Supplementary Data

The Core-Independent Promoter-Specific Interaction of Primary Sigma Factor

Hsin-Yi Yeh, Tsung-Ching Chen, Kung-Ming Liou, Hsiu-Ting Hsu, Kuei-Min Chung, Li-Ling Hsu and Ban-Yang Chang*

Institute of Biochemistry, National Chung-Hsing University, Taichung 40227, Taiwan, ROC

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

***To whom correspondence should be addressed. Tel: 886-4-2285-3486; Fax: 886-4-2285-3487; Email: bychang@mail.nchu.edu.tw**

SUPPLEMENTARY MATERIALS AND METHODS

In vitro transcription assay

For *in vitro* transcription, 10 μl of core RNAP (1.8 μg) was mixed with an equal volume of purified σ^A (2.2 μg) and incubated on ice for 10 min. The molar ratio of core to σ^A for σ^A-RNAP reconstitution was 1: 10. Afterwards, 20 μl (0.3 μg) of pCT24 plasmid (1) harboring the G3b promoter DNA of *B. subtilis* ϕ29 phage or the linear G3b promoter DNA fragment spanning from -115 to +67 of the promoter was mixed with the reconstituted σ^A-RNAP, and 40 μl of reaction cocktail (40 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 160 mM KCl, 0.4 mM DTT, 0.2 mM each of UTP, CTP, GTP, ATP, 3 μCi [α-^{32}P] ATP, and 5% of glycerol) pre-warmed at 37˚C was added to start transcription reaction. The reaction was allowed to proceed for 10 min before adding 160 μl of stop buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol). The mRNA products were run on a denaturing polyacrylamide/urea gel and autoradiography was performed after electrophoresis.

The non-promoter DNA used for EMSA and *in vitro* transcription

The two non-promoter DNA fragments used for analyzing the existence of non-specific binding of σ^A are the *tgbp2* and *tgbp3* cDNA fragments from *Bamboo mosaic virus*. The *tgbp2* cDNA fragment is 178-bp long, spanning from base 145 to base 322 of a mutant *tgbp2*. The nucleotide sequence for the *tgbp2* cDNA is CCGCACGGGGGTGCATACGTGGACGGCACCAAAGGAATTCTCTACAACAG CCCCACCTCCTCATACCCATCCTCATCTCTCCCATTCTCCATGGTTATCGCA CTAGCCACAACCCTTTTCCTCATCACCAGAAACATTCTCAACCCAGCCCCC ACCACACCTGCAATCTATGCGCCCC. The *tgbp3* cDNA fragment is 179-bp long and composed of the nucleotide sequence from base 4 to base 156 of *tgbp3* as well as the linker primer sequences at both ends. The nucleotide sequence for the cDNA fragment is AGACTACCATGGCTCTAAACACTGACACACTATGCATCATTCTGTTCATAC TAATATTAGGCATCCTATATAATATACTTCAACAGCATCTGCCCCCACCAT GTGAAATAATAATAAACGGGCACACTATATCCATTAGGGGCAACTGCTAC CACACCACCTCCAGCCTCGAGCACCAC. The underlined bases are linker primers used for the synthesis of *tgbp3* cDNA by polymerase chain reaction.

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Western blot and *in vitro* transcription analyses of purified σ^A. (**A**) Western blot analysis. (**B**) *In vitro* transcription assay. The Wt-σ^A, either overproduced in a soluble (S) or inclusion (R) form in *E. coli*, as well as the mutant σ^A (Mu-σ^A)

overproduced in a soluble form (S) in *E. coli* were purified and subjected to both Western blot and *in vitro* transcription analyses. The anti-β used for Western blotting is able to react efficiently with both the *B. subtilis* and *E. coli* β subunits of RNAP. Methods used for reconstitution of σ^A-RNAP and *in vitro* transcription are described in Supplementary materials and methods. In both (**A**) and (**B**), the Wt- and Mu-σ^A used is 32 μg which was as high as that used for detecting the footprint of σ^A on the G3b promoter DNA by DNase I footprinting. In (**A**), the core RNAP used is 50 ng.

