BJ Biomolecules



Biochem. J. (2015) **466**, 189–199 (Printed in Great Britain) doi:10.1042/BJ20140997



# Molecular mechanism of G-quadruplex unwinding helicase: sequential and repetitive unfolding of G-quadruplex by Pif1 helicase

Xi-Miao Hou*, Wen-Qiang Wu*, Xiao-Lei Duan*, Na-Nv Liu*, Hai-Hong Li*, Jing Fu*, Shuo-Xing Dou†[1], Ming Li† and Xu-Guang Xi*‡

*College of Life Sciences, Northwest A & F University, Xi'an, Shaanxi 712100, China
†Beijing National Laboratory for Condensed Matter Physics and CAS Key Laboratory of Soft Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China
‡Laboratoire de Biologie et Pharmacologie Appliquée, Ecole Normale Supérieure de Cachan, Centre National de la Recherche Scientifique, 61 Avenue du Président Wilson, 94235 Cachan, France

Recent advances in G-quadruplex (G4) studies have confirmed that G4 structures exist in living cells and may have detrimental effects on various DNA transactions. How helicases resolve G4, however, has just begun to be studied and remains largely unknown. In the present paper, we use single-molecule fluorescence assays to probe Pif1-catalysed unfolding of G4 in a DNA construct resembling an ongoing synthesis of lagging strand stalled by G4. Strikingly, Pif1 unfolds and then halts at the ss/dsDNA junction, followed by rapid reformation of G4 and 'acrobatic' re-initiation of unfolding by the same monomer. Thus, Pif1 unfolds single G4 structures repetitively. Furthermore, it is found that Pif1 unfolds G4 sequentially in two large steps. Our study has revealed that, as a stable intermediate, G-triplex (G3) plays an essential role in this process. The repetitive unfolding activity may facilitate Pif1 disrupting the continuously reforming obstructive G4 structures to rescue a stalled replication fork. The proposed mechanism for step-wise unfolding of G4 is probably applicable to other helicases that resolve G4 structures for maintaining genome stability.

Key words: G-quadruplex (G4), genomic stability, pif1 helicase, single-molecule fluorescence.

## INTRODUCTION

As a genetic material, DNA usually assumes a double-stranded B-form conformation in which two complementary strands are held together by Watson–Crick base pairs. Although the canonical double-helix structure is considered to predominate in cells, DNA can fold into various other inter- and intra-molecular secondary structures [1,2]. Guanine-rich oligodeoxynucleotides can fold into globular structures [G-quadruplex (G4)] characterized by quartets of stacked, hydrogen-bonded G residues surrounded by connecting loops that can adopt various topological arrangements [3–5]. Over the past years, the properties and structures of isolated G4 have been extensively studied by multi-disciplinary approaches from supramolecular chemistry, single-molecule manipulation and structural biology [6–10]. It is well established that G4 structures may be extremely stable or unstable, depending on factors such as the length and sequence composition of the G4 motif, the size of loops between the guanines, strand stoichiometry and alignment and nature of the binding cations [1,11].

Consistent with the proposal that G4 motifs in DNA might be biologically relevant [12], genome-wide analyses revealed that the distribution of the putative quadruplex motifs is not random, but instead, they generally cluster in specific regions of the genome such as chromosome ends, immunoglobulin switch regions [13,14], gene transcriptional regulatory regions [15] and also the promoters of certain oncogenes, such as c-MYC and BCL2 [16–18]. Recent advances in G4 studies have provided fascinating evidence that the thermodynamically stable G4 structures are formed under physiological conditions and are modulated during cell cycles in which DNA is being replicated [19,20]. The essential function of G4 unfolding in rescuing the genome from the negative effects of G4 has been demonstrated by the observation that in the *Saccharomyces cerevisiae* Pif1 deficient strain (*pif1*Δ), failure to unravel G4 leads to replication-fork impairment, unusual epigenetic silencing and gross chromosomal rearrangement [21,22]. These studies are also helpful for understanding the phenomenon that telomere lagging-strand synthesis is severely compromised in cells of a human patient suffering from Werner's syndrome in which WRN helicase possesses a strong G4 unfolding activity [23,24].

Understanding the molecular mechanism by which helicases unfold G4 is an issue of fundamental interest and medical importance. In contrast with duplex DNA unwinding, the detailed molecular mechanism for G4 recognition and unfolding by helicases is largely unknown and many fundamental questions remain to be answered. Among the various helicases that can resolve G4 structures *in vivo* and *in vitro*, including ATP-dependent DExH/D family of RNA helicases [25], RecQ family [26] and Pif family helicases [27], the *S. cerevisiae* Pif (Pif1) stands out for its potent unwinding activity [22,28].

In a recent study, Zhou et al. [29] have characterized general unwinding properties of Pif1 with different nucleic acid substrates including a G4 DNA. However, G4 structures may have different structures and conformations, such as more layers of G-tetrads and G4 linked with duplex DNA in the cellular environment. In the present work, we therefore focused on molecular mechanism of Pif1-mediated G4 DNA unwinding using different G4 constructions which mimic an ongoing synthesis of lagging-strand stalled by G4. Our studies revealed that Pif1 unfolds G4 in a sequential and repetitive manner that is linked strictly with the intrinsic properties of G4 and Pif1. The proposed mechanism for G4 unfolding by Pif1 is probably applicable to other helicases that resolve G4 structures for maintaining genome stability.

Abbreviations: C, G-column; CD, circular dichroism; Cy3/5, cyanine probe; G3, G-triplex; G4, G-quadruplex; L, lateral loop.

[1] To whom correspondence should be addressed (email xxi01@ens-cachan.fr or sxdou@iphy.ac.cn).
Xi-Miao Hou and Wen-Qiang Wu contributed equally to this work.

**Table 1 Sequences of substrates used in the experiments**

| | Substrates for single-molecule FRET |
|---|---|
| G4 strand | |
| 3L–G4*–T | AAGCAGTGGTATCAACGCAGAGAAAT(iCy3)GGGTTAGGGTTAGGGTTAGGGATGTATGACAAGGAAGG |
| 4L–G4*–T | AAGCAGTGGTATCAACGCAGAGAAAT(iCy3)GGGGTAGGGGTAGGGGTAGGGGATGTATGACAAGGAAGG |
| 5L–G4*–T | AAGCAGTGGTATCAACGCAGAGAAAT(iCy3)GGGGGTAGGGGGTAGGGGGTAGGGGGATGTATGTCAAGGAAGG |
| 3L–G4–T | AAGCAGTGGTATCAACGCAGAGAAATGGGTTAGGGTTAGGGTTAGGGATGTATGACAAGGAAGG |
| 4L–G4–T | AAGCAGTGGTATCAACGCAGAGAAATGGGGTAGGGGTAGGGGTAGGGGATGTATGACAAGGAAGG |
| 5L–G4–T | AAGCAGTGGTATCAACGCAGAGAAATGGGGGTAGGGGGTAGGGGGTAGGGGGATGTATGTCAAGGAAGG |
| 3L–G4* | (Cy3)GGGTTAGGGTTAGGGTTAGGGATGTATGACAAGGAAGG |
| 3L–G4 | GGGTTAGGGTTAGGGTTAGGGATGTATGACAAGGAAGG |
| 3L–G3–T | AAGCAGTGGTATCAACGCAGAGAAATGGGTTAGGGTTAGGGTTTTTTATGTATGACAAGGAAGG |
| 3L–G3 | GGGTTAGGGTTAGGGTTTTTTATGTATGACAAGGAAGG |
| Complementary strand | Biotin–CCTTCCTTGTCAT(iCy5)ACAT |
| | Substrates for equilibrium DNA binding |
| G4 | TGGGTTAGGGTTAGGGTTAGGG(Fluorescein) |
| ssDNA | CTCTGCTCGACGGATT(Fluorescein) |
| dsDNA | CTCTGCTCGACGGATT(Fluorescein) |
| G3 | GAGACGAGCTGCCTAA |
| Fork DNA | GGGTTAGGGTTAGGG(FAM) |
| | TTTTTTCGTCGAGCAGAG(Fluorescein) |
| | GCAGCTCGTCTC |
| 3L–G4–T | GCAGCTCGTCTC AAGCAGTGGTATCAACGCAGAGAAATGGGTTAGGGTTAGGGTTAGGGATGTATGACAAGGAAGG (Fluorescein) |
| | Substrates for CD |
| G4 | AGGGTTAGGGTTAGGGTTAGGG |
| G3 | GGGTTAGGGTTAGGG |

## EXPERIMENTAL

### Buffers

Pif1 reaction buffer contains 5 mM $MgCl_2$, 50 mM NaCl in 25 mM Tris/HCl, pH 7.5, with 2 mM DTT. For single-molecule measurements, 0.8 % D-glucose, 1 mg/ml glucose oxidase (266 600 units/g, Sigma), 0.4 mg/ml catalase (2000–5000 units/mg, Sigma) and 1 mM Trolox were added to the reaction buffer [30].

### DNA constructs

All oligonucleotides required to make the DNA substrates were purchased from Sangon Biotech. In single-molecule experiments, the DNA constructs were composed of a G4 strand and a complementary stem strand [9]. In CD and binding measurements, the substrate was an ssDNA or duplex DNA. Sequences and labelling positions of all the oligonucleotides are listed in Table 1. For DNA constructs used in single-molecule measurements, DNA was annealed with a 1:3 mixture of the stem and G4 strands by incubating the mixture at 95 °C for 5 min, then slowly cooling down to room temperature in about 7 h. The strand without biotin was used in excess to reduce the possibility of having non-annealed strand anchored at the coverslip surface. The concentration of stem strand was 2.5 $\mu$M and all annealing was carried out in annealing buffer containing 50 mM NaCl, 25 mM Tris/HCl, pH 7.5. For duplex DNA in bulk measurements, DNA was annealed with a 1:1 mixture of the two complementary strands.

