



HAL
open science

Caractérisation d'aptamères par électrophorèse capillaire couplée au séquençage haut-débit Illumina

Audrey Ric

► **To cite this version:**

Audrey Ric. Caractérisation d'aptamères par électrophorèse capillaire couplée au séquençage haut-débit Illumina. Chimie organique. Université Paul Sabatier - Toulouse III, 2017. Français. NNT : 2017TOU30388 . tel-02301523

HAL Id: tel-02301523

<https://theses.hal.science/tel-02301523v1>

Submitted on 30 Sep 2019

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.



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Si vous êtes en cotutelle internationale, remplissez ce champs en notant : Cotutelle internationale avec "nom de l'établissement", sinon effacer ce texte pour qu'il n'apparaisse pas à l'impression

Présentée et soutenue par :

Audrey Marie Amélie RIC

le 29 Septembre 2017

Titre :

Caractérisation d'aptamères par électrophorèse capillaire couplée au séquençage haut-débit Illumina.

École doctorale et discipline ou spécialité :

ED SDM : Chimie, Biologie, Santé - CO 042

Unité de recherche :

IMRCP/IPBS

Directeur/trice(s) de Thèse :

M. Couderc François

M. Paquereau Laurent

Jury :

M. Peyrin Eric - Université Grenoble Alpes - Rapporteur

Mme Nehmé Reine - Université d'Orléans - Rapporteur

Mme Gilard Véronique - Université de Toulouse - Examineur

Mme Boutonnet Audrey - Picometrics Technologies - Invitée

M. Ecochard Vincent - Université de Toulouse - Invité

A ma Mamie,

“ En vérité, le chemin importe peu, la volonté d'arriver suffit à tout.”
Albert Camus

“ Ma revendication en tant que femme c'est que ma différence soit prise en compte, que je ne sois pas contrainte de m'adapter au modèle masculin. ”
Simone Veil

Remerciements :

Les travaux faisant l'objet de cette thèse ont été en partie réalisés au sein de l'entreprise Picometrics Technologies dont M. Jean Charles Garcia en est le directeur. Je tiens à le remercier ainsi que M. Frédéric Ginot de m'avoir accueilli au sein de leurs locaux.

Je remercie Monsieur Eric Peyrin et Madame Reine Nehmé d'avoir accepté d'être les rapporteurs de cette thèse et d'avoir pris le temps de corriger mon manuscrit. Je remercie également Madame Véronique Gilard d'avoir bien voulu présider le jury.

Je remercie chaleureusement toutes les personnes qui m'ont aidé pendant l'élaboration de ma thèse et notamment mon directeur de thèse M. François Couderc, pour son intérêt et son soutien, sa grande disponibilité et ses nombreux conseils durant la rédaction de ma thèse.

Je tiens également à exprimer toute ma reconnaissance envers Madame Audrey Boutonnet pour sa gentillesse, sa constante disponibilité et son aide au cours de mon travail.

Je remercie particulièrement Messieurs Vincent Ecochard et Laurent Paquereau du laboratoire de l'IPBS pour leur implication au sein de ce projet.

Je souhaite particulièrement remercier Monsieur Jason Iacovoni de l'I2MC, pour sa grande disponibilité et pour tout le travail effectué, mais surtout pour avoir accepté de s'investir dans ce projet et analyser les données Illumina.

Je remercie Monsieur Jean José Maoret et Monsieur Frédéric Martins pour m'avoir accueilli au sein de leur laboratoire à l'I2MC, afin de résoudre les problèmes techniques liés à la PCR.

Je remercie également Madame Laure Buisson pour son aide précieuse sur le séquençage haut-débit Illumina. Merci d'avoir pris le temps de m'expliquer, répondre à mes questions et surtout corriger la partie Illumina de ce manuscrit.

Je souhaite remercier toute l'équipe de Picometrics Technologies, l'IPBS et les IMRCP, qui m'ont permis d'effectuer cette thèse dans les meilleures conditions possibles.

J'adresse également mes remerciements à mes amis et ma famille pour m'avoir soutenu durant ces 3 années.

Mes plus profonds remerciements vont à mes parents. Tout au long de mon cursus, ils m'ont toujours soutenu, encouragé et aidé. Ils ont su me donner toutes les chances et les meilleures conditions pour réussir. Qu'ils trouvent, dans la réalisation de ce travail, l'aboutissement de leurs efforts ainsi que l'expression de ma plus affectueuse gratitude et reconnaissance.

TABLE DES MATIERES

Introduction :	2
1. Les aptamères : structure et propriétés	6
1.1 L'architecture des aptamères	6
1.1.1 Structure primaire	6
1.1.2 Structure secondaire et tertiaire :.....	6
1.2 Etude de la structure en G-quadruplex de l'aptamère anti-thrombine	9
1.2.1 Définition des G-quadruplex :	9
1.2.2 Topologie des G-quadruplex :	11
1.2.3 Le rôle des ions sur la stabilité des G-quadruplex :.....	13
1.2.4 Etudes des aptamères anti-thrombine :.....	14
2 Modes de sélection des aptamères	15
2.1 SELEX	15
2.1.1 Principe général	15
2.1.2 Modification de la technologie SELEX	17
2.2 L'électrophorèse capillaire et le SELEX (CE-SELEX) :.....	26
2.2.1 Principe général de l'électrophorèse capillaire :.....	27
2.2.2 Différentes méthodes d'électrophorèse capillaire :	29
2.2.3 Applications de la CE-SELEX :.....	35
2.3 L'électrophorèse capillaire et le non-SELEX :.....	36
2.4 Conclusion :	39
3 Techniques de mesure de la constante de dissociation K_D à partir de l'électrophorèse capillaire :.....	40
3.1 Techniques de séparation :	40
3.1.1 La dialyse :	41
3.1.2 L'ultrafiltration ou filter binding.....	43
3.1.3 L'électrophorèse sur gel	44
3.1.4 La CLHP (chromatographie liquide haute performance).....	45
3.1.5 Anisotropie de fluorescence	46
3.1.6 Absorption UV-visible.....	48
3.1.7 Dichroïsme circulaire (DC).....	48
3.1.8 Résonance plasmonique de surface (SPR)	51
3.1.9 Titration calorimétrique isotherme (ITC)	53
3.1.10 Thermophorèse à micro-échelle (MST).....	54
3.1.11 L'électrophorèse capillaire :.....	55
4 Optimisation de l'amplification par PCR :.....	64
4.1 Après la sélection par électrophorèse capillaire, la PCR :	64
4.2 La PCR en temps réel :	67
4.3 L'émulsion PCR :.....	68
5 Next Generation sequencing (NGS) pour la sélection d'aptamères :	71
5.1 Principe :.....	71

5.2	Le pyroséquençage (Roche)	71
5.3	Principe du séquençage haut-débit Illumina :	72
5.4	Le séquençage par ligation :	74
5.5	Application du séquençage haut débit :	74
6.	Conclusion :	76
Chapitre 1 : L'électrophorèse capillaire et l'utilisation de la double détection UV- LEDIF pour sélectionner, caractériser les cibles, les aptamères et les complexes aptamères-cibles : ... 78		
1	Influence de l'utilisation du tampon de migration Tris Acétate sur la dégradation de l'ADN.....	78
1.1	Introduction.....	78
1.2	Le tampon TRIS, un piège pour la sélection d'aptamères ?.....	82
1.2.1	Introduction :	82
1.2.2	Article :	83
2	Mesure des paramètres de liaisons entre la thrombine et un aptamère par électrophorèse capillaire :	92
2.1	Caractérisation de la thrombine par électrophorèse capillaire	92
2.1.1	Capillaire en silice.....	92
2.1.2	Capillaire recouvert d'un revêtement en PVA (Polyvinylalcohol)	93
2.2	Caractérisation de l'aptamère T29 et du complexe T29.Thrombine par électrophorèse capillaire.....	94
2.2.1	Avec un capillaire en PVA.....	94
2.2.2	Avec un capillaire en silice.....	96
2.3	Détermination de la constante de dissociation du complexe T29. Thrombine.....	98
2.3.1	Par électrophorèse capillaire	98
2.3.2	Par filter binding.....	100
2.3.3	Conclusion	101
Chapitre 2 : Est-on capable d'identifier un aptamère G-quartet en CE- Illumina ?103		
1	Validation / Mise en place de la collecte à partir de l'exemple de 3 sucres.....	103
1.1	Injection et séparation de trois sucres : exemple	103
1.2	Méthode de collecte de trois sucres par électrophorèse capillaire :	104
1.3	Détermination du facteur de dilution et du rendement de collecte	105
2	Collecte d'aptamères :	105
2.1	Optimisation du diamètre du capillaire :	105
2.2	Injection de la banque d'ADN, de l'aptamère AT29B et de la thrombine :.....	106
2.2.1	La banque d'ADN (N=30) :	106
2.2.2	Injection de l'aptamère AT29B :	107
2.2.3	Injection de la thrombine :	107
2.2.4	Réalisation du complexe Banque d'ADN + thrombine + AT29B et détermination des fenêtres de collectes :	108

3	<i>Le contrôle qualité à l'issu de la collecte des échantillons par électrophorèse capillaire :</i>	109
1	<i>Principe de l'analyse par Illumina :</i>	112
1.1	Principales définitions utilisées en bioinformatique :	112
1.2	Principales étapes de l'analyse bioinformatique	112
1.3	Evaluation de la qualité des données grâce à Fastqc	113
1.4	Nettoyage des reads – filtrage des séquences d'adaptateurs	114
1.5	L'alignement des séquences	116
1.6	Profondeur de lecture = le nombre de fois où chaque nucléotide est lu	116
2	<i>Utilisation de la PCR - Illumina pour séquencer le T29</i>	116
2.1	Réalisation de la gamme étalon	116
2.2	Caractérisation du nombre de séquences de T29 obtenues – étude de la réponse de l'Illumina	117
2.3	Reproductibilité	118
2.4	Publication	122
	References	139
2.5	Précision sur l'analyse des nucléotides de T29 dans le brin anti-sens :	142
3	<i>La contamination des blancs :</i>	142
4	<i>Mesure de l'enrichissement de l'aptamère T29 par électrophorèse capillaire couplée au séquençage Illumina</i>	143
5	<i>Conclusion :</i>	144
	<i>Conclusion</i>	146
	<i>Annexes :</i>	148
1	<i>Instrumentation</i>	148
1.1	Le détecteur de fluorescence	148
1.2	La double détection absorbance UV - fluorescence	148
2	<i>Réactifs et produits utilisés :</i>	149
2.1	Réactifs :	149
2.2	Oligonucléotides :	150
2.3	Tampons :	150
3	<i>Capillaires :</i>	151
	Capillaire en Silice :	151
3.1	Capillaire en Poly-vinyl-alcool : PVA	151

3.2	Conditionnement du capillaire :	151
4	Principe de la PCR en temps réel :	151
4.1	Conditions opératoires	151
4.2	Utilisation de la PCR en temps réel pour amplifier l'aptamère T29 :	154
	Réalisation d'une gamme étalon de l'aptamère AT29B	154
4.2.1	Répétabilité de la qPCR	155
4.2.2	Diminution du nombre de cycles qPCR.	157
4.2.3	Effet des amorces – diminution de la concentration des amorces :	159
4.2.4	Augmentation de la température d'hybridation.....	159
4.3	Changement des constituants :	161
4.3.1	Changement des amorces :	161
4.3.2	Changement de la matrice : utilisation d'un aptamère sans G-quartet.....	162
4.4	Conclusion	163
5	Principe de la PCR conventionnelle :	164
6	Protocole d'amplification et de purification des molécules d'ADN	165
6.1	L'électroélution	165
6.2	Kits de purifications :	167
7	Quantification post-PCR des aptamères amplifiés	167
7.1	Le nanodrop.....	167
7.2	Par électrophorèse sur gel d'agarose	167
8	Représentation de la méthode globale de sélection par électrophorèse capillaire, amplification par PCR, purification des aptamères	168
9	Principe du séquençage haut-débit Illumina.....	168
9.1	Préparation de l'ADN : préparation de bibliothèques.	169
9.2	Amplification clonale : création de clusters	170
9.3	Séquençage.....	171
9.4	Analyse bioinformatique	172
	Bibliographie :	174

Introduction :

Introduction :

Les aptamères nucléotidiques sont des oligomères d'ADN ou d'ARN simple brin qui en se repliant sous formes de structures secondaires et tertiaires peuvent avoir des interactions fortes et spécifiques avec des cibles particulières (petites molécules, acides aminés naturels ou synthétiques, peptides, macromolécules mais également des cellules) (1). Cette propriété intrinsèque en fait des molécules efficaces pour reconnaître leur cible avec affinité et spécificité. A ce titre, ils sont donc considérés comme une alternative intéressante aux anticorps.

En 1990, Tuerk and Gold ont décrit une nouvelle méthode de sélection et d'amplification *in vitro* : la méthode SELEX (Systematic Evolution of Ligands by EXponential Enrichment). Ils ont utilisé une banque aléatoire afin de sélectionner des oligonucléotides d'ARN pouvant se lier spécifiquement et sélectivement à une cible (ADN T4 gp43 polymérase) (2). La même année, Ellington et Szostak ont mis au point une méthode de sélection similaire pour isoler des molécules d'ARN d'une banque aléatoire d'ARN avec comme cible des petits colorants organiques (1). Ces ligands nucléiques ont alors été appelés aptamères, nom dérivé du latin « aptus » signifiant « se lier à » et du grec « méros » désignant ainsi un polymère spécifique d'une cible. Deux ans plus tard, des oligonucléotides d'ADN ont été sélectionnés à partir d'une banque d'ADN par Ellington et Szostak (3). Depuis, le champ d'études des aptamères s'est considérablement développé, impliquant un regain d'intérêt de la part de la communauté scientifique (**Figure 0-1**).

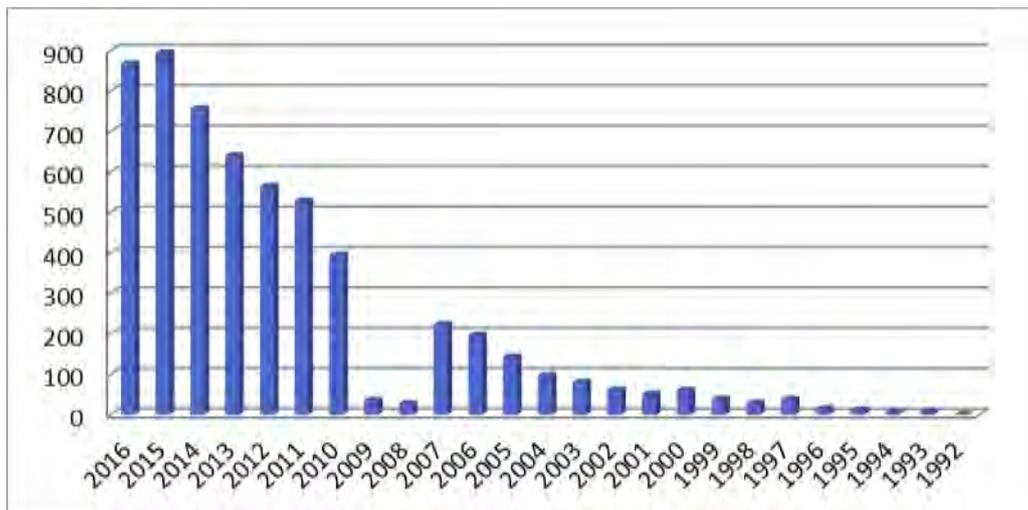


Figure 0-1 : Graphique représentant le nombre d'articles publiés sur les aptamères au cours du temps (1992 - 2016). Issu de <https://www.ncbi.nlm.nih.gov/pubmed/?term=aptamer>

Leur structure tridimensionnelle, leur longueur (inférieure à 100 nucléotides) et les conditions environnementales impliquent une structuration atypique en boucle, tige boucle, ou G-quadruplex. En présence de la cible, ils peuvent subir un changement de structure et la modification de leur structure permet de créer un site de liaison spécifique avec la cible. Plusieurs types d'interactions sont impliqués dans la liaison d'un aptamère avec sa cible : interaction électrostatique, hydrophobe ou liaisons hydrogène.

Les aptamères présentent de multiples avantages notamment avec une constante de dissociation de l'ordre du nanomolaire en fonction du type de cible. Ces caractéristiques uniques en font un choix plus avantageux, que les anticorps. Ils peuvent être sélectionnés à partir d'une banque d'oligonucléotides via un processus *in vitro* contre n'importe quelle molécule (protéines, peptides, acides nucléiques, molécules organiques etc...), sans avoir besoin de lignées cellulaires ou d'animaux comme c'est le cas pour les anticorps. De plus, ils peuvent subir une amplification grâce à la réaction en chaîne par la polymérase afin de pouvoir en produire une grande quantité avec une pureté élevée.

Ils peuvent subir des modifications en ajoutant des groupements fonctionnels (4), afin de permettre des applications ultérieures (**Figure 0-2**) (5). Leur structure chimique permet l'insertion de molécules fluorescentes ou d'agents de liaisons de surface sur des emplacements spécifiques. Par exemple, pendant la liaison sonde-cible, le changement de conformation de l'aptamère peut être utilisé pour mesurer des sécrétions cellulaires de protéines. (6)

Leur stabilité fait qu'ils peuvent notamment être conservés à température ambiante et leur régénération est réversible contrairement aux anticorps ou la sensibilité à la température entraîne une dénaturation irréversible. L'application des aptamères implique différents domaines tels que la thérapeutique, les biocapteurs ou le diagnostic. En 2004, la FDA (Food and Drug Administration) a approuvé un aptamère spécifique du facteur de croissance vasculaire endothélial (VEGF) pour le traitement de la dégénérescence maculaire liée à l'âge (DMLA). Connu sous le nom de Macugen® ou Pegaptanib sodium, ce médicament est commercialisé aux Etats-Unis depuis janvier 2005 et a obtenu une autorisation de mise sur le marché en France en mai 2006. (7)

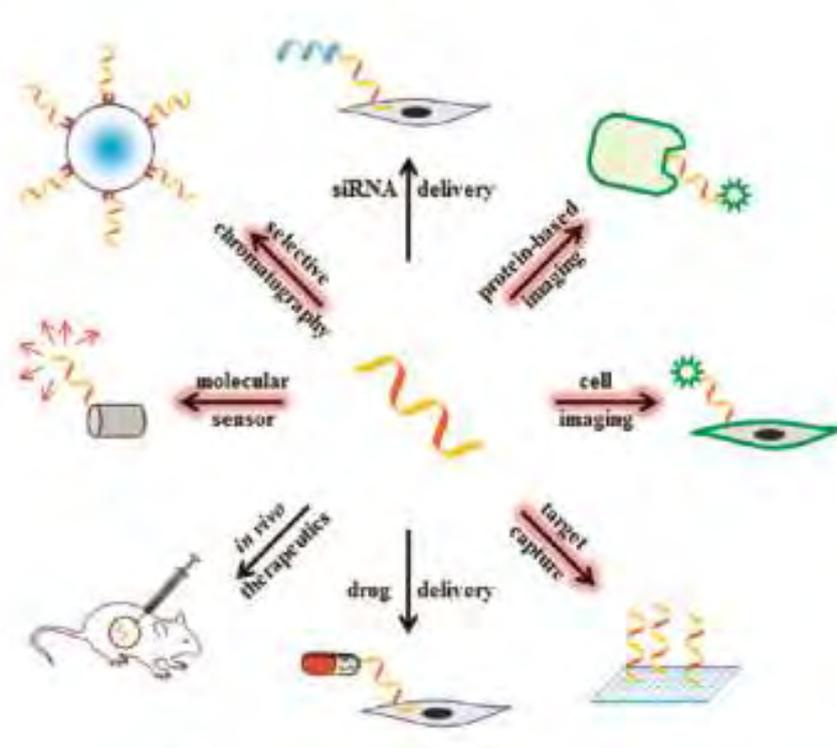


Figure 0-2 : Schéma indiquant les larges domaines d'applications des aptamères. Les applications bio analytiques sont surlignées en rouge. Issu de (5)

Dans ce domaine très prometteur et en croissance très importante, existe pourtant une limite méthodologique importante qui jugule cette progression : c'est l'étape de sélection à partir de la banque d'oligonucléotides. La méthode Systematic Evolution of Ligands by Exponentielle enrichment (SELEX) est la méthode la plus courante pour réaliser cette sélection (2). Elle a été utilisée sur de nombreuses classes de cibles comme des petites molécules organiques ou inorganiques, des peptides, des protéines ou des cellules (1). Cette méthode est rapidement devenue incontournable dans le domaine de la biologie moléculaire, de la pharmacie ou de la recherche médicale. La première étape consiste à utiliser une banque d'acides nucléiques. Les brins d'ADN ou d'ARN sont des séquences aléatoires composées de séquences aux extrémités 5' et 3' non dégénérées permettant d'utiliser des amorces de PCR afin d'initier la phase d'amplification après la sélection. Ces séquences aléatoires sont composées des 4 nucléotides naturels disposés aléatoirement. Pour n nucléotides dans la séquence, 4^n molécules différentes existeront dans la banque d'ADN ou d'ARN.

La seconde étape consiste à incuber cette banque d'ADN contenant généralement 10^{13} - 10^{15} séquences simple brin avec une cible d'intérêt pour initier le premier cycle de sélection. Elle met en œuvre la séparation des séquences nucléotidiques qui ont reconnu la cible par chromatographie d'affinité ou par séparation sur membrane, ou encore par l'utilisation de billes magnétiques. D'autres techniques ont été développées, comme la cytométrie de flux, l'électrophorèse capillaire... Les aptamères sélectionnés sont ensuite amplifiés par PCR et au bout de plusieurs cycles, les séquences potentiellement capables de se lier à la cible, sont clonées et séquencées. Cette méthode présente une alternative intéressante puisqu'une fois la séquence déterminée, les aptamères peuvent être synthétisés chimiquement. Cependant, elle reste une méthode longue puisque l'ensemble des étapes peut durer de quelques semaines à plusieurs mois. Par ailleurs, l'inconvénient majeur de l'approche SELEX est la prédiction incertaine de la réussite de la sélection. Dans le cas de la sélection par chromatographie d'affinité, il est nécessaire de faire une sélection négative, c'est-à-dire éliminer les séquences qui se lient de manière non spécifique au support. Bien que les paramètres intrinsèques des aptamères, tels que l'affinité et la spécificité, peuvent être influencés pendant le processus de sélection, il n'est pas possible de juger de l'aptitude d'une molécule cible à être appropriée pour le procédé de sélection et si un aptamère peut être obtenu avec succès contre la cible.

Beaucoup d'efforts ont été faits pour optimiser la méthode de sélection. L'équipe de Bowsert ainsi que l'équipe de Sergei Krylov quelques années plus tard ont mis au point une approche « hors équilibre » pour dépasser la contrainte temps en utilisant l'électrophorèse capillaire hors équilibre de mélanges à l'équilibre « Non Equilibrium Capillary Electrophoresis of an Equilibrium Mixture, NECEEM » (8). Dans un premier temps, on incube la banque d'ADN avec la cible. Les molécules qui ont une affinité avec la cible vont se lier tandis que celles qui ont peu d'affinités vont être non-liées. On injecte ensuite, un volume de ce mélange et on applique une tension pour séparer les fractions d'ADN liées à la cible du mélange. Ensuite on récolte ce complexe formé, on le réinjecte avec la cible à la même concentration de départ afin de récolter des aptamères complexés au bout de quelques cycles. Cette méthode est un puissant outil analytique qui facilite l'efficacité de sélection des aptamères, précise la détermination des paramètres d'interactions. Mais cette méthode est difficile à mettre en œuvre car peu de laboratoires utilisent cette méthode non-SELEX pour laquelle aucune standardisation existe et d'autre part, parce que la technique d'électrophorèse capillaire n'est toujours pas une technique largement répandue.

L'objectif de cette thèse a été de mettre au point une méthode de caractérisation d'aptamères par électrophorèse capillaire couplée à la fluorescence induite par laser et le séquençage haut-débit Illumina. La première étape a été de trouver des conditions de séparation qui permettaient d'obtenir une bonne répétabilité pour l'analyse de nos composés c'est-à-dire une banque d'ADN simple brin et une protéine. C'est ainsi que nous nous sommes aperçus que le choix du tampon Tris comporte quelques difficultés lorsqu'on veut faire des injections de petits volumes d'échantillons. Le Tris se dégradant électrochimiquement, nous avons opté pour la suite du travail pour un tampon de séparation phosphate de sodium. La thrombine a été utilisée comme modèle de sélection d'aptamère car elle a fait l'objet d'études par électrophorèse capillaire pour la détermination de la constante de dissociation, mais aucune de ces méthodes s'est intéressée à la sélection par non-SELEX/ NECEEM pour sélectionner un aptamère contre la thrombine. Deux aptamères, le T15 et le T29 se lient à cette protéine sur deux sites différents, avaient été découverts par la méthode SELEX. Ces aptamères T15 et T29 se structurent sous la forme d'un G-quadruplex. Ces structures peuvent présenter des difficultés à être amplifiées et séquencées, car elles sont construites par l'association de quatre tétrades de guanines et l'enzyme ADN polymérase peut avoir des difficultés à amplifier correctement les séquences proches du quadruplex. Dans ce contexte, nous avons vérifié que nous pouvions sélectionner par électrophorèse capillaire et séquençage haut-débit Illumina, un aptamère contenant un G-quadruplex. Ceci permettrait d'envisager l'utilisation de la méthode utilisée par Picometrics Technologies.

1. LES APTAMERES : STRUCTURE ET PROPRIETES

1.1 L'architecture des aptamères

1.1.1 Structure primaire

Les aptamères nucléotidiques sont des polymères biologiques qui résultent de l'association de nucléotides reliés entre eux par une liaison phosphodiester entre les atomes 3' et 5' de résidus de ribose ou déoxyribose adjacent (**Figure 1-1**). Les purines sont reliées au sucre sur l'atome N9 tandis que les pyrimidines sont reliées de par leur atome N1. La capacité de liaison des oligonucléotides simple brins à des cibles spécifiques est basée sur le principe de la conformation structurale.

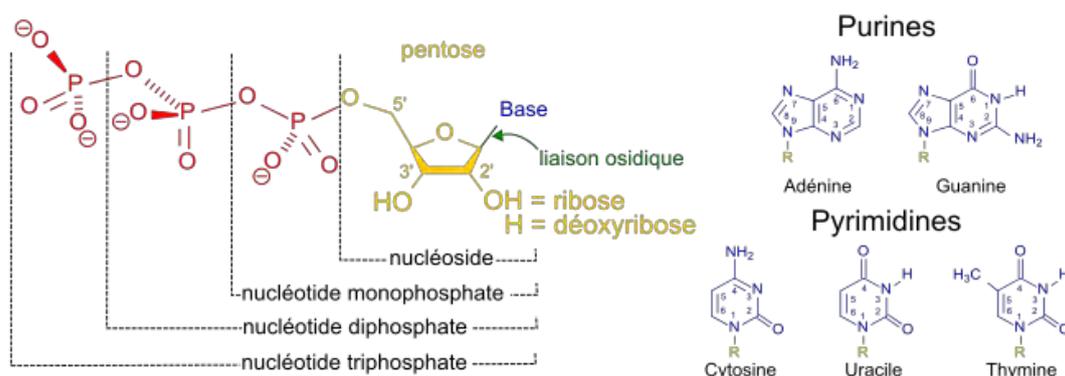


Figure 1-1 : Nucléotides et bases nucléiques. Issu de https://fr.wikipedia.org/wiki/Base_azot%C3%A9

Chaque motif unitaire possède une charge négative à pH neutre (due aux groupements phosphate), des sites donneurs/accepteurs de liaisons hydrogène et enfin un groupement aromatique (base nucléique) pouvant générer des interactions d'empilement. Par convention, les chaînes sont lues dans le sens 5' vers 3'. L'ARN possède sur le ribose un groupement hydroxyle en position 2'. Ce groupement le rend plus sensible à l'hydrolyse alcaline ce qui explique que l'ARN a une moins bonne stabilité chimique que l'ADN. De plus, l'amplification par PCR des brins d'ARN nécessite une étape supplémentaire de reverse-transcription. Ainsi, nous nous intéresserons surtout aux aptamères d'ADN.

1.1.2 Structure secondaire et tertiaire :

1.1.2.1 L'appariement des bases :

La structure secondaire résulte de la conformation des nucléotides et donc de l'interaction entre les nucléotides par des liaisons hydrogène et d'autres liaisons de type π -stacking. Ces différentes liaisons permettent ainsi à l'ADN d'adopter différentes structures.

L'adénine s'apparie avec la thymine via deux liaisons hydrogène, la cytosine et la guanine s'apparient quant à elles avec trois liaisons hydrogène (**Figure 1-2**). Ainsi, l'appariement entre deux brins contenant des liens guanine et cytosine génère une liaison plus stable (5.5 kcal vs 3.5 kcal).

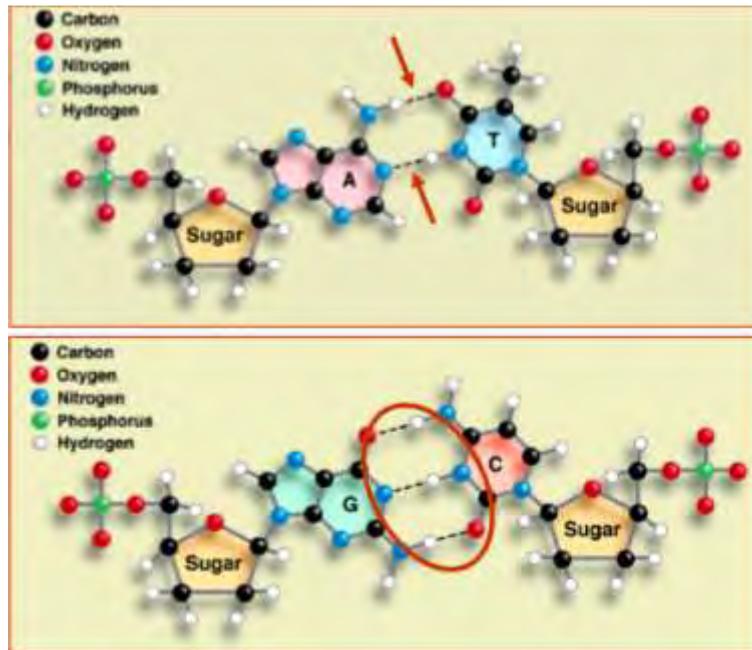


Figure1-2 : Structure secondaire de l'ADN et appariement des bases nucléotidiques. Issu de http://www.fsr.ac.ma/cours/biologie/amzazi/acide_nuc2.pdf

Cette structure est stabilisée par les interactions de type π - π entre les cycles aromatiques des bases (π -stacking) mais aussi grâce aux liaisons hydrogène entre les paires de bases. Elle est dépendante de la séquence primaire, de la longueur de l'acide nucléique (inférieur à 100 nucléotides), et des conditions environnementales (milieu aqueux et force ionique). Les bases qui ont un faible caractère hydrophobe sont empilées à l'intérieur selon un plan perpendiculaire à l'axe de la double hélice et l'extérieur (phosphate et sucre) est hydrophile.

1.1.2.2 La double hélice :

Dans le cas de l'ADN double brin et du fait de la dissymétrie des liaisons sucre-base, l'orientation entre les liaisons donne une structure en double hélice. La double hélice présente un grand sillon (à l'opposé des liaisons sucre-base) et un petit sillon (entre les liaisons sucre-base) dont les largeurs et profondeurs dépendent de la force ionique du milieu. Les protéines interagissent avec l'ADN en formant des liaisons spécifiques au fond des sillons par liaisons hydrogène et non spécifiques par interactions de van der Waals et interactions électrostatiques.

Trois formes naturelles de l'ADN ont été caractérisées (**Figure 1-3**), résultant de modifications mineures de la structure en double hélice, caractérisées par des techniques biophysiques telles que le dichroïsme circulaire (9). Parmi elles, la plus courante est l'ADN B (10), au grand sillon profond et large et petit sillon étroit. Puis en milieu déshydraté et riche en sels, on observe l'ADN A (11), où le grand sillon est plus large et le petit sillon plus étroit. Elle présente ainsi une hélice beaucoup plus condensée et se différencie par un déplacement de l'axe de l'hélice qui ne passe plus par le centre de l'appariement des bases mais à l'intérieur du grand sillon. Ceci dénote que la conformation de l'ADN dépend de la composition du milieu. La troisième forme observée est l'ADN Z (12), découverte par A.Rich en 1979 (13), elle présente une forme en zigzag de l'hélice. Cette forme est observée dans des fortes conditions salines et avec des

séquences riches en purines. Les sillons sont quasiment équivalents et l'hélice gauche est plus fine et allongée.

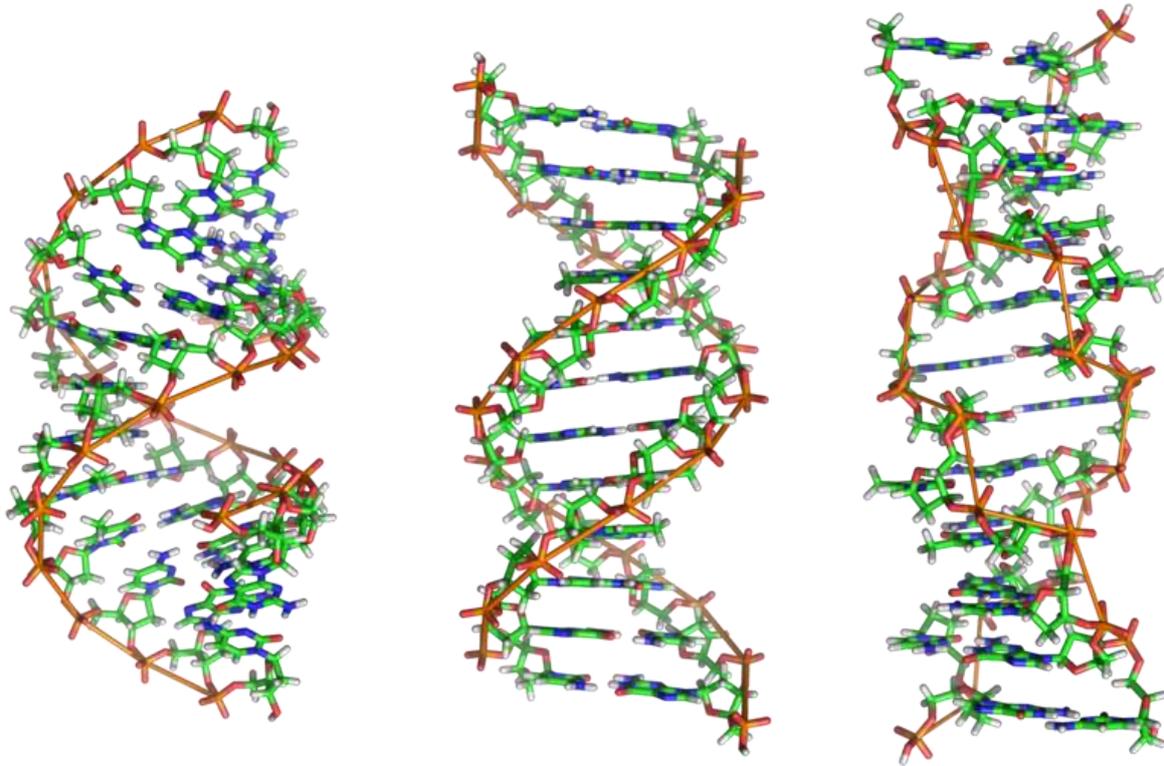


Figure 1-3 : les trois types de doubles hélices avec le petit sillon (0.34nm) et le grand sillon (3.4nm). Issu de <https://en.wikipedia.org/wiki/File:A-DNA, B-DNA and Z-DNA.png>

Dans le cas de simple brin d'ADN ou d'ARN, grâce aux liaisons hydrogène et à l'empilement des bases (stacking), ces molécules peuvent ainsi adopter une grande diversité structurale (**Figure 1-4**). Lorsqu'un brin se replie sur lui-même et qu'une ou plusieurs bases sont non-appariées, on parle de structure sous forme d'épingles à cheveux (hairpin). Cela peut concerner deux brins complémentaires et l'on parle alors de structure « bulge » ou boucle interne.

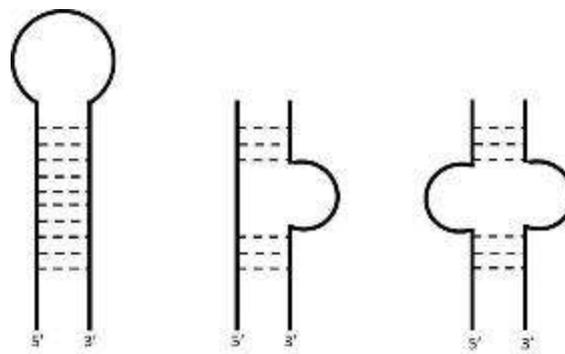


Figure 1-4 : Structures comportant une ou plusieurs bases désappariées. A gauche, structure de type hairpin ou « boucle à cheveux ». Au milieu, structure bulge et à droite structure en boucle interne.

D'autres motifs de structures secondaires et tertiaires existent dans les appariements entre base. La plus connue est le G quadruplex (**Figure 1-5**).

1.2 Etude de la structure en G-quadruplex de l'aptamère anti-thrombine

1.2.1 Définition des G-quadruplex :

D'autres motifs tertiaires des aptamères ont été décrits y compris la structure en G-quadruplex. Les G-quadruplex sont des structures tertiaires obtenus par des liaisons non Watson et Crick d'acides nucléiques, riches en guanine. Les nucléotides de guanine peuvent former au moins deux liaisons hydrogène sur chacune de leurs faces, ceci donne une capacité d'auto-assemblage de ces derniers et on parle alors de tétrade de guanines ou G-quartets lorsque quatre nucléotides de guanines s'assemblent (**Figure 1-5**). Dans une tétrade de guanines, les liaisons hydrogène se font entre la face donatrice de type Watson et Crick d'une guanine avec la face « acceptrice » de type Hoogsteen d'une autre guanine.

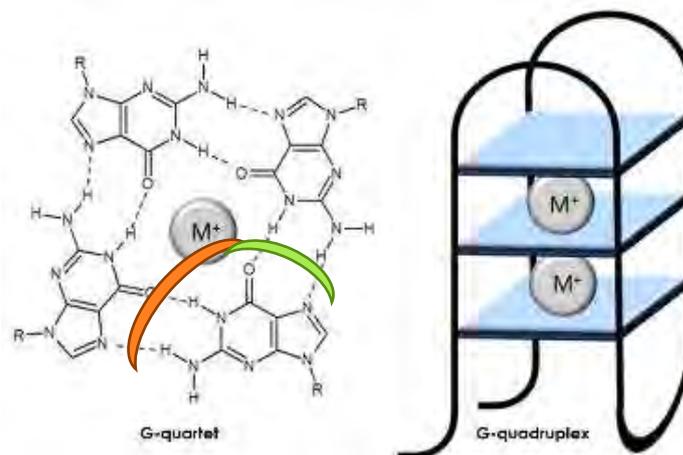


Figure 1-5 : Une tétrade de guanine (à gauche) et un G-quadruplex (à droite). Tracé orange face Crick Watson, vert face Hoogsteen. Issu de (14).

C'est l'équipe de Davies et al, qui en 1962, l'a découverte en analysant des gels d'acide guanylique. Quatre nucléotides guanines qui se reliaient par l'intermédiaire de liaison de type Hoogsten (**Figure 1-6**) formaient une structure hélicoïdale (15). L'appariement de type Hoogsteen implique le groupement purine et résulte de l'interaction de liaisons hydrogène entre ces bases dans le grand sillon de l'hélice.

Cette structure est stabilisée en présence de cations M^+ tels que le sodium ou le potassium qui se coordonnent au centre de la structure. Cet ensemble forme un motif plan et l'empilement de plusieurs tétrades de guanines constitue une structure secondaire appelée G-quadruplexes. La répétition du motif peut être composée d'un brin (G-quadruplex intramoléculaire) jusqu'à quatre brins (quadruplex intermoléculaires). De nombreuses méthodes peuvent être utilisées afin de mettre en évidence la structure en G-quartet, parmi elles, le dichroïsme circulaire est la

méthode la plus largement répandue pour analyser la topologie de cette structure. Contrairement, aux autres techniques telles que la RMN, ou la diffraction aux rayons X, qui donnent des informations précises sur la structure, le dichroïsme circulaire permet de mesurer la variation de structure des acides nucléiques induit par la fluctuation de la température et savoir si un G-quartet se forme. Cependant, la qualité des spectres obtenus est quelque fois peu sensible et dans ce contexte une approche computationnelle a été développée. Basée, sur la séquence primaire, un programme peut prévoir si des séquences peuvent potentiellement se structurer en G-quadruplex car toutes les séquences bien que riches en nucléotides de guanines ne présentent par la particularité de se structurer en G-quartet (15) .

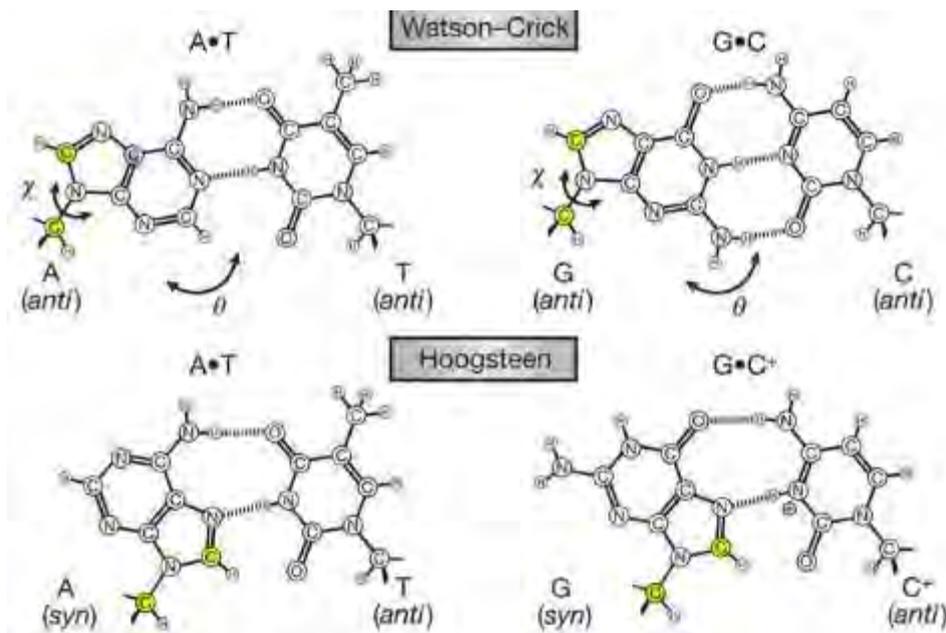


Figure 1-6 : Liaisons de type Watson et Crick (figures en haut) et liaisons de type Hoogsteen (figures en bas). Issu de https://commons.wikimedia.org/wiki/File:Hoogsteen_and_Watson%E2%80%93Crick_base_pairing.png

Le premier aptamère obtenu par SELEX et se structurant en G-quadruplex est l'aptamère anti-thrombine : T15, trouvé par l'équipe de Bock et al, en 1992. (16) Il est aussi nommé parfois TBA15 (Thrombin binding aptamer). C'est ce même aptamère qui fut pour la première fois analysé par RMN à une dimension ^1H en 1993, par l'équipe de Macaya et al. (17) Ils ont ainsi démontré qu'en présence de cations spécifiques tels que K^+ ou de la thrombine, ce dernier forme une structure intramoléculaire en G-quadruplex consistant en deux G-quartet connectés par deux boucles TT et une boucle TGT (Figure 1-7).

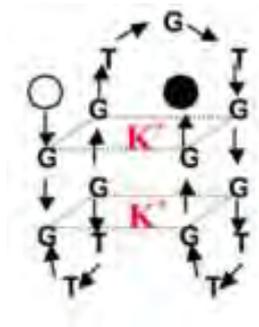


Figure 1-7 : Aptamère T15 stabilisé par les ions K^+

Depuis la découverte de cet aptamère avec une conformation en G-quadruplex, d'autres aptamères avec cette structure ont été observés (**Tableau 1** (18)).

Tableau 1: Exemple de quelques aptamères avec un G-quadruplex d'après (18)

Cible	Séquence	Topologie du G-quadruplex
Exemples de séquences de G-quadruplex interagissant avec des protéines		
Thrombine	d(GGTGGTGTGGTTGG)	AP, monomère
Nucléoline	d(GGTGGTGGTGGTTGTGGTGGTGGTGG)	AP, dimère
VEGF	d(TGTGGGGGTGGACGGGCCGGGTAGA)	AP
HIV.1 reverse transcriptase	d(GGGGGTGGGAGGGTAGGCCTTAGGTTTCTGA)	AP, monomère
Insuline	d(GGTGGTGGGGGGGGTTGGTAGGGTGTCTTC)	AP, monomère

1.2.2 Topologie des G-quadruplex :

Dans ces exemples de G-quadruplex nous voyons une caractéristique intéressante : la base nucléique la plus commune après la guanine G est la thymine T. On voit bien que beaucoup de structures en G-quadruplex sont formées de plusieurs nucléotides de guanines et de thymine $d(TGGG)_n$.

La thymine ne peut pas s'apparier avec les guanines contrairement à la cytosine (18). De plus, contrairement aux nucléotides d'adénine qui possèdent un noyau purique constitué de deux cycles pyrimidine et imidazole, la thymine ne possède qu'un seul cycle pyrimidique ce qui stabilise ainsi la structure en G-quadruplex. Les structures en G-quadruplex ont une double charge négative par unité de longueur contrairement à la structure en double hélice ce qui implique qu'en ayant un potentiel électrostatique plus élevé par unité de longueur, ceci provoque une forte liaison avec la surface chargée des protéines.

La conformation des aptamères G-quadruplex peut être déterminée par dichroïsme circulaire, mais il n'y a pas de règle précise concernant le mécanisme selon lequel un aptamère riche en séquences de guanines se structurera de façon parallèle (**Figure 1-8-C**) ou anti-parallèle (**Figure 1-8-A et B**). Plusieurs critères sont pris en compte dans l'orientation des brins incluant la nature de l'ion métallique, le nombre et le type de boucles, le nombre de tétrades et le nombre de brins les composants. Lorsque les brins sont orientés de façon parallèle, cela signifie que les 4 brins

d'acides nucléiques en G-quadruplex sont orientés dans la même direction. Dans le cas de structures à brins anti-parallèles, les 4 brins d'acides nucléiques composant le G-quadruplex sont orientés dans des directions différentes. Lorsqu'un brin est en anti-parallèle des trois autres, la structure est « hybride » (**Figure 1-8-D**). On peut aussi observer des réarrangements intermoléculaires et ainsi aboutir à une structure tétramérique (**Figure 1-8-F**), ou dimérique (**Figure 1-8-E**).

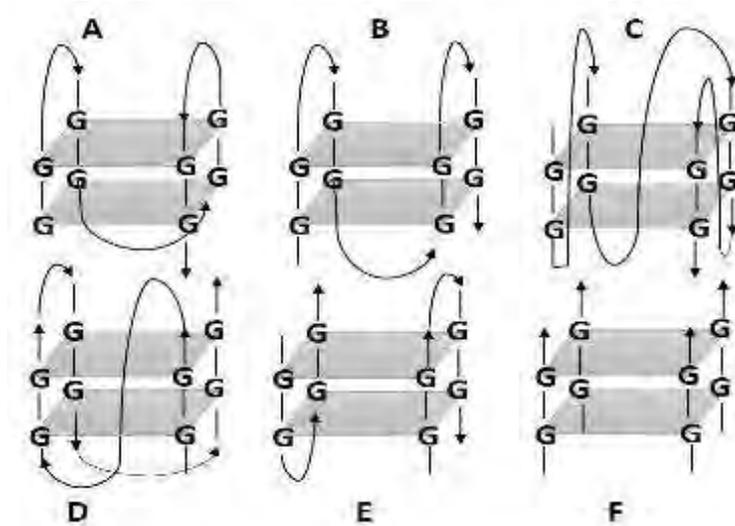


Figure 1-8 : Topologie des G-quadruplex : sous forme anti-parallèle (figure A et B), parallèle (figure C), hybride (figure D), dimérique (figure E) et tétramérique (figure F).

Les G-quadruplex doivent leur stabilité à l'ensemble de paramètres incluant les liaisons hydrogène, les liaisons de Van der Waals, les interactions de types ioniques, électrostatique et hydrophobe. Cependant, il existe une autre variable qui joue un rôle sur la stabilité des G-quadruplex : ce sont les angles glycosidiques adoptés par les bases, soit en conformation *anti* ou *syn* (**Figure 1-9**). Lorsque les brins sont parallèles entre eux, ils sont dans une conformation soit *syn* soit *anti*. Lorsque les angles sont différents entre deux brins anti-parallèles, l'un est en *syn* et l'autre est en *anti*.

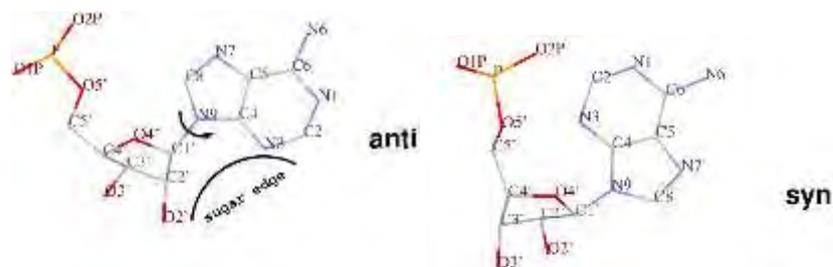


Figure 1-9 : Représentation des angles glycosidiques anti (figure à gauche) et syn (figure à droite). Issu de <http://x3dna.org/highlights/the-chi-x-torsion-angle-characterizes-base-sugar-relative-orientation>

Les techniques de spectroscopies telles que le dichroïsme circulaire en complément des techniques telles que la RMN ou la diffraction aux rayons X permettent d'étudier la formation de ces structures. Des oligonucléotides formant une structure parallèle avec des angles glycosidiques en *anti* sont caractérisés par un spectre en dichroïsme circulaire (**Figure 1-10**) présentant une large bande positive à 260 nm et une bande négative d'intensité plus faible à 240

nm ainsi qu'une autre bande d'intensité positive à 210 nm. Dans le cas de structures anti-parallèle, où les guanines alternent avec des angles glycosidiques de conformation *syn* et *anti* ont une bande négative à 260 nm et une bande positive à 290 nm. Cependant, comme le soulignent l'équipe de Kypr (19) ainsi que l'équipe de Spada et al. (20), c'est une relation empirique qui bien qu'elle puisse être appliquée à la plupart des G-quadruplex ne peut pas s'appliquer sur des systèmes plus complexes.

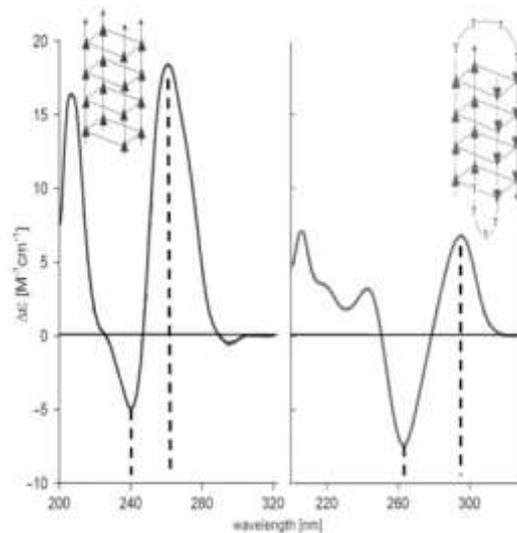


Figure 1-10 : Spectres obtenus par dichroïsme circulaire d'un G-quadruplex parallèle (à gauche), et anti-parallèle (à droite). Issu de (20)

1.2.3 Le rôle des ions sur la stabilité des G-quadruplex :

Les ions métalliques jouent un rôle dans la formation des G-quadruplex. La capacité des ions à induire la formation de cette structure varie selon l'ordre suivant : $K^+ > Rb^+ > Na^+ > Cs^+ > Li^+$.

La présence des ions mono et divalents stabilise la structure des acides nucléiques. Le repliement en solution et les changements de conformations sont très fortement dépendant de la présence de cations bivalents (21) (22) tels que les cations divalents (Mg^{2+} , Mn^{2+} , Ca^{2+}), qui ont une plus forte affinité pour les acides nucléiques que les cations monovalents. La stabilisation des G-quadruplex dépend des rayons atomiques des ions. Si l'ion est trop petit comme l'ion lithium, il ne pourra pas interagir au sein des tétrades pour les stabiliser. Dans le cas des ions sodium, ce dernier entraîne une perte de stabilité par rapport aux ions potassium ou chaque ion K^+ interagit avec les deux tétrades de guanines ce qui apporte plus de stabilité. Selon la disposition des cations, ces derniers influencent la topologie adoptée par les G-quadruplex et il arrive que deux G-quadruplex soient différents selon le cation présent en solution. Un certain nombre d'études notamment par dichroïsme circulaire, avec l'aptamère T15 ont été réalisées afin de mettre en évidence l'influence des cations et notamment le fait que les cations K^+ stabilisent mieux la structure en G-quadruplex contrairement aux ions Na^+ (23).

1.2.4 Etudes des aptamères anti-thrombine :

La thrombine est une sérine protéase multifonctionnelle de la famille des sérines protéases qui joue un rôle clé dans la réaction de l'hémostase. Cette enzyme déclenche la coagulation en hydrolysant le fibrinogène et en activant les plaquettes sanguines (24). Les caractéristiques structurales dominantes de la thrombine comprennent un site actif profond et deux surfaces chargées positivement, appelées exosites I et II. L'exosite I est le site de liaison au fibrinogène, tandis que l'exosite II est le site de liaison à l'héparine de la thrombine. La surproduction de thrombine peut causer des maladies dangereuses, comme l'apoplexie et les crises cardiaques. Cela justifie la très grande importance d'une méthode de sélection d'inhibiteurs de cette dernière dont les aptamères.

1.2.4.1 Le T15 :

Un des premiers aptamères à avoir été sélectionné par Bock *et al* (16), en 1992 à partir d'une banque d'ADN dégénérée sur 60 nucléotides, c'est-à-dire avec une diversité de 4^{60} nucléotides, fut l'aptamère T15 connu sous le nom de TBA : Thrombin binding aptamer avec une constante de dissociation K_D de 100 nM. Ils se sont intéressés à un ensemble de séquences de 14 à 17 nucléotides avec huit résidus de guanines qui forment un noyau central composé de deux quartets de guanines. Par la suite, le motif suivant de 15 nucléotides a été conservé : 5'-GGT-TGG-TGT-GGT-TGG-3' et de nombreuses études ont été menées car le T15 a une forte activité anticoagulante (25), il inhibe la coagulation du fibrinogène catalysé par la thrombine *in vitro* et réduit la formation de thrombus artériel. De plus il fut le premier à être utilisé pour le développement de biocapteurs appelés aptasensors (26). Le T15 consiste en deux tétrades de guanines qui sont connectées par deux boucles T-T et une boucle T-G-T selon les analyses aux rayons X de Padmanabhan (27) (Figure 1-11).

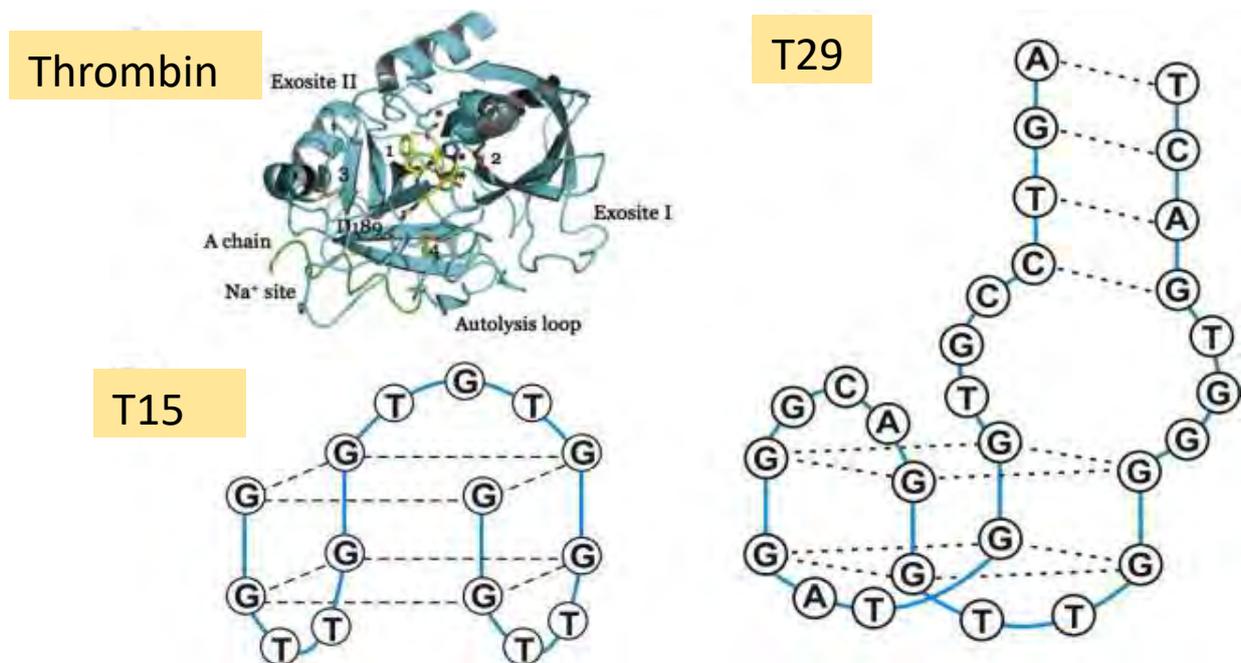


Figure 1-11 : liaison de l'aptamère T15 et de l'aptamère T29 avec la thrombine. Issu de 6 : Analytica Chimica Acta. 2014. Aptamer binding assays for proteins: The thrombin example. Deng B, Lin Y, Wang C, Li F, Wang Z, Zhang H, Li XF, Le XC.

Ce sont les boucles T-T qui se lient de façon spécifique sur l'exosite I (28), mais il existe aussi des interactions avec l'exosite II (29). En présence de cations spécifiques tels que K^+ , la structure est ainsi stabilisée (30), ainsi que lors de la réaction d'interaction avec la thrombine (31). La découverte de cet aptamère a abouti au développement de l'aptamère T15 en phase clinique de niveau I (ARC183) développé par Archemix corp et Nuvelo Inc en tant qu'anticoagulant pour une utilisation potentielle en chirurgie cardiaque. Il a ainsi été démontré qu'il y avait une réponse rapide de l'administration d'ARC183 mais la quantité d'aptamère nécessaire pour atteindre l'effet désiré était trop élevée ce qui n'a pas convaincu les entreprises de poursuivre leur développement (32). Néanmoins, cela a suscité un engouement pour la découverte d'autres aptamères de longueurs plus importantes et plus efficaces.

1.2.4.2 Le T29 ou HD22 :

Le T29 fut sélectionné en 1997 par l'équipe de Tasset et al (33) à partir de deux banques dégénérées comprenant 30 et 60 nucléotides. C'est un aptamère ayant pour séquence nucléotidique 29 bases dont un motif de 15 nucléotides similaires à l'aptamère T15 et avec huit guanines qui donnent la structure en G-quadruplex. L'aptamère T29 présente une affinité plus élevée pour la thrombine (caractérisée par un K_D de 0.5nM). Cet aptamère reconnaît le site de liaison de l'héparine (exosite II). La séquence du T29, où les guanines qui permettent d'aboutir à la structure en G-quartet sont surlignées en gras, est la suivante :



D'autres aptamères dérivés du T29 ont également été découverts, tels que TBA₂₇ ou TBA₃₁ qui ont une structure similaire à la séquence du T29 (34). Cependant, l'équipe de Gué a testé différents aptamères pour la mise au point d'un biocapteur pour la détection de la thrombine dans le plasma et en étudiant l'interaction aptamères-thrombine par SPR (surface plasmon resonance), ils ont montré que l'aptamère T29 était le seul qui permettait la détection efficace de la thrombine dans des échantillons réels de plasma sanguin (35). C'est pour cette raison que nous avons utilisé cet aptamère dans le cas de notre étude.

2 MODES DE SELECTION DES APTAMERES

2.1 SELEX

2.1.1 Principe général

La méthode SELEX (Systematic Evolution of Ligands by Exponential Enrichment) est une méthode très utilisée pour la sélection d'oligonucléotides envers une cible avec une forte affinité (**Figure 2-1**). Elle consiste à utiliser une banque d'oligonucléotides qui peut contenir jusqu'à 60 oligonucléotides aléatoires qui sont liés à de courtes régions constantes (primers), habituellement utilisées pour hybrider les amorces durant la PCR. La région centrale de la banque contient généralement entre 20 et 80 nucléotides dégénérés tandis que les extrémités non dégénérées servent à la fixation de primers de PCR et ont une longueur de 18 à 21 nucléotides. Plusieurs cycles de sélection sont alors réalisés où dans le cas des molécules d'ADN, on fait interagir la banque avec la cible. Pour la sélection d'aptamères ARN, la banque

d'ADN doit d'abord être convertie en banque d'ARN avant de commencer les cycles itératifs de sélection. Lors du premier cycle de sélection, l'ADN ou ARN est incubé avec la cible. Ensuite, les molécules non-liées sont séparées des molécules d'ADN ou d'ARN liées à la cible par des étapes de lavages. De l'efficacité des lavages, découle l'efficacité de sélection et donc les caractéristiques de liaisons des aptamères à la cible. Plusieurs techniques de séparations peuvent être utilisées : la chromatographie d'affinité qui utilise des colonnes de séparation en agarose ou sépharose (36), (37), la séparation magnétique (38). Ces méthodes sont basées sur l'immobilisation de la cible afin de générer des matrices d'affinités, capturer les molécules d'intérêts liées à la cible où les sélectionner selon la taille. D'autres méthodes de séparation sans immobilisation de la cible existent comme la filtration sur membrane de cellulose ((2) (39)) et la centrifugation (40). Ces techniques dépendent de la taille de la cible ainsi que de la liaison avec le complexe. Les oligonucléotides liés à la cible sont ensuite amplifiés par PCR (ADN) ou par PCR en temps réel (qPCR) (ARN). Les oligonucléotides doubles brins résultants sont purifiés en simple brin afin de réaliser un deuxième cycle de sélection (pour l'ADN, pour l'ARN il faut faire une transcription). Au total, plusieurs cycles (environ une vingtaine) sont effectués. Ce nombre varie en fonction de la cible, de la variété de la banque d'ADN initiale, des conditions de sélection et de la méthode de séparation. Par ailleurs, il est fortement recommandé de faire des cycles de sélection négatifs ou des étapes de soustractions afin d'éviter un enrichissement non-spécifique des oligonucléotides. En effet l'affinité de la cible pour la banque d'oligonucléotides peut être influencée par la stringence des conditions de sélection. Au fur et à mesure des cycles itératifs de SELEX, la stringence augmente et la banque d'ADN est ainsi réduite à quelques motifs ou séquences d'ADN avec la meilleure affinité et spécificité pour la cible. Ceci peut être réalisé en réduisant la concentration de la cible dans les cycles successifs de SELEX ou en changeant les conditions de lavages et de liaison de la cible (composition du tampon, volume). De même afin d'éviter des réactions croisées, et pour sélectionner des aptamères spécifiques d'une cible, une étape de pré-incubation appelée contre sélection est effectuée en présence de molécules cibles potentiellement interférentes (41). Le dernier cycle est stoppé juste après l'étape d'amplification par PCR. Les aptamères obtenus sont clonés et ensuite séquencés (42). Malgré que de nombreux aptamères aient pu être découverts par le procédé SELEX, cela reste toujours une méthode longue et compliquée. Les conditions de sélection décrites telles que les conditions de lavages, concentrations ioniques ou de cibles ou ajouts de compétiteurs sont susceptibles d'influencer le résultat final. Par ailleurs, le nombre possible de différentes séquences d'aptamères à cribler (4^n où n correspond aux nombres de nucléotides dans la séquence) est impossible à analyser complètement. A cela s'ajoute les biais d'amplification introduits lors de la PCR ce qui augmente l'amplification d'ADN non spécifique et a comme conséquence l'apparition de séquences abondantes mais qui n'ont pas pour autant la meilleure affinité. Mais ces mutations engendrent de la diversité dans les aptamères sélectionnés et constituent l'évolution darwinienne de la population de séquences. Elles peuvent permettre de créer de meilleurs aptamères qui n'étaient pas présent dans la banque d'origine. Depuis son développement en 1990, le SELEX-conventionnel a subi de nombreuses modifications afin d'améliorer son efficacité.

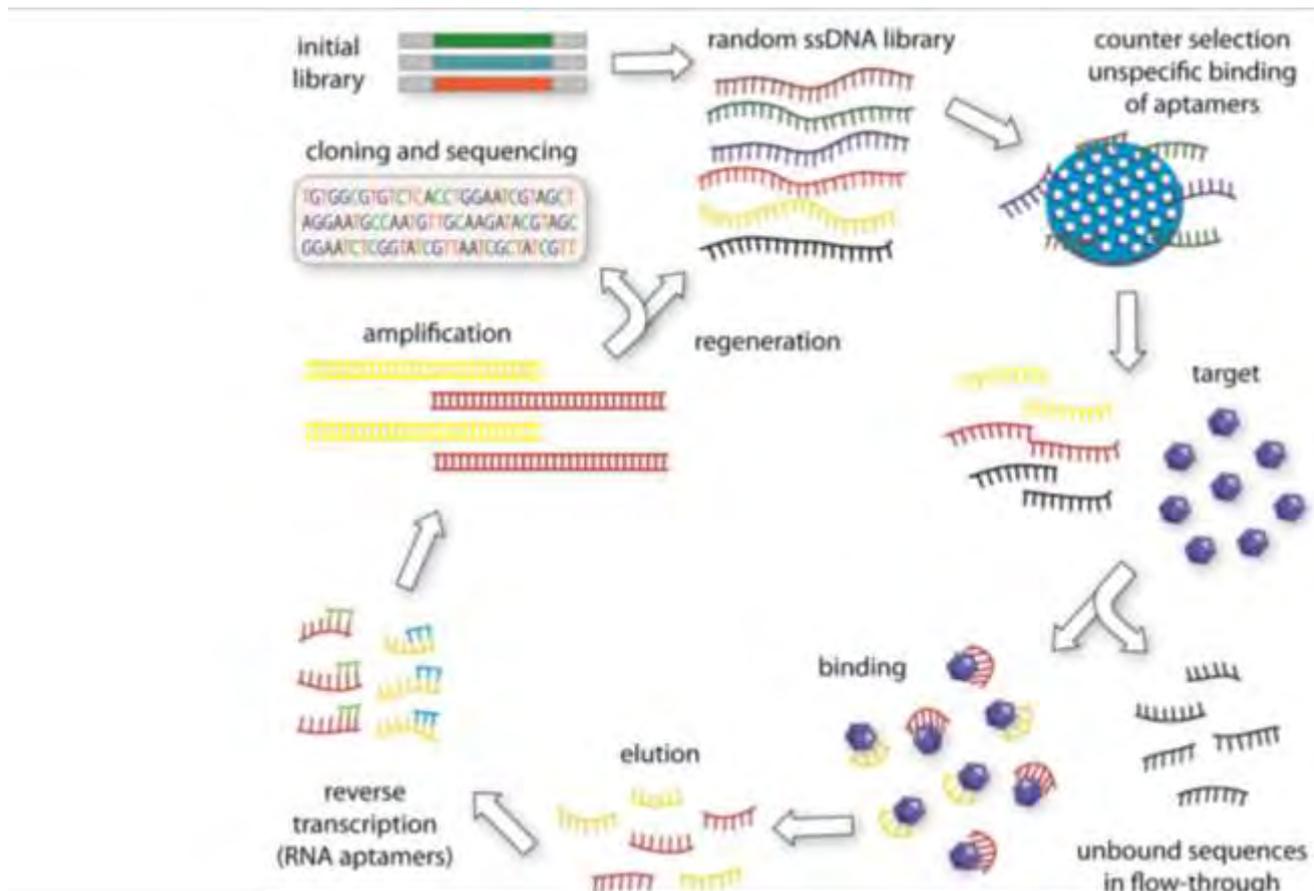


Figure 2-1 : Principe du SELEX d'après Darmostuk et al. (43)

2.1.2 Modification de la technologie SELEX

Afin d'améliorer soit la diversité des cibles ou le rendement du procédé SELEX, de nombreuses modifications y ont été apportées. Darmostuk et son équipe ont décrit toutes les améliorations de la méthode SELEX ayant été développées depuis la première utilisation de cette méthode (43).

Il y a peu de méthodes automatisées de SELEX à cause de la grande complexité à mettre en place un procédé d'automatisation qui intègre les différentes méthodes. Des variantes ont été développées. Parmi ces méthodes, la méthode Cell-SELEX qui permet à des cellules vivantes d'être utilisées comme cibles. Elle présente un intérêt dans des applications *in vivo* car bien que des aptamères aient été développés par la méthode conventionnelle SELEX contre des cibles telles que des peptides ou des protéines purifiées, ces derniers ne sont pas toujours efficaces pour des applications *in-vivo*.

2.1.2.1 Cell-SELEX

Cette méthode ne nécessite pas de connaissance de la cible pour la sélection des aptamères contrairement à la méthode SELEX (**Figure 2-2**). Des cellules entières, à la fois eucaryotes et procaryotes peuvent être utilisées pour générer des aptamères hautement spécifiques en utilisant cette technique. Ces aptamères peuvent être utilisés pour le diagnostic de cellules spécifiques ou dans le domaine thérapeutique pour la destruction ciblée de cellules. Comme dans la méthode SELEX classique, une banque d'oligonucléotides est incubée avec les cibles, qui sont des cellules entières. Les oligonucléotides non liés sont éliminés par lavage et les aptamères liés sont séparés de la cible et amplifiés par PCR.

La sélection n'est possible que si les aptamères qui se lient sur la surface des cellules des autres cibles sont retirés par l'intermédiaire d'une sélection négative au cours des cycles de sélection. Par conséquent, la méthode Cell-SELEX est souvent constituée d'une sélection positive contre les cellules cibles et d'une sélection négative contre les autres cellules. Elle peut exiger jusqu'à 35 cycles de sélection mais le plus souvent 8-10 cycles sont appliqués afin d'obtenir des aptamères avec une haute affinité. Certains aptamères utilisés contre divers cancers ont été générés par cette méthode (44) (45) (46) (47) (48) (49). Par exemple, les cellules cancéreuses et les cellules saines obtenues à partir du même tissu du même patient seraient un couple idéal de cellules.

Un des avantages majeurs de la méthode Cell-SELEX est que l'on n'a pas besoin de connaître l'identité de la cible des aptamères ou son état de modification. Toutefois, cette nature inconnue de la cible constitue également un obstacle pour l'utilisation thérapeutique de ces aptamères. L'identification de la cible et la vérification ultérieure de la spécificité prend beaucoup de temps et peut retarder de manière significative l'utilisation dans des applications de ces aptamères dérivés de ces cellules dans des applications diagnostiques et thérapeutiques (50) (51) (49).

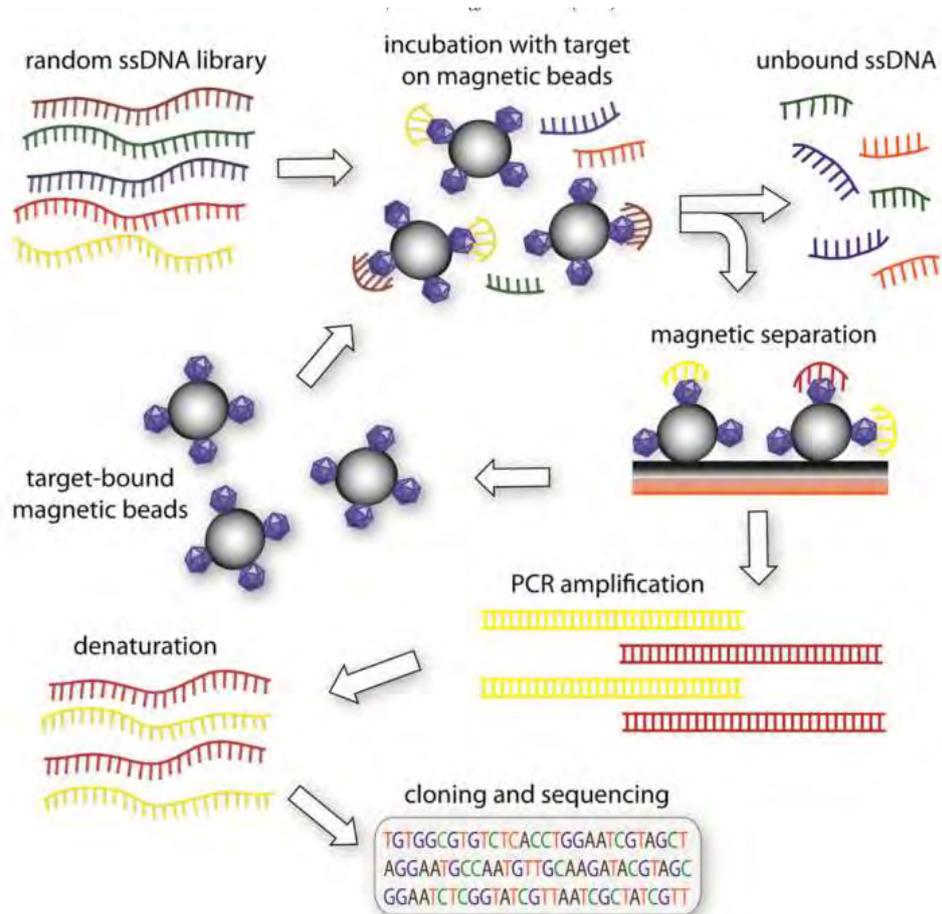


Figure 2-2 : Principe du Cell-SELEX. Issu de (43)

2.1.2.2 SELEX négatif

Utilisée pour la première fois par Ellington et Szostak en 1992 (3), cette méthode a été utilisée lors du troisième cycle de sélection dans le but d'éliminer les molécules d'ADN se liant sur la surface d'agarose où sont immobilisées les protéines. Ainsi les auteurs ont obtenu des séquences avec une affinité 10 fois plus élevées lors du 4ème cycle de sélection.

2.1.2.3 Contre SELEX

Cette méthode a été introduite dans le but d'augmenter l'efficacité de la sélection des aptamères. On introduit des analogues structuraux étroitement liés à la cible dans le processus de sélection afin d'éliminer les aptamères non-spécifiques et d'augmenter la spécificité des aptamères sélectionnés. Ce processus est analogue à celui du SELEX négatif sauf que l'on échange la cible par des dérivés analogues. Utilisée par Jenison et son équipe (52) lors de l'identification des aptamères d'ARN pour la théophylline. Les auteurs ont utilisé la caféine qui diffère de la théophylline uniquement par un groupe CH₃ à la place de l'atome H sur le N-7, ce qui leur a permis de sélectionner spécifiquement des aptamères anti théophylline.

2.1.2.4 SELEX soustractif

Le SELEX soustractif a, là aussi, comme but d'améliorer la sélectivité des aptamères sélectionnés comme dans le cas du contre SELEX, mais la différence entre ces deux méthodes se situe dans le procédé. Le SELEX soustractif élimine les molécules d'ADN qui se lient sur une autre surface que la cible d'intérêt avant que celle-ci ne soit mise en interaction. C'est l'équipe de Wang (53) qui a mis au point ce procédé en sélectionnant des aptamères qui distinguaient spécifiquement des cellules normales de cellules PC12 différenciées. Au cours de cette procédure, la banque d'oligonucléotides est incubée avec des cellules PC12 normales avant chaque cycle de sélection, où des cellules PC12 différenciées sont utilisées comme cible. Un protocole SELEX standard est appliqué, suivi d'une évaluation d'affinité des aptamères obtenus. La méthode aboutit à des aptamères sélectionnés avec une haute résolution pour des cibles similaires.

2.1.2.5 Photo SELEX

Le photo SELEX est une méthode de sélection d'aptamères basée sur la capacité des acides nucléiques photosensibles de réaliser une liaison covalente avec d'autres molécules sous une certaine longueur d'onde. Du fait que cette méthode utilise l'irradiation par un rayonnement UV induisant la formation d'une liaison covalente entre la cible et les acides nucléiques, la sélectivité des aptamères est ainsi améliorée. Mais le nombre de faux positif obtenu par cette méthode est également plus élevé de sorte que les conditions doivent être optimisées pour s'assurer de la validité de détection. Le processus est le suivant : on incorpore un atome d'iode ou de brome en position 5' du cycle aromatique des pyrimidines dans la banque d'ADN puis le pool est incubé avec la cible. Le mélange est ensuite irradié aux rayons UV ce qui induit la formation d'une liaison covalente par réticulation entre les acides nucléiques modifiés et la cible. Le mélange est ensuite séparé sur gel SDS PAGE. Cette méthode a été utilisée pour la première fois par Jansen en 1995 (54) qui a obtenu des aptamères pour la protéine Rev (régulateur de l'expression protéique de la vision) du virus VIH-1 en utilisant la 5-iodouridine (5-IU).

