Putative DNA G-quadruplex formation within the promoters of Plasmodium falciparum var genes

Citation for published version:
Rowe, JA, Smargiasso, N, Gabelica, V, Damblon, C, Rosu, F, De Pauw, E, Teulade-Fichou, M-P & Claessens, A 2009, 'Putative DNA G-quadruplex formation within the promoters of Plasmodium falciparum var genes' BMC Genomics, vol 10, 362. DOI: 10.1186/1471-2164-10-362

Digital Object Identifier (DOI):
10.1186/1471-2164-10-362

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
BMC Genomics

Publisher Rights Statement:
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Putative DNA G-quadruplex formation within the promoters of *Plasmodium falciparum* var genes

Nicolas Smargiasso*†1, Valérie Gabelica1, Christian Damblon2, Frédéric Rosu1, Edwin De Pauw1, Marie-Paule Teulade-Fichou3, J Alexandra Rowe4 and Antoine Claessens†4

Address: 1Mass Spectrometry Laboratory, GIGA-Research, University of Liege, Liege, Belgium, 2Structural Biological Chemistry Laboratory, University of Liege, Liege, Belgium, 3Institut Curie, Section Recherche, CNRS UMR176, Centre Universitaire Paris XI, Bat. 110, 91405 Orsay, France and 4Centre for Immunity, Infection and Evolution, Institute of Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

Email: Nicolas Smargiasso* - nsmargiasso@ulg.ac.be; Valérie Gabelica - V.Gabelica@ulg.ac.be; Christian Damblon - c.damblon@ulg.ac.be; Frédéric Rosu - F.Rosu@ulg.ac.be; Edwin De Pauw - E.DePauw@ulg.ac.be; Marie-Paule Teulade-Fichou - marie-paule.teulade-fichou@curie.fr; J Alexandra Rowe - Alex.Rowe@ed.ac.uk; Antoine Claessens - a.j.a.h.claessens@sms.ed.ac.uk

* Corresponding author    †Equal contributors

Abstract

**Background:** Guanine-rich nucleic acid sequences are capable of folding into an intramolecular four-stranded structure called a G-quadruplex. When found in gene promoter regions, G-quadruplexes can downregulate gene expression, possibly by blocking the transcriptional machinery. Here we have used a genome-wide bioinformatic approach to identify Putative G-Quadruplex Sequences (PQS) in the *Plasmodium falciparum* genome, along with biophysical techniques to examine the physiological stability of *P. falciparum* PQS in vitro.

**Results:** We identified 63 PQS in the non-telomeric regions of the *P. falciparum* clone 3D7. Interestingly, 16 of these PQS occurred in the upstream region of a subset of the *P. falciparum* var genes (group B var genes). The var gene family encodes PfEMP1, the parasite’s major variant antigen and adhesin expressed at the surface of infected erythrocytes, that plays a key role in malaria pathogenesis and immune evasion. The ability of the PQS found in the upstream regions of group B var genes (UpsB-Q) to form stable G-quadruplex structures in vitro was confirmed using 1H NMR, circular dichroism, UV spectroscopy, and thermal denaturation experiments. Moreover, the synthetic compound BOQ1 that shows a higher affinity for DNA forming quadruplex rather than duplex structures was found to bind with high affinity to the UpsB-Q.

**Conclusion:** This is the first demonstration of non-telomeric PQS in the genome of *P. falciparum* that form stable G-quadruplexes under physiological conditions in vitro. These results allow the generation of a novel hypothesis that the G-quadruplex sequences in the upstream regions of var genes have the potential to play a role in the transcriptional control of this major virulence-associated multi-gene family.
Background

Plasmodium falciparum is responsible for the majority of malaria cases worldwide and is the cause of an estimated 300–500 million infections and 1–2 million deaths per year [1]. The parasite invades circulating red blood cells and causes them to adhere to microvascular endothelial cells and sequester in blood microvessels, leading to vascular obstruction. The only proteins known to be responsible for this cytoadherence are members of the P. falciparum erythrocyte membrane protein one (PfEMP1) family (reviewed in [2]). These highly polymorphic parasite-derived erythrocyte surface proteins are encoded by a repertoire of 50 to 60 var genes. Crucially, each parasite expresses only one var gene at a time, with transcription sometimes being switched to a different var gene in subsequent generations, so allowing antigenic variation and immune evasion [2].

Despite their extreme sequence variability in the coding regions, var genes can be divided into 3 major groups (A, B and C) according to the presence of one of three conserved 5’ upstream (Ups) sequences (UpsA, UpsB and UpsC) [3]. Their chromosomal position further subdivides them into centromeric (C) or telomeric (T) locations [4]. Var gene groups have functional and clinical significance. For example, group B and C var genes are known to bind to the endothelial receptor CD36 [2], whereas group A var genes have been linked to the most severe clinical forms of malaria [5,6].

