



HAL
open science

Biophysical characterization of aptamer-ligand interactions by native mass spectrometry

Stefano Piccolo

► **To cite this version:**

Stefano Piccolo. Biophysical characterization of aptamer-ligand interactions by native mass spectrometry. Theoretical and/or physical chemistry. Université de Bordeaux, 2019. English. NNT: 2019BORD0276 . tel-02466639

HAL Id: tel-02466639

<https://theses.hal.science/tel-02466639>

Submitted on 4 Feb 2020

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 PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE SCIENCES CHIMIQUES
SPECIALITE CHIMIE PHYSIQUE

Par Stefano PICCOLO

**Biophysical characterization of aptamer-ligand
interactions by native mass spectrometry**

sous la direction de : Dr. Valérie Gabelica

Soutenue le 13 novembre 2019

Membres du jury :

Mme	TOKARSKI Caroline	Professeur à l'Université de Bordeaux (F)	Présidente
M.	SOBOTT Frank	Professeur à l'Université de Leeds (UK)	Rapporteur
Mme	SISSI Claudia	Professeur à l'Université de Padoue (IT)	Rapporteur
Mme	CIANFERANI Sarah	Directeur de recherche, CNRS, Strasbourg (F)	Examineur
M.	DARFEUILLE Fabien	Chargé de recherche, INSERM, Bordeaux (F)	Invité

Questa tesi è dedicata ad Adèle

<<[...] ogni studente in chimica, davanti ad un qualsiasi trattato, dovrebbe essere consapevole che in una di quelle pagine, forse in una sola riga o formula o parola, sta scritto il suo avvenire, in caratteri indecifrabili, ma che diverranno chiari “poi”: dopo il successo o l’errore o la colpa, la vittoria o la disfatta.>>

<<[...] every chemistry student, confronting any treatise, should be aware that on one of those pages, perhaps in a single line or formula or word, is written his future, in characters that are indecipherable but will become clear “later”: after success or error or guilt, victory or defeat.>>

Il Sistema periodico – Primo Levi

This work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 642738 (ITN "MetaRNA").

Index

I.	Abbreviations	xiii
II.	Résumé	xv
III.	Abstract.....	xvii
IV.	Acknowledgments.....	xix
1.	Introduction	1
1.1	Nucleic acids structure.....	1
1.1.1	Primary Structure (Sequence)	1
1.1.2	Secondary structure	2
1.1.3	Tertiary structure	6
1.1.4	Metal cations in nucleic acid structure	7
1.1.5	Structure-function relation and tertiary structure.....	8
1.2	Aptamers and SELEX.....	9
1.3	Standard state and equilibrium constant	12
1.4	Isothermal titration calorimetry (ITC).....	14
1.5	Mass spectrometry.....	16
1.5.1	Native mass spectrometry.....	17
1.5.2	Ion activation in mass spectrometry	18
1.5.3	How solution concentrations translate into MS intensities.....	19
1.5.4	Ion mobility mass spectrometry	21
1.5.5	Collision Induced Unfolding.....	22
2.	Objective of the thesis.....	24
2.1	Aptamers analysed in this thesis	24
3.	Materials and Methods.....	27
3.1	UV melting.....	27
3.2	Circular Dichroism	29
3.3	ITC instrumental set up and data treatment.....	29
3.4	Mass spectrometers	31
3.4.1	Agilent 6560 ESI-IM-Q-ToF	31
3.4.2	Thermo Exactive	37
3.5	Relative response factor determination	39

3.5.1	Determination of dissociation constant using MS titrations.....	40
3.5.2	Extraction of MS peak Intensity and background subtraction	43
3.6	Ion Mobility methods	45
3.6.1	CCS determination and CCS distribution reconstruction	45
3.6.2	Collision-induced unfolding experiments	47
3.7	Sample preparation	51
3.7.1	Nucleic acids.....	51
3.7.2	Ligands	53
Part I: Quantitative MS.....		57
4.	Quantitative comparison of Native MS and ITC	59
4.1	Why comparing native MS with ITC?.....	59
4.2	Malachite Green aptamer: a model RNA aptamer to compare ITC and MS	60
4.2.1	MS buffers TMAA and NH ₄ OAc do not alter malachite green binding mode...	61
4.2.2	Determination of MGA-MG binding constant with native MS titration.....	62
4.2.3	Native MS titration without and with response factor correction.....	64
4.3	RiboFMN aptamer: an Mg ²⁺ -dependent RNA aptamer.....	73
4.3.1	Does RiboFMN aptamer (1FMN) need Magnesium?.....	74
4.3.2	ITC of 1FMN in MgCl ₂ 200 μM and NH ₄ OAc 100 mM or TMAA 100 mM	75
4.3.3	1FMN native MS.....	76
4.3.4	MS titrations of 1FMN in presence of magnesium	80
4.4	Aminoglycoside RNA aptamers.....	85
4.4.1	Neomycin RNA aptamer (1NEM).....	86
4.4.2	Tobramycin RNA aptamer (1TOB)	88
4.5	Conclusions.....	90
Part II: Ion Mobility Spectrometry.....		93
5.	IM-MS of aptamers and their complexes	95
5.1	1TOB.....	95
5.2	1NEM.....	96
5.3	MGA.....	96
5.4	1FMN.....	97
5.5	Tetracycline RNA aptamer	100
5.6	Cocaine Binding Aptamers.....	104
5.6.1	Introduction	104

5.6.2	IM-MS to track adaptive binding	105
5.6.3	Maximizing the CBA ⁷⁻ signal by changing the [NH ₄ OAc]	110
5.6.4	Model constructs for comparison in gas phase: extra dA.....	112
5.6.5	Do the dA modifications alter aptamer structure in solution?	113
5.6.6	Adenine constructs in NH ₄ OAc 25 mM	118
5.6.7	IM-MS of CBA-dA aptamers.....	118
5.6.8	IM-MS of quinine complexes	121
5.6.9	Collisional induced unfolding of MN4 ⁷⁻ and CBA+dA ⁷⁻	122
5.6.10	CIU of aptamer-quinine complexes	129
5.7	Conclusions.....	136
6.	General conclusions and perspectives	139
6.1	Quantitative MS.....	139
6.1.1	The choice of the electrolyte	139
6.1.2	Relative response factor correction.....	140
6.1.3	One point measurement, full titration and data treatment	140
6.1.4	Saturated binding and detection limits	140
6.2	Ion Mobility Spectrometry.....	141
6.2.1	Tertiary structural elements prevent RNA compaction upon ESI	141
6.2.2	Some secondary structure aptamers can be discriminated via IM-MS.....	141
6.2.3	CIU/D breakdown curves.....	142
6.3	Outlook.....	143
7.	Annexes.....	145
7.1	Material and Methods.....	145
7.1.1	R-script file inputs for ATD extraction	145
7.1.2	R script for Integrals from Agilent 6560.....	146
7.1.3	Python Script for integrals Thermo EXACTIVE and CIU/CID plots	147
7.1.4	Sigma Plot Scripts for CIU	148
7.1.5	DynaFit Scripts.....	150
7.2	Native MS and Calorimetry.....	152
7.2.1	Malachite green aptamer - MS titrations	152
7.2.2	RiboFMN aptamer - MS titrations	168
7.2.3	Malachite green aptamer - ITC	177
7.2.4	RiboFMN aptamer - ITC	181

7.2.5	Aminoglycoside Aptamers 1NEM and 1TOB.....	185
7.3	IM-MS.....	188
7.3.1	IM-MS of quantified RNA aptamers	188
7.3.2	Tetracycline aptamer.....	192
7.3.3	CBA - UV-melting profiles	199
7.3.4	CBA - ¹ H-NMR of MN4 quinine complex.....	204
7.3.5	CBA - Circular Dichroism verify the B-DNA signature of dA constructs.....	205
7.3.6	CBA - CCS	208
7.3.7	Comparison pre-IMS tuning Agilent 6560.....	213
7.3.8	CIU/CID Breakdown diagram	214
8.	References	216

I. Abbreviations

NA – Nucleic Acid

DNA – Deoxyribonucleic acid

RNA – ribonucleic acid

ITC – Isothermal titration calorimetry

m/z – mass over charge ratio

ATD – Arrival Time Distribution

CCS – Collision Cross Section

ESI - electrospray ionization

nESI – nano electrospray ionization

IM-MS – Ion mobility mass spectrometry

r_x – Response factor

R_x – Relative response factor

NH_4^+ – Ammonium cation

TMA^+ – Trimethylammonium cation

TMAA – Trimethylammonium acetate

MALDI – matrix assisted laser desorption ionization

NH_4OAc – Ammonium acetate

NMR – Nuclear magnetic resonance

qNMR – Quantitative Nuclear Magnetic resonance

DSS – 2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt

MGA – Malachite green RNA aptamer

MG – Malachite green (ligand)

1FMN – RiboFlavinMonoNucleotide RNA aptamer

RiboFMN – RiboFlavinMonoNucleotide (ligand)

1NEM - Neomycin RNA aptamer

1TOB - Tobramycin RNA aptamer

CBA – Cocaine Binding DNA Aptamer

Q – Quinine (ligand)

TCA – Tetracycline RNA aptamer

TC – Tetracycline (ligand)

II. Résumé

Les aptamères sont des acides nucléiques capables de se lier sélectivement à un ligand ou à une famille de molécules. Les aptamères sont la partie sensible des riboswitches, qui sont des segments régulateurs de l'ARN messenger impliqués dans l'expression génétique. Les aptamères ont aussi des applications prometteuses comme sondes artificielles et capteurs. Pour ces technologies, il est crucial de comprendre comment la liaison se produit, de la quantifier, et de comprendre comment les changements conformationnels sont induits par les ligands. Les objectifs de cette thèse sont d'explorer l'applicabilité de la spectrométrie de mobilité ionique (IMS) couplée à la spectrométrie de masse (SM) native aux aptamères d'ADN et d'ARN, d'abord dans la quantification de liaison, ensuite dans la détection du changement conformationnel lors de la liaison du ligand.

Dans la première partie, nous avons évalué la détermination des valeurs de constantes d'équilibre de dissociation (K_D) par MS, en tenant compte des facteurs de réponse relatifs (R_x) des aptamères libres et liés. Les titrages en SM sont comparés, pour validation, avec la calorimétrie par titrage isotherme (ITC). Deux aptamères d'ARN sont pris comme modèles : l'aptamère du vert de malachite, largement étudié par ITC, et l'aptamère de la riboflavine mononucléotide, un cas réaliste d'ARN Mg^{2+} -dépendant pour la liaison du ligand. Nous avons observé que l'acétate d'ammonium et l'acétate de triméthyl ammonium conviennent à l'étude des aptamères et leurs complexes, et que les valeurs de K_D obtenues par ITC et SM native sont comparables. Les aptamères ARN de la néomycine et de la tobramycine ont été choisis pour tester la limite de détection en SM native. Nous concluons que la SM native est adaptée pour déterminer des valeurs de K_D comprises entre 50 nM et 30 μ M. La correction apportée par R_x est relativement modeste dans tous les cas, en suggérant que la liaison du ligand n'est pas associée à une différence conformationnelle significative lors de l'ionisation. Pour ces aptamères, nous concluons que l'hypothèse de R_x égaux est acceptable.

Dans la deuxième partie, nous avons évalué si le mécanisme de "liaison adaptative" des aptamères peut être révélé par IMS. À cette fin, en plus des systèmes énumérés ci-dessus, nous avons étudié l'aptamère ARN de la tétracycline et une série d'aptamères ADN capables de lier la cocaïne, pour lesquels le changement conformationnel par liaison du ligand est largement documenté dans la littérature. Pour tous les aptamères à l'exception de l'aptamère de la tétracycline, nous n'avons pas observé de différences significatives dans la conformation en phase gazeuse des ions liés aux ligands ou Mg^{2+} . Cependant, nous avons observé un changement significatif dans la mobilité des ions de l'aptamère de la tétracycline. Le Mg^{2+} (100 μ M) s'avère essentiel pour la liaison du ligand. Pour la série des aptamères de la cocaïne, même si nous avons observé dans des conditions douces de pré-IMS des ions compacts aussi bien pour les aptamères libres que pour les aptamères liés, une extension conformationnelle est visible à haute activation pré-IMS, bien révélée par l'état de charge 7-, qui suggère des réarrangements de phase gazeuse. Pour mieux étudier ces réarrangements, nous avons modifié les séquences avec des extensions dA, afin de comparer des systèmes ayant un nombre similaire de degrés de liberté sans modifier la structure cœur. Nous proposons également de nouvelles façons de présenter ces données, mieux adaptées quand la

dissociation du ligand, la perte d'adults et le dépliement d'ion arrivent dans les mêmes gammes d'énergie. L'augmentation graduelle de l'activation collisionnelle avant l'IMS, a révélé que l'énergie de dépliement est corrélée au contenu en paires de bases, ce qui suggère que les paires de bases sont conservées dans les structures en phase gazeuse. Nous avons également observé que le ligand se perd à des énergies inférieures à celles du dépliage.

Mots clés :

Spectrométrie de masse ; mobilité ionique ; ADN ; ARN ; aptamère.

III. Abstract

Aptamers are single-stranded nucleic acids capable to bind selectively to a ligand or to a family of molecules. Aptamers are the sensing part of riboswitches, which are regulatory segments of messenger RNA involved in gene expression. Aptamers are also promising artificial probes, sensors and stimuli-responsive elements. In the development of aptamer-based technology, it is crucial to understand how binding is occurring, to quantify affinities, and ligand-induced conformational changes. The objective of this thesis is to explore the applicability of native IM-MS to DNA and RNA aptamers to quantify binding and to detect conformational change upon binding.

In the first part, we evaluated the quantitative determination of equilibrium dissociation constants (K_D) by mass spectrometry (MS), and the necessity of including a correction for relative response factors of free and bound aptamers. We compared isothermal titration calorimetry and MS titrations to validate the quantifications. Two RNA aptamers were taken as models: the malachite green aptamer, extensively studied by ITC, and the riboflavin mononucleotide aptamer, a case of Mg^{2+} -dependent ligand binding. We observed that typical volatile electrolytes ammonium acetate and trimethyl ammonium acetate are suitable to study RNA aptamer binding, and that comparable K_D values are obtained from ITC and native MS. The neomycin and tobramycin RNA aptamers were chosen to test the limit of detection of native MS. We found that native MS is appropriate to determine K_D values in the range from 50 nM to 30 μ M. The relative response factor correction was relatively modest in all cases, suggesting that the ligand binding is not associated to a significant conformational difference upon ionization. For these aptamers, we conclude that assuming equal response factors is acceptable.

In the second part, we evaluated whether the aptamers' "adaptive binding" mechanism can be revealed by ion mobility spectrometry (IMS). To this aim, in addition to the systems listed above we studied the tetracycline RNA aptamer and a series of cocaine-binding DNA aptamers, for which the conformational change upon binding is reported in literature. For all aptamers except the tetracycline aptamer, we did not observe a significant difference in the shape of the gas-phase structure upon ligand or Mg^{2+} binding. However, a significant change was observed in tetracycline RNA aptamer's ion mobilities, at biologically relevant concentration of Mg^{2+} (100 μ M), and we found that Mg^{2+} is essential for ligand binding, in agreement with previous solution studies. For the cocaine-binding DNA aptamer series, although we observed similar compactness for the free and bound aptamers in soft pre-IMS conditions, a conformational extension occurs at high pre-IMS activation, best revealed by charge state 7-, suggesting gas-phase rearrangements. To better investigate whether the energetics of these rearrangements depend on pre-folding or on ligand binding, we modified the sequences with dA overhangs, to compare systems with similar numbers of degrees of freedom without altering the core structure. We also propose new ways of presenting the data, adapted to the cases where ligand dissociation, declustering and unfolding occur at similar voltages. The gradual increase of the pre-IMS collisional activation revealed that the unfolding energetics is correlated with the base pairs content, suggesting that base pairs are

conserved in the gas-phase structures. We also found that ligand is lost at lower energies than unfolding.

In summary, gas-phase compaction occur for both the free aptamers and bound aptamers, and memories of the solution-phase structures can only be revealed in some particular cases. However, the compaction towards similar shapes might constitute an advantage for the quantification, because molecular systems of similar shapes have similar electrospray responses. Consequently, native MS provides reliable estimations of K_D values.

Keywords:

Mass spectrometry; ion mobility; DNA; RNA, aptamer.

Unité de recherche

Acides nucléiques: régulations naturelles et artificielles (ARNA)
Université de Bordeaux, INSERM U 1212, CNRS UMR 5320

Institut européen de chimie et biologie (IECB)
2, rue escarpit 33607 Pessac (France)

IV. Acknowledgments

In first place, I would like to thank Pr. Calorine Tokarski, Dr. Sarah Cianferani, Dr. Fabien Darfeuille, Pr. Claudia Sissi and Pr. Frank Sobott to have accepted to evaluate this work.

I am very thankful to my mentor Dr. Valerie Gabelica for the great opportunity of this PhD. The encouragement to open collaborations and present my work at global level, was a precious gift on both professional and personal level. Your deep intellectual honesty, scientific rigour and love for constructive messages in science, guided me to coherently shape the fruits of my work in this manuscript and helped me in developing my *forma mentis*.

A sincere thank you goes to all MetaRNA fellows, Pr. Matthias Heinemman and all the P.I.s of this great ITN. Our meetings during these years were simply great and helped me through this journey on both professional and personal side.

Within the lab, I am in debt to Sandrine and Clémence, my two *colleagues en Or*, for the splendid atmosphere that I felt since my very first day here with you. I am very thankful to Adrian (our definitive cicerone of the lab), to Dr. Frédéric Rosu (your passion for mass spectrometry and physical chemistry is simply amazing) and to our post-docs Massimiliano, Josephine, Nina, Jorge, Steven, Anirban and (our associate professor) Eric. The numerous conversations have been a blossom of ideas to plan new (better) experiments and to keep the right critical view on my work. I cannot thank enough Dr. Jean-Louis Mergny, Dr. Carmelo Di Primo, Dr. Samir Amrane and Dr. Axelle Grelard for the very useful discussions and your technical help.

With Léonie and the 2017-2018 JJs committees friends I shared funny moments and a great work together in conference organization. Thank you very much guys.

I also thank Nassima, Laura, Britta, Pedro, Aurore, Oscar, Julien and all the IECB people that I have encountered during these years.

A lovely and warm thank goes finally to Serena, that by my side experienced the lights and the darks of a PhD. Your sustain throughout these years helped me in deliver this work to an end.

1. Introduction

1.1 Nucleic acids structure

One can describe nucleic acids structure in four level of complexity. The primary structure is the linear order of nucleic acid building blocks. The secondary structure is a two-dimensional representation of interaction between the building blocks (base pairing) that define some structural elements (e.g. double helices). Nucleic acids tertiary structure arise from the arrangements of two or more secondary structure in the three-dimensional space, through contacts called “tertiary interactions”. Finally the quaternary structure rises from the interaction of tertiary structures, often called domain because of their large size.

To visualize nucleic acids structure, all its diverse components and how they interact, is not easy. This chapter will introduce the basic principles of nucleic acid structural elements.

1.1.1 Primary Structure (Sequence)

Nucleic acids are linear sugar-phosphodiester biopolymers constituted by four nucleotides. Nucleic acid **primary structure** (or sequence) is defined as the order of the nucleotides. By convention this sequence is ordered from the 5' to the 3' end of the polyribose phosphodiester chain (Figure 1).

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) differ in sugar structure in that the 2'-hydroxyl group is missing on deoxyribose of DNA (Figure 1) nucleotides. The nucleoside composition (bases) is also different: the pyrimidine base thymine (thymidine (T)) is present in DNA and uracil in RNA, while Adenosine (A), cytosine (C), guanine (G) are shared in both DNA and RNA. The sequence is conveniently written with the letter of each nucleotide as d(GATC...) for DNA and r(GAUC...) for RNA.

The key interactions are the covalent bonds, such as the glycosidic bond that hold the nucleobase on the sugar scaffold and the phosphodiester bond that create the sequence. The Presence of the 2'-OH hydroxyl group makes RNA capable of three specific interactions involving the Sugar Edge (Figure 2), but RNA is less chemically stable than DNA^[1]. This makes DNA a better candidate for large genetic information, whereas RNA is more adapted to serve as dynamic (and transient) actor into transient cellular conditions.

A sequence of DNA or RNA not involved in any base pairing is commonly defined as **single stranded** region. Often these unpaired regions may occur, but not only, close to the sequence ends ^[2].

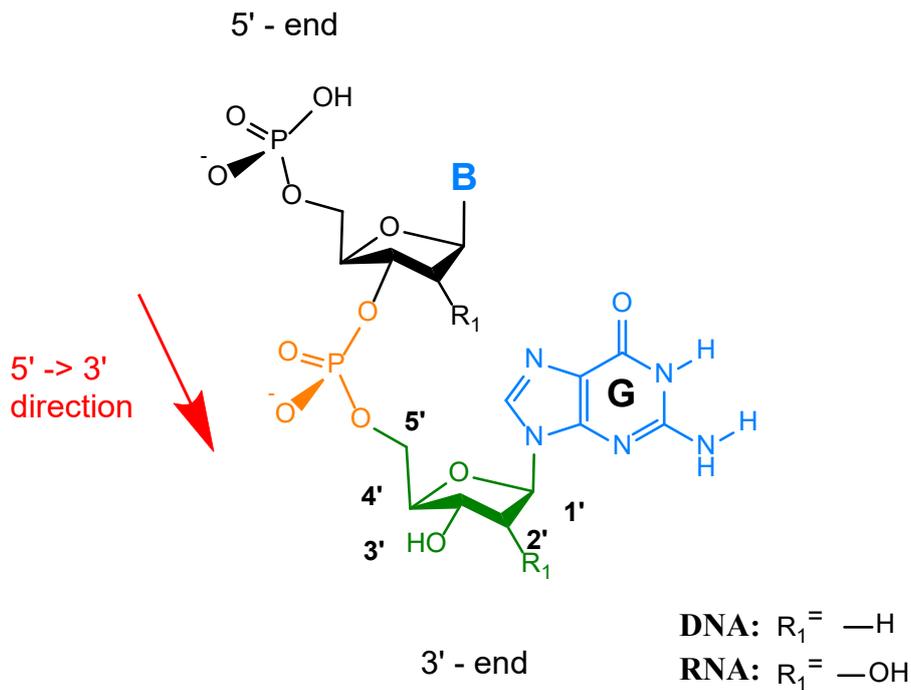


Figure 1 - Nucleic Acids sequence. Nucleoside is formed by a nitrogenous base (**blue**) and its sugar scaffold (**green**), ribose in RNA and deoxyribose in DNA. Nucleotide is the 5'-phosphorylated (**orange**) form. For convention, the direction of the chain is from the 5' to the 3'-end.

1.1.2 Secondary structure

Intramolecular non-covalent interactions among nucleotides define the secondary structure of the nucleic acid, normally represented in two dimensions. Base pairing is classified in three patterns (Figure 2A):

- The Watson-Crick (WC) “canonical” base pair (C=G, A-T in DNA and A-U in RNA) is the most common pattern. Double helices (Figure 3) are constituted by anti-parallel WC base pairs^[3].
- The non-canonical (non-WC) Hoogsteen base (e.g.: when the N7 and C6 edge is used to pair with the WC edge of the other nucleotide^[4]). This pattern is involved in secondary structures such as G-quartets^[5] and triplets^[6].
- The non-canonical Sugar edge base pair where the 2'-OH group of RNA nucleotides, can participate as an H-bond donor or acceptor.^[7] These structures are thermodynamically driven by the base stacking, which stems from the hydrophobic effect, but the specificity comes from the h-bonding pattern.

The non-canonical base pairing is an important difference between DNA and RNA helices. The RNA ability to adopt many conformations is due to the extra —OH group on the 2' C ribose, opening to various and diverse secondary structures (Figure 4).

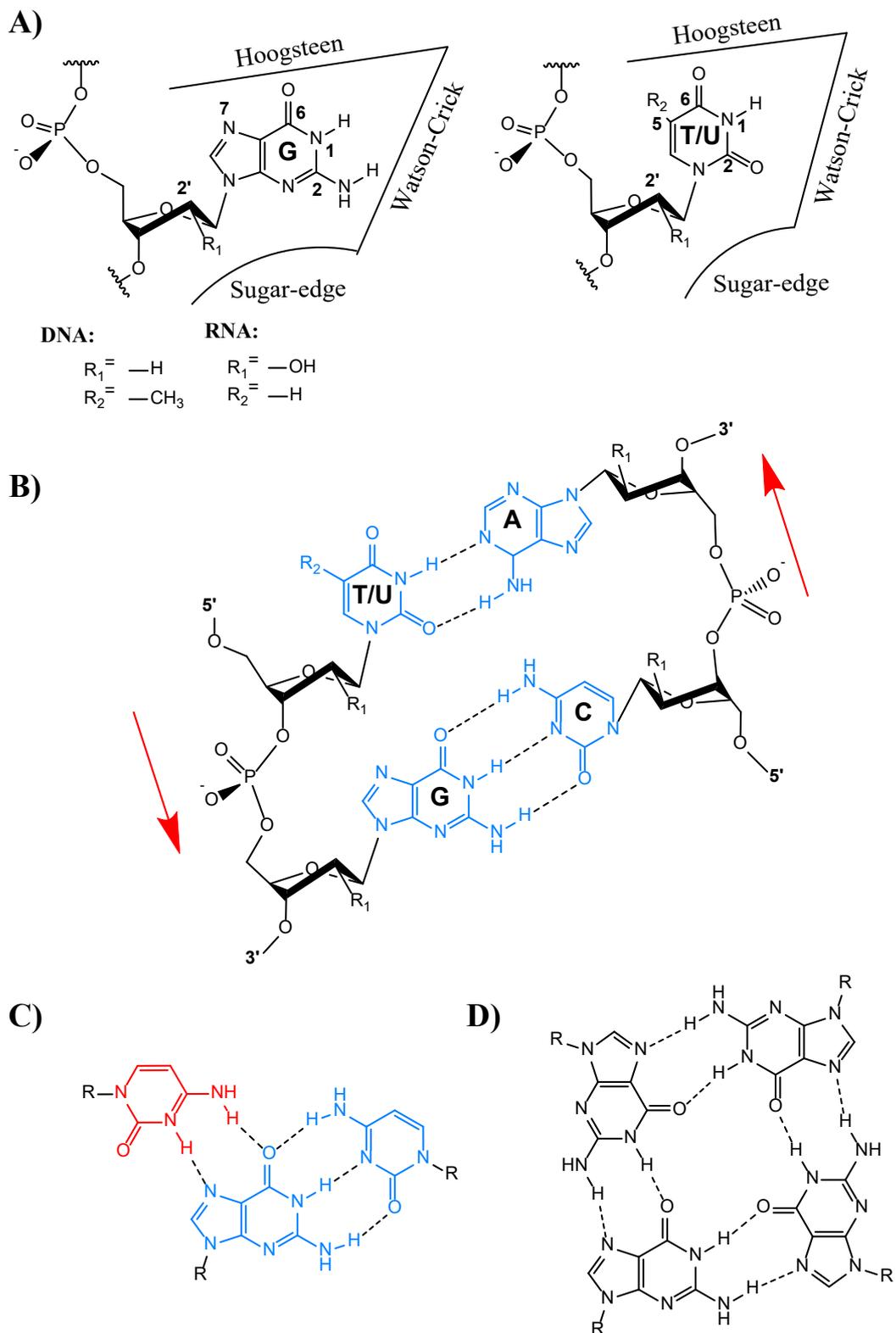


Figure 2 - **A)** Base pairing mode for purines (Guanine) and pyrimidines (Thymine or Uracil), WC, Hoogsteen, Sugar edge readapted from ref.^[8]; **B)** Watson-Crick "canonical" base pairing in double helix (e.g. B-DNA); **C)** Example of base triad, here a pyrimidine (C•G-C) motif. Here **light blue** G-C are part of a duplex and the **red** C belongs to the third strand; **D)** Guanine quartet. Ribose ring is represented in C2'-endo conformation (common in B-DNA).

Helices: From base stacking to three-dimensional structures

The conventional representation of WC pairing secondary structure in Figure 2B describes some structural elements like helices^[9]. The WC base pairing is accompanied by the stacking of the aromatic nucleic acid base, that brings a characteristic turn of nucleic acid helices^[10].

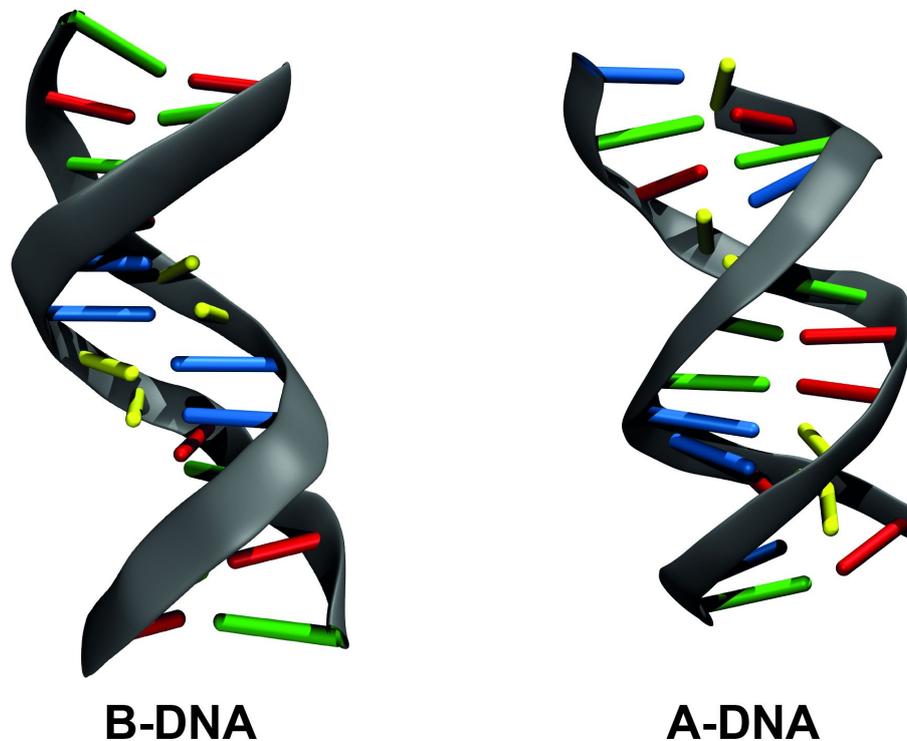


Figure 3 – Models of DNA double helices. The B form is a right-handed double helix of 10 nt per turn (0.34 nm per bp, 3.57 nm per turn, \varnothing 2.0 nm). The more compact A form of 11 nt per turn (0.26 nm per bp, 2.86 nm per turn, \varnothing 2.3 nm)^[11]. Both strands are antiparallel. Nucleobases color code: **A-Blue**; **G-Green**; **T-Yellow**; **C-Red**;

A double-helical tract is often referred to as “duplex region” or “stem”. The B-form duplex is the most common in DNA under physiological conditions (Figure 3). It forms when two single strands associate into a right-handed double helix structure wherein the opposing bases are paired along the WC edge (Figure 2A,B). A single turn of B-DNA about the axis of the molecule contains ten base pairs. The distance spanned by one turn of B-DNA is 3.4 nm (10 nt). The width (helical diameter) of the double helix in B-DNA is 2 nm. The A-form duplex is a more compact helix of 11 nt per turn with the base pairing aligned with helix axis. A-form duplex is found almost only in RNA^[12] or in dehydrated DNA^[13].

Other helices also exist. Triplex helices are formed when the base pairing follows triplet scheme (Figure 2C), for example when a duplex of two WC complementary polypurine and polypyrimidine strands, are bound by a third polypyrimidine strand via Hoogsteen intercalations. G-quadruplexes are formed in guanines-rich DNA or RNA^[14]. G-quadruplexes are organized by the consecutive stacking of G-quartets (Figure 2D). The base stacking across piled quartets is furthermore stabilized by the coordinated cation between the guanines^[5a, b]. G-quadruplex structures are found e.g. DNA telomeres, telomeric-transcribed RNA^[15], DNA aptamers^[16] and in fluorescent RNA aptamers^[5d, 17].

Loops

Loops are interruptions of a base pairing that lead to a local single stranded portion. "Interruption" may suggest a lack of structure but only the canonical WC pattern is interrupted. The unpaired bases in these portions can adopt various specific conformations and establish non-canonical base pairing or non-covalent interaction with other molecules (e.g. nucleic acids, proteins, small molecules, ions). Because of this variety, ligand binding is often detected in loops (Figure 4). The different types of loops are:

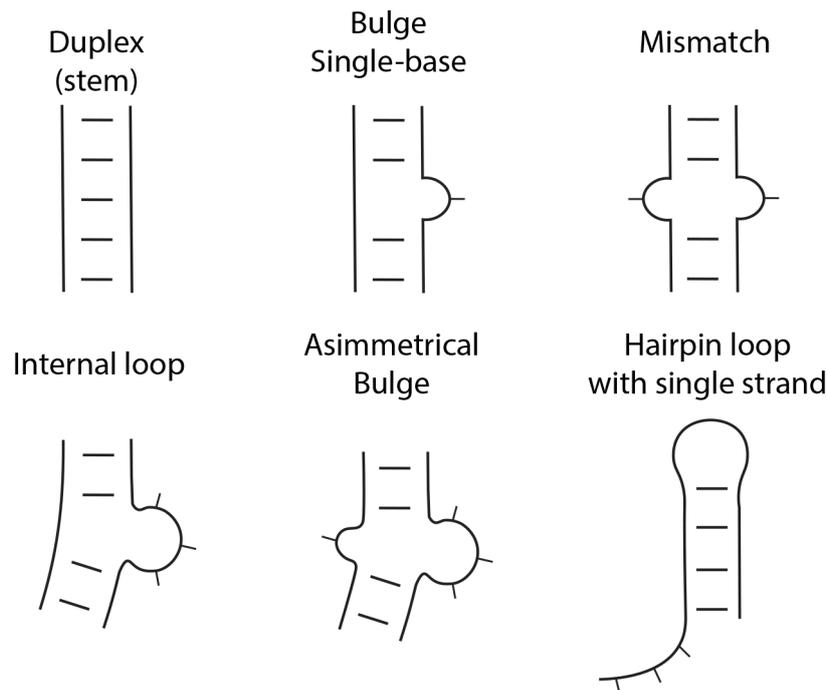


Figure 4 - Scheme of RNA secondary structure motifs.

Hairpin loops: A hairpin loop is a single strand portion that links the 5' and 3'-ends of a double helix, for instance when a single strand sequence is self-associating into a duplex region (Figure 4). This structural element is present in DNA but more common in RNA. It varies in length from 2 to 14 nucleotides. Tetraloops are the most frequent and studied^[18], e.g. the GNRA type (where N and R are exchangeable nucleotides)^[18]. Notable examples are the T-loop in ribosomal RNA (rRNA)^[19] and D-loop motifs of tRNA^[20].

Internal loops: When a double helix is not perfectly paired all the way, we have an internal loop. The internal loops can be symmetric or asymmetric, if the two opposing strands differ in length. Two special case of internal loops exists: **bulges**, when a single nucleotide is unpaired, and **mismatches** when this unpairing involves two opposed nucleotides on each strands^[21].

1.1.3 Tertiary structure

Kissing loops

When two hairpins interact with the loop nucleotides by canonical base pairings, we have “kissing” loops. This pairing starts from a minimum of two up to several base pairs^[22]. The formation of kissing loops is favoured by others structural elements or tertiary interactions (e.g. junctions) that put in contact the two hairpins (Figure 5A). An example of this kissing-loops is the HIV-1 dimerization site^[23].

Pseudoknot

A pseudoknot is formed when a hairpin loop is complementary (by canonical WC) to a single-stranded distal region of the sequence^[24] (Figure 5B). This element often shows coaxial stacking between the new hairpin and the new complementary portion^[25] and triplex characteristic (Figure 5B - orange-violet contact)

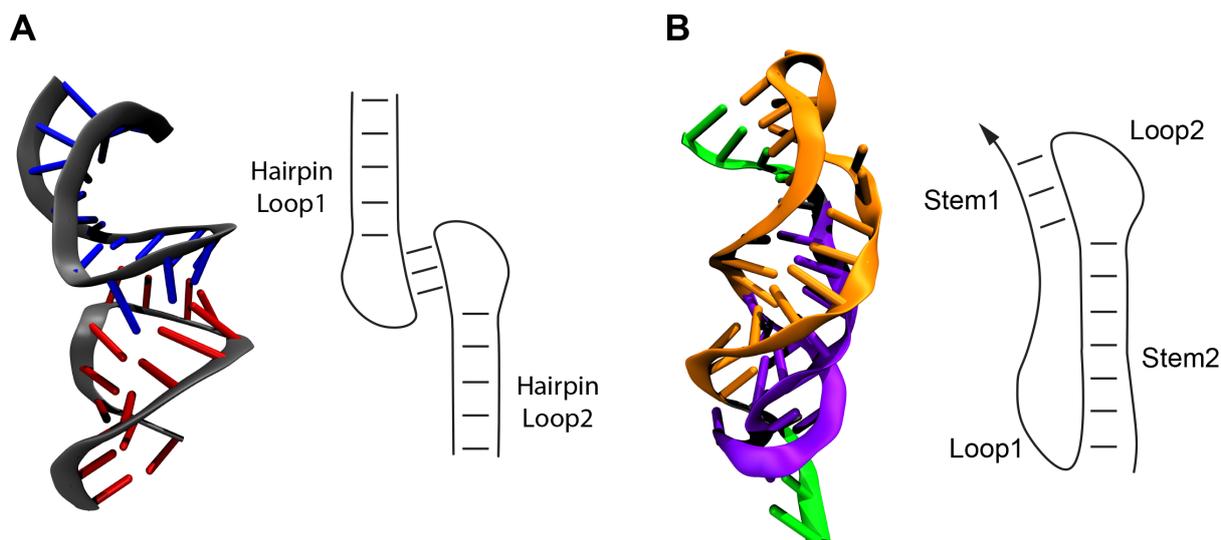


Figure 5 – A. Crystal structure of kissing loops (2jlt.pdb) and B. pseudoknot (1ymo.pdb) with the relative 2D representations. To notice, the triplex characteristic of the middle part of the pseudoknot (orange-violet contact).

Junctions (Multiloops)

When three or more duplexes intersect, a junction is formed (Figure 6). Single strand portions connect the helices^[26]. Three-way and four-way junctions are common in RNA^[27] and often promote the coaxial stacking of the duplexes. The increased base stacking of such arrangements^[10] contributes to the formation of other tertiary structures (some times called “long-range” contacts) such as the hairpin-duplex contact next to the three-way junction in tetracycline RNA aptamer^[28] (Figure 8B).

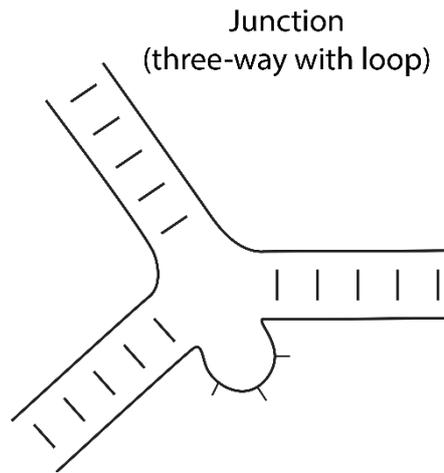


Figure 6 - 2D representation of a ternary junction. Often unpaired loops are located in the junction. A tridimensional example is visible in Figure 8B.

1.1.4 Metal cations in nucleic acid structure

Nucleic acids are large poly-anions and all the phosphates of the sequence require a surrounding ionic atmosphere. These ions have many roles in tertiary structure definition. Three cases can be outlined (Figure 7):

1. "Diffuse" ions ensure a global charge counter balancing. The hydrated nucleic acid and the hydrated ion interact with their reciprocal second (outer) coordination sphere of water molecules. The diffuse ions are not tightly bound and can readily exchange.
2. When the ions penetrate in the reciprocal second coordination sphere, ion and nucleic acid interact through some shared water molecules.
3. Finally, nucleic acid and ions can be in direct contact when they are in the reciprocal first (inner) coordination sphere contact and form an ion pair.

The direct contact with the nucleic acid may help in a specific charge balancing (Tetracycline - Figure 8B)^[28-29] or in tight structural contacts (as G-quartets)^[30].

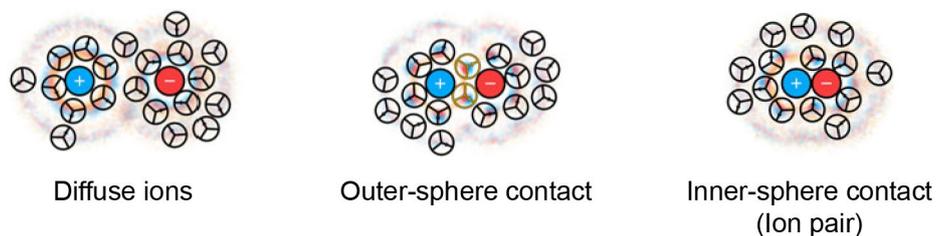


Figure 7 - Ionic contacts in function of the solvation sphere. Water molecules are represented by circled Y. Inner(first) and outer(second) coordination sphere are represented by the concentric shades. Readapted from ref ^[31].

1.1.5 Structure-function relation and tertiary structure

The role of each secondary structural element is clearer if we observe the whole RNA structure of two valid examples: Malachite green RNA aptamer and tetracycline RNA aptamer x-ray crystal structures in figure Figure 8.

Malachite green aptamer is an hairpin-loop shaped RNA with an asymmetrical bulge within the duplex region Figure 8A. The binding site for the ligand (**light blue**) occurs at this point. Here the “lack” of base pairing associated to the bulging (**orange**) is functional to host the ligand and preserving the hairpin shape of MGA^[32].

In tetracycline RNA aptamer the three-way junction (**Red**) allows to bend the hairpin loop (**cyan**) towards the adjacent duplex forming the ligand (shaded **green**) binding site through a triplex-like contact with the duplex. Metal interactions with Mg²⁺ also contributes to structure stabilization. One Mg²⁺ is chelated metal by the ligand and it is paired with one phosphate of the hairpin loop. The other Mg²⁺ cations acts as both outer-bound and diffused around the aptamer backbone to overall structure stabilization^[28].

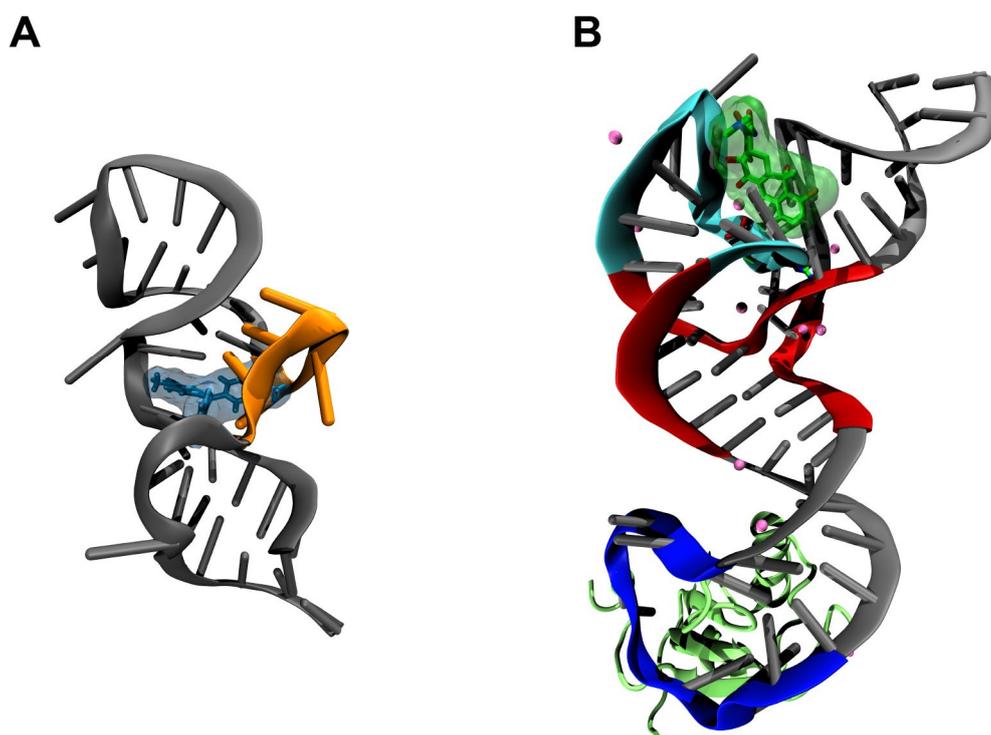


Figure 8 – Examples of structural elements in RNA aptamers. **A.** Malachite green aptamer is composed by a hairpin with an an internal loop (**orange**) mid-way in the duplex region, where ligand is hosted (**light blue**); **B.** In the tetracycline aptamer one finds a three-way junction (**Red**), a hairpin (**cyan**), a triplex with the adjacent duplex forms the ligand binding site (**green**) through a triplex-like contact. Several Mg²⁺ ions are highlighted in **pink**. A specific loop (**blue**), that binds selenomethionyl U1A protein (**lime**), has been added for crystal growth.

1.2 Aptamers and SELEX

Aptamers are short DNAs and RNAs capable to interact with their target molecules because of a unique three-dimensional structure. For this they are also called “chemical antibodies”, with the advantage of no immunogenicity, relatively cheap production (they can be chemically synthesized) and longer storage time compared to their protein-based counterparts^[33]. The targets can be very diverse, from metal cations^[34] small molecules^[28, 35]/ions to proteins, nucleic acids^[36] and cells^[37]. Natural aptamers also exist in riboswitches, a genetic control element, where they act as the recognition element for several metabolites^[38].

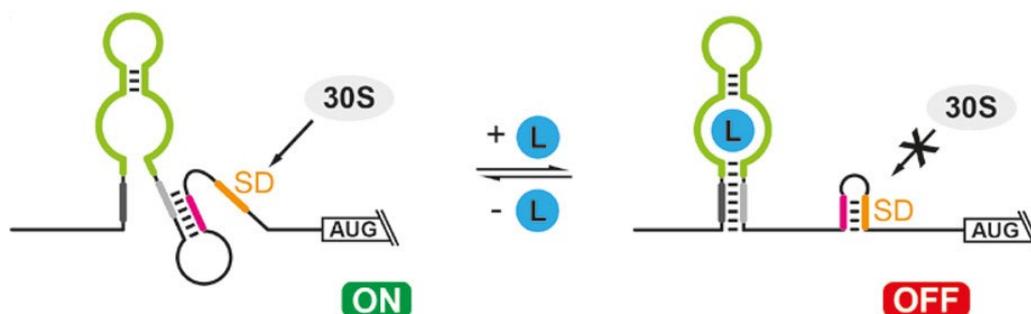


Figure 9 – Example of gene regulation in bacteria using engineered riboswitches. The aptamer (green) is separated from the Shine-Dalgarno (SD) sequence (magenta) by a hairpin-loop structure. Gene translation is active. Once added, the ligand (L) induces a conformational change into aptamer structure and gene translation is inactivated. SD sequence becomes paired into an hairpin and ribosomal unit 30S cannot bind. Readapted from^[39].

Aptamers often bind to their targets via an “adaptive binding” mechanism. This term aims to mark the difference compared to rigid and pre-structured ligand receptors such as often found in proteins (e.g. enzymes). Aptamers binding site Often the binding site comprise unpaired regions (e.g. loops), which are disordered and acquire a defined conformation only folding around the ligand and sometimes others co-factors like metal cations (e.g.: Mg^{2+} , K^+). Physically, one can see the “adaptive binding” mechanism as conformational selection. The “unstructured” free aptamer may populate many slightly different conformers at equilibrium by binding preferentially to one of these conformations, thereby displacing the conformational equilibria and seemingly forcing the aptamer to adopt only one structure, or a very limited population of shapes. (an example can be found in chapters 5 and 5.6).

Aptamers are artificially obtained *in vitro* by a combinatorial chemistry method called SELEX (Systematic Evolution of Ligands by Exponential enrichment)^[40]. Here we introduce its general principles.

The method starts with a large initial library of randomized DNA (or RNA) sequence, up to 10^{15} unique sequences. Each sequence is composed by a random region of arbitrary length (normally from 22 to 200 nt). To allow PCR (or RT-PCR) amplification, this region is flanked by two conserved primers. The large number of unique sequences is required to have the widest

sequences coverage (# unique sequences) and structural diversity (Length of single sequence...more likely to form complex structures), during the selection step.

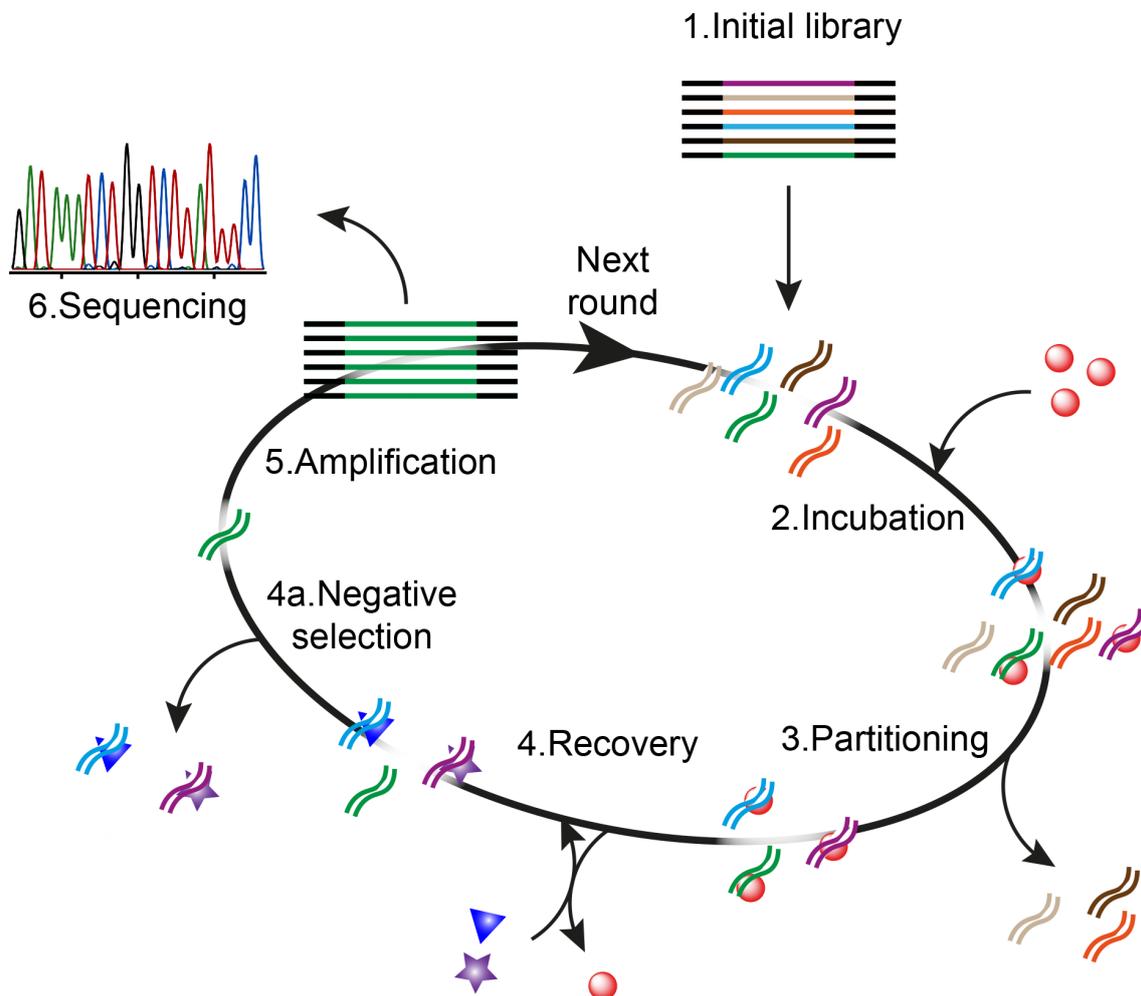


Figure 10 – Scheme of SELEX procedure. Target ligands are represented as red spheres; Negative targets are blue triangle and purple star.

The procedure follows an iterative procedure commonly divided in six steps:

1. Initial library design;
2. Incubation, the library and the target are incubated in the desired conditions (buffer composition, temperature, ...) for a given time, to allow the most adapted structure to bind the target molecule. Normally the target is immobilized (e.g. on a solid phase chromatography), so that non binding sequences can be easily removed;
3. Partitioning, putative aptamers are separated from the remaining pool (low or no affinity sequences), for example by affinity chromatography or other separation techniques (e.g. capillary electrophoresis, ligand decorated magnetic beads);
4. Recovery, chemical or physical dissociation of putative sequences from the target molecule. For instance, if the ligand is bound to a solid phase (e.g.: chromatography

column beads) and putative sequences are released via saturated ligand mobile phase elution or denaturation;

- 4a. Negative-selection, undesired non-specific sequences are eliminated using one or more negative-target molecules. This step can be added before re-amplification and sequencing (optional).
5. Amplification, the enriched library of candidate sequences is amplified via PCR (DNA library) or RT-PCR (RNA library); This library can initiate a new selection round.
6. Sequencing, enriched library is analyzed to find the best binders by Sanger or next-generation sequencing.

The procedure is re-iterated using the amplified enriched pool as starting library. Normally an enrichment is obtained within 15 rounds^[41].

Some limitations however are present: a standard method of SELEX does not exist. All the variations today explored on the first publications procedure proposed in Gold, Joyce and Szostak labs^[40, 42], were made to improve one or more aspects of the procedure depending on the function of the target or the conditions of use.

Choices in selection buffer composition, concentrations and temperature are likely to favour some secondary and tertiary structures compared to others. Thus, the reproducibility of the SELEX could be affected ^[43].

On the target side, enrichments on large size targets often result in lower K_D , compared to small molecule targets^[44]. The chemistry of targets may also affect the the selection process, as not all target are suitable: very hydrophobic (poor solubilization) or strongly negative-charged targets (strong repulsion forces aptamer-target) could be difficult or not amenable to induce an enrichment.

In general, common bottlenecks are the high time demand of the SELEX procedure and the characterization of aptamer functionality. Given the wide range of applications, a standardized method of analysis is not available ^[45].

1.3 Standard state and equilibrium constant

The parameter that describes the binding strength of two interacting partners (i.e.: an aptamer to its ligand) is the equilibrium constant of the binding reaction. This concept comes from thermodynamics. The direction of a process (i.e. binding reaction) is dictated by the changes in its Gibbs free energy. For a given chemical reaction as eq.(1), we can define this change, ΔG , as the first derivative of Gibbs energy with respect to the extent of reaction. The actual ratio between reactants and product activities, the reaction quotient Q , will tell us if the reaction is going rightward ($\Delta G < 0$) or leftward ($\Delta G > 0$)



$$\Delta G = \Delta G^o + RT \ln Q \quad (2)$$

$$Q = \frac{\text{products activities}}{\text{reactants activities}} = \prod_i a_i^{\nu_i} \quad (3)$$

At the equilibrium the ΔG is zero (no net reaction). This gives a useful relationship between is the standard Gibbs energy of reaction ΔG^o and the composition at the equilibrium:

$$\Delta G^o = -RT \ln K \quad ; \quad K = (\prod_i a_i^{\nu_i})_{equilibrium} \quad (4)$$

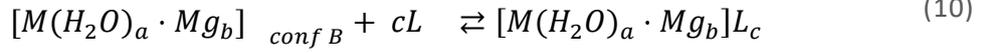
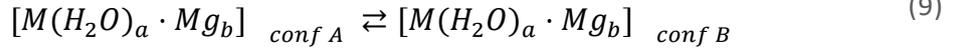
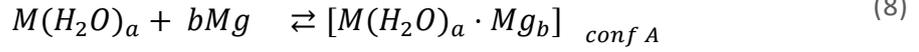
Activities (a_i) are dimensionless, thus the *thermodynamic* equilibrium constant K is a dimensionless number. However, in the biophysics and biochemistry literature, K values are often reported in molarity. This is an approximation where activities are replaced by concentrations, eq (5), assuming that all activity coefficients γ_i are equal to 1 (ideal solute), so the equilibrium constant is expressed in molar concentrations, eq (6).

$$a_i = \gamma_i M_i \quad (5)$$

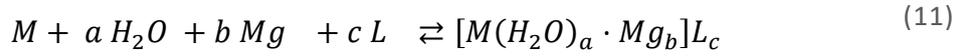
$$K \approx K_{M_i} \quad (6)$$

In this thesis the reference state is defined as $p=1$ atm, $a_i=1$ and pH=7, more appropriate for biomolecule description, and the approximation (6) is used for consistency with most of publications on nucleic acids, where solutes are considered ideal (same interaction energy between solute-solute and solute-solvent).^[46]

The concept of cumulative equilibria is often implied in biophysics. Thermodynamic functions and binding constants reported as “apparent” (i.e. ΔG_{app}^o) convey the message the reaction under study is instead the combination of many equilibria. For biomolecules as DNA or RNA we may imagine the binding of a ligand as the result of consecutive reactions like hydration of the molecule, specific Mg^{2+} binding, biomolecule conformational change to an active form and ligand binding:



Here the charges were omitted for clarity. The sum of these reactions is:



with a global cumulative equilibrium constant $\beta_4 = K_1K_2K_3K_4$, the ΔG of reaction in eq (11) could be written as

$$\Delta G = \sum_{n=1}^4 \Delta G_n^o + RT \ln K_n = \Delta G^o + RT \ln \beta_4 \quad (12)$$

The equilibrium constant β_4 is also influenced by the conditions such as temperature, pressure, ionic strength, Mg^{2+} concentration, pH, reactants and product concentrations (i.e.: when the assumption of ideal solute do not apply). For this, is hard to tell about the merit of a technique when it is taken alone. Even more challenging is to compare K_D obtained from different techniques when relying on literature values, especially if experimental conditions have to differ. That's the case of aptamers. One of the challenges of this work is the comparison of K_D obtained from different techniques, mass spectrometry and isothermal titration calorimetry.

1.4 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) allows one to directly measure the heat associated to a given reaction, and is frequently used for aptamers. A typical experiment consists in the successive injection of small volumes of titrant (i.e.: ligand) into the titrate solution in the sample cell (i.e.: nucleic acid or protein), and measuring the heat.

A

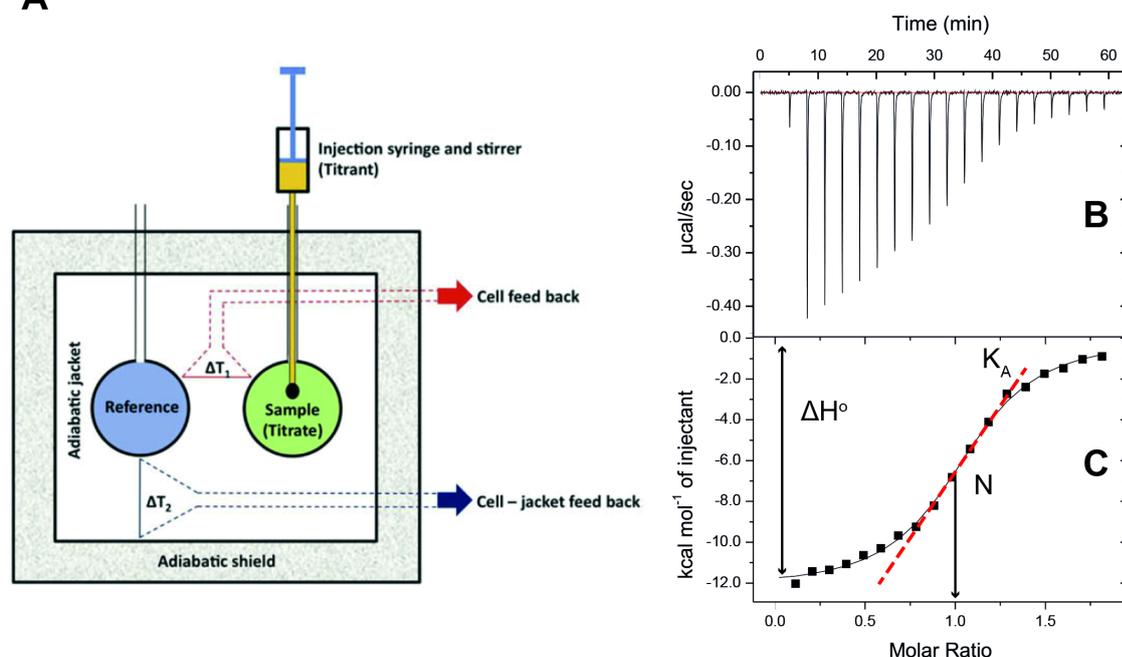


Figure 11 – **A**. Cells of an ITC calorimeter into an adiabatic jacket. The signal baseline is provided by the heater power applied to keep $\Delta T_1 = \Delta T_2 = 0$. Temporary variations in ΔT_1 trigger the heater compensation (heat pulse). These deflections from the power base line are plotted as a function of time and integrated to obtain reaction heat (readapted from [47]); **B**. Heat pulse vs time or thermograms; **C**. Wisemann isotherm fitting of integration points expressed as kcal mol^{-1} of injectant vs molar ratio of titrant and titrate.

The instrument records the amount of power needed to keep the sample cell at the same temperature as the reference cell (heat pulse), into their adiabatic jacket. The instrumental output is a heat pulse vs time [$\mu\text{cal s}^{-1}$]. This pulse corresponds the sum of all the heats measured upon ligand injection. In addition to reaction heat, other contributions are the partners dilution, eventual buffer dilution (in case of buffer mismatch) and additives within. Heats of dilution are taken into account by subtracting a control run with buffer into the reference cell. We will annotate thermodynamic function states and equilibrium constant with “apparent” (app) subscript because they will still depend on how M, L and ML interact with buffer components (see section 1.3).

Thanks to a direct measurement of the heat generated or absorbed, ITC is particularly convenient to study binding reactions on biomolecules through the enthalpy of reaction. Assuming the simplest case of macromolecule (M)-ligand (L) in molecularity 1:1, we have the binding equilibrium:



The heat associated to reaction (13) is related to the ligand concentration by the Wisemann isotherm (14), where $\frac{dQ}{d[L]_t}$ is the heat per ligand injection, L_R is the ratio of $[L]_t/[M]_t$ (total L and M concentration), and the parameter $a = NK_A[M]_t$, where N is the number of sites^[48].

$$\frac{dQ}{d[L]_t} = \Delta H V_0 \left[\frac{1}{2} + \frac{1 - L_R - a}{2\sqrt{(1 + L_R + a)^2 - 4L_R}} \right] \quad (14)$$

Points on the thermogram are the integrals of each heat pulse. By non linear least squares curve fitting, one can extract as parameters (assuming a 1:1 binding model) the enthalpy (ΔH_{app}) of binding, reaction stoichiometry (inflection point, N) and association constant ($K_{A,app}$)^[49]. Finally, the entropy (ΔS_{app}) and Gibbs free energy (ΔG_{app}) can be derived, giving a complete thermodynamic characterization of binding^[49-50]

$$\Delta G^o_{app} = -RT \ln K_{A,app} = RT \ln K_{D,app} \quad (15)$$

$$\Delta G^o_{app} = \Delta H^o_{app} - T\Delta S^o_{app} \quad (16)$$

In summary, ITC is a label-free technique that provides information on the thermodynamic parameters of the observed system and allows to measure K_D values in the range of 10^{-6} – 10^{11} ^[51]. However, some constraints should be considered^[50a, c]: additional heats may contribute significantly to the measured heats (e.g.: buffer differences); and the model of binding must be chosen *a priori* by the user.

1.5 Mass spectrometry

This thesis will evaluate mass spectrometry as a complementary method to evaluate aptamer binding stoichiometries, affinities and ligand-induced conformational changes. Mass spectrometry (MS) measures the mass-to-charge ratio (m/z) of ions generated from an ionization source. As ion movement is influenced by electric or magnetic fields, analytes are manipulated in the gas phase.

In the 1980's, innovative ionization methods were developed to analyse large synthetic polymers and biopolymers such as proteins and nucleic acids: electrospray ionization (ESI) from liquid samples, by Fenn *et al.*^[52], and matrix assisted laser desorption ionization (MALDI) from co-crystallized samples, by Karas *et al.*^[53].

ESI is nowadays the most popular ionization method for biomolecules analysis, because it handles liquid samples (suitable for most of proteins and nucleic acids) and is capable to generate multiply charged ions, easier to analyse by the mass analyser (many charges reduce the mass-to-charge ratio). When sample quantities are limited or tolerance for elevated salt concentration is required^[54], a miniaturized version of ESI, called nano ESI (nESI) can be used^[55].

The physical principle of ESI is summarized in Figure 12. The solution is delivered into a capillary in front of the mass spectrometer, usually at atmospheric pressure. A voltage is applied between the capillary and the entrance of the instrument. The charges in the solution migrate, and charges of one sign accumulate on the surface of the droplets pending at the extremity of the capillary, shaping the liquid into a Taylor cone from which smaller charged droplets are emitted. In electrospray, the flow rates are typically $> 1 \mu\text{L}/\text{min}$, and a coaxial flux of gas (usually nitrogen) comes around the capillary and/or from the entrance of the instrument (depending on instrument design) to facilitate solvent evaporation.

In their path towards the entrance, the charged droplets lose progressively their solvent until they reach a critical point where the charge density is too high for the surface tension to cope with it. Here ions are generated, and three different mechanisms are proposed:

- In the ion-evaporation model (IEM)^[56] the charged droplet reduce its charge ejecting ions, to contrast the rising charge density due to solvent removal. This “evaporation” of molecules as ions is accepted for small molecules and metal ions (Figure 12B).
- In the charged-residue model (CRM)^[57] the solvent evaporation and fission events produce final charged droplets that contain only a single molecule complex. When the solvent is removed, the residual charge is transferred to the molecule (Figure 12C).
- The unfolded biopolymers can be expelled in gas phase following the chain ejection model (CEM)^[58]. While exiting the droplet, excess charges are transferred to the chain end of the biopolymer until complete expulsion of the chain as gas phase ion (Figure 12D).

CRM and CEM models are commonly accepted for folded (e.g. globular proteins) or unfolded molecules (e.g. disordered proteins), respectively. For DNA and RNA, the situation is less clear,

as both folded and unfolded strands can be ionized by the CRM if analysed at high ionic strength^[59].

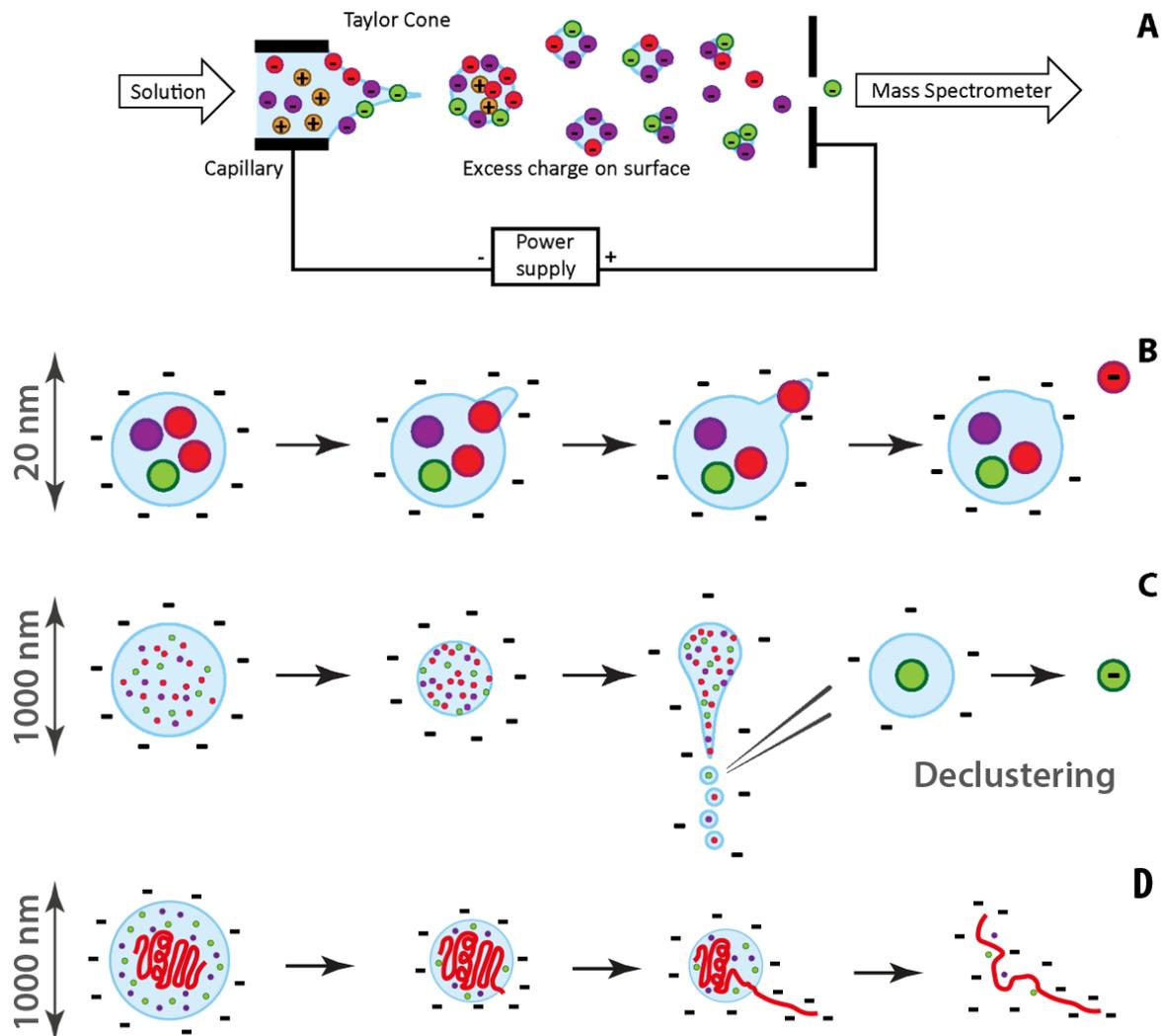


Figure 12 - Electro spray ionisation in negative mode. **A)** Scheme of the Electrolytic cell imposed between source capillary and mass spectrometer inlet; Ionization modes following **B)** ion evaporation model (IEM), **C)** charge residue model (CRM) with the charged droplet fission or **D)** chain ejection model (CEM). Solutes are represented as colored dots or chains.

1.5.1 Native mass spectrometry

When the analyte is a biomolecule (e.g. a nucleic acid) and one wants to analyse it by mass spectrometry while preserving the natural non-covalent interactions, particular care should be taken in sample preparation: ionic strength, additives and pH are carefully chosen to mimic the biological habitat of the analyte.

Commonly, NaCl/phosphate buffers used for *in vitro* studies are replaced with volatile salts (e.g. NH₄OAc) from 100 to 300 mM in total. The aim is twofold: **I.** volatile buffers avoid ion suppression and signal broadening due to clustering (p.18); **II.** this range of concentrations

mimics the ionic strength of physiological solutions, e.g. extracellular [NaCl] is 145 mM^[60]. Starting solutions in MS-friendly salts (i.e.: NH₄OAc 150 mM) have the advantage to be pH neutral^[61], but it is not buffered: that's a disadvantage.

Thus, the adjective “native” in mass spectrometry means that the sample is prepared to keep the native folded state of the analyte *before* the ionization (in solution) and preserve the relevant interaction that characterize this state until the ion detector.

1.5.2 Ion activation in mass spectrometry

An ion that moves in vacuum is an isolated system with no energy exchanges with the environment. A way to activate an ion (increase the ion internal energy) is called “collisional activation”. Via inelastic collisions with neutral gas (e.g. in “high pressures” portions like pre-IMS of Agilent 6560, collision quadrupoles), a fraction of the kinetic energy is converted in ion internal energy. This increase of internal energy lead to an excitation of the ion that can redistribute the converted energy in all the vibrational states availables, altering the balance between repulsive (Coulomb repulsion, ...) and attractive forces (intramolecular H-bonds, Van der Waals, ...)^[62].

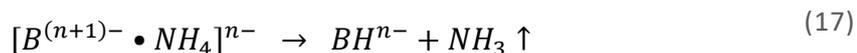
Commonly one can act on the strength of collisions in two ways:

- i) On the ions, varying their kinetic energy, using higher electric fields into the instrument (e.g.: High E, RFs, ...);
- ii) On the gas, controlling the kinetic energy of the collisions through its temperature, the gas mass and the pression^[63].

The magnitude of the converted energy can bring different effects. These effects are listed here for conceptual clarity and not to convey a consecutively order:

Desolvation/Declustering

The redistributed energy can break the weak interaction that hold neutral molecules (i.e. water) to the ion. If collisions are more energetic, also the buffer adducts can be removed after proton exchange. Equation (17) schematizes the removal/declustering of an ammonium ion from a biomolecule ion B in gas phase.



This pathway to remove adducts is not accessible for alkali and alkaline earths cations, which are thus detrimental for biomolecules MS signal and should be exchanged for volatiles cations during sample preparation. Often the declustering is followed by the *restructuring* of the DNA ion^[64], and in compaction of some globular protein complexes^[65].

Conformational changes

When the collisional activation overcomes the attractive intramolecular interactions, some ions can unfold due the Coulomb repulsion of like charges. This unfolding process is called Collision Induced Unfolding (CIU) and can be tracked by setting a so called CIU-experiment (ion mobility)^[59, 66] (pp.22,47).

Dissociation of the complexes

Considering an ion of a complex, further activation can lead to its dissociation. This occurs when the increase of internal energy overcomes the interaction between the binding partners, leading for instance to the loss of a neutral ligand. In the case of a charged ligand, the charge state of the complex ion and the ability of the ligand to undergo proton transfer, have an influence on the dissociation. Because of the strong Coulomb interaction between ions in gas phase, if the binding partners have opposite charges (e.g. RNAⁿ⁻ and cationic ligand), the complex is likely to be more stable in the gas phase compared to condensed phase and activation could lead to fragmentation (of covalent bonds) in one of the two partners (i.e. RNA) rather than complex dissociation. If partners have like charges, the Coulomb repulsion will contribute to complex dissociation. Thus, the higher the charge state the easier the dissociation.

1.5.3 How solution concentrations translate into MS intensities

How concentration in solution correlates with ESI-MS peak intensity is until now not completely understood^[67]. This relation is of strong interest for the basics of ESI technique and for quantitative MS methodologies, and we will explore it further in this thesis.

In Figure 13, decyltrimethylammonium ion (DTMA⁺) has a detected intensity five times higher compared to Cs⁺, even though they are injected at the same concentration^[67a].

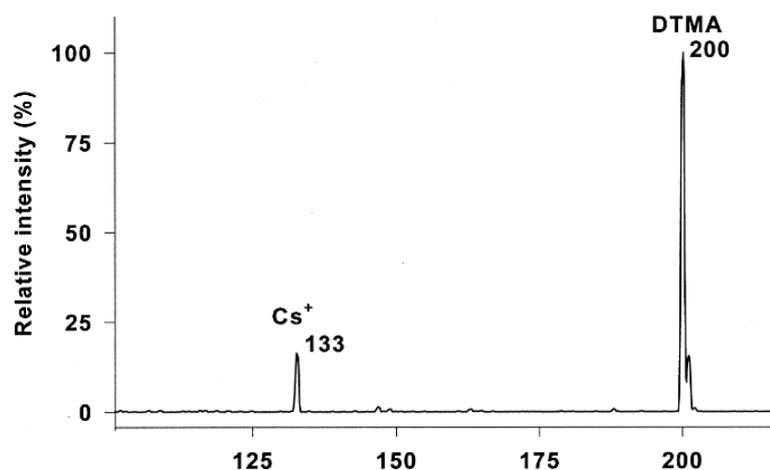


Figure 13 - Equimolar mixture (10 μ M) of Decyltrimethylammonium (DTMA) bromide and CsBr show different ESI-MS response when sprayed from a 50/50 water-methanol solution. Adapted from ref^[67a].

This difference is a consequence of the respective surface activities: the alkyl quaternary ammonium is likely to be on the surface of the droplet and it will be more competitive for surface during droplet fissioning in ESI. On contrast, the droplet surface will be less enriched in Cs⁺, more solvophilic. Thus, Cs⁺ will produce a lower signal after ESI process compared to DTMA⁺.

The common assumption that intensities directly correlate with concentration in solution (same response factors) can be adequate when analytes have similar chemical nature. In such case, surface activity is likely to be similar because of similar solute-solvent interactions (e.g.: net charge, dipole-moment, lipophilicity) . When some changes of surface activities are suspected, it is desirable to account for ionization response through response factor correction (p.39).

ESI response factor (r) is the proportionality constant that relates MS intensity (I_M) of an analyte M to its concentration in solution:

$$I_M = r_M[M] \quad (18)$$

The response factor is a term that depends on all factors that influence the intensity of M , observed on the ESI-MS spectrum. The intensity difference of Cs⁺ and DTMA⁺ in Figure 13 is an example of different partitioning of analytes in the droplet. Other factors are, for instance, ionization efficiency (e.g. ionic vs polar analyte, adduct formation), desolvation/declustering efficiency, the efficiency of transmission of M ions across the mass spectrometer and the detection efficiency.

Aptamers being DNA or RNA that bind their partner ligand (guest) through a shape-adaptive binding mode ^[40a, 68], we were interest om studying their ESI response: if a ligand induces a significant conformational change in its host and this new conformation results in a different exposure to the solvent, different ESI response factors (r_M) may occur for the bound and free-aptamer. Recently this effect has been studied on globular proteins, where the charge state distribution of protein and its complex can be influenced by the change of solvent-accessible surface areas upon ligand binding.^[69] However, for nucleic acids one should account for the phosphodiester backbone: the variations in exposure of hydrophobic and polar groups, and their influence on response factors, might be modest if the negatively charged backbone effects predominate.

Lastly, intensity ratios (M vs ML) are also affinity dependent. This means that an aptamer that has, hypothetically, the same K_D for two different ligands but different response factors (for ML_1 and ML_2) would show up with different complex intensities (hence different “conditional” K_D).

1.5.4 Ion mobility mass spectrometry

Ion mobility mass spectrometry (IM-MS) is a powerful analytical technique that is increasingly being used in the field of structural biology. Ion mobility separates ions according to shape and allows one to compare data measured by other biophysical methods, such as CD, NMR, x-ray crystallography^[70], microscopy^[71] and *in silico* calculated structures^[72].

An important feature of IM-MS is the capability to discern molecules with the same mass but different shape and arrival times. Common examples are topoisomers (e.g. leucine and iso-leucine) and conformational polymorphic molecules^[73].

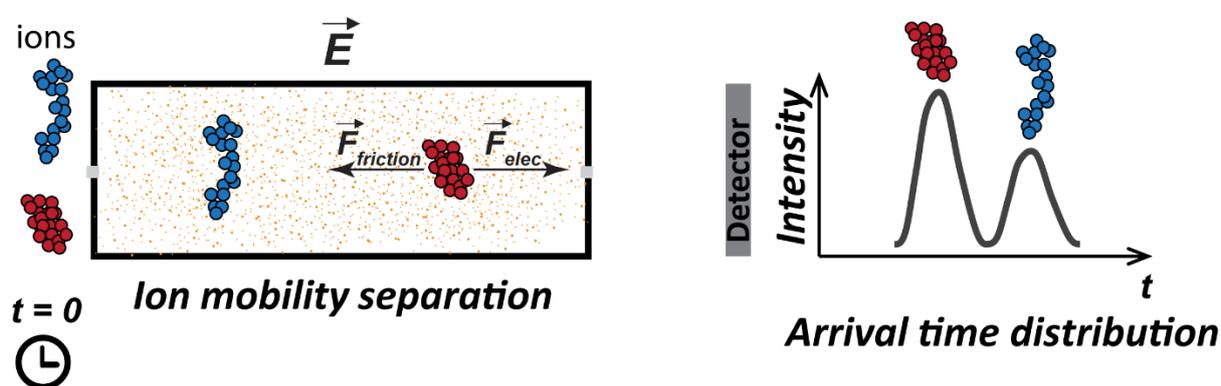


Figure 14 - Ions enter the ion mobility cell at the same time. Here the balance between acceleration from the electric field and the frictions with the buffer gas (i.e. He), move each ion with a different steady state velocity. This velocity difference results in different arrival times to the detector, hence in a separation according to the shape of each ion.

In an ion mobility experiment, in our case into a linear drift tube, ions are accelerated by a constant electric field and drift through a buffer gas. The ions reach a steady state velocity as a result of balance between the gas friction and the force applied by the electric field. The ratio between the resulting steady state drift velocity (v_d) and the applied electric field (E) defines the ion mobility (K) of the ion as follows:

$$K = \frac{v_d}{E} = \frac{L}{t_d E}, \text{ where } E = \frac{\Delta V}{L} \quad (19)$$

Where ΔV is the potential difference (drift voltage) between the two ends of the drift tube of length L .

K is related to the temperature (T , i.e. 296 K) and gas pressure (p , i.e. 3.89 mBar) of the experiment. The reduced ion mobility (K_0), scaled to standard temperature at pressures (273.15 K, 760 mBar), allows comparisons across different experiments. K_0 is defined as:

$$K_0 = K \frac{p T_0}{T p_0} \quad (20)$$

The knowledge of K or K_0 allows one to determine Ω (momentum transfer collision integral)^[74] as:

$$\Omega = \frac{3}{16} \sqrt{\frac{2\pi}{\mu k_B T} \frac{z e 1}{N K}} \quad , \text{where} \quad \mu = \frac{m_{ion} m_{gas}}{m_{ion} + m_{gas}} \quad (21)$$

Ω is thus related to t_d , the drift time (spent into the drift tube), by the following equation:

$$\Omega = \frac{t_d}{L} \frac{3}{16} \sqrt{\frac{2\pi}{\mu k_B T} \frac{z e E}{N}} \quad (22)$$

t_d is the difference between the arrival time (t_a), measured by the TOF detector, and the time spent outside the drift tube (t_0). So we can reformulate the eq (22) as:

$$\Omega = \frac{(t_a - t_0)}{L} \frac{3}{16} \sqrt{\frac{2\pi}{\mu k_B T} \frac{z e E}{N}} \quad , \text{where} \quad t_d = t_a - t_0 \quad (23)$$

Ω is the orientationally averaged momentum transfer collision integral between the buffer gas molecules and the ions, often called ‘‘Collision Cross Section’’ (CCS). Ω is a property of the gas-ion couple and has a dimension of an area (conveniently, \AA^2).

CCS is a derived property of the ion-gas pair, derived from K (ion mobility). However Ω itself is little informative about ion structure. Atomistic level of information on ion structure can be provided *in silico* by molecular modelling methods: candidate structures in the gas phase are generated and their CCS is then calculated with the appropriate model accounting for the gas nature, are compared with experimental CCS^[72b]. Helium at room temperature can be considered as a hard sphere colliding with our ion, and thus a larger CCS value indicates either more extended structures (larger rotationally averaged projected area), or structures with a rougher surface or deeper cavities (parachute effect). In the present work we did not tackle the interpretation by comparison with simulated structural models, so we will discuss only differences in CCS values (for example between a free aptamer and its complex, or depending on the instrumental conditions) and relative peak widths. When measured in helium, a smaller CCS will mean a more compact structure.

