Phenylacetyl Coenzyme A Is an Effector Molecule of the TetR Family Transcriptional Repressor PaaR from *Thermus thermophilus* HB8[▽][†]

Keiko Sakamoto,[1] Yoshihiro Agari,[1] Seiki Kuramitsu,[1,2] and Akeo Shinkai[1]*

RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan,[1] and Department of Biological Sciences, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan[2]

Phenylacetic acid (PAA) is a common intermediate in the catabolic pathways of several structurally related aromatic compounds. It is converted into phenylacetyl coenzyme A (PA-CoA), which is degraded to general metabolites by a set of enzymes. Within the genome of the extremely thermophilic bacterium *Thermus thermophilus* HB8, a cluster of genes, including a TetR family transcriptional regulator, may be involved in PAA degradation. The gene product, which we named *T. thermophilus* PaaR, negatively regulated the expression of the two operons composing the gene cluster *in vitro*. *T. thermophilus* PaaR repressed the target gene expression by binding pseudopalindromic sequences, with a consensus sequence of 5′-CNAACGNNCGTTNG-3′, surrounding the promoters. PA-CoA is a ligand of PaaR, with a proposed binding stoichiometry of 1:1 protein monomer, and was effective for transcriptional derepression. Thus, PaaR is a functional homolog of PaaX, a GntR transcriptional repressor found in *Escherichia coli* and *Pseudomonas* strains. A three-dimensional structure of *T. thermophilus* PaaR was predicted by homology modeling. In the putative structure, PaaR adopts the typical three-dimensional structure of the TetR family proteins, with 10 α-helices. A positively charged surface at the center of the molecule is similar to the acyl-CoA-binding site of another TetR family transcriptional regulator, *T. thermophilus* FadR, which is involved in fatty acid degradation. The CoA moiety of PA-CoA may bind to the center of the PaaR molecule, in a manner similar to the binding of the CoA moiety of acyl-CoA to FadR.

The phenylacetic acid (PAA) degradation pathway is a convergent route of several catabolic pathways for structurally related aromatic compounds containing an even number of carbon atoms, such as styrene, ethylbenzene, 2-phenylethylamine, phenylacetyl esters, phenylacetyl amides, and tropic acid (17). PAA, a common intermediate in their catabolic pathways, is converted into phenylacetyl coenzyme A (PA-CoA), which is also a common intermediate in the catabolic pathways of other aromatic compounds, such as phenylalkanoates, polyhydroxyphenylalkanoates, and *trans*-styrylacetic acid (17). PA-CoA is degraded to general metabolites by a set of enzymes referred to as the PA-CoA catabolon core (7, 17, 20, 27). The *paa* genes, encoding the enzymes for PAA degradation, exist in several bacteria (17, 20, 27). The transcriptional regulation of the *paa* genes has been extensively studied in *Escherichia coli* (6, 7) and *Pseudomonas* strains (8). In these strains, PaaX, a GntR family transcriptional regulator, whose gene is located in a *paa* gene cluster, is a transcriptional repressor of the *paa* genes. In the absence of PA-CoA, PaaX binds the cognate operators to repress the gene expression, while in the presence of PA-CoA, PaaX responds to this molecule and derepresses the gene. In *E. coli*, two more regulators, the cyclic AMP (cAMP) receptor protein and the integration host factor protein, are involved in controlling the expression of the *paa* genes (6). In *Pseudomonas* sp. strain Y2, PaaX and PA-CoA also control an upper pathway for the catabolism of

styrene with the StySR regulatory system (3). On the other hand, the *paa* gene clusters of several other bacteria, such as *Azoarcus evansii* (20), *Rhodococcus* sp. strain RHA1 (21), and *Burkholderia cenocepacia* K56-2 (10), lack *paaX* genes. Instead, each strain has a regulatory gene encoding a TetR family transcriptional regulator, named PaaR, in the *paa* gene cluster. Recently, the *Burkholderia* PaaR was reported to transcriptionally repress the *paa* operons *in vivo* (10).

Thermus thermophilus HB8, which belongs to the phylum *Deinococcus-Thermus*, is an extremely thermophilic bacterium isolated from the water at a Japanese hot spring and grows at an optimum temperature range of 65 to 72°C (22). Its genome is composed of the 1.85-Mbp chromosomal DNA, the 0.26-Mbp plasmid pTT27, and the 9.32-kbp plasmid pTT8, encoding 1,973, 251, and 14 open reading frames (ORFs), respectively (NCBI accession numbers NC_006461, NC_006462, and NC_006463, respectively). This strain has a putative *paa* gene cluster on the chromosomal DNA. *T. thermophilus* PaaI, which we renamed PaaD in this study, is encoded by *TTHA0965* in the gene cluster. It adopts a hotdog fold, as seen in certain CoA-binding enzymes, and displays thioesterase activity with several CoA compounds, including PA-CoA (14). The gene cluster contains a TetR family transcriptional regulator gene, *TTHA0973*, which is not a homolog of *paaX*. In this study, we biochemically characterized the TTHA0973 protein, which we named *T. thermophilus* PaaR.

MATERIALS AND METHODS

RT-PCR. Total RNA, isolated from the wild-type *T. thermophilus* HB8 strain cultured for 11.3 h in rich medium, was treated with DNase I, followed by ethanol precipitation, as described previously (26). Using the RNA (1 μg) as a template, the reverse transcription (RT) reaction was performed at 42°C for 20 min in a 20-μl reaction mixture, using a PrimeScript RT-PCR kit (Takara Bio)

---

* Corresponding author. Mailing address: 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.

according to the manufacturer's instructions. Using the reaction mixture (1 μl) as a template, PCR was performed in the presence of 0.2 μM (each) primer (see Table S1 in the supplemental material) in a 25-μl reaction mixture. Primers P01 and P02, which correspond to the genome positions 909204 to 909223 and 911723 to 911704, respectively, were used to detect the operon composed of *TTHA0963* to *TTHA0967*. The PCR analysis involved 40 cycles at 94°C for 30 s, 60°C for 30 s, and 72°C for 3.5 min in GC buffer I (Takara Bio). Primers P03 and P04, which correspond to the genome positions 916807 to 916788 and 914052 to 914071, respectively, were used to detect the operon composed of *TTHA0973* to *TTHA0968*. The PCR analysis involved 30 cycles at 94°C for 30 s, 63°C for 30 s, and 72°C for 3.5 min in GC buffer II (Takara Bio). In order to confirm the absence of genomic DNA contamination of the total RNA fraction used as the template, PCR was performed with no RT, as a control.

