WILEY
InterScience®
DISCOVER SOMETHING GREAT

Fast detection of quadruplex structure in DNA by the intrinsic fluorescence of a single-stranded DNA binding protein

Xin-ying Zhuang[1], Jun Tang[1], Yu-hua Hao[2] and Zheng Tan[1,2]*

[1]Laboratory of Biochemistry and Biophysics, College of Life Sciences, Wuhan University, Wuhan 430072, P. R. China
[2]State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, P. R. China

Single-stranded guanine-rich (G-rich) DNA can fold into a four-stranded G-quadruplex structure and such structures are implicated in important biological processes and therapeutic applications. So far, bioinformatic analysis has identified up to several hundred thousand of putative quadruplex sequences in the genome of human and other animal. Given such a large number of sequences, a fast assay would be desired to experimentally verify the structure of these sequences. Here we describe a method that identifies the quadruplex structure by a single-stranded DNA binding protein from a thermoautotrophic archaeon. This protein binds single-stranded DNA in the unfolded, but not in the folded form. Upon binding to DNA, its fluorescence can be quenched by up to 70%. Formation of quadruplex greatly reduces fluorescence quenching in a K^+-dependent manner. This structure-dependent quenching provides simple and fast detection of quadruplex in DNA at low concentration without DNA labelling. Copyright © 2007 John Wiley & Sons, Ltd.

INTRODUCTION

Single-stranded guanine-rich (G-rich) DNA can fold into a four-stranded G-quadruplex structure in the presence of K^+ and Na^+, but not of Li^+ (Gilbert and Feigon, 1999; Simonsson, 2001). Such structures have received significant attention in recent years because of their role in important biological processes and potential as therapeutic targets (Neidle and Read, 2001; Jing *et al.*, 2005). Quadruplex forming sequences have been found to exist in many essential regions of chromosomes (Catasti *et al.*, 1999), for example, telomeres (Blackburn, 1991), promoter of oncogenes (Simonsson *et al.*, 1998), immunoglobulin switch (Sen and Gilbert, 1988) and the insulin regulatory (Hammond-Kosack *et al.*, 1992) regions. Bioinformatic analysis has revealed up to several hundred thousand of putative quadruplex sequences in human and chicken genome (Huppert and Balasubramanian, 2005; Todd *et al.*, 2005; Du *et al.*, 2007).

Experimental recognition and verification of G-quadruplex structure in these sequences is a prerequisite for the design and development of drugs targeting quadruplex. Currently, many physical and chemical methods, i.e. UV melting, circular dichroism, surface plasmon resonance, NMR, chemical footprinting, gel electrophoresis and fluorescence resonance energy transfer have been used to probe G-quadruplex

formation (Hardin *et al.*, 1991; Feigon *et al.*, 1995; Mergny *et al.*, 1998; Mergny and Maurizot, 2001; Redon *et al.*, 2003; Zhao *et al.*, 2004; Dexheimer *et al.*, 2006). Some of these methods require complicated data analysis or sample processing, such as immobilization or labelling; many are time consuming. Given such a large number of putative quadruplex sequences, a fast and simple method would be desired. In the present work, we describe a method for fast detection of DNA quadruplex using a single-stranded DNA binding protein (SSB) from a thermoautotrophic archaeon, *Methanococcus jannaschii*. The fluorescence of this protein is quenched upon DNA binding and quadruplex structure in DNA can be recognized by the K^+-dependent reduction in fluorescence quenching.

MATERIALS AND METHODS

Oligonucleotides

All oligonucleotides used were listed in Table 1 and were synthesized and PAGE purified by Sangon Technology (Shanghai, China). $5'$-end-^{32}P labelling was carried out using T4 polynucleotide kinase (Fermentas, Lithuania).

Expression and purification of *M. jannaschii* single-stranded DNA binding protein (SSB)

The bacterial expression vector containing ORF MJ1159 of *M. jannaschii* was a kind gift from Professor Thomas J.

