Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Insights into the structural basis of substrate recognition by histidinol-phosphate aminotransferase from *Corynebacterium glutamicum*

Jan Marienhagen,[a]‡ Tatyana Sandalova,[b]‡ Hermann Sahm,[a] Lothar Eggeling[a] and Gunter Schneider[b]*

[a]Institute for Biotechnology, Research Center Juelich, 52425 Juelich, Germany, and [b]Department of Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden

‡ These authors contributed equally to this work.

Correspondence e-mail: gunter.schneider@ki.se

Received 31 January 2008
Accepted 7 April 2008

PDB References: apo-HisC, 3cq4, r3cq4sf; PMP-HisC, 3cq5, r3cq5sf; PLP-HisC, 3cq6, r3cq6sf.

Histidinol-phosphate aminotransferase (HisC) is a pyridoxal 5′-phosphate-dependent enzyme that catalyzes the reversible transamination reaction between histidinol phosphate (His-P) and 2-oxoglutarate (O-Glu). The crystal structures of apo histidinol-phosphate aminotransferase from *Corynebacterium glutamicum*, of the internal PLP aldimine adduct and of a pyridoxamine 5-phosphate–enzyme complex were determined at resolutions of 2.2, 2.1 and 1.8 Å, respectively. Residues important for substrate specificity were identified by modelling His-P into the active site and comparison with crystal structures of HisC from *Thermotoga maritima* and *Escherichia coli*. Four of the residues lining the substrate-binding pocket were studied by site-directed mutagenesis. Kinetic analysis of the Tyr21Phe mutant suggested that the hydrogen bond between the side chain of this residue and the phosphate group of His-P is important for recognition of the natural substrate and discrimination against other potential amino donors such as phenylalanine and leucine. The mutagenesis studies further indicated that residue Asn99 does not contribute to the specific recognition of the amino-acid donor, but may be involved in binding of the phosphate group of pyridoxal 5′-phosphate. The conserved residues Tyr123 and Tyr257 interact with the substrate through van der Waals interactions and their potential for hydrogen-bonding interactions is not utilized in substrate recognition, as the corresponding phenylalanine mutants show only a moderate effect on the catalytic efficiency k_{cat}/K_m.

1. Introduction

Pyridoxal 5′-phosphate-dependent aminotransferases catalyze transamination reactions, playing key roles in amino-acid metabolism (Christen & Metzler, 1985). During the first half of the reaction, pyridoxal 5′-phosphate (PLP) accepts the α-amino group of an amino acid, which is usually L-glutamate (L-Glu), to form pyridoxamine 5′-phosphate (PMP) and the corresponding 2-oxoacid. In the second half of the reaction, the amino group of PMP is then transferred to another 2-oxoacid to yield a new amino acid, thereby regenerating PLP (Fig. 1). An intriguing feature of the reactions catalyzed by these enzymes is that structurally rather different molecules have to be recognized at the active site. For instance, aromatic amino-acid aminotransferases recognize bulky and hydrophobic aromatic amino acids, but are also able to specifically bind a smaller charged substrate (L-Glu; Oue *et al.*, 1997; Kawaguchi & Kuramitsu, 1998). Similarly, branched-chain amino-acid transaminases bind L-Glu, but their active sites also accept small and hydrophobic amino acids (Goto *et al.*, 2003). Structural studies on these enzymes revealed that

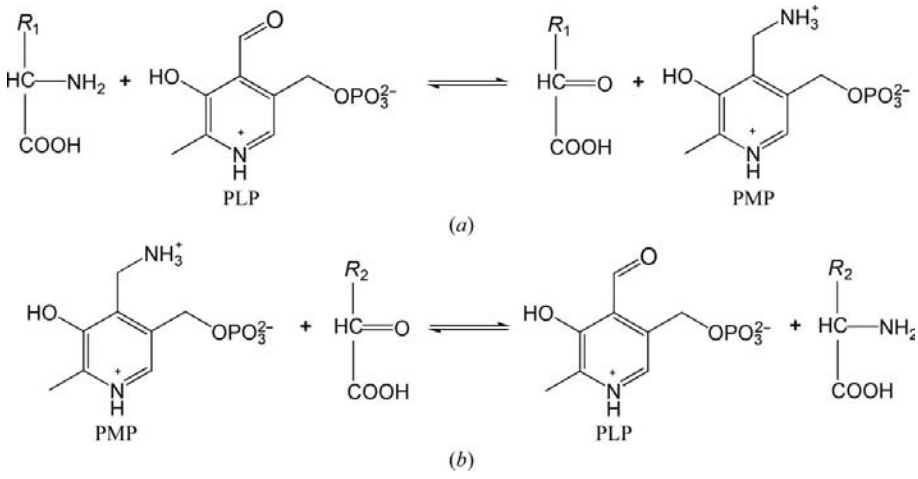

Figure 1
Schematic scheme of a transamination reaction. (*a*) Transfer of the α-amino group of an amino acid (amino donor) to the PLP cofactor of the aminotransferase, resulting in the formation of PMP and a 2-oxoacid. (*b*) Transfer of the amino group from PMP to another 2-oxoacid (amino acceptor) yields a new amino acid and regenerates PLP.

PLP-dependent aminotransferases have evolved two major strategies to deal with this problem (Hirotsu *et al.*, 2005).

In *Escherichia coli* aspartate aminotransferase, a large-scale rearrangement of the hydrogen-bond network at the active-site pocket enables accommodation of the two different amino acids (Malashkevich *et al.*, 1995). This is also the case for the aromatic amino-acid aminotransferase from *Paracoccus denitrificans* (Okamoto *et al.*, 1998). A different strategy of double substrate recognition is used by the branched-chain amino-acid aminotransferase from *E. coli* (Goto *et al.*, 2003). Here, the active site contains both hydrophobic and hydrophilic sites that are able to interact with either substrate without the necessity of a conformational rearrangement. Histidinol-phosphate aminotransferases (HisCs) recognize histidinol phosphate (His-P) and 2-oxoglutarate (O-Glu) in a similar fashion without conformational changes (Haruyama *et al.*, 2001; Sivaraman *et al.*, 2001; Fernandez *et al.*, 2004). The α-carboxylate of L-Glu forms a salt bridge with Arg335 (*E. coli* enzyme numbering), whereas the bulky α-phosphate of His-P interacts with Tyr20 and Arg322 in addition to Arg335. All these residues are conserved in the sequences of HisCs known to date. The γ-carboxylate of the side chain of L-Glu and the imidazole moiety of His-P occupy different binding pockets in the active site of the enzyme. Although these studies (Haruyama *et al.*, 2001; Sivaraman *et al.*, 2001; Fernandez *et al.*, 2004) have provided insights into the structural basis of substrate recognition in histidinol-phosphate aminotransferases, the extent to which individual active-site residues contribute to substrate binding is in most cases not known.

In this study, we addressed this question for the histidinol-phosphate aminotransferase HisC from *Corynebacterium glutamicum*. This Gram-positive bacterium is among the most thoroughly studied organisms used for amino-acid production today. Currently, about 2×10^6 tons of amino acids are produced annually, mostly by fermentation of *C. glutamicum* (Leuchtenberger *et al.*, 2005). Although a transamination reaction is involved in almost every amino-acid biosynthesis pathway, aminotransferases have generally not been the focus of metabolic engineering of *C. glutamicum* and other bacteria, probably since these reactions are close to thermodynamic equilibrium and therefore do not limit flux. In order to address problems of aminotransferase specificity at the molecular level, we have isolated and functionally characterized all the aminotransferases of *C. glutamicum*, together with other PLP-dependent enzymes (Marienhagen *et al.*, 2005). In addition to playing a vital role in the L-histidine pathway by aminating imidazole acetol 3-phosphate, HisC shows also activity with the precursors of aromatic amino acids in *in vitro* enzyme assays (Gu *et al.*, 1995). In such assays, HisC of *C. glutamicum* is also active with the L-Leu precursor 4-methyl-2-oxopentanoate, but lacks activity with the other two branched-chain amino-acid intermediates (Marienhagen *et al.*, 2005). Activity with 4-methyl-2-oxopentanoate was also demonstrated for the histidinol-phosphate aminotransferase of *Zymomonas mobilis in vitro* (Gu *et al.*, 1995). However, in *C. glutamicum* these activities can be regarded as minor activities of this enzyme since they do not contribute to *in vivo* synthesis of the respective amino acids (Marienhagen *et al.*, 2005).