Expression plasmid construction. The *T. thermophilus paaR* (*TTHA0973*) gene was amplified by genomic PCR, using the primers P05 and P06 (see Table S1 in the supplemental material), and the amplified fragment was cloned under the control of the T7 promoter (NdeI-BamHI sites) in the *E. coli* expression vector pET-11a (Merck), to construct pET-ttPaaR (RIKEN BioResource Center; http://www.brc.riken.jp/inf/en/index.shtml). The 8th codon of the open reading frame in the expression plasmid was CTA, but that in the current version of the genomic sequence is CTC. To construct the expression plasmid pET-ttPaaR_His, encoding PaaR fused with a His_{12} tag at its C terminus, PCR was performed using pET-ttPaaR as the template and P07 and P08 as primers (see Table S1 in the supplemental material), and the amplified fragment was cloned under the control of the T7 promoter (NdeI-BamHI sites) in the *E. coli* expression vector pET-11a.

Purification of recombinant PaaR proteins. *E. coli* BL21(DE3) cells (Merck) harboring pET-ttPaaR were cultured at 37°C in 9 liters of Luria-Bertani (LB) broth, containing 50 μg ampicillin ml^{-1}, for 16 h. The cells were resuspended in 60 ml of 20 mM Tris-HCl (pH 8.0), containing 50 mM NaCl, 5 mM β-mercaptoethanol, and 0.1 mM phenylmethylsulfonyl fluoride, and were disrupted by sonication in ice water. The same volume of buffer, preheated at 70°C, was added to the cell lysate. This mixture was incubated for 10 min at 70°C and then ultracentrifuged (200,000 × *g*) for 1 h at 4°C. The supernatant was applied to a Resource PHE column (GE Healthcare), preequilibrated with 50 mM sodium phosphate buffer (pH 7.0) containing 1.5 M ammonium sulfate, and then the bound protein was eluted with a linear gradient of 1.5 to 0 M ammonium sulfate. The target fractions were collected and desalted by fractionation on a HiPrep 26/10 desalting column (GE Healthcare). The sample was then applied to a Resource Q column (GE Healthcare), preequilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.1 M NaCl, and the bound protein was eluted with a linear gradient of 0.1 to 0.5 M NaCl. The target fractions were collected and desalted by fractionation on a HiPrep 26/10 desalting column. The sample was then applied to a HiTrap heparin column (GE Healthcare), preequilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.1 M NaCl, and the bound protein was eluted with a linear gradient of 0.1 to 0.5 M NaCl. The target fractions were collected and desalted by fractionation on a HiPrep 26/10 desalting column. The sample was applied to a HiTrap heparin column preequilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.1 M NaCl, and eluted with the buffer containing 2 M NaCl, to concentrate the sample. The sample was then applied to a HiLoad 16/60 Superdex 200-pg (GE Healthcare) column, preequilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl. The target fractions were collected and concentrated with a Centriprep YM-10 filter (Millipore).

PaaR with the C-terminal His_{12} tag (PaaR-His) was expressed in *E. coli* Rosetta (DE3)pLysS cells (Merck) harboring pET-ttPaaR_His. The cells were cultured at 37°C in 12 liters of LB broth, containing 50 μg ampicillin ml^{-1} and 30 μg chloramphenicol ml^{-1}, for 16 h and were disrupted as described above, except that 0.5 M NaCl was contained in the disruption buffer. The sample was applied to a HisTrap HP column (GE Healthcare), preequilibrated with 20 mM Tris-HCl (pH 8.0) containing 0.5 M NaCl and 20 mM imidazole, and the bound protein was eluted with a linear gradient of 20 to 500 mM imidazole. The target fractions were collected and fractionated on HiTrap heparin and HiLoad 16/60 Superdex 75-pg columns, as described above. The sample was concentrated with a Centricon YM-10 filter (Millipore). The protein concentration was determined by measuring the absorbance at 280 nm (15).

Genomic selex. The genomic SELEX search for the PaaR-binding sequence was performed in basically the same manner as described by Shimada et al. (25). Genomic DNA from *T. thermophilus* HB8 was fragmented by sonication, and then the ~0.1- to 0.3-kbp DNA fragments were excised from a 1% agarose gel. The DNA fragments were treated with T4 DNA polymerase and were then cloned into the plasmid pBR322 (EcoRV site) to construct the *E. coli* DNA library. PCR was performed using the DNA library plasmids as templates and P09 and P10 as primers (see Table S1 in the supplemental material). The

amplified fragments were excised from a 1% agarose gel. Approximately 10 pmol (1.3 μg) of the DNA fragment and 20 pmol of the PaaR-His dimer were mixed in 0.1 ml of buffer A, comprising 20 mM Tris-HCl, pH 7.7, 3 mM magnesium acetate, 0.1 M KCl, and 1.25 μg bovine serum albumin ml^{-1}, and incubated for 20 min at 55°C. The nickel resin (0.1 ml; ProBond resin; Invitrogen), preequilibrated with buffer A, was added to the sample, and the mixture was incubated for 20 min at 55°C. The mixture was transferred to a column, which was washed with 5 ml of buffer A containing 5 mM imidazole, and the bound sample was eluted with 0.25 ml of buffer A containing 250 mM imidazole. The DNA was extracted from the eluted sample with phenol and ether. PCR was performed with the extracted DNA as templates and P09 and P10 as primers. The amplified fragments were purified by using a QIAquick PCR purification kit (Qiagen) and were used for the next round of genomic selex. The selected DNA fragments were cloned into the plasmid pT7Blue (Merck), and the nucleotide sequences were analyzed.

BIAcore biosensor assay. A DNA fragment (0.1 mM), biotinylated at the 5′ end of one strand, was diluted to 50 nM in 10 mM HEPES-NaOH (pH 7.4) buffer, containing 0.5 M NaCl, 3 mM EDTA, and 0.005% surfactant P20, and then applied to an SA biosensor chip (GE Healthcare), as described previously (1). All experiments for measuring the interaction between the DNA and *T. thermophilus* PaaR were performed at 25°C, using buffer B (10 mM HEPES-NaOH, pH 7.4, 3 mM EDTA, 0.005% surfactant P20) containing 0.3 M NaCl. The PaaR was diluted with the same buffer and then injected over the DNA surface at a flow rate of 20 μl min^{-1}. Sensorgrams were recorded and normalized to a baseline of 0 response unit (RU). An equivalent volume of each protein dilution was also injected over a nontreated surface, to determine the bulk refractive index background. At the end of each cycle, the bound protein was removed by injecting 0.2 ml of 3 M NaCl, to regenerate the chip. The association and dissociation rate constants (k_{on} and k_{off}, respectively) and the dissociation constant (K_d) values were determined by 1:1 Langmuir local fitting with the BIAevaluation 3.0 software (GE Healthcare).

