Crystal Structure of TetR Family Repressor AlkX from *Dietzia* sp. Strain DQ12-45-1b Implicated in Biodegradation of *n*-Alkanes

Jie-Liang Liang,^{a,b} Yuan Gao,^c Zheng He,^c Yong Nie,^a Meng Wang,^a
Jing-Hong JiangYang,^a Xuejun C. Zhang,^c Wen-Sheng Shu,^b Xiao-Lei Wu^a

Department of Energy and Resources Engineering, College of Engineering, Peking University, Beijing, China^a;
College of Ecology and Evolution, Sun Yat-Sen University, Guangzhou, China^b; National Laboratory of
Macromolecules, National Center of Protein Science—Beijing, Institute of Biophysics, Chinese Academy of
Sciences, Beijing, China^c

ABSTRACT *n*-Alkanes are ubiquitous in nature and are widely used by microorganisms as carbon sources. Alkane hydroxylation by alkane monooxygenases is a critical step in the aerobic biodegradation of *n*-alkanes, which plays important roles in natural alkane attenuation and is used in industrial and environmental applications. The alkane oxidation operon, *alkW1-alkX*, in the alkane-degrading strain *Dietzia* sp. strain DQ12-45-1b is negatively autoregulated by the TetR family repressor AlkX via a product positive feedback mechanism. To predict the gene regulation mechanism, we determined the 3.1-Å crystal structure of an AlkX homodimer in a non-DNA-bound state. The structure showed traceable long electron density deep inside a hydrophobic cavity of each monomer along the long axis of the helix bundle, and further gas chromatography-mass spectrometry analysis of AlkX revealed that it contained the *Escherichia coli*-derived long-chain fatty acid molecules as a ligand. Moreover, an unusual structural feature of AlkX is an extra helix, $\alpha 6'$, forming a lid-like structure with $\alpha 6$ covering the inducer-binding pocket and occupying the space between the two symmetrical DNA-binding motifs in one dimer, indicating a distinct conformational transition mode in modulating DNA binding. Sequence alignment of AlkX homologs from *Dietzia* strains showed that the residues involved in DNA and inducer binding are highly conserved, suggesting that the regulation mechanisms of *n*-alkane hydroxylation are possibly a common characteristic of *Dietzia* strains.

IMPORTANCE With *n*-alkanes being ubiquitous in nature, many bacteria from terrestrial and aquatic environments have evolved *n*-alkane oxidation functions. Alkane hydroxylation by alkane monooxygenases is a critical step in the aerobic biodegradation of *n*-alkanes, which plays important roles in natural alkane attenuation and petroleum-contaminating environment bioremediation. The gene regulation of the most common alkane hydroxylase, AlkB, has been studied widely in Gram-negative bacteria but has been less explored in Gram-positive bacteria. Our previous study showed that the TetR family regulator (TFR) AlkX negatively autoregulated the alkane oxidation operon, *alkW1-alkX*, in the Gram-positive strain *Dietzia* sp. strain DQ12-45-1b. Although TFRs are one of the most common transcriptional regulator families in bacteria, the TFR involved in *n*-alkane metabolism has been reported only recently. In this study, we determined the crystal structure of AlkX, which implies a distinct DNA/ligand binding mode. Our results shed light upon the regulation mechanism of the common alkane degradation process in nature.

KEYWORDS AlkX, TetR family repressor, *n*-alkane biodegradation, crystal structure, *Dietzia*

Received 12 July 2017 Accepted 16 August 2017

Accepted manuscript posted online 18 August 2017

Citation Liang J-L, Gao Y, He Z, Nie Y, Wang M, JiangYang J-H, Zhang X-C, Shu W-S, Wu X-L. 2017. Crystal structure of TetR family repressor AlkX from *Dietzia* sp. strain DQ12-45-1b implicated in biodegradation of *n*-alkanes. *Appl Environ Microbiol* 83:e01447-17. <https://doi.org/10.1128/AEM.01447-17>.

Editor Maiva Kivisaar, University of Tartu

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Xiao-Lei Wu, xiaolei_wu@pku.edu.cn.

J.-L.L., Y.G., and Z.H. contributed equally to this article.

Numerous and diverse alkane-degrading microorganisms that utilize alkanes as carbon sources are widespread in nature, including in Gram-negative and Gram-positive bacteria and archaea (1, 2). Since microbial degradation processes are involved in natural alkane attenuation and can change the compositions of crude oil constituents in oil reservoirs, they have attracted wide attention for their industrial applications, such as biodegradation and microbially enhanced oil recovery (3). For aerobic alkane degradation, alkane hydroxylases are key and rate-limiting enzymes (4). Among the three major types of alkane hydroxylases that have been functionally identified, the best-characterized integral-membrane nonheme diiron monooxygenase (AlkB) is responsible for the degradation of medium- and long-chain *n*-alkanes (5, 6).

Dietzia sp. strain DQ12-45-1b is a Gram-positive bacterium capable of degrading a wide-range of *n*-alkanes (C_6 to C_{40}) and is used as the paradigm of biodegrading bacteria by our group (6–11). In this strain, an AlkB-type *n*-alkane hydroxylase, AlkW1, was found to degrade medium- and long-chain *n*-alkanes (C_{14} to C_{32}) (6). The expression of AlkW1 is regulated by a cotranscribed transcriptional repressor, AlkX (GenBank accession number [AEM66515.1](#)), a member of the TetR family regulators (TFRs) (11). AlkX forms a homodimer, and two AlkX dimers bind to an unusually long 48-bp inverted repeat (IR) in the *alkW1* promoter to repress the transcription of the *alkW1-alkX* operon. Additionally, *n*-alkane metabolites, which include fatty acids with alkyl chains ranging from C_{10} to C_{24} , modulate AlkX-DNA binding affinity, presumably by binding to the AlkX homodimer. Briefly, fatty acids regulate the alkane degradation pathway via an autoregulation positive feedback mechanism (11). As members of the third most common transcriptional regulator family in bacteria, TFRs play broad roles in prokaryotic physiology (12). However, the regulation of TFRs in *n*-alkane metabolism was only recently reported.

More than 200 crystal structures of TFRs are included in the Protein Data Bank (PDB). All TFRs form homodimers, and each monomer, consisting mostly of α -helices, has an N-terminal DNA-binding domain (DBD) and a C-terminal ligand-binding domain (LBD). Although the overall structures of TFRs are similar, they typically share few similarities in their amino acid sequences except in the helix-turn-helix (HTH) motifs of the DBDs. The LBDs typically possess variable hydrophobic cavities, which often contain highly diverse ligands. Known TFR ligands include antibiotics, bile acid and other toxic molecules, cell-cell signaling molecules, carbon sources, peptides, fatty acids, fatty acid derivatives, and metal ions (12). Binding of ligands to LBDs may modify the affinity of the TFRs to target DNA by adjusting the distance between the two HTH motifs in the DBDs in the TFR dimers.

Analysis of the available TFR-DNA complex structures revealed that TFR members can be divided into two major classes (13): (i) those that bind to their cognate DNA elements as a dimer and (ii) those that bind as a pair of dimers. For example, *Escherichia coli* TetR (14), *Pseudomonas aeruginosa* DesT (15), and *Staphylococcus antibioticus* SimR (16) bind their DNA as homodimers, whereas *Staphylococcus aureus* QacR (17), *Corynebacterium glutamicum* CgmR (18), *E. coli* SlmA (19), and *Mycobacterium smegmatis* Ms6564 (20) bind as two dimers. For the second class of TFRs, binding of the two dimers is typically cooperative and requires a defined distance between the two binding sites in the DNA double-strand helix (13, 18, 21). The TFRs generally bind to the cognate DNA sequences by widening the major groove to enable requisite protein-DNA contacts (13). In addition, they often introduce a bend in the cognate DNA to better stabilize the protein-DNA complex. However, the distortion is more significant for the second class of TFRs, such as QacR (21), CgmR (18), and CprB (13). AlkX binds DNA not only as a pair of dimers but also as a single dimer *in vitro*. Particularly, in contrast to the second class of TFRs, which bind a “half-site” IR by two subunits from separate dimers, two AlkX dimers contact the DNA side by side on opposite sides of the DNA (i.e., a “half-site” IR bound by one dimer) (for details, see Fig. S5 in the supplemental material) (11).

