Molecular Insights into Toluene Sensing in the TodS/TodT Signal Transduction System*

Serry Koh‡1, Jungwon Hwang‡, Koushik Guchhait†2, Eun-Gyeong Lee†4, Sang-Yoon Kim†, Sujin Kim†, Sangmin Lee†, Jeong Min Chung‡, Hyun Suk Jung‡, Sang Jun Lee‡,†5, Choong-Min Ryu‡,†6, Seung-Goo Lee‡,†7, Tae-Kwang Oh†,‡8, Oh Suk Kwon‡,†9, and Myung Hee Kim‡,‡10

From the ‡Infection and Immunity Research Center, the **Biochemicals and Synthetic Biology Research Center, and the **Molecular Phytobacteriology Laboratory, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Korea; the †Department of Biochemistry, College of Natural Sciences, Kangwon National University, Chuncheon, Gangwon-do 200-701, Korea, and the §Biosystems and Bioengineering Program, University of Science and Technology, Daejeon 305-350, Korea

TodS is a sensor kinase that responds to various monoaromatic compounds, which either cause an agonistic or antagonistic effect on phosphorylation of its cognate response regulator TodT, and controls tod operon expression in Pseudomonas putida strains. We describe a molecular sensing mechanism of TodS that is activated in response to toluene. The crystal structures of the TodS Per-Arnt-Sim (PAS) 1 sensor domain (residues 43–164) and its complex with toluene (agonist) or 1,2,4-trimethylbenzene (antagonist) show a typical B2a3β3 PAS fold structure (residues 45–149), forming a hydrophobic ligand-binding site. A signal transfer region (residues 150–163) located immediately after the canonical PAS fold may be intrinsically flexible and disordered in both apo-PAS1 and antagonist-bound forms and dramatically adapt an α-helix upon toluene binding. This structural change in the signal transfer region is proposed to result in signal transmission to activate the TodS/TodT two-component signal transduction system. Site-directed mutagenesis and β-galactosidase assays using a P. putida reporter strain system verified the essential residues involved in ligand sensing and signal transfer and suggest that the Phe46 residue acts as a ligand-specific switch.

As global fossil fuel consumption has increased, highly volatile aromatic hydrocarbons from petroleum by-products, such as toluene, benzene, and xylene isomers, have accumulated in the soil, water, and air. Growing public health concerns regarding prolonged exposure to these toxic substances have driven efforts to develop microbial remediation of contaminated soil and groundwater near petroleum refineries with Pseudomonas putida strains (1–6).

P. putida F1 (1, 7) and DOT-T1E (8) strains express genes encoding catabolic enzymes for the oxidation of toluene to 2-hydroxy-6-oxo-2,4-heptadienoate within the tod operon. TodS/TodT, a two-component signal transduction system, is responsible for regulation of tod operon expression in P. putida strains (1, 7, 8).

TodS is a member of the TodS-like family that is characterized by a unique domain architecture exclusively found in sensor kinases regulating degradation pathways (9, 10). It is an atypical histidine kinase composed of two supradomains, each containing a PAS4 domain and a histidine kinase (HK) domain, which are separated by a response regulator receiver (RRR) domain (10). TodS lacks a transmembrane region (7), and its N-terminal PAS domain (PAS1) binds a wide range of effectors, including toluene, to trigger increased basal autophosphorylation of TodS, leading to a phospho-signal relay to TodT and ultimately transcriptional stimulation of the tod operon genes by interacting with phosphorylated TodT in the P_{tod}{\text{X}} Promoter region (11–14). A study demonstrated that the N-terminal PAS1 domain, the histidine kinase (HK1) His190 residue, the central RRR Asp500 residue, and the C-terminal histidine kinase (HK2) His760 residue are all required for the multistep phospho-signal to the TodT Asp57 residue (9). The N-terminal PAS1 domain was identified as a signal sensor, whereas the role of the C-terminal PAS domain (PAS2) in the signal transduction system remains unclear (9–11).

Another Pseudomonas strain, Pseudomonas mendocina RK1, employs the TmoS/TmoT two-component signal transduction system to control expression of the toluene-4-monoxygenase pathway (15, 16). TmoS belonging to the TodS family binds to various aromatic compounds as agonists or antagonists to regulate TmoS autophosphorylation, further demonstrating the common functional features of this protein family (16).

‡This work was supported by Next-Generation Biogreen 21 Program Grants PJ00817002, PJ00952403, and PJ011118, Rural Development Administration, Korea; and Korea Research Institute of Bioscience and Biotechnology Research Initiatives Grants KGM4541622 and KGM2111622. The authors declare that they have no conflicts of interest with the contents of this article.

© 2016 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
Crystal Structure of the Toluene-sensing TodS PAS1 Domain

PAS domains have been identified in many proteins originating from archaea, bacteria, plants, and animals with highly divergent nucleotide sequences and ligand-binding capacities (17–19). As part of sensor kinases, PAS domains are often found at the N terminus, either in tandem or spaced in signal sensory modules (17, 20, 21). The functional roles of PAS domains include protein-protein interactions as well as ligand binding (17, 21, 22). PAS domains bind to a broad spectrum of ligands, such as light, oxygen, proton, heme, FMN, FAD, citrate, malate, and divalent metal cation (17, 18). In most cases, PAS functions either in a homodimer or a heterodimer with a parallel orientation, and a single ligand binds to each PAS domain (22–25).

Despite its importance in bioremediation, the structure of TodS has not yet been determined. A model of the TodS PAS1 structure predicted a hydrophobic ligand-binding pocket and suggested that four amino acid residues (Phe46, Ile74, Phe79, and Ile114) are involved in ligand binding (10), F46A, I74A, and I114A mutant variants of TodS showed significantly lower in vitro ligand-binding affinities, including for both agonists and antagonists, than that of wild-type (WT) TodS (10). Particularly, the F79A mutant did not respond to even high doses of ligands in both in vitro and in vivo assays, implying that the Phe79 residue of PAS1 is the most critical residue for ligand binding and autophosphorylation of TodS (10). Although the model structure of PAS1 is useful for predicting the ligand-binding residues, the key question of how the N-terminal PAS1 ligand-sensing domain transmits signals to the C-terminal TodS remains to be answered.

In the present study, we determined the crystal structures of the TodS PAS1 domain in complex with either toluene (an agonist) or 1,2,4-trimethylbenzene (1,2,4-TMB; an antagonist), as well that of apo-PAS1. On the basis of this structural information, we propose a mechanism for TodS molecular sensing and signal transfer to TodT using combined biochemical and in vivo physiological analyses.

