Article No. jmbi.1999.3215 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 293, 753–761

JMB

COMMUNICATION

Insights into Eukaryotic Multistep Phosphorelay Signal Transduction Revealed by the Crystal Structure of Ypd1p from *Saccharomyces cerevisiae*

Hyun Kyu Song†, Jae Young Lee†, Myong Gyong Lee†, Jinho Moon Kyeongsik Min, Jin Kuk Yang and Se Won Suh*

Department of Chemistry
College of Natural Sciences
Seoul National University
Seoul, 151-742, Korea

''Two-component'' phosphorelay signal transduction systems constitute a potential target for antibacterial and antifungal agents, since they are found exclusively in prokaryotes and lower eukaryotes (yeast, fungi, slime mold, and plants) but not in mammalian organisms. *Saccharomyces cerevisiae* Ypd1p, a key intermediate in the osmosensing multistep phosphorelay signal transduction, catalyzes the phosphoryl group transfer between response regulators. Its 1.8 Å structure, representing the first example of a eukaryotic phosphorelay protein, contains a four-helix bundle as in the HPt domain of *Escherichia coli* ArcB sensor kinase. However, Ypd1p has a 44-residue insertion between the last two helices of the helix bundle. The side-chain of His64, the site of phosphorylation, protrudes into the solvent. The structural resemblance between Ypd1p and ArcB HPt domain suggests that both prokaryotes and lower eukaryotes utilize the same basic protein fold for phosphorelay signal transduction. This study sheds light on the best characterized eukaryotic phosphorelay system.

Keywords: crystal structure; phosphorelay phosphotransferase; *Saccharomyces cerevisiae*; signal transduction; Ypd1p

**Corresponding author*

In order to sense changes in the environment and to respond appropriately a signal transduction strategy known as the ''two-component'' regulatory systems or His-Asp phosphorelay systems is commonly employed by prokaryotic organisms (Hoch & Silhavy, 1995; Eggar *et al*., 1997; Mizuno, 1998). This signaling machinery, ubiquitous in bacteria, consists of two basic units: a sensor histidine kinase and a response regulator. The more elaborate His-Asp-His-Asp four-step phosphorelay signal transduction pathways have also been identified such as the KinA/KinB/KinC-Spo0F-Spo0B-Spo0A system in *Bacillus subtilis* (Burbulys *et al*., 1991). The ''two-component'' phosphorelay systems also operate in lower eukaryotic organisms, including yeast, fungi, plants, and the slime mold *Dictyostelium* (Swanson & Simon, 1994; Chang & Meyerowitz, 1994; Swanson *et al*., 1994; Loomis *et al*., 1997; Wurgler-Murphy & Saito, 1997). However, no such system has been detected in mammalian organisms and this makes the phosphorelay signal transduction an attractive target for antibacterial and antifungal agents (Barrett *et al*., 1998).

The best characterized eukaryotic phosphorelay system is the His-Asp-His-Asp multistep phosphorelay that governs osmoregulation in the budding yeast *Saccharomyces cerevisiae* (Posas *et al*., 1996). It involves the four proteins Sln1p, Ypd1p, Ssk1p, and Skn7p. Sln1p is a hybrid or unorthodox histidine kinase that possesses both a conserved histidine kinase domain (H1 module) and an aspartate-containing receiver domain (D1 module) of response regulators (Ota & Varshavsky, 1993). Under normal growth conditions, it is active and autophosphorylates His576, from where the phosphate group is relayed to Asp1114. The phosphate group is then transferred to His64 of the phosphorelay intermediate protein Ypd1p (H2 module) and

†These authors contributed equally to this work.

Abbreviations used: HPt, histidine-containing phosphotransfer; MIR, multiple isomorphous replacement; NCS, non-crystallographic symmetry; rms, root-mean-square.

E-mail address of the corresponding author: sewonsuh@snu.ac.kr

0022-2836/99/440753–9 $30.00/0

passed along to Asp554 of the Ssk1p response regulator (D2 module). The phosphotransferase Ypd1p thus binds to the receiver domains of both Sln1p and Ssk1p to mediate the multistep phosphorelay. Phosphorylation of the C-terminal receiver domain in the Ssk1p response regulator renders it unable to activate the MAP kinase cascade, which is composed of Ssk2p/Ssk22p MAPKKKs, Pbs2p MAPKK, and Hog1p MAPK (Maeda *et al*., 1995; Boguslawski, 1992; Brewster *et al*., 1993). Specific phosphatases that act at the level of MAP kinase cascade have also been identified (Eggar *et al*., 1997). Under conditions of high osmolarity, the histidine kinase Sln1p is no longer active and the unphosphorylated Ssk1p leads to an enhanced level of intracellular glycerol so that the cells can continue to grow in the hyperosmotic environment. Sln1p-regulated osmotic-stress pathway is bifurcated (Ketela *et al*., 1998; Li *et al*., 1998), with Sln1p kinase transmitting signals *via* Ypd1p not only to the Ssk1p response regulator but also to Skn7p (Asp427). Therefore, Ypd1p plays a key role in this increasingly complex signal transduction pathway. In *S. cerevisiae*, there is also a second high osmolarity sensor, Sho1p, which is non-essential and is not a histidine kinase. Instead, it has an SH3 domain for interacting with the N-terminal proline-rich domain of Pbs2p MAPKK (Maeda *et al*., 1995).

Relatively little structural information is available on eukaryotic phosphorelay proteins. Therefore, we have determined the crystal structure of Ypd1p, a key intermediate in the *S. cerevisiae* osmosensing multistep phosphorelay signal transduction. Ypd1p, a 167-residue protein, is an interesting target for structural studies, since it interacts with at least three different receiver domains. We report here its crystal structure determined at 1.8 Å resolution. The structure shares a four-helix bundle with *Escherichia coli* ArcB HPt domain despite little overall sequence identity (Kato *et al*., 1997).

