Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural and functional analysis of the TetR-family transcriptional regulator SCO0332 from *Streptomyces coelicolor*

Ui Okada,[a] Kazuya Kondo,[a] Takeshi Hayashi,[b] Nobuhisa Watanabe,[a,c] Min Yao,[a,c] Tomohiro Tamura[d] and Isao Tanaka[a,c]*

[a]Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, [b]Department of Food and Bioscience, Faculty of Food and Nutrition, Beppu University, Beppu 874-8501, Japan, [c]Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan, and [d]Research Institute of Genome-based Biofactory, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062-8517, Japan

Correspondence e-mail: tanaka@castor.sci.hokudai.ac.jp

Received 31 October 2007
Accepted 16 November 2007

PDB Reference: SCO0332, 2zb9, r2zb9sf.

SCO0332 protein is a putative TetR-family transcriptional regulator from *Streptomyces coelicolor* A3(2). The crystal structure of SCO0332 was determined at 2.25 Å resolution by single-wavelength anomalous diffraction (SAD) phasing using the S atoms of the native protein. SCO0332 contains a helix–turn–helix (HTH) DNA-binding motif in its N-terminal region and forms a homodimer. The overall structure of SCO0332 shows significant similarity to other TetR-family regulators. A systematic evolution of ligands by exponential enrichment (SELEX) analysis indicated that SCO0332 has sequence-specific DNA-binding ability and determined the position of the operator element of SCO0332 on the chromosomal DNA of *S. coelicolor*. An electrophoretic mobility-shift assay (EMSA) showed that SCO0332 binds to the operator sequence upstream of the *sco0330* gene, which encodes a putative short-chain oxidoreductase. These results suggest that SCO0332 is a transcriptional repressor that regulates *sco0330* gene expression.

1. Introduction

In bacteria, adaptation to changes in living conditions is essential for survival. In many cases, these adaptive responses are mediated by transcriptional regulators that control the expression levels of functional proteins. TetR is a transcriptional regulator that controls the expression of the *tet* gene required for resistance to tetracycline. The three-dimensional structure of the TetR-family proteins consists of two domains: a DNA-binding domain and a regulatory domain that recognizes the signals *via* ligand binding (Ramos *et al.*, 2005; Hinrichs *et al.*, 1994; Kisker *et al.*, 1995). Members of this family regulate genes encoding biosynthetic enzymes for antibiotics, drug-efflux pumps and other proteins. TetR-family proteins show high sequence similarity in the DNA-binding domain, which contains a helix–turn–helix (HTH) motif (Wintjens & Rooman, 1996; Huffman & Brennan, 2002). However, there is no significant similarity in the regulatory domains, corresponding to their various signalling molecules (Ramos *et al.*, 2005).

The Gram-positive soil-inhabiting bacterium *Streptomyces coelicolor* (Bentley *et al.*, 2002) is known to produce various bioactive metabolites (Donadio *et al.*, 2002), the production of which is controlled by transcriptional regulators. Genome analysis revealed that *S. coelicolor* contains 150 genes belonging to the TetR family of transcriptional regulators (Ramos *et al.*, 2005), which is markedly greater than the number in *Escherichia coli* and *Bacillus subtilis*, which have 13

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution bin.

Data collection	
Wavelength (Å)	2.29
Temperature (K)	93
Space group	$P2_1$
Unit-cell parameters	
a (Å)	52.4
b (Å)	77.8
c (Å)	57.3
β (°)	104.3
Resolution (Å)	50.0–2.20 (2.28–2.20)
Unique reflections	21856 (1740)
Completeness (%)	96.4 (77.3)
Redundancy	7.1 (6.1)
$I/\sigma(I)$	46.3 (13.8)
$R_{\rm merge}$† (%)	4.8 (13.6)
Refinement	
Resolution (Å)	20.0–2.25 (2.33–2.25)
Unique reflections	20828 (2013)
Completeness (%)	98.2 (95.5)
$R_{\rm work}/R_{\rm free}$ (%)	19.5/24.1
No. of non-H atoms	
Protein	2881
Solvent	242
R.m.s. deviations	
Bond lengths (Å)	0.0052
Bond angles (°)	1.2
Mean B factor (Å^2)	
Protein	39.3
Solvent	35.3
Ramachandran plot (%)	
Most favoured	92.4
Additional allowed	7.6

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \overline{I(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\overline{I(hkl)}$ is the average intensity obtained from multiple observations of symmetry-related reflections.

and 20 TetR-family genes, respectively. However, the signalling molecules or transcriptional regulation mechanisms of these TetR-family genes, which were annotated based only on sequence information, are mostly unknown. Investigation of such uncharacterized transcriptional regulators may lead to the discovery of new antibiotics. Thus, functional analysis of putative transcriptional regulators is not only important to gain improved biological understanding, but also in medical, agricultural and industrial development.

The results of crystal structural analysis provide important insights into the functions of putative transcriptional regulators. Systematic evolution of ligands by exponential enrichment (SELEX) analysis (Ellington & Szostak, 1990; Tuerk & Gold, 1990) also provides significant information. The SELEX procedure explores the consensus sequence for nucleic acid-binding proteins through repeated cycles of binding selection and PCR amplification. In the standard SELEX procedure, the original random-sequence pool for selection is generated by PCR with synthesized oligonucleotides containing a random-sequence region as a template. In genomic SELEX, however, the sequence library is derived from the chromosomal DNA of the target organism (Shimada *et al.*, 2005; Itou *et al.*, 2005). Hence, genomic SELEX can directly determine the binding position of the target protein on its genome sequence.

