Solution Structure of a Low-Molecular-Weight Protein Tyrosine Phosphatase from *Bacillus subtilis*

Huimin Xu,1,2 Bin Xia,1,2,3 and Changwen Jin1,2,3*

Beijing Nuclear Magnetic Resonance Center1 and College of Life Sciences2 and College of Chemistry and Molecular Engineering,3 Peking University, Beijing 100871, China

Received 18 July 2005/Accepted 14 November 2005

The low-molecular-weight (LMW) protein tyrosine phosphatases (PTPs) exist ubiquitously in prokaryotes and eukaryotes and play important roles in cellular processes. We report here the solution structure of YwlE, an LMW PTP identified from the gram-positive bacteria *Bacillus subtilis*. YwlE consists of a twisted central four-stranded parallel β-sheet with seven α-helices packing on both sides. Similar to LMW PTPs from other organisms, the conformation of the YwlE active site is favorable for phosphotyrosine binding, indicating that it may share a common catalytic mechanism in the hydrolysis of phosphate on tyrosine residue in proteins.

Though the overall structure resembles that of the eukaryotic LMW PTPs, significant differences were observed around the active site. Residue Asp115 is likely interacting with residue Arg13 through electrostatic interaction or hydrogen bond interaction to stabilize the conformation of the active cavity, which may be a unique character of bacterial LMW PTPs. Residues in the loop region from Phe40 to Thr48 forming a wall of the active cavity are more flexible than those in other regions. Ala41 and Gly45 are located near the active cavity and form a noncharged surface around it. These unique properties demonstrate that this loop may be involved in interaction with specific substrates. In addition, the results from spin relaxation experiments elucidate further insights into the mobility of the active site. The solution structure in combination with the backbone dynamics provides insights into the mechanism of substrate specificity of bacterial LMW PTPs.

Protein phosphorylation and dephosphorylation are involved in the regulation of many cellular processes, such as cell growth, differentiation, and metabolism (12, 23, 35). The formation and hydrolysis of phosphate esters in proteins play key roles in the signal transduction through which the external environmental stimuli are converted to internal cellular response (35, 37). Protein tyrosine kinase (PTK) and protein tyrosine phosphatase (PTP) catalyze such reversible reactions. The cellular activities and cell growth are under the precise control of PTKs and PTPs. Defective or incorrect regulations in such systems result in many human diseases and thus they are ideal targets for drug design (24, 55).

PTPs belong to the protein phosphatase superfamily and catalyze the hydrolysis of phosphate ester on tyrosine residue in proteins. Based on their molecular masses and biological activities, PTPs can be divided into three families: the high-molecular-weight (HMW) PTP family, whose members contain a conserved catalytic domain with a molecular mass of around 30 kDa; the dual-specificity PTP family, which hydrolyzes phosphoserine or phosphothreonine as well as phosphotyrosine; and the low-molecular-weight (LMW) PTP family proteins, which are present in the cytoplasm with a molecular mass of around 18 kDa (11, 40, 50, 51). All these families of PTPs share very low sequence identities except the signature motif of the active site, C(X)5R(S/T), at the phosphate-binding site. In PTP families, this motif is a relatively flexible loop and is called the protein tyrosine phosphate-binding loop, or P-loop (51).

It is well known that LMW PTPs down-regulate the platelet-derived growth factor (PDGF)-stimulated cell proliferation by dephosphorylating the PDGF receptor in eukaryotes (1, 4). Phosphorylation on the tyrosine residue of LMW PTP itself in the process of PDGF signal transduction is involved in cytoskeleton rearrangement and cell adhesion (3, 5).

To date, many microbial LMW PTPs, such as from *Acinetobacter johnsonii*, *Escherichia coli*, *Klebsiella pneumoniae*, *Mycobacterium tuberculosis*, and *Staphylococcus aureus*, have been identified (19, 27, 39, 43, 46). Enzyme assays indicated that these LMW PTPs could specifically catalyze the hydrolysis of phosphotyrosine. In *K. pneumoniae*, the specificity was also confirmed by the observation that the LMW PTP could dephosphorylate an endogenous kinase that was autophosphorylated on tyrosine (39). The biological functions in cells are still poorly understood, though some functional roles have been proposed. The first identified yeast LMW PTP, Stp1, was found to induce a severe growth-defective phenotype and was considered to be involved in the regulation of cell growth (31).

