



Structural and biophysical aspects of L-asparaginases: a growing family with amazing diversity

Joanna I. Loch^{a*} and Mariusz Jaskolski^{b,c*}

Received 19 March 2021

Accepted 9 June 2021

Edited by Z.-J. Liu, Chinese Academy of Sciences, China

As a birthday tribute, this review is dedicated to Dr Alexander Wlodawer, who pioneered the field of structural biology of antileukemic L-asparaginases.

Keywords: leukemia; L-asparaginases; amidohydrolases; catalytic mechanism; active site; nucleophiles.

Supporting information: this article has supporting information at www.iucrj.org

^aDepartment of Crystal Chemistry and Crystal Physics, Faculty of Chemistry, Jagiellonian University, Cracow, Poland,

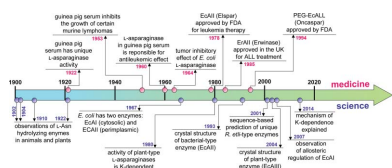
^bDepartment of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland, and ^cCenter for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland.

*Correspondence e-mail: loch@chemia.uj.edu.pl, mariuszj@amu.edu.pl

L-Asparaginases have remained an intriguing research topic since their discovery ~120 years ago, especially after their introduction in the 1960s as very efficient antileukemic drugs. In addition to bacterial asparaginases, which are still used to treat childhood leukemia, enzymes of plant and mammalian origin are now also known. They have all been structurally characterized by crystallography, in some cases at outstanding resolution. The structural data have also shed light on the mechanistic details of these deceptively simple enzymes. Yet, despite all this progress, no better therapeutic agents have been found to beat bacterial asparaginases. However, a new option might arise with the discovery of yet another type of asparaginase, those from symbiotic nitrogen-fixing Rhizobia, and with progress in the protein engineering of enzymes with desired properties. This review surveys the field of structural biology of L-asparaginases, focusing on the mechanistic aspects of the well established types and speculating about the potential of the new members of this amazingly diversified family.

1. Asparaginases: historical facts and structure-based classification

Asparaginases (EC 3.5.1.1) are amidohydrolases that catalyze hydrolysis at the side chain of L-asparagine to release L-aspartate and ammonia. This humble chemical reaction contrasts with the picturesque history of discovery of this amazingly diverse group of enzymes, which starts at the beginning of the 20th century with the observation of the existence of amide-splitting enzymes in suspensions of animal and plant tissues (Gonnermann, 1902). However, these early observations did not refer directly to the hydrolysis of L-asparagine (Geddes & Hunter, 1928; Fig. 1). The presence of enzymes hydrolyzing L-Asn (and L-Gln) was noticed in suspensions of beef liver (Lang, 1904), and horse and pig tissues (Furth & Friedmann, 1910). The serendipitous detection of the strong antileukemic properties of guinea pig serum (Kidd, 1953) was attributed by Broome to L-asparaginase activity (Broome, 1961), which had already been reported in guinea pig serum by Clementi (1922). The large-scale use of guinea pig serum in leukemia treatment was not possible, but in 1964 a similar tumor inhibitory effect was reported for *Escherichia coli* asparaginase (Mashburn & Wriston, 1964). These discoveries opened an era of intensive biochemical and structural studies of L-asparaginases, as L-Asn is vital for the survival of some cancer cells (Luo *et al.*, 2018). Most malignant lymphoblasts are unable to produce endogenous L-Asn, and in consequence these cancer cells are solely dependent on



OPEN ACCESS

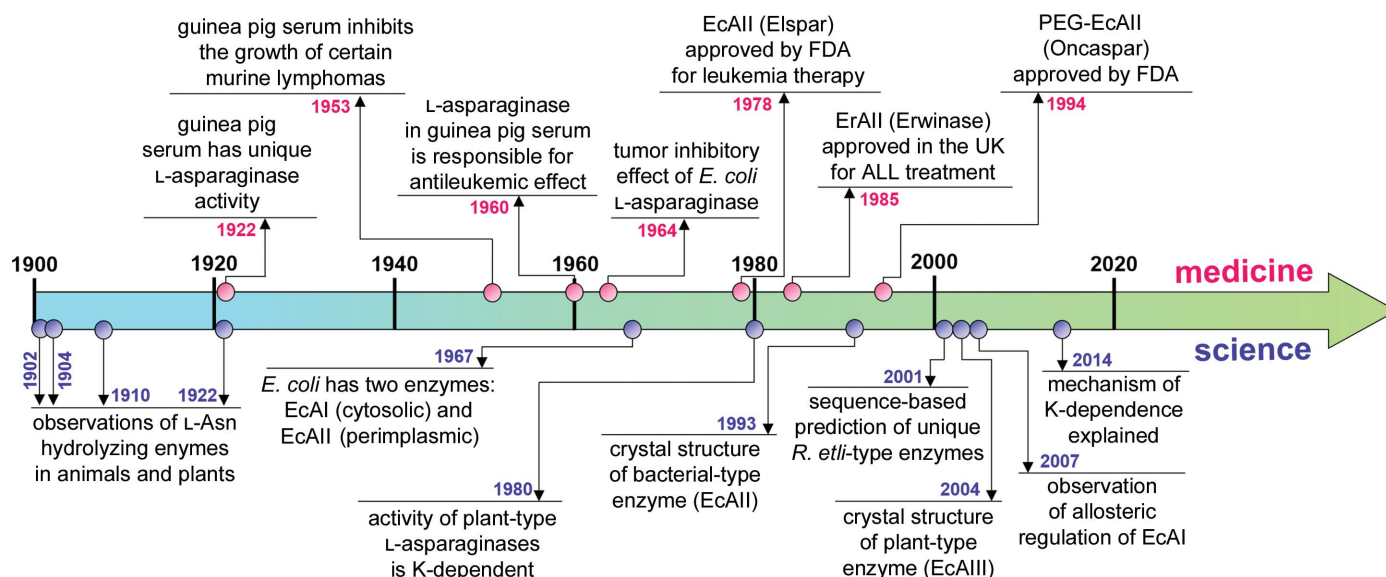


Figure 1

Overview of selected events in the history of L-asparaginases in science and medicine. Scientific discoveries intertwine with medical applications of L-asparaginases. References to the historical facts presented are included in the text.

extracellular L-Asn for survival (Abaji & Krajinovic, 2019). Therapeutic L-asparaginases work by depleting the pool of circulating L-Asn, leading to cancer cell starvation and death (Boyse *et al.*, 1967; Prager & Bachynsky, 1968; Radadiya *et al.*, 2020). The requirement for a low K_m (at least micromolar) of this unusual enzymatic cancer drug stems from the fact that the intravenous L-asparaginase must be able to clear the circulating pools of L-Asn, which are not very high (about 50 μM) (Ollenschläger *et al.*, 1988; Beckett & Gervais, 2019). The first therapeutic L-asparaginase (Elspar, EcAII) was approved by the US Food and Drug Administration in 1978. An improved, PEGylated version of EcAII (Oncaspar) with a prolonged circulation time and lower immunogenicity was approved in 1994 (Dinndorf *et al.*, 2007; Nunes *et al.*, 2020). For patients developing silent inactivation or strong allergic reactions to EcAII, an alternative bacterial enzyme derived from *Erwinia chrysanthemi*¹ (Erwinase, ErAII) was authorized in the UK in 1985 and in the USA in 2011 (Gervais *et al.*, 2013; Porton Biopharma, 1985).

In general, L-asparaginases are divided into three structural classes: bacterial-type, plant-type (Michalska *et al.*, 2006) and *Rhizobium etli*-type (Borek & Jaskólski, 2001) (Fig. 2). However, this classification, which is based on the sources of the first enzymes that were discovered, may be misleading (see Section 2). In fact, enzymes from all structural classes are distributed among different taxonomic groups, including microorganisms (for example bacteria and fungi), higher plants and animals (Borek *et al.*, 2004; Borek & Jaskólski, 2001). The bacterial-type enzymes can further be divided into two subclasses. This classification was initially introduced for the *E. coli* enzymes and was subsequently propagated to other

proteins of similar architecture. *E. coli* possesses two enzymes in this class, encoded by the *ansA* and *ansB* genes. The protein products of these genes are termed EcAI (type 1, cytosolic, low-affinity enzyme) and EcAII (type 2, secreted, high-affinity enzyme), respectively. EcAI is expressed constitutively in the cytoplasm, while EcAII is expressed under anaerobic conditions in the periplasmic space (Campbell *et al.*, 1967). The *ansB* gene is not organized in any operon, while its expression is regulated by cyclic AMP receptor and fumarate and nitrate reductase regulation (FNR) proteins (Jennings & Beacham, 1990). Interestingly, the eukaryotic yeast *Saccharomyces cerevisiae* also has two asparaginases: a cytoplasmic asparaginase and a secreted asparaginase (Dunlop & Roon, 1975; Dunlop *et al.*, 1978). The first pieces of structural information, describing single crystals and the low-resolution structures of bacterial-type asparaginases (or their close glutaminase-asparaginase, G-A, homologs), were published between 1969 and 1988 (North *et al.*, 1969; Epp *et al.*, 1971; Lee & Yang, 1973; Wlodawer *et al.*, 1977; Yonei *et al.*, 1977; Ammon *et al.*, 1985, 1988; Tanaka *et al.*, 1988;). The first high-resolution crystal structure of EcAII (Elspar) with serendipitously bound L-aspartate was published in 1993² (Swain *et al.*, 1993). It revealed the location of the active site and suggested a simple catalytic apparatus consisting of a Thr–Lys–Asp triad, reminiscent of the Ser–His–Asp triad of classic serine proteases (Seemüller *et al.*, 1996).

In the course of studies of L-asparagine (Atkins *et al.*, 1975) and nitrogen metabolism in symbiotic nitrogen-fixing legumes (Lees & Blakeney, 1970) such as lupins (Lea *et al.*, 1978), it was discovered that plants possess potassium-dependent and

¹ Currently, *Erwinia chrysanthemi* is classified as *Dickeya dadantii*. However, in view of the widespread use of the historical name, we will also use it in our article.

² It should be noted that the amino-acid sequence of Elspar (PDB entry 3eca) has four differences with respect to the sequence of EcAII from the UniProt database at positions 27 (V/A), 64 (N/D), 252 (S/T) and 263 (T/N). This is the reason why the sequence of the first EcAII model in the PDB is different from other depositions for the same protein.

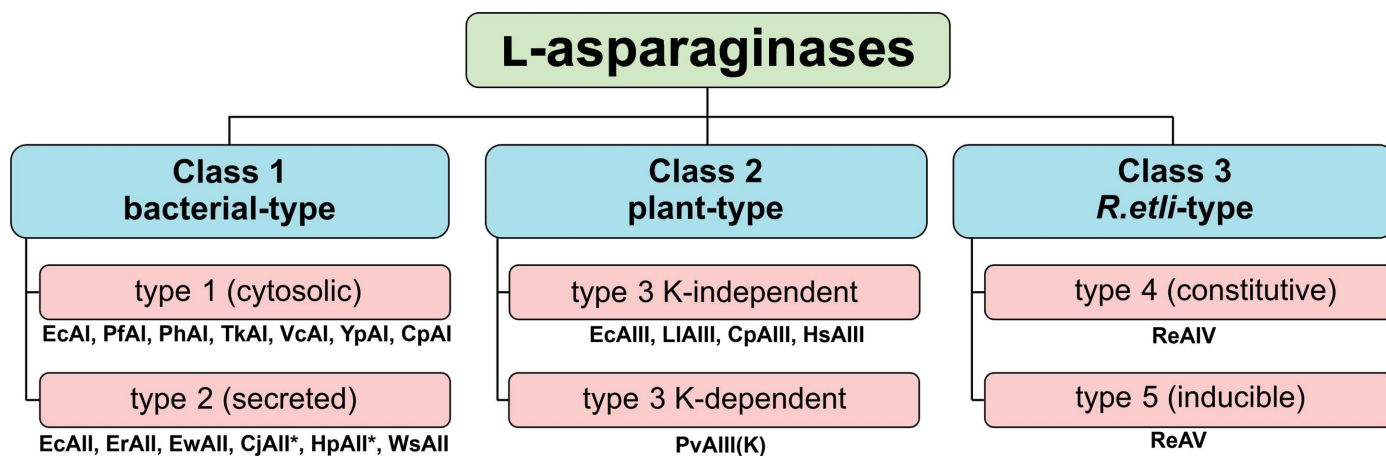


Figure 2

Structure-based classification of L-asparaginases. Historically, L-asparaginases were named according to the source organism of the first enzymes that were discovered. The new nomenclature divides L-asparaginases into three classes. Cytosolic type 1 enzymes are expressed constitutively, while the expression of type 2 enzymes, which are secreted to the periplasm, is induced under anaerobic conditions in *E. coli*. In the *R. etli*-type Class 3, constitutive type 4 enzymes are thermostable, while type 5 enzymes are considered thermolabile and their expression is induced by the presence of L-Asn. Examples of enzymes are listed below the boxes. The organism name abbreviations are as follows: Ec, *Escherichia coli*; Pf, *Pyrococcus furiosus*; Ph, *Pyrococcus horikoshii*; Tk, *Thermococcus kodakarensis*; Vc, *Vibrio cholerae*; Yp, *Yersinia pestis*; Cp, *Cavia porcellus*; Er, *Erwinia chrysanthemi*; Ew, *Erwinia carotovora*; Cj, *Campylobacter jejuni*; Hp, *Helicobacter pylori*; Ws, *Wolinnella succinogenes*; Ll, *Lupinus luteus*; Pv, *Phaseolus vulgaris*; Hs, *Homo sapiens*; Re, *Rhizobium etli*. Examples of PDB codes are summarized in Supplementary Table S1. An asterisk denotes enzymes that are annotated as products of an *ansA* gene; however, the architecture of the active site and substrate affinity suggest classification as Class 1 type 2 enzymes.

potassium-independent asparaginases (Sodek *et al.*, 1980). However, the molecular mechanism of their potassium dependence was only explained in 2014 on the basis of crystallographic studies (Bejger *et al.*, 2014). The cDNA encoding lupin L-asparaginase revealed that this enzyme differs from bacterial-type L-asparaginases but is similar to human aspartylglucosaminidase (Lough *et al.*, 1992). A few years later, human aspartylglucosaminidase (also termed, incorrectly, glycosylasparaginase), together with other amido-hydrolases, was classified as an N-terminal nucleophile hydrolase (Ntn-hydrolase; Brannigan *et al.*, 1995). Biochemical studies suggested that asparaginases from higher plants also belong to the Ntn-hydrolases (Oinonen & Rouvinen, 2000; Hejazi *et al.*, 2002). This hypothesis was confirmed by the first crystal structure to be determined of a plant-type L-asparaginase, that of LIA from yellow lupin (*Lupinus luteus*; Michalska *et al.*, 2006). Surprisingly, a gene (*yibK*) encoding a plant-type asparaginase was also found in the *E. coli* genome, and the protein product of this gene, with 70% sequence similarity to LIA, was termed EcAIII (Borek & Jaskólski, 2000). The crystal structure of EcAIII confirmed its classification as an Ntn-hydrolase (Borek & Jaskólski, 2000; Prah *et al.*, 2004; Michalska *et al.*, 2005). Plant-type (or type 3) L-asparaginases are also capable of hydrolyzing β -aspartyl peptides and this isoaspartyl aminopeptidase activity can even dominate (Borek *et al.*, 2004).