Here, we report three crystallographic structures of HisC of *C. glutamicum*: the apoenzyme, the covalent PLP–enzyme complex and a noncovalent enzyme–PMP complex. These structures formed the basis for modelling substrate binding into the active site and residues suggested to be involved in recognition of the substrate were probed by site-directed mutagenesis.

2. Experimental procedures

2.1. Expression and purification of HisC

Full-length histidinol-phosphate aminotransferase HisC (residues 1–366) from *C. glutamicum* (cgHisC) encoded by the *hisC* gene (NCgl2020) was amplified by standard PCR techniques and cloned into the pASK-IBA-3C and pET28a(+) vectors. Use of pASK-IBA-3C-*hisC* and expression in *E. coli* DH5αMCR yielded a fusion protein of HisC with an uncleavable Strep-tag (Strep-tag sequence WSHPQFEK) at the C-terminus. The cells were grown in 500 ml LB medium containing 25 µg ml^{-1} chloramphenicol at 303 K until A_{550} reached 0.5. Gene expression was induced by adding 10 µl anhydrotetracycline (2 mg ml^{-1}) and cultures were incubated for 4 h at 303 K before harvesting the cells.

Use of pET28a(+)-*hisC* and expression in *E. coli* BL21 (DE3) yielded HisC as a fusion protein with an N-terminal

Table 1
Data-collection and refinement statistics for HisC of *C. glutamicum*.

Values in parentheses are for the highest resolution shell.

	Apo-HisC	PMP-HisC	PLP-HisC
Data collection			
Resolution (Å)	54.9–2.2 (2.3–2.2)	78.0–1.8 (1.9–1.8)	88.0–2.1 (2.2–2.1)
Space group	$P3_221$	$C2$	$C2$
Unit-cell parameters			
a (Å)	102.3	195.3	191.8
b (Å)	102.3	85.5	80.0
c (Å)	140.1	89.4	88.5
α (°)	90.0	90.0	90.0
β (°)	90.0	93.6	94.8
γ (°)	120.0	90.0	90.0
No. of reflections			
Observed	266411	494758	193015
Unique	43528	133625	75438
$\langle I/\sigma(I)\rangle$	25.5 (4.7)	15.9 (2.7)	10.4 (1.8)
Completeness (%)	99.8 (99.7)	97.7 (98.0)	96.9 (96.8)
$R_{\rm merge}$ (%)	5.1 (53.3)	5.9 (45.6)	9.0 (45.0)
No. of monomers per ASU	2	3	3
Solvent content (%)	53.2	60.4	60
Wilson $\langle B\rangle$ (Å^2)	41.2	25.1	28.6
Refinement statistics			
$R_{\rm cryst}$ (%)	23.7	19.2	19.7
$R_{\rm free}$† (%)	27.6	21.9	23.8
No. of atoms			
Total	5656	9302	8838
Protein	5410	8406	8370
Solvent	238	774	361
Ions	8	30	30
PLP	—	48	48
B factors (Å^2)			
Average	44.3	27.6	30.7
Protein main chains	44.0	26.1	30.2
Protein side chains	45.0	27.6	31.3
Sulfate/phosphate	—	47.9	50.4
PLP	—	26.0	24.5
Solvent	40	34.5	30.0
R.m.s.d. from ideal geometry			
Bond lengths (Å)	0.011	0.010	0.010
Bond angles (°)	1.33	1.44	1.23
Ramachandran plot (%)			
Most favoured regions	92.5	92.3	90.8
Disallowed regions	0	0	0

† 5% of reflections were used to monitor the refinement.

His_6 tag and a thrombin-cleavage site for histidine-tag removal after protein purification. For expression, cultures of *E. coli* BL21 (DE3) pET28a(+)-*hisC* cells were grown in 500 ml LB medium including 30 µg ml^{-1} kanamycin at 303 K to A_{600} = 0.6. Gene expression was induced with 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside and cells were harvested after 5 h.

Cell pellets with the HisC-Strep-tag fusion protein were resuspended in 5 ml lysis buffer (100 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 1 m*M* EDTA) per 500 ml culture, whereas the cell pellets containing the HisC-His_6-tag fusion protein were resuspended in 6 ml of a different lysis buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 10 m*M* imidazole) per 500 ml culture. In both cases cells were disrupted by sonication under ice cooling and cytoplasmatic extracts were clarified by centrifugation at 277 K. The HisC-Strep-tag fusion protein was purified using Strep-Tactin Sepharose and the Strep-tag Protein Purification Buffer Set (IBA GmbH, Göttingen, Germany) according to the manufacturer's recommendations. Nickel-chelating metal-affinity chromatography was used for purification of the HisC-His_6-tag fusion protein. A column volume of 2 ml nickel-nitrilotriacetic acid resin per 500 ml expression culture was used for purification. After gravity flowthrough of the cytoplasmatic extract, impurities were washed off the column by applying 16 volumes of wash buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 20 m*M* imidazole). Elution with three columns of elution buffer (50 m*M* NaH_2PO_4 pH 8.0, 300 m*M* NaCl, 250 m*M* imidazole) was directly followed by a buffer exchange to 20 m*M* Tris–HCl pH 8.0 with 150 m*M* NaCl using PD-10 desalting columns (Amersham Biosciences). The N-terminal histidine tag was cleaved by protease treatment at 293 K for 2 h with 1.5 U biotinylated thrombin per milligram of fusion protein. The thrombin was captured by binding to streptavidin agarose (293 K, gentle shaking) and removed using spin filters. The purity of the full-length HisC protein including three additional residues, Gly-Ser-His, resulting from the engineered thrombin-cleavage site was judged by SDS–PAGE analysis. After further protein purification by preparative gel filtration, PLP was added to the HisC-Strep-tag fusion protein as well as to the thrombin-cleaved HisC protein to a final concentration of 1 m*M*, because analysis of the UV–Vis spectrum of purified HisC revealed the absence of PLP after purification. Solutions of HisC were concentrated to 12 mg ml^{-1} using Centricon 10 kDa devices (Millipore) and kept frozen at 193 K.

2.2. Site-directed mutagenesis

Site-directed mutagenesis was carried out in the pASK-IBA-3C-*hisC* vector using the QuikChange kit (Stratagene) with mutagenic oligonucleotides carrying the desired mutations following the manufacturer's recommendations. In all cases, the mutant genes were sequenced to verify the absence of unintended mutations. Expression and purification of the mutant enzymes was carried out as for the wild-type Strep-tagged HisC fusion protein.

2.3. Specific activites and determination of kinetic parameters

Specific activities of cgHisC were determined for the biosynthetic reactions using a previously described aminotransferase assay (Marienhagen *et al.*, 2005). Amino acids were quantified by high-pressure liquid chromatography as their *o*-phthaldialdehyde derivatives. Assays were linear over time and were proportional to the protein concentration. Kinetic parameters, such as the Michaelis–Menten constant ($K_{\rm m}$), the turnover number ($k_{\rm cat}$) and the catalytic efficiency ($k_{\rm cat}/K_{\rm m}$) with varying L-amino-acid and His-P concentrations were determined with the amino acceptor 2-oxoglutarate (O-Glu) held constant at a saturating concentration of 4.0 m*M*.

2.4. Crystallization of cgHisC

All crystallization experiments with cgHisC were carried out by vapour diffusion at 293 K. Crystals of Strep-tagged HisC were obtained under several different conditions, all of

Table 2
Specific activities of cgHisC with various amino donors and amino acceptors.

O-But, 2-oxobutyrate; O-Glu, 2-oxoglutarate; O-Ile, 3-methyl-2-oxopentanoate; O-Leu, 4-methyl-2-oxopentanoate; O-Phe, phenylpyruvate; O-Trp, indolepyruvate; O-Tyr, 4-hydroxyphenylpyruvate; O-Val, 3-methyl-2-oxobutanoate.

