JOURNAL OF BACTERIOLOGY, Dec. 2007, p. 8758–8764 Vol. 189, No. 23
0021-9193/07/$08.00+0 doi:10.1128/JB.01076-07
Copyright © 2007, American Society for Microbiology. All Rights Reserved.

Identification of Promoters Recognized by RNA Polymerase-σ^E Holoenzyme from *Thermus thermophilus* HB8▽†

Akeo Shinkai,[1]* Naomi Ohbayashi,[2] Takaho Terada,[2] Mikako Shirouzu,[2] Seiki Kuramitsu,[1,3] and Shigeyuki Yokoyama[1,2,4]*

RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan[1]; Genomic Sciences Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan[2]; Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan[3]; and Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan[4]

Received 7 July 2007/Accepted 17 September 2007

Thermus thermophilus **σ^E, an extracytoplasmic function σ factor from the extremely thermophilic bacterium *Thermus thermophilus* HB8, bound to the RNA polymerase core enzyme and showed transcriptional activity. With the combination of in vitro transcription assay and GeneChip technology, we identified three promoters recognized by σ^E. The predicted consensus promoter sequence for σ^E is 5′-CA(A/T)(A/C)C(A/C)-N_{15}-CCGTA-3′.**

Bacterial DNA-dependent RNA polymerase (RNAP) is composed of a core enzyme with the subunit structure $\alpha_2\beta\beta'\omega$ and a σ subunit that is responsible for promoter recognition (3, 5, 26). In general, most transcription in exponentially growing cells is initiated by an RNAP holoenzyme containing a housekeeping σ factor, a homolog of *Escherichia coli* σ^{70}. In contrast, alternative σ factors are involved in the transcription of specialized genes such as those that are activated under specific stress conditions and during growth transitions and morphological changes and those involved in pathogens (1, 6, 9, 11, 17, 21).

Most alternative σ factors, i.e., other than the σ^{54} family, are related in sequence to members of the housekeeping σ^{70} family, and they have been divided into three groups based on phylogenetic relatedness (9, 15, 16). Group 4 σ factors, also called extracytoplasmic function (ECF) σ factors, are distant, according to sequence similarity, from the σ^{70} family, and their sequences are divergent relative to those of most other σ factors. While a typical σ^{70}-type (e.g., group 1) σ factor consists of four regions, a typical ECF σ factor consists of two regions corresponding to regions 2 and 4 of a group 1 σ factor, which are involved in the interaction with the −10 and −35 regions of the promoter, respectively. Bacterial genomic sequences suggest that many bacteria have multiple ECF σ factors (18).

Thermus thermophilus HB8, which belongs to the phylum *Deinococcus-Thermus*, is an extremely thermophilic bacterium that was isolated from the water at a Japanese hot spring and can grow at 47 to 85°C, with an optimum temperature range of 65 to 72°C (20). The complete genomic sequence of this strain has been determined (NCBI accession numbers NC_006461, NC_006462, and NC_006463) and shows that the genome is composed of 1.85-megabase-pair chromosomal DNA, 0.26-megabase-pair plasmid pTT27, and 9.32-kilobase-pair plasmid pTT8, encoding 1,973, 251, and 14 open reading frames (ORFs), respectively. *T. thermophilus* HB8 has a typical bacterium-type RNAP, with the housekeeping σ^A (29, 32). In addition to σ^A, ORF TTHB211 (NCBI accession number YP_145450) has been found to be the sole alternative σ factor in the genome of this strain of *T. thermophilus*. This factor encodes a COG1595 family protein, which is designated RpoE, a DNA-directed RNAP-specialized σ subunit and a σ^{24} homolog, in the NCBI database of Clusters of Orthologous Groups (27).

The TTHB211 protein, which we refer to as *T. thermophilus* σ^E in this study, reportedly consists of 193 amino acid residues. During the course of our studies, we found that the N-terminal six residues were derived from the promoter region of the *sigE* gene (see below); therefore, *T. thermophilus* σ^E is most likely composed of 187 amino acid residues, with a predicted molecular mass of 20,919 Da (Fig. 1). According to a conserved domain database search (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml), the regions from ^{28}L to ^{85}H and ^{133}A to ^{182}E of *T. thermophilus* σ^E comprise conserved domains corresponding to regions 2 and 4 of σ^{70}-type σ factors, with the e values being 9e-7 and 3e-6 for the consensus sequences, respectively (Fig. 1). A BLAST search (http://www.ncbi.nlm.nih.gov/BLAST/) revealed that the TT_P0164 protein from *T. thermophilus* HB27 was the closest homolog of *T. thermophilus* σ^E (e value = 4e-83), with only a single amino acid substitution (S160R). The *T. thermophilus* σ^E homologs that showed high e values were from the phyla *Proteobacteria*, *Deinococcus-Thermus*, *Actinobacteria*, and *Chloroflexi*, and the homologs are all classified as members of the ECF σ factor family (Fig. 1). The *T. thermophilus sigE* gene is located on megaplasmid pTT27, on which many horizontally transferred genes are located (19). This might be the reason why homologs of σ^E are widely distributed even in phylogenetically distant bacteria.

We observed the expression profile of *T. thermophilus sigE*

* Corresponding author. Mailing address for Akeo Shinkai: RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan. Phone: 81 791 58 2891. Fax: 81 791 58 2892. E-mail: ashinkai@spring8.or.jp. Mailing address for Shigeyuki Yokoyama: Genomic Sciences Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. Phone: 81 45 503 9196. Fax: 81 45 503 9195. E-mail: yokoyama@biochem.s.u-tokyo.ac.jp.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 28 September 2007.

