Features of the Influence of a DNA Sequence on Its Adjacent Sequence

Lijuan Long, Xinxin Li, Hailang Wei, and Wei Li*

ABSTRACT: To explore the features of the influence of a DNA sequence (here called sequence A) on its adjacent sequence (here called sequence B), we linked some DNA repeated sequences to the 5′-end of the T7 promoter in the plasmid pET-42a (+) or the 5′- and/or 3′-end(s) of the EcoRI site in some DNA fragments using PCR and other molecular cloning methods. As a result, we found that the efficiency of the T7 promoter and EcoRI could be impacted by some flanking sequences, indicating that sequence B could be impacted by sequence A. The features of such influence include the following: (i) sequence A can directly impact sequence B without changing/modifying the base composition of sequence B or destroying the inherent connection between sequence B and its function-related sequences; (ii) such influence does not need the participation of trans-acting factors or products of sequence A (if any); (iii) such an influence might be undetectable when the activities of trans-acting factors of sequence B are normal but might become detectable when those are lower than the normal one; (iv) such an influence might be enhancive, inhibitory, or unobvious; (v) the influence of sequence A linked to the 5′-end of sequence B might be the same as or opposite to that of sequence A linked to the 3′-end; and (vi) the influences of sequence A linked to different ends of sequence B could enhance or partially offset each other when sequence A is linked to both 5′- and 3′-ends of sequence B. These findings might give us a further understanding of the interaction of two adjacent DNA sequences.

INTRODUCTION

Most human genes consist of exons, introns, cis-regulatory elements, and so forth, although 15% of all human genes do not contain introns.1 Cis-regulatory elements include the promoter, enhancer and silencer, and so forth. Trans-acting factors must recognize and bind to their corresponding cis-regulatory elements to initiate or inhibit the transcription of a gene. Therefore, the biological functions of a gene could be impacted by the mutations in the promoter, exon, splice site or enhancer, and so forth. For example, the biosynthesis of β-globin can be inhibited or completely terminated by the enhancer, and so forth. Additionally, the transcription of a gene can also be impacted by the enhancer and silencer far away from the gene, which requires the participation of trans-acting factors. There are many repeated sequences in a human genome.2 They were thought as the “junk” or “selfish” DNA.3 However, now it has been found that the repeated sequences participate in the regulation of gene expression, the formation and maintenance of the genome conformation, and so forth. The expansion of trinucleotides (such as CAG, CGG, CTG, GAA, and GCC) can change the base composition of genes and hence result in some diseases, for example, (i) fragile X syndrome caused by the expansion of CGG repeats in the 5′-untranslated region of the fragile X mental retardation (FMR1) gene; (ii) Friedreich ataxia caused by the expansion of GAA repeats in the intron of the X25 gene; (iii) myotonic dystrophy caused by the expansion of CTG repeats in the 3′-UTR of the protein kinase (DMPK) gene; and (iv) Huntington’s disease caused by the expansion of CAG repeats in the exon of the Huntingtin gene.4–10

DNA transposons are one type of repeated sequences. They account for about 44% of the human genome.11 Transposons can change their position in a genome. Thereby, sometimes they may create or reverse some mutations or damage a gene when they insert themselves in, leave, or increase the copies of a repeated sequence in the gene.12,13 In addition, they can result in aberrant expression of the linked genes if they contain promoters. Therefore, DNA transposons can increase cancer predisposition and even cause some diseases such as hemophilia A and B, severe combined immunodeficiency, porphyria, and so forth.14,15

In addition to the mutations or DNA sequences (mentioned above) that can change the base composition of a functional
DNA sequence or destroy the inherent connection between the functionally related DNA sequences, some in vitro experiments have shown that some DNA sequences could directly impact the biological function or activity of their adjacent sequences without changing the base composition of their adjacent sequences or destroying the inherent connection between the functionally related DNA sequences. For example, the affinity between the T7 RNA polymerase and T7 promoter or the transcription activities of some promoters could be impacted by the 5′-flanking sequences of promoters.16–22 Additionally, it has been reported that the cleavage rate of the EcoRI site could be influenced by the context of the EcoRI site.23–27 Rosa and colleagues reported that the sequences containing an adenine following GGATG could be digested more efficiently by the restriction enzymes BtsCI and BseGI.28 In 2015, we found that the silver staining of an oligonucleotide (oligo) containing T base could be inhibited by its own T base.29 Schöne and colleagues found that some DNA sequences directly flanking the core-binding site of the glucocorticoid receptor (GR) could change the three-dimensional structure of the DNA-binding site, the DNA-binding domain of the GR, and the quaternary structure of the dimeric complex.30 These findings mentioned above indicate that the biological functions or activity of sequence B could be directly impacted by sequence A. However, the features of such influence seem to have not yet been completely elucidated. To explore further the features of such influence, we linked some DNA repeated sequences to the 5′-end of the T7 promoter in the plasmid pET-42a (+) or one or two side(s) of the EcoRI site and then observed the changes in the efficiency of the T7 promoter and the cleavage rate of the EcoRI site. Some interesting things were found and reported below.

## RESULTS

### GST-Tag Expression in the Reconstructed Plasmid pET-42a (+)

Reconstructed Plasmid pET-42a (+). The DNA fragments a, c, g, and t (about 335 bp long) used to prepare the reconstructed plasmid pET-42a (+) could be amplified using overlap PCR (Figure S1). Their DNA bands in 2% agarose gel are shown in Figure S2A. They could be digested with Xba I and Sph I. The digested bands of the DNA fragment a are shown in Figure S2B. The plasmid pET-42a (+) could be digested with Xba I, Sph I, and SgrAI (Figure S2C). After purification, the digestion products of the DNA fragments a, c, g, and t were ligated with those of the plasmid pET-42a (+), which generated four different types of reconstructed plasmids pET-42a (+), namely, the plasmids Pa, Pc, Pg, and Pt, respectively. Figure S3 shows that 5′-ends of T7 promoters of plasmids Pa, Pc, Pg, and Pt had been linked by the repeated sequences 5′-AAAAA-3′, 5′-CCCCC-3′, 5′-GGGGG-3′, and 5′-TTTTT-3′, respectively.

Detection of GST-Tag mRNA. The total RNA from the Escherichia coli strain BL21 (DE3) transformed by different types of the reconstructed plasmids is shown in Figure S4. There was a little residual DNA in the total RNA before treatment with DNase I (Figure S4A). After being treated with RNase-free DNase I, the residual DNA was undetectable (Figure S4B). The results of qRT-PCR for GST-Tag mRNA are shown in Table 1 and Figure 1. There was no significant expression difference in GST-Tag mRNA between different types of plasmids (P > 0.05) when GST-Tag was induced with 1 mM IPTG for 60 min.

| plasmids | N | ΔCt | fold change | P value |
|----------|---|-----|-------------|---------|
| NC       | 5 | −2.046 ± 1.69 | 0.0034 | 0.00 |
| Po       | 5 | −10.96 ± 0.62 | 1.00 | 0.99 |
| Pa       | 5 | −10.95 ± 0.57 | 0.98 | 0.93 |
| Pc       | 5 | −10.91 ± 1.03 | 1.10 | 0.77 |
| Pg       | 5 | −11.11 ± 0.56 | 1.09 | 0.69 |
| Pt       | 5 | −11.16 ± 1.11 | 1.30 | 0.69 |

*Table 1. Results of qRT-PCR for GST-Tag mRNA*

(1) GST-Tag was induced with 1 mM IPTG for 60 min; (2) N represents the experimental number; (3) NC = negative control; (4) P < 0.05 between the negative control and the other types of plasmids; and (5) P > 0.05 among all types of plasmids except the negative control.

![Figure 1. Real-time PCR results of GST-Tag mRNA expressed by the E. coli strain BL21 (DE3) transformed by different types of plasmids.](https://dx.doi.org/10.1021/acsomega.0c02264)

Nc was the negative control of which total RNA was extracted from the E. coli strain BL21 (DE3) that was transformed with the plasmid Po but not induced with IPTG. Po, Pa, Pc, Pg, and Pt represent the real-time PCR results of GST-Tag mRNA expressed by the E. coli strain BL21 (DE3) transformed by the plasmids Po, Pa, Pc, Pg, and Pt, respectively.