The mechanisms regulating var gene transcription are not well understood and are currently the subject of intensive investigations. Var gene expression is thought to be regulated at the level of transcription initiation [7]. Many mechanisms have been suggested as being involved in the silencing of non-transcribed var genes including var intron sequences [8] and SPE and CPE motifs located in UpsB and UpsC sequences respectively [9,10]. The histone deacetylase PFSir2 is thought to be required for chromatin silencing in the subtelomeric regions [11], and histone methylation in the 5’ Ups region has been shown to regulate transcription of the var2csa gene [12]. Finally, a var-specific subnuclear expression site has been proposed recently [13]. How these pieces of evidence fit together is still unclear, and other mechanisms may be discovered before the full picture of var gene transcriptional control is obtained.

DNA usually maintains a double helix structure, however, recent evidence shows that in guanine-rich regions, DNA can adopt a more complex structure called a G-quadruplex [14] (Figure 1). G-quadruplexes are composed by the stacking of guanine tetrads, each one being stabilized by 8 Hoogsteen Hydrogen bonds (Figure 1A). Consequently, sequences containing four groups of three guanines are theoretically able to fold into a G-quadruplex containing three guanine tetrads (Figure 1B). Although there are also a few examples of G-quadruplexes formed from two guanine tetrads [15], these are much less stable and thus less likely to occur in vivo [16]. Hence, as in previous genome-wide analyses of potential G-quadruplex-forming sequences [17,18], we chose to investigate here sequences containing at least three tracks of four guanines.

G-quadruplexes are also stabilized by interactions with cations located between the tetrads, at the center of the structure. Potassium and sodium are the most commonly described G-quadruplex stabilizing cations, although ammonium and strontium can also assume this function [19-22]. It was previously reported that there are about 376,000 potential G-quadruplex structures in the human genome [17,18], and about 40% of human genes contain a putative G-quadruplex in their promoter [23]. Initial reports indicate a possible role for G-quadruplex sequences in the regulation of telomere length [24,25] and the transcriptional regulation of several genes such as c-MYC, c-kit, or KRAS [23,26-32]. For example, in the case of the c-MYC proto-oncogene, a single nucleotide mutation that destabilizes the G-quadruplex structure in the promoter region leads to a three-fold increase in basal transcription levels, suggesting that the G-quadruplex acts as a transcriptional repressor element [27]. Furthermore, a small ligand that binds to and stabilizes the G-quadruplex structure was shown to suppress further c-MYC transcriptional activity [27].

Given the increasing evidence for the importance of G-quadruplex sequences in gene regulation, we decided to investigate whether G-quadruplexes could be discovered in the genome of P. falciparum, and in particular to determine whether there are any G-quadruplex sequences in the upstream regions of var genes that have the potential to play a role in the transcriptional control of this major virulence-associated multi-gene family. In addition, the ability of potential G-quadruplex sequences from P. falciparum to form stable G-quadruplex structures under physiological conditions was examined using biophysical techniques.

Results and discussion

Identification of putative G-quadruplex forming sequences in the P. falciparum genome

The genome of P. falciparum clone 3D7 was searched for Putative Quadruplex Sequences (PQS) using QGRS-Mapper [33] on both the positive and negative strands. We set up the QGRS-mapper software to identify all PQS with four repeats of at least three guanines interrupted by loops of a maximum length of 11 nucleotides. As expected, most PQS were found in the telomeres (828 out of 891) due to their repetitive sequence: GGGTT(T/C)A (see Additional
file 1). These telomeric G-quadruplexes of *P. falciparum* have recently been described by De Cian et al [34]. Here we focused on the non-telomeric PQS because of their potential role in gene transcriptional regulation. We identified 63 non-telomeric PQS (listed in full in Additional file 2). This is an average of one PQS per 380 kb, which is a much lower ratio than that seen in the human genome (1 PQS per ~8 kb) [17]. This was expected due to the extreme AT-richness (80.6%) of the *P. falciparum* genome [35]. 37 of the 63 PQS are in intergenic regions, and of the 26 PQS within genes, 9 are on the coding strand and 17 on the non-coding strand.

**PQS in the upstreamB region of var genes**

Most importantly, 16 out of the 63 PQS were found in the upstreamB regions of *var* genes, 1612 to 1707 bp upstream of the initiation codon (Table 1 and Additional file 2). These 16 PQS contain only three distinct sequences that thus represent three slightly different putative G-quadruplexes that we named UpsB-Q-1, UpsB-Q-2 and UpsB-Q-3 in order of their frequency (sequences shown in Table 2). As the *var* gene repertoire varies from one clone to another, we also searched for PQS in the upstream regions of *var* genes in *P. falciparum* clone HB3 (available from the Broad Institute website, http://www.broad.mit.edu). Using the same parameters, we found 11 PQS in the upstream B region of *var* genes in HB3 (Table 1). Interestingly, UpsB-Q-1 was also the most common PQS (7 out of 11) in parasite clone HB3. One PQS was found in HB3 but not in 3D7 (named UpsB-Q-4, Table 1 and 2). These four types of PQS do not exist in any other sequenced organism to date (BLAST analysis, data not shown).