Amino acceptor	Amino donor	Specific activity (μmol min^{-1} mg^{-1})
O-Glu	His-P	2.8
O-Phe	L-Glu	1.5
O-Tyr	L-Glu	0.5
O-Trp	L-Glu	0.3
O-Ile	L-Glu	No activity
O-Val	L-Glu	No activity
O-But	L-Glu	0.2
O-Leu	L-Glu	1.4
O-Leu	L-Asp	0.4
O-Leu	L-Ala	0.1

which contained PEG and various salts. The best diffracting crystals were obtained in 9%(*v*/*v*) PEG 6K, 1 *M* lithium acetate, 100 m*M* MES pH 5.8–6.2. Typically, 1–2 µl of a 12 mg ml^{-1} protein solution in 20 m*M* Tris–HCl pH 8 and 0.15 *M* NaCl was mixed in a 1:1 ratio with the crystallization reservoir solution and equilibrated against 1 ml mother liquor. Bipyramidal-shaped crystals grew over several days to dimensions of 0.3 × 0.05 × 0.05 mm. Crystals of the untagged protein grew under two conditions: 1.7 *M* $(NH_4)_2SO_4$, 0.2 *M* Li_2SO_4, 0.1 *M* Tris pH 7.0 and 0.4 *M* NaH_2PO_4, 1.6 *M* K_2HPO_4 and 200 m*M* NaCl buffered in imidazole pH 6.0. For optimal crystal size, mixtures of 2 µl 12 mg ml^{-1} protein solution in 20 m*M* Tris–HCl pH 8, 0.15 *M* NaCl and 1.0 m*M* PLP with 3 µl mother liquor were equilibrated against 1.0 ml reservoir solution.

Subsequent analysis of the crystal structures showed that Strep-tagged HisC crystallized without PLP in the active site, even when 15 m*M* PLP was added to the protein solution. In the following, this structure will therefore be termed apo-HisC. Crystals grown in ammonium sulfate showed electron density for noncovalently bound cofactor and for reasons outlined below we will refer to this structure as PMP-HisC (HisC complex with pyridoxamine 5′-phosphate). Crystals obtained in phosphate buffer clearly showed electron density for the PLP cofactor covalently bound to the active-site Lys228 as the internal aldimine and this structure is denoted holo-HisC throughout the paper.

2.5. Data collection and processing

The data for apo-HisC were collected on beamline ID14-2 (ESRF Grenoble, France; λ = 0.933 Å) using an ADSC Q4 CCD detector. The crystals were soaked for 1–2 s in crystallization solution containing 25%(*v*/*v*) ethylene glycol and frozen directly in a nitrogen cryostream. The crystals of apo-HisC belonged to the trigonal space group $P3_221$, with unit-cell parameters $a = b = 102.3$, $c = 140.1$ Å. Data sets for PMP-HisC and holo-HisC were collected on beamline ID23-1, ESRF (λ = 0.939 Å). Prior to freezing in a cryostream, the crystals were quickly soaked in well solution containing 30%(*v*/*v*) glycerol. These crystals belonged to space group *C*2, with unit-cell parameters $a = 191.8$, $b = 80.0$, $c = 88.5$ Å, $\beta = 94.8^\circ$ for holo-HisC and $a = 195.3$, $b = 85.5$, $c = 89.4$ Å, $\beta = 93.6^\circ$ for PMP-HisC. All diffraction data were processed with the program *MOSFLM* and scaled and merged with *SCALA* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). The data-collection statistics are given in Table 1.

2.6. Structure determination and refinement

The structure of Strep-tagged apo-HisC was solved by molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997). The estimation of the solvent content in the crystals of apo-HisC suggested that the asymmetric unit contains one dimer. A polyserine model of the dimer of histidinol-phosphate aminotransferase from *Thermotoga maritima* (PDB code 1uu1; Fernandez *et al.*, 2004) was used as a search model. The best solution had a correlation coefficient of 0.195 (*R* factor of 0.577). Ten cycles of rigid-body refinement of the individual subunits using *REFMAC*5 and data in the 45–2.5 Å resolution interval (Murshudov *et al.*, 1997) resulted in a decrease of R_{free} by 4%. Tight noncrystallographic symmetry restraints for all backbone atoms were used at the beginning of the refinement. At this stage, the model was rebuilt using the programs *O* (Jones *et al.*, 1991) and *Coot* (Emsley & Cowtan, 2004). In subsequent refinement cycles, R_{free} fell to 30%. Seven residues from the C-terminal Strep-tag could be modelled into the electron-density map. At this stage of refinement, noncrystallographic symmetry restraints were only used for residues 1–300. Water molecules and one acetate ion were modelled with *ARP*/*wARP* (Cohen *et al.*, 2004).

The refined structure of apo-HisC was used as a template for the structure determination of PMP-HisC by molecular replacement. The Matthews coefficient suggested the presence of two dimers in the asymmetric unit, with a solvent content of 47%. However, *MOLREP* was unable to find a solution for the second dimer when the apo-HisC dimer was employed as a search model. When using only a monomer of apo-HisC as a template, *MOLREP* found a solution (*R* value 51%) with three monomers in the asymmetric unit, two of them forming a dimer and the third subunit related to a neighbouring subunit by the crystallographic twofold axis. Several rounds of refinement and model building produced an atomic model with good stereochemistry and *R* and R_{free} values of 22% and 26%, respectively. Inspection of the electron-density map revealed the presence of noncovalently bound cofactor in the active site, which was modelled as PMP (see below). In addition to the cofactor, two strong peaks (>5σ) per monomer were found, which were modelled as sulfate ions. Residual electron density in the active site was modelled as a Tris molecule from the crystallization buffer. The addition of 774 water molecules resulted in a final model with *R* and R_{free} values of 19.2% and 21.9%, respectively, for all diffraction data between 78 and 1.8 Å (Table 1).

The refined model of PMP-HisC (protein only) was used in *MOLREP* to a find solution for the holo-HisC data set. The

refinement procedure was identical to that outlined above for the PMP-HisC data set. Inspection of the electron density showed strong peaks at locations corresponding to those of the sulfate ions of PMP-HisC. These peaks were modelled as phosphate ions from the crystallization buffer. The electron density in the active site clearly indicated the presence of a covalent bond between PLP and the ε-amino group of Lys228. Refinement and model statistics for all data sets are summarized in Table 1.

All figures were made with *PyMOL* (DeLano, 2002). The crystallographic data have been deposited with the PDB with codes 3cq4 (apo-HisC), 3cq5 (PMP-HisC) and 3cq6 (PLP-HisC).

3. Results and discussion

3.1. General characteristics of HisC

Purified recombinant HisC of *C. glutamicum* was used to characterize the substrate-specificity of the enzyme by direct HPLC quantification of the amino acid formed over time. In these experiments the biosynthetic reaction was measured, with the exception of His-P because the substrate 3-(imidazol-4-yl)-2-oxopropyl phosphate was not accessible (Table 2). HisC is the only aminotransferase active in L-His formation in *C. glutamicum* since a *hisC*-deletion mutant of *C. glutamicum* requires L-His for growth (data not shown), similar to other bacteria such as *E. coli* or *Bacillus subtilis* (Garrick-Silversmith & Hartman, 1970; Weigent & Nester, 1976). The activity with phenylpyruvate (O-Phe) and the other aromatic oxoacids does not contribute to aromatic amino-acid synthesis in *C. glutamicum in vivo* (Marienhagen *et al.*, 2005), possibly owing to the low affinity of HisC (see below). Note the significant activity (1.4 U mg^{-1}) with 4-methyl-2-oxopentanoate (O-Leu) as amino acceptor and L-Glu as amino donor.

