ST1710–DNA complex crystal structure reveals the DNA binding mechanism of the MarR family of regulators

Thirumananseri Kumarevel1,*, Tomoyuki Tanaka1, Takashi Umehara2 and Shigeyuki Yokoyama2,3

1RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, 2Systems and Structural Biology Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045 and 3Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received April 24, 2009; Revised May 21, 2009; Accepted May 21, 2009

ABSTRACT

ST1710, a member of the multiple antibiotic resistance regulator (MarR) family of regulatory proteins in bacteria and archaea, plays important roles in development of antibiotic resistance, a global health problem. Here, we present the crystal structure of ST1710 from Sulfolobus tokodaii strain 7 complexed with salicylate, a well-known inhibitor of MarR proteins and the ST1710 complex with its promoter DNA, refined to 1.8 and 2.1 Å resolutions, respectively. The ST1710–DNA complex shares the topology of apo-ST1710 and MarR proteins, with each subunit containing a winged helix-turn-helix (wHtH) DNA binding motif. Significantly large conformational changes occurred upon DNA binding and in each of the dimeric monomers in the asymmetric unit of the ST1710–DNA complex. Conserved wHtH loop residues interacting with the bound DNA and mutagenic analysis indicated that R89, R90 and K91 were important for DNA recognition. Significantly, the bound DNA exhibited a new binding mechanism.

INTRODUCTION

Microbial antibiotic resistance is a result of either inactivation or reduced accumulation of antibiotics within an organism. Proteins belonging to the multiple antibiotic resistance regulators (MarR) family reportedly regulate the expression of proteins conferring resistance to multiple antibiotics, organic solvents, household disinfectants, oxidative stress agents and pathogenic factors (1–3). For example, in the absence of the appropriate stimulus, Escherichia coli MarR proteins negatively regulate the marRAB operon, and repression of this operon is alleviated by exposure to a variety of phenolic compounds, most notably sodium salicylate (1). Similarly, MexR negatively regulates an operon in Pseudomonas aeruginosa that, when expressed, encodes a tri-partite multi-drug efflux system that results in an increased resistance to multiple antibiotics, including tetracycline, β-lactams, chloramphenicol, novobiocin, trimethoprim, sulfonamides and fluoroquinolones (4,5). Some members of the MarR family of DNA-binding proteins, for example hypothetical uricase regulator (HucR) and organic hydroperoxide resistance regulator (OhrR), mediate a cellular response to reactive oxidative stress (ROS) (6,7). The Deinococcus radiodurans HucR was shown to repress its own expression as well as that of an uricase. This repression is alleviated both in vivo and in vitro upon binding uric acid, the substrate for uricase. As uric acid is a potent scavenger of reactive oxygen species, and D. radiodurans is known for its remarkable resistance to DNA-damaging agents, these observations indicate a novel oxidative stress response mechanism (8–10). Similar to HucR, the OhrR protein of Bacillus subtilis also mediates a response to oxidative stress; however, for OhrR, it is oxidation of a lone cysteine residue by organic hydroperoxides that abrogates DNA binding (11,12).

We have reported two different crystal forms of ST1710 (13) and others (14). The structure showed the winged helix-turn-helix (wHtH) motif at the DNA binding site that obviously belonged to the MarR family of proteins. The crystal structures of proteins in the MarR family have also been determined from a number of organisms including MarR from E. coli (15), MexR from P. aeruginosa (16), SarR from Staphylococcus aureus (17), SlyA-like protein from Enterococcus faecalis (18), OhrR from B. subtilis (19), HucR from D. radiodurans (20) and MTH313 from Methanobacterium thermoautotrophicum (21). Sequence comparisons of these proteins with ST1710 showed less...
The members of the MarR family of regulatory proteins recognize double stranded DNA by their wHtH motifs (15–19). Footprinting experiments revealed that MarR binds as a dimer at two different, but similar, sites in marO, protecting 21 bp of DNA on both strands at a single site without bending its target (29,30). One of the MarR families of proteins, the OhrR protein complexed with the ohrA operator with a 29-bp duplex was solved, which revealed the interactions between them. The protein–DNA contact region included the major groove of the –10 element, and indicated that OhrR, and probably MarR and MexR as well, repress transcription by blocking the access of RNA polymerase to this promoter element (19). In addition, the mutational analysis of the RNA polymerase binding site, the –10 element of the OhrR-, MarR- and MexR-regulated promoters, revealed the loss of DNA binding ability by ~10-fold when this region was altered (11,16,30,31).

