Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Refined structure of the histidine-containing phosphotransfer (HPt) domain of the anaerobic sensor kinase ArcB from *Escherichia coli* at 1.57 Å resolution

Masato Kato,[a] Takeshi Mizuno,[b] Toshiyuki Shimizu[a] and Toshio Hakoshima[a]*

[a]Department of Molecular Biology, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan, and [b]School of Agricultural Science, Nagoya University, Chikusaku, Nagoya 464-0814, Japan

Correspondence e-mail: hakoshima@bs.aist-nara.ac.jp

Received 16 February 1999
Accepted 28 July 1999

PDB Reference: HPt domain of ArcB, 2a0b.

The crystal structure of the histidine-containing phosphotransfer (HPt) domain of the anaerobic sensor kinase ArcB from *Escherichia coli* has been refined to 1.57 Å resolution, using the coordinates of the earlier 2.06 Å structure as a starting model. The final model contained 956 protein atoms, one zinc ion and 156 water molecules, with an *R* factor of 19.0%. The high-resolution electron-density maps clearly revealed additional solvent molecules and seven discrete rotamers in the protein side chains. One residue, Met755, was fully buried but was able to occupy the space in the hydrophobic core by means of the two-state conformation of its side chain. One water molecule was buried in the protein core and contributed to the rigidity of the HPt domain, cooperating in the coordination of the zinc ion.

1. Introduction

The two-component intracellular signal-transduction system is widespread in bacteria and is currently also being found in eukaryotes (Parkinson & Kofoid, 1992; Appleby *et al.*, 1996). The system is typically composed of two proteins: a sensor kinase and a response regulator. The sensor kinase contains a histidine protein-kinase (HPK) domain which has a histidine residue that is autophosphorylated with ATP upon stimulation. The response regulator contains a receiver domain which has an aspartate residue to be phosphorylated. It was previously believed that the signal was communicated by a single phosphotransfer reaction from the phospho-histidine residue of the HPK domain to the aspartate residue of the receiver domain, *i.e.* by the His–Asp phosphorelay (Egger *et al.*, 1997).

The *Escherichia coli* sensor kinase ArcB responds to anaerobic conditions by phosphorylating the response regulator ArcA. The phosphorylated ArcA directly or indirectly regulates some 30 operons. These regulations adapt the respiratory system of the cell to anaerobic conditions (Iuchi *et al.*, 1990; Lynch & Lin, 1996). ArcB has four domains: a receptor domain, an HPK domain, a receiver domain and a histidine-containing phosphotransfer (HPt) domain. The HPt domain contains the conserved histidine residue which can be phosphorylated. In this system, a phosphoryl group is transferred to the following domains in the following order: HPK domain, receiver domain, HPt domain and receiver domain of ArcA (Ishige *et al.*, 1994; Tsuzuki *et al.*, 1995; Georgellis *et al.*, 1997). Domains homologous to this HPt domain have been found not only in other prokaryote systems but also in some eukaryote systems, and have been shown to play an important role in the His–Asp–His–Asp phosphorelay (Uhl & Miller,

Table 1
Data-collection and refinement statistics.

Data-collection statistics	
Temperature (K)	277
Crystals	1
Resolution (Å)	1.57
Number of measured reflections	98202
Number of unique reflections	16338
Completeness (outer shell†) (%)	95 (89)
R_{merge}‡ (outer shell†)	4.9 (19)

Refinement statistics	
Resolution range (Å)	6.0–1.57
Number of reflections used	
Total [$F \geq \sigma(F)$]	16036 (94.8%)
Working set	15259
Test set (5% of total)	777
R factor§ (outer shell) (%)	19.0 (27.7, 1.63–1.57 Å)
Free R factor (outer shell) (%)	24.5 (29.9, 1.63–1.57 Å)
Number of residues	118¶
Number of protein atoms	956††
Number of zinc ions	1
Number of solvent molecules	156
Mean B factor (Å^2)	26.2

† Outer shell is 1.70–1.57 Å. ‡ $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum I_i(hkl)$, where $\langle I\rangle$ is the average intensity of the i observations of reflection hkl. § R factor (or free R factor) $= \sum_{hkl} \big||F_{obs}(hkl)| - |F_{calc}(hkl)|\big| / \sum_{hkl} |F_{obs}(hkl)|$, where $F_{obs}(hkl)$ is the observed structure factor and $F_{calc}(hkl)$ is the calculated structure factor from the refined model. ¶ Five residues at the N-terminus and two residues at the C-terminus were not included. †† The model includes seven residues which have discrete rotamers.

1996; Utsumi *et al.*, 1994; Nagasawa *et al.*, 1992; Hrabak & Willis, 1992; Tang *et al.*, 1991; Stevens *et al.*, 1992; Posas *et al.*, 1996; Kehoe & Grossman, 1997). We have previously determined the crystal structure of the HPt domain of ArcB at 2.06 Å resolution and, based on these results, have revealed the structure–function relationship of the HPt domain (Kato *et al.*, 1997). Our findings suggested that this module has a common structural motif and active site in prokaryotes and eukaryotes. In the present study, we further refined the crystal structure to 1.57 Å resolution. On the basis of the refined structure, we show detailed features of the structure of the HPt domain.