Model assessment

The structure of Ypd1p fused with a His_6 tag at its C terminus, was initially determined by multiple isomorphous replacement (MIR) using two heavy-atom derivatives of the triclinic crystals and was refined against 2.3 Å native2 data (Table 1). The model accounts for 652 amino acid residues in four copies of the His-tagged Ypd1p as well as 176 ordered solvent molecules in the crystallographic asymmetric unit. A four-residue segment (Gln-Val-Asp-Asp at positions 127-130) as well as N-terminal Met1 and five residues of the C-terminal His_6 tag were not visible in the electron density map.

Subsequently, intact Ypd1p without a C-terminal His_6 tag was crystallized into a tetragonal space group. And its structure was solved by molecular replacement and refined against 1.8 Å native3 data (Table 1). The model accounts for 332 amino acid residues in two copies of the intact Ypd1p as well as 56 ordered solvent molecules in the crystallographic asymmetric unit. Only Met1 was not visible in the electron density map for both copies of Ypd1p in the asymmetric unit. The 127-130 segment, invisible in the triclinic crystal, could be unambiguously traced. However, it shows a high average *B*-factor of 66.6 $Å^2$ and displays a large root-mean-square (rms) deviation between the two copies in the asymmetric unit (1.1 Å for 32 atom pairs). In comparison, a superposition of the two non-crystallographic symmetry (NCS)-related molecules gives an rms deviation of 0.6 Å for all 1339 atom pairs. The two molecules in the asymmetric unit show overall average *B*-factors of 26.5 and 29.2 $Å^2$, respectively. The Ypd1p model with a lower average *B*-factor is used for the following discussion, unless otherwise stated.

Overall structure

The yeast Ypd1p folds into a liver-shaped, compact domain with approximate dimensions of 60 Å × 32 Å × 26 Å. Its structure contains six α-helices and two 3_{10}-helices but no β-strands (Figure 1(a)). The α-helices B, C, D, and F form a four-helix bundle. The helices B, C, and D are amphiphilic as well as the C-terminal half of helix F. The N-terminal half of helix F is rich in hydrophobic residues and is surrounded by helices B, C, and D on one side and by irregular loops on the other side (Figure 1(a)). The longest helix F (43 Å long) is bent by about 26° at Arg149 and another long helix B (38 Å long) by about 14° at Asp37.

Two short α-helices A and E, two 3_{10}-helices G1 and G2, and connecting loops are the extra structural elements besides the four-helix bundle. The connection between helices B and C, and that between helices C and D are short (4 and 2 residues, respectively). Most linker regions between the two adjacent helices contain β-turns. However, the helices C and D are connected by an inverse γ-turn (Figure 1(a)), which is formed by the residues Gly-Leu-Gln at positions 74-76, with the C^{α} distance of 5.4 Å between Gly74 and Gln76. The hydrophobic residues (Leu73 and Leu75 in Ypd1p) and the glycine residue (Gly74) residing in the linker between helices C and D are nearly invariant in other homologous proteins (Kato *et al*., 1997). The helices D and F are connected by a 44-residue insertion (residues 93-136), which contains a one-turn α-helix E and a 3_{10}-helix G2 (Figure 1). The interaction between this mini-insertion domain and the four-helix bundle is mediated, in part, by the N-terminal loop region (residues 2-10). A highly negatively charged surface patch at the top of the molecule is formed by clustering of the residues, Asp22, Asp23, Asp24, and Asp26, in the connecting loop between the 3_{10}-helix G1 and α-helix B (Figure 1(c)). This may serve an important function such as protein-protein interaction but its exact role remains to be established.

Table 1. Data collection, phasing, and refinement statistics for Ypd1p

Data set	Native1	Native2	Native3	Thiomersal	SeMet
A. *Data collection statistics*					
Space group	*P*1	*P*1	$P4_32_12$	*P*1	*P*1
Resolution range (Å)	20-2.9	50-2.3	50-1.8	20-3.3	20-3.6
Unique reflections	15,796	35,254	33,318	12,693	9,567
Redundancy	1.9	1.7	9.6	3.5	1.8
Completeness (%)[a]	89.2 (48.6)	88.3 (85.9)	97.6 (94.0)	95.4 (94.6)	93.4 (92.6)
R_{merge}[b] (%)	6.4	4.4	4.8	7.6	9.8
B. *Phasing statistics*					
R_{iso}[c] (%)				26.0	12.5
Phasing power[d]				2.28	0.79
R_{Cullis}[e] (anomalous)				0.58 (0.79)	0.88
Heavy-atom sites				4	12
Figure of merit	0.54 for 12-3.3 Å data (before DM)/0.79 for 20-3.0 Å data (after DM)				
C. *Refinement and model statistics*					
R-factor/R_{free}[f]		0.206/0.248	0.201/0.247		
Resolution range (Å)		8.0-2.3	8.0-1.8		
No. of residues		163 × 4	166 × 2		
r.m.s. deviation from ideality					
Bond lengths (Å)		0.008	0.008		
Bond angles (deg.)		0.94	0.89		
Average *B*-factor (Å^2)					
Main-chain		39.8	24.8		
Side-chain		43.3	30.7		
Solvent		49.3	31.7		
Ramachandran plot (%)					
Most favorable region (outlier)		92.2 (0)	95.5 (0)		

