Vol. 12, pp. 47–56 (2015)
doi: 10.2142/biophysico.12.0_47

Biophysics and Physicobiology
https://www.jstage.jst.go.jp/browse/biophysico/

Regular Article

Structural comparison between the open and closed forms of citrate synthase from *Thermus thermophilus* HB8

Eiji Kanamori[1], Shin-ichi Kawaguchi[2], Seiki Kuramitsu[3], Tsutomu Kouyama[1] and Midori Murakami[1]

[1]*Department of Physics, Graduate School of Science, Nagoya University, Nagoya, Aichi 464-8602, Japan*
[2]*Graduate School Frontier Biosciences, Osaka University, Osaka 565-0871, Japan*
[3]*Graduate School of Science, Osaka University, Osaka 560-0043, Japan*

Received July 16, 2015; accepted September 16, 2015

The crystal structures of citrate synthase from the thermophilic eubacteria *Thermus thermophilus* HB8 (TtCS) were determined for an open form at 1.5 Å resolution and for closed form at 2.3 Å resolution, respectively. In the absence of ligands TtCS in the open form was crystalized into a tetragonal form with a single subunit in the asymmetric unit. TtCS was also co-crystallized with citrate and coenzyme-A to form an orthorhombic crystal with two homodimers in the asymmetric unit. Citrate and CoA are found in the active site situated between the large domain and the small domain in all subunit whereas the complex shows two distinct closed conformations, the fully closed form and partially closed form.
Structural comparisons are performed to describe conformational changes associated with binding of products of TtCS. Upon binding of citrate, basic residues in the active site move toward citrate and make a hydrogen bond network in the active site, inducing a large-scale rotation of the small domain relative to the large domain. CoA is sandwiched between the small and large domains and then the cysteamine tail is inserted into the active site with a cooperative rotation around mainchain dihedrals in the hinge region connecting helices M and N. According to this rotation these helices are extended to close the active site completely. The considerable flexibility and structural rearrangements in the hinge region are crucial for an ordered bibi reaction in catalysis for microbial CSs.

Corresponding author: Midori Murakami, Department of Physics, Graduate School of Science, Nagoya University, Furo-Cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan.
e-mail: midori@bio.phys.nagoya-u.ac.jp

Key words: X-ray crystallography, conformational change, domain motion, hinge motion, microbial citrate synthase

Citrate synthase (CS, E.C. 2.3.3.1), an enzyme involved in the first step of the tricarboxylic acid cycle, catalyzes the condensation of acetyl-coenzyme-A (CoA) with oxaloacetate via an intermediate citryl coenzyme A to yield citrate and CoA by an ordered bi-bi mechanism [1,2]:

$$\text{oxaloacetate} + \text{acetyl-CoA} + H_2O \rightarrow \text{citryl CoA} \rightarrow \text{citrate} + \text{CoA}$$

CS is found in almost all living cells and has been sequenced and characterized. Extensive structural studies of pig and chicken CSs have revealed their open and closed forms [3–6]. The enzyme is composed of two identical monomers of Mr ~50,000 with 20 helices. In comparison with the open and closed forms of pig CS, a monomer can be divided into two domains, a large domain (helices A through M, S, and T) and a small domain (helices N through R, following the nomenclature of Remington [3]). The active site is situated between the small and large domains. Upon binding of oxaloacetate/citrate to the active site, a large (~ 20°) rotation of the small domain is induced with respect to the large domain. The large domain is a core structure of the dimer and appears to be fixed during the rotation. The small domain appears to be translocated as a rigid body owing to distortion of two residues, His245 and Gly255 [7–10]. These residues lie in the MN loop connecting the large and small domains.

CS from microorganisms, archaebacteria, and bacteria has been intensively studied both kinetically and structurally.

X-ray crystallographic studies using CS from psychrophilic to thermophilic sources have revealed that the enzyme has a dimeric form, similar to that from higher animals [11–16]. It has 16 helices in each monomer of Mr ~45,000, indicating that it can play essential roles in catalytic reactions. Most of these structural studies on microbial CS have described the molecular basis of protein thermostability; however, the mechanism of conformational change between the open and closed states is unclear because structures in ligand-free and ligand-bound states are not available from the same species at high resolution [17]. In addition, the binding site of CoA that emerges after binding of citrate is unclear, given that the cysteamine tail of CoA is not visible in all closed conformations of microbial CSs in complex with the products although their active sites are closed [11–13].

CS from the thermophilic eubacterium *Thermus thermophilus* HB8 (TtCS) exists as a dimer of identical subunits of Mr 42,304 with 377 amino acid residues each. We obtained a tetragonal and an orthorhombic crystal of TtCS in the absence and presence of ligands and solved the open form of TtCS at 1.5 Å resolution and closed forms complexed with products at 2.3 Å resolution. In this study, we performed structural comparisons between the two forms of TtCS to identify the conformational change needed for catalytic reaction in the microbial enzyme.

Materials and Methods

Materials

The chemicals used in the final crystallization reaction were purchased from Wako Chemical.

PCR amplification of the TtCS gene

The TtCS gene was amplified from plasmid pT7Blue.

Subcloning of the TtCS gene into pET-3a expression vector

The PCR product was electrophoresed and purified from 1.2% low-melting agarose. The insert was then ligated into the pET-3a expression vector previously digested with *Nde*I and *Bgl*II. *Escherichia coli* strain BL21(DE3)pLysE was transformed with the ligation products by electroporation. A recombinant with the TtCS gene was identified by restriction analysis and DNA sequencing.

