Clerocidin-mediated DNA footprinting discriminates among different G-quadruplex conformations and detects tetraplex folding in a duplex environment

Biochimica et Biophysica Acta 1830 (2013) 4660–4668

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

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

Clerocidin-mediated DNA footprinting discriminates among different G-quadruplex conformations and detects tetraplex folding in a duplex environment

Matteo Nadai [a], Giovanna Sattin [a], Giorgio Palù [a], Manlio Palumbo [b], Sara N. Richter [a,*]

[a] Department of Molecular Medicine, University of Padua, via Gabelli 63, 35121 Padua, Italy
[b] Department of Pharmaceutical and Pharmacological Sciences, University of Padua, via Marzolo 5, 35131 Padua, Italy

ARTICLE INFO

Article history:
Received 9 January 2013
Received in revised form 10 May 2013
Accepted 27 May 2013
Available online 6 June 2013

Keywords:
G-quadruplex
Clerocidin
Footprinting
Protection assay
Alkylating agent
DNA conformation

ABSTRACT

Background: G-quadruplexes are polymorphic non-canonical nucleic acid conformations involved both in physiological and pathological processes. Given the high degree of folding heterogeneity and comparable conformational stabilities, different G-quadruplex forms can occur simultaneously, hence rendering the use of basic instrumental methods for structure determination, like X-ray diffraction or NMR, hardly useful. Footprinting techniques represent valuable and relatively rapid alternative to characterize DNA folding. The natural diterpenoid clerocidin is an alkylating agent that specifically reacts at single-stranded DNA regions, with different mechanisms depending on the exposed nucleotide.
Methods: Clerocidin was used to footprint G-quadruplex structures formed by telomeric and oncogene promoter sequences (c-myc, bcl-2, c-kit2), and by the thrombin binding aptamer.
Results: The easy modulability of CL reactivity towards DNA bases permitted to discriminate fully and partially protected sites, highlights stretched portions of the G-quadruplex conformation, and discriminate among topologies adopted by one sequence in different environmental conditions. Importantly, CL displayed the unique property to allow detection of G-quadruplex folding within a duplex context.
Conclusions: CL is a finely performing new tool to unveil G-quadruplex arrangements in DNA sequences under genomically relevant conditions.
General significance: Nucleic acid G-quadruplex structures are an emerging research field because of the recent indication of their involvement in a series of key biological functions, in particular in regulation of proliferation-associated gene expression. The use of clerocidin as footprinting agent to identify G-quadruplex structures under genomically relevant conditions may allow detection of new G-quadruplex-based regulatory regions.

1. Introduction

Nucleic acids that can fold in the G-quadruplex conformation are an emerging field of research because of the recent indication of their involvement in a series of key biological functions, such as telomerase-mediated maintenance of telomeres at the chromosome ends in tumor cells [1] and regulation of proliferation-associated gene expression [2–5]. G-quadruplexes are inter- or intramolecular four-stranded structures that may form in G-rich sequences. They are based on the formation of G-quartets, which are stabilized by Hoogsteen-type hydrogen bonds between guanines and by the interaction with monovalent cations located between the tetrads. G-quartets stack on top of each other to give rise to G-quadruplexes. G-quadruplexes can be very polymorphic, and the adopted structures are dependent on several factors, including the base sequence, strand concentration, loop connectivities, and cations present.

DNA strands in G-quadruplex may display antiparallel, parallel, or hybrid (Fig. 1A) orientation, and the nucleotide linkers between G-quartet stacks can adopt a multitude of loop structures. An example of the heterogeneity of G-quadruplex structures is the intramolecular quadruplex formed from human telomere, whose minimal repeated sequence containing 4 tracts of GGG has been solved to date by NMR in K^+ and Na^+ solutions. K^+ is much more abundant than Na^+ in intracellular environments, thus elucidation of the human telomere structure in K^+ solution has attracted intense research, even though the presence of multiple G-quadruplex conformations has required much effort [6]. While the crystal structure of the oligonucleotide d[AGGG(TTAGGG)$_3$] (hTel22 in Table 1) in K^+ has been represented as a parallel propeller type intramolecular G-quadruplex [7] (Fig. 1A), in solution this oligonucleotide was reported to form multiple G-quadruplexes [6], therefore hindering NMR determination of a stable conformation. The conformational heterogeneity was overcome by a variation of the flanking nucleotides and substitution of core guanines with guanine analogs [8]. The oligonucleotides d[AAAGGG(TTAGGG)$_3$AA] and d[TTGGG(TTAGGG)$_3$A] (hTel24 and hTel26, respectively, in Table 1) were found to form a hybrid

* Corresponding author. Tel.: +39 049 827 2346; fax: +39 049 8272355.
E-mail address: sara.richter@unipd.it (S.N. Richter).

0304-4165/$ – see front matter
http://dx.doi.org/10.1016/j.bbagen.2013.05.039

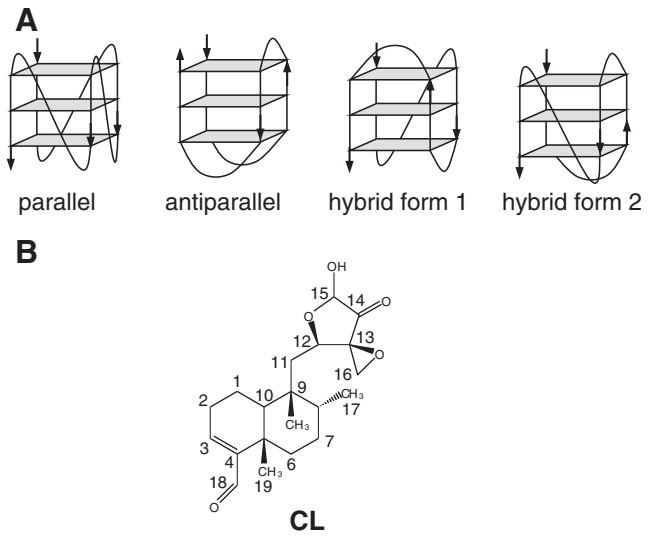

Fig. 1. A) Conformations that can be adopted by a monomolecular G-quadruplex. B) Chemical structure of CL. Numbering of positions is indicated.

structure that has both some parallel and antiparallel strands (hybrid 1 or form 1) [9,10] (Fig. 1A); the d[TTAGGG(TTAGGG)$_3$TT] and d [TAGGG(TTAGGG)$_3$TT] sequences form a similar structure (hybrid 2 or form 2) [11–13], with an identical G-quadruplex core and different connecting loops (Fig. 1A). More recently, an antiparallel, basket type human intramolecular G-quadruplex has also been demonstrated in K^+ solution for the d[(GGGTTA)$_3$GGGT] oligonucleotide; this conformation is more stable than the hybrid forms despite the presence of only two G-quartet layers in the core [14]. Conversely, in the presence of Na^+ ions the oligonucleotide d[AGGG(TTAGGG)$_3$] folds in an antiparallel basket-type quadruplex [15] (Fig. 1A).

It appears evident that the equilibrium between conformations is influenced by the flanking sequences of the oligonucleotide, outside the GGG(TTAGGG)$_3$ core.

Given the described heterogeneity, it is difficult to predict the propensity of a sequence to fold into a particular structure, and each sequence needs to be characterized empirically under different folding conditions. Sophisticated techniques capable of obtaining atomic-resolution structures, such as NMR and X-ray crystallography are available; however, each displays important limitations. X-ray crystallography relies on the formation of crystals that do not mimic the physiological environment [16]. NMR is obtained in solution, but the tested G-rich sequence does not usually exceed 30 nucleotides in length, its concentration needs to rest in the millimolar range, and it must be present in a highly prevalent conformation [17,18]. These restrictions have hindered high-resolution of recently identified large G-quadruplex structures, such as those present in the promoter region of bcl-2, VEGF, HIF-1α and hTERT [19]. In the case of the promoter regions of bcl-2 and hTERT, oligonucleotides shorter than the full-length G-rich sequence have been solved [20–23]: however it would be expected that the presence of the overlooked flanking nucleotides have a great impact on the complete sequence structure, in analogy to what was described for the telomeric region.