Expression, purification, crystallization, and X-ray data collection of Ypd1p fused with a His_6 tag at the C terminus were performed as described (Lee *et al.*, 1999). Crystals belong to the triclinic space group *P*1 with the cell parameters of $a = 66.34$ Å, $b = 66.48$ Å, $c = 66.49$ Å, $\alpha = 106.4°$, $\beta = 106.7°$, and $\gamma = 115.3°$. The crystallographic asymmetric unit contains four monomers of the recombinant Ypd1p. Subsequently, Ypd1p without a C-terminal His_6 tag was overexpressed in soluble form in C41(DE3) cells. Under exactly the same crystallization condition used to grow triclinic crystals of His_6 tagged Ypd1p, the intact protein unexpectedly crystallized into the tetragonal space group $P4_32_12$ (determined on the basis of translation function calculations) with unit cell parameters of $a = 53.01$ Å and $c = 244.9$ Å. Apparently isomorphous tetragonal crystals have been obtained under a different crystallization condition (Xu *et al.*, 1999). Native1 and heavy-atom derivative X-ray diffraction data were collected using CuKα radiation from a rotating anode generator. Native2 and native3 data have been collected at 20°C from triclinic and tetragonal crystals using 1.000 Å synchrotron radiation. Two large image plates (400 mm × 800 mm, Fuji BASIII) were placed at a distance of 573 mm from the crystal on a Weissenberg camera for macromolecular crystallography at the BL-6B experimental station of Photon Factory, Japan (Sakabe *et al.*, 1997). The diffraction patterns recorded on the image plates were digitized by the off-line scanner IPR4080. The data were processed and scaled using the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997). Mercury atoms were located by interpreting the difference Patterson maps with SHELX-97 (Sheldrick & Schneider, 1997). Selenium sites were located in the cross-phase difference Fourier maps. Heavy-atom parameters were refined with MLPHARE (CCP4, 1994). Initial MIR phases were improved and extended by density modification consisting of solvent-flattening, histogram-matching, and 4-fold NCS averaging with DM (CCP4, 1994). The 4-fold NCS matrices were determined using the mercury coordinates and by building a partial model of four NCS-related helices. They were further improved by IMP (Jones, 1992). The characteristic four-helix bundle was clearly visible in the resulting electron density map, which allowed fitting ~70% of the residues in a monomer using O (Jones *et al.*, 1991). The NCS relationship was applied to a model of one monomer to generate three other monomers in the asymmetric unit. Phases computed from the partially refined models were combined with the experimental MIR phases in each cycle of refinement using SIGMAA (CCP4, 1994). The SIGMAA-weighted $(2mF_o - DF_c)$ map in each refinement cycle showed a considerable improvement. The structure was refined with X-PLOR (Brünger, 1993), initially against 2.9 Å native1 data and later against 2.3 Å native2 data. An exact NCS relationship between the four independent monomers in the asymmetric unit of the triclinic crystal was maintained throughout all stages of the refinement. Subsequently, 1.8 Å X-ray diffraction data from tetragonal crystals of intact Ypd1p became available and its structure was solved by molecular replacement method using AMoRe (Navaza, 1994). The refinement against 1.8 Å tetragonal crystal data was performed similarly, except that the exact NCS relationship between the two independent monomers in the asymmetric unit that was maintained during the early stages of the refinement was relaxed in the final rounds.

[a] Values are for the reflections in the highest resolution bin of 0.1 Å thickness.

[b] $R_{merge} = \Sigma_h\Sigma_i|I(h,i) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I(h,i)$, where $I(h,i)$ is the intensity of the *i*th measurement of reflection *h* and $\langle I(h)\rangle$ is the average value over multiple measurements.

[c] $R_{iso} = \Sigma||F_{PH}| - |F_P||/\Sigma|F_P|$, where F_{PH} and F_P are the derivative and native structure factors, respectively.

[d] Phasing power $= \langle F_H\rangle/E$, where $\langle F_H\rangle$ is the r.m.s. heavy-atom structure factor and *E* is the residual lack of closure error.

[e] $R_{Cullis} = \Sigma||F_{PH} \pm F_P| - |F_{PH(calc)}||/\Sigma|F_{PH} - F_P|$, where F_{PH} and $F_{PH(calc)}$ are the observed and calculated structure factors of a heavy-atom derivative.

[f] $R = \Sigma||F_O| - |F_C||/\Sigma|F_O|$, where R_{free} is calculated for the 5% test set of reflections.