The LMW PTPs in prokaryotes were proposed to be involved in the production and transport of exopolysaccharides and capsular polysaccharides, which are important virulence factors of pathogenic microbes (8, 18, 39). In several bacterial species, PTKs and LMW PTPs constitute a pair and PTK is the specific endogenous substrate of LMW PTP, such as Yco6/Yor5 from *K. pneumoniae*, Wzc/Wzb from *E. coli*, and Ptk/Ptp from *A. johnsonii* (19, 39, 46). The encoding genes of the PTK and LMW PTPs are located in the same cluster that consists of genes related to biosynthesis of exopolysaccharide (8, 45).

Biochemical and structural studies have demonstrated that the eukaryotic LMW PTPs share a similar catalytic mechanism, and that common steps are involved in the process of
dephosphorylation (44, 48, 52). Bacterial LMW PTPs share low sequence identities with mammalian PTPs, whereas they exhibit high structural similarities.

The kinetics of the Bacillus subtilis PTK YwqD, the PTP YwqE, and the LMW PTPs YwlE and YfkJ were characterized (32, 33). The studies indicated that YwlE shows similar activity towards the substrate p-nitrophenylphosphate (pNPP), but a lower activity on the protein substrate of YwqE. A recent genetic and biochemical study suggested that YwlE plays roles in B. subtilis stress resistance (36). In order to explore the structural basis of bioactivity and further understand the binding mechanism of its specific endogenous substrate, we here report the solution structure and dynamic and kinetic properties of an LMW PTP, YwlE from B. subtilis. A detailed structural comparison with other LMW PTPs provides insights into the catalytic mechanism, especially the interactions with the specific substrates.

MATERIALS AND METHODS

Sample preparation. The ywlE gene from B. subtilis was cloned into vector pET21a (Novagen). Uniformly $^{15}$N/13C-labeled and $^{15}$N-labeled proteins were overexpressed in E. coli strain BL21(DE3) at 37°C in M9 minimal medium containing $^{15}$N-labeled ammonium chloride in the presence or absence of $^{13}$C-labeled glucose for $^{15}$N-labeled and $^{13}$C-labeled or $^{15}$N-labeled samples, respect-

FIG. 1. Sequence alignment of LMW PTPs with B. subtilis YwlE, Mycobacterium tuberculosis MptpA (PDB code 1U2P), Saccharomyces cerevisiae Ltp1 (PDB code 1D1P), bovine PTP (BPTP; PDB code 1PNT), and human HCPTPA (PDB code 5PNT) by the program CLUSTAL W (22). The conserved residues in the P-loop and others mentioned in the text are highlighted in boxes. The protein sequences were obtained from the PDB database.

FIG. 2. Solution structure of B. subtilis YwlE. (A). The stereo view diagram of the superimposition of 20 representative structures (blue) with the mean structure (orange) of YwlE. (B). The ribbon diagram of the mean structure shows the secondary-structure elements. The P-loop is in blue.
The soluble protein was purified by anion-exchange chromatography (Mono Q column) followed by gel filtration (Superdex-75 column) with the AKTA fast protein liquid chromatography (FPLC) system (Pharmacia). The protein purity is above 95% as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (PAGE). Nuclear magnetic resonance (NMR) samples were prepared in 50 mM Tris-HCl buffer (pH 7.5) and 50 mM sodium phosphate buffer (pH 7.5), respectively. Two-dimensional 1H-edited heteronuclear single quantum coherence (HSQC) spectra showed identical numbers but different relative intensities of cross peaks in the absence and presence of phosphate. The NMR samples for structure determinations, while samples with different concentrations (1.0, 0.5, 0.25, and 0.1 mM) were prepared for dynamics studies.