Expression and purification of CS

E. coli carrying the recombinant TtCS was grown in 4 l of LB medium containing 50 μg/ml ampicillin at 32°C to an absorbance of 0.4 at 650 nm. An equal volume of pre-heated (60°C) LB medium containing 50 μg/ml ampicillin was then added, the temperature of the culture was reduced to 42°C to induce TtCS transcription, and the culture was allowed to grow for 2.5 h in a water bath shaker. The cells were harvested by centrifugation (5 min at 8,000 rpm), suspended in 10 mM potassium phosphate buffer (pH 7.2) containing 1 mM EDTA and 10 mM 2-mercaptoethanol (buffer A), and disrupted by sonication. After removal of cell debris, the cell extract was heated at 70°C for 10 min to denature most of the host proteins and aggregated materials were removed by centrifugation at 14,000 rpm for 30 min. The supernatant was loaded onto an affinity column (RedTOYOPEARL) pre-equilibrated with buffer A. The column was then washed with buffer A and eluted with a linear NaCl gradient (0.3–1.5 M NaCl). Fractions containing the recombinant TtCS were collected and applied to a Sephacryl S-200 HR column pre-equilibrated with 10 mM potassium phosphate buffer (pH 7.2) containing 1 mM EDTA, 10 mM 2-mercaptoethanol, and 0.1 M NaCl. Fractions containing TtCS were diluted with 10 mM potassium phosphate buffer (pH 6.8) containing 10 mM 2-mercaptoethanol (buffer B) and applied to a column containing hydroxyapatite equilibrated in buffer B. Elution was achieved with a linear gradient (10–600 mM) of potassium phosphate buffer containing 10 mM 2-mercaptoethanol. Pooled fractions were dialyzed against 20 mM Tris-HCl (pH 7.5) containing 1 mM 2-mercaptoethanol.

Crystallization

For crystallization of TtCS in the open form, the recombinant protein solution (40 mg/ml) in 100 mM HEPES (pH 7.5) was used. Crystals were obtained by the hanging-drop vapor-diffusion method against 1.5–1.8 M lithium sulfate and 100 mM HEPES (pH 7.5) at 20°C. The final conditions employed a 3-μl drop containing 2 μl of protein solution and 1 μl of reservoir solution. A bipyramidal crystal grew to dimensions of $0.2 \times 0.2 \times 0.3$ mm^3 in several days.

For crystallization of the closed form, the protein solution (40 mg/ml) in 100 mM KCl, 10 mM CoA-3Na, and 10 mM Na-citrate (pH 5.6) was incubated for 6 h at room temperature. Crystals were obtained by the hanging-drop vapor-diffusion method against 1.5 M lithium sulfate and 100 mM Na-citrate (pH 5.8) at 20°C. The orthorhombic crystal grew to dimensions of $0.4 \times 0.2 \times 0.2$ mm^3 in several days.

Diffraction data collection, scaling, and refinement

Crystals of both forms were soaked in mother liquor containing 30% (v/v) glycerol for ~15 min and then flash-cooled in liquid propane. Diffraction data for the open form and closed form were collected at beamline BL44B2 and BL41XU of SPring-8, respectively, with crystals maintained at 100 K in liquid nitrogen. Diffraction data were processed with Mosflm [18] and SCALA incorporated in the CCP4 program suite [19].

The structures of both forms were solved by molecular replacement using CNS 1.0 [20] and XtalView [21]. To solve the structure of the open form, a search model was constructed from the structure of *Pyrococcus furiosus* CS (PfCS) (PDB ID; 1AJ8) in which the side chains of non-conserved residues were truncated to alanine. After a rotational and translational search, a marked drop in the R_{work} value (from 0.52 to 0.42) was observed when the simulated annealing

Table 1 Data collection and final refinement statistics

	Open form	Closed form
Data collection		
Resolution (Å)	20.0–1.5	24.1–2.3
Data completion (%) (outer shell)	96.6 (100)	90.7 (76.7)
Number of unique reflections	86558	66309
Multiplicity	8.3	4.2
Rsym[1] (%) (outer shell)	7.7 (44.9)	7.9 (21.7)
Unit cell (Å)	84.55 84.55 153.52	80.01 110.63 184.38
(°)	90 90 90	90 90 90
Space group	$P4_32_12$	$P2_12_12_1$
Refinement		
Resolution limit (Å)	20.0–1.5	14.79–2.3
Number of protein atoms	2929	11617
Number of water	469	517
Number of ligand atoms	27	318
Rwork[1] (%) (outer shell)	17.2 (18.6)	17.7 (18.6)
Rfree (%) (outer shell)	18.5 (21.3)	21.7 (25.2)
R.m.s.deviation of bond length (Å)	0.004	0.007
R.m.s.deviation of bond angle (º)	1.2	1
Estimated error from Luzzati Plot (Å)	0.15	0.22

[1] Rsym = ΣhklΣi |Ii - <I>| /ΣhklΣi Ii, where Ii is the intensity of an individual reflection and <I> is the mean intensity obtained from multiple observations of symmetry related reflections.
[2] Rwork = Σhkl (Fobs - | Fcalc |) / Σhkl | Fobs | (9.2% randomly omitted reflections were used for Rfree)

was refined on the assumption of space group $P4_32_12$. At this stage, a large part of the polypeptide main chain was clearly visible in the electron density map. Invisible parts of the polypeptide were truncated, and after the second simulated annealing, alanine residues that had been tentatively placed in the search model were replaced with correct residues by monitoring the omit map. The protein conformation was modified manually on the basis of the 2Fo-Fc map. Glycerol and bicarbonate molecules were placed into the active site after the electron densities had emerged there. Water molecules were added where stereochemically reasonable after the protein part was completed. Subsequent refinements, including simulated annealing and individual B-factor refinement, resulted in an R_{work} of 17.2 % and an R_{free} of 18.5% (Table 1).

To solve the structure of the closed form, calculation from the unit cell dimension was performed to show the Matthews number V_M, calculated as 4.8 if the asymmetric unit contains four dimers. A dimer model was created by rotating the open form structure of TtCS (PDB ID; 1IOM, one monomer per asymmetric unit) around a crystallographic two-fold axis and used as the search structure. First, rigid-body refinement was performed using the dimer model, showing both an R_{work} and R_{free} of 0.42. Then, each monomer was divided into two rigid groups (residues 3–219 and 324–376 as group 1 and residues 220–323 as group 2) and was again subjected to rigid-body refinement. The coordinate had an R_{work} and R_{free} of 0.34 and 0.41, respectively. The subsequent steps included refinement by simulated annealing. Citrate and CoA molecules were placed into the active site after the electron densities emerged there. Water molecules were added where stereochemically reasonable after the protein part was completed. The final model contained 1473 amino acid residues with four subunits, four citrate ions, four CoA molecules, 12 sulfonate ions, four glycerols, and water molecules in the asymmetric unit. The electron density map around the region corresponding to the PQ loop in the small domain was not visible and these residues were removed. The side chains of some residues located on the molecular surface were disordered and these atoms were also removed. The R_{work} was 17.7% and the R_{free} value was 21.7% for the final coordinate. A summary of the diffraction and refinement data statistics is given in Table 1.

