Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a conserved CoA-binding protein synthesized by a cell-free system

Takashi Wada,[a,b,c] Mikako Shirouzu,[a,b] Takaho Terada,[a,b] Yoshiko Ishizuka,[a] Takayoshi Matsuda,[a] Takanori Kigawa,[a] Seiki Kuramitsu,[b] Sam-Yong Park,[c] Jeremy R. H. Tame[c] and Shigeyuki Yokoyama[a,b]*

[a]RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, [b]RIKEN Harima Institute, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, and [c]Protein Design Laboratory, Yokoyama City University, 1-7-29 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

Correspondence e-mail: yokoyama@biochem.s.u-tokyo.ac.jp

Received 25 December 2002
Accepted 12 May 2003

PDB References: TT1466 from cell-free system, 1iul, r1iulsf; TT1466 expressed in *E. coli*, 1iuk, r1iuksf.

TT1466 is a hypothetical protein from the extremely thermophilic bacterium *Thermus thermophilus* HB8 and is highly conserved in bacteria and archaea. The selenomethionyl protein was synthesized by a cell-free system and the crystal structure was determined at 2.0 Å by MAD phasing. A native crystal was used for structure refinement to 1.7 Å. The structure is highly homologous to that of the CoA-binding domain of the succinyl-CoA synthetase from *Escherichia coli*, despite the protein having only 14% sequence identity to this domain. An isothermal titration calorimetry experiment was performed to investigate whether TT1466 binds CoA and revealed high-affinity CoA binding of TT1466.

1. Introduction

The recent progress of genome projects has generated a flood of sequence information. However, the majority of the identified proteins lack annotations regarding their functions (Zarembinski *et al.*, 1998; Christendat *et al.*, 2000). Thus, one of the crucial purposes of structural proteomics is to find clues to infer these unknown functions. A number of structures of hypothetical proteins have been determined for this purpose (Christendat *et al.*, 2000; Hwang *et al.*, 1999; Shin *et al.*, 2002; Zhang *et al.*, 2000; Teplova *et al.*, 2000; Saveanu *et al.*, 2002). For example, the crystal structure of a hypothetical protein, MJ0577 from *Methanococcus jannaschii*, contained a bound ATP, suggesting that the protein might be an ATPase or an ATP-mediated molecular switch. The deduced function was confirmed by biochemical experiments (Zarembinski *et al.*, 1998).

In this paper, we describe the crystal structure of a hypothetical protein consisting of 140 amino-acid residues, TT1466, from an extremely thermophilic bacterium, *Thermus thermophilus* HB8, and the structure-based analysis of its function. A BLAST database search revealed genes encoding TT1466 homologues in a wide range of bacterial and archaeal genomes (Altschul *et al.*, 1997) (Fig. 1*a*). These proteins consist of 130–173 residues and their sequence identities to TT1466 are between 26 and 48%. Selenomethionyl TT1466 protein was synthesized by a cell-free system (Kigawa *et al.*, 1999, 2001). The crystal structure was determined by multiple-wavelength anomalous dispersion (MAD) phasing (Hendrickson *et al.*, 1990; Hendrickson, 1991) and was refined using a native crystal. We discuss the possibility of CoA binding by TT1466 in terms of the structure and the results of isothermal titration calorimetry (ITC).

2. Materials and methods

2.1. Preparation of TT1466

The TT1466 gene from *T. thermophilus* HB8 was inserted into the pET11a plasmid (to give the pET-1466 plasmid). Selenomethionyl TT1466 was synthesized by the cell-free system using dialysis of 3 ml internal solution against 30 ml external solution with the same components as previously reported, except for the inclusion of 2 m*M* DL-selenomethionine (Wako) and 4 µg ml^{-1} pET-1466 plasmid as the template DNA (Kigawa *et al.*, 2001). Five of these reaction units were used. After incubation at 303 K for 8 h, the reaction solutions were collected and incubated at 341 K for 30 min. The following procedures were carried out at 277 K. After centrifugation, the supernatant was recovered and diluted to 100 ml with buffer *A* (20 m*M* Tris–HCl pH 8.0, 2 m*M* DTT). This was then applied to a Q-Sepharose column (Amersham Bioscience) and eluted with a linear gradient of sodium chloride (0–1000 m*M* NaCl in buffer *A*). The fractions containing TT1466 were pooled and diluted to 50 ml with buffer *A*. This solution was applied to a Mono-Q column (Amersham Bioscience) which was eluted with a linear gradient of sodium chloride (0–500 m*M* NaCl in buffer *A*). The target fractions were gathered and concentrated to 200 µl by centrifugation with a Centricon YM-3 filter (Millipore). This concentrated solution was applied to a Superdex75 gel-filtration column (Amersham Bioscience) run with a gel-filtration buffer (10 m*M* Tris–HCl pH 8.0, 100 m*M* sodium chloride, 1 m*M* DTT). The purity of the protein solution was checked by SDS–PAGE and a single band was observed. Native TT1466 was overexpressed in *Escherichia coli* BL21 (DE3) cells with the pET-1466 plasmid and grown in 2 l of LB medium with shaking at 310 K. Isopropyl-1-thio-β-D-galactopyranoside was added to 1 m*M* at A_{600} = 0.6. After 3 h induction, the cells were harvested and suspended in deionized water. The crude extract was treated by sonication, incubation and centrifugation. The recovered supernatant was subjected to the same chromatography procedures as described above.

