PLOS ONE

New Insights into Transcription Fidelity: Thermal Stability of Non-Canonical Structures in Template DNA Regulates Transcriptional Arrest, Pause, and Slippage

Hisae Tateishi-Karimata[1], Noburu Isono[1], Naoki Sugimoto[1,2]*

1 Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, Kobe, Japan, 2 Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, Kobe, Japan

Abstract

The thermal stability and topology of non-canonical structures of G-quadruplexes and hairpins in template DNA were investigated, and the effect of non-canonical structures on transcription fidelity was evaluated quantitatively. We designed ten template DNAs: A linear sequence that does not have significant higher-order structure, three sequences that form hairpin structures, and six sequences that form G-quadruplex structures with different stabilities. Templates with non-canonical structures induced the production of an arrested, a slipped, and a full-length transcript, whereas the linear sequence produced only a full-length transcript. The efficiency of production for run-off transcripts (full-length and slipped transcripts) from templates that formed the non-canonical structures was lower than that from the linear. G-quadruplex structures were more effective inhibitors of full-length product formation than were hairpin structure even when the stability of the G-quadruplex in an aqueous solution was the same as that of the hairpin. We considered that intra-polymerase conditions may differentially affect the stability of non-canonical structures. The values of transcription efficiencies of run-off or arrest transcripts were correlated with stabilities of non-canonical structures in the intra-polymerase condition mimicked by 20 wt% polyethylene glycol (PEG). Transcriptional arrest was induced when the stability of the G-quadruplex structure ($-\Delta G^{o}_{37}$) in the presence of 20 wt% PEG was more than 8.2 kcal mol^{-1}. Thus, values of stability in the presence of 20 wt% PEG are an important indicator of transcription perturbation. Our results further our understanding of the impact of template structure on the transcription process and may guide logical design of transcription-regulating drugs.
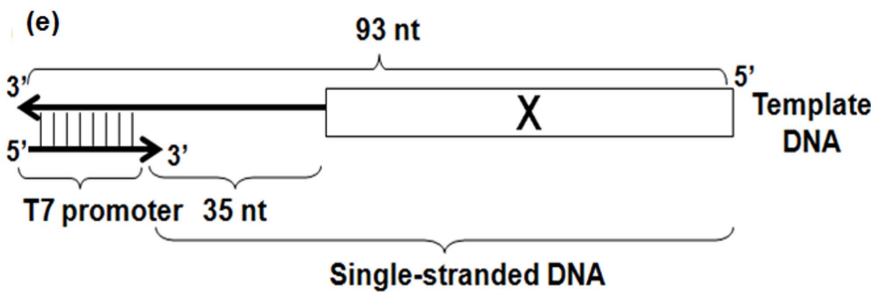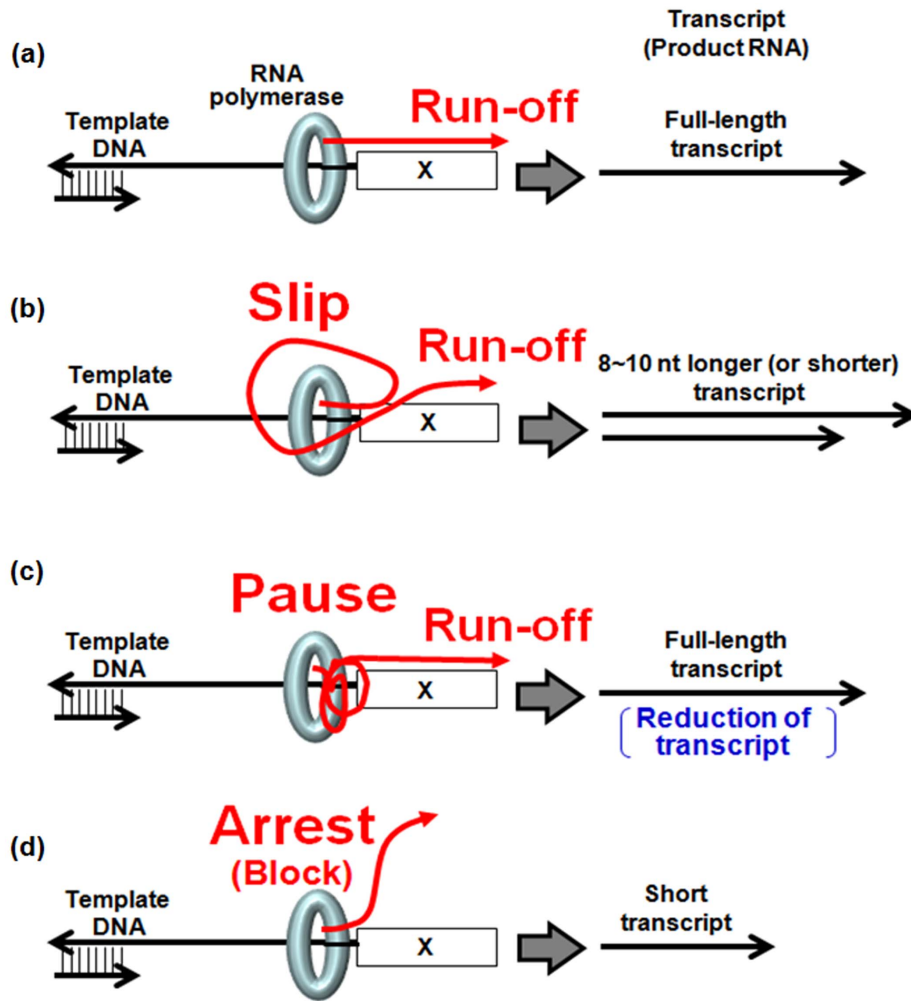
Introduction

Transcription is the first step in gene expression; it is highly regulated during both initiation and elongation.[1,2] Although fidelity of transcription elongation is critical for maintaining the accurate flow of genetic information, transcription elongation in cells and *in vitro* (Figure 1a) can be interrupted by certain sequences or structures.[3,4] For example, the RNA polymerase may slip back or forward on a template DNA during RNA synthesis at "slippage" sites such as the homopolymeric tract in human *amyloid precursor protein* gene.[2,5] In the *amyloid precursor protein* gene, the slippage results in a transcript 8- to 10-nt shorter or longer than the expected transcript, changing the coding capacity of mRNA (Figure 1b).[4,5,6] Moreover, the rate of transcription elongation is also dependent on sequence. The average rate of transcription elongation by RNA polymerase II on relatively unstructured DNA templates in the absence of any additional factors is ~5 nucleotides s^{-1} under optimal conditions.[7,8] At some template positions called "pause" sites such as the A-T-rich transactivation sequence in the HIV genome, the polymerase may halt for as long

as 1 min before continuing transcription. The pause results in decreased mRNA production and therefore less protein production (Figure 1c). At other locations called "arrest" sites, which can be a damaged DNA residue or a an A or T stretch,[5,6] a transcribing polymerase halts, and the 3′ end of the transcript is displaced from the active site of the polymerase and the short transcript is released (Figure 1d).[9] These arrested transcripts may misfold or result in translation of mutant protein. Like non-coding RNAs such as miRNAs and antisense RNAs,[10,11,12,13,14] these short transcripts may also have roles in regulation of gene expression. Transcription is important not only for cellular systems but is also in certain nano-materials such as logic devices.[15] If transcript mutations could be controlled, they might prove useful in medical, pharmaceutical, and materials sciences fields.

