SUPPLEMENTARY MATERIALS AND METHODS

Preparation of DNA substrates

Figure S1 illustrates the general strategy for the preparation of DNA substrates, similar to those previously used (1,2). A fragment containing the T7 promoter was obtained by EcoRI and BsgI restriction digest of pG4 plasmid (3). The restriction reaction was performed in a total volume of 50 μl with 55 μg of plasmid, 1 μl of 1x EcoRI buffer, 1x BSA, 30 units of EcoRI and 30 units of BsgI. The reaction was incubated at 37°C for 4 hours followed by addition of EDTA to a final concentration of 20 mM to stop the reaction. The restriction products were separated by agarose gel-electrophoresis, and the promoter-containing fragment was isolated using the QIAquick gel extraction protocol.

Single stranded oligonucleotides with the sequence of interest were purchased from Integrated DNA Technologies (IDT). The respective oligonucleotides were annealed together in a total volume of 100 μl containing 2 μM of each corresponding single stranded oligonucleotide, 10 mM TrisHCl (pH 7.9), and 10 mM $MgCl_2$. The oligonucleotides were annealed by placing the mixture in about 1 L of boiling water and allowing it to cool to room temperature.

The annealed oligonucleotides and the gel-purified promoter fragment were ligated together in reactions containing 8.5 μl of annealed oligonucleotide mixture, 85 ng of pure promoter fragment, 2 μl 10x ligase buffer, 1 μl of T4 DNA ligase (2000 units) and water to a total volume of 20 μl. The ligation reactions were incubated overnight at 16°C and stopped by heating at 65°C for 30 minutes to inactivate the DNA ligase. The resulting constructs were analyzed in a 1.5% agarose gel by running at 60 V for 4 hours, and visualized by staining with ethidium bromide and examination under UV light.

Figure S1 shows the complete ligation product plus possible products of incomplete ligation. Because the single-stranded oligonucleotides were not phosphorylated, a nick remains in the non-template strand. Previously in was shown that nick in this construct doesn't produce significant effect upon the transcription pattern (1). We chose to leave this nick non-ligated to avoid potential heterogeneity of the product dew to incomplete ligation. Note that the sticky ends produced by BsgI digest in this plasmid are not self-complementary, so the promoter fragments can ligate to each other only from the EcoRI side. From the scheme at the bottom of FigS1, it is seen that there are four possible ligation products, but they produce only two different run-off transcripts; one corresponds to the complete construct, and the other to the promoter fragment alone.

In vitro T7 transcription assays

The in vitro T7 transcription reaction was performed for 30 min at 37°C in 12 μl containing ~30 fmoles of DNA substrate, 10 μCi (α-^{32}P)CTP (which corresponds to the final concentration of about 0.0003 mM), 20 units of T7 RNA polymerase and 16 units of RNasin (both from Promega corp, Madison, WI) 0.17 mM each of ATP, GTP and UTP, and 0.017 mM CTP. Compositions of various transcription buffers are described in the main text. The reaction was stopped by adding 94 μl of buffer containing 1% SDS, 106 mM TrisHCl (pH 7.6), 13.2 mM EDTA, 160 mM NaCl, 25 μg tRNA (Invitrogen) and 10

μg Proteinase K (Invitrogen) and incubated at room temperature for 15 min. Next, 11 μl of 3 M NaOAc (pH 5.27) and 300 μl of 100 % ethanol (cooled to -20°C) were added, the mixture was incubated on dry ice for ~30 min and then centrifuged for 20 min at 14,000 rpm at 4°C. The supernatant was removed; then 500 μl of 75% ethanol (cooled to -20°C) was added, and the mixture was centrifuged for 5 min under the same conditions. The supernatant was removed, and the pellets were dried with a SpeedVac for 10 min and dissolved in 8 μl of the formamide loading solution (94 % formamide, 2 mM EDTA, 0.05 % bromophenol blue, 0.05 % xylene cyanol). Then, 2 μl of the resulting solution were mixed with 2 μl of formamide loading solution, incubated at 85°C for ~2 min, quickly chilled on a pre-cooled rack, and loaded on a 5% sequencing gel (acrylamide:bis-acrylamide 29:1) containing 8 M urea and run at ~70 V/cm. As size markers, 5'-end labeled denatured DNA ladders consisting of DNA fragments of sizes increasing in steps of 100 or 10 bases were used. Then, the gel was dried and exposed to a phosphoimager screen.

