STRUCTURAL BIOLOGY

ISSN: 2059-7983
journals.iucr.org/d

Structure and function of the thermostable L-asparaginase from *Thermococcus kodakarensis*

Jingxu Guo, Alun R. Coker, Steve P. Wood, Jonathan B. Cooper, Shahid Mahmood Chohan, Naeem Rashid and Muhummad Akhtar

Acta Cryst. (2017). D**73**, 889–895

STRUCTURAL BIOLOGY

ISSN 2059-7983

Structure and function of the thermostable L-asparaginase from *Thermococcus kodakarensis*

Jingxu Guo,[a] Alun R. Coker,[a] Steve P. Wood,[a] Jonathan B. Cooper,[a,b]* Shahid Mahmood Chohan,[c] Naeem Rashid[c] and Muhummad Akhtar[c]

[a]Division of Medicine, University College London, Gower Street, London WC1E 6BT, England, [b]Department of Biological Sciences, Birkbeck, University of London, Malet Street, Bloomsbury, London WC1E 7HX, England, and [c]School of Biological Sciences, University of the Punjab, Quaid-e-Azam Campus, Lahore 54590, Pakistan. *Correspondence e-mail: jon.cooper@ucl.ac.uk

Received 26 August 2017
Accepted 11 October 2017

Edited by P. Langan, Oak Ridge National Laboratory, USA

Keywords: protein crystallography; structural biology; L-asparaginase; *Thermococcus kodakarensis*.

PDB reference: L-asparaginase from *Thermococcus kodakarensis*, 5ot0

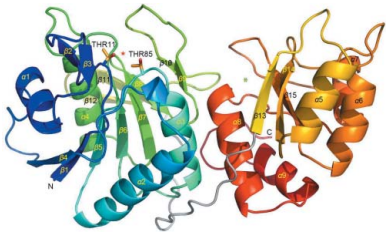
Supporting information: this article has supporting information at journals.iucr.org/d

L-Asparaginases catalyse the hydrolysis of asparagine to aspartic acid and ammonia. In addition, L-asparaginase is involved in the biosynthesis of amino acids such as lysine, methionine and threonine. These enzymes have been used as chemotherapeutic agents for the treatment of acute lymphoblastic leukaemia and other haematopoietic malignancies since the tumour cells cannot synthesize sufficient L-asparagine and are thus killed by deprivation of this amino acid. L-Asparaginases are also used in the food industry and have potential in the development of biosensors, for example for asparagine levels in leukaemia. The thermostable type I L-asparaginase from *Thermococcus kodakarensis* (TkA) is composed of 328 amino acids and forms homodimers in solution, with the highest catalytic activity being observed at pH 9.5 and 85°C. It has a K_m value of 5.5 m*M* for L-asparagine, with no glutaminase activity being observed. The crystal structure of TkA has been determined at 2.18 Å resolution, confirming the presence of two α/β domains connected by a short linker region. The N-terminal domain contains a highly flexible β-hairpin which adopts 'open' and 'closed' conformations in different subunits of the solved TkA structure. In previously solved L-asparaginase structures this β-hairpin was only visible when in the 'closed' conformation, whilst it is characterized with good electron density in all of the subunits of the TkA structure. A phosphate anion resides at the active site, which is formed by residues from both of the neighbouring monomers in the dimer. The high thermostability of TkA is attributed to the high arginine and salt-bridge content when compared with related mesophilic enzymes.

1. Introduction

L-Asparaginase (EC 3.5.1.1) catalyses the hydrolysis of asparagine to aspartic acid and ammonia and has a range of biological roles. For instance, plants transport nitrogen in the form of L-asparagine from their roots to growing tissues and thus have a high demand for this enzyme (Sieciechowicz *et al.*, 1988; Atkins *et al.*, 1975). In bacteria, when amino acids become the primary carbon source in anaerobic conditions, the expression level of asparaginase can be increased by 100-fold (Cedar & Schwartz, 1967, 1968). This is important since the metabolites of asparagine (and glutamine) can feed into the citric acid cycle. In contrast, the preferred carbon source, glucose, is a catabolite repressor of asparaginase expression. Thus, asparaginases and glutaminases are necessary for cell growth in ammonia-deficient media and their expression is activated by the presence of these amino acids in the medium.

Many asparaginases also have activity towards glutamine, producing glutamic acid, and can be classified accordingly. Type I enzymes are cytosolic and have a low glutaminase activity of around 2–10% of the L-asparagine hydrolysis



activity. In contrast, type II enzymes are usually periplasmic and have comparable L-asparaginase and L-glutaminase activities (Boyd & Phillips, 1971; Chohan & Rashid, 2013; Davidson *et al.*, 1977). Type I and type II L-asparaginases have low sequence similarity; for example, those from *Escherichia coli* have a sequence identity of only 24%. Generally, the enzymes form dimers or tetramers (dimers of dimers) with a subunit molecular mass of 35 kDa.

In humans, L-asparaginase has been widely used as a chemotherapeutic agent for the treatment of acute lymphoblastic leukaemia (ALL) and other haematopoietic malignancies. ALL is the most common childhood acute leukaemia, constituting approximately 80% of childhood leukaemias and 20% of adult leukaemias (Fullmer *et al.*, 2010). The history of asparaginase usage for the treatment of ALL can be traced back to the 1950s, when Kidd spotted that the progression of murine lymphoma was reduced by guinea-pig serum (Kidd, 1953). This discovery attracted broad interest and it was found that only guinea-pig serum had anti-lymphoma activity, in contrast to the sera of other animals such as horse or rabbit. In the 1960s, Broome identified that it was the L-asparaginase in guinea-pig serum which contributed mainly to the anti-lymphoma activity (Broome, 1961, 1963). Treatment with L-asparaginase has been shown to have improved event-free survival for ALL from <10% to >80% in recent years (Silverman *et al.*, 2001; Pui *et al.*, 2009; Möricke *et al.*, 2008).