### Pif1 purification

The plasmid encoding yeast Pif1 gene was a gift from Dr Zakian. The protein expression and purification were performed essentially according to the method of Boule and Zakian [31]. The purified Pif1 was >95 % pure according to analysis by SDS/PAGE (Supplementary Figure S1).

### Fluorescence labelling of Pif

Sequence alignment analyses show that among the five cysteine residues in yeast Pif1 sequence, only $Cys^{261}$ is relatively conserved and embedded into the interior of the protein. We therefore substituted the other three cysteine residues with alanine and kept $Cys^{199}$ and $Cys^{261}$ intact. The appropriate fluorophore [cyanine probe (Cy3)–maleimide, Amersham Biosciences] was then mixed with the mutant (in 10-fold molar excess) in buffer A [50 mM Tris (pH 8.0 at 4 °C), 350 mM NaCl, 25 % (v/v) glycerol] and incubated overnight at 4 °C, while slowly rotating the container [32]. Free fluorophores were removed by size exclusion chromatography with buffer A (10–15 column volumes). MS analyses confirmed that only $Cys^{199}$ was labelled (consistent with the one-step photobleaching phenomenon observed in experiments). The concentrations of the labelled protein were determined spectrophotometrically by recording spectra from 240 to 700 nm. The Cy3-labelled Pif1 mutant displayed very similar ATPase, DNA binding and unwinding activities as the wild-type protein.

### Circular dichroism spectropolarimetry

CD experiments were performed on a Bio-Logic MOS450/AF-CD optical system (BioLogic Science Instruments) equipped with a temperature-controlled cell holder, using a quartz cell with 1-mm path length. A 2.5 $\mu$M solution of G4 or G-triplex (G3) DNA was prepared in 25 mM Tris/HCl, pH 7.5, containing 50 mM NaCl, 50 mM LiCl or no salt. CD spectra were recorded in the UV (200–350 nm) region in 0.75 nm increments with an averaging time of 2 s at 25 °C.

### Single-molecule fluorescence data acquisition

Single-molecule FRET study was carried out with a home-built objective-type total-internal-reflection microscope [30]. Cy3 was excited by 532 nm Sapphire laser (Coherent Inc.). An oil

immersion objective (100×, numerical aperture 1.49) was used to generate an evanescent field of illumination. Fluorescence signal from Cy3 and Cy5 were split by a dichroic mirror and finally collected by an electron-multiplying charge-coupled device camera (iXON, Andor Technology). Fluorescence imaging process were controlled and recorded by MetaMorph (Molecular Device). The coverslips (Fisher Scientific) and slides were cleaned thoroughly by a mixture of sulfuric acid and hydrogen peroxide, acetone and sodium ethoxide, then the surfaces of coverslip were coated with a mixture of 99 % methoxy poly(ethylene glycol) (m-PEG-5000, Laysan Bio) and 1 % of biotin-PEG (biotin-PEG-5000, Laysan Bio, Inc.). Streptavidin (10 $\mu$g/ml) in buffer containing 50 mM NaCl, 25 mM Tris/HCl, pH 7.5, was added to the microfluidic chamber made of the PEG coated coverslip and incubated for 10 min. After washing, DNA was added to the chamber at a final concentration of 100 pM and immobilized for 10 min. Then free DNA was removed by washing with the reaction buffer. After that, the chamber was filled with reaction buffer with an oxygen scavenging system (0.8 % D-glucose, 1 mg/ml glucose oxidase, 0.4 mg/ml catalase and 1 mM Trolox). Imaging was initiated before Pif1 and ATP was flowed into the chamber. For the unfolding study with wild-type Pif1, 5 nM Pif1 and 25 $\mu$M ATP were flowed into the chamber simultaneously. For the unfolding study with Cy3–Pif1, we first incubated 2 nM Cy3–Pif1 with immobilized DNA for 5 min. Then, 50–100 $\mu$M ATP was flowed into the chamber to initiate the reaction. For the single-molecule binding study with Cy3–Pif1, Cy3–Pif1 and substrate 3L–G4* (Table 1) were pre-incubated for 5 min and then the unbound Cy3–Pif1 was washed away with the reaction buffer, after which FRET was measured for each molecule. We used an exposure time of 100 ms for all single-molecule measurements, carried out at a constant temperature of 22 °C.

### FRET data analyses

The raw fluorescence intensity trajectories were three-point averaged. Then the FRET efficiency was calculated using $I_A/(I_D + I_A)$, where $I_D$ and $I_A$ represent the intensity of donor and acceptor respectively. The FRET value above 1 is due to background subtraction from very low intensity in the donor channel, giving rise to negative donor intensity. Basic data analysis was carried out by scripts written in Matlab and all data fitting was performed by Origin 7.5. To create a FRET histogram, the area of FRET traces showing Pif1–G4 interaction were picked out by Matlab program and then all data from individual traces were combined. Histograms were fitted by multi-peak Guassian distributions, with the peak position unstrained. To create a dwell time histogram, $\delta t_1$ (or $\delta t_2$) was measured manually from individual FRET traces and the resulting histograms were fitted with a single exponential decay.

### Equilibrium DNA-binding assay with Pif1

Binding of Pif1 to DNA was analysed by fluorescence polarization assay [33] using Infinite F200 PRO (TECAN). DNA (ssDNA, duplex, G3 or G4) labelled only with fluorescein was used in the present study. Various amounts of protein were added to a 150 $\mu$l aliquot of binding buffer (25 mM Tris/HCl, pH 7.5, 50 mM sodium chloride, 2 mM magnesium acetate and 2 mM DTT) containing 2 nM DNA. Each sample was allowed to equilibrate in solution for 5 min, after which fluorescence polarization was measured. A second reading was taken after 10 min, in order to ensure that the mixture was well equilibrated. Less than 5 % change was observed between the 5- and 10-min measurements, indicating that equilibrium was reached in 5 min. The equilibrium dissociation constant was determined by fitting the data to the Michaelis–Menten equation using KaleidaGraph (Synergy Software).

### Equilibrium DNA-binding assay with G4 antibody

The plasmid encoding gene of G4 antibody was a gift from Dr Balasubramanian. The protein expression and purification were performed essentially according to the method of Biffi et al. [20]. Binding of G4 antibody to DNA was analysed by the fluorescence polarization assay described above. Various amounts of protein were added to a 150 $\mu$l aliquot of binding buffer (25 mM Tris/HCl, pH 7.5, 50 mM sodium chloride, 5 mM magnesium chloride and 2 mM DTT) containing 5 nM DNA. Each sample was allowed to equilibrate in solution for 5 min, after which fluorescence polarization was measured.

## RESULTS

### Design and preparation of unimolecular G4 substrates for FRET study

To explore the molecular mechanism by which Pif1 family helicases unfold G4, we used single-molecule FRET assays [34] and a DNA constructed with an unimolecular G4 adjacent to a DNA duplex (Figure 1A), which resembles an ongoing synthesis of lagging strand stalled by G4, thus mimicking Pif1-mediated G4 unfolding events in cells [35–37] (Figure 1B). The substrate was constructed with an ssDNA sequence containing donor fluorophore (Cy3) attached to the junction between the G4 motif and 5′ tail, and its 3′ tail was hybridized with a complementary stem strand modified at the 5′-end by biotin for immobilization and attached by an acceptor fluorophore (Cy5) at the fifth nucleotide from the 3′ end (Table 1). The fluorophores are so spaced that the FRET signal may report sensitively the conformational change of G4.

To confirm that the designed G4 DNA was well formed as we anticipated, we performed three control experiments. First, we measured CD spectra of the oligo that harbours G4-forming sequence in buffers containing 50 mM NaCl, 50 mM LiCl and no salt respectively. Whereas the CD spectrum in the first case shows a positive band around 290 nm and a negative band around 263 nm, characterizing the typical CD spectrum of antiparallel G4 [38], that in the other two cases exhibit significantly different bands, representing CD spectra of no structured oligonuclotide (Figure 1C). Secondly, being aware that the ss/dsDNA flanking the G4 sequence may influence the correct formation of the expected G4 structure, we then measured directly CD spectra with the constructed G4 DNA 3L–G4–T (Table 1) which was used for FRET study. The resulting CD spectra are not interpretable due to the presence of ss/dsDNA sequences at both sides of G4, which interfere strongly with the CD signals. We then determined the G4-binding curve with an engineered, structure-specific antibody that binds with high selectivity to G4 [20]. The results show that although both isolated G4 and 3L–G4–T bind equally well to the antibody, ssDNA and dsDNA fail to bind (Figure 1D), indicating that both the isolated G4 sequence and that in 3L–G4–T are folded correctly to G4 structures as expected. Finally, we analysed FRET spectra of the labelled G4 under conditions where G4 formation is favourable and unfavourable respectively. Under conditions that favour G4 formation (50 mM NaCl), we observed exclusive stable populations with high FRET value (0.937; Figure 1E), corresponding to the well-formed G4 structure due to the proximity of the two fluorophores. However, under