1.5.5 Collision Induced Unfolding

Compact ions (short arrival time/low CCS) are normally generated from structures that are folded in solution. That means that two hypothetical ions with different structures, but equally compact, are not distinguishable by simply native IM-MS. However, ions can be differentiated by their response to collisional activation. At high activation, the ion mobility spectrum of a DNA aptamer (M^7^-) changes completely, from one distribution centred to another more extended one (Figure 15). One can monitor how extension occurs as a function of a voltage serving to change the collisional activation prior to IM measurement, by setting a so-called collision-induced unfolding (CIU) experiment (p. 47). Once the voltage that induces

the ATD change is identified with a preliminary experiment like in Figure 15, one can perform an acquisition where the voltage is varied step-by-step, while all the other parameters kept constant.

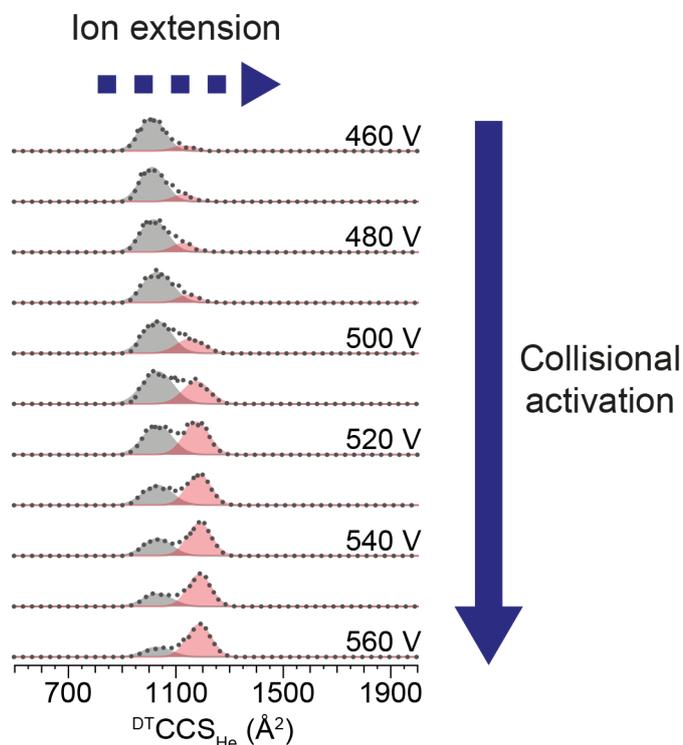


Figure 15 – Ion mobility of M^{7-} , a DNA aptamer, at low collisional activation (fragmentor 460 V) shows a CCS distribution (grey) that on the ion mobility spectrum. At higher activation (fragmentor 560 V) the extension to a second (red) population is visible.

The observation of the CIU plot is useful to understand how the ions respond to collisional activation. For example, to reveal if the extension pass through intermediate shapes (according to their CCS values), or to have an insight on the energetics of the transition (according to the voltage at which a shape change occurs). An unfolding that occurs at lower voltage suggests that weaker intramolecular interactions are involved in the compact structure.

2. Objective of the thesis

Aptamers are single-stranded nucleic acids that often change structure or shape upon ligand binding. This feature makes them promising probes, sensors and stimuli-responsive elements, i.e. in riboswitches for synthetic biology [75]. The major challenge is to understand how aptamers change conformation upon binding, alone or when included into larger molecules. Nowadays, there are no standardized procedures or preferred techniques for aptamer characterization. In this context, additional capabilities to screen a range of aptamer candidates *in vitro*, without labelling, is of high interest. Aptamers are routinely studied using fluorescence anisotropy [76], surface plasmon resonance [77], ITC [78]. Magnetic resonance techniques are used for structural and dynamic elucidations [79]. Here we explore the applicability of native IM-MS to DNA and RNA aptamers.

The first aim is to observe if binding properties of RNA aptamers are preserved in NH₄OAc or TMAA solutions, and whether the complexes survive upon transfer in gas phase in native ESI. To evaluate quantitative aspects, we will compare ITC and MS titrations in the exact same conditions. A challenge is the choice of intergation range and background estimation for the extraction of MS intensities, which could bias the relative response factor correction (if necessary) and data fitting for binding affinity determination. Another challenge is to observe and quantify aptamer binding in presence of Mg²⁺ cations, which are essential co-factor for some aptamers. We benefit also of MS stoichiometry determination to support ITC data fitting, where data interpretation may result biased by a given stoichiometry assumption.

The second aim is to determine whether the aptamers' "adaptive binding" mechanism can be observed based on the ion mobility of free and bound aptamers. In light of the "adaptive binding" we would expect significantly different conformational populations for bound and unbound aptamers, provided that the gas phase ions keep a memory of their in-solution structures. We want also to study if Mg²⁺ will give distinct ATDs, as reported for some RNA aptamers.

2.1 Aptamers analysed in this thesis

Five RNA and four DNA aptamers have been studied (Figure 16). Quantitative native MS, with relative response factors correction, has been used to measure the ligand affinity (expressed in K_D) of malachite green aptamer (MGA)^[32, 80], riboflavin mononucleotide aptamer (1FMN)^[35a, 81], neomycin aptamer (1NEM)^[82] and tobramycin aptamer (1TOB)^[83]. These aptamers covers a range of K_Ds from 10 to 10⁻³ μM (respectively 1; 0.5; 0.1; 0.009 μM) via intercalating binding mode (MGA, 1FMN) and electrostatic binding mode (1NEM, 1TOB). MGA has been studied extensively with ITC in conventional buffers, making it a convenient system to compare native MS and ITC in the same volatile salts. 1FMN is Mg²⁺ dependent and its reported K_D is in the same range of MGA's K_D^[80]. 1NEM is not sensitive to Mg²⁺ and has a K_D

(0.1 μM) similar to MGA and 1FMN^[81]. The K_D s of 1NEM and 1TOB (0.009 μM) make them a good system to test the limits of native MS quantification.

Table 1 – Ligands, K_D s and conditions available in literature for the aptamers considered in this work.

Type	Aptamer	nt	Ligand	K_D	Conditions
RNA	1NEM ^[82]	23	Neomycin	0.100 μM ^[84]	5 mM Mg^{2+} ; 50 mM Tris•HCl (pH 7.6); 250 mM NaCl ^[84]
	1TOB ^[83a]	27	Tobramycin	9 nM (1 Ligand); 2.7 μM (2 ligands) ^[83b]	20 mM TrisAc (pH 7.4); 1 mM MgCl_2 ; 1 mM CaCl_2 ; 5 mM KCl; 140 mM NaCl ^[83b]
	1FMN ^[35a]	36	RiboFMN	0.500 μM ^[81]	5 mM Mg^{2+} ; 50 mM Tris•HCl pH 7.6; 250 mM NaCl ^[81]
	MGA ^[85]	38	Malachite green	0.28 μM ^[80a]	25 °C, 10 mM sodium phosphate (pH 6.7), 150 mM NaCl ^[80a]
	TCA ^[28]	57	Tetracycline	0.770 nM ^[86]	25°C, 20 mM Tris•HCl (pH 7.5), 10 mM MgCl_2 , 100 mM NaCl ^[86]
DNA	MN4 ^[87]	36	Quinine	0.15 μM ^[80b]	10°C, 20 mM Tris (pH 7.4, 140 mM NaCl, 5 mM KCl) ^[80b]
	MN19 ^[87]	30		0.23 μM ^[80b]	
	OR8 ^[88]	28		11.2 μM ^[80b]	
	OR7 ^[88]	26		n.d. ^[80b]	

On the IM-MS side, we want to compare free and bound forms and see if they are differentiable by IM-MS. In addition to the previous list, we studied tetracycline (TC) aptamer and four DNA cocaine binding aptamers (CBA), for which a conformational change upon binding is reported in the literature. Supported by the x-ray coordinates for TC-Apt and NMR topologies for CBA, we want to see if these conformational differences are observable by IM-MS and in which extent.

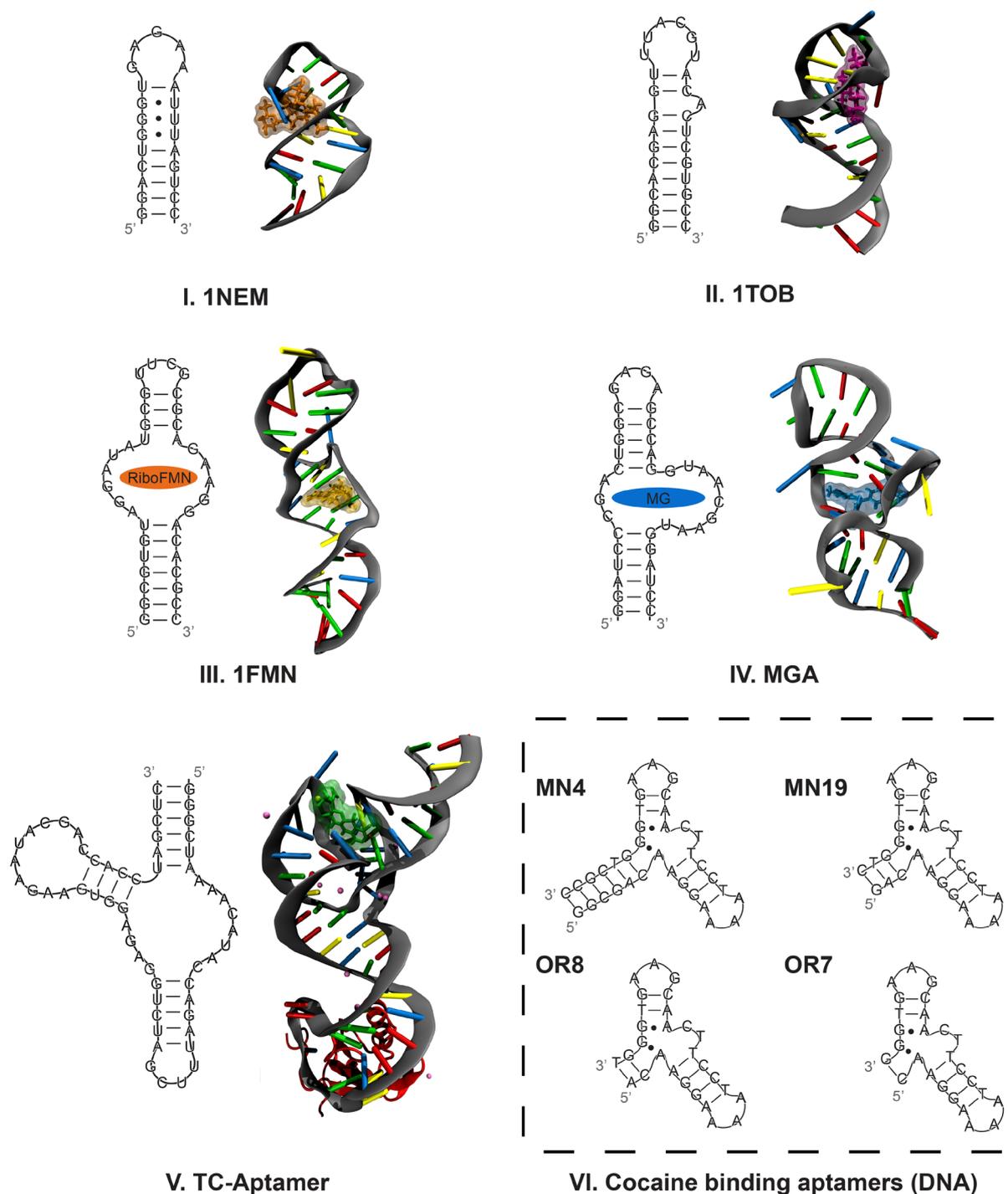


Figure 16 – I-IV. RNA aptamers studied by native MS to quantify their K_D in MS-friendly volatile salts; V. RNA tetracycline aptamer (x-ray crystal with U1A protein in red) and -VI. DNA cocaine binding aptamers, studied by IM-MS, for which large conformational change upon binding is expected (see text). Only NMR topologies are available for VI. PDB coordinates are available for aptamers 1NEM.pdb (I), 1TOB.pdb (II), 1FMN.pdb (III), 1Q8N.pdb (IV) and 3EGZ.pdb (V). Nucleobases color code: A-Blue; G-Green; U-Yellow; C-Red; Mg^{2+} rose.

3. Materials and Methods

3.1 UV melting

DNA and RNA have a strong absorption band at ~260 nm, due to the aromatic character of the nucleobases^[89]. This facilitates their quantification measuring their optical absorbance using the Beer-Lambert law:

$$A = \log_{10} \frac{I_0}{I} = \epsilon l c \quad (24)$$

I_0 and I are the incident and emerging light (to the sample), ϵ is the molar extinction coefficient [$\text{cm}^{-1} \text{M}^{-1}$], l is the optical path [cm] and c is the molarity of the sample. One can follow how the absorbance a given wavelength changes with the temperature, in order to monitor the structural transitions of DNA or RNA. This is called “thermal denaturation” and is a common method used to evaluate the stability of nucleic acids secondary structures (e.g.: self-complementary DNA duplex, quadruplex), or nucleic acids complexes. For example, the conversion of duplex regions to single strands is accompanied by a hyperchromism at 260 nm.

Assuming the case of a single stranded (*Ss*) DNA that self-associate into a hairpin (*Hp*) with a self-complementary duplex region, the curve will be sigmoidal shaped (Figure 17A) typical for a two-state transition (hairpin folding).



The melting temperature (T_m) is defined as the temperature at which half of the structure is unfolded. To determine folded fraction, first two baselines, at low ($L1_{(T)}$) and high ($L0_{(T)}$) temperature, should be found (red lines in Figure 17A). Then the ordinate $A_{(T)}$ can be converted into the fraction folded $\theta_{(T)}$:

$$\theta_{(T)} = \frac{L1_{(T)} - A_{(T)}}{L0_{(T)} - L1_{(T)}} \quad (26)$$

The T_m can determined from the intersection of data points and median of the two baselines ($\theta_{(T)} = 0.5$). However we will use the maximum of the first derivative of data points (dA_{260}/dT) that is unaffected by the choice of baselines. We will use the the sigmoidal fitting of $\theta_{(T)}$ vs T curve (Figure 17B) allows to to estimate the fraction folded at room temperature. The folded fraction is related to the association constant of the two partners, eq. (27). Then one can determine the ΔG° of the transition^[90] (p.12).

$$K_A = \theta_{(T)} / (1 - \theta_{(T)}) \quad (27)$$

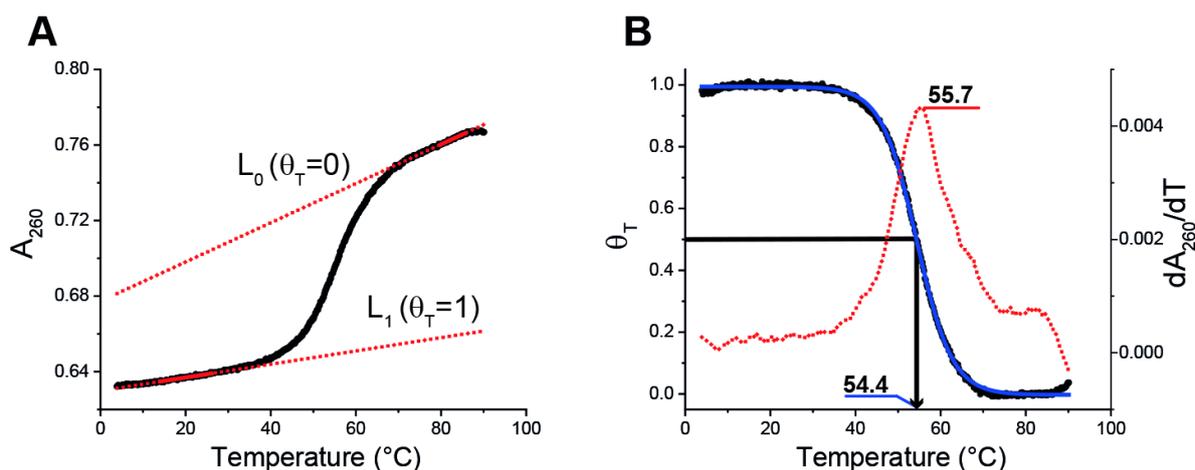


Figure 17 – **A**) Absorbance recorded at 260 nm is plotted vs temperature of the cell. The red straight lines are the baselines for the A_{260} transition associated to the melting. **B**) A_{260} is converted as folded fraction $\theta_{(T)}$ using eq. (26) . The melting temperature T_m is reported as sigmoid centroid (**blue solid curve**) and as maximum of A_{260} derivative (**red dotted curve**).

Nucleic acid aptamers have a secondary structure more complex rather than a simple duplex (e.g. hairpin, bulges, junctions at pp. 5-7). In this case, several successive or overlapping transitions may complicate the T_m determination, T_m a qualitative parameter about the global stability of a sequence. In those cases, plotting the first derivative of the melting curve can help to find the inflexion points to provide a rough estimate of the T_m (e.g. when is difficult to define one or both baselines).

In this work, we used 400 μ L QS quartz cuvettes (model 115B-10-40 - Hellma, Müllheim, Germany) of 10 mm path length, and a SAFAS UVmc1 (SAFAS, Monaco) UV-spectrophotometer. The temperature ramp was set to 0.2°C/min from 4°C to 90°C and back. Samples were prepared at 2 μ M concentration to have an A_{260} between 0.5 and 0.8 a.u. ($2 \cdot 10^5 < \epsilon_{260} < 4 \cdot 10^5$) at 25°C, to stay far from detector saturation.

3.2 Circular Dichroism

Circular dichroism is a visible technique that probes the asymmetry of a system. Nucleobases are achiral but their environment is chiral when included in DNA or RNA, because of the chiral sugar-phosphate backbone and the resulting helical structures. In addition, the spatial orientation (e.g. secondary structure) of the base chromophores contributes to chirality. B-DNA signature has a broad positive band at 275 nm and a negative band at 245 nm^[91].

A Jasco J-1500 spectrophotometer (JASCO Co., Ltd., Hachioji, Japan) was used to perform the CD experiments over 220-350 nm at 100 nm/min. A quartz cell of 2 mm path length (l) and 10 mm width was used. The temperature was set to 20 °C. Each spectrum was acquired in triplicate. The buffer was used as blank, recorded in the same cuvette with the same method, and subtracted by the instrument from each spectrum.

CD data were converted from ellipticity (θ - mdeg) to molar circular-dichroic absorption ($\Delta\varepsilon$ – cm² mmol⁻¹) based on DNA concentration (C) using the following equation^[92]:

$$\Delta\varepsilon = \theta / (32980 \cdot C \cdot l) \quad (28)$$

3.3 ITC instrumental set up and data treatment

Isothermal titrations were conducted at 25 °C using an ITC200 microcalorimeter from Microcal (Northampton, USA). All sequences were exchanged extensively on spin columns against NH₄OAc or TMAA 100 mM buffer (pH 6.8 were verified with an InoLab – Multi 9420 IDS (WTW, Xylem Analytics, Weilheim, Germany) and MgCl₂ 200 μM when needed, by dilution of Magnesium chloride BioUltra 1M from Sigma Aldrich (Saint-Quentin Fallavier, France). RNA and ligand concentrations were determined by UV-Vis spectroscopy on solutions before loading into ITC200. The concentrations used in the ITC cell are in the range of 10 to 15 μM for the RNA aptamer. The concentration of ligand inside the syringe was between 140 and 160 μM, prepared by dilution of concentrated ligand. Samples and buffers were degassed before use in a sonic bath for 5 minutes at 20°C.

Two injection methods (Table 2) have been used in order to explore a thermogram sampling between 19 and 25 points. Even though the methods are similar, the first (I) gave optimal (less noisy) heat pulse intensities for integration.

All ITC experiments were replicated from three to five times. Data sets are analysed using the embedded Origin 7.0 provided by the manufacturer. Manual correction of baseline on thermograms has been necessary only to exclude spikes from false peak identification and to correct baseline definition in spike proximity. Ligand dilution control runs were always subtracted to account for dilution heats. Origin 7.0 one binding site model, eq. (14), was used for point fitting after thermograms integration, discarding the first injection of 0.5 μL. Number of sites (N) was kept as an adjustable parameter, to can accommodate for errors on the

ligand/target concentration ratios. The standard deviation of fitted parameters across the replica is reported.

Table 2 – Instrumental settings for isothermal calorimetry. *a.* for exact values, refer to text and annexes

Injection Method - I			Injection Method - II		
Cell temperature		25°C	Cell temperature		25°C
Reference power		10 µcal/s	Reference power		6 µcal/s
Initial delay		300 s	Initial delay		300 s
Syringe concentration ^a		140-155 µM	Syringe concentration ^a		140-155 µM
Cell concentration ^a		10-15 µM	Cell concentration ^a		10-15 µM
Stirring speed		750 rpm	Stirring speed		750 rpm
Feedback Mode/Gain		High	Feedback Mode/Gain		Low
Total Injections		19	Total Injections		25
First injection	V	0.5 µL	First injection	V	0.5 µL
	<i>Duration</i>	1 s		<i>Duration</i>	1 s
	<i>Spacing</i>	180 s		<i>Spacing</i>	120 s
	<i>Filter</i>	5 s		<i>Filter</i>	5 s
Injections 2 to 19	V	2.0 µL	Injections 2 to 25	V	1.5 µL
	<i>Duration</i>	4 s		<i>Duration</i>	1 s
	<i>Spacing</i>	180 s		<i>Spacing</i>	120 s
	<i>Filter</i>	5 s		<i>Filter</i>	5 s
	<i>Run time</i>	~65 min		<i>Run time</i>	~55 min

3.4 Mass spectrometers

3.4.1 Agilent 6560 ESI-IM-Q-ToF

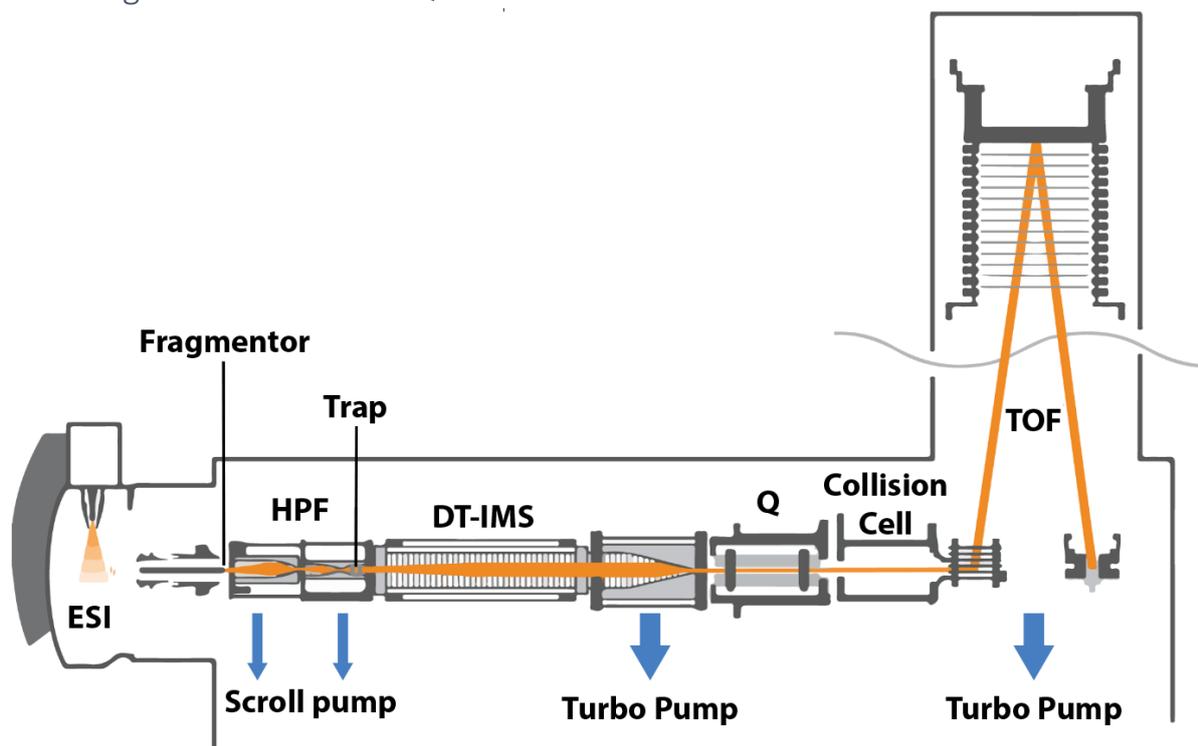


Figure 18 - Agilent 6560 is composed by an ESI source (Agilent Jet stream in the schema), the heated sampling capillary (Fragmentor parameter highlighted) before the pre-IMS funnels, the trapping (grids) region before the IMS drift tube; After the drift tube we have a focusing funnel, a quadrupole (Q), the collision cell filled with N₂ and the time of flight (TOF) region. Adapted from ref ^[93].

The Agilent 6560 is a quadrupole time-of-flight mass spectrometer coupled to anion mobility drift tube on its front end. This instrument is capable to separate ions by their mobilities (p.21) thanks to the drift tube, and then to separate them according to their m/z ratio by the time-of-flight (TOF) with resolution up to 42 000^[94]. The Q-TOF allows one to do MS/MS. In our case high purity helium is flowed into the ion mobility part using the “alternate gas kit” and with in-house modified pumping^[95], to perform IMS experiments in 100% helium.

The ion mobility part is divided in three regions: front funnels with trapping gate/region, the linear drift tube and the rear ion funnel.

Pre-IMS parameters

Here are summarized the parameters critical for our study. Values are summarized in Table 3 and Table 4 at p.34.

- Fragmentor

The “fragmentor” parameter is the voltage difference between the heated capillary (entrance of the instrument) and the first electric lens of the front funnels (High pressure funnel - HPF)

reported in Figure 18 and Figure 19. Because the gas pressure in this region is about 4-3 mbar (of a mixture of N₂ and He), this parameter acts directly on ion activation: how ions are “pulled” towards the front funnels will induce more or less energetic collisions with the gas. Depending on the set value, it is possible to control the extent of desolvation, declustering or even dissociation. The voltage range is 200-600 V.

- Front funnels

Ion funnels reduce ion losses when transferring ions. The focusing provided by ion funnels is particularly suitable to limit ion radial diffusion for high sensitivity and resolution in IM-MS experiments [96]. However, in typical IM-MS conditions (i.e. p_{He} 3.76 Torr) RF fields in ion funnels leads to collisional activation [97]. Voltages in this region may thus also influence the activation of ions [66a].

- Trapping region

The trapping region before the drift tube IM is necessary to form a discrete packet of ions to send into the drift tube, and define the start time (time zero) for the mobility drift. Three grids permit ion trapping in this region. Trapping time, release time and the voltages profile can also influence ion activation. The trap fill time range can vary from 100 μs (min. release time) to 60 ms (drift time duty cycle/max drift time). Using longer trap fill times does not influence ion activation but tends to lower the total ion count of the spectra. The reason is still not clear. [66a] The parameter Trap Entrance Grid Delta (TEGD) is the voltage difference between the trap entrance grid in trapping and release mode (Figure 19). A higher TEGD value gives a steeper voltage profile of trap region during release mode, and induces an higher activation [66a, 98]. Fragmentor and TEGD are the main parameters to control ion activation on our instrument and perform CIU experiments [59, 66a, 98].

Ion Mobility Drift Tube

Ion mobility separation occurs in this region. The drift tube is filled with Helium, at 3.89 Torr. One should prevent N₂ contamination into the tube (i.e.: from the front funnels), otherwise mobility measurements will be the result of collisions with a N₂-He mixture of undefined. Partial pressures. To do so, a pressure differential of minimum 0.18 Torr is maintained between the drift tube and the front funnels [72b].

Helium pressures and the differentials were adjusted on daily basis by verifying that the correct CCS of [d(TG₄T)₄•(NH₄)₃]⁵⁻ was within a tolerance of 1 Å² (DT_{He} = 788 Å²) [72b] and dT₆²⁻ within a tolerance of 1 Å² (DT_{He} = 306 Å²) [99], using a mixture of d(TG₄T)₄ 10 μM and dT₆ 5 μM in NH₄OAc 100 mM.

Post-IMS parameters

The quadrupole can be used either in RF-only mode as simple ion guides or as m/z filter to allow a specific m/z value to pass towards the TOF. The collision cell allows to do MS² experiments by accelerating ions (or selected ions) into the cell containing 2×10^{-3} mbar N₂. We also minimized collisional activation in all these regions (Table 3). This is necessary to avoid post-IMS dissociation of our complexes, and measure the same m/z ratio that were present into the DT-IMS.

The TOF analyser ($\sim 7 \times 10^{-8}$ bar) measures the time elapsed between the ion arrival at the detector and their start time from the pusher. One reflectron (electrical lens mirrors) is present to extend the ion flight path, time-focus the ions and thus achieve higher mass resolution. The detector is a microchannel plate that amplifies the hit of an ion with an electron cascade. The electric current is then converted to give the spectrum of flight times. Flight times are then converted to mass over charge ratio (m/z) by a calibration with Agilent Tunemix.

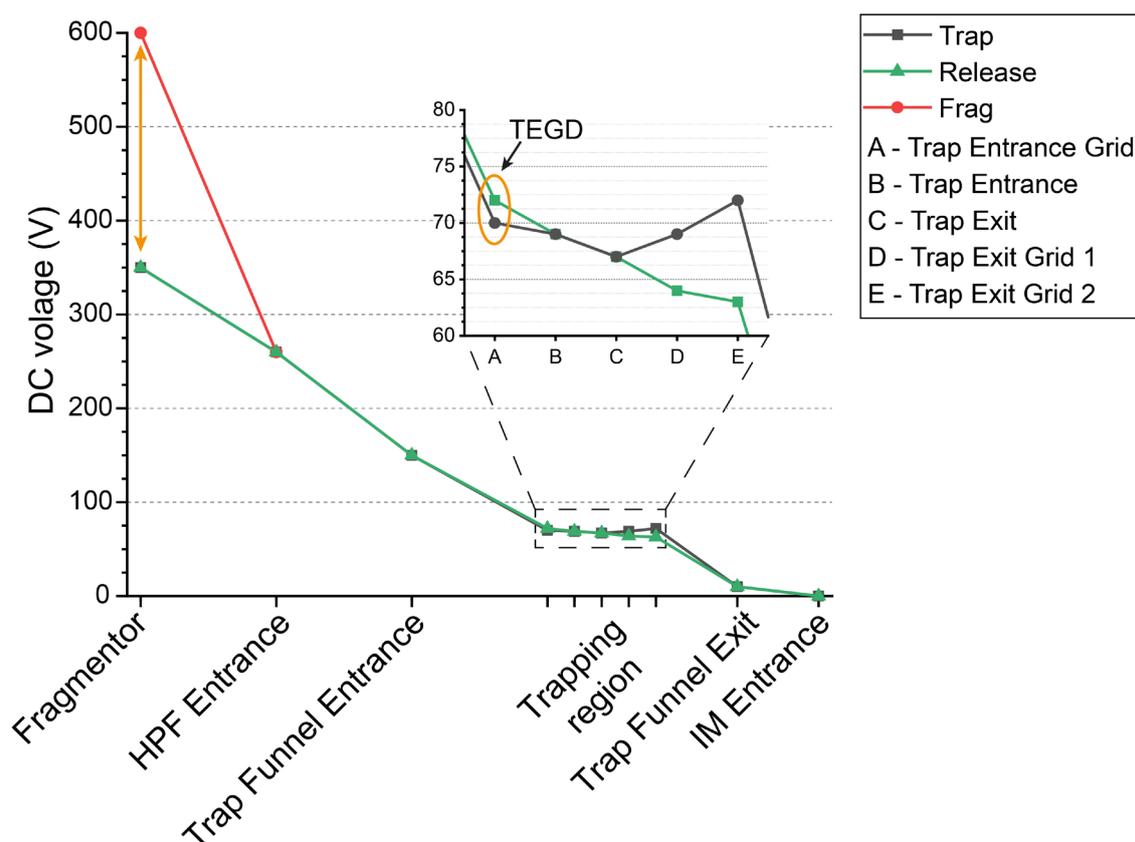


Figure 19 - Pre-IMS DC voltages of regular tuning and CIU tuning. The two activation parameter (fragmentor and Trap Entrance Grid Delta - TEGD) are marked in orange. Voltage of CIU are showed (Table 4). Voltages are in relative value to the IM entrance.

Agilent 6560 tuning

Table 3 - The following tuning parameters were used for all RNA and DNA samples acquired on Agilent 6560.

Pre-IMS Parameter	Value	post-IMS Parameter	Value
Source: gas temperature	200 °C	IM drift tube: Drift tube exit	-210 V
Source: drying gas	5 L/min	IM rear funnel: Rear funnel entrance	-200 V
Source: nebulizer pressure	13 psig	IM rear funnel: Rear funnel RF	-130 V _{p-p}
Source: capillary	-3500 V	IM rear funnel: Rear Funnel Exit	-35 V
Optics 1: Fragmentor	-350 V	IM rear funnel: IM Hex Entrance	-32 V
IM front funnel: high pressure funnel delta	-110 V	IM rear funnel: IM Hex Delta	-3 V
IM front funnel: high pressure funnel RF	-180 V _{p-p}	Optics 1: Oct Entrance Lens	-27 V
IM front funnel: trap funnel delta	-160 V	Optics 1: Oct 1 DC	-25 V
IM front funnel: trap funnel RF	-180 V _{p-p}	Optics 1: Lens 1	-23 V
IM front funnel: trap funnel exit	-10 V	Optics 1: Lens 2	Off
IM trap: trap entrance grid low	-82 V	Quad: Quad DC	-21 V
IM trap: trap entrance grid delta	-2 V	Quad: PostFilter DC	-21 V
IM trap: trap entrance	-79 V	Cell: gas flow	22 psi
IM trap: trap exit	-76 V	Cell: Cell Entrance	-16 V
IM trap: trap exit grid 1 low	-72 V	Cell: Hex DC	-16 V
IM trap: trap exit grid 1 delta	-6 V	Cell: Hex Delta	-3 V
IM trap: trap exit grid 2 low	-71 V	Cell: Hex2 DC	-14.6 V
IM trap: trap exit grid 2 delta	-13 V	Cell: Hex2 DV	-1.5 V
Acquisition: Trap fill time	1000 μs	Cell: Hex3 DC	-12.9 V
Acquisition: Trap release time	100 μs	Extractor: Ion Focus	-10 V

Table 4 – Pre-IMS Parameters for direct CCS_{He} and collisional induced unfolding on cocaine aptamers (p. 118). The key parameters to obtain the CIU are in **bold**. Fragmentor was varied up to 600V by steps of 10V; This method was obtained from a previous test using the optimized method for in ubiquitin 7+ in He from ref^[66a]; Comparison at p.213 - Figure 129.

Pre-IMS parameter	Value
Source: gas temperature	200 °C
Source: drying gas	5 L/min
Source: nebulizer pressure	13 psig
Source: capillary	-3500 V
Optics 1: Fragmentor	-350-600 V
IM front funnel: high pressure funnel delta	-110 V
IM front funnel: high pressure funnel RF	-160 V_{p-p}
IM front funnel: trap funnel delta	-140 V
IM front funnel: trap funnel RF	-160 V_{p-p}
IM front funnel: trap funnel exit	-10 V
IM trap: trap entrance grid low	-70 V
IM trap : trap entrance grid delta	-2 V
IM trap: trap entrance	-69 V
IM trap: trap exit	-67 V
IM trap: trap exit grid 1 low	-64 V
IM trap: trap exit grid 1 delta	-5 V
IM trap: trap exit grid 2 low	-63 V
IM trap: trap exit grid 2 delta	-9 V
Acquisition: Trap fill time	1000 μs
Acquisition: Trap release time	100 μs

Table 5 – Pre-IMS Parameters for direct CCS_{He} for Tetracycline RNA aptamer (p. 95). The key parameters to remove TMAA adducts and clustering in general are in **bold**; Fragmentor was fixed to 600V; This method was already reported from ref^[98]. † Source parameters refer to a nESI source and nebulizer pressure is used as back pressure to the emitter.

Pre-IMS parameter	Value
Source: gas temperature	200 °C
Source: drying gas†	3 L/min
Source: nebulizer pressure†	1 psig
Source: capillary†	-1050 V
Optics 1: Fragmentor	-600 V
IM front funnel: high pressure funnel delta	-160 V
IM front funnel: high pressure funnel RF	-180 V _{p-p}
IM front funnel: trap funnel delta	-160 V
IM front funnel: trap funnel RF	-180 V _{p-p}
IM front funnel: trap funnel exit	-10 V
IM trap: trap entrance grid low	-125 V
IM trap : trap entrance grid delta	-20 V
IM trap: trap entrance	-100 V
IM trap: trap exit	-99 V
IM trap: trap exit grid 1 low	-97 V
0.IM trap: trap exit grid 1 delta	-6 V
IM trap: trap exit grid 2 low	-96 V
IM trap: trap exit grid 2 delta	-10 V
Acquisition: Trap fill time	1000 μs
Acquisition: Trap release time	100 μs

3.4.2 Thermo Exactive

Orbitrap mass spectrometers are high-resolution instruments. The ions are trapped in stable orbits around the central spindle-like electrode, where the oscillation frequency is related to its mass-to-charge ratio. This design and its improvements allow resolution up to $>1\,000\,000$ ^[100], (only 50 000 on our instrument), large dynamic and m/z range ^[101].

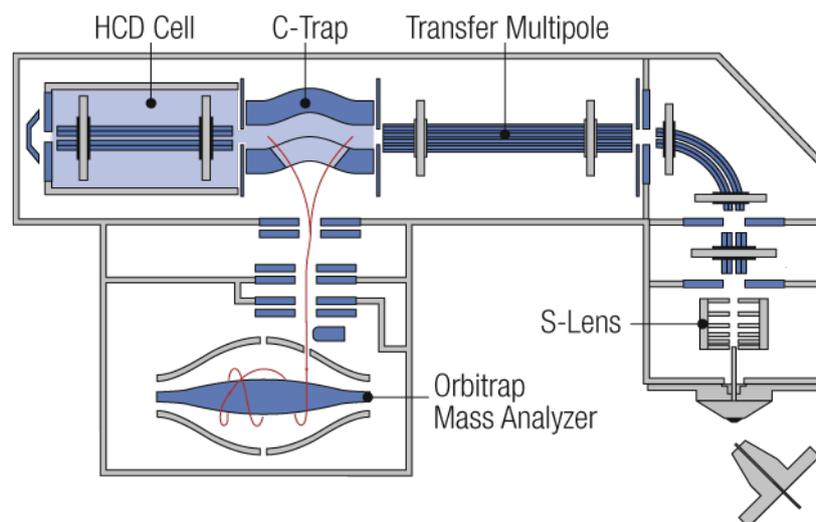


Figure 20 –The Thermo Exactive mass spectrometer and its components.

Samples are injected using a syringe pump connected to the ESI source. Ions are then transferred through four stages of differential pumping using radiofrequency (RF)-only multipoles into a RF-only trapping curved quadrupole filled with nitrogen, termed “C-trap”. Packets of ions are stored and thermalized before injection into the orbitrap mass analyser. In the analyser, ions oscillate between the spindle-like central electrode and the outer electrodes using only electrostatic fields. The high vacuum into the analyser chamber ($< 10^{-9}$ mbar) allows ions to oscillate during to 250 ms, inducing an image current that is detected and Fourier-transformed to oscillation frequencies (of each ion package) and finally converted to a mass spectrum^[101b, 102]. If needed, an octapole termed “High Collision energy Dissociation” (HCD) cell allows fragmentation before injection into the analyser.

The Exactive ESI-Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany) was used to titrate 1FMN in TMAA and NH_4OAc 100 mM with MgCl_2 200 μM . Parameters are summarized in Table 6.

Table 6 - Samples acquisition and native MS titrations were acquired with the following parameters. Fragmentation cell (HCD) was used only with 1FMN aptamer to reveal Mg²⁺ adducts (p.77).

Scan Parameters	Value
Scan Range	800 – 3200 m/z
Fragmentation (HCD)	Off
Only for 1FMN aptamer + Mg²⁺	15 ev
Resolution	High
Polarity	Negative
Microscans	2
Lock masses	Off
AGC target	High dynamic range
Maximum Inject time	250 ms
HESI Source	Value
Sheath gas flow rate	55
Aux gas flow rate	5
Sweep gas flow rate	0
Spray voltage	-3 kV
Capillary temperature	180 °C
Capillary voltage	-15 V
Tube lens voltage	-160 V
Skimmer voltage	-14 V
Heater temperature	25 °C

3.5 Relative response factor determination

In the introduction we defined the ESI response factor (r) as the proportionality constant that relates MS intensity (I_M) of an analyte M to its concentration in solution. In our work we will use the method proposed by Gabelica *et al.*^[103] to determine relative response factors using an internal standard. Here we briefly summarize the principles.

In the simplest case of a non-covalent complex formed by the binding of the host M with the guest L , the equilibrium dissociation constant (K_D) is:

$$M + L \rightleftharpoons ML$$

$$K_D = \frac{[M][L]}{[ML]} \quad (29)$$

and considering the response factors can be reformulated including eq. (18) :

$$K_D = \frac{r_{ML}}{r_M} \frac{I_M}{I_{ML}} [L] = R \frac{I_M}{I_{ML}} [L], \quad \text{where } I_M = r_M [M] \quad (30)$$

The method compares the intensity of each ion to that of an internal standard (Std), chemically similar enough that, if some factor influence the response, both will be affected in the same way. The internal standard in our case is the DNA hexamer dT₆, kept at constant concentration during all the experiments. Knowing the concentration of Std , we define the relative response factor (i.e. R_M, R_{ML}) for each with respect to the internal standard:

$$\frac{[M]}{[Std]} = \frac{r_{Std}}{r_M} \frac{I_M}{I_{Std}} = R_M \frac{I_M}{I_{Std}} \quad \text{and} \quad \frac{[ML]}{[Std]} = \frac{r_{Std}}{r_{ML}} \frac{I_{ML}}{I_{Std}} = R_{ML} \frac{I_{ML}}{I_{Std}} \quad (31)$$

and knowing that $[M]_{tot} = [M] + [ML]$ it is possible to reformulate eq. (31) as:

$$\frac{[M]_{tot}}{[Std]} = R_M \frac{I_M}{I_{Std}} + R_{ML} \frac{I_{ML}}{I_{Std}} = C \quad (32)$$

where C is the ratio between $[M]_{tot}$ and $[Std]$.

A titration experiment provides many points of measurable intensities I_M and I_{ML} , and each one gives a linear equation like (32) as part of a set, which can be written in matrixes as follows:

$$\begin{pmatrix} \frac{I_{M,1}}{I_{Std,1}} & \frac{I_{ML,1}}{I_{Std,1}} \\ \vdots & \vdots \\ \frac{I_{M,m}}{I_{Std,m}} & \frac{I_{ML,m}}{I_{Std,m}} \end{pmatrix} \begin{pmatrix} R_M \\ R_{ML} \end{pmatrix} = \begin{pmatrix} C \\ \vdots \\ C \end{pmatrix}$$

$$I \cdot R = C \quad (33)$$

The intensity matrix I is a $m \times n$ matrix where m are the experimental points (titration solutions, kinetics point, ...) and n are the different forms of the host M (M, ML, ...). If we have only species M and ML, n is equal to 2. The vector R has the same dimension as n (columns of I) and vector C has the same dimension as m (rows of I).

Albeit a square matrix is enough to solve the relation (33), it is desirable to use all the data points and have an overdetermined system (m rows $\geq n$ columns) in which the relative intensities of the ions vary significantly. Thus the relative response factor vector R is calculated as

$$R = I^+ \cdot C \quad (34)$$

Where I^+ is the Moore-Penrose pseudoinverse of the intensity matrix I , commonly used to compute a least squares solution of a system in linear algebra.

The conditions for method validity are that

- i. The relative response factors of each species remain constant along the experiment;
- ii. The complex(es) remains intact (no dissociation or further association) upon ESI-MS analysis (no ML converts to M after part of the process);
- iii. Relative intensities between analyte and complex(es) should vary enough (e.g. up to saturation of M by the ligand) to permit the calculation of response factors;
- iv. Vector C (ratio $[M]_{tot} / [Std]$) should be the most constant as possible. For example, M and Std pooled before preparing the samples at different ligand concentrations (titration points), will maintain their mixing ratio besides handling or errors.

Data treatment

Ion peaks to consider are integrated using an in-house developed R-script, for Agilent 6560, and a Python script for Thermo Exactive. In this work, MS peak intensities were extracted using in-house developed R or Python script. An example of script input is reported in Figure 91 at p.145.

3.5.1 Determination of dissociation constant using MS titrations

One of the basic assumptions in using mass spectrometers for K_D determination, is that during ESI process the reactions are “kinetically trapped”, because droplet Rayleigh fissions (μ s time scale) occurs faster than analytes diffusion rate constants, minimizing equilibrium displacement (dissociation of ML or association of M and L)^[104]. But also, multiple fission and evaporation of droplets helps in keeping the analytes concentration low^[105].

During a titration of a binding equilibrium such as eq. (29), one varies the concentration of one partner of the reaction (i.e. L) and monitors a physical change related to the concentration of at least one of the others reactant or products (i.e. M and or ML). In a MS titration the physical change are the intensities (ion currents) of the reactants and products (in our case, M and its complex ML). Concretely, several samples of the analyte (i.e. nucleic acid) are acquired with increasing ligand concentrations. Each sample is an equilibrium point of the binding reaction between nucleic acid and its ligand, from which it is possible to determine the relative response factors for analyte and complex(es) and reconstruct the binding isotherm(s).

To respect the balance of mass we combine equations (31) and (32) we define $[M]_{corr}$ and $[ML]_{corr}$ as

$$[M]_{corr} = [M]_{tot} \frac{R_M \frac{I_M}{I_{Std}}}{\left(R_M \frac{I_M}{I_{Std}} + R_{ML} \frac{I_{ML}}{I_{Std}}\right)} ; [ML]_{corr} = [M]_{tot} \frac{R_{ML} \frac{I_{ML}}{I_{Std}}}{\left(R_M \frac{I_M}{I_{Std}} + R_{ML} \frac{I_{ML}}{I_{Std}}\right)} \quad (35)$$

Remembering that we defined the relative response factors as $R_x = \frac{r_{Std}}{r_x}$, the equilibrium dissociation constant (30) can be expressed for each experimental point as

$$K_{D,corr} = \frac{[M]_{corr}[L]}{[ML]_{corr}} = \frac{R_M}{R_{ML}} \frac{I_M}{I_{ML}} [L]_{free} \quad (36)$$

And $[L]_{free}$ is obtained from the mass balance equation:

$$[L]_{free} = [L]_{tot} - [ML]_{corr} \quad (37)$$

Using these questions a K_D value can be determined for each titration point. To use the entire titration curve, we plot the binding isotherm, which consists of the values of $[ML]_{corr}$ as a function of $[L]_{tot}$. Dissociation constants, K_D , in our work have been determined by non-linear fitting of the binding isotherms using the software DynaFit (4.07.082, BioKin Ltd Watertown, MA)^[106]. We preferred this method because single “poor” data points can lead to biased equilibrium constant, whereas examining the entire titration curve allows to evaluate quality of points by examining how well the curves fit the data (Figure 21). Dynafit is also able to fit the titration data of more complex reaction stoichiometries if required.

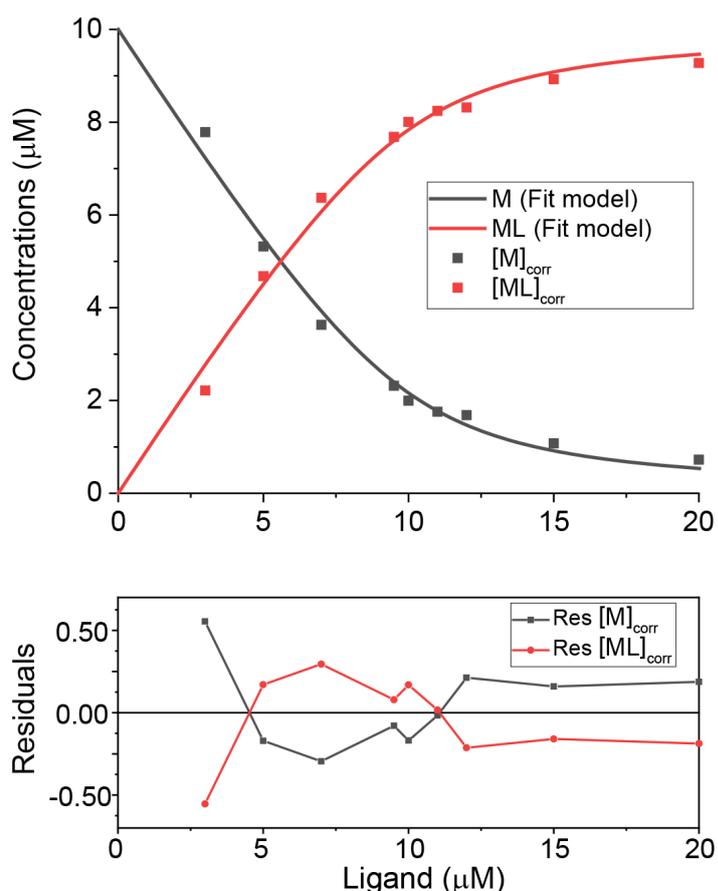


Figure 21 - Example of DynaFit output for a binding isotherm and residuals for a MS titration on a RNA aptamer (M) and its complex (ML). On ordinate, concentrations of M and ML after relative response factor correction; on abscissa, ligand concentration. $K_D = 0.598 \pm 0.046 \mu\text{M}$ ($R^2 = 0.994684$). Fit output is available in Table 23 at p.151.

Based on the reaction mechanism that we give as input, the DynaFit software computes the composition at equilibrium by solving simultaneously the mass balance equations for the component species by using the multidimensional Newton–Raphson method^[107]. Data are then fitted by the iteration method chosen by the user (e.g. by default a Levenberg–Marquardt algorithm for squared residuals minimization). A modified Monte-Carlo method^[108] is implemented to determine the confidence intervals of the parameters (i.e.: K_D), simulating 1000 data set permutations, without assuming a specific statistical distribution of the experimental data. However, the algorithm assumes no error on ligand concentrations (abscissa). DynaFit typical script and output are available at p.150.

In this work, we calculated the relative response factor correction for each charge state of each MS replica. The uncertainty of relative response factors is expressed as the standard deviation over the replicas. Both methods were repeated considering relative response factor equal to 1 (i.e.: fixing both at 1) for comparison purposes.

Dissociation constants (K_D) have been calculated using two methods:

- I) Concentrations (M and ML) of the single charge states per each replica are used as DynaFit input. Then an averaged K_D is calculated over the replicas per each charge state, with uncertainty expressed as standard deviation.
- II) Averaged concentrations of the single charge state over the replicas are used as input for DynaFit. The uncertainty on K_D of single charge state, is the confidence interval provided by Monte-Carlo iteration of DynaFit algorithm.

ΔG^o was calculated from these K_D s and we propagated the uncertainty using the following rule in quadrature on eq. (4):

$$\Delta G^o = RT \ln K_D \quad \delta \Delta G^o = |\Delta G^o| \sqrt{\left(\frac{\delta T}{T}\right)^2 + \left(\frac{\delta \ln K_D}{\ln K_D}\right)^2} \quad (38)$$

The basic assumption behind equation (38) are that the averaged values of variables are normally distributed, uncertainties (δT , $\delta \ln K_D$) are independent and random^[109].

3.5.2 Extraction of MS peak Intensity and background subtraction

For a given MS signal, one has to choose to integrate only the first fully desolvated peak only (+0 adducts - Figure 22) or a larger m/z -window that embraces adducts considered as belonging to the same species. In Figure 22, three spectra of M at 10 μ M in trimethylammonium acetate (TMAA) 100 mM, with 0, 0.5 and 2 equivalents of L are shown. A background contribution to the peak should be subtracted; however, the definition of peak background is a choice of the user and thus a source of uncertainty on the final values. One simple solution is to take an adjacent m/z -window (to the peak) as representative background. This is feasible for well-separated peaks, but overlap problems may arise in case of close peaks, or peaks with long adducts tails. A stratagem is to consider a narrow m/z -window with the minimal background intensity (i.e. orange and bright blue ranges in Figure 22) as representative background for the second.

The choice of only one peak (no adducts), like the M^{n-} in Figure 22, has the advantage of a better signal-to-noise ratio, but it induces a bias in the case of variations in the adduct distribution over the titration. The choice of summing up all adducts, like M^{n-} and 3 TMA⁺ in Figure 22, has the advantage of a higher intensity and it accounts for eventual variations in adducts distribution, at cost of a higher background that could lower the signal-to-noise ratio and eventually increase uncertainties after subtraction.

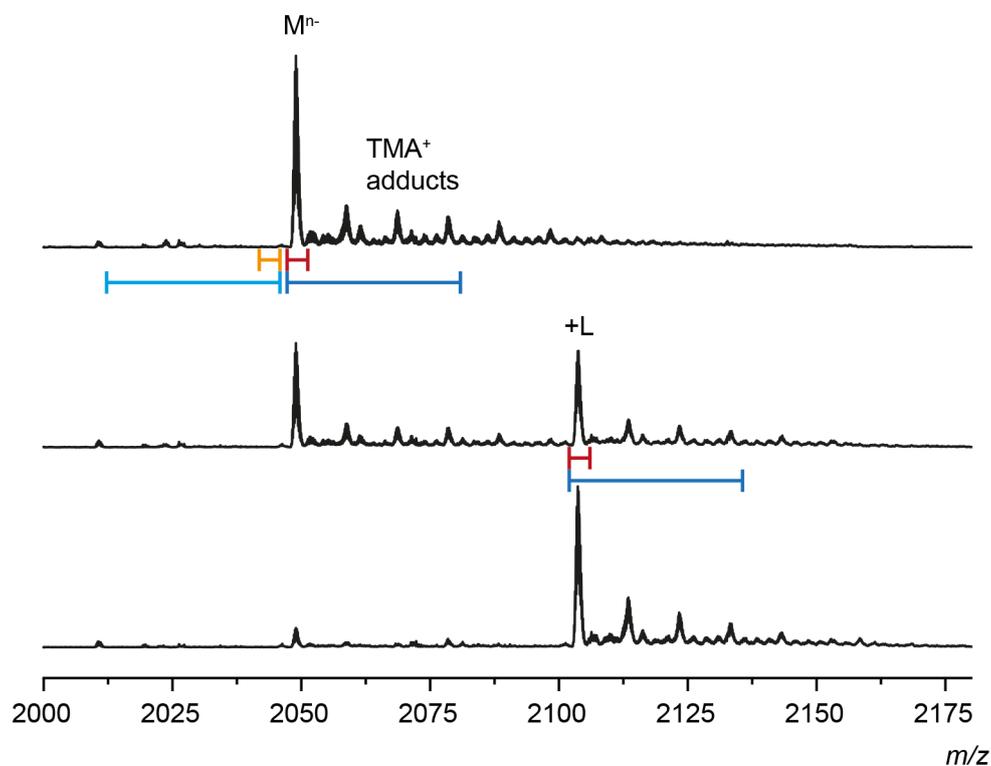


Figure 22 – Mass-to-charge range definition for MS peak intensity extraction, for an isolated peak (M) and two close-range peaks (M and ML). Fully desolvated peak only (**red**) and up to 3 adduct (**dark blue**) ranges are depicted. The background range is depicted in **orange** and **bright blue**, respectively. Spectra correspond to charge state 6⁻ of MGA (M^{n-}) 10 μ M in TMAA 100 mM with 0, 0.5 and 2 equivalents of MG (L).

3.6 Ion Mobility methods

3.6.1 CCS determination and CCS distribution reconstruction

An IM-MS spectrum can be represented as a three-dimensional plot with m/z on abscissa, arrival time (ms) on ordinate and intensity (counts) in z-coordinate, usually rendered with a heat map in false colours (Figure 23A). On the IM-MS spectra, one can readily read the typical information about of conventional MS spectra (m/z ratio of ions, isotopic pattern, stoichiometry) with the benefit of the ion mobility separation on the arrival time axis.

In Figure 23A we have a 7- ion of a DNA aptamer with its NH_4^+ adduct distribution. The ion mobility and its corresponding averaged CCS are determined by recording ion's arrival time distributions (ATDs) at many drift voltages. For direct CCS determination, IM-MS spectra are recorded using five drift voltages (390.5, 490.5, 590.5, 690.5 and 790.5 V). This method is termed "step-field". The arrival time (t_a) corresponds to the peak maximum determined by Gaussian fit of the ATD, in our case using the software PeakFit 4.11 (Systat Softwares, San Jose, CA-USA) using "GaussArea" with smoothing Savitzky–Golay 5% and no baseline or Origin 2016. The linear regression of the arrival times t_a as function of $1/\Delta V$, eq. (24), gives access to t_0 and K_0 (Figure 23B). Charge and reduced mass are known from the mass spectrum, and the CCS of the peak maximum can thus be calculated using eq. (24) (Figure 23C).

When subjected to pre-IMS collisional activation (with a single drift voltage of 390.5 V), the M^{7-} in Figure 23D loses part of its NH_4^+ adduct distribution. From the arrival time, we notice that the MS spectrum is the superimposition of two form of M^{7-} : One more compact (M_C^{7-}), with its NH_4^+ adduct distribution, the second more extended (M_E^{7-}) and with fewer NH_4^+ adducts.

In order to compare these two peaks in the same ATD, we used the empirical method of single-field reconstruction reported recently ^[99].

$$CCS = a \frac{z}{\sqrt{\mu}} t_a, \text{ where } a = \frac{CCS(\text{step-field}) \sqrt{\mu}}{t_a(390.5V) z} \quad (39)$$

Where $t_a(390.5 V)$ and CCS are determined from the previous IM-MS step-field acquisition (Figure 23A-C). The parameter a takes into account instrumental contributions of CCS in the total drift time. This reconstruction is valid only for ions of the same charge state (z) and have been done in SigmaPlot 12.5.0.38 (Systat Softwares, San Jose, CA-USA) using eq. (39) (script at p. 148).

Compared to step-field, the second method is more imprecise but is convenient to extract CCS values for the visualization of collisional induced unfolding experiments (described in next paragraph).

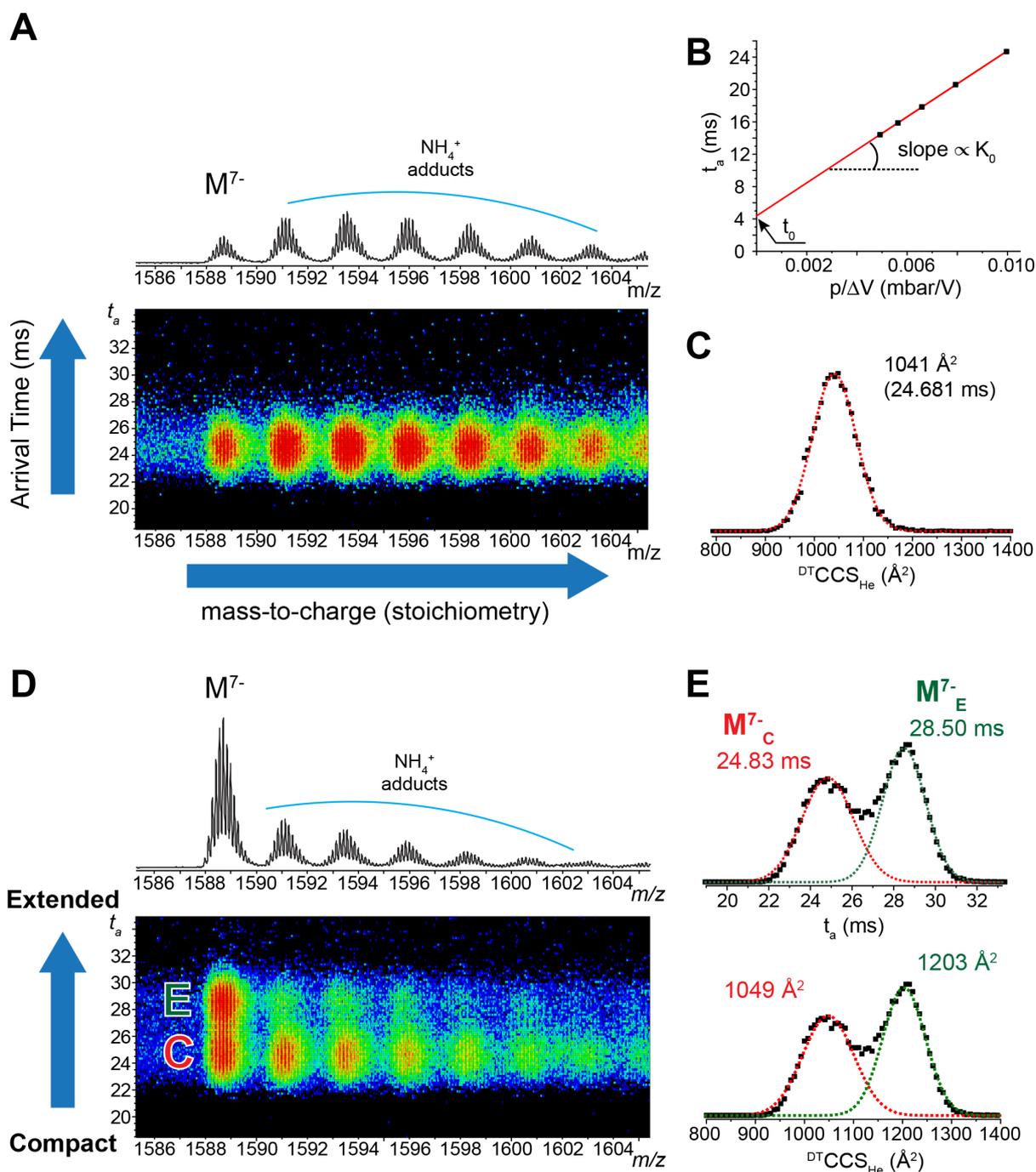


Figure 23 – **A**. IM-MS spectrum of a M^{7-} ion ($Frag = 350 \text{ V}$, $\Delta V = 390.5 \text{ V}$); **B**. Linear regression of t_a of step-field segments; **C**. derived CCS distribution with its center value. **D**. A second, more extended, conformation is visible for M^{7-} at fragmentor 530V. **E**. The two ATD peaks for M^{7-}_C and M^{7-}_E are reconstructed to CCS via eq. (39).

3.6.2 Collision-induced unfolding experiments

3.6.2.1 Principles

Instrumental parameters before IM drift tube affect the compactness of ions. Collision-induced unfolding is a popular pre-IMS activation experiment largely described in the literature^[59, 65-66, 110]. On our instrument, we need first a first step-field IM-MS acquisition, to determine the value of α to be used in eq.(39). We typically performed this at low activation energies (i.e. fragmentor at 350 V). This acquisition defines the $^{DT}CCS_{He}$ distribution of IM-spectra at each activation point

Two activation parameters are accessible on Agilent 6560: trap entrance grid delta (TEGD) and fragmentor (details at p.31). In this thesis the CIU are carried by varying the fragmentor voltage (if not otherwise specified) from 350 to 600 V. The fragmentor voltage is increased by steps of 10 V, and IM-MS spectra are recorded for 1 minute per voltage segment. The ATDs are extracted using Agilent IM-MS Browser v. B.07.01 (Build 7.01.147.0). Chosen the m/z range of a given ion, its ATD is normalized to 1, segment per segment, and converted in CCS. Segments converted this way can be represented as in the stacked plot in Figure 24A. This plot is convenient to fit (Gaussian area fit) the populations that contributes to the CCS distribution, and extract the peak areas to build the breakdown diagrams to define the transition point(s) between the populations (Figure 25). A common graphical visualization of CCS segments is the contour plot in Figure 24B (CIU plot), where axes of Figure 24A are transposed and the normalized intensity is rendered in false colours.

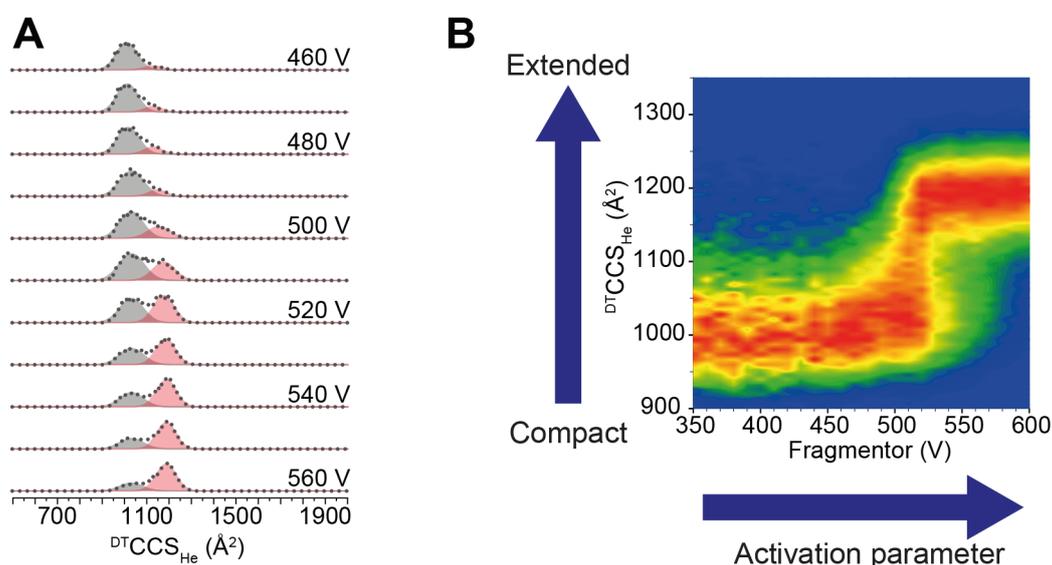


Figure 24 – **A.** Extract of ATDs converted to CCS distributions (normalized to 1) and gaussian fitting of the two distributions (black and red areas) that contribute to the signal. **B.** CIU contour plot obtained by transposed axes from A.

The observation of the CIU plot in Figure 24B is useful to have a glimpse of how the ions respond to collisional activation. For example, the plots reveal if the extension pass through intermediate shapes (according to their CCS values), or to have an insight on the energetics of the transition (according to the voltage at which a shape change occurs). An unfolding that occurs at lower voltage suggests that weaker intramolecular interactions are involved in the compact structure. However, the colour coding of CIU plots (Figure 24B) makes the transitions look sharper than when the actual breakdown curves are constructed (Figure 25). We used the CIU breakdown plots to extract the transition voltages.

3.6.2.2 How to construct CIU plots and breakdown diagrams in practice

IM-MS segments are exported using Agilent IM-MS Browser and ATDs extracted with an in-house developed R script. SigmaPlot 12.5 has been used to normalize the extracted ATDs (script at p. 148) and build the CIU plot in Figure 24B using a xyz-plot scheme for heat-maps (Figure 92 - p.149). Areas of each CCS distribution are obtained by fitting the ATD with Gaussian curves in Origin 2018 SR1 b9.5.1.195.

CIU breakdown plots

CIU breakdown plots are built with the MS relative intensities of compact and extended ions, per each fragmentor voltage. Normally only desolvated ions are used, except when specified (as discussed at p. 124 - Figure 79)

Data points can be m/z integration of a defined arrival time (or CCS) interval (Figure 25A), as described just above. Alternatively, Gaussian area fitting of CCS distributions (Figure 24A) can be used as data points to built the breakdown curve in Figure 25B. In principle, given the same m/z range, the two ways to built the plot are superimposable.

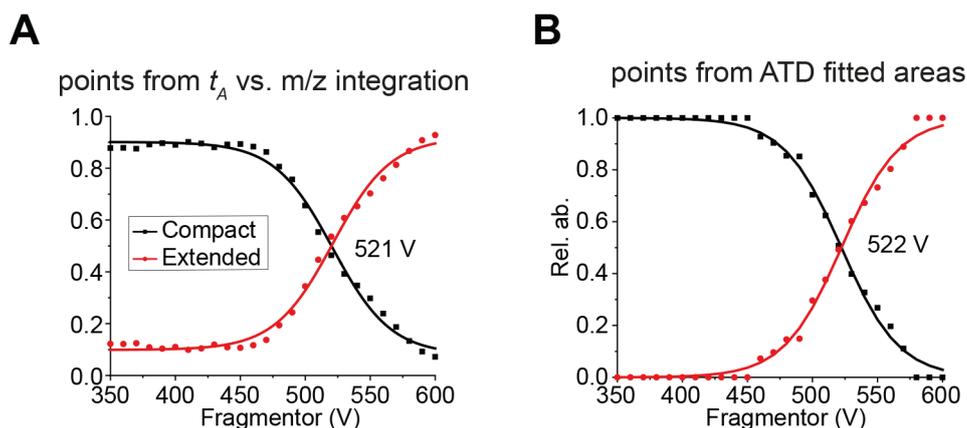


Figure 25 - CIU breakdown plot obtained from **A.** m/z integration of defined attival time intervals (see Figure 26); **B.** fitted area of ATDs vs activation parameter. Mid-point of transition is determined by sigmoidal fit following eq. (40).

However, the first method should be preferred because ATDs from noisy data may encounter some difficulties in fitting procedure. For instance, points below 450 V and above 560V in Figure 25B are forced to 0 and 1 (following the evolution of the curve). The characteristic mid-point of transition is then obtained by sigmoidal fit of the points.

A sigmoidal equation as in eq. (40) is appropriate to describe a phenomenon such the ion transition from one compact CCS distribution to a more extended one:

$$y(x) = \frac{A_1 - A_2}{1 + e^{(x-x_c)/dx}} + A_2 \quad (40)$$

Where A_1 and A_2 are the initial and final state relative abundancy, respectively. And x_c is the critical voltage value in the middle of the transition, for which $y(x_c) = 0.5$.

x_c , from here called transition voltage, is obtained via non-linear regression analyses such as the least squares method on data points of CIU plots. Relative abundancies A_1 and A_2 are not fixed to take into account the stochastic error of data points (if not mentioned).

CID/U diagrams

MS Intensities of compacted and extended population has been extracted via Agilent IM-Browser selecting two $t_a \times m/z$ range for the two CCS distributions, alike the example in Figure 26: For the two conformation a threshold value around 1100 \AA^2 as been considered to define the two $t_a \times m/z$ areas to extract (p.214) . The MS spectra corresponding of each area, is exported as .csv data file (m/z vs int). The procedure is repeated for each segment of the CIU experiment. From each single MS export is created a two column (m/z vs Intensity) .txt file to extract the integral corresponding to M and ML has been extracted using a python in-house developed script (p.145). $t_a \times m/z$ areas in Figure 26 and m/z integration ranges are reported in annexes at p.214. No background subtraction is applied for these m/z ranges.

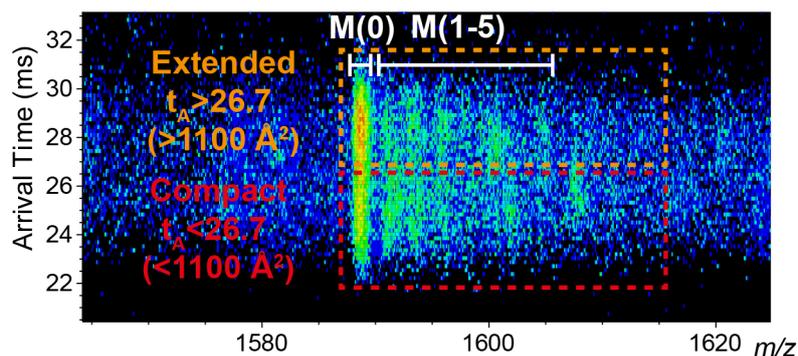


Figure 26 – IM-MS at fragmentor 520 V for charge state 7- for a mixture of MN4 with quinine (DNA and quinine 10 μM) in NH_4OAc 25 mM. Data selection for CIU breakdown plots and CID/U diagram. t_A for extended ion are in orange, compact are in red. m/z -ranges for aptamer and complex, are reported in white with the number of NH_4^+ adducts in brackets.

Integral values obtained as described, are used as data points to build the CIU breakdown plots and CID/U diagrams in Origin 2018. CID/U diagrams are built with MS relative intensities of all ions (i.e.: compact $M(0)$, $M(1-5)$ and extended $M(0)$, $M(1-5)$ - Figure 27) per each fragmentor voltage. The integral values of each ion is normalized over their sum per each fragmentor voltage.

These relative intensities are plotted as a stacked diagram on Origin 2018, normalized to 1 per each fragmentor voltage.

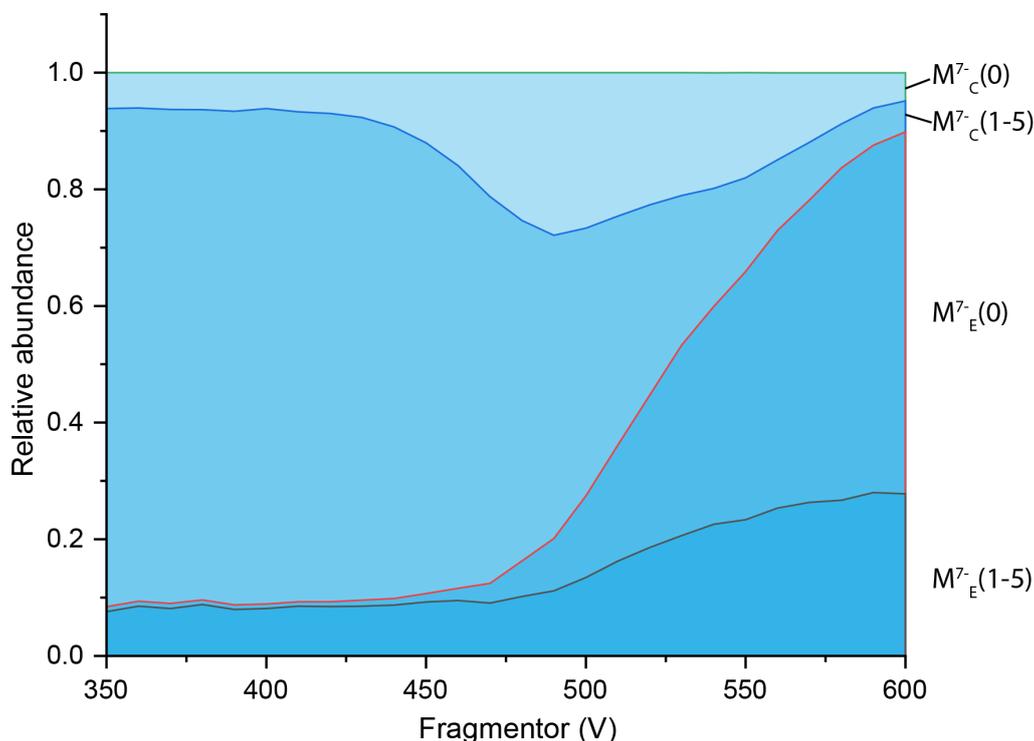


Figure 27 - Example of CID/U diagram. M^7 intensities from Figure 26 are normalized to 1 per each fragmentor voltage. Subscript "c" states for compact ions, "e" for extended ions. Their relative abundance is plotted as stacked diagram (see text).

3.7 Sample preparation

3.7.1 Nucleic acids

DNA and RNA sequences were obtained from Integrated DNA Technologies (Leuven - Belgium) in standard desalting purified lyophilized form and dissolved in RNase-Free water from Ambion™ (ThermoFischer Scientific, Illkirch, France) at 60°C for 10 min. The concentrations were determined using the Beer-Lambert law by measuring the absorbance at 260 nm and adjusted between 0.4 and 1.0 mM concentration. DNA and RNA extinction coefficients provided by Integrated DNA Technologies were obtained applying the Cavaluzzi-Borer correction^[111]. Samples were desalted on Amicon Ultra-0.5 mL centrifugal filters, 3 kDa cut-off, in successive steps with trimethylammonium acetate 500 mM (x2, 13 krpm, 30 min), ammonium acetate 500 mM (x2, 13 krpm, 30 min), and water (x5, 13 krpm, 30 min). Desalting allows to remove traces of contaminants (e.g.: alkali, deprotection reagents) that may interfere with sample folding or simply its analysis. After desalting, 100 µM DNA solutions were prepared in water for further sample preparation by dilution in desired buffer.

The d(TG₄T)₄ and dT₆ (10 and 5 µM) solution for daily helium pressure setting was prepared as follows: d(TG₄T) 800 µM single strand was incubated overnight at 4°C in NH₄OAc 150 mM to allow the formation of d(TG₄T)₄ G-quadruplex 200 µM. Final solution was obtained by dilution and adding appropriate aliquot of dT₆ 1mM. For quantitative MS, samples were prepared from solutions of 100µM DNA or RNA with 25 µM dT₆ used as internal standard, to ensure a constant ratio of dT₆ with the analytes for the relative (to dT₆) response factor determination (p.39).

Tetracycline RNA Aptamer preparation

TCA was provided from Dr. Vogel Marc (Pr. Suess Lab – Darmstadt, Germany). We report here their procedure to RNA transcription.

RNA was transcribed in vitro using a HindIII linearized plasmid as template. The plasmid codes for the T7 promoter followed by the aptamer sequence and a self-cleaving hepatitis delta virus (HDV) ribozyme, to ensure homogeneous 30-ends. In vitro transcription was performed at 37°C overnight in a total volume of 10 ml 20 mM magnesium acetate (0.2M Tris-HCl pH 8.0, 20 mM DTT, 2 mM spermidine, 0.2 mg/ml linearized plasmid, 4 mM of each NTP, 7.5 mg/ml of T7 polymerase (made in-house)). After transcription, precipitated pyrophosphate was pelleted by centrifugation and 10% (v/v) EDTA (0.5 mM, pH 8.0) was added to the supernatant. After ethanol precipitation, we performed a denaturing polyacrylamide gel electrophoresis (8% PAA, 8M urea). The RNA was detected via ultraviolet (UV) shadowing, cut out, and eluted from the gel in 0.3 M sodium acetate pH 6.5 at 4°C overnight. To remove the remaining gel slices, the supernatant was filtered using a 0.45 µm filter and again the RNA was precipitated using ethanol. Finally, the RNA was redissolved in double-distilled water and stored at 20°C

DNA Sequences used:

Name	Sequence (5' → 3')	MW (Da)	$\epsilon(260\text{nm})$ [cm ⁻¹ M ⁻¹]
T ₆	TTT TTT	1763.22	48406
TG ₄ T	TGG GGT	1863.3	57800
MN4	GGC GAC AAG GAA AAT CCT TCA ACG AAG TGG GTC GCC	11128.3	354516
MN19	GAC AAG GAA AAT CCT TCA ACG AAG TGG GTC	9273.1	303242
MN19-A	AAA GAC AAG GAA AAT CCT TCA ACG AAG TGG GTC AAA	11152.3	377106
OR8	ACA AGG AAA ATC CTT CAA CGA AGT GGG T	8654.7	285742
OR8-A	AAA AAC AAG GAA AAT CCT TCA ACG AAG TGG GTA AAA	11160.4	382496
OR7	CAA GGA AAA TCC TTC AAC GAA GTG GG	8037.3	263839
OR7-A	AAA AAC AAG GAA AAT CCT TCA ACG AAG TGG GAA AAA	11169.4	384911

RNA Sequences used:

Name	Sequence (5' → 3')	MW (Da)	$\epsilon(260\text{nm})$ [cm ⁻¹ M ⁻¹]
MGA	GGA UCC CGA CUG GCG AGA GCC AGG UAA CGA AUG GAU CC	12300.5	377500
1FMN	GGC GUG UAG GAU AUG CUU CGG CAG AAG GAC ACG CC	11337.9	347200
1NEM	GGA CUG GGC GAG AAG UUU AGU CC	7442.5	232700
1TOB	GGC ACG AGG UUU AGC UAC ACU CGU GCC	11337.9	347200
TCA	GGG CUA AAA CAU ACC AGA UUU CGA UCU GGA GAG GUG AAG AAU ACG ACC ACC UAG CUC A	18690.4	659600

3.7.2 Ligands

From preliminary ITC studies using the published molar extinction coefficient, ^[112] malachite green concentrations appeared overestimated (fitted stoichiometry $N > 1$ – Thermograms shifted toward the right), thus we re-determined $\epsilon(617)$. The purity grade of malachite green was verified by quantitative NMR (qNMR), using the relative quantification method on a Bruker 700 MHz NMR. In the relative quantification method, one can define the molar ratio of analyte X vs the internal standard (Std) as:

$$\frac{n_X}{n_{Std}} = \frac{I_X}{I_{Std}} \frac{N_{Std}}{N_X} \quad (41)$$

Where n are the number of moles, I the peak integral and N the number of nuclei of the integrated peak. This way, knowing the number of moles of the internal standard added into the tube, it is possible to deduce the right molar ratio with the analyte using the ^1H peaks ^[113].

Sample preparation qNMR

A 2,2-Dimethyl-2-silapentane-5-sulfonate (100% - Sigma 178837) (DSS) solution of known concentration has been used as internal standard, starting from the exact weight of the DSS sodium salt. 1.046 mg of DSS- d_6 (Prd n° 178837 – Batch n° BCBM2365V) were weighted using a Sartorius ME5 microbalance (Sartorius France S.A.S. - Aubagne), and dissolved into 1 mL of Nuclease-free water. From this stock solution (4791 μM) we obtained, by dilution, a 1500 μM “standard” solution used into relative quantification.

A Malachite green solution at exactly about 1.5 mM has been prepared based on the nominal purity (97.3%) of the batch (Prd n° 38800 – Batch n° 1405586): 0.751 mg of MG were weighted and dissolved into 1.3 mL of nuclease-free water. An aliquot of this stock (1540 μM) was then diluted to 1000 μM in nuclease-free water + 10% D_2O .

Finally 5-mm NMR tube was prepared with 400 μL of ligand solution in D_2O 10% + 50 μL of DSS (Std) 1500 μM , and Parafilm M (Bremis Company, Inc. – WI (USA)) sealed until ^1H -NMR spectra acquisition (Figure 28). From spectrum in Figure 28 we obtained a number of moles of 342 ± 10 nmol for malachite green in 400 μL . The stock solution was therefore deduced to be 1317 μM of malachite. Based on the Beer-Lambert law.

$$A(\lambda) = \epsilon(\lambda)l[C] \quad (42)$$

A series of solution from 8 to 1 μM has been prepared by dilution in water from the malachite green stock (1317 μM), determined by qNMR. $\epsilon(617\text{nm}) = 81675 \pm 543$ (M cm)⁻¹ has been determined by linear regression of $A(617 \text{ nm})$ vs Ligand concentration Figure 29.

