Biochimica et Biophysica Acta xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

What stoichiometries determined by mass spectrometry reveal about the ligand binding mode to G-quadruplex nucleic acids☆

Michael J. Lecours [a], Adrien Marchand [b], Ahdia Anwar [a], Corinne Guetta [c], W. Scott Hopkins [a], Valérie Gabelica [b,*]

[a] Department of Chemistry, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[b] INSERM, CNRS, Université de Bordeaux, Laboratoire Acides Nucléiques: Régulations Naturelle et Artificielle (ARNA, U1212, UMR5320), IECB, 2 rue Robert Escarpit, 33607 Pessac, France
[c] Institut Curie, CNRS, INSERM, Univ. Paris Sud, Laboratoire de Chimie, modélisation et imagerie pour la biologie (CMIB, U1196, UMR9187), Orsay, France

ARTICLE INFO

Article history:
Received 15 October 2016
Received in revised form 6 January 2017
Accepted 9 January 2017
Available online xxxx

Keywords:
G-quadruplex
Mass spectrometry
Ligand
PhenDC3
360A
Pyridostatin
Conformation
Ligand-induced isomerization

ABSTRACT

G-quadruplexes (G4s) have become important drug targets to regulate gene expression and telomere maintenance. Many studies on G4 ligand binding focus on determining the ligand binding affinities and selectivities. Ligands, however, can also affect the G4 conformation. Here we explain how to use electrospray ionization mass spectrometry (ESI-MS) to monitor simultaneously ligand binding and cation binding stoichiometries. The changes in potassium binding stoichiometry upon ligand binding hint at ligand-induced conformational changes involving a modification of the number of G-quartets. We investigated the interaction of three quadruplex ligands (PhenDC3, 360A and Pyridostatin) with a variety of G4s. Electrospray mass spectrometry makes it easy to detect K^+ displacement (interpreted as quartet disruption) upon ligand binding, and to determine how many ligand molecules must be bound for the quartet opening to occur. The reasons for ligand-induced conversion to antiparallel structures with fewer quartets are discussed. Conversely, K^+ intake (hence quartet formation) was detected upon ligand binding to G-rich sequences that did not form quadruplexes in 1 mM K^+ alone. This demonstrates the value of mass spectrometry for assessing not only ligand binding, but also ligand-induced rearrangements in the target sequence. This article is part of a Special Issue entitled "G-quadruplex" Guest Editor: Dr. Concetta Giancola and Dr. Daniela Montesarchio.

1. Introduction

Guanine-rich nucleic acids can form G-quadruplex (G4s) structures by the stacking of two or more G-quartets (tetrads) [1]. G-quadruplex structures are stabilized by hydrogen bonds between the guanines of a G-quartet, and by the intercalation of cations such as potassium (K^+) between the stacked G-quartets. G4 structures prevail in important regions of the human genome such as telomeres, oncogene promoters and transcription start sites [2], and have become potential anti-cancer drug targets [3].

G4s exhibit diverse structures. The topologies depend on many factors: the number of strands (intramolecular, bimolecular or tetramolecular G4s), the number of nucleobases in the strand, nucleobase orientation (*syn* or *anti*) in the G-quadruplex core, and the way the loops connect the G-tracts. Moreover, some sequences are polymorphic: they can form multiple topologies depending on the environmental conditions. For example, human telomeric sequences [4], consisting of TAGGGT repeats, can form antiparallel G4s in sodium-containing solution conditions, and hybrid G4s in potassium-containing solutions [5–7]. Owing to their high of degree structural polymorphism, G4s can also be used in switchable nanodevices [1,8,9].

At least four potential G4 ligand binding modes can be distinguished: intercalation, groove binding, loop binding, and stacking on external G-quartets. But although in principle one could design ligands that would be selective for a specific G4 structure, most ligands reported to date comprise large aromatic planes and hence bind mainly to external G-quartets [10,11]. As a result, ligand inter-G4 selectivity has not been well explored. Only a few examples G4-selective ligands have been reported, *e.g.*, the ligand NMM, which binds preferentially to parallel structures [12].