Besides circular dichroism, which provides hints on the relative strand orientations in the G-tetrad core [24], biochemical methods have been used to detect bases involved in the G-quadruplex folding. Historically, compounds able to alkylate specific nucleotide moieties have been used to footprint functional groups involved or exposed in the Hoogsteen G-base pairing: KMnO to detect the accessibility to oxidation of the 5,6-double bond in thymines; diethyl pyrocarbonate (DEPC) to assess the presence of exposed N7 in adenines; dimethylsulfate (DMS) to test the availability of N7 of guanines [25]. The latter has overcome the others and it is now the one alkylating agent widely used in the assessment of novel G-quadruplex structures [26,27].

Clerocidin (CL) (Fig. 1B) is a natural product [28] that has been shown to react at single-stranded (ss) DNA regions, with different mechanisms depending on the exposed nucleotide. In particular, CL electrophilic groups (i.e., a strained epoxy ring and an α-ketoaldehyde function in equilibrium with its hemi-acetalic form) target i) the nucleophilic N7 of guanine (G) inducing spontaneous depurination and DNA strand cleavage [29,30]; ii) the NH_2 and N3 of cytosine (C) with formation of a stable condensed ring system, which is degraded to induce DNA cleavage only after hot alkali treatment [31]; iii) the NH_2 and N1 of adenine (A) to generate an adduct that degrades upon alkali but does not result in DNA strand scission [32]. Due to lack of strong nucleophilic sites, thymine (T) does not react with CL. Based on these characteristics, when a DNA subjected to alkylation by CL is loaded on a sequencing gel, differential reactivity at DNA bases is manifested as reported in Table 2. In addition, the diterpenoid portion of CL restricts the accessibility of the epoxide and α-ketoaldehyde reactive groups towards the DNA and so double-stranded DNA is not alkylated [29].

The peculiar reactivity of CL has been exploited here to address structure solving by biochemical methods. We showed that CL was able to unambiguously detect regions of flexibility within a G-quadruplex, therefore distinguishing different quadruplex conformations. Moreover, CL was able to predict quadruplex versus duplex folding in a double-stranded environment. Experiments were clear-cut and easily afforded. Therefore, CL can be considered as a new tool to assess G-quadruplex conformations adding valuable structural information to currently used low- and high-resolution methods.

Table 1
Oligonucleotides used in this study.

Name	Sequence (5′ → 3′)
hTel22	AGGGTTAGGGTTAGGGTTAGGG
hTel24	TTGGGTTAGGGTTAGGGTTAGGGA
hTel26	AAAGGGTTAGGGTTAGGGTTAGGGAA
hTel26 hybrid 2	TTAGGGTTAGGGTTAGGGTTAGGGTT
c-kit2 Pu21	CGGGCGGGCGCGAGGGAGGGG
c-myc Pu27wt	TGGGGAGGGTGGGGAGGGTGGGGAAGG
c-myc Pu27I	TGGGGAGGGTGGIGAGGGTGGGGAAGG
bcl-2 Pu39	AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG
c-myc compl	CCTTCCCCACCCTCCCCACCCTCCCCA
bcl-2 compl	CAGCCCCGCTCCCGCCCCCTTCCTCCCGCGCCCGCCCCT
TBA	GGTTGGTGTGGTTGG

2. Materials and methods

2.1. Clerocidin and oligonucleotides

CL was a gift of Leo Pharmaceutical Products (Ballerup, Denmark). Molar extinction coefficients were experimentally determined to be 11,818 M^{-1} cm^{-1} for CL. Working drug solutions were obtained by diluting fresh stocks in the appropriate buffer. T4 polynucleotide kinase was purchased from Fermentas (Thermo Scientific, Italy) (Burlington, Canada). [γ−^{32}P]ATP was from PerkinElmer (Monza, Italy), while all oligonucleotides were from Sigma-Aldrich (Milan, Italy).

Oligonucleotides used in this study are reported in Table 1.

2.2. Reaction of CL towards G-quadruplex folded oligonucleotides

All oligonucleotides were gel-purified before use and prepared in desalted/lyophilized form. Oligonucleotides were 5′-end-labeled with [γ−^{32}P]ATP by T4 polynucleotide kinase and purified by MicroSpin G-25 columns (GE Healthcare, Europe). They were next resuspended in lithium cacodylate 10 mM, pH 7.4, with or without KCl or NaCl 100 mM, heat-denatured and folded. Drug reactions with the labeled G-quadruplex folded or unfolded oligonucleotides (4 pmol/sample) were performed at 37 °C in the resuspension buffer. These conditions were selected to maintain the stability of the target structure and to

Table 2
Effects of CL alkylation at DNA bases as manifested on a 20% polyacrylamide sequencing gel [30].

DNA base	Piperidine	
	−	+
G	- One band migrating slower than the full-length oligo = *alkylation* - One band migrating slower than the corresponding G base obtained in the marker lane = *alkylation + cleavage, with no loss of the adduct*	Cleavage band which migrated as the corresponding base obtained in the marker lane = *alkylation + cleavage and loss of the adduct*
C	One band running slower than the full-length oligo = *alkylation*	Cleavage band which migrated as the corresponding base obtained in the marker lane = *alkylation + cleavage and loss of the adduct*
A	One band running slower than the full-length oligo = *alkylation*	No cleavage band
T	No alkylating bands	No cleavage band

minimize the possible competition of buffer molecules for drug alkylation [33]. At the indicated time intervals, samples were precipitated with ethanol to eliminate non-reacted drug, then resuspended and either kept on ice, or treated at 90 °C for 30 min with 1 M piperidine to complete strand scission according to the Maxam and Gilbert protocol. Samples were then lyophilized, resuspended in formamide gel loading buffer, and heated at 95 °C for 2 min. Reaction products were analyzed on 20% denaturing polyacrylamide gels and visualized by Typhoon FLA 9000 phosphorimaging analysis (GE Healthcare). Quantification was performed by ImageQuant software (GE Healthcare).

2.3. DMS footprinting

Each labeled and annealed DNA was treated with DMS (0.5% in ethanol) for 5 min. Reactions were stopped by adding gel loading buffer containing 10% glycerol and β-mercaptoethanol. Samples were loaded on 16% native polyacrylamide gel and run until the desired resolution was obtained. DNA bands were localized via autoradiography, excised and eluted overnight. The supernatants were recovered, ethanol-precipitated and treated with 1 M piperidine at 90 °C for 30 min. Samples were dried in a speed vac, washed with water, dried again and resuspended in a formamide gel loading buffer.

2.4. Circular dichroism thermal unfolding

The c-myc Pu27wt and bcl-2 Pu39 oligonucleotides were diluted from stock to the final concentration (4 μM) in lithium cacodylate buffer (10 mM, pH 7.4) and KCl 50 mM. Samples were annealed by heating at 95 °C for 5 min, gradually cooled to room temperature and measured after 24 h. CD spectra were recorded on a Jasco-810 spectropolarimeter (JASCO, Inc., Easton, MD, USA) equipped with a Peltier temperature controller using a quartz cell of 5-mm optical path length and an instrument scanning speed of 100 nm/min with a response time of 4 s over a wavelength range of 230–320 nm. The reported spectrum of each sample represents the average of 2 scans at 20 °C and it is baseline-corrected for signal contributions due to the buffer. Observed ellipticities were converted to mean residue ellipticity $(\theta) = \text{deg} \times \text{cm}^2 \times \text{dmol}^{-1}$ (mol. ellip.). For the determination of T_m, spectra were recorded over a temperature range of 20–95 °C, with a temperature increase of 5 °C. T_m values were calculated according to the van't Hoff equation, applied for a two state transition from a folded to an unfolded state, assuming that the heat capacities of the folded and unfolded states are equal [34].