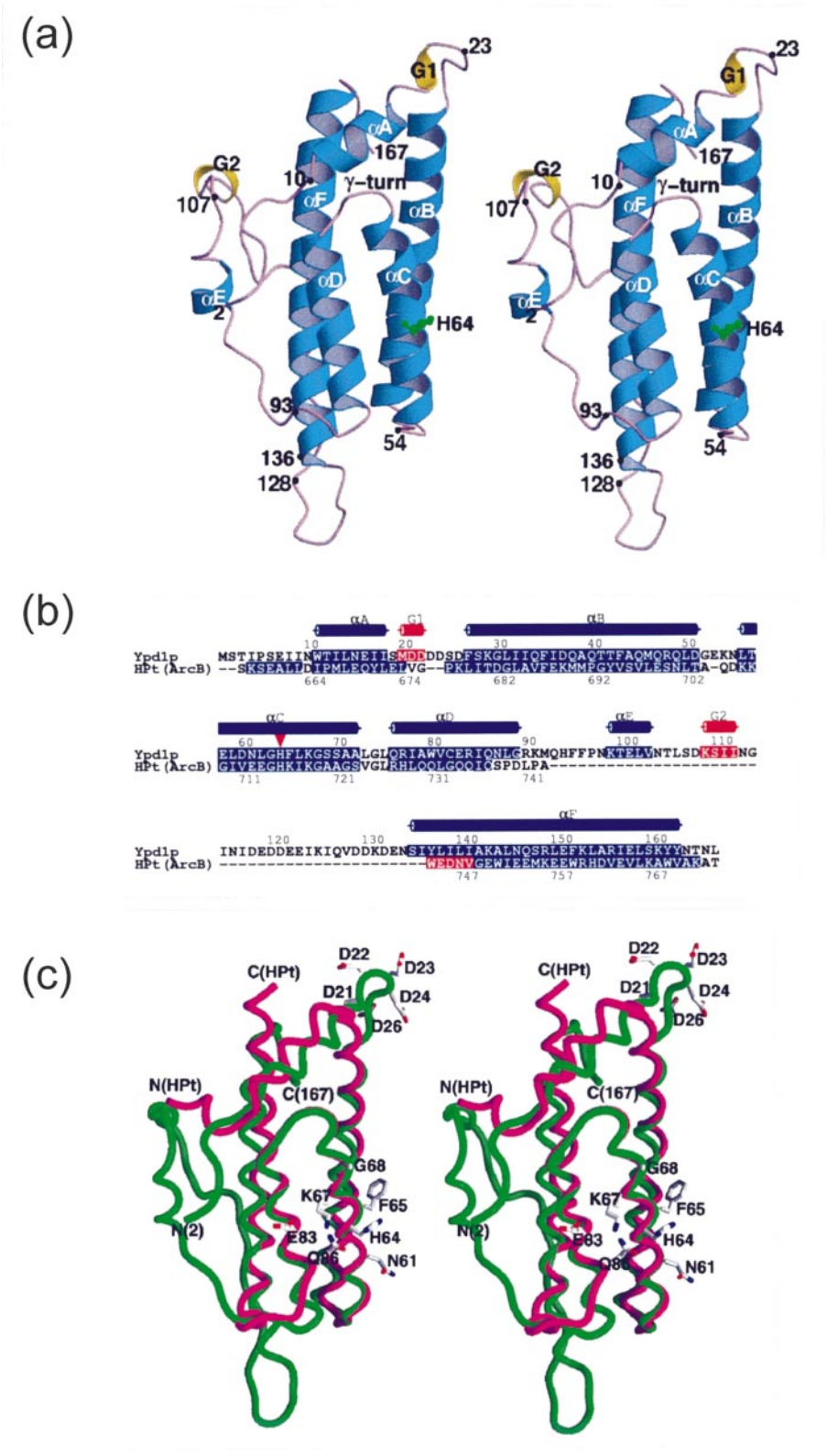

Figure 1 (*legend opposite*)

Structural comparisons and insights into eukaryotic multistep phosphorelay

When the three-dimensional structure of Ypd1p was compared with those in the database of DALI (Holm & Sander, 1993), a notable structural similarity (Z score >2.0) was found in a total of 120 proteins including many four-helix bundle proteins. Proteins showing highest structural similarity are the HPt domain of ArcB ($Z = 12.9$), cytochrome c' ($Z = 8.0$), cytochrome b_{562} ($Z = 7.4$), apolipoprotein E ($Z = 5.7$), and aspartate receptor ($Z = 5.6$). The P1 domain of CheA, although not included in the DALI database, shows also a notable structural similarity. The helix bundles of these proteins share an identical topology with that of Ypd1p. The following comparisons will be limited to the functionally related structures: the HPt domain of *E. coli* ArcB, the P1 domain of *E. coli* CheA, and *B. subtilis* Spo0B phosphotransferase (Figure 2).

The structure of Ypd1p is most similar to that of the HPt domain of *E. coli* ArcB despite very low overall sequence identity (Figure 1(c)). It indicates that the sequence alignment proposed by Kato *et al.* (1997) is not valid for the C-terminal part of Ypd1p due to the above-mentioned 44-residue insertion between helices D and F. A structure-based sequence alignment of the two proteins (Figure 1(b)) indicates that only 13 of 167 residues of Ypd1p are conserved (an overall sequence identity of 7.8 %). Eight of 13 identical residues of Ypd1p belong to the α-helical hairpin formed by α-helices C and D, and the connecting γ-turn. When the sequence alignment is limited to this region only (residues 56-99), the identity is increased to 23.5 %. In fact, the helix C with His64 (the site of phosphorylation), the inverse γ-turn, and the helix D of Ypd1p are structurally the most conserved region. A superposition of structurally equivalent residues in this α-helical hairpin of Ypd1p and the corresponding residues in the HPt domain of ArcB yields an rms deviation of 0.5 Å for 34 C^{α} atom pairs (excluding the last two residues of helix D, which show large deviations). Many of these residues are presumably crucial for the binding of phosphotransferases to their cognate receiver domains. This comparison suggests that the α-helical hairpin of Ypd1p formed by helices C, D and the connecting γ-turn or equivalent α-helical hairpins in other homologous proteins represent a minimal structural unit necessary for the phosphorelay phosphotransferase activity. Other structural elements are presumed to be required for the structural integrity and/or for other functions.