Detection of GST-Tag Protein. GST-Tag and Glyceraldheyde-3-phosphate dehydrogenase (GAPDH) proteins expressed by different types of plasmids were detected using the western blotting (WB) method. The WB bands of GST-Tag and GAPDH proteins are shown in Figure 2.

WB bands of GST-Tag and GAPDH are indicated with arrows. Lane M: protein marker [TureColor Tricolor predyed protein marker (Sangon Biotech, China)]. Lane N1: the negative control 1: the bacterial proteins extracted from the E. coli strain BL21 (DE3) which was transformed with the plasmid Po but not induced with IPTG. Lane N2: the negative control 2: the bacterial proteins extracted from the E. coli strain BL21 (DE3) which was induced with IPTG but not transformed with any type of plasmids. Lanes Po, Pt, Pa, Pg, and Pc: GST-Tag or GAPDH expressed by the E. coli strain BL21 (DE3) transformed by the plasmids Po, Pa, Pc, Pg, and Pt, respectively.

The GST-Tag protein consists of 220 amino acids. Its molecular weight should be about 26 kDa. However, because the termination codon of the GST-Tag gene in the plasmid pET-42a (+) did not occur until the 318th codon, the polypeptide chain of GST-Tag was extended to the 318th amino acid residues. The molecular weight of the extended
GST-Tag was about 36 kDa that was calculated according to the number of its amino acid residues. Therefore, WB bands of the GST-Tag protein in Figure 2 were a little slower than the 33 kDa band of the protein marker. The expression amount of the GST-Tag protein was represented with the ratios of ΔIntDen of GST-Tag to that of GAPDH (Table 2 and Figure 3). The higher the ratio was, the higher the expression amount of the GST-Tag protein was.

Po, Pt, Pa, Pg, and Pc represent the plasmids Po, Pt, Pa, Pg, and Pc, respectively. NRE represents the number of repeated experiments. (A) GST-Tag was induced with 1 mM IPTG for 60 min; (B) GST-Tag was induced with 0.33 mM IPTG for 30 min; (C) GST-Tag was induced with 0.11 mM IPTG for 30 min; (D) GST-Tag was induced with 0.05 mM IPTG for 20 min; and (E) GST-Tag was induced with 0.028 mM IPTG for 20 min.

The expression difference in the GST-Tag protein between different types of plasmids was unobvious when the GST-Tag protein was induced with 0.33−1 mM IPTG for ≥30 min but became obvious when the GST-Tag protein was induced with 0.05 mM IPTG for 20 min (Table 2, Figures 2 and 3). Under latter inducement conditions, the amount of the GST-Tag protein expressed by the plasmids Po, Pa, Pc, Pg, and Pt could be arranged from high to low as Po > Pt ≥ Pa > Pg ≥ Pc, even though the amounts of the GST-Tag protein expressed by the plasmids Po, Pa, Pc, Pg, and Pt could be arranged from high to low as Po > Pa > Pg ≥ Pc, even though the amounts of the GST-Tag protein were already pretty low at this moment. The amount of the GST-Tag protein expressed by the plasmids Po or Pg was even too little to be detected using WB. There was statistically significant difference in GST-Tag protein expression between the plasmids Pt and Pc or Pg (P < 0.05). These results indicated that the efficiency of the T7 promoter could be impacted by some repeated sequences linked to its 5′-end when the GST-Tag protein was induced with 0.028 or 0.05 mM IPTG for 20 min.

Digestion of Fragments 1−14 with EcoR I. To observe the influence of the flanking sequences of the EcoRI site on the cleavage rate of the EcoRI site, DNA fragments 1−14 (82−89 bp long) were prepared using 15 oligos (Figure S5). The undigested and digested bands of fragments 1−14 were about 82−89 and 41−46 bp long, respectively (Figure 4).

The digested and undigested DNA bands are indicated with arrows. Lane M: DNA marker [GeneRuler 50 bp DNA Ladder (Thermo, USA)]. (A) DNA fragments 1−10 without digestion. Lanes 1−10: Fragments 1−10; (B) DNA fragments 11−14 without digestion. Lanes 1−4: Fragments 11−14; (C) DNA fragments 1−8 were digested with 2.5 U/10 μL EcoR I for 30 min. Lanes 1−8: Fragments 1−8; (D) DNA fragments 9−14 were digested with 2.5 U/10 μL EcoR I for 30 min.
Table 2. Ratio of ΔIntDen of GST-Tag WB Bands to That of GAPDHfl

| N  | I/T (mM/min) | Po | Pt | Pa | Pg | Pc |
|----|--------------|----|----|----|----|----|
| 1  | 1/60         | 3.220 | 2.472 | 2.949 | 2.650 | 2.933 |
| 2  | 0.33/30      | 4.786 | 4.387 | 7.460 | 5.833 | 7.074 |
| 3  | 0.11/30      | 5.354 | 4.406 | 3.788 | 3.833 | 3.497 |
| 4  | 0.05/20      | 1.548 | 1.091 | 0.805 | 0.736 | w/o |
| 5  | 0.05/20      | 1.458 | 0.650 | 0.422 | 0.292 | w/o |
| 6  | 0.05/20      | 0.515 | 0.357 | 0.340 | 0.090 | 0.070 |
| 7  | 0.05/20      | 0.521 | 0.434 | 0.517 | 0.137 | 0.118 |
| 8  | 0.02/20      | 0.019 | 0.380 | 0.290 | w/o | w/o |
| 9  | 0.02/20      | 0.040 | 0.115 | 0.040 | 0.021 | 0.015 |
| 10 | 0.02/20      | 2.530 | 1.633 | 0.490 | 0.084 | 0.077 |
| 11 | 0.02/20      | 0.020 | 0.309 | 0.284 | w/o | w/o |
| 12 | 0.02/20      | 1.313 | 1.604 | 0.993 | 1.107 | 0.769 |
| 13 | 0.02/20      | 0.107 | 0.181 | 0.058 | w/o | w/o |

a (1) N represents the experimental number; (2) I/T = IPTG concentration/induction time; (3) w/o represents that the WB band of GST-Tag was invisible; (4) when GST-Tag was induced with 0.05 mM IPTG for 20 min, the average ratios of ΔIntDen of GST-Tag to that of GAPDH expressed by the plasmids Po, Pt, Pa, Pg, and Pc were 1.011, 0.633, 0.521, 0.314, and 0.047, respectively. There were significant differences in the average ratios of ΔIntDen of GST-Tag to that of GAPDH between (i) the plasmids Po and Pt (P < 0.05) and (ii) the plasmid Pc and other plasmids (P < 0.05); and (5) when GST-Tag was induced with 0.028 mM IPTG for 20 min, the average ratios of ΔIntDen of GST-Tag to that of GAPDH expressed by the plasmids Po, Pt, Pa, Pg, and Pc were 0.672, 0.704, 0.359, 0.202, and 0.144, respectively. There was significant difference between the plasmids Pt and Pg or Pc (P < 0.05).

Lanes 1−6: Fragments 9−14; (E) DNA fragments 1−4, 9, and 10 were digested with 1 U/10 μL EcoR I for 12 min. Lanes 1−8: Fragments 1, 2, 3, 4, 9, 10, 13, and 14, respectively; (F) DNA fragments 5−8 were digested with 1 U/10 μL EcoR I for 12 min. Lanes 1−6: Fragments 5, 6, 7, 8, 13, and 14, respectively; (G) fragments 9−14 were digested with 1 U/10 μL EcoR I for 12 min. Lanes 1−6: Fragments 9−14, respectively.

The cleavage rate of the EcoRI site was represented using the percentage of the gray value (PGV) of the digested DNA bands accounting for the total gray value of both the digested and undigested DNA bands. The higher the PGV was, the higher the cleavage rate of the EcoRI site was. When fragments 1−13 were digested with 2.5 U/10 μL EcoR I for 30 min, they were almost completely digested (Figure 4C,D), and the difference in the PGV among them was unobvious. However, when they were digested with 1 U/10 μL EcoR I for 12 min, the difference in the PGV among some of them became obvious (Tables 3−5 and Figure 5).

