservées, déjà présentes chez les bactéries et les virus, et qui se retrouvent chez l'humain. Leur substrat est l'ARN. Elles présentent une activité ATPase ARN-dépendante. Une fois fixées à l'ARN elles peuvent hydrolyser l'ATP. Certaines hélicases DEAD sont capables de déstabiliser de courts fragments d'ARN double brin ( $\sim 10$  paires de bases). Ce type d'activité permet entre autres le réarrangement de structures secondaires de l'ARN, et rend possible la directionnalité de réactions complexes (l'assemblage du ribosome en est un exemple). Le mécanisme de cette déstabilisation est toutefois encore débattu. En pratique, des protéines de ce type sont retrouvées dans presque tous les processus biologiques impliquant de l'ARN [47, 48].

#### 3.1.1 Exemples de protéines à motif DEAD

##### La protéine Ded1

Ded1 est une protéine issue de la levure *S. cerevisiae* (Figure 3.1) ; elle comporte le motif DEAD. Il s'agit de l'homologue de DDX3 chez l'humain. Elle intervient dans de nombreux processus impliquant des molécules d'ARN : épissage des ARNs messagers, transcription des ARNs issus de la polymérase III, initiation de la traduction, assemblage du ribosome 40S, dégradation de l'ARN... Ded1 a une activité ATP-ase assez importante, et une bonne capacité de liaison à l'ARN. Il a été montré qu'elle peut déstabiliser de courts fragments d'ARN double brin si ceux-ci sont adjacents d'un fragment simple brin (10-15 paires de bases), indifféremment dans les sens  $5' \rightarrow 3'$  ou  $3' \rightarrow 5'$ , mais elle est non processive. Sa fixation sur un ARN simple brin n'implique pas la reconnaissance d'une séquence spécifique. *In vivo*, son site de fixation est probablement guidé par la structure de son substrat et de protéines cofacteurs [49, 50].

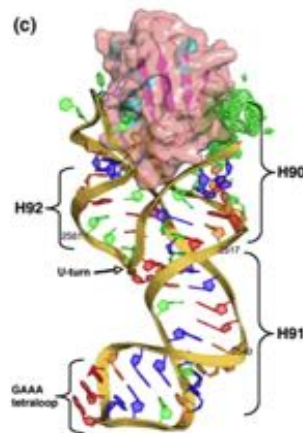


Figure 3.2 **Représentation de la fixation de DbpA au hairpin ARN H92.** Le domaine de fixation à l'ARN (RBD) de la protéine (rose) vient interagir avec le hairpin 92 du site de transfert peptidique du ribosome. Il s'agit sur ce schéma de l'analogue de DbpA chez *B. Subtilis*[54]. Figure reproduite à partir de : *Journal of Molecular Biology*, 402, Hardin et al., Copyright 2010, avec la permission de Elsevier.

### La protéine RhIE

RhIE est une protéine à motif DEAD de la bactérie *E. coli*. C'est une des rares protéines DEAD à pouvoir avoir pour substrat de l'ARN purement double brin, elle ne nécessite donc pas d'extension simple brin pour se fixer à une structure (*in vitro*). Elle n'a par ailleurs pas de séquence spécifique de fixation connue. Comme Ded1, c'est une protéine non processive, capable de déstabiliser seulement de courts fragments double brin (~ 15 paires de bases) [51]. RhIE présente une affinité pour les hélices ARN/ADN comparable à celle pour les doubles brins d'ARN.

### DbpA

DbpA est une protéine d'*E. coli* faisant partie de la famille des hélicases DEAD. Elle intervient dans l'assemblage du ribosome lors de la phase tardive de maturation de la sous-unité 23S [52, 53]. L'hélicase DbpA, non processive, est particulière en ce qu'elle possède (au moins *in vitro*) un site de fixation spécifique sur la boucle d'un hairpin au niveau du centre de transfert peptidique du ribosome (hairpin H92, Figure 3.2). Deux autres hairpins (H90 et H91) sont impliqués dans l'interaction avec la protéine, et augmentent la stabilité du complexe [54]. Ils ne sont cependant pas indispensables à l'activité de DbpA [55]. Le domaine de fixation à l'ARN (RBD) de DbpA est une partie fortement structurée de la protéine, de même que les domaines de type RecA, responsables de l'activité hélicase. Ces deux éléments sont reliés par une chaîne d'acides aminés moins structurée, plus flexible. Une fois le RBD fixé sur son site, cela permet de mettre en contact les domaines « RecA » avec une hélice d'ARN à déstabiliser. Contrairement à la fixation des domaines de type RecA, la fixation de DbpA à l'hélice H92 ne nécessite pas la présence d'ATP.

### 3.1.2 Étude des protéines hélicases en molécule unique

Avec les avancées de différentes techniques d'observation et de manipulation de molécules uniques (en particulier les pinces optiques et le FRET), les années 90 ont marqué les débuts d'études biomécaniques de protéines variées [56]. L'étude des protéines hélicases en molécule unique remonte aux années 2000. Elle concerne principalement les hélicase à ADN, plus processives, avec tout d'abord des expériences en FRET [57], puis par piégeage optique [58, 59, 60]. Des études plus récentes utilisent conjointement la mesure de force et les observations en fluorescence [61]. En ce qui concerne les hélicases à ARN, les résultats obtenus en molécule unique sont plus limités. Une étude parue en 2001 [62] rapportait l'observation par imagerie AFM d'une activité de séparation des doubles brins d'ARN par DbpA. Plus récemment, la déstabilisation de structures secondaires d'ARN par des hélicases DEAD (CYT-19 et Ded1) a été observée en FRET [63]. Enfin, un article publié en juin 2015 rapporte l'observation par mesure de force sur molécule unique de la déstabilisation d'un hairpin ARN par un complexe protéique constitué d'une hélicase DEAD (eIF4A) et de plusieurs cofacteurs [64]. La manipulation d'ARN par piégeage optique étant courante dans notre laboratoire, nous avons débuté une étude de l'activité d'hélicase à ARN (de la famille DEAD) par cette méthode.

## 3.2 Notre étude de l'activité des protéines à motif DEAD par mesures de forces

Des expériences préliminaires avec les protéines Ded1 et RhlE ont été réalisées sur les hairpins RNA10 et RNA18 décrits dans le chapitre précédent. La boucle (simple brin) du hairpin devait servir de site de fixation à Ded1, et nous nous attendions à une modification du signal d'ouverture des hairpins en présence de l'une ou l'autre de ces hélicases. Aucun résultat en ce sens n'ayant pu être observé avec Ded1 ou RhlE, nous avons décidé de nous concentrer sur DbpA, qui a l'avantage de posséder un site de fixation spécifique, et aussi une meilleure affinité pour l'ARN (du moins sur ce site) que les autres protéines DEAD.

### 3.2.1 Les constructions pour l'étude de DbpA

Afin d'observer l'activité de DbpA, nous avons tiré parti de la présence du site de fixation spécifique pour immobiliser la protéine à proximité de l'hélice dont nous souhaitons suivre la déstabilisation. Les constructions présentées ici ont été réalisées par Hala Bounihi et Thierry Bizebard, dans le cadre d'une collaboration avec l'Institut de Biologie Physico-Chimique (IBPC). Elles ont été inspirées par le travail de Diges et al. [55], qui ont étudié la déstabilisation d'une hélice ARN par DbpA sur un substrat bimoléculaire. Dans leurs expériences, un oligomère d'ARN (<10 bases) est hybridé à un ARN plus long contenant le hairpin H92, site de fixation de la protéine. Pour nos expériences en molécule unique, un fragment d'ARN a été ajouté afin de fermer la construction (Figure 3.3).

Les premières expériences sur ce substrat n'ont montré aucun signe de présence de DbpA. Afin d'améliorer l'affinité de DbpA pour la séquence, les hélices voisines impliquées dans sa fixation ont donc été ajoutées. La séquence d'ARN retenue, présentée figure 3.4, contient donc les hairpins H92 (site principal de fixation) et H90 de l'ARN ribosomique. L'hélice H91

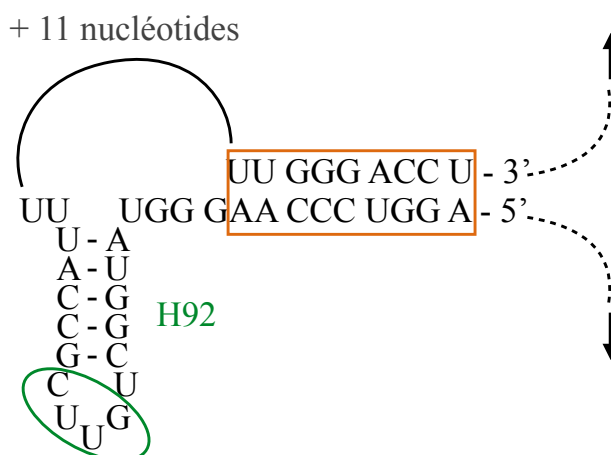


Figure 3.3 **Séquence pour l'étude de l'activité de DbpA en molécule unique.** Cette séquence contient le hairpin H92. Les quatre nucléotides du site de fixation spécifique de DbpA sont entourés en vert. Le substrat bimoléculaire utilisé par Diges et al. [55] (séquence explicite) a été fermé par une séquence arbitraire de 11 nucléotides. L'hélice cible de la déstabilisation est indiquée en orange.

a été prolongée afin d'ajouter une hélice cible pour la déstabilisation, tout en conservant une construction fermée. Il est ainsi possible d'effectuer, comme pour les hairpins du chapitre précédent, plusieurs cycles d'ouverture et de reformation sur une même molécule. De la même façon que pour les constructions à hairpin, la séquence d'intérêt est insérée entre deux poignées ARN d'environ 2000 bases. Ces poignées sont ensuite hybridées à leur complémentaires d'ADN, préalablement fonctionnalisés afin de pouvoir être liées à des billes micrométriques (fonctions biotine/ streptavidine et digoxygénine/anti-digoxygénine, figure 1.3). Afin de favoriser la formation de la structure secondaire attendue, les constructions sont placées dans un tampon contenant un sel divalent (50 mM KCl, 50 mM Hepes pH 7.5, 5 mM MgCl<sub>2</sub>). Ces conditions correspondent à celles utilisées majoritairement par l'équipe d'Uhlenbeck [55].

### 3.2.2 Conditions expérimentales

Les constantes cinétiques d'activité de DbpA dans les expériences d'enzymologie classiques sont bien définies [65]. DbpA peut être vue comme une ATPase activée par la présence de son substrat ARN. Dans une solution saturée en ATP, la constante apparente de fixation avec l'ARN mesurée est 10 nM. Les expériences en molécule unique sont toutefois très différentes des expériences d'enzymologie : ce n'est pas l'activité moyenne de la population de protéines qui est sondée, mais l'activité individuelle d'une protéine présente sur la molécule d'ARN immobilisée au moment de la mesure. Il faut donc se placer dans des conditions où statistiquement, toutes les molécules d'ARN sont occupées par une protéine. Pour cela, il est nécessaire d'utiliser des concentrations bien supérieures à celles suggérées par les constantes mesurées par les tests classiques. De plus, il est très difficile de connaître la concentration en

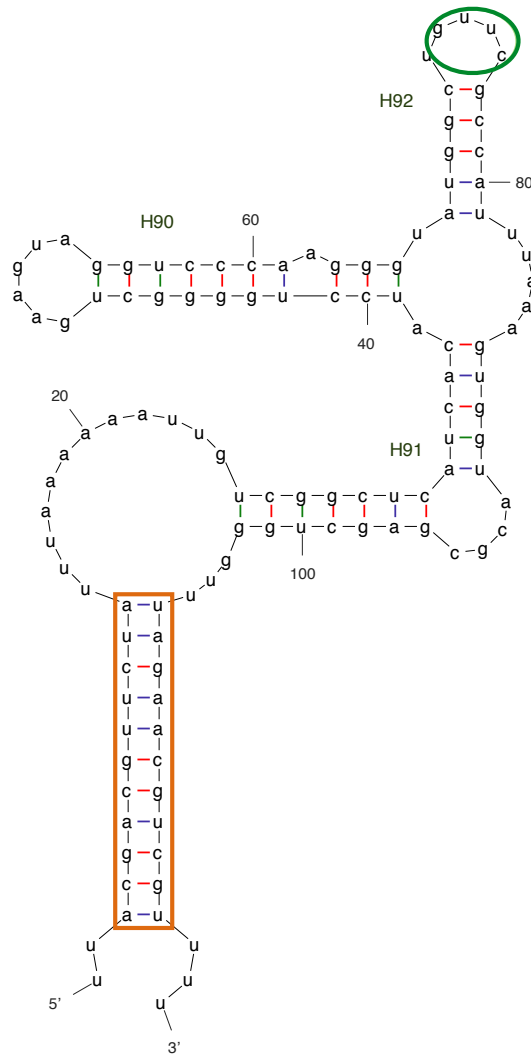


Figure 3.4 **Séquence pour l'étude de l'activité de DbpA en molécule unique.** Cette séquence contient les hairpins H90, H91, et H92 impliqués dans l'interaction du RBD de DbpA avec l'ARN ribosomique 23S. Les quatre nucléotides du site de fixation spécifique sont entourés en vert, et l'hélice cible de la déstabilisation est indiquée en orange. Cette construction est aussi adaptée à partir des études menées par l'équipe d'Uhlenbeck [55].

protéines réellement présente lors d'une mesure. En effet, les protéines ont fortement tendance à s'adsorber aux interfaces. Contrairement aux expériences réalisées « en tube », où le volume de réaction est grand et la surface de contact avec le récipient faible, nos expériences « sur lamelle » présentent le rapport d'aspect inverse. La surface de contact avec les lamelles est grande, et additionnée de la surface de contact avec les micro-billes utilisées pour le piégeage. La concentration effective de protéines est donc notablement plus faible dans la cellule expérimentale que dans la solution initiale. Nous avons donc utilisé de très fortes concentrations en protéines, mais les valeurs indiquées ne correspondent pas aux concentrations effectives. La protéine est ajoutée juste avant le début de l'expérience : après l'incubation des billes avec les constructions présentées précédemment, la solution est diluée dans un tampon favorable à l'activité de DbpA (50 mM KCl, 50 mM Hepes pH 7.5, 5 mM MgCl<sub>2</sub>, 5% glycerol, 0.1 mM DTT, 0.1 mg/mL BSA) [55]. Sont ensuite ajoutés, selon les expériences, 1 ou 2 μM de DbpA et 2 mM d'ATP.

### 3.3 Les résultats obtenus

#### 3.3.1 Expériences à vitesse constante

##### Caractérisation des constructions en l'absence de protéines

En l'absence de DbpA, la réponse des constructions à une rampe de déplacement (50 nm/s) est celle attendue pour la déstabilisation d'une structure secondaire de taille modérée (Figure 3.5). Comme décrit dans le chapitre 2, la première partie de la courbe montre une augmentation de la force avec le déplacement. Cela correspond aux réponses entropique, puis enthalpique des poignées hybrides ADN/ARN. Le modèle WLC décrit bien la réponse en force dans cette zone. Aux alentours de 9 pN, une chute brutale en force se produit, d'une amplitude d'environ 2 pN. Une certaine variabilité a été observée au niveau de ce saut en force : dans environ 1/3 des cas, celui-ci présente une étape intermédiaire, de durée variable (Figure 3.6). Ce saut ne correspond en général pas à l'ouverture complète de la construction, et des flips sont souvent observés entre l'état d'arrivée du saut et l'état d'ouverture complète (Figure 3.6). Une étude plus systématique du nombre de paires de bases ouvertes à chaque étape révèle une grande variabilité dans les états intermédiaires rencontrés. Ceci est probablement dû à la présence d'interactions non Watson Crick dans la grande boucle d'ARN. Par comparaison, dans le cas du hairpin RNA18 du chapitre précédent (qui présente une boucle de taille similaire), l'ouverture de la tige intervient aux alentours de 15 pN, sans étape intermédiaire. Avec une force d'ouverture presque deux fois moindre, des interactions non Watson-Crick dans la boucle de la construction DbpA peuvent être révélées. De plus, des interactions tertiaires sont prévues entre les différentes hélices de la construction, ce qui peut complexifier son signal d'ouverture. Enfin, il n'est pas à exclure que des structures secondaires alternatives à celle attendue soient présentes dans notre échantillon. L'occurrence du saut peut toutefois être attribué à la présence de la tige à la base de la construction. En effet, en présence d'un oligo compétiteur à cette séquence, le saut disparaît après une première ouverture de la construction. L'ouverture devient alors progressive, avec un signal assez mal défini.



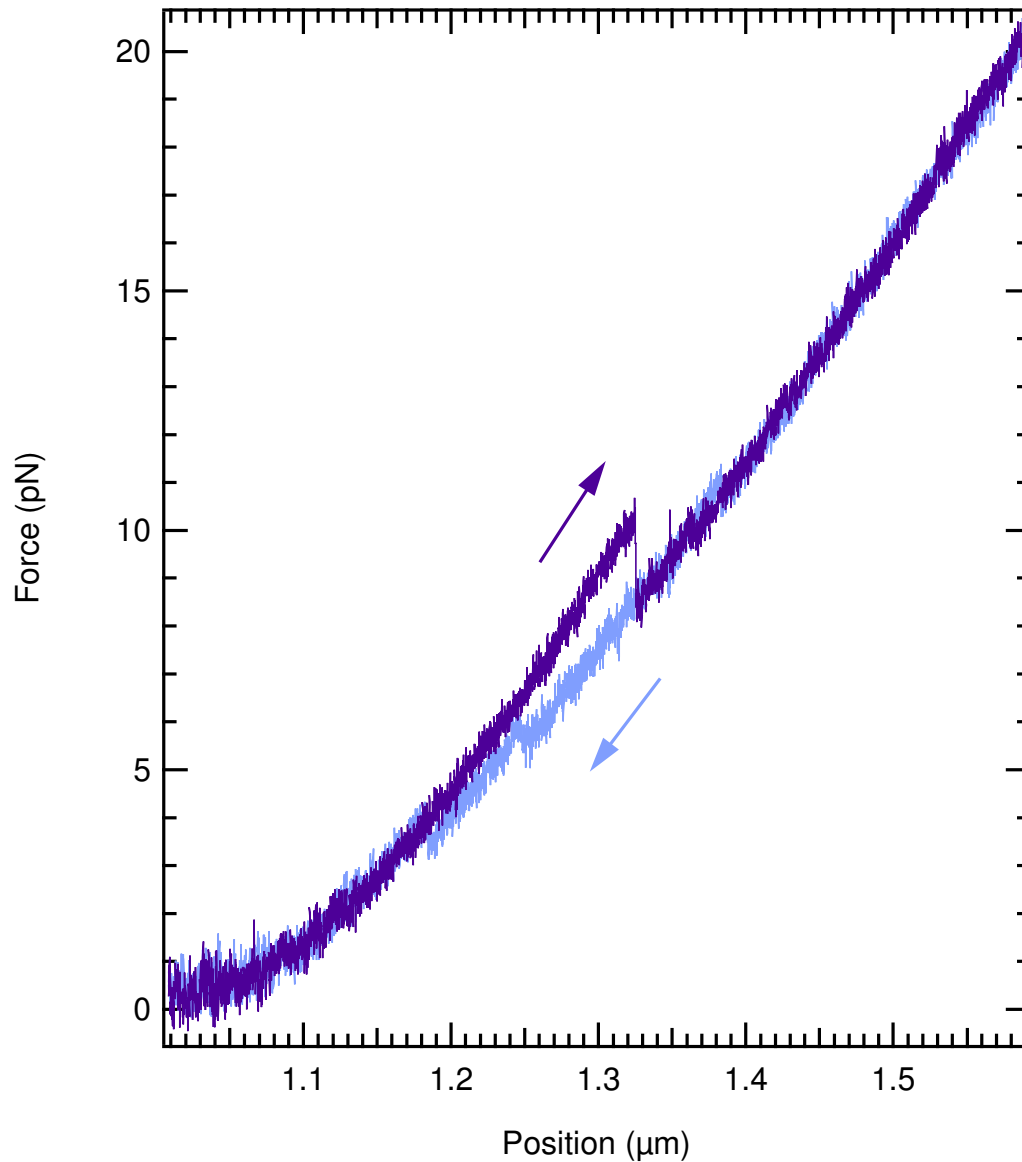


Figure 3.5 **Courbe de réponse en force des constructions pour l'étude de DbpA en l'absence de la protéine.** La courbe foncée correspond à l'extension de la construction. La première partie de la courbe correspond à la réponse élastique des poignées hybrides. Aux alentours de 10 pN, une chute brutale en force est observée. Il s'agit de la déstabilisation du fragment d'ARN ribosomique structuré contenant le site de fixation de DbpA. La courbe claire correspond à la relaxation de la construction. La re-structuration de la séquence est progressive.

Le signal de référence en l'absence de protéine sera donc défini par un saut net se produisant à une force toujours supérieure à 6 pN, et pouvant présenter une étape intermédiaire. Ce signal n'est pas modifié par l'ajout d'ATP dans la solution. Une hystérésis est visible entre les signaux d'ouverture et de reformation de la construction. Cette dernière intervient à basse force, généralement en plusieurs étapes.

### **Indices de la fixation de DbpA, variabilité**

D'après la conception des constructions de cette étude, DbpA devrait faciliter la déstabilisation de l'hélice à la base de la structure. Le signal attendu devrait donc présenter une ouverture à plus basse force, et potentiellement plus progressive que le signal de référence. L'ouverture du reste de la construction sera accompagné de l'éjection de la protéine. Selon l'effet stabilisant ou déstabilisant de la fixation de cette dernière, l'ouverture pourrait se produire à plus haute force, ou au contraire être confondue avec l'ouverture de la tige à la base de la structure.

Comme détaillé en introduction, DbpA peut se fixer à l'ARN en absence d'ATP (fixation au hairpin H92), mais la déstabilisation d'une double hélice se fait avec consommation d'ATP. Un signal spécifique de l'activité hélicase de DbpA ne sera donc observé que si la solution contient à la fois la protéine et de l'ATP. Aucune différence de ce type n'a été observée entre les expériences avec et sans ATP, en présence de DbpA.

Lorsque DbpA est ajoutée au mélange expérimental (avec ou sans ATP, donc), les signaux d'ouverture observés se diversifient. Les profils observés en l'absence de la protéine sont toujours présents, mais des ouvertures à des forces inférieures à 6 pN sont aussi observées. Dans ces cas, le signal peut comme précédemment présenter un saut net, avec ou sans étapes. Des profils nouveaux apparaissent, qui contiennent des zones d'ouverture progressive. Les signaux enregistrés sont donc très variables, aussi bien au niveau de leur forme que de la force d'ouverture, qui peut varier de 4 à plus de 10 pN. Bien évidemment, certains de ces signaux correspondent à des cas où la protéine est absente. Il est donc attendu d'obtenir une certaine proportion de profils similaires au cas sans protéine. La figure 3.7 présente un exemple d'ouverture de la construction en présence de DbpA où le profil observé est très différent du cas de référence. Aucun saut de grande amplitude n'est observé, mais l'ouverture semble se produire assez régulièrement à partir de 6 pN. Après cette zone, lorsque la construction est entièrement déstabilisée, les courbes avec et sans protéine se superposent parfaitement. Les hélices de fixation de DbpA ont donc elles-aussi été déstabilisées, et la protéine a été éjectée. Cet état final similaire au cas sans protéine est systématiquement observé (tous les enregistrements ont été faits jusqu'à une force finale d'au moins 15 pN).

### **Classification des profils d'ouverture observés**

Afin de quantifier les fréquences relatives des différents profils, ceux-ci ont été classés en 8 catégories, détaillées figure 3.8. Celles-ci sont basées sur la présence et l'ordre des étapes progressives et des sauts. Les deux dernières (6b et 7) représentent le cas sans DbpA. Des exemples correspondant aux différentes catégories sont donnés figure 3.9. Les 5 premières

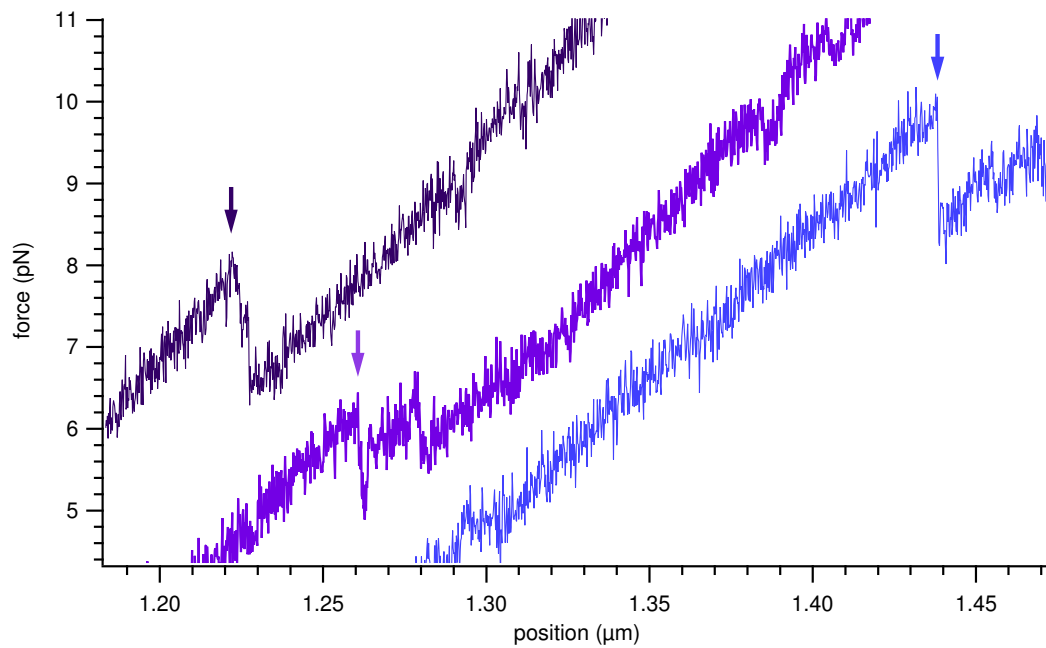


Figure 3.6 **Agrandissement de courbes de réponse en force des constructions pour l'étude de DbpA en l'absence de la protéine.** Les flèches indiquent le début de l'ouverture de la construction. Trois cas sont représentés ici. Les courbes ont été arbitrairement décalées en position par rapport à la courbe centrale pour une meilleure lisibilité. Courbe bleu clair : saut net sans étape. Courbe bleu foncé : Saut avec étape ; des flips sont observés  $\sim 0.05 \mu\text{m}$  après le saut. Courbe violette : saut avec étape, puis flip sur l'état intermédiaire du saut.

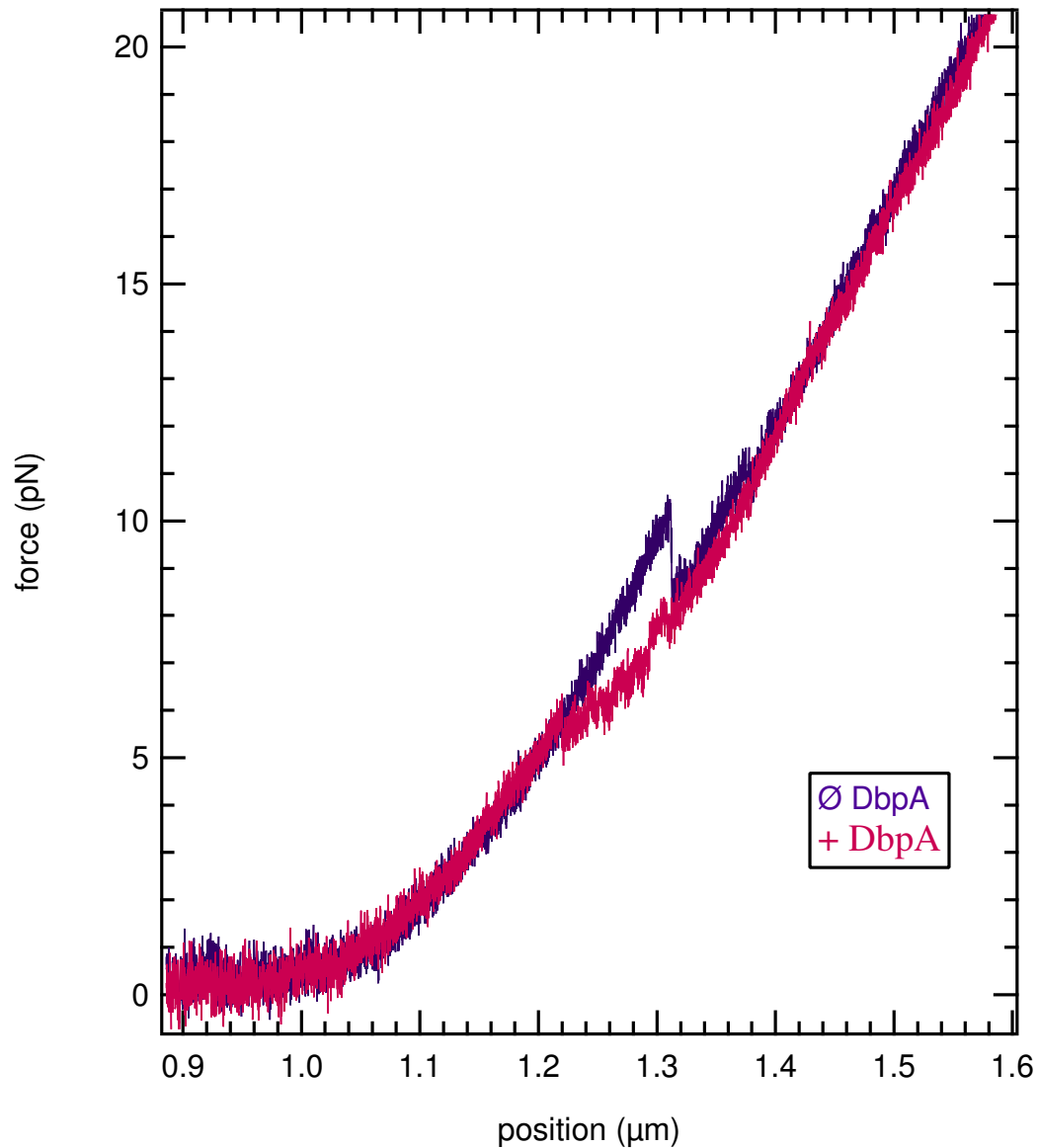


Figure 3.7 **Courbe de réponse en force des constructions pour l'étude de DbpA en absence et en présence de la protéine** La courbe violette correspond à la réponse en force des constructions en absence de la protéine, déjà décrite figure 3.5. La courbe rose présente la réponse obtenue en présence de DbpA. Sur cet exemple, les deux courbes sont très différentes : en présence de la protéine, l'ouverture est progressive, avec quelques sauts de faible amplitude. Avant et après la zone d'ouverture, les deux courbes sont parfaitement superposées.