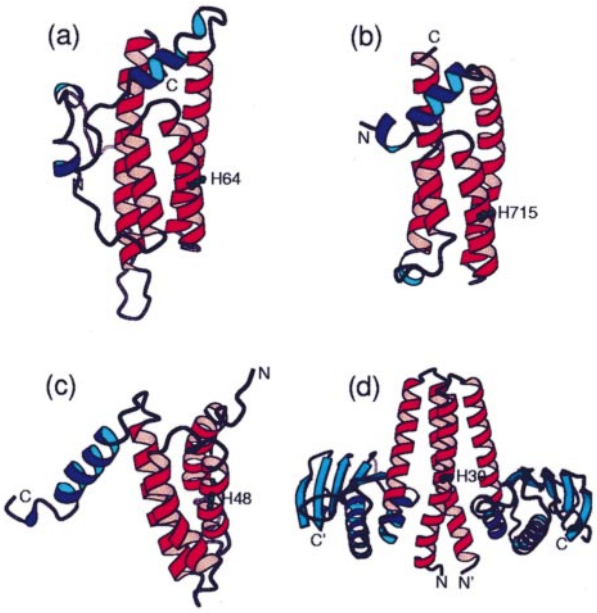

Figure 2. Ribbon diagram comparing the overall structures of (a) yeast Ypd1p, (b) the HPt domain of *E. coli* ArcB, (c) the P1 domain of *E. coli* CheA, and (d) Spo0B from *B. subtilis*. The common structural motif of the four-helix bundle is colored in red and the remaining secondary structure elements in blue, respectively. The phosphorylation sites (His64 in Ypd1p; His715 in ArcB HPt domain; His48 in CheA P1 domain; His30 in Spo0B) are also shown.

Despite sharing a common four-helix bundle motif, significant structural differences exist between Ypd1p and the HPt domain of ArcB. The helices C and F of ArcB HPt domain are severely kinked, whereas the corresponding helices B and F of Ypd1p are only slightly curved (Figure 2). When 22 residues belonging to helices B and F of Ypd1p (residues 38-51 and 141-148) are also included in the superposition of the two proteins in addition to the above-mentioned 34 C^{α} atom pairs, the rms deviation increases to 0.8 Å for 57 C^{α} atom pairs. A superposition of the remaining parts is not

Figure 1. (a) Stereo ribbon diagram showing the secondary structure elements of yeast Ypd1p. Six α-helices (blue ribbons), two 3_{10}-helices (yellow ribbons), and connecting loops (pink ropes) are drawn. The site of phosphorylation, His64, is shown in green. The Figure was drawn with MOLSCRIPT (Kraulis, 1991) and RASTER 3D (Merritt & Bacon, 1997). (b) Structure-based alignment of the sequences of yeast Ypd1p and *E. coli* ArcB HPt domain. This differs from the alignment reported by Kato *et al.* (1997) for the C-terminal part of Ypd1p. The site of phosphorylation (His64 in Ypd1p) is marked by a red triangle. Secondary structure elements of Ypd1p are shown above the sequence and the colored boxes also indicate the secondary structures for both Ypd1p and ArcB HPt domain. Six α-helices (blue) are labeled sequentially from A to F and two 3_{10}-helices (red) are labeled as G1 and G2. The Figure was produced with ALSCRIPT (Barton, 1993). (c) Stereo diagram showing the backbone superposition of Ypd1p (green) and ArcB HPt domain (magenta). Several residues of Ypd1p are also drawn and labeled. The Figure was drawn with GRASP (Nicholls *et al.*, 1991).

straightforward, due to large structural differences. The starting helix A of Ypd1p and the equivalent helix B of ArcB HPt domain show a shift of ~4.5 Å (Figure 1(c)). The loops connecting α-helices, except the γ-turn between helices C and D, are structurally more divergent than the secondary structural elements. The most pronounced difference is the presence in Ypd1p of the 44 residue insertion between helices D and F (Figure 1), which is missing from the HPt domain of ArcB. This insertion region of Ypd1p partly covers the four-helix bundle motif.

The P1 domain of CheA (Zhou & Dahlquist, 1997) has a global fold of five α-helices that includes an antiparallel four-helix bundle with the same topology as that of Ypd1p (Figure 2(c)). A superposition of structurally equivalent residues in the two proteins yields a relatively large rms deviation (2.5 Å for 26 C^{α} atoms corresponding to residues 60-85 of Ypd1p). The CheA P1 domain lacks the N-terminal region encompassing the helix A of Ypd1p but has an extra protruding helix at the C terminus (Figure 2(c)). This domain may provide the nucleophile for phosphoryl transfer (His48) and the activating glutamate (Glu70), thereby completing the ATPase catalytic center observed in GyrB (Bilwes *et al.*, 1999). But a detailed structural comparison is made difficult by the relatively low resolution of the solution structure of CheA P1 domain. Spo0B from *B. subtilis* (Varughese *et al.*, 1998), functionally equivalent to Ypd1p and the HPt domain of ArcB, is structurally unique among the phosphorelay H2 modules. Wherease the four-helix bundles of Ypd1p, the HPt domain of ArcB, and the P1 domain of CheA are all monomeric, that of Spo0B is formed by a dimeric association of two identical polypeptide chains (Figure 2(d)). Thus, the four-helix bundle of Spo0B has two active-site histidine residues on opposite sides of the bundle. To summarize, our structural comparisons suggest that the four-helix bundle architecture with the fully exposed active-site histidine residue is commonly utilized by both prokaryotic and lower eukaryotic organisms to serve as a key intermediate in multistep phosphorelay signal transduction.