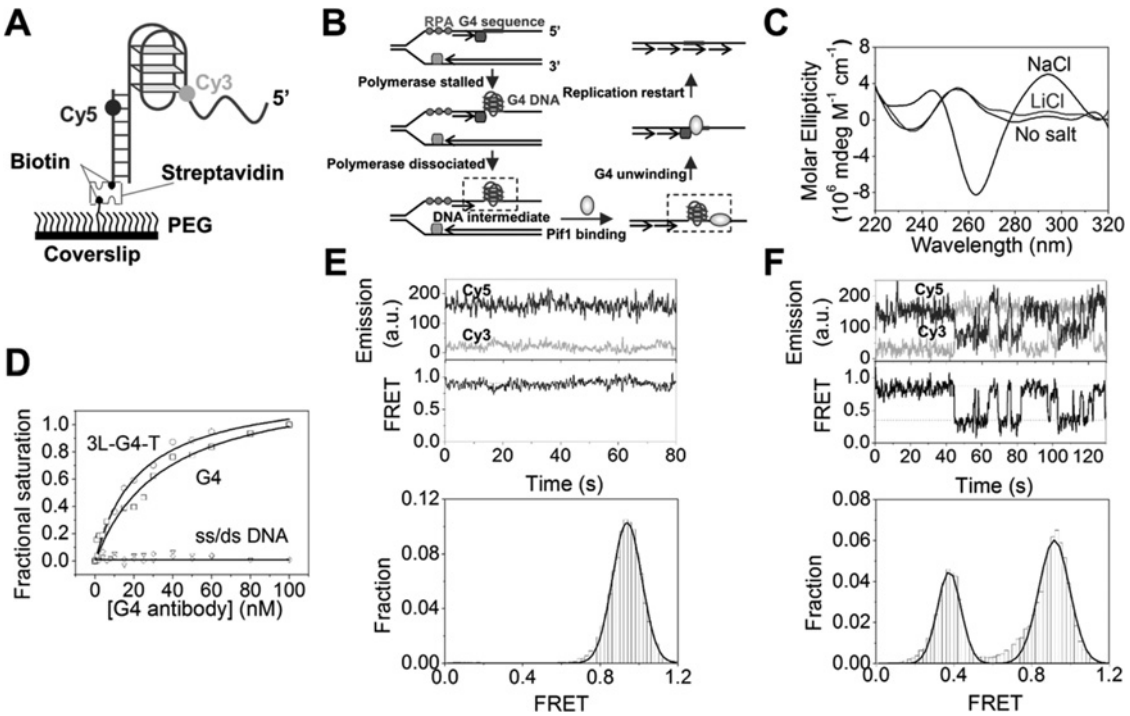


**Figure 1 DNA substrates used in the present study are biologically relevant and the G4 structure is well folded in 50 mM NaCl**

(**A**) Schematic diagram of Cy3- and Cy5-labelled DNA construct with G4 having three G-quartet planes. The G4 strand and the complementary stem strand are annealed to form a duplex stem. Biotin is used to immobilize the DNA to a streptavidin-coated coverslip surface. (**B**) Schematic diagram illustrating how G4 is formed from regions bearing a G4 motif during DNA replication. (**C**) CD spectrum from G4 in 50 mM NaCl has a peak around 290 nm and a valley around 263 nm, whereas in 50 mM LiCl or buffer without salt, no G4 structure has been formed. (**D**) Detection of G4 structures in different DNA substrates. The antibody was able to bind to isolated G4 and also G4 sequence with 3′ and 5′ tails with $K_d$ values of 35.4 and 22.5 nM respectively, suggesting G4 structure can also form in the presence of proximal overhangs. The antibody did not bind to ss/dsDNA. (**E**) Time trace of Cy3 and Cy5 emissions for 3L–G4*–T and the corresponding time trace of FRET in 50 mM NaCl buffer (upper panel). Gaussian fit of the FRET histogram (lower panel) gives one peak at $0.937 \pm 0.001$, corresponding to completely folded G4. (**F**) Time trace of Cy3 and Cy5 emissions for 3L–G4*–T and the corresponding time trace of FRET in 20 mM NaCl buffer (upper panel). Gaussian fit of the FRET histogram (lower panel) gives two peaks at $0.918 \pm 0.002$ and $0.376 \pm 0.002$, corresponding to completely folded and unfolded G4 respectively.

conditions that do not favour G4 formation (20 mM NaCl), we detected two conformations with high (0.918) and low (0.376) FRET values (Figure 1F), corresponding to the well-folded G4 and unfolded G4 respectively. Altogether, these experiments confirm that the single-molecule substrate was folded into G4 structure as we expected and suitable for FRET study.

### Repetitive unfolding and refolding of G4 catalysed by Pif1

We first analysed FRET trajectories at different concentration of Pif1 to determine the condition under which the FRET signals could be recorded with lowest Pif1 concentration. We found that 5–10 nM Pif1 was suitable to observe Pif1-mediated G4 unwinding under single-molecule conditions. Therefore, unless specified otherwise, all experiments were performed with 5 nM Pif1. Having established the optimal experimental conditions with the designed G4, we began our FRET experiments on three types of G4 substrates in which the G-tetrad layers were increased from three to five (Figure 2A).

Upon addition of 5 nM Pif1 and 25 $\mu$M ATP, FRET oscillation was observed that sometimes may last for more than 1 min (Figure 2B, top panel). This is surprising because Pif1 was expected to unfold firstly the G4 structure, then the duplex and finally the unwound strand will escape from the coverslip surface, which should be accompanied by abrupt disappearance of FRET signals. To better understand the observed oscillation phenomenon, we first confirmed that there was no such oscillation in the absence of either ATP or Pif1 or in the presence of Pif1 and adenosine 5′-[$\gamma$-thio] triphosphate indicating that the conformational changes of G4 indeed required ATP hydrolysis of Pif1. We then wondered whether the repetitive unfolding/refolding was caused by complete dissociation and rebinding of Pif1. By inspection of a wide range of data, we found that in some cases, the oscillations with low frequencies were well separated by a long waiting time ($\sim$50 s) (Supplementary Figure S2A). Such single-molecule FRET traces can be best interpreted by the fact that one enzyme binds and unfolds G4 with oscillation, then dissociates from G4 and another enzyme binds to the refolded G4 and begins another oscillation cycle. However, the above mentioned FRET trace is in sharp contrast with the FRET trace presented in Figure 2B in which the oscillations continue without a long waiting time, indicating that the observed oscillations were mediated continuously by the same Pif1 molecule.

To further confirm that the recorded FRET signals came from G4 unfolding that was catalysed by a single Pif1 molecule, we also pre-incubated Pif1 with immobilized G4 for 5 min and then the unbound proteins were washed out using 10 volumes of imaging buffer, followed by addition of 25 $\mu$M ATP. We found that the protocol gave similar results (Supplementary Figure S2B). As the buffer was devoid of free Pif1, the repetitive FRET oscillation should be induced by the same Pif1, rather than the

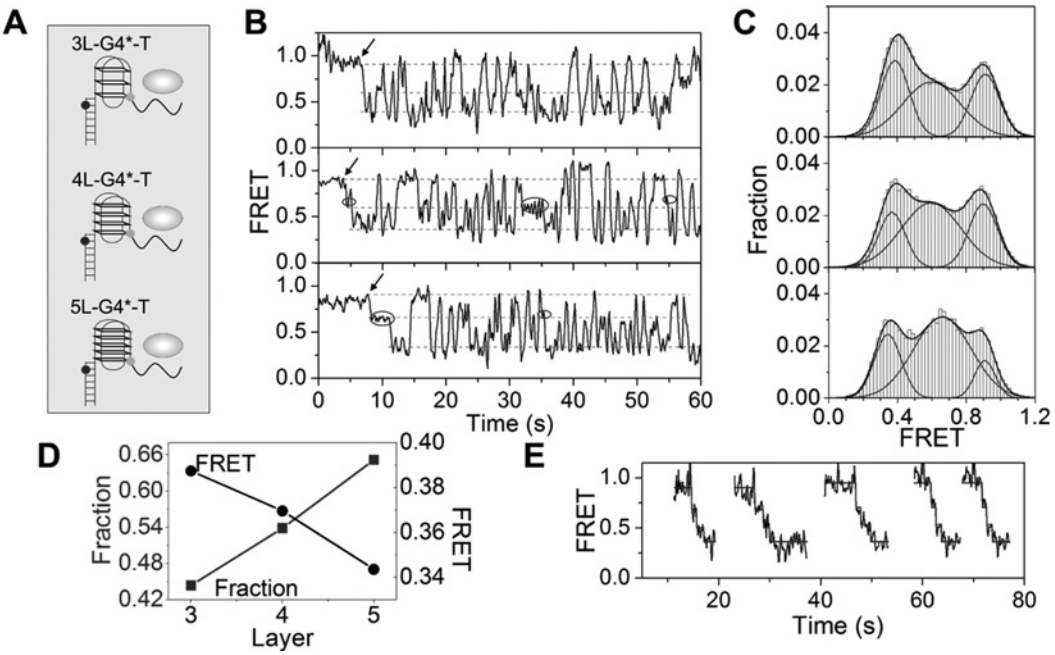


**Figure 2 Repetitive conformational changes of G4 induced by Pif1**

(**A**) Cy3- and Cy5-labelled DNA constructs with G4 of 3–5 layers. The lengths of the 5′ ssDNA tail and the duplex are 26 nt and 17 bp respectively. (**B**) FRET traces for the three DNA molecules obtained with 25 $\mu$M ATP and 5 nM Pif1 at 22 °C. The time when Pif1 and ATP were added is indicated by an arrow. Pauses of unfolding are labelled with black circles. (**C**) Histograms of FRET constructed from the oscillating parts of more than 100 single-molecule time traces. For each distribution, Gaussian fitting (thick black line) reveals three peaks with low, medium and high FRET. The three peak positions are 0.387 $\pm$ 0.002, 0.599 $\pm$ 0.014 and 0.912 $\pm$ 0.002 (3L–G4*–T, top); 0.370 $\pm$ 0.002, 0.598 $\pm$ 0.009 and 0.900 $\pm$ 0.002 (4L–G4*–T, middle) and 0.344 $\pm$ 0.003, 0.661 $\pm$ 0.005 and 0.908 $\pm$ 0.003 (5L–G4*–T, bottom) respectively. They are shown as dashed lines in the corresponding Figures in (**B**). (**D**) Effects of the number of G-quartet layers on FRET distribution. FRET value of completely unfolded G4, i.e. peak position of the low-FRET conformations and the population of intermediate conformations as a function of the number of G-quartet layers, obtained from (**C**). (**E**) Three steps may occasionally be observed during unfolding of 3L–G4*–T. The individual traces are shifted horizontally for clarity.