*Correspondence to: Z. Tan, State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, P.R. China.
E-mails: tanclswu@public.wh.hb.cn; z.tan@ioz.ac.cn

Table 1. Oligonucleotides used in this study

Oligonucleotide	Sequence (5′ to 3′)
G21	GGGTTAGGGTTAGGGTTAGGG
RG21	GTGTGAGTGAGGAGTGGTGTG
G24	TTAGGGTTAGGGTTAGGGTTAGGG
R24	TGGGATTGTGTGAAGTGGTGAGTG
C24	CCCTAACCCTAACCCTAACCCTAA
c-myc	AGGGTGGGGAGGGTGGGG
HIF1α	AGGGCGGGGGAGAGGGGAGGGG
PV	GGGGGGCCGGGGCGAGGGGAAGGGG
R19	CCTGCTGCCTTCCTTGGAT
H21	GTGGTCCTTTTTTTGGACCAC
H22	GTGGTCCACTTTTGTGGACCAC
RH21	GTTGTTGCTCTTGTGACTCCA
C21	CCCTAACCCTAACCCTAACCCC
RC21	CACACCACTCCTCACTCACAC

Kelly at Johns Hopkins University School of Medicine, Baltimore, USA. Purification of recombinant His-tagged SSB was carried out according to the previously published protocol with minor modification (Kelly *et al.*, 1998). The bacterial were harvested and lysed by lysozyme and sonication in buffer A (50 mM NaH$_2$PO$_4$, pH 8.0, 300 mM NaCl, 40 mM imidazole) plus 1 mM DTT, 2 mg/ml leupeptin, 1 mM phenylmethylsulfonyl fluoride, 1 mg/ml pepstatin, and 2 mg/ml aprotinin. After centrifugation for 20 min at 20 000 g, 4°C, the supernatant was collected and heated at 75°C for 15 min to denature the mesophilic proteins. The lysate was then kept on ice for 30 min, centrifuged for 20 min at 20 000 g, 4°C. The supernatant was loaded on affinity HisTrap HP Columns (GE, USA) equilibrated with buffer A. After washing with buffer A, the SSB was eluted with buffer A containing 300 mM imidazole and dialyzed against buffer B (10 mM Tris-HCl, pH 7.8, 3 mM EDTA, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride). Protein concentration was determined by the Bradford assay.

UV crosslinking and gel mobility shift assay

UV crosslinking of G24 G-quadruplex was carried out as described in Williamson *et al.* (1989), Zhao *et al.* (2004) and Zeng *et al.* (2005). ^{32}P-5′-end-labeled G24, crosslinked or uncrosslinked, were mixed with SSB or C24 at equimolar ratio in buffer C (25 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM DTT, 150 mM LiCl). The mixtures were incubated for 45 min at 25°C and then resolved on 12% native polyacrylamide gel. The gel was autoradiographed on X-ray film. G24 instead of G21 was used to facilitate the crosslinking that occurs on thymidine.

Fluorescence measurements

All fluorescence measurements were carried out on a Spex Fluorolog-3 spectrofluorometer (HORIBA Jobin Yvon, France) at room temperature. The slits of emission and excitation were both 5 nm. Excitation wavelength was set at

295 nm and the emission was detected at 348 nm. Binding reactions were performed with SSB in 10 mM HEPES buffer (pH 7.4) containing 150 mM K$^+$ or Na$^+$, or in 10 mM Tris-EDTA buffer (pH 7.4) containing 150 mM Li$^+$. Fluorescence was recorded before and after the addition of 1/40 volume of DNA stock solution in the same buffer. The oligonucleotides were heated at 95°C for 10 min, slowly cooled down to room temperature before use. For the detection of i-motif, measurements were performed in 10 mM Tris-acetate, 150 mM LiCl.