2.2. X-ray crystallography

Crystals of selenomethionyl TT1466 were grown by the hanging-drop method at 293 K, using 9.2 mg ml^{-1} protein

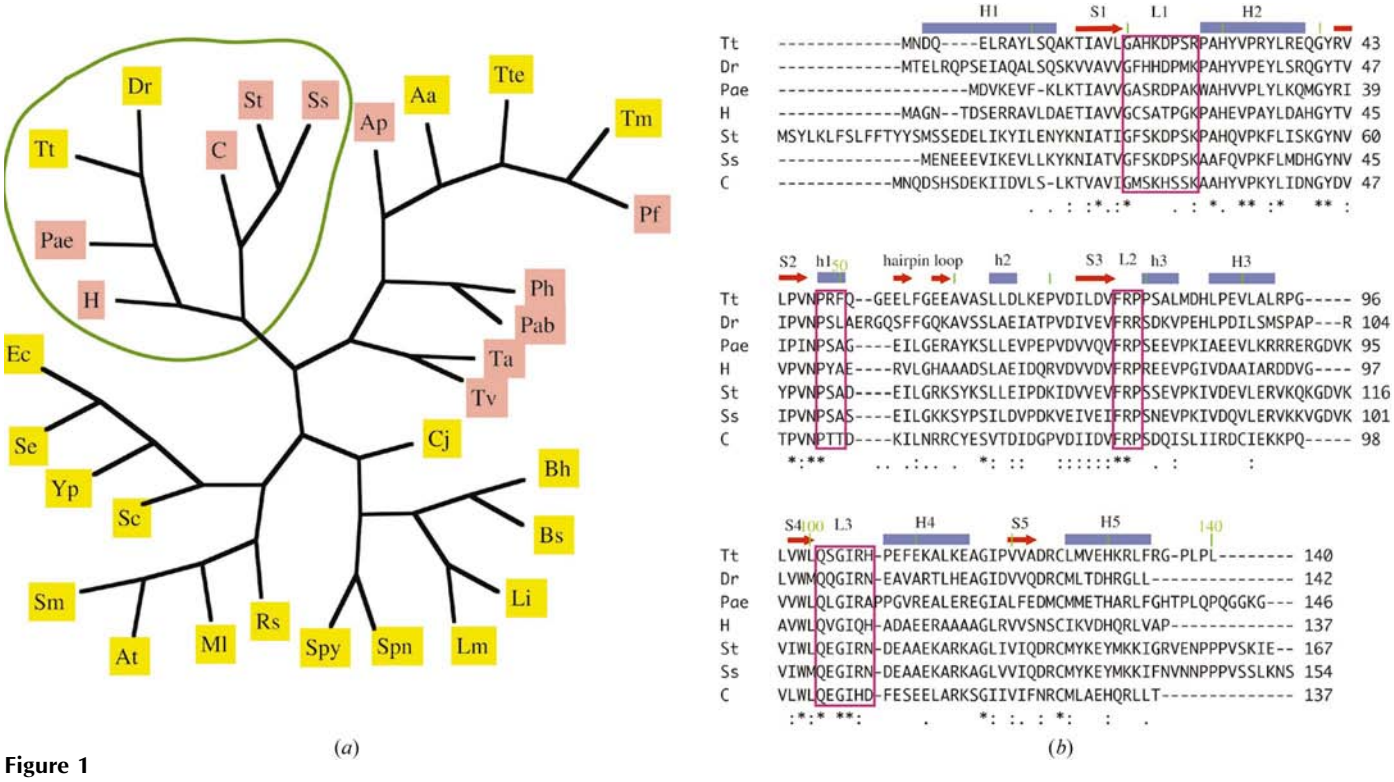

Figure 1
Analysis of TT1466-homologue sequences with the program *CLUSTALW* (Thompson *et al.*, 1994). (*a*) Phylogenic tree. Each abbreviation indicates the name of a species and the GenBank accession No. as follows: Aa, *Aquifex aeolicus*, AE000760; Ap, *Aeropyrum pernix*, AP000062; At, *Agrobacterium tumefaciens*, AE008053; Bh, *Bacillus halodurans*, AP001514; Bs, *Bacillus subtilis*, Z73234; C, uncultured crenarchaeote, AF393466; Cj, *Campylobacter jejuni*, AL139076; Dr, *Deinococcus radiodurans*, AE001913; Ec, *Escherichia coli*, AE000198; H, *Halobacterium* sp. AE005009; Li, *Listeria innocua*, AL596168; Lm, *Listeria monocytogenes*, AL591978; Ml, *Mesorhizobium loti*, AP003014; Pab, *Pyrococcus abyssi*, AJ248286; Pae, *Pyrobaculum aerophilum*, AE009888; Pf, *Pyrococcus furiosus*, AE010191; Ph, *Pyrococcus horikoshii*, AP000005; Rs, *Ralstonia solanacearum*, AL646074; Sc, *Streptomyces coelicolor*, AL592292; Se, *Salmonella enterica*, AL627269; Sm, *Sinorhizobium meliloti*, AL591786; Spn, *Streptococcus pneumoniae*, AE007320; Spy, *Streptococcus pyogenes*, AE006487; Ss, *Sulfolobus solfataricus*, AE006728; St, *Sulfolobus tokodaii*, AP000981; Ta, *Thermoplasma acidophilum*, AL445066; Tm, *Thermotoga maritima*, AE001795; Tt, *Thermus thermophilus*, AB097508; Tte, *Thermoanaerobacter tengcongensis*, AE013076; Tv, *Thermoplasma volcanium*, AP000993; Yp, *Yersinia pestis*, AJ414148. Bacterial and archaeal species are marked by yellow and pink backgrounds, respectively. (*b*) Alignment diagram for the branch enclosed by the green line in Fig. 1(*a*). The numbers at the right indicate the residue number at the right end. The four regions outlined in magenta correspond to the regions involved in CoA binding by SCS.

