Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure and activity of *Yersinia pestis* 6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase as a novel target for the development of antiplague therapeutics

Jaroslaw Blaszczyk,[a]‡§ Yue Li,[b]‡ Scott Cherry,[a]‡ Jerry Alexandratos,[a] Yan Wu,[b] Gary Shaw,[a] Joseph E. Tropea,[a] David S. Waugh,[a] Honggao Yan[b]* and Xinhua Ji[a]*

[a]Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702, USA, and [b]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA

‡ These authors contributed equally to this work.
§ Current address: RCSB-Rutgers, Department of Chemistry and Chemical Biology, Piscataway, NJ 08854, USA.

Correspondence e-mail: yanh@msu.edu, jix@ncifcrf.gov

Received 8 September 2007
Accepted 26 September 2007

PDB Reference: *Y. pestis* HPPK, 2qx0, r2qx0sf.

6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) is a key enzyme in the folate-biosynthetic pathway and is essential for microorganisms but absent from mammals. HPPK catalyzes Mg^{2+}-dependent pyrophosphoryl transfer from ATP to 6-hydroxymethyl-7,8-dihydropterin (HP). Previously, three-dimensional structures of *Escherichia coli* HPPK (EcHPPK) have been determined at almost every stage of its catalytic cycle and the reaction mechanism has been established. Here, the crystal structure of *Yersinia pestis* HPPK (YpHPPK) in complex with HP and an ATP analog is presented together with thermodynamic and kinetic characterizations. The two HPPK molecules differ significantly in a helix-loop area (α2-Lp3). YpHPPK has lower affinities than EcHPPK for both nucleotides and HP, but its rate constants for the mechanistic steps of both chemical transformation and product release are comparable with those of EcHPPK. *Y. pestis*, which causes plague, is a category A select agent according to the Centers for Disease Control and Prevention (CDC). Therefore, these structural and biochemical data are valuable for the design of novel medical countermeasures against plague.

1. Introduction

Yersinia pestis is the causative agent of bubonic, pneumonic and septicemic plague and is one of the most virulent pathogens known (Brubaker, 1991). It is estimated that the three recorded plague pandemics in history killed 200 million people (Perry & Fetherston, 1997). Plague outbreaks continue to occur throughout the world; the most recent occurred in Congo in June 2006 with 100 reported cases. The potential use of *Y. pestis* in biological weapons is of great concern in biodefense (Inglesby *et al.*, 2000). Because bioterrorist attacks are difficult to predict, detect or prevent, it is essential to have effective therapeutics to address the immediate health needs of the public in the event of such an attack. Although antimicrobial agents for treating *Y. pestis* exist, a new generation of antimicrobial agents is needed to combat drug-resistant variants that are engineered or emerge naturally. A multidrug-resistant *Y. pestis* strain resistant to all first-line and several prophylactic antibiotics was identified in the outbreak of bubonic plague in Madagascar in 1995 (Galimand *et al.*, 1997). Recently, another strain was identified to be resistant to streptomycin, the current antibiotic of choice for treating plague (Guiyoule *et al.*, 2001).

Folate cofactors are essential for life (Blakley & Benkovic, 1984). Mammals derive folates from their diet, whereas most

microorganisms must synthesize folates *de novo*. Therefore, enzymes in the folate-biosynthetic pathway are ideal targets for the development of antimicrobial agents. 6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase (HPPK) is a key enzyme in the folate pathway (Shiota, 1984). It catalyzes the transfer of pyrophosphate from ATP to 6-hydroxymethyl-7,8-dihydropterin (HP) and generates AMP and HP pyrophosphate (HPPP; Fig. 1*a*). Because HPPK is not the target of any existing antibiotics, it is very attractive for the development of novel antibiotics that are urgently needed to counteract worldwide antibiotic resistance. Recently, structural and functional data have been reported for several HPPK proteins, in particular *Escherichia coli* HPPK (EcHPPK). The catalytic mechanism of EcHPPK has been established (Fig. 1*b*) and the reaction trajectory of EcHPPK-catalyzed pyrophosphoryl transfer has been mapped out (Li *et al.*, 2002; Blaszczyk, Shi *et al.*, 2004).

Both structural and functional analyses indicate that protein dynamics are important for catalysis by EcHPPK, which involves large conformational changes in three catalytic loops: Lp1, Lp2 and Lp3. Among the three loops, the changes in Lp3 are not only dramatic but are also unusual (Xiao *et al.*, 2001). Within a catalytic cycle, Lp3 opens up twice: firstly upon the binding of MgATP, which is necessary for the enzyme to bring the side chains of two catalytic arginine residues into the active center, and then upon the completion of chemical transformation, which is necessary for product release (Blaszczyk, Li *et al.*, 2004; Blaszczyk, Shi *et al.*, 2004). For the pyrophosphoryl-transfer reaction to occur, the coupling of the three loops seals the active center of EcHPPK (Blaszczyk *et al.*, 2000). The function–dynamic correlation and ligand-induced conformational changes of EcHPPK have also been elucidated using a low-frequency normal-mode analysis approach, indicating that Lp2 and Lp3 exhibit the most concerted motions for ligand recognition and catalysis (Keskin *et al.*, 2002).

The *E. coli* enzyme is homologous to HPPK proteins from other pathogenic microbes (Xiao *et al.*, 1999). For example, it shares 61% identity with *Y. pestis* HPPK (YpHPPK; Fig. 2), 56% with *Haemophilus influenzae* HPPK, 40% with *Pneumocystis carinii* HPPK and 33% with *Mycobacterium tuberculosis* HPPK. Thus, the structural information and catalytic mechanism derived from EcHPPK may be adequate for the development of broad-spectrum antimicrobials. For the design of narrow-spectrum antimicrobial agents, however, detailed structural information about HPPKs from particular microorganisms is crucial. In addition to *E. coli*, crystal structures of HPPK have been determined from *H. influenzae* (Hennig *et al.*, 1999), *Streptococcus pneumoniae* (Garcon *et al.*, 2006) and *Saccharomyces cerevisiae* (Lawrence *et al.*, 2005). Here, we report the crystal structure of YpHPPK in complex with HP and α,β-methyleneadenosine triphosphate (AMPCPP, an ATP analog and dead-end inhibitor) together with the preparation and enzymatic studies of the enzyme. Our findings on YpHPPK will be helpful for the design of novel medical countermeasures against plague by targeting a key enzyme in the folate-biosynthetic pathway of *Y. pestis*, a category A select agent according to the Centers for Disease Control and Prevention (CDC).