ITC. Isothermal titration calorimetry (ITC) measurements were performed using a VP-ITC MicroCalorimeter (MicroCal) at 30°C. The ligand (50 μM) was dissolved in 20 mM Tris-HCl (pH 8.0) containing 0.15 M NaCl in the stirred syringe (300 rpm) and was titrated into 1.4 ml of 5 μM PaaR monomer in the same buffer. Aliquots (10 μl) of the ligand were injected 30 times at 0.42 μl s^{-1}, with a 4-min interval between the injections. The baseline was corrected by subtracting the trace of buffer injections from the raw calorimetric trace, and then the nonlinear one-site model was fitted to the baseline-corrected raw ITC data to determine the binding constant (K_b) value and the ligand-binding stoichiometry, using the ORIGIN (ver. 5.0) software (MicroCal).

In vitro **transcription assay. (i) Preparation of templates.** The construction of plasmids containing the upstream regions of the *TTHA0963* and *TTHA0973* genes was performed in basically the same manner as described previously (26), using oligonucleotides P11/P12 and P13/P14, respectively (see Table S1 in the supplemental material). Using each plasmid as the template, PCR was performed with the primers P15 and P16 (see Table S1 in the supplemental material) to prepare the template DNA for the transcription assay.

(ii) Runoff transcription. Assays were performed in 15-μl reaction mixtures, in the absence or presence of *T. thermophilus* PaaR, in basically the same manner as described previously (26). The template DNA was preincubated with or without PaaR at 55°C for 5 min. *T. thermophilus* RNA polymerase-σ^A holoenzyme (RNAP) was added, and the mixture was further incubated for 5 min. Transcription was initiated by the addition of 1.5 μCi [α-^{32}P]CTP and unlabeled ribonucleotide triphosphates. After further incubation for 10 min, the reaction was stopped, and the sample was fractionated on a 10% polyacrylamide gel containing 8 M urea and analyzed by autoradiography.

Identification of the transcriptional start site. Total RNA, isolated from wild-type *T. thermophilus* HB8 cells cultured at 70°C for 6 h in rich medium, was treated with DNase I, followed by ethanol precipitation, as described previously (26). Rapid amplification of cDNA ends (5′ RACE) was performed with a 5′-Full RACE Core Set (Takara Bio), according to the manufacturer's instructions. The first-strand cDNA was synthesized with 5′-phosphorylated primers P17 and P18 for the *TTHA0963* and *TTHA0973* genes, respectively (see Table S1 in the supplemental material). The RNA was digested by RNase H, and the single-stranded cDNA was ligated with T4 RNA ligase to construct DNA concatemers. PCR was performed using the DNA concatemers as templates and P19/P20 and P21/P22 (see Table S1 in the supplemental material) as primers for the *TTHA0963* and *TTHA0973* genes, respectively. The second PCR was performed using the primers P23/P24 and P25/P26 (see Table S1 in the supplemental material) for the *TTHA0963* and *TTHA0973* genes, respectively. The amplified DNA fragments were cloned into the plasmid pT7Blue (Merck), and the nucleotide sequences were analyzed.
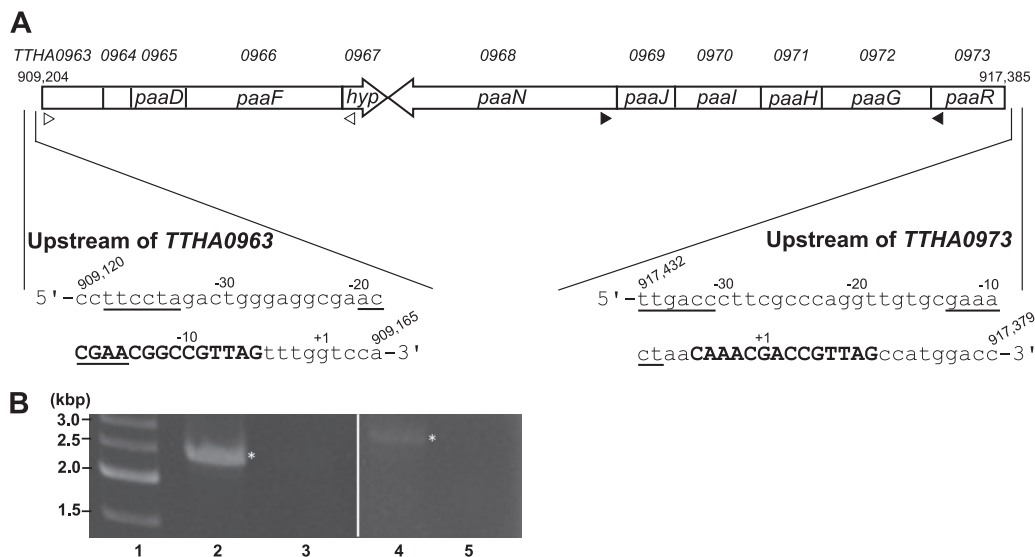
FIG. 1. Organization of the putative *paa* gene cluster in *T. thermophilus* HB8. (A) The locus tags (numerals preceded by *TTHA*) are denoted above the genes. The locations of the genes on the chromosomal DNA are shown. The putative *paa* genes, besides *paaR*, are named according to the consensus names defined by Luengo et al. (17); i.e., *paaD*, *paaF*, *paaN*, *paaJ*, *paaI*, *paaH*, and *paaG* encode a putative thioesterase, PA-CoA ligase, ring opening enzyme, ring-hydroxylating complex protein 4, ring-hydroxylating complex protein 3, ring-hydroxylating complex protein 2, and ring-hydroxylating complex protein 1, respectively. *paaR* and *hyp* encode a TetR family transcriptional regulator and a hypothetical protein, respectively. *TTHA0963* and *TTHA0964* encode ABC transporter ATP-binding protein-related proteins. The nucleotide sequences upstream of *TTHA0963* and *TTHA0973* are shown with their locations on the chromosomal DNA. Probable PaaR-binding pseudopalindromic sequences are indicated by bold capital letters. Possible -10 and -35 hexamer sequences of the promoters are underlined. The representative transcriptional start site ($+1$) of each gene, as determined by the 5′-RACE experiment, is indicated. In the case of the start site of *TTHA0963*, G ($+1$) was converted to C in all seven sequenced clones. As for *TTHA0973*, two of the seven clones were identified at the A (-2) site, and one was converted to C. The positions of the primers used for RT-PCR (see below) are indicated by open triangles and filled triangles for the *TTHA0963* to *TTHA0967* and *TTHA0973* to *TTHA0968* operons, respectively. (B) RT-PCR analysis to confirm the operon composed of *TTHA0963* to *TTHA0967* (lane 2) and that composed of *TTHA0973* to *TTHA0968* (lane 4). The positions of the primers are shown in panel A. PCR was also performed with no RT, as control using the same primers (lanes 3 and 5 are the controls for lanes 2 and 4, respectively). The samples were fractionated on a 1% agarose gel, followed by staining with ethidium bromide and photography. The expected bands are indicated with asterisks. Lane 1, 500-bp DNA ladder markers.

