Published online 29 August 2017

Nucleic Acids Research, 2017, Vol. 45, No. 17 ***9813–9822***
doi: 10.1093/nar/gkx765

Self-biotinylation of DNA G-quadruplexes via intrinsic peroxidase activity

Owen J. Einarson[1] and Dipankar Sen[1,2,*]

[1]Department of Chemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada and [2]Department of Molecular Biology & Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Received April 15, 2017; Revised August 09, 2017; Editorial Decision August 11, 2017; Accepted August 25, 2017

ABSTRACT

The striking and ubiquitous *in vitro* affinity between hemin and DNA/RNA G-quadruplexes raises the intriguing possibility of its relevance to biology. To date, no satisfactory experimental framework has been reported for investigating such a possibility. Complexation by G-quadruplexes leads to activation of the bound hemin toward catalysis of 1- and 2-electron oxidative reactions, with phenolic compounds being particularly outstanding substrates. We report here a strategy for exploiting that intrinsic peroxidase activity of hemin•G-quadruplex complexes for self-biotinylation of their G-quadruplex component. Such self-biotinylation occurs with good efficiency and high discrimination *in vitro*, being specific for G-quadruplexes and not for duplexes. The biotinylated DNA, moreover, remains amenable to polymerase chain reaction amplification, rendering it suitable for analysis by ChIP-Seq and related methods. We anticipate that this self-biotinylation methodology will also serve as a sensitive tool, orthogonal to existing ones, for identifying, labeling and pulling down cellular RNA and DNA G-quadruplexes in general, as well as proteins bound to or proximal to such quadruplexes.

INTRODUCTION

G-quadruplexes are a class of folded structures formed by single-stranded, guanine-rich DNAs and RNAs (1–5) under physiological solutions conditions. The fundamental folding unit within G-quadruplexes is the guanine quartet (6,7), in which four guanines from the same or different DNA/RNA strands hydrogen bond via Hoogsteen base pairing. Sequences identified to date to form stable G-quadruplexes range from those derived from organismal genomes to wholly artificial sequences such as aptamers and ribozymes/DNAzymes obtained from random-sequence DNA/RNA libraries by *in vitro* selection ('SELEX'). DNA G-quadruplexes are highly polymorphic in terms of strand orientation, structure and overall topology, whereas RNA quadruplexes are typically conservative, favoring a fully parallel strand orientation (8). G-quadruplexes are differentially stabilized by different alkali cations ($K^+ > Na^+ >> Li^+$), and Na^+ and K^+ often favor the formation of different G-quadruplex conformers from the same DNA sequence (1–5,8).

The search for G-quadruplexes in living cells has been the subject of intensive research in recent years. A number of different approaches have been taken; successful visualizations of G-quadruplexes in live cells using G4-specific fluorescent antibodies have, in particular, shown interesting temporal and spatial patterns (9–11). Techniques such as ChIP-Seq have been used to map natural G-quadruplex protein binding locations (12). For instance, Pif1 DNA helicase, known to unwind quadruplexes *in vitro*, was identified via ChIP-Seq on the *Saccharomyces cerevisiae* genome to target predicted quadruplex motifs (13). In yet another study, the known small molecule G-quadruplex stabilizer, pyridostatin, was allowed to form and/or stabilize G-quadruplexes within living cells (14); the resulting replicative pausing triggered a DNA damage response, and one of the proteins (γH2AX) triggered by the response was used as a target for ChIP-Seq analysis. In this study, regions of genomic DNA with higher predicted ability to form G-quadruplexes showed increased representation in the analysis (14). Higher resolution methods to map G-quadruplexes in cells have more recently been attempted (15). A 2015 study combined Illumina high-throughput sequencing of the human genome with a polymerase stop assay under conditions favoring G-quadruplex formation. Between 5×10^5 and 7.5×10^5 different quadruplex-capable sequences were found in the human genome, larger than the 3.6×10^5 sequences identified by earlier predictive algorithms (16). Such an abundance of potential quadruplex forming sequences in the human genome raises many questions regarding their possible function. Indeed, a sizable literature cites and postulates putative role(s) of DNA and RNA G-quadruplexes in the normal genetic program of the cell as well as in various disease states (5,17). Regarding the latter, there is evidence to suggest that G-quadruplexes may play key roles in certain abnormal repeat expansion-linked neu-

*To whom correspondence should be addressed. Tel: +1 778 782 4386; Fax: +1 778 782 5583; Email: sen@sfu.ca

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

rodegenerative diseases, such as familial amyotrophic lateral sclerosis and frontotemporal dementia. An expansion of the sequence GGGGCC in the *C9orf72* gene has been found to be present in neurons afflicted by these diseases, and both the DNA itself and RNA transcribed from this repeat expansion have been shown to form G-quadruplexes *in vitro* and *in vivo* (18–22).

Despite the above approaches for the identification of G-quadruplexes *in vivo*, a significant need remains for new, orthogonal and highly sensitive methodologies for identifying and sequencing G-quadruplex-forming motifs in the genomic DNA of living cells. G-quadruplexes have been shown to be excellent binding targets for small molecule cellular cofactors as well as for numerous synthetic ligands (4,23). Our lab originally showed that ferric heme or hemin [Fe(III)-protoporphyrin IX], the ubiquitous cellular cofactor responsible for multiple metabolic functions including oxidative catalysis, binds tightly to both RNA and DNA G-quadruplexes (24–28). The dissociation constants for such interactions can be as low as ∼10 nM (29). Primed by oxidants such as hydrogen peroxide, these complexes are excellent catalysts for oxidative reactions, both 1-electron oxidations (such as catalyzed by horseradish peroxidase (HRP)) and 2-electron oxidations (such as catalyzed by cytochrome P450 enzymes) (24–28). We refer to such complexes hereafter as 'heme-ribozymes' or 'heme-DNAzymes'.

In cell biology, immunohistochemistry protocols allow for visualization of cellular targets via use labeled antibodies. The specificity of these probes is often compromised by high background staining. One remedy for this has been the tyramide signal amplification assay. Imaging is obtained using of a secondary antibody conjugated with HRP and reactive HRP substrates that give rise to colored or fluorescent 'readout' products (30,31). Useful HRP substrates of this kind are often phenolic compounds (such as derivatized tyramides). In the presence of H_2O_2 and HRP-derivatized antibody, they react to form short-lived (lifetime <1 ms) phenoxyl radicals, which can react to form covalent adducts with nearby proteins and nucleic acids. If the supplied tyramide derivative is fluorescent, regular visualization methods are then enabled; if it is a biotin derivative, streptavidin-mediated pulldowns can be carried out to identify macromolecules present at or close to the targeted location (32). This basic premise of tyramides has been adapted to even more sophisticated use: Ting *et al*. engineered an ascorbate peroxidase, APEX2, that can be targeted intracellularly to specific regions of the cell (33,34). In the presence of a low concentration of H_2O_2 and a membrane-permeable biotin-derivatized tyramide, these authors were able to covalently biotinylate peroxidase-proximal intracellular proteins to be visualized or pulled down (33,34).