On the basis of the sequence of the ohrA promoter, we previously identified a putative promoter for ST1710 and showed binding ability by gel-mobility shift assays (13). This promoter is located immediately upstream of the first ATG of the ST1710 gene and downstream of the STS1709 gene. To understand the importance of MarR family members in antibiotic resistance and other biological processes, here, we solved the ST1710 in three different forms: (i) apo-form (native), (ii) complexed with its inhibitor, sodium salicylate (salicylate complex) and (iii) complexed with its promoter DNA. A slight conformational change on the side chains of protein residues’ was observed when bound to the salicylate ligand, compared to the apo-form. The DNA bound to the wHtH motif of one monomer on the dimeric ST1710, and specifically interacted with the residues R84, R89, R90 and K91. A significantly large conformational change was observed between the monomers of the dimeric protein bound to the DNA, and also with the apo/salicylate complex structures. Significantly, a distinct mode of DNA binding was observed with the bound DNA passing over the protein molecule rather than passing through the 2-fold related axis of the molecule as previously observed (19).

**MATERIALS AND METHODS**

**Cloning, expression and purification of ST1710**

The gene encoding the MarR family regulator protein (ST1710) from *Sulfolobus tokodaii* (*S. tokodaii*) strain 7 was amplified from genomic DNA by PCR, using the primers 5'–ggtaatCATATGTAGAATGATAAGAAA CAGAATAC-3' and 5'–ggatatGGATCCTATATTACTGA CTATTTCTCAATTTTTTCC-3'. The PCR fragment was digested with NdeI and BamHI and cloned into the pET21a(+) expression vector. The plasmid was transformed into the *E. coli* BL21-CodonPlus (DE3)—RIL-X strain (Stratagene), and the selenomethionine-containing ST1710 proteins were over-expressed and purified as described in our earlier studies (13).

**Crystallization and data collection**

Native crystals of ST1710 were produced to medium size (0.1 × 0.1 × 0.1 mm) within the period of a week at 20°C by the sitting drop vapor diffusion method (32), by adding 1 μl of protein solution to 1 μl of well solution, containing 18% PEG 8 K, 0.2 M calcium acetate, and 0.1 M sodium cacodylate, pH 6.5. For the ST1710–salicylate complex, crystals were soaked in the mother liquor containing 0.2 M of sodium salicylate for 3 min. Twenty percent ethylene glycol was used as a cryo-protectant and the complete data set was collected for the native and salicylate complex with the in-house R-axis VII system (RIGAGU MSC). These crystals belonged to the tetragonal space group, P41212, and the processed data statistics are given in Table 1. The 30 bp synthetic oligonucleotide containing the putative promoter sequence (5’–AATAATGTCATT GTTAAACAATAGCAATAAT-3’) and its complementary oligonucleotide (5’–ATTTTGTATTGTAAAACATG ACATTATT-3’) were annealed completely to form the DNA-duplex and complexed with the ST1710 protein at an equimolar ratio. Initial crystals of ST1710–DNA complex were produced at 20°C by the sitting drop vapor diffusion method (32), by adding 1 μl of protein–DNA complex solution to 1 μl of well solution, containing 30% (v/v) 2-Methyl-2,4-pentanediol (MPD), 0.02 M calcium chloride dehydrate and 0.1 M sodium acetate trihydrate, pH 3.8. Preliminary tiny crystals that were seeded in the equilibrated drops of the same condition grew up to a maximum dimension of 0.3 mm within a 2-week period. Complete Multiple Anomalous Dispersion (MAD) data sets were obtained at 100 K using a Jupiter210 CCD detector (RIGAGU MSC) on the RIKEN structural genomics beamline 1 (BL26B1) at SPring-8, Hyogo, Japan. These crystals belonged to the centered orthorhombic space group C2221, with cell dimensions a = 94.44, b = 106.73 and c = 82.26 Å. The native and Se-edge MAD data sets were processed up to 2.10 Å using the HKL2000 suite (33) (Table 1).
Structure determination and refinement