solution and a precipitant solution containing 50 m*M* sodium acetate pH 4.9, 100 m*M* ammonium sulfate, 15% PEG MME 2000. They belong to space group $P2_1$, with unit-cell parameters $a = 28.1$, $b = 64.9$, $c = 35.0$ Å, $\alpha = \gamma = 90.0$, $\beta = 106.5^\circ$, and contain one molecule per asymmetric unit. Crystals of native TT1466 were prepared by the hanging-drop method at 293 K, using 13.8 mg ml^{-1} protein solution and 20% PEG 400 as a precipitant solution. They belong to space group $P2_12_12_1$, with unit-cell parameters $a = 27.9$, $b = 64.6$, $c = 67.9$ Å, and contain one molecule per asymmetric unit. All data sets were collected at 100 K. For the MAD method, data from the selenomethionyl protein crystal were collected using synchrotron radiation at the BL44B2 station at SPring-8, Harima, Japan. Data from the native crystal were collected at the BL45XU station at SPring-8. All data sets were processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The positions of the Se atom were determined with the program *SOLVE* (Terwilliger & Berendzen, 1999) followed by density modification with the program *RESOLVE* (Terwilliger, 2000). The model was built with the program *TURBO-FRODO* (AFMB, CNRS). Structural refinement was performed using *X-PLOR* (Brünger, 1993). The data statistics are given in Table 1. The r.m.s. differences of the selenomethionyl and native TT1466 structures were calculated with *LSQKAB* and *DIFRES* (Collaborative Computational Project, Number 4, 1994).

Table 1
Statistics of the crystallographic analysis.

Values in parentheses refer to the highest resolution shell (2.07–2.00 Å for the data with a final resolution of 2.0 Å and 1.76–1.70 Å for that with 1.7 Å).