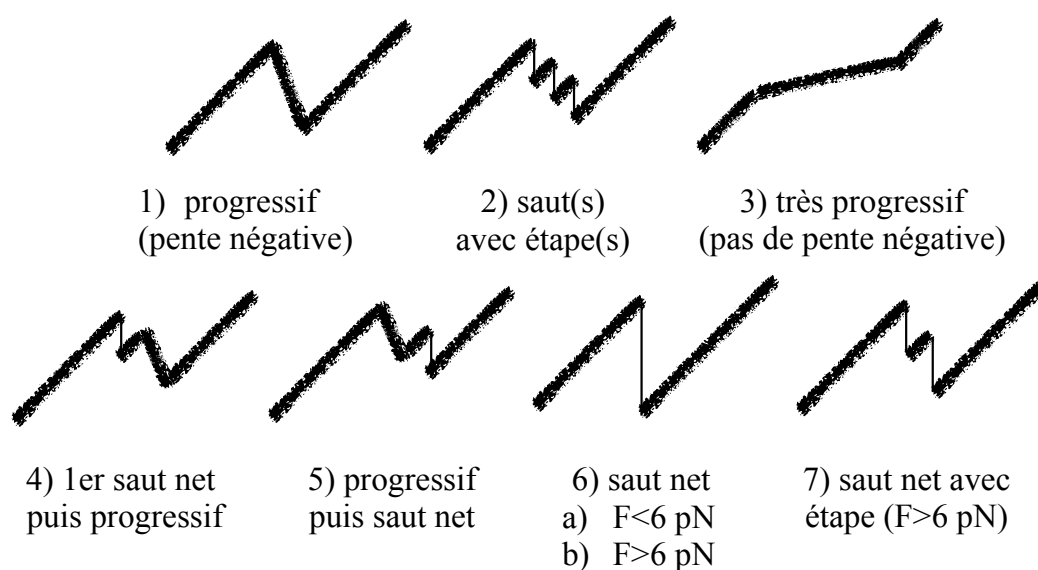


Figure 3.8 **Représentation schématique des différents profils d'ouverture en présence de DbpA.** Les différents profils ont été classés selon la présence ou l'absence de sauts nets, et le signe de la pente de la courbe dans la zone de déstabilisation de la structure secondaire.

correspondent à des profils observés uniquement en présence de la protéine, et sont caractérisées par la présence d'une zone d'ouverture progressive (3.9 zooms cat 1 et 4), ou de plusieurs sauts de faible amplitude. Ces 5 profils sont observés à des forces variables, à l'exception de la catégorie 3 (ouverture très progressive) où il n'est pas possible de définir une force d'ouverture. La catégorie 6, présentant un saut net sans étape, a été subdivisée pour isoler le cas déjà observé sans protéine : l'ouverture se produisait alors systématiquement à une force supérieure à 6 pN (6b), alors que des cas d'ouverture inférieure à 6 pN (6a) sont observés en présence de DbpA.

Les probabilités correspondant à l'apparition du profil de chaque catégorie sont représentées figure 3.10. Elles correspondent au cas où la solution expérimentale contient à la fois DbpA et de l'ATP. La présence d'un saut net à force supérieure à 6 pN (6b), profil le plus courant en l'absence de DbpA, voit sa fréquence diminuer fortement en présence de la protéine, au profit des autres catégories. Environ 40% des signaux avec DbpA présentent une ouverture progressive (catégorie 1) ou très progressive (catégorie 3), sans aucun saut net.

Comme annoncé au paragraphe précédent, ce type de signal est observé à la fois en présence et en absence d'ATP en solution. Il ne peut donc pas s'agir de l'activité hélicase de DbpA, qui est ATP-dépendante. Une modification du profil d'ouverture peut toutefois être due à la seule fixation de DbpA sur la construction. Les signaux de la catégorie 4, où un saut net est observé à une force supérieure à 6 pN et suivi d'une ouverture plus progressive, pourraient correspondre à ce cas : l'ouverture de l'hélice « cible » n'est pas modifiée (saut net), mais la

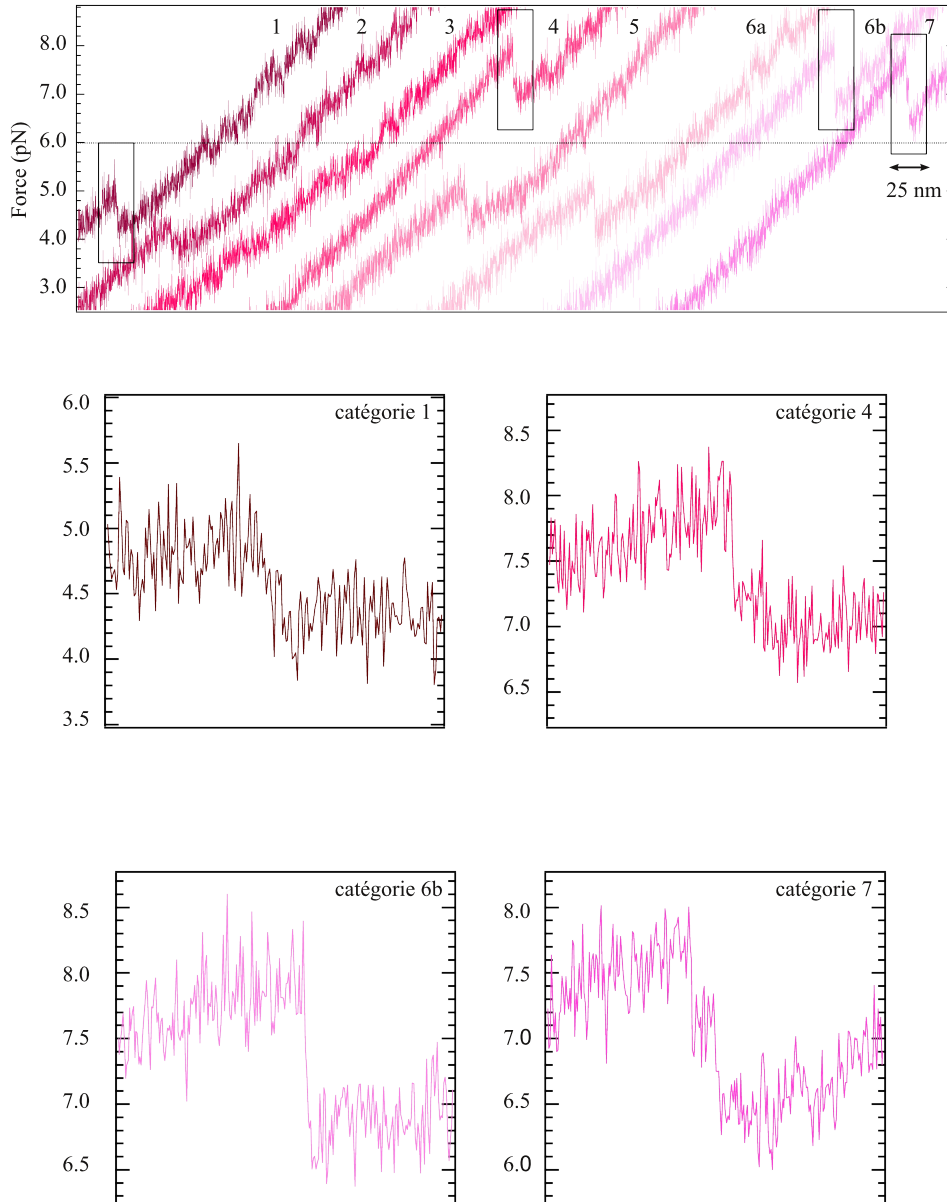


Figure 3.9 **Exemples des différents profils d'ouverture en présence de DbpA.** Le panneau supérieur présente un exemple pour chaque profil d'ouverture décrit dans la figure 3.8. Les courbes ont été arbitrairement décalées en position. Quatre agrandissements sont présentés dans les panneaux inférieurs. Les zones concernées sont matérialisées par des rectangles noirs dans le panneau supérieur.

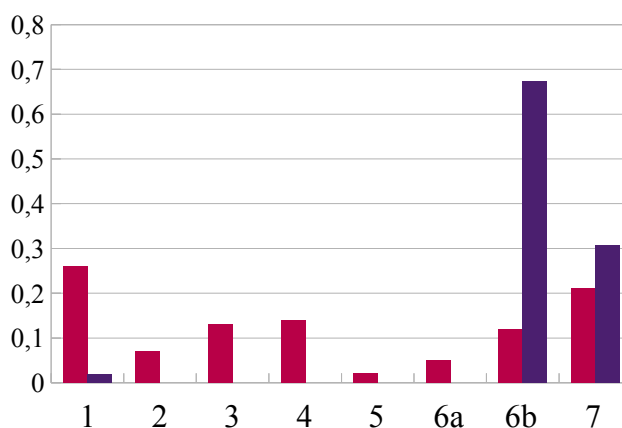


Figure 3.10 **Probabilités correspondant à l'apparition des différents profils observés en absence et en présence de DbpA.** Les barres violettes correspondent aux profils observés en l'absence de DbpA, les rouges au cas où la protéine est présente, dans une solution contenant de l'ATP. Environ 40% des signaux avec DbpA présentent une ouverture progressive (catégorie 1) ou très progressive (catégorie 3), sans aucun saut net. Dans ces mêmes conditions, la proportion d'un saut net à haute force (6b) diminue fortement.

déstabilisation des hélices impliquées dans la fixation de DbpA est modifiée par sa présence. Les cas d'ouverture à plus basses forces pourraient traduire l'interaction des domaines de type RecA de DbpA avec l'hélice cible, sans pour autant que cette interaction ne soit active. Il n'est pas à exclure que la présence de la protéine modifie la structuration de la construction, et que la géométrie obtenue ne soit donc pas celle attendue. Une structure secondaire avec des hélices moins stables présenterait une ouverture à plus basse force, éventuellement moins structurée. DbpA est connue pour être une hélicase peu processive, son activité est donc limitée à la déstabilisation d'hélices courtes. L'hélice cible qui lui est ici présentée a une énergie d'environ  $28 k_B T$  à  $29^\circ C$ , ce qui correspond aux stabilités maximales sur lesquelles une activité de DbpA a été rapportée [55]. Il est possible que pour une telle hélice, l'activité hélicase de DbpA soit systématiquement masquée par l'ouverture mécanique imposée. Quelques expériences complémentaires ont été réalisées afin de contourner ce problème.

### 3.3.2 Expériences complémentaires

#### Expériences à position constante

Afin de ne pas ouvrir mécaniquement la construction avant que DbpA n'ait pu déstabiliser son hélice cible, des expériences à position constante ont été réalisées. Dans ce cas, les pièges sont amenés à une distance telle que les poignées sont alignées et légèrement étirées, mais l'hélice cible n'est pas ouverte. Avec les constructions initiales, il faut pour cela imposer une force d'environ 3 pN. A une force si petite, le rapport signal/bruit n'est pas satisfaisant. Nous avons donc préparé de nouvelles constructions avec une hélice cible allongée de 3 nucléotides afin d'augmenter la stabilité de la première hélice à ouvrir. Ces constructions ont

été réalisées à l'IBPC par Virya Ok et Thierry Bizebard. Avec cette nouvelle version, il est possible de travailler à position constante sous une force d'environ 6 pN, et ainsi d'avoir un rapport signal/bruit acceptable. L'hélice cible a alors une stabilité supérieure à celle que peut ouvrir DbpA ( $\sim 37 k_B T$  à 29°C), mais l'application d'une force réduit sa stabilité effective. Mais même dans ces conditions, aucune activité ATP-spécifique de type hélicase n'a pu être observée.

### Hypothèses quand à l'absence d'activité hélicase

L'activité ATP-dépendante de notre échantillon de DbpA en conditions « classiques » a bien évidemment été vérifiée. Cela a été fait en FRET, sur un substrat bimoléculaire. Le protocole détaillé (proposé par Thierry Bizebard) est présenté en annexe (Test en FRET de l'activité de DbpA).

L'activité hélicase de DbpA n'a jamais été observée sur un substrat fermé, comme l'est notre construction. Il est possible que la contrainte imposée par la fermeture de la molécule entraîne une structuration incompatible avec l'activité de DbpA. Nous avons essayé de mettre au point une expérience en FRET sur une construction fermée, afin de tester par un autre moyen l'activité de DbpA sur notre substrat. Cela n'est pas chose facile puisque dans le cas d'une construction fermée, l'ouverture de l'hélice n'implique pas l'éloignement des deux brins. Ceux-ci sont en effet maintenus proches l'un de l'autre par le reste de la construction. Les fluorophores pour le FRET ne peuvent donc pas être directement placés sur la tige à ouvrir. De plus, la réhybridation intramoléculaire des deux brins est favorisée. Le test mis au point reposait donc sur un signal de FRET (ou de quench) entre un fluorophore placé sur un oligo compétiteur de l'hélice cible (en grand excès), et un autre placé sur un oligo complémentaire à la poignée prolongeant la boucle (Figure 3.11). Les résultats préliminaires de ce test n'ont pas fourni de réponse quant à l'activité de DbpA, et il n'a pas été poursuivi. Ce test est mentionné ici dans l'hypothèse d'une reprise ultérieure de ces travaux.

## 3.4 Conclusion

Malgré quelques résultats suggérant un effet de la protéine DbpA sur la déstabilisation des hélices ARN, nous n'avons pas obtenu de signature claire de son activité hélicase. L'objectif initial du projet était de tirer profit d'un outil qui paraissait parfaitement adapté, le double piège optique, pour visualiser simplement l'activité hélicase des protéines à motif DEAD sur une double hélice d'ARN. Il semble que notre protocole expérimental ne le permette pas. L'hypothèse la plus probable est que la construction préparée n'adopte pas la structure prévue, et ne permette donc pas la fixation correcte de DbpA. La séquence ajoutée pour obtenir la fermeture de la construction peut aussi avoir un effet défavorable pour la fixation de DbpA. Nos mesures semblent bien indiquer une interaction de DbpA avec la structure d'ARN, mais quelle qu'en soit la raison, cette interaction ne permet pas la fonctionnalité de la protéine. Une autre hypothèse est que la mise sous tension de l'hélice cible de l'hélicase empêche son activité. Le problème viendrait alors de la nature même de la mesure.



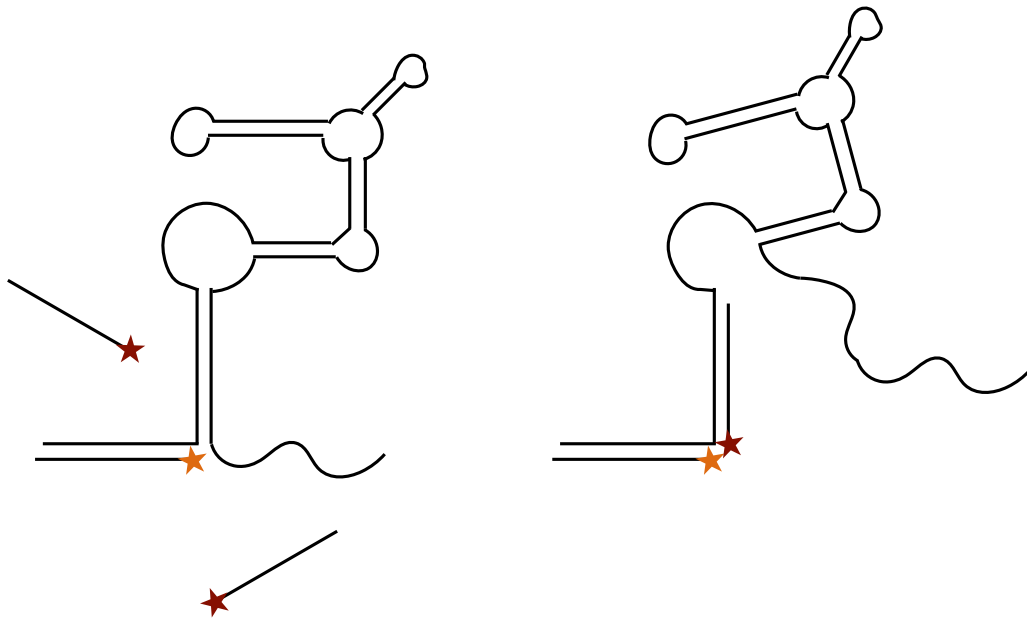


Figure 3.11 **Représentation schématique du test de l'activité hélicase de DbpA sur un substrat fermé.** La construction utilisée pour le test en FRET ne contient que la partie structurée de la construction utilisée en molécule unique, les poignées hybrides ont été retirées. Seules 20 bases environ sont conservées de chaque côté. Gauche : en l'absence de DbpA, l'hélice cible est hybridée. Un des oligonucléotides portant un fluorophore est hybridé à la courte poignée prolongeant la boucle. Droite : en présence de DbpA, l'hélice cible est ouverte, et le deuxième oligonucléotide (en grand excès) peut être hybridé. Les deux fluorophores sont alors suffisamment proches pour qu'un signal de FRET (ou de quench) ait lieu.

La publication très récente d'une étude de l'hélicase eIF4A (spécifique de l'ARN, appartenant à la famille DEAD) par mesure de force [64], dans une configuration de double piège optique proche de la nôtre, apporte une perspective différente. L'équipe de Block a montré qu'eIF4A n'a une activité hélicase processive ATP-dépendante que lorsqu'elle est associée à au moins deux de ses cofacteurs (eIF4G et eIF4B). Hors de ce complexe, la déstabilisation de l'ARN double brin est soit absente (si la concentration en eIF4A est inférieure à 1  $\mu\text{M}$ ), soit indépendante de la présence d'ATP (pour une concentration de 3-5  $\mu\text{M}$ ). Ce dernier résultat, très proche de ce que nous avons observé avec DbpA, est interprété comme une capture par la protéine d'une ouverture spontanée de quelques paires de bases. Il ne s'agit alors pas d'une activité catalytique, et ce processus ne nécessite pas d'ATP. Il requiert cependant une très grande concentration de protéines, et est bien moins efficace que l'ouverture catalytique en présence du complexe. Il est possible que de la même façon, l'activité de DbpA soit fortement dépendante de protéines cofacteurs. Contrairement au cas de eIF4A, peu de données à ce sujet sont disponibles.

Une étude fine visant d'une part à déterminer précisément la structure adoptée par l'ARN, et d'autre part à étudier l'activité de DbpA sur des substrats fermés, nous a semblé dépasser largement le cadre de cette thèse et sortir tout à fait de ses objectifs initiaux. En l'absence de données supplémentaires quant à d'éventuels cofacteurs de DbpA, la mise en place d'un système proche de celui utilisé pour eIF4A paraît prématurée. C'est pourquoi, après avoir consacré un temps beaucoup plus long qu'initialement prévu à l'étude des protéines à motif DEAD par mesure de force, nous avons décidé de ne pas poursuivre ce projet.



# Chapitre 4

## Surétirement d'hybrides ARN/ADN

*Il n'est point nécessaire que toutes les questions trouvent leur réponse. Pour les plus importantes, c'est déjà beaucoup qu'elles aient été posées.*

---

Friedrich Hebbel

### 4.1 La double hélice d'ADN sous haute force : le surétirement

#### 4.1.1 Les débuts

Dans les chapitres précédents, les forces appliquées sur les acides nucléiques n'ont jamais dépassé 20-30 pN. Dans ces conditions, la structure en double hélice du polymère n'est pas altérée. Cependant, les polyacides nucléiques peuvent supporter des forces bien supérieures sans rupture. La caractérisation des acides nucléiques sous haute force a débuté avec les expériences de peignage moléculaire (1994-1995) : des ADN double brin fixés par une extrémité à une surface étaient alors alignés, puis étirés par le recul d'une interface air-eau [66, 67]. Des forces supérieures à la centaines de piconewtons peuvent ainsi être appliquées, avec pour effet une extension des molécules de plus de 2 fois leur longueur cristallographique initiale. Ce phénomène a ensuite été caractérisé avec plus de précision, notamment grâce à des expériences de piégeage optique. Il a ainsi été montré que lorsque la force exercée sur un ADN double brin atteint  $\sim 65$  pN, une transition se produit : la molécule s'allonge alors à force quasi constante, jusqu'à atteindre une longueur de 70% supérieure à sa longueur cristallographique initiale [68, 69]. C'est ce qu'on appelle le phénomène de surétirement. Ces observations ont été interprétées comme une altération de la structure hélicoïdale de la molécule, soit par déroulement de l'hélice, soit par séparation des brins (Figure 4.1). Ce phénomène est réversible : si la force exercée sur la molécule est relâchée, celle-ci peut retrouver sa structure initiale. Dès ces premières expériences il a été observé que les conditions expérimentales, en particulier la

composition du tampon, influent à la fois sur la force à laquelle se produit la transition et sur la présence d'une hystérésis lors du retour à la structure initiale.

#### 4.1.2 Les différents mécanismes de surétirement de l'ADN

Selon la topologie du double brin d'ADN étudié, différents profils de surétirement ont été observés par mesure de force sur molécule unique. En cela, le mode d'attachement de la construction à son support (billes piégées ou lamelle) a un rôle important [70, 71, 72]. Lorsque la molécule est contrainte en torsion, c'est à dire qu'elle ne peut pas tourner sur elle-même librement, le plateau de surétirement intervient à des forces supérieures à 100 pN. Si cette contrainte est relâchée par la présence d'une liaison simple permettant la rotation (au moins une extrémité libre, « nicks », attachement sur une boucle d'ADN...), la transition intervient autour de 65 pN comme observé initialement. Différentes études ont été réalisées afin d'obtenir des informations détaillées sur le mécanisme moléculaire à la base du surétirement. En particulier, le couplage des expériences de mesure de force avec la mesure de fluorescence permet de repérer la présence d'ADN simple brin dans la molécule en surétirement. Cette technique a permis de mettre en évidence 3 modes de surétirement dont la présence dépend à la fois de la topologie de la molécule, de sa séquence (et donc des variations locales de stabilité), de la vitesse d'étirement, et de la composition du tampon expérimental [73].

Le premier mode correspond à la génération de simples brins à partir d'une discontinuité topologique (Figure 4.1). Les liaisons hydrogène entre les brins sont alors rompues. Devant l'absence de traduction satisfaisante, ce mode sera ici dénommé « peeling ». Dans le deuxième mode, des liaisons hydrogène sont aussi déstabilisées, mais les deux brins restent ancrés l'un à l'autre par des zones hybridées. Il se forme localement des « bulles » où deux simples brins se font face, mais la molécule reste topologiquement fermée. Le dernier mode correspond à un déroulement de la double hélice, mais les liaisons hydrogène sont maintenues. Cela correspond à la forme S de l'ADN, en opposition à la forme B de la double hélice classique. C'est le peeling qui induit la présence d'une hystérésis sur les signaux de mesure de force (Figure 4.2). Dans ce mode, l'un des brins est libre de former une pelote statistique favorable aux polymères en solution, alors que l'autre est maintenu sous étirement. Lorsque la force est diminuée, la réhybridation des brins est retardée par rapport au cas où ceux-ci sont maintenus en contact. En outre, lorsque l'extension de la molécule est réalisée à faible vitesse (typiquement jusqu'à quelques centaines de nm/s), le peeling présente une signature caractéristique en dents de scie.

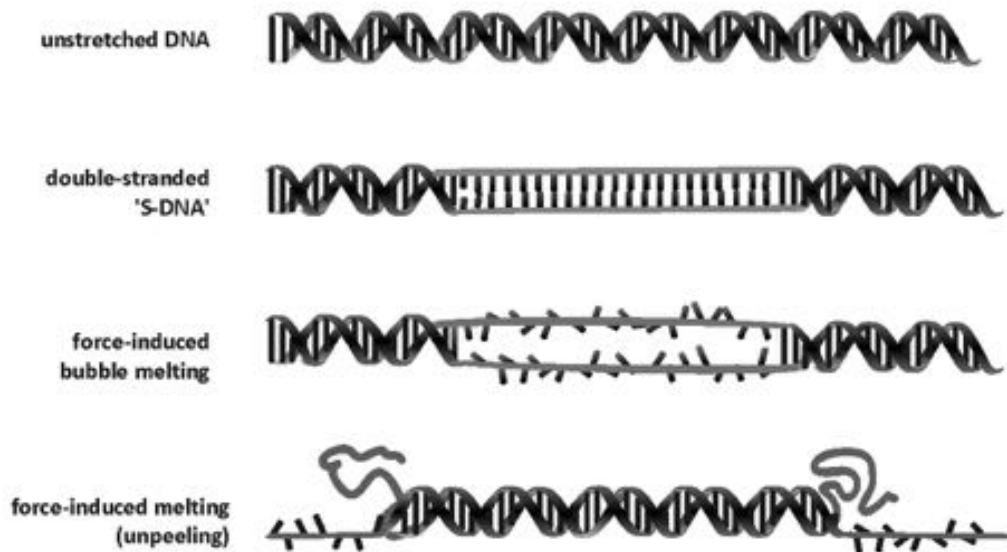


Figure 4.1 **Représentation schématique des différents modes de surétirement.** 1ère ligne : double hélice d'ADN libre de toute contrainte mécanique. Les bases sont appariées, et la structure hélicoïdale. 2ème ligne : ADN double brin sous forme S. La structure hélicoïdale est localement perdue. L'appariement des bases est maintenu. 3ème ligne : « bulle fondue ». Les liaisons Watson-Crick entre les bases sont rompues sur un domaine de la molécule. De part et d'autre de ce domaine, les liaisons sont conservées. Les deux simples brins ainsi générés sont donc maintenus sous tension. 4ème ligne : peeling de l'ADN. Les liaisons Watson-Crick entre les bases sont rompues à partir d'extrémités libres. Un des simples brins générés est ainsi libéré de toute contrainte mécanique, alors que l'autre est maintenu sous tension. Figure issue de la thèse de Peter Gross.

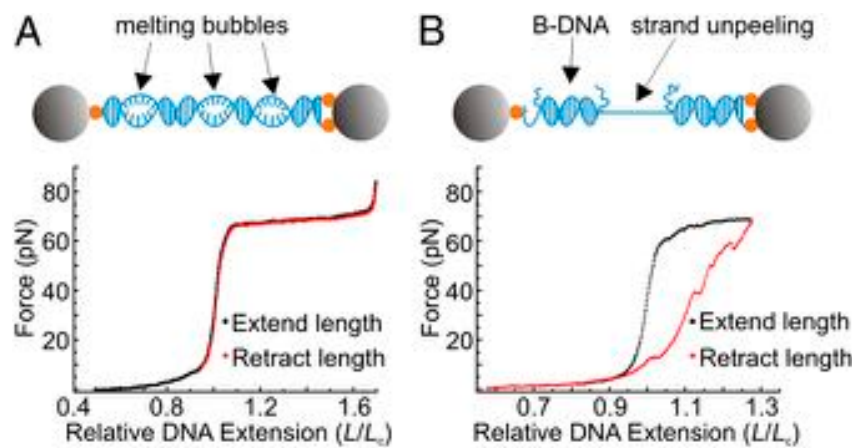


Figure 4.2 **Les deux modes de surétirement de l'ADN avec séparation des brins : bulles fondues versus peeling [73]** A. Surétirement d'un ADN double brin sans extrémité libre, non contraint en torsion. Dans ces conditions, des zones non hybridées apparaissent, isolées les unes des autres par des fragments où les liaisons hydrogènes sont conservées. Les « bulles fondues » ainsi formées peuvent éventuellement grandir, puis coalescer. Pour ce régime de surétirement, un plateau lisse est observé (courbe noire). La réhybridation se fait sans hystérésis (courbe rouge). B. Surétirement d'un ADN double brin avec une ou plusieurs extrémités libres. Dans ces conditions, le peeling peut avoir lieu. Des fragments simple brin sont libérés au niveau de ces extrémités libres. Cette libération n'est pas continue mais se fait par étapes successives (de type « stick and slip » [74]). Pour ce régime de surétirement, un plateau en dents de scie est observé (courbe noire). La courbe présentée ici correspond à une construction contenant plusieurs extrémités libres, et a été enregistrée à grande vitesse ( $3 \mu\text{m/s}$ ). Pour ces deux raisons, les dents de scie sont assez mal définies. La réhybridation se fait avec une hystérésis (courbe rouge) qui peut être de forte amplitude. Figure extraite de King *et al.*, PNAS, 2013 [73].

### 4.1.3 Modèle du « twistable worm » pour le surétirement de l'ADN.

Dans les expériences de surétirement, les chaînes d'acides nucléiques sont soumises à des forces bien supérieures à celles rencontrées dans les expériences décrites aux chapitres précédents. En particulier, le comportement des doubles brins ne peut plus être décrit par le modèle WLC extensible simple présenté au chapitre 2 et rappelé ci-dessous,

$$\langle x \rangle = L_c \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p} + \frac{F}{K}} \right)$$

où  $x$  représente la longueur bout-à-bout de la molécule,  $L_c$  et  $L_p$  ses longueurs cristallographique et de persistance, et  $F$  la force exercée.

Le couplage entre l'extension et la torsion des hélices d'acides nucléiques ne peut plus être négligé dans le terme enthalpique  $F/K$ . L'énergie élastique d'une chaîne présentant simultanément un module de torsion  $C$  et un module élastique  $K$  comporte 3 termes : deux correspondent aux énergies de torsion et d'étirement, et le troisième prend en compte le couplage entre ces deux effets, représenté par une fonction  $g(F)$ . En tenant compte du travail de la force  $F$ , l'hamiltonien enthalpique correspondant s'écrit donc :

$$H_{ent} = \frac{1}{2} \frac{C}{L_c} \theta^2 + g(F) \theta \frac{x}{L_c} + \frac{1}{2} \frac{K}{L_c} (x - L_c)^2 - xF$$

avec  $\theta$  représentant l'angle de torsion par rapport à la structure canonique de l'hélice. Le terme enthalpique  $x_{ent}$  est obtenu par minimisation de l'hamiltonien  $H_{ent}$  par rapport à la force  $F$  :

$$x_{ent} = \frac{C L_c}{-g(F)^2 + K C} \cdot F$$

La longueur de la chaîne sous tension peut donc être réécrite sous la forme :

$$\langle x \rangle = L_c \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p} + \frac{C}{-g(F)^2 + K C} F} \right)$$

Le modèle WLC fait l'hypothèse que les termes enthalpique et entropique sont indépendants. Ce dernier n'est donc pas affecté par la prise en compte d'un effet de torsion dans le terme enthalpique.

Cette description nécessite de connaître la relation  $g(F)$  entre la torsion et l'étirement de la molécule. Des mesures d'extension par piégeage magnétique [75] ont montré que pour un double brin d'ADN, la fonction  $g(F)$  est négative à basses forces, et change de signe autour de 30 pN. Cela traduit le fait qu'à basse force la torsion de l'hélice est amplifiée par l'étirement, ce qui peut s'expliquer par la non conservation de son diamètre. Une forme simplifiée de  $g(F)$  consiste à considérer que cette fonction est constante (et négative pour l'ADN) jusqu'à une force critique  $F_c$ , puis qu'elle augmente linéairement :

$$g(F) = \begin{cases} g_0 + g_1 F_c \\ g_0 + g_1 F \end{cases}$$



Cette forme décrit correctement le comportement des ADN double brin [23]. Un exemple d'ajustement d'une courbe de surétirement d'un ADN double brin par ce modèle est présenté figure 4.3. Il s'agit d'un double brin d'environ 8000 paires de bases, non contraint en torsion. Cette mesure a été réalisée au laboratoire de Nanobiophysique à partir d'une construction fournie gracieusement par le LaserLab de la VU University d'Amsterdam. Les paramètres obtenus (Tableau 4.1) sont similaires à ceux présentés dans d'autres travaux aussi bien pour l'étirement [76] que pour le couplage avec la torsion [23, 75].

#### 4.1.4 Le cas de l'ARN

Une étude comparative réalisée par l'équipe d'Arias-Gonzalez en mesure de forces par piégeage optique a montré que l'ARN double brin présente un comportement en surétirement comparable à celui d'un ADN de même séquence, avec toutefois des différences notables (Figure 4.4) [77]. Dans les deux cas, le surétirement engendre un allongement correspondant à un excès de 70% par rapport à la longueur cristallographique de la molécule au repos. Comme la longueur cristallographique de l'ARN est différente de celle de l'ADN à séquence identique, la longueur obtenue après surétirement est elle aussi différente. Cela indique que même sous haute force, la différence de structure entre les simples brins d'ADN et d'ARN perdure. Dans leurs conditions expérimentales (étirement à 500 nm/s, 150 à 500 mM NaCl), le plateau de surétirement observé est lisse, aussi bien pour l'ADN que pour l'ARN. Cependant, la force à laquelle advient le plateau de surétirement est toujours légèrement plus élevée pour l'ADN que pour l'ARN, avec une différence de l'ordre de 2 pN. En outre, l'augmentation de force entre le début et la fin du plateau est supérieure dans le cas de l'ARN ( $\sim 10$  pN) que dans celui de l'ADN ( $\sim 2$  pN) [77]. (Cette augmentation avait déjà été rapportée pour l'ADN[69], et est visible sur la courbe de la figure 4.2).

