bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

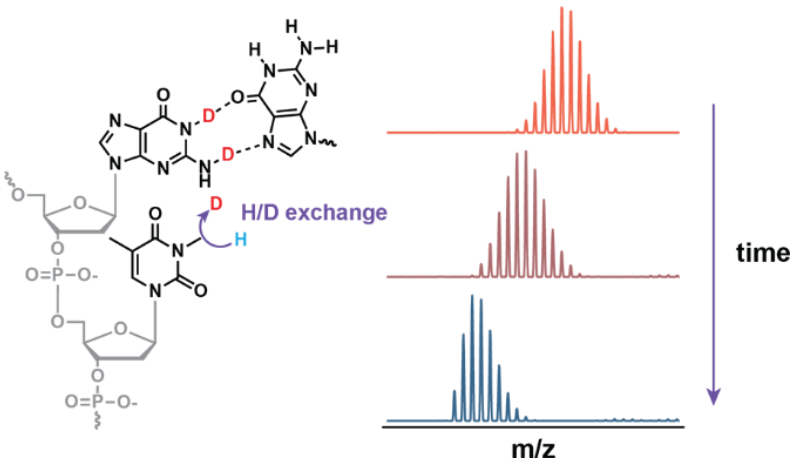
Native Hydrogen/Deuterium Exchange Mass Spectrometry of Structured DNA Oligonucleotides

Eric Largy,* and Valérie Gabelica

University of Bordeaux, INSERM and CNRS, Laboratoires Acides Nucléiques : Régulations Naturelle et Artificielle (ARNA, U1212, UMR5320), IECB, 2 rue Robert Escarpit, 33600 Pessac, France.

ABSTRACT

Although hydrogen-deuterium exchange mass spectrometry (HDX/MS) is well-established for the analysis of the structure and dynamics of proteins, it is currently not exploited for nucleic acids. Here we used DNA G-quadruplex structures as model systems to demonstrate that DNA oligonucleotides are amenable to in-solution HDX/MS in native conditions. In trimethylammonium acetate solutions and in soft source conditions, the protonated phosphate groups are fully back-exchanged in the source, while the exchanged nucleobases remain labeled without detectable gas-phase back-exchange. As a result, the exchange rates depend strongly on the secondary structure (hydrogen bonding status) of the oligonucleotides, but neither on their charge state nor on the presence of non-specific adducts. We show that native mass spectrometry methods can measure these exchange rates on the second to the day time scale with high precision. Such combination of HDX with native MS opens promising avenues for the analysis of the structural and biophysical properties of oligonucleotides and their complexes.



bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

INTRODUCTION

In-solution HDX/MS is increasingly used to study the structural dynamics and interactions of proteins, in particular since the advent of automated systems. It relies on measuring exchange rates of backbone amide hydrogens in a deuterium-rich buffer, because these rates are heavily dependent on hydrogen-bonding and solvent accessibility.[1–9] Since all amino-acid (except prolines) possess an amide hydrogen, HDX/MS can effectively cover the entire sequence. HDX/MS is particularly valuable for proteins that are not amenable to X-Ray diffraction or NMR experiments, because of *e.g.* their large size, high dynamics, or limited availability. In this manuscript, we show that HDX/MS is also particularly well suited to study structured oligonucleotides (Figure 1).

Oligonucleotides contain several exchangeable sites: the 5'- and 3'-OH termini, the 2'-OH of RNA riboses (not explored herein), the phosphates, and the amino and imino hydrogens of bases. The -OH termini obviously do not cover the whole sequence, and are not good markers of the structure. The phosphate groups do cover the full sequence, but given their very low pK_a (~ 1), all phosphate groups are fully deprotonated at physiological pH.[10] Some are neutralized during the electrospray process, but that may not provide a clear information on the oligonucleotide's structure and dynamics in the bulk solution. The nucleobases are involved in secondary structure formation via hydrogen bonds (Figure 1A), and given that each base contains at least one exchangeable site (dG: 3, dA and dC: 2, dT: 1), they theoretically suffice to cover the whole sequence.

Hydrogen exchange of nucleic acids with deuterium or tritium was in fact studied early on. Some of the earliest attempts at characterizing dsDNA breathing fluctuation on the minute time scale were made in the 1960s by H-T exchange quantified by scintillation counting.[11–13] The exchange was however too fast to resolve the surface sites accessible in the grooves. From the 1970s, NMR was used to obtain exchange rates at the nucleotide level. However, as nucleic acids contain only four distinct nucleotides, determining the exchange rate of a particular nucleotide is only possible for short analytes (e.g. 10 bp), or those selected for their particularly well-resolved NMR peaks.[14]

DNA secondary structures are not limited to the Watson-Crick double helix. Nucleic acids can form a variety of non-canonical structures such as G-quadruplexes, triplexes or i-motifs,[15,16] which either have biological roles[17,18] or can be exploited in artificial nucleic acid constructs, for example aptamers.[19–22] G-quadruplexes are formed in guanine-rich sequences by the stacking of a minimum of two G-quartets. A G-quartet (or tetrad) is a quasi-coplanar association of four guanines linked by a network of eight Hoogsteen hydrogen bonds through their imino and amino hydrogens (Figure 1B). A cation, typically potassium, is coordinated between each pair of G-quartets, and strongly contributes to the stability of the structure. In a G-quadruplex with *n* G-quartets, *n*-1 cations are generally specifically coordinated.[23] G-quadruplexes are extremely polymorphic: their topology is defined by the strand stoichiometry, the relative orientation of their strands (relative to their 5'-to 3'polarity), number of G-quartets, and the length and geometry of the loops bridging tracts of guanines.[16,24] The fact that not every guanine is necessarily part of a G-quartet makes structures difficult to predict, and is often causing polymorphism.[24,25]

HDX/^{1}H-NMR is currently used to detect protected protons in G-quadruplexes, but the technique is limited to structures that gives well-resolved NMR peaks.[26–28] More generally, both NMR and X-ray crystallography can provide high-resolution structures of oligonucleotides. Unfortunately, they are limited in their ability to deal with several conformers in equilibrium, may promote the formation of higher-order structures that might not be biologically relevant, and require expensive equipment. We and others have shown that native MS helps studying individual oligonucleotide species from mixtures of conformers.[10,29–31] It relies on the direct stoichiometry observation and quantification of each peak by m/z separation. In

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

particular, the measurement of cation stoichiometry provides information on structures and folding pathways of G-quadruplexes.[24,32,33]