We have shown that non-canonical structures such as triplexes, G-quadruplexes, and cruciform structures are stabilized drastically under cell-like conditions induced by cosolutes,[16] although a duplex is destabilized under the same conditions.[16,17] Recently, it has been reported that non-canonical G-quadruplex structures are present inside cells.[18] These results imply that non-canonical

(a) Template DNA, RNA polymerase, Run-off, Transcript (Product RNA), Full-length transcript

(b) Slip, Run-off, 8~10 nt longer (or shorter) transcript

(c) Pause, Run-off, Full-length transcript, Reduction of transcript

(d) Arrest (Block), Short transcript

(e) 93 nt, Template DNA, T7 promoter 35 nt, Single-stranded DNA, X

(f)

Sequence name	Sequence of X region (5'→3')
Linear:	GCCGTTTCGTAGTATTTGGGTTGTAACTATCGAGG
H1:	GCCGTTTCGTAGTATTTCTATCC*GGACTTCGGTCC*
H2:	GCCGTTTCGTAGT*CTATCGGACTTCGGTCCGATAG*
H3:	GCCGT*CCAACTATCGGACTTCGGTCCGATAGTTGG*
Q1:	GCCGTTTCGTAGTATTTC*GGTTAGGTTAGGTTAGG*
Q2:	GCCGTTTCGTAGTA*GGGTTAGGGTTAGGGTTAGGG*
Q3:	GCCGTTTCGT*GGGGTTAGGGGTTAGGGGTTAGGGG*
Q4:	GCCGTTTCGTAGTATTTCTA*GGTTGGTGTGGTTGG*
Q5:	GCCGTTTCGTAGTATT*GGGTTGGGTGTGGGTTGGG*
Q6:	GCCGTTTCGTAG*GGGGTTGGGGTGTGGGGTTGGGG*

Figure 1. Effects on RNA polymerase elongation by structures in template DNA (a–d) and illustration of the template DNA (e, f). (a) An unstructured template, (b) a template with a slippage site, (c) a template with a pause site, and (d) a template with an arrest site. (e) The region denoted by the box marked with an X contains the sequence designed to form a random coil or non-canonical structure. (f) Sequence names and sequences of X regions. Sequences expected to form non-canonical structures are highlighted by italic and bold.
doi:10.1371/journal.pone.0090580.g001

structures may play important roles in gene regulation. Additionally, we showed very recently that the formation of a G-quadruplex in an open reading frame (ORF) of mRNA suppresses its translation into protein.[19] It has been reported that quadruplex-forming sequences are enriched upstream and within gene promoters; structures formed in these regions may lead to transcription pausing via the inhibition of transcription initiation.[20] Moreover, hairpin structures formed by the nascent RNA may result in transcription pausing even if the hairpin does not form in template DNA.[21,22] Recent reports have indicated that transcription is arrested prior to stable structures in the DNA template such as a Z-form duplex,[23] triplex,[24,25] or G-quadruplex,[26,27] but there have been no quantitative and systematic reports the effects of various non-canonical structures on arrest, slippage or pause of transcription from the viewpoint of stability of non-canonical structures at the energy level.

The phenomena of arrest, pause, and slippage are ubiquitous aspects of transcriptional regulation. Arrest, pause, and slippage sites may result from unstable hybridization between template DNA and nascent RNA or upon formation of a structure in the DNA that interrupts and perturbs elongation. Although formation for a DNA/RNA hybrid can result in transcriptional interruption,[28] the mechanism of transcription interruption due to non-canonical structures may differ, because non-canonical structures like G-quadruplex have very different structural features and stabilities than DNA/RNA hybrids. Here, we evaluated the quantitative effects of non-canonical structures formed in the template DNA on transcription catalysed by T7 RNA polymerase.

Results

Structure and stability of non-canonical structures

Hairpins (template region sequences of cruciform structure) and G-quadruplexes are known to form in template DNA. We designed and synthesized ten different template DNAs (Figures 1e, 1f and Table S1 in File S1) in order to evaluate the effect of formation of hairpins and G-quadruplexes with different thermal stabilities on transcription elongation. The control sequence (Linear) should not form significant structure. Sequences H1 to H3 and Q1 to Q6 were designed to form hairpins or G-quadruplexes, respectively, at a site 35 bases downstream from the T7 promoter region as shown in Figure 1e. The G-quadruplex-forming sequences are based on the human telomeric sequence and the thrombin DNA aptamer sequence and have different numbers of G-quartets (Figure 1f).[16] To confirm the formation of the non-canonical structures in the template DNAs, oligonucleotides containing only the non-canonical structure region (linear, h1 to h3, and q1 to q6) were synthesized (Table S2 in File S1). We measured circular dichroism (CD) spectra of these DNA oligonucleotides at 37°C (Figure S1a in File S1) and also analyzed each by native gel electrophoresis (Figure S1b in File S1). These analyses indicated hairpin formation by h1, h2, and h3, and G-quadruplex formation by oligonucleotides q1-q6 (Table 1 and Figure S1c in File S1). Moreover, formation of G-quadruplexes in the template DNAs for Q1-Q6 was confirmed by florescent analysis using protoporphyrin IX, which binds specifically to G-quadruplexes and produces fluorescence (Figure S2 in File S1).[19,29]

The thermodynamic parameters for the formation of hairpins and G-quadruplexes were obtained from their thermal melting curves by monitoring UV absorbance at 295 or 260 nm (Figure 2a).[16] The mid-points of the thermal melting transitions (T_m values) and thermodynamic parameters for the structure formation are given in Table 1 and Table S3 in File S1. The values of T_m and $-\Delta G°37$ (the stability at 37°C) for h1, h2, and h3 increased with increasing the stem length (4, 9, and 13 base pairs, respectively). The T_m values for G-quadruplexes based on the human telomeric q1, q2, and q3 were 37.3, 62.9, 89.8°C, respectively. Those for G-quadruplexes q4, q5 and q6 based on the thrombin aptamer sequence were 46.8, 80.5, and >95°C, respectively. In both cases, the T_m values increased with the number of G-quartet stacks. The values of $-\Delta G°37$ for the G-quadruplex-forming oligonucleotides also increased with increasing the number of G-quartets. The G-quadruplexes formed by q4, q5, and q6 were much more stable than those formed by q1, q2, and q3 due to differences in stacking interactions of loop regions (Figure S1c in File S1). These thermal analyses indicated that all the non-canonical structures should form in a template DNA at 37°C.