L'équipe d'Arias-Gonzalez explique ces différences en force, à la fois la plus faible valeur de force d'initiation du plateau et sa plus grande pente pour l'ARN, par une plus faible coopérativité de la transition de surétirement pour l'ARN que pour l'ADN. Par des mesures calorimétriques, cette équipe a montré que de façon similaire, la transition thermique entre les formes double et simple brin a lieu sur une gamme de température notablement plus large pour l'ARN que pour l'ADN [77]. Il est intéressant de noter que notre étude sur les hairpins, rapportée au chapitre 2, arrive à la conclusion inverse : dans le cas d'expériences d'ouvertures d'une double hélice par l'application d'une force transverse à son axe, l'ARN présente un comportement plus coopératif que l'ADN.

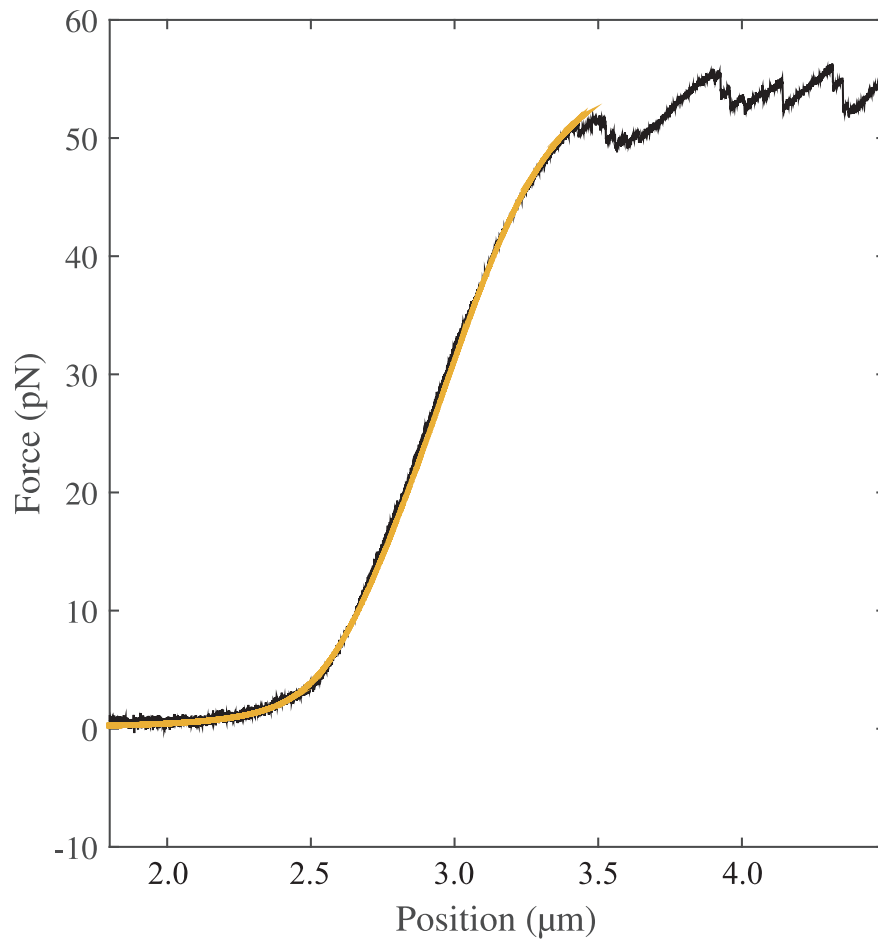


Figure 4.3 **Surétirement d'un ADN double brin en tampon PBS 1X (courbe noire).** Ajustement avec le modèle du twistable worm (courbe jaune). Les paramètres obtenus sont les suivants :  $F_c = 25.5$  pN ;  $g_0 = -605$  pN.nm ;  $g_1 = 21.85$  nm ;  $L_c = 2.79$   $\mu\text{m}$  ;  $L_p = 45$  nm ;  $K = 1150$  pN

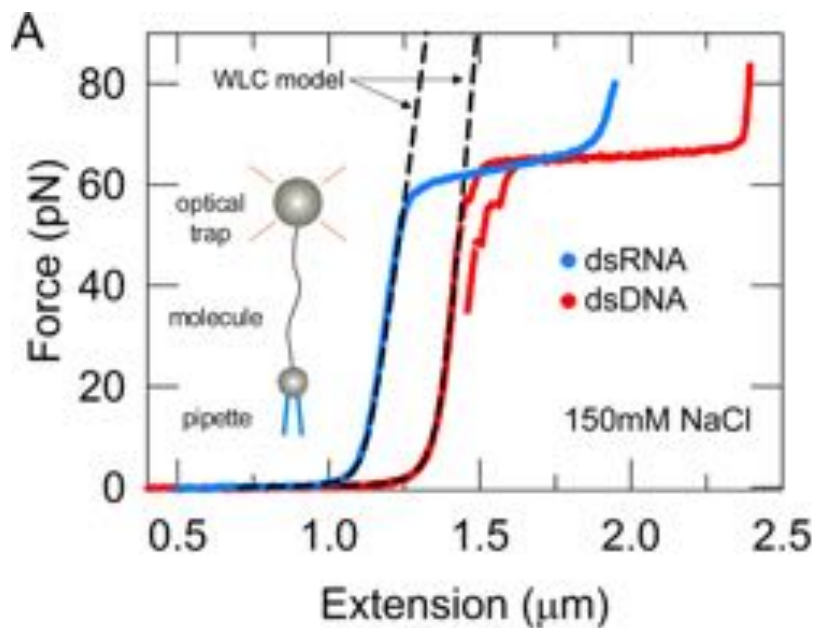


Figure 4.4 **Surétirement d'un ARN double brin (courbe bleue) et de l'ADN double brin de séquence identique (courbe rouge).** Le dispositif expérimental est schématisé sur la figure : la construction moléculaire est liée à deux billes micrométriques, l'une maintenue par une micropipette et l'autre piégée optiquement. Le déplacement de la pipette est contrôlé (500 nm/s), et la force est mesurée au niveau du piège optique. Le tampon expérimental contient 150 mM de sel monovalent (NaCl). Les deux courbes sont ajustées par le modèle WLC simple (pointillés noirs). Figure reproduite avec l'autorisation de Herrero-Galán et al., *J. Am. Chem. Soc.*, 2013. Copyright 2012 American Chemical Society.

## 4.2 Étude de molécules hybrides ARN/ADN en surétirement

### 4.2.1 Motivation de l'étude

Il n'existe jusque là pas d'étude publiée sur le surétirement de molécules hybrides ARN/ADN. Nous avons donc souhaité caractériser ce phénomène. Nous nous sommes en particulier interrogés sur la présence du mécanisme de peeling, celui-ci ayant été rapporté dans le cas de l'ADN mais pas de l'ARN. Une série d'expériences préliminaires, réalisées avec la construction hybride RNA10 présentée au chapitre 2, a permis d'observer le phénomène de peeling à une vitesse d'étirement de 50 à 200 nm/s, dans des tampons contenant 150 à 400 mM de NaCl. Les détails du processus restaient cependant inconnus, puisque la construction RNA10 possède 4 extrémités libres et susceptibles de générer le peeling (2 extrémités ARN, et 2 ADN). Afin de caractériser finement le surétirement des molécules hybrides, nous avons donc souhaité réaliser des constructions bien définies, ne présentant qu'une extrémité libre permettant le peeling éventuel du brin d'ARN.

Dans la continuité d'une étude sur l'assemblage du ribosome entreprise dans le laboratoire, nous nous sommes intéressés à l'ARN ribosomique 23S d'*E. coli*. Nous avons utilisé la configuration de surétirement dans l'optique de générer un simple brin d'ARN libre de toute contrainte mécanique. Dans des conditions expérimentales bien choisies, le mode de peeling permet d'obtenir un tel simple brin depuis un point connu de la construction, et de contrôler sa taille par l'étirement imposé. Le brin ainsi libéré peut alors se structurer sans contrainte extérieure, puisque la force est supportée par le deuxième simple brin, dépourvu d'extrémité libre.

L'ARN ribosomique 23S établit une structure secondaire complexe, présentée figure 4.5. Les constructions imaginées sont des hybrides ARN/ADN. L'ADN est fonctionnalisé aux deux extrémités, et ne présente donc pas de bout libre. Le brin d'ARN n'est par contre fonctionnalisé que du côté 3', ce qui laisse son extrémité 5' libre. Ainsi, dans des conditions favorables au peeling, le brin d'ARN sera progressivement libéré par son extrémité 5', de façon similaire à ce qu'il se produit durant la transcription (Figure 4.6). Cette configuration permettrait d'étudier le repliement du ribosome à partir d'un ARN partiel, mimant la situation de repliement co-transcriptionnel. Nous serions donc dans des conditions plus proches de l'*in vivo* que l'assemblage à partir d'un ARN complet. Les attachements sont réalisés avec le couple biotine/streptavidine. Afin d'accroître la stabilité des points d'attachements, plusieurs biotines sont intégrées à chaque extrémité fonctionnalisée.

Dans le cadre de cette thèse, j'ai développé un protocole pour la synthèse des constructions moléculaires adaptées à cette étude. Afin de tester la faisabilité de l'étude envisagée, j'ai ensuite caractérisé le comportement en surétirement de ces constructions (en particulier la présence de peeling) dans différentes conditions expérimentales (différents tampons et plusieurs vitesses d'étirement).

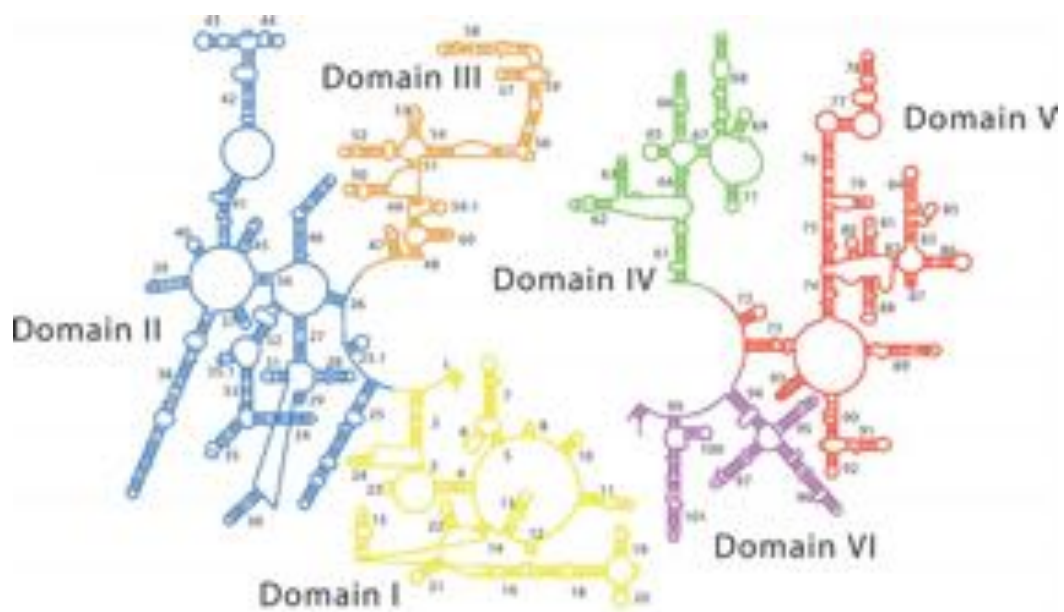


Figure 4.5 Schéma de la structure secondaire de l'ARN ribosomique 23S. Les six domaines (représentés ici de couleurs différentes) ont été définis arbitrairement afin de faciliter la lisibilité de la structure. Il sont composés de 600, 850, 400, 350, 500, et 260 nucléotides respectivement, dans l'ordre I → VI. Dans notre construction, l'extrémité libre correspond au début du domaine I. Il s'agit de l'extrémité 5'. Figure extraite de Kelin et al., *Journal of Molecular Biology*, 2004 [78]

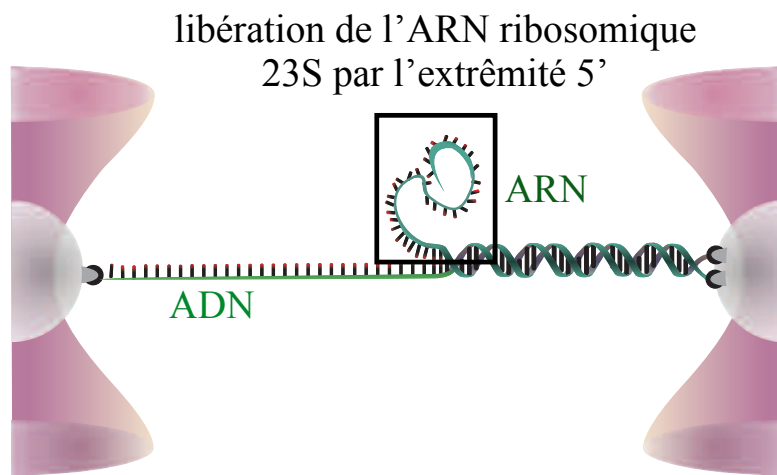


Figure 4.6 **Schéma de la configuration de peeling pour l'étude de structures secondaires** Une seule des extrémités de la construction est libre et permet la génération de simple brin. Pour les constructions de ce chapitre, il s'agit de l'extrémité 5' de l'ARN ribosomique 23S. Adapté avec la permission de Macmillan Publishers Ltd : Gross et al., *Nature Physics* ([www.nature.com/nphys/](http://www.nature.com/nphys/)), 2011 [23], copyright 2011

## 4.2.2 Protocole pour la réalisation des constructions moléculaires

Les constructions pour cette étude ont été réalisées à partir du plasmide PT7rrnB. Il contient la séquence complète des ARNs des deux sous-unités du ribosome d'*E.coli*, correspondant aux ARNs 16S, 23S, et 5S, successivement. Les différentes étapes du protocole sont représentées figure 4.7 et décrites dans cette partie.

### Protocole expérimental

1. Sélectionner par PCR la portion du plasmide contenant les séquences 23S et 5S. Introduire le promoteur de l'ARN polymérase T7 ainsi que le site de reconnaissance de l'enzyme de restrictions AflII par le primer « oligoF ». Introduire le site de reconnaissance de l'enzyme de restrictions FseI par le primer « oligoR ».
2. Transcrire *in vitro* la molécule obtenue précédemment. Purifier l'ARN, réserver.
3. Digérer le produit de PCR avec AflII. Traiter avec le fragment Klenow de l'ADN polymérase 1, en ajoutant les nucléotides suivants au mélange : biotine-dATP, dTTP, dCTP, dGTP. L'extrémité 3' du brin d'ADN est maintenant fonctionnalisée par 2 biotines.
4. Digérer le produit de l'étape précédente avec FseI.
5. Ajouter l'oligo DNAbiot, et incubé avec la ligase ADN T4. À ce stade, l'extrémité 5' du brin déjà fonctionnalisé en 3' présente elle aussi 3 biotines.
6. L'objectif de la prochaine étape est le remplacement du brin d'ADN non fonctionnalisé par le brin d'ARN obtenu à l'étape 2. Pour ce faire, incubé les produits des étapes 2 et 5 dans un solvant favorable aux hétéroduplexes, sous un cycle de température permettant dans un premier temps la déshybridation de tous les constituants, puis la réhybridation favorable des hétéroduplexes.
7. Digérer avec EcoRI afin de détruire les doubles brins d'ADN qui auraient pu se réhybrider.
8. Ajouter l'oligo RNAbiot et incubé avec l'ARN ligase T4-2.
9. La construction est alors directement utilisable. Conserver à -20 °C

## 4.2.3 Description d'une courbe de surétirement typique

### Paramètres mécaniques

Un exemple de courbe force-extension obtenue avec les constructions hybrides réalisées est présenté figure 4.8. Après une phase d'étirement de la molécule sous forme de double brin, un plateau de surétirement apparaît à une force proche de 50 pN, soit inférieure à celle observée pour l'ADN en conditions équivalentes (55-60 pN).

La phase d'étirement du double brin peut être modélisée par le modèle « twistable WLC » tenant compte du couplage entre l'étirement et la torsion, présenté précédemment. Comme dans le cas de l'ADN, nous avons considéré la fonction de couplage  $g(F)$  comme une constante à basse force, puis comme une fonction affine de la force [23].

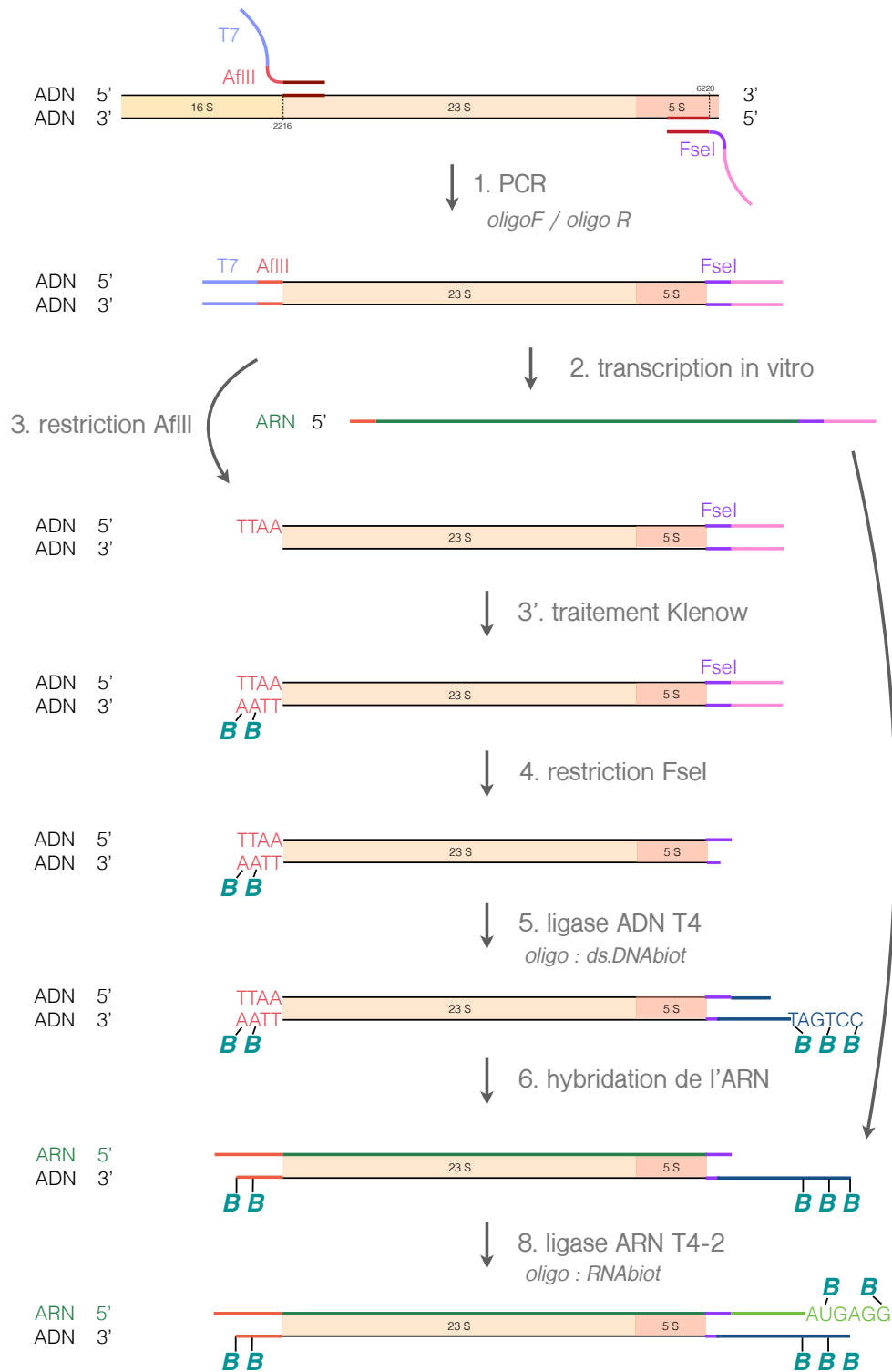


Figure 4.7 Schéma du protocole expérimental pour la fabrication des constructions hybrides ADN/ARN. Les différentes étapes du protocole sont détaillées ci-contre.



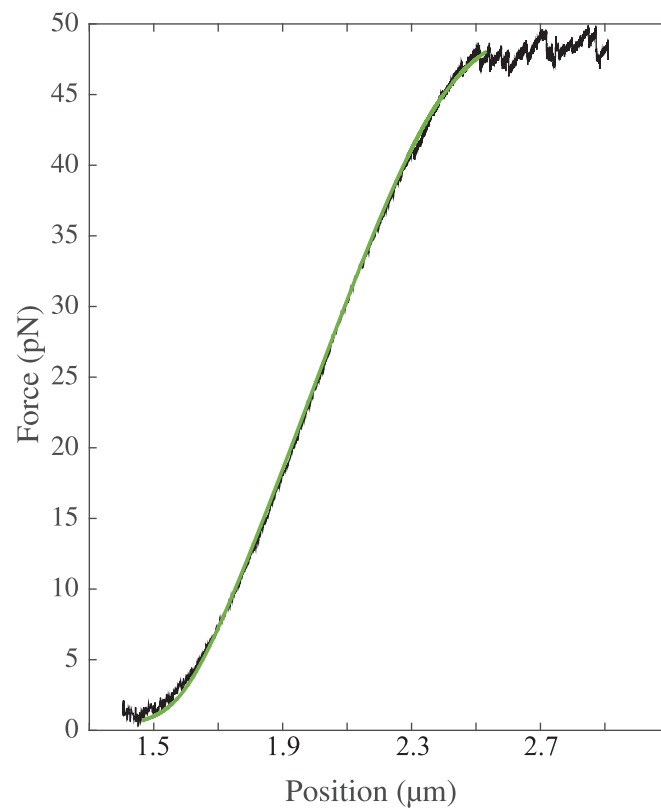


Figure 4.8 **Surétirement d'un double brin hybride ADN/ARN en tampon PBS 1X (courbe noire).** Ajustement avec le modèle du twistable worm (courbe verte). Les paramètres obtenus sont les suivants :  $F_c = 25$  pN ;  $g_0 = -525$  pN.nm ;  $g_1 = 22.7$  nm ;  $L_c = 1.175$  μm ;  $L_p = 55$  nm ;  $K = 1050$  pN

Les différents paramètres mécaniques de ce modèle correspondant aux ADN et ARN double brin, ainsi qu'aux hybrides, sont récapitulés dans le tableau 4.1. Ce tableau présente les valeurs obtenues par l'ajustement des courbes présentées figures 4.3 et 4.8, et également des valeurs issues de la littérature, à titre de comparaison [23, 44, 77]. La longueur cristallographique  $l_c$  (nm/nucléotide) obtenue pour la molécule hybride est similaire à celle d'un ARN double brin, plus faible que celle des ADN double brin. Ceci reflète le type d'hélice adopté par les différents acides nucléiques, A pour l'ARN et les hybrides, B pour l'ADN, l'hélice A étant en conditions équivalentes plus courte que l'hélice B. La valeur  $L_c = l_c \cdot n$  associée à une construction dépend évidemment aussi du nombre de nucléotides  $n$  qui la constitue.  $L_c$  influence la position à laquelle la réponse en force passe du régime entropique au régime élastique. Le module élastique  $K$  de l'hybride apparaît plus proche de celui des ADN double brin, avec une valeur nettement supérieure à celles rapportées pour l'ARN. Ce paramètre influence sur la pente de la courbe aux forces intermédiaires. Plus sa valeur est élevée, plus la pente est grande. La valeur obtenue pour la longueur de persistance  $L_p$  est intermédiaire à celles de l'ADN et de l'ARN. Son influence est sensible aux très basses forces ( $\sim 1$  pN), et a donc peu d'influence sur l'ajustement général des courbes présentées ici. Tous ces paramètres sont dépendants du tampon dans lequel est effectuée la mesure. Les résultats comparatifs entre les différents types d'acides nucléiques sont cependant conservés.

Le signe du couplage étirement-torsion des polyacides nucléiques n'est pas trivial. L'intuition suggère qu'une hélice se raccourcisse lorsque sa torsion est accrue (et inversement), ce qui correspond à un couplage positif. Cependant, l'allongement d'une hélice lorsque sa torsion augmente (couplage négatif) peut être compris en considérant le modèle suivant [75] : l'hélice est vue comme un fil rigide enroulé autour d'un coeur cylindrique déformable. L'allongement de l'hélice induit alors une diminution du diamètre du cylindre interne, et une augmentation de la torsion du fil.

Alors que l'ADN présente à basses forces un couplage négatif, il a récemment été montré par Lipfert et ses collaborateurs que le couplage de l'ARN double brin est positif [44] (Tableau 4.1). Cela pourrait être dû à la structure A de l'hélice d'ARN, différente de celle de l'ADN (hélice B). Cette hypothèse est soutenue par des résultats de simulations numériques comparant les réponses d'hélices d'ADN double brin sous forme A et B (à basses forces) : alors que la torsion de l'hélice B est amplifiée par l'étirement, celle de l'hélice A est atténuée [79]. Dans le cas de l'hélice A d'un hybride ADN/ARN, un couplage positif serait donc attendu. C'est le cas pour l'ajustement présenté figure 4.8, où pour  $F \leq F_c$ ,  $g = g_0 + g_1 \cdot F_c = +42.5$  pN.nm, valeur proche de celle indiquée pour l'ARN dans le travail de Lipfert *et al.* ( $g = 47.6 \pm 13.6$  pN.nm) [44]. Pour comparaison, l'ajustement de la courbe force-extension d'ADN de la figure 4.3 correspond à  $g = -47.9$  pN.nm. De façon générale, la force à laquelle le surétirement débute diminue lorsque  $g$  (donc  $g_0$  et / ou  $g_1$ ) augmente. C'est bien ce qui est observé ici, avec un début de surétirement autour de 45 pN pour la molécule hybride, valeur inférieure à celle de  $\sim 50$  pN obtenue pour l'ADN double brin (Figures 4.8 et 4.3).

Tableau 4.1 Comparaison des paramètres mécaniques des hélices de différents acides nucléiques

		$l_c$ (nm/nt)	$L_p$ (nm)	$K$ (pN)	$F_c$ (pN)	$g_0$ (pN.nm)	$g_1$ (nm)	$g_0 + g_1 \cdot F_c$ (pN.nm)
ADN	<sup>0</sup>	0.35	45	1150	25.5	-605	21.9	-47.9
	<sup>2</sup>	0.34	49	935				
	<sup>3</sup>	0.35	39	1500	30	-590	18	-50
hybride ARN	<sup>1</sup>	0.29	55	1050	25	-525	22.7	+42.5
	<sup>2</sup>	0.29	60	500				
	<sup>4</sup>	0.28	57	350				+47.6
erreur	<sup>0,1</sup>	$\pm 1\%$	$\pm 5\%$	$\pm 7.5\%$	$\pm 10\%$	$\pm 8\%$	$\pm 3\%$	$\pm 8\%$

<sup>0</sup> ce travail : valeurs issues de l'ajustement de la courbe présentée figure 4.3, en tampon 150 mM NaCl

<sup>1</sup> ce travail : valeurs issues de l'ajustement de la courbe présentée figure 4.8, en tampon 150 mM NaCl

<sup>2</sup> *Herrero-Galan et al.* [77] : valeurs obtenues par mesures de forces sur molécules uniques avec la technique de piégeage optique, en tampon 150 mM NaCl

<sup>3</sup> *Gross et al.* [23] : valeurs obtenues par mesures de forces sur molécules uniques avec la technique de piégeage optique, en tampon 50 mM NaCl.

<sup>4</sup> *Lipfert et al.* [44] : valeurs obtenues par mesures de forces sur molécules uniques avec la technique de piégeage magnétique, en tampon 100 mM NaCl

### Les différents modes de surétirement

Comme dans le cas de l'ADN, différents modes de surétirement sont observés sur nos molécules hybrides. Les premières expériences, réalisées en tampon peu salin et à 150 nm/s, ont révélé la présence d'un plateau en dents de scie, caractéristique du phénomène de peeling. Ce type de plateau est systématiquement associé à la présence d'une hystérésis très marquée (Figure 4.9). La réhybridation peut présenter des sauts dont l'amplitude dépasse 10 pN. Certains états intermédiaires sont parfois retrouvés lors de cycles successifs sur une même construction, mais le profil de réhybridation est globalement très variable. Cela est probablement lié à une forte structuration du brin d'ARN libéré, attendue pour la séquence du 23S (Figure 4.5). La formation de structures secondaires (voire tertiaires) d'une grande stabilité retarde fortement la reformation d'une hélice double brin. Comme présenté dans le chapitre 2, la dynamique associée aux structures secondaires d'ARN est relativement lente, et donne lieu à un phénomène d'hystérésis marqué même dans le cas des structures simples que sont les hairpins. Dans le cas du peeling de l'ARN, il arrive fréquemment que la réhybridation ne se produise pas dans la durée de l'expérience, même lorsque la force est ramenée à une valeur quasi nulle (Figure 4.10). Cela pourrait être associé au maintien de structures secondaires dynamiquement stables dans l'ARN simple brin libéré. Ce point sera discuté plus en détail par la suite.

Dans certaines conditions (détaillées par la suite), un plateau de surétirement lisse peut être observé. Cela peut concerner l'intégralité du plateau de surétirement, ou seulement certaines portions. Dans le cas où tout le plateau est lisse, une hystérésis peut, ou non, être présente lors de la relaxation de la molécule. Ces différents cas sont présentés figure 4.9. La présence d'un plateau lisse associé à une hystérésis, caractéristique d'une séparation des brins, peut s'interpréter comme la formation initiale de bulles fondues. La coalescence de ces bulles peut mener à la libération (partielle ou totale) de simple brin d'ARN, et donc à la présence d'une hystérésis. Un plateau lisse sans hystérésis correspondrait alors au cas où le brin d'ARN est resté solidaire du brin d'ADN, que ce soit en présence de bulles ou dans une forme conservant les liaisons hydrogène (équivalent de l'ADN S).

Le mode de surétirement a un effet notable sur la force à laquelle débute le plateau. Le phénomène de peeling débute à force plus basse (jusqu'à  $\sim 8$  pN) que celle à laquelle on observe un plateau lisse. Cela reste vrai sur toute la longueur du plateau. En conditions équivalentes, un plateau lisse marque la force limite que peuvent atteindre les dents de scie caractéristiques du peeling (Figure 4.15, insert). C'est aussi le cas lors du surétirement d'ADN double brin.