Figures were generated with Pymol [22] and LIGPLOT [23]

Results and Discussion

Overall structure of the open form

The crystal structure of TtCS in the open form contains 374 amino acid residues of one subunit in the asymmetric unit. One bicarbonate, one sulfonate, three glycerol, and 401 water molecules are also observed in it.

Figure 1a shows the dimer structure of the open form. Similar to many other bacterial CSs, TtCS consists of 16 α-helices labeled C through G and I through S. The monomer can be divided into two domains: a small domain consisting of helices N–R (residues 219–323 approximately) and a large domain consisting of helices C–G, I–M and S (residues 3–218 and 324–377) (Fig. 2b and 2c). The domains

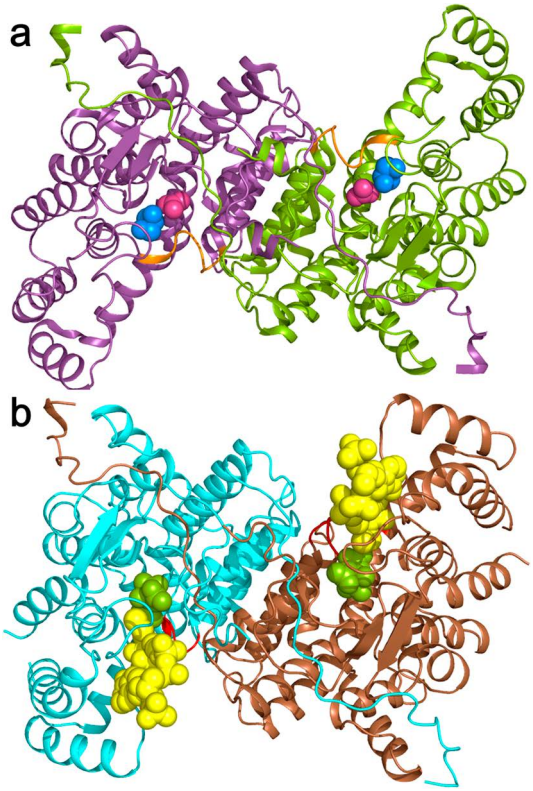

Figure 1 Overall structures of homodimer of TtCS. Polypeptide chains are shown with a ribbon model and ligand molecules are shown with a sphere model. The view is down a crystallographic 2-fold axis in the crystal of the open form, which relates the two subunits of the dimer. (a) Structure of the open form. Polypeptide chains A and B are colored in lime and purple, respectively, but the MN loop is colored in orange. Bicarbonate ion and glycerol are colored in red and blue, respectively. (b) Structure of the closed form. Polypeptide chains A and B are colored in brown and cyan, respectively, but the MN loop is colored in orange. Citrate and CoA are colored in green and yellow, respectively.

are linked by an MN loop as a hinge. The active site of the enzyme lies in the cleft between these two domains. Two subunits associate tightly to form a dimer, with the dimer interface comprising an eight α-helical "sandwich" consisting of four antiparallel pairs of helices, FF′, GG′, MM′, and LL′ (the prime indicates the two-fold symmetry operation) (Fig. 2b).

Active site

The active site of the open form contains one bicarbonate ion and one glycerol molecule, the latter added as a cryoprotectant (Fig. 3a). These heterogroups form an extensive hydrogen bond network containing the five basic residues His184, His219, His258, Arg267, and Arg337, mimicking citrate that is bound to the closed form (Fig. 3c).

Overall structure of the closed form

The crystal structure of TtCS in complex with the products citrate and CoA contains 1473 amino acid residues of four subunits in the asymmetric unit. Two homodimers in the asymmetric unit are roughly related by a local two-fold axis. Twelve sulfonate ions and four glycerol molecules contribute to stabilizing the crystal lattice by hydrogen bonds with the enzyme. Subunits A and B form a dimer as shown in Figure 1b, and subunits C and D also form a dimer in the asymmetric unit. Each subunit is composed of 16 α-helices and can be divided into a large domain and a small domain as in the open form. Citrate and CoA are found in the active site between the two domains of all subunits. Among the four subunits, subunits B, C and D are in the fully closed form in which the active site is closed completely. In contrast, subunit A has an alternative conformation in the partially closed form in which the active site is accessible to the bulk solvent. Superposition of the four subunits on one another yields a root-mean-square deviation (r.m.s.d.) for least-squares fitting of 0.44 Å and 0.33 Å for the 366 Cα atoms between subunits B and C or B and D, but 1.19 Å for the 370 Cα atoms in the superposition of subunit A onto subunit B. For convenience in the following sections, we will refer to the conformation of subunit B as the fully closed form and that of subunit A as the partially closed form.

This conformational differences among them are due to differences in crystal contacts (Fig. 2d). The large domains of both dimers contact each other via water molecules with average temperature factors <B>s of 19–22 Å^2. The small domains of subunits A (<B>=26 Å^2) and B (<B>=29 Å^2) are stabilized in the crystal lattice by hydrogen bonds with residues in the large domains of neighboring subunits. The small domain of subunit D (<B>=41 Å^2) interacts with that of subunit A, and the small domain of subunit C (<B>=51 Å^2) is not involved in crystal contact. It indicated that TtCS exists in the equilibrium between the fully and partially closed forms in the crystal, and it would favor the fully closed form over the partially closed form under the crystallization condition in the presence of products.