binding and release of different Pif1 proteins. Finally, it is still possible that the unexpected phenomenon results from the steric hindrance between the labelled G4 and the coverslip surface. To increase the steric distance between G4 and the coverslip surface, we prepared a modified G4 DNA in which the length of the biotin-labelled ssDNA was extended to 50 nt, thus the distance between G4 and the coverslip surface was increased significantly. In addition, to mitigate the steric interference between different substrates, the G4 substrate concentration was reduced to 10 pM. Experiments performed with the modified G4 substrates or reduced G4 concentration show that the previous observed oscillation phenomenon still persists (results not shown). These results altogether indicate that the observed oscillation phenomenon is an intrinsic property of Pif1-mediated G4 unwinding, but not an experimental artifact.

### Pif1 unfolds G4 DNA in discrete steps

To reveal the physical meaning of the oscillation, a FRET histogram was built from >100 single-molecule FRET trajectories. It consists of three Gaussian peaks at ~0.4, ~0.6 and ~0.9 respectively (Figure 2C, top panel). The low and high FRET peaks correspond respectively to unfolded and completely-folded conformations, as confirmed by measuring FRET under experimental conditions that disfavour formation of stable G4, as shown in Figure 1(F). The medium FRET peak might be attributed to unfolding/refolding intermediate conformation(s), most probably G3 structures which have been shown to be also mechanically and energetically stable structures [7,39–43]. To verify this conjecture, we carried out similar analysis with four- and five-layered G4. We expected that, if the intermediate conformations were indeed G3, they could be more easily trapped because they should be energetically more stable with more layers, just like G4 [44,45].

As expected, the population of intermediate conformations indeed increases with increasing number of tetrad layers (Figures 2C and 2D). Furthermore, even obvious unfolding pauses (FRET at ~0.6) that may last for several seconds appear frequently (Figure 2B, middle and bottom panels, black circles), indicating that Pif1 unfolds G4 in two large steps with a G3 intermediate. From Figure 2(B), it is clear that another important contribution to the population of intermediate conformations comes from the oscillation peaks of medium heights. These peaks may suggest that the refolding intermediate is unfolded again by Pif1 before further folding into G4. Consistent with the results from Zhou et al. [29], three steps may occasionally be observed during the G4 unfolding by Pif1 (Figure 2E). However, in most cases, only two steps can be distinguished in our experiments.

### G4 unfolding with Cy3-labelled Pif1

In the previous experiments, the FRET signal only revealed the change of DNA conformation but not that of Pif1 position. To follow Pif1 translocation on G4, we engineered a Pif1 mutant with Cy3 conjugated at amino acid position 199, Cy3–Pif1, and used the same DNA constructs as before, but without Cy3 labelling (Figure 3A). FRET occurs between two appropriately positioned fluorophores only when the distance separating them is 8–10 nm or less. Thus, FRET signals can only be observed when Cy3-labelled Pif1 binds to Cy5-labelled G4. In this case, the FRET

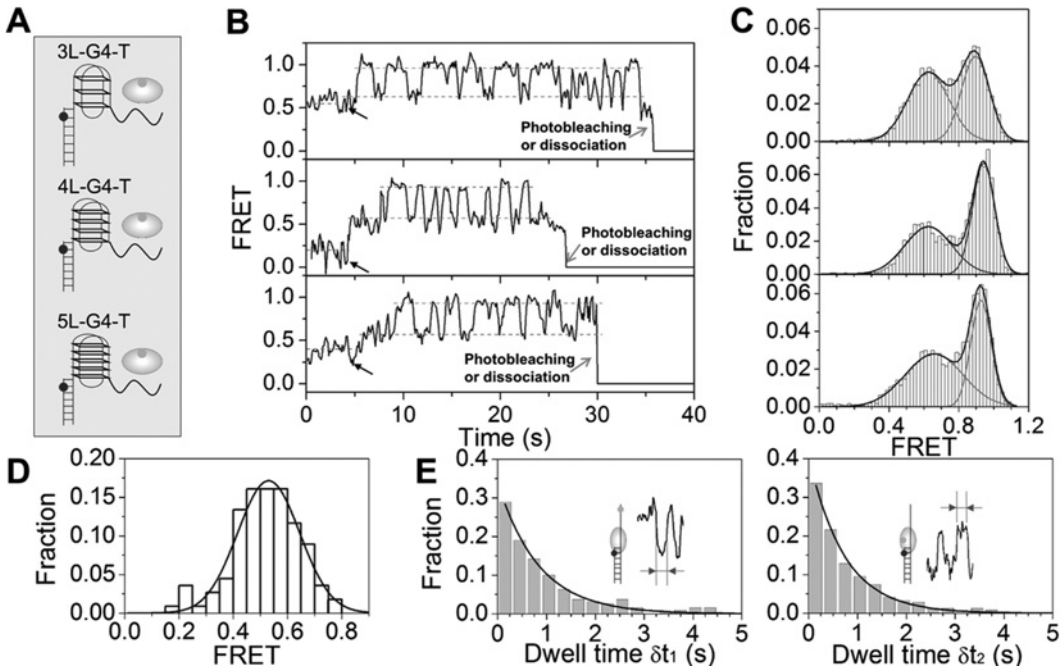


**Figure 3 Repetitive unfolding of G4 by Cy3-labelled Pif1 mutant**

(**A**) Schematic diagrams of Cy5-labelled DNA constructs with G4 having 3–5 G-quartet planes. (**B**) Time traces of FRET for the three DNA molecules obtained with 50–100 $\mu$M ATP and 2 nM Pif1. The time when ATP was added is indicated by a black arrow. As in all the following Figures, the abrupt drop of FRET as caused by photobleaching of Cy3 or dissociation of Cy3–Pif1 is indicated by a grey arrow. (**C**) Histograms of FRET constructed from the oscillating parts of 39–57 single-molecule time traces. For each distribution, Gaussian fitting (thick black line) reveals two peaks with medium and high FRET. The two peak positions are $0.625 \pm 0.005$ and $0.895 \pm 0.003$ (3L–G4–T, top panel); $0.626 \pm 0.007$ and $0.943 \pm 0.002$ (4L–G4–T, middle panel) and $0.658 \pm 0.008$ and $0.927 \pm 0.001$ (5L–G4–T, bottom panel) respectively. (**D**) Histogram of FRET [indicated by a dashed line in (**B**) before ATP was added] for initial binding of Cy3–Pif1 to G4 (3L–G4–T) with a Gaussian fit (solid line) having a peak at $0.530 \pm 0.007$. (**E**) Dwell time histograms for the ssDNA remaining in unfolded and partially folded conformations ($\delta t_1$, obtained from FRET traces of 3L–G4*–T unfolding by Pif1) and for Pif1 remaining stalled at the ss/dsDNA junction ($\delta t_2$, obtained from FRET traces of the translocation of Cy3–Pif1 on 3L–G4–T) during repetitive unfolding of three-layered G4. The solid lines are single exponential decays with time constants of $0.84 \pm 0.03$ and $0.73 \pm 0.03$ s respectively.

signal may yield information on both the initial binding position of Pif1 and its position change during G4 unfolding.

Since Pif1 binds well to ssDNA or G4 DNA in the absence of ATP (Supplementary Figures S3A and S3D), Cy3–Pif1 was first flowed into the chamber and then followed by ATP some time later. The initial FRET values determined from individual FRET traces (Figure 3B, top panel), which indicate the binding position of Cy3–Pif1 on G4, were used to construct FRET histograms (Figure 3D). We noticed that the initial FRET signal before ATP addition varied over a wide range from 0.2 to 0.8, indicating that Pif1 was initially bound at different positions on the 5′ tail (from the far 5′- end to the junction) and on G4. This is consistent with the fact that Pif1 displays similar affinity for G4 and ssDNA, but much lower affinity for dsDNA, as revealed by our binding assay (Supplementary Figures S3A, S3B and S3D). As before, FRET oscillation was observed after adding ATP (Figure 3B, top panel). The FRET histograms, however, consist of only two Gaussian peaks at 0.63 and 0.90 respectively (Figure 3C, top panel). The two peaks should correspond to two extreme positions of Pif1 relative to Cy5 during repetitive unfolding. According to the previous experimental results shown in Figure 2, the 0.90 peak should correspond to Pif1 halting at the ss/dsDNA junction after unfolding is completed and the 0.63 peak to Pif1 binding to refolded G3 or G4.

To get more insight into the repetitive unfolding activity of Pif1, we measured the dwell time for the ssDNA remaining in unfolded and partially folded conformations, $\delta t_1$, and that for Pif1 remaining stalled at the ss/dsDNA junction, $\delta t_2$, during repetitive unfolding (Figure 3E). Interestingly, the histograms can be well fitted by a single exponential decay with similar time constants (0.84 and 0.73 s), implying that the ssDNA is unable to refold completely into G4 when Pif1 is halting at the junction.