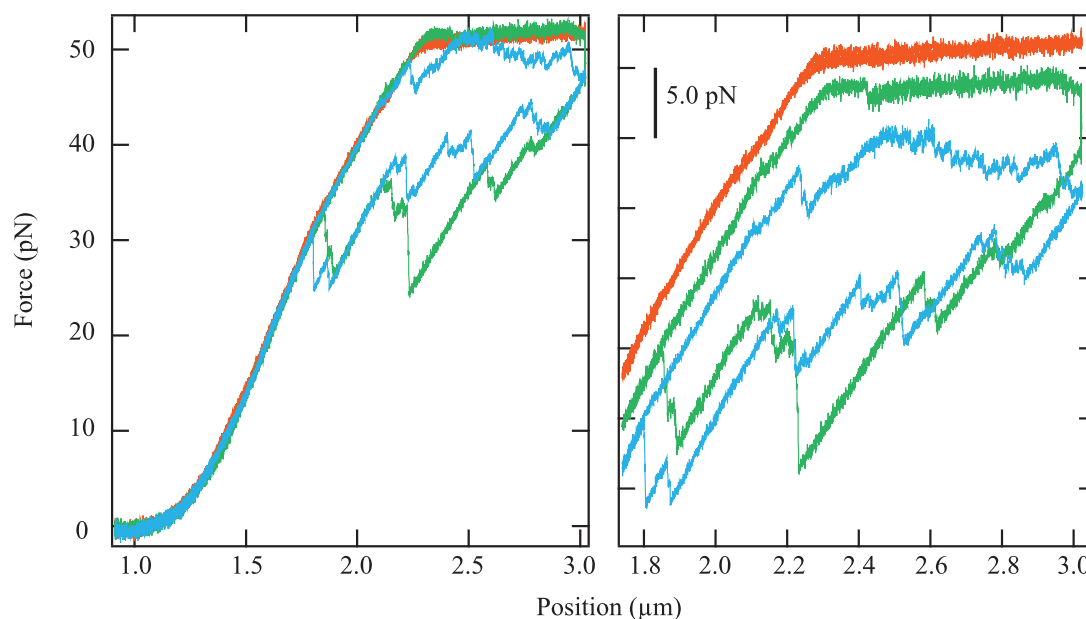


Figure 4.9 **Exemple des différents modes de surétirement observés pour les hybrides ARN/ADN.** Courbes de réponse en force des molécules hybrides ARN/ADN soumises à une rampe de déplacement (100 ou 150 nm/s), et agrandissement de la zone de surétirement (panneau de droite) (tampon expérimental : 150 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Hepes). Dans le panneau de droite, les différentes courbes ont été arbitrairement décalées en force. Trois modes de surétirement sont représentés. Le mode de peeling (courbe bleue) présente un plateau en dents de scie, et une forte hystérésis caractéristique de la libération d'un fragment simple brin. Le mode "mixte" (courbe verte), présente un plateau lisse et une hystérésis marquée. Cela peut s'interpréter comme l'apparition de "bulles" fondues, qui ont rejoint l'extrémité libre. Cela entraîne la libération de simple brin, et donc une hystérésis. Enfin, il arrive (rarement dans notre cas, observé sur moins de 5% des constructions étudiées) qu'un plateau lisse sans hystérésis soit observé (courbe orange). Il peut alors s'agir soit de bulles fondues restées localisées dans la molécule, soit d'un mode conservant les liaisons hydrogène.

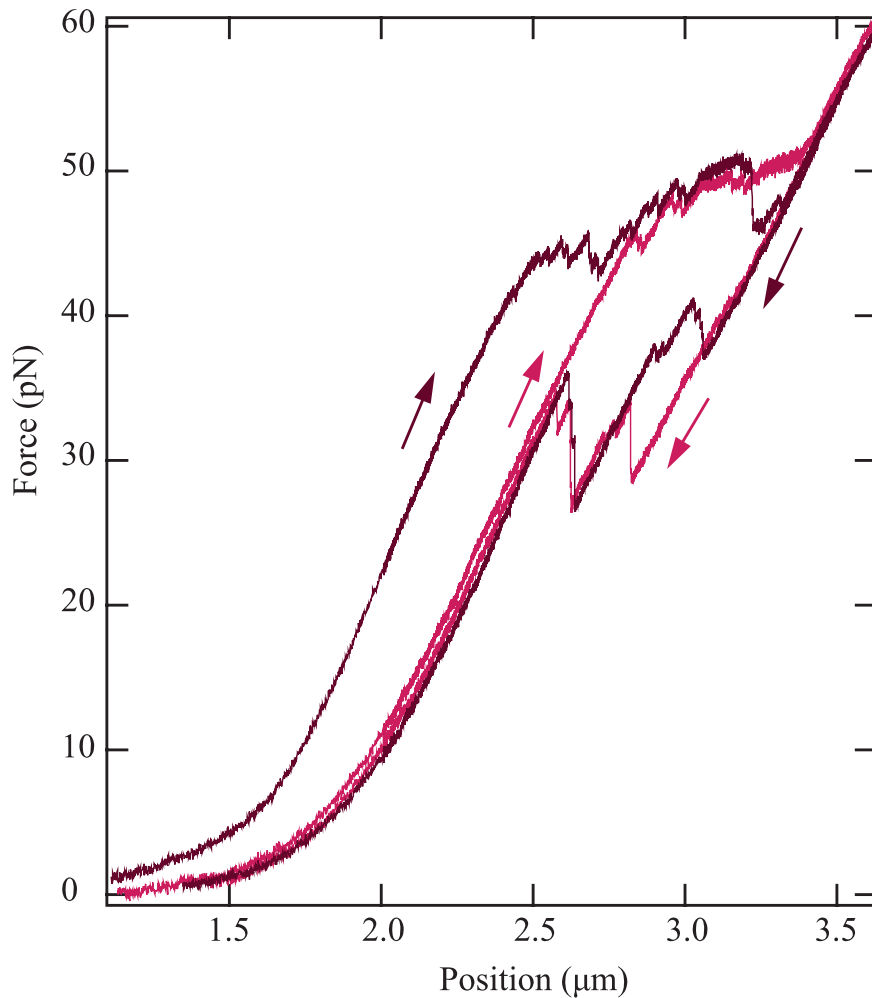


Figure 4.10 **Exemple de réhybridation partielle de l'hybride ARN/ADN.** La courbe marron représente le premier cycle étirement / relaxation de la molécule considérée, à 10 nm/s (tampon expérimental : 150 mM NaCl, 20 mM MgCl<sub>2</sub>, 10 mM Hepes). Un plateau mixte est observé : la majorité du plateau est un peeling, avec une zone de plateau lisse (3.1 - 3.3 µm). Lors de la relaxation de la construction, une hystérésis est observée. La courbe de relaxation ne rejoint pas celle d'étirement. Lors du cycle suivant (courbe rouge), la courbe à l'étirement se superpose à celle de relaxation du premier cycle, différente du premier étirement. Cela traduit le fait qu'à l'issue du premier cycle d'étirement / relaxation, la conformation initiale (totalement hybridée) n'a pas été retrouvée. Le second cycle présente donc la réponse d'une molécule mixte ADN simple brin / hybride.

#### 4.2.4 Analyse énergétique pour l'estimation de la force du plateau de peeling

La force du plateau de peeling peut être estimée en considérant un modèle énergétique relativement simple, inspiré de celui utilisé dans [23]. Le phénomène de peeling correspond à la conversion d'un polyacide nucléique double brin en deux simples brins, l'un seulement restant maintenu sous tension. Cette transition implique la rupture d'une liaison Watson-Crick, ainsi que l'étirement de la base simple brin libérée. La description de ce processus implique quatre termes énergétiques :

- Le premier terme,  $E_{ds}$ , correspond à l'énergie élastique par paire de bases du double brin non encore converti. L'élasticité du double brin est décrite par le modèle « twistable WLC », détaillé au début de ce chapitre.
- Le deuxième,  $E_{ss}$ , correspond à l'énergie élastique par nucléotide du simple brin sous tension déjà créé. Son élasticité est décrite par le modèle WLC simple, détaillé au chapitre 2.
- Le troisième terme,  $E_b$ , rend compte de l'énergie nécessaire à la rupture d'une liaison Watson-Crick. Sa valeur dépend du type d'acide nucléique, mais aussi de la paire de nucléotides considérée. Les valeurs correspondant aux paires A·T, A·U, et G·C, selon le type d'acide nucléique et selon la séquence locale (plus proche voisin), sont disponibles dans la littérature.
- Le quatrième terme est le travail correspondant à l'étirement de la base libérée,  $F\Delta L$ .  $\Delta L$  représente la différence de longueur entre une base de double brin et une base de simple brin à la force où se produit la transition.

À la transition,  $E_{ds} = E_{ss} + E_b - F\Delta L$ . Le calcul de ces différents termes énergétiques permet donc d'obtenir une estimation de la force à laquelle se produit le phénomène de peeling.

$E_{ds}$  est obtenue à partir de l'intégration de la solution proposée pour le modèle « twistable WLC » :

$$E_{ds}(F) = F \langle x \rangle_{ds}(F) - \int_0^F \langle x \rangle_{ds}(f) df \quad (4.2.1)$$

avec

$$\langle x \rangle_{ds}(F) = L_c^{ds} \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p^{ds}}} + \frac{C}{-g(F)^2 + K^{ds} C} F \right) \quad (4.2.2)$$

$L_c^{ds}$  et  $L_p^{ds}$  sont les longueurs cristallographique et de persistance relatives à l'acide nucléique sous forme double brin, à force nulle.  $K^{ds}$  est le module élastique de l'acide nucléique double brin.  $g(F)$  étant constante jusqu'à une force  $F_c$ , puis une fonction affine de  $F$ , l'expression de  $E_{ds}(F)$  est elle-aussi différente selon le domaine de forces considéré.

D'après les équations (4.2.1) et (4.2.2), les expressions suivantes sont obtenues :

Pour  $F \leq F_c$  :

$$E_{ds}(F) = \frac{1}{2}L_c^{ds} \sqrt{\frac{k_B T F}{L_p^{ds}}} + L_c^{ds} C \left[ \frac{F^2}{-(g_0 + g_1 F)^2 + K^{ds} C} - \frac{1}{2} \frac{F^2}{-(g_0 + g_1 F_c)^2 + K^{ds} C} \right]$$

Pour  $F > F_c$  :

$$E_{ds}(F) = E_{ds}(F_c) + \psi(F) - \psi(F_c)$$

où

$$\psi(F) = \frac{1}{2}L_c^{ds} \sqrt{\frac{k_B T F}{L_p^{ds}}} + L_c^{ds} C \left[ \frac{F^2}{-(g_0 + g_1 F)^2 + K^{ds} C} + \frac{1}{2g_1^2} \ln \left[ (g_0 + g_1 F)^2 - K^{ds} C \right] + \frac{g_0}{g_1^2 \sqrt{K^{ds} C}} \tanh^{-1} \left( \frac{g_0 + g_1 F}{\sqrt{K^{ds} C}} \right) \right]$$

$E_{ss}$  est obtenue à partir de l'intégration de la solution proposée pour le modèle WLC simple :

$$E_{ss}(F) = F \langle x \rangle_{ss}(F) - \int_0^F \langle x \rangle_{ss}(f) df \quad (4.2.3)$$

avec

$$\langle x \rangle_{ss}(F) = L_c^{ss} \left( 1 - 1/2 \sqrt{\frac{k_B T}{F L_p^{ss}} + \frac{F}{K^{ss}}} \right) \quad (4.2.4)$$

$L_c^{ss}$  et  $L_p^{ss}$  sont les longueurs cristallographique et de persistance relatives à l'acide nucléique sous forme simple brin, à force nulle.  $K^{ss}$  est le module élastique de l'acide nucléique simple brin. D'après les équations (4.2.3) et (4.2.4), l'expression obtenue pour  $E_{ss}$  est la suivante :

$$E_{ss}(F) = L_c^{ss} \left[ \frac{1}{2} \sqrt{\frac{k_B T F}{L_p^{ss}}} + \frac{F^2}{K^{ss}} \right]$$

Connaissant les paramètres mécaniques des différents acides nucléiques sous forme double et simple brin, la fonction énergétique  $E(F) = E_{ss}(F) + E_b - F \Delta L - E_{ds}(F)$  est parfaitement définie. Cette fonction est représentée figure 4.11 pour la molécule hybride étudiée, ainsi que pour les ADN et ARN double brin correspondants. Son annulation indique la conversion de double brin en simple brin, et donc la force moyenne du plateau de peeling.

Une analyse similaire peut être réalisée pour estimer la force de transition entre l'acide nucléique simple brin et le surétirement en mode de « bulles fondues ». La transition implique en ce cas la conversion d'une paire d'acides nucléiques sous forme double brin en deux simples brins sous tension. La contribution énergétique  $E_{ss}$  devient alors la somme des termes correspondant à chacun des deux simples brins sous tension. Dans le cas de l'hybride, où les deux simples brins générés sont de nature différente, le gain en longueur  $\Delta L$  doit aussi être adapté. Ces modifications sont détaillées en annexe. L'expression des autres termes est inchangée.



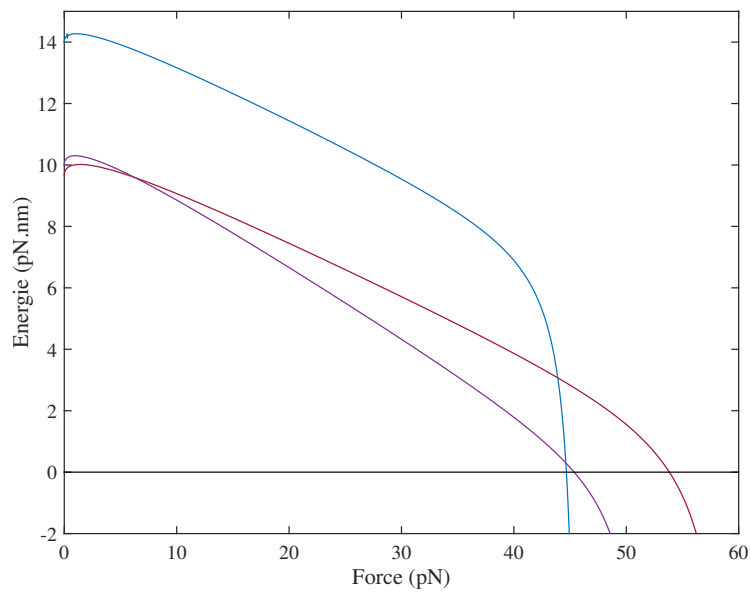


Figure 4.11 **Estimation théorique de la force du plateau de peeling.** Les courbes représentatives de la fonction énergétique  $E(F) = E_{ss}(F) + E_b - F\Delta L - E_{ds}(F)$  ont été représentées pour l'ADN double brin (courbe rouge), l'ARN double brin (courbe bleue), et l'hybride ARN/ADN (courbe violette). L'interaction de ces courbes avec la droite d'énergie nulle (noire) indique la force du plateau de peeling.

Tableau 4.2 Valeur des différents termes énergétiques à la force de transition vers le mode de peeling du surétirement.

	$E_{ss}$ (pN.nm/nt)	$E_{ds}$ (pN.nm/nt)	$E_b$ (pN.nm)	$F_{ds \rightarrow ss} \Delta L$ (pN.nm)	$\Delta L$ (nm)	$F_{ds \rightarrow ss}$ (pN)
ADN	6.1	2.3	9.66	13.4	0.25	53.7
hybride	5.3	1.3	10.0	13.6	0.30	45.3
ARN	4.6	7.6	14.0	11.2	0.25	44.7

Tableau 4.3 Valeur des différents termes énergétiques à la force de transition vers le mode de « bulles fondues » du surétirement.

	$E_{ss}$ (pN.nm/nt)	$E_{ds}$ (pN.nm/nt)	$E_b$ (pN.nm)	$F_{ds \rightarrow ss} \Delta L$ (pN.nm)	$\Delta L$ (nm)	$F_{ds \rightarrow ss}$ (pN)
ADN	7.6	4.3	9.66	13.0	0.23	56.5
hybride	6.7	7.3	10.0	9.1	0.18	50.5
ARN	5.8	8.3	14.0	11.2	0.25	44.8

Les paramètres mécaniques utilisés sont ceux rapportés dans le tableau 4.1. Il s'agit des valeurs obtenues dans ce travail pour l'hybride ARN/ADN et pour l'ADN, ainsi que pour les simples brins (chapitre 2). Pour l'ARN double brin, ce sont les valeurs issues de [77], avec la fonction  $g$  choisie identique à celle de l'hybride. Les énergies  $E_b$  associées à la rupture d'une liaison sont issues de la littérature [80, 81, 82] et rapportées dans les tableaux 4.2 et 4.3. Il s'agit de l'énergie moyenne obtenue pour les différentes paires et les différents voisins possibles. Cela revient à considérer une molécule contenant 50% de paires G·C, ce qui est proche du cas de nos constructions expérimentales (52% G·C). Les valeurs de  $\Delta L$  ont été obtenues à partir des courbes théoriques de l'élasticité de polyacides nucléiques double et simple brin de même longueur (modèle « twistable WLC » et WLC simple). La distance entre ces deux courbes a été mesurée à la force correspondant à la transition vers le surétirement. Les valeurs correspondantes sont rapportées dans les tableaux 4.2 et 4.3, et sont proches de celles obtenues à partir de courbes expérimentales.

Les valeurs obtenues pour la force de conversion  $F_{ds \rightarrow ss}^{peeling}$  correspondent quantitativement à celles mesurées expérimentalement, que ce soit pour l'ADN ( $F_{ds \rightarrow ss}^{peeling} \sim 54$  pN, à comparer avec la figure 4.3) ou pour l'hybride ( $F_{ds \rightarrow ss}^{peeling} \sim 45$  pN, à comparer avec les figures 4.8 et 4.15). En ce qui concerne l'ARN double brin, aucune courbe représentative de peeling n'a pour l'instant à ma connaissance été publiée. Il semble cependant cohérent d'observer le peeling à une force inférieure à celle de l'ADN double brin, comme rapporté pour un plateau

de surétirement lisse [77]. Étant donnée la similitude des hélices ARN double brin et hybride, il ne parait pas non plus surprenant d'observer le peeling à une force similaire dans ces deux cas.

L'analyse fine des différents termes énergétiques à la transition de peeling révèle cependant des différences notables entre l'ARN double brin et les autres acides nucléiques 4.2. En particulier, l'énergie  $E_{ds}$  associée au double brin d'ARN est bien plus élevée que celles des hélices contenant de l'ADN, alors que l'énergie  $E_{ss}$  associée au simple brin est plus faible. Cela induit une différence  $E_{ss}^{ARN} - E_{ds}^{ARN}$  négative, très favorable au simple brin ( $\sim -3$  pN.nm/nt), alors que celles associées à l'ADN et à l'hybride sont positives (3.8 et 4 pN.nm/nt respectivement). Cela compense une barrière énergétique  $E_b$  associée à la rupture d'une paire de base nettement plus élevée pour l'ARN (14 pN.nm pour l'ARN, et  $\sim 10$  pN.nm pour l'ADN et l'hybride).

En ce qui concerne l'ADN double brin et l'hybride, qui impliquent chacun une transition vers un simple brin d'ADN sous tension, la différence de force  $F_{ds \rightarrow ss}^{peeling}$  est à relier à la variation de longueur  $\Delta L$ . Lors de la conversion de double brin en simple brin, la transition « hybride  $\rightarrow$  ADN simple brin » implique un gain  $\Delta L = 0.30$  nm. Pour celle « ADN double brin  $\rightarrow$  ADN simple brin », le gain  $\Delta L$  n'est que de 0.25 nm. Pour un travail similaire, la force associée au peeling est donc inférieure dans le cas de l'hybride.

Quant à la transition vers le mode de « bulles fondues », les valeurs de  $F_{ds \rightarrow ss}^{melting}$  obtenues (tableau 4.3) sont toujours supérieures à celles de la transition de peeling. Cela correspond quantitativement à ce que nous mesurons lors du surétirement de molécules hybrides ( $F_{ds \rightarrow ss}^{peeling} = 45$  pN,  $F_{ds \rightarrow ss}^{melting} = 50.5$  pN). Dans les conditions où à la fois le mode de peeling et le mode lisse peuvent être observés, le plateau lisse a toujours lieu à une force supérieure d'environ 5 pN à celle de peeling. La différence est moins marquée pour l'ADN ( $F_{ds \rightarrow ss}^{peeling} = 53.7$  pN,  $F_{ds \rightarrow ss}^{melting} = 56.5$  pN), mais correspond toutefois aux observations expérimentales [73]. Dans le cas de l'ARN, la modélisation ne prévoit pas de différence entre les forces correspondant au peeling et aux bulles. Il n'existe cependant à ma connaissance pas de comparaison expérimentale publiée.

Cette analyse permet d'avoir une meilleure idée des influences relatives des composantes énergétiques de différents modes de surétirement, et leurs variations selon le type d'acide nucléique. Nos mesures sur les constructions hybrides ARN/ADN reflètent bien le comportement attendu par cette modélisation aussi bien pour le peeling que pour les « bulles fondues ». C'est sur cette dimension expérimentale que s'est concentré la suite de ce travail.

## 4.3 Caractérisation expérimentale du surétirement d'hybrides ARN/ADN

### 4.3.1 Effet de la vitesse sur le mode de surétirement observé

Aux faibles vitesses (quelques dizaines de nm/s), la sollicitation des hélices d'acides nucléiques est proche de l'équilibre. Ce n'est plus vrai lorsque la vitesse atteint plusieurs centaines de nm/s. Le mécanisme mis en jeu pour le surétirement pourrait donc être modifié par la vitesse à laquelle est réalisée l'expérience. Pour notre étude, des mesures ont été réalisées à des vitesses allant de 10 à 750 nm/s. La proportion de plateau lisse dans le plateau total augmente avec la vitesse d'étirement (Figure 4.12). Quel que soit le tampon expérimental, la probabilité d'observer une zone lisse à 10 nm/s est faible. Cette probabilité, ainsi que la proportion de plateau concerné, augmente avec la vitesse d'étirement. De façon contre-intuitive, l'amplitude de l'hystérésis (lorsque celle-ci est présente) ne dépend pas notablement de la vitesse. Deux diagrammes reprenant conjointement les effets de la vitesse et de la salinité du tampon (détaillés au paragraphe suivant) sont présentés figure 4.16.

### 4.3.2 Effet du tampon expérimental sur le surétirement

La composition du tampon expérimental influe sur la stabilité des doubles hélices d'acides nucléiques. Cet effet peut être stabilisateur (sels monovalents et divalents [24, 25, 83]), ou au contraire déstabilisateur (formamide [84, 85], figure 4.13). Une influence du tampon expérimental sur la force, et éventuellement le mécanisme de surétirement, est donc attendue. Afin de quantifier cet effet, des expériences ont été réalisées dans divers tampons. Deux tampons « classiques » utilisés en biologie moléculaire, le tampon phosphate (PB, pH 7.4) et le tampon phosphate salin (PBS, 150 mM NaCl); deux contenant de fortes concentrations en sel monovalent (1 M NaCl, 2 M NaCl); 3 contenant diverses concentrations en sel divalent (150 mM NaCl + 10 mM / 20 mM / 50 mM MgCl<sub>2</sub>); 3 contenant à la fois des sels divalents et de fortes concentrations en sel monovalent (10 mM MgCl<sub>2</sub> + 500 mM / 1M NaCl; 50 mM MgCl<sub>2</sub> + 1 M NaCl); et 4 contenant du formamide (150 mM NaCl + 1% / 2.5% / 5% / 10% en volume). Tous contiennent aussi 10 mM Hepes, pH 7.8. À titre de comparaison, l'eau de mer contient typiquement 0.5-0.7 M d'ions monovalents et 50 mM de divalents.

#### **Influence sur la force du début de peeling**

Des courbes de surétirement réalisées dans différents tampons sont présentées figure 4.14. La force du plateau de surétirement a été mesurée à 10 nm/s dans chacun des tampons présentés (Figures 4.15). Il s'agit systématiquement de plateau de peeling. Les forces représentées figure 4.15 correspondent à une moyenne sur toute la longueur du plateau, et sont issues de mesures sur 5 à 15 molécules selon les cas.

Une augmentation de la concentration en MgCl<sub>2</sub> de 0 à 50 mM induit une augmentation de 10% de la force où se produit le peeling, soit un décalage de 5 pN environ. L'effet des

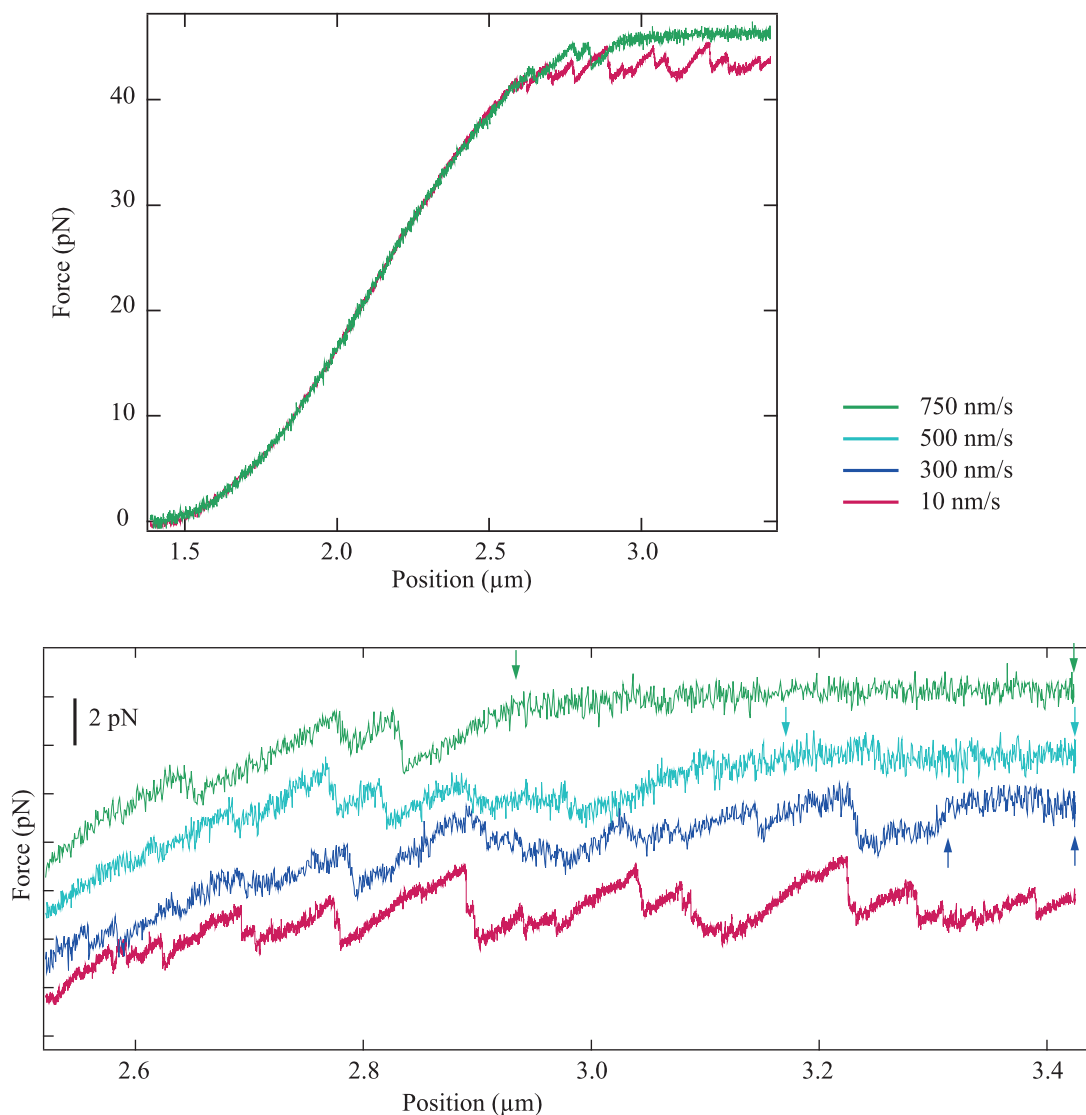


Figure 4.12 **Effet de la vitesse sur le surétirement d'hybrides ARN/ADN.** La vitesse d'étirement a une forte influence sur le mode de surétirement observé. Les courbes représentées ici correspondent à quatre extensions successives d'une même molécule (tampon expérimental : 150 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Hepes). Les deux vitesses extrêmes (10 et 750 nm/s) sont représentées dans le panneau supérieur. Un agrandissement sur la zone du plateau est présenté dans le panneau inférieur, pour toutes les vitesses. Dans ce panneau, les courbes sont décalées en force de 2, 4 et 6 pN par rapport à l'enregistrement à 10 nm/s. A 10 nm/s (courbe rouge), un plateau de peeling avec des dents de scie très marquées est observé. Lorsque la vitesse est augmentée à 300, 500, puis 750 nm/s (courbes respectivement bleue, turquoise, et verte), des zones de plateau lisse apparaissent. Celles-ci sont indiquées par des flèches sur les courbes du panneau inférieur. A 750 nm/s, la majeure partie du plateau est lisse. Lors de plusieurs enregistrements à une même vitesse sur une même molécule, la position et la longueur des zones lisses peuvent varier notablement.

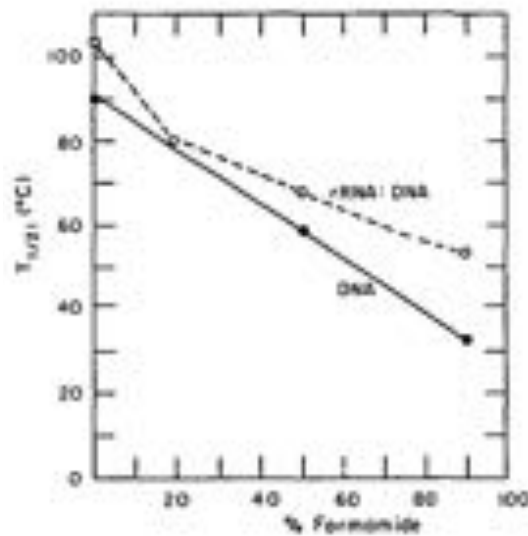


Figure 4.13 **Température de dénaturation d'hybrides ARN/ADN.** Il s'agit de la dénaturation thermique d'un fragment d'ARN ribosomique d'*E. coli* et des ADN double brins correspondants, en fonction de la proportion (en volume) de formamide dans le tampon. Pour cette figure, le tampon de référence est 0.3 M NaCl, 0.01 M Tris, pH 7.8. Figure issue de Casey et al., *Nucl. Acids. Res.*, 1977 [84], avec la permission de Oxford University Press.

sels monovalents est moins marqué. En revanche, l'ajout de formamide a un effet fort sur le peeling. Ainsi, à 10% de formamide en volume, la force du plateau de peeling baisse de 40% par rapport aux tampons qui en sont dépourvus (référence choisie à 45 pN, force caractéristique observée en tampon PB et à faible concentration en sels divalents). L'effet du formamide sur la température de dénaturation des acides nucléiques est étudié de longue date [84]. Dans le cas de l'ADN, un mécanisme expliquant l'affaiblissement des hélices double brin est proposé [85]. Le formamide étant un solvant protique, il tend à remplacer les molécules d'eau formant des liaisons hydrogènes avec acides nucléiques hydratés [85]. Cela induit une modification du potentiel ionique des hélices, tendant à les déstabiliser. Les hélices hybrides sont moins affectées que les ADN double brins par le formamide (Figure 4.13), mais ce mécanisme reste valable pour expliquer leur déstabilisation.

Bien que développé dans le cadre d'une dénaturation thermique, ce raisonnement reste pertinent dans notre cas de déstabilisation sous force. Il est toutefois difficile de comparer de façon quantitative ces deux types d'expériences tant les mécanismes impliqués sont différents. En particulier, les expériences sous force induisent une tension au sein de la molécule (longitudinale pour la configuration de surétirement, transversale dans le cas des hairpins du chapitre précédent), alors que la dénaturation thermique repose sur des fluctuations déstabilisant les paires de bases. Après dénaturation, la molécule déstabilisée thermiquement peut adopter une conformation de pelote statistique, alors que la tension générée par la présence d'une force reste appliquée au simple brin formé (au moins à l'un des brins en configuration de surétirement).