Table 3 and Figure 5A demonstrate that the average PGV of the fragments 2, 3, 4, 10, or 13 was higher than that of the fragment 1 (P < 0.05), indicating that the cleavage rate of the EcoRI site of which both 5′- and 3′-ends were linked by the sequence 5′-CCCCC-3′, 5′-GGGGG-3′, 5′-TTTTT-3′, or the original sequence and the cleavage rate of the EcoRI site of which the 3′-end was linked by the sequence 5′-AAAAA-3′ was higher than that of the EcoRI site of which both 5′- and 3′-ends were linked by the sequence 5′-AAA-3′ was higher than that of the EcoRI site of which only the 5′-end was linked by the sequence 5′-AAAAA-3′. Additionally, the average PGV of the fragment 3 was higher than that of the fragments 4, 9, or 13 (P < 0.05), indicating that the cleavage rate of the EcoRI site of which both 5′- and 3′-ends were linked by the sequence 5′-GGGG-3′ was higher than that of the EcoRI site of which both 5′- and 3′-ends were linked by the sequence 5′-TTTT-3′ or the original sequence or higher than that of the EcoRI site of which the 5′-end was linked by the sequence 5′-AAA-3′. There was no significant difference in the average PGV among the fragments 5−8 (P > 0.05, Table 5 and Figure 5B). Table 5 and Figure 5C show that the average PGV of the fragment 9 was lower than that of the fragment 10, 11, 12, or 13 (P < 0.05), meaning that the cleavage rate of the EcoRI site of which the 5′-end was linked by the sequence 5′-AAAAA-3′ was lower than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-CCCCC-3′ or the original sequence or lower than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-GGGGG-3′ or 5′-CCCCC-3′. Additionally, the average PGV of the fragment10 was higher than that of the fragments 2, 3, 9, and 12 (P < 0.05), indicating that the cleavage rate of the EcoRI site of which the 5′-end was linked by the sequence 5′-AAAA-3′ was higher than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-CCCC-3′ or the original sequence of the EcoRI site of which the 5′-end was linked by the sequence 5′-TTTT-3′ or the original sequence. The average PGV of the fragments 11 was higher than that of the fragment 13, meaning that the cleavage rate of the EcoRI site of which the 5′-end was linked by the sequence 5′-CCCCC-3′ was higher than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-AAAAA-3′.

The mechanism why the efficiency of the T7 promoter could be impacted by some repeated sequences linked to the 5′-end of the T7 promoter or the 5′- and/or 3′-end(s) of the EcoRI site.

DISCUSSION

The findings mentioned above indicated that the efficiency of the T7 promoter or the cleavage rate of the EcoRI site could be impacted by some repeated sequences linked to the 5′-end of the T7 promoter or the 5′- and/or 3′-end(s) of the EcoRI site.

The mechanism why the efficiency of the T7 promoter could be impacted by its 5′-flanking sequence is not very clear. T7 RNA polymerase consists of a single polypeptide chain (Mw 11,000) and requires Mg++ as its cofactor.51-55 According to the manufacture’s manual, the expression of T7 RNA polymerase in the E. coli strain BL21 (DE3) is also induced by IPTG. Therefore, the expression of T7 RNA polymerase in the E. coli strain BL21 (DE3) will reduce or increase as the concentration of IPTG used decreases or increases. It was supposed that the flanking sequences of a promoter could change the conformation of the promoter by kinking or bending and hence impact on the biochemical activity or function of the promoter.17-20 In addition, we found that the silver staining of an oligo containing T base could be inhibited by its own T bases, which might indicate that the ionic state of or around an oligo could be changed by its base composition because DNA silver staining involves the redox reaction of silver ions. However, we do not know if the ionic state of around the 5′-flanking sequence of the T7 promoter.
could impact the efficiency of the T7 promoter although T7 RNA polymerase requires Mg$^{2+}$ as its cofactor.

According to the manual of New England Biolabs (NEB), the activity of restriction enzymes can be impacted by DNA supercoil and the number of bases flanking to the restriction recognition sequence (https://international.neb.com/applications/cloning-and-synthetic-biology/dna-analysis/-/media/nebus/files/brochures/restendo_techguide.pdf). EcoRI can still retain more than 90% activity even if only one base is left at each side of the EcoRI site (https://international.neb.com/~media/NebUs/Files/Chart%20image/cleavage_olignucleotides_old.pdf). Additionally, it has been reported that the cleavage rates of the EcoRI site could be impacted by the sequence context of the EcoRI site.23−27 Our experimental results also suggested that the cleavage rate of the EcoRI site could be impacted by some flanking sequences of the EcoRI site (Tables 3−6, Figures 4 and 5), although we found something different from what had been reported. From Table 6, we could find that the effect of flanking sequences of the EcoRI site on the cleavage rate of the EcoRI site could be (i) enhanced; (ii) inhibitory; or (iii) unobvious. When a flanking sequence was linked to the 5′-end of the EcoRI site, its effect on the cleavage rate of the EcoRI site might be the same (or almost the same) as or opposite to that when it was linked to the 3′-end of the EcoRI site or vice versa. For example, the effect of the sequence 5′-GGGG-3′ linked to the 5′-end of the EcoRI site was almost the same as that linked to the 3′-end of the EcoRI site. However, the effect of the sequence 5′-AAAA-3′ linked to the 5′- or 3′-end of the EcoRI site was inhibitory or enhanced, respectively. It was interesting that the effects of the 5′- and 3′-flanking sequences of the EcoRI site on the cleavage rate of the EcoRI site could offset or enhance each
For example, the effect of the sequence 5′-AAAA-3′ linked to 5′- and 3′-ends of the EcoRI site was inhibitory and enhanced, respectively. However, the cleavage rate of the EcoRI site of which both 5′- and 3′-ends were linked by the sequence 5′-AAAA-3′ was higher than that of the EcoRI site of which the 5′-end was linked by the sequence 5′-AAAA-3′ but lower than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-AAAA-3′, indicating that the effects of the sequence 5′-AAAA-3′ on the EcoRI site could be partially offset when it was simultaneously linked to both 5′- and 3′-ends of the EcoRI site. The example of the

Figure 4. Digested and undigested DNA bands of fragments 1−14 in 3% agarose gels.

Table 3. PGV of DNA Fragments in Digestion Group 1

| N | frag 1 | frag 2 | frag 3 | frag 4 | frag 9 | frag 10 | frag 13 |
|---|--------|--------|--------|--------|--------|--------|--------|
| 1 | 73.8   | 83.6   | 82.1   | 73.2   | 49.9   | 81.4   | 73.9   |
| 2 | 48.0   | 73.8   | 84.5   | 58.5   | 38.1   | 83.8   | 58.6   |
| 3 | 45.6   | 73.4   | 67.8   | 55.8   | 33.7   | 81.4   | 45.4   |
| 4 | 47.8   | 75.7   | 78.7   | 50.9   | 33.0   | 81.4   | 59.7   |
| 5 | 40.7   | 77.0   | 83.9   | 58.7   | 35.8   | 92.3   | 62.3   |
| 6 | 34.2   | 57.2   | 66.2   | 40.8   | 25.0   | 81.3   | 43.4   |
| 7 | 49.3   | 78.8   | 90.9   | 60.6   | 35.6   | 89.1   | 50.2   |
| 8 | 53.3   | 70.5   | 75.7   | 54.5   | 54.0   | 90.3   | 57.3   |
| 9 | 48.3   | 71.5   | 78.5   | 55.2   | 37.4   | 77.9   | 56.5   |
| 10| 49.1   | 71.8   | 73.2   | 58.9   | 40.4   | 77.8   | 58.0   |
| mean| 49.0   | 73.3   | 78.2   | 56.7   | 38.3   | 83.7   | 56.5   |

(1) N represents the experimental number; (2) Frag = fragment; and (3) P < 0.05 (i) between the mean values of the PGV of the fragments 1 and 2, 3, 4, 9, 10, or 13; (ii) between the mean values of the PGV of the fragments 2 and 4, 9, 10, or 13; (iii) between the mean values of the PGV of the fragments 3 and 4, 9, or 13; (iv) between the mean values of PGV of the fragments 4 and 9 or 10; (v) between the mean values of PGV of the fragments 9 and 10 or 13; and (vi) between the mean values of PGV of the fragments 10 and 13.

