Chemistry & Biology
Article

A Matter of Location: Influence of G-Quadruplexes on *Escherichia coli* Gene Expression

Isabelle T. Holder[1] and Jörg S. Hartig[1,*]
[1]Department of Chemistry and Konstanz Research School Chemical Biology (KoRS-CB), University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany
*Correspondence: joerg.hartig@uni-konstanz.de
http://dx.doi.org/10.1016/j.chembiol.2014.09.014

SUMMARY

We provide important insights into secondary-structure-mediated regulation of gene expression in *Escherichia coli.* In a comprehensive survey, we show that the strand orientation and the exact position of a G-quadruplex sequence strongly influence its effect on transcription and translation. We generated a series of reporter gene constructs that contained systematically varied positions of quadruplexes and respective control sequences inserted into several positions within the promoter, 5′-UTR, and 3′-UTR regions. G-rich sequences at specific locations in the promoter and also in proximity to the ribosome-binding site (RBS) showed pronounced inhibitory effects. Additionally, we rationally designed a system where quadruplex formation showed a gene-activating behavior. Moreover, we characterized quadruplexes in proximity to the RBS that occur naturally in *E. coli* genes, demonstrating that some of these quadruplexes exert significant modulation of gene expression. Taken together, our data show strong position-dependent effects of quadruplex secondary structures on bacterial gene expression.

INTRODUCTION

Repetitive guanosine-rich nucleic acid sequences are prone to adopt G-quadruplex structures. Quadruplexes are composed of at least two tetrad layers that stack on each other via π-π interactions. Tetrads are composed of four guanines stabilized by Hoogsteen base pairs in a coplanar arrangement (Gellert et al., 1962) (Figure 1A). The stability of the structures varies with the sequence and size of the loops: the smaller the loop, the more stable the G-quadruplex (Huppert, 2010). Furthermore, G-tract length and the presence of monovalent cations, such as Na^+ and especially K^+, influence G-quadruplex stability (Burge et al., 2006; Huppert, 2008). Whereas DNA quadruplexes can adopt different topologies (parallel, antiparallel, or hybrid structures are known; Figures 1B–1D), RNA quadruplexes are exclusively found in the parallel conformation (Joachimi et al., 2009). We and others have shown that quadruplex structures adopted by RNA sequences are equally or even more stable than their respective DNA counterparts (Arora and Maiti, 2009; Joachimi et al., 2009; Saccà et al., 2005; Zhang et al., 2010). Several studies searched genomes for quadruplexes in human (Du et al., 2009; Huppert and Balasubramanian, 2005), yeast (Capra et al., 2010; Hershman et al., 2008), and bacteria (Rawal et al., 2006; Yadav et al., 2008) and found them overrepresented in certain regulatory elements such as promoter-proximal regions (Capra et al., 2010; Huppert and Balasubramanian, 2005; Rawal et al., 2006; Verma et al., 2008), nuclease hypersensitive sites (Huppert and Balasubramanian, 2007), CpG islands, enhancers, insulators (Du et al., 2009), and conserved elements, such as transcription factor binding sites (Eddy and Maizels, 2008; Nakken et al., 2009). They further occur within micro- and minisatellites (Bacolla et al., 2008; Sawaya et al., 2013), ribosomal DNA (Hanakahi et al., 1999), and telomeric DNA (Hershman et al., 2008; Wang and Patel, 1993). In two successive studies, Balasubramanian and colleagues described specific immunostaining of DNA (Lam et al., 2013) and RNA (Biffi et al., 2014) quadruplexes. These findings suggest that G-quadruplexes have important functions in cellular and genetic processes (Bochman et al., 2012). Potential quadruplex-forming sequences were identified in the promoter regions of many human proto-oncogenes. Specific effects on quadruplex-containing genes have been shown on treatment of mammalian cells with quadruplex-selective compounds (Halder et al., 2012b; Siddiqui-Jain et al., 2002; Verma et al., 2009). Recent studies investigated the influence of transcription-induced quadruplex formation in a double-stranded DNA (dsDNA) template. Quadruplexes can be induced far away from a transcription start site (TSS), functioning as silencer or enhancer of transcription (Perrone et al., 2013; Zhang et al., 2013). In addition to 5′ untranslated regions (UTRs), potential quadruplexes have also been identified in the 3′-UTRs near transcription termination, splicing, and polyadenylation sites (Huppert et al., 2008; Beaudoin and Perreault, 2013). Quadruplexes have also been described to interfere with translation. Different RNA quadruplexes have been linked to translational processing of human transcripts. We used a luciferase reporter assay to demonstrate that artificial RNA quadruplex sequences inserted into the 5′-UTR can act as translational suppressors in mammalian cells (Halder et al., 2009, 2012a).

Most studies describe the influence of quadruplexes on gene regulation in eukaryotic cells. Clues to potential functions of quadruplex sequences in bacteria are rare, although some very specific roles of quadruplexes have been described. In the pathogen *Neisseria gonorrhoeae*, pilin antigenic variation is necessary to evade the human immune system. Seifert and colleagues showed that a quadruplex forming upstream of the *pilE* gene

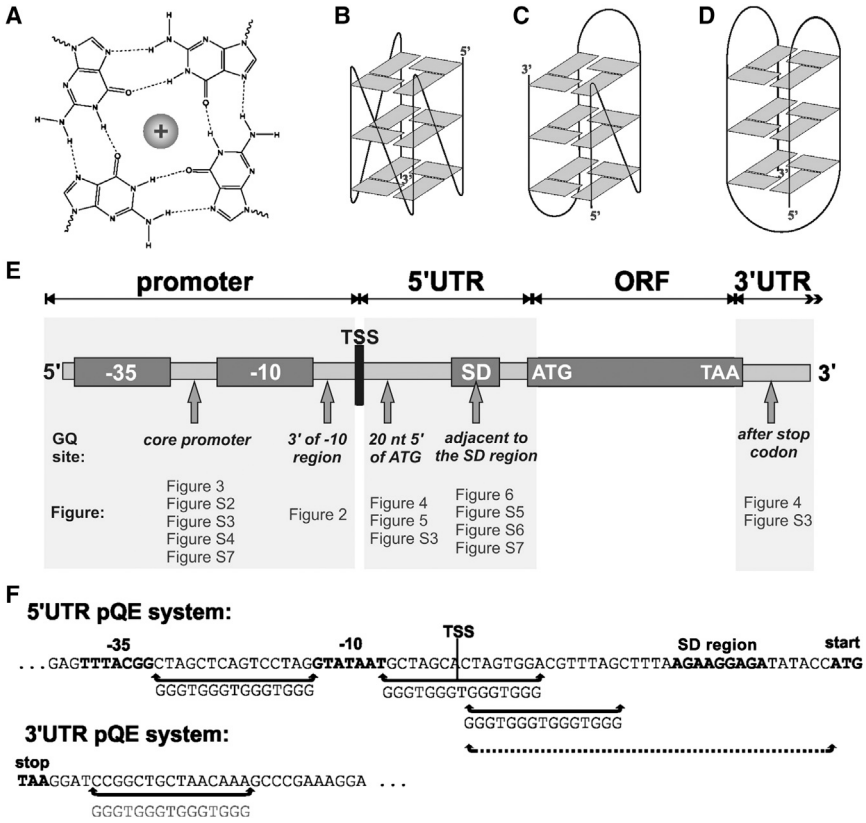


Figure 1. Schematic Representations

(A) G-quadruplex coplanar arrangement of guanines by noncanonical Hoogsteen base pairing. The tetrad can be stabilized by monovalent cations.

(B–D) Different quadruplex conformations: (B) parallel propeller type; (C) (3 + 1) hybrid; (D) antiparallel basket type.

(E) Schematic representation of G-quadruplex insertion sites investigated in this study. −35 and −10 represent the conserved promoter regions. ATG and TAA are representative for either the start or the stop codon of the reporter gene. Arrows indicate the sites that have been replaced by G-quadruplex-forming sequences or their respective controls in this study. G-quadruplexes have been investigated both on sense and on antisense strands.