2. Experimental

2.1. Cloning and expression of YpHPPK

The ORF encoding YpHPPK (GeneID 1145734) was amplified from genomic DNA by the polymerase chain reaction (PCR) using the following oligonucleotide primers: A, 5′-GAG AAC CTG TAC TTC CAG GGT CAT CAC CAC CAT CAC CAT ATG ATC CGG GTC TAT ATC GCG CTG-3′, and B, 5′-GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT ATT ACC AGA GAA CCA AGC CGT TC-3′. The resulting PCR amplicon was subsequently used as the template for a second PCR with PE-277 (Evdokimov *et al.*, 2002) and B as primers. The final PCR amplicon was inserted into pDONR201 (Invitrogen, Carlsbad, CA, USA) by recombinational cloning and the nucleotide sequence of the insert was confirmed experimentally. Next, the HPPK ORF, together with DNA encoding a tobacco etch virus (TEV) protease-cleavage site followed by a hexahistidine tag now joined in-frame to its N-terminus, was recombined into pKM596 (Fox & Waugh, 2003) to generate the plasmid expression vector pKM1016. The MBP-His_6-HPPK fusion protein was expressed in *E. coli* BL21(DE3)-RIL CodonPlus cells (Stratagene, Valencia, CA, USA) that also contained the TEV protease expression vector pRK603 (Kapust & Waugh, 2000). The cells were grown to mid-log phase ($OD_{600\,nm}$ = 0.5) at 310 K in Luria broth containing 100 µg ml^{-1} ampicillin, 30 µg ml^{-1} chloramphenicol and 25 µg ml^{-1} kanamycin, whereupon isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M* and the temperature was reduced to 303 K. 4 h after induction, the cells were pelleted by centrifugation and stored at 193 K.

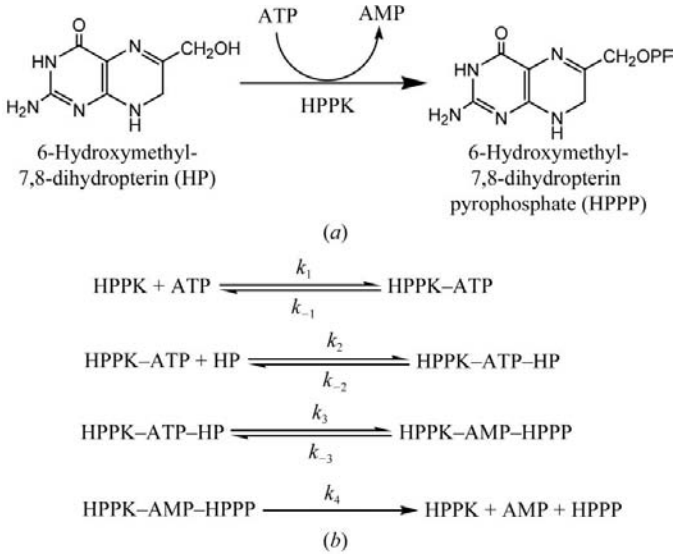

Figure 1
Reaction catalyzed by HPPK. (*a*) The substrates (HP and ATP) and products (HPPP and AMP) of the reaction. (*b*) The kinetic mechanism of the reaction. Mg^{2+} (not shown) is required for the reaction.

Table 1
X-ray data and refinement statistics for YpHPPK–MgAMPCPP–HP.

Values in parentheses are for the last shell.

Data statistics	
Resolution range (Å)	28.9–1.8 (1.86–1.80)
Completeness (%)	95.7 (90.9)
Redundancy	3.7 (2.9)
$I/\sigma(I)$	11.6 (2.8)
$R_{\rm scaling}$†	0.110 (0.288)
Refinement statistics	
Reflections used for refinement	30817 (2765)
Reflections used for $R_{\rm free}$	1525 (134)
Crystallographic R‡	0.230 (0.312)
$R_{\rm free}$	0.269 (0.348)
No. of protein atoms/average B factor (Å^2)	2516/21.4
No. of ligand atoms/average B factor (Å^2)	108/12.2
No. of water O atoms/average B factor (Å^2)	465/34.6
R.m.s.d. from ideal geometry	
Bond distances (Å)	0.006
Bond angles (°)	1.4
Ramachandran plot	
Most favored φ/ψ angles (%)	90.1
Disallowed φ/ψ angles (%)	0
Estimated coordinate error (Å)	0.24

† $R_{\rm scaling} = \sum |I - \langle I\rangle|/\sum I$. ‡ Crystallographic $R = \sum_{hkl} ||F_{\rm o}| - |F_{\rm c}||/\sum_{hkl} |F_{\rm o}|$.

2.2. Purification of YpHPPK

All chromatography steps were carried out at 277 K. *E. coli* cell paste was suspended in ice-cold 25 m*M* HEPES pH 8, 100 m*M* NaCl, 10% glycerol, 25 m*M* imidazole (buffer *A*) containing Complete EDTA-free protease-inhibitor cocktail (Roche Molecular Biochemicals, Indianapolis, IN, USA). The cells were lysed with an APV-1000 homogenizer (Invensys, Roholmsvej, Germany) at 69 MPa and centrifuged at 30 000*g* for 30 min at 277 K. The supernatant was filtered through a 0.45 µm polyethersulfone membrane and loaded onto an Ni–NTA superflow column (Qiagen, Valencia, CA, USA) equilibrated in buffer *A*. The column was washed to baseline with buffer *A* and eluted with a linear gradient of imidazole to 200 m*M*. Fractions containing recombinant protein were pooled, concentrated using a YM-10 membrane (Millipore Corporation, Bedford, MA, USA) and applied onto a 26/60 HiPrep Sephacryl S-100 prep-grade column (Amersham Biosciences, Piscataway, NJ, USA) equilibrated in 25 m*M* HEPES pH 8, 100 m*M* NaCl, 10% glycerol buffer. The peak fractions containing recombinant HPPK were pooled and concentrated to 8 mg ml^{-1} (determined spectrophotometrically using a molar extinction coefficient of 17 780 M^{-1} cm^{-1}). Aliquots were flash-frozen in liquid nitrogen and stored at 193 K. The final product was judged to be >95% pure by SDS–PAGE. The molecular weight was confirmed by electrospray mass spectrometry.

2.3. Crystallization and X-ray diffraction data collection

ATP and AMPCPP were purchased from Sigma. HP was purchased from Schircks Laboratories (Jona, Switzerland). The crystals of YpHPPK in complex with MgAMPCPP and HP were cocrystallized using the hanging-drop method at 292 ± 1 K. The drops contained equal volumes of protein and reservoir solution. The protein solution was composed of 8.0 mg ml^{-1} YpHPPK, 15 m*M* HP, 25 m*M* AMPCPP, 50 m*M* $MgCl_2$ and 10 m*M* Tris–HCl pH 8.0. The well solution contained 0.18 *M* ammonium acetate, 30%(*w*/*v*) PEG 4000, 20 m*M* imidazole and 0.1 *M* sodium acetate pH 4.6. The crystals appeared within a week and reached dimensions of 0.10 × 0.15 × 0.30 mm after two further weeks.