Table 4. PGV of DNA Fragments in Digestion Group 2

| N | frag 5 | frag 6 | frag 7 | frag 8 | frag 13 |
|---|--------|--------|--------|--------|--------|
| 1 | 99.2   | 98.8   | 99.2   | 91.9   | 97.9   |
| 2 | 91.8   | 94.1   | 88.2   | 86.6   | 70.2   |
| 3 | 93.2   | 91.2   | 91.8   | 89.7   | 33.5   |
| 4 | 96.0   | 95.3   | 85.5   | 94.4   | 76.3   |
| 5 | 90.5   | 87.5   | 78.0   | 84.8   | 53.9   |
| 6 | 67.3   | 87.4   | 86.5   | 75.9   | 77.0   |
| 7 | 92.7   | 91.1   | 93.1   | 90.8   | 92.3   |
| 8 | 92.6   | 90.3   | 94.7   | 89.3   | 87.6   |
| 9 | 86.3   | 89.6   | 89.0   | 88.6   | 82.4   |
| 10| 47.6   | 61.9   | 54.5   | 44.7   | 53.6   |
| 11| 36.0   | 46.3   | 48.8   | 48.7   | 42.7   |
| 12| 73.6   | 72.8   | 62.4   | 66.8   | 53.8   |
| 13| 32.1   | 56.1   | 49.8   | 29.8   | 32.2   |
| 14| 66.5   | 61.9   | 65.8   | 66.0   | 53.2   |
| mean| 76.1   | 80.3   | 77.7   | 74.9   | 64.8   |

(1) N represents the experimental number; (2) Frag = fragment; and (3) P > 0.05 among the mean values of PGV of DNA fragments 5−8 and 13.

which the 5′-end was linked by the sequence 5′-AAAA-3′ but lower than that of the EcoRI site of which the 3′-end was linked by the sequence 5′-AAAA-3′, indicating that the effects of the sequence 5′-AAAA-3′ on the EcoRI site could be partially offset when it was simultaneously linked to both 5′- and 3′-ends of the EcoRI site. The example of the
The enhancement effect was that the effect of the sequence 5'-GGGG-3' on the EcoRI site was unobvious when it was linked to either the 5'- or 3'-end of the EcoRI site but became enhanced when it was simultaneously linked to both 5'- and 3'-ends of the EcoRI site (Table 6).

It had been reported that the structure and flexibility of the EcoRI site could be impacted by its flanking sequences. Additionally, Mg⁺⁺ is required for the activity of EcoRI. Therefore, we supposed that the mechanism by which the cleavage rate of the EcoRI site was impacted by its flanking sequences was the same as that by which the T7 promoter was impacted by its 5' flanking sequences. Because restriction enzymes are one of the bacterial defense mechanisms against viruses, we infer that it might be possible that the invasiveness of a virus could change as the flanking sequences of viral restriction endonuclease sites change, especially when the activity of restriction endonuclease in the viral host decreases.

The findings mentioned above indicate that sequence A can directly impact the biological function of sequence B without the participation of trans-acting factors of sequence A. However, it has to be mentioned that such influence does not always occur or always inhibits sequence B. We supposed that such influence might occur only when the change of base composition of sequence A was great enough to alter the conformation and/or the ionic state of/around sequence B. Because the change of the conformation and/or the ionic state of/around sequence B caused by sequence A is usually mild, the influence of sequence A on the biological function or activity of sequence B is generally mild, too. Additionally, it was because the concentrations and actuation duration of IPTG and EcoRI used by us were lower/shorter than those recommended by the manufacturers that we could detect the change in the efficiency of the T7 promoter or the cleavage rate of the EcoRI site. For example, the final concentration of

Table 5. PGV of DNA Fragments in Digestion Group 3

| N | frag 9 | frag 10 | frag 11 | frag 12 | frag 13 |
|---|-------|--------|--------|--------|--------|
| 1 | 33.4  | 81.1   | 72.8   | 62.8   | 50.8   |
| 2 | 29.8  | 86.6   | 78.0   | 68.5   | 69.1   |
| 3 | 16.8  | 57.1   | 26.0   | 31.5   | 25.4   |
| 4 | 23.6  | 80.6   | 74.4   | 76.7   | 46.6   |
| 5 | 23.0  | 85.9   | 69.6   | 81.4   | 48.0   |
| 6 | 38.9  | 65.5   | 67.0   | 51.2   | 48.5   |
| 7 | 36.3  | 78.2   | 74.8   | 61.4   | 71.4   |
| 8 | 41.3  | 87.0   | 82.4   | 64.3   | 58.1   |
| 9 | 33.9  | 82.6   | 76.1   | 64.1   | 50.3   |
| 10| 30.7  | 69.5   | 70.6   | 67.7   | 18.8   |
| 11| 22.9  | 83.2   | 81.5   | 57.2   | 57.8   |
| mean| 30.1 | 77.9 | 70.3 | 62.4 | 49.5 |

*(1) N represents the experimental number; (2) Frag = fragment; and (3) P < 0.05 (i) between the mean values of PGV of the fragments 9 and 10, 11, 12, or 13; (ii) between the mean values of PGV of the fragments 10 and 11, 12, or 13; and (iii) between the mean values of PGV of the fragments 11 and 13.

Figure 5. Average PGV of different DNA fragments digested with EcoRI. NRE represents the number of repeated experiments. Frag = fragment. (A) Average PGV of the DNA fragments 1—4, 9, 10, and 13; (B) average PGV of the DNA fragments 5—8 and 13; and (C) average PGV of the DNA fragments 9—13.
IPTG and EcoRI recommended by the manufacturers was 1 mM and 2.5–10 units/10 μL, respectively. The actuation duration for IPTG and EcoRI recommended by the manufacturers was 8–10 h or 1–16 h, respectively. When the difference in the GST-Tag expression or the cleavage rate of the EcoRI site was detectable, the concentration of IPTG or EcoRI used by us was only about 1/30 or 1/5 of that recommended by the manufacturers, respectively, and the actuation duration for IPTG and EcoRI was only about 1/24–1/30 or 1/5–1/80 of that recommended by the manufacturers, respectively. These findings indicated that the influence of sequence A on sequence B might be negligible under normal conditions but might become noticeable under abnormal conditions. Considering that many biological effects in an organism can usually be accumulated or amplified through a series of steps (such as a cascade reaction, etc), we infer that the influence of sequence A on sequence B might be one of the causes of some individual differences (such as drug tolerance, disease susceptibility, etc).

Conclusion: The efficiency of the T7 promoter or EcoRI could be impacted by the flanking sequences of the T7 promoter or EcoRI site, respectively, indicating that the biological function or activity of sequence B could be impacted by sequence A. The features of such influence include the following: (i) sequence A can directly impact sequence B without changing/modifying the base composition of sequence B or destroying the inherent connection between sequence B and its function-related sequences; (ii) such influence does not need the participation of trans-acting factors or products of sequence A (if any); (iii) such influence might be undetectable when the activities of trans-acting factors of sequence B are normal but might become detectable when they are lower than the normal ones; (iv) such influence might be enhancive, inhibitory, or unobvious; (v) the influence of sequence A linked to the 5′ end of sequence B might be the same as or opposite to that of sequence A linked to the 3′ end and vice versa; and (vi) the influences of sequence A linked to different ends of sequence B could enhance or partially offset each other when sequence A is linked to both 5′ and 3′-ends of sequence B. These findings might give us a further understanding to the interaction of two adjacent DNA sequences.