Although a number of structure-based studies on TFRs have provided important insights into ligand-induced conformational changes and protein-ligand/protein-DNA contacts, it is difficult to apply the structural information to AlkX because of the low

TABLE 1 Data collection and refinement statistics the SeMet-AlkX crystal

Parameter	Value(s) for SeMet-AlkX ^a
Data processing statistics	
Space group	C121
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	109.6, 179.7, 56.6
Wavelength (Å)	0.9762
Resolution (Å)	45.68–3.1 (3.2–3.1)
Completeness (%)	99.4 (99.2)
<i>R</i> _{merge} (%) ^b	9.3 (48.2)
CC _{1/2}	0.922
<i>I</i> /σ (<i>I</i>)	18.5 (4.1)
No. of unique reflections	18,403 (1,755)
Redundancy	6.4 (6.5)
Refinement statistics	
Resolution (Å)	34.2–3.1
No. of reflections (test)	18,390 (918)
<i>R</i> _{work} / <i>R</i> _{free} (%) ^c	24.8/27.6
No. of atoms	4,182
Avg <i>B</i> factor (Å ²)	27
Root mean square deviation	
Bond length (Å)	0.009
Bond angle (°)	1.08
Ramachandran plot (%) ^d	
Favored regions	98.3
Allowed regions	1.2

^aValues in parentheses are for the highest-resolution shell.

^b $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i - \langle I \rangle| / \langle I \rangle$, where *I_i* is the intensity for the *i*th measurement of an equivalent reflection with indices *h*, *k*, and *l*.

^c $R = \sum |F_o - F_c| / \sum |F_o|$, where *F_o* and *F_c* are the observed and calculated structure factors, respectively.

^dCalculated using MolProbity.

amino acid sequence similarity (<30%). In addition, these structures cannot exactly explain the AlkX binding capacities of medium- and long-chain fatty acids (LCFAs) as well as the 48-bp-long operator (11). Therefore, in this study, we report the crystal structure of AlkX in complex with the ligand derived from *E. coli*, which was determined as LCFAs by gas chromatography-mass spectrometry (GC-MS). Structure-based mutagenesis analyses and functional assays suggested the binding mechanisms of AlkX with its target DNA as well as fatty acid inducers. Amino acid sequence analysis revealed that AlkX homologs represent a novel subfamily of TetR proteins involved in *n*-alkane degradation, and this regulation mode may be common in *Dietzia* strains.

RESULTS

Overall structure of AlkX. Recombinant AlkX protein from *Dietzia* sp. strain DQ12-45-1b was overexpressed in *E. coli* strain DH5α and then purified and crystallized. We used the single-wavelength anomalous diffraction method to determine the selenomethionine (SeMet)-derivative crystal structure of AlkX to a resolution of 3.1 Å and refined the *R*_{work} and *R*_{free} values to 24.8% and 27.6%, respectively (Table 1). One crystallographic asymmetric unit contained three AlkX monomers (referred to as molecules A, B, and C), which shared similar overall structure. Molecules A and B formed a symmetrical homodimer, while molecule C formed a similar homodimer with one of its symmetry mates. The Cα-atom root mean square deviations between the monomers were 0.50, 0.34, and 0.53 Å for the A/B, B/C, and A/C pairs, respectively.

While TFRs are commonly comprised of nine α-helices (12), some TFRs have 10 α-helices, such as FadR (22) and CprB (13). The AlkX monomer is also comprised of 10 α-helices, i.e., α1 (residues 30 to 45), α2 (52 to 61), α3 (63 to 71), α4 (73 to 100), α5 (103 to 122), α6 (122 to 130), α6' (135 to 145), α7 (146 to 162), α8 (168 to 189), and α9 (195 to 213) (Fig. 1A). These 10 helices can be divided into two domains, the N-terminal DBD (α1 to α3, residues 1 to 72) and the C-terminal LBD (α4 to α9, residues 73 to 219). Loop regions are labeled with the numbers of their preceding helices. To highlight the structural differences in the α6-α7 region between AlkX and its TetR homologs, we

FIG 1 Crystal structure of AlkX. (A) Cartoon representation of the protomer of AlkX. Helices are colored in a rainbow scheme (from the N terminus in red to the C terminus in blue). For clarity, the $\alpha 3$ helix is indicated with a blue arrow. (B) AlkX homodimer generated by crystallographic symmetry. The distance between the two recognition helices (defined by C α atoms of Tyr68 from each protomer) was measured to be 46 Å. The missing region is shown as dotted lines. Helix $\alpha 6'$ is highlighted with green color in both panels.

designated residues 135 to 145 helix $\alpha 6'$ (Fig. 1B). As discussed below, we hypothesized that this plays an important role in structural change in response to inducer binding. Because of their flexibility, neither the N terminus residues (1 to 25) nor the C terminus residues (215 to 219), as well as the N-terminal fused His tag, were visible in the electron density map, nor was part of the region between $\alpha 6$ and $\alpha 6'$ (sequence details are given in Fig. 2).

The three N-terminal helices, $\alpha 1$ to $\alpha 3$ (i.e., the DBD), presumably interact with target DNA via the canonical HTH motif. In this motif, the recognition helices $\alpha 2$ and $\alpha 3$ are

FIG 2 Sequence alignment of AlkX with TetR family transcriptional repressor DesT from *P. aeruginosa*, which possesses a similar ligand-binding pocket. Secondary-structure elements for AlkX are shown above the alignment, where α indicates the α -helix. Identical residues are highlighted with pink shading. Green shading denotes the phenylalanine sensor conserved in both AlkX and DesT. Three prolines located in unusual positions in AlkX are highlighted in orange shading. Invisible residues that were not modeled in the structures of AlkX and DesT are indicated with gray shading. Residues involved in the formation of the hydrophobic cavity are indicated with blue boxes. Red frames mark key residues conserved in TetR family proteins. The red bar indicates DNA recognition sites conserved in *Dietzia* strains and *Actinobacteria*.

FIG 3 Putative ligand-binding pocket of AlkX. (A) The tunnel-like pocket contains ligand density. The pocket was formed by hydrophobic residues shown as stick models. The traceable electron density from the $F_o - F_c$ map (contoured at 2.5σ) is shown in green, while the density from $2F_o - F_c$ (contoured at 1.0σ) is shown in purple. All density maps were calculated without modeling inside the pocket. (B to D) Top view of the central region (B), bottom view of the bottom entrance (C), and top view of the top (D) of the hydrophobic pocket. In addition, the cavity surface is shown, and residues referred to in the text are labeled.

thought to bind with the promoter at the positively charged surface patch formed by the basic residues Arg46, Lys58, Arg64, and Arg74 (see Fig. S1 in the supplemental material). The remaining seven helices constitute the LBD, with helix $\alpha 4$ serving as a linker connecting to the DBD. There is a large tunnel-like cavity in the LBD of each AlkX protomer, which is thought to be an inducer-binding pocket (Fig. 3A). This cavity is surrounded by helices $\alpha 4$, $\alpha 5$, $\alpha 7$, and $\alpha 8$ and is elongated in the direction of the long axis of the protomer. In addition, helices $\alpha 6$ and $\alpha 6'$ are located on top of helices $\alpha 4$, $\alpha 5$, $\alpha 7$, and $\alpha 8$ proximal to the DBD, blocking the binding pocket, which might contribute to pocket assembly.