Restriction Protection Assay:
The restriction protection assay was conducted by radiolabeling the 5' end of the top strand of the duplexes and the triplexes designed for the assay. The radiolabeled substrate was then annealed with the bottom strand and restricted with BsgI in 20 μl of reaction mixture containing 2 ng of the substrates, 1 μl of the BsgI (50 units), and 1x NEB Buffer4 (10 mM MgAc, 50 mM KAc, 20 mM TrisAc (pH 7.9), 1 mM DTT supplemented with 80 μM of S-adenosylmethionine (SAM) (NEB) plus extra 2.5 mM $MgCl_2$ and 2.5 mM TrisCl (pH 7.9) from the annealing buffer for 30 minutes. The restriction reaction was stopped by adding EDTA up to 25 mM. The resulting products were run on a 15% non-denaturing acrylamide gel to observe restriction by BsgI.

Quantitation of Data:
Intensities of the bands were quantified using the ImageLab software from BioRad. Transcription assays were performed in the presence of all four NTPs with a small amount of (α-^{32}P) CTP to obtain uniform labeling over their lengths. Thus, the intensity of each band was proportional to the number of cytosines in each molecule of the transcript. The molar percentage of truncated blockage products P_t was estimated by using the equation:

$$P_t = \frac{\frac{I_t}{N_t}}{\frac{I_t}{N_t} + \frac{I_r}{N_r}} \times 100\%$$

where I_t and I_r are intensities, and N_r and N_t are the numbers of cytosines in the truncated transcripts and the runoff product, respectively.

The intensity of each product in the image was measured by drawing a rectangle closely around a band of interest. Same-sized rectangles were drawn above and below the band of interest and the signals from the two rectangles were averaged to calculate the

background. This local background signal was subtracted from the signal obtained from the band of interest.

DNA melting experiments

Absorbance *versus* temperature profiles of oligonucleotide samples were measured in a 10 mm (300 ml) quartz microcuvette (Hellma) and a 1 mm path-length quartz cuvette (Starna GmbH) using a Hitachi U-2900 UV/Visible Spectrophotometer equipped with a thermoelectric temperature controller. Thermally-induced unfolding transitions were monitored between 40 and 90°C at 260 nm at the heating rate of 0.5°C /min. The spectra were smoothed in Origin 8.0 using a Savitzky-Golay filter.

To allow secondary structure formation, the starting samples dissolved in the buffers were first transformed into the single-stranded form by heating at 95°C for 5 min, cooled down fast to 4°C at the cooling rate of 5°C/min and then used in the UV experiments. Oligonucleotides were diluted in a buffer so that the A260 was 0.5–0.7. Absorption was normalized by the values at 50°C.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Preparation of DNA substrates. The pG4 plasmid was cut with restriction enzymes EcoRI and BsgI to produce a fragment containing T7 RNAP promoter. The promoter fragment was gel purified. Oligonucleotides were designed such that when annealed they produced an intramolecular triplex and its variants. Annealed oligonucleotides were ligated to the promoter fragment.

Figure S2. Restriction analysis of substrates.

Radioactively labeled DNA substrates were treated by BsgI and analyzed by gel-electrophoresis as described in Supplementary Materials and Methods. DNA substrates used for the analysis are shown at the top of the figure. For the mismatched triplex and the triplex-bulge substrates, only the part that is different from the "perfect" triplex is shown. Color-coding for triplex-forming area is the same as that in Fig.2 in the main text. Recognition site for BsgI enzyme is shown in bold, and position of cleavages created by this enzyme are shown by small triangles. Black asterisk symbolizes radioactive labeling of the 5' end of the top strand, which allows monitoring the non-cleaved substrate and the cleaved product 1 by gel-separation followed by autoradiography. Below the authoradiograph, quantitation of two independent experiments is shown. It is seen that the degree of cleavage for the duplex and the triplex-bulge substrates are virtually the same (around 90%; i.e. only about 10% remains uncleaved), while for the triplex, practically complete (around 99%) protection from the cleavage is observed. For the mismatched triplex, about three-fold more uncleaved product (comparing to the duplex and the triplex-bulge) is observed, indicating partial protection from cleavage that could be explained by the mismatched triplex formation.