**Evidence of G-quadruplex formation by PQS in the upstream B region of var genes**

In order to confirm the formation of G-quadruplexes by these sequences, two of them (UpsB-Q-1 and 2) were analyzed by $^1$H NMR in the presence of ammonium or potassium ions. Using this technique, it is well established that the presence of imino protons with a chemical shift between 10 and 12 ppm is characteristic of the formation of G-quadruplexes [36-38]. Spectra recorded on the presumably unstructured oligonucleotides in water did not show signals beyond 9 ppm, indicating that imino protons are in fast exchange with bulk water. The spectra
recorded after the addition of 150 mM cation are presented in Figure 2 and show clearly the presence of imino protons for the four samples. Most importantly, for UpsB-Q1 in ammonium and potassium, 11 imino peaks are clearly distinguished. In potassium, the peak at 11.93 ppm was resolved in two different peaks at 37°C (data not shown), indicating the formation of a single stable structure containing three quartets. In ammonium, the intensity of the peak at 11.55 ppm could also indicate a superposition of two different peaks. On the UpsB-Q-2 spectra, the imino peaks are less well resolved, indicating some structural polymorphism. Moreover, in both PQS sequences, peaks were observed at higher chemical shift, indicating the possible presence of additional structures like AT-rich hairpins on these G-quadruplexes. These NMR data show that UpsB-Q1 and UpsB-Q-2 do form stable G-quadruplex structures in the presence of physiological concentrations of potassium ions.

The four types of PQS found in the upstream B regions of var genes were also examined by circular dichroism (CD), which provides information about the orientation of strands within a G-quadruplex, because the CD signal changes with the syn/anti orientation about the glycosyl bonds. In parallel G-quadruplexes, the CD spectrum typically exhibits a positive peak around 260 nm and a negative peak around 240 nm, whereas in antiparallel G-quadruplexes, the CD spectrum displays a negative peak around 260 nm and a positive peak at 295 nm [39-41]. Results of CD for each UpsB-Q are showed in Figure 3. For UpsB-Q-2, 3 and 4 in potassium and ammonium, a minimum around 243 nm and a maximum around 295 nm were observed, while for UpsB-Q-1 in potassium the minimum was around 250 nm. These kinds of spectra are generally attributed to hybrid conformations (containing a mixture of both parallel and antiparallel strand orientations). As in sodium, the four sequences seem to adopt an antiparallel conformation, since a minimum near 260 nm and a maximum at 290 – 295 nm are observed. However, a shoulder was observed in the three cationic conditions at 270–275 nm for UpsB-Q-2, 3 and 4 but not for UpsB-Q-1. The absence of a shoulder in the case of UpsB-Q-1 could therefore indicate a different conformation of this G-quadruplex. Furthermore, the depth of the minimum observed for UpsB-Q-1 was systematically bigger than for the three other sequences. Together with NMR, this result suggests that the single structure adopted by UpsB-Q-1 could be systematically bigger than that of the three other sequences. Together with NMR, this result suggests that the single structure adopted by UpsB-Q-1 could therefore indicate a different conformation of this G-quadruplex.
bases for UpsB-Q-3 and a single base for UpsB-Q-4) are sufficient to induce structural polymorphism.

Finally, thermal difference spectra (TDS) were recorded [42]. Similarly to CD, this technique differentiates between the various potential structures adopted by DNA. It was shown previously by Mergny et al [42] that G-quadruplexes exhibit two positive peaks at 243 and 273 nm and one negative peak at 295 nm, while other DNA structures show different combinations of maxima and minima (see [42] for details of maxima and minima characterizing other DNA structures). The spectra obtained for the PQS from the var gene upstream B regions match expectations for G-quadruplex structures (Figure 4). There is a negative peak at 295 nm for all sequences in the three cations. Positive peaks were observed at 246 (only in

Figure 2

\(^1\)H NMR spectra of UpsB-Q-1 in potassium (A) and ammonium (C) and of UpsB-Q-2 in potassium (B) and ammonium (D) (cation concentration 150 mM). These spectra were acquired at 25°C. *: this peak was resolved in two peaks at 37°C. **: due to its intensity, this peak most likely corresponds to the superposition of two peaks.

Figure 3

Circular dichroism spectra of putative G-quadruplex sequences (PQS) from the upstream B regions of var genes in 150 mM potassium, sodium and ammonium cation. UpsB-Q-1 (blue), UpsB-Q-2 (red), UpsB-Q-3 (pink) and UpsB-Q-4 (green). Characteristic signatures of hybrid G-quadruplexes (i.e. containing a mixture of parallel and antiparallel strand orientations) are observed in potassium and ammonium. In sodium, G-quadruplexes are predominantly antiparallel.
potassium), 256 (except in ammonium) and 267 nm. These values are in good agreement with previously described results [42] and thus confirm the ability of these sequences to form G-quadruplexes. There is a small wavelength shift for positive peaks, which may be because of the presence of a long loop (10 bases) in the PQS of the var gene upstream regions (all the G-quadruplexes previously tested contained no more than four bases in their loops).