The interface of the symmetrical dimer is composed primarily of helices $\alpha 8$ and $\alpha 9$ from both subunits. Helices $\alpha 8$ and $\alpha 9$ form a pairwise perpendicular helix bundle with the same helices from the other monomer. The dimer interface is $2,800 \text{ \AA}^2$ in size, which includes both monomers. There is an extensive network of hydrogen bonding and electrostatic and hydrophobic interactions between helices $\alpha 8$ and $\alpha 9$ from both of the monomers, which might help stabilize the dimeric interface.

Inferred inducer-binding sites of AlkX. The putative inducer-binding pocket of AlkX structure was lined predominantly, albeit not exclusively, by hydrophobic residues, and this tunnel-like cavity measured approximately $32 \text{ \AA}$ long by 5 to $8 \text{ \AA}$ in diameter. Furthermore, we identified traceable long electron density in the cavity located deep in the core of each monomer (Fig. 3A). The electron density is parallel to the cavity. Since none of the buffer components from the crystallization setup fit the observed shape, the electron density likely corresponded to an organic molecule derived from the expression host *E. coli*. To further determine the true ligand in AlkX,

FIG 4 Structural homologs with distinct ligand-binding pockets. (A) Comparison between AlkX (cyan), FadR (green) (PDB ID 3ANG), and DesT (yellow) (3LSJ) with respect to the four-helix bundle ligand-binding pocket. The orientation of the phenylalanine-rich cluster and Leu148 in AlkX with its equivalents in DesT showed similar conformations. Helix $\alpha 7$ was moved for clarity. The ligand is shown as a stick representation. (B) Superposition of the subunits from structural homologs. Helix $\alpha 6$ and $\alpha 6'$ in AlkX (cyan) assemble a lid that is drastically different from other structures and serves as a platform for the transition of conformational changes in other TFRs, such as DesT (yellow), FadR (green), KstR (pink) (3MNL), and a putative TetR family transcriptional regulator (purple) (3BJB). The ligand is shown as a stick representation.

we performed GC-MS analysis and detected the presence of palmitic acid (see Fig. S3 in the supplemental material), which indicated AlkX crystalized in complex with *E. coli*-derived palmitic acid. The results were consistent with our previous electrophoretic mobility shift assay (EMSA) results that fatty acids ranging from C_{10} to C_{24} interfered with the AlkX-DNA binding affinity (11). The detection of palmitic acid is probably due to its relatively high protein binding affinity in comparison to those with shorter or longer alkyl chains (11). Actually, the length of a medium- and long-chain saturated fatty acid (C_{10} to C_{24}) was approximately 12 to 28 Å and fit into the hydrophobic cavity. However, the electron density could fit well with that of dodecanoic acid, so we analyzed the AlkX structure in complex with it.

The hydrophobic residues in the inducer-binding pocket are engaged in hydrophobic interactions with the buried electron density of the inducer; these residues included Ala82, Leu85, Phe89, etc. (indicated with blue boxes in Fig. 2). Actually, the alkyl chain appeared to have a strong influence on the packing of these hydrophobic core residues within the cavity, as the benzyl side chain of Phe115 and the isobutyl side chain of Leu148 seemed to deviate to accommodate the putative ligand (Fig. 3B and 4A). Furthermore, side chains of these residues around the cavity were positioned toward the top end ($\alpha 6$ and $\alpha 6'$ region), which may have also resulted from the ligand's introduction. Thus, the pocket may open to the bottom of the protomer, with Ile104 (from $\alpha 5$), Trp164, Met165, Ile167 (from loop 7), Leu209, and Val213 (from $\alpha 9$) located in the pocket entrance restricting the opening of the entrance tunnel (Fig. 3C). Additionally, $\alpha 6$ and $\alpha 6'$ assemble to form a lid-like structure, blocking the pocket from the top by using Leu124, Leu128, Leu138, and Ile141, with the aid of Trp48 from loop 1 (Fig. 3D). Among these, the side chain of Leu138 intimately contacted the density to prevent further invasion of the putative ligand.

Similar to AlkX, other TFRs, such as DesT and FadR, also possess a tunnel-like cavity lined by hydrophobic residues (Fig. 4A) (15, 23), which can accommodate strong hydrophobic molecules with long acyl chains. Previous structural analysis of DesT revealed that the shape of the acyl chain directly influenced the packing of hydrophobic core residues within the LBD and further triggered the rearrangement of the two HTH motifs (15). In detail, the acyl chain penetrates the Phe-rich cluster below $\alpha 6$ and induces several conformational changes to accommodate the ligand instead of DNA. Sequence alignment with DesT (Fig. 2) indicated a similar phenylalanine-rich cluster in the AlkX hydrophobic pocket, comprised of Phe89, Phe115, and Phe186 from helices

$\alpha 4$, $\alpha 5$, and $\alpha 8$, respectively. We compared the orientation of the Phe-rich cluster and Leu148 in AlkX with their equivalents in DesT, as shown in Fig. 4A, and found that they share a homologous conformation. We thus speculate that these residues adjust their side chain rotamers to accommodate the alkyl chains. Moreover, in contrast to the case for DesT, there are three proline residues located at the loops connecting 5, 6, 6', and 7 only in AlkX, which may have a different role in protein structure conformation changes (Fig. 2).

Although they possess a similar tunnel-like cavity, DesT and FadR are different from AlkX by using palmitoyl coenzyme A (palmitoyl-CoA) and lauroyl-CoA as ligands, the CoA moiety of which is too large to travel past the entrance of the ligand-binding pocket (15, 23). The pocket entrance of AlkX was approximately 4 Å by 8 Å, limited by the side chains of residues Ile104, Trp164, Met165, Ile 167, Leu209, and Val213 (Fig. 3A). Since the carboxyl group is approximately 3 Å in size, which is smaller than the size of the entrance, the carboxyl group can probably enter into the pocket. However, because of their strong hydrophobic effects, residues Ile93, Leu156, and Ile197, as well as the Phe-rich cluster located in the central region of the pocket, may prevent the carboxyl group from further penetration. Taking the data together, the hydrophobic effect of the long alkyl chain may further influence the packing of hydrophobic core residues, which must adjust their positions to accommodate the inducer, thus resulting in conformational changes not compatible for DNA.

Distinct DNA binding in AlkX. (i) Structural properties of the DBD. The N-terminal HTH motif involved in DNA binding is composed of helix $\alpha 2$, loop 2, and $\alpha 3$. As a general feature of TFRs, helix $\alpha 2$ acts as the “platform” helix spanning the cognate major groove of the target DNA, while $\alpha 3$ acts as the “recognition” helix penetrating the major groove. Long $\alpha 1$ is preceded by positively charged residues and is typically responsible for contacts with DNA phosphates (15, 18, 21). In AlkX, Arg32 is located at the N terminus of $\alpha 1$ and may play the same role (Fig. 2). In addition, the N terminus of $\alpha 4$, Lys48 in TetR (PDB ID [1QPI](#)) or Arg in other TFRs, linking the LBD to the DBD, participates in the binding of DNA by attaching to the phosphate backbone (14, 18, 21, 24). We also found the conserved residue Arg74 in $\alpha 4$ in AlkX, which may be involved in DNA binding. Superposition of the DBD and helix $\alpha 4$ of AlkX with other TFRs showed that they share high conformational similarities in DNA recognition helices, apart from $\alpha 4$ of AlkX, which showed a slight deviation (see Fig. S2 in the supplemental material).