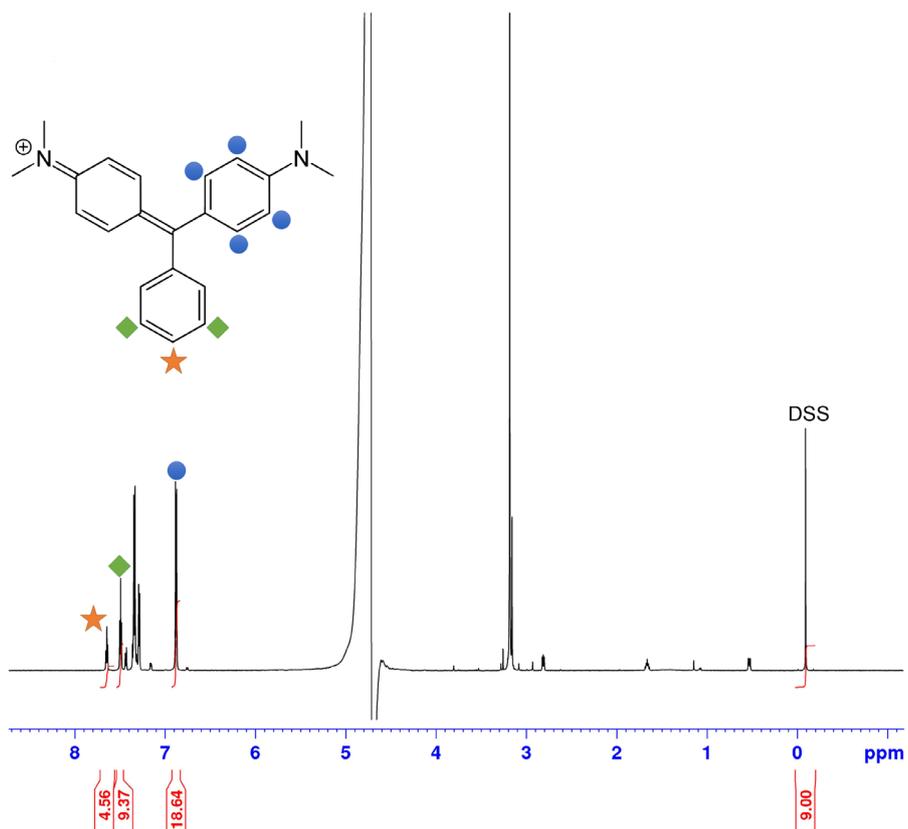


Figure 28 - Malachite green chloride resonances (and assignment) used for relative quantification to DSS methyl singlet. 400 μ L in H₂O/D₂O (9:1) in 1mM 50 μ L DSS – 1.5 mM at 298 K.

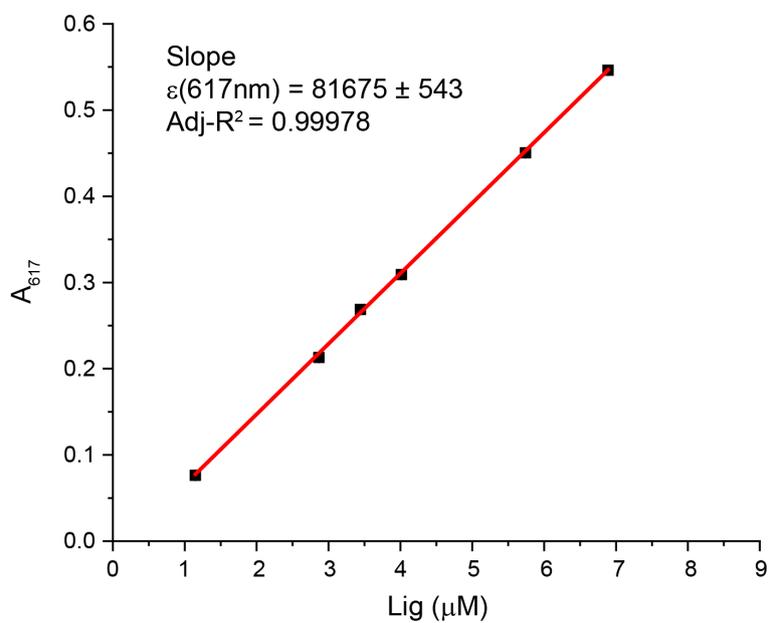


Figure 29 - Linear regression of extinction coefficient at 617nm for ligand malachite green chloride.

Ligands used

Ligands has been bought by Sigma Aldrich (Saint-Quentin Fallavier, France) and Janssen (Beerse, Belgium):

Name	Ref - ID	MW (Da)	$\epsilon(\lambda_{\max})$ [$\text{cm}^{-1} \text{M}^{-1}$]*
Malachite green chloride (analytical standard)	SA - 38800	364.91	$\epsilon(617 \text{ nm})= 81675 \pm 543$ (this work)
Riboflavin 5'-monophosphate sodium salt hydrate (synthetic)	SA - F2253	478.33	$\epsilon(347 \text{ nm})= 12500$ ^[114]
Quinine anhydrous	Janssen Chimica – 16.370.74	324.42	$\epsilon(349 \text{ nm})=5700$ ^[115]
Neomycin trisulfate salt	SA – 1458009 (USP std)	908.88	n.d.
Tobramycin standard	SA - T1500000 (EP std)	467.51	n.d.
Tetracycline hydrochloride	SA - T3383	480.90	$\epsilon(355 \text{ nm})=13320$ ^[116]
Ammonium acetate (BioUltra, ~5 M in H ₂ O)	SA - 09691	77.08	n.d.
trymethylammonium acetate (BioUltra, ~1 M in H ₂ O)	SA - 92286	119.16	n.d.
Magnesium chloride (BioUltra, ~1 M in H ₂ O)	SA - 63069	95.21	n.d.

Part I: Quantitative MS

4. Quantitative comparison of Native MS and ITC

4.1 Why comparing native MS with ITC?

We want to establish how mass spectrometry can complete the picture in aptamer characterization, first to resolve unambiguous stoichiometry assignment of ligands and co-factors (i.e.: Mg^{2+}); secondly, to quantify binding (K_D). We also want to establish the lower and higher K_D limits attainable.

Phosphate buffers with alkali and alkaline-earth counterions (i.e. K^+ , Mg^{2+}) are commonly employed in nucleic acid and aptamer studies, to mimic the ionic strength of physiological solutions (up to 150 mM)^[117]. These conditions are compatible with Isothermal titration calorimetry (ITC), NMR and other liquid phase techniques, but are unfortunately detrimental to ESI-MS analysis, because of abundant clustering on analyte ions and ion suppression due to salt and buffer cluster ions.

The first ESI-MS studies on aptamers were run in hydroalcoholic solutions containing ammonium acetate (50 mM aqueous NH_4OAc + 10 to 50 % iPr-OH) and binding was successfully detected^[118]. Recently Gulbakan *et al.*^[119] compared MS and ITC data of some DNA aptamers in NH_4OAc (up to 300 mM), however not in presence of Mg^{2+} .

NH_4^+ has an ionic radius similar to that of other common alkali cations^[120], and thus stabilizes nucleic acid structure. On the other hand, NH_4^+ may compete with additives such as K^+ or Mg^{2+} (e.g. detection of lower Mg^{2+} stoichiometries). In native MS, manganese (Mn^{2+}) is often used as a surrogate for Mg^{2+} ^[98, 118a] because its higher mass gives well-spaced adducts, and prevents overlapping with NH_4^+ or Na^+ adducts. Yet, although Mn^{2+} has a very similar ionic radius to Mg^{2+} , it belongs to the d-group and its different chemistry could make it a questionable surrogate of alkaline-earth metals^[121]. Our lab showed that trimethylammonium acetate (TMAA) is adapted to study specific interaction of alkali cations with DNA, in this case K^+ in G-quadruplex structure^[122]. We think that TMAA could be a good candidate electrolyte to study Mg^{2+} -aptamers interaction for three reasons: **i.** TMAA⁺ bulkiness^[123] prevents it from competing for specific Mg^{2+} sites^[122]; **ii.** TMAA⁺ adduct distribution is spaced enough to avoid superimposition with additives such as Mg^{2+} ; **iii.** TMAA⁺ can be easily removed through collisional activation (e.g. MS² cell) leaving only metal adducts on the adduct distribution of the analyte ion. However some precautions are needed, because TMAA buffer can be detrimental for kissing-loop interactions^[7].

In this chapter, we show that native MS buffers at biologically relevant ionic strength (i.e. 100 mM) are suitable to study RNA aptamers, including in presence of Mg^{2+} . Native MS and ITC are used to determine dissociation constants (K_D) in comparable experimental conditions ([RNA], [Ligand], buffers, [Mg^{2+}], ...), and ITC is used to validate the constants determined by MS.

4.2 Malachite Green aptamer: a model RNA aptamer to compare ITC and MS

The malachite green aptamer (MGA) is a 38-nt RNA with riboswitch activity that binds the malachite green (MG) dye^[35c, 85]. NMR and X-ray structures are available for its complexes, respectively with MG^[85] (Figure 30) and its analogue tetramethylrosamine (TMR)^[32].

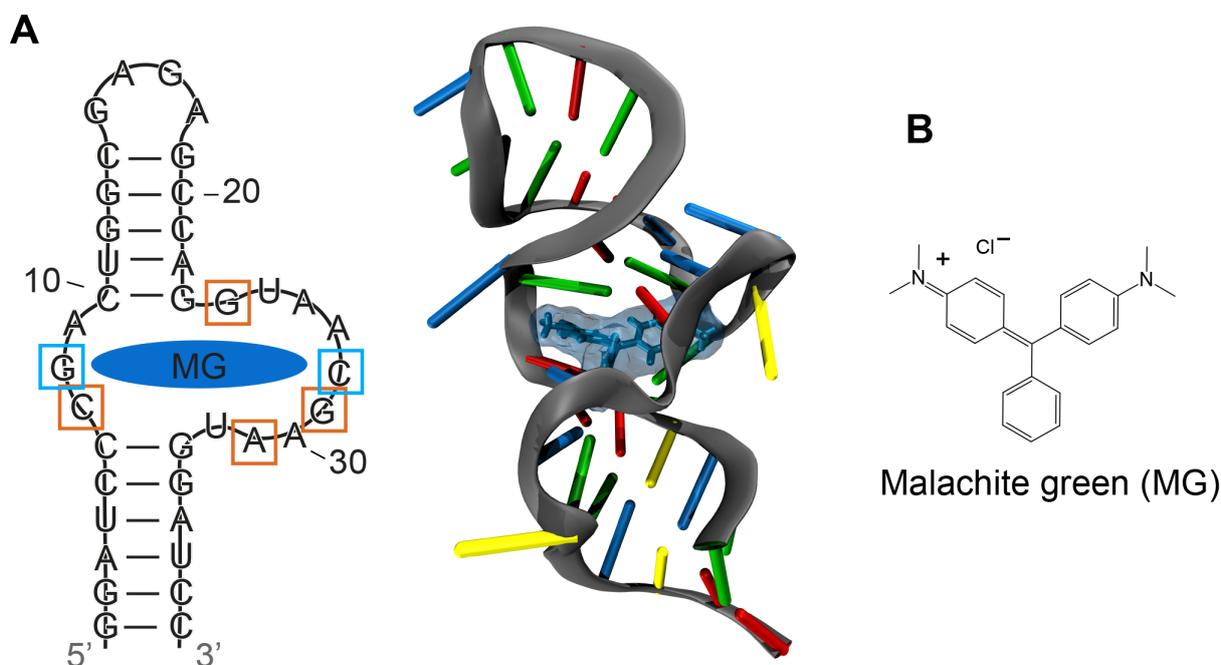


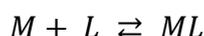
Figure 30 – **A**. Secondary structure topology^[80a] and NMR structure (1Q8N.pdb) of malachite green aptamer complexed with its ligand malachite green chloride (MG)^[85]; The binding pocket is constituted by a base quartet (orange) and a tilted WC base pair (light blue). Readapted from; **B**. Malachite green chloride structure; Color code: **A-Blue; G-Green; U-Yellow; C-Red**.

Both MG and TMR bind by intercalation, and the complex is stabilized by π -stacking with nucleobases and electrostatic interactions^[32, 85]. Compared to MG, the planarity of TMR leads to a stronger affinity with the aptamer^[32]. A K_D in the same range as the RNA concentration is preferable, in order to have informative binding isotherms^[124]. MG binding has been studied with ITC in four different buffers and Mg^{2+} was not essential for binding^[80b, 85]. MG showed K_D s of 0.1 - 0.8 μM in these buffers.^[80b]

MGA-MG binding is thus a good candidate complex for our aims: K_D in “ μM ” range is a suitable concentration range for both ITC and MS and the absence of Mg^{2+} dependence simplifies the binding and experiment design in MS.

4.2.1 MS buffers TMAA and NH₄OAc do not alter malachite green binding mode

With no participation of Mg²⁺, malachite green aptamer binding reaction follow the simplest scheme:



In 150 mM NaCl the K_D is 0.28 μM, ΔH° = -25.93 kcal/mol and -TΔS° = 16.95 kcal/mol (T = 298 K) [80b], the typical thermodynamic profile of intercalating ligands^[50b]. Our ITC runs were acquired in TMAA and NH₄OAc 100 mM at 25°C (Figure 31), in four or five replica (p.177) with the methods described at paragraph 3.3. The nature of the electrolyte does not alter the thermodynamic profile of binding, which remains enthalpically driven (ΔH°<0). The apparent K_D between NH₄OAc and TMAA is comparable within the uncertainty range (Table 7). This suggests that MGA binding mode to MG is unchanged from NaCl to TMAA or NH₄OAc, and that the intercalation binding mode is maintained.

Table 7 – ITC parameters obtained form OneBindingSite fitting model. Experiments were done in triplicate and uncertainties are the standard deviations across the 5 replicas in NH₄OAc and 4 replicas in TMAA (p.177).

Electrolyte	K _D (μM)	N	ΔH° (kcal/mol)	-(298)ΔS° (kcal/mol)	ΔG° (kcal/mol)
NH ₄ OAc 100 mM (A)	0.85 ± 0.48	0.99 ± 0.07	-14.5 ± 0.7	6.2 ± 2.4	-8.37 ± 0.32
TMAA 100 mM (B)	1.51 ± 0.58	1.04 ± 0.05	-14.6 ± 0.7	6.6 ± 2.8	-7.99 ± 0.25

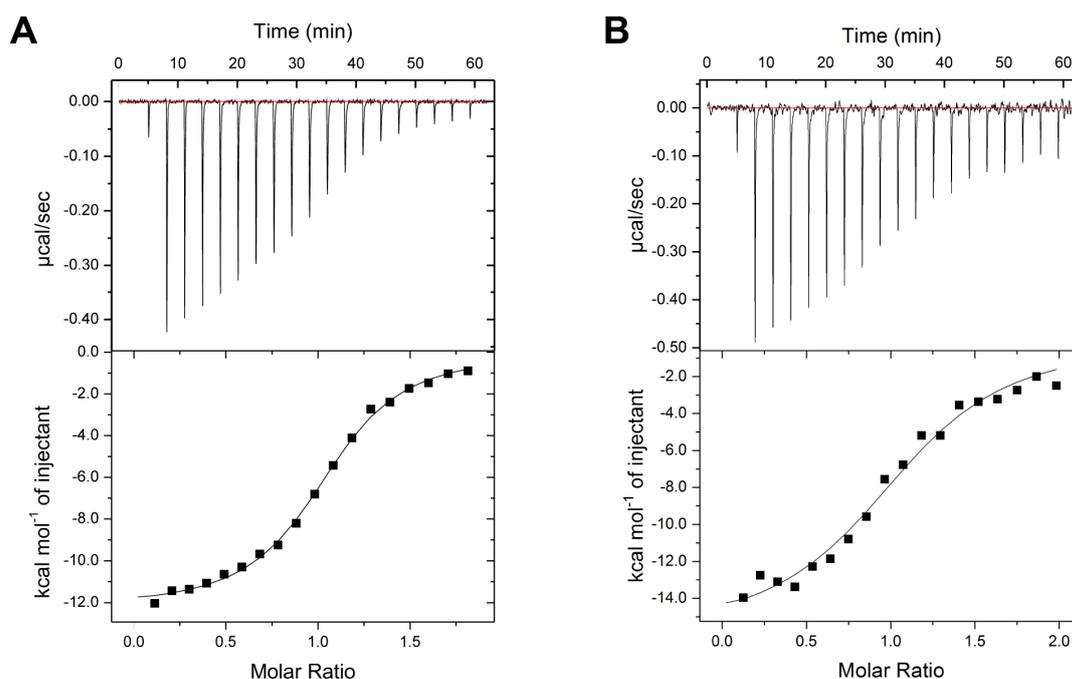


Figure 31 - ITC of A. malachite green aptamer 15μM vs 142 μM MG in NH₄OAc 100 mM and B. 139 μM MG TMAA 100 mM at 25 °C. The single replicas are at p.177.

4.2.2 Determination of MGA-MG binding constant with native MS titration

To verify the stoichiometry, we tested two mixtures of 10 μM MGA with 5 μM of MG in 100 mM TMAA and NH_4OAc (Figure 32). MGA is ionized to the 7-, 6- and 5- charge states in both buffers, and ligand binding is detected for all charge states. In TMAA, slightly lower charge states are favoured, as often for secondary and tertiary amines buffers ^[125].

The aptamer is approximately half-saturated in TMAA: MGA and MGA+Lig have nearly the same intensity. However in NH_4OAc the intensities ratios differs visibly across the charge states (i.e., compare 7- and 5-). If we were assuming that intensities reflect the concentrations in solution (ratio of relative response factors is equal to 1, p.39), we can define the equilibrium dissociation constant of the following binding reaction:

$$M + L \rightleftharpoons ML \quad \text{where} \quad K_D = \frac{[M][L]}{[ML]} = \frac{I_M}{I_{ML}} [L] \quad (43)$$

And calculate the actual concentrations:

$$[M]_{calc} = [M]_{tot} \frac{I_M}{I_M + I_{ML}} \quad \text{and} \quad [ML]_{calc} = [M]_{tot} \frac{I_{ML}}{I_M + I_{ML}} \quad (44)$$

$$[L]_{free} = [L]_{tot} - [ML]_{calc} \quad (45)$$

From the corresponding intensities of M and ML, we can calculate the K_D per each charge state in both buffers, listed in Table 8.

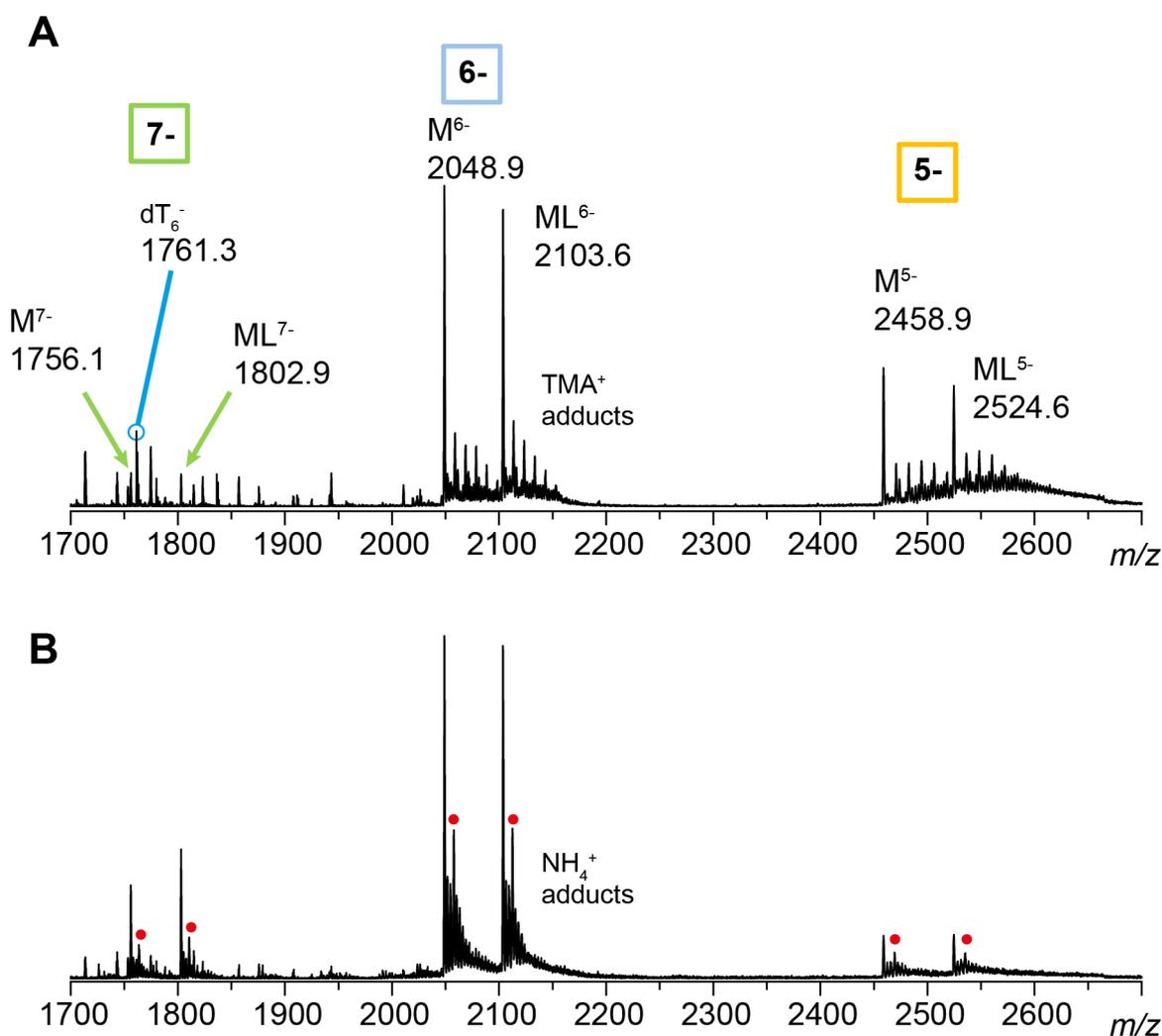


Figure 32 – MS spectra of MGA 10 μM in presence of 5 μM of MG chloride in **A**. TMAA 100 mM and **B**. NH_4OAc 100 mM. Notice that dT_6^- become part of M^{7-} background in spectrum B. The symbol “●” highlight an impurity coming from the stock solution of MGA aptamer.

Table 8 – Average binding partners concentrations from MS spectra of 10 μM MGA and 5 μM MG (Figure 32), calculated using eq. (35) over three replica. The reported uncertainty is the standar deviation over three replicas; *n.d.* are K_D not calculable because of $[L]_{free} < 0$.

	z	$[M]_{calc}$ (μM)	$[ML]_{calc}$ (μM)	$[L]_{free}$ (μM)	K_D (μM)
TMAA 100 mM	7-	5.84 ± 0.96	4.16 ± 0.96	0.84 ± 0.96	1.96 ± 2.47
	6-	5.01 ± 0.21	4.99 ± 0.21	0.01 ± 0.21	0.03 ± 0.23
	5-	4.69 ± 0.55	5.31 ± 0.55	-0.31 ± 0.55	<i>n.d.</i>
NH_4OAc 100 mM	7-	4.18 ± 0.04	5.82 ± 0.04	-0.82 ± 0.04	<i>n.d.</i>
	6-	4.96 ± 0.07	5.04 ± 0.07	-0.04 ± 0.07	<i>n.d.</i>
	5-	5.25 ± 0.75	4.75 ± 0.75	0.25 ± 0.75	0.57 ± 1.17

The problem is that the complex ML has a higher intensity compared to M, for 7- and 6- charge states in NH₄OAc and 5- in TMAA.

After checking across three replica, the [ML] over estimation is recurrent only for ligand concentrations below 5 μM. This could suggest either that the assumption of equal response factors for M and ML is not licit, but could also be related to non-random error like errors in the concentration ratios of ligand to aptamer (for low volumes of added ligand). Thus, a determination of response factors could be necessary, to correctly quantify the binding constant for this complex.

4.2.3 Native MS titration without and with response factor correction

4.2.3.1 Generalities

In order to determine the dissociation constant of MGA, we used the MS titration method (3.5.1 – p.40) up to 20 μM MG while MGA and dT₆ are kept at 10 μM and 2.5 μM, in 100 mM NH₄OAc or 100 mM TMAA (see annexes p.152). MS Thermo Exactive has been used for sample acquisition. The ion current from the last 5 min of 8-min acquisition was summed and exported using Thermo software Xcalibur. Peak areas were obtained by summing the ion counts over the appropriate *m/z* range (p.152). We compared two data treatment strategies: integrating only the desolvated ions (no adducts and less noise), or summing up to 3 adducts (higher integrals, more noise). In each case, the background was defined with the same *m/z* range of the observed peak as at p. 43. Each charge state was considered separately for relative response factors determination. K_D was determined with both methods described at p. 40, first assuming equal relative response factors, secondly with M and ML concentrations corrected for relative response factors. Finally, we evaluate the effect of this correction on the dissociation constants.

MS spectra of titration replica are reported in Figure 100 and Figure 101 at pp.166-167.

4.2.3.2 Native MS titration without response factor correction

Charge state 7- peaks have relevant superimposition with other background peaks, notably in TMAA 100 mM (Figure 32A and Figure 33A). This introduce a large noise on peaks signal, making charge state 7- a bad candidate for quantification. Charge state 5- is also a non-optimal charge state for quantification if one decide to consider the electrolyte adducts for integration. As we can see in Figure 32A, the TMA⁺ adducts distribution provide a significant background distribution. In a case such this, it is preferable to use only the desolvated peak for quantification. Charge state 6- is a good candidate for quantification because the strong peak intensity provides good signal-to-noise ratio for desolvated ions in both electrolytes. On top of this, charge state 6- has also the benefit of sharp peaks and wide TMA⁺ adducts distribution, when sprayed from TMAA (Figure 32A).

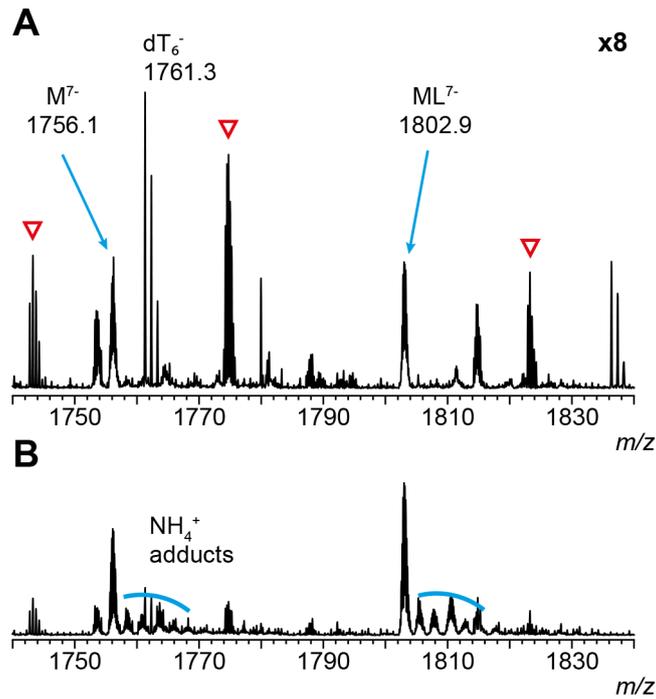


Figure 33 – Detail of charge state 7- for MGA 10 μ M with 5 μ M MG in **A.** TMAA and **B.** NH_4OAc 100 mM. Only the desolvate peaks could be considered for integration, because of intensity and superimposition issues. Spectrum in TMAA is zoomed (x8) compared to NH_4OAc . “ ∇ ” peaks are RNA degradation traces.

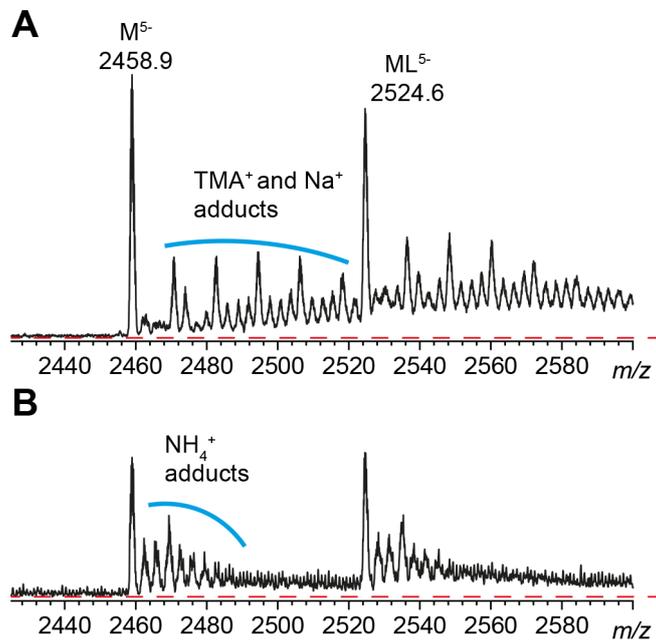


Figure 34 - Detail of background contribution for charge state 5- for MGA 10 μ M with 5 μ M MG in **A.** TMAA and **B.** NH_4OAc 100 mM.

4.2.3.3 On the evaluation of uncertainty

We introduced the two methods to obtain the dissociation constant (K_D) at p. 43. We compare averaged K_D from replicated experiments (Method 1 - Table 9), with K_D determined from averaged MGA and MGA+L concentrations (Method 2 - Table 10). The two ways to treat the data give almost similar results, meaning that both methods are valid. We suggest to follow the first method (independent data treatment on each replica) to better evaluate the quality of data on single replica level. Major advantages are the easier evaluation of errors and outliers. We highlight that standard deviations (outputs of DynaFit software) from method 2 in Table 10 tend to be lower, as expected (p.42).

Table 9 – Averaged K_D over 3 replicas (**Method 1** - Table 27 and Table 28 - p.155). Single charge state K_D is calculated from fitting of binding isotherm **assuming equal relative response factors** for MGA and MGA+L. Outliner values are highlighted in italic; **a.** K_D uncertainty is the standard deviation over 3 replicas; **b.** “add” states for buffer adducts in integrated m/z -ranges.

z	K_D (μM) in TMAA 100 mM ^a			
	0 add ^b		3 add ^b	
7-	0.52	± 0.09	<i>5.38</i>	<i>± 1.34</i>
6-	0.71	± 0.03	0.62	± 0.03
5-	0.78	± 0.07	0.36	± 0.05

z	K_D (μM) in NH_4OAc 100 mM ^a			
	0 add ^b		3 add ^b	
7-	0.65	± 0.24	0.77	± 0.30
6-	1.30	± 0.43	1.42	± 0.43
5-	1.78	± 0.56	1.48	± 0.29

Table 10 - K_D calculated from averaged concentrations (**Method 2** - Table 31 p. 165) of MGA and MGA+L over 3 replicas. Fitting of binding isotherm **assuming equal relative response factors**. Outlier values are highlighted in italic; **a.** K_D uncertainty is calculated from DynaFit statistic with a Monte-Carlo simulation ^[107-108]; **b.** “add” states for buffer adducts in integrated m/z -ranges.

z	K_D (μM) in TMAA 100 mM ^a			
	0 add ^b		3 add ^b	
7-	0.52	± 0.05	<i>5.84</i>	<i>± 0.64</i>
6-	0.73	± 0.05	0.65	± 0.08
5-	0.79	± 0.06	0.35	± 0.14

z	K_D (μM) in NH_4OAc 100 mM ^a			
	0 add ^b		3 add ^b	
7-	0.58	± 0.13	0.84	± 0.15
6-	1.22	± 0.16	1.39	± 0.20
5-	1.72	± 0.18	1.45	± 0.20

Besides important peak superimposition (i.e.: ch st 7-), the choice of m/z -range have a milder impact on the determination of K_D in both electrolytes. In TMAA we have a clear example of

how the choice of m/z -range can influence binding isotherms (Figure 35). While charge state 6- is not significantly affected by the change of the mass range, we observe that the same titrations give a worse curve if the first three TMAA⁺ adducts are included in peak integration for charge state 7-. This is a consequence of superimposed peaks close to MGA⁷⁻ and MGA+L⁷⁻ (Figure 33A). A similar effect is visible for first points of charge state 5-, but for strong background contribution of TMAA⁺ adducts (Figure 34A) ML⁵⁻ signal. However, desolvated ion works best across the three charge states with very similar curves and K_D s. In NH₄OAc 100 mM the two data processing strategies provides equivalent curves: we notice that the choice of m/z -range have a milder impact compared to TMAA.

Conversely, the choice of supporting electrolyte can play a role in the quality of data for which we notice (assuming relative resp fact equal to 1) a two-fold factor between the two, TMAA and NH₄OAc. For example, we attribute charge state 5- isotherms deviations in TMAA (Figure 35) to the background contribution in thus the lower signal-to-noise ratio, compared to 7- and 6- (Figure 34A).

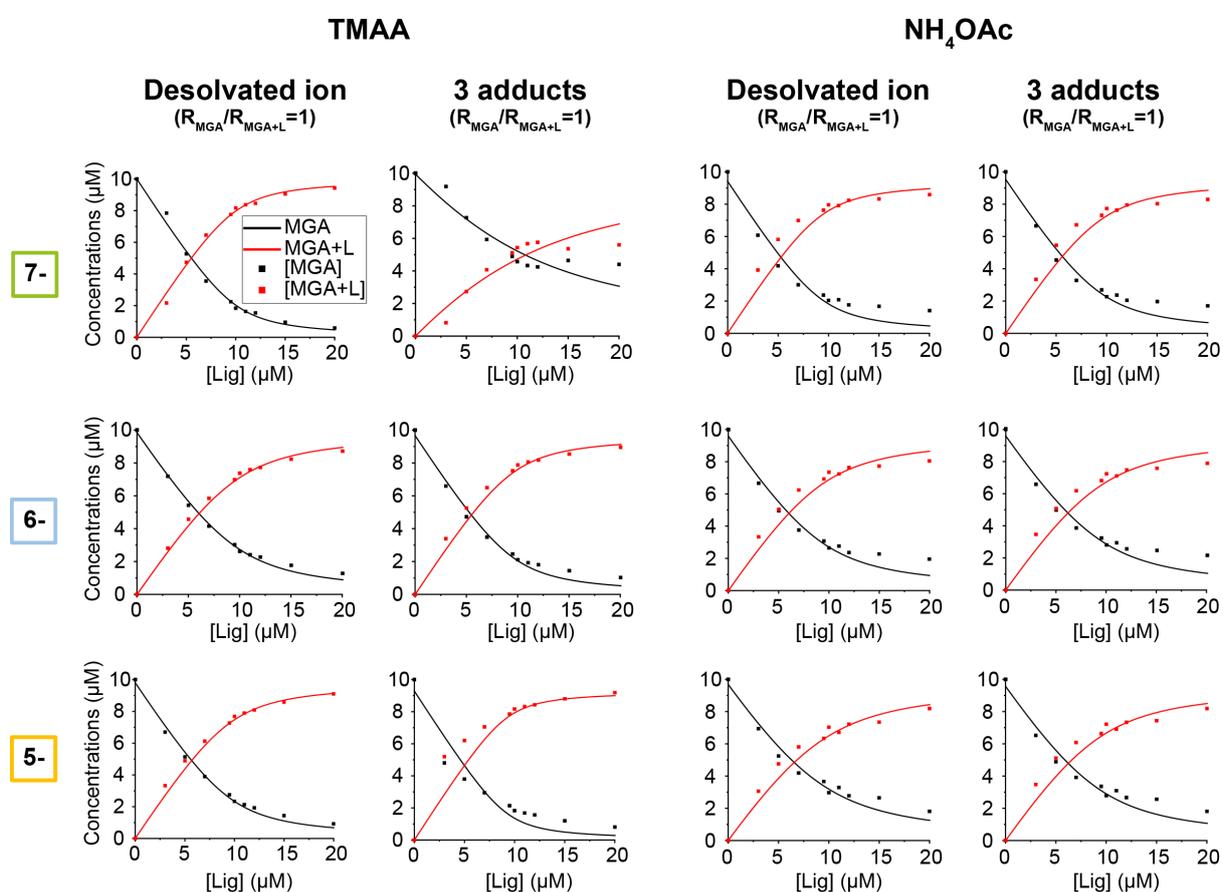


Figure 35 - Binding isotherm from averaged concentrations averaged over three replicas of MGA and MGA+L (**Method 2**) in NH₄OAc and TMAA 100 mM, assuming equal relative response factors for MGA and MGA+L. Desolvated ion or three buffer adduct are compared as peak integration strategies. (Figure 94 – pp.157-162)

The quality of the signal should be always take into account, like before data treatment such as the correction for relative response factor. As we observe in Figure 35, the correction

method is biased by the low signal quality of charge state 7- (considering 3 adducts) in TMAA and charge state 5- in NH₄OAc. In other words, the relative response factor correction method introduces a large error on the final concentrations.

4.2.3.4 Native MS titration with response factor correction

Isotherms are not changing much when taking into account the relative response factor correction. Relative response factors (to the Int. Std. dT₆) were determined as described at p. 39. K_D for the three charge states are listed in Table 11 and corresponding (average) relative response factors are listed in Table 12. (Rf TMAA replicas in Table 25 and Rf NH₄OAc replicas in Table 26).

For quantification purposes we suggest to use the charge state(s) with the highest signal-to-noise ratio, in this case 6-. This will minimize the uncertainty of the measures and, as consequence, on the final calculated values

Table 11 – Averaged K_D over 3 replicas (**Method 1**). Single charge state K_D (Table 27 and Table 28– p.155) is calculated from fitting of binding isotherm with relative response factors correction for MGA and MGA+L; **a**. K_D uncertainty is the standard deviation (and percent error) over 3 replicas; **b**. “add” states for buffer adducts in integrated *m/z*-ranges.

z	K_D (μM) in TMAA 100 mM ^a			
	0 add ^b		3 add ^b	
7-	1.05	± 0.24 (23%)	10.12	± 2.99 (30%)
6-	1.09	± 0.16 (15%)	1.10	± 0.11 (10%)
5-	1.16	± 0.21 (18%)	0.80	± 0.14 (17%)

z	K_D (μM) in NH₄OAc 100 mM ^a			
	0 add ^b		3 add ^b	
7-	1.72	± 0.79 (46%)	1.82	± 1.08 (59%)
6-	1.96	± 0.40 (20%)	2.10	± 0.41 (20%)
5-	4.63	± 1.13 (24%)	4.04	± 1.09 (27%)

Table 12 –Relative response factors used for fitting of binding isotherm; σ^a uncertainty is the standard deviation across the replicas.

TMAA 100 mM										
z	Desolvated					3 add				
	X	rSTD/rX	σ^a	rML/rM	σ^a	X	rSTD/rX	σ^a	rML/rM	σ^a
7-	M	59.62	± 21.85	1.45	± 0.14	M	20.61	± 6.72	1.55	± 0.08
	ML	41.19	± 13.81			ML	13.25	± 3.73		
6-	M	7.12	± 0.25	1.25	± 0.09	M	3.01	± 0.24	1.39	± 0.08
	ML	5.75	± 0.61			ML	2.29	± 0.28		
5-	M	15.60	± 1.01	1.22	± 0.07	M	4.99	± 0.61	1.42	± 0.08
	ML	12.81	± 1.61			ML	3.52	± 0.51		

NH ₄ OAc 100 mM										
z	Desolvated					3 add				
	X	rSTD/rX	σ^a	rML/rM	σ^a	X	rSTD/rX	σ^a	rML/rM	σ^a
7-	M	2.87	± 0.71	1.48	± 0.28	M	2.36	± 1.60	1.50	± 0.29
	ML	1.94	± 0.28			ML	1.58	± 0.86		
6-	M	0.80	± 0.21	1.31	± 0.21	M	0.45	± 0.25	1.26	± 0.25
	ML	0.61	± 0.05			ML	0.36	± 0.20		
5-	M	6.99	± 1.54	1.40	± 0.53	M	2.78	± 1.41	1.72	± 0.43
	ML	4.99	± 2.51			ML	1.62	± 0.71		

The dissociation constants determined from MS titrations vary between 0.3 and 2 μM , correcting either for relative response factor or not (respectively Table 9 and Table 11). Variations in absolute values of K_D s may appear moderate. These variations are very small for MGA^{6-} (Table 13) and are distributed within the same order of magnitude for charge states 7- and 5- (Table 29 and Table 30 – p.163).

4.2.3.5 Comparison with ITC

To understand if the MGA dissociation constant obtained from ITC and MS titrations are comparable, we can compare the obtained K_D values considering their expanded uncertainty (with 95% confidence). Interestingly we found that K_D s (and derived ΔG° - Table 13, Table 14) superimposes within their confidence interval of expanded uncertainty (Figure 36).

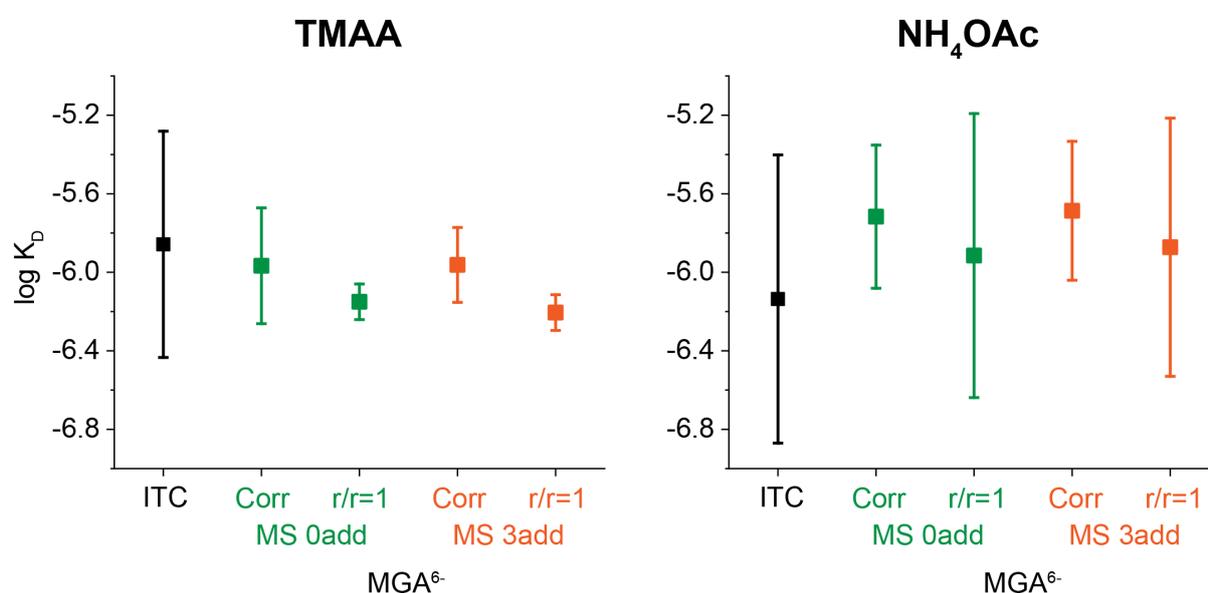


Figure 36 – Comparison between K_D determined via ITC and MS, on MGA^{6-} . The expanded uncertainties (for a 95% confidence interval) are reported as error bars. MS titration were acquired in triplicate (coverage factor = 4.303); ITC (TMAA) in quadruplicate (coverage factor = 3.182) and ITC (NH_4OAc) in pentaplicate (coverage factor = 2.776).

Table 13 - $\log K_D$ and ΔG° calculated from MS titrations of MGA⁶⁻ considering relative response factors correction and assuming equal response factors. “ σ ” is the standard deviation and “Exp. σ ” is the expanded uncertainty over three replicas (coverage factor = 4.303 for a 95% confidence interval) (Table 29 and Table 30 – p.163) .

MGA ⁶⁻ TMAA 100 mM	Corrected			Not corrected (r/r=1)		
		σ	Exp. σ		σ	Exp. σ
K_D (μM)	1.09	± 0.16	± 0.70	0.71	± 0.03	± 0.15
Desolvated LogK_D	-5.97	± 0.07	± 0.29	-6.15	± 0.02	± 0.09
ΔG° (kcal/mol)	-8.18	± 0.09	± 0.40	-8.43	± 0.03	± 0.12
K_D (μM)	1.10	± 0.11	± 0.48	0.62	± 0.03	± 0.13
3 adducts LogK_D	-5.96	± 0.04	± 0.19	-6.21	± 0.02	± 0.09
ΔG° (kcal/mol)	-8.17	± 0.06	± 0.26	-8.50	± 0.03	± 0.13
MGA ⁶⁻ NH ₄ OAc 100 mM	Corrected			Not corrected (r/r=1)		
	σ	Exp. σ		σ	Exp. σ	
K_D (μM)	1.96	± 0.40	± 1.72	1.30	± 0.43	± 1.84
Desolvated LogK_D	-5.72	± 0.08	± 0.36	-5.91	± 0.17	± 0.72
ΔG° (kcal/mol)	-7.83	± 0.12	± 0.50	-8.10	± 0.23	± 0.99
K_D (μM)	2.10	± 0.41	± 1.77	1.42	± 0.43	± 1.86
3 adducts LogK_D	-5.69	± 0.08	± 0.35	-5.87	± 0.15	± 0.66
ΔG° (kcal/mol)	-7.79	± 0.11	± 0.48	-8.05	± 0.21	± 0.90

Table 14 – K_D and ΔG° obtained from ITC on MGA with OneBindingSite fitting model. “ σ ” is the standard deviation and “Exp. σ ” is the expanded uncertainty. Experiments were done in 4 replicas in TMAA (coverage factor = 3.182 for a 95% confidence interval) and 5 replicas in NH₄OAc (coverage factor = 2.776 for a 95% confidence interval) (p.177).

Electrolyte	K_D (μM)	σ	Exp. σ	ΔG° (kcal/mol)	σ	Exp. σ
TMAA 100 mM	1.51	± 0.58	± 1.84	-7.99	± 0.25	± 0.80
NH ₄ OAc 100 mM	0.85	± 0.48	± 1.34	-8.37	± 0.32	± 0.88

Given these pros and cons of using each charge state, for MGA we consider the desolvated peak as best integration strategy. This way we can obtain the highest signal from each charge state. We decided to keep charge state 6- for quantification, because it has the best signal-to-noise ratio and do not present significant integration issue (i.e.: superimposed peaks, difficult background estimation).

The small difference of K_D between the two electrolytes suggest that both NH_4OAc and TMAA do not interfere with ligand binding between MG and MGA as showed in ITC (Table 7). From Table 12 we observe that the ratio r_{ML}/r_M remains between 1.2 and 1.7, which means that the response of MGA+L is not dramatically higher compared to MGA and is almost unchanged between the two electrolytes.

Besides, the discrepancy between K_D s with and without relative response corrected concentrations is limited. Thus, for MGA the assumption of equal relative response factors would not severely bias the quantification of binding partners. Interestingly, this result is in line with previous studies of our lab where a single ligand intercalating in pre-folded duplex, showed a low or negligible variation of complex relative response factors at low stoichiometries [103, 126]. However, it is the first time that a thorough uncertainty assessment and comparison with an independent method has been made.

Comparing results that we obtained from native MS titrations and ITC (Table 13 and Table 14) we have a good agreement, besides we are comparing K_D s obtained from two different techniques and in two different electrolytes. K_D s superimpose within the expanded uncertainty range. Interestingly, we observe the greater uncertainty on NH_4OAc titrations from the K_D s visual comparison with TMAA titrations (Figure 36).

Here the malachite green RNA aptamer is a simple model where only 1:1 binding stoichiometry is allowed and no particular additive, like Mg^{2+} , is needed. The relative response factor correction was useful to take into account the small differences in MGA and MGA+MG intensities during the MS titrations. Now we examine the case of a Mg^{2+} -dependent RNA aptamer selected to bind RiboFlavinMonoNucleotide^[81].

The 1FMN aptamer needs Mg^{2+} to bind riboFMN with a K_D of $0.5 \mu M$ ^[81], notably 4 mM Mg was needed during structure elucidation by NMR^[35a]. We decided to test this in MS-friendly electrolytes (TMAA and NH_4OAc) by ITC and native MS on 1FMN aptamer. In our lab a previous thesis showed that TMAA can interfere with some RNA-RNA contacts, such as kissing loop interactions^[98]. Given that 1FMN does not involve any particular RNA-RNA contact in binding, the use of TMAA should may however not interfere with ligand binding to the aptamer.

4.3.1 Does RiboFMN aptamer (1FMN) need Magnesium?

No significant binding is detected by ITC of $10 \mu M$ RNA by the ligand in 100 mM NH_4OAc (Figure 38): After ligand dilution heat subtraction, the thermogram appears close to instrument detection limit ($0.02 \mu cal/sec$). However, some differences between injection points are visible, last points are partially above the baseline. This trend could mean either very low binding enthalpy or no binding.

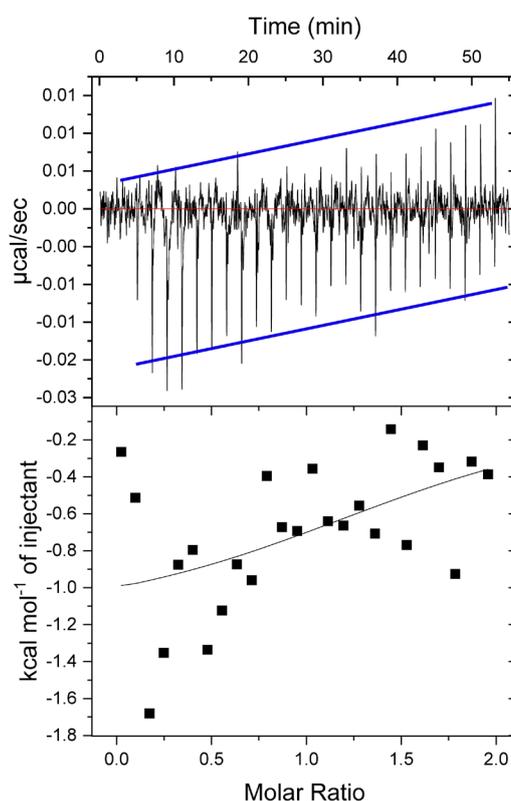


Figure 38 - ITC thermogram of 1FMN $10 \mu M$ by riboFMN in NH_4OAc 100 mM at $25 \text{ }^\circ C$. After ligand dilution heat subtraction, poor information is associated to heat pulses of each injection. The trend highlighted with the two blue parallel lines could be a low heat binding event either just noise associated to the injections (Table 48)

4.3.2 ITC of 1FMN in MgCl₂ 200 μM and NH₄OAc 100 mM or TMAA 100 mM

The thermograms changed completely once we carried out the ITC run in presence of 200 μM MgCl₂ in both NH₄OAc and TMAA 100 mM (Figure 39): we observe ligand binding in both electrolytes.

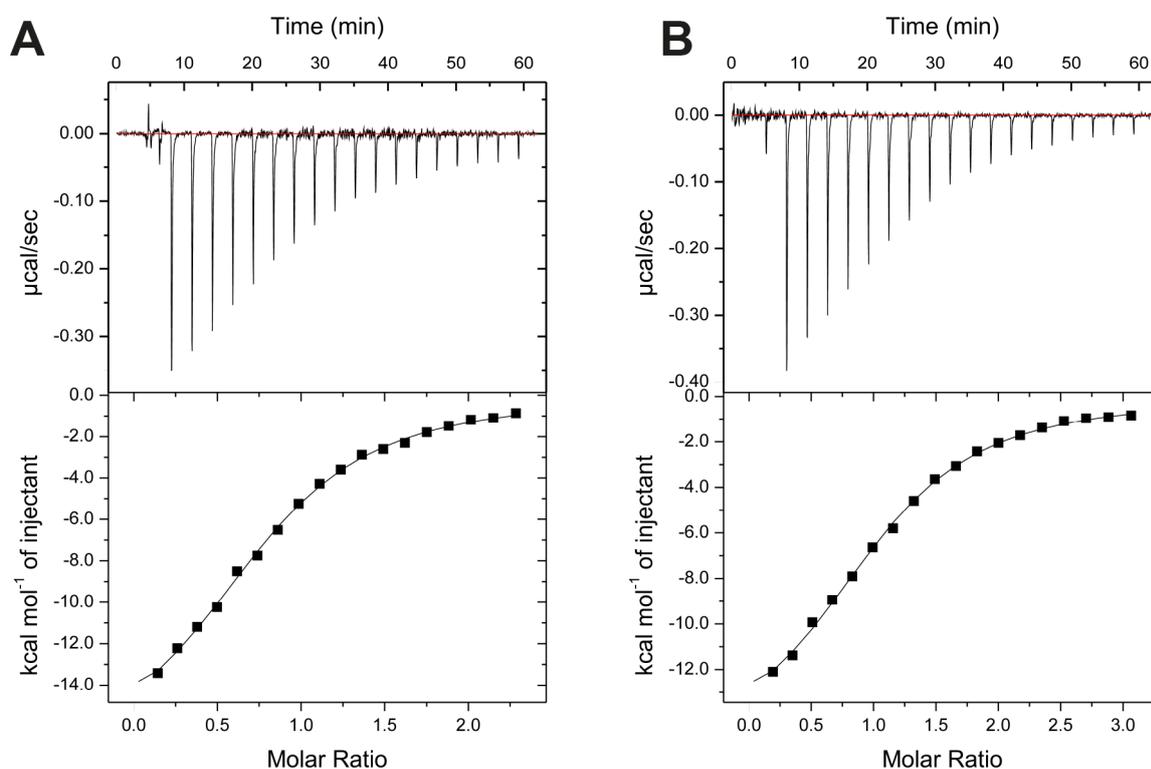


Figure 39 – ITC in 200 μM MgCl₂ and 25°C of **A.** 1FMN 10.19 μM and RiboFMN 150.0 μM in TMAA 100 mM **B.** 1FMN 10.20 μM and RiboFMN 159.6 μM in NH₄OAc 100 mM. Replica are at p.177.

On RNA aptamer 1FMN, no ITC data are available to our knowledge. From the fitted parameters in Table 15, we find a $\Delta H^\circ < 0$ and a derived $-T\Delta S^\circ > 0$ in both conditions. This thermodynamic profile is compatible with an intercalating binding mode^[50b], in agreement with the reported NMR structure^[35a]. The obtained K_D s are within the respective uncertainty range. We note that ΔH° and $-T\Delta S^\circ$ differs of 2.2 Kcal/mol between the two conditions.

Table 15 - ITC fit output of 1FMN vs RiboFMN in 200 μM MgCl₂. Uncertainties are the deviation standard across four replica in NH₄OAc and three replica in TMAA (Figure 110 and Figure 111 at p. 181).

Electrolyte	K_D (μM)	N	ΔH° (kcal/mol)	$-(298)\Delta S^\circ$ (kcal/mol)	ΔG° (kcal/mol)
NH ₄ OAc 100 mM, MgCl ₂ 200 μM	3.28 ± 0.51	0.97 ± 0.07	-16.8 ± 1.0	9.3 ± 1.1	-7.51 ± 0.10
TMAA 100 mM, MgCl ₂ 200 μM	3.36 ± 0.24	1.00 ± 0.03	-19.1 ± 0.5	11.6 ± 0.6	-7.47 ± 0.04

4.3.3 1FMN native MS

We decided to analyse an equimolar mixture of 1FMN (M) and RiboFMN (ML) at 10 μM in in NH_4OAc 100 mM by native MS. 1FMN is capable to bind up to 2 riboFMN in absence of Mg^{2+} (Figure 40).

In order to determine the K_D in absence of Mg^{2+} , we performed a MS titration with up to 3 equivalents of RiboFMN (in 5 replicas) using Agilent 6560 to monitor the IM-MS spectra of 1FMN and its complex with riboFMN. Only charge state 6- has been considered for quantification. The second complex with two ligands (ML_2) has too low intensity and extracting its relative response factor is not possible.

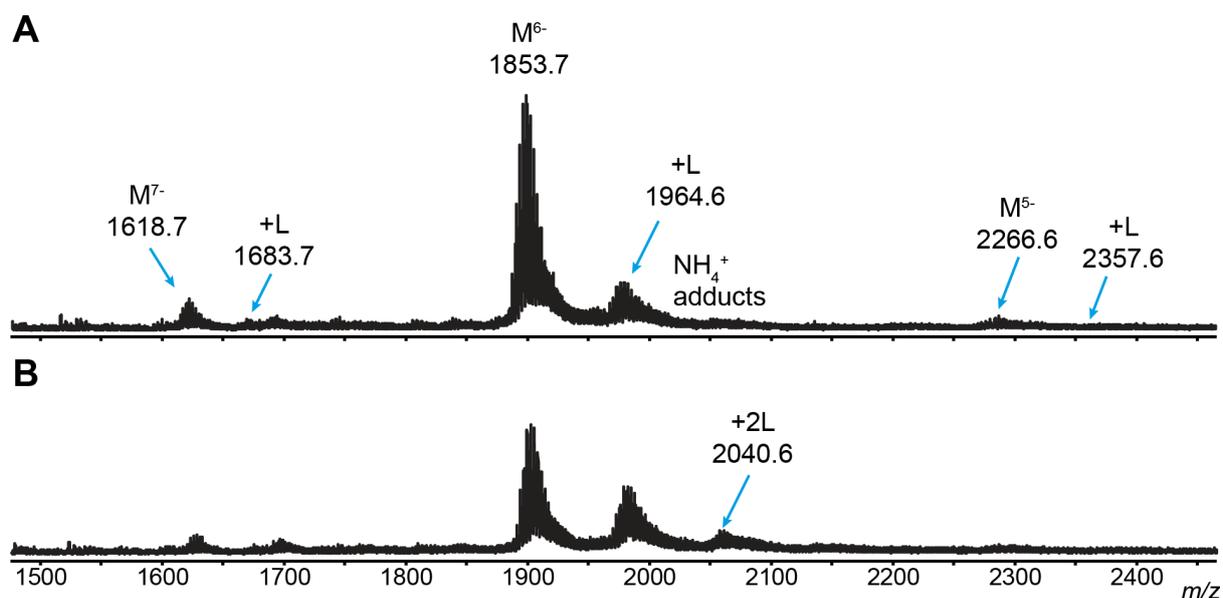


Figure 40 – **A.** Equimolar 10 μM mixture of 1FMN (M) and RiboFMN in NH_4OAc 100 mM. Charge states 7-, 6- and 5-, and relative 1:1 complexes are detected. **B.** A second complex ML_2^{6-} is detected in presence of 30 μM RiboFMN.

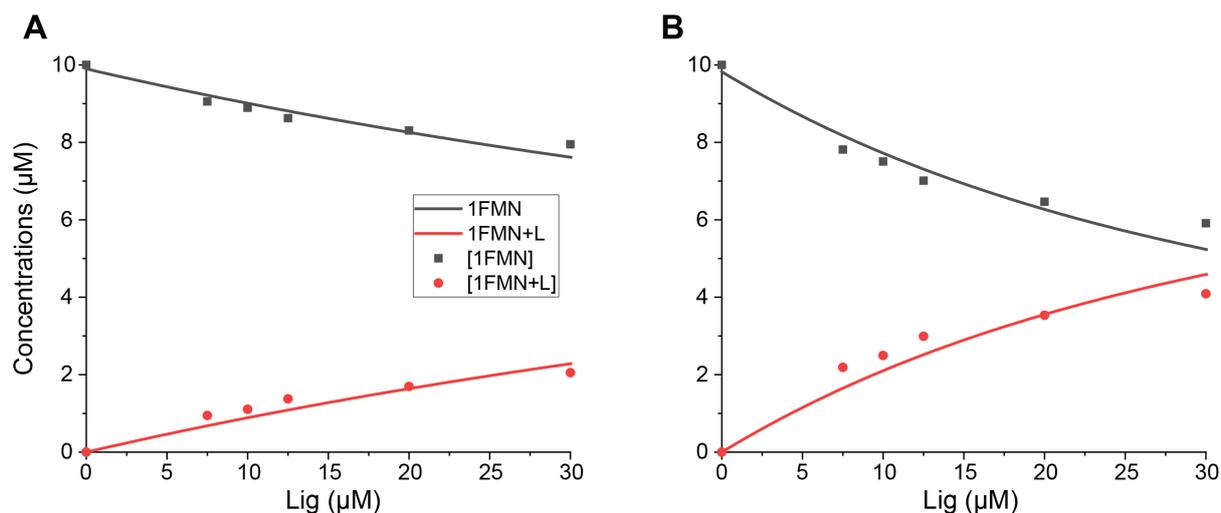


Figure 41 – **A.** Binding isotherm for 1FMN and its complex with RiboFMN in NH_4OAc 100 mM with relative response factor correction and **B.** assuming response factor ratio equal to 1. Points are averaged over 5 replicas (Figure 103 - p.171).

MS titration on 1FMN The K_D were determined per each replica following the method 2 at p.43: on the concentrations corrected for relative response factors. Then each K_D was averaged. The procedure was repeated assuming response factor ratio equal to 1. From the MS titrations of 1FMN^{6-} and $1\text{FMN}+L^{6-}$ we calculated the relative response factors ($r_{Std}/r_{1\text{FMN}}$) of 0.62 ± 0.03 and ($r_{Std}/r_{1\text{FMN}+L}$) 0.22 ± 0.08 for which we have a ratio of 3.21 ± 0.98 ($r_{1\text{FMN}+L}/r_{1\text{FMN}}$) (Table 37 – p.172).

The fittings on the corrected concentrations provides an average K_D of $110.8 \pm 40.7 \mu\text{M}$ (Figure 41A), whereas the same data sets, assuming relative response factor equal to 1 in Figure 41B, provides an average K_D of $27.6 \pm 2.4 \mu\text{M}$ (Table 34 and Table 35- p.170). The uncertainty on $r_{Std}/r_{1\text{FMN}+L}$ is very large (38% - Table 36 p.170). This is probably due to the low signal-to-noise ratio of the MS peak. This causes a large uncertainty on the corrected concentrations, thus on the fitted K_D s. Therefore we think that the assumption of relative response factor for these titrations is a better estimate of 1FMN and 1FMN+L concentrations and K_D .

However, we can affirm that riboFMN binds weakly to 1FMN in absence of Mg^{2+} , at least 50-fold lower to what reported in presence of magnesium ($\sim 0.5 \mu\text{M}$)^[81].

We then tested various concentrations of MgCl_2 to estimate how magnesium favours RiboFMN binding. We limited to $800 \mu\text{M}$ MgCl_2 , for two reasons: bioavailability of Mg^{2+} in the cytosol is normally not above 1mM ^[130]; Mg^{2+} concentrations from 1mM would lead to drastic signal suppression by Mg^{2+} clusters^[7]. In the MS spectra in Figure 42, the complex/free-apramer ratio is increasing with the Mg^{2+} concentration and a different adduct distribution for $1\text{FMN}+L^{6-}$ is observed: the desolvated signal of riboFMN complex disappears and the Mg^{2+} adducts distribution starts from 1 adduct. Conversely, adduct-free 1FMN^{6-} remains visible up to $400 \mu\text{M}$ Mg^{2+} . This different adducts distribution suggest that Mg^{2+} takes part in the formation of the complex.

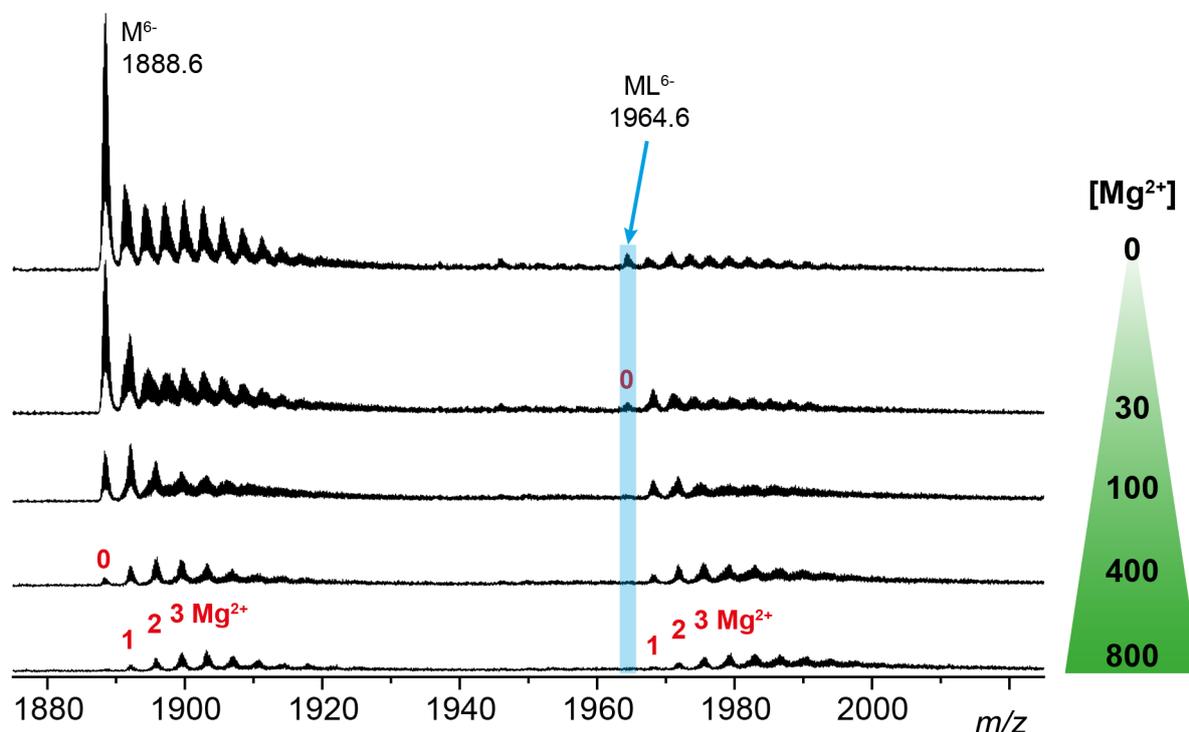
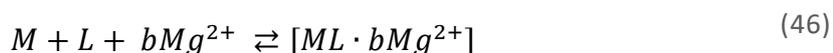


Figure 42 - MS spectra acquired on Agilent 6560 of 1FMN+riboFMN (5+5 μM) from 0 to 800 μM Mg^{2+} in NH_4OAc 100 mM.

The adducts distribution (i.e.: Mg^{2+} , NH_4^+) on M^{6-} and ML^{6-} looks similar from 100 to 800 μM MgCl_2 . We calculated on this Mg^{2+} concentration range the relative response factors of the peaks, with the entire adduct tails, obtaining ($r_{\text{Std}}/r_{1\text{FMN}}$) of **0.30** and ($r_{\text{Std}}/r_{1\text{FMN}+\text{L}}$) **0.37** for which we have a ratio of **0.82** ($r_{1\text{FMN}+\text{L}}/r_{1\text{FMN}}$) (Table 37 – p. 172). Interestingly, we found a ratio of response factors close to one, while without Mg^{2+} we found a pronounced difference (3.21). This difference in ESI response would suggest a significant conformational difference between M and ML in absence of Mg^{2+} .

The signal was integrated over the complete visible m/z -range of each ion, as for the MS titration in absence of MgCl_2 . Point-by-point K_D s are reported in Table 16.

The calculated K_D s as a function of MgCl_2 concentration allow to estimate the Mg^{2+} uptake stoichiometry in the riboFMN binding. We can write the following reaction where Mg^{2+} participates with unknown stoichiometry “ b ”:



If we assume that the point-by-point dissociation constant is independent from $[\text{Mg}^{2+}]^b$ (as if it is part of the milieu), we can separate the apparent dissociation constant K_D^{App} (where all forms of M and ML are considered) from $[\text{Mg}^{2+}]^b$:

$$K_D = \frac{[M][L][Mg^{2+}]^b}{[ML \cdot bMg^{2+}]} = K_D^{App} \cdot [Mg^{2+}]^b \quad (47)$$

and from its logarithmic form we have the following linear relation from which we can obtain Magnesium stoichiometry “*b*” by linear regression:

$$\log K_D^{app} = \log K_D - b \log[Mg^{2+}] \quad (48)$$

Table 16 - Estimated 1FMN+L K_D s for each $MgCl_2$ concentration in NH_4OAc 100 mM.

Mg^{2+} (μM)	K_D	$\log [Mg^{2+}]$	$\log K_D$
0	1.71E-05	--	-4.77
10	1.33E-05	1	-4.88
20	1.13E-05	1.30	-4.95
30	1.04E-05	1.48	-4.98
100	4.77E-06	2.00	-5.32
200	3.30E-06	2.30	-5.48
300	2.71E-06	2.48	-5.57
400	2.44E-06	2.60	-5.61
800	1.87E-06	2.90	-5.73

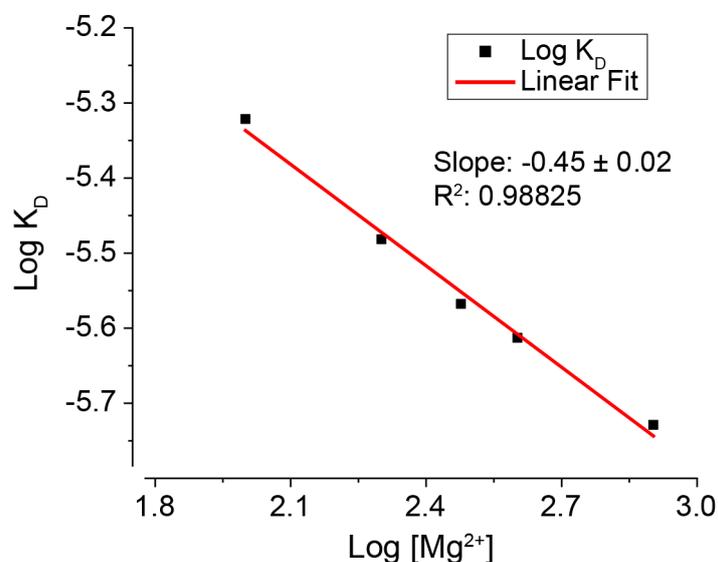


Figure 43 - Effect of Mg^{2+} concentration on 1FMN+L K_D^{App} , expressed in logarithmic scale. From the slope modulus of linear regression we have a stoichiometry of 0.45 ± 0.02 for Mg^{2+} (only points from Mg^{2+} 100 μM were considered).

From the linear regression in Figure 43 we obtain a Mg^{2+} stoichiometry of 0.45 ± 0.02 . A $b < 1$ suggests that Mg^{2+} cations are acting like “spectator” ions, diffused around the reactants but not influencing the binding reaction. This supposition is coherent with the weak binding detected in Mg^{2+} absence (Figure 40) but in contrast with the ITC experiments, where a significative heat of reaction is detected when Mg^{2+} is part of the medium. The linear

regression with the assumption of $(r_{1FMN+L}/r_{1FMN})=1$ give a similar stoichiometry (0.47 - Figure 104, p.172). On this result is difficult to speculate if magnesium cations favour the folding of a given conformation for 1FMN aptamer, acting as diffused ions (indirectly involved in binding) or if they binds specifically to the aptamer.

To have a better estimate of the K_D of 1FMN+L, we perform a full MS titration by ligand riboFMN with 200 μ M $MgCl_2$, because we expect a K_D in the same range of RNA concentration ($\sim 3 \mu$ M - Table 16) and acceptable signal intensities.

4.3.4 MS titrations of 1FMN in presence of magnesium

We did the MS titrations on the Thermo Exactive mass spectrometer (no need for IM-MS) in both NH_4OAc and TMAA to compare our results to ITC. To have an appreciable signal for 1FMN and its complex, we applied 15 eV HCD in the collision cell to remove TMA^+ or NH_4^+ as neutrals and reveal the Mg^{2+} adduct distribution. In Figure 44, 1FMN⁶⁻ and 1FMN+L⁶⁻ show different adduct distributions when we change the supporting electrolyte. In the spectra in Figure 44, 1FMN⁶⁻ (M^{6-}) carries two Mg^{2+} adducts minimum in TMAA whereas in NH_4OAc the non-adducted peak is visible. 1FMN+L⁶⁻ (ML^{6-}) in TMAA has a distribution starting from 3 Mg^{2+} ions while in NH_4OAc the adducts distribution start from 1 Mg^{2+} .

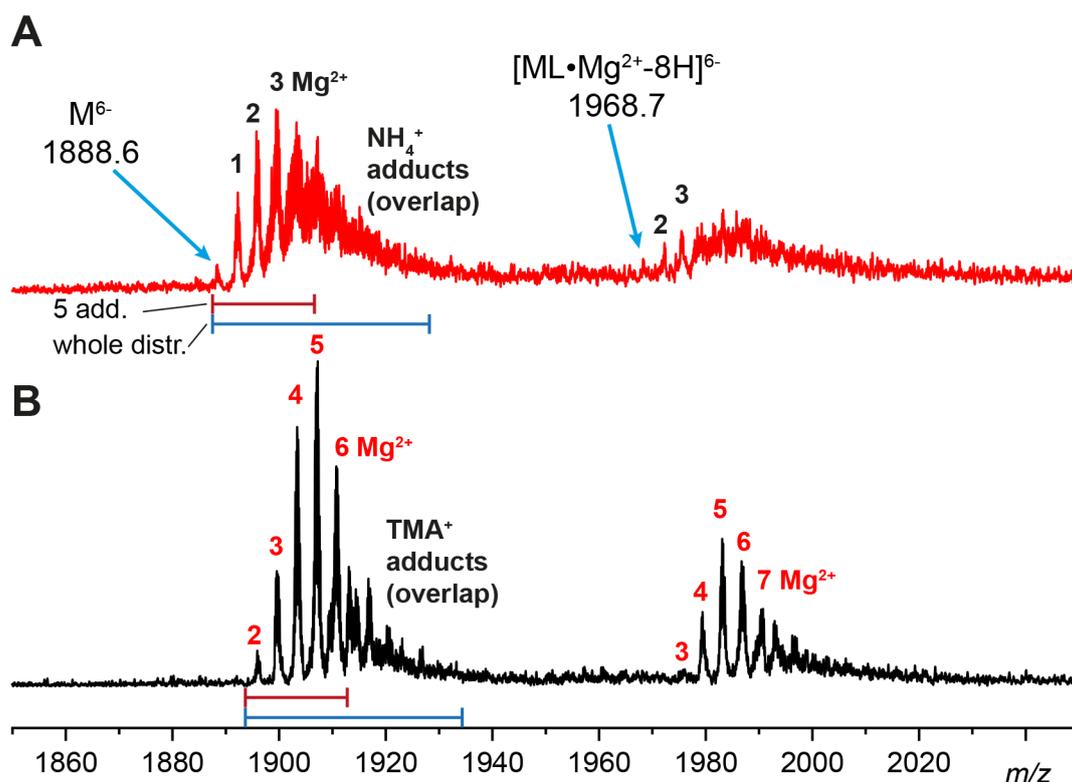


Figure 44 - 1FMN 10 μ M + 0.5 eq RiboFMN in **A.** 200 μ M $MgCl_2$ and NH_4OAc 100 mM and **B.** 200 μ M $MgCl_2$ and TMAA 100 mM buffer. Integration ranges: up to 5 adducts (red bar), whole adducts distribution (blue bar).

Comparing the two electrolytes, we notice two effects:

- i. TMAA facilitates the uptake of Mg^{2+} on both aptamer and complex, adding two Mg^{2+} on their adducts distribution;
- ii. riboFMN complex have one Mg^{2+} more compared to unbound aptamer;
- iii. The binding affinity seems higher in TMAA than in NH_4OAc .

We performed MS titrations up to 50 μM of riboFMN (Lig) in both buffers, considering the mass ranges for the first 5 adducts (red bar in Figure 44) of Mg^+ for peak integration (Figure 45 and Figure 47). Regarding data analysis, the choice of n° adducts (5 adducts or the whole distribution) does not affect the K_D for TMAA while for NH_4OAc a worse fit is observed (Figure 102 - p.169). This is the consequence of the increasing background along the titration leading to difficulties to the appropriate estimation of the background.

Table 17 - Averaged K_D over 3 replicas (**Method 1**). Apparent K_D for $1FMN^{6-}$ is calculated from fitting of binding isotherm of corrected concentrations for $1FMN$ and $1FMN+L$ and assuming equal relative response factors; †. K_D uncertainty are the standard deviation over 2 replicas for TMAA and 3 replicas for NH_4OAc (a third replica has been discarded) (Table 44).

Electrolyte	K_D apparent (μM) †			
	Correction		r/r=1	
TMAA 100 mM MgCl ₂ 200 μM	2.43	± 0.12	2.31	± 0.24
NH_4OAc 100 mM MgCl ₂ 200 μM	3.08	± 1.44	5.36	± 0.82

Table 18 - Relative response factors for $1FMN^{6-}$ and $1FMN+L^{6-}$ considering 5 adducts. α . Uncertainty on the relative response factors are the standard deviation over 2 replicas for TMAA (a third replica has been discarded) and 3 replicas for NH_4OAc (Table 39 - p.173).

Electrolyte	$r_{Std}/r_{1FMN} \pm \sigma^\alpha$	$r_{Std}/r_{1FMN+L} \pm \sigma^\alpha$	$r_{1FMN+L}/r_{1FMN} \pm \sigma^\alpha$
TMAA 100 mM MgCl ₂ 200 μM	7.49 ± 0.27	6.83 ± 0.93	1.11 ± 1.40
NH_4OAc 100 mM MgCl ₂ 200 μM	4.04 ± 2.35	3.15 ± 2.77	0.69 ± 1.09

We observe also that the ratio of relative response factors is nearly 1 in TMAA. Conversely, is 0.69 ± 1.09 in NH_4OAc : the complex $1FMN+L^{6-}$ has a lower response compared to the unbound aptamer. Not fully removed NH_4^+ might reduce the ESI response of the ions, as opposed to TMA^+ that get completely removed.

Comparing the K_D s issued from ITC (Table 19) with MS titrations ones, we found that K_D s (and derived ΔG° -Table 19 and tables from p.173) superimposes within their confidence interval of expanded uncertainty (with 95% confidence - Figure 46) Similarly to what obtained for MGA, we can say that ITC and MS dissociation constants are comparable.

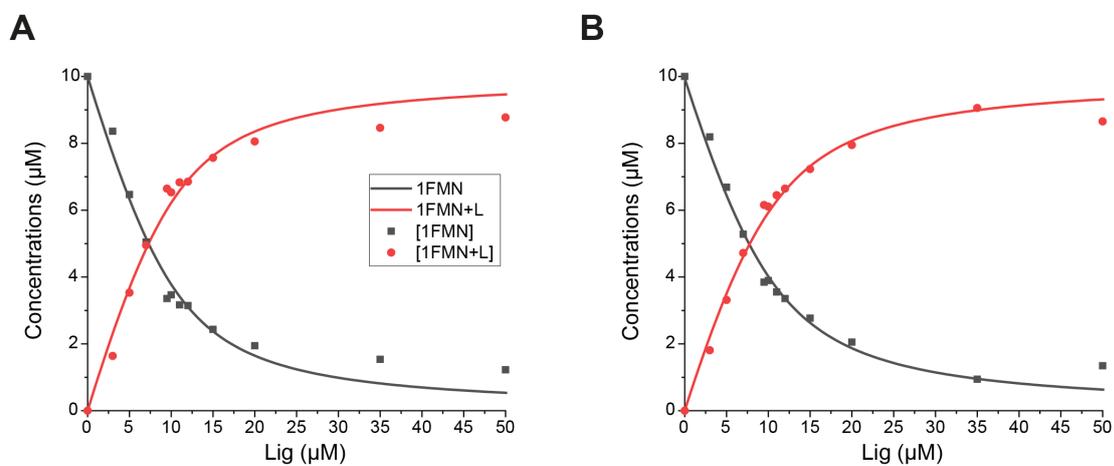


Figure 45 - Binding isotherm considering up to 5 Mg^{2+} adducts for $1FMN^{6-}$ and its complex with RiboFMN in $200 \mu M Mg^{2+}$ in **A.** TMAA and **B.** NH_4OAc 100 mM. Titrations are acquired in duplicate (TMAA) And triplicate (NH_4OAc).

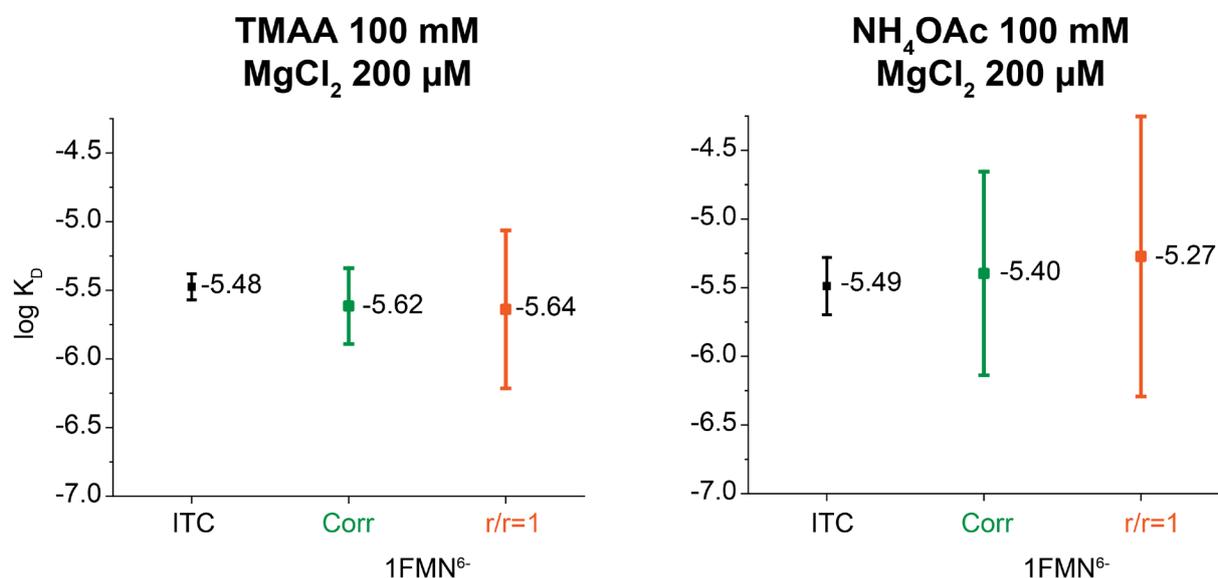


Figure 46 -Comparison between log K_D determined via ITC and MS, on 1FMN⁶⁻. The extended uncertainties (for a 95% confidence) are reported as error bars. MS values “Corr” are calculated with relative response correction, “r/r=1” assuming equal relative response factors. MS titrations were acquired in duplicate (coverage factor = 12.706); ITC (TMAA) in triplicate (coverage factor = 4.303) and ITC (NH₄OAc) in quadruplicate (coverage factor = 3.182). Single replicas are plotted in Figure 105 at p. 175.

Table 19 - K_D and ΔG° obtained from ITC on 1FMN with OneBindingSite fitting model. “ σ ” is the standard deviation and “Exp. σ ” is the expanded uncertainty. Experiments were done in 4 replicas in NH₄OAc (coverage factor = 3.182 for a 95% confidence interval) and 3 replicas in TMAA (coverage factor = 4.303 for a 95% confidence interval) (p.181).