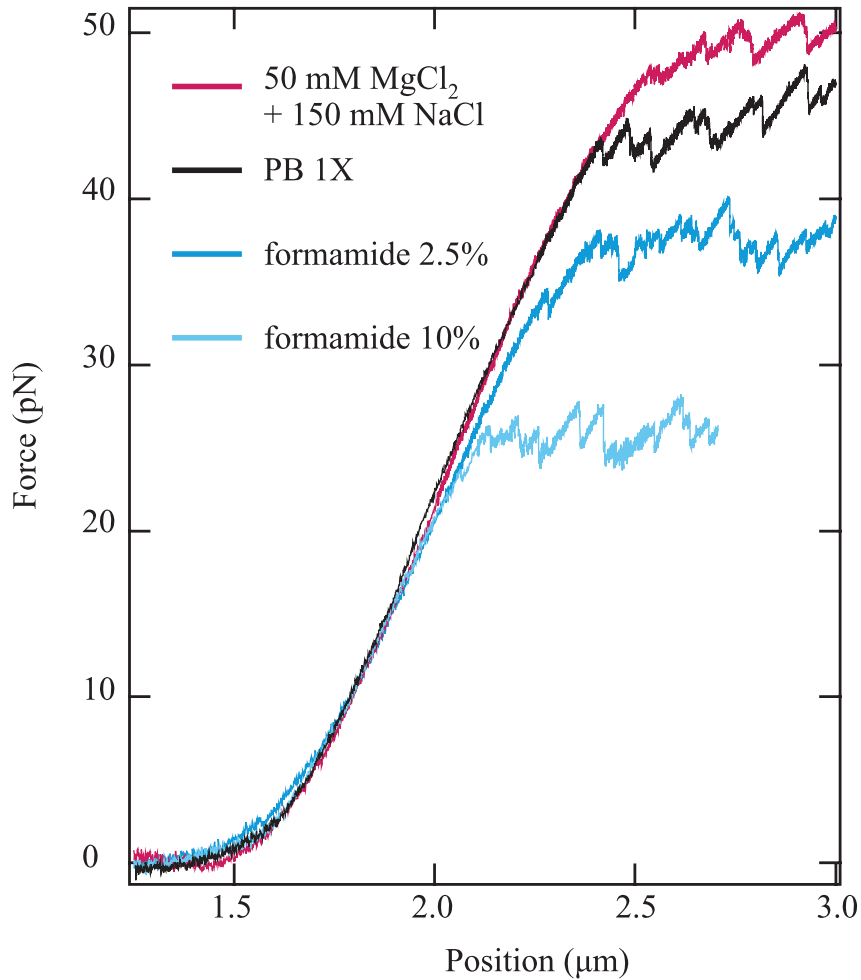


Figure 4.14 Exemples de courbes de réponse en force lors de l'étirement d'hybrides ARN/ADN dans différents tampons expérimentaux. Pour toutes les courbes, la vitesse d'étirement est de 10 nm/s. La force du plateau de peeling varie notablement avec le tampon expérimental utilisé. En considérant le tampon PB1X comme référence (courbe noire), une augmentation en force de 10% est observée en tampon divalent (50 mM MgCl<sub>2</sub> + 150 mM NaCl, courbe rouge). A l'inverse, la force diminue de 40% lorsque le tampon contient 10% de formamide (courbe bleu clair).

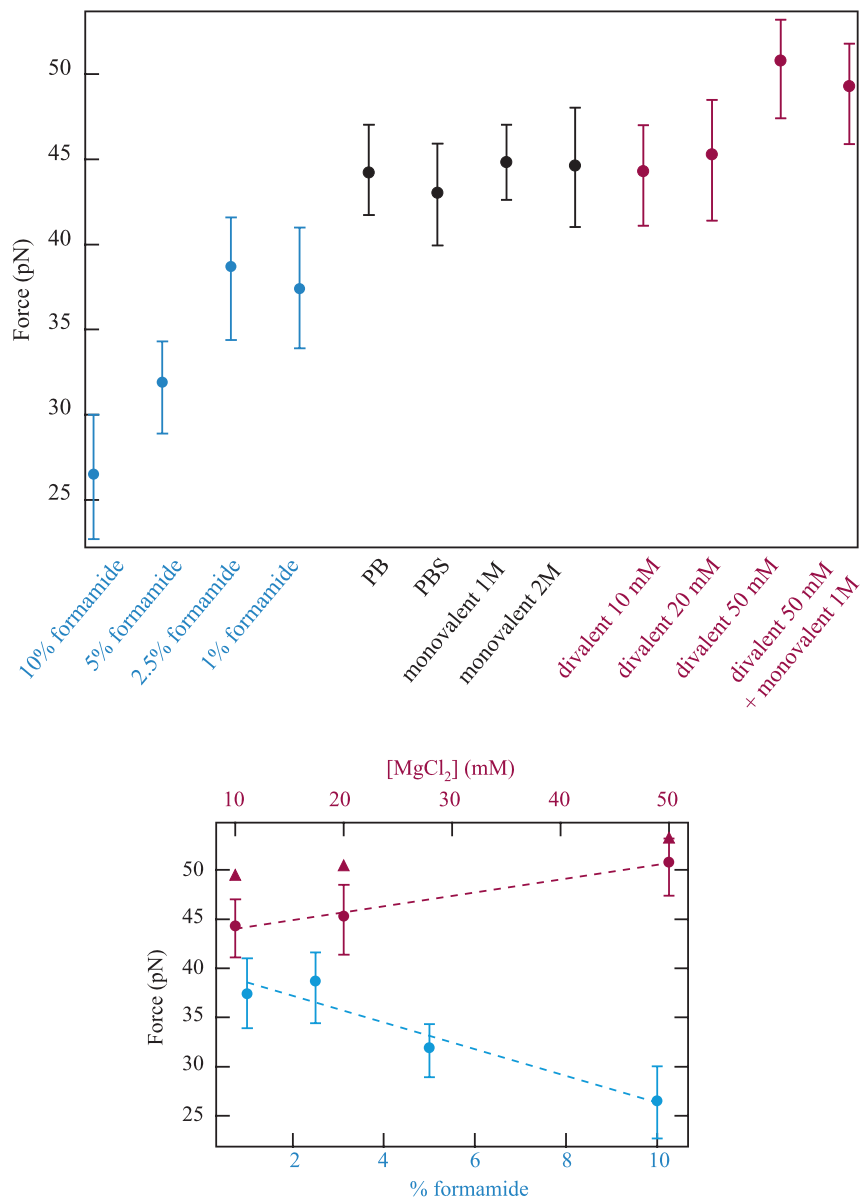


Figure 4.15 **Force du plateau de peeling mesurée pour un étirement à 10 nm/s dans différents tampons expérimentaux.** La valeur présentée (●) est la force moyenne sur l'intégralité du plateau de surétirement. Les barres d'erreur correspondent à la moyenne des maxima et des minima des dents de scie du plateau. Panneau supérieur : la force du plateau ne dépend pas de la concentration en sels monovalents du tampon expérimental (symboles noirs). Elle diminue par contre fortement lorsque du formamide est ajouté. A l'inverse, cette force augmente avec la concentration en sels divalents. Le panneau inférieur représente d'une part la variation de la force du plateau en fonction de la proportion (en volume) de formamide dans le tampon expérimental (courbe bleu, ajustement linéaire), et d'autre part cette variation en fonction de la concentration en sels divalents (MgCl<sub>2</sub>) dans le tampon (courbe rouge, ajustement linéaire). Les triangles (▲) représentent la force mesurée sur un plateau lisse dans le même tampon.



Comme représenté figure 4.15, l'amplitude du plateau de peeling (différence de force entre les extrémités supérieure et inférieure des dents de scie) atteint 5-8 pN, et ce quel que soit le tampon expérimental utilisé.

### **Influence sur le mode de surétirement**

La proportion de plateau lisse par rapport au plateau en dents de scie a été mesurée dans des tampons de concentration en sels monovalent et divalent variables. Les résultats sont présentés figure 4.16.

Dans les différents tampons dépourvus de sels divalents utilisés, aucun mode autre que le peeling n'a été observé à une vitesse d'étirement de 10 nm/s. Dans les tampons contenant  $MgCl_2$ , des zones de plateau lisse sont parfois visibles dès cette vitesse. Plus la concentration en sel bivalent est élevée, plus la fréquence et la longueur des zones lisses est grande. Pour une concentration de 50 mM en  $MgCl_2$ , le plateau lisse devient majoritaire dès 100 nm/s.

Pour des vitesses d'étirement supérieures à 10 nm/s, un plateau lisse partiel est parfois observé dans des tampons ne contenant que des sels monovalents. Il faut cependant atteindre 1M de NaCl, avec un étirement à 750 nm/s pour que le plateau lisse devienne majoritaire. À titre de comparaison, à une concentration en sels divalents de 50 mM, le plateau lisse est majoritaire dès 100 nm/s (Figure 4.16). La conjugaison de sels divalents et monovalents n'a pas une grande influence sur le mode de surétirement : pour une concentration de 10 mM en  $MgCl_2$ , la proportion de plateau lisse ne varie pas avec la concentration de NaCl (dans la gamme de 150 mM à 1M de NaCl). Les données présentées figure 4.16 proviennent de mesures sur différentes molécules, chaque molécule étant soumise à plusieurs cycles d'étirement et de relaxation.

### **4.3.3 Discussion**

L'apparition d'un plateau lisse a une importante composante stochastique : sur différents cycles d'une même molécule, dans le même tampon et à la même vitesse, une fraction de plateau lisse peut être présente ou non, de longueur et/ou de position variable, et engendrer ou non une hystérésis à la relaxation. Grâce aux deux paramètres (vitesse d'étirement,  $[MgCl_2]$ ), il est néanmoins possible de sélectionner des conditions où le peeling est dominant (basses vitesses, faible concentration en sels divalents), ou au contraire peu fréquent (hautes vitesses, hautes concentrations en sels divalents).

La stabilisation des doubles hélices d'acides nucléiques (ADN et ARN) par les cations provient d'un effet électrostatique. Les brins d'ADN et d'ARN étant des chaînes polymères chargées négativement, les cations (monovalents ou divalents) présents dans la solution viennent s'y agréger de façon non spécifique. L'épaisseur sur laquelle les cations s'accumulent, appelée longueur de Debye, est contrôlée par un équilibre entre l'attraction électrostatique et un contre-effet entropique. Ce dernier tend à limiter l'accumulation des cations dans une zone

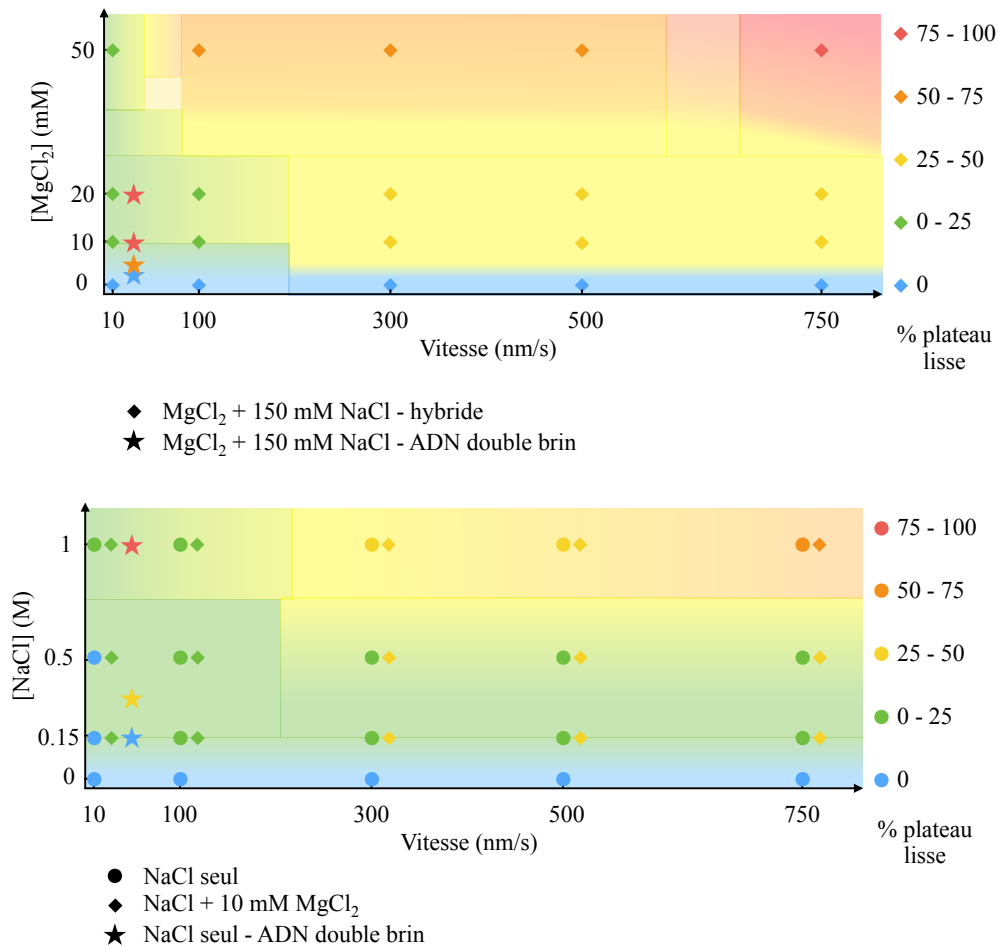


Figure 4.16 **Nature du plateau de surétirement (lisse ou en dents de scie) en fonction de la vitesse et de la concentration en sels divalents.** La proportion de plateau lisse observé est représentée par un code couleur. Chaque symbole représente une condition (vitesse/sels) où une mesure a été réalisée. Chaque point est une moyenne sur plusieurs étirements (5 à 25) et plusieurs molécules.

*Panneau supérieur :* influence de la concentration en sels divalents ( $MgCl_2$ ). Dans tous les cas (sauf la condition sans sels), le tampon expérimental contient aussi 150 mM de sels monovalents ( $NaCl$ ). La vitesse d'étirement et la teneur en  $MgCl_2$  du tampon expérimental influencent fortement la transition d'un plateau de peeling en dents de scie à un plateau lisse. En tampon dépourvu de sels (PB 1X), le régime de peeling est observé quelle que soit la vitesse. De façon similaire, la proportion de plateau lisse reste inférieure à 25% quelle que soit la concentration en  $MgCl_2$  lorsque la vitesse d'étirement est de 10 nm/s. Augmenter la vitesse d'étirement et/ou la concentration en sels divalents induit un passage du régime de peeling à un régime où le plateau lisse devient majoritaire.

*Panneau inférieur :* influence de la concentration en sels monovalents ( $NaCl$ ). A vitesse fixée, l'effet des sels monovalents est moins marqué que celui des sels divalents. Même avec une très forte concentration en  $NaCl$  (1M), le régime de plateau lisse ne devient majoritaire qu'à la vitesse de 750 nm/s. La conjugaison de sels monovalents et divalents (◆, 10 mM  $MgCl_2$  +  $NaCl$ ) n'a pas d'effet spectaculaire : à toutes les concentrations en  $NaCl$ , l'effet observé est le même qu'avec 150 mM de  $NaCl$  et la même concentration (10 mM) en  $MgCl_2$ .

À titre de comparaison, les étoiles (★) représentent le comportement de l'ADN en conditions identiques [73].

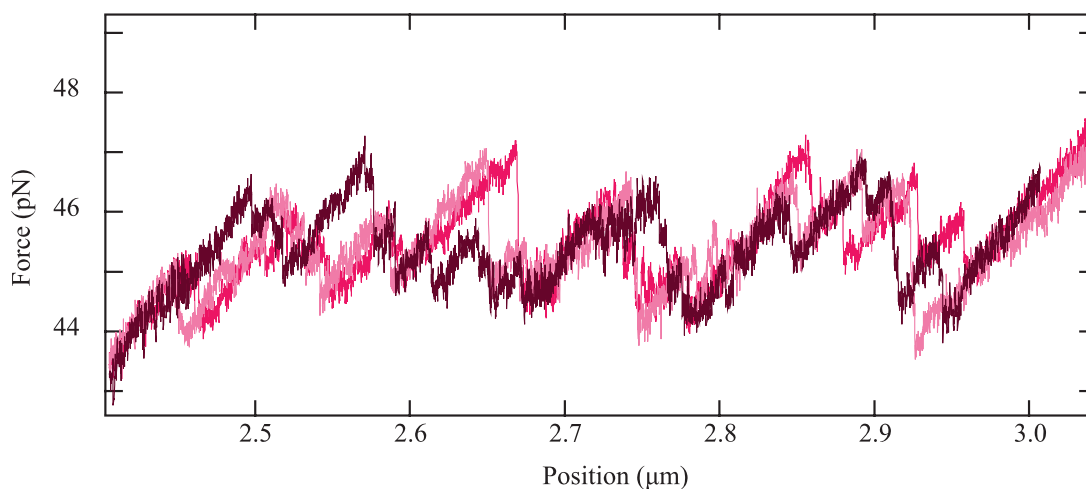


Figure 4.17 **Profil du plateau de peeling lors de plusieurs étirements successifs d'une même molécule.** Les trois étirements ont été réalisés à 10 nm/s (tampon expérimental : 150 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM Hepes). Certains motifs sont retrouvés sur les différents enregistrements. Le signal d'ouverture est relié au contenu local en G·C de la molécule, et donc à sa séquence nucléotidique.

localisée de l'espace. L'effet global est un écrantage des charges des polyacides nucléiques, sensible aux distances supérieures à quelques longueurs de Debye. Cela favorise le rapprochement des brins jusqu'à une distance où les liaisons H peuvent s'établir, et donc la formation d'hélices. Plus la salinité du tampon est grande, plus l'écrantage est fort à courtes distances, et donc plus deux simples brins chargés peuvent s'approcher l'un de l'autre. Ce raisonnement est aussi valable pour la formation d'hélices hybrides ARN/ADN. Une stabilité accrue des hélices est favorable aux modes de surétirement alternatifs au peeling. Elle permet en effet la stabilisation de bulles localisées par le maintien de zones double brin, ou même la conservation des liaisons H dans une forme déroulée de l'hélice.

La stabilité des hélices étant localement influencée par leur séquence nucléotidique, un effet de cette séquence sur le mode de surétirement a été rapporté pour l'ADN [23] et existe probablement aussi dans le cas de molécules hybrides. Une analyse fine du lien entre le mode de surétirement, plus particulièrement le passage du peeling à un plateau lisse, et la séquence de notre construction devra être réalisée afin de caractériser cet effet. Dans le cas d'unpeeling, le profil des dents de scie dépend lui aussi de la séquence. Dans le cas de l'ADN, ce profil peut être correctement modélisé à partir d'une analyse énergétique [23]. Cela devrait pouvoir être transposé dans le cas de nos molécules hybrides. Expérimentalement, nous avons observé que lors de peeling successifs d'une même construction, certains motifs des dents de scie sont retrouvés d'un enregistrement à l'autre (figure 4.17).

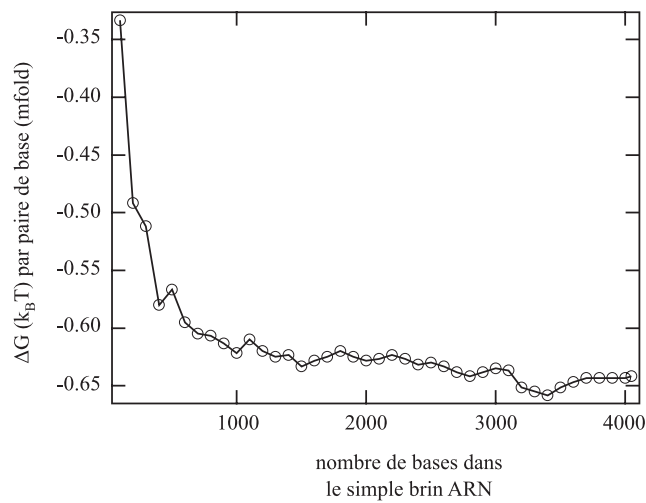


Figure 4.18 **Énergie des structures secondaires possibles dans le simple brin ARN généré par peeling.** La barrière énergétique à franchir pour reformer le double brin hybride ARN/ADN a été calculée pour différentes longueurs de simple brin libéré (mfold). Elle correspond à l'énergie stockée dans la structure secondaire la plus stable potentiellement formée, distribuée sur le nombre de bases à réhybrider.

De façon similaire, les différents états intermédiaires observés lors de la relaxation des constructions devraient pouvoir être associés à la réhybridation de fragments de longueur définie. Une telle analyse est cependant difficile à réaliser précisément, en particulier à cause du très grand nombre d'états intermédiaires observés expérimentalement, souvent proches les uns des autres. De plus, une analyse de l'énergie stockée dans les structures secondaires potentiellement formées dans le brin ARN généré par peeling (calculée avec mfold) révèle une stabilité assez homogène quelque soit la taille du brin libre, dès que celle-ci dépasse quelques centaines de bases (Figure 4.18). À partir d'environ 600 bases libérées, la barrière énergétique à franchir pour réhybrider le simple brin est supérieure à  $0.6 k_B T$  par paire de bases (en valeur absolue), soit environ deux fois plus que pour l'ADN étudié dans les articles utilisés pour comparaisons [23, 73]. La stabilité des structures augmente globalement lorsque le nombre de bases dans le simple brin augmente, avec des fluctuations inférieures à  $0.05 k_B T$  par paire de bases. À l'inverse, le profil calculé pour l'ADN présente des minimums locaux où la stabilité est deux fois supérieure à sa valeur moyenne [23]. Ces puits énergétiques ont été associés à des zones de retard à la réhybridation. Le paysage énergétique globalement lisse observé pour l'ARN ne permet pas une telle attribution. L'analyse du lien entre le profil de peeling et la séquence de la construction, mentionnée au paragraphe précédent, devrait permettre un contrôle fin de la longueur du brin libéré. Cela rendra donc possible une meilleure exploration expérimentale de la réhybridation, et donc une meilleure compréhension de son signal complexe.

Dans nos expériences, la relaxation se fait en grande majorité en présence d'une hysteresis après un plateau lisse, même à haute force ionique (150 mM NaCl + 20 mM MgCl<sub>2</sub>). Cela indique une séparation des brins, au moins partielle. À l'inverse, dans ces conditions, il a été

montré que le mode de surétirement prédominant pour l'ADN double brin est la forme S [73]. Cette forme implique la conservation des liaisons hydrogènes, et donc la non séparation des brins. L'influence de la salinité du tampon sur le surétirement est donc très différente selon le type d'acides nucléiques considéré.

De plus, en comparaison du cas de l'ADN double brin, le mode de peeling apparaît plus fréquemment pour la molécule hybride. Pour un ADN double brin, même à 10 nm/s, l'ajout de sels monovalents suffit à empêcher partiellement (250 mM NaCl) ou totalement (1 M NaCl) le peeling en faveur d'un plateau lisse [73]. L'effet des sels bivalents est lui aussi bien plus marqué : dès l'ajout de 5 mM de  $MgCl_2$ , un plateau lisse apparaît, et 20 mM suppriment complètement le peeling [73] (pour un tampon contenant aussi 150 mM NaCl). Le mode de surétirement impliquant une libération de simple brin apparaît donc bien plus robuste dans le cas de molécules hybrides (Figure 4.16). Cette observation peut être reliée à l'analyse énergétique développée dans la section précédente de ce chapitre (tableaux 4.2 et 4.3). En effet, dans le cas de l'ADN double brin, les modes de peeling et de « bulles fondues » se produisent à des forces plus proches (différence de 3 pN environ) que dans le cas de l'hybride ( $\sim 5$  pN). La proximité de ces deux états dans le cas de l'ADN favorise la transition de l'un à l'autre dans des conditions expérimentales proches. D'un point de vue énergétique, les deux types de transition apparaissent assez similaires pour l'ADN, avec une différence  $E_{ss}^{DNA} - E_{ds}^{DNA}$  proche de 3.5 pN.nm/nt. De même, le travail associé à l'étirement du/des simples brins générés varie peu. À l'inverse, dans le cas de l'hybride, les contributions énergétiques de la transition par « bulles fondues » sont très différentes du cas du peeling. La différence  $E_{ss}^{hyb} - E_{ds}^{hyb}$ , qui dans le cas du peeling est similaire à celle associée à l'ADN, devient légèrement négative dans le cas de la formation de bulles, et diminue donc le coût de la conversion du double au simple brin. Le gain en longueur  $\Delta L$  diminue fortement, devenant plus faible que dans le cas de l'ADN. Les transition de surétirement par peeling et par fonte de bulles sont donc très différentes dans le cas de l'hybride ARN/ADN. Cela peut expliquer la nécessité de modifications fortes des conditions expérimentales pour passer de l'une à l'autre.

Comme mentionné précédemment, il a été montré que pour les constructions ADN, le contenu en G-C a une influence sur le surétirement : en particulier, des zones stabilisées par un fort contenu en G-C sont défavorables au mode de peeling [23]. Il faut donc être prudent lors de la comparaison de résultats obtenus sur des constructions de séquences différentes. Nos constructions hybrides possèdent un taux moyen de 52% de G-C. Sur une fenêtre de 30 paires de bases, ce taux fluctue entre 35 et 75%. Les constructions ADN utilisées pour les comparaisons [23, 73] présentent un taux de G-C de 50%, avec des fluctuations entre 40 et 70% (sur une fenêtre identique). Bien que différentes, ces deux séquences ont donc un profil de stabilité similaire, ce qui justifie leur comparaison.

## 4.4 Perspectives

À la lumière des résultats obtenus dans cette partie, il semble tout à fait envisageable d'étudier la structuration de l'ARN ribosomique en configuration de surétirement. Une large gamme de tampons expérimentaux, à une vitesse bien choisie, permet d'obtenir le peeling du brin d'ARN. En particulier, le tampon recommandé pour l'assemblage du ribosome *in vitro* (4 mM Mg<sup>2+</sup> ; 400 mM K<sup>+</sup>) [86] en fait partie. Une modélisation du profil des dents de scie, comme réalisée pour le surétirement de l'ADN [23], est prévue. Elle devrait permettre de relier précisément le déplacement imposé à la position de la fourche de séparation des brins. Cela rendra possible une libération contrôlée d'une longueur choisie d'ARN simple brin (Figure 4.17).

La seule limitation prévisible est liée à la réhybridation du simple brin libéré. Dans certains cas, même à très faible vitesse, seule une réhybridation partielle a été observée (Figure 4.10). Cela ferait obstacle à une étude de la stabilité des structures secondaires et tertiaires formées. Cela arrive cependant principalement lorsqu'une grande partie du simple brin a été libérée. La majeure partie de l'ARN ribosomique 23S (qui constitue les trois premiers quarts du brin libéré) reste donc accessible.

Un autre axe d'étude consiste à réaliser une comparaison soignée du surétirement de notre molécule hybride avec son équivalent ADN double brin. La construction d'ADN correspondante, comportant trois extrémités fixées à des billes, devrait pouvoir être obtenue sans difficulté majeure. Cela permettrait de s'affranchir des effets de séquence. Il sera alors possible d'isoler la contribution de la libération d'ARN simple brin par rapport à de l'ADN, que ce soit lors du peeling ou de la réhybridation.



# Conclusion

*Alles hat eine Ende, nur die Wurst hat  
zwei!*

---

Stephan Remmler

Ce travail m'a permis d'explorer différents niveaux de structuration des molécules d'ARN simple brin : la forme la plus simple, le hairpin, système modèle pour l'étude des autres structures ; puis aux échelles intermédiaires, un fragment du coeur catalytique de l'ARN ribosomique, et son interaction avec une protéine hélicase ; enfin l'ARN ribosomique 23S dans son intégralité, support de l'étude du surétirement de molécules hybrides. La grande diversité des structures accessibles aux molécules d'ARN mérite d'être mise en valeur, puisque sa grande versatilité lui permet d'assumer aussi bien des fonctions catalytiques que la conservation et la transmission d'informations génétiques.

*In vitro*, ces deux types de fonctions sont aussi accessibles à l'ADN, mais son rôle dans les organismes vivants se cantonne aujourd'hui à la conservation du patrimoine génétique. Aucune fonction catalytique impliquant l'ADN n'a pour l'instant été observée *in vivo*. Il existe différentes hypothèses pour expliquer la séparation des fonctions aujourd'hui dévolues à l'un ou l'autre des acides nucléiques. Je voudrais souligner ici que notre étude sur les hairpins ajoute un argument à l'attribution de rôles catalytiques à l'ARN. Les fonctions catalytiques sont généralement assurées par des structures tridimensionnelles complexes. Afin d'avoir une activité biologique, ces structures doivent être robustes pour pouvoir se maintenir dans le temps. Nous avons montré que la dynamique de déstabilisation des structures secondaires, dans le cas le plus simple des hairpins, est notablement plus lente pour les structures d'ARN que pour celles d'ADN. Cela confère donc aux structures d'ARN une meilleure stabilité face aux forces, internes ou externes, en présence dans les cellules vivantes. Cela a pu contribuer à orienter une séparation des fonctions entre l'ARN et l'ADN.





# ANNEXE Calibration du piège optique : modèle de Langevin

La calibration du piège fixe, sur lequel repose la mesure de la force, est effectuée grâce à la caractérisation du mouvement brownien d'une bille piégée. Le piège agit comme un puits de potentiel harmonique, la bille est donc soumise à une force de type  $\vec{F} = -k_p \vec{x}$  où  $k_p$  est la raideur du piège dans la direction du déplacement.

Le mouvement de la bille (à une dimension) est alors décrit par une équation de Langevin de la forme :

$$m\ddot{x} = -k_p x - \alpha \dot{x} + F_L$$

où  $F_L$  est la force de Langevin, qui rend compte des chocs aléatoires des molécules environnantes sur la bille.  $-\alpha \dot{x}$  est la force de frottement visqueux exercée par le milieu aqueux sur la bille, et  $m$  la masse de la bille.

Comme la bille est piégée à  $5 \mu m$  de la surface, les interactions hydrodynamiques entre la bille et la lamelle ne sont pas négligeables. Le coefficient de frottement visqueux  $\alpha$  ne peut donc pas s'écrire sous la forme de Stokes, mais doit être corrigé. La forme choisie est la loi de Faxen [87] :

$$\alpha = \frac{6\pi\eta R}{1 - \frac{9}{16}\left(\frac{a}{h}\right) + \frac{1}{8}\left(\frac{a}{h}\right)^3 - \frac{45}{256}\left(\frac{a}{h}\right)^4 - \frac{1}{16}\left(\frac{a}{h}\right)^5}$$

où  $\eta$  désigne la viscosité du milieu environnant,  $R$  le rayon de la bille, et  $h$  la distance entre la surface et le centre de la bille.

L'équation de Langevin peut être résolue dans l'espace de Fourier. La mesure du spectre fréquentiel du mouvement brownien de la bille dans le piège permet alors d'obtenir les paramètres de calibration du piège, en particulier sa raideur.

L'équation du mouvement dans le domaine fréquentiel s'obtient par transformée de Fourier :

$$-m(2\pi)^2 f^2 \tilde{x}(f) = -k_p \tilde{x}(f) - \alpha i 2\pi f \tilde{x}(f) + \tilde{F}_L(f)$$

où  $\tilde{x}(f)$  et  $\tilde{F}_L(f)$  désignent les transformées de Fourier respectives de  $x$  et  $F_L$ .

On a donc :

$$|\tilde{x}(f)|^2 = \frac{|\tilde{F}_L(f)|^2}{|-m(2\pi)^2 f^2 + \alpha i 2\pi f + k_p|^2}$$

En pratique, dans la limite des faibles nombres de Reynolds, le terme inertiel en  $f^2$  est négligeable devant les termes  $k_p$  et  $\alpha 2\pi f$ . L'équation précédente peut donc s'écrire sous forme

simplifiée :

$$|\tilde{x}(f)|^2 = \frac{|\tilde{F}_L(f)|^2}{k_p^2} \frac{1}{1 + (f/f_0)^2}$$

où  $f_0 = k_p/2\pi\alpha$ .

