Supporting Information

Human DNA repair genes possess potential G-quadruplex sequences in their promoters and 5`-untranslated regions

Aaron M. Fleming,* Judy Zhu, Yun Ding, Josh A. Visser, Julia Zhu, and Cynthia J. Burrows*

Department of Chemistry, University of Utah, Salt Lake City, UT 84112-0850, USA

*To whom correspondence should be addressed.

E-mails: burrows@chem.utah.edu or afleming@chem.utah.edu

Additional materials and methods

Plasmid Preparation. The plasmids were constructed from the psiCHECK2 plasmid (Promega) that has coding sequences for the Renilla luciferase (Rluc) and firefly luciferase (luc) genes. The luc gene is regulated by the HSV-TK promoter that was not modified and used as the internal standard to conduct the Dual-Glo luciferase assay (Promega). The Rluc gene was originally regulated by the SV40 early enhancer/promoter, which we modified by replacing the TATA-box with the potential G-quadruplex sequences (PQS) of interest. Additionally, the PQS was flanked by recognition sequences for the Nt.BspQ1 nicking endonuclease. Insertion of the PQS and nicking endonuclease recognition sequences was achieved using restriction free cloning. The cloning experiment was conducted in a 25-μL reaction with 1x Phusion reaction buffer (NEB), 200 μM dNTPs, 1 μM each of the forward and reverse primers, 1,000 ng of psiCHECK2 plasmid, 3% DMSO, and 0.5 units of Phusion DNA polymerase. The PCR was initiated with a 98 °C denaturation step for 30 sec followed by 30 cycles of PCR consisting of a 98 °C denaturation step for 20 sec, a 55 °C annealing step for 30 sec, a 72 °C extension step for 5 min, and a post step consisting of 98 °C for 8 sec followed by 72 °C for 5 min. After the 30 cycles, a final extension step of 72 °C for 5 min was conducted. Following PCR, the samples were digested with Dpn1 by adding 2.3 μL of Dpn1 reaction buffer and 5 units of Dpn1 to the PCR reaction vessel. The Dpn1 reaction was conducted for 2 hr at 37 °C followed by heat quenching at 80 °C for 20 min.

The PCR and Dpn1 digested reaction was then transformed using NEB 5-alpha competent *E. coli* following the manufacturer's protocol. After transformation, 50-100 μL from the transformation was spread on a pre-warmed agar plate containing ampicillin (100 μg/mL). The plates were incubated overnight at 37 °C. Next, individual colonies were picked and each grown overnight at 37 °C in 5 mL of lysogeny broth media containing ampicillin (100 μg/mL). The plasmid DNA from each grown sample was then purified using a miniprep kit (Qiagen) following the manufacturer's protocols. Last, ~200 ng of plasmid and 1 μM sequencing forward primer in 15 μL of 10 mM Tris (pH 8.0) was submitted for Sanger sequencing at the DNA sequencing core at the University of Utah.

Insertion of site-specific modifications into the plasmids was achieved following literature protocols.[1,2] Specifically, 5 μg of plasmid with Nt.BspQ1 recognition sequences flanking the site in which the DNA base modification will be inserted was placed in 50 μL of 1x Nt.BspQ1 reaction buffer with 5 units of Nt.BspQ1. The reaction was placed at 50 °C for 60 min followed by heat quenching at 80 °C for 20 min. Next, 1 nmole (~1000x) of 5`-phosphorylated oligomer with the site-specific modification was added to the quenched reaction. The modified oligomers were made by solid-phase synthesis in the DNA/Peptide core facility at the University of Utah. They were cleaved, deprotected, and HPLC purified following protocols previously outlined.[3] The plasmid oligomer mixture was thermal cycled by heating at 80 °C for 2 min followed by cooling on ice for 2 min, which was repeated four times. The nicks were then sealed using T4 DNA ligase, by adding 6 μL of 10x ligase buffer and 800 units of ligase. The ligation reaction was left at 20 °C for 4 h. To induce supercoiling into the plasmids, they were treated with 5 units of gyrase for 1 h (see Supplementary Fig. 1 for a schematic of this method). The modified, supercoiled plasmids were purified from the protein and excess insert using an Ultra Clean PCR cleanup kit (Mo Bio) following the manufacturer's protocol. Plasmid concentrations were determined by nanodrop UV-vis measurements.

To confirm the DNA modifications were introduced into the plasmid, we applied a protocol established in our laboratory, in which the modification was removed by a DNA glycosylase to yield a ligatable gap.[4] By ligating the gap, Sanger sequencing provided a characteristic nucleotide loss at the modification site to confirm the presence of the modification. Complete details of the method can be found in our recent publication.[4]

PCR Primer Sequences

Insertion of the *NEIL1* PQS

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCG CTC TTC CCC GGG CCT GGG GGA GGG AAA GGG CCG CTC TTC TGT CAG CCA TGG GGC GGA GAA TGG GCG GAA CTG

5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACA GAA GAG CGG CCC TTT CCC TCC CCC AGG CCC GGG GAA GAG CGC AGA GGC CGA GGC CGC CTC GGC CTC TGA GCT

Insertion of the *BLM* PQS

5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC CGA GGG AGG GGC GGG AGG GAA GCT CTT CTG CAG AGG CCG AGG CCG CCT CGG CCT CTG AGC T

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CTT CCC TCC CGC CCC TCC CTC GGA AGA GCG TCA GCC ATG GGG CGG AGA ATG GGC GGA ACT G

Insertion of the *MGMT* PQS

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCG CTC TTC CGG GGG CGG GGC CGG GGC GCG CGG GGG CGG AGC TGG GGG CTC TTC TGT CAG CCA TGG GGC GGA GAA TGG GCG GAA CTG

5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACA GAA GAG CCC CCA GCT CCG CCC CCG CGC GCC CCG GCC CCG CCC CCG GAA GAG CGC AGA GGC CGA GGC CGC CTC GGC CTC TGA GCT

Insertion of the *NEIL3* PQS

5`- CAG TTC CGC CCA TTC TCC GCC CCA TGG CTG ACG CTC TTC TAG GGT GCT GTT TGG GCG GGG CCT GGG CGG GGC CGC TCT TCT GCA GAG GCC GAG GCC GCC TCG GCC TCT GAG CT

5`- AGC TCA GAG GCC GAG GCG GCC TCG GCC TCT GCA GAA GAG CGG CCC CGC CCA GGC CCC GCC CAA ACA GCA CCC TAG AAG AGC GTC AGC CAT GGG GCG GAG AAT GGG CGG AAC TG

Sequencing Primer

5`- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC ATG CAT CTC AAT TAG TCA GCA ACC ATA GT

Cell Studies. All cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 20 µg/mL gentamicin, 1x glutamax, and 1x non-essential amino acids. The cells were grown at 37 °C with 5% CO_2 at ~80% relative humidity and were split when they reached 70-80% confluence. The wild-type glioblastoma cells (U87 MG) were purchased from ATCC.

The transfection experiments were conducted in white, 96-well plates by seeding 2×10^4 cells per well and then allowing them to grow for 24 h. After 24 h, the cells were transfected with 200-400 ng of plasmid using X-tremeGene HP DNA transfection agent (Roch) following the manufacturer's protocol in Opti-MEM media. All transfection experiments were conducted at least 4 times. Next, 48 h post transfection, the Dual-Glo luciferase (Promega) assay was conducted following the manufacturer's protocol.

The data were analyzed by converting the luminescence measured into normalized relative response ratios (RRR), which is the luminescence of Rluc divided by the luminescence of luc (i.e., RRR = Rluc/luc). To obtain the normalized expression values reported, each RRR was divided by the RRR for the wild type sequence in that data set, for example, normalized expression = RRR_{OG12}/RRR_{WT}. The error bars represent 95% confidence intervals obtained from the data.

Table S1. PQSs in promoters and 5`-UTRs of human DNA repair genes

The sequences and their positions can be found on acs.figshare.com at the following address.

https://figshare.com/s/943e37d7549ad2b5960a

Table S2. The 37 human DNA repair genes without a PQS in their promoters or 5`-UTRs.

ACTR5
ASCC3
ASF1A
ATRX
CDC5L
CDC7
CDK7
DCLRE1B
ERCC3
ERCC5
ERCC6L2
EYA1
FANCL
GTF2H2
GTF2H3
HELQ
KLHL15
MSH4
PARG
PMS2
POLG2
POLI
POLQ
RAD51AP1
RAD54B
RFC3
SFR1
SPO11
SWSAP1
SYCP3
TEX15
TNP1
UBE2T
UBE2U
UBE2V2
XRCC2
XRCC4

Table S3. PQSs in the promoters and 5`-UTRs of all human genes.

The sequences and their positions can be found on acs.figshare.com at the following address.

https://figshare.com/s/943e37d7549ad2b5960a

Table S4. DNA Repair Promoter and 5`-UTR PQSs found in the BG4 G4 ChIP-Seq Data (Balasubramanian's lab Nat. Genet. 2016, 48, 1267).

Chr	Start	End	Strand	Gene Symbol
chr1	2205272	2205358	-	FAAP20
chr1	11262503	11262530	-	MTOR
chr1	11681179	11681217	-	MAD2L2
chr1	11681250	11681300	+	MAD2L2
chr1	23019402	23019425	+	KDM1A
chr1	28088745	28088788	-	EYA3
chr1	45340134	45340173	-	MUTYH
chr1	46247360	46247420	-	RAD54L
chr1	46247476	46247560	-	RAD54L
chr1	153728031	153728068	-	INTS3
chr1	182839184	182839222	-	DHX9
chr1	182839245	182839286	-	DHX9
chr1	226408170	226408247	+	PARP1
chr1	231338110	231338141	-	SPRTN
chr10	67884566	67884623	+	SIRT1
chr10	68527553	68527577	+	SLC25A16
chr10	97498468	97498498	+	MMS19
chr10	97498533	97498558	+	MMS19
chr10	97498639	97498685	+	MMS19
chr10	97498808	97498834	+	MMS19
chr10	100912915	100912947	-	SLF2
chr10	129467054	129467112	-	MGMT
chr10	129467148	129467211	-	MGMT
chr11	18322461	18322484	+	GTF2H1
chr11	46121470	46121542	-	PHF21A
chr11	46121585	46121685	+	PHF21A
chr11	47214847	47214872	-	DDB2
chr11	47214888	47214919	-	DDB2
chr11	47214931	47214958	+	DDB2
chr11	57311477	57311499	-	TNKS1BP1
chr11	57311614	57311642	-	TNKS1BP1
chr11	57311662	57311686	-	TNKS1BP1
chr11	57322381	57322438	-	TNKS1BP1
chr11	57322717	57322739	+	TNKS1BP1
chr11	57324955	57324999	-	TNKS1BP1
chr11	57324955	57324999	-	SSRP1
chr11	61792691	61792742	+	FEN1

chr11	63986529	63986566	+	OTUB1
chr11	63986583	63986614	+	OTUB1
chr11	63999396	63999423	+	OTUB1
chr11	63999442	63999477	-	OTUB1
chr11	63999551	63999585	+	OTUB1
chr11	63999604	63999637	-	OTUB1
chr11	65084238	65084268	+	CDCA5
chr11	65712294	65712326	+	KAT5
chr11	67118964	67119007	+	KDM2A
chr11	67119168	67119193	+	KDM2A
chr11	67239872	67239893	+	KDM2A
chr11	67248818	67248844	-	KDM2A
chr11	67353705	67353749	+	POLD4
chr11	67354857	67354891	+	POLD4
chr11	67374216	67374248	-	RAD9A
chr11	67391936	67391980	+	RAD9A
chr11	69675251	69675283	-	ORAOV1
chr11	73598112	73598148	+	FAM168A
chr11	73598161	73598195	+	FAM168A
chr11	73598522	73598555	-	FAM168A
chr11	73598619	73598663	-	FAM168A
chr11	74592551	74592578	+	POLD3
chr11	108222773	108222823	+	ATM
chr11	113875573	113875615	-	USP28
chr11	119121430	119121478	-	HINFP
chr11	125626298	125626328	+	CHEK1
chr12	991028	991082	-	RAD52
chr12	4321062	4321122	-	TIGAR
chr12	54188985	54189008	-	SMUG1
chr12	55966679	55966712	+	CDK2
chr12	55966751	55966790	+	CDK2
chr12	56221858	56221887	+	NABP2
chr12	56224252	56224289	+	NABP2
chr12	65824118	65824164	-	HMGA2
chr12	65824170	65824210	+	HMGA2
chr12	66302343	66302423	-	HELB
chr12	93442036	93442061	+	UBE2N
chr12	93442064	93442102	-	UBE2N
chr12	93442140	93442181	+	UBE2N
chr12	103965593	103965633	-	TDG
chr12	109098076	109098102	+	UNG
chr12	109098129	109098179	+	UNG
chr12	110501955	110502010	-	RAD9B

chr12	118372969	118372996	-	TAOK3
chr12	118373134	118373157	+	TAOK3
chr12	132687526	132687564	+	POLE
chr13	32586309	32586347	+	PDS5B
chr13	32586484	32586513	+	PDS5B
chr14	20262885	20262907	+	TTC5
chr14	64503828	64503855	+	ZBTB1
chr14	64503887	64503902	-	ZBTB1
chr14	64504126	64504178	+	ZBTB1
chr14	64505243	64505274	+	ZBTB1
chr14	89955839	89955867	+	TDP1
chr14	100238662	100238687	-	YY1
chr14	100240172	100240213	+	YY1
chr14	100240309	100240338	+	YY1
chr14	102592206	102592309	+	RCOR1
chr14	102592422	102592462	-	RCOR1
chr14	102592574	102592621	-	RCOR1
chr14	103715301	103715406	-	XRCC3
chr14	103715764	103715813	+	XRCC3
chr15	30903911	30903940	+	FAN1
chr15	40695141	40695173	+	RAD51
chr15	43510705	43510739	-	TP53BP1
chr15	43510772	43510799	+	TP53BP1
chr15	63504297	63504331	-	USP3
chr15	63504408	63504464	+	USP3
chr15	78872791	78872827	+	MORF4L1
chr15	78873547	78873575	+	MORF4L1
chr15	89334815	89334849	-	POLG
chr15	89334914	89334942	+	POLG
chr16	78159	78199	-	MPG
chr16	1782529	1782555	-	EME2
chr16	1782636	1782658	+	EME2
chr16	8963956	8964020	+	USP7
chr16	27268661	27268726	-	NSMCE1
chr16	29790697	29790753	+	KIF22
chr16	29805401	29805445	+	KIF22
chr16	29805619	29805665	-	KIF22
chr16	29805826	29805861	+	KIF22
chr16	29805866	29805913	-	KIF22
chr16	29806015	29806080	-	KIF22
chr16	29806127	29806153	-	KIF22
chr16	29806189	29806269	-	KIF22
chr16	29806286	29806322	-	KIF22

chr16	29996130	29996171	-	INO80E
chr16	53703765	53703792	-	FTO
chr16	85689247	85689284	-	GINS2
chr17	1829779	1829814	-	RPA1
chr17	1829890	1829914	+	RPA1
chr17	29390330	29390376	+	TAOK1
chr17	29390389	29390413	+	TAOK1
chr17	35119813	35119833	-	RAD51D
chr17	61863483	61863561	-	BRIP1
chr17	65561687	65561723	-	AXIN2
chr17	79797098	79797140	-	CBX8
chr17	79797156	79797191	-	CBX8
chr18	22933341	22933361	+	RBBP8
chr18	32091787	32091818	+	RNF138
chr18	32092366	32092403	+	RNF138
chr18	32092461	32092490	+	RNF138
chr18	32092508	32092547	-	RNF138
chr18	32092615	32092642	-	RNF138
chr18	32092687	32092715	+	RNF138
chr18	35498085	35498115	+	INO80C
chr19	1354732	1354763	+	MUM1
chr19	2163676	2163746	+	DOT1L
chr19	2163679	2163707	-	DOT1L
chr19	2163789	2163862	+	DOT1L
chr19	4909870	4909900	+	UHRF1
chr19	4909972	4910005	-	UHRF1
chr19	4910005	4910051	+	UHRF1
chr19	4910168	4910234	+	UHRF1
chr19	32971944	32971971	-	FAAP24
chr19	38388065	38388118	-	GGN
chr19	38388335	38388357	-	GGN
chr19	45424381	45424435	+	ERCC1
chr19	45450883	45450911	-	ERCC1
chr19	45451043	45451070	+	ERCC1
chr19	48169744	48169793	-	LIG1
chr19	48169807	48169866	-	LIG1
chr19	48170346	48170384	+	LIG1
chr19	48170441	48170473	+	LIG1
chr19	48170610	48170687	+	LIG1
chr19	49869284	49869356	+	PNKP
chr19	49869767	49869803	+	PNKP
chr19	50384134	50384181	+	POLD1
chr19	58544261	58544287	-	TRIM28

chr2	17754049	17754096	+	SMC6
chr2	17754049	17754096	+	GEN1
chr2	47403105	47403151	+	MSH2
chr2	68467283	68467305	-	APLF
chr2	68467440	68467479	+	APLF
chr2	74958909	74959007	+	POLE4
chr2	99490199	99490218	-	REV1
chr2	135531454	135531483	+	ZRANB3
chr2	135531566	135531590	-	ZRANB3
chr2	151409772	151409805	+	RIF1
chr2	191677889	191677921	+	NABP1
chr2	191677971	191678012	+	NABP1
chr2	191678013	191678037	-	NABP1
chr2	191678088	191678108	+	NABP1
chr2	214809683	214809730	+	BARD1
chr2	214809757	214809783	+	BARD1
chr2	214809809	214809830	+	BARD1
chr2	214809979	214810018	+	BARD1
chr2	214810050	214810083	+	BARD1
chr2	219160874	219160896	+	NHEJ1
chr20	1225660	1226027	-	RAD21L1
chr20	5126724	5126754	-	PCNA
chr20	5126835	5126866	+	PCNA
chr20	5126911	5126954	+	PCNA
chr20	5126986	5127010	+	PCNA
chr20	50113088	50113114	-	UBE2V1
chr21	31659535	31659578	-	SOD1
chr21	39349225	39349293	-	HMGN1
chr22	28741972	28742012	-	CHEK2
chr22	28742102	28742153	+	CHEK2
chr22	28742542	28742570	+	CHEK2
chr22	28742590	28742646	+	CHEK2
chr22	28742678	28742703	+	CHEK2
chr22	38570195	38570258	+	DMC1
chr22	38570272	38570401	+	DMC1
chr22	38570433	38570472	+	DMC1
chr22	41621125	41621168	+	XRCC6
chr3	14124778	14124811	-	CHCHD4
chr3	25665035	25665070	+	TOP2B
chr3	122564380	122564413	-	PARP9
chr3	122564380	122564413	-	DTX3L
chr3	129440449	129440482	+	MBD4
chr3	131026837	131026871	-	ASTE1