Figure S2. The functional properties of σ^A overproduced in a soluble form in *E. coli*. (**A**) The promoter DNA-binding activities of σ^A as analyzed by EMSA. Various concentrations (in μM) of the soluble σ^A (S) or refolded σ^A (R) as indicated above each lane were mixed with labeled G3b promoter DNA (1 nM). The detailed method for EMSA is described in Materials and methods of the main text. (**B**) The promoter DNA-binding patterns of σ^A as analyzed by DNase I footprinting assay. The σ^A-RNAP was reconstituted from core RNAP and 10-fold molar excess of σ^A. The σ^A-promoter DNA and σ^A-RNAP-promoter DNA complexes formed at designed temperature (20˚C) was digested with DNase I at the same temperature. The numbers shown under σ^A (S) and σ^A (R) are the concentrations of σ^A in μM. The concentration of σ^A-RNAP used is 200 nM. The region encompassed by the upward and downward arrowheads is the DNA with a clear change of DNase I digestion pattern in the presence of σ^A. Both the -10 and -35 elements protected by σ^A are indicated by dotted lines. (**C**) The transcription activities of σ^A as analyzed by *in vitro* transcription assay. The numbers shown under σ^A-RNAP (S) and σ^A-RNAP (R) are the molar ratios of σ^A to core RNAP and those shown on the right margin of panel are the lengths of mRNA transcripts in base. Methods used for purification of the soluble σ^A and the σ^A in inclusion bodies overproduced in *E. coli* are described in Materials and methods of the main text. S, the σ^A overproduced in *E. coli* in a soluble form. R, the σ^A overproduced in *E. coli* in inclusion bodies.

Figure S3. The promoter strengths and σ^A-binding activities of the G3b promoter DNA, *tgbp2* cDNA and *tgbp3* cDNA. (**A**) The transcription of G3b promoter DNA, *tgbp2* cDNA and *tgbp3* cDNA by σ^A-RNAP *in vitro*. Triplicate samples of *in vitro* transcription are shown for each tested DNA. The number shown on the right margin is the length (in base) of mRNA transcript. (**B**) The binding affinity of σ^A for the G3b promoter DNA, *tgbp2* cDNA and *tgbp3* cDNA as examined by EMSA. Various concentrations (in μM) of σ^A as indicated above each lane were mixed with labeled G3b promoter DNA (1 nM). The detailed method for EMSA was described in Materials and methods of the main text. The binding mixture was incubated at 37˚C

for 10 min before being run on a 5% non-denaturing polyacrylamide gel in a 4˚C-cold chamber.

Figure S4. The binding of σ^A or σ^A-RNAP to *trnS* spacing variants at 20˚C as analyzed by EMSA. Top panel: the binding of σ^A to each *trnS* spacing variant. Each binding mixture contains 10 μM σ^A, 0.5 nM *trnS* spacing variant and 0.1 μM heparin. Bottom panel: the binding of σ^A-RNAP to each *trnS* spacing variant. Each binding mixture contains 0.5 nM *trnS* spacing variant and 50 nM σ^A-RNAP reconstituted from 50 nM core RNAP and 10-fold molar excess of σ^A by incubation at 37˚C for 10 min. The binding mixture was incubated at 20˚C for 10 min before being electrophoresed on a 5% non-denaturing polyacrylamide gel in a 20˚C-cold chamber. The numbers above the horizontal lines are the spacer lengths of *trnS* spacing variants. The plus and minus signs denote, respectively, the addition and omission of σ^A or σ^A-RNAP.

Figure S5. The effect of σ^A concentration on the footprinting pattern of σ^A on the *trnS*-17 promoter. The footprinting patterns of σ^A on the template (left panel) and non-template strand DNA (right panel) of the *trnS*-17 promoter at 20˚C. Method used for DNase I footprinting assay was the same as that described in Materials and methods of the main text. In this assay, various concentrations of σ^A (5, 10, 20 and 40 μM) were used as indicated on top of each lane. The same assay was also performed at 37˚C; however, no footprint of σ^A was detected. The numbers shown on the left of each panel are positions relative to the transcription start site of the *trnS*-17 promoter. The DNA protected by σ^A is encompassed by the upward and downward arrowheads. The promoter -10 and -35 elements protected by σ^A are indicated by dotted lines.

