Supplementary Information for:

DNA template sequence control of bacterial RNA polymerase escape from the promoter

Ewa Heyduk and Tomasz Heyduk*

Edward A. Doisy Department of Biochemistry and Molecular Biology, St. Louis University Medical School,

1100 S. Grand Blvd., St. Louis, MO 63104, USA.

* To whom correspondence should be addressed. Tel: (314) 977 9238; Fax (314) 977 9205; Email: tomasz.heyduk@health.slu.edu

SI Methods

Preparation of DNA template constructs

We used three methods to prepare DNA templates:

1. PCR amplification of the synthetic oligonucleotide template corresponding to full-length nontemplate strand of the desired duplex.

10-12 cycles of PCR with Klentaq LA PCR Kit (DNA Polymerase Technology, Saint Louis, MO) using 10 nM 155 nt synthetic template were performed (10-20x 20 μl reactions). The products of amplification were purified using Wizard SV Gel and PCR Cleanup Kit (Promega, Madison, WI). The following DNA constructs were prepared using this approach:

a. DNA constructs with +2 to +10 randomized. O238-O240 (Table S1d) were used as templates for PCR with O101/O105, O102/O105, O103/O105 (Table S1a) as primers, respectively.

2. PCR amplification of products of a ligation reaction of synthetic oligonucleotides corresponding to promoter (-75 to -1) and transcribed regions (+1 to +60) of the construct.

The oligonucleotide (or mix of oligonucleotides) corresponding +1 to +60 of the construct was phosphorylated at the 5' end using T4 kinase (Promega, Madison, WI) in the 1x kinase buffer in the presence of ATP and BSA. Phosphorylation was stopped with 20 mM EDTA after 3.5hrs incubation at 37C. Phosphorylated oligonucleotides were purified on SEP-Pak Plus C18 Cartridge (Waters). The cartridge was washed with 10 ml of methanol and 20 ml of water before the use. Phosphorylation reaction mix was loaded on the cartridge, the cartridge was washed with 20 ml of water and 10 ml of 5% methanol. Phosphorylated oligonucleotides were eluted with 3 ml of 60% methanol and were dried in Speed Vac. For the ligation reaction, a 1 μM mix of oligonucleotide corresponding to -75 to -1 of the construct and phosphorylated +1 to +60 oligonucleotide(s) were incubated with 10 units of T4 DNA ligase in 1x T4 RNA buffer in the presence of 1mM ATP and 25% PEG for 24 hrs at room temperature. The reaction was stopped with 10 mM Tris (pH 8.0) buffer containing 2.5 mM EDTA. Reaction mixture was run on 10% Urea/TBE gel, the band corresponding to the ligated product was cut and eluted from the gel. DNA was recovered by ethanol precipitation and used as a template for PCR amplification with Klentaq LA PCR Kit. The products of amplification were purified using Wizard SV Gel and PCR Cleanup Kit. The following DNA constructs were prepared using this approach:

a. DNA constructs containing 96 variants of ITS sequence in a context of 4 promoters. O97, O98, O99 and O100 (Table S1a) were ligated with a mix of O1-O96 (Table S1a). The products of ligation were PCR amplified using O101/O105, O102/O105, O103/O105 and O104/O105 (Table S1a) as primers, respectively.

b. DNA constructs for libraries containing all single base substitutions at +1 to +40 of deoB, λP_R, and UV5 ITS in a context of deoB, λP_R, and UV5 promoters. O97 was ligated with a mixture of O118-O157, O158-O197 and O198-O237 (Table S1c). The products of ligation were PCR amplified using O101/O105, O102/O105 and O104/O105 (Table S1a) as primers, respectively. O98 was ligated with a mixture of O118-O157, O158-O197 and O198-O237 (Table S1c). The

products of ligation were PCR amplified using O101/O105, O102/O105 and O104/O105 (Table S1a) as primers, respectively. O3 was ligated with a mixture of O118-O157, O158-O197 and O198-O237 (Table S1c). The products of ligation were PCR amplified using O101/O105, O102/O105 and O104/O105 (Table S1a) as primers, respectively.

c. DNA templates for "slow" and "fast" mutants used in experiments depicted in Fig. S5. O245 was ligated with O241,O242, O243 and O244 (Table S1d). The products of ligation were PCR amplified using O103/O105 primers. O246 was ligated with O241, O242, O243 and O244 (Table S1d). The products of ligation were PCR amplified using O102/O105 primers.

3. *Extension by PCR of partial duplexes obtained by hybridizing appropriate synthetic oligonucleotides containing complementary overlapping sequences at their 3' ends.*

Partial duplexes were obtained by mixing 1 μM of -75 to -1 and 1.1 μM of -20 to +80 oligonucleotides in 250 μl (5x50 μl). The pre-hybridized partial duplexes were subjected to 2-3 cycles of PCR using Red AccuTaq LA DNA Polymerase (Sigma). Extension to full duplex DNA was checked on the 2% agarose gel. Extended DNA template was purified by FPLC chromatography on 1ml Resource Q column using 25 ml of 0.45M-0.9 M NaCl gradient in 25 mM Tris (pH 8.0) buffer. DNA was recovered from peak fractions by ethanol precipitation. The following constructs were prepared using this approach:

a. DNA templates with Cy3 probe at -4 position. O106 was labeled with Cy3 and HPLC purified as described previously (1). O106(Cy3)/O107, O106(Cy3)/O108, O106(Cy3)/O109, O106(Cy3)/O110 and O106(Cy3)/O111 partial duplexes were prepared, extended and purified as described above. O112 was labeled with Cy3 and HPLC purified as described previously (1). O112(Cy3)/O113, O112(Cy3)/O114, O112(Cy3)/O115, O112(Cy3)/O116 and O112(Cy3)/O117 partial duplexes were prepared, extended and purified as described above.

Time-dependent formation of full-length transcript measured using molecular beacon assay

Reaction mixture containing 150 nM promoter DNA and 225 nM RNAP holoenzyme in transcription buffer was incubated at room temperature for 5 min to allow open complex formation. Transcription was initiated by adding NTPs (100 μM) and heparin (200 mg/ml). At different time intervals, 18 μl samples of reaction mixture were withdrawn and reaction was stopped by adding 1 μl of 0.5 M EDTA followed by addition of 1 μl of 100 nM molecular beacon (O251, Table S1e) with recognition sequence complementary to target near the 5' end of full-length transcript. Fluorescence of the beacon (excitation at 490 nm and emission at 530 nm) was read after 30 min incubation on SpectraFluor Plus microplate plate reader (Tecan). Oligonucleotide complementary to the beacon (O252, Table S1e) was used to calibrate beacon signal. Fluorescence intensities measured at different time points were fitted by nonlinear regression to obtain rate constants.

Supplementary Tables

Table S1a. Oligonucleotides used to prepare libraries of DNA constructs containing 96 variants of ITS sequence. If there were multiple promoters associated with a gene, the identity of the promoter is identified by the index at the end of gene name. ITS sequences of σ70 promoters were obtained from EcoCyc database (https://ecocyc.org).