**NMR spectroscopy.** All NMR experiments were performed at 25°C on Bruker Avance 500-MHz (with cryoprobe) and 800-MHz spectrometers equipped with triple-resonance probes and pulsed-field gradients. The spectra were processed using software package NMRPipe (10) and analyzed using NMRView (25). Nearly complete backbone and side chain resonance assignments have been obtained and reported previously (49). Briefly, three-dimensional triple-resonance spectra HNCA, HN(CO)CA, CBCA(CO)NH, HBBa(CO)NH, HNCO, and HN(CA)CO were collected to obtain the backbone resonance assignments. The three-dimensional spectra H(CO)CO-NH-total correlation spectroscopy (TOCSY), (H)CC(CO)NH-TOCSY, HCCH-TOCSY, CCH-TOCSY, HCCH-correlation spectroscopy (COSY), HNC-COSY, and 15N-edited TOCSY-HSQC were recorded for the side chain resonance assignments. The three-dimensional 15N- and 1H-Edited nuclear Overhauser effect spectroscopy (NOESY)-HSQC (mixing times of 100 and 50 ms) were examined to confirm the assignments and generate distance restraints for structure calculations. The hydrogen-deuterium (H-D) exchange experiments were performed to obtain the hydrogen-bonding information.

**Structure calculations.** The structure was determined using NOE-derived distance restraints, in combination with dihedral angle restraints and hydrogen bond restraints. The backbone dihedral angle restraints were determined using the program TALOS (7). Hydrogen bond distance restraints were generated from the H-D exchange experiments in conjunction with the intermediate-range NOE-s and the secondary-structure information (41). The initial structures of YwlE were calculated using the module CANDID of the program CYANA by using NOE-derived distance restraints and dihedral angle restraints (20, 21). First, 50 structures were calculated and the 10 lowest-energy structures were selected as the filter models. The program SANE was employed to obtain further assignments semiautomatically from three-dimensional NOESY spectra based on the chemical shift assignments and predicted secondary-structure information (14).

The subsequently regenerated distance restraints by SANE were used as input for the structure calculations by the program CYANA. A total of 200 structures were calculated using CYANA. The 100 structures with the lowest target function values were selected and refined by AMBER using the force field 899 and Generalized Born solvation model (38). The same restraints for the last run of CYANA were used. A 30-picosecond simulated annealing from 1,000 to 0 K was performed. The force constants for the NOE distance, dihedral angle, and hydrogen bond restraints were 20 kcal mol⁻¹ Å⁻², 50 kcal mol⁻¹ rad⁻² and 40 kcal mol⁻¹ Å⁻², respectively. The chirality and transpeptide ω restraints with the force constants 150 kcal mol⁻¹ rad⁻² and 50 kcal mol⁻¹ rad⁻², respectively, were added to prevent chirality inversions or peptide bond flips during the high-temperature portions of simulated annealing (22).

During the structure calculations, several iterations of SANE-CYANA-AMBER were performed in conjunction with manual variation analysis. Finally, 20 structures with the lowest AMBER energies were selected to represent YwlE. The mean structure of the 20 structures was obtained by a 1,000-cycle energy minimization using AMBER7. The program PROCHECK-NMR was used to analyze the stereochemical geometry by Ramachandran statistics (28).

**PTP activity assay.** The phosphatase activity assays using pNPP as the substrate were performed in a total volume of 50 μl containing 0.1 mM sodium citrate (pH 6.5), 5 mM DTT, 1 mM EDTA, and 20 μg YwlE at the optimized temperature of 55°C. The initial rate of the reaction was measured by calculating the concentration of released 3-nitrophenolate (pNP) ion using an absorbance at 405 nm. The concentrations of pNPP used in the reaction ranged from 2 to 70 mM. The total reaction time was 4 min. The millimolar extinction coefficient was 18.