Experimental Procedures

**Bacterial Strains and Plasmids**—The bacterial strains used and plasmids constructed in this study are listed in Table 1. Briefly, the todS genes encoding TodS (residues 43–978) and dimerization helix-containing TodS (residues 23–978) were PCR-amplified using P. putida F1 genomic DNA as a template. The amplified TodS (residues 43–978) gene was subcloned into the Ncol and Xhol sites of pET28b (+) (Novagen), a C-terminal His-tagged protein expression vector. The TodS (residues 23–978) gene was subcloned into the BamHI and Xhol sites of the pGST-Parallel1 vector (26), an N-terminal GST-fused protein expression vector. The DNA fragments encoding PAS1 (residues 43–168), dimerization helix-containing PAS1 (residues 23–168), and PAS2 (residues 611–729) were PCR-amplified using the TodS expression plasmid and subcloned into the Ncol and Xhol sites of the pHis-Parallel1 vector (26), an N-terminal Hisα-tagged protein expression vector containing an recombinant tobacco etch virus protease cleavage site. The todST promoter and the todST gene region in P. putida F1 were PCR-amplified using TodSTpp-u and TodSTpp-m primers and TodST-F and TodST-R primers, respectively (Table 2). The todST promoter region was cloned into the NsiI and XbaI sites of the pBRRBB-eGFP vector (27), resulting in pBRR-PtodST. Subsequently, todST was subcloned into the SacI and HindIII sites of pBRR-PtodST to express todST under the native todST promoter (pBRR-PtodST-TodST) for in vivo assays. No green fluorescent protein gene is present in the resultant expression plasmid. The mutant variants were generated by site-directed mutagenesis using a QuikChange site-directed mutagenesis kit, according to the manufacturer’s instructions (Stratagene). The L71M mutation was introduced into PAS1 (residues 43–168) to facilitate selenomethionine (SeMet) incorporation for solving the crystal structure. All introduced mutations were confirmed by DNA sequencing (Macrogen, Seoul, Korea). All primers used to create plasmids and various mutant variants are listed in Table 2.

**Construction of the P. putida Reporter Strain**—To perform the β-galactosidase assay, a P. putida KT2440-PXZ reporter strain was generated, which contains the TodT-binding promoter P todX, lacZ, and the streptomycin resistance gene SmR (P todX-lacZ-SmR) cassette inserted into the mexC gene on the P. putida KT2440 chromosome (28). Multiple subcloning steps were performed for reporter strain construction. First, two 539- and 509-bp fragments of the P. putida KT2440 mexC gene and the SmR gene of pEXT21 plasmid (29) were PCR-amplified with primer combinations of mexCn-u and mexCn-d, mexCc-u and mexCc-d, and Sm-N and Sm-C, respectively. These three PCR products were digested with EcoRI and BamHI, BamHI and SphI, and BamHI, respectively. They were then subcloned into the EcoRI and SphI sites of pBRR32 (New England Biolabs, Inc.), resulting in pSYK123. Subsequently, the P todX-lacZ translational fusion cassette of pEXT20-todST-P todX was subcloned into the HindIII site of pSYK123 to generate pSYK124 with the P todX-lacZ-SmR cassette encompassed by the sequence of mexC. Finally, 5,528 bp of the mexC-P todX-lacZ-SmR-mexC cassette was PCR-amplified using the primers mexCn-u and mexCc-d and inserted into SmaI-digested pK18mob6ac (30), yielding pSYK135. P. putida KT2440 was then transformed with pSYK135 by electroporation (31). The Km5 and sucrose-sensitive *Pseudomonas* transformants were double-selected on Luria-Bertani (LB) medium supplemented with 50 μg/ml kanamycin and 10% sucrose. A single-crossover event in the selected transformants was further confirmed by diagnostic PCR. P. putida KT2440-PXZ, a reporter strain with a chromosomally integrated P todX-lacZ-SmR cassette generated by a second crossover round, was finally selected on LB plates supplemented with 100 μg/ml streptomycin.

**β-Galactosidase Assay**—The P. putida KT2440-PXZ reporter strain was transformed with pBRR-P todST-TodST or its TodS mutants by electroporation. The transformant strains were grown overnight and diluted in fresh LB medium supplemented with kanamycin (50 μg/ml). The strains were then grown at 30 °C at 500 rpm for 5 h in a gas phase saturated with a ligand supplied over a range of 10–400 μM. β-Galactosidase activity by TodST or its TodS mutants in response to each ligand at different concentrations was measured using E. G. Lee, S.-Y. Kim, and O. Kwon, unpublished data.
4-methylumbelliferyl $\beta$-galactopyranoside (MUG; Sigma). MUG units were then calculated, as described previously (32–34).

Expression and Purification of Proteins—The N-terminally His$_{6}$-tagged PAS1(43–168) and PAS2 proteins were expressed by induction with 0.5 mM isopropyl $\beta$-D-thiogalactopyranoside in the Escherichia coli BL21 Star (DE3) system (Invitrogen) at 18 °C overnight. All purification steps were performed with buffer A containing 50 mM Tris-Cl (pH 8.0) and 300 mM NaCl. Cultured cells were harvested, resuspended in buffer A, and ultrasonicated. The crude extracts were centrifuged at 16,000 $\times$ g at 4 °C for 1 h. The cell lysate was then loaded onto
an Ni-NTA (Qiagen) affinity column, and the His$_6$-tagged proteins were eluted with 200 mM imidazole. The eluted proteins were further purified by recombinant tobacco etch virus protease treatment, size exclusion chromatography (SEC) using a Superdex G75 column (GE Healthcare), and additional PAS1(23) NcoI-F