Cancer cells, such as lymphatic cells, have a high demand for asparagine for their survival and proliferation (Kiriyama *et al.*, 1989; Stams *et al.*, 2003). However, owing to the lack of L-asparagine synthetase required for L-asparagine synthesis, leukaemic lymphoblasts and some other tumour cells can only obtain this amino acid from blood serum. L-Asparaginase hydrolyses asparagine in the serum, leading tumours to a state of cell death (apoptosis), while healthy cells are not affected because they possess sufficient L-asparagine synthetase. In addition, studies have shown that L-asparaginase inhibits the mTOR pathway and induces an autophagic process which contributes to its anti-leukaemic activity and greatly affects leukaemia cells (Song *et al.*, 2015). Unlike conventional cancer therapy, L-asparaginase treatment is highly discriminatory.

Only the type II L-asparaginases from *E. coli* (EcAII) and *Erwinia chrysanthemi* (ErAII) have been approved for the treatment of ALL. The *E. coli* enzyme shows a higher activity, whilst the *E. chrysanthemi* enzyme has been used to treat patients who are allergic to the former (Albertsen *et al.*, 2001). In the USA, the most commonly used form of L-asparaginase is a covalent conjugation with polyethylene glycol, since PEGylation improves the bioavailability and biostability and reduces the immunological response. The elimination half-life of PEG-asparaginase (6 d) is five times longer than native EcAII preparations and nine times longer than ErAII preparations. This is an important improvement since the enzyme shows peak activity in the fifth day after intramuscular injection (Shrivastava *et al.*, 2016). In Europe, it is currently a second-line treatment only for patients who are allergic to native asparaginases. Side effects of L-asparaginase therapy such as immune responses, allergies and anaphylactic shock (Soares *et al.*, 2002) may be attributed to several reasons, including its L-glutaminase activity, which reduces the plasma L-glutamine level (Avramis *et al.*, 2002; Villa *et al.*, 1986). Thus, looking for alternative sources of L-asparaginase with fewer or no side effects is of great importance.

The enzyme is also widely applied in the food industry since treatment with L-asparaginase prior to cooking significantly reduces the formation of acrylamide. Acrylamide, also known as 2-propenamide, is a colourless and odourless crystalline solid that has potent neurotoxicity. It is largely produced from heat-induced reactions, *e.g.* frying or baking starchy foods at over 120°C, and is formed between the δ-amino group of asparagine and the carbonyl group of reducing sugars such as glucose (Friedman, 2003).

L-Asparaginase has been used to develop biosensors for the analysis of asparagine levels in leukaemia and in the food industry. Hydrolysis of asparagine produces ammonium ions, which induce a pH change that can be monitored spectrophotometrically using a suitable dye (Kumar *et al.*, 2013).

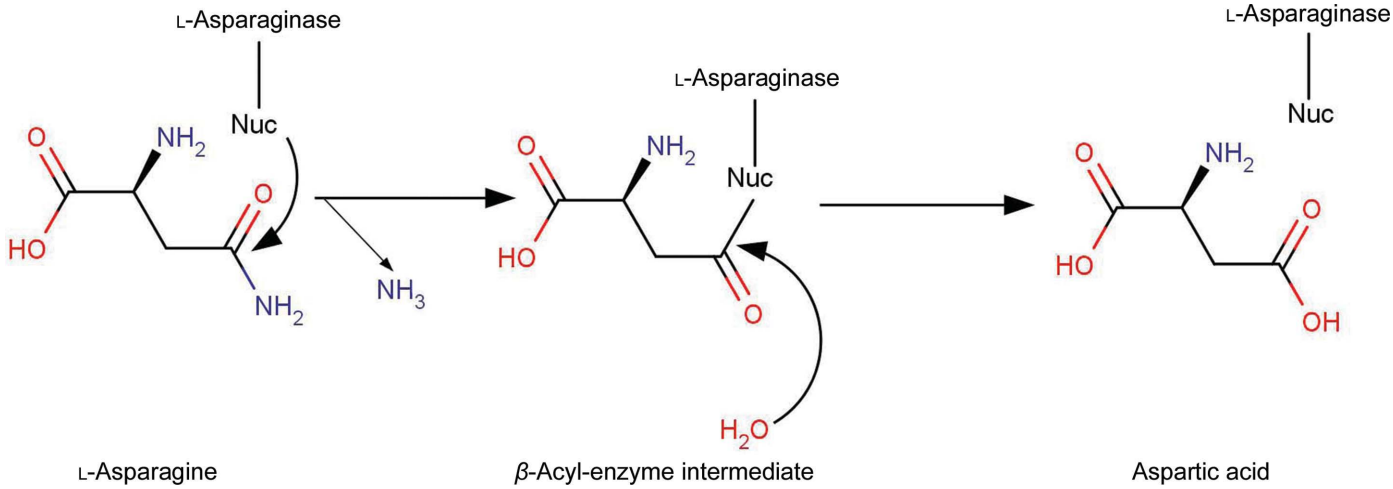


Figure 1
The reaction catalysed by L-asparaginase. The reaction involves nucleophilic attack on the asparagine side chain by one of the two active-site threonine residues, generating a β-acyl enzyme intermediate which is then hydrolysed to release the product.