**Titration experiments.** The protein sample (0.5 mM) was prepared in 50 mM Tris-HCl buffer containing 50 mM NaCl and 30 mM DTT at pH 7.5. Sodium phosphate was added at concentrations ranging from 1 to 50 mM. The phospho-binding site was determined using the chemical shift changes from a series of two-dimensional 15N-edited HSQC spectra at 25°C. The composite chemical shift changes of amide 15N and H were calculated using δH/δN = (δH/δN)comp = (δH/δN)ref + (δH/δN)dec + (δH/δN)exc, where (δH/δN)ref is the 1H, 15N, and composite chemical shift changes of the cross peaks in the HSQC spectra, respectively (34). The mean value of 20 structures with the lowest target function values out of 100 structures calculated with CYANA. The backbone heavy atoms include Cα, amide nitrogen atom, and carbonyl carbon atom. All heavy atoms include all nonhydrogen atoms in the protein.

### Table 1. Structural statistics of B. subtilis YwlE

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of NOE-derived proton-proton distance restraints</td>
<td>Intrasidechain unambiguous NOEs&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Sequential unambiguous NOEs&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Medium-range unambiguous NOEs&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Long-range unambiguous NOEs&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Total unambiguous NOEs</td>
</tr>
<tr>
<td></td>
<td>Total ambiguous NOEs&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>No. of H-bond distance restraints</td>
<td>48</td>
</tr>
<tr>
<td>No. of Dihedral angle restraints (φ and ψ)</td>
<td>162</td>
</tr>
<tr>
<td>No. of Restraint violations (20 structures)</td>
<td>NOE-derived distance violation (&gt;0.2 Å)</td>
</tr>
<tr>
<td></td>
<td>Torsion angle violation (&gt;5°)</td>
</tr>
<tr>
<td>Energy</td>
<td>CYANA target function&lt;sup&gt;f&lt;/sup&gt; (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
</tr>
<tr>
<td></td>
<td>Mean AMBER energy (kcal mol⁻¹)</td>
</tr>
<tr>
<td></td>
<td>NOE distance restraint violation energy (kcal mol⁻¹)</td>
</tr>
<tr>
<td></td>
<td>Torsion angle restraint violation energy (kcal mol⁻¹)</td>
</tr>
<tr>
<td>RMSD from mean structure (Å)</td>
<td>Backbone heavy atoms</td>
</tr>
<tr>
<td></td>
<td>Secondary structures</td>
</tr>
<tr>
<td></td>
<td>All heavy atoms&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>Secondary structures</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td>Residues in most favored regions (%)</td>
</tr>
<tr>
<td></td>
<td>Residues in additional allowed regions (%)</td>
</tr>
<tr>
<td></td>
<td>Residues in generously allowed regions (%)</td>
</tr>
<tr>
<td></td>
<td>Residues in disallowed regions (%)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Intrasidechain NOEs represent the distance restraint between two protons in the same residue.

<sup>b</sup> Sequential represents the distance restraint between two protons that belong to sequential residues.

<sup>c</sup> Medium range represents two protons belonging to residues separated by no more than five amino acids.

<sup>d</sup> Long range represents protons belonging to residues separated by more than five amino acids.

<sup>e</sup> Ambiguous NOEs represent NOE-derived distance restraints that could be assigned to more than one pair of protons.

<sup>f</sup> The mean value of 20 structures with the lowest target function values out of 100 structures calculated with CYANA.

<sup>a</sup> The backbone heavy atoms include Cα, amide nitrogen atom, and carbonyl carbon atom.

<sup>b</sup> All heavy atoms include all nonhydrogen atoms in the protein.
160, and 180 ms, and 64, 64, 16, and 16 transients were used at the corresponding sample concentrations as for the R1 measurements. The relaxation rate constants were determined by fitting the peak intensities using the single exponential functions (17). The (1H)-15N NOE experiments were performed in the presence and absence of a proton presaturation period of 3 seconds prior to the 15N excitation pulse and using recycle delays of 3.7 and 6.7 seconds, respectively (16). The numbers of transient used were 128, 96, 64, and 32 for the sample concentrations as in R1 and R2 experiments.

