Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of *Helicobacter pylori* L-asparaginase at 1.4 Å resolution

Prathusha Dhavala and Anastassios C. Papageorgiou*

Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Finland

Correspondence e-mail: tassos.papageorgiou@btk.fi

Received 9 July 2009
Accepted 21 September 2009

PDB References:
L-asparaginase, 2wlt, r2wltsd; 2wt4, r2wt4sd.

Bacterial L-asparaginases have been used in the treatment of childhood acute lymphoblastic leukaemia for over 30 years. Their therapeutic effect is based on their ability to catalyze the conversion of L-asparagine, an essential amino acid in certain tumours, to L-aspartic acid and ammonia. Two L-asparaginases, one from *Escherichia coli* and the other from *Erwinia chrysanthemi*, have been widely employed in clinical practice as anti-leukaemia drugs. However, L-asparaginases are also able to cause severe side effects owing to their intrinsic glutaminase activity. *Helicobacter pylori* L-asparaginase (HpA) has been reported to have negligible glutaminase activity. To gain insight into the properties of HpA, its crystal structure in the presence of L-aspartate was determined to 1.4 Å resolution, which is one of the highest resolutions obtained for an L-asparaginase structure. The final structure has an $R_{\rm cryst}$ of 12.6% ($R_{\rm free}$ = 16.9%) with good stereochemistry. A detailed analysis of the active site showed major differences in the active-site flexible loop and in the 286–297 loop from the second subunit, which is involved in active-site formation. Accordingly, Glu289, Asn255 and Gln63 are suggested to play roles in modulating the accessibility of the active site. Overall, the structural comparison revealed that HpA has greater structural similarity to *E. coli* L-asparaginase than to any other L-asparaginase, including *Er. carotovora* L-asparaginase, despite the fact that the latter is also characterized by low glutaminase activity.

1. Introduction

Bacterial L-asparaginases (amidohydrolases; EC 3.5.1.1) are effective anti-leukaemia drugs and established agents in the chemotherapy of blood malignancies (Avramis & Panosyan, 2005). They primarily catalyse the deamination of L-asparagine to L-aspartic acid and ammonia. L-Asn is a crucial amino acid for protein, DNA and RNA synthesis and appears to be a cell-cycle-specific requirement for the G1 phase of cell division (Stams *et al.*, 2005). Unlike normal cells, leucosis cells are dependent on an extracellular supply of L-Asn owing to generally low expression of the asparagine synthase gene (Richards & Kilberg, 2006). The administration of asparaginase depletes the concentration of L-Asn in the blood and thereby leads to apoptotic death of the cancer cells.

The anti-leukaemic properties of asparaginase from guinea pig serum were first discovered some 40 years ago (Mashburn & Wriston, 1964; Bridges, 1968) and led to a detailed evaluation of the biochemical properties of this family of enzymes. Further studies showed the existence of two types of bacterial L-asparaginases: I and II (Campbell *et al.*, 1967). Only type II asparaginases exhibit high specific activity against L-asparagine and are used as chemotherapeutics in acute

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the outermost resolution shell.

	In-house	EMBL Hamburg X13	
Data collection			
Wavelength (Å)	1.5418	0.8123	0.8081
Temperature (K)	100	295	100
Space group	*I*222	*I*222	*I*222
Unit-cell parameters			
a (Å)	65.3	64.9	63.6
b (Å)	96.2	96.4	94.8
c (Å)	101.3	101.9	100.3
Resolution (Å)	99.0–2.6 (2.64–2.60)	99–1.8 (1.83–1.80)	99.0–1.40 (1.42–1.40)
Measured reflections	66288	331595	1720324
Unique reflections	9860 (463)	29828 (1443)	59013 (2895)
Completeness (%)	97.5 (93.4)	99.9 (99.4)	99.7 (99.1)
Average redundancy per shell	4.0 (3.2)	11.1 (11.4)	7.1 (8.2)
$\langle I/\sigma(I)\rangle$	12.2 (2.4)	24.8 (3.5)	22.8 (7.1)
$R_{\rm merge}$† (%)	13.4 (42.3)	7.2 (50.2)	9.6 (27.2)
$R_{\rm meas}$‡ (%)		8.1	10.5
Protein molecules	1	1	1
Matthews coefficient (Å^3 Da^{-1})	2.1	2.1	2.1
Wilson *B* factor (Å^2)	41.6	19.9	15.5
Mosaicity (°)		0.13	0.56
Refinement			
Resolution range (Å)		70.0–1.80	20.0–1.40
Reflections (working/test)		28339/1474	56056/2939
$R_{\rm cryst}/R_{\rm free}$§ (%)		15.1/18.5	13.1/16.9
Protein atoms		2428	2531
Water molecules		168	391
L-Asp atoms		9	9
Average isotropic *B* factors (Å^2)			
Main chain		13.2	18.5
Side chain		14.9	22.6
Waters		33.8	36.1
L-Asp		23.3	21.0
R.m.s. deviation from ideality			
Bond lengths (Å)		0.014	0.012
Bond angles (°)		1.32	2.52
Ramachandran plot statistics (%)			
Residues in most favoured regions		90.3	91.3
Residues in additional allowed regions		8.6	8.4
Residues in disallowed regions		1.1	0.3

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*. ‡ Redundancy-independent *R* factor (Diederichs & Karplus, 1997) as obtained using *XDS* (Kabsch, 1993). § $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure factors, respectively. For $R_{\rm free}$, the sum is extended over a subset of reflections (5%) that were randomly selected and omitted from all stages of refinement. The same subset of reflections was used for both the cryogenic and room-temperature data.

lymphoblastic leukaemia (ALL; Duval *et al.*, 2002; Gökbuget & Hoelzer, 2006; Verma *et al.*, 2007). In contrast, type I asparaginases are characterized by enzymatic activity towards both L-asparagine and L-glutamine. Furthermore, type I asparaginases are expressed constitutively in the cytoplasm, whereas type II asparaginases are expressed under anaerobic conditions in the periplasmic space of the bacterial membranes.