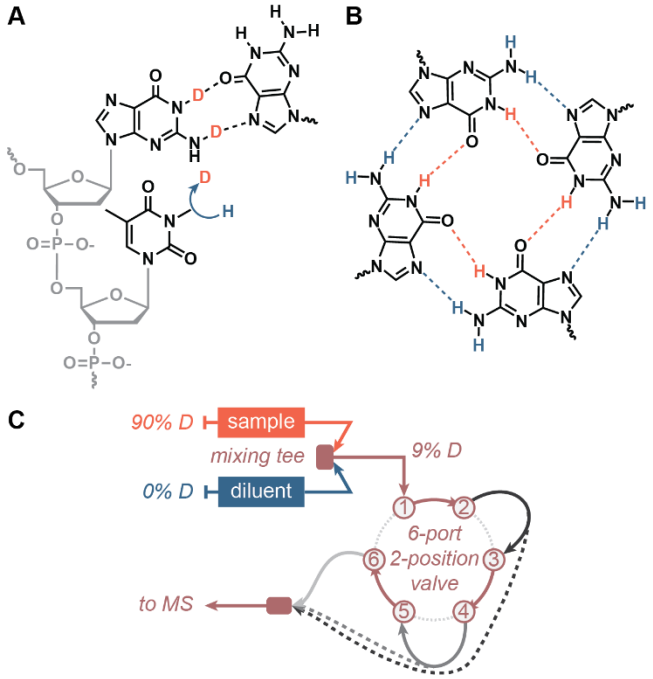


Figure 1. *HDX relies on the differential exchange of hydrogen and deuterium as a function of the hydrogen bonding status (herein a D to H exchange of nucleobases) (A). G-quadruplexes were used as model samples because the G-quartets they contain are formed by a network of eight imino (orange) and amino (blue) hydrogen bonds (B) that can be exchanged over a large time range. To monitor fast-exchanging sites (1—300 s), a continuous flow setup was used, making use of a 6-port, 2-position valve to easily cycle through capillary sizes (shades of grey) (C).*

Herein, we show that coupling HDX with native MS has the potential to combine the benefits of the two approaches. We validated that the bases, but not the phosphates, are isotopically exchanged proportionally to the solution deuterium content without in-source back-exchange. We developed analytical tools to precisely monitor the exchange over several orders of magnitude of time scales (seconds to days) (Figure 1C). Finally, we demonstrate on DNA G-quadruplexes that the measured exchange rates depend on the hydrogen bonding status of nucleobases, and thus allow to probe nucleic acid folding.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

EXPERIMENTAL SECTION

Materials

Oligonucleotides were purchased from Eurogentec (Seraing, Belgium) or IDT (Leuven, Belgium), in desalted and lyophilized form. Their sequence and melting temperatures are gathered in Table 1 and Table S1. Trimethyl ammonium acetate (TMAA) (1 M aqueous solution, HPLC grade), ammonium acetate (5 M aqueous solution, molecular biology grade), KCl (99.999% trace metal basis), and D_2O (99.9% D atom) were obtained from Sigma-Aldrich (Saint-Quentin Fallavier, France). Samples were prepared in nuclease-free water (Ambion, Life technologies SAS, Saint-Aubin, France). Data analysis was performed with OriginPro 2018 (OriginLab, Northampton, MA, USA) and RStudio 1.2.1335 (RStudio Team, Boston, MA, USA).

Table 1. Selected oligodeoxynucleotide sequences.[a]

Name	Sequence (5' to 3')	Structure[b]
T12	T_{12}	unfolded
T24	T_{24}	unfolded
23nonG4	$TG_3ATGCGACA(GA)_2G_2ACG_3$	unfolded[25]
N-myc	$TA\mathbf{G_3}C\mathbf{G_3}(A\mathbf{G_3})_2A_2$	G-quadruplex[34]
26CEB	$A_2\mathbf{G_3}T\mathbf{G_3}TGTA_2GTGT\mathbf{G_3}T\mathbf{G_3}T$	G-quadruplex[35]
T30177-TT	$T_2\mathbf{G}T\mathbf{G_2}(T\mathbf{G_3})_3T$	G-quadruplex[36]
TBA	$\mathbf{G_2}T_2\mathbf{G_2}TGT\mathbf{G_2}T_2\mathbf{G_2}$	G-quadruplex[37]
222T	$T(\mathbf{G_3}T)_4$	G-quadruplex[24]
ds26	$CA_2TCG_2ATCGA_2T_2CGATC_2GAT_2G$	dsDNA

[a] All sequences are reported in Table S1

[b] Guanines involved in G-quartet formation are bolded

Sample preparation

The deuterated (pre-exchange) samples contained 50 µM DNA (from 1 mM stock solutions in D_2O) and 90% D_2O, 1 mM KCl, 100 mM TMAA (from 10 mM KCl and 1 M TMAA stock solution in water). Melting temperatures were determined by UV-melting (see Supporting Information; Figure S1) The abundance of deuterium was adjusted by adding the appropriate amounts of D_2O or H_2O. The samples were annealed at 85°C for 5 min (ensuring proper folding while all sites are exchanged to 90% D), and stored at 4°C for at least a night before use. The diluent solution used to trigger the D-to-H exchange was composed of 1 mM KCl and 100 mM TMAA in H_2O. The 9%-D references were prepared from samples incubated at 85°C for 5 min to ensure that they reflect the exact same mixing ratio than in the D-to-H exchange measurements.

HDX/MS experiments

Electrospray ionization mass spectrometry (ESI-MS) experiments were performed on a Thermo Orbitrap Exactive mass spectrometer calibrated daily, and operated in the negative mode, on a 400 – 3000 m/z scan range, using the 50 000-resolution setting, with the following tuning parameters: 2.95 kV capillary, 200°C capillary temperature, sheath gas: 55, aux gas: 5, heater temperature: 25°C, tube lens voltage: -198 V, skimmer voltage: -14.5 V. For denaturing conditions, the capillary temperature was set to 240°C. For softer conditions, the capillary temperature was set at 170°C, and the tube lens voltage

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

at -190 V. The relative softness of all three settings was assessed by infusing a 10-µM solution of the $[G_4T_4G_4]_2$ G-quadruplex in 100 mM ammonium acetate (Figure S2).[38] All data was acquired with Thermo Xcalibur 2.2.

