Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 196 (2018) 185–195

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

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

Evaluation of the effect of polymorphism on G-quadruplex-ligand interaction by means of spectroscopic and chromatographic techniques

S. Benito [a], A. Ferrer [a], S. Benabou [a], A. Aviñó [b], R. Eritja [b], R. Gargallo [a,*]

[a] Department of Chemical Engineering and Analytical Chemistry, University of Barcelona, Martí i Franquès, 1-11, E-08028 Barcelona, Spain
[b] Institute for Advanced Chemistry of Catalonia (IQAC-CSIC), CIBER-BBN, Jordi Girona 18-26, E-08034 Barcelona, Spain

ARTICLE INFO

Article history:
Received 21 November 2017
Received in revised form 16 January 2018
Accepted 4 February 2018
Available online 06 February 2018

Keywords:
G-quadruplex
Crystal violet
Binding equilibria
Polymorphism
Multivariate analysis

ABSTRACT

Guanine-rich sequences may fold into highly ordered structures known as G-quadruplexes. Apart from the monomeric G-quadruplex, these sequences may form multimeric structures that are not usually considered when studying interaction with ligands. This work studies the interaction of a ligand, crystal violet, with three guanine-rich DNA sequences with the capacity to form multimeric structures. These sequences correspond to short stretches found near the promoter regions of *c-kit* and SMARCA4 genes. Instrumental techniques (circular dichroism, molecular fluorescence, size-exclusion chromatography and electrospray ionization mass spectrometry) and multivariate data analysis were used for this purpose. The polymorphism of G-quadruplexes was characterized prior to the interaction studies. The ligand was shown to interact preferentially with the monomeric G-quadruplex; the binding stoichiometry was 1:1 and the binding constant was in the order of 10^5 M^{-1} for all three sequences. The results highlight the importance of DNA treatment prior to interaction studies.

1. Introduction

Guanine-rich sequences of nucleic acids may fold into complex structures known as G-quadruplexes. The building blocks of such structures are known as G-tetrads, which are formed by the planar arrangement of four guanine bases stabilized by hydrogen bonding (Fig. 1a). The stability of the G-quadruplex depends heavily on the nucleic acid sequence (composition and length of both, guanine tracts and loops) and on external factors such as the nature of the counterion and the concentration, temperature and presence of appropriate ligands [1]. The biological importance of this structure lies in its presence in the promoter regions of several oncogenes and at the extreme ends of telomeres, which are related to cancer and aging [2]. The *in vivo* existence of these structures was demonstrated fairly recently [3]. In addition, the fact that their folding is strongly influenced by the presence of certain counterions, such as potassium, makes it a suitable tool for the development of analytical methods [4,5] and for nanotechnology applications [6,7].

In general, a 20–30-nucleotide long DNA sequence may fold intramolecularly into a monomeric G-quadruplex structure. This may show parallel, anti-parallel or 3 + 1 hybrid topologies (Fig. 1b) [1]. In addition, G-quadruplexes have been shown to form multimeric forms such as dimers and tetramers, depending on the experimental conditions [8–11]. In the presence of 150 mM potassium, sodium or ammonium cations, parallel topologies are favored in G-quadruplex structures that contain short loops. Moreover, these parallel G-quadruplexes may form stable multimeric structures such as dimers or even trimers, even at low strand concentrations. The correlation between multimer formation and parallel G-quadruplex formation strongly suggests that multimeric assemblies are parallel. On the other hand, the G-quadruplex structures tend to be more intramolecular and anti-parallel as loop length increases [12].

The *in vivo* presence of G-quadruplexes is related to their role in certain biological processes such as DNA replication. In this regard, interaction with ligands is a subject that requires extensive research to modulate the stability, and consequently, the functionality of these structures [13–15]. However, to the best of our knowledge, the potential formation of multimeric species other than the monomer is rarely considered in studies on G-quadruplexes and interaction with ligands. Thus, the main objective of this work was to study the interaction of a ligand with several guanine-rich sequences prone to form multimeric structures in order to ascertain whether the presence of multimeric structures should be considered in such research.