Conformational change between the open and closed forms

The large and small domains of each structure are nearly identical in conformation between both closed forms (Fig. 2). The large domain in the open form is also nearly identical to that in the fully closed form and partially closed form, with an r.m.s.d. of 0.39 Å and 0.30 Å, respectively, for 269 Cα atoms. We accordingly regarded the large domain, which is fixed during a conformational change, as the core structure of a dimer. Comparison of the small domain between the closed and open forms clearly shows the presence of two subgroups. The upper part constituted of helices N, O, and Q acts as a rigid body, which are identical in conformation with an r.m.s.d. of 0.32 Å and 0.30 Å for the 38 Cα atoms comprising α-helices in the fully and partially closed forms, respectively. Thus, variation in the full conformations of each subunit is characterized with respect to the position of the rigid-body group in the small domain relative to the large

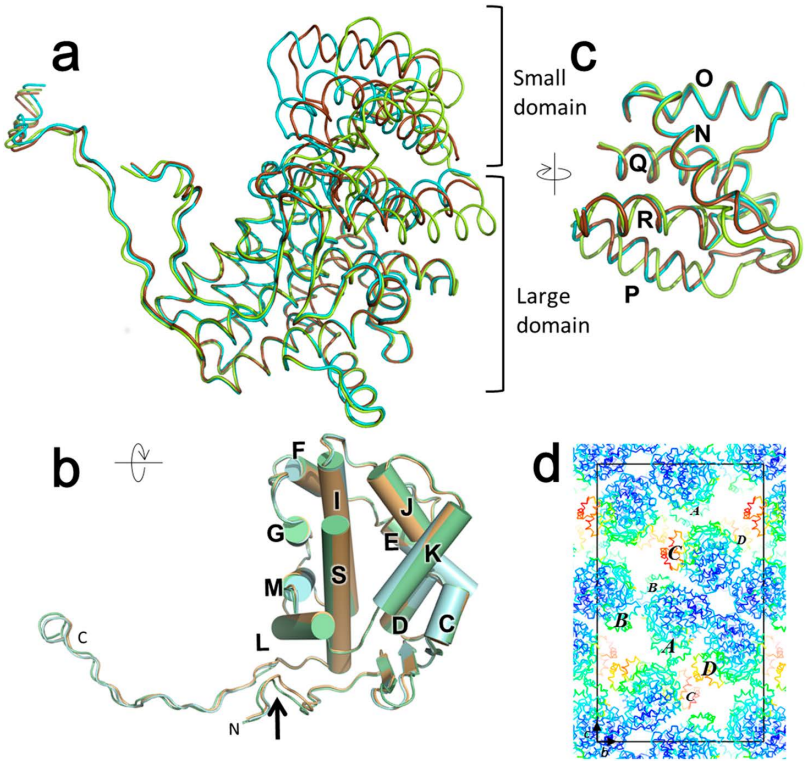

Figure 2 Structural comparison of the polypeptide chains between the open form and closed forms. (a) View of the subunits of the open form (lime), partially closed form (brown) and fully closed form (cyan) in which each large domains are superimposed. Polypeptide chains are shown with a ribbon model. (b) View of the large domains. Polypeptide chains are shown with a cylinder model. Black allow shows the two-fold axis. (c) View of the small domains in which helices N, O and Q in the small domain are superimposed. Polypeptide chains are drawn with a ribbon model. (d) Protein packing of four subunits in the orthorhombic crystal of TtCS, viewed along the a-axis. Polypeptide chains are drawn with a ribbon model and colored accordingly to the temperature factor of Cα atoms, from dark blue for low to red for high temperature factor. The unit cell is marked off by the black line. Large and small labels of four subunits indicate the small domain is situated at the front and back side of the page, respectively.

domain. A strand constituted by seven residues, the MN loop, connects the rigid-body group of the small and large domains.

We found that the large-scale motion of the rigid-body group in the small domain is described by a rotation and a translation relative to the large domain. The rigid body of helices N, O, and Q is rotated by 30° about a rotation axis and translated by 4.6 Å perpendicular to the axis and 0.6 Å downward along the axis in the fully closed form. In the partially closed form, it is rotated by 25° and translated by 3.0 Å perpendicular to the axis and by 0.1 Å along the axis. The rotation axes are inclined somewhat differently to each other. They pass close to the MN loop directly connecting the large and small domains. These observations suggest that distinct forces that originate around the MN loop, which is the active site, act on the enzyme during each transition from the open to the fully or partially closed forms.

Citrate-binding site

Citrate is bound deep in the cleft of the active site by hydrogen bonds of three carboxyl groups in three conserved histidines (His184, His219, and His258) as well as three conserved arginines (Arg267, Arg337, and Arg357′; the prime means belonging to the adjacent subunit) in the fully and partially closed forms (Fig. 3b and c). The schematic diagram of citrate is illustrated in Figure 4c. Three residues in the large domain (His184, Arg337, and Arg357′) show identical conformations to interact with the C2- and C3-carboxyl groups on citrate in both closed forms, respectively, but Arg357′ protrudes in the partially closed form (Fig. 3b). His219 in the MN loop also forms a hydrogen bond with the C4-carboxyl group of citrate identically in both closed forms. Two basic residues (His258 and Arg267) and one acidic residue (Asp312) from the small domain, which are on the opposite side from the large domain in the active site, are shifted greatly toward citrate during the conformational change from the open form to both closed forms.

It has been proposed that three ionizable residues, His274, His320, and Asp375, in the active site participate in acid–base catalysis by pig CS [3,6]. These residues are also conserved as His219, His258 and Asp312 in TtCS, respectively, and linked in this site with hydrogen bonds to citrate in both closed forms, suggesting that the catalytic mechanism of TtCS is the same as that of pig CS.