(F) Nucleotide sequence of the sense strand in the 5′-UTR and in the 3′-UTR of the pQE reporter system. Sequences that have been replaced by G-rich elements 15 nt long are indicated by lines; when only 11 nt were replaced (G2T), the first and the last 2 nt of the indicated sequence were not changed relative to wild-type. Dashed line indicates range that has been sequence modified for investigation of G-quadruplex influence adjacent to the SD region.

locus induces nonhomologous recombination between the *pilE* locus and many silent *pilS* donor loci, thereby facilitating antigenic variation (Cahoon and Seifert, 2009, 2011). An involvement of quadruplex structures in antigenic variation was also suggested in *Borrelia burgdorferi* (Walia and Chaconas, 2013) and *Treponema pallidum* (Giacani et al., 2012). One of those motifs was recently characterized by our group and found to be significantly enriched in bacteria (Rehm et al., 2014). Apart from intergenically occurring quadruplexes, four-stranded motifs also occur in open reading frames (ORFs). Sugimoto and colleagues described the influence of G-quadruplexes occurring in protein-coding sequences (Endoh et al., 2013a, 2013b, 2013c; Endoh and Sugimoto, 2013). It seems that quadruplex formation increases the potential of inducing ribosomal stalling and frameshifting (Murat et al., 2014; Yu et al., 2014). In an artificial setup, we showed that translation is strongly influenced by masking the ribosome-binding site of *E. coli* with a quadruplex motif (Wieland and Hartig, 2007). Furthermore, in a computational search, Chowdhury and colleagues found quadruplexes enriched in bacterial promoters across more than 140 bacterial species (Rawal et al., 2006; Yadav et al., 2008). Regarding *Deinococcus radiodurans*, putative G-quadruplex-forming sequences were found specifically in correlation to radioresistance genes. Quadruplex stabilization via addition of small molecule compounds led to the attenuation of radioresistance in vivo (Beaume et al., 2013).

The bacterial studies suggest that G-quadruplexes play important roles in regulatory processes. However, no comprehensive study of quadruplexes influencing bacterial gene expression is available. Here, we systematically investigate quadruplexes at different genetic positions in *E. coli*. We focus on the modulation of gene expression by inserting quadruplexes of different stabilities into promoter sites, 5′-UTRs, or 3′-UTRs of a reporter gene (scheme in Figure 1E). While quadruplexes in the 3′-UTR have no significant influence on gene expression, we found that G-quadruplexes in the 5′-UTR strongly affect expression. In addition to studying artificial model sequences, we investigate quadruplex-forming sequences found in the *E. coli* genome surrounding the Shine-Dalgarno region (SD region) and demonstrate their ability to modulate gene expression in the natural context.

RESULTS

We have designed a series of sequences that contain G-rich elements around the promoter and UTRs of a reporter gene coding for eGFP, schematically shown in Figure 1E. In our studies, we used potential G-quadruplex sequences of different stabilities and their respective nonquadruplex controls. We performed circular dichroism (CD) spectroscopy using synthetic DNA oligonucleotides (sequences listed in Figure 2A). Spectra were measured in buffers containing 100 mM KCl, a concentration in the physiological range for *E. coli* (Epstein and Schultz, 1965) (Figure 2B). The G_3T, G_3A and G_2T sequences fold into parallel G-quadruplexes, showing a typical maximum signal at 265 nm and a minimum at 240 nm. The G_2CT sequence folds into an antiparallel G-quadruplex structure, with a maximum peak at 290 nm and a minimum at 265 nm. CD signatures decrease from quadruplexes with three G-tetrads to those of two tetrads. The sequence G_2CT shows little CD intensities. In order to

A

Name	Sequence	Tm [°C]
G_3T	GGGTGGGTGGGTGGG	> 95
G_3A	GGGAGGGAGGGAGGG	> 95
G_2T	GGTGGTGGTGG	54.4
G_2CT	GGCTGGCTGGCTGG	51.4
ctrl 1	GGGTGTGTGTGTGTG	< 20
ctrl 2	CACTCACTCACTCCC	< 20

B

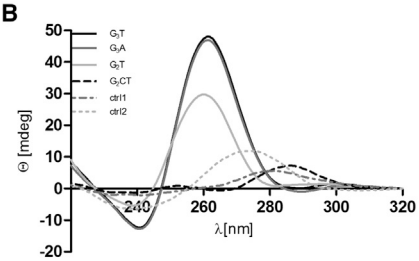


Figure 2. In Vitro Characterization of G-Quadruplex Sequences and Controls

(A) Name and sequence of quadruplex constructs. Guanines participating in G-quadruplex formation are underlined. Tm is also indicated. ctrl, control.

(B) CD spectra of 5 μM DNA in the presence of 100 mM KCl.

determine the stabilities of the quadruplexes, thermal denaturation was measured at 265 nm and 290 nm (Figure 2A; melting curves in Figure S2 available online). We determined melting temperatures (Tms) of 54.4°C and 51.4°C for the G_2T and G_2CT constructs. The Tms of G_3T and G_3A could not be accurately determined as they were very thermostable and started to denature only above 80°C. Controls 1 and 2 (ctrl1 and ctrl2, respectively) melted immediately with a Tm lower than 20°C, indicating that they are not able to form stable structures.

In the first set of constructs, the G_3T sequence was placed either in the core promoter (between the conserved −10 and −35 promoter regions) or immediately at the 3′ end of the −10 region (the sequence replaced in contrast to the wild-type promoter of the pQE vector system is shown in Figure 1F). For each set, G-tracts were placed once in the sense strand and once in the antisense strand. The total number of nucleotides between the conserved regions and the conserved sequences themselves were not changed compared to the wild-type promoter (Figures 1E and 1F). We investigated the influence of the quadruplex-forming sequence on eGFP expression in two different plasmid systems: (1) the pQE-J06-eGFP system with the constitutive J06 promoter (Figure 1F) and (2) the pBAD-eGFP plasmid (based on pBAD-18 with eGFP reporter gene; Guzman et al., 1995) with the arabinose-inducible araBAD promoter region (Figure S1). We chose to analyze two vector systems to better identify plasmid- or promoter-specific effects that are not necessarily triggered by secondary structure formation. Quadruplex insertion between the −10 and −35 regions of the J06 promoter resulted in a decrease of 86% in gene expression (Figure 3A) compared to the control sequence. A less extensive decrease (42%) was observed for the quadruplex construct compared to ctrl1 at the same position in the araBAD promoter (Figure 3B). In this case, the quadruplex-induced effect might be influenced by the arabinose-induced binding of AraC from position −35 to position −51 (Lee et al., 1987; Niland et al., 1996). However, in general, the same effects for both promoter systems were observed: quadruplexes between the −10 and −35 regions in the sense strand did not significantly influence gene expression; quadruplexes downstream of the −10 region in the sense strand decreased gene expression, and those in the antisense strand between the −10 and −35 regions showed the highest decrease in gene expression compared to the respective ctrl1. Interestingly, when G-rich sequences were located downstream of the −10 region in the antisense strand, we observed a significant increase in gene expression compared to the nonquadruplex control for the J06 promoter construct. However, the overall expression level for this construct decreased in comparison to the wild-type vector. As sequence changes have a huge influence in that region, it might be possible that they alone cause alterations in gene expression. In addition, the quadruplex inserted at the same position in the araBAD promoter region showed no significant change in gene expression when compared to control. Hence, the effect cannot clearly be attributed to G-quadruplex formation.

Influence of Quadruplexes on the Antisense Strand of the Core Promoter

The significant decrease in gene expression with the quadruplex located between the conserved promoter regions on the antisense strand (Figure 3A) motivated us to study this construct in more detail. We focused on the constitutively expressing J06 promoter. Given that sequence changes in these highly regulatory regions might have a huge influence on gene expression (Roberts and Roberts, 1996), we designed additional nonquadruplex-forming controls. Furthermore, we inserted G-quadruplexes of different stabilities ranging from G_3T (reported as being the most stable G-quadruplex structure; Bugaut and Balasubramanian, 2008) to less stable quadruplexes comprising two tetrads and longer loops (Figures 2A and 2B). As expected, less stable G-quadruplexes repressed gene expression less effectively compared to thermodynamically stable ones. However, all tested sequences resulted in reduced gene expression compared to ctrl1 and ctrl2, as well as the wild-type system. In comparison to ctrl1, the constructs G_3A, G_2T, and G_2CT repressed gene expression by 72.3%, 59.7%, and 37.6% (Figure 3C). As the quadruplex is located in front of the TSS, it should not be located on the messenger RNA (mRNA). To assay effects of the quadruplex on transcription, we next performed in vitro transcription reactions with *E. coli* RNA polymerase (Figure 3D). For the most stable quadruplex, G_3T, we observed almost no full-length transcription product. Also, the G_3A construct is less transcribed compared to G_2T and controls. Additionally, we analyzed eGFP mRNA levels in vivo by semiquantitative RT-PCR. Relative mRNA levels of quadruplex-containing constructs are reduced relative to controls: 93.7%, 76.0%, and 64% for G_3T, G_3A, and G_2T constructs, respectively, compared to ctrl1 (Figure 3E). Hence, it seems that, at the investigated position, quadruplex formation efficiently inhibits transcription.