2.1.2.6 Toggle SELEX

Cette méthode utilise plusieurs cibles incubées avec une banque d'ADN en même temps dans le procédé de sélection d'aptamères. Le but est d'obtenir des aptamères avec une réactivité croisée. Lors du premier cycle de sélection, une banque d'ADN est incubée avec un mélange de cibles. Les oligonucléotides enrichis sont utilisés dans les cycles de sélection suivants où ils sont incubés avec des cibles séparément. Tout d'abord, les oligonucléotides sont incubés avec une cible, puis on élue les séquences liées. Au prochain cycle, ces séquences sont incubées avec la deuxième cible. Les séquences liées sont à nouveau isolées et utilisées dans les cycles de sélection suivants. Après treize cycles de sélection, les aptamères sont clonés et séquencés.

Cette méthode peut être utilisée pour sélectionner des aptamères avec une spécificité large et une fonctionnalité pour des applications *in vivo* améliorée (55).

2.1.2.7 Design d'oligonucléotides

- Longueur de la banque

Un aspect important de la méthode SELEX tient compte de la variété de la banque d'ADN ; c'est-à-dire la taille de la région dégénérée, le type de dégénérescence ainsi que les éventuelles modifications chimiques réalisées durant la phase de synthèse de la banque.

Une banque d'ADN dégénérée sur 20 ou 60 nucléotides présentera une plus ou moins grande complexité et donc la possibilité de créer des nouvelles liaisons spécifiques avec la cible (56) (57). Par ailleurs des banques d'ADN chimiquement modifiées avec des groupes fonctionnels pourront augmenter la stabilité de conformation ou la résistance aux nucléases et donneront également une complexité importante.

Inversement, une courte banque est plus simple d'utilisation ce qui est préférable pour sélectionner de courts aptamères. C'est aussi moins coûteux à synthétiser.

- Diversité de la banque

La probabilité de trouver un aptamère d'intérêt sera d'autant plus grande que la banque aura une grande diversité, c'est-à-dire que les banques d'ADN auront le plus grand nombre de composés différents.

Si l'on considère une banque d'ADN dégénérée sur 50 nucléotides, la diversité sera de 10^{30} molécules. Ainsi pour disposer de chacune des 10^{30} molécules, il faudrait $10^{30} * \left(\frac{330}{6.022 * 10^{23}}\right) g$ d'échantillon soit 5.10^3 tonnes !

Par ailleurs durant les étapes d'amplifications par PCR, les polymérase bactériennes utilisées sont connues pour engendrer des mutations (environ une mutation tous les $2 * 10^4$ nucléotides incorporés). Des efforts ont été faits afin de réduire ce taux de mutation en modifiant les polymérase de manière à les rendre plus fidèles. Tout au long du processus SELEX, un candidat sur 100 sera modifié à chaque cycle d'amplification pour une banque avec une région aléatoire de 100 nucléotides, conduisant donc à l'apparition de 10^{13} candidats pour la totalité de la banque. Certains chercheurs ont même réalisé des amplifications dans des conditions de diminution de fidélité de la polymérase (41) (56) (58) (59) (60).

- Modification de l'unité sucre

Les aptamères sont capables de se structurer en trois dimensions afin de se lier à leurs cibles. Cependant, les nucléases compromettent l'utilisation des aptamères dans des applications *in-vitro* ou *in-vivo*. Par ailleurs, le comportement des aptamères dans les milieux biologiques ne sont pas encore bien appréhendés et l'on peut supposer que ces derniers du fait de leur forte affinité pourraient se fixer à n'importe quelles cibles. Par exemple, la présence de nucléases dans le sérum conduit à la digestion rapide des oligonucléotides naturels. Afin de rendre les acides nucléiques plus résistants aux nucléases ou développer des groupements fonctionnels sur les acides nucléiques afin d'améliorer leur affinité de liaison à des cibles, plusieurs modifications ont été proposées. Le but étant de les rendre plus résistants et assurer leur biostabilité.

Les premières modifications concernent l'amélioration de la résistance aux nucléases notamment en faisant des substitutions en position 2' du (déoxy)-ribose (61). Les autres substitutions concernent les 2'-amino pyrimidines (62), 2'-fluoro pyrimidines (63) et les 2'-

méthoxy nucléotides (**Figure 2-3**). La première utilisation des 2'-NH₂-pyrimidines a été reportée par Jayasena en 1994.

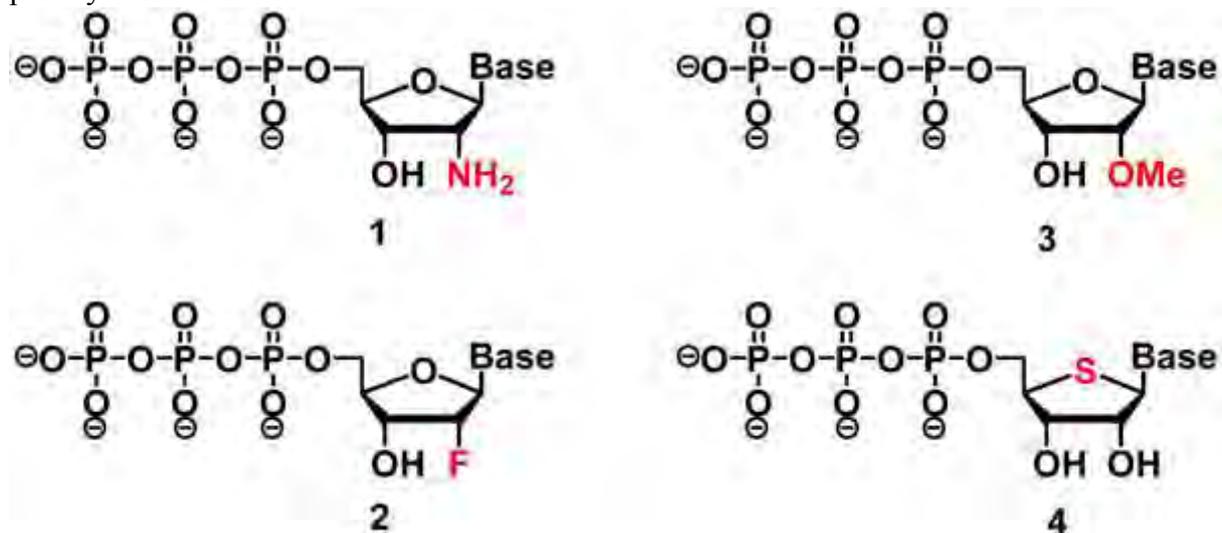


Figure 2-3 : Exemples de modifications de l'unité sucre par un groupement fonctionnel. Issu de (65).

Cependant les dérivés 2'-NH₂ ne sont plus utilisés car de nombreux problèmes ont été rencontrés durant la phase de synthèse (64). En effet, le groupement 2'-NH₂ peut déstabiliser les hélices par rapport aux groupements 2'-OH ou 2'-H en augmentant l'entropie de la liaison (64).

Les dérivés 2'-fluoro-pyrimidine ont été utilisés pour la banque d'ARN transcrite par l'ARN polymérase T7^{Y639F} utilisée pour des aptamères tels que la Tenascin C (65), mais également pour l'antigène membranaire spécifique de la prostate (PSMA) (66). Svobodova et son équipe ont généré depuis des aptamères ARN contre l'antigène spécifique de la prostate avec une constante d'affinité de 630nM. Ce dernier a été testé pour une application soit dans le domaine du diagnostic ou de la thérapeutique et donne des résultats plus prometteurs. (67) Les dérivés 2'-fluoro-pyrimidine et 2'-aminopyrimidine sont utilisés pour la production d'aptamères dans le but d'augmenter la résistance aux nucléases. La stabilité des dérivés 2'-fluoro-siRNAs a été estimée supérieure à 1 jour (68). La même polymérase a été utilisée pour générer des aptamères 2'-O-méthyl qui s'avèrent être moins toxiques que les autres banques proposées (69). Les dérivés 2'-methoxy et 2'-fluoro sont devenus les plus utilisés pour la sélection *in-vitro* d'aptamères. La combinaison des deux motifs a mené au développement du seul aptamère approuvé par la FDA : le pegaptanib (Macugen®, développé par les laboratoires Pfizer) (7) (70).

D'autres modifications consistent à utiliser des monothioaptamères, dans lequel une partie de l'oxygène est remplacé par un soufre.

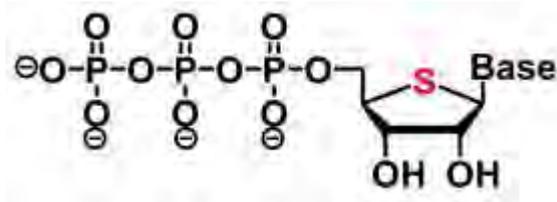


Figure 2: Un exemple de monothioaptamères. Issu de (65).

Les thioaptamères offrent d'autres avantages comme une augmentation de la stabilité, de l'affinité et de la spécificité. En 2000, le groupe de Matsuda a remplacé l'oxygène intracyclique de l'unité sucre par un atome de soufre. (71) Cette modification est également compatible avec une amplification enzymatique. Les derniers développements concernant la modification de l'unité sucre des acides nucléiques concernent les LNA (locked nucleic acids), HNA (hexitol nucleic acids), TNA (threose nucleic acids) et autres XNAs (xeno nucleic acids) utilisés comme blocs de construction (72).

- Modification de la base nucléique

Ces modifications concernent généralement le carbone 5 des bases pyrimidiques et l'azote 7 des bases purines car c'est à ces deux positions que l'amplification enzymatique se fait le mieux et l'architecture des oligonucléotides modifiés est bien stabilisée (Figure 2-4).

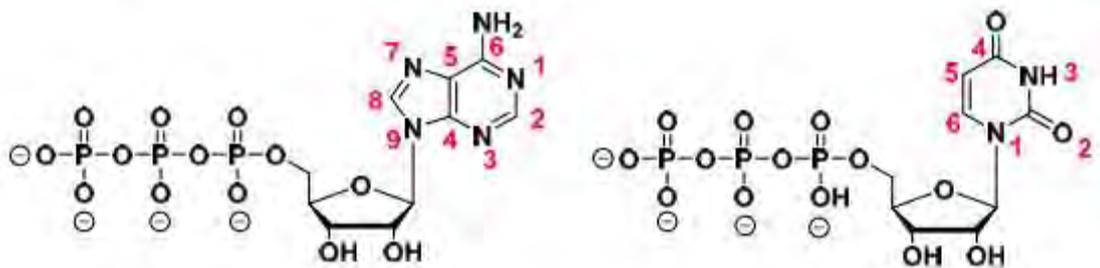


Figure 2-4 : Structure chimique des groupements purines et pyrimidines. Issu de (65).

L'introduction de ces modifications augmente les interactions entre les oligonucléotides et leur cible. Ceci augmente l'affinité des oligonucléotides pour leur cible et entraîne la création de structures secondaires ne pouvant pas être accessibles avec des acides nucléiques naturels.

Ainsi les différents groupements fonctionnels qui peuvent être ajoutés peuvent aider à l'amélioration de l'interaction entre un aptamère et sa cible. Par exemple, l'addition d'un groupement aromatique ou aliphatique (benzyl, naphthyl, pentinyl ou isobutyl) pourra augmenter la nature hydrophobe des oligonucléotides et donc favoriser l'interaction entre l'aptamère et des molécules hydrophobes. Par exemple, ce concept a été exploité par Latham et son équipe lorsqu'ils ont trouvé un aptamère anti-thrombine (cependant avec une faible affinité) en utilisant des dUTP avec un groupement hydrophobe (dU^{pen}TP) greffé sur le cycle pyrimidine (73). D'autres expérimentations ont également été réalisées avec des groupements 5-iodo-UTP pour générer un aptamère anti-HIV-1 (74). L'utilisation de groupements carbohydrates a également été réalisée. Par exemple, le 2'-deoxyuridine triphosphate peut aussi être synthétisé avec divers groupes fonctionnels positionnés en 5' tels que le groupe (benzyl, naphthyl, tryptamino, isobutyl, etc) (75).

Cependant, la plupart de ces modifications ne peuvent pas être introduites au cours du processus SELEX et ne peuvent pas être utilisées pour l'amplification par des enzymes polymérases classiques. L'équipe de Vaught et al (75), a synthétisé un dérivé carboxamide dUTP, modifié sur le carbone C5' et ils ont montré que l'utilisation de polymérases telles que *Thermus aquaticus* (Taq), *Thermus thermophilus* (Tth), n'étaient pas adaptés pour amplifier de l'ADN. Inversement, des enzymes telles que *Pyrococcus* (D.Vent) et *Thermococcus kodakaraensis* (KOD XL) donnent de bons résultats d'amplification et sont moins sensibles. Après avoir testé l'amplification enzymatique de ces dUTP modifiés, ils ont pu les faire réagir contre la protéine TNFRSF9, (tumor necrosis factor receptor super family member 9), protéine qui contribue à l'expansion clonale et pour laquelle la sélection d'aptamères avait échoué. En utilisant des

ADN modifiés, ils ont pu démontrer que la meilleure affinité était obtenue avec des ADN naturels modifiés ($K_D=5nM$ au lieu de $K_D=48nM$).

- Les SOMAmers

Inspirés par l'idée que la diversité chimique des protéines peut être étendue en ajoutant des groupements fonctionnels aux chaînes des acides nucléiques Larry Gold et son équipe ont développés une banque d'aptamères dégénérée comportant en position 5' des déoxyuridines modifiées. C'est ainsi qu'ont été développés les SOMAmers (Slow Off-rate Modified Aptamers) qui se structurent en une architecture compacte et présentent à leur surface une fonction hydrophobe, ce qui contraste fortement avec les interactions polaires, principalement observées dans les aptamères « traditionnels » se liant aux protéines (76). Ceci en fait des dérivés nucléiques particulièrement intéressants puisque le mode d'interaction est proche de celui des anticorps et donc un fort potentiel d'utilisation comme agents thérapeutiques (77). L'avantage majeur des SOMAmers est l'utilisation comme application pour la détection simultanée de plusieurs protéines dans le sérum sanguin pouvant servir à la découverte de marqueurs biologiques de maladies. En l'espace de 5 ans, des SOMAmers ont été identifiés pour 3 000 protéines humaines avec des propriétés uniques ce qui en fait d'excellents outils pour être employés dans les champs du domaine thérapeutique ou du diagnostic. Par exemple, pour résoudre le problème de la barrière de la filtration rénale, des améliorations ont été faites, notamment en créant des aptamères liés à des résidus polyéthylène glycol (PEGylated aptamers). Le fait de greffer des aptamères avec des nanomatériaux avec une bonne biodisponibilité tel que le PEG mais encore des nanoparticules d'or, des liposomes ou des copolymères, augmentent le poids moléculaire du complexe aptamère-PEG et donc sa rétention à travers la barrière rénale. Cela augmente également la résistance aux nucléases et la diminution de l'accumulation toxique dans les tissus non ciblés. Le plus connu des aptamères polyéthylène glycol est le Macugen. Après l'incorporation de dérivés 2'-fluoropyrimidine et 2'-O-méthyl purines et d'un nucléotide inversé en position terminale 3', une molécule de PEG à 40kDa a été conjuguée avec la séquence résultante. Grâce à ce rallongement, le Macugen prolonge ainsi sa résistance avec une demi-vie apparente de 10 jours. De plus aucune réponse indésirable n'a été détectée.

La plupart des SOMAmers contiennent un résidu hydrophobe tel qu'un groupement naphtyl ou benzyl qui augmente le nombre et la force hydrophobes entre les acides nucléiques et leurs cibles et donc l'affinité de l'aptamère pour la cible est améliorée (Figure 2-5).

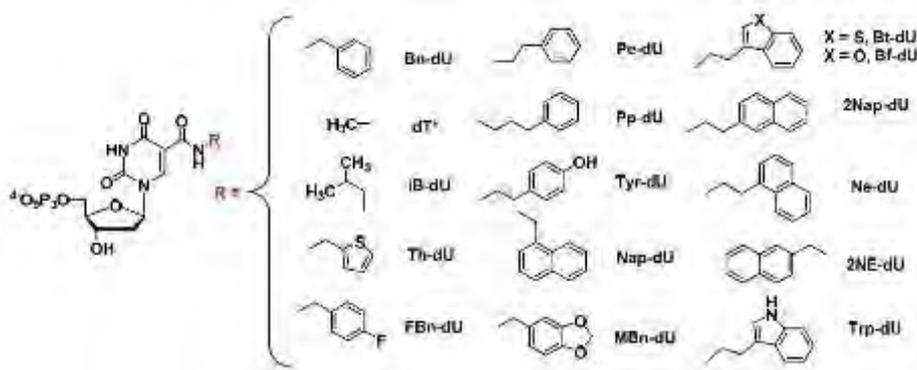


Figure 2-5 : Exemples de SOMAmers. Issu de (65).

Par ailleurs, il a été démontré que les SOMAmers améliorent la résistance aux nucléases. La sélection des SOMAmers implique une étape supplémentaire dans le protocole de sélection. Il

s'agit de l'interaction des SOMAmers liés à la cible avec un large excès de compétiteur polyanionique non-spécifique afin d'augmenter la spécificité de ces derniers pour la cible. Cette étape supplémentaire donne le nom de SOMAscan assay au procédé qui en découle (78).

- Spiegelmers

La modification la plus aboutie puisque trois acides nucléiques sont actuellement en phase d'essai clinique concerne les spiegelmers. Cela reprend le principe de la chiralité des acides nucléiques. Les spiegelmers sont synthétisés avec des nucléotides L-phosphoramidites après que le D-aptamère correspondant ait été identifié contre l'image miroir de la cible correspondante. Si un aptamère naturel (D) est sélectionné contre l'énantiomère de la cible visée, alors l'énantiomère de cet aptamère (L) reconnaîtra la cible naturelle, et sera en outre résistant aux nucléases tels que rapporté par Spiegelmer (79). Les spiegelmers sont utilisés comme de puissants agents anti-viraux grâce à leur grande stabilité. Ils peuvent être aussi utilisés en électrophorèse capillaire pour détecter des impuretés d'énantiomères (**Figure 2-6**). C'est ce qu'a réalisé l'équipe de Peyrin, en démontrant qu'un spiegelmer pouvait être utilisé avec succès pour concevoir un outil énantiosélectif pour la détection de traces d'énantiomères (80). Cet outil développé permet d'obtenir une détection de 0.01% d'un énantiomère minoritaire dans un mélange non-racémique et avec un court temps d'analyse (moins de 5 minutes).

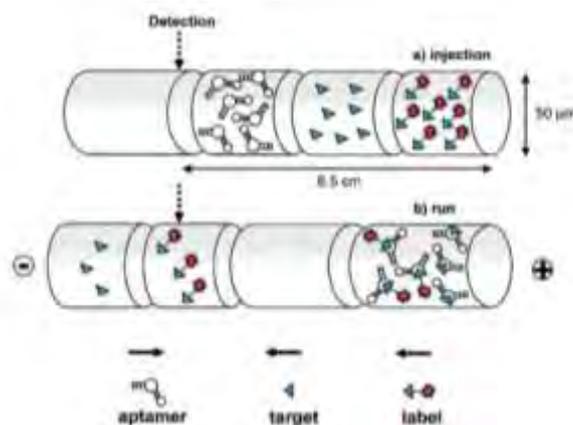


Figure 2-6: Principe de la détection basée sur la CE à l'aide du mélange des différentes espèces injectées individuellement. La première zone contient l'aptamère (L-ARN), la deuxième contient une quantité inconnue de cible : la D-arginine et la troisième contient une quantité connue de D-arginine greffée à un marqueur : le Dansyle. Les flèches indiquent la direction de migration des espèces en interaction lorsque le champ électrique est appliqué. Cette méthode a été améliorée ensuite en pré-incubant le spiegelmer avec l'arginine avant l'interaction avec la cible marquée. Cette amélioration de cette méthode appelée ACE (affinity CE) a permis d'améliorer la sensibilité du dosage par rapport à celle obtenue avec une approche de mélange "pleine" sur capillaire. Issu de (80).

Les 3 spiegelmers en cours de développement ont été développés par la société NOXXON Pharma en Allemagne (81). Il s'agit de l'Olaptesed pegol (NOX-A12) utilisés pour le traitement de la leucémie lymphatique chronique et des myélomes multiples réfractaires en se liant à la chimiokine CXCL-12.

- Les XNAs et les systèmes génétiques étendus

Avec 4ⁿ combinaisons possibles grâce aux 4 nucléotides naturels, A, G, T, C, on a vu combien il est possible de produire une diversité de banque d'oligonucléotides très importante. Cependant, certaines cibles ne sont pas compatibles pour se fixer avec des oligonucléotides chargés négativement.

Pour cela, certains chercheurs ont introduit de nouvelles bases nucléiques dans le but d'élargir le code génétique et de trouver des acides nucléiques avec de nouvelles propriétés (82) (83). Il existe également des formes non-naturelles d'ADN (AEGIS : artificially expanded genetic information systems) construites à partir de 6 nucléotides différents qui permettent de trouver un aptamère (ZAP-2012) avec une constante d'affinité de 30nM. La même expérience a été faite avec une banque d'ADN classique et n'a pas permis d'aboutir à trouver un aptamère pouvant se lier avec des cellules cancéreuses (84).

Hichiro Hirao et son équipe ont développé une nouvelle méthode de SELEX en appliquant une expansion de l'alphabet génétique (utilisation de paires de bases non-naturelles) (83). Ils ont introduit une troisième paire de bases dans le processus de réplication de l'ADN en utilisant deux oligonucléotides non naturels (paire Ds : Px) qui se lient exclusivement l'un à l'autre. Cette modification a abouti à la sélection d'aptamères contre les cibles INF- γ et VEGF-165 avec une grande valeur d'affinité comprise entre 1.69pM et 0.12nM.

Cependant, il est nécessaire que la base non-naturelle Ds soit placée au bon endroit afin de pouvoir faciliter l'amplification et le séquençage.

Une modification de l'unité sucre peut également conduire au développement de structures artificielles nucléiques orthogonales (XNA : Xeno-Nucleic Acids) capables de stocker des informations et de supporter la synthèse enzymatique et la résistance aux nucléases (86). Les Xeno Nucleic Acids sont des oligonucléotides avec un groupement sucre dénaturé qui présentent des caractéristiques similaires à celles des ADN ou ARN aptamères. L'incorporation d'acides nucléiques fermés (LNA : locked nucleic acids) est une autre stratégie de modifications qui a montré son efficacité pour augmenter l'amélioration de la résistance aux ribonucléases et l'affinité de la cible aux aptamères ARN modifiés. Bien que les aptamères ADN soient plus stables que les aptamères ARN en raison de leur absence de groupements 2'-OH, l'incorporation de groupement LNA peut aussi améliorer la bio-stabilité des aptamères ADN. L'équipe d'Hagiwara a utilisé une banque d'oligonucléotides contenant deux modifications dont une LNA et une modification 2'-fluoro appelé FNA (fluoro nucleic acid) afin de sélectionner des aptamères par électrophorèse capillaire-SELEX (CE-SELEX) (87).

Nous avons vu que les techniques de sélection reposent en grande partie sur le SELEX qui a gagné en spécificité et sélectivité grâce aux perfectionnements de différents auteurs. Une orientation importante est d'identifier des modifications chimiques des nucléotides pour arriver à toujours plus de spécificité et sélectivité. Le SELEX repose sur des techniques de purification des complexes acides nucléiques/ligand. Parmi les techniques utilisées, mais encore trop timidement, figure l'électrophorèse capillaire.

2.2 L'électrophorèse capillaire et le SELEX (CE-SELEX) :

Depuis plusieurs années, de nombreuses stratégies de méthodes de sélection par la voie SELEX ont été décrites, notamment la cytométrie de flux (88), la centrifugation ou la filtration sur nitrocellulose (40). D'autres méthodes impliquant l'électrophorèse capillaire ont été décrites par Mendonsa en 2004 (89) ou l'électrophorèse sur gel (EMSA-SELEX) (90). L'électrophorèse capillaire permet de réduire considérablement le nombre de cycles (91) (92). Elle a été publiée pour la première fois en 2004 par Mendosa et son équipe (89). Ils ont pu réduire le nombre de

cycles SELEX à 4 tout en maintenant l'affinité de la cible pour l'aptamère. Elle joue un rôle important dans la détermination des paramètres cinétiques et permet notamment la mesure des constantes d'affinités des aptamères générés. C'est l'une des variantes de la méthode SELEX, la plus utilisée. C'est une méthode de séparation efficace employée couramment pour les analyses d'ADN en raison du temps d'analyse rapide et du faible volume d'échantillon utilisé. Plusieurs modes d'électrophorèse capillaire ont été développées en fonction de l'application visée : la NECEEM (nonequilibrium capillary electrophoresis of equilibrium mixtures), la SweepCE (sweeping capillary electrophoresis) et l'ECEEM (equilibrium capillary electrophoresis of equilibrium mixtures). L'ensemble de ces méthodes a été regroupé sous le nom de KCE (kinetic capillary electrophoresis). Le principe du CE-SELEX est illustré dans la figure suivante (**Figure 2-7**) :

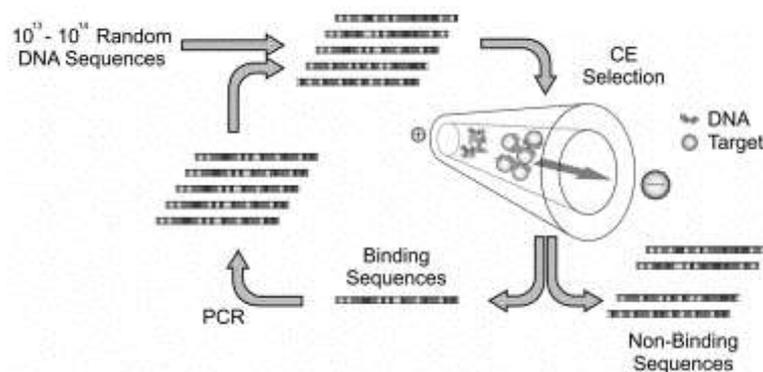


Figure 2-7 : Schéma du principe de l'électrophorèse capillaire (CE)-SELEX pour isoler des aptamères ADN. Issu de (89).

Premièrement, un mélange d'une cible avec une banque d'ADN contenant généralement entre 10^{13} et 10^{14} molécules est préparé puis quelques nanolitres de ce mélange sont injectés en électrophorèse capillaire. La migration et la séparation des espèces se fait par application d'un champ électrique, et les molécules d'ADN non liées à la cible vont avoir la même mobilité électrophorétique. Elles vont donc migrer à la même vitesse tandis que les molécules d'ADN liées à la cible vont avoir un rapport charge sur taille qui va changer en raison de la liaison avec la cible ce qui va augmenter la taille et modifier la charge du complexe et donc aboutir à une mobilité électrophorétique différente. Ainsi, les deux fractions peuvent être séparées et les molécules d'ADN liées à la cible peuvent être collectées en se basant sur les temps de migration. Bien que l'on augmente par CE la sélectivité des molécules d'ADN collectées, il est nécessaire ensuite d'effectuer une étape d'amplification par PCR afin d'amplifier la quantité de molécules collectées et ensuite de régénérer l'ADN simple brin par une étape de purification, afin qu'un nouveau cycle de sélection soit effectué.

D'autres applications issues de la KCE ont pu être créées, dont l'analyse quantitative de l'affinité de protéines, ADN ou ARN, la mesure de la température à l'intérieur du capillaire, et la sélection de ligands à partir d'une banque aléatoire.

2.2.1 Principe général de l'électrophorèse capillaire :

L'électrophorèse capillaire est une technique analytique permettant la séparation d'espèces chargées sous l'effet d'un champ électrique dans un capillaire de faible diamètre (10 à 100 μ m)

(Figure 2-8). Les molécules se déplacent la plupart du temps dans un milieu aqueux appelé électrolyte et servant de milieu conducteur.

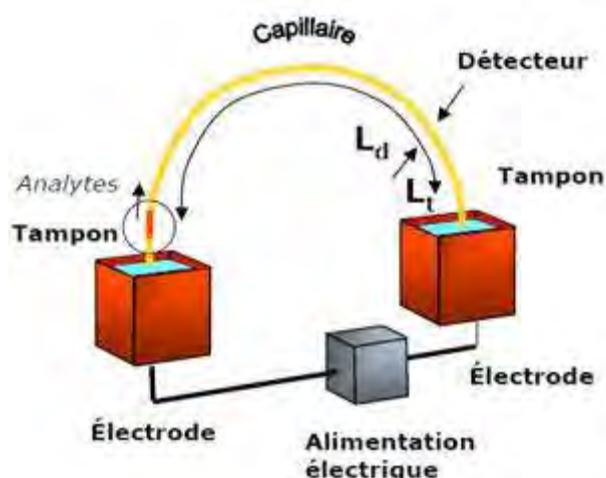


Figure 2-8 : Principe général de l'électrophorèse capillaire (issu de www.esi.umontreal.ca/~badiaa/CHM3102_CE1a.pdf)

La séparation des composés par électrophorèse capillaire résulte de deux mécanismes de transport, l'électromigration et l'électroosmose.

L'électromigration résulte du déplacement d'une espèce chargée lorsqu'elle est soumise à un champ électrique. La vitesse linéaire acquise (ou vitesse électrophorétique) est alors fonction du champ électrique et de la mobilité électrophorétique de l'ion selon la relation suivante :

$$v_{ep} = \mu_{ep} \cdot E \text{ où } v_{ep} = \text{vitesse électrophorétique (cm} \cdot \text{s}^{-1}\text{)}$$

$$\mu_{ep} = \text{mobilité électrophorétique (cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}\text{)}$$

$$E = \text{champ électrique appliqué (V} \cdot \text{cm}^{-1}\text{)}$$

L'électromigration s'effectue dans le sens du champ électrique (mobilité électrophorétique positive) pour les cations et dans le sens opposé pour les anions (mobilité électrophorétique négative). Une molécule neutre n'est pas soumise au phénomène d'électromigration, elle se déplace donc à la vitesse du flux électroosmotique (Figure 2-9).

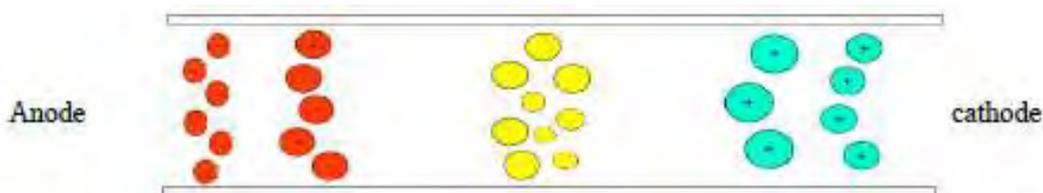


Figure 2-9 : Migration des espèces dans un capillaire.

La mobilité électrophorétique, μ_{ep} , et donc la vitesse électrophorétique, v_{ep} , dépendent du milieu, pH et composition du tampon (force ionique, viscosité, etc) et de la température.

L'électroosmose est responsable du flux électroosmotique. La paroi interne des capillaires de silice fondue est rincée avec de la soude ce qui permet l'hydrolyse des ponts siloxanes en groupements silanols. En fonction du pH de l'électrolyte, ces derniers seront plus ou moins déprotonés ($pK_a=5.5$). Les contre ions positifs de l'électrolyte vont neutraliser la paroi. Ces cations nombreux, attirés par le pôle négatif, entraînent avec eux le liquide : le flux électroosmotique. La vitesse électroosmotique s'exprime selon la relation suivante :

$$v_{eof} = \mu_{eof} * E \quad \text{où } v_{eof} = \text{vitesse électroosmotique (cm.s}^{-1}\text{)}$$

$$\mu_{eof} = \text{mobilité électrophorétique (cm.s}^{-1}\text{)}$$

$$E = \text{champ électrique appliqué (V.cm}^{-1}\text{)}$$

Lorsque le capillaire est rempli d'un électrolyte tampon, les cations du tampon sont attirés vers les charges négatives de la surface du capillaire, formant ainsi une double couche électrique, caractérisée par un potentiel de surface ou potentiel zéta (ζ).

Le flux électroosmotique est fonction des propriétés du tampon (nature, force ionique, pH et viscosité).

En raison des dimensions mêmes du capillaire et des faibles volumes injectés (quelques nanolitres), il n'est pas possible d'employer les méthodes classiquement utilisées en chromatographie pour l'injection des échantillons. Deux modes d'introduction de l'échantillon dans le capillaire existent : l'injection hydrodynamique et l'injection électrocinétique.

Le mode d'injection hydrodynamique consiste, par exemple, à créer pendant une durée de quelques secondes, une surpression au niveau du flacon de l'échantillon, qui le propulse alors à l'intérieur du capillaire. Le volume d'échantillon injecté dépend de plusieurs paramètres : différence de pression appliquée, temps d'injection, viscosité de l'échantillon, longueur et diamètre du capillaire. De même que pour l'injection électrocinétique où la pression est remplacée par la tension. Ce dernier type d'injection est surtout utilisé soit dans les travaux de préconcentration ou d'injection dans des capillaires remplis de gel.

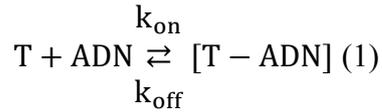
2.2.2 Différentes méthodes d'électrophorèse capillaire :

2.2.2.1 NECEEM

Krylov et son équipe ont décrit la méthode suivante : Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM (93)). C'est une méthode simple et efficace de séparation dans une solution d'une cible liée et non liée avec des oligonucléotides pour la sélection d'aptamères. De plus, le temps de sélection d'un aptamère serait considérablement réduit (temps = 1h). Cela permet la détermination aisée des paramètres cinétiques ; k_{on} ; k_{off} lors de l'interaction entre un aptamère et une protéine. Cette méthode permet de déterminer la valeur de K_D si la concentration de la cible et de l'ADN sont connus ou de déterminer les concentrations de la cible si K_D et la concentration des molécules d'ADN sont connus.

Le principe de cette méthode est basé sur deux processus majeurs :

- L'interaction entre une cible (T), des molécules d'ADN (L) et le complexe aptamère – cible (T. ADN)



k_{on} et k_{off} sont respectivement les constantes de vitesse de la formation et de la dissociation du complexe. Elles définissent la constante de dissociation K_D qui est le rapport des deux constantes de vitesse $k_{\text{off}}/k_{\text{on}}$.

Cette constante K_D , est caractérisée par l'équation suivante :

$$K_D = \frac{[\text{T}]_{\text{eq}}[\text{L}]_{\text{eq}}}{[\text{TL}]_{\text{eq}}}$$

- Et la séparation du complexe aptamère/cible des oligonucléotides non liés, selon la mobilité électrophorétique de chacun.

Dans un premier temps, on prépare un mélange de molécules d'ADN greffée à un fluorophore avec une cible et on incube ce mélange afin d'atteindre l'équilibre ADN-cible (**Figure 2-10**).

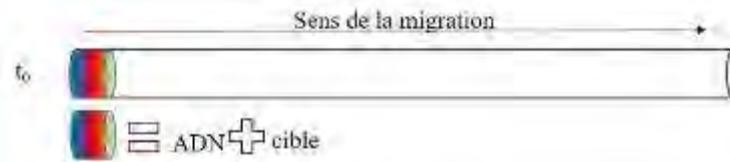


Figure 2-10 : Préparation et injection du mélange ADN + cible par électrophorèse capillaire, d'après (93)

Quand l'équilibre est atteint, le mélange est injecté par électrophorèse capillaire (CE) et séparé en appliquant un champ électrique. La séparation en CE entraîne une séparation des espèces en équilibre. L'équilibre entre les molécules d'ADN liées et non-liées est donc rompu et le complexe ADN-cible commence à se dissocier (**Figure 2-11**).



Figure 2-11 : Séparation du mélange à l'équilibre ADN.cible au début de l'analyse (t_1) par application d'un champ électrique. Le mélange à l'équilibre contient la cible, la cible liée aux molécules d'ADN et les molécules d'ADN libres (cas où $\mu_{\text{cible}} > \mu_{\text{ADN.cible}} > \mu_{\text{ADN}}$), d'après (93)

Les molécules d'ADN qui ont beaucoup d'affinités pour la cible restent liées majoritairement à cette dernière (k_{off} faible). Inversement, les molécules ayant peu d'affinité sont principalement non-liées à la cible (k_{off} important) (**Figure 2-12**).



Figure 2-12 : Distribution dans le temps des différents composés dans le capillaire. Toutes les molécules d'ADN ont la même mobilité électrophorétique, seule la mobilité de la cible est supérieure aux molécules d'ADN. Ainsi, la mobilité du complexe ADN.cible sera intermédiaire à celle de la cible et de l'ADN. Cet équilibre n'est pas maintenu longtemps et les molécules d'ADN.cible commencent à se dissocier, d'après (93)

On suppose que toutes les molécules ont la même mobilité électrophorétique indépendamment de leurs séquences et de leur structure tertiaire et donc ont le même temps de migration. De plus, pour notre exemple, nous considérons que la mobilité de la cible est positive, donc elle migre en premier. Donc nous illustrons notre propos en indiquant que la mobilité du complexe ADN-cible se situera dans une zone intermédiaire entre la cible et les molécules d'ADN libre, soit : $\mu_{\text{cible}} > \mu_{\text{ADN.cible}} > \mu_{\text{ADN}}$.

Il faut noter également que grâce à la grande efficacité de séparation, il n'y a pas de possibilité de re-complexation entre l'ADN dissocié et la cible. L'ordre de migration des composés est donc le suivant : 1° les molécules cibles, 2° les molécules cibles dissociées de l'ADN, 3° le complexe ADN-cible, 4° les molécules d'ADN dissociées de la cible et 5° les molécules d'ADN libres.

A partir de la connaissance des temps de migration, une fraction contenant le complexe ADN.cible est collectée. Connaissant les concentrations de la protéine et de l'ADN, la constante de dissociation K_D peut être déterminée. Inversement, à partir d'une constante de dissociation connue, on peut déterminer la concentration de la protéine. Les molécules d'ADN ayant un K_D inférieur à la concentration de la cible $[T_1]$ sont principalement liées à la cible tandis que celles $[T_2]$ avec un K_D supérieur à la concentration de la cible sont non-liées. Deux types de sélection peuvent être réalisés : positive ou négative. On appelle sélection positive, la collecte des molécules d'ADN correspondantes au complexe ADN-cible. Cette méthode est utilisée pour sélectionner les aptamères qui sont liés spécifiquement à la cible. Donc, dans ce cas, les molécules d'ADN ont une constante de dissociation K_D inférieure à la concentration de la cible. De plus, en fonction de la fenêtre de sélection utilisée pour la sélection positive (t_1 à t_2), on peut déterminer précisément la valeur de k_{off} selon l'équation suivante :

$$k_{\text{off}} = \frac{t_{\text{ADN}} - t_{\text{ADN.T}}}{t_{\text{ADN.T}} * (t_1 - t_2)} * \ln \left(\frac{t_{\text{ADN}} - t_2}{t_{\text{ADN}} - t_1} \right) \quad (2)$$

t_{ADN} et $t_{\text{ADN.T}}$ correspondent aux temps de migrations respectifs de l'ADN libre et du complexe ADN.cible.

t_1 et t_2 correspondent aux temps de migration des composants séparés dans le capillaire à différents moments ($t_0=0$, $t_2 > t_1 > t_0$) depuis le début de la séparation.

L'efficacité de séparation ne dépend pas seulement de la résolution entre le complexe ADN.cible et les molécules d'ADN libres mais également de l'emplacement de la fenêtre de sélection, du fait du décalage des temps de migration ou de la réassociation du complexe.

La sélection négative correspond à la collecte de la partie à l'équilibre de l'ADN libre. La collecte de cette fraction peut être intéressante pour éliminer les aptamères non spécifiques présents dans le tampon qui possèdent donc un K_D supérieur à la concentration de la cible. L'alternance de l'utilisation de ces deux modes de sélection conduit à la sélection d'aptamères dans une gamme d'affinité limitée, définie par : $[T_2] < K_D < [T_1]$ (**Figure 2-13**).

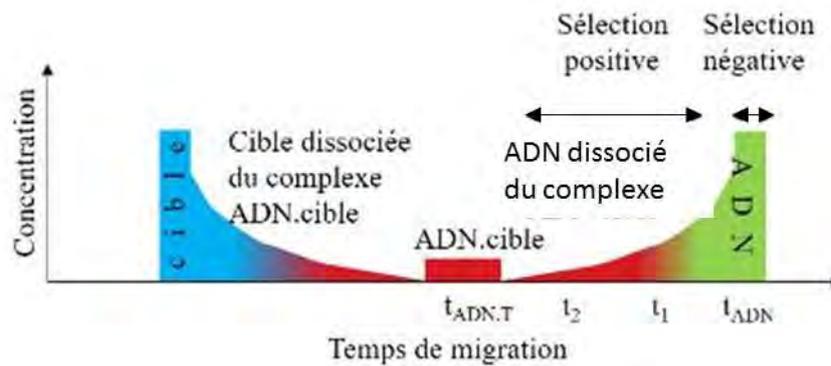


Figure 2-13 : Collecte des aptamères à la sortie du capillaire. La sélection positive est définie comme la collecte de la fraction à l'équilibre du complexe ADN.cible. La sélection négative correspond à la collecte des molécules d'ADN à l'équilibre non-liées. D'après (93).

2.2.2.2 SweepCE :

Le concept du Sweeping Capillary Electrophoresis est complémentaire de la sélection par NECEEM. Cette dernière permet de mesurer la constante de vitesse monomoléculaire k_{on} ainsi que la constante de dissociation K_D .

Le concept est basé sur la migration d'ADN par balayage (sweeping) continu d'une solution de protéine durant l'électrophorèse (94). Un capillaire est rempli d'une solution d'ADN (Figure 2-14), puis une solution de protéine est injectée par électrophorèse capillaire (Figure 2-15).

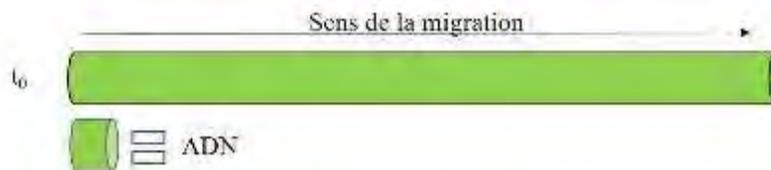


Figure 2-14 : Injection et remplissage du capillaire par une solution d'ADN greffée à un fluorophore. En l'absence de protéine, l'ADN s'écoule continuellement à l'outlet du capillaire tandis qu'à l'inlet, de l'ADN est injecté en continu. D'après (94).



Figure 2-15 : Injection de la cible. D'après (94) .

Du fait de la plus grande mobilité électrophorétique des protéines par rapport à l'ADN, ces dernières se mélangent continuellement avec l'ADN et forment les complexes ADN.protéines à l'équilibre (Figure 2-16). Le complexe migre avec une vitesse supérieure à celle de la banque d'ADN libre et provoque le balayage de l'ADN qui interagit avec la cible, il est donc préconcentré sur la cible.



Figure 2-16 : Schéma décrivant le phénomène de SweepCE à l'intérieur du capillaire. Comme la mobilité électrophorétique de la cible est supérieure à celle des molécules d'ADN, elle sera continuellement balayée avec les molécules d'ADN et formera les complexes ADN.cible. D'après (94) .

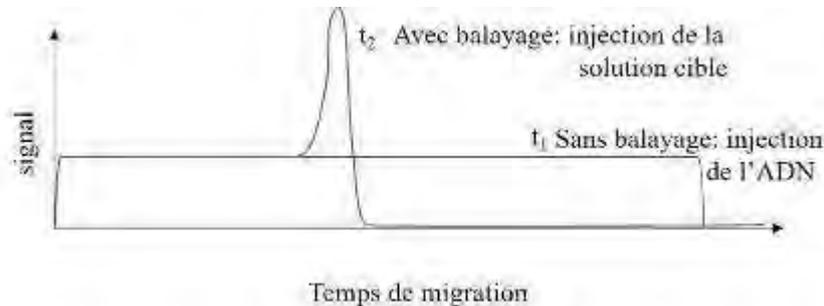


Figure 2-17 : Electrophérogramme décrivant en rouge le phénomène de balayage de l'ADN par injection d'une solution cible. Le contrôle de non balayage est décrit dans l'électrophérogramme (en rouge) où une solution d'ADN est injectée et remplit le capillaire. D'après (94) .

La valeur de la constante de vitesse de formation du complexe k_{on} peut être déterminée à partir du profil de concentration de l'ADN (**Figure 2-17**). Avec cette méthode, de faibles concentrations de réactifs peuvent être utilisés, permettant des mesures faibles de k_{on} de $10^9 \text{ M}^{-1} \text{ s}^{-1}$. Cette méthode donne non-seulement la constante de dissociation K_D mais aussi la constante de formation du complexe k_{on} qui permet d'étendre la connaissance de l'interaction du complexe aptamère-cible. La limitation de la méthode est évidemment les interactions non sélectives.

2.2.2.3 ECEEM

L'ECEEM (Equilibrium Capillary Electrophoresis of Equilibrium Mixtures) est une méthode de sélection des aptamères basée sur critères cinétiques ADN.cible prédéfinis (**Figure 2-18**). C'est la seule méthode pour laquelle l'équilibre ADN.cible est maintenu pendant la séparation du complexe ADN.cible / ADN libre par électrophorèse dans le capillaire (95). Pour maintenir l'équilibre, il faut que la cible soit présente dans le capillaire ce qui augmente la consommation de réactifs contrairement à la méthode NECEEM.

Dans cette approche, l'équilibre entre les molécules qui interagissent est maintenu pendant l'électrophorèse capillaire en ajoutant dans le tampon de migration l'un des deux composants. Un mélange ADN.cible à l'équilibre est préparé et injecté dans un capillaire pré-rempli avec un tampon de migration contenant les protéines à une concentration identique à celle du mélange. Les composants du mélange à l'équilibre sont séparés par électrophorèse capillaire pendant que l'équilibre est maintenu entre la cible et les aptamères. A cause de l'interaction dynamique entre les aptamères libres et complexés durant la séparation par ECEEM, les aptamères passent un certain temps complexés et d'autre fois sont sous forme libre non liés à la cible. Ceux qui ont la valeur de K_D la plus faible sont ceux qui sont sous forme complexés et leur temps de migration est proche de celui des molécules d'ADN complexées. Inversement, ceux qui ont la valeur la

plus élevée de K_D sont ceux qui sont sous forme libre et ont donc un temps de migration proche de celui de la banque d'ADN libre.

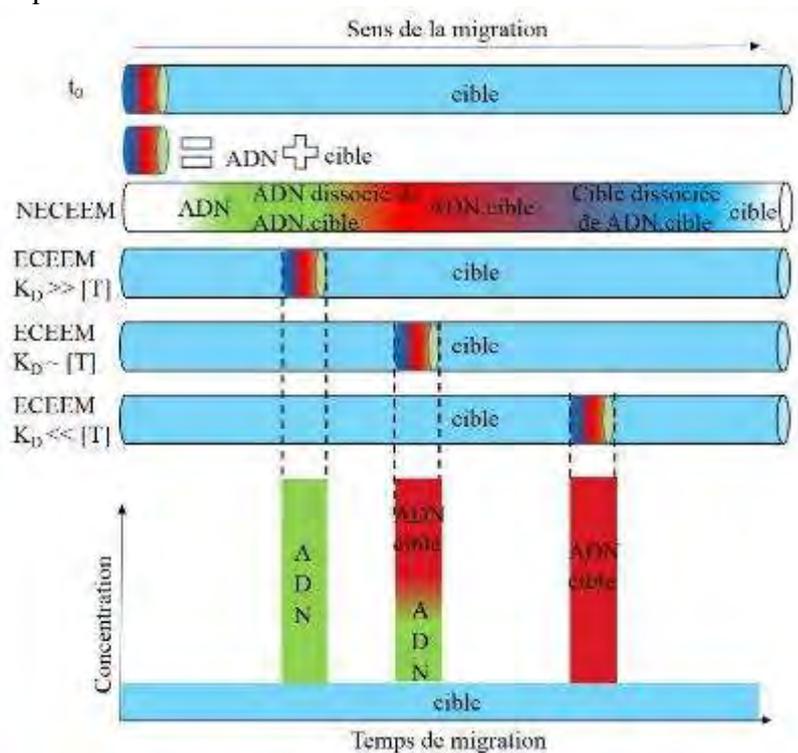


Figure 2-18 : Principe de l'ECEEM. La distribution spatiale des composés est montrée à deux instants. A l'instant t_0 correspondant à l'injection du mélange à l'équilibre, et à trois différents moments : $K_D \gg [T]$, $K_D \sim [T]$ et $K_D \ll [T]$; $[T]$ correspondant à la concentration de la cible. Si $K_D \gg [T]$, la plupart des molécules d'ADN seront à l'état libre et la zone du mélange à l'équilibre migrera avec la même vitesse effective que l'ADN libre dans le cas de la NECEEM. Si $K_D \ll [T]$, la plupart des molécules d'ADN seront complexées avec la cible et le mélange migrera avec la même vitesse que celle du mélange ADN.cible. Enfin, si $K_D \sim [T]$, les quantités de cible et de complexe seront similaires et la vitesse effective du mélange à l'équilibre sera intermédiaire entre celle de la cible et du complexe en NECEEM.

Le temps de migration des aptamères par la méthode ECEEM dépend de la concentration de la cible ainsi que de la valeur de K_D .

Cette dernière est donnée par la formule suivante :

$$K_D = [T] * \frac{t_{DNA} t - t_{T.DNA}}{t_{T.DNA} t_{DNA} - t}$$

Où t , t_{DNA} et $t_{T.DNA}$ correspondent au temps de migration respectivement de la cible, de l'ADN et du complexe [ADN.cible].

Un électrophérogramme en ECEEM contient un plateau correspondant à la cible avec un pic unique correspondant au mélange à l'équilibre dont le temps de migration dépend de K_D et de la concentration de la cible $[T]$.

$$\frac{1}{t} = \frac{1}{t_{T.DNA}} \frac{[T]}{[T] + K_D} + \frac{1}{t} \frac{K_D}{[T] + K_D}$$

Expérimentalement, cette méthode a été appliquée à la sélection d'aptamères avec un K_D déterminé préalablement par la méthode NECEEM. La cible utilisée était la protéine MutS. La banque d'ADN utilisée était dégénérée sur 39 nucléotides. La méthode NECEEM décrite précédemment a été utilisée pour trouver les temps de migrations des molécules d'ADN complexées et libres. 100nM de protéines MutS et $9 \cdot 10^{11}$ molécules de banques d'ADN ont été

injectées dans le capillaire. 3 cycles d'injections ont été réalisés. Ce nombre suffit à obtenir une valeur de K_D prévue. Les aptamères collectés ont été amplifiés par PCR afin d'obtenir des molécules d'ADN enrichies (95).

2.2.3 Applications de la CE-SELEX :

La CE-SELEX peut être appliquée dans la détermination des constantes de vitesses de dissociation et de formation, l'étude des interactions moléculaires, la détermination de la stoechiométrie de liaison. Dans le tableau suivant (**Tableau 2**), est comparé les constantes de dissociations pour des aptamères ayant été obtenus à partir de la méthode CE-SELEX et de la méthode SELEX.

Cible	CE-SELEX		SELEX		Réf
	K_D (nM)	Cycles	K_D (nM)	Cycles	
IgE	23 ± 12	4	6	15	(89)
HIVRT	0.18 ± 0.07	4	1	12	(91)
Neuropeptide Y	300 ± 200	4	370	12	(96)
PFTase	0.5	1	/		(93)
MutS	3.6 ± 0.5	3	/		(97)
Kinase C- δ	122	9	/	-	(98)
Ricine	58 ± 19	4	105 ± 41	9	(99)
NMM	$1.2 \pm 0.1 \mu\text{M}$	3	/		(100)
HE4	390nM	5			(101)
Toxine anthrax	112nM	6			(102)
Alphafetoprotéine	$17.41 \pm 4.32 \text{ nM}$	4	2.37nM		(103)
VEGF165	$50 \pm 60 \text{ nM}$	7			(104)
Virus H9N2 (A9)	46.23 ± 5.46	4			(105)
H9N2 (B4)	7.38 ± 1.09	4			(105)
Leptine	$0.4 \mu\text{M}$	4			(106)
Ara h 1	$450 \pm 60 \text{ nM}$	8			(107)

Tableau 3: Comparaison des aptamères sélectionnés par CE-SELEX et la méthode conventionnelle SELEX et détermination des K_D trouvés pour les aptamères obtenus par ces deux méthodes.

Comme l'indique ce tableau, la CE-SELEX a déjà été adoptée par plusieurs groupes pour identifier les aptamères envers un certain nombre de cibles. Introduit par Mendosa et son équipe, la CE-SELEX a été utilisée pour sélection des aptamères pour l'IgE humaine (200kDa) (89), la transcriptase inverse du virus de l'immunodéficience humaine du type 1 (VIH-1) (91), la protéine farnésyltransférase (45-48 kDa) (93), la toxine ricine (99), la protéine MutS (93), la protéine kinase C- δ , (98). La CE-SELEX permet de réduire la procédure de sélection des aptamères en raccourcissant ce procédé de plusieurs semaines à quelques jours (2 jours pour les aptamères anti IgE), (4 jours pour les deux aptamères A9 et B4 du virus H9N2). Krylov et son équipe sont mêmes parvenus à réduire cette méthode à un seul cycle de sélection pour les aptamères contre la protéine farnésyltransférase (93). L'efficacité de séparation par CE-SELEX permet de réduire le nombre de cycles SELEX de plus de 10 cycles avec la méthode conventionnelle SELEX à 5 -6 cycles en CE-SELEX. Les valeurs de K_D obtenues sont de l'ordre de la dizaine de nanomolaire, à l'exception de la transcriptase inverse du virus de l'immunodéficience humaine du type 1 (VIH-1) pour laquelle une valeur de 180pM a été

trouvée. La CE-SELEX a également été appliquée à des protéines plus petites (le neuropeptide Y=4 kDa), (96), pour laquelle la sélection est plus compliquée en raison du faible poids moléculaire de la cible donc une vitesse de migration du complexe [ADN.cible] proche de celle de la banque d'ADN libre. La sélectivité de la procédure CE-SELEX a été confirmée car bien souvent dans la méthode SELEX, les aptamères sélectionnés sont ceux qui sont liés au support et non à la cible. En présence d'albumine de sérums humains (BSA) et de caséine, Tang et son équipe n'ont pas observé d'interactions entre les aptamères sélectionnés et ces molécules (99). Par ailleurs, la même équipe (99) a mis en évidence qu'au bout de 4 cycles de sélection CE-SELEX, 87,2% de l'ensemble de l'échantillon avait une affinité pour la ricine mais seulement 38.5% du pool avait ensuite une affinité au bout de 9 cycles de sélections. Depuis, la méthode CE-SELEX a continué à être développée et grâce au développement des techniques de séquençage haut-débit, il est plus aisé d'obtenir des centaines de millions de séquences qui peuvent être testées pour leur capacité de liaison avec la cible. C'est ainsi, que plutôt que d'utiliser le clonage comme dans le cas de la sélection d'aptamère contre un allergène de la cacahuète (Ara h 1) (107), la toxine Anthrax (102) ou l'alpha-fétoprotéine (103), Whelan et son équipe (101), ont utilisé la technique de l'Illumina pour séquencer des aptamères pour le biomarqueur du cancer ovarien HE4, ce qui leur a permis à partir de $6.4 \cdot 10^6$ séquences, d'identifier des biais de séquençage et de regrouper les séquences dupliquées sous forme de familles avant d'aboutir à la sélection de 4 séquences d'aptamères jusqu'à l'identification de la séquence présentant la meilleure affinité. D'autres ont utilisé le pyroséquençage 454 (104), ce qui leur a permis d'obtenir $3 \cdot 10^4$ séquences à partir d'une banque d'ADN et d'une cible la protéine VEGF₁₆₅. Ils ont mis en évidence, l'hétérogénéité des séquences obtenues par la méthode CE-SELEX, puisque le motif le plus présent a été trouvé mais avec une abondance de seulement 0.8%, mais ont également observé l'évolution de la formation des sous-produits qui augmente au cours des cycles 6 et 7 de CE-SELEX. Ces sous-produits sont issus d'une erreur d'amplification par PCR. Ainsi, la CE-SELEX fournit un plus grand nombre d'aptamères indépendants par rapport à la méthode SELEX conventionnelle mais à cause du faible diamètre du capillaire, un nombre limité de molécules est injecté et donc seule une petite proportion de la banque d'ADN peut être séparée par électrophorèse capillaire.

2.3 L'électrophorèse capillaire et le non-SELEX :

Avec la méthode non-SELEX, on élimine l'étape intermédiaire d'amplification par PCR entre chaque cycle de sélection (**Figure 2-19**). Cette méthode est une variante dérivée de la méthode CE-SELEX qui implique plusieurs étapes répétitives de séparation par électrophorèse capillaire sans étape intermédiaire d'amplification par PCR entre chaque cycle de sélection. C'est en 2006, que Berezovski et son équipe ont utilisé cette méthode pour la première fois en électrophorèse capillaire (108). Ils ont ainsi pu sélectionner les aptamères pour la protéine

hRAS en effectuant 3 cycles de sélection par électrophorèse capillaire sans étape d'amplification par PCR entre les séparations.

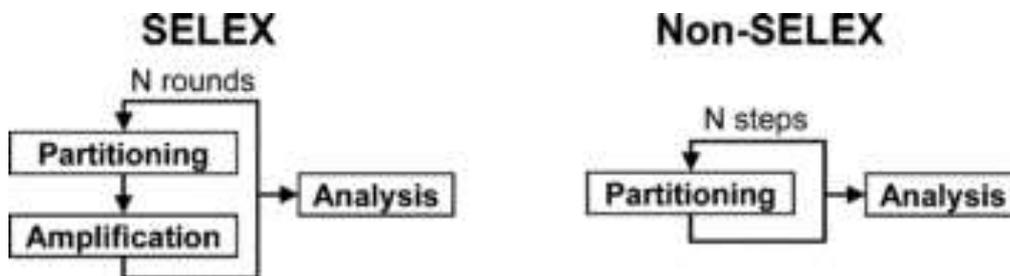


Figure 2-19 : Comparaison entre la méthode SELEX où plusieurs cycles de séparation et d'amplification sont réalisés et le non-SELEX où plusieurs cycles de séparations sans amplification sont réalisés. Issu de (108).

Un mélange de banque d'ADN et de molécules cibles est injectée par électrophorèse capillaire. Comme la quantité injectée est faible, la diversité du nombre de molécules d'ADN est plus faible en comparaison avec la méthode SELEX conventionnelle. Pour remédier à cela, l'équipe de Li (109) a optimisé les conditions d'injection en utilisant des capillaires de diamètre interne de 100 μ m. En augmentant le diamètre interne du capillaire, le volume d'injection est plus important (411 nL avec un capillaire de 100 μ m au lieu de 150nL avec un capillaire de 75 μ m) et donc le nombre de molécules injectées également (10¹³ molécules avec un capillaire de 100 μ m de diamètre et 2*10¹² molécules avec un capillaire de 75 μ m de diamètre) (Figure 2-20).

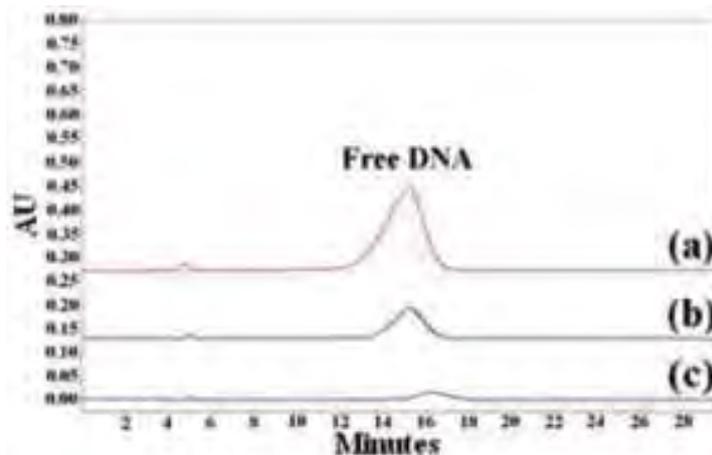


Figure 2-20 : Optimisation du volume du capillaire et donc du nombre de séquences injectées dans la méthode CE/non-SELEX. (a) Diamètre de 100 μ m, (b) diamètre de 75 μ m, (c) diamètre 50 μ m. Tampon de migration : tris (hydroxyamino) methane-glycine-potassium 3X, 333Vcm⁻¹, [ADN]= 50 μ M, injection 1 psi, 13s. Détection UV à 260 nm Issu de (109).

La fraction contenant le complexe [ADN.cible] est ensuite collectée dans un flacon contenant quelques microlitres de la protéine (5 μ L de protéine hRAS, 30 μ L de bovine catalase). Puis ce mélange est de nouveau injecté par électrophorèse capillaire et soumis à une seconde séparation. Une nouvelle fraction est de nouveau collectée, incubée avec la protéine contenue dans le flacon de collecte et ce mélange est réinjecté (Figure 2-21). Dans le cas de la protéine h-RAS, le calcul de la constante de dissociation K_D a été effectué pour chaque fraction collectée et l'affinité au cours des cycles de sélection s'est améliorée passant de 10² μ M à 0.3 μ M.

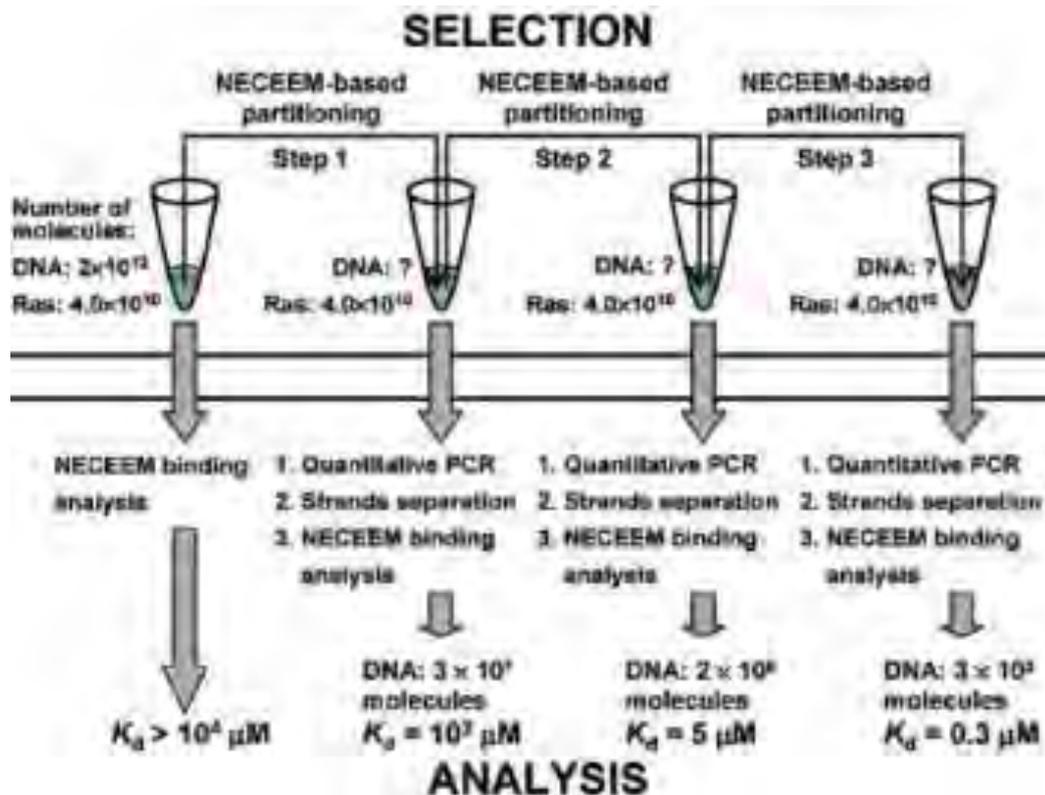


Figure 2-21 : Protocole de sélection d'aptamères par non-SELEX avec trois étapes d'injections Issu de (108).

Le mode de stockage des composés est aussi un paramètre important. Lors de la sélection d'aptamères pour les protéines de transduction du signal, les fractions collectées ont été stockées à -20°C avant leurs analyses (110). En fonction du type de protéines, les conditions de stockages et d'incubations varient. Pour les protéines de transduction du signal, des temps d'incubation d'une heure ont été utilisés à température ambiante afin d'assurer une forte stringence du processus de sélection ou seuls les aptamères qui peuvent se lier rapidement donc avec une sélectivité élevée ont été sélectionnés. Mais les temps d'incubation plus court ($t_{\text{incubation}} = 30\text{min}$ dans le cas de la sélection des aptamères pour la bovine catalase) permettent de maintenir l'activité de la protéine et de s'assurer que l'on sélectionne des aptamères se liant à la protéine et non des impuretés. Une autre approche a été proposée encore par l'équipe de Li (111) où une étape préliminaire de séparation sur filtre de nitrocellulose a lieu avant d'effectuer la séparation CE/ non-SELEX. Grâce à cette étape, cette méthode leur permet d'éliminer toutes les molécules d'ADN qui ne se lient pas à la cible, et qui sont retenues sur le filtre. Ils ont ainsi passé au crible, 10^{15} molécules, ce qui a permis de réduire la complexité de la banque d'ADN. Une des contraintes de cette méthode est qu'il faut tenir compte du facteur de dilution désigné par $f = (L_{\text{totale}} / L_{\text{détecteur}}) * t_{\text{migration du complexe}}$, où L_{totale} est la longueur totale du capillaire et $L_{\text{détecteur}}$ la longueur du capillaire depuis l'entrée de l'injection jusqu'à la fenêtre de détection. $t_{\text{migration du complexe}}$ représente le temps de migration du complexe et grâce à ce calcul, ils peuvent déterminer à quel moment la fraction de complexe sera éluée. Globalement, l'élimination de l'étape de PCR permet néanmoins de réduire considérablement le temps requis pour obtenir des aptamères puisque c'est cette étape qui est la plus longue du processus (112). En utilisant la méthode appelée NECEEM based non-SELEX, la sélection n'a duré qu'une heure et l'affinité de la protéine à l'aptamère est 4 fois plus importante dans le cas de la protéine hRAS. Enfin présentons un travail original conduit par Fiore et al (113). Les auteurs ont développé un

système de deux aptamères capables de se fixer en sandwich à l'Adénosine. L'un de ces aptamères est fluorescent tandis que l'autre est modifié en 5' par une fonction cholestérol. Une modification de la mobilité de ce dernier peut être ajustée lorsque le tampon de séparation en CE contient un tensio-actif. Ainsi deux types de séparation rapide utilisant comme tensio-actif neutre le Brij 35, sont conduites et permettent de séparer le complexe sandwich et les réactifs seuls. Le premier type de séparation ressemble grandement aux résultats du NECEEM, dont on sait que sa limite repose sur la dissociation du complexe durant la séparation en CE. Pour le deuxième type de séparation, les auteurs ajustent le flux électroosmotique afin qu'il permette de diriger le complexe (qui détient la partie cholestérol) et l'aptamère fluorescent non lié en sens opposé l'un de l'autre. Les auteurs collectent ainsi les complexes générés. La qPCR avec les amorces correspondant à l'aptamère fluorescent permet de vérifier que le complexe a bien été collecté et de doser son abondance en fonction de la quantité d'adénosine utilisé.

2.4 Conclusion :

Bien que la CE-SELEX et la CE-non-SELEX se révèlent être des moyens rapides et efficaces pour sélectionner des aptamères, il existe des limites à ces applications. Par ailleurs, il faut être sûr de la pureté de la cible pour laquelle on souhaite sélectionner des aptamères afin de ne pas identifier des aptamères liés à une impureté. Cependant, les protocoles de CE-SELEX et CE-non-SELEX peuvent être réalisés en une semaine, une fois les conditions de séparation optimisées : diamètre du capillaire, nature des tampons de migrations, concentrations des sels dans le tampon d'élution. En fonction de la cible, ces conditions peuvent évoluer. Si le tampon de migration a une concentration en sels trop élevées, cela va provoquer un échauffement du capillaire lors de l'application de la tension et donc entraîner une modification du flux électroosmotique et donc des temps de migration de la cible et des molécules d'ADN. Bien que ce procédé soit souvent utilisé, il n'y a pas de procédé universel pour l'analyse des aptamères. D'autres inconvénients comme le faible nombre de molécules injectées (correspondant à une quantité de quelques dizaines à centaines de nanolitres) limitent la probabilité de présence d'aptamères. Mais ces limites peuvent être surmontées en utilisant la technique d'électrophorèse à flux libre (micro free flow electrophoresis - μ FFE) dans laquelle la séparation est réalisée dans un canal avec un flux continu d'analytes (**Figure 2-22**). Contrairement à la CE où le champ électrique est appliqué en direction du mouvement du fluide, dans le μ FFE le champ électrique est appliqué perpendiculairement à l'écoulement sous pression afin de dévier les analytes latéralement selon leur mobilité (114).

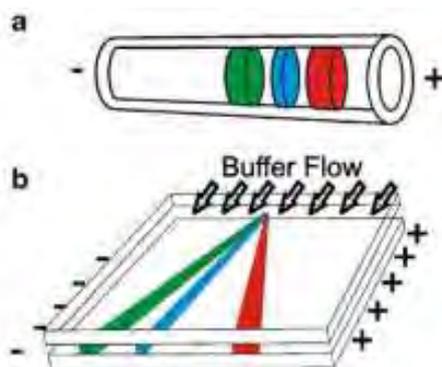


Figure 2-22 : Principe de la micro free flow electrophoresis - μ FFE. La figure en haut montre la séparation par électrophorèse capillaire. La figure du bas montre la séparation par μ FFE. Sous l'effet d'un flux de pression, les analytes se déplacent vers le bas de la puce et leur déplacement se fait sous l'effet du champ électrique basé sur leur mobilité électrophorétique.

Avec cette méthode, des aptamères avec des faibles constantes de dissociation (nM) ont pu être détectés lors d'un seul cycle de sélection. L'ensemble des techniques d'électrophorèse a pour avantages d'avoir réduit les temps de sélection des aptamères.

En conclusion, l'électrophorèse capillaire est un outil qui est peu utilisé pour la sélection d'aptamères. Le nombre faible d'études par rapport au nombre total d'études sur la sélection d'aptamères montre que sa limitation provient de la technique elle-même, souvent réservée à des spécialistes. Elle a du mal à progresser dans les milieux de biologistes. Mais les travaux ici présentés montrent sa pertinence, reposant sur la rapidité et l'automatisme de l'instrumentation. Par ailleurs on voit bien que tant en SELEX ou en CE SELEX ou CE non-SELEX, la preuve première de la sélection d'un aptamère repose sur la mesure d'un K_D de faible valeur.

3 TECHNIQUES DE MESURE DE LA CONSTANTE DE DISSOCIATION K_D A PARTIR DE L'ELECTROPHORESE CAPILLAIRE :

3.1 Techniques de séparation :

Depuis l'essor des aptamères, plusieurs techniques pour caractériser l'interaction entre un aptamère et une cible ont été étudiées. Les méthodes de détection utilisées doivent donc être sensibles. Pour comparer les avantages et inconvénients de chacune d'entre elles, l'équipe de DeRosa (115), a étudié les méthodes qui caractérisaient la liaison des aptamères à la cible ochratoxine, une petite molécule. Plusieurs techniques telles que la dialyse, l'électrophorèse sur gel, l'électrophorèse capillaire, la résonance plasmonique de surface existent parmi tant d'autres (116). Bowser et son équipe (116), ont répertorié ces différentes techniques de séparation :

Technique	Quantité d'échantillon	Durée et complexité de l'expérience	K_D	Equipements
Dialyse	100 μ L	\sim 48h	10^{-8} M	Appareil à dialyse et fluorimètre

Filtration sur nitrocellulose	200-500 μ L	~5min de séparation Marquage radioactif et comptage nécessaire	10^{-11} M à 10^{-12} M	Centrifugeuse, filtres, systèmes de filtration, compteur à scintillation
Electrophorèse sur gel	10 -50 μ L	~3h de séparation. Réalisation du gel, marquage radioactif, l'autoradiographie, découpage et le comptage sont nécessaires	10^{-13} M	Alimentation électrique et autoradiographie
CE	~10 μ L	~ 10 min	10^{-9} M	CE et détecteur LIF
HPLC	Fonction du type de colonne ~10 μ L à 18 mL	10 – 30 min par injection d'échantillons	10^{-6} M	HPLC et détecteur UV
Intensité de fluorescence	~150 μ L	~10 min par échantillon + temps d'attente entre chaque dosage	10^{-10} M	Spectrofluorimètre et cuvettes
Anisotropie de fluorescence	15 – 25 μ L ~150 μ L	~10 min ~10 min par échantillon + temps d'attente entre chaque mesure	10^{-9} M	Lecteur de microplaque et spectrofluorimètre avec polariseur et cuvettes à quartz
Absorption UV-visible	15 – 25 μ L ~2.5 mL	~10 min ~10 min par échantillon + temps d'attente entre chaque mesure	10^{-6} M	Spectrophotomètre, cuvettes à quartz
SPR (résonance plasmonique de surface)	10 – 20 μ L	~20 min par échantillon plus l'étape d'immobilisation de l'aptamère	10^{-12} M	Instrument biacore et puce à capteur
ITC (Titration calorimétrique isotherme)	~1.5mL dans une cellule d'échantillon 200 – 500 μ L volume d'injection total	1.5 – 3 h	10^{-8} M à 10^{-9} M	Calorimètre ITC

La variété de techniques utilisées sur cette petite molécule étant si importante, qu'elle nous permet d'entrevoir les différentes techniques utilisées dans ce type d'étude.

3.1.1 La dialyse :

Cette méthode de séparation est basée sur l'utilisation d'une membrane semi-perméable qui va séparer les composants en fonction de leur taille. D'un côté on place le mélange d'aptamères avec la cible tandis que de l'autre côté de la membrane, il y a du tampon. Si l'aptamère est lié à la cible, il ne traversera pas la membrane, seul l'aptamère libre pourra la traverser.

Cependant, cette technique simple d'utilisation a un certain nombre d'inconvénients (116). Le potentiel chimique de deux compartiments à l'équilibre est équilibré lorsque l'activité des analytes est égale, ce qui peut être significativement différent des concentrations d'analyte. La charge nette de la cible protéique est équilibrée par des contre-anticorps H⁺ ou OH⁻, donnant lieu à un gradient de pH et des concentrations différentes de ligands à travers la membrane. L'imperméabilité de la cible et du complexe donne lieu à un gradient de pression osmotique, entraînant un écoulement de solvant à travers la membrane et des erreurs dans les volumes de solvant. Des problèmes rencontrés dans les mesures de dialyse comprennent un temps d'incubation prolongé (48 heures dans le cas de l'étude de l'ochratoxine (115)), des besoins en volume important (300µL de chaque analytes (115)), et une adsorption non spécifique. Dans leur étude des aptamères de la tyramine, (117), ont utilisé la technique de la dialyse pour caractériser les aptamères obtenus pour SELEX. Ils ont comparé la valeur de la constante de dissociation trouvée avec une autre méthode : la thermophorèse à microéchelle. Pour deux aptamères sélectionnés, ils ont obtenu les valeurs suivantes :

K _D	Dialyse	Thermophorèse à microéchelle
Tyr_10*	0.2 ± 0.4 µM	0.097 ± 0.037 µM
Tyr_14**	0.9 ± 0.6 µM	1.25 ± 0.47 µM
*Tyr_10 correspond à la séquence de l'aptamère suivant : TCATGGTGGGGTTGGGCCTGGGAAGGGGGGAACCGTTT		
**Tyr_14 correspond à la séquence de l'aptamère suivant : CCACGGGGGGCTGGGCCTGGGAAGGGGAGGTACCGTGT		

Pour l'équipe de DeRosa (115), les valeurs obtenues sur l'ochratoxine étaient plus variables. Ils ont comparé dans un premier temps, les valeurs de K_D obtenues pour 3 techniques standards, la dialyse, l'ultrafiltration et la chromatographie d'affinité sur une résine d'agarose ou des billes magnétiques.