X-ray diffraction data were collected from a single crystal at 100 K using an ADSC Quantum-4 CCD detector mounted on the synchrotron beamline X9B at National Synchrotron Light Source, Brookhaven National Laboratory. The monoclinic crystal belonged to space group $P2_1$ (unit-cell parameters a = 43.40, b = 78.62, c = 52.60 Å, $\alpha = \gamma = 90$, $\beta = 100.72°$) and diffracted to 1.8 Å resolution. The solvent content of the crystal was 0.50, corresponding to a Matthews coefficient (Matthews, 1968) of 2.45 Å^3 Da^{-1}. Data processing was carried out with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). X-ray data statistics are summarized in Table 1.

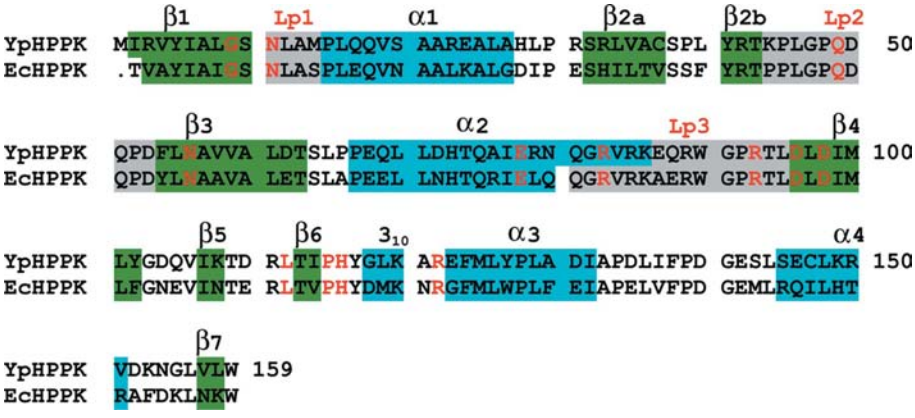

Figure 2
Structure-based sequence alignment of YpHPPK and EcHPPK. The SWISS-PROT accession codes for the sequences are Q7CKD7 and P26281, respectively, and the PDB accession codes for the structures are 2qx0 and 1q0n, respectively. The N-terminal Met of EcHPPK is removed *in vivo* as shown by N-terminal peptide sequencing of the native EcHPPK (Talarico *et al.*, 1992) and MALDI-MS analysis of the recombinant EcHPPK (Shi *et al.*, 2000). Secondary-structural elements are indicated by shading: helices in cyan, strands in green and loops (Lp1, Lp2 and Lp3) in gray. The 13 highly conserved amino-acid residues in HPPK sequences are highlighted in red.

2.4. Structure solution and refinement

The structure was solved by molecular replacement using *AMoRe* (Navaza, 1994). The search model was the crystal structure of EcHPPK–MgAMPCPP–HP (PDB code 1q0n; Blaszczyk *et al.*, 2000) without ligand and solvent molecules. The solution contained two YpHPPK molecules, with a correlation coefficient of 0.49 and a crystallographic R factor of 0.43 for X-ray data within the resolution range 10.0–4.0 Å.

The structure was refined using *CNS* (Brünger *et al.*, 1998) and *SHELXL*-97

(Sheldrick & Schneider, 1997). Difference electron density revealed, for each polypeptide chain, the position of the AMPCPP and HP molecules and two Mg^{2+} ions. Noncrystallographic symmetry restraints were applied at the early stage of refinement and were gradually released later. A total of 1525 (4.9%) reflections were randomly selected for cross-validation (R_{free}). Bulk-solvent correction was employed. During refinement, $2F_o - F_c$ and $F_o - F_c$ electron-density maps were regularly calculated and examined. Solvent molecules, as peaks $\geq 3\sigma$ on the $F_o - F_c$ map with reasonable hydrogen-bond networks, were included as water O atoms at the later stage of the refinement. A Cys residue (Cys37) was found to be oxidized, which might have occurred during crystallization as it was not detected by mass spectrometry earlier. Upon completion, the entire structure was verified with the composite annealed OMIT map (Brünger *et al.*, 1998) and assessed with *PROCHECK* (Laskowski *et al.*, 1993). The statistics for refinement and the final structure can be found in Table 1. All graphics work was carried out using *O* (Jones *et al.*, 1991). Illustrations were prepared with *PyMOL* (Delano Scientific, LLC).