The native and salicylate complex of ST1710 structures were determined by the molecular replacement method, using our previous ST1710 structure (PDB code, 2eb7), as a search model. The solution was found by automated-MOLREP, within the CCP4 program suite, and the refinement was carried out using CNS (34). The protein model was built using the programs Quanta (35) and Coot (36). The native and salicylate complexes were refined to resolutions 1.80 and 2.0 Å, respectively, (Table 1). Since the molecular replacement method was not successful for phasing the ST1710–DNA complex, we collected and processed the Se-MAD data sets. The ST1710–DNA complex structure was successfully phased by the MAD method with the three different wavelength data sets collected at the Se-edge, using the program SOLVE (37). Solvent flattening and initial model building were performed by RESOLVE (37). Improvement of the partial model derived from RESOLVE was performed with the ARP/wARP program (38). We observed unambiguous density for the DNA bases and built the DNA-protein model using the programs Quanta (35) and Coot (36) and Quanta (34). The final model with 285 protein residues and 23 nucleic acid bases, except for 4 and 3 residues in the N-terminal of A and B chains, respectively, was refined to a crystallographic R-factor of 0.237 (R_free = 0.287) at 2.10 Å resolution, using synchrotron radiation X-ray data collected at cryo temperature (see Table 1). Figures were prepared with the program Pymol (39). The coordinates and structure factors for the native, salicylate, and ST1710–DNA complex have been deposited in the Protein Data Bank, under the accession codes 3GEZ, 3GF2 and 3GFI, respectively.

Differential scanning calorimetric (DSC) analysis of salicylate binding to ST1710

DSC experiments were carried out using a VP-capillary DSC platform (Microcal, USA). For the DSC measurements, the protein concentration was fixed at 0.5 mg/ml in 20 mM Tris–HCl buffer (pH 8.0) containing 150 mM NaCl. Dialyzed protein sample and the sodium salicylate were filtered through a 0.22 μm pore size membrane and complexed with different concentrations of sodium salicylate (0–300 mM) and were loaded on to the capillary system. The scan rate was 90°C/hr for all experiments. We calculated the binding constant (K_d) of sodium salicylate using the DSC curves analysis using the Origin software (Microcal, USA) and the Graphpad Prism 2.0, a non-linear curve-fitting algorithm (GraphPad software).

Site-directed mutagenesis of ST1710 and gel-mobility shift assays

Initially, the ST1710 plasmid was prepared with a Qiagen miniprep kit. The Quickchange site-directed mutagenesis kit (Stratagene) was used to create the DNA-binding site mutants (R89A, R90A and K91A), and the resultant plasmids were transformed into JM109 cells. N-terminal sequencing was carried out for all the mutants, which were expressed and purified in a manner similar to the native protein (13). To evaluate the protein–DNA interactions in solution, a gel-mobility shift assay was used. The purified ST1710 was incubated in binding buffer (10 mM Tris–HCl, pH 8.0, 200 mM NaCl, 20 mM MgCl_2 and 5 mM β-ME) at predetermined concentrations, and the 30-mer DNA (100 nM) was added. The reaction mixture was incubated for 20 min at room temperature and mixed with 1 μl of 50% glycerol before loading onto the gel. The free DNA and ST1710–DNA complexes were resolved on a 10% polyacrylamide gel (running buffer, 1X TBE, constant voltage, 200 V; temperature, 4°C). Inhibition of ST1710 was analyzed in the presence of increasing concentrations of sodium salicylate by gel-mobility shift analysis. The free and complexed nucleic acids were stained by fluorescent SYBR Green (EMSA Kit, Invitrogen) and the bands visualized with a UV transilluminator (LAS-3000, FUJIFILM, Japan).

RESULTS AND DISCUSSION

Biophysical analysis of salicylate binding to ST1710 and inhibition assays

Previous in vitro and in vivo analyses of MarR family of proteins suggested that salicylate is a broad inhibitor for MarR activity at the millimolar concentration levels (1). To investigate whether salicylate binds to ST1710, we used a differential scanning calorimetric method for the binding analysis. Various concentrations of sodium salicylate (0–300 mM) were mixed with the protein and the heat capacities (C_p) measured with the scan rate of 90°C/hr (Figure 1A). The binding constant (K_d) was 20 ± 4.9 mM for sodium salicylate as calculated by the GraphPad Prism software, while the Origin program produced similar values. This analysis clearly suggests that salicylate binds to ST1710, and the binding constants were comparable with other members of MarR family (21).

Next, to visualize the protein–DNA inhibition by sodium salicylate, we used the gel-mobility shift assays. While the ST1710–DNA complex formed more compact than free DNA alone, the addition of salicylate inhibited the DNA–protein complex. At concentrations above 100 mM, the most of the DNA in the salicylate treated complex released from ST1710. This analysis clearly demonstrates that the ST1710–DNA complex formation was inhibited with increasing concentrations of salicylate (Figure 1B). Taken together, the data indicates that salicylate bound to the ST1710 in solution and inhibited the protein–DNA complex when it exceeded the intracellular concentration levels.