Figure S6. The specificity of holo RNAP binding to *trnS* promoter having an optimal spacing (17±1 bp) as analyzed by EMSA-based competition assay. Shown is the fractional retention of the preformed σ^A-RNAP-*trnS* promoter DNA complex as a function of the molar ratio of competitor DNA to ^{32}P-labeled *trnS* spacing variant ([Competitor]/[Labeled DNA]). The ^{32}P-labeled *trnS* spacing variant is indicated as *trnS*-X (X is the length of spacer DNA ranging from 16 to 18) above each chart. Each of the preformed binary complexes was obtained by mixing 0.5 nM ^{32}P-labeled *trnS* spacing variant and 50 nM σ^A-RNAP. For reconstitution of holo RNAP, the molar ratio of core RNAP to σ^A was 1: 10. The preformed binary complex was challenged with an indicated concentration of competitor DNA (0, 0.5, 2.5, 5, 12.5, 25, 50 nM) at 37˚C for 10 min, then analyzed by EMSA at room temperature. The competitor DNA used is indicated at lower right. Each retention value was the average of triplicate measurements with a standard deviation indicated by vertical bar.

Figure S7. Analysis of the naturally disordered region of *B. subtilis* σ^A by PONDR®. The predictor, PONDR® VL-XT, was used for prediction of the naturally disordered region of *B. subtilis* σ^A (2-4). Eight disordered regions were detected. The most extensive and disordered one (encompassing amino acid residues 229 ~ 297 and containing almost entire region 3 and a small N-terminal part of region 4.1 of σ^A) is indicated by thick line. The X-axis indicates the position of amino acid residue in σ^A. The bottom scale indicates the positions of the four conserved regions of σ^A. The Y-axis indicates the PONDR® value of each amino acid residue. It is considered to be disordered, if the value exceeds or matches a threshold of 0.5.

SUPPLEMENTARY REFERENCES

1. Liao,C.T., Wen,Y.D., Wang,W.H. and Chang,B.Y. (1999) Identification and characterization of a stress-responsive promoter in the macromolecular synthesis operon of *Bacillus subtilis*. *Mol. Microbiol.*, **33**, 377-388.
2. Li,X., Romero,P., Rani,M., Dunker,A.K. and Obradovic,Z. (1999) Predicting protein disorder for N-, C-, and internal regions. *Genome Inform. Ser. Workshop Genome Inform.*, **10**, 30-40.
3. Romero, Obradovic and Dunker,K. (1997) Sequence data analysis for long disordered regions prediction in the Calcineurin family. *Genome Inform. Ser. Workshop Genome Inform.*, **8**, 110-124.
4. Romero,P., Obradovic, Z., Li,X., Garner,E.C., Brown,C.J. and Dunker,A.K. (2001) Sequence complexity of disordered protein. *Proteins*, **42**, 38-48.