2. Materials and methods

2.1. Preparation and crystallization of the protein

E. coli K-12 strain DZ225 carrying plasmid pSU2DH encoding the HPt domain has been constructed previously (Nagasawa *et al.*, 1992; Ishige *et al.*, 1994). Overexpression and purification of the HPt domain were carried out according to the method described previously (Kato *et al.*, 1996). The HPt domain consists of 125 residues (Thr652–Lys776) of ArcB. The purified HPt domain was concentrated to 75 mg ml^{-1} using a Centricon-10 system (Amicon Inc., Beverly, MA) and was stored at 277 K. Protein purity was confirmed by 17.5% sodium dodecylsulfate–polyacrylamide gel electrophoresis, staining with Coomassie Blue G-250. The purified protein was crystallized as described previously (Kato *et al.*, 1996). A maximum crystal size of 0.3 × 0.5 × 0.5 mm was used for the data-collection experiment. The crystals belong to the orthorhombic system, space group $P2_12_12_1$, with unit-cell parameters a = 30.46, b = 34.92, c = 110.74 Å.

2.2. Data collection

X-ray diffraction data were collected at 277 K with an imaging-plate area detector (Rigaku R-AXIS IIc) using Cu $K\alpha$ radiation (λ = 1.54178 Å) generated by a rotaing-anode generator (Rigaku RU-300) operating at 40 kV and 100 mA. The focus size of the X-ray beam was 0.3 mm and the distance from the crystal to the imaging plate was 58.7 mm. Data were collected over a 213° rotation range with 1.5° increments and 40 min exposure per image. Intensities were evaluated with the program *PROCESS* (Rigaku). Diffraction data were collected at 1.57 Å resolution with a completeness of 94.7% and an R_{merge} of 4.9%. The data-collection and processing statistics are summarized in Table 1. A least-squares fit to a Wilson plot (Wilson, 1949) resulted in an estimate of the overall B of 24.3 Å^2.

2.3. Refinement procedure

The starting model was the structure of the HPt domain previously determined at 2.06 Å resolution by use of the MIR method phased with four heavy-atom derivatives (Kato *et al.*, 1997), which was refined using the simulated-annealing method of the program *X-PLOR* (Brünger, 1992) to an R factor of 18.5% in the resolution range 10.0–2.06 Å. Here, the new 1.57 Å data set was used for extension of the resolution and improvement of the structure. In the first step, refinement was carried out with the program *X-PLOR*. In each step, $2F_o - F_c$ and $F_o - F_c$ maps were calculated to improve the structure in the density maps using the program *O* (Jones *et al.*, 1991). The 2.06 Å model without solvent molecules was used as the starting model. The resolution was extended stepwise to 2.0, 1.8, 1.7, 1.6 and 1.57 Å. In the 1.7 Å refinement step, water molecules were temporarily picked out from both electron-density maps contoured at the 3σ level when corresponding electron-density peaks were positioned within hydrogen-bonding distance of the model coordinates. Upon refinement, water molecules were rejected if they moved beyond the hydrogen-bond distance. At the end of refinement using *X-PLOR*, the model contained 118 amino acids, one zinc ion and 134 water molecules, with an R factor of 21.1% and a free R factor of 27.0% in the resolution range 6.0–1.57 Å. This model was further refined using the program *REFMAC* (Murshudov *et al.*, 1997) in the same resolution range. In this refinement, the side chains of seven residues (Ile665, Met667, Ser696, Glu699, Lys716, Met755 and Arg760) were found to be discrete rotamers from inspection of both $2F_o - F_c$ and $F_o - F_c$ maps calculated with σ_A-weighted phases (Read, 1986). Each side chain of these residues was refined with an occupancy of 0.5, which was estimated from electron densities. A total of 156 water molecules were picked out at the temperature-factor limit of 80 Å^2 (the maximum was 75.7 Å^2). They were refined with unit occupancies and variable temperature factors. The root-mean-square (r.m.s.) deviations

of bond distances, angle distances and planar 1–4 distances are 0.024, 0.036 and 0.049 Å, respectively. The refinement statistics are given in Table 1.

3. Results and discussion

3.1. Model quality

The HPt domain used in this study consists of 125 residues (Thr652–Lys776) of ArcB. The numbering of the residues of the HPt domain was based on the recent correction to the SWISS-PROT file (code: P22763). The refined structure of the HPt domain contains 956 protein atoms from 118 residues plus 156 solvent molecules and one zinc ion (Table 1). Seven residues (five residues from the N-terminus and two from the C-terminus) were missing in the $2F_o - F_c$ and $F_o - F_c$ electron-density maps. The N-terminal analyses and mass spectrometry were carried out using crystals dissolved in water. The results showed that there was no deletion of residues at the N-terminus or at the C-terminus (data not shown). Therefore, these residues were probably highly mobile and disordered in the crystals. With the exception of these termini, the electron density was continuous and clearly interpretable for the entire polypeptide backbone. Seven residues (Ile665, Met667, Ser696, Glu699, Lys716, Met755 and Arg760) were modelled as two-state rotational conformers. Six residues (Ser657, Tyr671, Lys689, Gln705, Lys708 and Lys772) had weak densities for their side chains, suggesting that they are generally disordered. The average value of the residue-based real-space correlation (Jones *et al.*, 1991) of the final structure was 0.91.