Structure of ST1710 complexed with salicylate

Our DSC and gel-mobility shift analyses confirmed that salicylate bound to the ST1710 and inhibited the protein–DNA complex formation. To see how salicylate binds to ST1710, we determined the ST1710–salicylate complex at a resolution of 1.80 Å, and refined to a final R value of 23.3% and an R_free value of 26.5%. The overall structure
of ST1710 within the complex was similar to our recently reported native structure, showing that it belongs to the α/β family of proteins and resembles those of the MarR family of proteins. It consists of six α-helices and two β-strands, arranged in the order of α1-α2-α3-α4-β1-β2-α5-α6 in the primary structure (Figure 2A). The asymmetric subunit contains one molecule, with overall dimensions of 60 x 39 x 26 Å. Two monomers are related by a crystallographic 2-fold axis to form the dimer (Supplementary Figure S1), and this is consistent with our gel-filtration analysis (13) as well as with studies of other MarR family proteins (15–19). The N- and C-terminal equivalent residues of each monomer, located at the α1, α5 and α6 helices, are closely intertwined and form a dimerization domain, which is stabilized by hydrophobic and hydrogen bonding interactions between the residues located within these regions. Apart from the dimerization domain, as observed in many DNA binding transcriptional regulators, the residues located at the α2-α3-α4-β1-β2 formed a wHtH DNA binding motif. In the dimer, the distances between the recognition helix (α4) to the recognition helix and the loop to loop (connecting the β-strands) of the wHtH domains are ~30 and ~70 Å, respectively.

The fine quality of the electron density map allowed us to identify unambiguously the specific salicylate binding site in the complex structure. The bound salicylate is located at the interface between the helical dimerization and wHtH DNA-binding domains (Figures 2A and B). The large portion of the DNA-binding and dimerization domains formed a salicylate-binding pocket. The salicylate ligand has many interactions with the protein residues (Figure 2B). The O2’ of salicylate is bonded to the side chain oxygens of Y37 and Y111; in addition, side chain oxygen of Y37 is also bonded to the O2’ of the ligand molecule. The ligand oxygen O1’ is hydrogen bonded to the side chain amino group (NH2) of residue R20, while the O2 of ligand molecule is hydrogen bonded to the side chain nitrogen of K17. The latter two interactions are from the symmetrically related molecule. Thus, the bound salicylate has many interactions. All of the residues which interact with the ligand are highly conserved among the closely related species (Figure 1E).

Next, to observe if any conformational change occurred in the salicylate liganded complex when compared to the native structure, we collected the new native data set from the crystal grown under the same conditions, solved at 2.0 Å and refined to a final R value of 21.1% and an Rfree value of 25.2%. The overall conformation of the complex is very similar to the native structure, with an rmsd of 0.11 Å for superposition of 141 Cα atoms (Figure 2C). However, a minor conformational change was observed in the side chain orientations of the ligand interacting residues and the DNA-binding wHtH motifs. We believe that allosteric changes might occur at the molecular level in solutions in the presence of inhibiting ligand, salicylate, which is not seen in the crystal structure. A divalent metal ion, Ca2+, was observed in the native structure, and it interacted with the C-terminal Q146 and
E98 residues, which is closer to the DNA-binding loop of the symmetrically related molecule (Figure 2C). It is interesting to compare the salicylate complexes within the MarR family of regulators. There are two complex structures available; MarR from *E. coli* and the more recently solved MTH313 from *M. thermoautotrophicum*. Although these proteins displayed sequence similarity ~26% to ST1710, the superposition of the...
ST1710–salicylate complex with the two others revealed that the overall topology is similar, with an rmsd of 1.90 Å for 119 Cα atoms for *E. coli* MarR and *M. thermoautotrophicum* MTH313, respectively (Figure 2D). However, *E. coli* MarR was crystallized with two molecules of salicylate per dimer, both of which were highly exposed to the solvent (15), and these salicylate binding sites are not comparable to that with ST1710. The salicylate ligands in MarR hydrogen bonded to some of the amino acid residues (A70, T72, R77 and R86); however, the physiological relevance of either salicylate binding site could not be determined because the ligands were involved in interactions with protein molecules within the crystal which may stabilize the crystal lattice. On the other hand, the mode of salicylate binding between ST1710 and MTH313 is comparable, and the ligand adjusted up ~2 Å and ~3 Å towards the z5 helix upon binding. In contrast to the salicylate binding in ST1710, two direct and one water mediated protein residues were in contact with the ligand in MTH313. A comparison between the apo and salicylate complex of MTH313 revealed a significant asymmetrical conformational change that is mediated by the binding of sodium salicylate to two distinct locations in the dimer (21), whereas we did not observe such changes in the case of ST1710. The available in vivo and in vitro analyses suggest that the MarR family of regulators inhibits the activity in the presence of salicylate. Since we could see only the fixed side-chain orientation of the protein residues contacting the bound salicylate ligand, we believe possible dynamic or allosteric changes occurring upon salicylate binding to ST1710 to inhibit its activity would not be captured by crystallization.