SUPPLEMENTARY FIGURES

Figure S1

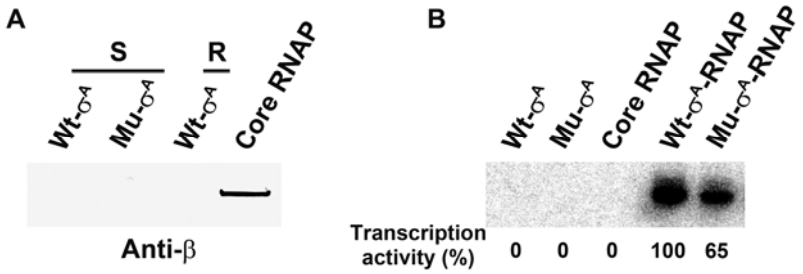

Figure S2

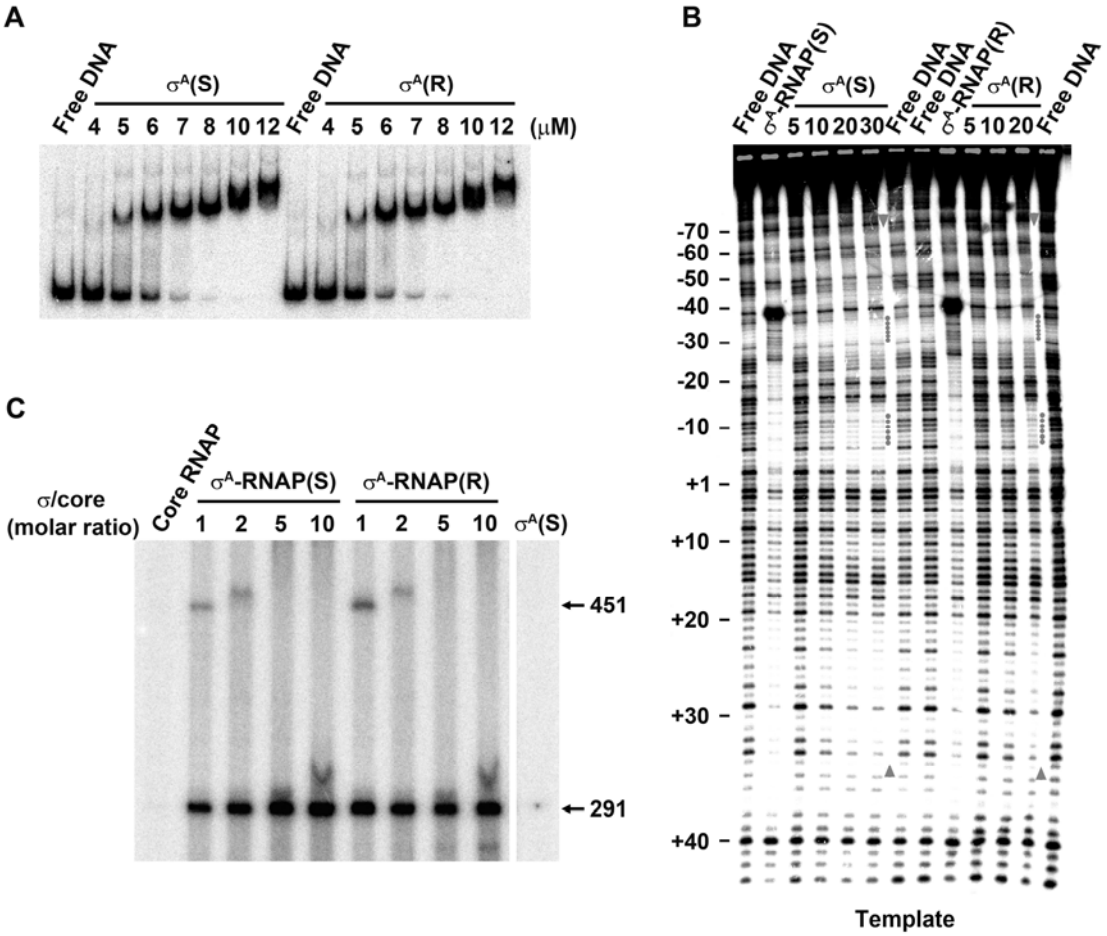
A
Free DNA
σA(S)
4 5 6 7 8 10 12
Free DNA
σA(R)
4 5 6 7 8 10 12
(μM)
B
Free DNA
σA-RNAP(S)
σA(S)
5 10 20 30
Free DNA
Free DNA
σA-RNAP(R)
σA(R)
5 10 20
Free DNA
-70
-60
-50
-40
-30
-20
-10
+1
+10
+20
+30
+40
Template
C
σ/core
(molar ratio)
Core RNAP
σA-RNAP(S)
σA-RNAP(R)
1 2 5 10 1 2 5 10
σA(S)
451
291