In order to characterize the kinetic properties, substrate-saturation kinetics with varying concentrations of histidinol phosphate, L-Phe or L-Leu were performed at constant O-Glu concentrations (Table 3). The K_m value for histidinol phosphate is similar to those of the enzymes from *T. maritima* (Fernandez *et al.*, 2004) and *Z. mobilis* (Gu *et al.*, 1995). In contrast, the K_m value for L-Phe is about two orders of magnitude less than that for histidinol phosphate, which again agrees with the properties of the enzymes from *Z. mobilis* and *T. maritima*. This suggests that an active-site pocket is present in HisC that specifically recognizes the imidazole ring. Furthermore, the enzyme has a catalytic efficiency, k_{cat}/K_m, with L-Leu of the same order of magnitude as determined for L-Phe, but it is inactive with L-Ile or L-Val. An activity with L-Leu is also known for HisC of *Z. mobilis*, but information on the other two branched-chain amino acids is not available (Gu *et al.*, 1995).

3.2. Quality of the electron-density map and the models

The quality of the final electron-density map is very good for all three structures and most of the residues are well defined in the electron-density map. The most complete model is chain *B* of holo-HisC, which contains all 366 residues. In the other subunits of PMP-HisC and holo-HisC all except the two N-terminal residues could be fitted to the electron-density map. One cofactor molecule and two sulfate/phosphate ions per monomer were found in every subunit of holo-HisC. One of the oxyanions is bound at the active-site entrance; the others are located at the surface of the protein, bound either to two arginine residues (Arg12*A* and Arg114*B*) from two different subunits of a dimer or between residues Arg333 and Trp322 from the same subunit. The three monomers in the asymmetric unit of the crystals of space group *C*2 are very similar to each other, with typical r.m.s.d. values of 0.2 Å for 364 equivalent C^{α} atoms.

The apoenzyme was crystallized as a fusion protein carrying a C-terminal Strep-tag, a construct that consists of 376 residues. The polypeptide chains of the dimer in the asymmetric unit are very similar to each other (r.m.s.d. 0.3 Å), except for the C-terminal Strep-tag, which displays differences in conformation that are most likely to be a consequence of crystal packing. Unlike holo-HisC, apo-HisC is less well ordered and contains two loops which are not visible in the electron-density map: one peptide stretch comprising residues 18–28 and a short connection of three residues (63–65) between two helices. Finally, the two C-terminal residues of chain *A* and the last ten residues of chain *B* are not well defined in the electron density. However, several residues from the Strep-tag at the C-terminus of HisC are clearly seen in the electron-density map in both subunits. The final model of apo-HisC contains 357 residues of subunit *A*, 347 residues of subunit *B*, 238 water molecules and two acetate ions.

3.3. Structure of the monomer

HisC from *C. glutamicum* belongs to the fold-type I subfamily of vitamin B_6-dependent enzymes (Grishin *et al.*, 1995; Schneider *et al.*, 2000). The HisC subunit consists of an N-terminal arm (residues Met1–Val28) which protrudes from the core of the monomer towards the second subunit, a large PLP-binding domain (residues Pro38–Arg272) and a small C-terminal domain (residues Asp29–Asn37 and His273–Leu366; Fig. 2*a*). The N-terminal arm is well ordered and the first five residues adopt an extended conformation followed by one helical turn (residues 6–10). These residues cover a hydrophobic patch of the neighbouring subunit (Phe109, Pro217, Ala218 and Phe246) and shield it from the solvent. The side chain of Arg16 serves as a C-cap of the one-turn helix (Leu6–Pro10). The N-terminal extension is connected to the main body of the protein by a long flexible loop which is invisible in the electron-density map of apo-HisC, *i.e.* 36 Å separate the last visible residues Gly17 and Asp29. However, this loop is well defined in the complexes of the enzyme with PLP.

The large domain of HisC is characterized by an $\alpha/\beta/\alpha$ architecture with a seven-stranded mixed β-sheet (strands β2–β8; strand 8 is antiparallel to all the others) flanked on both sides by several α-helices. The small domain consists of a four-

Table 3
Catalytic efficiencies of wild-type cgHisC and four cgHisC mutants with three different substrates as amino donors and O-Glu as acceptor.

	Wild type	Y21F	N99D	Y123F	Y257F
His-P					
K_m (mM)	0.89 ± 0.1	6.7 ± 0.4	9.1 ± 0.5	1.3 ± 0.3	2.2 ± 0.3
k_{cat} (s^{-1})	1.18 ± 0.1	1.93 ± 0.1	1.86 ± 0.1	0.39 ± 0.05	0.51 ± 0.05
k_{cat}/K_m (M^{-1} s^{-1})	1326 ± 94	139 ± 21	204.7 ± 15	313.9 ± 8	232.8 ± 38
L-Leu					
K_m (mM)	94.4 ± 12.2	80.8 ± 4.8	55.8 ± 6.1	53.4 ± 7.3	32.8 ± 5.2
k_{cat} (s^{-1})	3.27 ± 0.2	1.72 ± 0.1	1.6 ± 0.2	0.66 ± 0.05	0.69 ± 0.03
k_{cat}/K_m (M^{-1} s^{-1})	34.8 ± 5.4	21.3 ± 0.7	28.7 ± 3.9	12.4 ± 1.4	20.1 ± 1.9
L-Phe					
K_m (mM)	106 ± 17.8	25.5 ± 1.8	71.4 ± 9.7	27.5 ± 5.2	22.2 ± 2.3
k_{cat} (s^{-1})	4.1 ± 0.3	0.87 ± 0.1	2.22 ± 0.4	0.59 ± 0.04	0.68 ± 0.02
k_{cat}/K_m (M^{-1} s^{-1})	38.7 ± 4.2	34.1 ± 1.9	31.1 ± 5	21.7 ± 1.2	30.8 ± 1.2

stranded antiparallel β-sheet (strands β9–β12), with topology 1–2–4–3, and four α-helices. One side of the β-sheet is packed against the PLP-binding domain, while the second side is covered by α-helices. The β-sheets of the two domains are almost orthogonal to each other. Residues 30–32 (strand β1) form main-chain–main-chain hydrogen bonds with the N-terminal half of strand β11, but are not part of the sheet (Fig. 2*a*).

The closest relatives of known structure of HisC from *C. glutamicum* are the histidinol-phosphate aminotransferases from *T. maritima* (31% sequence identity; Fernandez *et al.*, 2004) and *E. coli* (29% sequence identity; Haruyama *et al.*, 2001). Superposition of the enzyme subunits gives typical r.m.s.d. values of 1.6–1.7 Å. HisC from *C. glutamicum* is 30 residues longer than the enzyme from *T. maritima*, with extensions at the N- and C-termini and insertions in loop regions. A multiple alignment of the sequences of HisC from *C. glutamicum* with those of the most similar aminotransferases of known three-dimensional structure is shown in Fig. 3. The most conserved parts of these sequences are located at the dimer interface and the PLP-binding site.