K _D	Dialyse	Ultrafiltration	Chromatographie d'affinité
OTA aptamer 1.12.2	287 ± 56 nM	255 ± 49 nM	374 ± 255 nM

Le protocole d'analyse de l'ochratoxine par dialyse utilisé par l'équipe de DeRosa a été adapté de celui de l'équipe de Penner dans lequel ils ont mis au point une méthode de détection de l'ochratoxine à partir d'un aptamère (118). D'un côté de la membrane, la chambre de chargement contient 300 µL d'une solution d'aptamères avec une concentration variable d'ochratoxine (100 nM), tandis que de l'autre côté de la membrane, la chambre réceptrice contient la même quantité de tampon. L'incubation se fait à température ambiante, sous agitation, et ceci pendant 48heures. Au total, dix concentrations d'aptamères variant de 12.5nM à 5µM ont été réalisées et dupliquées. A la fin de l'incubation, 50µL de solution contenant le complexe aptamère.ochratoxine et l'ochratoxine non liée dans la chambre de chargement, et 50µL de solution dans la chambre réceptrice contenant l'ochratoxine non liée sont prélevées afin d'être analysés par HPLC. Le pourcentage d'ochratoxine liée à l'aptamère (f), est déterminé selon la formule suivante :

$$f = \frac{F_l - F_r}{F_l} * 100$$

F_l et F_r représentent respectivement les aires des pics de fractions d'ochratoxine dans la chambre de chargement et la chambre réceptrice.

Ils ont ensuite tracé la courbe f en fonction de la concentration en aptamères afin de pouvoir ensuite déterminer la valeur de la constante de dissociation (**Figure 3-1**).

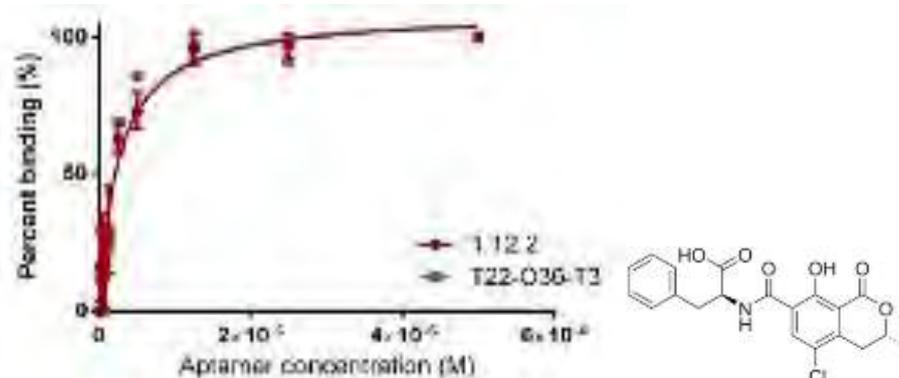


Figure 3-1 : Courbes de binding générées avec le logiciel graph prism pour deux aptamères anti ochratoxine obtenus par ultrafiltration. Chaque courbe est la moyenne de deux essais indépendants. La molécule d'ochratoxine à droite. Issu des supporting information de (117).

Penner et son équipe (119), ont également démontré que l'affinité entre l'aptamère anti ochratoxine et sa cible dépendait de la présence de cations divalents tels que Mg^{2+} , Ca^{2+} ou du pH. Lorsque des ions Mg^{2+} sont ajoutés au tampon de migration, la valeur de K_D est plus élevée que lorsque des ions Ca^{2+} sont utilisés. Ceci s'explique par le fait que l'ochratoxine forme un complexe de coordination avec les ions Mg^{2+} et Ca^{2+} et augmente l'affinité de la liaison du complexe avec l'ochratoxine en améliorant la formation de la liaison. En l'absence de ces cations, l'ochratoxine qui est chargée négativement ne se lie pas avec l'aptamère. D'autre part, ils ont aussi montré l'effet du pH et notamment l'effet d'un milieu acide qui réduit considérablement l'affinité de liaison. Ils ont observé qu'avec une concentration en calcium de 20 mM, lorsque le pH varie de 7.0 à 5.5, la valeur de K_D a augmenté de 50 à 700 nM. A pH acide, le groupement 8-hydroxyde de l'ochratoxine est protoné et donc la formation du complexe inhibé ce qui se traduit par une valeur de K_D meilleure.

3.1.2 L'ultrafiltration ou filter binding

L'ultrafiltration est une technique similaire à la dialyse mais elle utilise le marquage au phosphore 32 (^{32}P), pour quantifier la fraction d'aptamères liés. Une solution contenant la cible avec les aptamères marqués au ^{32}P est filtrée à travers une membrane en nitrocellulose sous l'effet d'une pression externe. Les aptamères marqués, liés à la cible sont retenus sur la membrane tandis que ceux qui sont non liés traversent le filtre (**Figure 3-2**). La fraction de molécules d'ADN liées à la cible est ensuite quantifiée en mesurant l'intensité de la radioactivité du ^{32}P . Par rapport à la dialyse, cette technique est plus rapide : on élimine le temps d'équilibre de la dialyse de plusieurs jours à quelques minutes. De plus, grâce à la radioactivité, on peut détecter des aptamères jusqu'à une concentration de l'ordre du picomolaire. Cependant, cela implique des mesures de sécurité supplémentaires, bien que des méthodes alternatives avec des marqueurs non-radioactifs aient été développées. Cette méthode possède aussi des inconvénients : la valeur de K_D peut parfois être mal estimée en raison d'une rétention incomplète du complexe aptamère-protéine à la surface de la membrane de cellulose. Cela

pourrait être dû à un effet compétiteur de la membrane. Parfois, il se peut aussi qu'il y est un effet compétiteur de l'aptamère avec la membrane.

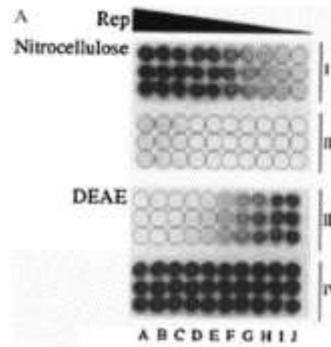


Figure 3-2 : Analyse d'une interaction entre l'ADN duplex HP (0,1 µM) et la protéine Rep. La nitrocellulose a conservé le complexe ADN-protéine alors que le DEAE contenait de l'ADN non lié. Les puits de gauche à droite (A-J) contiennent des concentrations de plus en plus faibles de protéine Rep. Chaque groupe de trois lignes, désigné par I-IV, contient des ensembles de données triplés. Les lignes II et IV sont les suites des lignes I et III respectivement. Issu de : (120)

3.1.3 L'électrophorèse sur gel

L'étude des interactions entre acides nucléiques et protéines en électrophorèse sur gel a été introduite par Fried et Crothers (121), Garner et Revzin (122). Les acides nucléiques sont des polyanions avec une mobilité électrophorétique négative tandis que les protéines ont généralement une charge et une mobilité plus faible. Ainsi, le complexe aptamère.cible aura une mobilité électrophorétique intermédiaire ce qui va faciliter la séparation sur gel. Le gel peut être soit un gel d'agarose ou un gel non-dénaturant de polyacrylamide. On garde fixe la concentration de la protéine ou de l'aptamère et dans chaque puits de gel, on va ajouter une quantité croissante de l'un des composés. Sous l'effet de l'application d'un champ électrique, les espèces vont migrer en fonction de leur rapport charge/taille. Les molécules d'ADN non liées vont migrer en premier suivi du complexe [aptamère. protéine] et de la protéine. Le gel est ensuite révélé par lecture aux rayons UV ou par marquage de l'ADN au ³²P ou par coloration au bleu de Coomassie R-250...

Comme la séparation peut durer de 30 min jusqu'à quelques heures, l'inconvénient est que le complexe peut se dissocier durant la séparation. Lorsque la constante de vitesse de dissociation est faible et que la demi-vie du complexe est significativement plus longue que l'échelle de temps de séparation, le ligand et le complexe d'ADN libres donneront naissance à des bandes distinctes sur le gel. Inversement, si la demi-vie du complexe est comparable ou plus petite que le temps de séparation, la séparation sera moins bonne et on pourra observer un smear entre les bandes. Cependant, la dissociation observée est généralement beaucoup plus lente que prévue. De plus, la structure réticulée du gel confine les ligands dissociés, donc donne la possibilité d'une réassociation du complexe, avant que l'aptamère et la protéine puissent diffuser sur une distance significative [35,43-46]. Cet effet stabilisant en combinaison avec la haute résolution des séparations par électrophorèse en gel rend l'étude des équilibres stœchiométriques multiples possibles [35]. Bien que cette méthode soit populaire en raison de sa haute sensibilité, de son faible coût et de ses différentes méthodes de détection, la réalisation du gel, la migration et la révélation de ce dernier est un procédé long.

3.1.4 La CLHP (chromatographie liquide haute performance)

La CLHP permet de séparer les composants en fonction de leur interaction avec la phase stationnaire. Ils sont plus ou moins retenus en fonction des interactions qu'ils développent avec chacune des phases. La chromatographie d'exclusion stérique, ou la chromatographie d'affinité sont des techniques utilisées pour l'analyse de petites molécules en présence d'ADN, qui permet de mesurer des constantes de dissociations. Mais c'est pour l'analyse comme phase stationnaire que les aptamères sont intéressants.

Les anticorps sont les ligands les plus utilisés dans la chromatographie d'affinité cependant du fait de leur taille, cela limite la capacité de liaison à la surface de la colonne. D'autres contraintes telles que l'immobilisation non-uniforme, et les conditions d'élutions peuvent réduire l'efficacité de cette méthode (123). Par analogie avec les anticorps, les aptamères peuvent être une alternative pour des applications de chromatographie par affinité analytique et préparative. Eric Peyrin (124), a décrit dans une revue, les caractéristiques de reconnaissance moléculaire des aptamères ainsi que les caractéristiques générales pour pouvoir les utiliser dans les techniques chromatographiques et électrophorétiques (**Figure 3-3**).

Features		Incidences in LC and CE
Affinity	K_d in the nM/mM range for small species K_d in the pM/nM range for macromolecules	High-affinity aptamers: type I assays ⁱⁱ Moderate-affinity aptamers: type II assays ⁱⁱ
Specificity	High to very high	Application to target purification, adaptable to real, biological conditions
Binding mechanism	Induced-fit mechanism, conformational changes, active conformers	Careful control of the denaturation-renaturation step of the aptamer, adequate binding buffer
Thermodynamics	Strong enthalpic and entropic contributions, significant heat capacity changes	Careful control of the temperature for modulating binding affinity and selectivity
Kinetics	Two steps with rate-limiting structural changes, relatively slow association-dissociation kinetic constants	Careful control of the medium composition (cation concentration notably) and temperature for improving binding affinity or efficiency performances and peak shape

Figure 3-3: Présentation des caractéristiques générales que doivent présenter les aptamères pour une application en chromatographie liquide et en CE. Type I assay correspond à l'utilisation d'aptamères caractérisés par une cinétique lente d'association-dissociation ($KD \sim nM/ \mu M$). Inversement, lorsque les aptamères ont une gamme d'affinité de l'ordre du $\mu M/mM$, ils sont caractérisés par une cinétique plus rapide. Tableau issu de : (124).

Pour cela ils doivent présenter une bonne affinité de liaison, une forte spécificité et ils doivent résister à de fortes conditions de lavages (par exemple 100 heures de trempage dans la soude à 1M). Les phases stationnaires utilisées sont des colonnes monolithiques ou des colonnes remplies ou à garnissage (packed), ou des colonnes capillaires (open tubular). Quel que soit le format, « un bras espaceur » est nécessaire pour maintenir les propriétés de liaison de l'aptamère lié à la phase stationnaire. Deux stratégies d'immobilisation à la surface chromatographique ont été décrites : par liaison covalente de l'oligonucléotide sur la surface chromatographique ou par liaison non-covalente. Ces deux modes de liaison, sont réalisés par l'utilisation de la biotine-streptavidine comme ligand d'immobilisation. Cette méthode implique le mélange du support chromatographique à la streptavidine avec la solution

d'aptamère pendant quelques heures à 48°C ou à température ambiante. Environ 15-20 nmol d'aptamère pour 100µL de milieu de séparation sont généralement liés par cette stratégie. Plus récemment, l'équipe de Perret et al, est parvenue en utilisant des aptamères à purifier trois protéines thérapeutiques provenant de diverses sources. En une seule étape, trois aptamères ont permis une purification efficace de leur protéines cibles avec des degrés de sélectivité allant de 0.5% à 98% de pureté en une seule étape (125).

3.1.5 Anisotropie de fluorescence

L'anisotropie de fluorescence est une technique puissante et rapide pour déterminer des mesures quantitatives d'interactions moléculaires diverses. Cependant cette méthode est difficilement utilisable contre les petites molécules dont la variation du signal de polarisation est peu quantifiable. L'aptamère est greffé à un marqueur fluorescent et lorsqu'une lumière polarisée illumine l'échantillon, seules les molécules dans la même orientation que les photons émis seront excitées. Si le marqueur tourne lentement par rapport à la cinétique d'émission, la polarité du photon entrant est retenue dans le photon de fluorescence. L'intensité du signal de fluorescence avec polarisation parallèle et perpendiculaire à celle de la source d'excitation peut être comparée pour évaluer la polarisation du signal. Lorsqu'un aptamère est lié à sa cible, le volume moléculaire augmente entraînant une moindre rotation et une augmentation de la polarisation. Peyrin et son équipe ont amélioré cette technique et ont proposé de nouvelles plates-formes de détection d'aptamères à anisotropie de fluorescence (FA) dédiées à la détection de petites molécules (126). Dans le cas de l'étude des aptamères pour le marqueur du cancer ovarien le CA125 (127), Whelan et son équipe ont déterminé la constante de dissociation de l'interaction aptamère.CA 125 par anisotropie de fluorescence. Pour cela, ils ont marqué les aptamères avec un marqueur : le Texas Red (5' TEX615). Puis 100 nM d'aptamères ont été chauffé à 97°C pendant 3min, laissé refroidir à température ambiante pendant 5 minutes. Ensuite les solutions d'aptamères ont été mélangées avec le CA125 à différentes concentrations mais en ajoutant de la BSA afin de prévenir les adsorptions à la surface du flacon. La solution est mise à incuber pendant 90 minutes dans le noir à température ambiante puis analysée. La longueur d'onde d'excitation était 585 nm, celle d'émission 635 nm et les valeurs d'anisotropies déterminées en présence (r) et en l'absence (r_0) du marqueur. La valeur de K_D est ensuite déterminée avec la formule suivante :

$$r - r_0 = \frac{(A * T)}{(K_D + T)}$$

T représente la valeur de la concentration de la cible CA125,

A représente la valeur de la concentration de l'aptamère qui est une constante.

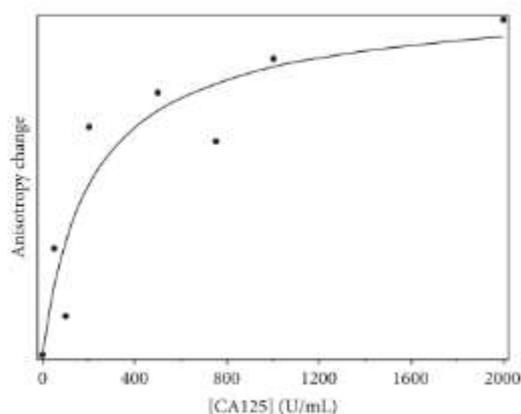


Figure 3-4 : Courbe de binding obtenue par anisotropie de fluorescence pour des échantillons contenant 100 nM de l'aptamère CA125_1 marqué au Texas Red avec des concentrations de cibles variant de 0 à 2000 U/mL. Excitation de la lumière polarisée à 585 nm, émission à 635 nm. Issu de (127).

La valeur obtenue par anisotropie de fluorescence était de 207 ± 109 U/mL (Figure 3-4). L'unité choisie est peu commune mais l'auteur justifie cette valeur car la protéine CA 125 présente une certaine hétérogénéité et donc une masse molaire mal définie. Utilisée pour les applications cliniques, cette unité de mesure a toujours été utilisée dans ce domaine-là. Les auteurs ont également, mesuré la valeur de K_D par électrophorèse capillaire appelée (APCE : Affinity Probe Capillary Electrophoresis). Ils ont obtenu une valeur de 80 ± 38 U/mL.

ID	Fluorescence anisotropy	Affinity probe CE
A1	2.2 μ M	390 nM
A3	9.1 μ M	500 nM
B10	280 nM	870 nM
D3	26 μ M	n/a
L1	>750 μ M	300 nM

n/a not available

Figure 3-5 : Comparaison de deux mesures de l'affinité de l'aptamère HE4 pour sa cible par anisotropie de fluorescence et par APCE (affinity probe capillary electrophoresis). Issu de : (101).

Bien qu'ils affirment que cette variation de K_D est dans la gamme observée par le laboratoire pour leurs autres études d'aptamères-protéines (Figure 3-5), (101), Lammertyn et son équipe ont également mesuré la liaison du lysozyme à un aptamère sélectionné par anisotropie de fluorescence mais également par électrophorèse capillaire (128). La valeur de K_D estimée par anisotropie de fluorescence était de 2.8 ± 0.3 nM tandis que la valeur trouvée par électrophorèse capillaire était 17 fois supérieure. L'explication donnée est qu'en électrophorèse capillaire, sous l'effet du champ électrique appliqué, l'effet joule entraîne une augmentation de température et donc durant la séparation l'équilibre du mélange est modifié. Donc au cours de la séparation, le complexe a été constamment dissocié (Figure 3-6).

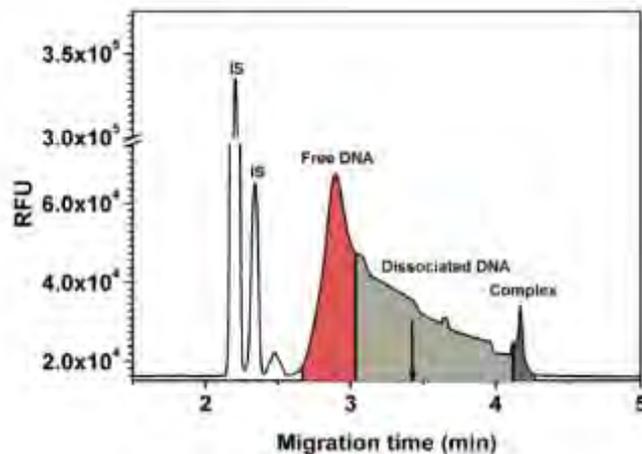


Figure 3-6 : Un exemple d'électrophérogramme de l'aptamère injecté avec le lysozyme en CE. Le mélange injecté contient 20nM d'aptamère marqué par fluorescence et 1600 nM de lysozyme. Conditions CE : Tampon TKG pH=8.3 Injection : 1 psi, 5s, Détection LIF. Capillaire : 50µm de diamètre, longueur : 40,2 cm. Séparation : 20kV. Le 2',7'-bis-(2-carboxyéthyl) -5-(et-6) - carboxyfluorescéine (BCECF) a été utilisé comme standard interne (IS). Les traits verticaux noirs correspondent aux points de collecte. D'après (128).

3.1.6 Absorption UV-visible

L'absorption UV-visible est une méthode largement utilisée du fait du faible coût et qu'il n'y a pas besoin de marqueurs pour détecter les composés. L'ADN a une absorption à 260 nm et les cibles telles que les protéines ont également une absorption à 210 et 280 nm. Ainsi, les constantes de liaison peuvent être déterminées par un changement de spectre de l'aptamère ou de la cible.

3.1.7 Dichroïsme circulaire (DC)

Le dichroïsme circulaire permet de détecter des changements de conformation et donc de caractériser des mesures d'affinité. A cause de la présence de carbones asymétriques dans les acides aminés, les sucres, les protéines et les acides nucléiques absorbent ainsi la lumière polarisée de façon différente. Les protéines avec des hélices α et des feuillets β anti-parallèles donnent des spectres distincts en dessous de 250 nm tandis que les acides nucléiques ont des spectres au-dessus de 250 nm. Ainsi il est possible de calculer la capacité de liaison en réalisant un dosage en ajoutant une concentration croissante de protéines et en évaluant le changement des signaux.

Dans l'étude de l'aptamère anti-thrombine, T15, l'équipe de Sica (129) a utilisé la technique du dichroïsme circulaire pour étudier la stabilité de l'aptamère T15 en interaction avec la thrombine en présence d'ions Na^+ et K^+ . Lorsque l'aptamère T15 est en interaction avec des ions Na^+ ou K^+ , le T15 se replie sous forme de quadruplex anti-parallèle. Le spectre présente un pic maximal à 295 nm, deux pics positifs à 210 et 250 nm et un pic minimum à 270 nm (Figure 3-7).

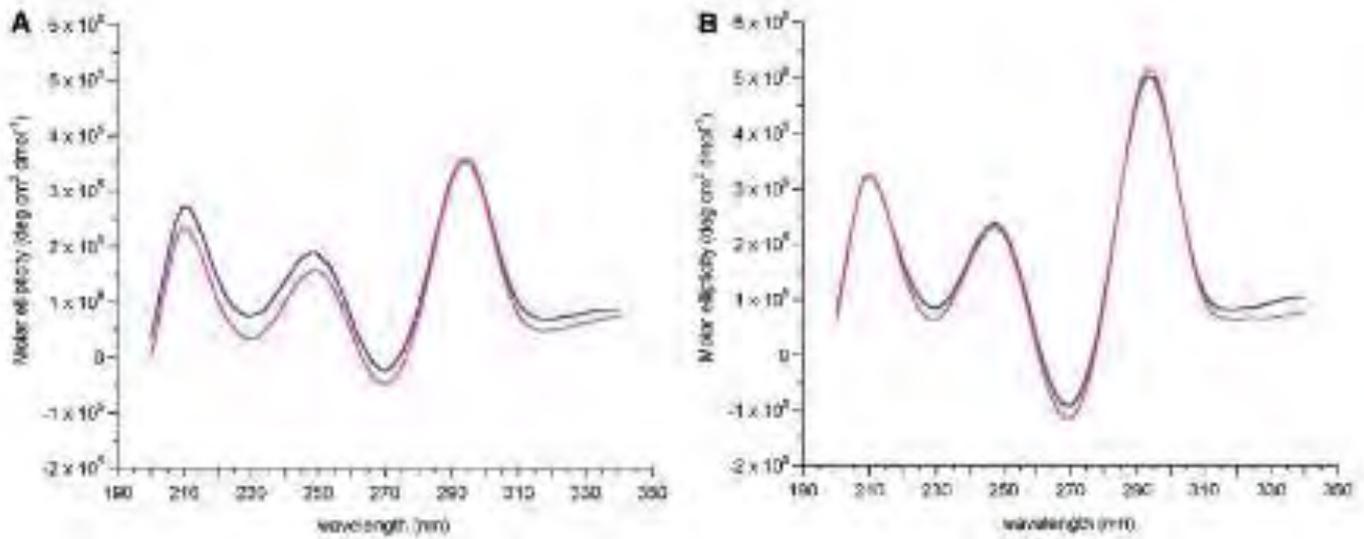


Figure 3-7 : Spectres dichroïques de l'aptamère TBA à $7\mu\text{M}$ dans du tampon Phosphate pH=7,1 et 100mM de sels, dans la figure A (100mM de Na^+) et dans la figure B (100 mM K^+). Les courbes bleues sont les résultats avant chauffage à 90°C , tandis que les courbes rouges représentent les spectres dichroïques après chauffage à 90°C . D'après (129).

Par ailleurs, la température de fusion de l'aptamère T15 en présence de Na^+ et K^+ est de respectivement 24°C et 53°C . Lorsqu'il sont mis en interaction l'aptamère et la thrombine, les courbes de fusion indiquent une augmentation considérable de la stabilité thermique de la structure en G-quadruplex. Les températures de fusion du T15 augmentent jusqu'à 58°C et jusqu'à 70°C en présence de Na^+ et K^+ (Figure 3-8).

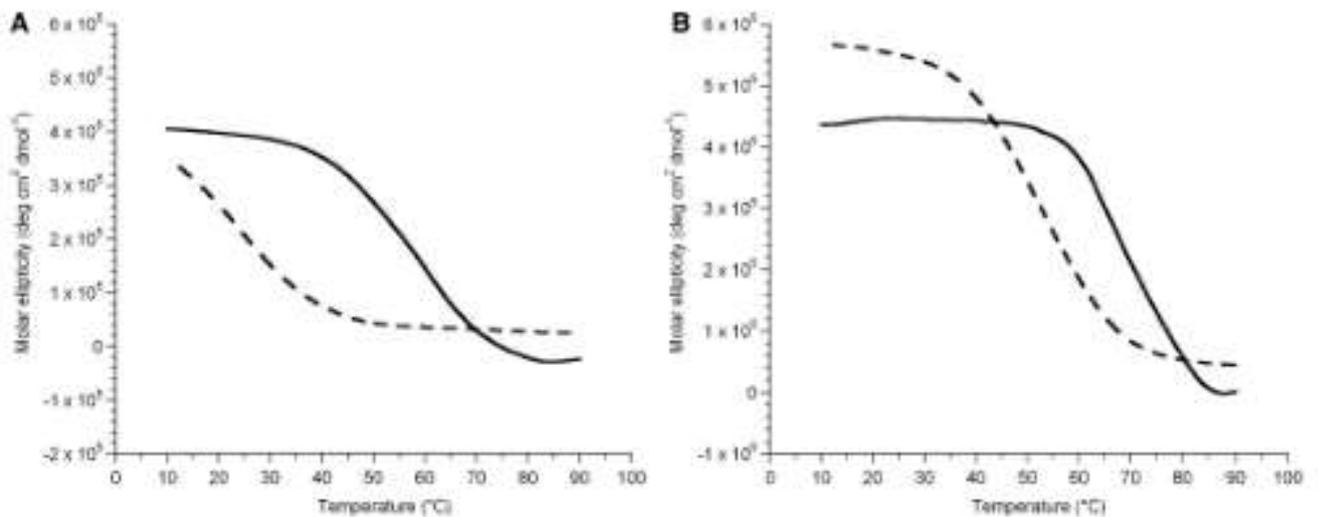


Figure 3-8 : Dénaturation thermique de l'aptamère TBA suivi d'une spectroscopie CD à 295 nm , en l'absence (ligne pointillée) et en présence (ligne en gras) d'une quantité équimolaire de thrombine à une vitesse de chauffage de $1^\circ\text{C} / \text{min}$. Les mesures ont été effectuées dans un tampon phosphate 25 mM à pH 7,1 avec $\text{MCl } 0,1\text{ M}$, où M représente soit Na^+ (A) et K^+ (B) en utilisant une concentration d'aptamère à $7\mu\text{M}$. D'après (129).

L'équipe de Chen (130), a combiné les données obtenues par dichroïsme circulaire, SPR et ITC afin de comparer les mécanismes de liaisons entre les aptamères T15, T29 et la thrombine. Le

dichroïsme circulaire a permis d'étudier les changements structuraux secondaires de l'aptamère T15 et T29 avec diverses concentrations de thrombine. Le spectre de l'aptamère T15 montre une bande négative à 250 nm et une bande positive à 280 nm (**Figure 3-9**). Ceci est caractéristique de la forme B de l'ADN. Lorsqu'il est mis en interaction avec la thrombine, une nouvelle bande positive apparaît à 254 nm, et les bandes précédentes sont décalées à 295 nm et 260 nm. Ceci indique que le complexe se structure sous forme de G-quadruplex. La thrombine induit le changement de conformation et l'intensité des deux bandes caractéristiques augmente avec la concentration croissante de thrombine.

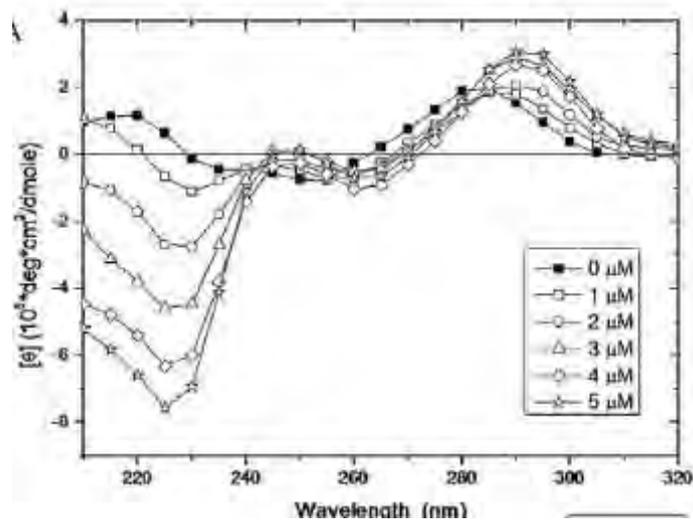


Figure 3-9: Spectre dichroïque de 5 μM de l'aptamère T15 dans 10 mM de TrisHCl à pH=7,4 et différentes concentrations de thrombine. D'après (130).

Le spectre de l'aptamère T29 est différent du T15 et présente une bande négative à 230 nm et une large bande vers 260 nm (**Figure 3-10**). Ce signal a une faible intensité lorsque la concentration de la thrombine est faible, ce qui indique que l'aptamère T29 forme une conformation en tige boucle. L'intensité de la bande à 265 nm continue d'augmenter avec la concentration de thrombine, ce qui indique que l'aptamère passe d'une forme B à une forme A de l'ADN.

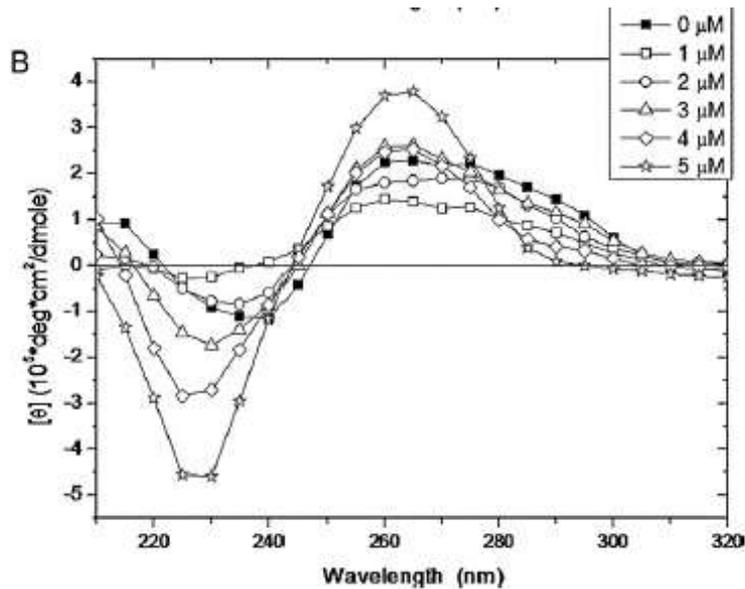


Figure 3-10 : Spectre dichroïque de 5 μM de l'aptamère T29 dans 10 mM de TrisHCl à pH=7,4 et différentes concentrations de thrombine. D'après (130).

3.1.8 Résonance plasmonique de surface (SPR)

Le phénomène de résonance plasmonique de surface correspond à l'oscillation collective d'électrons libres à l'interface entre deux milieux : un métal (Al, Ag, Au, Cu, ...) et un diélectrique. La résultante de ces oscillations est une onde électromagnétique évanescente, appelée onde plasmon, se propageant à cette même interface sur une épaisseur spécifique. Les aptamères sont habituellement choisis pour être immobilisés sur la surface métallique car les cibles comme les protéines sont plus grandes et induisent des changements de l'indice de réfraction plus importants. Cette méthode est bien adaptée pour mesurer des constantes d'affinités car le signal de résonance est lié linéairement à l'analyte immobilisé dans une large plage de concentration. Cette méthode a été appliquée à la mesure de l'aptamère contre l'IgE (**Figure 3-11**). Ainsi, lorsque cette dernière est injectée et s'écoule sur l'aptamère immobilisé, la liaison induit un changement de signal de l'indice de réfraction et le signal de résonance augmente. Ensuite, un tampon dénaturant est introduit à 900s et la dissociation se produit.

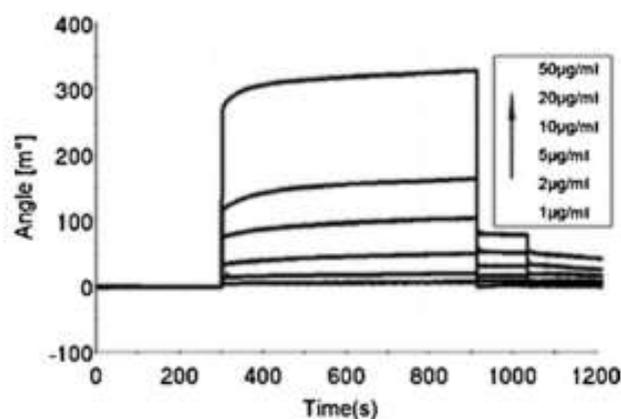


Figure 3-11 : Données SPR de l'interaction de l'aptamère avec différentes concentrations d'IgE. Issu de (116)

Lors de l'analyse des aptamères du lysozyme de l'œuf, les auteurs ont utilisé la mesure par SPR pour caractériser la liaison aptamère-lysozyme. Pour cela, l'aptamère analysé a été biotinylé et immobilisé sur une puce qui avait un revêtement de streptavidine. Ils ont ainsi pu caractériser la constante d'association $k_{on} = 6.99 \cdot 10^3 \pm 0.04 \cdot 10^3 \text{ M}^{-1} \text{ s}^{-1}$ et de dissociation $k_{off} = 4.3 \cdot 10^{-5} \pm 0.3 \cdot 10^{-1} \text{ M}^{-1} \text{ s}^{-1}$ et une valeur de K_D de $6.1 \pm 0.5 \text{ nM}$. Bien que cette valeur soit similaire à celle obtenue par anisotropie de fluorescence, les auteurs ont observé des interactions non-spécifiques entre le lysozyme et la puce de streptavidine. Ces interactions peuvent être dues à des liaisons électrostatiques entre la protéine et la surface de la puce, ou des liaisons de types hydrophobes entre la streptavidine et le lysozyme. Les auteurs ont ainsi conclu que les valeurs données par SPR doivent être prises avec précaution car le comportement cinétique d'un aptamère en solution est différent par rapport à un aptamère lié à une surface. Chen et son équipe (130), ont utilisé la SPR pour étudier les paramètres cinétiques de liaison de la thrombine avec les aptamères T15 et T29. Ils ont démontré notamment que la quantité de thrombine adsorbée diminuait lorsque la concentration en ions Na^+ augmentait (Figure 3-12). Ceci prouve que la liaison entre le T15 et la thrombine implique des interactions électrostatiques et donc la valeur de K_D augmente avec une concentration en sels NaCl plus élevée.

Table 2
The kinetic parameters of 15-mer aptamer binding with thrombin in different Na^+ concentrations.

NaCl conc.	k_{on}	k_{off}	K_D
50 mM NaCl	$6,627 \times 10^5$	0,036	$18,211 \times 10^6$
140 mM NaCl	$5,190 \times 10^5$	0,089	$5,845 \times 10^6$
200 mM NaCl	$3,538 \times 10^5$	0,173	$2,050 \times 10^6$

Table 3
The kinetic parameters of 29-mer aptamer binding with thrombin in different Na^+ concentrations.

NaCl conc.	k_{on}	k_{off}	K_D
50 mM NaCl	$8,704 \times 10^5$	0,104	$8,396 \times 10^6$
100 mM NaCl	$8,357 \times 10^5$	0,092	$9,103 \times 10^6$
200 mM NaCl	$8,222 \times 10^5$	0,075	$10,948 \times 10^6$

Figure 3-12 : Mesure des paramètres cinétiques de l'aptamère T15 (table 2) et de l'aptamère T29 (table 3) avec la thrombine avec différentes concentrations de NaCl. Issu de (130).

Cependant, la valeur de K_a (constante d'association) augmente lorsque la concentration en ions Na^+ augmente ce qui montre que ce n'est pas le même type d'interaction impliqué dans la liaison avec la thrombine. De plus, les valeurs des constantes k_{on} et k_{off} ne changent pas quel que soit la concentration en sels donc les auteurs ont conclu que des interactions hydrophobes étaient à l'origine de la liaison entre l'aptamère T29 et la thrombine.

3.1.9 Titration calorimétrique isotherme (ITC)

Cette technique repose sur la mesure précise des changements de température qui accompagnent en particulier les interactions inter-moléculaires, comme celles entre aptamère protéine en solution, sans devoir faire appel à un marqueur, puisque la quantité de chaleur produite ou absorbée est une propriété intrinsèque de pratiquement toutes les réactions biochimiques. L'ITC fournit des informations sur la stœchiométrie, l'enthalpie, l'entropie et la dynamique des liaisons entre un aptamère et une cible. On titre un échantillon contenant un volume constant d'aptamères ou de protéine avec une quantité croissante de l'un des deux composés. Lorsque le complexe se forme, cela se traduit par la libération (ou absorption) de chaleur soit une variation de température qui est enregistré. Le signal est mesuré comme la quantité d'énergie électrique nécessaire pour maintenir la température constante dans l'échantillon. Lorsque l'on atteint la saturation de la formation du complexe, on observe alors un palier (**Figure 3-13**).

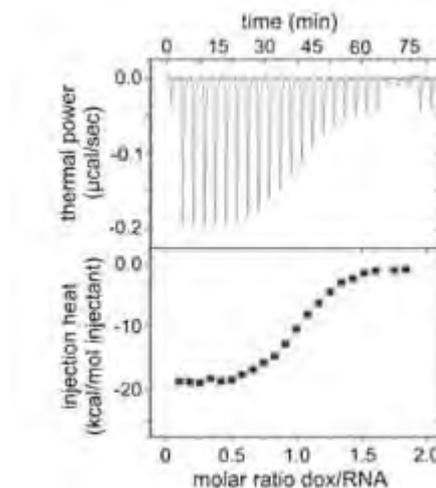


Figure 3-13 : Un exemple de thermogramme issu du dosage d'un aptamère ARN avec la doxycycline. Issu de (116).

Les travaux effectués par l'équipe de Chen (130), ont permis de comprendre quel était le type d'interactions impliqué dans la liaison entre l'aptamère T29 et la thrombine. Les deux aptamères T15 et T29 ont tous un comportement exothermique, cependant, lorsque l'aptamère T29 est lié à la thrombine ce dernier à un comportement moins exothermique que le T15 lié à la thrombine. L'entropie de liaison du T29 est meilleure que celle du T15. Par conséquent, c'est la raison pour laquelle le T29 se lierait avec la thrombine grâce à des interactions de types hydrophobe, en effet la variation d'entropie entre deux macromolécules est liée à la libération de molécules d'eau des surfaces hydrophobes.

3.1.10 Thermophorèse à micro-échelle (MST)

C'est une technique récente permettant l'analyse quantitative en solution à l'échelle du microlitre. Utilisée par Gao et son équipe (131), les auteurs ont analysé l'interaction aptamère thrombine en réalisant une série de courbes de fluorescences en analysant des gradients de concentrations de la thrombine dans les capillaires. Cette technologie est basée sur le mouvement des molécules marquées par fluorescence dans un gradient de température (2°C à 6°C) avec des fluides appelés thermophorèses dans un faible volume. La MST (microscale thermophoresis), n'a pas besoin d'immobilisation en surface ou de forte consommation d'échantillon comme la SPR ou l'ITC. C'est une méthode rapide qui utilise dans le cas de l'étude de l'interaction de la thrombine une douzaine de capillaires qui contiennent une série de gradients de concentration de la thrombine avec une concentration fixe d'aptamères. Le résultat obtenu montre que lorsque le laser infrarouge irradie le capillaire contenant de façon homogène la solution aptamère T15 et thrombine, un saut de température est tout d'abord observé (T jump) dès les 50ms après l'irradiation, suivi du phénomène de thermophorèse de l'aptamère hors de la zone chauffée, avec une diminution de la fluorescence qui atteint un équilibre. En effet, lorsque le laser est allumé, l'aptamère T15 et la thrombine diffusent dans la solution jusqu'à atteindre un état stable. La fluorescence dans le point chauffé diminue avec le mouvement de l'aptamère T15 de la zone de chaleur chaude vers la zone froide. L'arrêt du laser infrarouge entraîne de nouveau une augmentation de la fluorescence (**Figure 3-14**). En réalisant une série de courbes à différentes concentrations de thrombine, les auteurs ont utilisé la valeur du signal T-Jump pour calculer la valeur de la constante de dissociation. La constante de dissociation obtenue pour la liaison T15-thrombine était de $44,2 \pm 1,36$ nM tandis que pour la liaison T29-thrombine la valeur de K_D obtenue était de $15,5 \pm 0,637$ nM (**Figure 3-15**).

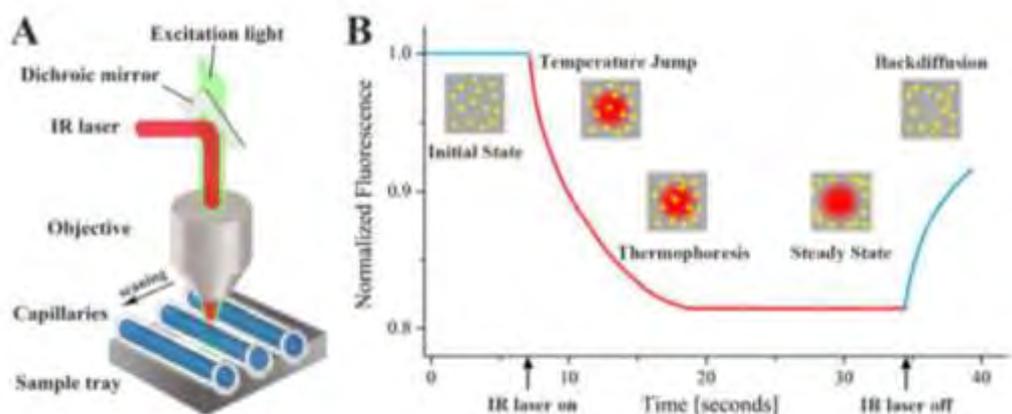


Figure 3-14 : Représentation schématique du système optique de la MST. (A) Une solution (4µL) dans un capillaire de diamètre 50 µm est localement chauffée. La fluorescence dans la solution est excitée et collectée à travers un objectif et un signal de thermophorèse est détecté dans un gradient de température induit par le rayon laser. En (B), est représentée une courbe typique de MST. Avant que le laser ne soit allumé, la thrombine, l'aptamère et le complexe thrombine.aptamère sont distribués de façon homogène. Dès que le laser est allumé, un saut rapide de température est observé (caractérisé par T-jump) en moins d'une seconde, suivi par la thermophorèse de l'aptamère marqué par fluorescence, hors de la zone chauffée. Quand le laser est éteint, l'aptamère revient dans la zone qui était éclairée par le laser et retrouve sa fluorescence. Issu de (131).

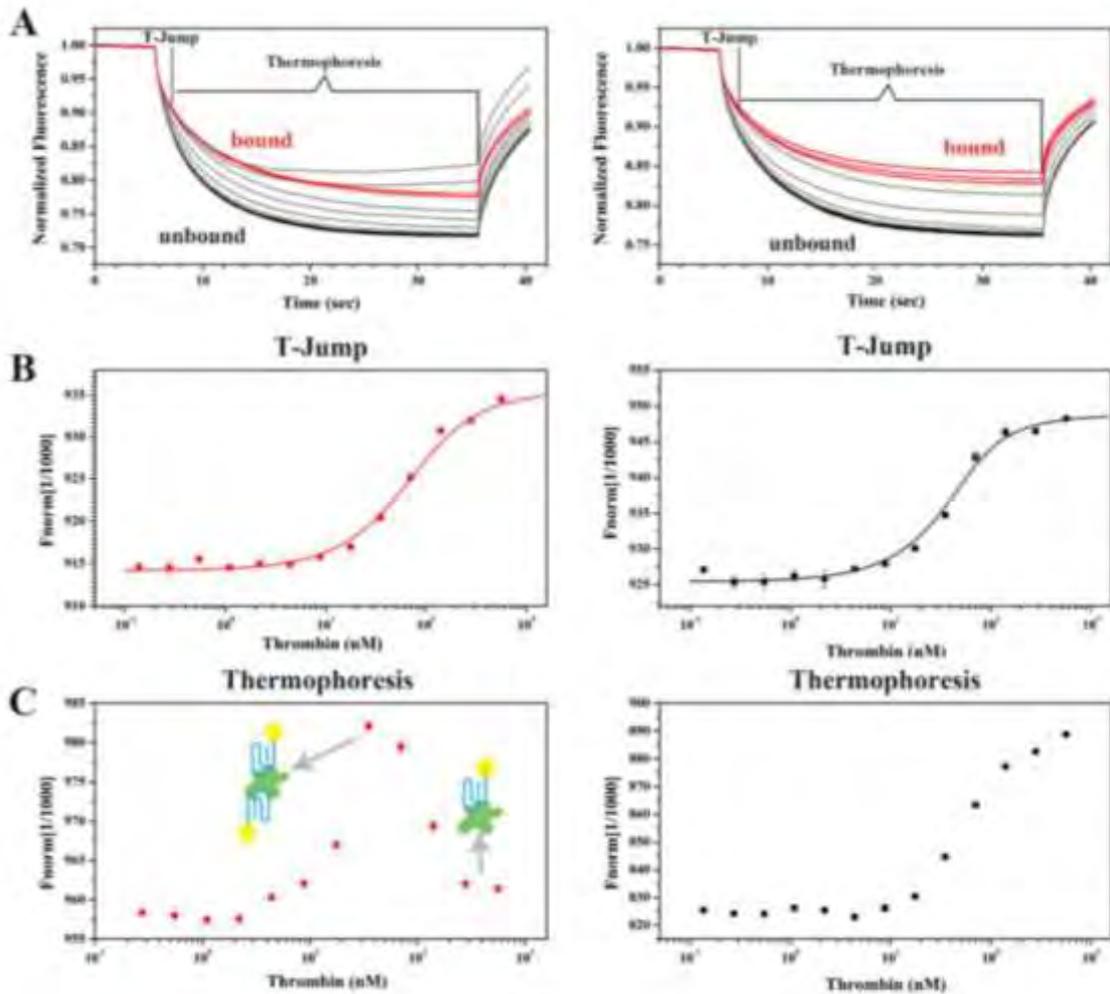


Figure 3-15 : Expériences de MST pour la détermination de la constante de dissociation pour l'aptamère T15 (à gauche) et l'aptamère T29 (à droite) avec la thrombine. Les courbes correspondants aux états liés et non liés sont colorées en noir et en rouge (A). Les signaux de ces courbes appelés « T-jump » sont extraits pour calculer les constantes de dissociation de la thrombine avec l'aptamère T15 (à gauche) et l'aptamère T29 (à droite) (B). Enfin, en (C) est représentée la ligne pointillée des signaux de thermophoresis de la thrombine pour l'aptamère T15 (à gauche) et l'aptamère T29 (à droite). Issu de (131).

3.1.11 L'électrophorèse capillaire :

Stuart et son équipe (132) ont utilisé ce mode de sélection pour améliorer l'affinité des séquences d'ADN se liant à la fibronectine. Cette protéine impliquée dans le phénomène de cicatrisation et de coagulation, présente à des taux élevés dans les tissus et sang de patients atteints d'un cancer du sang a une structure très similaire à la fibronectine. Pour être sûr de sélectionner les séquences ne se liant qu'à la fibronectine, ils ont effectué un cycle de sélection négative contre la fibronectine et un cycle de sélection positif pour la fibronectine. Ils ont ainsi augmenté l'affinité d'un facteur 30 en atteignant une valeur de constante de dissociation de $K_D = 500 \text{ nmol/L}$ (Figure 3-16).

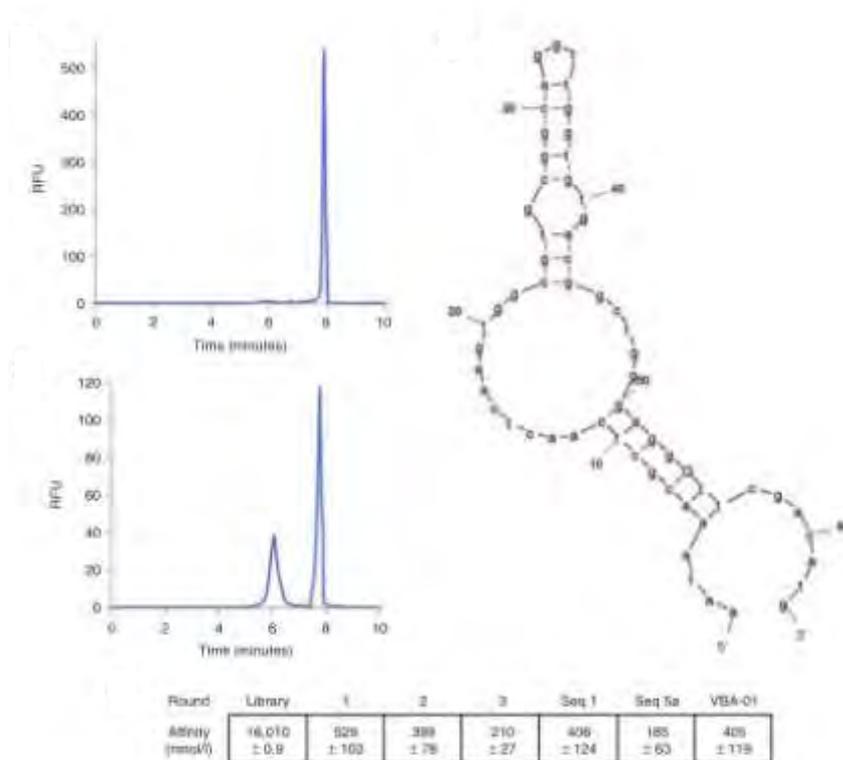


Figure 3-16 : Sélection d'aptamères par NECEEM. En (a), électrophérogramme de la banque d'ADN. (b) Electrophérogramme de la banque d'ADN après un cycle de SELEX. (c), KD obtenues à partir de plusieurs banques d'ADN. (d) Structure de l'aptamère VBA-01. Issu de (132).

3.1.11.1 Méthode de Krylov utilisant la NECEEM pour la détermination simultanée de la constante d'équilibre K_D et la concentration de la cible.

Le concept de NECEEM permet aussi de déterminer la fraction de molécules d'ADN non liées appelée R , à partir des aires des 3 pics et courbes (Figure 3-17) :

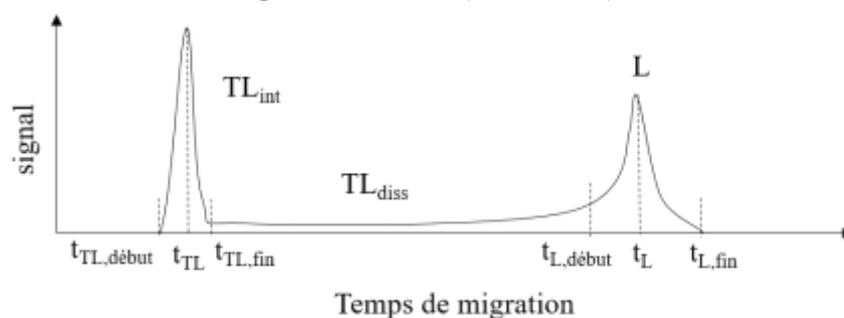


Figure 3-17 : Exemple d'un électrophérogramme NECEEM. Les intervalles de temps (TL) pour chacun des trois pics désignent respectivement le début ($t_{TL,début}$, $t_{L,début}$), le maximum (t_{TL} , t_L) et la fin ($t_{TL,fin}$, $t_{L,fin}$) des temps de migration de chacun des pics. TL_{int} correspond au complexe ADN.cible, TL_{diss} correspond au complexe qui a commencé à se dissocier avant d'atteindre le détecteur. Issu de (8).

$$R = \frac{A_L}{A_L + \varphi * (A_{TL_{int}} + A_{TL_{diss}})} \quad (5)$$

Où A_L , A_{TLint} , $A_{TL diss}$ correspondent respectivement aux aires du ligand, du ligand complexé avec la cible et du ligand non lié.

φ est le rapport du rendement quantique de fluorescence de la fraction L liée divisée par celui de la fraction L non liée. Par commodité, φ est égal à 1.

Ainsi, la valeur de K_D est donnée par l'équation suivante :

$$K_D = \frac{[T]_0 - [L]_0 * (1-R)}{\left(\frac{1}{R}\right) - 1} \quad (6)$$

Où $[T]_0$, $[L]_0$ représentent les concentrations totales de cible et de molécules d'ADN.

Ceci permet d'aboutir à la valeur de concentration de la cible si les valeurs de K_D et des molécules d'ADN sont connues :

$$[T]_0 = K_D * \left(\left(\frac{1}{R} \right) - 1 \right) + [L]_0 * (1 - R) \quad (7)$$

Cependant, dans certains cas, la détermination de K_D ne peut pas être déterminée avec l'équation ci-dessus car la concentration de la cible est mal connue. Krylov et son équipe ont proposé une méthode permettant la connaissance simultanée de K_D et de la concentration de la cible (8).

L'approche utilisée par l'équipe de Krylov se décompose en deux ensembles d'expériences :

- Un premier lot d'expériences pour lequel la concentration de molécules d'ADN $[L]_0$ est constante et la plus faible possible (atteinte du seuil de détection) et la concentration de la cible $[T]_0$ varie.

Le but de cette expérience étant de trouver la valeur du facteur de dilution α :

$$\alpha = \frac{[T]_{stock}}{[T]_0}$$

Où $[T]_{stock}$ correspond à la concentration stock (initiale) de la solution cible. La valeur de R (équation 5) est déterminée à partir des aires des multiples électrophérogrammes obtenus grâce au logiciel NAAP (NECEEM Area Analysis Program) développé par son équipe. Ce programme est dédié à l'analyse d'électrophérogrammes afin de calculer les aires des différents composants et d'en extraire les valeurs de K_D , k_{on} et k_{off} . Cette méthode a été développée à partir des travaux de Cherney et al (133). La courbe R est tracée en fonction du coefficient de dilution et la valeur du facteur de dilution donc par conséquent la valeur de $[T]_0$ correspondant à $R=0.5$ est ainsi déterminée (Figure 3-18).

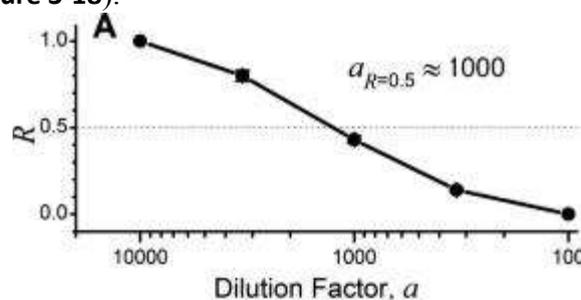


Figure 3-18 : Courbe $R = f(\alpha)$. A partir de cette courbe, détermination de la valeur alpha pour laquelle $R=0.5$. Grâce à cette valeur, détermination de la valeur de $[T]_0$ et comparaison avec la valeur de K_D . Issu de (8).

- Un deuxième lot d'expériences est ensuite réalisé à partir de la concentration de $[T]_0$ déterminée dans l'expérience 1 mais la concentration de $[L]_0$ varie. Le but est de déterminer la valeur de K_D et $[T]_0$. La valeur de R est également déterminée à partir des électrophérogrammes et le graphe R en fonction de $[L]_0$ est tracé (**Figure 3-19**).

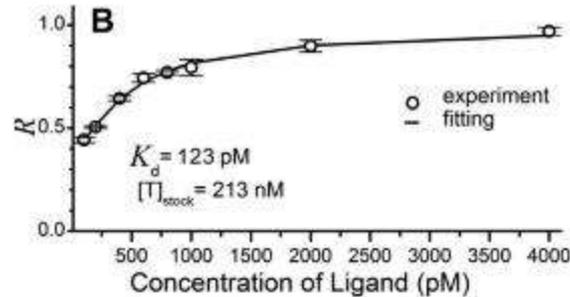


Figure 3-19 : Courbe $R = f(L_0)$ avec une concentration constante de T_0 déterminée dans la première expérience. Valeurs expérimentales de K_D et de T_0 obtenues pour la protéine MutS. Les points ronds représentent les valeurs obtenues à partir de R calculé. La courbe linéaire est celle ajustée à partir de l'équation 11. Issu de (8).

A partir de l'équation 6, on obtient une équation du second degré :

$$R^2 + \frac{K_D + [T]_0 - [L]_0}{[L]_0} * R - \frac{K_D}{[L]_0} = 0 \quad (10)$$

L'équation reliant R avec K_D , $[T]_0$, $[L]_0$ qui en découle est :

$$R = -\frac{K_D + [T]_0 - [L]_0}{2*[L]_0} + \sqrt{\left(\frac{K_D + [T]_0 - [L]_0}{2*[L]_0}\right)^2 + \frac{K_D}{[L]_0}} \quad (11)$$

Les valeurs expérimentales de R sont ensuite ajustées à partir de cette équation. Ils ont ainsi appliqué cette méthode à l'analyse de la protéine MutS et à son aptamère précédemment sélectionné (97). La procédure est relativement simple mais elle implique plusieurs considérations notamment le choix de la gamme de concentration de $[L]_0$, le nombre de points de $[L]_0$ à mesurer mais aussi la détermination de la concentration de la cible $[T]_0$. Ces paramètres sont importants car une mauvaise appréciation des valeurs peut diminuer la précision des mesures. En effet, la valeur de K_D obtenue précédemment était de 15nM, et la valeur obtenue avec cette méthode est de 123pM. Ils ont également pu comparer la valeur de la concentration de la cible de MutS mesurée au spectromètre ($[T]_0 = 226\text{nM}$) avec la valeur obtenue avec leur logiciel de simulation (COMSOL), ($[T]_0 = 213\text{nM}$) et avec la valeur obtenue expérimentalement ($[T]_0 = 202\text{nM}$).

Grâce à leur approche mathématique et l'utilisation de logiciels mise à la disposition de tous les utilisateurs, ils ont su développer avec précision la détermination de ces valeurs.

3.1.11.2 Modification de la technique NECEEM par Li et al.

L'équipe de Li (134) a développé une procédure permettant également de connaître les deux inconnues K_D et la concentration de la cible à partir de la concentration des molécules d'ADN. Pour cela, ils utilisent deux points de méthode pour réaliser un système de deux équations à deux inconnues :

$$K_D = \frac{[L]_{0,2} * (1 - R_2) - [L]_{0,1} * (1 - R_1)}{\frac{1}{R_1} - \frac{1}{R_2}} \quad (12)$$

$$[T]_0 = \frac{[L]_{0,2} * (1 - R_2) - [L]_{0,1} * (1 - R_1)}{\frac{1}{R_1} - \frac{1}{R_2}} * \left(\frac{1}{R_1} - 1 \right) + [L]_{0,1} * (1 - R_1) \quad (13)$$

Dans le cas de cibles difficilement quantifiables, ils ont su modifier la détermination de la constante de dissociation entre un aptamère et une concentration inconnue de cible. Dans un premier temps, ils se sont intéressés au modèle thrombine-T15 et une fois le modèle validé ont appliqué leur méthode à l'étude d'un virus : le H₉N₂. Pour la détermination de la valeur de la constante de dissociation, ils ne s'appuient que sur la concentration de l'aptamère et l'utilisation de deux équations à deux inconnues. Dans leur méthode, ils incubent deux mêmes volumes d'aptamères de différentes concentrations avec deux mêmes volumes de cible et en se basant sur l'équation de K_D suivante :

$$K_D = \frac{[T] * \left(1 + \frac{1}{R}\right) - [Apt]}{1 + R} \quad (14)$$

Où R = (A₁ + A₂) / A₃, [T] la concentration de la cible et [Apt] la concentration de l'aptamère. Ils obtiennent deux équations :

$$K_D * (1 + R_1) + [Apt]_1 = [T] * \left(1 + \frac{1}{R_1}\right) \text{ et } K_D * (1 + R_2) + [Apt]_2 = [T] * \left(1 + \frac{1}{R_2}\right)$$

Ce qui leur permet d'obtenir :

$$K_D = \frac{1}{R_1 - R_2} * \left(\frac{[Apt]_2}{1 + 1/R_2} - \frac{[Apt]_1}{1 + 1/R_1} \right) \quad (15)$$

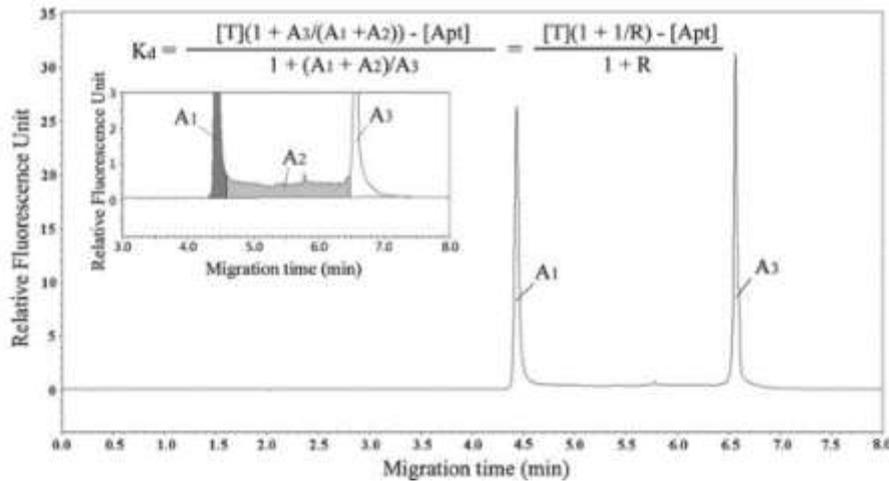


Figure 3-20 : Détermination de K_D par la méthode « classique ». Issu de (134). Injection par électrophorèse capillaire d'un mélange de 80nM d'aptamère 15 avec 100nM de thrombine. Les conditions en CE étaient les suivantes : injection : 0.5psi pendant 5s, séparation : +30kV, tampon de séparation TGK (Trisglycine potassium) à pH=8.5. L'électrophérogramme montre sur la vue détaillée 3 aires A₁, A₂ et A₃ correspondant respectivement aux aires du complexe (A₁), de l'aptamère dissocié du complexe pendant la séparation (A₂), et de l'aptamère libre (A₃).

En comparant les valeurs de K_D obtenues avec les deux méthodes (**Figure 3-20**), ils obtiennent des valeurs de K_D similaires ce qui leur a permis de valider leur modèle pour la thrombine et l'aptamère TBA mais également pour le virus H9N2 (**Figure 3-21**) (**Figure 3-22**).

Method	K_D (nM) determined in different thrombin concentrations		
	50 nM thrombin	200 nM thrombin	500 nM thrombin
Classical method ^a	33.65±0.96	36.73±1.77	56.98±2.85
Modified method ^b	42.75±4.58 ^c	28.98±0.94 ^d	64.89±5.10 ^e

Figure 3-21 : Comparaison des valeurs de K_D obtenues par la méthode "classique" et "modifiée". Issu de (125).

Group of different targets		Results for A9		Results for B4	
		20 nM aptamer	40 nM aptamer	20 nM aptamer	40 nM aptamer
Purified protein, 25 nM HA	<i>R</i>	0.55±0.02	0.43±0.01	1.20±0.08	0.60±0.03
	K_D in classical method (nM)	31.48±0.96 ^a		14.40±1.46 ^a	
	K_D in modified method (nM)	46.23±5.46		7.38±1.09	
Whole virus	<i>R</i>	0.49±0.03	0.36±0.02	0.58±0.02	0.30±0.01
	K_D in modified method (nM)	42.54±9.75 ^b		6.37±0.70 ^b	

A9 5'-GCT GCA ATA CTC ATG GAC AGC CTC CTG GGG TCA GGC TCA GAC ATT GAT AAA GCG ACA TCG GTC TGG AGT ACG ACC CTG AA-3', B4 5'-GCT GCA ATA CTC ATG GAC AGG GGC CGC GCC TGG TCG GTT GGG TGG GTG GCG CCC GGG ACG GTC TGG AGT ACG ACC CTG AA-3', HA haemagglutinin

^a Significant difference compared with the modified method using the purified protein as the target, $P < 0.01$

^b No difference compared with the modified method using the purified protein as the target, $P > 0.05$

Figure 3-22 : Capacité de liaison des aptamères A9 et B4 contre le virus H9N2. Issu de (125).

Cependant la détection expérimentale de la constante de dissociation K_D varie selon les données reportées dans la littérature. Par exemple, pour l'étude du modèle de la thrombine et de son aptamère, le T29, les valeurs de K_D en fonction de la technique utilisée varient de quelques nanomolaires à plusieurs centaines de nanomolaires (**figure 3-23**). L'explication la plus probable est le protocole expérimental utilisé. L'utilisation de cation comme Mg^{2+} favorise le folding du T29 à la thrombine. Dans les méthodes séparatives l'équilibre Mg^{2+} ADN est aussi perturbé voire remplacé par la contribution d'autres ions présents dans le tampon.

Méthode	K_D
Filter binding sur nitrocellulose	0.5nM
SPR	119.5nM
APCE	255nM
ctITP	128 nM

Figure 3-23 : Tableau représentant les différentes valeurs de K_D obtenues en fonction de la méthode utilisée.

Lammertyn et son équipe (128), ont également comparé plusieurs techniques de caractérisation de la constante de dissociation K_D sur des aptamères se liant au lysozyme. Durant le procédé CE-SELEX, ils ont amélioré l'affinité de leur aptamère en réalisant 5 cycles de SELEX. Cependant, ils ont trouvé des valeurs de constantes de dissociation variables :

Méthode	K_D
Anisotropie de fluorescence	2.8 +/- 0.3nM
SPR	6.1 +/- 0.5nM
ACE	52.9 +/- 9.1nM

Figure 3-24 : Tableau représentant les différentes valeurs de K_D obtenues en fonction de la méthode utilisée. Issu de (128)

La valeur de K_D obtenue par électrophorèse capillaire est supérieure aux valeurs obtenues par les autres méthodes. Krylov et son équipe (135) ont démontré que le vial contenant l'échantillon à l'inlet n'est pas thermostaté, donc il peut être soumis à la chaleur générée par le voltage, ce qui peut également être une source d'erreurs sur la détermination du K_D . Concernant, la mesure par SPR, l'aptamère est greffé à un groupement biotyl et immobilisé sur une bille de streptavidine, cependant, il peut y avoir des interactions électrostatiques entre la protéine et la surface des billes et des interactions hydrophobes entre la streptavidine et le lysozyme. Donc bien que la valeur de K_D obtenue par SPR soit proche de celle obtenue par anisotropie de fluorescence, le comportement de l'aptamère greffé est différent de celui en solution d'où des valeurs différentes.

3.1.11.3 Utilisation de la TDLFP : Transverse Diffusion of Laminar Flow Profile

L'électrophorèse capillaire requiert un faible volume d'échantillon dans le vial d'injection (une dizaine de microlitres). Cependant, en fonction du diamètre du capillaire, seuls quelques dizaines de nanolitres sont injectés. Aussi dans le but de réduire la consommation d'échantillon, Krylov (136) a proposé une autre approche en incubant directement la cible et l'aptamère dans le capillaire. Cette approche est appelée TDLFP (Transverse Diffusion of Laminar Flow Profile), elle permet d'injecter seulement quelques nanolitres de réactifs dans le capillaire par injection hydrodynamique et la réaction d'incubation se fait dans le capillaire. Les réactifs sont mélangés par diffusion laminaire (**Figure 3-22**). Les électrophérogrammes obtenus sont similaires à ceux trouvés en NECEEM, mais la valeur de K_D est différente : 18.0 ± 5.4 nM au lieu de 7.8 ± 0.7 nM, par NECEEM. Cette différence peut être attribuée à l'inexactitude du volume injecté par CE et/ou le temps d'incubation dans le capillaire qui n'est peut-être pas optimal pour avoir une bonne séparation des composés.

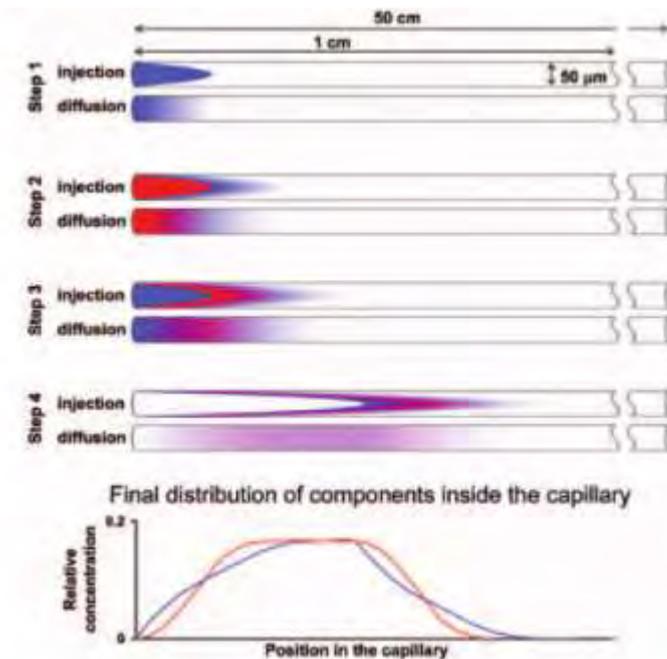


Figure 3-22 : Principe de la TDLFP. Issu de (136).

3.1.11.4 La MEKC: Micellar electrokinetic capillary chromatography

D'autres ont utilisé la chromatographie électrocinétique micellaire (MEKC : micellar electrokinetic capillary chromatography) (137) , pour mesurer les concentrations d'ADN et de 17β-estradiol pour obtenir la constante de dissociation d'un aptamère lié en position 5' à une biotine l'ensemble étant lié sur des billes magnétiques de streptavidine. Après ajout de la cible, le 17β-estradiol, le complexe est fixé via les interactions biotine-streptavidine, puis les billes sont rincées afin d'éliminer les interactions non-spécifiques c'est-à-dire, les molécules d'ADN non liées à la cible. Les billes contenant les molécules d'ADN liées à la cible sont lavées, chauffées à 70°C pour décrocher l'ADN et les complexe ADN.cible puis re-injectées en électrophorèse capillaire et la séparation entre l'estradiol et l'aptamère se fait après stacking en milieu micellaire en polarité inverse. L'ajout de SDS comme tensioactif dans le tampon de migration va permettre de séparer l'estradiol et l'ADN. Ainsi ils ont trouvé comme valeur de constante de dissociation : $K_D = 70 \pm 10 \mu\text{M}$.

3.1.11.5 La ctITP (capillary transient isotachopheresis) :

Colyer et son équipe (138), ont développé une méthode alternative de NECEEM, avec préconcentration en ctITP (capillary transient isotachopheresis) couplée avec le séquençage haut-débit pour sélectionner des aptamères anti-thrombine à partir d'une banque d'ADN. La ctITP, est une technique de concentration qui a permis d'améliorer les limites de détection de concentration et d'augmenter la quantité d'échantillon pouvant être injecté dans le capillaire. Les temps de migration des composées sont reproductibles ce qui leur a permis de mettre au point une méthode de collecte des fractions contenant le complexe ADN.cible dans le but d'être séquencé par le séquençage haut-débit NGS. Grâce à la méthode ctITP, et en réalisant le marquage de l'ADN à l'intérieur du capillaire en ajoutant un marqueur fluorescent le SYBR Gold dans le tampon de séparation, ils ont montré qu'ils obtenaient une meilleure résolution

des pics ce qui est nécessaire pour permettre la collecte de la fraction du complexe (**Figure 3-23**). Ils ont ainsi déterminé une valeur de K_D de 128 ± 9 nM pour l'aptamère T29 et 324 ± 6 nM pour l'aptamère T15, ces valeurs étant similaires à celles trouvées dans la littérature, ce qui a validé leur méthode de marquage de l'ADN.

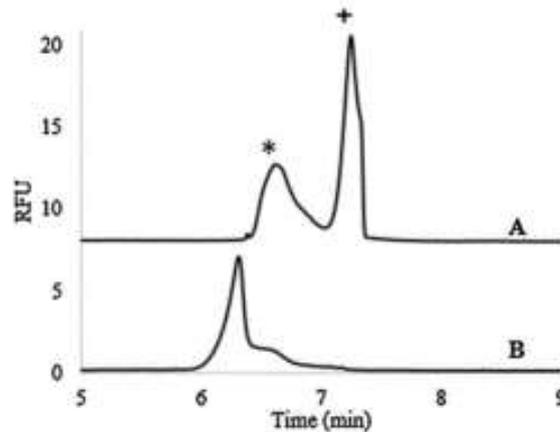


Figure 3-23 : Electrophérogrammes comparant la séparation du complexe [Thrombine.ADN] de l'aptamère libre (A) par ctITP (B) par CZE.

Un des inconvénients de la collecte est qu'il est nécessaire que le vial de collecte contienne un peu de tampon de séparation. Mais la fraction collectée ne constitue que quelques nanolitres, donc il ne faut pas que le facteur de dilution soit trop important car lors de la réinjection dans le capillaire de la fraction collectée, il ne serait plus possible de discerner les pics sur l'électrophérogramme. Colyer et son équipe (139), ont ainsi réalisé différentes collectes (1, 2, 8 et 16) dans différents vials contenant 5, 10 et 15 μL (**Figure 3-24**). Les échantillons collectés ont été vortexés et réinjectés. L'aire du pic de l'aptamère T29 a été mesurée et la courbe de calibration Aire du pic en fonction du nombre de réplicat a été tracée. Les courbes pour chacun des 3 volumes de vials ont donné une réponse linéaire mais c'est avec 10 μL de tampon de séparation qu'il y a la meilleure reproductibilité (RSD=9,4%).

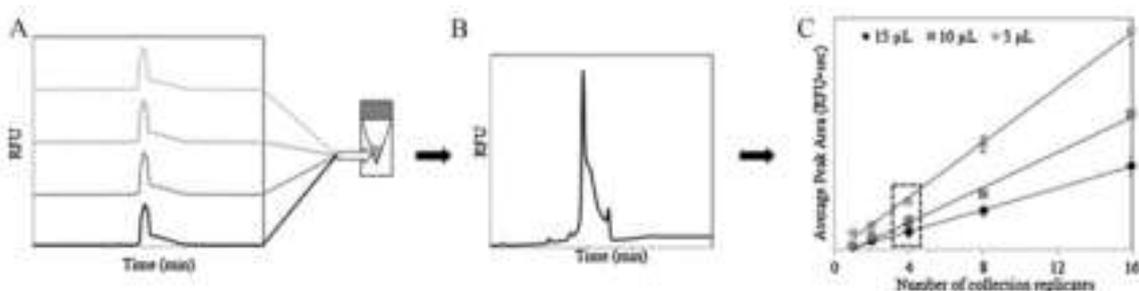


Figure 3-24 : Principe de la collecte en CE. (A) 4 collectes d'un échantillon d'ADN dans un vial contenant 5, 10 ou 15 μL de tampon. (B) Allure d'un électrophérogramme obtenu lors de la ré-injection d'une fraction. (C) Aire des pics du complexe représentée en fonction du nombre de collectes (1, 2, 8, et 16 collectes). Issu de (139).

Après avoir séquencé leurs échantillons par séquençage haut-débit (Ion Torrent Personal Genome Machine), ils ont ainsi démontré qu'un seul cycle d'électrophorèse capillaire où une banque d'ADN avec 10^{15} oligonucléotides mélangée avec 400nM de thrombine et contaminée avec 2nM de l'aptamère T29 (dans le but de séparer la fraction d'ADN libre de la fraction liée

à la thrombine) et collecté 16 fois, avait donné un enrichissement de l'aptamère T29 de 0,4% à 15%.

4 OPTIMISATION DE L'AMPLIFICATION PAR PCR :

4.1 Après la sélection par électrophorèse capillaire, la PCR :

Chaque fraction collectée par électrophorèse capillaire est ensuite amplifiée lors d'une étape de PCR (polymérase chain reaction). Un des challenges de la méthode NECEEM est que comme le nombre d'aptamères collectés est très faible, il faut réussir à les amplifier de façon spécifique et efficace. Par ailleurs, une réaction de PCR inefficace peut générer un important nombre de produits non-spécifiques et donc réduire l'efficacité de la méthode de sélection par électrophorèse capillaire. C'est ainsi que l'équipe de Sergei Krylov a suggéré une explication à l'échec du SELEX (140). Voici son raisonnement :

On considère qu'il y a deux catégories de molécules dans une banque d'ADN : celles se liant à la cible (Binders : B) et celles qui ne se lient pas (Non-binders : N).

Le nombre de molécules injectées par électrophorèse capillaire est $B_{in} + N_{in}$ et la proportion de molécules non-liées est plus importante que les molécules liées à la cible, donc $N_{in} \gg B_{in}$.

On attribue des coefficients de collection qui correspondent à l'efficacité de collection de chaque fraction : k_B et k_N et la fraction collectée contient normalement plus de molécules liées donc : $k_B \gg k_N$. Le cas idéal étant d'avoir $k_B=1$ et $k_N=0$.

Ainsi, le nombre de molécules collectées sera : $B_{in} * k_B + N_{in} * k_N$.

Le but de la réaction d'amplification par PCR est de réussir à amplifier préférentiellement la fraction $B_{in} * k_B$, plutôt que la fraction de molécules non liée $N_{in} * k_N$.

Par ailleurs, on comprend que l'efficacité d'amplification entre ces deux catégories de molécules sera différente car la fraction contenant le complexe [ADN.cible] contient des molécules d'ADN qui se sont structurées pour être liées à la cible. L'enzyme polymérase peut donc avoir plus de difficultés pour se fixer à ces molécules et ainsi réussir à les amplifier. Par exemple c'est le cas des G quadruplex.