Table 1
X-ray statistics for the TkA structure.

Values in parentheses are for the outer resolution shell.

Beamline	I03, DLS
Wavelength (Å)	0.9763
Space group	*P*1
Unit-cell parameters	
a (Å)	70.7
b (Å)	71.0
c (Å)	107.7
α (°)	72.1
β (°)	76.2
γ (°)	87.8
Resolution (Å)	68.65–2.18 (2.26–2.18)
R_{merge} (%)	6.3 (121.9)
R_{meas} (%)	7.5 (143.4)
$CC_{1/2}$ (%)	99.8 (58.8)
Completeness (%)	97.4 (97.2)
Average $I/\sigma(I)$	10.1 (1.3)
Multiplicity	3.5 (3.6)
No. of observed reflections	342457 (35236)
No. of unique reflections	98478 (9841)
Wilson plot *B* factor (Å^2)	50.0
Solvent content (%)	47.3
R factor (%)	19.8
R_{free} (%)	22.6
R.m.s.d., bond lengths (Å)	0.017
R.m.s.d., bond angles (°)	2.01
No. of reflections in working set	98,475
No. of reflections in test set	4,894
Mean protein *B* factor (Å^2)	47.7
PDB identifier	5ot0

The mechanism of L-asparaginase catalysis involves an intermediate known as a β-acyl-enzyme (Fig. 1), which is formed by the action of one of the catalytic threonine residues (Thr11 and Thr85 in TkA) nucleophilically attacking the side-chain amide C atom of the substrate L-asparagine. This is followed by nucleophilic attack on the intermediate by an enzyme-bound water molecule, which releases the product L-aspartate (Verma *et al.*, 2007).

The type I thermostable L-asparaginase from *Thermococcus kodakarensis* (TkA) is composed of 328 amino acids with a molecular mass of 35.5 kDa (Chohan & Rashid, 2013). Gel filtration demonstrated that the enzyme is active as a homodimer in solution, with the highest activity being observed at pH 9.5 and 85°C. It has a K_m value of 5.5 m*M* for L-asparagine, while no glutaminase activity was observed. In addition, TkA exhibited D-asparaginase activity, which was about 50% of the L-asparaginase activity. Here, we report the crystal structure analysis of TkA, which has been performed at a resolution of 2.18 Å.

2. Methods

Recombinant TkA was expressed and purified according to the method described by Chohan & Rashid (2013). In essence, the TK1656 gene was amplified by PCR and cloned into the pET-21a vector for expression in *E. coli* BL21-CodonPlus(DE3)-RIL cells. Purification involved heat treatment at 80°C for 20 min followed by the use of a Resource Q column (GE Healthcare) running a 0–1 *M* NaCl gradient. The purified enzyme was stored in 20 m*M* Tris–HCl pH 8.0. Screening for crystallization conditions was conducted using the sitting-drop method with a Mosquito robot. Crystals were obtained at a protein concentration of 10 mg ml^{-1} in MORPHEUS condition C6; for details, see Gorrec (2009). Optimum conditions were later found to be 0.03 *M* of each of $NaNO_3$, Na_2HPO_4 and $(NH_4)_2SO_4$, 0.1 *M* sodium HEPES/MOPS pH 7.5, 20% ethylene glycol, 10% PEG 8000 at 21°C. Selected crystals were transferred to a 10 µl drop containing 50% Paratone N MD2-08 and 50% paraffin oil to remove the mother liquor surrounding the crystals before flash-cooling and data collection.

X-ray data collection was carried out remotely at station I03 of Diamond Light Source (DLS) at 100 K using a PILATUS3 6M detector. Automatic data processing using *xia*2 (Winter, 2010) indicated that all of the crystals were triclinic and belonged to space group *P*1. Since this is an uncommon space group, other possibilities were checked by integration of the raw diffraction images with *DIALS* (Waterman *et al.*, 2013) and scaling with *AIMLESS* (Evans & Murshudov, 2013), which also suggested that *P*1 was the most likely symmetry. The best crystal diffracted to 2.18 Å resolution and the data were of good quality, as suggested by *phenix.xtriage* (Zwart *et al.*, 2005). Analysis using *MATTHEWS_COEF* (Kantardjieff & Rupp, 2003; Matthews, 1968) suggested that this crystal form possesses six monomers per asymmetric unit, with a solvent content of 47.32%.

The structure was determined by molecular replacement with *Phaser MR* (McCoy *et al.*, 2007) using the type I L-asparaginase from *Pyrococcus horikoshii* (PhA; PDB entry 1wls; 59% sequence identity to TkA; Yao *et al.*, 2005) as the search model. The structure-factor correlation coefficient was approximately 40% for the first correctly positioned monomer and it increased by a few percent for each monomer that was added to the structural model. Manual rebuilding and correction were accomplished using *Coot* (Emsley *et al.*, 2010) followed by TLS, local NCS and restrained refinement by the use of *REFMAC*5 (Murshudov *et al.*, 1997, 2011). The structure was then submitted to the *PDB_REDO* website (Joosten *et al.*, 2014) for further refinement and validation. Model validation was also performed using *MolProbity* (Chen *et al.*, 2010). All of the statistics for data collection, data processing, structure determination and refinement are shown in Table 1. The *VADAR* (Willard *et al.*, 2003) and *ESBRI* (Costantini *et al.*, 2008) online services were used to analyse hydrogen bonds, salt bridges and other factors related to the thermostability of the enzyme. The figures were prepared using *MarvinSketch* (ChemAxon), *PyMOL* (Schrödinger) and *CueMol* (http://www.cuemol.org).