In an attempt to detect quadruplex formation in the living bacterium, we performed in vivo DNA footprinting with dimethyl sulfate (DMS). Bacteria were incubated with DMS that selectively methylates guanines at the N7 position. The inserted G-rich stretch should be protected from methylation when engaged in Hoogsteen interactions in the quadruplex structure, whereas N7 should be accessible for methylation in the duplex form. Figure S4 shows the footprinting reaction of the G_3T construct in

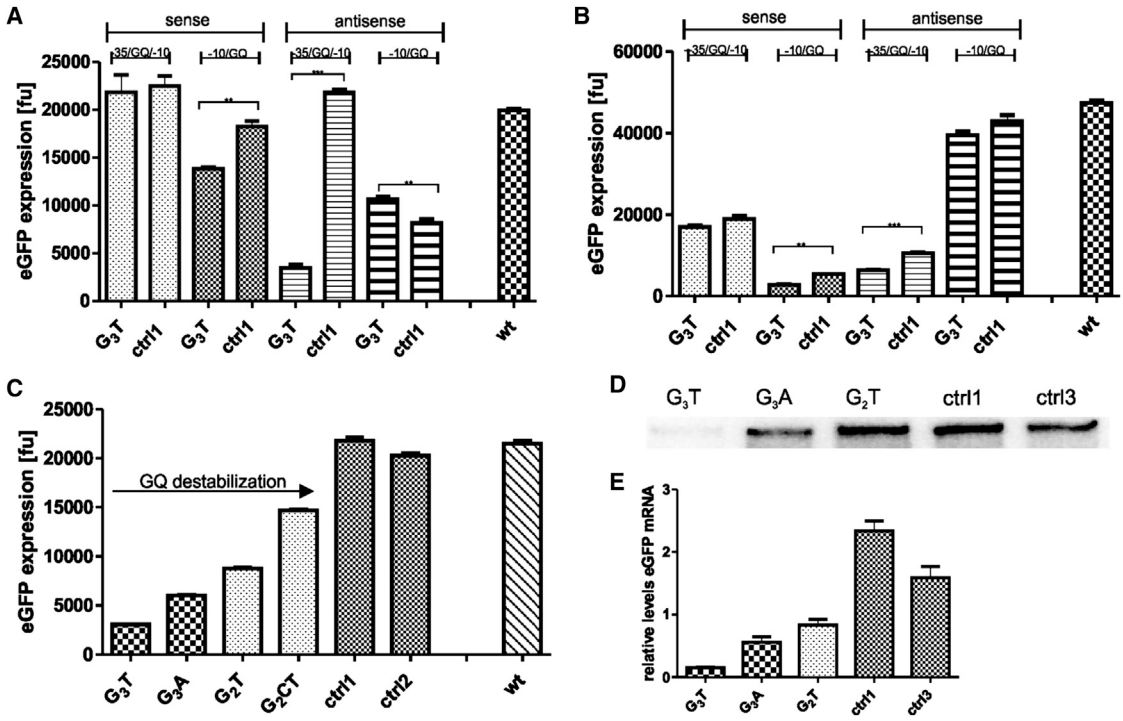


Figure 3. Influence of G-Quadruplexes in Bacterial Promoters on Gene Expression

(A and B) Quadruplex forming sequences were placed either between the −10 and −35 regions (−35/GQ/−10) or downstream of the −10 region (−10/GQ), each on the sense or antisense strand in an eGFP reporter system under the control of (A) the constitutive pQE promoter and (B) the inducible pBAD promoter. ctrl, control.

(C) G-quadruplex sequences placed in the core promoter of the pQE-eGFP system.

(D) In vitro transcription of different constructs with *E. coli* RNA polymerase.

(E) Analysis of eGFP mRNA levels by semiquantitative RT-PCR relative to the expression of the genomically encoded *ssrA* gene. All experiments were performed in triplicates. Error bars represent SD of three independent experiments.

According to the unpaired t test: **$p < 0.001$; ***$p < 0.0001$.

comparison to ctrl2. Cleavage is observed at the respective guanine sites, which might be explained by only temporary formation of the quadruplex during transcription. Interestingly, a strong band can be observed directly 5′ of the G-rich stretch for the untreated G3T construct (without addition of DMS) but not for the untreated ctrl2 construct. This finding likely results from inhibition of DNA polymerization during the primer extension reaction due to secondary structure formation. Furthermore, we tested whether the addition of the quadruplex-stabilizing ligand N-methyl mesoporphyrin (NMM) could enhance the influence of the quadruplex on gene expression. However, no change in the gene expression pattern was observed on addition of different concentrations of NMM (2 μM, 20 μM, and 100 μM; Figure S7).

Influence of Quadruplexes in the 5′-UTR

We next investigated whether similar effects as seen in the promoter region are also observed in transcribed regions. Chowdhury and colleagues reported quadruplex motifs to be enriched within 100 nt upstream of the start codon in *E. coli*. Some of these motifs were identified about 20 base pairs upstream of the start codon, e.g., in front of genes such as *yhiP* or *yabB* and their orthologs in other organisms (Rawal et al., 2006). Therefore, we decided to focus on quadruplexes located 20 nt upstream of the start codon and closely after the TSS (Figure 1E).

In our pQE-eGFP reporter system, we inserted the most stable G_3T quadruplex and the less stable G_2CT quadruplex and compared their eGFP expression with two different controls (Figure 4A). Intriguingly, the quadruplex sequences in the antisense strand enhanced gene expression in comparison to controls and wild-type vector. Expression increased by 37% and more than 100% for the G_2CT and the G_3T constructs, respectively, compared to ctrl1. When the motif was inserted in the sense strand instead, we observed decreased gene expression: 60% for G_3T and 49% for G_2CT, relative to ctrl1. Again, we investigated whether this modulation of gene expression occurs on the transcriptional or the translational level. By determination of mRNA levels, we observed an increase of 44% when comparing the eGFP mRNA level of G_3T on the antisense strand to that of ctrl1 (Figure 4C). However, the eGFP mRNA levels of G_3T inserted in the sense strand remained constant compared to the controls (Figure 4D).

Engineering of SD-Region-Adjacent Quadruplexes that Increase Gene Expression on Their Formation

Translational modulation of gene expression via quadruplexes has been shown earlier by our group. Artificially designed sequences were placed around the SD region so that secondary structure formation inhibits interaction of the 16S ribosomal RNA (rRNA) and initiation of translation (Wieland and Hartig,

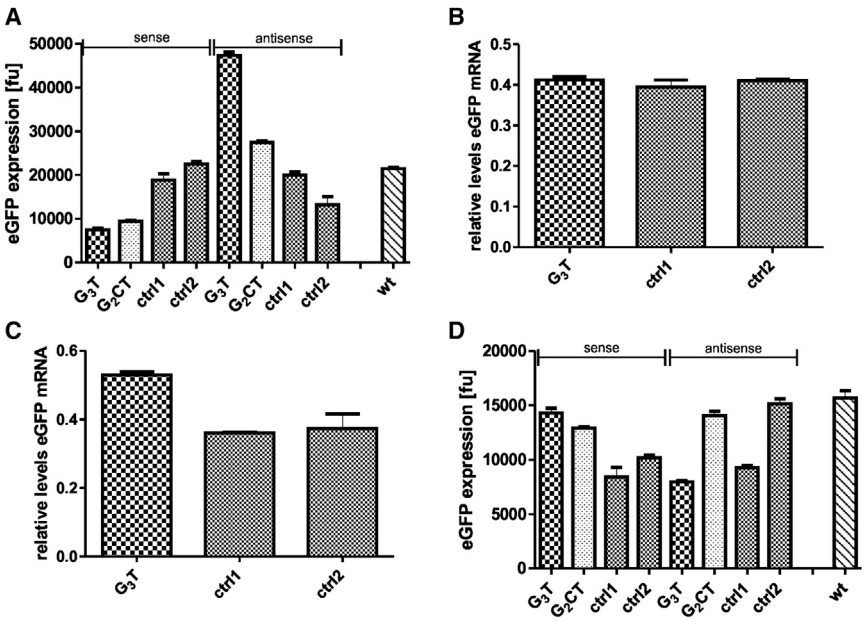


Figure 4. Influence of G-Quadruplexes on Gene Expression in the pQE-eGFP Reporter Construct

(A) G-quadruplex-forming sequences of different stabilities inserted 20 nt upstream of eGFP start codon, with G-tract either on sense or antisense strand. fu, fluorescence units.