| Primer       | Sequence (5’→3’) | Length |
|--------------|------------------|--------|
| TodST-F      | AAA GGA TCC AGG AGA CAT ATG AGC TCG TGT GAT A | 34 bp |
| TodST-R      | TAT AAG CTT CTA TCC AGG GCT ATC CTT TGT | 27 |
| TodStpp-u    | TAT ATG CAT CTC GAG AAA CTA GCC CAG TAC | 30 |
| TodStpp-m    | CCC TCT AGA AGC TTA CTA CCT CTC TTC TAC CAC C | 34 |
| mexCn-u      | AAA GGA TCC AGA AGC GAG GGG CTG TCA AAG GGG TG | 32 |
| mexCn-d      | AAA GGA TCC ATC TAA GCC GGA GAT CTT G | 27 |
| mexCc-u      | AAA GGA TCC GAT ATC CAG CAA TAT CAC TGC CCG | 30 |
| mexCc-d      | TTG GCC TCG GCT GCG CAG GCG CAA C | 29 |
| Sm-N         | AAA GGA TCC AAC GCT GAA CTA CGG AAC GCA | 33 |
| Sm-C         | ATG GAC TTA TTT GGG CAC TCT GAT GAT C | 34 |
| PAS1(23) Ncol-F | CA ACC ATG GAA AGG GAG AAA GGA TCT GAA G | 30 |
| PAS1(43) Ncol-F | CA ACC ATG GCG TCT TAC GAG TGT | 26 |
| PAS1(168) Xhol-R | CAA CTC GAG TCA CCA TCT CGT GCT TAC C | 31 |
| TodS52 BamHI-F | CCA GTC TAC GCT GAG GCC GTC GAT CTT G | 29 |
| TodS798 Xhol-R | CCG CTC GAG TGT GCC GGA GCC CTG TGT G | 29 |
| L71M-F       | TGG AGG GGG GGC GGT ACT ATG GAA GAA ATC GAG AGG | 41 |
| L71M-R       | CCT TCC TCG TAT TTT TTC CAT AAT CCC GGC CCC CTC CA | 41 |
| F46A-F2      | GAG GGG CTC TAC GAG GGC TCG CTC GCC CTT CTT CTT | 34 |
| F46A-R2      | C ATC AAG AAG GCC CAC GGC CTC GTA GAG CCC ATC | 41 |
| V59A-F       | CTG AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG GAG GAG | 43 |
| V59A-R       | C TCT TCC ACC TGA CAC CTG GTC GCT TGC TGC TGC TAC TGG | 43 |
| A67V-F       | G CTT CTT TTA TAA GCC AAC ATG CCA TAT CTA AAA AAA ACC GAG | 46 |
| A67V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG | 46 |
| A63V-F       | T AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG | 41 |
| A63V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG | 41 |
| TodS52 BamHI-F | CCA GTC TAC GCT GAG GCC GTC GAT CTT G | 29 |
| TodS798 Xhol-R | CCG CTC GAG TGT GCC GGA GCC CTG TGT G | 29 |
| L71M-F       | TGG AGG GGG GGC GGT ACT ATG GAA GAA ATC GAG AGG | 41 |
| L71M-R       | CCT TCC TCG TAT TTT TTC CAT AAT CCC GGC CCC CTC CA | 41 |
| F46A-F2      | GAG GGG CTC TAC GAG GGC TCG CTC GCC CTT CTT CTT | 34 |
| F46A-R2      | C ATC AAG AAG GCC CAC GGC CTC GTA GAG CCC ATC | 41 |
| V59A-F       | CTG AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG GAG GAG | 43 |
| V59A-R       | C TCT TCC ACC TGA CAC CTG GTC GCT TGC TGC TGC TAC TGG | 43 |
| A67V-F       | G CTT CTT TTA TAA GCC AAC ATG CCA TAT CTA AAA AAA ACC GAG | 46 |
| A67V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG | 46 |
| A63V-F       | T AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| A63V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| TodS52 BamHI-F | CCA GTC TAC GCT GAG GCC GTC GAT CTT G | 29 |
| TodS798 Xhol-R | CCG CTC GAG TGT GCC GGA GCC CTG TGT G | 29 |
| L71M-F       | TGG AGG GGG GGC GGT ACT ATG GAA GAA ATC GAG AGG | 41 |
| L71M-R       | CCT TCC TCG TAT TTT TTC CAT AAT CCC GGC CCC CTC CA | 41 |
| F46A-F2      | GAG GGG CTC TAC GAG GGC TCG CTC GCC CTT CTT CTT | 34 |
| F46A-R2      | C ATC AAG AAG GCC CAC GGC CTC GTA GAG CCC ATC | 41 |
| V59A-F       | CTG AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG GAG GAG | 43 |
| V59A-R       | C TCT TCC ACC TGA CAC CTG GTC GCT TGC TGC TGC TAC TGG | 43 |
| A67V-F       | G CTT CTT TTA TAA GCC AAC ATG CCA TAT CTA AAA AAA ACC GAG | 46 |
| A67V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG | 46 |
| A63V-F       | T AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| A63V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| TodS52 BamHI-F | CCA GTC TAC GCT GAG GCC GTC GAT CTT G | 29 |
| TodS798 Xhol-R | CCG CTC GAG TGT GCC GGA GCC CTG TGT G | 29 |
| L71M-F       | TGG AGG GGG GGC GGT ACT ATG GAA GAA ATC GAG AGG | 41 |
| L71M-R       | CCT TCC TCG TAT TTT TTC CAT AAT CCC GGC CCC CTC CA | 41 |
| F46A-F2      | GAG GGG CTC TAC GAG GGC TCG CTC GCC CTT CTT CTT | 34 |
| F46A-R2      | C ATC AAG AAG GCC CAC GGC CTC GTA GAG CCC ATC | 41 |
| V59A-F       | CTG AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG GAG GAG | 43 |
| V59A-R       | C TCT TCC ACC TGA CAC CTG GTC GCT TGC TGC TGC TAC TGG | 43 |
| A67V-F       | G CTT CTT TTA TAA GCC AAC ATG CCA TAT CTA AAA AAA ACC GAG | 46 |
| A67V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG | 46 |
| A63V-F       | T AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| A63V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| TodS52 BamHI-F | CCA GTC TAC GCT GAG GCC GTC GAT CTT G | 29 |
| TodS798 Xhol-R | CCG CTC GAG TGT GCC GGA GCC CTG TGT G | 29 |
| L71M-F       | TGG AGG GGG GGC GGT ACT ATG GAA GAA ATC GAG AGG | 41 |
| L71M-R       | CCT TCC TCG TAT TTT TTC CAT AAT CCC GGC CCC CTC CA | 41 |
| F46A-F2      | GAG GGG CTC TAC GAG GGC TCG CTC GCC CTT CTT CTT | 34 |
| F46A-R2      | C ATC AAG AAG GCC CAC GGC CTC GTA GAG CCC ATC | 41 |
| V59A-F       | CTG AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG GAG GAG | 43 |
| V59A-R       | C TCT TCC ACC TGA CAC CTG GTC GCT TGC TGC TGC TAC TGG | 43 |
| A67V-F       | G CTT CTT TTA TAA GCC AAC ATG CCA TAT CTA AAA AAA ACC GAG | 46 |
| A67V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG | 46 |
| A63V-F       | T AAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |
| A63V-R       | C TAT CCC GCC CCC CTC TAA TAC GGT CTT AAC TAC TGG AAC CAG CAG | 41 |

were concentrated to 15 mg/ml and stored at −80 °C for use. N-terminally His$_6$-tagged PAS1(23–168) and its mutant proteins were expressed in the E. coli C41(DE3) system (Lucigen) using the same method as for PAS1(43–168) and were purified in buffer A, supplemented with 5% glycerol and 2 mM DTT, using only Ni-NTA affinity chromatography for use in the isothermal titration calorimetry (ITC) experiments. S2Met-sub-
The C-terminal His6-tagged TodS(23–978) protein was expressed in *E. coli* B834 (DE3) (Novagen), a methionine auxotroph strain, in minimal medium containing 50 mg/ml SeMet. The purification procedure for substituted PAS1 (L71M) protein (10 mg/ml) and toluene complex in vapor diffusion methods at 21 °C. Crystals of the SeMet-substituted PAS1 (L71M) protein (10 mg/ml) and toluene complex in a 1:2 molar ratio were grown and optimized under conditions of 3.0 M ammonium phosphate and 0.1 M Tris-HCl (pH 8.0) and 0.1 M Tris-HCl (pH 8.0). The complex crystals were transferred to the crystallization solution functioning as a cryosolution and diffracted on a 173 °C nitrogen gas stream. Single-wavelength anomalous diffraction data for the SeMet-substituted complex was collected at a resolution of 1.65 Å at beamline 5C of the Pohang Accelerator Laboratory (Pohang, Korea). Crystals of apo-PAS1 and its complex with 1,2,4-TMB were diffracted on a 17A of the Photon Factory (Tsukuba, Japan), at a resolution of 1.5 and 2.0 Å, respectively. All data were processed using the HKL2000 software package (35). The structure of the SeMet-PAS1 and toluene complex was initially determined by utilizing the anomalous signals from selenium atoms with the AutoSol phasing module, and density modification and automatic model building were performed using AutoBuild from the PHENIX software package (36). The structures of apo-PAS1 and its complex with 1,2,4-TMB were solved by molecular replacement in MOLREP (37) using the SeMet-PAS1 and toluene complex structure as a template structure. All structures were refined with REFMAC5 (39). The refinement process included the translation-liberation-screw procedure. Crystallographic data are summarized in Table 3.

---

**Table 3**

| Data collection | Apo-PAS1 | SeMet-PAS1-toluene | PAS1–1,2,4-TMB |
|-----------------|----------|--------------------|------------|
| Space group     | P2₁      | P2₁,2₁             | P2₁,2₁     |
| Cell dimensions | a, b, c (Å) | 41.11, 49.31, 56.22 | 41.34, 47.74, 126.52 | 45.39, 51.02, 100.73 |
| α, β, γ (degrees) | 90, 109.22, 90 | 90, 90, 90 | 90, 90, 90 |
| Wavelength      | 0.9795   | 0.9794             | 0.9795     |
| Resolution (Å)  | 50–1.70 (1.73–1.70) | 50–1.65 (1.68–1.65) | 50–1.96 (1.99–1.96) |
| No. of total reflections | 174,101 | 398,442            | 235,138    |
| No. of unique reflections | 23,475 | 30,287             | 17,149     |
| Redundancy (%)  | 74 (7.3) | 13.2 (12.6)        | 13.7 (11.9) |
| Completeness (%) | 100.0 (100.0) | 97.9 (99.5)       | 98.8 (98.8) |
| Rfree (%)       | 10.2 (61.7) | 9.4 (31.7)        | 12.0 (52.9) |
| I/σ(I)          | 32.94 (5.00) | 48.62 (12.12)    | 49.98 (6.94) |

---

* Rfree is the intensity of the i-th observation and I is the mean intensity of the reflections.