We also carried out similar studies with four- and five-layered G4 (Figure 3B, middle and bottom panels). The resulting FRET histograms also consist of two Gaussian peaks with positions at 0.63 and 0.94 (Figure3C, middle panel), 0.66 and 0.93 (Figure 3C, bottom panel) respectively. Like three-layered G4, the 0.94/0.93 peak should correspond to Pif1 halting at the ss/dsDNA junction after unfolding is completed and the 0.63/0.66 peak to Pif1 binding to refolded G3 or G4. Those results further support that Pif1 repetitively shuttles between the ss/dsDNA junction and refolded G3/G4 structures.

### Repetitive unfolding is regulated by ATP concentration

ATP provides energy for helicase translocation and modulates DNA-binding through allosteric regulation. We then examined the effects of ATP concentration on the G4 unfolding process by varying ATP concentration from 25 $\mu$M to 3 mM. We found that Pif1-mediated repetitive folding/unfolding and its translocation were highly regulated by ATP concentration. A representative trace of Pif1-induced unfolding of three-layered G4 at 2 mM ATP shows that G4 was unfolded at first and then was kept in this unfolded state and seldom folded back (Figure 4A). This is radically different from the results with low concentration of ATP (25 $\mu$M), in which G4 goes through repetitive unfolding/folding processes (Figure 2B, top panel). Meanwhile, the Cy3–Pif1

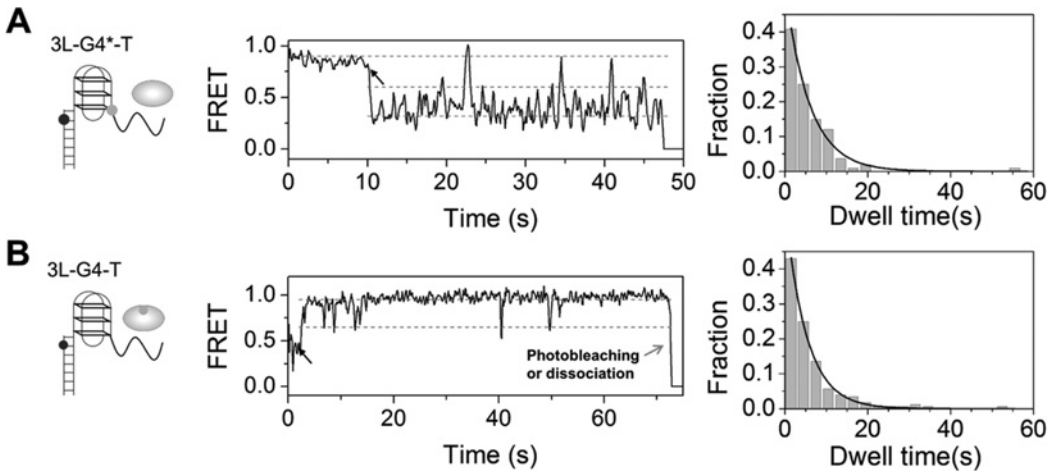


**Figure 4 Pif1 unfolds G4 in the presence of 2 mM ATP**

(**A**) Single-molecule trace was obtained in the presence of 5 nM Pif1 at 22 °C. The G4 structure no longer shows frequent folding/unfolding behaviour, but stays in the unfolded state for most of the time. Dwell time histogram (right panel) for the ssDNA remaining in unfolded and partially folded conformations ($\delta t_1$) follows an exponential decay with a time constant of $5.88 \pm 0.26$ s. (**B**) 2 nM Cy3–Pif1 were incubated with Cy5-labelled DNA substrate containing a three-layer G4 for 5 min. Then 2 mM ATP was added to initiate the unfolding. The representative FRET trace shows that Pif1 stays at the ss/dsDNA junction and only moves occasionally. Dwell time (right panel) of Pif1 remaining stalled at the ss/dsDNA junction ($\delta t_2$) follows an exponential decay with a time constant of $5.03 \pm 0.15$ s.

translocation trajectory further confirms that Pif1 moves towards the junction and stays there stably, with occasional hopping away (Figure 4B). We also measured the duration times similarly as in Figure 3(E). They follow single exponential decays, with decay constants of 5.9 and 5.0 s respectively (Figures 4A and 4B, right panels). Compared with the time constant of 0.84 and 0.73 s at the low concentration of ATP, we think high-concentration ATP can stabilize Pif1 at the 5′ ss/dsDNA junction, so G4 can be kept in an unfolded state for several seconds. Together, these results suggest that ATP concentration is important to regulate the Pif1-mediated G4 unfolding process.

### Unfolding of G4 without a 5′ tail

Previous single-molecule studies have revealed a repetitive shuttling of Rep helicase on DNA [46]. In that case, the helicase stalled at an ss/dsDNA junction will grab the far end of the ssDNA tail and initiate a new translocation cycle. This means that an ssDNA tail is necessary for the observed repetitive shuttling phenomenon with helicase unwinding duplex DNA. To verify whether the ssDNA tail is absolutely required for the oscillation phenomenon observed in our study, we used new DNA constructs the same as that in Figures 2 and 3, except that the 5′ tail was absent. As expected, similar repetitive unfolding was still observed (Figures 5A and 5B), indicating that a loading tail is not a pre-requisite for Pif1 to initiate G4 unfolding, as generally believed, and not used at all for repetitive unfolding. From binding measurement (i.e. in the absence of ATP) with Cy3–Pif1 and tail-less G4, we found that the FRET signal has a Gaussian distribution peaking at 0.62 (Figure 5C), which is quite similar to the low FRET peak values (0.63–0.66) in the above unfolding experiments (Figure 3C). This further shows that, during repetitive unfolding, Pif1 binds to refolded G3 or G4 rather than the ssDNA tail.

### Pif1 unfolds G4 as a monomer

There has been much debate as to whether a monomer or oligomers are responsible for DNA unwinding. It was well established that the number of associated subunits in a complex can be deduced by imaging single molecules and counting photobleaching steps of fluorophore-labelled proteins [47]. We took advantage of the fact that several experiments were performed with Cy3-labelled Pif1 under different conditions to determine the oligomeric state of Pif1 during G4 unwinding. By analysing the data shown in Figures 3B, 4B and 5B and in the Supplementary Figure S4, we concluded that the repetitive unfolding activity is achieved by a Pif1 monomer, because photobleaching of Cy3 or dissociation of Pif1 occurs only in a single step, but not in multiple steps, which is a typical feature of single-subunit-catalysed events. Now we might understand why Pif1 is stalled at the ss/dsDNA junction; a Pif1 monomer can unfold G4 but not a duplex. In fact, previous studies have shown that Pif1 unwinds partial duplex DNA with very low efficiency [31,48] and unwinds duplex as dimers [49]. We also observed in stopped-flow assays that Pif1 unwinds duplex DNA efficiently only at high concentrations (Supplementary Figure S5). This character of Pif1 is consistent with the previous observations and its G4-structure-resolving function in cells [22].

### Molecular mechanism for Pif1-catalysed repetitive unfolding of G4

Based on all of the above-mentioned experimental observations, we propose the following mechanism for sequential and repetitive unfolding of G4 by Pif1. The Pif1 monomer binds initially to the 5′ tail (Figure 6A). Upon ATP addition, Pif1 translocates rapidly to a G4-binding site, which includes probably the first G-column (C1) and the first lateral loop (L1) of G4 (Figure 6A', for clarity, Pif1 is shown to bind only a G-column in the Figure). Of course, Pif1 may alternatively bind directly to this position even at the very beginning.

From the position in Figure 6A', Pif1 moves forward, strips C1 and leaves a G3 as an on-pathway intermediate (Figure 6B), thus completing one large step in unfolding G4. Then Pif1 strips C2 in a similar way and completes disruption of the whole structure in another large step, followed by moving forward along the unfolded ssDNA until arriving at the ss/dsDNA junction (Figure 6C). Whereas Pif1 remains attached at the junction, the

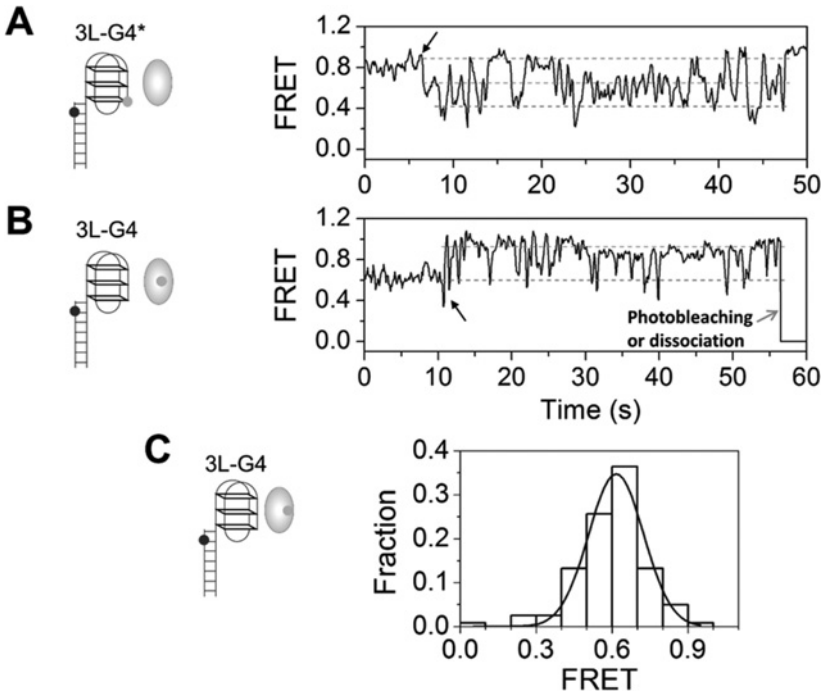


**Figure 5 The 5′ ssDNA tail of G4 is not required for repetitive unfolding**

(**A**) Unfolding of tailless G4 by Pif1 with a Cy3- and Cy5-labelled DNA construct (3L–G4*, left panel) and time trace of FRET (right panel). The arrow indicates the time when 5 nM Pif1 and 25 $\mu$M ATP were added. (**B**) Unfolding of tailless G4 by Cy3–Pif1 with a Cy5-labelled DNA construct (3L–G4, left panel) and time trace of FRET (right panel). 2 nM Cy3–Pif1 was incubated with DNA and then 50 $\mu$M ATP was added (black arrow). (**C**) FRET measurements for Cy3–Pif1 binding to tailless G4 (3L–G4, left panel) as described in 'Experimental'. FRET histogram (right panel, 121 molecules) with a Gaussian fit (solid line) having a peak at $0.616 \pm 0.008$.

rest of the unfolded G4 motif refolds spontaneously into a new G3 (Figure 6D).