Studies of *R. etli*, a nitrogen-fixing bacterium that associates symbiotically with the legume *Phaseolus vulgaris*, revealed that it possesses two types of asparaginase: a thermostable asparaginase and a thermolabile asparaginase (Huerta-Zepeda *et al.*, 1997). Most intriguingly, a comprehensive phylogenetic analysis of L-asparaginase sequences showed that the *R. etli* enzymes belong to a completely different class

of proteins, significantly distant from the bacterial-type and plant-type enzymes (Borek & Jaskólski, 2001). For this reason, a new class of asparaginases was established early on, but its structural confirmation has only just begun to emerge (Loch *et al.*, unpublished work).

2. L-Asparaginase nomenclature issues

Historically, enzymes with asparaginase activity were named according to the source organism, with the possible addition of a Roman numeral (I, II *etc.*) when more than one enzyme had a similar origin. Thus, we have the *E. coli* cytoplasmic EcAI and periplasmic EcAII asparaginases. The first organism in which a representative of a new structural class was discovered was then used to distinguish a family or type of asparaginases. Thus, the above two *E. coli* enzymes are bacterial-type. A terminological complication arose when it was realized that enzymes typical of one kingdom (for example plant-type) are also found in other kingdoms, as illustrated by EcAIII. In want of a better system, such proteins were still classified with the original type; this is why EcAIII is called plant-type. This has led to a Babel of nomenclature because enzymes of the three canonical types, *i.e.* bacterial-type, plant-type and *R. etli*-type, are widely found throughout several kingdoms of life. To rectify the situation, Schultz da Silva *et al.* (2021) have recently proposed that the class name should be detached from the biological taxonomy, and to use Class 1, Class 2 and Class 3 with reference to the three canonical classes above. We will use this reasonable proposal in our review, hoping that it will be universally adopted in the field. However, we do not feel that this change of global terminology should be transmitted down to concrete enzymes, which have historically sanctioned names that are firmly established in the literature. Thus, we

will say that EcAI belongs to type 1 of the Class 1 asparaginases. Analogously, EcAII is type 2 in the same Class 1, while EcAIII is a type 3 asparaginase in structural Class 2. The *R. etli* enzymes are tentatively allocated as type 4 (ReAIV) and type 5 (ReAV) in Class 3. This classification scheme is reflected in Fig. 2.

3. Class 1 asparaginases (formerly known as bacterial-type asparaginases)

3.1. 3D structure of the enzyme and its oligomeric state

To date, in addition to the *E. coli* enzymes, of which EcAII has an especially numerous representation in the PDB, the crystal structures of several other type 2 and type 1 (Class 1) L-asparaginases have been determined (Supplementary Table S1). A typical protomer of type 2 enzymes, for example EcAII, is ~ 330 residues long and comprises 14 β -strands and eight

α -helices that fold into two distinct, N- and C-terminal, domains (Aghaiypour *et al.*, 2001b) connected by an ~ 20 -residue linker (Fig. 3). The overall fold of protomers of type 1 (Class 1) asparaginases, for example EcAI, is very similar. The only available structure of a mammalian homolog of bacterial type 1 L-asparaginase is that of the guinea pig enzyme CpAI (PDB entry 4r8k). However, this guinea pig homolog (as well as the human counterpart, HsAI) has an additional ~ 200 -residue C-terminal domain which probably contains several ankyrin repeats and is not visible in the crystal structure of CpAI (Schalk *et al.*, 2014). The mammalian enzyme has several other minor changes in comparison to its bacterial counterparts, which include the length and conformation of the linker, the structure of the active-site flexible loop and the presence of an additional loop with a small helix in the N-terminal catalytic domain (Supplementary Fig. S1).

A bacterial type 2 (Class 1) L-asparaginase is a D_2 -symmetric homotetramer that can be described as a dimer of

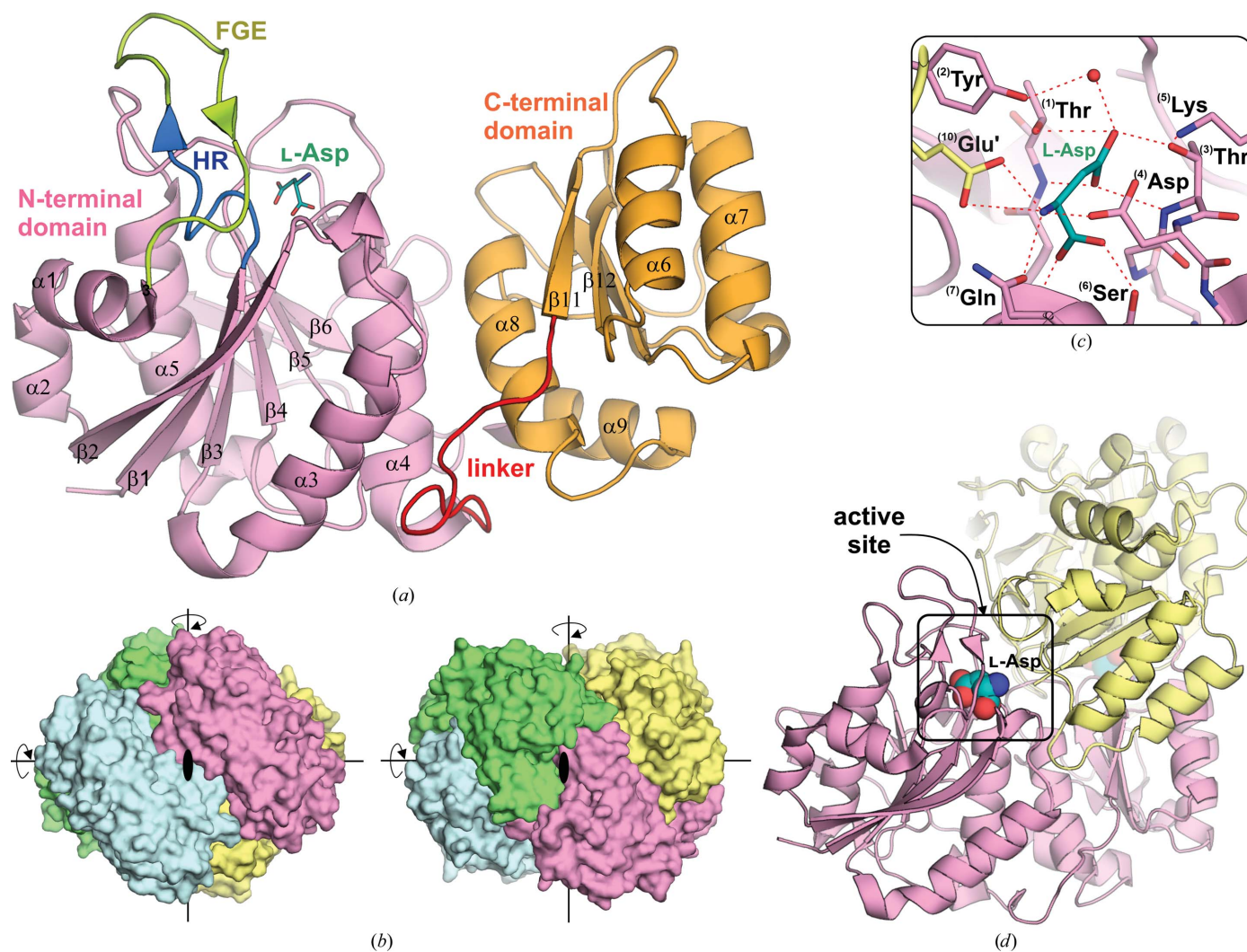


Figure 3 Architecture of bacterial type 2 asparaginases. (a) Protomer of EcAII (PDB entry 3eca) with the N-terminal domain in pink, the C-terminal domain in orange, the linker in red, the hinge region (HR) in blue and the flexible gating element (FGE) in green. (b) The EcAII homotetramer (a dimer of two intimate dimers) has D_2 (222) symmetry with intimate dimers A/B (pink/yellow) and C/D (green/blue). (c) Detailed view of the EcAII active site with L-Asp bound. (d) Location of the active site in the A/B dimer.

two intimate dimers (Swain *et al.*, 1993; Cerofolini *et al.*, 2019; Fig. 3). The bacterial type 1 (Class 1) enzyme EcAI is associated into a similar tetrameric assembly in the crystal packing, but it is not clear whether the protein exists as a tetramer in solution. The same doubt was expressed for the archaeal type 1 enzyme PhAI (from *Pyrococcus horikoshii*), which seems to function as a dimer rather than a tetramer (Yun *et al.*, 2007). Type 1 asparaginases from some bacterial organisms, for example *Thermus thermophilus*, may act as hexamers (trimers of dimers; Pritsa & Kyriakidis, 2001), but there is no structural evidence to support this hypothesis.

Sequence comparison of microbial L-asparaginases confirms that intracellular and extracellular enzymes evolved in prokaryotes and eukaryotes independently (Bonthon & Jaskólski, 1997). A conservative A176V mutation in the intracellular *Saccharomyces cerevisiae* L-asparaginase ScerAI completely abolished the enzyme activity. Sequence and structural comparisons with type 2 bacterial asparaginases show that the mutated residue is in a highly conserved region and plays a vital role in the cohesion of functional tetramers. The fact that type 1 asparaginases show less conservation in this region again suggests that they may have a different quaternary structure while adopting the same subunit fold and intimate dimer architecture.

3.2. Allosteric regulation

In the crystal structure of EcAI (PDB entry 2p2n), substrate molecules (L-Asn) were found not only in the active site but also in another region of the protein, strongly suggesting that L-Asn can act as an allosteric modulator (Yun *et al.*, 2007). The EcAI dimer possesses two allosteric sites located close to the dimer interface near Arg240 (Fig. 4). Signal transmission from the allosteric site to the active site probably involves subtle structural rearrangements at the dimer interface and relocation of Gln118 to facilitate the positioning of the catalytic water molecules (Yun *et al.*, 2007). EcAI is not the only allosteric asparaginase. The type 1 enzyme from *S. cerevisiae* is also allosteric (Costa *et al.*, 2016); however, the crystal structure of this yeast enzyme has not been determined to date. On the other hand, not all type 1 enzymes are allosteric. For

instance, *P. horikoshii* PhAI might not be allosteric because it seems to permanently be in the catalytically active state (Yun *et al.*, 2007). Allosteric properties were also not confirmed for *P. furiosus* PfAI, as the quaternary structure of PfAI is almost identical in the ligand-bound and unbound states (Tomar *et al.*, 2014; Bansal *et al.*, 2012). The mammalian CpAI enzyme is not allosteric as it follows classic Michaelis–Menten kinetics, while its human counterpart HsAI seems to be allosteric in enzymatic assays, although no crystal structure is available (Schalk *et al.*, 2014).

3.3. Domains and linker

In Class 1 enzymes, the N-terminal domain has all of the elements necessary for catalysis: the active site with the catalytic threonine residues, a loop acting as a flexible gating element (FGE) and the hinge region (HR) (Fig. 3). In the bacterial enzymes, the FGE does not comprise well defined secondary-structure elements and in the absence of a substrate it is usually disordered (Papageorgiou *et al.*, 2008). In contrast, in the mammalian CpAI the FGE is more rigid as it contains a short antiparallel β -sheet formed as a β -hairpin loop (Schalk *et al.*, 2014; Supplementary Fig. S1). In EcAII, the FGE includes residues 18–31, although some authors also consider the downstream helix to be part of the FGE (Nguyen *et al.*, 2017). The conserved HR region is formed by a glycine-rich octapeptide: residues 10–17 in EcAII (Lubkowski *et al.*, 2020).

To build the complete active site, the substrate-binding pocket from one protomer must be complemented by residues from the C-terminal domain of the intimate partner, which are necessary for the stabilization of the substrate in the active site. In contrast to the highly conserved sequences of the N-terminal domains, the sequences of the interdomain linker (Fig. 3) are highly variable. The length of the linker is correlated with the evolutionary development, and shorter linkers are found in archaeal proteins (Tomar *et al.*, 2014). As an illustration, the linker in the type 1 asparaginase from *P. horikoshii* (PhAI) is highly strained and seven residues shorter in comparison with that in EcAII (Yao *et al.*, 2005). Interestingly, the archaeal L-asparaginase from *P. furiosus* (PfAI) obtained from a linker-less recombinant expression of

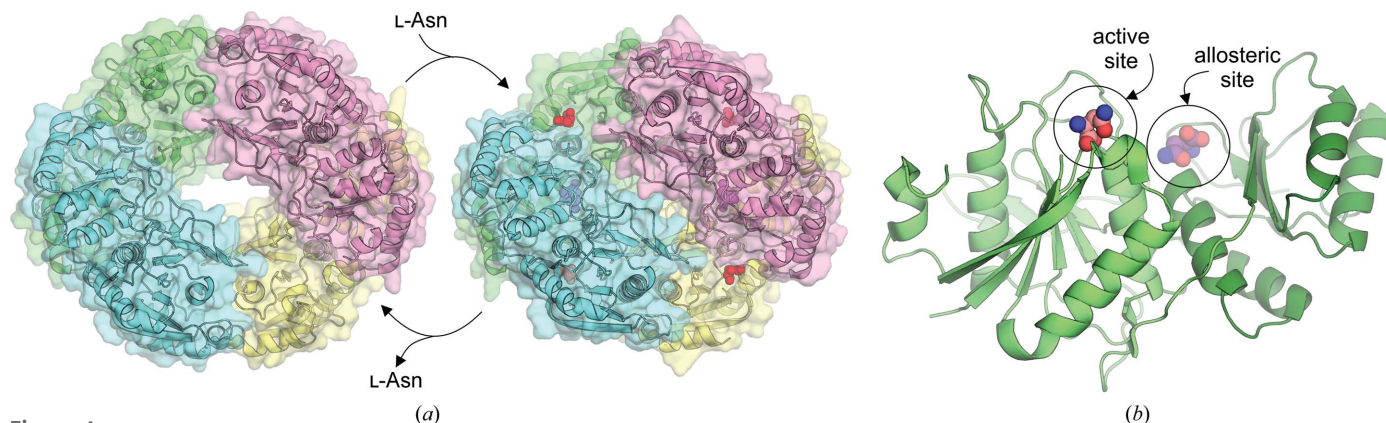


Figure 4

(a) Conformational changes accompanying the allosteric regulation of EcAI: left, without L-Asn (PDB entry 2p2d); right, in the presence of L-Asn (PDB entry 2p2n). (b) L-Asn located in the active site and allosteric site of the EcAI protomer.

the separate domains was active, indicating that the linker is dispensable for the enzymatic activity (Tomar *et al.*, 2013).