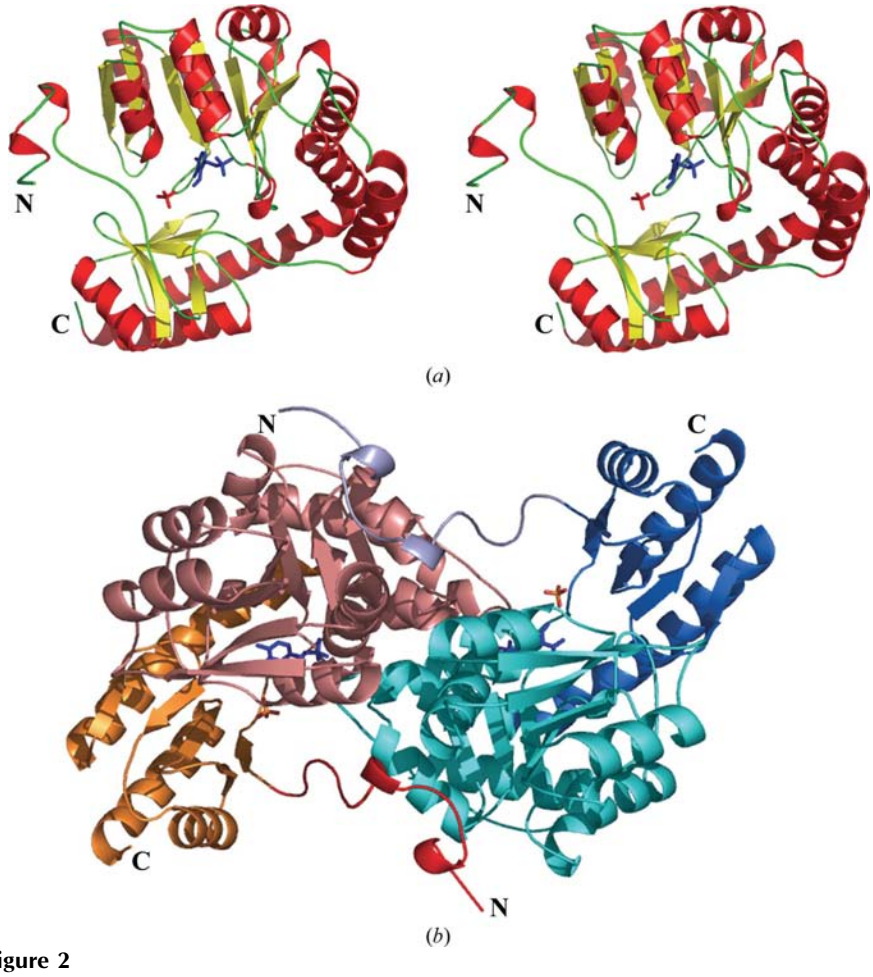

Figure 2
Cartoon representation of the monomer (*a*) and dimer (*b*) of HisC from *C. glutamicum*. PLP and bound phosphate ions are shown as blue and red stick models. (*a*) Stereoview of the monomer coloured according to secondary structure, with β-strands in yellow and α-helices in red. (*b*) The N-terminal arms of the subunits in the dimer are coloured red and light blue, the PLP domains in blue and salmon and the C-terminal domains in dark blue and orange.

3.4. Quaternary structure of HisC

HisC forms a dimer in the crystal (Fig. 2*b*), consistent with data for the protein in solution as shown by gel filtration, and the two subunits are related by a twofold noncrystallographic symmetry axis. The HisC dimer has approximate dimensions of 95 × 50 × 45 Å. Extensive hydrophobic interactions and 24 intersubunit hydrogen bonds including one salt bridge contribute to the large subunit–subunit interface (buried surface area 3650 $Å^2$). A significant part of the interface region is formed by the N-terminal arm that packs against the large domain of the second subunit. This polypeptide segment is responsible for 23% of the buried interface area and contributes four hydrogen bonds. Several other regions of the chain participate in dimerization. Four helical segments from the large domain, α3 including flanking loops (residues 36–66), α5 (97–110), α6 (127–134) and α9 (243–252), and one helix from the small domain, α10 (260–273), make contacts with the other monomer.

3.5. The Strep-tag contributes to crystal packing

Apo-HisC was crystallized as the Strep-tag construct. Five residues in subunit *A* and seven residues in subunit

B from the Strep-tag are clearly seen in the electron-density map (Fig. 4*a*). These residues are bound in extended conformation to the enzyme surface and are located between symmetry-related molecules such that the tag of subunit *A* is close to the entrance of the active site of the crystallographically related subunit *B*′ (Fig. 4*b*) and *vice versa*. The long loop region 18–28 of the N-terminal arm is not visible in the electron-density map for apo-HisC and the Strep-tag is located approximately at the same position as occupied by Tyr21 in holo-HisC. The orientation of helices α12 and α13 differs from that observed in the holoenzyme. The loop between α1 and α2 is flexible, as indicated by the lack of electron density. This stretch includes Tyr63, which forms a hydrogen bond with bound PLP in the adjacent subunit in the holoenzyme (see below). Formation of this crystal form on the one hand requires binding of the Strep-tag at the active-site surface, but on the other hand selects an enzyme conformation that is unable to sustain the interaction of Tyr63 with PLP. This particular crystal packing may thus explain the inability to obtain crystals of the holoenzyme under these conditions even with a high excess of PLP.

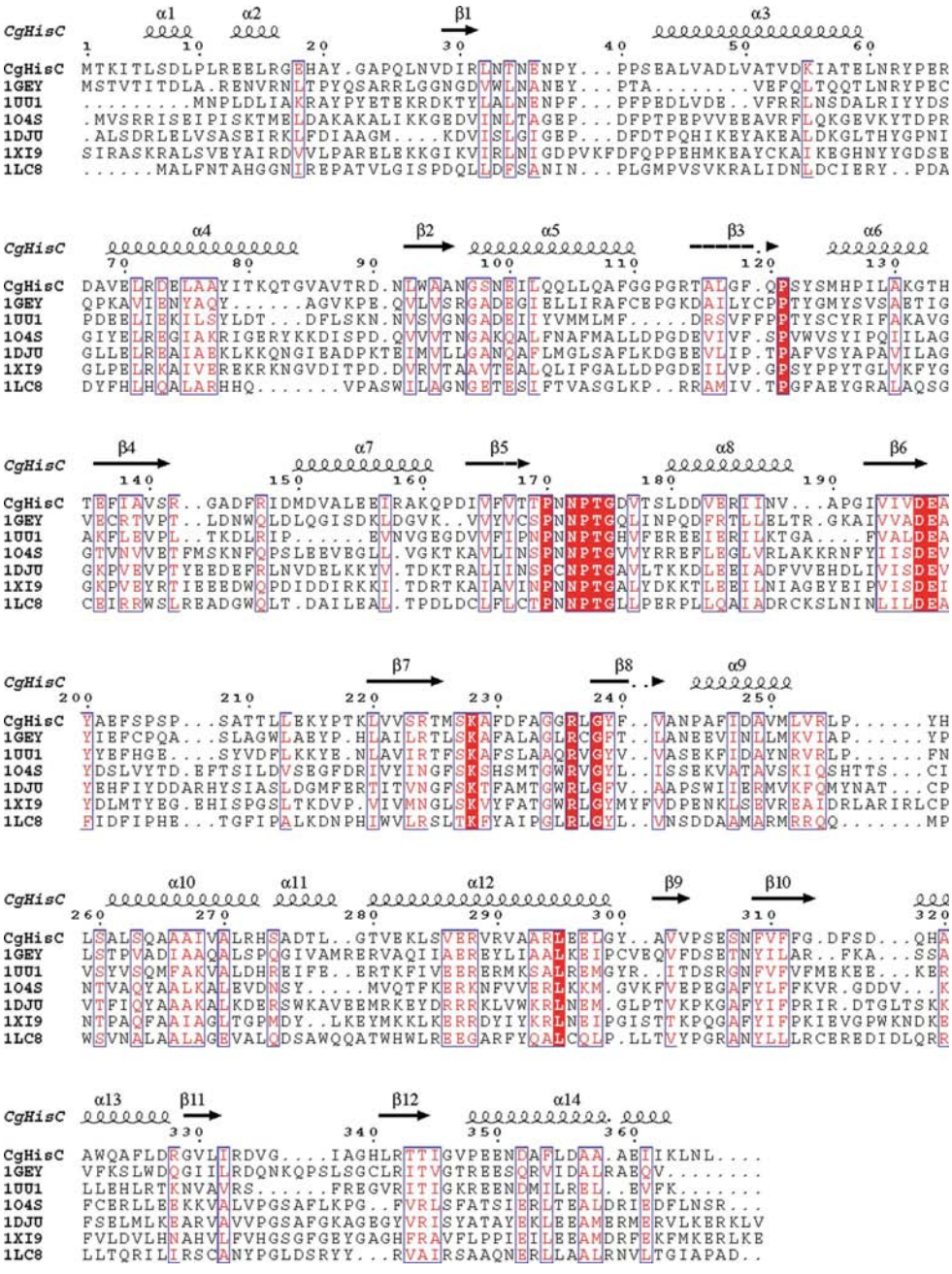

Figure 3
Structure-based amino-acid sequence alignment of HisC from *C. glutamicum* with the sequences of the closest structural relatives. 1UU1, HisC from *T. maritima* (Fernandez *et al.*, 2004); 1GEY, HisC from *E. coli* (Haruyama *et al.*, 2001); 1O4S, aspartate aminotransferase (TM1255) from *T. maritima* (Schwarzenbacher *et al.*, 2004); 1DJU, aromatic aminotransferase from *Pyrococcus horikoshii* (Matsui *et al.*, 2000); 1XI9, alanine aminotransferase from *P. furiosus* (Southeast Collaboratory for Structural Genomics, unpublished work); 1LC8, L-threonine-*O*-3-phosphate decarboxylase from *Salmonella enterica* (Cheong *et al.*, 2002). The secondary structure of HisC from *C. glutamicum* is shown on the top line.

