Acta Biochimica Polonica

Vol. 44 No. 3/1997

491–504

QUARTERLY

Why a "benign" mutation kills enzyme activity. Structure-based analysis of the A176V mutant of *Saccharomyces cerevisiae* L-asparaginase I*

David T. Bonthron[1] and Mariusz Jaskólski[2, 3]✉

[1]*Human Genetics Unit, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, U.K.*
[2]*Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznań, Poland*
[3]*Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland*

Keywords: L-asparaginase, *Saccharomyces cerevisiae*, A → V mutant, leukemia, molecular modelling, sequence analysis

A conservative and apparently harmless A176V mutation in intracellular *S. cerevisiae* L-asparaginase (ScerAI) completely abolishes the enzyme activity. Sequence and structural comparisons with type II bacterial L-asparaginases show that the mutated residue is in a very conservative region and plays a vital role in the cohesion of functional tetramers of these enzymes through participation in side-chain...main-chain (Ser) Oγ...O (Ala) hydrogen bonds across the tetramer interface. The fact that bacterial L-asparaginases of type I show less conservation in this region suggests that they may have different quaternary structure while adopting the subunit fold and intimate dimer architecture of type II enzymes. A comparison of all available sequences of microbial L-asparaginases confirms that separate intra- and extra-cellular enzymes evolved in prokaryotes and eukaryotes independently. However, an analysis of the available complete genome sequences reveals a surprising fact that *Haemophilus influenzae* possesses only a type II asparaginase while the archaebacterium *Methanococcus jannaschii* has a type I gene, but not a type II.

*The research of Mariusz Jaskólski was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute. The British Council and the State Committee for Scientific Research (KBN) provided partial support for this collaborative project.

✉Author for correspondence: Mariusz Jaskólski, Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704 Poznań, Poland, tel. (+061)+852-85-03, fax: (+061)+852-05-32, email: mariusz@krystal.amu.edu.pl

Abbreviations: AGA, *Acinetobacter glutaminasificans* L-glutaminase-asparaginase; BlA, *Bacillus licheniformis* L-asparaginase; BsA, *Bacillus subtilis* L-asparaginase; EcAI, *E. coli* intracellular L-asparaginase; EcAII, *E. coli* periplasmic L-asparaginase; ErA, *Erwinia chrysanthemi* L-asparaginase; HiA, *Haemophilus influenzae* Rd L-asparaginase; MjA, *Methanococcus jannaschii* L-asparaginase; MtA, *Mycobacterium tuberculosis* L-asparaginase; ORF, open reading frame; PDB, Protein Data Bank; PGA, *Pseudomonas* 7A L-glutaminase-asparaginase; r.m.s., root-mean-square; ScerAI, *S. cerevisiae* intracellular L-asparaginase; ScerAII, *S. cerevisiae* cell wall L-asparaginase; WsA, *Wolinella succinogenes* L-asparaginase.

The yeast *Saccharomyces cerevisiae*, in a way analogous to several bacterial organisms, produces two biochemically and genetically distinct L-asparaginase isoenzymes responsible for converting L-asparagine to L-aspartic acid and ammonia (Dunlop *et al.*, 1978). One of these (asparaginase II, ScerAII, the *ASP3* gene product) is a cell wall glycoprotein, while the other (asparaginase I, ScerAI, the product of *ASP1*) is a constitutively expressed intracellular enzyme. The ethylmethanesulphonate-induced, asparaginase I-deficient *S. cerevisiae* strain, XE101-1A, contains a conservative A176V change in the amino-acid sequence. Genetic evidence indicated that this substitution was responsible for the null *asp1*-12 allele carried by XE101-1A (Sinclair *et al.*, 1994). More recently, we have also directly confirmed this conclusion by expressing *S. cerevisiae ASP1* in *E. coli*; the A176V mutation abolishes activity of the recombinant ScerAI enzyme (unpublished).

It is quite surprising that this highly conservative A→V substitution, at a hydrophobic residue that is obviously not involved in the catalytic reaction, should render the enzyme completely inactive. Since the three-dimensional crystallographic structure of the ScerAI enzyme is not available, we decided to investigate this question using the crystallographically determined structures of several bacterial L-asparaginases, including the enzymes from *Escherichia coli* (EcAII) and *Erwinia chrysanthemi* (ErA) which for many years have been in clinical use as potent anti-leukemic agents (Gallagher *et al.*, 1989).

In addition to type II enzymes, such as EcAII and ErA, which are found in the periplasm, there are also biochemically and genetically separate cytosolic enzymes classified as type I, most notably EcAI, another isoenzyme found in *E. coli*.

All L-asparaginases that have been characterized structurally so far are class II bacterial enzymes that are active as tetramers and exist in their crystal structures as 222-symmetric homotetramers (Swain *et al.*, 1993; Miller *et al.*, 1993; Lubkowski *et al.*, 1994a, b, 1996; Jakob *et al.*, 1997). More precisely, an asparaginase tetramer should be viewed as a dimer of intimate dimers (AC + BD) (Swain *et al.*, 1993; subunit designation as in PDB file 3ECA). Basically, an active site (which in the EcAII sequence includes T89, K162, D90, and possibly other residues, Fig. 1a) is formed in the larger N-terminal domain of a subunit (for instance, A) with some additional contacts from the C-terminal domain of the intimately bound subunit (C) (Figs. 2, 3). The reason why the enzyme must assume a tetrameric form in order to be active remains elusive in spite of the accumulated structural data. The intriguing consequences of the conservative A176V mutation in ScerAI may provide a starting point for the understanding of the role of the tetrameric nature of the enzyme.

SEQUENCE ALIGNMENTS

Prior to structural comparisons, we aligned the amino-acid sequences of all complete members of the microbial L-asparaginase family identified in the EMBL and SWISSPROT sequence databases (http://www.ebi.ac.uk). Three of these thirteen sequences (those of *Haemophilus influenzae*, *Mycobacterium tuberculosis* and *Methanococcus jannaschii*) are derived from large-scale genome sequencing projects. These predicted enzymes have not been characterized biochemically, but the first of them is clearly very closely related to EcAII (Fig. 1a). The sequences were compared using the PILEUP program of the GCG package (Devereux *et al.*, 1984). In Fig. 1 they are ordered according to their degree of similarity, determined by the pairwise comparisons generated by PILEUP. Those enzymes for which three-dimensional structures have been determined have their names underlined.