Effect of non-canonical structures on transcription

T7 RNA polymerase transcription of the linear template under multi-turnover conditions was almost saturated at 90 min (data not shown). Inhibitory effects of non-canonical structures in the template DNA on transcription were estimated from the amount of transcript (product RNA) formed at this time point. Figure 2c shows the results of gel electrophoretic analysis of transcription carried out for 90 min at 37°C under multi-turnover conditions. RNA size was determined by analysis of samples of each reaction in parallel with size markers and a 35-nt RNA on a denaturing polyacrylamide gel (Figure 2c). When the template with no significant structure was used, the transcription proceeded to the end of the DNA template, resulting in the formation of a full-length transcript of 70 nt (Figure 2c, lane 3). To confirm that the Linear template produced mainly full-length transcript, we carried out additional experiments to quantitate formation of longer and shorter products from this template (Figure S3 in File S1 and Methods section). Interestingly, transcription of all template DNAs able to form non-canonical structures resulted in products in addition to the full-length transcript. The products of transcription of templates H1, H2, Q1, Q2, Q4, and Q5 yielded transcripts approximately 10-nt longer than the full-length transcript (Figure 2c, lanes 4, 5, 7, 8, 10, and 11). In contrast, transcripts from the reaction with H3, which formed the most stable hairpin, contained a minor product band that migrated at approximately 60 nt, about 10-nt shorter than the full-length transcript (Figure 2c, lanes 6). Previous studies indicate that elongating RNA polymerase can slip to produce transcripts longer and shorter by 8 to 10 nts. Thus, formation of non-canonical structures appears to induce slippage. The template DNAs Q3, Q5, and Q6 able to form the very stable G-quadruplexes ($-\Delta G°37$ values more than 14.3 kcal mol^{-1} in Table 1) induced production of a transcript of about 35 nt (Figure 2c, lanes 9, 11 and 12). This length suggests that T7 transcription was arrested at the G-quadruplex structure. Under the conditions used here, the hairpins induced slippage, whereas G-quadruplexes induced slippage and arrest, with extent of arrest

depending on the G-quadruplex stability. We also carried out the transcription using template DNA in the presence of complementary strand (Figures S4a and S4b in File S1). Template DNAs of Q3, Q5, and Q6 induced production of the arrested transcript even in the presence of complementary DNA (Figure S4c in File S1). Thus, our model systems revealed the correlation between the stability of structures adopted by the template and the transcription efficiency.

Amount of non-full length transcript depends on the stability of non-canonical structures

To understand how structural stability affected the production of slipped and arrested transcripts, we further analysed transcription from template Q5, which resulted in both slipped and arrested products (Figure 2c, lane 11). We evaluated transcription in buffer with KCl concentrations ranging from 0 to 70 mM (Figure 3), because the stability of G-quadruplex strongly depends on K+ concentration.[30,31,32] UV measurements showed the $-\Delta G°37$ values of q5 increased with increasing KCl concentration (Table S4 in File S1). In the absence of KCl, the $-\Delta G°37$ value for q5 was -0.1 kcal mol^{-1}, and slipped, but no arrested, product was produced from template Q5 (Figure 3a, lane 3). In the presence of 10 to 70 mM KCl, the arrested transcript was observed (Figure 3a, lanes 4 to 7, red arrows). The amount of slipped transcript decreased and the amount of arrested transcript increased with increasing KCl concentration. The transcription experiments were also performed in various concentrations of LiCl as G-quadruplexes are significantly destabilized in the presence of Li$^+$.[33] Only slipped transcript was produced in the presence of LiCl (Figure 3b, blue arrows). Thus, the slippage was induced even when the G-quadruplex was unstable, and increasing amounts of arrest were observed as the non-canonical structure was stabilized. Only full-length transcript was produced from the liner template in the presence of KCl and LiCl (Figure S5 in File S1).

To confirm that the production of slipped and arrested transcripts was caused by formation of a G-quadruplex, transcription experiments were also performed in the presence of G-quadruplex-stabilizing ligands, 5,10,15,20-tetrakis (N-methyl-4-pridyl) porphyrine (TMPyP4), berberine, and N-methyl mesoporphyrine IX (NMN).[34,35,36]As we expected, more of the arrested transcript was observed for template Q5 in the presence of each ligand than without ligand in 3 mM KCl (Figure S6 in File S1). The stability of the q5 structure in the presence of each of the ligand was significantly enhanced (Table S5 in File S1). We also investigated transcription from different template DNAs (Q3, Q6, and the platelet-derived growth factor B (PDFB) sequence,[37] which contain a G-quadruplex-forming sequence). In all case, the presence of G-quadruplex forming sequences induced transcription arrest (Table S5 in File S1).

Quantitative correlation between the stability of non-canonical structures and transcription efficiency

To quantify the effect of the non-canonical structure on the transcription, we estimated the correlation between the stability of non-canonical structures and transcription efficiency (TE). The full-length and slipped transcripts were defined as "run-off transcripts" (TE$_{run-off}$) because the RNA polymerase travels to the end of template DNA when these products are produced (Figure 1b). The TE of transcript (%) was calculated as the proportion of the gel band intensity for each transcript from structured template DNAs to the band intensity of full-length transcript from linear template DNA (Table S6 in File S1).

Table 1. The stabilities of non-canonical structures designed to form in template DNA [a]

Sequence[b]	Structure	T_m[c] (°C)	$-\Delta G°_{37}$ (kcal mol^{-1})	Arrest[d]
H1 (h1)	hairpin	55.5	2.7±0.3	No
H2 (h2)	hairpin	74.5	7.2±0.4	No
H3 (h3)	hairpin	82.4	12.7±0.2	No
Q1 (q1)	antiparallel G-quadruplex	37.3	0.1±0.1	No
Q2 (q2)	(3+1) mixed G-quadruplex	62.9	3.7±0.2	No
Q3 (q3)	(3+1) mixed G-quadruplex	89.8	n.d.[e]	Yes
Q4 (q4)	antiparallel G-quadruplex	46.8	1.4±0.2	No
Q5 (q5)	parallel G-quadruplex	80.5	14.3±0.3	Yes
Q6 (q6)	antiparallel G-quadruplex	>95	n.d.[e]	Yes

[a]All experiments were carried out in a buffer containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0), 8 mM MgCl$_2$, and 2 mM spermidine.
[b]The sequences of template DNAs are shown in Figure 1b and Table S1 in File S1. The sequences designated by lower case letters contain only the non-canonical structure region (see Table S2 in File S1).
[c]The melting temperature was determined at a strand concentration of 2 μM.
[d]Arrest was defined as more than 4% production of arrested product RNA.
[e]The $-\Delta G°_{37}$ value could not be determined because of very high stability.
doi:10.1371/journal.pone.0090580.t001

The TE$_{run-off}$ values for H1, H2, and H3 in 30 mM KCl solution were 87.3, 84.8, and 76.7%, respectively; thus, the transcription efficiency decreased with increasing hairpin stability (Table 1 and Table S6 in File S1). The TE$_{run-off}$ values for Q1, Q2, and Q3 in 30 mM KCl solution were 98.0, 66.9, 16.2%, and those for Q4, Q5, and Q6 were 99.3, 47.5, and 4.2%, respectively (Table S6 in File S1). Moreover, the TE$_{run-off}$ values for Q5 in 10 and 50 mM KCl solution were 68.5 and 45.0%, respectively (Table S4 in File S1). Thus, for these G-quadruplex-forming templates the transcription efficiency decreased with increasing G-quadruplex stability. In general, the TE$_{run-off}$ should not depend on DNA sequence because the rate-determining step of transcription in multi-turnover conditions is the initiation step (that is, RNA polymerase binding to promoter region). The TE$_{run-off}$ for templates with non-canonical structures clearly decreased depending on DNA sequence, which indicated that the pause induced by the non-canonical structures was the rate-determining step. For the hairpin-forming templates, H1, H2 and H3, the values of TE$_{run-off}$ decreased with increasing the values of $-\Delta G°37$ (Figure S7a in File S1). The values of TE$_{run-off}$ for Q1, Q2, Q4, and Q5 also decreased with increasing $-\Delta G°37$ (Figure S7a in File S1), but the decrement depended on sequence and topology of G-quadruplex.