RESULTS

The protein used is a SSB from a thermoautotrophic archaeon, *M jannaschii*. With five of the six tryptophan residues located at the DNA binding site, the protein is excited at 295 nm and emits fluorescence at 348 nm. The fluorescence is quenched by a maximum of 60–70% upon DNA binding (Figure 1a). This protein is monomeric and has high affinity and selectivity for single-stranded DNA. It binds DNA at 1:1 molar ratio for DNA of 12–30 nt (Kelly *et al.*, 1998). The protein was first tested with two structural forms of the G-rich strand of human telomeric DNA (G24) by gel electrophoresis (Figure 1b). Quadruplex does not form in Li$^+$ solution. The G24 loaded with SSB in the Li$^+$ solution resulted in retarded migration compared to the free G24 indicating that it was bound by the SSB. When the quadruplex formed in K$^+$ solution was UV-crosslinked to prevent it from unfolding (Williamson *et al.*, 1989; Zhao *et al.*, 2004; Zeng *et al.*, 2005), the DNA was not bound by the SSB since the inclusion of SSB did not affect its migration.

The ability of the SSB to distinguish the unfolded and the folded form of DNA by DNA-binding dependent fluorescence quenching provides fast detection of quadruplex structure in G-rich DNA. Figure 2 shows the detection of human telomeric DNA (G21) that was added into SSB solution after it has been equilibrated in the presence of K$^+$, Na$^+$ or Li$^+$. Quadruplex forms in K$^+$, Na$^+$ but not in the Li$^+$ solution (Balagurumoorthy and Brahmachari, 1994). Comparing with the randomized equivalent RG21 and the
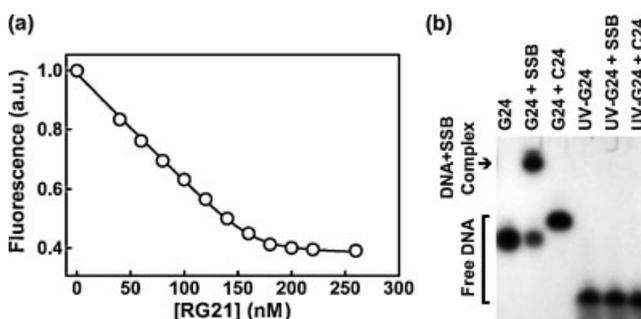
Figure 1. (a) Quenching of SSB (200 nM) fluorescence by a random 21 mer oligonucleotide RG21 at various concentrations in 150 mM K$^+$ solution. (b) Binding of SSB to the G-rich strand of human telomeric DNA G24 examined by gel electrophoresis. ^{32}P-labeled G24 or UV-crosslinked G24 quadruplex (UV-G24) were preincubated with SSB or C24 in solution containing 150 mM Li$^+$ before electrophoresis. The UV-crosslinked quadruplexes were more compact and migrated faster than the unstructured form.
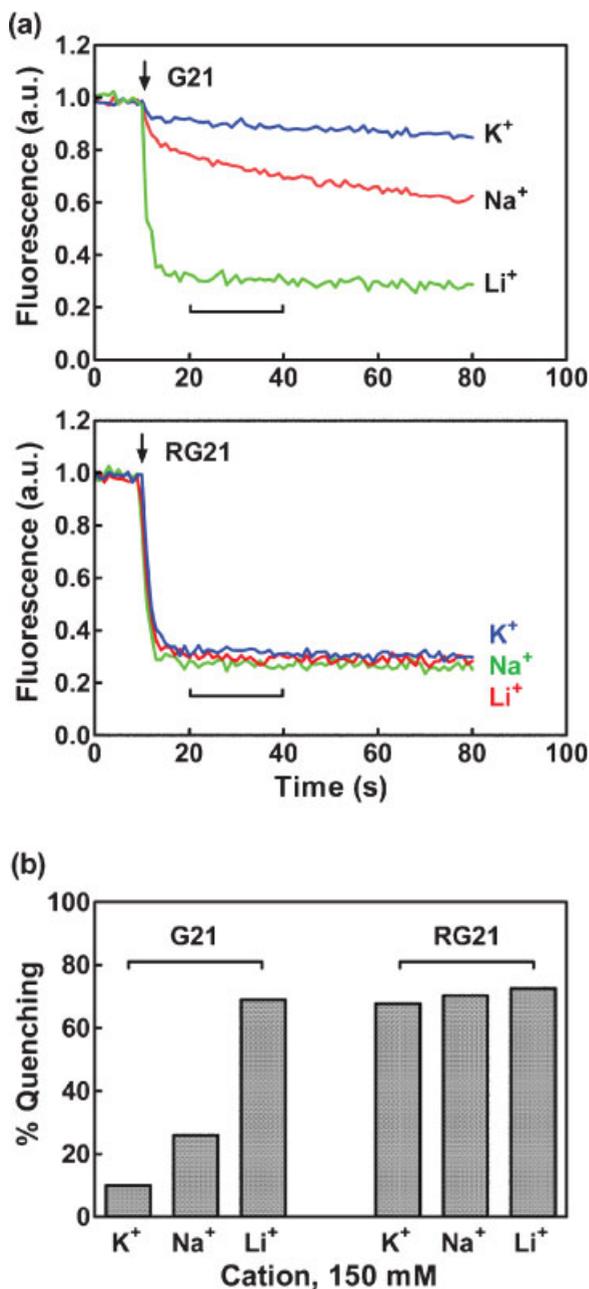
Figure 2. Recognition of quadruplex structure in G21 by fluorescence quenching of SSB. (a) Fluorescence quenching monitored over time by native (up panel, G21) and randomized (lower panel, RG21) human telomeric DNA in solution containing 150 mM of K^+, Na^+ or Li^+. (b) Per cent SSB fluorescence quenching reached by G21 and RG21 within the time interval of 10–30 sec after the addition of DNA. Results obtained using data from (a). 10 nM SSB and 20 nM DNA were used.