Structure modeling. Homology modeling of the three-dimensional structure of *T. thermophilus* PaaR was performed by the Modeller 9v8 program (5), using the X-ray crystal structure of the TetR family transcriptional regulator from *Thermobifida fusca* YX (Protein Data Bank code 3DCF), which shares 35% amino acid sequence identity (E value = 6e−23) (see Fig. 2), as the template structure. The electrostatic potentials were calculated using the adaptive Poisson-Boltzmann solver (APBS) (2) with the PyMol APBS tools.

Other methods. The N-terminal sequence analyses of the proteins were performed with a protein sequencer (Procise HT; Applied Biosystems). To estimate the molecular mass of *T. thermophilus* PaaR, gel filtration chromatography was performed with a Superdex 75 HR 10/30 column (GE Healthcare). BLAST and conserved domain database (CDD) searches were performed on the http://blast .ncbi.nlm.nih.gov/Blast.cgi and http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd .shtml websites, respectively. A nucleotide sequence motif search was performed with the GENETYX program, ver. 8.0 (Genetyx).

RESULTS

Organization of the putative *paa* gene cluster in *T. thermophilus* HB8. In the chromosomal DNA of *T. thermophilus* HB8 (NCBI accession number NC_006461), a putative *paa* gene cluster composed of two operons in opposite directions, which are possibly involved in PAA degradation, exists at the region spanning 909204 to 917385 (Fig. 1A). By RT-PCR analysis, we confirmed the operon structures, which are composed of the *TTHA0963* to *TTHA0967* genes and the *TTHA0973* to *TTHA0968* genes (Fig. 1B). The ABC transporter ATP-binding protein-related protein genes (*TTHA0963* and *-0964*) and

the hypothetical protein gene (*TTHA0967*) were found in the *paa* operons. The *TTHA0965* gene product, PaaD, which was previously named PaaI, displays thioesterase activity with several CoA compounds, including PA-CoA (14). According to the BLAST search, the *TTHA0973* gene encodes a TetR family transcriptional regulator, which shares homology to the *A. evansii* PaaR (20) (sequence identity = 25%, E value = 9e−11) (Fig. 2), and further experiments revealed that the TTHA0973 protein regulates the *paa* operons (see below); therefore, we named this gene *paaR*, as proposed by Mohamed Mel et al. (20). A homolog of the *paaX* gene, encoding a GntR family transcriptional repressor of the *paa* genes, which exists in the *paa* gene clusters of *E. coli* and *Pseudomonas* strains, was not found in close proximity to the *T. thermophilus paa* gene cluster. The *TTHA1876* gene product, which is annotated as a probable repressor for the PAA catabolic pathway and shares homology to the PaaX proteins from *E. coli* (E value = 1e−16) and *Pseudomonas putida* (E value = 8e−14), was found at a distant locus, but its function has not been identified.

Initial characterization of *T. thermophilus* PaaR. PaaR consists of 203 amino acid residues (NCBI accession number YP_144239), with a predicted molecular mass of 22.9 kDa. Based on the results of a CDD (18) search, the protein has a conserved domain of the TetR family transcriptional regulator, comprising residues I7 to V53, with an E value of 1.63e−09 for
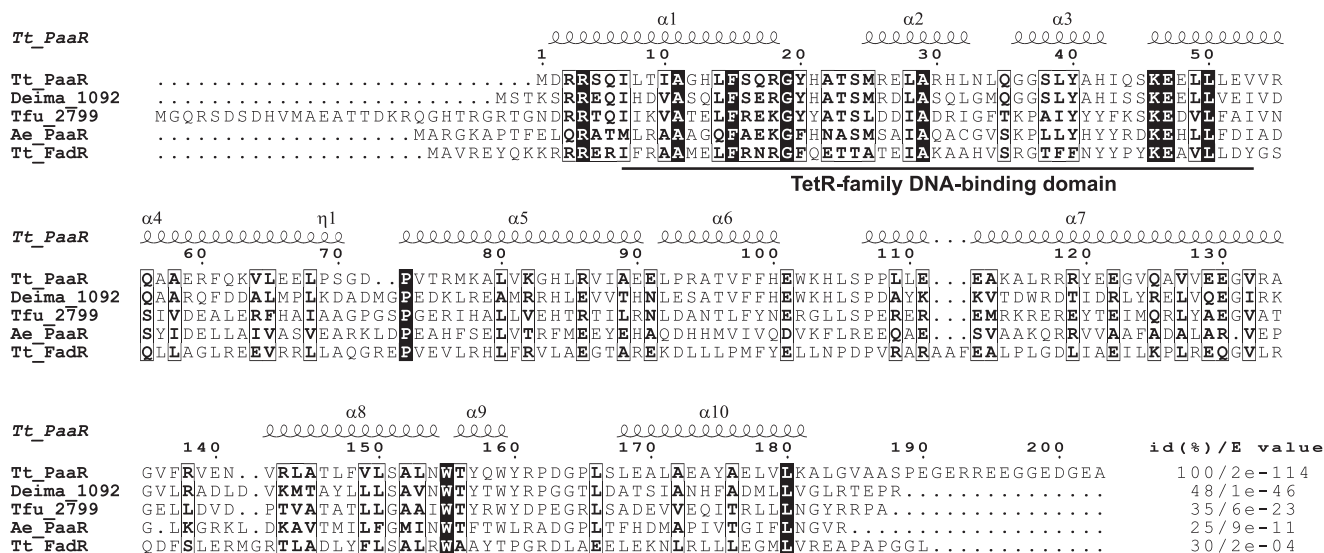
FIG. 2. Sequence alignment of *T. thermophilus* PaaR (Tt_PaaR) with representative homologous proteins. Strictly conserved residues are represented by white letters on a black background, and similar residues are depicted by boxed bold letters. The homologs are from *D. maricopensis* DSM 21211 (Deima_1092, YP_004170410), *Thermobifida fusca* YX (Tfu_2799, YP_290855), *A. evansii* (Ae_PaaR, CAC10611), and *T. thermophilus* FadR (Tt_FadR, YP_143367). The sequences were aligned using Clustal W2 (16). The secondary structure of Tt_PaaR was predicted with DSSP (13), and the figure was generated with ESPript 2.2 (9). α and η represent α-helix and 3₁₀-helix, respectively. The percent identities [id(%)] and E values relative to Tt_PaaR, determined by BLAST, are indicated on the right.

the consensus sequence (pfam00440). According to a BLAST search, the homologous protein most closely related to PaaR exists in *T. thermophilus* HB27 (E value = 6e−110), with the amino acid substitution E96K. The other homologs with lower E values were from *Thermus scotoductus* SA-01 (8e−88), *Thermus aquaticus* Y51MC23 (5e−82), *Meiothermus ruber* DSM 1279 (6e−55), *Truepera radiovictrix* DSM 17093 (4e−52), *Sphaerobacter thermophilus* DSM 20745 (1e−50), and *Deinococcus maricopensis* DSM 21211 (1e−46) (Fig. 2), and most of these homologs were annotated as TetR family transcriptional regulators.