The L-asparaginases from *Escherichia coli* (EcAII) and *Erwinia chrysanthemi* (ErA) have been used as drugs in the treatment of childhood ALL (Muller & Boos, 1998). Despite the success achieved, L-asparaginase therapy is often complicated by serious toxicity affecting the liver and pancreas in addition to hypersensitivity reactions and other drawbacks. Much of the toxicity is caused by the intrinsic glutaminase activity of L-asparaginase (Ollenschläger *et al.*, 1988). The glutaminase activities of EcAII and ErA amount to 2% and 10% of their asparaginase activities, respectively (Derst *et al.*, 2000; Aghaiypour *et al.*, 2001). Since glutamine is the main transport form of amino nitrogen in the blood and is also an amino-group donor in many biosynthetic reactions, a prolonged decline in plasma glutamine levels impairs a variety of biochemical functions, especially those of the liver. In view of these facts, the most suitable asparaginase for the treatment of ALL would be one with a high activity, a low $K_{\rm m}$ and a strong preference for L-Asn over L-Gln (Derst *et al.*, 2000). In the search for new L-asparaginases, an L-asparaginase from *Er. carotovora* (EwA) with a relatively low glutaminase activity (approximately 1.5% of its asparaginase activity) has recently been reported (Krasotkina *et al.*, 2004).

Bacterial L-asparaginases form homotetramers of 140–150 kDa molecular weight and are best described as dimers of intimate dimers. The X-ray crystallographic structures of a number of L-asparaginases, including EcAII (Swain *et al.*, 1993), ErA (Aghaiypour *et al.*, 2001), *Wolinella succinogenes* L-asparaginase (WsA; Lubkowski *et al.*, 1996), *Er. carotovora* L-asparaginase (EwA; Kravchenko *et al.*, 2008; Papageorgiou *et al.*, 2008) and related amidohydrolases from *Pseudomonas fluorescence* (PgA; Lubkowski, Wlodawer, Ammon *et al.*, 1994) and *Acinetobacter glutaminasificans* (Lubkowski, Wlodawer, Housset *et al.*, 1994), are available. An atomic resolution (1.0 Å) structure of ErA has been described in detail (Lubkowski *et al.*, 2003). The enzyme monomer consists of ∼330 amino-acid residues arranged in two domains. The active site in each monomer is located at the intersubunit interface between the N- and C-terminal domains of intimate dimers. Crystallographic data on bacterial asparaginases have revealed that their active site preserves a high degree of structural similarity (Aghaiypour *et al.*, 2001). Residues responsible for ligand binding form the rigid part of the active site. The flexible part of the active site (residues 12–30 in EcAII) controls access to the binding pocket and carries the catalytic nucleophile Thr12 that is highly conserved in all L-asparaginases.

Numerous studies of L-asparaginases have been conducted in order to understand the catalytic mechanism and substrate specificity of these enzymes (Ortlund *et al.*, 2000; Sanches *et al.*, 2007). The present report provides structural insights into an L-asparaginase (GenBank accession code NC_000921) from *Helicobacter pylori* strain J99 which has a high affinity for L-Asn (K_m of 19 $\pm$ 2 μM) and negligible glutaminase activity (0.01% of its asparaginase activity) under physiological conditions (Gladilina *et al.*, 2008). Interestingly, it was recently reported that an L-asparaginase from *H. pylori* strain CCUG 17874 with over 90% sequence identity to NC_000921 exhibits lower affinity towards L-Gln compared with EcAII, ErA and EwA (Cappelletti *et al.*, 2008). Importantly, the enzyme showed higher cytotoxicity towards various cell lines compared with EcAII, and a thermal stability that was between those of EcAII and ErA. Based on these results, L-asparaginases isolated from *H. pylori* could be promising for the development of chemotherapeutic drugs against ALL. The structure presented here is a step forward in better understanding the properties of *H. pylori* L-asparaginases. Analysis of the binding site and comparison of L-asparaginases from various sources would allow a better elucidation of their substrate specificity and activity.

2. Materials and methods

2.1. Crystallization conditions

HpA was expressed and purified using ion-exchange chromatography as reported previously (Dhavala *et al.*, 2008). Crystals were grown using the hanging-drop vapour-diffusion method at 289 K by mixing 2.5 µl enzyme solution (3.7 mg ml^{-1} in 10 m*M* HEPES–NaOH pH 7.0) with an equal volume of a reservoir solution consisting of 17.5%(*w*/*v*) PEG 4000, 0.1 *M* magnesium formate, 0.1 *M* HEPES–NaOH pH 7.0. Rod-like crystals appeared after 4 d and reached final dimensions of approximately 0.40 $\times$ 0.15 $\times$ 0.15 mm after 8 d.

2.2. Data collection and processing

An initial data set was collected from a single crystal to 2.6 Å resolution in-house under cryogenic conditions using a MAR imaging plate and a Rigaku RU-200B X-ray generator. Prior to data collection, the crystal was soaked for a few seconds in the crystallization solution supplemented with 25%(*v*/*v*) glycerol for cryoprotection. The crystal was subsequently flash-cooled to 100 K in a liquid-nitrogen stream. A higher resolution (1.4 Å) data set was subsequently collected on station X13 (λ = 0.8081 Å) at EMBL Hamburg, c/o DESY, Germany using a 165 mm MAR CCD detector. A total of 400 images were collected in dose mode using a rotation angle of 0.5° and an exposure time of 10–20 s per image. In addition, a room-temperature data set was collected to 1.8 Å resolution (300 images, 0.5° rotation angle, exposure time of 5–10 s per image) on X13 from a single crystal mounted on a quartz capillary. All data were indexed, processed and scaled using the *HKL* package (Otwinowski & Minor, 1997). Details of the data-collection statistics are given in Table 1. Assuming one molecule in the asymmetric unit, the Matthews coefficient V_M was 2.1 Å^3 Da^{-1}, corresponding to a solvent content of approximately 41%.

2.3. Structure determination and refinement

The structure was determined by molecular replacement using *Phaser* (McCoy *et al.*, 2007) in the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). A polyalanine model of subunit *A* of *W. succinogenes* L-asparaginase (PDB code 1wsa; 55% sequence identity to HpA; Lubkowski *et al.*, 1996) without water molecules was used as a search model. The low-resolution (2.6 Å) data were employed at this stage. The resulting top solution from *Phaser* showed a *Z* score of 22 (RFZ = 14.8, TFZ = 20.4, LLG = 406). After calculation of an initial electron-density map, most of the side chains were visible in the model and manual building was carried out. The program *Coot* (Emsley & Cowtan, 2004) was used for inspection of $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps and rebuilding of the structure. Refinement was initially carried out with *REFMAC* (Murshudov *et al.*, 1997). A randomly selected subset (5%) of reflections were set aside for cross-validation analysis in order to monitor the progress of refinement using the R_{free} factor. The coordinates were submitted to the *TLS Motion Determination* (*TLSMD*) server (http://skuld.bmsc.washington.edu/~tlsmd/) for protein partitioning to perform TLS and restrained refinement in *REFMAC*. The refinement was extended to 1.4 Å in steps of 0.2 Å. When R_{free} reached 19.9% the structure was further refined using *SHELXL*97 (Sheldrick, 2008) in conjugate-gradient least-squares (CGLS) mode against F^2 in the resolution range 20.0–1.4 Å.