We also estimated the correlation between the stability of G-quadruplex and transcription efficiency for arrested transcript (TF$_{arrest}$). Interestingly, we found a good correlation between the stability of G-quadruplex structures in the template DNA and the TF$_{arrest}$ (Figure S7b in File S1). The values of TF$_{arrest}$ increased with increasing stability of the G-quadruplex. We hypothesize that the polymerase cannot unwind a stable G-quadruplex, and, therefore, the polymerase stalls. The arrested transcripts are produced when the polymerase stalls long enough to dissociate from the template.
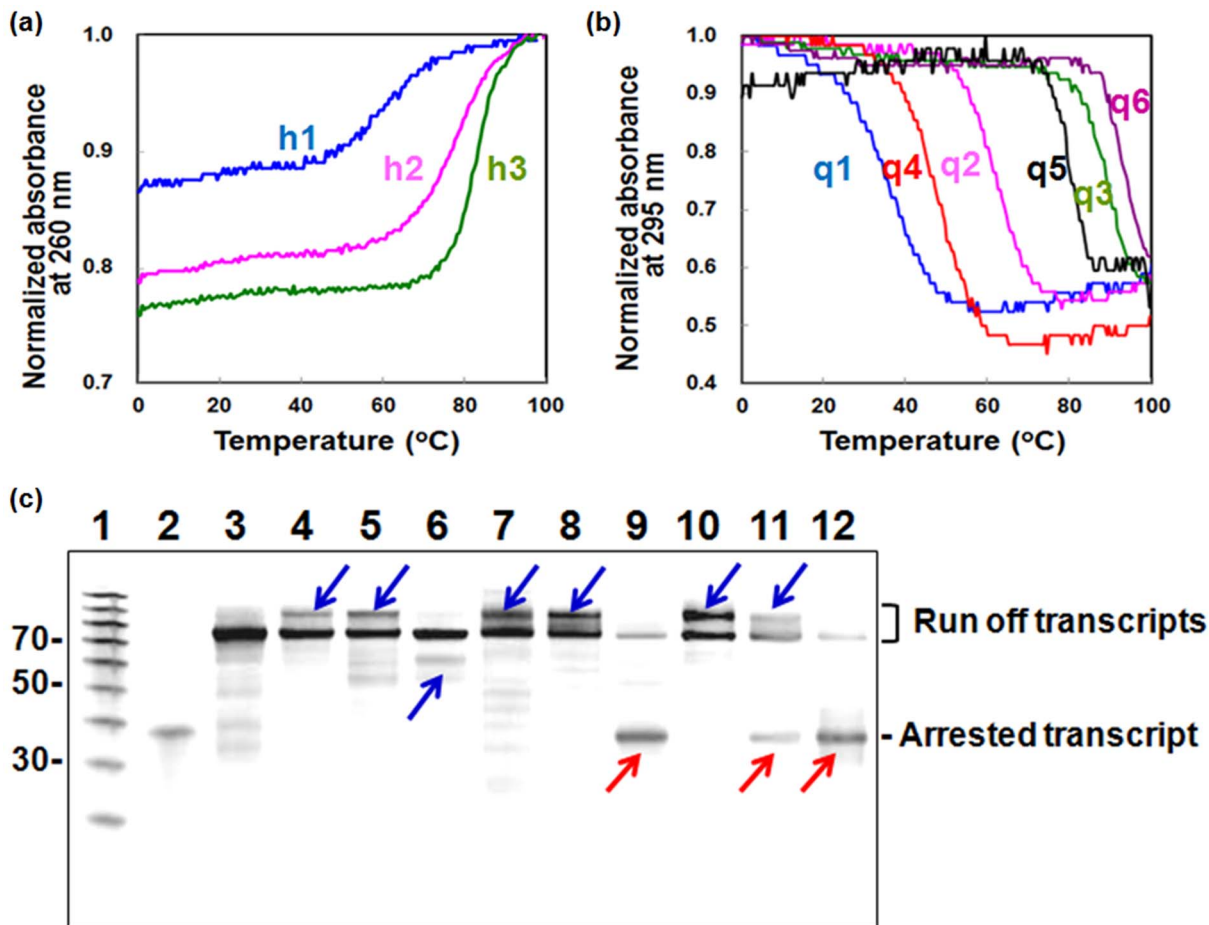
Figure 2. UV melting curves for oligonucleotides containing sequences in DNA templates (a, b) and effects of non-canonical structures in the template DNA on transcription (c). Normalized UV melting curves for 2 μM (a) h1 (blue), h2 (pink), and h3 (green), (b) q1 (blue), q2 (pink), q3 (green), q4 (red), q5 (black), and q6 (purple) in a buffer containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, and 2 mM spermidine. The T_m values for h1, h2, h3, q1, q2, q3, q4, q5, and q6 were 55.5, 74.5, 82.4, 37.3, 62.9, 89.8, 46.8, 80.5 and >95°C. (c) Denaturing gel electrophoresis of products of transcription reactions carried out for 90 min at 37°C. Reaction mixtures contained 0.3 μM T7 polymerase and 1.5 μM DNA template in a buffer containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, and 2 mM spermidine. Lane 1, size marker; lane 2, 35-nt RNA; lanes 3 to 12, transcription products for Linear, H1, H2, H3, Q1, Q2, Q3, Q4, Q5, and Q6 templates, respectively. Blue and red arrows indicate the slipped and arrested transcripts, respectively.
doi:10.1371/journal.pone.0090580.g002

Correlation between the stability of non-canonical structures in the presence of polyethylene glycol and transcription efficiency

Intracellular environments are highly crowded with various biomolecules; therefore, *in vitro* studies under molecular crowding conditions provide important information on how biomolecules behave in cells.[16,38,39] We previously showed that the structures of DNA immobilized on a gold surface are totally different from those in dilute bulk solution due to molecular crowding of immobilized DNAs.[40] Moreover, the presence of discrete nanoscale spaces such as those inside micelles and reverse micelles induces molecular crowding and alters the stability and structures of biomolecules.[41,42] During transcription, the template DNA duplex is taken into the active site of RNA polymerase, therefore, the environment surrounding DNA is very different from a dilute solution. We hypothesized that the polymerase essentially induces crowding conditions on the template DNA and therefore alters the stability of the structures adopted by the template. We measured the values of $-\Delta G°37$ for the non-canonical structures formed by the templates in the

molecular crowding condition induced by 20 wt% polyethylene glycol of molecular weight 200 (PEG 200).[16,38,39] As expected based on previous studies,[16,17] hairpins formed by h1, h2, and h3 were destabilized by addition of 20 wt% PEG 200. In contrast, the G-quadruplexes formed by q1, q2, q4, and q5 were stabilized, although the stabilities of q3 and q6 were too high to allow accurate determination of thermodynamic parameters. In principle, formation of DNA structures is accompanied by the formation of a hydrogen-bonding network of water surrounding the DNA surface that is highly sensitive to the water activity of the solution.[16,17,43] A decline in the water activity in the presence of PEG 200 disfavors formation of structures such as hairpins that require uptake of water molecules, and favors formation of structures such as G-quadruplex that are accompanied by the release of water. [16,43]