Figure S3

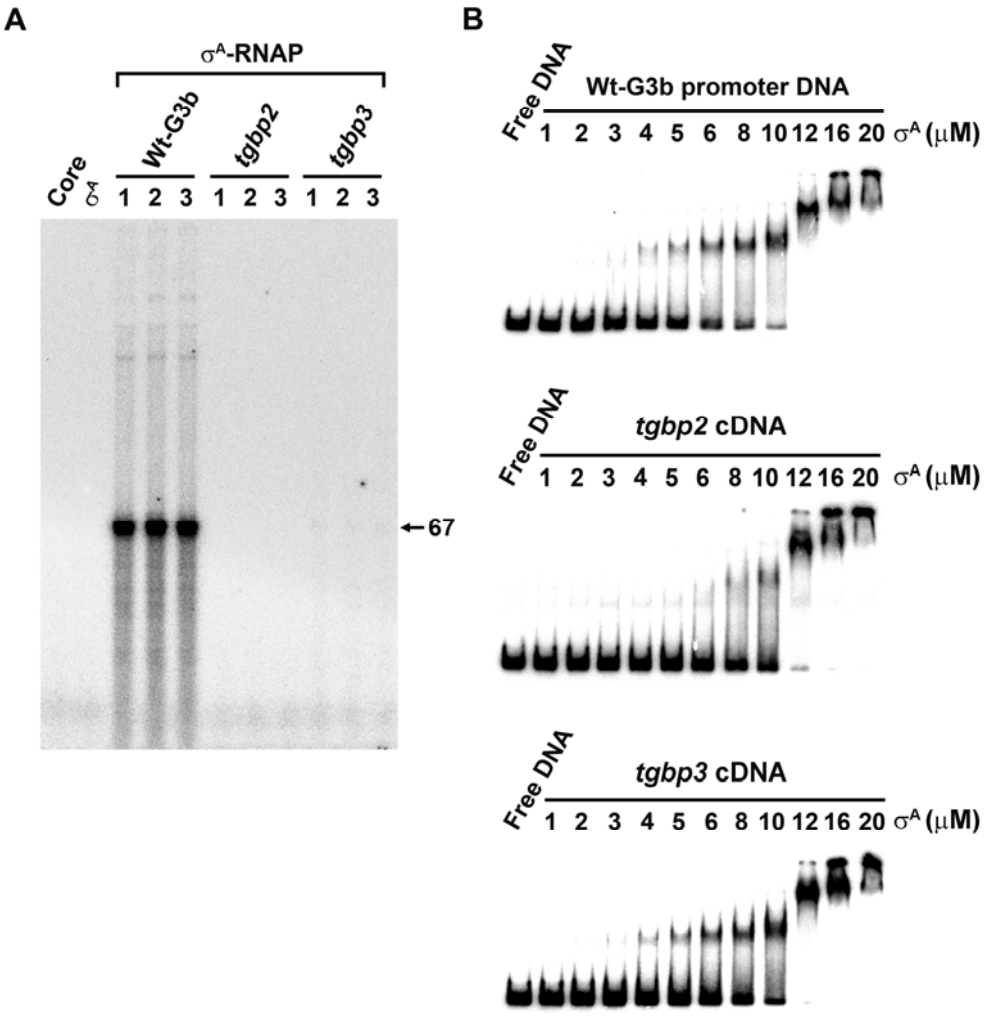
A
σA-RNAP
Core σA
Wt-G3b
tgbp2
tgbp3
1 2 3 1 2 3 1 2 3
67
B
Free DNA
Wt-G3b promoter DNA
1 2 3 4 5 6 8 10 12 16 20 σA (μM)
Free DNA
tgbp2 cDNA
1 2 3 4 5 6 8 10 12 16 20 σA (μM)
Free DNA
tgbp3 cDNA
1 2 3 4 5 6 8 10 12 16 20 σA (μM)