Performing the experiments in the D-to-H direction allows working in a mostly non-deuterated solvent (9% deuterium), which is closer to physiological conditions and reduces the consumption of D_2O. The exchange was triggered by mixing the deuterated sample (90% D) with nine volumes of the same buffer, albeit non-deuterated (0% D), *i.e.* 100 mM TMAA, 1 mM KCl in pure H_2O. In manual mixing experiments, the sample and diluent were thoroughly mixed with a vortex and the resulting solution (9% D) was transferred in a 500-µL glass syringe, which was then used to flush the tubing dead volume, before being installed on a syringe pump operated at 5 µL/min. In continuous-flow experiments (Fig. 1C), the deuterated sample and the diluent were pumped at a 1:9 ratio towards a PEEK high pressure static mixing tee, equipped with a UHMWPE frit (IDEX Health & Science, Oak Harbor, WA, USA). The resulting mixed solution was brought to the MS source through PEEK tubings of different volumes (by changing both the length and internal diameter) to generate different mixing times. To vary the mixing time in a robust and simple way, an MX Series II 2 Position/6 Port UltraLife Switching Valve (IDEX Health & Science, Oak Harbor, WA, USA) was used (see Figure S3). The accuracy of the flow rates was verified by measuring the volumes delivered in two minutes. All experiments were performed at room temperature, i.e 22±1°C.

Raw HDX/MS data were analyzed manually, by visually selecting the isotope distributions for all exchange time points then calculating the centroid mass (Figure S4). Exchange curves can be plotted as the *m/z* centroid of the isotopic distributions of interest against time *t* (Figure S5). Centroids were converted to a number of unexchanged sites (*NUS*) using Equation 1.

$$NUS_t = \frac{[(m/z)_t - (m/z)_\infty] \times |z|}{\Delta m_s} \quad (1)$$

It is defined as the measured mass shift between the sample centroid mass $(m/z)_t$ at a time *t* and that of the fully-exchanged reference centroid $(m/z)_\infty$, converted to mass units by multiplying by $|z|$, and divided by the apparent isotopic mass shift for a single site (Δm_s). The value of $(m/z)_\infty$ will be discussed in the text. The value of Δm_s depends on the pre- and post-mixing isotopic compositions of the solvent, because the apparent isotopic mass of each site is the sum of the isotopic masses of hydrogen and deuterium (m_H and m_D, respectively), multiplied by their respective molar fractions in the medium (*H* or *D*),

$$\Delta m_s = (H_0 \times m_H + D_0 \times m_D) - (H_\infty \times m_H + D_\infty \times m_D) \quad (2)$$

Which simplifies into Equation 3 because $H = 1 - D$:

$$\Delta m_s = (D_0 - D_\infty)(m_D - m_H) \quad (3)$$

Unless otherwise mentioned, all experiments were performed here with $D_0 = 0.9$ and $D_\infty = 0.09$, yielding $\Delta m_s \approx 0.815$.

The exchange of a single site can typically be fitted with an exponential decay model. All sites are independent from one another and, unless there is evidence of cooperative unfolding, exchange at their respective rates.[4,6] Exchange curves acquired at the intact level therefore consists of the combination of the simultaneous exponential decay of all *n* sites,[5] and thus might not be fitted with a single exponential decay model. Some sites nevertheless display similar exchange rates, which reduces the total number of terms necessary to fit the data. Non-linear fitting of the *NUS* = f(*t*) exchange curves was therefore carried out with Equation 4, where NUS_0 is the offset, *k* is the rate constant, *N* is the number of sites in each group, and *j* the number of exponentials.

$$NUS_t = NUS_0 - \sum_{i=1}^{j} N_i e^{-k_i t} \quad (4)$$

Over large time ranges, non-linear fitting was typically performed with *j* = 3, and was reduced to *j* = 2 or 1 for fast exchanging species and/or when the exchange time range was restricted to either 1 s—5 min or 2 min—90 min, as is also usually observed for peptides in protein bottom-up experiments.[4]

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

RESULTS AND DISCUSSION

The exchangeable sites are located on the nucleobases

In DNA, the acquisition of the final charge is a partial neutralization of the phosphate groups by protons. The total number of sites *nX* that are potentially exchanged in a DNA oligonucleotide detected by ESI-MS in negative mode is therefore equal to the sum of all bases imino and amino groups, neutralized phosphate groups *nP*, and the -OH termini *T*. In an *n*-mer oligonucleotide capped by two OH termini $[dX_n]^{z-}$, there are *n*-1 phosphates and therefore $nP = n - 1 - z$. Furthermore, in MS of oligonucleotides, it is frequent to observe non-specific cation binding to the phosphates. We will suppose that the charge balance is ensured by phosphate neutralization. The number of neutralized phosphates of an *n*-mer oligonucleotide binding *m* cations *C* of absolute charge *y* is given by Equation 5.

$$nP_{[\mathrm{dX}_n+mC^{p+}-m\times p\times \mathrm{H}^+]^{z-}} = n - 1 - z - m \times p \qquad (5)$$

To determine the number of actually exchanged sites, we acquired the mass spectra of several oligonucleotides (dT_{12}, dT_{24}, 23nonG4, and N-myc) across a large range of deuterium content (10 to 90%; Figures S6 to S9). These sequences differ by their length (12 to 24 nucleotides), composition (polydT vs. mixed-base nucleotides), and only the latter is folded into a specific structure at room temperature, in 100-mM TMAA, 1-mM KCl solutions, specifically a G-quadruplex of T_m = 55.3 ± 0.3°C (*n* = 5; Figure S1). To ensure a quantitative isotopic exchange, all oligonucleotides were incubated 5 min at 85°C prior to MS measurements (N-myc is entirely unfolded at equilibrium above 72°C). These measurements were performed on different days by two distinct operators, using the MS tune used in all HDX experiments presented herein, as well as both a softer tune and a denaturing tune (see the experimental section) to check for possible gas-phase back-exchange. The centroid masses were plotted as a function of solution deuterium content and overlaid with the theoretical centroid values calculated for all possible number of exchangeable sites (0 to *nX*) and deuterium content (0—100%) to readily visualize the number of exchanging sites (Figure 2A and S10).

A linear relationship between the solution deuterium abundance and the centroid mass is observed for all oligonucleotides, for data replicated with the standard and softer tunes. The slopes are consistent with a statistical exchange of a defined number of unprotected sites, in solution, which follows the binomial law assuming the exchange of different sites is uncorrelated.[5,39,40] However, the number of actually exchanging sites is lower than the total number of theoretically exchangeable sites listed above. It is closest to the number of bases' exchangeable sites ($nX - nP - T$), although a few more sites are exchanged for T12 (one) and T24 (four). These results suggest that the phosphate groups are not (or only weakly) deuterated, regardless of the solution D_2O content.