Ligand binding can also induce conformational switching in polymorphic sequences. For example, the human telomeric G4 structures convert to more hybrid G-quadruplexes when bound to the TMPyP4 ligand [13], or to antiparallel structures when bound to Cu-tolylterpyridine [14]. We have previously shown that the human telomeric G4 structures change to antiparallel geometries with one K^+, therefore containing presumably two G-quartets [15], upon binding of ligands 360A [16,17], PhenDC3 [18] or Pyridostatin (PDS) [19]. The strand arrangement was characterized by circular dichroism, and the cation binding stoichiometry was determined by electrospray ionization mass spectrometry (ESI-MS). Because these first results revealed

☆ This article is part of a Special Issue entitled "G-quadruplex" Guest Editor: Dr. Concetta Giancola and Dr. Daniela Montesarchio.

* Corresponding author.
E-mail address: v.gabelica@iecb.u-bordeaux.fr (V. Gabelica).

http://dx.doi.org/10.1016/j.bbagen.2017.01.010
0304-4165/

intriguing properties for these three ligands, which are among the most widely used in the community, we decided to expand our ESI-MS study of PhenDC3, 360A and PDS to many more G-rich sequences. We report here several types of conformational changes that can be easily inferred from the mass spectra, based on the detected changes in K^+ stoichiometries upon ligand binding. The existence of these structural changes in solution is validated by circular dichroism (CD) spectroscopy.

2. Materials and methods

2.1. Materials

Oligonucleotides were purchased from Eurogentec (Seraing, Belgium) in reverse-phase purified lyophilized form (RP cartridge-Gold quality). Solutions were prepared in nuclease-free water (Ambion, Life technologies SAS, Saint-Aubin, France). Table 1 includes the short name and sequence for the oligonucleotides used in this work.

2.2. Solution preparation

Concentrations of stock solutions in H_2O were measured by UV absorption at 260 nm on a Uvikon XS. Molar absorption coefficients were obtained from the IDT Website using the Cavaluzzi-Borer correction [33]. Stock solutions were then diluted to 50–200 μM of single-stranded DNA in 100 mM trimethylammonium acetate (TMAA, Ultra for HPLC, Fluka analytical), and 1 mM of potassium chloride was added (>99.999%, Sigma). Intramolecular G4s were allowed to fold for at least 18 h. Bimolecular and tetramolecular G4s were allowed one week to fold at 200 μM single strand. Solutions of intramolecular and bimolecular quadruplexes were prepared at 10 μM quadruplex concentrations in a 1:1 ligand to G4 equivalents, while $[TG4T]_4$ was prepared at 5 μM of G4 (20 μM single strand) and analyzed at 2:1 ligand to G4 concentrations. Solutions were stored in a fridge at 4 °C for the duration of the screening preparation, and were allowed to remain overnight with the ligand before injection. Solutions for CD were prepared at 5 μM of G4 and up to 15 μM of ligand.

Ligands 360A (iodide salt) and PhenDC3 (trifluoromethyl sulfonate salt) were donated by Marie-Paule Teulade-Fichou [18], and Pyridostatin (trifluoroacetate salt) was purchased from Sigma-Aldrich. The concentrations were determined using molar ellipticity coefficients of 40,000 cm^{-1} M^{-1} at 260 nm for 360 A, 62,400 cm^{-1} M^{-1} at 320 nm for PhenDC3 and 67,500 cm^{-1} M^{-1} at 227 nm for PDS. It should be noted that when older solutions (~18 months) of PDS (stored at −20 °C at 2 mM concentration in H_2O) and 360A (stored at 4 °C at 200 μM concentration in H_2O) were analyzed, G4 binding of partially degraded ligands was observed (see Fig. S1). All results reported here were carried out with solutions prepared less than a month before analysis.

Table 1
Oligonucleotide sequences used in this screening, where available articles containing NMR structures of these sequences have been referenced.