The recognition helix $\alpha 3$ of TFRs always contains a conserved Tyr residue (Tyr42 in TetR), which forms important interactions with DNA. Separation of the two recognition helices in a homodimer is represented by the distance between the C α atoms of the two Tyr residues. In known structures of DNA-TFR complexes, the distance varies by 36.6, 36.7, 36.9, and 38.2 Å in TetR (PDB ID [1QPI](#)) (14), DesT ([3LSP](#)) (15), QacR ([1JTO](#)) (21), and CprB ([4PXI](#)) (13) (36.3 Å designated by Ile42 in CgmR [[2YVH](#)] [18]), respectively. Generally, in inducer-bound or apo-TFR structures, the separation is much wider and is commonly not compatible with DNA binding. The equivalent residue in AlkX is Tyr68, and the C α distance is 46 Å (Fig. 1B), which is much wider than that of the canonical DNA-binding mode, indicating that AlkX in our current structure assumed a conformation that was not compatible with DNA binding. Furthermore, it also suggests that the binding of AlkX to the operator DNA might induce a conformational change in AlkX to shrink the distance between its $\alpha 3$ and $\alpha 3'$ helices (see Fig. S5 in the supplemental material).

(ii) Determination of residues related to DNA binding. As helix $\alpha 3$ acts as the “recognition” helix which interacts mainly with DNA major groove, we replaced the residues in $\alpha 3$ (Ser63, Arg64, Gln65, Thr66, Leu67, and Tyr68) with glycine individually to investigate the possible roles of these residues in the DNA-binding process. The AlkX variants were measured for DNA-binding affinity in an EMSA, using wild-type (WT) AlkX as a positive control (Fig. 5). The R64G mutation resulted in the complete loss of DNA-binding affinity, indicating that the interaction with DNA bases was disrupted. Mutating Thr66 and Leu67 to glycine also caused the complete loss of binding affinity,

FIG 5 Point mutation analysis of the DNA-binding domain of AlkX. The EMSA results are shown for the DNA-binding activity of wild-type (WT) AlkX and its various single point mutants. The two AlkX-DNA shift bands are indicated with black arrows. The upper shifted band is the result of two dimers interacting with the operator, whereas the lower shifted band is one dimer interacting with the operator. Free DNA is indicated with a red arrow.

whereas mutating Ser63 and Tyr68 to glycine impaired the binding affinity. These amino acid residues may interact with the phosphate groups of nucleotides via hydrogen bonds. However, the mutation of Gln65 to glycine resulted in a binding affinity similar to that of the WT AlkX, indicating that this residue is not directly involved in DNA binding.

(iii) Mode of binding between repressor and DNA. Our previous study showed that two AlkX dimers bound to the unusually long 48-bp IR through two pairs of palindromic sequences, designated OP1/2 and OP3/4 (11) (sequence details, are given in Fig. 6). In order to identify the specific nucleotides that interact with AlkX, we carried out three-point mutations on the IR and estimated the effects by EMSA (Fig. 6, lanes

FIG 6 Identification of specific nucleotides involved in AlkX-DNA binding. Mutated DNAs (OP1-1 to C+6) designed for EMSA are shown with the mutated sequences in red. There were three kinds of mutations, including three-point mutations in OP1 and OP3 (i.e., OP1-1, OP1-2, OP1-3, OP3-1, and OP3-2), combined three-point mutations in OP1 and OP2 and in OP1 and OP4 (i.e., OP1-3-OP2-3 and OP3-1-OP4-1), one base pair added in the nonpalindromic sequence between OP1 and OP3 (i.e., 70+1), and five or six base pairs added at the center of the operator (i.e., C+5 and C+6). A red asterisk indicates the location of the 2-fold axis. The nonpalindromic sequence between OP1 and OP3 is highlighted in blue. The EMSA results for the DNA-binding activity of AlkX to various DNA fragments are shown. Two AlkX-DNA shift bands are indicated with black arrows. The upper shifted band is the result of two dimers interacting with the operator, whereas the lower shifted band is one dimer interacting with the operator. Free DNA is indicated with a red arrow.

OP1-1 to OP3-2). Densitometry data for all shifted AlkX-DNA binding bands and free DNA bands were further documented to compare the relative binding affinities of AlkX to different mutated operators (see Fig. S4 in the supplemental material). Compared with that of the control, the AlkX binding affinity to the mutated operators was generally weaker. In detail, our results showed that the 5'-AAA-3' in OP1 (lane OP1-3) and 5'-TGT-3' in OP3 (lane OP3-1) were more important than the other nucleotides in OP1 and in OP3, respectively. This suggested that the nucleotides in the palindromic regions proximal to the nonpalindromic region were vital for specific recognition by DBDs. Moreover, simultaneous mutations of 5'-AAA-3' in OP1 and 5'-TTT-3' in OP2 greatly decreased AlkX-DNA binding activity, producing more free DNA (lane OP1-3-OP2-3). However, insertion of one base pair into the 11-bp nonpalindromic sequence between OP1 and OP3 could not significantly reduce the AlkX-DNA binding affinity (lane 70+1).

In most available structures of TFR-DNA complexes, such as QacR, CgmR, and CprB, two consecutive major grooves are contacted by two HTH motifs from separate dimers. However, this well-known binding fashion is quite different in AlkX, where one dimer binds to OP1 and OP3, while another binds to OP2 and OP4. Thus, the two dimers attach to the IR side by side on opposite sides of the DNA. In order to confirm this observation, we mutated the operator by inserting five or six additional base pairs at the center of the operator. The EMSA result showed that a 5-bp or 6-bp insertion sequence between OP3 and OP4 did not significantly alter AlkX-DNA binding affinity (lanes C+5 and C+6) compared to that of the control. The results were preliminary; however, our results showed that AlkX might be the first reported example where each dimer binds a "half-site" IR. It is different from the way that two subunits from separate dimers bind a "half-site" IR adopted as QacR or CprB.

Highly conserved sequences of AlkX homologs in *Dietzia* strains. The corresponding TFRs in *Dietzia* sp. strain UCD-THP, *Dietzia* sp. strain E1, *Dietzia alimentaria* 72, *Dietzia cinnamea* P4, and other 22 additional *Dietzia* strains originally isolated from geographically distant places around the world share high amino acid sequence similarities in the range of 82 to 99% with AlkX in *Dietzia* sp. strain DQ12-45-1b (11). In particular, the residues Arg32 in $\alpha 2$, Ser63, Arg64, Gln65, Thr66, and Tyr68 in $\alpha 3$ of AlkX are highly conserved in all *Dietzia* strains and even among *Actinobacteria* (Fig. 7), indicating a similar DNA-binding mode among these strains. In addition, the hydrophobic residues (indicated with blue boxes in Fig. 2) and the Phe-rich cluster in the AlkX inducer-binding pocket, as well as the three proline residues at the loops connecting $\alpha 5$, $\alpha 6$, $\alpha 6'$, and $\alpha 7$, are conserved in all *Dietzia* strains (Fig. 7), indicating that they share a ligand-binding mode that is specific among *Dietzia* strains.

DISCUSSION

Numerous crystal structures of TFRs in the unbound or DNA/inducer-binding form are now available in the PDB, increasing our understanding of the regulation mechanisms of TFRs in antibiotic production, multidrug resistance, amino acid metabolism, osmotic stress, pathogenicity, and development (12, 25). However, we report that AlkX is a TFR involved in *n*-alkane hydroxylation (11); the crystal structure of AlkX can be used to predict the gene transcription regulation of alkane degradation. Our structural results revealed that AlkX might represent a novel subfamily of TetR proteins.

