STRUCTURAL BIOLOGY
ISSN 2059-7983

Rhizobium etli has two L-asparaginases with low sequence identity but similar structure and catalytic center

Joanna I. Loch,[a]‡ Paulina Worsztynowicz,[b]‡ Joanna Sliwiak,[b]‡ Marta Grzechowiak,[b] Barbara Imiolczyk,[b] Kinga Pokrywka,[b] Mateusz Chwastyk,[c] Miroslaw Gilski[b,d] and Mariusz Jaskolski[b,d]*

[a]Department of Crystal Chemistry and Crystal Physics, Faculty of Chemistry, Jagiellonian University, Krakow, Poland, [b]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland, [c]Institute of Physics, Polish Academy of Sciences, Warsaw, Poland, and [d]Department of Crystallography, Faculty of Chemistry, Adam Mickiewicz University, Poznań, Poland. *Correspondence e-mail: mariuszj@amu.edu.pl

Received 16 May 2023
Accepted 26 June 2023

Edited by Z. S. Derewenda, University of Virginia, USA

‡ These authors made equal contributions.

Keywords: L-asparaginases; *Rhizobium etli*; zinc-binding proteins; enzymatic mechanism; occluded water molecules; amidohydrolases; structural homology; enzyme kinetics.

PDB references: ReAIV, monoclinic form R4mC-1, 8osw; monoclinic form R4mC-2, 8clz; tetragonal form R4tP, 8cly; orthorhombic form R4oP-1, 8ori; orthorhombic form R4oP-2, 8col

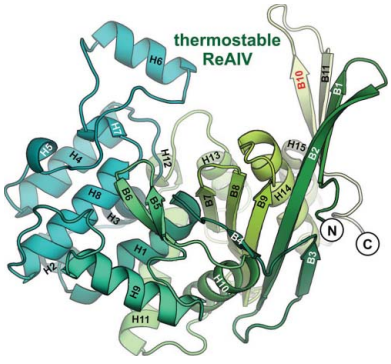
Supporting information: this article has supporting information at journals.iucr.org/d

The genome of *Rhizobium etli*, a nitrogen-fixing bacterial symbiont of legume plants, encodes two L-asparaginases, ReAIV and ReAV, that have no similarity to the well characterized enzymes of class 1 (bacterial type) and class 2 (plant type). It has been hypothesized that ReAIV and ReAV might belong to the same structural class 3 despite their low level of sequence identity. When the crystal structure of the inducible and thermolabile protein ReAV was solved, this hypothesis gained a stronger footing because the key residues of ReAV are also present in the sequence of the constitutive and thermostable ReAIV protein. High-resolution crystal structures of ReAIV now confirm that it is a class 3 L-asparaginase that is structurally similar to ReAV but with important differences. The most striking differences concern the peculiar hydration patterns of the two proteins, the presence of three internal cavities in ReAIV and the behavior of the zinc-binding site. ReAIV has a high pH optimum (9–11) and a substrate affinity of ∼1.3 m*M* at pH 9.0. These parameters are not suitable for the direct application of ReAIV as an antileukemic drug, although its thermal stability and lack of glutaminase activity would be of considerable advantage. The five crystal structures of ReAIV presented in this work allow a possible enzymatic scenario to be postulated in which the zinc ion coordinated in the active site is a dispensable element. The catalytic nucleophile seems to be Ser47, which is part of two Ser–Lys tandems in the active site. The structures of ReAIV presented here may provide a basis for future enzyme-engineering experiments to improve the kinetic parameters for medicinal applications.

1. Introduction

The structural repertoire of L-asparaginases, *i.e.* enzymes that catalyze the simple but metabolically essential reaction of amidohydrolysis at the side chain of L-asparagine, has recently been extended to include a class 3 protein (Loch *et al.*, 2021), formerly known as *Rhizobium etli*-type asparaginase (Borek & Jaskólski, 2001; Loch & Jaskolski, 2021). Class 3, which is completely unrelated to the well established classes 1 (formerly known as bacterial type) and 2 (formerly known as plant type), adds an extra level of complexity to the enzymology of this simple, ammonia-producing reaction. The *R. etli* enzyme for which the crystal structure has been reported (Loch *et al.*, 2021) is a thermolabile and inducible protein named ReAV.

Curiously, *R. etli*, which is a nitrogen-fixing bacterial symbiont of the leguminous common bean, does not possess any class 1 or class 2 L-asparaginases. However, its genome encodes another L-asparaginase, ReAIV, which is constitutive



and thermostable, and which despite rather low sequence identity (31.1%) and similarity (44.6%) has tentatively been included in class 3 together with ReAV. Since ReAIV and ReAV are the only L-asparaginases in *R. etli*, their original classification as *R. etli*-type seemed quite logical. However, in view of a recent pan-genomic analysis of the distribution of L-asparaginases in bacteria (Zielezinski *et al.*, 2022) this term is at least a partial misnomer, as the ReAV-type enzymes are more characteristic of other taxonomic groups of bacteria. Indeed, it is quite likely that *R. etli* acquired the ReAV gene via horizontal gene transfer (HGT) from *Burkholderia*.

The structural canon of the dimeric class 3 L-asparaginases, established by the crystal structure of ReAV, is based on a subunit fold that is reminiscent of some serine β-lactamases, penicillin-binding proteins (PBPs) and glutaminases, with an active site built around two Ser–Lys tandems, in which Ser48 appears to be the key nucleophile. However, at variance with the structural homologs, the active site of ReAV includes a zinc coordination site formed by one lysine residue, two cysteine residues and a water molecule (Loch *et al.*, 2021).

In this work, we determined several crystal structures of the constitutive ReAIV enzyme, showing that it is indeed a structural class 3 ortholog of ReAV. We also report the kinetic parameters and chemical factors, such as pH, that influence the ReAIV-catalyzed L-asparaginase reaction. Systematic investigations of L-asparaginases are important from the medicinal point of view, as enzymes with a sufficiently high (micromolar) substrate affinity are potential drug candidates for the treatment of some forms of leukemia, most notably acute lymphoblastic leukemia (ALL). To date, only enzymes of class 1 (subtype II) derived from *Escherichia coli* and *Erwinia chrysanthemi*, with suitable kinetic properties, have been used in clinical practice, although their administration is often accompanied by severe side effects.

2. Methods

2.1. Cloning, expression and purification

A synthetic gene (GeneArt) encoding the constitutive wild-type (WT) *R. etli* (strain CFN 42) L-asparaginase ReAIV was cloned into pET151/D-TOPO expression vector carrying a tobacco etch virus (TEV) protease-cleavable N-terminal 6×His tag. Production of the ReAIV protein was induced in *E. coli* strain BL21 Gold (DE3) with 0.3 m*M* isopropyl β-D-1-thiogalactopyranoside. Expression was carried out overnight at 16°C. The cells were centrifuged and resuspended in 50 m*M* Tris buffer pH 8.0, 500 m*M* NaCl, 20 m*M* imidazole, 1 m*M* TCEP, 5% glycerol. The cells were disrupted by sonication with the addition of nuclease (OMNI Nuclease, EURx) and the clarified cell lysate was loaded onto a chromatography column filled with HisTrap Ni resin (GE Healthcare). ReAIV was eluted from the column using 400 m*M* imidazole and fractions containing ReAIV were pooled together. The 6×His tag was cleaved off during dialysis at 4°C against 50 m*M* Tris buffer pH 8.0 containing 500 m*M* NaCl in the presence of TEV protease. The 6×His-tag debris and the His-tagged TEV protease were separated in another nickel-affinity chromatography step. The protein was then concentrated, loaded onto a HiLoad 16/600 Superdex 200 (GE Healthcare) size-exclusion (SEC) column and eluted using 25 m*M* Tris pH 8.0 buffer containing 100 m*M* NaCl and 1 m*M* TCEP. The purity and homogeneity of ReAIV were checked by SDS–PAGE.

Protein localization *in vivo* was predicted using *SignalP* (Almagro Armenteros *et al.*, 2019) and cPSORTdb (Yu *et al.*, 2010).

2.2. Determination of thermal stability and molecular size

Prior to nanoDSF measurements, protein samples (ReAIV and ReAV as a reference) were applied onto a Superdex 75 10/300 GL column (GE Healthcare) to exchange the buffer for 50 m*M* phosphate pH 7.5 supplemented with 20 m*M* NaCl. The thermal stability of ReAIV (and ReAV) was determined by nanoDSF, while the hydrodynamic radius (R_h) was determined by nanoDLS at 25°C using a Prometheus Panta instrument (NanoTemper Technologies) and a sample volume of ∼10 µl per capillary. Melting scans were recorded as the fluorescence emission of protein samples subjected to a temperature ramp of 60°C h^{-1} from 25 to 95°C.

2.3. Determination of enzyme kinetics and the effect of pH on ReAIV activity

The effect of pH on the activity of ReAIV (and ReAV as a reference) was determined at values ranging from 7.0 to 11.0 by the Nessler method, as proposed by Simas *et al.* (2021), using appropriate buffers (30 m*M* phosphate at pH 7.0–9.0 and 30 m*M* carbonate at pH 9.0–11.0). The assay mixture consisted of 10 µl 100 m*M* L-Asn (prepared in the appropriate buffer), 80 µl buffer and 10 µl enzyme at a concentration in the range 0.6–0.8 µ*M*. Reactions were carried out in a 96-well plate for 5 min at room temperature (RT) in a final volume of 100 µl. After this time, 10 µl of each reaction mixture was transferred to a new 96-well plate and the reaction was quenched using 3 µl 1.5 *M* trichloroacetic acid (TCA). Next, 10 µl of a stabilizing solution [4 m*M* sodium tartrate with 10 mg ml^{-1} poly(vinyl alcohol) (PVA)], 10 µl Nessler reagent and 67 µl ultrapure water were added to each well. Samples were vigorously stirred and the optical density was recorded immediately at 420 nm using a microplate reader (Hidex). A calibration curve for an ammonium ion concentration range of 0–65 m*M* was prepared in parallel to all reactions in the same 96-well plate. Measurements were made in triplicate for both the enzyme assays and the calibration curve.

Before kinetic analysis, ReAIV and ReAV were concentrated to ∼3.0 mg ml^{-1} using Amicon 30 kDa spin columns (Merck–Millipore) and then dialyzed against 30 m*M* carbonate buffer pH 9.0. Kinetic parameters for L-Asn hydrolysis by ReAIV and ReAV were determined using isothermal titration calorimetry (ITC; Loch *et al.*, 2021) by the multiple injection method (MIM; Wang *et al.*, 2020) under experimental conditions similar to those described previously for ReAV (Loch *et al.*, 2021). A MicroCal PEAQ-ITC (Malvern) calorimeter was used. All measurements were taken in 10 m*M* carbonate buffer pH 9.0 at 37°C with stirring at 700 rev min^{-1} and the differential power set to 10 µcal s^{-1}. In the first step, the

enthalpy of total conversion of all of the substrate into product (total molar enthalpy of the reaction; H_{app}) was determined by injecting 2 µl 10 m*M* substrate into a reaction cell containing 2 µ*M* enzyme. The four injections were separated by 10 min intervals to ensure total substrate conversion. After integration of each peak, the enthalpies of all injections appeared to be comparable, and the four values were averaged to obtain H_{app}. Next, the differential power change (dQ/dt) arising from turnover of the substrate into product was determined in a heat-rate shift experiment, in which the substrate at a concentration of 100 m*M* (in the syringe) was injected in 20 1.8 µl aliquots at short 60 s intervals (to minimize substrate depletion) into the cell, with the enzyme kept at a concentration of ∼12–15 n*M*. The raw rate data were analyzed using the *MicroCal PEAQ-ITC Analysis Software*.