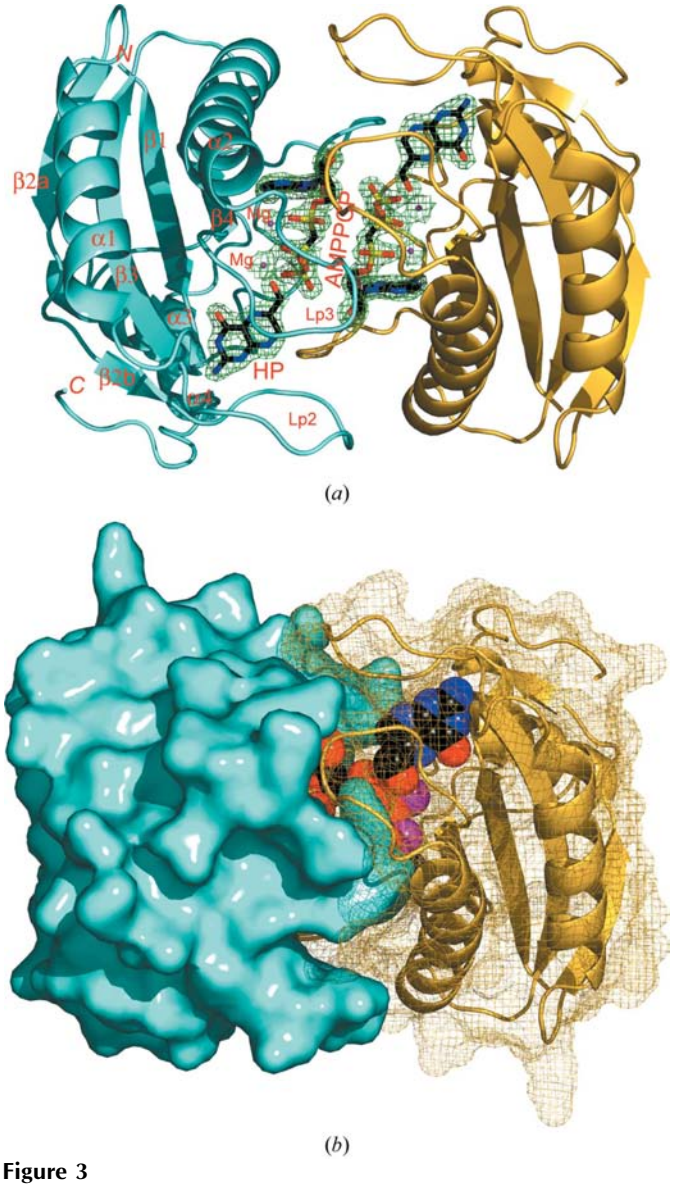

Figure 3
The dimeric structure of YpHPPK–MgAMPCPP–HP. (*a*) Schematic illustration of the YpHPPK–MgAMPCPP–HP dimer. The two subunits of the dimer are indicated in different colors (cyan and orange). For clarity, secondary-structural elements are labeled for one subunit only. Helices, β-strands and loops are drawn as spirals, arrows and pipes, respectively. The AMPCPP and HP molecules are shown as stick models and the Mg^{2+} ions as balls in atomic colors (carbon in black, nitrogen in blue, oxygen in red, phosphorus in yellow and magnesium in purple). The composite annealed OMIT map ($2F_o - F_c$, contoured at 1.5σ) is shown for the ligands as green nets. (*b*) Surface representation of the dimer. One subunit is shown as a solid surface and the partner subunit as a mesh surface with enclosed ribbon diagram for the subunit. The ligands are illustrated as van der Waals spheres in the same colors as shown in (*a*).

2.5. K_d measurements

The K_d values for various ligands were measured by fluorometric titration at 297 K on a Spex FluoroMax-2 fluorometer, with the use of a fluorescent analogue of MgATP, 3′(2′)-*O*-anthraniloyladenosine 5′-triphosphate (Ant-ATP), as described previously (Shi *et al.*, 2000; Li *et al.*, 2002). The dissociation constant for MgAnt-ATP was determined by measuring the change in fluorescence upon the formation of the complex between MgAnt-ATP and HPPK; those of other nucleotides were determined by a competitive binding assay measuring the fluorescent changes caused by the displacement of MgAnt-ATP by the other nucleotides. The dissociation constant of HP was measured in the presence of MgAMPCPP.

2.6. Stopped-flow analysis

Stopped-flow experiments were performed in an Applied Photophysics SX.18MV-R stopped-flow spectrofluorometer as described previously (Li *et al.*, 2002). To measure the binding of MgAnt-ATP, a solution containing HPPK was mixed rapidly with a solution containing Ant-ATP and $MgCl_2$ at 298 K. YpHPPK and $MgCl_2$ were fixed at 1 μ*M* and 10 m*M*, respectively. Ant-ATP was varied from 6 to 24 μ*M*. To measure the binding of HP, a solution containing YpHPPK, AMPCPP and $MgCl_2$ was mixed rapidly with a solution containing HP at 298 K. YpHPPK, AMPCPP and $MgCl_2$ were fixed at 0.25 μ*M*, 30 μ*M* and 10 m*M*, respectively. HP was varied from 0.25 to 4 μ*M*. All concentrations refer to the concentrations immediately after the rapid mixing. The rate constants were evaluated by numerical analysis of the stopped-flow data using the program *DYNAFIT* (Kuzmic, 1996). The initial values for the iterative numerical analysis were obtained by a nonlinear least-squares fit of the stopped-flow data to an exponential equation using the program *Origin*.

2.7. Quench-flow analysis

Quench-flow experiments were carried out with a KinTek RQF-3 rapid quench-flow instrument as described previously (Li *et al.*, 2002). All reaction components were dissolved in 100 m*M* Tris–HCl buffer pH 8.3. The reaction mixture contained 10 μ*M* YpHPPK, 100 μ*M* ATP, 110 μ*M* HP, 10 m*M* $MgCl_2$, 25 m*M* DTT, 0.5 m*M* EDTA and a trace amount of

[α-^{32}P]-ATP. The reaction was initiated with Mg^{2+} at 303 K and quenched with 500 m*M* EDTA. The substrates and products were separated by thin-layer chromatography on a plastic sheet coated with PEI-cellulose developed with 0.25 *M* NaH_2PO_4. The radioactivity was quantified using a Molecular Dynamics Storm 820 Phosphor-Imager. The pre-steady-state data were first analyzed by a nonlinear least-squares fit of the data to (1),

$$[\mathrm{AMP}]/[\mathrm{HPPK}] = A[1 - \exp(\lambda t)] + k_{\mathrm{cat}}t. \qquad (1)$$

The amplitudes and rate constants were then used to set the initial values for fitting the data to the complete mechanism by numerical analysis using the program *DYNAFIT* (Kuzmic, 1996).

3. Results and discussion

3.1. Crystal structure of the YpHPPK–MgAMPCPP–HP complex

The statistics for the final structure are listed in Table 1. The entire structure is well defined except for a few solvent-exposed side chains. The $2F_o - F_c$ annealed OMIT map (contoured at 1.5σ) is shown for the ligands (Fig. 3*a*). The asymmetric unit contains a dimer of the YpHPPK–MgAMPCPP–HP complex, including two polypeptide chains (each containing 159 amino-acid residues), four Mg^{2+} ions, two AMPCPP molecules, two HP molecules and 465 water O atoms. Each subunit has a three-layered $\alpha\beta\alpha$ fold as found in EcHPPK (Fig. 4*a*). The secondary-structure assignment for YpHPPK is virtually the same as that of the *E. coli* homolog, except that the α2 helix is longer by ~1.5 turns and the Lp3 loop is shorter by six residues (Fig. 2).

The two crystallographically independent YpHPPK–MgAMPCPP–HP complexes in the asymmetric unit are almost identical, with an r.m.s.d. of 0.16 Å for all 159 pairs of C^{α} positions. They face each other in the dimer, with their AMPCPP phosphate bridges in contact (Fig. 3*a*). A high degree of complementarity exists between the surfaces of the two subunits (Fig. 3*b*) and the total buried surface area between the two YpHPPK–MgAMPCPP–HP complexes is 2851 $Å^2$, with 22 inter-subunit hydrogen bonds. According to an analysis of 122 homodimeric proteins, the average total buried surface area between the two subunits is 3900 $Å^2$, with 18.5 inter-subunit hydrogen bonds (Bahadur *et al.*, 2003; Janin *et al.*, 2007). Comparatively, the crystal-packing interfaces are small (1510 $Å^2$) with a low number of inter-subunit hydrogen bonds (5.4), on the basis of an analysis of 188 large crystal-packing interfaces (Janin & Rodier, 1995; Bahadur *et al.*, 2004; Janin *et al.*, 2007). Although YpHPPK–MgAMPCPP–HP crystallizes as a dimer, it is not clear whether the dimerization is catalytically relevant. The solution state of the YpHPPK–MgAMPCPP–HP complex is under investigation.