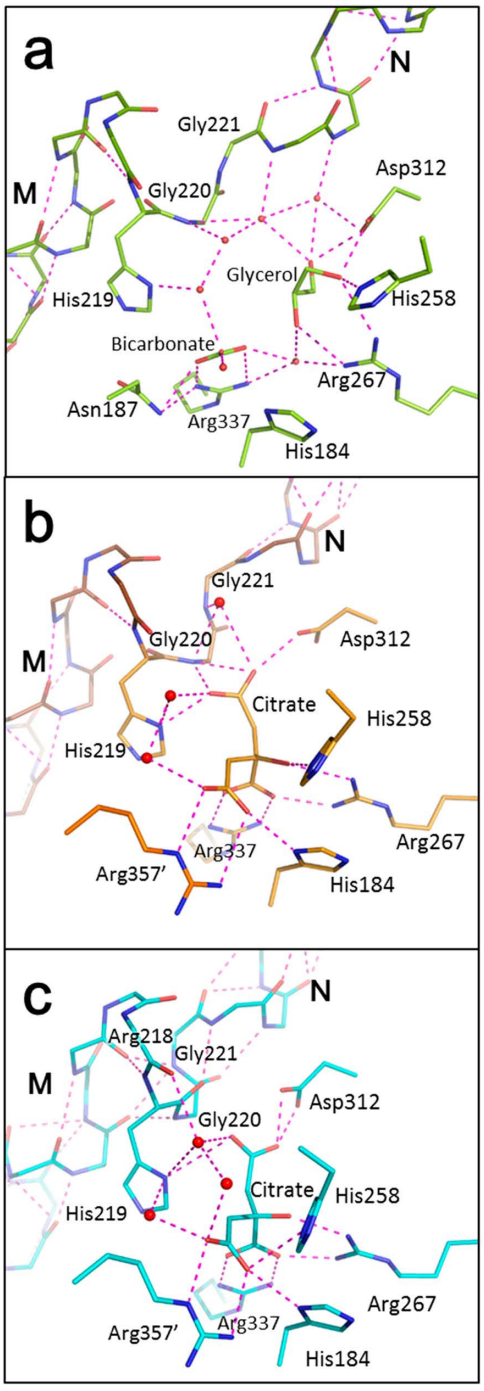

Figure 3 View in the citrate binding site of the open form (a), the partially closed form (b) and the fully closed form (c), respectively. Carbon atoms are shown in lime, orange and cyan. Oxygen, nitrogen and sulfur atoms are shown in red, blue and orange, respectively. Dotted lines indicate hydrogen bonds.

CoA-binding site

CoA is bound at the upper part of the active site just above the citrate-binding site in a compact form because of an intramolecular hydrogen bond between AN7 of the adenine ring and PO9 of the pantetheine tail in both closed forms (The schematic diagram of CoA was illustrated in fig. 4c). As shown in Figure 4a and b, the electron density maps are clearly visible in both closed forms but show some differences around the cysteamine tail.

The CoA molecule is attached along the OP loop by hydrogen bonds of the main chain of the protein for the adenine ring (AN6-Ile253O, AN6-M256O) and for the pantetheine part (PO10-Met256O, PO5-Gly255N, PN8-Met256O). These main-chain interactions are common in the closed structures of CS with CoA or its analog in microorganisms and higher animals [3,7–9]. The interaction between the carbonyl oxygen PO5 of CoA and the side chain of Asn310 of the QR loop is observed in both closed forms and is also conserved in all known sequences of CS, suggesting that these main-chain and side-chain interactions are essential for the recognition of CoA. The adenine ring is sandwiched between the extended side chain of Arg252 of the OP loop and the side chains of Leu302 in the carboxy terminal of the Q helix and Ile307 in the QR loop, and the guanidino group of Arg252 forms salt bridges to the α-phosphate of CoA.

In the fully closed form, Arg354′ of the carboxy-terminal loop of the neighboring subunit and Arg259 from the OP loop bind to α- and β-phosphate, respectively, resulting in a 5′-phosphate region of CoA setting in the middle between the large and small domains. The upper part of CoA is thus surrounded and held by three arms of arginines and the OP loop in the fully closed form. In contrast, an α-phosphate moiety is bound to Lys262 of the OP loop and there is no interaction with CoA by residues from the neighboring subunit in the partially closed form. This region of CoA in the partially closed form remains at the surface of the small domain.

In the tail region, an additional hydrogen bond between the nitrogen (PN4) of the cysteamine and main-chain oxygen of Arg218 of MN loop is observed in the fully closed form (Fig. 4d). This leads the terminal sulfur atom to be situated just above citrate near the C4-carboxyl moiety (Fig. 4a). In the partially closed form, the partner of the hydrogen bond of PN4 in the cysteamine is replaced with a water molecule, which is one of continuous bulk solvent (Fig. 4e). The terminal sulfur atom directs oppositely from Asn310 toward the opening of the cleft in the partially closed form (Fig. 4e).

MN loop

Figure 5 shows that the region from His219 through Gly221 is primarily responsible for twisting between the large and small domains during conformational change of TtCS. In the open form, the carbonyl oxygen of His219 directs outside of the active site (Fig. 3a). Upon formation of the fully closed form, it turns oppositely toward citrate (Fig. 3c). This turn is accomplished by an increase in the Cα-C angle of 110° in a right-handed direction during the domain