**Overall structure of the ST1710–DNA operator complex**

In our earlier studies, we identified the putative promoter for ST1710, which is located upstream of the *st1710* gene (13). The gel-mobility shift analysis suggested that protein–DNA interactions were competitive and concentration dependent (Supplementary Figure S2). Here, we complexed the ST1710 and 30-mer duplexed DNA (Figure 3A), crystallized, and collected the data set up to a resolution of 2.1 Å, under the space group C2221. Initially, we were unsuccessful in solving the structure by molecular replacement using the coordinates of our native molecule of the full-length promoter DNA. The full-length (30-mer) of the DNA-duplex in the crystals were also confirmed by dissolving the complex crystals (ST1710–DNA) in water, after washing them a few times in reservoir solution, and then analyzed DNA-duplex, using agarose gel electrophoresis. These analyses indicated that the DNA-duplex present in the crystal was, indeed, the full-length DNA (Supplementary Figure S3). To clarify further the DNA binding to ST1710, we prepared an additional DNA duplex with exactly 2-fold related sequence based on the observed fragments in the present crystal (5′-TTGCTATT GTTAAACATGGAAAT-3′) (Figure 3A, bottom sequence), crystallized, and solved the structure. The overall structure of this new complex resembles the present one, and it recognizes the DNA fragment irrespective of the sequence heterogeneity (our unpublished data). As shown in Figure 3E, one full-length duplex-DNA was interacted with four molecules of the ST1710, which was coming from four different asu of the dimeric molecules. Thus, the analysis of the ST1710–DNA complex in the present study suggests that the protein:DNA ratio is 2:1. Interestingly, the bound DNA passes over the wHtH motif, making contacts only at the loop regions.