$|\tilde{F}_L(f)|^2$  s'obtient grâce au théorème d'équipartition de l'énergie :

$$\frac{1}{2}k_B T = \frac{1}{2}k_p \langle x^2 \rangle = \frac{1}{2}k_p \int_0^\infty |\tilde{x}(f)|^2 df$$

soit :

$$k_B T = k_p \int_0^\infty \frac{|\tilde{F}_L(f)|^2}{k_p^2} \frac{1}{1 + (f/f_0)^2} df$$

La force de Langevin étant due à des chocs aléatoires, son temps caractéristique de variation est très faible ( $\sim 10^{-13}s$ ) par rapport aux temps caractéristiques de l'électronique de détection. Son spectre de fréquence peut donc être assimilé à un bruit blanc. Par conséquent,

$$\frac{k_B T}{k_p} = \frac{F_L^2}{k_p^2} \frac{1}{2} \pi f_0$$

et donc

$$|\tilde{x}(f)|^2 = \frac{k_B T}{k_p} \frac{2}{\pi f_0} \frac{1}{1 + (f/f_0)^2} = \frac{4\alpha k_B T}{k_p^2} \frac{1}{1 + (f/f_0)^2}$$

La chaîne de détection électronique peut être modélisée par un filtre passe bas du premier ordre de fréquence de coupure  $f_e = 40$  kHz. Le signal électronique détecté  $U$  est proportionnel au déplacement de la bille, et donc de la forme suivante :

$$|\tilde{U}(f)|^2 = \frac{4\alpha k_B T}{k_p^2} \frac{1}{1 + (f/f_0)^2} \frac{S_0}{1 + (f/f_e)^2}$$

où  $S_0$  est la sensibilité du détecteur.

La fonction  $|\tilde{U}(f)|^2$  est mesurée expérimentalement avec un analyseur de spectre. La courbe obtenue est ensuite ajustée par le modèle ci-dessus, avec pour paramètres  $f_0$  et  $S = 4\alpha k_B T S_0 / k_p^2$ , et  $f_e = 40$  kHz.

La raideur du piège se déduit de  $f_0$  :  $k_p = 2\pi\alpha f_0$ .

Le coefficient de conversion tension/force  $\gamma$  est obtenu à partir de la mesure de  $S$ , dans la limite où  $f$  est petit devant  $f_e$  :

$$|F| = \gamma |\tilde{U}| = k_p |\tilde{x}| \quad \text{soit} \quad \gamma = k_p \frac{|\tilde{x}|}{|\tilde{U}|}$$

par conséquent

$$\gamma = \sqrt{\frac{k_p^2}{S_0}} = \sqrt{\frac{4\alpha k_B T}{S}}$$

La précision sur la mesure de la force est limitée par les incertitudes sur la température, la distance bille/surface, et le diamètre de la bille. En tenant compte de tous ces facteurs, la précision sur la force mesurée est de l'ordre de 5%.

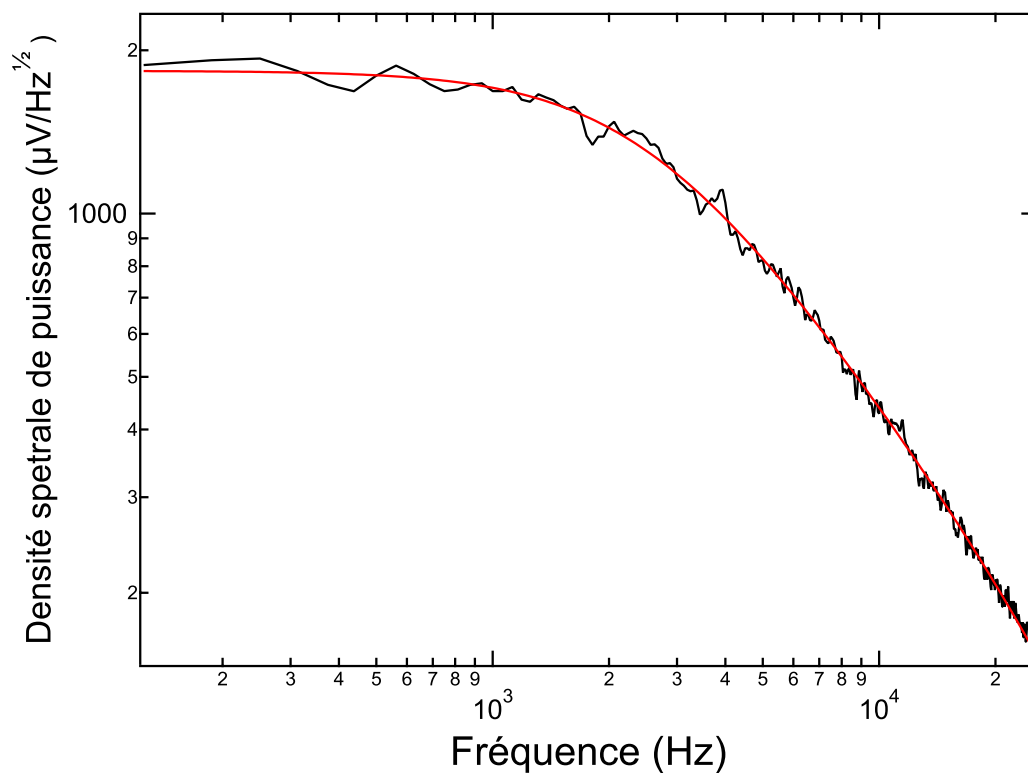


Figure 4.19 **Spectre de calibration du piège optique.** Il s'agit d'une mesure de densité spectrale de puissance d'une bille de polystyrène de  $1.87 \mu\text{m}$  de diamètre piégée à  $5 \mu\text{m}$  de la surface (courbe noire). La courbe rouge représente son ajustement théorique.



# ANNEXE Covalent coupling of proteins to carboxylated microspheres

« Polylink Protein Coupling Kit », Polysciences  
Ref : Technical Data Sheet 644

## MATERIAL

### **PolyLink Protein Coupling Kit :**

- PolyLink Coupling Buffer (50mM MES, pH 5,2 ; 0,05% Proclin 300)
- PolyLink Wash/Storage Buffer (10mM Tris, pH 8,0 ; 0,05% BSA ; 0,05% Proclin 300)
- PolyLink EDAC (Carbodiimide) : 750mg  
(Store powder dessicated at -20°C. Flood headspace with N<sub>2</sub> gas for best preservation. Warm the sealed vial to room temperature closed until condensation disappears. *Make working solution just before use.*)

### **Carboxylated Polystyrene or Silica microspheres :**

- Carboxylated Silica Microspheres, Bangs Laboratories, d=1,01 µm. Dry.
- Carboxylated Polymer Microspheres, Bangs Laboratories, d=2,19 µm. Suspension 10%.

### **Proteins :**

- Anti-digoxigenin, Fab-fragments from sheep, Roche
- Streptavidin, recombinant from *Streptomyces avidinii*, Roche

### **Additional reagents :**

- Tween 20
- BSA
- PVP
- Phosphate Buffer x10

**Centrifuge****Test tubes (1.5µL)****NOTES**

Centrifugation times will vary with the size and density of the particles :

Silica microspheres 1,01 µm : 5 minutes at 500 xG

Polymer microspheres 2,19 µm : 5 minutes at 5000 xG

Sonicate briefly : about 5-10 s, mode power

Resuspend pellet : energetic pipet mixing and vortexing until the solution is homogenous

**PROCEDURE**

1. Prepare 4 mL of *Tween Coupling Buffer* by adding 0,05% of Tween 20 in *PolyLink Coupling Buffer*.
2. Prepare 1 mL of *BSA Storage Buffer* by adding 0,5 mg/mL of BSA and 0,1% of Tween 20 in *PolyLink Wash/Storage Buffer*.
3. Warm microparticles, *Tween Coupling Buffer*, *PolyLink Wash/Storage Buffer* (1mL), and *BSA Storage Buffer* to room temperature.
4. Transfer 12,5 mg of microparticles into a 1,5 mL polypropylene microcentrifuge tube.
5. Pellet the Polymer microspheres (suspension) via centrifugation. Remove supernatant.
6. Resuspend microparticles in 0,4 mL of *Tween Coupling Buffer*. Sonicate briefly.
7. Pellet again via centrifugation. Remove supernatant.
8. Resuspend the microparticles pellet in 0,17 mL of *Tween Coupling Buffer*. Sonicate briefly.
9. Warm EDAC to room temperature. Just before use, prepare a 200 mg/mL EDAC solution by dissolving 10 mg PolyLink EDAC in 50 µL *Tween Coupling Buffer*. Use within 15 minutes.
10. Add 20 µL of the EDAC solution to each microparticle suspension. Allow the activation step to proceed for 15 minutes with mixing end-over-end. Sonicate briefly.
11. Add protein equivalent to 350-400 µg at 1-5 µg/µL in *Tween Coupling Buffer*. This may be prepared by dissolving 350 µg of protein in 100 µL of *Tween Coupling Buffer*.  
Note : The amount of protein bound to the microparticles is dependent on the concentration of protein in solution and the size of the microparticles. Refer to Tech Data Sheet 644.
12. Incubate for 2 hours at room temperature with mixing end-over-end. Sonicate briefly every 15-20 minutes.
13. Pellet microparticles via centrifugation.  
If needed, save the supernatant for determination of the amount of bound proteins.
14. Resuspend microparticles pellet in 0,4 mL *PolyLink Wash/Storage Buffer*.

15. Pellet microparticles via centrifugation.  
If needed, combine the supernatant for use in bound protein calculation.
16. Resuspend microparticles pellet in 0,4 mL *BSA Storage Buffer*. Allow the blocking step to proceed overnight at 4°C with mixing end-over-end.
17. Prepare 4 mL of PB/PVP Buffer by adding 10 mg/mL PVP, 1 µL/mL Tween 20, 100 µL/mL PB ×10 to *PolyLink Wash/Storage Buffer*. Warm to room temperature.
18. Pellet the Polymer microspheres via centrifugation. Remove supernatant.
19. Resuspend the microparticles pellet in 1,5 mL of *PB/PVP Buffer*. Sonicate briefly.
20. Pellet the Polymer microspheres via centrifugation. Remove supernatant.
21. Resuspend the microparticles pellet in 0,4 mL of *PB/PVP Buffer*.
22. Sonicate 15 minutes, mode power.
23. Dilute as needed in *PB/PVP Buffer*. Store at 4°C.





# ANNEXE Hairpins under tension : RNA versus DNA

# Hairpins under tension: RNA versus DNA

Mathilde Bercy and Ulrich Bockelmann\*

Laboratoire de Nanobiophysique, ESPCI ParisTech, 10 rue Vauquelin, Paris 75005, France

Received May 28, 2015; Revised August 04, 2015; Accepted August 15, 2015

## ABSTRACT

**We use optical tweezers to control the folding and unfolding of individual DNA and RNA hairpins by force. Four hairpin molecules are studied in comparison: two DNA and two RNA ones. We observe that the conformational dynamics is slower for the RNA hairpins than for their DNA counterparts. Our results indicate that structures made of RNA are dynamically more stable. This difference might contribute to the fact that DNA and RNA play fundamentally different biological roles in spite of chemical similarity.**

## INTRODUCTION

RNA molecules assemble into a large variety of structures, ranging from a simple hairpin to huge cellular machinery like the ribosome. The structural versatility of RNA is of critical importance to many of the complex functions of RNA, like enzymatic activity, sensing of metabolites and regulated processes, such as transcription termination and attenuation. RNA hairpins participate in the regulation of messenger RNA translation and degradation (1,2). This is made possible by the ability of a hairpin to switch between two distinct conformations, folded and unfolded, where one conformation allows a process and the other one inhibits it. Single stranded DNA can adopt similar secondary structure as RNA and has even been reported to act like an enzyme (deoxyribozyme) *in vitro* (3). It may thus appear surprising that DNA is almost exclusively found in double-stranded form in the biological cell, being only transiently unfolded as for instance during replication or transcription. As an exception single-stranded DNA occurs in certain viruses, but still mainly serves as a hereditary support (4). RNA occupies the domain of complex three-dimensional nucleic acid structures.

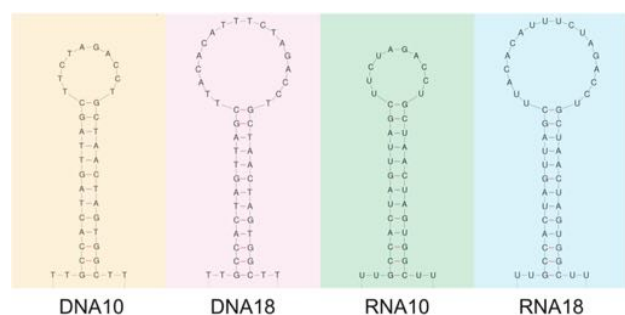
With one Watson–Crick paired double helix stem and one single-stranded loop, the hairpin is the simplest secondary structure element and a building block for more complex structures (1). The double helix structures of RNA and DNA are different. RNA double strands adopt an A-form helix that is known to be wider and shorter than the B-form DNA helix (11 nucleotides per turn in the A helix versus 10 in the B helix). RNA and DNA hairpins have been studied separately since the seventies. The first experi-

ments were done with bulk calorimetric techniques (5). Single molecule fluorescence (6) and force (7,8) measurements more recently provided further insight into the kinetics and thermodynamics of hairpins. To the best of our knowledge, no systematic comparison of the folding dynamics of RNA and DNA hairpins has been published so far.

We performed single molecule force measurements with a dual-beam optical trapping setup on a series of molecular constructs containing a DNA or RNA hairpin. Pulling in a direction transverse to the stem, we can impose a mechanical constraint and follow the folding and unfolding of individual hairpins in time. In the main series of experiments, dynamic stretch and release cycles are applied. We used different pulling speeds (corresponding to force loading rates in the range 1–25 pN/s) and extracted kinetic and thermodynamic parameters of the hairpins by comparing the experimental data with a theoretical description based on non-equilibrium statistical physics. Analysing these results, it is crucial to keep in mind the distinction between the activation energy barrier and the equilibrium free energy difference between the two hairpin states (resp.  $E$  and  $\Delta G$  in the following). There is no general relationship between these two quantities. In addition, the attempt frequency and the force loading rate are the key parameters in this case of non-equilibrium conditions.

We find that the RNA hairpins unfold at higher average force than the DNA hairpins. The force hysteresis, i.e. the difference between transition forces measured during stretching and releasing, increases with pulling speed for all hairpins, in accordance with the theoretical predictions. The RNA hairpins exhibit a much larger force hysteresis than their DNA counterparts. Characteristic transition rates, fitted to the DNA hairpin data, exceed the corresponding RNA rates. This quantitative difference is confirmed and further illustrated by experiments with immobile traps, where the constructs are maintained at constant extension. In this configuration, we find that a large-loop DNA hairpin flips with sub-second dwell-times between its folded and unfolded states, while such flipping was not observable in the corresponding RNA case. As described in the discussion section, our results suggest that RNA structures are less perturbed by mechanical load than DNA structures. The latter tend to switch more rapidly between alternative conformations under the same experimental condition.

\*To whom correspondence should be addressed. Tel: +33 1 40794761; Fax: +33 1 40795120; Email: ulrich.bockelmann@espci.fr



**Figure 1.** The four studied hairpins. From left to right: DNA10, DNA18, RNA10 and RNA18. They exhibit a 13 base pair stem of the same sequence and loops of either 10 (DNA10 and RNA10) or 18 (DNA18 and RNA18) nucleotides. Apart from the replacement of DNA's Thymine by RNA's Uracil, the base sequences are the same for the DNA and RNA hairpins. The hairpins are presented with their 5' ends at the bottom left side.

## MATERIALS AND METHODS

### Dual-beam optical trapping interferometer

Our dual-beam optical trapping setup relies on a CW Nd:YVO<sub>4</sub> laser, emitting 2W at 1064 nm (Millenia IR, Spectra Physics, Irvine, CA). The laser beam is split in two beams of orthogonal polarisation. One of the two beams encounters only immobile optical compounds, while the second one is reflected by a mirror mounted on a piezoelectric stage. To avoid issues caused by residual interference between the polarized beams, the mobile beam is frequency shifted with an acousto-optic modulator (9). The two beams are recombined and focused by a 100× oil immersion microscope objective (NA 1.4, Nikon, Tokyo, Japan). This generates two optical traps: a fixed and a mobile one. The setup allows to adjust the distance between the two traps over a range of several micrometers with nanometric precision.

Behind the sample, the beam is collected by a 60× water immersion microscope objective (NA 1.2, Olympus, Tokyo, Japan). The polarisation corresponding to the mobile trap is rejected with a Glan polariser, and the displacement of a bead in the fixed trap is measured by back focal plane interferometry using a PSD detector (Pacific Silicon, Westlake Village, CA, USA). After calibration by analysis of the power spectral density of a captured bead, force is obtained with piconewton precision. More details on the setup have been presented elsewhere (9–11).

### Hairpin molecular constructs

We designed two hairpins with the same stem of 13 base pairs and a loop of 10 or 18 nucleotides, both as an RNA and a DNA molecule. The GC-content of the stem is close to 50%. The resulting four hairpins are depicted in Figure 1. The colour code used here to distinguish the four hairpins is conserved throughout the article. A distribution of hairpin loop length was determined from the atomic resolution RNA structures available in the Protein Data Bank in January 2013 and it was found that loop sizes between 4 and 20 nucleotides are most frequent (12).



**Figure 2.** Schematic representation of a molecular construct linked to functionalized beads. A DNA (resp. RNA) hairpin is inserted between two DNA/DNA (resp. RNA/DNA) handles. Each handle has a crystallographic length of about 0.5  $\mu\text{m}$  (resp. 0.4  $\mu\text{m}$ ).

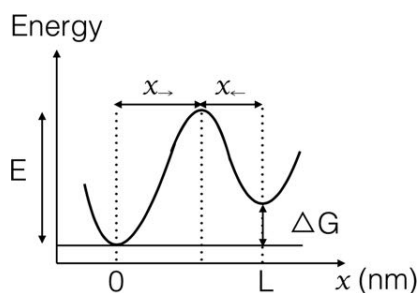
DNA oligonucleotides carrying the hairpin sequence of 36 or 44 bases have been inserted between two double-stranded handles of about 2000 base pairs each. The original DNA corresponds to a sequence coding for 23S ribosomal RNA in an *Escherichia coli* plasmid, which serves as a template for two polymerase chain reactions (PCRs). In addition to the part complementary to the template, one primer of each PCR includes one half of the sequence encoding the hairpin. The PCR products are then digested with XbaI, which generates cohesive ends in the middle of the hairpin sequence. After a ligation step, the complete hairpin sequence appears in the DNA. The resulting  $\approx 4000$  nucleotide double-stranded DNA is amplified by PCR. For the RNA hairpin constructs, an *in vitro* transcription of this DNA is performed.

To generate the handles, the DNA strands complementary to the sequence before and after the hairpin part are synthesized by PCR. A digoxigenin functionalized nucleotide is inserted at the 5' end of one of the handles through the PCR primer. The other handle undergoes an enzymatic treatment (XhoI and Klenow, New England Biolabs, Ipswich, MA, USA) after PCR to add a biotin functionalized nucleotide at the 3' end (10).

To hybridise the handles with the hairpin containing strand, each double-stranded handle and the complete double-stranded DNA (resp. single-stranded RNA) are incubated for 30 s at 95°C, and the sample is cooled down to 4°C over 2 h. Carboxylated polystyrene (resp. silica) microspheres (Bangs Laboratories, Fishers, IN, USA) of 2.19  $\mu\text{m}$  diameter (resp. 1.01  $\mu\text{m}$ ) are functionalized (PolyLink Protein Coupling Kit, Polysciences, Warrington, PA, USA) with biotin (resp. anti-digoxigenin) from Roche (Basel, Switzerland). Prior to each experiment, the beads are linked to the molecular constructs by incubation at room temperature (11) (Figure 2).

### Single molecule force measurements

After incubating the molecular constructs with the beads, the sample is diluted in the working buffer (20 mM Tris, 50 mM K acetate, 5 mM Mg acetate, pH7.5) and put in a  $\approx 100$   $\mu\text{m}$  deep chamber between two microscope cover slips (11). The chamber is placed between the microscope objectives. The temperature inside the sample during an experiment has been measured. It is 29°C. The experimenter spots a pair of beads linked by a molecular construct, and captures one bead in each trap. The mobile trap displacement is monitored by a custom software programmed in the Labview environment (National Instrument, Austin, TX, USA). After an appropriate anti-aliasing filter, the voltage that is proportional to the force is acquired and displayed in real-time.



**Figure 3.** Schematic representation of the simplified energy landscape used in our theoretical description. The state of the folded hairpin occurs at  $x = 0$  and the state of the unfolded hairpin at  $x = L$ . The distances between the transition state and the folded and unfolded states are denoted by  $x_{\rightarrow}$  and  $x_{\leftarrow}$ , respectively. By definition, we have  $L = x_{\rightarrow} + x_{\leftarrow}$ .

### Theoretical description and data analysis

Hairpin folding and unfolding are complex dynamical processes. They depend on several characteristics of the molecular construct. The base sequence of the hairpin stem induces a complex energy landscape and the size of the loop and the elasticity of the DNA/RNA handles also are important. Moreover, the force measurement by itself influences the experimental result, via the trap stiffness and the externally imposed distance versus time protocol. The imposed distance between the centres of the optical traps is subject to position noise arising from piezo-control electronics, mechanical vibration and acoustic perturbations.

We use an out-of-equilibrium statistical physics description that is closely related to earlier publications (8,13–17) (usually referred to as Bell-Evans model), and including a minimal set of parameters. We replace the complex molecular energy landscape by the single-barrier landscape presented in Figure 3 and the elasticities of the molecular construct and optical traps are lumped together in one effective stiffness  $k_{eff}$ . A free energy barrier of height  $E$  separates the folded ( $x = 0$ ) and unfolded ( $x = L$ ) states. The free energies of the two states differ by  $\Delta G$ . In the absence of force, the folded to unfolded transition rate is  $k_0 = k(F = 0) = \nu_0 \exp(-E/k_B T)$  (it includes the barrier height  $E$  and the attempt frequency  $\nu_0$ ). The reverse process is associated to a transition rate  $k_1 = k_0 \exp(\Delta G/k_B T) = \nu_0 \exp(-(E - \Delta G)/k_B T)$ . Applying an external force  $F$  tilts this landscape. In the limit of high barrier ( $E \gg k_B T$ ) and small applied force ( $F \ll E/x_{\rightarrow}$ ), the position of the transition state  $x_{\rightarrow}$  does not vary significantly and the barrier for unfolding is reduced by an amount  $F x_{\rightarrow}$ . Increasing (resp. decreasing) the external force linearly with time within this model, the unfolding probability  $p_{\rightarrow}(F)$  (resp. folding probability  $p_{\leftarrow}(F)$ ) can be derived analytically. More details about the model and a complete derivation of these probabilities are presented in Supplementary Section S1. Both probabilities explicitly depend on the pulling speed.

The distribution of the force hysteresis between unfolding and folding corresponds to the convolution of these probabilities. To obtain the experimental hysteresis distributions, for each molecule the unfolding and folding forces measured are paired to generate all possible hysteresis values ( $F_{unfold} - F_{fold}$ ). The obtained values are weighted accord-

ing to the number of force cycles actually performed on the molecule, and normalized. More details about the building of the histograms are presented in Supplementary Section S3. We numerically calculate  $p_{\rightarrow}(F)$ ,  $p_{\leftarrow}(F)$  and their convolution, compare the theoretical result with the measured hysteresis distribution and fit the position of the transition state  $x_{\rightarrow}$  and a characteristic transition rate  $k_0$ . The fit is globally performed on the data of the four velocities in order to obtain a single set of parameters for each hairpin (see Supplementary Section S1). We analyse hysteresis values rather than separate unfolding and folding forces because the uncertainties of the hysteresis values are smaller in our measurements. This is explained by a simple estimation of the experimental errors, presented in Supplementary Section S2. All other parameters are measured, taken from the literature or estimated with the structure prediction program *mfold* (18). More details about the theoretical description, the parameters determination and the fitting procedure are provided in the Supplementary Sections S1–S3.

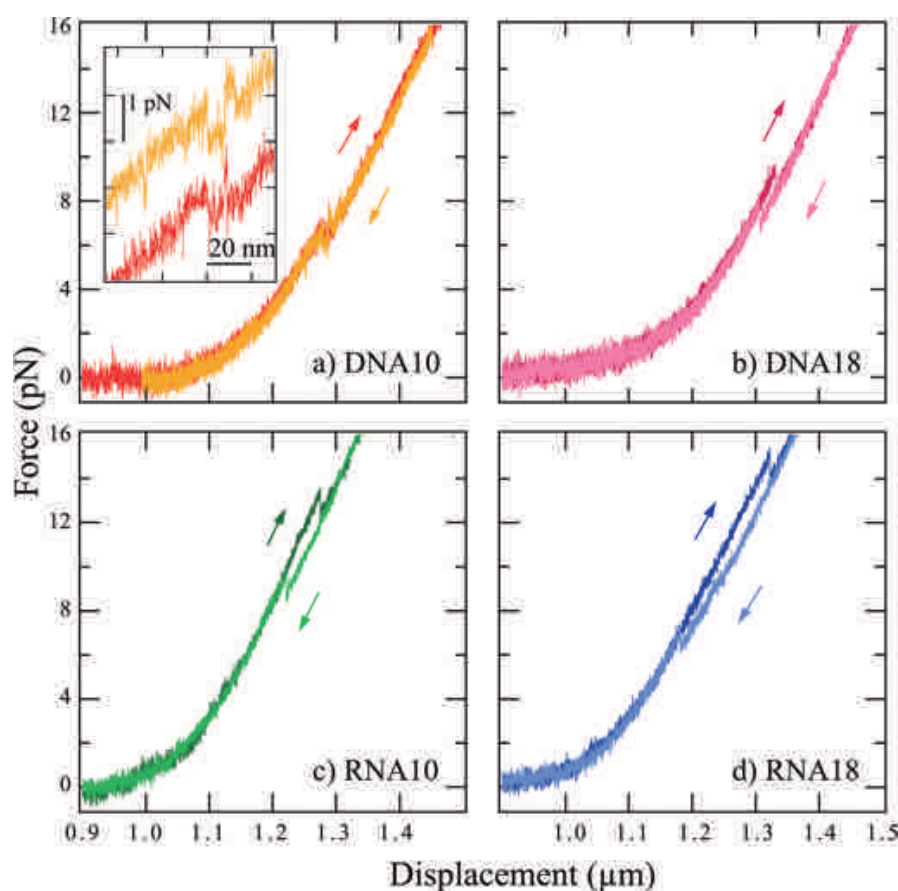
### RESULTS

In the beginning of a measurement with varying extension, the mobile trap is moved apart from the fixed trap with a constant velocity. This movement is stopped at a relative extension of the construct (ratio of total length under tension to crystallographic length) of  $\approx 1.4$ , where the hairpin is always in the unfolded state. The movement is then inverted, so that the traps reapproach at the same speed and the hairpin can refold. This cycle is performed several times on the same molecule.

In Figure 4, we present typical force versus extension curves for the four hairpins at 50 nm/s. The left part of each curve (small displacement range not entirely shown) corresponds to the entropic response of the double-stranded handles where the force remains small. Force increases with a rising slope when the displacement approaches the crystallographic length of the molecular construct. This regime corresponds to the elastic response of the double-stranded handles, well described by the Worm Like Chain model (19) (see Supplementary Section S4, Supplementary Figure S1). For the DNA hairpin constructs, the handles are double-stranded DNA, whereas the RNA hairpin constructs have hybrid DNA/RNA handles. This explains the length difference at the starting point of the elastic response. At a certain extension, a sudden drop in force by about 1 pN is observed, which corresponds to the unfolding of the hairpin. After this event, the force increases again, now stretching the whole construct with the double-stranded handles and the unfolded hairpin in series. No intermediate state is observed during hairpin unfolding.

When releasing the strain by inverting the direction of displacement and bringing the traps closer to each other, the hairpin can fold back into its initial structure. As no intermediate state is observed during the corresponding sudden upwards rise in force, this process also appears to be cooperative.

The stretch and release curves are indistinguishable, except for the region where the unfolding (resp. folding) occurs. For a velocity of 50 nm/s and for all constructs except DNA10, a force hysteresis between unfolding and folding is



**Figure 4.** Force versus extension curves at 50 nm/s for the four different hairpins. The darker curves correspond to the stretching of the construct, whereas the lighter ones represent the release process. For both the DNA and RNA hairpins, the hysteresis between unfolding and folding increases with loop length. For the same hairpin sequence, the hysteresis is larger for the RNA construct than for the DNA one. The inset provides a closer look to the force flips observed with DNA10. In this inset, the unfolding and folding curves are shifted vertically for clarity.

observed. The hysteresis indicates that transitions occur out of thermal equilibrium. On the other hand, DNA10 shows several force flips between the folded and unfolded states during dynamic stretching and releasing of the construct (Figure 4a, inset). These force flips are a signature of a close-to-equilibrium dynamics (20). They are no longer observed and hysteresis becomes notable when the displacement velocity is increased to 150 nm/s and above (Figure 5). No force flips are observed for the other three hairpins.

Experimental data presented in Table 1 show that, both for the DNA and RNA hairpins, the most probable folding force decreases with increasing loop size, while the unfolding force is less sensitive to this parameter. Both the unfolding and the folding forces of the RNA hairpins are higher than the corresponding forces of the DNA hairpins.

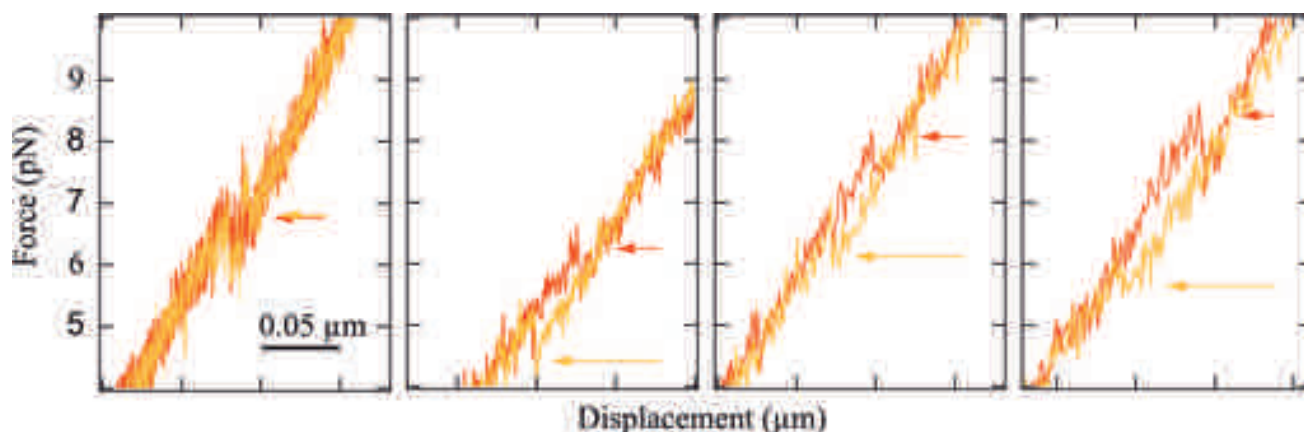
Measuring the successive unfolding (resp. folding) forces during stretch/release cycles on different molecules allowed us to build the hysteresis histograms for the four hairpins at the different displacement velocities. For DNA10 at 50 nm/s, when flips occur, only the first unfolding and first refolding forces are measured, according to the model used. Figure 6 shows such histograms for the four hairpins at 150 nm/s. The histograms corresponding to the other velocities are presented in Supplementary Section S6 and Supplemen-

**Table 1.** Most probable folding and unfolding forces, determined from measurements at a velocity of 50 nm/s. See Supplementary Figure S2 for an example of unfolding and folding forces distribution

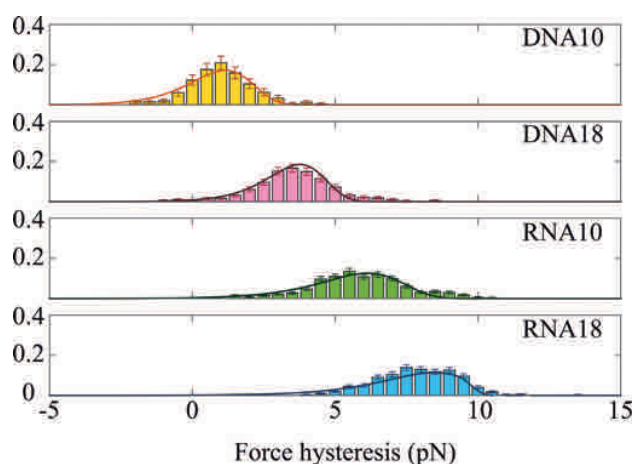
Hairpin	Unfolding force (pN)	Folding force (pN)
DNA10	6.4	6.2
DNA18	6.1	3.6
RNA10	13.6	8.8
RNA18	14.1	5.8

The presented force values correspond to the maxima of a Gaussian fit of the measured distributions of the forces where unfolding or folding occurs. The unfolding force distributions are built from 77 measured unfolding events for DNA10, 93 events for DNA18, 179 events for RNA10 and 146 events for RNA18. The folding force distributions are based on 79 measured folding events for DNA10, 70 events for DNA18, 148 events for RNA10 and 114 events for RNA18. The widths of these distributions are about 2.5 pN (FWHM).

tary Figure S3. Our data show that for a given hairpin structure and under identical experimental conditions (same experimental buffer, same temperature, same pulling speed), the hysteresis is notably higher for RNA than for DNA. We find that the hysteresis increases with loop length for both DNA and RNA.