2.4. Mass spectrometry

Protein solutions were subjected to overnight trypsin digestion at 37°C with 1 µg trypsin (Promega). Cysteines were reduced by 1 h incubation with 5 m*M* tris(2-carboxyethyl) phosphine (TCEP) at 37°C followed by 10 min incubation at room temperature with 10 m*M* methyl methanethiosulfonate (MMTS). After digestion, the peptides were acidified with 0.1% formic acid. Samples were analyzed using an Evosep One chromatograph (Evosep Biosystems) coupled to an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). 20 µl samples were loaded onto disposable Evotips C18 trap columns (Evosep Biosystems). The bound peptides were washed twice with 100 µl 0.1% formic acid and covered with 200 µl 0.1% formic acid. MS/MS spectra were acquired in positive mode. Raw data were pre-processed with the *Mascot Distiller* software (version 2.4.2.0; Matrix Science); the obtained peptide masses and fragmentation spectra were then matched to the NCBIprot database or user database using the *Mascot* search engine (*Mascot Server* version 2.4.1; Matrix Science). The protein mass was unrestricted; the mass values were treated as monoisotopic with two missed cleavages allowed. Methylation of cysteine was set as fixed, while the oxidation of methionine and, additionally, phosphorylation of serine, threonine and tyrosine were set as variable modifications.

2.5. Crystallization

Crystals of ReAIV were grown at 19°C using vapor diffusion in a hanging-drop setup. Prior to crystallization, ReAIV was concentrated to 16–17 mg ml^{-1}. Five different crystal forms of ReAIV were obtained as follows: the monoclinic crystals R4mC-1 and R4mC-2, the orthorhombic crystals R4oP-1 and R4oP-2, and the tetragonal crystal R4tP. The R4tP and R4mC-2 crystals were grown in 0.2 *M* ammonium formate, 20%(*w*/*v*) PEG 3350 pH 6.7, 0.2% lauryldimethylamine oxide (LDAO). R4mC-1 crystals were obtained after pre-incubation of the protein with 100 m*M* L-Asn (1:50 molar ratio) and crystallization in 0.1 *M* Tris pH 8.5, 0.2 *M* $MgCl_2$, 17.5%(*w*/*v*) PEG 8000, 0.3% dodecyl-β-D-maltoside (DDM). The R4oP-1 and R4oP-2 crystals were obtained after pre-incubation of the protein with 100 m*M* L-Asp (1:50 molar ratio) and crystallization in 0.1 *M* HEPES pH 7.5, 0.2 *M* $MgCl_2$, 17.5%(*w*/*v*) PEG 8000, 0.3% DDM. Prior to X-ray data collection, crystals were cryoprotected with mother liquor supplemented with ethylene glycol and vitrified in liquid nitrogen.

2.6. X-ray data collection, structure solution and refinement

X-ray diffraction data were collected using synchrotron radiation on EMBL beamline P13 at PETRA III, DESY, Hamburg. The diffraction images were indexed and integrated with *XDS* and the intensity data were scaled using *XSCALE* (Kabsch, 2010). Structures were solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007) using ReAV (PDB entry 7os5) as the starting model. Structures were refined with *REFMAC*5 (Murshudov *et al.*, 2011) or *phenix.refine* (Afonine *et al.*, 2012) using anisotropic or TLS protocols (see Table 1). *Coot* (Emsley *et al.*, 2010) was used for model adjustment in the electron-density maps and to divine the solvent structure. The statistics of data collection and structure refinement are summarized in Table 1. The structures of ReAIV were analyzed and visualized in *PyMOL* (DeLano, 2002) and *UCSF Chimera* (Pettersen *et al.*, 2004). Sequence similarity between ReAIV and selected structural homologs was assessed by pairwise sequence alignments in *EMBOSS Needle* (Needleman & Wunsch, 1970; Madeira *et al.*, 2019). The distribution of electrostatic surface potential was calculated for dimers of representative structures of ReAIV (R4oP-1) and ReAV (PDB entry 7os5) using the *APBS-PDB2PQR* software (Jurrus *et al.*, 2018).

2.7. Detection and volume calculation of protein cavities

The cavities within the ReAIV and ReAV proteins were analyzed using the *SPACEBALL* algorithm (Chwastyk *et al.*, 2016; Chwastyk, Jaskolski *et al.*, 2014). *SPACEBALL* is one of several approaches that allow users to detect, define and calculate the volume of cavities within biological structures. Its advantage is that it does not require manual determination of the position of the cavity, as is necessary in many other algorithms. In the definition in *SPACEBALL*, a cavity is a region buried inside the protein structure.

The cavity description starts with placing the considered structure in a cuboid box with walls made of beads on a lattice with constant a. The beads play the role of probes that can travel in all directions. When a probe, which represents a solvent molecule, moves in a given direction it visits lattice points. It stops when it overlaps with any of the van der Waals (vdW) spheres associated with atoms of the protein. The vdW radii are taken from the classic book by Pauling (1960). Different sets of vdW radii (Chwastyk *et al.*, 2016; Zhao *et al.*, 2017; Chwastyk, Galera-Prat *et al.*, 2014) are also available, but for the purpose of this project we only used the radii from Pauling's compilation. The grid points that are not visited by the probes are verified in the next step by checking whether the probe can be placed at a given point without overlapping with protein atoms. All grid points that satisfy these criteria are counted as belonging to the cavity. Their total number, multiplied by a^3, gives the total cavity volume.

Table 1
X-ray data-collection and structure-refinement statistics.

Values in parentheses are for the highest resolution shell.

Structure	R4mC-1	R4mC-2	R4tP	R4oP-1	R4oP-2
Data collection					
Source	P13, PETRA III	P13, PETRA III	P13, PETRA III	P13, PETRA III	P13, PETRA III
Wavelength (Å)	0.9762	0.9770	0.9700	0.9762	0.9763
Temperature (K)	100	100	100	100	100
Space group	*C*2	*C*2	$P4_22_12$	$P2_12_12_1$	$P2_12_12_1$
a, *b*, *c* (Å)	95.76, 85.89, 82.72	96.24, 85.30, 82.95	125.34, 125.34, 115.23	83.20, 89.88, 169.80	83.10, 89.75, 169.20
α, *β*, *γ* (°)	90, 108.72, 90	90, 109.18, 90	90, 90, 90	90, 90, 90	90, 90, 90
Oscillation range (°)	0.20	0.20	0.10	0.20	0.20
No. of images	1800	800	1200	1800	800
Resolution range (Å)	78.35–1.30 (1.38–1.30)	78.34–1.50 (1.59–1.50)	88.63–2.50 (2.65–2.50)	59.42–1.35 (1.43–1.35)	84.60–2.00 (2.12–2.00)
No. of collected reflections	943997 (89972)	299264 (46321)	284391 (45079)	3663476 (277080)	510945 (84068)
No. of unique reflections	150517 (19954)	100056 (16066)	32235 (5074)	578452 (44283)	86068 (13748)
Completeness (%)	96.6 (79.4)	99.1 (99.1)	99.9 (99.7)	99.9 (99.5)	99.7 (99.7)
Multiplicity	6.27 (4.51)	2.99 (2.88)	8.82 (8.88)	13.22 (13.06)	5.94 (6.11)
R_{merge} (%)	7.9 (58.2)	8.8 (84.9)	17.6 (101.6)	11.3 (155.8)	13.5 (111.5)
R_{meas} (%)	8.6 (65.5)	10.7 (104.4)	18.7 (107.9)	11.8 (162.0)	14.8 (122.0)
Wilson *B* factor (Å^2)	19.5	26.7	45.0	13.6	34.8
$\langle I/\sigma(I)\rangle$	14.27 (2.22)	6.60 (1.12)	11.41 (2.16)	15.00 (1.72)	10.22 (1.45)
$CC_{1/2}$	99.8 (79.1)	99.4 (51.7)	99.6 (73.9)	99.9 (72.2)	99.8 (68.1)
Refinement					
Refinement program†	*phenix.refine*	*REFMAC*5	*REFMAC*5	*phenix.refine*	*REFMAC*5
Protein chains in asymmetric unit	2	2	2	4	4
Matthews volume (Å^3 Da^{-1})	2.14	2.27	3.20	2.21	2.24
Solvent content (%)	42.5	45.9	61.6	44.3	45.0
Unique/test reflections	149517/1001	99046/1000	31215/1000	276997/2769	85052/1000
R_{work}/R_{free} (%)	13.3/16.3	25.3/30.2	17.4/21.7	13.0/17.7	17.6/23.4
No. of atoms					
Protein	5370	4931	4939	10652	9912
Metal‡	2	2	2	4 + 1	4 + 2
Solvent	871	377	163	1786	888
B-factor model	Aniso	TLS	TLS	Aniso	TLS
Average *B* factor (Å^2)					
Protein	17.7	24.0	43.6	16.8	32.8
Metal	14.9	25.7	68.8	15.8	38.1
Solvent	34.4	29.6	35.9	31.4	41.6
R.m.s.d., bond lengths (Å)	0.012	0.008	0.008	0.013	0.011
R.m.s.d., angles (°)	1.25	1.48	1.52	1.32	1.73
Ramachandran plot					
Favored (%)	99	100	99	99	100
Allowed (%)	1	0	1	1	0
Outliers (%)	0	0	0	0	0
PDB code	8osw	8clz	8cly	8ori	8col
Raw diffraction data	https://doi.org/10.18150/NOBOWP	https://doi.org/10.18150/WFJHW8	https://doi.org/10.18150/SWISNO	https://doi.org/10.18150/PLH5BS	https://doi.org/10.18150/7RCKMY

† Refinement was carried out in *phenix.refine* (Afonine *et al.*, 2012) or in *REFMAC*5 (Murshudov *et al.*, 2011). ‡ Zinc ions + magnesium ions.

The algorithm uses two parameters: the lattice constant *a* and the probe radius *r*. The former value was set based on a number of tests (Chwastyk *et al.*, 2020; Chwastyk & Cieplak, 2021). The probe radius was set to 1.42 Å, which is close to the size of a water molecule. The final result may depend on the orientation of the structure in the box. To avoid this bias, we repeated the analysis for 25 different rotations of the model, as recommended in Chwastyk, Jaskolski *et al.* (2014) and Chwastyk *et al.* (2016), and averaged the results.

3. Results

3.1. Sequence (dis)similarity of the *R. etli* L-asparaginases

The cytoplasmic constitutive L-asparaginase ReAIV has a sequence of 335 residues and is shorter than ReAV by 32 residues (Fig. 1). A comparison of the two sequences revealed that the shortening of the ReAIV sequence is not an effect of random residue deletions, but rather arises from the removal of three segments from the ReAV sequence (Fig. 1): 211–217, 275–284 and 350–367 (the C-terminal fragment). Alignment of the sequences showed that the residues forming the two catalytic Ser–Lys tandems in ReAV (Loch *et al.*, 2021) are also preserved in the ReAIV sequence. The same observations are made for the residues forming the metal coordination site and the cysteine residue carrying a post-translational modification.

In our previous report (Loch *et al.*, 2021) we presented a topology diagram of the inducible ReAV enzyme, where we proposed a nomenclature for the secondary-structure elements of class 3 L-asparaginases. However, the above comparison of the ReAIV and ReAV sequences strongly suggests that the previous labeling scheme should be modified and that in fact the ReAIV protein should serve as the reference as it is

shorter and does not contain some accessory fragments that are present in the ReAV sequence. For example, in the ReAV sequence, additional short helices H0, H9a, H13a and H15a (labeled blue in Figs. 1 and 2*c*) are present. On the other hand, β-strand B10 of ReAIV (labeled in light green in Fig. 1 and in red in Fig. 2*b*) does not have a counterpart in the ReAV sequence.

3.2. Biophysical properties of ReAIV

The gel-filtration experiment revealed that, similarly to ReAV, the ReAIV protein is dimeric (Supplementary Fig. S1). The nanoDSF experiment confirmed the high thermal stability of ReAIV, with a T_m of 72.2°C (Fig. 2*a*). The T_m for ReAV determined under the same conditions was only 51.1°C. The hydrodynamic radius (R_h) determined by nanoDLS for the ReAIV dimer is 3.49 $\pm$ 0.01 nm; that for the ReAV dimer was 4.00 $\pm$ 0.02 nm. These data indicate that ReAIV has a more compact structure than ReAV, in agreement with the crystallographic data (see Section 3.5).