**Interactions between the ST1710 and promoter DNA**

As explained above, Figure 4A represents the full length duplex-DNA bound to four monomers of the symmetry-related dimeric molecules. Although we used only a 30-mer duplexed DNA for our crystallization studies, we could see the duplexed-DNA consisting of T5 to A27 and T5′ to A27′ of the bases bound to the protein (Figures 4A and B). The 4 and 3 bases at the 5′- and 3′-end, respectively, were highly disordered in both of the DNA-strands and hence not modeled. The bound DNA adapted a B-form right handed structure, passing over the protein molecule by only contacting at the wHtH loop regions. The protein–DNA interactions seen in the asu of the complex (Figures 3B and C) were essentially same in all of the four symmetry-related molecules. Interestingly, as observed in the OhrR–*ohrA* operator complex, the –10 region (TAACAAT) of the promoter DNA (15–21) was recognized by the wHtH domains (Figures 4B–F). Of the bound 54 nucleotides, only 22 nucleotides make 36 contacts with six protein residues (Figures 4B–F). The side chain oxygen of S65 was bonded to the O2′ of Thy5′. Interestingly, the side chain (NH1) of residue R84 formed water-mediated hydrogen bonds to the N3 of bases G13′ and Ade17. In addition, side chain (NH1) of R89 hydrogen bonded to the backbone phosphate oxygen (O3′) of Thy14′. The residue R90 hydrogen bonded to the O3′ and O2′ of base Cyt15 and the same residue made two salt-bridge contacts with D88. This salt bridge may assist in fixing the conformation of residue R90 in order to make contact with the nucleic acid base, Cyt15. Besides, the side chain atom (CD) bonded to the bases of Gua13N2 and Thy14O2. The side chain of K91 interacted with backbone phosphate of Ade19 and Ile91 to C5 of Ade29. Thus, the following residues S65, R84, E88, R89, R90, K91 and I92 interacted with the bound promoter DNA. To evaluate the protein–DNA interactions at the loop region,
**Figure 3.** Structure of ST1710–DNA complex. (A) DNA sequence used for crystallization (30-mer and 26-mer duplex). The blue background shows the 2-fold related sequence. The red boxed residues were observed in the crystal structure. (B) The final 2Fo-Fc omit electron density map with the nucleic acids contoured at 1σ level is shown in blue mesh. The template strand and its complementary strands are shown in blue and yellow stick models. (C) Stereo view of the asymmetric unit of ST1710–DNA complex. The secondary structural assignments, N- and C-terminal ends are labeled in one of the dimeric monomers. The bound nucleic acids are shown as stick representations. (D) Stereo view of the two dimers of the asymmetric unit related by 2-fold axis to form the tetramer for recognition of one promoter DNA. (E) Part of the crystal packing is shown (stereo view). Each asymmetric unit of the complex is shown in a unique color. The 5’- and 3’- ends of each nucleotide chain is labeled.
Figure 4. ST1710–promoter DNA interactions. (A) A ribbon diagram of ST1710–DNA complex colored as in Figure 4E. The full-length promoter DNA is shown in stick model. The T5-A27 and T5'-A27' strands are in blue and yellow, respectively. (B) Schematic representation showing all ST1710–DNA contacts. The bases are shown in rectangles. The protein–DNA contacts in each of the ST1710 monomer are represented by the same color in (A). (C–F) The close-up view of the critical protein–DNA interactions in the complex is shown in A–D. The residues of nucleic acids and protein are shown in stick models. The hydrogen bonds are shown in broken lines.
we prepared three mutant proteins (R89A, R90A and K01R) and analyzed the binding ability by gel-mobility shift assays. All three mutants failed to bind to DNA, suggesting that these three residues are important for protein–DNA interactions (Supplementary Figure S4). We also crystallized all three mutant proteins under the native protein conditions and solved their structures by molecular replacement method as explained previously (Table 1). All three mutant structures resembled the native ST1710, except very little changes were observed in the wHtH loop regions (Supplementary Figure S5). Furthermore, DNA-binding residues in ST1710 were highly conserved among the closely related proteins and OhrR (Figure 2E). The winged loop region connecting the strands β1 and β2 apparently plays a major role in modulating their conformation for binding to the DNA molecule, and this mode of recognition is anticipated for the proteins closely related to ST1710 as well as the family of MarR regulators. We observed Ca$^{2+}$ ions in all of the mutant and native structures, but not in the salicylate and DNA-complex structures. Superimposition of the native, salicylate and DNA-complex structures suggested that the C-terminal helix (z6) in the DNA complex deviated greatly from the metal ion binding site and in the salicylate complex slight changes in side chain orientations were observed (Supplementary Figure S6). However, the functions of this metal ion observed in the native and mutant structures of ST1710 will need to be further investigated.

Conformational changes

In our recent report, we noticed a small change only at the loop region connecting strands β1 and β2 in the protein conformers crystallized in two different space groups and the overall structure was identical with an rmsd of 0.519 Å for 141 Cα atoms (13). In a similar way, when we compared the present ST1710–salicylate complex and native structure crystallized under the same conditions, a similar structural conformation was revealed with an rmsd of 0.11 Å for superposition of 141 Cα atoms. Additionally, the subunits in the dimer were identical. In contrast, the superimposition of the ST1710–DNA complex subunits (A and B chains) on one another revealed a large local conformational change all along the structure, excluding the helices z1 and z5 with an rmsd of 2.85 Å for 142 Cα atoms; however, the overall structural topology was similar (Figure 5A). A displacement of ~3.5–5.5 Å was seen all along the winged HtH motif region and the C-terminal helix showed the displacement of around 2–3 Å. The winged HtH motif of A-chain where the DNA is recognized was elevated up compared to the B-chain, while the C-terminal helix z6 was shifted down.

As seen in Figure 5B, significant changes were also observed between the subunits of the DNA-complex and the native and salicylate complex. Superposition of the native and salicylate complexes on to the A-chain showed greater differences than on the B-chain: an rmsd of 2.6 and 2.9 Å with the A-chain; and 0.9 and 1.0 Å with the B-chain, for the native structure and salicylate complex, respectively. One of the wHtH motifs may have been relocated to establish contact with the DNA, and by this mode of recognition, the wHtH loop regions were stabilized. The temperature factor for this loop region in the B-chain increased significantly, although it formed a dimer related by a non-crystallographic 2-fold axis, observed in the asymmetric unit of the cell (Supplementary Figure S7).

It is noteworthy to mention that the temperature factor was higher for the DNA-complex B-chain when not only compared to the DNA-complex A-chain, but also to the native, salicylate complex and all three of the mutant structures. Apparently, when protein binds to the DNA, the wHtH loop becomes stabilized and the temperature factor of that region is lowered; this observation was similar to that of the salicylate-complex which was solved at the highest resolution. It is also interesting to note that in the dimer, the distances between the loop to loop (connecting the β-strands) of the wHtH domains were reduced by ~10 Å for the ST1710–DNA complex, compared to the native and salicylate complexes.