Crystal violet (CV), the ligand used in this work, is a fluorescent triphenylmethane dye (Fig. 1c) that has been shown to bind selectively to the G-quadruplex motif over duplex or single-stranded DNA [16]. Research has been carried out on the interaction of CV with the anti-parallel G-quadruplex formed by repeated subunits of the *Oxytricha* telomere [17]. Based on these studies, a binding mode was proposed in which the dye stacks onto the two external tetrads of the G-quadruplex structure. Many sensors have been developed based on the interaction of CV with

* Corresponding author.
E-mail address: raimon_gargallo@ub.edu (R. Gargallo).

https://doi.org/10.1016/j.saa.2018.02.006
1386-1425/

DNAs in order to detect a plethora of analytes, such as potassium ions [18,19], lead ions [20], copper ions [21] and ATP [22], and the activity of proteins, such as 3′ → 5′ exonuclease [23].

The guanine-rich DNA sequence models used in this work are shown in Fig. 1d. A guanine-rich sequence near the promoter region of the *c-kit* gene has previously been studied (ckit21) [24–27]. This sequence shows three tracts of three guanines and one tract of four guanines near the 3′ end. Because of this asymmetric distribution of guanines, there is an equilibrium between two conformers that hinders the structural study of this sequence. To overcome this problem, appropriate G-T mutations in the fourth tract were proposed (ckit21T21 and ckit21T18), and the resulting G-quadruplex structures were shown to be more homogeneous than those formed from the ckit21 sequence. Whereas ckit21T21 formed a parallel topology, ckit21T18 formed mainly parallel G-quadruplex with some antiparallel G-contribution [24]. Later studies on ckit21T21 using NMR spectroscopy led to the proposal that this sequence adopts either a dimeric or a monomeric G-quadruplex structure, at high or low potassium concentrations, respectively [25,26]. SMG03 corresponds to a guanine-rich sequence located between bases −71 and −28 upstream of the promoter region of the SMARCA4 gene, and has not yet been studied. Lastly, T21 was selected as a linear (*i.e.*, non-compact) structure model.

Instrumental techniques such as circular dichroism (CD), molecular absorption and molecular fluorescence spectroscopies, size-exclusion chromatography (SEC) and mass spectrometry, as well as multivariate analysis methods, were used in this work.

2. Material and Methods

2.1. Reagents

The DNA sequences were synthesized on an Applied Biosystems 3400 DNA synthesizer using the 200 nmol scale synthesis cycle. Standard phosphoramidites were used. Ammonia deprotection was performed overnight at 55 °C. The resulting products were purified using a Glen-Pak Purification Cartridge (Glen Research, Sterling, VA, USA). DNA strand concentration was determined by absorbance measurements (260 nm) at 90 °C using the extinction coefficients calculated using the nearest-neighbor method, as implemented on the OligoCalc website [28]. KCl, KH_2PO_4, K_2HPO_4, $NaCH_3COO$, HCl and NaOH (a.r.) were purchased from Panreac (Barcelona, Spain). MilliQ® water was used in all experiments.

2.2. Instruments and Apparatus

Absorbance spectra were recorded on an Agilent 8453 diode array spectrophotometer (Agilent Technologies; Waldbronn, Germany). Temperature was controlled by means of an Agilent 89090A Peltier device (Agilent Technologies). CD spectra were recorded on a Jasco J-810 spectropolarimeter equipped with a JULABO F-25-HD temperature control unit (Seelbach, Alemania). Molecular fluorescence spectra were measured with an AMINCO-Bowman AB2 spectrofluorometer. The temperature was controlled by means of a water bath. Measurements were taken at 570 nm excitation wavelength, with 900 V sensitivity, and 4 nm excitation and emission slides. In all spectroscopic studies, Hellma quartz cells (10 mm path length, and 350, 1500 or 3000 μL volume) were used.