The (mutually cross-referenced) EcAI sequences in the EMBL and SWISSPROT data banks are different. While the former (accession code M26934) gives the N-terminal sequence as reported in the original reference (Jerlstroem *et al.*, 1989), the latter (accession code P18840) is derived by translating from a different initiation codon (16 amino acids farther upstream). Originally, the initiation codon for *ansA* was assigned on the basis of

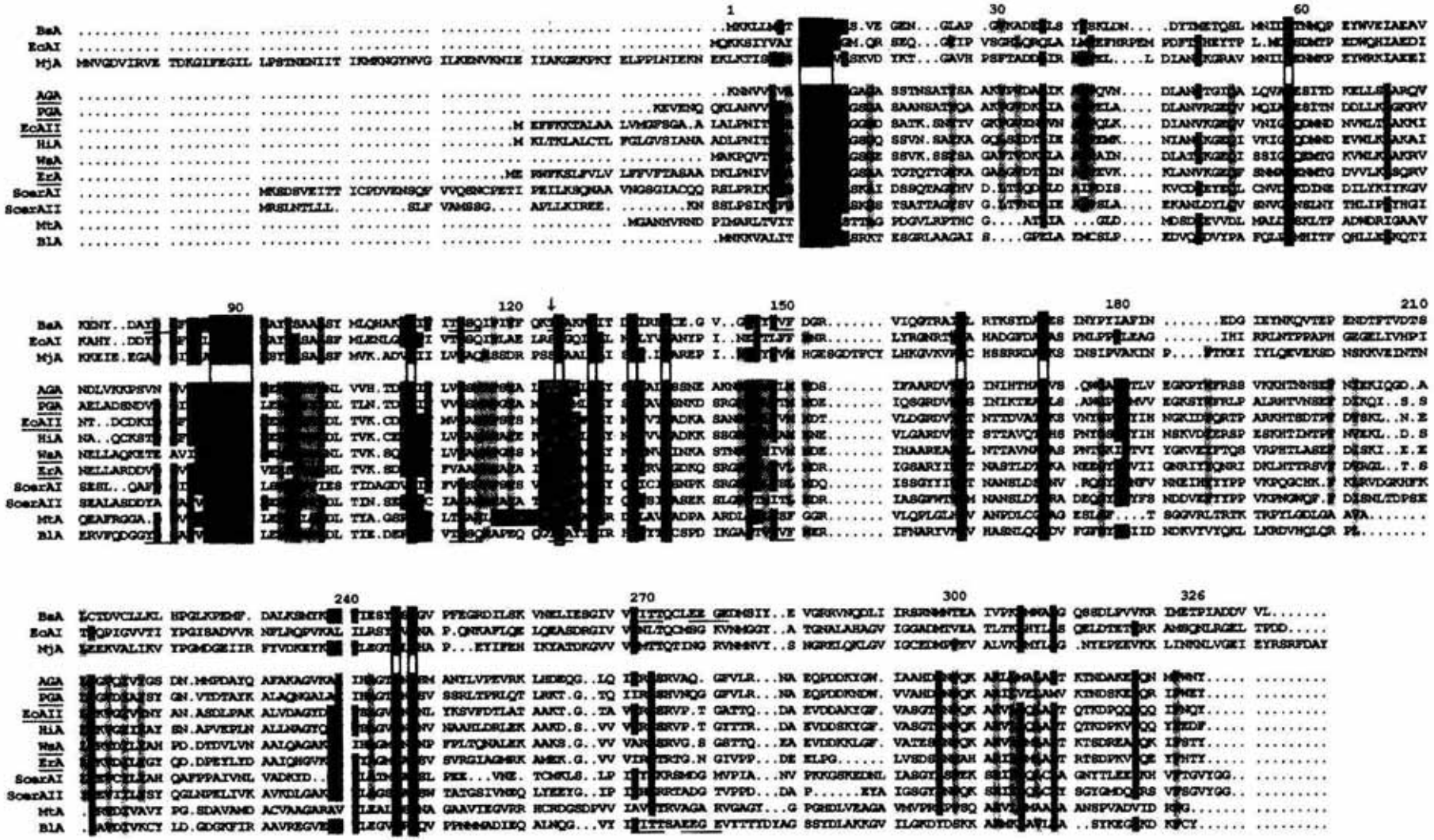
being the only one within the cloned region which had a reasonable Shine-Dalgarno consensus adjacent to it (Jerlstroem *et al.*, 1989). However, the reading frame was continuously open upstream of this point to the end of the Jerlstroem *et al.* clone, and a TGGTI sequence is encoded immediately upstream of their proposed initiator codon. Since this motif is highly conserved in all other asparaginases (Fig. 1a) and since the second T is essential for catalysis (Swain *et al.*, 1993; Palm *et al.*, 1996), the originally proposed EcAI sequence is certainly truncated at its N-terminus. We translated the relevant region of the longer DNA sequence D90820 (Kohara clone No. 329; Aiba *et al.*, 1996) to define the upstream limits of the ORF. Only one upstream methionine is present in this ORF, and the resulting protein has 16 additional N-terminal amino acids, as in P18840; this revised sequence was used in the alignment of Fig 1a.

Five of the bacterial enzymes (EcAII, ErA, PGA, AGA, WsA) are recognized as type II microbial asparaginases because of a high level of sequence similarity and localisation in the periplasmic compartment. In contrast EcAI, the constitutive intracellular *E. coli* enzyme, is classified as type I. The two *S. cerevisiae* enzymes have not been placed within this grouping scheme. We previously suggested that the separate intra- and extracellular isoenzymes arose in the prokaryotic and eukaryotic lineages by independent ancestral asparaginase gene duplication events (Sinclair *et al.*, 1994). The ScerAII sequence we used in our comparison at that time was C-terminally truncated compared to other asparaginases. However, a more recent cosmid sequence from the *S. cerevisiae* genome project (Goffeau *et al.*, 1996) which includes four identical copies of the *ASP3* gene (U51921) differs from the previous sequence (J03926) by a single frameshift. This alters and extends the C-terminal part of ScerAII, and our alignment of the new sequence shows clear similarities to ScerAI extending right to the C terminus. A sequencing error in J03926 (Kim *et al.*, 1988) is thus most likely. In the new alignment, the conclusion regarding the independent duplication events in bacteria and yeast is reinforced, since ScerAI and ScerAII are much closer to each other than to any of the bacterial enzymes (Fig. 1a). Furthermore, both of these enzymes appear to belong to the class II asparaginase family, even though ScerAI is intracellular.

The multiple sequence alignment and associated dendrogram (Fig. 1b) also suggest classifications for the *Mycobacterium* and *Bacillus* enzymes (though the dendrogram does not represent a true evolutionary tree). The *B. subtilis* asparaginase is grouped with EcAI rather than EcAII, and thus appears to be a type I enzyme. The *B. licheniformis* enzyme, though, appears most closely related to the *Mycobacterium* asparaginase, which is a type II enzyme (*vide infra*).

For a more objective assessment of the relationship of the *Bacillus* and *Mycobacterium* asparaginases to other members of the family, we eliminated the variable-length N- and C-terminal fragments and performed multiple pairwise similarity comparisons over residues 2–326 of Fig. 1, using the DISTANCES GCG program. By this method, BsA is most similar to the archaebacterial *Methanococcus* asparaginase and then to EcAI; it is much less similar to all other sequences. We therefore classify BsA (and MjA) as type I enzymes, though it should be emphasized that the divergence of the three type I enzymes from each other is much greater than that within the Gram negative bacterial type II group, and is in fact comparable to the distance of the two *S. cerevisiae* enzymes from any of the bacterial type II enzymes. MtA, in contrast, is clearly much more like a type II than a type I enzyme in the DISTANCES analysis.