The known peroxidase activity of G-quadruplex-heme complexes suggests a possibility for using an analogous approach to achieve *in vitro* as well as *in vivo* tagging of G-quadruplexes as well as of G-quadruplex-bound or -proximal proteins. Phenolic compounds, including tyramine itself, have been shown to be exceptionally good substrates for G-quadruplexes complexed to hemin (heme-ribozymes and -DNAzymes) (35,36). Furthermore, there is evidence from phenolic natural products, such as Ochratoxin A, that activated phenolic radicals attack and form covalent adducts with good efficiency with mainly the C8 position of the guanine base (37). The concept is illustrated, schematically, in Figure 1.

MATERIALS AND METHODS

Oligonucleotide purification

All oligonucleotides were purchased from Integrated DNA Technologies. They were first treated with 10% aqueous piperidine at 90°C for 30 min to cleave oligonucleotides containing chemical lesions from synthesis. Following lyophilization, the treated oligonucleotides were size-purified in 8–12% denaturing gels, eluted, ethanol precipitated and dissolved in TE (10 mM Tris, pH 8.0, 0.1 mM ethylenediaminetetraacetic acid (EDTA)) buffer to give stock solutions.

CatG4 (' G4 ') *: 5′-TGG GTA GGG CGG GTT GGG AAA-21 nt* (38)

Hum4 *: 5′-TTA GGG TTA GGG TTA GGG TTA GGG-24 nt*

c-Myc *: 5′-TGA GGG TGG GGA GGG TGG GGA A-22 nt*

BLD: *5′-AAT ACG ACT CAC TAT AGG AAG AGA TGG-27 nt*

Uncat *: 5′-GTG AGG AGT GCG TGG TGA GAG-21 nt*

Duplex *: 5′-TTAGCTCACGAGACGCTCCCATAGT GA-28 bp,* bound to its complementary strand.

Footprinting G4 *: 5′-ACA TAG CTG ACT GGC TTG ATT TTG GGT AGG GCG GGT TGG GAA ATA TCG AAT TCT CAG CCT ACA CTG CAG TAC TA-74 nt*

G4-ext *: 5′-ACA TAG CTG ACT GGC TTG ATT TTG GGT AGG GCG GGT TGG GAA ATA TCG AAT TCT CAG CCT ACA CTG CAG TAC TAG TAC ATA TCA-84 nt*

PCR reverse primer *: 5′-TAG TAC TGC AGT GTA GGC TG-20 nt*

PCR forward primer *: 5′-ACA TAG CTG ACT GGC TTG AT-20 nt*

As required, oligonucleotides were ^{32}P-5′-labeled using OptiKinase (Affymetrix) in 50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT). The kinased DNA was ethanol precipitated and gel purified, eluted overnight into TE buffer and recovered by ethanol precipitation.

Biotin tyramide (bio-tyr) reaction

Biotin tyramide ('bio-tyr'; Toronto Research Chemicals) was dissolved in dimethyl sulfoxide as a 100 mM stock. Hemin or Fe(III)-heme (Frontier Scientific) stock solutions were prepared fresh in dimethyl formamide to 1 mM. H_2O_2 working solutions were diluted from a 17.4 M stock to give 100 mM in ddH_2O. DNA stock solutions were diluted in reaction buffer (final concentration: 40 mM HEPES pH 8.0, 20 mM potassium chloride, 1% dimethylformamide, 0.05% Triton X-100) and heated to 95°C for 3 min, then cooled to room temperature. Hemin was added, followed bio-tyr and the solution rested for 10 min prior to initiating the peroxidase reaction by addition of H_2O_2 to the various final concentrations described below. Reactions were allowed to proceed for 30 min, generally followed by quenching of the reaction by ethanol precipitation or by addition of 10 U bovine liver catalase enzyme (Sigma) followed by

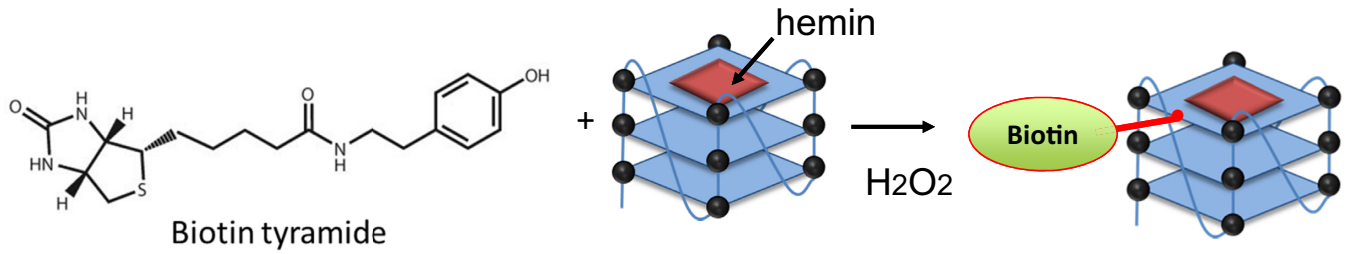


Figure 1. Schematic for covalent biotinylation of G-quadruplexes via the intrinsic peroxidase activity of G-quadruplex•heme complexes.

ethanol precipitation (as indicated). For the streptavidin gel shift experiments the recovered DNA was dissolved in a streptavidin-containing water solution for 5 min, followed by the addition of gel loading buffer. For footprinting analysis, the samples were redissolved in 10% piperidine and heated at 90°C for 30 min, vacuum dried and then treated with streptavidin and loading dye, as above.