For SEC, the chromatographic system consisted of an Agilent 1100 Series HPLC instrument equipped with a G1311A quaternary pump, a G1379A degasser, a G1392A autosampler, a G1315B photodiode-array detector with a 13-μL flow cell, and an Agilent ChemStation for data acquisition and analysis (Rev. A 10.02), all from Agilent Technologies. Two chromatographic columns were used for separation at room temperature: a BioSep-SEC-S3000 column (300 × 7.8 mm, particle size 5 μm and pore size 290 Å) from Phenomenex (Torrance, CA, USA) and an Acclaim SEC-300 column (300 × 7.8 mm, particle size 5 μm and pore size 300 Å) from Thermo Scientific (Waltham, MA, USA).

Electrospray ionization mass spectrometry (ESI-MS) spectra were recorded with an Agilent G1969A LC/MSD TOF Mass Spectrometer (Agilent Technologies, Santa Clara, CA, USA). The capillary voltage was set to 3.5 kV, the voltage range to 200 or 300 V, the gas temperature to 325 °C, the pressure of nebulizer N_2 gas to 15 psi, and the N_2 flow rate for drying was set to 7.0 $L \cdot min^{-1}$.

2.3. Procedures

The DNA solutions for measurements were prepared by dilution of DNA stock solutions in the appropriate medium. In accordance with the procedure usually applied when studying the conformational equilibria of DNA, all samples for measurement were subjected to a preliminary treatment that consisted of heating the samples at 90 °C for 10 min and slowly cooling them overnight to room temperature (from now on, referred to as the thermal treatment). As described in the Results section, another treatment based on the addition of LiOH 1 M to pH 11.0 and the subsequent addition of HCl 1 M to pH 7 was also carried out (from now on, referred to as the acid-base treatment).

The melting experiments were monitored using the Agilent 8453 spectrophotometer or the J-810 spectropolarimeter equipped with a Peltier unit. The DNA solution was transferred to a covered cell and spectra were recorded at 1 °C intervals with a hold time of 3 min at each temperature, which yielded an average heating rate of approximately 0.3 $°C \cdot min^{-1}$. The medium consisted of 20 mM phosphate buffer (pH 7.0) and 50 mM KCl. Each sample was allowed to equilibrate at the initial temperature for 30 min before the melting experiment began.

For binding studies monitored by CD spectroscopy, small aliquots of stock CV solutions were added to a diluted DNA solution. In the case of molecular fluorescence-monitored titrations, the procedure consisted of the addition of small aliquots of DNA stock solution to CV diluted solutions. In all cases, the experimental conditions were 25 °C, 20 mM phosphate buffer and 50 mM KCl. Spectra were recorded 5 min after each addition of titrant.

For SEC analysis, the mobile phase was 300 mM KCl and 20 mM phosphate buffer (pH 7.0). The flow rate was set to 1.0 mL min^{-1}. The injection volume was 15 μL. Absorbance spectra were recorded between 200 and 500 nm. T_{15}, T_{20}, T_{25}, T_{20} and T_{45} sequences were used as standards to construct the plot of logarithm of the molecular weight *vs.* retention time (t_R). Standards were injected twice to assess the reproducibility of the t_R values, and the relative difference between t_R values for a given standard was lower than 0.5%. SEC profiles were normalized to equal length (Euclidean normalization) to eliminate potential variations in the DNA concentration of samples that could hinder the comparison of chromatograms. Normalization was carried out using Eq. 2 [29]:

$$Normalized\ chromatogram = \frac{raw\ chromatogram}{\sqrt{\sum_1^n d_i^2}} \quad (2)$$

For ESI-MS measurements, the sample (10 μL volume, 150 mM NH_4AcO, 20% CH_3OH) was introduced into the source by means of an Agilent 1100 HPLC pump. The flow rate was set to 200 $\mu L \cdot min^{-1}$ (H_2O:CH_3CN, 1:1).

2.4. Data Analysis

2.4.1. Univariate Analysis of Melting Data

For melting experiments, the absorbance or ellipticity data measured at one single wavelength (univariate data) as a function of temperature were analyzed, as described elsewhere [30]. For the unfolding of intramolecular structures such as those studied here, the chemical equation and corresponding equilibrium constant may be