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

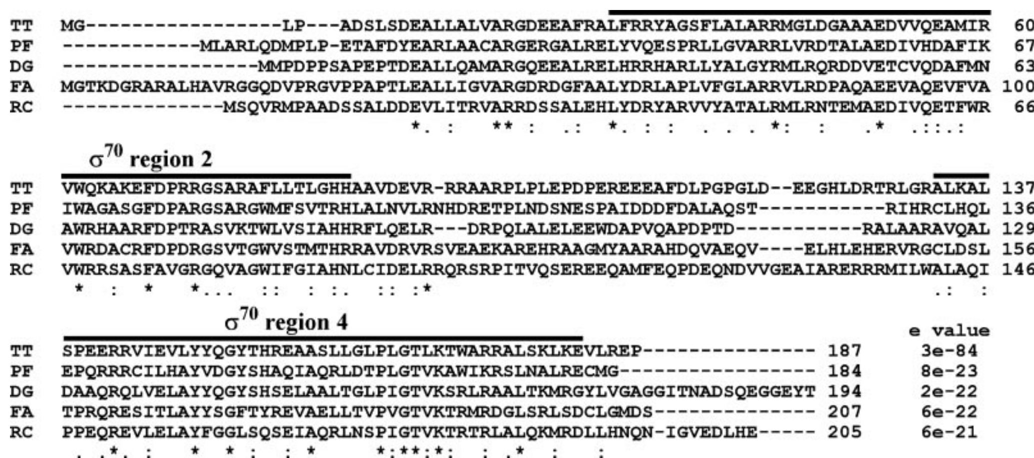


FIG. 1. Amino acid sequence alignment of *T. thermophilus* σ^E with representative homologous proteins. The amino acid sequences of *T. thermophilus* σ^E (TT), *Pseudomonas fluorescens* PfO-1 Pfl_3746 (PF), *Deinococcus geothermalis* DSM 11300 Dgeo_0606 (DG), *Frankia alni* ACN14a FRAAL3632 (FA), and *Roseiflexus castenholzii* DSM 13941 RcasDRAFT_1056 (RC) were aligned using Clustal W (28). Identical residues in all sequences (*), conserved substitutions (:), and semiconserved substitutions (.) are indicated. The e values obtained with the BLAST search are indicated on the right. The regions similar to regions 2 and 4 of σ^{70}-type σ factors, identified in the conserved domain database search, are indicated.

mRNA in vivo during cultivation in a rich medium consisting of 0.8% polypeptone, 0.4% yeast extract, 0.2% NaCl, 0.4 mM $CaCl_2$, and 0.4 mM $MgCl_2$, which was adjusted to pH 7.2 with NaOH at 70°C, using GeneChip technology as described previously (23). The mRNA was constitutively expressed under the experimental conditions used, and the level varied ~1.5-fold (Fig. 2). The σ^E-disruptant (Δ*sigE*) strain isolated in basically the same way as that described previously (7, 23) was not lethal, although the lag phase was longer than that in the case of the wild type, indicating that this gene is not essential for this strain (Fig. 2).

In order to determine the RNAP-binding ability of *T. thermophilus* σ^E, the *sigE* gene was cloned and expressed in an *E. coli* cell-free system (12, 31) in the absence or presence of the *T. thermophilus* RNAP core enzyme (Fig. 3). The *sigE* gene was amplified by genomic PCR and cloned into the pET-15b′ vector, which was constructed by replacing the NcoI-NdeI small fragment of the pET-15b vector (Novagen, Madison, WI) with a linker, 5′-CATGGGCCATCATCATCATCATCA-3′ and 5′-TATGATGATGATGATGATGGCC-3′. The resulting plasmid, pET-His-sigE, carrying a σ^E with an N-terminal uncleav-

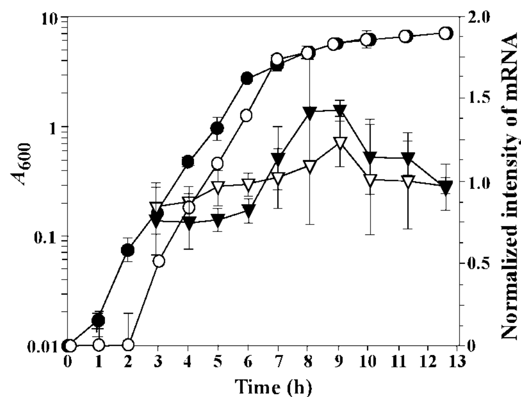


FIG. 2. Three clones of the wild-type (closed circles) and of the Δ*sigE* (open circles) strains of *T. thermophilus* were each grown in a rich medium, and the A_{600} values obtained at the indicated times are expressed as means ± standard deviations (SD), respectively. The expression levels of the *T. thermophilus sigE* (TTHB211) (open triangles) and TTHB212 (closed triangles) mRNAs in the wild-type strain were investigated at the indicated times using GeneChip technology, as described previously (23), and are expressed as normalized intensities ± SD.

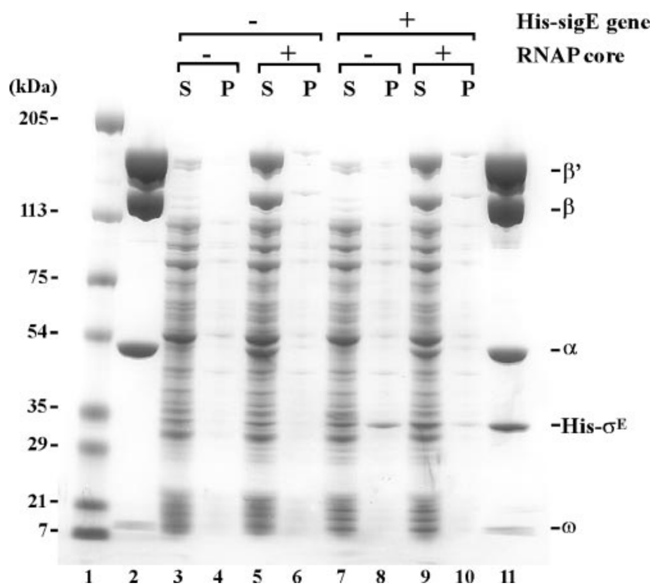


FIG. 3. Expression of *T. thermophilus* σ^E and purification of *T. thermophilus* RNAP-σ^E holoenzyme ($E\sigma^E$). N-terminal His-tagged *T. thermophilus* σ^E (His-σ^E) was synthesized by means of *E. coli* coupled transcription-translation cell-free protein synthesis in the absence (lanes 7 and 8) or presence (lanes 9 and 10) of the *T. thermophilus* RNAP core enzyme. After the reaction, each sample was fractionated into a supernatant (S) and a precipitate (P), which were analyzed by SDS-PAGE. Control samples without the *sigE* gene are shown (lanes 3 to 6). The $E\sigma^E$ synthesized in vitro was purified and analyzed by SDS-PAGE (lane 11). Lane 1, molecular mass markers in kDa; lane 2, *T. thermophilus* RNAP core enzyme.