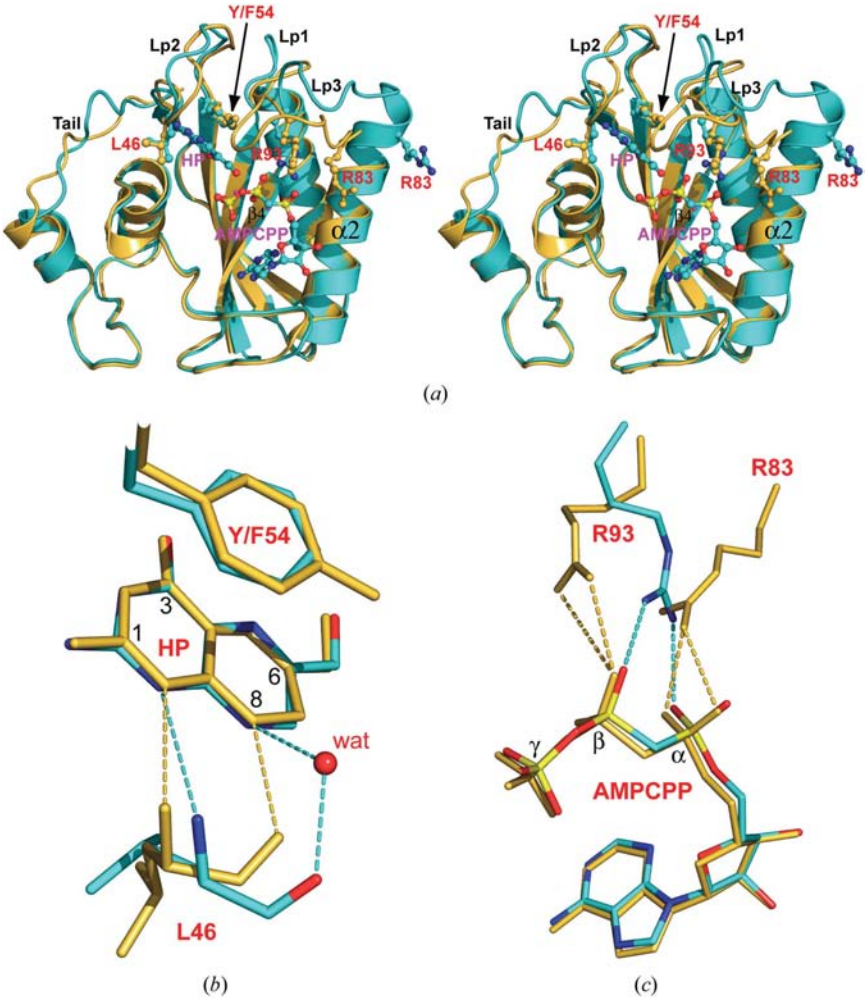

Figure 4
Structural comparison of the HPPK–MgAMPCPP–HP complexes of the *Y. pestis* and *E. coli* proteins. (*a*) Stereoview showing the YpHPPK (in cyan) and EcHPPK (in orange) structures. Helices, β-strands and loops are drawn as spirals, arrows and pipes, respectively. Selected side chains and the ligands are shown as ball-and-stick models. An atomic color scheme (carbon in cyan, nitrogen in blue, oxygen in red, phosphorus in yellow and magnesium in purple) is used for the YpHPPK structure. Only α2, Lp2 and Lp3 are labeled among the secondary-structure elements. (*b*) The differences between the HP-binding modes in the two HPPK complexes. (*c*) The differences between the AMPCPP-binding modes in the two complexes.

3.2. Substrate-binding sites

Superposition of YpHPPK–MgAMPCPP–HP and EcHPPK–MgAMPCPP–HP (PDB code 1q0n) shows that the helices and strands in the two structures superimpose well whereas the loops do not, especially the three catalytic loops and the C-terminal tail (Fig. 4*a*). Owing to the different conformations of Lp2, the HP–protein interactions in the two complexes differ significantly. In the

Table 2
Comparison of the ligand-binding properties of YpHPPK and EcHPPK.

	EcHPPK	YpHPPK
K_d(MgAnt-ATP) (μ*M*)	1.6 ± 0.05†	7.5 ± 0.6
K_d(MgATP) (μ*M*)	2.6 ± 0.06†	7.6 ± 0.8
K_d(MgAMPCPP) (μ*M*)	0.077 ± 0.006‡	0.29 ± 0.008
K_d(HP) (μ*M*)	0.17 ± 0.01‡	0.78 ± 0.03

† Data from Shi *et al.* (2000). ‡ Data from Li *et al.* (2002).

EcHPPK complex, residue Leu46 forms two hydrogen bonds with HP, one between the Leu46 amide N atom and the HP N1 atom and the other between the Leu46 carbonyl O atom and the HP N8 group, whereas in the YpHPPK complex the interaction between the Leu46 carbonyl O atom and the HP N8 group is bridged by a water molecule (Fig. 4*b*). Residue Leu46 is not a catalytic residue and its interaction with HP should not have any impact on catalysis, although it may affect the affinity for HP. Another minor difference between the two HP-binding sites is that residue Tyr54 in EcHPPK becomes a Phe in YpHPPK. The stacking interaction of the benzyl-ring system of this residue with HP is conserved in the two structures (Fig. 4*a*). Residue 54 is also not a catalytic residue. Thus, the Tyr/Phe variation should not have any impact on either the HP binding or the catalytic properties of the enzyme.

The positions and coordination systems of the two Mg^{2+} ions are the same as previously observed for EcHPPK (Blaszczyk *et al.*, 2000). Although the AMPCPP molecule also occupies the same position in the two complexes (Fig. 4*a*), its interactions with the protein are not identical. In the *E. coli* enzyme residue Arg83 interacts with the α-phosphate while residue 93 interacts with the β-phosphate; in YpHPPK, however, Arg83 is far away from the AMPCPP molecule (Fig. 4*a*) and Arg93 interacts with both the α-phosphate and β-phosphate groups (Fig. 4*c*).

Residues Arg83 and Arg93 are highly conserved in HPPK sequences. They are located within the Lp3 loop of the *E. coli* enzyme, whereas Arg83 becomes part of the extended α2 helix in YpHPPK (Figs. 2 and 4*a*). In EcHPPK, the two Arg residues play not only important but also dynamic roles in catalysis (Blaszczyk *et al.*, 2003; Li *et al.*, 2003). Structural information suggests that Arg93 first binds to the α-phosphate of ATP and then shifts to interact with the β-phosphate, while Arg83, which initially does not bind to ATP, moves in and binds to the α-phosphate when pyrophosphoryl transfer is about to occur. Biochemical analysis of the site-directed mutants EcHPPK(R83A) and EcHPPK(R93A) suggests that both Arg residues are dispensable for the binding of MgATP, but that Arg93 is much more important for catalysis. The structure of YpHPPK–MgAMPCPP–HP suggests that Arg83 of YpHPPK may not be involved in either binding or catalysis.