After the first two rounds of refinement, the non-H atoms were refined anisotropically in all subsequent refinement cycles. During the first round of anisotropic refinement R_{free} dropped considerably (approximately 4%) compared with isotropic refinement. Each round of refinement consisted of ten cycles of CGLS refinement followed by inspection of the model and the $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps using *Coot*. Manual modelling of various aspects of the structure, such as the flipping of Asn and Gln residues, alternative conformations and the addition of solvent molecules, was carried out. Water molecules with *B* factors higher than 65 Å^2 were deleted from subsequent refinement rounds.

Default restraints and weights from *SHELX*97 were used for protein geometry. The restraints for L-Asp were created using the *PRODRG* server (http://davapc1.bioch.dundee.ac.uk/prodrg/) and were manually added to the command file of *SHELX*97.

Refinement of the room-temperature structure was initiated using the final coordinates of the 1.4 Å structure after removal of the anisotropic temperature factors. The structure was refined using *REFMAC*.

The final structures were assessed with *MolProbity* (Lovell *et al.*, 2003), *PROCHECK* (Laskowski *et al.*, 1993) and validation tools in *Coot*. Structure superposition was performed using *Secondary Structure Matching* (*SSM*; Krissinel &

Henrick, 2004). Graphic representations of the superimposed structures were visually inspected. Contacts were measured using the program *CONTACT* in *CCP*4. Interface analysis was performed using the EBI *PISA* server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html).

3. Results and discussion

3.1. Quality of the structure

The structure of HpA in the presence of L-Asp was refined to 1.4 Å resolution with anisotropic *B* factors for all non-H atoms. In addition, an HpA structure was refined to 1.8 Å resolution using data collected at room temperature in order to better assess the flexibility of surface regions. However, all discussion and results are based on the 1.4 Å resolution structure except where stated otherwise. The refinement process converged to an $R_{\rm cryst}$ of 12.6% for all 58 995 reflections. The last recorded $R_{\rm free}$ was 16.9% for 2939 reflections [15.8% for 2660 reflections with $F > 4\sigma(F)$]. The final model contained 2531 protein atoms, one L-Asp molecule and 391 water molecules. The first four residues at the N-terminus and residues 22 and 23 could not be modelled in the structure owing to a lack of sufficient density. Residues 21 and 24–45 are characterized by poor density and were refined with an occupancy of 0.65 to obtain reasonable *B* factors and better electron density. The r.m.s. deviations (r.m.s.d.s) in bond lengths and bond angles were 0.012 Å and 2.52°, respectively. The average *B* factors for protein atoms and water molecules were 20.5 and 36.1 Å^2, respectively. The coordinate error was 0.07 Å according to a Luzzati plot. A similar value (0.05 Å) was obtained using the diffraction-component precision indicator (Cruickshank, 1999).

The Ramachandran plot (φ, ψ) showed that 99.7% of non-Gly and non-Pro residues were in most favoured and additionally allowed regions and that no residues were in generously allowed regions. Although the overall stereochemical quality of the model was very good, Thr204 was identified as being in the disallowed region (φ = 37.9°, ψ = −108.4°) of the Ramachandran plot. Thr204 is highly conserved in all L-asparaginases and contributes to folding of the protein in the correct conformation (Swain *et al.*, 1993; Lubkowski, Wlodawer, Housset *et al.*, 1994). The electron density of Thr204 is of very high quality, leaving no doubt about its conformation. The participation of its main-chain atoms in several strong hydrogen bonds could explain its strained conformation (Lubkowski *et al.*, 2003). 16 residues (Gln53, Gln56, Gln75, Ser125, Ser140, Asn149, Glu165, Val167, Lys178, Glu208, Ser241, Lys262, Gln265, Lys307, Asn320 and Glu327) were modelled in alternative conformations. Of these residues, Gln53, Gln56, Gln75, Ser125, Asn149, Val167, Lys307 and Asn320, as well as Leu10, Ser182, Glu198 and Glu206, are found in alternative conformations in the 1.8 Å room-temperature structure. The statistics of data collection and structure refinement are summarized in Table 1.