Cette différence d'amplification peut s'exprimer par : Z_B et Z_N .

Une amplification idéale produit deux copies d'ADN donc $Z=2$ mais dans notre cas, il est plus difficile d'amplifier la fraction liée que non liée donc $Z_B < Z_N$. Au bout de n cycles de PCR, le nombre de molécules amplifié se traduit par :

$$B_{out} = B_{in} * k_B * Z_B^n \quad (1)$$

$$N_{out} = N_{in} * k_N * Z_N^n$$

Pour enrichir la fraction de molécules liées à la cible, B_{out} il faut que cette fraction soit plus importante que donc $\frac{B_{out}}{N_{out}} > \frac{B_{in}}{N_{in}}$, et en simplifiant l'équation (1), on obtient :

$$\frac{k_B}{k_N} > \left(\frac{Z_N}{Z_B}\right)^n$$

Si cette équation n'est pas vérifiée, cela signifie que la réaction de sélection et d'amplification ne peut pas fonctionner. Par exemple, si l'amplification Z_N des molécules non liées est de 1% supérieure à Z_B et si l'on considère $n=40$ cycles de PCR, alors l'efficacité de séparation $\frac{k_B}{k_N}$ devra être supérieure à 1,5. Si Z_N est 10% supérieur à Z_B alors il faudra que $\frac{k_B}{k_N}$ soit supérieur à 50, c'est-à-dire qu'il faudra collecter 50 fois plus de molécules d'ADN liées à la cible que les molécules non liées.

L'application mathématique peut se poursuivre en considérant que pour qu'une sélection par SELEX soit efficace il faut que : $\frac{B_{out}}{N_{out}} > 1$

Si l'on fait m cycles de SELEX on obtient alors :

$$B_{out} = B_{in} * (k_B * Z_B^n)^m \text{ et } N_{out} = N_{in} * (k_N * Z_N^n)^m$$

Ce qui permet d'obtenir la valeur de m :

$$m > \frac{\log(B_{in}/N_{in})}{\log\left(\frac{k_N}{k_B} * \left(\frac{Z_N}{Z_B}\right)^n\right)}$$

Grâce à la valeur de m , on peut ainsi déterminer le nombre de cycles de SELEX nécessaires. Ce nombre varie en fonction du nombre de biais PCR. Par exemple, pour un biais PCR

$$\frac{Z_N}{Z_B} = 1, \text{ le nombre minimal de cycles SELEX sera égal à 4, en considérant } \frac{B_{in}}{N_{in}} = 10^{-8},$$

$$\frac{k_N}{k_B} = 0.01 \text{ et } n=40.$$

Pour 10% de biais PCR, il faudrait 23 cycles de SELEX.

Ce calcul permet d'obtenir tous les enjeux de l'efficacité de la sélection et de la collecte, en effet la limitation des biais de PCR est souvent difficile. L'émulsion PCR où chaque réaction est réalisée dans une microgouttelette devrait diminuer certains de ces biais. Par ailleurs on voit tout l'intérêt de pouvoir choisir un nombre de cycles de PCR le plus important possible, mais ce nombre est intimement lié aux biais. Il est évident que dans la proposition mathématique de Krylov les facteurs Z ne sont pas des constantes et vont variés selon le nombre de cycle de PCR.

Il est donc important d'estimer le nombre de cycle de PCR optimal pour la sélection d'un aptamère.

Lors de la sélection de l'aptamère se liant au marqueur du cancer ovarien HE4, l'équipe de Whelan (101) a amplifié chaque pool issu de la CE-SELEX (pour laquelle ils avaient effectué, 3 cycles de sélection positive et 2 cycles de sélection négative). Lors de la préparation du mix nécessaire à la réaction d'amplification, ils ont mélangé les réactifs suivants : 484µL d'eau, 16µL de dNTPS, 20µL de primers sens et anti-sens, 96µL de MgCl₂ et 160µL de tampon. Après avoir été mélangés, l'enzyme Taq Polymérase a été incorporée puis 5µL de fractions collectées en CE ont été ajoutées au mélange final. Dans leur étude, afin d'atteindre une amplification optimale c'est-à-dire de trouver le bon nombre de cycles de PCR afin d'amplifier les fractions collectées, mais pour éviter de réaliser trop de cycles PCR qui aboutiraient à la production d'amplicons non spécifiques, ils ont réalisé des duplicats des fractions collectées en CE qui ont servi de témoin afin de prévoir le nombre de cycles PCR nécessaires. Chaque fraction amplifiée a ensuite été déposée sur un gel d'agarose et analysée en UV. Le nombre de cycle optimal a été choisi comme étant celui qui avait la bande la plus visible et le moins de sous-produits possibles sur le gel d'agarose (**Figure 4-1**).

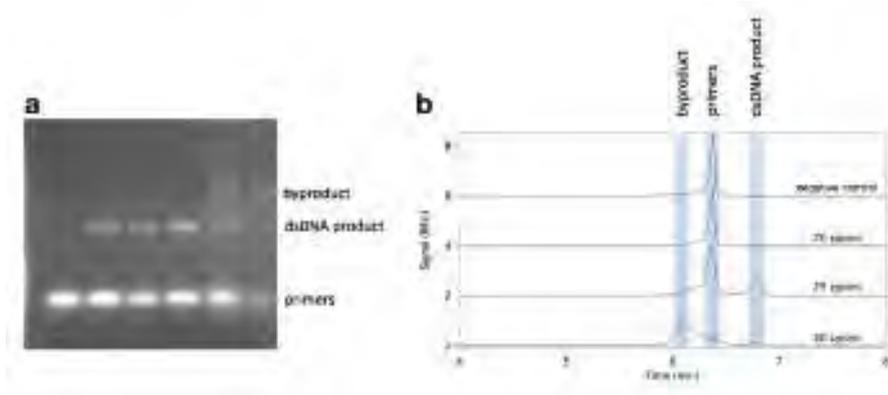


Figure 4-1 : (a) Gel montrant l'effet d'une augmentation du nombre de cycles PCR et l'apparition de sous-produits. (b) Superposition d'électrophérogrammes montrant l'effet d'une augmentation du nombre de cycles de PCR. Issu de (101).

Krylov et son équipe ont étudié la protéine MutS (136), pour laquelle, ils avaient déjà sélectionné des aptamères. Le but de leur étude était de prouver que l'on pouvait amplifier par PCR asymétrique des aptamères sélectionnés par CE, clonés et séquencés lors d'une étape de CE/ non-SELEX (**Figure 4-2**), et que ces mêmes aptamères pouvaient être directement injectés en CE pour être caractérisé sans que leur constante de dissociation en soit altérée. Ils ont ainsi montré que les meilleures conditions en PCR asymétrique étaient lorsque la quantité de primer sens était 10 fois supérieure à celle de l'anti-sens. De plus, lorsqu'ils ont incubé l'aptamère non purifié avec la cible afin de déterminer sa valeur de K_D , ils ont obtenu une valeur de $8,1 \pm 0,6$ nM, ce qui est très proche de la valeur obtenue avec un aptamère purifié ($7,8 \pm 0,7$ nM).

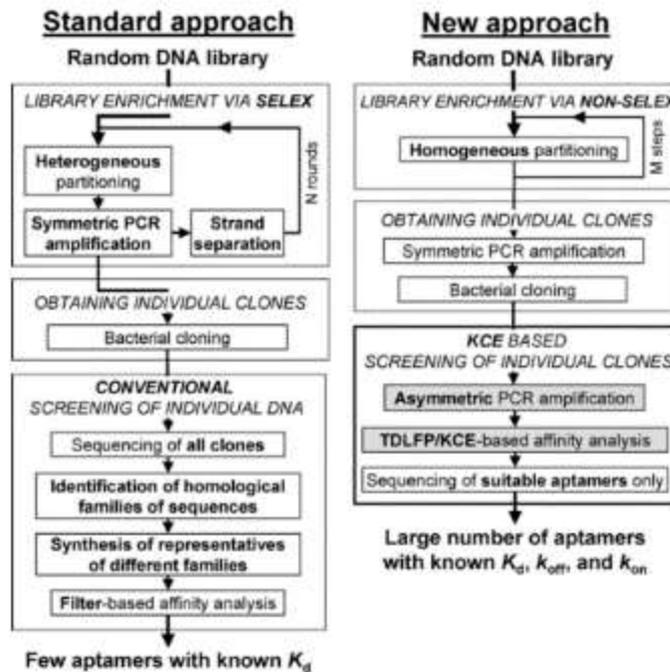


Figure 4-2 : Comparaison des deux méthodes SELEX et non-SELEX utilisée par Krylov pour sélectionner des aptamères pour la protéine MutS. Issu de (136)

4.2 La PCR en temps réel :

D'autres auteurs ont utilisé la PCR en temps réel (141) pour amplifier les échantillons collectés issus de la CE-SELEX. Pour cela, en électrophorèse capillaire, ils ont déterminé les temps de migrations de la cible (ici la BSA : bovin serum albumin) et de la banque d'ADN par absorbance UV, et lors de l'injection du mélange, ils ont réalisé 7 fenêtres de collection de deux minutes entre la fin du pic de la BSA et le début de la banque d'ADN. La particularité de leurs travaux réside dans le fait qu'ils n'utilisent pas de détecteur LIF couplée à la CE et que seule la PCR en temps réel (qPCR) leur fournit des informations sur la nature de la fraction collectée. Le principe de celle-ci est similaire à la PCR conventionnelle sauf que l'on rajoute une sonde fluorescente (le SYBR green) qui ne fluoresce qu'une fois liée à l'ADN. En comparant la quantité de cycles nécessaires pour amplifier chaque fraction, ils ont remarqué que la deuxième fraction était celle qui contenait le plus de matrice et que cette fraction avait été collectée bien après la migration de la BSA et avant la migration de l'ADN. Donc ils ont choisi cette fraction à séquencer.

Cependant, avec la PCR conventionnelle l'enrichissement d'une banque d'ADN peut se traduire par la conversion des aptamères en sous-produits d'amplifications non spécifiques (142). Il est intéressant de remarquer que l'amplification d'une banque d'ADN dégénérée sur 80 nucléotides, diffère d'un d'ADN classique. En effet, avec une banque d'ADN aléatoire (dégénérée sur 80 nucléotides), des sous-produits de réaction sont formés à partir du 15^{ème} cycle de PCR alors qu'il n'en est rien pour un fragment d'ADN de même taille que la banque. Le produit PCR est converti totalement en sous-produits dès le 20^{ème} cycle (Figure 4-3).

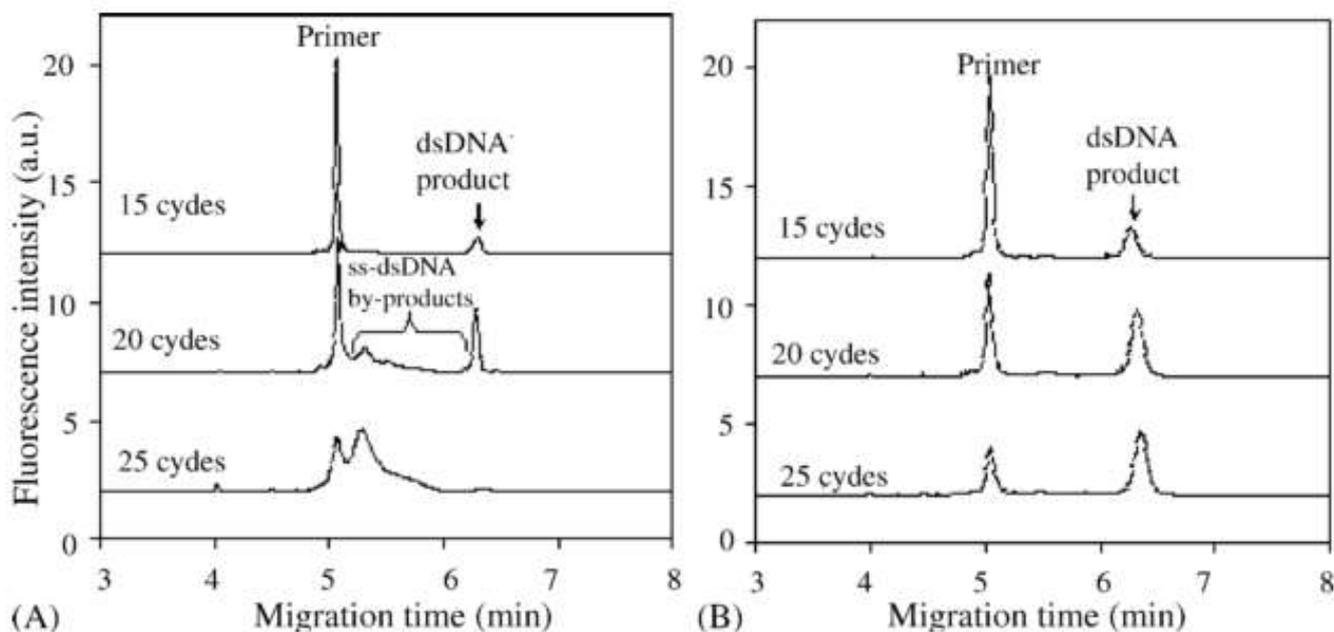


Figure 4- 3 : Electrophérogammes de mélange de PCR pour l'amplification d'une banque d'ADN dégénérée sur 80 nucléotides (à gauche), et d'une banque d'ADN homogène de 80 nucléotides (à droite). Le nombre initial de molécules était de 108, pour une concentration en Taq Polymérase de 0.05 unité d'activité par microlitre. Issu de (142).

Les auteurs ont aussi étudié quel était l'impact de la longueur de la banque d'ADN et de sa concentration sur l'amplification et ont noté que plus la longueur de la banque d'ADN est importante, moins bon est le rendement PCR et plus importante est la conversion en sous-produits. Ils ont également démontré qu'une concentration trop importante de l'enzyme Taq Polymérase pouvait engendrer la formation de sous-produits PCR. Les primers sens et anti-sens placés à chaque extrémité de la banque d'ADN peuvent aussi amplifier des sous-produits de par la nature de leur séquence, ou interférer dans la réaction de PCR.

4.3 L'émulsion PCR :

On l'a déjà mentionné, en vue de réduire la formation de sous-produits lors de la réaction de PCR, Krylov et son équipe ont proposé l'utilisation de l'émulsion PCR (ePCR) (140) couplée à l'électrophorèse capillaire qui permet de réduire la formation de sous-produits d'amplification. Cette méthode permet de miniaturiser l'amplification PCR en réalisant dans un mélange eau / huile, des microgouttelettes contenant les billes et l'ensemble des réactifs de la PCR. Dans chacune des microgouttelettes on imagine disposer d'un seul fragment d'ADN à amplifier. En vortexant le mélange, les billes et l'ensemble des réactifs de la PCR sont réunis, l'émulsion est ensuite cassée et les billes portant de multiples copies d'un fragment sont enrichies (**Figure 4-4**). Afin de réaliser une amplification, on travaille en conditions limitante où des sondes complémentaires d'un adaptateur particulier sont fixées sur les billes magnétiques. Seuls les fragments ayant cet adaptateur à leur extrémité seront fixés sur les billes et amplifiés.

Les autres fragments avec des adaptateurs différents seront donc éliminés. L'échantillon est dilué afin d'obtenir dans chaque bille une molécule d'ADN cible. La distribution statistique de l'ADN est alors donnée par la loi de Poisson, qui décrit les probabilités d'événements rares. Chaque échantillon est soumis à une amplification par PCR et analysé individuellement. En fin de PCR, on aura alors des réactions montrant une amplification (1) ou pas (0). Une sonde reconnaît les produits d'amplification (mutés et non mutés), et une deuxième sonde émettant une couleur différente s'hybride aux séquences non mutées. À la fin de l'analyse, seuls les puits de la plaque émettant un signal de fluorescence sont analysés. Chaque compartiment produit ainsi, en fonction de la présence ou non de la molécule d'intérêt, un signal positif ou négatif, d'où le terme de « PCR digitale ». En calculant le rapport du nombre de puits émettant chaque couleur, il devient possible de quantifier la fraction d'ADN mutée contenue dans l'échantillon établir une courbe de calibration. Grâce à cette discrimination binaire, on peut détecter des événements rares, à l'échelle de la molécule unique. La sensibilité de l'ePCR dépend principalement du volume d'échantillon et du nombre total de compartiments que l'on peut analyser.

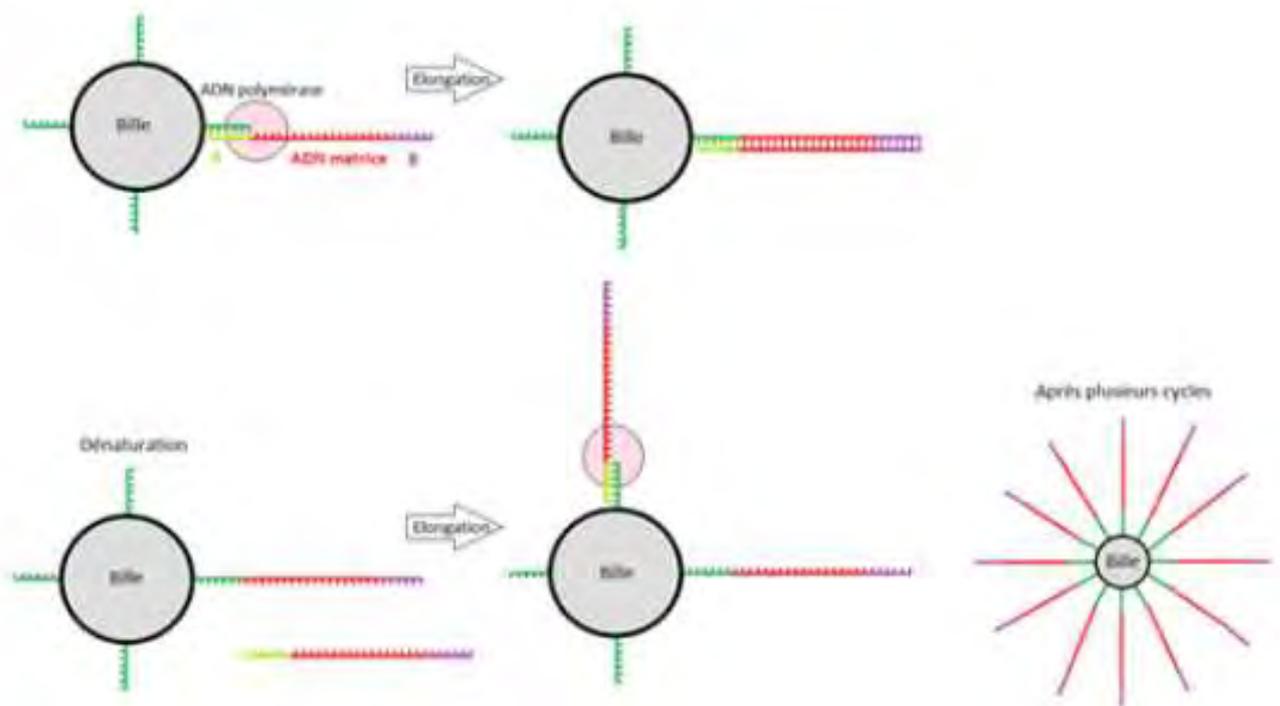


Figure 4- 4 : Principe de l'émulsion PCR (ePCR).

Appliquée à la protéine ABH2, l'équipe de Krylov a montré la limite de la NECEEM-PCR qui malgré 3 tentatives avec 5 cycles de SELEX d'interaction entre la protéine ABH2 et une banque d'ADN de 80 nucléotides dégénérée sur 20 nucléotides, ils n'ont obtenu aucun enrichissement. Au bout de 15 cycles d'amplification de PCR, ils ont montré l'apparition de sous-produits d'amplification qui aboutit vers le 30^{ème} cycle à la conversion quasi-totale des produits en sous-produits. Ces produits peuvent être dus à l'amplification de dimères d'amorces lorsque ces dernières sont présentes en trop grande quantité où alors dus à la formation d'hybrides d'ADN lorsque c'est la concentration de la matrice qui est trop élevée (Figure 4-5).

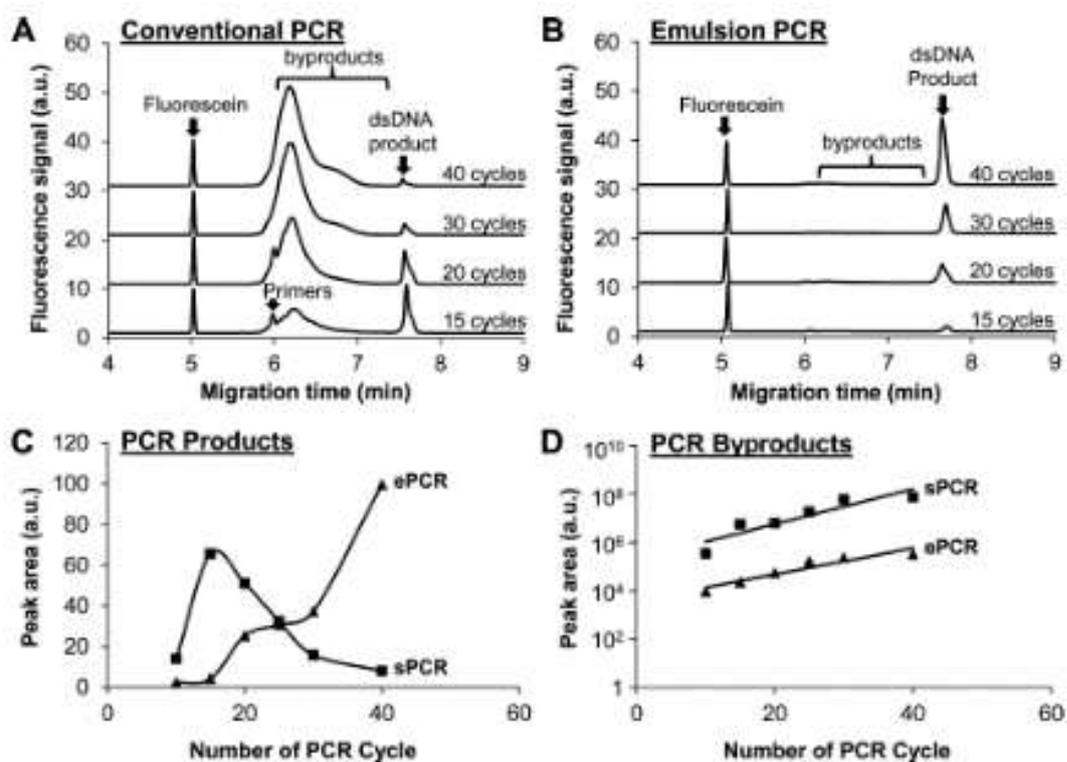


Figure 4-5 : Mise en évidence de la formation des sous-produits lors de la PCR conventionnelle et lors de l'émulsion PCR Issu de (140).

Avec l'ePCR, aucune conversion de produits n'a été observée et donc il n'y a pas eu de perte du produit d'ADN spécifique à amplifier. En utilisant l'émulsion PCR, ils ont ainsi pu obtenir dès le 4^{ème} cycle de sélection des valeurs de K_D allant de 0.1 μ M à 38 nM (Figure 4-6).

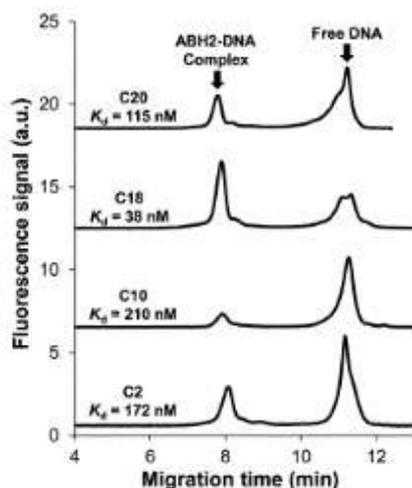


Figure 4-6 : Sélection par NECEEM de l'aptamère pour la cible ABH2. 200nM d'aptamères ont été mélangées avec 200nM de cibles. Issu de (140).

5 NEXT GENERATION SEQUENCING (NGS) POUR LA SELECTION D'APTAMERES :

5.1 Principe :

La récente accessibilité de techniques de séquençage à haut débit connue sous le nom de NGS (Next Generation Sequencing) a permis de caractériser des aptamères pouvant être sélectionnés avec moins de cycles de sélection. Les avantages de cette approche comprennent la très grande réduction de temps et des matériaux requis. Traditionnellement, le séquençage de Sanger était utilisé révélant 30 à 100 clones, qu'il fallait déjà produire et qui prenait plusieurs semaines. Avec le séquençage NGS, jusqu'à 100 millions de séquences peuvent être séquencées en une seule expérience, le multiplexage de plusieurs échantillons dans une séquence de séquençage permettant un séquençage en profondeur. Apparue en 2007, 3 technologies ont été développées mais globalement elles suivent toutes le même processus :

- 1) Préparation et amplification des molécules d'ADN à analyser
- 2) Incorporation des bases complémentaires du brin à analyser
- 3) Lecture de la séquence

5.2 Le pyroséquençage (Roche)

L'ADN est tout d'abord fragmenté, puis subit une étape d'amplification par PCR réalisée en émulsion. Le mélange obtenu est constitué de microbilles liées à deux enzymes : la luciférase et la sulfurylase ; Ensuite, ce mélange est déposé sur une plaque qui contient 1.6 millions de puits afin que chaque emplacement ne contienne qu'une unique molécule d'ADN. La réaction de séquençage par synthèse est alors initiée base par base par des réactions de pyroséquençage : Un flux de nucléotides triphosphate (dNTP) traverse la plaque et lorsque l'un d'entre eux est incorporé par la polymérase, un pyrophosphate (PPi) est libéré. Deux enzymes, réalisent ensuite une cascade enzymatique qui permet d'obtenir de l'ATP qui permet ensuite la conversion de luciférine en oxyluciférine par une luciférase. Cette réaction produit de la lumière et l'intensité lumineuse est capturée par une caméra qui permet de déduire la séquence en fonction de l'intensité de fluorescence (**Figure 5-1**). Cette technologie produit des séquences longues (330 bases en moyenne). Mais la méthodologie a été abandonnée, dépassée par d'autres modes plus récents de séquençage.

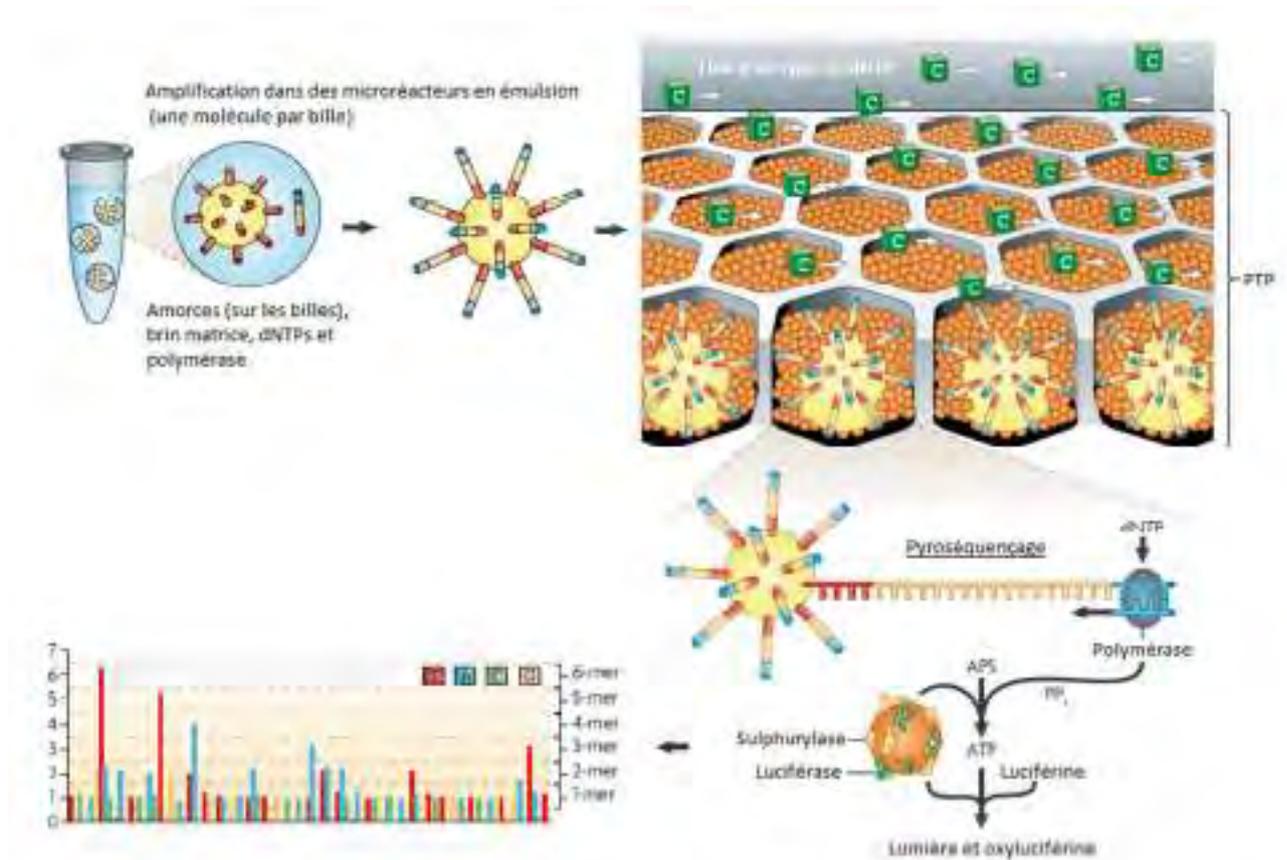


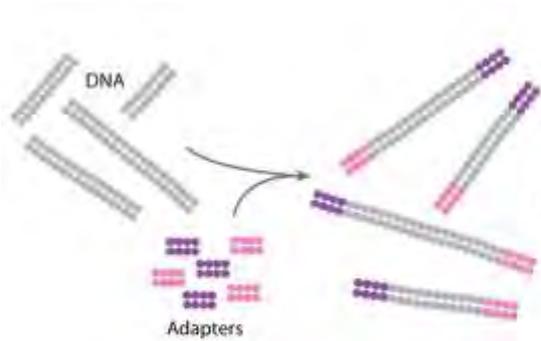
Figure 6: Aperçu de la technologie de séquençage 454

PTP : PicoTiter Plate, PPI : pyrophosphate inorganique, APS : adénosine phosphosulphate, ATP : adénosine triphosphate. Modifié d'après Metzker 2010 [98].

Figure 5-1: Principe de la technique du pyroséquencage. Issu de http://bebatut-edu.github.io/metagenomique_cm/sequencage.html#/2/8

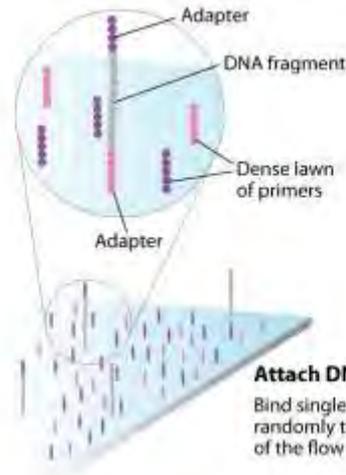
5.3 Principe du séquençage haut-débit Illumina :

Développée en 2006 par la société Solexa, cette technologie repose sur l'amplification sous forme de ponts (bridge) de fragments à séquencer. Pour cela, on utilise une lame de verre appelée *flow cell* divisée en 8 lignes. Cette dernière contient des adaptateurs qui vont se lier à chaque extrémité de la séquence d'ADN à amplifier. Lors de l'étape d'amplification par PCR, la polymérase synthétise le nouveau brin complémentaire pour former un pont d'ADN double brin puis les deux copies sont éliminées par dénaturation. Ce type de séquençage (contrairement au précédent) permet d'incorporer un nucléotide à la fois et répéter ainsi le cycle jusqu'à la taille attendue du fragment. La lecture du brin se fait en incorporant les 4 différents nucléotides marqués par fluorescence, chacun ayant une couleur différente après lavage et acquisition de l'image. Cette méthode permet l'acquisition en parallèle de plus de 3 milliards de séquences de 100 bases de long. Chaque position étant lue l'une après l'autre, les erreurs principales de cette technologie sont des erreurs de substitution d'une base par une autre (Figure 5-2).



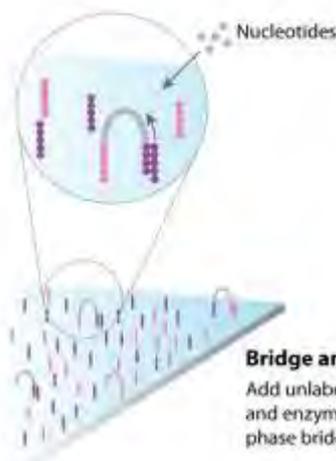
Prepare genomic DNA sample

Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.



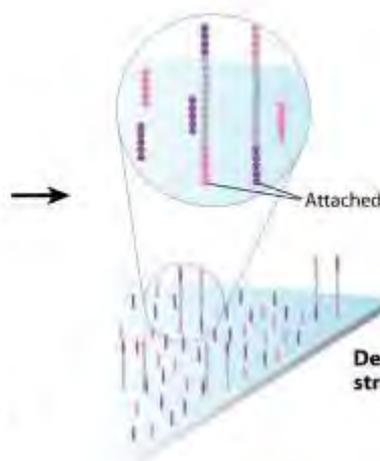
Attach DNA to surface

Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

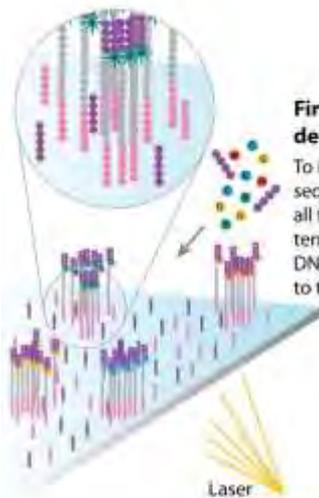


Bridge amplification

Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.



Denature the double stranded molecules



First chemistry cycle: determine first base

To initiate the first sequencing cycle, add all four labeled reversible terminators, primers, and DNA polymerase enzyme to the flow cell.

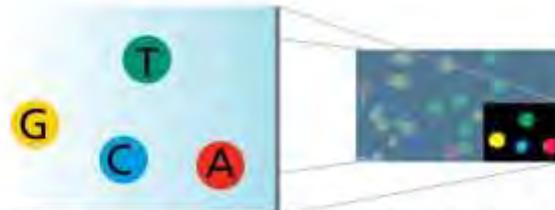


Image of first chemistry cycle

After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

Before initiating the next chemistry cycle

The blocked 3' terminus and the fluorophore from each incorporated base are removed.



Sequence read over multiple chemistry cycles

Repeat cycles of sequencing to determine the sequence of bases in a given fragment a single base at a time.

5.4 Le séquençage par ligation :

La première étape d'amplification de la méthode de séquençage d'Applied Biosystems est identique à celle du pyroséquençage. Par contre, les séquences amplifiées sont fixées sur un support solide au lieu d'une plaque. La réaction de séquençage s'effectue ensuite par un système assez complexe de cycles de ligation et de clivage. Cette technique va permettre non seulement la lecture de la séquence mais inclut un système de correction des erreurs d'incorporations. Grâce à cette méthode, il est possible de lire jusqu'à un milliard et demi de séquences en parallèle de 75 bases de long. Le système incorporé de correction d'erreurs associé à l'utilisation de la ligase rend cette technologie très fiable.

5.5 Application du séquençage haut débit :

Plutôt que d'utiliser le clonage, l'équipe de Whelan (101) a utilisé le séquençage haut-débit Illumina, qui génère un pipeline bioinformatique pour sélectionner des aptamères pour le biomarqueur HE4 du cancer ovarien. En effet, jusqu'à présent, l'étape finale après la sélection par électrophorèse capillaire consistait à cloner les oligonucléotides sélectionnés puis à les séquencer par la méthode de Sanger et à caractériser l'affinité des séquences identifiées pour la cible. Cependant, cette méthode a des limites (143), notamment le rendement avec une centaine de clones mais n'aboutissant qu'à l'obtention de quelques séquences. Ils ont réalisé plusieurs cycles d'incubation d'une banque d'ADN dégénérée à 10 μ M avec le marqueur HE4 présent dans le cancer ovarien, injection séparation en NECEEM et collecte. Ensuite, chaque échantillon a été amplifié lors d'une étape de PCR et purifié avec des colonnes de streptavidine afin de convertir le double brin en simple brin. Lors de l'étape de PCR, chaque échantillon a été assigné d'un primer unique dans l'amorce anti-sens permettant son identification lors de l'étape de séquençage. Après avoir été séquençé par Illumina, ils ont utilisé un programme bioinformatique disponible librement sur internet, utilisant le programme « Python »¹. Dans un premier temps, ils se sont d'abord assurés que le séquençage avait réussi en analysant les données brutes FastQC en envoyant ces données dans un pipeline Biopieces. Ensuite, chaque donnée a été lue et classée individuellement en utilisant la fonction *read_fastqc*. Puis les amorces ont été enlevées en utilisant la fonction *remove_primers*, en laissant une tolérance de non concordance de 5% et une tolérance de 0% pour les insertions et délétions. C'est lors de cette étape également qu'ont été éliminées les adaptateurs et les amorces de séquençage. Ils ont ensuite conservé les séquences ayant une taille de 25 nucléotides (+/- 2 nucléotides) en utilisant la fonction *grab*, et avec la fonction *uniq_seq*, chaque donnée a été enregistrée sous un fichier. Un tableau a été composé donnant les informations sur l'enrichissement par ordre décroissant avec la fonction *sort_records*. A partir de ce tableau, des analyses plus en profondeur ont été réalisées en déterminant l'enrichissement des parties de séquences dégénérées en fonction du nombre de cycles de sélection. Ceci a permis d'aboutir à la sélection des 1000 principales séquences les

¹ Ce programme est disponible à l'adresse suivante : <https://github.com/rebeccawhelan/PythonEnrichment>.

plus enrichies pour chaque cycle de sélection. Puis ces séquences ont été regroupées en fonction de leur homologie de séquences afin de déterminer l'apparition d'éventuels motifs. Un score a été attribué à chaque séquence en utilisant un logiciel MPBind qui calcule un score « meta-Z » en prenant en compte plusieurs paramètres. Ainsi 4 séquences ont été retenues pour tester leur affinité. Cependant on remarque que pour une de ces séquences sélectionnées, la corrélation entre ce score et l'affinité des séquences est faible, puisque le score « meta-Z » était négatif. L'ensemble du pipeline informatique utilisé est illustré dans la figure suivante (**Figure 5-3**). Ce même pipeline a été aussi utilisé pour identifier des aptamères pour un autre biomarqueur du cancer ovarien, le CA125 (127).

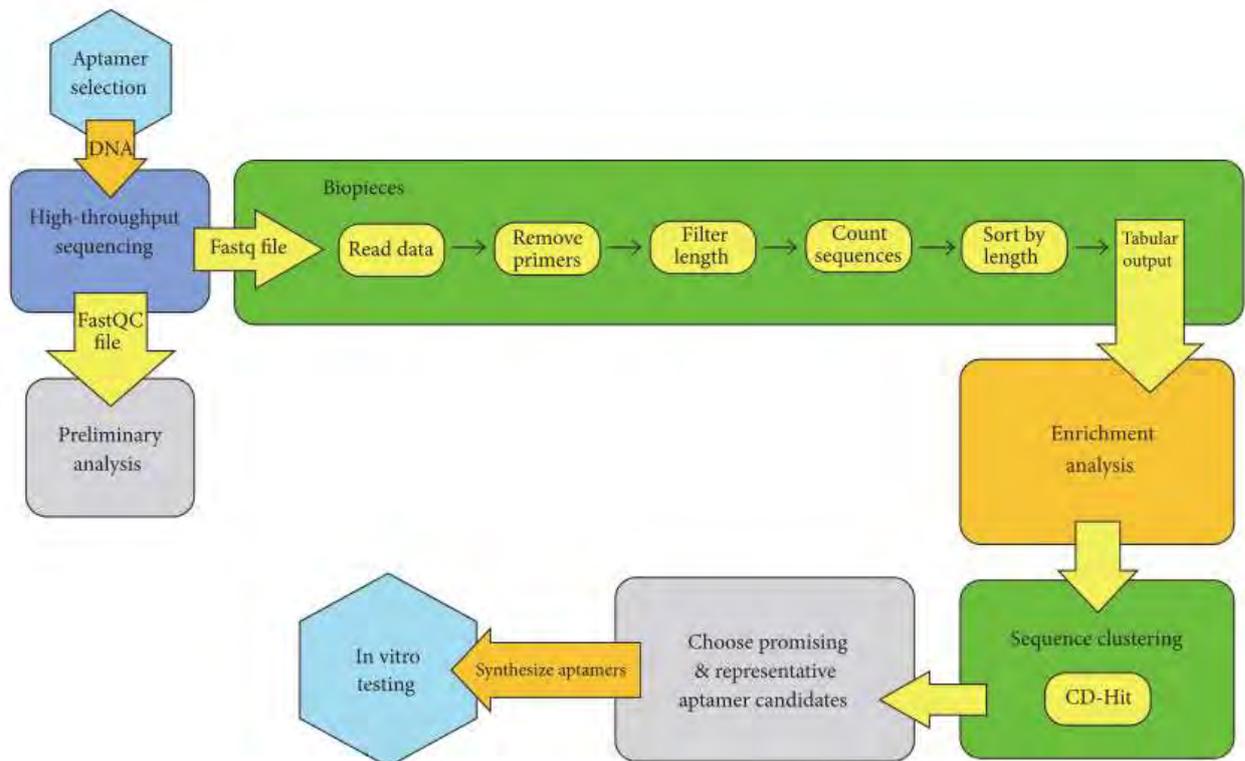


Figure 5-4 : Principales étapes de l'analyse bio-informatique des données du séquençage haut-débit. Issu de (127).

6. CONCLUSION :

L'ensemble de cette partie bibliographique nous rappelle bien que si l'électrophorèse capillaire est utilisée pour la sélection d'aptamère, elle ne fait partie que des méthodes possibles d'utilisation. Son application reste restreinte à un petit nombre de laboratoires dans le monde. Nous avons souvent cité le nom de Krylov qui travaille depuis plus de 20 ans sur les interactions entre macromolécules en électrophorèse capillaire. Mais nous pouvons aussi noter d'autres noms comme celui de Li ou de Coyle. Nous voyons que la NECEEM est la technique qui est surtout utilisée et que les progrès concernent l'avant ou l'après séparation. C'est-à-dire d'un côté l'influence des conditions d'obtention des complexes avec le rôle des ions et de l'autre les conditions de PCR et de séquençage.

Notre pierre à ce nouvel édifice sera d'examiner les conditions d'analyse en NECEEM, qui n'ont pas fait l'objet d'une attention particulière sur le choix de l'électrolyte tampon. Et d'un autre côté de répondre à la question si nous utilisons ces nouveaux modes de séquençage avec un minimum de réaction de PCR en CE-non SELEX, peut-on sélectionner un aptamère contenant un G quartet ?

Chapitre 1 : L'électrophorèse capillaire et l'utilisation de la double détection UV- LEDIF pour sélectionner, caractériser les cibles, les aptamères et les complexes aptamères-cibles :

Chapitre 1 : L'électrophorèse capillaire et l'utilisation de la double détection UV- LEDIF pour sélectionner, caractériser les cibles, les aptamères et les complexes aptamères-cibles :

La plupart des travaux effectués pour la sélection des aptamères en électrophorèse capillaire sont basés sur l'unique détection du complexe [ADN.cible]. Peu d'articles utilisent l'électrophorèse capillaire pour détecter la cible et l'aptamère simultanément (144). L'analyse des protéines et des petites molécules (le ligand) par électrophorèse capillaire peut être compliquée du fait qu'il faut regarder en absorbance UV leur temps de migration, changer de détecteur, espérer que leur temps de migration ne changera pas et détecter l'aptamère fluorescent en fluorescence, sans « voir » le ligand et en espérant voir le complexe. Dans le cas de la cocaïne, Zhou et al, (145) ont utilisé un aptamère greffé à un fluorophore (le FAM : carboxyfluoroscéine) et son brin partiellement complémentaire et dans un premier temps ils ont optimisé la séparation entre ces deux composés. Quand la cible a ensuite été incubée avec le mélange, l'équilibre a été déstabilisé ce qui s'est traduit par un changement d'intensité de fluorescence entre l'aptamère et le brin complémentaire. Le détecteur de fluorescence de Picometrics permet de réaliser une double détection par UV et par fluorescence. Deux fenêtres de détections sont réalisées, ce qui permet de détecter à la longueur d'onde de 210 nm la thrombine et par fluorescence, le complexe [T29.thrombine] et l'aptamère T29 libre. Dans le cas de la thrombine, la liaison de cette protéine à cet aptamère fluorescent conduit au changement du rapport charge/taille des complexes, ce qui rend les mobilités électrophorétiques du complexe, de la protéine et de l'aptamère différentes.

1 INFLUENCE DE L'UTILISATION DU TAMPON DE MIGRATION TRIS ACETATE SUR LA DEGRADATION DE L'ADN

1.1 Introduction

La nature et la force ionique du tampon peuvent influencer la migration des composés dans le capillaire et la séparation du complexe [aptamère.cible].

L'utilisation de l'électrophorèse capillaire permet d'obtenir les paramètres spécifiques des interactions entre protéines et ADN. Très souvent, ces analyses se font dans un tampon utilisant du Tris. Le tampon Tris-acétate est actuellement utilisé dans la sélection et la caractérisation de l'ADN simple brin par électrophorèse capillaire (CE). S'il existe une véritable constante dans l'analyse d'ADN par les techniques électrophorétiques, c'est d'utiliser un tampon TRIS alors que les tampons phosphates sont réservés aux analyses de protéines. Nous avons mis en évidence une difficulté réelle d'utiliser un tampon TRIS pour la sélection d'aptamères.

Nous avons également utilisé ce tampon lors des premières analyses de la thrombine et de l'aptamère T29 par électrophorèse capillaire. Cependant, nous avons remarqué que lors de l'analyse de T29, il est apparu des pics parasites en LIF au bout de plusieurs jours d'analyses

successives. De plus le pic de l'aptamère diminue jusqu'à disparaître complètement au bout de six jours (**Figure 1-1**).

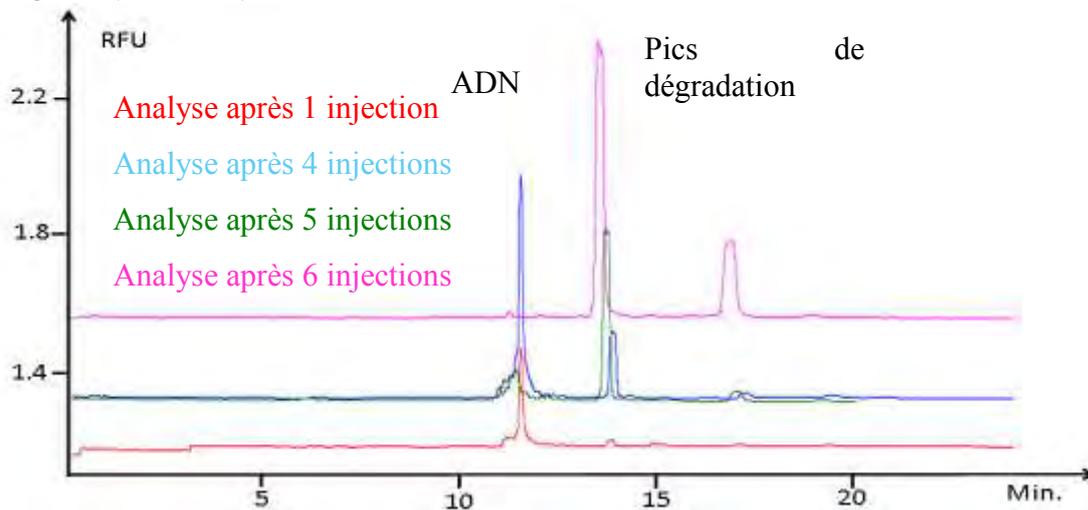


Figure 1-1 : Électrophérogramme d'un échantillon de T29 50 nM analysé au bout de plusieurs jours (6 jours) à l'aide d'un capillaire recouvert d'un polymère : le PVA (diamètre 50 µM – longueur totale 90 cm). Injection : 50 mBAR – 30 s Séparation : -30 kV Tampon de migration : tris-Acétate 50 mM pH= 8.2, On note l'apparition de pics de dégradation à t=14 min et t=17 min.

Ces pics peuvent être contraignants, car on peut les confondre avec des pics de complexe d'aptamères avec la cible et donner lieu à une fausse interprétation des électrophérogrammes. Pour tenter d'expliquer l'apparition de ce phénomène de dégradation, plusieurs hypothèses ont été formulées :

- L'aptamère T29 subit un changement de conformation lors de l'activation thermique ce qui masque le fluorophore (phénomène de quenching). En effet, lorsque nous réalisons l'interaction entre l'aptamère et la thrombine, ce dernier est chauffé à 95°C pendant 5 min afin de pouvoir le défolder, puis après être revenu à température ambiante, il est mélangé avec la cible afin de pouvoir réaliser l'interaction. Comme ce dernier a une conformation particulière en G-quadruplex, nous avons voulu vérifier la structure de l'aptamère T29 que nous utilisons pour nos analyses.

Pour vérifier cette hypothèse, nous avons analysé l'aptamère T29 par dichroïsme circulaire en chauffant l'aptamère à 90°C (**Figure 1-2**).

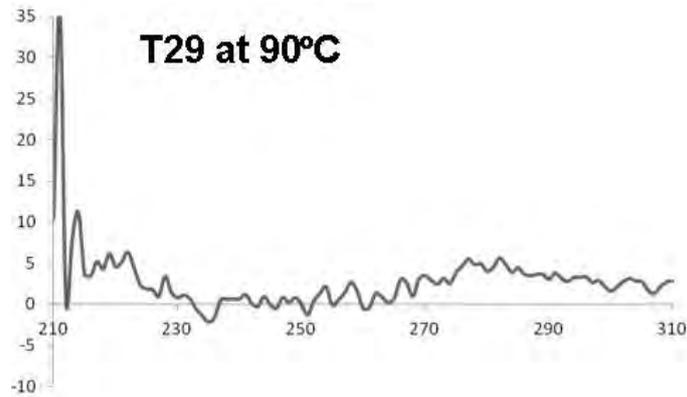


Figure 1-2 : Analyse du T29 chauffé à 90°C par dichroïsme circulaire. Lorsqu'il est chauffé à 90°C, l'aptamère se déplie et perd sa conformation initiale.

Puis l'aptamère a été de nouveau analysé par dichroïsme circulaire mais cette fois-ci à la température de 25°C qui est la température d'analyse en électrophorèse capillaire (Figure 1-3).

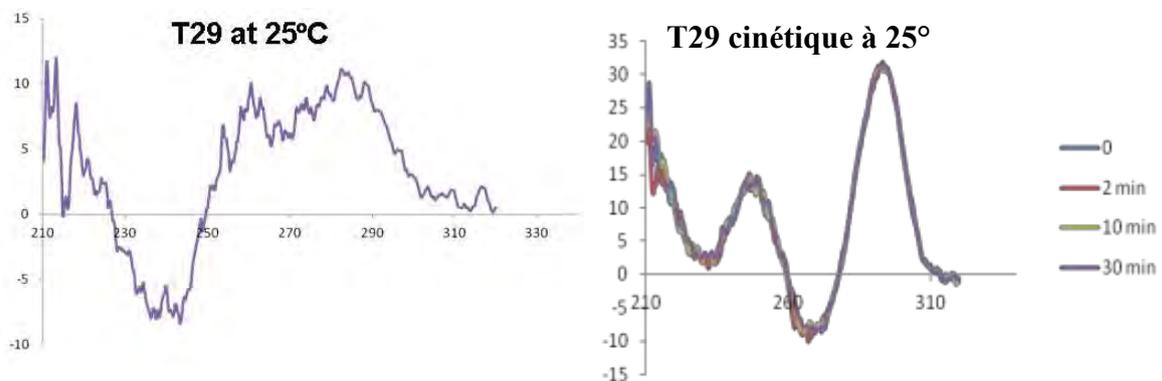


Figure 1-3 : Analyse du T29 à la température de 25°C par dichroïsme circulaire (figure à gauche). Ce dernier retrouve instantanément sa conformation initiale après avoir été chauffé à 90°C et cette structure est stable dans le temps (figure à droite).

L'analyse par dichroïsme circulaire montre que ce n'est pas un problème de conformation de l'ADN qui est à l'origine de sa dégradation. En effet, l'aptamère T29 chauffé à 90°C n'a pas de conformation structurale particulière et il retrouve instantanément sa structure en G-quadruplex à 25°C (Figure 1-3). On peut donc rejeter l'hypothèse selon laquelle T29 change de structure au cours du temps.

– Le tampon de dilution de l'échantillon (TrisHCl) utilisé pour diluer la solution mère de l'aptamère T29 dégrade l'aptamère T29 injecté en CE.

Pour vérifier si la dégradation de T29 était due au tampon Tris-HCl, deux solutions de T29 5 µM ont été préparées dans deux tampons différents : le tampon Tris-HCl 50 mM pH= 7.5 / NaCl 100 mM / MgCl₂ 1 mM et le tampon phosphate sodium dibasique 50 mM pH= 8.2

Ces solutions ont été analysées ensuite tous les jours. Il s'avère que dans le tampon phosphate (Figure 1-5), le pic de T29 reste le même et qu'il ne se dégrade pas alors qu'on observe une dégradation de

T29 dans le tampon TrisHCl+NaCl+MgCl₂ en LIF à 488 nm (Figure 1-4).

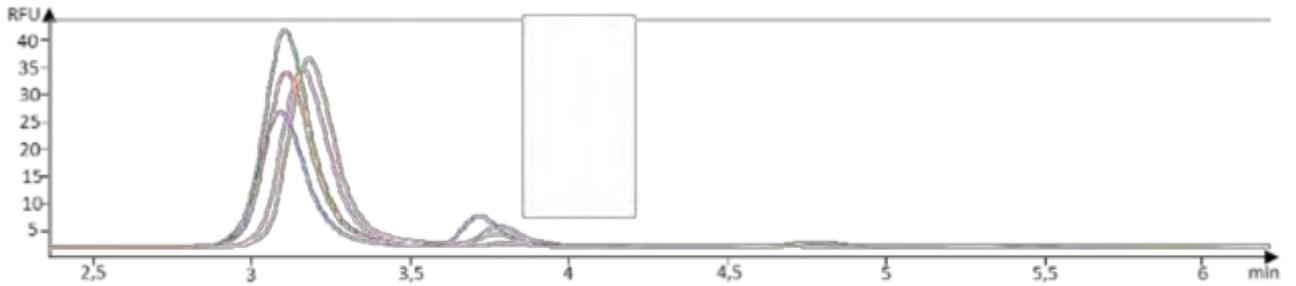


Figure 1-4 : suivi de la dégradation de T29 à 5 μM dans le tampon TrisHCl/NaCl/MgCl₂ avec une détection LIF 488 nm. En rose analyse de T29 5 μM lors de la première injection / En vert foncé analyse de T29 5 μM à lors de la 6^{ème} injection / En rouge analyse de T29 5 μM à la 7^{ème} injection / En bleu analyse de T29 5 μM à la 8^{ème} injection / En vert clair analyse de T29 5 μM à la dixième injection.

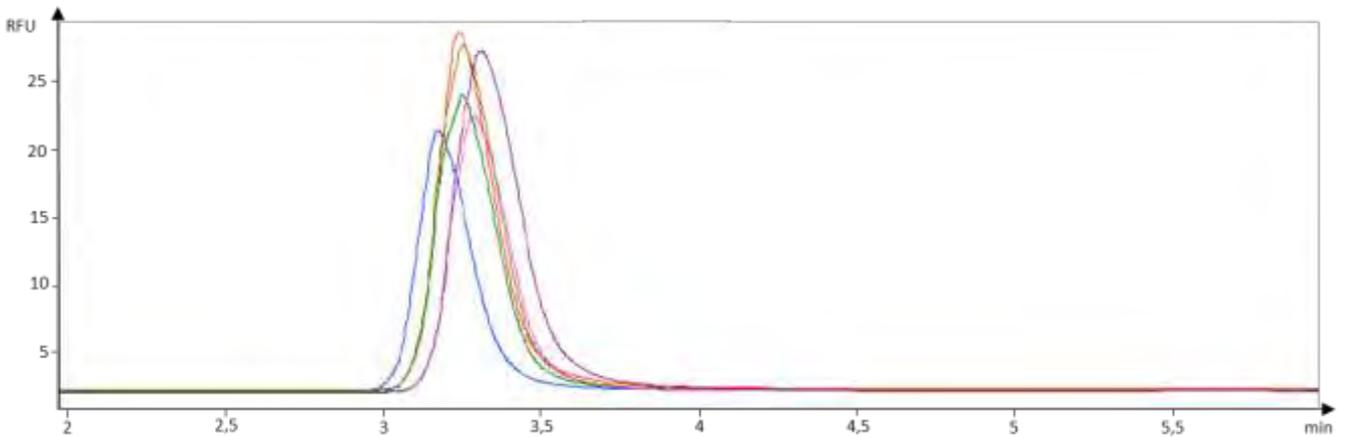


Figure 1-5 : suivi de la dégradation de T29 à 5 μM dans le tampon PBS avec une détection LIF 488 nm. En bleu analyse de T29 5 μM lors de la première injection / en rouge analyse de T29 5 μM lors de la 2^{ème} injection / en vert foncé analyse de T29 5 μM lors de la 5^{ème} injection / en rose analyse de T29 5 μM lors de la 6^{ème} injection / en vert clair analyse de T29 5 μM lors de la 7^{ème} injection / en violet analyse de T29 5 μM lors de la 8^{ème} injection.

– Le marqueur se détache de l'aptamère et le pic que nous détectons correspond à celui du marqueur fluorescent.

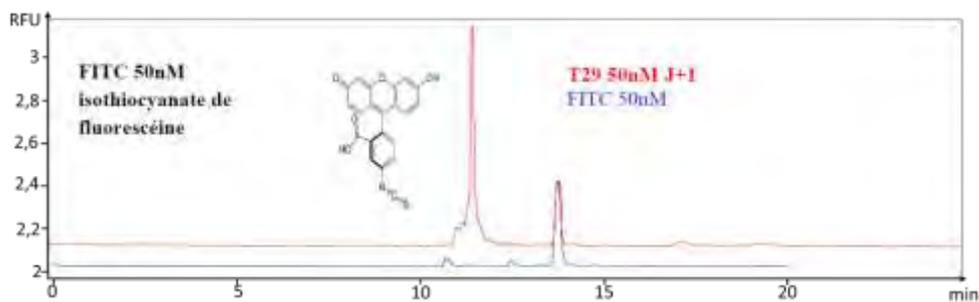


Figure 1-6 : superposition des électrophérogrammes de FITC à 50 nM et T29 à 50 nM à lors de la 2^{ème} injection. Mise en évidence du pic du FITC à $t = 13.8$ min (en bleu) se superposant au pic de dégradation de T29 (en rouge).

Lorsque l'on superpose l'électrophérogramme d'une solution de FITC à 50nM à celui du complexe greffé en 5' par le FAM,, les pics sont similaires. On peut donc assimiler le pic parasite au fluorophore (Figure 1-6).

Nous avons également déposé sur gel d'acrylamide 10% l'aptamère T29 FAM à 25 ng dilué dans le tampon Tris et un autre échantillon d'aptamère T29 FAM à 25 ng dilué dans le tampon phosphate de sodium après avoir été injecté 6 fois par électrophorèse capillaire.

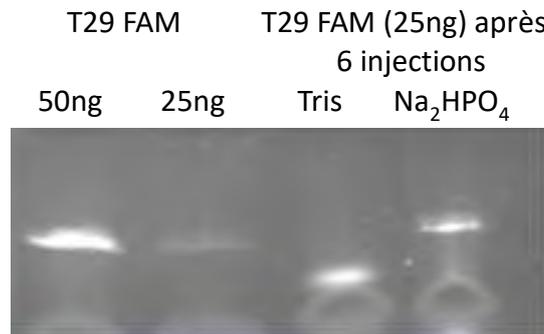


Figure 1-7 : gel de polyacrylamide 10% sur lequel a été déposé le T29 FAM à 50 ng et 25 ng non dégradés ainsi que le T29 FAM 25 ng dilué dans deux tampons différents (Tris et Na₂HPO₄) et injectés 6 fois en CE.

Seul le marqueur FAM est visible aux UV à 260 nm. Dans le tampon phosphate de sodium, le gel montre que le marqueur FAM est toujours présent, tandis que dans le tampon Tris, une bande est visible mais proche du front de migration, correspondant au FAM seul.

Ainsi, en injectant dix fois par électrophorèse capillaire de l'ADN simple brin (50 nM), l'oligonucléotide entier a été dégradé. Dans cet article, nous avons montré que le tampon Tris dans les flacons de migration est modifié le long du processus électrophorétique par des réactions électrochimiques. Nous avons également observé que la composition des ions métalliques change dans les flacons de tampon de fonctionnement. Ce phénomène, qui n'a jamais été décrit en CE, est important pour l'analyse par fluorescence de l'ADN simple brin utilisant du tampon Tris. Les oligonucléotides sont dégradés par des espèces synthétisées par voie électrochimique (présentes dans les flacons Tris en cours) jusqu'à leur disparition, même si le tampon de séparation dans le capillaire est neuf. Pour résoudre ces problèmes, nous proposons d'utiliser un tampon phosphate de sodium que nous démontrons être électrochimiquement inactif (**Figure 1-7**).

L'équipe de Sergeï Krylov a développé une autre méthode d'analyse des aptamères dans un tampon biologique : le PBS (phosphate buffered saline) (146). Leur principe consiste à développer l'utilisation de tampon biologique, pourrait trouver de nombreuses applications dans l'étude fondamentale des interactions biomoléculaires, la mise en place d'un diagnostic clinique et le développement d'agents thérapeutiques avec une cinétique définie (147). Les auteurs ont tout d'abord démontré que le PBS rendait l'ADN indétectable car le flux électroosmotique était instable ce qui rendait les pics d'ADN très large et avec de long temps de migration, du fait d'une faible mobilité de ce dernier dans le capillaire. Cependant, en ajoutant un peu de pression lors de la séparation, ils ont surpassé ces inconvénients et montré que l'utilisation du tampon PBS pouvait avoir des applications *in vitro*. Ils ont validé leur méthode et l'utilisation de ce tampon à la détection de la cible PDGF (platelet-derived growth factor) et de son aptamère. La valeur de la constante de dissociation, K_D obtenue pour l'aptamère de la PDGF a été de 6.2 ± 1.2 nM. Cette valeur est nettement plus faible que la valeur obtenue par la méthode NECEEM, pour laquelle les auteurs avaient trouvé une valeur de K_D de 56 nM.

1.2 Le tampon TRIS, un piège pour la sélection d'aptamères ?

1.2.1 Introduction :

Le tampon Tris-acétate est actuellement utilisé dans la sélection et la caractérisation de l'ADN simple brin par électrophorèse capillaire (CE). En appliquant une haute tension, la migration

d'espèces ioniques dans le capillaire génère un déplacement d'ions qui se traduit par la mesure d'un courant induit aux électrodes par l'électrolyse de l'eau. Ce phénomène peut être suivi par la modification du pH dans les flacons anodes et cathodes. Cette électrolyse de l'eau était bien connue et mesurée par divers auteurs (148). Dans l'étude de la répétabilité de l'analyse de la banque d'ADN nous nous sommes rendu compte que la production de dérivés du Tris avait aussi lieu. En effet, en injectant dix fois par électrophorèse capillaire de l'ADN simple brin (50 nM), l'oligonucléotide entier a été dégradé. Dans l'article qui suit, nous avons montré que le tampon Tris dans les flacons de migration est modifié le long du processus électrophorétique par des réactions électrochimiques. Nous avons également observé que la composition des ions métalliques change dans les flacons de tampon de fonctionnement. Ce phénomène, qui n'a jamais été décrit en CE, est important pour l'analyse par fluorescence de l'ADN simple brin utilisant du tampon Tris. Les oligonucléotides sont dégradés par des espèces synthétisées par voie électrochimique (présentes dans les flacons Tris en cours) jusqu'à leur disparition, même si le tampon de séparation dans le capillaire est propre. Les espèces produites à partir du Tris dégradent l'ADN et les produits de dégradation de l'ADN pourraient être confondus avec un possible complexe [Aptamère.ligant]. Pour résoudre ces problèmes, nous proposons d'utiliser un tampon phosphate de sodium que nous démontrons être électrochimiquement inactif (**Figure 1-8**).

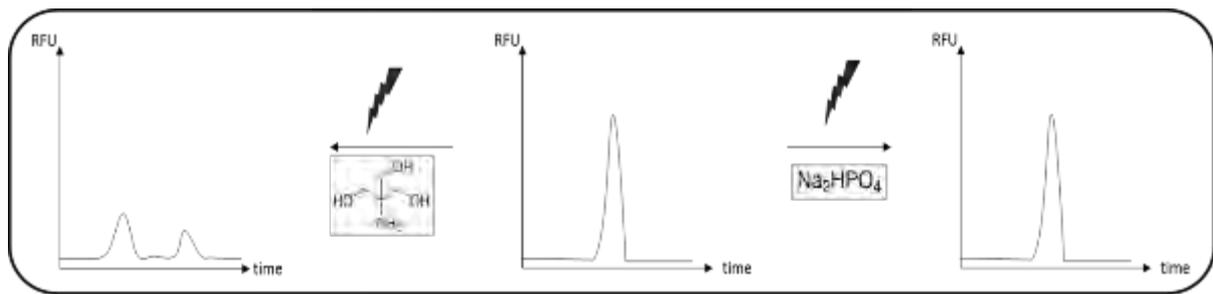


Figure 1-8: Résumé graphique du phénomène de dégradation de l'ADN simple brin lors de l'utilisation du tampon Tris par électrophorèse capillaire

1.2.2 Article :

Audrey Ric^{1,2,3}
 Varravaddheay Ong-Meang¹
 Verena Polinsot¹
 Nathalie Martins-Froment⁴
 Fabien Chauvet⁵
 Audrey Boutonnet⁵
 Frédéric Ginot³
 Vincent Ecochard²
 Laurent Paquereau²
 François Couderc¹

¹Laboratoire des IMRCP, UMR 5623, Université de Toulouse, Université Paul Sabatier, Toulouse, France

²CNRS, Institut de Pharmacologie et de Biologie Structurale, UMR 5089, Université de Toulouse, France

³Picometrics Technologies, Labège, France

⁴Service commun de spectrométrie de masse, Institut de chimie de Toulouse, Université de Toulouse, Université Paul Sabatier, Toulouse, France

⁵Laboratoire de Génie Chimique, UMR 5503, Université de Toulouse, Université Paul Sabatier, Toulouse, France

Received December 22, 2016

Revised February 17, 2017

Accepted February 18, 2017

1 Introduction

It is well established that electrochemical phenomena happen beside electromigration and electroosmosis during capillary electrophoresis (CE). As the CE separation requires high voltage, to induce the migration of ionic species, an electric current is generated. On the electrodes, immersed in the buffer, electrolysis of water happens. In consequence, the pH in the buffer vials changes. This phenomenon is called buffer depletion and has been described by many authors [1–4]. The modification of the buffer pH is mainly due to the poor buffering capacity of the background electrolyte and also the small volumes of the vials. This can be overcome by a frequent refreshing of the buffer vials or, when possible, by changing the buffer composition. Among the commonly used buffers

Correspondence: Prof. François Couderc, Laboratoire des IMRCP, UMR 5623, Université de Toulouse, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France
 E-mail: couderc@chimie.ups-tlse.fr

Abbreviations: CRB, CE running buffer; ECGS, electrochemically generated species; HRMS, high resolution mass spectrometry; NCRB, non-CE running buffer; Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol

Research Article

ssDNA degradation along capillary electrophoresis process using a Tris buffer

Tris-Acetate buffer is currently used in the selection and the characterization of ssDNA by capillary electrophoresis (CE). By applying high voltage, the migration of ionic species into the capillary generates a current that induces water electrolysis. This phenomenon is followed by the modification of the pH and the production of Tris derivatives. By injecting ten times by capillary electrophoresis ssDNA (50 nM), the whole oligonucleotide was degraded. In this paper, we will show that the Tris buffer in the running vials is modified along the electrophoretic process by electrochemical reactions. We also observed that the composition of the metal ions changes in the running buffer vials. This phenomenon, never described in CE, is important for fluorescent ssDNA analysis using Tris buffer. The oligonucleotides are degraded by electrochemically synthesized species (present in the running Tris vials) until it disappears, even if the separation buffer in the capillary is clean. To address these issues, we propose to use a sodium phosphate buffer that we demonstrate to be electrochemically inactive.

Keywords:

Biomolecules / Capillary electrophoresis / Oligonucleotides / ssDNA degradation / Tris-Acetate buffer
 DOI 10.1002/elps.201600561

2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris) is very popular and can be employed for many CE experiments run at pH around 8.5. The main application of this buffer concerns slab gels and CE applications in nucleic acids separations. Actually, the sentence “for reasons not fully evident today, Tris became established as the favored cation for DNA electrophoresis” opens the “Tris reigned” paragraph of an article concerning the history of DNA separations [5]. Today, most of DNA separations by gel and capillary electrophoresis use Tris based buffers. It is still commonly employed for ssDNA studies [6], PCR using CE [7], or on-line preconcentration studies of DNA fragments [8]. Twenty years ago, Ray et al [9, 10] have shown that during gel electrophoresis the DNA of two *Streptomyces* species undergo a Tris-dependent strand scission. The authors indicated that the nucleolytic activity acts at the modification sites and implies a peracid derivative of Tris. Formed at the anode, it can be chemically mimicked by addition of peracetic acid to Tris buffers. The aim of this “ancient” paper (only cited 46 times) was to prevent the electrophoretic community to neither use the Tris buffer nor amine containing buffers, but to prefer buffers using 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES). This recommendation was poorly followed by the analysts using CE. Indeed,

Colour Online: See the article online to view Fig. 3 in colour.

they claimed that the capillary diameter is small compared to the buffer vial volume, and that the electro-synthesized species of the back-ground electrolyte cannot interfere with the injected DNA.

In this paper, we will show that the same ssDNA degradation phenomenon can be observed and becomes very important for fluorescent ssDNA like aptamers analysis in CE using Tris buffer. We will show that this buffer is modified along the electrophoretic process by electrochemical reactions, as it was observed that the composition of the metal ion changes in the buffer vials. The electrochemically synthesized molecules are produced and pollute the sample even if the separation buffer in the capillary is clean.

2 Materials and methods

2.1 Chemicals

Tris (two different lots were studied), acetic acid, sodium phosphate dibasic anhydrous, sodium chloride, hydrochloric acid, magnesium chloride, sodium hydroxide, ammonium persulfate, xylene cyanol, and bromophenol blue were purchased from Sigma-Aldrich - Fluka (St Quentin Fallavier, France). Urea was from Euromedex (Souffelweyersheim, France), TEMED (N,N,N',N' Tetramethyl-1,2-diaminoethane) from Eurobio (Courtaboeuf, France), acrylamide (Fischer Scientific Illkirch, France). The Sequence of ssDNA was the following: 5'-AGTCCGTGGT AGGGCAGGTTGGGGTGACT-3'. ssDNA was labeled with a fluorophore, 6-carboxyfluorescein: FAM in 5' end (Sigma-Aldrich—Fluka).

2.2 Instrumentation

An Agilent Technologies CE 7100 instrument was equipped with a 50 μm id fused silica capillary (Polymicro Technologies, Phoenix, Arizona) or a PVA coated capillary (Agilent, Walbronn, Germany) effective length 69 cm, total length 90 cm. Zetlif LED detector from Picometrics Technologies (Toulouse, France) was equipped with a LED at 480 nm. For some experiments, the electrodes were cleaned using ultrasonic bath in ethanol/water (50/50, v/v). CE experiments were performed with 50 mM Tris-Acetate pH 8.2 as BGE, the volume of the vials at anode and cathode was 1.5 mL (the volume of buffer that can be contained in a vial is 1.5 mL maximum). +30 kV for an uncoated capillary (current intensity: 30 μA) or -30 kV (current intensity: -30 μA) for a coated capillary. The 50 mM Tris-Acetate buffer at the inlet and outlet was formerly described for ssDNA analysis by CE [11].

Samples of ssDNA were prepared in 50 mM Tris-HCl with 100 mM NaCl, 1 mM MgCl_2 . The volume of ssDNA sample is 20 μL . The ten injections of ssDNA (50 nM) were

performed by pressure (50 mbar during 30 s). The duration of the entire process (injections and separations) was 300 min.

Polyacrylamide gel was prepared by weighing 10.5 g of urea diluted in 5 mL of Tris-Borate EDTA (TBE) 5 \times . A 6.25 mL of acrylamide 40% was added in the mixture. The reaction of polymerization was then initiated by the formation of free radicals by the persulfate ammonium (150 μL of 10% solution) and catalyzed by the addition of TEMED (15 μL). A 20 μL of ssDNA sample (50 nM) was hydrodynamically injected six times in CE and then dropped in the polyacrylamide gel. A fresh ssDNA sample is used as the reference. Samples of ssDNA were diluted in a buffer consisting of 95% of formamide, 20 mM of EDTA, 0.05% of xylene cyanol, 0.05% bromophenol blue, and glycerol. A 20 μL of ssDNA at 50 nM was dropped on the polyacrylamide gel. Samples of ssDNA migrated in a TBE - 7M urea, buffer with 100 V. After migration, ssDNA samples were revealed with ethidium bromide under UV illumination or by fluorescent detection (λ_{exc} : 480 nm).

LC/MS experiments: UPLC Acquity Xevo G2 QTOF Waters, a column BEH C18 (1.7 μm ; 2.1 \times 15 mm) and a gradient ACN 3–30% in water with 0.1% formic acid in 4 min with a flow rate of 0.3 mL/min, bearing of 1 min at the beginning, were used at $T = 35^\circ\text{C}$. Mass spectrometry calculations were done using ChemCalc, [12].

A Thermo-Finnigan Neptune MCICP-MS was employed to determine the relative concentrations of elements in the Tris solutions. These solutions were diluted with precision between 4.5 and 5.1 times in 2% HNO_3 .

3 Results and discussion

3.1 ssDNA degradation in Tris buffer during CE

During our work, we washed the uncoated or coated capillary with a Tris buffer which was never submitted to the high voltage (HV). The anode and cathode vials during the separation (with the HV) were changed every ten runs. When we injected several times (around ten injections) a sample containing 50 nM solution of a FAM labeled ssDNA, we observed a decrease of the ssDNA peak until its extinction was raised. A concomitant increase of two peaks was observed (Fig. 1A). Meanwhile, the pH of the anodic vial decreased slightly (from 8.7 to 8.3) after ten separations of 30 min each. The ssDNA degradation phenomenon was also observed using the PVA coated capillary. In these analyses, the electroosmotic flow is negligible and doesn't change the migration time of the ssDNA. These analyses confirm that the ssDNA is degraded.