Protein structure accession numbers. The atomic coordinates of Bacillus subtilis YwlE have been deposited in the RCSB under PDB code 1ZGG. The chemical shift assignments have been deposited with BMRB accession number 6460.

RESULTS AND DISCUSSION

Solution structure of YwlE. B. subtilis YwlE contains 150 amino acid residues and the sequence alignment of YwlE with other PTPs is shown in Fig. 1. The solution structure of YwlE has been determined using interproton NOE-derived distance restraints in combination with the dihedral angle restraints and hydrogen bond restraints. The 20 lowest-energy structures were selected to represent YwlE and are shown in Fig. 2A, together with the ribbon diagram showing the secondary-structure elements in Fig. 2B. The structural statistics are shown in Table 1. The root-mean-square deviation (RMSD) of backbone heavy atoms from the mean structure is 0.38 ± 0.07 Å, 0.26 ± 0.04 Å for those in regular secondary structures. Only two NOE distance restraints show violation greater than 0.2 Å and no angle restraints show violation greater than 5°. For the 20 structures, 87.2% of the residues are in the most favored regions of the Ramachandran plot using PROCHECK_NMR (28), 11.2% are in the additionally allowed regions, 1.0% are in generously allowed regions, and 0.6% are in disallowed regions. The residue in the disallowed region is Asn10 in the P-loop, which is relatively flexible. Only few NOE-derived distance restraints were obtained for this residue.

YwlE comprises four β-strands (residues 2 to 6 [β1], 33 to 37 [β2], 79 to 82 [β3], and 102 to 104 [β4]), which form a twisted central parallel β-sheet, with six α-helices (residues 14 to 28 [α1], 51 to 57 [α2], 71 to 76 [α3], 85 to 95 [α4], 107 to 110 [α6], and 124 to 145 [α7]) and a 310 helix (residues 98 to 101 [α5]) flanking both sides. The highly conserved signature sequence from residues Cys7 to Ser14 forms the P-loop connecting β1 and α1. Although YwlE shares a low sequence identity with LMW PTPs of known structures (26% with human HCPTPA, 29% with bovine PTP [BPTP], 23% with Saccharomyces cerevisiae Ltp1, and 31% with Mycobacterium tuberculosis MptpA), the core packing of YwlE is highly similar to that of those PTPs. The RMSD values for the Cα atom in the regular secondary-structure elements of YwlE from those of HCPTPA, BPTP, Ltp1, and MptpA are 1.32, 1.26, 1.24, and 1.22 Å, respectively.

Mapping of the phosphate-binding site. The phosphate ion is a competitive inhibitor of LMW PTPs when pNPP is used as a substrate. To map the inhibitor-binding site, phosphate ion
titration experiments were performed and monitored by a series of two-dimensional $^1$H-15N HSQC spectra. Significant chemical shift changes were observed for residues Cys7, Asn10, Ser14, Met16, Glu18, Gly38, Val39, Ala41, Ser67, Leu69, Thr84, Val116, Leu117, Asp118, and Phe120. As shown in Fig. 3, these residues are located in the P-loop and neighboring regions surrounding the cavity formed by the P-loop. Among them, Asp118 is the putative proton donor, a key catalytic residue for the enzyme reaction. The results are in general agreement with similar observations on other LMW PTPs, indicating that the P-loop region of YwlE is the catalytic site (29, 48).

**Active site.** The active site of YwlE contains the P-loop and an aspartic acid residue, Asp118, that is nearby. The amino acid sequence of the P-loop in YwlE is CTGNTCRS, in which residues Cys7, Gly9, Asn10, Cys12, Arg13, and Ser14 are conserved in all LMW PTPs (Fig. 1). As shown in Fig. 4, most of the backbone N-H bonds in the P-loop point to the center of the phosphate-binding pocket, and Arg13 is located at the bottom of this pocket. This conformation suggests that strong hydrogen bond interactions might occur between the amide protons and the phosphate oxygen atoms of substrate.