BlA (van Dijl *et al.*, 1991) is the most difficult enzyme to classify. It has the greatest similarity to ScerAII and MtA (type II enzymes) in DISTANCES, but there are nonetheless some short motifs (underlined in Fig. 1a) uniquely shared with the *B. subtilis* enzyme, including the *TDA* in place of the *ADG* (residues 123–125 in EcAII sequence) motif discussed in detail below. Analysis using PILEUP and DISTANCES of four sequences alone (EcAI, EcaII, BsA, BlA) again places EcAI and BsA in one group and EcAII and BlA in another. Compared pairwise



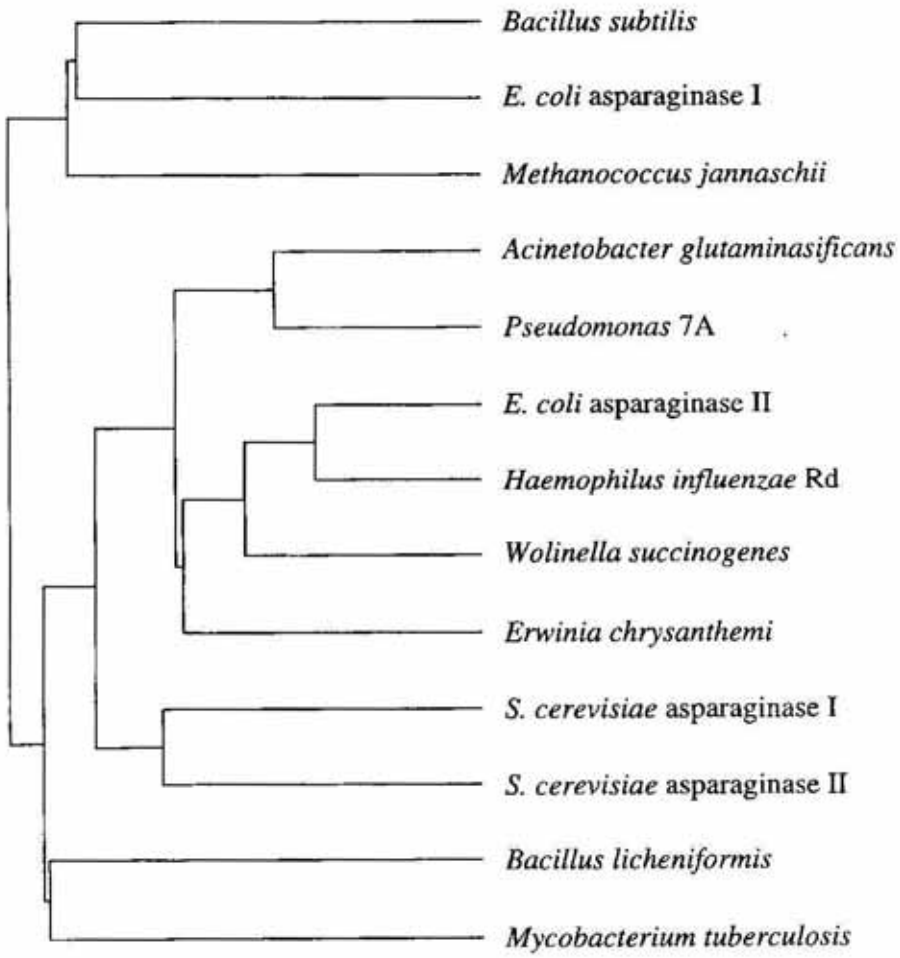


Figure 1. a. Alignment of the amino-acid sequences of bacterial and fungal L-asparaginases identified in the EMBL and SWISSPROT sequence data banks (http://www.ebi.ac.uk).

The identifying codes and the accession numbers of the aligned L-asparaginase sequences are as follows: BsA, *Bacillus subtilis*, M63264 (Sun & Setlow, 1991); MjA, *Methanococcus jannashii*, Q60331 (Bult *et al.*, 1996); EcAI, *E. coli*, cellular, type I, (*ansA*) M26934 (Jerlstroem *et al.*, 1989); AGA, *Acinetobacter glutaminasificans*, P10172 (Tanaka *et al.*, 1988); PGA, *Pseudomonas 7A*, P10182 (Lubkowski *et al.*, 1994a); EcAII, *E. coli*, periplasmic, type II, (*ansB*) M34234 (Bonthron, 1990); HiA, *Haemophilus influenzae*, P43843 (Fleischmann *et al.*, 1995); WsA, *Wolinella succinogenes*, P50286 (Lubkowski *et al.*, 1996); ErA, *Erwinia chrysanthemi*, M14741 (Minton *et al.*, 1986); ScerAI, *S. cerevisiae*, cellular asparaginase I, (*ASP1*) Z27406 (Sinclair *et al.*, 1994); ScerAII, *S. cerevisiae*, cell-wall asparaginase II, (*ASP3*) U51921 (Goffeau *et al.*, 1996); MtA, *Mycobacterium tuberculosis*, Q10759; BlA, *Bacillus licheniformis*, Z11497 (van Dijl *et al.*, 1991). This multiple alignment was produced in GCG PILEUP (Devereux *et al.*, 1984). The order of the last two sequences (MtA and BlA) in the alignment has been reversed for ease of indicating the highlighted residues in the SADGP critical region. Sequences classified as type I are separated from the others by a single empty line. Numbering above the BsA sequence refers to the mature EcAII subunit. The critical SADGP region is shown in white on black lettering, as is the five-residue repeat in the MtA sequence. The shading of conserved residues was performed according to the following scheme. BOXSHADE v.3.0 (Hofmann & Baron, unpublished) was used to identify positions where 11 or more sequences are identical or similar (according to the usual groupings GA;VLIM;ST;DE;KR;FYW;NQ). Then at each of these positions, green shading was used to indicate identity (if 8 or more sequences were identical) and yellow similarity. Positions where all 13 sequences show identity were further highlighted by boxes; in many cases (e.g. T12, S58, T89, K162) these are elements postulated to be involved in the catalytic process (Swain *et al.*, 1993). Next, a separate BOXSHADE analysis of the eight clearly type II enzymes (excluding MtA and BlA) was used to identify positions where all eight are identical or similar, and such additional positions were shaded in blue, also shading any residues in the other five sequences which happen to match. The A176 residue mutated in *S. cerevisiae* XE101-1A is indicated by an arrow. In this alignment it is matched with A123 in the EcAII sequence. Underlined residues are motifs of 3 or more amino acids which are uniquely conserved between the two *Bacillus* enzymes. **b.** Dendrogram generated by PILEUP showing the clustering relationships between the asparaginase sequences.

(GAP or BESTFIT) to BlA, EcAI has slightly more similar but EcAII slightly more identical residues. Overall, the difficulty in classifying BlA indicates that further information (including the crystal structure of a type I enzyme) is required to define clearly the relationship between the families.