(B and C) Respective eGFP mRNA levels have been analyzed by semiquantitative RT-PCR and are shown in (B) for the sense and (C) for the antisense strand. RNA levels were calculated relative to the expression of the genomically encoded *ssrA* gene.

(D) eGFP expression of constructs with G-quadruplex-forming sequences of different stabilities inserted 4 nt downstream of eGFP stop codon, with G-tract on either the sense or antisense strand. Error bars represent SDs of three independent experiments.

2007). We were curious as to whether the opposite effect—activation of gene expression—can be accomplished by quadruplex formation in the 5′-UTR. Natural RNA-regulatory mechanisms such as cold-sensing thermometers exist where a secondary structure formation liberates SD region and activates gene expression (Giuliodori et al., 2010). In order to implement a similar system based on quadruplex formation, we designed another set of 5′-UTRs where the SD region can be masked by the formation of a stem-loop structure in the mRNA (Figure 5A). We included a G-rich sequence that competes via quadruplex formation with the hairpin fold. Insertion of single-nucleotide mismatches destabilizes the stem and should simplify quadruplex formation. The formed quadruplex should prevent stem-loop formation, and the SD region should become accessible for the ribosome, thus facilitating translation. Predicted Mfold structures (Zuker, 2003) of our designs showing the mRNA region upstream of the start codon are depicted in Figure S6. We chose the G_3U quadruplex for our investigations, as it is a short sequence that folds into a remarkably stable RNA G-quadruplex structure (Zhang et al., 2011). In the first design, the G_3U quadruplex sequence was inserted 21 nt upstream of the eGFP start codon. By comparing the G_3U quadruplex construct to its non-quadruplex control (G_3U ctrl), gene expression slightly increased (15%). Destabilization of the stem-loop structure by insertion of five mismatches allowed easier quadruplex formation. For this construct (G_3U_2), we observed an activation of gene expression of more than 100% compared to the control (G_3U_2 ctrl). Destabilization of the G-quadruplex by introduction of longer loops (G_3CU_2) still increased gene expression, compared to the respective control (G3CU_2 ctrl), but with less efficiency (89%; Figure 5B). The similarity of eGFP-mRNA levels in G-quadruplex constructs and respective controls confirmed the regulation on the translational level (Figure S5E). Our results show that freeing the masked ribosome-binding site by the formation of a G-quadruplex in the mRNA is a potential mechanism of translational regulation.

The demonstrated system of activating quadruplexes somehow contradicts the results observed earlier after inserting the G_3T quadruplex into the 5′-UTR on the sense strand 20 nt upstream of the start codon, where a repression of gene expression was observed. Real-time PCR analysis showed that this decrease was related to translational modulation. In this structure, the SD region should be easily accessible for the ribosome (see Figure S6G for Mfold prediction). In the engineered system, we observed the opposite effect: quadruplexes inserted 21 nt upstream of the start codon activated gene expression. However, in this design, the whole 5′-UTR was modified (nucleotide composition and length between SD region and start codon) in order to mask the SD region when no quadruplex formation occurs. Hence, the two designs are not comparable, and both results reflect possible influences of quadruplex sequences located close to the SD region. Furthermore, the overall gene expression of both the controls and the quadruplex constructs decreased in this system in comparison to the wild-type sequence, indicating an influence of sequence changes. Changes in the 5′-UTR alter the mRNA translation rate (Espah Borujeni et al., 2014), as the 30S ribosomal complex also interacts with sequences upstream of the SD region (Simonetti et al., 2008; Yusupova et al., 2001).

Naturally Occurring Quadruplexes in the SD Region in E. coli

The observation of drastic quadruplex-mediated effects in 5′-UTRs raises the question of whether quadruplexes in natural genetic contexts exert similar control over gene expression. We investigated the occurrence of quadruplexes surrounding the SD region in the genomic sequence of *E. coli* MG1655. Using the ProQuad Pattern Algorithm (Yadav et al., 2008), we searched for quadruplexes with two to five tetrads and loops of 1–5 nt that overlap with the anticipated SD region sequences (~10–12 nt upstream of the start codon) and identified 46 sequences (Table S5). Gene functions were categorized utilizing the KEGG database (Kanehisa and Goto, 2000; Kanehisa et al., 2014). We found quadruplexes widely distributed with most sequences in categories of metabolic pathways, microbial metabolism, and

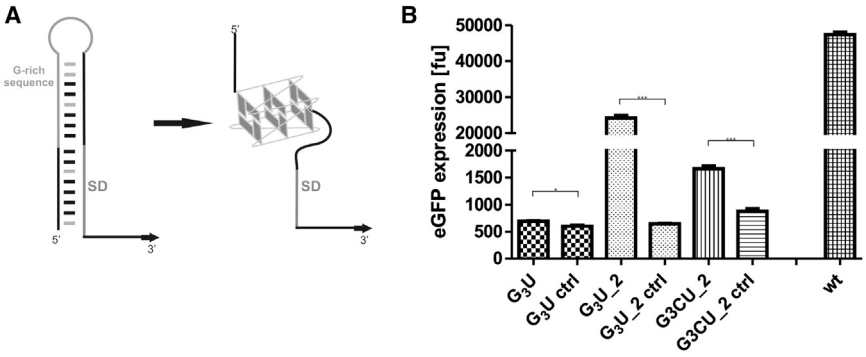


Figure 5. Artificial System Comprising SD-Region-Adjacent Quadruplexes

(A) Mechanism suggested for enhancing gene expression via G-quadruplex formation. The gray sequence is G-rich and able to form a quadruplex but can also partly pair with the black sequence immediately upstream of the SD region. Access to the SD region can be blocked by formation of a stem loop structure. G-quadruplex formation leads to the breakup of the stem-loop structure and freeing of the SD region. Gray base pairs indicated mismatches introduced for facilitating quadruplex formation.

(B) Modulation of eGFP expression. G_3U: GG GUGGGUGGGUGGG; G_3U ctrl: GGGUGGGUGU GUGUG; G_3CU_2:GGGCUGGGCUGGGCUGGG; G_3CU_2 ctrl: GGGCTGGGCTGTGCTGTG. Error bars represent SDs of three independent experiments: *$p < 0.05$; ***$p < 0.0001$. wt, wild-type.

biosynthesis of secondary metabolites. All identified quadruplexes are anticipated to form structures with two tetrads and do not show a conserved sequence pattern.

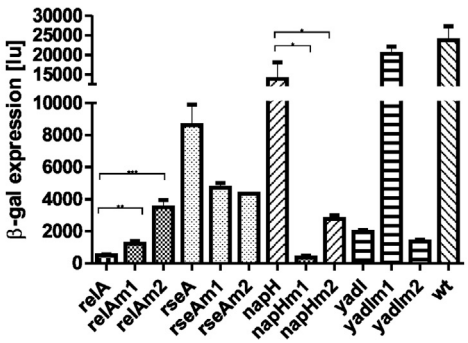
We placed the identified quadruplexes including the whole natural 5′-UTR in front of a β-galactosidase reporter gene under control of the araBAD promoter. We randomly chose five different genes with quadruplex sequences overlapping the ribosome-binding site from our list of sequences identified in the *E. coli* genome: *oxyR*, *relA*, *rseA*, *napH*, and *yadI* (Figures 6A and 6B).The quadruplexes differ in loop length and distance from the start codon (Figures 6A and 6C; Table S2). For each construct, we designed two controls which should not be able to form a quadruplex (Figure 6). Mutating Gs outside the core SD region should not alter the efficiency of 16S rRNA interactions but, instead, reduce quadruplex-based secondary structure formation. However, in this region, it is likely that even small sequence changes influence expression (Salis et al., 2009). Some of the mutants showed very high gene expression patterns that might be explained by sequence changes that facilitate ribosomal interactions. Especially yadIm1 showed an unexpectedly high gene expression, although containing only two G to U mutations. However, A/U-rich sequences upstream of the SD region have been reported to serve as mRNA-stabilizing elements (Komarova et al., 2005). Unfortunately, both controls for the yadI construct behaved differently in the gene expression studies, so that no conclusions with regard to quadruplex formation can be drawn (Figure 6B). For three other constructs (relA, oxyR, and napH), an effect on gene expression was observed that seems to be related to secondary structure formation. For the sequences upstream of the *E. coli oxyR* (Figure 6D) and *relA* (Figure 6B) genes, expression increased in both mutants: 87.1% for oxyRm1 and 92.0% for oxyRm2, as well as 59.2% for relAm1 and 85.9% for relAm2 (Figures 6B and 6D). Regarding the napH construct, we observed a decrease of expression for both controls compared to the napH construct (more than 100%). However, gene expression of napHm1 is decreased 7-fold compared to napHm2. As the two mutants differ considerably, it is difficult to associate this with secondary structure formation. Also, addition of the quadruplex-stabilizing compound NMM did not change the gene expression levels significantly (Figure S7). For the *rseA* 5′-UTR G-rich sequence, β-galactosidase expression is increased compared to both mutants (Figure 6D). In this case, the effect was not significant according to the unpaired t test.