A Ramachandran plot of the main-chain torsion angles (φ, ψ) produced with the program *PROCHECK* (Ramachandran & Sasisekharan, 1968; Laskowski *et al.*, 1993) showed that all the non-glycine and non-proline residues fell into the most favoured regions (97%) or the additional allowed regions (3%). Three residues (Gly678, Gln707 and Gly727) have positive φ angles and two of these (Gln707 and Gly727) are in α_L conformations. The mean coordinate error was estimated to be 0.17 Å for the working set of reflections and 0.23 Å for the test set, according to the method of Luzzati (1952).

The HPt domains were tightly packed within the unit cell. The HPt domain interacted with 12 symmetrically related molecules, using 54% of the total molecular surface (7300 Å^2). The interactions were mainly made by electrostatic contacts, involving direct or indirect (water-mediated) hydrogen bonds, salt bridges and intermolecular coordination to the zinc ion. In addition, hydrophobic interactions were observed on the surface of helices *B* and *C*. Consequently, the temperature factors of the molecules were rather low; the averaged temperature factors for the main-chain atoms and for the side-chain atoms were 24.0 and 28.4 Å^2, respectively. There were four regions which had main-chain temperature factors above 30 Å^2: residues 657–661 (the N-terminus), residues 676–680, residues 700–707 and residues 771–774 (the C-terminus). The second and third regions correspond to the loops located at either termini of helix *C*. One water molecule (W6) was buried inside the protein molecule (see below). A single zinc ion was identified in the crystal structure of the HPt domain, based on this ion's high electron density and on inspection of its coordination geometry. As mentioned above, it was coordinated between symmetrical molecules. The temperature factor of the zinc ion was 24.1 Å^2. The refinement statistics demonstrate the high quality and good stereochemistry of the final structure.

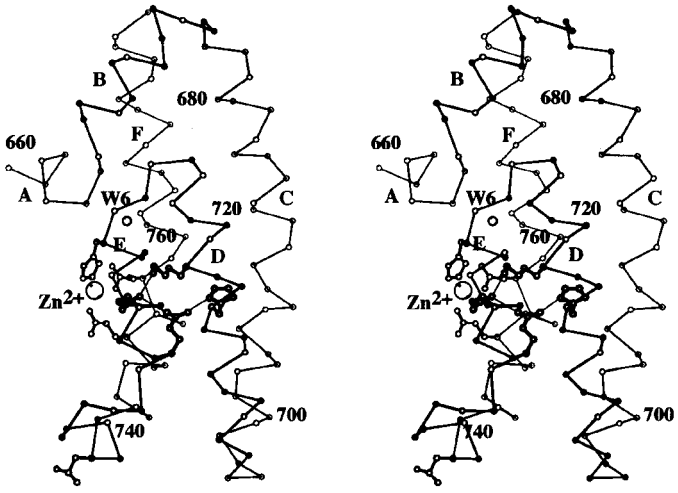

Figure 1
Stereo drawing of the overall structure of the HPt domain of ArcB as the C_α backbone. The side chain of the active residue His715 is drawn as a ball-and-stick model, together with those of Lys718, Gln734 and Gln737 at the active site. Hydrogen bonds are indicated by solid lines. The Zn^{2+} ion is represented by a large sphere and its coordinated residues His728, Asp746, Glu754 and Glu758 are shown. Asp746 of the symmetry-related molecule participated in Zn^{2+} coordination. The small sphere indicates the buried water molecule W6.

3.2. Discrete rotamers and additional solvent molecules

The refined structure of the HPt domain determined at 1.57 Å resolution was compared with the 2.06 Å structure. The HPt domain has a kidney-shaped all-α structure at both resolutions, and the overall folding was also the same in the two structures (Fig. 1). The r.m.s. deviation of the C_α positions between the structures was 0.18 Å. There was little difference

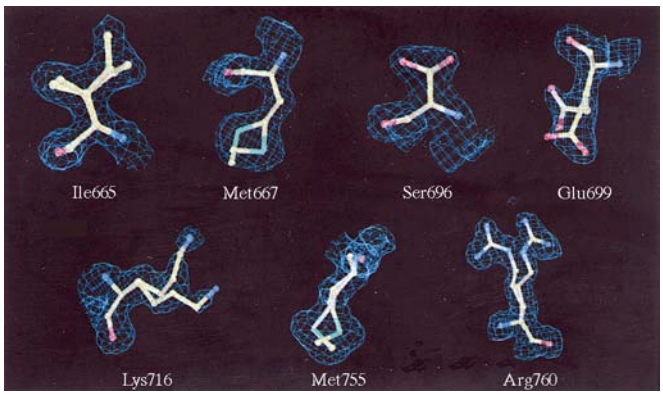

Figure 2
Discrete rotamers with their electron densities. The residues are shown as ball-and-stick models. The densities are contoured at the 1.0σ level.

Table 2
Range and geometry of helices.

Helix ranges and geometry values were calculated with the program *PROMOTIF* (Hutchinson & Thornton, 1996).