3.3. Effect of pH on enzyme activity and determination of kinetic parameters of ReAIV at optimum conditions

The effect of pH on the L-asparaginase activity of the constitutive ReAIV and the inducible ReAV (serving as a reference) was determined at pH values ranging from 7.0 to 11.0 using the Nessler method and is shown in Fig. 3(*a*). Both proteins exhibit the highest enzymatic activity at pH 11.0; however, ReAIV has a greater than twofold higher activity than ReAV under these conditions (82 versus 40 m*M* s^{-1} mg^{-1}). ReAIV has an almost 50-fold lower activity at pH 7.0 than at 11.0. A similar, albeit less pronounced, effect was found for ReAV, the activity of which is ~30 times lower at pH 7.0 in comparison to pH 11.0. It was also found that ReAIV retained only 4% of its maximum activity (at pH 11.0) at physiological pH and 10% at pH 8.0, whereas ReAV retained 8% and 16% of its maximum activity (also observed at pH 11.0) at pH 7.5 and 8.0, respectively. Although the activity of ReAIV in the alkaline pH range (9.0–11.0) adjusted with carbonate buffer was over two times higher than the activity of ReAV in the same buffer, this effect was not observed when the alkaline pH 9.0 was adjusted with phosphate buffer. The kinetic parameters determined by ITC-MIM in 10 m*M* carbonate buffer pH 9.0 (average values from three separate experiments) and the Michaelis–Menten equation fit to the sample kinetics data of ReAIV and ReAV are presented in Fig. 3(*b*). ReAIV is a fast enzyme (k_{cat} = 411 $\pm$ 21 s^{-1}) with a K_m for L-Asn of 1.34 $\pm$ 0.02 m*M*.

3.4. Multiple crystal structures of ReAIV

Here, we present the structures of five crystals of ReAIV grown under different crystallization conditions. These structures represent three different crystallographic space groups

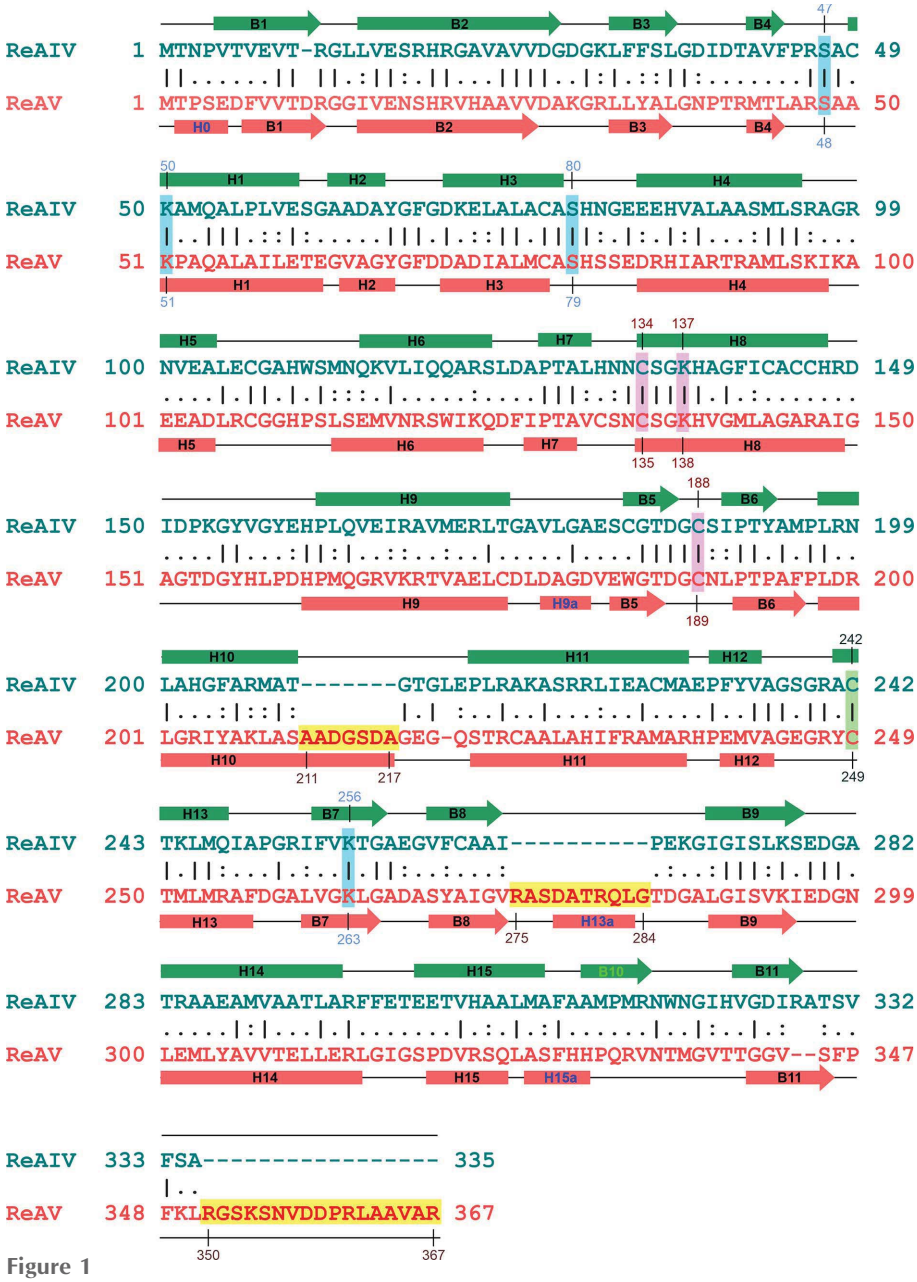


Figure 1
Alignment of the ReAIV and ReAV sequences. The secondary-structure elements are marked by rectangles (helices) and arrows (β-strands). Blue boxes mark the catalytic Ser–Lys tandems. Pink boxes mark residues involved in zinc coordination. The green box marks the cysteine residue with post-translational modification. The yellow color highlights fragments of the ReAV sequence that are absent in the ReAIV protein. The numbers above (ReAIV) and below (ReAV) the sequences mark the positions of important residues (highlighted by colored boxes).

(Table 1). The monoclinic structures R4mC-1 and R4mC-2 contain one ReAIV dimer in the asymmetric unit but differ in the occupancy of the zinc ions, in the occupancies and the conformations of the residues involved in zinc coordination and in the hydration of the nucleophilic Ser47 residue (Supplementary Table S1).

The orthorhombic structures R4oP-1 and R4oP-2 contain two dimers in the asymmetric unit. They differ in Zn^{2+} occupancy and in the conformation of the cysteine residues involved in zinc coordination. In the orthorhombic structure R4oP-1 the dimers form further intermolecular interactions, in which Mg^{2+} ions are coordinated in the vicinity of Glu84 and Glu85 from two symmetry-related dimers. Another Mg^{2+} ion coordination site was found in R4oP-2 in the close vicinity of Asp149 (Supplementary Fig. S2). The tetragonal structure R4tP contains one dimer in the asymmetric unit. It differs from the other structures in the oxidation state of Cys242 and in the conformation of the cysteine residues involved in zinc coordination (Supplementary Table S1).

3.5. Overall fold of the ReAIV molecule and the architecture of the dimer interface

Even though the ReAIV polypeptide is significantly shorter compared with ReAV and has only a low level of sequence identity, the five different crystal structures presented here, with as many as seven independent instances of the ReAIV dimer, all unambiguously revealed that the overall fold of ReAIV follows the topology pattern of ReAV (Loch *et al.*, 2021; Fig. 1). There is, however, an interesting difference in the C-terminal β-structure, which in ReAIV consists of two consecutive β-strands B10 and B11 connected by an extended loop. Strand B10 together with helix H6 is involved in dimer stabilization in ReAIV (Fig. 4*c*). The β-structure B10 was

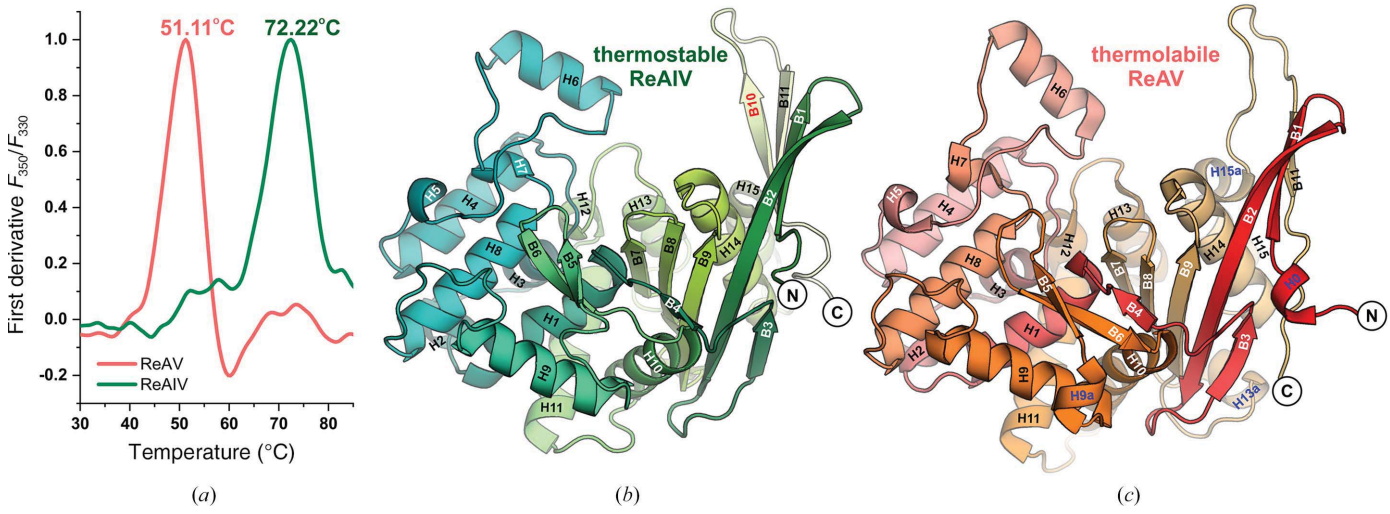


Figure 2
R. etli has two L-asparaginases: thermostable ReAIV and thermolabile ReAV. (*a*) Thermal denaturation profiles of ReAIV and ReAV determined by nanoDSF. (*b*, *c*) Comparison of the subunit structures of (*b*) ReAIV and (*c*) ReAV. As the ReAIV sequence is shorter, some elements present in ReAV are missing in ReAIV, *i.e.* helices H0, H9a, H13a and H15a [labeled in blue in (*c*)]; ReAIV has an additional β-strand B10 [labeled in red in (*b*)].

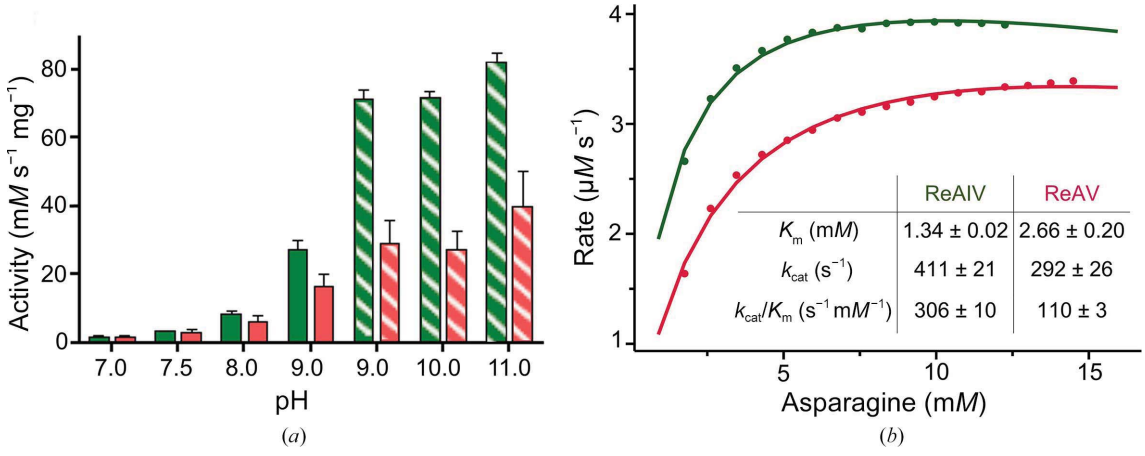


Figure 3
Enzyme activity of *R. etli* L-asparaginases. (*a*) pH dependence of the L-asparaginase activity of ReAIV (green) and ReAV (red) assessed by the Nessler method in the presence of 10 m*M* L-Asn. Filled bars correspond to phosphate buffer, whereas hatched bars correspond to carbonate buffer. (*b*) Michaelis–Menten fit to sample ITC-MIM kinetic data measured in 10 m*M* carbonate buffer pH 9.0. Kinetic parameters for both enzymes (calculated as the average values derived from three separate experiments) are collected in the inset.