Data set	Edge	Peak	High remote	Low remote	Native
Wavelength (Å)	0.9804	0.9800	0.9747	0.9824	1.0200
Resolution range (Å)	13.0–2.0	13.0–2.0	13.0–2.0	13.0–2.0	50.0–1.7
Reflection (measured/unique)	31854/8044	31394/8037	32011/8107	31932/8007	62895/24119
Completeness (%)	98.7 (88.6)	98.9 (91.5)	99.8 (100.0)	98.2 (84.0)	92.4 (80.1)
$R_{\rm merge}$† (%)	6.2 (9.1)	6.5 (9.7)	8.8 (9.8)	8.3 (9.9)	6.0 (12.2)
Multiplicity	4.0	3.9	3.9	4.0	2.6
Phasing statistics					
Resolution (Å)		13.0–2.0			
Mean figure of merit after *RESOLVE* phasing (overall/centric/acentric)		0.68/0.61/0.68			
Refinement statistics					
Resolution (Å)		13.0–2.0			50.0–1.7
σ cutoff		0.0			0.0
$R_{\rm cryst}$ (%)		18.7			21.9
$R_{\rm free}$‡ (%)		27.0			28.4
R.m.s. deviations					
Bond length (Å)		0.004			0.007
Bond angles (°)		1.0			1.2
Water molecules		121			85
Ramachandran plot					
Residues in most favorable regions (%)		91.2			91.3
Residues in additional allowed regions (%)		8.8			8.7

† $R_{\rm merge} = \sum\sum_i |I(h) - I(h)_i| / \sum\sum_i I(h)$, where $I(h)$ is the mean intensity after rejections. ‡ $R_{\rm free}$ was calculated with 5% of data set aside for refinement.

2.3. Isothermal titration calorimetry

The TT1466 solution was dialyzed against a buffer consisting of 10 m*M* Tris–HCl pH 8.0 and 100 m*M* sodium chloride. CoA (Wako) was dissolved in the buffer used for the above dialysis. The lysozyme (Nakalai) solution and the CoA solution were prepared in the same manner. The ITC experiments were performed at 298 K using a CSC isothermal titration calorimeter model 4200 (Calorimetry Sciences Corp.). A 1270 µl aliquot of 128 µ*M* TT1466 solution was incubated in the cell and 10 µl aliquots of 2.34 m*M* CoA solution were then injected into the cell 20 times. To measure the dilution heat of the titrant, a blank titration was performed in the same way with no protein in the cell. An ITC experiment using 120 µ*M* lysozyme and 2.50 m*M* CoA was performed in the same way as a control. The data were analyzed using the program *BindWorks* (Calorimetry Sciences Corp.)

3. Results and discussion

3.1. Preparation of selenomethionyl TT1466 by cell-free synthesis

Recent improvements in cell-free synthesis systems now allow the production of several milligrams of a target protein (Kigawa *et al.*, 1999; Madin *et al.*, 2000). We have previously produced selenomethionyl Ras protein by cell-free synthesis and determined the crystal structure (Kigawa *et al.*, 2001). We therefore synthesized selenomethionyl TT1466 by the cell-free system. Five reaction units (3 ml internal and 30 ml external solutions at dialysis) produced 5.2 mg of the protein with almost 100% efficiency of selenomethionine incorporation (data not shown) in 8 h. The sample yield per volume of the reaction solution was 32 µg ml^{-1}. Native TT1466 was prepared by the conventional method, using overexpression in *E. coli* BL21 (DE3) cells. From 2 l of the LB culture medium, 19 mg of protein was obtained. The yield per volume of medium was 9.5 µg ml^{-1}. Thus, the cell-free system is efficient for protein preparation.

3.2. Crystal structure of TT1466

The structure of TT1466 synthesized by the cell-free system (CF-TT1466) was determined at 2.0 Å resolution by MAD phasing using selenomethionine. The molecular model fitted

well to the electron-density map (Fig. 2*a*) and was built from Met1 to Phe134 (Fig. 2*b*). The program *PROCHECK* (Laskowski *et al.*, 1993) showed that it comprised Asn2–Gln12 (helix H1), Thr15–Leu19 (strand S1), Pro28–Arg37 (H2), Arg42–Val46 (S2), Pro48–Phe50 (short helix h1), Glu54–Glu59 (hairpin loop), Leu64–Asp66 (h2), Ile73–Val76 (S3), Pro80–Ala82 (h3), Met84–Asp85 (h4), Leu87–Leu93 (H3), Val98–Leu100 (S4), Pro107–Glu115 (H4), Val120–Val121 (S5) and Leu126–Leu133 (H5) and that they formed a Rossmann fold (Rao & Rossmann, 1973). All the short helices h1–h4 are 3_{10}-helices. Since analytical ultracentrifugation showed that TT1466 is a monomer in solution (data not shown), oligomeric models were not considered. The structure of TT1466 over-expressed in *E. coli* cells (cell-based; CB-TT1466) was refined to 1.7 Å by molecular replacement using the CF-TT1466 model. The model was built from Met1 to Gly136 (Fig. 2*b*). It comprised Asp3–Gln12 (H1), Thr15–Leu19 (S1), Pro28–Glu38 (H2), Arg42–Val46 (S2), Pro48–Phe50 (h1), Glu54–Glu 59 (hairpin loop), Leu64–Asp66 (h2), Ile73–Val76 (S3), Pro80–Leu83 (h3), Leu87–Leu93 (H3), Val98–Leu100 (S4), Pro107–Glu115 (H4), Val120–Ala122 (S5) and Leu126–Phe134 (H5). The short helices h1 and h2 are 3_{10}-helices. Although the short helix h3 is located at Pro80–Asp85, there are two short helices, h3 and h4, in the same region of of CF-TT1466. Substituting selenomethionine for methionine at residue 84 may cause a small change in the local structure. However, the secondary and tertiary structures are almost the same as those of CF-TT1466 (Fig. 2*b*). The overall r.m.s. differences in the main chain and side chains were 0.27 ± 0.23 and 0.58 ± 0.61 Å, respectively.