Here, we report the crystal structure of the putative TetR-family transcriptional regulator SCO0332 from *S. coelicolor* at 2.25 Å resolution. The primary functional characterization of SCO0332 was carried out using further information obtained by genomic SELEX analysis and electrophoretic mobility-shift assay (EMSA).

2. Materials and methods

2.1. Protein preparation

The *sco0332* gene encoding SCO0332, a 206-amino-acid protein, was amplified by PCR using *S. coelicolor* A3(2) genomic DNA as a template. The PCR product was introduced into the pTip vector (Nakashima & Tamura, 2004*a*,*b*); this expression vector works in several *Rhodococcus* species. The pTip vector contains a *tipA* promoter, from which protein expression is induced by thiostrepton. SCO0332 protein was expressed in *R. erythropolis* L88 (Mitani *et al.*, 2005) as a C-terminal six-histidine-tagged fusion protein. Cells were grown in Luria–Bertani (LB) medium at 303 K until the OD_{600} reached 0.4. Protein expression was then induced by the addition of 1 µg ml^{-1} thiostrepton and the culture was continued for an additional 20 h. Harvested cells were resuspended in buffer *A* (50 m*M* sodium phosphate pH 8.0, 300 m*M* NaCl), treated with lysozyme at a final concentration of 1.0 mg ml^{-1} for 1 h at 277 K and disrupted by sonication. After centrifugation at 40 000*g* for 1 h at 283 K, the supernatant was applied onto HiTrap Chelating HP (GE Healthcare) charged with nickel ions. After extensive washing, the target protein was eluted with a linearly increasing concentration of imidazole. Further purification was performed on HiLoad Superdex 200 pg columns (GE Healthcare) equilibrated with buffer *B* (50 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 10% glycerol).

2.2. Crystallization and data collection

SCO0332 protein at a concentration of 15 mg ml^{-1} was crystallized at 293 K using the hanging-drop vapour-diffusion method with a reservoir solution consisting of 0.1 *M* HEPES pH 7.0, 7–10% polyethylene glycol 3350 and 0.05–0.1 *M* magnesium chloride. Diffraction data were collected to 2.2 Å resolution at 93 K using an in-house Cr $K\alpha$ X-ray source (Rigaku FR-E SuperBright) with the bufferless crystal-mounting method (Kitago *et al.*, 2005; Watanabe, 2006). These diffraction images were processed using the *HKL*-2000 program (Otwinowski & Minor, 1997). The SCO0332 crystal belongs to space group $P2_1$, with unit-cell parameters a = 52.4, b = 77.8, c = 57.3 Å, β = 104.3°; the asymmetric unit contained two molecules of SCO0332 protein.

2.3. Structure determination and refinement

The structure of SCO0332 was solved by single-wavelength anomalous diffraction (SAD) phasing using the S atoms of the native SCO0332 protein as anomalous scatterers. Four of the six sulfur sites in the asymmetric unit were found by the program *SHELXD* (Schneider & Sheldrick, 2002) using the

HKL2MAP (Pape & Schneider, 2004) interface; the sites were refined and primary phases were calculated using the program *SOLVE* (Terwilliger & Berendzen, 1999). Phase improvement and automated model building were carried out with *RESOLVE* (Terwilliger, 1999) and *REFMAC*5 (Murshudov *et al.*, 1997) using the *RESOLVE BUILD* script. After iterative cycles of automatic model refinement using the program *LAFIRE* (Yao *et al.*, 2006) with *CNS* (Brünger *et al.*, 1998), the model was checked and fitted manually with the molecular-graphics program *Coot* (Emsley & Cowtan, 2004). The final refinement statistics are summarized in Table 1.

2.4. Genomic SELEX analysis

The DNA library for the search for aptamers of SCO0332 was prepared as follows. The genomic DNA of *S. coelicolor* was digested into short fragments of approximately 50–200 bp in length with the restriction enzyme *Hae*III. These blunt-ended fragments were ligated into the vector pBR322 at the *Eco*RV restriction site. The resulting vectors with inserts of various sequences were transformed into *E. coli* strain XL1-Blue. Over 200 000 colonies were collected and plasmids were extracted using a QIAprep Spin Miniprep Kit (Qiagen). The library was amplified by PCR with primers *Eco*RV-F (5′-CT-TGGTTATGCCGGTACTGC-3′) and *Eco*RV-R (5′-GCGA-TGCTGTCGGAATGGAC-3′), corresponding to sequences around the *Eco*RV restriction site of pBR322.