We estimated the correlation between the stability of non-canonical structures in the presence of 20 wt% PEG 200 and $TE_{run-off}$ (Figure 4a). The values of $TE_{run-off}$ decreased with increasing stability of non-canonical structures in the presence of 20 wt% PEG 200. The decrement of $TE_{run-off}$ was not dependent

on topology of non-canonical structures, although values of $TE_{run-off}$ seemed to strongly depend on the topology of non-canonical structures in the absence of PEG. Furthermore, the amount of arrested product increased linearly with the increasing stability of non-canonical structures in the presence of 20 wt% PEG 200 (Figure 4b). Thus, the $-\Delta G°37$ values for the non-canonical structures in the presence of PEG are reflective of the stabilities of the non-canonical structures during transcription.

Discussion

During transcription, pause or arrest may result from unstable hybridization between template DNA and nascent RNA or upon formation of a structure in the DNA that interrupts and perturbs elongation. The decrease of the transcription efficiency in the templates studied here results from a halt in transcription elongation at structures formed in the DNA templates. Significantly lower transcription efficiencies (as measured based on the amount of full-length transcript produced) were observed with templates able to form G-quadruplex structures relative to templates able to form hairpins (Figure 2c and Table S6 in File S1). In addition, transcription arrest was induced only by stable G-quadruplexes (Figures 2c, 3 and S6 in File S1, and Tables S4, S5, and S6 in File S1). Thus, G-quadruplex structures are more effective inhibitors of elongation than are hairpin structures, as suggested by the correlation between transcript production and the stability of structures formed by oligonucleotides of the same sequence (Figure S7 in File S1).

Our experiments in the presence of PEG 200 indicate that the conditions within the polymerase active site might be more similar to those induced by molecular crowding agents than to those in an aqueous buffer. The dielectric constant is also likely low in the region of the polymerase active site: There were several reports that the dielectric constant is very low (less than 10) inside proteins such as RNase T1.[44] We and the other groups have shown that molecular crowding conditions stabilize G-quadruplexes and destabilize hairpins.[16,39,45] Moreover, G-quadruplexes are stabilized when the dielectric constant is low, but duplex stabilities are unchanged or slightly destabilized. Therefore, the stabilities of the G-quadruplexes during the transcription could be higher than stabilities estimated based on our thermal melting analyses in the absence of PEG 200 and that of hairpins could be lower. Additionally, Hoogsteen base pairs are stabilized under the molecular crowding conditions[16,39]. Thus, the PEG 200 added

as the molecular crowding inducer may mimic not only molecular crowding but also intra-polymerase conditions. We observed correlations between the values of $TE_{run-off}$ or TE_{arrest} and stabilities of non-canonical structures in the presence of 20 wt% PEG 200 (Figure 4).

A model based on the transcription elongation complex (TEC) that consists of RNA polymerase, template DNA, and nascent RNA at an abasic site suggests that transcription can be interrupted by formation of a hybrid between template DNA and nascent RNA.[8,46,47] We hypothesize that the mechanism of transcription interruption due to non-canonical structures differs from this model. When the polymerase encounters a non-canonical structure that is not readily denatured, the TEC complex will be destabilized. Due to the destabilization of TEC, the polymerase may slip backward or forward on the template DNA. Alternatively, the TEC may dissociate. Transcriptional arrest was induced when the stability of the structure ($-\Delta G°37$) in the presence of 20 wt % PEG 200 was more than 8.2 kcal mol^{-1} (the X-intercept value in Figure 4b), indicating that the polymerase is unable to melt structures more stable than 8.2 kcal mol^{-1}. Patel and co-workers determined experimentally that the observed free energy changes for opening (melting) the template DNA duplex ($-\Delta G_{open}$) in the complex of T7 polymerase and template DNA in the promoter during transcription initiation were in the range of 7.0 to 8.0 kcal mol^{-1}, different from the predicted value ($\sim$6 kcal mol^{-1}) of observed free energy changes by nearest neighbour analysis the under physiological ionic conditions.[48] Interestingly, the $-\Delta G_{open}$ value from the Patel's report is similar to the value we estimate based on G-quadruplex stability. Thus, the stable non-canonical structures might stall the polymerase long enough to result in its dissociation from the template. Our estimated values of non-canonical stability in the presence of 20 wt% PEG 200 are an indicator of stability necessary to result in transcription perturbation by the non-canonical structures.

Pause and a slippage is known to occur in homopolymeric A/T or C/T tracts in the template DNA[2] and can lead to production of mutant protein. It was reported that polymerase backtracking appears to pass through a paused, transcriptionally inactive intermediate state.[28,49] The molecular basis for backtracking is not clear, but it is correlated either with mispairing[50,51,52] or very weak pairing of the 3′ end of the transcript and the template.[49,53] Current models require that the RNA-DNA hybrid at the point of arrest be very weak to initiate upstream translocation.[50] During transcription elongation, structures
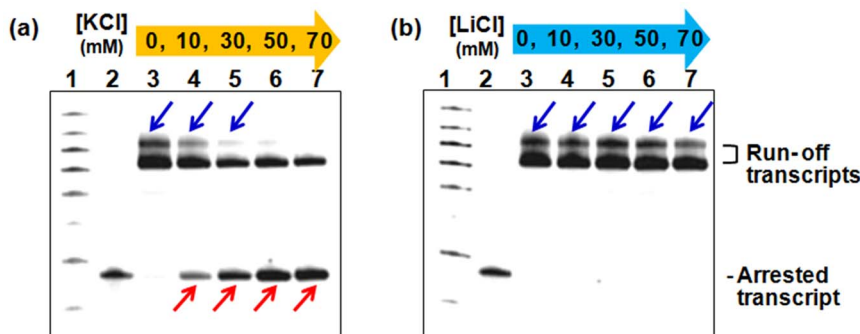
Figure 3. Effect of the G-quadruplex stability on the production of slipped and arrested transcripts. Denaturing gel electrophoresis of products of transcription reactions carried out for 90 min at 37°C using Q5 template. Reaction mixtures contained 0.3 µM T7 polymerase and 1.5 µM DNA template in a buffer containing 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, and 2 mM spermidine and various concentrations of (a) KCl or (b) LiCl. Lane 1 shows 10-nt size marker, lane 2 shows 35-nt RNA, and lanes 3 to 7 show transcription products in the presence of 0, 10, 30, 50, and 70 mM (a) KCl or (b) LiCl. Blue and red arrows indicate the slipped and arrested transcripts, respectively.
doi:10.1371/journal.pone.0090580.g003

formed by the template DNA had not been thought to induce slippage, although G-quadruplexes have been shown to induce transcription arrest and pause. [27,46] It has also been shown that G-quadruplex formation in a promoter region of template DNA may inhibit transcription initiation. It was also suggested that a purine tract can slippage, but existing models do not account for this effect since weak hybridization at the 3′end of purine-containing transcripts are not necessary to provoke arrest.[54] We hypothesize that these purine tracts form G-quadruplex structures. Our data imply that the effects resulting from the presence of non-canonical structures depend on the energy necessary to dissociate the non-canonical structures. If the polymerase encounters a non-canonical structure at least 8–10 nt upstream of a slippery sequence, the polymerase may slip and return to the transcriptionally competent position (resulting in a long mutated transcript) or may dissociate (resulting in an arrested protein product). Hence, our findings suggest another mechanism for mutation induced by non-canonical structures.