3.3. Insight into CoA binding by TT1466 from the crystal structure

According to the DALI server (Holm & Sander, 1993), the structure of TT1466 is highly homologous to that of the CoA-binding domain of the α-subunit of succinyl-CoA synthetase (SCS) from *E. coli* (PDB code 2scu), with a *Z* score of 13, despite having only 14% sequence identity with this domain (Buck *et al.*, 1985; Wolodko *et al.*, 1994; Fraser *et al.*, 1999) (Figs. 3*a* and 3*b*). This result strongly suggests that TT1466 binds CoA.

The sequence of TT1466 has implicit identity (24%) to that of the CoA-binding domain of the acetyl-CoA synthetase (ADP-forming; ACS) from the amitochondriate eukaryote *Giardia lamblia*, which belongs to the acyl-CoA synthetase superfamily (NDP-forming; Sanchez *et al.*, 2000). Although the tertiary structure of this enzyme has not yet been determined, the secondary structure of the ACS CoA-binding domain could be predicted using that of TT1466 (Fig. 3*c*). This prediction should be more precise than that using the structure of SCS (Sanchez *et al.*, 2000), as the sequence identity between the CoA-binding domain of ACS and that of SCS is only 19% (Fig. 3*c*).

Here, we attempted to explore the CoA-binding mode of TT1466 with reference to that of SCS as no crystal of TT1466 with CoA has yet been obtained. The adenosine end of CoA, which forms an extended conformation, is anchored in SCS (Wolodko *et al.*, 1994; Fraser *et al.*, 1999) (Fig. 3*a*). Four regions are involved in CoA binding by SCS (Fraser *et al.*, 1999) (Fig. 3*b*). The first region is the loop Gly14–Gly20, which contains the CoA-binding motif G*X*S/TG*XX*G. The side chain of Thr16 strongly interacts with the 3′-phosphate of CoA *via* a hydrogen bond. The amide groups in the main chain of Ser18 and Gln19 interact with the 5′-diphosphate of CoA. However, these residues are not conserved in the corresponding loop L1 of the TT1466 homologues (Fig. 1*b*). The second region is the loop Thr39–Gly44. The side chain of Lys42 strongly interacts with the 3′-phosphate of CoA. Pro40 provides a hydrophobic interaction with the adenine ring of CoA. This loop corresponds to the short helix h1 of the TT1466 homologues, where Pro is conserved. The third region, Tyr71–Pro73, contains Val72, which interacts hydrophobically with the adenine ring of CoA and corresponds to the L2 loop of the TT1466 homologues. Although Val is not present in L2, there is the highly conserved sequence FRP which might be involved in CoA binding by TT1466. The fourth region, Ile95–Pro100, contains Ile95 and Glu97, whose amide groups at the main chain interact with the amide groups in a part of the panto-