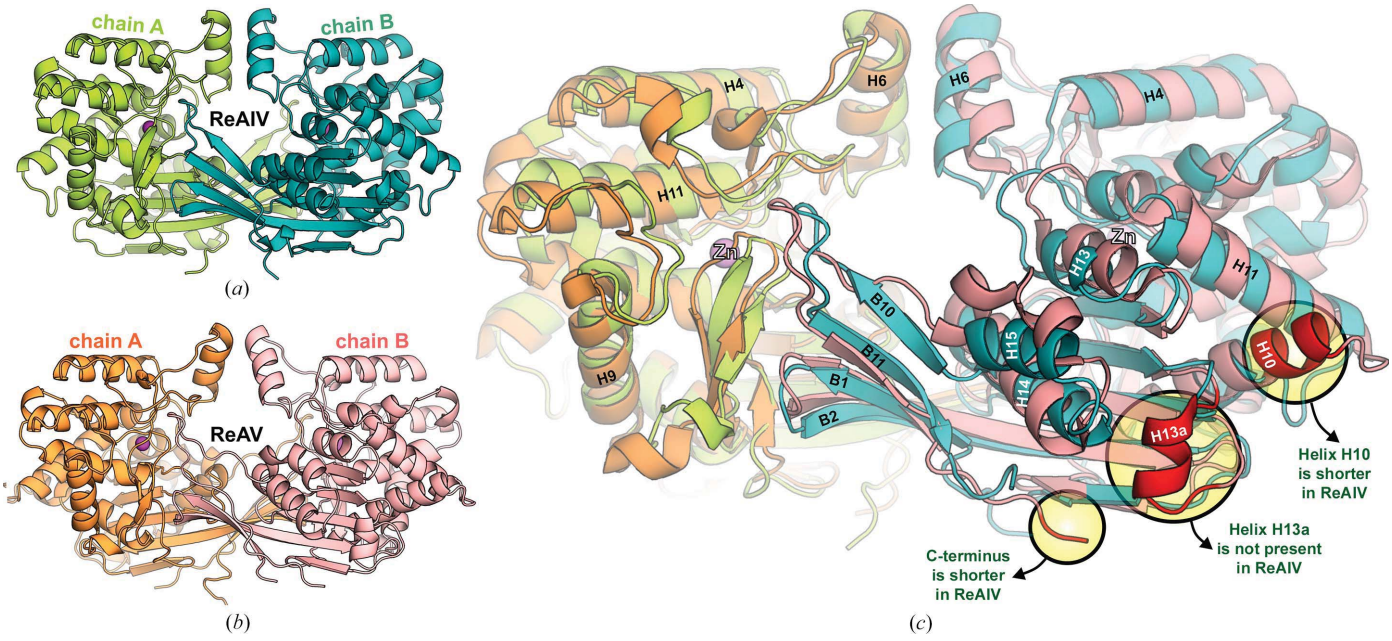


Figure 4
Comparison of ReAIV and ReAV dimers. The zinc ion located in the active site is marked by a light violet sphere. (*a*) ReAIV dimer (chain *A*, lime; chain *B*, dark green; structure R4tP). (*b*) ReAV dimer (PDB entry 7os5; chain *A*, orange; chain *B*, pink). (*c*) Superposition of ReAIV and ReAV dimers; the sequence fragments that are present in ReAV but absent in ReAIV are colored red and marked with yellow circles. The absence of these fragments makes the ReAIV dimer more compact.

physically absent in the topology diagram of ReAV and was represented by an unstructured chain (Loch *et al.*, 2021). As a consequence, the very terminal β-strand of ReAV was annotated B10, while it should have been labeled B11, with B10 missing in the topology diagram of this prototypical class 3 L-asparaginase. It is therefore interesting that the shorter and more compact ReAIV protein can serve to amend the canonical topology diagram of class 3 L-asparaginases (see Section 3.1).

Nevertheless, all of the determined structures show that the biological assembly of ReAIV is a C_2-symmetric homodimer with two identical active sites accentuated by coordinated zinc ions (Figs. 4 and 5). The structure of the ReAIV dimer is similar to the previously described dimer of ReAV. The dimer interface of ReAIV is stabilized by hydrogen bonds formed between residues from β-strands B2 and helices H6 (Fig. 4). However, in comparison with ReAV the structure of the ReAIV dimer is more compact due to the presence of the additional β-strand B10, which is absent in ReAV. The ReAIV dimer is also much more

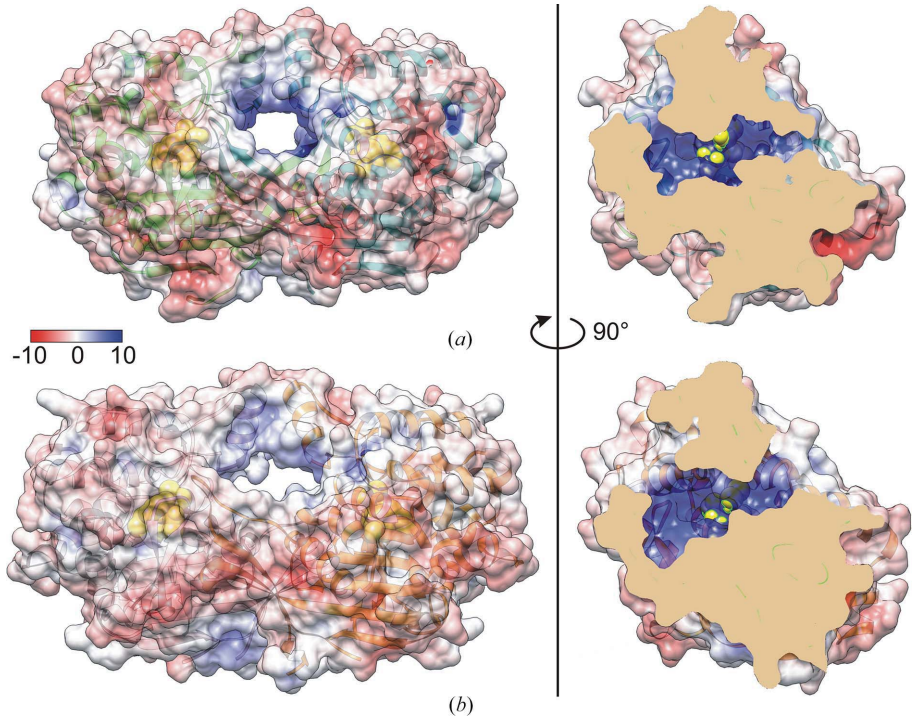


Figure 5
Comparison of the electrostatic potential surfaces of ReAIV and ReAV. The electrostatic potential mapped onto the surface of (*a*) the ReAIV dimer (R4oP-1) and (*b*) the ReAV dimer (PDB entry 7os5). The panels on the left present the overall view of the dimers oriented as in Fig. 4. The panels on the right are cutaway side views (after 90° rotation around the vertical line) looking into the active site from the direction of the complementary subunit. Strong positive potential (blue) is visible in the cleft of the active-site area. The solid yellow spheres represent the active-site residues Ser47/48, Lys50/51, Ser79/80 and Lys256/263 in ReAIV/ReAV. The electrostatic potential was calculated using the *APBS*/*PDB2PQR* software (Jurrus *et al.*, 2018) for pH 8.0 with a range from $-10\ kT$ (red) to $+10\ kT$ (blue).

tightly bound, as the dimer interface is stitched by as many as ten salt bridges, compared with only six salt bridges in the ReAV dimer interface (Table 2). Being of predominantly electrostatic character, reinforced by hydrogen bonding, salt bridges are the strongest associative intermolecular interactions. The interface salt bridges are undoubtedly an important factor contributing to the thermal stability of the ReAIV protein.

The calculated electrostatic potential mapped onto the dimer surfaces of ReAIV and ReAV shows regions of strong positive charge in the central part of each protomer (Fig. 5) in the area of the active site. This area is surrounded by several arginine and lysine residues, which are suitable for attracting negatively charged moieties (such as the α-carboxylate group of the substrate) into the active site.

3.6. The architecture of the active site: conformation and hydration of the serine residues

The active site with the zinc coordination sphere and the catalytic Ser–Lys tandems is located in the central part of each protomer. In ReAIV, the catalytic tandems are formed by Ser47–Lys50 and Ser79–Lys256 (Fig. 6*b*). While the conformations of Lys50 and Lys256 are conserved in all of the structures, Ser47 and Ser79 assume alternative conformations (Fig. 6*a*) in some cases, accompanied by different hydration patterns. A dual conformation of Ser79 is observed in all subunits of the R4oP-1 structure (Fig. 6*a*), while in all other cases no additional conformation of Ser79 was detected. The major conformer (*A*) of Ser79 has its O^{γ} atom hydrogen-bonded to the N^{ζ} atoms of Lys50 and Lys256, while the minor rotamer (*B*) of Ser79 is hydrogen-bonded to the side chain of Asp133 (Fig. 6*a*) and directed towards the protein surface (specifically towards Met112, which is located at the entrance to the active site).

In the R4mC-2 and R4oP-1 structures, high difference electron-density peaks were found around the side chain of Ser47, suggesting strong coordination (O···O distances of ∼2.3–2.4 Å) of three water molecules (Fig. 6*a*) or possible phosphorylation of this hydroxyl group. A similar situation was previously found in some ReAV structures (Loch *et al.*, 2021). As mass-spectrometric analysis showed the absence of Ser47 phosphorylation in the ReAIV molecule beyond any doubt, we must interpret the electron-density peaks close to the Ser47 OH group as water molecules, possibly mimicking the substrate molecule (or a covalent intermediate) bound to the nucleophilic Ser47 residue during the catalytic event.

In the R4mC-2 structure, Ser47 (with tightly bound water molecules) has a single conformation, while in the R4oP-1 structure two alternative conformations of Ser47 were found (Fig. 6*a*): a higher occupancy conformer *A* with tightly bound water molecules and the O^{γ} atom directed towards the zinc ion, and a lower occupancy unhydrated conformer *B* with the O^{γ} atom hydrogen-bonded to the carbonyl O atom of Thr257. In the R4oP-2 structure, a single conformation (*A*) of unhydrated Ser47 is observed in three of the four subunits. A single conformation of Ser47 without strongly associated water molecules was also observed in R4tP. In the R4oP-2 and R4tP structures, Ser47 formed several O···O hydrogen bonds (∼2.4–3.0 Å) to other water molecules located in the close vicinity. It is of note that the catalytic Ser48 of ReAV can also assume two conformations in cases when it is not hydrated (Loch *et al.*, 2021). Importantly, in all of the discussed structures we observe a hydrogen bond (with an O···O distance in the range 2.5–3.2 Å) between the nucleophilic Ser47 (in conformation *A*) and water molecule w1 from the zinc coordination sphere (Figs. 6*c* and 6*d*).

Table 2
Comparison of the hydrogen bonds between subunits *A* and *B* at the dimer interfaces of ReAIV (structure R4mC-1; 24 hydrogen bonds and ten salt bridges) and ReAV (PDB entry 7os5; 25 hydrogen bonds and six salt bridges).

Distances in bold indicate salt bridges.