3.6. PLP binding in HisC

The PLP-binding sites are located at the bottom of deep cavities formed at the interfaces between the two domains of each subunit (Fig. 2) and are lined by conserved residues such as Gly97, Ser98, Ala199, Tyr200, Thr225, Lys228 and Arg236 (Fig. 5*a*). The active sites are built up mostly by residues located on the central β-strands of the β-sheet, loops connecting these strands and the N-terminal parts of α-helices α5 and α6. The PLP-binding sites are formed almost exclusively from residues within each subunit, although (for example) the conserved residue Tyr63 from the adjacent subunit participates in a hydrogen bond to the cofactor across the subunit–subunit interface. In addition, the tight packing of the dimer shields the active sites from the solvent.

The cofactor, including the charged phosphate group, is buried in the cavity and is bound to the enzyme by extensive noncovalent interactions with active-site residues (Fig. 5*a*). In the holoenzyme, the side chain of Lys228 also forms a covalent linkage to the cofactor, as observed in other PLP-dependent enzymes (Fig. 5*b*). Surprisingly, in the data set collected from crystals that were obtained using

ammonium sulfate as precipitant this covalent linkage is not observed (Fig. 5*c*). The reasons for the inability of the enzyme to form the internal aldimine under these conditions are not completely clear. The binding mode of PLP and its interactions with the enzyme are very similar in this complex compared with those observed in holo-HisC and the pH of crystallization (pH 7.0) should actually favour aldimine formation compared with that of the holoenzyme, which was crystallized at pH 6.0. Radiation damage seems unlikely as the cause of the lack of the covalent bond between the enzyme and PLP because both data sets were collected on the same beamline using similar exposure times. A major significant difference in the crystallization conditions for the two species is the presence of 3.4 *M* ammonium ions in the crystals lacking the covalent linkage. It is quite possible that ammonium ions at this high concentration compete successfully with the lysine residue and lead to ammonolysis, resulting in the formation of pyridoxamine 5′-phosphate. For this reason, we have modelled the cofactor as PMP rather than PLP.

Comparison of PMP-HisC with holo-HisC shows that Schiff-base formation between PLP and Lys228 did not change the conformation of the protein; superposition of the two structures gives an r.m.s.d. of 0.25 Å for 365 C^{α} atoms. This is in contrast to the much larger r.m.s.d. of 1.26 Å when the structures of the apo-HisC and holo-HisC subunits were compared. Unfortunately, it is difficult to separate the effect of the Strep-tag on the one hand and the absence of PLP in the active site on the other hand on the structure of the protein. The positions of many residues involved in PLP binding differ significantly in apo-HisC and holo-HisC. The side chain of Tyr123, which forms a stacking interaction with PLP, points away from the active site in apo-HisC and the entire loop connecting β3 and α6 has a different conformation, resulting in a shift of the position of Ser124 of about 5 Å between apo-HisC and holo-HisC. Two other loops close to the active site show significantly different conformations in apo-HisC and holo-HisC (residues 229–236) or are disordered in the apoenzyme (residues 58–66).

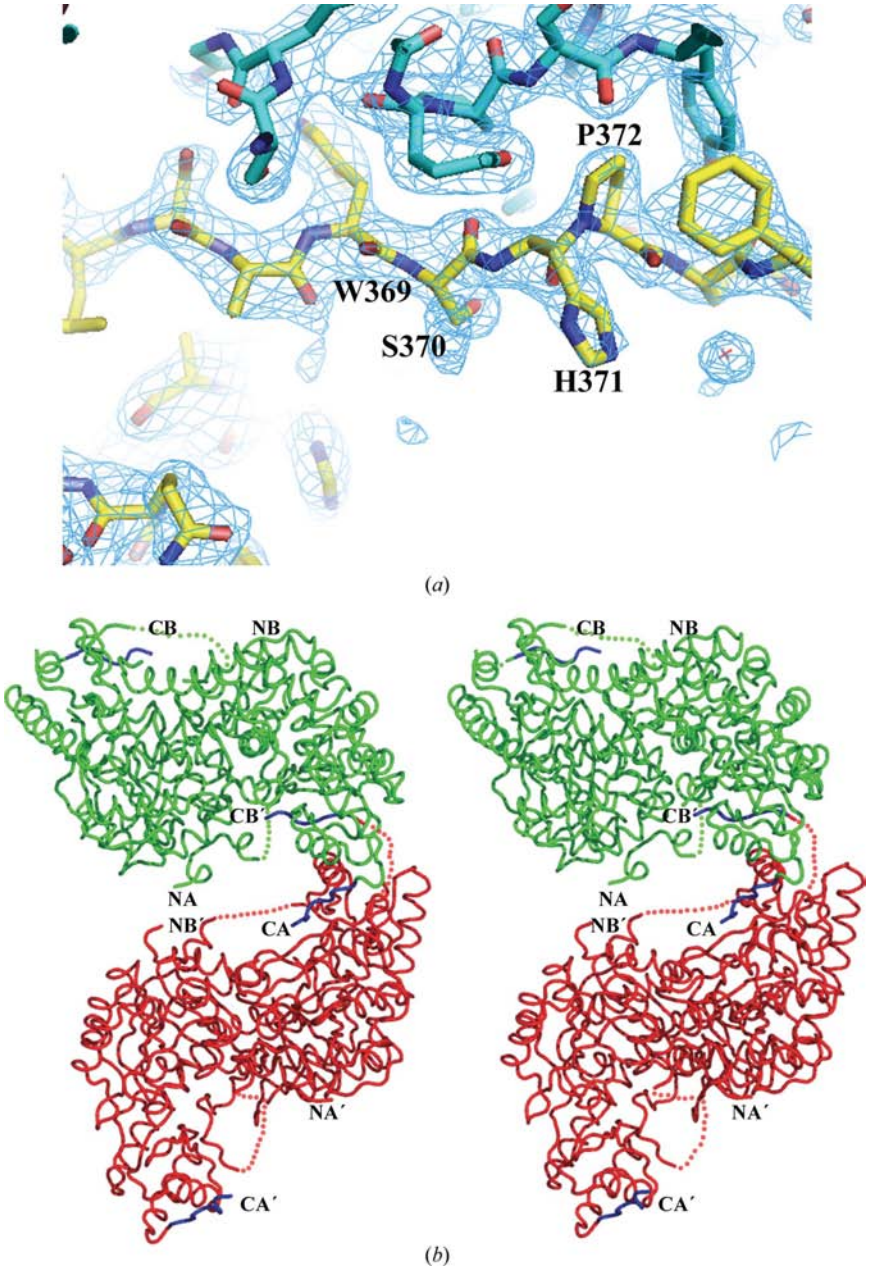

Figure 4
Crystal-packing interactions mediated by the C-terminal Strep-tag. (*a*) Electron density for the C-terminal residues of the Strep-tag. The $2F_o - F_c$ map is contoured at 1.2σ. Symmetry-related molecules are shown in cyan. (*b*) Stereoview of the interactions of the Strep-tag with a symmetry-related molecule in the crystal lattice. Two symmetry-related HisC dimers (labelled *A* and *B*′) are shown in green and red, respectively. The N- and C-termini are indicated by the letters N and C. The Strep-tags are shown in blue.