3.4. Conformational states

The active site of Class 1 asparaginases is located between the subunits of the intimate dimer between strands $\beta 1$ and $\beta 3$, which form an unusual left-handed β - α - β crossover (Miller *et al.*, 1993; Swain *et al.*, 1993). Each intimate dimer has two complete active sites. The active site can adopt two stable but essentially different conformational states, cat(+) and cat(-), determined by the status of the HR and the FGE and the presence or absence of the substrate (Supplementary Fig. S1). In the absence of the L-Asn substrate, the active site adopts its noncatalytic cat(-) conformation (open conformation, inactive state), with wide-ranging disorder in the FGE region. Upon substrate binding, the FGE changes its conformation to cat(+) (closed conformation, active state), which brings the hydroxyl group of ⁽²⁾Tyr³ towards the substrate and enables the formation of a hydrogen bond between the substrate and ⁽¹⁾Thr (Supplementary Table S2) from the HR (Lubkowski *et al.*, 2020). The FGE can prevent the enzyme from binding other ligands before the reaction is completed, so it also acts as ‘post-substrate entry’ gatekeeper.

It seems that the HR and FGE may spontaneously switch conformation between cat(-) and cat(+), but stabilization of the cat(+) state can only be achieved in the presence of a substrate molecule. When a nonhydrolyzable ligand, for example citrate, is bound in the active site, the FGE remains permanently locked, abrogating enzyme turnover (Tomar *et al.*, 2014). However, it was found that a specific (ordered), rigid β -hairpin structure of the FGE can be present even with an empty active site, for example in the crystal structure of *P. horikoshii* type 1 asparaginase (PhAI), where the FGE with a rigid β -hairpin motif partially covered the active site, limiting substrate access; thus, it can also act as a ‘pre-substrate entry’ gatekeeper (Tomar *et al.*, 2014; Supplementary Fig. S1). Not only a proper substrate can induce the cat(+) conformation; the same effect is achieved by inorganic ions, for example phosphate (Tomar *et al.*, 2014; Guo *et al.*, 2017), sulfate (Lubkowski *et al.*, 2003) or ammonium ions (Jakob *et al.*, 1997), that mimic a substrate bound in the active site. However, in some enzymes, proper loop closure is impossible if the ligand in the active site is too large. For instance, L-Gln bound in the active site of ErAII prevents proper conformational transition to the cat(+) state (Nguyen *et al.*, 2016; Supplementary Fig. S1). Structures of WsAII from *Wolinella succinogenes* revealed that the enzyme may treat the substrate (L-Asn or L-Gln) and the product (L-Asp or L-Glu) in a different way. When L-Asn (or L-Gln) is bound in the active site, a network of hydrogen bonds keeps the FGE in the cat(+) conformation. When the substrate is hydrolyzed to L-Asp (or L-Glu), the hydrogen-bond network is broken to facilitate loop opening

³ To avoid confusing sequence numbers, the catalytic residues are labeled ⁽ⁿ⁾Xxx according to their importance ($n = 1, 2 \dots$) in catalysis. Mapping of n onto concrete sequence numbers is provided in Supplementary Table S2. For plant-type asparaginases the system is similar, but the numbering starts with $n = 100$, as explained in Supplementary Table S4.

and release of the product (Nguyen *et al.*, 2017). An open state with bound reaction product, representing the terminal catalytic state, has been observed not only in WsAII (Nguyen *et al.*, 2017) but also in EcAII (Borek *et al.*, 2014), where such a conformation was interpreted as a crystal-packing artifact.

3.5. The catalytic mechanism of Class 1 L-asparaginases

On the basis of structural and kinetic data, two general catalytic mechanisms have been proposed for Class 1 asparaginases: a single-displacement (or direct-displacement) mechanism was proposed for the mammalian CpAI (Schalk *et al.*, 2016) and a double-displacement (ping-pong) mechanism was proposed for bacterial enzymes *sensu stricto* (Lubkowski & Wlodawer, 2019), which is supposed to proceed via a covalent acyl-enzyme intermediate (Palm *et al.*, 1996). Recent papers unequivocally show that hydrolysis of L-Asn by EcAII follows the double-displacement mechanism, including a sequence of two nucleophilic substitutions. In the ping-pong process, two tetrahedral intermediates (TI1 and TI2) are created in the course of the catalytic cycle. Two important features, namely (i) activation of the nucleophilic threonine by a distant general base (proton sink) and (ii) an unusual structure of the oxyanion hole, including a water molecule, differentiate EcAII, and probably all type 2 and type 1 asparaginases, from other hydrolytic enzymes (Lubkowski *et al.*, 2020). These studies have identified five residues that are crucial for catalysis by the model Class 1 asparaginase EcAII: Thr12, Tyr25, Thr89, Asp90 and Lys162, with additional residues playing a role in substrate binding and specificity (Lubkowski *et al.*, 2020). As these residues are conserved among other bacterial enzymes, a unified nomenclature is used in this review: ⁽¹⁾Thr, ⁽²⁾Tyr, ⁽³⁾Thr, ⁽⁴⁾Asp and ⁽⁵⁾Lys [Fig. 3(c)]. A summary of the residue numbers in the different enzymes is given in Supplementary Table S2.

In a bacterial type 2 enzyme such as EcAII, initiation of the catalytic process requires precise orientation of the substrate molecule, completed by the conformational change of the HR and transition of the FGE from the cat(-) to the cat(+) state. The conformational changes are driven by electrostatic forces and lead to the polarization of the ⁽¹⁾Thr hydroxyl group (Lubkowski *et al.*, 2020). ⁽¹⁾Thr acts as a weak nucleophile without the direct assistance of a general base which could enhance its nucleophilic properties. The role of general base can be attributed to the carboxylate of ⁽⁴⁾Asp; however, this residue is distant from ⁽¹⁾Thr. The first nucleophilic attack is launched by ⁽¹⁾Thr and the reaction proceeds by covalent bond formation between ⁽¹⁾Thr (O^γ) and the C^γ atom of the substrate [Fig. 5(a)]. The proton abstracted from the nucleophile is transferred through ⁽²⁾Tyr, which acts as a proton conveyor, and a network of water molecules to ⁽⁴⁾Asp. The α -carboxylate group of the substrate and the negatively charged side chain of ⁽⁴⁾Asp form a proton sink which increases the efficiency of the first nucleophilic attack, leading to the formation of the negatively charged tetrahedral intermediate TI1 (Fig. 6), stabilized by the oxyanion hole [Fig. 5(b)]. The oxyanion hole is formed by the backbone

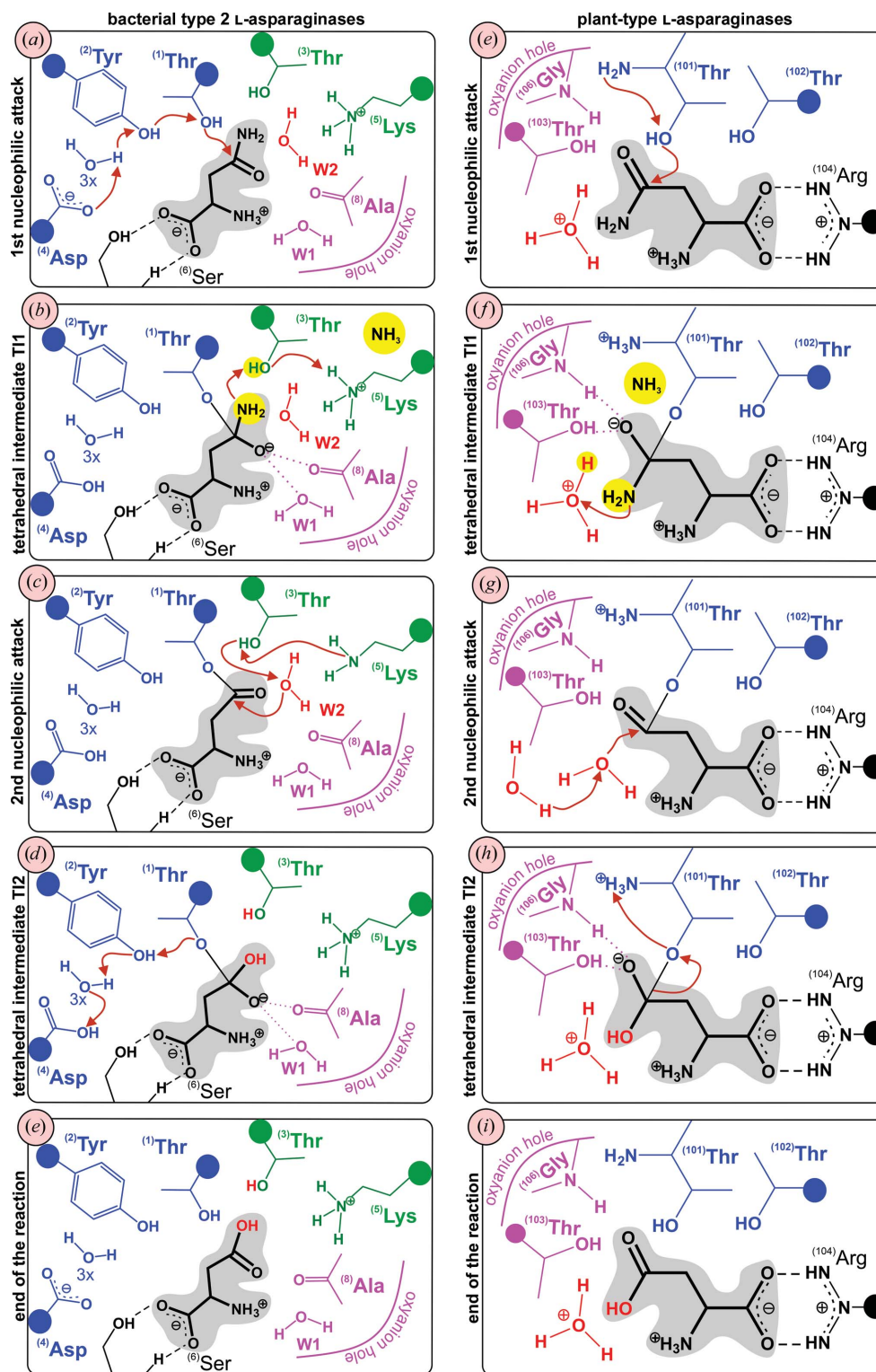


Figure 5
 (a)–(e) Steps in the catalytic mechanism of type 2 bacterial (Class 1) L-asparaginases based on the reactions proposed by Lubkowski *et al.* (2020). Residues and waters (3x) involved in proton shuttling from ⁽¹⁾Thr to ⁽⁴⁾Asp via ⁽²⁾Tyr are colored blue. ⁽³⁾Thr and ⁽⁵⁾Lys that exchange protons with conserved water W2 (red) are colored green. ⁽⁴⁾Ala (carbonyl O atom) and W1 creating the oxyanion hole are colored magenta (the carbonyl O atom of ⁽¹⁾Thr, which is also part of the oxyanion hole, is not shown). ⁽⁶⁾Ser participating in substrate/product anchoring, the L-Asn substrate, intermediate products and the L-Asp end product are colored black (substrate/product is also hydrogen-bonded to ⁽⁷⁾Gln, ⁽⁸⁾Ala, ⁽⁹⁾Asn' and ⁽¹⁰⁾Glu', but these interactions are not shown in the figure). (f)–(j) Steps of the catalytic mechanism of Class 2 L-asparaginases based on the reactions proposed by Nomme *et al.* (2012). Protein residues involved in catalysis are colored blue, residues creating the oxyanion hole are colored magenta and water molecules important for catalysis are colored red, ⁽¹⁰⁴⁾Arg important for substrate anchoring is colored black (substrate/product is also hydrogen-bonded to ⁽¹⁰³⁾Thr, ⁽¹⁰⁶⁾Gly, ⁽¹⁰⁵⁾Asp and ⁽¹⁰⁷⁾Gly, but these interactions are not shown in the figure). In all panels, the directions of the nucleophilic attacks are shown by brown arrows. For clarity, only the functional groups of the protein residues are shown in atomic detail; the remainder is marked by a circle. In all panels, the substrate/product molecules are colored black and shaded.

amides of ⁽¹⁾Thr and ⁽⁸⁾Ala, and by a conserved water molecule W1 (Figs. 5 and 6). Next, the molecule of ammonia is liberated from the substrate, which remains covalently attached to ⁽¹⁾Thr [Fig. 5(b)].

The release of NH₃ initiates the second displacement step. The leaving –NH₂ group of the substrate accepts a proton from the OH group of ⁽³⁾Thr, which in turn accepts one from ⁽⁵⁾Lys. After the release of ammonia, ⁽⁵⁾Lys remains uncharged [Fig. 5(c)] and recovers protonation from water W2 [Fig. 5(c)], which like water W1 is conserved in all type 2 L-asparaginases. The shuttle of protons activates water molecule W2 (Fig. 6), making it a strong nucleophile (OH[−]) to start the second displacement reaction, which leads to the second tetrahedral intermediate TI2 [Fig. 5(d)]. The proton from the carboxylic group of ⁽⁴⁾Asp is transferred back to ⁽¹⁾Thr via the same network of water molecules and ⁽²⁾Tyr. The covalent bond between ⁽¹⁾Thr and the product becomes weaker and finally breaks, leading to the formation of the L-Asp product [Fig. 5(e)]. The departure of L-Asp recreates the original state of the active site: the HR and FGE return to the cat(−) conformations and the enzyme becomes ready for the next catalytic cycle. In EcAII, the rate of the net enzymatic reaction is controlled by the rate of the first displacement step (Lubkowski *et al.*, 2020). This step is irreversible due to the dissociation of NH₃ from the active site, while the second nucleophilic substitution, carried out by a water molecule, is fully reversible.