The altered residue in the A176V mutant of ScerAI is located in a conserved region, SADGP, residues 175–179 in ScerAI or 122–126 in EcAII (Fig. 1). This sequence is shared by both *S. cerevisiae* enzymes and by all the clearly type II bacterial asparaginases, with a minor exception in PGA, which has a methionine instead of the proline. In contrast, the archetypal class I enzyme, EcAI, has an RSDGQ pattern in this region, sharing only the DG dipeptide with the class II family. Both *Bacillus* sequences have a *TDA* pattern, homologous with the central SDG tripeptide in EcAI, which is in favour of a type I classification (though see above). The most intriguing pattern is seen in MtA which in the alignment of Fig. 1 has a sequence characteristic of type II enzymes except for the first residue (G) which lacks a hydroxyl group.

MODELLING STUDIES

In the EcAII structure, residue A123, the equivalent of A176 in ScerAI, is located in a loop on the surface of a subunit and in the quaternary structure of the protein is involved in intersubunit interactions that bring the two intimate dimers (i.e. dimers A/C and B/D in EcAII labelling) together within a tetramer. The interactions occur between a loop segment in subunit A and its equivalent in subunit B, as well as between subunits C and D. The situation in the tetramer interface region has been analysed using the computer graphics program O (Jones & Kjeldgaard, 1994). The analysis was carried out for EcAII and was later repeated using all available crystallographic structures of bacterial asparaginases. For easier visualisation and numerical estimate of the agreement between the three-dimensional structures of the compared enzymes, the program ALIGN was used to superimpose the corresponding Cα atoms by the method of least squares (Satow *et al.*, 1986). The results of the superpositions were displayed and further analysed in O.

RESULTS AND DISCUSSION

The discussion below is mainly based on the three-dimensional structure of the periplasmic enzyme from *E. coli* (EcAII), the first L-asparaginase structure to be determined (Swain *et al.*, 1993), which serves as a structural reference in this field. In the EcAII structure, the A123 equivalent of ScerAI A176 is located far enough from the active site and the bound reaction product (Figs. 2, 3) to exclude any possibility of its direct involvement in the catalytic mechanism. It is located in a broad loop (115–125) which in the tertiary structure is adjacent to another, very long flap-like loop (9–31) which has been implicated as a mobile gate important for substrate's entry and possibly orientation in the active site (Swain *et al.*, 1993; Lubkowski *et al.*, 1994a). Figure 3 suggests that the 115–125 loop itself, through its proximity to the flap loop might also play a role in controlling substrate's access to the active site. However, a much more obvious and possibly more important role of A123 emerges from the analysis of the interface between the subunits forming a functional enzyme.

The EcAII A123 residue is found at the A/B (and C/D) intersubunit interface where the intimate dimers come into contact to form the tetramer (Figs. 2, 3, 4). It forms a short O...H–O hydrogen bond with the hydroxyl group of S122 in the complementary subunit. Due to the symmetric nature of the tetramer, two such hydrogen bonds exist at the A/B (and also C/D) junction (Figs. 3, 4). The O...O distances in these bonds are between 2.53 and 2.92 Å in EcAII (Table 1) and are, therefore, in the most preferred range for such interactions in proteins (Jeffrey & Saenger, 1991). We have identified the same mode of interaction at the tetrameric interface in all other structures of bacterial asparaginases (Table 1). The O...O distances in Table 1 range from 2.46 to 2.93 Å confirming the conclusion derived from the EcAII structure.

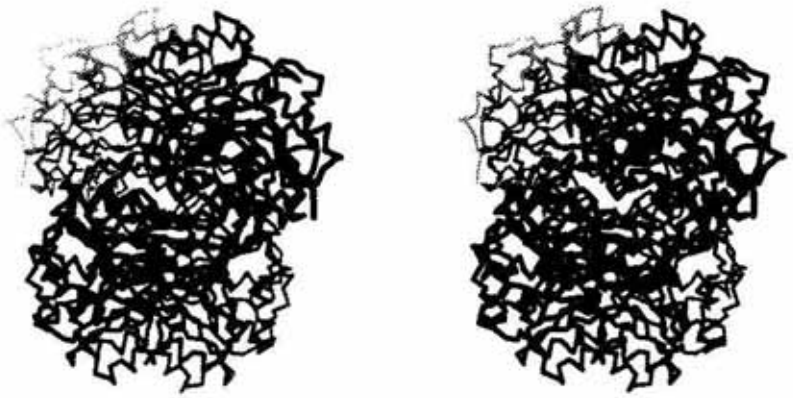


Figure 2. Stereoview of the 222-symmetric EcAII tetramer along the dyad relating dimer AC to BD (coordinates as in PDB entry 3ECA).

The subunits are shown in colour Cα traces (A red, B green, C gold, D blue). The active sites are formed within the intimately bound dimers (AC and BD). One (of the two) active sites within the AC dimer is shown, and one in BD. The active sites are marked by red spheres representing the residues involved in catalysis (and residing in the N-terminal domains of the A and B subunits) and by yellow spheres representing the bound aspartate ligands. The blue spheres mark the A123 residues at the A/B interface and show their spatial relationship to the active sites in the N-terminal domains of A and B.

The average O...O distance in the hydrogen bonds listed in Table 1 is 2.772(2) Å, nearly identical to the most preferred O...O distance (2.762 Å) observed in ice (Wells, 1990). We stress that this pattern of Ser Oγ(H)...O (Ala) interactions is present in all known crystallographic structures of L-asparaginase tetramers, i.e. is invariant at both truly symmetric and pseudosymmetric dimer-dimer interfaces. We therefore conclude that this pair of side-chain...main-chain hydrogen bond interactions seems to be very important, perhaps even critical, for the cohesion of the tetramer of type II L-asparaginases, where the crucial

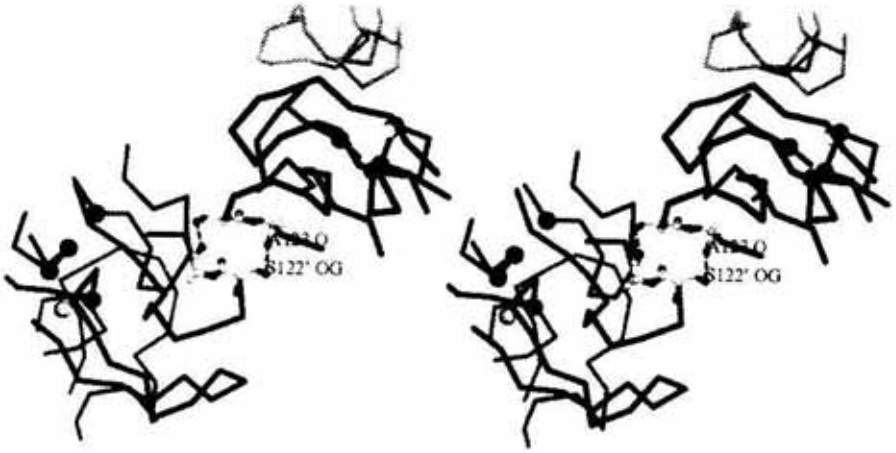


Figure 3. Stereo close-up view of the interface area between subunits A and B, with active sites marked by red and yellow spheres as in Fig. 2.