2.5. Three-dimension view of G-quadruplex conformations

The TBA 3D-structure (PDB 148D) was visualized with the Discovery Studio 3.5 Visualizer software (Accelrys Software Inc., CA, USA).

2.6. Band shift assay

Labeled ss and duplex substrates, obtained by annealing the labeled G-quadruplex forming oligonucleotide with small excess (1.1 folds) of the complementary cold oligonucleotide, were loaded on 16% non-denaturing polyacrylamide gels, run for ~16 h at 50 volts. Gels were visualized by phosphorimaging.

3. Results

3.1. CL-mediated cleavage pattern identifies different G-quadruplex conformations

To assess whether CL was able to discriminate between the reported G-quadruplex folding of the human telomeric repeat in K^+ and Na^+ solutions, the oligonucleotide hTel22 (Table 1) was denatured and refolded in the presence of either K^+ or Na^+ and subjected to a reaction with CL. After the reaction, samples were either immediately precipitated, or treated with hot piperidine to induce strand cleavage at susceptible alkylation sites, before precipitation. Samples were next loaded on a 20% sequencing gel to monitor the presence of CL:DNA adducts and cleavage sites according to CL peculiar reactivity (Table 2) [29–32].

Alkylation of the hTel22 oligonucleotide in both folding conditions was manifested as a ladder of bands migrating slower than the full length oligonucleotide (symbol §, lanes 3, 5, 7 and 9, Fig. 2A). In the presence of K^+, the reaction at G4, G10, G16 and G22 was manifested as cleavage bands at the relevant bases, both before (symbol ¤, lanes 3, 5, 7, 9 K^+, Fig. 2A) and after piperidine treatment (symbol *, lanes 4, 6, 8, 10 K^+, Fig. 2A). In Na^+ solution, cleavage was observed again at G4 and G10 (symbols ¤ and *, lanes 7–10 Na^+, Fig. 2A); in addition G14 was also exposed to the reaction. In contrast, CL+ piperidine cleaved all Gs of a scrambled hTel22 oligonucleotide and of hTel22 folded in the absence of K^+, confirming appropriate folding of the hTel sequence in the presence of K^+ (Supplementary Fig. 1).

While the fully parallel G-quadruplex topology of hTel22 in Na^+ has been solved by NMR (PDB code 143D) [15], its conformational heterogeneity in K^+ has prevented high resolution structural definition [6]. Slightly longer analogues (hTel24 and hTel26) [9,10] have been reported to be present in K^+ solution in a highly prevalent structure (>95%), allowing their characterization as hybrid forms 1, with three parallel and one anti-parallel strands (PDB codes 2GKU and 2HY9).

To test whether hTel adopts a conformation similar to that of hTel24 and hTel26 and to assess if CL displays a similar reactivity towards oligonucleotides that fold in the same G-quadruplex type, the molecule was reacted with hTel22, hTel24 and hTel26. As shown in Fig. 2B, after treatment with piperidine the last G base of each G repeat in the 5′′ → 3′ direction (corresponding to positions G10, G16 and G22 in hTel22) was more exposed to CL, as attested by the more intense cleavage bands (symbols *, lane 4, Fig. 2B). A quantification of cleavage at G15 and G16 in samples treated with CL and piperidine, minus cleavage obtained in the control treated with piperidine is provided in Supplementary Fig. 2. Loading of 2× amount of oligonucleotides did not affect results. This behavior was consistent among the three tested oligonucleotides, indicating that hTel adopts a hybrid structure analogously to hTel24 and hTel26 and that CL reacts alike towards similar G-quadruplex structures.

The reactivity of CL against hTel22 in the presence of K^+ or Na^+ was next compared to that of DMS. As shown in Fig. 2C, in the presence of Na^+, DMS reacted with G10 and G14, analogously to what was

observed with CL (lanes 4 and 6, Fig. 2C, and lanes 3–10, Na^+, Fig. 2A). In the presence of K^+, similarly to CL, DMS revealed exposures of G10, G16 and G20; however, differently from CL, it reacted also with G8, G14 and not with G4. In addition, all As (A7, A13, A19) were also targeted by DMS (compare lanes 3 and 5, Fig. 2C, and lanes 3–10 K^+, Fig. 2A).

As emerged by both CL and DMS footprinting of the hTel sequence, bases that were involved in G-quartet formation were nonetheless exposed to alkylation by chemicals. It is conceivable that strained Gs do not adopt a planar tetrad, therefore exposing reactive groups to the solvent, as previously found for a G-quadruplex forming oligonucleotide [3]. In our case, in the presence of Na^+, Gs that connect the diagonal loop in the hTel oligonucleotide partially exposed their N7 function to the solvent (scheme in Fig. 2A, bottom). In the presence of K^+, all Gs with the 3′-hydroxyl termini close to the loop region were subjected to alkylation, indicating a higher flexibility of these bases in the hybrid model (scheme in Fig. 2A, bottom). Based on these strained regions, CL- and DMS-footprinting methods were able to discriminate between the two conformations adopted by hTel22 in K^+ and Na^+ salt solutions.

Hybrid form 2 was also footprinted with CL; however, cleavage was similar to that obtained in hybrid form 1 (data not shown), likely due to the incomplete prevalence (75%) of hybrid form 2 over hybrid form 1. Three non-telomeric G-quadruplex forming sequences, whose structures had been solved by NMR, were next employed to confirm the ability of CL to correctly identify G-quadruplex conformations.

The nuclease hypersensitivity region of the promoter segment of the human *c-kit* gene contains two guanine-rich sequences, designated *c-kit1* and *c-kit2*, that form G-quadruplex structures. Both sequences, solved by NMR, have been shown to adopt a parallel conformation in K^+ solution [35,36]. In particular, the 21-nt *c-kit2* sequence assumed a