Mode of nucleic-acid binding

There are several MarR families of regulators reported to date from different organisms including E. coli (15), P. aeruginosa (16), S. aureus (17), E. faecalis (18), B. subtilis (19), D. radiodurans (20) and M. thermotolerans (21). However, only the structure of OhrR from B. subtilis is available with its promoter and revealed their interactions. It is intriguing to compare our ST1710–DNA complex with OhrR-ohrA operator binding to clarify the binding mechanism. The superimposition of the ST1710–DNA complex on the OhrR-ohrA operator revealed large conformational changes, with an rmsd of 4.4 Å for 215 Cα atoms when compared to ST1710 before binding to its cognate promoter (rmsd 2.80 for 238 Cα atoms) (Figures 6A and B). Compared to the ST1710–DNA complex with its native structure and salicylate complex in the previous section, we found a similar change in the protein and unexpectedly found conformational changes in the mode of DNA recognition also. In the OhrR-ohrA complex, the distance between the subunits loop-to-loop was around 67 Å and the recognition helix (z4)–helix was about 20 Å (Figure 6A), although the wings of the subunits translated about 16 Å compared to the structure of reduced OhrR (19). The bound 2-fold related promoter sequence recognized the protein wHtH loop-to-loop, with substantial widening and deepening of the major groove that resulted from insertion of the recognition helix (z4) of the HtH motif. In contrast to this mode of binding, the bound DNA in ST1710 passed over the wHtH motif without deepening the structure through the 2-fold axis, even though the protein contacting residues are highly conserved between these two proteins and among the MarR family of regulators (Figures 2E, 4A and 6B). This unexpected mode of DNA-binding originated due to the translocation of one of the subunits around 13 Å towards the 2-fold axis, reducing the distance between the recognition helix of the subunits to 13 Å (Figures 6A and B). Thus, the DNA passing through the 2-fold axis deepening the recognition helices as observed in OhrR-ohrA operator complex would be impossible for that of ST1710.
<table>
<thead>
<tr>
<th>Data collection</th>
<th>ST1710-salicylate</th>
<th>ST1710-native</th>
<th>ST1710–DNA complex</th>
<th>R89A</th>
<th>R90A</th>
<th>K91A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>1.54178</td>
<td>1.54178</td>
<td>0.97896</td>
<td>1.00</td>
<td>1.54178</td>
<td>1.54178</td>
</tr>
<tr>
<td>Space group</td>
<td>P4,2,2</td>
<td>P4,2,2</td>
<td>C2,2</td>
<td>P4,2,2</td>
<td>P4,2,2</td>
<td>P4,2,2</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
<td>a = b = 46.14, c = 137.64</td>
<td>a = b = 46.04, c = 138.45</td>
<td>a = b = 94.44, b = 106.73, c = 82.26</td>
<td>a = b = 45.86, c = 138.17</td>
<td>a = b = 45.90, c = 137.42</td>
<td>a = b = 45.96, c = 138.29</td>
</tr>
<tr>
<td>No. of molecules/asu</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Resolution Range (Å)</td>
<td>40–1.80 (2.07–2.0)</td>
<td>40–2.0 (2.07–2.0)</td>
<td>50–2.10 (2.18–2.20)</td>
<td>40–2.20 (2.28–2.20)</td>
<td>40–1.90 (1.97–1.90)</td>
<td>50–2.10 (2.18–2.20)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>13 381</td>
<td>10 380</td>
<td>24 204</td>
<td>24 231</td>
<td>8143</td>
<td>12 085</td>
</tr>
<tr>
<td>Redundancy</td>
<td>13.4 (7.6)</td>
<td>18.5 (19.1)</td>
<td>9.6 (9.8)</td>
<td>12.5 (13.0)</td>
<td>13.0 (12.6)</td>
<td>12.2 (12.9)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.3 (99.8)</td>
<td>96.0 (94.4)</td>
<td>99.1 (98.7)</td>
<td>99.0 (98.6)</td>
<td>100 (100)</td>
<td>97.7 (95.1)</td>
</tr>
<tr>
<td>Rmerge $^b$</td>
<td>0.100 (0.264)</td>
<td>0.104 (0.321)</td>
<td>0.080 (0.294)</td>
<td>0.072 (0.284)</td>
<td>0.067 (0.281)</td>
<td>0.091 (0.389)</td>
</tr>
<tr>
<td>Refinement statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution Range (Å)</td>
<td>20–1.80</td>
<td>20–2.0</td>
<td>24 204</td>
<td>24 231</td>
<td>8143</td>
<td>12 085</td>
</tr>
<tr>
<td>Reflections used in the refinement</td>
<td>13 138</td>
<td>10 244</td>
<td>24 202</td>
<td>7877</td>
<td>11 906</td>
<td>8997</td>
</tr>
<tr>
<td>Total no. of reflections used for working set</td>
<td>12 246</td>
<td>9423</td>
<td>24 276</td>
<td>7292</td>
<td>10 943</td>
<td>8258</td>
</tr>
<tr>
<td>$R_{work}^c$</td>
<td>0.233</td>
<td>0.211</td>
<td>0.237</td>
<td>0.201</td>
<td>0.205</td>
<td>0.218</td>
</tr>
<tr>
<td>Total no. of reflections used for Rfree $^d$</td>
<td>892</td>
<td>821</td>
<td>1726</td>
<td>585</td>
<td>963</td>
<td>739</td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>1142</td>
<td>1142</td>
<td>2305</td>
<td>1136</td>
<td>1136</td>
<td>1136</td>
</tr>
<tr>
<td>No. of nucleic acid atoms</td>
<td>–</td>
<td>–</td>
<td>466</td>
<td>1136</td>
<td>1136</td>
<td>1136</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>152</td>
<td>156</td>
<td>154</td>
<td>125</td>
<td>160</td>
<td>127</td>
</tr>
<tr>
<td>No. ligand molecules</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Average B factor (Å$^2$)</td>
<td>29.6</td>
<td>31.8</td>
<td>51.2</td>
<td>28.1</td>
<td>33.5</td>
<td>36.9</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most favored regions (%)</td>
<td>97.7</td>
<td>98.5</td>
<td>98.1</td>
<td>97.0</td>
<td>97.7</td>
<td>96.2</td>
</tr>
<tr>
<td>Allowed regions (%)</td>
<td>2.3</td>
<td>1.5</td>
<td>1.9</td>
<td>3.0</td>
<td>2.3</td>
<td>3.8</td>
</tr>
<tr>
<td>PDB code</td>
<td>3GEZ</td>
<td>3GF2</td>
<td>3GFJ</td>
<td>3GFL</td>
<td>3GFM</td>
<td>3GFM</td>
</tr>
</tbody>
</table>