The red spheres mark the Cα atoms of residues T89, D90, K162, and T12. The yellow sphere marking the Cα atom of the Asp ligand is flanked by T89 and T12, the potential nucleophiles. The S122, A123 residues crucial for the formation of the A/B interface are shown in ball-and-stick fashion and the (123 Oγ)-H...O(122) hydrogen bonds across the interface are shown as broken lines. As in Fig. 2, the view is nearly along the two-fold axis between A and B. The active site formed in the N-terminal domain of B (green) is partly covered by the C-terminal domain of D (blue). An analogous relation exists between A (red) and C (gold).

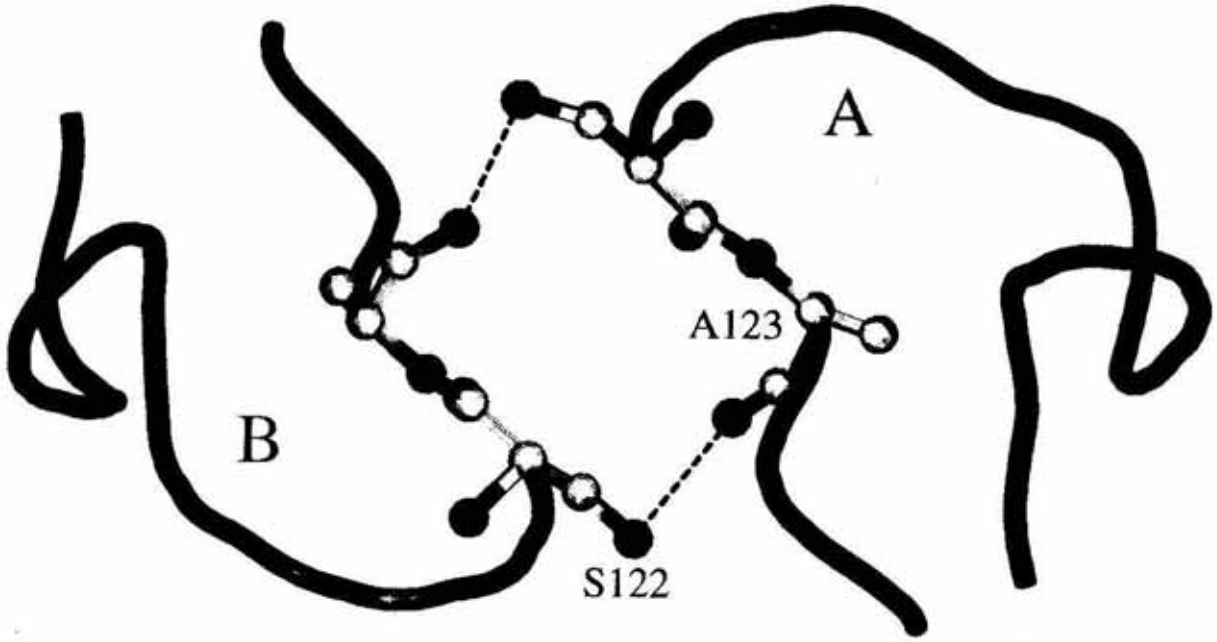


Figure 4. View of the 115–125 loops with the S122, A123 residues forming the interface between subunits A and B of the EcAII tetramer (coordinates as in PDB entry 3ECA).

The hydrogen bonds between the side chain of S122 and the main chain carbonyl of A123 across the interface are shown as broken lines. Due to the symmetric nature of the tetramer, two such bonds are formed around the interface two-fold axis. Note the peculiar shape of the 115–125 loops, which fold over the Cβ atoms of their A123 residues.

amino acids, SA, are present in an absolutely conserved SADG sequence (Fig. 1).

The reason why the character of the amino acid at position 123 is so important is explained by Fig. 5. The 115–125 loop (of which A123 is a part) has a peculiar shape, forming in its N-terminal segment a sort of a niche which covers residue 123. The niche is just

Table 1. Hydrogen bonding distances (Å) between S122 Oγ(H) and A123 O (EcAII numbering) at the tetramer interface in the structurally characterised bacterial L-asparaginases.

Primed residues are in the complementary subunit across the tetramer interface (e.g. A–B in EcAII).

		(Ser) Oγ...O (Ala′)	(Ser′) Oγ...O (Ala)
EcAII[a]	A/B	2.53	2.79
	C/D	2.63	2.92
ErA[b]	1/2	2.74	2.75
	3/4	2.79	2.75
PGA[c]	1/2	2.93	2.77
	3/4	2.82	2.92
PGA[d]	A/A′	2.85	2.85[e]
	B/B′	2.84	2.84[e]
AGA[f]		2.85	2.85[e]
WA[g]	A/A′	2.46	2.46[e]
	B/B′	2.78	2.78[e]

[a]PDB entry 3ECA; [b]Coordinates provided by Dr. Maria Miller; [c]PDB entry 3PGA; [d]Coordinates provided by Drs. Krzysztof Lewiński and Łukasz Lebioda for a high resolution (1.7 Å) structure of PGA; [e]Crystallographic symmetry; [f]PDB entry 1AGX; [g]Coordinates provided by Dr. Alexander Wlodawer.

big enough to accommodate the Cβ atom of A123. In addition, it is lined with hydrophilic groups (at least four O atoms and partial contributions from amide N atoms) which may position A123 through repulsive interactions. It is obvious from Fig. 5 that replacing A123 with any bigger residue would destroy the precise fit in this area and might lead to a disruption of the tetrameric interface. Possibly, only two A123X mutations could be tolerated. (i) X = G — there would be no steric hindrance between G123 and the tight niche. However, without the 123 Cβ, the cavity itself might collapse, disrupting the precise geometrical tuning in this crucial area. (ii) X = S — although bulkier than alanine, serine could use its OH group to form H-bond(s) with the hydrophilic groups on the niche wall to compensate for the unfavourable steric interactions.

Since the hydrogen bond formed across the tetramer interface is formed between A123 and S122, the latter residue should be similarly important for enzymatic competence of the protein. This is confirmed by the conserved SA pattern seen in sequences 4–11 in Fig. 1a. However, since S122 is located in a less constricted area (Fig. 4), its role will reside in the chemical nature of the side chain which must be capable of forming a hydrogen bond with the carbonyl O atom of residue 123′. One might hypothesise that S122T (or S175T in ScerAI) would be the only mutation tolerated at this site. We are currently testing this hypothesis by site-directed mutagenesis of ScerAI.