Electrolyte	K_D (μM)	σ	Exp σ	ΔG° (kcal/mol)	σ	Exp σ
TMAA 100 mM MgCl ₂ 200 μM	3.28	± 0.51	± 1.61	-7.51	± 0.10	± 0.32
NH ₄ OAc 100 mM MgCl ₂ 200 μM	3.36	± 0.24	± 1.02	-7.47	± 0.04	± 0.17

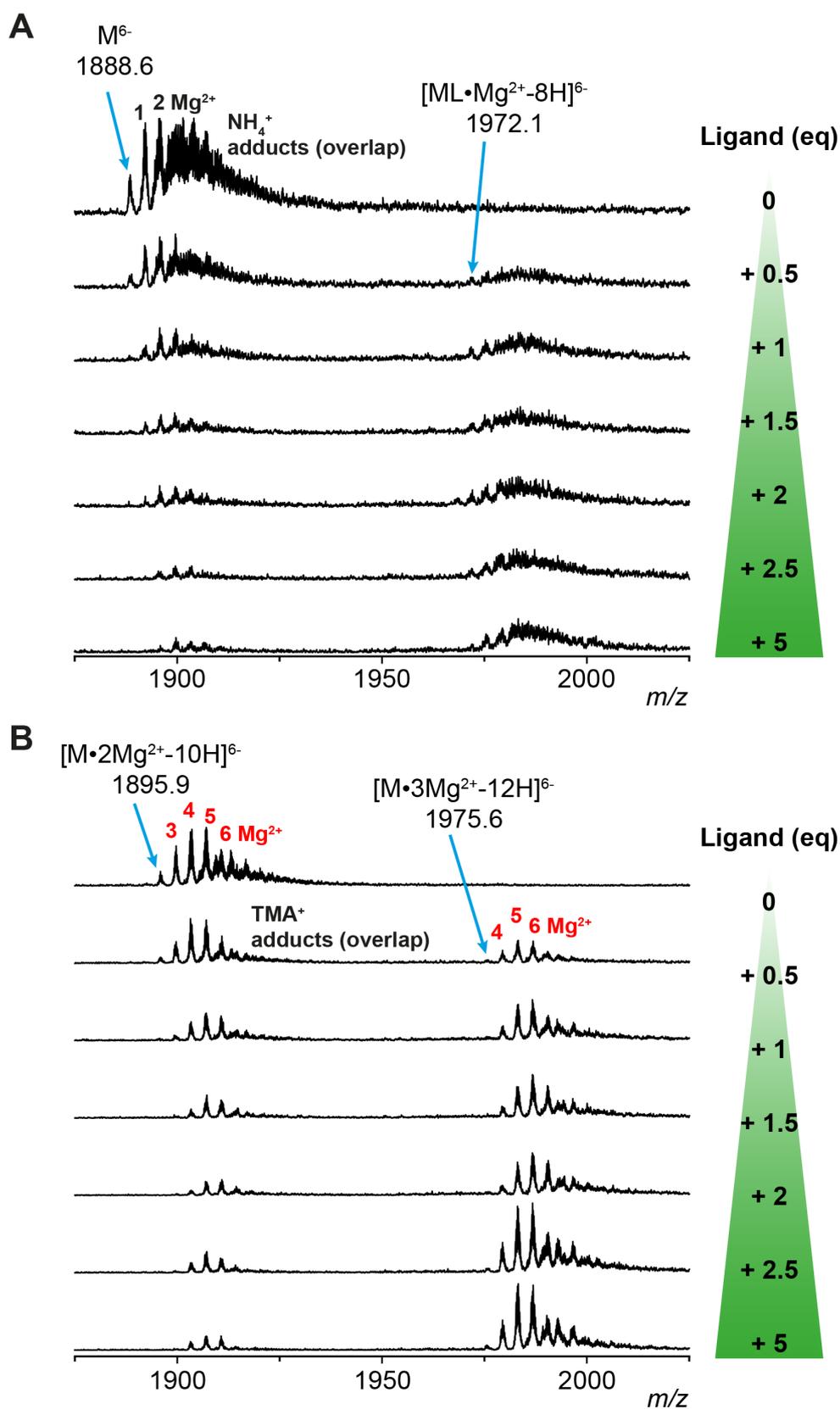


Figure 47 – NativeMS titrations of 1FMN 10 μM up to 5 eq of RiboFMN in 200 μM Mg^{2+} and **A.** TMAA and **B.** NH_4OAc 100 mM.

4.4.1 Neomycin RNA aptamer (1NEM)

For the Neomycin RNA aptamer (1NEM) a K_D of 100 nM is reported (in 5 mM Mg^{2+} ; 50 mM Tris•HCl pH 7.6; 250 mM NaCl) [84]. When dissociation constants are particularly low, like in the case of neomycin to 1NEM, is preferable to have a $[RNA]_{tot} \approx K_D$ to have an informative titration curve (with a significant non-linear portion)[124]. Thus, we decided to titrate at 50 nM RNA, on an expected K_D of 100 nM.

When sprayed 1NEM at 50 nM in NH_4OAc 100 mM, the MS spectrum shows two charge states, a dominant 5- and a low intensity 4- (Figure 49). We decided to consider the charge state 5- for quantification, because it provides the best signal intensity.

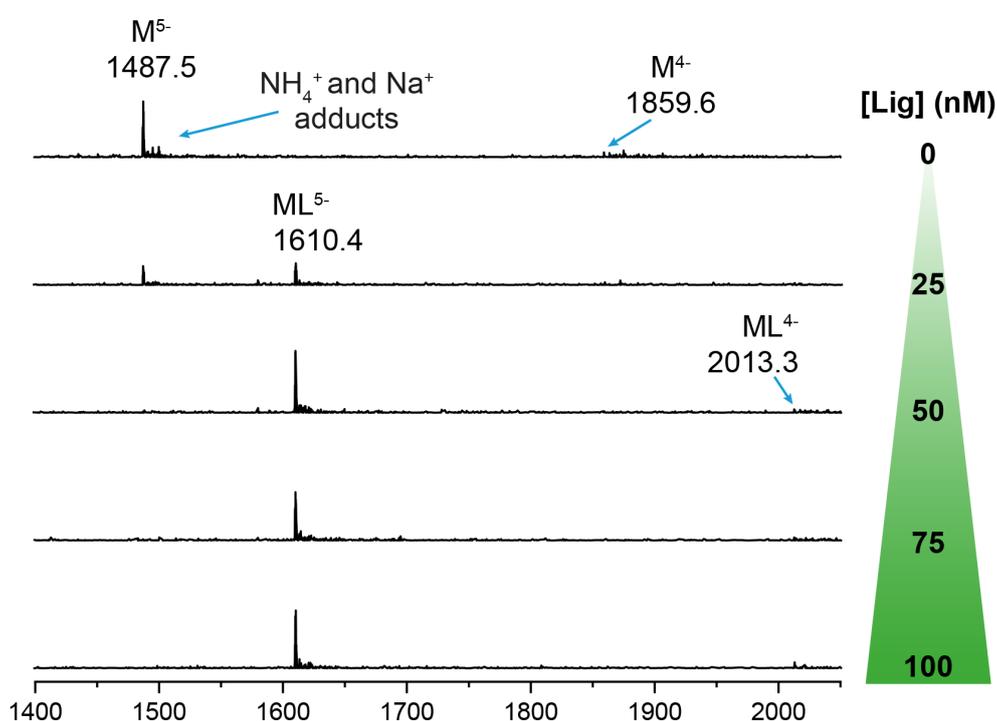


Figure 49 – Neomycin RNA aptamer 1NEM (RNA 50 nM, dT_6 50 nM) in NH_4OAc 100 mM show the charge states 5- and 4-. Only 5- is used for response factor determination and quantification.

From our titrations, the neomycin is more tightly bound compared to what expected: the binding isotherm for first complex at 50 nM RNA (Figure 50) show the saturation of the aptamer with an estimated K_D of 0.13 nM. Unfortunately below 50 nM RNA no signal is detectable and for this aptamer we can conclude only that neomycin complex has a $K_D \ll 100$ nM. As we can see in Figure 50, we do not have any non-linear portion of the binding curve near the equivalence point. Such situation is the consequence of a $[RNA]_{tot} \gg K_D$, for which K_D estimation becomes too uncertain to be trustable[124]. A second titration at 0.5 μM was done only for response factor determination (Figure 113 and Figure 114).

However, we can calculate the relative response factors for the aptamer and its complex. For M we have 0.84 ± 0.05 and ML we have 0.68 ± 0.04 (over four replica). Their ratio R_M/R_{ML} is

1.23 (over four replica). This means that neomycin complex has response (a little) higher compared to the free-aptamer. Even if neomycin binding mode is not an intercalator, we found a similar ratio to MGA and 1FMN aptamers.

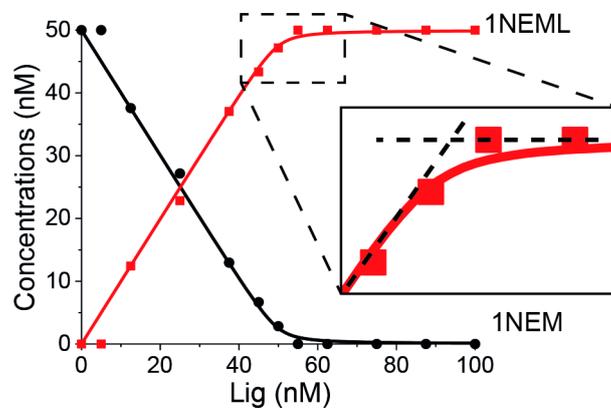


Figure 50 - 1NEM aptamer titrated at 50 nM in NH_4OAc 100 mM . The non-linear portion of the binding isotherm is almost absent, suggesting the saturation of the RNA by the neomycin at this concentration.

4.4.2 Tobramycin RNA aptamer (1TOB)

For the tobramycin RNA aptamer (1TOB), two complexes are reported, the first with a K_{D1} of 9 nM, the second with K_{D2} 2.7 μM [83b]. Both complexes are detectable by native MS (Figure 51). 1TOB shows a pronounced difference in aptamer/complex ratio between the charge states 6- and 5-, which suggest a marked difference in response factor ratios.

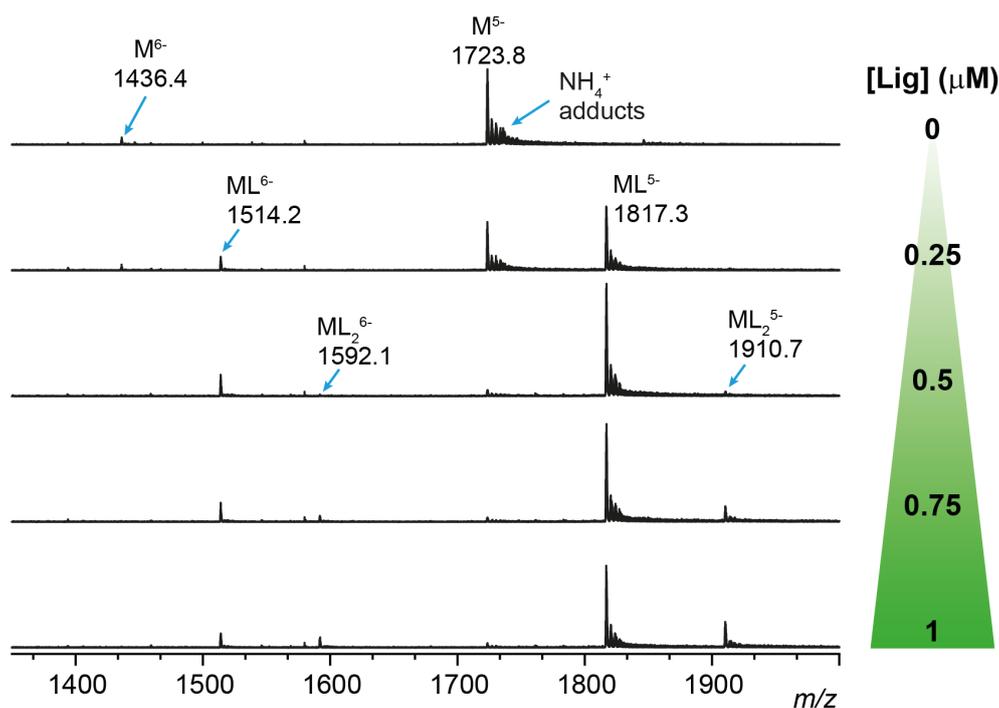


Figure 51 - MS titration of 1TOB 0.5 μM (dT_6 0.5 μM) with tobramycin up to 1 μM in NH_4OAc 100 mM. Charge states 6- and 5- are visible.

From the titration up to 2 equivalents of tobramycin (1 μM) we obtain fairly different relative response factors between the two charge states, in line with the expectations (Table 20). The correction is minimal for 5-, between 0.45 and 1.10, while more pronounced for 6-. Here we deduce that both complexes ML and ML_2 have an higher response, because the required correction is lower compared to M (eq. (31)– p. 39; $R_M = \frac{r_{Std}}{r_M}$).

From binding isotherms at 0.5 μM RNA (Figure 52) we obtain $K_{D1} = 7.8 \pm 1.8 \text{ nM}$; $K_{D2} = 0.25 \pm 0.09 \mu\text{M}$ (charge state 6-) and $K_{D1} = 11.1 \pm 3.4 \text{ nM}$; $K_{D2} = 0.18 \pm 0.07 \mu\text{M}$ (charge state 5-), over two replicas. Interestingly, we found a K_{D1} similar to what reported in literature and a 10-fold discrepancy on K_{D2} . Unfortunately, this aptamer was no more detectable for $[\text{RNA}] < 0.5 \mu\text{M}$, so we cannot further dilute $[\text{1TOB}]_{\text{tot}}$ to improve titration points.

Table 20 - Relative response factors for 1TOB and its tobramycin complexes. α . the uncertainty is the standard deviation over two replicas.

6-			5-		
r_{Std}/r_M	r_{Std}/r_{ML}	r_{Std}/r_{ML2}	r_{Std}/r_M	r_{Std}/r_{ML}	r_{Std}/r_{ML2}
15.01 ± 0.34^a	3.63 ± 0.45^a	5.11 ± 1.22^a	0.65 ± 0.06^a	0.45 ± 0.01^a	1.10 ± 0.24^a

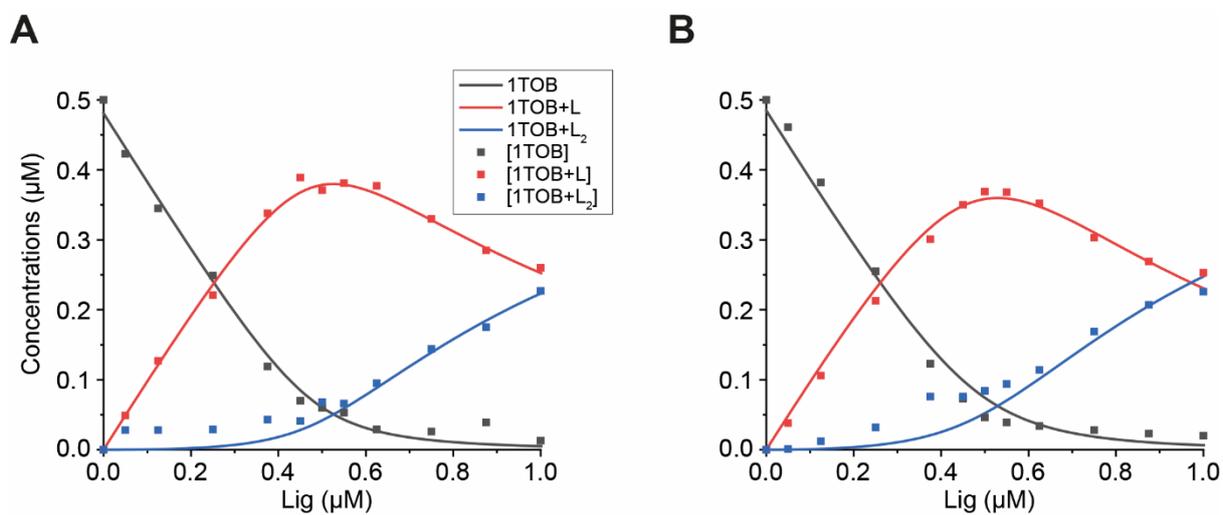


Figure 52 - Binding isotherms of 1TOB **A.** ch st 6- and **B.** ch st 5-. 1TOB aptamer was titrated at 0.5 μM up to 1 μM of tobramycin in NH₄OAc 100 mM.

4.5 Conclusions

Typical native MS electrolytes, NH_4OAc and TMAA, are suitable also for ITC studies. The thermodynamic profile for MGA was in line with the literature in various non-volatile buffers [80].

K_D values determined by MS and ITC were comparable for both MGA and 1FMN (same order of magnitude and values within the expanded uncertainty for a 95% confidence interval). These two RNA have a 1:1 stoichiometry exclusively, and ITC can benefit of the unambiguous binding stoichiometry (from native MS) for the choice of the fitting model. In this sense, native MS can help to unravel multiple stoichiometries or binding events with very low reaction heats in ITC (Figure 38).

The use of TMAA on MGA and 1FMN didn't bring any significant drawback in comparison with NH_4OAc [98]. On the contrary, TMAA helps maximizing the signal-to-noise ratio (on both MGA and 1FMN) because of a lower number of TMA^+ adducts and more spacing between them compared to NH_4^+ . In TMAA, the 1FMN binding activity is preserved and we revealed an interesting Mg^{2+} distribution for 1FMN, higher than in NH_4OAc . We think that NH_4^+ could compete with other cations (in solution) for RNA because it has a similar ionic radius to alkali metals [123]. TMA^+ is too bulky to compete or simply shield the aptamer from other cations, such as Mg^{2+} , so in that case a higher number of "small" cation binding sites can be filled by Mg^{2+} [135]. Although we tested the buffers for only two RNA aptamers, we think that TMAA could be an useful volatile electrolyte to study cations interactions on RNA and DNA.

From our preliminary results, Mg^{2+} showed a stoichiometry of $\frac{1}{2}$ (0.47) that would suggest a role of diffused ions in the ligand binding. But the riboFMN complex appeared to be always bound to at least 1 Mg^{2+} in both electrolytes and also ITC thermograms changed significantly in presence of MgCl_2 . A further verification about the presence of un-specific Mg^{2+} adducts could be addressed in two ways: by using a control sequence (e.g. of conveniently randomized sequence) [98] and by nano ESI using sub-micrometer emitters [136]. Also more titrations, in the K_D range of 100-1000 μM , are necessary to extrapolate the Mg^{2+} stoichiometry in riboFMN binding (on logarithmic scale $\log K_D/\log \text{Mg}^{2+}$).

We also recommend to choose the charge state(s) with higher signal-to-noise ratio for relative response factor correction and quantification, in order to minimize the uncertainty on the signals and finally on the results. The user could decide which electrolyte choose (and its concentration) depending on how the aptamer charge state distribution changes, to improve signal-to-noise ratio of minor charge states and the uncertainty related to their data sets (MGA⁷⁻). However, this is very analyte (RNA and DNA) dependent.

The relative response factors of free and bound aptamer were analysed. Even for the scenario where conformational changes were the most probable (1FMN in presence of Mg^{2+}), the response factor correction was relatively modest. The pronounced difference in relative response factors observed for 1FMN, in absence of Mg^{2+} , is probably related to the low signal-to-noise of the complex ($\sim 38\%$ on $r_{\text{std}}/r_{1\text{FMN}+\text{L}}$) and therefore high uncertainty. We think that

the ligand binding was not associated to a significant conformational change that would lead to a substantial variation of ESI-MS response between aptamers and their complexes. For these aptamers, we believe that the assumption of equal response factors (ratio=1) is acceptable. Besides, the property that matters most is the $\log(K_D)$ (related to ΔG°).

We quantified K_{DS} from 30 μM for 1FMN (Figure 41B) to 0.1 μM for 1TOB (Figure 52). Potentially K_{DS} about 50 nM would be accessible in our conditions, but the K_D of 1NEM go beyond our spectrometer detection limit (Figure 50). In general, for quantification it is recommended to have $[\text{RNA}]_{\text{tot}}$ at least in the same order of magnitude of dissociation constant, and spectrometer detection limit may be passed. That is the case of 1NEM, but we think that is very analyte (RNA) dependent.

In summary:

- 1- ITC and native MS give comparable K_D values
- 2- We preserved aptamer binding in a “non conventional” electrolyte such as TMAA, with the benefit of clear metal adduct stoichiometries on aptamer and its complex. However, the adducts distribution was different than in NH_4OAc , suggesting a different competition for the cation binding sites.
- 3- Relative response factor correction is not required for MS aptamer quantification, simplifying the use of this technique to non-experts.
- 4- The range of K_D values that can be quantified by native MS depends on the total signal response. This is aptamer-dependent. In our conditions, we could have quantified K_D values from 50 nM to 30 μM .



Part II: Ion Mobility Spectrometry

5. IM-MS of aptamers and their complexes

Our aim is to observe the aptamer conformers *via* their ion mobilities, for example to discriminate free and bound aptamer conformations based on their CCS distributions, and see if ion mobility spectrometry can detect “adaptive binding” mechanisms (p. 9). In this chapter we present the CCS distributions observed for the RNA aptamers studied above (MGA, 1FMN, 1NEM, 1TOB). We also study tetracycline RNA aptamer (a well known aptamer with a successful riboswitch implementation in yeast) and a group of cocaine binding DNA aptamer with a modulable secondary structure (by ligand binding and chain length).

RNA aptamers MGA, 1FMN, 1TOB and 1NEM were analyzed by IM-MS, in NH₄OAc and TMAA 100 mM, in order to differentiate free and complexed forms by their ion mobilities. Unfortunately, CCS distributions did not change significantly to infer a conformational change upon ligand or Mg²⁺ binding (for 1FMN). Upon collisional activation in pre-IMS with fragmentor at 600V, all aptamers and complexes undergo a slight compaction if present.

We observe a small increase in CCS distribution of each complex, compared to the relative unbound aptamers. Such small increase is probably due to a global increase of overall volume of the aptamer after ligand incorporation, rather than a significant shape change upon binding. All CCS values are listed in Table 21.

5.1 1TOB

1TOB shows CCS distributions of a single distribution for both aptamer and tobramycin complexes in charge states 5- and 6-. Aptamer CCS increases upon tobramycin binding, but this increase is limited (Table 21). We describe this by a volume increase consequent to ligand encapsulation (Figure 48). Pre-IMS activation induces a modest compaction effect both charge states for which CCS is substantially unchanged.

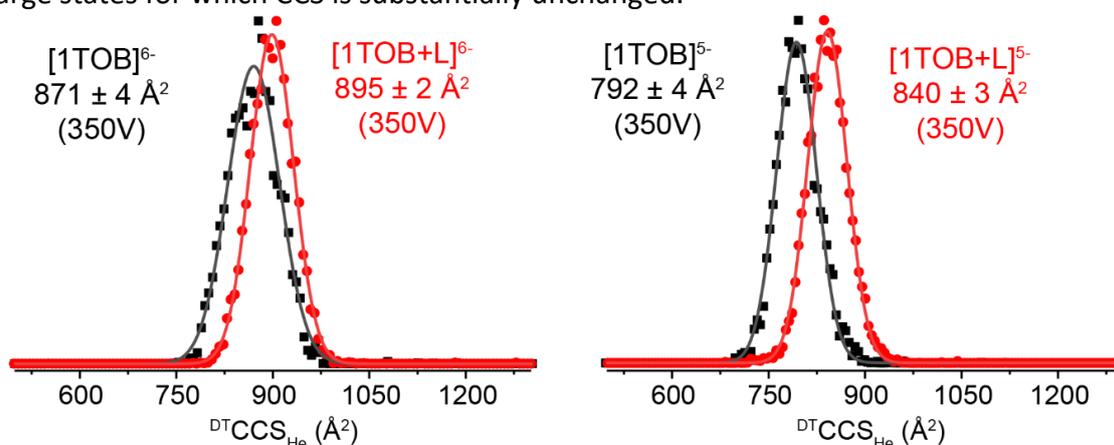


Figure 53 – CCS distributions of charge states 6- and 5- of 1TOB and its tobramycin complex at low pre-IMS activation (fragmentor 350 V).

5.2 1NEM

1NEM⁵⁻ in Figure 54 shows a little increase in CCS upon ligand binding, ascribable to the encapsulation of neomycin (see NMR structure in Figure 48). For the single 1NEM⁵⁻ and 1NEM+L⁵⁻ CCS distributions, a minor compaction is observed upon pre-IMS activation with fragmentor at 600 V.

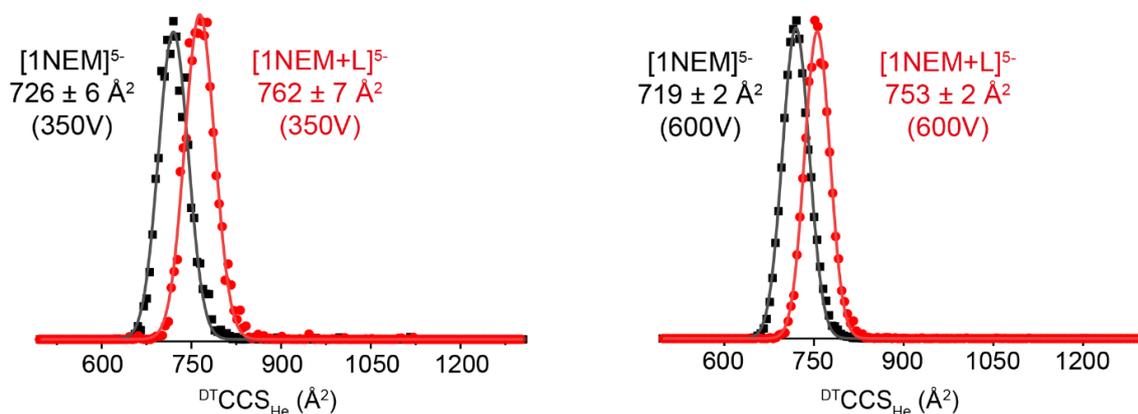


Figure 54 - CCS distributions for 1NEM⁵⁻ and its relative complex at low and high pre-IMS collisional activation (respectively fragmentor 350 and 600V). A minor CCS increase is observed upon ligand binding at both activation conditions.

5.3 MGA

At low activation, the large adduct distributions on MGA aptamer shades the declustered ion, so we have IMS data only at high activation (fragmentor 600V). CCS increase between free-aptamer and its complex remain modest for each charge state (Figure 55). On these results, we cannot differentiate free and bound aptamers solely on their ion mobilities.

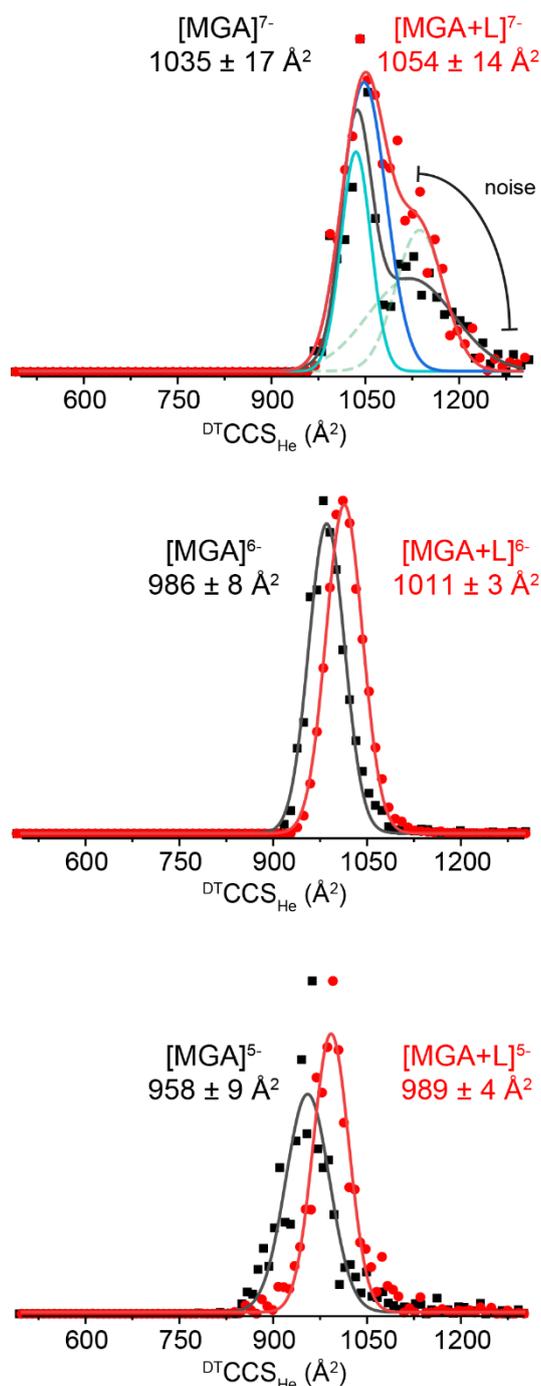


Figure 55 - CCS distributions (fragmentor 600 V) for charge states 7-, 6- and 5- of MGA and its complex with MG. Distributions remains almost unchanged by the ligand binding.

5.4 1FMN

For 1FMN aptamer ion mobilities were analyzed only for charge state 6-, because the low intensity of 7- and 5-. Similarly to 1TOB, for 1FMN⁶⁻. We observe a modest CCS increase upon binding, if present, and collisional activation induces a compaction on both 1FMN⁶⁻ and complex (Figure 56).

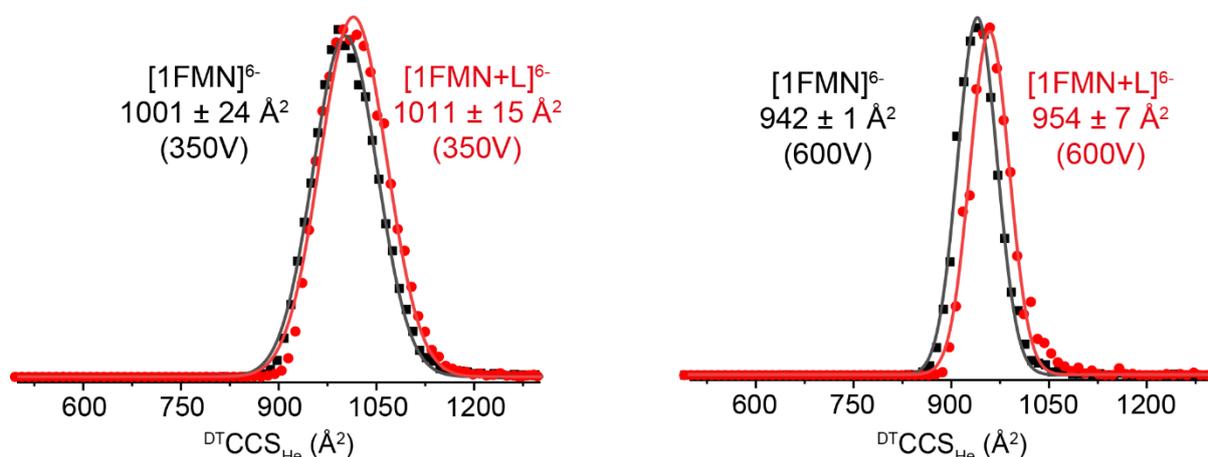


Figure 56 - CCS distributions for 1FMN^{6-} and relative complexes at low and high pre-IMS activation. 1FMN^{6-} CCS remains almost unchanged by the ligand binding.

1FMN^{6-} does not show a significant change in its CCS distribution when bound to magnesium (Table 21). At low activation, CCS distributions are broad and superimposed. At high activation (fragmentor 600V) the distributions undergo a little compaction and become sharper, but desolvated 1FMN become slightly more compact compared to Mg^{2+} adducts (Figure 57).

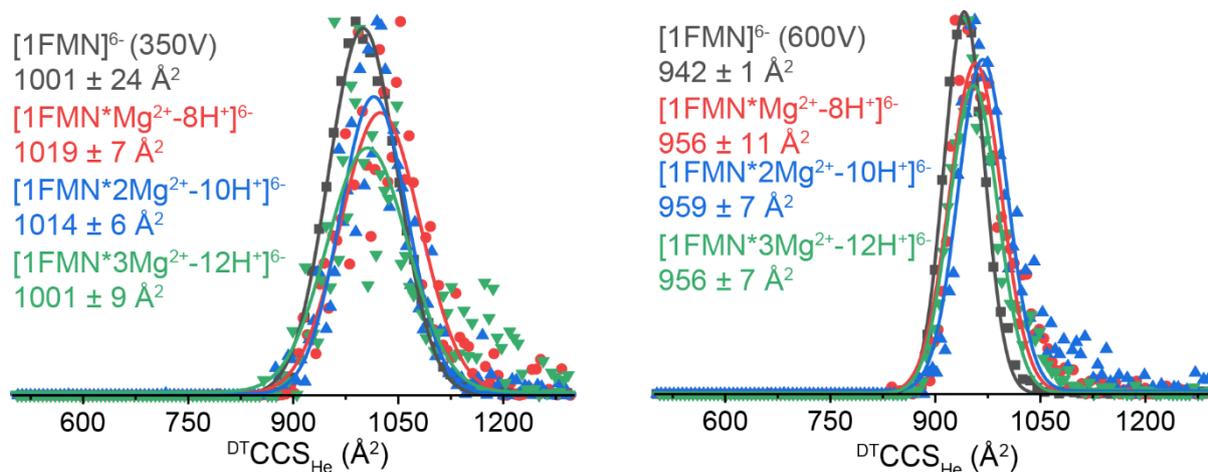


Figure 57 - CCS distributions for 1FMN^{6-} and its first three Mg^{2+} adducts at low activation (fragmentor 350 V) and high activation (fragmentor 600V).

The complex with riboFMN is visible up to $30 \mu\text{M}$ MgCl_2 . $1\text{FMN}+\text{L}^{6-}$ CCS distribution is lightly compact and narrower, compared to its first three Mg^{2+} adducts (Figure 58). We suggest that Mg^{2+} adducts stabilize a conformational population that does not compact at high pre-IMS activation, whereas desolvated $1\text{FMN}+\text{L}^{6-}$ compacts following a mechanism similar to what observed for B-DNA duplexes^[64]. A similar but smaller effect is observed for 1FMN^{6-} and its Mg^{2+} adducts in Figure 57.

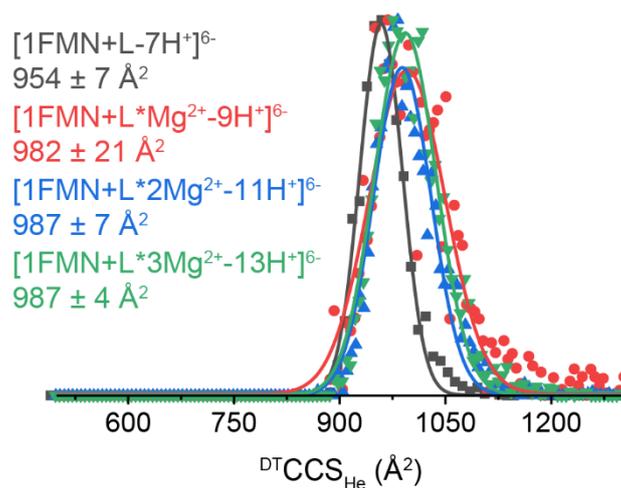


Figure 58 - CCS distributions for 1FMN+L⁶⁻ (acquired in absence of Mg²⁺) and its first three Mg²⁺ adducts at high pre-IMS activation (fragmentor 600V).

Table 21 - ^{DT}CCS_{He} of 1TOB, 1NEM, 1FMN and MGA aptamer and relative complexes. Uncertainty is the standard deviation over the replicas (Table 54 and Table 55- p.189).

Fragmentor	Ion	z	350 V		600 V	
			M	ML	M	ML
	[1TOB-zH+]z-	6	871 ± 4	895 ± 2	840 ± 3	864 ± 2
		5	792 ± 4	840 ± 3	786 ± 2	820 ± 1
	[1NEM-zH+]z-	5	726 ± 6	762 ± 7	719 ± 2	753 ± 2
	[MGA-zH+]z-	7			1035 ± 17	1054 ± 14
		6			986 ± 8	1011 ± 3
		5			958 ± 9	989 ± 4
	[1FMN-zH+]z-	6	1001 ± 24	1011 ± 15	942 ± 1	954 ± 7
		+1 Mg ²⁺	1019 ± 7	1023 ± 2	956 ± 11	980 ± 21
		+2 Mg ²⁺	1014 ± 6	1033 ± 11	959 ± 7	984 ± 7
		+3 Mg ²⁺	1001 ± 9	1029 ± 5	956 ± 7	985 ± 4

5.5 Tetracycline RNA aptamer

The RNA tetracycline aptamer (TCA) has been selected *in vitro* by Berens et al^[137] and is among the few artificial aptamers with riboswitch activity in yeast^[138]. The aptamer secondary structure is composed by a three-way junction of two hairpin-loop and the duplex at 3',5' end. Two single-stranded portions forms the junction (Figure 59). The very high binding affinity reported TCA ($K_D = 0.770$ nM in $[Mg^{2+}] = 10$ mM)^[139] is due to the formation of a loop-loop interaction between the loop (in cyan) and the junction (in purple) in Figure 59A. This interaction is mediated by the ligand, chelated to one Mg^{2+} cation, and by an unusual triple helix pattern around the circled adenines in Figure 59A.

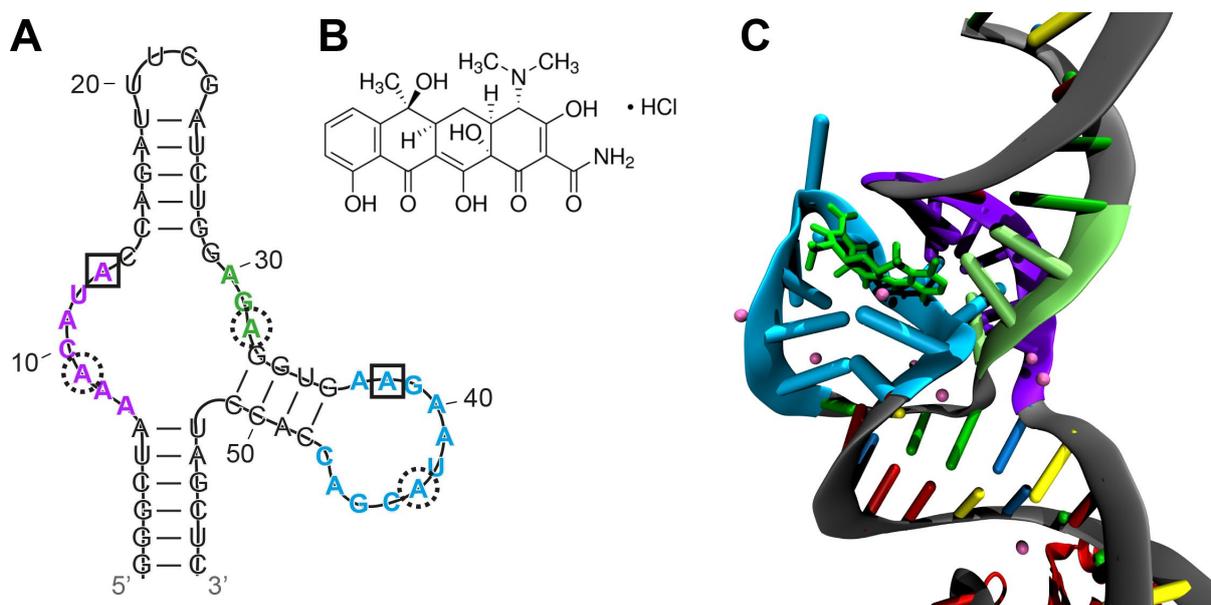


Figure 59 - **A.** RNA tetracycline aptamer sequence, **B.** tetracycline hydrochloride (ligand) structure and **C.** x-ray crystal structure of the complex with focus on tertiary contacts (from pdb 3EGZ)^[28]. Squared bases are directly involved in ligand binding. Circled bases are involved in the irregular triplex pattern. Pink spheres are Mg^{2+} cations. Bases color code: **A-Blue; G-Green; U-Yellow; C-Red.**

Ligand binding is possible only in presence of Mg^{2+} ^[137] (Figure 60). For Mg^{2+} concentrations between 0.2 and 0.8 mM^[140], the aptamer populates many conformational states at the equilibrium. Once added, the ligand binds and perturbs this equilibrium towards the complex with the appropriate tertiary interactions. For Mg^{2+} concentrations above 1.5 mM, the aptamer already forms the tertiary contacts already in absence of the ligand^[79a, 137], resulting in the very high affinity reported for this RNA aptamer in high Mg^{2+} conditions^[140-141].

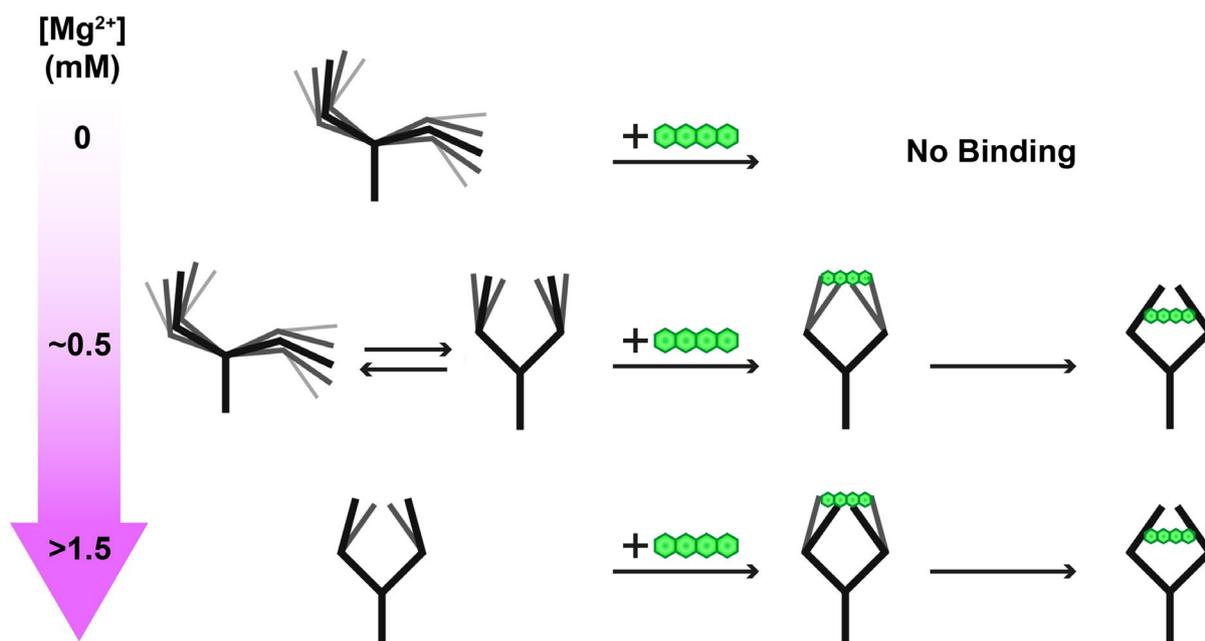


Figure 60 – TCA binding scheme at three different concentration range of Mg^{2+} , proposed by Reuss et al^[140]. No binding occurs in absence of Mg^{2+} . Between 0.2-0.8 mM, Mg^{2+} allow the aptamer to adopt multiple conformations and ligand (green) binding can occur. For Mg^{2+} above 1.5 mM TCA tertiary contacts almost fully formed and the final complex conformation is obtained by ligand binding.

According to previous studies^[140], at $[Mg^{2+}]$ 100 μM , tertiary contacts of unbound aptamer are partially stabilized. The complex, instead, is restrained to the structure showed in Figure 59C made by long-range interactions. We expect to observe two different conformational ensembles for complex and unbound aptamer, by IM-MS. The relative small size of TCA (58 nt) and its ability to form a tertiary structure in presence of Mg^{2+} ^[140], make this aptamer an interesting candidate for MS analysis.

In this work TCA was transcribed in vitro using a Hind-III linearized plasmid. A solution of 10 μM TCA was analysed, by nESI-IM-MS, in a mixture 5 μM tetracycline, TMAA 100 mM, 100 μM $MgCl_2$ and 100 μM KCl. Regardless of repeated desalting steps with NH_4OAc and TMAA, the TCA signals are partially suppressed by a large currents of alkali clusters, presumably from the precipitation steps after transcription. However, the aptamer is visible from 7- to 17- charge state and complex is visible on charge states from 9- to 12-. We found that tetracycline binding is visible only in presence of Mg^{2+} , as expected (Figure 61). Agilent 6560 pre-IMS parameter are listed in Table 5.

M and ML maxima in Figure 61 correspond to the ions with two Mg^{2+} and two K^+ . Interestingly, a minor signal of tetracycline complex with only one Mg^{2+} is visible for charge states 9-, 10- and 11-. These ions with the minimal Mg^{2+} required for binding, could be a complex not fully folded as consequence of ligand binding on (at least) one of the free-aptamer conformers possible at this Mg^{2+} concentration.

However the intense noise on the spectrum prevent for unambiguous adduct stoichiometry assignment, and for further analysis totally synthetic aptamers would be needed.

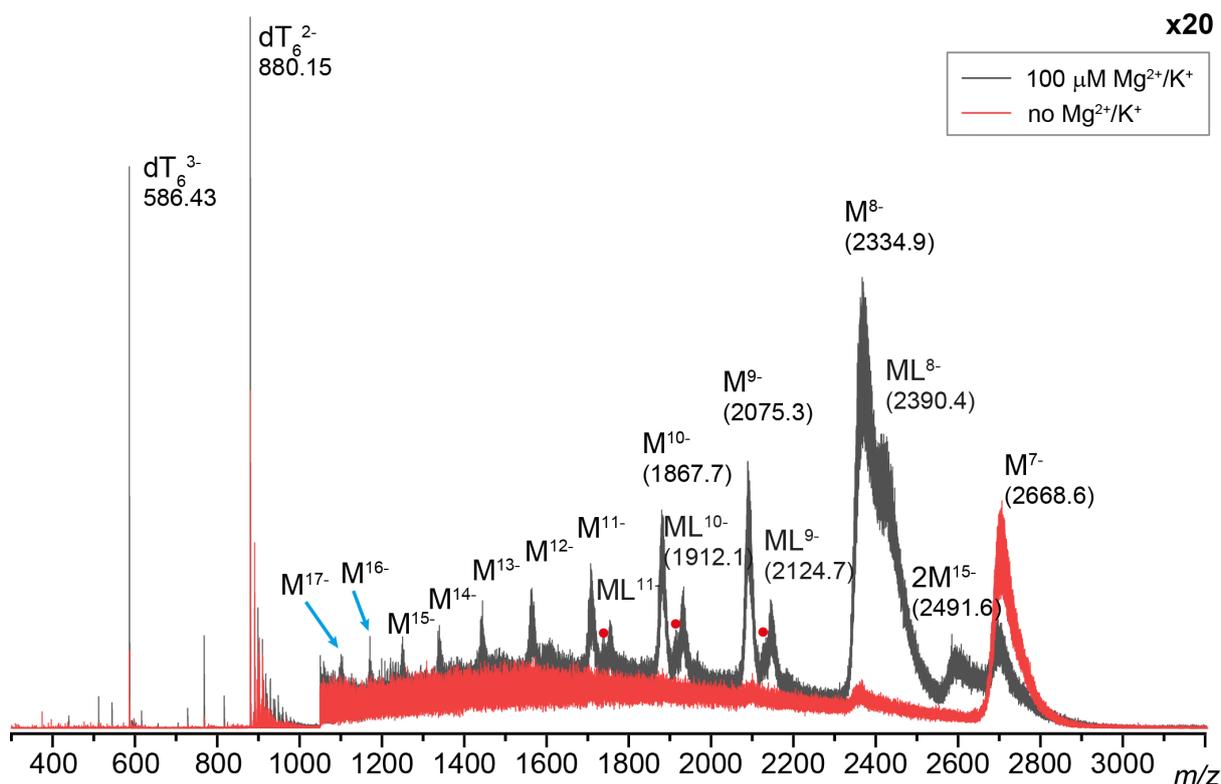


Figure 61 – **Black spectrum**: TCA 10 μM and 5 μM tetracycline in TMAA 100 mM with 100 μM MgCl_2 and KCl. Spectrum is 20-fold zoomed from 1050 m/z . Charge states are from 7- to 18-. Binding is visible up to 12-. M and ML peak maxima are with 2 Mg^{2+} and 2 K^+ for all charge states, in brackets are reported the m/z values of declustered ion. The symbol “●” highlight a complex with only one Mg^{2+} adduct. Superimposed Mg^{2+} and K^+ distributions are visible on dT_6^{2-} . **Red spectrum**: no binding is detected and a very low charge state distribution is visible in absence of MgCl_2 and KCl.

Interestingly, from the CCS distribution of M and ML we observe at least two broad conformational ensembles. In Figure 61 we report the CCS distribution of charge state 9- because it provides the best compromise between intensity and clustering (Figure 117 – p.193). For TCA^{9-} we observe a main broad peak centred at 1400 \AA^2 , followed by the background alkali currents from 1600 \AA^2 . The unbound aptamer shows two distributions: the main one at 1359 \AA^2 and a broader and more extended one at 1488 \AA^2 . For $[\text{TCA}+\text{L}\cdot\text{Mg}_2]^{9-}$ is possible to deconvolute three CCS distributions: the most intense at 1360 \AA^2 , 1456 and 1497 \AA^2 . For $[\text{TCA}+\text{L}\cdot\text{Mg}_2\text{K}_2]^{9-}$ we found a reversed situation to TCA^{9-} : we have a minor distribution at 1377 \AA^2 and the main distribution at 1487 \AA^2 .

A priori, one would expect to find a compact CCS distribution for fully folded complex and a more extended distributions for the free-aptamer conformers, rather than what we see in Figure 62. But in ESI-MS we would expect that the not fully folded conformers are prone to compact upon transfer into gas phase^[64], whereas some native conformation would be retained thanks to the tertiary contacts of its characteristic structure (X-ray crystal in Figure 60C).

The broad distributions for TCA^{9-} and $[\text{TCA}+\text{L}\cdot\text{Mg}_2]^{9-}$ are coherent with the anticipated mechanism of Figure 60 at relatively low Mg^{2+} : the unbound aptamer explores multiple conformations at the equilibrium and tertiary contacts are not formed without the appropriate number of metal adducts^[141-142], even with ligand bound^[79a, 140]. The reversed proportions of CCS distributions for $[\text{TCA}+\text{L}\cdot\text{Mg}_2\text{K}_2]^{9-}$, compared to TCA^{9-} and $[\text{TCA}+\text{L}\cdot\text{Mg}_2]^{9-}$, suggest that the distribution centred at 1487 \AA^2 is the complex fully folded thanks to the minimal Mg^{2+} and K^+ necessary for (backbone) charge shielding. Conversely, the minor distribution at 1377 \AA^2 is a fraction of not fully folded complex or simply a tailoring effect of $[\text{TCA}+\text{L}\cdot\text{Mg}_2]^{9-}$ adducts tail that superimposes.

IM-MS spectra of charge states 8- and up to 12- are in line with this interpretation (Figure 119 and Figure 120 at p.195). CCS values are listed in Table 57 at p.197.

However, some precautions should be taken: a better Mg^{2+} and K^+ adduct stoichiometry should be addressed. A synthetic RNA aptamer would provide better-resolved peaks, avoiding the intense alkali contamination coming from yeast plasmid transcription and successive purification and precipitation steps.

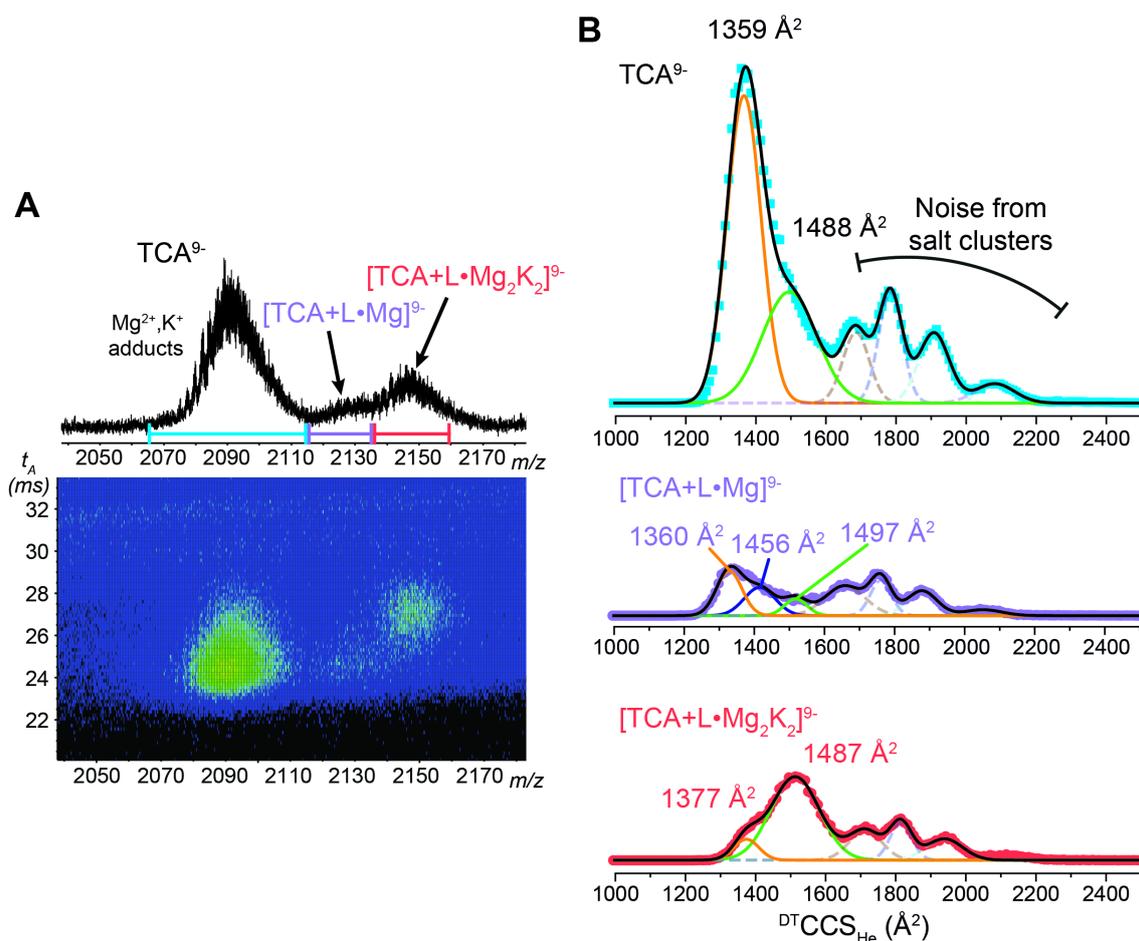


Figure 62 – **A**. IM-MS spectrum of TCA^{9-} and ATDs extraction ranges reported as colored bars **B**. CCS distribution of TCA^{9-} (cyan), $\text{TCA}+\text{L}^{9-}$ (violet) and $[\text{TCA}+\text{L}\cdot\text{Mg}_2\text{K}_2]^{9-}$ (red). Dashed distributions are from salt clusters.

5.6 Cocaine Binding Aptamers

5.6.1 Introduction

Cocaine Binding Aptamers (CBA) were selected during the early 2000s to design molecular beacons for cocaine and its metabolites^{[15] [35b]}. These aptamers share a relatively simple secondary structure constituted by two hairpin-loops (stem 2, stem 3), a duplex/helical segment (stem 1) with a three-way junction. The cocaine binding site is located between the junction and the bulge (Figure 63)

This aptamer family is also capable to bind other alkaloids and steroids such as cocaine metabolites, quinine and some of its analogs^[87, 143], and sodium deoxycholate^[144]. This large array of potential targets is atypical for an aptamer, normally intended to be specific for a ligand. During SELEX protocols, the “evolutionary” pressure is supposed to select a sequence that will bind the target specifically. At the time, cocaine presence in SELEX induced an ‘evolutionary’ stimulus to specifically bind either a saturated polycyclic moiety or a bicyclic aromatic ring, rather than cocaine itself^[143]. This peculiarity leaves room to conceive sensors based on molecular beacons, for which CBAs are still widely studied today.

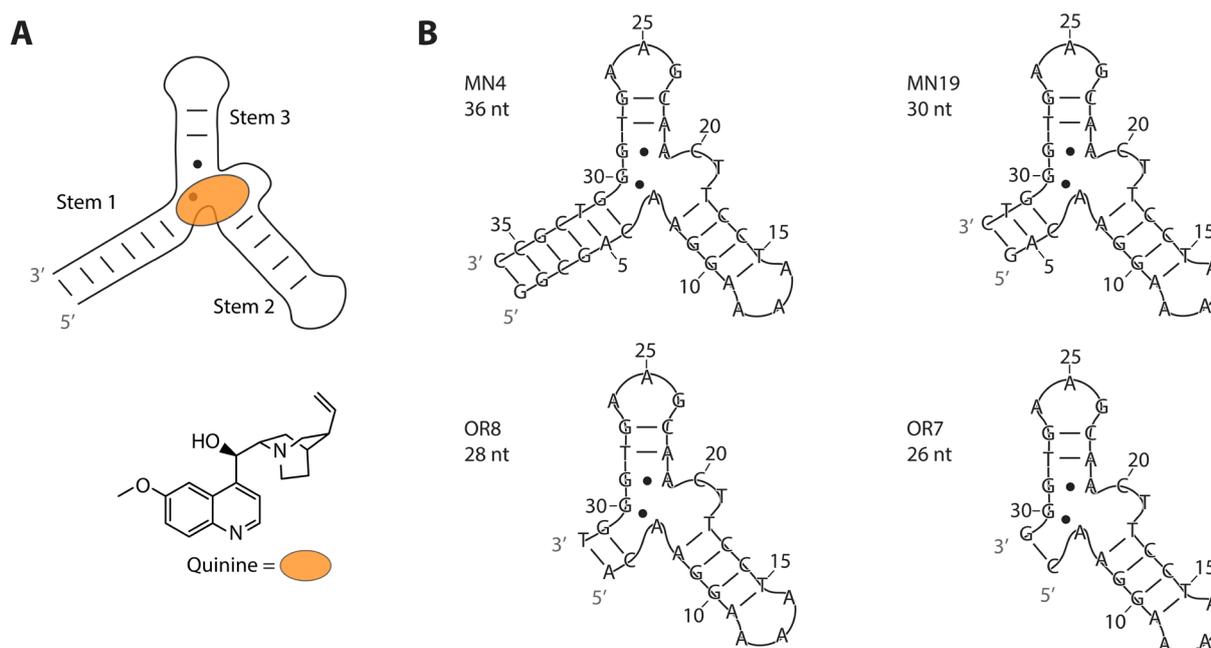


Figure 63 - **A)** Topology for DNA cocaine binding aptamers^[145]. The “•” represent a non-WC base pairing. Quinine (Ligand) binding site is proposed in the junction. **B)** Aptamers analyzed in this chapter. The topologies are Sequence numbering and topology is maintained from Johnson and co-workers^[88, 145] for consistency.

The junction in CBAs (Figure 63A) is required to form the binding site and bind quinine. Stem 1 length is critical in the appropriate folding of this junction: in MN4 stem 1 is 6-bp long, is folded and participates to the junction in free and quinine bound form^[145]. When stem1 is 3-bp or less (i.e. MN19, OR8, OR7), it is less folded. Consequently, the junction is less folded and the aptamer has a lower affinity for quinine (K_D between 11.2 and 0.14 μM in function of stem 1 length)^[88]. However, quinine binding contributes to the formation of the junction and of

stem1, following the concept of “adaptive binding” [68, 88]. This hypothesis is supported by ITC, NMR^[145], spin-labelled EPR^[79c] and enzyme digestion^[146].

We will focus on the sequences MN4, MN19, OR8 and OR7 (Figure 63), a series of CBAs reported by Johnson’s lab [88], and their quinine complexes. The sequences differ in length of stem 1, from 1 bp (OR7) to 6 bp (MN4), keeping the same core sequence.

The stabilization of secondary (stem1) and tertiary structure (junction) makes CBAs good candidates to explore whether the adaptive binding of aptamers can be studied by ion mobility mass spectrometry (IM-MS). Thus, for complexes we would expect CCS values corresponding to more compact conformations, compared to the free aptamers.

5.6.2 IM-MS to track adaptive binding

MN4 is expected to be fully folded both in its free and quinine-bound forms. We thus expect similar collision cross sections for free and bound MN4. Analysing by ESI-IM-MS a solution of 10 μM MN4 with 5 μM quinine in NH_4OAc 100 mM, we observe three dominant charge states (7-, 6- and 5-) (Figure 64), with the same arrival times and thus collision cross section distributions for free aptamer and its quinine complex (Figure 65).

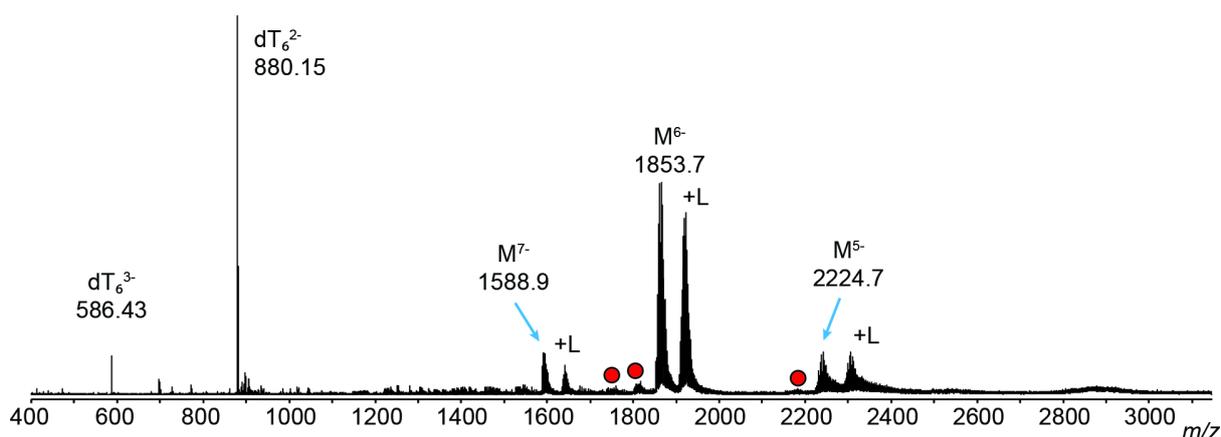


Figure 64 - MS spectra of MN4 (10 μM + 1 μM dT₆) with 5 μM of quinine in NH_4OAc 100 mM. Low charge states distribution of 7-, 6- and 5- is visible with quinine binding on each and NH_4^+ adduct distribution. Impurities “●” from missing nucleotide coupling (up to M-2nt) are visible on the dominant charge states. Fragmentor 350V and parameters as at p.35.

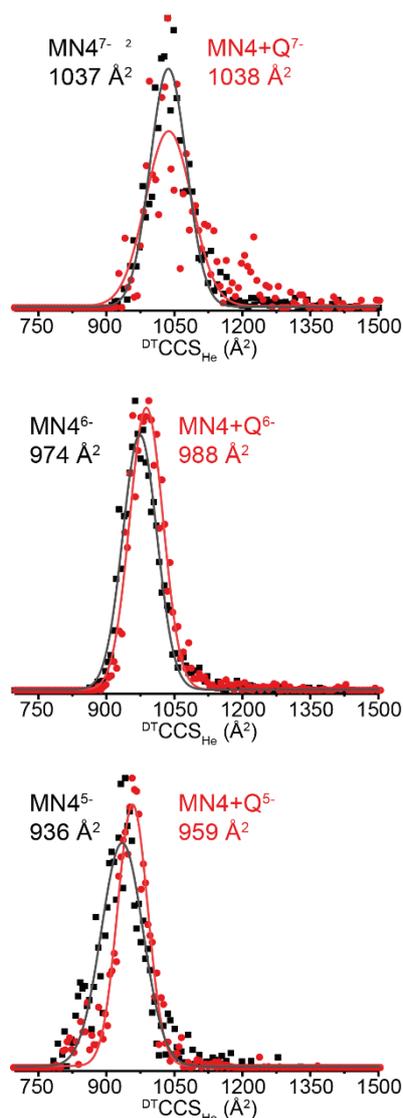


Figure 65 - Normalized $^{DT}CCS_{He}$ distributions for MN4 ch st 7-, 6-, 5- and relative quinine complexes at low pre-IMS activation (350V) in NH_4OAc 100 mM. Only desolvated ions (no NH_4^+ adducts) are considered for CCS determination.

At low activation, $MN4^{7-}$ shows a single conformational ensemble for the free-aptamer and its quinine complex, and the same goes for charge states 6- and 5- (Figure 65). The CCS distributions of bound and unbound MN4 are very close. This observation is compatible with the accommodation of quinine without a significant shape change of the aptamer, but is not in line with our expectations for an adaptive binding. One cannot differentiate quinine complex from free-aptamer solely on the CCS distributions as in Figure 65.

Upon activation (Fragmentor 600 V) we observe the dissociation of quinine complex ($MN4+Q^{7-}$ peak is absent at this voltage) and a significant change in ion mobility spectrum of $MN4^{7-}$ (Figure 65B): a transition from the conformational ensemble centred at 1037 \AA^2 , at low activation, to a second centred at 1197 \AA^2 . This CCS increase of 11.5% indicates a shape change towards an extended conformational ensemble. This extension is compatible with the

unfolding of the ion, such as either the stem1 opening proposed for this CBA^[79c] or the opening of the whole base pairing pattern.

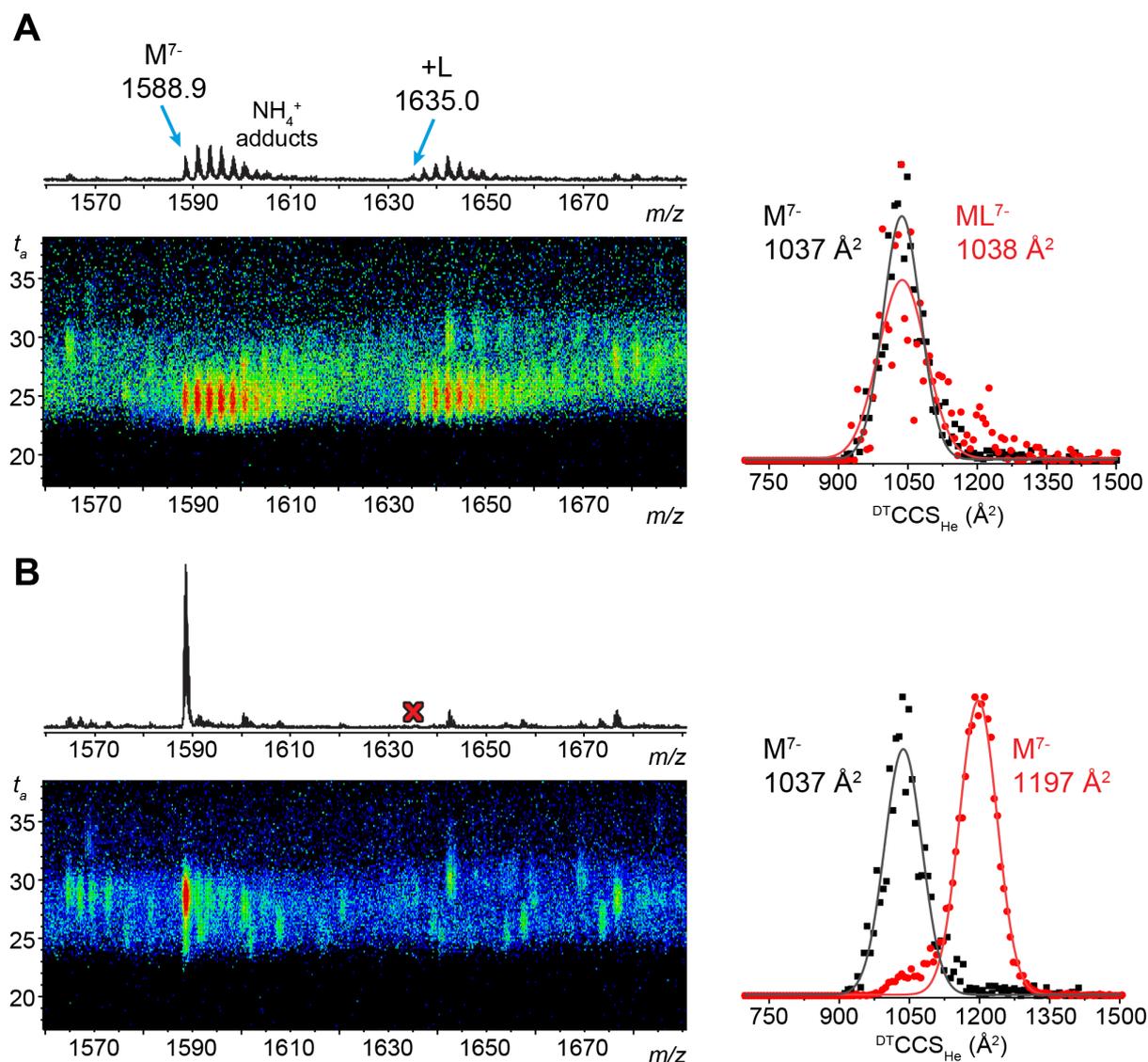


Figure 66 – IM-MS of MN4 10 μ M and 5 μ M quinine in NH_4OAc 100mM (Ch St 7-). **A.** Spectrum acquired with fragmentor at 350V and reconstructed CCS distribution for $MN4^{7-}$ and its quinine complex. Very similar CCS are consistent with a pre-formed aptamer that binds to its ligand. **B.** At 600V we observe complex dissociation “X”, NH_4^+ declustering of $MN4^{7-}$ with a significant change of CCS distribution. Normalized CCS distribution for $MN4^{7-}$ show the transition from one ensemble, centered at 1037 \AA^2 (350V), to a second ensemble, centered at 1197 \AA^2 (600V) upon ion activation.

The pre-IMS activation do not alter the CCS distributions of charge states 6- and 5-, presumably because the lower charge bring an higher kinetic stability of the ions.

In contrast to MN4, the aptamers MN19 and OR8 are not pre-structured before ligand binding, but upon binding their secondary structure (at least for MN19) should change towards a topology similar to MN4 aptamer^[87-88, 145]. In this framework, if MN19 and OR8 in gas phase keep a memory of their solution structures, these aptamers are more likely to show

differences in arrival time distributions, e.g. a more compact quinine complex compared to the free aptamer.

Before direct gas phase comparison, we notice that MN19 and OR8 have a different charge state distribution compared to MN4 (Figure 67):. This shift is a consequence of the different sequence length of the aptamers, MN4 (36 nt), MN19 (30 nt) and OR8 (28 nt). We also observe a progressive decrease of quinine complex (Figure 67). This is consistent with the lower extent of pre-folding, being associated with lower binding affinity as suggested by Neves *et al.*^[88].

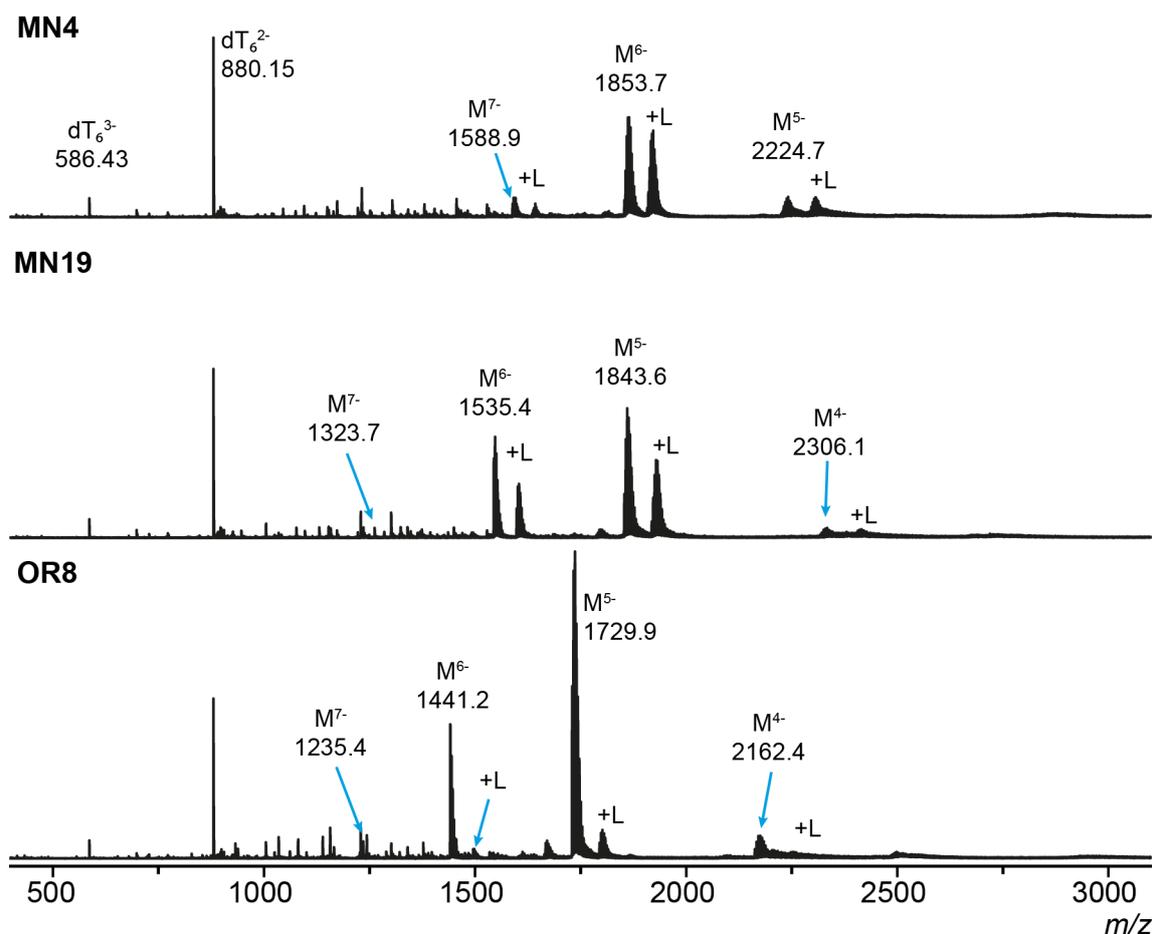


Figure 67 - MS spectra of MN4, MN19 and OR8 (10 μ M + 1 μ M dT₆) + 5 μ M quinine in NH₄OAc 100 mM (Fragmentor 350V). All CBA show a distribution of low charge state, 7- to 5- for MN4 and 6- to 4- for MN19 and OR8. This difference in charge uptake is due to the different chain length of the aptamers. Quinine binding is detected for all the sequences with progressive lower ratio, as reported from Neves *et al.*^[88].

Besides the difference in the number of nucleotides, we can qualitatively compare MN4, MN19 and OR8 at charge states 7- and 6-. Upon collisional activation, charge state 7- passes from one single distribution to a second one (Figure 68). However, charge state 6- does not (Figure 68BDF)

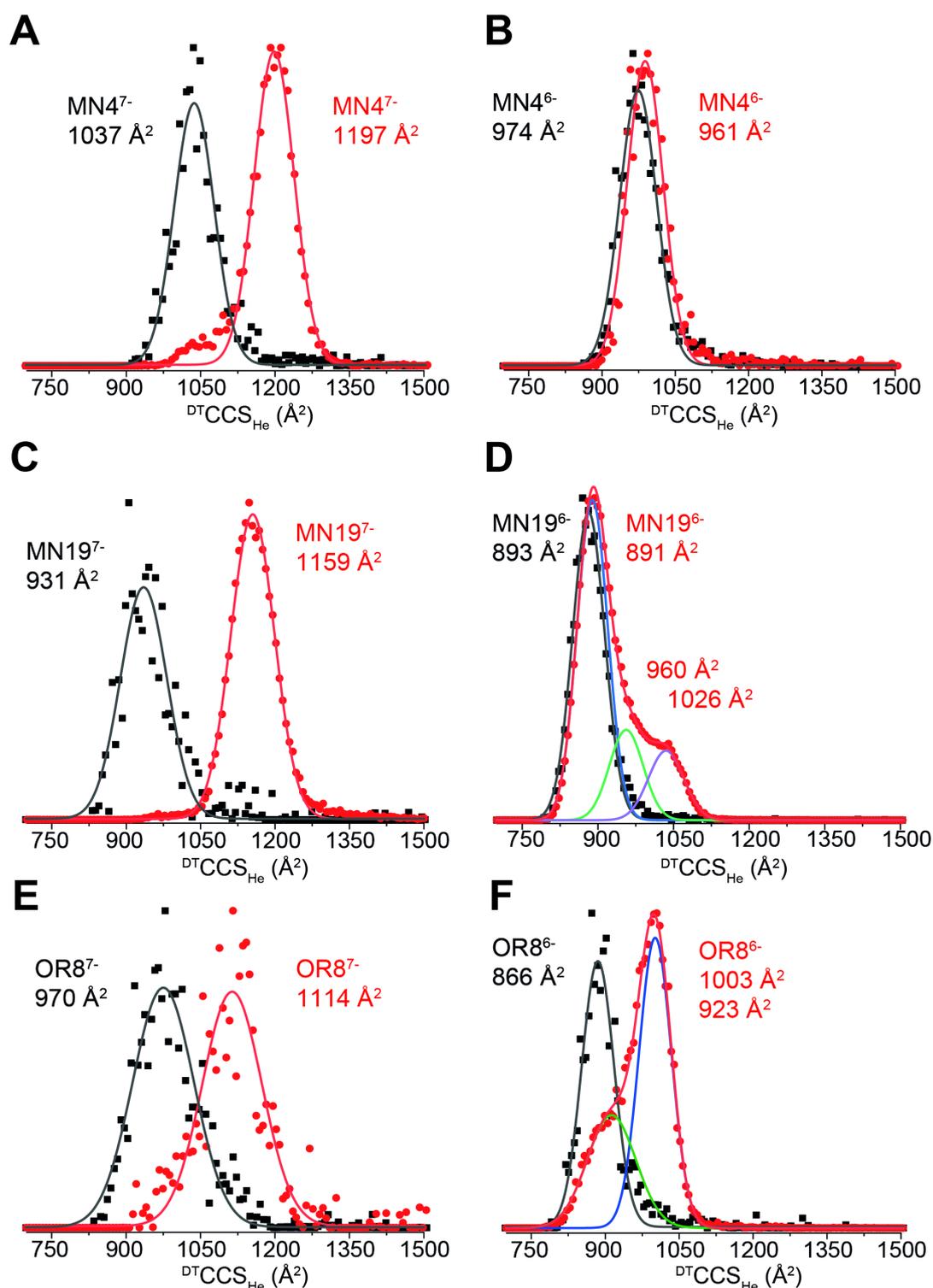


Figure 68 – Normalized reconstructed CCS distributions of MN4 MN19 and OR8 charge states 7- and 6-. **Black** curve are at low activation (F350V) and **red** curves are at high activation (F600V). Charge state 7- is reported on the left and charge state 6- on the right. CCS values with uncertainty are reported in Table 59, Table 60 - p.209

Charge state 7- could be a good candidate for further investigations if this CCS extension is related to the specific secondary (and tertiary) structure common to these aptamers. However the different chain length complicates a direct comparison of the energetics

between these ions, because aptamers of different lengths have different numbers of degrees of freedom and thus different internal energies at a given fragmentor voltage. Also, their internal Coulomb repulsion forces differ if the length differs, because the charge density differs. CCS distributions for charge state 6⁻ exemplify these concepts: at low activation we observe that aptamers have a single distribution.

But discrepancies appear with activation. Setting fragmentor at 600 V, we observe a minor extension for MN19⁶⁻ with two minor distributions at 960 and 1026 Å² (Figure 68D) while for OR8⁶⁻ we observe a more significant fraction of the population extended at 1003 Å² (Figure 68F).

5.6.3 Maximizing the CBA⁷⁻ signal by changing the [NH₄OAc]

In parallel to sequence modifications with dA overhangs, we decided also to optimize the intensity of charge state 7⁻, which gives the most prominent unfolding upon activation. One can vary the concentration of NH₄OAc to shift the charge state distribution^[147]. At 25 mM NH₄OAc the distribution shifted to higher charge states (Figure 69B - up to 9⁻) and charge state 7⁻ intensity is increased.

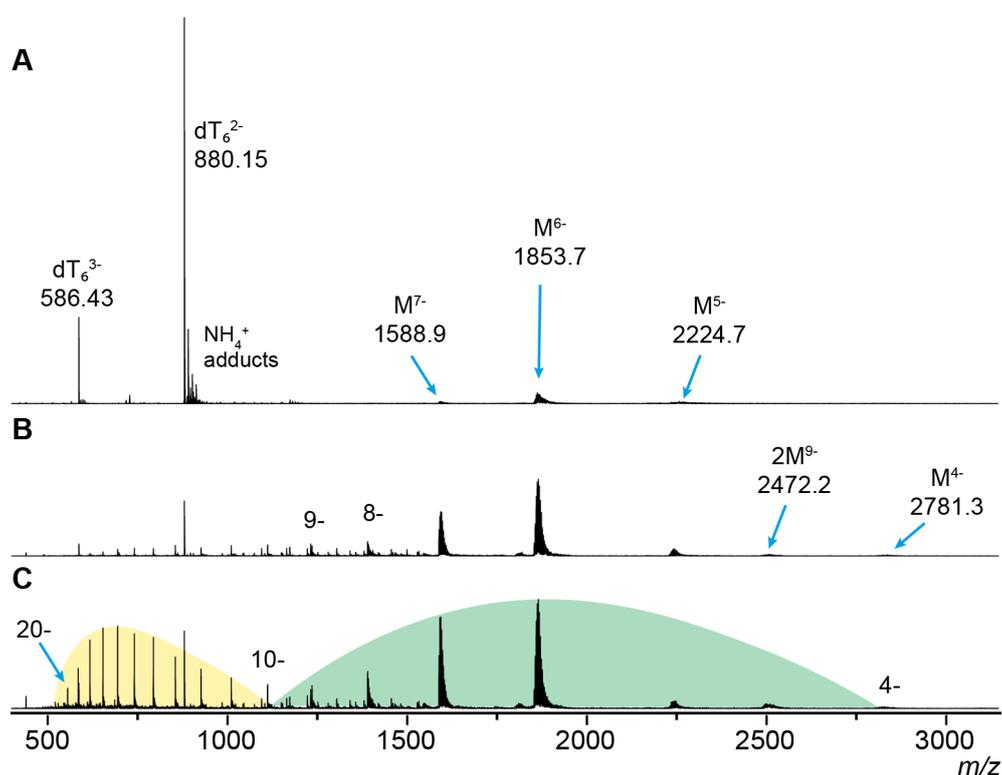
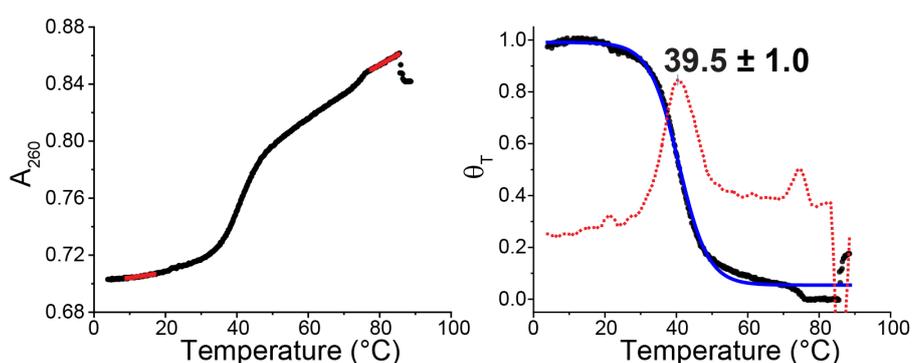


Figure 69 - MN4 10 μM and dT₆ 2.5 μM at 100 (A), 25 (B) and 10 (C) mM NH₄OAc. Decreasing the ionic strength MS spectrum shifts from one charge low state distribution, 9⁻ to 4⁻ (green), to a coexistence with a second distribution of high charge states, 20⁻ to 10⁻ (yellow).