The PCR products were mixed with 60 µg purified SCO0332 protein captured using magnetic beads (MagExtractor-His-tag; Toyobo) *via* the six-histidine tag in binding buffer (20 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 10 m*M* imidazole). The mixture was incubated at room temperature for 1 h. After washing to remove nonspecific binding, protein–DNA complexes were eluted with binding buffer containing 500 m*M* imidazole. Eluted DNA fragments were extracted by phenol–chloroform extraction and ethanol precipitation and amplified by PCR with primers *Eco*RV-F and *Eco*RV-R. The PCR products were used as a DNA pool for the next cycle of selection. This process was repeated for two cycles. After two cycles of selection, selected DNA fragments were cloned into

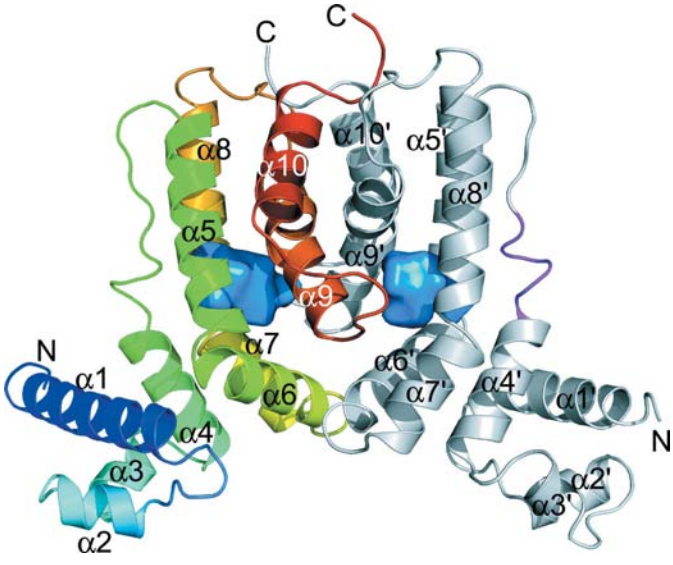

Figure 1
Ribbon representation of the SCO0332 dimer. Colour coding in chain *A* is from blue (N-terminus) to red (C-terminus). Chain *B* is shown in grey. The putative ligand-binding cavities are shown in marine blue and the loop (Val79′–Ala84′) covering the putative ligand-binding cavity of chain *B* is coloured magenta. This figure and the following structural figures were generated using *PyMOL* (DeLano Scientific LLC).

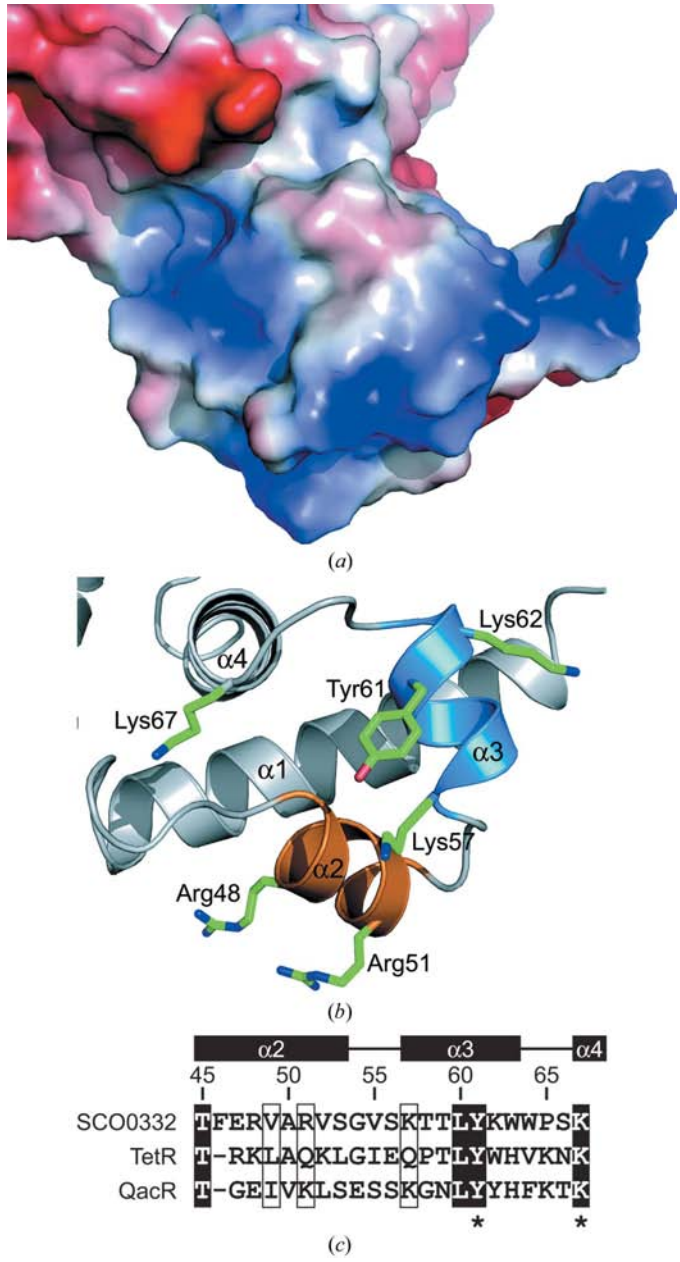

Figure 2
DNA-binding domain of SCO0332. (*a*) Electrostatic surface representation coloured according to surface potential (blue, positive; red, negative) of the DNA-binding domain in chain *A*. (*b*) Close-up view of the HTH motif in chain *A*. Positively charged side chains are shown as stick models coloured as follows: carbon, green; nitrogen, blue; oxygen, red. The α2 and α3 helices are coloured orange and marine blue, respectively. (*c*) Amino-acid sequence alignment of SCO0332, TetR from *E. coli* and QacR from *S. aureus* around the HTH motif with secondary structure. Conserved resides are boxed. Residues that are highly conserved among TetR-family members are indicated by asterisks.

the pGEM-T Easy Vector (Promega) and transformed into *E. coli* strain JM109 for sequence analysis.