3. Results and discussion

3.1. Quality of the model

The electron density for the first five chains (*A*, *B*, *C*, *D* and *E*) is of good quality, whilst that for chain *F* suggests that this monomer is slightly more disordered than the others since a number of regions have poor electron density. Indeed, the first

four chains are characterized by all-atom average *B* factors of around 43.6 Å^2, while chains *E* and *F* have higher *B* factors of 53.3 and 59.3 Å^2, respectively. Data analysis suggested that the data were of good quality and that no anisotropy or translational NCS was present. Analysis by *MolProbity* indicated that 95.5% of the residues are in the Ramachandran favoured region, with an additional 1.0% being classed as outliers.

3.2. Overall structure

The six chains in the asymmetric unit of TkA share a very similar structure, with r.m.s.d. values ranging from 0.08 to 0.15 Å for the C$^\alpha$ atoms. When considering chain *A* only, the r.m.s.d. values between TkA and the *E. coli* type I and II L-asparaginases are 1.11 and 1.67 Å, respectively. Thus, TkA has a similar overall fold to both classes of L-asparaginase.

Each subunit of TkA consists of an N-terminal and a C-terminal α/β domain connected by a linker loop formed by residues 185–203 (Fig. 2). The N-terminal domain contains an eight-stranded mixed β-sheet (β1, β4–β8, β11 and β12) flanked by four α-helices (α1–α4). It has been shown that the β-hairpin composed of strands β2 and β3 is highly flexible and is often characterized by poor or no electron density. This β-hairpin is involved in substrate binding and catalysis, and adopts 'open' and 'closed' conformations (Nguyen *et al.*, 2016). However, this often disordered area is characterized by good electron density in all of the subunits of TkA. Chains *B* and *D* show a more 'open' conformation compared with EcA (PDB entry 2p2d; Yun *et al.*, 2007), whilst the other chains adopt a more 'closed' conformation, leaving this region of the *E. coli* enzyme somewhere between them (Fig. 3). Of greater significance is the fact that the α1 helix has moved towards the active site in all of the subunits. The α4 helix has also moved slightly towards the active site. Thus, it is very likely that these two helical segments participate in substrate recognition in addition to the flexible β-hairpin. Part of the main sheet is formed by a pronounced β-hairpin, involving β9 and β10, which resides between the two domains of the enzyme and participates in subunit adhesion within the physiological dimer. The active site of TkA is located in a pocket in which the two catalytically important threonine residues, Thr11 and Thr85, occur and consists of residues from the two neighbouring monomers of the dimer.

The relatively small C-terminal domain is formed mainly by a three-stranded parallel β-sheet (β13–β15) and five α-helices (α5–α9). There is also a putative allosteric site which is located between the β15 strand and the α8 helix (Fig. 2) and is involved in asparagine binding (Yun *et al.*, 2007), giving rise to a cooperative conformational switch from an inactive to an active form of the enzyme.

The sequence alignment of TkA with several homologues is shown in Fig. 4. TkA shares 58.3 and 60.6% sequence identity over all residues with the L-asparaginase from *P. furiosus* (PfA) and an L-asparaginase I homologue from *P. horikoshii* (PhA), respectively. In contrast, it has sequence identities of below 30% with EcA, EcAII and ErAII. However, there are many highly conserved regions in all of these proteins, including the active site, which suggests that they are likely to share a similar tertiary structure and catalytic mechanism.

3.3. Active site

As mentioned before, TkA is active as a homodimer and many residues from both subunits are involved in substrate recognition and catalysis. Structural comparison identified Thr11, Tyr21, Ser54, Thr55, Thr85, Asp86 and Lys156, as well as Tyr233′ and Glu275′ from the neighbouring subunit, as being both important and highly conserved in L-asparaginases (Figs. 3 and 4). One of the two key residues participating in catalysis, Thr11, resides in a β-hairpin whose flexibility is considered to be deeply involved in the activity of the enzyme. The other key residue, Thr85, which is located in the loop between α3 and β5, mediates sequential 'ping-pong' nucleophilic attacks together with Thr11 during amidohydrolysis

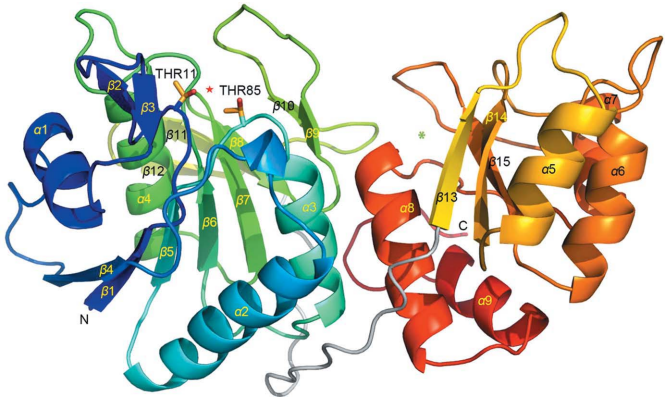


Figure 2
The overall structure of TkA. Each monomer is composed of an N-terminal and a C-terminal domain connected by a loop (coloured grey). The two key threonine residues involved in catalysis are shown in ball-and-stick representation. The active site and the putative allosteric site are indicated by a red star and a green asterisk, respectively.