chr3	131381536	131381561	-	NUDT16
chr3	131381595	131381629	-	NUDT16
chr3	142578877	142578921	+	ATR
chr3	196503826	196503858	+	RNF168
chr3	196503869	196503884	-	RNF168
chr4	1346801	1346841	-	UVSSA
chr4	2242295	2242462	+	POLN
chr4	102827904	102827932	-	UBE2D3
chr4	102868948	102868977	-	UBE2D3
chr4	143513334	143513364	-	SMARCA5
chr4	143513387	143513410	+	SMARCA5
chr4	184649488	184649518	-	PRIMPOL
chr4	184649538	184649722	-	PRIMPOL
chr5	55233838	55233874	+	CCNO
chr5	69369693	69369724	+	RAD17
chr5	75511619	75511707	-	POLK
chr5	80655060	80655131	+	MSH3
chr5	134370963	134370986	-	UBE2B
chr5	134371143	134371198	-	UBE2B
chr6	2765182	2765210	-	WRNIP1
chr6	18264581	18264675	+	DEK
chr6	18264748	18264798	+	DEK
chr6	24666828	24666854	-	TDP2
chr6	30717155	30717184	-	MDC1
chr6	30717517	30717552	-	MDC1
chr6	30914357	30914420	+	GTF2H4
chr6	31739660	31739694	-	MSH5
chr6	31814602	31814670	-	HSPA1A
chr6	31815149	31815176	-	HSPA1A
chr6	31815189	31815228	-	HSPA1A
chr6	34235917	34235955	-	HMGA1
chr6	43576106	43576129	+	POLH
chr6	118935173	118935243	+	MCM9
chr6	158168258	158168293	-	GTF2H5
chr7	2242299	2242322	+	NUDT1
chr7	36152767	36152887	-	EEPD1
chr7	36153046	36153089	-	EEPD1
chr7	36153134	36153176	-	EEPD1
chr7	44082526	44082541	+	POLM
chr7	44082765	44082891	+	POLM
chr7	44123474	44123504	-	POLD2
chr7	44123527	44123557	-	POLD2
chr7	44123837	44123891	-	POLD2

chr7	44123921	44123956	-	POLD2
chr7	47979266	47979317	-	HUS1
chr7	47979333	47979382	-	HUS1
chr7	50450439	50450468	+	FIGNL1
chr7	55018793	55018825	-	EGFR
chr7	55018969	55019007	-	EGFR
chr7	73522749	73522806	+	BAZ1B
chr7	101154381	101154432	+	AP1S1
chr8	27772753	27772784	+	ESCO2
chr8	27774432	27774450	+	ESCO2
chr8	31033468	31033534	+	WRN
chr8	31033665	31033696	+	WRN
chr8	31033722	31033790	+	WRN
chr8	42338410	42338437	-	POLB
chr8	47960136	47960183	-	PRKDC
chr8	73878896	73878958	+	UBE2W
chr8	73878973	73879011	+	UBE2W
chr8	89984242	89984281	+	NBN
chr8	102239058	102239087	-	RRM2B
chr8	102239567	102239616	-	RRM2B
chr8	102412741	102412792	+	UBR5
chr8	102412988	102413012	+	UBR5
chr8	116874876	116874897	-	RAD21
chr8	144291346	144291375	-	HSF1
chr8	144291444	144291463	-	HSF1
chr8	144291470	144291491	+	HSF1
chr8	144291515	144291542	+	HSF1
chr8	144428429	144428482	-	TONSL
chr8	144428630	144428657	-	TONSL
chr8	144428722	144428744	-	TONSL
chr8	144517879	144517908	+	RECQL4
chr8	144518034	144518077	+	RECQL4
chr9	33025142	33025172	-	APTX
chr9	33025205	33025237	+	APTX
chr9	70258881	70258898	+	SMC5
chr9	95317012	95317044	+	FANCC
chr9	95317224	95317285	-	FANCC
chr9	95317770	95317802	+	FANCC
chr9	97697428	97697470	+	XPA
chr9	107282814	107282862	-	RAD23B
chr9	107283222	107283243	-	RAD23B
chr9	112718129	112718180	-	INIP
chr9	112718281	112718348	+	INIP

chr9	113410539	113410567	-	POLE3
chr9	127785049	127785095	-	CDK9
chr9	127785589	127785624	-	CDK9
chr9	127785666	127785693	-	CDK9
chr9	127785969	127786007	+	CDK9
chr9	128275943	128275970	+	SWI5
chr9	128276221	128276251	+	SWI5
chr9	132355255	132355285	+	SETX
chrX	14873248	14873283	+	FANCB
chrX	24693627	24693658	-	POLA1
chrX	71283351	71283405	-	NONO
chrX	152830649	152830684	+	CETN2
chrX	152830771	152830803	-	CETN2
chrX	152831023	152831052	+	CETN2
chrX	153470637	153470683	-	TREX2
chrX	153470710	153470744	-	TREX2
chrX	153470766	153470896	+	TREX2
chrX	155071053	155071096	+	BRCC3
chrX	155071148	155071196	-	BRCC3
chrX	155071281	155071311	-	BRCC3
chrX	155071327	155071405	-	BRCC3

Table S5. DNA Repair Promoter and 5`-UTR PQSs found in the XPB/XPD ChIP-Seq Data (Maizel's lab Nat. Chem. Biol. 2014, 10, 313).

Chr	Start	End	Strand	Gene Symbol
chr1	2212739	2212779	+	FAAP20
chr1	11681688	11681714	-	MAD2L2
chr1	11692679	11692734	-	MAD2L2
chr1	11692947	11692965	+	MAD2L2
chr1	11692978	11693000	+	MAD2L2
chr1	23019402	23019425	+	KDM1A
chr1	27915164	27915192	-	RPA2
chr1	28088745	28088788	-	EYA3
chr1	45340134	45340173	-	MUTYH
chr1	62435971	62435997	+	USP1
chr1	2.31E+08	2.31E+08	-	SPRTN
chr1	2.42E+08	2.42E+08	+	EXO1
chr10	49539841	49539873	+	ERCC6
chr10	60778568	60778625	+	CDK1
chr10	60778672	60778731	+	CDK1
chr10	97499549	97499585	+	MMS19
chr10	97500336	97500384	+	MMS19
chr10	1.11E+08	1.11E+08	+	SMC3
chr11	6402857	6402886	+	APBB1
chr11	6403381	6403416	+	APBB1
chr11	18322461	18322484	+	GTF2H1
chr11	43880471	43880512	-	ALKBH3
chr11	43880608	43880648	-	ALKBH3
chr11	47214847	47214872	-	DDB2
chr11	47214888	47214919	-	DDB2
chr11	47214931	47214958	+	DDB2
chr11	57304109	57304134	-	TNKS1BP1
chr11	57323563	57323590	-	TNKS1BP1
chr11	61333353	61333396	-	DDB1
chr11	61792408	61792431	-	FEN1
chr11	61792691	61792742	+	FEN1
chr11	65712294	65712326	+	KAT5
chr11	65859918	65859954	-	MUS81
chr11	65860163	65860193	-	MUS81
chr11	65860635	65860666	-	MUS81
chr11	67239872	67239893	+	KDM2A
chr11	67317458	67317502	-	RAD9A

chr11	67354857	67354891	+	POLD4
chr11	67355440	67355464	+	POLD4
chr11	67355468	67355498	-	POLD4
chr11	67357225	67357266	-	POLD4
chr11	67374216	67374248	-	RAD9A
chr11	67374701	67374734	+	RAD9A
chr11	67374744	67374771	-	RAD9A
chr11	69670101	69670150	+	ORAOV1
chr11	73598522	73598555	-	FAM168A
chr11	73598619	73598663	-	FAM168A
chr11	75404880	75404909	+	RPS3
chr11	1.14E+08	1.14E+08	-	USP28
chr11	1.26E+08	1.26E+08	+	CHEK1
chr11	1.26E+08	1.26E+08	-	CHEK1
chr12	2877326	2877391	+	FOXM1
chr12	56221135	56221163	-	NABP2
chr12	56221273	56221299	+	NABP2
chr12	56221388	56221452	-	NABP2
chr12	56221858	56221887	+	NABP2
chr12	56222284	56222336	-	NABP2
chr12	65824170	65824210	+	HMGA2
chr12	65824636	65824664	-	HMGA2
chr12	65824873	65824903	-	HMGA2
chr12	66302343	66302423	-	HELB
chr12	1.02E+08	1.02E+08	-	PARPBP
chr12	1.04E+08	1.04E+08	-	TDG
chr12	1.18E+08	1.18E+08	-	TAOK3
chr12	1.18E+08	1.18E+08	+	TAOK3
chr13	30618153	30618182	+	HMGB1
chr14	20457954	20457990	+	APEX1
chr14	20458194	20458224	-	APEX1
chr14	1.03E+08	1.03E+08	-	RCOR1
chr15	40694830	40694864	-	RAD51
chr15	40695141	40695173	+	RAD51
chr15	40695329	40695356	+	RAD51
chr15	63157494	63157529	+	RPS27L
chr15	63157561	63157613	+	RPS27L
chr15	63503564	63503604	-	USP3
chr15	63504066	63504125	-	USP3
chr15	63504297	63504331	-	USP3
chr15	63504408	63504464	+	USP3
chr15	74043808	74043839	+	PML
chr15	89334914	89334942	+	POLG

chr16	23641657	23641691	-	PALB2
chr16	27268547	27268608	-	NSMCE1
chr16	27268661	27268726	-	NSMCE1
chr16	29790697	29790753	+	KIF22
chr16	29814471	29814516	+	PAGR1
chr16	29814587	29814623	-	PAGR1
chr16	29814691	29814741	-	PAGR1
chr16	29815239	29815267	+	PAGR1
chr16	29815358	29815386	+	PAGR1
chr16	30075483	30075513	+	PPP4C
chr16	30076134	30076167	+	PPP4C
chr16	70523293	70523334	-	SF3B3
chr16	89921476	89921514	+	MC1R
chr17	28335928	28335963	+	TNFAIP1
chr17	58691874	58691907	-	RAD51C
chr17	58692004	58692042	+	RAD51C
chr17	58692125	58692162	+	RAD51C
chr17	75634227	75634254	+	RECQL5
chr17	75634383	75634412	+	RECQL5
chr17	81552623	81552648	+	FAAP100
chr17	81552939	81552979	+	FAAP100
chr17	81553173	81553195	-	FAAP100
chr17	81553243	81553276	+	FAAP100
chr17	81554301	81554327	+	FAAP100
chr18	35478563	35478595	-	INO80C
chr18	35498085	35498115	+	INO80C
chr18	35498295	35498319	-	INO80C
chr19	4910339	4910408	+	UHRF1
chr19	7629196	7629227	-	XAB2
chr19	7629878	7629900	+	XAB2
chr19	7630016	7630078	+	XAB2
chr19	12949710	12949732	+	RAD23A
chr19	17267550	17267582	-	BABAM1
chr19	32971757	32971808	-	FAAP24
chr19	32971944	32971971	-	FAAP24
chr19	38388893	38388918	-	GGN
chr19	38389127	38389171	-	GGN
chr19	38389283	38389315	-	GGN
chr19	38389486	38389514	+	GGN
chr19	38389530	38389563	-	GGN
chr19	38389699	38389721	-	GGN
chr19	38389770	38389803	+	GGN
chr19	38389915	38390062	+	GGN

chr19	45406722	45406749	+	ERCC1
chr19	45422107	45422139	-	ERCC1
chr19	45424381	45424435	+	ERCC1
chr19	48999245	48999279	+	RUVBL2
chr19	49865100	49865147	-	PNKP
chr19	49865353	49865382	-	PNKP
chr19	50383119	50383154	+	POLD1
chr19	54115437	54115464	-	TFPT
chr19	54115808	54115849	+	TFPT
chr19	58547235	58547274	-	TRIM28
chr19	58547721	58547751	+	TRIM28
chr19	58548525	58548563	-	TRIM28
chr2	53970787	53970842	-	PSME4
chr2	1.01E+08	1.01E+08	-	NPAS2
chr2	1.51E+08	1.51E+08	+	RIF1
chr2	1.92E+08	1.92E+08	+	NABP1
chr2	1.92E+08	1.92E+08	+	NABP1
chr2	1.92E+08	1.92E+08	+	NABP1
chr2	1.92E+08	1.92E+08	-	NABP1
chr2	1.92E+08	1.92E+08	+	NABP1
chr2	1.92E+08	1.92E+08	-	NABP1
chr2	1.92E+08	1.92E+08	-	NABP1
chr20	50086237	50086286	-	UBE2V1
chr20	50113155	50113192	-	UBE2V1
chr20	63657835	63657877	+	RTEL1
chr20	63657920	63657956	+	RTEL1
chr20	63658105	63658138	+	RTEL1
chr21	31659535	31659578	-	SOD1
chr22	28741972	28742012	-	CHEK2
chr22	28742102	28742153	+	CHEK2
chr22	28742542	28742570	+	CHEK2
chr22	28742590	28742646	+	CHEK2
chr22	28742678	28742703	+	CHEK2
chr22	41621125	41621168	+	XRCC6
chr3	4303529	4303559	+	SETMAR
chr3	9749339	9749372	-	OGG1
chr3	14124778	14124811	-	CHCHD4
chr3	14125234	14125268	-	CHCHD4
chr3	14125734	14125765	+	CHCHD4
chr3	14125992	14126031	-	CHCHD4
chr3	48465360	48465386	+	TREX1
chr3	48465919	48465965	+	TREX1
chr3	48466099	48466136	+	TREX1

chr3	48466180	48466214	+	TREX1
chr3	51942882	51942925	-	PARP3
chr3	51943215	51943248	+	PARP3
chr3	51943271	51943306	+	PARP3
chr3	51943833	51943886	-	PARP3
chr3	51944040	51944086	-	PARP3
chr3	1.23E+08	1.23E+08	-	DTX3L
chr3	1.23E+08	1.23E+08	+	DTX3L
chr3	1.23E+08	1.23E+08	-	DTX3L
chr3	1.23E+08	1.23E+08	+	PARP9
chr3	1.23E+08	1.23E+08	+	DTX3L
chr3	1.23E+08	1.23E+08	-	PARP9
chr3	1.23E+08	1.23E+08	-	DTX3L
chr3	1.28E+08	1.28E+08	-	RUVBL1
chr3	1.28E+08	1.28E+08	-	RUVBL1
chr3	1.29E+08	1.29E+08	+	MBD4
chr3	1.31E+08	1.31E+08	-	ASTE1
chr3	1.31E+08	1.31E+08	-	NUDT16
chr3	1.31E+08	1.31E+08	-	NUDT16
chr3	1.87E+08	1.87E+08	+	RFC4
chr3	1.87E+08	1.87E+08	+	RFC4
chr4	1.03E+08	1.03E+08	-	UBE2D3
chr4	1.03E+08	1.03E+08	-	UBE2D3
chr4	1.03E+08	1.03E+08	-	UBE2D3
chr4	1.85E+08	1.85E+08	-	PRIMPOL
chr5	892151	892184	-	TRIP13
chr5	892257	892295	+	TRIP13
chr5	892419	892495	-	TRIP13
chr5	892512	892575	-	TRIP13
chr5	34915334	34915363	-	RAD1
chr5	60945606	60945640	-	ERCC8
chr5	60945764	60945789	+	ERCC8
chr5	69369416	69369458	-	RAD17
chr5	69369577	69369647	-	RAD17
chr5	69369693	69369724	+	RAD17
chr5	69370030	69370047	-	RAD17
chr5	1.33E+08	1.33E+08	-	RAD50
chr5	1.34E+08	1.34E+08	-	UBE2B
chr5	1.34E+08	1.34E+08	-	UBE2B
chr5	1.34E+08	1.34E+08	-	UBE2B
chr6	30717155	30717184	-	MDC1
chr6	30717517	30717552	-	MDC1
chr6	30742265	30742297	-	IER3

chr6	30742405	30742452	-	IER3
chr6	30742554	30742605	+	IER3
chr6	30742636	30742670	-	IER3
chr6	30742715	30742775	-	IER3
chr6	30743075	30743111	+	IER3
chr6	31739660	31739694	-	MSH5
chr6	31740327	31740351	-	MSH5
chr6	31740537	31740565	-	MSH5
chr6	34235632	34235655	-	HMGA1
chr6	34235717	34235761	+	HMGA1
chr6	34235917	34235955	-	HMGA1
chr6	34236462	34236504	-	HMGA1
chr6	34236701	34236744	-	HMGA1
chr6	1.11E+08	1.11E+08	+	REV3L
chr6	1.11E+08	1.11E+08	+	REV3L
chr7	2244767	2244792	+	NUDT1
chr7	4775368	4775393	-	AP5Z1
chr7	44072467	44072498	+	POLM
chr7	44072739	44072765	-	POLM
chr7	44073179	44073208	+	POLM
chr7	73522749	73522806	+	BAZ1B
chr8	47960136	47960183	-	PRKDC
chr8	1.02E+08	1.02E+08	+	UBR5
chr9	1.13E+08	1.13E+08	-	POLE3

Table S6. Genes with PQSs common to the BG4 ChIP-Seq and XPB/XPD ChIP-Seq Data Sets.