3.3. Thermodynamic and kinetic properties

The HPPK-catalyzed reaction follows an ordered bi-bi mechanism (Bermingham *et al.*, 2000). Previously, we established a thermodynamic and kinetic framework for the EcHPPK-catalyzed reaction (Fig. 1*b*) using a combination of thermodynamic and transient kinetic analysis (Li *et al.*, 2002). The same approach was used for the characterization of YpHPPK. The results of the thermodynamic analysis are summarized in Table 2 and representative data are shown in Fig. 5. The K_d values for both nucleotides and HP were threefold to fourfold higher than those for EcHPPK. The results showed that YpHPPK has lower affinities for both nucleotides and HP, in agreement with the water-bridged and thus weakened interaction between Leu46 and HP, and the elimination of the interaction between Arg83 and AMPCPP, when compared with those for the *E. coli* enzyme (Fig. 4).

The assembly of the ternary Michaelis complex involves large conformational changes in EcHPPK (Blaszczyk *et al.*, 2000) and product release is the rate-determining step in the EcHPPK-catalyzed reaction (Li *et al.*, 2002). Thus, it was of interest to measure the kinetics of the YpHPPK-catalyzed reaction. The rate constants for substrate binding were

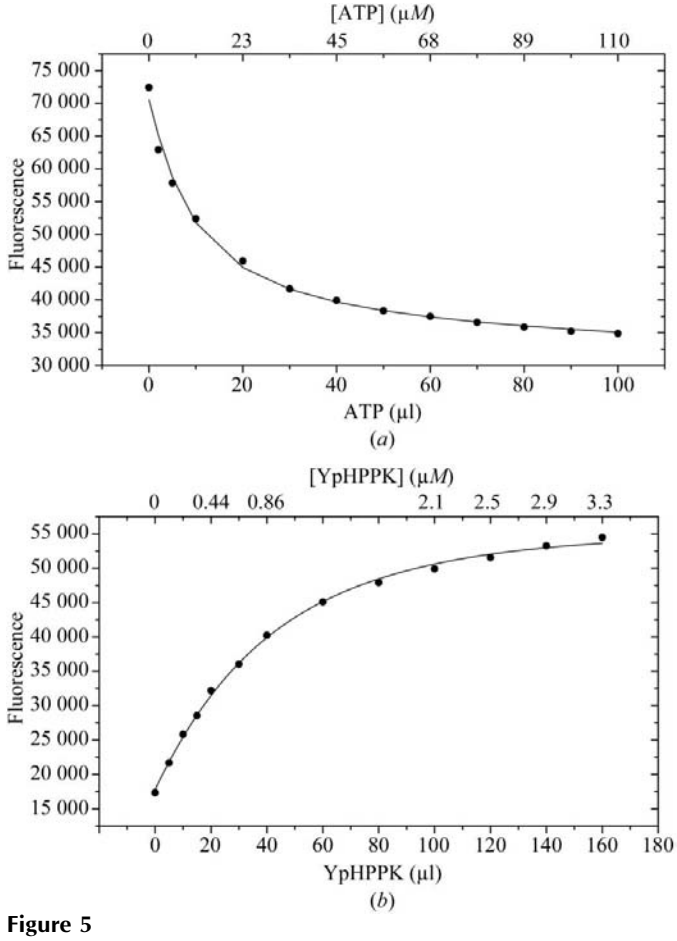

Figure 5
Binding of MgATP and HP to YpHPPK measured by fluorometry. (*a*) A 2 ml solution containing 10 μ*M* YpHPPK and 10 μ*M* Ant-ATP in 100 m*M* Tris–HCl, 10 m*M* $MgCl_2$ pH 8.3 was titrated with ATP by adding aliquots of a 2.3 m*M* ATP stock solution at 297 K. The top axis indicates the concentrations of ATP during the titration. (*b*) A 2 ml solution containing 0.55 μ*M* HP and 100 μ*M* AMPCPP in 100 m*M* Tris–HCl, 10 m*M* $MgCl_2$, 25 m*M* DTT pH 8.3 was titrated with YpHPPK by adding aliquots of a 44 μ*M* YpHPPK stock solution at 297 K. The top axis indicates the concentrations of YpHPPK during the titration. The solid line in both panels was obtained as described previously (Li *et al.*, 2002).

measured by stopped-flow fluorometry. The fluorescent MgATP analogue MgAnt-ATP was used to measure the kinetics of nucleotide binding, as MgATP induces little change in the fluorescence of the protein. The rate constants for the binding of HP were measured in the presence of the dead-end inhibitor MgAMPCPP. The results (Table 3) showed that the decrease in the affinity for the nucleotide in comparison to that of EcHPPK is mainly a consequence of an increase in the dissociation rate constant. On the other hand, the relative decrease in the affinity for HP is mainly a consequence of an increase in the association rate constant.

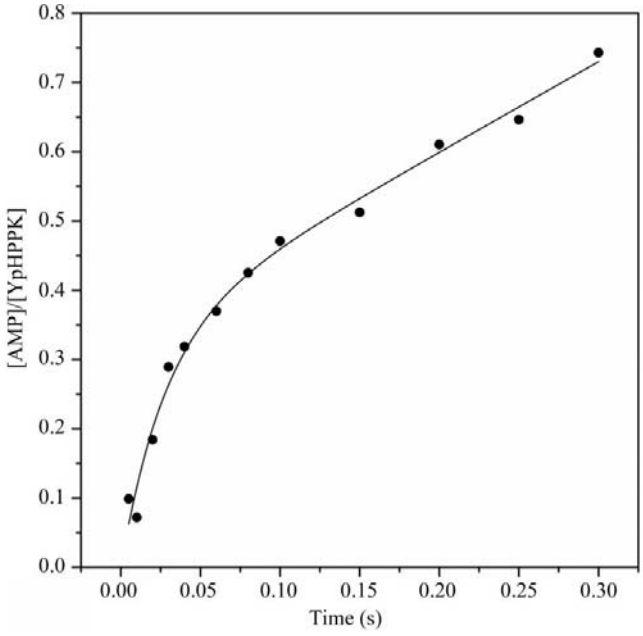

Figure 6
Quench-flow analysis of YpHPPK. The reaction mixture contained 10 µ*M* YpHPPK, 100 µ*M* ATP, 110 µ*M* HP, 25 m*M* DTT, 0.5 m*M* EDTA and 10 m*M* $MgCl_2$ in 100 m*M* Tris–HCl buffer pH 8.3. The reaction was initiated by mixing with $MgCl_2$ at 303 K and quenched by mixing with EDTA. The solid lines were obtained by a nonlinear least-squares fit to (1) as described in §2.7.

The rate constants for the chemical step were measured by quench-flow analysis as shown in Fig. 6. The time course consists of two phases: a burst phase and a steady-state phase described by (1), indicating that product release is rate-limiting in the enzymatic reaction. The values of the individual rate constants obtained from the transient kinetic analysis are listed in Table 3 together with the k_{cat} values calculated from the individual rate constants. The results showed that the rate constants for the chemical step and product release are very similar for YpHPPK and EcHPPK.