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

able His_6 tag, Met-Gly-$(His)_6$, under the control of the T7 promoter, was added to the reaction mixture at the concentration of 8 μg/ml, with or without 2 mg/ml of the core enzyme that was purified from *T. thermophilus* HB8 cells, as described previously (30). After the reaction mixture was incubated at 30°C for 4 h, each sample was centrifuged at 16,000 × *g* for 10 min, and then the supernatant and precipitate were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in a precast 5 to 20% (wt/wt) polyacrylamide gel (ATTO Corp., Tokyo, Japan), according to the method of Laemmli (13), and the gel was stained with Coomassie brilliant blue R-250 (Fig. 3). In the absence of the core enzyme, the N-terminal His-tagged σ^E (His-σ^E) was present in the precipitate of the expression system (Fig. 3, lanes 7 and 8). In contrast, when the protein was expressed in the presence of the core enzyme, it was not precipitated and instead was observed mostly in the supernatant (Fig. 3, lanes 9 and 10). The reaction mixture (6 ml) containing both the His-σ^E and the core enzyme was dialyzed against buffer A (20 mM Tris-HCl [pH 7.7], 1 mM EDTA, 5% glycerol, 5 mM 2-mercaptoethanol, 0.1 mM phenylmethanesulfonyl fluoride, 1 mM benzamidine hydrochloride hydrate) and then centrifuged at 16,000 × *g* for 10 min at 4°C. The supernatant was applied to a 120-ml Q Sepharose (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) column preequilibrated with buffer A. The column was washed with buffer A, and then the bound proteins were eluted with a 1-liter linear gradient of 0 to 1 M NaCl in buffer A. The fractions containing both the His-σ^E and the core enzyme were collected and dialyzed against buffer A and then applied to a MonoQ HR (10/10) column (GE Healthcare Bio-Sciences Corp.) preequilibrated with buffer B (20 mM Tris-HCl [pH 7.7], 5% glycerol, 5 mM 2-mercaptoethanol, 0.1 mM phenylmethanesulfonyl fluoride, 1 mM benzamidine hydrochloride hydrate). The column was washed with buffer B, and then the bound proteins were eluted with a 165-ml linear gradient of 0 to 0.8 M NaCl in buffer B. The fractions containing both the His-σ^E and the core enzyme were collected and applied to a 5-ml nickel resin (ProBond resin; Invitrogen, Carlsbad, CA) column preequilibrated with buffer C (20 mM Tris-HCl [pH 7.5], 0.4 M NaCl) containing 5 mM imidazole. The column was washed with the same buffer, and then the bound proteins were eluted with buffer C containing 0.2 M imidazole. The fractions containing both the His-σ^E and the core enzyme were collected and then applied to a Hi-Load Superdex 200 16/60 HR column (GE Healthcare Bio-Sciences Corp.) preequilibrated with 20 mM Tris-HCl (pH 7.7), containing 0.3 M NaCl. As a result, the RNAP complexed with His-σ^E was purified >95% (Fig. 3, lane 11). These results suggest that *T. thermophilus* σ^E binds to the *T. thermophilus* RNAP core enzyme to form the holoenzyme $E\sigma^E$.

In order to determine the activity of the *T. thermophilus* $E\sigma^E$ synthesized in vitro, a runoff transcription assay was performed (Fig. 4). An 85-bp fragment derived from the upstream region of the *sigE* (TTHB211) gene (position 218783 to 218867 of pTT27) was cloned into the BamHI-EcoRI sites of pUC19. Then, various lengths of template DNA containing the upstream region of the *sigE* gene were prepared using 5′-GATCGGTGCGGGCCTCTTCG-3′ and 5′-ACCCATCCCTACGGCCGGAG-3′ (for template a), 5′-GGTGGTGGATGAGGCCAAAC-3′ (for template b), 5′-GAGGCCAAACCGGCCGGGCT-3′ (for template c), 5′-CAAACCGGCCGGGCTTTCCG-3′ (for template d), and 5′-CGGCCGGGCTTTCCGCCCGT-3′ (for template e) as primers (Fig. 4A). The amplified fragments were purified as described previously (23) and used for the following in vitro transcription assay. The assay was performed in a 15-μl reaction mixture as reported previously (23), except that the composition of the reaction mixture was buffer D [50 mM glycine-NaOH (pH 8.5), 0.1 mM EDTA, 5 mM dithiothreitol, 18 mM $MgCl_2$, 50 mM KCl, 1 mM each of spermidine and *N*,*N*′-bis(3-aminopropyl)-1,3-propanediamine (thermine)], 0.3 mM each ribonucleotide triphosphate (rNTP), 1.5 μCi of [α-^{32}P]CTP (GE Healthcare Bio-Sciences Corp.), 0.2 μM template DNA, 50 nM *T. thermophilus* $E\sigma^E$, and 50 μg/ml bovine serum albumin. After the reaction at 55°C for 10 min, samples were analyzed on a 10% polyacrylamide gel containing 8 M urea, with visualization by autoradiography. We found that template a was efficiently transcribed by $E\sigma^E$ (Fig. 4B). The same extent of transcription was observed even when the 5′-terminal 34 base pairs were deleted from the full-length template (Fig. 4B). The efficiency of the transcription was dramatically reduced when five more base pairs were deleted from the 5′ terminus; furthermore, transcription was hardly observed at all when five more base pairs were deleted (Fig. 4B). These results suggest that the promoter recognized by $E\sigma^E$ resides within 42 bp upstream of the transcription start site (Fig. 5B).