2.5. Electrophoretic mobility-shift assay (EMSA)

The synthetic oligonucleotides P1′ (5′-GTTATATGAACGTGCCGTTCATTAGAG-3′) and P2′ (5′-CGACTTCCAGGAGATCCCCATGAGCAC-3′), which were the candidates for the operator sequence of SCO0332 determined by genomic SELEX, 5′-fluorescein isothiocyanate (FITC) labelled P1′ and their complementary oligonucleotides were obtained commercially (Hokkaido System Science). The complementary strands were mixed in equimolar amounts and the mixtures were heated at 358 K for 10 min and then cooled slowly to 277 K overnight for annealing.

Purified SCO0332 and annealed P1′ or P2′ were mixed in buffer (20 m*M* Tris–HCl pH 7.5). In some cases, the non-specific competitor DNA poly(dI-dC)-(dI-dC) was added. The mixtures were incubated for 2 h at room temperature and loaded onto 10% nondenaturing polyacrylamide gels in TBE buffer (50 m*M* Tris–HCl pH 7.8, 50 m*M* boric acid, 2 m*M* EDTA). Electrophoresis was performed in TBE buffer at 100 V for 90 min at 277 K. The DNA in the polyacrylamide gels was detected by ethidium bromide staining and the protein was detected by CBB staining. FITC-labelled DNA was visualized using an LED light source (470 nm).

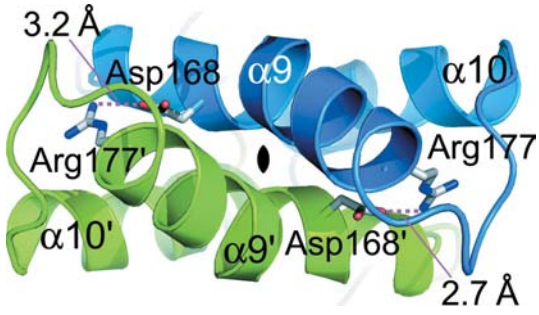

Figure 3
Close-up view of the dimer formation of helices α9 and α9′. Chains *A* and *B* are coloured marine blue and lime green, respectively. The side chains of Asp168 and Arg177 in chain *A* and of Asp168′ and Arg177′ in chain *B* are shown as stick models coloured as follows: carbon, grey; nitrogen, blue; oxygen, red. The black symbol represents the pseudo-twofold axis of the dimer.

3. Results

3.1. Overall structure

The crystal structure of SCO0332 was determined at 2.25 Å resolution and refined to a crystallographic *R* factor (R_{work}) and free *R* factor (R_{free}) of 19.5% and 24.1%, respectively. The asymmetric unit of the crystal contains two SCO0332 monomers. The final model contains chain *A* (Gly17–Arg204), chain *B* (Glu22′–Pro203′) and 246 water molecules. The N- and C-terminal regions were not visible in the electron-density map. A Ramachandran plot calculated using the *PROCHECK* program (Laskowski *et al.*, 1993) showed that 92.4% of the residues are in most favoured regions and 7.6% are in additional allowed regions.

The SCO0332 monomer is comprised of ten helices (Fig. 1): α1 (Pro20–Glu39), α2 (Phe46–Ser53), α3 (Lys57–Trp63), α4 (Lys67–Val79), α5 (Val91–Arg108), α6 (Pro110–Ala122), α7 (Ala126–Leu135), α8 (Ser137–Glu153), α9 (Val163–Leu179) and α10 (Asp187–Asp199). The monomer structure is divided into an N-terminal DNA-binding domain (α1–α3) and a C-terminal dimerization domain (α4–α10). The DNA-binding domain contains a typical HTH motif (α2–α3) in which the α3 helix is the DNA-recognition helix. The DNA-binding domain of SCO0332 has a positively charged surface consisting of Arg48, Arg51, Lys57, Lys62 and Lys67 in the α2 and α3 helices and the amino end of the α4 helix (Fig. 2). This positively charged molecular surface may be used to interact with the phosphate backbone of DNA.

Two monomers are related by a pseudo-twofold rotation and the resulting dimer is the functional unit of SCO0332, which is similar to other TetR regulators. The results of size-exclusion chromatography also indicated that SCO0332 forms a dimer in solution (data not shown). The dimer structure is mainly stabilized by hydrophobic interactions and two intermolecular salt bridges between the α9 and α9′ helices from each monomer (Fig. 3). Salt bridges were observed between Asp168′ in α9 and Arg177′ in α9′ with a distance of 3.2 Å and between Asp168 in α9′ and