From this stage, further unfolding/folding may occur in two different cycles. Once G3 is reformed quickly from the unwound sequence, Pif1 dissociates immediately from the ss/dsDNA junction and rebinds to G3. This hopping process is feasible because the refolded G3 may, driven by thermal fluctuations, swing rapidly around Pif1 and, furthermore, Pif1 has a much higher affinity for G3 than ss/dsDNA (Supplementary Figures S3C and S3E). If Pif1 binds to G3 through C1 and L1 (Figure 6E), G3 will instantaneously fold into G4 with the last run of GGGTTA repeat that is just released (Figure 6A'), completing an unfolding/refolding cycle (Cycle I).

Alternatively, Pif1 may bind to G3 through C2/ L2 (Figure 6E') or C3/L3 (result not shown). In this case, G3 is unable to further fold into G4 because Pif1 may occupy partially the position originally left for G-column C4. On the other hand, Pif1 will initiate unfolding of the G3 by striping C2 (or C3) and disrupt the whole G3 structure, finally halting at the ss/dsDNA junction again (Figure 6C) and allowing ssDNA to refold into G3 (Figure 6D), thus completing a small G3 unfolding/folding cycle (Cycle II).

It should be noted that Pif1 may initially bind to one of the other three G-columns (C2–4) of G4 at the very beginning. If Pif1 binds initially to C2 (or C3), we think it cannot unfold G4 because the G-strand is not releasable from the G4 structure due to the topological constraint imposed by the two lateral loops at both ends of the G-strand. If binding initially to C4, Pif1 can only unfold G4 initially to G3 in a configuration the same as that in Figure 6D. This situation has thus not been discussed separately. Our observation that only about half of the Pif1-bound tailless G4 molecules (57.9 %) have shown repetitive unfolding and refolding in our experiments (Figure 5B) seems to support the above points of view. It should be noted that, for complete unfolding of G4 in a single cycle, Pif1 needs to initiate the unfolding from the first G-column (C1) of G4, disrupt G4 to G3 and then G3 to ssDNA in a sequential way.

In the above-mentioned model, we have assumed that G3 (Figure 6E) may refold into G4 (Figure 6A') before it is unfolded by Pif1. This is justified by the fact that G4 refolding is very rapid (~0.25 s, Figure 1F). We have also assumed that Pif1 may bind to any of the three G-columns (and loops) of G3 considering their structural similarity [22]. Our proposal that G3 can be unfolded by Pif1 in any of the three binding modes is supported by our G3 unfolding experiment (see below).

Now we can use the proposed model to explain the observed FRET oscillations quantitatively. In the experiments with Cy3- and Cy5-labelled DNA (Figure 2), Figure 6A' should correspond to the high FRET (~0.9), Figures 6B, 6D, 6E and 6E' to the medium FRET (~0.6) and Figure 6C to the low FRET (~0.4) DNA conformations respectively. Thus the refolding/unfolding of G4 in Cycle I and that of G3 in Cycle II yield high and low oscillation peaks respectively. With Cy3–Pif1 (Figure 3), Figures 6C and 6D should correspond to the high FRET (~0.90–0.94) states and the rest to the medium FRET (~0.63–0.66) ones. Both Cycle I and Cycle II yield similar oscillation peaks.

In addition to Cycle I and Cycle II, we think when Pif1 stalls at ss/dsDNA junction stably, G3 structure could possibly unfold and fold spontaneously, which might also contribute to the medium height peaks (Figure 4A).

### Repetitive unfolding of G3 by Pif1

To further confirm the above model and probe the function of G3 structure in the repetitive cycling, we first checked the

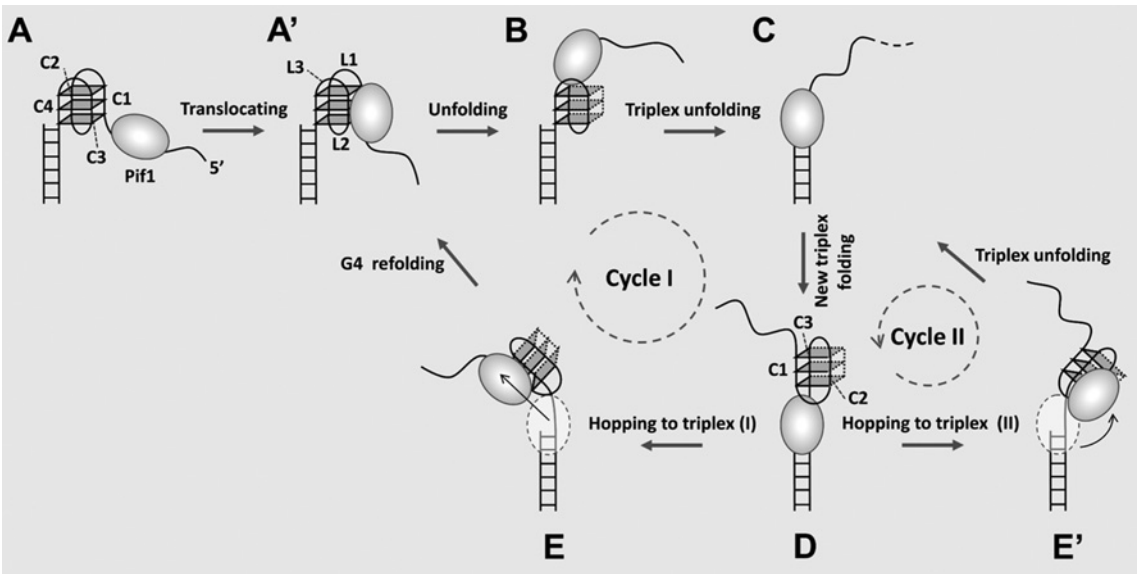


**Figure 6 Proposed model for Pif1-mediated G4 unfolding and repetitive cycling**

The model shows that Pif1 unfolds G4 through G3 as an on-pathway intermediate. (**A**) Pif1 monomer binds initially to the 5′ tail. C1–4 denote the four G-columns of G4. (**A'**) Pif1 binds to G4 through C1 (and L1). L1–3 denote the three lateral loops that connecting the four G-columns. (**B**) Pif1 is in a position ready to strip C2. (**C**) Pif1 is stalled at the ss/dsDNA junction with the ssDNA in unfolded conformation. (**D**) The free part of the unfolded G4 motif refolds spontaneously into a new G3 structure. (**E**) Pif1 hops and binds to G3 through C1. (**E'**) Pif1 hops and binds to G3 through C2, accompanied by a rotation of about 180° for a correct binding orientation. See text for detailed description of the model.

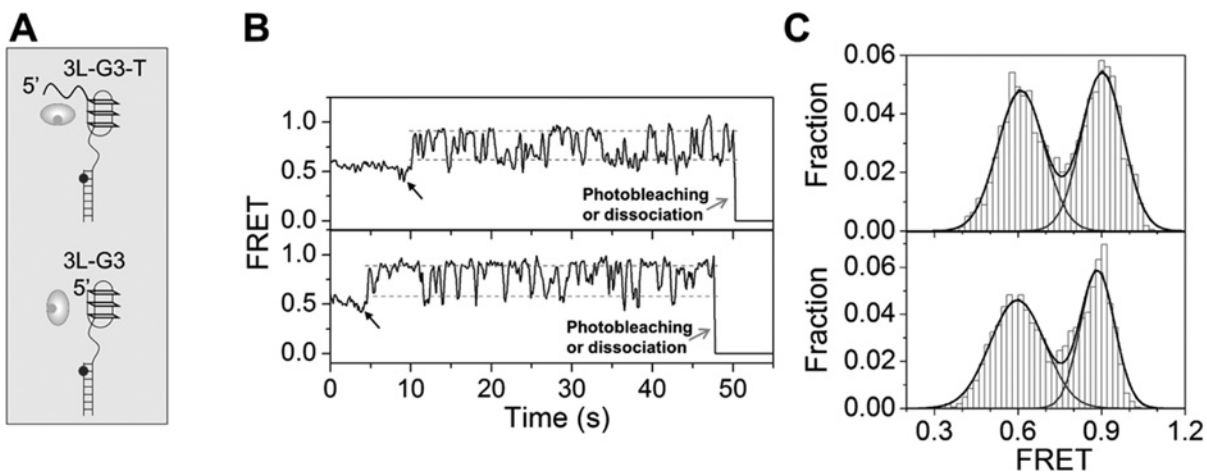


**Figure 7 Repetitive unfolding of G3**

(**A**) Schematic diagrams of Cy5-labelled DNA constructs with G3 having a 26 nt 5′ ssDNA tail (upper panel) or no tail (lower panel). (**B**) Time traces of FRET for the two DNA molecules obtained with 50 $\mu$M ATP and 2 nM Pif1. The time when ATP was added is indicated by a black arrow. (**C**) Histograms of FRET constructed from the oscillating parts of 55 (upper panel) and 60 (lower panel) single-molecule time traces. For each distribution, Gaussian fitting (thick black line) reveals two peaks with medium and high FRET. The two peak positions are $0.610 \pm 0.004$ and $0.903 \pm 0.003$ (3L–G3–T, upper) and $0.598 \pm 0.005$ and $0.885 \pm 0.003$ (3L–G3, lower) respectively.