Table 6. Effect of Flanking Sequences of the EcoRI Site and the Differences in PGV

| Dg | fragments | Fs (5′−3′) | location | effect | differences in PGV |
|----|-----------|-----------|----------|--------|---------------------|
| 1  | AAAA      | 5′-and 3′-ends | ↓ | >fragments 9; <fragments 2, 3, 4, 10, 13† |
| 2  | CCCC      | 5′-and 3′-ends | ↑ | >fragments 1, 4, 9, 13; <fragments 10‡; =fragment 3‡ |
| 3  | GGGG      | 5′-and 3′-ends | ↑ | >fragments 1, 4, 9, 13‡; =fragments 2, 10‡ |
| 4  | TTTT      | 5′-and 3′-ends | unobvious | >fragments 1, 9†; <fragments 2, 3, 10‡; =fragment 13‡ |
| 9  | AAAA      | 5′-end | ↓ | <fragments 1, 2, 4, 9, 13‡; =fragment 3‡ |
| 10 | AAAA      | 3′-end | ↑ | >fragments 1, 2, 4, 9, 13‡; =fragment 3‡ |

*(1) Dg represents the digestion group; (2) Fs represents the flanking sequence; (3) the comparison of PGV was performed only among the different DNA fragments in the same digestion group; (4) arrows ↑ and ↓ represent enhanced and inhibitory, respectively, comparing with the positive control. Equality sign (=) represents no significant difference; and (5) †(P < 0.05). ‡(P > 0.05).

Table 7. PCR Primers Used to Amplify the DNA Fragments a, c, g, and t

| primers | sequences (5′−3′) | frag |
|---------|------------------|------|
| T7UP    | GTTAAACAAAATATTATCtagGGGGAGATTGTTATCCG | T |
| T7DP    | AGCCCCGATGGGTTAGGCCCG | c |
| T7UPa   | GAGTTGTATCCGCTCAACAATTCCTATTAGTAGTGCATTAAAGGGCGGATCGAG | a |
| T7UPc   | GAGTTGTATCCGCTCAACAATTCCTATTAGTAGTGCATTACCCCGCGGAT | a |
| T7UPg   | GAGTTGTATCCGCTCAACAATTCCTATTAGTAGTGCATTAGGGGCGGATCGAG | a |
| T7UPt   | GAGTTGTATCCGCTCAACAATTCCTATTAGTAGTGCATTATTCCCCCGGAT | a |

*(1) The sequence written in lowercase letters is the Xba I site; (2) the underlined sequences are the antisense chain of the T7 promoter; (3) The sequences written in bold italic letters are the repeated sequences linked to the 3′-end of the antisense chain of the T7 promoter; and (4) Frag = fragment.

Expression of mRNA and Protein of GST-Tag in the E. coli Strain BL21 (DE3). Reconstruction of the Plasmid pET-42a (+). In our previous experiment, we found that the intensity of the cyanine staining of some oligos could be significantly altered when the changes in base composition of the oligos accumulated to a certain extent (usually 4 nt). Additionally, the silver staining of the oligos (A1T4, G1T4) could be completely inhibited by their own T bases.29 Therefore, to observe the influence of a short repeated
sequence on the efficiency of the T7 promoter, the repeated sequence S'-AAAAA-3', S'-CCCCC-3', S'-GGGGG-3', or S'-TTTTT-3' was linked to the 5'-end of the T7 promoter in the plasmid pET-42a (+) by the following four steps.

**Amplification of the DNA Fragments a, c, g, and t.**

1. PCR Primers: The DNA fragments a, c, g, and t were amplified using overlap PCR (Figure S1). Each of them contained a Sph I site in 5'-side, an Xba I site in 3'-side, and a T7 promoter near Xba I site. The 5'-ends of T7 promoters in the DNA fragments a, c, g, and t were linked by the repeated sequences S'-AAAAA-3', S'-CCCCC-3', S'-GGGGG-3', and S'-TTTTT-3', respectively. The PCR primers used are listed in Table 1 and were synthesized by Sangon Biotech (China).

The primer T7UP with an Xba I site was the common forward primer of the first- and third-round PCR. Primer T7DP was the common reverse primer. The primers T7UPa, T7UPc, T7UPg, and T7UPt contained an antisense chain of the T7 promoter of the forward primer of the specific fragment.

2. PCR reaction system: The volume of the first-, second-, and third-round PCR was 50 μL, 10 μL, and 50 μL, respectively. PCR mixture contained 2X phanta turbo master mix (Vazyme Biotech, China), 0.2 μM each of forward and reverse primers, and 1 μL of PCR templates per 10 μL of PCR mixture. The template of the first-round PCR was 2.5 pg/μL plasmid pET-42a (+). The first- and second-round PCR products were diluted 10–30 folds with ddH2O and used as the templates of the second- and third-round PCR, respectively.

3. PCR cycle conditions: (i) the first-round PCR: 95 °C, 3 min, followed by 15 cycles of 94 °C, 30 s, and 67 °C, 60 s; (ii) the second-round PCR: 95 °C for 3 min, followed by 15 cycles of 94 °C, 30 s, and 65 °C, 60 s; and (iii) the third-round PCR: 95 °C for 3 min, followed by 30 cycles of 94 °C, 30 s, and 64 °C, 90 s. PCR cycles were performed on a PCR system 9700 (ABI, USA).

4. Electrophoresis: The third-round PCR products were checked using 2% agarose gel electrophoresis. Briefly, 5 μL of the third-round PCR products was mixed with 1 μL of 6X Loading Dye, loaded on 2% miniagarose gel containing 0.01% (w/v) 4S Red Plus Nucleic Acid Stain (Sangon Biotech, China). Electrophoreses were run in 1X TAE buffer, at 100 V for 20–30 min, and then, DNA bands in the gels were analyzed using a Molecular Imager GEL system (Universal Hood II, Bio-Rad, USA).

5. The purification and concentration determination of PCR products: 100 μL of the third-round PCR products of each DNA fragment was purified using a universal DNA purification kit (Tiangen biotech, China) following the manufacturer’s instruction. DNA fragments were re-suspended with ddH2O. The concentrations of the purified DNA fragments a, c, g, and t were about 95, 66, 94, and 79 ng/μL, respectively, which were determined using a NanoDrop 2000 (Thermo Scientific, USA).

**Digestion of the DNA Fragments a, c, g, and t.**

1. The digestion reaction: The purified DNA fragments a, c, g, and t were digested with both QuickCut Xba I and Sph I (Takara Bio, China) following the manufacturer’s instruction. Briefly, these DNA fragments were digested in 155 μL of digestion reaction mixture containing 1× quickcut green buffer, 30 μL of the purified DNA fragment a, c, g, or t, and 5 μL of each of Sph I and Xba I, incubated at 37 °C for 5 min.

2. Electrophoresis: The digestion products of the DNA fragment a, c, g, or t were checked using 2% agarose gel electrophoresis at 100 V for 20 min. DNA bands in the gels were analyzed as mentioned above.

3. The purification and concentration determination of the digestion products: The digested DNA fragments were purified using a universal DNA purification kit as mentioned above. After purification, the concentrations of the digested DNA fragments a, c, g, and t were 6.9, 6.6, 6.9, and 6.6 ng/μL, respectively, which were determined as mentioned above.

**Digestion of the Plasmid pET-42a (+).**

1. The digestion reaction: The plasmid pET-42a (+) was digested with SgrA1 (NEB, USA) and QuickCut Xba I and Sph I (Takara Bio, China) following the manufacturer’s instruction. Briefly, the plasmid pET-42a (+) was digested in 50 μL of digestion reaction mixture containing 1× NEB buffer, 820 ng of plasmid pET-42a (+), 1 μL of each of Sph I and Xba I, and 2 U of SgrA1. Because there was a SgrA1 site between Xba I and Sph I in 3' end, we added 50 μL of digestion reaction mixture containing 1× NEB buffer, 370 μL of ddH2O, 2 U of SgrA1, and 50 units of T4 DNA ligase, in 1× T4 DNA ligase buffer, 132–138 ng of digested DNA fragments, 69 ng of the digested plasmid pET-42a (+), and 28 Weiss units of T4 DNA ligase. After digesting the plasmid pET-42a (+), the digestion products were purified using an Amicon Ultra-0.5 centrifugal filter unit. The digestion reaction mixture was incubated at 37 °C for 60 min.

2. Electrophoresis: The digestion products of the plasmid pET-42a (+) were checked using 1.2% agarose gel electrophoresis at 100 V for 25 min. DNA bands in the gels were analyzed as mentioned above.