A striking feature of the AlkX structure compared to other TFR structures. As reported previously, the distance and relative orientation of the HTH motifs distinguish between the repressed and derepressed status of TFRs, for the center-to-center separation of the recognition helices $\alpha 3$ and $\alpha 3'$ in the operator complex increases upon inducer binding (14–16, 18, 21, 26). Many TFRs have only one helix, $\alpha 6$, (Fig. 4B), which serves as a platform to transmit conformational changes between the repressed and derepressed status. For example, QacR and TetR undergo a transition from coil to helix upon ligand binding (14, 21, 27), thus switching the DNA-bound conformation to the ligand-bound conformation. In this case, the flexibility of the changeable coil region is vital for DNA binding, and the flexibility is hindered upon ligand binding (28). In

FIG 7 Amino acid sequence alignment of AlkX in *Dietzia* sp. strain DQ12-45-1b with the corresponding TFRs in other Gram-positive strains, including *Dietzia*, *Rhodococcus*, *Mycobacterium*, *Gordonia*, *Nocardioides*, and *Präuserella*. The degree of conservation associated with each position is indicated by the bar graph shown above the alignment. Conserved amino acid residues in all sequences are indicated on a dark gray background. The 5'-SRQT-Y-3' conserved region is indicated with a red line. The hydrophobic residues in the ligand-binding pocket (same as those indicated with blue boxes in Fig. 2) are indicated with blue arrows. The Phe-rich sites are indicated with green arrows. The Pro sites are indicated with orange arrows.

contrast to the overall structures of other TFRs, the most striking feature of the AlkX structure is the extra helix $\alpha 6'$, which blocks the binding pocket together with $\alpha 6$ (Fig. 1B and 4B). Helices $\alpha 6$ and $\alpha 6'$ occupy the space between two DBDs, making it difficult to adopt conformational changes similar to those of other TFRs (12, 25). In order to offset this allosteric inhibition, the two monomers might shrink the distance between the $\alpha 3$ and $\alpha 3'$ helices by rotation along the dimeric axis, similar to what is observed in CprB, which utilizes the disulfide bond as a fulcrum to rotate between the repressed and derepressed state (13). We assumed that the region between helices $\alpha 6$ and $\alpha 6'$ was flexible for rotation, for the electron density of this region was disordered due to its high flexibility. We further noticed three proline residues located at the loops connecting helices $\alpha 5$, $\alpha 6$, $\alpha 6'$, and $\alpha 7$ (Fig. 2), which were highly conserved in *Dietzia* strains (Fig. 7). These proline residues might contribute to rotations between two monomers. When the medium- or long-chain saturated fatty acids enter the binding pocket and occupy the cavity, the interactions between the ligand and the residues in the cavity decrease the flexibility of the entire molecule and restrict the rotation of the two monomers. In this way, helices $\alpha 6$ and $\alpha 6'$ play an important role in the conformational transition between the repressed and derepressed status.

The AlkX ligand-binding pocket. In contrast to the case for DesT and FadR, the entire ligand could enter the binding pocket of AlkX due to the relatively small carboxyl group. Therefore, there could be a balance between the hydrophobic (AlkX hydrophobic residues)/hydrophobic (alkyl chain of LCFA) attraction and hydrophobic (AlkX hydrophobic residues)/hydrophilic (carboxyl group of LCFA) repulsion forces. In general, the shorter the alkyl chains, the less the hydrophobic/hydrophobic attraction force. Given the same repulsion force between the carboxyl group and the AlkX hydrophobic residues, there should be a critical alkyl chain length with which the attractive force is balanced with the repulsion force. If the alkyl chain of LCFA is shorter than the critical length, the ligand probably could not remain inside the cavity even it enters into it. This agrees with our previous results showing that small organic acids such as acetic acid and propanoic acid cannot modulate the AlkX-DNA binding affinity (11). The obtaining of the AlkX-palmitic acid structure is in process for further verification of these assumptions.

In general, different ligand-binding pockets are needed to accommodate specific ligands. The multidrug-binding TFRs, such as QacR and TgtR, are often featured with their relatively large ligand-binding pockets and several polar residues in the pockets, which are used to bind diverse molecules, such as cationic lipophilic drugs, with the polar moieties (22, 27). These ligand-binding pockets are therefore less specific and can bind diverse ligands (12), while other TFRs, such as AlkX, DesT, and FadR, possess a tunnel-like cavity lined by hydrophobic residues specific for strong hydrophobic molecules with long alkyl or acyl chains (Fig. 3) (15, 28). The distinct feature of these ligand-binding pockets may result from the long-term adaptations of these bacteria to the environment. For example, the binding pocket of DesT accommodates saturated and unsaturated acyl-CoA, enabling the *P. aeruginosa* strain to adjust the membrane lipid composition for environmental adaptation (15). In contrast, the multidrug-binding pockets of QacR and TgtR protect the strains against the deleterious effects of toxic molecules encountered in the environment (22, 27). For *Dietzia* sp. strain DQ12-45-1b, isolated from an alkane-rich environment, AlkX accommodates medium- and long-chain fatty acids in the binding pocket, which derepresses the alkane hydroxylating gene and initiates the utilization of *n*-alkanes.

Possible DNA-binding mode of AlkX. When we inserted a 5-bp or 6-bp sequence at the center of the operator to change the location of the nucleotide pairs (i.e., OP4 and OP2) to the opposite side of the DNA double helix, AlkX could still bind to the operator (Fig. 6; see Fig. S4 in the supplemental material). It is indicated that the two monomers might not interact with each other in DNA-binding process, as illustrated in Fig. S5 in the supplemental material. Based on the present and previous results (Fig. 6) (11), the detailed binding mode of AlkX differs from the binding mode observed in other TFRs or protein families, such as DtxR (29). The two AlkX dimers contact the cognate operator side by side by binding to OP1/OP3 and OP2/OP4 (Fig. S5) without cross-dimer interactions, while other TFRs bind to the respective OP1/OP4 and OP2/OP3 (13). In this way, AlkX can bind to longer DNA up to 48 bp, which is in sharp contrast to the case for other dimers of DNA-binding proteins (QacR, 28 bp [21]; CgmR, 32 bp [18]; Ms6564, 31 bp [20]; and CprB, 22 bp [13]). There are 16 bp between OP1 and OP4, which is about 51 Å in the B-form DNA. If the AlkX dimer binds to OP1 and OP4, a large amount of energy is required to form a large bend in the DNA to shorten the distance and accommodate the protein structure. Instead, the distance between OP1 and OP3 is much shorter (36 Å in the B-form DNA), which requires less energy to form the protein-DNA complex. This hypothesis agrees with our previous isothermal titration calorimetry results, indicating the binding of AlkX to the operator and repression of gene transcription at a low protein concentration (11). Furthermore, our structural data indicated that the molecular ratio of inducer to protein was 1:1, and the binding affinity of the inducer appears to be higher than that of the DNA (11). Since *Dietzia* sp. strain DQ12-45-1b survived in a petroleum-related environment where nutrients and energy are less abundant, these bacteria typically maintain low protein concentrations and low energy cost to better adapt to these environments, although *n*-alkanes are ubiquitously present. This binding character may increase the rate of the transcription response (30) and allow for a fast switch-off when the concentration of alkanes is low.