A···*B*/*B*···*A* (Å)					
ReAIV			ReAV		
Arg11 $N^{\eta1}$	2.77/2.76	Thr185 O	Arg12 $N^{\eta1}$	2.90/2.91	Thr186 O
Ser17 O^{γ}	2.69/2.69	Glu260 $O^{\varepsilon1}$	Glu17 O	2.98/2.95	Arg42 $N^{\eta2}$
Ser17 O^{γ}	3.01/3.02	Lys277 N^{ζ}	Glu17 O	3.10/3.13	Lys294 N^{ζ}
Arg18 N	2.89/2.97	Glu279 O	Asn18 $N^{\delta2}$	3.00/2.96	Asp267 $O^{\delta1}$
Glu105 O	2.83/2.85	Asn320 $N^{\delta2}$	Asn18 $O^{\delta1}$	2.87/2.85	Lys294 N^{ζ}
Ala121 O	2.94/2.89	Arg122 N^{ε}	Ser19 N	3.08/3.14	Glu296 O
Arg122 $N^{\eta2}$	2.84/2.93	Leu124 O	Ser19 O^{γ}	2.62/2.66	Glu296 $O^{\varepsilon2}$
Arg122 $N^{\eta2}$	3.05/2.99	Asp125 O	Gly108 N	3.13/3.04	Thr336 $O^{\gamma1}$
Asp186 O	3.05/3.10	Asn318 $N^{\delta2}$	Lys123 N^{ζ}	NA†/3.04	Ile122 O
Ser189 O^{γ}	2.89/2.90	Asn318 $N^{\delta2}$	Asp187 O	3.17/3.13	Asn335 $N^{\delta2}$
Ser189 O^{γ}	2.93/2.96	Asn320 $N^{\delta2}$	Gly188 O	2.63/2.75	Thr336 $O^{\gamma1}$
Asp280 O	3.24/3.09	Arg284 $N^{\eta2}$	Gly188 O	3.04/3.01	Thr336 N
Glu16 $O^{\varepsilon1}$	**3.39/3.32**	Arg46 $N^{\eta1}$	Asn190 $O^{\delta1}$	2.99/2.97	Asn335 $N^{\delta2}$
Glu16 $O^{\varepsilon2}$	**2.82/2.85**	Arg46 $N^{\eta1}$	Glu17 $O^{\varepsilon1}$	**3.17/3.19**	Arg47 $N^{\eta1}$
Arg20 N^{ε}	**2.92/2.90**	Glu279 $O^{\varepsilon2}$	Glu17 $O^{\varepsilon2}$	**2.82/2.93**	Arg47 $N^{\eta1}$
Glu260 $O^{\varepsilon1}$	**2.85/2.89**	Arg284 $N^{\eta2}$	Lys123 N^{ζ}	**2.71/2.98**	Asp125 $O^{\delta1}$
Glu260 $O^{\varepsilon2}$	**2.77/2.79**	Arg284 $N^{\eta1}$			

† The hydrogen bond between subunits *A*···*B* is absent due to a different side-chain orientation of Lys123*A*.

3.7. The zinc coordination site: occupancy of the metal ion and conformation of the cysteine residues

The zinc coordination sphere in ReAIV is formed by Lys137, two cysteine residues, Cys134 and Cys188, and water molecule w1 (Fig. 6*c*). Water w1 is hydrogen-bonded to the nearby side chains of Asp186 and Ser47. However, we note that in the gallery of 14 independent views of the ReAIV active site presented in this work, not all of the active sites are fully occupied by the Zn^{2+} ion. When the metal occupancy is less than one, alternative conformers of Cys134 and/or Cys188 are also observed, with the additional possibility of direct Cys134–Cys188 disulfide-bridge formation (Fig. 6*c*). The electron-density maps of the high-resolution structure R4oP-1 revealed that partial removal of Zn^{2+} creates a void that is filled by two water molecules, with a specific network of hydrogen bonds: one water (w1a) interacts with Lys50 and Asp133, while the second water (w1b) is hydrogen-bonded to Lys137, Asp186 and His138 (Fig. 6*c*).

We also observed that the conformation of Asp186 is correlated with the presence or absence of the Zn^{2+} ion. When Zn^{2+} is present, the side chain of Asp186 (conformer *A*) is

rotated towards Arg148 and hydrogen-bonded to water w1 from the zinc coordination sphere (Fig. 6*e*). When Zn^{2+} is absent, the side chain of Asp186 (conformer *B*) is flipped into a position that allows its $O^{\delta1}$ atom to form hydrogen bonds to the main-chain N atoms of Gly187, Cys188, Ser189 and Ile190 (Figs. 6*e* and 6*f*). It appears that this alternative position of Asp186, which is assumed to be cooperative with S–S bridge formation between Cys134 and Cys188, stabilizes the zinc coordination region in the absence of the Zn^{2+} ion. Interestingly, when the Zn^{2+} ion is absent the Lys137 side chain in its coordination sphere does not change its conformation and position in most of the structures and remains hydrogen-bonded to Gln53 and neighboring water molecules.

3.8. An oxidized cysteine flanking the active site

In most of the structures, we observed that (analogously to Cys249 of ReAV) Cys242 is oxidized to sulfenic acid and adopts two conformations of the side chain: *A*, directed towards Lys256, and *B*, directed towards Arg240 (Fig. 6*g*). The former rotamer is engaged in a water-mediated hydrogen bond to the N^{ζ} atom of Lys256, the carbonyl O atoms of Ser79 and Val235, and the $N^{\eta2}$ side-chain atom of Arg240. In the R4tP structure, Cys242 is not oxidized (Fig. 6*h*) and its S^{γ} atom is hydrogen-bonded to the main-chain N atom of Thr257 and two water molecules, which mediate interactions with the N^{ζ} atom of Lys256 and the $N^{\eta2}$ and N^{ε} atoms of Arg240.

3.9. Occluded water molecules

Careful inspection of the electron-density maps at all stages of structure refinement revealed the presence of three clusters of structural water molecules buried in three internal cavities in the protein core. The first internal cavity (cavity I), with a volume of 93 ± 2 Å^3, is found near Phe141 and the second cavity (cavity II; 69.8 ± 0.3 Å^3) is near Phe254, while the largest cavity, cavity III (149 ± 1 Å^3), is near Gln53 (Fig. 7*a*). Cavity I is surrounded by Lys137, Lys50, Ala54, Gln53, Pro56, Ala76, Leu57, Gln53, Ile166, Leu162 and Phe141 (Fig. 7*c*) and cavity II is surrounded by Ala48, Ala51, Cys77, Phe204, Cys228, Val235, Met246, Val255, Lys256 and Phe254 (Fig. 7*c*), while cavity III is lined by Pro45, Cys49, Gln53, Lys137, Gln163, Ile166, Arg167, Met170, Asp186, Ile190, Pro191, Thr192, Tyr193 and Met195 (Fig. 7*b*). In cavities I and II there

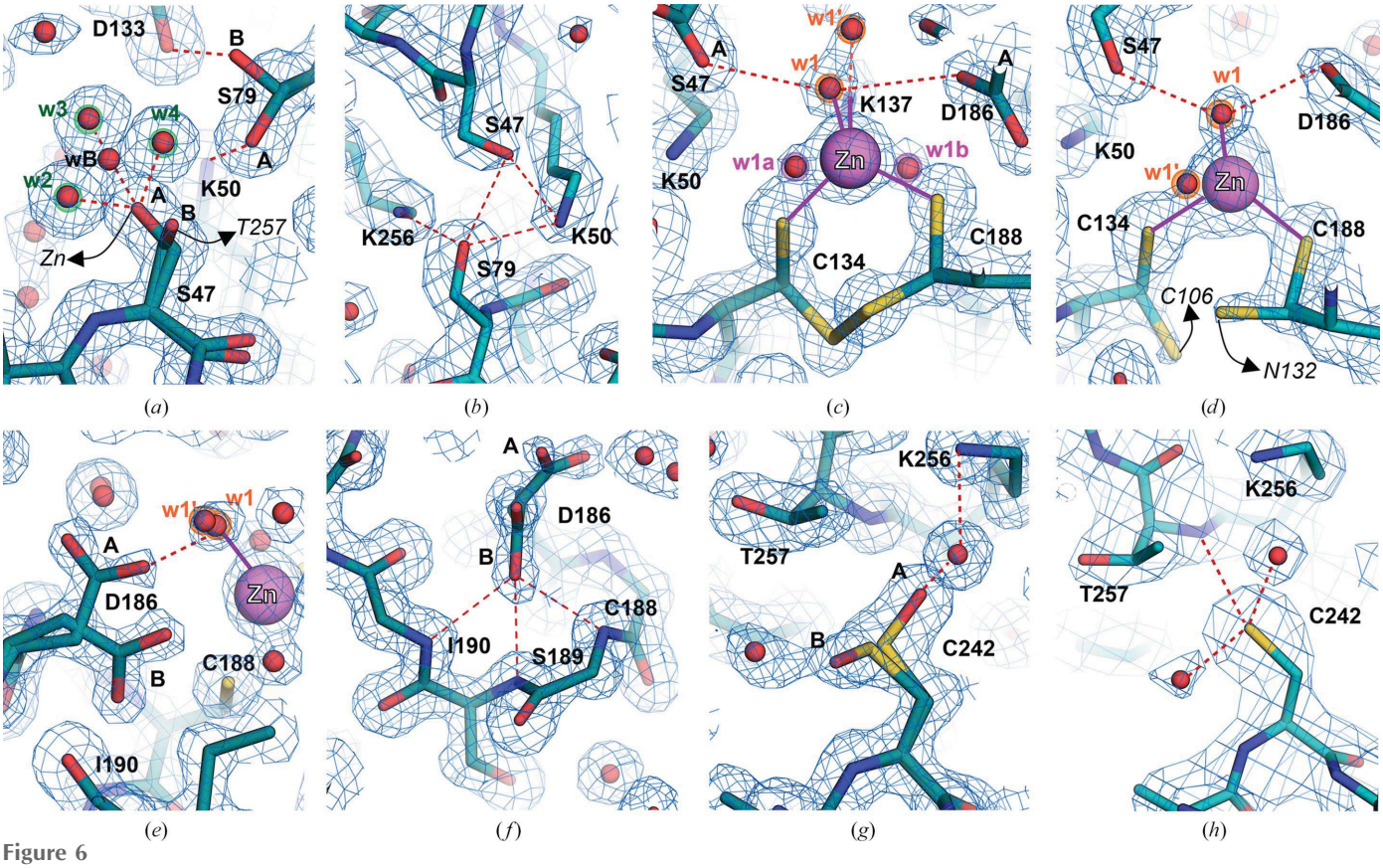


Figure 6
$2F_o - F_c$ electron-density maps (contoured at 1.0σ) for different ReAIV structures. (*a*) Ser47 and Ser79 in two conformations (*A*/*B*) in structure R4oP-1 (chain *D*); water molecules w2, w3 and w4 are hydrogen-bonded to conformer *A* of Ser47, while water wB appears together with conformer *B* of Ser47. (*b*) The catalytic tandems Ser47–Lys50 and Ser47–Lys256 in the active site of structure R4oP-2 (chain *B*). (*c*) The zinc coordination sphere and its close neighborhood in structure R4oP-1 (chain *D*). (*d*) The zinc coordination site and the dual-conformation Cys134 and Cys188 residues in structure R4mC-2 (chain *B*). (*e*) Double conformation of Asp186 in structure R4oP-1 (chain *D*). (*f*) Interactions of conformer *B* of Asp186 in structure R4oP-1 (chain *D*). (*g*) Double conformation of the oxidized Cys242 in structure R4oP-1 (chain *D*). (*h*) Single conformation of Cys242 in structure R4tP (chain *A*). In all figures, red dashed lines mark hydrogen bonds, while solid violet lines mark coordination bonds. Water molecules are shown as small red balls and the Zn^{2+} cation is shown as a purple sphere.