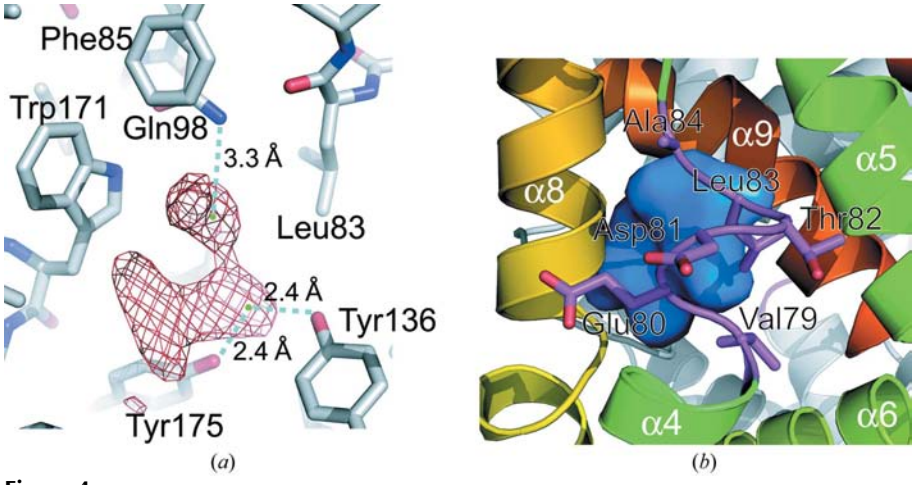

Figure 4
Putative ligand-binding cavity of SCO0332. (*a*) Close-up view of the cavity in chain *A*. The $F_o - F_c$ difference electron-density map of the uncharacterized ligand contoured at 2.5σ is shown as a red mesh. A similar density map was also observed in chain *B* at the corresponding position. The distances between the centre points of the electron-density spheres and the N atom of Gln98 and the O atoms of Tyr136 and Tyr175 are indicated. (*b*) Side view of the cavity. The helices are coloured in the same way as in Fig. 1. The main chain of the loop (Val79–Ala84) is coloured magenta and the side chains are shown as stick models coloured as follows: carbon, magenta; oxygen, red.

Arg177 in α9 with a distance of 2.7 Å. Hydrophobic interactions between α6 and α6′ or α10 and α10′ also stabilize dimer formation.

A structural similarity search using the DALI (Holm & Sander, 1998) server indicated that the overall structure of SCO0332 resembles those of the TetR family proteins, such as QacR (PDB code 1rkw; Murray *et al.*, 2004), a TetR-family regulator from *Staphylococcus aureus* involved in regulation of multidrug resistance, with a Z score of 14.3 and an r.m.s. deviation of 3.0 Å for 175 C^{α} atoms, and TetR from *E. coli* (PDB code 2trt; Hinrichs *et al.*, 1994), with a Z score of 10.6 and an r.m.s. deviation of 4.1 Å for 159 C^{α} atoms.

3.2. Putative ligand-binding site

SCO0332 has a cavity of dimensions 5 × 5 × 10 Å at the position corresponding to the ligand-binding site in QacR (Schumacher *et al.*, 2001, 2004; Fig. 1). Some electron density was observed in this cavity (Fig. 4*a*). However, we could not interpret the map as none of the reagents used in the process of protein purification and crystallization fitted the difference map. The uncharacterized ligand may have been captured by SCO0332 from living *R. erythropolis* cells during fermentation. The surface of this cavity is mainly covered with hydrophobic residues: Leu83, Phe85, Phe102 and Trp171. The three hydrophilic residues Gln98, Tyr136 and Tyr175 are also at interacting distances. The distances from the characteristic points in the density map to the N atom of the Gln98 side chain and the O atoms of the Tyr136 and Tyr175 side chains are 3.3, 2.4 and 2.4 Å, respectively (Fig. 4*a*). These are sufficiently close to interact with some types of ligand. The cavity is covered by the loop region of Val79–Ala84 that connects the α4 and α5 helices (Figs. 1 and 4*b*). Thus, a conformational change is necessary for the ligand to approach the putative ligand-binding site.

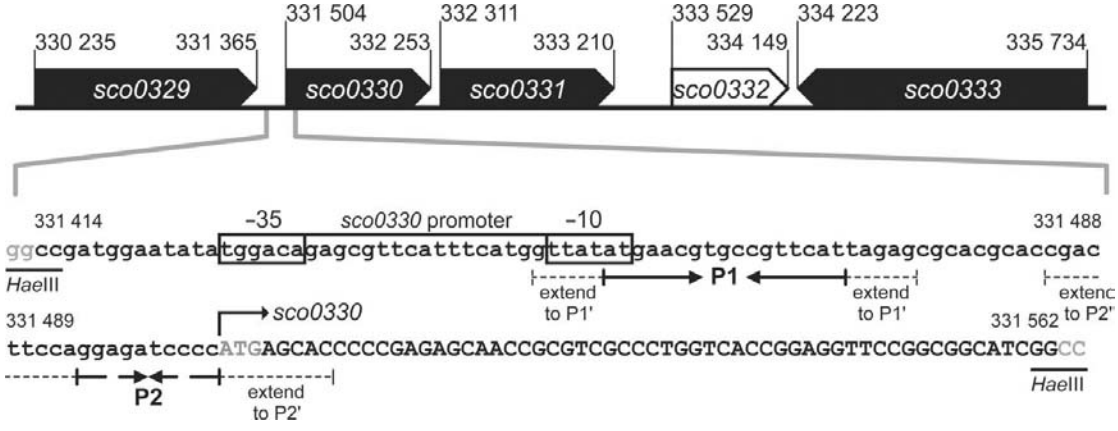

Figure 5
Top, genetic organization of *sco0329* to *sco0333* from *S. coelicolor*. Bottom, sequence presentation of the DNA fragment isolated by genomic SELEX analysis. Putative −35 and −10 promoter elements of *sco0330* are boxed. The imperfect palindromic sequences P1 and P2 are shown as arrows facing each other. The DNA fragments P1′ and P2′ extended for EMSA are shown as dotted lines. The bent arrow indicates the translational start site of the *sco0330* gene.