G21 in Li^+ solutions, the formation of G21 quadruplex in K^+ and Na^+ solution resulted in a lower quenching of SSB fluorescence. The much lower quenching in K^+ vs. Na^+ solution agrees with the known fact that K^+ is a much stronger quadruplex stabilizer than Na^+ (Balagurumoorthy and Brahmachari, 1994). The K^+ concentration-dependent quadruplex formation was also well reflected by the quenching of SSB fluorescence (Figure 3). With increase in the K^+ concentration, the quenching by G21 decreased but that by RG21 did not.
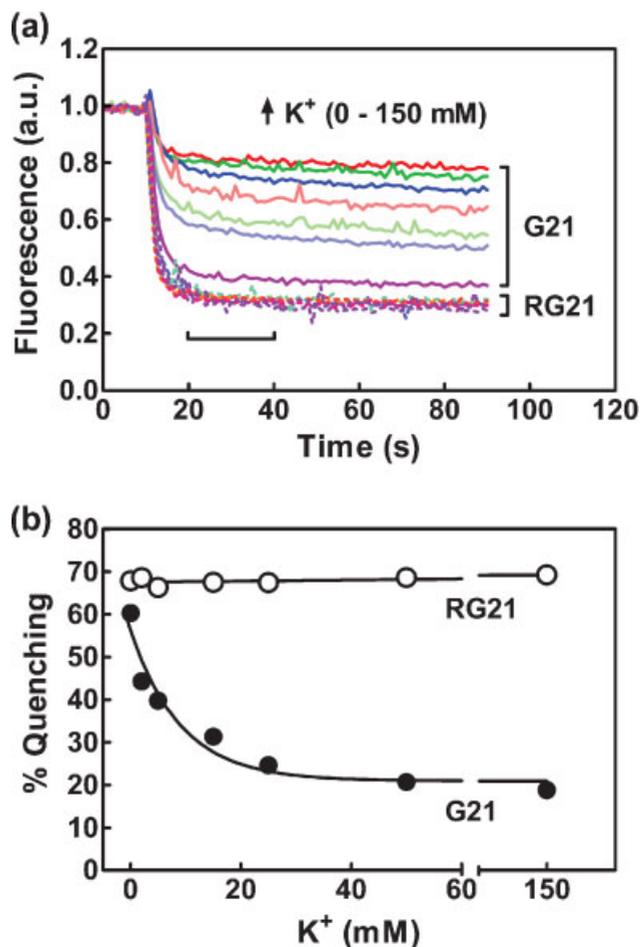
Figure 3. Quadruplex formation as a function of K^+ concentration detected by quenching of SSB fluorescence. (a) Fluorescence quenching monitored over time by native (G21) and randomized (RG21) human telomeric DNA in solution containing 0–150 mM of K^+. Li^+ was added to make the concentration of total mono cations to 150 mM. (b) Per cent SSB fluorescence quenching reached by G21 and RG21 within the time interval of 10–30 sec after the addition of DNA. Results obtained using data from (a). 10 nM SSB and 20 nM DNA were used.