A purine tract found in the HIV-1 sequence[55] and trinucleotide repeat (TNR) sequences[56] impact polymerase elongation. It was also shown that repeat sequence of $(CG)_{14}$ induces large deletions in transcripts in mammalian cells.[57] The effects of these sequences are likely caused by non-canonical structures such as Z-form, G-quadruplexes, and hairpins (cruci-forms). The molecular crowding conditions present inside living cells varies with the cell cycle, and non-canonical structure formation will thus be regulated by conditions such as concentration of cosolutes and cations.[38] Our observations suggest how slipped and arrested transcript formation is controlled by conditions inside cells from the viewpoint of stability of non-canonical structures at the energy level.

The regulation of transcription elongation impacts the expression of many eukaryotic and viral genes. Our results indicate that transcriptional arrest by G-quadruplex structures within the template DNA depends on the stability of the structures formed. Moreover, the transcription arrest could be induced by the addition of ligands that bind to and stabilize G-quadruplexes. These ligands have potential as drugs for cancer treatment since they inhibit gene expression via transcription regulation. Our results further our understanding of the impact of template structure on the transcription process and may guide design of transcription-regulating drugs.

Conclusions

We evaluated the quantitative effect of the non-canonical structures (G-quadruplexes and hairpins) in template DNA on transcription. Previously, it had been thought that G-quadruplexes block elongation. Our results indicated quantitatively that transcriptional arrest, pause, and slippage can be induced by non-canonical structures and hairpins and that the effect is dependent on the structure formed and its thermal stability. We speculate that G-quadruplex structures are more effective inhibitors of full-length product formation than are hairpin structures because the intra-polymerase conditions preferentially stabilize the G-quadruplex. Thus, the perturbation of transcription fidelity due to non-canonical structures can be regulated by the stability of the structures formed in the polymerase.

Materials and Methods

Materials

All oligodeoxynucleotides used in this study were purified by high-performance liquid chromatography (Japan Bio Service). Single-strand concentrations of DNA oligonucleotides were calculated from the absorbance measured at 260 nm and 80°C using single-strand extinction coefficients calculated from the mononucleotide and dinucleotide data according to the nearest-neighbor approximation model.1 The absorbance was measured using a Shimadzu 1700 spectrophotometer connected to a thermoprogrammer. The concentration of T7 polymerase was determined by its absorbance at 280 nm by using $\varepsilon_{280} = 1.4 \times 10^5$ M^{-1} cm^{-1}. [58]
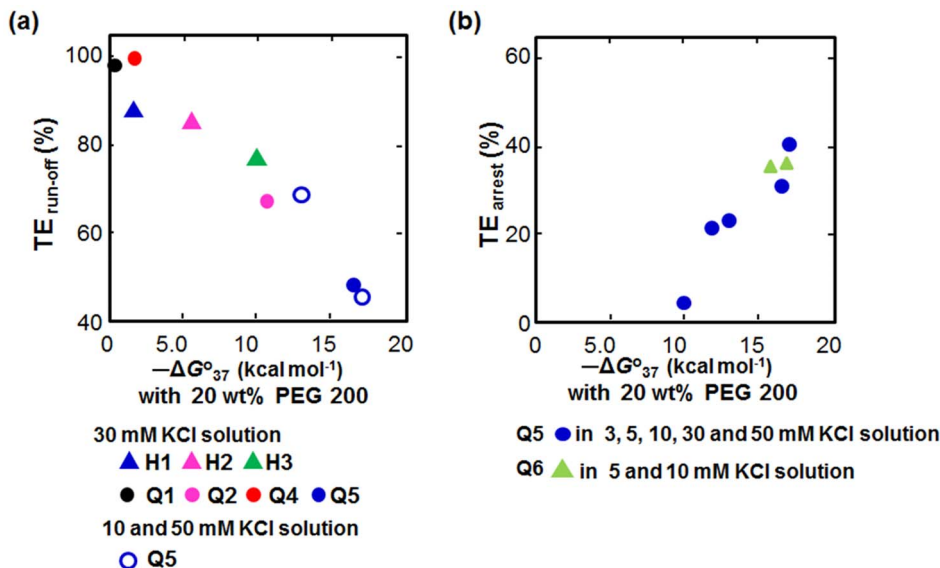
Figure 4. Correlation between $-\Delta G^o{}_{37}$ values obtained in 20 wt% PEG200 and transcription efficiency (TE). (a) TE of run-off transcripts and (b) TE of arrested transcript. Reaction mixtures contained 0.3 µM T7 polymerase and 1.5 µM DNA template in a buffer containing 40 mM Tris-HCl (pH 8.0), 8 mM MgCl$_2$, and 2 mM spermidine, and reactions were incubated for 90 min at 37°C.
doi:10.1371/journal.pone.0090580.g004

Circular dichroism (CD) measurements

CD measurements were made on a JASCO J-820 spectropolarimeter at 2 μM total DNA strand concentration in buffers containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0 at 37°C), 8 mM MgCl$_2$, and 2 mM spermidine. The spectra at 37°C were obtained by taking at least three scans from 200 to 350 nm in a cuvette with a pathlength of 0.1 cm. The temperature of the cell holder was regulated by a JASCO PTC-348 temperature controller, and the cuvette-holding chamber was flushed with a constant stream of dry N$_2$ gas to avoid condensation of water on the cuvette exterior. Before the measurement, the samples were heated to 95°C, cooled at a rate of -1°C min^{-1}, and incubated at 37°C for 30 min.

Gel electrophoresis

Native gel electrophoresis was carried out on 10% nondenaturing polyacrylamide gels in buffer containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0 at 37°C), 8 mM MgCl$_2$, and 2 mM spermidine at 37°C. Loading buffer (1 μL of 40% glycerol and 1% blue dextran) was mixed with 2 μL of 2 μM DNA sample. Gels were stained with SYBR® Gold (PerkinElmer Life Sciences) and imaged using a fluorescent imager (FUJIFILM, FLA-5100). Before the measurement, the samples were heated to 95°C, cooled at a rate of -1°C min^{-1}, and incubated at 37°C for 30 min. Unstructured DNAs of 12, 15, 25, and 30 nt were electrophoresed in parallel. When hairpin or G-quadruplex is formed, migration of the DNA should be fast relative to unstructured DNA of the same length. The migration of all template DNAs with hairpins were faster than that of the 12-nt unstructured DNA and migration of templates designed to form G-quadruplexes were faster than that of unstructured DNA of 15 nt indicating the all DNA sequences formed hairpin or G-quadruplexes.