PCR

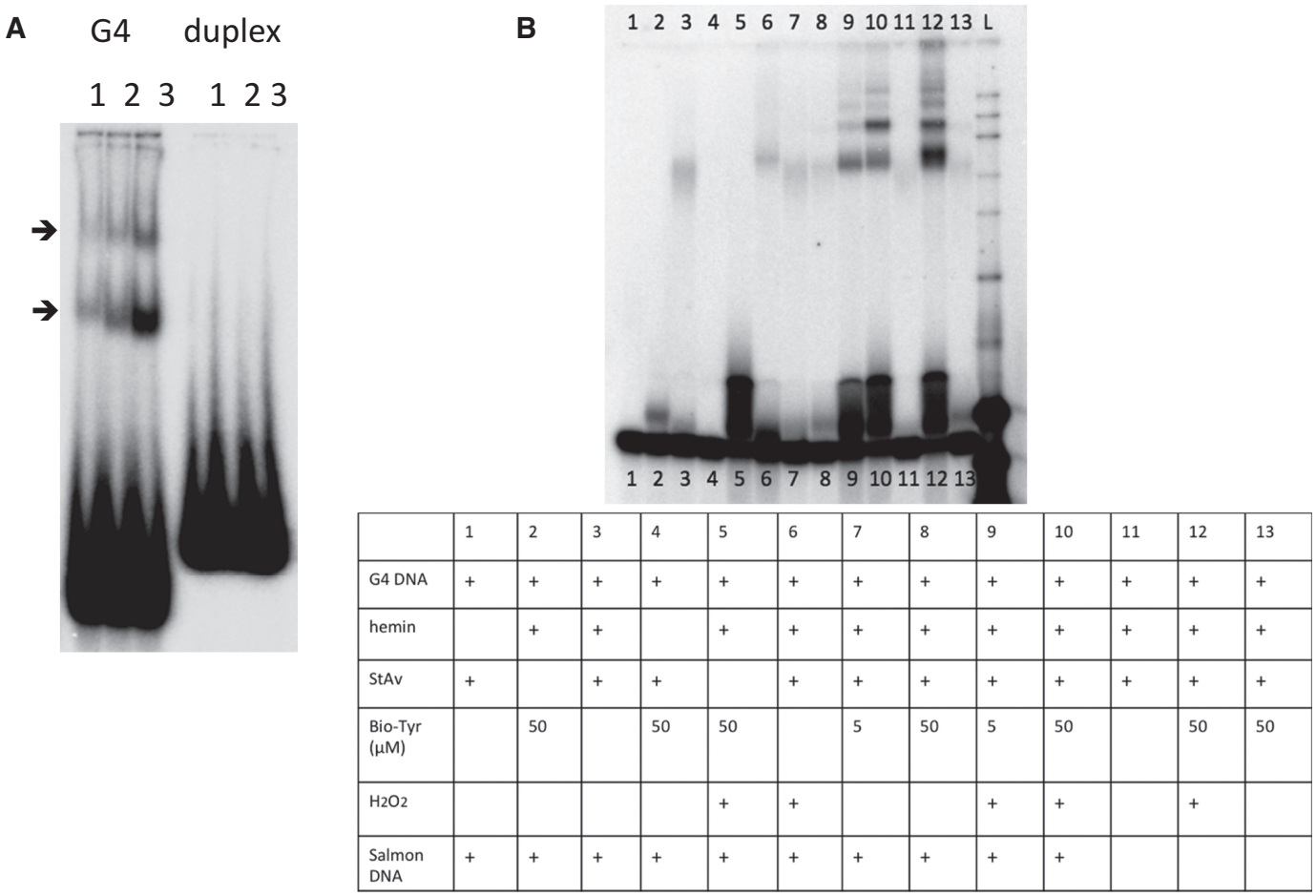
Solutions of lightly ^{32}P-5′-labeled G4-ext DNA were subjected (or not) to the bio-tyr reaction, following which the DNA, mixed with streptavidin as described above was loaded into a preparative denaturing gel. Various classes of unshifted or streptavidin-shifted DNA bands were excised from the gel, and the DNA (or streptavidin complex) eluted by crush-and-soak into TE buffer. These solutions were treated to successive butanol extractions to reduce volume, and the concentrated DNA was recovered by ethanol precipitation. To separate streptavidin-shifted biotinylated DNA from the bound streptavidin prior to polymerase chain reaction (PCR), these solutions were placed into removal solution (10 mM EDTA, 95% formamide pH 8.0) and heated to 90°C for 5–10 min prior to running in a second denaturing gel. Streptavidin-free biotinylated DNA so recovered was eluted and purified from this second gel. DNA of various categories so obtained were now subjected to a standard PCR protocol. In each case ∼20 pmol of DNA was added to a 20 μl PCR reaction containing 1 × Taq reaction buffer, 1 U of Taq polymerase enzyme and 0.2 μM each of forward and reverse primer (the reverse primer being ^{32}P-5′-labeled) and 200 μM of each of the four dNTPs. Duplicate samples were made that lacked either the forward primer or were left off the PCR machine at room temperature during the cycling. Cycling conditions were: initial denaturation for 30 s at 95°C; 20 cycles at 95°C for 30 s, 52°C for 30 s; 68°C for 45 s; final extension of 2 min at 68°C. The DNA solutions were then ethanol precipitated and the purified DNA run on a 10% denaturing gel.

RESULTS

G4 self-biotinylation using bio-tyr

Figure 2A shows the results of an experiment to test whether the peroxidase activity of heme•G-quadruplex complexes (heme-DNAzymes) can be use to biotinylate the G-quadruplex itself. The experiment was carried out with a widely studied parallel quadruplex, CatG4 (referred to as 'G4' hereafter) and a duplex DNA control. Reactions contained ^{32}P-end-labeled DNA (1 μM); 5 μM hemin; 0, 5 or 50 μM bio-tyr; and, 1 mM H_2O_2. Following 30 min of reaction at 22°C, 60 μM streptavidin (final) was added to each solution. An equal volume of 95% formamide was added, and the mixture heated to 95°C for 3 min prior to loading on an 8% denaturing polyacrylamide gel (the biotin-streptavidin interaction has been shown to be stable even under the highly denaturing conditions of the loading buffer and the 8M urea denaturing gel- 39). Figure 2A, lanes 1 show the reaction with no added bio-tyr, whereas lanes 2 and 3 show reactions containing 5 and 50 μM bio-tyr, respectively. Whereas none of the duplex lanes in Figure 2A shows any trace of a streptavidin-retarded (i.e. of significantly lower gel mobility than the free duplex DNA) band or bands, all three G4 reactions show two such gel-retarded species each. Curiously, even lane 1, which shows the products of a bio-tyr-free incubation, shows low levels of the two retarded bands, while lanes 2 and 3 show progressively higher levels of these two products, consistent with their being streptavidin-retarded biotinylated G4. The existence of two low mobility bands suggests either one streptavidin molecule has bound to one or more tagged DNA molecules, or multiple biotinylations on a single G4 molecule have recruited one or more streptavidin molecules. The mobility separation of the two shifted bands supports multiple biotinylations on a single G4; subsequent gels run against reference polynucleotide ladders show the separation corresponds to greater than the 21-nt size of G4 itself. The curious presence of low levels of the retarded bands in the 0 μM bio-tyr condition, however, suggest that in the absence of bio-tyr but in the presence of streptavidin, a secondary means for forming low levels of a denaturation-stable streptavidin-G4 composite exist. Given there has been no prior description of any intrinsic affinity between streptavidin and G-quadruplex DNAs it is likely these bands represent covalent adducts between G4 and streptavidin, enabled by the transfer of reactive radicals from the bound and activated heme to G4. Prior Electron paramagnetic resonance (EPR) studies on heme-DNAzymes have indeed shown evidence for such carbon-centered radicals on the DNA (26).

To investigate the above initial evidence in greater detail, a more systematic analysis involving varying conditions and components was carried out. The default reaction conditions for the experiments shown in Figure 2B were 1 μM G4; 1 μM heme; ∼100× mass excess of duplex salmon sperm DNA; 5 or 50 μM bio-tyr; and, 500 μM H_2O_2, in 20



	1	2	3	4	5	6	7	8	9	10	11	12	13
G4 DNA	+	+	+	+	+	+	+	+	+	+	+	+	+
hemin		+	+		+	+	+	+	+	+	+	+	+
StAv	+		+	+		+	+	+	+	+	+	+	+
Bio-Tyr (μM)		50		50	50		5	50	5	50		50	50
H_2O_2					+	+			+	+		+	
Salmon DNA	+	+	+	+	+	+	+	+	+	+			

Figure 2. (**A**) Self-biotinylation of G-quadruplex DNA (G4). Lane 1: Control (5 μM hemin; 1 μM DNA; 60 μM streptavidin added following 30 min reaction). Lane 2: 5 μM hemin; 1 μm DNA; 5 μM bio-tyr; 60 μM streptavidin added after 30 min. Lane 3: 5 μM hemin; 1 μM DNA; 50 μM bio-tyr; 500 μM H_2O_2; 60 μM streptavidin added after 30 min. (**B**) Self-biotinylation of G-quadruplex DNA. Determination of the key components for successful bio-tyr reaction-mediated G4 biotinylation.