---

**Crystallization, Diffraction, and Structure Determination—** Crystallization of PAS1(43–168) was performed by the sitting drop vapor diffusion method with commercially available sparse matrix screening kits at 21 °C. Initial crystals were grown in several different conditions, and the crystals for x-ray diffraction were optimized in conditions of 2.8–3.0 M ammonium phosphate, and 0.1 M Tris-HCl (pH 8.0–8.5) with PAS1 (5–10 mg/ml). For complex formation, PAS1 and each ligand were mixed in a 1:2 molar ratio and incubated on ice for 1 h. Crystals of PAS1 complexed with toluene and 1,2,4-TMB, respectively, were produced under the same conditions as apo-PAS1. To avoid mixing of volatile organic compounds, crystallization plates were kept in separate growth incubators. Crystals of the complex were optimized by both haging and sitting drop vapor diffusion methods at 21 °C. Crystals of the SeMet-substituted PAS1 (L71M) protein (10 mg/ml) and toluene complex in a 1:2 molar ratio were grown and optimized under conditions of 3.0 M ammonium phosphate and 0.1 M Tris-HCl (pH 8.0). The complex crystals were transferred to the crystallization solution functioning as a cryosolution and diffracted on a 173 °C nitrogen gas stream. Single-wavelength anomalous diffraction data for the SeMet-substituted complex was collected at a resolution of 1.65 Å at beamline 5C of the Pohang Accelerator Laboratory (Pohang, Korea). Crystals of apo-PAS1 and its complex with 1,2,4-TMB were diffracted at beamline 17A of the Photon Factory (Tsukuba, Japan), at a resolution of 1.5 and 2.0 Å, respectively. All data were processed using the HKL2000 software package (35). The structure of the SeMet-PAS1 and toluene complex was initially determined by utilizing the anomalous signals from selenium atoms with the AutoSol phasing module, and density modification and automatic model building were performed using AutoBuild from the PHENIX software package (36). The structures of apo-PAS1 and its complex with 1,2,4-TMB were solved by molecular replacement in MOLREP (37) using the SeMet-PAS1 and toluene complex structure as a template structure. All structures were revised using COOT (38) and refined with REFMAC5 (39). The refinement process included the translation-liberation-screw procedure. Crystallographic data are summarized in Table 3.
**Crystal Structure of the Toluene-sensing TodS PAS1 Domain**

**ITC**—The interactions between dimerization helix-containing PAS1(23–168) and ligand were analyzed by performing ITC experiments. The purified N-terminally His₆-tagged PAS1(23–168) and its mutant proteins were dialyzed against a buffer containing 50 mM Tris–HCl (pH 7.5), 200 mM KCl, 2 mM MgCl₂, 1 mM DTT, and 0.1 mM EDTA. These samples were then degassed by vacuum aspiration for 15 min, and ethanol was added to the samples (final concentration of 0.2%) prior to loading. Ligands containing 0.2% ethanol were diluted into the degassed buffer (10, 11). Titration was carried out at 25 °C. The calorimetric assay was performed using a VP-ITC (MicroCal Inc.). The stirring speed was 300 rpm, and the thermal power was recorded every 10 s. A concentration of 2–6 mM toluene or m-xylene in 0.2% ethanol in a syringe was titrated against 50 mM DTT. The absolute molecular weight of purified apo-TodS(23–978) protein in solution was determined using a Wyatt miniDAWN TREOS, a three-angle static light scattering detector (Wyatt Technology), coupled to an AKTA Purifier FPLC protein purification system equipped with a Superdex G200 column. PBS (pH 7.3) containing 10% glycerol and 5 mM DTT was filtered through a 0.20-μm filter (Toyo Roshi Kaisha Ltd., Tokyo, Japan), degassed, and used to equilibrate the Superdex G200 column at a constant rate of 0.25 ml/min at 25 °C. Calibration of the buffer solution was performed by adjusting the refractive index of 10% glycerol to 1.3448 and viscosity to 1.2125 at 25 °C. After stabilization, two reference proteins (1.5 mg/ml each), β-amylase (~220 kDa) and bovine serum albumin (BSA; 66.4 kDa), were used to evaluate the accuracy of light scattering signals. TodS(23–978) protein (1 mg/ml) was then loaded onto the Superdex G200 column and monitored. The Rayleigh ratio was calculated based on the collected UV absorbance at 280 nm and LS2 data at 659.2 nm (refractive index), and the molecular mass was determined by Astra software version 6.0 (Wyatt Technology).

**Results**

**Structure of the TodS PAS1 Sensor**—There are two PAS1 molecules in each asymmetric unit (Fig. 1A), resulting in an antiparallel dimer structure. The structure of each PAS1 domain is very similar to those described for other PAS domain structures, which consist of a five-stranded antiparallel β-sheet and three α-helices spanning residues 45–149 (Fig. 1A). An additional α-helix (α4, residues 150–163) is located immediately outside the canonical PAS fold in molecule A (Fig. 1A). Remarkably, the corresponding region of α4 is completely disordered in molecule B (Fig. 1A). The structure of the PAS1 backbone is most similar to those of the light-, oxygen-, voltage-sensitive (LOV) domain (Protein Data Bank code 1N9L, 11% sequence identity, root mean square deviation of 1.52 Å for 82 α-carbon

---

**FIGURE 1. Overall structure of apo-PAS1.** A, ribbon representation of the structure of TodS PAS1 in an asymmetric configuration. There are two PAS1 molecules in the asymmetric unit. The β-sheets and α-helices are shown in cyan and green, respectively. The C-terminal disordered region in molecule B corresponding to α4 in molecule A is displayed in red. B, apo-PAS1 (green) superimposed onto the LOV/PAS domain of the *C. reinhardtii* photoreceptor (Protein Data Bank code 1N9L; cyan) and the LOV1 domain of the Arabidopsis blue light receptor protein phototropin-2 (Protein Data Bank code 226D; gray).
pairs) of the *Chlamydomonas reinhardtii* photoreceptor (41) and the LOV1 domain (Protein Data Bank code 2Z6D, 11% sequence identity, root mean square deviation of 1.30 Å for 58 α-carbon pairs) of the *Arabidopsis* blue light receptor protein phototropin-2 (42), both of which bind FMN (Fig. 1B).

Dimerization of many PAS domains occurs mainly through an amphipathic α-helix located in the immediate upper region of the core PAS fold (42, 43). The TodS PAS1 domain constructed in this study did not include the corresponding helix region; thus, the antiparallel face-to-face dimerization observed in the TodS PAS1 structure may be an artificial consequence of the symmetry of crystal packing. In fact, secondary structure prediction suggests an α-helical region (residues 32–43) in the immediate upper region of the canonical PAS fold of TodS, which may be involved in the dimerization of PAS1 (see "Dimerization of PAS1").