Helix	Range	φ (°)	ψ (°)	Length (Å)	Unit rise (Å)	Residues per turn	Pitch (Å)
A	658–662	−73	−37	7.8	1.5	3.7	5.5
B	665–674	−66	−41	15.1	1.5	3.6	5.3
C	677–703	−66	−38	40.7	1.5	5.0	7.5
D	707–723	−63	−41	26.3	1.5	3.7	5.5
E	727–736	−66	−38	15.3	1.5	3.6	5.3
F	744–773	−67	−38	42.5	1.4	5.1	7.2

in the backbone between the two structures. However, large differences in the side-chain conformations were found in several residues. These large differences resulted mainly from the seven discrete residues described above (Fig. 2). The hydrophobic residues Ile665 and Met667 were located at a boundary of the protein core and their side chains could move more freely than those of residues within the protein core. The hydrophilic residues Ser696, Glu699, Lys716 and Arg760 were located on the surface of the molecule and their side chains were exposed to the solvent region. Each rotamer of their side chains formed hydrogen bonds with solvent molecules to stabilize its conformation. Met755 was fully buried in the protein core. However, there was a small space around its side chain. In order to compensate for this space in the core, Met755 had a two-state side-chain conformation. Furthermore, six residues having weak side-chain electron densities as described above were poorly defined and contributed to the large side-chain deviations. Hydrophilic residues on the molecular surface (Glu673, Lys678, Glu688, Glu712 and Glu731) had clear side-chain electron densities, but their side chains showed relatively large differences compared with those of the 2.06 Å structure. The side chains of the 2.06 Å structure were thought to be refined as one of the local minimum conformations.

A large number of the water molecules in the 2.06 Å structure were also found in the 1.57 Å structure. However, high-resolution electron density also revealed that the 1.57 Å structure had 35 additional water molecules not contained in the 2.06 Å structure. While most of these additional water molecules were located in the second hydration shell, some were located in the first hydration shell, forming a hydrogen bond with a hydrophilic side chain. These new water molecules in the first hydration shell had relatively high temperature factors and were unable to be defined at 2.06 Å resolution. 20 water molecules in the 2.06 Å structure had no equivalent molecules in 1.57 Å structure. They also had relatively high temperature factors and were located in the second hydration shell, near the N or C-terminus and near the amino acids which contributed to the large side-chain differences. This means that these 20 water molecule were artefacts of the refinement of the 2.06 Å structure.

3.3. Secondary-structure features

The secondary-structure assignment was carried out with the program *PROMOTIF* (Hutchinson & Thornton, 1996), which uses the modified algorithm of Kabsch & Sander (1983). The range and geometry of the six α-helices are summarized in Table 2. Four of these, *A*, *B*, *D* and *E*, showed a typical α-helix hydrogen-bonding pattern and a normal α-helix geometry. In helix *C*, Pro692 caused a kink of 54°, which is almost twice as large as the average kink (26°) induced by proline (Barlow & Thornton, 1988; Sankararamakrishnan & Vishveshwara, 1993). A water molecule (W1) was attached to this kink and completed the hydrogen-bonding bridge between the main-chain atoms (Fig. 3; Table 3). Helix *F* had two kinks, one located near its centre and the other located at the N-terminus. Two water molecules, W2 and W3, coupled by a hydrogen bond, formed hydrogen-bonding bridges in the central kink. In the N-terminal kink of helix *F*, another two water molecules, W4 and W5, also formed hydrogen-bonding bridges (Fig. 3; Table 3). Because of these kinks, helices *C* and *F* showed irregular averaged helical parameters but normal unit rises. Half of helix *C* and half of helix *F* participated in forming a four-helix bundle with helices *D* and *E*, and the remaining halves formed a hydrophobic interaction with helix *B*. As a result, the hydrophobic core of the bundle extended throughout the whole molecule. Helices *D* and *E* interacted closely with each other at an interhelical distance of 7.5 Å. This value is smaller than the mean value (9.6 Å) for interacting helices in four-helix bundles (Weber & Salemme, 1980; Sheridan *et al.*, 1982). The hydrogen-bonding network, including

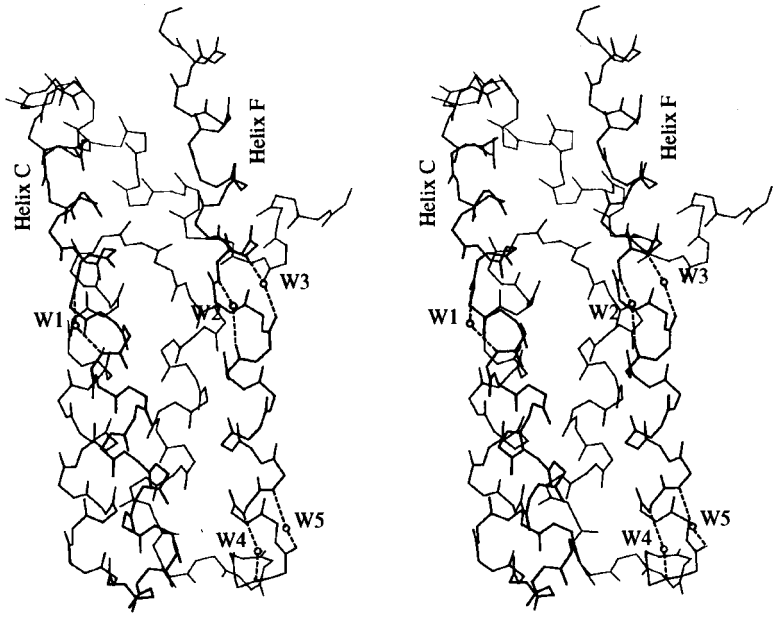

Figure 3
Stereo drawing of water molecules inserted into the kinks in α-helices. The HPt domain is illustrated as a backbone model. The water molecules are shown as spheres. Hydrogen bonds are indicated by dashed lines. The viewing angle is the opposite of that in Fig. 1.