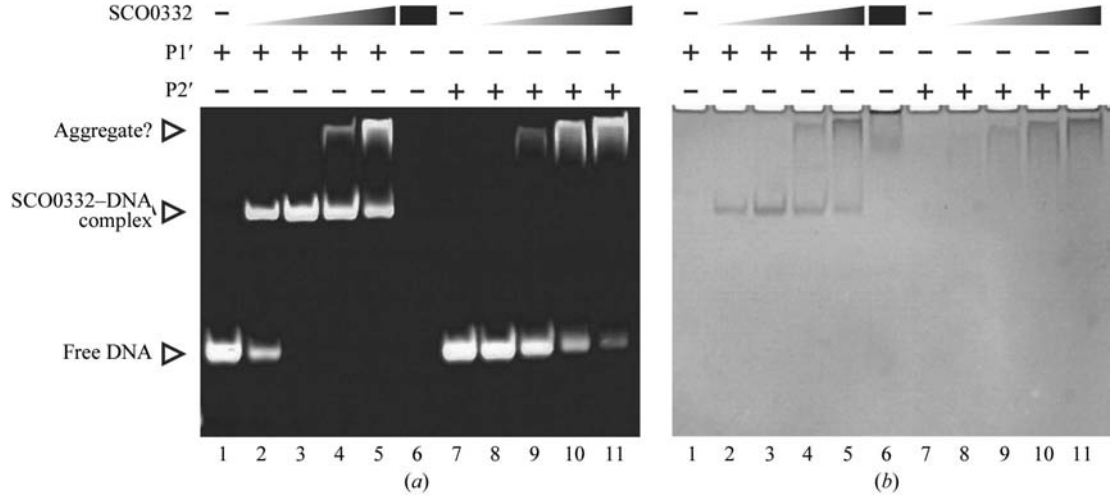

Figure 6
EMSA of SCO0332 and P1′ and P2′ DNA fragments. (*a*) Aliquots of 50 pmol annealed P1′ (lanes 1–5) or P2′ (lanes 7–11) were mixed with 0 mol (lanes 1 and 7), 50 pmol (lanes 2 and 8), 100 pmol (lanes 3 and 9), 150 pmol (lanes 4 and 10) and 200 pmol (lanes 5 and 11) of SCO0332 dimer. After electrophoresis, the polyacrylamide gels were stained with ethidium bromide. (*b*) The same gel as shown in (*a*) stained with CBB.

3.3. Binding region on the chromosomal DNA

After two cycles of genomic SELEX selection, 21 of the 23 isolated DNA fragments contained the same region of the chromosomal DNA: a sequence 149 bp in length extending from position 331 414 to position 331 562 of *S. coelicolor* chromosomal DNA, which contains putative −35 and −10 promoter elements and the first 59 nucleotides of the open reading frame of the *sco0330* gene (Fig. 5). In general, the members of the TetR-family of transcriptional regulators act as homodimers and consequently their consensus sequence shows perfect or imperfect palindrome patterns. Our structural analysis revealed that SCO0332 forms a homodimer and thus it is highly likely that SCO0332 recognizes a palindrome-like sequence. The chromosomal region determined by genomic SELEX analysis contains two imperfect palindromic sequences, indicated P1 and P2 in Fig. 5, which should be considered to be candidates for the recognition sequence of SCO0332. The imperfect palindromic sequences of P1 and P2 are 17 and 10 bp in length, respectively.

The binding affinities of SCO0332 to the two candidate sequences, P1 and P2, were tested by EMSA. The binding assay was carried out with 27 bp extended oligonucleotides P1′ or P2′ corresponding to the genomic sequence around P1 and P2, respectively. SCO0332 did not show binding affinity for any shorter fragments of 17 or 21 bp containing P1 or P2, respectively (data not shown). EMSA using the 27 bp fragments showed that SCO0332 and P1′ formed a protein–DNA complex, which was detected by both ethidium bromide and Coomassie Brilliant Blue (CBB) staining, but P2′ did not (Fig. 6). In the presence of a high concentration of SCO0332, some aggregation bands containing protein and DNA were observed, but these may not be related to the function of SCO0332. Moreover, addition of P2′ or poly(dI-dC)-(dI-dC) as a competitor did not inhibit formation of the SCO0332–P1′ complex (Fig. 7). Thus, the complex formation of SCO0332 was sequence-specific and we concluded that the P1′ region corresponding to the genomic sequence from position 331 449 to position 331 475 (or a shorter region) is the operator sequence of SCO0332.

4. Discussion

The amino-acid sequence around the DNA-binding motif is highly conserved in the TetR family (Fig. 2*c*). The most highly conserved residue is the lysine present at the amino end of the α4 helix next to the HTH DNA-binding motif. Surprisingly, this lysine residue is present in 77% of 2353 TetR-family regulators (Ramos *et al.*, 2005). Lys67 appeared at the same position in the structure of SCO0332 (Fig. 2*b*). Crystal structure analysis of the TetR–DNA or QacR–DNA complexes revealed this lysine residue to interact with the phosphate of the DNA backbone (Orth *et al.*, 2000; Schumacher *et al.*, 2002) and thus it seems that Lys67 in SCO0332 interacts with the operator DNA in the same way.