RMSD in bond angles and bond lengths are varied from 0.005 to 0.007 (Å) and 1.0–1.1 (°). Solvent content is about 55%.

Values in parentheses are for the highest resolution shell.

$^b$Rmerge $= \left( \sum_{h} \sum_{i} |I(h,i)| - \langle |I(h,i)| \rangle \right) / \left( \sum_{h} \sum_{i} |I(h,i)| \right)$, where $|I(h,i)|$ is the intensity value of the $i$th measurement of $h$ and $\langle |I(h)| \rangle$ is the corresponding mean value of $|I(h)|$ for all measurements.

$^c$Rwork $= \left( \sum_{h} \sum_{i} |F_{obs}\rangle - |F_{calc}\rangle \right) / \left( \sum_{h} \sum_{i} |F_{obs}\rangle \right)$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively.

$^d$Rfree is the same as R factor, but for a 5–7% subset of all reflections.
In conclusion, this report describes the crystal structure of ST1710 in three different forms: apo-form, ST1710–salicylate and ST1710–DNA complex. We showed that the salicylate binding does not affect the overall structure. In addition we found that all the residues interacting with salicylate ligand are highly conserved among the closely related proteins. The ST1710 conformation changed dramatically upon binding to the DNA, and these changes might be necessary for its recognition. Structural analyses of the MarR family of regulators along with the bacterial transcription regulators such as BmrR (40) and CRP (41) and the eukaryotic transcription regulators like RFX1 (42) and histone H5 (43) suggested that orientations of the wHTH motifs were similar despite having little differences in loop lengths and its orientations (19). We believe most of the regulators containing wHTH motifs are prone to bind DNA through their acidic/basic residues located in their wings. However, the mode of DNA recognition depends on the subunit organization of the regulators as observed in the MarR family of proteins (ST1710, MarR).

ACCESSION NUMBERS
3GEZ, 3GF2, 3GFI, 3GFJ, 3GFL, 3GFM.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The authors thank Dr K. Yutani and Ms. M. Takehira for DSC measurements and Dr T. Ishikawa for his moral support and encouragement.
REFERENCES