ASTE1
BAZ1B
CHCHD4
CHEK1
CHEK2
DDB2
DTX3L
ERCC1
EYA3
FAAP20
FAAP24
FAM168A
FEN1
GGN
GTF2H1
HELB
HMGA1
HMGA2
INO80C
KAT5
KDM1A
KDM2A
KIF22
MAD2L2
MBD4
MDC1
MMS19
MSH5
MUTYH
NABP1
NABP2
NSMCE1
NUDT1
NUDT16
ORAOV1
PARP9
PNKP
POLD1
POLD4

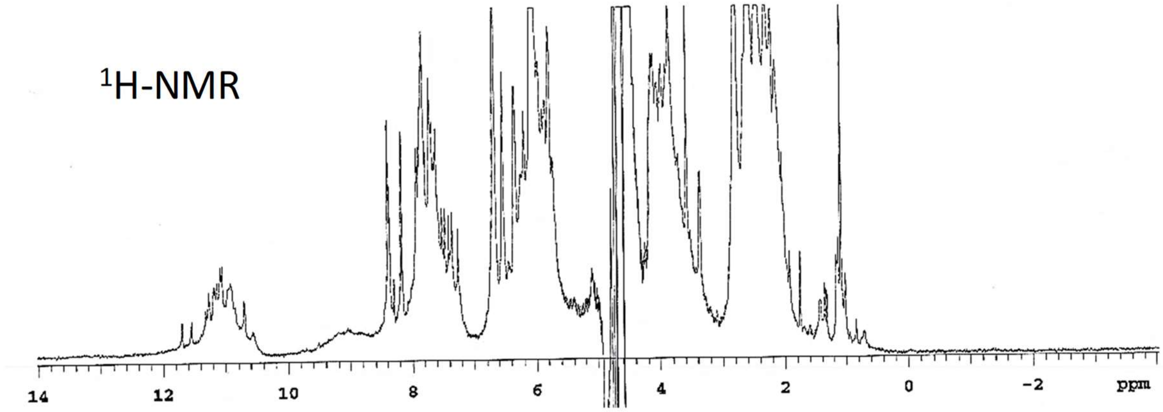
POLE3
POLG
POLM
PRIMPOL
PRKDC
RAD17
RAD51
RAD9A
RCOR1
RIF1
SOD1
SPRTN
TAOK3
TDG
TNKS1BP1
TRIM28
UBE2B
UBE2D3
UBE2V1
UBR5
UHRF1
USP28
USP3
XRCC6

Figure S1. ^{1}H-NMR, CD, TDS, T_m, and ThT data for the 30 PQSs studied.

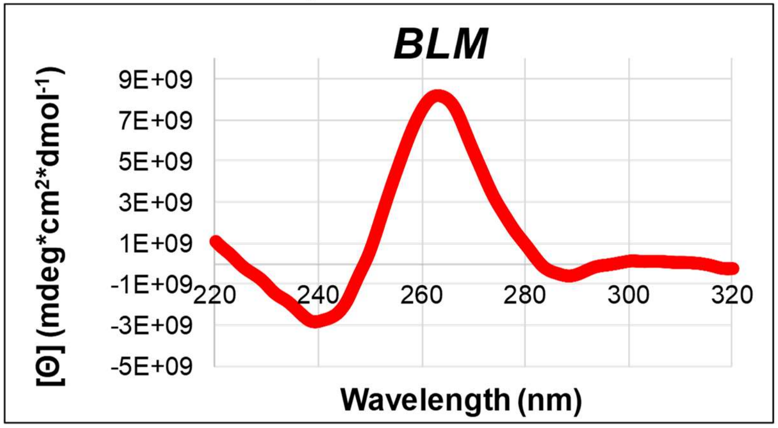
BLM (coding strand)

Sequence
5`-GA GGG A GGGG C GGG A GGG AA

^{1}H-NMR



CD Spectrum



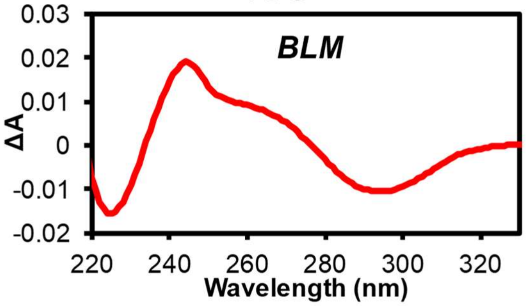
T_m Value

81.8 ± 0.3 °C

ThT Data

Sample	ThT	ssDNA	dsDNA	*cMYC* G4	*BLM* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	32 ± 4

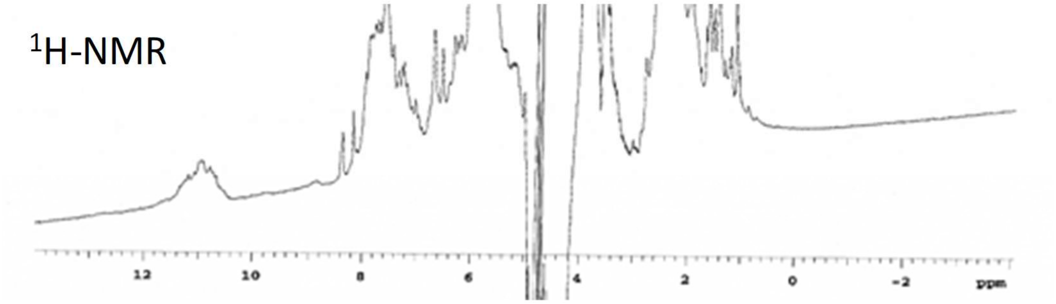
TDS



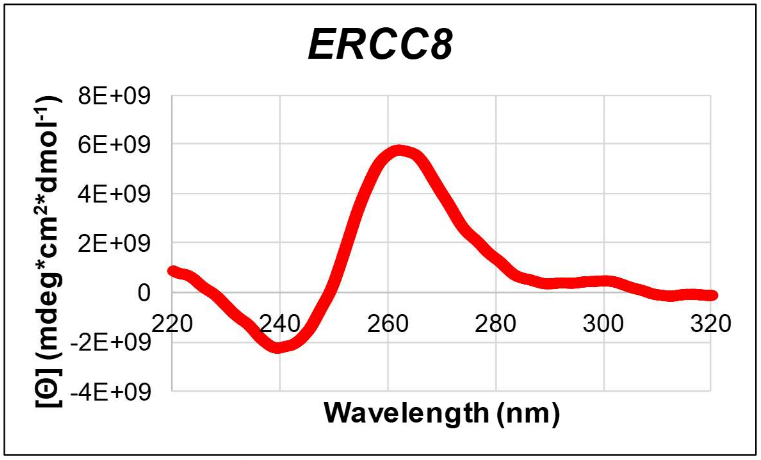
ERCC8
(template strand)

Sequence
5`-AA GGG CTAGAA GGG CCA GGGG A GGGG TT

^{1}H-NMR



CD Spectrum



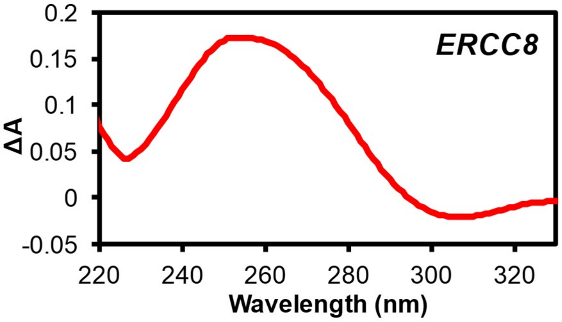
T$_m$ Value

62.3 ± 0.4 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*ERCC8* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	41 ± 5

TDS

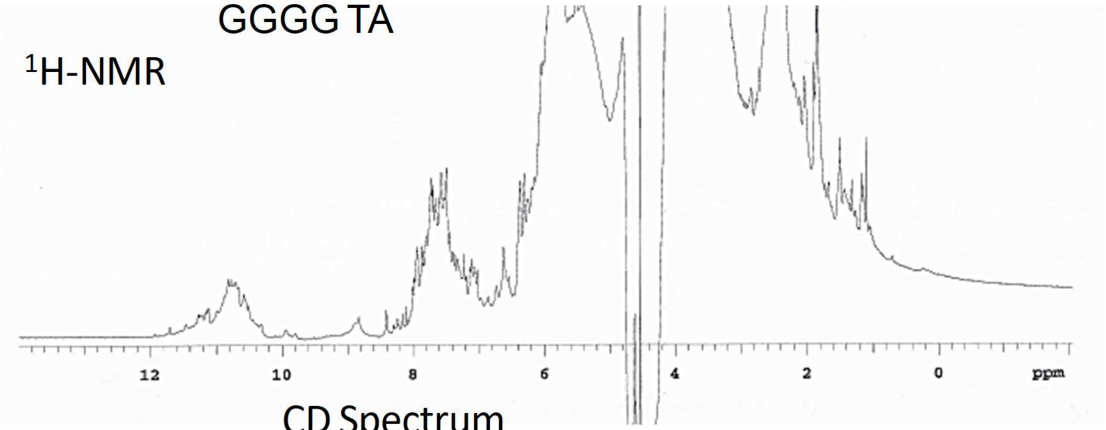


FAAP24 (coding strand)

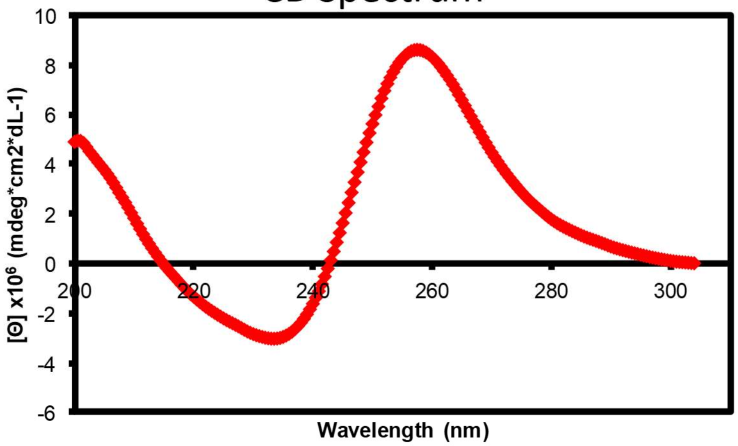
Sequence

5`-GT GGGG CTCTGT GGGG CC GGG AATTA GGGG TA

^{1}H-NMR



CD Spectrum



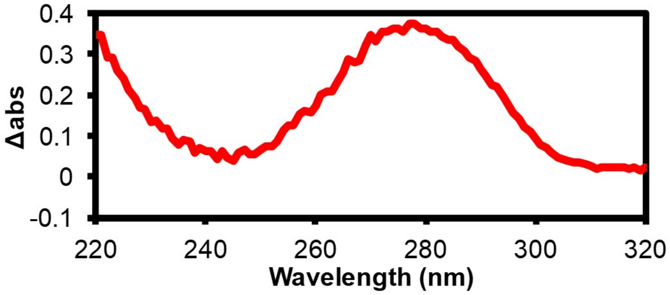
T$_m$ Value

57.1 ± 0.8 °C

ThT Data

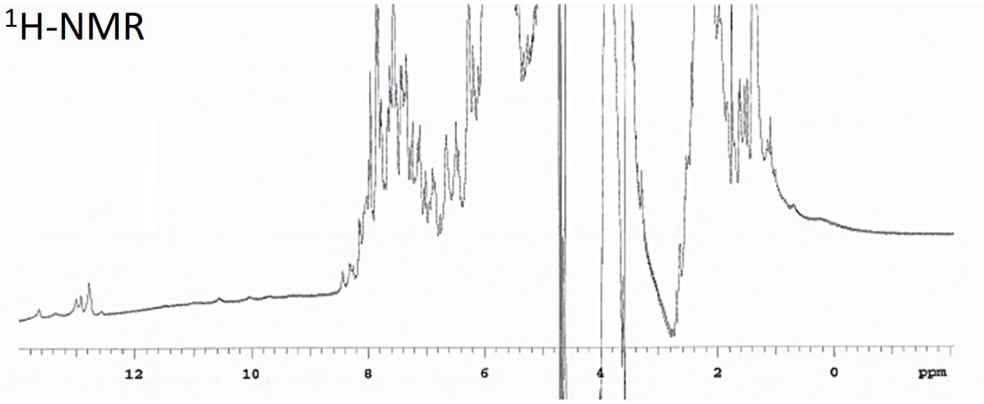
Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*C19orf40* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	53 ± 2.2

TDS

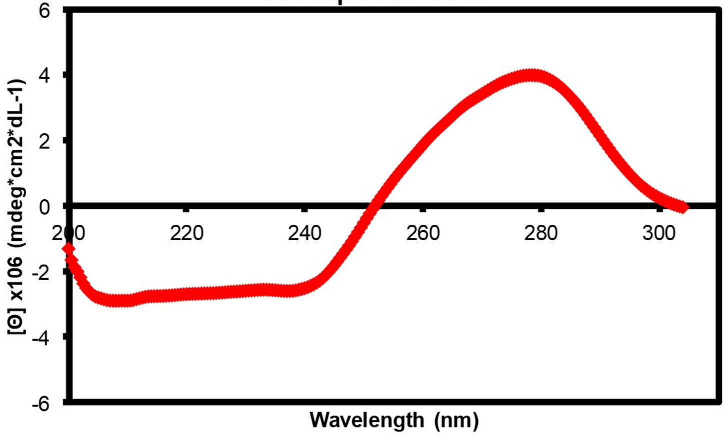


FANCA	Sequence
(coding strand)	5`-GC GGG CTC GGG CGCA GGG AGCCGCCGCC GGGG CT

^{1}H-NMR



CD Spectrum



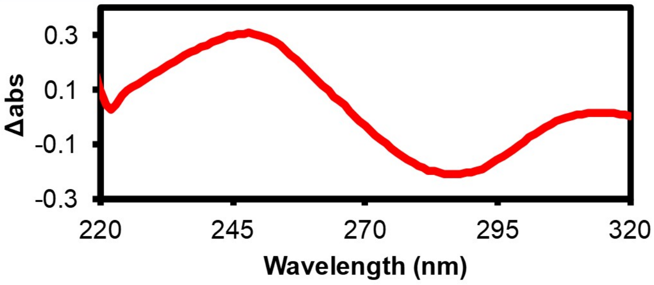
T_m Value

A temperature-dependent transition was not observed when monitoring the Abs at 295 nm.

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*FANCA* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	15.9 ± 0.2

TDS

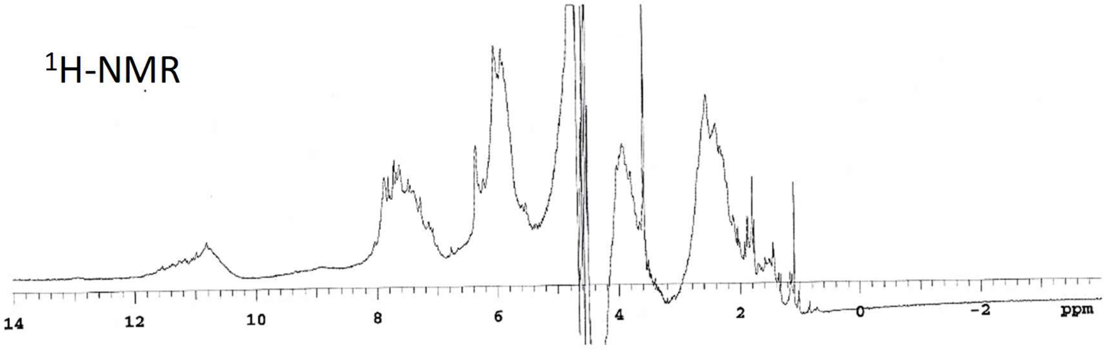


FANCC (template strand)

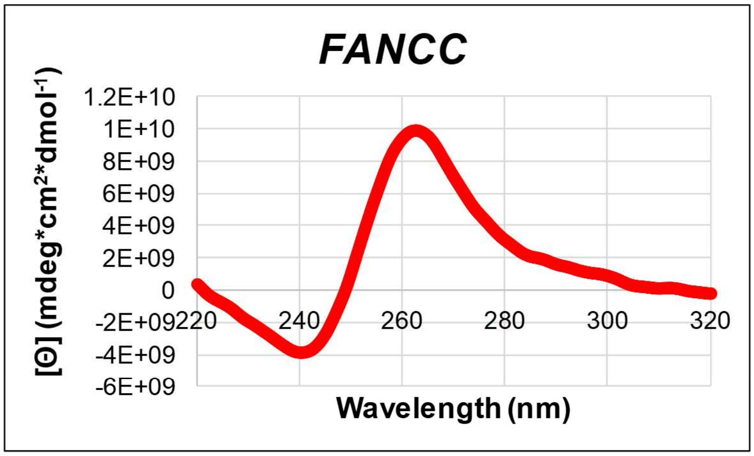
Sequence

5`-TC GGG TCCGT GGGG C GGGG C GGG CG

^{1}H-NMR



CD Spectrum



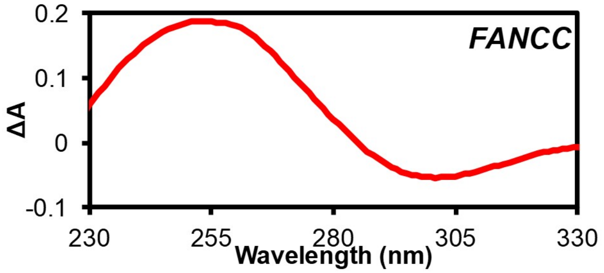
T_m Value

73.1 ± 0.6 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*FANCC* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	25 ± 2