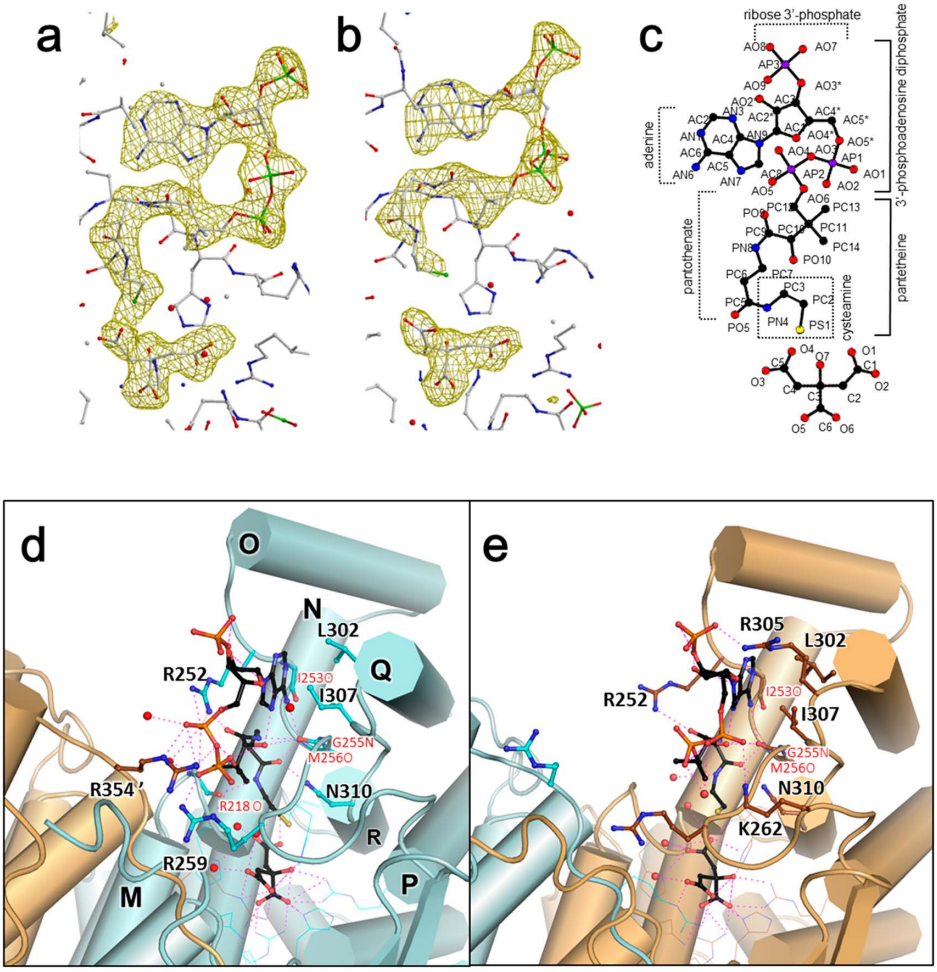

Figure 4 View around the active site with citrate and CoA. (a) Omit maps of citrate and CoA in the fully closed form (b) Omit maps of citrate and CoA in the partially closed form. Maps are contoured at 3σ superimposed with the final model shown in a wire model. (c) Schematic diagram of CoA (upper) and citrate (lower). (d, e) View around the binding site of CoA in the fully closed form (d) and partially closed form (e), respectively. Polypeptide chains are drawn with a cylinder model. Sidechains are drawn with a wire model. CoA and bounded residues are drawn with a stick model. The model are Dotted lines indicate hydrogen bonds. Carbon atoms are shown in white (a,b), black (c) , orange (d) and cyan (e). Oxygen, nitrogen and sulfur atoms are shown in red, blue and orange, respectively. Dotted lines indicate hydrogen bonds.

motion from the open to the fully closed form (Fig. 5a, 5c). Consequently, these three residues harmonically generate a large twist in the backbone required by the 30° rotation of the rigid-body group consisting of helices N, O, and Q.

It is noteworthy that the sequence of Gly216-Pro217-Arg218-His219 makes a rigid turn, owing to a main-chain hydrogen bond between the oxygen of Gly216 and nitrogen of His219. Because of this fixation of the backbone, the side chain of His219 is oriented toward the catalytic site in a similar manner in all forms in order to receive citrate/oxaloacetate. In the fully closed form, there is a main-chain hydrogen bond between the carbonyl of Arg218 and PN4 of CoA, leading the tail to enter the catalytic site near the citrate. In this case, the active site cleft is narrowed between the large and small domains so that the main-chain Cα-C bond in His219 rotates and the nitrogen of Gly220 is directed toward the carboxy-terminal end of helix M for the entry of CoA. This rotation orients the oxygens and nitrogens of the next two glycines (Gly220 and Gly221) to align with helices M and N. Consequently, the main-chain carbonyl oxygens of His219 and G221 form hydrogen bonds with the nitrogens of Gly222 and Ala225, respectively, to extend the amide-terminal end of helix N. Moreover, the main-chain nitrogens of Gly220 and Gly221 are hydrogen bonded with the main-chain carbonyl to Leu214, extending the carboxyl-terminal end of helix M in the fully closed form. Thus, a stabilized MN loop comprising the turn bound to CoA and the helical hinge closes the active site completely.

In the partially closed form, the profile of the dihedral angle around the hinge region is similar to that of the open form rather than that of the fully closed form. The main-chain nitrogen of Gly220 forms a hydrogen bond to the C4-carboxyl group of citrate in the partially closed form and some water molecules are observed around this region in the

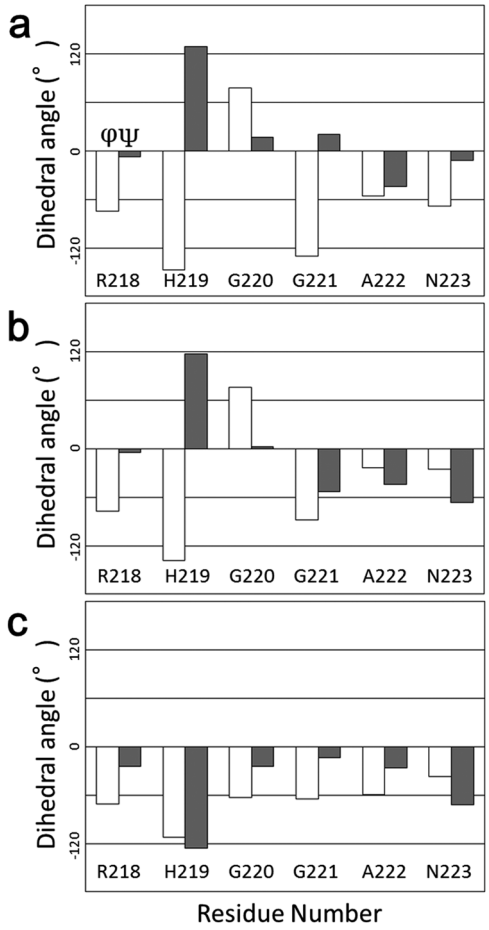

Figure 5 Dihedral angles around the N-Cα (φ) and Cα-C (ψ) bonds in the hinge region of the open form (a), the partially closed from (b) and fully closed form (c), respectively. Open bars indicate φ-angles, and gray bars indicate ψ-angles.

open and partially closed forms. There are no interactions with the cysteamine tail of CoA in the MN loop, and the cleft of the active site remains open to the bulk solvent. Thus, the hinge region appears to retain its formation in open to receive the cysteamine tail into the catalytic site.