Further test was performed using three other sequences (Figure 4). The HIF1α and c-myc sequences, which have been reported to form quadruplex (Simonsson *et al.*, 1998; De Armond *et al.*, 2005), only resulted in weak SSB fluorescence quenching as the G21 did. The PV sequence has been predicted to form quadruplex (Weitzmann *et al.*, 1996) though this not has been experimentally verified. Its low quenching on SSB fluorescence indicates that it formed quadruplex, which was further confirmed by its increased mobility in native gel electrophoresis in the presence of 150 mM K^+ (Figure 4b). Such reduced fluorescence quenching was dependent on structure rather than sequence because of the same sequences in the Li^+ solution (Figure 4c) and the other three random sequences in the K^+ solution (Figure 4d) all showed strong quenching on SSB fluorescence.

Since the SSB binds DNA in the unfolded form, it should also detect other forms of secondary structures. To demonstrate the ability to distinguish quadruplex from other structures, two DNA hairpins with a loop of 4 and 7 nt
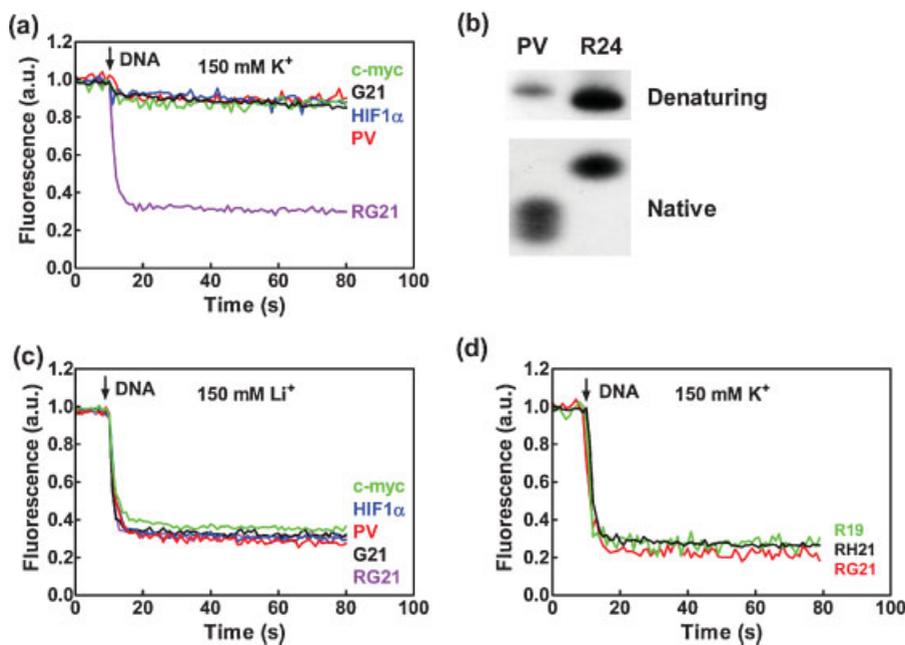
Figure 4. Recognition of quadruplex structure in different sequences by fluorescence quenching of SSB. (a) Fluorescence quenching by quadruplex-forming sequences (c-myc, G21, HIFα, PV) in 150 mM K^+ solution. (b) Electrophoresis of ^{32}P-labeled PV DNA (left lane) and a randomized 24 mer (R24, right lane) in denaturing (top) and native (bottom, with 150 mM K^+) gel. The faster migration of PV in native gel shows it formed quadruplexes that might present in more than one isofoms as it appeared smearing. (c) Fluorescence quenching by quadruplex-forming sequences (c-myc, G21, HIFα, PV) in 150 mM Li^+ solution. (d) Fluorescence quenching by sequences (R19, RH21, RG21) that do not form quadruplex in 150 mM K^+ solution. SSB fluorescence was monitored over time with 10 nM SSB and 20 nM DNA.