3.4. Conformational changes and dimerization

As shown in Fig. 4(*a*), there are significant conformational differences between the two HPPKs in Lp1, Lp2, α2-Lp3 and the C-terminal tail. The α2-Lp3 region exhibits the most obvious differences because the $_{80}$QGRVRK$_{85}$ sequence folds differently, resulting in a longer α2, a shorter Lp3 and the removal of the side chain of Arg83 from the active center of YpHPPK (Fig. 4*a*). Of the four flexible motifs, Lp2 and Lp3 are directly involved in the dimerization of YpHPPK–MgAMPCPP–HP, forming the complementary subunit interface and sealing the active centers of both subunits (Fig. 3). Therefore, the fold of the $_{80}$QGRVRK$_{85}$ sequence appears to be critical for the dimerization of YpHPPK. The new fold also prohibits the guanidinium group of Arg83 from reaching the active center, but allows it to form two hydrogen bonds across the subunit interface with the carbonyl O atom of Trp90 (Fig. 7). Trp90 is located in the middle of Lp3. Therefore, Arg82 may play a new role in YpHPPK by stabilizing the dimeric ternary complex.

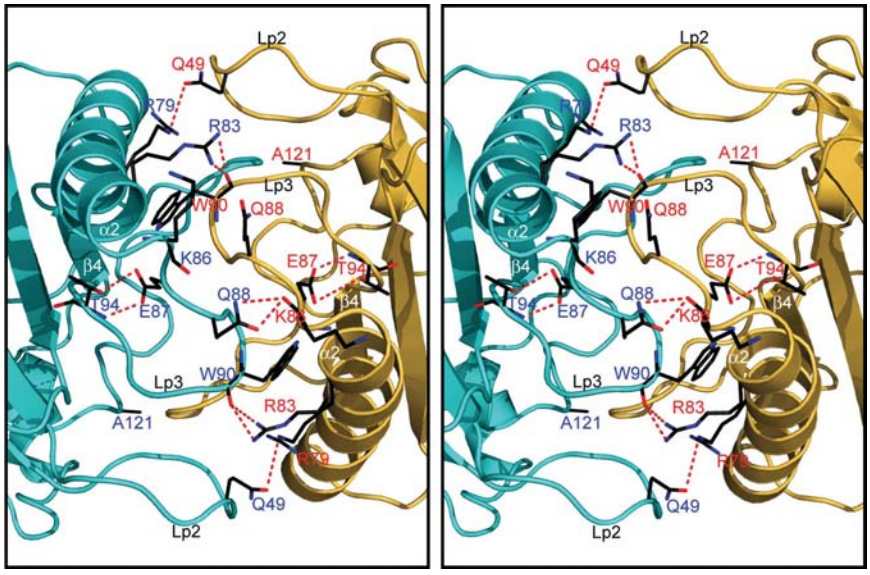

Figure 7
Stereoview showing the subunit interface of dimeric YpHPPK–MgAMPCPP–HP. The two subunits are shown in cyan (with blue labels) and orange (with red labels), respectively. Helices, β-strands and loops are drawn as spirals, arrows and pipes, respectively. Selected side chains are shown as stick models in an atomic color scheme (carbon in black, nitrogen in blue and oxygen in red). Hydrogen bonds are indicated by dashed lines in red. For clarity, ligands are not shown and only four secondary-structure elements are labeled.

3.5. Amino-acid substitutions and dimerization

The residues that constitute the subunit interface of YpHPPK–MgAMPCPP–HP include four Lp2 residues (Leu46, Gly47, Pro48 and Gln49), five α2 residues (Asp72, Glu78, Arg79, Arg83 and Lys85), five Lp3 residues (Glu87, Gln88, Arg89, Trp90 and Arg93) and residues Thr109, Asp110, Arg111, Thr113, Trp117, Ala121 and Arg122. Of these 21 residues, 15 are identical in Yp and EcHPPK. The remaining six residues that differ between the two enzymes are Asp72, Arg79, Glu87, Gln88, Asp110 and Ala121 in YpHPPK and are Asn72, Leu79, Ala87, Glu88, Glu110 and

Table 3
Kinetic parameters of YpHPPK and EcHPPK.

	EcHPPK	YpHPPK
k_1 ($\mu M\ s^{-1}$)	0.37 ± 0.007†	0.85 ± 0.04
k_{-1} (s^{-1})	0.51 ± 0.07†	5.1 ± 0.08
k_2 ($\mu M\ s^{-1}$)	11 ± 0.03‡	3.6 ± 0.2
k_{-2} (s^{-1})	2.0 ± 0.04‡	2.1 ± 0.1
k_3 (s^{-1})	16 ± 1‡	17 ± 1
k_{-3} (s^{-1})	20 ± 1‡	18 ± 1
k_4 (s^{-1})	1.8 ± 0.1‡	2.5 ± 0.07
k_{cat} (s^{-1})	0.76‡	0.61

† Measured for MgAnt-ATP (this work). ‡ Data from Li *et al.* (2002).

Asn121 in EcHPPK (Fig. 2). Of the six substitutions, two appear to be critical, two are helpful and two are not important for the dimerization of YpHPPK.

Glu87 and Ala121 may be critical for the dimerization of YpHPPK. Glu87 is the first residue following the extended $\alpha2$ helix. As shown in Fig. 7, the side chain of Glu87 in YpHPPK points towards the active center and its carboxyl group forms two hydrogen bonds with Thr94: one with NH (2.7 Å between Glu87 OE2 and Thr94 N) and the other with the hydroxyl group (3.2 Å between Glu87 OE1 and Thr94 OG1). Thr94 is located at the beginning of the $\beta4$-strand that follows Lp3. The relatively long side chain of Glu87 and its two hydrogen bonds to Thr94 play critical roles, not only stabilizing Lp3 but also stabilizing the C-terminal portion of $\alpha2$ formed by the $_{80}$QGRVRK$_{85}$ sequence, which is essential for the dimerization of YpHPPK. In EcHPPK, the residue at this position is Ala87, the side chain of which has neither the length nor the properties to interact favorably with Thr94. Unlike residue 87, the short side chain of Ala121 appears to also be necessary for dimerization because any bulkier side chain in this position, such as Asn121 in EcHPPK, would disturb the subunit interface with unfavorable steric interactions (Fig. 7). Additionally, the side chain of Ala121 may contribute to the stabilization of the dimer by making favorable hydrophobic interactions with the aliphatic portions of the side chains of Thr109 and Asp110 of the partner subunit.