Fluorescence measurement of protoporphyrin IX (PPIX)

Template DNA (1 μM) were mixed with PPIX (1 μM) in buffer containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0 at 37°C), 8 mM MgCl$_2$, and 2 mM spermidine for 30min. Fluorescence intensity of PPIX was measured at 37°C using a spectrofluorometer (JASCO, F6500) with 400 nm excitation and 630 nm emission. Before the measurement, the samples were heated to 95°C, cooled at a rate of -1°C min^{-1}, and incubated at 37°C for 30 min.

Thermodynamic analyses

Ultraviolet (UV) absorbance was measured on a Shimadzu 1700 spectrophotometer equipped with a temperature controller. Melting curves at 260 or 295 nm were measured in buffers containing 30 mM KCl, 40 mM Tris-HCl (pH 8.0 at 37°C), 8 mM MgCl$_2$, and 2 mM spermidine. Samples were heated at a rate of 0.5 or 0.1°C min^{-1}; melting curves of certain oligonucleotides differed at 0.1 and 0.5°C min^{-1} (data not shown). Thermodynamic parameters (Tables 1 and S3 in File S1) were calculated using the average values obtained from curve fitting at the different DNA concentrations (2, 5, 10, and 20 μM). Before the measurements, the DNA samples were heated to 95°C, cooled to 0°C at a rate of -0.5°C min^{-1}, and incubated at 0°C for 30 min.

Transcription assays

Unless otherwise noted, transcription reactions were carried out at 37°C in a total volume of 20 μl. T7 RNA polymerase was present at 0.3 μM, with DNA template at 1.5 μM. After incubating at 37°C for 10 min, NTPs were added to a final concentration of 1 mM each to initiate the reaction. The final reaction buffer contained 30 mM KCl, 40 mM Tris-HCl (pH 8.0 at 37°C), 8 mM MgCl$_2$, 2 mM spermidine, and 5 mM DTT. Reactions were quenched after incubation at the time indicated by addition of DNase I. After incubation for 10 min, a 20-fold excess volume of transcription stop solution (80 wt% formamide, 10 mM Na$_2$EDTA, and 0.1% blue dextran) was added. The samples were then heated to 90°C for 5 min, cooled rapidly, and then loaded onto a 10% polyacrylamide, 7 M urea gel run at 60°C. After electrophoresis, the gels were stained by SYBR® Gold (PerkinElmer Life Sciences), and levels were quantified with a fluorescent imager (FUJIFILM, FLA-5100). The transcription efficiency of product RNA (%) was calculated as the proportion of the gel band intensity for product RNA of each template DNAs to full-length product of linear DNA.

To confirm that transcription from Linear was not affect by conditions, we examined the transcript RNA produced by template DNA with (CT)$_7$ (sequence name: CT). CT was designed to have a typical slippage sequence of (CT)$_7$ at a site 35 bases downstream from the T7 promoter region. From the CT template, the intensity of the full-length transcript relative to the full-length transcript from Linear was very low. Moreover, transcripts migrating slower and faster relative to full-length transcript were produced from CT template. The Linear template produced mainly full-length transcript and the production amount full-length transcript from Linear was very high suggesting no perturbation of transcript fidelity. In addition, the intensities in the region of transcripts migrating slower (area A in Figure S3 in File S1) and faster (area B in Figure S3 in File S1) than full-length relative to the intensity of full-length transcript were very low from the Linear template (less than 30% relative to CT). These measurements indicate that little slippage occurs on the Linear template.

Supporting Information

File S1
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: HT-K NS. Performed the experiments: HT-K NI. Analyzed the data: HT-K NI. Contributed reagents/materials/analysis tools: HT-K NI. Wrote the paper: HT-K NS. Supervised all experiments and discussion: NS.

References

1. Jeon C, Agarwal K (1996) Fidelity of RNA polymerase II transcription controlled by elongation factor TFIIS. Proc Natl Acad Sci U S A 93: 13677–13682.

2. Wagner LA, Weiss RB, Driscoll R, Dunn DS, Gesteland RF (1990) Transcriptional slippage occurs during elongation at runs of adenine or thymine in Escherichia coli. Nucleic Acids Res 18: 3529–3535.

3. Howe KJ, Kane CM, Ares M, Jr. (2003) Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA 9: 993–1006.

4. Strathern J, Malagon F, Irvin J, Gotte D, Shafer B, et al. (2013) The fidelity of transcription: RPB1 (RPO21) mutations that increase transcriptional slippage in S. cerevisiae. J Biol Chem 288: 2689–2699.

5. Macdonald LE, Zhou Y, McAllister WT (1993) Termination and slippage by bacteriophage T7 RNA polymerase. J Mol Biol 232: 1030–1047.

6. Kassavetis GA, Chamberlin MJ (1981) Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro. J Biol Chem 256: 2777–2886.

7. Anand VS, Patel SS (2006) Transient state kinetics of transcription elongation by T7 RNA polymerase. J Biol Chem 281: 35677–35685.

8. Rudd MD, Luse DS (1996) Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. J Biol Chem 271: 21549–21558.

9. McLuckie KI, Waller ZA, Sanders DA, Alves D, Rodriguez R, et al. (2011) G-quadruplex-binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. J Am Chem Soc 133: 2658–2663.

10. Bassi GS, Weeks KM (2003) Kinetic and thermodynamic framework for assembly of the six-component bI3 group I intron ribonucleoprotein catalyst. Biochemistry 42: 9980–9988.

11. Nakano S, Karimata HT, Kitagawa Y, Sugimoto N (2009) Facilitation of RNA enzyme activity in the molecular crowding media of cosolutes. J Am Chem Soc 131: 16881–16888.

12. Sket P, Crnugelj M, Kozminski W, Plavec J (2004) 15NH4+ ion movement inside d(G4T4G4)2 G-quadruplex is accelerated in the presence of smaller Na+ ions. Org Biomol Chem 2: 1970–1973.

13. Tse MT (2012) Antisense therapeutics: Nuclear RNA more susceptible to knockdown. Nat Rev Drug Discov 11: 674.

14. Torres AG, Fabani MM, Vigorito E, Gait MJ (2011) MicroRNA fate upon targeting with anti-miRNA oligonucleotides as revealed by an improved Northern-blot-based method for miRNA detection. RNA 17: 933–943.

15. Bonnet J, Yin P, Ortiz ME, Subsoontorn P, Endy D (2013) Amplifying genetic logic gates. Science 340: 599–603.

16. Miyoshi D, Karimata H, Sugimoto N (2006) Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions. J Am Chem Soc 128: 7957–7963.

17. Nakano S, Karimata H, Ohmichi T, Kawakami J, Sugimoto N (2004) The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability. J Am Chem Soc 126: 14330–14331.

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

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

20. Bochman ML, Paeschke K, Zakian VA (2012) DNA secondary structures: stability and function of G-quadruplex structures. Nat Rev Genet 13: 770–780.

21. Toulme F, Mosrin-Huaman C, Artsimovitch I, Rahmouni AR (2005) Transcriptional pausing in vivo: a nascent RNA hairpin restricts lateral movements of RNA polymerase in both forward and reverse directions. J Mol Biol 351: 39–51.