In order to find other candidates for σ^E-regulated genes, we selected genes that showed decreased expression in the Δ*sigE* strain compared with that in the wild type by using GeneChip technology. The three wild-type and the three Δ*sigE* strains were cultured for 8 h (A_{600} = 3 to 4) in the rich medium. From each strain, crude RNA was extracted, and then cDNA was synthesized from 10 μg of the total RNA, followed by fragmentation and labeling with biotin-dideoxy UTP, as described previously (23). The 3′-terminal-labeled cDNA (2 μg) was hybridized to a TTHB8401 GeneChip (Affymetrix, Santa Clara, CA), and then the array was washed, stained, and scanned as described previously (23). The image data for the three wild-type and three Δ*sigE* strains were scaled to a 2,238 ORF target intensity as described previously (23). Then, these data were normalized through the following three normalization steps using a GeneSpring GX 7.3.1 program (Agilent Technologies, Santa Clara, CA), data transformation (set measurements of less than 0.01 to 0.01), followed by per chip normalization (normalize as to median), and per gene normalization (normalize using the wild-type data as the control sample) (according to the GeneSpring manual; http://www.bimcore.emory.edu/Services/GeneSpring/Docs/GeneSpringManual.pdf; Agilent Technologies). The false discovery rate (*q* value) (24, 25) of the differences observed between the normalized intensities of the wild-type and those of the Δ*sigE* strains was calculated using an R program (http://cran.us.r-project.org/). The expression level of the *sigE* mRNA in the Δ*sigE* strain relative to that in the wild-type strain was 0.012 (*q* value = 0.067), indicating that the *sigE* gene was disrupted. Upstream of the genes and gene clusters that showed decreased expression in the Δ*sigE* strain (*q* value < 0.13) (see Table S1 in the supplemental material), we searched for sequences similar to that of the σ^E-dependent promoter upstream of TTHB211 (position 218817 to 218862 of pTT27), by using GENETYX software (GENETYX