We next investigated the *oxyR* quadruplex in more detail. We characterized the stability of the *oxyR* quadruplex RNA sequence (Figure 6C, oxyR and oxyRm1) via CD spectroscopy and thermal denaturation (Figure S5), finding a parallel four-stranded structure with a Tm of 56.2°C. The control sequence oxyRm1 (Figure 6A) containing two G-to-U mutations showed a shifted CD signal and a much lower stability, with a Tm of 38.6°C. As described earlier, the *oxyR* sequence showed significant reduction of gene expression compared to the controls oxyRm1 and oxyRm2. For further analysis, we included sequences that were mutated beyond the G-tract (Figures 6C and 6D) and should still be able to form a quadruplex structure (oxyRm3, oxyRm4, and oxyRm5). With these constructs, we wanted to support our assumption of changes in gene expression resulting from secondary structure formation and not being the effect of sequence changes in this regulatory region. Accordingly, we expected reduced gene expression for controls able to form quadruplexes with respect to the nonquadruplex controls. For oxyRm3 and oxyRm4, the respective A was changed into U 14 nt and 10 nt in front of the start codon. Gene expression increased significantly (80%) compared to the naturally *oxyR* sequence but still remained repressed in comparison to the mutants that were not able to form a G-quadruplex (oxyRm1 and oxyRm2). When U was changed into A 13 nt upstream of the start codon, gene expression decreased even more than in the natural *oxyR* sequence. However, in oxyRm6, the last G-tract was mutated, so no G-quadruplex formation should be possible. In this case, gene expression was repressed compared to the other mutants but still significantly increased (77%) compared to the natural *oxyR* sequence. Presumably, both effects (secondary structure formation and single-nucleotide changes in the SD region) contribute to the observed changes in gene expression. To exclude the influence of sequence mutations on mRNA stability or altered transcription rates for the oxyR constructs, we determined mRNA levels via semiquantitative RT-PCR. We found similar mRNA abundances for G-quadruplex constructs and mutants (Figure S5D), pointing at differential translation initiation as the likely cause of the observed differences in gene expression.

A

Name	Sequence
relA	GGAGAGGACGAUGGUUGCGG
relAm1	GGAGAGGACGAUGGUUGCUG
relAm2	UUAGAGGACGAUGGUUGCGG
rseA	GGAUACUGGAUAAGGGUAUUAGGCAUG
rseAm1	GGAUACUGGAUAAGGGUAUUAUGCAUG
rseAm2	GUAUACUGGAUAAGGGUAUUAGGCAUG
napH	GGCUGGCUGGAGGGGAACAAUG
napHm1	GGCUGUCUGGAGGGGAACAAUG
napHm2	UUCUGGCUGGAGGGGAACAAUG
yadI	GGCUAAGGAGGAAGGCGAUG
yadIm1	UUCUAAGGAGGAAGGCGAUG
yadIm2	GGCUAAUGAGGAAGGCGAUG

B



C

Name	Sequence
oxyR	GGCGAUGGAGGAUGGAUAAUG
oxyRm1	GUCGAUGGAGGAUUGAUAAUG
oxyRm2	UUCGAUGGAGGAUGGAUAAUG
oxyRm3	GGCGUUGGAGGAUGGAUAAUG
oxyRm4	GGCGAUGGUGGAUGGAUAAUG
oxyRm5	GGCGAAGGAGGAUGGAUAAUG
oxyRm6	GGCGAUGGAGGAUGAAUAAUG
wU	UAAAGAGGAGAAAUUAACUAUG

D

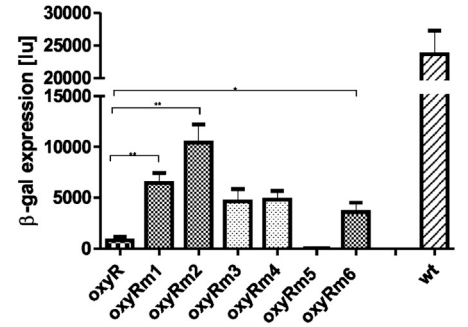


Figure 6. Naturally Occurring Quadruplexes in *E. coli* SD Regions

(A) Sequences of quadruplexes occurring in the SD region of the *E. coli relA*, *rseA*, *napH*, and *yadI* genes with their respective control mutants. wt, the SD region sequence in the wild-type pBAD vector.

(B) β-galactosidase expression of constructs listed in (A).

(C) Sequence of the G-quadruplex in front of the *oxyR* gene and the respective controls are listed.

(D) β-galactosidase expression of oxyR constructs. Error bars represent SDs of three independent experiments.

*p < 0.05; **p < 0.001; ***p < 0.0001.

Furthermore, we show that this modulation is not selective for a specific plasmid or readout system, as the insertion of the same SD region background in front of the eGFP gene in the pQE vector led to comparable results (Figure S5C).

Influence of Quadruplexes in the 3′-UTR

We also analyzed whether a quadruplex sequence inserted in the bacterial 3′-UTR influences gene expression. For eukaryotes, functions of 3′-UTR G-quadruplexes as *cis*-regulatory elements have been reported (Beaudoin and Perreault, 2013). In our study, we inserted the G-rich sequence 4 nt downstream of the eGFP stop codon in the pQE-eGFP reporter plasmid (Figures 1E and 1F). Comparing quadruplexes with different stabilities to different controls, we found no consistent modulation related to G-quadruplex stability when inserting the quadruplex sequence into the antisense strand. In addition, control sequences inserted into the antisense strand showed a different behavior, and results varied when tested in the second plasmid system (Figure 4B). In conclusion, it seems that quadruplexes inserted into the immediate 3′-UTR do not influence gene expression in a consistent manner in *E. coli*.

DISCUSSION

Several studies have described quadruplex-mediated alterations in eukaryotic gene expression. Although, in computational searches, quadruplex-forming sequences have been found enriched in regulatory regions of prokaryotes, studies that investigate their influence in vivo are rare. We systematically analyzed the influence of G-rich sequences in bacterial gene-regulatory regions. We show that the effect of G-quadruplexes in the bacterial promoter region on gene expression of the downstream gene is position dependent. Recently, the influence of strand asymmetry on quadruplex-mediated alteration of transcription was described for eukaryotes by Maiti and colleagues (Agarwal et al., 2014). In their study, a quadruplex sequence in the 5′-UTR only repressed transcription efficiency when placed into the antisense strand. In contrast, translational repression of gene expression was also possible when the quadruplex was found in the sense strand. However, the comparison of prokaryotic and eukaryotic systems in this context might prove difficult, as genetic mechanisms differ significantly. Hence, conclusions drawn from studies in eukaryotic contexts cannot be transferred to bacteria; instead, separate investigations are necessary.

Our results illustrate that G-quadruplexes can be involved in bacterial gene regulation on both transcriptional and translational levels. A significantly decreased gene expression by transcriptional modulation was observed for constructs bearing the quadruplex on the antisense strand of the core promoter. We found that transcriptional repression at this position correlates with quadruplex stability. Next, we inserted quadruplexes into the sense and antisense strands after the TSS but 20 nt upstream of the start codon. An increase of gene expression was observed for quadruplexes placed into the antisense strand, and a repression of gene expression was observed for those in the sense strand. When analyzing mRNA levels, we found that the quadruplex on the antisense strand influences transcription, possibly by interfering with polymerase binding or elongation. The antisense strand serves as template for the *E. coli* RNA polymerase, which reads this strand in 3′ to 5′ orientation. Although the TSS is located downstream of the promoter region, polymerase binding to the promoter is essential for transcription initiation. The σ^{70} factor of the *E. coli RNA* polymerase core enzyme is responsible for promoter recognition, binding the −10 and −35 regions of the double-stranded promoter and unwinding