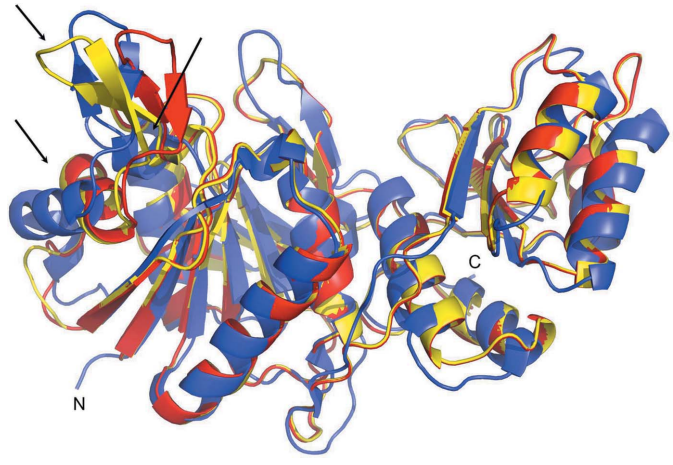


Figure 3
Structural superposition of TkA with the *E. coli* type I L-asparaginase. The structures are in approximately the same orientation as that shown in Fig. 2. Chains *A* and *B* of TkA and both subunits of the *E. coli* enzyme (PDB entry 2p2d) are coloured red, yellow and blue, respectively. The arrows indicate the segments which adopt different conformations in response to substrate binding.

(Harms *et al.*, 1991). Since some unexpected electron density was identified in the active site of each monomer, many different molecules were fitted in an effort to identify it, including the substrate asparagine and the product aspartic acid. However, this feature was finally interpreted and successfully refined as a phosphate ion. This was one of the components of the crystallization buffer and has been reported in other asparaginase structures (Tomar *et al.*, 2014; Wehner *et al.*, 1992). The phosphate anion occupies the binding site for the substrate and forms many interactions with the side chains of the neighbouring residues, including Thr11 and Thr85. Tomar *et al.* (2014) also suggested that the binding of a ligand (asparagine, citrate or phosphate) stabilizes the flexible β-hairpin, which acts as a gatekeeper and prevents further substrate entry. However, Yun *et al.* (2007) reported that this hairpin is still not visible in the Asp- or Asn-bound structures. In addition, the *B* and *D* chains in the TkA enzyme adopt a more 'open' conformation than that of the ligand-free *E. coli* type I L-asparaginase (PDB entry 2p2d).

Yun *et al.* (2007) indicated that L-asparaginases which have high L-glutaminase activity possess smaller residues at the equivalent position to residue Gly237′ in TkA. They predicted that those with a glycine at this position should have substantial glutaminase activity. However, whilst TkA does possess a glycine at this position, no glutaminase activity has been observed.