The stoichiometries of the G-quadruplexes formed by the UpsB-Q were also examined to determine whether these structures are likely to form intra-molecular bonds (unimolecular structures) or inter-molecular bonds (multimolecular structures) [43]. Mass spectrometry showed only monomeric DNA (Additional file 3), indicating that the UpsB-Q form intra-molecular G quadruplex structures (inter-molecular structures would have been indicated by the presence of multimers by mass spectrometry).

**Stability of G-quadruplexes formed by PQS in the upstreamB region of var genes**

When the UV absorbance is monitored at 295 nm [44], thermal denaturation experiments allow determination of the $T_m$ of G-quadruplexes, i.e. the temperature at which the half of the signal is lost and thus indicate the relative stability of structures adopted by the oligonucleotides. Because G-quadruplexes with long loops are usually less stable that those with small loops [43,45,46], it was necessary to check if the UpsB-Q are able to form G-quadruplexes that are stable under physiological conditions ($37^\circ C$ and $[K^+] \approx 150 \text{ mM}$). These results are shown in Figure 4 and the $T_m$'s are listed in Table 3. In potassium, the four PQS sequences from the var gene upstream B regions have a $T_m$ about 50°C, and the transition curves (Figure 5) show that, at 37°C, the proportion of folded G-quadruplexes is higher than 85% for the four sequences, confirming their potential to form G-quadruplexes in living cells. As expected, $T_m$'s observed in sodium and ammonium are lower than in potassium. This is generally attributed to the weaker stabilization capacity of these two cations, due to their smaller ionic radius [47-49]. The $T_m$ ranking $K^+ > Na^+ > NH_4^+$ is also characteristic of antiparallel G-quadruplexes (parallel G-quadruplexes have $T_m$ ranking $K^+ > NH_4^+ > Na^+$) [43].

**Interactions of UpsB-Q G-quadruplexes with a ligand**

In addition to the potential transcriptional repressor activity of G-quadruplex sequences themselves [27], it has been shown previously that G-quadruplex ligands can further suppress transcription of genes containing potential G-quadruplexes in their promoters, by impeding the binding of proteins needed for initiation of transcriptional activity on DNA [50,51]. Moreover, these molecules are also able to interfere with telomere structure and to indirectly induce their shortening [52-55]. These molecules are thus promising weapons in the fight against cancer, since this disease needs both a high expression of oncogenes and stable telomere length to develop and survive [56-62]. With the discovery of G-quadruplex forming sequences in the genome of *P. falciparum*, it can be hypothesized that these ligands may also have the potential to affect parasite gene expression by stabilizing G-quadruplexes located in gene promoter regions.

It was thus decided to evaluate the equilibrium dissociation constant of each UpsB-Q with the G-quadruplex ligand BOQ1 (Figure 6A). BOQ1 is a synthetic compound that exhibits a good selectivity for G-quadruplex versus duplex DNA [63,64]. A relatively high association constant with the UpsB-Q sequences would therefore provide additional evidence for the presence of G-quadruplex structures. Binding constants were determined by electrospray mass spectrometry. A typical spectrum of a G-quadruplex-ligand mixture is shown in Figure 6B. The peaks correspond to the free DNA sequence, and to 1:1 com-

**Figure 4**

Thermal denaturation spectra of PQS from the upstream B regions of var genes in potassium, sodium and ammonium (cation concentration 150 mM). UpsB-Q-1 (green), UpsB-Q-2 (blue), UpsB-Q-3 (red) and UpsB-Q-4 (pink).
The equilibrium dissociation constants of all UpsB-Q with BOQ 1 were deduced from the relative intensity of peaks of free DNA and complexes, as described previously [65]. For the four sequences, the values are around 2 μM (Table 3). They are lower than those obtained by mass spectrometry for the binding of BOQ1 to telomeric G-quadruplexes (5.7 μM) or to model duplex sequences (57 μM) (unpublished data). These results confirm the ability of the PQS in the upstream B regions of the var genes to fold in G-quadruplexes, and show that G-quadruplex ligands are likely to bind to these structures within the P. falciparum genome, and could therefore be tested for biological activity against the parasite.

### Table 3: Melting temperature ($T_m$) of G-quadruplex sequences from the upstream B regions of var genes and the equilibrium dissociation constant of the PQS with the G-quadruplex ligand BOQ 1 (shown in the $K_d$ column)

| PQS          | $T_m$ (°C) | $K^+$ (μM) | $Na^+$ (μM) | $NH_4^+$ (μM) | $K_d$ (μM) |
|--------------|------------|------------|-------------|--------------|------------|
| UpsB-Q-1     | 47.2 ± 0.6 | 35.1 ± 1.2 | 31.6 ± 0.7  | 2.6 ± 0.5    |            |
| UpsB-Q-2     | 49 ± 0.5   | 36.1 ± 2   | 27.7 ± 0.5  | 1.4 ± 0.4    |            |
| UpsB-Q-3     | 50 ± 1.2   | 36.9 ± 1.1 | 34.2 ± 1.1  | 1.7 ± 0.7    |            |
| UpsB-Q-4     | 49.3 ± 1.4 | 39.4 ± 0.9 | 32.3 ± 0.3  | 2.7 ± 1.1    |            |

$^1$ Thermal denaturation experiments were performed at the cation concentration of 150 mM. Thermal denaturation experiments were repeated twice and $T_m$ values shown are the mean of four values obtained from four curves (two heatings and two coolings).