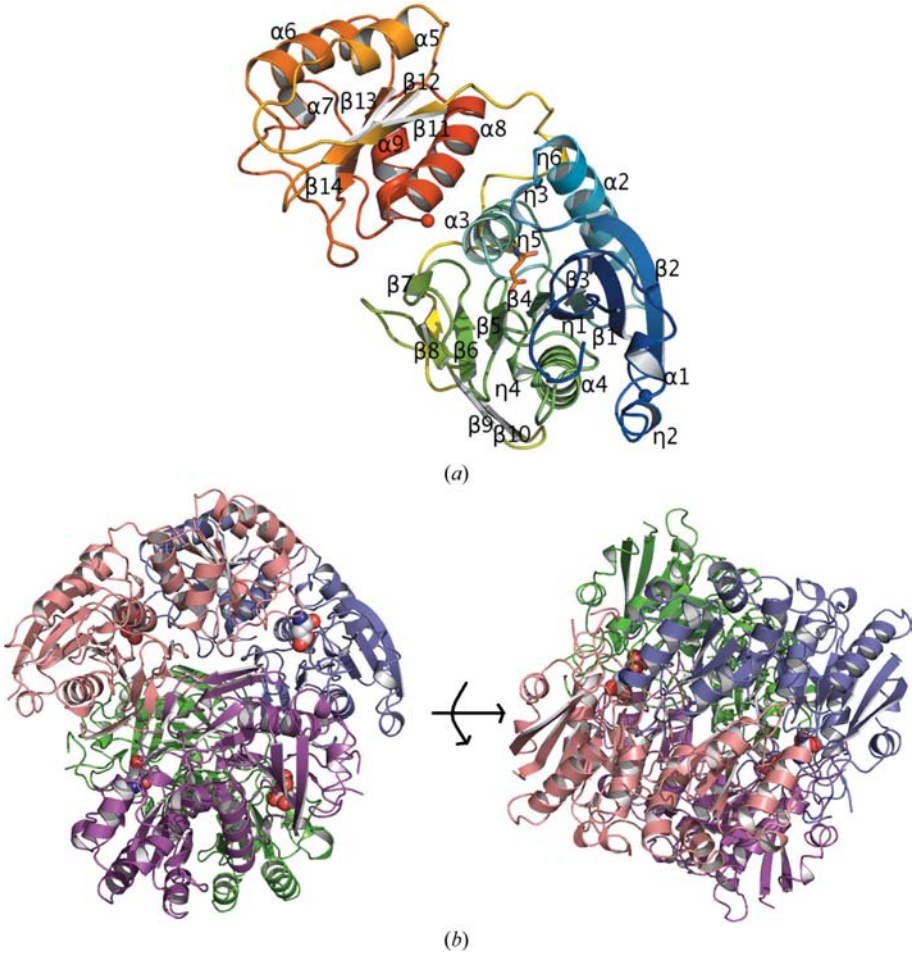

Figure 1
(*a*) Cartoon diagram of HpA. The colouring scheme is from blue (N-terminus) to red (C-terminus). The active-site location is indicated by the bound L-Asp (shown in stick representation). Secondary-structure elements were assigned with *DSSP* (Kabsch & Sander, 1983). (*b*) Cartoon representation of the HpA tetramer. Each monomer is shown in a different colour. The orientation of the blue-coloured monomer is the same as in (*a*). The blue–salmon and green–magenta pairs correspond to intimate dimers. L-Asp is depicted as a space-filling model. The same structure after a 90° rotation is shown on the right. This figure was created with *PyMOL* v.0.99 (DeLano Scientific, Palo Alto, California, USA).

3.2. Description of the structure

The structure of HpA is characterized by an α/β fold (27.9% α-helices, 5.5% 3_{10}-helices and 20.2% β-strands) similar to those of other L-asparaginases. The monomer consists of 327 amino acids that form nine α-helices, 14 β-strands

and six 3_{10}-helices and is divided into a large N-terminal domain and a smaller C-terminal domain connected by a linker of ~26 residues (Fig. 1). The overall dimensions of the molecule are approximately 48 × 58 × 62 Å. The accessible solvent area (ASA) is 13 880 Å^2, similar to those of other L-asparaginases. A distinguishing feature of all L-asparaginases is the presence of a rare left-handed crossover between strands $\beta4$ and $\beta5$ (residues 119–153 in HpA) that contributes to the activity of the enzyme (Miller *et al.*, 1993). The residues involved in the crossover form part of the active

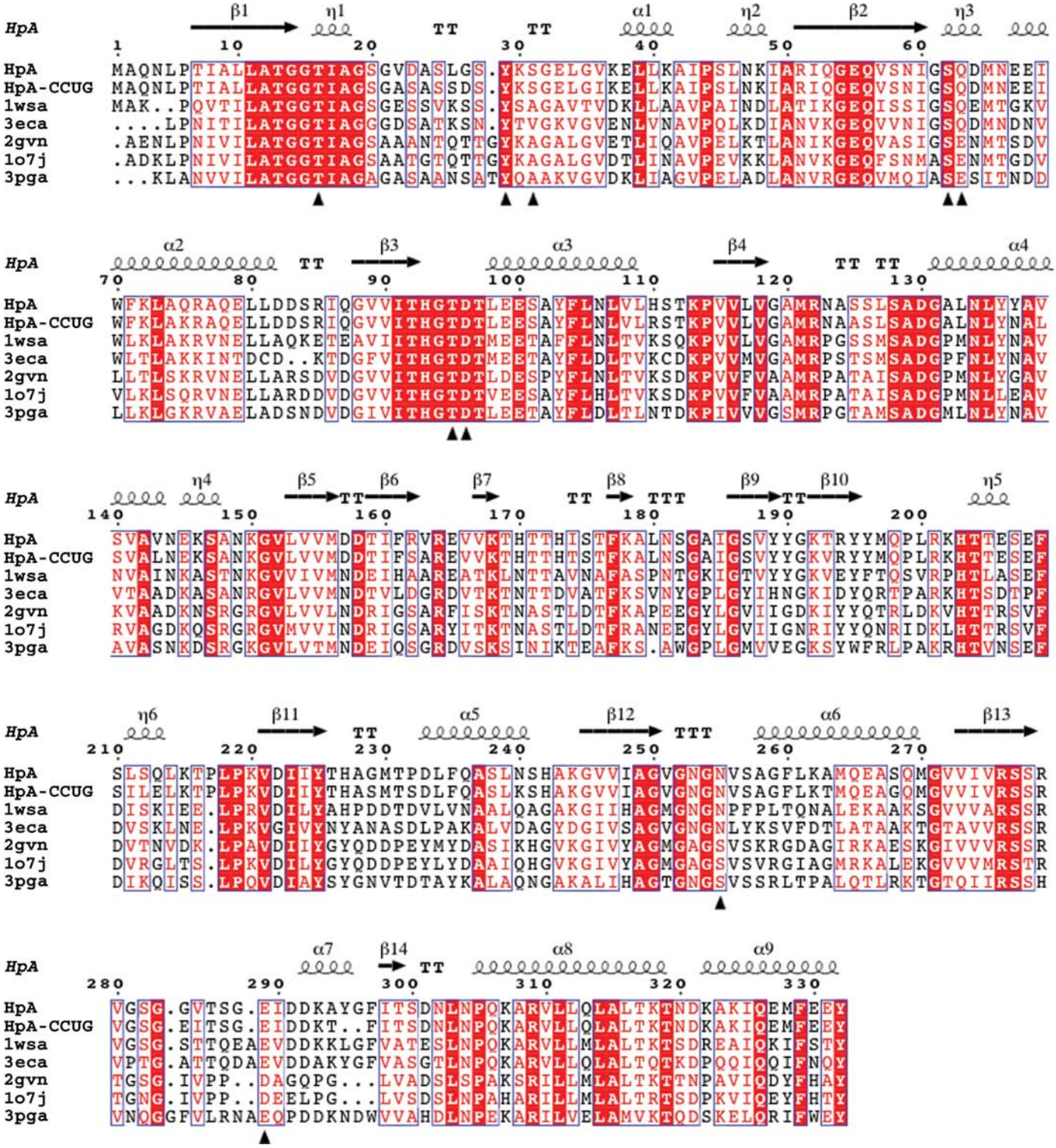

Figure 2
Sequence alignment of HpA, HpA from strain CCUG 17874 (HpA-CCUG), WsA (1wsa; NCBI code CAA61502), EwA (2gvn; AAS67027.1), EcAII (3eca; AAA23445.1), ErA (1o7j; CAA32884.1) and PGA (3pga). Active-site residues are marked with black arrows. This figure was drawn with *ESPript* (Gouet *et al.*, 1999). 3_{10}-Helices are indicated by η, strict β-turns by TT and strict α-turns by TTT.