Table 3
Distances of hydrogen bonds and coordinations.

Hydrogen bonds were considered to be present if the distance was <3.45 Å.

Donor	Acceptor	Distance (Å)	Comment
W1 O	Lys689 O	2.8	Central kink of helix *C*
Gly693 N	W1 O	3.0	Central kink of helix *C*
W2 O	Lys756 O	2.9	Central kink of helix *F*
Arg760 N	W2 O	3.0	Central kink of helix *F*
W3 O	Glu757 O	2.7	Central kink of helix *F*
His761 N	W3 O	3.1	Central kink of helix *F*
W4 O	Glu745 O	2.9	N-terminal kink of helix *F*
Gly749 N	W4 O	3.0	N-terminal kink of helix *F*
W5 O	Asp746 O	2.7	N-terminal kink of helix *F*
Glu750 N	W5 O	3.1	N-terminal kink of helix *F*
W6 O	Leu662 O	2.8	Hydrogen-bond cluster of W6
His728 N	W6 O	3.2	Hydrogen-bond cluster of W6
Leu729 N	W6 O	3.2	Hydrogen-bond cluster of W6
W6 O	Asp762 OD2	2.7	Hydrogen-bond cluster of W6
Asp664 N	Gly725 O	2.8	Between T1 and T4
Met667 N	Asp664 OD1	3.3	N-cap of helix *B*
Gly709 N	Asp706 OD1	3.0	N-cap of helix *D*
His715 ND1	Gln737 OE1	2.6	Hydrogen-bond network of active site
Lys718 NZ	Gln734 OE1	2.6	Hydrogen-bond network of active site
Gln737 NE2	Gln734 OE1	3.1	Hydrogen-bond network of active site
Trp744 NE1	Ile736 O	2.9	Both ends of T5
Zn^{2+}	His728 NE2	2.1	Coordinated to zinc ion
Zn^{2+}	Asp746 OD1	2.2	Coordinated to zinc ion
Zn^{2+}	Glu754 OE1	2.1	Coordinated to zinc ion
Zn^{2+}	Glu758 OE2	2.0	Coordinated to zinc ion

the active residue His715 and Lys718 on helix *D* and Gln734 and Gln737 on helix *E*, contributed to this close interaction of helices *D* and *E* (Fig. 1; Table 3).

The five segments connecting the helices were all well ordered. Four short segments (T1–T4) consisted of two or three residues. Loop T1, consisting of Leu663 and Asp664, interacted with loop T4 *via* a hydrogen bond (Table 3). The side chain of Asp664 formed a hydrogen bond with the main chain of Met667 to form the N-terminal cap of helix *B*. The side chain of Asp706 of loop T3 also formed the N-terminal cap of helix *D*. Glu705 and Asp706 of loop T3 formed hydrogen bonds through their main-chain N atoms with Asn701 O of helix *C*. The segment (loop T5) between helices *E* and *F* consisted of seven residues, including two type-I β-turns, of which the second residue was proline. Thus, the interhelical segments contained many intramolecular interactions which stabilize their conformations.

3.4. Zinc-binding site and a buried water molecule

The Zn^{2+} ion formed a coordination sphere located between helices *E* and *F* (Figs. 1 and 4). His728, Glu754, Glu758 and the symmetry-related Asp746 participated in the coordination. In the 2.06 Å structure, both carboxyl O atoms of Glu758 seemed to coordinate to the Zn^{2+} ion, creating the appearance of a variation on pyramidal coordination. However, the high-resolution data revealed more detailed features of the coordination. His728 NE2, Glu754 OE1 and Glu758 OE2 occupied three of the four apices of the tetrahedral coordination (Table 3). The atom occupying the remaining apex was Asp746 OD1 of the symmetry-related molecule. This coordination of Asp746, which is located at the N-terminus of helix *F*, may cause the N-terminal kink in helix *F*. The binding site was remote from the phosphotransfer active site and the zinc binding was not demonstrated to have any effect on this activity. This site seems to be one of the anchors contributing to the tight folding of the four-helix bundle, since the temperature factors of helices *E* and *F* were relatively low compared with those of helices *C* and *D*.

There is a buried water molecule (W6) between the C-terminus of helix *A* and the N-terminus of helix *E* (Figs. 1 and 4). This water molecule also exists in the 2.06 Å structure, although it was not discussed in our previous paper (Kato *et al.*, 1997). It has the lowest temperature factor of the water molecules, owing to its formation of multiple hydrogen bonds with Leu662 O, His728 N, Leu729 N and Asp762 OD2 and to its being tightly fixed (Table 3). The other side-chain OD1 atom of Asp762 formed a hydrogen bond to Glu758, which coordinated to the zinc ion. This hydrogen bond linked the zinc-binding site to the buried water site. These interactions seem to lock the helices *A* and *B* to the four-helix bundle.

3.5. Active site and mutational analyses

The high-resolution structure enables us to identify six additional water molecules surrounding the active-site residue His715. The protein structure itself shows no significant conformational change around His715. These newly detected water molecules, in conjunction with the water molecules previously reported, form the hydrogen-bond network around the active site. However, no water molecules directly make hydrogen bonds with His715. The water molecules could easily move away from His715 when the phosphotransfer reaction takes place, as they have relatively high temperature factors. In addition, one water molecule was identified as making a hydrogen bond with Lys718. His715 forms hydrogen-bond networks with Gln737, Gln734, Lys718 and this water. This network would contribute to the activity and stability of the HPt domain.