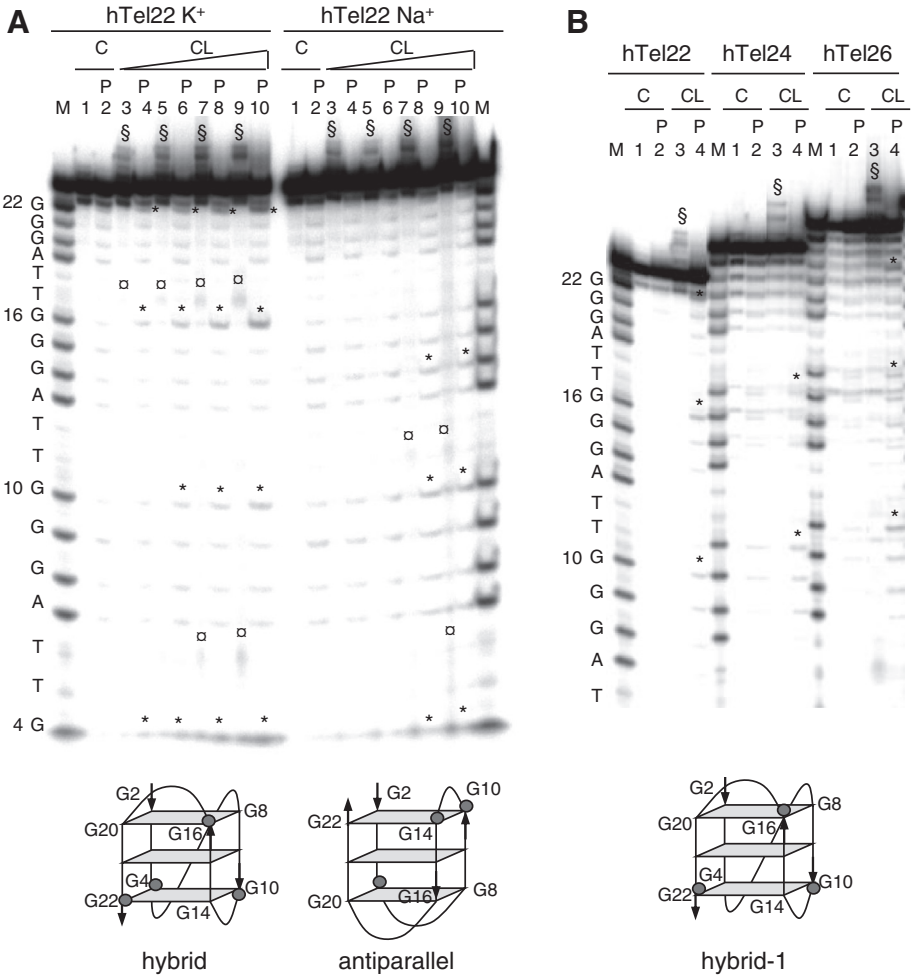

Fig. 2. CL footprinting of human telomeric sequences. A) CL footprinting of hTel22 in K^+ and Na^+ solution. B) CL footprinting of hTel22, hTel24 and hTel26 in K^+ solution. The oligonucleotides indicated above each set of lanes were heat denaturated and folded in the presence of 100 mM of K^+ or Na^+, as indicated. The folded hTel oligonucleotides were incubated with increasing concentrations [25–200 μM (A) or 100 μM (B)] of CL for 24 h at 37 °C. After reaction, samples were precipitated and either kept on ice (samples indicated by odd numbers) or treated with hot piperidine and lyophilized (samples indicated by even numbers and by the symbol P) and loaded on a 20% denaturing polyacrylamide gel. C indicates samples that were either kept on ice or treated with piperidine, but not with CL. The symbol § indicates CL/full-length DNA adducts which migrate slower than the full-length DNA. The symbol ¤ indicates bands that correspond to the oligonucleotide alkylated and cleaved by CL. CL is still bound to the cleaved oligonucleotide, thus the cleavage band runs slower than the corresponding band in the Maxam and Gilbert marker lane (M lanes). The symbol * indicates bands that correspond to the oligonucleotide alkylated and cleaved by CL, with loss of CL. Position of alkylation is evinced by comparison of cleavage bands after piperidine treatment and the marker lane. The hTel22 nucleotide sequence is indicated on the left of the marker lane. Base numbering has been assigned in the 5′ → 3′ direction. Relevant bases exposed to CL-mediated cleavage are indicated by numbers. Schematic models of the most abundant conformation adopted by each hTel sequence, as solved by NMR analysis, are provided below the set of lanes corresponding to each oligonucleotide in the gels. Gray circles indicate G bases that were cleaved by CL. C) Comparison of DMS and CL footprinting of hTel in K^+ and Na^+ solutions. HTel was heat denaturated and folded in the presence of 100 mM of K^+ or Na^+, as indicated. The folded hTel oligonucleotide was treated with CL or DMS according to the description in the Materials and methods section. After purification, samples were precipitated and treated with hot piperidine, lyophilized and loaded on a 20% denaturing polyacrylamide gel. C indicates samples that were treated with piperidine but not with CL. Position of alkylation is evinced by comparison with the Maxam and Gilbert marker lane. The nucleotide sequence is indicated on the left of the marker lane. Base numbering has been assigned in the 5′ → 3′ direction. Relevant bases exposed to DMS- and CL-mediated cleavage are indicated by numbers. Schematic models of the conformations adopted by the hTel sequence in K^+ and Na^+ are provided below the gel image. Gray circles indicate G bases that were cleaved only by CL. White circles indicate G bases that were cleaved only by DMS. Gray and white circles indicate G bases that were cleaved by both CL and DMS.

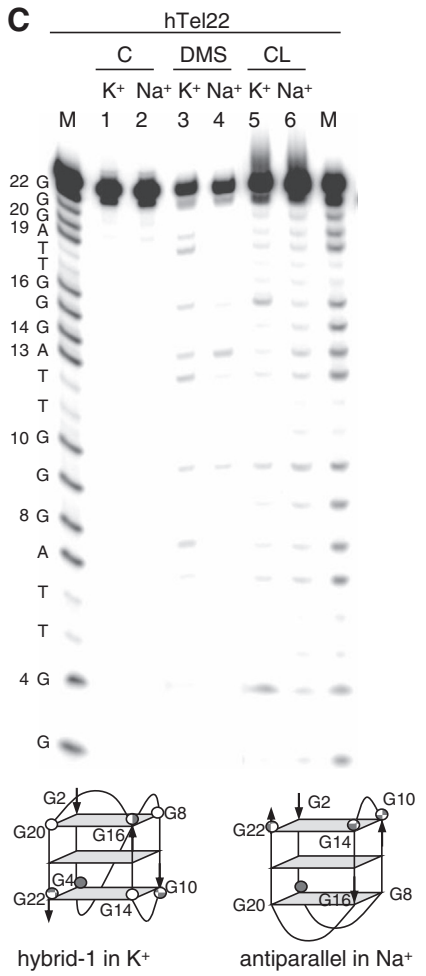

Fig. 2 (*continued*).

standard parallel topology [37]. The c-kit2 Pu21 sequence was folded in the absence or presence of K^+ and treated with CL to confirm its ability to recognize a G-quadruplex conformation. DMS footprinting was performed for comparison. By matching up CL-mediated cleavage of the folded sample (K^+) with the unfolded one (no K^+), it was possible to clearly estimate 4 regions that were protected from CL-mediated cleavage in the folded conformation. In particular, the bands corresponding to G3–G4, G6–G8, G14–G16 and G18–G20 were less intense with respect to the control sample (compare lanes 6 and 5, Fig. 3A). This is in prefect agreement with the proposed structure (Fig. 3A, scheme at the bottom). To note that, in the case of CL-mediated cleavage, non-protected bases displayed band intensity identical to the corresponding bands in the control sample. In contrast, DMS-mediated cleavage displayed protection of fewer bases, namely G7, G14–G15 and G19–G20 (lanes 3 and 4, Fig. 3A). In this case, G that were not involved in G-quartets, were cleaved to a much higher extent with respect to control lanes, so that a comparison had to be performed both among bands within the same lane and among bands in the sample and control lanes. CL and DMS were also used to characterize a non-human G-quadruplex forming sequence, the CD and UV spectra of which indicated a clear G-quadruplex folding (unpublished results): both CL and DMS recognized four G-tracts that were clearly protected from alkylation (Supplementary Fig. 3). Thus, CL-mediated cleavage was at least as informative and easy to be interpreted as cleavage with DMS.

The nuclease hypersensitivity element III_1 (NHE III_1) is the most important control region in the *c-myc* promoter. The 27-nt purine-rich strand of this element, which contains six G-tracts, forms alternate G-quadruplex structures in equilibrium (oligo c-myc Pu27wt, Table 1). Deletion of the 5′ G-tract allowed NMR resolution of several shortened oligonucleotides [38–40], which, based on the experience with the hTel sequence, are expected to be conformationally different from the full length sequence.

When CL was reacted with c-myc Pu27wt, three non-protected regions clearly emerged: positions G5, G13–G14, G23–G26 (lines 3–6 c-myc Pu27wt, Fig. 3B). In particular, the high reactivity of G13 and G14 indicated that these two bases were particularly exposed, probably stretched to connect other G bases involved in G-quartets. Among the solved NMR conformations of *c-myc*, the most recent one, Pu24I (PDB code 2A5P) [41], showed that G13–G14 were not part of G-quartets but indeed were involved in connection loops. Therefore, in our conditions, c-myc Pu27wt adopted a conformation similar to that solved for the shorter and mutated Pu24I oligonucleotide. Mutation of G13 to I13 had been introduced to increase the presence of a single conformation to allow NMR analysis [41]; to check the influence of this mutation on the overall G-quadruplex structure, Pu27I was assayed with CL. We found that in this case G11 and G12 were also exposed to the reaction, I13 and G14 were less exposed, and G5 and G23–26 were not affected by the mutation (lines 3–6 Pu27I, Fig. 3B). This data indicate that the Pu27 G-quadruplex topology is slightly modified by the G13 → I13 mutation.