Table 2
Comparison of interfaces in L-asparaginase structures.

Protein	Assembly	Interface area (Å^2)	No. of residues	No. of hydrogen bonds	No. of nonbonded contacts	No. of salt bridges
HpA	Tetramer	*A*/*C* = 2005/2006	64	32	307	6
EcAII	Tetramer	*A*/*C* = 2297/2272	61	42	339	7
ErA	Tetramer	*A*/*C* = 2267/2280	67	26	352	8
WsA	Dimer	*A*/*B* = 1991/1990	55	33	274	8
EwA	Dimer of tetramers	*A*/*E* = 2177/2171	65	43	403	11

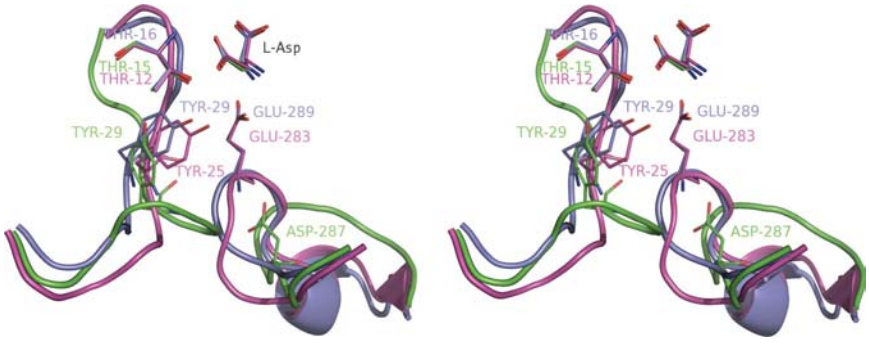

Figure 3
Close-up stereodiagram of the active-site flexible loop. Residues are shown in stick representation. HpA, EwA and EcAII are coloured blue, green and magenta, respectively. L-Asp is labelled.

site and are evolutionarily conserved. The formation of a hydrogen bond between the carbonyl O atom of Ala25 from the active-site flexible loop and the main-chain N atom of Leu127 from the crossover in HpA is characteristic of all L-asparaginases.

3.3. Comparison with other L-asparaginases

Structure superposition was carried out using the coordinates of HpA against the structures of WsA, EcAII, EwA, ErA and PgA. R.m.s.d.s of 0.98 Å with WsA (PDB code 1wsa; 53% sequence identity), 0.94 Å with EcAII (3eca; 51%), 1.13 with EwA (2gvn; 48%), 1.01 Å with ErA (1o7j; 47%) and 1.01 with PgA (3pga; 43%) in C^{α} positions were found, respectively, confirming the similar architecture to other type II L-asparaginases. However, significant variations that may contribute to different binding properties or stabilities are seen in flexible loops in all the enzymes, as explained below.

The active-site flexible loop in HpA consists of about 27 residues (19–46). This loop exhibits conformational variability in all L-asparaginases despite the presence of several strictly conserved residues (Tyr29, Val36, Leu39 and Pro44 in HpA; Fig. 2). Some degree of stabilization of the flexible loop is observed in L-asparaginases upon substrate binding. The orientation of the flexible loop upon L-Asp binding is similar in HpA and EcAII with only minor differences, suggesting an analogous mechanism of accessibility to the binding site in these two enzymes (Fig. 3). On the other hand, residues 25–31 of the flexible loop of EwA have a considerably different position from the equivalent region in HpA and EcAII.

Apart from the flexible loop, the parts of the L-asparaginases that deviate most include residues 210–218 and the loop 286–297 close to the active site. The former exists in a similar conformation in EcAII, ErA, EwA and WsA but is different in PgA and HpA. The 286–297 loop is also seen in two conformations, one found in both ErA and EwA and the other in HpA, WsA, EcAII and PgA.

It has previously been reported that cryogenic structures can show subtle but significant changes compared with room-temperature structures, especially at the surface of proteins (Dunlop *et al.*, 2005). Accordingly, a room-temperature data set was collected in order to further investigate the conformational variability of the active-site loop. However, inspection of the electron-density maps showed that most of the loop could not be located. Electron density was only found for the C-terminal end of the loop after residue 35. The position of this part of the loop was the same in both structures. These results confirmed the high flexibility of the active-site loop and ruled out the possibility that it adopts a stable conformation at room temperature. Furthermore, the r.m.s.d.s in main-chain and side-chain atom positions between the two structures were found to be 0.5 and 0.7 Å, respectively, suggesting no major changes upon freezing.

3.4. Description of the HpA tetramer

The HpA tetramer was generated by the symmetry operators $(-x, y, -z)$, $(-x, -y, z)$, $(x, -y, -z)$ of space group *I*222 (Fig. 1*b*). The arrangement of the monomers is the same as in other L-asparaginase tetramers and obeys 222 symmetry as anticipated. The intimate dimers are defined by monomers *A* and *C*, and by monomers *B* and *D*. Each monomer contributes 64 residues at the interface that account for 19.6% of the total number of residues. The interface residues are classified as 40.7% polar, 37.3% nonpolar and 22.0% charged. About 2005 Å^2 of ASA corresponding to 14.6% of the total ASA of the monomer are buried in the interface (Table 2). 32 hydrogen bonds and six salt bridges are seen at the dimer interface of HpA. Three of the 32 hydrogen bonds are formed by strictly conserved amino acids (Asp96, Glu100 and Tyr225); of these, Asp96 also contributes to the binding of the substrate. Residues that are partly conserved in HpA, WsA and EcAII and are found at the interface are Gln63, Thr171 and Gln306, where Gln63 is an active-site residue.

3.5. Active site

The active site of the enzyme is located between the N-terminal and C-terminal domains of adjacent subunits in the

Table 3
Enzyme–L-Asp interactions at the active site (cutoff distance of 3.5 Å).