The general derepression mechanism of AlkX may be common among *Dietzia* strains. Sequence alignment revealed a consensus sequence motif, "SRQT-Y" in $\alpha 3$ of AlkX and some TFRs in *Actinobacteria* (Fig. 7). Additionally, our previous study showed that *Dietzia* strains share the same IR (5'-TGGACAAA-N₁₁-TGTCTAGACA-N₁₁-TTTGTCTA-3') in the putative *alkB* promoter region. A conserved IR (5'-TTTACAAA-N₉-TTTGTCTA-3') was also identified in *Actinobacteria* such as *Mycobacterium*, *Rhodococcus*, and *Gordonia*, which was similar to the OP1 and OP2 motifs in *Dietzia* species (11). The similarity of the operators and the HTH motifs among the strains indicated a potential common DNA-binding mechanism among these *Actinobacteria*. However, the Phe-rich region of the ligand pocket is conserved only in *Dietzia* strains, suggesting that fatty acids may be a common inducer only in *Dietzia* species (Fig. 7). Since *Dietzia* strains live

TABLE 2 Strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s) ^a	Reference or source
Strains		
<i>E. coli</i> DH5 α	Cloning strain	TaKaRa
<i>Dietzia</i> sp. strain DQ12-45-1b	Grows on C ₁₀ –C ₄₀ alkanes	8
<i>E. coli</i> BL21(DE3)	Expression strain	New England BioLabs
<i>E. coli</i> BL21(pXL2801)	BL21 expressing AlkX; Km ^r	11
Cloning and expression vectors		
pET-28a	Expression vector in <i>E. coli</i> ; Km ^r	TaKaRa
pXL2801	pET-28a carrying <i>alkX</i> from the strain DQ12-45-1b for expression and purification of AlkX protein; Km ^r	11
pXL2803	pXL2801 with Ser63 replaced with glycine	This study
pXL2804	pXL2801 with Arg64 replaced with glycine	This study
pXL2805	pXL2801 with Gln65 replaced with glycine	This study
pXL2806	pXL2801 with Thr66 replaced with glycine	This study
pXL2807	pXL2801 with Leu67 replaced with glycine	This study
pXL2808	pXL2801 with Tyr68 replaced with glycine	This study

^aKm^r, kanamycin resistance.

in alkane-rich environments, similar selection pressures may have resulted in the evolution of similar regulation mechanisms.

Taken together, structural observation of AlkX has provided new insight into the inducer-regulator recognition and binding mode as well as the regulator-DNA binding mode. Based on the crystal structure of AlkX and the sequence alignment of TFRs in *Dietzia* strains, we hypothesize that AlkX represents a distinct subfamily of TetR proteins. Members of this subfamily likely have the following features: (i) they use fatty acids as effectors and recognize longer DNA sequences; (ii) the distance between the two HTH motifs increases upon ligand binding; and (iii) they function as transcriptional repressors in bacterial catabolism for *n*-alkane degradation. These hypotheses require further investigation in other members of the proposed subfamily. Future structural studies on the AlkX-DNA and AlkX-inducer complexes are required to gain a comprehensive understanding of the mechanism.

MATERIALS AND METHODS

Protein expression and purification. The native AlkX protein was prepared as described previously (11). Seleno-methionine (SeMet) protein was overexpressed in the *E. coli* BL21(DE3) strain containing plasmid pXL2801 (Table 2) for 18 h at 16°C. The recombinant strain was cultured in modified E9 minimal medium with 60 μ g/ml seleno-L-methionine and 50 μ g/ml kanamycin, as described previously (31). The harvested cell pellets were resuspended in lysis buffer (30 mM Tris-HCl [pH 8.3], 500 mM NaCl, 5% (wt/vol) glycerol, and 2 mM β -mercaptoethanol). Cells were lysed using a high-pressure breaking method (JN-R2C; Juneng Biology & Technology Co., Ltd., Guangzhou, China), and the lysate was centrifuged at 100,000 $\times g$ (Himac CP90WX; Hitachi, Tokyo, Japan) for 1 h at 4°C to remove unbroken cells and debris. The supernatant was applied three times to a nickel affinity column (Qiagen, Hilden, Germany) and then washed with lysis buffer containing 40 mM imidazole to remove nonspecifically bound proteins. The SeMet-AlkX protein was eluted with morpholineethanesulfonic acid (MES) buffer (30 mM MES [pH 6.3], 500 mM NaCl, and 2 mM β -mercaptoethanol) containing 300 mM imidazole. The target protein was desalted on Sephadex G-25 M (GE Healthcare, Little Chalfont, UK) in buffer containing 30 mM MES (pH 6.3), 500 mM NaCl, and 2 mM β -mercaptoethanol. Target protein fractions were collected and concentrated to 13 mg/ml for crystallization.

Crystallization. Initial crystallization trials were carried out at 20°C using commercial crystallization screening kits from Hampton Research (Aliso Viejo, CA, USA) and the sitting-drop method. Hits were optimized by handmade hanging drops. To obtain *Dietzia* sp. strain DQ12-45-1b SeMet-AlkX crystals, 1.5 μ l protein (0.5 mM) was mixed with 1 μ l reservoir solution containing 0.01 M magnesium acetate, 0.05 M sodium cacodylate (pH 6.5), and 1.3 M lithium sulfate over a 200- μ l reservoir at 20°C. Irregular triangle-shaped crystals grew to maximum dimensions of 200 by 120 by 60 μ m³ in 1 day. Crystals were harvested at 20°C and subjected to flash cooling in liquid nitrogen for storage and data collection.

Data collection and structure determination. All diffraction data sets were collected at beamlines 18U and 19U at the Shanghai Synchrotron Radiation Facility and processed using HKL2000 (32). One data set from an SeMet-AlkX crystal was collected at a wavelength of 0.9762 Å to a resolution of 3.1 Å. The structure was determined using the single-wavelength anomalous diffraction method. Identification of 14 out of the 21 expected selenium sites, the phasing, primary model building, and structure refinement were performed using the Phenix software suite (33). The models were rebuilt manually using Coot

TABLE 3 Oligonucleotides used in this study

Oligonucleotide	Sequence (5'→3') ^a	Usage
alkXFN	GGAATTCATATGAGATCAGGAACAATGGACCGC	Mutation
alkXRH	CCCAAGCTTTACTGCTGGACGTCGACCTG	Mutation
S63GF	GCGGGACTGGGCCGAGACG	Mutation
S63GR	CGTCTGCCGGCCAGTCCCGC	Mutation
R64GF	GGACTGAGCGGCAGACGCTC	Mutation
R64GR	GAGCGTCTGCCGCTCAGTCC	Mutation
Q65GF	CTGAGCCGGGGACGCTCTAC	Mutation
Q65GR	GTAGAGCGTCCCCGGCTCAG	Mutation
T66GF	AGCCGGCAGGGCTCTACAGC	Mutation
T66GR	GCTGTAGAGCCCTGCCGGCT	Mutation
L67GF	CGGCAGACGGGTACAGCACC	Mutation
L67GR	GGTGCTGTAGCCGCTGCGG	Mutation
Y68GF	CAGACGCTCGGCAGACCTTC	Mutation
Y68GR	GAAGGTGCTGCCGAGCGTCTG	Mutation
W-DIG	TAGCCACGCTGGACAAAGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
W-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGACACGATCTCTCACTTTGTCCAGACGTGGCTA	EMSA
OP1-1-DIG	TAGCCACGCTACTACAAAGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP1-1-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGACACGATCTCTCACTTTGTAGTGACGTGGCTA	EMSA
OP1-2-DIG	TAGCCACGCTCTGGGTCAAGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP1-2-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGACACGATCTCTCACTTGACCCAGACGTGGCTA	EMSA
OP1-3-DIG	TAGCCACGCTCTGGACCGCGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP1-3-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGACACGATCTCTACGCGGTCCAGACGTGGCTA	EMSA
C+5-DIG	TAGCCACGCTCTGGACAAAGTGAGAGATCGTGTCTCCCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
C+5R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAAGGGAGACACGATCTCTCACTTTGTCCAGACGTGGCTA	EMSA
OP3-1-DIG	TAGCCACGCTCTGGACAAAGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP3-1R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGGAGCGATCTCTCACTTTGTCCAGACGTGGCTA	EMSA
OP3-2-DIG	TAGCCACGCTCTGGACAAAGTGAGAGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP3-2R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGCAGCGATCTCTCACTTTGTCCAGACGTGGCTA	EMSA
OP13-23-DIG	TAGCCACGCTCTGGACCGCGTGAGAGATCGTGTCTAGACAACGCCACGGTTCGCGTCTAACCTCGACCACA	EMSA
OP13-23-R	TGTGGTCGAGGTTAGACGCGAACCCTGGCGTTGTCTAGACACGATCTCTACGCGGTCCAGACGTGGCTA	EMSA
OP31-41-DIG	TAGCCACGCTCTGGACAAAGTGAGAGATCGTGTCTAGCTACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
OP31-41-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTGAGCTAGGAGCGATCTCTCACTTTGTCCAGACGTGGCTA	EMSA
70+1-DIG	TAGCCACGCTCTGGACAAAGTGAGACGATCGTGTCTAGACAACGCCACGGTTTTGTCTAACCTCGACCACA	EMSA
70+1-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAGACACGATCGTCTCACTTTGTCCAGACGTGGCTA	EMSA
C+6-DIG	TAGCCACGCTCTGGACAAAGTGAGAGATCGTGTCTCCCTTTAGACAACGCCACGGTTTTGTCTAACCTCGACCA	EMSA
C+6-R	TGTGGTCGAGGTTAGACAAAAACCGTGGCGTTGTCTAAGGGAGACACGATCTCTCACTTTGTCCAGACGTGGC	EMSA