To test this hypothesis, we monitored the exchange in ambient air, with additional H_2O vapor in the source (Figures S11 and S12), and with additional D_2O vapor in the source (Figures S13 and S14), coming from droplets lying at the bottom of the source during the experiment. Enriching the source with D_2O leads to increased deuteration (Figure 2B and Figures S15 to S18). Moreover, although there is still a linear relationship between the centroid mass and the solution D_2O content, the slopes are not consistent anymore with a defined number of solution-exchanging sites, highlighting the presence of a gas-phase contribution. However, the number of exchanged sites does not change when increasing the source H_2O partial vapor pressure.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

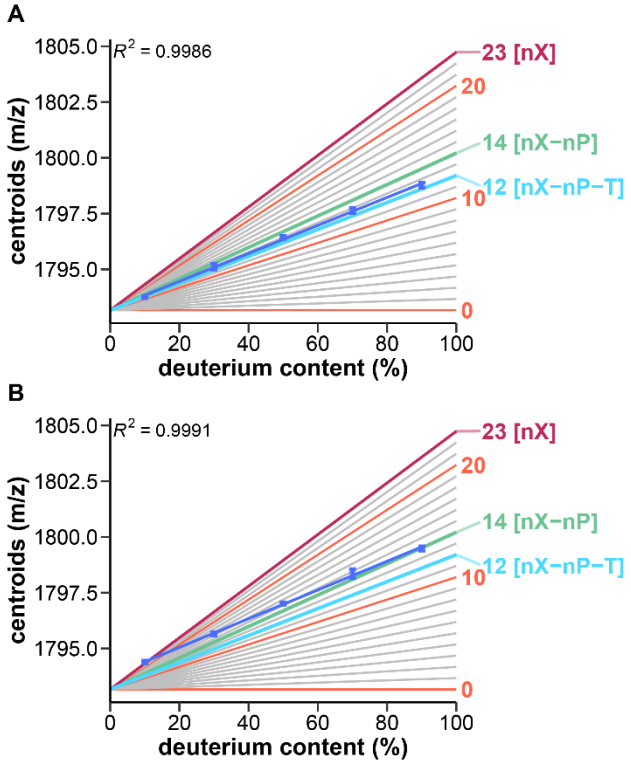


Figure 2. *Exchangeable sites in T12 for* z *= 2-, in 100 mM TMAA, 1 mM KCl, using the HDX tune in A) ambient conditions and B) in a D_2O-doped source atmosphere: Measured centroid masses against the solution deuterium content (standard deviation on duplicate measurements as error bars, linear fit as dark blue lines), overlaid over an array of theoretical masses across the full deuterium (0—100%) and exchangeable site ranges (0—*nX*).* nX *is the maximum number of potentially exchangeable sites (purple),* nX–nP *the case where the phosphates are not exchanged (green), and* nX–nP–T *the case for which the termini do not exchange either (cyan).*

We also noted that using ammonium acetate instead of trimethylammonium acetate yields 15% more exchanged sites for both T12 and T24 (Figures S19 to S21). The buffer choice thus influences gas-phase back-exchange. Trimethylammonium is a protic, tertiary ammonium ionic liquid (IL), less volatile and acidic than ammonium (boiling point ≈ 3 vs. -33°C, pK_a = 9.25 vs. 9.80).[41] The cation and anion of an IL can for ion pair via hydrogen bonding, and some (including ammonium-based ILs such as TMAA) can H-bond with water as well, leading to altered viscosities and acidities.[41–44] These differences in properties might alter the droplets' composition during the electrospray mechanism, leading to changes in the neutralization of phosphates.

A possible explanation to the virtual absence of phosphate deuteration is that, in ambient conditions, the humidity in the source transfers protons (or deuterons in presence of D_2O vapors) to remaining TMA (or ammonium) cations capping DNA, and then to the DNA phosphates, provided that the de-clustering finishes well after the ions escape the droplets (Figure 3B). As a result, the phosphate sites are deuterated to an extent that depends on both the atmosphere isotope composition and buffer cation nature. This is consistent with the strong acidity of phosphates that makes them prone to rapid back-exchange to H, likely with vapors coming from ambient air. It is further corroborated by the fact that, although the number of neutralized (and potentially exchangeable) phosphates increases which decreasing charges (Equation 5), the number of effectively exchanged sites remains constant across charge states for all oligonucleotides.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

When infusing the samples in harsher source conditions, the number of exchanged sites decreases for all oligonucleotides, below the number of bases's sites in some cases (N-myc and 23nonG4; Figures S22 to S26). Back-exchanging the bases is therefore possible as well, but is totally avoided when using sufficiently soft source conditions.

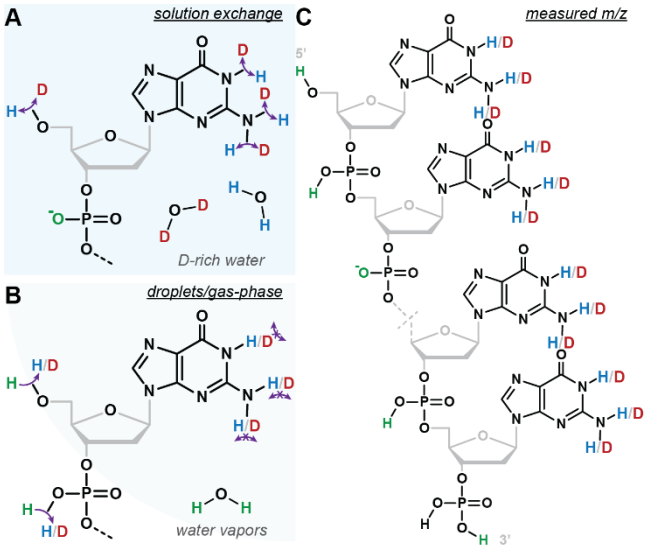


Figure 3. *Exchangeable sites in native HDX/MS of TMAA-solutions of DNA oligonucleotides. A) Unpaired nucleobases exchange to reflect the isotopic solvent composition, whereas phosphates remain ionized until B) the electrospray process during which some (depending on the charge) are neutralized by either H or D. Ultimately they are back-exchanged by the protons from the ambient H_2O vapors (unless the source was doped with D_2O vapors as in Figure 2B). Nucleobases do not back-exchange, leading to C) the final analyte in which the nucleobases reflect the solution exchange and the phosphate are either ionized or protonated.*

In summary, we have set up a robust approach and robust conditions to quantify the number of exchanging sites by MS, and demonstrated that it is consistent with the exchangeable sites being located on the nucleobases (Figure 3A), without gas-phase back-exchange if using relatively soft ionization conditions (Figure 3B). The virtual absence of phosphate deuteration on the final analyte using TMAA is not detrimental, and may in fact be considered as an advantage, as it streamlines the data analysis without information loss (Figure 3C).