3. The purification and concentration determination of the digested plasmids: The digested plasmids were purified with an Amicon Ultra-0.5 centrifugal filter unit with a Ultracel-100 membrane (Merck Millipore, USA) following the manufacturer’s instruction. Briefly, we added 50 μL of digestion reaction mixture into an Amicon Ultra-0.5 centrifugal filter unit, mixed with 350 μL of 1X TE, centrifuged at 14,000 rpm for 5 min, removed the liquid from the collection tube, washed with 370 μL of ddH2O, centrifuged at 14,000 rpm for 10 min, removed the liquid from the collection tube, washed one more time, turned the filter tube upside down and put it into a new collection tube, centrifuged at 12,000 rpm for 2 min, and collected the solution (about 40 μL) in a new centrifuge tube. The concentration of the digested plasmid pET-42a (+) was determined as mentioned above.
(Takara Bio, China) and it was incubated at 16 °C for 16 h. After ligation, we got four different types of the reconstructed plasmids, namely, the plasmids pET-42a (+)a, pET-42a (+)c, pET-42a (+)g, and pET-42a (+)t. Here, we called plasmids pET-42a (+)a, pET-42a (+)c, pET-42a (+)g, pET-42a (+)t, and the original plasmid as plasmids Pa, Pc, Pg, Pt and Po, respectively. Because the plasmids Pa, Pc, Pg, and Pt were linked by the repeated sequences 5′-AAAAA-3′, 5′-CCCCC-3′, 5′-GGGGG-3′, and 5′-TTTTT-3′, respectively.

Sequencing and Amplification of the Reconstructed Plasmids.

1 The transformation of the E. coli strain DH5α: DH5α competent cells (Tiangen biotech, China) were transformed using the plasmids Po, Pa, Pc, Pg, and Pt following the manufacturer’s instruction. Briefly, 10 μL of the plasmid Po (75 ng/μL) or the ligation products of each reconstructed plasmid described above were mixed gently with 100 μL of DH5α competent cells, placed in an ice-bath for 30 min, incubated in a water bath at 42 °C for 90 s, and quickly placed in an ice bath for 2–3 min. A total of 900 μL of sterilized LB culture medium was added to each reaction tube, mixed, then incubated at 37 °C and 200 rpm for 45 min to allow the bacteria to recover, and centrifuged at 12,000 rpm for 2 min. The supernatant was removed and bacterial pellets were resuspended in 300 μL of LB culture medium containing 100 μg/mL kanamycin. A total of 100 μL of this bacterial suspension was spread on the surface of LB agar containing 100 μg/mL kanamycin and then incubated at 37 °C for 16 h.

2 The sequencing of the reconstructed plasmids: A total of 100 μL of the bacterial suspension mentioned above was spread on the surface of LB agar containing 100 μg/mL kanamycin and incubated at 37 °C for 16 h. Five well-isolated colonies were picked up and transferred into LB culture medium (5 mL per colony) containing 100 μg/mL kanamycin, incubated at 37 °C and 200 rpm for 6 h until the OD 600 value reached 0.4–1.0, and then sequenced as described in the Sequencing and Amplification of the Reconstructed Plasmids.

3 The preparation of bacterial glycerol stocks: Bacterial glycerol stocks of each type of the reconstructed plasmids containing the correct target sequence were prepared by adding 160 μL of sterilized 50% (v/v) glycerol to 840 μL of bacterial suspension mentioned above and stored at −80 °C until use.

Induced Expression and Quantitative Analysis of GST-Tag mRNA. Induced Expression of GST-Tag mRNA. A total of 50 μL of bacterial glycerol stocks mentioned above was added to 3 mL of LB culture medium containing 100 μg/mL kanamycin and incubated at 37 °C and 200 rpm for about 4 h until the OD 600 value reached 0.4–1.0, and then, IPTG (1 mM final concentration) was added into LB culture medium and incubated at 37 °C and 200 rpm for 60 min.

Extraction of Total bacterial RNA. After IPTG induction, the total bacterial RNA in 2 mL of bacterial suspension mentioned above was extracted with an RNAiso Plus kit (Takara, China) following the manufacturer’s instruction. The total bacterial RNA was digested with DNAse I (Takara, China) following the manufacturer’s instruction. The concentration of total bacterial RNA was about 300–800 ng/μL, which were determined using a NanoDrop 2000 (Thermo Scientific, USA). Finally, the total bacterial RNA was stored at −80 °C until use.

Reverse Transcription and qRT-PCR for GST-Tag mRNA.

1 Primers: The primers used for the reverse transcription and qRT-PCR are listed in Table S1 and were synthesized by Sangon Biotech (China). Of them, the primers GSTU and GSTD were the forward and reverse primers of qRT-PCR for GST-Tag mRNA, respectively. Additionally, the primer GSTVR was also the reverse transcription primer of GST-Tag mRNA. The primers GAPDH and GAPDHV were the forward and the reverse primers of qRT-PCR for GAPDH mRNA, respectively. The primer GAPDHR was the reverse transcription primer of GAPDH mRNA.

2 The reverse transcription: The reverse transcription of GST-Tag and GAPDH mRNA were performed using the HiScript II QRT SuperMix for qPCR (+gDNA wiper) kit (Vazyme, China) following the manufacturer’s instruction. Briefly, to further remove the residual DNA in the total bacterial RNA before reverse transcription, the total bacterial RNA was treated in 16 μL of gDNA wiper mixture containing 4 μL of 4× gDNA wiper Mix (Vazyme, China), 1 μg of total bacterial RNA, 0.125 μM each of primers GSTU and GAPDH, and the 20 U RNase inhibitor (Takara, China). This mixture was incubated at 42 °C for 2 min, and then, 4 μL of 5X
**Table 8. Oligos Used to Prepare the DNA Fragments 1−14**

| oligos | sequences (5′−3′) | frag |
|--------|-------------------|------|
| 1      | ATCTCTTCCTGAGCCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAAAAAagattcAAAGAACTACTGTGCGGCATGGCACAGGAAACCAGGA   | 1    |
| 2      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 2    |
| 3      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 3    |
| 4      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 4    |
| 5      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 5    |
| 6      | ATCTCTTCCTGAGCCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 6    |
| 7      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 7    |
| 8      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 8    |
| 9      | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 9    |
| 10     | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 10   |
| 11     | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 11   |
| 12     | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 12   |
| 13     | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 13   |
| 14     | ATCTCTTCCTTGGCGAAATAGTACAGCTGCGGATTTGCCACACTCCTCAGGGAAGGAAGGAATGAGGATACCACTGTCGGGCAAGGAA   | 14   |

**Article**

ACS Omega 2020, 5, 23631−2364

---

**Detection of the GST-Tag Protein by Using WB.**

1 Electrophoresis and transfer of the bacterial protein: A total of 84 μg of bacterial protein was mixed with 18 μL of SX protein loading dye (Sangon Biotech, China), ddH₂O was added to 90 μL, and denatured at 99 °C for 5 min. A total of 15 μL of this mixture was loaded on 12% (w/v) precast-GelTris-Glycine PAGE (Sangon Biotech, China). Electrophoreses were run in 1× Tris-Glycine SDS PAGE running buffer (pH 8.3, Sangon Biotech, China) at 120 V for about 110 min. The proteins in the gels were transferred to PVDF membranes (Sangon Biotech, China) in EZ-Buffers C 1× western transfer buffer (Sangon Biotech, China) at 100 V for 1 h.

2 Controls: The GAPDH protein was used as the internal control. Additionally, two negative controls were set up: (i) The negative control 1 (N1): the bacterial proteins extracted from the E. coli strain BL21 (DE3) which were transformed with the plasmid Po but not induced with IPTG. (ii) The negative control 2 (N2): the bacterial proteins extracted from the E. coli strain BL21 (DE3) that was transformed with the plasmid Po but not induced with IPTG was used as the negative control.

---

**Induced Expression and Detection of the GST-Tag Protein. Induced Expression of the GST-Tag Protein.**

1 Inducement: A total of 150 μL of bacterial glycerol stocks prepared in Transformation of the E. coli Strain BL21 (DE3) was added to 4 mL of LB culture medium containing 100 μg/mL kanamycin, incubated at 37 °C and 200 rpm for about 3 h until the OD₆₀₀ value reached 0.4−1.0, then mixed with IPTG (0.028−1 mM final concentration), and incubated at 37 °C and 200 rpm for 20−60 min.