Figure S4

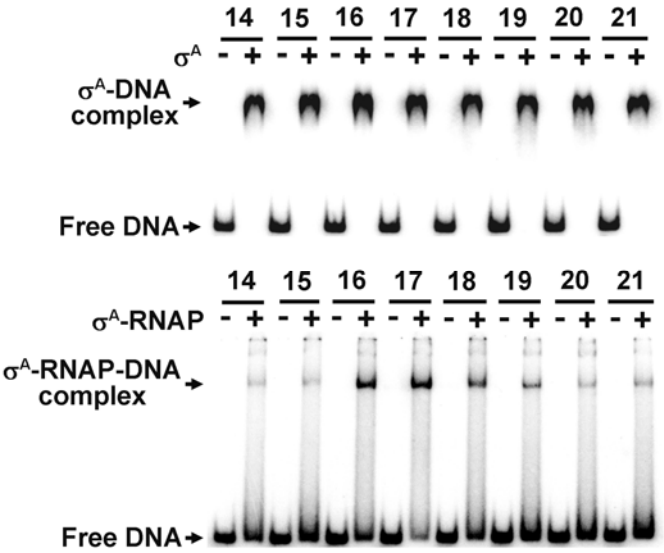

Figure S5

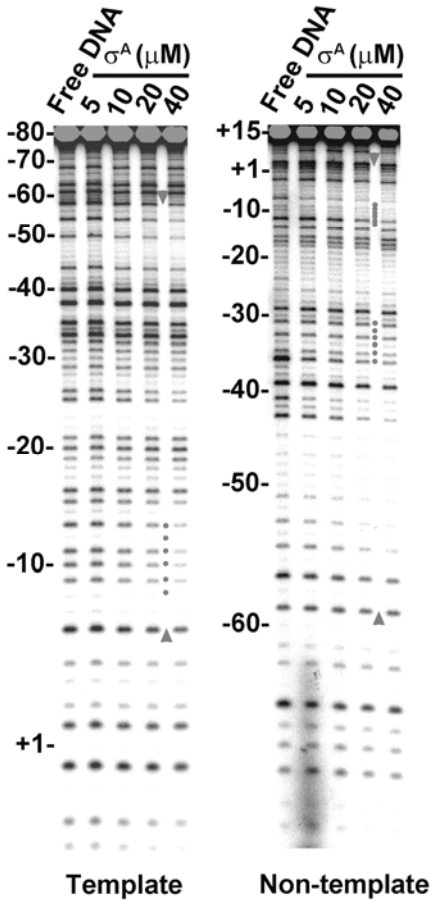

Figure S6

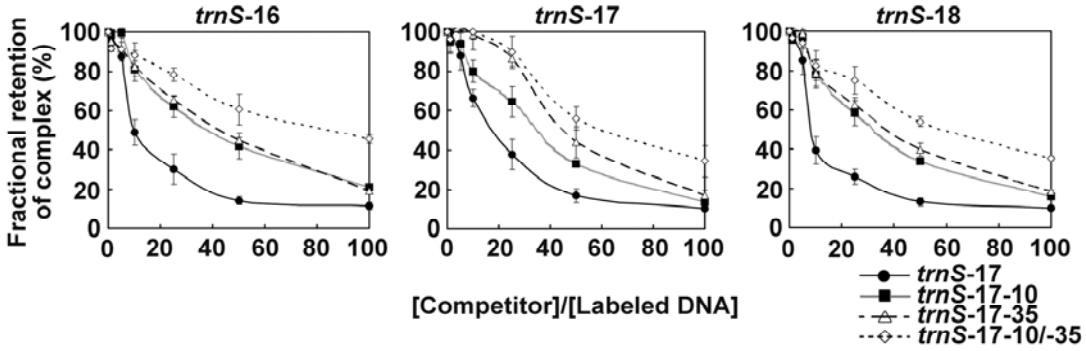

Figure S7

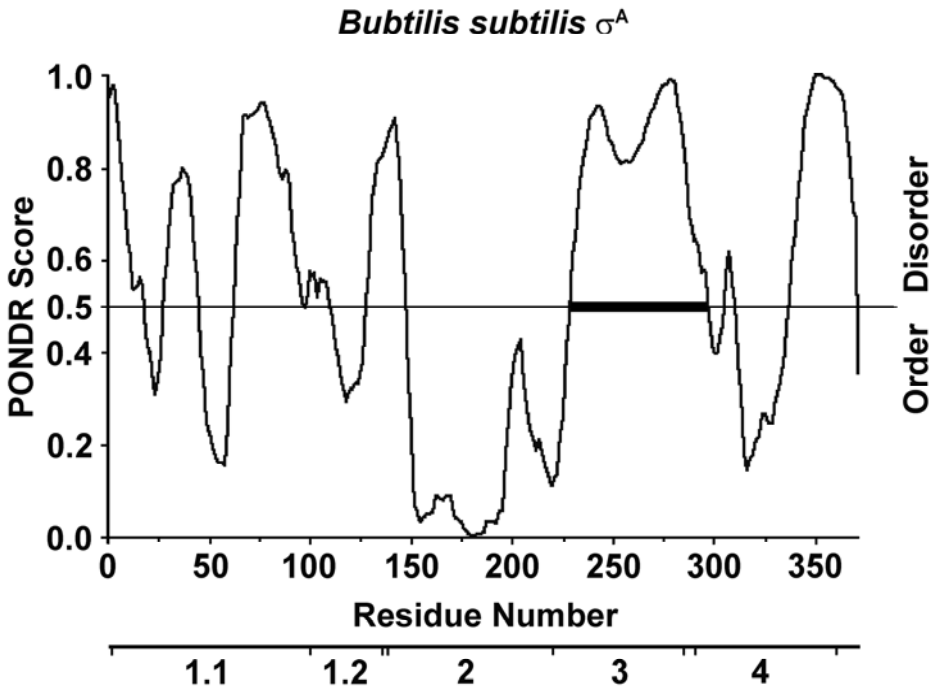

Table S1. The templates and primers used for the construction and synthesis of promoter variants by polymerase chain reaction