Implications for converting the *m/z* distribution into unexchanged sites equivalent and interpreting *NUS* values

In HDX/MS experiments, monitoring *m/z* shifts as a function of time is sufficient to obtain relative trends, but does not provide an absolute number-of-site information, and complicates the comparison across species of different masses. Instead, we opted to convert the y-axis to a number of unexchanged sites (*NUS*) using equation 1. The *NUS* reports on the absolute number of non-exchanged sites. This differs from the "deuterium level" often used in the protein HDX/MS literature, which calculates the absolute number of *deuterium* incorporated, using 0%- and 100%-D references.[4–6] The main advantages of the *NUS* value are (1) that it provides a direct readout of the apparent number of sites that are involved in secondary structure formation, and (2) only requires measuring the centroid of the fully exchanged reference, which is straightforward. Note however that a *NUS* of 1 does not mean that all sites are fully exchanged except for one entirely protected, but rather that the overall, apparent amount of unexchanged is *equivalent* to one site (e.g. five 20%-unexchanged sites would give a *NUS* of 1).

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

The conversion to *NUS* values requires the acquisition of a mock end-point, which mimics the state of the oligonucleotides upon exchange completion, as described in the experimental section. Following our findings on phosphate and terminal-OH back-exchange, we compared the experimentally acquired mock-end point spectra of a dozen oligonucleotides to the corresponding theoretical isotopic distributions, following the postulate that only the exchange of nucleobases is monitored (Figures S27 and S28). Mass errors down to the 10-ppm range were obtained, which highlights the satisfactory accuracy in deuterium content in the HDX/MS experiments described below, and further validates the localization of the exchanged sites on the bases.

Exchange rates can be measured precisely on a large time scale, regardless of the charge state and adduct stoichiometry

To monitor the isotopic distributions upon exchange, the sample and diluent solutions can be manually mixed to trigger the exchange, then directly infused in the mass spectrometer (Figure 4 - squares). This gives access to time scales from a few minutes to several hours (Figure S5) or even days.

Relying on HDX/MS to study oligonucleotide structures and biophysics requires sufficient precision in measuring the exchange rates, and ideally no dependence on artifacts related to ionization such as the charge state or the formation of non-specific adducts. We identified several potential sources of errors on both the time and mass dimensions. Errors in masses may arise from the mass spectrometry measurement itself (mass spectrometer accuracy, inaccurate peak selection and integration, particularly in the case of partially overlapped isotopic distributions, weak signal and poor signal-to-noise ratio) as well as from inaccurate deuterium content before (sample preparation errors) and during exchange (inaccurate mixing ratio). The measurement of the mixing dead time (for manual-mixing experiments), and temperature variations during the exchange may also be sources of errors on the HDX rate.

To evaluate the precision of the method, two operators ran triplicate experiments on three G-quadruplex-forming oligonucleotides, namely N-myc, T30177-TT, and 26CEB. These are three-tetrad G-quadruplexes of different lengths (19—26 nucleotides) and number of exchanging sites (45—62). However, they have similar thermal stabilities (T_m = 52—55°C; Table S1 and Figure S1), and apparent exchange rates. For each replicate, the data of the three most abundant potassium stoichiometries from the three most intense charge states were examined (Figures S29 to S38). The mean *NUS* values, and their corresponding standard deviations and RSDs across replicates were calculated for each time point (Figure 4A, Figure S39A,B). Overall, the method has an excellent precision, with low mean RSD, whether all species (3.7%, 3.1%, and 4.4% for N-Myc, T30177-TT, and 26CEB, respectively) or only the most abundant ones (2.7%, 3.0%, and 4.6%) are considered (Table S2). The corresponding mass errors are inferior or equal to 0.2 Da (Table S3), which is on par with well-controlled protein HDX experiments (precise sample preparation, stable mass spectrometer calibration).[6,45–47]

Despite being quite similar (three-tetrad parallel quadruplexes of close stability), all three oligonucleotides can therefore be discriminated from each other using the exchange data. This can be achieved by visual inspection (Figure 4A), pairwise statistical comparison of individual time points (Table S4) or fitting parameters (Table S5). Comparing apparent exchange rates may not be reliable because it reduces numerous sites exchanging at their own rates to a few apparent rates, which masks part of the information and can lead to nonconclusive results (the 95% confidence intervals can overlap). The two former approaches are straightforward and the large number of data points acquired (compared to classical bottom-up approaches) allows to confidently conclude that these are three different exchange behaviors. The results on this set of

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

oligonucleotides suggest that HDX/MS is well-suited for the quality control analysis of oligonucleotide folding, for instance during the production of structured oligonucleotides such as aptamers (batch-to-batch consistency, stability and stress testing), or when introducing modifications that may alter the structure (e.g. modified nucleotides, biotinylation, fluorophores).

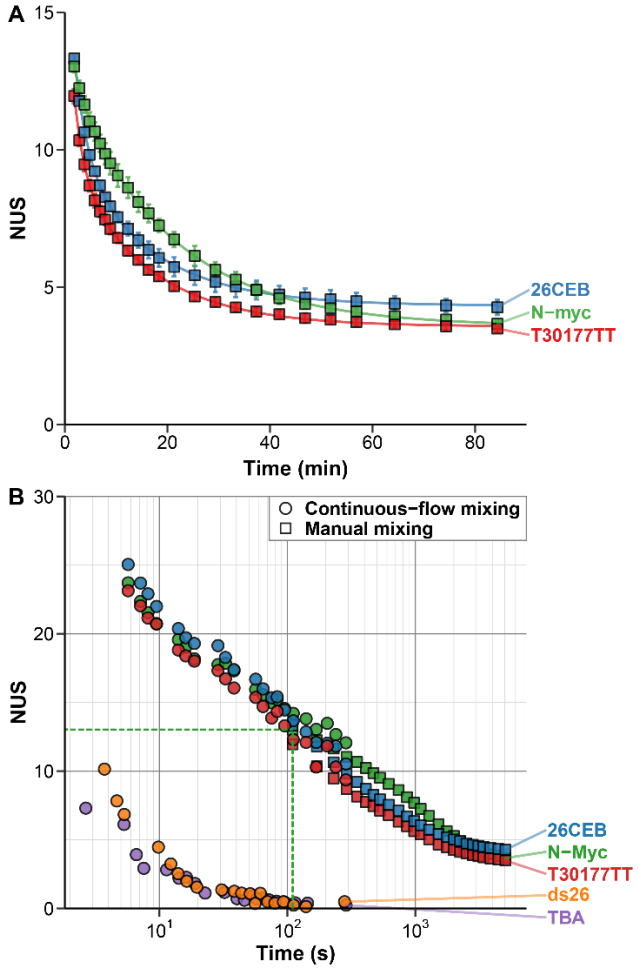


Figure 4. *Exchange kinetics of oligonucleotides acquired by manual mixing (A and B; squares) and continuous-flow (B; circles) native HDX/MS (90% to 9% D) for five oligonucleotides. A) The error bars show the standard deviation across all species (z = 3 to 5- for N-myc and T30177TT, and 4 to 6- for 26CEB,* K^+ *= 2 to 4) for three replicates. The lines are the results of the non-linear fitting using Equation 4 with* j *= 2. B) The dashed green lines indicate the position of the earliest N-Myc point acquired by manual mixing (110 s).*