The sequence alignments showed that Glu18 and His64 of...
YwlE were conserved among LMW PTPs with known structures (Fig. 1). Based on the structure of YwlE, Glu18 is close to the imidazole ring of His64 and the distance between O1 of Glu18 and N6 of His64 is 3.71 ± 0.68 Å, and it is 3.72 ± 0.80 Å between O2 of Glu18 and N8 of His64. The spatial locations of Glu18 and His64 suggest that the electrostatic interaction or hydrogen bond interaction might occur between the two side chains and stabilize the conformation of the active cavity for binding with the phosphate group of the substrate. As shown in Fig. 4, the side chain hydroxide group of Asp118 is located at the top of the pocket, and the catalytic thiol of Cys7 is located at the bottom, which is geometrically similar to other LMW PTPs.

The structure of the active site provides insights into the enzyme activity, which was further confirmed by the enzyme activity assay. Previous studies demonstrated that both C7S and D118A mutations completely abolished enzyme activity (36). The highly conserved sequences and low RMSD values of Cα atoms of YwlE and those of HCPTPA, BPTP, Ltp1, and MptPA in the P-loop region (0.79, 0.83, 0.86, and 0.90 Å, respectively) further support that YwlE may share a similar catalytic mechanism with those LMW PTPs.

Residue Arg13 of YwlE in the P-loop is a signature of all LMW PTPs and plays an important role in enzyme activity. Previous studies indicated that the mutation R13K inactivated the enzyme completely (36). The corresponding residue in BPTP, Arg18, was proposed to be responsible for binding with the phosphate group and mutation of R18A abolished enzyme activity completely (9). Based on the structure of BPTP, the side chain atoms of Arg18 along with the backbone of the P-loop form a complete circle in which all amide N-H bonds and N-H of Arg18 point to the center of the active cavity (53). The distances between Sα of Cys12 and the amide nitrogen atoms of P-loop residues and Nα of Arg18 are no more than 4.8 Å. However, the positions of the side chain atoms of Arg13 in YwlE are relatively far from the sulfur atom of Cys7. The distance between Sα of Cys7 and Nα of Arg13 is 6.54 ± 0.53 Å. No significant chemical shift perturbations of the side chain Nα and Hα of Arg13 were observed upon phosphate binding from the titration experiments. These results imply that Arg13 may not directly participate in the phosphate binding process in YwlE.

From the structure of YwlE, Asp115 is located near the guanidinium group of Arg13. The distance between Nα of Arg13 and O2 of Asp115 is 2.85 ± 0.52 Å, and that between Nε1 of Arg13 and Oε1 of Asp115 is 2.76 ± 0.50 Å. It is possible that the electrostatic or hydrogen bond interaction occurs between the guanidinium group of Arg13 and the carboxylate group of Asp115. The local conformation of the active cavity may be stabilized by the salt-bridge or hydrogen bond interaction between them. In all identified bacterial LMW PTPs, residue Asp115 is either conserved or replaced with a glutamate acid residue, whereas nonpolar residues are located at this position in mammalian ones. This fact suggests that the electrostatic or hydrogen bond interaction between Arg13 and Asp115 is likely a unique property of bacterial LMW PTPs. A similar electrostatic interaction in stabilizing the active pocket was found between Arg18 and Asp92 in BPTP (54). These results suggest that electrostatic and/or hydrogen bond interactions between the conserved Arg in P-loop and nearby Asp or Glu residues may play roles in the formation of the active site of LMW PTPs.

**Substrate specificity.** It is well known that substrate specificity is highly related to the local conformation of the active site and/or nearby regions. In LMW PTPs, loops between β2 and ß2 and that between the last two helices are considered to be involved in the recognition of and binding with phosphotyrosine-containing specific substrates (30, 52). The D-P-Y/W/F motif in the loop between the last two helices is also a conserved sequence in LMW PTPs. The hydrophobic nature and the orientation of the aromatic residue in this motif are important for the affinity of the enzyme towards different substrates. In YwlE, the aromatic ring of Phe120 extends the active cavity and forms a wall of the crevice for the incoming phosphotyrosine substrates. The loop between residues Phe40 and Thr48 forms another wall of the crevice (Fig. 5).