μl reactions (a lower exposure of this gel is shown as Supplementary Figure S1). In all cases, following 30 min of reaction, streptavidin was added to ∼50 μM and the mixture left to rest for 5 min. In these experiments, as in those shown in Figure 2A, there was no explicit quenching of any residual peroxidase activity prior to the streptavidin addition step, as part of our effort to examine the nature of streptavidin-retarded G4 bands generated in the absence of added bio-tyr. Following the streptavidin incubation, 1 μl of the final solution was mixed with 5 μl of denaturing loading solution consisting of 95% formamide and dyes, heated for 2 min at 90°C, prior to loading in a denaturing polyacrylamide gel. Such a relatively harsh denaturing treatment was intended to eliminate any adventitious interaction between G4 and streptavidin. A number of the reactions were carried out in the presence of ∼100-fold mass excess of sheared salmon sperm DNA, to observe the impact of nonspecific DNA on the bio-tyr reaction.

The higher resolution of this 12% gel (Figure 2B) identifies that the streptavidin-retarded banding pattern (seen most prominently in lane 12) is not restricted to the two retarded bands identified in Figure 2A. In lane 12 at least 4 distinct retarded bands induced by the addition of streptavidin, with 2 further and faint bands visible (all streptavidin-retarded bands running above the 100-nt reference band). The presence of excess salmon sperm DNA does impact somewhat on the yield of biotinylated G4 (lane 10 versus 12, in which the shifted bands constitute 13 versus 27%, respectively, of the total DNA in those lanes). The impact of the excess duplex DNA, however, appears not be as a competitor with G4 for biotinylation. Supplementary Figure S2 (discussed in detail below) shows an experiment in which the bio-tyr reaction, carried out in solutions containing 1:1 or 1:10 G4/duplex mixtures, results in biotinylation exclusively of the G4 and not discernably of the duplex. Figure 2B, lanes 1 and 4 reconfirm the lack of any DNA mobility shift in the absence of hemin. These lanes also confirm the lack of any intrinsic binding affinity between G4 and streptavidin. The various minor products seen in lanes 2 and 3 show that even in the absence of H_2O_2, with G4, hemin and either streptavidin or bio-tyr present, low levels (1–5%) of different reaction products can be seen—in the absence of streptavidin (lane 2), this product moves with just slightly slower mobility than the unmodified G4 itself and likely represents biotinylated G4s (here, the bio-tyr reaction is likely initiated by dissolved O_2 gas, a poor but sufficient initial oxi-

dant for heme-DNAzymes—40,41). In lane 3 (lacking both H_2O_2 and bio-tyr, but with streptavidin added at 30 min) and in lane 6 (with H_2O_2 present but no bio-tyr, and with streptavidin added at 30 min) minor streptavidin-shifted products (∼1%) are seen. These products have survived the denaturing gel as well as heating in 95% formamide to 95°C prior to gel loading, and likely represent direct covalent adducts of streptavidin with G4 (EPR data have shown the presentation of DNA-localized radicals in heme-DNAzymes, which could react with streptavidin—*vide infra*). Comparison between lanes 2 and 5 shows up the variety of primary biotinylated products (unshifted by streptavidin, which is lacking in these two lanes) created in the absence and presence of H_2O_2. In the absence of H_2O_2 (lane 2), the altered mobility species run as a tight bands close the unmodified G4 itself, representing 5% of the total DNA; in the presence of H_2O_2 (lane 5) a smear representing 58% of the total lane counts is visible just above the unmodified DNA, suggesting that H_2O_2-activated bio-tyr reaction gives rise to a larger variety of biotinylated DNA species.

Reaction optimization

To characterize the efficiency of the reaction and to determine the ideal concentrations of the different reactants, the following experiments were done. First, Figure 3, upper gel, shows the bio-tyr and hemin dependencies of the reaction. For these experiments the default condition was: 1 μM G4, 1 μM hemin, 5 μM bio-tyr and 250 μM H_2O_2 (for the bio-tyr dependence measurement, bio-tyr concentrations varied from 0–25 μM; for the hemin dependence measurement, hemin concentrations varied between 0 and 5 μM). The reaction shows a strong dependence on hemin at low (0–0.5 μM) concentrations, whereas bio-tyr concentrations of >5 μM appears to give rise to quite high (20–30%) biotinylation levels. All the above dependencies were measured at more or less arbitrary *in vitro* reactant concentrations. If the bio-tyr reaction were to be carried out within live cells, the conditions required for that would need to be evaluated systematically, with perhaps the conditions used for the APEX2 and bio-tyr mediated intracellular proteomic labeling used as a starting point (32,33).

The time-scale for the bio-tyr reaction was measured for G4, under conditions of 1 μM G4, 1 μM hemin, 5 μM bio-tyr and 250 μM H_2O_2. Figure 3, lower gel, shows the data from this reaction, measured over a 0.1–30 min time interval. To calculate the level of G4 biotinylation achieved as a function of time as accurately as possible, two changes were introduced in the way that the results were assessed. First, the bio-tyr reaction was quenched at each time point by the addition of 100 U of residual H_2O_2-destroying catalase enzyme. Second, analysis was done on a native (non-denaturing) gel. The 0.1 min time point shows a reaction quenched immediately after the addition of H_2O_2. The figure shows that under these conditions the reaction plateaus within the first 5 min, as judged by a quantification of the percentage of G4 DNA shifted by streptavidin as a function of time. Under these conditions, biotinylation impacts ∼25% of the total DNA present at the start of the experiment.

Specificity and locus of biotinylation

Figure 2B (lanes 10 versus 12) showed that the presence of a large mass-excess of non-specific duplex DNA did somewhat impact the relative distribution of different biotinylated and streptavidin-shifted G4 DNA species. However, whether that effect was due to a competing and promiscuous labeling of duplex DNA was not directly explored in that experiment. We therefore carried out the bio-tyr reaction in mixtures containing a 28-bp duplex and G4 (at 1:1 μM and 10:1 μM duplex/G4 ratios), with both DNAs ^{32}P-labeled. Supplementary Figure S2 shows the results. The streptavidin shifting pattern of G4 remains the same whether it was treated in isolation (lane 8), treated in a 1:1 duplex:G4 mixture (lane 12) or in a 10:1 mixture (lane 14). Similarly, the duplex band is unchanged under the same three conditions (lanes 9, 12, 14), suggesting that there is little or no promiscuous tagging of non-quadruplex DNA, at least in these relatively dilute DNA solutions.

We attempted to determine the location of the biotin tag or tags on the 5′-^{32}P-labeled G4-ext oligonucleotide (G4-ext is a DNA oligomer that incorporates the entire sequence of G4 but has single-stranded extensions on both 5′ and 3′ ends). The goal was to generate a nucleotide ladder from this oligomer subsequent to bio-tyr mediated biotinylation, then incubating that ladder with streptavidin to determine which bands shifted up with streptavidin. As such, simply treating 5′-^{32}P-labeled G4-ext with 10% v/v piperidine at 95°C provided a low but sufficient breakdown ladder for such an analysis (heating with base leads DNA strand breakage at abasic sites generated during chemical synthesis). Supplementary Figure S3 shows that treated with streptavidin, the mobility shift of DNA bands involves those G4-ext truncations that still include the entire G4 sequence (the 5′-edge of this motif within G4-ext is indicated with an arrow in the figure). This provides evidence that bio-tyr indeed has a preference for reacting with the guanines of the G-quadruplex, and likely acts within a very short radius from the G-quadruplex-bound hemin moiety.