In a typical native mass spectrum of oligonucleotide, a few charge states are detected (here, 3- to 6- depending on the sequence; Figure S29). We did not observe any significant differences in exchange rates between different charge states (Figure S39C). Hence, any or all charge states can be selected for data treatment; those of higher intensity and/or with fewer non-specific cation binding should thus typically be favored. Additionally, all four oligonucleotides bind specifically two K^+ to form 3-tetrad quadruplexes, whereas all higher stoichiometries (i.e. 3 and 4 K^+) arise from additional non-specific binding. No significant differences in the exchange rates between the 2, 3, and 4 K^+ species were observed, across all replicates and charge states (Figure S39D), and therefore any or all can be used for data treatment. In absence of knowledge on cation specificity, it is however necessary to process each stoichiometry separately to assess which can be used, and possibly detect the coexistence of several conformers in solution.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In manual-mixing experiments, fast exchanging phenomena can be missed within the dead time. For instance, the N-myc sequence has 50 exchanging sites, but at the earliest continuous flow time point (110 s; signal integrated from 50 s to 170 s) only the equivalent of 13 sites remain unexchanged (Figure 4, dashed green line). To access faster exchange events, a continuous flow setup was employed, wherein the sample is continuously mixed with the exchanging buffer for a fixed duration. Continuous-flow mass spectrometry is used to study time-resolved events,[7,48] and has already been employed in HDX/MS experiments by Wilson and Konermann.[49–51] Another advantage is to allow recording a given mixing time point for as long as necessary. This is particularly interesting for analyzing low-concentration samples, as well as low-abundance and/or fast-exchanging species, for which real time analysis would require the extraction of mass spectra over a narrow time range, which could lead to poor S/N. Here, continuous flow was used to access exchange events ranging from around 1 second to 5 minutes (Figure 4B - circles). The exchange of 10 additional sites could be monitored for N-myc. Importantly, this enables the analysis of fast-exchanging species such as 2-tetrad G-quadruplexes (e.g. TBA; Figure S40) and dsDNA (e.g. ds26; Figure S41) that is not possible by manual-mixing. The results of the manual and continuous flow mixing techniques can be combined to obtain exchange curves over labeling times spanning several orders of magnitude (Figure 4B), as is recommended in protein HDX/MS,[52] giving access to both weakly and strongly protected sites.

Exchange rates on the second-to-day time scale are sensitive to oligonucleotide folding

To be used in structural and biophysical studies, HDX/MS of oligonucleotides must depend on site exchange availability, or *hydrogen bonding status*. A simple comparison can be made with 24TTG and 24nonG4, two sequences of close composition and same length but of different structures. The unfolded 24nonG4 oligonucleotide, whose sites are all available for exchange, is fully exchanged within a second, likely at the chemical exchange rates of the individual sites (Figure 5A). Conversely, the G-quadruplex 24TTG is still partially unexchanged after 5 minutes of incubation, most probably because some sites are engaged in hydrogen bonding and therefore protected from exchange. In the next example, we tested the sequence 222T, which forms different G-quadruplex conformers depending on the K^+ concentration.[24] First, we observed clearly different exchange profiles for 222T•2K^+, a three-tetrad quadruplex, in 0.1 or 1.0 mM KCl solutions (Figure 5B). Second, such differences can be easily measured for different species in equilibrium by operating in native MS conditions. Here, a two-tetrad quadruplex 222T•1K^+ and an unfolded conformer 222T•0K^+ are also detected (Figure 5C): 222T•1K^+ exchanges quicker than 222T•2K^+, and 222T•0K^+ exchanges within a second owing to the absence of base pairing (Figure 5B). The result also means that the three complexes do not interconvert faster than the time scale of the experiment. This showcases the advantage of coupling native MS to HDX to characterize several conformers in a mixture, which can be challenging with conventional methods. More generally, all the unfolded oligonucleotides we essayed systematically exchanged fully in less than a second (and likely close to their intrinsic exchange rates owing to the lack of hydrogen bonding). Conversely, structured oligonucleotides were all found to be protected from exchange to an extent that depends on their precise conformations. Significant differences in exchange rates were indeed observed in the examples above (Figures 4 and 5), as well as in a larger panel of G-quadruplex structures, even across structures of similar numbers of protected sites (Figures S42), further demonstrating that the exchange rates are strongly dependent on the precise conformers at hand. This opens promising avenues for the use of native HDX/MS nucleic acids biophysics by native HDX/MS.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