### Conclusion

Increasing evidence suggests that G-quadruplexes play a role in gene transcriptional regulation in humans and other organisms. We identified 63 potential G-quadruplex sequences in the non-telomeric regions of the genome of P. falciparum clone 3D7. 16 of these PQS occurred in the upstream region of group B var genes. The var gene-related PQS were shown to form stable G-quadruplex structures in vitro under physiological conditions and bind with high affinity to a known G-quadruplex ligand. It is noteworthy that the most prevalent sequence UpsB-Q-1 (dCAGGGT-TAAGGTTATAACITTTAGGGT-AGGGTT) adopts a single structure which is stable in physiological conditions.

---

**Figure 5**

Thermal denaturation curves (heating) of putative G-quadruplex sequences (PQS) from the upstream B regions of var genes in 150 mM potassium cation.

UpsB-Q-1 (blue), UpsB-Q-2 (red), UpsB-Q-3 (pink) and UpsB-Q-4 (green). For the four sequences, $T_m$'s are about 50°C and the proportion of folded G-quadruplexes at 37°C is above 85%.
This discovery allows us to generate a new hypothesis concerning var gene regulation mechanisms in P. falciparum, in which a helicase such as PFI0910w could be involved in G-quadruplex unwinding and thus facilitate RNA polymerase transcriptional activity. The role of G-quadruplexes in Plasmodium gene regulation, the structure of these G-quadruplexes, and their use as potential drug targets merits further research.

Methods
Bioinformatic analysis
Both strands of each chromosome of the P. falciparum 3D7 clone (Plasmodb_5.4 [73]) were analyzed using QGRS-Mapper [33]. The parameters used were: Max length: 33; Min G-group: 3; loop size: 0 to 11. The P. falciparum HB3 genome was downloaded from the Broad Institute http://www.broad.mit.edu. Upstream sequences of var genes were analyzed using QGRS-Mapper with the same parameters.

Materials
All oligonucleotides were ordered from Eurogentec (Seraing, Belgium) with Oligold quality. The oligonucleotide sequences used are shown in the Table 2. Oligonucleotides were received lyophilized and stock solutions were prepared in bi-distilled water with 300 μM total strand concentration. For all experiments, the stock solution was heated at 80°C for 5 minutes, diluted using a cold aqueous solution containing either KCl, NaCl or NH₄OAc to reach the desired DNA concentration in 150 mM cation, and then cooled rapidly on ice. 10 mM lithium cacodylate, pH 7.4 was added in thermal denaturation and circular dichroism experiments. The molecule BOQ1 was synthesized as described previously [74].
Circular dichroism
Experiments were performed on a Jasco J-810 spectropolarimeter using 1-cm path length cells (Hellma, type No. 114-QS, France). The final concentration of oligonucleotide was 5 μM in a buffer containing 150 mM salt and 10 mM lithium cacodylate, pH 7.4. For each sample, five spectra were recorded from 220 nm to 350 nm with a scan rate of 100 nm/min.

NMR
NMR samples were prepared by dissolving the oligonucleotides in H₂O/D₂O 90/10, lithium cacodylate 10 mM, pH 7.4 to get a oligonucleotide final concentration of 270 μM. Ammonium acetate or potassium chloride were progressively titrated in to a final cation concentration of 150 mM. NMR data were collected at 500 MHz on a Bruker Avance spectrometer (fitted with a TXI triple resonance probe with z-axis gradient). 1D 1H spectra were recorded at a temperature of 25°C using a WATERGATE sequence with a water flip-back pulse [75, 76].

Thermal denaturation
Thermal denaturation experiments were carried out on a Uvikon XS spectrophotometer (Secomam), using 1-cm path length quartz cells (Hellma, type No. 115B-QS, France). The final oligonucleotide concentration was 5 μM in 150 mM salt and 10 mM lithium cacodylate, pH = 7.4. Absorbance was monitored as a function of the temperature at 295, 240, 260 nm for the determination of the melting temperature (T_m) [44] and at 405 nm as control wavelength. Gradient was 0.2°C/min between 10 and 90°C. Melting temperatures were determined using the method described by Marky and Breslauer [77]. Before heating and after the cooling, spectra were recorded from 220 to 440 nm, to allow thermal difference spectra (TDS) to be obtained. TDS were obtained by subtracting the low temperature curve from the high temperature curve and normalization, as described previously by Mergny et al. [42].