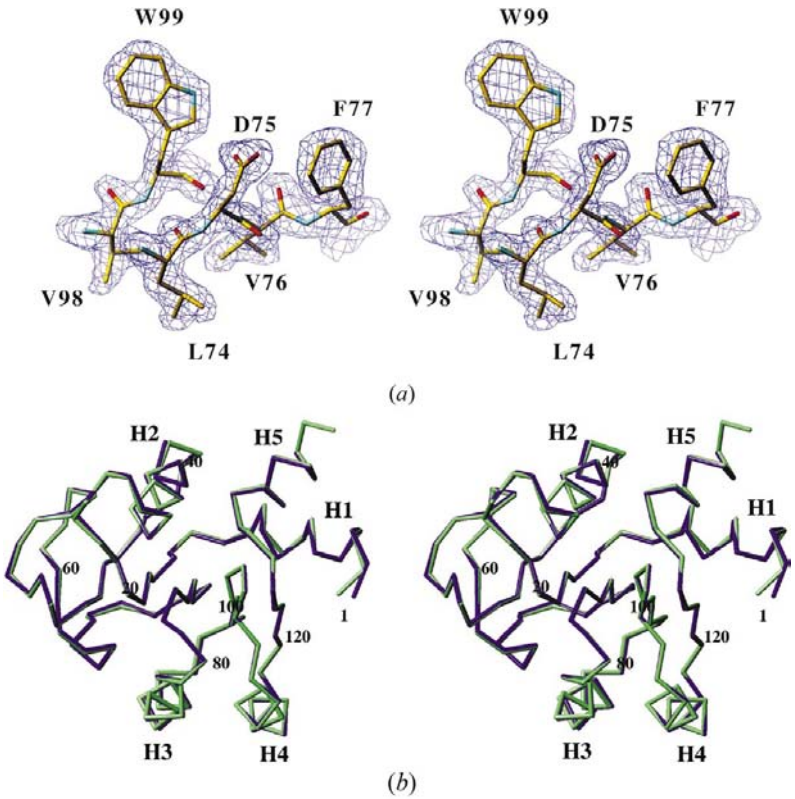

Figure 2
Crystal structures of TT1466 in stereoview. (*a*) A portion of the $2F_o - F_c$ electron-density map of CF-TT1466 at 2.0 Å, with the well fitted molecular models of Leu74–Val76 in S3 and Phe77, Val98 and Trp99 in S4. The contour level is 1.5σ. (*b*) Superimposition of the C^{α} backbone structures of CF-TT1466 and CB-TT1466, indicated by blue and green lines, respectively. Every 20th residue is labelled.

theine of CoA *via* hydrogen bonds. The corresponding loop L3 of the TT1466 homologues has the conserved sequence Q*X*GI, which might be involved in CoA binding.

These comparisons of the secondary structures do not explicitly show the CoA-binding mode of TT1466. However, several conformations of CoA exist among other CoA-binding proteins because of the flexibility of CoA (Engel & Wierenga, 1996). If TT1466 binds to CoA, then the CoA-binding mode of TT1466 may be different from that of SCS.

3.4. CoA-binding experiment for TT1466 by ITC

We examined the CoA binding of TT1466 by ITC. The experiments revealed the binding isotherm of CoA for TT1466 at 298 K (Fig. 4), with an enthalpy change $\Delta H = -32.3 \pm 1.9$ kJ mol^{-1}, an equilibrium constant $K = (1.2 \pm 0.5) \times 10^5\ M^{-1}$ and 1.7 ± 0.1 binding sites. The K value is about twofold less than that for the binding of CoA to SCS from *E. coli*, the apparent K_M of which is $2.5 \times 10^5\ M^{-1}$ (Joyce *et al.*, 1999). However, the ΔH and K of TT1466 are comparable to those for the binding of 2′-CMP to RNaseA [$\Delta H = -44.7 \pm 0.3$ kJ mol^{-1} and $K = (6.1 \pm 0.4) \times 10^4\ M^{-1}$ at 298 K], a system that is frequently used for ITC test reactions (Horn *et al.*, 2001). An ITC experiment on CoA binding to lysozyme under the same conditions as that on CoA binding to TT1466 showed only a small enthalpy change, $\Delta H = -1.2 \pm 0.2$ kJ mol^{-1}, indicating that CoA does not have high affinity for every protein. Thus, we conclude that CoA binds to TT1466 with high affinity.