Figure S3. Transcription assay under the same metal ions concentration as in the restriction analysis.
Transcription was performed as in the standard conditions, except transcription buffer was replaced by 1x NEB Buffer4 (10 mM MgAc, 50 mM KAc, 20 mM TrisAc (pH 7.9), 1 mM DTT), supplemented with additional $MgCl_2$ to make concentration of Mg iones the same as in restriction assay (i.e. 12.5 mM total). To test possible effect of BsgI site upon the pattern of transcription blockage, the versions of duplex, triplex and triplex-bulge containing BsgI site in the same position as substrates for the restriction assay (see the scheme below the figure; for BsgI substrates, only the downstream flank that contains BsgI site is shown) were also tested. It is seen (see also quantitation in Fig.4S), that the results for respective substrates with and without BsgI site are practically the same). Similar to other conditions (see the main text), triplex-forming substrate produces three well pronounced blockage bands (dn, m, up). In the duplex substrate, a much weaker m-band is seen, while in the triplex-bulge substrate, significantly weaker up-band is seen. In the mismatched triplex substrates, both up-and m-bands can be seen, the latter, though much weaker than in the perfect triplex, is about twice stronger than in the duplex. That might indicate some contribution of the mismatched triplex (which formation is consistent with partial protection from BsgI digest, Fig.2S) to the blockage. In addition to well-pronounced blockage bands, there is a diffusive area of minor blockage signals further downstream (designated as fdn, and mapped as a gray oval in the scheme below) that is more pronounced in substrates containing single-stranded regions comparing to the duplex. For example, the intensity of this area normalized upon run-off product was about two-fold larger for triplex-bulge substrate (both for usual and BsgI-containing substrates) than for the respective duplex substrates. As it is discussed in the main text, that might indicate R-loop formation.

Figure S4. Quantitation of the results from Fig.S3.

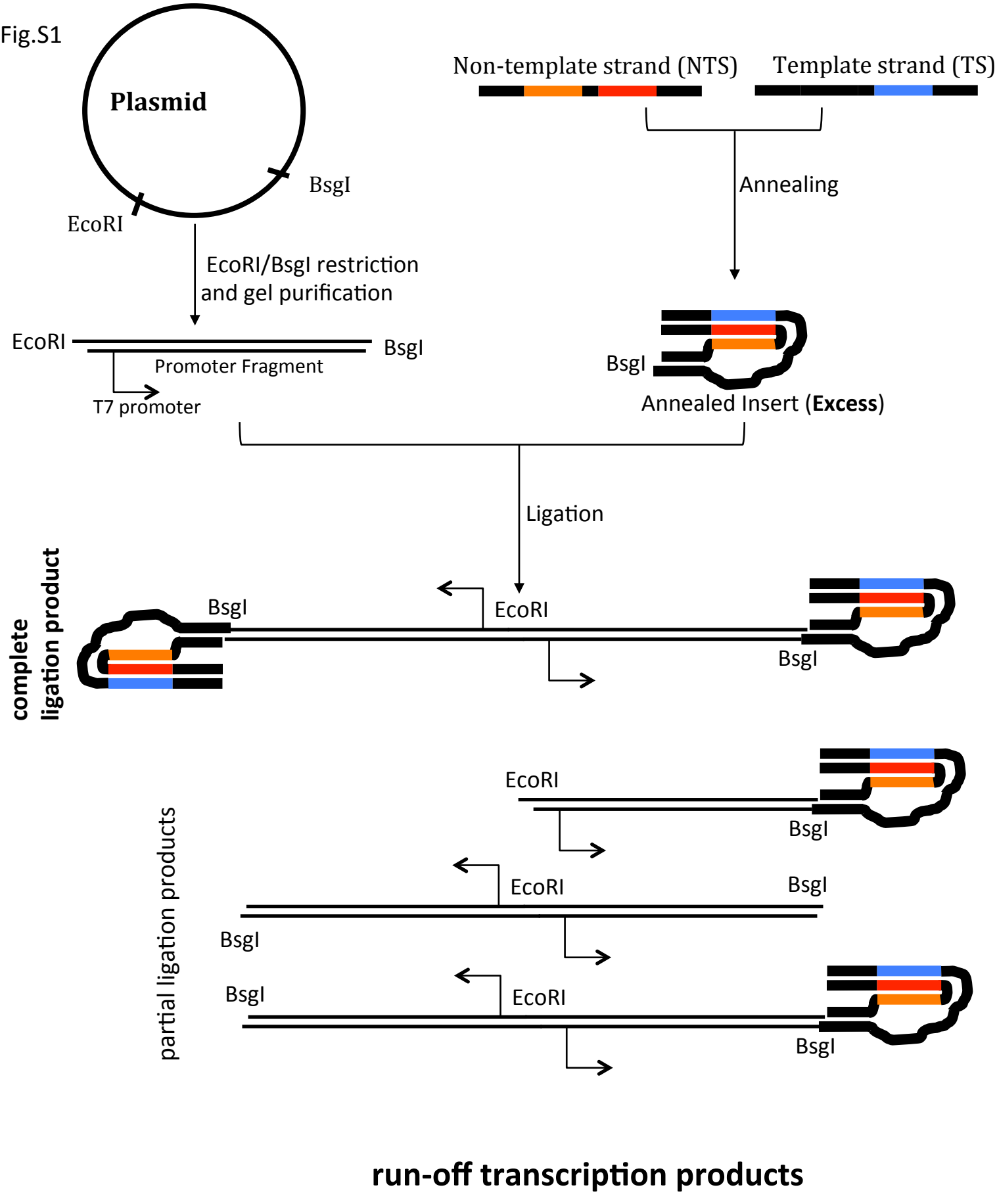
Figure S5. Melting experiments for duplex and triplex under various conditions
At the top of the Figure, duplex (hairpin)- and triplex-forming substrates used in the melting experiments are shown. All buffers contain 4 mM $MgCl_2$, 37.8 mM TrisHCl (pH 7.9), and 1.7mM spermidine. In addition, "standard" buffer contains 8.3 mM NaCl, "K+" buffer contains 80 mM KCl, and "Li+" buffer contains 80 mM LiCl.