folding of a G3 motif with CD measurement. It was seen that the folded DNA shows two positive peaks around 289 and 253 nm and two valleys around 235 and 272 nm (Supplementary Figure S6), indicating the stacking of nucleobases and formation of G3 structure [39,50–52]. In addition, we measured the binding affinity of Pif1 for G3, finding Pif1 displays significant preference for G3 over ss/dsDNA junction (Supplementary Figures S3C and S3E), as already mentioned. We then modified the three-layered G4 in Figures 3(A) (top panel) and 5(B) by replacing the last GGG strand near the ss/dsDNA junction with TTT. In this case, only G3 can be folded whereas a six-nt linker between G3 and the duplex is left (Figure 7A). According to the above model, repetitive cycling from Figures 6E to 6C to 6D or from Figures 6E' to 6C to 6D should occur with the new DNA constructs. As expected, when Cy3–Pif1 was added, the FRET was stable at ~0.6 (Figure 7B), indicating G3 was formed and bound with Pif1. Upon addition of ATP, FRET oscillation appeared, indicating that repetitive unfolding of G3 indeed occurred. FRET histograms have two peaks at ~0.6 and ~0.9 (Figure 7C). In the experiments, we have found that almost all of the Pif1-bound tailless G3 molecules (89 %) are unfolded, indicating that G3 can indeed be unfolded by Pif1 bound to any of the three G-columns.

## DISCUSSION

It is a long-standing question and a much debated issue about the occurrence and biological consequence of G4 in cells. Although recent advances have provided convincing evidence that G4 structures are extremely common elements in human genome and the failure of their untangling leads to drastic consequence for cells [20,22,37,53], the question of how helicases process these unusual DNA structures at the molecular level has not yet been addressed. Pif family helicases have been shown to unwind G4 quickly and efficiently in genetic and biochemical studies [22]. In our present work, we used a single-molecule FRET technique to study the molecular mechanism by which one of the Pif family helicase, Pif1, unfolds G4. Overall, our study provides not only an integrated view on how Pif1 resolves G4 structures, but also great details of spontaneous refolding and repetitive unfolding of G4.

First, our studies established a comprehensible basic framework for G4 unfolding by helicases. G4 is radically different from the Watson–Crick double-helix structure of DNA in biochemical and biophysical properties and structural features. Compared with the apparent simplicity and regularity of canonical double-helix DNA, G4 is tiny and compact, usually composed of 3–5 tetrads. Since Pif1 binds well to such structured DNA (Supplementary Figure S3D), it is reasonable to think that Pif1 binds firstly the whole structured G4 and recognizes the first G-tetrad, then disrupts the individual planar G•G•G•G tetrad through breaking simultaneously eight hydrogen bonds [4]. The unwound G-tetrad may be released concomitantly or be released finally when all G-tetrads are disrupted. Alternatively, Pif1 uses a mechanism similar to that for unwinding of canonical double-helix DNA, in which Pif1 binds to the 5′-ssDNA tail and translocates along the ribose-phosphate backbone to unfold the four vertical GGG columns sequentially. Our study has excluded the former mechanism and revealed that Pif1 unfolds a G4 structure sequentially in two large steps, first striping the G-column nearest to the 5′-end and leaving G3 as an on-pathway intermediate, then striping the next G-column and completing the whole unfolding process.

Secondly, we revealed a new type of repetitive activity of helicase during DNA unwinding, which is biologically relevant and radically different from that previously reported with over-hanging duplex DNA [46]. For the previous repetitive shuttling to occur, the protein requires a free ssDNA end to snap back through thermodynamic movement. Since a free ssDNA overhang is neither available nor stable in cells, it is unlikely that this shuttling process actually occurs. In our present study, however, the DNA substrate is reminiscent of an ongoing synthesis of lagging strand stalled by G4. In addition, the repetitive cycling observed here is linked strictly with the intrinsic properties of G4 (and G3) and Pif1, i.e. the spontaneous formations of G3 and G4 as well as the different binding affinities of Pif1 for ss/dsDNA, G3 and G4. The new type of repetitive cycling may allow Pif1 to effectively unfold spontaneously- and rapidly-refolding G4 and create more opportunities for DNA polymerase to rescue stalled replication forks promptly.

Thirdly and maybe more surprisingly, we have identified that G3 structures, which just begin to be recognized although their biological roles remain unclear, play an essential role in repetitive cycling. As revealed from the present study, G3 structure exists as a stable intermediate during G4 unwinding, thus we speculate that G3 structures may have important biological functions that were unimagined before. This is supported by the observed higher binding affinity of Pif1 for G3 than G4 and the strong unfolding activity of Pif1 with G3. Previously, based on a quadruplex folding rule, bioinformatics analysis of human genome has revealed that there are over 370000 potential quadruplex sequences [54]. Now, the probability of DNA sequences forming G3 structures has been increased by more than 10–30-fold based on G3 folding sequences (Xi et al. unpublished results), suggesting that G3 may also need to be resolved by some helicases and be as important as G4 with regard to genome stability in cells. Indeed, our ongoing experiments have shown that a polymerase can be stalled not only by G4, but also by G3. The established link, G4–G3-repetitive cycling, may deepen our understanding of G3/G4 regulation in cells and prompt us to revisit the G4 'only' concept. Whether G3 structures, which should be more abundant than G4 in cells, play similar or different physiological roles to G4 might become a new DNA research field of great interest.

Finally, it should be noted that, Zhou et al. [29] used DNA constructs different from ours and observed that a Pif1 monomer preferentially anchors itself to a 3′-tailed DNA junction and periodically reels in the 3′ tail. This periodic patrolling activity of Pif1 is used to unwind RNA–DNA hybrids in its path or unfold an intramolecular G4 structure on every encounter. As the Pif1 monomer always remains fixed at the 3′-tailed DNA junction, the FRET signal changes due to the repetitive unfolding of G4 are relatively simple and regular. In our experiments, however, the Pif1 monomer does not have a fixed position and may bind to the reformed G3 or G4 at different sites during the repetitive unfolding, thus the FRET signal changes that we observed are more complicated and irregular. Based on all our experimental observations, we have revealed the flexible and strong G3- and G4-unfolding activity of Pif1 as well as the two cycles in the repetitive unfolding processes.

## AUTHOR CONTRIBUTION

Xi-Miao Hou and Xu-Guang Xi are responsible for the experiments design. Wen-Qiang Wu, Xiao-Lei Duan, Na-Nv Liu, Hai-Hong Li and Jing Fu performed experiments. Xi-Miao Hou and Wen-Qiang Wu contributed equally to the present work. All authors analysed and interpreted the data. Figures were compiled by Xi-Miao Hou, Wen-Qiang Wu and Shuo-Xing Dou. Shuo-Xing Dou, Xi-Miao Hou and Xu-Guang Xi wrote and edited the manuscript.

## FUNDING

This work was supported by the National Natural Science Foundation of China [grant numbers 31370798, 11304252 and 31301632]; and Northwest A&F University startup funding for Xu-Guang Xi [grant number Z101021102] and Xi-Miao Hou [grant number Z111021205].

## REFERENCES

1 Bochman, M.L., Paeschke, K. and Zakian, V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. Nat. Rev. Genet. **13**, 770–780 CrossRef PubMed
2 Maizels, N. and Gray, L.T. (2013) The G4 genome. PLoS Genet. **9**, e1003468 CrossRef PubMed
3 Keniry, M.A. (2000) Quadruplex structures in nucleic acids. Biopolymers **56**, 123–146 CrossRef PubMed
4 Davis, J.T. (2004) G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. Angew. Chem. Int. Ed. **43**, 668–698 CrossRef
5 Lane, A.N., Chaires, J.B., Gray, R.D. and Trent, J.O. (2008) Stability and kinetics of G-quadruplex structures. Nucleic Acids Res. **36**, 5482–5515 CrossRef PubMed
6 de Messieres, M., Chang, J.C., Brawn-Cinani, B. and La Porta, A. (2012) Single-molecule study of G-quadruplex disruption using dynamic force spectroscopy. Phys. Rev. Lett. **109**, 058101 CrossRef PubMed
7 Rajendran, A., Endo, M., Hidaka, K. and Sugiyama, H. (2014) Direct and single-molecule visualization of the solution-state structures of G-hairpin and G-triplex intermediates. Angew. Chem. Int. Ed. **53**, 4107–4112 CrossRef
8 Dhakal, S., Schonhoft, J.D., Koirala, D., Yu, Z., Basu, S. and Mao, H. (2010) Coexistence of an ILPR i-motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level. J. Am. Chem. Soc. **132**, 8991–8997 CrossRef PubMed

9 Lee, J., Okumus, B., Kim, D. and Ha, T. (2005) Extreme conformational diversity in human telomeric DNA. Proc. Natl. Acad. Sci. U.S.A. **102**, 18938–18943 CrossRef PubMed