L-Asp atom	HpA	EwA	EcAII
OXT	Ser31*A* OG (2.8), Ser62*A* N (2.8), Gly61*A* CA (3.3), Gly15*A* CA (3.2), Gly94*A* CA (3.2)	Ser62*A* N (2.8), Wat824 O (3.0), Glu63*A* OE1 (3.4), Gly14*A* CA (3.3)	Ser58*A* OG (2.5), Thr89*A* N (3.2), Asp90*A* N (3.0), Asp90*A* CG (3.2), Asp90*A* CB (3.0), Gly88*A* CA (3.3), Gly88*A* C (3.4)
C	Ser62*A* N (3.5), Gly94*A* CA (3.4)	Ser62*A* OG (3.3), Glu63*A* OE1 (3.3)	Ser58*A* OG (3.4), Ser58*A* N (3.4), Gly88*A* CA (3.4)
O	Ser62*A* OG (2.6), Thr95*A* N (3.2), Asp96*A* N (2.9), Asp96*A* CG (3.2), Asp96*A* OD2 (3.4), Ser62*A* CB (3.5), Asp96*A* CB (3.2), Gly94*A* CA (3.3), Gly94*A* C (3.3)	Thr95*A* N (3.3), Asp96*A* N (3.1), Asp96*A* CB (3.3), Asp96*A* CG (3.3), Asp96*A* OD2 (3.3), Ser62*A* OG (2.3), Ser62*A* CB (3.4), Gly94*A* CA (3.4)	Ser58*A* N (2.7), Val27*A* CG2 (2.9), Gly88*A* CA (3.2), Gly11*A* CA (3.3), Gly57*A* CA (3.3), Gly57*A* C (3.4)
CA	Thr16*A* OG1 (3.3)	Asp96*A* OD2 (3.5), Thr15*A* OG1 (3.3)	Thr12*A* OG1 (3.3), Glu283*C* OE2 (3.3), Val27*A* CG2 (3.4)
N	Gln63*A* OE1 (3.1), Asp96*A* OD2 (2.8), Glu289*C* OE1 (3.1), Glu289*C* OE2 (2.7), Glu289*C* CD (3.2), Asn255*C* ND2 (3.4), Asp96*A* CG (3.4)	Asp96*A* OD2 (2.8), Glu63*A* OE1 (3.0), Wat2172 O (2.9), Wat2269 O (3.3)	Gln59*A* OE1 (3.0), Asp90*A* OD2 (3.0), Asn248*C* ND2 (3.5), Glu283*C* OE2 (2.5), Glu283*C* CD (3.1), Glu283*C* OE1 (3.1)
CB	Thr16*A* OG1 (3.1), Asp96*A* OD1 (3.3)	Thr15*A* OG1 (3.4), Asp96*A* OD1 (3.4), Asp96*A* OD2 (3.4)	Thr12*A* OG1 (3.1), Asp90*A* OD1 (3.4)
CG	Thr16*A* OG1 (2.8), Thr16*A* CB (3.3), Thr95*A* OG1 (2.9), Thr95*A* N (3.4)	Thr95*A* OG1 (3.0), Thr95*A* N (3.4), Thr15*A* OG1 (3.1), Thr15*A* CB (3.4)	Thr12*A* OG1 (2.9), Thr12*A* CB (3.4), Thr89*A* N (3.1), Thr89*A* OG1 (2.9)
OD2	Thr16*A* OG1 (3.1), Thr16*A* CB (3.3), Thr95*A* OG1 (2.6), Ala120*A* O (3.1), Wat100 O (2.8)	Thr95*A* OG1 (2.7), Ala120*A* O (3.1), Thr95*A* CG2 (3.4), Thr15*A* CB (3.5), Thr15*A* OG1 (3.4), Wat307 O (2.6), Gly61*A* CA (3.3)	Thr89*A* OG1 (2.4), Thr89*A* CB (3.4), Thr12*A* OG1 (3.3), Thr12*A* CB (3.4), Ala114*A* O (3.1), Wat206 O (2.8)
OD1	Thr16*A* OG1 (3.1), Thr16*A* CB (3.2), Thr16*A* N (3.0), Thr95*A* N (2.7), Thr95*A* OG1 (3.3), Wat37 O (3.0), Gly94*A* CA (3.2), Gly94*A* C (3.5)	Thr95*A* N (3.0), Thr15*A* N (3.1), Thr15*A* OG1 (3.3), Wat15 O (3.0), Gly94*A* CA (3.3), Thr15*A* CB (3.3)	Thr12*A* OG1 (3.2), Thr12*A* N (3.2), Thr12*A* CB (3.5), Thr89*A* OG1 (3.4), Thr89*A* N (2.6), Gly88*A* CA (3.0), Gly88*A* C (3.2)

intimate dimers (*i.e. AC* and *BD*). Strong electron density corresponding to bound L-Asp (up to 8σ and 4σ in the $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ electron-density maps, respectively) was observed in the active site of HpA. The similar *B* factors for the active-site residues and L-Asp (21.4 and 21.0 Å^2, respectively) suggest stable binding of the latter. The active-site architecture is shown in Fig. 4. L-Asp is stabilized by strong hydrogen bonds (<3.1 Å) to Thr16, Ser31, Ser62, Gln63, Thr95, Asp96 and Glu289 from the second subunit of the dimer (Table 3). Ser31, in particular, is seen only in HpA and corresponds to Val in EcAII and to Ala in WsA, EwA and ErA. Thus, no analogous hydrogen bond is formed in these L-asparaginases as no suitable side chain is present. Amongst the rest of the residues involved in substrate binding, Ser62 and Thr95 are conserved in all L-asparaginases, whereas Gln63 is conserved in WsA and EcAII but not in EwA, ErA and PgA, where it is substituted by a Glu residue.

Tyr29 from the flexible part of the active site is close to Thr16, with a distance of 2.5 Å between Tyr29 OH and Thr16 OG1; this distance is similar to that observed for Tyr25 in EcAII (2.9 Å) but significantly different from that for Tyr29 in EwA (3.8 Å; Fig. 3). This is mainly a consequence of the different orientation of the Tyr29 side chain in EwA. Consequently, the flexible loop adopts a more open conformation in HpA and EcAII compared with that in EwA. Notably, the flexible loop in HpA could be partly built owing to the appearance of electron density for the aromatic ring of Tyr29, indicating stabilization of this residue in the presence of the substrate compared with the rest of the loop. This is in agreement with previous reports that Tyr25 is essential for loop closure and catalysis in EcAII (Derst *et al.*, 1994; Aung *et al.*, 2000).

The two catalytic Thr residues in L-asparaginases are expected to be in a hydrophobic environment that serves to restrict their rotation (Miller *et al.*, 1993). Indeed, Met121 provides a hydrophobic environment to Thr95 as in other family members. In contrast, Thr16 in HpA is flanked by Asn123 instead of the Pro found in all other L-asparaginase structures. This change is noteworthy and might have some implications in the catalytic role of Thr16 as a result of the reduced hydrophobicity.