Anticipated structure	Short name	DNA sequence 5′ to 3′
Intramolecular parallel G4s	222T [8]	TGGGTTGGGTTGGGTTGGGT
	Pu24 [20]	$TGAG_3TG_3GAG_3TG_4A_2G_2$
	26CEB [21]	$A_2G_3TG_3TGTA_2GTGTG_3TG_3T$
	Bulge-TB1 [22]	TTGTGGTGGGTGGGTGGGT
Intramolecular telomeric G4s	22GT [23]	$(GGGTTA)_3GGGT$
	22CTA [24]	$A(GGGCTA)_3GGG$
	22AG	$A(GGGTTA)_3GGG$
	23TAG [25]	$TA(GGGTTA)_3GGG$
	23AG [26]	$A(GGGTTA)_3GGGT$
	24TTG [27]	$TT(GGGTTA)_3GGGA$
	25TAG [25]	$TA(GGGTTA)_3GGGTT$
Bimolecular G4s	$[G3T4G4]_2$ [28]	2 × GGGTTTTGGGG
	$[G4T4G3]_2$ [28]	2 × GGGGTTTTGGG
	$[G4T4G4]_2$ [29,30]	2 × GGGGTTTTGGGG
	$[G4T3G4]_2$ [30]	2 × GGGGTTTGGGG
	$[12TAG]_2$ [31]	2 × TAGGGTTAGGGT
Tetra-molecular G4	$[TG4T]_4$ [32]	4 × TGGGGT
Artificial sequences classified by G4Hunter scores	22non105 [15]	$G_3ATGCGACAGAGAGGACG_3$
	23non100 [15]	$TG_3ATGCGACAGAGAG_2ACG_3$
	24non096 [15]	$TG_3ATGCGACAGAGAG_2ACG_3A$
	26non088 [15]	$T_2G_3ATGCGACAGAGAG_2ACG_3AT$
	22non059	TGCGAGCGAGAGAGGAGCGAGG
	22non068	TGGCGAGCGAGAGAGGACGAGG
	22non077	TGCGAGCAGGAGAGGACGGAGG
	22non086	TGGCGAGACAGAGAGGAGCGGG
	22non095	TGCGACAGGAGGAGGGACGAGG
	22non105b	TGCGACGAAGAGGAGGGACGGG
	22non114	TGCGACAGGGAGGGAGGACGGG
	22non123	TGCGGGACAAGAGAGGGACGGG
	22non132	TGGCACAGAGGGAGGGAACGGG

2.3. Circular dichroism (CD)

CD experiments were performed with a JASCO J-815 spectropolarimeter equipped with a JASCO CDF 426S Peltier temperature controller, using quartz cells with a 1 cm path length. Reported spectra are a sum of 3 accumulations at 20 °C with a scan speed of 50 nm/min and integration time of 0.5 s in the range of 220 nm to 350 nm. Data were normalized to molar circular-dichroic absorption Δε based on DNA concentrations using $\Delta\epsilon = \theta/(32980 \times c \times l)$ where θ is the CD ellipticity in millidegrees, c is the DNA concentration in mol/L and l is the pathlength in cm (here, $l = 0.2$ cm). Baselines were subtracted using a 100 mM TMAA and 1 mM KCl solution.

2.4. Electrospray mass spectrometry (ESI-MS)

ESI-MS spectra were obtained using a Thermo-Exactive Orbitrap mass spectrometer in the negative ion mode. We used the standard ESI source, and the samples were injected by a syringe pump at 4 μL/min. The full scan mass range was [500–4000]. The Exactive was tuned to "soft" conditions using the bimolecular quadruplex $[G4T4G4]_2$ in 100 mM ammonium acetate [30]. Conditions are considered "soft" when the dominating stoichiometry detected is $[(G4T4G4)_2 + 3(NH_4)\text{-}8H]^{5-}$ at $m/z = 1524.6$. Fig. S2 in the supporting information depicts mass spectra for "soft" conditions, which can be tuned mainly by setting the HCD off, and adjusting MP_0 offset parameter to −8 V. Annotated mass spectra recorded under "soft" conditions are available in the supporting information for all of the oligonucleotides studied in this screening.

2.5. Methodology to infer conformational changes from ESI-MS data

A 10 μM solution of DNA sequence is prepared in a buffer that needs to be electrospray-compatible, have close to physiological ionic strength, and in which canonical 3-quartet G-quadruplexes are folded. We used a buffer solution consisting of 100 mM trimethylammonium acetate (TMAA) to fix the ionic strength of the solution, and 1 mM of KCl to supply the K^+ ions needed to fold the G4s [34]. The KCl concentration is limited because if higher than 1 mM, the mass spectrum is dominated by peaks that correspond to $(KCl)_nCl^-$ clusters. The solution is injected into the mass spectrometer and the ion intensity is measured for each mass-to-charge ratio (m/z). An example of full scan mass spectrum is shown in Supporting Fig. S3, and example zooms on a charge state are shown in Fig. 1 to illustrate the discussion. One can then unambiguously assign stoichiometries to each m/z peak.

To assign the peaks in the mass spectrum, we must first determine the charge state (z) of the ion signal. To do this, we zoom in on the isotopic distribution for the peak of interest, which arises predominantly from the naturally occurring abundance of ^{13}C isotopologues (see inset C in Fig. 1). The charge can be determined from the apparent separation between the isotopologue peaks; for a singly-charged ion, the