Electrospray mass spectrometry
All measurements were carried out on a Q-TOF Ultima Global mass spectrometer (Micromass, now Waters, Manchester, U.K.), using the electrospray ionization (ESI) source in negative mode, as described previously [78]. Source conditions were optimized to avoid in-source fragmentation: capillary voltage = -2.2 kV, cone voltage = 50 V, RF, source block temperature = 80°C, and desolvation gas temperature = 100° C. Source backing pressure was set to 3.5 mbar. Oligonucleotide samples were first prepared at 50 μM final concentration in NH₄OAc 150 mM. Just before injection in the mass spectrometer, they were further diluted to 10 μM in 150 mM NH₄OAc and 20% methanol. The role of methanol is to increase ion signals.

Figure 7
Alignment of PFI0910w (P. falciparum) with two helicases of the RecQ family, BLM (human) and Sgs1p (Saccharomyces cerevisiae). The seven conserved motifs of the helicase domain are indicated with a red bar.
Authors' contributions

NS designed research, performed biophysical experiments, analyzed data and wrote the manuscript. AC designed research, performed the bioinformatics analysis, analyzed data and wrote the manuscript. CD performed NMR experiments and wrote the manuscript. FR contributed to biophysical experiments and analyzed data, EDP wrote the manuscript, MPTF contributed new reagent, JAR wrote the manuscript, VG analyzed data and wrote the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Figure S1. Distribution of telomeric and non-telomeric Putative G-Quadruplex Sequences (PQS) in Plasmodium falciparum 3D7. Click here for file

[http://www.biomedcentral.com/content/full/1471-2164-10-362-S1.doc]

Additional file 2

Figure S2 Full list of non-telomeric PQS in the Plasmodium falciparum 3D7 genome. Click here for file

[http://www.biomedcentral.com/content/full/1471-2164-10-362-S2.xls]

Additional file 3

Supplementary Materials. Stoichiometry of G quadruplexes formed from the UTR-Q. Click here for file

[http://www.biomedcentral.com/content/full/1471-2164-10-362-S3.doc]

Acknowledgements

We thank Prof. Pierre Colson for UV experiments, Dr Pascal De Tullio for NMR experiments and Dr. Andre Magne for CD experiments. We acknowledge the FNRS and FRIA (VG is a FNRS research associate, FR is a FRIA doctoral fellow) and the financial contributions of the Fonds de la Recherche Scientifique-FNRS (FRFC 2.4.623.03 to EDP, CC 1.5.096.08 to VG) and the University of Liège (Starting Grant D-08/10 to VG). JAR is funded by a Wellcome Trust Senior Research Fellowship in Biostatistics.