The next most highly conserved residue in the TetR family is the tyrosine residue (Tyr61 in SCO0332) at the middle of the α3 recognition helix, which is conserved in 74% of family members (Ramos *et al.*, 2005). This tyrosine residue interacts with the phosphate backbone *via* its hydroxyl group and recognizes the methyl group of a thymine base with the hydrophobic region of its side chain. These interactions are common in the two complex structures mentioned above. The functional units of TetR and QacR are homodimers and two conserved tyrosine residues from each monomer recognize thymine bases in the sense and antisense strands of DNA individually. Thus, the operator sequence is restricted to the

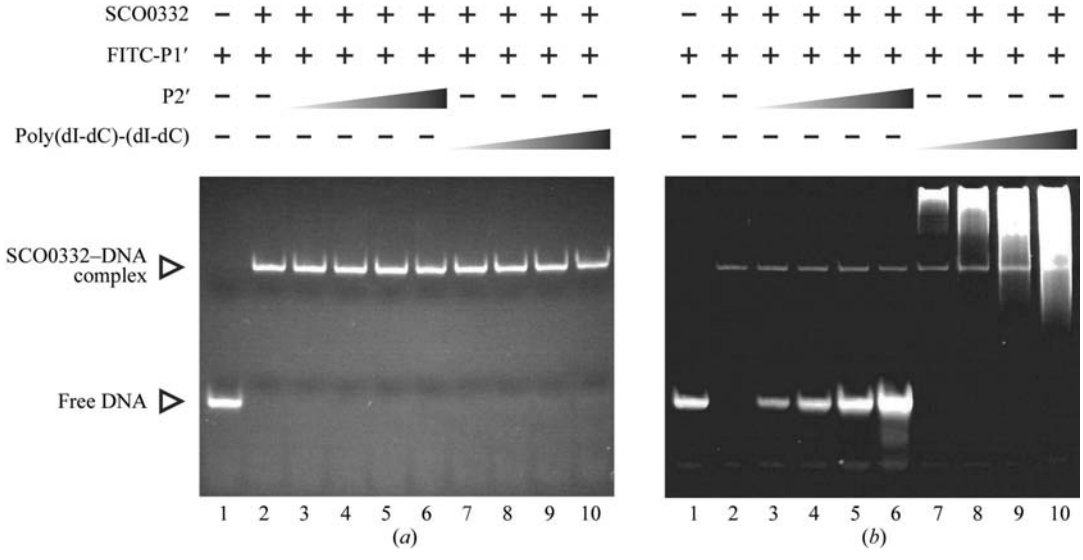

Figure 7
EMSA of SCO0332 and P1′. (*a*) Aliquots of 10 pmol FITC-labelled P1′ (all lanes) and 40 pmol SCO0332 dimer (lanes 2–10) were mixed followed by the addition of 5 pmol (lane 3), 10 pmol (lane 4), 20 pmol (lane 5) or 50 pmol (lane 6) nonlabelled P2′ or 0.5 µg (lane 7), 1 µg (lane 8), 2 µg (lane 9) or 5 µg (lane 10) poly(dI-dC)-(dI-dC). After electrophoresis, FITC-labelled P1′ was visualized using an LED light source. (*b*) The same gel as shown in (*a*) stained with ethidium bromide.

pattern 5′-A-X_n-T-3′, where n = 7 in TetR and n = 8 in QacR. This adenine–thymine separation is in accordance with the pitch of double-stranded DNA and two recognition helices bind individually to the major grooves. Structure analysis of SCO0332 revealed that the conserved residue Tyr61 is located in the α3 DNA-recognition helix and the side chain of Tyr61 is oriented towards the outside of the protein (Fig. 2*b*). These observations suggest that the operator sequence of SCO0332 should also show the same adenine–thymine pattern as the TetR or QacR operator. The P1′ sequence, which is the operator sequence of SCO0332 determined by genomic SELEX and EMSA, contains this pattern at three locations (Fig. 8). The structural evidence corresponds well to these functional analyses. In the case where one SCO0332 dimer binds to one operator DNA sequence as in TetR, Tyr61 of SCO0332 will recognize any one of three adenine–thymine separation pairs. It is also possible that two dimers bind to one operator sequence, as in the case of QacR–DNA interaction, because there are two A-X_8-T pairs with a 4 bp shift and they will appear at the opposite sides of the double-stranded DNA structure.

The DNA-binding affinity of TetR-family transcriptional regulators is controlled by ligand-induced conformational changes. The distances between the two DNA-recognition helices in the DNA-bound forms of TetR (PDB code 1qpi) and QacR (PDB code 1jt0) are 34.8 and 36.6 Å, respectively (Orth *et al.*, 2000; Schumacher & Brennan, 2002) and these values are consistent with the pitch of double-stranded DNA. The distance was measured between two C^{α} atoms of conserved tyrosine residues (Tyr42 in TetR and Tyr40 in QacR) on the α3 recognition helix from each monomer. Ligand binding increases this distance to 37.3 Å in TetR with tetracycline (PDB code 2trt) or 45.5 Å in QacR with pentamidine (PDB code 1rkw) (Hinrichs *et al.*, 1994; Murray *et al.*, 2004); they then leave from the operator element and the repressed target genes are transcribed. Considering the structural similarity, the DNA-binding affinity of SCO0332 is also controlled in the same way by the conformational change. In the structure of SCO0332, the distance measured between the C^{α} atoms of Tyr61 is 45.9 Å and so the present structure should be the ligand-bound (non-DNA-binding) form. This was supported by the observation of some electron density in the putative ligand-binding site. It is probable that the purified SCO0332 is in equilibrium between ligand-bound and nonbound forms. This is the reason why the DNA-binding affinity of SCO0332 was detected by genomic SELEX analysis and EMSA using the same protein solution as used in the crystallization experiment.