Mutational analyses of the HPt domain have recently been performed (Matsushika & Mizuno, 1998*a*,*b*). These results and the high-resolution struc-

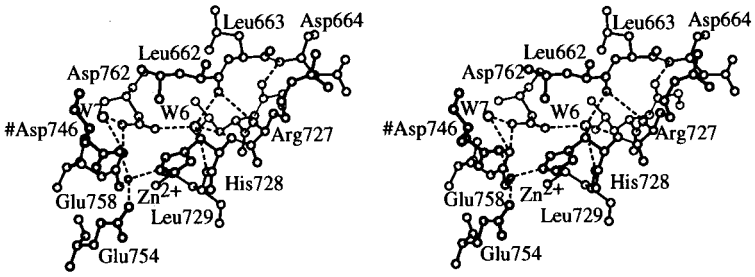

Figure 4
Stereo diagram of the Zn^{2+}-binding site and buried water molecule. All residues are shown as ball-and-stick models, with water molecules (W6 and W7) and the Zn^{2+} ion as spheres. Dashed lines indicate hydrogen bonds or coordinations to Zn^{2+}.

ture reported here allow us to consider the phosphotransfer activity of the HPt domain. There are eight mutants which showed markedly reduced or little activity: H715A, H715Y, G683R, G719D, G722E, K716E, L732P and W744R. The residue numbers are corrected in accordance with the SWISS-PROT file described above.

H715A and H715Y should have no activity because of the absence of the active histidine residue. These mutants were found to show no activity.

The G719D and G722E mutants exhibited an approximately fivefold lower activity than the wild type. Gly719 and Gly722 are located near His715 on helix *D* (Fig. 5). These mutations might loosen the interaction of the HPt domain with the cognate receiver domain ArcA. Gly719 is located on the adjacent ridge of helix *D* with His715 so as to carve out an empty space for the histidine residue. If an amino acid with a large side chain exists at this point, the ability of the histidine residue to accept or to transfer a phosphoryl group would be disturbed.

G683R showed little activity. Gly683 is located in a deep hole surrounded by the hydrophobic patch on the molecular surface between helix *C* and the C-terminal of helix *D* (Figs. 5 and 6). This hydrophobic hole, which is formed by highly conserved residues in HPt domains and corresponds to residues 683, 684, 687, 720 and 724 of the the HPt domain of ArcB (Kato *et al.*, 1997), could accommodate a hydrophobic residue. Interestingly, the side chain of Leu674 of a symmetrical molecule was inserted into this hole in the crystal. The feature suggests that this hole might be an important site for molecular recognition. The G683R mutant loses its activity because the large side chain of arginine can cap the hole and interfere with the insertion of a hydrophobic residue of a receiver domain.

The mutants at Lys716, Leu732 and Trp744 were very unstable (Matsushika & Mizuno, 1998*a*), which resulted in the observation of no activity. Leu732 and Trp744 are located in the protein core and are important in protein folding. In particular, the side chain of Trp744, which was located at the N-terminus of helix *F*, formed a hydrogen bond with the main chain of Ile736, which was located at the C-terminus of helix *E* (Table 3). The longest loop in the HPt domain, T5, was stabilized by this hydrogen bond. Substitution of these residues could induce the collapse of the folding. In the crystal structure, Lys716 had two-state side chain; one conformation made a hydrogen bond with Glu713 and the other made a hydrophobic interaction between its methylene group and the aromatic ring of Tyr694. This might affect the protein folding.

The other mutants, L679A, D682A, D706A and E712K, exhibited a twofold lower phosphotransfer activity of the HPt domain (Fig. 5). Leu679 was positioned in the N-terminal of helix *C* and its side chain caps the protein core between helix *B* and the four-helix bundle. The absence of this side chain might allow solvent molecules to access the protein core and

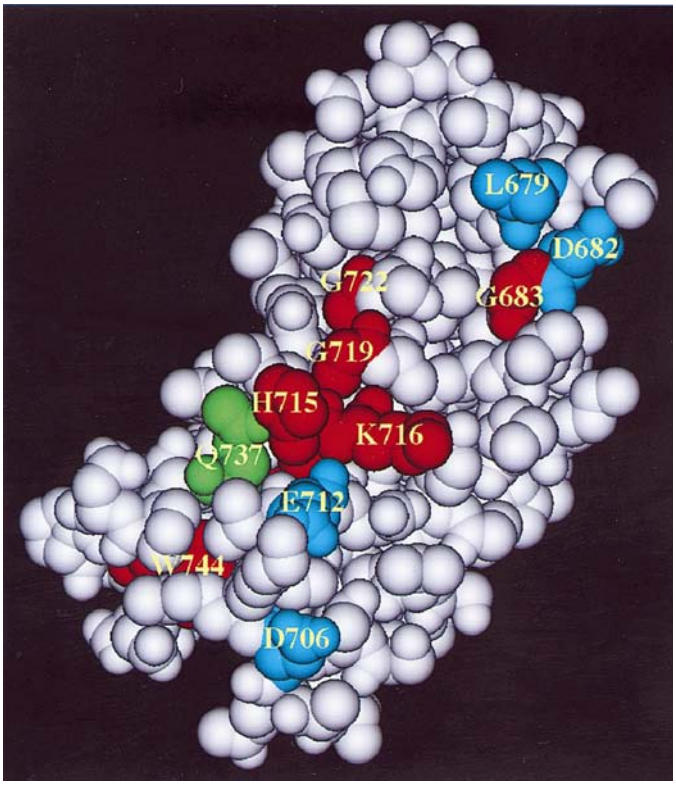

Figure 5
Space-filling model of the HPt domain. The amino acids changed in the mutational analyses are coloured red for those mutants which have no activity, cyan for those with decreased activity and green for those having normal activity. Leu732 is fully buried in the protein core and cannot be seen. It is located under G719.