Insights into the conformational change of TtCS

Here we showed an open and two closed forms of TtCS, with the conformational change described by a 30° rotation of the small domain relative to the large domain from the open to the fully closed form and by a 25° rotation from the open to the partially closed form. Comparison of the open and two closed structures of TtCS show that there is only a small difference of 5° in rotation angles between both closed forms, but a large difference in the hinge-twisting motion between them. We can distinguish these two motions and define distinct mechanisms underlying them: binding of citrate induces the domain movement, but twisting of the hinge region accompanied by binding of the tail of CoA evokes domain closure, which constrains the active site to be closed completely. The sequence His-Gly-Gly in the hinge region is conserved in microbial CSs, and in known structures, helical conformation of the hinge region is observed not only in the fully closed form of PfCS complexed with products in which terminal two atoms of cysteamine tail of CoA are not visible (Fig. 6d, h and l) but also in the closed form of CS from *Thermolasma acidophilum* (TaCS) with an intermediate analog, citryl dethia CoA (CDCoA) (PDB ID: 2R9E) (Fig. 6c, g and k).

The MN loop is apparently more rigid than that seen in the open and partially closed forms, because its conformation is additionally stabilized by main-chain hydrogen bonds. However, the helical conformation may be rather flexible, given that two glycines in this region make 3–10 helical turns. Although the hinge region is fully extended in the partially closed and open forms of TtCS, the open form of CS form *Sulfolbus tokodaii* (StCS, PDB ID: 1VGM) already makes a main-chain hydrogen bond between Leu211 and Gly220 to extend helix M with a small rotation in the main-chain carbonyl of His219 (Fig. 6a, e and i). This conformation is also observed in the partially closed form of TaCS binary complexed with oxaloacetate with a hydrogen bond between Leu217 and Gly223 (PDB ID: 2IFC) (Fig. 6b, f and j). In these two forms the dihedral angles in the hinge region are quite similar to each other (Fig. 6i and j), but a large difference in rotation angles of the small domain relative to the large domain between them as shown in TtCS (Fig. 6a and b, Fig. 2a). Since the dihedral angles in the hinge region of TaCS in the fully closed form are quite similar to those in the fully closed form of TtCS (Fig. 6i and Fig. 5c), it is suggested that rearrangements in the hinge region are needed to keep various molecules including substrates, products and reaction intermediates in the active site during the reaction. Thus, flexibility in the sequence His-Gly-Gly and its structural rearrangements are crucial during the ordered bi-bi catalysis reaction of microbial CSs.

Conclusion

In this study, we determined and compared crystal structures of CS from *T. thermophilus* HB8 in open and closed states. Structural comparison between the open form and two closed forms revealed that a large domain movement is evoked upon citrate binding to the active site. In addition, a twisting motion in the hinge region connecting the large and small domains is crucial for closing the active site. We propose that these distinct motions are essential for catalytic function and that the conservation of the three-residue His-Gly-Gly array around the hinge region would help elucidate the sequential ordered bi-bi catalytic mechanism of CS from microorganisms.

Acknowledgements

We thank the beam line staff of SPring-8 (Harima, Japan)