We decided to load ssDNA samples on a denaturing urea polyacrylamide gel, to control its size after one injection and after six injections by CE. Two samples of ssDNA diluted in Tris-Acetate buffer were analyzed: the first one is the control ssDNA CE sample (50 μL) with the non-CE run buffer (NCRB) (Fig. 1B - lane a); this sample was injected

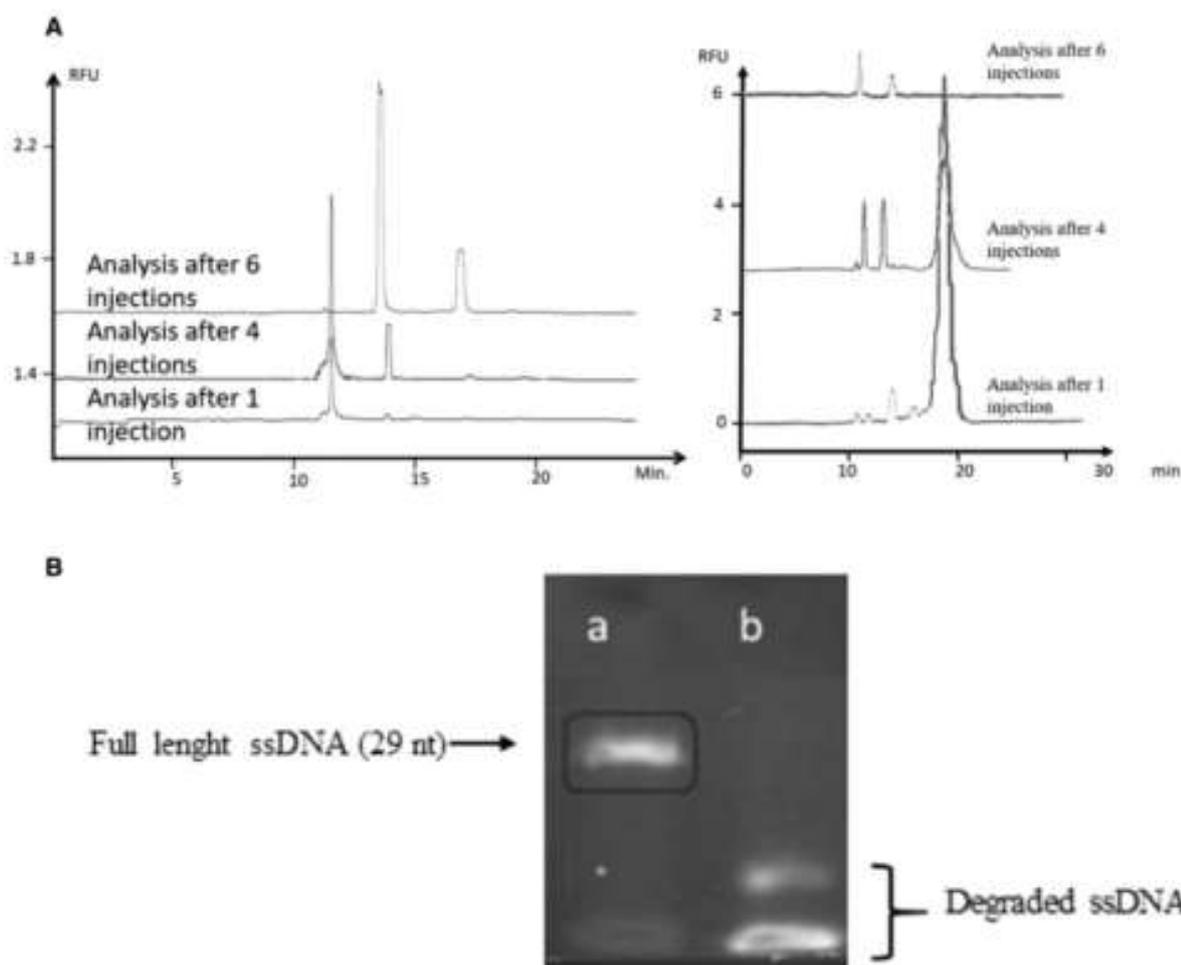


Figure 1. (A) Analysis of ssDNA in Tris-Acetate migration buffer (50 mM pH 8.2) during six injections with a 50 μm id, PVA coated capillary (Fig. 1A on the left). Identification of degradation peaks at 14 and 17 min. The same experience was realized with an uncoated capillary with a 50 μm id (Fig. 1A on the right) and ssDNA injected six times. Identification of degradation peaks at 11 and 13 min. (B) polyacrylamide gel 7 M urea of (a) the ssDNA sample injected six times in CE. The revelation of the fluorophore (FAM) is performed with UV at 480 nm. Two fluorescence spots are observed at low size, but no smear of ssDNA is observed which indicates that the ssDNA is affected by electrochemical radicals and degraded. (b) ssDNA-FAM not injected in CE.

one time and the electropherogram showed any degradation. The second one is the same ssDNA CE sample but injected six times in CE, with the same CE run buffer (CRB) at anode and cathode: the electropherogram confirmed its degradation and the gel-run with the residual sample in the vial (injected 6 times) demonstrated a complete alteration (Fig. 1B - lane b). We could only detect, by fluorescence at 480 nm two fluorescent signals with low DNA size. These two fluorescent signals corresponded to free fluorophore and to very small DNA fragments. When ethidium bromide was used to detect the ssDNA after 6 CE injections, no band could be identified, while the control ssDNA can be observed (data not shown) indicating the smallness of DNA fragments. This observation highlights that the ssDNA sample was affected by the CE buffer and degraded into the vial. We assumed that

electrochemical generated species (ECGS) were produced from the Tris inside the electrode vials and proceed on ssDNA in the vial. Furthermore, we have shown that the ssDNA degradation was associated with its concentration. When the concentration of ssDNA was increased (from 50 nM to 1 μM), the degradation process was lengthened. This was probably due to the fact that we used the same final concentration of Tris and therefore the amounts of ECGS were the same. Thus, with a larger amount of ssDNA (1 μM injected), the ECGS mediated degradation occurs but slower: the ssDNA peak is constant during 13 injections and then decreases rapidly (data not shown).

To better identify the electrochemical processes in the buffer vials, we measured the influence of the electrode material in the buffer submitted to electrophoresis.

3.2 Analysis of Tris buffer before and after CE separation by ICP/MS

What is the behavior of the electrode along the CE run? The electrodes of the CE instruments are made of platinum/iridium. Nothing is published about how the Pt/Ir electrode can act during the electrophoresis. Therefore, we wanted to know how the electrode is affected during the run. We measured the different amounts of sodium, potassium, platinum, and iridium in Tris vials (cathode and anode) after 200 min of the high voltage CRB, compare to the same vials before the CE running NCRB. Compared to the initial value (NCRB), we observed an increase of 55% Na⁺ (70 μM) in the anode CRB due to traces of NaOH which was used to wash the capillary. As washes were done with an empty vial at the cathode, we did not have any increase of Na⁺ in the cathode CRB compare to the NCRB. An increase of 33% of potassium (35 μM) at the anode compared to the cathode was

more surprising, because we used potassium buffer several weeks ago (no increase at the cathode). In consequence, the behavior of sodium and potassium on the platinum/iridium electrode showed a retentive effect. It seems that these ions are adsorbed on the electrode and are only slowly released.

We also noticed an increase of 5 and 6-fold of the Si at the anode and cathode respectively, compared to NCRB. The huge increase of the concentration of Pt (estimated concentration at the anode 25 nM, more than 2220- and 1200-fold at the anode and the cathode respectively) and Ir (8 nM, 2000, 800) showed that an oxidation was realized from the surface of each electrode. It was known that when the Pt was used for electrochemical reactions, it was first oxidized. Then the PtO surface electro-reacted with the electrolyte [13]. In our case the electro-osmotic flow and electromobility help the metallic species to move to the cathode.

We conclude that during the CE process, even if the current was very low (30 μA), we observed some amounts of Pt

Table 1. List of the principal chemical formulas of molecules (ECGS) identified, analyzed in ESI^{+/−} HRMS at the anode and cathode vials of Tris. DB, number of double bonds

Experimental mass	Absolute intensity	Theoretical mass	Proposed chemical formula	DB	Δ (mDalton)
Anode ESI+					
103.0623 ^c	1.3 E4	103.0633	C4H9N02	1	−1.0
113.0851 ^b	4.0 E3	113.0840	C6H11N0	2	1.1
121.0749 ^a	2.11 E4	121.0739	C4H11N03	0	1.0
124.0644	3.3 E4	124.0636	C6H8N20	4	0.8
131.0593	1.10 E4	131.0582	C5H9N03	2	1.10
139.0630	2.7 E3	139.0633	C7H9N02	4	−0.30
145.0752	5.2 E3	145.0738	C6H11N03	2	1.40
157.0751	2.4 E4	157.0739	C7H11N03	3	1.21
173.0684	3.8 E3	173.0688	C7H11N04	3	−0.40
175.0861	9.8 E3	175.0844	C7H13N04	2	1.7
205.0285	1.7 E4	205.0276	C9H5N204	8.5	0.9
210.1026	2.8 E3	210.1034	C10H14N203	5	2.2
228.1132	3.2 E3	228.1110	C10H16N204	4	2.2
244.1064	2.2 E3	244.1059	C10H16N205	4	0.5
294.147	2.2 E4	294.1440	C12H18N603	8.5	3.0
a: Tris; b: impurity in Tris; c: this molecule corresponds to an ion after the loss of H ₂ O of Tris, in the ESI+ source.					
Cathode ESI+					
117.0588	0.9 E3	117.0578	C8H7N	6	1.0
131.0596	3.5 E3	131.0582	C5H9N03	2	1.4
133.0749	2.6 E3	133.07389	C5H11N03	1	1.0
157.0753	6.7 E3	157.0738	C7H11N03	3	1.4
159.0225	3.2 E3	159.01945	C8H3N202	8.5	3.0
196.0843	5.0 E3	196.0848	C9H12N203	5	0.5
206.0146	1.7 E3	206.0175	C5H6N207	4	2.9
214.0955	4.2 E3	214.0953	C9H14N204	4	0.1
244.1064	2.5 E3	244.1059	C10H16N205	4	0.5
Anode ESI−					
113.9928 ^b	4.0 E4	113.9953	C4H204	4	2.5
121.0753 ^a	1.9 E4	121.0749	C4H11N03	0	0.4
135.0554	4.9 E4	135.0531	C4H9N04	1	−2.2
149.0698	4.3 E4	149.0688	C5H11N04	1	−1.0
258.0978	3.1 E5	258.0978	C11H16N06	4.5	−0.0
260.0951	1.0 E5	260.2852	C14H14N04	8.5	−2.8
268.1265	2.1 E5	268.1271	C9H20N207	1	0.5

a: Tris; b: impurity in the Tris, greyed lined are identical chemical formulas in anodes and cathode

(or PtO) and Ir which were displaced into the buffer. This highlights that the electrochemical processes can be important on the electrode during CE.

3.3 Analysis of ECGS by HRMS

Using UPLC/ESI⁺ HRMS or UPLC/ESI⁻ HRMS we identified different compounds in the Tris buffer. The major one was the Tris (C₄H₁₁NO₃) itself.

After 10 runs of the same Tris CRB from the anode and cathode vial, were analyzed in ESI⁺/HRMS and MS/MS. Its composition was slightly changed. We identified at least 24 new compounds. Table 1 lists the different chemical formulas of the molecules which were identified, while we proposed some structures considering the reaction between different Tris molecules, in Fig. 2. We noticed that we obtained unsaturated molecules. The UV spectrum at an absorption wavelength of 280 nm confirmed the formation of aromatic molecules.

Using ESI⁻ ionization, new compounds were identified, and presented in Table 1. The C₂H₂O₄ acid corresponded to the oxidation of a Tris molecule and the two other compounds resulted in the condensation of at least three Tris molecules. One corresponded to the mass of a radical which can be stabilized thanks to electron conjugation. None peracetic species could be identified.

In the literature, to our best knowledge, we did not identify a work reporting on the electrochemical modification of the Tris molecules. It was well known that phenolic compounds can be oxidized at the platinum anode. Actually, when phenol is adsorbed on the electrode it becomes conductive and its oxidation involves ring cleavage, leading to simple organic acids. It can also be oxidized in hydroquinone, benzoquinone, phenolic ether, biphenol, and polymers, induced by phenoxy radicals [14]. Diols were also oxidized and resulted in lactones [15]. Likewise, another well-known electrochemical oxidation was the Shono oxidation. Shono's protocol generates N-acyliminium ions from the oxidation of amides [16]. Polyaniline was reported to be a very good conducting

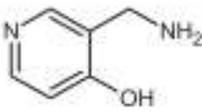
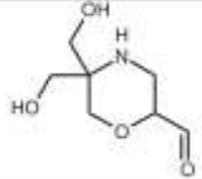
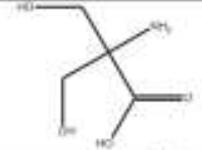
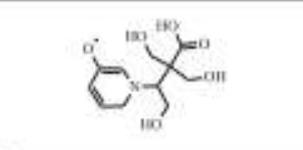
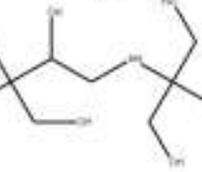
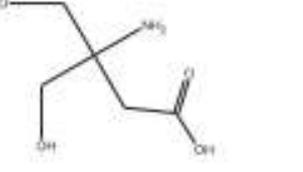
Chemical structure	M _{exp}	M _{exp}	Chemical structure
Cathode			
	124.0644	157.0751	
	175.0861	173.0684	
Anode			
	135.0531	258.0978	
	268.1265	149.0688	

Figure 2. Proposal of some chemical structures of molecules identified in ESI⁺ and ESI⁻.



Figure 3. Proposal of chemical reactions of degradation of ssDNA following the scheme proposed by [21]. R = nucleic base; R2 and R3 are the following nucleotides.

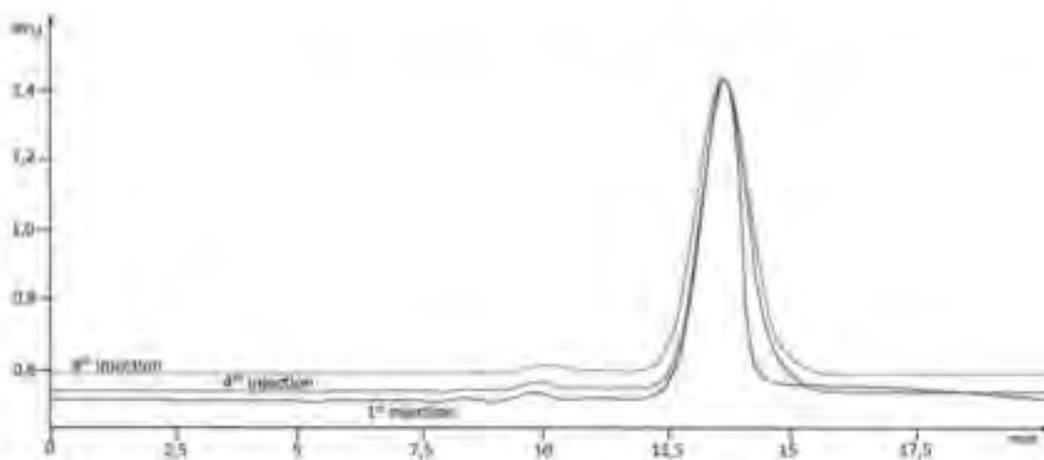


Figure 4. Electropherogram of ssDNA in 50 mM phosphate buffer (pH 8.6) after one, four, and eight injections.

organic material, and can be synthesized on the Pt electrode. It has been shown that with amines, the reactivity of radicals can also have induced such polymerization reactions to obtain polyaniline [17]. Such electrochemical reactions were also observed with pyrrole [18] or 1, 5-diaminonaphthalene [19]. In consequence, many compounds can be synthesized using electrophoresis in the buffer. All these studies highlighted that more or less complex electrochemical reactions can occur on aromatic compounds. From our results, even if Tris was not an aromatic molecule, we showed that we had also a very high number of different electrochemical reactions.

Among the molecules which were electrochemically synthesized, two radicals can be identified at the anode while we had only one at the cathode. We can imagine that the radicals presented in Fig. 2 could initiate diverse reactions on the ssDNA.

We supposed that we can have the following reactions at the anode:

Reactions of Tris at the platinum anode

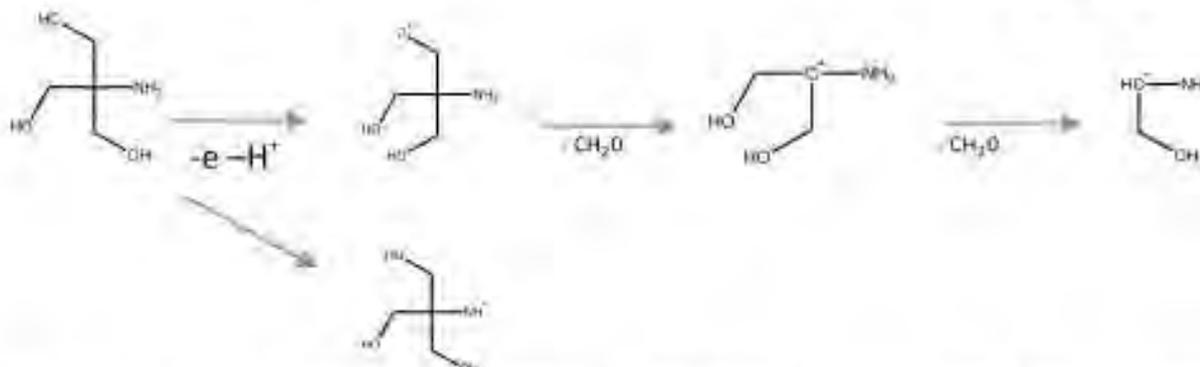
The four radicals can react on Tris and form different adducts able to lose H_2O , CO , CO_2 , NH_3 , and other simple molecules. We expected that these radicals reacted also with Tris molecules and result in products which can evolve in the

different C5 to C11 compounds of Table 1. These molecules can as well lose electrons and undergo other electrochemical reactions.

The cathode CRB composition also differed from the initial Tris sample. It contained low quantities of the compounds identified at the anode vial, certainly thanks to the electroosmotic flow. We found molecules ranging from C5 to C10, and an aromatic molecule (C_6H_7N). In the literature, the cathodic reduction of quinones was well documented [20, 21], it showed that ECGS can also be synthesized at the cathode.

Finally, we remarked that some impurities were present in the commercial Tris. Low amounts of C_6H_7NO (an amine which can be seen in ESI^+) and $C_6H_7O_4$ (an acid visible in ESI^-) were present even before electrophoresis. These molecules could also have some importance in the described electrochemical synthesis.

Because radicals were produced as well in the anodic as in the cathodic vials, several side reactions will be undergone in presence of substrate. These radicals can react on ssDNA in the presence of metals (Pt, Ir) and degrade it totally, as mentioned in many papers [22]. As reported elsewhere, we present in Fig. 3, a proposal of the chemical reactions which explains the total degradation of the 50 nM ssDNA sample.



A question remains: how long can the ECGS radicals, produced in the running buffer, react on ssDNA?

3.4 How stable are the active ECGS molecules?

To investigate the stability of the ECGS molecules, a Tris-Acetate buffer was run with the CE during 400 min. A first separation of the ssDNA was run with this 400 min buffer by injecting a 50 nM ssDNA sample. The second separation was run 24 h later with this same Tris buffer (run during 400 min, 24 h sooner) and a novel ssDNA (50 nM). In the first run, we only observed the degradation peak, the ssDNA peak disappeared. In the second run, we observed the degradation peak at the same migration time but the ssDNA peak decreased only of a factor 2 compared to the ungraded ssDNA.

This experience showed that ECGS were quite stable molecules because the half of ssDNA was still degraded when the buffer was rested.

3.5 The use of phosphate buffer

Our results obtained with Tris-Acetate buffer showed that this buffer was not appropriate for separating low concentrations of ssDNA. So, we decided to use a buffer without nitrogen atoms. We chose a phosphate buffer (pH 8.6) as propose by Kanoatov and Krylov [25]. To prove the efficiency of phosphate buffer, we prepared different concentrations (50, 25, 10, and 5 mM). As the different concentrations of phosphate buffer did not induce the ssDNA degradation (Fig. 4), we preferred to use 50 mM concentration because the buffer capacity is better. In these conditions, we did not observe the degradation of ssDNA after eight injections in phosphate buffer corresponding to 160 min of CE runs while it was observed after six injections in Tris buffer (corresponding to 180 min of CE runs).

4 Concluding remarks

Tris buffer is degraded during the CE experiments; many different compounds are electrochemically synthesized. Among

them we were able to identify radical species. These compounds can react on the ssDNA and induce its total degradation. When the ssDNA from the sample vial is observed using slab gel separation after 10 runs, no ssDNA can be observed; solely the fluorescein is detected. This shows that the degradation of the ssDNA is induced in situ. Using β phosphate buffer, no ssDNA degradation is observed.

We thank Mr. Frédéric Candau (Service ICP-MS, Géosciences Environnement Toulouse, Université de Toulouse) for his kind help in ICP-MS experiments.

We are very grateful to Dr André Saul for the discussions about the electrochemical synthesis.

We thank The Conseil Régional Occitanie for its financial support.

The authors have declared no conflict of interest.

5 References

- 1) Kelly, M. A., Alfria, E. D., Clark, B. J., *J. Chromatogr. A* 1997, 768, 73–80.
- 2) Shafiqati, A., Clark, B. J., *Anal. Proc.* 1993, 30, 481.
- 3) Zhu, T., Sun, Y., Zhang, C., Ling, D., Sun, Z., *J. High Resolut. Chromatogr.* 1994, 17, 563–584.
- 4) Rose, D. J., Jorgenson, J. W., *J. Chromatogr.* 1988, 430, 23–34.
- 5) Brody, J. R., Kern, S. E., *Anal. Biochem.* 2004, 333, 1–13.
- 6) Kanoatov, M., Galievsky, V. A., Krivova, S. M., Chernov, L. T., Jankowski, H. K., Krylov, S. N., *Anal. Chem.* 2015, 87, 3099–3106.
- 7) Chen, H., Zhu, Z., Lu, J. J., Liu, S., *Anal. Chem.* 2015, 87, 1518–1522.
- 8) Liu, D. S., Zhao, S. J., *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 2015, 26, 978–979.
- 9) Ray, T., Mills, A., Dyson, P., *Electrophoresis* 1995, 16, 884–884.
- 10) Ray, T., Weedan, J., Dyson, P., *FEMS Microbiol. Lett.* 1992, 96, 247–252.
- 11) Berezovski, M., Mischenko, M., Drabovich, A., Krylov, S. N., *J. Am. Chem. Soc.* 2006, 128, 1410–1411.
- 12) Pariny, L., Borel, A., *J. Chrom. Inform. Mass.* 2013, 53, 1223–1228.

- [13] Laitinen, H. A., Enke, C. G., *J. Electrochem. Soc.* 1960, *107*, 773–781.
- [14] Gattrell, M., Kirk, D. W., *J. Electrochem. Soc.* 1993, *140*, 1534–1540.
- [15] Yanagisawa, Y., Kashiwagi, Y., Kurashima, F., Anzal, J., Osa, T., Bobbitt, J. M., *Chem. Lett.* 1996, 1043–1044.
- [16] Shono, T., *Curr. Chem.* 1988, *148*, 131–151.
- [17] Camalet, J. L., Lacroix, J. C., Dung Nguyen, T., Aeliyach, S., Pham, M. C., Petitjean, J., Lacaze, P. C., *J. Electroanal. Chem.* 2000, *485*, 13–20.
- [18] Ferreira, C. A., Aeliyach, S., Aaron, J. J., Lacaze, P. C., *Electrochim. Acta.* 1996, *41*, 1801–1809.
- [19] Meneguzzi, A., Pham, M. C., Lacroix, J. C., Piro, B., Adenier, A., Ferreira, C. A., Lacaze, P. C., *J. Electrochem. Soc.* 2001, *148*, B121–B126.
- [20] Saveant, J. M., *Acc. Chem. Res.* 1993, *26*, 455–461.
- [21] Sperry, J. B., Wright, D. L., *Chem. Soc. Rev.* 2006, *35*, 605–621.
- [22] Pratviel, G., in: Sigel, A., Sigel, H., Sigel, R. K. O. (Eds.), *Oxidative DNA Damage Mediated by Transition Metal Ions and Their Complexes*, Life Sciences, Springer Netherlands, Dordrecht 2012 pp. 201–216.
- [23] Kanoatov, M., Krylov, S. N., *Anal. Chem.* 2016, *88*, 7421–7428.

2 MESURE DES PARAMETRES DE LIAISONS ENTRE LA THROMBINE ET UN APTAMERE PAR ELECTROPHORESE CAPILLAIRE :

Dans le cadre de cette étude, le système utilisé était une CE Agilent 7100 avec un détecteur DAD et un détecteur de fluorescence Picometrics avec un laser 488 nm ou LED 480 nm.

Les conditions opératoires étaient les suivantes : Longueur du capillaire = 90cm
Diamètre interne : 50 μ m

Les conditions de séparations consistaient en :

➤ Lavages :

1) NaOH – 3 min

2) Eau – 3 min

3) Tampon Phosphate de sodium dibasique 50mM pH = 8.2 – 3 min

➤ Injection : 50 mBAR, 30 secondes

➤ Séparation : +/- 30 kV

Les échantillons (thrombine et aptamère T29) ont été dilués dans le tampon suivant composé de Phosphate de sodium dibasique 50 mM pH=8.2 + NaCl=100 mM + MgCl₂=1 mM.

Avant d'être injecté en électrophorèse capillaire, l'aptamère a été chauffé à 95°C pendant 5 min puis refroidi à température ambiante. 15 μ L de cet échantillon ont été mélangés avec 15 μ L de thrombine, puis injectés en CE.

2.1 Caractérisation de la thrombine par électrophorèse capillaire

La thrombine est une sérine-protéase de poids moléculaire de 36 700 Da, et de point isoélectrique (pI) compris entre 7,0 et 7,6. Elle est constituée de deux chaînes de respectivement 6 000 Da (36 acides aminés) et 31000 Da (258 acides aminés) reliées par un pont disulfure. La thrombine possède différentes conformations dont la forme α correspondant à la forme active de la protéine mais pouvant conduire aux formes β et γ thrombine qui sont des produits dégradés. Dans un premier temps, nous avons caractérisé la détection de cette protéine par détection UV.

2.1.1 Capillaire en silice

La thrombine est une molécule amphiphile qui contient des charges positives pouvant créer des interactions électrostatiques en interagissant avec les groupements silanols du capillaire. Ceci aboutit à l'adsorption de cette dernière à la surface du capillaire, ce qui peut devenir un obstacle majeur pour l'étude du complexe [Aptamère.cible] par électrophorèse capillaire. Nous avons donc décidé de travailler à un pH= 8.2, supérieur au pI de la thrombine afin que cette dernière et l'aptamère soient tous les deux chargés négativement.

Cependant, après plusieurs injections successives de thrombine en CE et malgré une série de lavages avec de l'eau, le pic de la thrombine persistait et ne disparaissait pas sur les électrophérogrammes de blanc. Donc, pour prévenir l'adsorption de la protéine sur le capillaire qui aurait pu compromettre l'étude de l'interaction protéine-aptamère et notamment la détermination de la constante de dissociation K_D , nous avons décidé d'utiliser un capillaire

recouvert d'un revêtement en PVA (polyvinylalcool) qui permet de masquer les charges négatives à la surface de ce capillaire.

2.1.2 Capillaire recouvert d'un revêtement en PVA (Polyvinylalcool)

L'utilisation d'un capillaire en PVA va masquer les groupements silanols du capillaire et donc quasiment supprimer le flux électroosmotique. Ainsi seule la mobilité électrophorétique est impliquée dans le processus de migration. La thrombine, étant globalement négative à pH = 8.2, nous avons donc inversé la polarité du champ électrique et appliqué un voltage négatif de -30 kV (**Figure 2-1**).

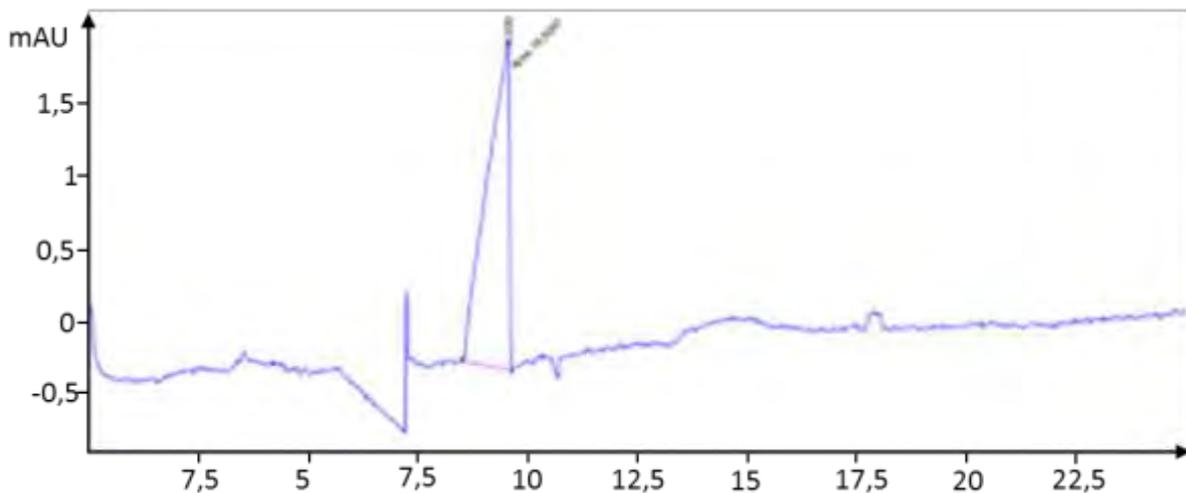


Figure 2-1 : Détection de la thrombine à 1 μ M en absorbance UV à 210nm. $T_m=9,6$ min. Capillaire en PVA : diamètre = 50 μ m. Longueur totale L = 90cm ; longueur effective (l=21 cm) ; Tampon de séparation : Phosphate de sodium dibasique 50 mM pH = 8.2. Injection : 50 mBAR, 30s. Séparation : -30kV.

Que ce soit avec un capillaire en silice ou avec un capillaire en PVA, nous observons une forme de pic légèrement asymétrique. Ceci peut s'expliquer par le fait que la mobilité de la thrombine est très différente de la mobilité du tampon phosphate de sodium dibasique. Pour un échantillon de faible conductivité, l'analyte migrera dans la région du tampon moins rapidement car le champ électrique est plus faible (**Figure 2-2**).

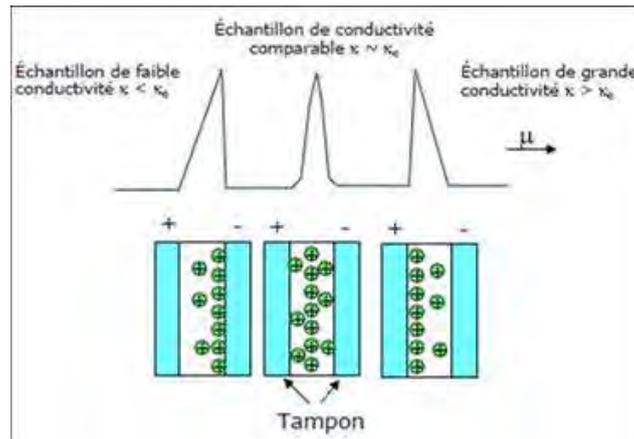


Figure 2-2 : Principe de l'électro dispersion et différence de conductivité. La différence de conductivité entre la zone de l'analyte et celle du tampon est un facteur pouvant entraîner un profil de pic différent. Si l'échantillon à analyser a une faible conductivité, alors il migrera moins facilement dans le tampon car le champ électrique est plus faible. Inversement, pour un échantillon avec une forte conductivité, l'analyte entrera plus facilement dans la région du tampon car il y aura un fort champ électrique.

2.2 Caractérisation de l'aptamère T29 et du complexe T29.Thrombine par électrophorèse capillaire

2.2.1 Avec un capillaire en PVA

L'aptamère T29 est greffé à un fluorophore en position 5' de ce dernier. Le fluorophore utilisé est le 6-carboxyfluorescéine (6-FAM).



Ainsi, grâce au greffage d'un marqueur fluorescent, on peut détecter l'aptamère T29 (**Figure 2-3**), mais également le complexe [T29. Thrombine] lorsque l'on injecte le mélange T29.thrombine. Dans un premier temps, nous avons caractérisé la détection du T29, en utilisant un capillaire en PVA.

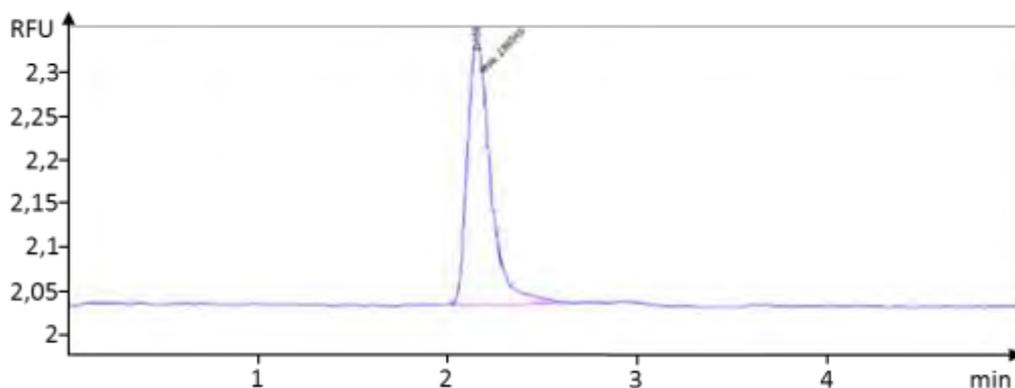


Figure 2-3 : Détection de l'aptamère T29 à 50 nM par fluorescence à 488 nm. Capillaire en PVA - diamètre : 50 μm . Longueur totale $L=90$ cm (longueur effective $l=21$ cm). $T_m=2.2$ min. Tampon de séparation : Phosphate de sodium dibasique pH= 8.2. Injection : 50 mBAR, 30 s. Séparation : -30 kV

Dans ces conditions, lorsque nous avons voulu réaliser le complexe T29.thrombine, nous n'avons pas observé le pic de complexe. L'analyse du complexe T29 50 nM + Thrombine 500 nM, nous a donné un électrophérogramme (**Figure 2-4**), pour lequel nous observons une diminution de la fluorescence du pic d'aptamère libre ($H_{T29 \text{ complexé}} = 0.05 \text{ RFU}$ et $H_{T29 \text{ libre}} = 0.2 \text{ RFU}$)

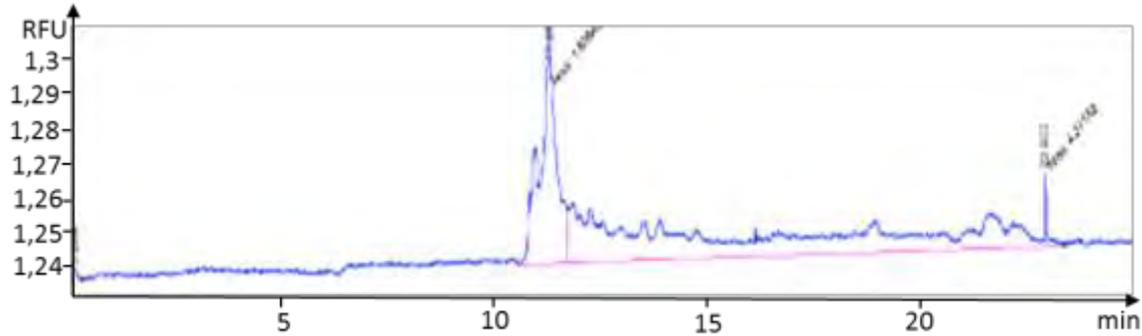
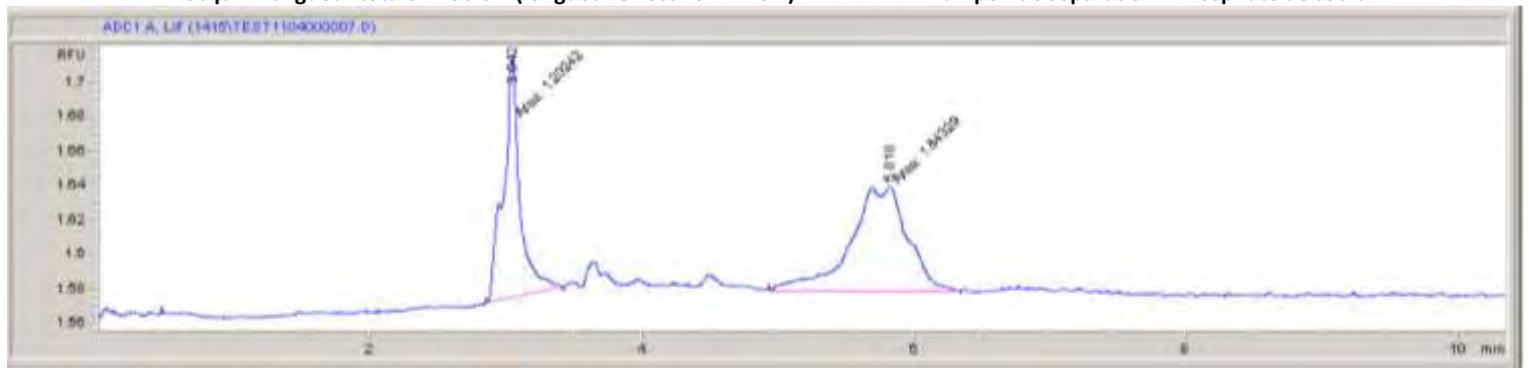


Figure 2-4: Analyse du complexe T29 50 nM + thrombine 500 nM par fluorescence à 488 nm. Capillaire en PVA - diamètre : 50 μm . Longueur totale L= 90 cm (longueur effective l=21 cm). $T_m=12$ min. Tampon de séparation : Phosphate de sodium



dibasique 50 mM pH= 8.2. Injection : 50 mBAR, 30s. Séparation : -30 kV

Cette observation peut traduire le fait que le complexe injecté serait instable au cours de la séparation dans le capillaire, et donc le complexe [T29. Thrombine] se dissocierait au cours de la migration. Cela aboutirait à une mauvaise séparation et la traînée observée (**Figure 2-4**).

Pour vérifier cette hypothèse, nous avons décidé d'inverser le sens d'injection du complexe [T29 + thrombine] afin de raccourcir la longueur du trajet de séparation dans le capillaire et donc le temps de migration du complexe (**Figure 2-5**). Pour pouvoir réaliser cette méthode, il faut veiller à ce que l'électrode côté injection soit au pôle négatif.

Figure 2-5: Analyse du complexe T29 50 nM + thrombine 500 nM par fluorescence à 488 nm. Capillaire en PVA - diamètre : 50 μm . Longueur totale L= 90 cm (longueur effective l=19 cm). $T_{mT29}=2.2$ min. $T_{m\text{complexe}}= 5.8$ min. Tampon de séparation : Phosphate de sodium dibasique 50 mM pH= 8.2. Injection à l'outlet : 50 mBAR, 30 s. Séparation : +30 kV

Avec un capillaire en PVA, le revêtement diminue de façon importante le flux électroosmotique et ce sont les molécules qui ont le rapport (charge/ taille) le plus élevé qui migrent en premier. Donc c'est l'aptamère sous forme non complexé qui migre en premier. En inversant le sens d'injection, le complexe parcourt une distance plus courte avant d'atteindre la fenêtre de détection LIF. Alors qu'il parcourait l= 71 cm et avait un temps de migration de 12 min, en étant injecté à l'outlet, la distance est réduite (l= 19 cm) et le temps de migration de l'aptamère

libre correspond à celui prévu par le calcul pour une distance effective quatre fois plus courte ($t = 3$ min). En ce qui concerne le complexe, on observe toujours une dissociation de ce dernier car il n'y a pas de retour à la ligne de base entre le pic d'aptamère libre et celui du complexe. Ainsi pour amoindrir l'effet de la dissociation, nous avons mesuré l'effet de la pression sur la taille du pic du complexe (**Figure 2-6**). Pour cela, nous avons appliqué une dépression en plus de la séparation par voltage. Lorsque nous appliquons une dépression de 50 mBar, nous avons une meilleure séparation entre le T29 libre et le complexe. Cependant, bien que la détection et la séparation soient optimales avec une dépression de 50 mBAR, cela pourrait signifier dans le cas de différents aptamères injectés que ceux qui sont détectés à 50 mBAR ont une moins bonne affinité que ceux détectés avec une dépression de 10 mBAR.

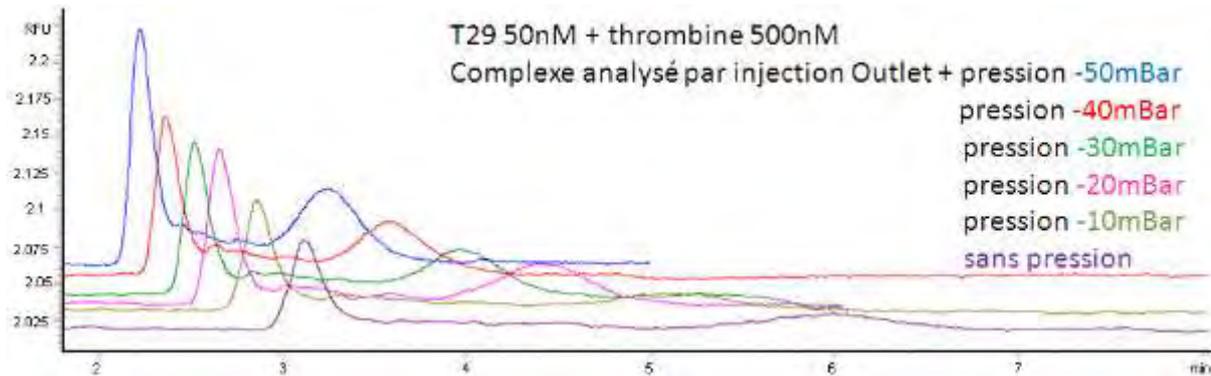


Figure 2-6 : Détection du complexe aptamère T29 50 nM + thrombine 500 nM par fluorescence à 488 nm. Capillaire en PVA - diamètre : 50 μ m. Longueur totale L = 90 cm (longueur effective l=21 cm). Tampon de séparation : Phosphate de sodium dibasique 50 mM pH= 8.2. Injection à l'OUTLET : 50 mBAR, 30 s. Séparation : +30 kV + pression.

2.2.2 Avec un capillaire en silice

Nous avons décidé de changer de diamètre de capillaire afin d'injecter une plus grande quantité d'ADN. D'après l'article de M.Toulmé, je cite « on peut penser que dans une population présente en quantité importante, se trouvera une séquence adaptée à la reconnaissance de la cible choisie. La coexistence, dans la même entité moléculaire, de la forme et de la séquence permet, par alternance d'une étape de sélection (liée à la forme) et d'amplification (liée à la séquence), d'extraire de la famille, à l'issue d'un processus d'évolution moléculaire, le(s) candidat(s) doué(s) des propriétés attendues » (149). Mais n'ayant pas de capillaire greffés PVA de 100 μ m de diamètre, nous avons été obligés d'utiliser un capillaire en silice nue de diamètre 100 μ m. Dans un premier temps, les conditions de séparation ont été optimisées en utilisant un capillaire de diamètre 50 μ m (**Figure 2-7**).

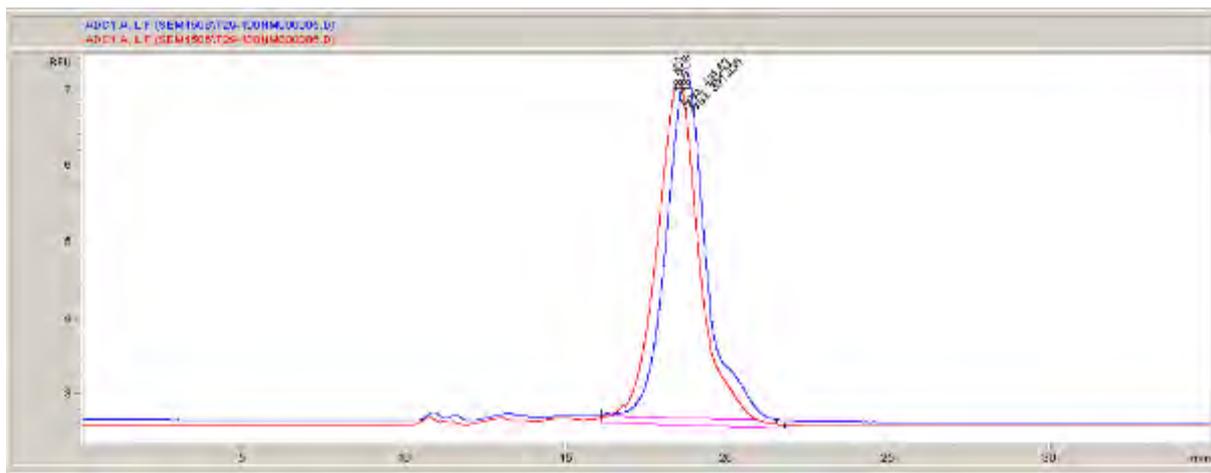


Figure 2-7 : Détection de l'aptamère T29 à 100 nM par fluorescence à 488 nm. Capillaire en silice - diamètre : 50 μ m. Longueur totale L= 90 cm (longueur effective l=21 cm). T_m=18 min. Tampon de séparation : Phosphate de sodium dibasique 50 mM pH= 8.2. Injection : 50 mBAR, 30 s. Séparation : + 20 kV, + pression 5 mBAR.

Les conditions de séparation + 20 kV, + pression 5 mBAR, sont différentes par rapport aux conditions utilisées pour injecter la thrombine et l'aptamère T29 (+ 30 kV) avec différentes pressions. Avec un capillaire de diamètre 100 μ m, l'échauffement de ce dernier si l'on applique un voltage de + 30kV serait trop important ($I > 80\mu$ A), ce qui pourrait nuire à la répétabilité de la séparation.

Nous avons ensuite réalisé l'interaction entre l'aptamère T29 à 50 nM et la thrombine à 100 nM (**Figure 2-8**). Avec un capillaire en silice, nous modifions la nature du capillaire et la manière d'injecter et ce sont les molécules qui ont le rapport (masse/ charge) le plus faible qui migrent en premier. Donc le complexe migre en premier suivi de l'aptamère sous forme libre.

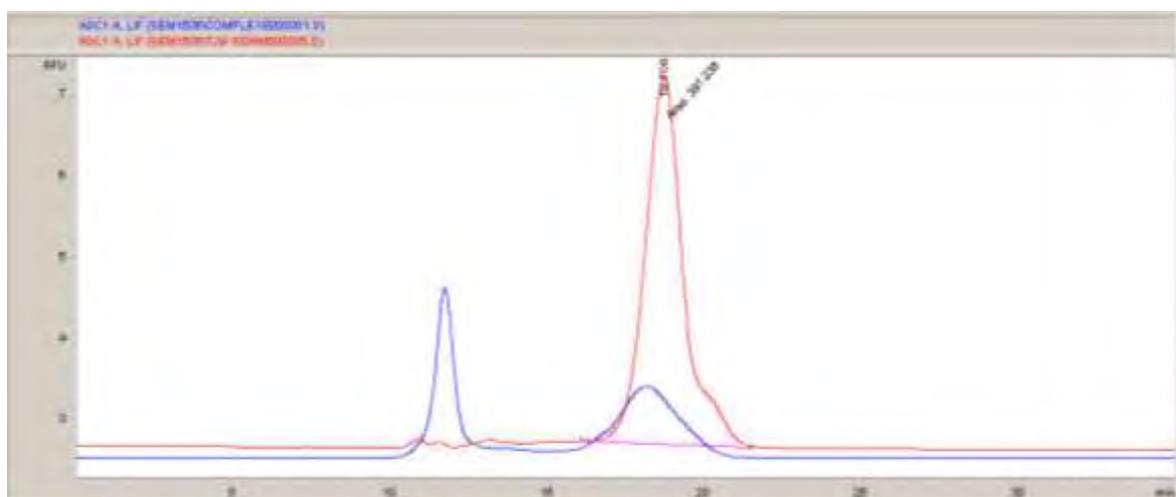


Figure 2-8 : Analyse de l'aptamère T29 100 nM (électrophérogramme en rouge) et du complexe T29 50 nM + thrombine 100 nM (électrophérogramme bleu). L'aptamère a un temps de migration de 18.1 min et le complexe migre à t= 11.7 min. Capillaire en silice - diamètre : 50 μ m. Longueur totale L= 90 cm (longueur effective l= 21 cm). T_m= 18 min. Tampon de séparation : Phosphate de sodium dibasique 50 mM pH= 8.2. Injection : 50 mBAR, 30 s. Séparation : + 20 kV, + pression 5 mBAR.

Nous avons également mesuré l'effet de la pression sur la séparation du complexe avec un capillaire en silice. Pour cela, nous avons injecté le complexe [T29. Thrombine] à différentes pressions (**Figure 2-9**).

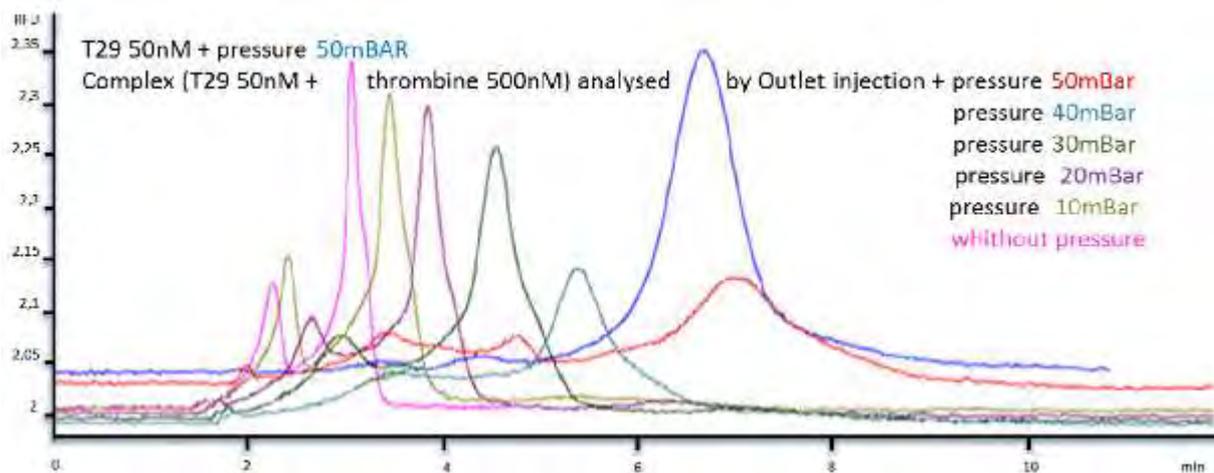


Figure 2-1 : Analyse de l'aptamère T29 100 nM (électrophérogramme en rouge) et du complexe T29 50 nM + thrombine 100 nM (électrophérogramme bleu). L'aptamère a un temps de migration de 18.1 min et le complexe migre à $t = 11.7$ min. Capillaire en silice - diamètre : 50 μm . Longueur totale $L = 90$ cm (longueur effective $l = 21$ cm). $T_m = 18$ min. Tampon de séparation : Phosphate de sodium dibasique 50 mM pH=8.2. Injection : 50 mBAR, 30 s. Séparation : + 20 kV, + pression 5 mBAR

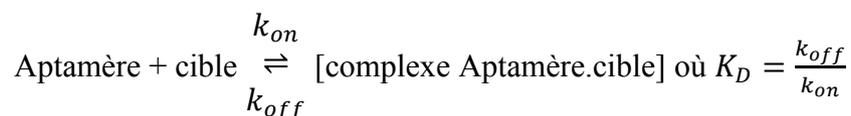
T29.thrombine. Sans pression, les aptamères peuvent être choisis avec un k_{off} plus lent avec une meilleure séparation entre le complexe et le T29 libre. En revanche, plus la pression augmente, moins la séparation entre le complexe et le T29 libre est atteinte. Cependant, là aussi, le signal du complexe augmente avec une pression plus élevée et conduit ainsi à une détection plus élevée de ce dernier.

2.3 Détermination de la constante de dissociation du complexe T29. Thrombine

Ce travail a été réalisé par de nombreux auteurs.

2.3.1 Par électrophorèse capillaire

On rappelle que l'équilibre d'un aptamère avec sa cible est régi selon l'équation suivante :



Nous avons réalisé une gamme étalon de thrombine allant de 0 à 25nM et en utilisant pour l'aptamère T29, une concentration constante de 25 nM. Nous avons choisi cette concentration en nous basant sur les travaux de S. Krylov, qui préconise lorsque la valeur de K_D n'est pas connue précisément, de maintenir la valeur de concentration de l'aptamère la plus faible possible afin que les aires mesurées du complexe [T29.thrombine] soient les plus fiables.

Le capillaire utilisé était un capillaire de silice, de longueur totale $L = 90$ cm (longueur effective $l = 19$ cm). Le tampon de migration utilisé était le tampon Na_2HPO_4 à une concentration de 50 mM, pH= 8.2. Le tampon de dilution des échantillons était un tampon composé de Na_2HPO_4 à 50 mM + NaCl 100 mM + MgCl_2 1 mM. Le conditionnement du capillaire était le suivant : Lavage 3 min NaOH 1M / 3 min eau / 3 min tampon de migration. Injection : 50 mBar, 30 secondes. Séparation : + 15 kV, + pression 5 mBar. A partir d'une solution mère de T29 à 100 μM , ce dernier a été dilué à 25 nM. La thrombine a été diluée à partir d'une solution mère à 20 μM . Pour réaliser l'interaction avec la thrombine, il faut chauffer l'échantillon d'aptamère à 95°C pendant 5 min, puis le laisser revenir à température ambiante. Ensuite, on incube 10 μL

de thrombine avec 10 μ L d'aptamère et on injecte le mélange par électrophorèse capillaire (Figure 2-10).

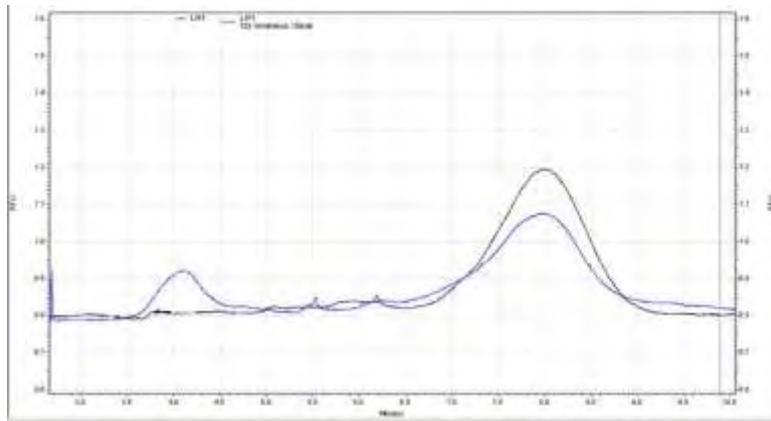


Figure 2--10: Détermination du temps de migration de l'aptamère T29 25 nM $t = 8.5$ min (égram noir) pour la mesure de K_D par détection LIF. L'électrophérogramme bleu montre l'interaction entre l'aptamère à 25 nM et la thrombine à 150 nM. On observe une diminution du pic de l'aptamère libre et l'apparition d'un pic à $t = 4$ min correspondant au complexe [T29. Thrombine].

Pour déterminer la constante de dissociation K_D , nous avons mesuré l'aire du pic correspondant au complexe pour chaque concentration de thrombine (Figure 2-13). La courbe Aire du pic de complexe = f [thrombine] a ensuite été tracée (Figure II-12).

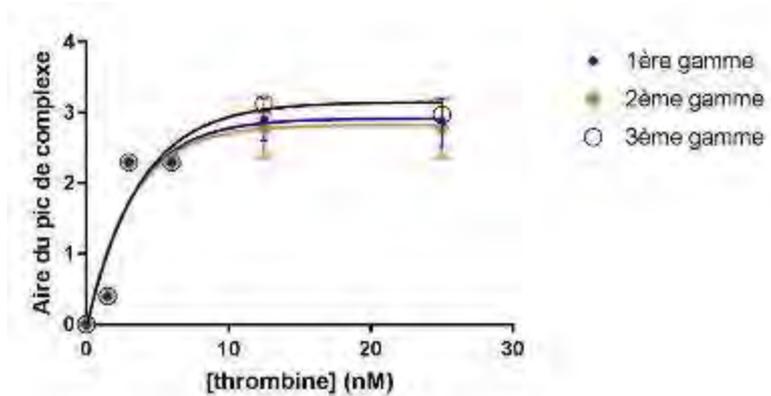


Figure 2-2 : Aire du pic de complexe = f [thrombine] (nM)

[Thrombine] (nM)	Aire du pic [T29.thrombine]	Aire du pic [T29.thrombine]	Aire du pic [T29.thrombine]
0	0	0	0
1.5	0.4	0.4	0.4
3	2.9	2.5	2.3
6	2.2	2.3	2.6
12.5	2.6	2.6	2.6
25	3	3.2	3

Figure II-3 : Valeur des aires obtenues pour le complexe [T29. Thrombine] en réalisant un complexe de T29 à 25 nM + différentes concentrations de thrombine à 0 - 1,5 - 3 - 6 - 12,5 et 25 nM. L'aire de chaque pic de complexe est déterminée grâce au logiciel Chemstation.

Ainsi la valeur obtenue pour la constante de dissociation est $K_D = 2.62 \pm 0.83 \text{ nM}$. Cette valeur est dans le champ des valeurs trouvées dans la littérature qui varient entre 0.5nM et 200nM selon la méthode utilisée.

2.3.2 Par filter binding

La valeur de K_D a également été déterminée par filter binding afin de pouvoir comparer les valeurs obtenues par les deux méthodes (**Figure 2-14**).

Le tampon utilisé était composé de : Tris HCl 50 mM, pH= 7.5 + NaCl 100 mM + MgCl₂ 1 mM. Le tampon a ensuite été filtré avec un filtre de diamètre 0.2 µm. Pour réaliser l'interaction entre l'aptamère et la thrombine, l'aptamère marqué au P³² a été chauffé à 95°C pendant 5 minutes, puis la température est abaissée jusqu'à température ambiante pendant 40 min. Ce mélange est ensuite passé à travers un filtre de nitrocellulose. L'ADN libre passe à travers le filtre de nitrocellulose et sera retenu par le filtre de nylon (**Figure 2-15**). Le complexe aptamère/thrombine sera retenu à la surface du filtre de nitrocellulose.

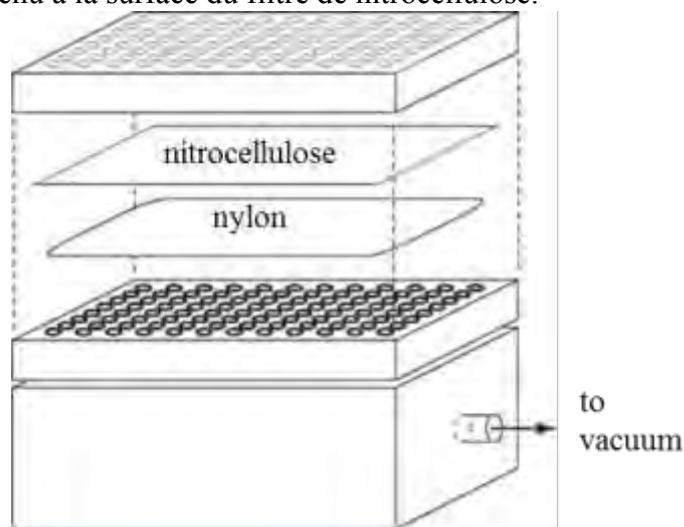


Figure 2-4 : Schéma de principe du filter binding avec le filtre en nitrocellulose qui retient la thrombine et le filtre en nylon qui permet de piéger tout l'ADN qui n'est pas retenu par la membrane.

Pour quantifier le nombre de molécules liées à la thrombine, le mélange est ensuite lavé, séché et les filtres sont exposés au phosphoImager (Storm 840) (**Figure II-15**).

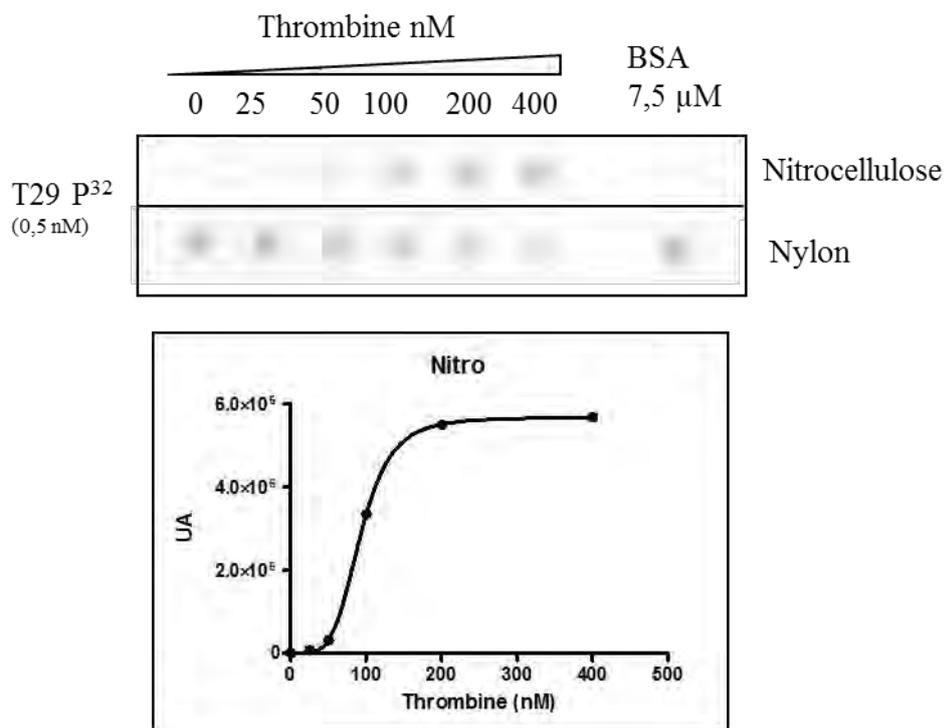


Figure 2-5 : Analyse par filter binding de l'aptamère T29. Le filtre en nitrocellulose va permettre de retenir les protéines, tandis que les molécules d'ADN non liées vont passer à travers la nitrocellulose et seront retenus sur le filtre de nylon. Un lavage est ensuite fait avec une solution de KOH afin d'éliminer les molécules d'ADN liées à la cible qui auraient pu rester sur la membrane de nitrocellulose.

La valeur de la constante de dissociation obtenue est de 100nM.

2.3.3 Conclusion

Les valeurs obtenues par ces deux méthodes sont très différentes, bien qu'elles soient obtenues avec la même thrombine. L'analyse par électrophorèse capillaire s'effectue en solution donc la surface de liaison de l'aptamère T29 à la cible est maximale contrairement à la technique de filter binding où l'on fixe l'aptamère sur une surface. On peut donc supposer que ce paramètre peut avoir une influence sur l'interaction entre la thrombine et le T29. Le point important à noter est que les compositions de chaque tampon sont aussi différentes : tampon sodium de phosphate pour l'électrophorèse capillaire et tampon Tris HCl pour le filter binding mais chaque tampon a une concentration de 50 mM, avec un pH= 8.2. On peut supposer que l'utilisation du Tris HCl par rapport au tampon sodium de phosphate augmente la concentration en sels et peut favoriser une interaction défavorable entre la thrombine et l'aptamère T29 et donc une valeur de la constante de dissociation plus élevée.

Chapitre 2 : Est-on capable d'identifier un aptamère G-quartet en CE- Illumina ?

Chapitre 2 : Est-on capable d'identifier un aptamère G-quartet en CE- Illumina ?

1 VALIDATION / MISE EN PLACE DE LA COLLECTE A PARTIR DE L'EXEMPLE DE 3 SUCRES

La méthode de collecte est une étape importante dans le processus de sélection d'aptamères car c'est durant cette étape que l'on va pouvoir déterminer les paramètres qui vont permettre de collecter les fractions des complexes ADN-cibles avec une bonne résolution et une bonne séparation. Les quantités collectées sont faibles (quelques nanogrammes), il faut donc pouvoir collecter une quantité suffisante afin que si l'on réinjecte le complexe, ce dernier ne soit pas trop dilué et par conséquent indétectable par fluorescence. Colyer et son équipe (150), ont décrit une méthode de collecte afin de limiter les dilutions de la zone collectée. Pour cela, ils ont regardé l'influence de la quantité de tampon contenu dans le vial de collecte (5, 10 ou 15 μ L). Avec 5 μ L de tampon contenu dans le flacon de collecte, ils ont trouvé que la sensibilité de détection était meilleure mais la répétabilité de collecte est moins efficace. Avec 15 μ L de tampon contenu dans le flacon, la répétabilité de l'expérience est légèrement meilleure mais avec une sensibilité plus faible. Ils en ont déduit qu'il fallait travailler avec un volume de tampon contenu dans le vial de collecte de 10 μ L et ceci en utilisant un capillaire de diamètre 75 μ m.

1.1 Injection et séparation de trois sucres : exemple

Dans le but d'optimiser une méthode de séparation et de collecte des aptamères, nous avons décidé d'utiliser comme modèle trois sucres présentant des temps de migration distincts et reproductibles permettant de facilement les collecter et les réinjecter. Le mélange injecté était composé de glucose (DP 1), de maltose (DP 2) et de maltotriose (DP 3). Ces trois sucres étaient greffés à un marqueur : l'APTS (acide 8-amino-1, 3,6-pyrenetrisulfonique) afin de pouvoir être détectés par fluorescence à 488 nm. Dans un premier temps les temps de migration de chacun des trois composés ont été détectés (**Figure 3-1**).

La méthode d'analyse est la suivante :
LAVAGES : 3 min NaOH 1M
 3min HCl 0.1M
 3min Eau
 3min Tampon Na₂HPO₄

INJECTION : 30mbars, 10s

SEPARATION : +20kV

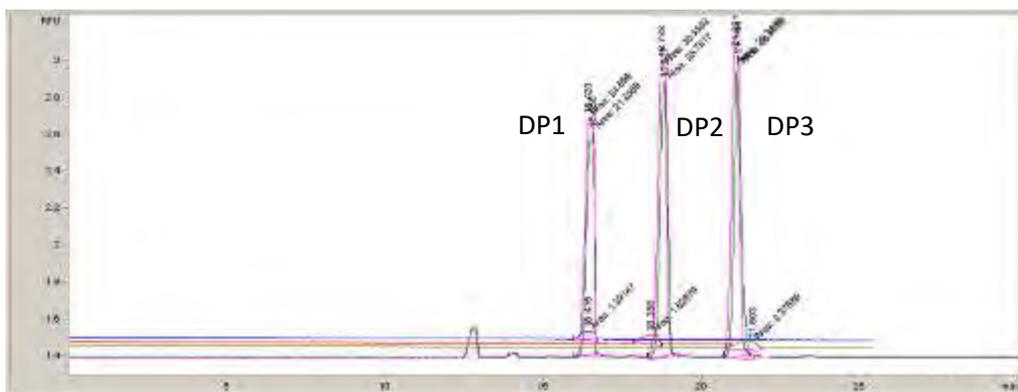


Figure 3-1 : Analyse de 3 sucres (DP1 : $t_m=16.4\text{min}$ - $A=24.9$), (DP2 : $t_m=18.7\text{min}$ - $A=30.6$), (DP3 : $t_m=21.1\text{min}$ - $A=21.1$). Injection : 30mBARS, 10s / Voltage : +20Kv

1.2 Méthode de collecte de trois sucres par électrophorèse capillaire :

Une fois les temps de migration de chaque composé déterminé, nous avons mis au point la méthode de collecte de chaque sucre (**Figure 3-2**). Pour cela, nous avons déterminé le moment à partir duquel le composé arriverait à la sortie du capillaire et dans le tube de collecte. Ainsi, nous avons réalisé des fenêtres de collecte aux temps de migration suivant :

DP1 $t_{collecte}$: [20.3;21.3]

DP2 $t_{collecte}$: [23.3;24.3]

DP3 $t_{collecte}$: [26.4;27.3]

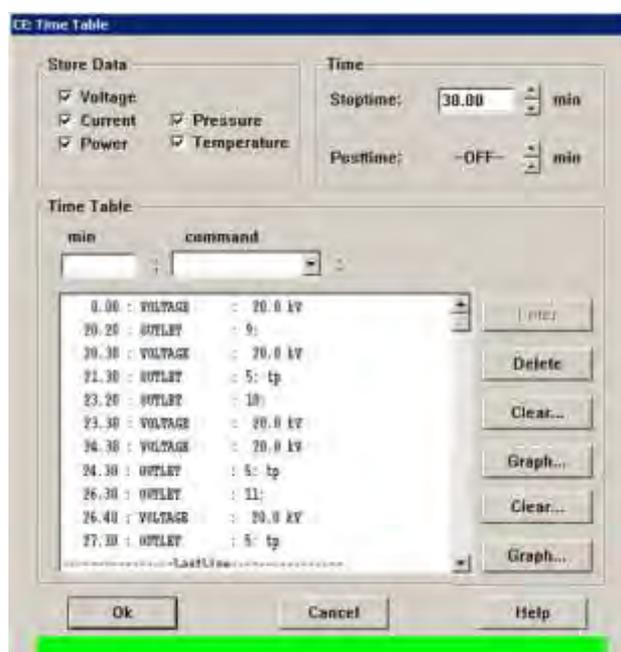


Figure 3-2 : Onglet time table sur l'instrument d'électrophorèse capillaire d'Agilent permettant de réaliser des collectes. Le vial 9 correspond au vial de collecte du sucre DP1 (désigné ci-dessus par outlet 9), le vial 10 correspond au vial de collecte

du sucre DP2 (désigné ci-dessus par outlet 10) tandis que le vial 11 correspond au vial de collecte du sucre DP3 (désigné ci-dessus par outlet 11).

Le tube de collecte contenait 15µL de tampon. C'est la quantité minimale que l'on peut mettre afin que l'électrode soit au contact du tampon. Une fois ces paramètres déterminés, nous avons réalisé 8 injections et collectes des trois sucres dans 3 tubes distincts. Ensuite, chaque sucre a été réinjecté séparément (**Figure 3-3**).

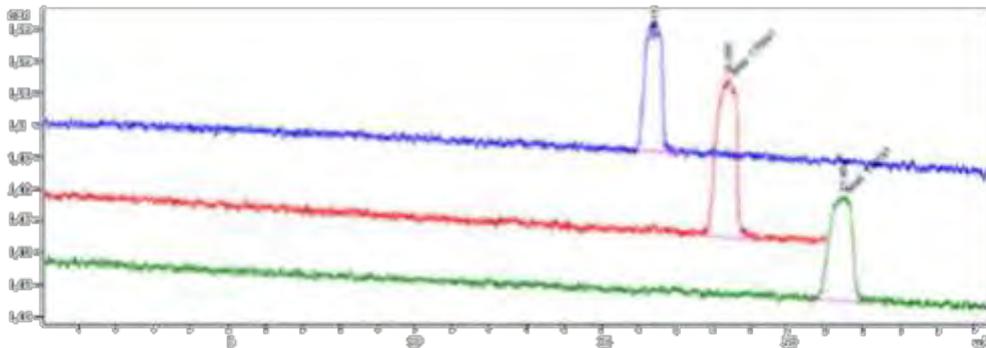


Figure 3-3 : Réinjection des 3 sucres après 8 collectes DP1 : $t_m = 16.4\text{min}$ - $A = 1.2$), (DP2 : $t_m = 18.4\text{min}$ - $A = 1.8$), (DP3 : $t_m = 21.5\text{min}$ - $A = 1.4$). Injection : 30 mBAR, 10s / voltage : + 20kV

1.3 Détermination du facteur de dilution et du rendement de collecte

Afin d'estimer le rendement des échantillons collectés, nous avons déterminés le facteur de dilution (DF) :

$$DF = \frac{nb \text{ de collectes} * v_i}{v_f} = \frac{8 * 100 * 10^{-9}}{15 * 10^{-6}} = 0.05$$

A partir de l'aire des trois sucres injectés (**Figure 3-1**), nous avons calculé l'aire théorique que nous devons obtenir après 8 collectes en tenant compte du facteur de dilution.

$$A_{1 \text{ théorique}} = DF * A_{i1} = 0.05 * 24.9 = 1.3$$

$$A_{2 \text{ théorique}} = DF * A_{i2} = 0.05 * 30.6 = 1.5$$

$$A_{3 \text{ théorique}} = DF * A_{i3} = 0.05 * 29.3 = 1.5$$

Si l'on compare les aires théoriques calculées avec celles obtenues après avoir injecté les trois sucres collectés ($A_1 = 1.2$, $A_2 = 1.8$ et $A_3 = 1.4$), celles-ci ont des valeurs similaires, ce qui signifie que le rendement de collecte est proche de 100%. La méthode de collecte des trois sucres a ainsi été validée.

2 COLLECTE D'APTAMERES :

2.1 Optimisation du diamètre du capillaire :

L'électrophorèse capillaire utilise des faibles quantités d'échantillons. Or, si l'on souhaite sélectionner des aptamères, il faut pouvoir injecter la plus grande quantité de molécules d'ADN

afin d'avoir la plus grande variété de séquences. Plus le volume de banque d'ADN injecté est important et plus la population d'aptamères sera enrichie. Donc nous avons décidé d'utiliser un capillaire en silice de diamètre 100 μ m pour réaliser les collectes d'aptamères. Le volume injecté avec ce capillaire est de 200 nL. Dans un premier temps, nous avons réalisé l'interaction entre l'aptamère T29 à 50nM et la thrombine à 100nM (**Figure 3-4**).

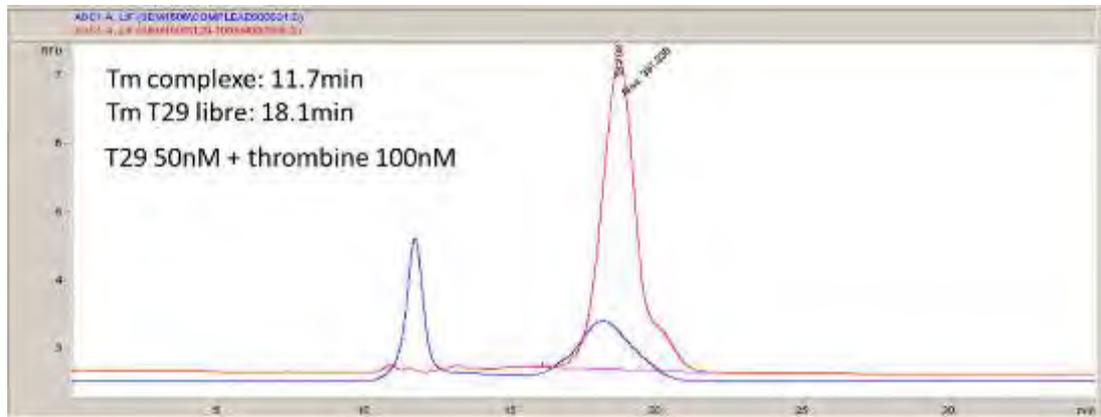


Figure 3-4: Un exemple d'injection de l'aptamère T29 50nM (électrophérogramme rouge) et du complexe aptamère T29 50nM + thrombine 100nM (électrophérogramme bleu).

Pour cette expérience, 50 μ L d'aptamère T29 à une concentration de 100nM a été incubé à 95 $^{\circ}$ C pendant 5 min, puis laisser revenir à température ambiante et mélangé avec 50 μ L de thrombine à 200nM.

$$v = 50\mu L \text{ de T29 à } 100nM \text{ correspondent à } n = 5 * 10^{-12} \text{ mol}$$

$$v = 50\mu L \text{ de Thrombine à } 200nM \text{ correspondent à } n = 2.5 * 10^{-12} \text{ mol}$$

Avec un capillaire de diamètre 100 μ m, le volume d'injection est de 200nL, donc le nombre de molécules injectées est de $5 * 10^{-12} * 6.02 * 10^{23} = 2 * 10^{10}$ molécules de T29 et $2.5 * 10^{-12} * 6.02 * 10^{23} = 2 * 10^{10}$ molécules de thrombine. Tandis qu'en utilisant un capillaire de 50 μ m de diamètre, le volume injecté serait donc d'environ 100nL donc nous injecterions 2 fois moins de molécules dans le capillaire.

2.2 Injection de la banque d'ADN, de l'aptamère AT29B et de la thrombine :

2.2.1 La banque d'ADN (N=30) :

Nous avons décidé d'utiliser une banque d'ADN (L3 : Library 3) dégénérée sur 30 nucléotides (N = 30), marquée au FAM en position 5'.



La banque d'ADN a été injectée à une concentration de 10 μ M, ce qui correspond à 10^{12} molécules. A cette concentration-là, le pic de banque d'ADN sature et la largeur du pic est importante (**Figure 3-5**). Mais comme le but est d'injecter le plus grand nombre de molécules, nous avons conservé cette concentration de 10 μ M et la banque d'ADN a un temps de migration de 18 min par détection LIF en électrophorèse capillaire.

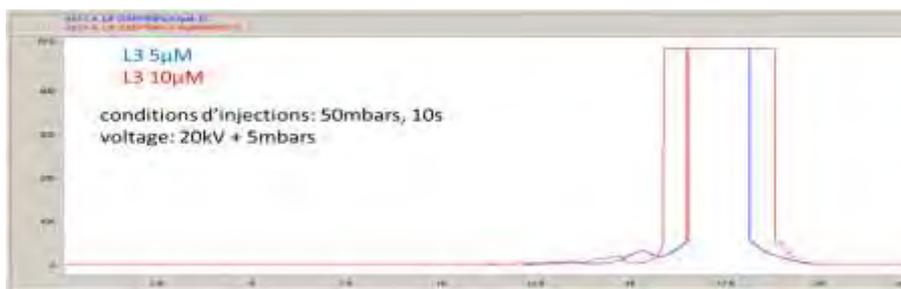


Figure 3-5 : Injection de la banque dégénérée L3 10 μ M. Sens de l'injection : Inlet 50mbars, 10s Voltage : +20kV +5mbars

2.2.2 Injection de l'aptamère AT29B :

Nous avons décidé d'injecter l'aptamère T29 mais avec les deux séquences qui permettent de réaliser l'amplification par PCR (70 paires de bases au lieu de 29 paires de bases) et que l'on appelle aptamère AT29B. Nous avons vérifié que le temps de migration de cet aptamère était identique à celui de l'aptamère T29 de 29 paires de bases. Pour cela, nous avons injecté l'aptamère AT29B à 100nM (Figure 3-6) :



Figure 3-6 : Injection de l'aptamère AT29B à 100nM

L'aptamère AT29B a le même temps de migration que l'aptamère T29, $t_m = 18$ min, mais surtout les temps de migration sont similaires à celui de la banque d'ADN, donc on pourra plus facilement déterminer la fin de la fenêtre de collecte.