Asn255 from the rigid side of the active site is also involved in key interactions around the active site. In subunit *C*, its ND2 atom is located close to L-Asp N (3.4 Å) from subunit *A* and participates in four additional hydrogen bonds with Glu289 OE2 (2.9 Å) and Wat76 (2.9 Å) from subunit *C* and Asp96 OD2 (2.9 Å) and Wat98 (2.6 Å) from subunit *A*. Asn255 is conserved in EcAII and WsA, but is replaced by Ser in ErA, EwA and PgA. The structural equivalent Ser254 in EwA has its OG1 atom oriented away from L-Asp N at a distance of 4.1 Å and thus is unable to make contacts with L-Asp.

Glu289 may also play a key role in the catalytic properties of HpA. This residue belongs to the 286–297 loop that shows significant variations in L-asparaginases (Fig. 5). The shift in C^{α} positions between Glu289 in HpA and the corresponding Glu283 in EcAII is only 0.7 Å, suggesting a similarity between the two L-asparaginases upon substrate binding, with Glu289 in HpA closer to the active site than Glu283 in EcAII. The structurally equivalent residue in ErA and EwA is Asp287,

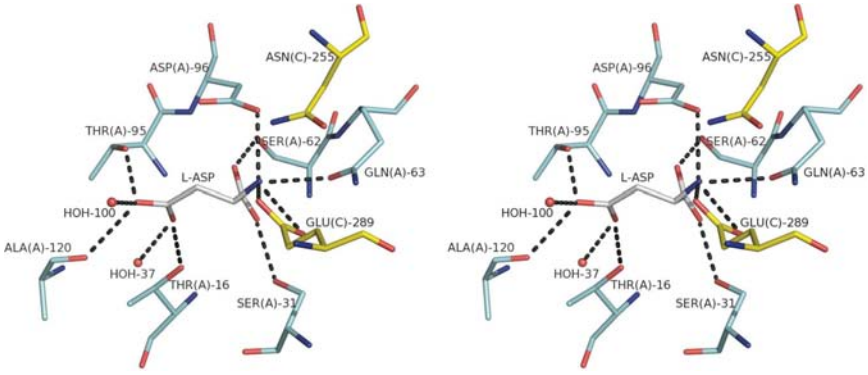

Figure 4
Stereoview of the active site of HpA with bound L-Asp (in grey). Residues from the second subunit of the intimate dimer are shown in yellow. The active-site residues involved in interactions with the substrate and major hydrogen bonds are depicted (a full list of contacts is given in Table 3).

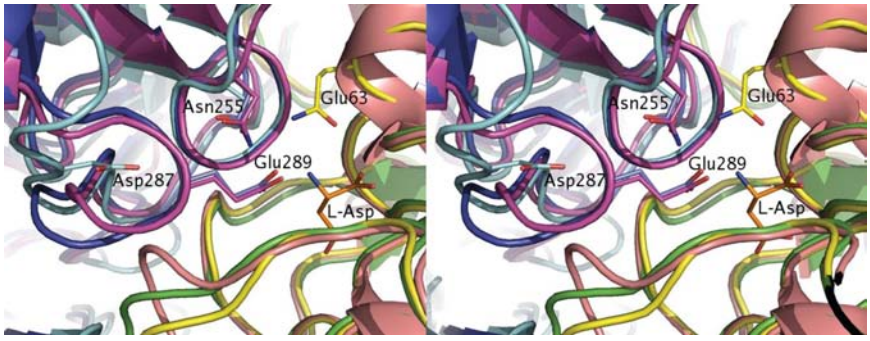

Figure 5
Close-up stereoview of the differences around Glu289 in HpA. HpA subunits *A* and *C* are coloured yellow and magenta, respectively. EcAII is shown in green (subunit *A*) and blue (subunit *C*); ErA is shown in salmon (subunit *A*) and cyan (subunit *C*). Residues are shown in stick representation according to the C^{α} colouring of the structures. Residues Glu289, Asn255 and Gln63 of HpA are labelled.

which is found ∼8.5 Å away from the active site in both enzymes. Thus, this region shows similarities between EwA and ErA, suggesting that substrate binding may not significantly affect the position of the loop in EwA. Possible reasons for this could be attributed to the presence of Ser in place of Asn255 in both ErA and EwA, which leads to elimination of the hydrogen bond present between the glutamate residue and asparagine in HpA and EcAII. Furthermore, Glu289 of HpA interacts with the side chain of the catalytically important Gln63 (Glu289 OE1–Gln63 NE2, 2.9 Å). Therefore, the presence of three residues (Glu289, Asn255 and Gln63) around the active site may contribute to the kinetic properties of HpA and help in excluding larger amino acids, such as L-glutamate, from entering. A similar cluster of residues and interactions is found in EcAII. However, the reduced glutaminase activity of EwA cannot be explained by the lack of this cluster. Other parts of the active site may therefore be responsible for the catalytic properties of EwA.

3.6. Role of water molecules

As in other L-asparaginases, water molecules also play an important role in ligand binding in HpA. Waters found in the vicinity of the active site of HpA include Wat100, Wat74, Wat76 and Wat37. These waters are characterized by low temperature factors (Wat100, 20.2 $Å^2$; Wat74, 14.8 $Å^2$; Wat76, 14.6 $Å^2$; Wat37, 16.7 $Å^2$) that increase the probability of their role in stabilizing the ligand. Wat100, in particular, is involved in a hydrogen bond to OD2 of L-Asp (2.8 Å) and interacts indirectly with Tyr29 OH through another water molecule (Wat265). All four waters are also found in the atomic resolution structure of ErA. The equivalents of HpA Wat100, Wat74, Wat76 and Wat37 in ErA are Wat142, Wat198, Wat115 and Wat110, respectively. Two additional waters (Wat112 and Wat116) found in ErA are displaced in HpA by the L-Asp OD2 and N atoms, respectively.

3.7. Comparison between *H. pylori* L-asparaginases

HpA has 92% sequence identity with the L-asparaginase from *H. pylori* strain CCUG 17874. The residues involved in contacts with L-Asp are all conserved in the two enzymes, in agreement with their similarity in kinetic properties. Only four residues in the flexible loop (Val22Ala, Asp23Ser, Leu26Ser and Gly27Asp) are found to differ. HpA from strain CCUG 17874 has also been reported to form tetramers with a molecular weight of 140 kDa. Inspection of the residues found at the interface revealed only three minor differences: Gly229Ser, Gly284Glu and Val285Ile. The two *H. pylori* L-asparaginases are, therefore, expected to share similar properties at the structural and functional levels.