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

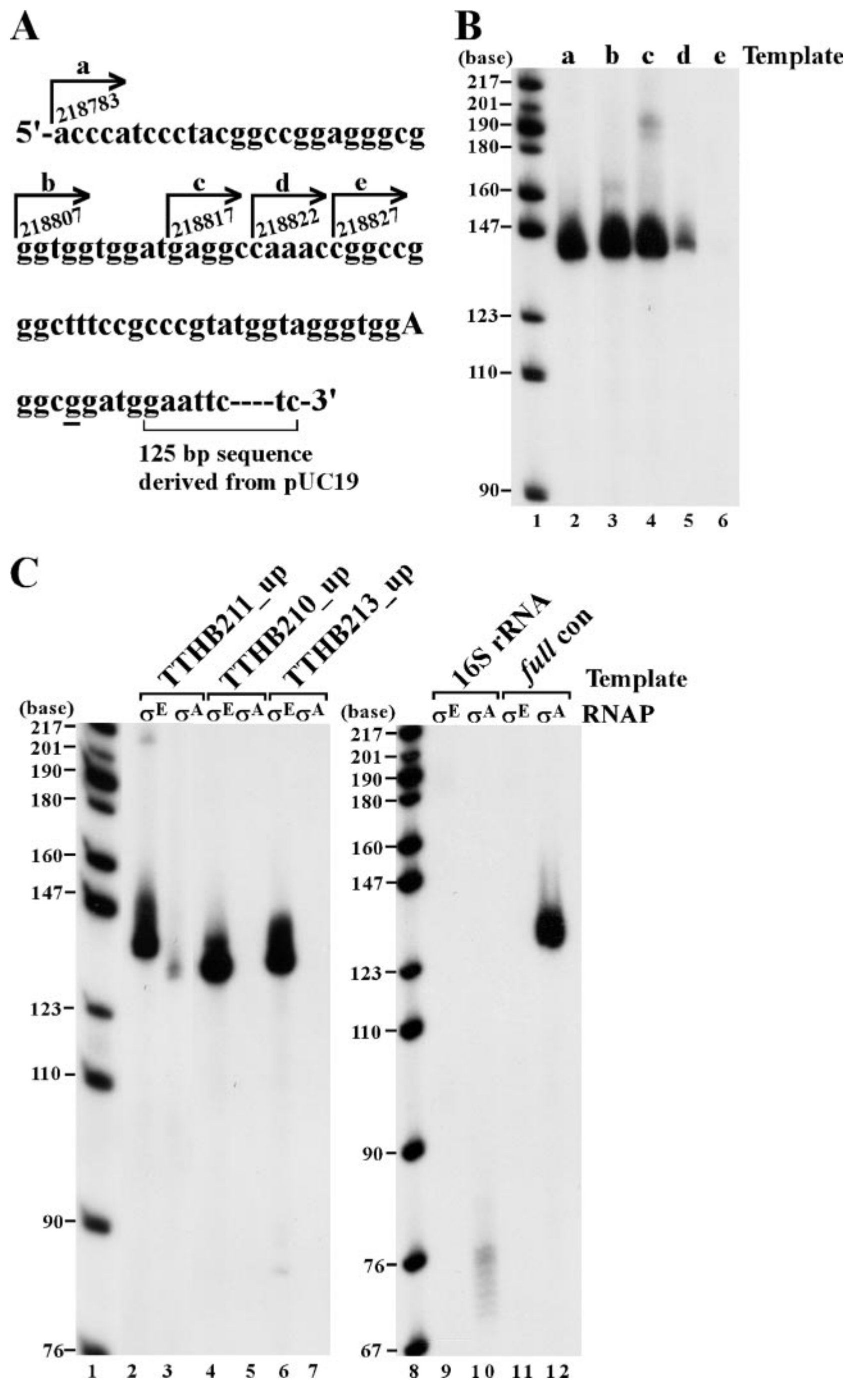


FIG. 4. (A) Upstream sequence of the *T. thermophilus sigE* (TTHB211) gene, which was used as the template for the runoff transcription assays. The numerals represent genome positions in megaplasmid pTT27. The transcriptional start site of Eσ^E is capitalized and that of Eσ^A is underlined. (B) Runoff transcription assay performed with *T. thermophilus* Eσ^E and templates a, b, c, d, and e is shown in panel A. After the reaction, samples were analyzed by PAGE, followed by autoradiography. Lane 1, [α-^{32}P]dCTP-labeled MspI fragments of pBR322. (C) The runoff transcription assay was performed with *T. thermophilus* Eσ^E (σ^E) (lanes 2, 4, 6, 9, and 11) or Eσ^A (σ^A) (lanes 3, 5, 7, 10, and 12) involving the DNA fragment containing the region upstream of TTHB211 (TTHB211_up), TTHB210 (TTHB210_up), or TTHB213 (TTHB213_up), the *T. thermophilus* 16S rRNA gene (16S rRNA), or the *E. coli* consensus promoter (*full* con) as a template. After the reaction, samples were analyzed by PAGE, followed by autoradiography. Lanes 1 and 8, [α-^{32}P]dCTP-labeled MspI fragments of pBR322.

Corp., Tokyo, Japan). Among 41 sequences, we found two such sequences upstream of the TTHB210 and the TTHB213 genes, the expression levels of which in the Δ*sigE* strain relative to that in the wild type were 0.095 (q value = 0.125) and 0.157 (q value = 0.079), respectively. In order to confirm the possibility of a σ^E-dependent promoter, we performed an in vitro transcription experiment using DNA containing the region upstream of TTHB210 or TTHB213 as

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

```
TTHB210_up  5'-ggcctcatccaccacccgccctccggccgtagggatgggtggagcCg-3'
TTHB211_up  5'-gaggccaaaccggccgggctttccgcccgtatggtagggtggAggcg-3'
TTHB213_up  5'-cttgacatccaaactccccctcgggccgtagggaagggcgagggGg-3'
Consensus   5'-snnsncawmcmnnnmyssscyytcsgsccgtakggwwgggy-3'
                     -35                     -10
```

FIG. 5. (A) Identification of the transcription start site. The RNA transcribed by *T. thermophilus* Eσ^E (lane 5) or Eσ^A (lane 6) from the genes containing the sequences upstream of TTHB210 (TTHB210_up), TTHB211 (TTHB211_up), and TTHB213 (TTHB213_up), respectively, was reverse transcribed, and the nucleotide sequence of the template DNA was determined by the dideoxy-mediated chain termination method (lanes 1 to 4). After the reaction, samples were analyzed by PAGE, followed by autoradiography. The sample volume applied to lane 6 for TTHB211_up was twofold more than that applied to lane 5 for TTHB211_up. The 3′ terminus of the cDNA is indicated by an asterisk. (B) The nucleotide sequence alignment of predicted promoter regions recognized by *T. thermophilus* Eσ^E found in TTHB210_up, TTHB211_up, and TTHB213_up is shown. The transcription start site of Eσ^E is capitalized and that of Eσ^A is underlined. The conserved bases are boxed, and the consensus sequence (Consensus) is indicated. The predicted −10 and −35 regions are shown in bold. Nucleotides y, k, m, w, and s represent t or c, g or t, a or c, a or t, and g or c, respectively. Nucleotide n represents g, a, t, or c. (C) Schematic view of the *T. thermophilus* σ^E regulon found in this study. The arrowhead indicates the ORF. The arrows indicate the promoter positions and directions of transcription by Eσ^E (σ^E; solid line) or Eσ^A (σ^A; broken line).

the template in the same way as that described above (Fig. 4C). We observed that *T. thermophilus* Eσ^E transcribed these templates, indicating that each template contains a σ^E-dependent promoter. Preparation of the DNA templates used for the transcription assays was performed as follows. For the construction of plasmids containing the upstream regions of TTHB210 and TTHB213, oligonucleotides 5′-ATATAggatccCCGGTTTGGCCTCATCCACCACCCGCCCTCCGGCC-3′ (the BamHI site is in lower case) and 5′-ATATGgaattcGGCTCCACCCATCCCTACGGCCGGAGGGCGGGTGGT-3′ (the EcoRI site is in lower case) and 5′-ATATAggatccTAAAGGCTTGACATCCAAACTCCCCCCTCGGGCCG-3′ and 5′-ATATGgaattcCCCCTCGCCCTTCCCTACGGCCCGAGGGGGGAGTTTG-3′, respectively, were annealed, and the partially duplex oligonucleotides were extended, digested with BamHI-EcoRI, and cloned into pUC19, respectively, as described previously (23). Using each plasmid as a template, PCR was performed to prepare template DNA for the transcription assay using 5′-TGCATGCCTGCAGGTCGACT-3′ and 5′-GATCGGTGCGGGCCTCTTCG-3′ as primers.

Based on nucleotide sequence alignment of the three promoter regions and a comparison with known ECF-type promoters (9), a consensus sequence of the promoter for *T. thermophilus* σ^E, 5′-CA(A/T)(A/C)C(A/C)-N_{15}-CCGTA-3′, was proposed (Fig. 5B). The "CGTA" sequence in the −10 region is the same as that in the consensus promoter sequence recognized by one of the *Bacillus subtilis* ECF σ factors, σ^w (9). The "AAC" motif, which is often found in the −35 region of a promoter recognized by ECF σ factors (9), is present in the region of the TTHB211 promoter (Fig. 5B).