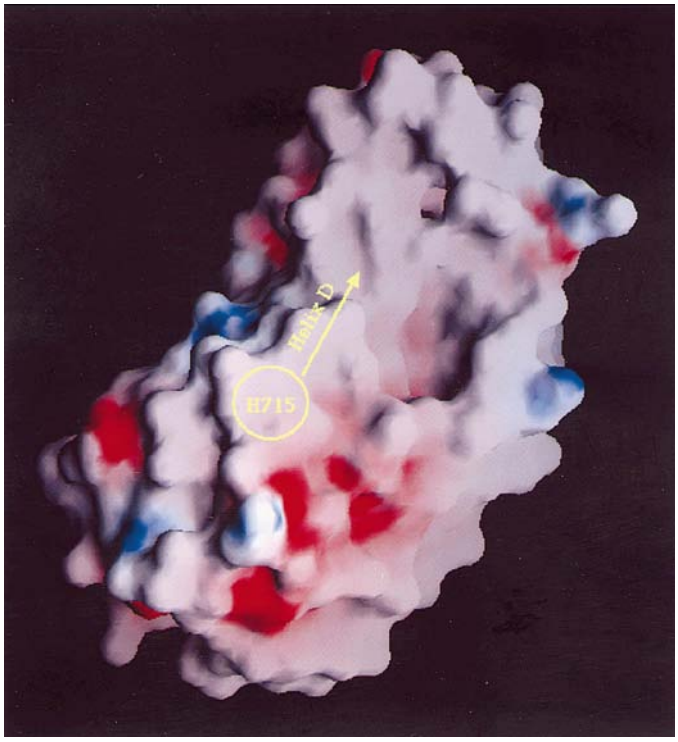

Figure 6
Electrostatic potential surface of the HPt domain. The viewing angle is the same as in Fig. 5. The electrostatic potential is coloured red for negative charge and blue for positive charge with a range of $-20\ k_BT$ to $+20\ k_BT$; k_B and T are the Boltzmann constant and temperature, respectively. The active residue His715 and helix *D* are labelled.

induce changes in the HPt domain folding. The side chains of Asp682 and Glu712 form hydrogen bonds with Lys678 and Lys708, respectively. Highly conserved Asp706 made a hydrogen bond with the main chain of Gly709 to form the N-terminal cap of helix *D* (Table 3). Therefore, the substitution of these residues removes the interaction and makes the structure unstable. Interestingly, *E. coli* hybrid sensors BarA, EvgS and TorS maintain hydrophilic residues at the positions corresponding to Lys678, Asp682, Lys708 and Glu712 of the ArcB HPt domain (Matsushika & Mizuno, 1998*a*). In these sensor proteins, hydrogen bonds could be formed in a similar manner to ArcB.

Gln737 makes a hydrogen bond with the side chain of the active residue His715 and fixes the imidazole ring so that it is exposed to the solvent region (Fig. 1). Surprisingly, however, the mutant Q737A had the same activity as the wild type. These residues also form a hydrogen-bond network with Lys718 and Gln734. In the mutant Q737A, the hydrogen bond between Lys718 in helix *D* and Gln734 in helix *E* still remained and could stabilize the close interaction between helices *D* and *E*. This indicated that the hydrogen bond between His715 and Gln737 may not be important for activity and that the close packing of helices *D* and *E* contribute most to the activity.

The electrostatic potential surface showed no significant feature on the concave surface around the active site (Fig. 6). Similarly, the chemotaxis response regulator CheY, which is able to be phosphorylated by the HPt domain (Yaku *et al.*, 1997), showed no significant electrostatic potential feature on the surface around the active pocket. On the other hand, we observed the hydrophobic hole, which was surrounded by the conserved residues as described previously (Fig. 6). CheY also has a hydrophobic patch near its active pocket. This suggests that the hole is one of the contact sites. The complex structure of the histidine-containing phosphocarrier protein (HPr) and the N-terminal domain of enzyme I (EIN) reveals that the majority of the contacts are hydrophobic (Garrett *et al.*, 1999). Some specific electrostatic interactions also contribute to the contacts between the two proteins. It is difficult to rule out specific electrostatic interactions between the HPt domain and the receiver domain, but hydrophobic interactions may be important for the interaction as in the HPr–EIN complex. There is a negatively charged patch on the opposite side where the zinc-binding site is positioned; the zinc ion neutralizes the negative charge. This site is far from the active site, and the relationship between the active site and the zinc-binding site is unclear, as mentioned previously. The structure of the complex between the HPt domain and the chemotaxis response regulator CheY must be obtained for further discussion, although we have recently determined the non-binding mode of the two proteins (Kato *et al.*, 1998, 1999).