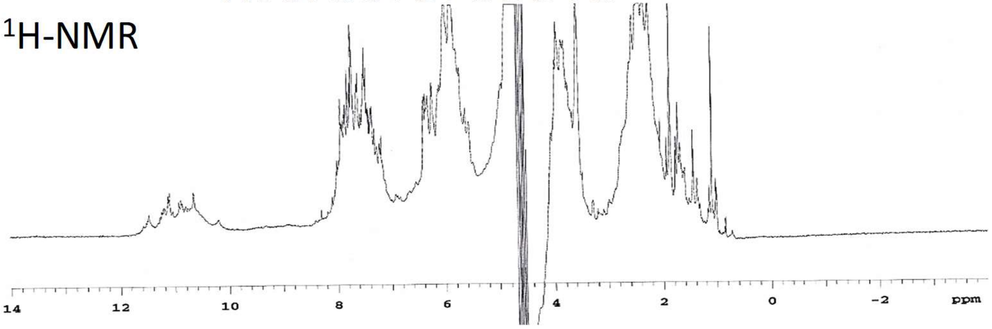
TDS



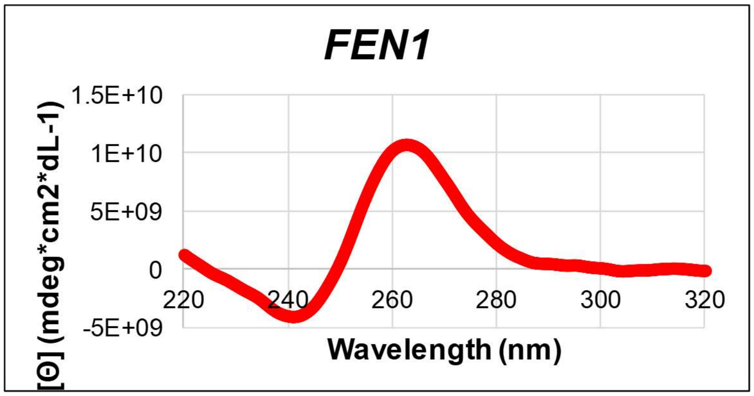
FEN1
(template strand)

Sequence
5`-GA GGG AGCA GGGG CCTGC GGGG
ACCTGGTGT GGG TG

^{1}H-NMR



CD Spectrum



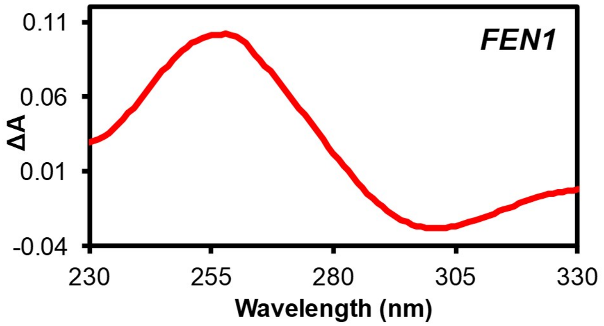
T_m Value

69.0 ± 0.4 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*FEN1* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	108 ± 4

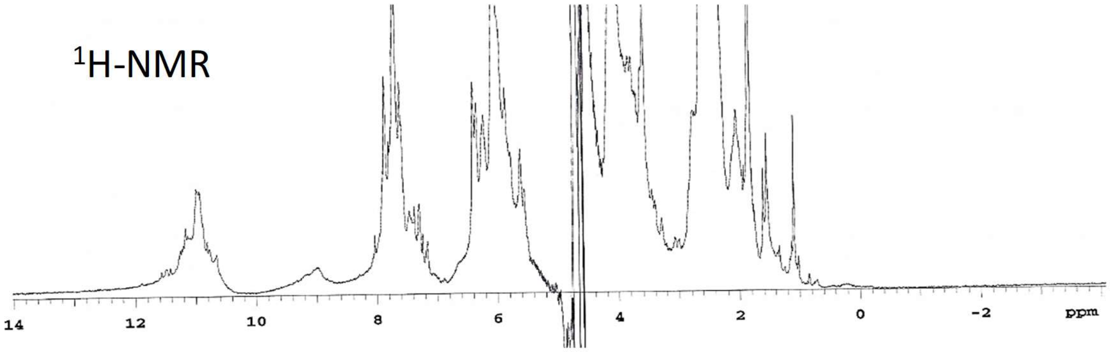
TDS



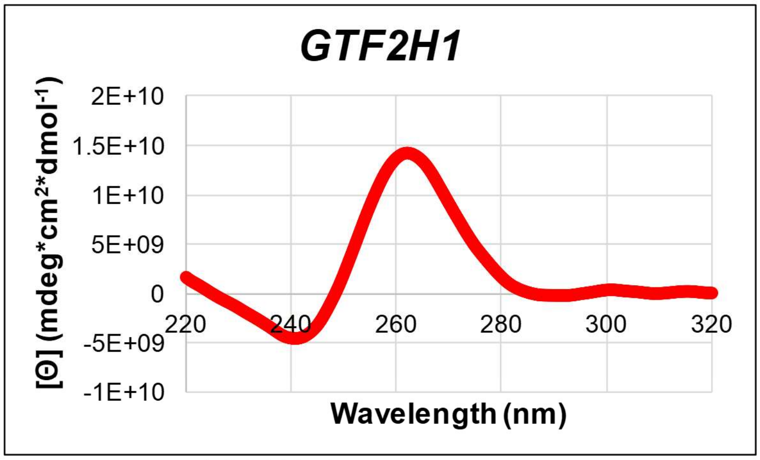
GTF2H1
(coding strand)

Sequence
5`-GC GGG AACCCGT GGGGG A GGG A GGG AA

^{1}H-NMR



CD Spectrum



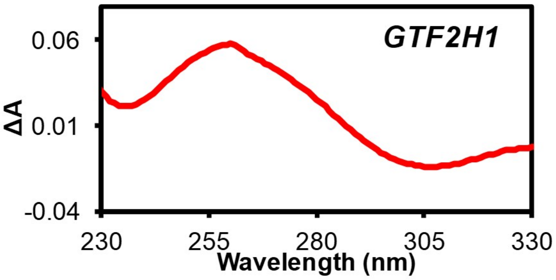
T$_m$ Value

68.1 ± 0.4 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*GTF2 H1* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	15.0 ± 0.1

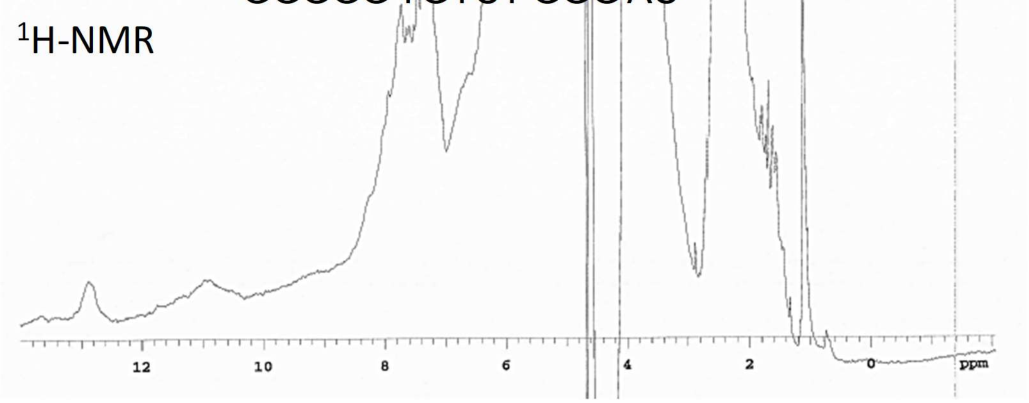
TDS



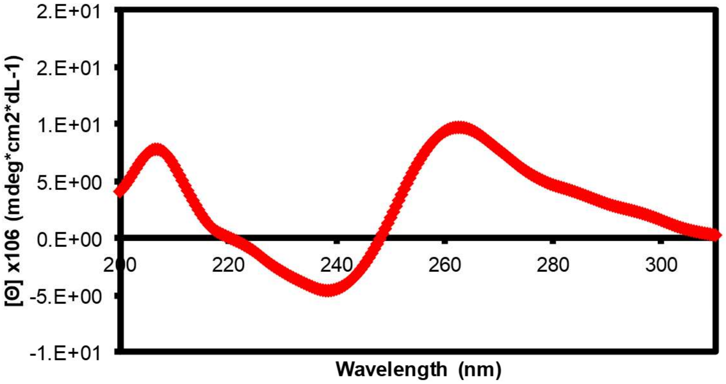
LIG4
(coding strand)

Sequence
5`-TT GGGGG TCT GGGGG ATCCGGTCGT GGGGG TGTCT GGG AC

^{1}H-NMR



CD Spectrum



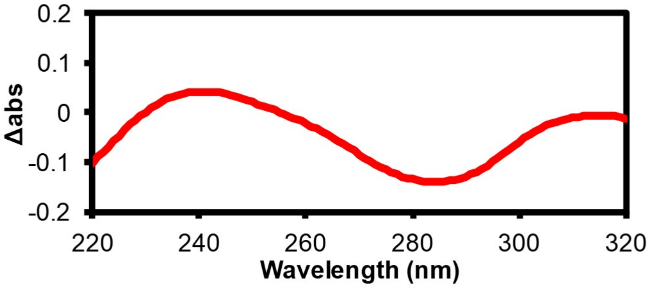
T_m Value

62.6 ± 1.2 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*LIG4* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	8.7 ± 0.1

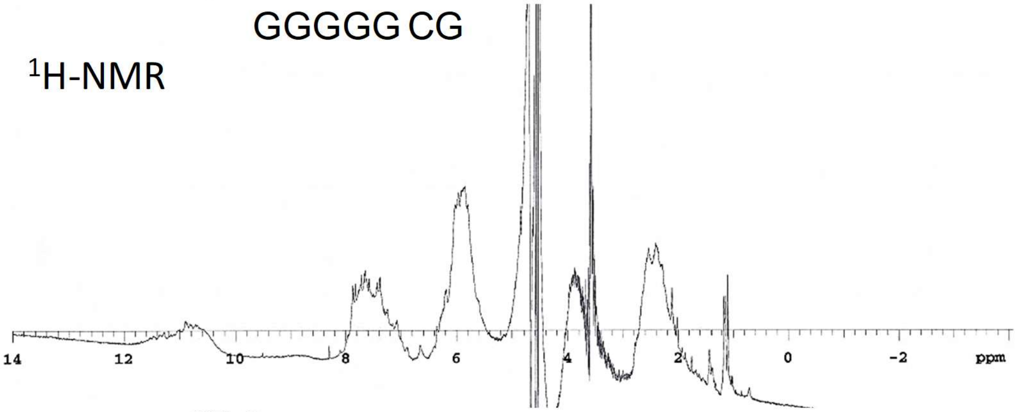
TDS



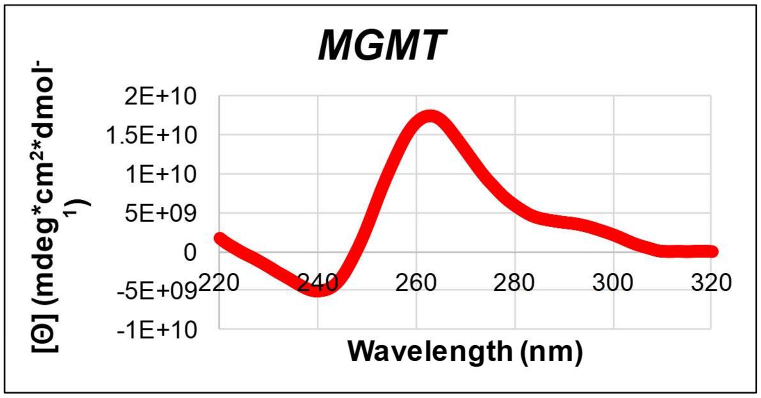
MGMT
(template strand)

Sequence
5`-CC GGGGG C GGGG CC GGGG CGCGC GGGGG CG

^{1}H-NMR



CD Spectrum



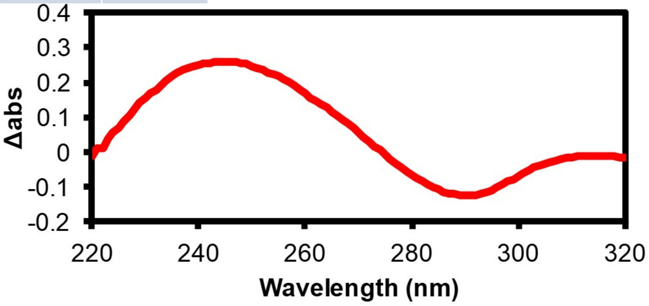
T$_m$ Value

75.4 ± 0.5 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*MGMT* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	27.0 ± 0.1

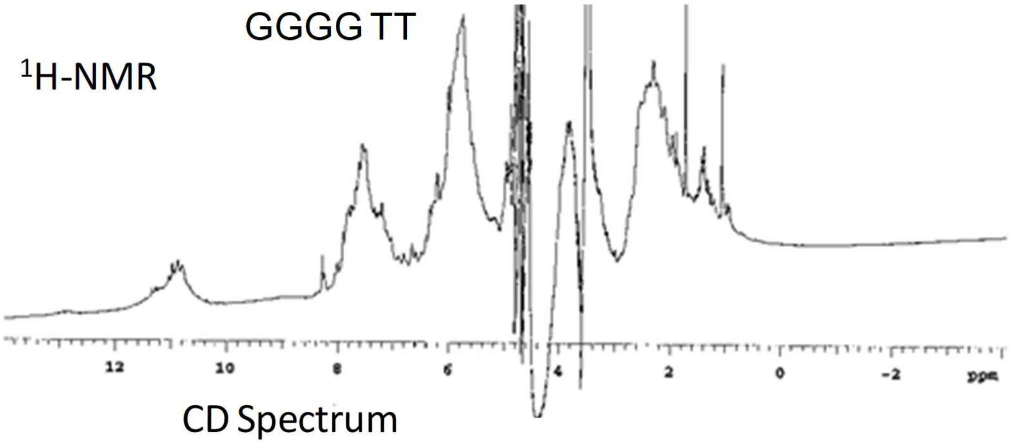
TDS



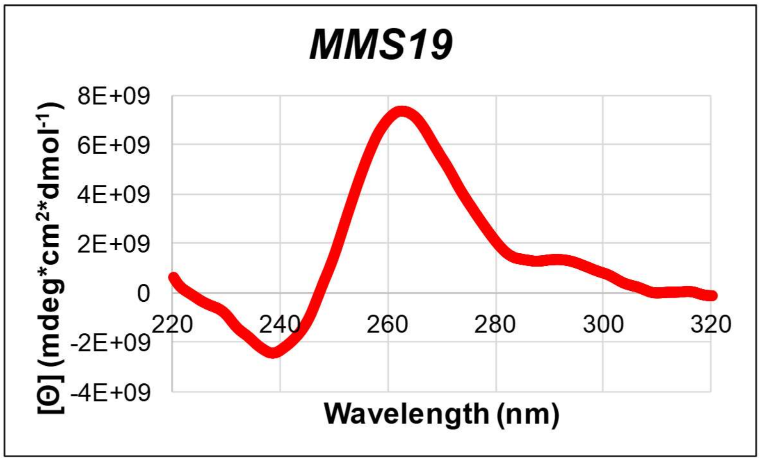
MMS19
(template strand)

Sequence
5`-AA GGG AGA GGGG CCGGCCT GGGGG C GGGG TT

^{1}H-NMR



CD Spectrum



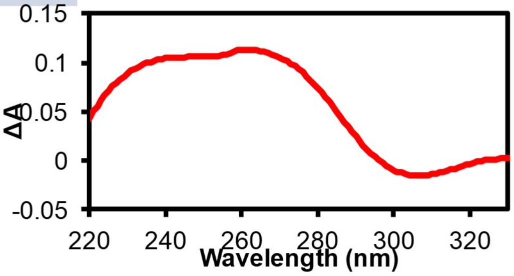
T_m Value

72.3 ± 1.0 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*MMS 19* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	38 ± 3

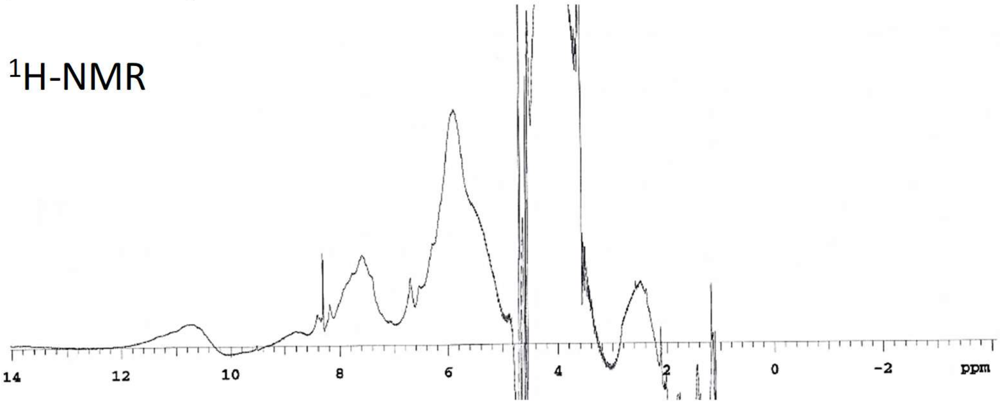
TDS



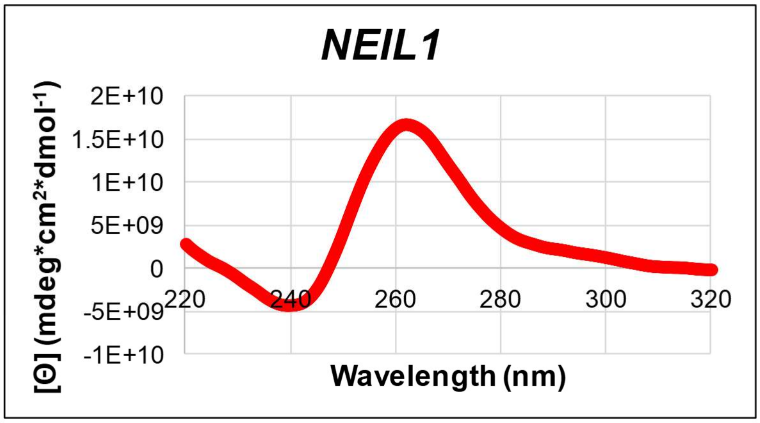
NEIL1
(template strand)