The thrombin binding aptamer (TBA) is a well characterized chair-like, antiparallel quadruplex structure (PDB code 148D) that binds specifically to thrombin showing interesting anticoagulant properties [42]. TBA displays one G base (G8) in a chair-type loop (Table 1 and the G-quadruplex model in Fig. 3C, below the gel image), which, based on these structural characteristics, was expected to promptly react with CL. The TBA oligonucleotide was denatured and folded in the absence or presence of 100 mM K^+, to assess CL reactivity towards the unfolded (no K^+) and G-quadruplex folded (K^+) DNA. Increasing amounts of CL induced increasing cleavage at all Gs in the unfolded oligonucleotide (Fig. 3C, left side). However, contrary to what was expected, no G base was susceptible to CL-induced cleavage in the folded oligonucleotide (Fig. 3C, right side). Considering the NMR structure reported for TBA (PBD 148D) [43], we observed that G8 was stacked on the purine ring of G6, therefore burying the G8–N7 nucleophilic moiety inside the G-quadruplex structure (Fig. 3C, bottom). In addition, G8–N7 was closely positioned to G6–N9: the two atoms may thus be hydrogen-bonded through a water molecule, which would reduce the overall negative charge on the two purine nitrogens, and consequently lower their nucleophilicity. Hence, CL was likely hindered in its possibility to react with G8–N7 by both steric and electronic factors.

CL was also reacted with the sequence of the promoter region of the oncogene *bcl-2*, which contains six guanine-tracts with three or more contiguous guanines (Table 1). This G-rich 39-bp long sequence had been reported to be present as a mixture of at least three G-quadruplex conformers in K^+ solution [44], and as such, only shorter and mutated forms of the wild-type sequence have been solved by NMR [20,21].

The wild-type 39-nt-long *bcl-2* sequence was subjected to the reaction with CL. As shown in Fig. 3D, a complex pattern of bases exposed to CL (darker bands) and bases excluded from the reaction with CL (no or faint bands) were found (compare lanes 3–10 to lanes 1–2, Fig. 3D). Four regions were effectively protected, i.e., G4–G5, G7–G9, G22–G24, G33–G35, while two large regions, C10–A20 and G25–G31, were exposed to cleavage. In particular, a strong band was obtained at G31, indicating a significant exposure of the N7 group of this base, probably due to local strand stretching. These results are in part consistent with those obtained with DMS footprinting [44]; however, the C10–C12 region appeared less protected than the corresponding region treated with DMS. This apparent discrepancy is likely due to the ability of CL to cleave both at G and C bases.