We investigated the specificity of the promoter recognition of the *T. thermophilus* Eσ^E and the *T. thermophilus* Eσ^A holoenzyme that transcribes housekeeping genes (Fig. 4C). Eσ^A was purified from *T. thermophilus* HB8 cells as described previously (30). The DNA fragment containing the region upstream of the TTHB210 or the TTHB213 gene was hardly transcribed by Eσ^A, whereas Eσ^A transcribed the DNA fragment containing the region upstream of the TTHB211 gene, although the efficiency was lower than that when Eσ^E was used. The nucleotide sequence of the predicted −35 region of the TTHB211 promoter is 5′-CAAACC-3′, which differs from the 5′-CATCCA-3′ sequence of TTHB210 and TTHB213; thus, the sequence of the −35 region might determine the specificity for σ factors. Conversely, Eσ^E did not transcribe the DNA template containing the *T. thermophilus*-derived 16S rRNA gene and the *E. coli* consensus promoter named *full* con (4), which were recognized by Eσ^A (Fig. 4C). These results indicate that the specificity of promoter recognition differs between Eσ^E and Eσ^A. Preparation of a DNA template containing the 16S rRNA gene or *full* con was performed as follows. Oligonucleotides 5′-ATATGgatccAAATGTGATCTAGATCACATTTGAGATCCGCTGCCCTCGCAA-3′ and 5′-TATATATgcgcAGAAAAGCCATGCTATCAATCCCCCTCCTTTTTGTCAAGGCTTGCGAGGGCAGCGGATCTC-3′ (the HhaI site is in lower case) were annealed, and the partially duplex oligonucleotide was extended and digested with BamHI

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

and HhaI. The fragment was cloned into the BamHI-SacI sites of pUC19 together with an oligonucleotide linker, 5′-CGGC GAAAGCCCAAGCAGGGGGATCTTGAAAAGGGGAG AAGGCAGGCCAAGCATCTGAGCT-3′ and 5′-CAGATG CTTGGCCTGCCTTCTCCCCTTTTCAAGATCCCCCTGC TTGGGCTTTCGCCGCG-3′, to construct a plasmid containing the 16S rRNA gene. Using this plasmid as a template, PCR was performed to prepare a template DNA using 5′-TGCATGCC TGCAGGTCGACT-3′ and 5′-CGAGCTCAGATGCTTGGC CT-3′ as primers. Oligonucleotides 5′-ATATggatccTAATGT GAGTTAGCTCACTCATTAGGCACCCCAGGCTT-3′ and 5′-ATATgaattcTTCCACACAGTATAGCACAATTTAACACTT TTGTCAAGCCTGGGGTGCCTAATGA-3′ were annealed, and the partially duplex oligonucleotide was extended, digested with BamHI-EcoRI, and then cloned into pUC19 to construct a plasmid containing the *full* con. Using this plasmid as a template, PCR was performed to prepare template DNA, using 5′-TGCATGCCTGCAGGTCGACT-3′ and 5′-GATCGG TGCGGGCCTCTTCG-3′ as primers.

We determined the transcription start sites of the TTHB210, TTHB211, and TTHB213 genes by primer extension analysis with RNA transcribed in vitro from the DNA templates containing these promoters, using basically the same method as described previously (23), except for the composition of the reaction mixture for the initial transcription reaction, which was buffer D, 0.2 μM template DNA, 0.1 μM *T. thermophilus* $E\sigma^E$ or $E\sigma^A$, 0.3 mM each rNTP, and 50 μg/ml bovine serum albumin (Fig. 5). The nucleotide sequence of the template DNA was determined by the dideoxy-mediated chain termination method (22), as described previously (23). Samples were analyzed on an 8% polyacrylamide gel containing 8 M urea, with visualization by autoradiography (Fig. 5A). The transcription start sites are shown in Fig. 5B.

In many cases, an ECF σ factor is autoregulated and cotranscribed with a transmembrane anti-σ factor with an extracytoplasmic sensory domain and an intracellular inhibitory domain that directly binds to the ECF σ factor and controls its activity (8–10). In the case of *T. thermophilus*, σ^E also induces its own synthesis. Genome analysis of *T. thermophilus* HB8 (NCBI accession number NC_006462) indicated that the *sigE* gene forms an operon with the downstream TTHB212 gene (Fig. 5C). The expression profile of TTHB212 mRNA during cultivation in rich medium was similar to that of the *sigE* gene (Fig. 2), supporting the idea that the two genes form an operon. The transmembrane prediction program TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) revealed that residues 1 to 77, 78 to 100, and 101 to 205 of the TTHB212 protein were cytoplasmic, transmembrane, and extracytoplasmic domains, respectively (data not shown). Therefore, the TTHB212 protein might be an anti-σ factor for σ^E, although known anti-σ factors were not found with a BLAST search with the TTHB212 protein as a query.

In order to determine the cellular role of *T. thermophilus* σ^E, we examined the amino acid sequence features of the products of the other two σ^E-regulated genes because these gene products are also biochemically or biophysically uncharacterized. According to a BLAST search, the TTHB213 protein is a hypothetical protein that shows the highest e value, 5e-96, for the hypothetical protein TT_P0169 from *T. thermophilus* HB27. Interestingly, the TTHB213 protein contains a twin arginine translocation pathway signal-like sequence, ^{6}ARRQFM$(X)_{26}$P, at the N terminus (14) and exhibits a high e value for the RxyI_1665 (6e-18) and RxyI_1666 (3e-11) proteins from *Rubrobacter xylanophilus*, which harbors a similar twin arginine motif-like sequence. The functions of many ECF σ factors are associated with some aspect of the cell surface or transport (1, 2, 9, 17, 21); thus, the TTHB213 protein might be a cell surface protein. By use of the TTHB210 protein as a query, poorly conserved proteins other than the uncharacterized ones, i.e., *T. thermophilus* HB27 TT_P0163 (1e-54) and *Mesorhizobium* sp. strain BNC1 Meso_1245 (2e-15), were found; thus, the TTHB210 protein is an unknown protein characteristic of *T. thermophilus*.