respectively were assayed. As is shown in Figure 5, the formation of hairpin did reduce quenching on the fluorescence of SSB. But the quenching was not K^+-dependent and occurred in the presence of either K^+ or Li^+. Therefore, the K^+-dependence of SSB fluorescence quenching associated with quadruplex will provide a criterion to distinguish quadruplex from other forms of structures. The SSB requires
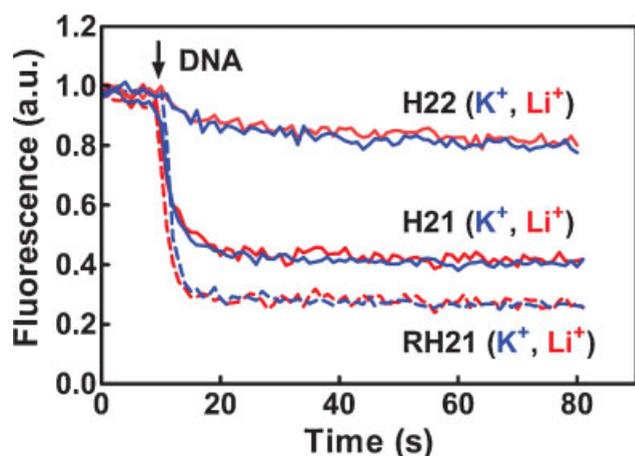
Figure 5. Quenching of SSB fluorescence by oligonucleotide that forms (H21, H22) or does not (RH21) form hairpin. The loop sizes for H21 and H22 are 7 and 4 nt respectively. Fluorescence quenching monitored over time with 10 nM SSB and 20 nM DNA in solution containing either 150 mM K^+ or Li^+ as indicated.

>10 nt for binding simple single-stranded DNA (Kelly *et al.*, 1998). The presence of the double-stranded stem may relax the length requirement in this case since the hairpin with a 7 nt loop showed significant quenching.

Single-stranded cytosine-rich (C-rich) DNA can fold into an intercalated four-stranded i-motif structure at acidic pH, which is stabilized by hemi-protonated cytosine–cytosine base pairing (Gueron and Leroy, 2000). We also used the SSB to detect the pH-induced structural transition in $(C_3TA_2)_3C_3$, the core sequence of the C-rich strand of human telomeric DNA. From neutral to pH 5.0, the fluorescence quenching of the SSB clearly indicated a pH-dependent structural transition of the DNA (Figure 6) similar to those observed by other techniques (Manzini *et al.*, 1994; Zeng *et al.*, 2005). Within this range, pH change did not significant affect the fluorescence and binding activity of the SSB. The reduced fluorescence quenching is explained by the formation of i-motif that cannot be bound with the SSB.