Name	Sequence	Description
O1	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	+1 to +60 of deoB
O2	AATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	+1 to +60 of UV5
O3	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	+1 to +60 of λP_R
O4	AATTCAATAAGTGGCGTTCGCCATGCGAGGATAAAATGTCCTGGACTCATTCGGCATCGG	+1 to +60 of argF
O5	TGCGGTACTGGGCATTTACCCTACAAACTGCTGTCTCACACTGGACTCATTCGGCATCGG	+1 to +60 of acnB
O6	ACACTTTAAACGCCACCAGATCCCGTGGAATTGAGGTCGTCTGGACTCATTCGGCATCGG	+1 to +60 of cyoA
O7	GTTTTCCTCCAGCGGGTTTAACAGGAGTCCTCGCATGAAACTGGACTCATTCGGCATCGG	+1 to +60 of nfo
O8	CATATCAAATGGTTAATTTTTGCACAGGACTGGTGGGTTTCTGGACTCATTCGGCATCGG	+1 to +60 of rfe
O9	ATAACCAGCATTCGGAGTCAACAGTGGTACGGCGTTTAACCTGGACTCATTCGGCATCGG	+1 to +60 of cho
O10	CTGTTGCATTATTCGCCTGAAACCACAATATTCAGGCGTTCTGGACTCATTCGGCATCGG	+1 to +60 of proVp1
O11	CGTTTTTTCGCTATCTTTGACAAAAAATATCAACTTTCTCCTGGACTCATTCGGCATCGG	+1 to +60 of proVp2
O12	AGTGACTATTTCCATTGGGTAATATATCGACATAGACAAACTGGACTCATTCGGCATCGG	+1 to +60 of proVp3
O13	CTCATTGTGTTTTATTTCTCACATTGATGACGGTCGCATGCTGGACTCATTCGGCATCGG	+1 to +60 of yfiRp5
O14	GTAGCCAACAAACAATGCTTTATGAATCCTCCCAGGAGACCTGGACTCATTCGGCATCGG	+1 to +60 of pstS
O15	TCGTCAACACGGCACGCTACTTAAGAAAGCCGTAATAAATCTGGACTCATTCGGCATCGG	+1 to +60 of spy
O16	GTTCACGTAACTGGAGGAATGAAATGGAGTTTTTCAAAAACTGGACTCATTCGGCATCGG	+1 to +60 of ansBp2
O17	ATCTCCGGCAATATTGCCCCTTTGAAGGCTGGCGAATAAGCTGGACTCATTCGGCATCGG	+1 to +60 of parCp4
O18	ATATCCCCTTAAGCGGATAAACTTGCTGTGGACGTATGACCTGGACTCATTCGGCATCGG	+1 to +60 of hipB
O19	GATATTTTCGATAGGTTTGGGGTTATGGATCTGCGCCGTTCTGGACTCATTCGGCATCGG	+1 to +60 of yfiEp7
O20	AAAATCACGCTTCAGGGCTAAGTAAAACGCTGTCTCTGCCCTGGACTCATTCGGCATCGG	+1 to +60 of cvrAp6
O21	CCTACAAGGAGAACAAAAGCATGAGCCAAATTCACAAACACTGGACTCATTCGGCATCGG	+1 to +60 of acsp2
O22	TTCGGGTGAACAGAGTGCTAACAAAATGTTGCCGAACAACCTGGACTCATTCGGCATCGG	+1 to +60 of rpoS
O23	CGGATGCAAATCCGCACACAACATTTCAAAAGACAGGATTCTGGACTCATTCGGCATCGG	+1 to +60 ofserAp1
O24	GATAACTCATATAACGCAGGGCTGTTTATCGTGAATTCACCTGGACTCATTCGGCATCGG	+1 to +60 of fliAp1
O25	GCTTCAGTGGGAAAGATTAAAAACTCCCGCTTTATTGGTTCTGGACTCATTCGGCATCGG	+1 to +60 of secG
O26	ATATCAATTACGGCTTGAGCAGACCTATGATCCCGGAAAACTGGACTCATTCGGCATCGG	+1 to +60 of fnr
O27	GCTACATGGGTGCTAAATCTTTAACGATAACGCCATTGAGCTGGACTCATTCGGCATCGG	+1 to +60 of ssrAp2
O28	ACACATTGGGGCTGATTCTGGATTCGACGGGATTTGCGAACTGGACTCATTCGGCATCGG	+1 to +60 of ssrAp
O29	GTATCCCAATTTCATACAGTTAAGGACAGGCCATGAGTAACTGGACTCATTCGGCATCGG	+1 to +60 of cysKp1
O30	TGTCCATATCATAAATAAGTTACGTATTTTTTCTCAAGCACTGGACTCATTCGGCATCGG	+1 to +60 of fimA
O31	GACGTACGAAACGTCAGCGGTCAACACCCGCCAGCAATGGCTGGACTCATTCGGCATCGG	+1 to +60 of ptsHp2
O32	ACCTGTGGTCGCAATCGATTGACCGCGGGTTAATAGCAACCTGGACTCATTCGGCATCGG	+1 to +60 of serC
O33	TAAGCAAACGGTGATTTTGAAAAATGGGTAACAACGTCGTCTGGACTCATTCGGCATCGG	+1 to +60 of purA
O34	AATATTTGTTGATAAGGATAGTAACATGAACATGACAAAACTGGACTCATTCGGCATCGG	+1 to +60 of slp
O35	ACACTTCCACTGTGTGTGACTGTCTGGTCTGACTGAGACACTGGACTCATTCGGCATCGG	+1 to +60 of frrp3
O36	ACTTCCACTGTGTGTGACTGTCTGGTCTGACTGAGACAAGCTGGACTCATTCGGCATCGG	+1 to +60 of frrp
O37	TGCCCGGATAGCTCAGTCGGTAGAGCAGGGGATTGAAAATCTGGACTCATTCGGCATCGG	+1 to +60 of pheV
O38	ATGTATTAACATGCTGAGTTTATACGAAAAGATAAAGATACTGGACTCATTCGGCATCGG	+1 to +60 of yliEp6
O39	GTTGGAACACGATGTTGCACAGGCTGTGGTGTTAGGCCTGCTGGACTCATTCGGCATCGG	+1 to +60 of yehLp3
O40	GTTACAGCGACTTACAAAATTAAGTCATTCGCTCACTGACCTGGACTCATTCGGCATCGG	+1 to +60 of yhflp5
O41	GGGAATGTAAATGAACAAATTTATTAAAGTTGCACTGGTACTGGACTCATTCGGCATCGG	+1 to +60 of yhflp4
O42	ATGAGAGATGATGCTTAACTTATCGTATTTAATCTATTGTCTGGACTCATTCGGCATCGG	+1 to +60 of ydeP
O43	GGTGCCACATGCAGCGCCTGCTAATGTTGTAAGCTTGTAGCTGGACTCATTCGGCATCGG	+1 to +60 of yneEp7
O44	ATAACAACGCAGTCGCACTATTTTTCACTGGAGAGAAGCCCTGGACTCATTCGGCATCGG	+1 to +60 of mtrp2
O45	ATGAAAATCCTCAGTAAGCTGCCCGCCCTTTTTTACACTTCTGGACTCATTCGGCATCGG	+1 to +60 of yiaYp4
O46	TTTTATTAGATTAATTTTGATAATCAATAATACTGATCATCTGGACTCATTCGGCATCGG	+1 to +60 of yiaYp6
O47	ATCCACAATTTTAATATGGCCTTGTTTAATTGCTTCAAAACTGGACTCATTCGGCATCGG	+1 to +60 of caiF
O48	GTTTTTATTTTTTAATGTATTTGTACATGGAGAAAATAAACTGGACTCATTCGGCATCGG	+1 to +60 of phoA
O49	AATCAATAGCTAAAATAAGTAACATCAAAAATAACGCGACCTGGACTCATTCGGCATCGG	+1 to +60 of cspEp1
O50	TTTCTGGTTCCTTTTTTAACCATATAGCAAGGGATCGCCACTGGACTCATTCGGCATCGG	+1 to +60 of smtA
O51	AAAGGGATCTACAACCTACAGATTGGTGTAGCTTTATGGACTGGACTCATTCGGCATCGG	+1 to +60 of ycdT
O52	ATAACGAAACGCACTGCCGGACAGACAAATGAACTTATCCCTGGACTCATTCGGCATCGG	+1 to +60 of ydeEp1
O53	GTTACTTAATTTAAGTGACGATCGCTAAAAACGACTGTCACTGGACTCATTCGGCATCGG	+1 to +60 of pykAp2
O54	TCATGACTACTGCAAGACTAAAATTAACATGACAAGTCTGCTGGACTCATTCGGCATCGG	+1 to +60 of znuC

O55	ACAACACAACAAGGAGCCACGCATGGAAATGCTCGAAGAGCTGGACTCATTCGGCATCGG	+1 to +60 of yeiGp6
O56	ATACACCCGGCCCCTCGCCGGGTTTTTTGTGATCTGCGTCCTGGACTCATTCGGCATCGG	+1 to +60 of yffSp4
O57	GCATTCTGAAAAAATATCGATGCATTTCGAGCGAAGATGGCTGGACTCATTCGGCATCGG	+1 to +60 of yfiRp3
O58	GCAATCAGATCAACGGCCTGTAATACCTCTAACGCACGCTCTGGACTCATTCGGCATCGG	+1 to +60 of lpoAp2
O59	AAATAATCTTTCTAACTCCTGAACACATCTCTGGAGATTTCTGGACTCATTCGGCATCGG	+1 to +60 of def
O60	GTTCACGCGGCATAATCTCCCGCCACGGAACCCGTGGCAACTGGACTCATTCGGCATCGG	+1 to +60 of yiaDp3
O61	ATCTGTATATATACCCAGCTTTTTGGCGGAGGGCGTTGCGCTGGACTCATTCGGCATCGG	+1 to +60 of uvrD
O62	ACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCGCTGGACTCATTCGGCATCGG	+1 to +60 of rrsB
O63	AAAAGACGAGTTTTTACGGGCGTATTTAAAGTGATAATCACTGGACTCATTCGGCATCGG	+1 to +60 of lptFp5
O64	GTGATCACCCGGTTCGCGGTTATTTGATCAAGAAGAGTGGCTGGACTCATTCGGCATCGG	+1 to +60 of yjtDp6
O65	GGTCGCGTGCGCAAATACGCTTTTCCTCACACAGTTGTCACTGGACTCATTCGGCATCGG	+1 to +60 of dksAp3
O66	ACTTTTTTTTGTCCCAGGCTCGCCTTGAGCCTGTTCTACCCTGGACTCATTCGGCATCGG	+1 to +60 of proS
O67	CTATTTGCATTTGCAATAGCGTAATGGCGCGCCGTGGGAACTGGACTCATTCGGCATCGG	+1 to +60 of fepA
O68	TGAAGTCGGTATTTCACCTAAGATTAACTTATGTAACAGTCTGGACTCATTCGGCATCGG	+1 to +60 of gltA
O69	ACATCGCCAGGGGTGCTCGGCATAAGCCGAAGATATCGGTCTGGACTCATTCGGCATCGG	+1 to +60 of ompA
O70	ATCGGCCCGGTCGAATTTTCTTTGCTGTAGAAAAAATAACCTGGACTCATTCGGCATCGG	+1 to +60 of rihB
O71	ATCAAAATGTGAATTGTAGCTGACCTGGGACTTGTACCCGCTGGACTCATTCGGCATCGG	+1 to +60 of greA
O72	AGTGAATGTATCTTAGGTAAATAATATATATTATTTATTTCTGGACTCATTCGGCATCGG	+1 to +60 of envR
O73	GTTGCCTGAAGCGCTGGATGCTGTCGGAGCTTTCTCCACACTGGACTCATTCGGCATCGG	+1 to +60 of damp2
O74	CCGATAACAGTTACCCGTAACATTTTTAATTCTTGTATTGCTGGACTCATTCGGCATCGG	+1 to +60 of gntK
O75	ATGAAAATCCTCAGTAAGCTGCCCGCCCTTTTTTACACTTCTGGACTCATTCGGCATCGG	+1 to +60 of yiaYp4
O76	GCAAGGGCGGCTGACAGAGTAAAACGTAATGGATGACTGTCTGGACTCATTCGGCATCGG	+1 to +60 of hypTp4
O77	ATTGTCGCTGGCGTTCAGGCCATGCAGGATTCTGGCCTTGCTGGACTCATTCGGCATCGG	+1 to +60 of rnd1
O78	TGATTCTGCCAGCGTATTAATAAACAATCGCTGAAACAGACTGGACTCATTCGGCATCGG	+1 to +60 of rnd2
O79	AGGCCATCAGAAATAACGACCTGTACGTCCGGATTGGCAACTGGACTCATTCGGCATCGG	+1 to +60 of rnd3
O80	GATGCGAGAAACACATTACAATGCGGGTATATTCGTTTGGCTGGACTCATTCGGCATCGG	+1 to +60 of rnd4
O81	TTCGCTGACATCATCCTGGTGGCTGACCACATCACTGAAACTGGACTCATTCGGCATCGG	+1 to +60 of rnd5
O82	CTGCTGTTAGTGATATTTAAAAGTGTGACTGGCGAACCATCTGGACTCATTCGGCATCGG	+1 to +60 of rnd6
O83	GGTACAGGCTCGGCGCTTTATAAGCACGGGCGATGCCCATCTGGACTCATTCGGCATCGG	+1 to +60 of rnd7
O84	GAAGAAGGCACCCGGGTCAGCATTAATAATGCCAGCGCGGCTGGACTCATTCGGCATCGG	+1 to +60 of rnd8
O85	CAGCAGTGATGCCGGGTGGATCAAGCCTTCATCCAGCGCCCTGGACTCATTCGGCATCGG	+1 to +60 of rnd9
O86	TGCGGATTAGTTTTTGCGCGTAAATGTTGCAATGTAATATCTGGACTCATTCGGCATCGG	+1 to +60 of rnd10
O87	CGCGGCAAGTGGCGACCGAGTTAATATTTGCGTAGCGAAACTGGACTCATTCGGCATCGG	+1 to +60 of rnd11
O88	TGTGTTGAGGATCACAAAACGAATAATTGCTGATCGCCGCCTGGACTCATTCGGCATCGG	+1 to +60 of rnd12
O89	CAAACTTCAGCATCATGCCAGGGTAGAGTGTGTCTAATCGCTGGACTCATTCGGCATCGG	+1 to +60 of rnd13
O90	ATGGTGAGCATCAGGCAGGGATCCTGACGCCACAACAGGCCTGGACTCATTCGGCATCGG	+1 to +60 of rnd14
O91	CGGAGAAGCCCCGTCGCGCGCCATCACCACGATGTACGGGCTGGACTCATTCGGCATCGG	+1 to +60 of rnd15
O92	CGCACCGATGCAAAAGCTCTCTGGTGGCGAAACGCAGCGTCTGGACTCATTCGGCATCGG	+1 to +60 of rnd16
O93	GGCACGAGGATCAATATTGCCTGCTGCCCCCTGCAAGGCCCTGGACTCATTCGGCATCGG	+1 to +60 of rnd17
O94	CGTTTTTCTCGATCGCGTGCAGAATTTCACCCAGTTCGTTCTGGACTCATTCGGCATCGG	+1 to +60 of rnd18
O95	CGGTGCTTTGTCGCCGGGAATTATTCTGACGGCGGTGATTCTGGACTCATTCGGCATCGG	+1 to +60 of rnd19
O96	GCGAGTTGCTCATATTCATCTGGCAAATGAGAATAATCCACTGGACTCATTCGGCATCGG	+1 to +60 of rnd20
O97	AAACTTGCCGATAAAGCACCGGAAAGCACACCAACTGTCTATCGCCGTATCAGCGAATAA CGGTATACTGATCGT	-75 to -1 of deoB
O98	AATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCT CGTATAATGTGTGG	-75 to -1 of UV5
O99	ATCTATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTGGCG GTGATAATGGTTGC	-75 to -1 of λP_R
O100	CGATTCAGCCACTTTTTTATGTTGCTTTTTTGTAAACAGATTAACACCTCGTCAAAATCCTG CTATTCTGCCCGT	-75 to -1 of acnB
O101	AAACTTGCCGATAAAGCACCG	upstream primer for deoB promoter
O102	AATTAATGTGAGTTAGCTCAC	upstream primer for UV5 promoter
O103	ATCTA**T**CACCGCAAG**G**GATAA	upstream primer for λP_R promoter
O104	CGATTCAGCCACTTTTTTATG	upstream primer for acnB promoter
O105	CCGATGCCGAATGAGTCCAG	universal downstream primer