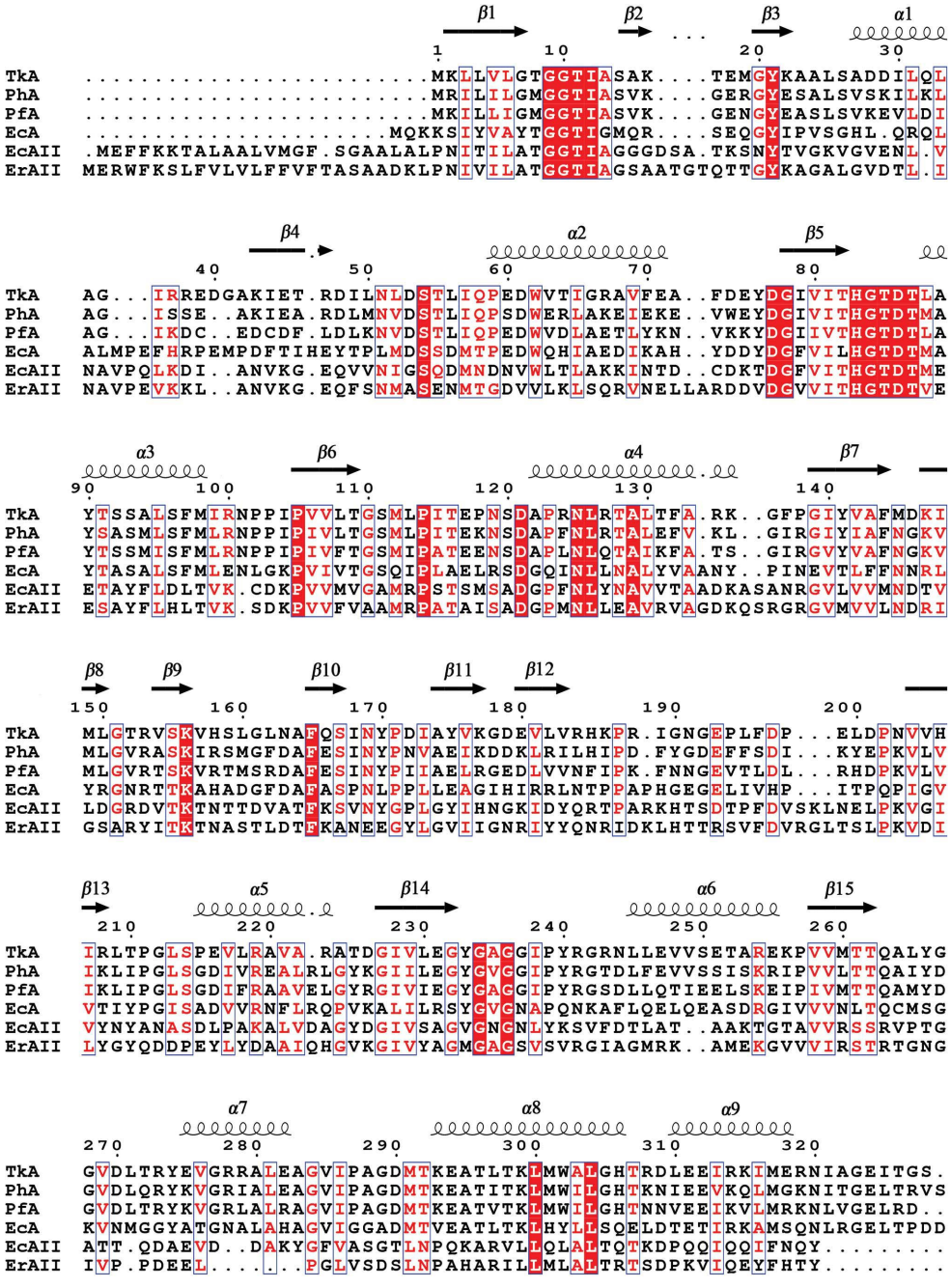


Figure 4
A sequence alignment showing the secondary-structure characteristics of TkA. Helices and strands are labelled according to the TkA structure. All of the conserved residues are boxed and the fully conserved residues are coloured white with a red background, while the less conserved residues are coloured red. The alignment was performed using *ESPript* 3.0 (Robert & Gouet, 2014).

3.4. Thermostability

The great thermostability of hyperthermophilic proteins can be attributed to several factors. These proteins tend to have greater hydrophobicity (Haney *et al.*, 1997), more hydrogen bonds (Vogt *et al.*, 1997; Vogt & Argos, 1997) and salt bridges (Yip *et al.*, 1995, 1998; Haney *et al.*, 1997; Kumar, Ma *et al.*, 2000), increased helical content, low occurrence of thermolabile residues such as Cys and Ser (Russell *et al.*, 1997), high occurrence of Arg, Tyr and Pro (Watanabe *et al.*, 1997; Bogin *et al.*, 1998; Haney *et al.*, 1997), amino-acid substitutions within and outside the secondary structures (Zuber, 1988; Haney *et al.*, 1997; Russell *et al.*, 1997), better packing, smaller and less numerous cavities, deletion or shortening of loops (Russell *et al.*, 1997), increased surface area buried upon oligomerization (Salminen *et al.*, 1996) and increased polar surface area (Haney *et al.*, 1997; Vogt *et al.*, 1997; Vogt & Argos, 1997). However, it should be noted that no single factor proposed to

contribute towards protein thermostability is 100% consistent in all thermophilic proteins. Kumar, Tsai *et al.* (2000) observed that the most consistent trend is shown by side chain–side chain hydrogen bonds and salt bridges. They may rigidify a thermophilic protein in the room-temperature range and the protein may still be flexible enough at high temperature in order to function (Jaenicke & Böhm, 1998).

A comparison of some of these factors for several thermophilic and mesophilic L-asparaginases is shown in Table 2, and those which probably contribute to the thermostability of the TkA dimer are indicated in bold. The increased content of salt bridges and arginine residues along with the lower number of Cys and Ser residues are notable features of TkA.

The monomer–monomer contacts in the dimer are predominantly mediated by the C-terminal domains of each subunit, which form a continuous β-sheet involving strands $\beta13$–$\beta15$ of both monomers (Fig. 5). The occurrence of significantly higher than average B factors in the outer-facing helical segments of the C-terminal domains of the dimer formed by the last two monomers in the asymmetric unit (chains E and F) suggests that the dimerization interface of the protein may have greater flexibility at higher temperatures, perhaps leaving the structure of the catalytic domains in the dimer relatively unaffected.