To define the biotinylated product better, MALDI-TOF mass spectrometry was carried out (Supplementary Figure S4). There is a clear evidence of higher molecular weight peaks in the biotinylated G4 (CatG4) sample; although, possibly owing to complex fragmentation patterns, it is not yet possible to assign specific peaks to refer unambiguously to mono-biotinylated, di-biotinylated or other species.

To what extent are other DNA sequences known to fold to G-quadruplexes capable of self-biotinylation? Supplementary Figure S5 shows results for the oligonucleotides CatG4, c-Myc (from the human c-Myc gene promoter) and Hum4 (from the human telomeric repeat), all known to form G-quadruplexes and to bind and activate hemin (28). In all cases, clear streptavidin-shifted bands can be seen, which are contingent on the presence in the reaction mixture of both hemin and bio-tyr. By contrast, BLD, an oligonucleotide incapable of forming a G-quadruplex (28), and Uncat, a permuted version of the CatG4 sequence that is also incapable of forming a G-quadruplex, do not show streptavidin-retarded bands.

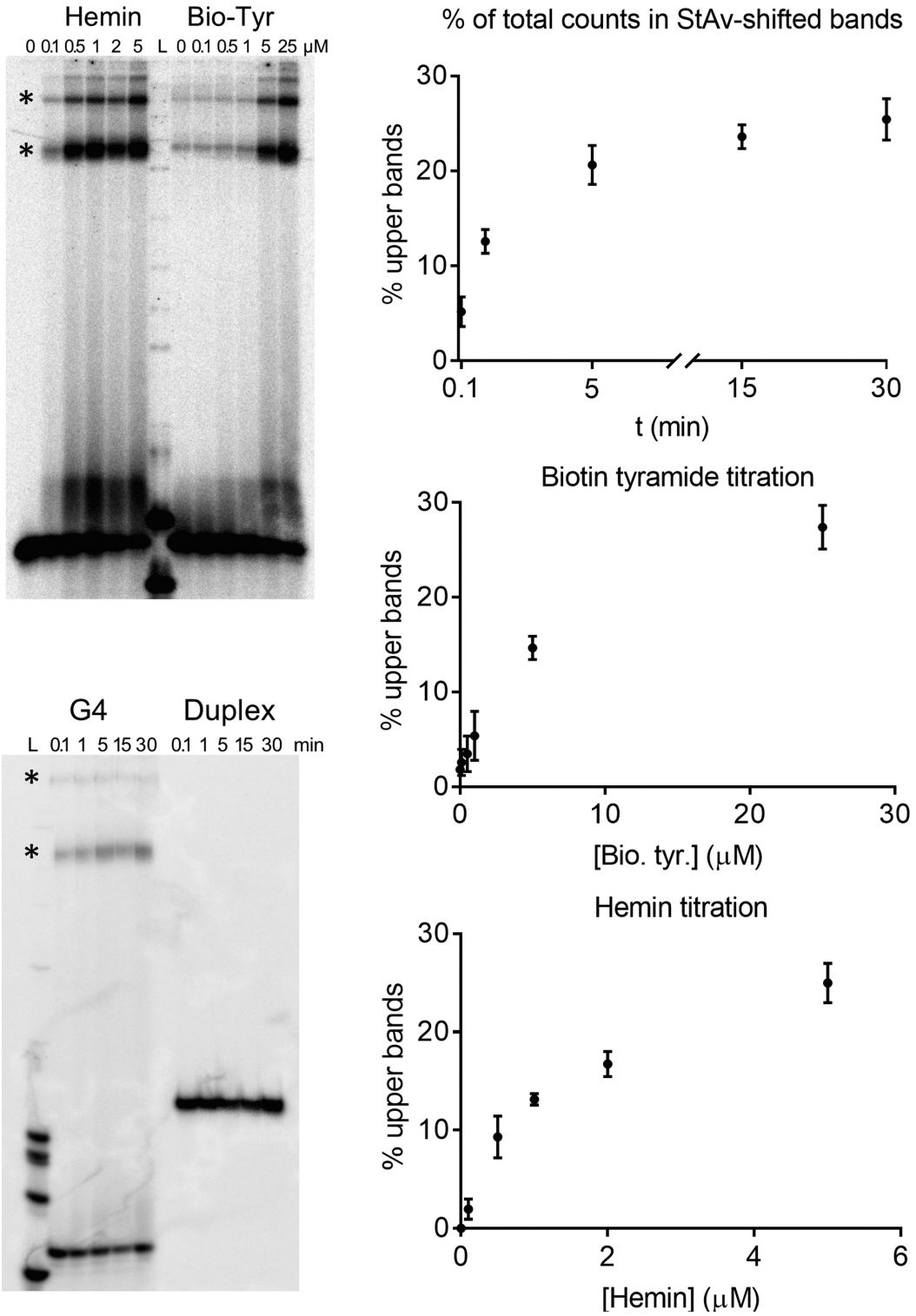


Figure 3. Hemin-dependence, biotin-tyramide dependence and time dependence for G4 biotinylation. (**Upper gel**) The default condition was 30 min reactions of 1 μM CatG4, 1 μM hemin, 5 μM bio-tyr and 250 μM H_2O_2 (for bio-tyr dependence, [bio-tyr] = 0–25 μM (plotted in the middle graph); for hemin dependence, [hemin] = 0–5 μM (plotted in the lowest graph). (**Lower gel**) the reaction contained 1 μM CatG4, 1 μM hemin, 5 μM bio-tyr and 250 μM H_2O_2. Time aliquots were quenched by the addition of catalase enzyme to degrade residual H_2O_2. These data are plotted in the top graph. All data report the mean values obtained from two independent experiments. The error bar indicates one standard deviation from the mean. The asterisks show streptavidin-shifted G4 bands.

Is bio-tyr mediated biotinylated G-quadruplex DNA amenable to PCR amplification?

In order to carry out a successful *in vivo* pulldown of a G-quadruplex forming DNA, it would be advantageous to be able to amplify such sequences given that they may be present in low abundance in the cell. The plausibility of using PCR to amplify bio-tyr generated biotinylated DNA fragments (given that this biotinylation is radical-generated, likely resulting in bio-tyr attachment at multiple and varied sites on the DNA) was investigated. This was done by subjecting the G4-ext oligomer to the bio-tyr reaction, separating the biotinylated DNA from the non-biotinylated DNA and testing PCR amplification of both (using 5′- and 3′-sequence extensions built into G4-ext as primer binding sites). For this experiment the reaction conditions were adjusted to maximize the yield of biotin-tagged sequences: 1 μm DNA, 5 μM hemin, 1 mM H_2O_2 and either low ('L': 20 μM) or high ('H': 200 μM) bio-tyr. Figure 4A and B show schematics for purification of the various biotinylated DNAs as well as unbiotinylated and bio-tyr-untreated DNA controls. An important feature of the purification protocol was its stringency, designed to cleanly separate away streptavidin-shifted DNA bands from unshifted DNA bands under denaturing conditions (i.e. heating each DNA or DNA/streptavidin sample in 95°C for 3 min in 95% formamide followed by running in a denaturing polyacrylamide gel). Following excision and elution of the various gel bands, the streptavidin shifted DNAs were stripped of streptavidin via a standard protocol of heating with formamide and EDTA, followed by re-purification from a second denaturing gel in which the biotinylated DNA now ran free of streptavidin. The various DNAs recovered in this way were then subjected to PCR amplification.