T. thermophilus PaaR was overexpressed in *E. coli*, and the recombinant protein was purified from the cell lysate. The lysate was subjected to a heat treatment at 70°C for 10 min, and after centrifugation, the supernatant was fractionated on a hydrophobic column. This was followed by anion exchange, heparin, and gel filtration column chromatography steps. The protein was purified to >95%, based on an analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (see Fig. S1A in the supplemental material). We confirmed the N-terminal amino acid sequence of the purified protein as M-D-R-R-S. The molecular mass of PaaR, as estimated by gel filtration chromatography, was ~43 kDa, suggesting that *T. thermophilus* PaaR exists as a homodimer in solution (see Fig. S1B in the supplemental material).

Identification of *T. thermophilus* PaaR-binding sequences. Genomic selex was performed to screen DNA fragments that bind PaaR. The *T. thermophilus* HB8 genomic library used for this assay comprised 5×10^4 independent clones, carrying ~0.1- to 0.3-kbp DNA fragments, and thus, the redundancy of the clones in this library relative to the total genomic DNA is ~4.5. The DNA fragments that bound PaaR_His were selected from this library by using a nickel resin column. We cloned the

DNA fragments selected after three rounds of genomic selex, because the DNAs isolated after more than four rounds showed smear patterns on the agarose gel after electrophoresis and were difficult to clone. Among the selected clones, the nucleotide sequences of 63 clones were analyzed (see Table S2 in the supplemental material). The clones were categorized into three groups: those from ORFs (51 clones), that from an intergenic region (1 clone), and those containing both ORFs and intergenic regions (11 clones). Seven clones out of the 11 in the third category contained an upstream portion of the *TTHA0963* gene, which included a pseudopalindromic sequence, 5′-CGAACGGCCGTTAG-3′, a potential transcription factor-binding site (Fig. 1A). A similar pseudopalindromic sequence, 5′-CAAACGACCGTTAG-3′, was found upstream of the *TTHA0973* gene, which was also isolated by genomic selex (Fig. 1A).

DNA-binding ability of *T. thermophilus* PaaR. We confirmed the ability of PaaR to bind double-stranded DNAs (dsDNAs) with the aforementioned pseudopalindromic sequences, by means of a BIAcore analysis. The dsDNA fragments, containing the upstream regions of the *TTHA0963* (*TTHA0963p*; 5′-GAGGCGAACCGAACGGCCGTTAGT TTGGTCCA-3′) and *TTHA0973* (*TTHA0973p*; 5′-CGA AACTAACAAACGACCGTTAGCCATGGACC-3′) genes, were each immobilized on the streptavidin surface of a sensor chip, through biotin conjugated at the 5′ end of one of the strands, and then PaaR was injected over the DNA surface at 25°C. We found that PaaR bound both DNAs (Fig. 3A and B). PaaR did not bind the upstream region of the *TTHA0706* gene (5′-AG AAAAACACTTGACCAGTTGCTCAAATGATGCTACCCT GG-3′), which was not selected by the genomic selex experiment (data not shown), suggesting that PaaR does not nonspecifically bind DNA under these experimental condi-
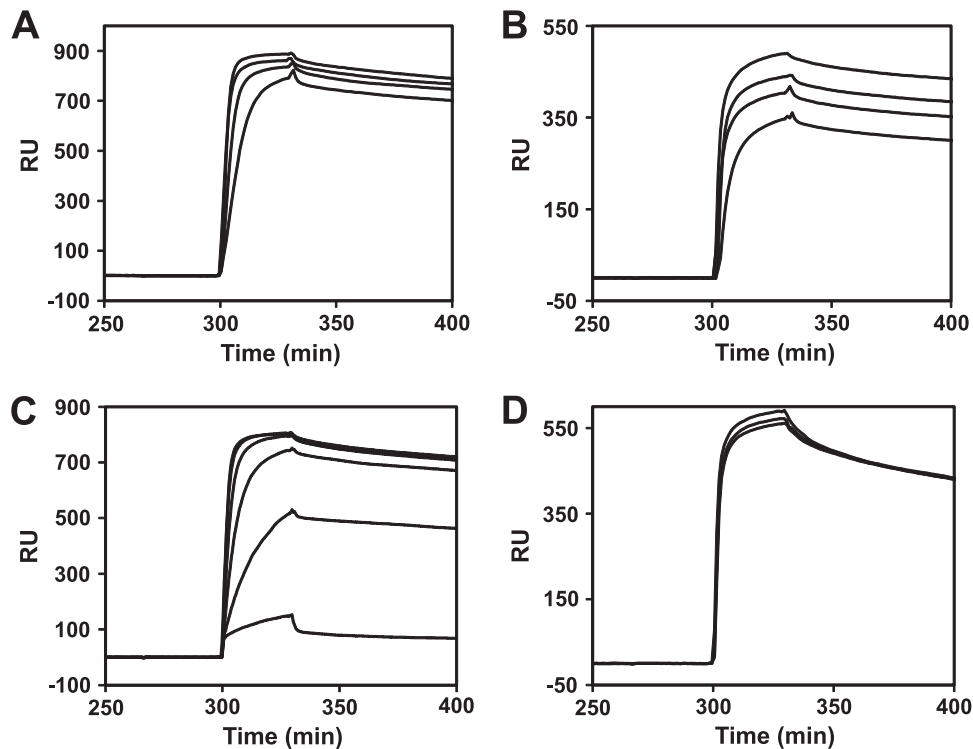
FIG. 3. BIAcore biosensor analysis of the *T. thermophilus* PaaR-DNA interaction. (A) A dsDNA corresponding to the upstream region of the *TTHA0963* gene (*TTHA0963p*) (see text) was immobilized on the sensor chip, and then PaaR was injected over the DNA surface at concentrations of 0.4, 0.3, 0.2, and 0.1 μM dimer, in buffer B containing 0.3 M NaCl. (B) A dsDNA corresponding to the upstream region of the *TTHA0973* gene (*TTHA0973p*) (see text) was used under the same experimental conditions as those of panel A. (C) The PaaR dimer (0.4 μM), along with 0, 0.2, 0.4, 0.6, 1, or 2 μM PA-CoA, was injected over the *TTHA0963p*-immobilized surface, in buffer B containing 0.3 M NaCl. (D) The PaaR dimer (0.4 μM) plus nothing (middle line), 2 μM Bz-CoA (bottom line), or 2 μM Ac-CoA (top line) was injected over the *TTHA0963p*-immobilized surface, as in panel C. Representative sensorgrams, minus the bulk refractive index background, were recorded and normalized to the baseline of 0 RU.