It is important to note that positioning of the catalytic residues in Class 1 type 1 enzymes differs from that in type 2, especially for tyrosine ⁽²⁾Tyr. In type 1 enzymes, ⁽²⁾Tyr' comes from the complementary subunit (denoted by ') of the inti-

mate dimer (Supplementary Fig. S2). The type 1 enzymes also lack a ⁽¹⁰⁾Glu' counterpart (Supplementary Table S2). These arrangements of residues also occur in the mammalian enzyme CpAI, where ⁽²⁾Tyr' closing the active site in the cat(+) state is likewise located in a C-terminal loop of the complementary protomer, not in the FGE [Supplementary Fig. S2(e)]. Upon ligand binding this C-terminal loop becomes more ordered and the new conformation locates ⁽²⁾Tyr' in close vicinity of the catalytic ⁽¹⁾Thr residue (Schalk *et al.*, 2014).

Interestingly, careful analysis of the sequences and crystal structures has allowed the detection of possible errors in asparaginase classification. For example, the enzyme from *Campylobacter jejuni* (PDB entry 3nxk; Center for Structural Genomics of Infectious Diseases, unpublished work) was assigned in UniProt as the product of the *ansA* gene, while the architecture of the active site strongly suggests that it belongs to the bacterial type 2 enzymes (Supplementary Table S2). The same mistake was made for *Helicobacter pylori* asparaginase (UniProt gene *ansA*), where the active-site arrangement and substrate affinity suggest that it is a type 2 enzyme (PDB entry 2wlt; Dhavala & Papageorgiou, 2009). As there are different origins of ⁽²⁾Tyr/⁽²⁾Tyr' in type 2/type 1 asparaginases, the structure of the enzyme in the cat(+) state can be treated as an important criterion (together with substrate affinity and sequence similarity) of bacterial asparaginase classification.

The positioning of ⁽²⁾Tyr close to ⁽¹⁾Thr is necessary to initiate the catalytic process. However, this event will have slightly different consequences in type 2 and type 1 enzymes. In type 1 enzymes, ⁽²⁾Tyr' is hydrogen-bonded only to ⁽¹⁾Thr [Supplementary Figs. S2(b)–S2(d)], while in type 2 enzymes it

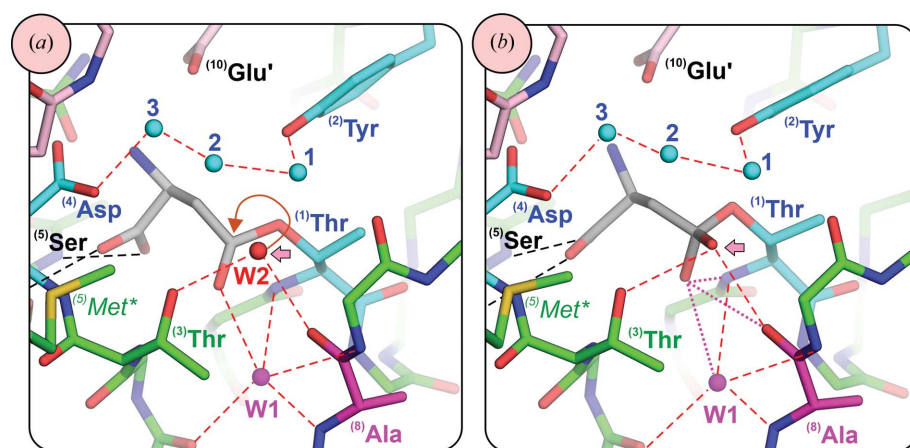


Figure 6

Details of the catalytic mechanism of bacterial (Class 1) type 2 asparaginases. The panels show structures of the K162M mutant of EcAII, in which ⁽⁵⁾Lys was replaced by ⁽⁵⁾Met*. (a) Acyl-enzyme intermediate (PDB entry 6v2g). (b) Tetrahedral intermediate TI2 (PDB entry 6v25). Residues and waters (W1, W2, W3) participating in proton shuttling from ⁽¹⁾Thr to ⁽⁴⁾Asp via ⁽²⁾Tyr are colored cyan. ⁽³⁾Thr and ⁽⁵⁾Met* (corresponding to ⁽⁵⁾Lys in the wild-type protein) are colored green. Red dashed lines represent hydrogen bonds between waters and residues involved in catalysis. Magenta dotted lines show the interactions stabilizing the TI2 intermediate in the oxyanion hole formed by the carbonyl O atom of ⁽⁸⁾Ala and the conserved water molecule W1. A fat pink arrow marks the position of water W2 in (a), which is incorporated into TI2 in (b). Black dashed lines represent hydrogen bonds anchoring the substrate/product in the active site. Some less important hydrogen bonds were omitted for clarity. A brown curved arrow indicates the direction of the nucleophilic attack of water W2 on the acyl-enzyme intermediate.

is also hydrogen-bonded to a chain of three conserved water molecules [Supplementary Fig. S2(a)]. In this context, the scheme of the catalytic reaction described for EcAII is probably true for the whole family of microbial type 2 L-asparaginases (Lubkowski *et al.*, 2020) and glutaminase-asparaginases (G-As) (Supplementary Fig. S3). However, docking of L-Gln in a G-A enzyme seems to proceed via two steps and is more complex than in the case of L-Asn (Strzelczyk *et al.*, 2020). The probable difference in the catalytic mechanisms of Class 1 type 2 and type 1 asparaginases is related to the mode of proton transfer between ⁽¹⁾Thr and ⁽⁴⁾Asp, but this mechanism has not been clarified to date. Interestingly, the structural analysis of the mammalian CpAI and its mutants unambiguously revealed that the substrate does not form a covalent acyl-enzyme intermediate with ⁽¹⁾Thr, in contrast to the scenario described for EcAII. This observation argues in favor

of the direct-displacement mechanism in the mammalian enzymes. Another piece of evidence is the absence of so-called ‘burst kinetics’ in CpAI, which is typical for enzymes utilizing the ping-pong mechanism (Schalk *et al.*, 2016). An excellent review of the mechanism of Class 1 asparaginases has been recently published by Lubkowski & Wlodawer (2021).

3.6. Substrate specificity and inhibitors of Class 1 asparaginases

Another classification of asparaginases is based on enzymatic activity and substrate specificity. In this scheme, Class 1 L-amidohydrolases can be divided into two major groups. The first includes enzymes that primarily utilize L-asparagine as their substrate and have low or very low L-glutaminase activity (for example EcAII and ErAII). The second group includes enzymes, referred to as glutaminase-asparaginases, which hydrolyze L-Asn and L-Gln with a similar efficiency [for example *Pseudomonas* 7A glutaminase-asparaginase (PGA) and *Acinetobacter glutaminasificans* glutaminase-asparaginase] (Aghaiypour *et al.*, 2001b; Jakob *et al.*, 1997).

In addition to L-Asn, bacterial type 1 and type 2 asparaginases can also accept other substrates, such as L-Gln, D-Asn,⁴ and succinic monoamide (Supplementary Fig. S4). Structures of complexes of ErAII with L-Glu, D-Asp and succinic acid (Suc) revealed a similar binding mode for all of these ligands. These compounds are products of the enzymatic reaction; however, they also can serve as substrates when protonated (Aghaiypour *et al.*, 2001b). Well known irreversible inhibitors of asparaginases (and glutaminases) can be found among diazo derivatives of amino acids, e.g. L-DON (6-diazo-5-oxo-L-norleucine) and L-DONV (5-diazo-4-oxo-L-norvaline), but inhibitory activity is also attributed to citrate ions (Fig. S4) (Lubkowski *et al.*, 2019). However, the chemistry of the reactions of the enzymes with diazo inhibitors might be significantly different from their mechanisms with natural substrates because of the high reactivity of diazo compounds (Aghaiypour *et al.*, 2001a). The crystal structures revealed that the diazo inhibitors are covalently bound to the nucleophilic ⁽¹⁾Thr but also to ⁽²⁾Tyr from the FGE [Supplementary Fig. S3(b)], suggesting that the mode of inhibition is different from that of citrate ion, which only reacts with ⁽¹⁾Thr [Supplementary Fig. S3(c)]. The crystal structures of PGA inhibited with L-DON and L-DONV showed that the covalent linkage to ⁽¹⁾Thr is critical for the inhibition process, while the covalent linkage to ⁽²⁾Tyr permanently locks the FGE in the closed conformation (Ortlund *et al.*, 2000). Similar observations were made for ErAII complexed with L-DON (or D-DON). Both inhibitors were covalently bound to ⁽¹⁾Thr and ⁽²⁾Tyr, but they did not affect the conformations of residues in the active site or the positions of the conserved water molecules (Aghaiypour *et al.*, 2001a).

Some residues located in the active site play a role in defining substrate specificity. The natural substrate, L-Asn, fits perfectly to the active site of type 2 and type 1 asparaginases.

⁴ Detection of weak D-asparaginase activity may be controversial because optically pure D-Asn is virtually impossible to obtain.

It has also been shown that an alternative mode of L-Asn orientation is possible in EcAI (Yun *et al.*, 2007). Binding the larger L-Gln requires more sophisticated gymnastics within the active site of bacterial asparaginases (Nguyen *et al.*, 2016). The specificity of L-asparaginases towards L-Asn or L-Gln seems to be determined by two residues in the close vicinity of the active site, positions ⁽⁷⁾Gln and ⁽⁹⁾Asn' (Supplementary Table S2; Aghaiypour *et al.*, 2001b; Yun *et al.*, 2007). Similarly to EcAII, the *H. pylori* L-asparaginase HpAII, with negligible L-glutaminase activity, has a triad, ⁽⁷⁾Gln, ⁽⁹⁾Asn' and ⁽¹⁰⁾Glu', near the catalytic region that is typical of enzymes with dominating L-asparaginase activity. These residues modulate the accessibility of the active site, excluding larger substrates such as L-Gln. Enzymes with significant glutaminase co-activity, as in all G-A enzymes, have Glu and Ser at positions corresponding to ⁽⁷⁾Gln and ⁽⁹⁾Asn', respectively (Supplementary Fig. S3 and Table S2). Other substitutions (Supplementary Table S2) at positions ⁽⁷⁾Gln/⁽⁹⁾Asn' in bacterial type 2 enzymes include ⁽⁷⁾Gln/⁽⁹⁾Ser' in ErAII and EwAII (both have a relatively low glutaminase co-activity; Nguyen *et al.*, 2016; Kravchenko *et al.*, 2008) and ⁽⁷⁾Ser/⁽⁹⁾Ser' in CjAII (no kinetic data). In type 1 enzymes other combinations are present: ⁽⁷⁾Ser/⁽⁹⁾Asn' in EcAI (no detectable glutaminase activity; Yun *et al.*, 2007), YpAI and VcAI (no kinetic data) and ⁽⁷⁾Thr/⁽⁹⁾Gly' in PfAI, PhAI and TkAI (all have negligible glutaminase activity; Saeed *et al.*, 2020; Bansal *et al.*, 2010; Chohan *et al.*, 2020).

Another residue affecting glutaminase co-activity is proline located in close proximity to the catalytic ⁽¹⁾Thr. In *H. pylori* HpAII ⁽¹⁾Thr is flanked by Asn123 instead of the proline present at the corresponding position of EcAII. This change reduces the hydrophobicity around ⁽¹⁾Thr and may have an additional influence on the substrate preferences of HpAII, which has negligible glutaminase activity (Dhaval & Pappageorgiou, 2009). On the other hand, WsAII from *W. succinogenes* has two variants with Pro/Ser121 flanking ⁽¹⁾Thr. It was found that the proline variant has L-glutaminase activity, while the serine variant is free of glutaminase activity. Crystallographic studies revealed that residue 121 affects the conformation of the conserved ⁽²⁾Tyr of FGE. When L-Gln is docked in the active site, steric clashes with the extra methylene group prevent loop closure in the Ser121 variant. In contrast, the circular side chain of Pro121 participates in CH... π aromatic interactions with ⁽²⁾Tyr, stabilizing the cat(+) state even in the presence of L-Gln (Nguyen *et al.*, 2017).