In *T. thermophilus* HB8, the aforementioned genes are basically expressed under the experimental conditions used. The genes that showed altered expression in the Δ*sigE* strain (see Table S1 in the supplemental material) may include those that are not directly regulated by σ^E but are affected by the gene products that are under direct control of σ^E, although it is hard to distinguish them from false positives. As observed for the cases of many ECF σ factors (9, 21), the expression level of *T. thermophilus* σ^E might increase in response to some environmental signal. Identification of such a signal that leads to the increased expression of *T. thermophilus* σ^E may shed light on the roles of the σ factor and the σ^E-regulated gene products in the physiology of *T. thermophilus*.

Microarray data accession number. The microarray data discussed in this study have been deposited in the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) and are accessible under GEO Series accession number GSE8781.

We thank Aiko Kashihara, Satoshi Kira, and Noriko Nakagawa for excellent support with the GeneChip analysis.

This work was supported in part by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), by the National Project on Protein Structural and Functional Analyses, and by the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. **Bashyam, M. D., and S. E. Hasnain.** 2004. The extracytoplasmic function sigma factors: role in bacterial pathogenesis. Infect. Genet. Evol. **4:**301–308.
2. **Braun, V., and S. Mahren.** 2005. Transmembrane transcriptional control (surface signalling) of the *Escherichia coli* Fec type. FEMS Microbiol. Rev. **29:**673–684.
3. **Ebright, R. H.** 2000. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. J. Mol. Biol. **304:**687–698.
4. **Gaal, T., W. Ross, S. T. Estrem, L. H. Nguyen, R. R. Burgess, and R. L. Gourse.** 2001. Promoter recognition and discrimination by $E\sigma^S$ RNA polymerase. Mol. Microbiol. **42:**939–954.
5. **Gross, C. A., C. Chan, A. Dombroski, T. Gruber, M. Sharp, J. Tupy, and B. Young.** 1998. The functional and regulatory roles of sigma factors in transcription. Cold Spring Harbor Symp. Quant. Biol. **63:**141–155.
6. **Gruber, T. M., and C. A. Gross.** 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. **57:**441–466.
7. **Hashimoto, Y., T. Yano, S. Kuramitsu, and H. Kagamiyama.** 2001. Disruption of *Thermus thermophilus* genes by homologous recombination using a thermostable kanamycin-resistant marker. FEBS Lett. **506:**231–234.
8. **Helmann, J. D.** 1999. Anti-sigma factors. Curr. Opin. Microbiol. **2:**135–141.
9. **Helmann, J. D.** 2002. The extracytoplasmic function (ECF) sigma factors. Adv. Microb. Physiol. **46:**47–110.
10. **Hughes, K. T., and K. Mathee.** 1998. The anti-sigma factors. Annu. Rev. Microbiol. **52:**231–286.
11. **Kazmierczak, M. J., M. Wiedmann, and K. J. Boor.** 2005. Alternative sigma factors and their roles in bacterial virulence. Microbiol. Mol. Biol. Rev. **69:**527–543.
12. **Kigawa, T., T. Yabuki, N. Matsuda, T. Matsuda, R. Nakajima, A. Tanaka,**

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

and S. Yokoyama. 2004. Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. J. Struct. Funct. Genomics **5:**63–68.
13. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
14. **Lee, P. A., D. Tullman-Ercek, and G. Georgiou.** 2006. The bacterial twin-arginine translocation pathway. Annu. Rev. Microbiol. **60:**373–395.
15. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843–3849.
16. **Lonetto, M. A., K. L. Brown, K. E. Rudd, and M. J. Buttner.** 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proc. Natl. Acad. Sci. USA **91:**7573–7577.
17. **Missiakas, D., and S. Raina.** 1998. The extracytoplasmic function sigma factors: role and regulation. Mol. Microbiol. **28:**1059–1066.
18. **Mittenhuber, G.** 2002. An inventory of genes encoding RNA polymerase sigma factors in 31 completely sequenced eubacterial genomes. J. Mol. Microbiol. Biotechnol. **4:**77–91. (Erratum, **4:**514.)
19. **Omelchenko, M. V., Y. I. Wolf, E. K. Gaidamakova, V. Y. Matrosova, A. Vasilenko, M. Zhai, M. J. Daly, E. V. Koonin, and K. S. Makarova.** 2005. Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans*: divergent routes of adaptation to thermophily and radiation resistance. BMC Evol. Biol. **5:**57.
20. **Oshima, T., and K. Imabori.** 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) com. nov., a non-sporulating thermophilic bacterium from a Japanese thermal spa. Int. J. Syst. Bacteriol. **24:**102–112.
21. **Raivio, T. L., and T. J. Silhavy.** 2001. Periplasmic stress and ECF sigma factors. Annu. Rev. Microbiol. **55:**591–624.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
23. **Shinkai, A., S. Kira, N. Nakagawa, A. Kashihara, S. Kuramitsu, and S. Yokoyama.** 2007. Transcription activation mediated by a cyclic AMP receptor protein from *Thermus thermophilus* HB8. J. Bacteriol. **189:**3891–3901.
24. **Storey, J. D.** 2002. A direct approach to false discovery rates. J. R. Stat. Soc. B **64:**479–498.
25. **Storey, J. D., and R. Tibshirani.** 2003. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA **100:**9440–9445.
26. **Sweetser, D., M. Nonet, and R. A. Young.** 1987. Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. USA **84:**1192–1196.
27. **Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin, and D. A. Natale.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:**41.
28. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
29. **Vassylyev, D. G., S. Sekine, O. Laptenko, J. Lee, M. N. Vassylyeva, S. Borukhov, and S. Yokoyama.** 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. Nature **417:**712–719.
30. **Vassylyeva, M. N., J. Lee, S. I. Sekine, O. Laptenko, S. Kuramitsu, T. Shibata, Y. Inoue, S. Borukhov, D. G. Vassylyev, and S. Yokoyama.** 2002. Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from *Thermus thermophilus*. Acta Crystallogr. D **58:** 1497–1500.
31. **Wada, T., M. Shirouzu, T. Terada, Y. Ishizuka, T. Matsuda, T. Kigawa, S. Kuramitsu, S. Y. Park, J. R. Tame, and S. Yokoyama.** 2003. Structure of a conserved CoA-binding protein synthesized by a cell-free system. Acta Crystallogr. D **59:**1213–1218.
32. **Wnendt, S., R. K. Hartmann, N. Ulbrich, and V. A. Erdmann.** 1990. Isolation and physical properties of the DNA-directed RNA polymerase from *Thermus thermophilus* HB8. Eur. J. Biochem. **191:**467–472.