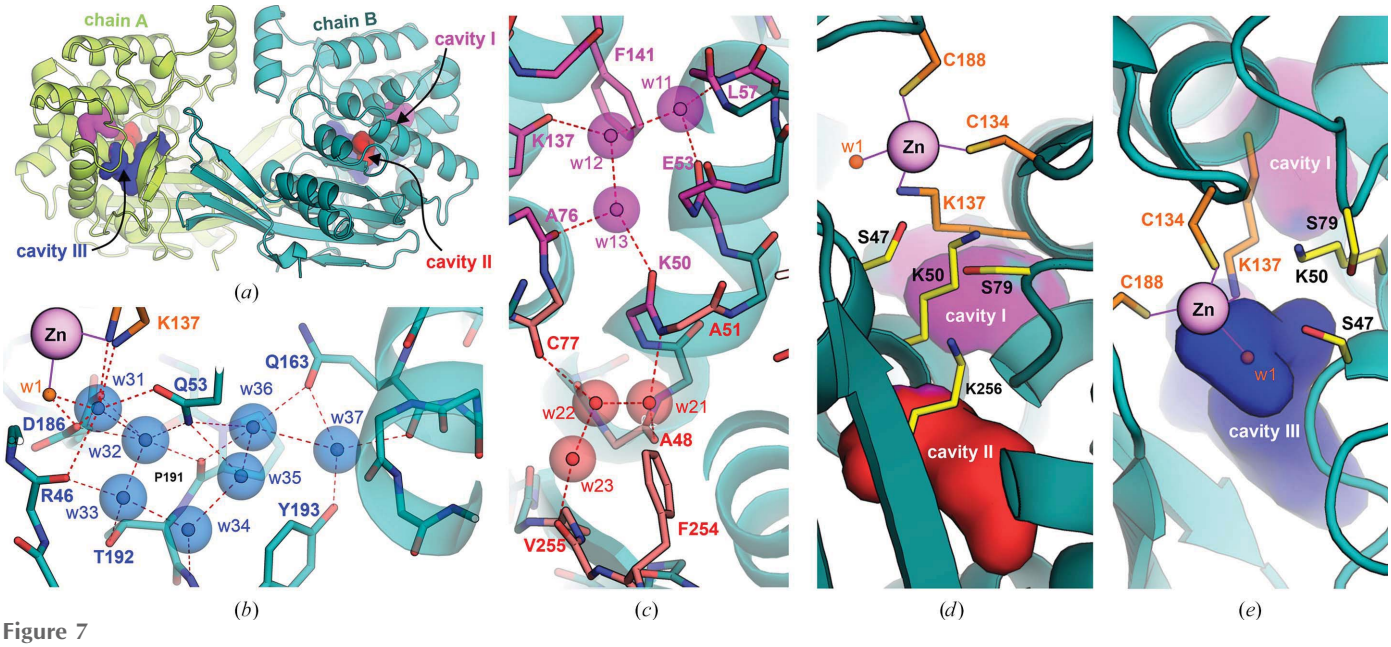


Figure 7
Structural water molecules in the hydrophobic core of ReAIV. (*a*) Three cavities, shown in both subunits of the ReAIV dimer, are found in the core of each ReAIV subunit: near Phe141 (cavity I, magenta), near Phe254 (cavity II, red) and near the Zn^{2+} cation (cavity III, blue). (*b*) Cavity III is occupied by seven structural water molecules w31–w37. (*c*) The water molecules in cavity I near Phe141 (w11–w13) and cavity II near Phe254 (w21–w23) are hydrogen-bonded to each other and to the protein backbone atoms but do not interact with the side chains. Although cavities I and II are separated from each other, the network of hydrogen bonds involving the occluded water molecules extends along the hydrophobic core of the protein, stabilizing the internal, mostly hydrophobic region of the ReAIV molecule. (*d*) Location of cavities I and II with the occluded water molecules in relation to the catalytic Ser–Lys tandems and the zinc coordination site (the Zn^{2+} cation is in purple and the coordinated water molecule w1 is in orange). (*e*) Location of cavity III in relation to cavity I and the Zn^{2+} cation coordinated near the catalytic Ser–Lys tandems.

are three water molecules apiece, hydrogen-bonded to each other and to the protein main chain only (Fig. 7*d*, Supplementary Table S3). In cavity III there is a chain of seven hydrogen-bonded water molecules, with additional hydrogen-bonding contacts to both the main chain and side chains of protein residues located in the close vicinity (Fig. 7*b*, Supplementary Table S3). Interestingly, the side chain of Asp186 bridges water w32 in this cavity to water w1 from the zinc coordination sphere. The water molecules in all three cavities act as 'drops of glue', cementing the protein core. Thus, all of these water molecules are likely to play an important role in the increased thermal stability of ReAIV. Interestingly, all water-filled cavities in ReAIV are located very close to regions that are important for the catalytic apparatus, namely the zinc coordination site (cavities I and III) and the two catalytic Ser–Lys tandems (cavities I and II; Figs. 7*d* and 7*e*). It could be envisioned that the three water-filled cavities act as stabilizing 'water cushions' that protect the active site from damage.

4. Discussion

4.1. Structural comparison of ReAIV and ReAV

The levels of sequence identity and similarity between ReAIV and ReAV are rather low (31.1% and 44.6%, respectively, as determined using the *EMBOSS Needle* alignment algorithm; Needleman & Wunsch, 1970) but, nevertheless, the overall folds of the thermostable L-asparaginase ReAIV and its thermolabile counterpart ReAV (Loch *et al.*, 2021) are very similar (C^{α} r.m.s.d. of 1.38 Å). Both enzymes also display a similar architecture of the dimer, although the ReAIV dimer interface is stabilized by the additional β-strand B10 that is not present in the ReAV molecule and by a significantly higher number of salt bridges (Table 2). Although the sequence of the constitutive ReAIV enzyme is markedly shorter, the deletions (with respect to ReAV) have only a minor effect on the topology of the polypeptide chain. The most conspicuous deletions are visible as a shortening of the solvent-exposed helix H10 and a surface loop carrying the small helix H13a, as well as truncation at the C-terminus of the protein (Fig. 1). As a consequence of the reduction of these peripheral elements, the ReAIV structure is more compact in comparison with ReAV, in agreement with the nanoDLS measurements (Supplementary Fig. S1).

Similarly to ReAV, the active site of ReAIV is located in the central part of a subunit. Despite the different sequence lengths, the catalytic sites of the two proteins are identical and are built around two catalytic tandems, Ser47–Lys50 and Ser79–Lys256 in ReAIV (Ser48–Lys51 and Ser80–Lys263 in ReAV). A zinc ion is located in the close neighborhood of these tandems in both enzymes. The coordination sphere of Zn^{2+} in ReAIV is the same as in ReAV and consists of two cysteine residues, a lysine residue and a water molecule (Supplementary Fig. S3*a*). It is therefore a so-called 'open' coordination sphere with a single water molecule that could potentially be activated for a catalytic process (McCall *et al.*, 2000; Zastrow & Pecoraro, 2014).

Although the architecture of the active sites of the two proteins is very similar, some differences are discernible at the entrance to the catalytic pocket. For example, in ReAIV a bulky Trp319 side chain (contributed by the complementary subunit of the dimer) is present close to the Zn^{2+} coordination site, while in ReAV the corresponding position is occupied by the much smaller Thr336 (Supplementary Fig. S3*c*). Furthermore, in ReAIV the active-site entrance is guarded by Trp110 and Met112, whereas in ReAV the less bulky Pro111 and Leu113 are found at these positions (Supplementary Fig. S3*b*). Additional differences are observed in the region of the oxidized cysteine residue (Cys242 in ReAIV), which in ReAIV is surrounded by Thr257, Glu287 and Phe263, while the corresponding positions in ReAV are occupied by Leu264, Tyr304 and Tyr270, respectively (Supplementary Fig. S3*d*).

Intriguingly, despite its shorter sequence and more compact overall structure, the ReAIV fold is still able to enclose three internal cavities (I, II and III) filled with occluded water molecules (Fig. 7). Cavities I, II and III do not have obvious counterparts in the structure of the thermolabile ReAV enzyme. Superposition of the ReAIV molecule, for example chain *A* from structure R4mC-2, on the PDB model 7os5 of ReAV (chain *A*) revealed that cavity I is present in ReAV but is significantly reduced in size and accommodates only a single water molecule corresponding to w13 in ReAIV (Supplementary Fig. S4*b*). The volume of cavity I in ReAV is reduced due to positional shifts of helices H1 (by ∼1.5 Å) and H8 (by ∼2.5 Å) with respect to their positions in ReAIV. The presence of aliphatic side chains in ReAV, Met77 (instead of Ala76 in ReAIV), Met142 (instead of Phe141) and Ile58 (instead of Pro56), additionally constricts the space in the region corresponding to cavity I in ReAIV (Supplementary Fig. S4*a*). Cavity II is absent in the ReAV molecule altogether, as the space available in this region is not only constricted by positional shifts (by ∼1.0–1.5 Å) of helices H1, H11 and H13, as well as β-strands B5, B6 and B7, but also by the presence of Tyr205 (instead of Phe204 in ReAIV) and Met235 (instead of Cys228). Remnants of cavity III can be recognized in the ReAV molecule, but the volume is smaller due to the presence of more bulky residues, for example Trp184 instead of Tyr193 (Supplementary Fig. S4*c*). These observations indicate that the interior of ReAIV is stabilized by occluded water molecules that contribute to the increased thermal stability of this constitutive ReAIV enzyme.

4.2. A gallery of ReAIV structures

In this work, we report five different crystal structures of ReAIV (Table 1): two monoclinic structures (R4mC-1 and R4mC-2) with a dimer in the asymmetric unit, two orthorhombic structures with two dimers in the asymmetric unit (R4oP-1 and R4oP-2) and one tetragonal structure (R4tP) with a dimer in the asymmetric unit. In the orthorhombic structures, a magnesium cation was found to be coordinated between the two dimers in the asymmetric unit, near Glu84 and Glu85 at the protein surface (Supplementary Fig. S2). In total, these structures present snapshots of 14 ReAIV chains assembled into seven dimers, with different conformational states of Ser47, Ser79, Cys134, Cys188 and Asp186 as well as different conformational and oxidation states of Cys242 (Supplementary Table S1).

In structure R4oP-1 (1.35 Å resolution), very high quality electron-density maps showed a similar conformational pattern in all four protein chains *A*, *B*, *C* and *D*, namely a 0.5 occupancy of the Zn^{2+} ion and of the cysteine side chains involved in its coordination, a 0.5 occupancy of the S–S bridge formed by these cysteines in the alternative conformation, half occupancy of Asp186 and additionally double conformational states of Ser47 and Ser79 from the Ser–Lys tandems. Ser47 was refined with a higher occupancy (∼0.75) hydrated major rotamer *A* and a lower occupancy (∼0.25) minor rotamer *B*, which is not hydrated. In this structure, Cys242 is oxidized and has two equally occupied conformers. A different picture of the active site was observed in the tetragonal structure R4tP (2.50 Å resolution), where in both chains (*A* and *B*) the zinc site is fully occupied, Ser47 has a single conformation with no hydration and Ser79 and Asp186 also exist in single conformations (Supplementary Table S1). Structure R4mC-2 (1.50 Å resolution) is similar to R4tP, the only differences being Ser47 hydration (in a single conformation) and the presence of two rotamers of the oxidized Cys242.

Structures R4mC-1 and R4oP-2 illustrate different variations of the conformational states defined in structures R4oP-1, R4tP and R4mC-2. In the monoclinic structure R4mC-1 (1.30 Å resolution) we observed half occupancy of the zinc cation, an S–S bond between the Cys134 and Cys188 residues that are not involved in zinc coordination, a dual conformation of Asp186, two conformations of Ser47, both without the tightly bound water molecules w2–w4, and a single conformation of Ser79. In structure R4oP-2 (2.00 Å resolution) there are again no water molecules hydrating the side chain of Ser47 and an alternative conformation *B* of Ser47 was only found in protein chain *C*. In the R4oP-2 structure the occupancies of the Zn^{2+} cations are 0.6–0.7. Although some fraction of the ReAIV molecules in this structure are not occupied by Zn^{2+}, for the cysteines that are not involved in zinc coordination an S–S bond was observed only in chains *A* and *B*, while in chains *C* and *D* the distance between the S atoms (2.90 and 2.95 Å) was too long for S–S bridge formation (Supplementary Table S1).