Active site features and conserved residues

The site of phosphorylation, His64, resides on the surface of helix C (Figure 1). Its side-chain, well defined by the electron density (Figure 3(a)), protrudes into the solvent for efficient phosphorelay (Figure 1(c)). Its side-chain atoms exhibit higher than average *B*-factors in both tetragonal and triclinic crystal forms (32.2 and 52.2 $Å^2$, respectively). Between the two crystal forms, there is a shift of about 0.4 Å in the position of His64 side-chain and a rotation of about 20° around the C^{β}-C^{γ} bond axis in the orientation. Residues whose side-chains are spatially adjacent to His64 include Asn61, Phe65, and Lys67 on the same helix C as well as Glu83 and Gln86 located near the C-terminal side of helix

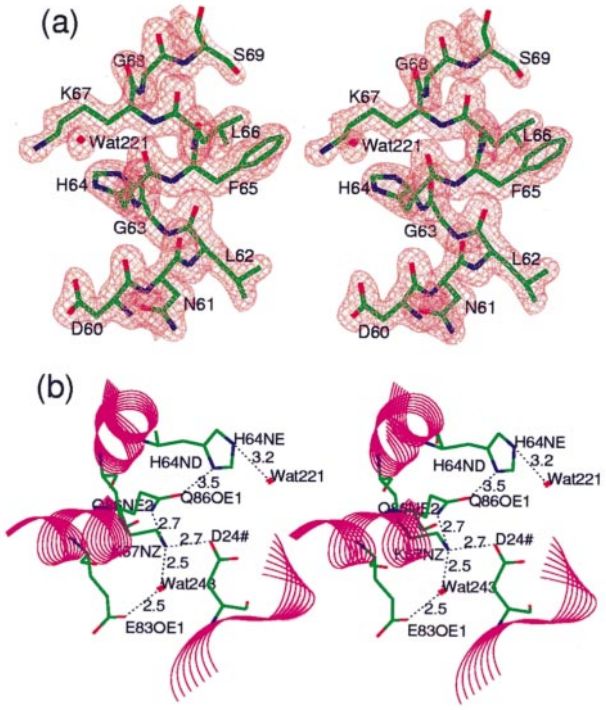

Figure 3. (a) Final ($2F_o - F_c$) electron density map around His64 in α-helix C of the Ypd1p structure. The map was calculated using 15-1.8 Å data and contoured at 1.4 σ (0.28 e/$Å^3$). (b) Hydrogen bonding network around His64 in the active site of Ypd1p. The distances indicated are an average value for the two molecules in the crystallographic asymmetric unit of the tetragonal crystal.

D (Figure 1(c)). A hydrogen-bonding network is formed by the side-chains of His64, Gln86, Lys67, a water molecule (Wat243), and Glu83 in this sequential order in the tetragonal crystal (Figure 3(b)). This network also includes Asp24 from a symmetry-related molecule in both crystal forms. A similar but not identical hydrogen-bonding network is present in ArcB HPt domain, where His715 bonds directly with Gln737 (Kato *et al.*, 1997). Here, the residues of ArcB HPt domain are numbered following a recent correction of the SWISS-PROT file (code P22763). Gln737 of ArcB HPt domain is equivalent to Gln86 of Ypd1p. It was suggested that this glutamine residue plays a crucial role in the activity by possibly orienting the imidazole ring so as to expose the imidazole N^{ε} atom to the solvent through a direct hydrogen-bonding to the imidazole N^{δ} atom (Kato *et al.*, 1997). Our structure further supports this notion. The imidazole N^{ε} atom of His64 in Ypd1p is hydrogen bonded to a water molecule as in ArcB HPt domain (Figure 3(b)). This geometry allows the N^{ε} atom of His64 to be more favorable for phosphorylation. In fact, the NMR studies indicated that the N^{ε} atom of the active histidine resi-

due of the P1 domain of CheA is phosphorylated (Zhou & Dahlquist, 1997).

Gly68 of Ypd1p is highly conserved among HPt domains (Kato *et al.*, 1997). The corresponding residue Gly719 of the ArcB HPt domain, which is located on the adjacent ridge of helix D with His715, seems to carve out an empty space for the histidine residue so as to expose the imidazole ring to the solvent region (Kato *et al.*, 1997). It appears that Gly68 of Ypd1p plays the same role. The side-chain of Asn61, located adjacent to His64, points away from His64 so that it would not block the side-chain of His64 (Figure 1(c)). This was also the case for the corresponding residue Glu712 of ArcB (Kato *et al.*, 1997). It may participate in interaction with the receiver domains. Leu66 of Ypd1p and its equivalent Ile717 of ArcB are well conserved (Kato *et al.*, 1997). Their side-chains are mostly buried and, therefore, they do not appear to be part of the interface for interaction with the receiver domains.

For the ArcB HPt domain (Kato *et al.*, 1997), it was suggested that a zinc ion, coordinated to His728, Glu754, Glu758, and a symmetry-related Asp746, may play a possible role in interactions with other parts of ArcB or the receiver domain of its response regulator. Corresponding residues of Ypd1p are Arg77, Gln147, Glu151, and Ile139 (Figure 1(b)). Their side-chains adopt unfavorable orientations for coordinating a metal ion and are blocked by the 44-residue insertion region between helices D and F. Therefore, the metal binding is unique to the HPt domain of ArcB and its role may be restricted to interactions with other parts of ArcB. It is not likely to be essential for interactions with the receiver domain.