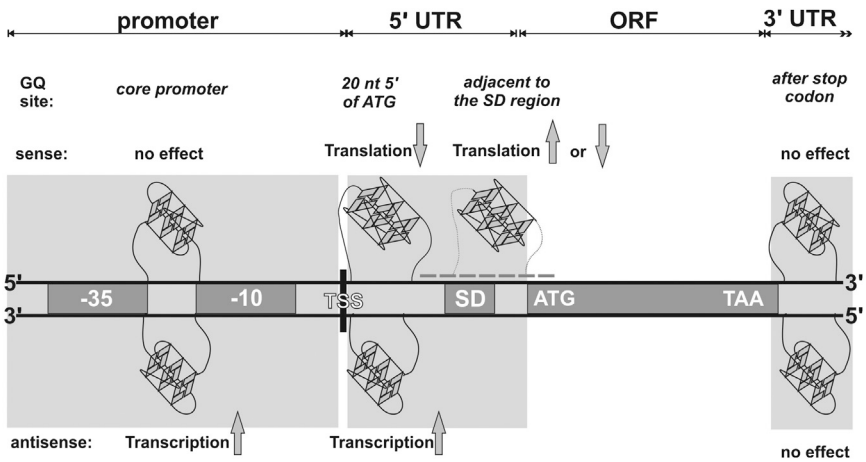


Figure 7. Overview of Quadruplex-Induced Influence on Gene Expression

Different quadruplex insertion sites and their effects are shown. Arrows pointing up, increased gene expression; arrows pointing down, decreased gene expression. Dashed line indicates sequence range that was modified for investigation of the SD adjacent region.

dsDNA at the −10 region, followed by binding to the sense strand. An initial transcript of ~10 nt causes the release of the σ^{70} factor from the core RNA polymerase, which leaves the promoter and enters the elongation phase by moving along the antisense strand (Mekler et al., 2011; Roberts and Roberts, 1996). Promoter recognition can be strongly influenced by the nucleotide composition of the adjacent 5′-UTR. The formation of a secondary structure such as a G-quadruplex could create a physical barrier that hinders polymerase binding or complicates promoter recognition by σ^{70}. Polymerase binding might also facilitate quadruplex formation, which ultimately might hamper transcriptional initiation or entering the elongation phase. A formed quadruplex structure might simplify strand separation supporting the helicase activity of RNA polymerase. In the region downstream of the −10 region, the polymerase actively separates the double helix (Mekler et al., 2011; Roberts and Roberts, 1996). This might explain why the insertion of a quadruplex in the promoter results in transcriptional repression, whereas location downstream of the promoter increases gene expression. Notably, we observed a similar behavior in two different σ^{70}-dependent promoters, the constitutive J06 promoter and the arabinose-inducible araBAD promoter.

We also tested the influence of the quadruplex-stabilizing compound NMM, but no effects on gene expression were observed in our systems. Although the use of quadruplex-stabilizing compounds has been developed into a valuable tool in quadruplex research, most studies are concerned with eukaryotic systems (Hershman et al., 2008). There are only few reports about quadruplex-targeting compounds in bacteria. NMM was shown to stabilize G-quadruplex structures in *Neisseria gonorrhoeae* (Cahoon and Seifert, 2009) and *Deinococcus radiodurans* (Beaume et al., 2013). However, effects of NMM addition on quadruplexes in *E. coli* have not been shown so far. In fact, we have never observed effects with a series of other quadruplex-targeting compounds in *E. coli* in the currently described system, as well as other systems characterized before (data not shown). Hence, in our opinion, we are not able to draw conclusions from the absence of effects on gene expression on treatment with NMM or other quadruplex-binding compounds.

In addition to promoter-based control of gene expression, masking of the ribosome-binding site is a common mechanism for translational regulation, e.g., in riboswitches, RNA thermometers, sRNA-mediated regulation and in artificially designed riboregulators. Different systems have been described using engineered devices as sensitive switches of gene expression in prokaryotic organisms (Klauser and Hartig, 2013; Wieland et al., 2009; Wieland and Hartig, 2007; Winkler et al., 2002). In this study, we successfully constructed a system where the SD region was masked via a hairpin structure. A quadruplex sequence was incorporated into the stem-loop structure so that quadruplex formation destabilizes the hairpin and liberates the SD region, resulting in activation of gene expression. We summarized the influences of G-quadruplex sequences in these different artificial systems in Figure 7.

With regard to naturally occurring sequences, we identified 46 quadruplexes occurring on the coding strand within the SD region in the *E. coli* K-12 genome. A significant quadruplex-mediated repression of expression for two of five investigated natural 5′-UTR regions was observed. As suggested earlier (Wieland and Hartig, 2007), the secondary structure of a G-quadruplex might complicate the binding of the ribosome to the SD region and thereby decrease gene expression. Given that naturally occurring quadruplexes in bacteria mediate gene expression efficiently, the possibility of quadruplexes playing functional roles in controlling gene expression cannot be excluded. In such a scenario, it might be possible that these distinct structures are specifically induced under certain conditions. It is of importance to note that intracellular K^+ concentrations increase in response to osmotic upshock and general stress responses (Epstein and Schultz, 1965; Huo et al., 2008). Intriguingly, G-quadruplexes are stabilized by K^+ ions. We found several quadruplexes in 5′-UTRs of genes related to stress responses (Table S5). OxyR, the oxidative stress regulator, is a transcriptional regulator in the oxidative and nitrosative stress response (Anjem et al., 2009; Seth et al., 2012). *RelA* encodes for an enzyme involved in the stringent response (Magnusson et al., 2005). RseA is an antagonist of the sigma E factor, which is involved in heat shock, osmotic shock, and other stress responses (Ades et al., 1999). Other genes related to stress have been identified in our search (Table S5). One could speculate that the identified quadruplex motifs function as regulatory units responding to environmental changes. However, bacterial adaptations and lifestyle changes are regulated by several complex and overlaying pathways. This makes the formation and influence of a potentially temporarily induced quadruplex structure difficult to prove. We carried out initial experiments with osmotic upshock that should result in temporarily increased intracellular K^+ levels but found no conclusive influence in reporter gene assays (data not shown). However, further experiments along these lines

utilizing even better suited reporter gene assays might be able to shed more light on the possibility of quadruplex formation as a natural mechanism for conditional gene regulation.

SIGNIFICANCE

For eukaryotes, the influence of quadruplexes on gene expression is well established. Research investigating the role of quadruplexes in prokaryotic genetic contexts is sparse. We present a systematic analysis of the influence of quadruplex sequences in bacterial gene-regulatory regions. We demonstrate that drastic effects are observed in a strand- and position-dependent manner. We designed model sequences that elicit gene activation and inactivation on formation of quadruplex structures. In addition, we study the influence of natural quadruplex sequences occurring in crucial positions. The finding of pronounced effects of four-stranded structures provokes the question of whether nature makes use of such simple measures for controlling gene expression in a conditional manner. In general, the present study significantly broadens the insights into the effects of nucleic acid secondary structure formation on gene expression.

EXPERIMENTAL PROCEDURES

Materials, Plasmids, and *E. coli* Strain

Primers and oligonucleotides for CD measurements and melting assays were synthesized by Sigma-Aldrich at the 1 μmol scale with high-performance liquid chromatography purification. Used plasmids are either based on the pQE-Tri-System (QIAGEN) or the pBAD18a (Guzman et al., 1995) vector systems. Promoters used are the J06 promoter (modified from the Anderson promoter library; http://parts.igem.org/Promoters/Catalog/Anderson) and the araBAD promoter (Guzman et al., 1995) (Figure 1; Figure S1). Standard molecular cloning procedures were performed as described in the literature (Sambrook and Russel, 2001) and in the Supplemental Information. To introduce G-quadruplex sequences and their respective controls in the regulatory regions, we used primers listed in Tables S1–S4 with assigned vectors. All experiments were conducted with the *E. coli* XL10 gold strain (Invitrogen); endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tetR F′[proAB lacIqZΔM15 Tn10(TetR Amy CmR)]. Plasmids were introduced by electroporation. Bacteria were grown aerobically in Luria-Bertani (LB) medium supplemented with 100 μg/ml carbenicillin.

Gene Expression Studies

Gene expression of constructs with G-quadruplexes in the promoter region and constructs with G-quadruplexes liberating the SD region was examined using enhanced green fluorescent protein (eGFP) readout for both pBAD and pQE vector systems. One hundred microliters of each culture were transferred into 96-well microplates, and the eGFP fluorescence was determined with a TECAN Infinite M200 plate reader (excitation wavelength = 288 nm, emission wavelength = 535 nm). Fluorescence values were corrected by dividing by the optical density of 600 (OD_{600}). Gene expression of constructs bearing the naturally occurring SD-G-quadruplexes was determined by β-galactosidase assay. Outgrown cultures were diluted 1:2,000 with LB medium and induced by addition of 1mM arabinose. When grown to OD_{600} = 0.5, *lacZ* expression was determined through the Gal-Screen β-galactosidase reporter gene assay system (Life Technologies) and luminescence was measured with a TECAN Infinite M200 plate reader. Luminescence values were corrected by dividing by the OD_{600} values. All experiments were performed in triplicates, and error bars represent SD.