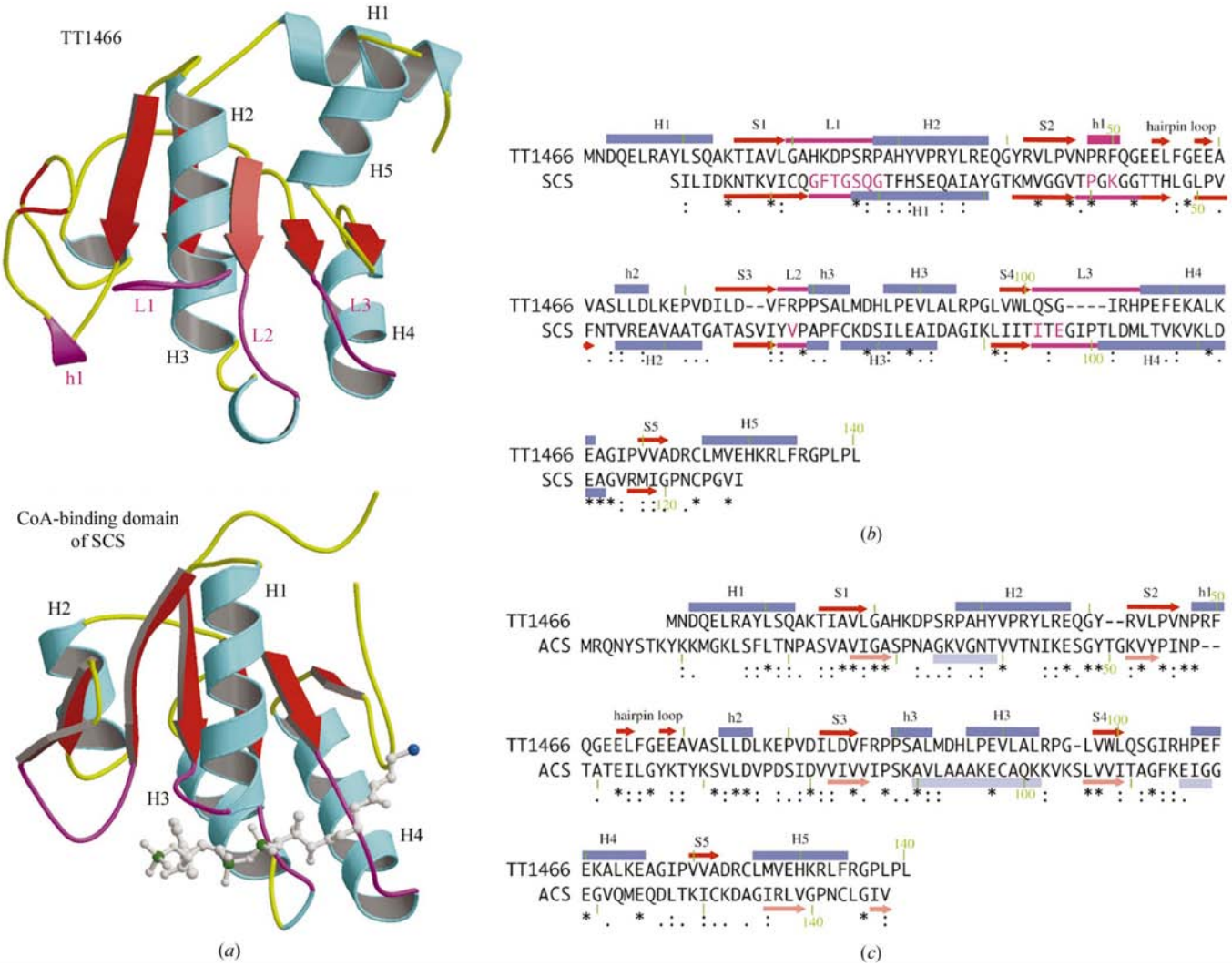

Figure 3
Structural comparisons of TT1466 with the CoA-binding domains of SCS and ACS. (*a*) Ribbon models of TT1466 and the CoA-binding domain of SCS. The α-helices and the β-sheets are colored cyan and red, respectively. The magenta loops in SCS contact the molecule of CoA, shown as a white ball-and-stick model. In the model, the P atoms in the phosphate groups and the S atom in the thiol group are coloured green and blue, respectively. Part of the adenosine 3′-phosphate is located at the left end of the model. Each magenta helix and loop in TT1466 corresponds to the CoA-binding loops of SCS. (*b*) Secondary structures of TT1466 and the CoA-binding domain of SCS, determined with the program *PROCHECK* (Laskowski *et al.*, 1993). The α-helices and the β-sheets are indicated by thick blue bars and red arrows, respectively. The regions involved in CoA binding are shown by magenta bars. Asterisks and dots indicate identical and homologous residues, respectively. Every 10th residue is marked by a green vertical bar. (*c*) Secondary structures of TT1466 and the CoA-binding domain of ACS. Predicted α-helices and β-sheets within ACS (Sanchez *et al.*, 2000) are indicated by light blue bars and light red arrows, respectively.