Sequence
5`-CC GGG CCT GGGGG A GGG AAA GGG CC

^{1}H-NMR



CD Spectrum



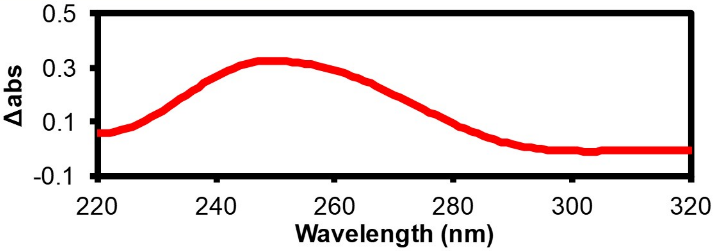
T$_m$ Value

70.1 ± 0.6 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*NEIL1* G4
λ$_{em}$ at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	76 ± 6

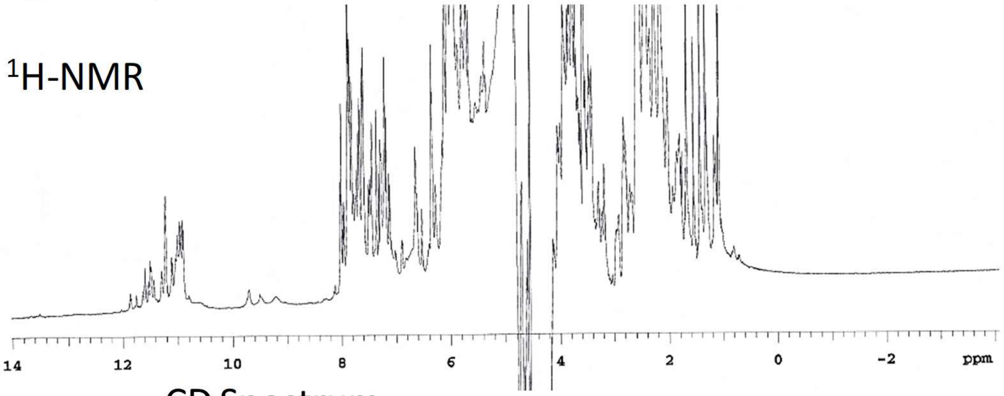
TDS



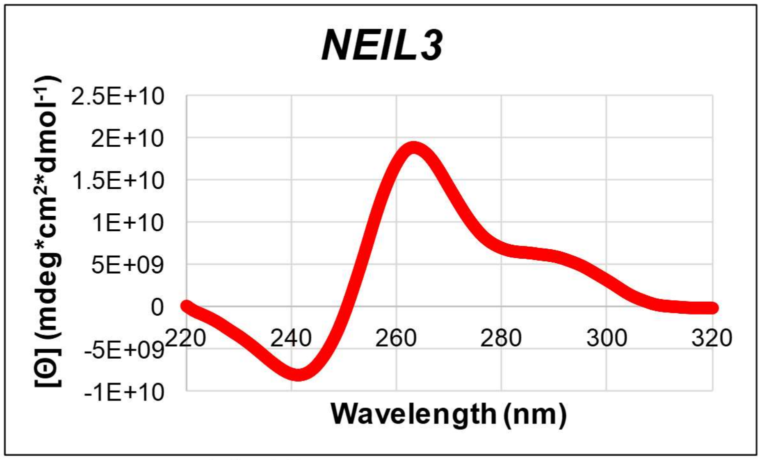
NEIL3
(coding strand)

Sequence
5`-TT GGG C GGGG CCT GGG C GGGG CC

^{1}H-NMR



CD Spectrum



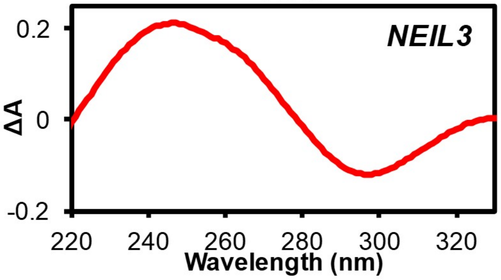
T_m Value

81.5 ± 0.7 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*NEIL3* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	28 ± 1

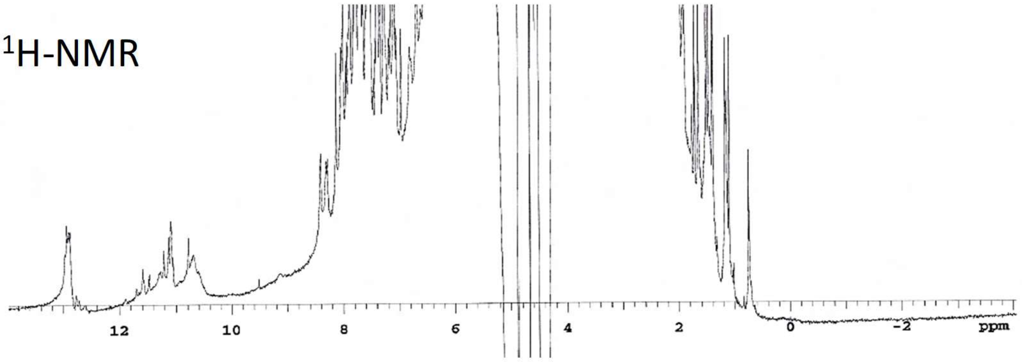
TDS



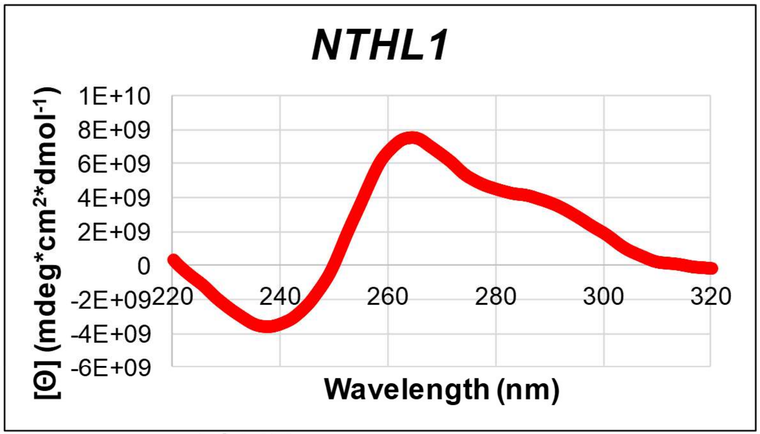
NTHL1
(coding strand)

Sequence
5`-GT GGG CGC GGG TGA GGG CCC GGG AC

^{1}H-NMR



CD Spectrum



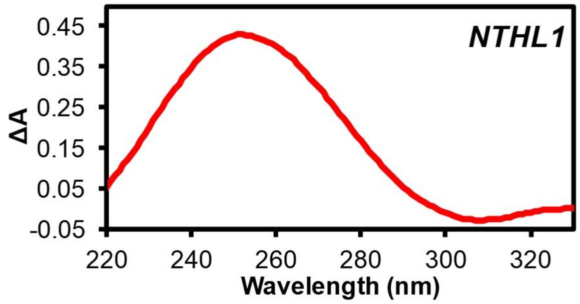
T$_m$ Value

70.9 ± 0.8 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*NTHL 1* G4
λ$_{em}$ at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	34 ± 1

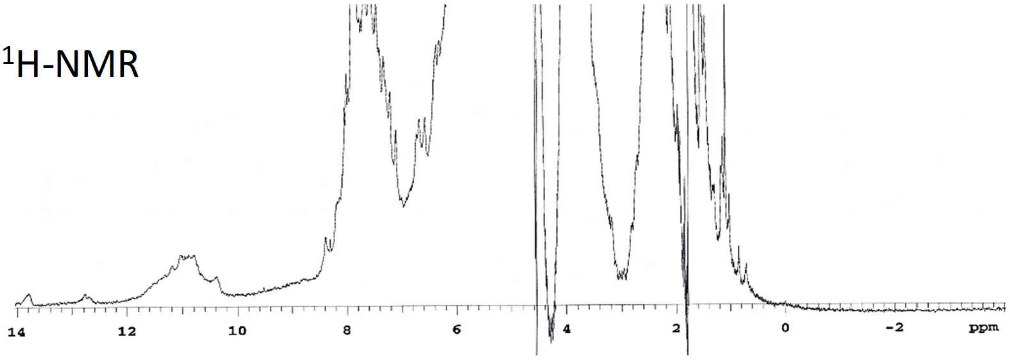
TDS



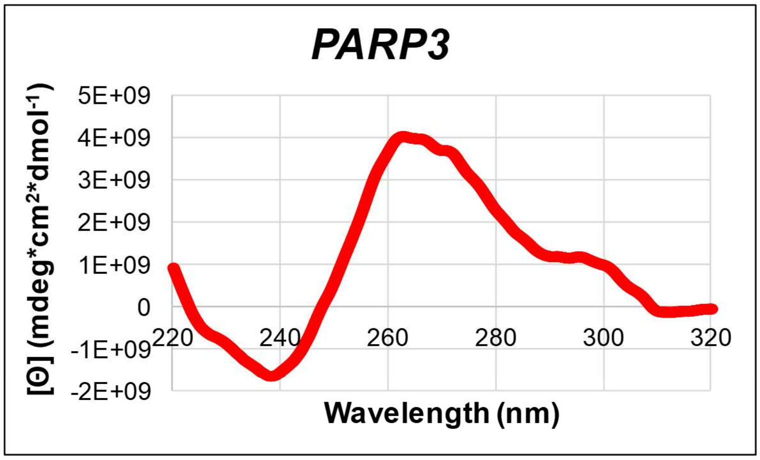
PARP3
(coding strand)

Sequence
5`-AA GGG CT GGGG AA GGG CC GGG AC

^{1}H-NMR



CD Spectrum

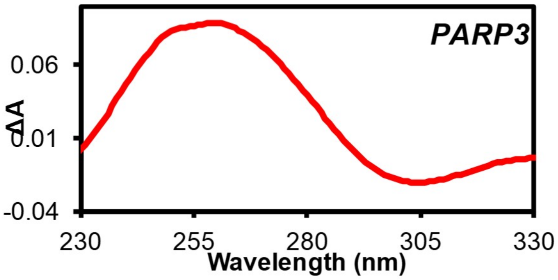


T_m Value
64.3 ± 1.0 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*PARP 3* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	27 ± 1

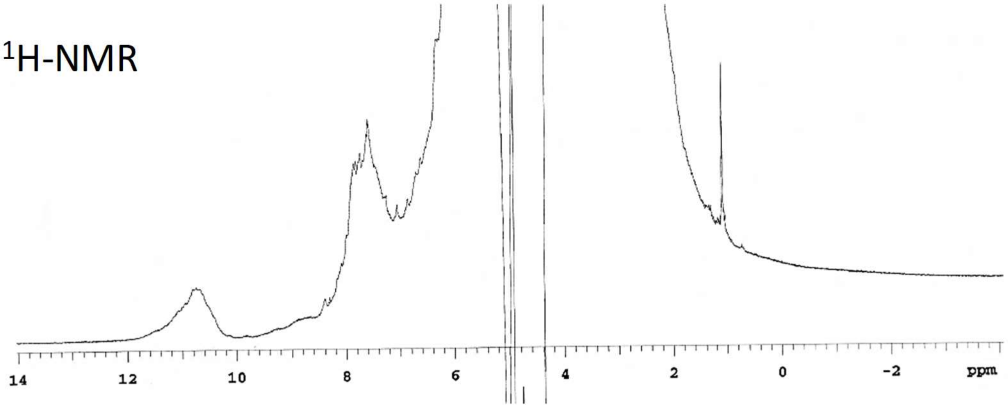
TDS



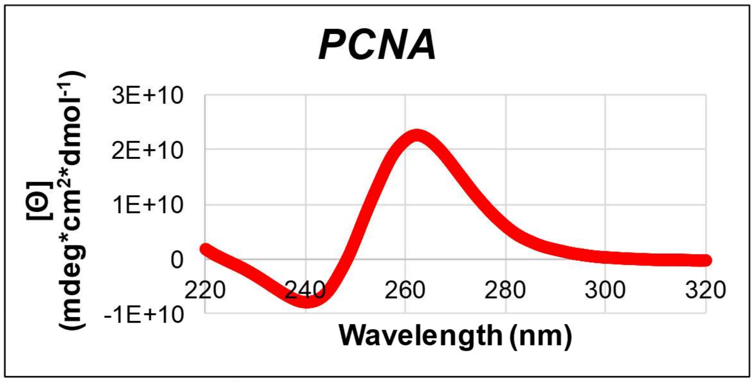
PCNA
(coding strand)

Sequence
5`-CA GGG CGAC GGGGG C GGGG C GGGG CG

^{1}H-NMR



CD Spectrum



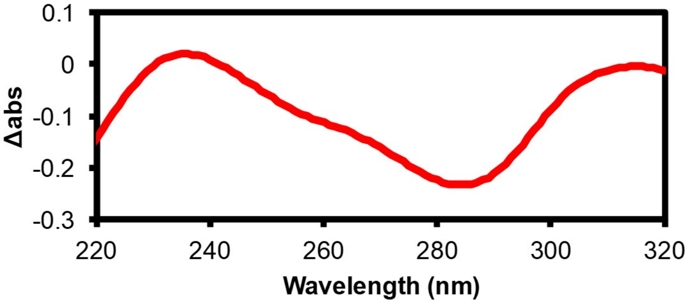
T_m Value

72.7 ± 0.6 °C

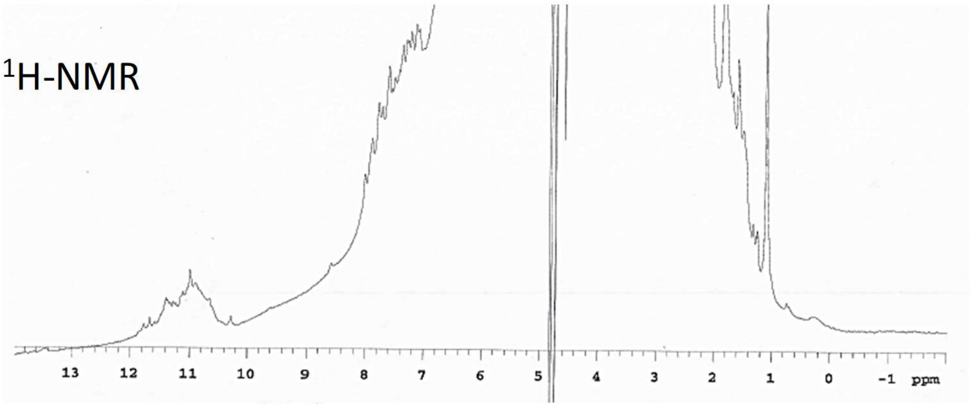
ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*PCNA* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	58.0 ± 0.6

TDS

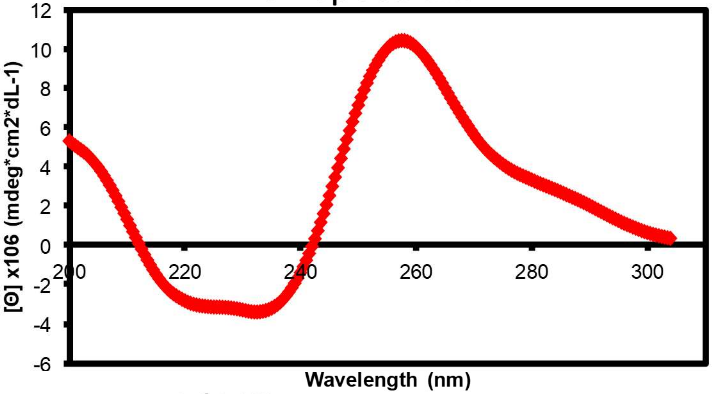


PMS1	Sequence
(coding strand)	5`-CT GGG TGC GGG TGC GGG TGC GGGG TT



^{1}H-NMR

CD Spectrum



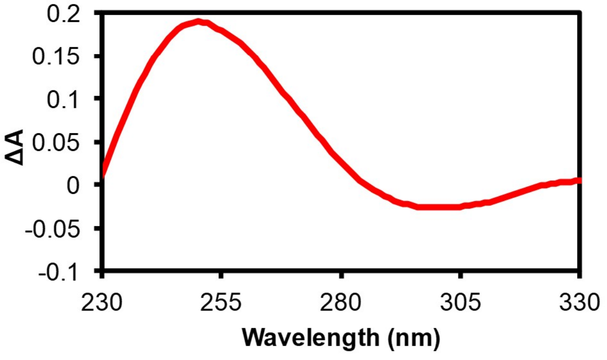
T_m Value

70.1 ± 0.7 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*PMS1* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	59 ± 1