^aA restriction enzyme site is underlined.

software (34). Data collection and refinement statistics are summarized in Table 1. Structural figures were generated using PyMOL (35).

Determination of the ligand in AlkX. We performed GC-MS to determine the ligand in AlkX. Briefly, the purified AlkX (15 mg) was freeze-dried overnight and then transferred to glass test tubes with caps. Methyl esterification of protein-bound fatty acid was accomplished using a mixture of 1 ml 100% (vol/vol) toluene and 2 ml 1% (vol/vol) sulfuric acid in methanol overnight at 50°C. Heptadecanoic acid (0.8 mg ml⁻¹) was added for an internal standard. The mixtures were added in 5 ml water containing 5% NaCl and extracted with hexane (1:1, vol/vol) twice. The resulting samples were washed with 2% KHCO₃ and evaporated with N₂. The resulting products were then dissolved in 1 ml hexane and analysis by GC-MS using an Agilent 7890C gas chromatograph and a 5975C mass spectrometer (Agilent, Santa Clara, CA, USA) equipped with an Agilent HP-5AS capillary column (30 m by 0.25 mm). Helium was the carrier gas at a flow rate of 2.0 ml min⁻¹. One microliter of sample was injected via a split injection port at a split ratio of 1:30. The injection port temperature was set at 260°C. The column temperature was kept initially at 100°C for 1 min and then increased at 20°C min⁻¹ to 260°C and held for 5 min, and then the temperature was increased to 280°C at 20°C min⁻¹ and held for 5 min. The fatty acid methyl esters were identified by chromatographic comparison with authentic standards (National Institute of Standards and Technology). The experiment was performed at least three times. A negative-control experiment (sample without AlkX) was also performed to show that there were no contaminants in the GC-MS samples.

Site-directed mutagenesis. Genes of the S63G, R64G, Q65G, T66G, L67G, and Y68G mutants were constructed using fusion PCR methods (Table 3) and then confirmed by DNA sequencing. The mutant proteins were expressed (Table 2) and purified following the method described for the recombinant protein AlkX. Expression and purification of the mutant proteins were similar to those of wild-type (WT) protein as described previously (11).

EMSA. For electrophoretic mobility shift assay (EMSA), 70-bp oligonucleotides of two complementary strands, including the WT and mutated 48-bp operator labeled with digoxigenin at the 5' terminus as the top strands, were synthesized directly by Sangon Biotech Corporation (Shanghai, China) (Table 3). There were three kinds of mutations, including three-point mutation in OP1 to OP4, respectively (to determine the more important binding sites), one base pair added in the nonpalindromic sequence between OP1

and OP3 (to investigate the role of the nonpalindromic sequence), and five and six base pairs added at the center of the operator (to explore the AlkX-DNA binding mode). Annealing of DNA and EMSA were conducted as previously described (11). Briefly, 0.1 μ M labeled WT or mutated DNA fragment and 1 μ M WT or mutant AlkX were mixed and incubated in binding buffer [10 mM Tris-HCl (pH 7.6), 100 mM KCl, 1 mM EDTA, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 1 mM dithiothreitol, and 0.2% (wt/vol) Tween 20]. After incubation at room temperature for 30 min, the samples were separated by nondenaturing 5% polyacrylamide gel electrophoresis in 0.05 M Tris-borate-EDTA buffer (pH 8.3) for 2 h at 10 V/cm and transferred to positively charged nylon membranes for 30 min at 300 mA. Labeled DNA was visualized using a digoxigenin detection kit (Roche, Mannheim, Germany). Gels from three independent EMSAs were photographed, and the bands (AlkX-DNA binding bands and free DNA bands) were quantified by densitometric analysis using NIH Image J software.

DNA sequences analysis. Primer design and sequence analysis were performed using Vector NTI Advanced 11 (Invitrogen, Carlsbad, CA, USA). The initial alignment of AlkX and DesT in Fig. 2 was generated using the MultAlin server (<http://multalin.toulouse.inra.fr/multalin/>) (36) and subsequently formatted using ESPrnt (37). Multiple-sequence alignments of the TFRs in Fig. 7 were generated using ClustalW2 (38).

Accession number(s). The atomic coordinates and structure factors of the SeMet-substituted AlkX have been deposited in the Protein Data Bank with accession code 5YDP.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.01447-17>.

SUPPLEMENTAL FILE 1, PDF file, 0.6 MB.

ACKNOWLEDGMENTS

We thank the staff of the Protein Research Core Facility at the Institute of Biophysics, Chinese Academy of Sciences (CAS), for their excellent technical assistance; particular thanks go to Xiaoxia Yu. We are grateful to staff members of the SSRF (China) and Photon Factory (Japan) synchrotron facilities. We thank Hai-tao Li for helpful discussion on the manuscript.

This work was supported by the National Natural Science Foundation of China (31225001 to X.-L.W., 31300108 to Y.N., 31470745 to X.C.Z., and 31600082 to J.-L.L.) and the National High Technology Research and Development Program of China (2012AA02A703).

All authors declare no conflict of interest.