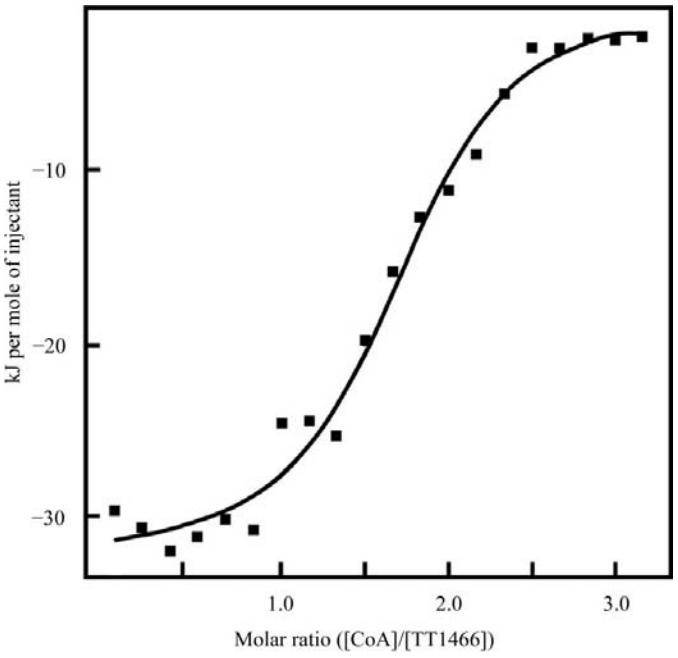

Figure 4
Binding isotherm of CoA for TT1466 at 298 K. Filled boxes show exothermic heat at the molar ratio of CoA per TT1466. The solid line was fitted to the data and provides the ΔH and K values.

We would like to thank K. Tajima for assistance in protein preparation at RIKEN Genomic Sciences Center and Dr S. Adachi for advice on data collection at SPring-8. This work was supported by the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
Brünger, A. T. (1993). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR*. New Haven, Connecticut, USA: Yale University Press.
Buck, D., Spencer, M. E. & Guest, J. R. (1985). *Biochemistry*, **24**, 6245–6252.
Christendat, D. *et al.* (2000). *Nature Struct. Biol.* **7**, 903–909.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Engel, C. & Wierenga, R. (1996). *Curr. Opin. Struct. Biol.* **6**, 790–797.
Fraser, M. E., James, M. N., Bridger, W. A. & Wolodko, W. T. (1999). *J. Mol. Biol.* **285**, 1633–1653.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Holm, L. & Sander, C. (1993). *J. Mol. Biol.* **233**, 123–138.
Horn, J. R., Russell, D., Lewis, E. A. & Murphy, K. P. (2001). *Biochemistry*, **40**, 1774–1778.
Hwang, K. Y., Chung, J. H., Kim, S.-H., Han, Y. S. & Cho, Y. (1999). *Nature Struct. Biol.* **6**, 691–696.
Joyce, M. A., Fraser, M. E., Brownie, E. R., James, M. N., Bridger, W. A. & Wolodko, W. T. (1999). *Biochemistry*, **38**, 7273–7283.
Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. & Yokoyama, S. (1999). *FEBS Lett.* **442**, 15–19.
Kigawa, T., Yamaguchi-Nunokawa, E., Kodama, K., Matsuda, T., Yabuki, T., Matsuda, N., Ishitani, R., Nureki, O. & Yokoyama, S. (2001). *J. Struct. Funct. Genom.* **2**, 27–33.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Madin, K., Sawasaki, T., Ogasawara, T. & Endo, Y. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 559–564.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.
Sanchez, L. B., Galperin, M. Y. & Muller, M. (2000). *J. Biol. Chem.* **275**, 5794–5803.
Saveanu, C., Miron, S., Borza, T., Craescu, C. T., Labesse, G., Gagyi, C., Popescu, A., Schaeffer, F., Namane, A., Laurent-Winter, C., Barzu, O. & Gilles, A. M. (2002). *Protein Sci.* **11**, 2551–2560.
Shin, D. H., Yokota, H., Kim, R. & Kim, S.-H. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 7980–7985.
Teplova, M., Tereshko, V., Sanishvili, R., Joachimiak, A., Bushueva, T., Anderson, W. F. & Egli, M. (2000). *Protein Sci.* **9**, 2557–2566.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Wolodko, W. T., Fraser, M. E., James, M. N. & Bridger, W. A. (1994). *J. Biol. Chem.* **269**, 10883–10890.
Zarembinski, T. I., Hung, L.-W., Mueller-Dieckmann, H.-J., Kim, K.-K., Yokota, H., Kim, R. & Kim, S.-H. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 15189–15193.
Zhang, H., Huang, K., Li, Z., Banerjei, L., Fisher, K. E., Grishin, N. V., Eisenstein, E. & Herzberg, O. (2000). *Proteins*, **40**, 86–97.