SUPPLEMENTARY REFERENCES

1. Salinas-Rios, V., Belotserkovskii, B.P. and Hanawalt, P.C. (2011) DNA slip-outs cause RNA polymerase II arrest in vitro: potential implications for genetic instability. *Nucleic Acids Res*, **39**, 7444-7454.

2. Neil, A.J., Belotserkovskii, B.P. and Hanawalt, P.C. (2012) Transcription blockage by bulky end termini at single-strand breaks in the DNA template: differential effects of 5' and 3' adducts. *Biochemistry*, **51**, 8964-8970.
3. Belotserkovskii, B.P., Liu, R., Tornaletti, S., Krasilnikova, M.M., Mirkin, S.M. and Hanawalt, P.C. (2010) Mechanisms and implications of transcription blockage by guanine-rich DNA sequences. *Proc Natl Acad Sci U S A*, **107**, 12816-12821.

Fig.S1



run-off transcription products

run-off from the promoter fragment

full size run-off

Fig.S2

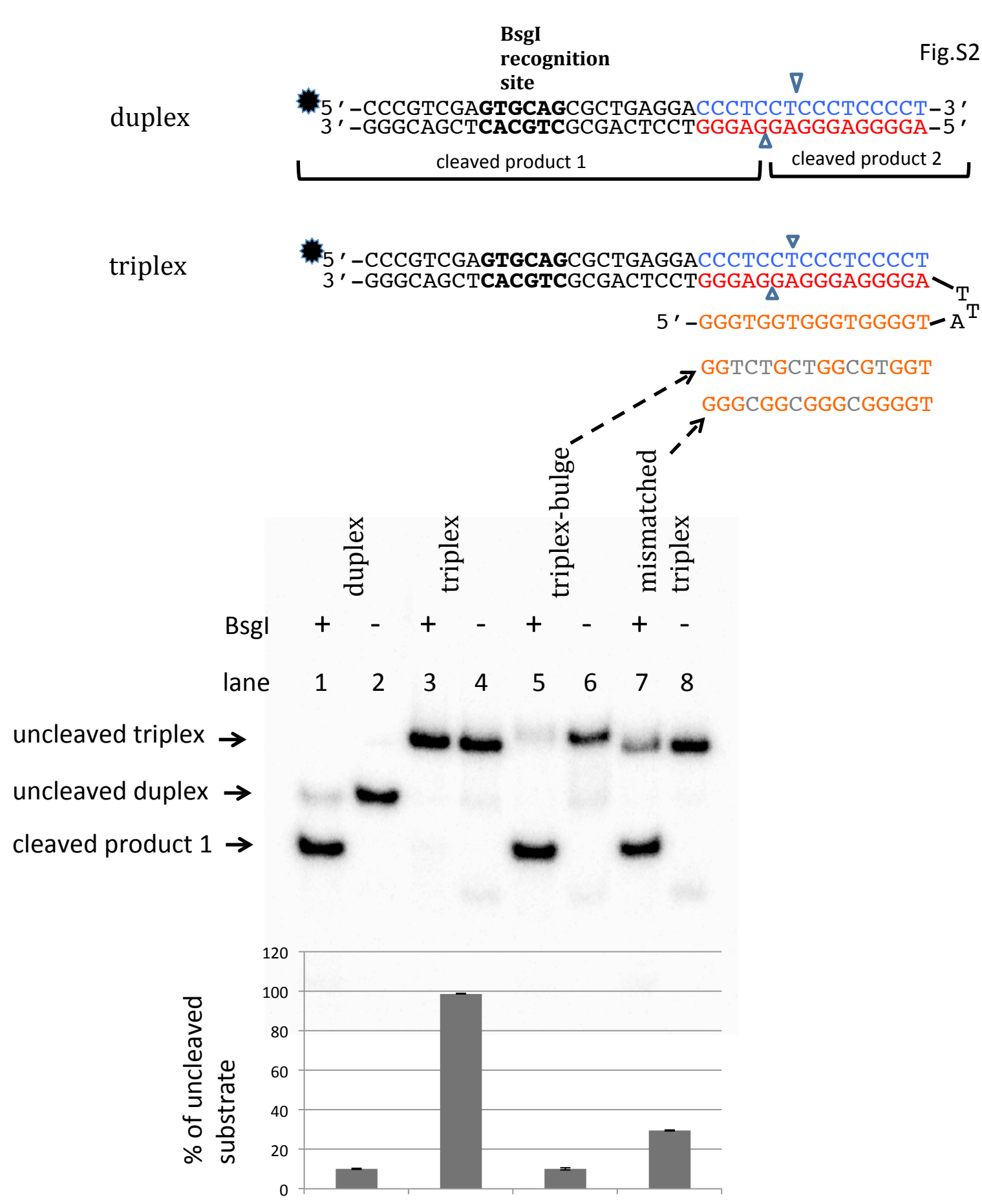
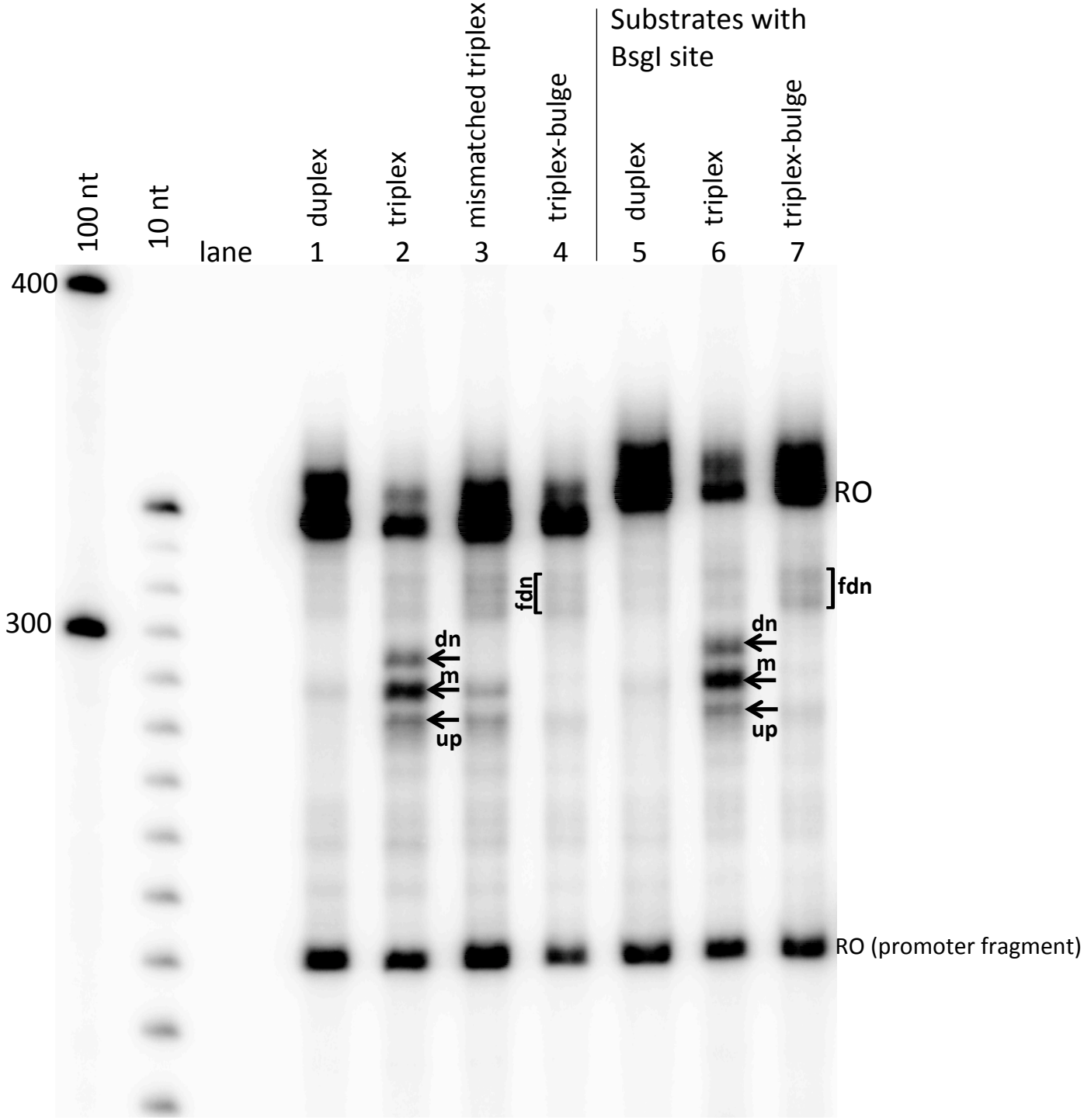