3.7. Structural basis for substrate specificity in histidinol-phosphate aminotransferases

Attempts to determine the crystal structures of complexes of cgHisC with substrates were unsuccessful. Electron-density maps calculated from data collected from crystals that had been cocrystallized or soaked with substrate did not show any bound ligands at the active site. We therefore modelled these complexes (Fig. 6) based on comparisons with the structures of HisC from *T. maritima* with bound His-P (PDB code 1uu1; Fernandez *et al.*, 2004) and of *E. coli* HisC in complex with His-P (PDB codes 1fg3 and 1gex) and L-Glu (PDB code 1gey) (Haruyama *et al.*, 2001; Sivaraman *et al.*, 2001), with subsequent energy minimization using *REFMAC*. The position of the phosphate group of the His-P coincides with that of the phosphate/sulfate/acetate ions

observed in the *C. glutamicum* HisC structures. The phosphate group is bound in an anion-binding pocket defined by the side chains of the conserved residues Arg333, Arg342, Tyr21 and Asn172. The amino-acid side chains of the substrates are bound in a pocket lined by the side chains of Tyr21, Tyr63, Tyr123, Tyr257, Asn99 and Arg236. The imidazolium ring of His-P is held in place by a hydrogen bond from the NE2 atom to the side chain of Asn99 (Fig. 6*a*), while the side chains of the L-Leu and L-Phe substrates interact predominantly with the hydrophobic elements of this pocket, in particular the aromatic ring systems of Tyr21, Tyr63 and Tyr257 (Figs. 6*b* and 6*c*).

HisC from *C. glutamicum* can accept L-Leu and L-Phe as substrates, although with K_m values that are two orders of magnitude higher than that of the natural substrate His-P (Table 3). The models of the external aldimine of PLP with L-Leu and L-Phe, respectively, suggest that the carboxyl groups of the amino-acid substrates are bound in the anion-binding site, but form fewer hydrogen-bonding interactions compared with the phosphate group of His-P (Figs. 6*b* and 6*c*). Only the side chains of Arg342 and Asn172 are within hydrogen-bonding distance of the carboxyl group of the substrate and the decrease from four to two hydrogen bonds may in part be responsible for the lower affinity for these amino-acid substrates, which is consistent with the observed increase in the K_m values.

In order to further explore the underlying structural basis for the specificity of HisC with respect to amino-acid substrates, we carried out site-directed mutagenesis experiments of several residues involved in substrate binding and recognition (Table 3). The side chain of Tyr21, which is conserved among histidinol-phosphate aminotransferases, appears to be particularly well placed for the recognition of His-P owing to the hydrogen bond with the phosphate group (Fig. 6*a*). Substitution of tyrosine by phenylalanine led to a mutant with a tenfold reduction in $k_{cat}/K_{m,\ app}$ using His-P as substrate (Table 3), consistent with a role of this residue in substrate recognition. This hydrogen bond is apparently not important for the binding of the L-Leu and L-Phe substrates, since the $k_{cat}/K_{m,\ app}$ values are essentially unchanged upon residue replacement. This observation is consistent with the models (Fig. 6), which suggest that the side chain of Tyr21 interacts with the phosphate group of His-P but is unable to form a similar hydrogen bond to the carboxyl groups of the nonphosphorylated substrates L-Leu and L-Phe owing to the longer distance. Removal of this hydrogen bond by mutagenesis is thus expected to influence the reaction with His-P but not L-Leu or L-Phe, which indeed is observed.

Residue Asn99 is in contact with the phosphate group of PLP as well as with the substrate His-P (Figs. 6*a* and 6*b*). The HisC sequences of *E. coli* and *T. maritima* naturally have an aspartate residue at this position (Sivaraman *et al.*, 2001) that is positioned similarly as observed in the structure of the *C. glutamicum* enzyme. The exchange of Asn99 by aspartate in cgHisC by site-directed mutagenesis results in a variant that retains significant activity with His-P as substrate. Remarkably, the introduction of the negative charge as part of the pocket that binds hydrophobic aromatic or aliphatic side chains does not cause a significant loss of activity with the amino acids L-Phe and L-Leu (Table 3). However, removal of the asparagine side chain in the Asn99Gly mutant resulted in a complete loss of activity with all substrates, emphasizing the critical role of this residue, most likely in anchoring the PLP phosphate group.

The pyridine ring of PLP is stacked with the aromatic side chain of Tyr123. Sequence alignments suggest that this interaction is essential since an aromatic amino acid, mostly Tyr, is conserved at this location (Fig. 3). In the amino-acid sequences of *Z. mobilis* and *B. subtilis* HisC a phenylalanine residue is located at this position and Sivaraman *et al.* (2001) proposed that this residue substitution might

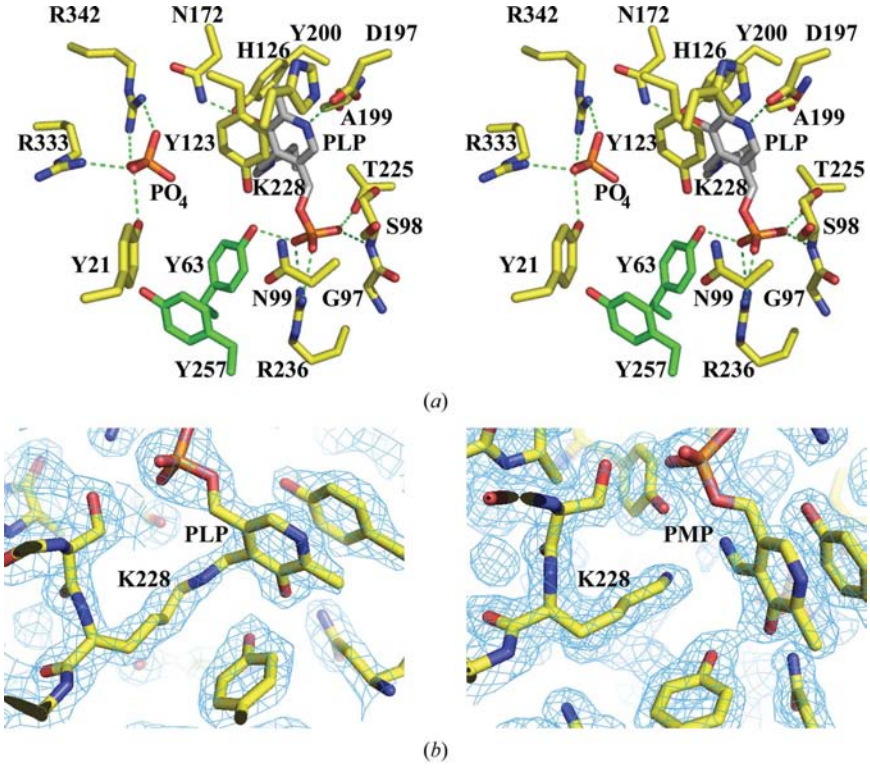

Figure 5
(*a*) Stereoview of the PLP-binding site in HisC of *C. glutamicum* as observed in the holo-HisC complex. The anion-binding site is indicated by the position of the bound phosphate ion. Dotted lines indicate hydrogen bonds (<3.2 Å distance). Residues from the second subunit are shown in green. (*b*) OMIT map contoured at 1.2σ at the PLP-binding site showing the covalent linkage between the catalytic lysine residue and the PLP cofactor. (*c*) OMIT map contoured at 1.2σ in the PMP complex, illustrating the absence of a covalent bond between the cofactor and the lysine residue.

contribute to the observed broader substrate specificity for aromatic amino acids in these two enzymes (Gu *et al.*, 1995; Weigent & Nester, 1976). Amino-acid replacement by site-directed mutagenesis at this position (Tyr123Phe) led to only moderate changes in kinetic constants with the natural substrate and the reactions with aromatic substrates were influenced even less (Table 3). These findings suggest that amino-acid exchanges occurring at position 123 in His-P aminotransferases are not a significant determinant of substrate-specificity, because larger effects on the kinetic parameters would have been expected.

Tyr257 is conserved in most HisC sequences and forms part of the walls of the specificity pocket. The effects of amino-acid replacement at position 257 were similar to that of the Tyr123Phe mutant; *i.e.* removal of the hydroxyl group of the tyrosine side chain by substitution by phenylalanine has only minor effects on the specificity of HisC (Table 3). This outcome confirms that the interactions of this residue with the substrate are nonspecific van der Waals interactions and that the hydrogen-bonding potential of the tyrosine side chain is not exploited in substrate recognition by HisC.