The R4mC-1, R4oP-1 and R4oP-2 crystals were grown after pre-incubation of the ReAIV protein with the L-Asn substrate in an attempt to generate an enzyme–product (or enzyme–substrate) complex. Unfortunately, this main goal has not been achieved, as no convincing electron density for an L-Asp/L-Asn ligand was found near the active site or in any other region of the protein. However, the process of pre-incubation of the enzyme with L-Asn (and ultimately with the L-Asp product of the hydrolysis reaction) had a decisive effect on the crystallization outcome, as in all of these crystals the occupancy of the Zn^{2+} cation decreased from 1.0 (as found in crystals grown without L-Asn additive) to ∼0.5–0.7. It may therefore be hypothesized that the partial occupancy of the Zn^{2+} cation could somehow be related to the catalytic process

that has occurred during the incubation process. One could envision, for example, that at sufficiently high concentration the L-aspartic acid product could start a precipitation reaction with the Zn^{2+} of the poorly water soluble zinc aspartate (Cortex Chemicals; https://cortexch.com/produkt/zinc-aspartate/; Ishizuka *et al.*, 1973). This would simultaneously explain why we did not observe any L-Asp bound in the active site of ReAIV. For an enzyme that critically depends on the Zn^{2+} cation, such a suicide cofactor–product crystallization would be disastrous. However, as shown for ReAV, the Zn^{2+} cation plays a structural rather than a catalytic role and is dispensable for the L-asparaginase reaction. Based on the similarity of the active sites of ReAV and ReAIV, we can reasonably assume that the situation in ReAIV is analogous. Moreover, the crystal structures of ReAIV elegantly demonstrate that the active site does not fall apart even in the absence of the Zn^{2+} cation, because under oxidizing conditions the two cysteine residues in the zinc coordination sphere (Cys134 and Cys188) will form a disulfide bond which can stabilize the active-site region just as well. Parenthetically, it is interesting to recall that a mutual replacement of a structural metal coordination with a structurally equivalent disulfide bridge is known from class 2 L-asparaginases. Specifically, the role of the sodium-binding (stabilization) loop of EcAIII (Michalska & Jaskolski, 2006) is played by a similarly located S–S bond in human aspartylglucosaminidase (Riikonen *et al.*, 1996).

4.3. Functional states of the ReAIV molecule

Analysis of the five crystal structures of ReAIV determined in this work, in which as many as 14 snapshots of the protein chain and the active site are visualized (Supplementary Table S1), allowed us to identify several distinct conformational states of the active-site residues and to arrange them in a sequence illustrating the changes of the enzyme during a conjectural catalytic cycle. Although our classification of the functional states of the ReAIV enzyme might be arbitrary, we postulate that the principal state of the enzyme is the 'resting state', in which the zinc-binding site is fully saturated with Zn^{2+} cations and the side chain of Ser47 is hydrogen-bonded to water w1 from the zinc coordination sphere (Fig. 8*a*). In the 'resting state' the enzyme is between the catalytic events and is ready to accept a substrate molecule.

At the beginning of the catalytic cycle, when the substrate approaches the active site, the nucleophilic Ser47 must become polarized for nucleophilic attack, and thus the ReAIV molecule goes into the 'active state'. The most conspicuous symptom of Ser47 polarization is the strong association of three water molecules (w2, w3 and w4) with its hydroxyl group (Fig. 8*b*). As the distance between water w1 from the Zn^{2+} coordination sphere and O^{γ} of Ser47 decreases when Ser47 is polarized, it might be hypothesized that the potential role of water w1 is to contribute to the polarization of the nucleophilic Ser47 residue. It is worth noting that despite numerous attempts we have been unable to crystallize an ReAIV–product/substrate complex. However, in the previous analysis of ReAV we showed that the positions of waters w2, w3 and w4 superpose well with the covalent intermediate of serine β-lactamases, which have an analogous active site built around two Ser–Lys tandems (Loch *et al.*, 2021). After the catalytic event, when the reaction product leaves the active site, the distance between Ser47 O^{γ} and Zn^{2+} increases and the enzyme returns to the 'resting state'.

The Ser47 residue in ReAIV can change its conformation as it rotates away from water w1 in the Zn^{2+} coordination sphere. This movement of Ser47 is associated with a simultaneous rotation of Ser79, leading to a 'relaxed state' in which the enzyme is not ready to initiate a new catalytic event. As can be inferred from analysis of the occupancy factors of Ser47 and Ser79, for example in the R4oP-1 structure (Supplementary Table S1), the presence of alternative conformers of Ser47 and Ser79 is not related to the presence (occupancy) of the Zn^{2+} cation. It can therefore be assumed that the conformational changes of Ser47 and Ser79 occur independently of the presence of the metal cation near the active site.

Zinc-binding sites with cysteine residues in the coordination sphere may be involved in various physiological processes, such as redox (oxidative) activation, enzymatic catalysis, protein inhibition or structure stabilization (Maret, 2004; Maret & Li, 2009). The Zn^{2+} coordination sphere of ReAIV (and also of ReAV) is 'open', *i.e.* contains a water molecule (w1), which might suggest a catalytic role. However, from the previous analysis of ReAV a rather different picture emerged, namely that the role of Zn^{2+} would be structural rather than catalytic and that the role of water w1 would be to polarize the Ser47 nucleophile (Loch *et al.*, 2021; see below). In most of the ReAIV subunits described here we observed fractional occupancy of the Zn^{2+} cation. This phenomenon is clearly correlated with the pre-incubation process of ReAIV with L-Asn prior to the crystallization experiment (see Section 4.2)

The situation in which the Zn^{2+} ion has partially dissociated from its binding site of ReAIV could be called an 'intermediate state' (Fig. 8*c*) because at this stage significant conformational changes are triggered at the zinc-binding site. The cysteine residues involved in zinc coordination (Cys134 and Cys188) assume alternative conformations; namely, Cys134 rotates towards the nearby Cys106, while Cys188 rotates towards Asn132. Analysis of the crystal structures (especially of atomic occupancies) suggests that the movement of the cysteine residues starts at Cys188, while the side chain of Cys134 is rotated in the next step.

When the Zn^{2+} ion has completely dissociated from the active site, the ReAIV enzyme is most likely to enter an alternative 'resting state' (Fig. 8*d*) in which an S–S bridge between Cys134 and Cys188 is formed. This bridge stabilizes the active-site region in the absence of the Zn^{2+} ion. In this interpretation, the ReAIV enzyme with an S–S bond instead of the Zn^{2+} cation can imitate the same functional states as the enzyme with a fully occupied Zn^{2+} site. To differentiate these S–S states from those described in the presence of zinc, we label them 'S–S-resting', 'S–S-active' and 'S–S-relaxed' in Supplementary Table S2. Interestingly, in the close vicinity of Cys134 and Cys188 there is another cysteine residue, Cys106, located at a distance that allows the formation of disulfide

bonds with each of the cysteine residues from the Zn^{2+} coordination sphere. However, in none of the analyzed ReAIV structures was such a conformational change of Cys106 observed. This might suggest that this residue (which is conserved in ReAV) only plays the role of a 'succorer', protecting the integrity of the zinc-binding site and the active site at large in scenarios when more dramatic structural changes around the zinc-binding site could induce irreversible damage and unfolding.

When the Zn^{2+} ion is absent altogether, the remaining empty space is filled with two water molecules, w1a and w1b, which are anchored by hydrogen bonds to neighboring residues (Fig. 8*d*). Interestingly, we do not observe any conformational changes of Lys137 in the coordination sphere in the absence of zinc. On the other hand, the absence of zinc is correlated with a movement of the Asp186 side chain towards Gly187, Cys188, Ser189 and Ile190 (Fig. 6*f*). With the zinc coordination site intact, the $O^{\delta 2}$ atom of Asp186 is hydrogen-bonded to water w1. When the zinc cation is absent, water w1 may adopt an alternative position (w1′) in which it is hydrogen-bonded to Lys137. However, in this alternative location w1′ is too far from Ser47 to form a hydrogen bond to its O^{γ} atom (Fig. 6*c*). Since in our interpretation of the catalytic mechanism water w1 is important for the polarization of Ser47, we may hypothesize that in the absence of Zn^{2+} polarization of the Ser47 nucleophile occurs simultaneously with substrate binding and rearrangement of the water molecules in the active site (Isupov *et al.*, 1996; Ronning *et al.*, 2010). This speculation of course requires experimental verification by the structure of an enzyme–substrate/product complex. So far, our crystallization efforts in this direction have been unsuccessful, but further experiments are under way. We are encouraged by the fact that the presence of the L-Asn additive clearly influences the crystallization process, leading to a depletion of Zn^{2+} occupancy and the collateral structural changes.

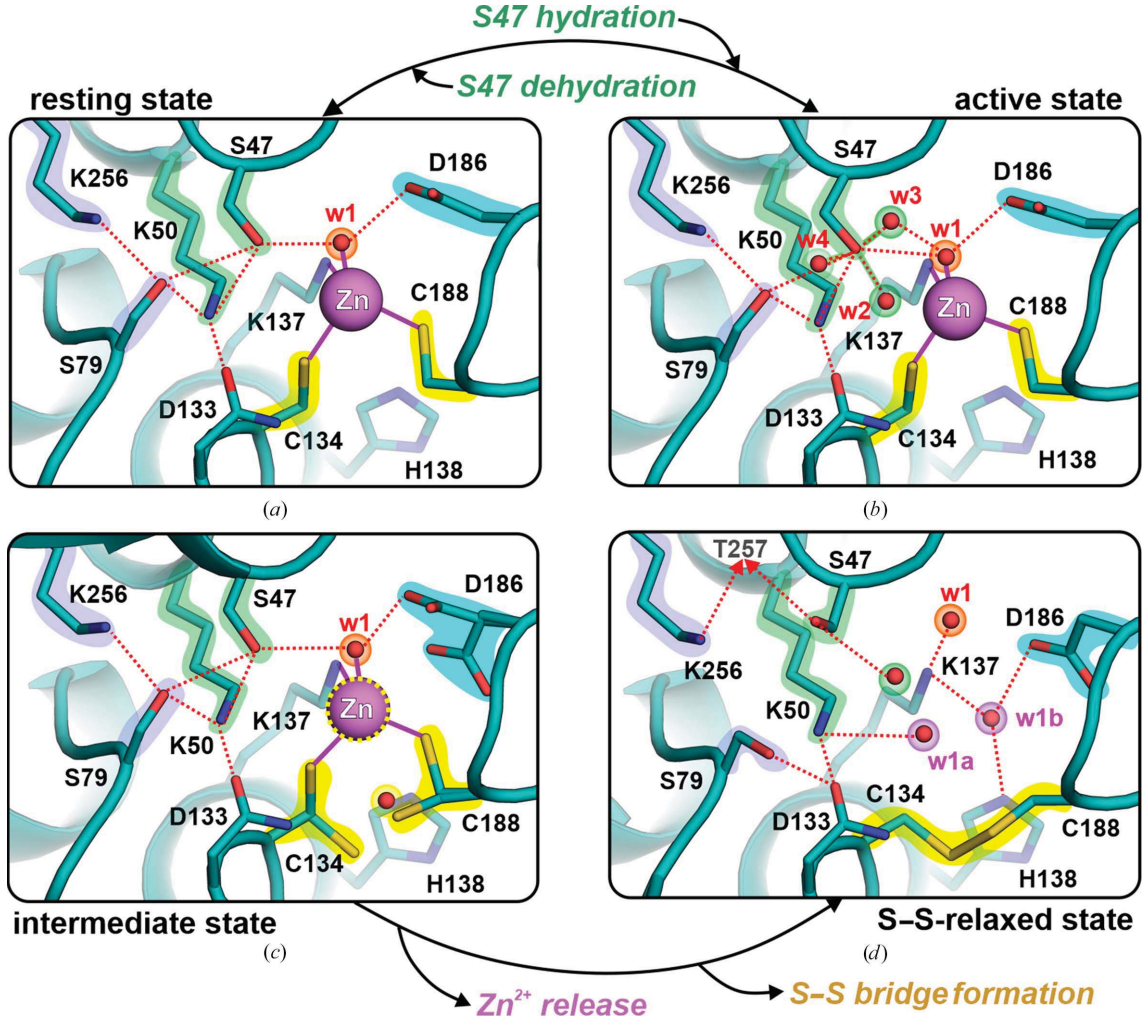


Figure 8
Selected functional states of the ReAIV molecule. (*a*) The resting state (enzyme between catalytic cycles) as observed in structure R4tP (chain *A*). (*b*) Active state with Ser47 polarized and hydrogen-bonding three tightly coordinated water molecules (as observed, for example, in structure R4oP-1, chain *D*). (*c*) Intermediate state observed during Zn^{2+} release from the active site (as observed in R4oP-2, chain *D*). (*d*) S–S-relaxed state, where the S–S bridge replaces the Zn^{2+} cation and alternative conformations (conformers *B*) of Ser47 and Ser79 are present (as observed, for example, in structure R4oP-1, chain *D*). The states in (*b*) and (*d*) are from the same structure (R4oP-1, chain *D*) presented after separation into two alternative conformational states.