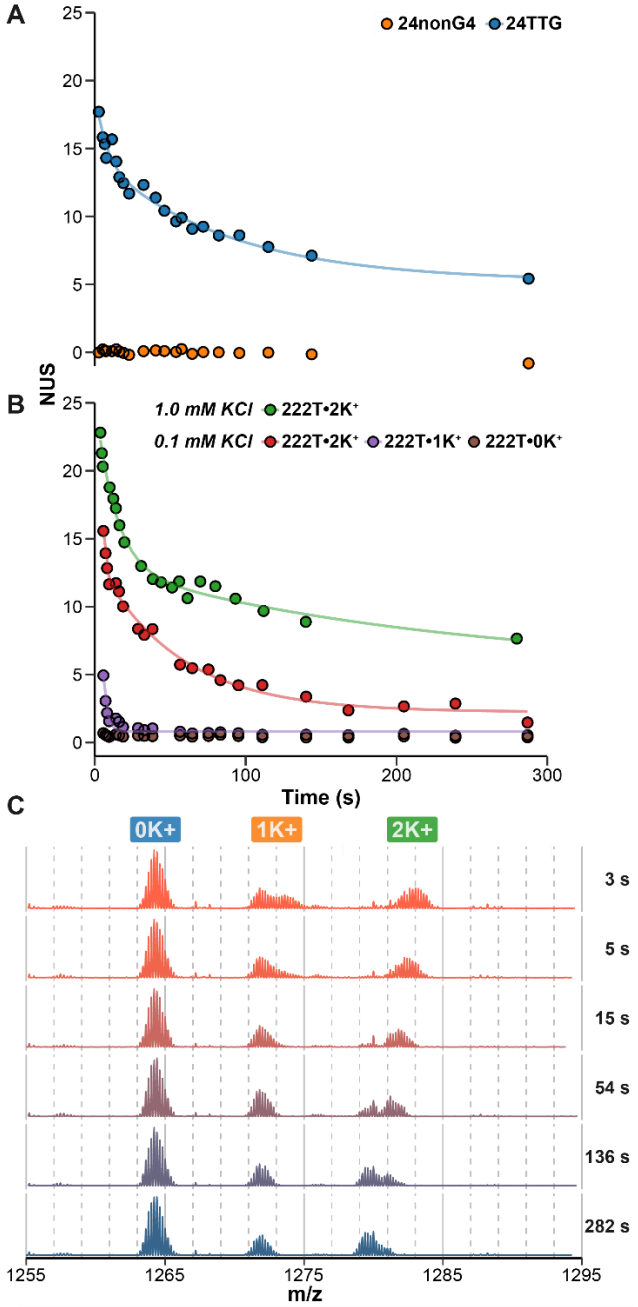


Figure 5. *Hydrogen bonding status dependency of oligonucleotide HDX/MS. Exchange kinetics acquired by continuous-flow HDX (90% to 9% D) of A) unstructured (24nonG4) and structured (24TTG•2K⁺) oligonucleotides, and B) the same oligonucleotide (222T) in presence of 1.0 or 0.1 mM KCl, which leads to the formation of three distinct species (the G-quadruplexes 222T•2K⁺ and 222T•1K⁺, and the unfolded conformer 222T•0K⁺). The lines are the results of nonlinear fitting using Equation 4 with* j *= 1 or 2 for 2- and 3-tetrad species, respectively. C) Corresponding spectra for 222T in 0.1 mM KCl solution, zoomed on the 5- charge state. The K⁺-binding stoichiometry is indicated with labels. No exchange occurs in the explored time range for unfolded species (0 K⁺), unlike the G-quadruplexes species (≥ 1K⁺).*

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CONCLUSION

We demonstrated that oligonucleotides are perfectly amenable to HDX/native MS, by using DNA G-quadruplexes as models. Most notably, we have established that:

i) Nucleobases are isotopically exchanged proportionally to the solution deuterium content without in-source back-exchange, whereas phosphates are virtually entirely back exchanged, using TMAA as a buffer and relatively soft ionization conditions.

ii) The apparent exchange rates depend strongly on the specific hydrogen bonding status of DNA bases, and therefore on oligonucleotides' secondary structure dynamics.

iii) Mass spectrometry can precisely measure these exchange rates over a wide time range (seconds to minutes with continuous flow setups, minutes to days with manual mixing), without influence from the charge state or the presence of non-specific adducts.

The technique will find numerous applications in nucleic acids biophysics. The relatively low concentrations of oligonucleotides (here 5 µM) allows avoiding the formation of non-physiologically-relevant multimers.[16] It does not require digestion, quenching and chromatographic separation steps, found routinely in bottom-up HDX/MS analysis of proteins. Native HDX/MS also does not require expensive or highly specialized equipment other than the mass spectrometer itself. The use of native MS gives concomitant access to other biophysical parameters (*e.g.* binding constants, stoichiometries), opening up promising avenues for the analysis of oligonucleotide structures and interactions. In particular, coupling to ion mobility mass spectrometry to further disentangle complex conformational ensembles and small-molecule ligand binding analysis that cannot be straightforwardly deciphered by NMR or X-ray crystallography, will be presented in further communications.

The method can also find applications in the quality control of structured oligonucleotides, such as aptamers. Thanks to its simple and high precision acquisition of a large number of data points, native HDX/MS of oligonucleotide is well suited for typical quality control procedures such as structure batch-to-batch consistency, stability over time, and stress testing (e.g. from oxidant, high temperature, non-physiological pH). Modified sequences are often used in oligonucleotide therapeutics (modifications on e.g. the nucleobase to increase binding affinities, phosphodiester backbone for resisting nuclease degradation, PEGylation for lowering renal clearance),[53] probes (e.g. fluorophore coupling or fluorescent base analogues),[54] PCR primers[55] and affinity tags[56] (biotinylation). Another possible application of native HDX/MS is therefore the assessment of the influence of modifications on oligonucleotide structural dynamics.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ASSOCIATED CONTENT

Supporting Information

Figures S1—S42, Tables S1—S5, UV-melting protocol

AUTHOR INFORMATION

Corresponding Author

* eric.largy@u-bordeaux.fr

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

This work was supported by a grant from the FR TECSAN. We thank Frédéric Rosu at the structural biophysico-chemistry platform of the Institut Européen de Chimie et Biologie for the access to mass spectrometers and his invaluable support. We gratefully acknowledge Laura Fricot, Anaïs Ferrer, Emile Feugas, and Benjamin Lienard for their technical assistance.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

REFERENCES

(1) Oganesyan, I.; Lento, C.; Wilson, D. J. *Methods* **2018**, *144*, 27–42.

(2) Trabjerg, E.; Nazari, Z. E.; Rand, K. D. *TrAC Trends Anal. Chem.* **2018**, *106*, 125–138.

(3) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. *Expert Rev. Proteomics* **2011**, *8* (1), 43–59.

(4) Konermann, L.; Pan, J.; Liu, Y.-H. *Chem. Soc. Rev.* **2011**, *40* (3), 1224–1234.

(5) Zhang, Z.; Smith, D. L. *Protein Sci.* **1993**, *2* (4), 522–531.

(6) Wales, T. E.; Eggertson, M. J.; Engen, J. R. In *Mass Spectrometry Data Analysis in Proteomics*; Matthiesen, R., Ed.; New Jersey, 2013; pp 263–288.

(7) Urban, P. L.; Chen, Y.-C.; Wang, Y.-S. In *Time-Resolved Mass Spectrometry: From Concept to Applications*; John Wiley & Sons, Ltd: Chichester, UK, 2016; pp 291–314.

(8) Kostyukevich, Y.; Acter, T.; Zherebker, A.; Ahmed, A.; Kim, S.; Nikolaev, E. *Mass Spectrom. Rev.* **2018**, *37* (6), 811–853.

(9) Kaltashov, I. A.; Eyles, S. J. In *Mass Spectrometry in Structural Biology and Biophysics*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; pp 127–159.

(10) Gülbakan, B.; Barylyuk, K.; Schneider, P.; Pillong, M.; Schneider, G.; Zenobi, R. *J. Am. Chem. Soc.* **2018**, *140* (24), 7486–7497.

(11) Printz, M. P.; Von Hippel, P. H. *Biochemistry* **1968**, *7* (9), 3194–3206.

(12) Printz, M. P.; Hippel, P. H. v. *Proc. Natl. Acad. Sci.* **1965**, *53* (2), 363–370.