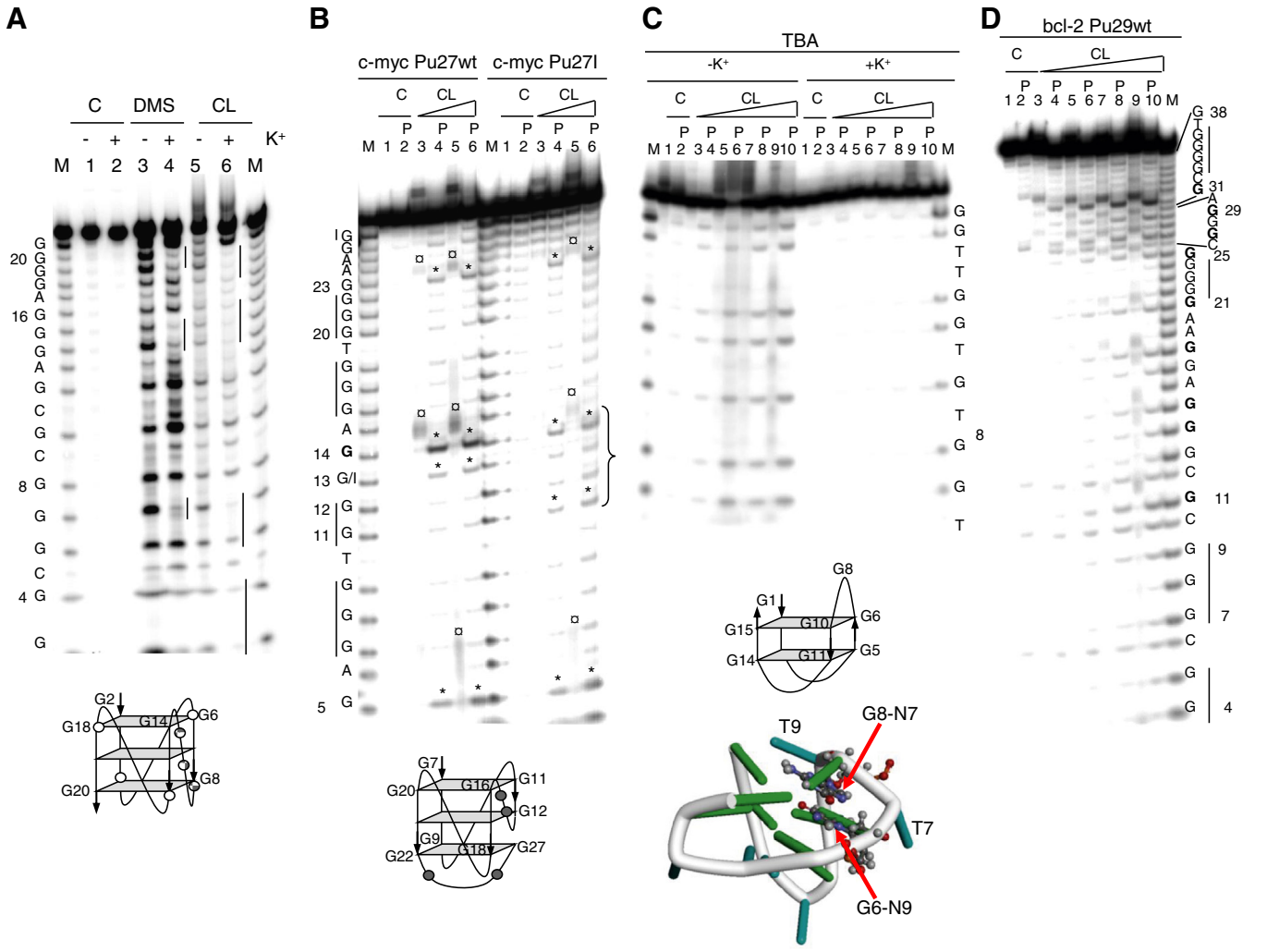

Fig. 3. CL footprinting of non-telomeric G-quadruplex forming sequences. A) DMS and CL-mediated footprinting of the c-kit2 Pu21 sequence. The oligonucleotide was folded in the absence or presence of 100 mM of K^+, as indicated. The folded oligonucleotides were incubated with DMS or CL, according to the Materials and methods section. After purification, samples were treated with hot piperidine, lyophilized and loaded on a 20% denaturing polyacrylamide gel. C indicates samples that were treated with piperidine but not with CL. Position of alkylation is evinced by comparison with the Maxam and Gilbert marker lane. The nucleotide sequence is indicated on the left of the marker lane. Base numbering has been assigned in the 5′ → 3′ direction. Bases not exposed to CL-mediated cleavage are highlighted by vertical lines. A schematic model of the conformation adopted by the c-kit2 Pu21 sequence is provided below the gel image. White circles indicate G bases that were cleaved only by DMS. Gray and white circles indicate G bases that were cleaved by both CL and DMS. B) CL footprinting of c-myc Pu27wt and c-myc Pu27I. The oligonucleotides indicated above each set of lanes were folded in the presence of 100 mM of K^+. The folded oligonucleotides were incubated with increasing concentrations (25–200 μM) of CL for 24 h at 37 °C. After reaction, samples were precipitated and either kept on ice (samples indicated by odd numbers) or treated with hot piperidine and lyophilized (samples indicated by even numbers and by the symbol P) and loaded on a 20% denaturing polyacrylamide gel. C indicates samples that were not treated with CL and were either kept on ice or treated with piperidine. The symbol ¤ indicates bands that correspond to the oligonucleotide alkylated and cleaved by CL. Because CL is still bound to the cleaved oligonucleotide, the cleavage band runs slower than the corresponding band in the Maxam and Gilbert marker lane (M lanes). The symbol * indicates bands that correspond to the oligonucleotide alkylated and cleaved by CL, with loss of CL. Position of alkylation is evinced by comparison of cleavage bands after piperidine treatment and the Maxam and Gilbert marker lane. The nucleotide sequence is indicated on the left of the marker lane. Base numbering has been assigned in the 5′ → 3′ direction. Bases not exposed to CL-mediated cleavage are highlighted by a vertical line. The brace on the right of the gel indicates the region that is differently exposed to CL in the two oligonucleotides. C) CL footprinting of TBA. On the left, the oligonucleotide was folded in the absence of K^+, on the right, in the presence of K^+. The folded oligonucleotides were incubated with increasing concentrations (25–200 μM) of CL. C indicates samples that were not treated with CL and were either kept on ice or treated with piperidine. Below the gel image, a schematic model and a 3D view of the NMR-determined TBA quadruplex conformation are shown. The buried position of the G8–N7 moiety is indicated and its vicinity to G6–N9 is highlighted. D) CL footprinting of bcl-2 Pu39. The folded oligonucleotide was incubated with increasing concentrations (25–200 μM) of CL. After reaction, samples were precipitated and either kept on ice (samples indicated by odd numbers) or treated with hot piperidine and lyophilized (samples indicated by even numbers and by the symbol P) and loaded on a 20% denaturing polyacrylamide gel. C indicates samples that were not treated with CL and were either kept on ice or treated with piperidine. Bases not exposed to CL-mediated cleavage are highlighted by vertical lines.

3.2. CL discriminates between G-quadruplex and double-stranded nucleic acid conformations

It has been reported that CL is inert towards fully double stranded DNA [29]. We argued that this feature could be exploited to discriminate between duplex and quadruplex structures. To test this hypothesis, the c-myc Pu27wt and bcl-2 Pu39 oligonucleotides were mixed with their complementary sequences, denatured and renatured in the presence of 100 mM K^+. Folded oligonucleotides were next subjected to CL reaction. As shown in Fig. 4A, CL-mediated cleavage pattern in the presence of the c-myc Pu27wt complementary sequence was almost identical to that obtained with the corresponding single-stranded G-quadruplex folded oligonucleotide. The only major difference was an exposure of the 3′-end nts in the two-strand sample which was not present in the one-strand sample. In contrast, in the case of the bcl-2 Pu39 (Fig. 4B), no cleavage was obtained in the two-strand sample, indicating a fully annealed structure. These results indicate that, in the presence of the complementary strand, c-myc Pu27wt was mostly present in a G-quadruplex conformation, while bcl-2 Pu39 was fully double-stranded. To evaluate the bio-physical reasons of the different folding

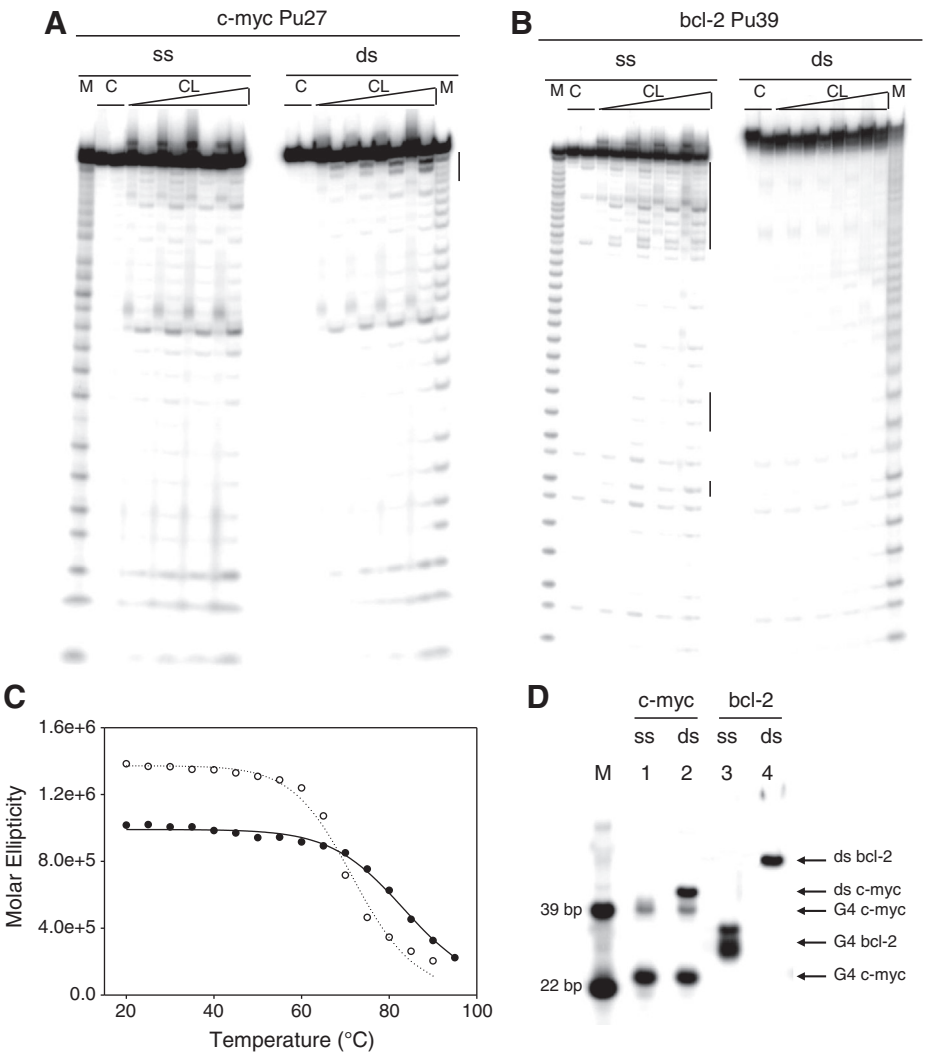

Fig. 4. Comparison of CL footprinting on G-quadruplex and duplex DNA. The c-myc Pu27wt (A) and bcl-2 Pu39 (B) oligonucleotides were denatured and refolded in 100 mM K^+ either alone (ss samples) or in the presence of their complementary oligonucleotides (ds samples). Samples were subjected to incubation with increasing amounts of CL (25–200 μM) and treated or not with piperidine. In A) the vertical line indicates the only region where bases are exposed to CL in the duplex DNA and not in the quadruplex DNA. In B) the vertical lines highlight regions that are exposed to CL in the quadruplex DNA and are protected in the duplex DNA. C) Circular-dichroism-detected thermal unfolding of c-myc Pu27 (black circles) and bcl-2 Pu39 (white circles). The spectra of the two oligonucleotides were measured from 20 °C to 95 °C, with increasing temperature steps of 5 °C. Molar ellipticity values obtained at the most intense CD peak were plotted against temperature. The van't Hoff equation was applied to calculate T_m. The theoretical values (solid lines) were fitted to the experimental points. D) Band shift assay of c-myc Pu27 and bcl-2 Pu39 in the presence/absence of an excess (1.1 folds) of their complementary strands. Lanes 1 and 3 represent the G-quadruplex folded c-myc Pu27 and bcl-2 Pu39, respectively. Lanes 2 and 4 represent c-myc Pu27 and bcl-2 Pu39 in the presence of their complementary strands. M is a marker lane containing 39- and 22-bp oligonucleotides.

behavior observed for the c-myc Pu27wt and bcl-2 Pu39 sequences, the T_m values of the two G-quadruplex structures were calculated by circular-dichroism (CD)-monitored thermal unfolding experiments (Fig. 4C). The resulting T_m values, calculated according to the van't Hoff equation, were 84.0 ± 0.4 °C and 72.0 ± 0.7 °C for c-myc Pu27wt and bcl-2 Pu39, respectively. In contrast, theoretical T_m values for the duplex structures, in the presence of the appropriate salt concentration, obtained by the IDT calculator (http://eu.idtdna.com), were 71 °C and 82 °C, respectively. Therefore, the G-quadruplex conformation of c-myc Pu27wt is more stable than its double-stranded counterpart, while the opposite holds true for the bcl-2 Pu39 sequence. In addition, c-myc Pu27 and bcl-2 Pu39 were assayed in the absence or presence of their complementary strand in a gel shift assay (Fig. 4C). As shown, c-myc Pu27 separated in one predominant fast-migrating band, likely the intramolecular G-quadruplex conformation, and one minor band which, based on the apparent MW, was attributed to a bimolecular G-quadruplex structure. When a molar excess of the complementary strand was added, a duplex conformation formed; however, the ss G-quadruplex structures were still largely present. Bcl-2 Pu39 was also present as two G-quadruplex folded bands; however, when the complementary strand was added, the folded ss DNA disappeared and only the ds conformation was observed. These data are thus in perfect agreement with the cleavage pattern observed in Fig. 4A and B, indicating that CL efficiently discriminates between G-quadruplex and duplex topologies.