BsgI recognition site
duplex
5′-CCCGTCGAGTGCAGCGCTGAGGACCCTCCTCCCTCCCCT-3′
3′-GGGCAGCTCACGTCGCGACTCCTGGGAGGAGGGAGGGGA-5′
cleaved product 1
cleaved product 2
triplex
5′-CCCGTCGAGTGCAGCGCTGAGGACCCTCCTCCCTCCCCT
3′-GGGCAGCTCACGTCGCGACTCCTGGGAGGAGGGAGGGGA
T
5′-GGGTGGTGGGTGGGGT-A
T
GGTCTGCTGGCGTGGT
GGGCGGCGGGCGGGGT
triplex-bulge
mismatched triplex
duplex
triplex
BsgI + - + - + - + -
lane 1 2 3 4 5 6 7 8
uncleaved triplex
uncleaved duplex
cleaved product 1
% of uncleaved substrate
120
100
80
60
40
20
0

Fig.S3


100 nt
10 nt
lane
duplex
triplex
mismatched triplex
triplex-bulge
Substrates with BsgI site
duplex
triplex
triplex-bulge
1
2
3
4
5
6
7
400
300
RO
fdn
fdn
dn
m
up
dn
m
up
RO (promoter fragment)

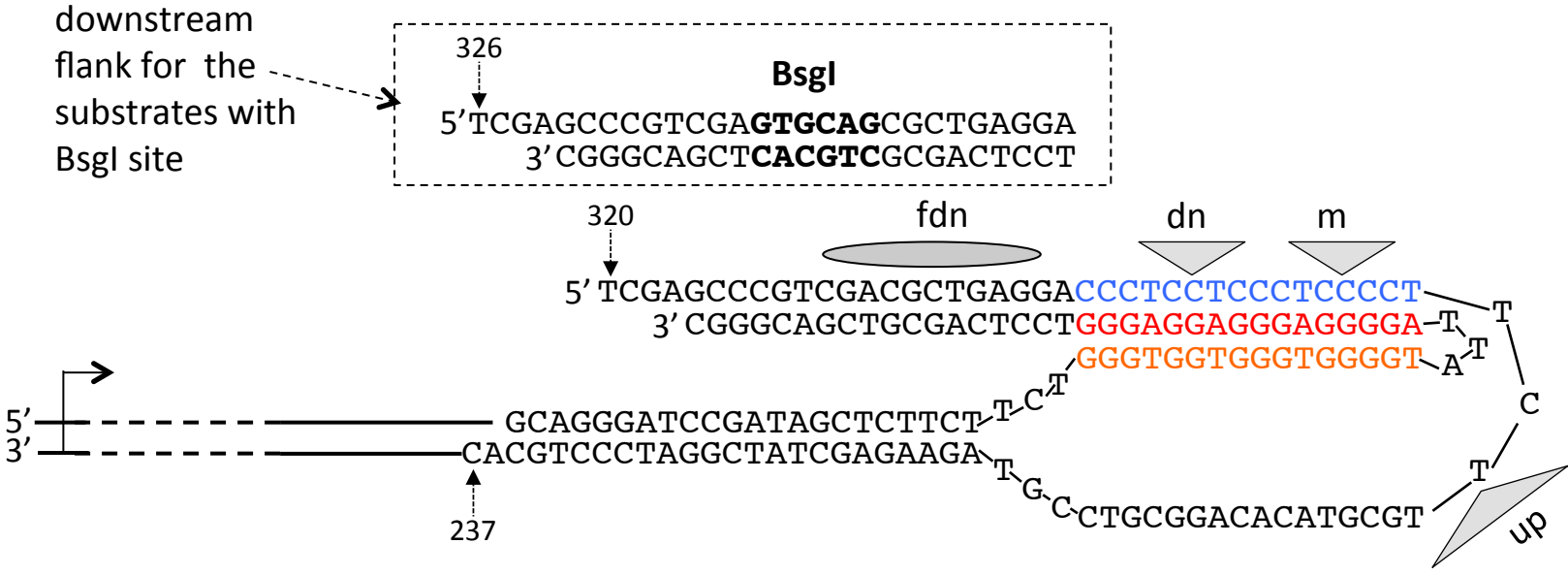

downstream flank for the substrates with BsgI site
326
BsgI
5'TCGAGCCCGTCGAGTGCAGCGCTGAGGA
3'CGGGCAGCTCACGTCGCGACTCCT
320
fdn
dn
m
5' TCGAGCCCGTCGACGCTGAGGACCCTCCTCCCTCCCCT
3' CGGGCAGCTGCGACTCCTGGGAGGAGGGAGGGGA
GGGTGGTGGGTGGGGT
5'
3'
GCAGGGATCCGATAGCTCTTCT
CACGTCCCTAGGCTATCGAGAAGA
CTGCGGACACATGCGT
237
up