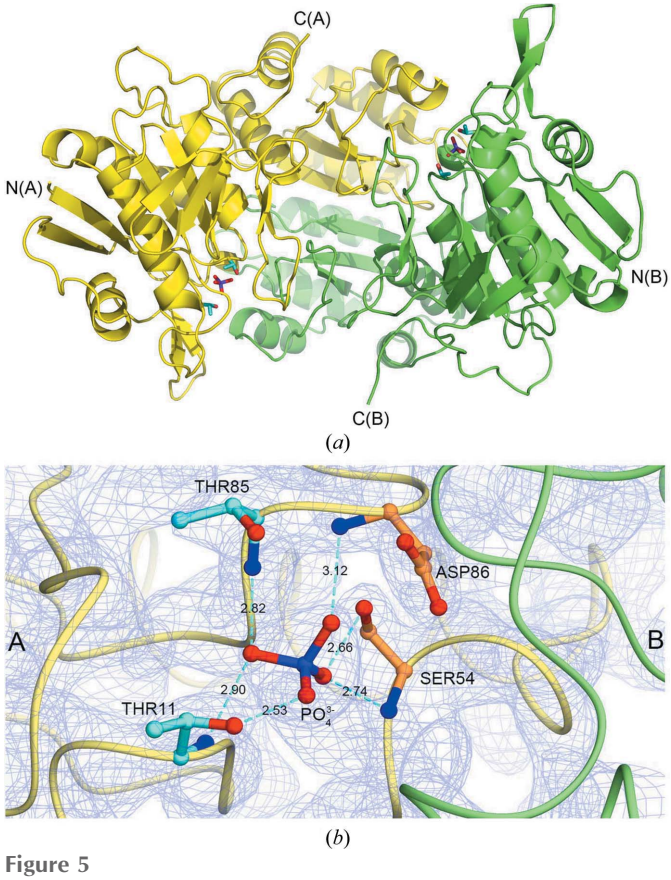


Figure 5
Dimer assembly and the active site of TkA. (*a*) The dimer assembly formed between chains *A* (yellow) and *B* (green). (*b*) The active site formed by residues from both chain *A* (yellow) and chain *B* (green). The two key threonine residues involved in catalysis are coloured cyan and other residues that participate in substrate recognition are coloured orange. A phosphate (purple) ion has been identified tightly bound in the active site of each subunit. The $2F_o - F_c$ electron-density map is shown in pale blue, contoured at 1.5 r.m.s.

Table 2
Thermostability-related factors for several thermophilic and mesophilic L-asparaginases.

Factors which probably contribute to the thermostability of the TkA dimer are indicated in bold.

	Thermophilic			Mesophilic		
Enzyme†	TkA	1wls	4q0m	2p2d	3ntx	2ocd
Salt bridges (%)	20.0	17.5	19.0	13.9	9.2	9.9
Hydrogen bonds (%)	73	72	73	71	73	73
Helix content (%)	29	30	28	29	29	28
Pro content (%)	5.5	4.0	4.0	6.5	5.3	6.2
Arg content (%)	7.0	4.6	4.6	4.4	4.1	3.0
Tyr content (%)	3.0	3.7	3.4	4.1	3.3	3.9
Cys content (%)	0	0	0.6	0.3	0.3	0.6
Ser content (%)	4.3	6.7	4.9	4.7	7.4	5.9

† All of the enzymes are represented by their PDB code (except for TkA) as follows: 1wls, L-asparaginase I from *P. horikoshii*; 4q0m, L-asparaginase I from *P. furiosus*; 2p2d, L-asparaginase I from *E. coli*; 3ntx, L-asparaginase from *Yersinia pestis* (Center for Structural Genomics of Infectious Diseases, unpublished work); 2ocd, L-asparaginase I from *Vibrio cholerae* (Center for Structural Genomics of Infectious Diseases, unpublished work).

4. Summary

The crystal structure of TkA has been determined at 2.18 Å resolution in space group $P1$ with six monomers in the asymmetric unit, forming three dimeric pairs. Each subunit of TkA consists of an N-terminal and a C-terminal α/β domain connected by a linker loop. TkA is active as a homodimer and many residues from the neighbouring molecules in a dimer are involved in substrate recognition as well as in catalysis. The N-terminal domain contains a highly flexible β-hairpin which adopts 'open' and 'closed' conformations in different subunits of the TkA structure. This region is usually only visible in L-asparaginase structures that adopt a 'closed' conformation, whilst it is characterized with good electron density in all of the subunits in the TkA structure. One phosphate anion has been built in the active site. The great thermostability of TkA may be attributed to the higher arginine content, the lower numbers of Cys and Ser residues and the increased content of salt bridges.

Acknowledgements

We gratefully acknowledge Diamond Light Source for X-ray beam time and travel support for data collection (award MX12342).

References

Albertsen, B., Jakobsen, P., Schrøder, H., Schmiegelow, K. & Carlsen, N. T. (2001). *Cancer Chemother. Pharmacol.* **48**, 77–82.
Atkins, C. A., Pate, J. S. & Sharkey, P. J. (1975). *Plant Physiol.* **56**, 807–812.
Avramis, V. I. *et al.* (2002). *Blood*, **99**, 1986–1994.
Bogin, O., Peretz, M., Hacham, Y., Burstein, Y., Korkhin, Y., Kalb (Gilboa), A. J. & Frolow, F. (1998). *Protein Sci.* **7**, 1156–1163.
Boyd, J. W. & Phillips, A. W. (1971). *J. Bacteriol.* **106**, 578–587
Broome, J. (1961). *Nature (London)*, **191**, 1114–1115.