4. Discussion and conclusions

NMR methods have proven to be valuable to dissect G-quadruplex conformations in solutions. However, not all sequences can be solved by NMR. In particular, a successful analysis is seldom obtained with oligonucleotides longer than 30 nt, concentrations must be in the millimolar range, therefore far from physiological conditions, and just one conformation must be present in large excess with respect to others [17,18]. These conditions are rarely fulfilled, especially with the discovery of the long G-tracts present in many oncogene promoters, such as bcl-2, h-TERT, VEGF, HIF-1α, to name a few [19]. Stabilization of one conformation is usually obtained deleting terminal parts and introducing mutations in the target sequence. However, the human telomeric

sequence has taught a harsh lesson: small modifications such as addition of a single terminal nucleotide may have great impact on the overall G-quadruplex conformation [6,8–14].

Therefore, in the presence of tough sequences, classic biological methods, such as chemical footprinting, may still help to spread light on the regions principally involved in loop and quartet formation.

We have shown here that the natural compound CL was able to efficiently discriminate among different G-quadruplex conformations. By using sequences whose conformations had been solved by NMR, we proved that CL-mediated cleavage pattern reflects the actual G-quadruplex structure. In addition, when the same G bases were involved in the quadruplex structure, such as in the case of hTel22, it was still possible to discriminate the different conformations adopted in two salt solutions based on the biased stretching of the G bases.

DMS is the chemical normally used to detect G-quadruplex topologies. We have shown that also CL provided accurate information in determining a G-quadruplex folding (see c-kit2 Pu21, and bcl-2 Pu39 compared to [44]). To note that the intrinsic lower reactivity of CL with respect to DMS confers to CL a higher degree of modulability. DMS reaction times are in the range of minutes, while CL's are hours which make it easier to adjust the footprinting protocol. Moreover, after the methylation step, DMS has to be removed by gel electrophoresis before treatment with piperidine to avoid unspecific reactions. In contrast, a purification step based on DNA precipitation is sufficient with CL, making the footprinting protocol simpler and less time-consuming. In addition, at 25 °C CL is a stable non-volatile solid and in aqueous solution hydrolysis rates are slow and products are non-toxic, whereas DMS is an unstable, hygroscopic and volatile liquid which degrades fast in solution generating toxic hydrolysis products (Table 3). Therefore CL is easier to handle and less toxic than DMS. However, it should be noted that the long reaction times required for CL would not allow detection of multiple conformations in a slow dynamic equilibrium, resulting in a protection pattern averaged over the various conformations. In any event, should CL footprinting suggest the presence of dynamically interconverting G-quadruplexes, a DMS assay may follow to assess the single conformation.

We have also shown that CL was able to discriminate between G-quadruplex and duplex structures, proving that, in the presence of the complementary strand, c-myc Pu27wt maintains a G-quadruplex conformation, while bcl-2 Pu39 adopts a fully duplex structure. In line with our results, based on the unfolding and folding rates of the *c-myc* G-quadruplex, it had been demonstrated that in the presence of duplex DNA at low, equivalent molar strand concentrations, duplex conformation was significantly less favorable than G-quadruplex formation, even at neutral pH [45]. This discriminatory power is unique to CL, because DMS is equally reactive towards single and double-stranded DNA and, as such, it is still used in the Maxam and Gilbert chemical sequencing [46]. Detecting a quadruplex conformation within a double-stranded environment could be of valuable importance to assess the DNA folding state in physiological conditions. For example, at the cellular level, CL might be useful to detect the topological state of oncogene promoters, some of which have been reported to exist in alternative duplex and quadruplex conformations depending on the exerted promotorial activity [4,47].

In this connection, additional useful information could be obtained by examining CL footprinting data prior or after piperidine treatment. In fact, this latter would evidentiate protection both at Gs and Cs and allow a structural assessment also in C-rich regions.

A concern might arise considering the large size of CL compared to DMS, which could locally affect the conformation of the G-quadruplexes. Given the shape and the electronic distribution of the molecule, non-covalent binding of the diterpenoid to DNA can hardly be expected. The bulky moiety essentially appears to modulate access of the reactive epoxide to nucleophilic sites. This is confirmed by the remarkable matching of footprinting location with the structural information obtained by NMR or X-ray methods. To note that these latter methods are always necessary to confirm the initial structural assessment obtained by footprinting techniques.

Since CL is a natural product isolated from a fungus, its availability may be limited by the necessity of purification from the microorganism. However, a complete synthesis of CL has been reported [48] and a structural analog exhibiting a naphthalene group in place of the diterpenoid moiety has been shown to possess identical base selectivity and reactivity per se towards the DNA [30,32,33].

Moreover, CL appears to be superior to DMS as a footprinting agent for a number of reasons as summarized in Table 3.

In conclusion, our results indicate that CL is an effective and versatile tool to discriminate among different DNA G-quadruplex conformations and selectively recognize tetraplex arrangements in a duplex environment.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2013.05.039.

Table 3
Comparison between selected physical/chemical properties of CL and DMS.

	CL	DMS
Physical state at 25 °C	Solid	Liquid
Stability	Good	Poor
Volatility	None	High
Chirality	Present	Absent
Reactivity vs. G	Yes	Yes
Reactivity vs. C, A	Yes	No
Reactivity vs. ds-DNA	No	Yes
Reaction times	Hours	Minutes
Hydrolysis rate	Slow	Fast
Toxicity	Low	High
Hydrolysis products	Non-toxic	Toxic
Handling simplicity	High	Low

Acknowledgements

This work was supported by the Italian Ministry of University and Research [FIRB-Ideas RBID082ATK_001], and by the University of Padua.

References

[1] M. Folini, L. Venturini, G. Cimino-Reale, N. Zaffaroni, Telomeres as targets for anticancer therapies, Expert Opin. Ther. Targets 15 (2011) 579–593.
[2] S. Balasubramanian, L.H. Hurley, S. Neidle, Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? Nat. Rev. Drug Discov. 10 (2011) 261–275.
[3] A. Siddiqui-Jain, C.L. Grand, D.J. Bearss, L.H. Hurley, Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 11593–11598.
[4] R.K. Thakur, P. Kumar, K. Halder, A. Verma, A. Kar, J.L. Parent, R. Basundra, A. Kumar, S. Chowdhury, Metastases suppressor NM23–H2 interaction with G-quadruplex DNA within c-MYC promoter nuclease hypersensitive element induces c-MYC expression, Nucleic Acids Res. 37 (2009) 172–183.
[5] P. Rawal, V.B. Kummarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S.K. Das, S. Chowdhury, Genome-wide prediction of G4 DNA as regulatory motifs: role in *Escherichia coli* global regulation, Genome Res. 16 (2006) 644–655.
[6] A.T. Phan, D.J. Patel, Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics, J. Am. Chem. Soc. 125 (2003) 15021–15027.
[7] G.N. Parkinson, M.P. Lee, S. Neidle, Crystal structure of parallel quadruplexes from human telomeric DNA, Nature 417 (2002) 876–880.
[8] Y. Xu, Y. Noguchi, H. Sugiyama, The new models of the human telomere d [AGGG(TTAGGG)3] in K^+ solution, Bioorg. Med. Chem. 14 (2006) 5584–5591.
[9] A. Ambrus, D. Chen, J. Dai, T. Bialis, R.A. Jones, D. Yang, Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution, Nucleic Acids Res. 34 (2006) 2723–2735.
[10] K.N. Luu, A.T. Phan, V. Kuryavyi, L. Lacroix, D.J. Patel, Structure of the human telomere in K^+ solution: an intramolecular (3 + 1) G-quadruplex scaffold, J. Am. Chem. Soc. 128 (2006) 9963–9970.
[11] J. Dai, M. Carver, C. Punchihewa, R.A. Jones, D. Yang, Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence, Nucleic Acids Res. 35 (2007) 4927–4940.