4. Conclusions

The crystal structure of HpA in the presence of L-Asp was determined to 1.4 Å resolution, one of the highest resolutions obtained for an L-asparaginase. Anisotropic refinement of non-H atoms resulted in a final structure with good crystallographic statistics and stereochemistry. Despite the high quality of the structure and the binding of an L-Asp molecule in the active site, part of the flexible loop involved in substrate binding could not be seen. However, Tyr29 from the active-site flexible loop was well defined, suggesting that its position may reflect initial steps of the loop-stabilization mechanism upon substrate binding. Subtle but probably structurally significant differences that could modulate the activity of the enzyme were found in the active site. However, the most striking differences were found in Asn255 and Glu289. The former is replaced by Ser in ErA and EwA as well as in PgA. The presence of Glu289 at a similar location to Glu283 in EcAII and close to L-Asp suggests a more restricted active site in

HpA and EcAII compared with those of EwA and ErA. Consequently, the lack of key interactions in EwA and ErA might be the reason for the open conformation adopted by their respective loops. The structure of HpA presented here could form the basis for further studies to improve the current properties of the enzyme in order to obtain a better therapeutic for leukaemia treatment.

This work was supported by the Sigrid Jusélius Foundation and the Academy of Finland (Grant No. 121278). We thank Julya Krasotkina for providing us with pure preparations of the enzyme and the staff at EMBL Hamburg (c/o DESY) for assistance with data collection. Access to EMBL Hamburg (c/o DESY) was provided by the European Community through the Research Infrastructure Action under FP6 (contract No. RII3/CT/2004/5060008 to EMBL Hamburg Outstation).

References

Aghaiypour, K., Wlodawer, A. & Lubkowski, J. (2001). *Biochemistry*, **40**, 5655–5664.
Aung, H. P., Bocola, M., Schleper, S. & Rohm, K. H. (2000). *Biochim. Biophys. Acta*, **1481**, 349–359.
Avramis, V. & Panosyan, E. (2005). *Clin. Pharmacokinet.* **44**, 367–393.
Bridges, J. (1968). *Lancet*, **2**, 1299.
Campbell, H. A., Mashburn, L. T., Boyse, E. A. & Old, L. J. (1967). *Biochemistry*, **6**, 721–730.
Cappelletti, D., Chiarelli, L. R., Pasquetto, M. V., Stivala, S., Valentini, G. & Scotti, C. (2008). *Biochem. Biophys. Res. Commun.* **377**, 1222–1226.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
Derst, C., Henseling, J. & Rohm, K. H. (2000). *Protein Sci.* **9**, 2009–2017.
Derst, C., Wehner, A., Specht, V. & Röhm, K. H. (1994). *Eur. J. Biochem.* **224**, 533–540.
Dhavala, P., Krasotkina, J., Dubreuil, C. & Papageorgiou, A. C. (2008). *Acta Cryst.* F**64**, 740–742.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Dunlop, K. V., Irvin, R. T. & Hazes, B. (2005). *Acta Cryst.* D**61**, 80–87.
Duval, M., Suciu, S., Ferster, A., Rialland, X., Nelken, B., Lutz, P., Benoit, Y., Robert, A., Manel, A. M., Vilmer, E., Otten, J. & Philippe, N. (2002). *Blood*, **99**, 2734–2739.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gladilina, I. A., Sokolov, N. N. & Krasotkina, I. V. (2008). *Biomed. Khim.* **54**, 482–486.
Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
Gökbuget, N. & Hoelzer, D. (2006). *Hematology. American Society of Hematology Education Program Book*, pp. 133–141.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Krasotkina, J., Borisova, A. A., Gervaziev, Y. V. & Sokolov, N. N. (2004). *Biotechnol. Appl. Biochem.* **39**, 215–221.
Kravchenko, O. V., Kislitsin, Y. A., Popov, A. N., Nikonov, S. V. & Kuranova, I. P. (2008). *Acta Cryst.* D**64**, 248–256.
Krissinel, E. & Henrick, K. (2004). *Acta Cryst.* D**60**, 2256–2268.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Lubkowski, J., Dauter, M., Aghaiypour, K., Wlodawer, A. & Dauter, Z. (2003). *Acta Cryst.* D**59**, 84–92.
Lubkowski, J., Palm, G. J., Gilliland, G. L., Derst, C., Röhm, K. H. & Wlodawer, A. (1996). *Eur. J. Biochem.* **241**, 201–207.
Lubkowski, J., Wlodawer, A., Ammon, H. L., Copeland, T. D. & Swain, A. L. (1994). *Biochemistry*, **33**, 10257–10265.
Lubkowski, J., Wlodawer, A., Housset, D., Weber, I. T., Ammon, H. L., Murphy, K. C. & Swain, A. L. (1994). *Acta Cryst.* D**50**, 826–832.
Mashburn, L. T. & Wriston, J. C. Jr (1964). *Arch. Biochem. Biophys.* **105**, 450–452.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Miller, M., Rao, J., Wlodawer, A. & Gribskov, M. (1993). *FEBS Lett.* **328**, 275–279.
Muller, H. J. & Boos, J. (1998). *Crit. Rev. Oncol. Hematol.* **28**, 97–113.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Ollenschläger, G., Roth, E., Linkesch, W., Jansen, S., Simmel, A. & Mödder, B. (1988). *Eur. J. Clin. Invest.* **18**, 512–516.
Ortlund, E., Lacount, M. W., Lewinski, K. & Lebioda, L. (2000). *Biochemistry*, **39**, 1199–1204.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Papageorgiou, A. C., Posypanova, G. A., Andersson, C. S., Sokolov, N. N. & Krasotkina, J. (2008). *FEBS J.* **275**, 4306–4316.
Richards, N. & Kilberg, M. (2006). *Annu. Rev. Biochem.* **75**, 629–654.
Sanches, M., Krauchenco, S. & Polikarpov, I. (2007). *Curr. Chem. Biol.* **1**, 75–86.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Stams, W., den Boer, M., Beverloo, H., van Wering, E. & Pieters, R. (2005). *Leukemia*, **19**, 318–319.
Swain, A. L., Jaskólski, M., Housset, D., Rao, J. K. & Wlodawer, A. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 1474–1478.
Verma, N., Kumar, K., Kaur, G. & Anand, S. (2007). *Crit. Rev. Biotechnol.* **27**, 45–62.