How MS high/low charge state distributions is related to DNA or RNA folding in solution, is still not fully understood^[148]. When the NH₄OAc concentration is further decreased to 10 mM

a second charge state distribution emerges below 1200 m/z (Figure 69C). The appearance of a distribution of high charge states can be interpreted as consequence of partial unfolding of the aptamer, if one borrows the conceptual framework from protein mass spectrometry, where unfolded proteins give high charge states ^[149]. However, UV-melting profiles of MN4 suggest a fully folded state (at 20°C) in NH_4OAc 25 mM and 96 % folded in NH_4OAc 10 mM (Figure 70). Thus the abundance of a high-charge distribution isn't necessarily indicative of MN4 unfolding in solution. We think that DNA ions can easily compete for the droplet surface during ESI when using the lower ionic strength of 10 mM NH_4OAc . The bimodal charge state distribution could be interpreted in the framework of co-existing ionization via the CRM (low charge states) and CEM (high charge states) pathways^[59] (see p.16). Evidence of such apparent bimodal distribution have been reported in our group on i-motif structures and unstructured controls, for which the supporting electrolyte concentration similarly influenced the charge state distribution. One should bear in mind that DNA carries a negatively charged phosphodiester backbone, in contrast to proteins, and thus could have a different behaviour in ESI.

NH_4OAc 25 mM



NH_4OAc 10 mM

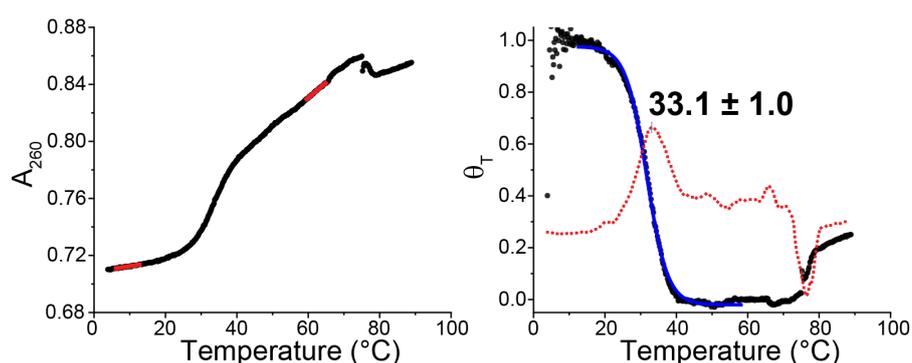


Figure 70 – MN4 aptamer result fully folded at 20°C in NH_4OAc 25 mM and 96% folded in 10 mM.

5.6.4 Model constructs for comparison in gas phase: extra dA

To circumvent the problem of shifted charge distribution and different number of degrees of freedom, we designed some constructs of the same length of MN4 (36nt). By adding 6 (in MN19), 8 (in OR8) and 10 (in OR7) extra deoxyadenosines on the 3' and 5'-ends, we have nearly the same molecular weight while keeping the core sequence un-altered. This leads to the same charge distribution across the constructs. We added "-A" to distinguish these modified constructs.

Deoxyadenosine was chosen because is one of the weakest homo base pairs mismatches, for which we have the following stability trend: GG > TT \geq AA > CC^[150]. Overhangs of dG and dC repeats were excluded *a priori* to avoid secondary structure formation like G-quadruplexes^[91a, 151] or i-motifs^[152]. We tested dT repeats initially, but discarded this modification because the overhangs seemed to perturb significantly the secondary structure.

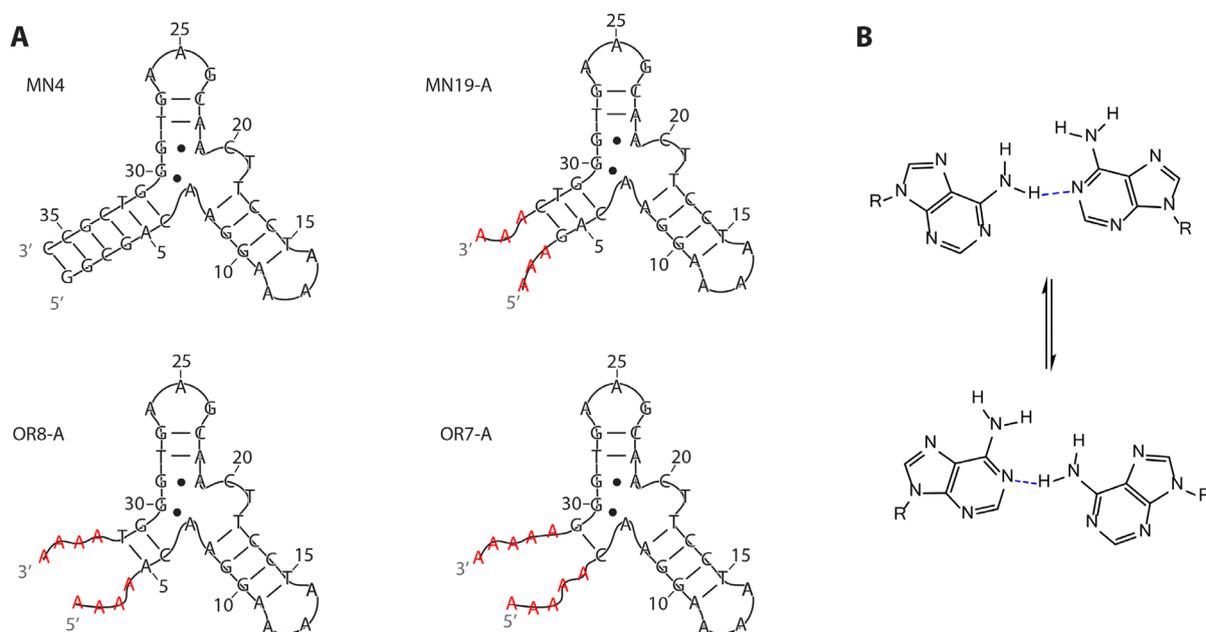


Figure 71 – **A.** Constructs of cocaine of cocaine binding aptamers with extra adenosines overhangs (red); **B.** dA-dA mismatch isomerism from ref^[153].

5.6.5 Do the dA modifications alter aptamer structure in solution?

To be sure that the aptamers with dA constructs correctly mimic the non-modified aptamers in solution (MN4, MN19, OR8 and OR7 - Figure 63), we compared their UV-melting, CD spectra (annexes p.205) and $^1\text{H-NMR}$ of original sequences [88, 144a, 145] in our MS sample conditions (25 mM NH_4OAc).

5.6.5.1 UV-Melting profiles

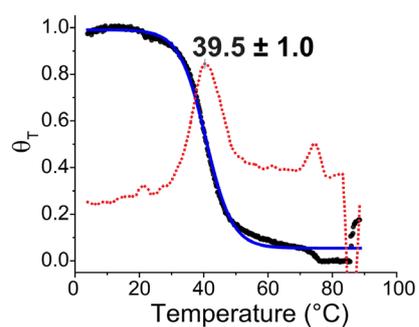
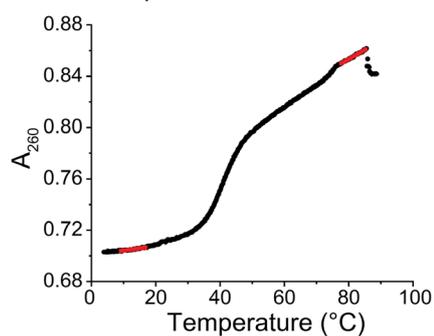
UV-melting experiments for MN4, MN19-A, OR8-A and OR7-A were carried at $[\text{DNA}] = 2 \mu\text{M}$ in NH_4OAc 25 mM as described at p. 27. We found a melting temperature of $39.5 \pm 1.0^\circ\text{C}$ for MN4 (Figure 72) and no melting for MN19-A, OR8-A and OR7-A (p.201 - Figure 122). In presence of one equivalent of quinine, the melting temperature of MN4 is little increased to $43.4 \pm 0.1^\circ\text{C}$, MN19-A shows a melting temperature at $24.3 \pm 0.1^\circ\text{C}$ (Figure 72) and no melting for OR8-A and OR7-A (p.201 - Figure 122).

The relative stability order that we found in NH_4OAc 25 mM, is the same reported by Neves *et al.* (in NaCl 140 mM and NaH_2PO_3 20 mM buffer) [88]. We highlight that the absolute values of melting transitions in NH_4OAc 25 mM are lower to by Neves *et al.* [88] because of the lower ionic strength of the solution [154].

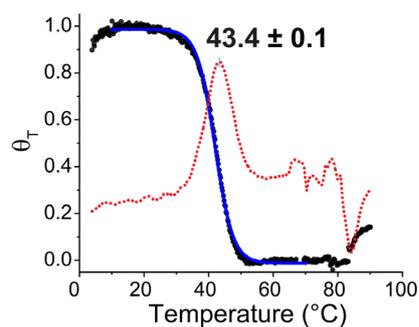
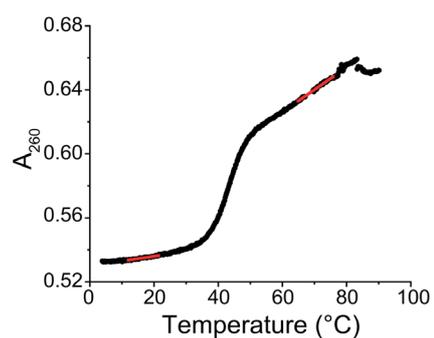
The absence of UV-melting for unbound aptamers MN19-A, OR8-A and OR7-A suggests the absence of a folded structure in this range of temperature (Figure 122, Figure 123), as expected. Notably, MN19-A does not show significant difference compared to MN19 that melts only in presence of quinine (Figure 72 and p.202 - Figure 123).

So we conclude that aptamers structure and their binding mode are unaltered by the dA overhangs. More detail about the base pairing can be addressed by $^1\text{H-NMR}$ in the next paragraph. UV-meltings in NH_4OAc 100 mM are reported at p. 188.

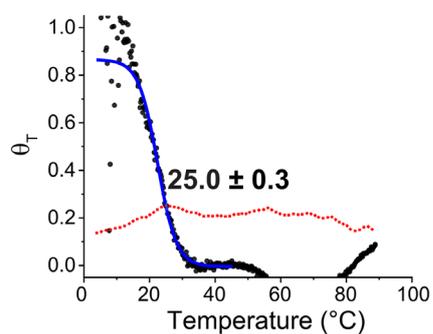
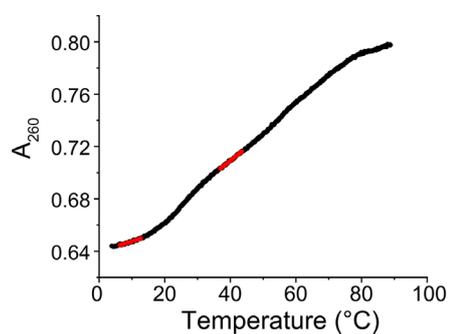
MN4 in NH₄OAc 25 mM



MN4 +quinine



MN19 +quinine



MN19-A +quinine

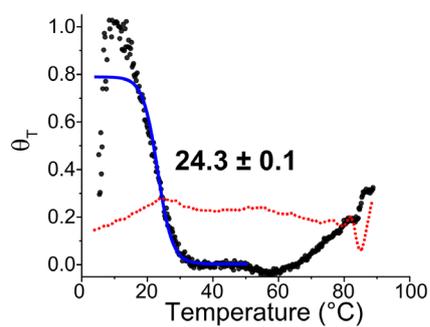
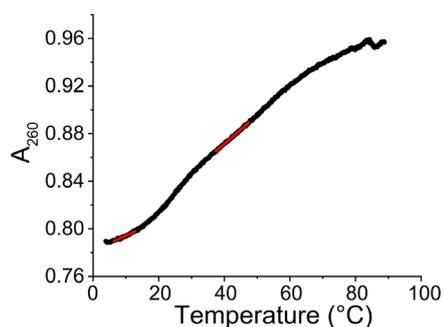


Figure 72 - UV-melting plots in NH₄OAc 25 mM of MN4 2 μ M, MN4 with 2 μ M quinine, MN19 and MN19-A 2 μ M with 2 μ M quinine. MN19 and MN19-A do not show any melting transition in this buffer and temperature range (p.202 - Figure 123). Baselines for curve fitting are reported in red filled line, first derivative of A₂₆₀ is reported in red dotted line.

5.6.5.2 $^1\text{H-NMR}$ in NH_4OAc

Imino protons $^1\text{H-NMR}$ resonances of nucleobases (and base pairs) fall in the window of 10-14 ppm and a complete peak assignment by 2D NMR in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) at 278 K [87, 145] is available for MN4 and MN19 (quinine complex only). First we can check if MN4 and MN19 secondary structures [87, 145] are present in NH_4OAc 25 mM. Secondly, we can qualitatively estimate if the construct MN19-A has a similar secondary structure to MN19.

Spectra are acquired on a 700 MHz Bruker spectrometer in [DNA] 100 μM / 450 μL , NH_4OAc 25 mM, 5% D_2O at 278K. The low temperature is necessary to populate folded structures as much as possible.

From the MN4 spectrum (Figure 73B), we found all the characteristic imino resonances assigned for the “ γ -shaped” topology proposed from Johnson and co-workers[145]: the two hairpins are represented, T15,16,28 and G9,10,27, the 3',5'-duplex with G2,4 and T32, and finally the non-canonical pairing on G29,30.[145]

For MN19 the base pairing assignment is available only for its quinine complex in water[87]. In NH_4OAc 25 mM and absence of quinine, we detect a partial folding: the highly conserved imino resonances T15, T18 and for G9, G10, G27, G24 and G30 support the folding of the two hairpins and the duplex pattern between the junction and the 3',5' ends. However, stem 1 doesn't seem to be folded, as there is no resonance for G31.

For MN19-A, we do not have a NOESY spectrum for fine assignment of bases, but the strong similarities with MN19 spectrum suggest that secondary structure is similar. T18 and G29,30 are present, suggesting that the two hairpins are well preserved. Caution should be taken for the junction and the 3',5'-duplex: if conserved G29,30 resonances would support the junction as depicted in Figure 73A, the attribution to G4, G31 and T28 (asterisks on Figure 73B) remains tentative, and their confirmation would require further 2D-NMR investigation.

Globally, for MN19 and MN19-A, imino proton peaks are less sharp compared to MN4. This means that the relative protons are either in chemical exchange (more exposed to the solvent) or in conformational exchange (protons in fast equilibrium between different nucleobases). This is compatible with a less rigid secondary structure compared to MN4, coherently with literature [87, 155]. Thus we could reasonably expect a similar secondary structure for MN19 and MN19-A in our conditions.

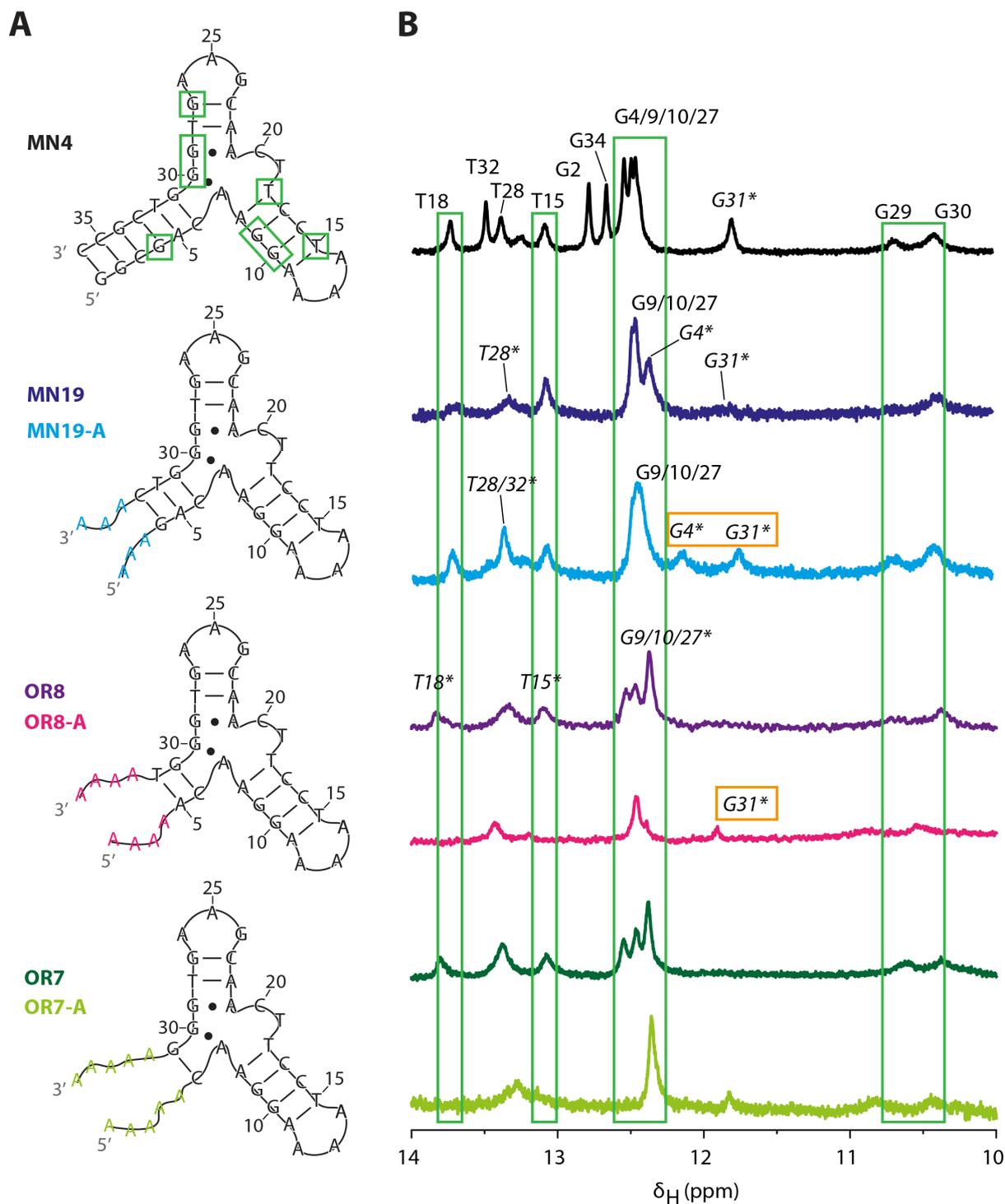


Figure 73 – **A**. Proposed topologies for MN4, MN19^[87, 145] OR8, OR7 and relative adenine constructs; **B**. Comparison of imino proton resonances. A shared secondary structure is suggested by the characteristic resonances (green boxes). Orange boxed peaks suggest a 3',5'-end stabilization for MN19-A and OR8-A. Base assignment (*) are indicative and deduced by comparisons with MN4 and MN19 (see text). Spectra are acquired in [DNA] 100 μ M / 450 μ L, NH_4OAc 25 mM, 5% D_2O at 278K.

For OR8 and OR7 sequences no 2D-NMR are available so the comparison remains qualitative. Compared to MN4 and MN19, the spectra show fewer peaks and many are broad (Figure 73 – pink, violet and green), in particular for OR8-A and OR7-A constructs. This means that the imino protons can exchange with the solvent because they are little engaged in base pairing. So we may expect a structure with two hairpins (Figure 74), similarly to what Cekan *et al*^[79c] reported.

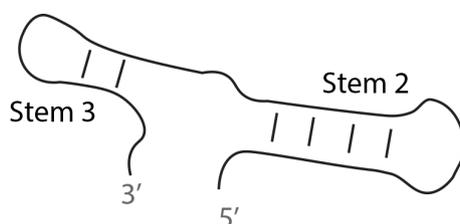


Figure 74 - Hypothetical structure of unbound OR8, OR7 and relative extra dA constructs OR8-A and OR7-A.

NMR spectra in Figure 73B and the hypothetical structure in Figure 74 agrees with the results of UV-Melting (Figure 72 and p.201 - Figure 122, Figure 123). Thus, the absence of UV-melting in OR8, OR7 and OR8-A, OR7-A constructs is ascribable to the low base pair content (thus low hyperchromicity – p.27) of these sequences.

From these spectra we can suggest that G4,31* in MN19-A and G31* in OR8-A (Figure 73) contributes to 3',5'-ends stabilization, thus to the second CCS distribution at 1045 and 1069 Å² for MN19-A⁷⁻ and OR8-A⁷⁻ as visible in Figure 76. We attribute it to dA overhangs that induce a minor, but visible, stem 1 stabilization. However, a definitive evidence dA contribution should be addressed by further NMR experiments (e.g. NOESY).

5.6.6 Adenine constructs in NH₄OAc 25 mM

We observed that MN4 is fully folded at 25 mM NH₄OAc and show a dominant low charge states distribution after ESI. Constructs MN19-A, OR8-A and OR7-A are fully unfolded at 25 mM NH₄OAc but their MS spectra have a bimodal charge states distribution. Thus, the unfolded structure do not shift completely the MS spectrum to high charge states, as anticipated for MN4 in Figure 69

On the other hand, we obtain the expected optimization for gas phase comparison:

- Same charge state distribution from MN4 to OR8-A, with similar m/z ratio for charge states 7- (sequence length effect);
- Similar intensities for charge states 7- (NH₄OAc concentration effect).

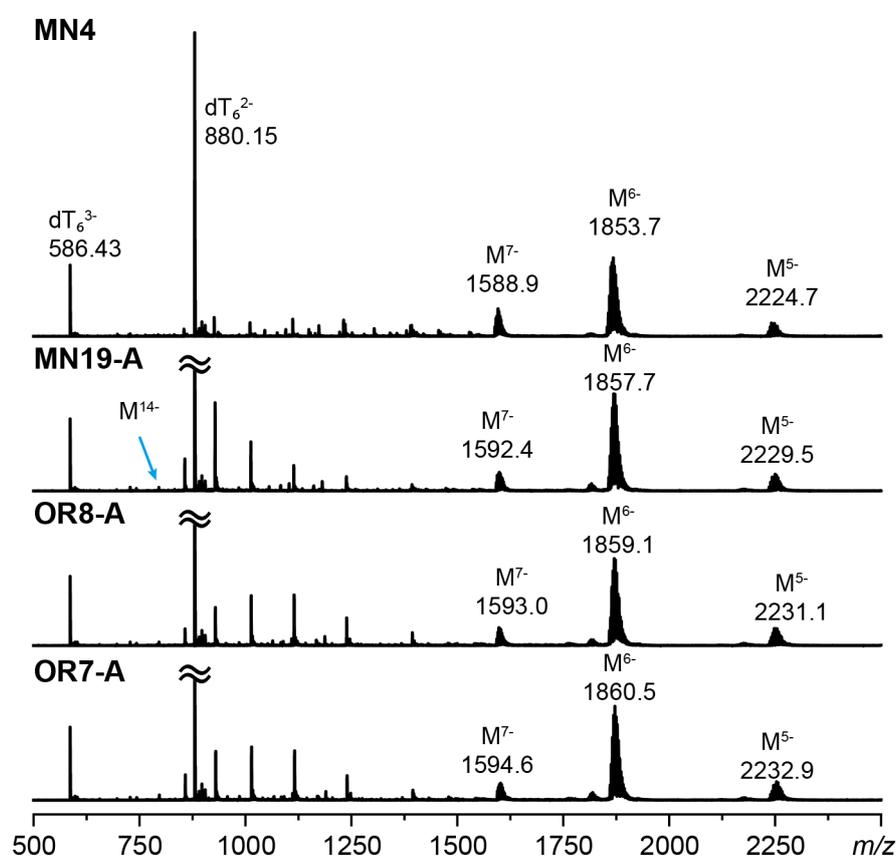


Figure 75 -MN4 and CBA+extra dA constructs at 10 μM + dT6 2.5 μM in NH₄OAc 25 mM. The length lead to the same low charge states distribution across the sequences. This simplify gas phase comparisons on ions of the same charge state across different sequences. dT₆²⁻ as been cutted for clarity.

5.6.7 IM-MS of CBA-dA aptamers

Focusing our attention on the CCS distributions of charge state 7- for each aptamer (Figure 76), we observe that all the modified constructs pass from a single distribution at low pre-IMS

activation (Fragmentor 350V) to a single and more “extended” distribution at high activation (Fragmentor 600V), similarly to what we observed for unmodified aptamers (Figure 68). Secondly, the reduced ionic strength has no influence on MN4⁷⁻ CCS distribution: compact and extended distribution are centred, respectively, at 1037 and 1197 Å² from NH₄OAc 100 mM (Figure 68) and 1036 Å² and 1205 Å² NH₄OAc 25 mM (Figure 76). Besides MN4⁷⁻, the difference between the CCS distribution is not significant, therefore, we can exclude ESI effects on charge state 7- CCS values related to a low ionic strength. For MN4⁷⁻ the center position of the CCS distribution does not vary significantly in NH₄OAc 25 mM, but it becomes broader.

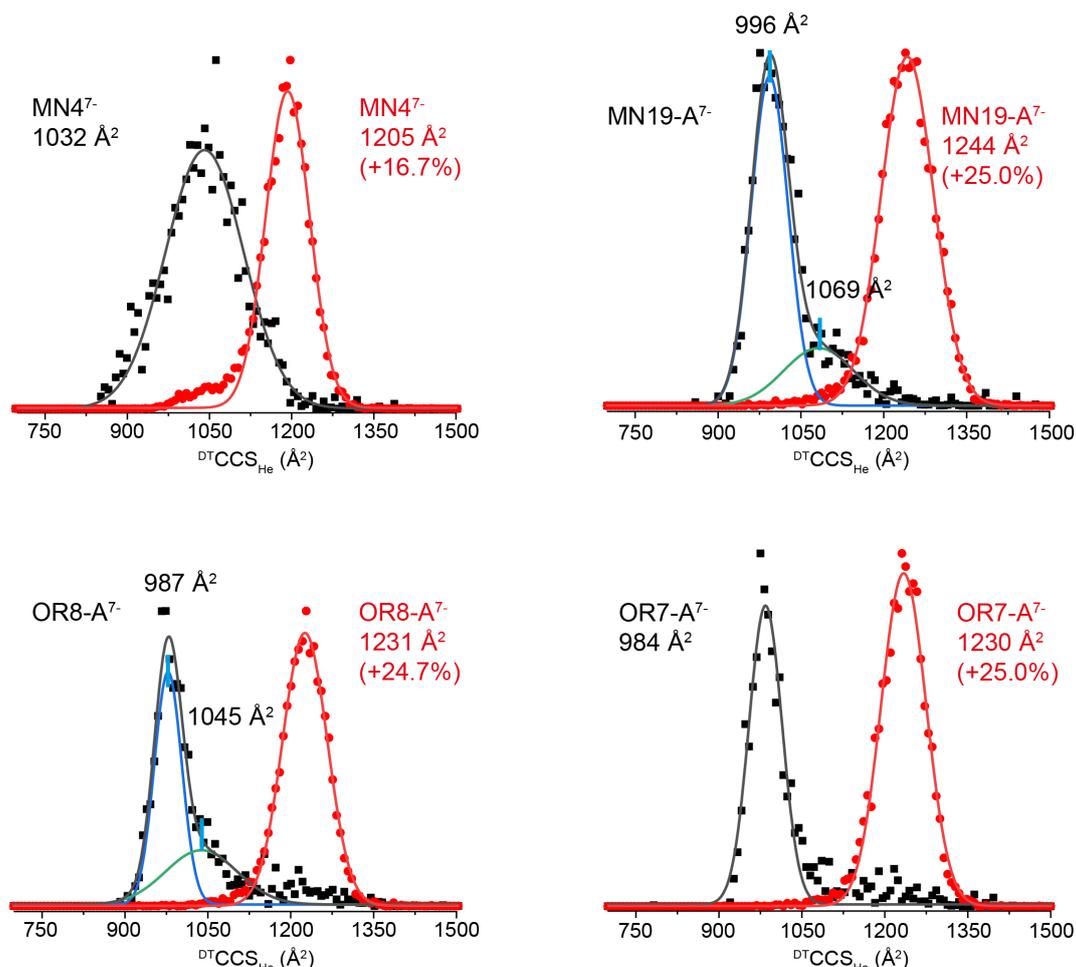


Figure 76 - Normalized CCS distributions of CBA+dA⁷⁻. Each anion has a single distribution at low activation (F350V - **black**) that extends towards a second one at high activation (F600V – **red**), similarly to MN4⁷⁻.

At low activation we found that MN19-A⁷⁻ and OR8-A⁷⁻ have two CCS distributions: one major and narrow between 984 and 996 Å²; a second between 1045 and 1069 Å² (for OR7-A⁷⁻ it is not resolved or absent).

We speculate that the dA overhangs (stem1 -Figure 71), that may compact upon transfer in gas phase^[64], are compatible with the major and compact distribution. In this sense, the second distribution (close to the CCS distribution of MN4⁷⁻) could be attributed to a memory (in gas phase) of folded aptamer in solution. From ¹H-NMR spectra (Figure 73) we may support

this hypothesis of the minor CCS distributions for MN19-A⁷⁻ and OR8-A⁷⁻ in Figure 76, based on resonances of G4,31* in MN19-A and G31* in OR8-A, that contributes to 3',5'-ends stabilization. However, a definitive evidence of dA contribution should be addressed by further NMR experiments (e.g. NOESY). Unfortunately, UV-melting are not informative because MN19-A and OR8-A melting profiles do not have a transition that would suggest a minor folded fraction (pp. 113,188).

Conversely, at high activation all sequences show a single distribution between 1205-1244 Å² (Table 22). Thus, the starting shape for MN19-A⁷⁻, OR8-A⁷⁻, OR7-A⁷⁻ is more compact compared to MN4⁷⁻ and the end point has a comparable shape for all the sequences.

Table 22 – ^{DT}CCS_{He} of MN4 and dA constructs. Reported values are the average values with their relative standard deviation (tables at pp. 209-211). † CCS obtained from CIU experiments with CCS reconstruction at 350V, otherwise we used the step-field method.

Ion	z	350 V		600 V	
		^{DT} CCS _{He} (Å ²)		^{DT} CCS _{He} (Å ²)	
[MN4-zH+]z-	8	1158 ± 22		1351 ± 13†	
	7	1032 ± 13		1205 ± 5	
				1204 ± 8†	
	6	974 ± 7		966 ± 4	
				963 ± 2†	
	5	919 ± 16		934 ± 6	
[MN19-A-zH+]z-	8	1163 ± 9		1388 ± 22†	
	7	996 ± 9	Main	1244 ± 3†	
		1069 ± 18	Shoulder		
	6	962 ± 3		971 ± 5†	
	5	924 ± 14		938 ± 3†	
[OR8-A-zH+]z-	8	1356 ± 18		1374 ± 1†	
	7	987 ± 11	Main	1231 ± 3†	
		1045 ± 20	Shoulder		
	6	966 ± 1		973 ± 3†	
	5	944 ± 12		951 ± 8†	
[OR7-A-zH+]z-	8	1344 ± 4		1364 ± 10†	
	7	984 ± 7		1230 ± 3†	
	6	957 ± 1		967 ± 2†	
	5	935 ± 7		938 ± 2†	

To know if the differences found at low activation between MN4⁷⁻ and MN19-A⁷⁻, OR8-A⁷⁻, OR7-A⁷⁻ are significant, we can look for similar trends in charge states 8-, 6- and 5- (Table 22).

For 6- and 5- we observe that CCS values are all grouped to similar values, from which we expect a similar conformation in gas phase for all the ions. This would suggest that either no memory of conformation in solution is retained in gas phase or no conformational differences are present initially.

Oppositely, for 8- we have a more compact CCS for MN4⁸⁻ and MN19-A⁸⁻, between 1158-1163 Å², and a significantly extended CCS for OR8-A⁸⁻ and OR7-A⁸⁻, between 1344-1356 Å². This divergence is compatible with the lower base pairs content of the unfolded OR8-A and OR7-A in solution. In support of this hypothesis, we have the absent UV melting and the missing peaks of stem1 nucleobases (G4,31*) in ¹H-NMR spectra, for (see UV and NMR pp.113,115). Therefore, the Coulomb repulsion is probably enough to extend OR8-A⁸⁻ and OR7-A⁸⁻ to higher values of CCS. Conversely, MN4⁸⁻ and MN19-A⁸⁻ have more base pairs (attractive forces) that oppose to the Coulomb repulsion, resulting in lower CCS values for MN4⁸⁻ and MN19-A⁸⁻ compared to OR8-A⁸⁻ and OR7-A⁸⁻. This suggests that these base pairs are conserved in the gas phase.

5.6.8 IM-MS of quinine complexes

Comparing the CCS distributions of unbound aptamers with their complexes, we notice a double distribution for quinine complex of MN19-A⁷⁻, OR8-A⁷⁻ and OR7-A⁷⁻, an intense one between 984 and 996 Å², and a minor one between 1098-1116 Å² (Figure 77). We can attribute the main distribution to a population where the dA overhangs (at the unpaired 3',5'-ends) compact during the transfer in gas phase^[64]. The second distribution to a minor population that keeps a conformation more similar to MN4+Q⁷⁻, alike to what we concluded for free-aptamers low activation (Figure 76).

Interestingly, we notice that quinine binding make visible the two distributions for OR7-A+Q⁷⁻, whereas the second distribution is not well resolved for OR7-A⁷⁻.

Based solely on IM-MS results at low pre-IMS activation we can differentiate MN4⁷⁻ (1036 Å²) from the others dA constructs, because they appear to be more compact with CCS values between 984 and 996 Å². For MN19-A⁷⁻, OR8-A⁷⁻ and OR7-A⁷⁻ and relative quinine complexes we observe two distributions and a minimal increase in CCS likely due to the presence of ligand.

These findings at low pre-IMS activation, taken alone, do not suggest a significant change upon binding according to the "adaptive binding" mode commonly attributed to aptamers, and cocaine binding aptamers in this case.

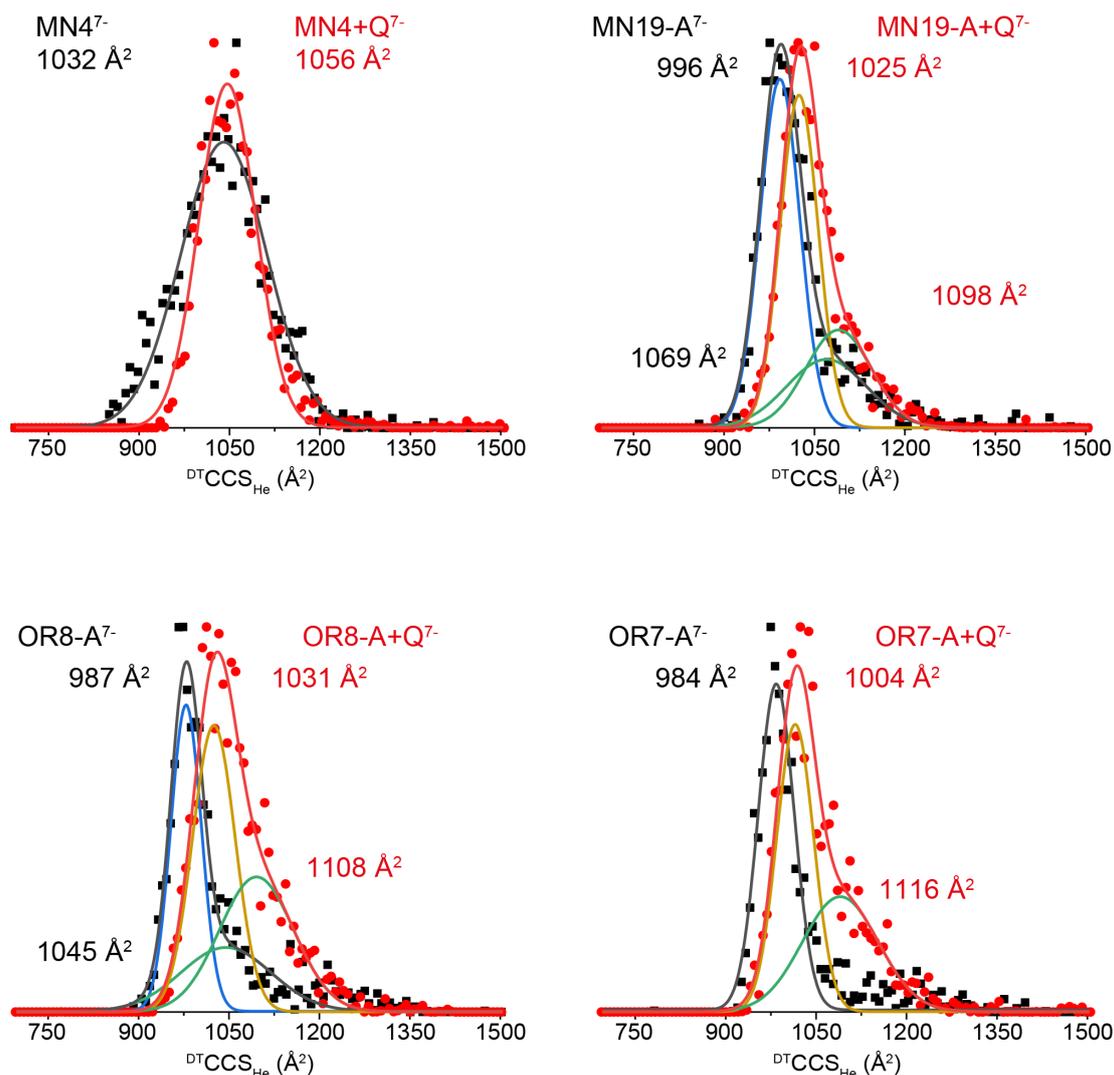


Figure 77 - Normalized CCS distributions of CBA+dA⁷⁻ and quinone complexes at low pre-IMS activation (fragmentor 350V). Two distributions are visible for MN19-A⁷⁻ and OR8-A⁷⁻ (**blue** and **green**) and for MN19-A⁷⁻, OR8-A⁷⁻ and OR7-A⁷⁻ quinone complexes (**yellow** and **green**). The minor distribution in **green** is attributed to a folded complex alike to MN4⁷⁻ and MN4+Q⁷⁻, whereas **blue** and **yellow** are compact populations due to dA overhangs (see text).

However, comparing the CCS distributions at 600V of MN4⁷⁻ and CBA-dA⁷⁻ (Table 22), we think that base pairs are preserved in gas phase. If this is true we can discriminate the sequences over their stability in gas phase of these ions, by observing how each sequence reacts to the pre-IMS collisional activation.

5.6.9 Collisional induced unfolding of MN4⁷⁻ and CBA+dA⁷⁻

To check if the IM-MS preserves a memory of the fact that MN4 is folded, MN19-A is partially folded in solution whereas OR8-A and OR7-A are unfolded, we used the IM-method “collision-induced unfolding” (CIU - p.22). If a memory of the conformations in solution is preserved in

gas phase, we expect a different sensitivity of each sequence to pre-IMS activation. We performed these experiments on MN_4^{7-} and $CBA+dA^{7-}$ by varying the pre-IMS activation (fragmentor voltages between 350 and 600V by steps of 10 V).

Charge state 7- of aptamer MN_4 is used as main example: its CCS distribution change completely from a single ensemble at 1037 \AA^2 to one at 1197 \AA^2 (Figure 66). However, at the same voltages MN_4^{7-} still undergoes a progressive NH_4^+ declustering (lost as NH_3) (Figure 78). Thus, collisional activation causes both NH_4^+ adduct removal and the extension of MN_4^{7-} . We recall that incomplete declustering at low activation energies is typically more acute in DNA and RNA than for proteins of the same size, because of the physiological counterion cloud that surrounds nucleic acids (i.e. phosphate backbone) [156].

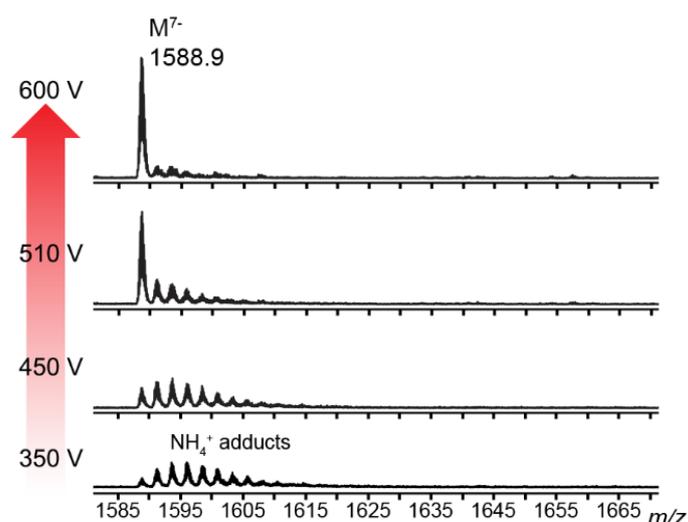


Figure 78 - Collisional activation on MN_4^{7-} cause NH_4^+ declustering, that accompan CCS extension.

When building a CIU plot for MN_4^{7-} we can integrate the signals of either the declustered ion or the entire NH_4^+ adduct distribution (Figure 79). The 2D-plot on MN_4^{7-} without NH_4^+ with false colour rendering between 350 and 450V (intensities normalized per voltage segment) exhibits a dotted distribution due to the low signal-to-noise ratio of declustered MN_4^{7-} at low voltage. If one considers the entire NH_4^+ adduct distribution (Figure 79A-right), the 2D plot has a more uniform rendering. However, this latter choice also results in a shifted transition point, because aptamers with adducts are more compact. This shift is clearer in the breakdown diagrams: the transition voltage is shifted at 529V if one account for NH_4^+ adducts (compared to 521 V for the fully desolvated MN_4^{7-}). We conclude that, for the CIU plot to render only the unfolding effect, it is better to consider just the declustered ion.

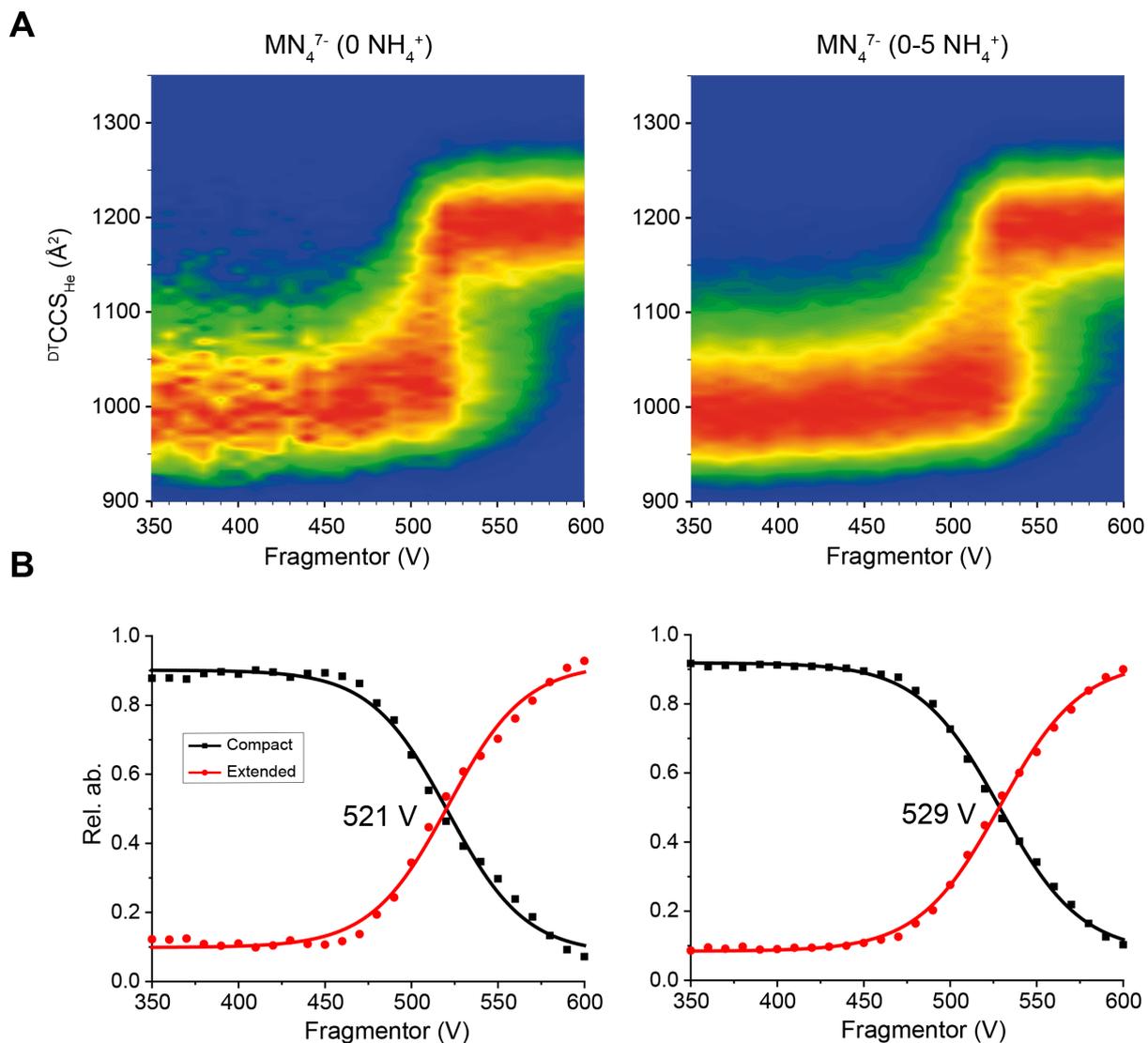


Figure 79 - **A.** CIU plots considering declustered MN_4^{7-} (left) or with 5 NH_4^+ adducts (right). **B.** Breakdown diagram from the CIU plots. A sigmoidal fit gives 521 V as CCS mid-point transition.

However, the information of NH_4^+ declustering is missing on the single CIU plot in itself, as well as on the derived breakdown diagram. In Figure 80 we propose a way to depict together the dissociation of NH_4^+ and the extension of MN_4^{7-} .

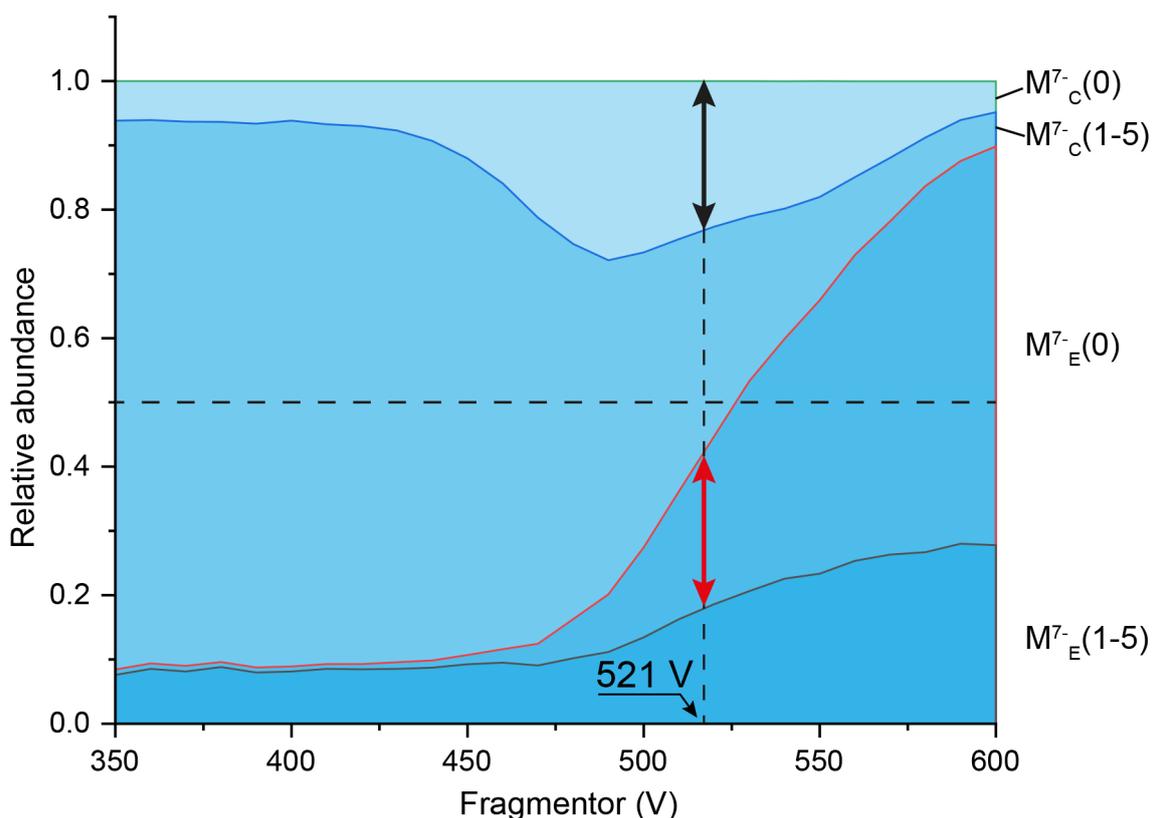


Figure 80 – CID/U breakdown curves of $MN4^{7-}$ (M^{7-}). Relative abundance of each population is reported for each voltage (method at p. 48). Compact and extended form are marked as C and E, and the number of adducts are reported in brackets. CCS transition considering only desolvated ion is marked with a vertical dotted line, transition considering the NH_4^+ series cross the 0.5 horizontal dotted line (from Figure 79).

We underline that no background subtraction has been done, resulting in a significant portion of extended $MN4^{7-}$ (M^{7-E}) at low voltages, as well as compact $MN4^{7-}$ (M^{7-C}) at high voltages. This is due to the background currents at those arrival times.

This CID/U breakdown curve is suited to describe all the species involved in both unfolding (U) and dissociation (D), and allow us to extract:

1. the removal of NH_4^+ adducts from compact $MN4^{7-}$ (M^{7-C}) starts at 425 V, and precedes the formation of extended $MN4^{7-}$ (M^{7-E}) which starts at 450V;
2. The CIU considering the NH_4^+ adducts is shifted to 529 V because of the intensity contribution of M^{7-E} (1-5);
3. M^{7-C} (0) and M^{7-E} (0) intensities are equal at 521 V, the transition voltage obtained from CIU of only declustered ions (0 NH_4^+);

More generally, we can follow the reactions (in gas phase), dissociation of NH_4^+ and conformational extension, reported in the scheme of Figure 81.

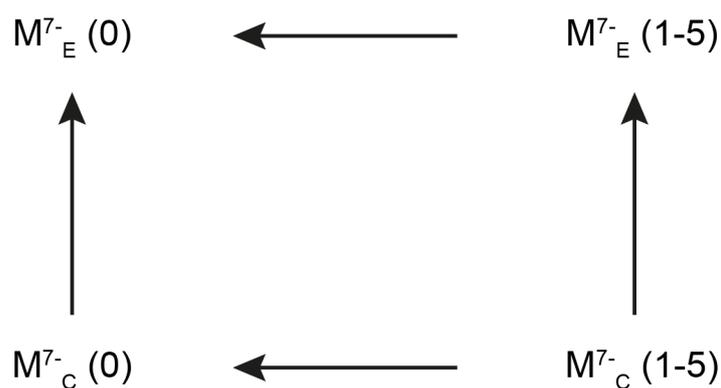


Figure 81 - Reaction scheme for NH_4^+ desolvation and unfolding (Compact, Extended) on MN4^{7-} (M^{7-}). $\text{M}_C^{7-}(\mathbf{0}) \rightarrow \text{M}_E^{7-}(\mathbf{0})$ is the extension followed in CIU in Figure 79A. These are reactions in gas phase, thus the use of single arrows because are not equilibria.

The aptamers with fewer base pairs but same total length (MN19-A, OR8-A and OR7-A) show a similar removal of NH_4^+ adducts accompanied by the extension of the charge state 7- (Figure 82), and a two-state CCS transition with no intermediate population (Figure 82A and B).

Interestingly, the aptamers are in the same relative kinetic stability order in the gas phase as the stability order in solution (UV-meltings at p.113 and 188). This suggests the preservation of base pairing interaction from solution to gas phase and allow to discriminate MN4 from the three constructs MN19-A, OR8-A and OR7-A.

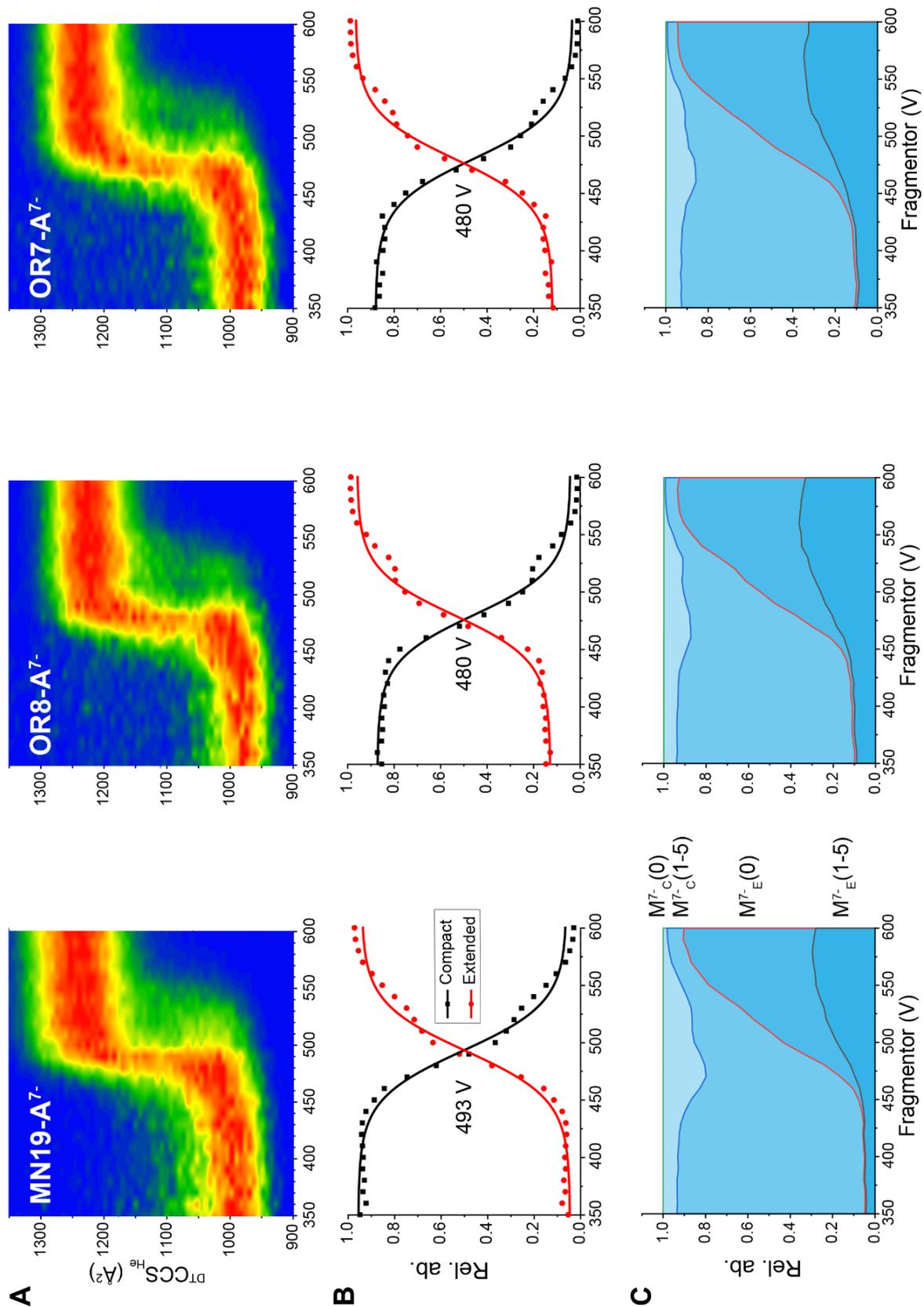


Figure 82 - **A.** CIU plots of declustered MN19-A⁷⁻, OR8-A⁷⁻ and OR7-A⁷⁻; **B.** CIU breakdown diagram of M(0); **C.** CID/U breakdown curve of each ion shows the loss of NH₄⁺ adducts during the transition. Boltzmann fit fails for these data sets. Transition points (crossing of follow line with $\pm 5V$ estimation) are reported with double arrows on CID/U breakdown curves.

The CIU plots of Figure 79 and Figure 82 highlight that it is not possible to differentiate whether aptamers were folded or unfolded in solution based on their gas-phase unfolding pathway: folded aptamers (MN4, MN19-A) unfolded ones (OR8-A, OR7-A) unfold via a two-state transition. In addition, the orders of magnitudes of the collision cross sections of the states are similar. However, we can differentiate the sequences based on the observed relative gas phase stabilities. The breakdown diagrams, from CIU plots, show that the lower is the base pair content, the lower is the transition voltage.

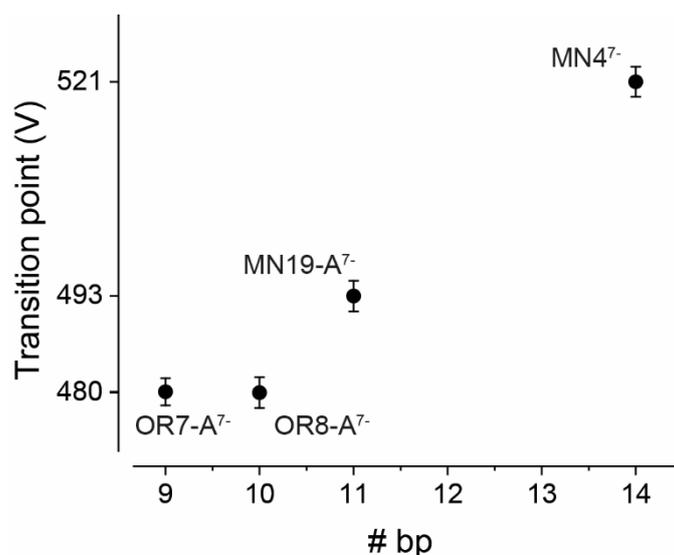


Figure 83 - The CIU transition points of charge state 7- increase with the base pair content of the ion. Note that all ions have the same chain length of 36 nt (Figure 71 *Errore. L'origine riferimento non è stata trovata.*).

We therefore propose the following mechanism:

- I. Aptamer structures undergo a compaction during their transfer in gas phase (ESI process). We suggest by phosphates auto-solvation alike to B-DNA helices, previously reported from our lab ^[64];
- II. Collisional activation induces the unfolding and dissociation of base pairing.



Figure 84 – Putative unfolding mechanism for cocaine binding aptamer ions. After compaction upon ESI (alike B-DNA helices), pre-IMS collisional activation breaks the base pairing pattern in a unique step.

5.6.10 CIU of aptamer-quinine complexes

Following the same approach for charge state 7- of the aptamers alone, we also carried out the CIU experiments in presence of one equivalent of quinine for MN4 (10 μ M DNA and 10 μ M quinine). This little shift is due to the lower intensity of MN4⁷⁻. However, we do not think that is significant, especially if the others sequences would not change.

For MN4+Q⁷⁻ (ML⁷⁻ in Figure 85) we observe that the pre-IM collisional activation induces **I)** NH₄⁺ dissociation from both M⁷⁻ and ML⁷⁻ and **II)** the ligand dissociation (conversion of ML⁷⁻ to M⁷⁻). These two types of information are missing on the CIU plot reconstructed on the ML⁷⁻ ion (Figure 85B). From the derived breakdown diagram (Figure 85C) we have a CCS transition at 574 V that suggests a partial extension of MN4+Q⁷⁻ just before complete dissociation, and a stabilization compared to MN4⁷⁻ without ligand.

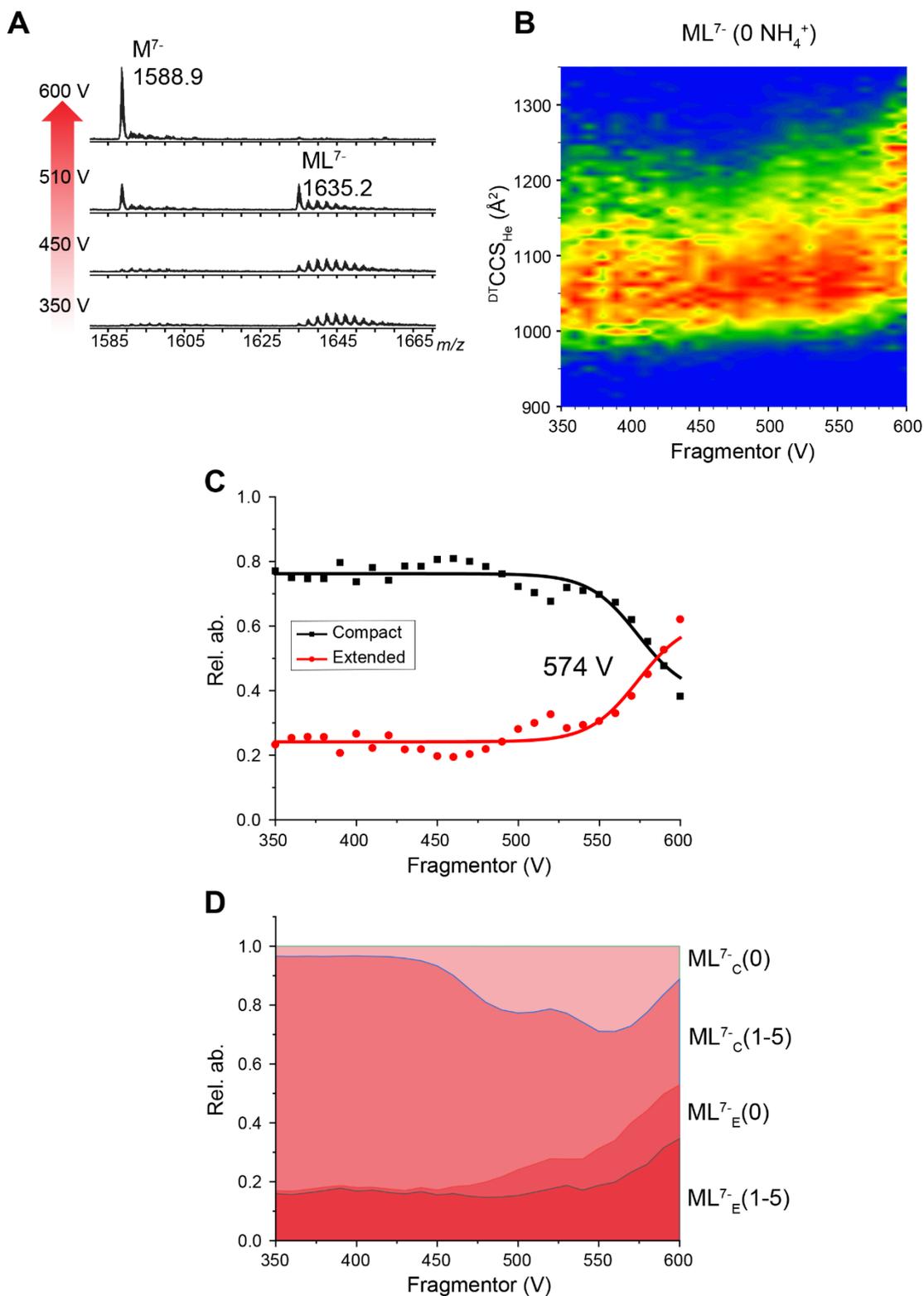


Figure 85 - **A.** Collisional activation on MN_4+Q^{7-} cause NH_4^+ removal and ligand dissociation; **B.** CIU plot considering declustered MN_4+Q^{7-} ; **C.** The sigmoidal fit of intensity points diagram gives 574 V as CCS transition point (filled arrows); **D.** CID/U diagram of $MN_4+Q^{7-}(ML)$. Compact and extended form are marked as C and E, and the number of adducts are reported in brackets.

However, this still does not represent all of the concomitant events, as the ligand dissociation is not shown. We thus constructed the CID/U breakdown diagram of $MN4^{7-}$ and $MN4+Q^{7-}$ (Figure 85D) to portray together the dissociation of NH_4^+ , ligand, and the unfolding of all ions (Figure 86).

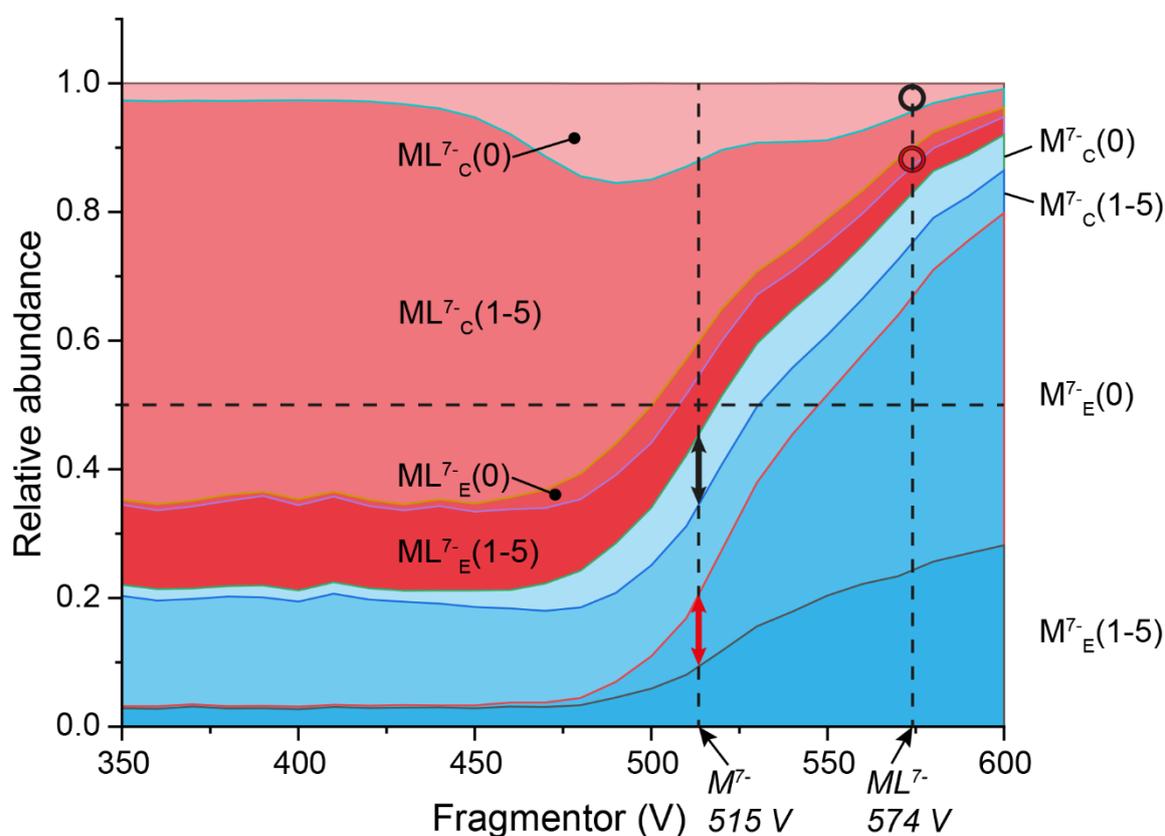


Figure 86 - CID/U breakdown diagram of $MN4^{7-}$ (M - cyan) and $MN4+Q^{7-}$ (ML - red). Compact and extended form are marked as C and E, and the number of adducts are reported in brackets. CCS transition point for M^{7-} and ML^{7-} are reported as vertical dotted lines.

The CID/U breakdown diagram of $MN4^{7-}$ (M^{7-}) and $MN4+Q^{7-}$ (ML^{7-}) gives a global vision on how the relative abundances vary with the “pre-IMS” activation (fragmentor voltage):

- I. At low voltages (350-475 V) the ratio M/ML remains substantially unchanged. Only partial NH_4^+ declustering (i.e. $ML_c(1-5)$ to $ML_c(0)$) occurs between 450 and 475 V.
- II. From 475 V on, the ligand dissociation becomes substantial, and free-aptamer M is progressively populated, mainly in the extended form M_e .
- III. The range from 475 to 600 V, ligand dissociation and NH_4^+ declustering occurs at the same time. For this “middle” ions of the plot may result graphically distorted over the total intensity, as is the case of ML_e and M_c (relative variations remains unaltered).

Thus we conclude that: **a.** dissociation (i.e. NH_4^+ , ligand) and extension occur in parallel; **b.** ML^{7-} ($MN4+Q^{7-}$) does not extend significantly before dissociation; **c.** M^{7-} ($MN4^{7-}$) extends following the dissociation of NH_4^+ .

The scheme in Figure 87 summarizes the reactions that may occur between M^{7-} and ML^{7-} (NH_4^+ solvated or declustered) ions.

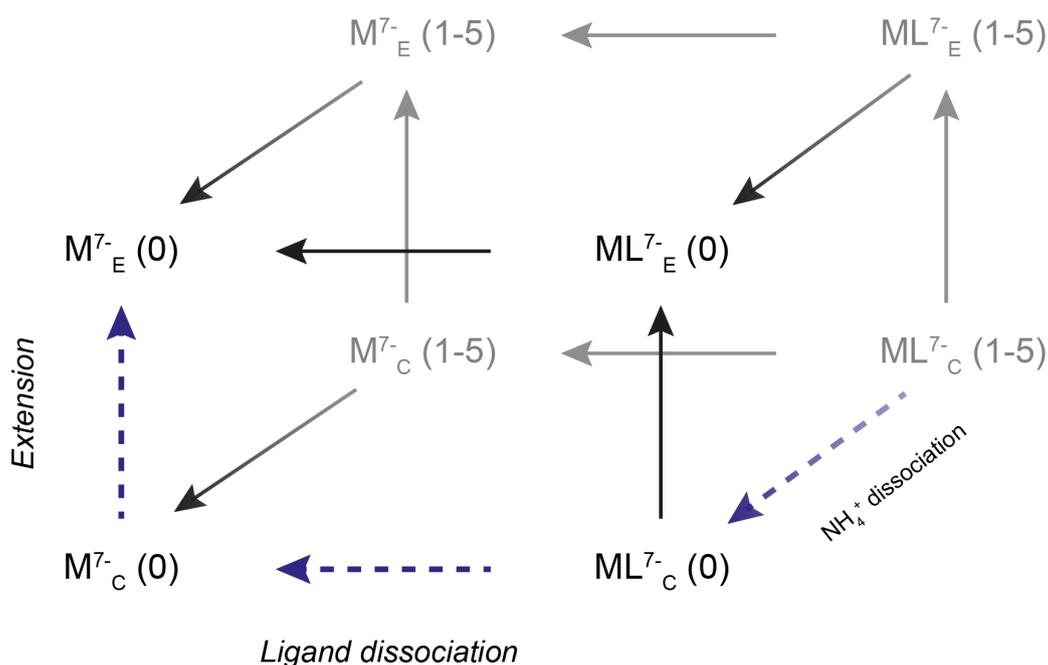


Figure 87 - Reaction scheme of dissociation and extension of ions $MN4^{7-}$ and $MN4+Q^{7-}$. The most probable events, deduced from CIU and CID/U are marked with **blue** dashed arrows.

Observing the intensities distributions (Figure 86) we propose the mechanism highlighted in blue in Figure 87. $ML_C^{7-}(1-5)$ loses its NH_4^+ adducts and dissociate, without showing any extension. Finally the $M_C^{7-}(0)$ populated this way, undergoes the extension observed in CIU plot and CID/U diagram (Figure 86). The CCS transition at 574 V for ML_C^{7-} , in Figure 85D, could represent a useful information to differentiate M^{7-} and ML^{7-} in terms of gas phase stability.

With this reaction scheme, one can interpret CIU plots and CID/U diagrams of MN19-A, OR8-A and OR7-A in presence of Q (Figure 88, Figure 89 and Figure 90). We know that these aptamers have a lower affinity for quinine, because of the missing base pairs at 3',5'-end (stem1) compared to $MN4^{[88]}$. Thus, the complex abundance is lower (red areas of CID/U diagrams - Figure 88C, Figure 89C and Figure 90C). The ML^{7-} CIU plots are difficult to interpret because quinine dissociation is not represented (false rendering from the intensity normalization and the concomitant ligand dissociation), whereas we included this information into CID/U plot. We see on the CID/U breakdown diagrams in diagrams (Figure 88C, Figure 89C and Figure 90C) that the intensity of ML^{7-} is significantly lower compared to M^{7-} . The "two overlaid CIU plots" is a signal-to-noise problem.

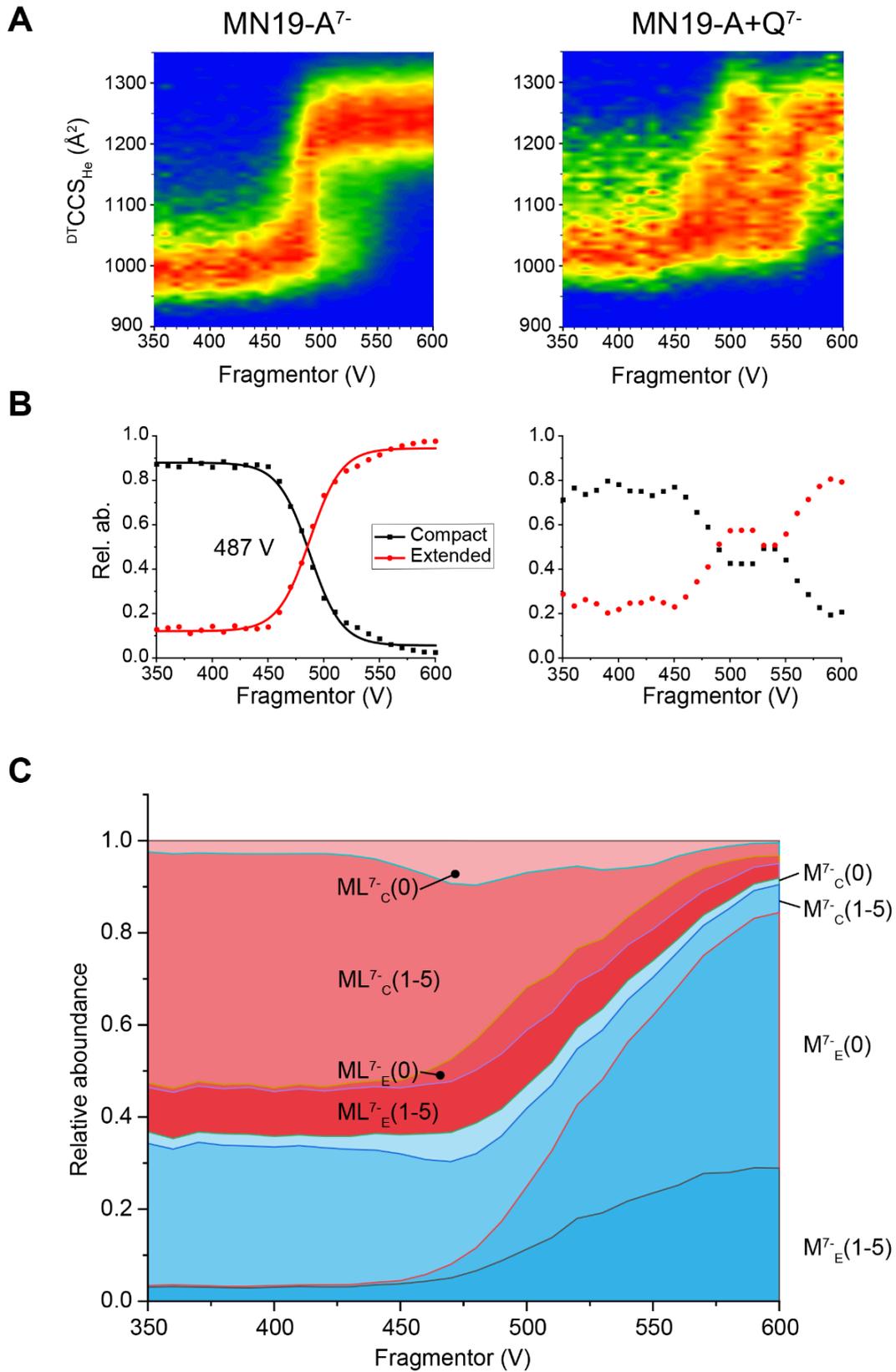


Figure 88 – Ion extension represented as **A**. CIU plot of declustered MN19-A⁷⁻ (M^{7-}) and MN19-A+Q⁷⁻ (ML^{7-}), and relative **B**. breakdown plots (as relative abundances of extended and compact form); **C**. CID/U diagram for MN19-A⁷⁻ and MN19-A+Q⁷⁻; (DNA 10 μ M, quinine 10 μ M, NH₄OAc 25 mM).

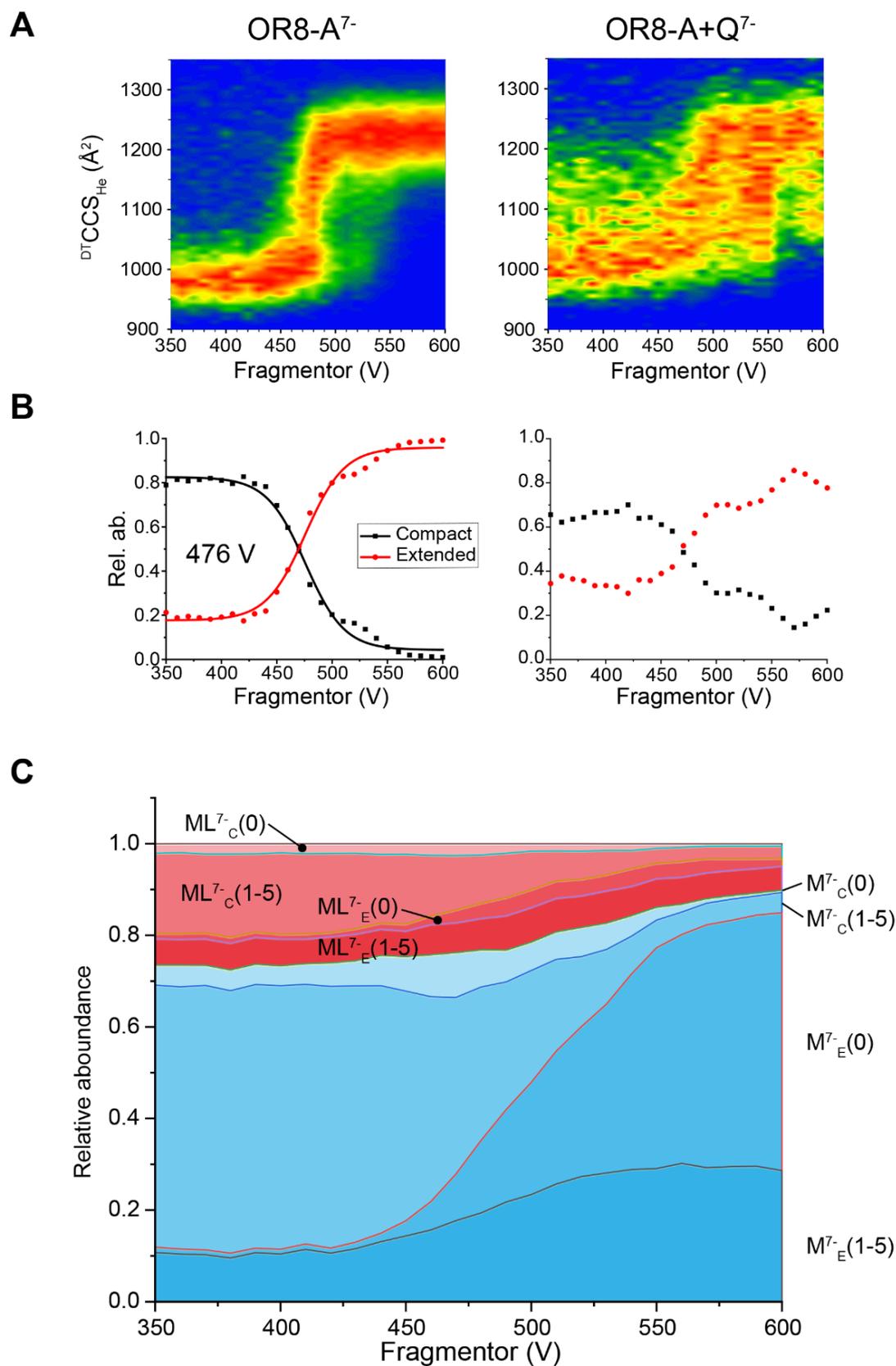


Figure 89 – Ion extension represented as **A**. CIU plot of declustered OR8-A⁷⁻ (M⁷⁻) and OR8-A+Q⁷⁻ (ML⁷⁻), and relative **B** breakdown plots (as relative abundances of extended and compact form); **C**. CID/U diagram for OR8-A⁷⁻ and OR8-A+Q⁷⁻; (DNA 10 μ M, quinine 20 μ M, NH₄OAc 25 mM).