To enable and monitor consistency of template G4-ext concentration in the various PCR solutions, the following strategy was adopted: prior to the bio-tyr reaction the G4-ext DNA was ^{32}P end-labeled lightly, such that even subsequent to the various purification steps it could be ensured that in each PCR reaction approximately equal counts of template were present. To prevent confusion of this radiolabeled template with a putative, also radiolabeled PCR product, the reverse primer was designed to be slightly nested (and was strongly ^{32}P-labeled), such that the labeled PCR product was shorter (74 nt) and ran faster on a denaturing gel than the labeled G4-ext template (84 nt). Using this strategy both template and PCR product could be visualized together in the same gel.

Figure 4C shows the results of the PCR reactions. It can be seen from the lanes labeled 4.0 that successful amplification was indeed achieved with the different biotinylated G4-ext species categorized in Figure 4B. In Figure 4C, the blue arrow to the right of the gel indicates the 84-nt template band, and the red star shows the expected 74-nt PCR product. The template band (blue arrow) is smeared in the lanes showing biotinylated DNA, 4.X (where X refers to the four classes of DNA purifications, 1-4, referred to in Figure 4A), while it remains tight in the control 'untreated' group (1.X) as well as in the streptavidin-unshifted group (3.X). The smear of the template in the 4.X lanes is consistent with it being variously and multiply biotinylated. Interestingly, the yield of PCR product is relatively constant across all lanes in which such a product may be expected. Comparing the PCR product yields from 1.0, 4.0H-mono, 4.0H-dimer and 4.0H-well ('mono' signifies the highest mobility streptavidin-shifted products, 'dimer' is the band above that and 'well' is product recovered from the well of the gel—Figure 4B), no substantive difference is seen. That no major change in PCR product yield is seen from template incorporating zero to multiple biotinylations suggests that PCR protocols that use Taq polymerase are indeed compatible with covalent modifications generated by the bio-tyr reaction.

Some additional observations from this figure are: lane 1.5 shows that there are some undefined primer-only products generated even in the absence of template; however, all such minor products have mobilities distinguishable from the PCR product. A closer look at amplification of the stock G4-ext (1.0), which has not been subjected to the bio-tyr reaction or to any gel purifications, identifies an almost complete consumption of the labeled primer, which is not the case for any of the 4.X labeled conditions. The controls categorized as X.4 are effectively primer extension assays; they lack the forward primer, with extension occurring only from the reverse primer annealed to the template. These lanes should give some indication of sites on unmodified as well as biotinylated templates where extension pausing occurs. In lane 1.4 we can see a region of bands extending upward from the site marked by the black arrow. These pause can be seen across all of the PCR lanes, and appear to correspond to the G-quadruplex-forming sequence element within G4-ext. In the 2.X series, especially, pauses at the individual guanine-blocks within this region can be seen. A comparison of lane 2.4 (showing a 'bulk' bio-tyr reaction) with lane 1.4 (not subjected to the bio-tyr reaction) shows higher pausing intensities at the quadruplex guanines in lane 2.4, supporting the premise that this reflects additional impediments to primer extension linked to biotinylation at these guanines. Interestingly, the 2.4 pattern does not identify any specific biotinylation hotspots within the G-quaruplex motif.

In the above PCR reactions, we had deliberately chosen a higher template:primer ratio than is typical for a PCR reaction. In Supplementary Figure S4, we show that PCR also occurs efficiently under standard conditions from a very low mass of biotinylated template (∼150 fmol; ∼100 pg).

DISCUSSION

We have shown here a strategy for exploiting the intrinsic peroxidase activity of G-quadrupex•hemin complexes, in the presence of hydrogen peroxide and biotin tyramide, to self-biotinylate the DNA. Such DNA-biotinylation occurs with good efficiency *in vitro*, and is specific to G-quadruplexes and not to duplex DNA. G-quadruplexes so biotinylated nevertheless remain amenable to PCR amplification.

In what ways could such a strategy be useful *in vivo*? Two distinct if related questions could potentially be addressed using this approach. The first would be to explore the hypothesis that cellular heme does indeed bind to, and is activated by, RNA and DNA G-quadruplexes in the cell—whether under homeostatic conditions or in differ-