2.2.3 Injection de la thrombine :

La thrombine est détectée en UV à 210 nm (Figure 3-7). Comme la sensibilité en détection UV est moins précise qu'en LIF, nous l'avons injecté à de fortes concentrations, 5 et 10 μ M afin de calculer son temps de migration à la fenêtre de détection de fluorescence afin de déterminer le début de la fenêtre de collecte des aptamères.

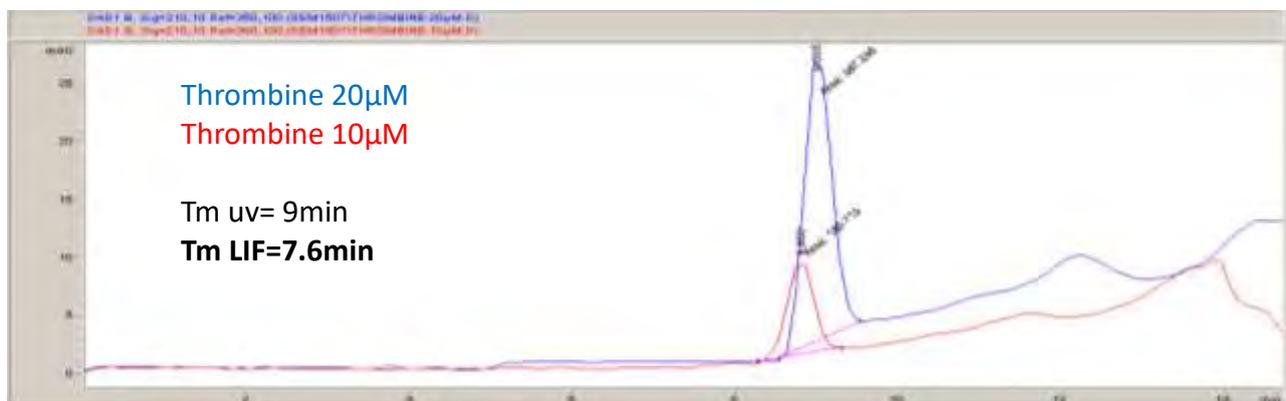


Figure 3-7 : Détection par absorbance UV à 210nm de la thrombine à 10 et 20 μ M par électrophorèse capillaire.

2.2.4 Réalisation du complexe Banque d'ADN + thrombine + AT29B et détermination des fenêtres de collectes :

Nous avons réalisé un mélange de thrombine à 500nM + banque d'ADN à 1 μ M + AT29B à 1 μ M. Le mélange a été préparé de la façon suivante : nous avons fait chauffer la banque d'ADN et l'aptamère AT29B dans deux Eppendorf distincts à 95°C pendant 5 minutes, puis laissé revenir à température ambiante. Ensuite, on prélève 10 μ L de chaque réactif que l'on ajoute avec 10 μ L de thrombine et le mélange est injecté par électrophorèse capillaire (**Figure 3-8**).

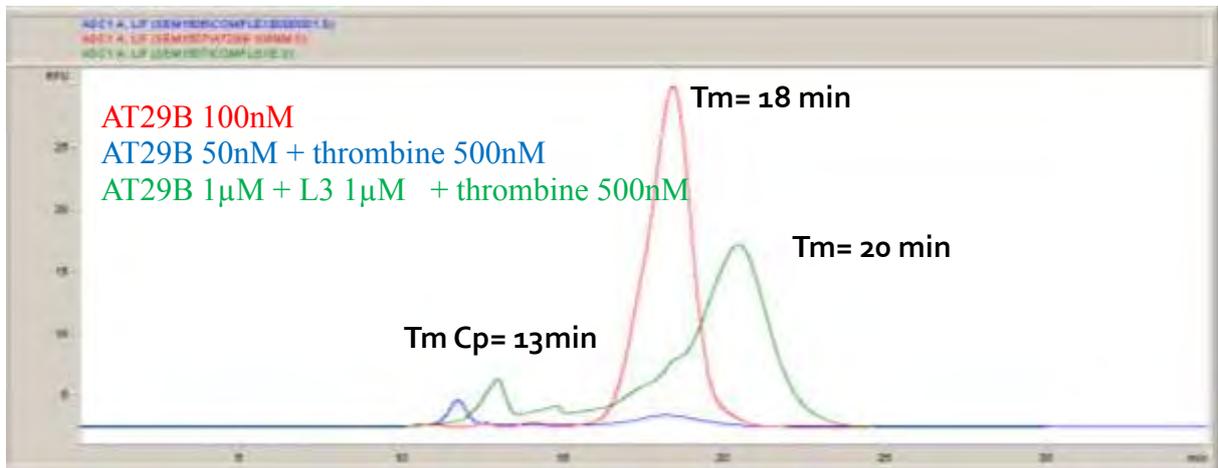


Figure 3-8 : Electrophérogramme de l'aptamère AT29B à 100nM (courbe rouge), du mélange AT29B à 50 nM et de la thrombine à 500 nM (courbe bleue) et du mélange AT29B à 1 μ M + banque d'ADN L3=10 μ M + thrombine à 500 nM (courbe verte).

Nous avons aussi réalisé un mélange composé de l'aptamère AT29B et de thrombine afin d'être sûr que les temps de migration étaient identiques à ceux de la banque qui était ajoutée en plus dans le complexe. Le temps de migration de la banque 20 min et de l'aptamère libre 18 min est de 20 min, tandis que l'on voit le complexe apparaître vers 13 min. Nous remarquons que les temps de migration sont décalés, mais nous distinguons le complexe du mélange banque d'ADN et aptamère T29. Une fois les temps de migration du complexe et de la banque + AT29B déterminés, nous avons mis au point un procédé de collecte

Nous avons ensuite utilisé la méthode d'analyse suivante :

LAVAGES : 3 min NaOH 1M
3min HCl 0.1M
3min Eau
3min Tampon Na₂HPO₄

INJECTION : 50mbars, 30s

SEPARATION : +15kV + Pression 5mbars

0.00	: VOLTAGE	: 15.0 kV
0.00	: PRESSURE	: 5 mbar
7.50	: VOLTAGE	: 15.0 kV
7.50	: PRESSURE	: 5 mbar
7.80	: VOLTAGE	: 15.0 kV
7.80	: PRESSURE	: 5 mbar
7.80	: OUTLET	: 9:
9.80	: VOLTAGE	: 15.0 kV
9.80	: PRESSURE	: 5 mbar
9.80	: OUTLET	: 6: tp
12.60	: VOLTAGE	: 15.0 kV
12.60	: PRESSURE	: 5 mbar
12.60	: OUTLET	: 10:
16.90	: VOLTAGE	: 15.0 kV
16.90	: PRESSURE	: 5 mbar
-----LastLine-----		

Figure 3-10 : Onglet time-table crée pour la collecte des fractions A et B lors de l'analyse d'un échantillon de L3 + AT29B + thrombine. L'outlet 9 correspond à la collecte de la fraction A, l'outlet 6 correspond au vial de tampon de migration tandis que l'outlet 10 correspond au vial de collecte de la fraction B.

A partir du temps de migration des molécules, il suffit de calculer le temps qu'il leur reste à parcourir pour sortir du capillaire. Les molécules d'ADN ont un temps de migration compris entre 15 et 20 min. Le complexe migre vers 13min. Dans l'onglet time table de la méthode, il suffit d'indiquer à quel moment l'on doit changer de vials afin de pouvoir collecter la fraction désirée. (**Figure 3-10**).

3 LE CONTROLE QUALITE A L'ISSU DE LA COLLECTE DES ECHANTILLONS PAR ELECTROPHORESE CAPILLAIRE :

Lorsque la société GATC qui a réalisé le séquençage reçoit les échantillons, elle effectue un contrôle qualité initial afin de mesurer la taille attendue (84 paires de bases) et la quantité de l'échantillon. Cette étape se fait par analyse au fragment analyzer de l'échantillon reçu. Après l'étape de ligation des adaptateurs, sélection de la taille et amplification, l'échantillon est de nouveau soumis à un contrôle qualité final (**Figure 3-14**).

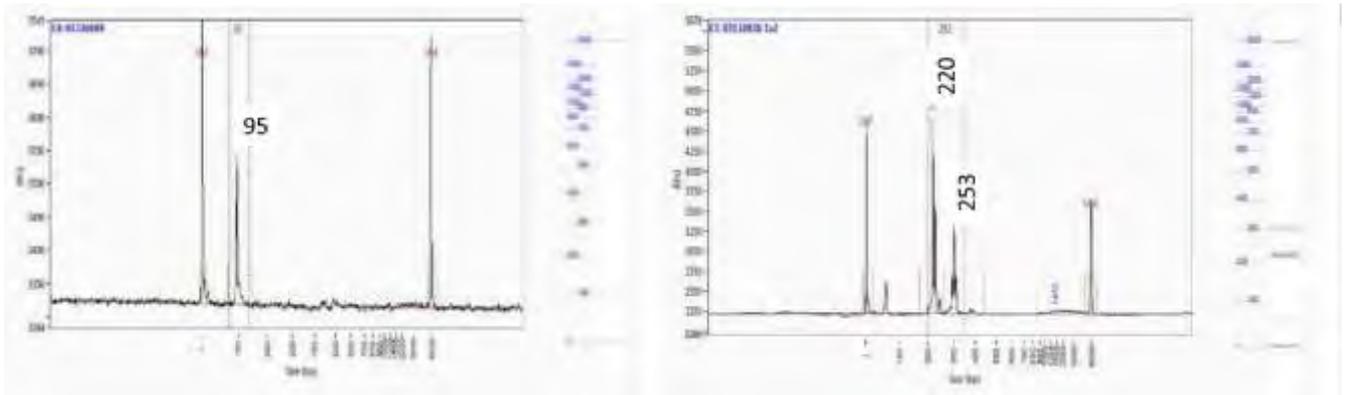


Figure 3-14 : Analyse au fragment analyzer de l'échantillon avant le séquençage (figure à gauche), où la taille de l'échantillon est de 84 paires de bases et après l'étape de ligation des adaptateurs, sélection de la taille et amplification. La taille attendue de l'échantillon attendue après cette étape est de 225 paires de bases.

L'analyse au fragment analyzer nous permet ainsi de confirmer que l'échantillon envoyé est de bonne qualité avant et après le séquençage

Pour chaque séquençage réalisée, l'analyse des échantillons au fragment analyzer par la société GATC qui a réalisé le séquençage, nous montre que l'échantillon reçu a un pic aux alentours de 95 paires de bases, ce qui montre que notre échantillon a bien la taille attendue. Après séquençage donc ajout des adaptateurs, un seul pic est toujours observé à une taille de 229 paires de bases. Nous pouvons donc affirmer que le séquençage s'est déroulé correctement (Figure 3-17).



Figure 3-17 : Contrôle qualité des échantillons avant séquençage (figure à gauche) et après séquençage (figure à droite). Celui-ci indique de par la présence d'un seul pic sur chacun des graphes que l'échantillon envoyé est pur et de bonne qualité.

Cependant, bien que l'échantillon envoyé à séquencer soit de bonne qualité (Figure 3-25), il s'avère que lors des étapes suivantes, un problème est survenu et après séquençage, le contrôle qualité de l'échantillon a donné de multiples pics. Ces pics non-spécifiques sont apparus durant l'amplification de la banque. L'explication donnée par GATC, est que des fragments non spécifiques se sont liés sur la flow cell et « ont pu » gêner la lecture des reads sur le fragment approprié (Figure 3-26).

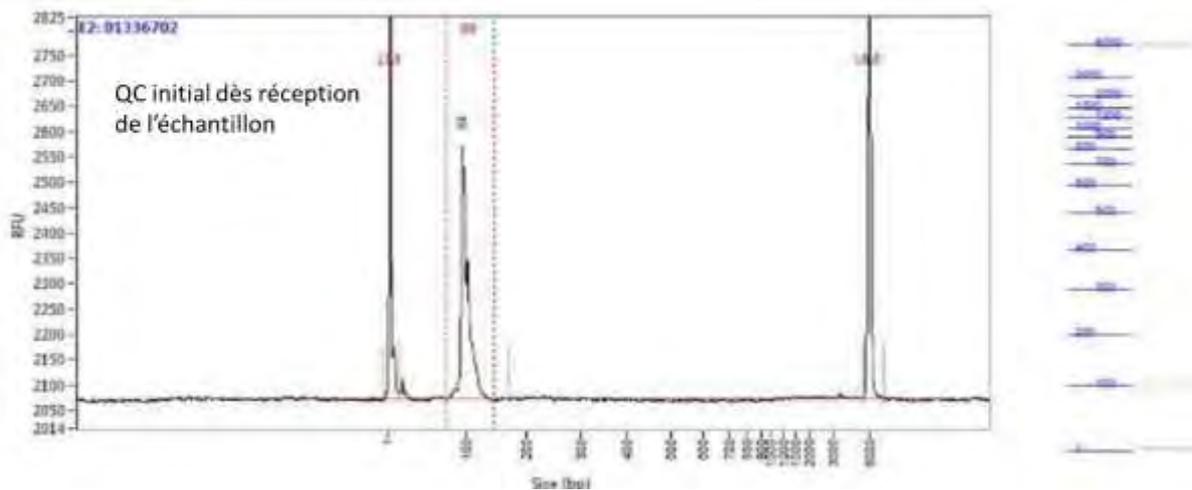


Figure 3-25 : Contrôle qualité de l'échantillon dès réception avant le séquençage au fragment analyzer.

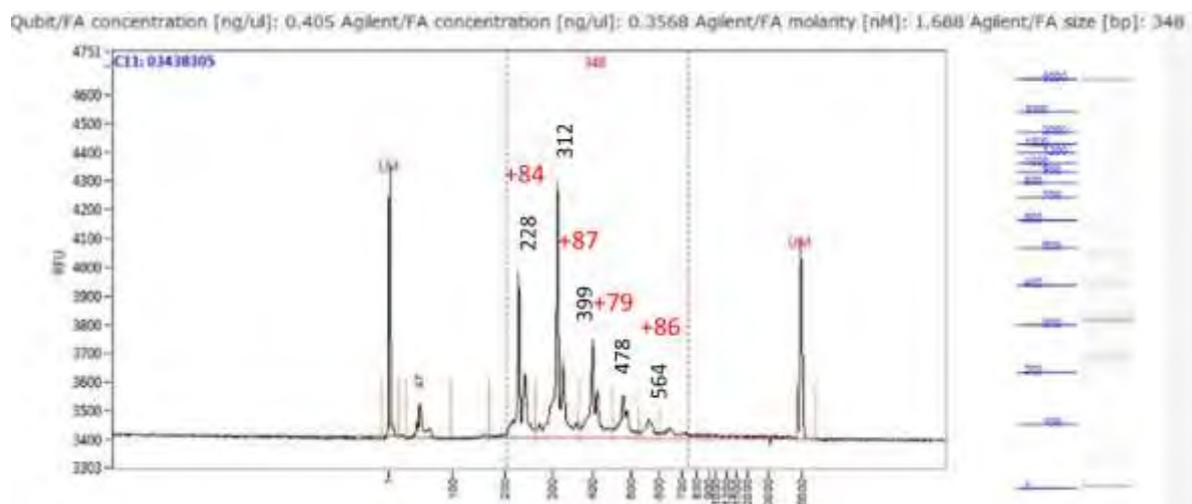


Figure 3-26 : Contrôle qualité après le séquençage.

1 PRINCIPE DE L'ANALYSE PAR ILLUMINA :

1.1 Principales définitions utilisées en bioinformatique :

Noms	Définitions
Lecture ou « read »	Séquences issues du séquençage Illumina
Mismatch	Décalage/inadéquation entre 2 nucléotides
Profondeur de lecture	Nombre moyen de lectures par base à une position donnée
.Fasta	Format permettant de représenter une ou plusieurs séquences. Dans ce cas, la séquence est précédée d'une ligne de titre commençant par le caractère >. Cela permet de mettre plusieurs séquences dans un même fichier.
.Fastq	Fichier texte combinant les séquences et les scores de qualité de séquençage associés à chaque position nucléotidique. Il est décrit par un @ pour l'identification de la séquence, la séquence, une ligne débutant par + et suivi de l'identifiant et les scores de qualités associés à chacune des bases.
Qualité des séquences	Analyse des données brutes à la sortie du séquençage pour s'assurer qu'il n'y a pas de problème ou de biais dans les données.
Multiplexer	Différents échantillons marqués avec leur barcode et qui sont poolés (mélangés) ensemble pour être séquencés sont dit multiplexer.
Barcodes ou « clés »	Courtes séquences d'oligonucléotides (6 ou 8) ajoutés aux échantillons d'ADN à séquencer
Gaps	Parties non séquencées dont les séquences sont très peu diversifiées
Contigs	Séquences continues générées par l'alignement de séquences de fragments qui se chevauchent
Paired-end	Le séquençage paired-end consiste à séquencer les reads par paires, ces reads étant séparés par une distance connue
Mi-seq	Il existe différents séquenceurs qui ont des capacités de séquençages différentes. Parmi ces séquenceurs, le Mi-seq est destiné à séquencer des reads de 0.3 à 15Gb en donnant 15 à 25 millions de séquences. Le temps de séquençage varie entre 5 et 60 heures, et les reads obtenus font jusqu'à 2*300bp. Le séquençage Hi-seq séquence des reads de 2*100bp pour une gamme comprise entre 100 et 600Gb. Le nombre de séquences est plus important (3000 millions de reads) et le temps de séquençage varie entre 2 et 11 jours.

1.2 Principales étapes de l'analyse bioinformatique

L'analyse bio-informatique est un processus complexe qui se déroule en plusieurs étapes. En effet, les logiciels de bureautique classiques ne sont pas compatibles pour stocker les données (nécessité d'une grande capacité de stockage) et pour réaliser les étapes d'analyses informatiques sur de telles quantités de données.

Les principales étapes du tri des données brutes sont résumées dans la figure suivante (**Figure 4-1**) :

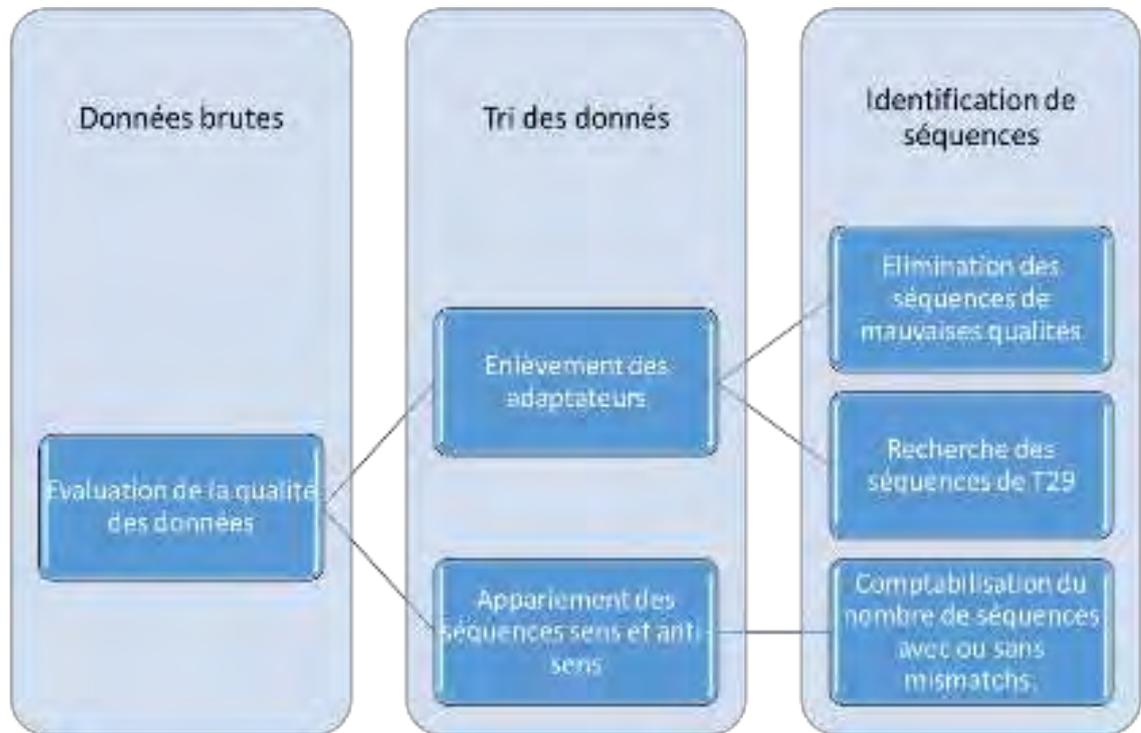


Figure 4-1 : Principales étapes de l'analyse bioinformatique.

A la fin du séquençage, un fichier «fasta» contenant les séquences issues du séquençage Illumina est téléchargé. Il contient plusieurs dizaines de millions de séquences qui doivent être traitées par des logiciels spéciaux.

1.3 Évaluation de la qualité des données grâce à Fastqc

La première étape consiste à vérifier la qualité des séquences c'est-à-dire vérifier qu'il n'y a pas eu de problème ou de biais de séquençage. Pour cela, on utilise un logiciel Fastqc qui lit un ensemble de fichiers de séquences et produit à partir de chacun d'eux un rapport de contrôle de la qualité des séquences, composé d'un certain nombre de modules différents.

Il est disponible gratuitement et permet de générer un rapport de la qualité moyenne des séquences sous forme html.

(<http://www.bioinformatics.bbsrc.ac.uk/projects/download.html#fastqc>)

Ainsi, nous obtenons des informations sur le type d'encodage utilisé pour les valeurs de qualité (**Figure 4-2**), le nombre de séquences, leurs longueurs, le pourcentage de GC, le nombre de séquences surreprésentées etc...



Figure 4-2 : Un exemple d'un rapport Fastqc représentant la qualité des séquences lors du séquençage. Pour chaque position, un tracé de ce type est obtenu. La ligne rouge pointillée est la valeur médiane. La ligne bleue représente la qualité moyenne. L'axe des ordonnées sur le graphique montre les scores de qualité. Plus le score est élevé, plus la base est de qualité, c'est-à-dire qu'elle est à la même position. L'arrière-plan du graphique divise l'axe des y en très bonne qualité (vert), qualité raisonnable (orange) et mauvaise qualité (rouge). La qualité des séquences sur la plupart des plateformes se dégrade au fur et à mesure que le séquençage progresse, de sorte qu'il est courant de voir les dernières bases séquencées chuter dans la zone orange vers la fin d'une lecture.

1.4 Nettoyage des reads – filtrage des séquences d'adaptateurs

Nous utilisons une banque d'ADN dégénérée sur N=30 nucléotides (Figure 4-3).



Figure 4-3 : Banque d'ADN (N= 30) avec les adaptateurs contenant les clés sens (6 nucléotides) et les clés anti-sens (8 nucléotides). Cette étape a lieu durant la PCR. Au cours de l'étape de séquençage Illumina, des adaptateurs sont de nouveaux ajoutés afin que les brins d'ADN puissent se fixer sur la plaque Illumina.

La banque d'ADN et l'aptamère T29 sont encadrés par des séquences permettant une amplification par PCR à l'aide d'amorces auxquelles on rajoute des clés ou barcodes : (Barcodes sens : 6 nucléotides / barcodes antisens : 8 nucléotides) qui permettent de multiplexer plusieurs échantillons (Figure 4-4).

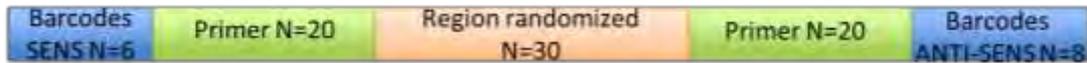


Figure 4-4 : Détail de la banque d'ADN (N=30) avec les primers contenant les clés sens et anti-sens

La somme des barcodes + primers et de la région d'intérêt correspond à un total de 84 ou 83 nucléotides. Grâce à ces barcodes, on peut attribuer facilement une séquence à un échantillon d'ADN (Figure 4-5).

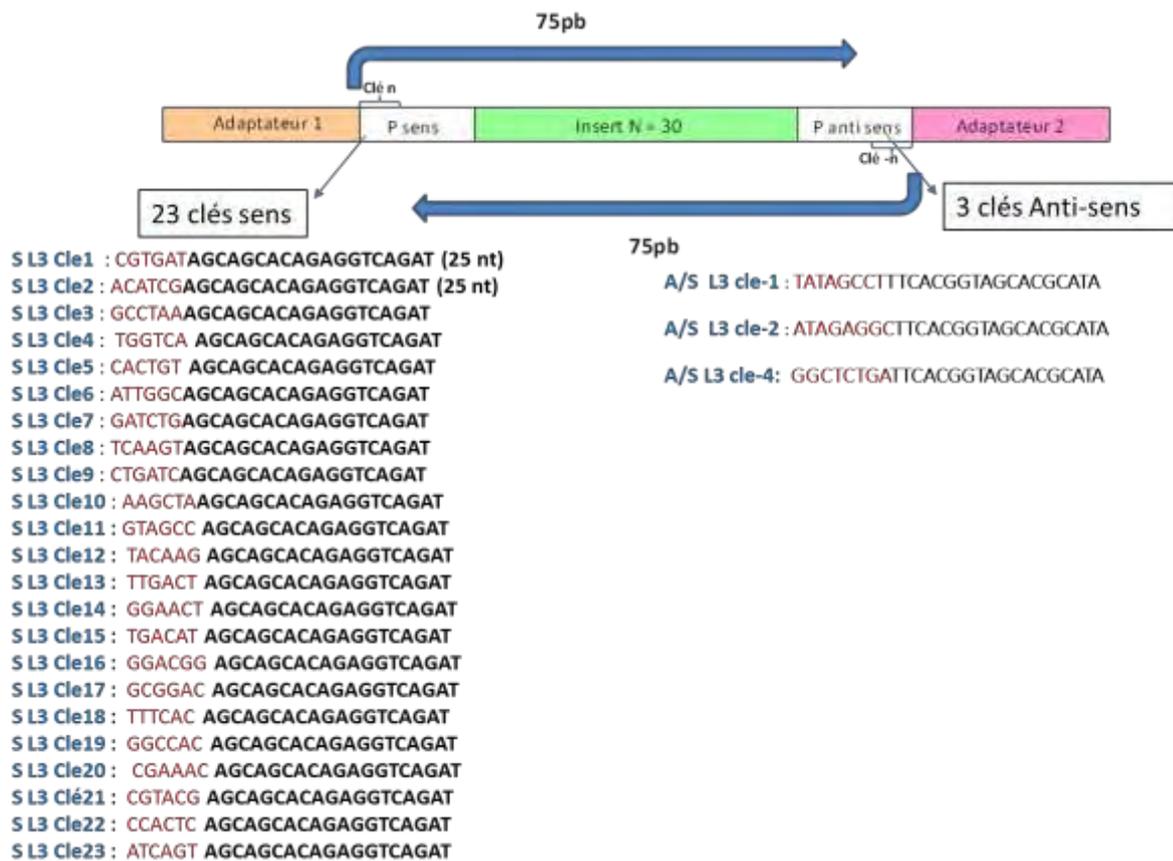


Figure 4-5 : Tableau de l'ensemble des barcodes utilisés (sens et antisens)

Lorsque l'on trie les séquences des adaptateurs, il faut donc veiller à comparer chaque séquence (et son reverse complément) avec chacune des autres séquences (Figure IV-5). L'identification correcte des adaptateurs est facilitée en maximisant l'autocorrélation entre les 2 reads par le fait qu'il est nécessaire d'avoir des positions identiques pour les débuts d'adaptateurs. Les paramètres importants sont la longueur de chevauchement, la stringence (le pourcentage de bases identiques) et la longueur de la séquence.

Une fois que le tri des barcodes est réalisé, deux fichiers sont créés. Ils contiennent chacun toutes les données associant un barcode sens dans une direction de lecture (read 1) et dans le sens inverse de lecture (read 2). Un troisième fichier contenant les données n'ayant pu être associées à aucun barcode est créée, représentant entre 1 et 10 % des données.

1.5 L'alignement des séquences

L'alignement constitue une des étapes les plus importantes car il transforme une information simple (les données triées en fonction de leurs barcodes) en une information globale et structurée (les données regroupées et alignées par rapport à une séquence de référence). La séquence de référence est celle de l'aptamère T29. Lorsque nous réalisons l'alignement des reads sur la séquence du T29, nous pouvons faire varier la stringence de l'alignement en faisant varier le nombre de mismatches (nombre d'erreurs de séquences). Nous pouvons faire varier cette valeur jusqu'à 6 nucléotides. Les traitements des données ont été réalisés avec des mismatches allant de 0 jusqu'à 3 nucléotides.

1.6 Profondeur de lecture = le nombre de fois où chaque nucléotide est lu

La profondeur de lecture du séquençage Illumina (1000 copies de chaque fragment) permet de séquencer plusieurs fois la même molécule d'ADN (au moins 10 fois) afin d'être sûr de la base de chaque séquence. Le séquençage haut-débit NGS est aussi appelé « deep sequencing » car même si la fiabilité de chaque séquence individuelle est variable, elle est compensée par le nombre important de séquences (reads) du même nucléotide à une position donnée correspondant à la profondeur du séquençage.

2 UTILISATION DE LA PCR - ILLUMINA POUR SEQUENCER LE T29

L'objectif de cette partie était de déterminer le rapport entre le nombre de molécules et le nombre de séquences obtenues par le séquençage haut-débit Illumina. Nous voulions aussi connaître le nombre minimal de molécules d'aptamères AT29B que le séquenceur Illumina pouvait détecter, c'est-à-dire quelle était la limite basse de détection d'un aptamère avec une structure en G-quadruplex par le séquençage Illumina.

Pour répondre à ces questions, nous avons décidé de réaliser une gamme étalon de l'aptamère T29, afin de déterminer si le séquençage Illumina était capable de séquencer une gamme étalon de 3 molécules jusqu'à $4,5 \cdot 10^{12}$ molécules sans injection par électrophorèse capillaire.

2.1 Réalisation de la gamme étalon

Pour réaliser la gamme étalon, plusieurs aliquots de l'aptamère AT29B ont été amplifiés par PCR en réalisant 28 cycles d'amplifications (Figure 4-6). Des amorces comportant des clés différentes ont été introduites pour chaque aliquot afin de pouvoir les identifier lors du séquençage de chaque échantillon. Une PCR avec un nombre de cycles poussé (28 cycles) de l'aptamère AT29B avec des clés différentes a été réalisée. Après purification et quantification, 1 µl de chaque échantillon a été ajouté avec les échantillons injectés et collectés par électrophorèse capillaire afin d'être séquencés. Cette gamme a été réalisée trois fois, en rajoutant dans le pool final différentes concentrations de l'aptamère AT29B.

Après l'étape d'amplification par PCR, ces aliquots ont été purifiés, puis déposés sur gel d'agarose 2% afin de vérifier que l'intensité de chaque bande d'ADN était la même (Figure 4-7).

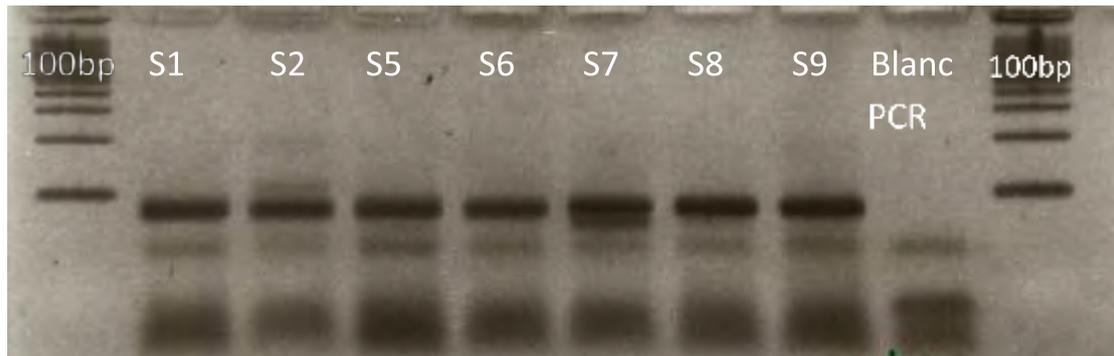


Figure 4-6 : Dépôt de 2 μ L sur gel d'agarose 2% de l'aptamère AT29B amplifié après 28 cycles de PCR. Chaque aptamère T29 a été amplifié avec des clés différentes S1, S2, S5, S6, S7, S8, S9. Le blanc PCR contient du mix de PCR, les clés sens et anti-sens ainsi que de l'eau à la place de la matrice à amplifier. Cela permet de contrôler qu'il n'y a pas eu de contamination.

Une deuxième mesure au nanodrop a permis de quantifier en nanogramme/ μ L la concentration, puis la quantité de chaque aliquot. Ensuite des dilutions successives ont été faites jusqu'à obtenir une gamme étalon allant de 9 molécules jusqu'à $4.5 \cdot 10^{12}$ molécules. 1 μ L de chaque échantillon a été introduit avec les autres échantillons collectés par électrophorèse capillaire et l'ensemble a été envoyé à séquencer. Si le coût du nucléotide est faible, chaque expérience de séquençage est coûteuse (jusqu'à 10 000 euros pour 4 séquençages). Il était trop onéreux de réaliser un séquençage de la gamme étalon seule. C'est la raison pour laquelle, cette gamme a été ajoutée à d'autres échantillons. Pour chaque séquençage, entre 16 et 18 échantillons différents ont été séquencés. Entre quatre et huit points de la gamme étalon ont été ajoutés au pool à séquencer. Au total, trois gammes étalons ont été réalisées afin de mesurer la reproductibilité des résultats.

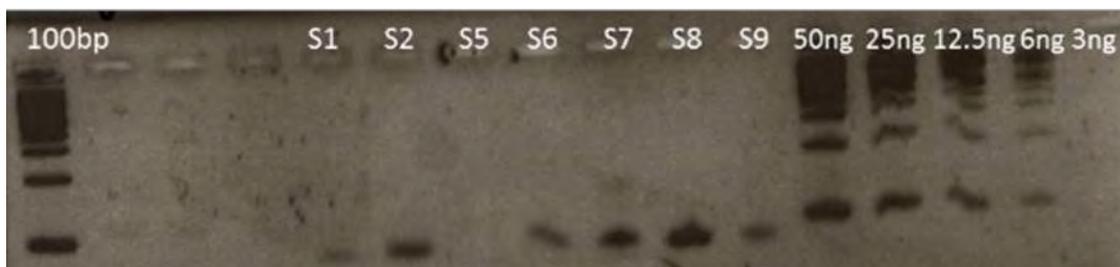


Figure 4-7 : Dépôt de 1 μ L d'aptamère AT29B après amplification et purification sur gel d'agarose 2%.

2.2 Caractérisation du nombre de séquences de T29 obtenues – étude de la réponse de l'Illumina

Après traitement informatique, chaque séquence ayant un barcode identique correspondant à un nombre de molécules donné, a pu être attribué à la séquence du T29. Ainsi, parmi les séquences contenant les barcodes d'un nombre de molécules connu de T29, celles-ci ont été comparées avec la séquence de référence de l'aptamère T29. Le rapport entre le nombre de séquences de T29 et le nombre de séquences contenant la même clé pour un échantillon donné correspond au rendement de séquençage.

Nombre de molécules	Nombre de séquences	Nombre de séquences de T29	Rendement de séquençage
9	31	31	100%
100	8554	7783	91%
10 ⁶	87 767	82 765	94%
9*10 ⁶	272	231	85%
9*10 ⁷	937	822	88%
10 ⁸	72 397	68 883	95%
9*10 ⁸	2 850	1 895	66%
10 ¹⁰	3 136	2 735	87%
4.5*10 ¹¹	146 660	127 676	87%
9*10 ¹¹	161 255	153 467	95%
1*10 ¹²	148 109	141 665	96%
4.5*10 ¹²	1 283 019	1 218 588	95%

Figure 4-8 : Tableau indiquant pour chaque nombre de molécules d'aptamères T29 séquencées, le nombre de séquences avec la même clé (2ème colonne), le nombre de séquences identiques à la séquence de l'aptamère T29 (3ème colonne) et enfin la 4ème colonne indique le rendement c'est-à-dire le rapport entre le nombre de séquences de T29/ le nombre de séquences avec la clé. Ce rapport est donné en pourcentage.

En comparant les séquences contenant les barcodes avec la séquence du T29, on s'aperçoit que les séquences de taille N=83 sont quasiment toutes que du T29 (Figure 4-8). Il y a donc eu très peu de mutation de séquences au cours de la PCR et de l'analyse Illumina. En effet, en calculant le rendement, on remarque que ce dernier est compris entre 84% et 100%, excepté pour 9*10⁸ molécules où l'on obtient un rendement de (66%).

On peut donc étudier la relation entre le nombre de séquences de T29 obtenues après Illumina et le nombre de molécules de T29 introduites dans les échantillons. En comparant les valeurs des trois séquençages, on trouve que cette relation est du type « log-log ». La même série observée en repère cartésien conduit à un tassement des premiers points pour permettre le placement du dernier point et montre des points se plaçant vaguement sur une courbe polynomiale. Ce type de graphe « log-log » est utilisé pour représenter des phénomènes où y est une fonction puissance de x. On montre dans ce graphe que le nombre de séquences de T29 suit une fonction puissance par rapport au nombre de molécules de T29 séquencées. Cette évolution se fait selon l'équation $y=157,28x^{0,2259}$. On peut donc dire que le nombre de séquences de l'aptamère T29 est relativement proportionnel aux nombres de molécules.

On peut aussi voir que le séquençage Illumina permet de détecter au minimum 9 molécules (31 séquences obtenues) jusqu'à 4.5*10¹² molécules (1 218 588 séquences).

2.3 Reproductibilité

3 gammes étalons ont été réalisées (Figure 4-9).

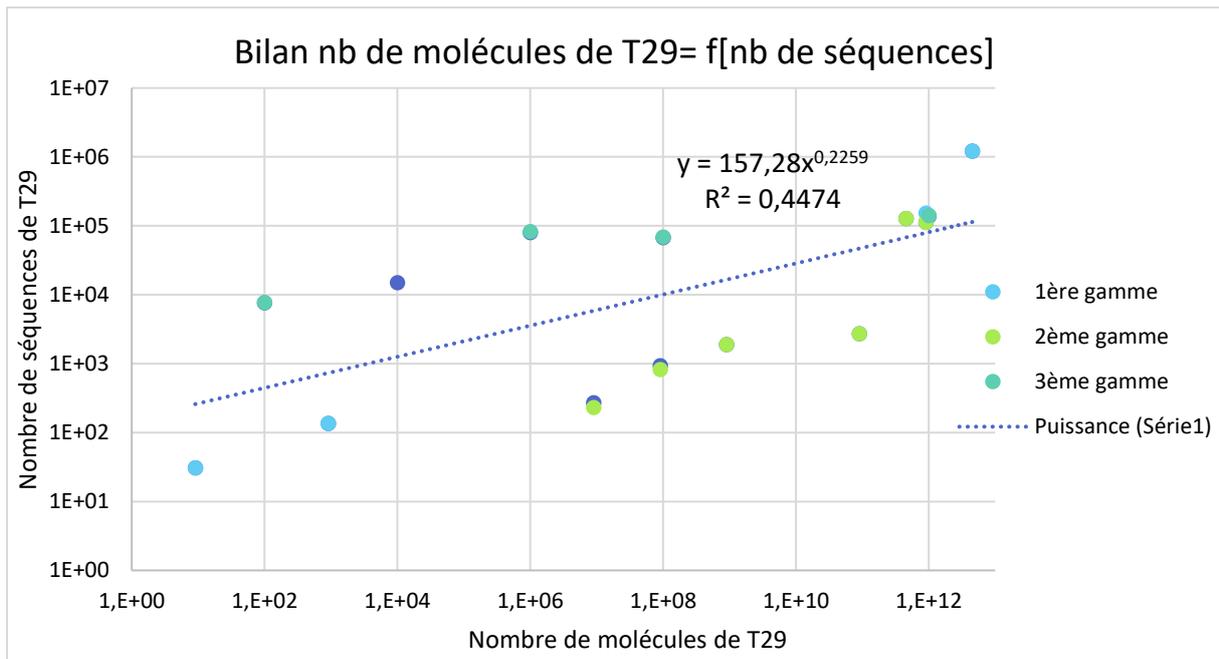


Figure 4-9 : Courbe nombre de molécules de T29= f (nombre de séquences du T29) avec tous les points des 3 séquençages.

La première gamme comprenait 8 points allant de 9 jusqu'à $4.5 \cdot 10^{12}$ molécules (Figure 4-10). Pourtant après le séquençage, seuls 4 points (9, 900, $9 \cdot 10^{11}$ et $4.5 \cdot 10^{12}$ molécules) ont été retrouvés. L'explication la plus plausible serait une mauvaise purification sur gel de chaque bande. La détermination de la concentration des échantillons aurait été biaisée donc il y a un doute sur la quantification du nombre de molécules. Néanmoins, nous trouvons un rendement de séquences correctes (entre 95% et 100%) sur les échantillons ayant pu être séquençés.

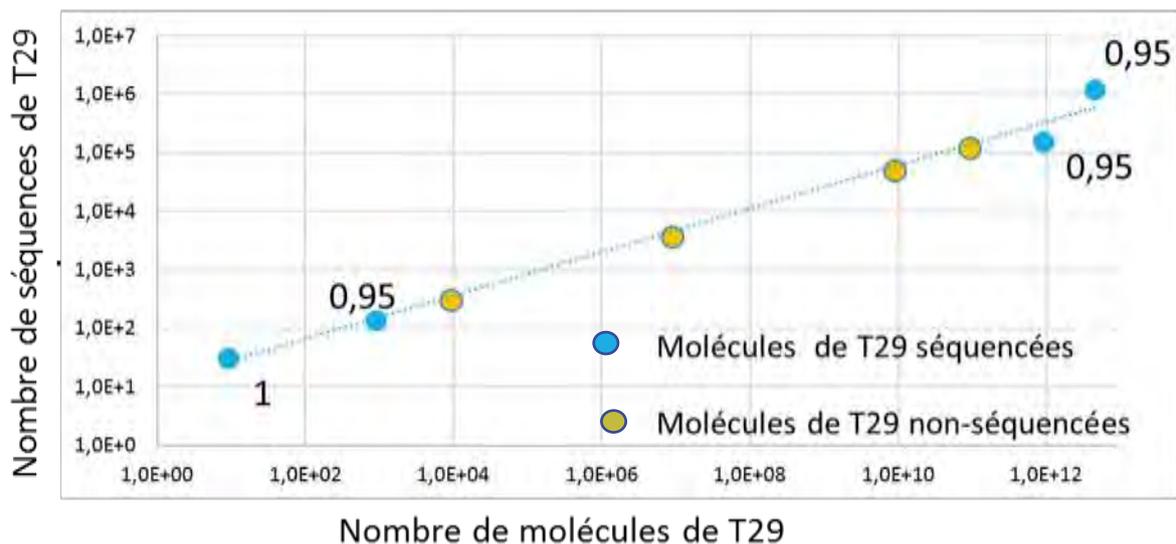


Figure 4-10 : Première gamme étalon de l'aptamère T29 séquençé par Illumina. 9, 900, $9 \cdot 10^{11}$ et $4.5 \cdot 10^{12}$ molécules d'aptamères ont été séquençés. Seuls 4 points ont été détectés lors de l'analyse bioinformatique (points bleus). Les points en jaunes correspondent aux molécules de T29 séquençées mais dont les clés n'ont

pas été retrouvées lors de l'analyse informatique. La valeur indiquée pour chaque point correspond au rendement de séquences Les points en jaune ont été séquencé mais pas retrouvés dans les datas.

La deuxième gamme a été réalisée en reprenant les points intermédiaires non-séquencés de la première gamme, c'est-à-dire 6 points : $9 \cdot 10^6$ – $9 \cdot 10^7$ – $9 \cdot 10^8$ – $9 \cdot 10^{10}$ – $4,5 \cdot 10^{11}$ – $9 \cdot 10^{11}$ molécules (Figure 4-11). Ces points ont été préparés à partir d'un nouvel échantillon de T29 amplifié avec différents barcodes par PCR puis purifiés, quantifiés et dilués jusqu'à atteindre la concentration voulue.

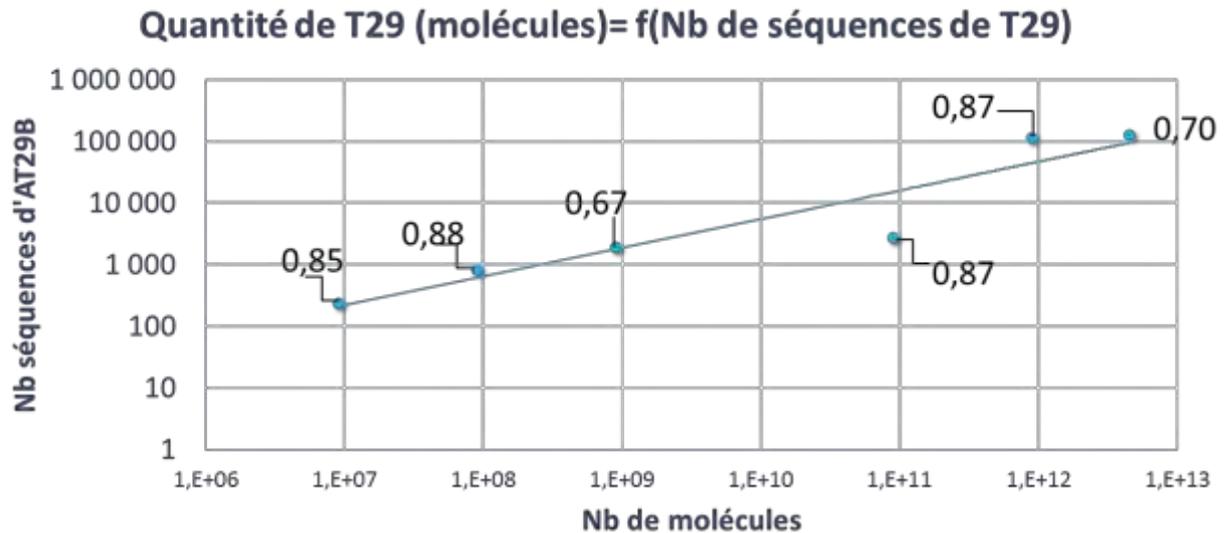


Figure 4-11 : Deuxième gamme étalon réalisée à partir des points suivants : $9 \cdot 10^6$ – $9 \cdot 10^7$ – $9 \cdot 10^8$ – $9 \cdot 10^{10}$ – $4,5 \cdot 10^{11}$ – $9 \cdot 10^{11}$ molécules. Tous les points de la gamme ont été séquencés. Les valeurs à côté de chaque point correspondent au rendement de séquences.

La troisième gamme étalon comprend 4 points : 100 molécules – 10^5 molécules – 10^7 molécules – $9 \cdot 10^{11}$ molécules (Figure 4-12).

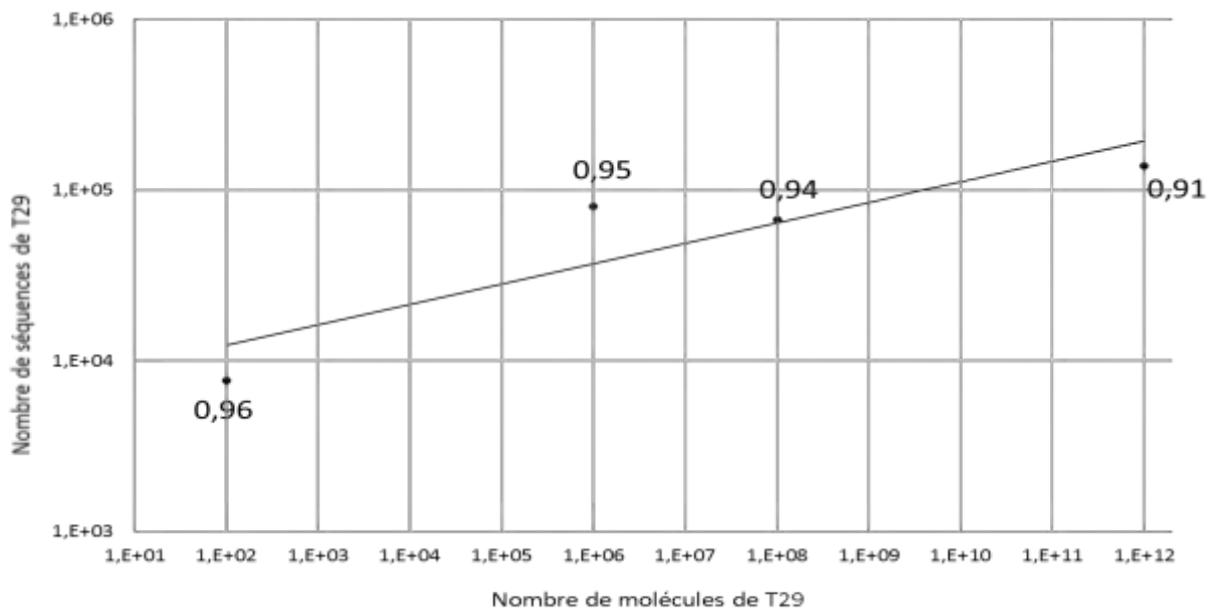


Figure 4-12 : Troisième gamme étalon réalisée à partir des points suivants : 100 molécules – 10⁵ molécules – 10⁷ molécules – 9*10¹¹ molécules. La pente de la courbe est différente de celles des autres séquençages car le nombre de séquences est plus important pour chaque point de la gamme.

En comparant les trois gammes (Figure 4-13), on remarque qu'il n'y a pas de reproductibilité entre chaque expérience. Cependant, les pentes des gammes 1 et 2 sont proches ($x_1 = 0.3706$ et $x_2 = 0.5011$). La troisième gamme correspondant au 4^{ème} séquençage nous donne une pente différente des deux précédentes gammes ($x_3 = 0.1196$). Cela nous montre bien que la mauvaise qualité lors de l'étape de préparation de la librairie gène dans l'interprétation des résultats.

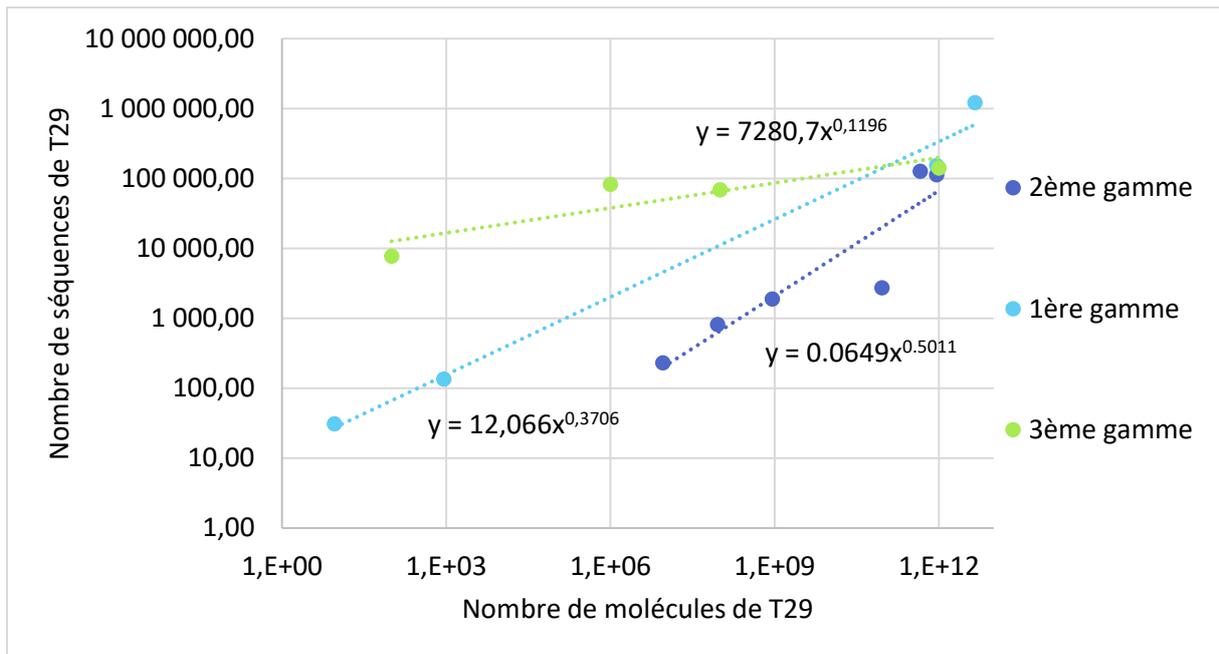


Figure 4-13 : Superposition des courbes nombre de molécules de T29 = f (nombre de séquences de T29) pour chaque gamme étalon. Le séquençage Illumina d'un nombre connu de molécules d'aptamères T29 n'est pas reproductible.

2.4 Publication



410:1991–2000
<https://doi.org/10.1007/s00216-018-0865-5>

G-quadruplex aptamer selection using capillary electrophoresis-LED-induced fluorescence and Illumina sequencing

Audrey Ric^{1,2,3} & Vincent Ecochard² & Jason S. Iacovoni⁴ & Audrey Boutonnet³ & Frédéric Ginot³ & Varravaddheay Ong-Meang¹ & Véréna Poinso¹ & Laurent Paquereau² & François Couderc¹

Received: 27 October 2017 / Revised: 28 December 2017 / Accepted: 8 January 2018 / Published online: 29 January 2018

Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

One of the major difficulties that arises when selecting aptamers containing a G-quadruplex is the correct amplification of the ssDNA sequence. Can aptamers containing a G-quadruplex be selected from a degenerate library using non-equilibrium capillary electrophoresis (CE) of equilibrium mixtures (NECEEM) along with high-throughput Illumina sequencing? In this article, we present some mismatches of the G-quadruplex T29 aptamer specific to thrombin, which was PCR amplified and sequenced by Illumina sequencing. Then, we show the proportionality between the number of sequenced molecules of T29 added to the library and the number of sequences obtained in Illumina sequencing, and we find that T29 sequences from this aptamer can be detected in a random library of ssDNA after the sample is fractionated by NECEEM, amplified by PCR, and sequenced. Treatment of the data by the counting of double-stranded DNA T29 sequences containing a maximum of two mismatches reveals a good correlation with the enrichment factor (f_E). This factor is the ratio of the number of aptamer sequences found in the collected complex sample divided by the total number of sequencing reads (aptamer and non-aptamer) plus the quantity of T29 molecules (spiked into a DNA library) injected into CE.

Keywords Aptamer · Capillary electrophoresis · Biomolecule · Thrombin

Introduction

Berezovski [1] and Krylov [2] described the selection of aptamers by the non-equilibrium capillary electrophoresis (CE) of equilibrium mixtures (NECEEM). Since these studies were completed, many authors have used this method to select aptamers [3, 4] or have tried to describe the method using a mathematical model [5]. This aptamer selection method

* François Couderc
couderc@chimie.ups-tlse.fr

¹ Laboratoire des IMRCP, UMR 5623, Université de Toulouse, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, France

² Institut de Pharmacologie et de Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, 205 Route de Narbonne, 31077 Toulouse, France

³ Picometrics Technologies, 478 Rue de la Découverte, 31670 Labège, France

⁴ I2MC, UMR1048, 1 Avenue du Professeur Jean Poulhès, 31432 Toulouse Cedex 4, France

involves injecting a CE into a mixture of a degenerate DNA library with the target to obtain an aptamer specific to the target molecule (M). If the complex is sufficiently stable, three possible peaks should be obtained: one for the free ssDNA, one for the molecule M alone, and one for the {ssDNA.M} complex, which should be very weak, even undetectable, due to its very low concentration. However, even if the peak is not detected, the complex can be collected between the migration time of the molecule M and that of the ssDNA library. After this collection step, the aptamers with affinities for M can be enriched by re-injecting the sample into CE at a constant concentration of M. Finally, the enriched specific aptamers can be sequenced. This analytical strategy does not use the classical SELEX protocol [6] but is a non-SELEX method, in contrast to the methods of Luo et al., who developed a fraction collection approach involving capillary electrophoresis and SELEX (FCE-SELEX) [7].

G-quadruplexes are high-order DNA structures formed from G-rich sequences. They consist of parallel, antiparallel, or mixed parallel/antiparallel-stranded configurations with loops and tails of various sequences. They are built around nearly planar tetrads of hydrogen-bonded guanine bases.

These structures may be difficult to amplify by PCR and to sequence because polymerase stalling at G-tetrad sites affects the specific incorporation of bases [8]. Very few authors have selected aptamers containing a G-quadruplex using CE [9]. In the above paper, Kasahara et al. published a 15-mer aptamer of thrombin (Tbn) that contains a G-quadruplex.

Thrombin is a multifunctional serine protease that plays a key role in anticoagulant function. Two different thrombin aptamers were identified in the 1990s, and a recent review described their intensive use [10]. The 29-nucleotide aptamer (T29), which was selected by SELEX 20 years ago, was reported to have a K_D of approximately 0.5 nM [11]. The sequence of T29 is 5'-A GTCCGTG G TAGG G CAGGT4TGGGGTACT-3', and this aptamer contains a G-quadruplex. A very interesting study was performed on the dynamic binding between thrombin (Tbn) and T29 in the presence of K^+ and Na^+ in the sample buffer [12]. The study showed that the presence of these ions at concentrations above 5 mM contributes greatly to the formation of the {T29.Tbn} complex (Tbn concentration 20 μ M). An optimum reaction time above 5 min was found for a T29/Tbn ratio of 1/10.

A transient isotachopheresis (tITP) pre-concentration step and the use of SYBR Gold as a non-covalent fluorescent dye were proposed by Riley et al. [13] to study the separation of Tbn and T29. This method shows two distinct peaks for the T29 ssDNA and the complex, and a K_D of 128 nM was measured. This experiment was extended by using a known amount of T29 in a randomized DNA library (randomer). As expected, a single peak was observed after an injection of a 1:100 DNA mixture (2 nM aptamer, 200 nM randomer). However, when Tbn (400 nM) was added to the previous 1:100 mixture, two new signals were observed. One of them was the {T29.Tbn} complex peak, whereas the second peak had the same migration time as the non-specific complex peak observed between the random library and thrombin [14]. These authors used the Ion Torrent Personal Genome Machine to sequence the DNA of the CE-collected {T29.Tbn} complex fractions to determine the selection efficiency. They called the measurement the enrichment factor, f_E , which is the ratio of the number of aptamer sequences found in the collected complex sample to the total number of sequencing reads (aptamer and non-aptamer). The value of f_E was shown to vary between 6- and 93-fold, with an average enrichment of 38 for 5 sequencing runs of the same sample.

Huge et al. [15] tested their drop-on-demand ink-jet printer interfaced with a CE instrument to collect fractions at the output of the capillary. The collector deposits one 10- μ L drop of a sheath liquid in each well along with a few picoliters of solution from the capillary (fraction width duration 7 s). With CE/SELEX conditions in the presence of 10 μ M of a random ssDNA library degenerated on 30 nucleotides (L30) and 30 μ M Tbn, the fractions were collected from a 50- μ m i.d. capillary. Forty cycles of PCR were run, and Illumina

sequencing was used to characterize the DNA binders. With the Burrows-Wheeler alignment tool, over 17,606 sequences with close matches to T29 were identified.

No pre-existing study has yet answered the question of whether a non-SELEX/NECEEM experiment can be used to select an aptamer against Tbn. In this study, we evaluated the possibility of selecting a known aptamer containing a G-quadruplex by CE and high-throughput DNA sequencing (Illumina). We chose T29 and Tbn as a model to show how the sequence of T29 DNA can be identified when spiked at low quantities into an ssDNA library.

Materials and methods

Materials and reagents

Sodium phosphate dibasic anhydrous, sodium chloride, hydrochloric acid, magnesium chloride, sodium hydroxide, Tris/borate/EDTA (TBE) buffer, ammonium persulfate, agarose for gel electrophoresis, xylene cyanol, bromophenol blue, and thrombin were purchased from Sigma-Aldrich (St. Quentin Fallavier, France). The sequence of the T29 ssDNA (capital letters) and the PCR primers (lowercase letters) was 5'-agcagcacagaggtcagatgAGTCCGTGGTAGGGCAGGTTGGGGTACTcctatgcgtgctaccgtgaa-3' (T29₇₀); thus, T29₇₀ is T29 with the 5' and 3' primers that are used for PCR. The ssDNA random library was 5'-agcagcacagaggtcagatg-N30-cctatgcgtgctaccgtgaa-3' (L30). The PCR primers were labeled at the 5' end with a fluorophore, 6-carboxyfluorescein (6-FAM) (Sigma-Aldrich). The forward primer was 5'-XXXXXXAGCAGCACAGAGGTCAGATG-3', and the reverse primer was 5'-YYYYYYYYTTTCACGGTAGCACGCATAGG-3'. Sequences X6 and Y8 were the barcodes to allow researchers to identify the PCR work performed.

CE-UV-LIF and fraction collection

An Agilent Technologies CE 7100 instrument (Waldbronn-Germany) was equipped with a 100- μ m i.d. fused silica capillary (Polymicro Technologies, Phoenix, Arizona) with a total length of 90 cm. A Zetalif LED detector from Picometrics Technologies (Toulouse, France) was equipped with an LED with a wavelength of 480 nm. Two detection windows were defined, one at 8.5 cm from the capillary outlet for the UV window and the other at 21 cm at the capillary outlet for fluorescence detection. To prepare the mixture, the T29₇₀ aptamer was prepared from a 100 μ M stock solution. Next, the aptamer sample was heated at 95 °C for 5 min and then cooled to room temperature. Finally, 10 μ L of Tbn at 10 nM was incubated with 10 μ L of aptamer at various concentra-

tions, and the mixture was injected into the CE instrument. For the spiking experiments, 200 nL of a mixture containing

10 μ M L30, 10 nM thrombin, and T29₇₀ aptamer at various quantities (from 3 to 10⁸ molecules) was added to a buffer composed of 25 mM phosphate buffer (pH 8.2) with 100 mM NaCl and 1 mM MgCl₂; then, the samples were injected by CE at 50 mbar for 30 s. Due to the double detection by UV absorbance and fluorescence, thrombin was detected by 210 nm UV at a migration time of 9 min instead of the library. The {T29.Tbn} complexes were detected by fluorescence at 480 nm at 18 and 13 min, respectively. The complexes could not be detected by UV because of their absorption in relation to their abundance. We used a 50 mM phosphate buffer, pH 8.2. Phosphate buffer does not use nitrogen molecules as a buffer; Tris buffer was excluded from this type of experiment [16]. Species migration was performed with a voltage of + 15 kV and a pressure of 5 mbar. Fraction collection was performed according to a timetable in OpenLAB Chemstation (software of CE 7100), which allowed us to collect {T29.Tbn} complex samples between 12.6 and 17.9 min.

PCR and DNA purification

The reagents used for the PCR amplification were 29 μ L of water, 5 μ L of 10 \times buffer (500 mM KCl, 100 mM Tris-HCl pH 8.3, 15 mM MgCl₂), 5 μ L of forward primer at 5 μ M, 5 μ L of reverse primer at 5 μ M, 5 μ L of 2.5 mM dNTPs, 1 unit of the enzyme Taq DNA Polymerase, and 1 μ L of the ssDNA sample. The PCR program applied was 95 $^{\circ}$ C for 3 min, 95 $^{\circ}$ C for 30 s, 55 $^{\circ}$ C for 30 s, 72 $^{\circ}$ C for 30 s, 72 $^{\circ}$ C for 3 min, and a return to 20 $^{\circ}$ C. All samples were amplified for the first time with 10 cycles. Then, 5 μ L of the first PCR was diluted 10 times in a novel PCR mixture, and a second PCR of 8 cycles was performed. After this amplification step, 2 μ L of each sample was loaded on a 2% agarose gel, and after migration, the dsDNA bands were revealed with ethidium bromide. The amplified samples were 84 bp long. The Nucleospin Purification Kit (Macherey Nagel, Hoerd, France) was used to purify the ssDNA.

Quantification of the T29₇₀ sequence in Illumina sequencing

The T29₇₀ aptamer was amplified with the same PCR program but with 28 cycles. To determine the proportional response between the number of T29₇₀ molecules subjected to Illumina high-throughput sequencing and the number of sequences obtained by this sequencing method, we conducted different PCR runs with different barcode primers (six forward nucleotides and eight reverse nucleotides). Each forward barcode corresponded to a known quantity used in Illumina. The aim of these amplifications was to obtain a high quantity of T29₇₀ labeled with a specific forward barcode. After PCR amplification, the T29₇₀ samples with different barcodes were loaded on 2% agarose gel. After migration, the DNA band

was cut and dissolved at 55 $^{\circ}$ C during 5 min in the buffer

NT1 from the Nucleospin Purification Kit. Then, this mixture was loaded onto the silica membrane of the kit. Through centrifugation, the PCR reactants were removed. The purified dsDNA was eluted with 5 mM Tris-HCl, pH 8.5, to a final volume of 15 μ L. All the amplified and purified samples were then quantified with a NanoDrop 2000 UV spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Then, 2 μ g of the mixture of all the dsDNAs was sent to GATC Biotech (Konstanz, Germany) for Illumina sequencing. Each sequencing reaction contained at least 24 different samples prepared under different conditions.

For each run, a unique 5' and 3' end sequence was used to assign each read in the multiplexed Illumina sequence data to a sample. We required any reads used in downstream analyses to be the appropriate length and contain perfect matches to the barcodes and primers to assign a read to a sample. This matching was performed in R with standard string comparison and manipulation. The central sequence regions (length 29 or 30 bp) were extracted for each sample.

CE collection and quantification of the T29₇₀ sequence in Illumina

After collection from CE, each fraction was amplified with 10 cycles of PCR, diluted 10-fold and amplified again with 8 cycles of PCR. Purification was performed with the Nucleospin Purification Kit. Quantification was also conducted with the NanoDrop. Next, 2 μ g of the mixture of all the dsDNAs was sent to GATC Biotech for Illumina sequencing. Figure 1 presents the results of these experiments.

Bioinformatic treatment of the data

Cleaning and merging of paired-end sequences

Each Illumina paired-end sequence was run through a three-step pipeline to generate a single fastq file for subsequent downstream analyses. Sickle was used to trim low-quality reads with the command "sickle pe -f R1.fastq -r R2.fastq --qual-type sanger --qual-threshold 20 --length-threshold 70 --truncate-n -o cleanR1.fastq -p cleanR2.fastq". Then, spades were run in error correction mode with the command "spades.py --only error-correction --pe-1 cleanR1.fast --pe-2 cleanR2.fastq." The resulting fastq files from spades were used with usearch to merge R1/R2 pairs into a single sequence with the command "usearch -fastq_mergepairs correctedR1.fastq -reverse correctedR2.fastq -fastqout cleanedMerged.fastq -fastq_merge_maxee 1 -fastq_minovlen 20".

Splitting the barcoded sequences

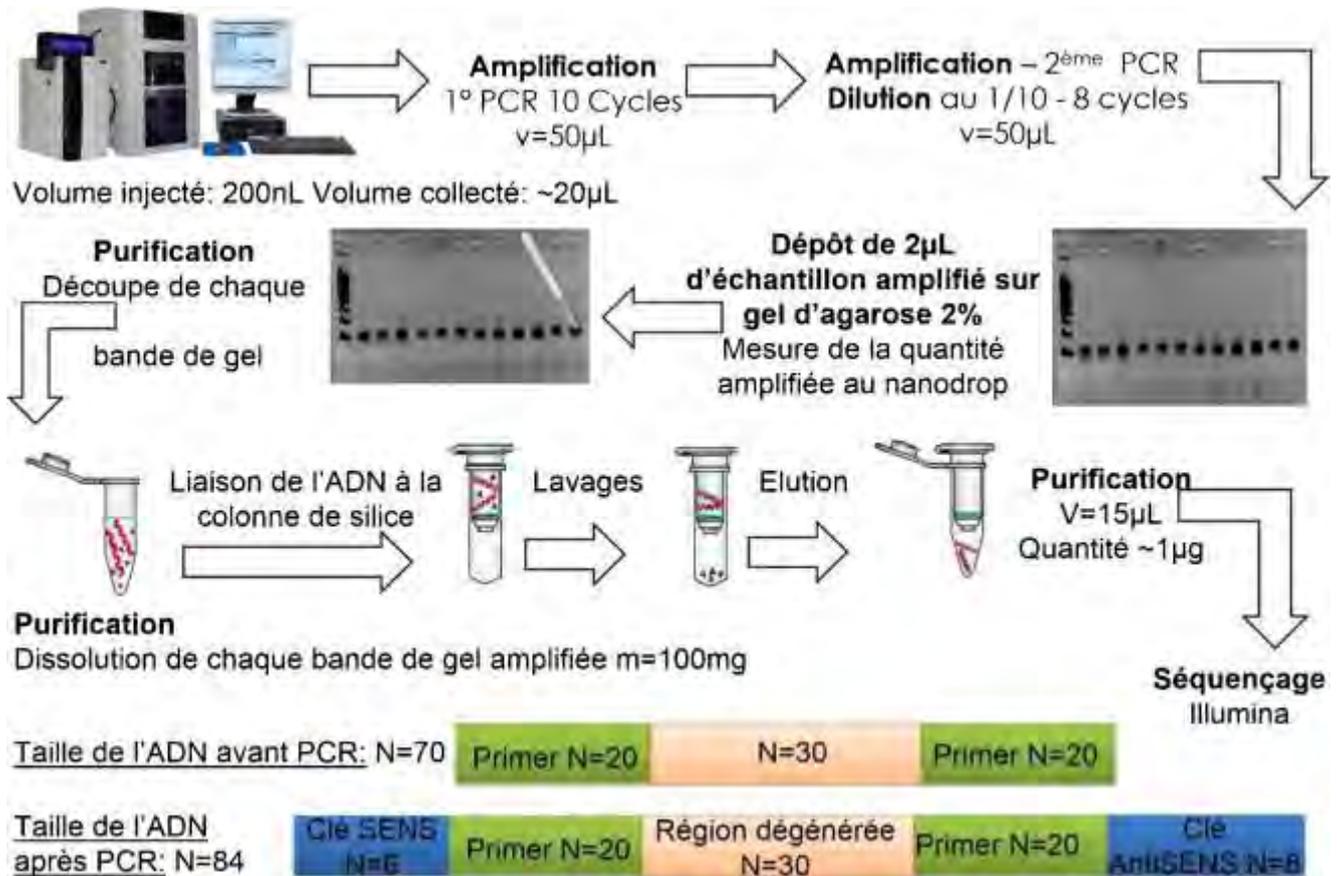


Fig. 1 The analytical strategy used to separate, collect, amplify, and sequence in Illumina with spiked T29₇₀

T29₇₀ sequence matching

Two methods were employed to find ssT29₇₀ in the Illumina sequencing data. The method `vmatchPattern` was used to search a complete set of sequences from a sample for T29-like sequences with an allowance of up to three mismatches. Alternatively, pairwise alignment was used with `type = "local-global"` to obtain an alignment score of T29₇₀ versus each Illumina read.

To trim the whole data set by searching the forward single strand of T29₇₀, only one computer command is required, namely, the `grep` command. In Linux, this command is commonly used to search for a word in a file and to display the lines in which that word is found. The command is easy to use, and by adding special characters called regular expressions, one can improve the precision of the search. With the sequences containing the keys to be searched, this command can be easily applied to search for the sequence of aptamer T29₇₀. For example, `Grep '(^ forward primer \| $ reverse primer \)' filename | Grep -c AGTCACCCCAACCTGCCCTACCAC` translates as follows: a file is searched for sequences that contain the sense key at the beginning of the line (^) and the antisense keys at the end of the line (\$). Once these data are sorted, a pipe (|) is used to consolidate the two commands. In our case, the result of the left com-

mand will go into the pipe, while the right command will extract

the data for processing. Thus, we search the key-sorted data with the `grep` command for sequences that contain the T29₇₀ sequence, and then we count the number of occurrences of the T29₇₀ sequence with the `-c` option (count).

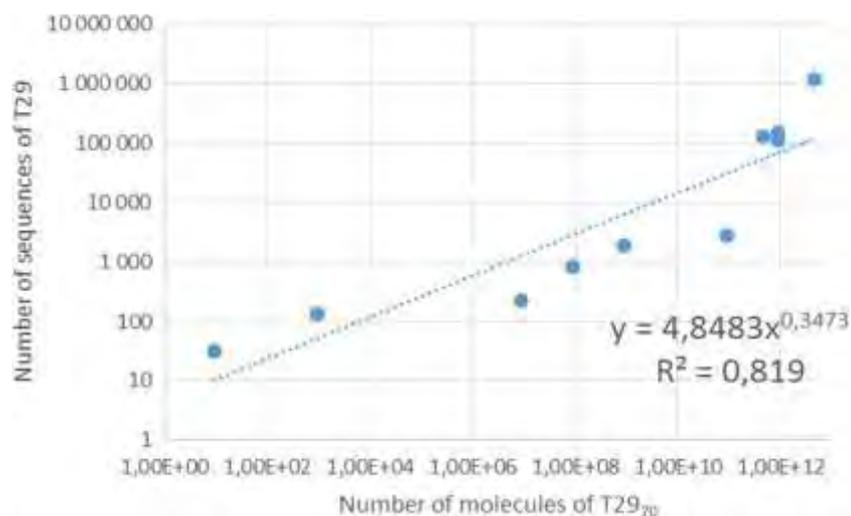
Results and discussion

We began this study by analyzing the correlation between the number of molecules used for the Illumina sequencing and the number of reads obtained by this high-throughput sequencing method. In these first experiments, we did not use CE.

The relationship between known quantities of T29₇₀ and the number of sequences detected in Illumina

Our aim was to estimate the proportionality between the number of molecules of the aptamer (Q_{apt}) and the number of sequences (Seq) obtained in Illumina sequencing in this first step without using CE collections. We spiked the DNA library (always at the same concentration) with known quantities of T29₇₀ (each quantity corresponding to a specific barcode) varying from 900 molecules to 4.5×10^{12} molecules and then used Illumina sequencing. In Fig. 2, we see that the correlation between Q_{apt} and Seq is not

Fig. 2 Relationship between the number of molecules of T29 and the number of detected sequences for pair end T29₇₀



linear but is related by a power of 1/3, $\text{Log}(\text{Seq}) = 4.8 \text{Log}(\text{Qapt})^{0.35}$. On the basis of two different Illumina sequencing runs, we have a coefficient of determination of $R^2 = 0.82$. Why is the value of R^2 between the T29₇₀ quantities and the number of sequences unusually low? When we identify T29₇₀ by the barcode sequences or by the T29₇₀ sequences, we obtain different numbers. In all cases, increasing the amount of T29₇₀ in the samples increased the number of sequences obtained (Table 1). If the PCR amplification and then the Illumina sequencing had no influence on the amplification and sequencing of the T29, the same number should be obtained by both approaches. This outcome could indicate that DNA polymerase makes important modifications to the sequence of our G-quadruplex-containing aptamer.