[12] A.T. Phan, K.N. Luu, D.J. Patel, Different loop arrangements of intramolecular human telomeric (3 + 1) G-quadruplexes in K^+ solution, Nucleic Acids Res. 34 (2006) 5715–5719.
[13] A.T. Phan, V. Kuryavyi, K.N. Luu, D.J. Patel, Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution, Nucleic Acids Res. 35 (2007) 6517–6525.
[14] K.W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D.J. Patel, K.N. Luu, A.T. Phan, Structure of the human telomere in K^+ solution: a stable basket-type G-quadruplex with only two G-tetrad layers, J. Am. Chem. Soc. 131 (2009) 4301–4309.
[15] Y. Wang, D.J. Patel, Solution structure of a parallel-stranded G-quadruplex DNA, J. Mol. Biol. 234 (1993) 1171–1183.
[16] N.H. Campbell, G.N. Parkinson, Crystallographic studies of quadruplex nucleic acids, Methods 43 (2007) 252–263.
[17] M. Adrian, B. Heddi, A.T. Phan, NMR spectroscopy of G-quadruplexes, Methods 57 (2012) 11–24.
[18] M. Webba da Silva, NMR methods for studying quadruplex nucleic acids, Methods 43 (2007) 264–277.
[19] D.J. Patel, A.T. Phan, V. Kuryavyi, Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics, Nucleic Acids Res. 35 (2007) 7429–7455.
[20] J. Dai, D. Chen, R.A. Jones, L.H. Hurley, D. Yang, NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region, Nucleic Acids Res. 34 (2006) 5133–5144.
[21] J. Dai, T.S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R.A. Jones, D. Yang, An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution, J. Am. Chem. Soc. 128 (2006) 1096–1098.
[22] K.W. Lim, L. Lacroix, D.J. Yue, J.K. Lim, J.M. Lim, A.T. Phan, Coexistence of two distinct G-quadruplex conformations in the hTERT promoter, J. Am. Chem. Soc. 132 (2010) 12331–12342.
[23] A. Virno, L. Mayol, A. Ramos, F. Fraternali, B. Pagano, A. Randazzo, Structural insight into the hTERT intron 6 sequence d(GGGGTGAAAGGGG) from ^{1}H NMR study, Nucleosides Nucleotides Nucleic Acids 26 (2007) 1133–1137.
[24] R.D. Gray, J.B. Chaires, Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures, Nucleic Acids Res. 36 (2008) 4191–4203.
[25] P. Balagurumoorthy, S.K. Brahmachari, Structure and stability of human telomeric sequence, J. Biol. Chem. 269 (1994) 21858–21869.
[26] S.L. Palumbo, S.W. Ebbinghaus, L.H. Hurley, Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands, J. Am. Chem. Soc. 131 (2009) 10878–10891.
[27] D. Sun, L.H. Hurley, Biochemical techniques for the characterization of G-quadruplex structures: EMSA, DMS footprinting, and DNA polymerase stop assay, Methods Mol. Biol. 608 (2010) 65–79.
[28] N. Andersen, P. Rasmussen, The constitution of clerocidin. A new antibiotic isolated from Oidiodendron truncatum, Tetrahedron Lett. 25 (1984) 465–468.
[29] B. Gatto, S. Richter, S. Moro, G. Capranico, M. Palumbo, The topoisomerase II poison clerocidin alkylates non-paired guanines of DNA: implications for irreversible stimulation of DNA cleavage, Nucleic Acids Res. 29 (2001) 4224–4230.
[30] S. Richter, B. Gatto, D. Fabris, K. Takao, S. Kobayashi, M. Palumbo, Clerocidin alkylates DNA through its epoxide function: evidence for a fine tuned mechanism of action, Nucleic Acids Res. 31 (2003) 5149–5156.
[31] S.N. Richter, I. Menegazzo, D. Fabris, M. Palumbo, Concerted bis-alkylating reactivity of clerocidin towards unpaired cytosine residues in DNA, Nucleic Acids Res. 32 (2004) 5658–5667.
[32] S.N. Richter, I. Menegazzo, M. Nadai, S. Moro, M. Palumbo, Reactivity of clerocidin towards adenine: implications for base-modulated DNA damage, Org. Biomol. Chem. 7 (2009) 976–985.
[33] S. Richter, D. Fabris, M. Binaschi, B. Gatto, G. Capranico, M. Palumbo, Effects of common buffer systems on drug activity: the case of clerocidin, Chem. Res. Toxicol. 17 (2004) 492–501.
[34] N.J. Greenfield, Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, Nat. Protoc. 1 (2006) 2527–2535.
[35] A.T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D.J. Patel, Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter, J. Am. Chem. Soc. 129 (2007) 4386–4392.
[36] H. Fernando, A.P. Reszka, J. Huppert, S. Ladame, S. Rankin, A.R. Venkitaraman, S. Neidle, S. Balasubramanian, A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene, Biochemistry 45 (2006) 7854–7860.
[37] V. Kuryavyi, A.T. Phan, D.J. Patel, Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter, Nucleic Acids Res. 38 (2010) 6757–6773.
[38] A. Ambrus, D. Chen, J. Dai, R.A. Jones, D. Yang, Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization, Biochemistry 44 (2005) 2048–2058.
[39] A.T. Phan, Y.S. Modi, D.J. Patel, Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter, J. Am. Chem. Soc. 126 (2004) 8710–8716.
[40] P.A. Rachwal, I.S. Findlow, J.M. Werner, T. Brown, K.R. Fox, Intramolecular DNA quadruplexes with different arrangements of short and long loops, Nucleic Acids Res. 35 (2007) 4214–4222.
[41] A.T. Phan, V. Kuryavyi, H.Y. Gaw, D.J. Patel, Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter, Nat. Chem. Biol. 1 (2005) 167–173.
[42] A. Avino, C. Fabrega, M. Tintore, R. Eritja, Thrombin binding aptamer, more than a simple aptamer: chemically modified derivatives and biomedical applications, Curr. Pharm. Des. 18 (2012) 2036–2047.
[43] P. Schultze, R.F. Macaya, J. Feigon, Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG), J. Mol. Biol. 235 (1994) 1532–1547.
[44] T.S. Dexheimer, D. Sun, L.H. Hurley, Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter, J. Am. Chem. Soc. 128 (2006) 5404–5415.
[45] S. Kendrick, L.H. Hurley, The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements, Pure Appl. Chem. 82 (2010) 1609–1621.
[46] A.M. Maxam, W. Gilbert, Sequencing end-labeled DNA with base-specific chemical cleavages, Methods Enzymol. 65 (1980) 499–560.
[47] D. Sun, L.H. Hurley, The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression, J. Med. Chem. 52 (2009) 2863–2874.
[48] A.X. Xiang, D.A. Watson, T. Ling, E.A. Theodorakis, Total synthesis of clerocidin via a novel, enantioselective homoallenylboration methodology, J. Org. Chem. 63 (1998) 6774–6775.