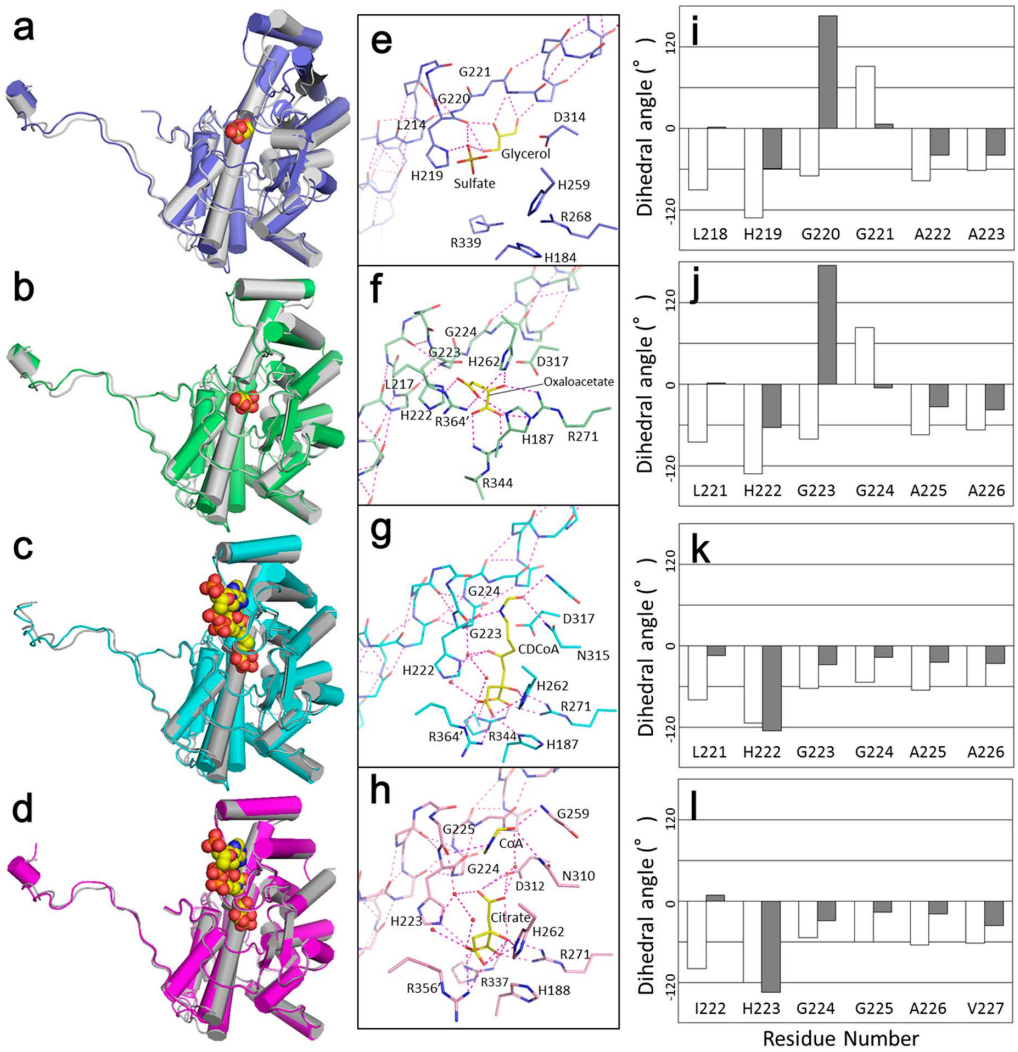

Figure 6 Structural comparisons between TtCS and other microbial CSs. (a–d) View of the subunits of StCS in the open form (a, purple), TaCS in the partially closed form (b, green), TaCS in the fully closed form (c, cyan) and PfCS in the fully closed form (d, magenta) in which Cα atoms are superimposed on those of TtCS (white) with r.m.s.d.s of 2.17 Å, 1.75 Å, 1.26 Å and 1.51 Å, respectively. (e–h) View in the active site of StCS in the open form (e), TaCS in the partially closed form (f), TaCS in the fully closed form (g) and PfCS in the fully closed form (h), respectively. Carbon atoms are shown in purple, green, cyan, magenta and yellow. Oxygen, nitrogen and sulfur atoms are shown in red, blue and orange, respectively. Dotted lines indicate hydrogen bonds. (i-l) Dihedral angles around the N-Cα (φ) and Cα-C (ψ) bonds in the hinge region of StCS in the open form (i), TaCS in the partially closed from (j), TaCS in the fully closed form (k) and PfCS in the fully closed form (l), respectively. Open bars indicate φ-angles, and gray bars indicate ψ-angles.

for technical assistance during data collection. We also thank Prof. Kunio Ihara for valuable comments.

Accession Number

Coordinates and structural parameters had been published with Protein Data Bank under accession code: 1IOM for the open form and 1IXE for the closed form.

Conflicts of Interest

The authors have declared that no competing interests exist.

Author Contributions

S. K. directed the project. T. K. supervised experiences and wrote the very first draft on the structure of the open form. S.-i. K. performed expression and purification. E. K. and M. M. carried out crystallization, data collection, and structural analyses, and wrote the paper.

References

[1] Nelson, D. L. & Cox, M. M. Reaction of the citric acid cycle. in *Lehninger's Principles of Biochemistry: Fourth Edition*. pp. 606-614 (W. H. Freeman and Co., New York, 2005).

[2] Remington, S. J. Structure and Mechanism of Citrate Synthase. *Curr. Top. Cell. Regul*. **33**, 209–229 (1992).
[3] Remington, S. J., Wiegand, G. & Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2·7 and 1·7 Å resolution. *J. Mol. Biol.* **158**, 111–152 (1982).
[4] Liao, D. I., Karpusas, M. & Remington, S. J. Crystal structure of an open conformation of citrate synthase from chicken heart at 2.8 Å resolution. *Biochemistry* **30**, 6031–6036 (1991).
[5] Wiegand, G., Remington, S. J., Deisenhofer, J. & Huber, R. Crystal structure analysis and molecular model of a complex of citrate synthase with oxaloacetate and S-acetonyl-coenzyme A. *J. Mol. Biol.* **174**, 205–219 (1984).
[6] Karpusas, M., Branchaud, B. & Remington, S. J. Proposed mechanism for the condensation reaction of citrate synthase: 1.9 Å structure of the ternary complex with oxaloacetate and carboxymethyl coenzyme A. *Biochemistry* **29**, 2213–2219 (1990).
[7] Wiegand, G. & Remington, S. J. Citrate synthase: structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 97–117(1986).
[8] Roccatano, D., Mark, A. E. & Hayward, S. Investigation of the mechanism of domain closure in citrate synthase by molecular dynamics simulation. *J. Mol. Biol.* **310**, 1039–1053 (2001).
[9] Gerstein, M., Lesk, A. M. & Chothia, C. Structural mechanisms for domain movements in proteins. *Biochemistry* **33**, 6739–6749 (1994).
[10] Taylor, D., Cawley, G. & Hayward, S. Quantitative method for the assignment of hinge and shear mechanism in protein domain movements. *Bioinformatics* **30**, 3189-3196 (2014).
[11] Russell, R. J., Ferguson, J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. The crystal structure of citrate synthase from the hyperthermophilic archaeon *pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry* **36**, 9983–9994 (1997).
[12] Russell, R. J. M., Gerike, U., Danson, M. J., Hough, D. W. & Taylor, G. L. Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium. *Structure* **6**, 351–361 (1998).
[13] Russell, R. J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. The crystal structure of citrate synthase from the thermophilic archaeon, Thermoplasma acidophilum. *Structure* **2**, 1157–1167 (1994).
[14] Bell, G. S., Russell, R. J. M., Connaris, H., Hough, D. W., Danson, M. J. & Taylor, G. L. Stepwise adaptations of citrate synthase to survival at life's extremes. From psychrophile to hyperthermophile. *Eur. J. Biochem.* **269**, 6250–6260 (2002).
[15] Ferraris, D. M., Spallek, R., Oehlmann, W., Singh, M. & Rizzi, M. Structures of citrate synthase and malate dehydrogenase of Mycobacterium tuberculosis. *Proteins* **83**, 389–394 (2015).
[16] Boutz, D. R., Cascio, D., Whitelegge, J., Perry, L. J. & Yeates, T. O. Discovery of a thermophilic protein complex stabilized by topologically interlinked chains. *J. Mol. Biol*. **368**, 1332–1344 (2007).
[17] Wells, S. A., Crennell, S. J., & Danson, M. J. Structures of mesophilic and extremophilic citrate synthases reveal rigidity and flexibility for function. *Proteins* **82**, 2657–2670 (2014).
[18] Steller, I., Bolotovsky, R. & Rossmann, M. G., An algorithm for automatic indexing of oscillation images using Fourier analysis. *J. Appl. Cryst*. **30**, 1036–1040 (1997).
[19] Collaborative Computational Project, Number 4., The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **D50**, 760–763 (1994).
[20] Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., *et al. Crystallography & NMR System*: A new software suite for macromolecular structure determination. *Acta Cryst.* **D54**, 905–921(1998).
[21] McRee, D. E. XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165(1999).
[22] The PyMOL Molecular Graphics System, Version 0.99. Schrodinger LLC.
[23] Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134(1995).