**Signal Sensing by PAS1—Toluene** is a key signal effector required to activate the TodS/TodT signal transduction system. We determined the structure of PAS1 in complex with the agonist toluene. The overall structure of the complex is generally identical to that of apo-PAS1 (Fig. 2A). The agonist is positioned in a hydrophobic pocket formed by the hydrophobic residues Phe46, Gly48, Val59, Ala63, Phe79, Trp84, Trp85, Ile114, Val126, Phe128, Ala145, and Gly147 in each PAS1 molecule (Fig. 2B). The buried surface area of the pocket is 247.27 Å², as calculated using PISA software (44). In detail, the benzene ring of toluene is placed into the area formed mainly by the aromatic amino acids Phe46, Phe79, Trp84, Trp85, and Phe128. Ala63, Ile114, Val126, and Gly147 are involved in the interaction as well. The methyl group of toluene is surrounded by the residues Gly48, Val59, Ala63, and Ala145.

We further determined the structure of PAS1 in complex with the antagonist 1,2,4-TMB (Fig. 2C). The same amino acid residues as those in the hydrophobic pocket are involved in the interaction with the antagonist (Fig. 2D). The only difference between PAS1 interactions with the agonist and those with the antagonist was found at the Phe46 position. The aromatic ring of the residue in complex with the antagonist is tilted ~80° relative to that of Phe46 in complex with toluene in both molecules A and B (Fig. 2, B and D; see "Discussion" for details).

The residues critical for binding to ligand were further evaluated by ITC analysis using purified PAS1(23–168) and its mutant proteins, which exist as a dimer in solution (Fig. 2E). The PAS1 protein was titrated with toluene, and the ITC data were integrated and fitted into a "one set of binding sites" model, which yielded a stoichiometry of 1:1 and a binding affinity ($K_d$, dissociation constant) of 8.5 μM for each molecule in the PAS1 dimer (Fig. 2F). The binding affinity of toluene was significantly reduced with mutations at positions F46A and F79Y and was mildly reduced with mutations at V59A and W85H in PAS1 (Fig. 2F). The binding of toluene to the mutants revealed a ~1.5–11-fold weaker interaction than that of WT. The thermodynamic parameters of toluene binding to the proteins determined by ITC are summarized in Fig. 2F.

Strikingly, the structural conformation of the C-terminal region (residues 150–163), which is directly connected to the C-terminal kinase domain, differs according to which ligand is present. As described, in the apo-PAS1 structure, the helix is completely disordered in molecule B (Fig. 1A). In the PAS1 structure with toluene, the corresponding region is dramatically reconstructed into an α-helix (Fig. 2A). In the case of the antagonist 1,2,4-TMB, the equivalent region is completely disordered, and there was no electron density for amino acid residues after Lys155 in molecule B (Fig. 2C). It is important to note that molecule A in all structures exhibits an identical α-helix (α4) conformation in the C-terminal region (residues 150–163). These results imply molecular signal transduction to the C-terminal kinase domain of TodS by effector-dependent PAS1 sensing. Hereafter, we refer to the PAS1 C-terminal region (residues 150–163) as the signal transfer region (STR).

**In Vivo Validation of the Ligand-binding Residues**—We designed the *P. putida* KT2440-PXZ reporter strain, which was constructed by fusing the promoter region $P_{todX}$, the transcriptional regulator TodT-binding promoter, to lacZ encoding β-galactosidase for further validation of the ligand-binding residues, using a high-throughput assay system (Fig. 3A). The reporter strain was transformed with pBBR-$P_{todST}$-TodST and its variants, and their ligand-binding and multistep-signal relaying capacities were evaluated, either upon agonist (toluene, styrene, and m-xylene) or antagonist (1,2,4-TMB and o-xylene) binding. The *P. putida* KT2440-PXZ reporter strain harboring WT TodST or its mutant-expressing plasmid was grown at 30 °C at 500 rpm for 5 h in a gas phase saturated with each ligand supplied over a range of 10–100 μM. Maximum β-galactosidase activity for WT TodS was obtained in response to 10 μM toluene (Fig. 3B). However, other known agonists, such as styrene and m-xylene, were ineffective signal effectors compared with toluene in TodS/TodT signal transduction. Styrene (Fig. 3C) and m-xylene (Fig. 3D) did not activate TodS/TodT signaling below 100 μM and only showed maximum β-galactosidase activities at concentrations over 200 and 400 μM, respectively (data not shown). Thus, these results further indicate that toluene is the primary signal effector of the TodS/TodT signal transduction system (10, 11). No signal was transmitted with the antagonist o-xylene (Fig. 3E) or 1,2,4-TMB (Fig. 3F), consistent with previously reported results (10, 11).

The PAS1 hydrophobic pocket residues critical for signal sensing were evaluated in the presence or absence of agonists or antagonists. All mutations in the pocket, except for I114V, blocked signal relay from PAS1 to HK and ultimately to TodT and completely abolished enzyme activity in the toluene environment (Fig. 3B). The V126A mutant showed a response similar to that of the WT in the presence of 100 μM toluene (Fig. 3B). A similar pattern was observed with the weak agonist styrene, considering its responsiveness to PAS1. Interestingly, the I114V mutation increased sensitivity to 50 and 100 μM m-xylene, whereas the sensitivity of I114V mutant to styrene and toluene remained nearly the same as that of WT TodS (Fig. 3, B and C). As mentioned, m-xylene acts as a very weak agonist (responding at ~400 μM) for WT TodS. We further evaluated the mutant binding with toluene and m-xylene using ITC analysis (Fig. 3G). The binding affinity of the I114V mutant with toluene was similar (calculated as $K_d = 8.4$ μM) to that of WT PAS1 ($K_d = 8.5$ μM) and correlated with β-galactosidase activity (Fig. 3B). The binding affinity ($K_d$) of WT PAS1 with m-xylene was calculated to be 9.5 μM, which is comparable with that
Crystal Structure of the Toluene-sensing TodS PAS1 Domain

A

Molecule A

C

N

Molecule B

B

Molecule A

C

N

Molecule B

C

Molecule A

D

Molecule A

C

N

Molecule B

B

Molecule A

C

N

Molecule B

E

WT PAS1 (23-168) (38 kDa)

Bovine Serum Albumin (66 kDa)

Carbonic anhydrase (29 kDa)

Retention volume (mL)

0 4 6 8 10 12 14 16 18

Fraction No.

(kDa)

8 9 10 11 12 13 14 15

F

| Protein | $K_a$ (10$^{-3}$ M$^{-1}$) | $K_d$ (µM) | $\Delta H$ (kcal mol$^{-1}$) | $T\Delta S$ (kcal mol$^{-1}$) | $\Delta G$ (kcal mol$^{-1}$) |
|---------|-----------------|-------|-----------------|-----------------|-----------------|
| WT      | 118.0 ± 39.0    | 8.5   | -1.7 ± 0.04     | 5.2             | -6.9            |
| F46A    | 10.5 ± 7.7      | 95.2  | -2.0 ± 0.18     | 3.5             | -5.5            |
| V59A    | 68.4 ± 3.6      | 14.6  | -2.2 ± 0.04     | 4.4             | -6.6            |
| F79Y    | 14.0 ± 1.1      | 71.4  | -1.5 ± 0.10     | 4.2             | -5.7            |
| W85H    | 94.0 ± 8.3      | 10.6  | -1.8 ± 0.33     | 5.0             | -6.8            |

Time (min)

0 10 20 30 40 50 60 70 80

μcal/sec

WT PAS1 (23-168)