The above hypothesis about the role of the SA dipeptide in stabilising type II asparaginase tetramers through side-chain...main-chain O...O interactions appears to be violated by MtA, which on the grounds of sequence homology is a type II enzyme (*vide supra*), and yet in Fig. 1a has a GADAP sequence aligned with the SADGP region in EcAII. However, inspection of the MtA sequence reveals that immediately before the GADAP pattern (MtA residues 128–132) aligned with the EcAII SADGP sequence by PILEUP, there is a similar pattern, SADAP (MtA residues 123–127), which could perfectly fulfil our predictions. We therefore hypothesise that MtA has an unusual situation in the analysed region, in that there is a duplication of a similar pentapeptide that could cause gaps with respect to the EcAII sequence. It will be most interesting to check this prediction when the crystallographic structure of the *M. tuberculosis* enzyme becomes available.

To confirm the conservation of the structural features of the 115–125 region in asparaginase II tetramers, we have superimposed all available A/B type dimers, using the A/B dimer of EcAII as a target. Each super-

Figure 5. Stereoview of the 115–125 loop with S122, A123 shown as ball-and-stick, and the cavity-forming residues 116–120 shown in van der Waals representation (coordinates as in PDB entry 3ECA).

Note the scoop-like shape of this cavity with its concave surface lined with oxygen atoms (red), and the position of the A123 Cβ atom (shown as a grey van der Waals sphere) in it.

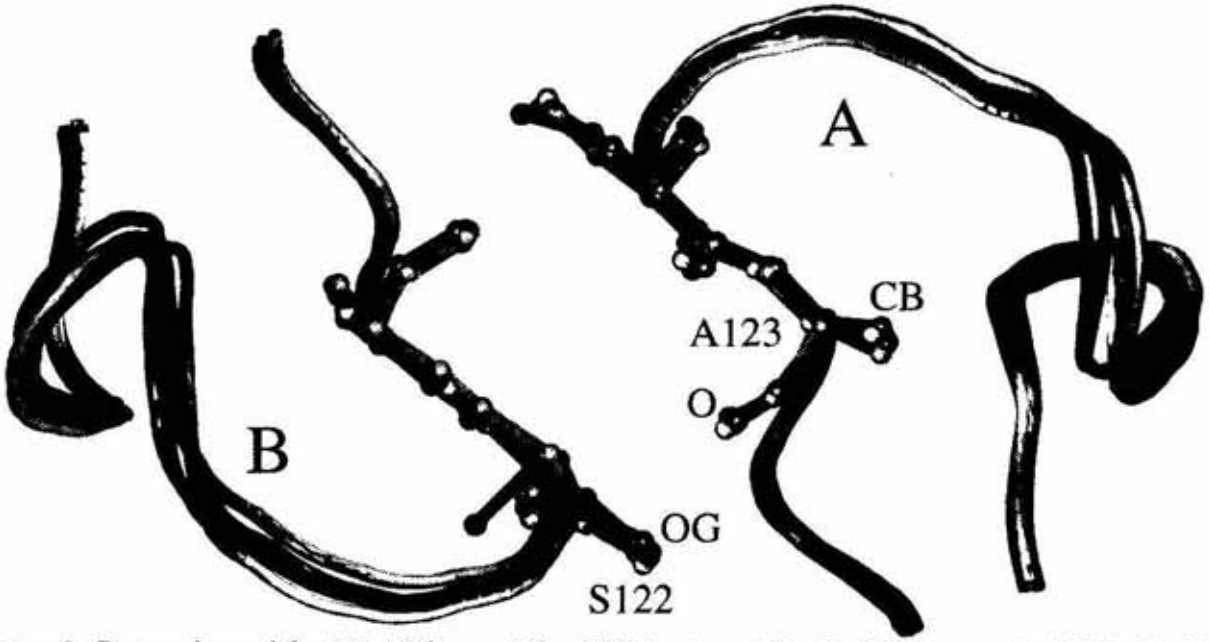


Figure 6. Comparison of the 115–125 loops at the A/B interface of the EcAII tetramer (as in Fig. 4) with corresponding segments in several structurally characterized L-asparaginases.

The structures illustrated in this Figure have been superimposed (ALIGN, Satow *et al.*, 1986) on the EcAII A/B dimer using all Cα atoms of the corresponding dimers. The loops are represented as tubes following the main chain atoms, and the crucial residues forming the interface link, S122 and A123 in EcaII sequence, are shown at atomic resolution. Note the close agreement of the shape of the loops and nearly identical positions of the atoms critical to the formation of the interface: S122 Oγ, A123 O and Cβ (EcAII numbering). Colour code: red EcAII subunits A/B, orange EcAII C/D, grey ErA 1/2, light blue ErA 3/4, green PGA 1/2, light green PGA 3/4, yellow AGA. Source of coordinates: EcAII, PDB entry 3ECA; PGA, PDB entry 3PGA; AGA, PDB entry 1AGX; ErA, Dr. Maria Miller.

position was based on all Cα atoms in the corresponding dimers. The calculations were done in ALIGN (Satow *et al.*, 1986) and included the following dimers: EcAII C/D (Swain *et al.*, 1993), ErA 1/2 and 3/4 (Miller *et al.*, 1993), PGA 1/2 and 3/4 (Lubkowski *et al.*, 1994a), WsA A/A′ and B/B′ (Lubkowski *et al.*, 1996), and AGA (Lubkowski *et al.*, 1994b). Figure 6 shows that the 115–125 loops (EcAII numbering) trace in all cases an essentially identical path, and that the position of A123 Cβ with respect to the rest of the loop is always the same. Also, the hydrogen-bonded groups, S122 (Oγ)-H and A123 O, are located in all dimers in identical positions. The agreement between the Cα positions of the S122 and A123 residues in the compared dimers is of the order of 0.4 Å, and usually better than the r.m.s. deviation for all Cα pairs in the superimposed dimers, which is of the order of 0.9 Å (Table 2).

Figure 1a reveals that the type I asparaginases differ significantly from the type II enzymes in the region corresponding to the SADGP sequence (EcAII 122–126). There is only one residue, the central aspartate, which appears with absolute conservation in the two families (even in the Archaebacterial *Methanococcus* sequence). We have analysed its structural role in detail, and conclude that it has an exclusively intrasubunit function stabilising the conformation of the peculiar 115–125 loop (EcAII numbering) depicted in Figs. 4 and 6. These interactions include hydrogen bonds between the carboxylic group of the aspartate and the main-chain amide groups of two residues at the entry to the 115–125 loop, M115 and R116 (EcAII numbering). In the EcAII structure, these interactions (illustrated in Fig. 7) are characterised by the following N...O distances: (M115) N...Oδ2 (D124) 2.81 Å, (M115) N...Oδ1 (D124) 3.16 Å, (R116) N...Oδ1 (D124) 2.98 Å.