4. Class 2 asparaginases (formerly known as plant-type asparaginases)

4.1. Architecture of Class 2 L-asparaginases

Plant-type (type 3, Class 2) asparaginases are present not only in plants but also in microorganisms, insects and mammals, including humans (Borek & Jaskólski, 2001). They belong to the Ntn-hydrolase family (Brannigan *et al.*, 1995) and usually have dual isoaspartyl aminopeptidase/L-asparaginase activity (EC 3.5.1.1/3.4.19.5). Ntn-hydrolases are expressed as inactive precursors that develop enzymatic

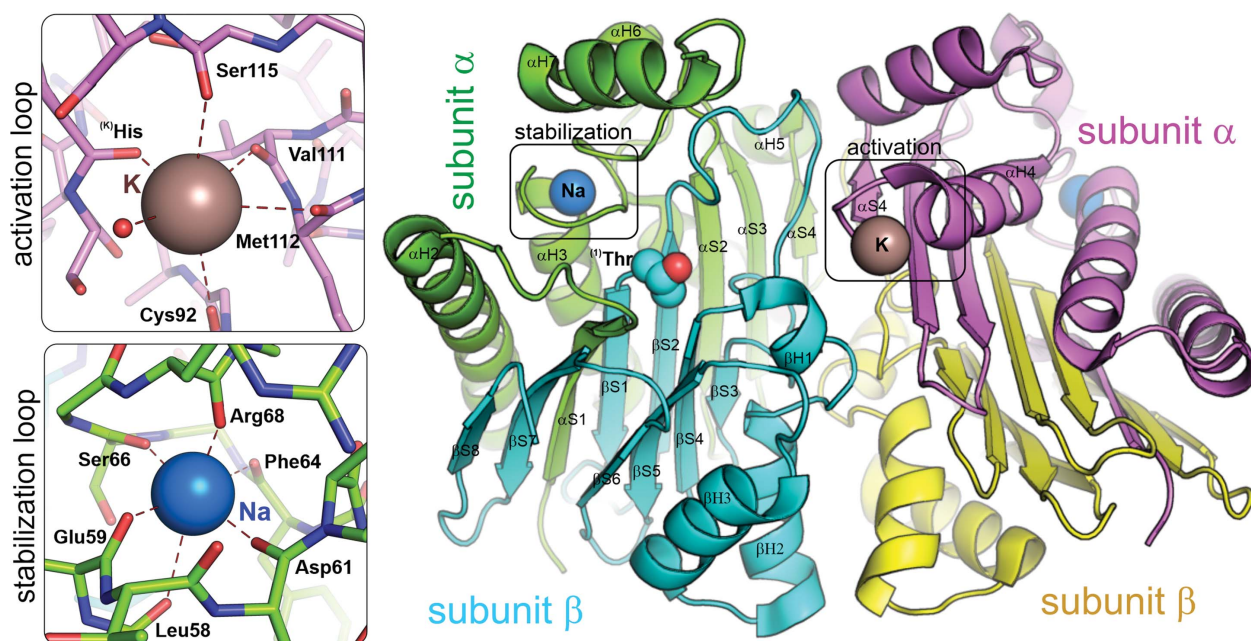


Figure 7

Crystal structure of the mature form of the Class 2 potassium-dependent L-asparaginase from *P. vulgaris* (PDB entry 4pv2). The mature protein is a heterotetramer built of two α subunits (green and magenta) and two β subunits (cyan and yellow). All Class 2 L-asparaginases possess a stabilization loop with a coordinated Na^+ ion (blue sphere). Potassium-dependent enzymes also possess an activation loop that is capable of coordinating a K^+ ion (brown sphere).

activity upon autoproteolytic maturation. The autoproteolysis involves either the removal of a pro-peptide or cleavage of the precursor chain (Michalska *et al.*, 2008). The mechanism of autocleavage is diversified between different Ntn-enzymes and also among Class 2 asparaginases (Michalska *et al.*, 2008; Nomme *et al.*, 2014).

The most thoroughly studied Class 2 L-asparaginases are human HsAIII (Nomme *et al.*, 2014, 2012; Su *et al.*, 2013) and *E. coli* EcAIII (Michalska *et al.*, 2005, 2008; Borek & Jaskólski, 2000). These proteins come from very distant taxonomic groups but share high structural similarity (r.m.s.d. of 0.6 Å). This structural homology is also shared by other Class 2 enzymes. However, except for EcAIII, HsAIII and their mutants, the PDB only contains structures of the enzymes from yellow lupin (*L. luteus*; LIAIII; Michalska *et al.*, 2006), kidney bean (*P. vulgaris*; PVAIII; Bejger *et al.*, 2014) and guinea pig (*Cavia porcellus*; CpAIII; Schalk & Lavie, 2014), and all of these structures are almost identical (r.m.s.d. of 0.6–0.8 Å; Supplementary Table S3). All Class 2 asparaginases have the same architecture arising from two copies of a 33–35 kDa precursor assembled into an inactive dimer (Supplementary Fig. S5). During the maturation process, each precursor is autocatalytically cleaved into the N-terminal subunit α (~20 kDa) and C-terminal subunit β (~15 kDa). After maturation, the whole assembly can be viewed as a heterotetramer or as a dimer of two $\alpha\beta$ heterodimers. Structurally, the heterodimer has $\alpha\beta\alpha^5$ sandwich topology, with

the core composed of two β -sheets flanked on each side by a layer of α -helices. The smaller β -sheet consists of four antiparallel β -strands belonging to chain β , while the larger sheet has mostly antiparallel character and is made of eight strands contributed by both chains (Fig. 7).

Class 2 asparaginases share common structural elements that are crucial for their enzymatic performance: a sandwich-like fold of two distinct subunits connected by a linker (Supplementary Fig. S5), an autoproteolytic and catalytic center with a conserved threonine triad and other conserved elements, the His-Gly-Gly (HGG) loop and binding site for autoprocessing accelerators, and two alkali-metal binding loops: the stabilization loop and activation loop (Fig. 7). In the following description of the structure and catalytic mechanisms, similarly to the system adopted in Section 2, a unified residue numbering will be used, with the corresponding numbers in each particular enzyme listed in Supplementary Table S4.

Chains α and β of Class 2 enzymes, forming the two structural domains of the precursor protein, are connected via a linker or variable loop. The linker has a highly variable sequence and is usually not visible in crystal structures (Supplementary Fig. S5), which suggests its high dynamic disorder or its processive C-terminal degradation after the incision in front of the catalytic $^{(101)}\text{Thr}$ (Michalska *et al.*, 2008). However, in the structure of the EcAIII T179A mutant (with the catalytic threonine replaced by an alanine) it was possible to model as many as 11 residues upstream of the $^{(100)}\text{Gly}$ - $^{(101)}\text{Thr}$ cleavage point (PDB entry 3c17; Michalska *et al.*, 2008). The visible part of the linker in this structure tightly covers access to the active site (Supplementary Fig. S5). This

⁵ In the literature, the same Greek letters (α , β) are used to denote both the subunits and secondary-structure elements. In most cases, in this review we qualify these symbols to avoid ambiguity.

might suggest that the linker protects the immature protein from degradation by other proteases or premature auto-activation under suboptimal conditions. Interestingly, in the crystal structure of HsAIII in the pre-cleaved state co-crystallized with glycine (which was used as an autoprocessing activator; see Section 4.2), it was possible to model a large fragment of the linker (PDB entry 4osx). Its conformation is different from that in uncleaved EcAIII; however, this might be due to the presence of the glycine molecule close to the scissile $^{(100)}\text{Gly}$ - $^{(101)}\text{Thr}$ bond and its interactions with the linker residues (Supplementary Fig. S6). This suggests that the glycine molecule not only activates the nucleophilic $^{(101)}\text{Thr}$ but possibly also affects the conformation of the linker itself, helping to 'uncover' the active site, thus facilitating the auto-activation process.

4.2. Autoproteolytic maturation

The funnel-shaped active site of Class 2 asparaginases is located between the two β -sheets of the $\alpha\beta$ heterodimer, with an entrance surrounded by loops from both subunits (Michalska *et al.*, 2005). The active site is formed as a conserved threonine triad $^{(101)}\text{Thr}$ - $^{(102)}\text{Thr}$ - $^{(103)}\text{Thr}$ (Supplementary Table S4). This site has dual activity: it hydrolyzes its asparagine (and isoaspartyl dipeptide) substrates, and it is also responsible for the cleavage of the inactive precursor in the self-maturation process. The precursor is split into subunits α and β to liberate the catalytic $^{(101)}\text{Thr}$ at the N-terminus of subunit β (Michalska *et al.*, 2008; Nomme *et al.*, 2012). A series of mutants were generated to elucidate the role of the threonine triad in autoprocessing (Nomme *et al.*, 2014). It was established that correct positioning of the $^{(101)}\text{Thr}$ hydroxyl group is critical for autocleavage, while $^{(102)}\text{Thr}$ and $^{(103)}\text{Thr}$ are not essential for this process. It seems that a torsional strain at the $^{(101)}\text{Thr}$ side chain in the immature precursor is the driving force of the autoprocessing reaction (Li *et al.*, 2012). The cleavage relieves the torsional restriction of $^{(101)}\text{Thr}$ and this process is facilitated by a conformational

change of the HGG loop [Supplementary Fig. S6(b)], which is conserved in Class 2 asparaginases (Nomme *et al.*, 2014). It is also possible that a conformational flip of the first glycine residue may promote structural changes that lead to the initiation of the autocleavage reaction (Nomme *et al.*, 2012). Substitution of the glycine residues in the HGG loop by alanines resulted in a sharp decrease in the autoprocessing rate. It was found that the cleaved and uncleaved states may also differ in the position of the whole fragment harboring the HGG element (Li *et al.*, 2012).

The autocleavage starts with a nucleophilic attack of the side chain of $^{(101)}\text{Thr}$ on the carbonyl C atom of the peptide bond in front of it (Li *et al.*, 2012; Fig. 8). $^{(101)}\text{Thr}$ is involved in a network of hydrogen bonds that make its OH proton labile. The most important interaction is the hydrogen bond to the hydroxyl group of $^{(102)}\text{Thr}$, the position of which is stabilized by hydrogen bonds [Supplementary Fig. S6(a)] to a conserved tandem of glycine residues in front of it. In the next step, a tetrahedral transition state is built and stabilized by the oxyanion hole formed by $^{(\text{Na})}\text{Asn}$ from the sodium-binding loop and a water molecule. This oxyanion hole is different from that utilized by the mature enzyme for the asparaginase reaction. In the final step, a nucleophilic water molecule hydrolyzes the tetrahedral transition state to complete the reaction (Michalska *et al.*, 2008). The self-maturation process not only liberates the catalytic $^{(101)}\text{Thr}$ but also provides the environment for substrate binding. The crystal structure of L-Asp in complex with uncleaved HsAIII revealed that in the immature protein the substrate (Fig. 9) cannot be correctly oriented with respect to $^{(101)}\text{Thr}$ and therefore no hydrolysis can take place (Nomme *et al.*, 2012).

Several studies have indicated that different Class 2 asparaginases have different autoproteolysis rates and that in some cases the process can be accelerated by small-molecule activators, at least in the case of human and guinea pig asparaginases (Schalk & Lavie, 2014). The first such activator to be reported was glycine (Su *et al.*, 2013), which accelerates autoprocessing *in vitro* and *in vivo* (Li *et al.*, 2016). Crystal

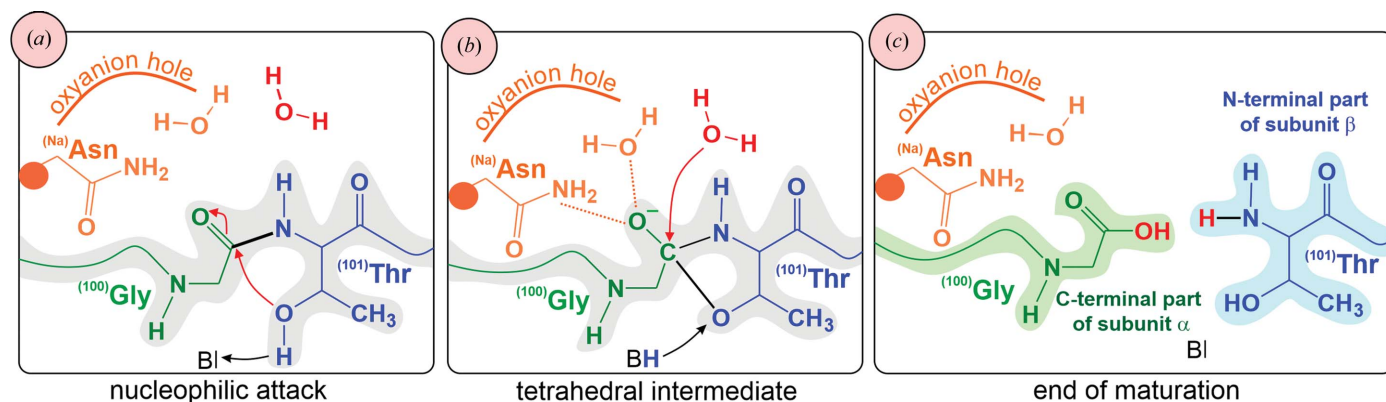


Figure 8

(a)–(c) Schematic representation of the mechanism of maturation of Class 2 asparaginases. The process is started by the nucleophilic attack of $^{(101)}\text{Thr}$ on the scissile bond $^{(100)}\text{Gly}$ - $^{(101)}\text{Thr}$ (black). The reaction needs a water molecule (red) and assistance of the general base **B**. The oxyanion hole (orange) is made by $^{(\text{Na})}\text{Asn}$ from the stabilization loop and another water molecule (orange). At the end of the maturation process, the protomer is cleaved into two subunits: α (green) and β (blue).

structures of HsAIII–Gly complexes show that two glycine molecules can bind to the enzyme in the pre-cleaved state [Supplementary Fig. S6(c)]. One mimics the L-Asn substrate in the active site and is hydrogen-bonded to the conserved ⁽¹⁰⁴⁾Arg residue, while the other is positioned near the scissile ⁽¹⁰⁰⁾Gly–⁽¹⁰¹⁾Thr bond (Su *et al.*, 2013). It is possible that in the interaction with ⁽¹⁰¹⁾Thr the glycine activator acts as a base for proton abstraction from the hydroxyl group, activating its nucleophilic character and promoting its attack on the preceding peptide bond.

Amino acids similar to glycine, such as alanine or serine, seem to be too large to fit the binding site near the scissile ⁽¹⁰⁰⁾Gly–⁽¹⁰¹⁾Thr bond and do not accelerate autoproteolysis (Su *et al.*, 2013). It has been suggested that the glycine activator is first bound to ⁽¹⁰⁴⁾Arg and only later near the scissile bond, as the ⁽¹⁰⁴⁾Arg/Gln mutant (R196Q in HsAIII) requires a very high glycine concentration to rescue the autoprocessing, while without glycine no autoprocessing of the HsAIII R196Q variant takes place (Li *et al.*, 2016). Tests of other compounds revealed that sarcosine, carbamate and phosphate also promoted the cleavage, albeit much less efficiently. Sarcosine, for instance, accelerated the cleavage, while its isostere β -alanine did not, indicating that the position of the glycine amino group provides the interactions necessary for the cleavage reaction. Ammonium, β -alanine, betaine, O-phosphorylethanolamine, glycinamide, bicarbonate, taurine and L-asparagine (the substrate) do not accelerate the autoproteolysis reaction (Li *et al.*, 2016).