Table S1b. Oligonucleotides used to prepare DNA constructs with Cy3 probe at -4 of the nontemplate strand.

Name	Sequence	Description
O106	ATCTATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCT CTGGCGGTGATAATGG(aminodT)TGC	-75 to -1 of λP_R with aminodT at -4
O107	TCCAGTGCTGACTGCTTAATCGCTTCCAGGGTTATGCGTTGTTCCATACAACCTCCTT AGTACATGCAACCATTATCACCGCCAG	-20 to +80 of λP_R with overlap sequence (-20 to -1) for λP_R
O108	TCCAGTGCTGACTGCTTAATCGCTTCCATATGTTCTCCGTACTCAGTGCTTCAAATTT AAATGATGCAACCATTATCACCGCCAG	-20 to +80 of deoB with overlap sequence (-20 to -1) for λP_R
O109	TCCAGTGCTGACTGCTTAATCGCTTCTAGCTGTTTCCTGTGTGAAATTGTTATCCGCT CACAATTGCAACCATTATCACCGCCAG	-20 to +80 of UV5 with overlap sequence (-20 to -1) for λP_R
O110	TCCAGTGCTGACTGCTTAATCGCTTCCAATAGATTAAATACGATAAGTTAAGCATCAT CTCTCATGCAACCATTATCACCGCCAG	-20 to +80 of ydep with overlap sequence (-20 to -1) for λP_R
O111	TCCAGTGCTGACTGCTTAATCGCTTCGCAACGCCCTCCGCCAAAAAGCTGGGTATA TATACAGATGCAACCATTATCACCGCCAG	-20 to +80 of uvrD with overlap sequence (-20 to -1) for λP_R
O112	AAACTTGCCGATAAAGCACCGGAAAGCACACCAACTGTCTATCGCCGTATCAGCG AATAACGGTATACTGA(aminodT)CGT	-75 to -1 of deoB with aminodT at -4
O113	TCCAGTGCTGACTGCTTAATCGCTTCCAGGGTTATGCGTTGTTCCATACAACCTCC TTAGTACATACGATCAGTATACCGTTATT	-20 to +80 of λP_R with overlap sequence (-20 to -1) for deoB
O114	CCAGTGCTGACTGCTTAATCGCTTCCATATGTTCTCCGTACTCAGTGCTTCAAATT TAAATGATACGATCAGTATACCGTTATT	-20 to +80 of deoB with overlap sequence (-20 to -1) for deoB
O115	TCCAGTGCTGACTGCTTAATCGCTTCTAGCTGTTTCCTGTGTGAAATTGTTATCCG CTCACAATTACGATCAGTATACCGTTATT	-20 to +80 of UV5 with overlap sequence (-20 to -1) for deoB
O116	TCCAGTGCTGACTGCTTAATCGCTTCCAATAGATTAAATACGATAAGTTAAGCATCA TCTCTCATACGATCAGTATACCGTTATT	-20 to +80 of ydep with overlap sequence (-20 to -1) for deoB
O117	TCCAGTGCTGACTGCTTAATCGCTTCGCAACGCCCTCCGCCAAAAAGCTGGGTATA TATACAGATACGATCAGTATACCGTTATT	-20 to +80 of uvrD with overlap sequence (-20 to -1) for deoB

Table S1c. Oligonucleotides (+1 to +60) used to prepare DNA constructs for libraries containing all single base substitutions at +1 to +40 of deoB, λP_R, and UV5 ITS.

Name	Sequence	Description
O118	BTCATTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +1 mutants
O119	AVCATTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +2 mutants
O120	ATDATTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +3 mutants
O121	ATCBTTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +4 mutants
O122	ATCAVTTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +5 mutants
O123	ATCATVTAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +6 mutants
O124	ATCATTVAAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +7 mutants
O125	ATCATTTBAATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +8 mutants
O126	ATCATTTABATTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +9 mutants
O127	ATCATTTAABTTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +10 mutants
O128	ATCATTTAAAVTTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +11 mutants
O129	ATCATTTAAATVTGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +12 mutants
O130	ATCATTTAAATTVGAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +13 mutants
O131	ATCATTTAAATTTHAAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +14 mutants
O132	ATCATTTAAATTTGBAGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +15 mutants
O133	ATCATTTAAATTTGABGCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +16 mutants
O134	ATCATTTAAATTTGAAHCACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +17 mutants
O135	ATCATTTAAATTTGAAGDACTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +18 mutants
O136	ATCATTTAAATTTGAAGCBCTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +19 mutants
O137	ATCATTTAAATTTGAAGCADTGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +20 mutants