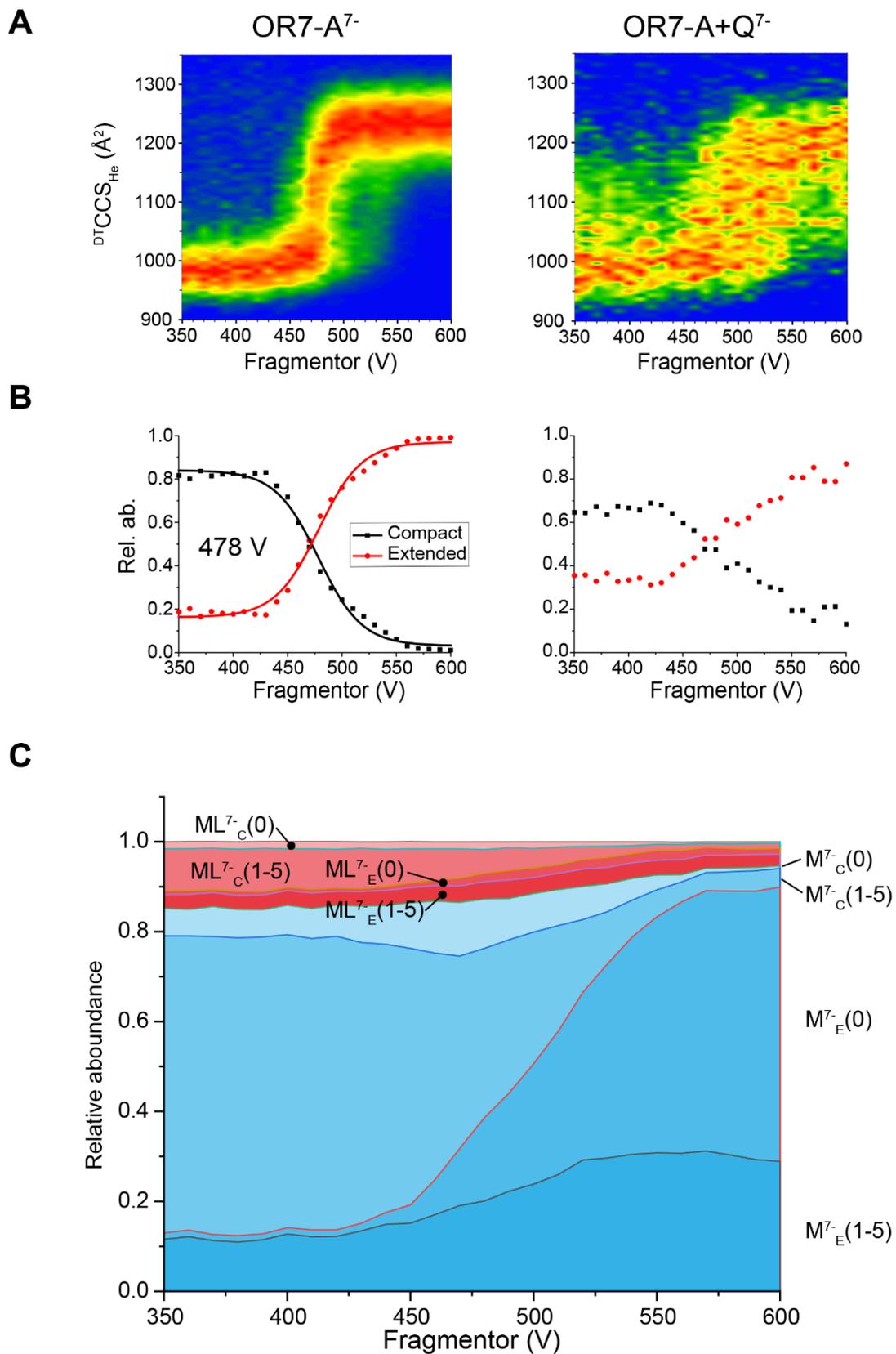


Figure 90 - Ion extension represented as **A**. CIU plot of declustered OR7-A⁷⁻ (M⁷⁻) and OR7-A+Q⁷⁻ (ML⁷⁻), and relative **B**. breakdown plots (as relative abundances of extended and compact form); **C**. CID/U diagram for OR7-A⁷⁻ and OR7-A+Q⁷⁻; (DNA 10 μ M, quinine 20 μ M, NH₄OAc 25 mM).

Even if acquired in excess of ligand, the very low signal intensity of quinine complex for MN19-A, OR8-A and OR7-A prevents us to say if free-aptamer have a different kinetic stability and quinine complex.

Here we underline a limitation of our instrument design: we can't m/z -select prior to the drift tube. Consequently, we encounter a misleading CIU contour plot for ML^{7-} because signal intensity varies. Additionally, without m/z -selection we cannot be certain if the complex dissociates before extension, generating a compact free-aptamer (filled arrows - Figure 87). We cannot exclude *a priori* that dissociation may occur after extension, generating an extended aptamer (dotted arrows - Figure 87). Actually, none of the intermediates ($M_{compact}$ or $ML_{extended}$) is very abundant at any of the intermediate voltages. Further investigations are needed using a different instrumental design that allows m/z -selection (to avoid dissociation) before activation and IMS.

5.7 Conclusions

We studied a series of RNA and DNA aptamers and investigated whether the conformational changes upon ligand binding are transposed into the gas phase when probed by ion mobility spectrometry.

For the four RNA aptamers of the quantitative MS part, (MGA, 1FMN, 1NEM and 1TOB) we observed similar CCS distributions for both free and bound aptamers. This may suggest that either aptamers do not significantly change their shape upon ligand binding in solution, or that we observe ions that went through a compaction during the transfer in gas-phase. In this regard, our lab previously showed that DNA anions, at low charge states, undergo a compaction upon ESI and transfer in gas phase^[64]. A similar phenomenon is probably responsible for the impossibility to differentiate the single unbound aptamer from its complex solely on their ion mobilities (at low pre-IMS activation).

High pre-IMS activation induces only a small compaction on the CCS distributions of both free and bound aptamers. We suggest that the small CCS difference (increase) observed for these aptamers upon binding (Table 21) are compatibles with ligand binding modes: via intercalation for MGA^[32] and 1FMN^[35a], and encapsulation for 1NEM^[82] and 1TOB^[83a]. These binding modes are likely to induce small change in aptamer conformation. Phenomenologically, we can describe the free-aptamer alike a sphere that increases its volume to accommodate the ligand.

Only between 1FMN bound aptamer and its Mg^{2+} adducts, we observe a small but significant difference in CCS distributions, that support the hypothesis of different conformational ensembles for desolvated and Mg -bound complex. The Mg^{2+} ions may contribute to shield the negative charges on the glucophosphate backbone, reduce the Coulomb repulsions, and facilitate some chain contacts within the secondary structure of the bound 1FMN.

A clearer example of RNA conformational change, in presence of metal additives, is tetracycline RNA aptamer (TCA). We observed a pronounced difference between free TCA

and its metal bound complex. The tertiary contacts mediated by Mg^{2+} in bound TCA give two extended populations corresponding to the partially and fully folded complex^[140]. From IM-MS we observe three distinct populations: a compact CCS population for free TCA and two CCS populations for its complex, as function of the Mg^{2+} and K^+ adducts (Figure 62). This suggests the absence of tertiary structure in free TCA that may compact upon ESI and transfer in gas phase^[64], whereas the tertiary contacts prevent the compaction in the tetracycline complex. These observations are in agreement with the TCA binding scheme documented in literature at low ($\sim 200 \mu M$) Mg^{2+} concentrations (Figure 60)^[140].

For DNA cocaine binding aptamers the conformational changes upon quinine binding are widely reported in literature^[35b, 79c, 87-88, 144a, 145, 157]. In the experimental design, it was important to add dA overhangs at 3',5'-ends to have ions of comparable degrees of freedom, without altering the core secondary structure.

At low collisional activation, we observed similar CCS distributions for both free and quinine-bound aptamers, similarly to what observed for MGA, 1FMN, 1NEM and 1TOB RNA aptamers. However, comparing different aptamers within the series, we noticed that known folded sequences (MN4, MN19-A) and unfolded ones (OR8-A, OR7-A) all show a significant compaction, which depends on the charge state (5.6.6 – p.118). Charge state 7- is the most interesting one: at low activation MN19-A⁷⁻, OR8-A⁷⁻, OR7-A⁷⁻ aptamers have a main CCS distribution that suggests a conformation more compact to MN4⁷⁻. A second CCS distribution, comparable to MN4⁷⁻, is observed and would suggest a minor fraction of folded aptamer. We believe that dA overhangs may help in 3',5'-ends stabilization but compacts more in gas phase than MN4 unmodified ends, in light of the absent (or minimal) dA-dA base pairing. Speculations on the characteristic NMR imino resonances of stem1 and junction nucleobases in MN19-A, OR8-A (5.6.5.2 – p.115) support this thesis, but further two-dimensional NMR investigation in our conditions are required.

Charge state 7- shows equally extended CCS distributions at high pre-IMS activation, suggesting the same extended conformational ensemble for all aptamers. Unfortunately, in quinine complexes the ligand is completely lost at maximum pre-IMS activation, so no comparison can be made. However, aptamers can be differentiated following the extension pathway of charge state 7- (MN4⁷⁻, MN19-A⁷⁻, OR8-A⁷⁻, OR7-A⁷⁻). Gradually increasing the pre-IMS collisional activation, we revealed a common two-state transition, which occurs at lower voltages if the number of base pairs is lower. This tells us that base pairing interactions prior to ESI are preserved in the gas phase.

It was also possible to observe a different transition voltage for MN4+Q⁷⁻ from free aptamer (Figure 80 and Figure 85). However, this was not possible for the other complexes, because of the low quinine affinity that accompanies the removal of base pairs in the structure (Figure 88, Figure 89 and Figure 90). We suggest to test higher excess of quinine or ligands with higher affinity that may not undergo complete dissociation, or to work with a different instrument allowing precursor ion m/z selection. The absence of mass selection before the IM drift tube obliges us to activate all ions generated from the source. This could be a disadvantage when

fragile ions dissociate during the pre-IMS activation, because other reactions take place in gas phase during the “CIU” experiment in the same energy range. In our case, we couldn’t address unambiguously the sequential order of dissociation/extension for quinine complexes. In this regard, further investigations with instruments capable to perform m/z -selected IMS would be useful to assess if quinine-bound aptamers (i.e. MN_4+Q^{7-}) do extend before dissociation or not.

Finally, we also highlight a limitation in the CIU plots. The intensity normalization per segment could provide a false rendering because does not account for the variations of global signal intensity, especially when collisional activation provoke dissociation events that vary the relative abundance of different ions. This was the case of cocaine binding aptamers’ adducts and aptamers’quinine complexes. To include this information, we propose a new plot for breakdown curves called “CID/U” (Figure 80).

In summary:

- 1- We observed only small ΔCCS upon ligand binding on RNA aptamers as MGA, 1FMN, 1NEM and 1TOB. Ligand may be associated to small conformational change (intercalation, encapsulation) to be detected in gas phase.
- 2- Aptamer complexes with a tertiary structure are likely to give conformational ensembles that prevent gas phase compaction. Thus we can differentiate unstructured (free) aptamers from their complexes by IM-MS.
- 3- Sequence modifications that do not alter aptamer’s core sequence are useful in comparing a series of similar aptamers.
- 4- CIU experiments can be used to discriminate a series of similar aptamers based on their different base pairs content (i.e.: CBA-dA).
- 5- We propose a new CID/U plot to show unfolding and dissociation events on the same graph.

6. General conclusions and perspectives

6.1 Quantitative MS

In chapter 4, we showed that NH_4OAc and TMAA electrolytes could be used in ITC and do not alter the binding capabilities of RNA aptamers in solution. For example, the thermodynamic parameters derived for the malachite green aptamer (MGA) are in line with the literature values in other non-volatile buffers. This suggests that the ligand binding mode remains unchanged (enthalpically driven). In support of this observation, the K_D values obtained in NH_4OAc and TMAA are comparable within the expanded uncertainty range (95% confidence interval).

More importantly, the K_{DS} derived from native MS titrations are also comparable to those derived from ITC within the expanded uncertainty range. These observations prove that ITC and native MS, when carried out in similar experimental conditions ($[\text{RNA}]$, $[\text{electrolyte}]$, T), report comparable K_{DS} . In this work we detected K_{DS} from 30 μM , and estimated that the lower limit is ~ 50 nM.

In addition to reliable K_D determination, aptamers analysis via native MS can be of great interest to rapidly screen aptamer binding. The stoichiometry of ligands and additives (M^{n+} , co-factors) can be unambiguously assigned, and this is important to guide the choice of data fitting models in other biophysical methods on the same molecular system.

6.1.1 The choice of the electrolyte

The outcome of the ESI process (charge state distribution, peak intensities) is very analyte dependent. The choice of the supporting electrolyte (NH_4OAc , TMAA) can also influence some aspects of this process and change the appearance of the MS spectrum. The charge state distribution changes with the nature of the supporting electrolyte and its concentration. While ammonium salts favour higher charge states distribution (especially at high concentrations) ^[148], lower charge states are favoured in alkyl-functionalized ammonium salts, like TMAA ^[125].

TMAA provides also more spaced and less abundant adducts compared to NH_4OAc , resulting in better signal-to-noise ratios. This reduces the uncertainty on peak integration when quantification is required, and Mg^{2+} adduct series are more easily revealed. On the 1FMN aptamer, we observed a higher number of Mg^{2+} adducts and higher ligand affinity (lower K_D) compared to the same conditions in NH_4OAc . Our interpretation is that NH_4^+ may compete with Mg^{2+} for RNA binding, whereas the bulky TMA^+ cations does not.

For these reasons, we conclude that TMAA is adapted to study aptamers and their interactions with ligands and additives such as Mg^{2+} . We recommend to explore TMAA as supporting electrolyte for aptamer analysis, and nucleic acids in general, but also to check that the molecular process is identical to that in NH_4OAc .

6.1.2 Relative response factor correction

The relative response factor correction on the analysed RNA aptamers was relatively small. This means that either the conformational change upon ligand binding is not sufficient to significantly change the ESI response of aptamers and complexes, or that even if the free and bound aptamers have different conformation in the bulk solution, they converge to more similar ones before their final charging/declustering. Therefore, the assumption of equal response factors (ratio of response factor equal to 1) is justified. Indeed, when ranking aptamer or ligand series, what matters is to have significant differences in the order of magnitude of K_D values.

MS peaks' signal-to-noise ratio influence the calculation of relative response factors. These factors are themselves source of uncertainty that sums to the integral uncertainty (i.e.: intensity ratios) in K_D determination. Therefore, is highly recommended to use MS peaks with the best signal-to-noise ratio for both relative response factor calculation and the eventual concentration correction.

6.1.3 One point measurement, full titration and data treatment

Whether relative response factor correction is taken into account or not, K_D s could be determined with a single point measurement. This method have minimum time and sample consumption but should be chosen on a very well-known ligand-aptamer couple, to minimize the sources of uncertainty ([ligand], [aptamer], model assumption...). A full titration provides a larger set of information and lower uncertainty, compared to a single point measurement.

In general, on MS titration data elaboration we recommend to do a complete data treatment of each replica until non-linear fitting for K_D determination (method 1). This afford a better control on uncertainties associated on K_D s, the presence of outliers (or replicas) and the eventual relative response factors (if any).

6.1.4 Saturated binding and detection limits

However, in case the aptamer concentration was too high compared to the K_D (binding reaction too displaced toward the products), the uncertainty on K_D would be high as well. In this case, full titrations should be repeated at lower aptamer concentrations, to afford a reliable K_D determination. A disadvantage of aptamer dilution is the risk to reach the detection limit of the instrument: while equilibrium is displaced towards the reactants, the overall signal response is also lowered. It is the case of aminoglycosides aptamer, where the dilution below 50 nM give signal loss on our instrument.

6.2 Ion Mobility Spectrometry

6.2.1 Tertiary structural elements prevent RNA compaction upon ESI

The comparable CCS distributions observed for free and bound form of MGA, 1FMN, 1NEM and 1TOB aptamers, suggest that either their conformational change upon binding in solution is not significant, or that ions went through a compaction during ESI. These four RNA aptamers have a hairpin-loop secondary structure and ligands are bound via intercalation or encapsulation, with very little conformational change. We think that RNAs without a rigid secondary (e.g.: G-quartets)^[122] or a tertiary structure that preserves their native conformation during ESI and transfer in gas phase, are likely to compact with the mechanism of phosphodiester pairing (auto-solvation) recently documented in our lab for low charged anions of DNA double helices^[64]. Only the 1FMN complex in presence of Mg^{2+} gave a small but significant difference in CCS distributions. We think that Mg^{2+} may stabilize a conformational population that resists gas phase compaction.

In support to this hypothesis for 1FMN complex, we observe the conformational change of tetracycline RNA aptamer (TCA) upon ligand binding. In presence of ligand, $MgCl_2$ and KCl (100 μM each), we discern three conformational populations: the unbound aptamer, the partially folded complex and the fully folded TCA, in line with previous in solution studies^[140]. The CCS values of unbound TCA are significantly lower, indicating a more compact unbound aptamer compared to the complex (both partially and fully folded). This would appear counterintuitive to what we expect from unfolded RNAs in solution, but if we account for gas phase compaction we have a meaningful picture: The tertiary contacts formed with ligand and metal cations (Mg^{2+}) prevent the full compaction of the TCA complex, whereas the unbound aptamer undergoes a full compaction alike to what described for DNA double helices. This explains the three population observed and their relative order in terms of CCS.

Interestingly, we observe that TCA was unable to bind its ligand in the absence of $MgCl_2$ and that its charge state distribution was much lower in the same supporting electrolyte (TMAA) without $MgCl_2$ and KCl.

6.2.2 Some secondary structure aptamers can be discriminated via IM-MS

The cocaine binding aptamer series is marked by differences in secondary structures. Their different sequence length obliged us to add some extra dA hangovers on the 3',5' ends to have ions of comparable numbers of degrees of freedom. At low pre-IMS activation, the charge state 7- allows to distinguish only the fully folded aptamer MN4 from the others. Shorter aptamers MN19-A, OR8-A and OR7-A all show a more compact CCS distribution.

Contrarily to aptamers in chapter 4, it was possible to discriminate cocaine-binding aptamers progressively increasing the pre-IMS activation: charge state 7- the four aptamers show a two-state transition, from their respective compact CCS distribution to a similarly extended one.

Interestingly, the relative order at which the transition occurs is correlated with the base pair content, and the same as observed in solution with UV-melting. We conclude that the nucleobase pairing is preserved during ESI and that the stability order reported in gas phase and in solution depicts the separation of base pairs.

Unfortunately, at low activation the quinine complexes show just a small increase of CCS with similar distributions to their free-aptamer, so we could not infer any conformational change. Moreover, only MN4 quinine complex was distinguishable with the increasing pre-IMS activation from its free-aptamer. The other complexes undergo dissociation are lost during collisional activation. We think that cocaine-binding aptamers do have different gas phase stabilities but this observation was possible only for MN4. On the other aptamers, the lower quinine affinity and the ligand dissociation, prevent to observe the extension to a population of higher CCS.

6.2.3 CIU/D breakdown curves

In chapter 5 we observed for the cocaine binding aptamers that ligands and volatile salts adduct could dissociate from the aptamer-ligand complexes in similar pre-IMS activation that cause unfolding. This results in a false rendering on conventional “CIU plots”^[158], wherein the normalization per segment does not account for the variation of global intensity. When the unfolding of fragile ions is studied at variable pre-IMS activation, we recommend to build full breakdown curves, including both CCS and m/z partitioning of the ion signal. In this work we called such diagram as “CID/U” breakdown curves and used color codes to distinguish each ion among the stacked breakdown curves.

6.3 Outlook

In this work we showed that aptamers and their complexes can be analysed in volatile electrolytes such as NH_4OAc and TMAA. What matters is the ionic strength provided by the electrolyte and the eventual presence of metal additives.

A valuable outlook of the first part this work would be to:

- Determine whether TMAA systematically favours the Mg^{2+} binding to RNA, by comparing the number of specific Mg^{2+} adducts visible in TMAA and NH_4OAc on 1FMN and other Mg^{2+} dependent aptamers. This work can be done using a control sequence, similarly to what described by Klassen ^[159] and Rabin ^[98];
- Extend the ITC and native MS comparisons to other aptamers of diverse structure and size, to strengthen or broaden our conclusions. In particular, we would be curious to test the response factors for other aptamer systems with large structural change and some difference in shape seen by ion mobility. The TCA systems could be a good candidate, but we would need a synthetic TCA sequence, to minimize salt contaminations.
- Competitive titrations methods could be envisaged to overcome the lower determination limit of 50 nM, and better estimate very low K_D values of complexes such as aminoglycosides and tetracycline aptamer.

A valuable outlook of the second part of this work would be to:

- Simulate the gas phase transfer of RNA double helices and hairpins, to understand if compaction upon ESI and gas phase transfer is applicable to such RNA structures, alike to DNA double helices.
- Expand the variety of secondary and tertiary structures to test the resistance of such nucleic acid structures capable to resist to compaction upon ESI. Some examples of aptamers with tertiary contacts are TPP ^[160], SAM ^[161] and guanine ^[162] aptamers domains of the relatives riboswitches.
- Test the gas phase stability of cocaine binding aptamers bound to other ligands than quinine, and eventually on other instruments with m/z -selection prior IMS (e.g. Waters Synapt series) to determine unambiguously if extension occurs before or after dissociation of the complex.
- Generalize the use of "CID/U" breakdown diagram to multiple ions with more than two CCS transitions. This could be particularly valuable for proteins.

7. Annexes

7.1 Material and Methods

7.1.1 R-script file inputs for ATD extraction

Ranges for Peak (fully desolvated only)		
(m/z)_i	(m/z)_f	Output file
879.8	882.4	001PeakT6.txt
876.8	879.4	002BackgroundT6.txt
<hr/>		
2048	2050.5	003PeakM6m.txt
2042.5	2045	004BackgroundM6m.txt
<hr/>		
2102.5	2105.5	005PeakM+L6m.txt
2042.5	2045	006BackgroundM+L6m.txt
<hr/>		
Ranges for Peak + 3 adducts		
(m/z)_i	(m/z)_f	Output file
2048	2080	003PeakM6m.txt
2015	2047	004estrattoBgM6m.txt
<hr/>		
2102.5	2134.5	005PeakM+L6m.txt
2015	2047	006BackgroundM+L6m.txt

Figure 91 - Example of *m/z* ranges for peak integration (MGA aptamer)

7.1.2 R script for Integrals from Agilent 6560

```
#####INPUT#####
#####

#you should set the work directory as the same of files:
#window on the right 1) find the directory; 2)More>set as working directory

#####Give the name of species
#Exemple:
#If one: Name_Species = "1NEM5m"
#If more: Name_Species = c("1NEM5m","dT6")

#these are just the labels of each integral.
#NOTA BENE: write them in the right order!
Name_Species = c("T6", "1NEM5m", "1NEM-neo5m")

#these are the patterns to insert: 1st - 2nd
#the script see them as couples...so the first file red with "pattern_file" - the second
with "pattern_background"
pattern_file = "Si"
pattern_background = "Bg"

#name of the output file
Name_File = "Intensities1NEM+0"
#click "Source" & be happy
#####

files = list.files(path= getwd(), pattern = pattern_file)
files_background = list.files(path= getwd(), pattern = pattern_background)

name = 1

for(i in 1:length(files)) {

  file = read.table(files[i], header=TRUE)
  background = read.table(files_background[i], header=TRUE)

  Intensity = sum(file$Int, file$Int.1, file$Int.2, file$Int.3, file$Int.4) -
    sum(background$Int, background$Int.1, background$Int.2, background$Int.3,
background$Int.4)
```

7.1.3 Python Script for integrals Thermo EXACTIVE and CIU/CID plots

```
testlist = list()

def PeakIntegrator(file,lomz,himz):

    mz = []
    intensity = []

    File = open(file)
    lines = File.readlines()
    size = len(lines)
    File.close()

    for j in range(size):
        splitline = lines[j].split()
        #print(splitline[0])
        mz.append(float(splitline[0]))
        intensity.append(float(splitline[1]))

    prange = []
    low = 0
    hi = 0
    n=-1
    for j in range(len(mz)):
        #print(mz[j])
        if(mz[j] >= lomz and mz[j] <= himz):
            #print(prange)
            prange.append(j)
            n=n+1

    low = prange[0]
    hi = prange[n]
    sumfrag = 0
    sumtotal = 0
    for j in range(low, hi):
        sumfrag = sumfrag+intensity[j]

    return sumfrag

def PeakSum(filelist,lomz,himz):# filelist = list containing file names+directory, lomz
= low range for sum, himz = high range for sum.
    for i in range(len(filelist)):
        holder = PeakIntegrator(filelist[i],lomz,himz)
        print(holder)
```

7.1.4 Sigma Plot Scripts for CIU

ATD to CCS conversion

```
jsv5G.;
'col(1)=tA distribution@600V'
'col(2)=counts of the distribution@600V'
'cell(3,2)=calculated CCS'
'cell(3,4)=peak center of the tA@390.5V'
'cell(3,6)=z'
'cell(3,8)=Mass of the ion'

'define  $\mu=(M*m)/(M+m)$ '
cell(3,10)=(cell(3,8)*4)/(cell(3,8)+4)

'define conversion factor -->  $a=(CCS/tA@390.5V)*(sqrt(\mu)/z)$ '
cell(3,12)=(cell(3,2)/cell(3,4))*((sqrt(cell(3,10))/cell(3,6)))

'CCS reconstruction -->  $CCS=tA*a*(z/sqrt(\mu))$ '
col(4)=col(1)*cell(3,12)*(cell(3,6)/sqrt(cell(3,10)))
col(5)=col(2)
```

ATDs intensities normalization segment per segment

```
jsv5G.,
col(3)=col(3)/(max(col(3)))
col(4)=col(4)/(max(col(4)))
col(5)=col(5)/(max(col(5)))
col(6)=col(6)/(max(col(6)))
col(7)=col(7)/(max(col(7)))
[...]
col(28)=col(28)/(max(col(28)))
```

	1	2	3	4	5	6
1	350.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	360.0000	6.7261	0.0000	0.0000	0.0000	0.0000
3	370.0000	13.8725	0.0000	0.0000	0.0000	0.0000
4	380.0000	20.5985	0.0000	0.0000	0.0000	0.0000
5	390.0000	27.3246	0.0000	0.0000	0.0000	0.0000
6	400.0000	34.4710	0.0000	0.0000	0.0000	0.0000
7	410.0000	41.1971	0.0000	0.0000	0.0000	0.0000
8	420.0000	47.9231	0.0000	0.0000	0.0000	0.0000
9	430.0000	55.0696	0.0000	0.0000	0.0000	0.0000
10	440.0000	61.7956	0.0000	0.0000	0.0000	0.0000
11	450.0000	68.5217	0.0000	0.0000	0.0000	0.0000
12	460.0000	75.6681	0.0000	0.0000	0.0000	0.0000
13	470.0000	82.3941	0.0000	0.0000	0.0000	0.0000
14	480.0000	89.1202	0.0000	0.0000	0.0000	0.0000
15	490.0000	96.2666	0.0000	0.0000	0.0000	0.0000
16	500.0000	102.9927	0.0000	0.0000	0.0000	0.0000
17	510.0000	109.7187	0.0000	0.0000	0.0000	0.0000
18	520.0000	116.8652	0.0000	0.0000	0.0000	0.0000
19	530.0000	123.5912	0.0000	0.0000	0.0000	3.9916e-5
20	540.0000	130.3173	0.0000	0.0000	0.0000	0.0000
21	550.0000	137.4637	0.0000	0.0000	0.0000	0.0000
22	560.0000	144.1897	0.0000	0.0000	0.0000	0.0000
23	570.0000	150.9158	1.3453e-5	0.0000	0.0000	0.0000
24	580.0000	158.0622	0.0000	0.0000	0.0000	0.0000
25	590.0000	164.7883	0.0000	0.0000	0.0000	0.0000
26	600.0000	171.5143	0.0000	5.3180e-6	0.0000	0.0000
27		178.6608				
28		185.1173				
29		191.5738				
30		197.5303				
31		203.4868				
32		212.7114				
33		219.8578				
34		226.6230				

Figure 92 - Axes pattern for CIU plot rendering in SigmaPlot 12.5 (filled contour plot).

7.1.5 DynaFit Scripts

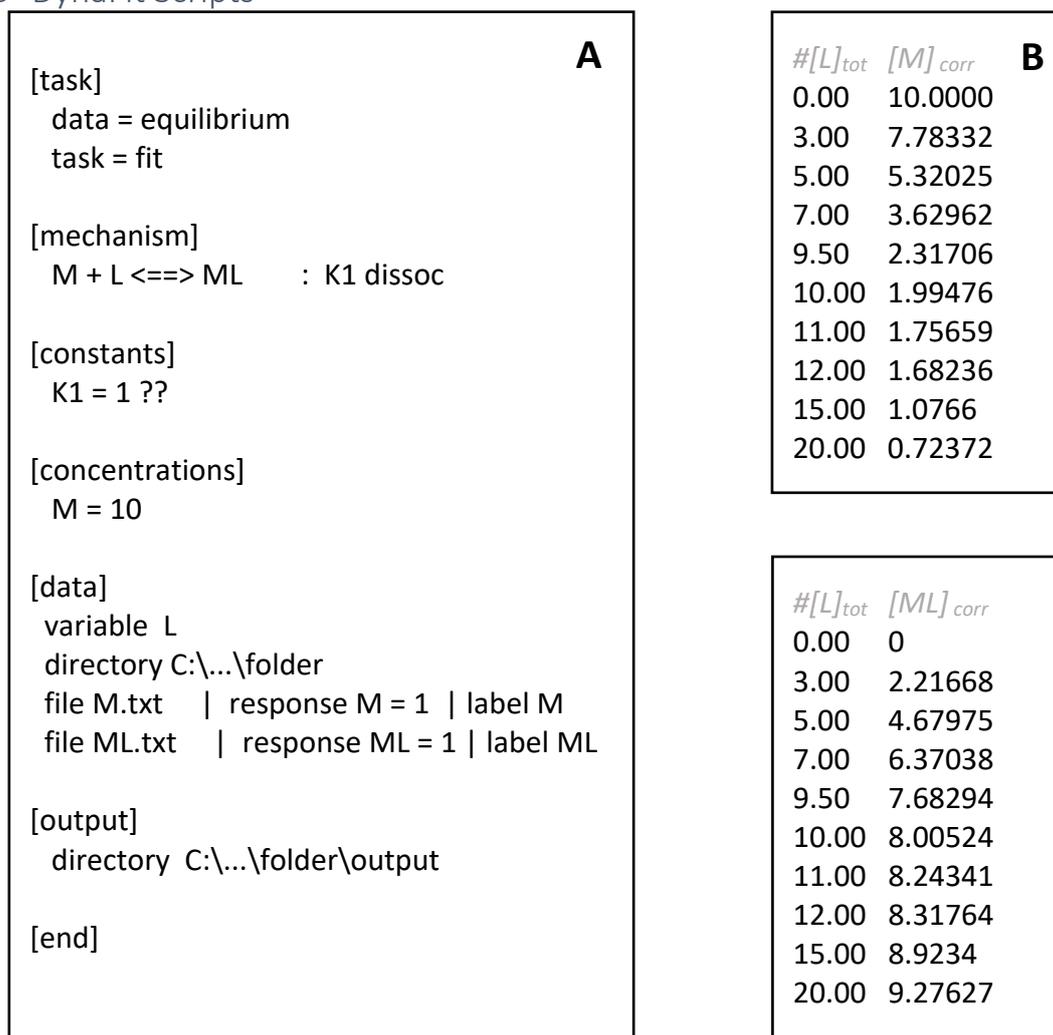


Figure 93 - **A**. DynaFit script to fit MS titration points. We highlight the instruction “??” that compute the full confidence interval for the corresponding parameter. $K1 = 1$ is a realistic seed value to start the fitting algorithm. **B**. Example of input (plain text) files with $[L]_{tot}$ versus $[M]_{corr}$ and $[ML]_{corr}$ per each titration point.

Table 23 - Dynafit output summary of isobar of Figure 21 (p.42)

Regression summary						
Levenberg-Marquardt Algorithm		Par#Set	Initial	Final	Std. Err	CV (%)
		K _D	1	0.598	0.046	7.7
sum of squares	1.12877					
mean square	0.0564383					
r.m.s. deviation	0.237567					
relative r.m.s. (%)	2.37567					
R ²	0.994684					
R ² _{adj}	0.994684					
log(determinant)	1.4429					
data points	20					
optimized parameters	1					
iterations	23					
subiterations	0					
elapsed time (sec)	0.062					
error status	1					
error message						

7.2 Native MS and Calorimetry

7.2.1 Malachite green aptamer - MS titrations

MGA Mass ranges for data extraction

Table 24 – m/z ranges for signal integration in MS titration and for IM-MS ATDs extraction.

TMAA 100 mM				
desolvated		3 adducts		Ion
m/z_i	m/z_f	m/z_i	m/z_f	
879.8	882.4	879.8	882.4	T6 ²⁻
876.8	879.4	876.8	879.4	Background T6 ²⁻
1755	1757	1755	1789.5	MGA ⁷⁻
1750	1752	1720.5	1755	Background MGA ⁷⁻
1802.25	1804.25	1802.3	1836.75	MGA+Lig ⁷⁻
1750	1752	1720.5	1755	Background MGA+Lig ⁷⁻
2048	2050.5	2048	2069.5	MGA ⁶⁻
2042.5	2045	2026.5	2048	Background MGA ⁶⁻
2102.75	2105.25	2102.75	2124.25	MGA+Lig ⁶⁻
2042.5	2045	2026.5	2048	Background MGA+Lig ⁶⁻
2457.75	2460.75	2457.8	2484	MGA ⁵⁻
2454	2457	2430.8	2457	Background MGA ⁵⁻
2523	2526	2523	2549.3	MGA+Lig ⁵⁻
2454	2457	2430.8	2457	Background MGA+Lig ⁵⁻
NH ₄ OAc 100 mM				
desolvated		3 adducts		Ion
m/z_i	m/z_f	m/z_i	m/z_f	
879.8	882.4	879.8	882.4	T6 ²⁻
876.8	879.4	876.8	879.4	Background T6 ²⁻
1755	1757.3	1755	1764.5	MGA ⁷⁻
1750	1752.3	1745.5	1755	Background MGA ⁷⁻
1802	1804.3	1802.3	1811.8	MGA+Lig ⁷⁻
1750	1752.3	1745.5	1755	Background MGA+Lig ⁷⁻
2048	2050.5	2048	2059	MGA ⁶⁻
2042.5	2045	2037	2048	Background MGA ⁶⁻
2102.7	2105.2	2102.7	2113.7	MGA+Lig ⁶⁻
2042.5	2045	2037	2048	Background MGA+Lig ⁶⁻
2457.7	2461	2457.7	2471	MGA ⁵⁻
2454.4	2457	2443.7	2457	Background MGA ⁵⁻
2523	2526.3	2523	2536.3	MGA+Lig ⁵⁻
2454.4	2457	2443.7	2457	Background MGA+Lig ⁵⁻

Table 25 - - ESI relative response factors in TMAA 100 mM of MGA and MGA+L.. σ^a "σ" is the standard deviation across the replicas.

0 TMA⁺ adducts						
7m	rSTD/rM	σ^a	rSTD/rML	σ^a	rML/rM	σ^a
I	45.82		28.98		1.58	
II	42.57		34.10		1.25	
III	90.46		60.49		1.50	
Average	59.62	± 21.85	41.19	± 13.81	1.44	± 0.14
6m						
I	7.02		5.21		1.35	
II	6.88		5.44		1.26	
III	7.47		6.59		1.13	
Average	7.12	± 0.25	5.75	± 0.61	1.25	± 0.09
5m						
I	14.98		11.54		1.30	
II	14.80		11.79		1.26	
III	17.03		15.08		1.13	
Average	15.60	± 1.01	12.81	± 1.61	1.23	± 0.07
3 TMA⁺ adducts						
7m	rSTD/rM	σ^a	rSTD/rML	σ^a	rML/rM	σ^a
I	16.77		11.73		1.43	
II	15.00		9.65		1.55	
III	30.06		18.39		1.63	
Average	20.61	± 6.72	13.25	± 3.73	1.54	± 0.08
6m						
I	2.90		2.02		1.44	
II	2.79		2.18		1.28	
III	3.35		2.67		1.25	
Average	3.01	± 0.24	2.29	± 0.28	1.32	± 0.08
5m						
I	4.69		3.04		1.54	
II	4.44		3.29		1.35	
III	5.84		4.22		1.38	
Average	4.99	± 0.61	3.52	± 0.51	1.43	± 0.08

Table 26 - ESI relative response factors in NH₄OAc 100 mM of MGA and MGA+L. σ^a "σ" is the standard deviation across the replicas.

0 NH₄⁺ adducts						
7-	rSTD/rM	σ^a	rSTD/rML	σ^a	rML/rM	σ^a
I	2.10		1.62		1.29	
II	11.10		5.62		1.98	
III	3.58		2.22		1.61	
Average	5.59	± 3.94	3.15	± 1.76	1.63	± 0.28
6-						
I	0.58		0.57		1.02	
II	2.54		1.90		1.33	
III	1.00		0.66		1.53	
Average	1.37	± 0.84	1.04	± 0.61	1.29	± 0.21
5-						
I	5.09		2.51		2.03	
II	18.97		7.93		2.39	
III	8.53		7.50		1.14	
Average	10.86	± 5.90	5.98	± 2.46	1.85	± 0.53
3 NH₄⁺ adducts						
7-	rSTD/rM	σ^a	rSTD/rML	σ^a	rML/rM	σ^a
I	1.18		0.98		1.20	
II	6.44		3.39		1.90	
III	1.78		1.21		1.47	
Average	3.13	± 2.35	1.86	± 1.09	1.53	± 0.29
6-						
I	0.24		0.25		0.96	
II	1.14		0.86		1.32	
III	0.41		0.26		1.56	
Average	0.60	± 0.39	0.46	± 0.29	1.28	± 0.25
5-						
I	1.71		1.00		1.71	
II	7.36		3.08		2.39	
III	2.93		2.16		1.36	
Average	4.00	± 2.43	2.08	± 0.85	1.82	± 0.43

Ch. St. 7- of MGA in TMAA 100 mM

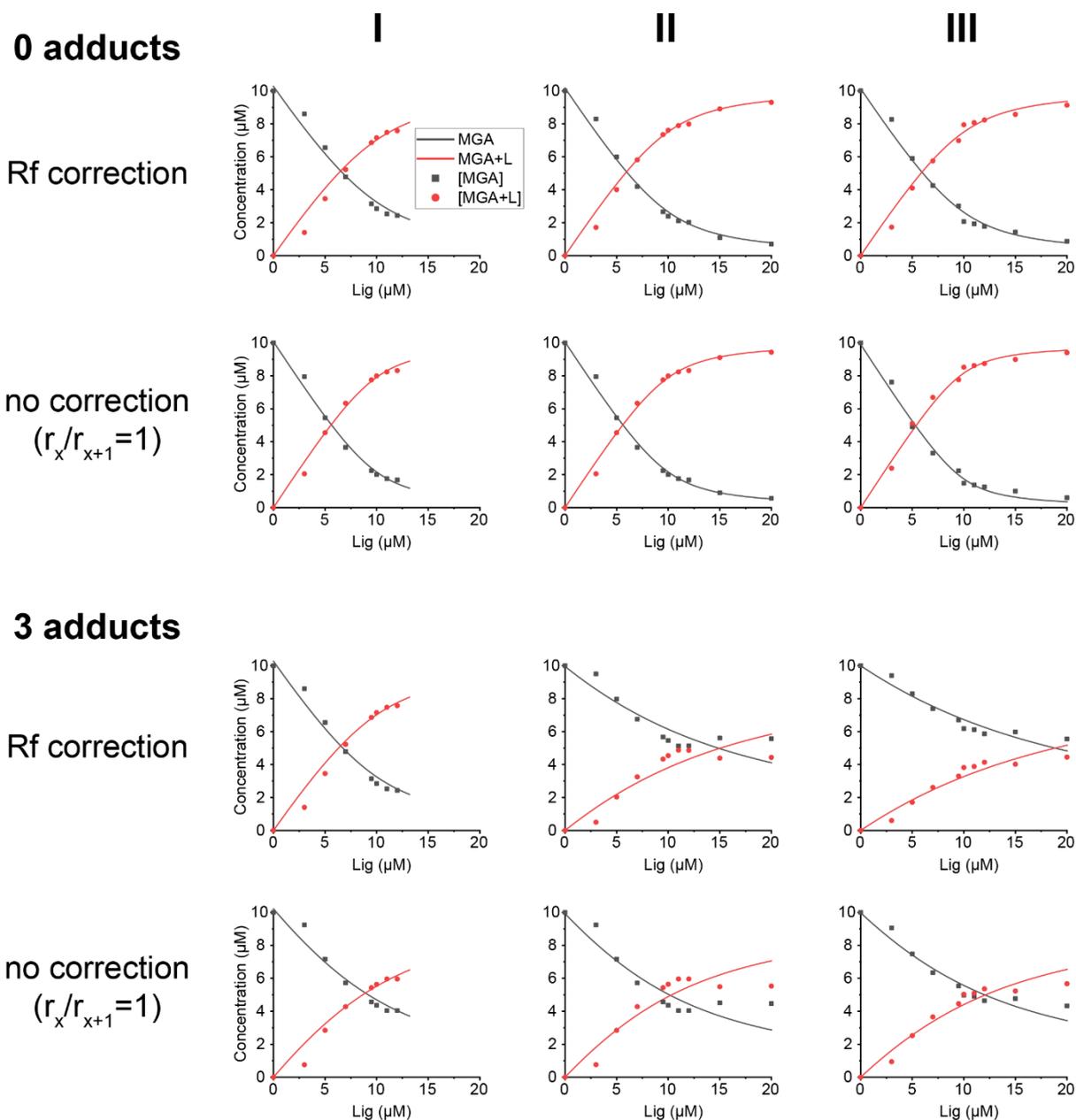


Figure 94 – Fitted isotherm binding curve of each MS replica for MGA ch st 7- in TMAA 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

Ch. St. 6- of MGA in TMAA 100 mM

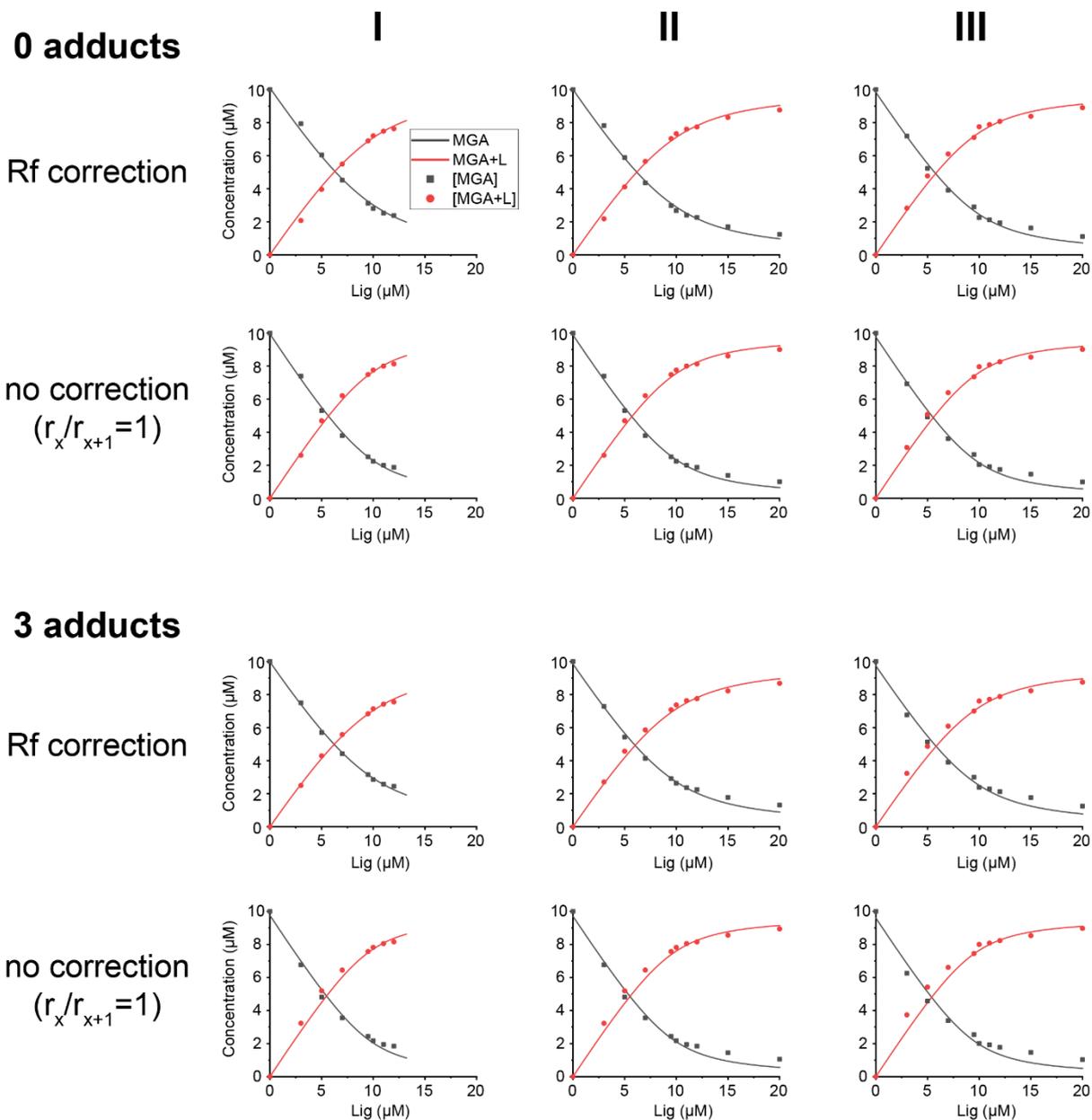


Figure 95 – Fitted isotherm binding curve of each MS replica for MGA ch st 6- in TMAA 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

Ch. St. 5- of MGA in TMAA 100 mM

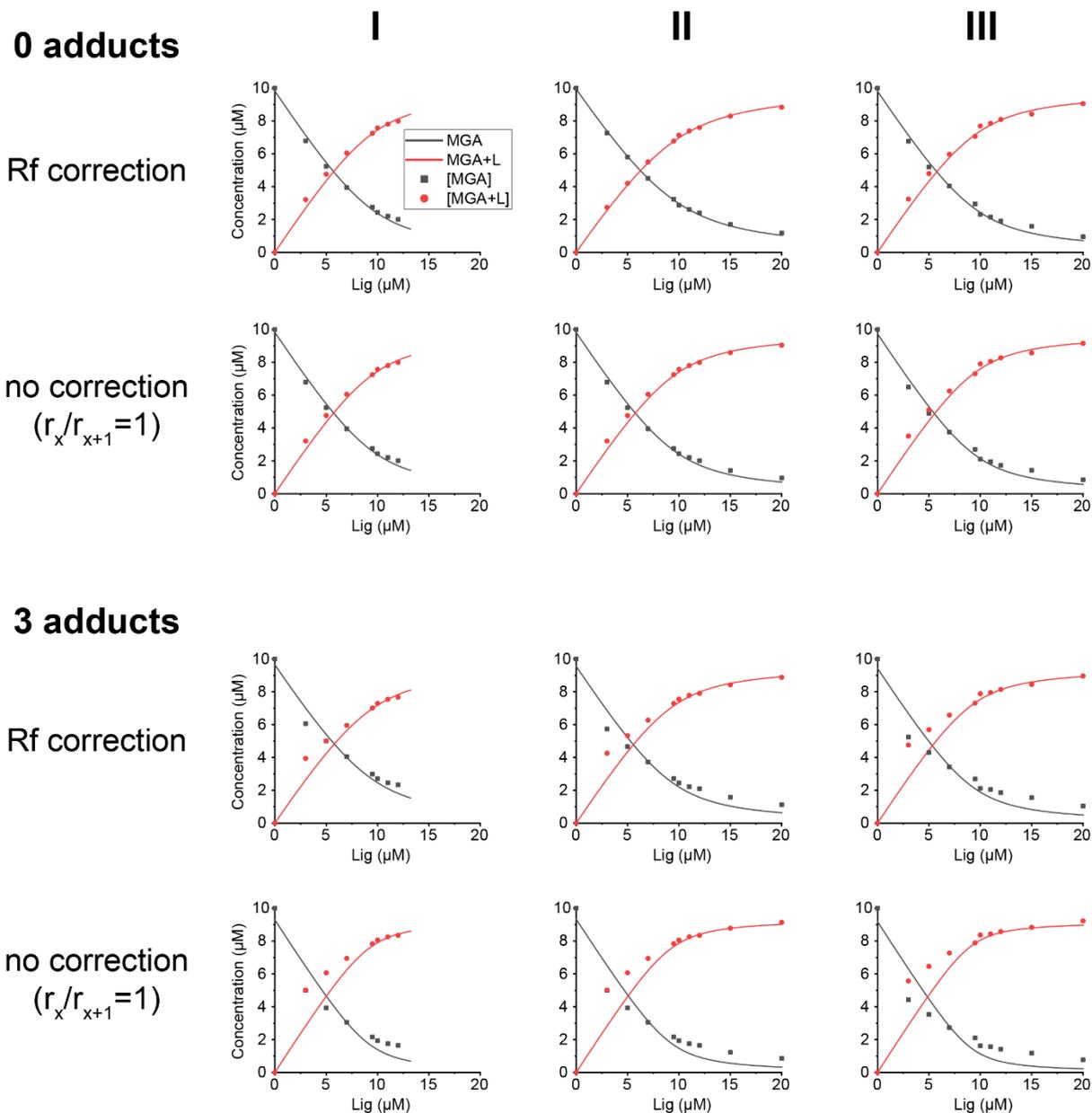


Figure 96 – Fitted isotherm binding curve of each MS replica for MGA ch st 5- in TMAA 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

Ch. St. 7- of MGA in NH₄OAc 100 mM

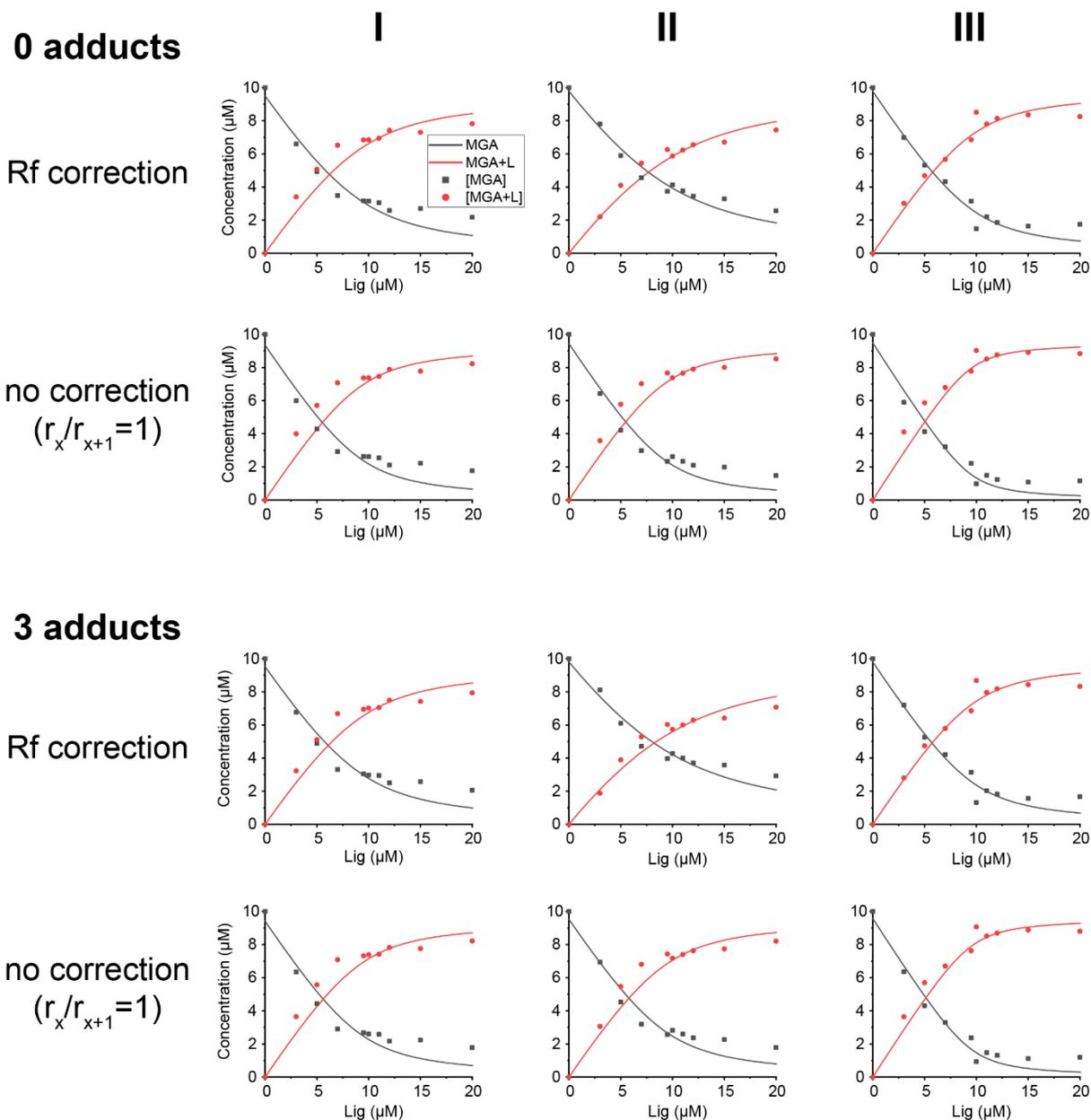


Figure 97 - Fitted isotherm binding curve of each MS replica for MGA ch st 7- in NH₄OAc 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

Ch. St. 6- of MGA in NH₄OAc 100 mM

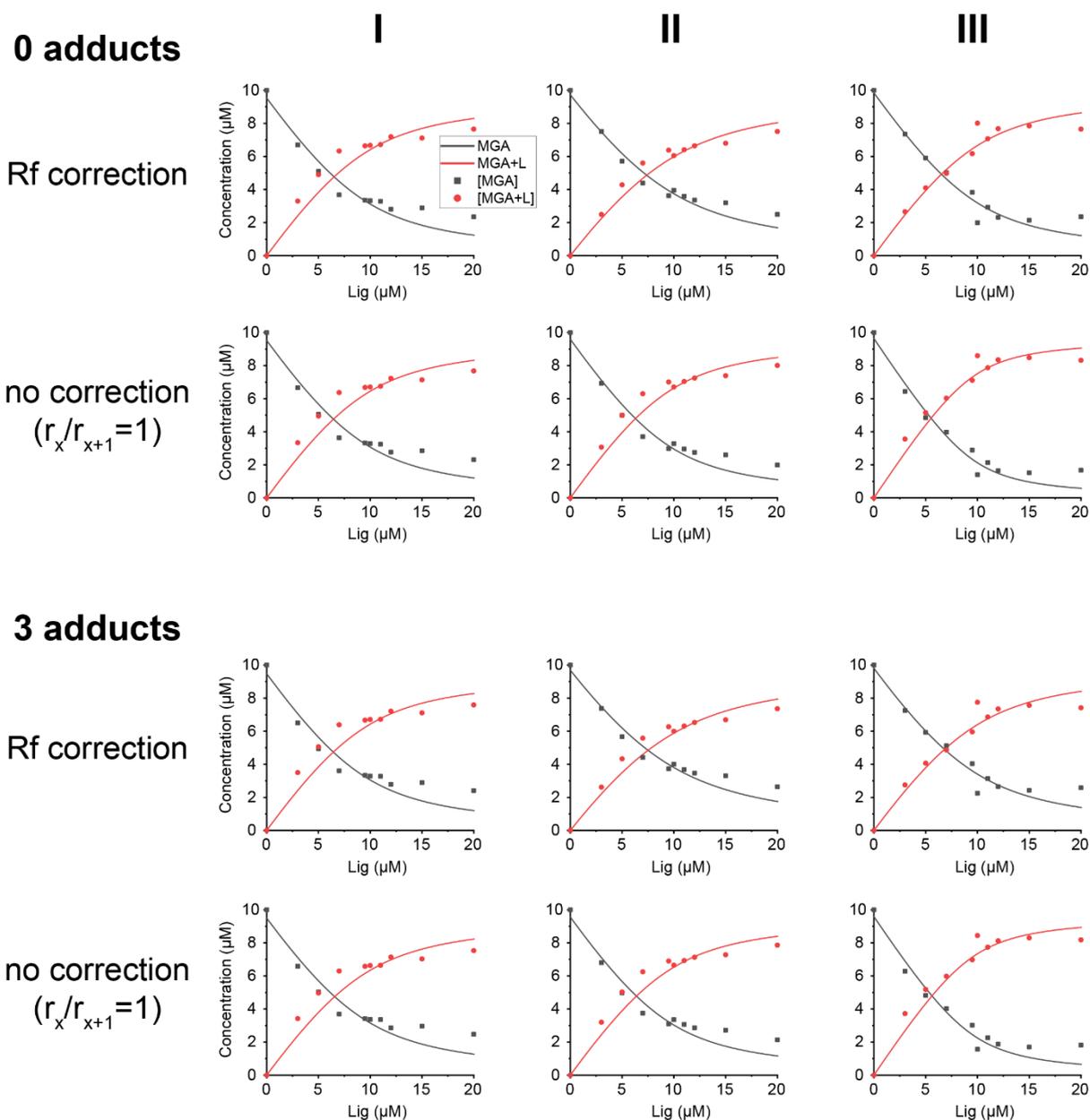


Figure 98 - Fitted isotherm binding curve of each MS replica for MGA ch st 6- in NH₄OAc 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

Ch. St. 5- of MGA in NH₄OAc 100 mM

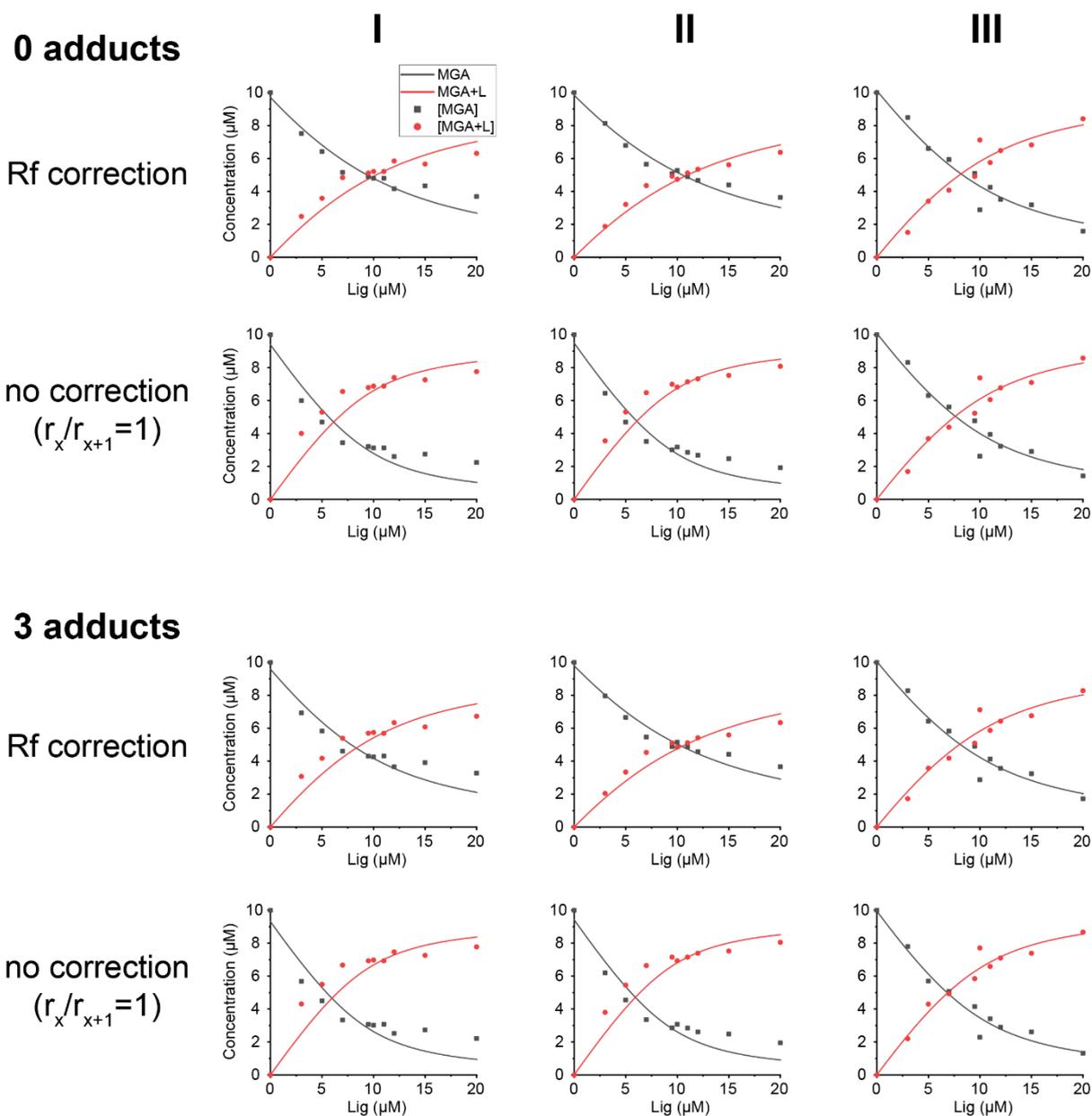


Figure 99 - Fitted isotherm binding curve of each MS replica for MGA ch st 5- in NH₄OAc 100 mM. Integration strategies of 0 or 3 electrolyte adducts are reported, considering relative response factor (Rf) correction or assuming equal response factor ($r_x/r_{x+1}=1$).

MS log K_D and ΔG° - TMAA

Table 29 – Averaged K_D , and calculated Log K_D and ΔG° , from MS titration of MGA in TMAA 100 mM; “ σ ” is the standar deviation over the replica; “Exp. σ ” is the expanded uncertainty considering a 95% confidence interval with coverage factor of 4.303.

		0 TMA ⁺ adducts (Desolvated)					
		Corrected			Not corrected ($r_x/r_{x+1}=1$)		
		σ	Exp. σ		Σ	Exp. σ	
7-	K_D (μM)	1.05	± 0.24	± 1.01	0.52	± 0.09	± 0.37
	Log K_D	-5.99	± 0.09	± 0.40	-6.29	± 0.08	± 0.33
	ΔG° (kcal/mol)	-8.21	± 0.13	± 0.54	-8.62	± 0.11	± 0.46
6-	K_D (μM)	1.09	± 0.16	± 0.70	0.71	± 0.03	± 0.15
	Log K_D	-5.97	± 0.07	± 0.29	-6.15	± 0.02	± 0.09
	ΔG° (kcal/mol)	-8.18	± 0.09	± 0.40	-8.43	± 0.03	± 0.12
5-	K_D (μM)	1.16	± 0.21	± 0.92	0.78	± 0.07	± 0.32
	Log K_D	-5.94	± 0.09	± 0.37	-6.11	± 0.04	± 0.18
	ΔG° (kcal/mol)	-8.14	± 0.12	± 0.51	-8.37	± 0.06	± 0.25
		3 TMA ⁺ adducts					
		Corrected			Not corrected ($r_x/r_{x+1}=1$)		
		σ	Exp. σ		σ	Exp. σ	
7-	K_D (μM)	10.12	± 2.99	± 12.87	5.38	± 1.34	± 5.76
	Log K_D	-5.02	± 0.13	± 0.58	-5.28	± 0.11	± 0.47
	ΔG° (kcal/mol)	-6.87	± 0.18	± 0.79	-7.24	± 0.15	± 0.65
6-	K_D (μM)	1.10	± 0.11	± 0.48	0.62	± 0.03	± 0.13
	Log K_D	-5.96	± 0.04	± 0.19	-6.21	± 0.02	± 0.09
	ΔG° (kcal/mol)	-8.17	± 0.06	± 0.26	-8.50	± 0.03	± 0.13
5-	K_D (μM)	0.80	± 0.14	± 0.59	0.36	± 0.05	± 0.22
	Log K_D	-6.11	± 0.08	± 0.33	-6.45	± 0.07	± 0.28
	ΔG° (kcal/mol)	-8.37	± 0.11	± 0.45	-8.84	± 0.09	± 0.38

MS log K_D and ΔG° - NH₄OAc

Table 30 – Averaged K_D, and calculated Log K_D and ΔG°, from MS titration of MGA in NH₄OAc 100 mM; “σ” is the standar deviation over the replica; “Exp. σ” is the expanded uncertainty considering a 95% confidence interval with coverage factor of 4.303.

		0 NH ₄ ⁺ adducts (Desolvated)					
		Corrected			Not corrected r _x /r _{x+1} =1)		
		σ	Exp. σ		σ	Exp. σ	
7-	K _D (μM)	1.72	± 0.79	± 3.41	0.65	± 0.24	± 1.05
	Log K _D	-5.81	± 0.20	± 0.87	-6.23	± 0.20	± 0.87
	ΔG° (kcal/mol)	-7.96	± 0.28	± 1.19	-8.54	± 0.28	± 1.19
6-	K _D (μM)	1.96	± 0.40	± 1.72	1.30	± 0.43	± 1.84
	Log K _D	-5.72	± 0.08	± 0.36	-5.91	± 0.17	± 0.72
	ΔG° (kcal/mol)	-7.83	± 0.12	± 0.50	-8.10	± 0.23	± 0.99
5-	K _D (μM)	4.63	± 1.13	± 4.87	1.78	± 0.56	± 2.39
	Log K _D	-5.35	± 0.12	± 0.50	-5.77	± 0.13	± 0.54
	ΔG° (kcal/mol)	-7.33	± 0.16	± 0.68	-7.90	± 0.17	± 0.74
		3 NH ₄ ⁺ adducts					
		Corrected			Not corrected (r _x /r _{x+1} =1)		
		σ	Exp. σ		σ	Exp. σ	
7-	K _D (μM)	1.82	± 1.08	± 4.63	0.77	± 0.30	± 1.29
	Log K _D	-5.81	± 0.25	± 1.08	-6.16	± 0.21	± 0.91
	ΔG° (kcal/mol)	-7.97	± 0.34	± 1.48	-8.44	± 0.29	± 1.25
6-	K _D (μM)	2.10	± 0.41	± 1.77	1.42	± 0.43	± 1.86
	Log K _D	-5.69	± 0.08	± 0.35	-5.87	± 0.15	± 0.66
	ΔG° (kcal/mol)	-7.79	± 0.11	± 0.48	-8.05	± 0.21	± 0.90
5-	K _D (μM)	4.04	± 1.09	± 4.70	1.48	± 0.29	± 1.23
	Log K _D	-5.41	± 0.11	± 0.48	-5.84	± 0.08	± 0.34
	ΔG° (kcal/mol)	-7.41	± 0.15	± 0.66	-8.00	± 0.11	± 0.47

MGA Method 2 (Non-linear fitting with averaged [Conc] titration points as input)

Table 31 – Ratios of relative response factors (defined as $R_{MGA}/R_{MGA+Lig}$) and the corresponding K_D calculated from fitting of binding isotherm are reported; **a.** uncertainty is the standard deviation across the replicas; **b.** K_D uncertainty is calculated from DynaFit statistic with a Monte-Carlo simulation [107-108].

TMAA 100 mM											
z		Desolvated			3 adducts						
		Rx^a	R_M/R_{ML}	K_D (μM)^b	Rx^a	R_M/R_{ML}	K_D (μM)^b				
7-	M	59.62	±21.85	1.45± 0.14	1.01	±0.09	20.61	±6.72	1.55± 0.08	10.90	±1.10
	ML	41.19	±13.81				13.25	±3.73			
6-	M	7.12	±0.24	1.25± 0.09	1.10	±0.05	3.01	±0.21	1.39± 0.08	1.11	±0.06
	ML	5.70	±0.60				2.17	±0.25			
5-	M	15.60	±1.01	1.22± 0.07	1.15	±0.04	4.99	±0.61	1.42± 0.08	0.80	±0.14
	ML	12.81	±1.61				3.52	±0.51			

NH₄OAc 100 mM											
z		Desolvated			3 adducts						
		Rx^a	R_M/R_{ML}	K_D (μM)^b	Rx^a	R_M/R_{ML}	K_D (μM)^b				
7-	M	2.87	±0.71	1.48± 0.28	1.61	±0.17	2.36	±1.60	1.50± 0.29	1.64	±0.17
	ML	1.94	±0.28				1.58	±0.86			
6-	M	0.80	±0.21	1.31± 0.21	1.93	±0.19	0.45	±0.25	1.26± 0.25	2.07	±0.22
	ML	0.61	±0.05				0.36	±0.20			
5-	M	6.99	±1.54	1.40± 0.53	4.49	±0.23	2.78	±1.41	1.72± 0.43	3.93	±0.26
	ML	4.99	±2.51				1.62	±0.71			

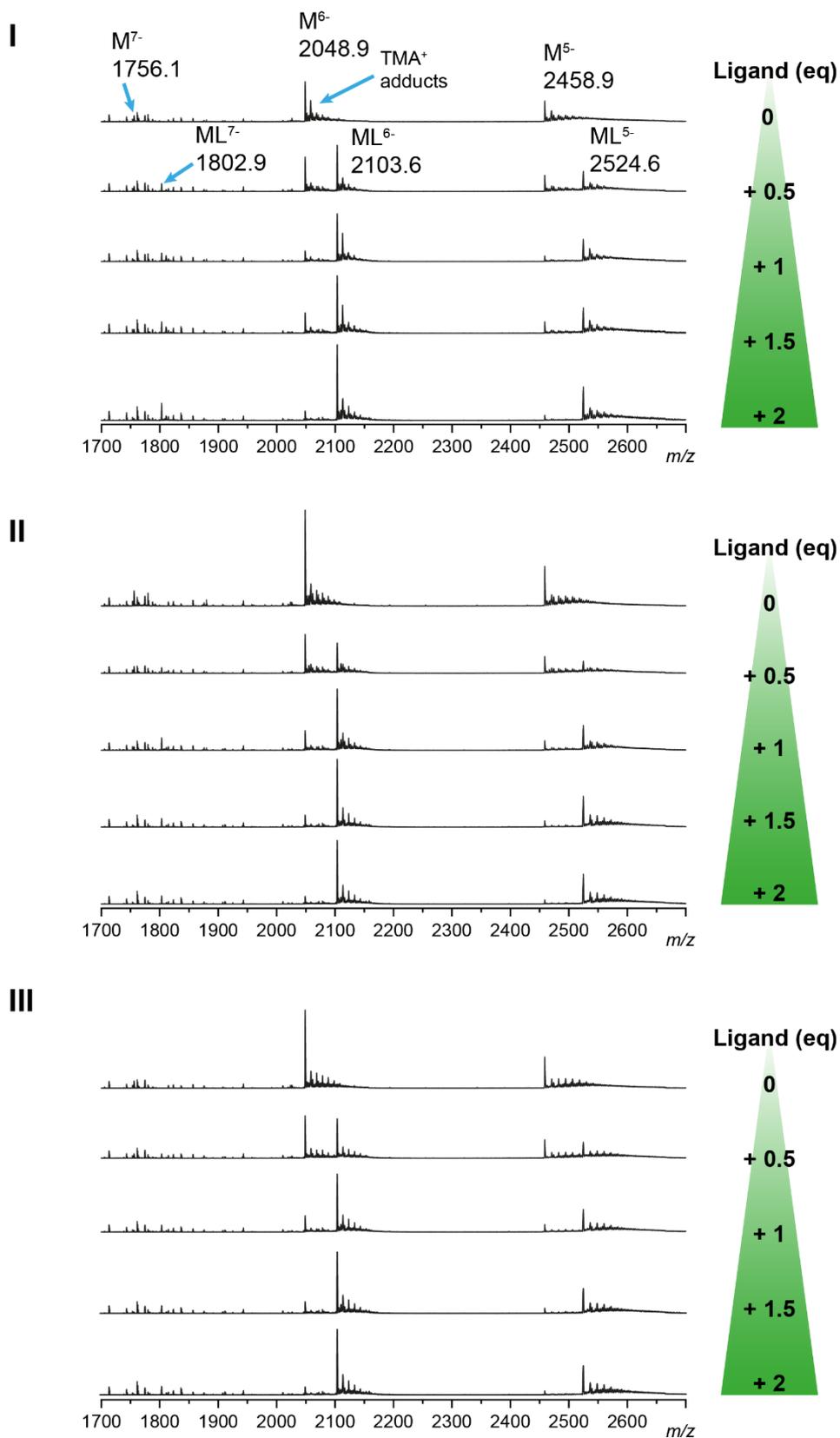


Figure 100 – MS replica of MGA (10 μ M) titrated up to 2 eq. of MG (20 μ M) in TMAA 100 mM.

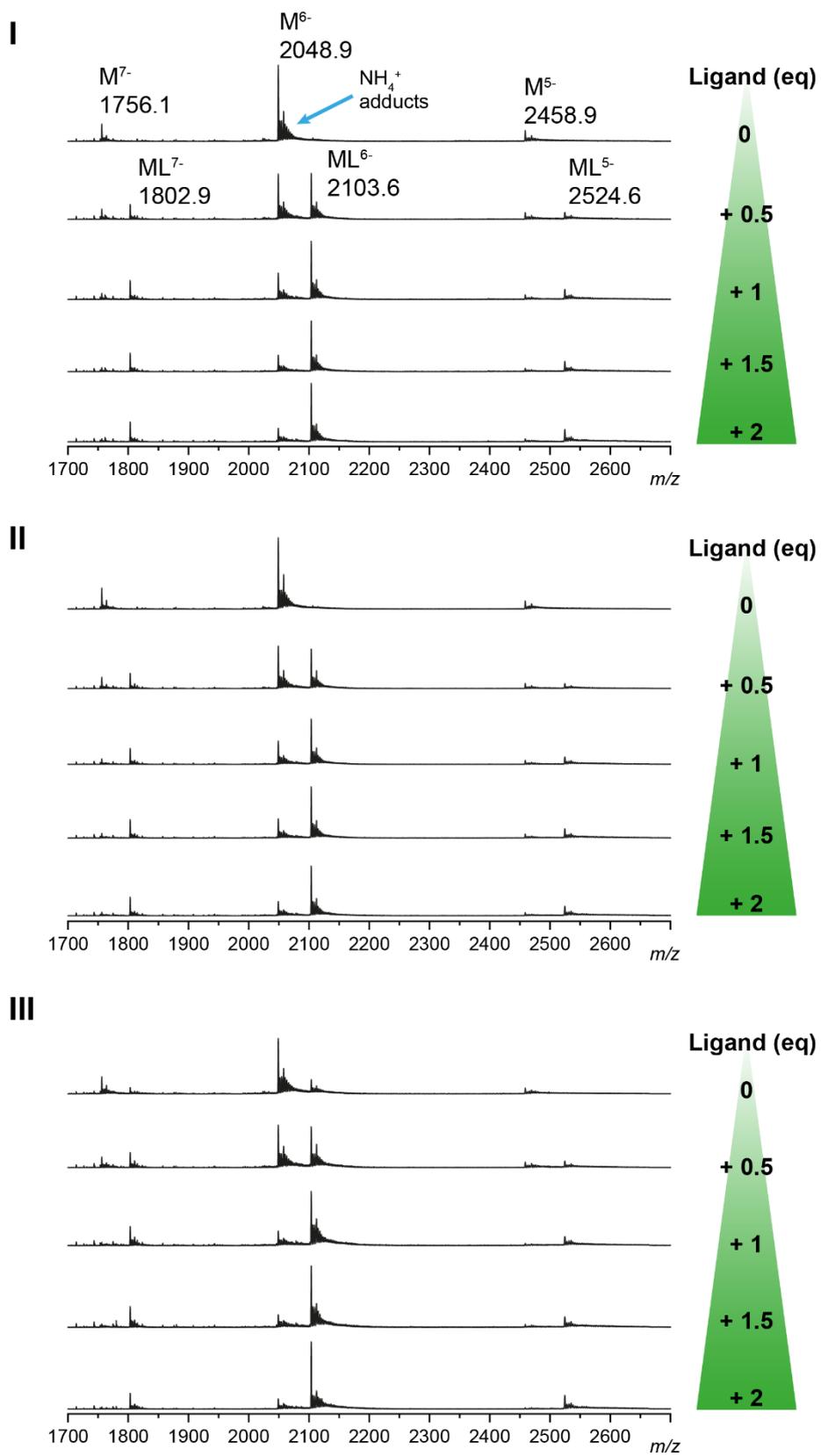


Figure 101 – MS replica of MGA (10 μ M) titrated up to 2 eq. of MG (20 μ M) in NH_4OAc 100 mM.

7.2.2 RiboFMN aptamer - MS titrations

Table 32 - m/z ranges for MS peak integration 1FMN aptamer in NH_4OAc and TMAA 100 mM.

1FMN in TMAA 100 mM, MgCl_2 200 μM				
All adducts		5 adducts		Ion
m/z_i	m/z_f	m/z_i	m/z_f	
879.8	882.4	879.8	882.4	T6^{2-}
876.8	879.4	876.8	879.4	Background T6^{2-}
1894	1924	1894	1912.5	1FMN^{6-}
1864	1894	1875.5	1894	Background 1FMN^{6-}
1978	2008	1978	1996.5	$1\text{FMN}+\text{Lig}^{6-}$
1864	1894	1875.5	1894	Background $1\text{FMN}+\text{Lig}^{6-}$

1FMN in NH_4OAc 100 mM, MgCl_2 200 μM				
All adducts		5 adducts		Ion
m/z_i	m/z_f	m/z_i	m/z_f	
879.8	882.4	879.8	882.4	T6^{2-}
876.8	879.4	876.8	879.4	Background T6^{2-}
1890	1920	1890	1908.5	1FMN^{6-}
1860	1890	1871.5	1890	Background 1FMN^{6-}
1971	2001	1971	1989.5	$1\text{FMN}+\text{Lig}^{6-}$
1860	1890	1871.5	1890	Background $1\text{FMN}+\text{Lig}^{6-}$

Table 33 - m/z ranges for MS peak integration 1FMN in NH_4OAc 100 mM and MgCl_2 concentration test.

NH ₄ OAc 100 mM (all visible adducts)		
m/z_i	m/z_f	
448.6	452.0	Background riboFMN
454.8	458.2	riboFMN
876.8	879.4	Background T6^{2-}
879.8	882.4	T6^{2-}
1875	1893	Background
1894	1912.5	1FMN^{6-}
1978	1996.5	$1\text{FMN}+\text{L}^{6-}$

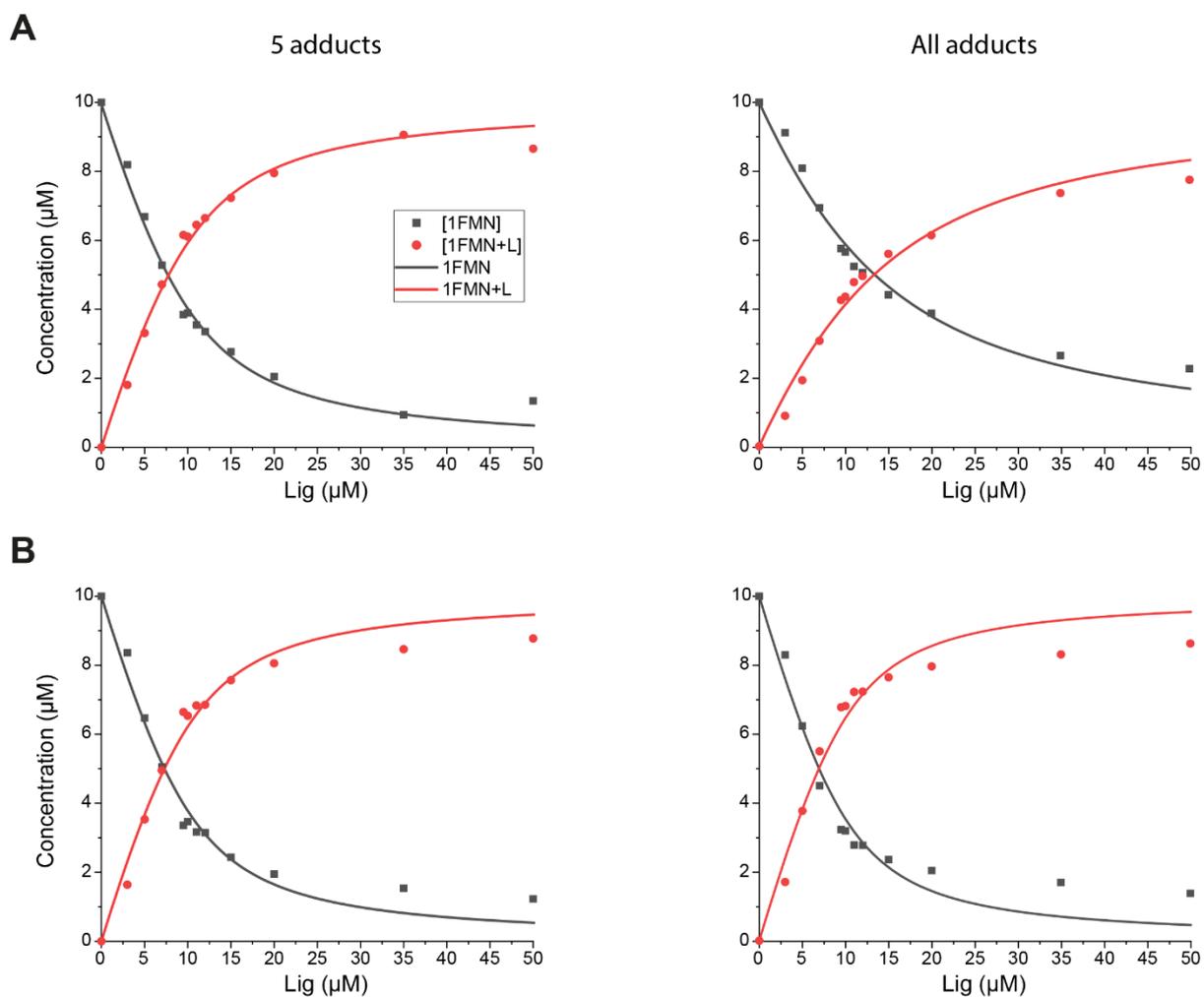


Figure 102 - Comparison of binding isotherm considering only 5 Mg^{2+} or the whole visible adduct serie, for 1FMN^{6-} and its complex with RiboFMN^{6-} in $200 \mu\text{M Mg}^{2+}$ and **A.** NH_4OAc or **B.** TMAA 100 mM. Points are averaged over 3 replica.

MS titr NH₄OAc - 1FMN

Table 34 – 1FMN+L K_D of each replica of MS titration in NH₄OAc 100 mM with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r/r=1$). “ σ ” is the standard deviation of DynaFit’s Monte-Carlo algorithm.

Replica	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
	K _D (μ M)	σ	Log K _D	K _D (μ M)	σ	Log K _D
I	128.7	± 5.20	-3.89	29.33	± 1.50	-4.53
II	167.2	± 6.40	-3.78	30.01	± 1.50	-4.52
III	92.94	± 4.00	-4.03	29.22	± 1.50	-4.53
IV	44.61	± 2.10	-4.35	24.23	± 1.30	-4.62
V	120.5	± 6.30	-3.92	25	± 1.60	-4.60

Table 35 – Averaged K_D, log K_D and calculated ΔG° for 1FMN MS titration in NH₄OAc 100 mM with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r/r=1$). “ σ ” is the standard deviation over the replicas, “**Exp.** σ ” is the expanded uncertainty over 5 replicas (coverage factor = 2.776 for a 95% confidence interval).

Averaged values	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
		σ	Exp. σ		σ	Exp. σ
K _D (μ M)	110.79	± 40.73	± 113.09	27.56	± 2.43	± 6.75
Log K _D	-3.99	± 0.20	± 0.54	-4.56	± 0.04	± 0.11
ΔG° (kcal/mol)	-5.48	± 0.27	± 0.75	-6.25	± 0.05	± 0.15

Response Factors replicas

Table 36 - 1FMN⁶⁻ and 1FMN+L⁶⁻ relative response factors for MS titration in NH₄OAc 100 mM. “ σ ” is the standard deviation over the replicas. Instrument: Agilent 6560.