(13) von Hippel, P. H.; Johnson, N. P.; Marcus, A. H. *Biopolymers* **2013**, *99* (12), 923–954.

(14) Lee, J.-H.; Pardi, A. *Nucleic Acids Res.* **2007**, *35* (9), 2965–2974.

(15) Mergny, J.-L.; Sen, D. *Chem. Rev.* **2019**, *119* (10), 6290–6325.

(16) Largy, E.; Mergny, J.-L. *Nucleic Acids Res.* **2014**, *42* (19), e149.

(17) Rhodes, D.; Lipps, H. J. *Nucleic Acids Res.* **2015**, *43* (18), 8627–8637.

(18) Millevoi, S.; Moine, H.; Vagner, S. *Wiley Interdiscip. Rev. RNA* **2012**, *3* (4), 495–507.

(19) O. Tucker, W.; T. Shum, K.; A. Tanner, J. *Curr. Pharm. Des.* **2012**, *18* (14), 2014–2026.

(20) Platella, C.; Riccardi, C.; Montesarchio, D.; Roviello, G. N.; Musumeci, D. *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861* (5), 1429–1447.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(21) Yaku, H.; Murashima, T.; Miyoshi, D.; Sugimoto, N. *Sensors* **2015**, *15* (4), 9388–9403.

(22) Park, J. Y.; Cho, Y. L.; Chae, J. R.; Moon, S. H.; Cho, W. G.; Choi, Y. J.; Lee, S. J.; Kang, W. J. *Mol. Ther. - Nucleic Acids* **2018**, *12*, 543–553.

(23) Largy, E.; Mergny, J.-L.; Gabelica, V. In *Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; Springer International Publishing, 2016; Vol. 16, pp 203–258.

(24) Largy, E.; Marchand, A.; Amrane, S.; Gabelica, V.; Mergny, J.-L. *J. Am. Chem. Soc.* **2016**, *138* (8), 2780–2792.

(25) Marchand, A.; Granzhan, A.; Iida, K.; Tsushima, Y.; Ma, Y.; Nagasawa, K.; Teulade-Fichou, M.-P.; Gabelica, V. *J. Am. Chem. Soc.* **2015**, *137* (2), 750–756.

(26) Adrian, M.; Heddi, B.; Phan, A. T. *Methods* **2012**, *57* (1), 11–24.

(27) Hsu, S.-T. D.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. *J. Am. Chem. Soc.* **2009**, *131* (37), 13399–13409.

(28) Kuryavyi, V.; Cahoon, L. A.; Seifert, H. S.; Patel, D. J. *Structure* **2012**, *20* (12), 2090–2102.

(29) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V. *J. Am. Chem. Soc.* **2008**, *130* (31), 10208–10216.

(30) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; Pauw, E. De. *Rapid Commun. Mass Spectrom.* **2002**, *16* (18), 1729–1736.

(31) Marchand, A.; Rosu, F.; Zenobi, R.; Gabelica, V. *J. Am. Chem. Soc.* **2018**, *140* (39), 12553–12565.

(32) Marchand, A.; Gabelica, V. *Nucleic Acids Res.* **2016**, *44* (22), 10999–11012.

(33) Marchand, A.; Gabelica, V. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1146–1154.

(34) Trajkovski, M.; Webba Da Silva, M.; Plavec, J. *J. Am. Chem. Soc.* **2012**, *134* (9), 4132–4141.

(35) Amrane, S.; Adrian, M.; Heddi, B.; Serero, A.; Nicolas, A.; Mergny, J.-L.; Phan, A. T. *J. Am. Chem. Soc.* **2012**, *134* (13), 5807–5816.

(36) Mukundan, V. T.; Phan, A. T. *J. Am. Chem. Soc.* **2013**, *135* (13), 5017–5028.

(37) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235* (5), 1532–1547.

(38) Balthasart, F.; Plavec, J.; Gabelica, V. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (1), 1–8.

(39) Guttman, M.; Weis, D. D.; Engen, J. R.; Lee, K. K. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (12), 1906–1912.

(40) Palmblad, M.; Buijs, J.; Håkansson, P. *J. Am. Soc. Mass Spectrom.* **2001**, *12* (11), 1153–1162.

bioRxiv preprint doi: https://doi.org/10.1101/848598; this version posted November 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(41) Greaves, T. L.; Drummond, C. J. *Chem. Rev.* **2008**, *108* (1), 206–237.

(42) Umapathi, R.; Attri, P.; Venkatesu, P. *J. Phys. Chem. B* **2014**, *118* (22), 5971–5982.

(43) Dong, K.; Zhang, S. *Chem. - A Eur. J.* **2012**, *18* (10), 2748–2761.

(44) Greaves, T. L.; Drummond, C. J. *Chem. Rev.* **2015**, *115* (20), 11379–11448.

(45) Houde, D.; Berkowitz, S. A.; Engen, J. R. *J. Pharm. Sci.* **2011**, *100* (6), 2071–2086.

(46) Iacob, R. E.; Engen, J. R. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (6), 1003–1010.

(47) Hudgens, J. W.; Gallagher, E. S.; Karageorgos, I.; Anderson, K. W.; Filliben, J. J.; Huang, R. Y. C.; Chen, G.; Bou-Assaf, G. M.; Espada, A.; Chalmers, M. J.; et al. *Anal. Chem.* **2019**, *91* (11), 7336–7345.

(48) Urban, P. L.; Chen, Y.-C.; Wang, Y.-S. In *Time-Resolved Mass Spectrometry: From Concept to Applications*; John Wiley & Sons, Ltd: Chichester, UK, 2016; pp 89–156.

(49) Wilson, D. J.; Konermann, L. *Anal. Chem.* **2003**, *75* (23), 6408–6414.

(50) Resetca, D.; Wilson, D. J. *FEBS J.* **2013**, *280* (22), 5616–5625.

(51) Konermann, L.; Simmons, D. A. *Mass Spectrom. Rev.* **2003**, *22* (1), 1–26.

(52) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; et al. *Nat. Methods* **2019**, *16* (7), 595–602.

(53) Ni, S.; Yao, H.; Wang, L.; Lu, J.; Jiang, F.; Lu, A.; Zhang, G. *Int. J. Mol. Sci.* **2017**, *18* (8), 1683.

(54) Boutorine, A.; Novopashina, D.; Krasheninina, O.; Nozeret, K.; Venyaminova, A. *Molecules* **2013**, *18* (12), 15357–15397.

(55) Bengtström, M.; Harju, L.; Syvhen, A.-C. *Nucleosides Nucleotides* **1991**, *10* (1–3), 507–509.

(56) Largy, E.; Hamon, F.; Teulade-Fichou, M.-P. *Methods* **2012**, *57* (1), 129–137.