4. Conclusions

The high-resolution structure provides detailed information about the HPt domain. Together with the hydrophobic core extending through the whole molecule, the wealth of additional intramolecular hydrogen-bonding, ion-pairing and zinc-coordination interactions with intra- and inter-secondary elements endowed the HPt domain with rigidity. Indeed, the rigidity of the HPt domain was previously confirmed in the crystal structure of the HPt domain complexed with CheY (Kato *et al.*, 1998, 1999). The rigidity of the HPt domain demonstrated here may be important in the development of the rapid phosphotransfer activity. The active residue His715 is found to be hydrated, and one water molecule can be detected which participates in the hydrogen-bond networks around the active site. We can clarify the relationship between the mutational analyses and the roles of the residues which caused a loss of or a decrease in the phosphotransfer activity of the HPt domain. The feature of the electrostatic potential surface of the HPt domain suggests that the majority of the contacts are hydrophobic. Moreover, a hydrophobic hole was found which could participate in the interaction with the receiver domain. In order to confirm these surface features, we require the structure of the complex between the HPt domain and the receiver domain, which will show the phosphotransfer reaction mode.

This work was supported by Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan to TM (06276105) and TH (06276104, 08249225). MK was supported by a research fellowship from the Japan Society for the Promotion of Science.

References

Appleby, J. L., Parkinson, J. S. & Bourret, R. B. (1996). *Cell*, **86**, 845–848.
Barlow, D. J. & Thornton, J. M. (1988). *J. Mol. Biol.* **201**, 601–619.
Brünger, A. T. (1992). *X-PLOR. A System for Crystallography and NMR.* Yale University Press, New Haven, CT, USA.
Egger, A. L., Park, H. & Inouye, M. (1997). *Genes Cells*, **2**, 167–184.
Garrett, D. S., Seok, Y.-J., Peterkofsky, A., Gronenborn, A. M. & Clore, G. M. (1999). *Nature Struct. Biol.* **6**, 166–173.
Georgellis, D., Lynch, A. S. & Lin, E. C. C. (1997). *J. Bacteriol.* **179**, 5429–5435.
Hrabak, E. M. & Willis, D. K. (1992). *J. Bacteriol.* **174**, 3011–3020.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Ishige, K., Nagasawa, S., Tokishita, S. & Mizuno, T. (1994). *EMBO J.* **13**, 5195–5202.
Iuchi, S., Matsuda, Z., Fujiwara, T. & Lin, E. C. C. (1990). *Mol. Microbiol.* **4**, 715–727.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Kato, M., Ishige, K., Mizuno, T., Shimizu, T. & Hakoshima, T. (1996). *Acta Cryst.* D**52**, 1214–1215.
Kato, M., Mizuno, T. & Hakoshima, T. (1998). *Acta Cryst.* D**54**, 140–142.
Kato, M., Mizuno, T., Shimizu, T. & Hakoshima, T. (1997). *Cell*, **88**, 717–723.
Kato, M., Shimizu, T., Mizuno, T. & Hakoshima, T. (1999). *Acta Cryst.* D**55**, 1257–1263.
Kehoe, M. D. & Grossman, R. A. (1997). *J. Bacteriol.* **179**, 3914–3921.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Luzzati, P. V. (1952). *Acta Cryst.* **5**, 802–810.

Lynch, A. S. & Lin, E. C. C. (1996). *J. Bacteriol.* **178**, 6238–6249.
Matsushika, A. & Mizuno, T. (1998*a*). *J. Biochem.* **124**, 440–445.
Matsushika, A. & Mizuno, T. (1998*b*). *Biosci. Biotechnol. Biochem.* **62**, 2236–2238.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nagasawa, S., Tokishita, S., Aiba, H. & Mizuno, T. (1992). *Mol. Microbiol.* **6**, 799–807.
Parkinson, J. S. & Kofoid, E. C. (1992). *Annu. Rev. Genet.* **26**, 71–112.
Posas, F., Wurgler-Murphy, M. S., Maeda, T., Witten, A. E., Thai, C. T. & Saito, H. (1996). *Cell*, **86**, 865–875.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein. Chem.* **23**, 283–438.
Read, R. (1986). *Acta Cryst.* A**42**, 140–149.
Sankararamakrishnan, R. & Vishveshwara, S. (1993). *Proteins*, **15**, 26–41.
Sheridan, R. P., Levy, R. M. & Salemme, F. R. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 4545–4549.
Stevens, A. M., Sanders, J. M., Shoemaker, N. B. & Salyers, A. A. (1992). *J. Bacteriol.* **174**, 2935–2942.
Tang, J. L., Liu, Y. N., Barber, C. E., Dow, J. M., Wootton, J. C. & Daniels, M. J. (1991). *Mol. Gen. Genet.* **226**, 409–417.
Tsuzuki, M., Ishige, K. & Mizuno, T. (1995). *Mol. Microbiol.* **18**, 953–962.
Uhl, A. M. & Miller, F. J. (1996). *J. Biol. Chem.* **271**, 33176–33180.
Utsumi, R., Katayama, S., Taniguchi, M., Horie, T., Ikeda, M., Igaki, S., Nakagawa, H., Miwa, A., Tanabe, H. & Noda, M. (1994). *Gene*, **140**, 73–77.
Weber, P. C. & Salemme, F. R. (1980). *Nature (London)*, **287**, 82–84.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Yaku, H., Kato, M., Hakoshima, T., Tsuzuki, M. & Mizuno, T. (1997). *FEBS Lett.* **408**, 337–340.