Molar Ratio

0 1 2 3 4 5 6 7 8 9 10

kcal/mole
Crystal Structure of the Toluene-sensing TodS PAS1 Domain

FIGURE 3. In vivo validation of PAS1 in the TodS/TodT signal transduction system. A, schematic diagram showing the high-throughput β-galactosidase assay system. The P. putida KT2440-PXZ strain expressing WT-TodS/TodT or its variant mutants was grown in 96-deep well plates in the presence of a gas phase ligand in the range of 10–400 μM for β-galactosidase induction. The activity of the induced β-galactosidase was measured using the hydrolyzed fluorescent substrate, 7-hydroxy-4-methylcoumarin, at 465 nm after a 1-h incubation with MUG substrate and calculated as MUG units. B–F, the PAS1 ligand-binding residues were evaluated for their ligand sensing and phospho-signal-relaying capacities, either upon agonist (toluene, styrene, and m-xylene) or antagonist (1,2,4-TMB and o-xylene) application in a gas phase saturated with each ligand supplied over a range of 10–100 μM. G, thermodynamic parameters of toluene or m-xylene binding to WT PAS1(23–168) and its mutant (I114V) proteins. The best fitting results were obtained with a “one set of binding sites” model using the ORIGIN software package (MicroCal). The heat data generated by the toluene or m-xylene addition to the reaction buffer were subtracted from the heat data generated from the reaction of each protein variant either with toluene or m-xylene. H, dimerization of PAS1. The dimeric structure of PAS1 was modeled on the structure of the A. vinelandii NifL LOV domain (Protein Data Bank code 2GJ3). Hydrophobic residues predicted to be involved in dimerization are displayed in green and gray in molecules A and B, respectively. I, the residues involved in dimerization were evaluated using the β-galactosidase assay system in a toluene environment. All results were obtained with three independent experiments. Error bars, S.D.

FIGURE 2. Toluene sensing by TodS PAS1. A, TodS PAS1 structure in complex with the agonist toluene. There are two toluene-bound PAS1 molecules in the asymmetric unit. Toluene molecules are shown as red carbon atoms. The C-terminal α-helix in molecule B corresponding to the disordered region in molecule B of apo-PAS1 is displayed in red. B, the toluene-binding pocket. The pocket is shown in black in an electrostatic surface representation. The detailed environment of toluene sensing by PAS1 is magnified for both molecules A and B. Toluene (red) is surrounded by hydrophobic PAS1 residues (green). C, PAS1 structure in complex with the antagonist 1,2,4-TMB. There are two antagonist-bound PAS1 molecules in the asymmetric unit. The 1,2,4-TMB-binding pocket is shown in black in an electrostatic surface representation. The detailed environment of 1,2,4-TMB-binding by PAS1 is magnified for both molecules A and B. Notably, 1,2,4-TMB (red) is surrounded by the same hydrophobic residues (green) as those in the toluene-binding pocket. The corresponding residues (yellow) responsible for forming the hydrophobic pocket in apo-PAS1 are superimposed onto the toluene- and 1,2,4-TMB-binding residues, respectively. Gly48 is not shown in B and D, because only the side chains of all residues are displayed. E, SEC analysis of purified His6-tagged PAS1(23–168). Eluted PAS1 was compared with the molecular weight standard markers BSA (66 kDa) and carbonic anhydrase (29 kDa) (left) and analyzed by SDS-PAGE (right). F, thermodynamic parameters of toluene-binding to PAS1(23–168) and its mutants. The best fitting results were obtained with a “one set of binding” sites model using the ORIGIN software package (MicroCal). The heat data generated by the toluene addition to the reaction buffer were subtracted from the heat data generated from the reaction of each protein variant with toluene. The typical ITC profile for the binding of toluene molecules to WT PAS1(23–168) is displayed.

with toluene ($K_D = 8.5 \mu M$). However, the binding affinity of $m$-xylene with the I114V mutant was 6.1 μM, which is 1.5-fold higher than that of the WT protein. The replacement of isoleucine with smaller side-chain-containing valine may induce better fit of $m$-xylene into the ligand-binding pocket. The mildly increased binding affinity of $m$-xylene to the mutant may enhance signal transduction, suggesting that Ile114 plays a role not only in ligand binding but also in ligand selectivity.

Dimerization of PAS1—As discussed above, the dimerization helix located in the immediate upper region (residues 32–43) of the core PAS domain was removed from the PAS1 domain (residues 43–168) for crystallization and resulted in an artificial antiparallel face-to-face dimer. Thus, we modeled the dimeric structure of PAS1 based on the structure (Protein Data Bank code 2GJ3) of the Azobacter vinelandii NifL LOV domain (43). Hydrophobic residues, such as Val37, Leu38, and Leu131, on the outside of the β-sheet in molecule A may interact with Ile39 and Leu40, His35 and Ile36, and Gly42, respectively, in molecule B, and are likely to be involved in dimerization and the maintenance of structural stability (Fig. 3H). The V47L, L49D, and...
L131D mutations completely abolished β-galactosidase activity in the toluene environment (Fig. 3I), suggesting that these residues are essential for PAS1 dimerization. By contrast, mutation of the charged residue Glu58 located outside of the β-sheet did not affect PAS1 dimerization (Fig. 3I).

Role of PAS2 in the TodS/TodT System—TodS comprises two modules, each containing a PAS-type sensor domain and an autokinase domain, and is thus classified as a “double sensor kinase” family member (9). As suggested by Lau et al. (7), the complexity of the TodS domain structure may provide the capacity for fine tuning by the two-component signal transduction system via integrating additional signals. The tod pathway is oxygen-dependent, and the PAS2 domain is not responsible for sensing aromatic signal molecules; therefore, they proposed that PAS2 might be able to sense oxygen (7). However, no direct evidence supporting this hypothesis has yet been reported.

To better understand the phosphorelay of TodS, we attempted to crystallize PAS2 to determine the structure and uncover its function. Although it was well purified and analyzed as a dimer in solution (Fig. 4A), crystallization failed. The sequence analysis predicted that PAS2 may have a similar structure to that of the A. vinelandii NifL LOV domain (Protein Data Bank code 2GJ3) (43). Thus, we modeled the structure of PAS2 (residues 611–729) using the structure of the NifL LOV domain as a template (Fig. 4B). Among the residues corresponding to the FAD (ligand) binding residues (Thr778 and Leu674) of NifL, Glu666 and Leu674 of PAS2 (Fig. 4C) were mutated to alanine, and their activity was compared with that of WT TodS in a β-galactosidase assay. The results showed no difference in β-galactosidase activity between the mutant variants and the WT (Fig. 4D). The structure of PAS1 complexed with toluene was superimposed onto the modeled structure (Fig. 4E). Then we tested residues Tyr691 and Ala703 in PAS2, corresponding to the critical toluene-binding residues Ile114 and Val126 in PAS1. The TodS mutants at the Tyr691 and Ala703 positions did not exhibit a toluene signal transfer capacity significantly different from that of WT TodS (Fig. 4D). These results are consistent with those of previous reports that the PAS2 domain does not function either as a toluene or FAD sensor (9, 10). It is worth mentioning that all results of the β-galactosidase activity assay in this study were obtained in culture medium containing limited ligand sources, in which PAS2 might not have the opportunity to bind to an unknown ligand(s). Thus, we do not rule out the possibility that PAS2 may bind to an unknown ligand(s) in order to regulate TodS/TodT signal transduction.
On the other hand, given that TodS functions as a dimer (see “SEC-MALS and TEM Analyses of TodS”), we hypothesized that dimerization of PAS2 is also important for tightly controlled phosphorelay. The modeled PAS2 structure revealed that the first α-helix surrounding residues Ser611–Ser622 (Fig. 4B) may be involved in PAS2 dimerization. Interestingly, deletion of the dimerization region of TodS (residues 617–623) significantly increased basal levels of β-galactosidase activity in the presence of weak agonists, such as m-xylene and styrene (note that these agonists at a concentration of 10 µM in a gas phase could not be sensed by WT TodS), as well as in the absence of ligand (Fig. 4D). The same results were obtained with TodS with a double mutation (Y619A and E620A) (Fig. 4D). These results indicate that maintaining the correct dimerization of PAS2 of TodS is critical for finely tuning phosphorelay to the C-terminal HK2 and TodT.