References

1. Snow RW, Guerria CA, Noor AM, Myint HY, Hay SI: The global distribution of clinical episodes of Plasmodium falciparum malaria. Nature 2005, 434:214-217.
2. Kyes SA, Kraemer SM, Smith JD: Antigenic variation in Plasmodium falciparum: gene organization and regulation of the var multitope family. Eukaryot Cell 2007, 6:151-11528.
3. Lavstsen T, Salani A, Jensen AT, Amot DE, Theander TG: Subgrouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. Malar J 2003, 2:27
4. Kraemer SM, Kyes SA, Agarwal G, Springer AL, Nelson SO, Christodoulou Z, Smith LM, Wang W, Levin E, Newbold CI, et al.: Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates. BMC Genomics 2007, 8:45.
5. Jensen AT, Magistrado P, Sharp S, Ijegboghen L, Lavstsen T, Chiu R, Salani A, Vestergaard LS, Lusingu JP, Hermsen R, et al.: Plasmodium falciparum associated with severe childhood malaria preferentially expresses PIEMP1 encoded by group A var genes. J Exp Med 2004, 199:1179-1190.
6. Krypaciou HM, Stone GN, Chalis RJ, Raza A, Lyke KE, Thera MA, Kone AK, Dombou OK, Powe CW, Rowe JA: Differential var gene transcription in Plasmodium falciparum isolates from patients with cerebral malaria compared to hyperparasitemia. Mol Biochem Parasitol 2006, 150:211-218.
7. Kyes S, Christodoulou Z, Piches R, Kriek N, Horrocks P, Newbold C: Plasmodium falciparum var gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. Mol Microbiol 2007, 63(1213-1247.
8. Dzikowski R, Frank M, Deitsch K: Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. PLoS Pathog 2006, 2:e22.
9. Voss TS, Kaestli M, Vogel D, Bopp S, Beck HP: Identification of nuclear proteins that interact differentially with Plasmodium falciparum var gene promoters. Mol Microbiol 2003, 49:1296-1307.
10. Voss TS, Tonkin CJ, Marty AJ, Thompson JK, Healer J, Crabb BS, Cowman AF: Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. Mol Microbiol 2007, 66:139-150.
11. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Theander TG, Freitas-Junior LH, Scherf A, Crabb BS, Cowman AF: Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell 2005, 121:13-24.
12. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A: 5’ flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol 2007, 66:1296-1305.
13. Dzikowski R, Li F, Amulic B, Lisberg A, Frank M, Patel S, Wellens TE, Deitsch KW: Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO Rep 2007, 8:959-965.
14. Simonsson T: G-quadruplex DNA structures – variations on a theme. Biochim Biophys Acta 2001, 138:621-628.
15. Macaya RF, Schultze P, Smith FW, Roe JA, Feigon J: Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. Proc Natl Acad Sci USA 1993, 90:3745-3749.
16. Darby RA, Sologoub M, McNamee C, Brown N, Barton C, Brown T, Fox KR: High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. Nucleic Acids Res 2002, 30(9):e39.
17. Huppert JL, Balasubramanian S: Prevalence of quadruplexes in the human genome. Nucleic Acids Res 2003, 31:2908-2916.
18. Todd AK, Johnston M, Neidle S: Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res 2005, 33:2901-2906.
19. Williamson JR, Raghuraman MK, Cech TR: Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 1989, 59:871-880.
20. Nagesh N, Chatterji D: Ammonium ion at low concentration stabilizes the G-quadruplex formation by telomeric sequence. J Biochem Biophys Methods 1995, 30:1-8.
21. Chen FM: Sr2+ facilitates intermolecular G-quadruplex formation of telomeric sequences. Biochemistry 1992, 31:3769-3776.
22. Pedrosot M, Duarte LF, Yaney G, Baker AM, Fletcher TM: Induction of parallel human telomeric G-quadruplex structures by Sr2+. Biochim Biophys Acta 2007, 1778:298-303.
23. Huppert JL, Balasubramanian S: G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res 2007, 35:406-413.
24. Sen D, Gilbert W: Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 1988, 334:364-366.
25. Henderson E, Hardin CC, Walk SK, Tinoco I Jr, Blackburn EH: Telomeric DNA oligonucleotides form novel intramolecular....
structures containing guanine-guanine base pairs. Cell 1987, 31:899-908.
26. Vries AI, Lilley DM: Retinoblastoma susceptibility genes contain S’ sequences with a high propensity to form guanine-tetrad structures. Nucleic Acids Res 1992, 20:49-53.
27. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH: 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 USA 2002, 99:11593-11598.
28. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S: Putative DNA quadruplex formation within the human c-kit oncogene. J Am Chem Soc 2005, 127:10504-10509.
29. Cogo S, Xodo LE: G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. Nucleic Acids Res 2006, 34:2536-2549.
30. Phan AT, Kurywayi V, Burge S, Neidle S, Patel DJ: Structure of an Antiparallel G-Quadruplex Scaffold in the Human c-Kit Promoter. J Am Chem Soc 2007, 129:4386-4392.
31. Simonsson T, Pecinka P, Kubista M: DNA tetraplex formation in the control region of c-myc. Nucleic Acids Res 1998, 26:1167-1172.
32. Maiuz N: Dynamic roles for G4 DNA in the biology of eukaryotic cells. Nat Struct Mol Biol 2003, 10:1055-1059.
33. Klini O, D’Antoni L, Bagga PG: QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. Nucleic Acids Res 2006, 34:W676-682.
34. Rio JD, Amat A, Gauger P, Mouray D, Bernaud H, Monchaud D, Teulea-Fichou MP, Mergrny JL, Alberti P: Plasmid G-Quadruplex Sequences: Structure, Stability and Quadruplex Targeting by Small Compounds. ChemBioChem 2008, 9:2730-2739.
35. Gardiner MJ, Hall N, Fung E, White O, Berriman H, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, et al.: Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 2002, 419:498-511.