tions. PaaR bound the two DNAs in a similar manner, with association rate constant, dissociation rate constant, and dissociation constant values of $(9.3 \pm 1.0) \times 10^5$ (M^{-1} s^{-1}), $(1.0 \pm 0.1) \times 10^{-3}$ (s^{-1}), and 1.1 ± 0.1 (nM), respectively, for *TTHA0963p* and $(9.8 \pm 2.8) \times 10^5$ (M^{-1} s^{-1}), $(0.9 \pm 0.1) \times 10^{-3}$ (s^{-1}), and 0.9 ± 0.2 (nM), respectively, for *TTHA0973p*. These results suggest that the consensus sequence of the PaaR-binding site is 5′-CNAACGNNCGTTNG-3′. To find other possible *T. thermophilus* PaaR-binding sites, we searched for potential binding sites in the whole genome of *T. thermophilus* HB8, by using the consensus sequence as a query. However, no other sequences besides the two described above were found. Several short sequences similar to the PaaR-binding site were found in the other clones selected by the genomic selex. Among them, 41 clones derived from the *TTHA0647* ORF contained 5′-CGAACGTTAG-3′, five clones contained only one half of the consensus sequence (5′-CNAACG-3′ or 5′-CGTTNG-3′), five clones contained 5′-CNAAC-3′ or 5′-GTTNG-3′, two clones contained 5′-AACG-3′, and two clones contained 5′-AAC-3′ or 5′-GTT-3′ (see Table S2 in the supplemental material). PaaR may also bind these sites under the experimental conditions used for the genomic selex.

Next, we investigated the effects of several CoA compounds on the DNA-binding ability of *T. thermophilus* PaaR. The DNA-binding ability of PaaR gradually decreased with increased concentrations of PA-CoA (Fig. 3C), suggesting that it

is a ligand of *T. thermophilus* PaaR and that ligand binding releases PaaR from DNA. Benzoyl (Bz)-CoA and acetyl (Ac)-CoA did not significantly alter the DNA-binding ability of PaaR (Fig. 3D).

Ligand-binding ability of *T. thermophilus* PaaR. ITC experiments were performed to investigate the ligand-binding ability of PaaR. A typical thermogram was obtained when PA-CoA was titrated into PaaR (Fig. 4A). The K_b value (4.1×10^7 [M^{-1}]) and the number of binding sites on the protein (~1.3 sites per monomer PaaR) were obtained from the nonlinear one-site model to the normalized fitting curve. Under the same experimental conditions, a qualitatively different thermogram was obtained when Ac-CoA was titrated into PaaR (Fig. 4C). A reasonable fit was achieved with the nonlinear one-site model, with a K_b value of 2.8×10^5 (M^{-1}) (data not shown). The clear thermogram which indicates a molecular interaction was not observed when Bz-CoA was titrated (Fig. 4B). These results, together with those of the BIAcore assays (Fig. 3C and D), indicate that PA-CoA tightly binds PaaR and its binding releases PaaR from DNA.

Effects of *T. thermophilus* PaaR on transcription. The DNA fragments containing the probable PaaR-binding sites identified above were cloned and used as templates for *in vitro* transcription assays (Fig. 5A). We found that the DNA fragments were transcribed by *T. thermophilus* RNAP. The transcription reaction was repressed in the presence of *T. thermo-*
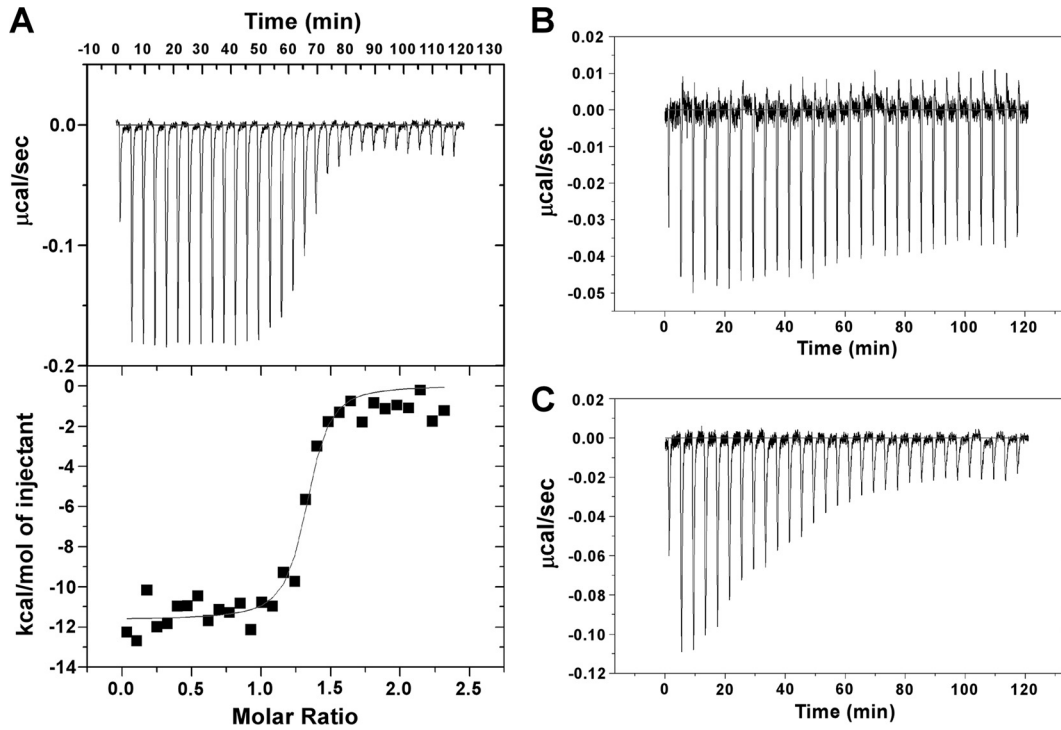
FIG. 4. ITC profiles of the titrations of *T. thermophilus* PaaR with PA-CoA (A), Bz-CoA (B), and Ac-CoA (C) at 30°C. The raw thermogram of each experiment is shown. The lower panel in panel A indicates the titration curve fitted to the one-site model.