SEC-MALS and TEM Analyses of TodS—We noticed that full-length TodS protein was unstable and that it aggregated easily. It was therefore difficult to purify and use in biochemical studies in vitro. The C-terminal His₆-tagged TodS(43–978) and the N-terminal GST-fused TodS(23–978) protein were, however, relatively soluble and could be purified. SEC analysis showed that purified TodS(23–978) protein was eluted in the same fraction as the size marker protein apoferritin (443 kDa), indirectly indicating its flexible properties (Fig. 5A). We further employed SEC-MALS to determine the absolute molecular mass of TodS(23–978) in solution based on the angular dependence of scattered light intensity, which is independent of the molecular shapes. The retention volume (~12 ml) of TodS(23–
Crystal Structure of the Toluene-sensing TodS PAS1 Domain

978) in SEC-MALS analysis was similar to the SEC result of TodS (23–978) (Fig. 5A); however, MALS analysis indicated that the molecular mass of this fraction was 233.1 ± 5.99 kDa (Fig. 5B), which corresponds to dimeric TodS (23–978). It is worth noting that aggregation of TodS (23–978) was observed during SEC-MALS analysis (Fig. 5B). Collectively, these results suggest that the TodS (23–978) protein exists primarily as a flexible dimeric structure in solution.

To define the molecular arrangements of TodS proteins, we performed immunogold labeling electron microscopy analysis. The neighboring position of the Ni-NTA gold particles, as shown in Fig. 5C, corresponding to the position of the C-terminal His6-tagged region of the TodS (43–978) proteins, indicated a parallel arrangement of each protein in a dimeric assembly (Fig. 5C). Some monomeric gold particle labeling of TodS (Fig. 5C) may be caused by the single focal plane of the image presented or artificial assemblies by the TEM sample preparation procedure.

Furthermore, averaged images taken from negatively stained single particles of apo-TodS (23–978) protein provided information about the characteristics of TodS in the presence of different ligands. The apo-TodS molecule had a length of 14.2 nm and a width of 7.0 nm (Fig. 5D, top). In the structure, toluene sensing by PAS1 induced a structural change (especially residues 150–163) and affected TodS/TodT signal transduction. Thus, we assumed that the overall features of TodS might be affected by the presence of toluene. The toluene-sensing TodS proteins had a similarly shaped appearance to that of apo-TodS at the molecular level; however, they had a length and a width of 16.2 and 9.4 nm, respectively, demonstrating a structural change to a more straightened configuration (Fig. 5D, middle). The ligand-dependent conformational change of TodS was also supported by the changed overall features of TodS complexed with the antagonist 1,2,4-TMB, which showed mainly oligomerization (Fig. 5D, bottom). Thus, these results suggest that TodS undergoes ligand-dependent structural changes that are responsible for the activation of signal transduction.

Discussion

The P. putida tod operon, comprising genes encoding enzymes that catabolize toluene, is tightly regulated by TodS/TodT two-component multistep-signal transduction in a ligand-dependent manner (9–11, 45). The N-terminal PAS1 domain in TodS acts as a sensor for benzene-derived hydrocarbons with high affinity, whereas the C-terminal PAS2 domain lacks a signal-sensing function (9–11, 45).

In this study, we determined the structure of the PAS1 sensor as an apoprotein and as a complex with either with toluene (an agonist) or 1,2,4-TMB (an antagonist) to explore the molecular mechanisms underlying how PAS1 senses effector molecules and delivers signals to C-terminal HK to regulate TodS/TodT signal transduction.

Unlike most PAS homologues that bind cofactors, such as FAD and FMN, the TodS PAS1 binds hydrophobic benzene-derived compounds. Accordingly, the ligand-binding pocket of PAS1 is surrounded by completely hydrophobic residues. The phototropin-2 LOV1 domain (Fig. 1B), which has the structure most similar to that of PAS1, regulates light sensitivity through binding to FMN (42). The binding pocket for FMN in LOV1 is greater than that for toluene in TodS PAS1 and is surrounded by polar residues. Bulky and hydrophobic residues, including Trp64, Trp85, Phe46, Phe79, and Phe128, are involved in the interaction with toluene in TodS PAS1, whereas charged and polar residues, including Arg133, Asn136, Asn164, Glu265, and Glu272, are associated with FMN in LOV1. Such properties of the ligand-binding pocket of PAS1 are consistent with a recent study showing that TodS does not bind cofactors involved in redox sensing, including FAD, FMN, heme, NAD, NADH, NADP, and NADPH (46).

The structural comparison suggests that the residues ranging from 150 to 163 (STR) in PAS1 are essential for relaying phospho-signals to HK1, RRR, and HK2. The dramatic conformational change in STR from a flexible conformation to an α-helix seems to be controlled by the aromatic ring of Phe46 (Fig. 2, B and D). According to the type of ligand that PAS1 senses, Phe46 may be positioned differently to turn on or off signal transfer. In molecule A, the Phe46 aromatic ring is tilted ~80° in response to the antagonist 1,2,4-TMB relative to that of Phe46 either in apo-PAS1 or in complex with toluene. This change led to the formation of a salt bridge between two amino acid residues, Arg148 and Glu146, on the β5 strand, whereas no such bridge was observed in either apo- or agonist-bound PAS1 (Fig. 6A). Nevertheless, molecule A in all structures maintains the STR region as an α4 helix (Fig. 6A). By contrast, in molecule B, tilting of the Phe46 aromatic ring occurred only in response to the agonist toluene, and this was linked to the formation of a salt bridge between Arg48 and Glu146 and the transformation of the flexible STR region (in apo-PAS1) to an α4 helix (Fig. 6B). Interestingly, residue Arg48 in molecule B of PAS1 complexed with the antagonist 1,2,4-TMB is very flexible and shows no visible electron density, whereas the electron density map of the corresponding residues in other structures is seen clearly (Fig. 6C). The phenomenon is probably related to the flexibility and lack of electron density in the STR region of molecule B of PAS1 complexed with 1,2,4-TMB (Figs. 2C and 6B). It is not clear whether the conformational changes are mutually affected by ligand binding to molecules A and B. The signal switching of Phe46 upon binding different types of ligands is summarized in Fig. 6D. We further assessed the importance of the Glu146 and Arg148 residues in signal transfer. In the β-galactosidase assay systems, E146A and R148A (or R148M) TodS mutations completely abolished β-galactosidase activity even in the presence of 100 μM toluene, whereas WT TodS responded to >10 μM toluene in a gas phase (Fig. 6E). The toluene-binding affinity of the E146A mutant dimer (Kd = 9.6 μM) was very similar to that of WT PAS1 (Fig. 6F). Collectively, these results suggest that the Phe46 residue of PAS1 acts as a molecular switch to selectively control TodS/TodT signaling. Another member of the TodS-like family, TmoS, also binds to and distinguishes a wide range of aromatic carbons either as an agonist or as an antagonist (16). In fact, amino acid sequence alignments of TodS-like family sensor kinases, such as TodS, TmoS, and StyS, revealed that residues Phe46, Glu146, and Arg148 are strictly conserved (16). Thus, it could be hypothesized that Phe46 may also play a key role in the
ligand sensing of TmoS. Biochemical characterizations of Phe46 residues in TodS-like family members should be further investigated.