A recent study has revealed that the autoprocessing of Class 2 asparaginases depends on the oligomeric state (Li *et al.*, 2016). In the studies reported for HsAIII, measurements of T_m as a function of time during autoproteolysis allowed three different molecular species to be distinguished: $\alpha\beta|\beta\alpha$ (unprocessed homodimer), $\alpha|\beta|\beta\alpha$ (half-processed hetero-

trimer) and $\alpha|\beta|\beta|\alpha$ (fully processed heterotetramer). Autoprocessing removes the structural strain accumulated at ⁽¹⁰¹⁾Thr in the uncleaved protein and increases the thermal stability (melting temperature T_m) from 61°C for $\alpha\beta|\beta\alpha$ to 70°C for the fully cleaved $\alpha|\beta|\beta|\alpha$ species, while the half-processed $\alpha|\beta|\beta\alpha$ form has a T_m of ~65°C. The same study revealed that the cleavage rate of the first $\alpha\beta$ protomer is significantly faster than the cleavage of the second protomer (Li *et al.*, 2016). It is thus possible that cleavage of the first protomer results in allosteric rearrangements at the dimer interface that may reduce the torsional strain at ⁽¹⁰¹⁾Thr of the second protomer or shift the equilibrium of the unprocessed active site away from the cleavable state. Also, cleavage of the first protomer may result in an entropy increase that could lower the probability of a strained conformation in the second protomer (Li *et al.*, 2016). It was also suggested that cleavage of the second protomer is supported by the association of an additional monomer, leading to the formation of a transient trimeric state. The trimerization would trigger a conformational switch in the half-processed dimer, rebuilding the torsional strain in the second protomer (Morais *et al.*, 2020). However, this hypothesis needs to be verified by additional studies, as to date there are no other reports to suggest the possibility of HsAIII trimerization.

Site-directed mutagenesis at the dimer interface of HsAIII showed that dimer formation is also critical for autoproteolysis (Li *et al.*, 2016). In mutants such as C202A or C202S the dimerization process was weaker but autoprocessing recovery was possible in the presence of an activator (glycine), and the fully cleaved variants had nearly wild-type kinetics of substrate hydrolysis. It was thus concluded that L-asparaginase activity and autoprocessing are separate molecular events and are unrelated (Li *et al.*, 2016), even if carried out by the same catalytic center.

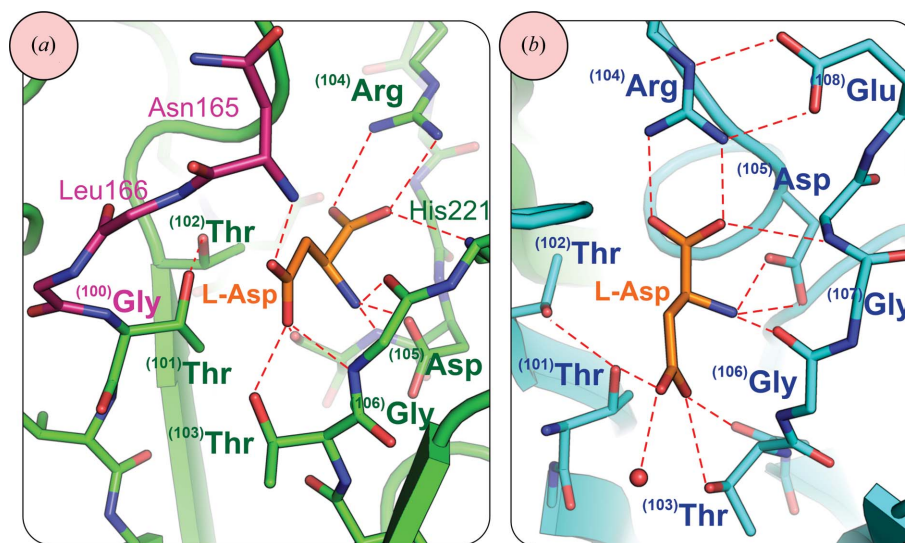


Figure 9

Details of L-Asp (orange) binding in Class 2 asparaginases. (a) L-Asp bound in the active site of immature HsAIII (PDB entry 4pvr); a fragment of the linker region is shown in dark pink. (b) L-Asp in the active site of mature ECAIII and the residues stabilizing its binding (PDB entry 2zal).

4.3. Substrate binding and architecture of the binding site

The dominant enzymatic activity of Class 2 asparaginases is attributed to the hydrolytic decomposition of isoaspartyl (β -aspartyl) dipeptides, while their affinity for L-asparagine is relatively weak (millimolar range; Michalska *et al.*, 2005; Cantor *et al.*, 2009). Isoaspartyl residues are formed spontaneously at L-Asn sites upon protein aging. They introduce unnatural kinks in the protein chain, disrupting normal folding and leading to toxic effects (Aswad *et al.*, 2000). Proteins with isoaspartyl residues are repaired or degraded by proteases. The latter process produces β -peptides, which are further hydrolyzed by specialized isoaspartyl peptidases. The enzymatic degradation of β -aspartyl dipeptides by Class 2 asparaginases has never been investigated in detail, so the catalytic mechanism is not known. On the other hand, the mechanism of L-Asn hydrolysis has been intensively studied for ~20 years due to the potential application of the plant-type enzymes in leukemia therapy (Nomme *et al.*, 2012; Schalk & Lavie, 2014; Michalska *et al.*, 2005, 2006).

Two of the conserved threonines in the $^{(101)}\text{Thr}$ - $^{(102)}\text{Thr}$ - $^{(103)}\text{Thr}$ triad are connected by a network of hydrogen bonds [Supplementary Fig. S6(a)]. The same arrangement of hydrogen bonds exists in the free enzyme and in enzyme-product complexes, so the binding of a substrate/product in the active site does not seem to affect the conformation of the catalytic residues. $^{(101)}\text{Thr}$ acts as the Ntn-nucleophile, *i.e.* is the N-terminal residue of subunit β (β -strand βS1 ; Fig. 7), $^{(102)}\text{Thr}$ is located in β -strand βS2 and $^{(103)}\text{Thr}$ is positioned at β -strand βS3 . In the mature enzyme, the nucleophilic $^{(101)}\text{Thr}$ is the acceptor of a hydrogen bond from the $^{(102)}\text{Thr}$ OH group, while $^{(102)}\text{Thr}$ is an acceptor of two N-H...O hydrogen bonds from the neighboring glycine residues that secure the proper conformation of $^{(102)}\text{Thr}$ (Michalska *et al.*, 2008) [Supplementary Fig. S6(a)]. The catalytic $^{(101)}\text{Thr}$ and other conserved residues, $^{(103)}\text{Thr}$, $^{(104)}\text{Arg}$, $^{(105)}\text{Asp}$, $^{(106)}\text{Gly}$ and $^{(107)}\text{Gly}$ (Supplementary Table S4), located in the β subunit are involved in substrate (product) binding [Fig. 9(b)]. Residue $^{(104)}\text{Arg}$ forms a salt bridge with the α -carboxylate group of the substrate/product. This interaction is responsible for the anchoring of the substrate/product in the active site and explains the specificity of the enzyme, which is limited to substrates with a free α -carboxyl group. The side chain of $^{(104)}\text{Arg}$ is held in place by $^{(108)}\text{Glu}$. $^{(107)}\text{Gly}$ forms a hydrogen bond to one of the O atoms of the α -carboxyl group, while $^{(105)}\text{Asp}$ and $^{(106)}\text{Gly}$ are responsible for anchoring the α -amino group of the substrate/product [Fig. 9(b)]. Residues $^{(103)}\text{Thr}$ and $^{(106)}\text{Gly}$, via their N-H groups, form the oxyanion hole that stabilizes the negative charge developing on the substrate β carbonyl group in the tetrahedral transition state (Michalska *et al.*, 2005).

4.4. Catalytic mechanism

The $^{(101)}\text{Thr}$ nucleophile is the catalytic residue poised for a ping-pong mechanism of L-Asn hydrolysis, which, despite all of the structural differences, suggests a mechanistic similarity to Class 1 asparaginases. The structural differences include a

fundamentally different quaternary structure of Class 2 and Class 1 enzymes, but the situation in the active site looks surprisingly similar: there are also two threonine residues, $^{(101)}\text{Thr}$ and $^{(103)}\text{Thr}$, which could perform a ping-pong mechanism, and likewise there is a (geometrically) problematic proton abstractor (the N-terminus of subunit β) that is supposed to activate the primary nucleophile of $^{(101)}\text{Thr}$. The identity of the final water nucleophile also needs confirmation.

The hydrolysis of L-asparagine is dependent on the free amino group of $^{(101)}\text{Thr}$, which in Ntn-hydrolases is supposed to activate its hydroxyl group, while conserved water molecules located in the close vicinity (and held in position by the adjacent residues from the sodium-binding loop) aid in proton transfer. The function of these water molecules is to stabilize the protonated amino group once it has abstracted the proton from the $^{(101)}\text{Thr}$ hydroxyl. The mechanism of L-Asn hydrolysis proposed by Michalska *et al.* (2005) was expanded by Nomme *et al.* (2012). It can be described as a sequence of two nucleophilic substitutions [Figs. 5(f)–5(j)].

The reaction starts with the nucleophilic attack of $^{(101)}\text{Thr}$ on the substrate and this initiates the first displacement step. The negatively charged tetrahedral transition state at the C γ atom of the substrate is stabilized by the oxyanion hole formed by $^{(103)}\text{Thr}$ and $^{(106)}\text{Gly}$ [Figs. 5(f)–5(j)] (Michalska *et al.*, 2005; Nomme *et al.*, 2012). The covalent transition state also affects the conformation of the conserved HGG loop, suggesting that the HGG motif is involved not only in autoprocesing but also in the L-asparaginase activity. Collapse of the tetrahedral transition state releases a molecule of ammonia and leaves the substrate covalently attached (via β -acylation) to $^{(101)}\text{Thr}$. A water molecule coordinated in the close vicinity of the β -acyl linkage probably acts as a proton donor for the leaving $-\text{NH}_2$ group. The release of NH_3 initiates the second displacement event. The attack of a water molecule on the covalent intermediate results in the formation of another tetrahedral transition state and the reaction is terminated by release of the L-Asp product and the return of the enzyme to the original state (Nomme *et al.*, 2012).

4.5. Regulation of the enzymatic activity: metal-binding loops and catalytic switch

Plant-type asparaginases have a conserved structural element referred to as the sodium-binding loop (or stabilization loop) located near the active site in the α chain. The stabilization loop creates a positional lock on the catalytic $^{(101)}\text{Thr}$ residue by forming a hydrogen bond between its free amino group and the adjacent $^{(\text{Na})}\text{Asn}$ residue (Fig. 7). This interaction is necessary for proper orientation of $^{(101)}\text{Thr}$. The conformation of the stabilization loop is maintained by a sodium cation with an octahedral coordination sphere formed by six carbonyl groups from the main chain within the stabilization loop sequence (Michalska & Jaskolski, 2006).

Depending on their activity profiles, Class 2 L-asparaginases can be further divided into potassium-dependent and potassium-independent enzymes (Fig. 2; Sodek *et al.*, 1980; Bejger *et al.*, 2014). Potassium-dependent plant-type asparaginases

possess yet another loop, which acts to control the enzymatic activity (Bejger *et al.*, 2014). This activation loop can switch the enzyme between ON/OFF states in the presence/absence of potassium cations (Supplementary Fig. S7). When K^+ is coordinated in the activation loop, the side chain of $^{(104)}\text{Arg}$, which is responsible for substrate binding, has a conformation that enables anchoring of the substrate/product in the active site (ON state). When potassium is not bound or is replaced by Na^+ , the loop undergoes a rearrangement that switches the enzyme to the OFF state (Supplementary Fig. S7). This conformational change has a dramatic effect on three side chains forming a ‘catalytic switch’: $^{(K)}\text{His}$ – $^{(K)}\text{Arg}$ – $^{(K)}\text{Glu}$ (Supplementary Table S4). During this transformation, the side chain of $^{(K)}\text{His}$ moves into the protein core, taking up the space usually occupied by the side chain of $^{(104)}\text{Arg}$. To avoid steric clashes, the side chain of $^{(104)}\text{Arg}$ rotates by 180° and in this new position is no longer able to bind to the substrate/product in the active site (OFF state). The movement of $^{(104)}\text{Arg}$ to the OFF state triggers a rotation of the $^{(108)}\text{Glu}$ side chain, which no longer stabilizes $^{(104)}\text{Arg}$ in the ON state (Bejger *et al.*, 2014). All of these changes also affect the position of the whole helix at the C-terminal fragment of chain α and of the neighboring loops. The catalytic switch can be described as an allosteric process, as coordination of K^+ in the activation loop of one protomer affects the position of $^{(104)}\text{Arg}$ in the second protomer.

In potassium-independent asparaginases, the structural element that could be viewed as the activation loop also exists; however, it is unable to bind metal ions. It was confirmed that the $^{(K)}\text{Ser}$ residue is crucial for metal coordination in the potassium-dependent enzymes (Supplementary Table S4). If $^{(K)}\text{Ser}$ is replaced by isoleucine, the enzyme loses its potassium dependence. Conversely, introduction of serine at the corresponding position in the potassium-independent enzyme from *P. vulgaris* endowed it with potassium sensitivity (Ajewole *et al.*, 2018).

Interestingly, it was also discovered that in the crystal structure of the D90E active-site mutant of the Class 1 EcAII enzyme there is a conserved zinc-binding site formed by Asp100, His197 and Asp200 (Borek *et al.*, 2014; Supplementary Fig. S8). This binding site seems to be quite specific despite the lack of an obvious biological context (Cerofolini *et al.*, 2019).

5. Class 3 asparaginases (formerly known as *R. etli*-type asparaginases)

5.1. What do we know about *R. etli* asparaginases?

The most mysterious class of L-asparaginases are the enzymes produced by *R. etli*. These soil bacteria belonging to the genus *Rhizobium* colonize the root nodules of the symbiotic legume *P. vulgaris* (common bean), in which they fix atmospheric nitrogen for its ultimate use in the form of NH_4^+ by the plant cells. On the other hand, organic carbon compounds (including amino acids) are supplied to the bacteria by the plant symbiont as the source of the energy

necessary for nitrogen fixation (Huerta-Zepeda *et al.*, 1996). It has been reported that L-asparaginase activity in *R. etli* cultures was highest when the bacteria were grown on L-Asn as a nitrogen and carbon source, while the activity decreased in the presence of other carbon sources (glycerol, glucose or succinate; Huerta-Zepeda *et al.*, 1996). On the other hand, *R. etli* cultures grown on L-Gln showed only 10% of the asparaginase activity in the presence of L-Asn. Asparaginase activity in *R. etli* is not regulated by oxygen concentration, in contrast to EcAII, which is expressed under anaerobic conditions.