Promoter variants	Templates	Primers
The *trnS* spacing variants		
trnS-14	pOtrnS[a]	14F: 5′-cctcttgacactgcaaaaggctgatataataagtc-3′
		14R: 5′-gacttattatatctgccttttgcagtgtcaagagg-3′
trnS-15	pOtrnS	15F: 5′-cctcttgacactgcaaataggctgatataataagtc-3′
		15R: 5′-gacttattatatcagcctatttgcagtgtcaagagg-3′
trnS-16	pOtrnS	16F: 5′-cctcttgacactgcaaataaggctgatataataagtc-3′
		16R: 5′-gacttattatatcagccttatttgcagtgtcaagagg-3′
trnS-18	pOtrnS	18F: 5′-cctcctgacactgcaaatctaaggctgatataataagtc-3′
		18R: 5′-gacttattatatcagccttagatttgcagtgtcaagagg-3′
trnS-19	pOtrnS	19F: 5′-cctcttgacactgcaaatactaaggctgatataataagtc-3′
		19R: 5′-gacttattatatcagccttagtatttgcagtgtcaagagg-3′
trnS-20	pOtrnS	20F: 5′-cctattgacactgcaaatacataaggctgatataataagtc-3′
		20R: 5′-gacttattatatcagccttatgtatttgcagtgtcaagagg-3′
trnS-21	pOtrnS	21F: 5′-cctcttgacactgcaaattacataaggctgatataataagtc-3′
		21R: 5′-gacttattatatcagccttatgtaatttgcagtgtcaagagg-3′
The *trnS*-17 promoter with base substitutions at -10, -35 or both elements		
trnS-17-10	pOtrnS	-10F: 5′- cactgcaaatcaaggctgaggtaacaagtcttgtctcattattcacagtagctc-3′
		-10R: 5′-gagctactgtgaataatgagacaagacttgttacctcagccttgatttgcagtg-3′
trnS-17-35	pOtrnS	-35F: 5′-cgtcacgagagataaataaaaacatttacctcggaacactgcaaatcaaggctg-3′
		-35R: 5′-cagccttgatttgcagtgttccgaggtaaatgtttttatttatctctcgtgacg-3′
trnS-17-10/-35	pOtrnS-35[b]	-10F: 5′- cactgcaaatcaaggctgaggtaacaagtcttgtctcattattcacagtagctc-3′
		-10R: 5′-gagctactgtgaataatgagacaagacttgttacctcagccttgatttgcagtg-3′
The G3b promoter with single or pairwise substitutions at -10, -35 element and TG motif		
G3b-10	pCT-G3b[c]	g-10F: 5′-gaaaagtgttgaaaattgtcgaacagggtgaggtaacaaaagagtagaagag-3′
		g-10R: 5′-ctcttctactcttttgttacctcaccctgttcgacaattttcaacacttttc-3′
G3bTG	pCT-G3b	TG1F: 5′-gaaaagtgttgaaaattgtcgaacaggggaatataataaaagagtagaagag-3′
		TG1R: 5′-ctcttctactcttttattatattcccctgttcgacaattttcaacacttttc-3′
G3b-35	pCT-G3b	g-35F: 5′-gtcgaacttttttatagaaaagtgggaaaaattgtcgaacagggtgatataataaaagagtagaagag-3′
		g-35R: 5′-ctcttctactcttttattatatcaccctgttcgacaatttttcccacttttctataaaaagttcgac-3′
G3b-10/TG	pCT-G3b-10[d]	TG2F: 5′-gaaaagtgttgaaaattgtcgaacaggggaaggtaacaaaagagtagaagag-3′

		TG2R: 5′-ctcttctactcttttgttaccttcccctgttcgacaattttcaacacttttc-3′
G3b-10/-35	pCT-G3b-35[e]	g-10F-1: 5′-gaaaagtgggaaaaattgtcgaacagggtgaggtaacaaaagagtagaagag-3′
		g-10R-1: 5′-ctcttctactcttttgttacctcaccctgttcgacaatttttcccacttttc-3′
G3b-35/TG	pCT-G3b-35	TG3F: 5′-gaaaagtgggaaaaattgtcgaacaggggaatataataaaagagtagaagag-3′
		TG3R: 5′-ctcttctactcttttattatattcccctgttcgacaatttttcccacttttc-3′

[a] pOtrnS, the wild-type *trnS* promoter DNA-containing plasmid

[b] pOtrnS-35, the pOtrnS-derived plasmid with a mutation at the -35 element of *trnS*-17 promoter

[c] pCT-G3b, the wild-type G3b promoter DNA-containing plasmid

[d] pCT-G3b-10, the pCT-G3b-derived plasmid with a mutation at the -10 element of G3b promoter

[e] pCT-G3b-35, the pCT-G3b-derived plasmid with a mutation at the -35 element of G3b promoter