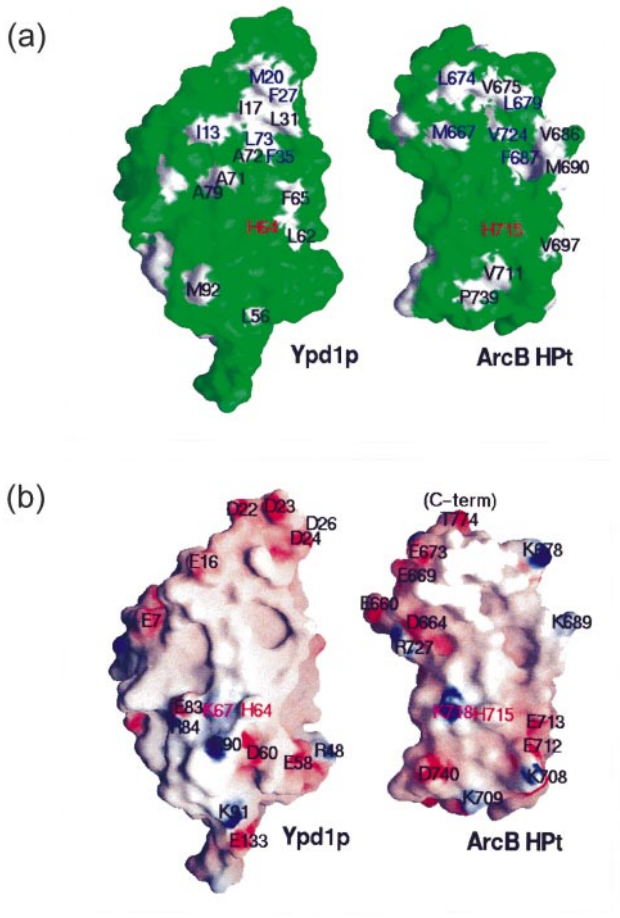

Figure 4. (a) Diagram comparing the hydrophobic surfaces of Ypd1p and ArcB HPt domain. This view is obtained by a clockwise 30° rotation of Figure 1(a) around a vertical axis. The residues forming the hydrophobic surface are white and are labeled. (b) Diagram comparing the electrostatic potential at the molecular surfaces of Ypd1p and the ArcB HPt domain. This is the same view as in (a). Negatively charged regions are red and positively charged regions blue. Charged residues are labeled. The Figure was drawn with GRASP (Nicholls *et al.*, 1991).

Specificity of phosphotransfer

One of the crucial questions in phosphorelay signal transduction is how a specific recognition among the components is achieved. In order to understand the mode of interaction between Ypd1p and its cognate response regulators, structural information on the complexes is needed. This study reveals the structure of Ypd1p but no structural data on its interacting counterparts, the receiver domains of Sln1p, Ssk1p, and Skn7p are available. Since these receiver domains may be expected to fold into the same globular $(\alpha/\beta)_5$ fold (Volz, 1993) as that of CheY, an attempt was made to construct the homology models using the 3D modeling server†. Preliminary docking studies indicate that there exists a general shape complementarity between Ypd1p and the receiver domains. However, homology models are not accurate enough to allow proposing specific models of the complexes.

Recently, the solution structure of a complex between the N-terminal domain of enzyme I and the histidine-containing phosphocarrier protein of the *E. coli* sugar phosphotransferase system provided a direct view of a protein-protein phosphoryl transfer complex (Garrett *et al.*, 1999). It showed that no large conformational changes accompany the complex formation and that the majority of the protein-protein interface contacts are hydrophobic, with additional electrostatic interactions. Therefore, we have compared the distribution of hydrophobic residues and the electrostatic potential at the molecular surfaces of Ypd1p and ArcB HPt domain around the phosphorylation sites in Figure 4. In both proteins, the upper half is more hydrophobic than the lower half. Some residues are conserved between the two proteins, while others are unique to each protein. Ile13, Met20, Phe27, Phe35, and

† http://www.expasy.ch/swissmod/SWISS-MODEL.html

Leu73 of Ypd1p correspond to Met667, Leu674, Leu679, Val724, and Phe687 of the ArcB HPt domain (labeled in blue in Figure 4(a)). Side-chains of all these residues are located in generally the same position relative to the key histidine residues, except Met20 of Ypd1p and Leu674 of ArcB HPt domain whose side-chains point in different directions. The hydrophobic surface of Met20 side-chain in Ypd1p is instead provided by the side-chain of Val675 in the HPt domain of ArcB. A notable difference is the presence of a small hydrophobic patch formed by residues Leu62 and Phe65 to the right of His64 in Ypd1p, whereas a hydrophobic patch of similar size, formed by Val711 and Pro739, is located below His715 in ArcB HPt domain (Figure 4(a)). In both Ypd1p and the ArcB HPt domain, the charged residues are more abundant in the lower half of the molecules, with Lys67 of Ypd1p and Lys718 of the HPt domain occupying similar positions (labeled in magenta in Figure 4(b)). The different locations of other charged residues relative to the key histidine residues make the electrostatic potential surfaces of these two proteins much different from each other. We suggest that these distinct surface features are the key determinants of specificity in recognizing the cognate response regulators.

Protein Data Bank accession numbers

The coordinates have been deposited in the Protein Data Bank for immediate release: accession number 1C02 for the tetragonal form and 1C03 for the triclinic form.

Acknowledgments

We thank the Inter-University Center for Natural Science Research Facilities for providing the X-ray equipment. We thank Professor N. Sakabe, Dr N. Watanabe, Dr M. Suzuki, and Dr N. Igarashi for assistance during data collection at BL-6A and BL-6B (TARA beamline) of Photon Factory, Japan. We thank Dr F. W. Dahlquist and Dr J.-P. Samama for kindly providing us the coordinates of the P1 domain of CheA and the complex between the P2 domain of CheA and CheY. H.K.S. is supported by the Postdoctoral Fellowship from Korea Ministry of Education. This work was supported by a grant from Korea Science and Engineering Foundation through the Center for Molecular Catalysis at Seoul National University.

References

Barrett, J. F., Goldschmidt, R. M., Lawrence, L. E., Foleno, B., Chen, R., Demers, J. P., Johnson, S., Kanojia, R., Fernandez, J., Bernstein, J., Licata, L., Donetz, A., Huang, S., Hlsta, D. J. & Macielag, M. J., *et al*. (1998). Antibacterial agents that inhibit two-component signal transduction systems. *Proc. Natl Acad. Sci. USA,* **95**, 5317-5322.