5.2. *R. etli* has two asparaginases: constitutive and inducible

A report published in 1997 indicated that *R. etli*, similarly to other bacteria, also has two asparaginases: constitutive (thermostable) and inducible (thermolabile). The constitutive enzyme was stable at 50°C , while the inducible enzyme had a half-life of only 11 min at 50°C with two melting slopes (Huerta-Zepeda *et al.*, 1997). The two forms differ not only in thermostability but also in electrophoretic mobility: the constitutive protein migrated faster on native electrophoretic gels than the inducible protein. The activity (expression) of the thermolabile enzyme was positively regulated by its substrate L-Asn and was negatively regulated by carbon sources (Huerta-Zepeda *et al.*, 1997). Searching the European Nucleotide Archive revealed that the genome of *R. etli* CFN 42 (GCA_000092045.1; González *et al.*, 2006) encodes at least two L-asparaginases: one related to the *ansA* gene available under accession No. RHE_PE00350 and another deposited with accession No. RHE_CH01144 (Ensembl Bacteria repository; Howe *et al.*, 2020). These data are in good agreement with previous reports (Huerta-Zepeda *et al.*, 1997). The *ansA* gene encoding the thermolabile enzyme is part of the *ansPAB* operon containing four adjacent ORFs: *ansR* (transcriptional repressor), *ansP* (L-asparagine permease), *ansA* (inducible L-asparaginase) and *ansB* (L-aspartase)⁶ (Ortuño-Olea & Durán-Vargas, 2000). The expression of the genes encoded by *ansPAB* is induced by the presence of L-Asn, and the products of this gene expression are responsible for the conversion of L-Asn to L-Asp and ammonia (thermolabile L-asparaginase) and the degradation of L-Asp to fumarate and ammonia in the next step (L-aspartase) (Huerta-Zepeda *et al.*, 1997). We abbreviate the two *R. etli* asparaginases as ReAIV (constitutive) and ReAV (inducible).

Very little is known about the constitutive *R. etli* asparaginase. It was only noted that ReAIV plays a minor role in the utilization of L-Asn as a nitrogen or carbon source (Huerta-Zepeda *et al.*, 1997). The level of sequence homology between ReAIV and ReAV is rather low, with 29.7% identity (Supplementary Fig. S9). Searching UniProt with *BLAST* (Johnson *et al.*, 2008; Altschul *et al.*, 1990) revealed that ReAIV has close homologs (50–80% sequence identity) in

⁶ It should be noted that in *E. coli* (and in most bacteria) the *ansA* gene encodes EcAI (the constitutive enzyme), while the *ansB* gene encodes EcAII (the inducible enzyme). Although the names of the genes are identical, in *R. etli* they encode completely different proteins.

hundreds of bacterial species, predominantly in Proteobacteria (including Rhizobiales). It must be emphasized that the sequences of the constitutive/inducible *R. etli* enzymes are completely different from the bacterial-type (Class 1) or plant-type (Class 2) asparaginases, with identities of only about 13/14% (EcAI), 15/16% (EcAII) and 15/15% (EcAIII). A comprehensive search of the *R. etli* genome available via GenBank (NCBI reference number NC_007761.1; González *et al.*, 2006) using *BLAST* and the sequences of EcAI/II/III as targets revealed that this bizarre organism does not possess any type 1, type 2 or type 3 asparaginases. This justifies classifying the unique *R. etli*-type asparaginases as type 4 and type 5 and naming them ReAIV (constitutive) and ReAV (inducible) in this paper. Conversely, ReAIV and ReAV have no homologs in the *E. coli* genome.

5.3. Properties of the inducible *R. etli* asparaginase (ReAV)

It was noticed early on that the sequence of ReAV is essentially different from the sequences of other microbial asparaginases and that ReAV has probably evolved independently of typical bacterial and fungal asparaginases (Borek & Jaskólski, 2001; Ortuño-Olea & Durán-Vargas, 2000). The lack of sequence similarity between ReAV and bacterial type 2 asparaginases opens the possibility of a different mechanism for asparagine degradation. It was found that ReAV has specific affinity for L-Asn and negligible affinity for L-Gln (Moreno-Enríquez *et al.*, 2012). Biophysical experiments show that recombinant ReAV is most active at alkaline pH (9.0–10.5), while its activity is about threefold lower at the physiological pH of 7.5 and is also reduced by about 27% at pH 11.0. It was also observed that at pH 9.0 the asparaginase activity of ReAV increased with temperature and reached a maximum at 50°C, but a further increase of temperature to 60°C resulted in enzyme inactivation. It was also found that the activity of ReAV was reduced in the presence of the reaction products, which acted as inhibitors: L-Asp (competitive) and NH_4^+ (noncompetitive). Divalent metal cations, such as Mn^{2+} , Zn^{2+} , Ca^{2+} and Mg^{2+} , have also been reported as inhibitors (Moreno-Enríquez *et al.*, 2012), but in view of our own results this information may not be very reliable, at least with regard to some of the metals (Loch *et al.*, unpublished work). ReAV is also not reactive with D-Asn and β -aspartylhydroxamate, indicating a narrow substrate specificity for L-Asn. It was reported that ReAV shows a typical Michaelis–Menten profile, and the kinetic parameters (Supplementary Table S6) indicated that the enzyme has a lower affinity for L-Asn than EcAII.

In silico trials have been performed to model the structure of ReAV and predict its epitopes based on sequence analysis (Huerta-Saquero *et al.*, 2013). The computational model of ReAV has a completely different fold to known Class 1 or Class 2 asparaginases, as well as quite distant immunogenic properties. The computed ReAV structure has significant similarity to proteins with a β -lactamase/transpeptidase fold (Huerta-Saquero *et al.*, 2013).

A search of the UniProt database shows that most ReAV homologs appear in Firmicutes bacteria and in Proteobacteria, although some homologs are also present in eukarya, including the fungi Ascomycetes, often referred to as ‘green and blue molds’. The latter observation was noticed previously by Moreno-Enríquez *et al.* (2012), who suggested horizontal gene transfer between bacteria and fungi. Sequences similar to ReAV were found in fungi belonging to the genera *Aspergillus*, *Penicillium*, *Trichoderma*, *Colletotrichum* and *Fusarium* (50–63% identity). Some of these ubiquitous fungi are pathogenic to humans, but several of them also cause serious plant diseases.

6. Conclusions and outlook

The simple amidohydrolysis at the side chain of asparagine has an amazing variety of biochemical realizations. It is a perplexing question as to why so many different enzymes have evolved to catalyze a very simple, almost trivial, reaction. This question does not have a plausible overarching answer, especially with the *R. etli* enzymes (and their evolution) very likely to complicate the situation even further. The typical bacterial asparaginases (types 1 and 2) found in Gram-negative bacteria serve different purposes in the cytoplasm (type 1) and in the periplasmic space (type 2). The low-affinity EcAI is expressed constitutively and plays a role in the physiological utilization of L-Asn as a nitrogen source. It degrades L-Asn when it is accumulated at an appropriate intracellular (cytoplasmic) concentration. The expression of the high-affinity EcAII is activated by anaerobiosis (by FNR and cyclic AMP receptor). Recent studies suggest that EcAII enables anaerobic respiration and is part of a metabolic pathway leading to the production of fumarate as a terminal electron acceptor. EcAII is localized in the periplasm to be able to utilize exogenous (extracellular) L-Asn to produce fumarate, as under anaerobic conditions transport of L-Asn from the cytoplasm to the periplasm is abolished (Srikhanta *et al.*, 2013). The role of Class 2 enzymes in plants is quite clear, particularly in symbiotic legumes. The atmospheric nitrogen assimilated in the roots of these plants is transported to metabolically active parts (shoots, leaves and flowers) in the form of L-Asn, where it is liberated by hydrolysis. Plant asparaginases are particularly important (and abundant) in seeds, where they have to supply NH_3 for protein synthesis upon germination. Moreover, the second (isoaspartyl aminopeptidase) activity of these enzymes protects dormant seeds from the spontaneous accumulation of highly toxic β -asparagine dipeptides.

Although first detected in animal tissue almost 120 years ago, interest later focused on bacterial L-asparaginases with micromolar substrate affinity sufficient for successful application in the treatment of leukemia. However, even in bacteria another subtype of L-asparaginase is present, structurally similar but with a poor K_m . The plant-specific enzymes that were discovered later are now known to have very close homologs in bacteria. While their different antigenic proper-

ties would make plant-type asparaginases excellent drug candidates, their K_m (millimolar) is too high for this purpose.

Structurally, bacterial-type (Class 1) and plant-type (Class 2) enzymes are completely different, but their active sites have intrinsically similar features, consisting of an abundance of threonine residues that function in a double-displacement ping-pong mechanism. Threonine as an unusual nucleophile was first indicated in the crystal structure of EcAII (Swain *et al.*, 1993) and was confirmed in this role by the crystal structure of archaeal proteasome (Löwe *et al.*, 1995). It would be tempting to speculate why threonine is the ‘nucleophile of choice’ in L-asparaginases. One reason might be that the nucleophile for asparagine hydrolysis does not have to be particularly strong, but it is of advantage if it has a structural feature (such as the methyl group) that would allow stereochemical control during the course of the reaction. However, we should be cautious with such speculation because the limited information that we have about *R. etli* asparaginases, for example about their tentative structural homology to β -lactamases, seems to indicate that in this class the nucleophile might be the more classical serine.

The asparaginases found in *R. etli*, a nitrogen-fixing bacterium living in symbiosis with legumes, are a relatively new addition to the asparaginase family. The full structural and mechanistic properties of these enzymes have yet to be elucidated. However, preliminary kinetic data indicate a submillimolar K_m , just one order of magnitude shy of the range required for cancer therapy. It is thus possible that with suitable structure-based engineering the *R. etli*-type enzymes (which also have homologs in eukaryotes) could become the long-sought alternative antileukemics. Research on L-asparaginases has a fascinating history but possibly an even more exciting future.

7. Related literature

The following references are cited in the supporting information for this article: Ajewole (2016), Anderson (2009), Chohan & Rashid (2013), Derst *et al.* (2000), Jaskólski *et al.* (2001), Kotzia & Labrou (2005, 2009), Lubkowski *et al.* (1996), Lubkowski, Wlodawer, Ammon *et al.* (1994), Lubkowski, Wlodawer, Housset *et al.* (1994), Maggi *et al.* (2017), Nakamura *et al.* (1971), Roberts (1976), Sanches *et al.* (2003), Steckel *et al.* (1983) and Tollersrud & Aronson (1992).

Funding information

This work was supported by National Science Centre (NCN, Poland) grants 2020/38/E/NZ1/00035 to JIL and 2020/37/B/NZ1/03250 to MJ.

References

Abaji, R. & Krajcinovic, M. (2019). *Cancer Drug Resist.* **2**, 242–255.
 Aghaiypour, K., Wlodawer, A. & Lubkowski, J. (2001a). *Biochim. Biophys. Acta*, **1550**, 117–128.
 Aghaiypour, K., Wlodawer, A. & Lubkowski, J. (2001b). *Biochemistry*, **40**, 5655–5664.

Ajewole, E. (2016). PhD thesis. University of Western Ontario, Canada. <https://ir.lib.uwo.ca/etd/3872/>.
 Ajewole, E., Santamaria-Kisiel, L., Pajak, A., Jaskólski, M. & Marsolais, F. (2018). *FEBS J.* **285**, 1528–1539.
 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
 Ammon, H. L., Murphy, K. C., Chandrasekhar, K. & Wlodawer, A. (1985). *J. Mol. Biol.* **184**, 179–181.
 Ammon, H. L., Weber, I. T., Wlodawer, A., Harrison, R. W., Gilliland, G. L., Murphy, K. C., Sjölin, L. & Roberts, J. (1988). *J. Biol. Chem.* **263**, 150–156.
 Anderson, W. (2009). *Infect. Disord. Drug Targets*, **9**, 507–517.
 Aswad, D. W., Paranandi, M. V. & Schurter, B. T. (2000). *J. Pharm. Biomed. Anal.* **21**, 1129–1136.
 Atkins, C. A., Pate, J. S. & Sharkey, P. J. (1975). *Plant Physiol.* **56**, 807–812.
 Bansal, S., Gnanewari, D., Mishra, P. & Kundu, B. (2010). *Biochemistry*, **75**, 375–381.
 Bansal, S., Srivastava, A., Mukherjee, G., Pandey, R., Verma, A. K., Mishra, P. & Kundu, B. (2012). *FASEB J.* **26**, 1161–1171.
 Beckett, A. & Gervais, D. (2019). *World J. Microbiol. Biotechnol.* **35**, 152.
 Bejger, M., Imiolczyk, B., Clavel, D., Gilski, M., Pajak, A., Marsolais, F. & Jaskólski, M. (2014). *Acta Cryst.* **D70**, 1854–1872.
 Bonthron, D. T. & Jaskólski, M. (1997). *Acta Biochim. Pol.* **44**, 491–504.
 Borek, D. & Jaskólski, M. (2000). *Acta Cryst.* **D56**, 1505–1507.
 Borek, D. & Jaskólski, M. (2001). *Acta Biochim. Pol.* **48**, 893–902.
 Borek, D., Kozak, M., Pei, J. & Jaskólski, M. (2014). *FEBS J.* **281**, 4097–4111.
 Borek, D., Michalska, K., Brzezinski, K., Kisiel, A., Podkowinski, J., Bonthron, D. T., Krowarsch, D., Otlewski, J. & Jaskólski, M. (2004). *Eur. J. Biochem.* **271**, 3215–3226.
 Boyse, E. A., Old, L. J., Campbell, H. A. & Mashburn, L. T. (1967). *J. Exp. Med.* **125**, 17–31.
 Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J. L., Tomchick, D. R. & Murzin, A. G. (1995). *Nature*, **378**, 416–419.
 Broome, J. D. (1961). *Nature*, **191**, 1114–1115.
 Campbell, H. A., Mashburn, L. T., Boyse, E. A. & Old, L. J. (1967). *Biochemistry*, **6**, 721–730.
 Cantor, J. R., Stone, E. M., Chantranupong, L. & Georgiou, G. (2009). *Biochemistry*, **48**, 11026–11031.
 Cerofolini, L., Giuntini, S., Carlon, A., Ravera, E., Calderone, V., Fragai, M., Parigi, G. & Luchinat, C. (2019). *Chem. Eur. J.* **25**, 1984–1991.
 Chohan, S. M. & Rashid, N. (2013). *J. Biosci. Bioeng.* **116**, 438–443.
 Chohan, S. M., Sajed, M., Naeem, S. & Rashid, N. (2020). *Int. J. Biol. Macromol.* **147**, 131–137.
 Clementi, A. (1922). *Arch. Int. Physiol.* **19**, 369–398.
 Costa, I. M., Schultz, L., De Araujo Bianchi Pedra, B., Leite, M. S. M., Farsky, S. H. P., De Oliveira, M. A., Pessoa, A. & Monteiro, G. (2016). *Sci. Rep.* **6**, 1–11.
 Derst, C., Henseling, J. & Röhm, K.-H. (2000). *Protein Sci.* **9**, 2009–2017.
 Dhavala, P. & Papageorgiou, A. C. (2009). *Acta Cryst.* **D65**, 1253–1261.
 Dinndorf, P. A., Gootenberg, J., Cohen, M. H., Keegan, P. & Pazdur, R. (2007). *Oncologist*, **12**, 991–998.
 Dunlop, P., Meyer, G., Ban, D. & Roon, R. (1978). *J. Biol. Chem.* **253**, 1297–1304.
 Dunlop, P. C. & Roon, R. J. (1975). *J. Bacteriol.* **122**, 1017–1024.
 Epp, O., Steigemann, W., Formanek, H. & Huber, R. (1971). *Eur. J. Biochem.* **20**, 432–437.
 Furth, O. & Friedmann, M. (1910). *Biochem. Z.* **26**, 435–440.
 Geddes, W. F. & Hunter, A. (1928). *J. Biol. Chem.* **77**, 197–229.
 Gervais, D., Allison, N., Jennings, A., Jones, S. & Marks, T. (2013). *Bioprocess Biosyst. Eng.* **36**, 453–460.
 Gonnermann, M. (1902). *Pflügers Arch.* **89**, 493–516.