**Figure 5.** Force versus extension curves for DNA10 at 50, 150, 300 and 450 nm/s (from left to right). The darker curves correspond to the stretching of the construct, whereas the lighter ones represent the release process. The arrows indicate the force of the unfolding and folding events. The force flips observed with DNA10 at 50 nm/s vanishes at 150 nm/s and above. The hysteresis between unfolding and folding increases with the pulling velocity.



**Figure 6.** Histograms of the hysteresis measured at 150 nm/s on the four different hairpins. The experimental distributions are fitted to calculated convolution of the unfolding and folding probability distributions (solid line), as explained in Materials and Methods and Supplementary Section S1 (Equations (4–6)). For each fitted histogram, the root mean squared error (*rmse*) is in the range  $(2.0\text{--}3.0) \times 10^{-3}$ . For DNA10 the data come from 46 stretch/release cycles on 12 molecules. DNA18: 95 cycles on 13 molecules. RNA10: 152 cycles on 44 molecules. RNA18: 150 cycles on 21 molecules.

For all four hairpins, the force hysteresis increases with displacement velocity. Supplementary Figure S4 (Supplementary Section S7) presents histograms of the hysteresis recorded for RNA10 at 50, 150, 300 and 450 nm/s as an example. A clear increase of the mean hysteresis is visible, with its value going from 5 pN at 50 nm/s to 7.5 pN at 450 nm/s. The hysteresis values for the four hairpins are given in Supplementary Table S2. For each hairpin and pulling speed, variability in the unfolding and folding forces is seen, inducing a variability in the hysteresis. This behaviour is expected from the stochastic character of the unfolding and folding processes.

Theoretical analysis of the hysteresis histograms enables us to extract thermodynamic and kinetic information about the hairpin structures. The experimental distributions were

fitted to calculated probability distributions, as defined in Materials and Methods and supplementary information. For each hairpin, the fitting procedure is applied over the global data set, combining the four different speeds. The calculated force hysteresis distribution is presented by the solid line in Figure 6.

Measured mean transition forces  $F_t$  (average of the mean opening force and the mean folding force at 50 nm/s) and transition lengths  $L$ , as well as theoretical free energy values that were obtained by simulation (see Supplementary Section S1 for details) are presented in Table 2. Two parameters, the distance  $x_{\rightarrow}$  between folded state and transition state and the characteristic transition rate  $k_0$  were fitted to the experimental data and are presented in Table 3. The quality of the fits is good, with a root mean squared error in the range  $(1.0\text{--}5.0) \times 10^{-3}$  for each fitted histogram (*rmse*, see Supplementary Section S1, Equation (7) for definition) and allows us to univocally determine the two parameters.

The results for  $x_{\rightarrow}$  indicate a clear asymmetry in the energy landscape for both DNA and RNA hairpins (all ratio  $x_{\rightarrow}/L < 0.5$ ). The DNA values of  $x_{\rightarrow}/L \approx 0.3$  are higher than the RNA values of  $x_{\rightarrow}/L \approx 0.17$ . The transition rates  $k_0$  differ by a factor of about 10 between DNA18 and RNA18. The ratio is also about 10 between DNA10 and RNA10. These results confirm and quantify the faster dynamics of DNA as compared to RNA, already suggested by the appearance of force flips in the DNA10 curves at 50 nm/s.

The observation that the transition rates and hence the dynamics differ between the DNA and RNA hairpins is further confirmed by additional measurements performed with the four hairpins at constant extension. Here, the molecular construct is brought to a defined extension, which is then maintained constant during 10 s to 2 min. If the energy barrier is not too high, there are extension values where the hairpin can spontaneously flip between its folded and unfolded states (7,20); the measured force flips are induced by thermal fluctuations. If the corresponding dwell times do not significantly exceed the time-scale of the measurement,

**Table 2.** Data obtained from measurements at 50 nm/s (a) or theoretical prediction (b)

Hairpin	(a) Transition force $F_t$ (pN)	(a) Transition length $L$ (nm)	(b) $\Delta G$ mfold (kJ/mol)	(b) $\Delta G$ stretch (kJ/mol)	(b) $\Delta G$ total (kJ/mol)
DNA10	6.3	14.8 ± 2	63 ± 3	20 ± 1	83 ± 4
DNA18	4.85	16.6 ± 2	60 ± 3	18 ± 1	78 ± 4
RNA10	11.2	17.7 ± 2	89 ± 4	32 ± 2	121 ± 6
RNA18	9.95	18.4 ± 2	87 ± 4	35 ± 2	122 ± 6

mfold: <http://mfold.rna.albany.edu>, see Materials and Methods and Supplementary Information for parameter determination.

**Table 3.** Parameters deduced from hysteresis measurements

Hairpin	Transition state $x_{\rightarrow}$ (nm)	Transition state $x_{\rightarrow}/L$	Transition rate $k_0$ ( $s^{-1}$ )
DNA10	4.4 ± 0.2	0.30	$(1.0-3.0) \times 10^{-4}$
DNA18	4.7 ± 0.2	0.28	$(3.8-8.5) \times 10^{-5}$
RNA10	3.0 ± 0.2	0.17	$(2.9-5.5) \times 10^{-5}$
RNA18	3.1 ± 0.2	0.17	$(3.0-10) \times 10^{-6}$

By definition, we have  $x_{\leftarrow} = L - x_{\rightarrow}$  and  $x_{\leftarrow}/L = 1 - x_{\rightarrow}/L$ ,  $k_0 = k(F=0) = v_0 \exp(-E/k_B T)$

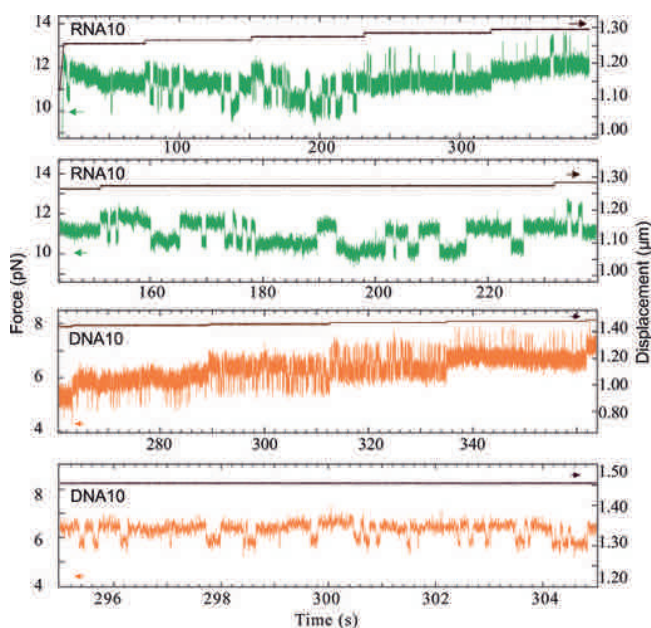
these spontaneous transitions can be observed experimentally.

Figure 7 (top) shows the force recorded on RNA10 as a function of time, while extension is held constant apart from stepwise increases. At the smallest extension, the hairpin remains folded most of the time, but already displays some brief passages to the unfolded state. The mean lifetime in the unfolded state progressively increases with extension, and at a given point equals the one in the folded state (this situation is presented in the second panel). At even larger extensions, the unfolded state becomes energetically favorable, and the hairpin only folds sporadically. The same global behaviour is observed for DNA10 (two bottom panels), but the transitions occur more frequently and the dwell-times of the folded and unfolded states are shorter than for RNA10. This observation is in line with the different  $k_0$  values of RNA10 and DNA10, presented in Table 3.

DNA10, DNA18 and RNA10 present spontaneous force flips at constant extension. For RNA18 however, no extension was observed where both spontaneous unfolding and folding occurs, although the two RNA hairpins are close in terms of the theoretically predicted  $\Delta G$  values. This difference between RNA10 and RNA18 is consistent with the observation that the  $k_0$  value of RNA18 obtained from the varying extension data is an order of magnitude smaller than the one of RNA10.

## DISCUSSION

We observe that the RNA hairpins unfold and fold at higher force than the DNA hairpins. This may be expected from the known fact that in terms of free energy difference  $\Delta G$  an RNA duplex is more stable than the DNA duplex of equivalent sequence. Note however that there are three problems with this simplified view. First, the denaturation mechanisms under force and by thermal melting are different (21). Whereas thermal melting involves independent and stochastic fluctuations of the stem bonds until denaturation



**Figure 7.** Spontaneous transitions between folded and unfolded hairpin states at constant extension. (Top panel) Force recorded on RNA10 (green) as a function of time, while the extension is constant apart from stepwise increases (black). Force flipping is observed for each extension value. (Second panel) Zoom on an extension where both states have a similar probability of occupation: the average dwell times are 1.77 s in the unfolded state and 3.1 s in the folded state. Presented time range: 80 s. (Third panel) Force recorded on DNA10 (orange) as a function of time, while the extension is constant apart from stepwise increases (black). Force flipping is observed for each extension value. (Bottom panel) Zoom on an extension where both states have a similar probability of occupation: the average dwell times are 0.11 s in the unfolded state and 0.25 s in the folded state. Presented time range: 10 s. The transition rate is much higher for DNA10 than for RNA10.

occurs, force-induced unfolding is a directed process. Even if the bonds still fluctuate, the stem is sequentially opened



from the bottom of the hairpin to the top. Similarly, the initial conditions for hairpin thermal folding are different from the ones for force reduction. Whereas folding of a thermally denatured molecule starts from a relatively compact conformation, the force measurement imposes a fully stretched initial construct. Second, the unfolding force can be estimated by a relation of the type  $F = \Delta G/L$  only if the transition occurs sufficiently close to thermal equilibrium. In the non-equilibrium case the actual barrier height, the attempt frequency and the force loading rate are the main parameters. There is no general relation between the equilibrium  $\Delta G$  and the barrier height. Third, the elastic energy usually plays a significant role in force-induced unfolding. If for the sake of simplicity we consider the equilibrium case, the estimation of the opening force is of the form  $F = \Delta G_{tot}/L$ , but the molecular elasticity influences both the characteristic energy ( $\Delta G_{tot} = \Delta G_{mfold} + \Delta G_{stretch}$ , in the notations of Table 2) and the characteristic length  $L$  ( $L$  is the sum of the molecule length at zero-force and the force-induced extension). As can be seen from Table 2, the stretching energy contributions are sizeable for both DNA and RNA.

The hysteresis is larger for RNA18 than for RNA10 and the same holds for DNA18 compared to DNA10. This mainly arises from a loop-length dependence of the folding force. As shown in Table 1, the folding forces significantly decrease with increasing loop size, while the unfolding forces change weakly. To initiate hairpin folding a loop must form to approach the bases of the stem for base pairing. As folding occurs against an external force  $F$  in the present case, this initiation requires a mechanical work of the order  $W = FL$ , where  $L$  denotes the length of the unfolded loop. For similar  $W$  (energy fluctuations of similar amplitude), the different  $L$  thus implies that folding of the small-loop hairpin occurs at higher  $F$  than folding of the large-loop one. Of note, it was found by fluorescence techniques that even without external force the rate of hairpin closing decreases with increasing loop size (6).

We find that the transition rate  $k_0$  is smaller for the RNA hairpins than for their DNA counterparts. As shown in Figure 7, the RNA hairpin that exhibits smaller  $k_0$  also shows longer dwell times in both hairpin states. Smaller  $k_0$  leads to larger hysteresis in the constant velocity measurement, since both unfolding and folding happen statistically later on the force ramp. This relation between  $k_0$  on the one hand and magnitude of the force hysteresis on the other hand is observed in our measurements. RNA10 (resp. 18) exhibits smaller  $k_0$  and larger hysteresis than DNA 10 (resp. 18) at given velocity. For example, the transition length  $L$  corresponding to RNA10 unfolding (resp. folding) is  $\approx 20$  nm. Considering a typical equilibrium dwell time of 1 s for this hairpin (estimated from Figure 7, second panel), a pulling velocity of 50 nm/s induces a displacement of 50 nm during this time. This displacement is larger than  $L$ , and the folding (resp. unfolding) transition is thus highly unlikely. The unfolding transition (resp. folding) thus occurs out of equilibrium at this speed. On the contrary, the shorter dwell time of DNA10 ( $\approx 0.1$  s) induces a displacement ( $\approx 5$  nm) that is smaller than  $L$  ( $\approx 15$  nm) and allows for flipping between the two states. In this case, the measurement can be considered as close to equilibrium (Figure 4a, inset).

As shown in Table 3, the transition state is closer to the folded state than to the unfolded state. We convert the values of  $x_{\rightarrow}$  to the number of unfolded nucleotides, using a procedure described earlier (27). For DNA10 (resp. DNA18) the transition state corresponds to 9 (resp. 11) unfolded nucleotides, or base pair 5–6 of the hairpin stem counted from the bottom. For RNA10 and RNA18 we obtain 5 (resp. 6) nucleotides, corresponding to base pair 3 from the bottom. We tentatively attribute these positions to the stem sequence: the first base pairs from the bottom are three G-C (Figure 1), conferring a greater stability to the bottom compared to the top of the stem. This interpretation is consistent with the fact that no intermediate state is observed during hairpin unfolding. Simply speaking, when the mechanical work becomes sufficient to unfold the initial G-C rich regions, the whole stem bursts open cooperatively. Within our model that assumes a single transition state, the folding of the hairpin is correspondingly expected when the G-C rich part of the stem reforms. The fact that no intermediate states were observed in our measurements indicates that the folding and unfolding transitions occur with a high degree of cooperativity. For comparison, we have calculated unfolding free energy landscapes of the hairpins. This study is described in section S8 of the supplementary information. The calculations are based on mfold free energies, include the sequential energy contributions of the different base pairs of the stem and the single-stranded loop, but do not include cooperativity. Interestingly the calculated landscapes exhibit their maxima at the top of the hairpin stem, in contrast to the measurements indicating transition states closer to the bottom of the stem. The differences are significant. First, they provide additional support to the idea that cooperativity is important for all investigated hairpins. Moreover, the comparison suggests a higher degree of cooperativity in RNA than in DNA, since we experimentally observe that the RNA hairpins exhibit smaller  $x_{\rightarrow}$  than the DNA hairpins.

There are structural arguments both for the stem and the loop that support a stronger cooperativity of the RNA hairpin as compared to the DNA hairpin. The stem of the RNA hairpin exhibiting a A-form helical structure is shorter and therefore more compact than the stem of the DNA hairpin, which is a B-form double helix. In the RNA helix, the constraint caused by the opening fork may thus more easily extend over several base pairs than in the DNA helix. It is expected that interactions between nucleotides in the hairpin loop contribute in a non-negligible manner to the overall stability. Indeed RNA is known (and DNA suspected) to build non Watson–Crick base pairs and we note that mfold does not fully take the corresponding free-energy contributions into account (22,23). The 2'-OH group of the RNA nucleotide forms hydrogen bonds, while the 2'-H of DNA does not. In this regard, nuclear magnetic resonance studies of chimeric RNA stem/DNA loop hairpins revealed that 2'-H in the hairpin loop reduces the overall stability and also suggests that the structure of the stem, A-form or B-form, strongly influences the loop structure (24).

We tentatively attribute the slower dynamics of the RNA hairpins to a stronger cooperativity. For strong cooperativity, barriers against folding and unfolding may be localized and high, since they can cumulate energy contributions of

many structural elements. With weak cooperativity, on the other hand, the energy contributions of the structural elements distribute more equally over the reaction coordinate, high local barriers are less probable to occur and therefore passage of this landscape by a random walk that is biased by the external force is faster.

Measurements on more complex RNA structures showed pronounced structures in the force versus extension curves (or length versus force curves for force-clamp measurements) and hysteresis occurred even at low pulling speed (25–27), in qualitative agreement with the present findings. Recently, an investigation of the mechanical properties of double-stranded RNA under force and torque has been published (28). The authors applied torque to a torsionally constraint double-stranded RNA construct using magnetic tweezers. They investigated in particular a plectonemic buckling transition. They find that the characteristic transition rate of this buckling transition is two orders of magnitude smaller for double-stranded RNA than for double-stranded DNA. It is interesting to mention this study in the present context because it shows a slower dynamics for RNA than for DNA, although the buckling transition is different from hairpin folding.

A typical RNA molecule has a complex three-dimensional shape stabilized by several duplexes. The individual duplexes are connected by single strands of varying length, which allows for propagation of conformational changes within the RNA structure. This propagation of strain and displacement is influenced by the elasticity of the connecting single strands. Forces of the order  $k/\Delta x$  develop, where  $k$  is the stiffness of the molecular linker and  $\Delta x$  the relative displacement. This global picture suggests that the configuration studied in the present paper represents an elementary building block of complex three-dimensional structures formed by a single-stranded nucleic acid chain. Specifically, our result that the dynamics of folding and unfolding under mechanical load is slower in RNA hairpins than in the corresponding DNA hairpins translates itself into predicting that RNA structures are dynamically more stable. In other words, structures formed by single-stranded DNA rather than RNA exhibit significantly faster interconversion between alternative configurations. There are of course good reasons of evolutionary origin to explain why the domain of biologically-active three-dimensional nucleic acid structures is traditionally occupied by RNA. Our suggestion that RNA structures are dynamically more stable than DNA structures may be an interesting additional argument to understand why single-stranded DNA was not selected to become a molecule that assures its biological role by a complex internal structure.

## CONCLUSION

Two RNA hairpin structures of different loop size were compared with their DNA equivalents. Under the same experimental conditions, the forces needed to unfold the DNA hairpins are systematically lower. Hysteresis between unfolding and folding is more pronounced for RNA than for DNA. For all studied hairpin types, the hysteresis increases with the pulling speed. Imposing a constant distance, spontaneous force flips are observed for the large loop DNA

hairpin, but are absent for the equivalent RNA structure. For small loop hairpins, flipping between the folded and unfolded states occurs with both DNA and RNA. Comparison of the experimental data with a theoretical description allowed us to estimate the position of the transition state and a characteristic transition rate. The latter is found to be smaller for RNA than DNA. From all these results emerges the picture that RNA exhibits a more pronounced out-of-equilibrium character than DNA. The force-induced conformational dynamics of hairpins is slower in RNA than in DNA. The results suggest that structures made of RNA are more stable against internal and external forces than structures made of single-stranded DNA and therefore the former are more apt to form complex structures that are required for an active biological function.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Thierry Bizebard, Josette Banroque, Kyle Tanner and Marc Dreyfus for advice and assistance in the preparation of the biological samples, Emilie Gontrand and Emel Çetin for recording part of the data and David Lacoste for stimulating discussions.

## FUNDING

Human Frontier Science Program [RGP008/2014]. Funding for open access charge: Human Frontier Science Program [RGP008/2014].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Bevilacqua, P.C. and Bloise, J.M. (2008) Structures, kinetics, thermodynamics, and biological functions of RNA hairpins. *Annu. Rev. Phys. Chem.*, **59**, 79–103.
2. Adam, C.C. and Stern, D.B. (1990) Control of mRNA stability in chloroplasts by 3' inverted repeats: effects of stem and loop mutations on degradation of *psbA* mRNA in vitro. *Nucleic Acids Res.*, **18**, 6003–6010.
3. Silverman, S.K. (2004) Deoxyribozymes: DNA catalysts for bioorganic chemistry. *Org. Biomol. Chem.*, **2**, 2701–2706.
4. Astell, C.R., Smith, M., Chow, M.B. and Ward, D.C. (1979) Structure of the 3' hairpin termini of four rodent parvovirus genomes: Nucleotide sequence homology at origins of DNA replication. *Cell*, **17**, 691–703.
5. Porschke, D. (1974) Thermodynamic and kinetic parameters of an oligonucleotide hairpin helix. *Biophys. Chem.*, **1**, 381–386.
6. Bonnet, G., Krichevsky, O. and Libchaber, A. (1998) Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 8602–8606.
7. Liphardt, J., Onoa, B., Smith, S.B., Tinoco, I. and Bustamante, C. (2001) Reversible unfolding of single RNA molecules by mechanical force. *Science*, **292**, 733–737.
8. Woodside, M.T., Behnke-Parks, W.M., Larizadeh, K., Travers, K., Herschlag, D. and Block, S.M. (2006) Nanomechanical measurement of the sequence dependent folding landscapes of single nucleic acid hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 6190–6195.
9. Mangeol, P. and Bockelmann, U. (2008) Interferences and crosstalk in double optical tweezers using a single laser source. *Rev. Sci. Instrum.*, **79**, 83–103.

10. Mangeol,P., Côte,D., Bizebard,T., Legrand,O. and Bockelmann,U. (2006) Probing DNA and RNA single molecules with a double optical tweezer. *Eur. Phys. J. E Soft Matter*, **19**, 311–317.
11. Cisse,I., Mangeol,P. and Bockelmann,U. (2011) DNA unzipping and force measurements with a dual optical trap. In: Peterman,EJG and Wuite,GJL (eds). *Single Molecule Analysis: Methods and Protocols*, Humana press, c/o Springer Science+Business Media, NY, pp. 45–61.
12. Petrov,A.I., Sweeney,B.A. and Leontis,N.B. (2013) Analyzing, searching and annotating recurrent RNA three-dimensional motifs. In: Klostermeier,D and Hammann,C (eds). *RNA Structure and Folding*, De Gruyter, Berlin, pp. 363–398.
13. Kramers,H.A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **7**, 284–304.
14. Bell,G.I. (1978) Models for the specific adhesion of cells to cells. *Science*, **200**, 618–627.
15. Evans,E. and Ritchie,K. (1997) Dynamic strength of molecular adhesion bonds. *Biophys. J.*, **72**, 1541–1555.
16. Dudko,O.K., Hummer,G. and Szabo,A. (2008) Theory, analysis and interpretation of single-molecule force spectroscopy experiments. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 15755–15760.
17. Manosas,M., Collin,D. and Ritort,F. (2006) Force-dependent fragility in RNA hairpins. *Phys. Rev. Lett.*, **96**, 218301–218304.
18. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
19. Rivetti,C., Walker,C. and Bustamante,C. (1998) Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibilities. *J. Mol. Biol.*, **280**, 41–59.
20. Bockelmann,U., Thomen,P., Essevez-Roulet,B., Viasnoff,V. and Heslot,F. (2002) Unzipping DNA with optical tweezers: high sequence sensitivity and force flips. *Biophys. J.*, **82**, 1537–53.
21. Hyeon,C. and Thirumalai,D. (2006) Force-unfolding and force-quench refolding of RNA hairpins. *Biophys. J.*, **90**, 3410–3427.
22. Leontis,N.B., Stombaugh,J. and Westhof,E. (2002) The non Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.*, **30**, 3497–3531.
23. Steger,G. and Giegerich,R. (2013) RNA structure predication. In: Klostermeier,D and Hammann,C (eds). *RNA Structure and Folding*, De Gruyter, Berlin, pp. 335–362.
24. Sakata,T., Hiroaki,H., Oda,Y., Tanaka,T, Ikehara,M. and Uesugi,S. (1990) Studies on the structure and stabilizing factor of the CUUCGG hairpin RNA using chemically synthesized oligonucleotides. *Nucleic Acids Res.*, **18**, 3831–3839.
25. Onoa,B., Dumont,S., Liphardt,J., Smith,S.B., Tinoco,I. and Bustamante,C. (2003) Identifying kinetic barriers to mechanical unfolding of the T. thermophila ribozyme. *Science*, **299**, 1892–1895.
26. Harlepp,S., Marchal,T., Robert,J., Leger,J.F., Xayaphoummine,A., Isambert,H. and Chatenay,D. (2003) Probing complex RNA structures by mechanical force. *Eur. Phys. J. E Soft Matter*, **12**, 605–615.
27. Mangeol,P., Bizebard,T., Chiaruttini,C., Dreyfus,M., Springer,M. and Bockelmann,U. (2011) Probing ribosomal protein-RNA interactions with an external force. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 18272–18276.
28. Lipfert,J., Skinner,G.M., Keegstra,J.M., Hengsgens,T., Jager,T., Dulin,D., Köber,M., Yu,Z., Donkers,S.P., Chou,F. *et al.* (2014) Double-stranded RNA under force and torque: similarities to and striking differences from double-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 15408–15413.

# ANNEXE Test en FRET de l'activité de DbpA

## Protocole pour le test en FRET de l'activité hélicase de DbpA en solution

- Préparer le duplexe d'oligo à séparer (dbpa9/dbpa32) à 2 $\mu$ M en tampon 75 mM KCl / 10 mM Hepes (Figure 4.20).
- Préparer l'échantillon à tester : duplexe dbpa9/dbpa32 (20 nM), oligo compétiteur dbpa11 (300 nM), ATP/Mg (2 mM), en tampon DbpA (glycérol 5% ; 50 mM KCl ; 50 mM Hepes pH 7.5 ; 5 mM MgCl<sub>2</sub>).
- Prendre un spectre de fluorescence en émission avec excitation à 500 nm (excitation de Cy3) sur la plage 520-700 nm (Figure 4.21). Sur ce spectre de référence (figure 4.21), deux pics sont attendus : celui de Cy3 en émission (~560nm), et celui de Cy5 en FRET (~660nm).
- Ajouter DbpA à 320 nM.
- Prendre un deuxième spectre de fluorescence en émission avec excitation à 500 nm (excitation de Cy3) sur la plage 520-700 nm (Figure 4.21). On attend cette fois-ci un seul pic, celui de Cy3 en émission, qui a une plus grande amplitude que le pic de référence. Le pic de FRET a disparu puisque l'oligo compétiteur dbpa11 (en excès) a remplacé l'oligo9 portant Cy5, déplacé par DbpA.

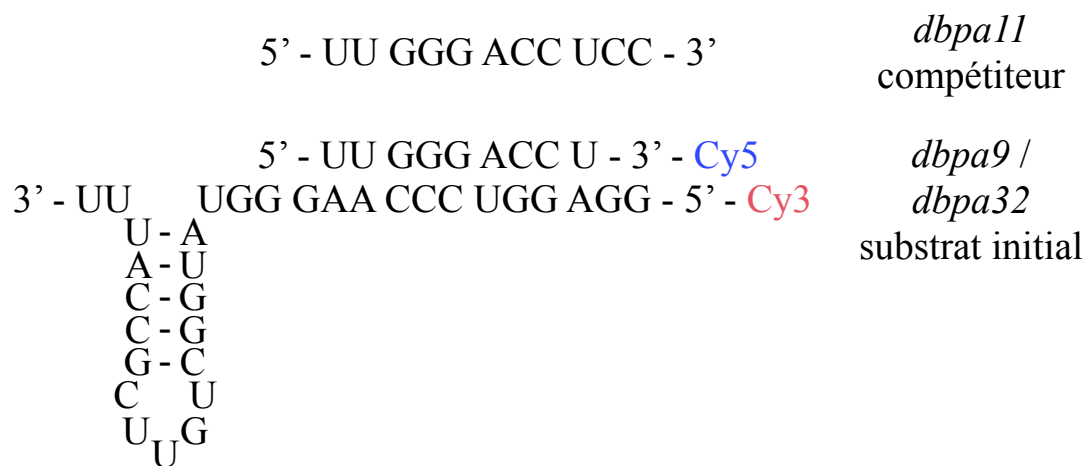


Figure 4.20 **Oligomères utilisés pour le test en FRET de l'activité de DbpA en solution.** Le substrat initial est composé du duplexe *dbpa9/dbpa32* hybridé. Chacun de ces oligomères porte un fluorophore, le substrat initial est actif en FRET. L'oligomère *dbpa11* est un compétiteur qui peut prendre la place de *dbpa9*. Si *dbpa9* est déplacé par *DbpA*, c'est *dbpa11*, placé en excès, qui viendra s'hybrider à *dbpa32* avec la plus grande probabilité.

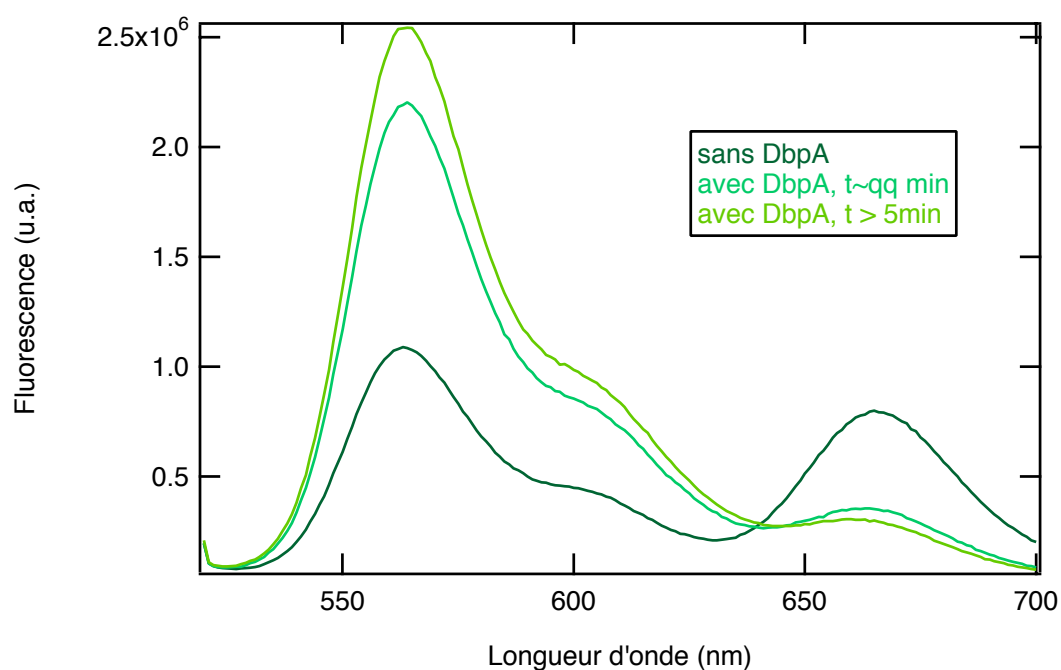


Figure 4.21 **Spectres de fluorescence en émission pour le test d'activité de DbpA en solution.** La courbe foncée correspond à la mesure de référence en l'absence de DbpA. Les deux pics correspondants à Cy3 (560 nm) et Cy5 (660 nm) sont bien définis. Les deux autres spectres ont été réalisés environ 1 et 6 minutes après l'ajout de DbpA. Le pic correspondant à Cy5 est très atténué, et celui de Cy3 a une amplitude plus que double par rapport au cas sans DbpA. La protéine est donc active.