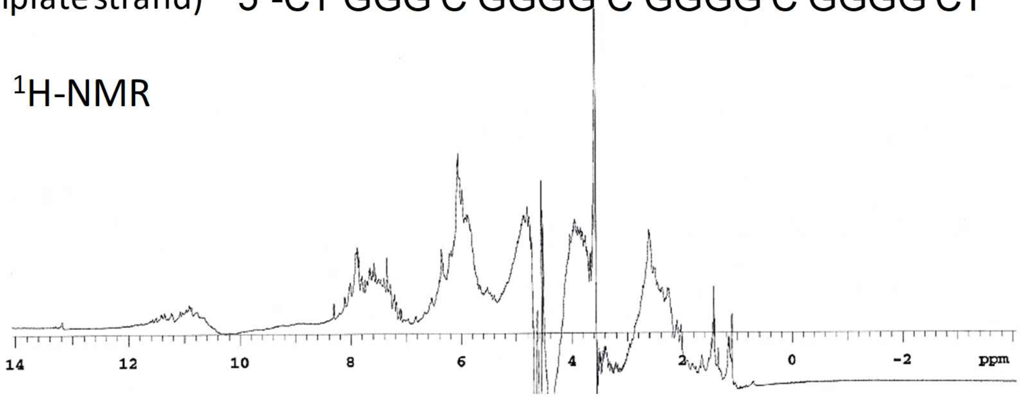
TDS



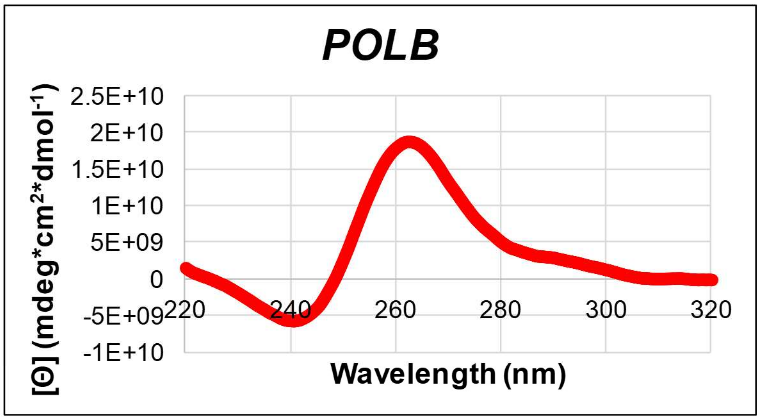
POLB Sequence

(template strand) 5`-CT GGG C GGGG C GGGG C GGGG CT

^{1}H-NMR



CD Spectrum



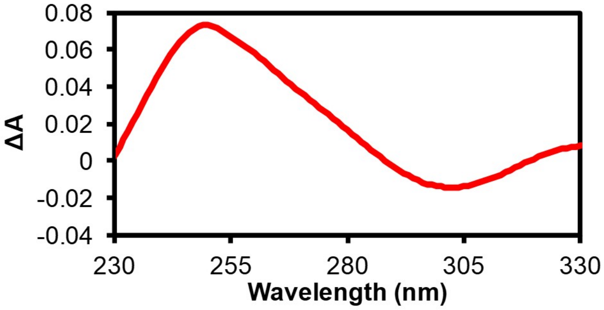
T_m Value

84.7 ± 1.1 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*POLB* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	29 ± 4

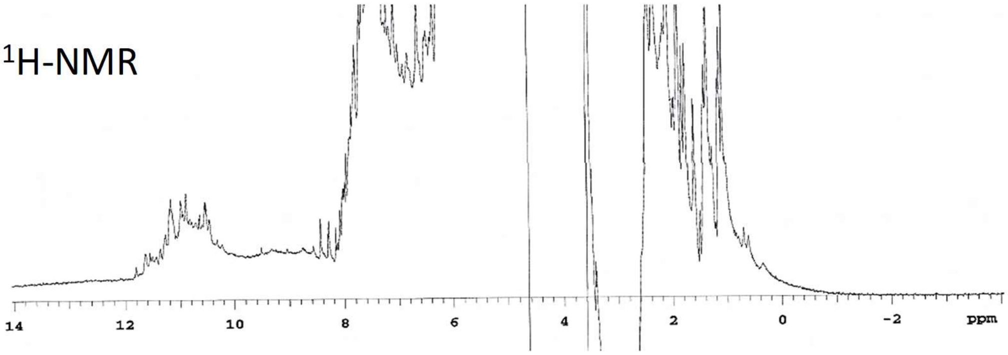
TDS



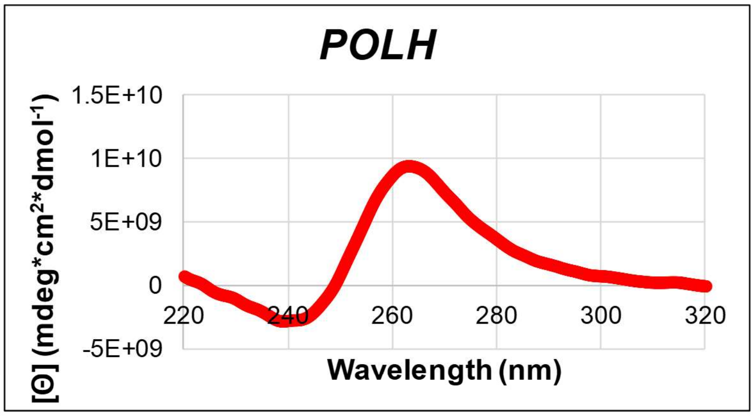
POLH
(coding strand)

Sequence
5`-CT GGGG CT GGG AGA GGG TGTC GGG AC

^{1}H-NMR



CD Spectrum

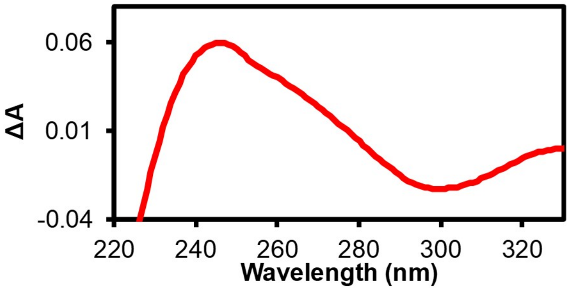


T_m Value
75.6 ± 0.8 °C

ThT Data

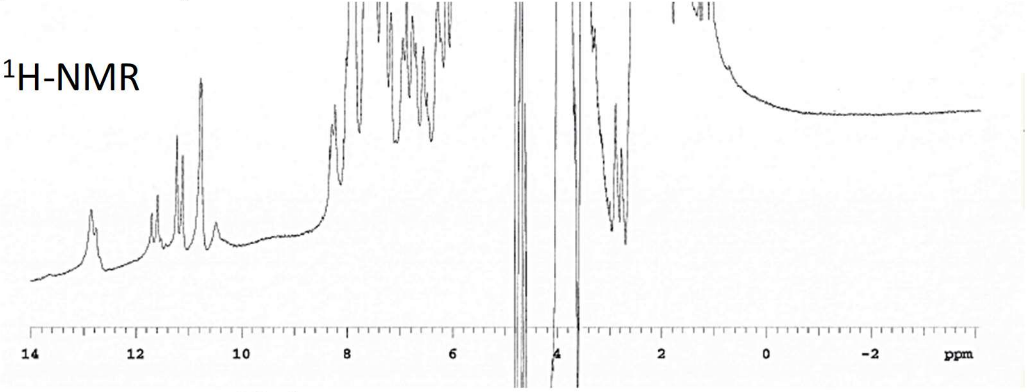
Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*POLH* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	95 ± 10

TDS

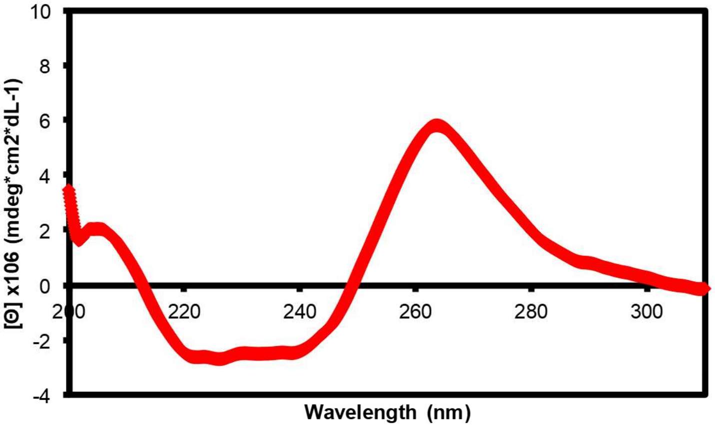


POLL (template strand)	Sequence 5`-CC GGG AA GGG CCTCA GGG CCT GGG TT

^{1}H-NMR



CD Spectrum



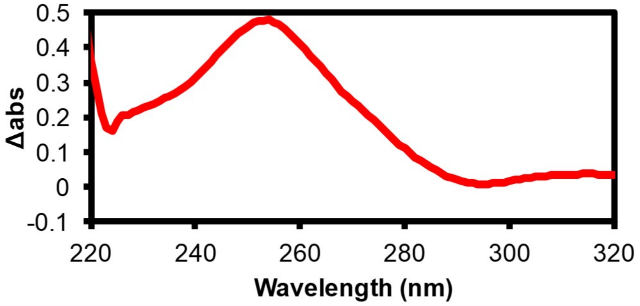
T_m Value

68.2 ± 0.9 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*POLL* G4
λ_{em} at 490 nm	0.09 ± 0.01	0.15 ± 0.01	0.30 ± 0.01	6.1 ± 0.3	46.5 ± 2.0

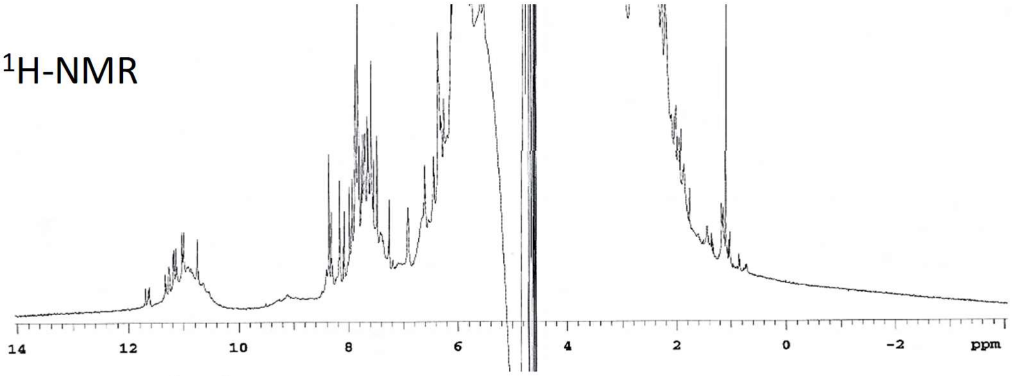
TDS



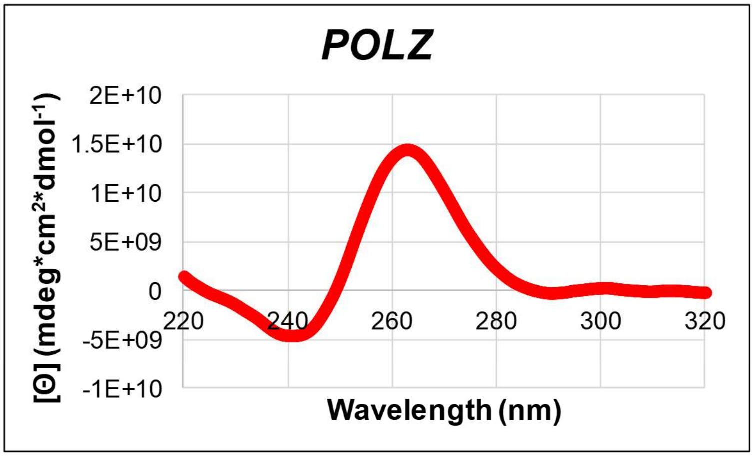
POLZ
(template strand)

Sequence
5`-GA GGG AA GGG C GGG C GGG CG

^{1}H-NMR



CD Spectrum

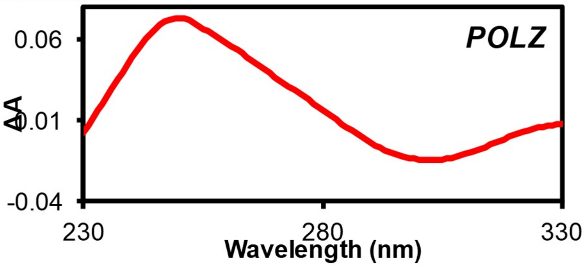


T_m Value
83.8 ± 0.5 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*POLZ* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	59 ± 9

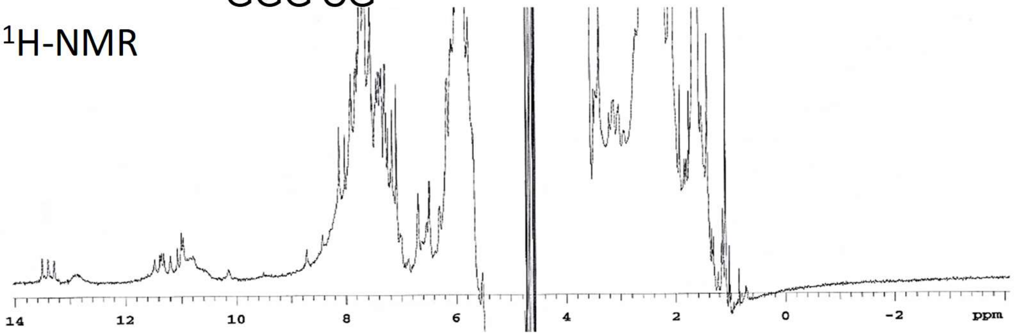
TDS



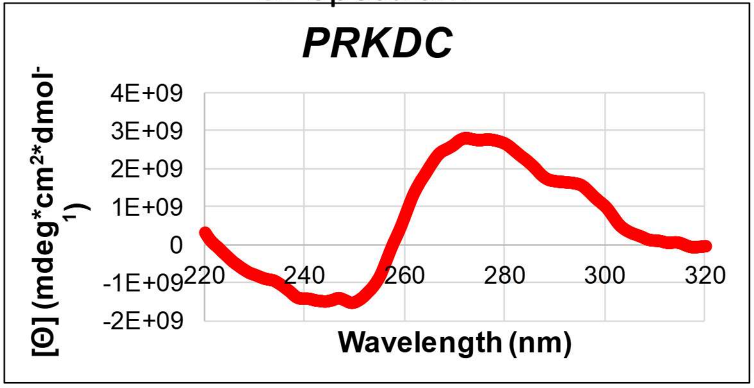
PRKDC
(coding strand)

Sequence
5`-TA GGGG CATTTCC GGG TCC GGG CCGAGC GGG CG

^{1}H-NMR



CD Spectrum



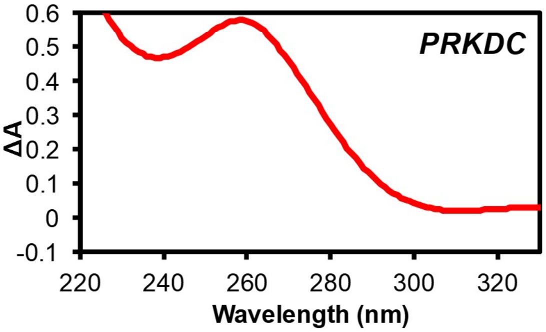
T_m Value

54.1 ± 0.5 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*PRKDC* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	29.0 ± 0.1

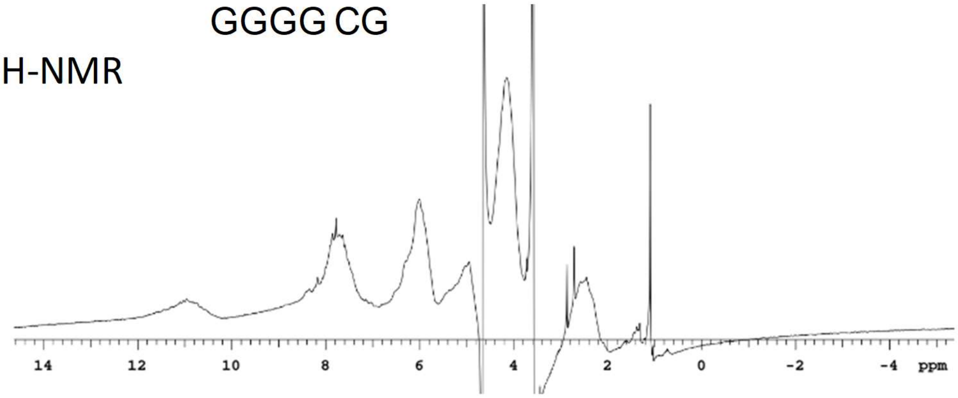
TDS



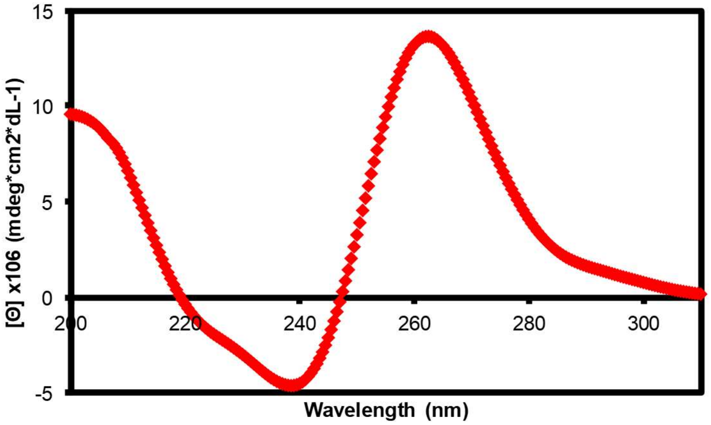
RAD9A
(coding strand)

Sequence
5`-AT GGGG A GGC GGGGG C GGGG CCGGCA GGGG CG

^{1}H-NMR



CD Spectrum

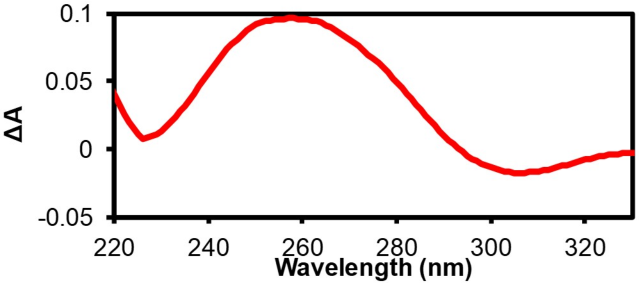


T_m Value
77.6 ± 1.1 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*RAD9* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	55 ± 14