- González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S. C., Ramírez, M. A., Jiménez-Jacinto, V., Collado-Vides, J. & Dávila, G. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 3834–3839.
- Guo, J., Coker, A. R., Wood, S. P., Cooper, J. B., Chohan, S. M., Rashid, N. & Akhtar, M. (2017). *Acta Cryst. D* **73**, 889–895.
- Hejazi, M., Piotukh, K., Mattow, J., Deutzmann, R., Volkmer-Engert, R. & Lockau, W. (2002). *Biochem. J.* **364**, 129–136.
- Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., Alvarez-Jarreta, J., Barba, M., Bolser, D. M., Cambell, L., Carbajo, M., Chakiachvili, M., Christensen, M., Cummins, C., Cuzick, A., Davis, P., Fexova, S., Gall, A., George, N., Gil, L., Gupta, P., Hammond-Kosack, K. E., Haskell, E., Hunt, S. E., Jaiswal, P., Janacek, S. H., Kersey, P. J., Langridge, N., Maheswari, U., Maurel, T., McDowall, M. D., Moore, B., Muffato, M., Naamati, G., Naithani, S., Olson, A., Papatheodorou, I., Patricio, M., Paulini, M., Pedro, H., Perry, E., Preece, J., Rosello, M., Russell, M., Sitnik, V., Staines, D. M., Stein, J., Tello-Ruiz, M. K., Trevanion, S. J., Urban, M., Wei, S., Ware, D., Williams, G., Yates, A. D. & Flicek, P. (2020). *Nucleic Acids Res.* **48**, D689–D695.
- Huerta-Saquero, A., Evangelista-Martínez, Z., Moreno-Enriquez, A. & Perez-Rueda, E. (2013). *Bioengineered*, **4**, 30–36.
- Huerta-Zepeda, A., Durán, S., Du Pont, G. & Calderón, J. (1996). *Microbiology*, **142**, 1071–1076.
- Huerta-Zepeda, A., Ortuño, L., Du Pont, G., Durán, S., Lloret, A., Merchant-Larios, H. & Calderón, J. (1997). *J. Bacteriol.* **179**, 2068–2072.
- Jakob, C. G., Lewinski, K., LaCount, M. W., Roberts, J. & Lebioda, L. (1997). *Biochemistry*, **36**, 923–931.
- Jaskólski, M., Kozak, M., Lubkowski, J., Palm, G. & Wlodawer, A. (2001). *Acta Cryst. D* **57**, 369–377.
- Jennings, M. P. & Beacham, I. R. (1990). *J. Bacteriol.* **172**, 1491–1498.
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. & Madden, T. L. (2008). *Nucleic Acids Res.* **36**, W5–W9.
- Kidd, J. G. (1953). *J. Exp. Med.* **98**, 565–582.
- Kotzia, G. A. & Labrou, N. E. (2005). *J. Biotechnol.* **119**, 309–323.
- Kotzia, G. A. & Labrou, N. E. (2009). *FEBS J.* **276**, 1750–1761.
- Kravchenko, O. V., Kisilitsin, Y. A., Popov, A. N., Nikonov, S. V. & Kuranova, I. P. (2008). *Acta Cryst. D* **64**, 248–256.
- Lang, S. (1904). *Beitr. Chem. Physiol. Pathol.* **5**, 321–345.
- Lea, P. J., Fowden, L. & Mifflin, B. J. (1978). *Phytochemistry*, **17**, 217–222.
- Lee, B. & Yang, H. J. (1973). *J. Biol. Chem.* **248**, 7620–7621.
- Lees, E. M. & Blakeney, A. B. (1970). *Biochim. Biophys. Acta*, **215**, 145–151.
- Li, W., Cantor, J. R., Yogesha, S. D., Yang, S., Chantranupong, L., Liu, J. Q., Agnello, G., Georgiou, G., Stone, E. M. & Zhang, Y. (2012). *ACS Chem. Biol.* **7**, 1840–1847.
- Li, W., Irani, S., Crutchfield, A., Hodge, K., Matthews, W., Patel, P., Zhang, Y. J. & Stone, E. (2016). *Biochemistry*, **55**, 960–969.
- Lough, T. J., Reddington, B. D., Grant, M. R., Hill, D. F., Reynolds, P. H. S. & Farnden, K. J. F. (1992). *Plant Mol. Biol.* **19**, 391–399.
- Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. & Huber, R. (1995). *Science*, **268**, 533–539.
- Lubkowski, J., Chan, W. & Wlodawer, A. (2019). *Sci. Rep.* **9**, 11070.
- 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., Vanegas, J., Chan, W. K., Lorenzi, P. L., Weinstein, J. N., Sukharev, S., Fushman, D., Rempe, S., Anishkin, A. & Wlodawer, A. (2020). *Biochemistry*, **59**, 1927–1945.
- Lubkowski, J. & Wlodawer, A. (2019). *Protein Sci.* **28**, 1850–1864.
- Lubkowski, J. & Wlodawer, A. (2021). *FEBS J.*, <https://doi.org/10.1111/febs.16042>.
- 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.
- Luo, M., Brooks, M. & Wicha, M. S. (2018). *Cell Metab.* **27**, 947–949.
- Maggi, M., Mittelman, S. D., Parmentier, J. H., Colombo, G., Meli, M., Whitmire, J. M., Merrell, D. S., Whitelegge, J. & Scotti, C. (2017). *Sci. Rep.* **7**, 14479.
- Mashburn, L. T. & Wriston, J. C. (1964). *Arch. Biochem. Biophys.* **105**, 450–453.
- Michalska, K., Borek, D., Hernández-Santoyo, A. & Jaskolski, M. (2008). *Acta Cryst. D* **64**, 309–320.
- Michalska, K., Brzezinski, K. & Jaskolski, M. (2005). *J. Biol. Chem.* **280**, 28484–28491.
- Michalska, K., Bujacz, G. & Jaskolski, M. (2006). *J. Mol. Biol.* **360**, 105–116.
- Michalska, K., Hernandez-Santoyo, A. & Jaskolski, M. (2008). *J. Biol. Chem.* **283**, 13388–13397.
- Michalska, K. & Jaskolski, M. (2006). *Acta Biochim. Pol.* **53**, 627–640.
- Miller, M., Rao, J. K. M., Wlodawer, A. & Gribskov, M. R. (1993). *FEBS Lett.* **328**, 275–279.
- Morais, S. B., Pirolla, R. A. S., Frota, N. F., Lourenzoni, M. R., Gozzo, F. C. & Souza, T. A. C. B. (2020). *J. Proteomics*, **224**, 103818.
- Moreno-Enriquez, A., Evangelista-Martínez, Z., González-Mondragón, E. G., Calderón-Flores, A., Arreguín, R., Pérez-Rueda, E. & Huerta-Saquero, A. (2012). *J. Microbiol. Biotechnol.* **22**, 292–300.
- Nakamura, N., Morikawa, Y., Fujio, T. & Tanaka, M. (1971). *Agric. Biol. Chem.* **35**, 219–225.
- Nguyen, H. A., Durden, D. L. & Lavie, A. (2017). *Sci. Rep.* **7**, 41643.
- Nguyen, H. A., Su, Y. & Lavie, A. (2016). *Biochemistry*, **55**, 1246–1253.
- Nomme, J., Su, Y., Konrad, M. & Lavie, A. (2012). *Biochemistry*, **51**, 6816–6826.
- Nomme, J., Su, Y. & Lavie, A. (2014). *J. Mol. Biol.* **426**, 2471–2485.
- North, A. C. T., Wade, H. E. & Cammack, K. A. (1969). *Nature*, **224**, 594–595.
- Nunes, J. C. F., Cristóvão, R. O., Freire, M. G., Santos-Ebinuma, V. C., Faria, J. L., Silva, C. G. & Tavares, A. P. M. (2020). *Molecules*, **25**, 5827.
- Oinonen, C. & Rouvinen, J. (2000). *Protein Sci.* **9**, 2329–2337.
- Ollenschläger, G., Roth, E., Linkesch, W., Jansen, S., Simmel, A. & MÖdler, A. (1988). *Eur. J. Clin. Invest.* **18**, 512–516.
- Ortlund, E., Lacount, M. W., Lewinski, K. & Lebioda, L. (2000). *Biochemistry*, **39**, 1199–1204.
- Ortuño-Olea, L. & Durán-Vargas, S. (2000). *FEMS Microbiol. Lett.* **189**, 177–182.
- Palm, G. J., Lubkowski, J., Derst, C., Schleper, S., Röhm, K. H. & Wlodawer, A. (1996). *FEBS Lett.* **390**, 211–216.
- Papageorgiou, A. C., Posypanova, G. A., Andersson, C. S., Sokolov, N. N. & Krasotkina, J. (2008). *FEBS J.* **275**, 4306–4316.
- Porton Biopharma (1985). *Erwinase*. <https://www.portonbiopharma.com/products-2/erwinase/>.
- Prager, M. D. & Bachynsky, N. (1968). *Arch. Biochem. Biophys.* **127**, 645–654.
- Prahl, A., Pazgier, M., Hejazi, M., Lockau, W. & Lubkowski, J. (2004). *Acta Cryst. D* **60**, 1173–1176.
- Pritsa, A. A. & Kyriakidis, D. A. (2001). *Mol. Cell. Biochem.* **216**, 93–101.
- Radadiya, A., Zhu, W., Coricello, A., Alcaro, S. & Richards, N. G. J. (2020). *Biochemistry*, **59**, 3193–3200.
- Roberts, J. (1976). *J. Biol. Chem.* **251**, 2119–2123.
- Saeed, H., Hemida, A., El-Nikhely, N., Abdel-Fattah, M., Shalaby, M., Hussein, A., Eldoksh, A., Ataya, F., Aly, N., Labrou, N. & Nematalla, H. (2020). *Int. J. Biol. Macromol.* **156**, 812–828.
- Sanches, M., Barbosa, J. A. R. G., de Oliveira, R. T., Abrahão Neto, J. A. & Polikarpov, I. (2003). *Acta Cryst. D* **59**, 416–422.
- Schalk, A. M., Antansijevic, A., Caffrey, M. & Lavie, A. (2016). *J. Biol. Chem.* **291**, 5088–5100.

- Schalk, A. M. & Lavie, A. (2014). *Biochemistry*, **53**, 2318–2328.
- Schalk, A. M., Nguyen, H. A., Rigouin, C. & Lavie, A. (2014). *J. Biol. Chem.* **289**, 33175–33186.
- Seemüller, E., Lupas, A. & Baumeister, W. (1996). *Nature*, **382**, 468–470.
- Silva, L. S. da, Doonan, L. B., Pessoa, A., de Oliveira, M. A. & Long, P. F. (2021). *Biotechnol. Appl. Biochem.*, <https://doi.org/10.1002/bab.2127>.
- Sodek, L., Lea, P. J. & Mifflin, B. J. (1980). *Plant Physiol.* **65**, 22–26.
- Srikhanta, Y. N., Atack, J. M., Beacham, I. R. & Jennings, M. P. (2013). *Biochem. Biophys. Res. Commun.* **436**, 362–365.
- Steckel, J., Roberts, J., Philips, F. S. & Chou, T.-C. (1983). *Biochem. Pharmacol.* **32**, 971–977.
- Strzelczyk, P., Zhang, D., Dyba, M., Wlodawer, A. & Lubkowski, J. (2020). *Sci. Rep.* **10**, 17516.
- Su, Y., Karamitros, C. S., Nomme, J., McSorley, T., Konrad, M. & Lavie, A. (2013). *Chem. Biol.* **20**, 533–540.
- Swain, A. L., Jaskólski, M., Housset, D., Rao, J. K. M. & Wlodawer, A. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 1474–1478.
- Tanaka, S., Robinson, E. A., Appella, E., Miller, M., Ammon, H. L., Roberts, J., Weber, I. T. & Wlodawer, A. (1988). *J. Biol. Chem.* **263**, 8583–8591.
- Tollersrud, O. K. & Aronson, N. N. (1992). *Biochem. J.* **282**, 891–897.
- Tomar, R., Garg, D. K., Mishra, R., Thakur, A. K. & Kundu, B. (2013). *FEBS J.* **280**, 2688–2699.
- Tomar, R., Sharma, P., Srivastava, A., Bansal, S., Ashish & Kundu, B. (2014). *Acta Cryst. D* **70**, 3187–3197.
- Wlodawer, A., Roberts, J. & Holcenberg, J. S. (1977). *J. Mol. Biol.* **112**, 515–519.
- Yao, M., Yasutake, Y., Morita, H. & Tanaka, I. (2005). *Acta Cryst. D* **61**, 294–301.
- Yonei, M., Mitsui, Y. & Iitaka, Y. (1977). *J. Mol. Biol.* **110**, 179–186.
- Yun, M.-K., Nourse, A., White, S. W., Rock, C. O. & Heath, R. J. (2007). *J. Mol. Biol.* **369**, 794–811.