# ANNEXE Analyse énergétique dans le cas des « bulles fondues »

## Estimation de la force du plateau de surétirement dans le cas des « bulles fondues »

Le cas du surétirement en mode de « bulles fondues » est représenté sur la figure 4.22. Contrairement au cas du peeling, les deux simples brins générés par rupture d'une paire de bases sont maintenus sous tension. La force  $F$  appliquée à la construction est donc répartie sur ces deux brins en parallèle. Les forces  $F_1$  et  $F_2$  supportées par chacun des brins vérifient la relation  $F_1 + F_2 = F$ . Si les deux simples brins sont de même nature, alors  $F_1 = F_2 = F/2$ . Dans le cas d'une construction hybride, les simples brins d'ARN et d'ADN sont soumis à une force différente. De part la géométrie du problème, la longueur  $l_{ss}$  de simple brin généré est identique pour les deux brins.

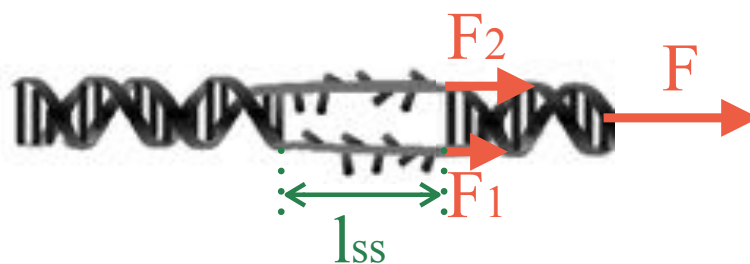


Figure 4.22 **Schema d'une « bulle fondue ».** De part et d'autre de la bulle, l'acide nucléique est sous forme double brin. Au niveau de la bulle les liaisons H entre les brins sont rompues. Deux simples brins se font donc face. Ils peuvent être de même nature (ADN ou ARN), ou bien de nature différente dans le cas d'une construction hybride. La longueur de simple brin  $l_{ss}$  est la même pour les deux brins. La force  $F$  imposée à la construction se répartit en  $F_1$  et  $F_2$  sur chacun des simples brins, avec  $F_1 + F_2 = F$ .



Les différentes composantes énergétiques utilisées pour la description détaillée au chapitre 4 sont donc les suivantes :

- $E_{ds}$  est inchangée par rapport au cas du peeling.
- $E_{ss}$  est décomposée en deux termes, correspondant à chacun des simples brins :  $E_{ss}(F) = E_{ss}(F_1) + E_{ss}(F_2)$ , avec  $F_1$  et  $F_2$  les forces appliquées sur chacun des brins. Dans le cas des molécules d'ARN ou d'ADN double brin,  $E_{ss}(F) = 2E_{ss}(F/2)$ . L'expression d' $E_{ss}(F)$  est celle donnée au chapitre 4.
- $E_b$  est inchangée par rapport au cas du peeling.
- L'expression du travail  $F\Delta L = (F_1 + F_2)\Delta L$  est inchangée. Dans le cas où les deux simples brins sont de nature différente, la valeur de  $\Delta L$  doit être adaptée. L'obtention de cette valeur est détaillée par la suite.

Comme dans le chapitre 4, la différence de longueur  $\Delta L = l_{ss} - l_{ds}$  est obtenues à partir de la distance entre les courbes théoriques d'étirement de simple et double brin d'acides nucléiques. La longueur  $l_{ss}$  est ici fixée par la relation  $F_1 + F_2 = F_{ADN} + F_{ARN} = F$ . Les courbes d'extension de simple brin ARN et de simple brin ADN imposent une valeur de  $l_{ss}$  permettant de vérifier la relation sur les forces. C'est cette valeur qui a été choisie, et  $l_{ds}$  est mesurée à la force correspondante. Les valeurs utilisées sont rapportées dans le tableau suivant.

$F$ (pN)	$F_{ADN}$ (pN)	$F_{ARN}$ (pN)	$l_{ss}$ (nm/nt)	$l_{ds}$ (nm/nt)	$\Delta L$ (nm/nt)
50	19.5	30.5	0.56	0.38	0.18

# Bibliographie

- [1] J. D. Watson et F. H. C. Crick, “Molecular structure of nucleic acids,” *Nature* **171**, 737–738 (1953). 10634.
- [2] N. B. Leontis, J. Stombaugh, et E. Westhof, “The non watson-crick base pairs and their associated isostericity matrices,” *Nucleic Acids Research* **30**, 3497–3531 (2002). URL <http://nar.oxfordjournals.org/content/30/16/3497,00499>.
- [3] Yikrazuul, “Own work. licensed under cc by-sa 3.0 via wikimedia commons,” <http://commons.wikimedia.org/wiki>.
- [4] A. Ashkin, “Acceleration and trapping of particles by radiation pressure,” *Phys. Rev. Lett.* **24**, 156–159 (1970). URL <http://link.aps.org/doi/10.1103/PhysRevLett.24.156>.
- [5] A. Ashkin, “Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime,” *Biophys J* **61**, 569–82 (1992).
- [6] B Maheu and G Gouesbet and G Grehan, “A concise presentation of the generalized Lorenz-Mie theory for arbitrary location of the scatterer in an arbitrary incident profile,” *Journal of Optics* **19**, 59 (1988). URL <http://stacks.iop.org/0150-536X/19/i=2/a=002>.
- [7] P. Mangeol et U. Bockelmann, “Interference and crosstalk in double optical tweezers using a single laser source,” *Review of Scientific Instruments* **79**, 083103 (2008). URL <http://link.aip.org/link/RSINAK/v79/i8/p083103/s1&Agg=doi>.
- [8] E. J. G. Peterman et G. J. L. Wuite, “Single molecule analysis : Methods and protocols,” (2011).
- [9] J. Wolters, “The nature of preferred hairpin structures in 16s-like rRNA variable regions,” *Nucleic Acids Research* **20**, 1843–1850 (1992). URL <http://nar.oxfordjournals.org/content/20/8/1843.abstract>.
- [10] R. R. Gutell, “Collection of small subunit (16s- and 16s-like) ribosomal RNA structures : 1994,” *Nucleic Acids Research* **22**, 3502–3507 (1994). URL <http://nar.oxfordjournals.org/content/22/17/3502.abstract>.
- [11] P. Babitzke, J. Schaak, A. V. Yakhnin, et P. C. Bevilacqua, “Role of RNA structure in transcription attenuation in *Bacillus subtilis* : the trpEDCFBA operon as a model system,” *Methods in enzymology* **371**, 392–404 (2003). URL <http://europepmc.org/abstract/MED/14712717>.

- [12] K. A. LeCuyer et D. M. Crothers, "Kinetics of an RNA conformational switch." *Proceedings of the National Academy of Sciences* **91**, 3373–3377 (1994). URL <http://www.pnas.org/content/91/8/3373.abstract>.
- [13] P. Wenter, B. Fürtig, A. Hainard, H. Schwalbe, et S. Pitsch, "Kinetics of Photoinduced RNA Refolding by Real-Time NMR Spectroscopy," *Angewandte Chemie International Edition* **44**, 2600–2603 (2005). URL <http://dx.doi.org/10.1002/anie.200462724>.
- [14] C. Höbartner et R. Micura, "Bistable Secondary Structures of Small RNAs and Their Structural Probing by Comparative Imino Proton NMR Spectroscopy," *Journal of Molecular Biology* **325**, 421–431 (2003). URL <http://www.sciencedirect.com/science/article/pii/S0022283602012433>.
- [15] P. C. Bevilacqua et J. M. Blose, "Structures, Kinetics, Thermodynamics, and Biological Functions of RNA Hairpins," *Annual Review of Physical Chemistry* **59**, 79–103 (2008). URL <http://dx.doi.org/10.1146/annurev.physchem.59.032607.093743>.
- [16] C. C. Adams et D. B. Stern, "Control of mrna stability in chloroplasts by 3' inverted repeats : effects of stem and loop mutations on degradation of psba mrna in vitro," *Nucleic Acids Research* **18**, 6003–6010 (1990). URL <http://nar.oxfordjournals.org/content/18/20/6003.abstract,00083>.
- [17] C. R. Astell, M. Smith, M. B. Chow, et D. C. Ward, "Structure of the 3' hairpin termini of four rodent parvovirus genomes : Nucleotide sequence homology at origins of DNA replication," *Cell* **17**, 691–703 (1979). URL <http://www.sciencedirect.com/science/article/pii/0092867479902769,00120>.
- [18] S. K. Silverman, "Deoxyribozymes : DNA catalysts for bioorganic chemistry," *Organic & Biomolecular Chemistry* **2**, 2701–2706 (2004). URL <http://pubs.rsc.org/en/content/articlehtml/2004/ob/b411910j>.
- [19] M. Bercy et U. Bockelmann, "Hairpins under tension : RNA versus DNA," *Nucleic Acids Research* (2015). URL <http://nar.oxfordjournals.org/content/early/2015/08/31/nar.gkv860.abstract,00000>.
- [20] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, et S. M. Block, "Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins," *Proceedings of the National Academy of Sciences* **103**, 6190–6195 (2006). URL <http://www.pnas.org/content/103/16/6190.short>.
- [21] T. Odijk, "Stiff chains and filaments under tension," *Macromolecules* **28**, 7016–7018 (1995). URL <http://pubs.acs.org/doi/abs/10.1021/ma00124a044,00306>.
- [22] C. Bustamante, J. Marko, E. Siggia, et S. Smith, "Entropic elasticity of lambda-phage DNA," *Science* **265**, 1599–1600 (1994). URL <http://www.sciencemag.org/content/265/5178/1599.short,00002>.
- [23] P. Gross, N. Laurens, L. B. Oddershede, U. Bockelmann, E. J. G. Peterman, et G. J. L. Wuite, "Quantifying how DNA stretches, melts and changes twist under tension," *Nature Physics* **7**, 731–736 (2011). URL <http://www.nature.com/doifinder/10.1038/nphys2002>.

- [24] A. Bosco, J. Camunas-Soler, et F. Ritort, "Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions," *Nucleic Acids Research* **42**, 2064–2074 (2014). URL <http://nar.oxfordjournals.org/content/42/3/2064>.
- [25] C. V. Bizarro, A. Alemany, et F. Ritort, "Non-specific binding of Na<sup>+</sup> and Mg<sup>2+</sup> to RNA determined by force spectroscopy methods," *Nucleic Acids Research* **40**, 6922–6935 (2012). URL <http://nar.oxfordjournals.org/content/40/14/6922.abstract>.
- [26] D. B. McIntosh, G. Duggan, Q. Gouil, et O. A. Saleh, "Sequence-Dependent Elasticity and Electrostatics of Single-Stranded DNA : Signatures of Base-Stacking," *Biophysical Journal* **106**, 659–666 (2014). URL [http://www.cell.com/biophysj/abstract/S0006-3495\(13\)05810-4,00000](http://www.cell.com/biophysj/abstract/S0006-3495(13)05810-4,00000).
- [27] G. Bell, "Models for the specific adhesion of cells to cells," *Science* **200**, 618–627 (1978). URL <http://www.sciencemag.org/content/200/4342/618.abstract>.
- [28] H. Kramers, "Brownian motion in a field of force and the diffusion model of chemical reactions," *Physica* **7**, 284–304 (1940). URL <http://www.sciencedirect.com/science/article/pii/S0031891440900982>.
- [29] E. Evans et K. Ritchie, "Dynamic strength of molecular adhesion bonds," *Biophysical Journal* **72**, 1541–1555 (1997). URL [http://dx.doi.org/10.1016/S0006-3495\(97\)78802-7](http://dx.doi.org/10.1016/S0006-3495(97)78802-7).
- [30] O. K. Dudko, G. Hummer, et A. Szabo, "Theory, analysis, and interpretation of single-molecule force spectroscopy experiments," *Proceedings of the National Academy of Sciences* **105**, 15755–15760 (2008). URL <http://www.pnas.org/content/105/41/15755>.
- [31] M. Manosas, D. Collin, et F. Ritort, "Force-Dependent Fragility in RNA Hairpins," *Physical Review Letters* **96** (2006). URL <http://link.aps.org/doi/10.1103/PhysRevLett.96.218301>.
- [32] O. Dudko, G. Hummer, et A. Szabo, "Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments," *Physical Review Letters* **96** (2006). URL <http://link.aps.org/doi/10.1103/PhysRevLett.96.108101,00431>.
- [33] C. A. Pierse et O. K. Dudko, "Kinetics and energetics of biomolecular folding and binding," *Biophysical Journal* **105**, L19–L22 (2013). URL <http://www.sciencedirect.com/science/article/pii/S0006349513010710,00001>.
- [34] C. Rivetti, C. Walker, et C. Bustamante, "Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility," *Journal of Molecular Biology* **280**, 41–59 (1998). URL <http://www.sciencedirect.com/science/article/pii/S0022283698918307>.
- [35] U. Bockelmann, P. Thomen, B. Essevez-Roulet, V. Viasnoff, et F. Heslot, "Unzipping dna with optical tweezers : High sequence sensitivity and force flips," *Biophysical journal* **82**, 1537–1553 (2002). URL <http://linkinghub.elsevier.com/retrieve/pii/S0006349502755069>.

- [36] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, et C. Bustamante, “Reversible Unfolding of Single RNA Molecules by Mechanical Force,” *Science* **292**, 733–737 (2001). URL <http://www.sciencemag.org/content/292/5517/733>.
- [37] C. Hyeon et D. Thirumalai, “Forced-Unfolding and Force-Quench Refolding of RNA Hairpins,” *Biophysical Journal* **90**, 3410–3427 (2006). URL <http://linkinghub.elsevier.com/retrieve/pii/S0006349506725255>.
- [38] G. Bonnet, O. Krichevsky, et A. Libchaber, “Kinetics of conformational fluctuations in DNA hairpin-loops,” *Proceedings of the National Academy of Sciences* **95**, 8602–8606 (1998). URL <http://www.pnas.org/content/95/15/8602.short>.
- [39] P. Mangeol, T. Bizebard, C. Chiaruttini, M. Dreyfus, M. Springer, et U. Bockelmann, “Probing ribosomal protein-rna interactions with an external force,” *Proceedings of the National Academy of Sciences* **108**, 18272–18276 (2011). URL <http://www.pnas.org/content/108/45/18272.short>.
- [40] D. Klostermeier et C. Hammann, *RNA structure and folding : biophysical techniques and prediction methods* (De Gruyter, Berlin [etc.], 2013). URL <http://www.degruyter.com/view/product/182520>.
- [41] T. Sakata, H. Hiroaki, Y. Oda, T. Tanaka, M. Ikehara, et S. Uesugi, “Studies on the structure and stabilizing factor of the CUUCGG hairpin RNA using chemically synthesized oligonucleotides.” *Nucleic Acids Research* **18**, 3831–3839 (1990). URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC331083/>.
- [42] B. Onoa, S. Dumont, J. Liphardt, S. B. Smith, I. Tinoco, et C. Bustamante, “Identifying Kinetic Barriers to Mechanical Unfolding of the T. thermophila Ribozyme,” *Science* **299**, 1892–1895 (2003). URL <http://www.sciencemag.org/content/299/5614/1892.abstract>.
- [43] S. Harlepp, T. Marchal, J. Robert, J.-F. Léger, A. Xayaphoummine, H. Isambert, et D. Chatenay, “Probing complex RNA structures by mechanical force,” *The European Physical Journal E* **12**, 605–615 (2003). URL <http://dx.doi.org/10.1140/epje/e2004-00033-4>.
- [44] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F.-C. Chou, R. Das, et N. H. Dekker, “Double-stranded RNA under force and torque : Similarities to and striking differences from double-stranded DNA,” *Proceedings of the National Academy of Sciences* **111**, 15408–15413 (2014). URL <http://www.pnas.org/content/111/43/15408>, 00000 PMID : 25313077.
- [45] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1 (Elsevier, 1992).
- [46] U. Bockelmann et V. Viasnoff, “Theoretical Study of Sequence-Dependent Nanopore Unzipping of DNA,” *Biophysical Journal* **94**, 2716–2724 (2008). URL <http://www.sciencedirect.com/science/article/pii/S0006349508705241>.
- [47] N. K. Tanner et P. Linder, “Dexd/h box rna helicases : from generic motors to specific dissociation functions,” *Molecular Cell* **8**, 251–62 (2001).

- [48] P. Linder et E. Jankowsky, "From unwinding to clamping - the dead box rna helicase family," *Nature reviews. Molecular cell biology* **12**, 505–16 (2011).
- [49] W.-Y. Tarn et T.-H. Chang, "The current understanding of ded1p/ddx3 homologs from yeast to human," *RNA biology* **6**, 17–20 (2009).
- [50] J. Banroques, M. Doère, M. Dreyfus, P. Linder, et N. K. Tanner, "Motif iii in superfamily 2 "helicases" helps convert the binding energy of atp into a high-affinity rna binding site in the yeast dead-box protein ded1," *Journal of molecular biology* **396**, 949–66 (2010).
- [51] T. Bizebard, I. Ferlenghi, I. Iost, et M. Dreyfus, "Studies on three e. coli dead-box helicases point to an unwinding mechanism different from that of model dna helicases," *Biochemistry* **43**, 7857–7866 (2004). URL <http://pubs.acs.org/doi/abs/10.1021/bi049852s>, pMID : 15196029.
- [52] L. M. S. Elles et O. C. Uhlenbeck, "Mutation of the arginine finger in the active site of escherichia coli dbpa abolishes atpase and helicase activity and confers a dominant slow growth phenotype," *Nucleic Acids Research* **36**, 41–50 (2008). URL <http://nar.oxfordjournals.org/content/36/1/41.abstract>.
- [53] L. M. Sharpe Elles, M. T. Sykes, J. R. Williamson, et O. C. Uhlenbeck, "A dominant negative mutant of the e. coli rna helicase dbpa blocks assembly of the 50s ribosomal subunit," *Nucleic Acids Research* (2009). URL <http://nar.oxfordjournals.org/content/early/2009/09/04/nar.gkp711.abstract>.
- [54] J. W. Hardin, Y. X. Hu, et D. B. McKay, "Structure of the RNA binding domain of a DEAD-box helicase bound to its ribosomal RNA target reveals a novel mode of recognition by an RNA recognition motif," *Journal of molecular biology* **402**, 412–427 (2010). URL <http://www.sciencedirect.com/science/article/pii/S0022283610008053>.
- [55] C. M. Diges et O. C. Uhlenbeck, "Escherichia coli DbpA is an RNA helicase that requires hairpin 92 of 23s rRNA," *The EMBO journal* **20**, 5503–5512 (2001). URL <http://www.nature.com/emboj/journal/v20/n19/abs/7594045a.html>.
- [56] A. D. Mehta, M. Rief, J. A. Spudich, D. A. Smith, et R. M. Simmons, "Single-Molecule Biomechanics with Optical Methods," *Science* **283**, 1689–1695 (1999). URL <http://www.sciencemag.org/content/283/5408/1689.abstract>.
- [57] T. Ha, I. Rasnik, W. Cheng, H. P. Babcock, G. H. Gauss, T. M. Lohman, et S. Chu, "Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase," *Nature* **419**, 638–641 (2002). URL <http://dx.doi.org/10.1038/nature01083>.
- [58] D. S. Johnson, L. Bai, B. Y. Smith, S. S. Patel, et M. D. Wang, "Single-molecule studies reveal dynamics of dna unwinding by the ring-shaped t7 helicase," *Cell* **129**, 1299 – 1309 (2007). URL <http://www.sciencedirect.com/science/article/pii/S0092867407005843>.
- [59] X. Qu, L. Lancaster, H. F. Noller, C. Bustamante, et I. Tinoco, "Ribosomal protein S1 unwinds double-stranded RNA in multiple steps," *Proceedings of the National Academy of Sciences* **109**, 14458–14463 (2012). URL <http://www.pnas.org/content/109/36/14458.abstract>.

- [60] D. J. Koslover, F. M. Fazal, R. A. Mooney, R. Landick, et S. M. Block, “Binding and Translocation of Termination Factor Rho Studied at the Single-Molecule Level,” *Journal of Molecular Biology* **423**, 664–676 (2012). URL <http://www.sciencedirect.com/science/article/pii/S0022283612006286>.
- [61] M. J. Comstock, K. D. Whitley, H. Jia, J. Sokoloski, T. M. Lohman, T. Ha, et Y. R. Chemla, “Direct observation of structure-function relationship in a nucleic acid-processing enzyme,” *Science* **348**, 352–354 (2015). URL <http://www.sciencemag.org/content/348/6232/352.abstract>.
- [62] A. Henn, O. Medalia, S.-P. Shi, M. Steinberg, F. Franceschi, et I. Sagi, “Visualization of unwinding activity of duplex RNA by DbpA, a DEAD box helicase, at single-molecule resolution by atomic force microscopy,” *Proceedings of the National Academy of Sciences* **98**, 5007–5012 (2001). URL <http://www.pnas.org/content/98/9/5007>.
- [63] C. Pan, J. P. Potratz, B. Cannon, Z. B. Simpson, J. L. Ziehr, P. Tijerina, et R. Russell, “DEAD-Box Helicase Proteins Disrupt RNA Tertiary Structure Through Helix Capture,” *PLoS Biol* **12**, e1001981 (2014). URL <http://dx.doi.org/10.1371/journal.pbio.1001981>.
- [64] C. García-García, K. L. Frieda, K. Feoktistova, C. S. Fraser, et S. M. Block, “Factor-dependent processivity in human eIF4a DEAD-box helicase,” *Science* **348**, 1486–1488 (2015). URL <http://www.sciencemag.org/content/348/6242/1486.short>, 00000.
- [65] C. A. Tsu et O. C. Uhlenbeck, “Kinetic Analysis of the RNA-Dependent Adenosine-triphosphatase Activity of DbpA, an Escherichia coli DEAD Protein Specific for 23s Ribosomal RNA,” *Biochemistry* **37**, 16989–16996 (1998). URL <http://pubs.acs.org/doi/abs/10.1021/bi981837y>.
- [66] A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, et D. Bensimon, “Alignment and Sensitive Detection of DNA by a Moving Interface,” *Science* **265**, 2096–2098 (1994). URL <http://www.jstor.org/stable/2885861>, articleType : research-article / Issue Title : Genome Issue / Full publication date : Sep. 30, 1994 / Copyright © 1994 American Association for the Advancement of Science.
- [67] D. Bensimon, A. J. Simon, V. Croquette, et A. Bensimon, “Stretching DNA with a Receding Meniscus : Experiments and Models,” *Physical Review Letters* **74**, 4754–4757 (1995). URL <http://link.aps.org/doi/10.1103/PhysRevLett.74.4754>.
- [68] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, et F. Caron, “DNA : An Extensible Molecule,” *Science* **271**, 792–794 (1996). URL <http://www.sciencemag.org/content/271/5250/792>.
- [69] S. B. Smith, Y. Cui, et C. Bustamante, “Overstretching B-DNA : The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules,” *Science* **271**, 795–799 (1996). URL <http://www.sciencemag.org/content/271/5250/795>.
- [70] J. v. Mameren, P. Gross, G. Farge, P. Hooijman, M. Modesti, M. Falkenberg, G. J. L. Wuite, et E. J. G. Peterman, “Unraveling the structure of DNA during overstretching

- by using multicolor, single-molecule fluorescence imaging,” *Proceedings of the National Academy of Sciences* **106**, 18231–18236 (2009). URL <http://www.pnas.org/content/106/43/18231>.
- [71] M. C. Williams, I. Rouzina, et M. J. McCauley, “Peeling back the mystery of DNA overstretching,” *Proceedings of the National Academy of Sciences* **106**, 18047–18048 (2009). URL <http://www.pnas.org/content/106/43/18047>.
- [72] D. H. Paik et T. T. Perkins, “Overstretching DNA at 65 pN Does Not Require Peeling from Free Ends or Nicks,” *Journal of the American Chemical Society* **133**, 3219–3221 (2011). URL <http://pubs.acs.org/doi/abs/10.1021/ja108952v>.
- [73] G. A. King, P. Gross, U. Bockelmann, M. Modesti, G. J. L. Wuite, et E. J. G. Peterman, “Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy,” *Proceedings of the National Academy of Sciences* (2013). URL <http://www.pnas.org/content/early/2013/02/13/1213676110>.
- [74] U. Bockelmann, B. Essevaz-Roulet, et F. Heslot, “Molecular Stick-Slip Motion Revealed by Opening DNA with Piconewton Forces,” *Physical Review Letters* **79**, 4489–4492 (1997). URL <http://link.aps.org/doi/10.1103/PhysRevLett.79.4489>.
- [75] J. Gore, Z. Bryant, M. Nöllmann, M. U. Le, N. R. Cozzarelli, et C. Bustamante, “DNA overwinds when stretched,” *Nature* **442**, 836–839 (2006). URL <http://www.nature.com/nature/journal/v442/n7104/abs/nature04974.html>, 00161.
- [76] M. D. Wang, H. Yin, R. Landick, J. Gelles, et S. M. Block, “Stretching DNA with optical tweezers,” *Biophysical Journal* **72**, 1335–1346 (1997). URL <http://www.sciencedirect.com/science/article/pii/S0006349597787800>, 00870.
- [77] E. Herrero-Galán, M. E. Fuentes-Perez, C. Carrasco, J. M. Valpuesta, J. L. Carrascosa, F. Moreno-Herrero, et J. R. Arias-Gonzalez, “Mechanical Identities of RNA and DNA Double Helices Unveiled at the Single-Molecule Level,” *Journal of the American Chemical Society* **135**, 122–131 (2013). URL <http://pubs.acs.org/doi/abs/10.1021/ja3054755>.
- [78] D. Klein, P. Moore, et T. Steitz, “The Roles of Ribosomal Proteins in the Structure Assembly, and Evolution of the Large Ribosomal Subunit,” *Journal of Molecular Biology* **340**, 141–177 (2004). URL <http://linkinghub.elsevier.com/retrieve/pii/S0022283604003845>.
- [79] K. M. Kosikov, A. A. Gorin, V. B. Zhurkin, et W. K. Olson, “DNA stretching and compression : large-scale simulations of double helical structures,” *Journal of Molecular Biology* **289**, 1301–1326 (1999). URL <http://www.sciencedirect.com/science/article/pii/S0022283699927985>, 00000.
- [80] T. Xia, J. SantaLucia, M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox, et D. H. Turner, “Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs,” *Biochemistry* **37**, 14719–14735 (1998). URL <http://pubs.acs.org/doi/abs/10.1021/bi9809425>, 00000.



- [81] J. SantaLucia, "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics," *Proceedings of the National Academy of Sciences* **95**, 1460–1465 (1998). URL <http://www.pnas.org/content/95/4/1460.short>, 02277.
- [82] N. Sugimoto, S.-i. Nakano, M. Katoh, A. Matsumura, H. Nakamuta, T. Ohmichi, M. Yoneyama, et M. Sasaki, "Thermodynamic parameters to predict stability of rna/dna hybrid duplexes," *Biochemistry* **34**, 11211–11216 (1995). URL <http://pubs.acs.org/doi/abs/10.1021/bi00035a029>, 00535.
- [83] R. Owczarzy, B. G. Moreira, Y. You, M. A. Behlke, et J. A. Walder, "Predicting Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations," *Biochemistry* **47**, 5336–5353 (2008). URL <http://dx.doi.org/10.1021/bi702363u>.
- [84] J. Casey et N. Davidson, "Rates of formation and thermal stabilities of RNA :DNA and DNA :DNA duplexes at high concentrations of formamide," *Nucleic Acids Research* **4**, 1539–1552 (1977).
- [85] R. D. Blake et S. G. Delcourt, "Thermodynamic effects of formamide on DNA stability," *Nucleic Acids Research* **24**, 2095–2103 (1996).
- [86] K. Nierhaus, "The assembly of prokaryotic ribosomes," *Biochimie* **73**, 739–755 (1991). URL <http://www.sciencedirect.com/science/article/pii/0300908491900545>, 00209.
- [87] K. Svoboda et S. M. Block, "Biological Applications of Optical Forces," *Annual Review of Biophysics and Biomolecular Structure* **23**, 247–285 (1994). URL <http://dx.doi.org/10.1146/annurev.bb.23.060194.001335>, 01640.
- [88] G. A. Mackie et J. L. Genereaux, "The Role of RNA Structure in Determining RNase E-dependent Cleavage Sites in the mRNA for Ribosomal Protein S20 in vitro," *Journal of Molecular Biology* **234**, 998–1012 (1993). URL <http://www.sciencedirect.com/science/article/pii/S0022283683716542>.
- [89] U. Bockelmann, B. Essevaz-Roulet, et F. Heslot, "Dna strand separation studied by single molecule force measurements," *Physical Review E* **58**, 2386–2394 (1998). URL <http://link.aps.org/doi/10.1103/PhysRevE.58.2386>.
- [90] P. Mangeol, D. Côte, T. Bizebard, O. Legrand, et U. Bockelmann, "Probing DNA and RNA single molecules with a double optical tweezer," *The European Physical Journal E* **19**, 311–317 (2006). URL <http://www.springerlink.com/index/10.1140/epje/i2005-10060-4>.
- [91] D. Pörschke, "Thermodynamic and kinetic parameters of an oligonucleotide hairpin helix," *Biophysical Chemistry* **1**, 381–386 (1974). URL <http://www.sciencedirect.com/science/article/pii/0301462274850088>.







---

**Sujet : Les structures secondaires dans l'ARN :  
une étude par mesure de forces sur molécules uniques**

---

**Résumé :** L'ARN s'est longtemps vu attribuer un simple rôle de transmission entre l'ADN, garant de l'information génétique, et les protéines, assurant les fonctions et donc la survie cellulaire. Ce n'est qu'avec les découvertes des ARNs de transfert dans les années 70, puis des ribozymes dans les années 80, qu'il a été réalisé que l'ARN pouvait assurer ces deux rôles : l'information génétique est stockée dans sa séquence linéaire, et l'adoption de structures tridimensionnelles complexes rend possible une activité catalytique.

Depuis, de nouvelles fonctions de l'ARN n'ont cessé d'être découvertes, à tous les niveaux de régulation de l'expression génique entre autres. La majorité de ces fonctions repose sur la structuration tridimensionnelle d'ARNs simple brin.

Dans ce travail, différents aspects de la structuration de l'ARN sont abordés, toujours en utilisant la technique de mesure de forces sur molécules uniques par piégeage optique. Dans un premier temps, une étude comparative d'une structure secondaire modèle, le hairpin dans ses formes ARN et ADN, a été réalisée. La question de l'interaction d'une structure secondaire avec une protéine hélicase (DbpA) a ensuite été abordée. Enfin, dans le cadre plus général d'une étude sur l'assemblage du ribosome, nous avons débuté le développement d'une nouvelle méthode d'analyse des structures secondaires. Cette méthode repose sur le surétirement d'un hybride ARN ribosomique / ADN.

**Mots clés :** molécule unique, pièges optiques, ARN, ADN, hairpins, hélicases, surétirement

---

**Subject : RNA secondary structure:  
a single molecules force measurements study**

---

**Résumé :** Traditionally, RNA has been considered as a mere intermediate between DNA, keeper of the genetic information, and proteins, which assume cells self-sustenance. With the discoveries of the transfert RNA in the 70s, and of the ribozymes in the 80s, RNA took on both roles: it can store information in its linear sequence, and tridimensional structuration enables catalytic functions.

Since then, numerous roles devoted to RNA have been discovered, particularly for gene expression regulation. Most of these functions rely on tridimensional structuration of single stranded RNA.

In this work, we used an optical tweezers setup to study several aspects of RNA structuration by single molecule force measurement. In a first part, we compared the dynamic behaviour of a model secondary structure made of either RNA or DNA, the hairpin. Then we considered the interaction of a secondary structure with a protein, the RNA helicase DbpA. Finally, within a wider study of ribosome assembly, we worked on the development of a new method to study tridimensional structuration. This method relies on the overstretching of a hybrid ribosomal RNA / DNA molecule.

---

**Keywords :** single molecule, optical tweezers, RNA, DNA, hairpins, helicases, overstretching