The positions of the hydrophobic and aromatic residues lining the walls of the active cavity suggest the enzyme may
specifically select the phosphotyrosine substrate instead of phosphoserine or phosphothreonine. Recent studies showed that YwlE was only active on phosphotyrosine-containing substrates when different substrates, including phosphotyrosine-, phosphoserine-, and phosphothreonine-containing peptides, were used, which is consistent with our proposal (36). In comparison with Trp49 in BPTP or Tyr48 in Ltp1, a nonaromatic residue, Ser42, occupies this position in YwlE. Though an aromatic residue, Phe40, is close to the active cavity, the phenylalanyl group of Phe40 is too far away from that of Phe120 to stack for the hydrophobic interaction. The lower stacking interaction may be consistent with the high \( K_m \) value of YwlE when pNPP is used as the substrate, which is several orders of magnitude higher than that of BPTP and Ltp1. Due to the hydrophilic residue Ser42, it is possible that sodium pyrophosphate can function as an inhibitor of YwlE, but not other LMW PTPs. Therefore, the concentration of pyrophosphate in \( B. subtilis \) may act as a regulator of YwlE activity.

The charge distribution around the active site is mainly contributed from the residues located in the loop between the \( \beta_2 \) and \( \alpha_2 \) (residues Phe40 to His50). Ala41 and Gly45 are close to the active site and may be important for the selection of specific substrates. In eukaryotic LMW PTPs, residues closed to the active cavity are polar so that the surface near the active site is positive (BPTP) or negative (HCPTPA), whereas Ala41 and Gly45 in YwlE are nonpolar residues and contribute a noncharged surface around the active site (Fig. 5). Studies in mammalian LMW PTPs demonstrated that the charge distribution on the surface around the active site has a significant role in determining the specificities of the enzymes toward the substrates (6, 52). Thus, the presence of Ala41 and Gly45 indicates that YwlE might have natural peptides or protein substrates with fewer charges distributed on their surface.

**Structural basis of the activity.** The phosphatase activity was measured using pNPP, a nonphysiological phosphotyrosine anlogue, as the substrate. The Michaelis-Menten constant \( K_m \) is about 28 mM and the rate constant \( k_{cat} \) is about 0.14 s\(^{-1}\), which are much lower than that of eukaryotic LMW PTPs (0.38 mM for BPTP and 0.017 mM for Ltp1).
In the structure of YwIE, the position of the aromatic residue in other LWM PTPs that was lining the wall of the active cavity is replaced by serine, and in addition the side chain ring of Phe40 is not oriented favorably to form a closed crevice with Phe120. These structural differences may result in less stacking interaction between YwIE and pNPP. In addition, the conformation of the P-loop is different from that of BPTP and Ltp1. The N-H bond of Thr11 points away from the center of the pocket, which probably results in a smaller number of amide protons interacting with oxygen atoms of the phosphate group of the substrate, such as pNPP (Fig. 4). In comparison with Cys17 in BPTP and Cys18 in Ltp1, the side chain of Cys12 in YwIE flips into the active pocket, resulting in resistance in the space when the substrate interacts with the enzyme.

Internal dynamics. The longitudinal and transverse relaxation rate constants R$_1$ and R$_2$ and heteronuclear $(^1$H)$-^{13}$N NOE values were determined to investigate the motional properties of B. subtilis YwIE. The R$_1$, R$_2$, and NOE values versus the protein sequence are shown in Fig. 6A. We have calculated the $\tau_0$ values at sample concentrations of 0.1, 0.25, 0.5, and 1.0 mM with the program TENSOR2 (13) and the corresponding correlation time $\tau_0$ values were 8.83 ± 0.16, 9.37 ± 0.13, 9.65 ± 0.15, and 11.24 ± 0.23 ns, respectively, suggesting a weak oligomerization (dimmer) of YwIE at higher concentrations. In addition, YwIE showed a slightly axial anisotropy around 1.1, calculated using the structure and dynamics data.