During refinement of the apo-PAS1 structure, we noted a region of weak electron density in the hydrophobic ligand-binding pocket. The shape of the region is somewhat similar to those of benzene-derived compounds, including toluene and 1,2,4-TMB (Fig. 7A). It is important to note that volatile organic compounds, including benzene derivatives, are common chemical contaminants that readily evaporate to some extent in air, depending on the environmental conditions. Particularly, laboratory environments are exposed to a wide variety of chemicals and regularly contain volatile organic compounds (47), even in incubators in which protein crystals are grown. Thus, we thought that the region of weak electron density observed in the apo-PAS1 structure may have originated from contaminants in the incubator (or the laboratory environment). It may also be possible that a cellular ligand, such as an aromatic compound, was co-purified with PAS1 protein. Nonetheless, the level of contamination was probably not sufficiently high to affect PAS1-mediated TodS signaling. It is worth mentioning that PAS1 was crystallized with different ligands in separate incubators to avoid cross-contamination.

Our results showed that maintaining the correct dimerization of PAS2 is crucial to finely tune phosphoryl. PAS2 exists as a dimer in solution (Fig. 4A), and the region (residues 617–623) of PAS2 involved in dimerization appears to be important for tight regulation of the basal levels of β-galactosidase activity in ligand-specific TodS/TodT signal transduction (Fig. 4). The Krell group constructed a minimal form of TodS, termed Min-TodS, which contains only the N-terminal PAS1 and the C-terminal HK2 of TodS, by removal of the central HK1, RRR, and PAS2 domains (45). They demonstrated that Min-TodS binds effector molecules with affinities similar to those observed for WT TodS. They also showed that both Min-TodS and WT TodS can classify effector molecules into agonists and antagonists, indicating that the molecular determinants of effector recognition and their agonistic or antagonistic action are located in the PAS1 domain (45). However, they observed much higher basal activity of Min-TodS, whereas WT TodS was almost silent in the absence of toluene. Busch et al. (9) reported that TodS mutants, in which the phosphoryl-accepting amino acids His190, Asp500, and His760 are replaced with alanine, exhibit low basal activity that is comparable with the WT TodS system, suggesting that the low basal activity occurs irrespective of whether there is functional phosphoryl. Thus, these data suggest that all domains within TodS are required to maintain the low basal activity of the TodS system for the final accurate response of TodT because tod operon expression in the absence of effector molecules is useless.

Finally, taken together with the results of previous structural, biological, and in vivo studies, the results of TEM analysis, which revealed changes in TodS into aligned or misaligned
structures, depending on the ligand employed, allowed us to propose an outline of the molecular mechanism of TodS/TodT signal transduction (Fig. 7, B and C). Structural flexibility between the sensor PAS1 and HK1 as well as the dimeric features of TodS are essential for efficient signal relay under certain environmental conditions. Toluene binding to the PAS1 sensor domain induces conformational changes in TodS that result in a functional form that is proficient for phospho-signal relay (Fig. 7, B and C). All structures of the STR of molecule A in both apo- and ligand-bound PAS1 exhibited a well structured α-helix, whereas the corresponding region in molecule B was structured differently (in all structures), depending on which ligand was employed. In the absence of a signal effector, the PAS1 sensor domain cannot provide a signal to HK1 via the flexible nature of STR. Thus, this conformation could not induce the functional dimeric conformation of HK1 to enable successful autophosphorylation. Upon toluene sensing at an effective level, the PAS1 STRs may be reorganized to transmit signals and induce conformational changes in TodS to align the HK1-RRR-PAS2-HK2 domains for efficient multistep phosphorelay. See “Discussion” for a full description. Molecules A and B in the PAS1 dimer are shown in light green and dark green, respectively. TodT dimers are displayed in light and dark yellow. Toluene molecules are shown in red.

Figure 7. Proposed molecular mechanism of TodS/TodT signal transduction. A, the Fo – Fc electron density map of apo-PAS1 contoured at 3.0 σ (left), corresponding to the maps for toluene contoured at 3.0 σ (center) and 1,2,4-TMB contoured at 2.5 σ (right) of PAS1-ligand complexes, respectively. B and C, proposed models of TodS/TodT signal transduction. TodS exists as a dimer with a flexible nature, which might possess a basal level of autokinase activity. In this condition, the PAS1 sensor domain would not effectively deliver signals to the C-terminal HK1 via the flexible nature of STR. Thus, this conformation could not induce the functional dimeric conformation of HK1 to enable successful autophosphorylation. Upon toluene sensing at an effective level, the PAS1 STRs may be reorganized to transmit signals and induce conformational changes in TodS to align the HK1-RRR-PAS2-HK2 domains for efficient multistep phosphorelay.

Author Contributions—S. K. performed most of the biochemical experiments. J. H., K. G., and M. H. K. carried out the structural study. E.-G. L., S.-Y. K., and O. K. constructed the P. putida reporter strain. S. L., J. M. C., and H. S. J. performed EM analysis. S. K., S. J. L., C.-M. R., S.-G. L., and T.-K. O. contributed to the discussion and provided reagents. S. K., J. H., and M. H. K. designed the study and wrote the manuscript. All authors helped with data analysis.

Acknowledgments—We thank the beamline staff at Pohang Light Source (beamline 5C) (Pohang, Korea) and the Photon Factory (beamline 17A) (Tsukuba, Japan) for help with data collection. We are very grateful to Dr. Hyeon Ho Park and Jae Young Choi (Yeungnam University, Korea) for help with SEC-MALS analysis. We also appreciate the kind advice from Dr. Verna Frasca (MicroCal) about ITC analysis.
Crystal Structure of the Toluene-sensing TodS PAS1 Domain

41. Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M., and Hegemann, P. (2003) Crystal structures and molecular mechanism of a light-induced signaling switch: the Phot-LOV1 domain from Chlamydomonas reinhardtii. Biophys. J. 84, 2474–2482
42. Nakasako, M., Zikihara, K., Matsuoka, D., Katsura, H., and Tokutomi, S. (2008) Structural basis of the LOV1 dimerization of Arabidopsis phototropins 1 and 2. J. Mol. Biol. 381, 718–733
43. Key, J., Hefti, M., Purcell, E. B., and Moffat, K. (2007) Structure of the redox sensor domain of Azotobacter vinelandii NifL at atomic resolution: signaling, dimerization, and mechanism. Biochemistry 46, 3614–3623
44. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797
45. Silva-Jiménez, H., Ramos, J. L., and Krell, T. (2012) Construction of a prototype two-component system from the phosphorelay system TodS/TodT. Protein Eng. Des. Sel. 25, 159–169
46. Silva-Jiménez, H., Ortega, Á., García-Fontana, C., Ramos, J. L., and Krell, T. (2015) Multiple signals modulate the activity of the complex sensor kinase TodS. Microb. Biotechnol. 8, 103–115
47. Cohen, J., Gilligan, A., Esposito, W., Schimmel, T., and Dale, B. (1997) Ambient air and its potential effects on conception in vitro. Hum. Reprod. 12, 1742–1749