O138	ATCATTTAAATTTGAAGCACVGAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +21 mutants
O139	ATCATTTAAATTTGAAGCACTHAGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +22 mutants
O140	ATCATTTAAATTTGAAGCACTGBGTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +23 mutants
O141	ATCATTTAAATTTGAAGCACTGAHTACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +24 mutants
O142	ATCATTTAAATTTGAAGCACTGAGVACGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +25 mutants
O143	ATCATTTAAATTTGAAGCACTGAGTBCGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +26 mutants
O144	ATCATTTAAATTTGAAGCACTGAGTADGGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +27 mutants
O145	ATCATTTAAATTTGAAGCACTGAGTACHGAGAACATATGACTGGACTCATTCGGCATCGG	deoB +28 mutants
O146	ATCATTTAAATTTGAAGCACTGAGTACGHAGAACATATGACTGGACTCATTCGGCATCGG	deoB +29 mutants
O147	ATCATTTAAATTTGAAGCACTGAGTACGGBGAACATATGACTGGACTCATTCGGCATCGG	deoB +30 mutants
O148	ATCATTTAAATTTGAAGCACTGAGTACGGAHAACATATGACTGGACTCATTCGGCATCGG	deoB +31 mutants
O149	ATCATTTAAATTTGAAGCACTGAGTACGGAGBACATATGACTGGACTCATTCGGCATCGG	deoB +32 mutants
O150	ATCATTTAAATTTGAAGCACTGAGTACGGAGABCATATGACTGGACTCATTCGGCATCGG	deoB +33 mutants
O151	ATCATTTAAATTTGAAGCACTGAGTACGGAGAADATATGACTGGACTCATTCGGCATCGG	deoB +34 mutants
O152	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACBTATGACTGGACTCATTCGGCATCGG	deoB +35 mutants
O153	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACAVATGACTGGACTCATTCGGCATCGG	deoB +36 mutants
O154	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACATBTGACTGGACTCATTCGGCATCGG	deoB +37 mutants
O155	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACATAVGACTGGACTCATTCGGCATCGG	deoB +38 mutants
O156	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACATATHACTGGACTCATTCGGCATCGG	deoB +39 mutants
O157	ATCATTTAAATTTGAAGCACTGAGTACGGAGAACATATGBCTGGACTCATTCGGCATCGG	deoB +40 mutants
O158	BTGTACTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +1 mutants
O159	AVGTACTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +2 mutants
O160	ATHTACTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +3 mutants
O161	ATGVACTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +4 mutants
O162	ATGABCTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +5 mutants
O163	ATGTADTAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +6 mutants
O164	ATGTACVAAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +7 mutants
O165	ATGTACTBAGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +8 mutants
O166	ATGTACTABGGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +9 mutants
O167	ATGTACTAAHGAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +10 mutants
O168	ATGTACTAAGHAGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +11 mutants
O169	ATGTACTAAGGBGGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +12 mutants
O170	ATGTACTAAGGAHGTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +13 mutants
O171	ATGTACTAAGGAGHTTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +14 mutants
O172	ATGTACTAAGGAGGVTGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +15 mutants
O173	ATGTACTAAGGAGGTVGTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +16 mutants
O174	ATGTACTAAGGAGGTTHTATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +17 mutants
O175	ATGTACTAAGGAGGTTGVATGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +18 mutants
O176	ATGTACTAAGGAGGTTGTBTGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +19 mutants
O177	ATGTACTAAGGAGGTTGTAVGGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +20 mutants
O178	ATGTACTAAGGAGGTTGTATHGAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +21 mutants
O179	ATGTACTAAGGAGGTTGTATGHAACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +22 mutants
O180	ATGTACTAAGGAGGTTGTATGGBACAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +23 mutants
O181	ATGTACTAAGGAGGTTGTATGGABCAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +24 mutants
O182	ATGTACTAAGGAGGTTGTATGGAADAACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +25 mutants
O183	ATGTACTAAGGAGGTTGTATGGAACBACGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +26 mutants
O184	ATGTACTAAGGAGGTTGTATGGAACABCGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +27 mutants
O185	ATGTACTAAGGAGGTTGTATGGAACAADGCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +28 mutants
O186	ATGTACTAAGGAGGTTGTATGGAACAACHCATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +29 mutants
O187	ATGTACTAAGGAGGTTGTATGGAACAACGDATAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +30 mutants
O188	ATGTACTAAGGAGGTTGTATGGAACAACGCBTAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +31 mutants
O189	ATGTACTAAGGAGGTTGTATGGAACAACGCAVAACCCTGGCTGGACTCATTCGGCATCGG	λP_R +32 mutants
O190	ATGTACTAAGGAGGTTGTATGGAACAACGCATBACCCTGGCTGGACTCATTCGGCATCGG	λP_R +33 mutants
O191	ATGTACTAAGGAGGTTGTATGGAACAACGCATABCCCTGGCTGGACTCATTCGGCATCGG	λP_R +34 mutants
O192	ATGTACTAAGGAGGTTGTATGGAACAACGCATAADCCTGGCTGGACTCATTCGGCATCGG	λP_R +35 mutants
O193	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACDCTGGCTGGACTCATTCGGCATCGG	λP_R +36 mutants
O194	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACCDTGGCTGGACTCATTCGGCATCGG	λP_R +37 mutants
O195	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACCCVGGCTGGACTCATTCGGCATCGG	λP_R +38 mutants
O196	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACCCTHGCTGGACTCATTCGGCATCGG	λP_R +39 mutants
O197	ATGTACTAAGGAGGTTGTATGGAACAACGCATAACCCTGHCTGGACTCATTCGGCATCGG	λP_R +40 mutants
O198	BATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +1 mutants
O199	ABTTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +2 mutants
O200	AAVTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +3 mutants
O201	AATVGTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +4 mutants

O202	AATTHTGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +5 mutants
O203	AATTGVGAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +6 mutants
O204	AATTGTHAGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +7 mutants
O205	AATTGTGBGCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +8 mutants
O206	AATTGTGAHCGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +9 mutants
O207	AATTGTGAGDGGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +10 mutants
O208	AATTGTGAGCHGATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +11 mutants
O209	AATTGTGAGCGHATAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +12 mutants
O210	AATTGTGAGCGGBTAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +13 mutants
O211	AATTGTGAGCGGAVAACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +14 mutants
O212	AATTGTGAGCGGATBACAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +15 mutants
O213	AATTGTGAGCGGATABCAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +16 mutants
O214	AATTGTGAGCGGATAADAATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +17 mutants
O215	AATTGTGAGCGGATAACBATTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +18 mutants
O216	AATTGTGAGCGGATAACABTTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +19 mutants
O217	AATTGTGAGCGGATAACAAVTTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +20 mutants
O218	AATTGTGAGCGGATAACAATVTCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +21 mutants
O219	AATTGTGAGCGGATAACAATTVCACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +22 mutants
O220	AATTGTGAGCGGATAACAATTTDACACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +23 mutants
O221	AATTGTGAGCGGATAACAATTTCBCACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +24 mutants
O222	AATTGTGAGCGGATAACAATTTCADACAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +25 mutants
O223	AATTGTGAGCGGATAACAATTTCACBCAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +26 mutants
O224	AATTGTGAGCGGATAACAATTTCACADAGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +27 mutants
O225	AATTGTGAGCGGATAACAATTTCACACBGGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +28 mutants
O226	AATTGTGAGCGGATAACAATTTCACACAHGAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +29 mutants
O227	AATTGTGAGCGGATAACAATTTCACACAGHAAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +30 mutants
O228	AATTGTGAGCGGATAACAATTTCACACAGGBAACAGCTATCTGGACTCATTCGGCATCGG	UV5 +31 mutants
O229	AATTGTGAGCGGATAACAATTTCACACAGGABACAGCTATCTGGACTCATTCGGCATCGG	UV5 +32 mutants
O230	AATTGTGAGCGGATAACAATTTCACACAGGAABCAGCTATCTGGACTCATTCGGCATCGG	UV5 +33 mutants
O231	AATTGTGAGCGGATAACAATTTCACACAGGAAADAGCTATCTGGACTCATTCGGCATCGG	UV5 +34 mutants
O232	AATTGTGAGCGGATAACAATTTCACACAGGAAACBGCTATCTGGACTCATTCGGCATCGG	UV5 +35 mutants
O233	AATTGTGAGCGGATAACAATTTCACACAGGAAACAHCTATCTGGACTCATTCGGCATCGG	UV5 +36 mutants
O234	AATTGTGAGCGGATAACAATTTCACACAGGAAACAGDTATCTGGACTCATTCGGCATCGG	UV5 +37 mutants
O235	AATTGTGAGCGGATAACAATTTCACACAGGAAACAGCVATCTGGACTCATTCGGCATCGG	UV5 +38 mutants
O236	AATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTBTCTGGACTCATTCGGCATCGG	UV5 +39 mutants
O237	AATTGTGAGCGGATAACAATTTCACACAGGAAACAGCTAVCTGGACTCATTCGGCATCGG	UV5 +40 mutants

Table S1d. Oligonucleotides used to prepare constructs with randomized +2 to +10.

Name	Sequence	Description
O238	AAACTTGCCGATAAAGCACCGGAAAGCACACCAACTGTCTATCGCCGTATCAGCGAA TAACGGTATACTGATCGTANNNNNNNNNTTTGAAGCACTGAGTACGGAGAACATATG GAAGCGATTAAGCAGTCAGCA CTGGACTCATTCGGCATCGG	deoB with +2 to +10 randomized
O239	AATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCC GGCTCGTATAATGTGTGGANNNNNNNNNGGATAACAATTTCACACAGGAAACAGCTA GAAGCGATTAAGCAGTCAGCA CTGGACTCATTCGGCATCGG	UV5 with +2 to +10 randomized
O240	ATCTATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTG GCGGTGATAATGGTTGCANNNNNNNNNGAGGTTGTATGGAACAACGCATAACCCTGG AAGCGATTAAGCAGTCAGCA CTGGACTCATTCGGCATCGG	λP_R with +2 to +10 randomized

Table S1d. Oligonucleotides used to prepare “slow” and “fast” mutants for experiments depicted in Fig. S5.

Name	Sequence	Description
O241	5’ Phosphate-CCTTTCCCAGAGGTTGTATGGAACAACGCATAACCCTGGAAGCG ATTAAGCAGTCAGCACTGGACTCATTCGGCATCGG	+2 to +80 of “slow”
O242	5’ Phosphate-TAAACAAAC GAGGTTGTATGGAACAACGCATAACCCTGGAAGCG ATTAAGCAGTCAGCACTGGACTCATTCGGCATCGG	+2 to +80 of “fast”
O243	5’ Phosphate-TCTTTCCCA GAGGTTGTATGGAACAACGCATAACCCTGGAAGCG ATTAAGCAGTCAGCA CTGGACTCATTCGGCATCGG	+2 to +80 of C to T“slow”
O244	5’ Phosphate-GAAACAAACGAGGTTGTATGGAACAACGCATAACCCTGGAAGCG ATTAAGCAGTCAGCA CTGGACTCATTCGGCATCGG	+2 to +80 of T to G “fast”

O245	ATCTATCACCGCAAGGGATAAATATCTAACACCGTGCGTGTTGACTATTTTACCTCTGGCGGTGATAATGGTTGCA	-75 to +1 of λP_R
O246	AATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGA	-75 to +1 of UV5

Table S1e. Oligonucleotides used to prepare NGS sequencing libraries.

Name	Sequence	Description
O247	UCGAUAAGGGUAGGUGAGG	5' RNA adapter
O248	TAAGCGTGACAGTCTCCGATGCCGAATGAGTCCAG	Primer for reverse transcriptase
O249	CCTCTCTATGGGCAGTCGGTGATTAAGCGTGACAGTCTCCG ATG	Downstream Ion Torrent sequencing primer
O250	CCATCTCATCCCTGCGTGTCTCCGACTCAGXXXXXXXXXXTCGATAAGGGTAGGTGAGG	Upstream Ion Torrent indexing primer (X's correspond to one of Ion Torrent barcode sequences)
O251	5'-fluorescein-CCATCCTGCTGACTGCTTAATCGCTTCGGATGG–3'dabcyl	Molecular beacon specific to a common sequence of RNA products
O252	GAAGCGATTAAGCAGTCAGCA	Oligonucleotide for calibrating beacon signal

Table S2. Correlation coefficients between sequence encoded energy parameters for all 262,144 sequence variants of 9 nt long DNA/RNA.