philus PaaR in both cases. PaaR did not affect the transcription of the *TTHA0890* gene (Fig. 5A), which is regulated by another TetR family repressor, FadR (1), indicating that PaaR specifically represses the transcription of the DNAs containing the PaaR-binding sites. Note that the transcription efficiency differs depending on the template, i.e., the *TTHA0973* gene was

transcribed with much higher efficiency than the *TTHA0963* gene (see legend to Fig. 5A). The *TTHA0973* gene has a typical promoter sequence, with a 17-bp space between the −10 and −35 hexamer sequences, but the *TTHA0963* gene does not (Fig. 1A) (19). PaaR lost the ability to function as a transcriptional repressor in the presence of PA-CoA (Fig. 5A). Bz-CoA
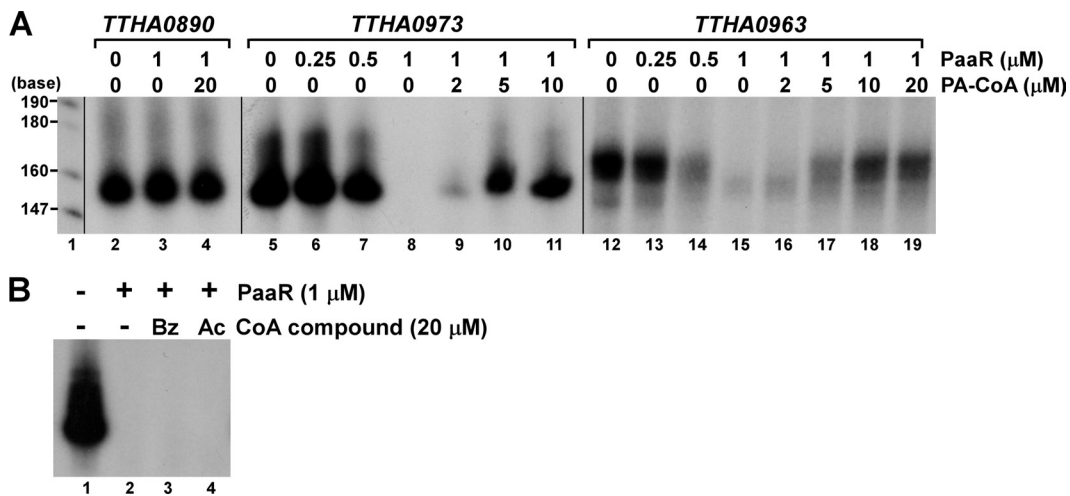

FIG. 5. Effects of *T. thermophilus* PaaR on transcription *in vitro*. (A) Runoff transcription assays were performed with templates containing the upstream sequences of *TTHA0890*, *TTHA0963*, and *TTHA0973* genes, in the absence or presence of 0.25, 0.5, or 1 μM PaaR as a dimer and in the absence or presence of 2, 5, 10, or 20 μM PA-CoA. In lanes 2 to 11, equivalent volumes were fractionated on the polyacrylamide gel, followed by autoradiography. In lanes 12 to 19, 4-fold-higher sample volumes were applied on the gel, which was exposed six times longer than in lanes 2 to 11. Lane 1, [α-^{32}P]dCTP-labeled MspI fragments of pBR322. (B) Runoff transcription assays were performed with a template containing the upstream sequence of the *TTHA0973* gene, in the absence or presence of 1 μM dimer PaaR, and in the absence or presence of 20 μM Bz-CoA or Ac-CoA. After the reaction, equivalent volumes of samples were fractionated on the polyacrylamide gel, followed by autoradiography.

or Ac-CoA had no effects on the function of PaaR, even at 20-fold-higher concentrations relative to the PaaR dimer (Fig. 5B). These results are consistent with those of the BIAcore analysis (Fig. 3) and the ITC experiments (Fig. 4). The effective concentration of PaaR on transcriptional repression and that of PA-CoA on transcriptional derepression of PaaR differed between the two promoters, perhaps reflecting their different strengths.

The transcriptional start sites of the PaaR-regulated genes *in vivo* were determined by means of 5′ RACE. They were identified downstream of and within the probable PaaR-binding sites, in the cases of the *TTHA0963* and *TTHA0973* genes, respectively (Fig. 1A). Note that the transcriptional start site of *TTHA0963*, G (+1), was converted to C in all seven sequenced clones. As for that of *TTHA0973*, two of the seven clones were identified at the A (−2) site, and one was converted to C. The potential promoter sequences were found within or in close proximity to the probable PaaR-binding sites (Fig. 1A). These results agree well with those from the *in vitro* transcription assays and support the proposed function of PaaR as a transcriptional repressor of the *paa* genes.

Putative three-dimensional structure of *T. thermophilus* PaaR. Homology modeling was performed, using a previously determined three-dimensional structure as the template. We selected the X-ray crystal structure of a TetR family transcriptional regulator from *Thermobifida fusca* YX as the template, since it showed the highest amino acid sequence identity and E value of the amino acid sequence (Fig. 2) to that of *T. thermophilus* PaaR, among the proteins with solved three-dimensional structures. The putative *T. thermophilus* PaaR structure adopts the typical three-dimensional structure of the TetR family proteins (1, 12, 23, 24, 28), with 10 α-helices (see Fig. S2A in the supplemental material). The N-terminal domain, composed of a typical helix-turn-helix motif (α2-3) with a positively charged surface (see Fig. S2B in the supplemental material), may be a DNA-binding domain. A second positively charged surface, found at the center of the predicted PaaR model (see Fig. S2B in the supplemental material), is similar to the acyl-CoA-binding site of another TetR family transcriptional regulator, the *T. thermophilus* FadR protein (1), despite the low sequence homology (E value = 2e−04) (Fig. 2).

DISCUSSION

The *T. thermophilus* HB8 genome encodes a TetR family transcriptional regulator within a gene cluster composed of putative *paa* genes, which are possibly involved in PAA catabolism. We found that the product of the TetR family gene, which we named *T. thermophilus* PaaR, is a repressor that controls the expression of the two operons composing the *paa* gene cluster *in vitro*. PA-CoA functioned as a ligand of PaaR, with an estimated binding stoichiometry of 1:1 protein monomer, and the PaaR–PA-CoA complex was released from the DNA, to allow transcription. Although weak binding by Ac-CoA was observed in the ITC experiment, this molecule had no observable effect on the activity of PaaR. In *E. coli* and *Pseudomonas* strains, the GntR family protein PaaX, whose gene is located in the *paa* gene clusters, is a repressor of the *paa* genes, and its ligand is PA-CoA; thus, PaaR is a functional homolog of PaaX. Since the transcripts of the *paa* operons

were observed in *T. thermophilus* HB8 cultivated in rich medium, the operons may be slightly expressed even under noninducing conditions. The transcription efficiency, the effective concentration of PaaR on transcriptional repression, and that of PA-CoA on transcriptional derepression of PaaR differed between the two promoters *in vitro*; thus, they may not be equally regulated *in vivo*. PaaR probably binds a pseudopalindromic sequence covering the −10 hexamer of the promoter or transcriptional start site. The consensus sequence is different from those of other TetR family proteins (1, 10, 23), indicating that the DNA-binding mechanisms differ among the TetR family proteins. The genomic selex experiment isolated large amounts of DNA fragments derived from ORFs, including part of the probable PaaR-binding consensus sequence. Whether PaaR actually acts on these genes *in vivo* remains to be elucidated.

The homologs of the PAA transport system genes (*paaL* and *paaM*) and those of the β-oxidation-like system genes for PAA metabolism (*paaA*, *paaB*, *paaC*, and *paaE*) (17) have not been found in the *T. thermophilus* genome. In this strain, other genes, including *TTHA0963*, *TTHA0964*, and *TTHA0967* in the putative *paa* operons, might be involved in such systems. Since the genome size of *T. thermophilus* HB8 is relatively small, i.e., about half of those of *E. coli*, *Pseudomonas*, and *Azoarcus* strains, PAA might be metabolized in an alternative manner in this strain. *T. thermophilus* HB8 has a functionally unknown gene distant from the putative *paa* gene cluster, which encodes a homolog of PaaX. This protein might be a transcriptional regulator of the other unidentified genes involved in PAA metabolism.