Taken as a subset, the type I enzymes (and BlA) have a variable pattern in the EcAII region SADGP and homology is limited to the central tripeptide which, in this class (though based on a limited number of cases), appears to be (S/T)D(A/G). Even though there is a

Table 2. Cα–Cα deviations in A/B-type dimer superpositions on the EcAII A/B target

Superposition	r.m.s. dev. (Å)	Cα atoms	Cα–Cα dev. (Å)	
			S122 (A) S122 (B)	A123 (A) A123 (B)
EcAII C/D[a]	0.24	623	0.44	0.22
			0.26	0.19
ErA 1/2[b]	0.87	609	0.42	0.34
			0.70	0.37
ErA 3/4[b]	0.85	607	0.39	0.26
			0.59	0.33
PGA 1/2[c]	1.02	588	0.47	0.29
			0.32	0.25
PGA 3/4[c]	0.99	588	0.61	0.26
			0.50	0.30
AGA[d]	0.98	613	0.35	0.09
			0.37	0.25
WsA A/A'[e]	0.97	600	0.59	0.56
			0.49	0.36
WsA B/B'[e]	0.99	600	0.56	0.59
			0.52	0.44

[a]PDB entry 3ECA; [b]Coordinates provided by Dr. Maria Miller; [c]PDB entry 3PGA; [d]PDB entry 1AGX; [e]Coordinates provided by Dr. Alexander Włodawer

hydroxyl-bearing residue in this pattern, it is unknown if it could play an intersubunit cohesive role similar to that played by S122 in EcAII and other type II enzymes. The doubt stems from the fact that, in the absence of a more extensive sequence similarity, it is questionable whether a carbonyl group analogous to A123 C = O would be correctly positioned across the interface to accept a hydrogen bond. In addition, the S/T residue is now adjacent to the absolutely conserved aspartate, which is quite certain to be primarily responsible for stabilising the tertiary structure in its region. On the other hand, the (S/T)D(A/G) tripeptide occurring in the three examples of type I asparaginase could be responsible for a different mode of intersubunit interactions. We, therefore, hypothesise that the arrangement of intimate dimers into a tetramer in type I enzymes might be different, if such tetramers exist at all. The molecular mass of EcAI, estimated by gel filtration (64±5 kDa) has in fact previously been taken to indicate that this enzyme may be a dimer in its native state (Jerlstroem *et al.*, 1989).

CONCLUSIONS

Sequence alignments for the family of microbial L-asparaginases, with the two *E. coli* isoenzymes (EcAI and EcAII) as standards for type I ("cytosolic") and type II ("periplasmic") asparaginases, allow classification of the currently known amino-acid sequences into two broad categories. Based on these alignments, the type I group includes

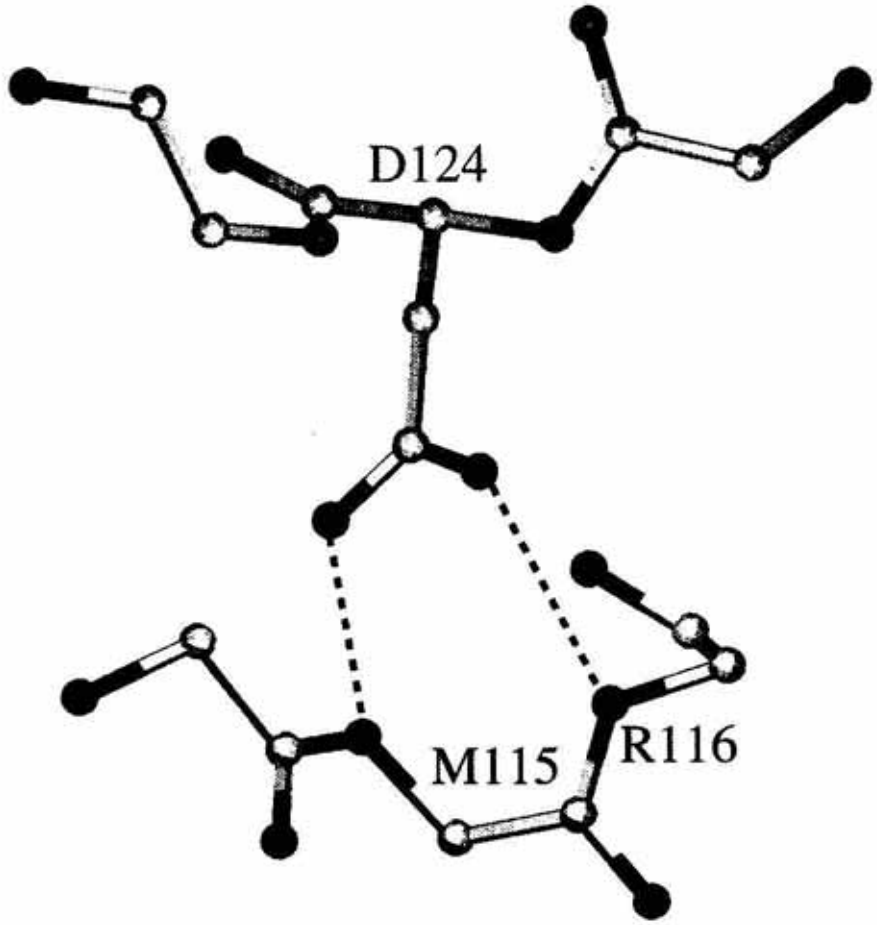


Figure 7. Hydrogen bonding interactions between D124 and the residues at the beginning of the 115–125 loop (EcAII model and sequence, PDB entry 3ECA).

Except for D124, which is shown complete with the side chain, for the rest of the illustration (residues 114–116 and 123–125) only main-chain atoms are shown.

asparaginases from *B. subtilis* and *Methanococcus jannaschii*, while the L-asparaginase from *Mycobacterium tuberculosis* belongs to class II. Only *B. licheniformis* asparaginase is uncertain in its placement; it has closest similarity to MtA. The two *S. cerevisiae* enzymes, intracellular (ScerAI) and cell-wall (ScerAII), are both clearly class II, in line with our suggestion that the asparaginase genes were independently duplicated in prokaryotes and eukaryotes after the divergence of these lineages. In the case of *Haemophilus influenzae*, we have a type II enzyme very close in sequence to EcAII. However, *H. influenzae* does not appear to have an equivalent of EcAI; searching of the complete *H. influenzae* genome (http://www.tigr.org/tdb/mdb/hidb/hidb.html) using the EcAI peptide sequence reveals the previously identified sequence (HI0754) as the only significant match. Conversely, *M. jannaschii* does not have an EcAII homologue; only the type I sequence analysed above (MJ0020) is identified by searching with the EcAII sequence (http://www.tigr.org/tdb/mdb/mjdb/mjdb.html).

The A176V mutation (which completely kills enzymatic activity) occurs in ScerAI in a region, SADGP, that is nearly absolutely conserved in type II asparaginases. This sequence is in a loop region of very peculiar shape, located on the surface of a subunit in all crystallographically characterised (type II) L-asparaginases. In the tetramers (which are found in all known structures of type II asparaginases), two residues from this region, the SA dipeptide, are involved in a pair of side-chain...main-chain Oγ(H)...O interactions that seem to be crucial for the formation of tetramers from intimate (catalytic-centre competent) dimers. Curiously, in the MtA sequence, the SADGP pattern is repeated in two consecutive, closely related variants (SADAPGADGP). From the sequence and structural conservation in type II asparaginases of the region in which the ScerAI A176V mutation occurs, as well as from the constancy with which this area participates

in the tetrameric interface, we infer that this mutation abolishes enzymatic activity by disrupting tetramer formation. We also predict that similar consequences would be observed by mutating A123 in EcAII or the corresponding alanine in other type II asparaginases. The lack of a more extensive homology between type II and type I asparaginases in the discussed region may indicate that, if they exist at all, tetramers of type I enzymes could have a different intersubunit interface. There is an absolutely conserved residue in the centre of the SADGP pentapeptide, the aspartate, in both type II and type I enzymes. The present analysis shows that this residue is important for the conformation of a peculiar loop at the surface of an asparaginase monomer. In summary, asparaginases of type I and type II may share similar subunit fold and intimate dimer organisation but not necessarily the same quaternary structure.