Replica	rSTD/rM		rSTD/rML		rML/rM	
I	0.61		0.17		3.59	
II	0.63		0.14		4.50	
III	0.58		0.22		2.68	
IV	0.61		0.37		1.62	
V	0.69		0.19		3.65	
	Avrg.	σ	Avrg.	σ	Avrg.	σ
	0.62	± 0.03	0.22	± 0.08	3.21	± 0.98

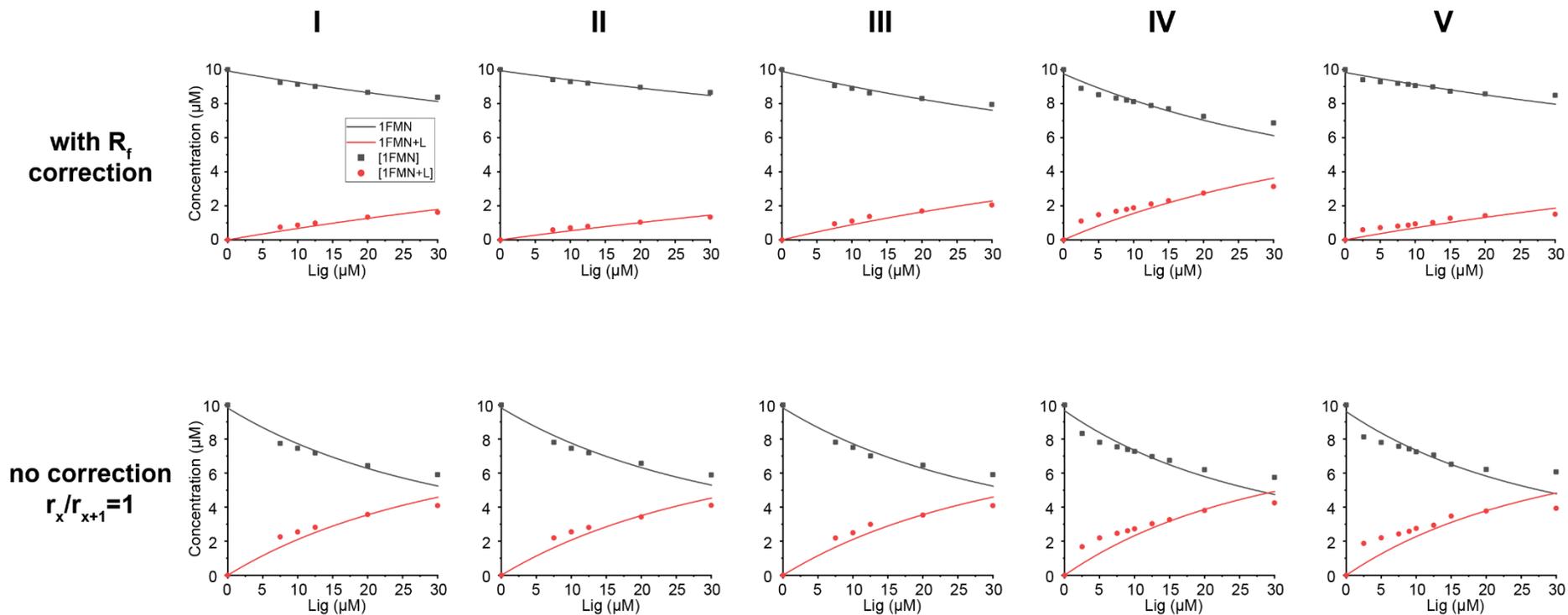


Figure 103 – Replicas of MS titration binding isotherm of 1FMN^{6-} in NH_4OAc 100 mM. Corrected concentrations and assumption of equal response factor ($r_x/r_{x+1}=1$) are compared.

Table 37 - 1FMN⁶⁻ and 1FMN+L⁶⁻ Averaged relative response factors for MS titration in NH₄OAc 100 Mm with 100, 200, 300, 400 and 800 μM MgCl₂.

Mg ²⁺ dependency Rfs from 100 to 800 μM Mg ²⁺			
1FMN ⁶⁻	rSTD/rM	rSTD/rML	rML/rM
All adducts	0.30	0.37	0.82

Table 38 Estimated 1FMN+L K_Ds for each MgCl₂ concentration in NH₄OAc 100 mM, assuming ratio of relative response factors equal to 1.

Mg ²⁺ (μM)	K _D	Log [Mg ²⁺]	Log K _D
0	2.21E-05	--	-4.65
10	1.73E-05	1	-4.76
20	1.48E-05	1.30	-4.83
30	1.36E-05	1.48	-4.87
100	6.44E-06	2.00	-5.19
200	4.51E-06	2.30	-5.35
300	3.72E-06	2.48	-5.43
400	3.37E-06	2.60	-5.47
800	2.60E-06	2.90	-5.59

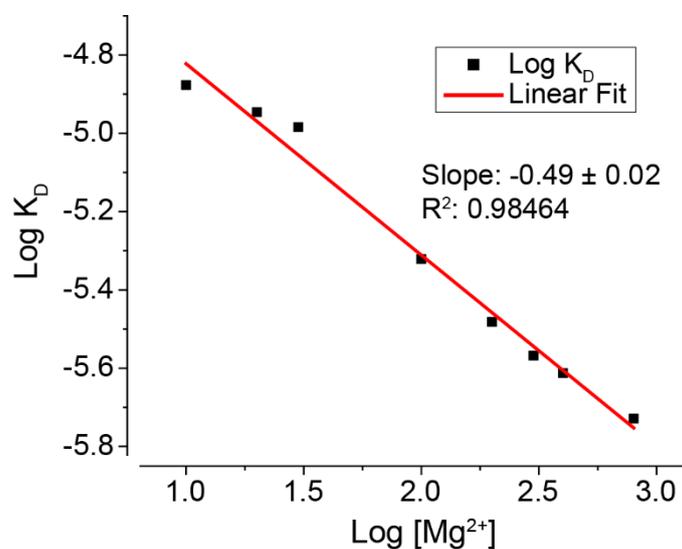


Figure 104 - Effect of Mg²⁺ concentration on 1FMN+L K_D^{App}, expressed in logarithmic scale. From the slope modulus of linear regression we have a stoichiometry of 0.49 ± 0.02 for Mg²⁺.

MS titr Response factors in NH₄OAc+Mg and TMAA+Mg - 1FMN

Table 39 -1FMN⁶⁻ and 1FMN+L⁶⁻ relative response factors for MS titration in NH₄OAc and TMAA 100 Mm with 200 μM MgCl₂; First replica (italic) in TMAA has been discarded. Instrument: Thermo Exactive.

5 Mg ²⁺ adducts in NH ₄ OAc						
6-	rSTD/rM	σ ^a	rSTD/rML	σ ^a	rML/rM	σ ^a
I	4.47		6.84		0.65	
II	5.11		5.00		1.02	
III	0.98		2.42		0.41	
Average	3.52	± 1.82	4.75	± 1.81	0.69	± 0.25

5 Mg ²⁺ adducts in TMAA						
6-	rSTD/rM	σ ^a	rSTD/rML	σ ^a	rML/rM	σ ^a
I	9.10		4.26		2.14	
II	7.22		5.91		1.22	
III	7.77		7.76		1.00	
Average	7.49	± 0.27	6.83	± 0.93	1.11	± 0.11

MS titr NH₄OAc+Mg - 1FMN

Table 40 -1FMN+L K_D^{app} of each replica of MS titration in NH₄OAc 100 mM and MgCl₂ 200 μM with relative response factor correction (**Corrected**) and assuming ratio of response factor equal to 1 (r_x/r_{x+1}=1). “σ” is the standard deviation of DynaFit’s Monte-Carlo algorithm. Signal were integrated only the 5 first Mg²⁺ adducts.

Replica	1FMN in NH ₄ OAc 100 mM and MgCl ₂ 200 μM (5 Mg ²⁺ adducts)			1FMN in NH ₄ OAc 100 mM and MgCl ₂ 200 μM (5 Mg ²⁺ adducts)		
	K _D ^{app}	σ	Log K _D	K _D ^{app}	σ	Log K _D
I	3.51	0.18	-5.45	6.43	0.30	-5.19
II	4.59	0.18	-5.34	4.44	0.18	-5.35
III	1.15	0.06	-5.94	5.23	0.17	-5.28

Table 41 - Averaged K_D, log K_D and calculated ΔG° for 1FMN MS titration in NH₄OAc 100 mM and MgCl₂ 200 μM (considering 5 Mg²⁺ adducts) with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 (r_x/r_{x+1}=1). “σ” is the standard deviation over the replicas, “Exp.σ” is the expanded uncertainty over 3 replicas (coverage factor = 4.302 for a 95% confidence interval).

	1FMN in NH ₄ OAc 100 mM and MgCl ₂ 200 μM (5 Mg ²⁺ adducts)			1FMN in NH ₄ OAc 100 mM and MgCl ₂ 200 μM (5 Mg ²⁺ adducts)		
	σ	Exp.σ		σ	Exp.σ	
K_D (μM)	3.08	± 1.44	± 6.19	5.36	± 0.82	± 3.51
Log K_D	-5.58	± 0.26	± 1.12	-5.28	± 0.07	± 0.28
ΔG° (kcal/mol)	-7.64	± 0.36	± 1.54	-7.23	± 0.09	± 0.39

Table 42 -1FMN+L K_D^{app} of each replica of MS titration in NH_4OAc 100 mM and $MgCl_2$ 200 μM with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r/r=1$). " σ " is the standard deviation of DynaFit's Monte-Carlo algorithm. Signal were integrated only the entire range of Mg^{2+} adducts.

1FMN in NH_4OAc 100 mM and $MgCl_2$ 200 μM (All Mg^{2+} adducts)

Replica	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
	K_D^{app}	σ	Log K_D	K_D^{app}	σ	Log K_D
I	4.80	0.26	-5.32	6.32	0.32	-5.20
II	3.32	0.18	-5.48	4.09	0.17	-5.39
III	1.23	0.06	-5.91	4.71	0.16	-5.33

Table 43 - Averaged K_D , log K_D and calculated ΔG° for 1FMN MS titration in NH_4OAc 100 mM and $MgCl_2$ 200 μM (considering All Mg^{2+} adducts) with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r_x/r_{x+1}=1$). " σ " is the standard deviation over the replicas, "**Exp. σ** " is the expanded uncertainty over 3 replicas (coverage factor = 4.302 for a 95% confidence interval).

1FMN in NH_4OAc 100 mM and $MgCl_2$ 200 μM (All Mg^{2+} adducts)

	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
	K_D	σ	Exp. σ	K_D	σ	Exp. σ
K_D (μM)	3.55	± 1.64	± 7.07	5.04	± 0.94	± 4.03
Log K_D	-5.52	± 0.27	± 1.18	-5.30	± 0.08	± 0.34
ΔG° (kcal/mol)	-7.57	± 0.38	± 1.62	-7.27	± 0.11	± 0.46

MS titr TMAA+Mg - 1FMN

Table 44 - 1FMN+L K_D^{app} of each replica of MS titration in TMAA 100 mM and $MgCl_2$ 200 μM with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r/r=1$). " σ " is the standard deviation of DynaFit's Monte-Carlo algorithm. Signal were integrated only the 5 first Mg^{2+} adducts.

Replica	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
	K_D^{app}	σ	Log K_D	K_D^{app}	σ	Log K_D
I	1.12	0.08	-5.95	0.19	0.04	-6.72
II	2.31	0.12	-5.64	2.07	0.10	-5.68
III	2.55	0.12	-5.59	2.55	0.12	-5.59

Table 45 - Averaged K_D , log K_D and calculated ΔG° for 1FMN MS titration in TMAA 100 mM and $MgCl_2$ 200 μM (considering 5 Mg^{2+} adducts) with relative response factor correction (**Corr**) and assuming ratio of response factor equal to 1 ($r_x/r_{x+1}=1$). " σ " is the standard deviation over the replicas, "**Exp. σ** " is the expanded uncertainty over 2 replicas (coverage factor = 12.706 for a 95% confidence interval).

	Corrected			Not corrected ($r_x/r_{x+1}=1$)		
	K_D (μM)	σ	Exp. σ	K_D (μM)	σ	Exp. σ
K_D (μM)	2.43	± 0.12	± 1.54	2.31	± 0.24	± 3.04
Log K_D	-5.62	± 0.12	± 1.54	-5.64	± 0.24	± 3.04
ΔG° (kcal/mol)	-7.69	± 0.17	± 2.12	-7.73	± 0.33	± 4.17

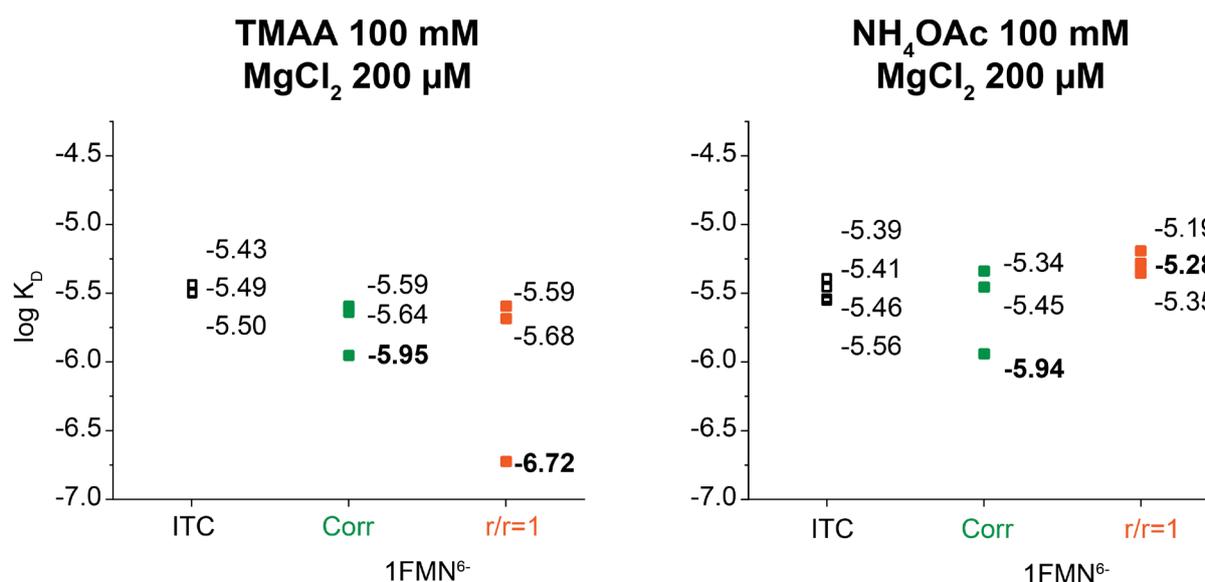
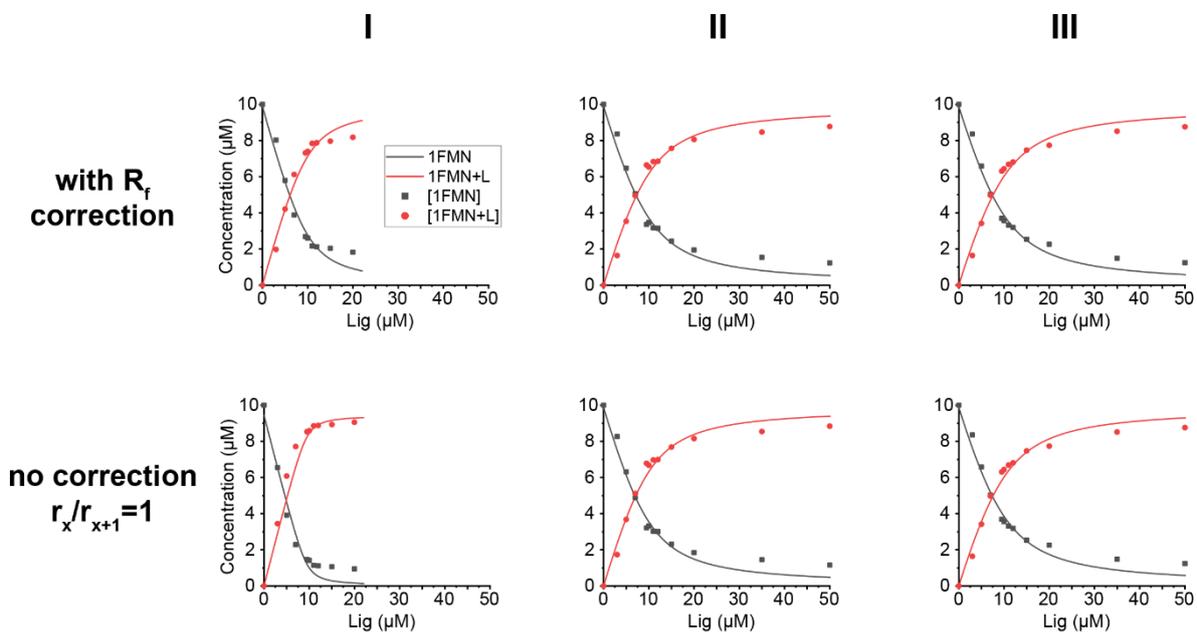


Figure 105 -Comparison between replica of log K_D determined via ITC and MS on 1FMN⁶⁻. MS values "**Corr**" are calculated with relative response correction, " $r/r=1$ " assuming equal response factors. Values in **bold** were discarded for the average calculation in picture Figure 46.

1FMN in TMAA 100 mM + 200 μM MgCl_2



1FMN in NH_4OAc 100 mM + 200 μM MgCl_2

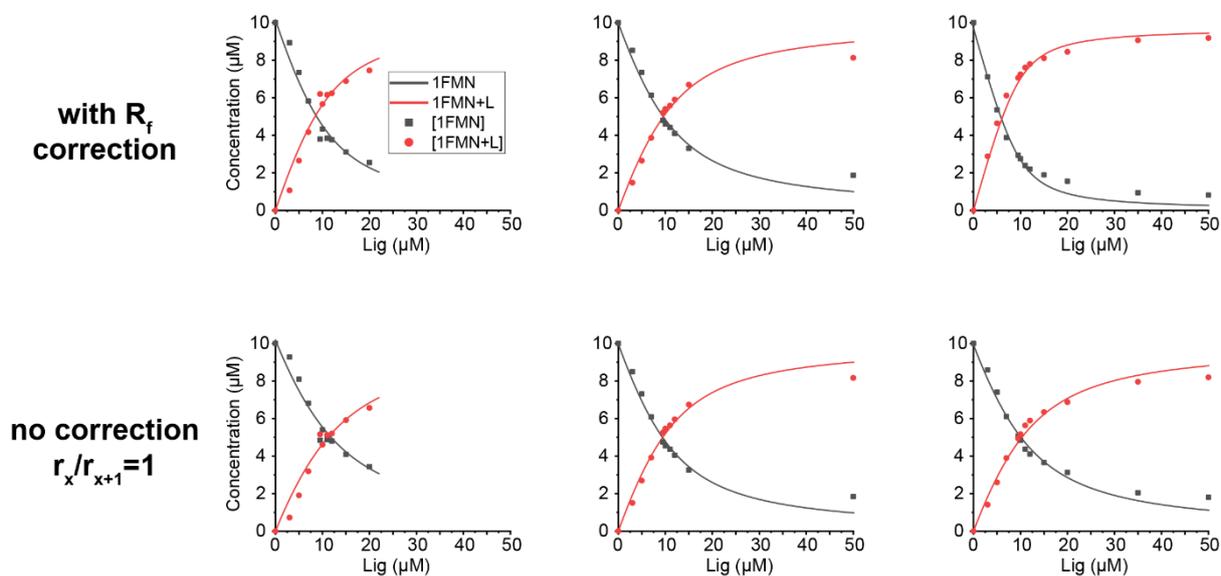


Figure 106 – Replicas of MS titration binding isotherm of 1FMN⁶⁻ in presence of MgCl_2 200 μM in TMAA 100 mM and NH_4OAc 100 mM. Corrected concentrations and assumption of equal response factor ($r_x/r_{x+1}=1$) are compared.

7.2.3 Malachite green aptamer - ITC

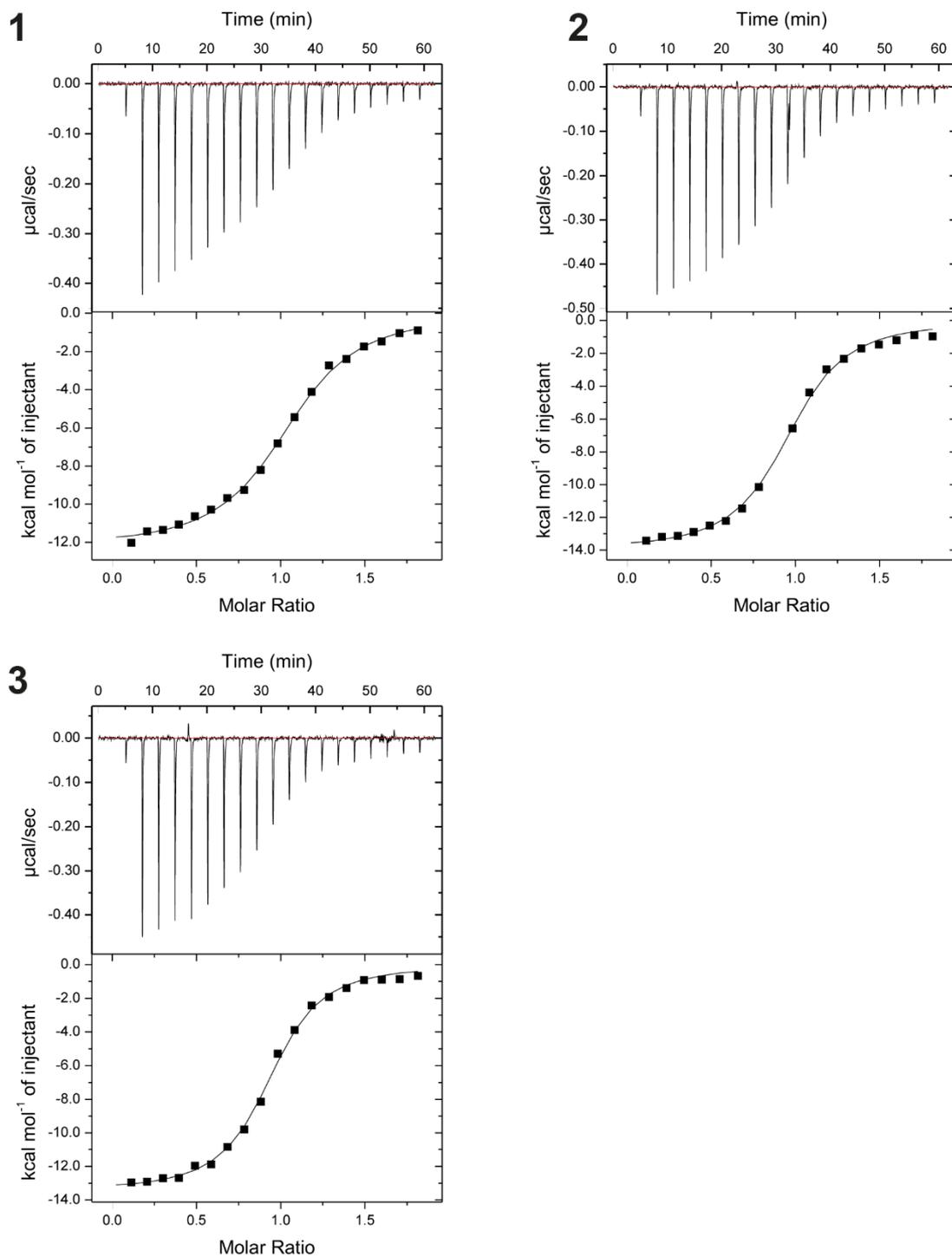


Figure 107 - MGA ITC NH_4OAc 100 mM [RNA]= 15.35 μM [Lig]=142.3 μM . Injection method n.II at p. 27. See Table 46 for fit output.

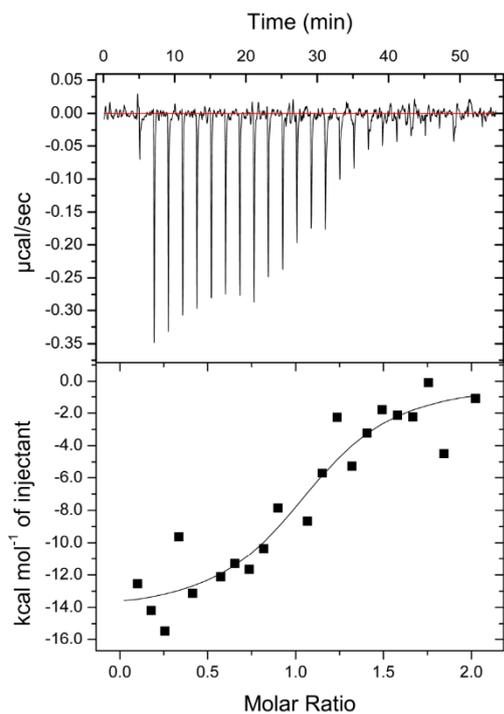
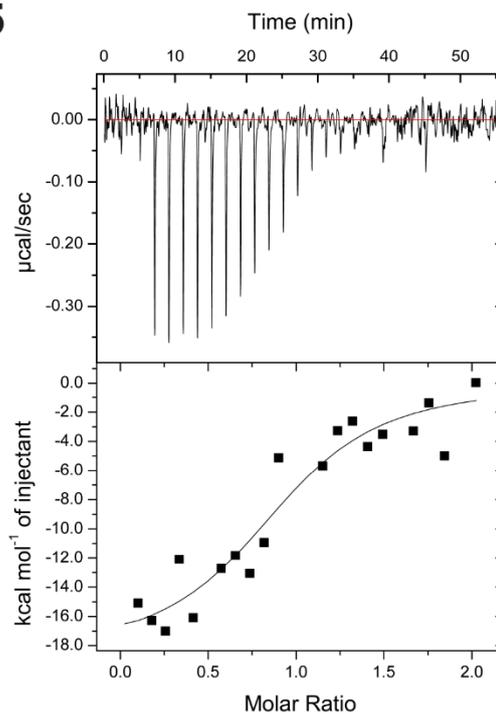
4**5**

Figure 108 - MGA ITC NH_4OAc 100mM [RNA]= 15.19 μM [Lig]=139.2 μM . Injection method n.I at p. 27. See Table 46 for fit output.

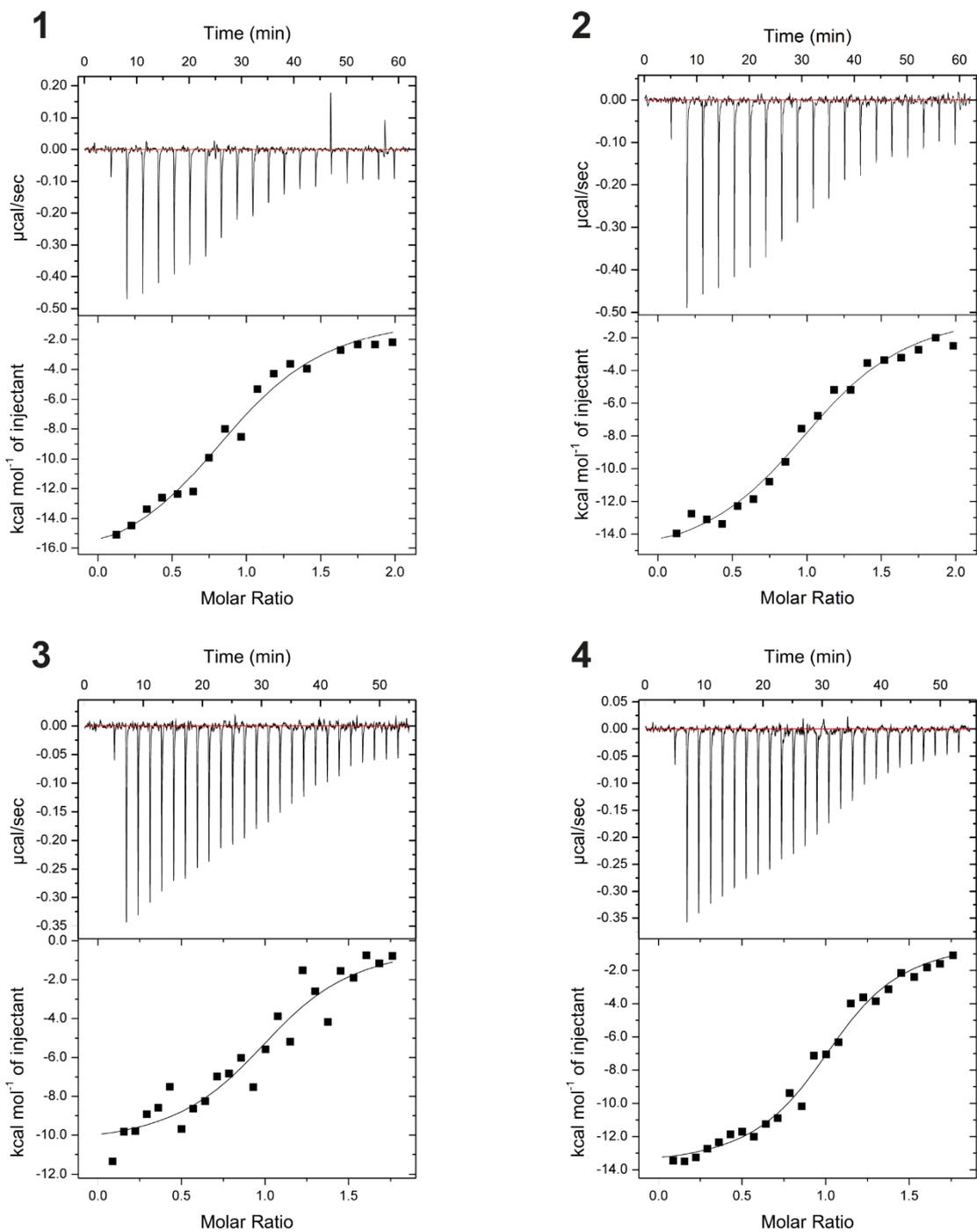


Figure 109 - - MGA ITC TMAA 100 mM 1-2. [RNA]= 15.02 μM [Lig]=135.1 μM with method 1 and 3-4. [RNA]= 15.20 μM [Lig]=152.3 μM with method 2 Second Series at p. 27. See Table 47 for fit output.

ITC MGA vs Ligand in NH₄OAc 100 mM

Replica 1-3 were acquired with method (II) (Figure 107) and method (I) for replica 4-5 (Figure 108), reported at p. 27. Uncertainties reported come from the one binding site fitting model of Origin 7, embedded in ITC200 Microcal.

Table 46 – Replicas of MGA in NH₄OAc 100 mM. [MGA] = 15.35 μM and [MG] = 142.3 μM (replica 1-3); [MGA] = 15.19 μM and [MG] = 139.2 μM (replica 4-5); *a*. K_D has been derived from K_A. Run temperature for each replica is 298 K.

#replica	K _A	K _D (μM) ^a	N	ΔH ⁰ (kcal/mol)	ΔS ⁰ (cal/mol/K)
1	1.46E+06 ± 1.04E+05	0.68 ± 0.05	1.04 ± 0.01	-12.24 ± 0.12	-12.9
2	2.12E+05 ± 1.86E+05	0.47 ± 0.04	0.95 ± 0.01	-14.00 ± 0.15	-18.0
3	2.57E+05 ± 2.27E+05	0.39 ± 0.03	0.92 ± 0.01	-13.48 ± 0.14	-15.9
4	1.04E+06 ± 5.10E+05	0.96 ± 0.47	1.09 ± 0.06	-14.40 ± 1.16	-20.8
5	5.78E+05 ± 2.80E+05	1.73 ± 0.84	0.95 ± 0.08	-18.55 ± 2.13	-35.9

ITC MGA vs Ligand in TMAA 100 mM

The replicas in Figure 109 were acquired with method (II) for replica 1-3 and method (I) for replica 4-5, reported at p. 27.

Table 47 - Replicas of MGA in TMAA 100 mM. [MGA] = 15.20 μM and [MG] = 152.3 μM (replica 1-2); [MGA] = 15.02 μM and [MG] = 135.1 μM (replica 4-5); *a*. K_D has been derived from K_A. Run temperature for each replica is 298 K.

# replica	K _A	K _D (μM) ^a	N	ΔH ⁰ (kcal/mol)	ΔS ⁰ (cal/mol/K)
1	4.47E+05 ± 1.06E+05	2.24 ± 0.53	0.96 ± 0.04	-17.76 ± 1.10	-33.7
2	5.34E+05 ± 1.00E+05	1.87 ± 0.35	1.09 ± 0.03	-15.91 ± 0.66	-27.2
3	8.77E+05 ± 3.42E+05	1.14 ± 0.44	1.06 ± 0.05	-10.67 ± 0.72	-8.59
4	1.29E+06 ± 2.63E+05	0.78 ± 0.16	1.04 ± 0.02	-13.92 ± 0.42	-18.7

7.2.4 RiboFMN aptamer - ITC

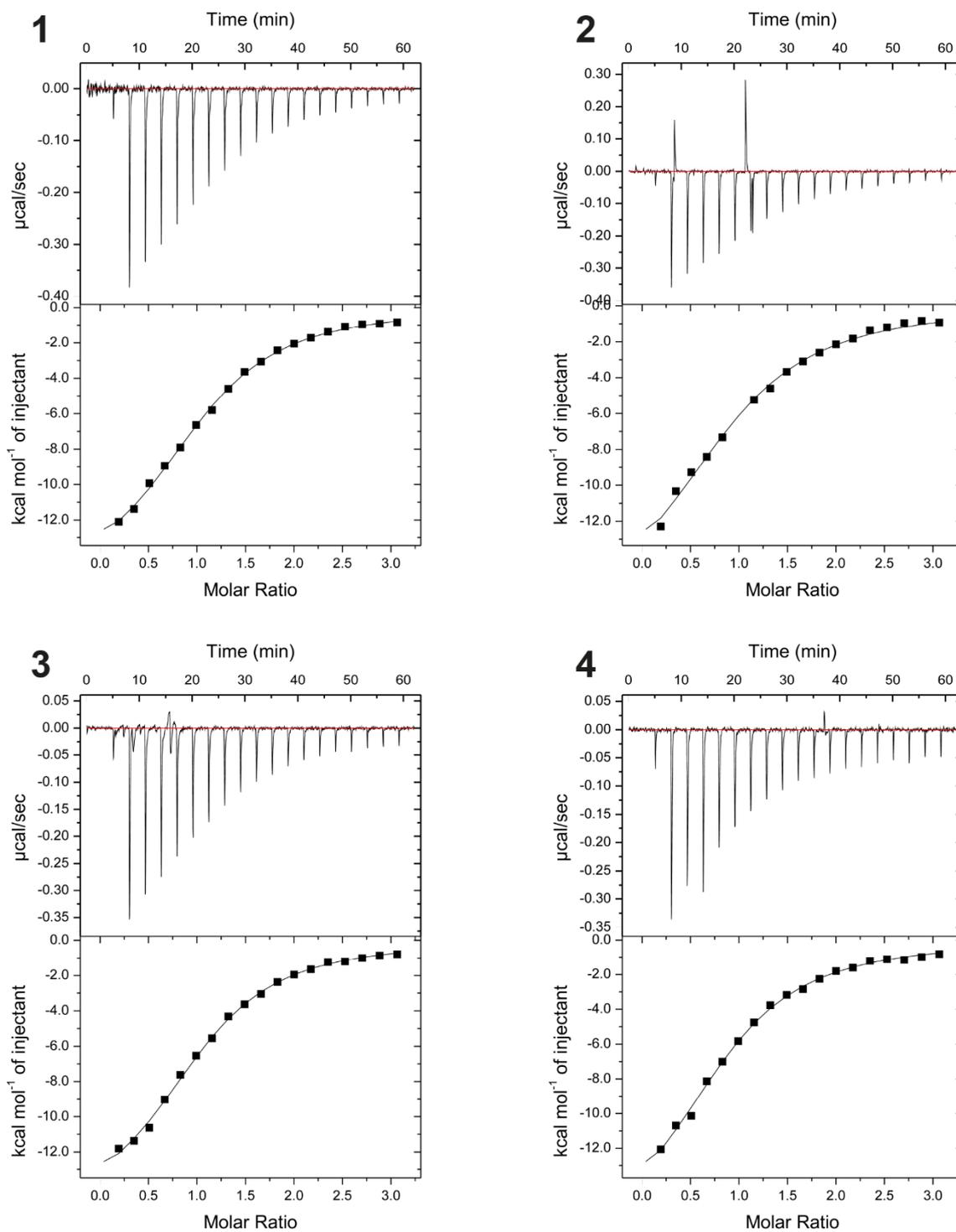


Figure 110 - 1FMN ITC in NH_4OAc 100mM + MgCl_2 200 μM , $[\text{RNA}] = 10.20 \mu\text{M}$ $[\text{Lig}] = 159.6 \mu\text{M}$. Injection method 1 (p. 27). See Table 49 for fit output. Note: the spike in second run is the typical signal of an air bubble into the cell.

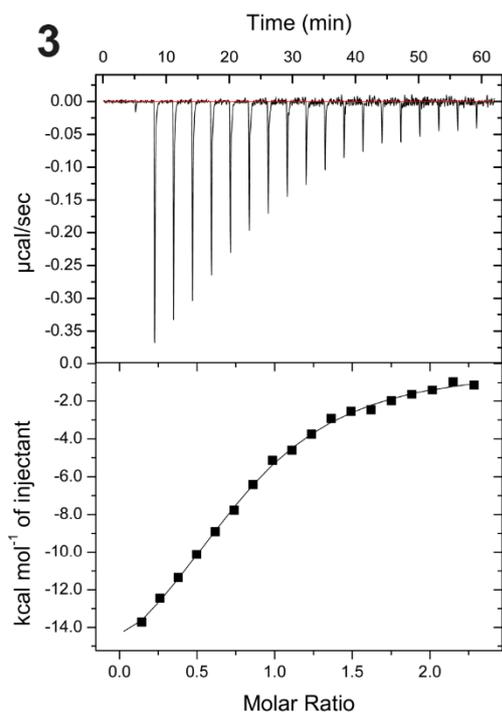
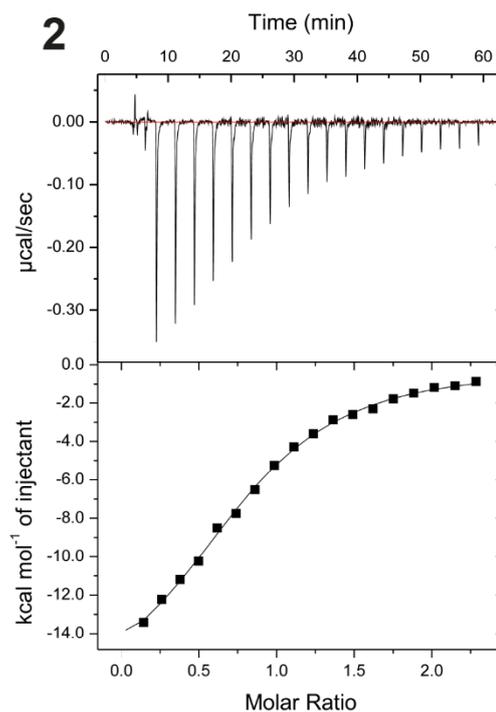
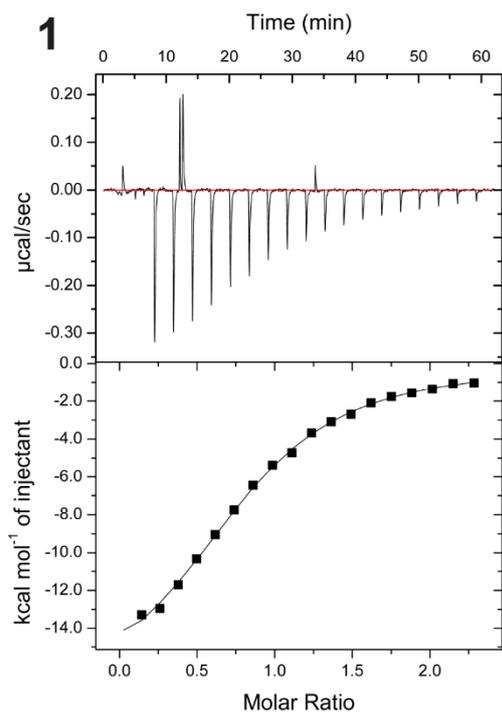


Figure 111 - 1FMN ITC in TMAA 100 mM + MgCl₂ 200 µM, [RNA]=10.19 µM [Lig]=150.0 µM. Injection method 1 at p. 27. See Table 50 for fit output.

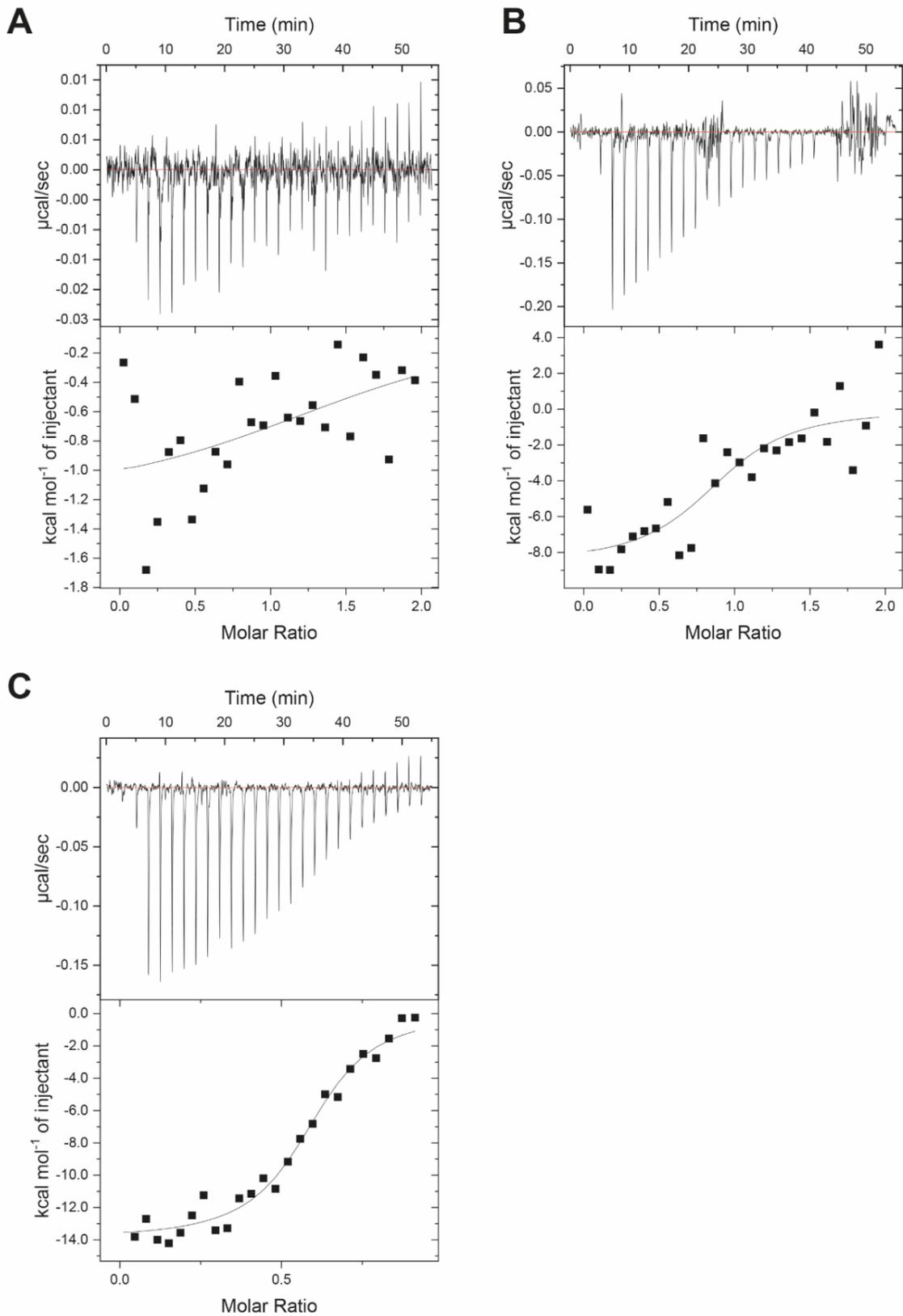


Figure 112 - 1FMN ITC test in Ammo 100mM without Mg [RNA]=15.10 µM [Lig]=149.5 µM; Qualitative tests adding 200µM and 1000 µM Mg²⁺ directly into the initial solution. Injection method p. 27. See Table 48 for fit output.

ITC 1FMN vs Ligand in NH₄OAc 100 mM + Mg²⁺ 0/0.2/1 mM

The replica in Figure 112 were acquired with method (I) reported at p. 27. Uncertainties reported come from the one binding site fitting model of Origin 7, embedded in ITC200 Microcal.

Table 48 - Replicas of 1FMN in NH₄OAc 100 mM at three concentrations of MgCl₂. [1FMN] = 15.10 μM and [riboFMN] = 149.5 μM; *a*. K_D has been derived from K_A; *n.d.* no binding detected; Run temperature for each replica is 298 K.

[MgCl ₂]	K _A	K _D (μM) ^a	N	ΔH ⁰ (kcal/mol)	ΔS ⁰ (cal/mol/K)
0 μM	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>	<i>n.d.</i>
200 μM	8.69E+05 ± 7.36E+05	1.15 ± 3.45	0.92 ± 0.01	-8.59 ± 1.34	-1.63
1000 μM	1.24E+06 ± 2.51E+05	0.81 ± 0.17	1.42 ± 0.01	-7.44 ± 0.34	-16.5

ITC 1FMN vs Ligand in NH₄OAc 100 mM + Mg²⁺ 200 μM

The replica in Figure 110 were acquired with method (I) reported at p. 27.

Table 49 - Replicas of 1FMN in NH₄OAc 100 mM and MgCl₂ 200 μM. [1FMN] = 10.12 μM and [riboFMN] = 159.6 μM; *a*. K_D has been derived from K_A; Run temperature for each replica is 298 K.

# replica	K _A	K _D (μM) ^a	N	ΔH ⁰ (kcal/mol)	ΔS ⁰ (cal/mol/K)
1	3.61E+05 ± 1.95E+04	2.77 ± 0.15	1.04 ± 0.02	-15.88 ± 0.33	-27.8
2	2.48E+05 ± 2.76E+04	4.03 ± 0.45	0.95 ± 0.04	-14.40 ± 1.08	-35.0
3	3.48E+05 ± 2.67E+04	2.87 ± 0.22	1.03 ± 0.02	-18.55 ± 0.41	-27.3
4	2.89E+05 ± 2.34E+04	3.46 ± 0.22	0.88 ± 0.03	-18.55 ± 0.75	-35.0

ITC 1FMN vs Ligand in TMAA 100 mM + Mg²⁺ 200 μM

The replica in Figure 111 were acquired with method (I) reported at p. 27.

Table 50 - Replicas of 1FMN in TMAA 100 mM at and MgCl₂ 200 μM. [1FMN] = 10.19 μM and [riboFMN] = 150.0 μM; *a*. K_D has been derived from K_A; Run temperature for each replica is 298 K.

# replica	K _A	K _D (μM) ^a	N	ΔH ⁰ (kcal/mol)	ΔS ⁰ (cal/mol/K)
1	3.15E+05 ± 3.25E+04	3.18 ± 0.328	1.040 ± 0.034	-18.43 ± 0.38	-36.7
2	2.48E+05 ± 2.76E+04	3.21 ± 0.27	0.991 ± 0.030	-19.03 ± 0.71	-38.7
3	3.48E+05 ± 2.67E+04	3.69 ± 0.40	0.975 ± 0.040	-19.75 ± 1.04	-41.4

7.2.5 Aminoglycoside Aptamers 1NEM and 1TOB

1NEM

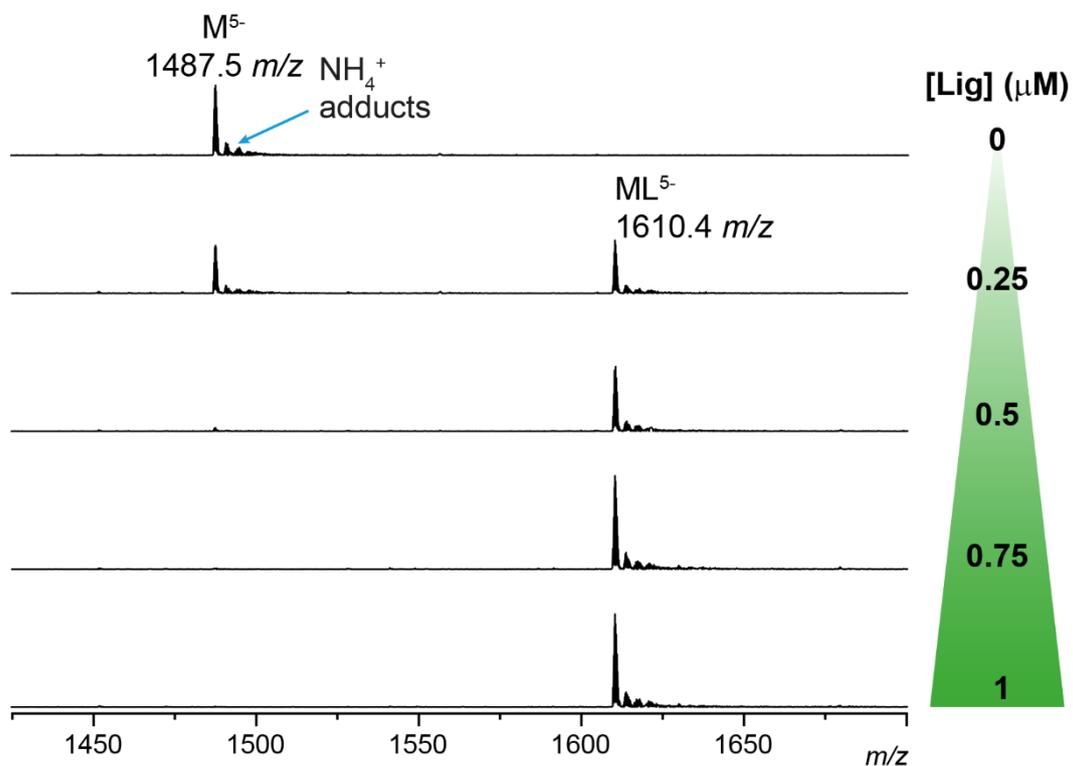


Figure 113 – MS titration of 1NEM 0.5 μM (dT₆ 0.5 μM) with neomycin up to 1 μM in NH₄OAc 100 mM.

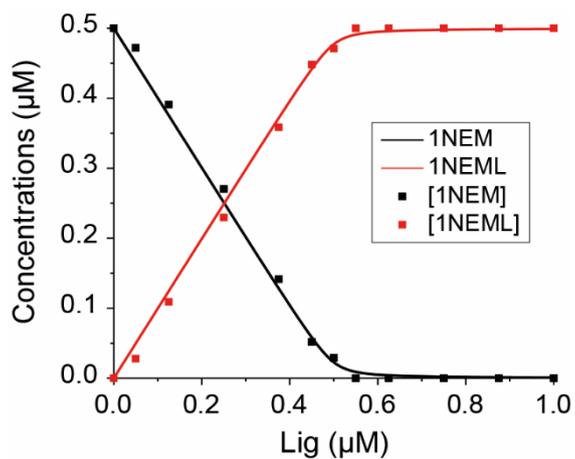


Figure 114 - 1NEM aptamer titrated at 0.5 μM in NH₄OAc 100 mM. The non-linear portion of the binding isotherm is almost absent. The estimated K_D is 1 nM.

1TOB

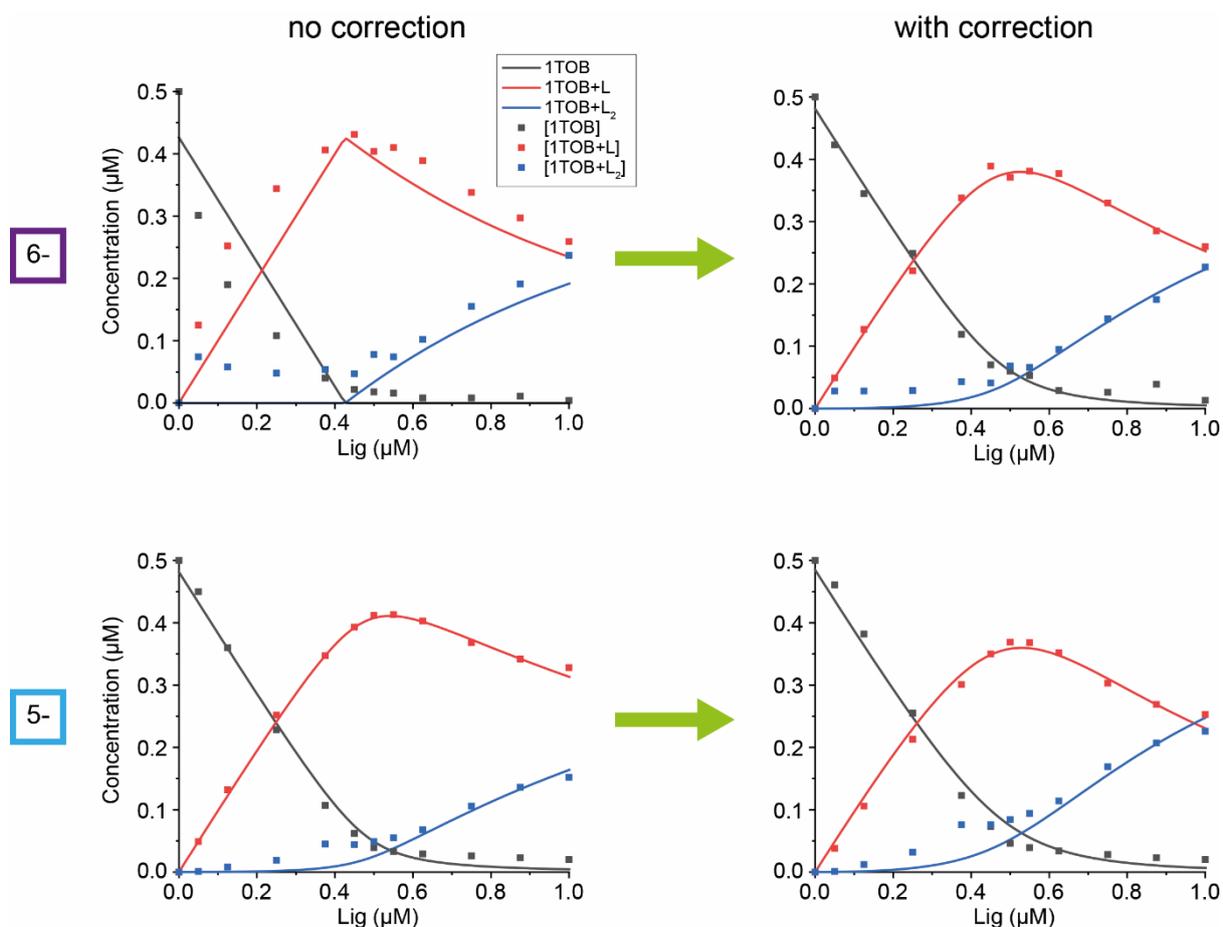


Figure 115 – Binding isotherm for 1TOB aptamer for charge states 6⁻ and 5⁻. On the left is assumed that ratio of relative response factor is equal to 1. On the right, the relative response factor correction is applied. Titration obtained at 1TOB 0.5 µM, tobramycin up to 1 µM in NH₄OAc 100 mM.

Table 51 – Dissociation constant obtained by DynaFit without (assuming R=1) and with correction for relative response factors. *a*. uncertainties are obtained from DynaFit, based only on a monte-carlo resampling of the data points.

K _D (nM)	6 ⁻		5 ⁻	
	1TOB+L	1TOB+L ₂	1TOB+L	1TOB+L ₂
No correction (<i>r_x/r_{x+1}</i> =1)	0.18 ± 0.7 ^a	380 ± 130 ^a	2.6 ± 0.8 ^a	338 ± 40 ^a
Corrected	3.8 ± 1.0 ^a	201 ± 32 ^a	2.4 ± 0.8 ^a	380 ± 48 ^a

Table 52 - 1NEM MS titration m/z integration ranges.

1NEM MS titration in NH₄OAc 100 mM		
m/z_i	m/z_f	Ion
879.8	882.4	Background T6 ²⁻
876.8	879.4	T6 ²⁻
1486.5	1516.5	1NEM ⁵⁻
1450	1480	Background 1NEM ⁵⁻
1609.1	1639.1	1NEM+L ⁵⁻
1450	1480	Background 1NEM+L ⁵⁻

Table 53 – 1TOB MS titration m/z integration ranges

1TOB MS titration in NH₄OAc 100 mM		
m/z_i	m/z_f	Ion
879.8	882.4	Background T6 ²⁻
876.8	879.4	T6 ²⁻
1486.5	1516.5	1TOB ⁶⁻
1435.5	1447.5	Background 1TOB ⁶⁻
1414.5	1426.5	1TOB+L ⁶⁻
1513	1525	Background 1TOB+L ⁶⁻
1414.5	1426.5	1TOB+2L ⁶⁻
1591.2	1603.2	Background 1TOB+2L ⁶⁻
1722.5	1752.5	1TOB ⁵⁻
1691.5	1721.5	Background 1TOB ⁵⁻
1816	1846	1TOB+L ⁵⁻
1691.5	1721.5	Background 1TOB+L ⁵⁻
1909.8	1939.8	1TOB+2L ⁵⁻
1691.5	1721.5	Background 1TOB+2L ⁵⁻

7.3 IM-MS

7.3.1 IM-MS of quantified RNA aptamers

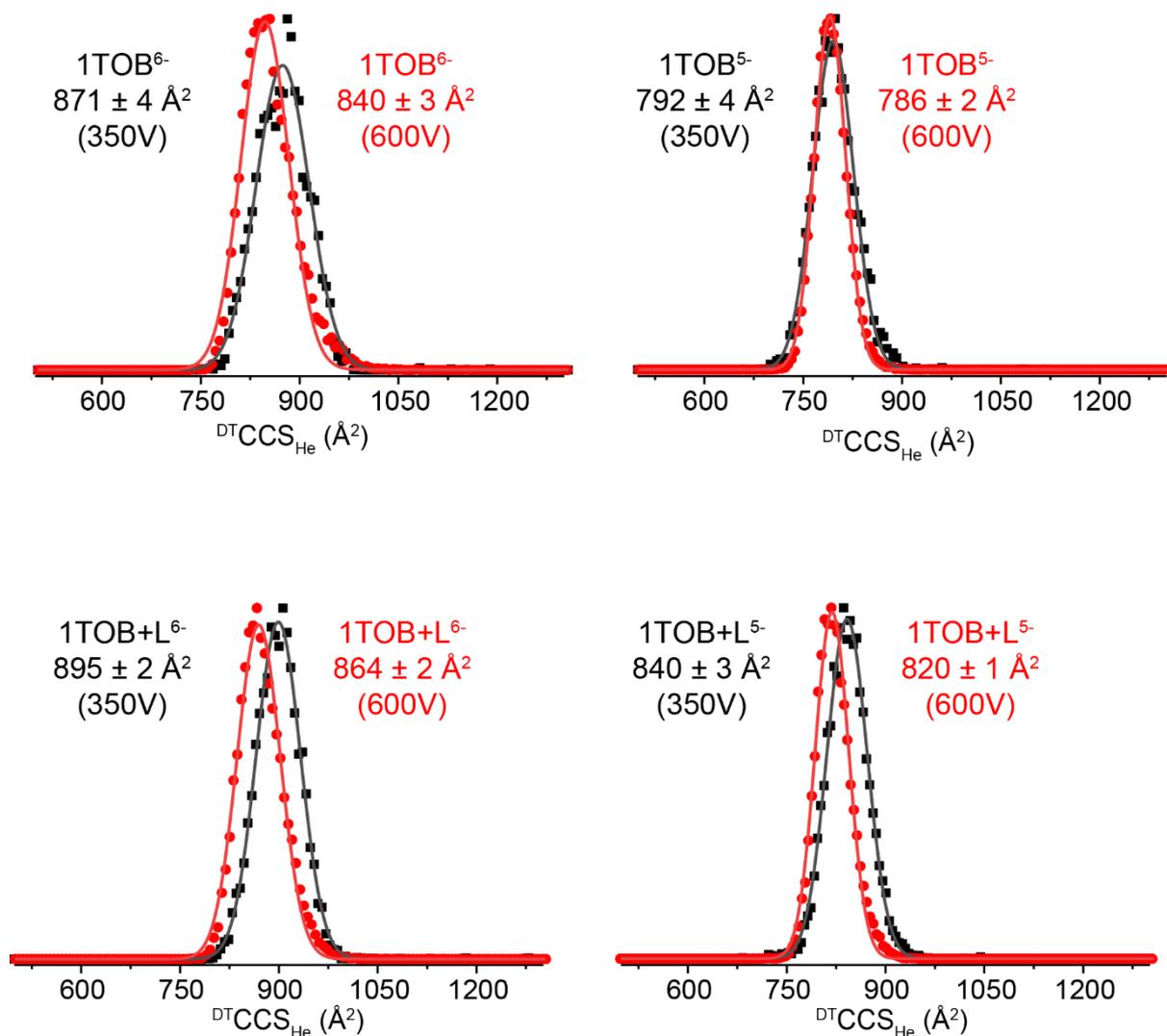


Figure 116 - $^{DT}CCS_{He}$ for charge states 6- and 5- of 1TOB aptamer and its complex with tobramycin. Black curves are at low activation (Fragmentor 350V), red curves are at high activation (fragmentor 600V).

Table 54 - t_a , t_0 , K_0 and CCS value, with their standard deviation for 1NEM, 1TOB and 1FMN aptamers at fragmentor 350V. Only t_a at $\Delta V=390.5$ V is shown; # are the number of replica.

Ion	MW (Da)	z	T (K)	t_a (ms)	t_0 (ms)	K_0 (cm ² /Vs)	^{DT} CCS _{He} (Å ²)	#
[1NEM-zH+]z-	7442.5	5-	298.3 ± 0.4	23.76 ± 0.22	3.95 ± 0.14	3.695 ± 0.029	726 ± 6	8
[1NEM+L-zH+]z-	8057.1	5-	298.4 ± 0.4	24.94 ± 0.20	4.13 ± 0.13	3.519 ± 0.031	762 ± 7	8
[1TOB-zH+]z-	8624.2	6-	298.4 ± 0.5	24.07 ± 0.06	4.25 ± 0.08	3.693 ± 0.015	871 ± 4	4
		5-	298. ± 0.5	26.34 ± 0.05	4.74 ± 0.04	3.386 ± 0.020	792 ± 4	4
[1TOB+L-zH+]z-	9091.7	6-	298.5 ± 0.5	24.78 ± 0.02	4.43 ± 0.03	3.592 ± 0.011	895 ± 2	3
		5-	298.5 ± 0.5	27.79 ± 0.02	4.87 ± 0.01	3.190 ± 0.013	840 ± 3	3
[1FMN-zH+]z-	11337.9	6-	298.6 ± 0.7	27.12 ± 0.56	4.39 ± 0.05	3.212 ± 0.076	1001 ± 24	5
[1FMN+1Mg2+-(2+z)H+]z-	11817.6		298.5 ± 0.6	27.66 ± 0.20	4.42 ± 0.19	3.143 ± 0.021	1022 ± 7	3
[1FMN+2Mg2+-(4+z)H+]z-	11841.9		298.5 ± 0.6	27.53 ± 0.17	4.48 ± 0.04	3.157 ± 0.017	1017 ± 6	3
[1FMN+3Mg2+-(6+z)H+]z-	11866.2		298.5 ± 0.6	27.28 ± 0.14	4.54 ± 0.18	3.200 ± 0.026	1003 ± 8	3
[1FMN+L-zH+]z-	11452.7	6-	298.8 ± 0.7	27.36 ± 0.46	4.39 ± 0.23	3.180 ± 0.049	1011 ± 15	5
[1FMN+L+1Mg2+-(2+z)H+]z-	11477.0		298.5 ± 0.6	27.96 ± 0.12	4.69 ± 0.09	3.131 ± 0.008	1025 ± 3	3
[1FMN+L+2Mg2+-(4+z)H+]z-	11501.3		298.5 ± 0.6	28.05 ± 0.22	4.59 ± 0.12	3.101 ± 0.034	1035 ± 11	3
[1FMN+L+3Mg2+-(6+z)H+]z-	11525.6		298.5 ± 0.6	28.09 ± 0.18	4.71 ± 0.12	3.111 ± 0.013	1032 ± 4	3

Table 55 - t_a , t_0 , K_0 and CCS value, with their standard deviation for 1NEM, 1TOB, 1FMN and MGA aptamers at fragmentor 600V. Only t_a at $\Delta V=390.5$ V is shown; # are the number of replica.

Ion	MW (Da)	z	T (K)	t_a (ms)	t_0 (ms)	K_0 (cm ² /Vs)	^{DT} CCS _{He} (Å ²)	#
[1NEM-zH+]z-	7442.5	5-	298.1 ± 0.2	23.65 ± 0.03	4.02 ± 0.04	3.732 ± 0.012	719 ± 2	4
[1NEM+L-zH+]z-	8057.1	5-	298.5 ± 0.3	24.78 ± 0.02	4.20 ± 0.04	3.561 ± 0.012	753 ± 2	3
[1TOB-zH+]z-	8624.2	6-	298.4 ± 0.5	23.39 ± 0.08	4.27 ± 0.03	3.832 ± 0.016	840 ± 3	4
		5-	298.4 ± 0.5	26.08 ± 0.02	4.63 ± 0.02	3.413 ± 0.010	786 ± 2	4
[1TOB+L-zH+]z-	9091.7	6-	298.3 ± 0.5	23.90 ± 0.05	4.25 ± 0.06	3.725 ± 0.010	864 ± 2	3
		5-	298.3 ± 0.5	27.17 ± 0.01	4.80 ± 0.01	3.272 ± 0.007	820 ± 1	3
[1FMN-zH+]z-	11337.9	6-	298.9 ± 0.2	25.38 ± 0.01	4.00 ± 0.01	3.413 ± 0.006	942 ± 1	4
[1FMN+1Mg2+-(2+z)H+]z-	11817.6		298.5 ± 0.6	26.15 ± 0.39	4.34 ± 0.13	3.348 ± 0.038	959 ± 11	3
[1FMN+2Mg2+-(4+z)H+]z-	11841.9		298.5 ± 0.6	26.00 ± 0.02	4.22 ± 0.12	3.338 ± 0.024	962 ± 7	3
[1FMN+3Mg2+-(6+z)H+]z-	11866.2		298.5 ± 0.6	26.12 ± 0.11	4.36 ± 0.10	3.349 ± 0.025	958 ± 7	3
[1FMN+L-zH+]z-	11452.7	6-	298.8 ± 0.7	25.85 ± 0.13	4.21 ± 0.07	3.370 ± 0.026	954 ± 7	3
[1FMN+L+1Mg2+-(2+z)H+]z-	11477.0		298.5 ± 0.6	26.86 ± 0.33	4.67 ± 0.41	3.270 ± 0.068	982 ± 21	3
[1FMN+L+2Mg2+-(4+z)H+]z-	11501.3		298.5 ± 0.6	26.88 ± 0.11	4.59 ± 0.18	3.252 ± 0.024	987 ± 7	3
[1FMN+L+3Mg2+-(6+z)H+]z-	11525.6		298.5 ± 0.6	26.98 ± 0.05	4.55 ± 0.12	3.251 ± 0.013	987 ± 4	3
[MGA-zH+]z-	12300.5	7-	298.7 ± 0.2	24.02 ± 0.13	3.93 ± 0.28	3.625 ± 0.059	1035 ± 17	3
		6-	298.6 ± 0.2	26.66 ± 0.13	4.27 ± 0.06	3.263 ± 0.026	986 ± 8	5
		5-	298.6 ± 0.2	31.12 ± 0.16	4.94 ± 0.18	2.797 ± 0.027	958 ± 9	4
[MGA-zH+]z-	12665.4	7-	298.9 ± 0.2	24.69 ± 0.13	4.18 ± 0.24	3.558 ± 0.047	1054 ± 14	3
		6-	298.8 ± 0.3	27.39 ± 0.04	4.43 ± 0.06	3.180 ± 0.010	1011 ± 3	4
		5-	298.8 ± 0.3	32.06 ± 0.11	5.12 ± 0.06	2.708 ± 0.012	989 ± 4	4

7.3.2 Tetracycline aptamer

Table 56 - m/z ranges for IM-MS intensities extraction from TCA spectra.

m/z_i	m/z_f	Ion
879.8	882.4	T6 ²⁻
1430	1455	TCA ¹³⁻
1550	1582	TCA ¹²⁻
1582	1597	TCA+Lig ¹²⁻ —1
1597	1620	TCA+Lig ¹²⁻ —2
1695	1730	TCA ¹¹⁻
1735	1747	TCA+Lig ¹¹⁻ —1
1747	1767	TCA+Lig ¹¹⁻ —2
1865	1900	TCA ¹⁰⁻
1905	1922	TCA+Lig ¹⁰⁻ —1
1922	1950	TCA+Lig ¹⁰⁻ —2
2065	2115	TCA ⁹⁻
2115	2135	TCA+Lig ⁹⁻ —1
2135	2160	TCA+Lig ⁹⁻ —2
2330	2405	TCA ⁸⁻
2406	2510	TCA+Lig ⁸⁻

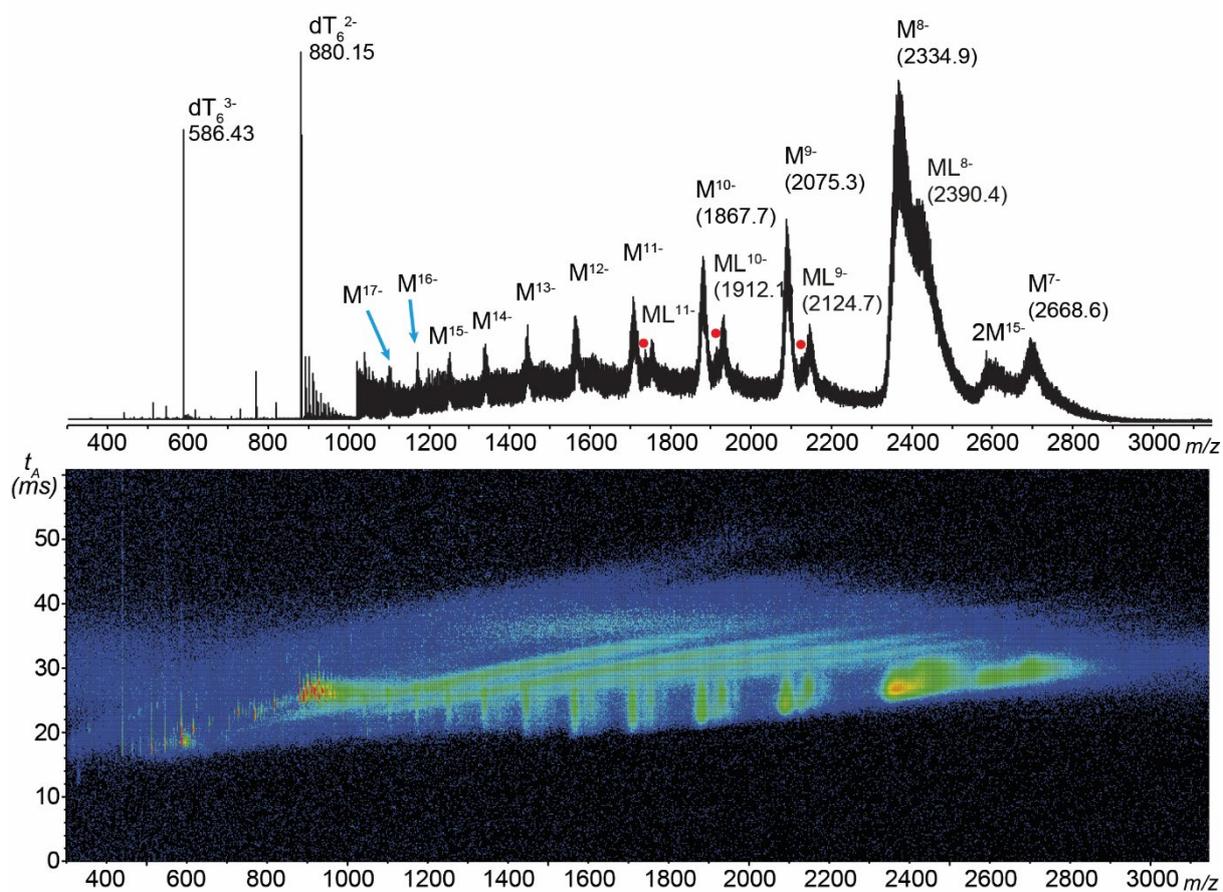


Figure 117 – Full scan IM-MS spectrum of TCA 10 μM and 5 μM tetracycline in TMAA 100 mM with 100 μM MgCl_2 and KCl. The large salts cluster contaminations are visible on the ion mobility spectrum. MS spectrum is 20-fold zoomed from 1050 m/z . In brackets are reported the m/z values of declustered ion. The symbol “●” highlight a complex with only one Mg^{2+} adduct.

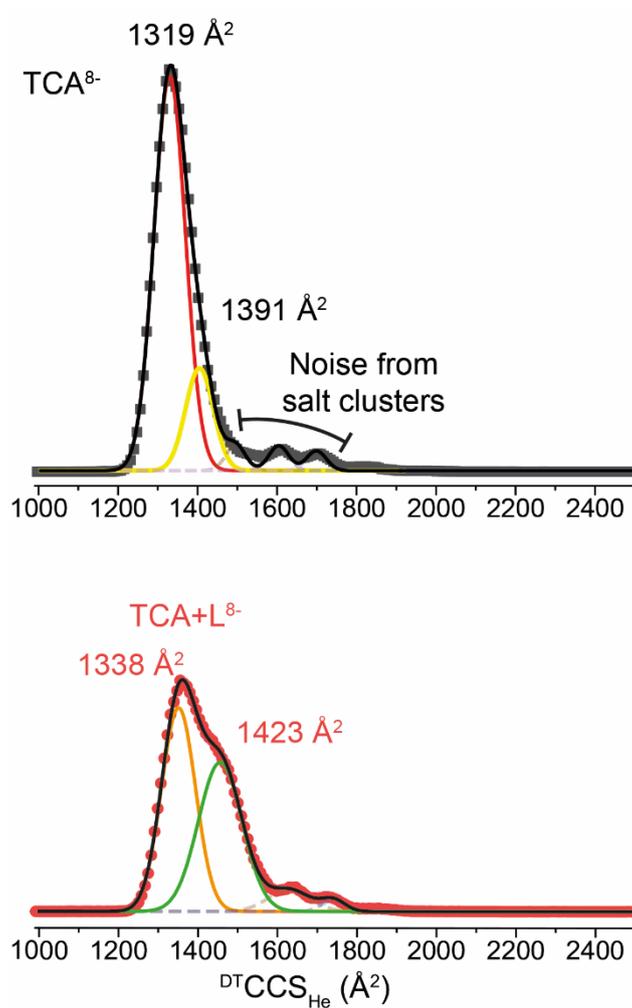


Figure 118 -- Two conformational ensembles are deducibles from CCS distributions of charge states $8--$ of TCA and TCA+L.

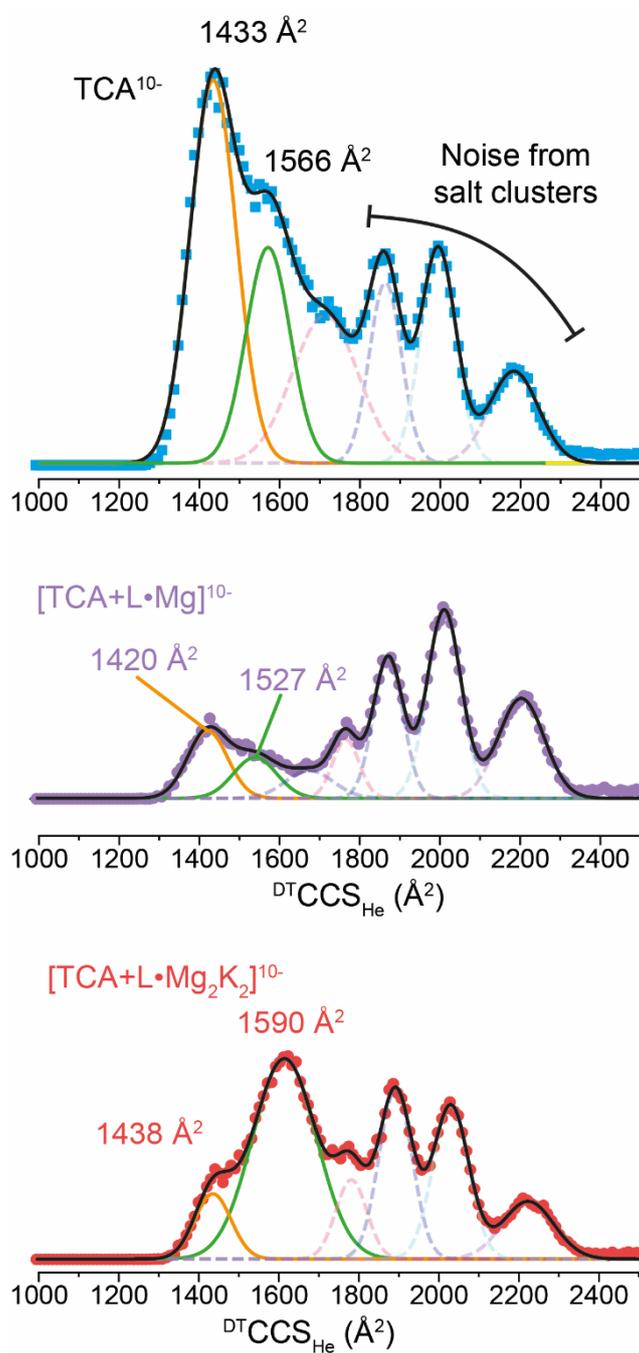


Figure 119 – Two conformational ensembles are deducibles from CCS distributions of charge state 10- TCA and TCA+L, similarly to charge state 9-.

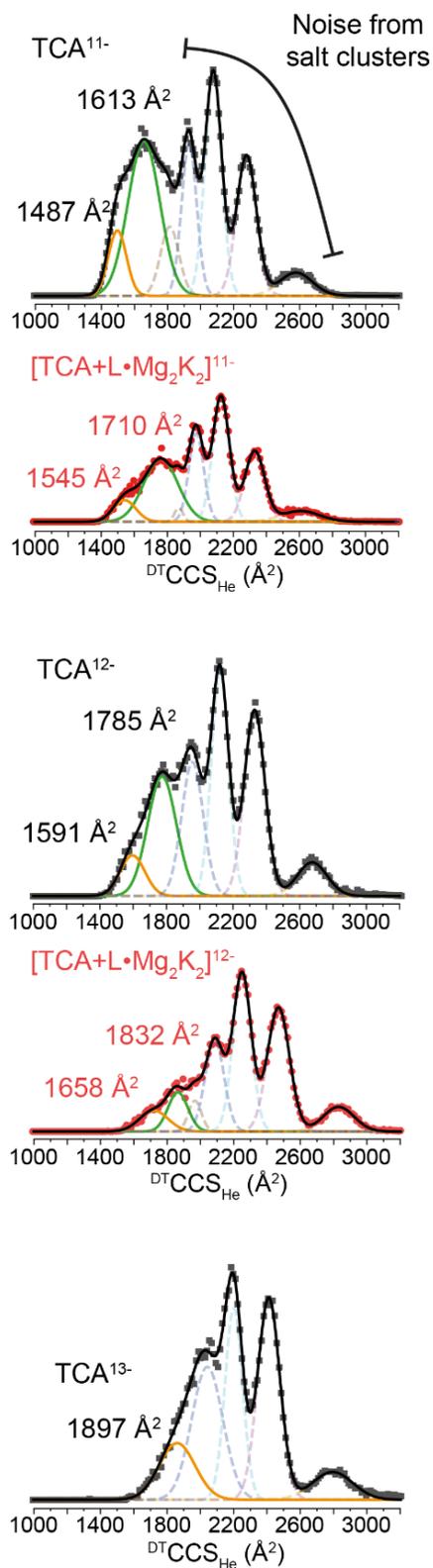


Figure 120 – Two conformational ensembles are deducible from CCS distributions of charge states 11- and 12- of TCA and TCA+L. Charge state 13- low intensity and superimposition with background noise make difficult to identify more than one distribution solely for TCA.

Table 57 - t_a , t_0 , K_0 and CCS value, with standard deviation (over two replica) for tetracycline apramer. Only t_a at $\Delta V=390.5$ V is shown.