An X-ray crystal structure analysis revealed that another *T. thermophilus* TetR family protein, FadR, which is involved in fatty acid degradation, bound lauroyl-CoA (1). The lauroyl-CoA binds to the FadR molecule with its acyl chain moiety in the center of the molecule, enclosed within a tunnel-like substrate-binding pocket surrounded by hydrophobic residues, and the CoA moiety interacting with basic residues on the protein surface. The putative three-dimensional structure of *T. thermophilus* PaaR revealed that it adopts the typical structure of the TetR family proteins. A positively charged surface, found at the center of the PaaR molecule, is similar to the acyl-CoA-binding site of FadR. The CoA moiety of PA-CoA may bind to the center of the PaaR molecule, in a manner similar to the interaction of the CoA moiety of acyl-CoA with FadR. Further biochemical and structural studies will be needed to elucidate the regulatory mechanism of PaaR and the PA-CoA catabolic system of this thermophilic bacterium.

Interestingly, *E. coli* FadR, which is a functional homolog of *T. thermophilus* FadR, is the same GntR family protein (4, 11) as *E. coli* PaaX. Thus, the structures of the molecules that bind the intermediate CoA compounds in the catabolic pathways and regulate the expression of the genes involved in the cognate catabolism might be similar within a bacterial species but may not always be conserved among bacteria, even though the functions of the molecules are conserved.

REFERENCES

1. **Agari, Y., K. Agari, K. Sakamoto, S. Kuramitsu, and A. Shinkai.** 2011. TetR family transcriptional repressor *Thermus thermophilus* FadR controls fatty acid degradation. Microbiology **157**:1589–1601.

2. **Baker, N. A., D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon.** 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. **98**:10037–10041.

3. **del Peso-Santos, T., et al.** 2006. Coregulation by phenylacetyl-coenzyme A-responsive PaaX integrates control of the upper and lower pathways for catabolism of styrene by *Pseudomonas* sp. strain Y2. J. Bacteriol. **188**:4812–4821.

4. **DiRusso, C. C., T. L. Heimert, and A. K. Metzger.** 1992. Characterization of FadR, a global transcriptional regulator of fatty acid metabolism in *Escherichia coli*. Interaction with the *fadB* promoter is prevented by long chain fatty acyl coenzyme A. J. Biol. Chem. **267**:8685–8691.

5. **Eswar, N., et al.** 2006. Comparative protein structure modeling using Modeller. Curr. Protoc. Bioinformatics **15**:5.5.1–5.6.30.

6. **Ferrandez, A., J. L. Garcia, and E. Diaz.** 2000. Transcriptional regulation of the divergent *paa* catabolic operons for phenylacetic acid degradation in *Escherichia coli*. J. Biol. Chem. **275**:12214–12222.

7. **Ferrandez, A., et al.** 1998. Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. J. Biol. Chem. **273**:25974–25986.

8. **Garcia, B., et al.** 2000. Phenylacetyl-coenzyme A is the true inducer of the phenylacetic acid catabolism pathway in *Pseudomonas putida* U. Appl. Environ. Microbiol. **66**:4575–4578.

9. **Gouet, P., X. Robert, and E. Courcelle.** 2003. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. **31**:3320–3323.

10. **Hamlin, J. N., R. A. Bloodworth, and S. T. Cardona.** 2009. Regulation of phenylacetic acid degradation genes of *Burkholderia cenocepacia* K56-2. BMC Microbiol. **9**:222.

11. **Haydon, D. J., and J. R. Guest.** 1991. A new family of bacterial regulatory proteins. FEMS Microbiol. Lett. **63**:291–295.

12. **Hinrichs, W., et al.** 1994. Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance. Science **264**:418–420.

13. **Kabsch, W., and C. Sander.** 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers **22**:2577–2637.

14. **Kunishima, N., et al.** 2005. A novel induced-fit reaction mechanism of asymmetric hot dog thioesterase PAAI. J. Mol. Biol. **352**:212–228.

15. **Kuramitsu, S., K. Hiromi, H. Hayashi, Y. Morino, and H. Kagamiyama.** 1990. Pre-steady-state kinetics of *Escherichia coli* aspartate aminotransferase catalyzed reactions and thermodynamic aspects of its substrate specificity. Biochemistry **29**:5469–5476.

16. **Larkin, M. A., et al.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.

17. **Luengo, J. M., J. L. Garcia, and E. R. Olivera.** 2001. The phenylacetyl-CoA catabolon: a complex catabolic unit with broad biotechnological applications. Mol. Microbiol. **39**:1434–1442.

18. **Marchler-Bauer, A., et al.** 2002. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res. **30**:281–283.

19. **Maseda, H., and T. Hoshino.** 1995. Screening and analysis of DNA fragments that show promoter activities in *Thermus thermophilus*. FEMS Microbiol. Lett. **128**:127–134.

20. **Mohamed Mel, S., W. Ismail, J. Heider, and G. Fuchs.** 2002. Aerobic metabolism of phenylacetic acids in *Azoarcus evansii*. Arch. Microbiol. **178**:180–192.

21. **Navarro-Llorens, J. M., et al.** 2005. Phenylacetate catabolism in *Rhodococcus* sp. strain RHA1: a central pathway for degradation of aromatic compounds. J. Bacteriol. **187**:4497–4504.

22. **Oshima, T., and K. Imahori.** 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a non-sporulating thermophilic bacterium from a Japanese thermal spa. Int. J. Syst. Bacteriol. **24**:102–112.

23. **Ramos, J. L., et al.** 2005. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. **69**:326–356.

24. **Schumacher, M. A., et al.** 2001. Structural mechanisms of QacR induction and multidrug recognition. Science **294**:2158–2163.

25. **Shimada, T., N. Fujita, M. Maeda, and A. Ishihama.** 2005. Systematic search for the Cra-binding promoters using genomic SELEX system. Genes Cells **10**:907–918.

26. **Shinkai, A., et al.** 2007. Transcription activation mediated by a cyclic AMP receptor protein from *Thermus thermophilus* HB8. J. Bacteriol. **189**:3891–3901.

27. **Teufel, R., et al.** 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. Proc. Natl. Acad. Sci. U. S. A. **107**:14390–14395.

28. **Willems, A. R., et al.** 2008. Crystal structures of the *Streptomyces coelicolor* TetR-like protein ActR alone and in complex with actinorhodin or the actinorhodin biosynthetic precursor (S)-DNPA. J. Mol. Biol. **376**:1377–1387.