We wish to thank Drs. Alexander Włodawer, Maria Miller, Krzysztof Lewiński and Łukasz Lebioda for making the coordinates of several bacterial asparaginases available to us prior to their release through the Protein Data Bank.

REFERENCES

Aiba, H., Baba, T., Fujita, K., Hayashi, K., Inada, T., Isono, K., Itoh, T., Kasai, H., Kashimoto, K., Kimura, S., Kitakawa, M., Kitagawa, M., Makino, K., Miki, T., Mizobuchi, K., Mori, H., Mori, T., Motomura, K., Nakade, S., Nakamura, Y., Nashimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampei, G., Seki, Y., Sivasundaram, S., Tagami, H., Takeda, J., Takemoto, K., Takeuchi, Y., Wada, C., Yamamoto, Y. & Horiuchi, T. (1996) A 570-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 28.0–40.1 min region on the linkage map. *DNA Res.* **3**, 363–377.

Bonthron, D.T. (1990) L-Asparaginase II of *Escherichia coli* K-12. Cloning, mapping and sequencing of the *ansB* gene. *Gene* **91**, 101–105.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D. *et al.* (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073.

Devereux, J., Haeberli, P. & Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387–395.

van-Dijl, J.M., de-Jong, A., Smith, H., Bron, S. & Venema, G. (1991) Lack of specific hybridization between the *lep* genes of *Salmonella typhimurium* and *Bacillus licheniformis*. *FEMS Microbiol. Lett.* **65**, 345–351.

Dunlop, P.C., Meyer, G.M., Ban, D. & Roon, R.J. (1978) Characterization of two forms of asparaginase in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **253**, 1297–1304.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M. *et al.* (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512.

Gallagher, M.P., Marshall, R.D. & Wilson, R. (1989) Asparaginase as a drug for treatment of acute lymphoblastic leukaemia. *Essays Biochem.* **24**, 1–40.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M. *et al.* (1996) Life with 6000 genes. *Science* **274**, 546, 563–7.

Jakob, C.G., Lewinski, K., LaCount, M.W., Roberts, J. & Lebioda, L. (1997) Ion binding induces closed conformation in *Pseudomonas* 7 Å glutaminase-asparaginase (pga): Crystal structure of the PGA-SO_4^{2-}-NH_4^+ complex at 1.7 Å resolution. *Biochemistry* **36**, 923–931.

Jeffrey, G.A. & Saenger, W. (1991) *Hydrogen Bonding in Biological Macromolecules*. Springer-Verlag, Berlin.

Jerlstroem, P.G., Bezjak, D.A., Jennings, M.P. & Beacham, I.R. (1989) Structure and expression in *Escherichia coli* K-12 of the L-asparaginase I-encoding *ansA* gene and its flanking regions. *Gene* **78**, 37–46.

Jones, T.A. & Kjeldgaard, M. (1994) *O — The Manual*. Uppsala University, Uppsala, Sweden.

Kim, K.-W., Kamerud, J.Q., Livingston, D.M. & Roon, R.J. (1988) Asparaginase II of *Saccharomyces cerevisiae*. Characterization of the *ASP3* gene. *J. Biol. Chem.* **263**, 11948–11953.

Lubkowski, J., Włodawer, A., Ammon, H.L., Copeland, T.D. & Swain, A.L. (1994a) Structural characterization of *Pseudomonas* 7 Å glutaminase-asparaginase. *Biochemistry* **33**, 10257–10265.

Lubkowski, J., Włodawer, A., Housset, D., Weber, I.T., Ammon, H.L., Murphy, K.C. & Swain, A.L. (1994b) Refined crystal structure of *Acinetobacter glutaminasificans* glutaminase-asparaginase. *Acta Crystallogr.* **D50**, 826–832.

Lubkowski, J., Palm, G.J., Gilliland, G.L., Derst, Ch., Roehm, K.-H. & Włodawer, A. (1996) Crystal structure and amino acid sequence of *Wolinella succinogenes* L-asparaginase. *Eur. J. Biochem.* **241**, 201–207.

Miller, M., Rao, J.K.M., Wlodawer, A. & Gribskov, M. (1993) A left-handed crossover involved in amidohydrolase catalysis. Crystal structure of *Erwinia chrysanthemi* L-asparaginase with bound L-aspartate. *FEBS Lett.* **328**, 275–279.

Minton, N.P., Bullman, H.M.S., Scawen, M.D., Atkinson, T. & Gilbert, H.J. (1986) Nucleotide sequence of the *Erwinia chrysanthemi* NCPPB 1066 L-asparaginase gene. *Gene* **46**, 25–35.

Palm, G.J., Lubkowski, J., Derst, Ch., Schleper, S., Roehm, K.-H. & Wlodawer, A. (1996) A covalently bound catalytic intermediate in *Escherichia coli* asparaginase: Crystal structure of a Thr-89-Val mutant. *FEBS Lett.* **390**, 211–216.

Satow, Y., Cohen, G.H., Padlan, E.A. & Davies, D.R. (1986) Phosphocholine binding immunoglobulin Fab McPC603: An X-ray diffraction study at 2.7 Å. *J. Mol. Biol.* **190**, 593–604.

Sinclair, K., Warner, J.P. & Bonthron, D.T. (1994) The *ASP1* gene of *Saccharomyces cerevisiae*, encoding the intracellular isozyme of L-asparaginase. *Gene* **144**, 37–43.

Sun, D.X. & Setlow, P. (1991) Cloning, nucleotide sequence, and expression of the *Bacillus subtilis ans* operon, which codes for L-asparaginase and L-aspartase. *J. Bacteriol.* **173**, 3831–3845.

Swain, A.L., Jaskólski, M., Housset, D., Rao, M.J.K. & Włodawer, A. (1993) Crystal structure of *Escherichia coli* L-asparaginase, an enzyme used in cancer therapy. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1474–1478.

Tanaka, S., Robinson, E.A., Appella, E., Miller, M., Ammon, H.L., Roberts, J., Weber, I.T. & Włodawer, A. (1988) Structures of amidohydrolases. Amino-acid sequence of a glutaminase-asparaginase from *Acinetobacter glutaminasificans* and preliminary crystallographic data for an asparaginase from *Erwinia chrysanthemi*. *J. Biol. Chem.* **263**, 8583–8591.

Wells, A.F. (1990) *Structural Inorganic Chemistry*. Oxford University Press.