Quantification of mRNA Levels

Bacterial cultures were grown as for gene expression studies. Total RNA was extracted using RNAeasy Mini Kit (QIAGEN). Isolated RNA was digested with DNase I (New England Biolabs) and further purified by phenol-chloroform extraction (Aqua Phenol, Carl Roth GmbH). The reverse transcription reaction was performed with 1 μg total RNA and random hexamer priming using the Superscript III Reverse Transcriptase (Invitrogen) in a total volume of 20 μl for 60 min at 50°C. Real-time PCR analysis was performed on a TOptical Thermocycler (Biometra). Each reaction mixture was prepared using Phusion Hot-Start Polymerase II (New England Biolabs) for amplification and SYBRGreen (Sigma-Aldrich) for detection in a total volume of 13 μl. The primers in parentheses were used for the amplification reaction of the eGFP mRNA (forward: 5′-AAGCTGACCCTGAAGTTCATCTGC-3′; reverse: 5′-TTCACCTCGGCGCG GGTCTTGTAG-3′), the β-galactosidase mRNA (forward: 5′-ATGACCATGAT TACGGATTCACTG-3′; reverse: 5′-GCGATCGGTGCGGGCCTCTTC-3′), and the *ssrA* mRNA (reference gene; forward: 5′-ACGGGGATCAAGAGAGGTC AAAC-3′; reverse: 5′-GGACGGACACGCCACTAAC-3′). RNA levels were calculated assuming a static PCR efficiency of two for each primer pair and determined relative to the expression of the genomically encoded *ssrA* gene.

CD Measurements

CD spectra were measured on a JASCO-J815 spectropolarimeter equipped with a MPTC-490S/15 multicell temperature unit using a 1 cm optical path. Oligonucleotide samples were prepared as a 5 μM solution in 10 mM Tris-HCl, pH 7.5, supplemented with 100 mM KCl in a reaction volume of 600 μl. DNA folding was facilitated by heating to 98°C for 5 min, followed by slow cooling to 20°C overnight. Scans were performed at 20°C over a wavelength range of 220–320 nm (five accumulations) with a scanning speed of 500 nm/min, 0.5 s response time, 0.5 nm data pitch, and 1 nm bandwidth. The buffer spectrum was subtracted, and all spectra were zero-corrected at 320 nm. For thermal denaturation, oligonucleotides were prepared as for CD measurements. Folded samples were heated from 20°C to 100°C with a heating rate of 0.5°C min^{-1}. The CD signal at the respective maximum was recorded every 0.5°C (G3T, G3A, and G2T: 260 nm; G2CT: 290 nm; ctrl1: 280 nm; ctrl2: 265 nm; ctrl3: 270 nm). The Tm was obtained from the normalized ellipticity decrease.

In Vitro Transcription

In vitro transcription was performed with PCR products as templates (forward: 5′-AGTGCCACCTGACGTCTAAGAAACC-3′; reverse: 5′-GATGATGATGATG ATGATGGC-3′). PCR products had a length of 187 base pairs containing the pQE-J06 promoter and the eGFP gene start. They were gel purified, and 0.5 μg was used for the following experiment. Before in vitro transcription, templates were folded by heating to 95°C and cooling down to 4°C within 50 min. In vitro transcription was performed with α-^{32}P-GTP and *E. coli* RNA polymerase according to the manufacturer's protocol (New England Biolabs). Samples were ethanol precipitated and analyzed by 10% denaturing PAGE.

In Vivo DMS Footprint

Fifty microliters of a 10% DMS solution was added to a 5 ml overnight bacterial culture in LB medium and incubated for 5 min at 37°C before being placed on ice. Two milliliters of bacterial culture was pelleted at 4°C, and plasmid DNA was extracted with a commercial Plasmid Miniprep Kit (Zymo Research). DNA was digested with NcoI (New England Biolabs) to generate a full-length product in the primer extension assay and purified via phenol/chloroform extraction (Roti-Phenol, Carl Roth GmbH). Sequencing controls were generated by treating isolated plasmid DNA formic acid for purine sequencing or hydrazine for pyrimidine sequencing, as described by Maxam and Gilbert (1977). Cleavage at the modified sites was performed by addition of 10% piperidine at 94°C for 30 min. Piperidine was removed in a vacuum concentrator. For primer extension, the primer (5′-AGGCGTATCACGAGGCCCTTTC-3′) was radioactively 5′ end-labeled with γ-^{32}P-ATP. Primer extension was performed with VENTexo- polymerase (New England Biolabs) and analyzed on a 10% denaturating PAGE gel.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2014.09.014.

ACKNOWLEDGMENTS

We are grateful to Astrid Joachimi, Sandra Gunesch, and Alexander Reim for assistance. We thank the members of the J.S.H. group for helpful discussions.

Received: July 15, 2014
Revised: August 13, 2014
Accepted: September 17, 2014
Published: October 30, 2014

REFERENCES

Ades, S.E., Connolly, L.E., Alba, B.M., and Gross, C.A. (1999). The Escherichia coli sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. Genes Dev. *13*, 2449–2461.

Agarwal, T., Roy, S., Kumar, S., Chakraborty, T.K., and Maiti, S. (2014). In the sense of transcription regulation by G-quadruplexes: asymmetric effects in sense and antisense strands. Biochemistry *53*, 3711–3718.

Anjem, A., Varghese, S., and Imlay, J.A. (2009). Manganese import is a key element of the OxyR response to hydrogen peroxide in Escherichia coli. Mol. Microbiol. *72*, 844–858.

Arora, A., and Maiti, S. (2009). Differential biophysical behavior of human telomeric RNA and DNA quadruplex. J. Phys. Chem. B *113*, 10515–10520.

Bacolla, A., Larson, J.E., Collins, J.R., Li, J., Milosavljevic, A., Stenson, P.D., Cooper, D.N., and Wells, R.D. (2008). Abundance and length of simple repeats in vertebrate genomes are determined by their structural properties. Genome Res. *18*, 1545–1553.

Beaudoin, J.D., and Perreault, J.P. (2013). Exploring mRNA 3′-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. Nucleic Acids Res. *41*, 5898–5911.

Beaume, N., Pathak, R., Yadav, V.K., Kota, S., Misra, H.S., Gautam, H.K., and Chowdhury, S. (2013). Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of D. radiodurans involves G4 DNA-mediated regulation. Nucleic Acids Res. *41*, 76–89.

Biffi, G., Di Antonio, M., Tannahill, D., and Balasubramanian, S. (2014). Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. Nat. Chem. *6*, 75–80.

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.

Bugaut, A., and Balasubramanian, S. (2008). A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes. Biochemistry *47*, 689–697.

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.

Cahoon, L.A., and Seifert, H.S. (2009). An alternative DNA structure is necessary for pilin antigenic variation in Neisseria gonorrhoeae. Science *325*, 764–767.

Cahoon, L.A., and Seifert, H.S. (2011). Focusing homologous recombination: pilin antigenic variation in the pathogenic Neisseria. Mol. Microbiol. *81*, 1136–1143.

Capra, J.A., Paeschke, K., Singh, M., and Zakian, V.A. (2010). G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae. PLoS Comput. Biol. *6*, e1000861.

Du, Z., Zhao, Y., and Li, N. (2009). Genome-wide colonization of gene regulatory elements by G4 DNA motifs. Nucleic Acids Res. *37*, 6784–6798.

Eddy, J., and Maizels, N. (2008). Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. Nucleic Acids Res. *36*, 1321–1333.

Endoh, T., and Sugimoto, N. (2013). Unusual -1 ribosomal frameshift caused by stable RNA G-quadruplex in open reading frame. Anal. Chem. *85*, 11435–11439.

Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013a). Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor α. Nucleic Acids Res. *41*, 6222–6231.

Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013b). Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability. Angew. Chem. Int. Ed. Engl. *52*, 5522–5526.

Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013c). Translational halt during elongation caused by G-quadruplex formed by mRNA. Methods *64*, 73–78.

Epstein, W., and Schultz, S.G. (1965). Cation transport in Escherichia coli: V. Regulation of cation content. J. Gen. Physiol. *49*, 221–234.

Espah Borujeni, A., Channarasappa, A.S., and Salis, H.M. (2014). Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. Nucleic Acids Res. *42*, 2646–2659.