36. Phan AT, Patel DJ: Two-repeat human telomeric d(TAGGTT-TAGGTT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: distinct topologies, thermodynamic properties, and folding/unfolding kinetics. J Am Chem Soc 2003, 125:15021-15027.
37. Phan AT, Modi YS, Patel DJ: Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. J Am Chem Soc 2004, 126:8710-8716.
38. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D: Human telomeric sequence forms a hybrid type-1 intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solutions. Nucleic Acids Res 2006, 34:2723-2735.
39. Balagurumoorthy P, Brahmachari SK: Intramolecular G-quadruplex DNA assemblies: length, cation identity, and multimer formation. J Am Chem Soc 2008, 130:10202-10204.
40. Mergny JL, Li J, Lacroix L, Amrane S, Chaires JB: Thermal difference spectra: a specific signature for nucleic acid structures. Nucleic Acids Res 2005, 33:e138.
41. Smargiasso N, Rosu F, Hsia W, Colson P, Baker ES, Bowers MT, De Pauw E, Gabelica V: G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation. J Am Chem Soc 2008, 130:10202-10204.
42. Mergny JL, Phan AT, Lacroix L: Following G-quartet formation by UV-spectroscopy. FEBS Lett 1998, 435:74-78.
43. Bugaut A, Balasubramanian S: A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. Biochemistry 2008, 47:689-697.
44. Hazel P, Huppert J, Balasubramanian S, Neidle S: Loop-length-dependent folding of G-quadruplexes. J Am Chem Soc 2004, 126:16405-16415.
45. Wlodarczyk A, Grzybowsk P, Patkowski A, Dobek A: Effect of ions on the Polymorphism, Effective Charge, and Stability of Human Telomeric DNA. Photon Correlation Spectroscopy and Circular Dichroism Studies. J Phys Chem B 2005, 109:3594-3605.
46. Ritsan A, Fox KR: Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. Biochemistry 2003, 42:6507-6513.
47. Ross WS, Hardin CC: Ion-Induced Stabilization of the G-DNA Quadruplex – Free-Energy Perturbation Studies. J Am Chem Soc 1994, 116:6070-6080.
48. Qin Y, Rezler EM, Gokhale V, Sun D, Hurley LH: Characterization of the G-quadruplex structure in the duplex nuclear hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4. Nucleic Acids Res 2007, 35:7698-7713.
49. Bejugm M, Sewitz S, Shirude PS, Rodrigues R, Shahid R, Balasubramaniam S: Trisubstituted Isoalloxazines as a New Class of G-Quadruplex Binding Ligands: Small Molecule Regulation of c-kit Oncogene Expression. J Am Chem Soc 2007, 129:12926-12927.
50. Zhou JM, Zhu XF, Lu Y, Deng R, Huang ZS, Mei YP, Wang Y, Huang WH, Hu ZC, Gu QL, et al.: Sequence-specific base-pairing induced by novel potent G-quadruplex interactive agents, quinodine derivatives, in human cancer cell lines. Oncogene 2006, 25:503-511.
51. Tahara H, Shin-Ya K, Seimiya H, Yamada H, Tsurow T, Ide T: G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3’ telomeric overhang in cancer cells. Oncogene 2006, 25:1953-1966.
52. Phatak P, Cookson JC, Dai F, Smith V, Gartenhaus RB, Stevens MF, Burger AM: Telomerase activity overhang and telomeric sequence containing guanine-guanine base pairs. Proc Natl Acad Sci USA 2002, 99:10584-10589.
53. Tahara H, Shin-Ya K, Seimiya H, Yamada H, Tsurow T, Ide T: G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3’ telomeric overhang in cancer cells. Oncogene 2006, 25:1953-1966.
54. Phatak P, Cookson JC, Dai F, Smith V, Gartenhaus RB, Stevens MF, Burger AM: Telomerase activity overhang and telomeric sequence containing guanine-guanine base pairs. Proc Natl Acad Sci USA 2002, 99:10584-10589.
66. Zaug AJ, Podell ER, Cech TR: Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. Proc Natl Acad Sci USA 2005, 102:10864-10869.
67. Nakagama H, Higuchi K, Tanaka E, Tsuchiya N, Nakashima K, Katahira M, Fukuda H: Molecular mechanisms for maintenance of G-rich short tandem repeats capable of adopting G4 DNA structures. Mutat Res 2006, 598:120-131.
68. Sun H, Bennett RJ, Maiizels N: The Saccharomyces cerevisiae Sgs1 helicase efficiently unwinds G-G paired DNAs. Nucleic Acids Res 1999, 27:1978-1984.
69. Sun H, Karow JK, Hickson ID, Maiizels N: The Bloom's syndrome helicase unwinds G4 DNA. J Biol Chem 1998, 273:27587-27592.
70. Hershman SG, Chen Q, Lee JY, Kozak ML, Yue P, Wang LS, Johnson FB: Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae. Nucleic Acids Res 2008, 36:144-156.
71. Huppert JL: Hunting G-quadruplexes. Biochimie 2008, 90:1140-1148.
72. Killoran MP, Keck JL: Sit down, relax and unwind: structural insights into RecQ helicase mechanisms. Nucleic Acids Res 2006, 34:4098-4105.
73. Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Harb OS, et al.: PlasmoDB: a functional genomic database for malaria parasites. Nucleic Acids Res 2008, 37:D539-543.
74. Baudoin O, Teulade-Fichou MP, Vigneron JP, Lehn JM: Cyclobisintercaland Macrocycles: Synthesis and Physicochemical Properties of Macrocyclic Polyamines Containing Two Crescent-Shaped Dibenzophenanthroline Subunits. J Org Chem 1997, 62:5458-5470.
75. Piotto M, Saudak V, Sklenar V: Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR 1992, 2:661-665.
76. Sklenar V, Piotto M, Leppik R, Saudak V: Gradient-Tailored Water Suppression for H-1-N-15 Hsqc Experiments Optimized to Retain Full Sensitivity. Journal of Magnetic Resonance Series A 1993, 102:241-245.
77. Marky LA, Breslauer KJ: Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. Biopolymers 1987, 26:1601-1620.
78. Rosu F, Gabelica V, Houssier C, Colson P, Pauw ED: Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. Rapid Commun Mass Spectrom 2002, 16:1729-1736.