Broome, J. (1963). *J. Exp. Med.* **118**, 121–148.
Cedar, H. & Schwartz, J. H. (1967). *J. Biol. Chem.* **242**, 3753–3755.
Cedar, H. & Schwartz, J. H. (1968). *J. Bacteriol.* **96**, 2043–2048.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Chohan, S. M. & Rashid, N. (2013). *J. Biosci. Bioeng.* **116**, 438–443.
Costantini, S., Colonna, G. & Facchiano, A. M. (2008). *Bioinformation*, **3**, 137–138.
Davidson, L., Brear, D. R., Wingard, P., Hawkins, J. & Kitto, G. B. (1977). *J. Bacteriol.* **129**, 1379–1386.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* D**69**, 1204–1214.
Friedman, M. (2003). *J. Agric. Food Chem.* **51**, 4504–4526.
Fullmer, A., O'Brien, S., Kantarjian, H. & Jabbour, E. (2010). *Expert Opin. Emerg. Drugs*, **15**, 1–11.
Gorrec, F. (2009). *J. Appl. Cryst.* **42**, 1035–1042.
Haney, P., Konisky, J., Koretke, K., Luthey-Schulten, Z. & Wolynes, P. (1997). *Proteins*, **28**, 117–130.
Harms, E., Wehner, A., Aung, H.-P. & Röhm, K. (1991). *FEBS Lett.* **285**, 55–58.
Jaenicke, R. & Böhm, G. (1998). *Curr. Opin. Struct. Biol.* **8**, 738–748.
Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ*, **1**, 213–220.
Kantardjieff, K. A. & Rupp, B. (2003). *Protein Sci.* **12**, 1865–1871.
Kidd, J. G. (1953). *J. Exp. Med.* **98**, 565–582.
Kiriyama, Y., Kubota, M., Takimoto, T., Kitoh, T., Tanizawa, A., Akiyama, Y. & Mikawa, H. (1989). *Leukemia*, **3**, 294–297.
Kumar, K., Kataria, M. & Verma, N. (2013). *Artif. Cell. Nanomed. Biotechnol.* **41**, 184–188.
Kumar, S., Ma, B., Tsai, C.-J. & Nussinov, R. (2000). *Proteins*, **38**, 368–383.
Kumar, S., Tsai, C.-J. & Nussinov, R. (2000). *Protein Eng.* **13**, 179–191.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
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.
Möricke, A. *et al.* (2008). *Blood*, **111**, 4477–4489.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nguyen, H. A., Su, Y. & Lavie, A. (2016). *Biochemistry*, **55**, 1246–1253.
Pui, C.-H. *et al.* (2009). *New Engl. J. Med.* **360**, 2730–2741.
Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
Russell, R. J., Ferguson, J. M., Hough, D. W., Danson, M. J. & Taylor, G. L. (1997). *Biochemistry*, **36**, 9983–9994.
Salminen, T., Teplyakov, A., Kankare, J., Cooperman, B. S., Lahti, R. & Goldman, A. (1996). *Protein Sci.* **5**, 1014–1025.
Shrivastava, A., Khan, A. A., Khurshid, M., Kalam, M. A., Jain, S. K. & Singhal, P. K. (2016). *Crit. Rev. Oncol. Hematol.* **100**, 1–10.
Sieciechowicz, K. A., Joy, K. W. & Ireland, R. J. (1988). *Phytochemistry*, **27**, 663–671.
Silverman, L. B., Gelber, R. D., Dalton, V. K., Asselin, B. L., Barr, R. D., Clavell, L. A., Hurwitz, C. A., Moghrabi, A., Samson, Y., Schorin, M. A., Arkin, S., Declerck, L., Cohen, H. J. & Sallan, S. E. (2001). *Blood*, **97**, 1211–1218.
Soares, A. L., Guimarães, G. M., Polakiewicz, B., de Moraes Pitombo, R. N. & Abrahão-Neto, J. (2002). *Int. J. Pharm.* **237**, 163–170.
Song, P., Ye, L., Fan, J., Li, Y., Zeng, X., Wang, Z., Wang, S., Zhang, G., Yang, P., Cao, Z. & Ju, D. (2015). *Oncotarget*, **6**, 3861–3873.
Stams, W. A., den Boer, M. L., Beverloo, H. B., Meijerink, J. P., Stigter, R. L., van Wering, E. R., Janka-Schaub, G. E., Slater, R. & Pieters, R. (2003). *Blood*, **101**, 2743–2747.
Tomar, R., Sharma, P., Srivastava, A., Bansal, S., Ashish & Kundu, B. (2014). *Acta Cryst.* D**70**, 3187–3197.
Verma, N., Kumar, K., Kaur, G. & Anand, S. (2007). *Crit. Rev. Biotechnol.* **27**, 45–62.
Villa, P., Corada, M. & Bartosek, I. (1986). *Toxicol. Lett.* **32**, 235–241.
Vogt, G. & Argos, P. (1997). *Fold. Des.* **2**, S40–S46.
Vogt, G., Woell, S. & Argos, P. (1997). *J. Mol. Biol.* **269**, 631–643.
Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y. & Suzuki, Y. (1997). *J. Mol. Biol.* **269**, 142–153.
Waterman, D., Winter, G., Parkhurst, J., Fuentes-Montero, L., Hattne, J., Brewster, A., Sauter, N. & Evans, G. (2013). *CCP4 Newsl. Protein Crystallogr.* **49**, 13–15. http://www.ccp4.ac.uk/newsletters/newsletter49/articles/CCP4Dispatchers.pdf.
Wehner, A., Harms, E., Jennings, M. P., Beacham, I. R., Derst, C., Bast, P. & Röhm, K. H. (1992). *FEBS J.* **208**, 475–480.
Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D. & Wishart, D. S. (2003). *Nucleic Acids Res.* **31**, 3316–3319.
Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
Yao, M., Yasutake, Y., Morita, H. & Tanaka, I. (2005). *Acta Cryst.* D**61**, 294–301.
Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D. & Rice, D. W. (1998). *FEBS J.* **255**, 336–346.
Yip, K., Stillman, T., Britton, K., Artymiuk, P., Baker, P., Sedelnikova, S., Engel, P., Pasquo, A., Chiaraluce, R. & Consalvi, V. (1995). *Structure*, **3**, 1147–1158.
Yun, M.-K., Nourse, A., White, S. W., Rock, C. O. & Heath, R. J. (2007). *J. Mol. Biol.* **369**, 794–811.
Zuber, H. (1988). *Biophys. Chem.* **29**, 171–179.
Zwart, P., Grosse-Kunstleve, R. & Adams, P. (2005). *CCP4 Newsl. Protein Crystallogr.* **43**, contribution 10. http://www.ccp4.ac.uk/newsletters/newsletter42/articles/CCP4_2005_PHZ_RWGK_PDA.doc.