22. Toulokhonov I, Artsimovitch I, Landick R (2001) Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. Science 292: 730–733.

23. Ditlevson JV, Tornaletti S, Belotserkovskii BP, Teijeiro V, Wang G, et al. (2008) Inhibitory effect of a short Z-DNA forming sequence on transcription elongation by T7 RNA polymerase. Nucleic Acids Res 36: 3163–3170.

24. Ebbinghaus SW, Gee JE, Rodu B, Mayfield CA, Sanders G, et al. (1993) Triplex formation inhibits HER-2/neu transcription in vitro. J Clin Invest 92: 2433–2439.

25. Belotserkovskii BP, De Silva E, Tornaletti S, Wang G, Vasquez KM, et al. (2007) A triplex-forming sequence from the human c-MYC promoter interferes with DNA transcription. J Biol Chem 282: 32433–32441.

26. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH (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 U S A 99: 11593–11598.

27. Broxson C, Beckett J, Tornaletti S (2011) Transcription arrest by a G quadruplex forming-trinucleotide repeat sequence from the human c-myb gene. Biochemistry 50: 4162–4172.

28. Palangat M, Landick R (2001) Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. J Mol Biol 311: 265–282.

29. Li T, Wang E, Dong S (2010) Parallel G-quadruplex-specific fluorescent probe for monitoring DNA structural changes and label-free detection of potassium ion. Anal Chem 82: 7576–7580.

30. Smargiasso N, Rosu F, Hsia W, Colson P, Baker ES, et al. (2008) G-quadruplex DNA assemblies: loop length, cation identity, and multimer formation. J Am Chem Soc 130: 10208–10216.

31. Podbevsek P, Sket P, Plavec J (2008) Stacking and not solely topology of T3 loops controls rigidity and ammonium ion movement within d(G4T3G4)2 G-quadruplex. J Am Chem Soc 130: 14287–14293.

32. Miyoshi D, Matsumura S, Li W, Sugimoto N (2003) Structural polymorphism of telomeric DNA regulated by pH and divalent cation. Nucleosides Nucleotides Nucleic Acids 22: 203–221.

33. Miyoshi D, Inoue M, Sugimoto N (2006) DNA logic gates based on structural polymorphism of telomere DNA molecules responding to chemical input signals. Angew Chem Int Ed 45: 7716–7729.

34. Yaku H, Murashima T, Tateishi-Karimata H, Nakano SI, Miyoshi D, et al. (2013) Study on effects of molecular crowding on G-quadruplex-ligand binding and ligand-mediated telomerase inhibition. Methods.

35. Rodriguez R, Pantos GD, Goncalves DP, Sanders JK, Balasubramanian S (2007) Ligand-driven G-quadruplex conformational switching by using an unusual mode of interaction. Angew Chem Int Ed Engl 46: 5405–5407.

36. Ma Y, Ou TM, Tan JH, Hou JQ, Huang SL, et al. (2009) Synthesis and evaluation of 9-O-substituted berberine derivatives containing aza-aromatic terminal group as highly selective telomeric G-quadruplex stabilizing ligands. Bioorg Med Chem Lett 19: 3414–3417.

37. Chen Y, Agrawal P, Brown RV, Hatzakis E, Hurley L, et al. (2012) The major G-quadruplex formed in the human platelet-derived growth factor receptor beta promoter adopts a novel broken-strand structure in K+ solution. J Am Chem Soc 134: 13220–13223.

38. Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct 22: 27–65.

39. Spink CH, Chaires JB (1999) Effects of hydration, ion release, and excluded volume on the melting of triplex and duplex DNA. Biochemistry 38: 496–508.

40. Kobayashi K, Tateishi-Karimata H, Tsutsui K, Wada Y, Sugimoto N (2011) DNA morphologic changes induced by spermine on a gold surface under DNA crowding conditions. ChemLett 40: 855–857

41. Van Horn WD, Ogilvie ME, Flynn PF (2009) Reverse micelle encapsulation as a model for intracellular crowding. J Am Chem Soc 131: 8030–8039.

42. Tsuji A, Yoshikawa K (2010) ON-OFF switching of transcriptional activity of large DNA through a conformational transition in cooperation with phospho-lipid membrane. J Am Chem Soc 132: 12464–12471.

43. Tateishi-Karimata H, Nakano S, Sugimoto N (2013) Quantitative analyses of nucleic Acid stability under the molecular crowding condition induced by cosolutes. Curr Protoc Nucleic Acid Chem Chapter 7: Unit7 19.

44. Fukunaga Y, Tamaoki H, Sakiyama F, Narita K (1982) The role of the single tryptophan residue in the structure and function of ribonuclease T1. J Biochem 92: 143–153.

45. Miller MC, Buscaglia R, Chaires JB, Lane AN, Trent JO (2010) Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3′ sequence of d(AG3(TTAG3)3). J Am Chem Soc 132: 17105–17107.

46. Tornaletti S, Maeda LS, Hanawalt PC (2006) Transcription arrest at an abasic site in the transcribed strand of template DNA. Chem Res Toxicol 19: 1215–1220.

47. Kuzmine I, Gottlieb PA, Martin CT (2001) Structure in nascent RNA leads to termination of slippage transcription by T7 RNA polymerase. Nucleic Acids Res 29: 2601–2606.

48. Bandwar RP, Patel SS (2002) The energetics of consensus promoter opening by T7 RNA polymerase. J Mol Biol 324: 63–72.

49. Neuman KC, Abbondanzieri EA, Landick R, Gelles J, Block SM (2003) Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. Cell 115: 437–447.

50. Nudler E (1999) Transcription elongation: structural basis and mechanisms. J Mol Biol 288: 1–12.

51. Thomas MJ, Platas AA, Hawley DK (1998) Transcriptional fidelity and proofreading by RNA polymerase II. Cell 93: 627–637.

52. Shaevitz JW, Abbondanzieri EA, Landick R, Block SM (2003) Backtracking by single RNA polymerase molecules observed at near-base-pair resolution. Nature 426: 684–687.

53. Liu X, Martin CT (2009) Transcription elongation complex stability: the topological lock. J Biol Chem 284: 36262–36270.

54. Hawryluk PJ, Ujvari A, Luse DS (2004) Characterization of a novel RNA polymerase II arrest site which lacks a weak 3′ RNA-DNA hybrid. Nucleic Acids Res 32: 1904–1916.

55. Brinson RG, Turner KB, Yi-Brunozzi HY, Le Grice SF, Fabris D, et al. (2009) Probing anomalous structural features in polypurine tract-containing RNA-DNA hybrids with neomycin B. Biochemistry 48: 6988–6997.

56. Caskey CT, Pizzuti A, Fu YH, Fenwick RG, Jr., Nelson DL (1992) Triplet repeat mutations in human disease. Science 256: 784–789.

57. Wang G, Christensen LA, Vasquez KM (2006) Z-DNA-forming sequences generate large-scale deletions in mammalian cells. Proc Natl Acad Sci U S A 103: 2677–2682.

58. King GC., Martin, C T., Pham, T T., and Coleman, J E. (1986) Transcription by T7 RNA polymerase is not zinc-dependent and is abolished on amidomethylation of cysteine-347. Biochemistry 25: 36–40.