2 Extraction of bacterial proteins: The bacterial proteins were immediately extracted with a bacterial protein extraction kit (Sangon Biotech, China) following the manufacturer's instruction. In brief, 4 mL of bacterial suspension was centrifuged at 4 °C and 4000 rpm for 5 min. The supernatant was removed and the bacterial pellet was resuspended in 1.5 mL of PBS, and the bacterial suspension was transferred to a 2 mL centrifuge tube and centrifuged at 4 °C and 4000 rpm for 5 min. The supernatant was removed, and the bacterial pellet was resuspended in 320 μL of 1× cell lysis buffer. A total of 3.2 μL of PMSF and 6.4 μL of lysozyme were added and incubated at 37 °C for 30 min and then 37 °C and 200 rpm for 10 min, mixed with 2 μL of DNAse I/ RNase, incubated at 37 °C and 200 rpm for 10 min, centrifuged at 4 °C and 10,000 rpm for 30 min. A total of 190 μL of the supernatant was collected and mixed with 10 μL of 20× protease inhibitor complex in a new 1.5 mL centrifuge tube.

3 Concentration determination of bacterial proteins: The protein concentrations were about 2009−3998 ng/μL, which were determined with a Micro BCA protein assay kit (Sangon Biotech, China) following the manufacturer's instruction.
3 Block of the PVDF membrane: The transferred PVDF membranes mentioned above were blocked with BSA blocking buffer (Beijing CoWin Biotech, China) on a horizontal shaking platform at room temperature (RT) for 1 h.

4 Hybridization and washing: The primary antibodies for GST-Tag and GAPDH proteins were anti-GST (Cell Signaling Technology, USA) and anti-GAPDH (Abcam, UK), respectively. They were diluted 1:1000 with primary antibody dilution buffer. The secondary antibody was goat anti-mouse IgG H&L (Alexa Fluo 680) (Abcam, UK). It was diluted 1:10,000 with secondary antibody dilution buffer. The PVDF membranes blocked were hybridized with anti-GST, anti-GAPDH, and goat anti-mouse IgG H&L, respectively, following the manufacturer’s instruction. Briefly, one of two PVDF membranes blocked was hybridized with anti-GST (1:1000), while another was hybridized with anti-GAPDH (1:1000), incubated on a horizontal shaking platform at RT for 3 h, washed with 1× TBST (Sangon Biotech, China) four times, 5 min/time, then incubated with goat anti-mouse IgG H&L (1:10,000) on a horizontal shaking platform at RT for 1 h, and washed with 1× TBST for twice (5 min/time).

5 Detection of WB bands: Finally, the PVDF membranes were scanned using an infrared fluorescence imager (ODYSSEY, LI-COR, USA).

When the GST-Tag protein was induced with 0.028 and 0.05 mM IPTG, the experiments from IPTG inducement to SDS-PAGE electrophoresis of total RNA; and schematic diagram showing overlap PCR; three photos of agarose gel electrophoresis of DNA; four DNA sequencing results; two photos of agarose gel electrophoresis of total RNA; and schematic diagram showing the preparation of the DNA fragments 1–14 (PDF).

Digestion of the DNA Fragments 1–14 with EcoRI.

Preparation of the DNA Fragments 1–14. Design and Synthesis of Oligos. The DNA fragments 1–14 (82–89 bp long) were prepared using 15 oligos (Figure S5). These oligos are listed in Table 8 and were synthesized by Sangon Biotech (China).

The Oligos 1–14 were used as the template of oligo R. The extension of oligo R on the oligos 1–14 could generate the DNA fragments 1–14, respectively. There was an EcoRI site in the middle of the fragments 1–13. The 5′- and/or 3′-end(s) of the EcoRI site in the fragments 1–12 were linked by the 4 nt repeated sequence 5′-AAAA-3′, 5′-CCCC-3′, 5′-GGGG-3′, or 5′-TTTT-3′. The fragment 13 was used as the positive control of which 5′- and 3′-ends of the EcoRI site were not linked by any 4 nt repeated sequences. The fragment 14, without an EcoRI site, was used as the negative control.

Extension Reaction. The extension reaction used to prepare the fragments 1–14 was performed in 100 μL of hi-fi PCR mixture containing 1X pfu buffer, 0.15 mM dNTPs, 0.66 μM each of oligos 1–14 and oligo R, and 10 U pfu DNA polymerase (Sangon Biotech, China). This hi-fi PCR mixture was incubated at 95 °C for 3 min, 55 °C for 30 s, and 71 °C for 5 min. Then, the hi-fi PCR mixture of each DNA fragment was purified and resuspended using a universal DNA purification kit (Tiangen biotech, China) following the manufacturer’s instruction. The concentrations of the purified DNA fragments were 48–65 ng/μL, which were determined as mentioned above.

Digestion of the DNA Fragments 1–14 with EcoRI. Digestion Groups. It would be pretty hard for us to ensure that the digestion time was equal for every DNA fragment if all 14 DNA fragments were digested in the same batch of digestion reactions. Therefore, the DNA fragments 1–14 were divided into three digestion groups, namely, groups 1, 2, and 3. Group 1 contained the DNA fragments 1–4, 9, 10, 13, and 14. Group 2 included the DNA fragments 5–8, 13, and 14. Group 3 contained DNA the fragments 9–14. Every DNA fragment in the same group was digested in the same batch of digestion reactions.

Digestion Reaction. The volume of the digestion reaction mixture for each DNA fragment was 10 μL containing 1X EcoRI buffer, 50–100 ng of the purified DNA fragment, and 1–2.5 U EcoRI (Sangon Biotech, China). The digestion reaction mixture was incubated at 37 °C for 12–30 min, and then, the digestion reaction was stopped at 75 °C for 20 min.

Electrophoresis. The digestion products were checked using 3% agarose gel electrophoresis. Electrophoresis was run in 1× TAE buffer at 100 V for 20–30 min. DNA bands in the gels were analyzed as mentioned above.

The digestion reactions were repeated at least 10 times when the final concentration of EcoRI used was 1 U/10 μL.

Analyses of Integrated Density and Percentage of Grey Value. Integrated density (IntDen) of WB bands of GST-Tag and GAPDH proteins was analyzed using the software ImageJ (NIH, USA). The PGV of the digested DNA bands accounting for the total grey value of both digested and undigested DNA bands was analyzed using the software Image Lab (Bio-Rad, USA). The intensity of a WB band was represented with ΔIntDen. ΔIntDen = (IntDen of a WB band) − (IntDen of the background with the same area as the WB band).

Statistical Analysis. Ct values of qRT-PCR, the ratios of ΔIntDen of GST-Tag to that of GAPDH, and the PGV mentioned above were analyzed using one-way analysis of variance (One-way ANOVA) of software SPSS. The comparison of PGV was performed only for the DNA fragments in the same digestion group. P values less than 0.05 (LSD method) were deemed statistically significant.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.0c02264.

Primers of qRT-PCR and reverse transcription; schematic diagram showing the overlap PCR; three photos of agarose gel electrophoresis of DNA; four DNA sequencing results; two photos of agarose gel electrophoresis of total RNA; and schematic diagram showing the preparation of the DNA fragments 1–14 (PDF).

AUTHOR INFORMATION

Corresponding Author

Wei Li – Medical Scientific Research Center, Guangxi Medical University, Nanning 530021, Guangxi, China; orcid.org/0000-0001-7213-2211; Email: liwei60@yahoo.com

Authors

Lijuan Long – Department of Pediatrics, First Affiliated Hospital of Guangxi Medical University, Nanning 530021, Guangxi, China

Xinxin Li – Department of Nuclear Medicine, First Affiliated Hospital of Guangxi Medical University, Nanning 530021, Guangxi, China
Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c02264

Author Contributions
L.L. and X.L. contributed equally to this work. W.L. designed and supervised the study, prepared and digested DNA fragments 1−14, analyzed the experimental results, and wrote and edited the manuscript. L.L. reconstructed the plasmid pET-42a (+), performed qRT-PCR and WB, and edited a part of the manuscript. X.L. prepared and digested DNA fragments 1−14 and wrote and edited a part of the manuscript. H.W. prepared and digested DNA fragments 1−14.

Funding
This work is funded by the National Natural Science Foundation of China, Grant no. 31360218.