Fig.S4

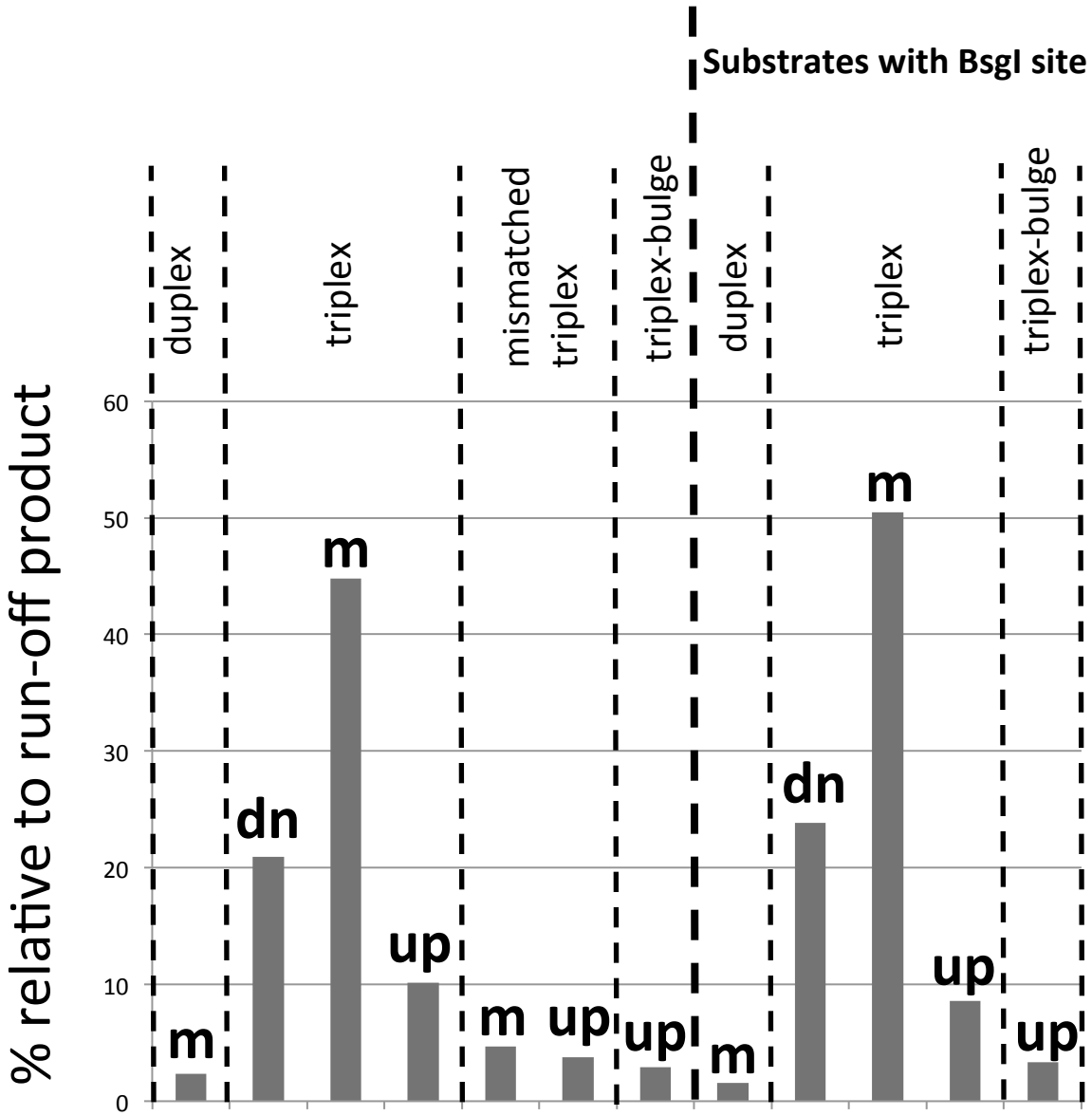
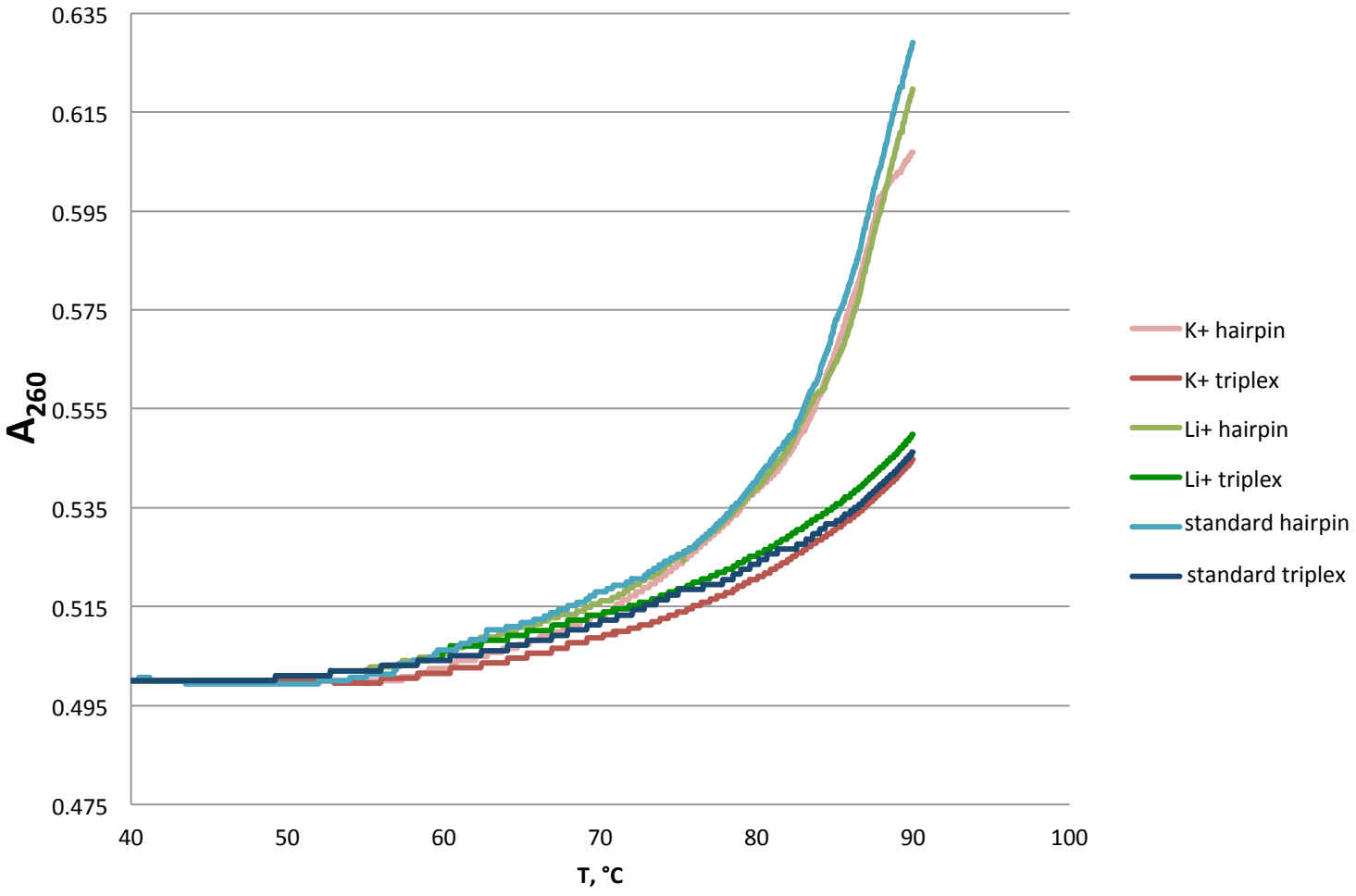

Substrates with BsgI site
duplex
triplex
mismatched triplex
triplex-bulge
duplex
triplex
triplex-bulge
% relative to run-off product
0
10
20
30
40
50
60
m
dn
m
up
m
up
up
m
dn
m
up
up

Fig.S5


hairpin
T–T
CCCTCCTCCCTCCCCT 3’
A
A–GGGAGGAGGGAGGGGA 5’
triplex
T–T
CCCTCCTCCCTCCCCT 3’
A
A–GGGAGGAGGGAGGGGA–T
5’GGGTGGTGGGTGGGGT
A
T


0.635
0.615
0.595
0.575
0.555
0.535
0.515
0.495
0.475
A260
40
50
60
70
80
90
100
T, °C
K+ hairpin
K+ triplex
Li+ hairpin
Li+ triplex
standard hairpin
standard triplex