DISCUSSION

The data presented in this work demonstrated a spectroscopic method for the detection of tetraplex structure in G- and C-rich DNA using a single-stranded DNA binding protein. The assay utilizes the DNA binding-dependent quenching of the intrinsic fluorescence of the SSB and its
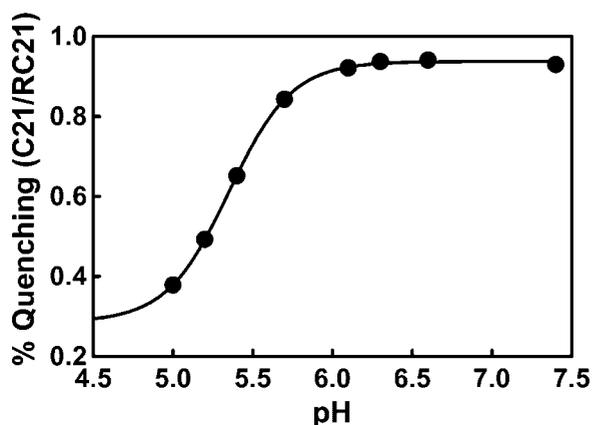
Figure 6. pH-dependent SSB fluorescence quenching showing i-motif formation in the C-rich strand of human telomeric DNA (C21). At each pH, two oligonucleotides, C21 and its randomized equivalent RC21, were assayed for quenching on SSB fluorescence and the ratio of their % fluorescence quenching was plotted as a function of pH.

structural selection towards unfolded single-stranded DNA. To detect the quadruplex structure in putative sequence, the DNA is equilibrated in 150 mM K^+ solution before it is mixed with SSB in the same solution. In comparison with DNA that does not form quadruplex, formation of quadruplexes in putative sequences will reduce the amount of unstructured species and, as a result, fluorescence quenching (Figures 2 and 4). Quadruplex formation is identified by reduction of SSB fluorescence quenching in K^+ but not in Li^+ solution. Likewise, the formation of i-motif is identified by the pH-dependence of fluorescence quenching.

The method possesses several advantages. It is simple, fast (can be done well within 1 minute) and sequence-universal. It uses native DNA that does not involve DNA modifications, such as immobilization or labelling as in some other methods, which may alter the property of quadruplexes (Green *et al.*, 2003; Zhao *et al.*, 2004). The quantity of DNA required (20 nM) is roughly 100–1000-fold lower than those in other spectroscopic methods. Because of the fast speed and that the protein can be prepared in large quantity by expression in *Escherichia. coli*, the assay may especially be suitable for high-throughput and large-scale initial analysis of potential quadruplex or i-motif forming sequences.

Upon mixing with SSB, the unfolded DNA molecules are captured by the SSB within seconds. This will move the equilibrium between the folded and unfolded species towards the unfolded one, thus resulting in additional quenching of fluorescence over time. The rate of this process depends on the stability of quadruplex or i-motif. Successful detection may rely on the stability of the structures. K^+ is a much stronger quadruplex stabilizer than Na^+ (Balagurumoorthy and Brahmachari, 1994). For this reason, the quadruplex formed in the Na^+ solution displayed faster quenching than in the K^+ solution (Figure 2a). Therefore, it is recommended to perform the assays in K^+ solution for quadruplex to achieve maximum stability. Quadruplexes are very stable and their unfolding is usually a slow process. For instance, the quadruplex formed by human telomeric DNA has a folding equilibium constant of several tens to several hundreds (Zhao *et al.*, 2004) and takes hours or longer to unfold (Green *et al.*, 2003). For the several sequences tested, the fluorescence quenching accompanying quadruplex unfolding is much slower than that produced by the DNA species that is unfolded at the time of mixing (Figures 2 and 4). However, it is still possible that very weak quadruplex may unfold quickly and thus may not be easily identified. In this case, other techniques may be required to verify the structure. In our assays, one-fold excess of DNA over SSB was used to ensure maximal quenching of SSB fluorescence by randomized DNA. For less stable structures, the molar DNA concentration may need to be adjusted to match that of the SSB to improve sensitivity. On the other hand, sequences flanking a quadruplex are potential binding targets for the SSB. Thus it may be desired to use only core sequence to avoid such binding. The SSB fluorescence is quenched by a maximum of 70%. Further improvements in assay sensitivity may be achieved by deleting or replacing the tryptophan residue outside of the DNA binding site.

Acknowledgments

REFERENCES

Balagurumoorthy P, Brahmachari SK. 1994. Structure and stability of human telomeric sequence. *J. Biol. Chem.* **269**: 21858–21869.

Blackburn EH. 1991. Structure and function of telomeres. *Nature* **350**: 569–573.