Arg79 and Gln88 may help to stabilize the homodimer. The long side chain of Arg79 in YpHPPK forms many hydrophobic interactions with Trp90 and Gln49 from the partner subunit and a weak hydrogen bond with the side chain of Gln49, which may help to stabilize the dimeric molecule (Fig. 7). In EcHPPK, this residue is Leu79, the side chain of which does not have the length or properties to do so. The side-chain amide of Gln88 in YpHPPK (Glu88 in EcHPPK) forms two hydrogen bonds with the main-chain carbonyl O atom of Lys86 of the partner subunit, which also helps to stabilize the dimer. Glu88 of EcHPPK could function similarly in this regard should dimerization occur.

Asp72 and Asp110 may not be important for dimerization. Asp72 in YpHPPK and Gln72 in EcHPPK make similar interactions with other residues (not shown). The carboxylate of Asp110 in YpHPPK forms one hydrogen bond with the main-chain NH of Arg111, whereas that of Glu110 in EcHPPK only interacts with solvent.

4. Conclusions

YpHPPK and EcHPPK are homologous proteins with 61% identity in their amino-acid sequences. The folds of the two HPPKs are almost identical except for the $_{80}$QGRVRK$_{85}$ sequence. In EcHPPK this sequence is part of Lp3, whereas in YpHPPK it is part of $\alpha2$. The conversion of $_{80}$QGRVRK$_{85}$ from a loop to a helical structure results in the complementary interface between the two subunits. The amino-acid substitutions in the subunit interface favor dimerization by stabilizing Lp3 and increasing the number of hydrogen bonds across the subunit interface. Our thermodynamic data indicate that YpHPPK has lower affinities for both nucleotides and HP than EcHPPK, which is in agreement with the structural data in that the conformational change of Lp2 results in a weakened interaction between Leu45 and HP, and that the different fold of $_{80}$QGRVRK$_{85}$ prevents the side chain of Arg83 from interacting with the nucleotide. Nonetheless, the rate constants for the mechanistic steps of both chemical transformation and product release are comparable between the two HPPKs. It is not clear, however, whether dimerization is involved in the catalysis by YpHPPK. Additional structural and biochemical data are required before the catalytic mechanism can be derived.

We thank Karen Routzahn for construction of the clone and Zbigniew Dauter for help during data collection. X-ray diffraction data were collected at the X9B beamline of the National Synchrotron Light Source, Brookhaven National Laboratory. This work was supported by NIH grant GM51901 (to HY) and the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

Bahadur, R. P., Chakrabarti, P., Rodier, F. & Janin, J. (2003). *Proteins*, **53**, 708–719.
Bahadur, R. P., Chakrabarti, P., Rodier, F. & Janin, J. (2004). *J. Mol. Biol.* **336**, 943–955.
Bermingham, A., Bottomley, J. R., Primrose, W. U. & Derrick, J. P. (2000). *J. Biol. Chem.* **275**, 17962–17967.
Blakley, R. L. & Benkovic, S. J. (1984). Editors. *Chemistry and Biochemistry of Folates.* New York: John Wiley & Sons.
Blaszczyk, J., Li, Y., Shi, G., Yan, H. & Ji, X. (2003). *Biochemistry*, **42**, 1573–1580.
Blaszczyk, J., Li, Y., Wu, Y., Shi, G., Ji, X. & Yan, H. (2004). *Biochemistry*, **43**, 1469–1477.
Blaszczyk, J., Shi, G., Li, Y., Yan, H. & Ji, X. (2004). *Structure*, **12**, 467–475.
Blaszczyk, J., Shi, G., Yan, H. & Ji, X. (2000). *Structure*, **8**, 1049–1058.
Brubaker, R. R. (1991). *Clin. Microbiol. Rev.* **4**, 309–324.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Evdokimov, A. G., Tropea, J. E., Routzahn, K. M. & Waugh, D. S. (2002). *Acta Cryst.* D**58**, 398–406.
Fox, J. D. & Waugh, D. S. (2003). *Methods Mol. Biol.* **205**, 99–117.
Galimand, M., Guiyoule, A., Gerbaud, G., Rasoamanana, B., Chanteau, S., Carniel, E. & Courvalin, P. (1997). *N. Engl. J. Med.* **337**, 677–680.

Garcon, A., Levy, C. & Derrick, J. P. (2006). *J. Mol. Biol.* **360**, 644–653.
Guiyoule, A., Gerbaud, G., Buchrieser, C., Galimand, M., Rahalison, L., Chanteau, S., Courvalin, P. & Carniel, E. (2001). *Emerg. Infect. Dis.* **7**, 43–48.
Hennig, M., Dale, G. E., D'Arcy, A., Danel, F., Fischer, S., Gray, C. P., Jolidon, S., Muller, F., Page, M. G., Pattison, P. & Oefner, C. (1999). *J. Mol. Biol.* **287**, 211–219.
Inglesby, T. V. *et al.* (2000). *JAMA*, **283**, 2281–2290.
Janin, J. & Rodier, F. (1995). *Proteins*, **23**, 580–587.
Janin, J., Rodier, F., Chakrabarti, P. & Bahadur, R. P. (2007). *Acta Cryst.* D**63**, 1–8.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kapust, R. B. & Waugh, D. S. (2000). *Protein Expr. Purif.* **19**, 312–318.
Keskin, O., Ji, X., Blaszcyk, J. & Covell, D. G. (2002). *Proteins*, **49**, 191–205.
Kuzmic, P. (1996). *Anal. Biochem.* **237**, 260–273.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lawrence, M. C., Iliades, P., Fernley, R. T., Berglez, J., Pilling, P. A. & Macreadie, I. G. (2005). *J. Mol. Biol.* **348**, 655–670.
Li, Y., Gong, Y., Shi, G., Blaszczyk, J., Ji, X. & Yan, H. (2002). *Biochemistry*, **41**, 8777–8783.
Li, Y., Wu, Y., Blaszczyk, J., Ji, X. & Yan, H. (2003). *Biochemistry*, **42**, 1581–1588.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perry, R. D. & Fetherston, J. D. (1997). *Clin. Microbiol. Rev.* **10**, 35–66.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Shi, G., Gong, Y., Savchenko, A., Zeikus, J. G., Xiao, B., Ji, X. & Yan, H. (2000). *Biochim. Biophys. Acta*, **1478**, 289–299.
Shiota, T. (1984). *Chemistry and Biochemistry of Folates*, edited by R. T. Blakley & S. J. Benkovic, pp. 121–134. New York: John Wiley & Sons.
Talarico, T. L., Ray, P. H., Dev, I. K., Merrill, B. M. & Dallas, W. S. (1992). *J. Bacteriol.* **174**, 5971–5977.
Xiao, B., Shi, G., Chen, X., Yan, H. & Ji, X. (1999). *Structure*, **7**, 489–496.
Xiao, B., Shi, G., Gao, J., Blaszczyk, J., Liu, Q., Ji, X. & Yan, H. (2001). *J. Biol. Chem.* **276**, 40274–40281.