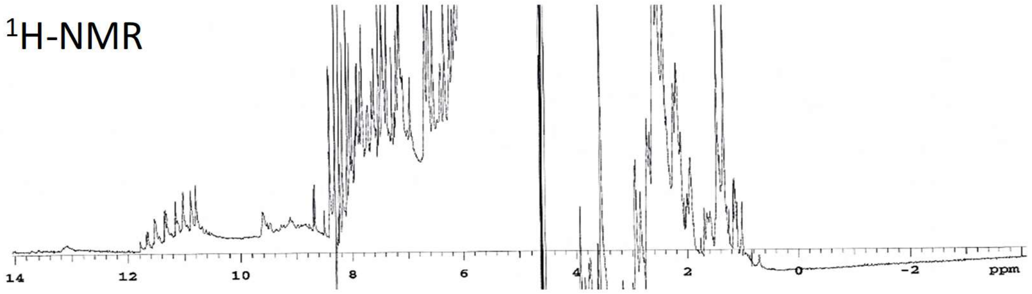
TDS



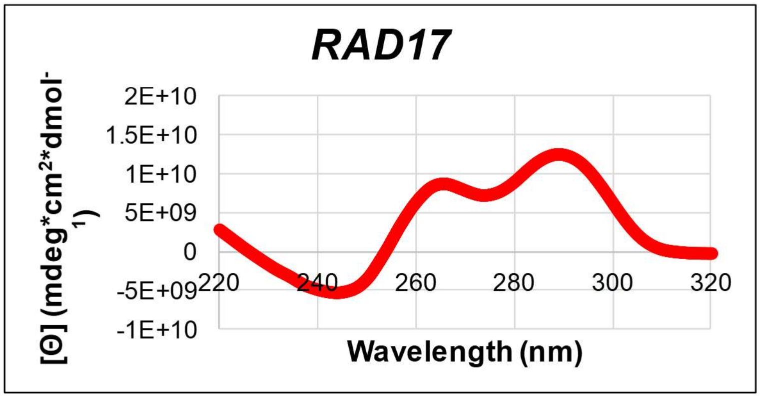
RAD17
(template strand)

Sequence
5`-CC GGG A GGG ACT GGG CT GGGG CA

^{1}H-NMR



CD Spectrum



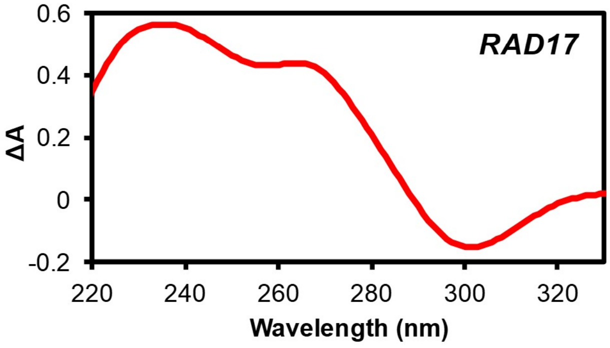
T_m Value

69.5 ± 0.8 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*RAD17* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	20.6 ± 0.3

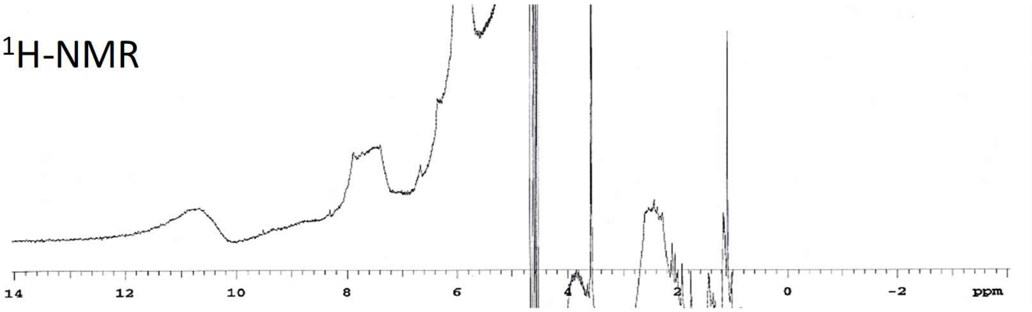
TDS



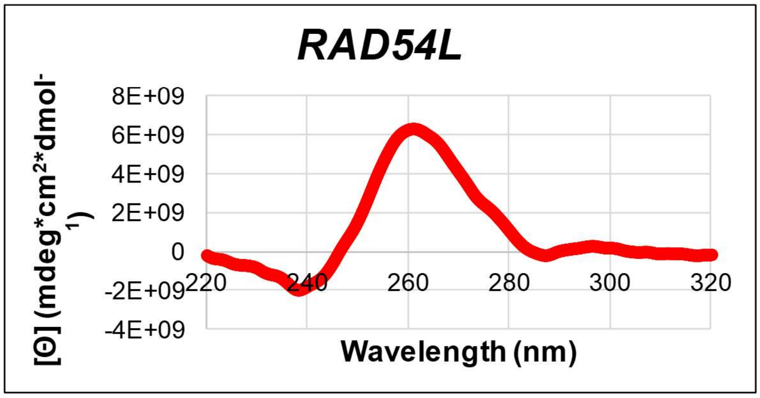
RAD54L
(template strand)

Sequence
5`-GA GGGG C GGGG C GGGGG C GGGGG TG

^{1}H-NMR



CD Spectrum



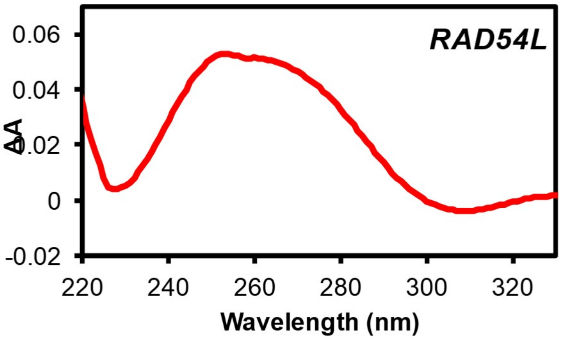
T$_m$ Value

87.3 ± 1.1 °C

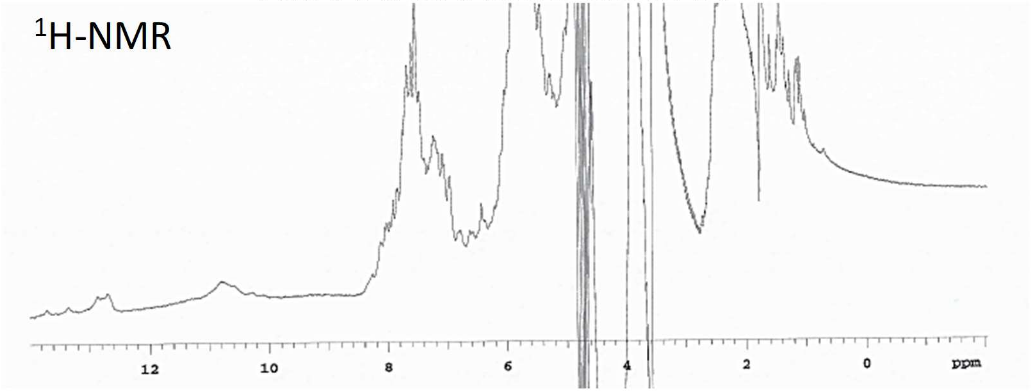
ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*RAD54L* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	10 ± 1

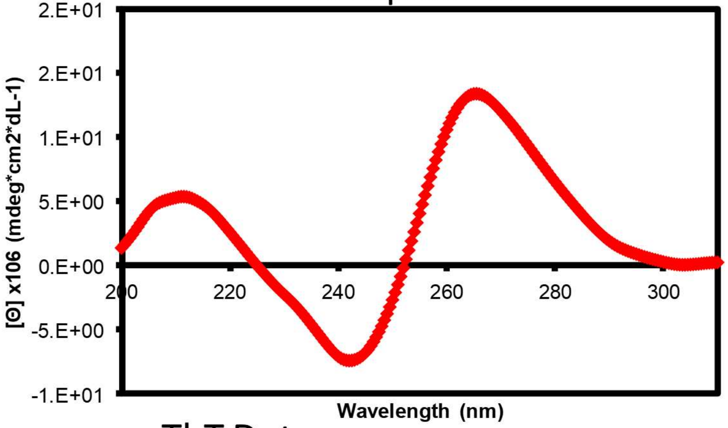
TDS



RECQL (coding strand)	Sequence 5`-CA GGG T GGG AAGCTGAGTC GGG AGAAATGAAGCC GGG AA



CD Spectrum



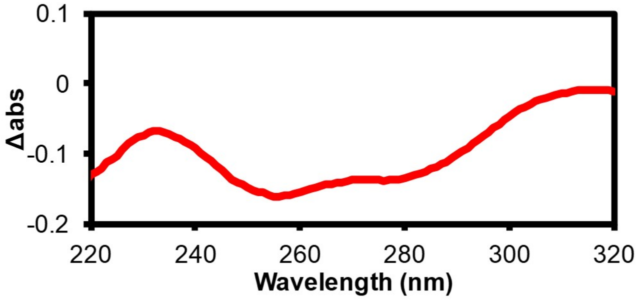
T_m Value

51.7 ± 1.1 °C

ThT Data

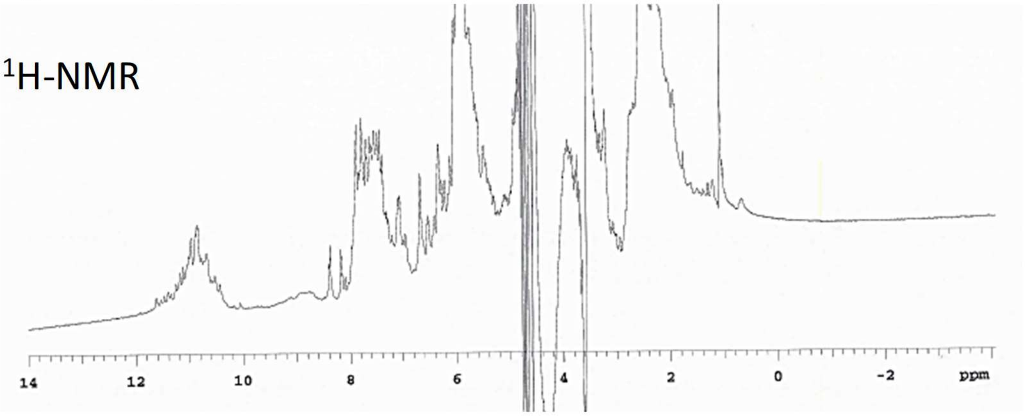
Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*RECQ L* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	.30 ± 0.1	61 ± 3	105 ± 3

TDS

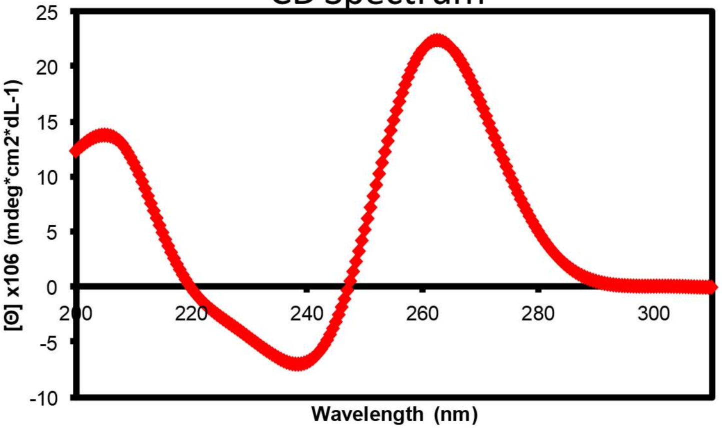


REV1
(coding strand)

Sequence
5`-CA GGG C GGGG CC GGGG A GGGG AG



CD Spectrum



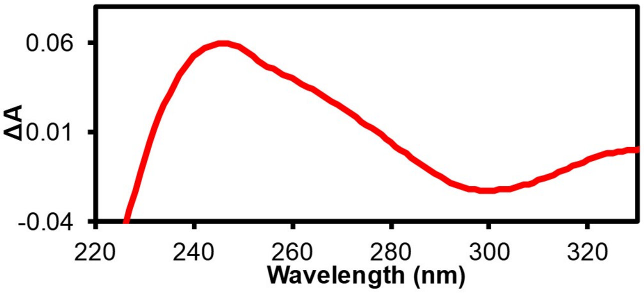
T_m Value

81.6 ± 1.2 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*REV1* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	93 ± 1

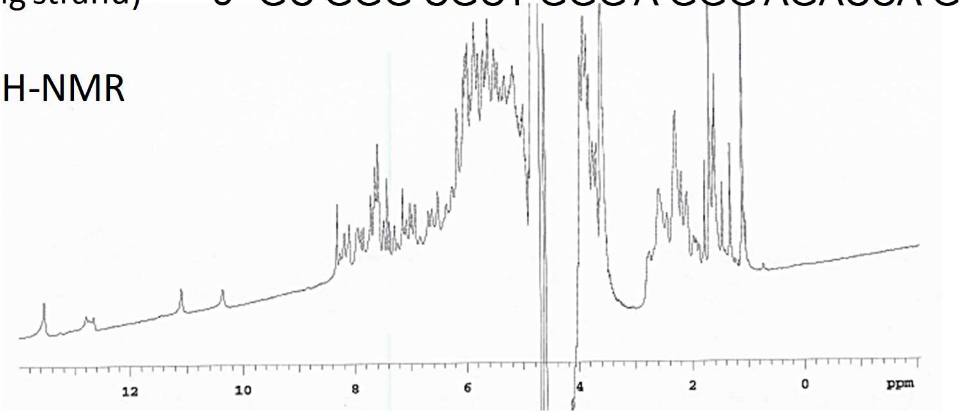
TDS



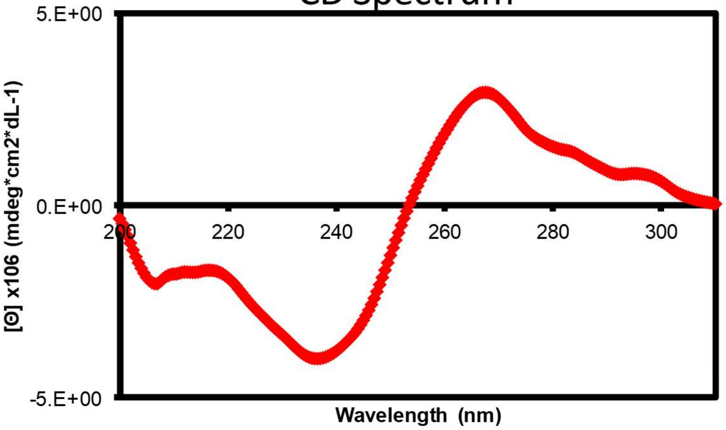
RPA1
(coding strand)

Sequence
5`-GC GGG CGCT GGG A GGG AGACCA GGG CG

^{1}H-NMR



CD Spectrum



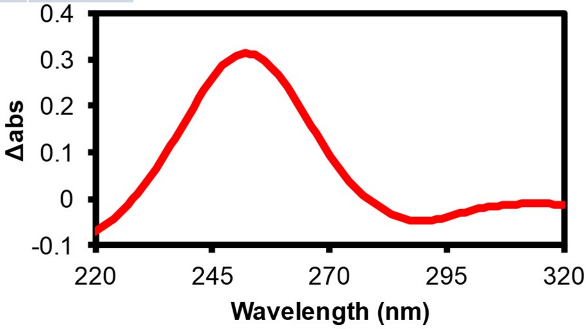
T_m Value

This sequence failed to provide a temperature-dependent transition at 295 nm.

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*RPA1* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	103 ± 6

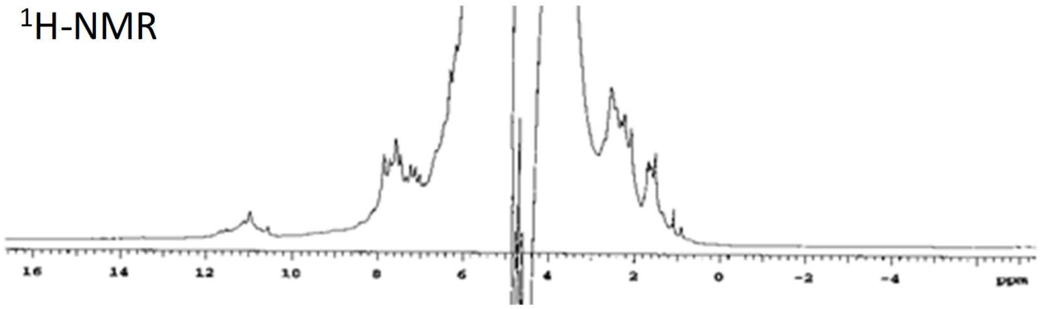
TDS



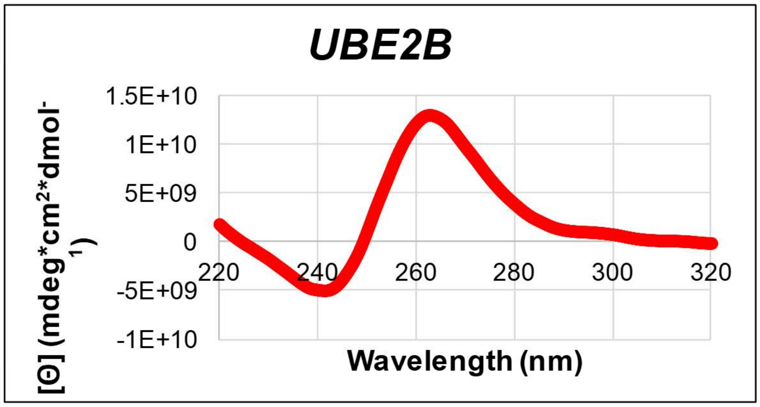
UBE2B (template strand)