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Brosius, J. The fragmented gene. Ann. N. Y. Acad. Sci. 2009, 1178, 186–193.
(2) Kazazian, H. J., Jr.; Boehm, C. Molecular basis and prenatal diagnosis of beta-thalassemia. Blood 1988, 72, 1107–1116.
(3) Blackwood, E. M.; Kadonaga, J. T. Going the distance: a current view of enhancer action. Annu. Rev. Genomics Hum. Genet. 2006, 7, 29–59.
(4) de Koning, A. P. J.; Gu, W.; Castoe, T. A.; Batzer, M. A.; Pollock, D. N. Repetitive elements may comprise over two-thirds of the human genome. Nucleic Acids Res. 2010, 38, 6645–6646.
(5) Ohno, S. So much “junk” DNA in our genome. Evolution of Genetic Systems, Brookhaven Symp. Biol., 1972; Vol. 23, pp 366–370.
(6) Orgel, L. E.; Crick, F. H. C.; Sapienza, C. Selfish DNA. Nature 1980, 288, 645–646.
(7) Uslin, K.; House, N. C. M.; Freudenreich, C. H. Repeat instability during DNA repair: Insights from model systems. Crit. Rev. Biochem. Mol. Biol. 2015, 50, 142–167.
(8) Kim, S. D.; Fung, Y. S. C. An update on Huntington’s disease. Curr. Opin. Neurol. 2014, 27, 477–483.
(9) Cummings, C. J.; Zoghbi, H. Y. Fourteen and counting: unraveling trinucleotide repeat diseases. Hum. Mol. Genet. 2000, 9, 909–916.
(10) Mills, R. E.; Bennett, E. A.; Iskow, R. C.; Devine, S. E. Which transposable elements are active in the human genome? Trends Genet. 2007, 23, 183–191.
(11) Bourque, G.; Burns, K. H.; Gehring, M.; Gorbunova, V.; Seluanov, A.; Hammell, M.; Imbeault, M.; Izsák, Z.; Levin, H. L.; MacFarlan, T. S.; et al. Ten things you should know about LTR retrotransposons: for better or worse, in sickness and in health. Genome Res. 2008, 18, 343–358.
(12) Rabin, J. LINE-1 retroelements: a wonder macromolecule in biological realm. Int. J. Biol. Macromol. 2018, 118, 49–56.
(13) Kahl, T. R.; Fong, K. K.; Jordan, B.; Lek, J. C.; Levitan, R.; Patel, S. S. Kinetic mechanism of transcription initiation by bacteriophage T7 DNA polymerase: A wonder macromolecule in biological realm. Int. J. Biol. Macromol. 2018, 118, 49–56.
(14) Kazazian, H. H.; Goodier, J. L. LINE Drive. Trends Genet. 2002, 18, 744–750.
(15) Kapitonov, V. O.; Pavlicek, A.; Jurka, J. Anthology of Human Repetitive DNA. Encyclopedia of Molecular Cell Biology and Molecular Medicine. Wiley, 2006.
(16) Jablonov, M.; Gollubova, L. N.; Malyn, E. G. Effect on DNA transcription of nucleotide sequences upstream to T7 promoter. Nucleic Acids Res. 1996, 24, 3659–3660.
(17) Libri, M. D.; Brenowitz, M.; Willis, I. M. The TATA element and its context affect the cooperative interaction of TATA-binding protein with the TFIIB-related factor, TFIIBβ. Biopol. Chem. 1998, 273, 4563–4568.
(18) Barek-Samish, A.; Cohen, I.; Haran, T. E. Signals for TBP/TATA box recognition. J. Mol. Biol. 2000, 299, 965–977.
(19) Wohler, B. S.; Gralla, J. D. TATA-flanking sequences influence the rate and stability of TATA-binding protein and TFIIB binding. J. Biol. Chem. 2001, 276, 6260–6266.
(20) Tang, G.; Q.; Bandwar, R. P.; Patel, S. S. Extended upstream A−T sequence increases T7 promoter strength. J. Biol. Chem. 2005, 280, 40707–40713.
(21) Faiger, H.; Ivanchenko, M.; Cohen, I.; Haran, T. E. TBP flanking sequences: asymmetry of binding, long-range effects and consensus sequences. Nucleic Acids Res. 2006, 34, 104–119.
(22) Yella, V. R.; Bhihm, D. G.; Ghoshdastidar, D.; Rodriguez-Martinez, J. A.; Ansari, A. Z.; Bansal, M. Flexibility and structure of flanking DNA impact transcription factor affinity for its core motif. Nucleic Acids Res. 2018, 46, 11863–11879.
(23) Thomas, M.; Davis, R. W. Studies on the cleavage of bacteriophage lambda DNA with EcoRI restriction endonuclease. J. Mol. Biol. 1975, 91, 315–328.
(24) Halford, S. E.; Johnson, N. P.; Grinsted, J. The EcoRI restriction endonuclease with bacteriophage lambda DNA. Kinetic studies. Biochem. J. 1980, 191, 581–592.
(25) Berkner, K. L.; Folk, W. R. An assay for the rates of cleavage of specific sites in DNA by restriction endonucleases: Its use to study the cleavage of phage γ DNA by EcoRI and phage P22 DNA containing thymine or 5-bromouracil by HindIII. Anal. Biochem. 1983, 129, 446–456.
(26) Alves, J.; Pina, M. J.; Haupt, W.; Langowski, J.; Peters, F.; Mass, G.; Wolff, C. The influence of sequences adjacent to the recognition site on the cleavage of oligodeoxynucleotides by the EcoRI endonuclease. Eur. J. Biochem. 1984, 140, 83–92.
(27) Windolph, S.; Fritz, A.; Oelgeschläger, T.; Wolfs, H. Sealing of the cleavage site and DNA strand break by Escherichia coli endonuclease IV. J. Biol. Chem. 1984, 259, 11897–11898.
(28) Rosa, J.; Fernandez-Gonzalez, E.; Ducani, C.; Högberg, B. BsoCl and BseGI display sequence preference in the nucleotides flanking the recognition sequence. Plas. One 2018, 13, No. e020257.
(29) Tang, W.; Zhou, H.; Li, W. Silver and Cyanine Staining of Oligonucleotides in Polyacrylamide Gel. Plas. One 2015, 10, No. e014422.
(30) Schöne, S.; Jurt, M.; Helabad, M. B.; Dör, I.; Lebars, I.; Kieffer, B.; Imhof, P.; Rois, R.; Vingron, M.; Thomas-Chollier, M.; et al. Sequences flanking the core-binding site modulate glucocorticoid receptor structure and activity. Nat. Commun. 2016, DOI: 10.1038/ncomms12621.
(31) Sousa, R.; Chung, Y. J.; Rose, J. P.; Wang, B.-C. Crystal structure of bacteriophage T7 DNA polymerase at 3.3 Å resolution. Nature 1993, 364, 593–599.
(32) Cheetham, G. M.; Steitz, T. A. Structure of a transcribing T7 DNA polymerase initiation complex. Science 1999, 286, 2309–2309.
(33) Baus, Y.; Patel, S. S. Kinetic mechanism of transcription initiation by bacteriophage T7 DNA polymerase. Biochemistry 1997, 36, 4223–4232.
(34) Temiakov, D.; Mentesana, P. E.; Ma, K.; Mustaev, A.; Borukhov, S.; McAllister, W. T. The specificity loop of T7 DNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. Proc. Natl. Acad. Sci. U.S.A. 2000, 97, 14109–14114.
(35) Borkotoky, S.; Muráli, A. The highly efficient T7 DNA polymerase: A wonder macromolecule in biological realm. Int. J. Biol. Macromol. 2018, 118, 49–56.
(36) Khan, T. R.; Fong, K. K.; Jordan, B.; Lek, J. C.; Levitan, R.; Mitchell, P. S.; Wood, C.; Hatcher, M. E. An FTR investigation of flanking sequence effects on the structure and flexibility of DNA binding sites. Biochemistry 2009, 49, 1345–1357.
(38) Krüger, D. H.; Bickle, T. A. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol. Rev.* 1983, 47, 345–360.

(39) Youell, J.; Firman, K. Mechanistic insight into Type I restriction endonucleases. *Front. Biosci.* 2012, 17, 2122–2139.