10 Parkinson, G.N., Lee, M.P.H. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. Nature **417**, 876–880 CrossRef PubMed

11 Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res. **34**, 5402–5415 CrossRef PubMed

12 Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature **334**, 364–366 CrossRef PubMed

13 Larson, E.D., Duquette, M.L., Cummings, W.J., Streiff, R.J. and Maizels, N. (2005) MutSalpha binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. Curr. Biol. **15**, 470–474 CrossRef PubMed

14 Dempsey, L.A., Sun, H., Hanakahi, L.A. and Maizels, N. (1999) G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, a role for G-G pairing in immunoglobulin switch recombination. J. Biol. Chem. **274**, 1066–1071 CrossRef PubMed

15 Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. Nucleic Acids Res. **35**, 406–413 CrossRef PubMed

16 Brooks, T.A. and Hurley, L.H. (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. Nat. Rev. Cancer. **9**, 849–861 PubMed

17 Freyer, M.W., Buscaglia, R., Kaplan, K., Cashman, D., Hurley, L.H. and Lewis, E.A. (2007) Biophysical studies of the c-MYC NHE III1 promoter: Model quadruplex interactions with a cationic porphyrin. Biophys. J. **92**, 2007–2015 CrossRef PubMed

18 Dai, J., Dexheimer, T.S., Chen, D., Carver, M., Ambrus, A., Jones, R.A. and Yang, D. (2006) 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**, 1096–1098 CrossRef PubMed

19 Lam, E.Y.N., Beraldi, D., Tannahill, D. and Balasubramanian, S. (2013) G-quadruplex structures are stable and detectable in human genomic DNA. Nat. Commun. **4**, 1796 CrossRef PubMed

20 Biffi, G., Tannahill, D., McCafferty, J. and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. Nat. Chem. **5**, 182–186 CrossRef PubMed

21 Boulé, J.B. and Zakian, V.A. (2006) Roles of Pif1-like helicases in the maintenance of genomic stability. Nucleic Acids Res. **34**, 4147–4153 CrossRef PubMed

22 Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C. and Zakian, V.A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. Nature **497**, 458–462 CrossRef PubMed

23 Crabbe, L., Verdun, R.E., Haggblom, C.I. and Karlseder, J. (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. Science **306**, 1951–1953 CrossRef PubMed

24 Crabbe, L., Jauch, A., Naeger, C.M., Holtgreve-Grez, H. and Karlseder, J. (2007) Telomere dysfunction as a cause of genomic instability in Werner syndrome. Proc. Natl. Acad. Sci. **104**, 2205–2210 CrossRef

25 Booy, E.P., Meier, M., Okun, N., Novakowski, S.K., Xiong, S., Stetefeld, J. and McKenna, S.A. (2012) The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. Nucleic Acids Res. **40**, 4110–4124 CrossRef PubMed

26 Huber, M.D., Duquette, M.L., Shiels, J.C. and Maizels, N. (2006) A conserved G4 DNA binding domain in RecQ family helicases. J. Mol. Biol. **358**, 1071–1080 CrossRef PubMed

27 Bessler, J.B., Torre, J.Z. and Zakian, V.A. (2001) The Pif1p subfamily of helicases: Region-specific DNA helicases? Trends Cell Biol. **11**, 60–65 CrossRef PubMed

28 Bochman, M.L., Sabouri, N. and Zakian, V.A. (2010) Unwinding the functions of the Pif1 family helicases. DNA Repair **9**, 237–249 CrossRef PubMed

29 Zhou, R., Zhang, J., Bochman, M.L., Zakian, V.A. and Ha, T. (2014) Periodic DNA patrolling underlies diverse functions of Pif1 on R-loops and G-rich DNA. eLife **3**, e02190 PubMed

30 Roy, R., Hohng, S. and Ha, T. (2008) A practical guide to single-molecule FRET. Nat. Methods **5**, 507–516 CrossRef PubMed

31 Boulé, J.B. and Zakian, V.A. (2007) The yeast Pif1p DNA helicase preferentially unwinds RNA-DNA substrates. Nucleic Acids Res. **35**, 5809–5818 CrossRef PubMed

32 Rasnik, I., Myong, S., Cheng, W., Lohman, T.M. and Ha, T. (2004) DNA-binding orientation and domain conformation of the *E. coli* Rep helicase monomer bound to a partial duplex junction: Single-molecule studies of fluorescently labeled enzymes. J. Mol. Biol. **336**, 395–408 CrossRef PubMed

33 Dou, S.-X., Wang, P.-Y., Xu, H.Q. and Xi, X.G. (2004) The DNA binding properties of the *Escherichia coli* RecQ helicase. J. Biol. Chem. **279**, 6354–6363 CrossRef PubMed

34 Ha, T. (2001) Single-molecule fluorescence methods for the study of nucleic acids. Curr. Opin. Struct. Biol. **11**, 287–292 CrossRef PubMed

35 Han, H., Hurley, L.H. and Salazar, M. (1999) A DNA polymerase stop assay for G-quadruplex-interactive compounds. Nucleic Acids Res. **27**, 537–542 CrossRef PubMed

36 Ribeyre, C., Lopes, J., Boulé, J.B., Piazza, A., Guédin, A., Zakian, V.A., Mergny, J.L. and Nicolas, A. (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences *in vivo*. PLoS Genet. **5**, e1000475 CrossRef PubMed

37 Paeschke, K., Capra, John A. and Zakian, Virginia A. (2011) DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. Cell. **145**, 678–691 CrossRef PubMed

38 Wang, Y. and Patel, D.J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. Structure **1**, 263–282 CrossRef PubMed

39 Limongelli, V., De Tito, S., Cerofolini, L., Fragai, M., Pagano, B., Trotta, R., Cosconati, S., Marinelli, L., Novellino, E. and Bertini, I. (2013) The G-triplex DNA. Angew. Chem. Int. Ed. **52**, 2269–2273 CrossRef

40 Koirala, D., Ghimire, C., Bohrer, C., Sannohe, Y., Sugiyama, H. and Mao, H. (2013) Long-loop G-quadruplexes are misfolded population minorities with fast transition kinetics in human telomeric sequences. J. Am. Chem. Soc. **135**, 2235–2241 CrossRef PubMed

41 Li, W., Hou, X.M., Wang, P.Y., Xi, X.G. and Li, M. (2013) Direct measurement of sequential folding pathway and energy landscape of human telomeric G-quadruplex structures. J. Am. Chem. Soc. **135**, 6423–6426 CrossRef PubMed

42 Hwang, H., Buncher, N., Opresko, Patricia L. and Myong, S. (2012) POT1-TPP1 regulates telomeric overhang structural dynamics. Structure **20**, 1872–1880 CrossRef PubMed

43 Mashimo, T., Yagi, H., Sannohe, Y., Rajendran, A. and Sugiyama, H. (2010) Folding pathways of human telomeric type-1 and type-2 G-quadruplex structures. J. Am. Chem. Soc. **132**, 14910–14918 CrossRef PubMed

44 Rachwal, P.A., Brown, T. and Fox, K.R. (2007) Effect of G-tract length on the topology and stability of intramolecular DNA quadruplexes. Biochemistry **46**, 3036–3044 CrossRef PubMed

45 Rachwal, P.A. and Fox, K.R. (2007) Quadruplex melting. Methods **43**, 291–301 CrossRef PubMed

46 Myong, S., Rasnik, I., Joo, C., Lohman, T.M. and Ha, T. (2005) Repetitive shuttling of a motor protein on DNA. Nature **437**, 1321–1325 CrossRef PubMed

47 Jiang, Y., Douglas, N.R., Conley, N.R., Miller, E.J., Frydman, J. and Moerner, W.E. (2011) Sensing cooperativity in ATP hydrolysis for single multisubunit enzymes in solution. Proc. Natl. Acad. Sci. **108**, 16962–16967 CrossRef

48 Ramanagoudr-Bhojappa, R., Chib, S., Byrd, A.K., Aarattuthodiyil, S., Pandey, M., Patel, S.S. and Raney, K.D. (2013) Yeast Pif1 helicase exhibits a one-base-pair stepping mechanism for unwinding duplex DNA. J. Biol. Chem. **288**, 16185–16195 CrossRef PubMed

49 Galletto, R. and Tomko, E.J. (2013) Translocation of *Saccharomyces cerevisiae* Pif1 helicase monomers on single-stranded DNA. Nucleic Acids Res. **41**, 4613–4627 CrossRef PubMed

50 Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A. and Spada, G.P. (2010) A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. Org. Biomol. Chem. **8**, 2683–2692 CrossRef PubMed

51 Karsisiotis, A.I., Hessari, N.M.a., Novellino, E., Spada, G.P., Randazzo, A. and Webba da Silva, M. (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. Angew. Chem. Int. Ed. **50**, 10645–10648 CrossRef

52 Randazzo, A., Spada, G. and da Silva, M. (2013) Circular dichroism of quadruplex structures. In Quadruplex nucleic acids (Chaires, J.B. and Graves, D., eds), pp. 67–86, Springer, Heidelberg

53 Murat, P. and Balasubramanian, S. (2014) Existence and consequences of G-quadruplex structures in DNA. Curr. Opin. Genet. Dev. **25**, 22–29 CrossRef PubMed

54 Stegle, O., Payet, L., Mergny, J.-L., MacKay, D.J.C. and Huppert, J.L. (2009) Predicting and understanding the stability of G-quadruplexes. Bioinformatics **25**, i374–i1382 CrossRef PubMed


Received 4 August 2014/2 December 2014; accepted 4 December 2014
Published as BJ Immediate Publication 4 December 2014, doi:10.1042/BJ20140997