Sequence encoded energy parameters	**Correlation coefficient**
DNA/DNA duplex stability versus RNA/DNA duplex stability	0.88
DNA/DNA duplex stability versus base stacking	0.06
DNA/DNA duplex stability versus posttranslocated state bias	0.51
RNA/DNA duplex stability versus base stacking	0.37
RNA/DNA duplex stability versus posttranslocated state bias	0.70
Posttranslocated state bias versus base stacking	0.48

Supplementary Figures

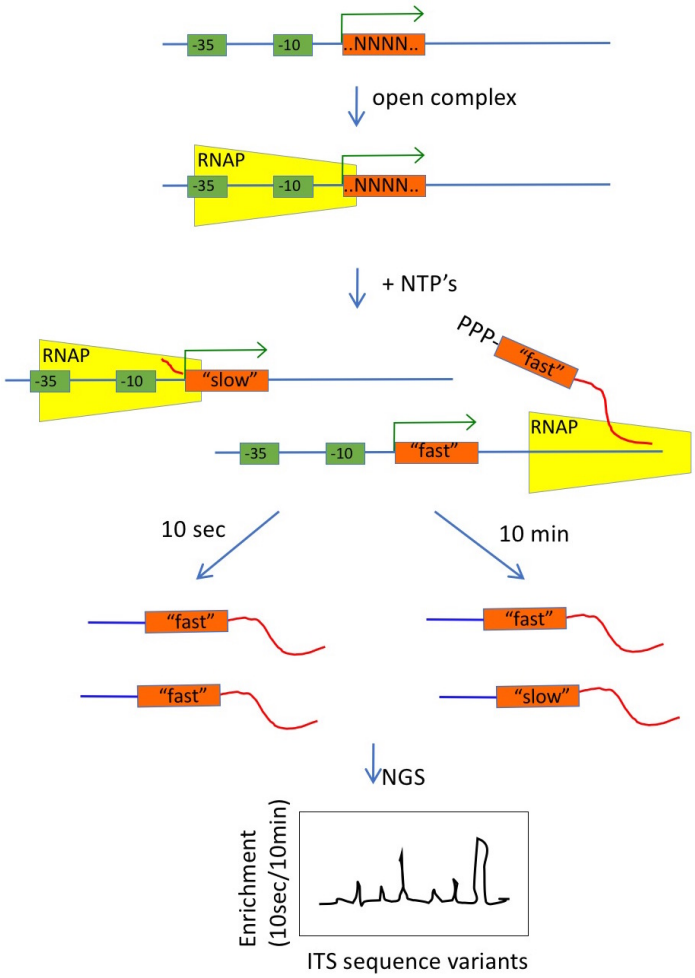

Figure S1. NGS-based approach to analyze in parallel promoter escape of DNA template library containing many variants of ITS. RNAP is incubated with the template library to form open complexes and transcription is initiated by the addition of NTP and heparin (to prevent reinitiation of transcription). Transcription is stopped by the addition of EDTA after 10 sec and 10 min. RNA products are isolated, reversed transcribed into DNA and are subjected to NGS sequencing. Sequencing data is analyzed to determine number of reads for each of the ITS variant in the analyzed library. Enrichment, the ratio of reads at 10sec and 10 min (each normalized to the total number of reads in a sequencing run) is calculated.

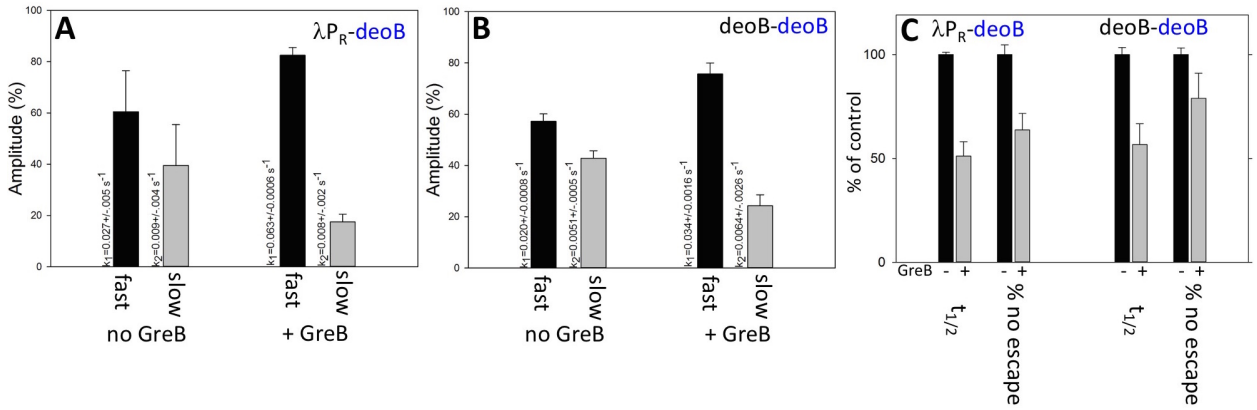

Figure S2. Effect of GreB on the amplitudes of fast and slow kinetic components. (A) λP_R promoter with deoB ITS. (B) deoB promoter with deoB ITS. The values of fast (k_1) and slow (k_2) rate constants are shown as well. (C) Effect of GreB on overall escape kinetics ($t_{1/2}$) and on the amount of unproductive "open" complexes. Y axis (% control) depicts the values of $t_{1/2}$ or fraction of no escape complexes expressed as % of values measured in the absence of GreB.

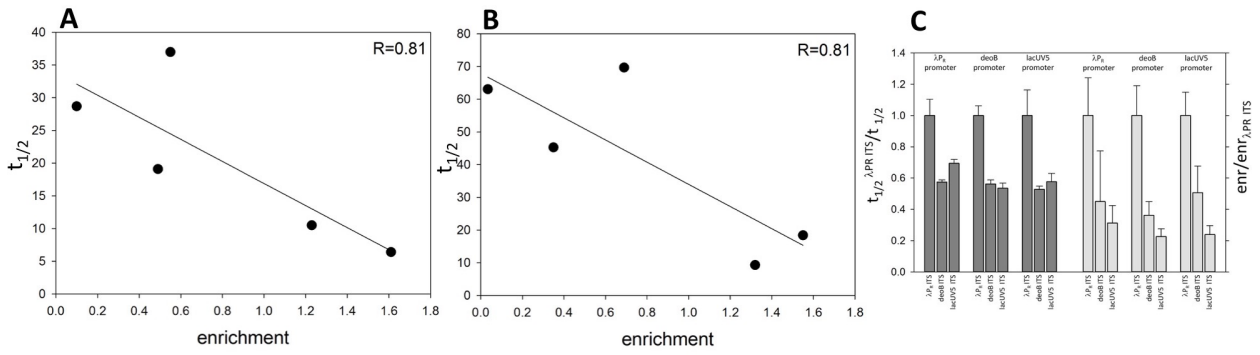

Figure S3. Correlation between half-time values for escape and enrichment values for 5 ITS from Fig. 3 in a context of λP_R (A) and (B) deoB promoters. (C) Comparison of $t_{1/2}$ of escape reaction (calculated from time-dependence of full-length transcript production measured by molecular beacon assay) and corresponding enrichments from NGS-based experiment.

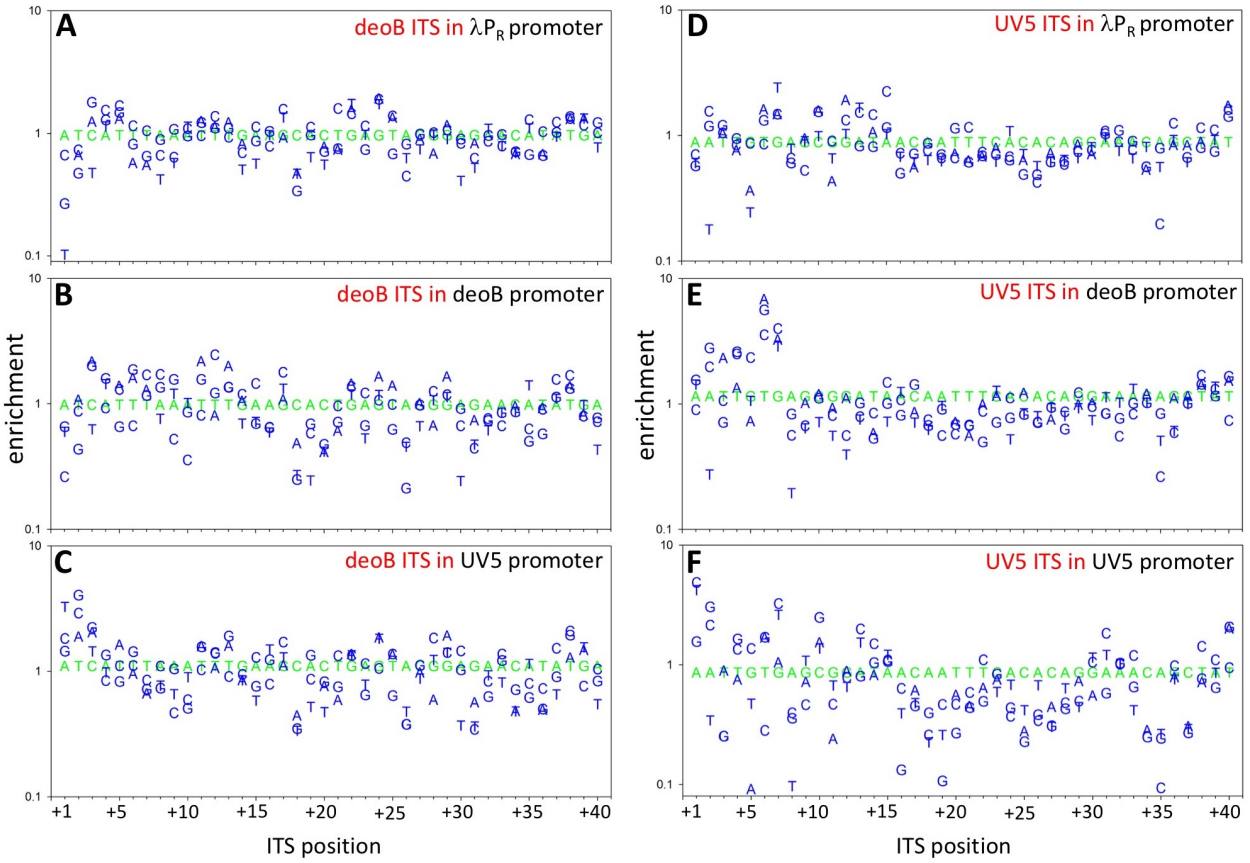

Figure S4. Effects of single base substitutions in 40 bp deoB (A-C) and 40 bp UV5 (D-F) ITS on promoter escape in a context of λP$_R$ promoter (A and D), deoB promoter (B and E) and UV5 promoter (C and F). Wt sequences are shown in green.