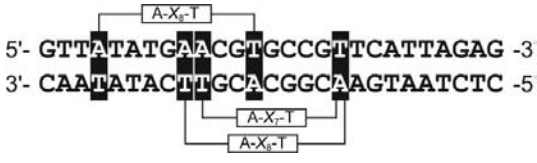

Figure 8
Nucleotide sequence of P1′. Conserved 5′-A-$X_{7,8}$-T-3′ patterns are indicated.

The operator region of SCO0332 (P1′) overlaps the putative −10 promoter sequence of the *sco0330* gene. Binding of SCO0332 inhibits the approach of RNA polymerase to the promoter site of *sco0330*. These results suggest that SCO0332 represses transcription of the *sco0330* gene in the same way as TetR or QacR. The *sco0330* gene is predicted to encode a short-chain dehydrogenase/reductase (SDR; Jornvall *et al.*, 1995; Kallberg *et al.*, 2002; Oppermann *et al.*, 2003). Members of the SDR family catalyze the NAD- or NADP-dependent oxidation/reduction reaction with a wide range of substrates, including sugars, alcohols, fatty acids and steroids. The *sco0331* gene, which is located between *sco0330* and *sco0332*, is also annotated as an SDR gene with unknown substrate. The interval between *sco0330* and *sco0331* is only 57 bp, which raises the hypothesis that the gene cluster consisting of *sco0330*, *sco0331* and *sco0332* may comprise a functional operon. SCO0330 and SCO0331 may catalyze continuous reactions and both may be regulated by the transcription repressor SCO0332.

In conclusion, the structural properties and sequence-specific DNA-binding ability of SCO0332 strongly suggest that this protein is a transcriptional repressor involved in regulation of expression of *sco0330* (and *sco0331*). Further analyses of the ligand of SCO0332 or of the functions of the SCO0330 and SCO0331 proteins will facilitate the comprehensive functional characterization of SCO0332.

This work was supported by a research grant from the National Project on Protein Structural and Functional Analyses from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Bentley, S. D. *et al.* (2002). *Nature (London)*, **417**, 141–147.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Donadio, S., Sosio, M. & Lancini, G. (2002). *Appl. Microbiol. Biotechnol.* **60**, 377–380.
Ellington, A. D. & Szostak, J. W. (1990). *Nature (London)*, **346**, 818–822.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Hinrichs, W., Kisker, C., Duvel, M., Muller, A., Tovar, K., Hillen, W. & Saenger, W. (1994). *Science*, **264**, 418–420.
Holm, L. & Sander, C. (1998). *Nucleic Acids Res.* **26**, 316–319.
Huffman, J. L. & Brennan, R. G. (2002). *Curr. Opin. Struct. Biol.* **12**, 98–106.
Itou, H., Okada, U., Suzuki, H., Yao, M., Wachi, M., Watanabe, N. & Tanaka, I. (2005). *J. Biol. Chem.* **280**, 38711–38719.
Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.
Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. (2002). *Protein Sci.* **11**, 636–641.
Kisker, C., Hinrichs, W., Tovar, K., Hillen, W. & Saenger, W. (1995). *J. Mol. Biol.* **247**, 260–280.
Kitago, Y., Watanabe, N. & Tanaka, I. (2005). *Acta Cryst.* D**61**, 1013–1021.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Mitani, Y., Meng, X., Kamagata, Y. & Tamura, T. (2005). *J. Bacteriol.* **187**, 2582–2591.
Murray, D. S., Schumacher, M. A. & Brennan, R. G. (2004). *J. Biol. Chem.* **279**, 14365–14371.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nakashima, N. & Tamura, T. (2004*a*). *Appl. Environ. Microbiol.* **70**, 5557–5568.
Nakashima, N. & Tamura, T. (2004*b*). *Biotechnol. Bioeng.* **86**, 136–148.
Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jornvall, H. (2003). *Chem. Biol. Interact.* **143–144**, 247–253.
Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. (2000). *Nature Struct. Biol.* **7**, 215–219.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Ramos, J. L., Martinez-Bueno, M., Molina-Henares, A. J., Teran, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R. & Tobes, R. (2005). *Microbiol. Mol. Biol. Rev.* **69**, 326–356.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Schumacher, M. A. & Brennan, R. G. (2002). *Mol. Microbiol.* **45**, 885–893.
Schumacher, M. A., Miller, M. C. & Brennan, R. G. (2004). *EMBO J.* **23**, 2923–2930.
Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A. & Brennan, R. G. (2001). *Science*, **294**, 2158–2163.
Schumacher, M. A., Miller, M. C., Grkovic, S., Brown, M. H., Skurray, R. A. & Brennan, R. G. (2002). *EMBO J.* **21**, 1210–1218.
Shimada, T., Fujita, N., Maeda, M. & Ishihama, A. (2005). *Genes Cells*, **10**, 907–918.
Terwilliger, T. C. (1999). *Acta Cryst.* D**55**, 1863–1871.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Tuerk, C. & Gold, L. (1990). *Science*, **249**, 505–510.
Watanabe, N. (2006). *Acta Cryst.* D**62**, 891–896.
Wintjens, R. & Rooman, M. (1996). *J. Mol. Biol.* **262**, 294–313.
Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.