The reduced spectral density function mapping approximation was used to analyze the experimentally determined relaxation data (15). In this approach, the calculated spectral density functions $J(0)$, $J(\omega_0)$, and $J(0.87\omega_0)$ values can be used to analyze the overall and internal motions, where $J(0)$, $J(\omega_0)$, and $J(0.87\omega_0)$ represent the spectra density functions at frequencies of 0, $\omega_0$, and $0.87\omega_0$, respectively. The calculated $J(0)$ values for YwIE are shown in Fig. 6B. Since the values of $J(0)$ are normally dominated by that of $R_2$, it is reasonable that the plots of $J(0)$ shows the trends of $R_2$ versus residues. In addition, the $J(0)$ values are also contributed by the chemical exchanges $R_\infty$.

The fairly constant values of $J(\omega_0)$ and $J(0.87\omega_0)$ imply that the protein is overall rigid except residues in the loop regions of Phe40-Ala51 and Leu62-Ser67 and the C terminus, where the residues show systematically higher $J(0.87\omega_0)$ values than those in other regions, indicative of a degree of flexibility on the picosecond to nanosecond timescales. Residues Thr48, Ala51, and Val65 and residues Ser14 and Met16 in the P-loop region show significantly large $J(0)$ values, suggesting conformational exchanges on a millisecond timescale. In addition, some residues in the P-loop were missing or unused due to the signal broadening and being undetectable or of poor quality for analysis, which arose from the intermediate conformational changes. Therefore, the $J(0)$ and $J(0.87\omega_0)$ values suggested that the residues in the P-loop region and loop regions of Phe40-Ala51 and Leu62-Ser67 are involved in both subnano-second internal motions and millisecond conformational exchanges (Fig. 6B). It is well known that enzyme activity is correlated with flexibility (26, 42, 47). Flexibility of the P-loop implies that it may influence the substrate-binding specificities.

Conclusions. The solution structure of YwIE shows the conserved structure and catalytic residues of the active site that catalyzes the hydrolysis of phosphate on tyrosine residue in its targets. Due to the conservations YwIE may function in a similar way in which the intermediate phosphocysteine enzyme is formed. The backbone structure of the P-loop is similar to that of mammalian LWM PTPs, whereas the side chain conformations are different. Unlike Arg18 in BPTP, Arg13 may not be directly involved in initial phosphate binding in YwIE. However, the sequence conservation and its possible stabilization by Asp115 suggest its importance in forming a phosphate-binding pocket.

The pair Arg13/Asp115 exists in all known bacterial LWM PTPs and is likely a hallmark of these enzymes. The flexibility of loops around the active site and noncharged surface resulting from Ala41 and Gly45 indicate that the hydrophobic interaction may occur between YwIE and the specific endogenous protein substrate. Recent studies on B. subtilis PTPs indicate that YwIE shows activity similar to that of YwQE towards pNPP but a lower activity on the substrate proteins of YwQE (33). However, YwIE is not located in the same operon as YwQE and the operon encoding YwIE remains largely unknown. Reduced resistance to ethanol stress resulting from genomic deletion of YwIE was observed and provided a clue in exploring the physiological roles of B. subtilis LWM PTPs (36). Further investigations are expected to characterize the specific endogenous protein substrates for YwIE and elucidate the role of Arg13 and Asp115.

ACKNOWLEDGMENTS

All NMR experiments were carried out at the Beijing Nuclear Magnetic Resonance Center (BNMRC), Peking University. We thank Xiao-Dong Su and Xiaofeng Zheng (Peking University) for help in the activity assays. We thank Ewen Lescop, Yunfei Hu, and Xianrong Guo (BNMRC) for discussions. We thank Xiubei Liao of the University of Illinois at Chicago for critical reading of the manuscript and suggestions.

This work was supported by the National Natural Science Foundation of China grant 30125009 (B.X.), grant 30325010 (C.J.), and grant 30570354 (C.J.).

REFERENCES