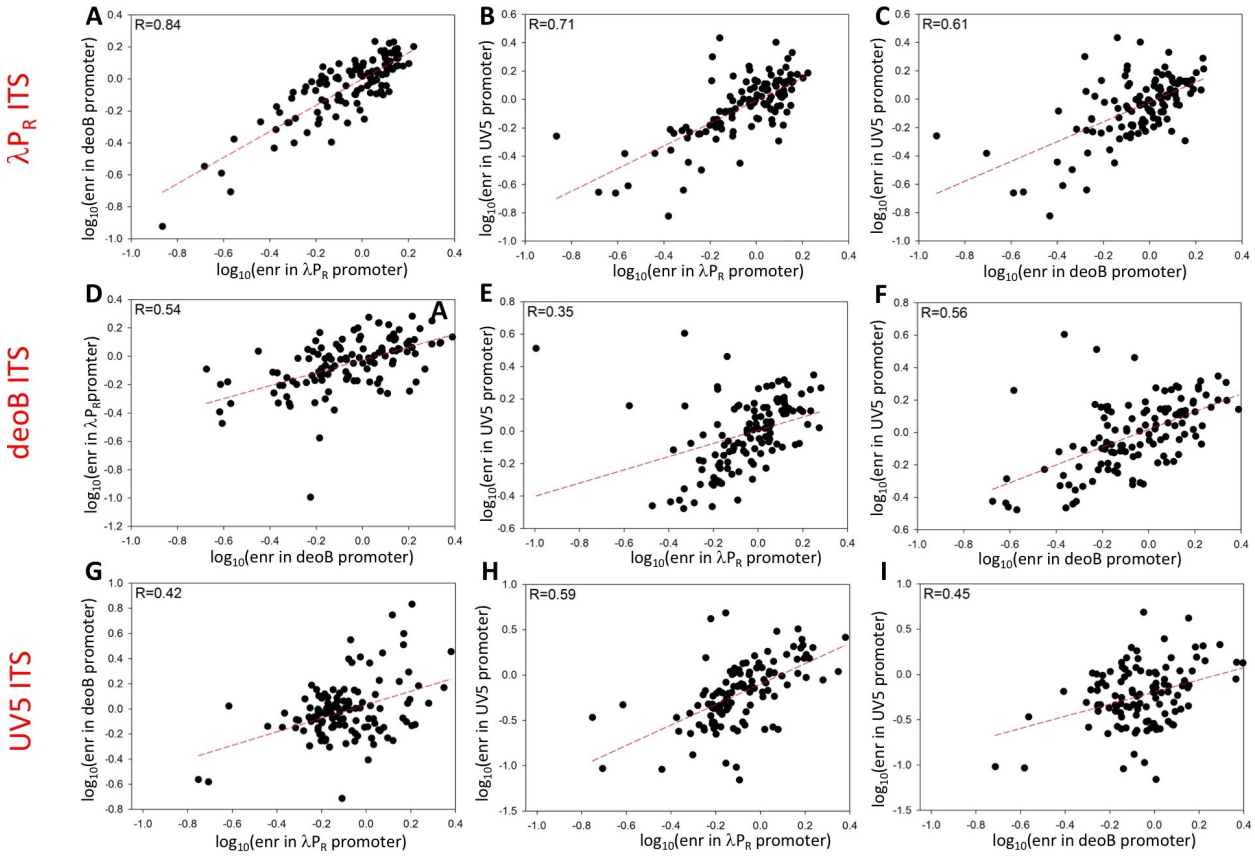

Figure S5. Correlations between the effects of single base substitutions in ITS ((λP_R ITS, panels A-C; deoB ITS, panels D-F; UV5 ITS, panels G-I) on promoter escape in different promoter contexts.

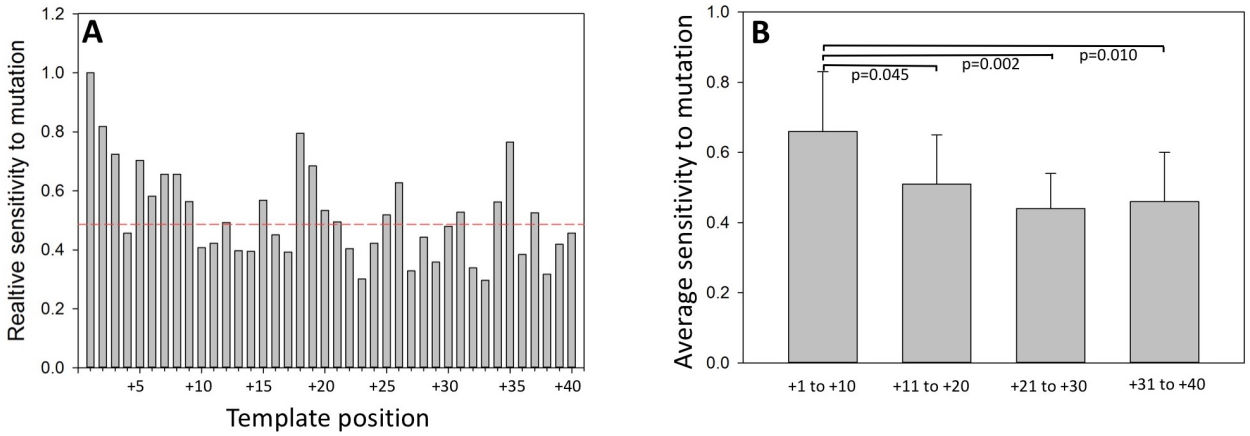

Figure S6. (A) Relative sensitivity of promoter escape to mutations at each template position. Red line marks the average value. Relative sensitivity was calculated by summing the differences between the enrichment values for each base substitution and the enrichment value for the wt base. All nine combinations of promoter (λP_R, deoB, UV5) and ITS (λP_R, deoB, UV5) were included in this calculation. The resulting values were normalized to the value calculated for position +1. (B) Average values (calculated data from panel A) of relative sensitivity to mutation for 10 bp long segments of ITS. Statistical significance (p-values) of differences between the average for +1 to +10 region and averages for the regions further downstream are shown.

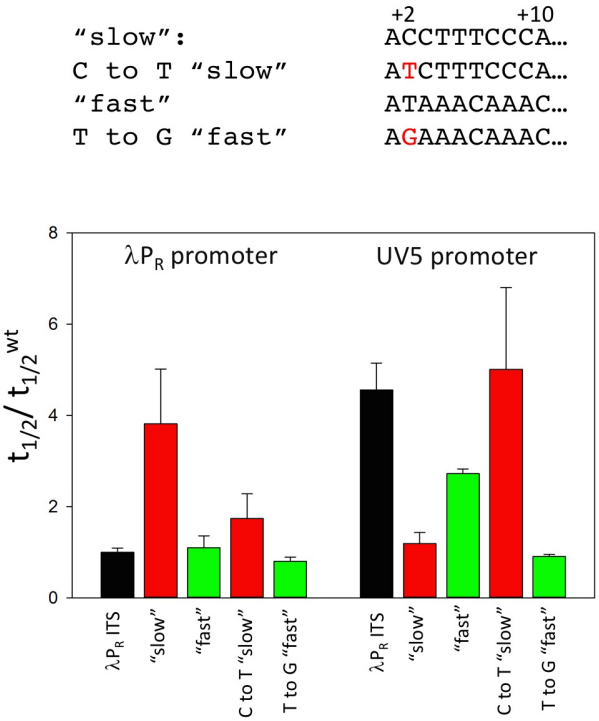

Figure S7. Experiments confirming strong preference for G and T at +2 for UV5 promoter for fast and slow promoter escape, respectively. "Slow" and "fast" sequences correspond to examples of slow and fast escaping sequences selected from the data for λP_R ITS in λP_R promoter context which at +2 position had the bases inconsistent with UV5 promoter preferences at this position for slow and fast escape, respectively (Fig. 6 C&D). Promoter escape kinetics for these sequences and their mutants (with bases at +2 changed to those preferred in UV5 promoter context) was measured by monitoring time dependence of full transcript production using fluorescent beacon assay. Y-axis shows relative values of promoter escape reaction half-times ($t_{1/2}$) normalized to $t_{1/2}$ for λP_R ITS in λP_R promoter. The "slow" and "fast" sequences in the context of UV5 promoter exhibit opposite behavior (i.e. "slow" is fast and "fast is slow) which is reversed by the mutations that change the vases at +2 (to T and G in "slow" and "fast", respectively).

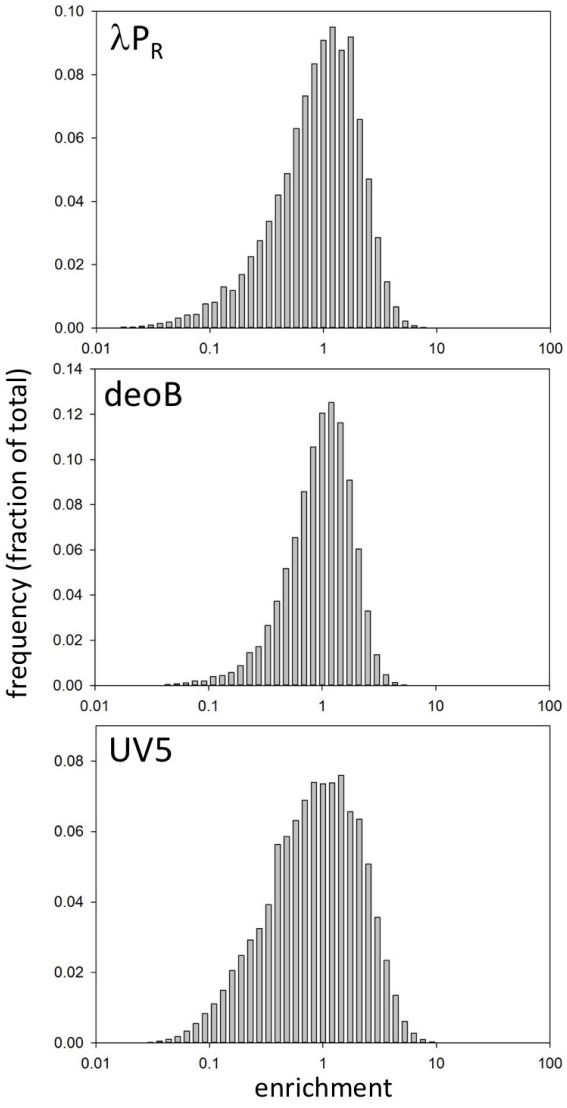

Figure S8. Histograms of the distribution of enrichment values in experiments with indicated promoters with randomized sequence from +2 to +10. Y-axis shows the ratio of the counts in each bin of the histogram to total number of date in points in a data file.