Systematic errors in the sequences of T29₇₀

Table 1 reports the number of reads of T29₇₀ sequenced in Illumina with the use of different quantities. As we reported previously, we observed that for each quantity of T29₇₀, the number of reads containing the corresponding barcodes is higher than the number of reads containing the exact sequence of T29₇₀. Illumina sequencing has been shown to have fewer than 0.1% errors in sequencing [17], and such high error for a

small ssDNA length is surprising. Because this quantity-dependent study was performed with T29 synthesized by PCR, PCR errors could be present in the amplified T29₇₀ molecules. Therefore, we wanted to examine the different sequence errors that can be indirectly caused by the G-quadruplex. Figure 3 shows the ratio ($R\%$) of the number of reads containing the correct base at position N of the T29 molecule divided by the number of reads containing the correct base at position $N + 1$.

$$R\% = \frac{\text{reads } N_i \text{ p1} = \delta \text{nb reads } N_i \text{ p} \times 100}{\text{reads } N_i \text{ p} \times 100}$$

$R\%$ for each base should be near 99%. Instead, we see that in all the experiments, $R\%$ is below 90% for specific positions. These base positions are highlighted in Fig. 3B. The most important error occurred at the 3' terminus, where 60% mismatches are recorded for the last C. These errors can arise either from the PCR runs or from the Illumina sequencing. Examining the sequences of the antisense strand revealed mismatches at the same position in the sense strand. This result indicates that PCR induces a high number of mismatches in G-quadruplex DNA. Otherwise, the G-quadruplex nucleotides would not be more affected by the mismatches. The main

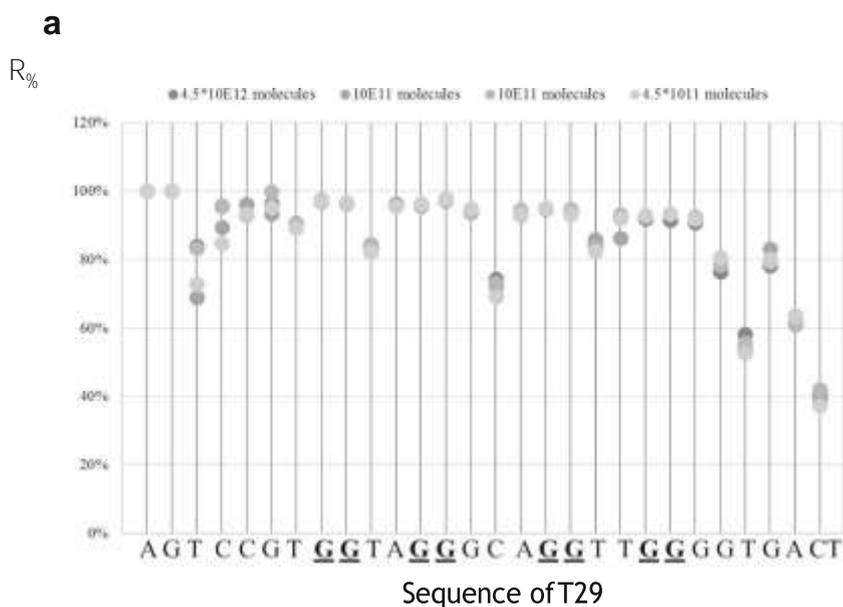
Table 1 T29 quantities used in the Illumina samples, 5' barcodes used to identify the different samples, and number of reads obtained in Illumina containing the barcodes or the exact T29 sequence, or the pair end, i.e., the δ LT29 (<3 mismatches)

Barcode	CGTGAT	ACAT	GATC	ATTG	CTGA	TCAA	CTGA	CGTG	CTGA	AAGC
Number of molecules of T29 ₇₀	4.5E + 12	CG 9E + 11	TG 9E + 11	GC 4.5E + 11	TC 9E + 10	GT 9E + 08	TC 9E + 07	AT 9E + 06	TC 9E + 02	TA 9E + 00
Number of reads containing the barcode	716,228	98,061	113,041	103,775	2251	1496	634	183	82	30
Number of ssT29 sequences	628,335	84,309	92,328	85,431	1837	1256	514	150	68	24
Number of reads corresponding	1,218,588	153,467	113,262	127,676	2735	1895	822	231	136	31

to a δ LT29 (< 3 mismatches)

The 3' barcode was always the same one for all these experiments

Fig. 3 Systematic error for the T29 sequences using Illumina. A $R\%$ following the different nucleotides of the sequence of T29 (from the T29₇₀) at different quantities sent for Illumina sequencing. The measurements were done on data of four different quantities of T29₇₀ and two samples at 9×10^{11} molecules are the same quantities sent for two different Illumina sequencing runs. B A scheme of the tertiary structure of the T29 in which the red cycles present the most mismatched nucleotides



base mismatches occur at the 3' terminus of the T29 DNA, immediately after the G-quadruplex, and at the C14 between the two G nucleotides of the quadruplex.

Overall, to fit the quantities of T29₇₀ to the number of sequences, we used a different approach. We preferred to be more selective and use the number of reads of T29 that were complementary to the T29 antisense. As a result, we counted the dsT29 with a pair end counting method that allowed a

maximum of three mismatches.

Separation by CE-UV-LIF of T29, the {T29.Tbn} complex, and Tbn

Figure 4 presents the UV absorption (for Tbn detection) and LEDIF (for T29 detection) traces in CE after the mixing of Tbn (500 nM) with T29₇₀ (500 nM). The complex was well separated from the free DNA. For this experiment, we added 5 mbar of pressure to decrease the migration times of the different species and to further

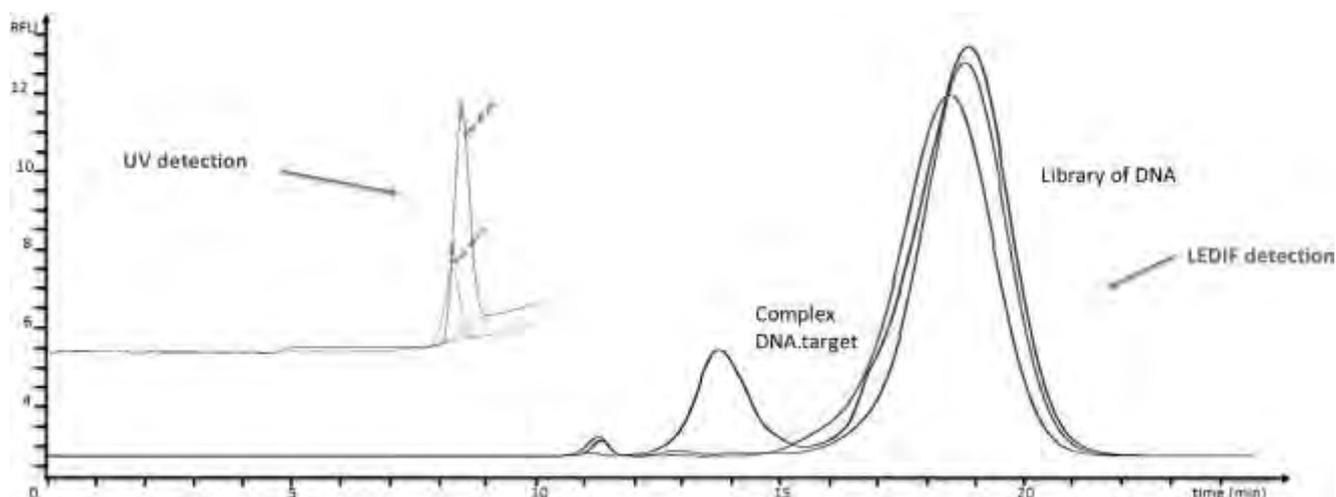


Fig. 4 Electropherograms obtained in UV for Tbn and LIF for the L30 alone (200 nM, red), the T29₇₀ (200 nM, green) alone, and T29₇₀ (200 nM) with Tbn (50 nM) (black)

decrease the collection times of the complex. By decreasing the migration time using pressure, we increased the {T29.Tbn} peak height (data not shown). This effect was certainly due to the lower influence of the k_{off} of the complex because of the shorter migration time. The apparent migration time (t_{ap}) of the protein in the fluorescence electropherogram was calculated with Eq. 1.

$$t_{ap} \approx l_{ef} = l_{LIF} \times t^{UV} \approx 0,45 \times t^{UV} \quad \delta 1 \text{p}$$

where t^{UV} is the migration time of protein with UV detector, l_{ef} effective length for UV detection, 8.5 cm, and l_{LIF} effective length for LIF detection, 19 cm.

In this case, $t_{ap} = 7.6$ min. This migration time was shorter than that of the {T29.Tbn} complex.

Collection of the fractions containing the {T29.Tbn} complex and determination of the number of T29 sequences

In these experiments, we spiked the L30 (10 μM) with different amounts of T29₇₀ from 3 to 10⁸ molecules. We performed 10 collections of each sample. As shown in Fig. 1, the fraction collection was conducted in vials containing 15 μL of phosphate buffer. We injected 200 nL of the T29₇₀/Tbn (10 nM) mixture into the 100- μm i.d. capillary. This injected volume and the 10 injections of each sample statistically determined a minimum amount of three molecules of T29₇₀ for the most diluted spiked sample. After a migration time of 15.6 min, we collected the fractions over 2 min using the same migration conditions.

Relationship between known amounts of T29₇₀ molecules injected into CE and the number of paired-end sequences of T29

In this experiment, the main problem was the high diversity of reads obtained for each sample, which were identified by barcodes corresponding to the {T29.Tbn} concentration. The insert in Fig. 5 shows the number of reads corresponding to the injection of different quantities of T29₇₀ into CE. The number of sequences was not proportional to the T29₇₀ quantity. However, for each run, the injected L30 quantity was always the same. Therefore, we used the enrichment factor f_E introduced by Riley et al. [13]. Figure 5 presents the correlation between the numbers of molecules injected in CE and the enrichment factor f_E . The correlation curve has the equation $f_E = 10^{-9} (\text{paired-end sequences})^{0.92}$. The value 0.92 is close to 1, which shows a good correlation between the number of T29₇₀ molecules spiked into the DNA library and the number subjected to CE separation. This result shows that by using this NECEEM approach along with Illumina sequencing, we can obtain the sequence of an aptamer containing a G-quadruplex. This result confirms that the method proposed 10 years ago by Berezovski and Krylov [1, 2] is universal and can be used for G-quadruplex aptamer selection.

The K_D measurement of T29

To validate our NECEEM experiment, we measured the K_D of Tbn and T29 as described by Kanoatov et al. [18]. The equilibrium of an aptamer with its target is governed by Eq. 2:

The collected samples were treated as described in the experimental section.

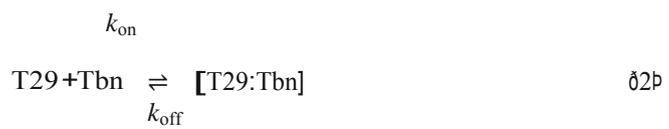
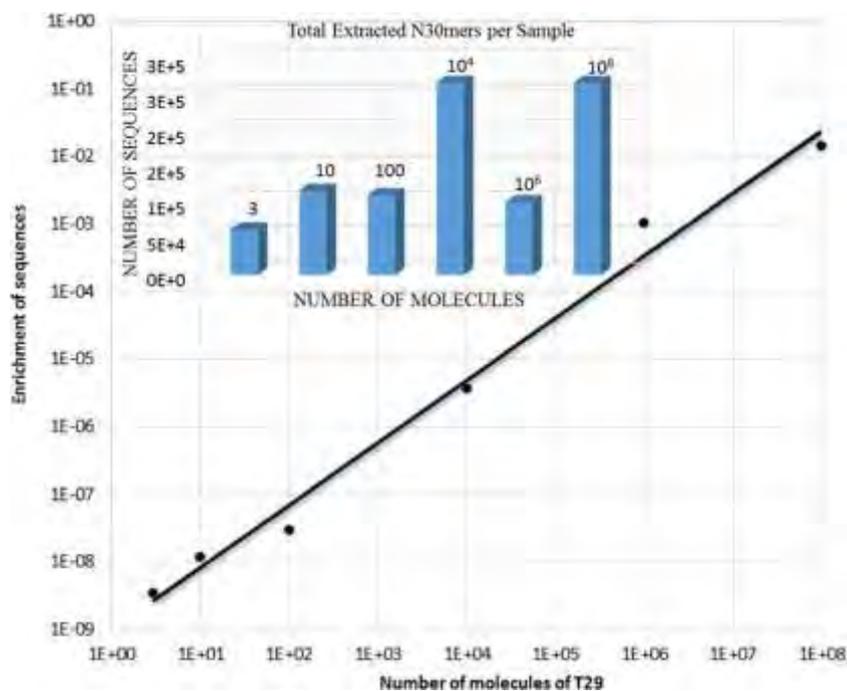


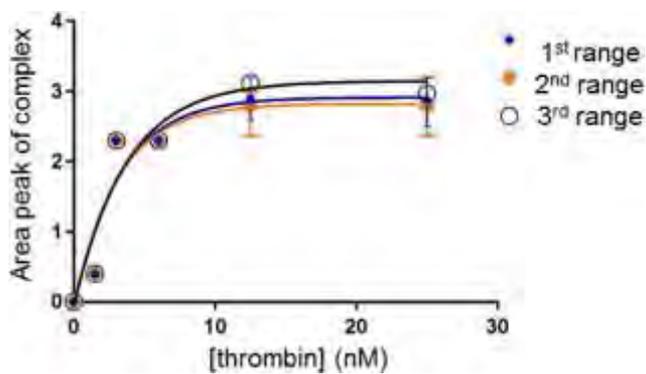
Fig. 5 Correlation between the enrichment factor f_E and the numbers of T29₇₀ molecules injected in CE. In the inset, the number of reads obtained for the different quantities of T29₇₀ injected in CE



where $K_D = k_{off} / k_{on}$.

We tested a standard range of concentrations from 0 to 1 μ M Tbn with a constant concentration of 25 nM for the aptamer T29. This concentration was chosen based on the results of [17], in which the authors claimed that when the K_D value is not known precisely, the concentration of the aptamer must be kept as low as possible so that the measured areas of {T29.Tbn} remain reliable. Three experiments were carried out to measure the repeatability of the experiment. The area of the complex peak was measured for each concentration of thrombin, and we plotted the following curve: Area of the {T29.Tbn} peak = f [thrombin]. Figure 6 presents the results. The deduced value of K_D was 2.62 ± 0.83 nM, which is very close to the value (0.5 nM) found by Tasset et al. [11]. For comparison with the values found in previous studies, Riley

Fig. 6 The fit between the area of the {T29.Tbn} peak and the concentration of Tbn. The $K_D = 2.62 \pm 0.83$ nM



et al. [13] found a value of 128 nM using their tITP in CE experiment, while Li et al. [19] found values of approximately 200 nM. Even if the sample preparation and separation conditions for this study were not identical to those of previous studies (i.e., no alkaline ion in the tITP study, while K^+ ions are present in the study by Li et al.), the K_D values have similar ranges.

Using filter binding, Bock et al. [20] studied different sequences containing a similar G-quadruplex. These authors found K_D to be approximately 200 nM (data not shown), which was very far from the value determined by Tasset et al. [11]. The very large differences in K_D (ranging from

1. to 200 nM) measured for the same interaction indicate the dependency of the measurements on the analytical conditions. For method validation, the community must determine standard conditions to minimize such discrepancies. We compared the materials and methods of these articles in search of useful ideas (Table 2). Li et al. [19, 21] postulated that the differences in the K_D values reflect the dissociation/ association equilibrium under the strong electric field of the CE. Riley et al. did not compare their values with those of the nitrocellulose filter bindings and reported good agreement of these K_D values with the literature, providing evidence that the use of SYBR Gold as a fluorescent probe and cITP does not interfere with the binding between thrombin and its known aptamers.

In analyzing the published works, we observed important differences in K_D values between the filter binding experiments and the CE experiments, independently of the field. In these last cases, low salt concentrations did not result in

2. Krylov SN, Berezovski M. Non-equilibrium capillary electrophoresis of equilibrium mixtures—appreciation of kinetics in capillary electrophoresis. *Analyst*. 2003;128:571–5.
3. Ashley J, Ji K, Li SF. Selection of bovine catalase aptamers using non-SELEX. *Electrophoresis*. 2012;33:2783–9.
4. Yufa R, Krylova SM, Bruce C, Bagg EA, Schofield CJ, Krylov SN. Emulsion PCR significantly improves nonequilibrium capillary electrophoresis of equilibrium mixtures-based aptamer selection: allowing for efficient and rapid selection of aptamer to unmodified ABH2 protein. *Anal Chem*. 2015;87:1411–9.
5. Yu X, Yu Y. A mathematical analysis of the selective enrichment of NECEEM-based non-SELEX. *Appl Biochem Biotechnol*. 2014;173:2019–27.
6. Gold L. The SELEX process: a surprising source of therapeutic and diagnostic compounds. *Harvey Lect*. 1995-1996;91:47–57.
7. Luo Z, Zhou H, Jiang H, Ou H, Li X, Zhang L. Development of a fraction collection approach in capillary electrophoresis SELEX for aptamer selection. *Analyst*. 2015;140:2664–70.
8. Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol*. 2015;33:877–81.
9. Kasahara Y, Irisawa Y, Fujita H, Yahara A, Ozaki H, Obika S, et al. Capillary electrophoresis-systematic evolution of ligands by exponential enrichment selection of base- and sugar-modified DNA aptamers: target binding dominated by 2'-O,4'-C-methylene-bridged/locked nucleic acid primer. *Anal Chem*. 2013;85:4961–7.
10. Deng B, Lin Y, Wang C, Li F, Wang Z, Zhang H, et al. Aptamer binding assays for proteins: the thrombin example—a review. *Anal Chim Acta*. 2014;837:1–15.
11. Tasset DM, Kubik MF, Steiner W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol*. 1997;272:688–98.
12. Wang J, Gu Y, Liu L, Wang C, Wang J, Ding S, et al. Novel application of fluorescence coupled capillary electrophoresis to resolve the interaction between the G-quadruplex aptamer and thrombin. *J Sep Sci*. 2017; <https://doi.org/10.1002/jssc.201700456>.
13. Riley KR, Saito S, Gagliano J, Colyer CL. Facilitating aptamer selection and collection by capillary transient isotachopheresis with laser-induced fluorescence detection. *J Chromatogr A*. 2014;1368: 183–9.
14. Riley KR, Gagliano J, Xiao J, Libby K, Saito S, Yu G, et al. Combining capillary electrophoresis and next-generation sequencing for aptamer selection. *Anal Bioanal Chem*. 2015;407:1527–32.
15. Hüge BJ, Flaherty RJ, Dada OO, Dovichi NJ. Capillary electrophoresis coupled with automated fraction collection. *Talanta*. 2014;130: 288–93.
16. Ric A, Ong-Meang V, Poinot V, Martins-Froment N, Chauvet F, Boutonnet A, et al. ssDNA degradation along capillary electrophoresis process using a Tris buffer. *Electrophoresis*. 2017;38:1624–31.
17. Glenn TC. Field guide to next-generation DNA sequencers. *Molec Ecol Res*. 2011;11:759–69.
18. Kanoatov M, Mehrabanfar S, Krylov SN. Systematic approach to optimization of experimental conditions in nonequilibrium capillary electrophoresis of equilibrium mixtures. *Anal Chem*. 2016;88:9300–8.
19. Li Y, Guo L, Zhang F, Zhang Z, Tang J, Xie J. High-sensitive determination of human alpha-thrombin by its 29-mer aptamer in affinity probe capillary electrophoresis. *Electrophoresis*. 2008;29: 2570–7.
20. Bock LC, Griffin LC, Latham JA, Vermaas EH, Toole JJ. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*. 1992;355:564–6.
21. Li H Y, Deng QP, Zhang DW, Zhou YL, Zhang XX. Chemiluminescently labeled aptamers as the affinity probe for interaction analysis by capillary electrophoresis. *Electrophoresis*. 2010;31:2452–60.

2.5 Précision sur l'analyse des nucléotides de T29 dans le brin anti-sens :

Pour éliminer l'hypothèse selon laquelle ce serait le G-quadruplex qui serait à l'origine de la difficulté de séquençage, nous avons analysé de la même façon le brin anti-sens.

Dans un premier temps, nous avons compté le nombre de nucléotides correspondants à la séquence du T29. Contrairement au brin sens, ou environ 80 % des séquences correspondaient au T29, on s'aperçoit que dans le brin anti-sens, nous perdons la moitié des séquences dès le 2^{ème} nucléotide. Par exemple, pour le point à 10⁵ molécules, nous avons 330 358 séquences contenant le début de la séquence de T29 A, mais dès que l'on cherche la suite A-G, le nombre de séquences chute à 129 820 séquences. Cette différence est vraie pour le début de la séquence car malgré tout, nous obtenons 91 514 séquences de T29.

Lorsque nous regardons la proportion de nucléotides dans la séquence du T29, nous remarquons également un fort taux de déchets de séquences supérieur à celui du brin sens (Figure 4-14).

3 LA CONTAMINATION DES BLANCS :

Maîtrisant, les informations sur le nombre d'échantillons maximal pouvant être séquençé, le nombre de cycles PCR optimal, nous avons toujours une contrainte : les blancs pour lesquels un grand nombre de séquences est obtenu. L'origine de la contamination n'étant pas identifiée, plusieurs blancs ont été séquençés : l'un avant la réalisation de la collecte en CE, l'autre à la fin des collectes en CE et enfin un dernier lors du dépôt sur gel des échantillons (Figure 3-18).

blc avt CE	blc post CE	blc Eau	BLANCS		Pureté		Nb de cycles PCR
			Nombre de séquences		Avant CE	Après CE	
10+8cycles			1 ^{er} séquençage	100 000		1*10 ⁻⁴	25
Blancs: Avant CE / Après CE / gel			2 ^{ème} séquençage	350 000		0,11	18
			3 ^{ème} séquençage	Avant CE	Après CE	Avant CE	Après CE
				8 900	10 500	0,90	0,90

Figure 3-18 : dépôt sur gel après amplification par PCR, des blancs avant CE, après CE et après dépôt sur gel d'agarose 2%. Le but étant de séquençer ces trois échantillons afin d'identifier l'origine de la contamination.

Cependant, malgré l'absence de bande visible sur le gel et les précautions prises, nous observons toujours une contamination estimée à 25nM. Le blanc du gel n'est pas représenté car il n'a pas pu être identifié lors de l'analyse bio-informatique.

Nous avons également séquençé des « blancs » en prenant certaines précautions comme changer le tampon de la cuve de migration des gels entre la migration sur gel après PCR des échantillons (Figure 3-22). Nous avons réalisé 3 « blancs », un blanc avant d'injecter nos échantillons en électrophorèse capillaire, un blanc à la fin des analyses en électrophorèse capillaire afin de mesurer l'effet du capillaire sur la contamination en AT29B et un blanc lors

de la migration des échantillons sur gel d'agarose 2% après PCR. Le but est de situer à quel moment se produit la contamination par l'aptamère T29. Ainsi, on peut voir que la contamination se produit lors des injections par électrophorèse capillaire car le blanc réalisé lors du dépôt des échantillons sur gel ne révèle qu'un faible nombre de séquences (420). Ceci est surprenant, car lors de chaque collecte, l'électrophorèse capillaire et les électrodes sont nettoyées avec précaution.

BLANCS	Nombre de séquences		Pureté		Nb de cycles PCR		
1 ^{er} séquençage	350 000		0,11		18 ou 25		
2 ^{ème} séquençage	50 000		0,11		18		
3 ^{ème} séquençage	Avant CE	Après CE	Avant CE	Après CE	18		
	8 900	10 500	0,90	0,90			
4 ^{ème} séquençage	Avant CE	Après CE	EAU	Avant CE	Après CE	EAU	18
	15 251	31 461	470	0,77	0,78	0,21	



Figure 3-22 : Blancs séquencés lors du 4^{ème} séquençage. 3 échantillons dont l'un avant les collectes par électrophorèse capillaire, l'autre après les collectes et le dernier "eau" correspondant au dépôt sur gel ont été réalisés. La contamination a lieu lors de l'électrophorèse capillaire, puisque le blanc « eau » ne contient qu'un très faible nombre de séquences.

Néanmoins, nous avons vérifié nos électrophérogrammes (Figure 3-23), qui ne montrent pas de pic de contamination. Nous n'observons que du bruit de fond.

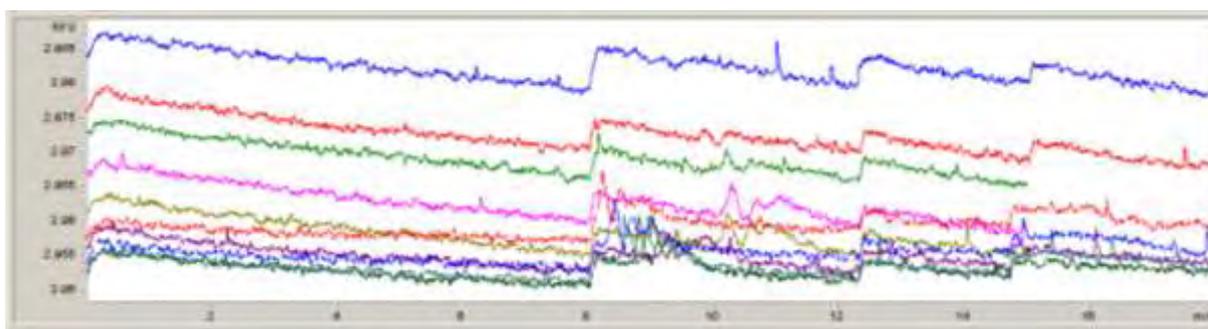


Figure 3-23 : électrophérogramme des "blancs" lors du 2^{ème} séquençage

La contamination est peut-être due à une erreur d'amplification soit lors de la réaction PCR ou lors du séquençage qui utilise aussi une réaction PCR. Comme l'analyse au fragment analyzer l'indique, il y a eu une amplification de produits non-spécifique lors de ce dernier séquençage

4 MESURE DE L'ENRICHISSEMENT DE L'APTAMERE T29 PAR ELECTROPHORESE CAPILLAIRE COUPLEE AU SEQUENÇAGE ILLUMINA

Enfin, nous avons réalisé une gamme étalon d'aptamère AT29B seul qui a été amplifié par PCR afin de vérifier la reproductibilité de l'expérience. Nous avons également mesuré l'effet de l'électrophorèse capillaire sur l'enrichissement des séquences et pour cela, nous avons réalisé une gamme étalon d'un mélange de banque d'ADN à 10 μ M, de thrombine à 1 μ M et différentes molécules d'aptamères à 10⁶, 10⁸ et 10¹⁰ molécules (Figure 3-24). Sans électrophorèse capillaire, nous remarquons que la pureté ne varie pas alors que le nombre de molécules d'aptamères T29 augmente. Avec électrophorèse capillaire, pour une même complexité d'échantillons, la pureté varie d'un facteur 100 entre 10⁷ molécules et 10⁹ molécules. Mais l'enrichissement augmente avec électrophorèse capillaire pour une même pureté d'échantillon (pour 10¹⁰ molécules,

l'enrichissement passe de 1.6 à 34 avec l'électrophorèse capillaire ce qui montre bien l'intérêt de cette dernière dans le processus de sélection).

• Sans CE:

AT29B	Nb de séquences d'AT29B	Nb de séquences totales	Complexité	Pureté	Enrichissement
10 ⁵ molécules	1 894	70 045	1,7*10 ⁻⁵	0,03	16 000
10 ⁶ molécules	3 093	114 458	1,7*10 ⁻⁴	0,03	160
10 ¹⁰ molécules	1 626	64 178	1,7*10 ⁻³	0,025	1,6

• Avec CE:

	Nombre d'AT29B	Nb de seq AT29B	Nb de seq totales	Pureté	Enrichissement
Complexité 1,7*10 ⁻⁴	10 ⁷ molécules	251	460 000	5*10 ⁻⁴	300
Complexité 1,7*10 ⁻⁴	10 ⁷ molécules	19 256	379 000	0,05	300
Complexité 1,7*10 ⁻²	10 ¹⁰ molécules	46 186	81 165	0,57	34

Figure 3-24 : Mesure de l'enrichissement dû à l'électrophorèse capillaire.

5 CONCLUSION :

Le protocole de détection d'un aptamère lié à la thrombine et la double détection UV-LIF, permettent de déterminer précisément les temps de migration de chacun des composés afin de pouvoir collecter les fractions d'intérêts. Nous avons émis l'hypothèse que collecter deux fractions l'une avant la migration du complexe (la fraction A) et l'autre correspondant à la fraction du complexe (fraction B) aurait permis de distinguer des séquences d'aptamères non visibles par électrophorèse capillaire mais présentant une affinité plus importante pour la cible. Or, cette expérience réalisée lors du premier séquençage n'a pas été concluante car la diversité des échantillons envoyés à séquencer (environ 40 échantillons) était trop importante ce qui en a résulté un faible nombre de séquences permettant de tirer une information de ces données. La contamination par les blancs interpelle sur l'origine d'une telle contamination, qui est invisible sur les électrophérogrammes, n'a pas lieu durant l'étape de PCR et pourtant si l'on corrèle le nombre de séquences des blancs au nombre de séquences d'aptamères T29, cela correspond à une contamination de l'ordre du nanomolaire. Néanmoins, nous montrons que l'électrophorèse capillaire permet d'enrichir les échantillons d'un facteur 20, mais face aux difficultés rencontrées lors du dernier séquençage nous ne pouvons affirmer ce résultat. Nous avons pu comparer les différentes gammes étalons d'aptamères de T29 sans CE et avec CE afin de savoir si nous pouvions détecter un aptamère structuré en G-quartet par notre méthode. Ces résultats seront l'objet du chapitre suivant.

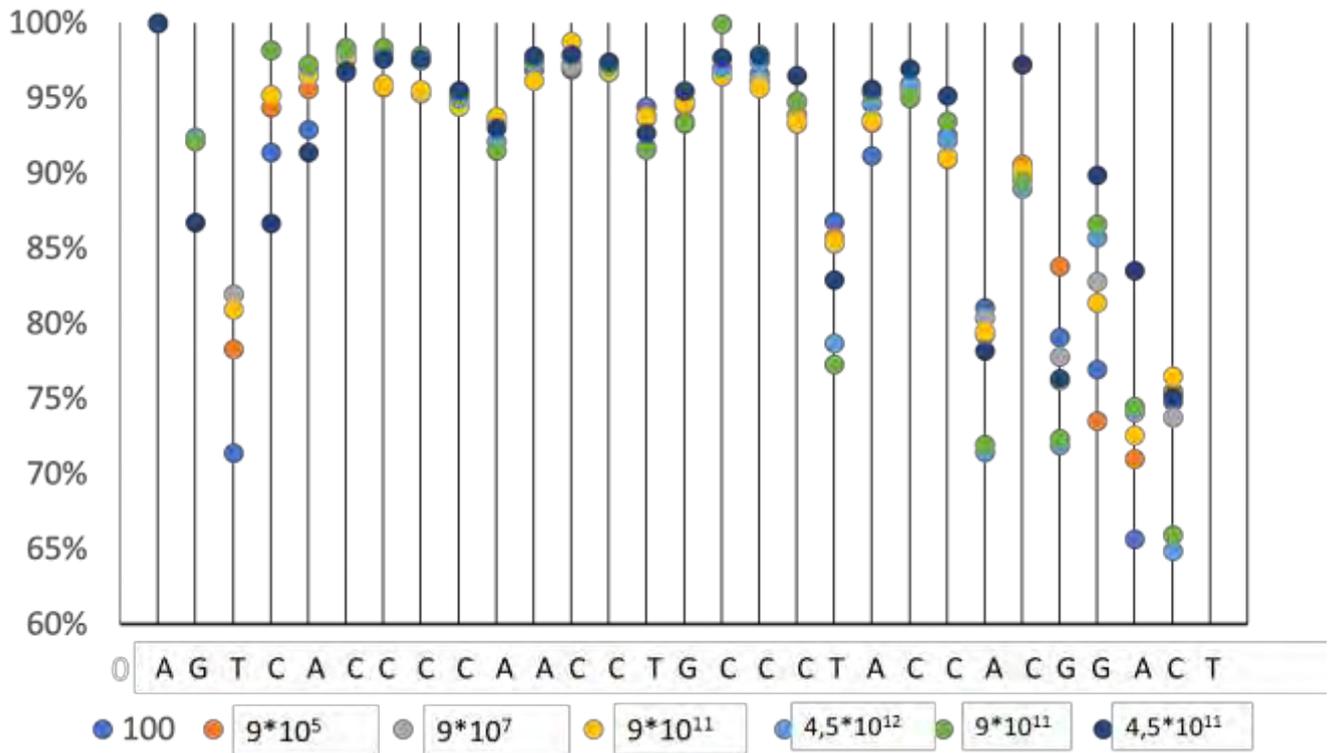


Figure 4-14 : Abondance de chaque nucléotide dans la séquence du T29 dans le brin anti-sens.

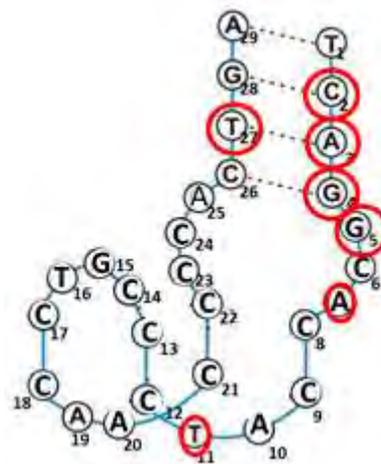


Figure 4-15 : Représentation de la structure de l'aptamère T29 dans le brin anti-sens

On remarque que ce sont toujours les nucléotides en fin de séquences du T29 qui sont affectés par un taux d'erreur de 40% (Figure 4-15). Hormis, les nucléotides en position 2 et 3, le reste de la séquence et particulièrement les nucléotides C n'ont pas de taux d'erreurs inférieur à 80%.

Conclusion

Les différents séquençages nous ont montré qu'il existait une limite de variabilité d'échantillons à analyser s'il l'on voulait pouvoir faire une analyse en profondeur des séquences. Lors du 1^{er} séquençage, près de 40 échantillons correspondants à des conditions différentes ont été analysés mais le nombre de séquences obtenu pour chaque condition était très faible du fait de la trop grande diversité des échantillons présents sur la plaque du séquenceur. C'est ainsi que nous avons limité le nombre d'échantillons pour les séquençages suivants à 18.

Cependant, nous nous sommes intéressés à la pureté des séquences de l'aptamère T29 avec les données des trois gammes réalisées. A partir du fichier de données brutes, nous avons directement triés les données en cherchant la séquence du T29 avec les clés sens et anti-sens. Nous ne nous sommes pas intéressés à la complémentarité des séquences par clés. Nous avons regardé la position de chaque nucléotide (Ni) par rapport à la séquence du T29 et nous avons tracé le graphe ci-dessus représentant : $\%_i = \left(\frac{Nb_{N_i}}{Nb_{A_i}} \right) * 100$. Ainsi 80% des séquences contenant la clé sont des séquences de l'aptamère T29. Pour autant, cela signifie que nous avons 20% de déchets ce qui pour un aptamère d'une taille de 29 paires de bases est relativement important. Ces résultats ont également été confirmés dans l'analyse du brin anti-sens. Nous avons également mesuré l'abondance des nucléotides dans la séquence de l'aptamère T29. Cette analyse a permis d'éliminer l'hypothèse selon laquelle ce serait la structure en G-quadruplex qui empêcherait l'amplification et le séquençage correct de l'aptamère T29. Nous remarquons que les tétrades de guanines ne sont pas affectées par une substitution et que leur abondance se situe toujours aux alentours de 90 – 100%. Cependant, les nucléotides en début de séquences ou en fin de séquences ont une abondance très faible (de 50% jusqu'à 20%). On peut donc supposer que la Taq Polymerase amplifie mal ces séquences, puisque ce sont les mêmes nucléotides que l'on retrouve avec le plus faible taux d'abondance dans le brin sens ou anti-sens. Cette dernière gênée par la structuration des nucléotides de guanines sous forme de G-quadruplex peut « sauter » des nucléotides et conduire à une amplification de produits non-spécifiques. Dans les séquences qui ont été injectées par électrophorèse capillaire, ce sont les nucléotides de début de séquences qui ont l'abondance la plus faible. Ces résultats sont à corrélés avec le fait que nous n'avons pas pu mettre au point une amplification efficace par qPCR. Si notre travail se termine sur cette validation de la technique de NECEEM associée à l'Illumina pour les G-quartet, il ouvre la porte en grand sur le traitement des données de l'Illumina. En effet si pour notre travail nous avons voulu savoir si l'on était capable d'identifier un aptamère connu dans une banque d'ADN, nous n'avons pas pu identifier d'autres aptamères, ni tester si les aptamères générés à partir des modifications du T29 avaient des KD proches de celui du T29. Ainsi, nous voyons qu'entre le travail qui était effectué à partir des clonages puis des séquençages de Sanger, un petit nombre de données était accessible et aisément triable par la détermination de KD afin de sélectionner le meilleurs aptamères. Aujourd'hui le travail de traitement des données Illumina impose de développer des outils bio informatiques spécialisés afin de trouver automatiquement des motifs de séquences, qui permettraient de ne tester que certains aptamères utilisés. Mais tout ceci reste à construire dans l'avenir aussi bien pour les techniques utilisées en SELEX ou celles utilisant la CE-non SELEX. Si l'on croit que les aptamères ont un avenir, il s'agit ici d'un vrai enjeu industriel...

Annexes :

1 INSTRUMENTATION

L'ensemble du dispositif qui a été utilisé dans ce travail se compose du système d'électrophorèse capillaire 7100/ 1600 d'Agilent Technologies couplée à un détecteur LEDIF Picometrics.

1.1 Le détecteur de fluorescence

Les détecteurs commercialisés sont équipés d'une bille (**Figure 1-1**), qui permet de focaliser le rayon d'excitation et de récolter le signal d'émission en un même point.

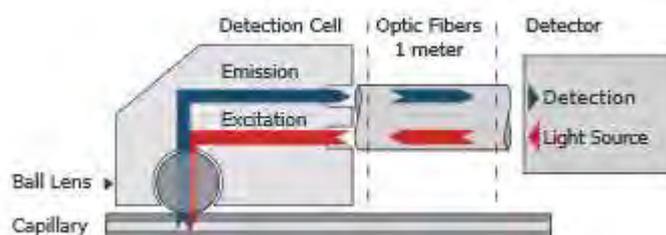


Figure 1- 1 : Schéma de fonctionnement du détecteur ZetaLIF de Picometrics.

Pour cette étude, nous avons utilisé soit un laser 488 nm, soit une LED 480 nm. Ce rayonnement d'excitation est focalisé par la bille qui excite le fluorophore de l'échantillon à son maximum d'absorption. La molécule va fluorescer en retournant à l'état fondamental et la fluorescence sera récoltée par la bille placée le long du capillaire et qui focalise le faisceau sur le photomultiplicateur.

1.2 La double détection absorbance UV - fluorescence

L'analyse des molécules qui ne fluorescent pas est réalisée par absorbance UV. Par exemple, les protéines absorbent dans la zone UV (les aptamères aussi). Dans notre étude, la thrombine a été détectée à 210nm. L'absorbance UV est une méthode de détection intéressante pour l'étude des interactions entre un aptamère et une molécule cible en raison de sa simplicité d'utilisation et de son faible coût. En effet, les protéines et les acides nucléiques ont des bandes d'absorption dans la région de l'UV proche (de 200 à 400 nm). Ces biomolécules peuvent donc être analysées par la spectrophotométrie d'absorption UV-visible.

Grâce à la double détection par absorbance UV et par fluorescence, on peut ainsi détecter simultanément la thrombine, le complexe aptamère-thrombine et l'aptamère libre. Pour réaliser la double détection, il suffit de créer deux fenêtres de détection sur le capillaire et de le positionner dans la cassette (**Figure 1-2**).



Figure 1-2 : Schéma du capillaire avec la double fenêtre UV et LIF ci-dessus et schéma de la cassette modifiée par Picometrics ci-dessous.

Le pilotage de l'appareil et du logiciel chemstation se fait à l'aide d'un ordinateur. Le driver Picometrics permet une intégration complète de tous les aspects du détecteur de fluorescence : le contrôle du détecteur LIF, le stockage des méthodes et de données et un large éventail d'options supplémentaires (Figure 1-3).

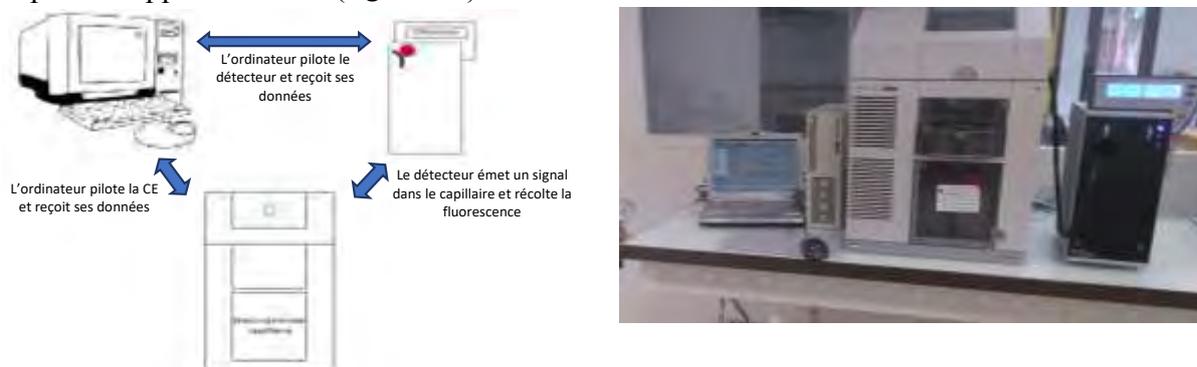


Figure 1-3 : Appareillage complet avec l'électrophorèse capillaire, le détecteur LIF et le PC.

2 REACTIFS ET PRODUITS UTILISES :

2.1 Réactifs :

La plupart des réactifs utilisés proviennent de Sigma Aldrich.

Nom	Fournisseur
Acide acétique	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Phosphate de sodium dibasique anhydre	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Chlorure de sodium	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Acide chlorhydrique	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)

Chlorure de magnésium	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Hydroxyde de sodium	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Persulfate d'ammonium	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Xylène cyanol	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Bromophénol	Sigma-Aldrich - Fluka (St Quentin Fallavier, France)
Urée	Euromedex (Souffelweyersheim, France)
TEMED*	Eurobio (Courtaboeuf, France)
Acrylamide	Fischer Scientific Illkirch, France

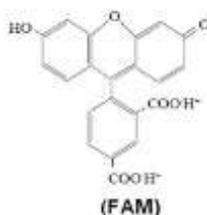
* TEMED: (N, N, N', N' tetramethyl-1-, 2-diaminoethane)

2.2 Oligonucléotides :

La séquence de l'aptamère étudié, le T29 est la suivante :



Cet aptamère est greffé à un fluorophore, le 6-carboxyfluorescein (FAM) en position 5' afin de pouvoir le détecter par fluorescence. (Sigma-Aldrich - Fluka).



La banque d'ADN utilisée est dégénérée sur 30 nucléotides et a pour séquence :



2.3 Tampons :

Les solutions tampons ont été préparées à Picometrics, avec de l'eau ultra-pure et conservée à température ambiante.

Deux types de tampons ont été utilisés en électrophorèse capillaire comme tampon de migration :

- Le tampon Tris-Acétate à 50 mM pH=8.2
- Le tampon Sodium Phosphate dibasique anhydre à 50 mM pH=8.2

Les tampons dans lesquels ont été dilués la thrombine et les échantillons d'ADN ont également été préparés au laboratoire à Picometrics.

Lorsque le tampon de migration était le tampon Tris-Acétate, le tampon de dilution des échantillons était: Tris-HCl 50mM, NaCl 100mM, MgCl₂ 1mM.

Lorsque le tampon de migration était le tampon Sodium Phosphate dibasique anhydre, le tampon de dilution était : Na₂HPO₄ 50mM, NaCl 100mM, MgCl₂ 1mM.

3 CAPILLAIRES :

Capillaire en Silice :

Un capillaire de silice de Polymicro a été employé : il est recouvert d'une gaine de polyimide excepté aux niveaux des fenêtres de détection.

Les fenêtres de détection sont réalisées en enlevant le polyimide au niveau de l'endroit où l'on place le détecteur c'est-à-dire à 8,5cm pour la fenêtre UV et 19cm pour la fenêtre LIF grâce au brûle capillaire et on nettoie délicatement la fenêtre avec de l'éthanol. On vérifie ensuite la propreté de la fenêtre à l'aide d'une loupe binoculaire.

3.1 Capillaire en Poly-vinyl-alcool : PVA

Un autre capillaire de silice recouvert de poly-vinyl-alcool (PVA) a été utilisé lors de la mise au point de la méthode de détection de la thrombine car celle-ci s'adsorbait sur la surface du capillaire en silice. Il existe de nombreuses gammes de capillaires recouverts d'un polymère à la surface des silanols. Le PVA est celui qui est le plus efficace pour éliminer les interactions entre la thrombine et la surface du capillaire. Avec ce capillaire le flux électroosmotique est supprimé et on a une plus grande efficacité de séparation. La fenêtre de détection du PVA ne peut être réalisée avec le brûle capillaire car on risquerait de dégrader le PVA. Elle est réalisée en faisant une fenêtre à l'acide, c'est-à-dire que l'on place le capillaire sur une plaquette de plastique à l'endroit où l'on veut faire la fenêtre de détection et l'ensemble est mis sur une plaque chauffante. On verse ensuite une goutte d'acide sulfurique sur la fenêtre en veillant à avoir une tenue de protection (lunettes, gants, blouse) et en activant la hotte. On nettoie ensuite la fenêtre du capillaire au méthanol.

3.2 Conditionnement du capillaire :

On vérifie tout d'abord que le capillaire n'est pas cassé ou bouché en injectant de l'eau pendant 3 minutes, et en vérifiant que la goutte tombe bien dans le vial de récupération. Ensuite, on effectue un test de courant en lavant le capillaire avec le tampon pendant 3 minutes puis on place deux vials de tampons aux extrémités du capillaire et on applique une tension. Le courant doit rester constant durant la manipulation. Une fois ces paramètres réglés, on réalise un focus, c'est-à-dire que l'on vérifie que le détecteur fonctionne bien, en injectant une solution fluorescente témoin et en mesurant la valeur de fluorescence. On peut aussi modifier la valeur du photomultiplicateur pour avoir une meilleure sensibilité et un maximum de fluorescence.

4 PRINCIPE DE LA PCR EN TEMPS REEL :

4.1 Conditions opératoires

Les quantités d'ADN collectées par électrophorèse capillaire sont de l'ordre de 100-200ng. Or, le séquençage Illumina nécessite 2µg d'ADN double brin. Il faut donc amplifier la quantité d'échantillons d'ADN collectés en électrophorèse capillaire.

Deux méthodes d'amplification ont été expérimentées, dont la PCR en temps réel. L'objectif est de pouvoir suivre en continu le processus d'amplification PCR en détectant la fluorescence émise par les produits de PCR néo formés.

Différents kits ont été utilisés :

Nom du kit	Description du kit	Préparation du Mix de réaction	Programme qPCR
Takyon SYBR MasterMix Blue	<ul style="list-style-type: none"> - Takyon DNA Polymerase - MgCl₂ 2.5mM - SYBR green - dNTPs - Inert blue dye - Stabilizers 	<ul style="list-style-type: none"> - Mix Takyon 10μL - Clé sens 2μL (100nM) - Clé Anti-sens 2μL (100nM) - Eau 3.5μL 	<ul style="list-style-type: none"> - 50°C – 2 min X { - 95°C – 3 min - 95°C – 3 secs - 60°C – 25secs X= 40 cycles
Takyon SYBR MasterMix Blue + Uracil-N-Glycosylase (UNG) / dUTP Mix	<ul style="list-style-type: none"> - Takyon DNA Polymerase - MgCl₂ 2.5mM - SYBR green - dNTPs - Inert blue dye - Stabilizers - dUTP 2.5μL 	<ul style="list-style-type: none"> - Mix Takyon 10μL - Clé sens 2μL (100nM) - Clé Anti-sens 2μL (100nM) - Eau 3.5μL 	<ul style="list-style-type: none"> - 50°C – 2 min X { - 95°C – 3 min - 95°C – 3 secs - 60°C – 25 secs X= 40 cycles
Sso Fast Evagreen Supermix Biorad	<ul style="list-style-type: none"> - 2X reaction buffer with dNTPs - Sso7d-fusion polymerase - MgCl₂ 2.5mM - EvaGreen dye - Stabilizers 	<ul style="list-style-type: none"> - 50μL Mix - Clé sens 10μL (500nM) - Clé Anti-sens 10μL (500nM) - Eau 10μL 	<ul style="list-style-type: none"> - 95°C – (1min ou 30 secs) X { - 95°C – 10 secs - 60°C – 15 secs - 65°C → 95°C X= 35 – 45 cycles

Un kit de qPCR est principalement composé de l'enzyme la Taq Polymérase nécessaire à l'amplification du brin d'ADN à amplifier, de désoxyribonucléotides (dNTPs), ainsi qu'une sonde fluorescente (SYBR green) qui se fixe sur l'ADN double brin et ne fluoresce qu'une fois fixée à ce dernier. (Tableau ci-dessus : description du kit)

Avant la réaction, on mélange ces réactifs avec l'ADN à amplifier, ainsi qu'avec des amorces spécifiques du fragment d'ADN à amplifier. (Tableau ci-dessus : Préparation du Mix de réaction).

L'ensemble est déposé dans un petit tube, lui-même introduit dans un thermocycleur qui programme des cycles de températures. Cet appareil réalise des cycles de trois températures différentes pendant des durées différentes. Ces étapes sont :

- La dénaturation (**Figure 1-4**) : Elle se produit à 95°C et durant cette étape l'ADN double brin est séparé en deux simples brins.

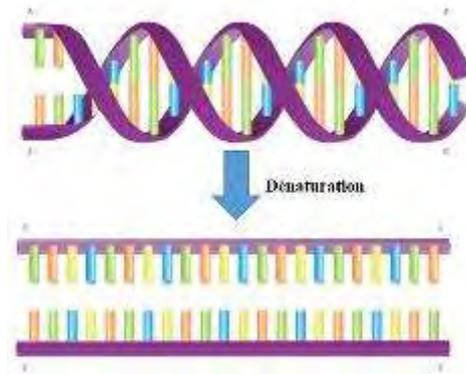


Figure 4-1

Figure 1-4 : Dénaturation de l'ADN. Durant la dénaturation, le SYBR Green I libre exhibe peu de fluorescence.

- L'hybridation (**Figure 1-5**) : La réaction descend aux alentours de 55°C-60°C (dépend du T_m des amorces- en général $T_m-5^\circ\text{C}$) et les amorces ou clés se fixent sur chaque brin d'ADN. La réaction est initiée par l'enzyme Taq Polymérase qui commence à incorporer les dNTPS.



Figure 1-5 : Hybridation de l'ADN. À la température d'appariement, quelques molécules se lient au double brin d'ADN naissant résultant en une émission de fluorescence lors de l'excitation.

- L'élongation (**Figure 1-6**) : À chaque étape d'hybridation et d'élongation d'un cycle PCR, le SYBR-Green s'intercale entre les bases nucléotidiques de l'ADN double brin. Cette molécule émet très peu de fluorescence à l'état libre, tandis qu'une fois incorporée dans l'ADN il émet fortement à 520 nm. Cette capacité de se fixer à l'ADN conduit à la limite de cette méthode car il y a un risque de détecter des dimères d'amorces ou une amplification non spécifique. Pour vérifier la spécificité d'amplification une courbe de fusion est réalisée après le dernier cycle de PCR. La fluorescence est mesurée à la fin de chaque étape d'élongation de chaque cycle, l'augmentation de fluorescence traduit donc la formation d'amplicons. La mesure de l'intensité du signal émis à la fin de chaque étape d'élongation permet le suivi cycle par cycle de la réaction PCR.



Figure 1-6 : élongation de l'ADN. Durant la phase de polymérisation, de plus en plus de molécules se lient au brin naissant et l'accroissement de la fluorescence est suivi en temps réel.

L'intensité de la fluorescence émise par l'échantillon amplifié augmente à chaque cycle PCR. Il est alors possible de mesurer la quantité de produits d'amplification générée lors de la phase exponentielle en traçant la courbe suivante (**Figure 1-7**) :

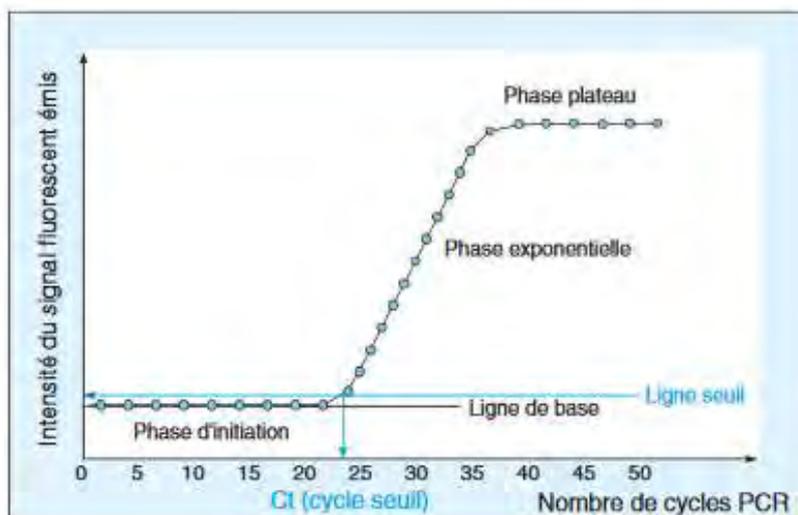


Figure 1-7 : Suivi en temps réel d'une réaction de qPCR. En ordonnée, on trouve l'intensité de fluorescence et en ordonnée le nombre de cycle PCR. La ligne seuil correspond au seuil de détection optique au-delà duquel la variation de fluorescence suit une loi exponentielle. Le point d'intersection de la courbe avec la ligne seuil désigne le Ct (cycle seuil), c'est-à-dire, le point de départ de la phase exponentielle qui est directement lié à la quantité initiale de cible présente dans l'échantillon d'après (151).

Le cycle seuil correspond au nombre de cycles à partir duquel la quantité de produit formé est distinct du "bruit de fond". Ce Ct marque le début de la phase exponentielle d'amplification (= phase reproductible non limitante).

Le Ct est inversement proportionnel au logarithme de base 2 du nombre de copies initiales présentes de la cible dans le milieu réactionnel : plus le Ct d'un échantillon est petit, plus le nombre de copies initiales de la cible est important.

4.2 Utilisation de la PCR en temps réel pour amplifier l'aptamère T29 :

L'intérêt de la qPCR pour ce projet était de pouvoir quantifier directement l'aptamère T29 collecté par électrophorèse capillaire. Pour ce faire, nous souhaitons établir une gamme étalon de la matrice AT29B à différentes concentrations (100nM \rightarrow 10⁻³fM).

Réalisation d'une gamme étalon de l'aptamère AT29B

Une gamme avec 7 points étalons d'aptamères AT29B a été préparée dans un premier temps : 10nM, 1nM, 0.1nM, 10pM, 1pM, 0.1pM, et un point de contrôle ne contenant que le mix, les amorces et la matrice étant remplacée par de l'eau (**Figure 1-8**). Ce point de contrôle est appelé « blanc ». Cette gamme a été dupliquée afin d'évaluer la répétabilité des résultats.

Le programme de températures utilisé fut le suivant :

Dénaturation : 95°C – 1 min.

Hybridation : 55°C – 5 secs.

Elongation : 65°C – 5 secs.

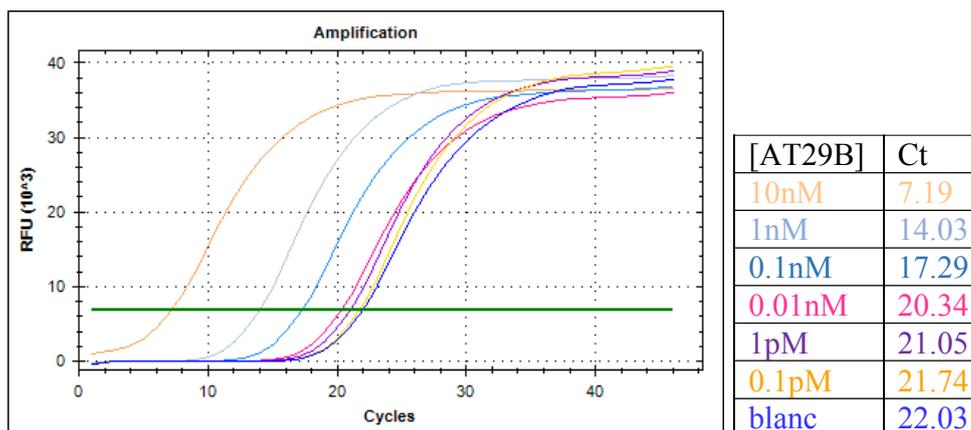


Figure 1-8 : Quantification absolue – détermination du nombre de Ct (nombre de cycles requis pour qu’une réaction de fluorescence atteigne le seuil de référence).

La courbe d’amplification obtenue nous montre que pour les plus fortes concentrations (1nM, 0,1nM, 0,01nM), nous avons à peu près une relation linéaire entre la quantité de molécules cibles présentes dans l’échantillon et le Ct obtenu puisqu’entre les points à 1nM, 0.1nM et 0.01nM on a un écart de 3 Ct entre chaque concentration. Le Ct est la valeur de cycle correspondant au signal seuil. Il est appelé cycle threshold ou cycle seuil.

Chaque échantillon a ensuite été déposé sur gel d’agarose 2% (Figure 1-9).

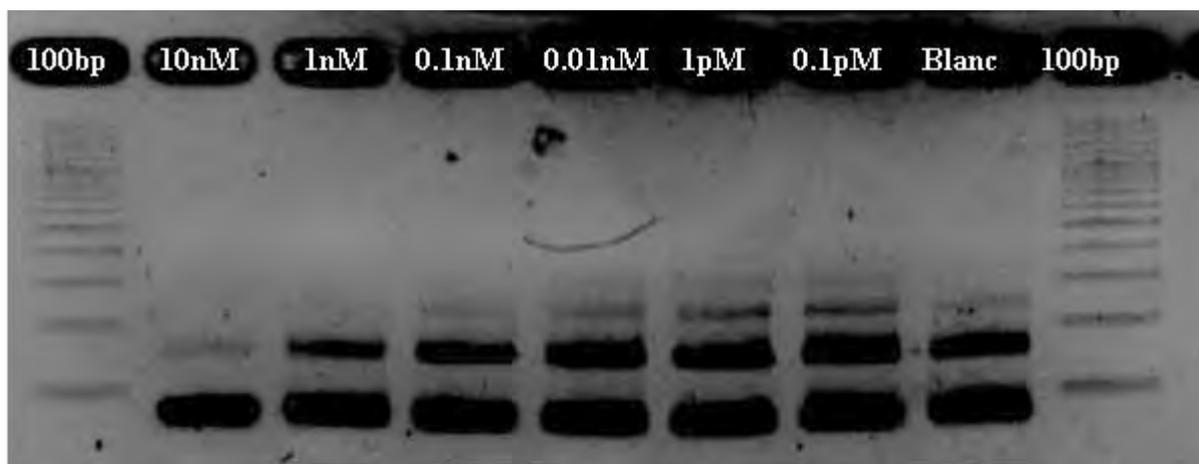


Figure 1-9 : Gel d’agarose 2% réalisé en complément de la qPCR.

De part et d’autre du gel a été déposé un marqueur qui permet d’estimer la taille de l’ADN. La bande la plus loin du point de dépôt correspond à une bande d’ADN de taille 100 paires de bases. Donc comme le produit amplifié a une taille de 84 paires de bases, nous nous attendons à voir les échantillons migrer à la même hauteur que le marqueur d’ADN. Or, on note ici l’apparition de sous-produits de PCR, très visible pour la gamme de 1nM à 0.01pM. Par ailleurs, le contrôle blanc ne contenant que de l’eau indique qu’il y a eu une contamination puisque nous observons les mêmes sous-produits que les échantillons de T29 amplifiés.

4.2.1 Répétabilité de la qPCR

Une répétabilité de cette expérience a été réalisée afin de confirmer ou infirmer l’hypothèse d’une contamination en aptamère T29 qui aurait pollué tous nos échantillons. Afin d’améliorer le procédé, nous avons réalisé une deuxième expérience de qPCR avec une gamme étalon allant

de 100nM à 0.01pM et en utilisant le même programme de température que précédemment (**Figure 1-10**).

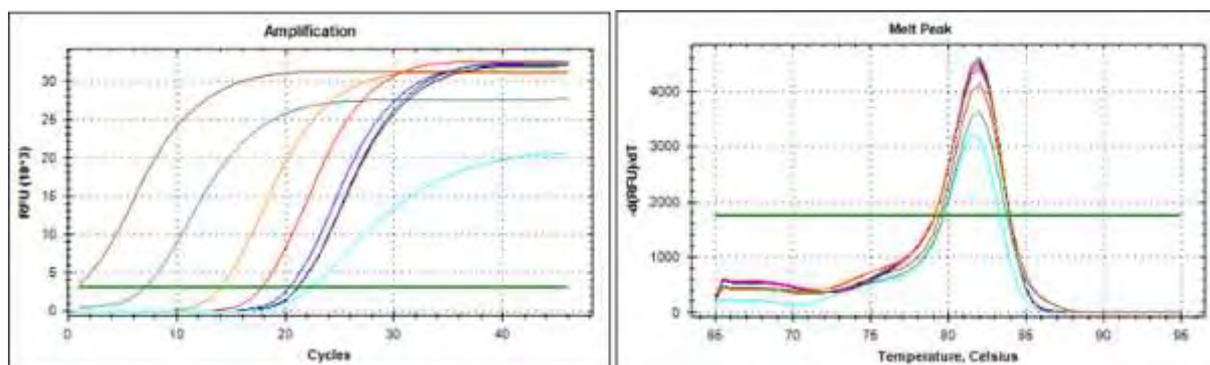


Figure 1-10 : Courbe d'amplification et courbe de melting T_m - température à laquelle la moitié de l'ADN est sous forme double-brin et l'autre sous forme simple brin.

Pour les deux expériences, nous obtenons des températures de fusion de $T_m = 82^\circ\text{C}$. Le T_m est la température à laquelle la moitié brins d'ADN formés par la réaction PCR se dissocie. Cette température est en lien directe avec la composition de l'ADN formé à étudier. Ainsi, un amplicon aura un T_m différent selon sa composition en base et sa longueur essentiellement. Nous observons un seul T_m ce qui indique qu'à la fin de la réaction de PCR, un seul produit de PCR est formé. Cependant, la valeur du T_m observée est différente de la valeur attendue. D'après les données du fournisseur, l'aptamère AT29B a un T_m de $94,6^\circ\text{C}$. Donc, on peut supposer que si le T_m obtenu est plus bas que la valeur théorique, cette valeur étant dépendante de la composition des amorces cela signifie peut-être que l'on a une mauvaise amplification de l'aptamère due à une mauvaise hybridation de l'enzyme sur les amorces.

[AT29B]	Ct
100nM	/
10nM	7.13
1nM	14.10
0,1nM	17.89
0,01nM	20.39
1pM	21.17
0,1pM	21.18
0,01pM	21.22
Blanc	22.53

A partir de la concentration d'1pM, nous atteignons le seuil de la qPCR, c'est-à-dire que l'on obtient le même nombre de Ct quel que soit la concentration de l'aptamère utilisé à partir d'1 pM et ceci jusqu'à 0.01pM.

Chaque échantillon amplifié a également été déposé sur gel d'agarose 2% afin de visualiser le résultat de l'amplification de l'aptamère AT29B (**Figure 1-11**).

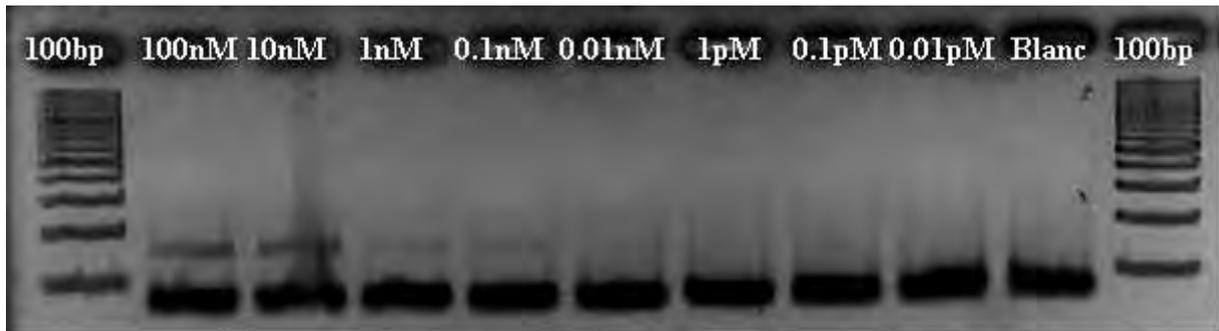


Figure 1-11 : Dépôt sur gel d'agarose 2% de la gamme de répétabilité de l'aptamère AT29B amplifiée par qPCR.

Inversement, nous obtenons pour les fortes concentrations, des sous-produits de PCR. Ceci concerne notamment les points à 100nM, 10nM, 1nM et 0,1nM. Plusieurs hypothèses ont été émises pour justifier l'apparition de ce smear :

- L'augmentation du nombre de cycles de qPCR implique la formation de sous-produits. Il ne faut pas augmenter les cycles PCR au-delà de 45 cycles, c'est la limite expérimentale au-delà il n'y a plus d'amplification.
- L'enzyme Taq Polymérase n'a pas le temps de voir les substrats et/ou n'a pas le temps de les amplifier car la quantité de matrice à amplifier est trop importante.
- Il y a un excès d'amorces qui entraînent un effet compétiteur. Elles ne peuvent pas voir la matrice pour les faibles concentrations et donc on a la même allure pour les Ct. Pour les fortes concentrations, on observe une amplification non-spécifique.

Les « blancs » considérés comme les contrôles de la qPCR présentent un Ct élevé équivalent à une contamination d'AT29B aux alentours du picomolaire.

4.2.2 Diminution du nombre de cycles qPCR.

En conservant le même programme de températures, nous avons réalisé deux expériences en réalisant 25 cycles (**Figure 1-12**) et 35 cycles (**Figure 1-13**) de qPCR.

4.2.2.1 Cas de 25 cycles

[AT29B]	Ct
1nM	14,59
0.1nM	18,68
0.01nM	20,31
1pM	20,67
0.1pM	17,61
0.01pM	20,63
Blanc	22,53

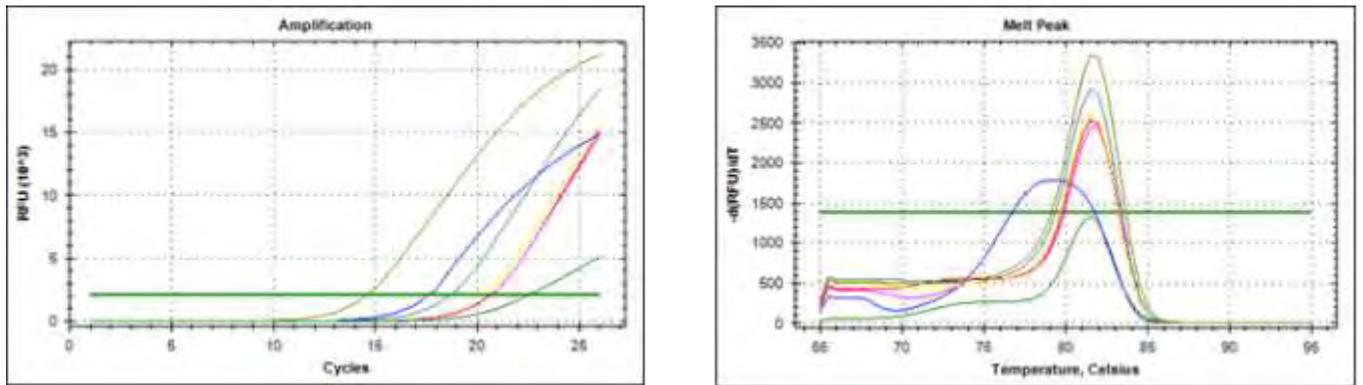


Figure 1-12 : En haut à gauche, courbe de quantification absolue - détermination du nombre de Ct en réalisant 25 cycles de qPCR. En haut à droite, courbe de melting (T_m) obtenues lors de l'analyse qPCR - 25 cycles. La courbe bleue correspondant au point 0.01pM a un T_m de 79°C. Les autres points ont un T_m de 82°C. Le tableau ci-dessus donne les valeurs des Ct pour chaque valeur.

En réalisant 25 cycles de qPCR, le plateau n'est pas atteint. Cette dernière étape correspond à la fin de la réaction, c'est-à-dire que les constituants de la PCR deviennent limitants et donc la réaction s'achève.

4.2.2.2 Cas de 35 cycles

[AT29B]	Ct
1nM	11,24
0.1nM	15,78
0.01nM	16,74
1pM	17,51
0.1pM	17,37
0.01pM	17,39
Blanc	18,09

A partir de 35 cycles, on commence à voir apparaître la phase de plateau. Pour autant, nous n'observons pas une linéarité entre tous les points de la gamme. Le nombre de cycles de qPCR idéal est donc celui compris entre 25 et 45 cycles. De plus, le blanc est toujours contaminé.

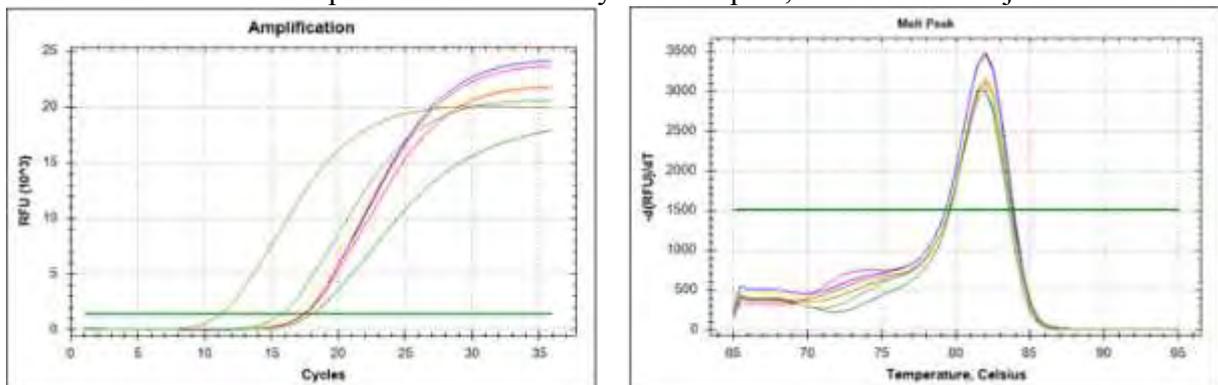


Figure 1-13 : En haut à gauche, courbe de quantification absolue - détermination du nombre de Ct en réalisant 35 cycles de qPCR. En haut à droite, courbe de melting ($T_m=82^\circ\text{C}$) obtenues lors de l'analyse qPCR - 35 cycles.

4.2.3 Effet des amorces – diminution de la concentration des amorces :

Pour vérifier que les amorces en excès n'étaient pas responsables de la mauvaise linéarité de l'amplification, nous avons décidé de diminuer la concentration des amorces de 500nM à 100nM.

En réalisant 45 cycles de qPCR, nous obtenons la courbe d'amplification suivante ainsi qu'un $T_m=82^\circ\text{C}$ (Figure 1-14) :

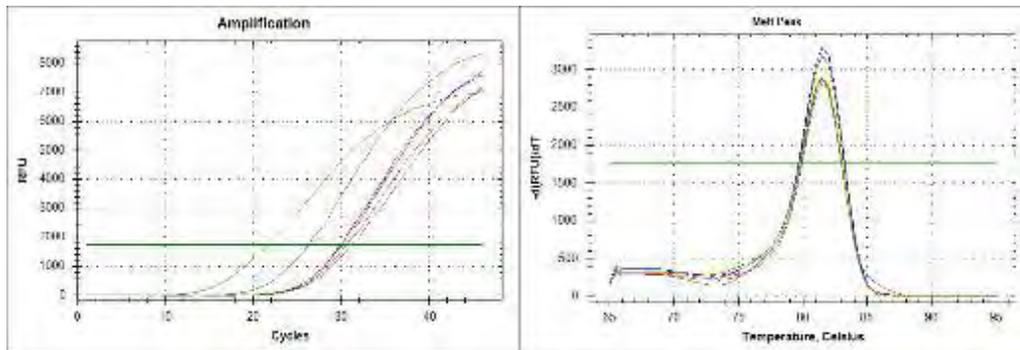


Figure 1-14 : En haut à gauche, courbe de quantification absolue - détermination du nombre de Ct en réalisant 45 cycles de qPCR en diminuant la concentration des amorces de 500nM à 100nM.

[AT29B]	Ct
1nM	15.01
0.1nM	19.07
0.01nM	20.24
1pM	20.72
0.1pM	20.47
0.01pM	20.87
Blanc	20.62

En diminuant la concentration des amorces, nous n'atteignons pas la phase de plateau indiquant que les produits de PCR ont été formés. La valeur de Ct pour le premier point de la gamme à 0.1nM est de 19.07 ($C_t=14$ pour le point à 1nM à 45 cycles) mais cela ne suffit pas pour que la fluorescence atteigne le seuil de fluorescence déterminant le nombre de molécules amplifiées.

4.2.4 Augmentation de la température d'hybridation.

En conservant le couple d'amorces et le nombre de cycles à 45, nous avons supposé qu'en ayant une température d'hybridation de 55°C , la température d'hybridation était trop basse, et les amorces risquaient de s'hybrider non seulement à la position choisie mais aussi de manière non spécifique avec d'autres régions d'ADN de séquence proche de celle qui a été sélectionnée. On risquait alors une amplification de fragments d'ADN non-spécifique. En effet, à une température aussi basse d'hybridation, la stringence étant moins forte il peut y avoir une formation de structures secondaires empêchant la réaction de PCR. Donc nous avons réalisé une nouvelle gamme en partant de $1\mu\text{M}$ jusqu'à 0.01pM et en augmentant la température

d'hybridation à 62°C. Cette gamme a été dupliquée afin d'évaluer la répétabilité (Figure 1-15).

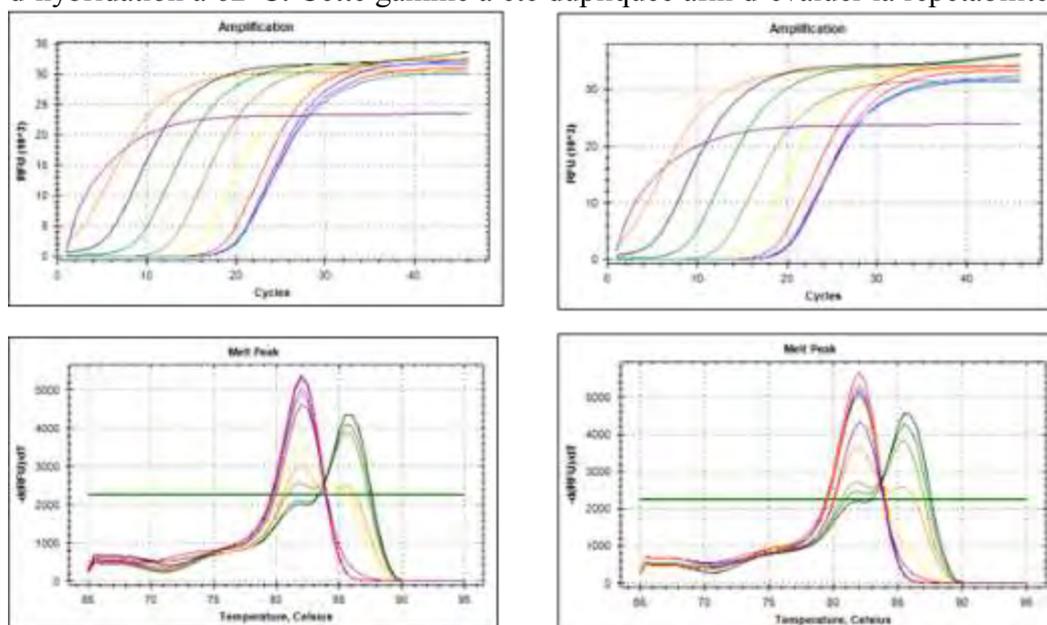


Figure 1-15 : gamme étalons 1μM - 0.1μM - 10nM – 1nM – 0.1nM – 0.01nM – 1pM – 0.1pM – 0.01pM – blanc amplifiée par qPCR en augmentant la température d'hybridation à 62°C. Deux gammes ont été réalisées dans les mêmes conditions afin d'évaluer la répétabilité des expériences.

[AT29B]	Ct
1μM	1,63
0,1μM	3,08
10nM	5,59
1nM	8,48
0,1nM	12,09
0,01nM	15,23
1pM	18,30
0,1pM	19,46
0,01pM	19,63
blanc	19,59

[AT29B]	Ct
1μM	1,60
0,1μM	3,00
10nM	5,45
1nM	8,45
0,1nM	11,97
0,01nM	15,43
1pM	18,30
0,1pM	19,36
0,01pM	19,79
blanc	19,79

On observe une linéarité de la gamme pour les valeurs de Ct jusqu'au point 1pM puisque nous obtenons des valeurs similaires pour les deux gammes. Au-delà, tous les points à 0.1pM, 0.01pM ainsi que le blanc ont un Ct de 19. On peut donc attribuer une contamination de 0.1pM au blanc puisqu'il a la même valeur de Ct que le point à 1pM.

De plus, pour certains points de la gamme, nous observons 2 T_m indiquant la formation de sous-produits que nous n'avions observés précédemment que sur gel d'agarose. Cela concerne les points à 10nM, 1nM et 0.1M qui ont un T_m de 85,5°C.

Paradoxalement, les points les plus concentrés de la gamme 1μM, 0.1μM, et les points les moins concentrés de la gamme 0.01nM, 1pM, 0.1pM, 0.01pM ont le même T_m de 82°C.

La réaction de qPCR est une réaction en chaîne par la répétition des trois cycles dénaturation – hybridation – élongation. Ceci aboutit à une amplification exponentielle de 2ⁿ fois par molécule d'ADN. A chaque cycle d'amplification la quantité d'ADN double selon la relation suivante :

$$N = N_0 * 2^n$$

Où N : nombre de produits PCR présents au Ct ;

N_0 le nombre produit PCR initial et n le nombre de cycles d'amplification.