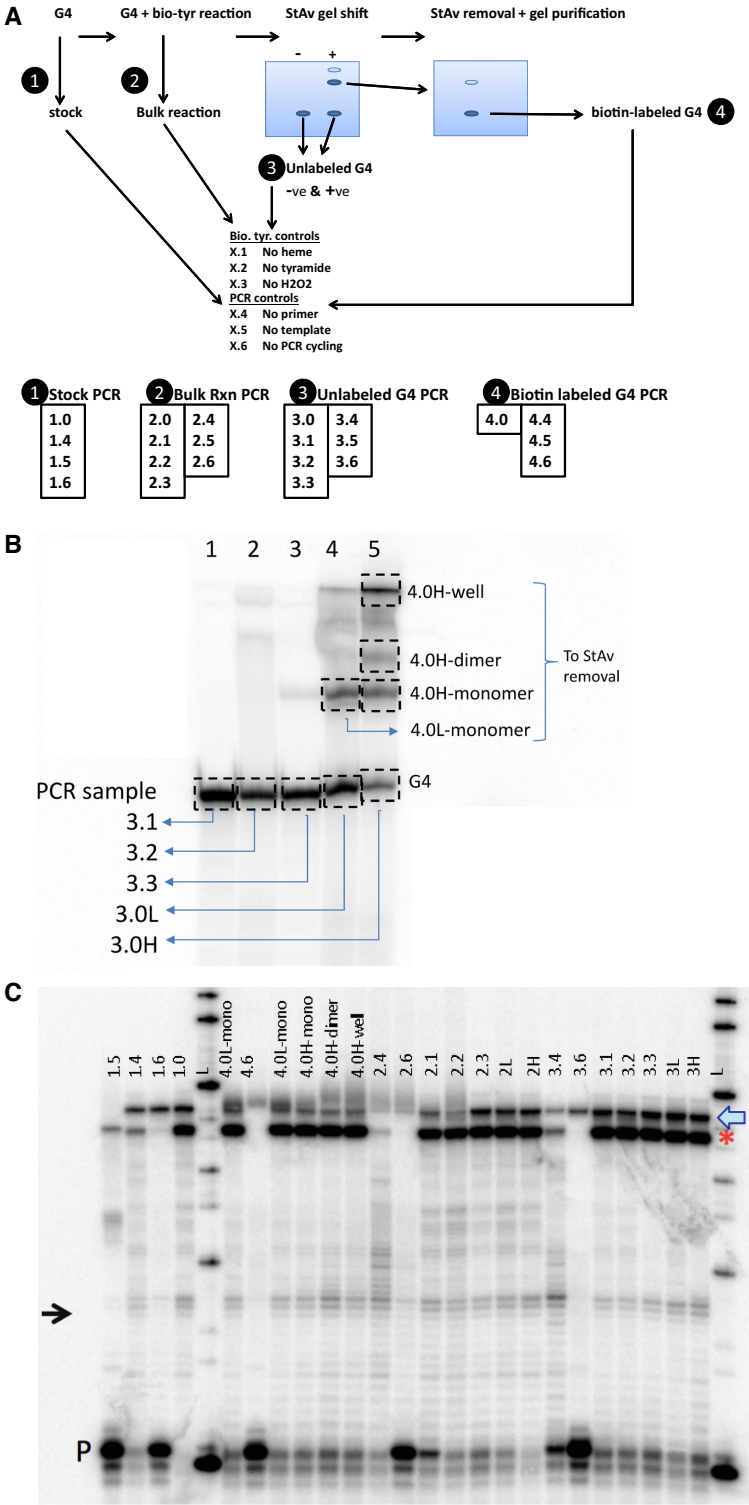


Figure 4. Investigation of whether bio-tyr-mediated covalent biotinylation of G4 DNA enables such DNA to be amplified by PCR. (**A**) A scheme for treatment of G4 with bio-tyr to generate biotinylated DNAs and also various unbiotinylated controls. In block (1) G4 'stock' refers to G4 DNA that has not been subjected to a bio-tyr-mediated biotinylation reaction. Block (2) 'bulk reaction' refers to G4 DNA that has subjected to the bio-tyr reaction, but not subsequently subjected to a purification protocol to separate out biotinylated from unbiotinylated DNA. shows 'L' and 'H' refer to low (20 μM) and high (200 μM) concentrations of bio-tyr, respectively. (**B**) A further elaboration on blocks (3) and (4). In this preparatory gel, lane 1 shows bio-tyr reaction conditions but excluding hemin; lane 2 shows the reaction excluding bio-tyr; lane 3 excludes H_2O_2; lane 4 shows the full reaction with all components, including 20 μM bio-tyr; lane 5 shows full reaction but with 200 μM bio-tyr. (**C**) Gel showing the results of experiment schematized in A and B. The thick arrow to the right of the gel indicates the lightly ^{32}P-labeled template (84 nt) whereas the star to the right indicates the PCR-amplified product (74 nt). 'P' refers to the ^{32}P-labeled reverse primer. 'L' shows a reference ladder. The arrow on the left shows the 5′ end of the guanine-rich motif that forms the G-quadruplex.

ent disease states. This hypothesis is based on the observed ubiquity of heme binding to G-quadruplexes *in vitro*, the demonstrated existence of DNA and RNA G-quadruplexes *in vivo*, and the presence of low concentrations of heme in most intracellular compartments. Increasingly it has been clear that heme's involvement in different aspects of biology is in fact very diverse (42). In addition to serving as an enzymatic cofactor, participating in electron transfer, and as a gas transporter, it participates in gene regulation, intracellular sensing, as well as perhaps other not yet elucidated functions in the cell. While intracellular traffic of heme is tightly controlled, low concentrations of exchangeable heme ('regulatory heme', defined as heme that is transported via transient protein or other carriers) are found in most intracellular compartments and organelles, including the cytosol and the nucleus (43–45). Given the submicromolar dissociation constants of heme• G-quadruplex complexes, it is plausible that cytoplasmic, nuclear or mitochondrial DNA/RNA G-quadruplexes could also participate in the traffic of regulatory heme within the cell.

A second general utility for this methodology could be to target and label intracellular G-quadruplexes as a whole, for either visualization or pull-down. DNA or RNA sequences pulled down via self-biotinylation should be amenable to analysis by ChIP-Seq or RNA-Seq (although with RNA, it might be necessary to test which reverse transcriptases are able to efficiently process biotinylated RNAs). It is possible to imagine that if our first hypothesis (see above) of heme-G-quaduplex complexation *in vivo* is either incorrect or found to occur at insufficient levels, this second application, dependent on the first, may be compromised. However, there are valid experimental approaches to circumvent that potential problem. Strong evidence exists that supplementation of cell growth media with either heme itself or heme complexed with bovine serum albumin leads to significant heme uptake into cells in culture; indeed, such supplementation has been shown to bolster different heme-utilizing cellular enzymes from a state of poor or non-existent catalytic activity to high catalytic activity (46,47). Analogously, therefore, it should be possible to enhance complexation of hemin with cellular G-quadruplexes by means of supplementation of growth media with hemin.

One final issue concerns the generality of this method for capturing all possible DNA G-quadruplex topologies (RNA forms exclusively all-parallel strand orientations that favor hemin binding) that may occur in living cells. Overall, all-parallel or mixed-orientation DNA G-quadruplexes have been shown to be best capable of binding and activating hemin (36). Therefore, it may be that certain antiparallel quadruplex topologies with loop structures that sterically interfere with hemin-binding will be under-represented in pulldowns from cells. However, it is likely that adjusting the concentrations of the active reactants for the bio-tyr reaction may show up these particular G-quadruplexes as well, since they are still better binders and activators of hemin compared to standard duplex DNA (36).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Natural Sciences and Engineering Research Council of Canada (NSERC). Funding for open access charge: NSERC.

Conflict of interest statement. None declared.

REFERENCES

1. 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.
2. Lipps,H.J. and Rhodes,D. (2009) G-quadruplex structures, in vivo evidence and function. *Trends Cell Biol.*, **19**, 414–422.
3. Wu,Y. and Brosh,R.M. Jr (2010) G-quadruplex nucleic acids and human disease. *FEBS J.*, **277**, 3470–3488.
4. Hänsel-Hertsch,R., Antonio,M.D. and Balasubramanian,S. (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell. Biol.*, **18**, 279–284.
5. Rouleau,S., Jodoin,R., Garant,J.M. and Perreault,J.P. (2017) RNA G-quadruplexes as key motifs of the transcriptome. *Adv. Biochem. Eng. Biotechnol.*, doi:10.1007/10_2017_8.
6. Gellert,M., Lipsett,M.N. and Davies,D.R. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 2013–2018.
7. Ralph,R.K., Connors,W.J. and Khorana,H.G. (1962) Secondary structure and aggregation in deoxyguanosine oligonucleotides. *J. Am. Chem. Soc.*, **84**, 2265–2266.
8. Sen,D. and Gilbert,W. (1990) A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature*, **344**, 410–414.
9. Schaffitzel,C., Berger,I., Postberg,J., Hanes,J., Lipps,H.J. and Plückthun,A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 8572–8577.
10. 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.
11. Henderson,A., Wu,Y., Huang,Y.C., Chavez,E.A., Platt,J., Johnson,F.B., Brosh,R.M. Jr, Sen,D. and Lansdorp,P.M. (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.*, **42**, 860–869.
12. Hänsel-Hertsch,R., Beraldi,D., Lensing,S.V., Marsico,G., Zyner,K., Parry,A., Di Antonio,M., Pike,J., Kimura,H., Narita,M. *et al.* (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, **48**, 1267–1272.
13. Paeschke,K., Capra,J.A. and Zakian,V.A. (2011) DNA replication through G-quadruplex motifs is promoted by the S. cerevisiae Pif1 DNA helicase. *Cell*, **145**, 678–691.
14. Rodriguez,R., Miller,K.M., Forment,J.V., Bradshaw,C.R., Nikan,M., Britton,S., Oelschlaegel,T., Xhemalce,B., Balasubramanian,S. and Jackson,S.P. (2012) Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.*, **8**, 301–310.
15. Chambers,V.S., Marsico,G., Boutell,J.M., Di Antonio,M., Smith,G.P. and Balasubramanian,S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.*, **33**, 877–881.
16. Huppert,J.L. and Balasubramanian,S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
17. Rhodes,D. and Lipps,H.J. (2015) G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.*, **43**, 8627–8637.
18. DeJesus-Hernandez,M., Mackenzie,I.R., Boeve,B.F., Boxer,A.L., Baker,M., Rutherford,N.J., Nicholson,A.M., Finch,N.A., Flynn,H., Adamson,J. *et al.* (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*, **72**, 245–256.
19. Renton,A.E., Majounie,E., Waite,A., Simon-Sanchez,J., Rollinson,S., Gibbs,J.R., Schymick,J.C., Laaksovirta,H., van Swieten,J.C., Myllykangas,L. and ITALSGEN Consortium (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, **72**, 257–268.
20. Fratta,P., Mizielinska,S., Nicoll,A.J., Zloh,M., Fisher,E.M., Parkinson,G. and Isaacs,A.M. (2012). C9orf72 hexanucleotide repeat

associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.*, **2**, 1016.
21. Reddy,K., Zamiri,B., Stanley,S.Y., Macgregor,R.B. Jr and Pearson,C.E. (2013). The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.*, **288**, 9860–9866.
22. Haeusler,A.R., Donnelly,C.J., Periz,G., Simko,E.A., Shaw,P.G., Kim,M.S., Maragakis,N.J., Troncoso,J.C., Pandey,A., Sattler,R. *et al.* (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, **507**, 195–200.
23. Li,Y., Geyer,C.R. and Sen,D. (1996) Recognition of anionic porphyrins by DNA aptamers. *Biochemistry*, **35**, 6911–6922.
24. Travascio,P., Li,Y. and Sen,D. (1998) DNA-enhanced peroxidase activity of a DNA aptamer-hemin complex. *Chem. Biol.*, **5**, 505–517.
25. Travascio,P., Bennet,A.J., Wang,D.Y. and Sen,D. (1999) A ribozyme and a catalytic DNA with peroxidase activity: active sites versus cofactor-binding sites. *Chem. Biol.*, **6**, 779–787.
26. Travascio,P., Witting,P.K., Mauk,A.G. and Sen,D. (2001) The peroxidase activity of a hemin–dna oligonucleotide complex: free radical damage to specific guanine bases of the DNA. *J. Am. Chem. Soc.*, **123**, 1337–1348.
27. Poon,L.C., Methot,S.P., Morabi-Pazooki,W., Pio,F., Bennet,A.J. and Sen,D. (2011) Guanine-rich RNAs and DNAs that bind heme robustly catalyze oxygen transfer reactions. *J. Am. Chem. Soc.*, **133**, 1877–1884.
28. Sen,D. and Poon,L.C. (2011) RNA and DNA complexes with hemin [Fe(III) heme] are efficient peroxidases and peroxygenases: how do they do it and what does it mean? *Crit. Rev. Biochem. Mol. Biol.*, **46**, 478–492.
29. Li,T., Dong,S. and Wang,E. (2009) G-quadruplex aptamers with peroxidase-like DNAzyme functions, which is the best and how does it work? *Chem. Asian J.*, **4**, 918–922.
30. Faget,L. and Hnasko,T.S. (2015) Tyramide signal amplification for immunofluorescent enhancement. *Methods Mol. Biol.*, **1318**, 161–172.
31. Bobrow,M.N., Litt,G.J., Shaughnessy,K.J., Mayer,P.C. and Conlon,J. (1992) The use of catalyzed reporter deposition as a means of signal amplification in a variety of formats. *J. Immunol. Methods*, **150**, 145–149.
32. Stack,E.C., Wang,C., Roman,K.A. and Hoyt,C.C. (2014) Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods*, **70**, 46–58.
33. Hung,V., Udeshi,N.D., Lam,S.S., Loh,K.H., Cox,K.J., Pedram,K., Carr,S.A. and Ting,A.Y. (2016) Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.*, **11**, 456–475.
34. Rhee,H.-W., Zou,P., Udeshi,N.D., Martell,J.D., Mootha,V.K., Carr,S.A. and Ting,A.Y. (2013) Proteomic mapping of mitochondria in living cells via spatially-restricted enzymatic tagging. *Science*, **339**, 1328–1331.
35. Rojas,A.M., Gonzalez,P.A., Antipov,E. and Klibanov,A.M. (2007) Specificity of a DNA-based (DNAzyme) peroxidative biocatalyst. *Biotechnol. Lett.*, **29**, 227–232.
36. Nakayama,S., Wang,J. and Sintim,H.O. (2011) DNA-based peroxidation catalyst—what is the exact role of topology on catalysis and is there a special binding site for catalysis? *Chemistry*, **17**, 5691–5698.
37. Dai,J., Wright,M.W. and Manderville,R.A. (2003) Ochratoxin a forms a carbon-bonded c8-deoxyguanosine nucleoside adduct: implications for C8 reactivity by a phenolic radical. *J. Am. Chem. Soc.*, **125**, 3716–3717.
38. Kong,D.M., Cai,L.L., Guo,J.H., Wu,J. and Shen,H.X. (2009) Characterization of the G-quadruplex structure of a catalytic DNA with peroxidase activity. *Biopolymers*, **91**, 331–339.
39. Ulanovsky,L., Drouin,G. and Gilbert,W. (1990) DNA trapping electrophoresis. *Nature*, **343**, 190–192.
40. Golub,E., Freeman,R. and Willner,I. (2011) A hemin/G-quadruplex acts as an NADH oxidase and NADH peroxidase mimicking DNAzyme. *Angew. Chem. Int. Ed.*, **50**, 1–6.
41. Grigg,J.C., Shumayrikh,N. and Sen,D. (2014) G-quadruplex structures formed by expanded hexanucleotide repeat RNA and DNA from the neurodegenerative disease-linked C9orf72 gene efficiently sequester and activate heme. *PLOS One*, **9**, e106449.
42. Smith,A.G., Raven,E.L. and Chernova,T. (2011) The regulatory role of heme in neurons. *Metallomics*, **3**, 955–962.
43. Atamna,H., Killilea,D.W., Killilea,A.N. and Ames,B.N., (2002) Heme deficiency may be a factor in the mitochondrial and neuronal decay of aging. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 14807–14812.
44. Atamna,H. (2004) Heme, iron, and the mitochondrial decay of ageing. *Ageing Res. Rev.*, **3**, 303–318.
45. Atamna,H. (2006). Heme binding to Amyloid-beta peptide: mechanistic role in Alzheimer's disease. *J. Alzheimers Dis.*, **10**, 255–266.
46. Richards,M.K. and Marletta,M.A. (1994) Characterization of neuronal nitric oxide synthase and a C4 15H mutant, purified from a baculovirus overexpression system, *Biochemistry*, **33**, 14723–14732.
47. Martell,J.D., Deerinck,T.J., Sancak,Y., Poulos,T.L., Mootha,V.K., Sosinsky,G.E., Ellisman,M.H. and Ting,A.Y. (2012) Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotech.*, **30**, 1143–1148.