Ion	MW (Da)	z	T (K)	t_a (ms)	t_0 (ms)	K_0 (cm ² /Vs)	^{DT} CCS _{He} (Å ²)	
[TCA-zH ⁺] ^{z-}	18687	8-	298.4 ± 0.3	26.81 ± 0.08	4.30 ± 0.14	3.248 ± 0.033	1319 ± 13	
	<i>Minor</i>	8-	298.4 ± 0.3	28.47 ± 0.17	4.67 ± 0.24	3.082 ± 0.059	1391 ± 27	
		9-	298.4 ± 0.3	24.45 ± 0.01	3.89 ± 0.08	3.547 ± 0.018	1359 ± 7	
	<i>Minor</i>	9-	298.4 ± 0.3	26.65 ± 0.06	4.19 ± 0.34	3.240 ± 0.047	1488 ± 22	
		10-	298.4 ± 0.3	23.09 ± 0.05	3.58 ± 0.01	3.738 ± 0.005	1433 ± 2	
	<i>Minor</i>	10-	298.4 ± 0.3	25.02 ± 0.32	3.77 ± 0.19	3.423 ± 0.082	1566 ± 38	
		11-	298.4 ± 0.3	24.13 ± 0.10	4.13 ± 0.15	3.653 ± 0.044	1613 ± 20	
	<i>Minor</i>	11-	298.4 ± 0.3	22.03 ± 0.02	3.63 ± 0.02	3.961 ± 0.003	1487 ± 1	
		12-	298.4 ± 0.3	24.50 ± 0.31	4.22 ± 0.10	3.600 ± 0.044	1785 ± 22	
	<i>Minor</i>	12-	298.4 ± 0.3	22.01 ± 0.32	3.84 ± 0.04	4.040 ± 0.060	1591 ± 24	
		13-	298.4 ± 0.3	23.24 ± 0.41	3.39 ± 0.61	3.677 ± 0.164	1897 ± 85	
	[TCA+Lig+Mg ²⁺ -(2+z)H ⁺] ^{z-}	19156	9-	298.4 ± 0.3	24.43 ± 0.06	3.85 ± 0.02	3.543 ± 0.005	1360 ± 2
		<i>Minor-I</i>	9-	298.4 ± 0.3	25.96 ± 0.01	3.93 ± 0.23	3.311 ± 0.034	1456 ± 15
<i>Minor-II</i>		9-	298.4 ± 0.3	27.19 ± 0.22	4.54 ± 0.17	3.219 ± 0.009	1497 ± 4	
		10-	298.4 ± 0.3	23.19 ± 0.09	3.63 ± 0.03	3.741 ± 0.032	1432 ± 12	
<i>Minor</i>		10-	298.4 ± 0.3	25.39 ± 0.37	3.99 ± 0.01	3.660 ± 0.312	1563 ± 37	
<i>Minor</i>		11-	298.4 ± 0.3	22.08 ± 0.19	3.87 ± 0.12	3.993 ± 0.056	1476 ± 21	
		11-	298.4 ± 0.3	23.83 ± 0.09	4.86 ± 0.01	3.814 ± 0.029	1545 ± 12	

(Table 57 continued)

Ion	MW (Da)	z	T (K)	t _a (ms)	t ₀ (ms)	K ₀ (cm ² /Vs)	^D TCCS _{He} (Å ²)
[TCA+Lig+2Mg ²⁺ +2K ⁺ -(10+z)H ⁺] ^{z-}	19258	8-	298.4 ± 0.3	27.35 ± 0.17	4.54 ± 0.15	3.204 ± 0.047	1338 ± 19
	<i>Minor</i>	8-	298.4 ± 0.3	29.23 ± 0.22	4.99 ± 0.32	3.014 ± 0.069	1423 ± 33
		9-	298.4 ± 0.3	27.08 ± 0.04	4.55 ± 0.30	3.242 ± 0.047	1487 ± 21
	<i>Minor</i>	9-	298.4 ± 0.3	24.80 ± 0.13	3.90 ± 0.15	3.501 ± 0.007	1377 ± 3
		10-	298.4 ± 0.3	25.94 ± 0.07	4.28 ± 0.28	3.369 ± 0.053	1590 ± 25
	<i>Minor</i>	10-	298.4 ± 0.3	23.23 ± 0.04	3.63 ± 0.18	3.725 ± 0.022	1438 ± 8
		11-	298.4 ± 0.3	25.27 ± 0.31	4.01 ± 0.41	3.448 ± 0.102	1710 ± 51
	<i>Minor</i>	11-	298.4 ± 0.3	22.52 ± 0.08	3.45 ± 0.16	3.814 ± 0.009	1545 ± 4
		12-	298.4 ± 0.3	24.48 ± 0.33	3.69 ± 0.63	3.510 ± 0.046	1832 ± 24
	<i>Minor</i>	12-	298.4 ± 0.3	22.23 ± 0.11	3.27 ± 0.61	3.878 ± 0.101	1658 ± 43

7.3.3 CBA - UV-melting profiles

UV-melting experiments for MN4, MN19 and OR8 was also carried in NH_4OAc 100 mM and 2 μM DNA as described at p. 27. We found a melting temperature of $54.9 \pm 0.7^\circ\text{C}$ for MN4 and no melting for MN19 (Figure 121), OR8 and OR7 (Figure 122). In presence of one equivalent of quinine, the melting temperature of MN4 do not change significantly ($55.1 \pm 0.2^\circ\text{C}$), MN19 shows a melting temperature at $31.6 \pm 0.2^\circ\text{C}$ (Figure 121). OR8 and OR7 show no melting (Figure 122). These melting profiles are very similar to those reported by Neves *et al.*^[88] in 140 mM NaCl + 20 mM sodium phosphate buffer (pH 7.4) for which MN4 melts at $57.6 \pm 0.4^\circ\text{C}$, MN19 and OR8 have no melting, whereas in presence of quinine MN19 melts at $35.6 \pm 0.3^\circ\text{C}$, OR8 at $28 \pm 1^\circ\text{C}$ and MN4 is unchanged.

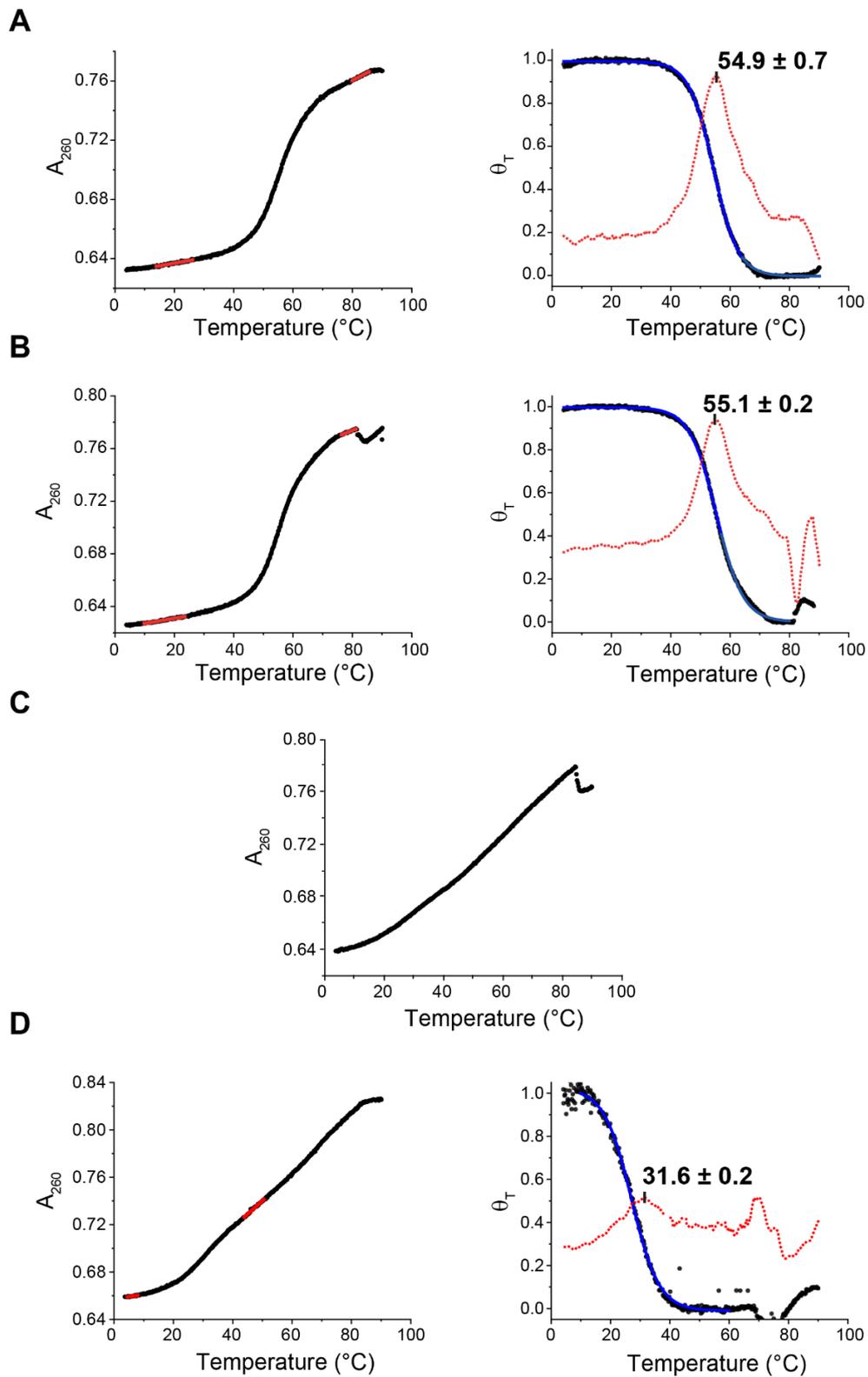


Figure 121- UV-melting plots in NH₄OAc 100 mM of **A.** MN4 2 μ M and **B.** with 2 μ M quinine; **C.** MN19 2 μ M and **D.** with 2 μ M quinine. Baselines for curve fitting are reported in filled red, first derivative of A_{260} is reported in dotted red. OR8 and OR7 sequences didn't show a melting transition in this buffer and temperature range (p.201 - Figure 122)

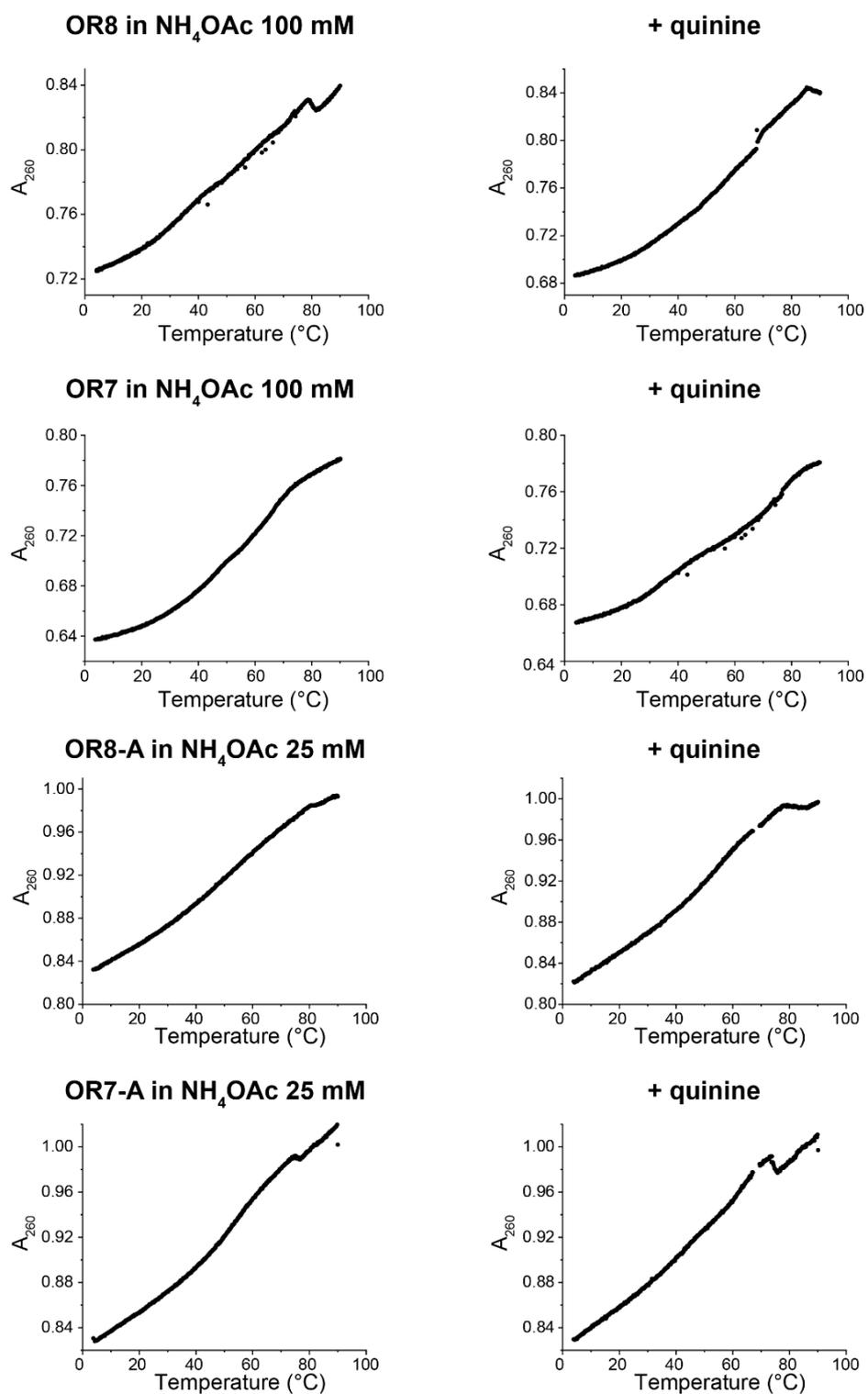
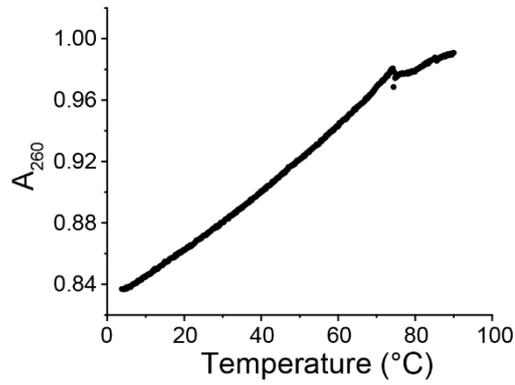
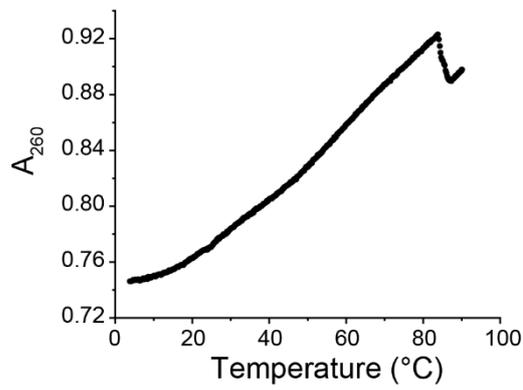


Figure 122 - UV-melting profiles of OR8 and OR7 (2 μM) in NH_4OAc 100 mM. No melting detectable even in presence of 2 μM of quinine.

MN19-A in NH₄OAc 25 mM



MN19-A in NH₄OAc 100 mM



MN19-A +quinine in NH₄OAc 100 mM

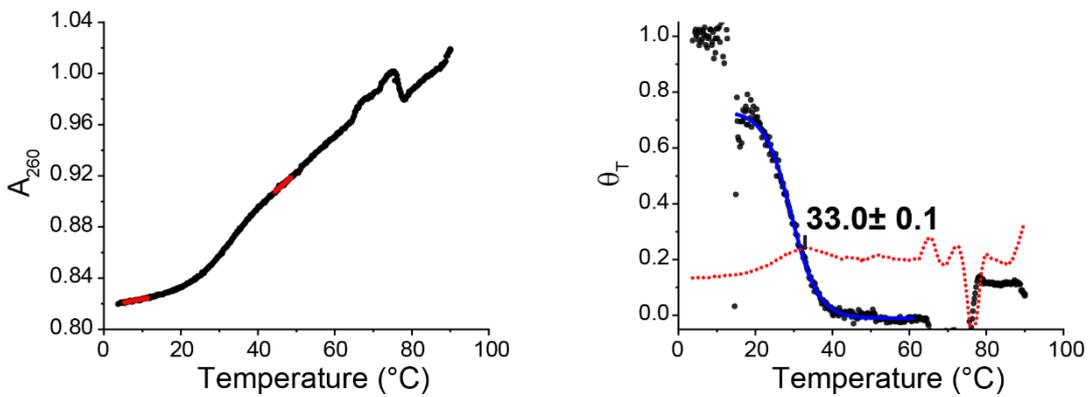


Figure 123 - UV-melting profile of MN19-A (2 μM) in NH₄OAc 25 and 100 mM. Only in presence of quinine (2 μM) a melting at 33 $^{\circ}\text{C}$ is detectable in NH₄OAc 100 mM.

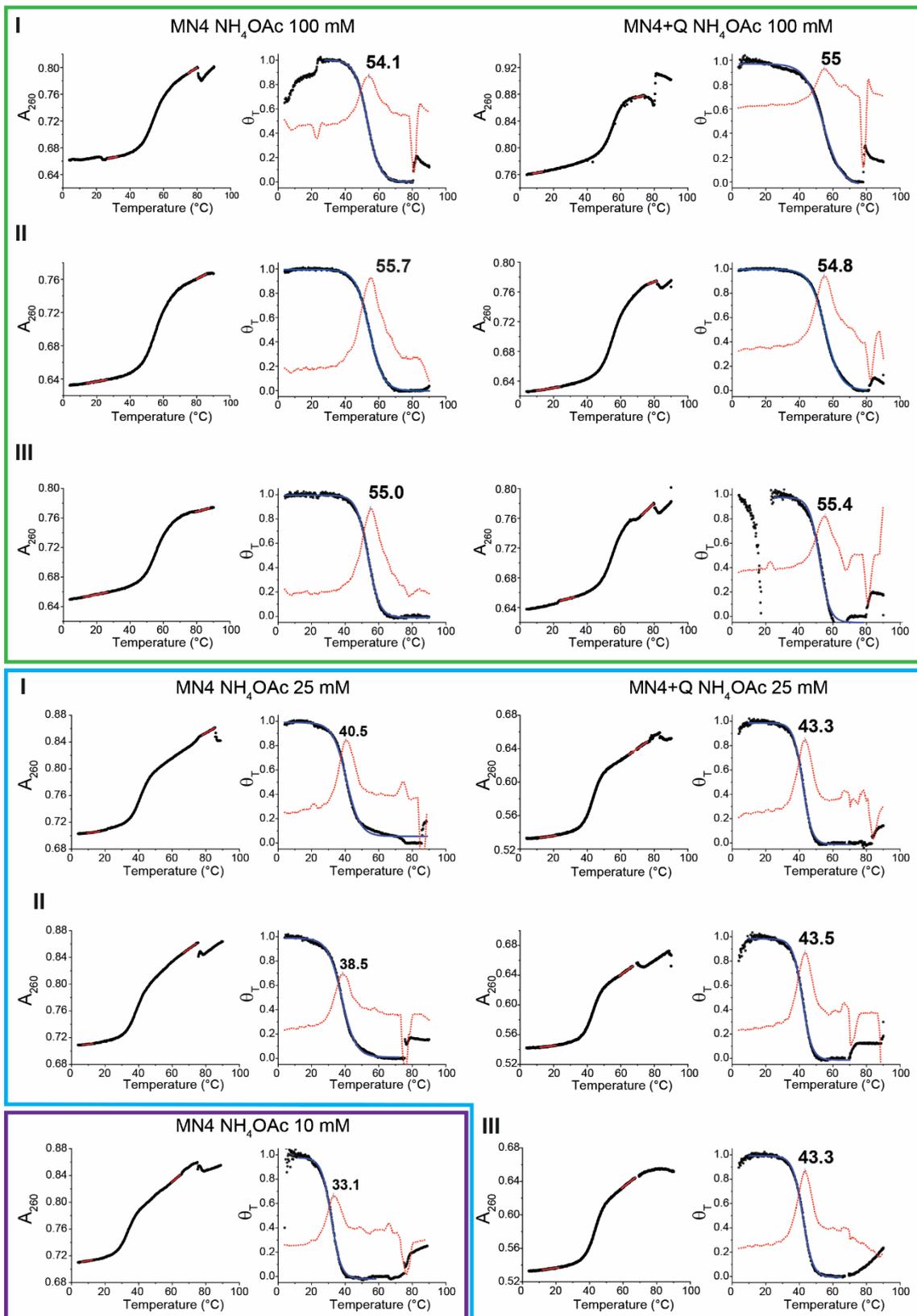


Figure 124 - Replicas of MN4 at different NH_4OAc concentrations, with and without quinine. MN4 at 2 μM and quinine at 2 μM .

7.3.4 CBA - ^1H -NMR of MN4 quinine complex

Aptamer binding mode towards quinine can be verified comparing diagnostic ^1H -NMR resonances of un-modified aptamers, previously assigned in Johnson's Lab, to our CBA+dA constructs in NH_4OAc 25 mM and 1 equivalent of quinine. If NH_4OAc does not alter the binding mode, we should find the same peaks position (or very similar) for quinine binding. Base numbering of Johnson's papers is kept for consistency.

Many peaks are visible as reported from 11.1 ppm (G31 – Stem 3 and junction), 12.4 (G4/9/10/27), 13.1 (T15 - Stem2) and 13.3 ppm (T19/28/32) [87, 155] (Figure 125). These peaks suggest that MN4 is capable to bind quinine in NH_4OAc 25 mM with the expected topology.

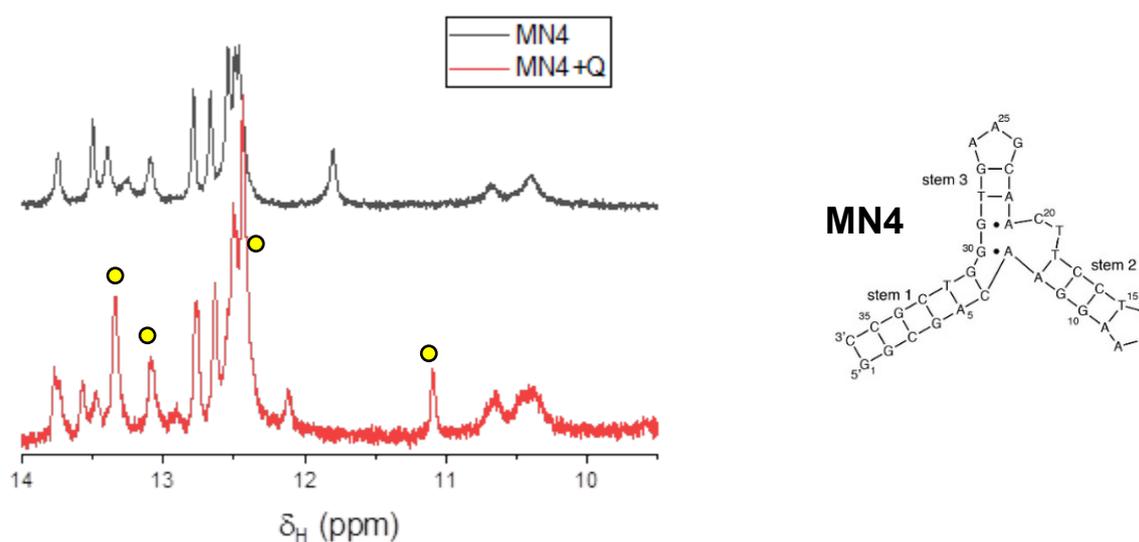


Figure 125 - ^1H -NMR imino resonances of MN4 and MN4 with 1 equivalent of quinine. Characteristic peaks “●” at 11.1, 12.4, 13.1 and 13.3 ppm denote quinine binding in NH_4OAc 25 mM in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) at 278 K.

7.3.5 CBA - Circular Dichroism verify the B-DNA signature of dA constructs

In the case of cocaine binding aptamers we can follow the B-DNA signals, proportional to hairpin formation and stem closing (i.e. MN4 in Figure 63). For an anti-parallel B-DNA we expect two bands, a positive one around 280 nm and a negative band at 245 nm.^[163]

Our aim is to verify if the aptamers MN4, MN19-A, OR8-A and OR7-A show a B-DNA pattern in NH_4OAc 25 mM, if this signature remains after ligand binding (or in other words, to exclude a dramatic change to others structures like A-DNA, Z-DNA, ...), and if the signature is increased (which could indicate additional stem closing upon ligand binding).

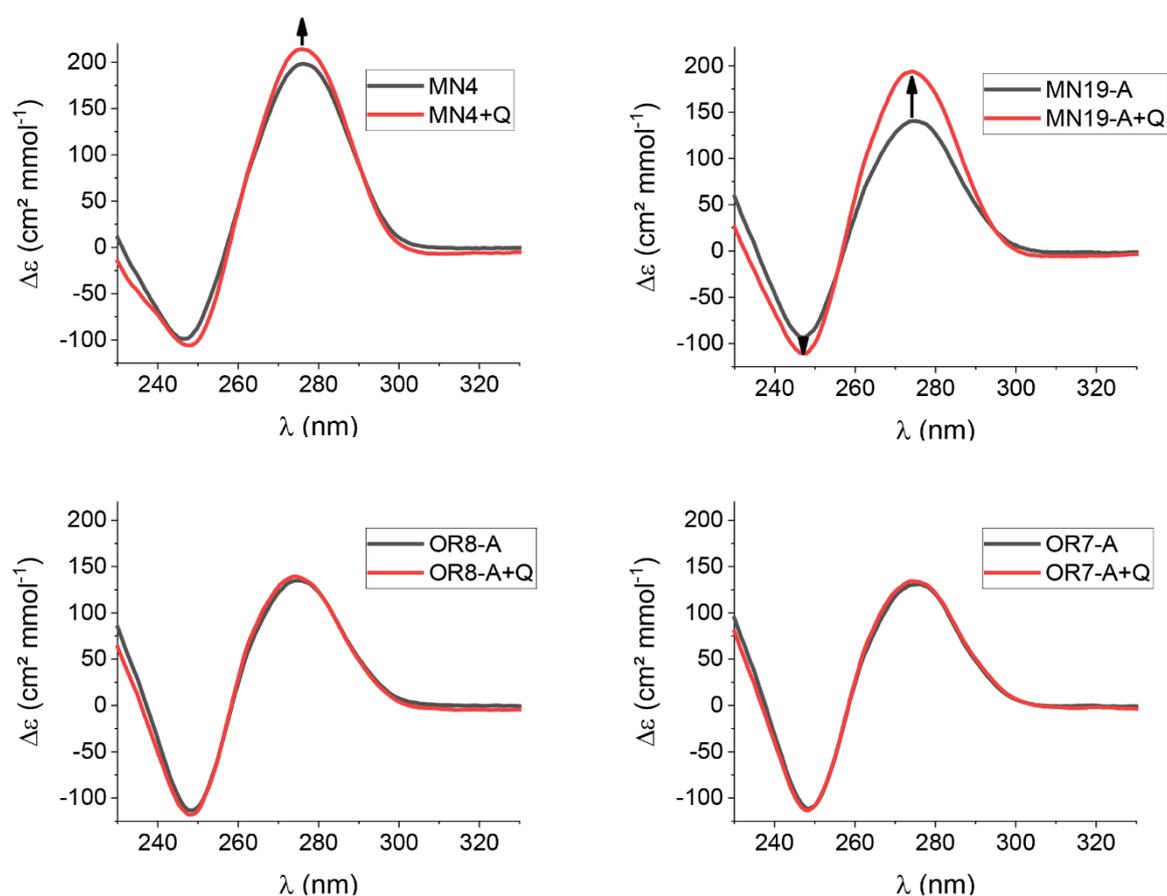


Figure 126 - CD spectra of free MN4, MN19-A, OR8-A and OR7-A (red curves) and in presence of 1 eq of quinine (black curves). Spectra were acquired in NH_4OAc 25 mM at 25°C.

From Figure 126 shows a typical B-DNA pattern for all the sequences. Upon ligand binding we observe an increase of B-DNA pattern for MN19-A, and a little for MN4, whereas for OR8-A and OR7-A the spectra remain unchanged. First we can exclude that quinine induce a transition towards totally different structures.

The low increase on MN4 spectrum upon quinine binding is consistent with a pre-structured aptamer, where the minor changes on bands (i.e. the little increase at 275 nm) could be interpreted as minor structural rearrangements after binding.

The unchanged spectra of OR8-A and OR7-A tell us that in presence of quinine there is no structural change (detectable by CD). This support the hypothesis that these sequences, deprived of stem 1 nucleotides/cut at 3',5' end, in free state have only the two hairpins stem 2 and 3, and binds quinine very weakly. The increase of B-DNA bands amplitude in MN19-A spectrum is similar to the hairpin-duplex transitions^[164]. This increase can be described with the pairing of nucleotides in stem 1 that are now contributing with their relative orientation (like a B-DNA).

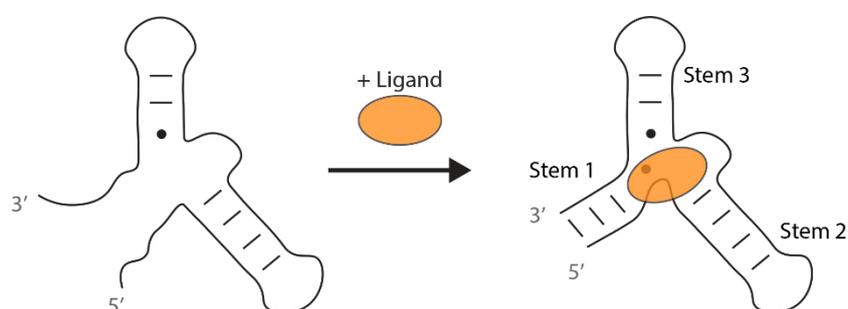


Figure 127 - Hypothetical folding of 3',5' end in B-DNA duplex after quinine (ligand) binding in MN19 and MN19-A proposed in ref ^[87].

The shape of the spectra also changes across the four sequences. This could be due to a different relative orientation between the base pairs that form the structure of these aptamers. For instance, stem 1 is very short in OR8-A and OR7-A and probably they are not represented by the topology proposed for MN4 (Figure 74). By superimposing the CD spectra of the original aptamers (MN4, MN19, OR8, OR7) and adenine-modified (MN19-A, OR8-A, OR7-A) in Figure 128, we notice a similar deeper negative band (245 nm) and steeper spectrum for $\lambda < 240$ nm, common for adenine-modified constructs. This two features could be a contribution of dA overhangs. Although they do not super impose perfectly (275-nm band of OR8 looks higher), we suggest that the adenine modification do not alter dramatically aptamers structure. We deduce that the B-DNA secondary structure character of original aptamers (MN4, MN19, OR8, OR7) is maintained in adenine-modified constructs. . Jointly with UV and NMR comparisons at paragraphs 5.6.5 (p.113), we consider the constructs MN19-A, OR8-A and OR7-A as good model sequences for gas phase comparisons.

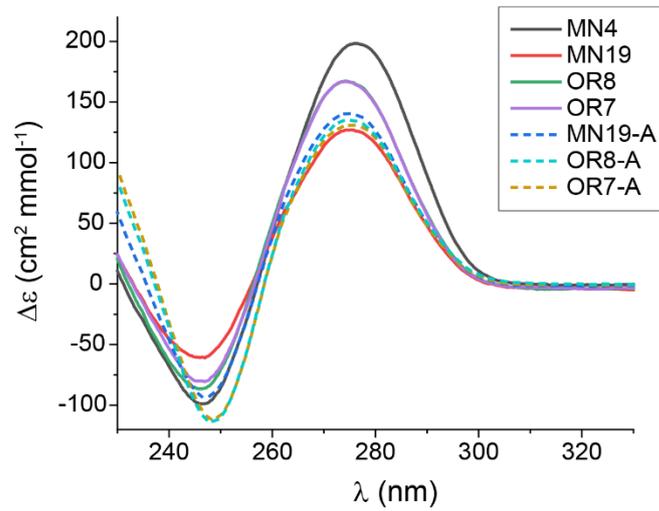
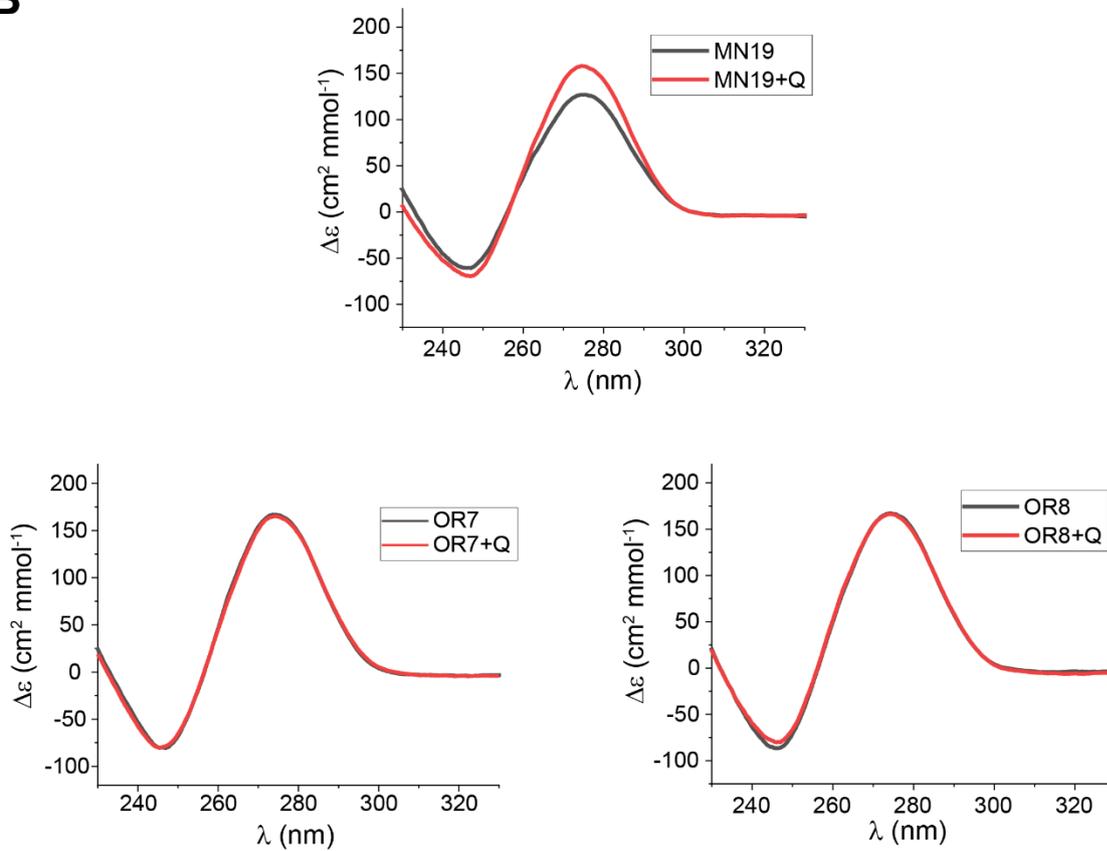
A**B**

Figure 128 – **A**. Superimposition of MN4, MN19, OR8 OR7 to MN19-A, OR8-A and OR7-A show the contribution of extra dA construct to the negative band and lower λ (<240 nm); **B**. Controls on MN19 show a minor increase of band 275 nm upon quinie binding whereas OR8 and OR7 remains unchanged. Spectra were acquired in NH_4OAc 25 mM at 25°C.

7.3.6 CBA - CCS

Table 58 - *m/z* ranges for IM-MS ATDs extraction of DNA cocaine binding aptamers.

<i>m/z_i</i>	<i>m/z_f</i>	Ion	<i>m/z_i</i>	<i>m/z_f</i>	Ion
879.8	882.4	T6 ²⁻	1234.5	1236.5	OR8 ⁷⁻
1389.5	1391	MN4 ⁸⁻	1280.7	1282.7	OR8+Lig ⁷⁻
1430	1432.5	MN4+Lig ⁸⁻	1440.5	1443	OR8 ⁶⁻
1587.8	1590	MN4 ⁷⁻	1494.5	1497	OR8+Lig ⁶⁻
1634	1636.2	MN4+Lig ⁷⁻	1728.5	1732	OR8 ⁵⁻
1852.5	1855.5	MN4 ⁶⁻	1793.8	1797.3	OR8+Lig ⁵⁻
1906.5	1909.5	MN4+Lig ⁶⁻	1393.2	1395	OR8-A ⁸⁻
2223.5	2226.5	MN4 ⁵⁻	1592.5	1594.5	OR8-A ⁷⁻
2288.5	2291.5	MN4+Lig ⁵⁻	1638.5	1641.5	OR8-A+Lig ⁷⁻
1157.5	1159	MN19 ⁸⁻	1858	1860.5	OR8-A ⁶⁻
1197.7	1200.2	MN19+Lig ⁸⁻	1912	1914.5	OR8-A+Lig ⁶⁻
1322.8	1324.8	MN19 ⁷⁻	2230.5	2233	OR8-A ⁵⁻
1369	1372	MN19+Lig ⁷⁻	2294.5	2297	OR8-A+Lig ⁵⁻
1543.5	1546	MN19 ⁶⁻	1394.3	1395.7	OR7-A ⁸⁻
1597.5	1600	MN19+Lig ⁶⁻	1593.6	1595.6	OR7-A ⁷⁻
1852	1855.5	MN19 ⁵⁻	1640.3	1642.3	OR7-A+Lig ⁷⁻
1917.2	1920.5	MN19+Lig ⁵⁻	1859.5	1862	OR7-A ⁶⁻
1392	1394	MN19-A ⁸⁻	1913.5	1916	OR7-A+Lig ⁶⁻
1432.6	1434.6	MN19-A+Lig ⁸⁻	1968	1970.5	OR7-A ⁵⁻
1591.2	1593.2	MN19-A ⁷⁻	2231.5	2234	OR7-A+Lig ⁵⁻
1637.5	1639.5	MN19-A+Lig ⁷⁻			
1856.5	1859.2	MN19-A ⁶⁻			
1910.5	1913.2	MN19-A+Lig ⁶⁻			
2228.5	2231.5	MN19-A ⁵⁻			
2293.5	2295.5	MN19-A+Lig ⁵⁻			

Table 59 - t_a , t_0 , K_0 and CCS value, with their standard deviation, for cocaine binding apramers at fragmentor 350V. Only t_a at $\Delta V=390.5$ V is shown; # is the number of replica.

Ion	MW (Da)	z	T (K)	t_a (ms)	t_0 (ms)	K_0 (cm ² /Vs)	$^{DT}CCS_{He}$ (Å ²)	#	
[MN4-zH+]z-	11128.3	8-	298.4 ± 0.2	24.08 ± 0.28	4.34 ± 0.13	3.706 ± 0.069	1158 ± 22	4	
		7-	298.5 ± 0.3	24.49 ± 0.31	4.41 ± 0.08	3.636 ± 0.047	1032 ± 13	7	
		6-	298.4 ± 0.2	26.90 ± 0.20	4.79 ± 0.09	3.302 ± 0.025	974 ± 7	6	
		5-	298.6 ± 0.2	30.51 ± 0.32	5.50 ± 0.15	2.915 ± 0.053	919 ± 16	2	
[MN4+Q-zH+]z-	11452.72	7-	298.5 ± 0.1	24.00 ± 0.08	4.40 ± 0.29	3.551 ± 0.049	1056 ± 15	4	
		6-	298.5 ± 0.1	27.52 ± 0.06	5.05 ± 0.04	3.247 ± 0.005	990 ± 2	2	
		5-	298.5 ± 0	31.74 ± 0	5.63 ± 0	2.794 ± 0	959 ± 0	1	
[MN19-zH+]z-	11169.4	8-	298.5 ± 0.1	25.35 ± 0.61	4.88 ± 0.11	3.564 ± 0.099	1204 ± 32	3	
		7-	298.5 ± 0.1	22.45 ± 0	4.37 ± 0.24	4.032 ± 0.039	931 ± 9	2	
		6-	298.6 ± 0.1	24.51 ± 0.01	4.37 ± 0.10	3.603 ± 0.058	893 ± 14	3	
		5-	298.5 ± 0.1	28.02 ± 0.03	5.04 ± 0.01	3.183 ± 0.007	842 ± 2	2	
[MN19+Q-zH+]z-	11493.82	7-	298.5 ± 0.2	23.32 ± 0.25	4.40 ± 0.03	3.868 ± 0.043	970 ± 11	2	
		6-	298.6 ± 0.1	25.27 ± 0.03	4.60 ± 0.02	3.536 ± 0.008	910 ± 2	2	
		5-	298.8 ± 0.3	28.95 ± 0.06	4.87 ± 0.38	3.032 ± 0.066	885 ± 19	2	
[MN19-A-zH+]z-	11152.3	8-	298.6 ± 0.1	24.56 ± 0.44	4.65 ± 0.27	3.686 ± 0.028	1163 ± 9	2	
		Main	7-	298.5 ± 0.1	23.59 ± 0.17	4.18 ± 0.10	3.768 ± 0.035	996 ± 9	5
		Second		298.5 ± 0.1	25.79 ± 0.29	4.85 ± 0.27	3.510 ± 0.059	1069 ± 18	5
		6-	298.6 ± 0.1	26.46 ± 0.12	4.63 ± 0.05	3.345 ± 0.010	962 ± 3	3	
		5-	298.6 ± 0.1	30.42 ± 0.32	5.28 ± 0.05	2.902 ± 0.044	924 ± 14	2	
[MN19-A+Q-zH+]z-	11476.72	7-	298.6 ± 0.1	24.47 ± 0.12	4.39 ± 0.11	3.635 ± 0.005	1032 ± 2	2	
		7-	298.6 ± 0.1	26.82 ± 0.50	4.39 ± 0.03	3.419 ± 0.081	1098 ± 26	2	
		6-	298.7 ± 0	27.20 ± 0	4.78 ± 0	3.253 ± 0	988 ± 0	1	
		5-	298.7 ± 0	31.57 ± 0	5.03 ± 0	2.749 ± 0	975 ± 0	1	

(Table 59 continued)

ion	MW (Da)	z	T (K)	t _a (ms)	t ₀ (ms)	K ₀ (cm ² /Vs)	^{DT} CCS _{He} (Å ²)	#	
[OR8-zH+]z-	11169.4	7-	298.5 ± 0.1	23.12 ± 0.52	4.22 ± 0.09	3.872 ± 0.110	970 ± 27	2	
		<i>shoulder</i>	298.6 ± 0.1	25.45 ± 0.65	4.79 ± 0.29	3.543 ± 0.059	1059 ± 18	2	
		6-	298.5 ± 0.1	24.02 ± 0.24	4.32 ± 0.08	3.713 ± 0.055	866 ± 13	3	
		5-	298.5 ± 0	27.09 ± 0.02	4.98 ± 0.10	3.307 ± 0.014	811 ± 3	3	
[OR8+Q-zH+]z-	11493.82	7-	298.7 ± 0	22.94 ± 0	4.33 ± 0	3.930 ± 0	955 ± 0	1	
		6-	298.5 ± 0.2	24.65 ± 0.04	4.55 ± 0.02	3.637 ± 0.003	884 ± 1	2	
		5-	298.8 ± 0.2	28.06 ± 0.04	5.08 ± 0.01	3.180 ± 0.004	843 ± 1	2	
[OR8-A-zH+]z-	11160.4	8-	298.6 ± 0.1	28.63 ± 0.36	5.53 ± 0.05	3.164 ± 0.044	1356 ± 18	2	
		<i>main</i>	7-	298.4 ± 0.2	23.46 ± 0.22	4.24 ± 0.07	3.801 ± 0.042	987 ± 11	4
		<i>shoulder</i>		298.4 ± 0.2	24.99 ± 0.40	4.71 ± 0.15	3.594 ± 0.069	1045 ± 20	4
		6-	298.5 ± 0.2	26.53 ± 0.07	4.56 ± 0.08	3.330 ± 0.003	966 ± 1	3	
		5-	298.6 ± 0.1	31.19 ± 0.51	5.44 ± 0.19	2.839 ± 0.038	944 ± 12	3	
[OR8-A+Q-zH+]z-	11484.82	7-	298.3 ± 0.2	24.73 ± 0.01	4.45 ± 0.24	3.601 ± 0.028	1042 ± 8	2	
			298.3 ± 0.2	26.94 ± 0.38	5.29 ± 0.06	3.389 ± 0.051	1108 ± 16	2	
[OR7-A-zH+]z-	11169.4	8-	298.8 ± 0.2	28.78 ± 0.12	5.76 ± 0.15	3.188 ± 0.009	1344 ± 4	2	
		7-	298.7 ± 0.2	23.37 ± 0.18	4.22 ± 0.08	3.811 ± 0.028	984 ± 7	3	
		6-	298.7 ± 0.2	26.37 ± 0.09	4.63 ± 0.09	3.358 ± 0.002	957 ± 1	3	
		5-	298.7 ± 0.2	30.73 ± 0.19	5.25 ± 0.09	2.865 ± 0.021	935 ± 7	3	
[OR7-A+Q-zH+]z-	11493.82	7-	298.7 ± 0.2	24.18 ± 0.05	4.73 ± 0.30	3.739 ± 0.058	1004 ± 16	2	
		7-	298.7 ± 0	26.58 ± 0	5.06 ± 0	3.362 ± 0	1116 ± 0	1	
		6-	298.7 ± 0	27.12 ± 0	4.68 ± 0	3.256 ± 0	988 ± 0	1	

600V

Table 60 - t_a , t_0 , K_0 and CCS values, with their standard deviation, for CBAs at fragmentor 600 V. Only t_a at $\Delta V=390.5$ V is shown; † obtained from CIU; # number of replica

ion	MW (Da)	z	T (K)	t_a (ms)	t_0 (ms)	K_0 (cm ² /Vs)	$^{DT}CCS_{He}$ (Å ²)	#
[MN4-zH+]z-	11128.3	8-	298.5 ± 0.1				1351 ± 13 [†]	3
		7-	298.5 ± 0.1	28.81 ± 0.17	5.34 ± 0.05	3.115 ± 0.012	1205 ± 5	3
		7-	298.6 ± 0.3				1204 ± 8 [†]	5
		6-	298.6 ± 0.1	26.74 ± 0.10	4.78 ± 0.04	3.328 ± 0.011	966 ± 4	3
		6-	298.5 ± 0.1				963 ± 2 [†]	3
		5-	298.6 ± 0.1	31.03 ± 0.17	5.58 ± 0.03	2.869 ± 0.017	934 ± 6	3
[MN4+Q-zH+]z-	11452.72	6-	298.5 ± 0.1	27.40 ± 0.13	4.92 ± 0	3.252 ± 0.014	989 ± 5	2
		5-	298.5 ± 0.1	31.77 ± 0.13	5.73 ± 0.12	2.805 ± 0.001	956 ± 1	2
[MN19-zH+]z-	11169.4	8-	298.5 ± 0	28.25 ± 0	5.45 ± 0	3.252 ± 0.014	1337 ± 0	1
		7-	298.6 ± 0.1	27.86 ± 0.08	5.29 ± 0.02	2.805 ± 0.001	1159 ± 3	2
		6-	298.6 ± 0.1	24.73 ± 0.01	4.55 ± 0.03	3.252 ± 0.014	891 ± 3	3
				26.63 ± 0.01	4.97 ± 0.30	3.352 ± 0.055	960 ± 16	3
				28.72 ± 0.11	5.46 ± 0.23	3.136 ± 0.047	1026 ± 15	3
		5-	298.5 ± 0.1	28.11 ± 0.10	5.05 ± 0.05	2.805 ± 0.001	846 ± 5	2
[MN19+Q-zH+]z-	11493.82	6-	298.5 ± 0	25.33 ± 0	4.76 ± 0	3.552 ± 0	905 ± 0	1
		5-	298.5 ± 0	28.93 ± 0	5.28 ± 0	3.091 ± 0	867 ± 0	1
[MN19-A-zH+]z-	11152.3	8-	298.5 ± 0.1				1388 ± 22 [†]	3
		7-	298.5 ± 0.1				1244 ± 3 [†]	5
		6-	298.5 ± 0.1				971 ± 5 [†]	3
		5-	298.8 ± 0.2				938 ± 3 [†]	2

600V

(Table 60 continued)

ion	MW (Da)	z	T (K)	t _a (ms)	t ₀ (ms)	K ₀ (cm ² /Vs)	^{DT} CCS _{He} (Å ²)	#
[OR8-zH+]z-	11169.4	7-	298.5 ± 0.1	27.35 ± 0.27	5.36 ± 0.05	3.322 ± 0.048	1130 ± 16	2
	Main	6-	298.6 ± 0.1	27.90 ± 0.07	5.15 ± 0.12	3.207 ± 0.017	1003 ± 5	3
	second		298.6 ± 0.1	25.29 ± 0.04	4.45 ± 0.42	3.486 ± 0.070	923 ± 19	3
		5-	298.5 ± 0.1	27.17 ± 0.05	4.90 ± 0.02	3.280 ± 0.013	817 ± 3	3
[OR8+Q-zH+]z-	11493.82	5-	298.7 ± 0.3	28.14 ± 0.09	5.14 ± 0.02	3.183 ± 0.009	843 ± 2	2
[OR8-A-zH+]z-	11160.4	8-	298.5 ± 0.1				1374 ± 1 [†]	2
		7-	298.5 ± 0.1				1231 ± 3 [†]	4
		6-	298.5 ± 0.1				973 ± 3 [†]	3
		5-	298.5 ± 0.1				951 ± 8 [†]	3
[OR7-A-zH+]z-	11169.4	8-	298.5 ± 0.1				1364 ± 10 [†]	3
		7-	298.5 ± 0.1				1230 ± 3 [†]	3
		6-	298.5 ± 0.1				967 ± 2 [†]	3
		5-	298.5 ± 0.1				938 ± 2 [†]	3

7.3.7 Comparison pre-IMS tuning Agilent 6560

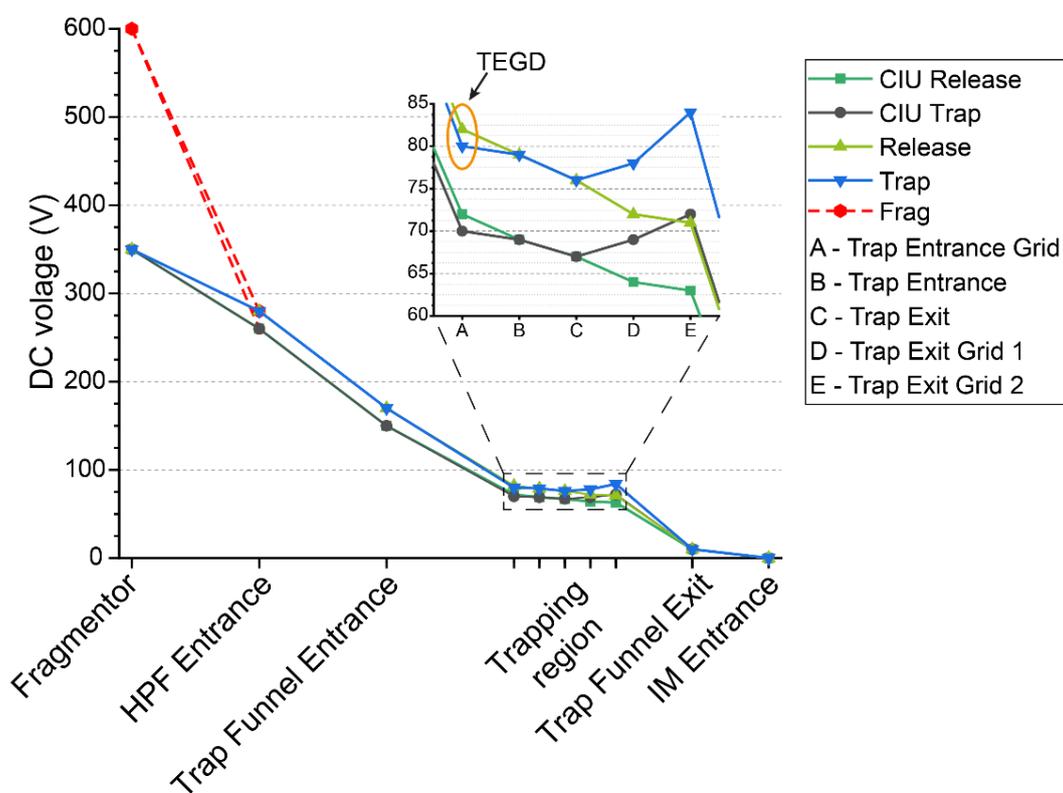


Figure 129 - Comparison between pre-IMS DC voltages of regular tuning (Table 3) and CIU tuning (Table 4). Trap Entrance Grid Delta (TEGD) is marked with an orange circle. Voltages are in absolute value.

7.3.8 CIU/CID Breakdown diagram

Table 61 - MS and t_a ranges for CIU/D diagram. α . CCS scale calibrated at fragmentor 350V.

Ion	MS ranges (m/z)				Conf A		ConfB	
	0 add.	0-5 add.	1-5 add.	tot range	t_a (ms)	CCS (\AA^2) $^\alpha$	t_a (ms)	CCS (\AA^2) $^\alpha$
[MN4-7H ⁺] ⁷⁻	1587-1590	1587-1602	1950-1602	1586-1610	22-26.4	935-1120	26.4-31	1120-1308
[MN4+Q-7H ⁺] ⁷⁻	1634-1636.2	1634-1648.5	1636.2-1648.5	1586-1658	22.3-26.7	935-1120	26.7-31.2	1120-1308
[MN19-A-7H ⁺] ⁷⁻	1591.2-1593.2	1591.2-1605.7	1593.2-1605.7	1589-1611	21.5-26.7	915-1135	26.7-32.2	1135-1370
[MN19-A+Q-7H ⁺] ⁷⁻	1637.5-1639.5	1637.5-1652	1639.5-1652	1589-1661	21.5-26.75	915-1135	26.75-32.2	1135-1370
[OR8-A-7H ⁺] ⁷⁻	1592.2-1594.2	1592.2-1607	1594.2-1607	1591-1615	21.6-26.1	911-1097	26.1-32	1907-1344
[OR8-A+Q-7H ⁺] ⁷⁻	1638.5-1641.5	1639.5-1655	1641.5-1655	1591-1670	21.62-26.14	911-1103	26.14-32	1103-1351
[OR7-A-7H ⁺] ⁷⁻	1593.6-1595.6	1593.6-1607.1	1595.6-1607.1	1591-1615	21-26	910-1100	26-32	1100-1335
[OR7-A+Q-7H ⁺] ⁷⁻	1640.5-1642.5	1640.5-1651	1642.5-1651	1591-1670	21-26	884-1090	26-32	1090-1344

8. References

- [1] J. E. Sponer, J. Leszczynski, V. Sychrovsky and J. Sponer, *J Phys Chem B* **2005**, *109*, 18680-18689.
- [2] M. Chastain and I. Tinoco in *Structural Elements in RNA, Vol. 41* Eds.: W. E. Cohn and K. Moldave), Academic Press, **1991**, pp. 131-177.
- [3] T. Hermann and D. J. Patel, *J Mol Biol* **1999**, *294*, 829-849.
- [4] K. Hoogsteen, *Acta Crystallographica* **1963**, *16*, 907-916.
- [5] a) M. A. Keniry, *Biopolymers* **2000**, *56*, 123-146; b) S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res* **2006**, *34*, 5402-5415; c) S. Neidle, *Current Opinion in Structural Biology* **2009**, *19*, 239-250; d) K. D. Warner, M. C. Chen, W. Song, R. L. Strack, A. Thorn, S. R. Jaffrey and A. R. Ferre-D'Amare, *Nat Struct Mol Biol* **2014**, *21*, 658-663.
- [6] a) P. Beal and P. Dervan, *Science* **1991**, *251*, 1360-1363; b) J. L. Mergny, J. S. Sun, M. Rougee, T. Montenay-Garestier, F. Barcelo, J. Chomilier and C. Helene, *Biochemistry* **1991**, *30*, 9791-9798.
- [7] N. B. Leontis, J. Stombaugh and E. Westhof, *Nucleic Acids Research* **2002**, *30*, 3497-3531.
- [8] A. Lescaute and E. Westhof, *Nucleic Acids Res* **2006**, *34*, 6587-6604.
- [9] N. B. Leontis and E. Westhof, *RNA* **2001**, *7*, 499-512.
- [10] A. E. Walter, D. H. Turner, J. Kim, M. H. Lyttle, P. Muller, D. H. Mathews and M. Zuker, *Proc Natl Acad Sci U S A* **1994**, *91*, 9218-9222.
- [11] R. E. Dickerson in *[5] DNA structure from A to Z, Vol. 211* Academic Press, **1992**, pp. 67-111.
- [12] W. Saenger, *Principles of Nucleic Acid Structure*, **1984**, p.
- [13] H. L. Ng, M. L. Kopka and R. E. Dickerson, *Proc Natl Acad Sci U S A* **2000**, *97*, 2035-2039.
- [14] M. D. Frank-Kamenetskii and S. M. Mirkin, *Annu Rev Biochem* **1995**, *64*, 65-95.
- [15] a) G. W. Collie, S. Sparapani, G. N. Parkinson and S. Neidle, *Journal of the American Chemical Society* **2011**, *133*, 2721-2728; b) G. W. Collie, S. M. Haider, S. Neidle and G. N. Parkinson, *Nucleic Acids Res* **2010**, *38*, 5569-5580.
- [16] P. J. Bates, J. B. Kahlon, S. D. Thomas, J. O. Trent and D. M. Miller, *J Biol Chem* **1999**, *274*, 26369-26377.
- [17] E. V. Dolgosheina, S. C. Jeng, S. S. Panchapakesan, R. Cojocar, P. S. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins and P. J. Unrau, *ACS Chem Biol* **2014**, *9*, 2412-2420.
- [18] N. B. Leontis and E. Westhof, *Current Opinion in Structural Biology* **2003**, *13*, 300-308.
- [19] U. Nagaswamy and G. E. Fox, *RNA* **2002**, *8*, 1112-1119.
- [20] G. J. Quigley and A. Rich, *Science* **1976**, *194*, 796-806.
- [21] T. Hermann and D. J. Patel, *Structure* **2000**, *8*, R47-R54.
- [22] W. Stephenson, P. N. Asare-Okai, A. A. Chen, S. Keller, R. Santiago, S. A. Tenenbaum, A. E. Garcia, D. Fabris and P. T. Li, *J Am Chem Soc* **2013**, *135*, 5602-5611.

- [23] E. Ennifar, P. Walter, B. Ehresmann, C. Ehresmann and P. Dumas, *Nat Struct Biol* **2001**, *8*, 1064-1068.
- [24] R. C. O. Alexander P Gultyaev, Cornelis WA Pleij, Eric Westhof in *RNA Structure: Pseudoknots*, **2012**.
- [25] F. H. van Batenburg, A. P. Gultyaev and C. W. Pleij, *Nucleic Acids Res* **2001**, *29*, 194-195.
- [26] D. M. J. Lilley, *Quarterly Reviews of Biophysics* **2001**, *33*, 109-159.
- [27] D. M. Lilley, *Biopolymers* **1998**, *48*, 101-112.
- [28] H. Xiao, T. E. Edwards and A. R. Ferre-D'Amare, *Chem Biol* **2008**, *15*, 1125-1137.
- [29] T. Yamauchi, D. Miyoshi, T. Kubodera, A. Nishimura, S. Nakai and N. Sugimoto, *FEBS Lett* **2005**, *579*, 2583-2588.
- [30] K. Phillips, Z. Dauter, A. I. Murchie, D. M. Lilley and B. Luisi, *J Mol Biol* **1997**, *273*, 171-182.
- [31] D. Laage, T. Elsaesser and J. T. Hynes, *Chem Rev* **2017**, *117*, 10694-10725.
- [32] C. Baugh, D. Grate and C. Wilson, *J Mol Biol* **2000**, *301*, 117-128.
- [33] a) F. Tolle, G. M. Brandle, D. Matzner and G. Mayer, *Angew Chem Int Ed Engl* **2015**, *54*, 10971-10974; b) M. Famulok, G. Mayer and M. Blind, *Accounts of Chemical Research* **2000**, *33*, 591-599.
- [34] H. P. Hofmann, S. Limmer, V. Hornung and M. Sprinzl, *RNA* **1997**, *3*, 1289-1300.
- [35] a) P. Fan, A. K. Suri, R. Fiala, D. Live and D. J. Patel, *J Mol Biol* **1996**, *258*, 480-500; b) M. N. Stojanovic, P. de Prada and D. W. Landry, *J Am Chem Soc* **2000**, *122*, 11547-11548; c) D. Grate and C. Wilson, *Proc Natl Acad Sci U S A* **1999**, *96*, 6131-6136.
- [36] L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas and J. J. Toole, *Nature* **1992**, *355*, 564-566.
- [37] D. Shangguan, Y. Li, Z. Tang, Z. C. Cao, H. W. Chen, P. Mallikaratchy, K. Sefah, C. J. Yang and W. Tan, *Proc Natl Acad Sci U S A* **2006**, *103*, 11838-11843.
- [38] J. E. Barrick and R. R. Breaker, *Genome Biol* **2007**, *8*, R239.
- [39] C. Berens, F. Groher and B. Suess, *Biotechnol J* **2015**, *10*, 246-257.
- [40] a) C. Tuerk and L. Gold, *Science* **1990**, *249*, 505-510; b) A. D. Ellington and J. W. Szostak, *Nature* **1990**, *346*, 818-822.
- [41] S. K. Silverman in *Artificial Functional Nucleic Acids: Aptamers, Ribozymes, and Deoxyribozymes Identified by In Vitro Selection*, Eds.: L. Yingfu and L. Yi), Springer New York, New York, NY, **2009**, pp. 47-108.
- [42] D. L. Robertson and G. F. Joyce, *Nature* **1990**, *344*, 467-468.
- [43] E. J. Cho, J. W. Lee and A. D. Ellington, *Annu Rev Anal Chem (Palo Alto Calif)* **2009**, *2*, 241-264.
- [44] J. M. Carothers, J. A. Goler, Y. Kapoor, L. Lara and J. D. Keasling, *Nucleic Acids Res* **2010**, *38*, 2736-2747.
- [45] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T. R. Keeney, N. Kim, T. H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S. K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R. M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D. I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J. D. Vaught, M.

- Vrkljan, J. J. Walker, M. Watrobka, S. Waugh, A. Weiss, S. K. Wilcox, A. Wolfson, S. K. Wolk, C. Zhang and D. Zichi, *PLoS One* **2010**, *5*, e15004.
- [46] P. D. P. Atkins, J. *Atkins' Physical Chemistry*, Oxford University Press, **2006**, p.
- [47] N. N. Salim and A. L. Feig, *Methods* **2009**, *47*, 198-205.
- [48] L. Indyk and H. F. Fisher in [17] *Theoretical aspects of isothermal titration calorimetry, Vol. 295* Academic Press, **1998**, pp. 350-364.
- [49] T. Wiseman, S. Williston, J. F. Brandts and L.-N. Lin, *Analytical Biochemistry* **1989**, *179*, 131-137.
- [50] a) J. Tellinghuisen, *Anal Biochem* **2003**, *321*, 79-88; b) J. B. Chaires, *Annu Rev Biophys* **2008**, *37*, 135-151; c) W. B. Turnbull and A. H. Daranas, *J Am Chem Soc* **2003**, *125*, 14859-14866.
- [51] P. Hensley, *Structure* **1996**, *4*, 367-373.
- [52] a) J. Fenn, M. Mann, C. Meng, S. Wong and C. Whitehouse, *Science* **1989**, *246*, 64-71; b) M. Yamashita and J. B. Fenn, *The Journal of Physical Chemistry* **1984**, *88*, 4451-4459.
- [53] M. Karas, D. Bachmann, U. Bahr and F. Hillenkamp, *International Journal of Mass Spectrometry and Ion Processes* **1987**, *78*, 53-68.
- [54] A. C. Susa, Z. Xia and E. R. Williams, *Angew Chem Int Ed Engl* **2017**, *56*, 7912-7915.
- [55] a) M. S. Wilm and M. Mann, *International Journal of Mass Spectrometry and Ion Processes* **1994**, *136*, 167-180; b) M. Karas, U. Bahr and T. Dülcks, *Fresenius' Journal of Analytical Chemistry* **2000**, *366*, 669-676.
- [56] J. V. Iribarne, *The Journal of Chemical Physics* **1976**, *64*, 2287.
- [57] a) L. L. Mack, P. Kralik, A. Rheude and M. Dole, *The Journal of Chemical Physics* **1970**, *52*, 4977-4986; b) M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson and M. B. Alice, *The Journal of Chemical Physics* **1968**, *49*, 2240-2249.
- [58] a) E. Ahadi and L. Konermann, *J Phys Chem B* **2012**, *116*, 104-112; b) L. Konermann, A. D. Rodriguez and J. Liu, *Anal Chem* **2012**, *84*, 6798-6804; c) H. Metwally, Q. Duez and L. Konermann, *Analytical Chemistry* **2018**, *90*, 10069-10077.
- [59] N. Khristenko, J. Amato, S. Livet, B. Pagano, A. Randazzo and V. Gabelica, *J Am Soc Mass Spectrom* **2019**.
- [60] B. A. Lodish, H. Zipursky, S. L., et al. in *Intracellular Ion Environment and Membrane Electric Potential*, (Ed. W. H. Freeman), **2000**.
- [61] L. Konermann, *J Am Soc Mass Spectrom* **2017**.
- [62] S. A. McLuckey, *Journal of the American Society for Mass Spectrometry* **1992**, *3*, 599-614.
- [63] a) J. Ujma, K. Giles, M. Morris and P. E. Barran, *Anal Chem* **2016**, *88*, 9469-9478; b) D. E. Clemmer and M. F. Jarrold, *Journal of Mass Spectrometry* **1997**, *32*, 577-592; c) M. Kwasnik, K. Fuhrer, M. Gonin, K. Barbeau and F. M. Fernández, *Analytical Chemistry* **2007**, *79*, 7782-7791.
- [64] a) S. M. Swasey, F. Rosu, S. M. Copp, V. Gabelica and E. G. Gwinn, *J Phys Chem Lett* **2018**, *9*, 6605-6610; b) M. Porrini, F. Rosu, C. Rabin, L. Darre, H. Gomez, M. Orozco and V. Gabelica, *ACS Cent Sci* **2017**, *3*, 454-461.
- [65] B. T. Ruotolo, K. Giles, I. Campuzano, A. M. Sandercock, R. H. Bateman and C. V. Robinson, *Science* **2005**, *310*, 1658-1661.

- [66] a) V. Gabelica, S. Livet and F. Rosu, *J Am Soc Mass Spectrom* **2018**; b) T. M. Allison, E. Reading, I. Liko, A. J. Baldwin, A. Laganowsky and C. V. Robinson, *Nat Commun* **2015**, *6*, 8551.
- [67] a) S. Zhou and K. D. Cook, *Journal of the American Society for Mass Spectrometry* **2001**, *12*, 206-214; b) N. B. Cech and C. G. Enke, *Mass Spectrom Rev* **2001**, *20*, 362-387.
- [68] T. Hermann and D. J. Patel, *Science* **2000**, *287*, 820-825.
- [69] K. Chingin and K. Barylyuk, *Anal Chem* **2018**, *90*, 5521-5528.
- [70] a) M. Bakhtiari and L. Konermann, *J Phys Chem B* **2019**, *123*, 1784-1796; b) F. Lanucara, S. W. Holman, C. J. Gray and C. E. Eyers, *Nat Chem* **2014**, *6*, 281-294.
- [71] E. G. Marklund, M. T. Degiacomi, C. V. Robinson, A. J. Baldwin and J. L. Benesch, *Structure* **2015**, *23*, 791-799.
- [72] a) T. Wyttenbach, G. von Helden and M. T. Bowers, *Journal of the American Chemical Society* **1996**, *118*, 8355-8364; b) V. D'Atri, M. Porrini, F. Rosu and V. Gabelica, *J Mass Spectrom* **2015**, *50*, 711-726.
- [73] a) S. J. Valentine, M. D. Plasencia, X. Liu, M. Krishnan, S. Naylor, H. R. Udseth, R. D. Smith and D. E. Clemmer, *Journal of Proteome Research* **2006**, *5*, 2977-2984; b) M. D. Plasencia, D. Isailovic, S. I. Merenbloom, Y. Mechref and D. E. Clemmer, *Journal of the American Society for Mass Spectrometry* **2008**, *19*, 1706-1715; c) K. Jeanne Dit Fouque, C. Afonso, S. Zirah, J. D. Hegemann, M. Zimmermann, M. A. Marahiel, S. Rebuffat and H. Lavanant, *Anal Chem* **2015**, *87*, 1166-1172; d) C. Manz and K. Pagel, *Curr Opin Chem Biol* **2018**, *42*, 16-24; e) J. Hofmann, H. S. Hahm, P. H. Seeberger and K. Pagel, *Nature* **2015**, *526*, 241-244.
- [74] W. F. Siems, L. A. Viehland and H. H. Hill, Jr., *Anal Chem* **2012**, *84*, 9782-9791.
- [75] J. E. Weigand and B. Suess, *Appl Microbiol Biotechnol* **2009**, *85*, 229-236.
- [76] E. Goux, S. Lisi, C. Ravelet, G. Durand, E. Fiore, E. Dausse, J. J. Toulme and E. Peyrin, *Anal Bioanal Chem* **2015**, *407*, 6515-6524.
- [77] G. Durand, S. Lisi, C. Ravelet, E. Dausse, E. Peyrin and J. J. Toulme, *Angew Chem Int Ed Engl* **2014**, *53*, 6942-6945.
- [78] M. Vogel and B. Suess in *Label-Free Determination of the Dissociation Constant of Small Molecule-Aptamer Interaction by Isothermal Titration Calorimetry*, (Ed. G. Mayer), Springer New York, New York, NY, **2016**, pp. 113-125.
- [79] a) D. Wunnicke, D. Strohbach, J. E. Weigand, B. Appel, E. Feresin, B. Suess, S. Muller and H. J. Steinhoff, *RNA* **2011**, *17*, 182-188; b) E. Duchardt-Ferner, J. E. Weigand, O. Ohlenschlager, S. R. Schmidtke, B. Suess and J. Wohnert, *Angew Chem Int Ed Engl* **2010**, *49*, 6216-6219; c) P. Cekan, E. O. Jonsson and S. T. Sigurdsson, *Nucleic Acids Res* **2009**, *37*, 3990-3995.
- [80] a) J. B. Da Costa, A. I. Andreiev and T. Dieckmann, *Biochemistry* **2013**, *52*, 6575-6583; b) J. Bernard Da Costa and T. Dieckmann, *Mol Biosyst* **2011**, *7*, 2156-2163.
- [81] P. Burgstaller and M. Famulok, *Angewandte Chemie International Edition in English* **1994**, *33*, 1084-1087.
- [82] L. Jiang, A. Majumdar, W. Hu, T. J. Jaishree, W. Xu and D. J. Patel, *Structure* **1999**, *7*, 817-817.
- [83] a) L. Jiang, A. K. Suri, R. Fiala and D. J. Patel, *Chemistry & Biology* **1997**, *4*, 35-50; b) Y. Wang, J. Killian, K. Hamasaki and R. R. Rando, *Biochemistry* **1996**, *35*, 12338-12346.
- [84] M. G. Wallis, U. von Ahsen, R. Schroeder and M. Famulok, *Chemistry & Biology* **1995**, *2*, 543-552.

- [85] J. Flinders, S. C. DeFina, D. M. Brackett, C. Baugh, C. Wilson and T. Dieckmann, *ChemBiochem* **2004**, *5*, 62-72.
- [86] M. Müller, J. E. Weigand, O. Weichenrieder and B. Suess, *Nucleic Acids Research* **2006**, *34*, 2607-2617.
- [87] O. Reinstein, M. Yoo, C. Han, T. Palmo, S. A. Beckham, M. C. Wilce and P. E. Johnson, *Biochemistry* **2013**, *52*, 8652-8662.
- [88] M. A. D. Neves, A. A. Shoara, O. Reinstein, O. Abbasi Borhani, T. R. Martin and P. E. Johnson, *ACS Sens* **2017**, *2*, 1539-1545.
- [89] I. Tinoco, *Journal of the American Chemical Society* **1960**, *82*, 4785-4790.
- [90] J. L. Mergny and L. Lacroix, *Oligonucleotides* **2003**, *13*, 515-537.
- [91] a) G. R. Bishop and J. B. Chaires, *Current Protocols in Nucleic Acid Chemistry* **2002**, *11*, 7.11.11-17.11.18; b) N. C. Garbett, P. A. Ragazzon and J. B. Chaires, *Nature Protocols* **2007**, *2*, 3166.
- [92] R. W. Woody in [4] *Circular dichroism, Vol. 246* Academic Press, **1995**, pp. 34-71.
- [93] K. I. Ruwan Kurulugama, Lester Taylo in *The Agilent Ion Mobility Q-TOF MassSpectrometer System, Vol. Agilent Technologies, Inc., 2013*.
- [94] I. Agilent Technologies in *Agilent 6560 Ion Mobility Q-TOF Specifications, Vol. 2016*.
- [95] D. Paul, A. Marchand, D. Verga, S. Bombard, M.-P. Teulade-fichou, F. Rosu and V. Gabelica, *The Analyst* **2019**.
- [96] K. Tang, A. A. Shvartsburg, H. N. Lee, D. C. Prior, M. A. Buschbach, F. Li, A. V. Tolmachev, G. A. Anderson and R. D. Smith, *Anal Chem* **2005**, *77*, 3330-3339.
- [97] R. T. Kelly, A. V. Tolmachev, J. S. Page, K. Tang and R. D. Smith, *Mass Spectrom Rev* **2010**, *29*, 294-312.
- [98] C. Rabin in *Investigation of RNA kissing complexes by native electrospray mass spectrometry : magnesium binding and ion mobility, Vol. 2017*.
- [99] A. Marchand, S. Livet, F. Rosu and V. Gabelica, *Anal Chem* **2017**, *89*, 12674-12681.
- [100] E. Denisov, E. Damoc, O. Lange and A. Makarov, *International Journal of Mass Spectrometry* **2012**, *325-327*, 80-85.
- [101] a) J. Gault, J. A. Donlan, I. Liko, J. T. Hopper, K. Gupta, N. G. Housden, W. B. Struwe, M. T. Marty, T. Mize, C. Bechara, Y. Zhu, B. Wu, C. Kleanthous, M. Belov, E. Damoc, A. Makarov and C. V. Robinson, *Nat Methods* **2016**, *13*, 333-336; b) Q. Hu, R. J. Noll, H. Li, A. Makarov, M. Hardman and R. Graham Cooks, *J Mass Spectrom* **2005**, *40*, 430-443.
- [102] R. A. Zubarev and A. Makarov, *Analytical Chemistry* **2013**, *85*, 5288-5296.
- [103] V. Gabelica, F. Rosu and E. De Pauw, *Anal Chem* **2009**, *81*, 6708-6715.
- [104] M. Peschke, U. H. Verkerk and P. Kebarle, *J Am Soc Mass Spectrom* **2004**, *15*, 1424-1434.
- [105] P. Kebarle and U. H. Verkerk, *Mass Spectrom Rev* **2009**, *28*, 898-917.
- [106] P. Kuzmic, *Anal Biochem* **1996**, *237*, 260-273.

- [107] P. Kuzmič in *Chapter 10 - DynaFit—A Software Package for Enzymology*, Vol. 467 Academic Press, **2009**, pp. 247-280.
- [108] M. Straume and M. L. Johnson in *[7] Monte Carlo Method for determining complete confidence probability distributions of estimated model parameters*, Vol. 210 Academic Press, **1992**, pp. 117-129.
- [109] J. R. Taylor, *An Introduction to Error Analysis*, University Science Books, **1982**, p.
- [110] a) D. A. Polasky, S. M. Dixit, S. M. Fantin and B. T. Ruotolo, *Anal Chem* **2019**, *91*, 3147-3155; b) S. Niu and B. T. Ruotolo, *Protein Sci* **2015**, *24*, 1272-1281; c) B. T. Ruotolo, J. L. Benesch, A. M. Sandercock, S. J. Hyung and C. V. Robinson, *Nat Protoc* **2008**, *3*, 1139-1152.
- [111] a) M. J. Cavaluzzi and P. N. Borer, *Nucleic Acids Res* **2004**, *32*, e13; b) C. R. Cantor, M. M. Warshaw and H. Shapiro, *Biopolymers* **1970**, *9*, 1059-1077.
- [112] F. J. Green, *The Sigma-Aldrich Handbook of Stains, Dyes and Indicators*, Aldrich Chemical Co., **1990**, p.
- [113] S. K. Bharti and R. Roy, *TrAC Trends in Analytical Chemistry* **2012**, *35*, 5-26.
- [114] S. A. Catalog in *F1392*, Vol.
- [115] J. L. Irvin and E. M. Irvin, *J Biol Chem* **1948**, *174*, 589-596.
- [116] M. Takahashi, L. Altschmied and W. Hillen, *J Mol Biol* **1986**, *187*, 341-348.
- [117] E. J. Cho, J. R. Collett, A. E. Szafranska and A. D. Ellington, *Anal Chim Acta* **2006**, *564*, 82-90.
- [118] a) K. M. Keller, M. M. Breeden, J. Zhang, A. D. Ellington and J. S. Brodbelt, *J Mass Spectrom* **2005**, *40*, 1327-1337; b) K. A. Sannes-Lowery, R. H. Griffey and S. A. Hofstadler, *Anal Biochem* **2000**, *280*, 264-271; c) L. A. Cassiday, L. L. Lebruska, L. M. Benson, S. Naylor, W. G. Owen and L. J. Maher, 3rd, *Anal Biochem* **2002**, *306*, 290-297.
- [119] B. Gulbakan, K. Barylyuk, P. Schneider, M. Pillong, G. Schneider and R. Zenobi, *J Am Chem Soc* **2018**.
- [120] Y. Marcus, *Journal of the Chemical Society, Faraday Transactions* **1993**, *89*, 713-718.
- [121] a) F. Leonarski, L. D'Ascenzo and P. Auffinger, *RNA* **2018**; b) F. Leonarski, L. D'Ascenzo and P. Auffinger, *Nucleic Acids Res* **2017**, *45*, 987-1004.
- [122] A. Marchand and V. Gabelica, *J Am Soc Mass Spectrom* **2014**, *25*, 1146-1154.
- [123] D. H. Aue, H. M. Webb and M. T. Bowers, *Journal of the American Chemical Society* **1976**, *98*, 318-329.
- [124] a) C. S. Wilcox in *Frontiers in Supramolecular Organic Chemistry and Photochemistry*, (Ed. Frontie), ed. H.-J. Schneider and H. Dürr, VCH, Weinheim, **1991**, pp. 123-144; b) P. Thordarson, *Chem Soc Rev* **2011**, *40*, 1305-1323.
- [125] B. Ganisl, M. Taucher, C. Rimpl and K. Breuker, *Eur J Mass Spectrom (Chichester)* **2011**, *17*, 333-343.
- [126] V. Gabelica, E. De Pauw and F. Rosu, *Journal of Mass Spectrometry* **1999**, *34*, 1328-1337.
- [127] S. J. Klug and M. Famulok, *Molecular Biology Reports* **1994**, *20*, 97-107.
- [128] T. K. Misra, *Plasmid* **1992**, *27*, 4-16.
- [129] P. Macheroux, B. Kappes and S. E. Ealick, *FEBS J* **2011**, *278*, 2625-2634.

- [130] R. D. Grubbs, *Biometals* **2002**, *15*, 251-259.
- [131] M. G. Wallis and R. Schroeder, *Progress in Biophysics and Molecular Biology* **1997**, *67*, 141-154.
- [132] a) D. Fourmy, M. I. Recht, S. C. Blanchard and J. D. Puglisi, *Science* **1996**, *274*, 1367-1371; b) J. Haddad, L. P. Kotra, B. Llano-Sotelo, C. Kim, E. F. Azucena, M. Liu, S. B. Vakulenko, C. S. Chow and S. Mobashery, *Journal of the American Chemical Society* **2002**, *124*, 3229-3237.
- [133] J. Cho and R. R. Rando, *Biochemistry* **1999**, *38*, 8548-8554.
- [134] T. K. Stage, K. J. Hertel and O. C. Uhlenbeck, *RNA* **1995**, *1*, 95-101.
- [135] Y. Marcus, *Journal of Molecular Liquids* **2005**, *118*, 3-8.
- [136] T. Kenderdine, Z. Xia, E. R. Williams and D. Fabris, *Anal Chem* **2018**, *90*, 13541-13548.
- [137] C. Berens, A. Thain and R. Schroeder, *Bioorg Med Chem* **2001**, *9*, 2549-2556.
- [138] P. Kotter, J. E. Weigand, B. Meyer, K. D. Entian and B. Suess, *Nucleic Acids Res* **2009**, *37*, e120.
- [139] M. Muller, J. E. Weigand, O. Weichenrieder and B. Suess, *Nucleic Acids Res* **2006**, *34*, 2607-2617.
- [140] A. J. Reuss, M. Vogel, J. E. Weigand, B. Suess and J. Wachtveitl, *Biophys J* **2014**, *107*, 2962-2971.
- [141] U. Forster, J. E. Weigand, P. Trojanowski, B. Suess and J. Wachtveitl, *Nucleic Acids Res* **2012**, *40*, 1807-1817.
- [142] V. Aladin, M. Vogel, R. Binder, I. Burghardt, B. Suess and B. Corzilius, *Angew Chem Int Ed Engl* **2019**, *58*, 4863-4868.
- [143] S. Slavkovic, M. Altunisik, O. Reinstein and P. E. Johnson, *Bioorganic & Medicinal Chemistry* **2015**, *23*, 2593-2597.
- [144] a) O. Reinstein, M. A. Neves, M. Saad, S. N. Boodram, S. Lombardo, S. A. Beckham, J. Brouwer, G. F. Audette, P. Groves, M. C. Wilce and P. E. Johnson, *Biochemistry* **2011**, *50*, 9368-9376; b) M. N. Stojanovic, E. G. Green, S. Semova, D. B. Nikic and D. W. Landry, *J Am Chem Soc* **2003**, *125*, 6085-6089.
- [145] M. A. Neves, O. Reinstein and P. E. Johnson, *Biochemistry* **2010**, *49*, 8478-8487.
- [146] J. Canoura, Z. Wang, H. Yu, O. Alkhamis, F. Fu and Y. Xiao, *J Am Chem Soc* **2018**, *140*, 9961-9971.
- [147] M. J. Greig, H.-J. Gaus and R. H. Griffey, *Rapid Communications in Mass Spectrometry* **1996**, *10*, 47-50.
- [148] X. Guo, M. F. Bruist, D. L. Davis and C. M. Bentzley, *Nucleic Acids Res* **2005**, *33*, 3659-3666.
- [149] L. Konermann and D. J. Douglas, *J Am Soc Mass Spectrom* **1998**, *9*, 1248-1254.
- [150] a) A. Tikhomirova, I. V. Beletskaya and T. V. Chalikian, *Biochemistry* **2006**, *45*, 10563-10571; b) N. Peyret, P. A. Seneviratne, H. T. Allawi and J. SantaLucia, Jr., *Biochemistry* **1999**, *38*, 3468-3477.
- [151] F. Balthasart, J. Plavec and V. Gabelica, *J Am Soc Mass Spectrom* **2013**, *24*, 1-8.
- [152] F. Rosu, V. Gabelica, L. Joly, G. Gregoire and E. De Pauw, *Phys Chem Chem Phys* **2010**, *12*, 13448-13454.
- [153] A. Granzhan, N. Kotera and M. P. Teulade-Fichou, *Chem Soc Rev* **2014**, *43*, 3630-3665.
- [154] N. C. Stellwagen and E. Stellwagen, *J Phys Chem B* **2019**, *123*, 3649-3657.

- [155] Z. R. Churcher, M. A. D. Neves, H. N. Hunter and P. E. Johnson, *J Biomol NMR* **2017**, *68*, 33-39.
- [156] V. K. Misra and D. E. Draper, *Biopolymers* **1998**, *48*, 113-135.
- [157] M. N. Stojanovic, P. de Prada and D. W. Landry, *J Am Chem Soc* **2001**, *123*, 4928-4931.
- [158] J. D. Eschweiler, J. N. Rabuck-Gibbons, Y. Tian and B. T. Ruotolo, *Analytical Chemistry* **2015**, *87*, 11516-11522.
- [159] N. Sun, J. Sun, E. N. Kitova and J. S. Klassen, *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 1242-1250.
- [160] A. Haller, R. B. Altman, M. F. Souliere, S. C. Blanchard and R. Micura, *Proc Natl Acad Sci U S A* **2013**, *110*, 4188-4193.
- [161] C. D. Stoddard, R. K. Montange, S. P. Hennelly, R. P. Rambo, K. Y. Sanbonmatsu and R. T. Batey, *Structure* **2010**, *18*, 787-797.
- [162] S. D. Gilbert, C. D. Stoddard, S. J. Wise and R. T. Batey, *J Mol Biol* **2006**, *359*, 754-768.
- [163] M. Vorlíčková, I. Kejnovská, K. Bednářová, D. Renčíuk and J. Kypr, *Chirality* **2012**, *24*, 691-698.
- [164] J. Kypr, I. Kejnovska, D. Renciuk and M. Vorlickova, *Nucleic Acids Res* **2009**, *37*, 1713-1725.