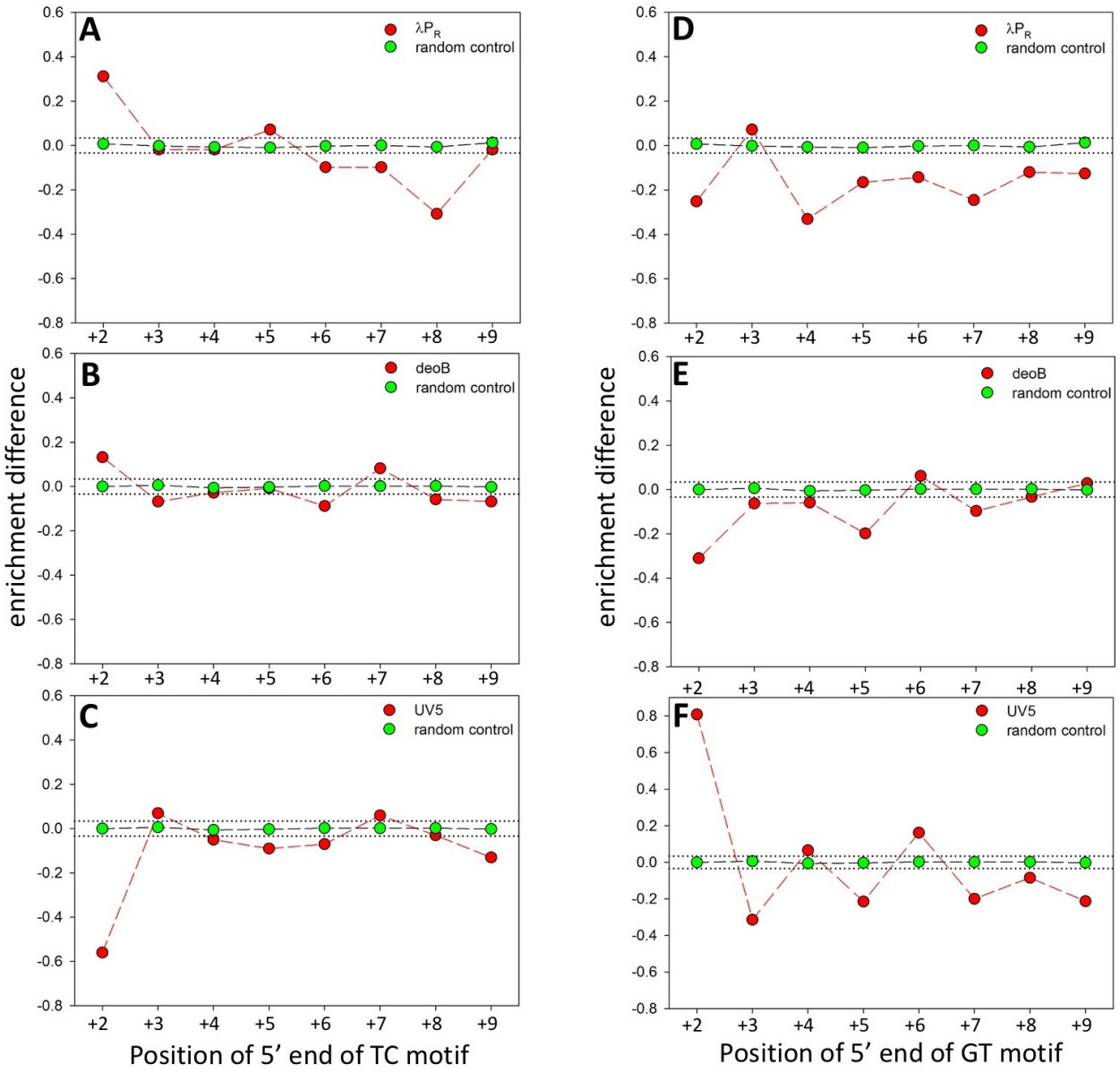

Figure S9. Effect of TC motif (A-C) or GT (D-F) at a specific position within +2 to +10 region in a context of λP_R promoter (A, D), deoB promoter (B, E) and UV5 promoter (C, F) on escape kinetics averaged over all base combinations at remaining positions (16,384 sequences for each TC or GT position) . Enrichment difference (y axis) corresponds to a difference between calculated averaged enrichment with TC or GT at a given position and the average of all enrichment values in the dataset. Random control (green symbols) corresponds to calculations performed on 16,384 randomly selected sequences. Dotted lines depict boundary for statistical significance of enrichment differences over the random control (i.e. enrichment differences between and outside the dotted lines have p values >0.0001 and <0.0001, respectively).

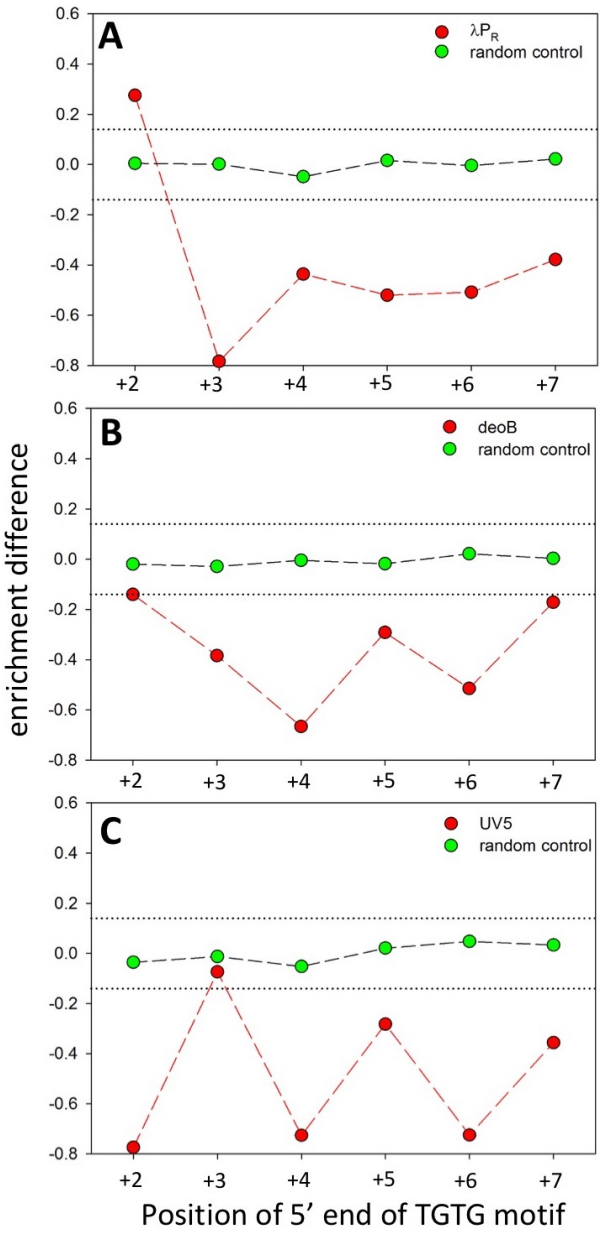

Figure S10. Effect of TGTG motif at a specific position within +2 to +10 region in a context of λP_R promoter (A), deoB promoter (B) or UV5 promoter (C) on escape kinetics averaged over all base combinations at remaining positions (1024 sequences for each TGTG position). Enrichment difference (y axis) corresponds to a difference between calculated averaged enrichment with TGTG at a given position and the average of all enrichment values in the dataset. Random control (green symbols) corresponds to calculations performed on 1024 randomly selected sequences. Dotted lines depict boundary for statistical significance of enrichment differences over the random control (i.e. enrichment differences between and outside the dotted lines have p values >0.0001 and <0.0001, respectively).

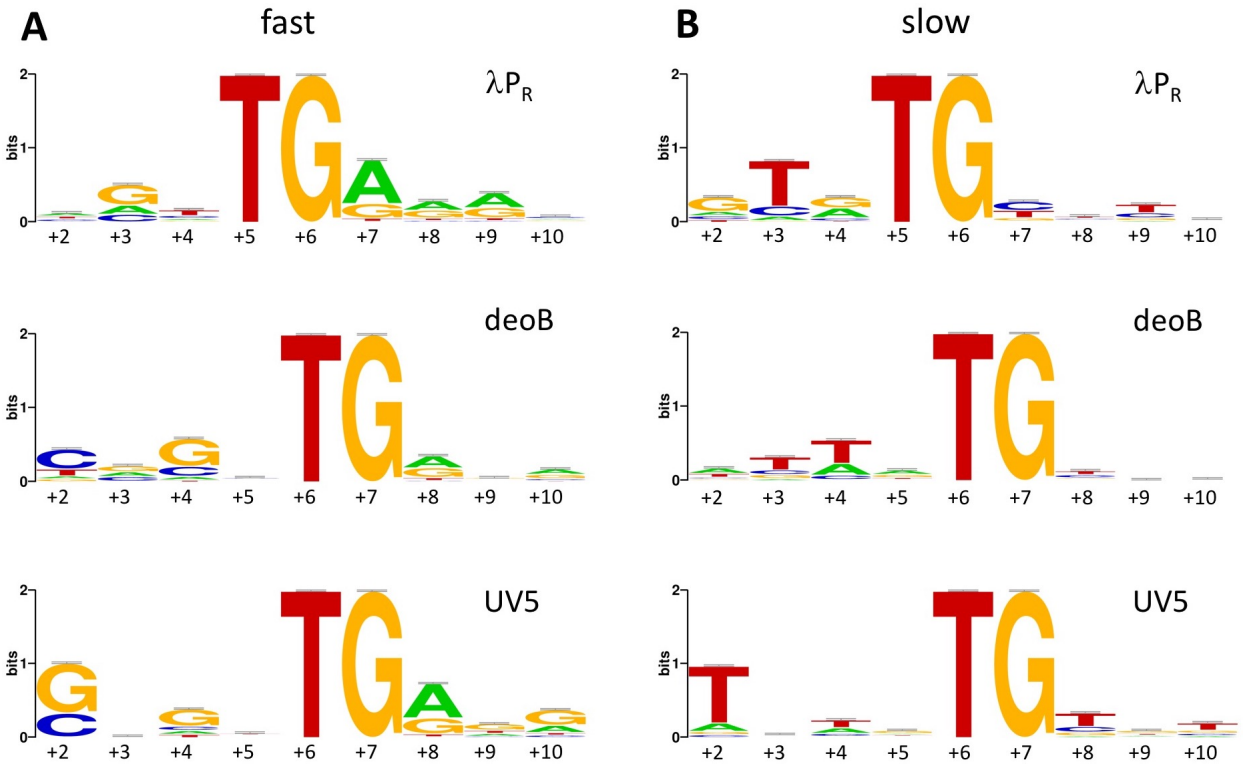

Figure S11. Sequence logos for fast (200 sequences with highest enrichment values) (A) and slow (200 sequences with lowest enrichment values) (B) among 16,384 sequences containing TG motif at +5 (λP_R) or at +6 (deoB or UV5 promoters).