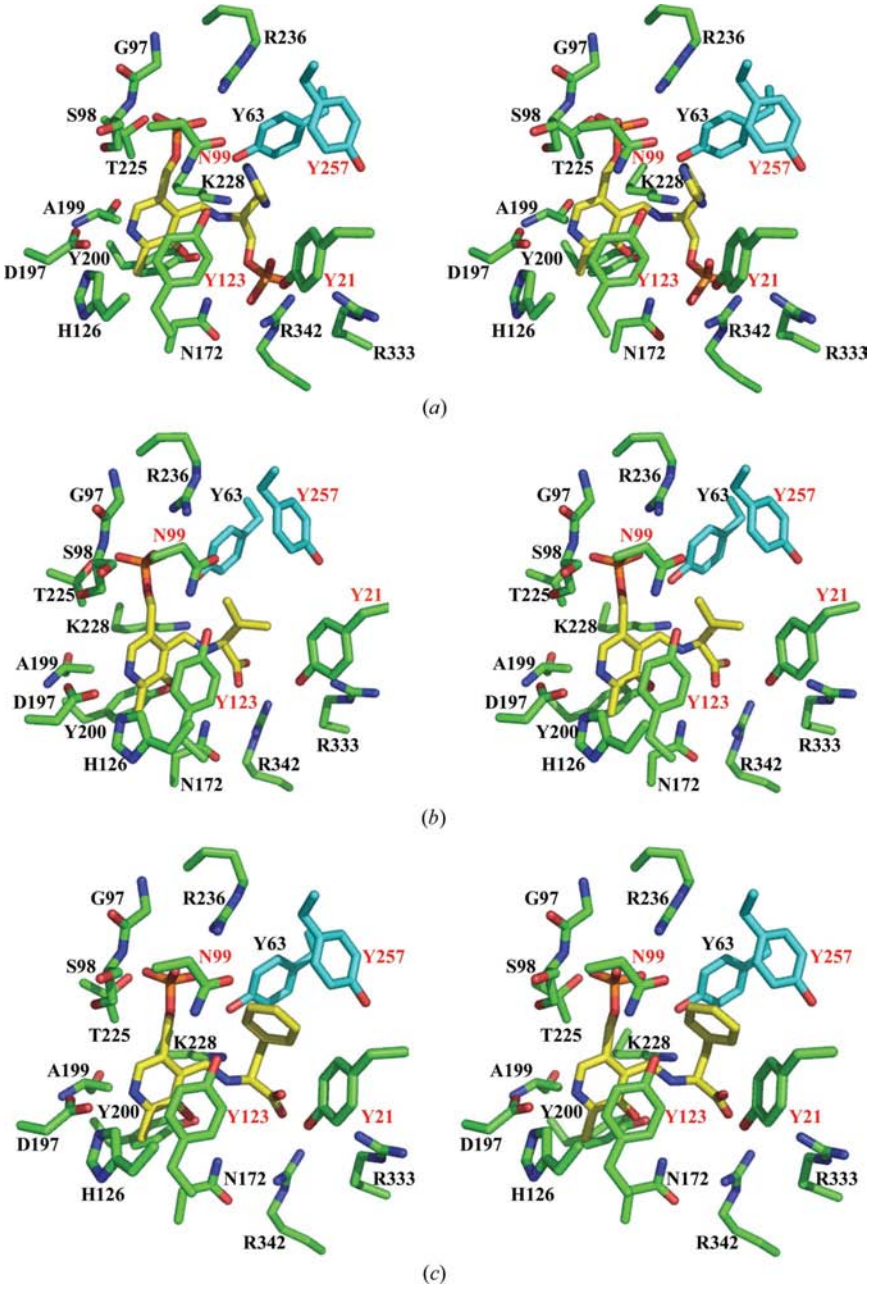

Figure 6
Models of the external aldimine formed in the active site of HisC of *C. glutamicum*. The models are based on a comparison with the structures of the His-P complexes with HisC from *E. coli* (Sivaraman *et al.*, 2001; Haruyama *et al.*, 2001) and *T. maritima* (Fernandez *et al.*, 2004). (*a*) Model of the external aldimine formed with PLP and His-P in *C. glutamicum* HisC, (*b*) external aldimine formed with L-Leu and (*c*) external aldimine formed with L-Phe. Residues from different subunits are shown in green or cyan, respectively. Residues that were mutated are labelled in red.

We gratefully acknowledge access to synchrotron radiation at the European Synchrotron Radiation Facility, Grenoble, France. This work was supported by the Deutscher Akademischer Austauschdienst (DAAD), the Deutsche Bundestiftung Umwelt (DBU) Projekt 13158 and the Swedish Research Council.

References

Cheong, C. G., Escalante-Semerena, J. C. & Rayment, I. (2002). *Biochemistry*, **41**, 9079–9089.
Christen, P. & Metzler, D. E. (1985). *Transaminases.* New York: John Wiley & Sons.
Cohen, S. X., Morris, R. J., Fernandez, F. J., Ben Jelloul, M., Kakaris, M., Parthasarathy, V., Lamzin, V. S., Kleywegt, G. J. & Perrakis, A. (2004). *Acta Cryst.* D**60**, 2222–2229.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fernandez, F. J., Vega, M. C., Lehmann, F., Sandmeier, E., Gehring, H., Christen, P. & Wilmanns, M. (2004). *J. Biol. Chem.* **279**, 21478–21488.
Garrick-Silversmith, L. & Hartman, P. E. (1970). *Genetics*, **66**, 231–244.
Goto, M., Miyahara, I., Hayashi, H., Kagamiyama, H. & Hirotsu, K. (2003). *Biochemistry*, **42**, 3725–3733.
Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. (1995). *Protein Sci.* **4**, 1291–1304.
Gu, W., Zhao, G., Eddy, C. & Jensen, R. A. (1995). *J. Bacteriol.* **177**, 1576–1584.
Haruyama, K., Nakai, T., Miyahara, I., Hirotsu, K., Mizuguchi, H., Hayashi, H. & Kagamiyama, H. (2001). *Biochemistry*, **40**, 4633–4644.
Hirotsu, K., Goto, M., Okamoto, A. & Miyahara, I. (2005). *Chem. Rec.* **5**, 160–172.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kawaguchi, S. & Kuramitsu, S. (1998). *J. Biol. Chem.* **273**, 18353–18364.
Leuchtenberger, W., Huthmacher, K. & Drauz, K. (2005). *Appl. Microbiol. Biotechnol.* **69**, 1–8.
Malashkevich, V. N., Onuffer, J. J., Kirsch, J. F. & Jansonius, J. N. (1995). *Nature Struct. Biol.* **2**, 548–553.
Marienhagen, J., Kennerknecht, N., Sahm, H. & Eggeling, L. (2005). *J. Bacteriol.* **187**, 7639–7646.
Matsui, I., Matsui, E., Sakai, Y., Kikuchi, H., Kawarabayasi, Y., Ura, H., Kawaguchi, S., Kuramitsu, S. & Harata, K. (2000). *J. Biol. Chem.* **275**, 4871–4879.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K. & Kagamiyama, H. (1998). *J. Mol. Biol.* **280**, 443–461.
Oue, S., Okamoto, A., Nakai, Y., Nakahira, M., Shibatani, T., Hayashi, H. & Kagamiyama, H. (1997). *J. Biochem. (Tokyo)*, **121**, 161–171.
Schneider, G., Käck, H. & Lindqvist, Y. (2000). *Structure*, **8**, 1–6.
Schwarzenbacher, R. *et al.* (2004). *Proteins*, **55**, 759–763.
Sivaraman, J., Li, Y., Larocque, R., Schrag, J. D., Cygler, M. & Matte, A. (2001). *J. Mol. Biol.* **311**, 761–776.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Weigent, D. A. & Nester, E. W. (1976). *J. Biol. Chem.* **251**, 6974–6980.