Downloaded from http://jb.asm.org/ on January 18, 2021 by guest

JB Journals.ASM.org

AUTHOR CORRECTION

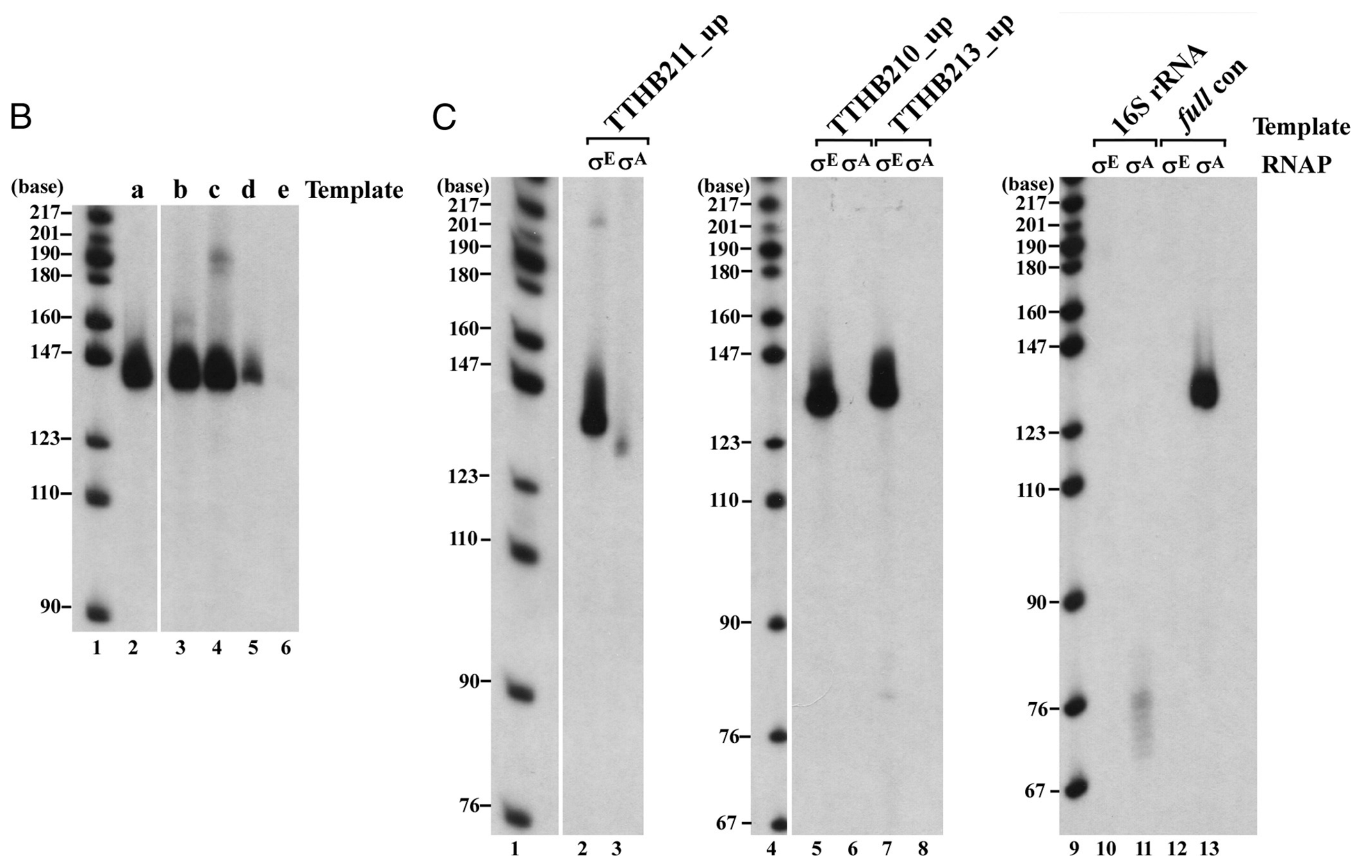
Correction for Shinkai et al., Identification of Promoters Recognized by RNA Polymerase-σ^E Holoenzyme from *Thermus thermophilus* HB8

Akeo Shinkai,[1] Naomi Ohbayashi,[2] Takaho Terada,[2] Mikako Shirouzu,[2] Seiki Kuramitsu,[1,3] and Shigeyuki Yokoyama[1,2,4]

RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan[1]; Genomic Sciences Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan[2]; Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan[3]; and Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan[4]

Volume 189, no. 23, pages 8758-8764, 2007. We corrected Fig. 4B and C of this article, to clarify that these figures were created by the assembly of several images. The raw data used for the creation of these figures are shown in Fig. S1 in the supplemental material.

Page 8761: The legend to Fig. 4C should read as follows and Fig. 4B and C should appear as shown below. "(C) Runoff transcription assay performed with *T. thermophilus* Eσ^E (σ^E) (lanes 2, 5, 7, 10, and 12) or Eσ^A (σ^A) (lanes 3, 6, 8, 11, and 13) involving the DNA fragment containing the region upstream of TTHB211 (TTHB211_up), TTHB210 (TTHB210_up), or TTHB213 (TTHB213_up), the *T. thermophilus* 16S rRNA gene (16S rRNA), or the *E. coli* consensus promoter (*full* con) as a template. After the reaction, samples were analyzed by PAGE followed by autoradiography. Lanes 1, 4, and 9, [α-^{32}P]dCTP-labeled MspI fragments of pBR322. The raw data are shown in Fig. S1 in the supplemental material."



Supplemental material: Figure S1 in the supplemental material is posted at http://dx.doi.org/10.1128/JB.01076-07.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JB.01953-14