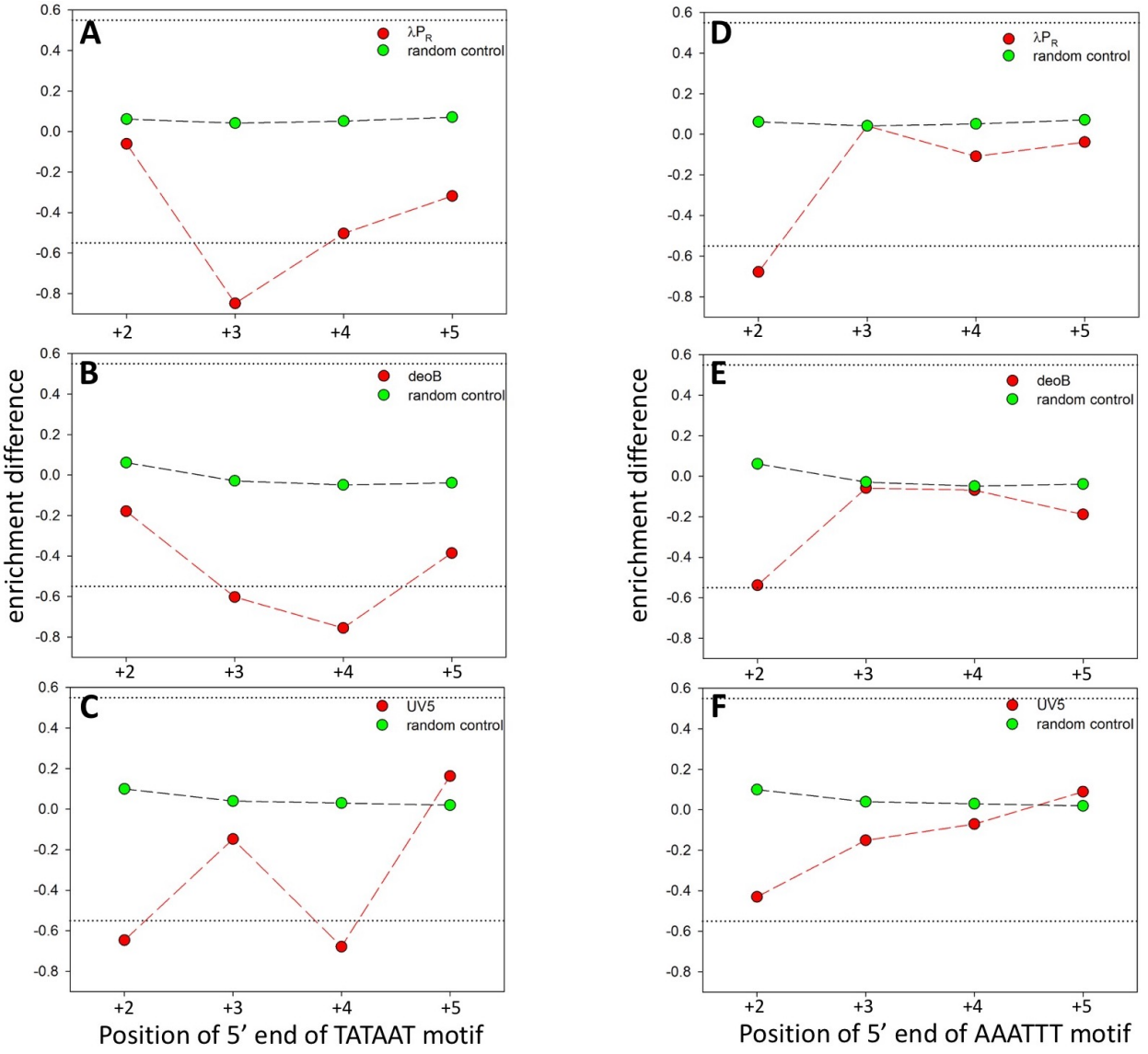

Figure S12. Effect of TATAAT motif (A-C) or AAATTT motif (D-F) at a specific position within +2 to +10 region in a context of λP_R promoter (A,D), deoB promoter (B,E) or UV5 promoter (C,F) on escape kinetics averaged over all base combinations at remaining positions (64 sequences for each TATAAT or AAATTT position). Enrichment difference (y axis) corresponds to a difference between calculated averaged enrichment with TATAAT or AAATTT at a given position and the average of all enrichment values in the dataset. Random control (green symbols) corresponds to calculations performed on 64 randomly selected sequences. Dotted lines depict boundary for statistical significance of enrichment differences over the random control (i.e. enrichment differences between and outside the dotted lines have p values >0.0001 and <0.0001, respectively).

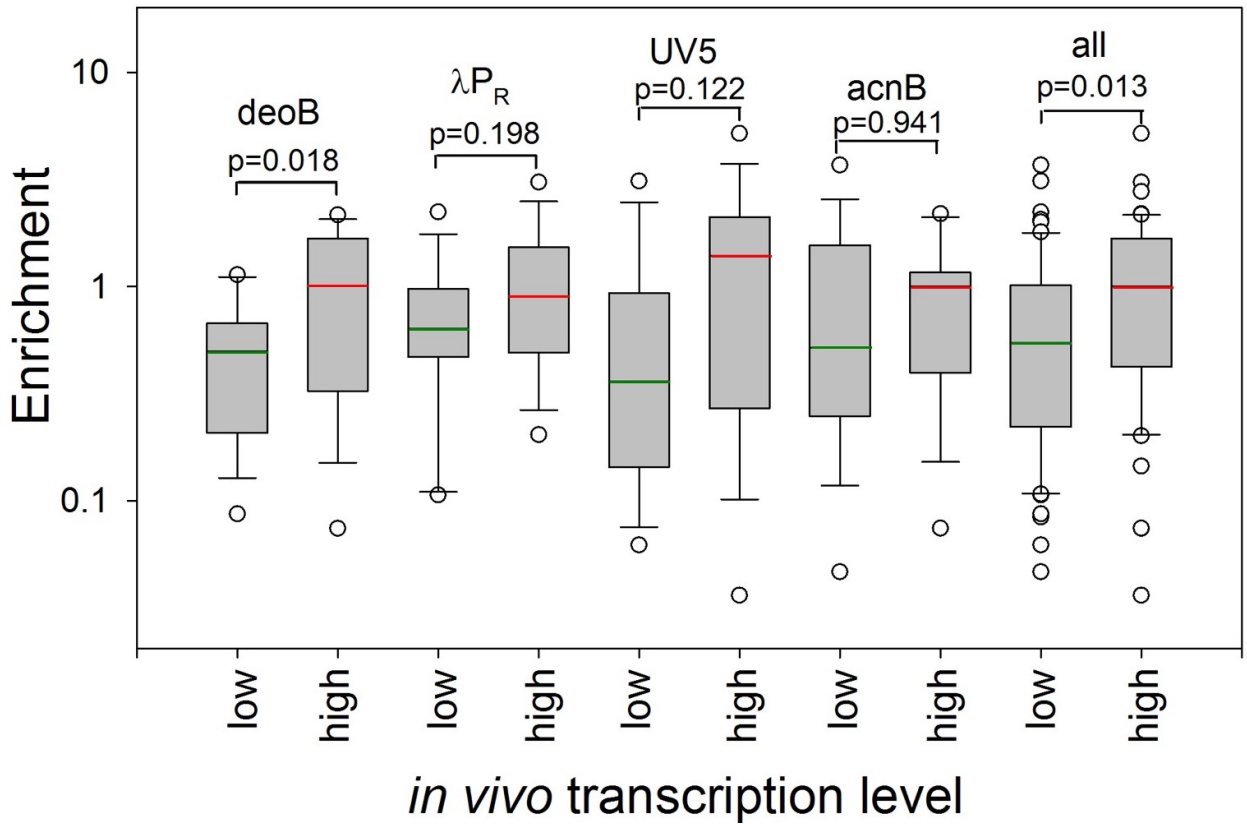

Figure S13. Box plot comparing the enrichment values from NGS-based *in vitro* experiment with the *in vivo* transcript levels. Low and high in *vivo* expression ITS are ITS # 7-21 (Fig. 1B; O7-O27, Table S1a) and ITS # 22-36&6 (Fig. 1B; O22-O36&O6, Table S1a) (2). Green and red lines depict average enrichment values for low and high expression ITS, respectively.

Supplementary References

1. Ko, J. and Heyduk, T. (2014) Kinetics of promoter escape by bacterial RNA polymerase: effects of promoter contacts and transcription bubble collapse. *The Biochemical journal*, **463**, 135-144.
2. Reppas, N.B., Wade, J.T., Church, G.M. and Struhl, K. (2006) The transition between transcriptional initiation and elongation in E. coli is highly variable and often rate limiting. *Mol Cell*, **24**, 747-757.