On définit l'efficacité d'amplification (E_{ff}) comme étant l'efficacité de la réaction de qPCR. Si l'efficacité de la réaction est de 100%, cela signifie qu'à chaque cycle, 2^n molécules ont été générées. Cette valeur est généralement comprise entre 78 et 97% selon le gène amplifié. Elle est caractérisée par l'équation suivante :

$$N = N_0 * (1 + E_{eff})^n$$

Sous forme logarithmique, on l'a simplifiée selon cette expression :

$$\log N = \log N_0 + n * \log(1 + E_{eff})$$

La représentation graphique des variations du logarithme du nombre N de molécules amplifiées (Log N) en fonction du nombre de cycles d'amplification (n) est une droite dont la pente est $\log(1 + E)$ et l'ordonnée à l'origine le nombre initial de molécules (N_0).

Connaissant la valeur de la pente, il est alors possible de déterminer l'efficacité de la réaction : $E = 10^{-1/pente}$.

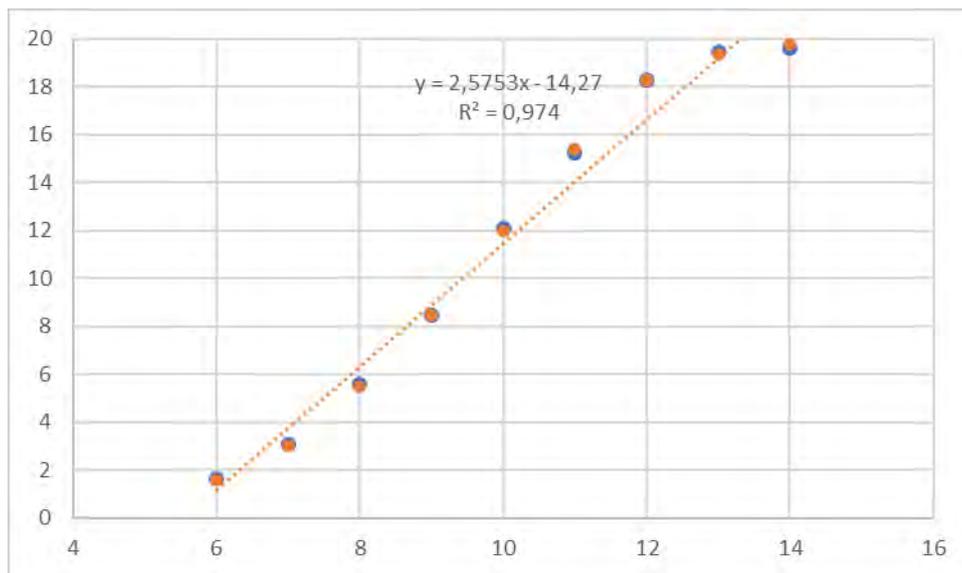


Figure 1-16 : Tracé de la courbe - $\log[C_0] = Ct$

En utilisant $p=2.5753$, on trouve une valeur d'efficacité de la réaction de qPCR de 41%, ce qui montre bien que nous n'arrivons pas à amplifier un aptamère. D'autre part, la courbe $\log C_0 = Ct$ (Figure 1-16) ne nous donne pas une droite linéaire mais une droite avec un plateau vers 20 Ct.

4.3 Changement des constituants :

4.3.1 Changement des amorces :

On peut estimer la contamination des blancs aux alentours du pM. L'analyse au fragment analyser de tous les échantillons de qPCR nous indique une taille de produits d'amplifions de 80 paires de bases. Cela correspond à la taille du T29 (Figure 1-17).

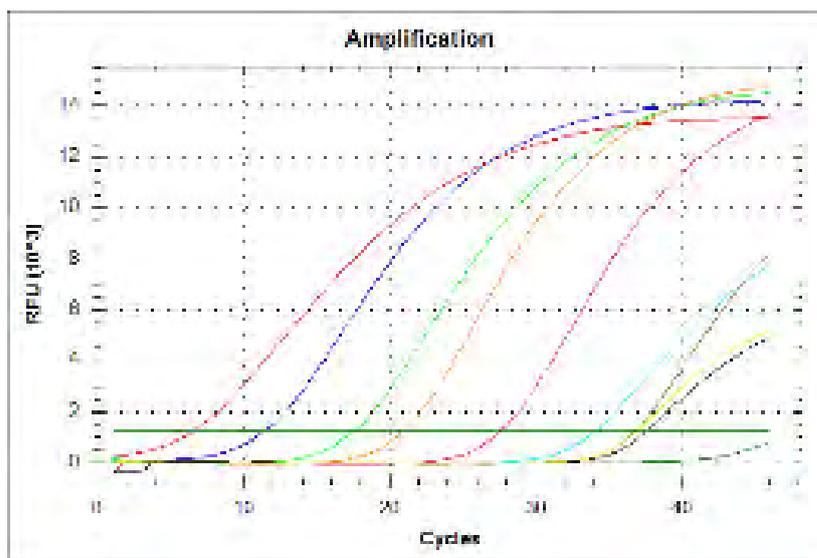


Figure 1-19 : Réalisation d'une qPCR sur un aptamère ne se structurant pas sous forme de G-quadruplex

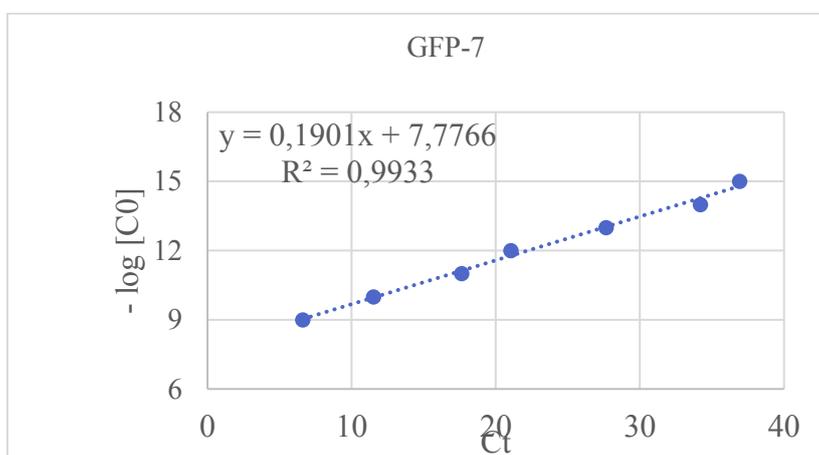


Figure 1-20 : Tracé de la courbe $-\log[C0] = Ct$

Contrairement à l'aptamère T29, nous obtenons une bonne linéarité de notre gamme (Figure 1-20). De plus, le blanc est négatif.

4.4 Conclusion

En diminuant le nombre de cycles qPCR, nous pensions diminuer l'amplification non-spécifique de l'aptamère AT29B et ainsi obtenir une linéarité pour la gamme étalon. Il s'avère que la diminution/ augmentation du nombre de cycles qPCR tout comme la diminution de la concentration des amorces conduisent à diminuer l'amplification des molécules cibles ce qui se traduit par un nombre de cycles d'amplifications plus élevé pour atteindre la phase de plateau. Or, dans les deux cas, cette phase n'est pas atteinte ce qui montre qu'il faut conserver un nombre de cycles de 45 et une concentration d'amorces à 500nM. Par ailleurs, nous n'avons pas testé différents couples d'amorces en qPCR afin de vérifier leurs spécificités. Les courbes de fusion obtenues nous ont permis de vérifier qu'il n'y a qu'un « seul produit amplifié » et donc que les couples d'amorces utilisés sont bien spécifiques. Nous n'avons pas réussi à mettre au point une gamme étalon de l'aptamère T29 par qPCR. De plus, l'origine de la contamination n'a pas pu être établie. Nous supposons que le T29 doit être présent en quantité suffisante (de l'ordre du picomolaire) dans la pièce où nous réalisons les qPCR. Seuls l'utilisation d'un kit d'UTP, où la

manipulation très précautionneuse (utilisation de gants – cônes à filtres – matériel dédié à chaque pièce et poste de travail) permettrait d'éliminer cette contamination.



D'autre part, nous pensons que l'amplification non-linéaire de l'aptamère T29 est dû à sa structure puisque dès lors que nous utilisons un autre aptamère, les blancs sont négatifs et la gamme étalon est linéaire avec un coefficient de corrélation de $R^2=0,99$.

5 PRINCIPE DE LA PCR CONVENTIONNELLE :

Le principe de la PCR conventionnelle est sensiblement identique à celui de la PCR en temps réel. La différence entre ces deux méthodes réside dans le fait que l'on ne peut pas suivre en temps réel la quantité de matrices amplifiées

Les réactifs utilisés étaient :

- 29 μ L Eau
- 5 μ L Tampon 10 X
- 5 μ L Amorces sens 5 μ M
- 5 μ L Amorces anti-sens 5 μ M
- 5 μ L dNTPs 2,5mM
- 1 μ L Taq Polymerase
- 1 μ L de matrice ADN

Le programme de température utilisé était :

- 10 cycles
- Cycle 1 : 95°C pendant 3 min (dénaturation)
 - Cycle 2 : 95°C pendant 30 secs
 - Cycle 3 : 55°C pendant 30 secs (hybridation)
 - Cycle 4 : 72°C pendant 30 secs (élongation)
 - Cycle 5 : 72°C pendant 3 min
 - Cycle 6 : Retour à 20°C

Nous avons ensuite dilué au 1/10^{ème} les échantillons et réalisé un deuxième cycle de PCR en faisant varier le nombre de cycles de 2 à 10 selon la quantité à amplifier.

Chaque amplicon de PCR a été déposé sur un gel d'agarose 2%. Une fois que le gel a été coulé et réalisé, on dépose 2 μ L de chaque amplicon.

Les échantillons d'ADN dénaturés sont déposés dans les puits. Les échantillons sont dilués dans un tampon de charge (95%formamide, 20mM d'EDTA, 0.05% de xylène cyanol, 0.05% de bleu de bromophénol, glycérol).

Les 2 colorants permettent de suivre la migration des fragments d'ADN :

- la migration du bleu de bromophénol est "comparable" à celle d'un fragment d'ADN de 300 paires de base
- la migration du xylène cyanol est "comparable" à celle d'un fragment d'ADN de 4000 paires de base.

Le gel est placé dans une cuve d'électrophorèse avec du tampon d'électrophorèse (TBE 1X) qui est mis dans la cuve et la migration s'effectue sous l'action d'un champ électrique.

Les échantillons d'ADN dénaturés sont déposés dans les puits. Après migration, le gel est démoulé. On observe au transilluminateur le marqueur FAM sur les échantillons d'ADN marqués. La révélation des bandes d'ADN se fait grâce au bromure d'éthidium BET. C'est un intercalant de l'ADN qui se fixe entre les paires de bases.

. La visualisation des bandes révélées avec le bromure d'éthidium se fait aussi au transilluminateur UV (**Figure 1-21**). (Longueur d'onde : 254 nm).

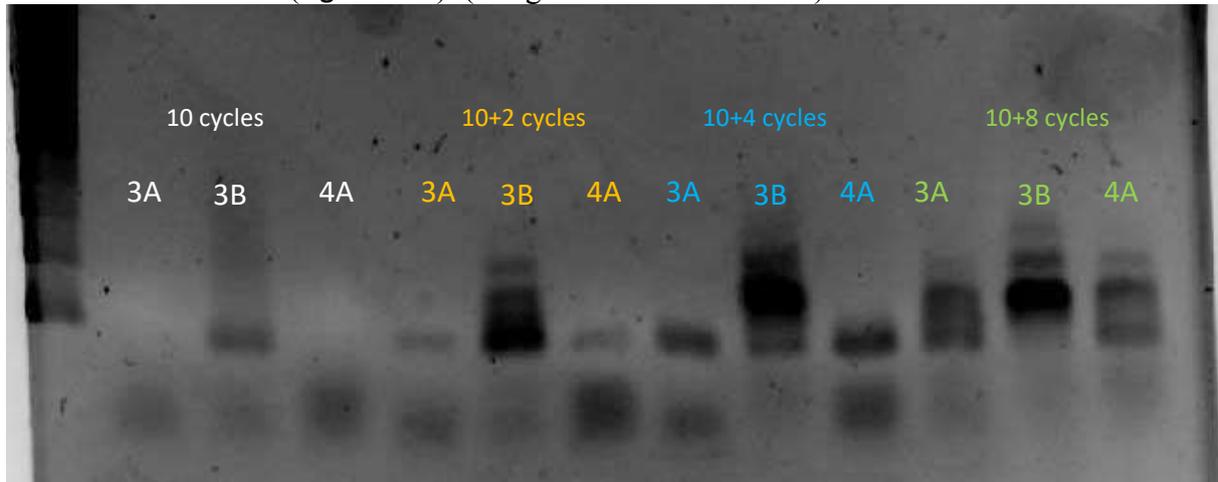


Figure 1-21: exemple d'amplification par PCR sur trois échantillons 3A, 3B et 4A, en réalisant 10 cycles de PCR puis en diluant les échantillons au 1/10ème et en réalisant 2 cycles, 4 cycles et 8 cycles de PCR. L'amplification qui donne la meilleure bande

6 PROTOCOLE D'AMPLIFICATION ET DE PURIFICATION DES MOLECULES D'ADN

6.1 L'électroélution

Le fragment issu de la PCR est ensuite purifié en déposant le reste du volume réactionnel sur un autre gel agarose 2% puis, après migration, en découpant au scalpel la bande fluorescente correspondante. Le morceau d'agarose contenant notre ADN est placé dans une cuve d'électrophorèse pour y subir une électroélution.

L'électroélution a été la première technique de purification que nous avons expérimentée.

Elle s'effectue dans un montage permettant au courant de passer à travers des morceaux de gel contenant l'échantillon à purifier. L'ADN est retenu en bas du canal par les sels d'acétate de sodium et les réactifs sont entraînés par le courant électrique (**Figure 1-22**).

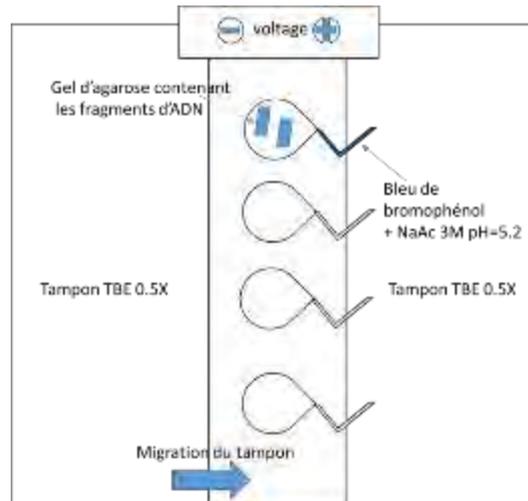


Figure 1-22: Principe de la migration par électroélution

À la fin de l'électroélution, la majorité des échantillons élués se retrouvent dans le fond du canal. Environ 100µL sont récupérés et on ajoute 2,5 volumes d'éthanol à 100% + 1µL de glycogène afin de faire précipiter l'ADN. La solution est placée au congélateur pendant une demi-journée. L'aliquot est ensuite centrifugé à 13 000 rpm pendant 25min. Des lavages successifs avec des solutions d'éthanol à 75% sont effectués tout en centrifugeant plusieurs fois. Le tube est retourné afin de le laisser sécher à l'air libre. Le culot est repris dans 10µL d'eau (Figure 1-23).

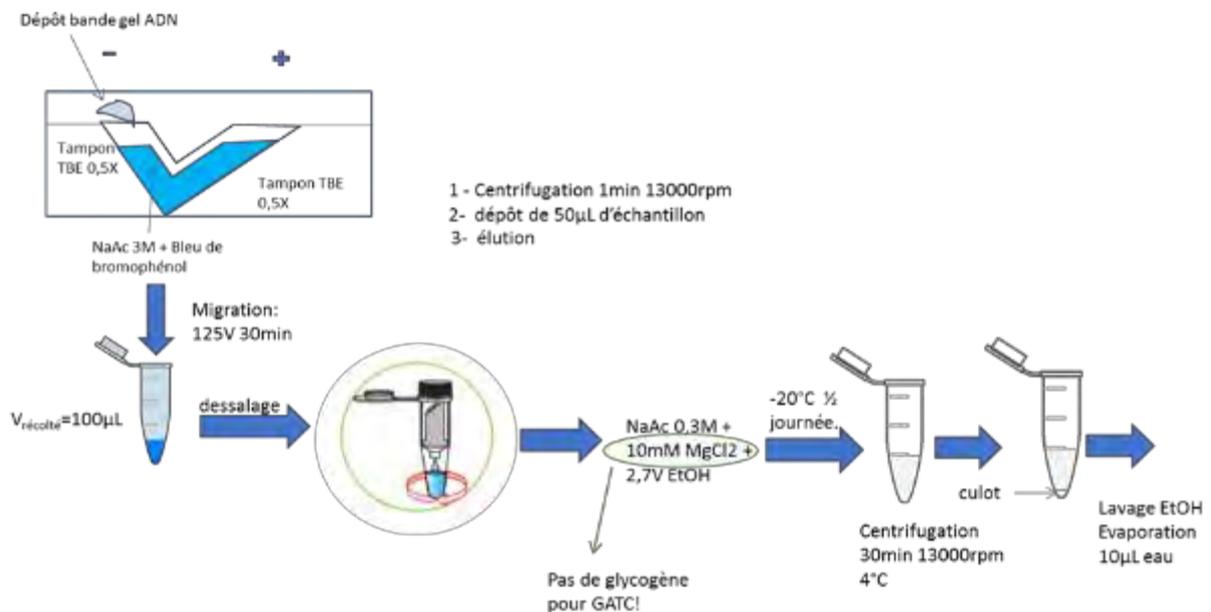


Figure 1-23: Purification post-migration par électroélution

Cette méthode a été utilisée pour la 1^{ère} collecte (1^{er} séquençage) pour laquelle nous avons 36 échantillons à purifier. Le pool final contenait environ 1µg. C'est un procédé long, le rendement n'a pas été déterminé avec exactitude mais nous perdons énormément d'ADN entre le dépôt sur gel et la purification. Donc, nous avons décidé d'utiliser des kits commerciaux de purification standards.

6.2 Kits de purifications :

Nous avons utilisé des kits de purification de la société Macherey Nagel : « NucleoSpin ».

L'échantillon est déposé sur gel. Puis on découpe au scalpel la bande fluorescente. L'échantillon est mélangé avec un tampon afin de dissoudre le gel. Avec la présence de sels chaotropiques, l'ADN va se lier à la membrane de silice. Les résidus sont enlevés en effectuant plusieurs lavages avec une solution éthanolique. Finalement, l'ADN purifié est élué avec une solution alcaline (5mM TrisHCl pH=8.5). Cette méthode est plus rapide et plus facile à réaliser. Néanmoins, on estime le rendement de purification à 50%.

Etapes :	Principe de kits de purification
EXTRACTION DU GEL	Ajout du tampon. Incubation 5-10 min à 50°C
LIAISON ADN-MEMBRANE	Ech placé dans une colonne + éppendorf. Centrifugations
LAVAGE MEMBRANE	Ajout de tampon TrisHCl. Centrifugations.
SECHAGE MEMBRANE	Centrifugations à sec.
ELUTION ADN	Ajout de 15 - 30µL de tampon éluant. Incubation puis centrifugation.

7 QUANTIFICATION POST-PCR DES APTAMERES AMPLIFIES

7.1 Le nanodrop

Le nanodrop est un spectrophotomètre UV qui à partir d'un faible volume d'échantillon déposé permet de quantifier les échantillons d'ADN. 1 à 2 µL d'échantillons, est déposé sur un socle à l'extrémité d'un câble de fibre optique. Lorsque l'échantillon est mis au contact de la fibre optique, la goutte est maintenue par une tension de surface et traversée par un signal lumineux. Ainsi, nous pouvons mesurer des ADN double brins comme des ADN simple brins dans une gamme de 2 à 15 000 ng/µL pour de l'ADN double brin.

7.2 Par électrophorèse sur gel d'agarose

En complément de la quantification par le nanodrop, nous avons aussi déposé sur gel d'agarose 2%, une petite quantité de nos échantillons purifiés afin de comparer les deux méthodes. Pour cela, nous avons déposé sur un gel, un ladder dont nous connaissons la quantité, et qui a été dilué et déposé dans différents puits. Puis nous avons déposé 1µL de notre échantillon afin de pouvoir le comparer avec la bande de 100 bp de notre ladder et d'estimer une gamme de concentration (**Figure 1-24**).

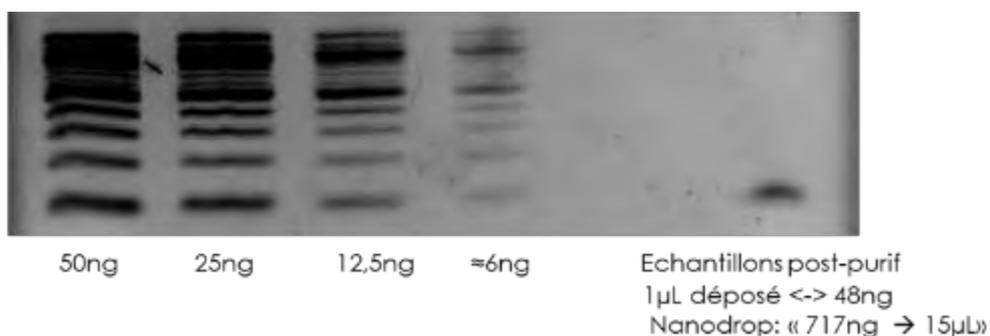


Figure 1-24 : Quantification d'un pool de 18 échantillons sur gel d'agarose 2%. Estimation de la quantité déposée en comparant 3 ladders dilués à 50ng, 25ng, 12.5ng et 6ng.

8 REPRESENTATION DE LA METHODE GLOBALE DE SELECTION PAR ELECTROPHORESE CAPILLAIRE, AMPLIFICATION PAR PCR, PURIFICATION DES APTAMERES

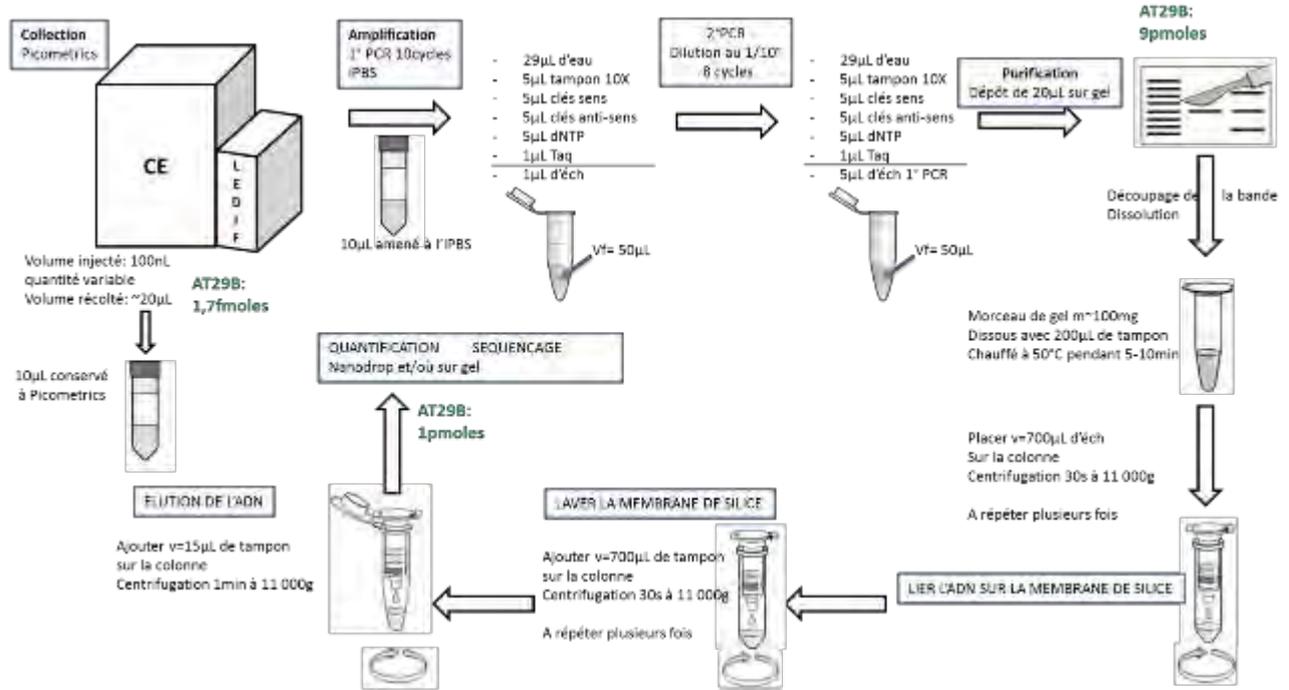


Figure 1-25 : Représentation de la méthode globale collecte + amplification + purification

Ci-dessus (**Figure 1-25**), nous représentons la méthode globale de collecte / amplification / purification de nos échantillons. Par électrophorèse, on injecte un volume de 100nL de banque L2 + AT29B + thrombine. Le volume de complexe récolté est compris entre 20 et 30µL. Une partie est conservée à Picometrics puis l'autre est amenée à l'IPBS afin de procéder à l'étape d'amplification par PCR. Nous faisons 10 cycles de PCR puis une dilution de l'échantillon préamplifié et nous faisons 8 cycles de PCR supplémentaires. Nous voulions ainsi nous assurer que nous ne produisons pas de sous-produits non-spécifiques. L'étape de purification consiste à déposer sur un gel d'agarose 20µL d'échantillon amplifié, découper au scalpel la bande fluorescente et utiliser le kit de purification commercial afin de recueillir 15µL de produits purifiés et quantifiés au nanodrop ou sur gel d'agarose 2%. A titre d'exemple, pour un échantillon d'aptamères à 1.7fmol, nous récupérons finalement 1pmol.

9 PRINCIPE DU SEQUENÇAGE HAUT-DEBIT ILLUMINA

Le séquençage haut-débit Illumina, représente 56% du marché des NGS. Le séquençage NGS permet de séquencer un nombre colossal de nucléotides à moindre coût et en un temps réduit (jusqu'à 15 milliards de nucléotides en 56 h sur le séquenceur MiSeq Illumina par exemple). Des millions de brins d'ADN sont séquencés en même temps (jusqu'à 25 millions sur le MiSeq) sur une lame en verre (flowcell). Les brins séquencés sont courts, entre 30 et 600 paires de bases selon le modèle de séquenceur Illumina et le kit de séquençage retenu. Par exemple, le séquenceur MiSeq peut séquencer 2*300 paires de bases et le HiSeq 2*150 paires de bases.

Une préparation de librairie doit donc être réalisée au préalable pour chaque échantillon afin d'obtenir des fragments à la taille adéquate et pouvant s'hybrider sur la flowcell. Puis une amplification clonale de cette librairie sera réalisée dans le séquenceur afin d'avoir plusieurs copies de chaque fragment et ainsi permettre la détection du signal fluorescent émis lors de l'incorporation de chaque nucléotide durant le séquençage.

9.1 Préparation de l'ADN : préparation de librairies.

L'ADN doit être fragmenté si nécessaire, et des séquences « adaptateurs » doivent être ajoutées aux extrémités de chaque fragment. Ainsi les brins d'ADN pourront s'hybrider aux adaptateurs présents également sur la flowcell. Les séquences adaptateurs comportent également une séquence index différente pour chaque échantillon, ce qui permet de séquencer plusieurs échantillons au cours d'un même run de séquençage si besoin.

Plusieurs kits contenant tous les réactifs nécessaires existent afin de réaliser la préparation des librairies. Ce procédé comprend généralement 3 étapes qui sont : la réparation des extrémités après sonication, la ligation des adaptateurs et le sizing (sélection des fragments par la taille). Nous avons voulu connaître les recommandations du fournisseur GATC, mais ces derniers n'ont pas souhaité étayer leur protocole. Néanmoins, après discussion avec plusieurs spécialistes, nous pouvons détailler le principe général de ces kits qui utilisent entre 1 et 5 µg d'échantillon. L'ADN est dans un premier temps fragmenté par sonication (inutile dans notre cas car les fragments à séquencer sont de 84pb) : on génère ainsi une librairie de fragments (150 – 200 bp) et les extrémités de ces fragments sont préparées pour la ligation des adaptateurs. Pour cela, on utilise des exonucléases afin d'obtenir des fragments à bords francs et permettre la phosphorylation des brins en 5' et l'adénylation des brins en 3'. La construction des adaptateurs pour la ligation est telle qu'un adaptateur P5 sera lié en 5' et un adaptateur P7 en 3' de chaque brin d'ADN. (**Figure 1-26**). Les fragments avec adaptateurs sont ensuite amplifiés (quelques cycles) et peuvent ensuite être isolés par extraction sur gel ou le plus souvent par purification et sélection par la taille par des billes magnétiques et des lavages avec une solution d'éthanol. Après dénaturation et dilution de la librairie finale (en pM) selon le protocole lié au kit de séquençage et au séquenceur utilisé, la librairie est déposée dans le séquenceur.

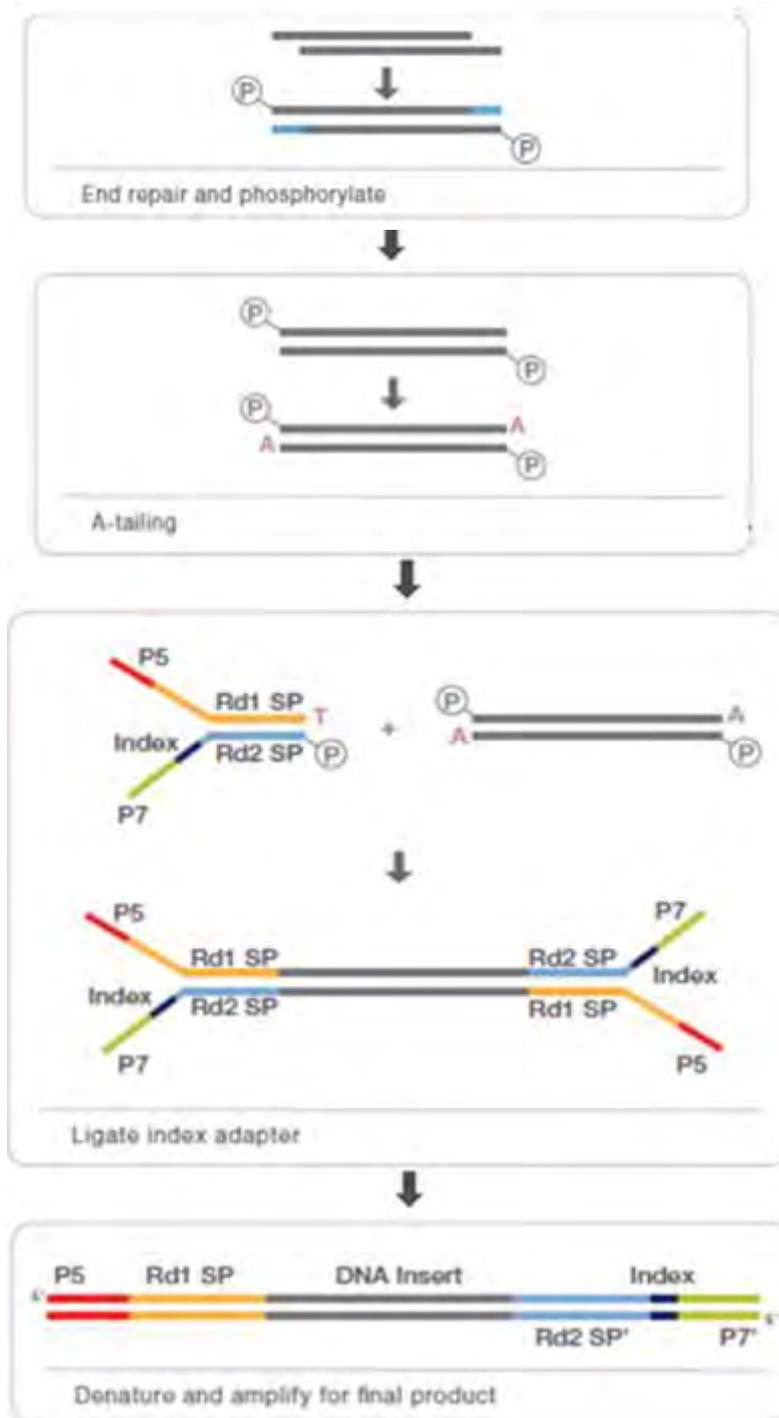


Figure 1-26 : Préparation de l'ADN pour le séquençage Illumina. [Source Illumina].

9.2 Amplification clonale : création de clusters

Une amplification clonale de chaque fragment est alors réalisée dans le séquenceur générant des millions de clusters comportant chacun environ 1000 copies d'un même fragment afin de permettre par la suite la détection du signal fluorescent lors du séquençage.

Cette amplification en phase solide à la surface de la flowcell se fait de manière automatisée dans le séquenceur MiSeq Illumina : les fragments préalablement dénaturés sont hybridés, grâce aux adaptateurs liés à leurs extrémités, aux adaptateurs présents également sur la flowcell

puis ils subissent une amplification en pont. La dénaturation et l'extension répétées entraînent l'amplification localisée de molécules simples dans des millions d'endroits à la surface de la flowcell. On génère ainsi des amas du même fragment d'ADN qu'on appelle un *cluster* (**Figure 1-27**) (les brins complémentaires sont clivés).

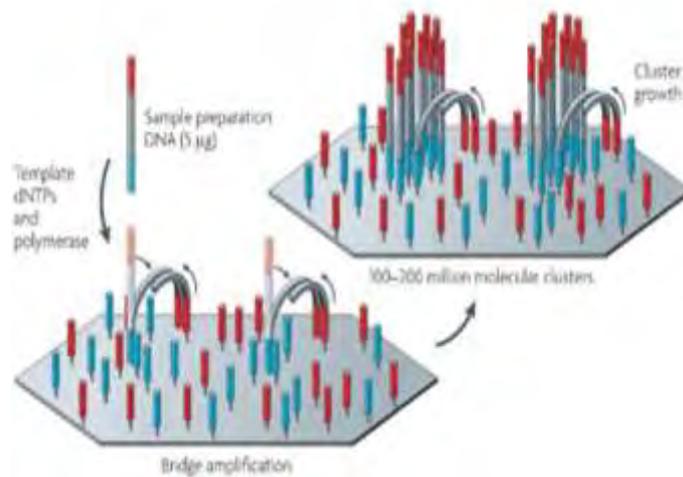


Figure 1-27 : Amplification "en pont" par la technique du séquençage Illumina

9.3 Séquençage

Une fois la génération des clusters effectuée, le séquençage peut commencer. Une amorce de séquençage est hybridée sur chaque brin et à chaque cycle, en présence des réactifs adéquats, un nucléotide est incorporé et identifié. En effet, chaque nucléotide porte un fluorochrome (différent pour chacun des 4 nucléotides) et un groupement protecteur de manière à ce que l'élongation soit bloquée pour permettre la détection du signal par une caméra après excitation du fluorochrome par un laser. L'image de ce premier cycle est récupérée puis le groupement protecteur avec le fluorochrome est clivé permettant l'incorporation du nucléotide suivant. Et ainsi de suite *cluster* (**Figure 1-27**). Si le séquençage choisi est paired-end (**Figure 1-28**), un cluster est à nouveau généré pour séquençer le brin dans l'autre sens avec une deuxième amorce de séquençage.

Paired end sequencing

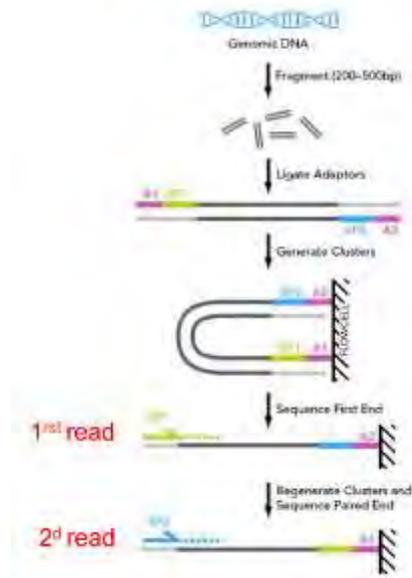


Figure 1-28 : Principe du séquençage paired-end. Les deux extrémités d'un fragment d'ADN sont séquencées dans les deux sens. Ainsi, en plus de l'information apportée par la séquence, la distance séparant les deux lectures est connue (cette distance ne peut dépasser ~800 pb).

Pour le séquenceur Mi-Seq, plusieurs kits de séquençage sont disponibles selon la taille des séquences (de 25 à 300 pb), le type de séquençage (single ou paired-end) et la quantité de séquences (1 à 50 millions) souhaités. Dans notre cas, nous avons choisi un séquençage *paired-end* afin que le fragment soit séquencé dans les deux sens, ce qui permet de réduire les erreurs dues au séquençage. Nos échantillons ont été séquencés en paired-end (2*75pb) avec le séquenceur MiSeq d'Illumina. Environ 16 millions de séquences (reads) ont été obtenues par échantillon.

9.4 Analyse bioinformatique

Si plusieurs échantillons ont été séquencés au cours d'un même run, les données sont tout d'abord démultiplexées, à savoir triées grâce à l'index propre à chaque échantillon (ou couple d'index si un index sur chaque adaptateur).

Puis les séquences sont « nettoyées » en enlevant les séquences correspondant aux adaptateurs :

Les adaptateurs ont la séquence suivante :

CAAGCAGAAGACGGCATAACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTG
CTCTTCCGATCT : 66 nucléotides

AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCTACACG
ACGCTCTTCCGATCT : 70 nucléotides

A l'issue du séquençage, les séquences obtenues sont assemblées en contigs (séquences continues générées par l'alignement de séquences de fragments qui se chevauchent). Environ 16 millions de paires de séquences ont été obtenues (**figure 1-29**).

...

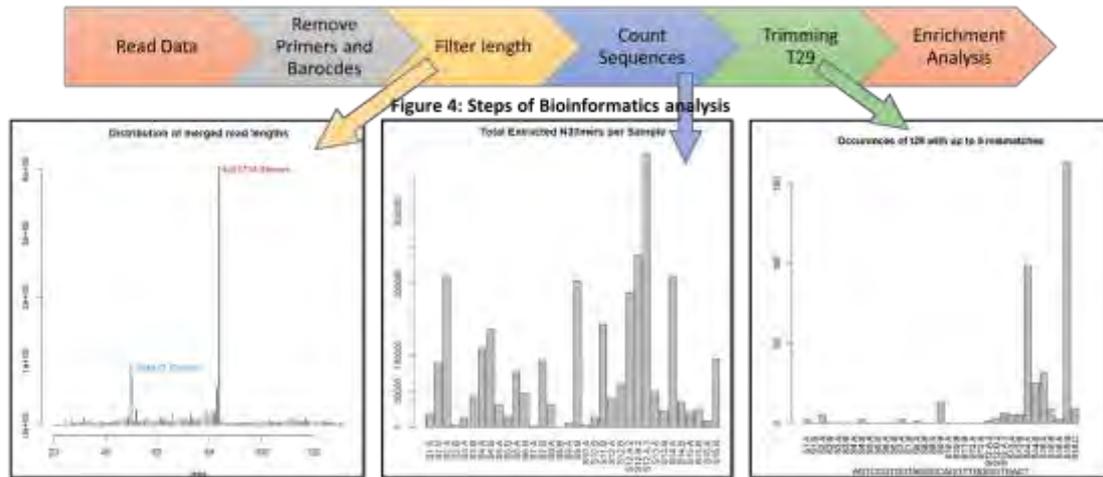


Figure 1-29 : Schéma représentant les étapes bio-informatiques utilisées pour le traitement des données.

Bibliographie :

1. Wilson DS, Szostak JW. In vitro selection of functional nucleic acids. *Annual review of biochemistry*. 1999, Vol. 68, pp. 611-47.
2. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*. 1990, pp. 505-510.
3. Szostak, Ellington et. Selection in vitro of single-stranded DNA molecules that fold into specific ligand-binding structures. *Nature*. 27 Février 1992, pp. 850-2.
4. Balamurugan S, Obubuafo A, Soper SA, Spivak DA. Surface immobilization methods for aptamer diagnostic applications. *Analytical and bioanalytical chemistry*. 2008, Vol. 390, pp. 1009-21.
5. Iliuk AB, Hu L, Tao WA. Aptamer in bioanalytical applications. *Analytical chemistry*. 15 Juin 2011 , pp. 44440-52.
6. Liu Y, Tuleouva N, Ramanculov E, Revzin A. Aptamer-based electrochemical biosensor for interferon gamma detection. *Analytical chemistry*. 2010, Vol. 82, pp. 8131-6.
7. Ng EW, Shima DT, Calias P, Cunningham ET Jr, Guyer DR, Adamis AP. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature reviews. Drug discovery*. 2006, Vol. 5, pp. 123-32.
8. Kanoatov M, Galievsky VA, Krylova SM, Cherney LT, Jankowski HK, Krylov SN. Using nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) for simultaneous determination of concentration and equilibrium constant. *Analytical chemistry*. 2015, Vol. 87, pp. 3099-106.
9. Circular dichroism and conformational polymorphism of DNA. *Nucleic acids research*. 2009, Vol. 37, pp. 1713-1725.
10. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*. 1953, Vol. 171, pp. 737-738.
11. Lu XJ, Shakked Z, Olson WK. A-form conformational motifs in ligand-bound DNA structures. *Journal of Molecular Biology*. . 2000, Vol. 300, pp. 819-40.
12. Gessner RV, Frederick CA, Quigley GJ, Rich A, Wang AH. The molecular structure of the lefthanded Z-DNA double helix at 1.0-Å atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *Journal of Biological chemistry*. 1989, Vol. 264, pp. 7921-7935.
13. Wang A. H., Quigley G. J., Kolpak F. J., Crawford J. L., van Boom J. H., van der Marel G., Rich A. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*. 1979, Vol. 282, pp. 680-686.
14. GELLERT M, LIPSETT MN, DAVIES DR. Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences*. 15, 1962 15 Dec, Vol. 48, pp. 2013-8.
15. Kikin O, D'Antonio L, Bagga PS. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res*. 34, 2006, pp. 676-82.

16. Bock LC, Griffin LC, Latham JA, Vermaas EH, Toole JJ. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*. 6, 1992, Vol. 355(6360), pp. 564-6.
17. Macaya RF, Schultze P, Smith FW, Roe JA, Feigon J. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proceedings of the National Academy of Sciences of the United States of America*. 90, 1993, Vol. 8, pp. 3745-9.
18. Tucker WO, Shum KT, Tanner JA. G-quadruplex DNA aptamers and their ligands: structure, function and application. *Current pharmaceutical design*. 2012, Vol. 18, pp. 2014-26.
19. Vorlíčková, Michaela Kejnovská, Iva Sagi, Janos Renčiuk, Daniel Bednářová, Klára Motlová, Jitka Kypr, Jaroslava. Circular dichroism and guanine quadruplexes. *Methods*. 1, 2012, Vol. 57, pp. 64-75.
20. Stefano Masiero, Robert Trotta, Silvia Pieraccini, Stefano De Tito, Rosaria Perone, Antonio Randazzo. A non-empirical chromophoric interpretation of CD spectra of DNA. *Organic & biomolecular chemistry*. 12, 2010, Vol. 8, pp. 2683-2692.
21. Girardot M, Gareil P, Varenne A. Interaction study of a lysozyme-binding aptamer with mono- and divalent cations by ACE. *Electrophoresis*. 2010, Vol. 31, pp. 546-55.
22. Gelinis AD, Davies DR, Edwards TE, Rohloff JC, Carter JD, Zhang C, Gupta S, Ishikawa Y, Hirota M, Nakaishi Y, Jarvis TC, Janjic N. Crystal structure of interleukin-6 in complex with a modified nucleic acid ligand. *The journal of biological chemistry*. J Biol Chem. 2014 Mar 21;289(12):8720-34. doi: 10.1074/jbc.M113.532697. Epub 2014 Jan 12., 2014, Vol. 289, 12.
23. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic acids research*. 9, 2006, Vol. 34, pp. 2723-2735.
24. Mann, K. G. Biochemistry and physiology of blood coagulation. *Thromb Haemost*. 1999, Vol. 82.
25. Griffin, L. C., Tidmarsh, G. F., Bock, L. C., Toole, J. J., & Leung, L. L. In vivo anticoagulant properties of a novel nucleotide-based thrombin inhibitor and demonstration of regional anticoagulation in extracorporeal circuits. *Blood*. 1993, Vol. 81, pp. 3271-3276.
26. Deng B, Lin Y, Wang C, Li F, Wang Z, Zhang H, Li XF, Le XC. Aptamer binding assays for proteins: the thrombin example--a review. *Analytica chimica acta*. 2014, Vol. 837, pp. 1-15.
27. Padmanabhan K, Padmanabhan KP, Ferrara JD, Sadler JE, Tulinsky A. The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *The journal of biological chemistry*. 24, 1993, Vol. 268, pp. 17651-17654.
28. Wu Q, Tsiang M, Sadler JE. Localization of the single-stranded DNA binding site in the thrombin anion-binding exosite. *The journal of biological chemistry*. 1992, Vol. 267, pp. 24408-12.
29. Paborsky, L.R., McCurdy, S.N., Griffin, L.C., Toole, J.J., and Leung, L.L. The single-stranded DNA aptamer-binding site of human thrombin. *The journal of biological chemistry*. 1993, Vol. 268, pp. 20808-20811.
30. Russo Krauss I, Merlino A, Giancola C, Randazzo A, Mazzarella L, Sica F. Thrombin-aptamer recognition: a revealed ambiguity. *Nucleic acids research*. 17, 2011, Vol. 39, pp. 7858-7867.
31. Baldrich E, O'Sullivan CK. Ability of thrombin to act as molecular chaperone, inducing formation of quadruplex structure of thrombin-binding aptamer. *Analytical Biochemistry*. 1, 2005, Vol. 341, pp. 194-197.

32. Nimjee SM, Rusconi CP, Sullenger BA. Aptamers: an emerging class of therapeutics. *Annual review of medicine*. 2005, Vol. 56, pp. 555-83.
33. Tasset, D. M., Kubik, M. F., & Steiner, W. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *Journal of molecular biology*. 1997, Vol. 272, pp. 688-98.
34. —. Oligonucleotide inhibitors of human thrombin that bind distinct epitopes. *J Mol Biol*. 1997, Vol. 272, pp. 688–698.
35. Ana Trapaidze, Jean-Pascal Héroult, Jean-Marc Herbert, Aurélien Bancaud, Anne-Marie Gué. Investigation of the selectivity of thrombin-binding aptamers for thrombin titration in murine plasma. *Biosensors and Bioelectronics*. 2016, Vol. 78, pp. 58-66.
36. Liu J, Stormo GD. Combining SELEX with quantitative assays to rapidly obtain accurate models of protein-DNA interactions. 141, 2005, Vol. 33.
37. Tombelli S, Minunni M, Luzi E, Mascini M. Aptamer-based biosensors for the detection of HIV-1 Tat protein. *Bioelectrochemistry*. 2005, Vol. 67, pp. 135-141.
38. Stoltenburg R, Reinemann C, Strehlitz B. FluMag-SELEX as an advantageous method for DNA aptamer selection. *Analytical and bioanalytical chemistry*. 2005, Vol. 383, pp. 83-91.
39. Bianchini M, Radrizzani M, Brocardo MG, Reyes GB, Gonzalez Solveyra C, Santa-Coloma TA. Specific oligobodies against ERK-2 that recognize both the native and the denatured state of the protein. *Journal of immunological methods*. 2001, Vol. 252, pp. 191–197.
40. Rhie A, Kirby L, Sayer N, Wellesley R, Disterer P, Sylvester I, Gill A, Hope J, James W, Tahiri-Alaoui A. Characterization of 2'-fluoro-RNA aptamers that bind preferentially to disease-associated conformations of prion protein and inhibit conversion. *The journal of biological chemistry*. 2003, Vol. 278, pp. 39697-705.
41. Fitzwater T, Polisky B. A SELEX primer. *Methods in enzymology*. 1996, Vol. 267.
42. Mayer G, Ahmed MS, Dolf A, Endl E, Knolle PA, Famulok M. Fluorescence-activated cell sorting for aptamer SELEX with cell mixtures. *Nature Protocols* 5. 5, 2010, pp. 199-2004.
43. Darmostuk M, Rimpelova S, Gbelcova H, Ruml T. Current approaches in SELEX: An update to aptamer selection technology. *Biotechnology advances*. 2015, Vol. 33, pp. 1141-1161.
44. Bayrac AT, Sefah K, Parekh P, Bayrac C, Gulbakan B, Oktem HA, Tan W. In vitro Selection of DNA Aptamers to Glioblastoma Multiforme. *ACS Chemical Neuroscience*. 2011, Vol. 2, pp. 175-181.
45. Ara MN, Hyodo M, Ohga N, Hida K, Harashima H. Development of a novel DNA aptamer ligand targeting to primary cultured tumor endothelial cells by a cell-based SELEX method. *PLoS One*. 2012, Vol. 7, p. e50174.
46. Kunii T, Ogura S, Mie M, Kobatake E. Selection of DNA aptamers recognizing small cell lung cancer using living cell-SELEX. *The analyst*. 2011, Vol. 136, pp. 1310-2.
47. Jiménez E, Sefah K, López-Colón D, Van Simaëys D, Chen HW, Tockman MS, Tan W. Generation of lung adenocarcinoma DNA aptamers for cancer studies. *PLoS One*. 2012, Vol. 7.
48. Yuanyuan Wang, Yun Luo , Tao Bing, Zheng Chen, Minhua Lu, Nan Zhang, Dihua Shangguan , Xin Gao. DNA Aptamer Evolved by Cell-SELEX for Recognition of Prostate Cancer. *PLoS One*. 2014, Vol. 9.

49. Li WM, Bing T, Wei JY, Chen ZZ, Shangguan DH, Fang J. Cell-SELEX-based selection of aptamers that recognize distinct targets on metastatic colorectal cancer cells. *Biomaterials*. 2014, Vol. 35, pp. 6998-7007.
50. Ozer A, Pagano JM, Lis JT. New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Molecular therapy. Nucleic acids*. 2014, Vol. 3, p. e183.
51. Daniels DA, Chen H, Hicke BJ, Swiderek KM, Gold L. A tenascin-C aptamer identified by tumor cell SELEX: systematic evolution of ligands by exponential enrichment. *Proceedings of the National Academy of Sciences*. 2014, Vol. 3.
52. Jenison RD, Gill SC, Pardi A, Polisky B. High-resolution molecular discrimination by RNA. *Science*. 1994, Vol. 263, pp. 1425-9.
53. Wang C, Zhang M, Yang G, Zhang D, Ding H, Wang H, Fan M, Shen B, Shao N. Single-stranded DNA aptamers that bind differentiated but not parental cells: subtractive systematic evolution of ligands by exponential enrichment. *Journal of biotechnology*. 2003, Vol. 102, pp. 15-22.
54. Jensen KB, Atkinson BL, Willis MC, Koch TH, Gold L. Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proceedings of the National Academy of Sciences of the United States of America*. 1995, Vol. 92, pp. 12220-4.
55. White R, Rusconi C, Scardino E, Wolberg A, Lawson J, Hoffman M, Sullenger B. Generation of species cross-reactive aptamers using "toggle" SELEX. *Molecular therapy: the journal of the American Society of Gene Therapy*. 2001, Vol. 4, pp. 567-73.
56. Sabeti PC, Unrau PJ, Bartel DP. Accessing rare activities from random RNA sequences: the importance of the length of molecules in the starting pool. *Chemistry and biology*. 1997, Vol. 4, 10, pp. 767-774.
57. Osborne SE, Ellington AD. Nucleic Acid Selection and the Challenge of Combinatorial Chemistry. *Chemical reviews*. 1997, pp. 349-370.
58. Beaudry AA, Joyce GF. Directed evolution of an RNA enzyme. *Science*. 1992, Vol. 257, pp. 635-641.
59. Cadwell RC, Joyce GF. Randomization of genes by PCR mutagenesis. *PCR methods and applications*. 1992, pp. 28-33.
60. Surugiu-Wärnmark I, Wärnmark A, Toresson G, Gustafsson JA, Bülow L. Selection of DNA aptamers against rat liver X receptors. *Biochemical and biophysical research communications*. 2005, Vol. 332, pp. 512-517.
61. G., Mayer. The chemical biology of aptamers. *Angewandte Chemie International Edition in English*. 2009, Vol. 48, pp. 2672-2689.
62. Jellinek D, Green LS, Bell C, Lynott CK, Gill N, Vargeese C, Kirschenheuter G, McGee DP, Abesinghe P, Pieken WA, et al. Potent 2'-amino-2'-deoxyuridine RNA inhibitors of basic fibroblast growth factor. *Biochemistry*. 1995, Vol. 34, pp. 11363-72.
63. Pagratis NC, Bell C, Chang YF, Jennings S, Fitzwater T, Jellinek D, Dang C. Potent 2'-amino-, and 2'-fluoro-2'-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nature Biotechnology*. 1997, Vol. 15, pp. 68-73.

64. Rohloff JC, Gelinas AD, Jarvis TC, Ochsner UA, Schneider DJ, Gold L, Janjic N. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Molecular therapy. Nucleic Acids*. 2014, Vol. 3, p. e201.
65. Hicke BJ, Marion C, Chang YF, Gould T, Lynott CK, Parma D, Schmidt PG, Warren S. Tenascin-C aptamers are generated using tumor cells and purified protein. *The journal of biological chemistry*. 2001, Vol. 276, pp. 48644-54.
66. Lupold SE, Hicke BJ, Lin Y, Coffey DS. Identification and characterization of nuclease-stabilized RNA molecules that bind human prostate cancer cells via the prostate-specific membrane antigen. *Cancer research*. 2002, Vol. 62, pp. 4029-4033.
67. Svobodova M, Bunka DH, Nadal P, Stockley PG, O'Sullivan CK. Selection of 2'F-modified RNA aptamers against prostate-specific antigen and their evaluation for diagnostic and therapeutic applications. *Analytical and bioanalytical chemistry*. 2013, Vol. 405, pp. 9149-57.
68. Layzer JM, McCaffrey AP, Tanner AK, Huang Z, Kay MA, Sullenger BA. In vivo activity of nuclease-resistant siRNAs. *RNA*. 2004, Vol. 10, pp. 766-771.
69. Smith CM, Steitz JA. Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell*. 1997, Vol. 89, pp. 669-672.
70. 2'-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. 1998, Vol. 273, pp. 20556-20567.
71. Kato Y, Minakawa N, Komatsu Y, Kamiya H, Ogawa N, Harashima H, Matsuda A. New NTP analogs: the synthesis of 4'-thioUTP and 4'-thioCTP and their utility for SELEX. *Nucleic acids research*. 2005, Vol. 33, pp. 2942-2951.
72. Pinheiro VB, Holliger P. Towards XNA nanotechnology: new materials from synthetic genetic polymers. *Trends in biotechnology*. 2014, Vol. 32, pp. 321-328.
73. Latham JA, Johnson R, Toole JJ. The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentynyl)-2'-deoxyuridine. 1994. Vol. 22, pp. 2817-22.
74. Jensen KB, Atkinson BL, Willis MC, Koch TH, Gold L. Using in vitro selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proceedings of the National Academy of Sciences of the United States of America*. 1995, Vol. 92, pp. 12220-4.
75. Vaught JD, Bock C, Carter J, Fitzwater T, Otis M, Schneider D, Rolando J, Waugh S, Wilcox SK, Eaton BE. Expanding the chemistry of DNA for in vitro selection. *Journal of the American chemical society*. 2010, Vol. 132, pp. 4141-51.
76. Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. *Annual review of biochemistry*. 2011, Vol. 80, pp. 273-299.
77. Davies DR, Gelinas AD, Zhang C, Rohloff JC, Carter JD, O'Connell D, Waugh SM, Wolk SK, Mayfield WS, Burgin AB, Edwards TE, Stewart LJ, Gold L, Janjic N, Jarvis TC. Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proceedings of the National Academy of Sciences*. 2012, Vol. 109, pp. 19971-6.

78. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, Carter J, Dalby AB, Eaton BE, Fitzwater T, Flather D, Forbes A, Foreman T, Fowler C, Gawande B, Goss M, Gunn M, Gupta S, Halladay D, Heil J, Heilig J, Hicke B, Husar G, Janjic N, Jarvis T, Jennings S. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*. 2010, Vol. 5, p. e15004.
79. *Mirror-image RNA that binds D-adenosine*. Klusmann S, Nolte A, Bald R, Erdmann VA, Fürste JP. 1996, Vol. 14, pp. 1112-5.
80. Ruta J, Ravelet C, Baussanne I, Décout JL, Peyrin E. Aptamer-based enantioselective competitive binding assay for the trace enantiomer detection. *Analytical chemistry*. 2007, Vol. 79, pp. 4716-9.
81. Vater A, Klusmann S. Turning mirror-image oligonucleotides into drugs: the evolution of Spiegelmer(®) therapeutics. *Drug discovery today*. 2015, Vol. 20, pp. 147-155.
82. Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew SY, McLaughlin SH, Herdewijn P, Holliger P. Synthetic genetic polymers capable of heredity and evolution. *Science*. Vol. 336, pp. 341-4.
83. Kimoto M, Yamashige R, Matsunaga K, Yokoyama S, Hirao I. Generation of high-affinity DNA aptamers using an expanded genetic alphabet. *Nature Biotechnology*. 2013, Vol. 31, pp. 453-7.
84. Sefah K, Yang Z, Bradley KM, Hoshika S, Jiménez E, Zhang L, Zhu G, Shanker S, Yu F, Turek D, Tan W, Benner SA. In vitro selection with artificial expanded genetic information systems. *PNAS, Proceedings of the National Academy of Sciences of the United States of America*. 2014, Vol. 111, pp. 1449-54.
85. ollenstein M1, Hipolito CJ, Lam CH, Perrin DM. Toward the combinatorial selection of chemically modified DNAzyme RNase A mimics active against all-RNA substrates. *ACS combinatorial science*. 2013, Vol. 15, pp. 174-82.
86. He W, Elizondo-Riojas MA, Li X, Lokesh GL, Somasunderam A, Thivyanathan V, Volk DE, Durland RH, Englehardt J, Cavasotto CN, Gorenstein DG. X-aptamers: a bead-based selection method for random incorporation of druglike moieties onto next-generation aptamers for enhanced binding. *Biochemistry*. 2012, Vol. 51, pp. 8321-3.
87. Hagiwara K, Fujita H, Kasahara Y, Irisawa Y, Obika S, Kuwahara M. In vitro selection of DNA-based aptamers that exhibit RNA-like conformations using a chimeric oligonucleotide library that contains two different xeno-nucleic acids. *Molecular bioSystems*. 2015, Vol. 11, pp. 71-76.
88. Davis KA, Abrams B, Lin Y, Jayasena SD. Use of a high affinity DNA ligand in flow cytometry. *Nucleic Acids Research*. 1996, Vol. 15, pp. 702-6.
89. Mendonsa SD, Bowser MT. In vitro selection of high-affinity DNA ligands for human IgE using capillary electrophoresis. *Analytical chemistry*. 2004, Vol. 76, pp. 5387-92.
90. Tsai RY1, Reed RR. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Molecular and cellular biology*. 1998, Vol. 18, pp. 6447-6456.
91. Mosing R.K, Mendonsa S.D, Bowser M.T. Capillary Electrophoresis-SELEX selection of aptamers with affinity for HIV-1 reverse transcriptase. *Analytical chemistry*. 2005, Vol. 77, pp. 6107-6112.
92. Tang J.J, Xie J.W, Shao N.S, Yan Y. In vitro selection of DNA aptamer against abrin toxin and aptamer-based abrin direct detection. *Biosensors & Bioelectronics*. 2006, Vol. 27, pp. 1303-1311.

93. Berezovski M, Drabovich A, Krylova SM, Musheev M, Okhonin V, Petrov A, Krylov SN. Nonequilibrium capillary electrophoresis of equilibrium mixtures: a universal tool for development of aptamers. *Journal of the American Chemical Society*. 2005, Vol. 127, pp. 3165-71.
94. Okhonin V, Berezovski M, Krylov SN. Sweeping capillary electrophoresis: a non-stopped-flow method for measuring bimolecular rate constant of complex formation between protein and DNA. *Journal of the American Chemical Society*. 2004, Vol. 126, pp. 7166-7.
95. Drabovich A, Berezovski M, Krylov SN. Selection of smart aptamers by equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM). *Journal of the American Chemical Society*. 2005, Vol. 127, pp. 11224-5.
96. Mendonsa SD, Bowser MT. In vitro selection of aptamers with affinity for neuropeptide Y using capillary electrophoresis. *Journal of the American Chemical Society*. 2005, Vol. 127, pp. 9382-3.
97. Drabovich AP, Berezovski M, Okhonin V, Krylov SN. Selection of smart aptamers by methods of kinetic capillary electrophoresis. *Analytical Chemistry*. 2006, Vol. 78, pp. 3171-8.
98. Mallikaratchy P, Stahelin RV, Cao Z, Cho W, Tan W. Selection of DNA ligands for protein kinase C-delta. *Chemical communications (Cambridge, England)*. 2006, Vol. 30, pp. 3229-31.
99. Hesselberth J, Robertson MP, Jhaveri S, Ellington AD. In vitro selection of nucleic acids for diagnostic applications. *Journal of biotechnology*. 2000, Vol. 74, pp. 15-25.
100. Yang J, Bowser MT. Capillary electrophoresis-SELEX selection of catalytic DNA aptamers for a small-molecule porphyrin target. *Analytical chemistry*. 2013, Vol. 85, pp. 1525-30.
101. Eaton RM, Shallcross JA, Mael LE, Mears KS, Minkoff L, Scoville DJ, Whelan RJ. Selection of DNA aptamers for ovarian cancer biomarker HE4 using CE-SELEX and high-throughput sequencing. *Analytical and bioanalytical chemistry*. 2015, Vol. 23, pp. 6965-73.
102. Cella LN, Sanchez P, Zhong W, Myung NV, Chen W, Mulchandani A. NANO APTASENSOR FOR PROTECTIVE ANTIGEN TOXIN OF ANTHRAX. *Analytical chemistry*. 2010, Vol. 82, pp. 2042-2047.
103. Dong L, Tan Q, Ye W, Liu D, Chen H, Hu H, Wen D, Liu Y, Cao Y, Kang J, Fan J, Guo W, Wu W. Screening and Identifying a Novel ssDNA Aptamer against Alpha-fetoprotein Using CE-SELEX. *Scientific reports*. 2015.
104. Jing M, Bowser MT. Tracking the Emergence of High Affinity Aptamers for rhVEGF165 during CE-SELEX Using High Throughput Sequencing. *Analytical Chemistry*. 2013, Vol. 85, pp. 10761-70.
105. Zhang Y, Yu Z, Jiang F, Fu P, Shen J, Wu W, Li J. Two DNA aptamers against avian influenza H9N2 virus prevent viral infection in cells. *PLoS one*. 2015.
106. Ashley J, Li SF. Three-dimensional selection of leptin aptamers using capillary electrophoresis and implications for clone validation. *Analytical biochemistry*. 2013, Vol. 434, pp. 146-52.
107. Dinh T.T, Karel K, Kris P.J, Jeroen P, Dragana S, Jeroen L. Selection of aptamers against Ara h 1 protein for FO-SPR biosensing of peanut allergens in food matrices. *Biosensors and Bioelectronics*. 2013, Vol. 43, pp. 245-251.
108. Berezovski MV, Musheev MU, Drabovich AP, Jitkova JV, Krylov SN. Non-SELEX: selection of aptamers without intermediate amplification of candidate oligonucleotides. *Nature Protocols*. 2006, Vol. 1, pp. 1359-69.

109. Ashley J, Ji K, Li SF. Selection of bovine catalase aptamers using non-SELEX. *Electrophoresis*. 2012, Vol. 33, pp. 2783-9.
110. Tok J, Lai J, Leung T, Li SF. Selection of aptamers for signal transduction proteins by capillary electrophoresis. *Electrophoresis*. 2010, Vol. 31, pp. 2055-62.
111. Ashley J, Ji K, Li SF. Selection of cholesterol esterase aptamers using a dual-partitioning approach. *Electrophoresis*. 2015, Vol. 36, pp. 2616-21.
112. Berezovski M, Musheev M, Drabovich A, Krylov SN. Non-SELEX selection of aptamers. *Journal of the American Chemical Society*. 8 Feb 2006, Vol. 128, pp. 1410-1.
113. Fiore E, Dausse E, Dubouchaud H, Peyrin E, Ravelet C. Ultrafast capillary electrophoresis isolation of DNA aptamer for the PCR amplification-based small analyte sensing. *Frontiers in chemistry*. 2015, Vol. 3:49.
114. Micro free-flow electrophoresis: theory and applications. *Analytical and Bioanalytical Chemistry*. 2009, Vol. 394, pp. 187–198.
115. McKeague M, De Girolamo A, Valenzano S, Pascale M, Ruscito A, Velu R, Frost NR, Hill K, Smith M, McConnell EM, DeRosa MC. Comprehensive analytical comparison of strategies used for small molecule aptamer evaluation. *Analytical chemistry*. 2015, Vol. 87, pp. 8608-12.
116. Jing M, Bowser MT. Methods for measuring aptamer-protein equilibria: A review. *Analytica chimica acta*. 2011, Vol. 686, pp. 9-18.
117. Valenzano S, De Girolamo A, DeRosa MC, McKeague M, Schena R, Catucci L, Pascale M. Screening and Identification of DNA Aptamers to Tyramine Using in Vitro Selection and High-Throughput Sequencing. *ACS combinatorial science*. 2016, Vol. 18, pp. 302-13.
118. Cruz-Aguado, JA. et Penner, G.,. Determination of ochratoxin a with a DNA aptamer. *Journal of Agricultural and Food Chemistry*. 2008, Vol. 56, pp. 10456-10461.
119. Cruz-Aguado JA, Penner G. Determination of Ochratoxin A with a DNA Aptamer. *Journal of agricultural and food chemistry*. 2008, Vol. 56, pp. 10456-10461.
120. Wong I, Lohman TM. A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A*. 1993, Vol. 90, pp. 5428-32.
121. Fried M, Crothers DM. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research*. 1981, Vol. 9, pp. 6505-25.
122. Garner MM, Revzin A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucleic Acids research*. 1981, Vol. 9, pp. 3047-60.
123. Ravelet C, Grosset C, Peyrin E. Liquid chromatography, electrochromatography and capillary electrophoresis applications of DNA and RNA aptamers. *Journal of Chromatography A*. 2006, Vol. 1117, pp. 1-10.
124. E., Peyrin. Nucleic acid aptamer molecular recognition principles and application in liquid chromatography and capillary electrophoresis. *Journal of separation science*. 2009, Vol. 32, pp. 1531-6.

125. Forier C, Boschetti E, Ouhammouch M, Cibiel A, Ducongé F, Nogré M, Tellier M, Bataille D, Bihoreau N, Santambien P, Chtourou S, Perret G. DNA aptamer affinity ligands for highly selective purification of human plasma-related proteins from multiple sources. *Journal of chromatography. A*. 2017, Vol. 1489, pp. 39-50.
126. Perrier S, Bouilloud P, De Oliveira Coelho G, Henry M, Peyrin E. Small molecule aptamer assays based on fluorescence anisotropy signal-enhancer oligonucleotides. *Biosensors & bioelectronics*. 2016, Vol. 82, pp. 155-61.
127. Scoville DJ, Uhm TK, Shallcross JA, Whelan RJ. Selection of DNA Aptamers for Ovarian Cancer Biomarker CA125 Using One-Pot SELEX and High-Throughput Sequencing. *Journal of nucleic acids*. 2017, Vol. 2017, p. 9.
128. Tran DT, Janssen KP, Pollet J, Lammertyn E, Anné J, Van Schepdael A, Lammertyn J. Selection and characterization of DNA aptamers for egg white lysozyme. *Molecules*. 2010, Vol. 15, pp. 1127-40.
129. Russo Krauss I, Merlino A, Randazzo A, Novellino E, Mazzarella L, Sica F. High-resolution structures of two complexes between thrombin and thrombin-binding aptamer shed light on the role of cations in the aptamer inhibitory activity. *Nucleic acids research*. 2012, Vol. 40, pp. 8119-28.
130. Lin PH, Chen RH, Lee CH, Chang Y, Chen CS, Chen WY. . Studies of the binding mechanism between aptamers and thrombin by circular dichroism, surface plasmon resonance and isothermal titration calorimetry Lin PH1, Chen RH, Lee CH, Chang Y, Chen CS, Chen WY. *Colloids and surfaces. B, Biointerfaces*. 2011, Vol. 88, pp. 552-8.
131. Liu Y, Liu N, Ma X, Li X, Ma J, Li Y, Zhou Z, Gao Z. Highly specific detection of thrombin using an aptamer-based suspension array and the interaction analysis via microscale thermophoresis. *The Analyst*. 2015, Vol. 140, pp. 2762-70.
132. Stuart CH, Riley KR, Boyacioglu O, Herpai DM, Debinski W, Qasem S, Marini FC, Colyer CL, Gmeiner WH. Selection of a Novel Aptamer Against Vitronectin Using Capillary Electrophoresis and Next Generation Sequencing. *Molecular therapy. Nucleic acids*. 2016, Vol. 5, p. e386.
133. Cherney LT, Kanoatov M, Krylov SN. Method for determination of peak areas in nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical chemistry*. 2011, Vol. 83, pp. 8617-22.
134. Zhang Y.W, Yan H.Y, Fu P, Jiang.F, Zhang Y, Wu W.X, Li J.X. Modified capillary electrophoresis based measurement of the binding between DNA aptamers and an unknown concentration target. *Analytical and Bioanalytical Chemistry*. 2013, Vol. 405, pp. 5549-55.
135. Musheev MU, Filiptsev Y, Krylov SN. Noncooled capillary inlet: a source of systematic errors in capillary-electrophoresis-based affinity analyses. *Analytical chemistry*. 2010, Vol. 82, pp. 8637-41.
136. Yunusov D, So M, Shayan S, Okhonin V, Musheev MU, Berezovski MV, Krylov SN. Kinetic capillary electrophoresis-based affinity screening of aptamer clones. *Analytica chimica acta*. 2009, Vol. 631, pp. 102-7.
137. Langan TJ, Nyakubaya VT, Casto LD, Dolan TD, Archer-Hartmann SA, Yedlapalli SL, Sooter LJ, Holland LA. Assessment of aptamer-steroid binding using stacking-enhanced capillary electrophoresis. *Electrophoresis*. 2012, Vol. 33, pp. 866-9.
138. Riley KR, Saito S, Gagliano J, Colyer CL. Facilitating aptamer selection and collection by capillary transient isotachopheresis with laser-induced fluorescence detection. *Journal of chromatography. A*. 2014, Vol. 1368, pp. 183-9.

139. —. Facilitating aptamer selection and collection by capillary transient isotachopheresis with laser-induced fluorescence detection. *Journal of chromatography A*. 2014, Vol. 1368, pp. 183-9.
140. Yufa R, Krylova SM, Bruce C, Bagg EA, Schofield CJ, Krylov SN. Emulsion PCR Significantly Improves Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures-Based Aptamer Selection: Allowing for Efficient and Rapid Selection of Aptamer to Unmodified ABH2 Protein. *Analytical chemistry*. 2015, Vol. 87, pp. 1411-9.
141. Selection., Real-Time PCR-Coupled CE-SELEX for DNA Aptamer. Ruff P, Pai RB, Storic F. *ISRN molecular biology*. Vol. 2012, p. 9.
142. Musheev MU, Krylov SN. Selection of aptamers by systematic evolution of ligands by exponential enrichment: Addressing the polymerase chain reaction issue. *Analytica chimica acta*. 2006, Vol. 564, pp. 91-6.
143. ChoM, XiaoY, Nie J, Stewart R, CsordasAT, Oh SS, Thomson JA, Soh HT. Quantitative selection of DNA aptamers through microfluidic selection and high-throughput sequencing. *Proc Natl Acad Sci*. 2010, Vol. 107, pp. 15373–15378.
144. Li Y, Guo L, Zhang F, Zhang Z, Tang J, Xie J. High-sensitive determination of human AE -thrombin by its 29-mer aptamer in affinity probe capillary electrophoresis. *Electrophoresis*. 2008, Vol. 29, pp. 2570-7.
145. Deng QP, Tie C, Zhou YL, Zhang XX. Cocaine detection by structure-switch aptamer-based capillary zone electrophoresis. *Electrophoresis*. 2012, Vol. 33, pp. 1465-70.
146. Kanoatov M, Krylov SN. Analysis of DNA in Phosphate Buffered Saline Using Kinetic Capillary Electrophoresis. *Analytical chemistry*. 2016, Vol. 88, pp. 7421-8.
147. Petrov A, Okhonin V, Berezovski M, Krylov SN. Kinetic capillary electrophoresis (KCE): a conceptual platform for kinetic homogeneous affinity methods. *Journal of the American Chemical Society*. 2005, Vol. 127, pp. 17104-10.
148. Ray T, Mills A, Dyson P. Tris-dependent oxidative DNA strand scission during electrophoresis. *Electrophoresis*. 1995, Vol. 16, pp. 888-94.
149. JJ, Toulmé et R, Giegé. Les aptamères: des ligands et des catalyseurs oligonucléotidiques obtenus par sélection in vitro. *Médecine/sciences*. 1998, Vol. 14, pp. 155-66.
150. Riley KR, Saito S, Gagliano J, Colyer CL. Facilitating aptamer selection and collection by capillary transient isotachopheresis with laser-induced fluorescence detection. *Journal of chromatography A*. 2014, Vol. 1368, pp. 183-9.
151. Tse C, Capeau J. Quantification des acides nucléiques. *Annales de biol clinique*. 2003, Vol. 61, pp. 279-93.

Résumé:

Les aptamères sont des oligomères d'ADN ou d'ARN simple brin qui, en se repliant sous forme de structures tridimensionnelles peuvent avoir des interactions fortes et spécifiques envers un certain nombre de cibles. L'objectif de cette thèse a été de compléter les études existantes sur l'utilisation de l'électrophorèse capillaire (CE) et les aptamères afin de mettre au point une méthode de sélection d'aptamères par CE couplée à la fluorescence induite par laser et le séquençage haut-débit Illumina. Dans un premier temps, nous avons mis au point une méthode de détection et de séparation par électrophorèse capillaire couplée à la double détection UV-LEDIF d'une banque d'ADN en interaction avec une cible : la thrombine. C'est un modèle déjà étudié pour lequel deux aptamères ont fait l'objet de publications. Nous avons utilisé l'aptamère T29 dans le cadre de notre étude car c'est celui qui présente la meilleure affinité. L'électrophorèse capillaire est un puissant outil analytique qui facilite l'efficacité de sélection des aptamères et précise la détermination des paramètres d'interactions. Nous avons ainsi pu déterminer la constante d'affinité K_D par CE-UV-LEDIF sur le modèle de base : la thrombine. Par ailleurs, nous montrons également comment l'utilisation du tampon Tris peut dégrader un ADN simple brin en électrophorèse capillaire et nous proposons comme alternative l'utilisation d'un tampon sodium phosphate dibasique qui évite ce phénomène de dégradation. Enfin, nous expliquons la difficulté d'amplification par qPCR et PCR d'un aptamère comme le T29 ayant une structure en G-quadruplex. Nous avons montré que le séquençage haut-débit Illumina nous a permis de trouver une corrélation entre le nombre de molécules séquencées et le nombre de séquences obtenues. L'analyse des séquences obtenues montre une quantité importante (20%) de séquences de T29 qui ne correspondent pas à la séquence de cet aptamère. Cela prouve que les étapes de PCR et de séquençage haut débit pour la détection de G-quadruplex peuvent induire un biais dans l'identification de ces molécules.

Abstract:

Aptamers are oligomers of small single-stranded DNA or RNA which can have strong and specific interactions with some targets when they fold into three-dimensional structures. The objective of this thesis was to complete existing studies on the use of capillary electrophoresis to develop a method for the selection of aptamers by CE coupled to laser induced fluorescence and Illumina high-throughput sequencing. In a first step, we developed a method of detection and separation by capillary electrophoresis coupled with the double detection UV-LEDIF of a DNA library interacting with a target: thrombin. It is a model already studied and for which two aptamers have been published. We used aptamer T29 as part of our study because it has the best affinity. Capillary Electrophoresis is a powerful analytical tool that facilitates the selection efficiency of aptamers and specifies the determination of the interaction parameters. We thus were able to determine the affinity constant K_D by CE-UV-LEDIF on the basic model: thrombin. Moreover, we also show how the use of Tris buffer can degrade single-stranded DNA during capillary electrophoresis and we propose as an alternative the use of a dibasic sodium phosphate buffer which avoids the phenomenon of degradation. Finally, we explain the difficulty of amplification by qPCR and PCR of an aptamer such as T29 with a G-quadruplex structure. We showed that the Illumina high-throughput sequencing allowed us to find a correlation between the number of sequenced molecules and the number of sequences obtained. Analysis of the sequences obtained shows a significant amount (20%) of T29 sequences which do not correspond to the sequence of this aptamer. This shows that the PCR and high-throughput sequencing steps for the detection of G-quadruplex can induce bias in the identification of these molecules.