Gellert, M., Lipsett, M.N., and Davies, D.R. (1962). Helix formation by guanylic acid. Proc. Natl. Acad. Sci. USA *48*, 2013–2018.

Giacani, L., Brandt, S.L., Puray-Chavez, M., Reid, T.B., Godornes, C., Molini, B.J., Benzler, M., Hartig, J.S., Lukehart, S.A., and Centurion-Lara, A. (2012). Comparative investigation of the genomic regions involved in antigenic variation of the TprK antigen among treponemal species, subspecies, and strains. J. Bacteriol. *194*, 4208–4225.

Giuliodori, A.M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., Gualerzi, C.O., and Pon, C.L. (2010). The cspA mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. Mol. Cell *37*, 21–33.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. *177*, 4121–4130.

Halder, K., Wieland, M., and Hartig, J.S. (2009). Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes. Nucleic Acids Res. *37*, 6811–6817.

Halder, K., Benzler, M., and Hartig, J.S. (2012a). Reporter assays for studying quadruplex nucleic acids. Methods *57*, 115–121.

Halder, R., Riou, J.F., Teulade-Fichou, M.P., Frickey, T., and Hartig, J.S. (2012b). Bisquinolinium compounds induce quadruplex-specific transcriptome changes in HeLa S3 cell lines. BMC Res. Notes *5*, 138.

Hanakahi, L.A., Sun, H., and Maizels, N. (1999). High affinity interactions of nucleolin with G-G-paired rDNA. J. Biol. Chem. *274*, 15908–15912.

Hershman, S.G., Chen, Q., Lee, J.Y., Kozak, M.L., Yue, P., Wang, L.S., and Johnson, F.B. (2008). Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae. Nucleic Acids Res. *36*, 144–156.

Huo, Y.X., Rosenthal, A.Z., and Gralla, J.D. (2008). General stress response signalling: unwrapping transcription complexes by DNA relaxation via the sigma38 C-terminal domain. Mol. Microbiol. *70*, 369–378.

Huppert, J.L. (2008). Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chem. Soc. Rev. *37*, 1375–1384.

Huppert, J.L. (2010). Structure, location and interactions of G-quadruplexes. FEBS J. *277*, 3452–3458.

Huppert, J.L., and Balasubramanian, S. (2005). Prevalence of quadruplexes in the human genome. Nucleic Acids Res. *33*, 2908–2916.

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

Huppert, J.L., Bugaut, A., Kumari, S., and Balasubramanian, S. (2008). G-quadruplexes: the beginning and end of UTRs. Nucleic Acids Res. *36*, 6260–6268.

Joachimi, A., Benz, A., and Hartig, J.S. (2009). A comparison of DNA and RNA quadruplex structures and stabilities. Bioorg. Med. Chem. *17*, 6811–6815.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. *42*, D199–D205.

Klauser, B., and Hartig, J.S. (2013). An engineered small RNA-mediated genetic switch based on a ribozyme expression platform. Nucleic Acids Res. *41*, 5542–5552.

Komarova, A.V., Tchufistova, L.S., Dreyfus, M., and Boni, I.V. (2005). AU-rich sequences within 5′ untranslated leaders enhance translation and stabilize mRNA in Escherichia coli. J. Bacteriol. *187*, 1344–1349.

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

Lee, N., Francklyn, C., and Hamilton, E.P. (1987). Arabinose-induced binding of AraC protein to aral2 activates the araBAD operon promoter. Proc. Natl. Acad. Sci. USA *84*, 8814–8818.

Magnusson, L.U., Farewell, A., and Nyström, T. (2005). ppGpp: a global regulator in Escherichia coli. Trends Microbiol. *13*, 236–242.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA *74*, 560–564.

Mekler, V., Minakhin, L., and Severinov, K. (2011). A critical role of downstream RNA polymerase-promoter interactions in the formation of initiation complex. J. Biol. Chem. *286*, 22600–22608.

Murat, P., Zhong, J., Lekieffre, L., Cowieson, N.P., Clancy, J.L., Preiss, T., Balasubramanian, S., Khanna, R., and Tellam, J. (2014). G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. Nat. Chem. Biol. *10*, 358–364.

Nakken, S., Rognes, T., and Hovig, E. (2009). The disruptive positions in human G-quadruplex motifs are less polymorphic and more conserved than their neutral counterparts. Nucleic Acids Res. *37*, 5749–5756.

Niland, P., Hühne, R., and Müller-Hill, B. (1996). How AraC interacts specifically with its target DNAs. J. Mol. Biol. *264*, 667–674.

Perrone, R., Nadai, M., Frasson, I., Poe, J.A., Butovskaya, E., Smithgall, T.E., Palumbo, M., Palù, G., and Richter, S.N. (2013). A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter. J. Med. Chem. *56*, 6521–6530.

Rawal, P., Kummarasetti, V.B., Ravindran, J., Kumar, N., Halder, K., Sharma, R., Mukerji, M., Das, S.K., and Chowdhury, S. (2006). Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation. Genome Res. *16*, 644–655.

Rehm, C., Holder, I.T., Groß, A., Wojciechowski, F., Urban, M., Sinn, M., Drescher, M., and Hartig, J.S. (2014). A bacterial DNA quadruplex with exceptional K^+ selectivity and unique structural polymorphism. Chem. Sci. *5*, 2809–2818.

Roberts, C.W., and Roberts, J.W. (1996). Base-specific recognition of the nontemplate strand of promoter DNA by E. coli RNA polymerase. Cell *86*, 495–501.

Saccà, B., Lacroix, L., and Mergny, J.L. (2005). The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. Nucleic Acids Res. *33*, 1182–1192.

Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. *27*, 946–950.

Sambrook, J.F., and Russel, D.W. (2001). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Sawaya, S., Bagshaw, A., Buschiazzo, E., Kumar, P., Chowdhury, S., Black, M.A., and Gemmell, N. (2013). Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS ONE *8*, e54710.

Seth, D., Hausladen, A., Wang, Y.J., and Stamler, J.S. (2012). Endogenous protein S-nitrosylation in E. coli: regulation by OxyR. Science *336*, 470–473.

Siddiqui-Jain, A., Grand, C.L., Bearss, D.J., and Hurley, L.H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc. Natl. Acad. Sci. USA *99*, 11593–11598.

Simonetti, A., Marzi, S., Myasnikov, A.G., Fabbretti, A., Yusupov, M., Gualerzi, C.O., and Klaholz, B.P. (2008). Structure of the 30S translation initiation complex. Nature *455*, 416–420.

Verma, A., Halder, K., Halder, R., Yadav, V.K., Rawal, P., Thakur, R.K., Mohd, F., Sharma, A., and Chowdhury, S. (2008). Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species. J. Med. Chem. *51*, 5641–5649.

Verma, A., Yadav, V.K., Basundra, R., Kumar, A., and Chowdhury, S. (2009). Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. Nucleic Acids Res. *37*, 4194–4204.

Walia, R., and Chaconas, G. (2013). Suggested role for G4 DNA in recombinational switching at the antigenic variation locus of the Lyme disease spirochete. PLoS ONE *8*, e57792.

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

Wieland, M., and Hartig, J.S. (2007). RNA quadruplex-based modulation of gene expression. Chem. Biol. *14*, 757–763.

Wieland, M., Benz, A., Klauser, B., and Hartig, J.S. (2009). Artificial ribozyme switches containing natural riboswitch aptamer domains. Angew. Chem. Int. Ed. Engl. *48*, 2715–2718.

Winkler, W., Nahvi, A., and Breaker, R.R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature *419*, 952–956.

Yadav, V.K., Abraham, J.K., Mani, P., Kulshrestha, R., and Chowdhury, S. (2008). QuadBase: genome-wide database of G4 DNA—occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. Nucleic Acids Res. *36*, D381–D385.

Yu, C.H., Teulade-Fichou, M.P., and Olsthoorn, R.C. (2014). Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. Nucleic Acids Res. *42*, 1887–1892.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. Cell *106*, 233–241.

Zhang, D.H., Fujimoto, T., Saxena, S., Yu, H.Q., Miyoshi, D., and Sugimoto, N. (2010). Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. Biochemistry *49*, 4554–4563.

Zhang, A.Y., Bugaut, A., and Balasubramanian, S. (2011). A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology. Biochemistry *50*, 7251–7258.

Zhang, C., Liu, H.H., Zheng, K.W., Hao, Y.H., and Tan, Z. (2013). DNA G-quadruplex formation in response to remote downstream transcription activity: long-range sensing and signal transducing in DNA double helix. Nucleic Acids Res. *41*, 7144–7152.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.