Barton, G. J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37-40.

Bilwes, A., Alex, L. A., Crane, B. R. & Simon, M. I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell,* **96**, 131-141.

Boguslawski, G. (1992). PBS2, a yeast gene encoding a putative protein kinase, interacts with the RAS2 pathway and affects osmotic sensitivity of *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **138**, 2425-2432.

Brewster, J. L., de Valoir, T., Dwyer, N. D., Winter, E. & Gustin, M. G. (1993). An osmosensing signal transduction pathway in yeast. *Science,* **259**, 1760-1763.

Brünger, A. T. (1993). *X-PLOR version 3.1*, Yale University Press, New Haven, USA, CT.

Burbulys, D., Trach, K. A. & Hoch, J. A. (1991). Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell,* **64**, 545-552.

CCP4 (1994). Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallog. sect. D,* **50**, 760-763.

Chang, C. & Meyerowitz, E. M. (1994). Eukaryotes have ''two-component'' signal transducers. *Res. Microbiol.* **145**, 481-486.

Eggar, L. A., Park, H. & Inouye, M. (1997). Signal tranduction *via* the histidyl-aspartyl phosphorelay. *Genes Cells,* **2**, 167-184.

Garrett, D. S., Seok, Y.-J., Peterkofsky, A., Gronenborn, A. M. & Clore, G. M. (1999). Solution structure of the 40,000 M_r phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. *Nature Struct. Biol.* **6**, 166-173.

Hoch, J. A. & Silhavy, T. J. (1995). *Two-component Signal Transduction*, ASM Press, Washington, USA, DC.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.

Jones, T. A. (1992). A set of programs. In *Molecular Replacement* (Dodson, E. J., Gover, S. & Wolf, W., eds), pp. 91-105, SERC Daresbury Laboratory, Daresbury, UK.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.

Kato, M., Mizuno, T., Shimizu, T. & Hakoshima, T. (1997). Insights into multistep phosphorelay from the crystal structure of the C-terminal HPt domain of ArcB. *Cell,* **88**, 717-723.

Ketela, T., Brown, J. L., Stewart, R. C. & Bussey, H. (1998). Yeast Skn7p activity is modulated by the Sln1p-Ypd1p osmosensor and contributes to regulation of the HOG pathway. *Mol. Gen. Genet.* **259**, 373-378.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Lee, M. G., Lee, J. Y., Song, H. K. & Suh, S. W. (1999). Crystallization and preliminary X-ray analysis of *Saccharomyces cerevisiae* Ypd1p, a key intermediate in phosphorelay signal transduction. *Acta Crystallog. sect D,* **55**, 1219-1221.

Li, S., Ault, A., Malone, C. L., Raitt, D., Dean, S., Johnston, L. H., Deschenes, R. J. & Fassler, J. S. (1998). The yeast histidine protein kinase, Sln1p, mediates phosphotransfer to two response regulators, Ssk1p and Skn7p. *EMBO J.* **17**, 6952-6962.

Loomis, W. F., Shaulsky, G. & Wang, N. (1997). Histidine kinases in signal transduction pathways of eukaryotes. *J. Cell Sci.* **110**, 1141-1145.

Maeda, T., Takekawa, M. & Saito, H. (1995). Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. *Science,* **269**, 554-558.

Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Mizuno, T. (1998). His-Asp phosphotransfer signal transduction. *J. Biochem.* **123**, 555-563.

Navaza, J. (1994). AMoRe: an automatic package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-293.

Ota, I. M. & Varshavsky, A. (1993). A yeast protein similar to bacterial two-component regulators. *Science,* **262**, 566-569.

Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Thai, T. C. & Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multistep phosphorelay mechanism in the SLN1-YPD1-SSK1 ''two-component'' osmosensor. *Cell,* **86**, 865-875.

Sakabe, K., Sasaki, K., Watanabe, N., Suzuki, M., Wang, Z. G., Miyahara, J. & Sakabe, N. (1997). Large format imaging plate and Weissenberg camera for accurate protein crystallographic data collection using synchrotron radiation. *J. Synchrotron Rad.* **4**, 136-146.

Sheldrick, G. M. & Schneider, R. M. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319-343.

Swanson, R. V. & Simon, M. I. (1994). Bringing the eukaryotes up to speed. *Curr. Biol.* **4**, 234-237.

Swanson, R. V., Alex, L. A. & Simon, M. I. (1994). Histidine and aspartate phosphorylation: two-component systems and the limits of homology. *Trends Biochem. Sci.* **19**, 485-490.

Varughese, K. I., Madhusudan, , Zhou, X. Z., Whiteley, J. M. & Hoch, J. A. (1998). Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase. *Mol. Cell,* **2**, 485-493.

Volz, K. (1993). Structural conservation in the CheY superfamily. *Biochemistry,* **32**, 11741-11753.

Wurgler-Murphy, S. M. & Saito, H. (1997). Two-component signal transducers and MAPK cascades. *Trends Biochem. Sci.* **22**, 172-176.

Xu, Q., Nguyen, V. & West, A. H. (1999). Purification, crystallization, and preliminary X-ray diffraction analysis of the yeast phosphorelay protein YPD1. *Acta Crystallog. sect. D,* **55**, 291-293.

Zhou, H. & Dahlquist, F. W. (1997). Phosphotransfer site of the chemotaxis-specific protein kinase CheA as revealed by NMR. *Biochemistry,* **36**, 699-710.

Edited by I. A. Wilson

(Received 15 July 1999; received in revised form 16 September 1999; accepted 16 September 1999)