REFERENCES

- Bertrand JC, Almallah M, Acquaviva M, Mille G. 1990. Biodegradation of hydrocarbons by an extremely halophilic archaeobacterium. *Lett Appl Microbiol* 11:260–263. <https://doi.org/10.1111/j.1472-765X.1990.tb00176.x>.
- Nie Y, Chi CQ, Fang H, Liang JL, Lu SL, Lai GL, Tang YQ, Wu XL. 2014. Diverse alkane hydroxylase genes in microorganisms and environments. *Sci Rep* 4:4968. <https://doi.org/10.1038/srep04968>.
- Brown LR. 2010. Microbial enhanced oil recovery (MEOR). *Curr Opin Microbiol* 13:316–320. <https://doi.org/10.1016/j.mib.2010.01.011>.
- Rojo F. 2009. Degradation of alkanes by bacteria. *Environ Microbiol* 11:2477–2490. <https://doi.org/10.1111/j.1462-2920.2009.01948.x>.
- van Beilen JB LZ, Duetz WA, Smits THM, Witholt B. 2003. Diversity of alkane hydroxylase systems in the environment. *Oil Gas Sci Technol* 58:427–440. <https://doi.org/10.2516/ogst.2003026>.
- Nie Y, Liang J, Fang H, Tang YQ, Wu XL. 2011. Two novel alkane hydroxylase-rubredoxin fusion genes isolated from a *Dietzia* bacterium and the functions of fused rubredoxin domains in long-chain *n*-alkane degradation. *Appl Environ Microbiol* 77:7279–7288. <https://doi.org/10.1128/AEM.00203-11>.
- Liang JL, JiangYang JH, Nie Y, Wu XL. 2016. Regulation of the alkane hydroxylase gene CYP153 in a Gram-positive alkane-degrading bacterium, *Dietzia* sp. strain DQ12-45-1b. *Appl Environ Microbiol* 82:608–619. <https://doi.org/10.1128/AEM.02811-15>.
- Wang XB, Chi CQ, Nie Y, Tang YQ, Tan Y, Wu G, Wu XL. 2011. Degradation of petroleum hydrocarbons (C6–C40) and crude oil by a novel *Dietzia* strain. *Bioresour Technol* 102:7755–7761. <https://doi.org/10.1016/j.biortech.2011.06.009>.
- Lu SL, Nie Y, Tang YQ, Xiong GM, Wu XL. 2014. A critical combination of operating parameters can significantly increase the electrotransforma-
- tion efficiency of a Gram-positive *Dietzia* strain. *J Microbiol Methods* 103:144–151. <https://doi.org/10.1016/j.mimet.2014.05.015>.
- Nie Y, Liang JL, Fang H, Tang YQ, Wu XL. 2014. Characterization of a CYP153 alkane hydroxylase gene in a Gram-positive *Dietzia* sp. DQ12-45-1b and its “team role” with *alkW1* in alkane degradation. *Appl Microbiol Biotechnol* 98:163–173. <https://doi.org/10.1007/s00253-013-4821-1>.
- Liang JL, Nie Y, Wang M, Xiong G, Wang YP, Maser E, Wu XL. 2016. Regulation of alkane degradation pathway by a TetR family repressor via an autoregulation positive feedback mechanism in a gram-positive *Dietzia* bacterium. *Mol Microbiol* 99:338–359. <https://doi.org/10.1111/mmi.13232>.
- Cuthbertson L, Nodwell JR. 2013. The TetR family of regulators. *Microbiol Mol Biol Rev* 77:440–475. <https://doi.org/10.1128/MMBR.00018-13>.
- Bhukya H, Bhujbalrao R, Bitra A, Anand R. 2014. Structural and functional basis of transcriptional regulation by TetR family protein CprB from *S. coelicolor* A3(2). *Nucleic Acids Res* 42:10122–10133. <https://doi.org/10.1093/nar/gku587>.
- Orth P, Schnappinger D, Hillen W, Saenger W, Hinrichs W. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* 7:215–219. <https://doi.org/10.1038/73324>.
- Miller DJ, Zhang YM, Subramanian C, Rock CO, White SW. 2010. Structural basis for the transcriptional regulation of membrane lipid homeostasis. *Nat Struct Mol Biol* 17:971–975. <https://doi.org/10.1038/nsmb.1847>.
- Le TB, Schumacher MA, Lawson DM, Brennan RG, Buttner MJ. 2011. The crystal structure of the TetR family transcriptional repressor SimR bound to DNA and the role of a flexible N-terminal extension in minor groove

- binding. *Nucleic Acids Res* 39:9433–9447. <https://doi.org/10.1093/nar/gkr640>.
17. Grkovic S, Brown MH, Roberts MJ, Paulsen IT, Skurray RA. 1998. QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. *J Biol Chem* 273:18665–18673. <https://doi.org/10.1074/jbc.273.29.18665>.
 18. Itou H, Watanabe N, Yao M, Shirakihara Y, Tanaka I. 2010. Crystal structures of the multidrug binding repressor *Corynebacterium glutamicum* CgmR in complex with inducers and with an operator. *J Mol Biol* 403:174–184. <https://doi.org/10.1016/j.jmb.2010.07.042>.
 19. Tonthat NK, Milam SL, Chinnam N, Whitfill T, Margolin W, Schumacher MA. 2013. SlmA forms a higher-order structure on DNA that inhibits cytokinetic Z-ring formation over the nucleoid. *Proc Natl Acad Sci U S A* 110:10586–10591. <https://doi.org/10.1073/pnas.1221036110>.
 20. Yang S, Gao Z, Li T, Yang M, Zhang T, Dong Y, He ZG. 2013. Structural basis for interaction between *Mycobacterium smegmatis* Ms6564, a TetR family master regulator, and its target DNA. *J Biol Chem* 288:23687–23695. <https://doi.org/10.1074/jbc.M113.468694>.
 21. Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG. 2002. Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* 21:1210–1218. <https://doi.org/10.1093/emboj/21.5.1210>.
 22. Alguel Y, Meng C, Terán W, Krell T, Ramos JL, Gallegos MT, Zhang X. 2007. Crystal structures of multidrug binding protein TtgR in complex with antibiotics and plant antimicrobials. *J Mol Biol* 369:829–840. <https://doi.org/10.1016/j.jmb.2007.03.062>.
 23. Agari Y, Agari K, Sakamoto K, Kuramitsu S, Shinkai A. 2011. TetR-family transcriptional repressor *Thermus thermophilus* FadR controls fatty acid degradation. *Microbiology* 157:1589–1601. <https://doi.org/10.1099/mic.0.048017-0>.
 24. Le TB, Stevenson CE, Fiedler HP, Maxwell A, Lawson DM, Buttner MJ. 2011. Structures of the TetR-like simocyclinone efflux pump repressor, SimR, and the mechanism of ligand-mediated derepression. *J Mol Biol* 408:40–56. <https://doi.org/10.1016/j.jmb.2011.02.035>.
 25. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R. 2005. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* 69:326–356. <https://doi.org/10.1128/MMBR.69.2.326-356.2005>.
 26. Sawai H, Yamanaka M, Sugimoto H, Shiro Y, Aono S. 2012. Structural basis for the transcriptional regulation of heme homeostasis in *Lactococcus lactis*. *J Biol Chem* 287:30755–30768. <https://doi.org/10.1074/jbc.M112.370916>.
 27. Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG. 2001. Structural mechanisms of QacR induction and multidrug recognition. *Science* 294:2158–2163. <https://doi.org/10.1126/science.1066020>.
 28. Müller G, Hecht B, Helbl V, Hinrichs W, Saenger W, Hillen W. 1995. Characterization of non-inducible Tet repressor mutants suggests conformational changes necessary for induction. *Nat Struct Mol Biol* 2:693–703. <https://doi.org/10.1038/nsb0895-693>.
 29. Chen CSY, White A, Love J, Murphy JR, Ringe D. 2000. Methyl groups of thymine bases are important for nucleic acid recognition by DtxR. *Biochemistry* 39:10397–10407. <https://doi.org/10.1021/bi0009284>.
 30. Rosenfeld N, Elowitz MB, Alon U. 2002. Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* 323:785–793. [https://doi.org/10.1016/S0022-2836\(02\)00994-4](https://doi.org/10.1016/S0022-2836(02)00994-4).
 31. Jiang D, Zhao Y, Wang X, Fan J, Heng J, Liu X, Feng W, Kang X, Huang B, Liu J, Zhang XC. 2013. Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proc Natl Acad Sci U S A* 110:14664–14669. <https://doi.org/10.1073/pnas.1308127110>.
 32. Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276:307–326. [https://doi.org/10.1016/S0076-6879\(97\)76066-X](https://doi.org/10.1016/S0076-6879(97)76066-X).
 33. Echols N, Grosse-Kunstleve RW, Afonine PV, Bunkóczi G, Chen VB, Headd JJ, McCoy AJ, Moriarty NW, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Adams PD. 2012. Graphical tools for macromolecular crystallography in PHENIX. *J Appl Crystallogr* 45:581–586. <https://doi.org/10.1107/S0021889812017293>.
 34. Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132. <https://doi.org/10.1107/S0907444904019158>.
 35. Schrodinger LLC. 2010. The PyMOL molecular graphics system, version 1.3r1. Schrodinger LLC, New York, NY.
 36. Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16:10881–10890. <https://doi.org/10.1093/nar/16.22.10881>.
 37. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320–W324. <https://doi.org/10.1093/nar/gku316>.
 38. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>.