Catasti P, Chen X, Mariappan SV, Bradbury EM, Gupta G. 1999. DNA repeats in the human genome. *Genetica* **106**: 15–36.

De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW. 2005. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 alpha promoter. *Biochemistry* **44**: 16341–16350.

Dexheimer TS, Sun D, Hurley LH. 2006. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. *J. Am. Chem. Soc.* **128**: 5404–5415.

Du Z, Kong P, Gao Y, Li N. 2007. Enrichment of G4 DNA motif in transcriptional regulatory region of chicken genome. *Biochem. Biophys. Res. Commun.* **354**: 1067–1070.

Feigon J, Koshlap KM, Smith FW. 1995. 1H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol.* **261**: 225–255.

Gilbert DE, Feigon J. 1999. Multistranded DNA structures. *Curr. Opin. Struct. Biol.* **9**: 305–314.

Green JJ, Ying L, Klenerman D, Balasubramanian S. 2003. Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap. *J. Am. Chem. Soc.* **125**: 3763–3767.

Gueron M, Leroy JL. 2000. The i-motif in nucleic acids. *Curr. Opin. Struct. Biol.* **10**: 326–331.

Hammond-Kosack MC, Dobrinski B, Lurz R, Docherty K, Kilpatrick MW. 1992. The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res.* **20**: 231–236.

Hardin CC, Henderson E, Watson T, Prosser JK. 1991. Monovalent cation induced structural transitions in telomeric DNAs: G-DNA folding intermediates. *Biochemistry* **30**: 4460–4472.

Huppert JL, Balasubramanian S. 2005. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **33**: 2908–2916.

Jing N, Sha W, Li Y, Xiong W, Tweardy DJ. 2005. Rational drug design of G-quartet DNA as anti-cancer agents. *Curr. Pharm. Des.* **11**: 2841–2854.

Kelly TJ, Simancek P, Brush GS. 1998. Identification and characterization of a single-stranded DNA-binding protein from the archaeon Methanococcus jannaschii. *Proc. Natl. Acad. Sci. USA* **95**: 14634–14639.

Manzini G, Yathindra N, Xodo LE. 1994. Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res.* **22**: 4634–4640.

Mergny JL, Maurizot JC. 2001. Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat. *Chembiochem* **2**: 124–132.

Mergny JL, Phan AT, Lacroix L. 1998. Following G-quartet formation by UV-spectroscopy. *FEBS Lett.* **435**: 74–78.

Neidle S, Read MA. 2001. G-quadruplexes as therapeutic targets. *Biopolymers* **56**: 195–208.

Redon S, Bombard S, Elizondo-Riojas MA, Chottard JC. 2003. Platinum cross-linking of adenines and guanines on the quadruplex structures of the AG3(T2AG3)3 and (T2AG3)4 human telomere sequences in Na$^+$ and K$^+$ solutions. *Nucleic Acids Res.* **31**: 1605–1613.

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

Simonsson T. 2001. G-quadruplex DNA structures—variations on a theme. *Biol. Chem.* **382**: 621–628.

Simonsson T, Pecinka P, Kubista M. 1998. DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* **26**: 1167–1172.

Todd AK, Johnston M, Neidle S. 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **33**: 2901–2907.

Weitzmann MN, Woodford KJ, Usdin K. 1996. The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. *J. Biol. Chem.* **271**: 20958–20964.

Williamson JR, Raghuraman MK, Cech TR. 1989. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**: 871–880.

Zeng ZX, Zhao Y, Hao YH, Tan Z. 2005. Tetraplex formation of surface-immobilized human telomere sequence probed by surface plasmon resonance using single-stranded DNA binding protein. *J. Mol. Recognit.* **18**: 267–271.

Zhao Y, Kan ZY, Zeng ZX, Hao YH, Chen H, Tan Z. 2004. Determining the folding and unfolding rate constants of nucleic acids by biosensor. Application to telomere G-quadruplex. *J. Am. Chem. Soc.* **126**: 13255–13264.