Sequence
5`-GC GGG TTTAAGA GGG TGA GGG C GGG TA

^{1}H-NMR



CD Spectrum



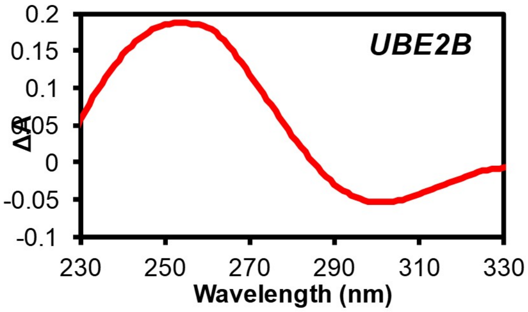
T_m Value

63.2 ± 0.4 °C

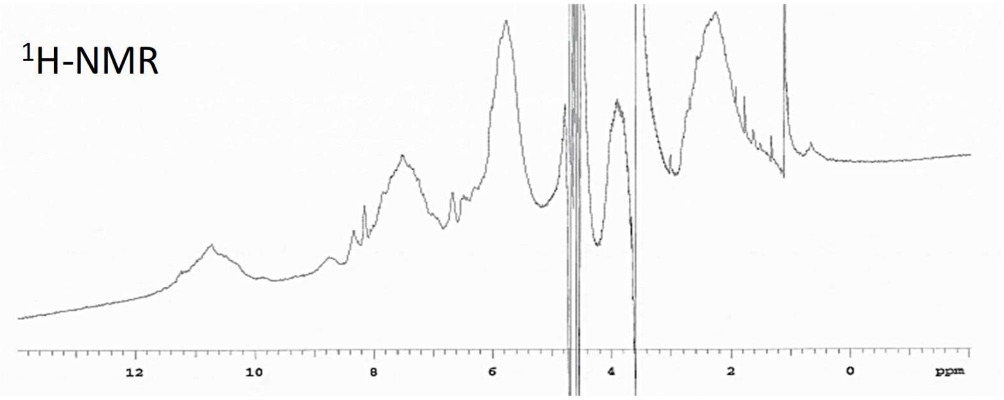
ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*UBE2B* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	97 ± 7

TDS

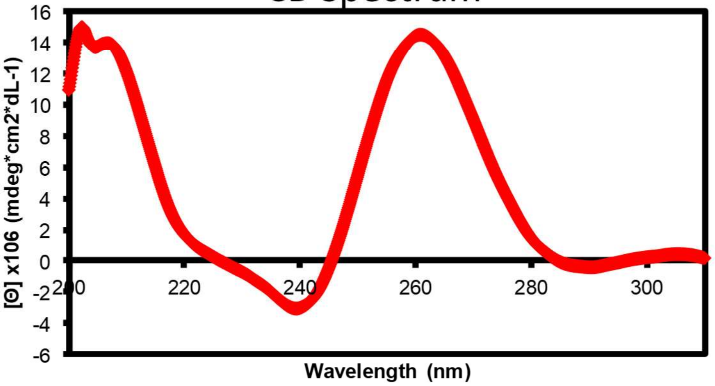


WRN (template strand)	Sequence 5`-GA GGGG A GGG AA GGGG AGGC GGGG AG



^{1}H-NMR

CD Spectrum



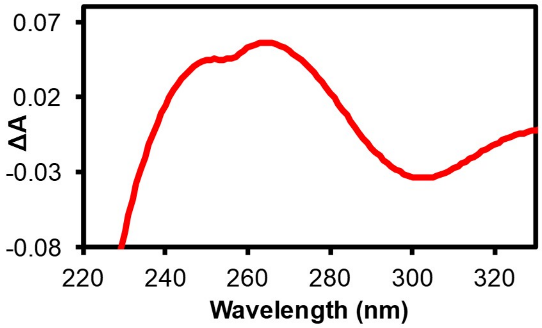
T$_m$ Value

69.4 ± 1.0 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*WRN* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	77 ± 3

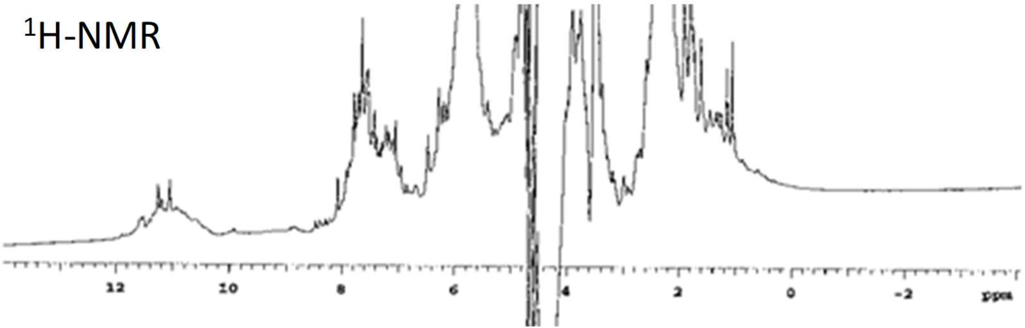
TDS



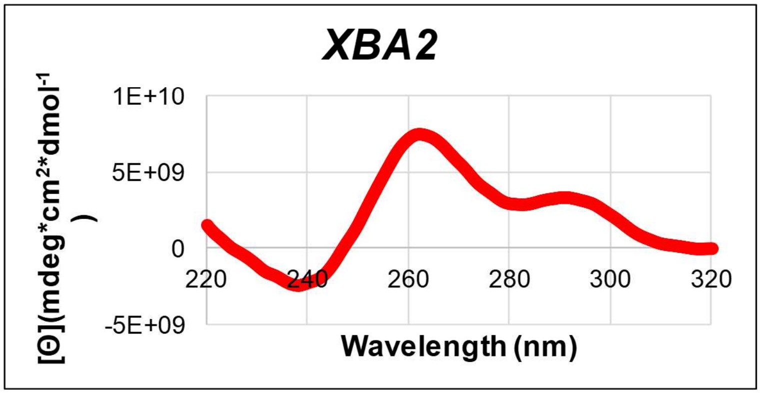
XAB2 (template strand)

Sequence
5`-GT GGG TT GGG AGGCT GGG CA GGGG AT

^{1}H-NMR



CD Spectrum



T_m Value

67.1 ± 1.0 °C

ThT Data

Sequence	ThT	ssDNA	dsDNA	*cMYC* G4	*XAB2* G4
λ_{em} at 490 nm	0.9 ± 0.1	1.5 ± 0.1	3.0 ± 0.1	61 ± 3	83 ± 1

TDS

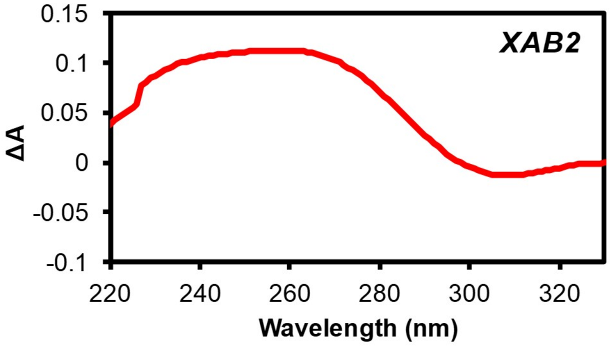
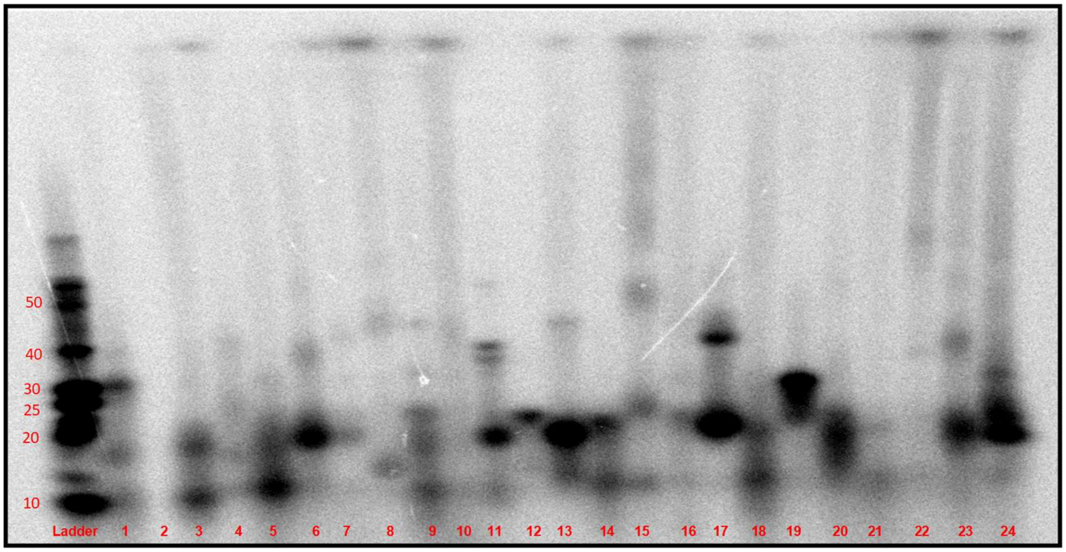
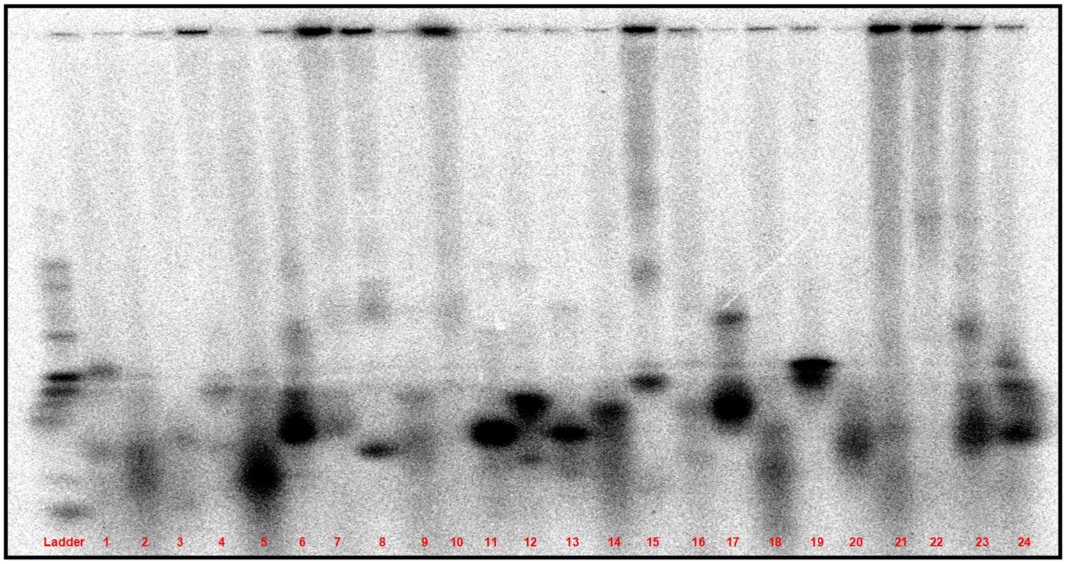


Figure S2. Native PAGE for the 30 PQSs.

Analysis at 20 °C and 10 μM DNA.

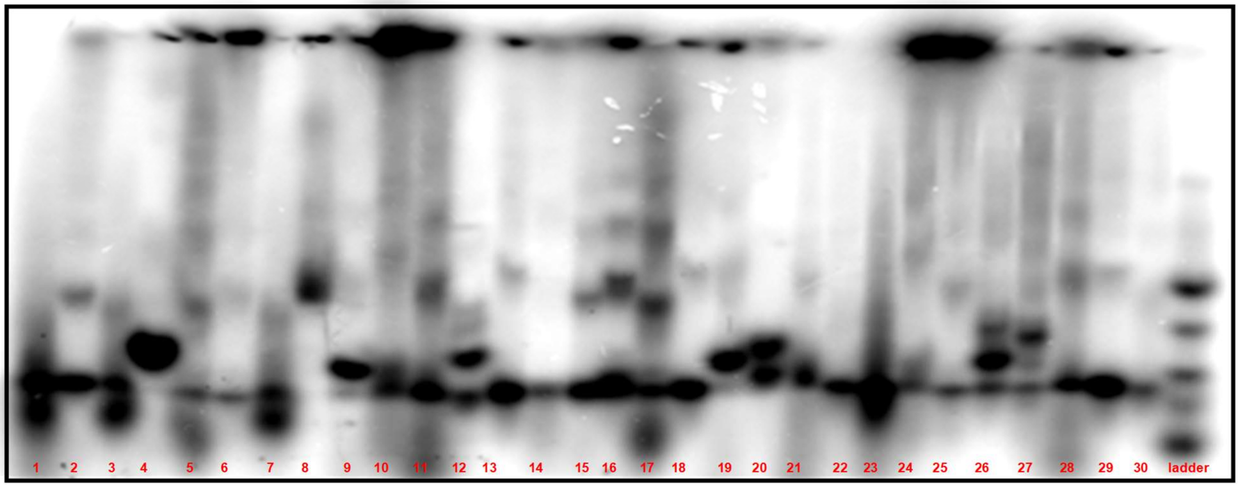


Analysis at 20 °C and 300 μM DNA.



1 = *FAAP24 (31 nts)*, 2 = *ERCC8 (28 nts)*, 4 = *FANCA (34 nts)*, 5 = *FANCC (25 nts)*, 6 = *FEN1 (26 nts)*, 7 = *GTF2H1 (28 nts)*, 8 = *MGMT (30 nts)*, 9 = *NEIL1 (25 nts)*, 10 = *NEIL3 (23 nts)*, 11 = *NTHL1 (25 nts)*, 12 = *PARP3 (23 nts)*, 13 = *PCNA (26 nts)*, 14 = *POLH (26 nts)*, 15 = *POLZ (20 nts)*, 16 = *PRKDC (33 nts)*, 17 = *RAD9A (32 nts)*, 18 = *RAD17 (23 nts)*, 19 = *RAD54L (25 nts),* 20 = *REV1 (23 nts)*, 21 = *SHPRH*, 22 = *UBE2B (27 nts)*, 23 = *WRN (26 nts)*, 24 = *XAB2 (26 nts)*, and ladder = poly-dC$_n$ (n = 10, 20, 25, 30, 40, and 50 nts).

Analysis at 4 °C and 10 µM DNA.



1 = *BLM (20 nts)*, 2 = *FAAP24 (31 nts)*, 3 = *ERCC8 (28 nts)*, 4 = *FANCA (34 nts)*, 5 = *FANCC (25 nts)*, 6 = *FEN1 (26 nts)*, 7 = *GTF2H1 (28 nts)*, 8 = *LIG4 (40 nts)*, 9 = *MGMT (30 nts)*, 10 = *MMS19 (31 nts)*, 11 = *NEIL1 (25 nts)*, 12 = *NEIL3 (23 nts)*, 13 = *NTHL1 (25 nts)*, 14 = *PARP3 (23 nts)*, 15 = *PCNA (26 nts)*, 16 = *PMS1 (26 nts)*, 17 = *POLB (22 nts)*, 18 = *POLH (26 nts)*, 19 = *POLZ (20 nts)*, 20 = *PRKDC (33 nts)*, 21 = *RAD9A (32 nts)*, 22 = *RAD17 (23 nts)*, 23 = *RAD54L (25 nts)*, 24 = *RECQL (39 nts)*, 25 = *REV1 (23 nts)*, 26 = *RPA1 (27 nts)*, 27 = *SHPRH*, 28 = *UBE2B (27 nts)*, 29 = *WRN (26 nts)*, 30 = *XAB2 (26 nts)*, and ladder = poly-dC$_n$ (n = 10, 15, 20, 30, 40 nts).

Note: Some wells have material that did not migrate that cannot be interpreted as unimolecular or multimolecular G-quadruplex folds. The material in the wells dramatically increased when the gel was conducted at 4 °C.

Figure S3. Probability of folding for the 30 PQSs studied.

PQS	**^{1}H-NMR Signals**	**G4-specific CD Signature**	**TDS Profile**	***G4-specific* T_m (295 nm)**	**ThT**	**Probability of Folding**
BLM	+	+ (Parallel G4)	+	+	+	High
ERCC8	+	+ (Parallel G4)	+	+	+	High
FAAP24	+	+ (Parallel G4)	-	+	+	High
FANCA	-	-	-	-	-	Low
FANCC	+	+ (Parallel G4)	+	+	+	High
FEN1	+	+ (Parallel G4)	+	+	+	High
GTF2H1	+	+ (Parallel G4)	+	+	-	High
LIG4	+	-	-	+	-	Low
MGMT	+	+ (Mixed G4)	+	+	+	High
MMS19	+	+ (Mixed G4)	+	+	+	High
NEIL1	+	+ (Parallel G4)	+	+	+	High
NEIL3	+	+ (Mixed G4)	+	+	+	High
NTHL1	+	+ (Mixed G4)	+	+	+	High
PARP3	+	+ (Mixed G4)	+	+	+	High
PCNA	+	+ (Parallel G4)	+	+	+	High
PMS1	+	+ (Parallel G4)	+	+	+	High
POLB	+	+ (Parallel G4)	+	+	+	High
POLH	+	+ (Parallel G4)	+	+	+	High
POLL	+	+ (Parallel G4)	+	+	+	High
POLZ	+	+ (Parallel G4)	+	+	+	High
PRKDC	+	-	-	+	+	Medium
RAD9A	+	+ (Parallel G4)	+	+	+	High
RAD17	+	+ (Hybrid G4)	+	+	+	High
RAD54L	+	+ (Parallel G4)	+	+	-	High
RECQL	+	-	-	+	-	Low
REV1	+	+ (Parallel G4)	+	+	+	High
RPA1	-	-	+	-	-	Low
UBE2B	+	+ (Parallel G4)	+	+	+	High
WRN	+	+ (Parallel G4)	+	+	+	High
XAB2	+	+ (Mixed G4)	+	+	+	High

References

1. Allgayer, J., Kitsera, N., Bartelt, S., Epe, B., and Khobta, A. (2016) Widespread transcriptional gene inactivation initiated by a repair intermediate of 8-oxoguanine. *Nucleic Acids Res. 44*, 7267-7280.
2. You, C., Dai, X., Yuan, B., Wang, J., Brooks, P. J., Niedernhofer, L. J., and Wang, Y. (2012) A quantitative assay for assessing the effects of DNA lesions on transcription. *Nat. Chem. Biol. 8*, 817-822.
3. Fleming, A. M., Zhou, J., Wallace, S. S., and Burrows, C. J. (2015) A role for the fifth G-track in G-quadruplex forming oncogene promoter sequences during oxidative stress: Do these "spare tires" have an evolved function? *ACS Cent. Sci. 1*, 226-233.
4. Riedl, J., Fleming, A. M., and Burrows, C. J. (2015) Sequencing of DNA lesions facilitated by site-specific excision via base excision repair DNA glycosylases yielding ligatable gaps. *J. Am. Chem. Soc. 138*, 491-494.