Near the active site of ReAIV there is a conspicuous Cys242 residue. In most of the structures (except R4tP) we observed oxidation of this residue and two alternative conformations of the side chain (Supplementary Table S1). However, analysis of the ReAIV structures did not elucidate the role of this residue. As it has a counterpart in the ReAV molecule (Cys249 in ReAV numbering), where it is also oxidized, we maintain our tentative suggestion that it might function to protect the sensitive active site from oxidative damage (Loch *et al.*, 2021). The observation of a reduced Cys242 in the R4tP structure demonstrates that its oxidation is independent of the presence of Zn^{2+} in the zinc-binding site; it is also independent of the functional state of the active site. While the role of Cys242 oxidation is still unclear, we may speculate that Cys242 could act as a suicide protector preventing the zinc-containing active site from oxidative damage, as rhizobia have a very intensive metabolism that generates a large amount of reactive oxygen species.

4.4. Comparison of the enzymatic activity of ReAIV and ReAV

It has been known since 1996 that the activity of L-asparagine-degrading enzymes is high when *R. etli* is grown on L-Asn as the sole nitrogen and carbon source (Huerta-Zepeda *et al.*, 1996). These findings gave rise to the hypothesis that the L-asparaginase enzymatic system in these bacteria can play a catabolic role in the degradation of L-asparagine. The presence of a system of two intracellular L-asparaginases in *R. etli* was reported in 1997 (Huerta-Zepeda *et al.*, 1997), with one of the enzymes being constitutive (ReAIV according to the current convention) and the other inducible (ReAV) (Ortuño-Olea & Durán-Vargas, 2000). Expression of ReAV is positively regulated (induced) by the presence of its substrate L-Asn in the culture medium and negatively regulated by a carbon source. The ReAIV enzyme appeared to be constitutive as its expression was not affected by the presence of L-Asn, L-Asp, ammonia or nitrogen/carbon sources or by the amount of oxygen in the culture medium (Huerta-Zepeda *et al.*, 1997). It was concluded that the constitutive ReAIV enzyme plays a rather minor role in the utilization of L-Asn as a nitrogen or carbon source and physiologically might function to maintain a proper balance between L-Asn and L-Asp.

The biochemical and structural characterization of the ReAIV L-asparaginase described in this work is the first report to date presenting the properties of the constitutive L-asparaginase activity in *R. etli*. Our previous work showed that the inducible enzyme ReAV retains its L-asparaginase activity over a wide but rather alkaline pH range (Loch *et al.*, 2021). A similar survey carried out for ReAIV confirmed that the constitutive enzyme has an almost identical activity profile and is most active at pH 11 (Fig. 3*a*). It was also found, however, that ReAIV has a somewhat narrower pH range than ReAV (Fig. 3*a*). The relatively high and similar pH optimum of the two *R. etli* L-asparaginases is undoubtedly a result of the peculiar and nearly identical architecture of the active sites of the two enzymes, which are comprised of two catalytic Ser–Lys pairs and a Lys-containing Zn^{2+}-binding site sitting in a neighborhood rich in positively charged Arg and Lys residues. Physiologically, *R. etli*, which is a nitrogen-fixing symbiotic soil bacterium, prefers slightly acidic and neutral soils of pH 6.5–7.5 and does not grow under strongly alkaline or acidic conditions (Verástegui-Valdés *et al.*, 2014). Interestingly, low soil pH (about 4.0) represses nodule development (*i.e.* successful symbiotic association) on the roots of the legume symbionts by more than 90%, although the precise mechanism of this phenomenon is not known (Lin *et al.*, 2012). On the other hand, the amino-acid metabolism of *Rhizobium* leads to alkalinization of the soil (or unbuffered culture medium), partly from hydroxide exchange and partly from the release of significant amounts of NH_3 (Howieson & Dilworth, 2016). These results might suggest that by releasing sufficient amounts of NH_3, *R. etli* might promote nodule development, especially in acidic soils. Moreover, high L-asparaginase activity at alkaline pH might be beneficial to the bacteria, helping them to acquire the necessary amino-acid nutrients when the pH in the cytoplasm is temporarily elevated during intensive nitrogen fixation ($N_2 \rightarrow NH_4^+$). However, this hypothesis should be further verified, as the *R. etli* cytoplasm is resistant to pH shocks and its pH is maintained at about 6.0–6.6 (Graham *et al.*, 1994).

Moreover, we noticed that ReAIV (and also ReAV) demonstrated variable enzymatic activity depending on the buffer system used. These results clearly show that the type of buffering ions has a strong influence on the enzymatic activity of ReAIV, suggesting that the active site of ReAIV is highly sensitive to the reaction conditions. Specifically, the hydrolysis of L-Asn is more efficient in the presence of carbonate ions than phosphate ions, especially in the case of ReAIV. The problem of external factors affecting assessment of the activity of *R. etli* L-asparaginases will be discussed in detail in a separate paper (Sliwiak *et al.*, in preparation).

A full kinetic analysis showed that under selected optimal pH and buffer conditions (10 m*M* carbonate buffer pH 9.0) ReAIV is almost three times more efficient than ReAV (306 versus 110 s^{-1} mM^{-1}, respectively), with the following kinetic parameters: K_m = 1.34 ± 0.02 m*M* and k_{cat} = 411 ± 21 s^{-1} (Fig. 3*b*). Despite the high turnover number, the rather low substrate specificity makes ReAIV about ten times less efficient than the EcAII or ErAII enzymes that are already used in the treatment of ALL (with efficiencies of 2960 and 4370 s^{-1} mM^{-1}, respectively; Nguyen *et al.*, 2016). Since the concentration of L-Asn in human blood is low, ∼50 µ*M*, efficient hydrolysis of this pool of substrate requires an enzyme with a K_m in the low micromolar range, such as the EcAII and ErAII proteins already approved for ALL therapy, or other promising candidates, such as the enzymes from *Wolinella succinogenes* (K_m = 58 µ*M*; Nguyen *et al.*, 2017) or from guinea pig (K_m = 22 µ*M*; Schalk *et al.*, 2014), which are ∼20–50 times more specific than ReAIV. Moreover, the optimum of ReAIV activity lies too far above the physiological pH range for medicinal applications. Nevertheless, the lack of glutaminase activity makes ReAIV a potential target for site-directed

mutagenesis aimed at enhancing its L-asparagine specificity at physiological pH.

4.5. Comparison of the thermal stability of ReAIV and ReAV

To find the structural origins of the thermal stability of ReAIV, we analyzed the available literature to identify factors that regulate the thermal behavior of proteins. Adaptation of protein folds to higher temperature can be achieved by several strategies, such as enrichment of the protein sequence in particular amino acids (Zhou *et al.*, 2008; Farias & Bonato, 2003) or a high number of hydrogen bonds or electrostatic salt bridges, which greatly stabilize the protein fold (Vogt *et al.*, 1997; Vogt & Argos, 1997; Kumar *et al.*, 2000). As a rule, thermostable proteins have shorter and more rigid loops (to protect the entire fold from thermal excitation) and optimized (compact) packing of the N- and C-terminal parts, as well as a compact hydrophobic core (Dumina & Zhgun, 2023).

The optimal temperature for the growth of most *Rhizobium* species is 25–30°C, although many taxons are exposed to more extreme (cold or hot) temperatures (Alexandre & Oliveira, 2013). The same 25–30°C temperature range is also optimal for nodulation and atmospheric N_2 fixation in *Phaseolus vulgaris* (common bean), which is the symbiotic partner of *R. etli* (Michiels *et al.*, 1994). It has been shown that some *Rhizobium* species that nodulate *P. vulgaris* can survive even at 47°C, but they are unable to induce nodule formation at high soil temperatures, for example in tropical regions. These conditions restrict agricultural nitrogen fixation by legume crops (Alexandre & Oliveira, 2013). It was found early on that the constitutive *R. etli* ReAIV enzyme is thermostable and migrates faster in native electrophoretic gels than the thermolabile and inducible L-asparaginase ReAV (Ortuño-Olea & Durán-Vargas, 2000). As it is evident that temperature determines the metabolic state of rhizobia, it seems logical that two L-asparaginases with different thermal stabilities may be necessary for metabolic regulation under different conditions. The constitutive enzyme ReAIV is responsible for maintaining cell homeostasis even at high temperatures when the bacteria are unable to fix atmospheric nitrogen, while the inducible ReAV is active at lower temperatures and its activity is correlated with the conditions under which nitrogen fixation is most intensive.

In oligomeric proteins the domain interfaces are optimized to enhance intramolecular cohesive forces (Rahaman *et al.*, 2015). A comparison of the crystal structures of the thermostable ReAIV and thermolabile ReAV proteins generally confirms these rules: the sequence of ReAIV is shorter, ReAIV has a more compact fold (as confirmed by nanoDLS) as a result of the reduction of the length of the solvent-exposed structural elements (for example, ReAIV has no helix H13a and has a shorter helix H10; Fig. 4) and the dimer interface of ReAIV is stitched by significantly more salt bridges. The structure of ReAIV is also more ordered as, for example, it features an extra B10 β-strand involved in dimer stabilization, while the structure of ReAV in the same region is disordered. The presence of the additional B10 β-strand also stabilizes the dimer interface.

Protein thermal stability is also dependent on interactions with water molecules. As a rule, an increased number of polar contacts between protein residues and water molecules will improve the thermal stability (Gribenko *et al.*, 2009; Sanchez-Ruiz & Makhatadze, 2001). These interactions involve not only water molecules surrounding the protein surface, but also internal water molecules (Takano *et al.*, 1997; Pechkova *et al.*, 2007). The presence of water molecules inside internal cavities or superficial pockets usually improves the molecular packing of protein side chains and extends the network of hydrogen bonds in the protein core. Inner hydration reduces the volume of internal voids and makes the proteins resistant not only to temperature but usually also to elevated pressure due to the presence of an extended system of hydrogen bonds that function as internal 'springs' (Chakraborty *et al.*, 2015). An interesting case of five internalized water molecules stabilizing the core of a small polypeptide (bovine pancreatic trypsin inhibitor; BPTI; 58 residues) was reported early on and was described in detail at very high resolution (Czapinska *et al.*, 2000; Addlagatta *et al.*, 2001). In the ReAIV molecule we observe three completely sealed, relatively large internal cavities, I, II and III, occluding volumes of 93, 70 and 149 Å^3, respectively, filled with three, three and seven occluded water molecules, respectively (Fig. 7), which greatly stabilize the entire protein fold, functioning as protective 'water cushions'. The occluded water molecules are linked by hydrogen bonds (Supplementary Fig. S3 and Table S3) and form hydrogen bonds to the protein main chain and, sporadically in cavity III, also to side chains. In contrast, the ReAV molecule has only one recognizable cavity (corresponding to cavity III of ReAIV), but it is significantly smaller (∼91 Å^3) and holds only three occluded water molecules (Supplementary Fig. S4*c*). In the area of cavity I, there is just one water molecule in an occluded volume of ∼39 Å^3. There is no void corresponding to cavity II of ReAIV.

Based on structural analyses of the mesophilic class 1 L-asparaginase EcAII from *E. coli* and its homolog PyAII from the thermophile *Pyrococcus yayanosii*, it has recently been postulated that differences in protein thermostability might be linked to surface-charge distribution (Howieson & Dilworth, 2016; Dumina & Zhgun, 2023). The similarity of the surface-charge distributions of ReAIV and ReAV indicates that this argument cannot be invoked to explain the drastically different thermal stabilities of these two *R. etli* L-asparaginases.

It seems that the ReAIV molecule has been perfectly adapted during evolution to survive over a wide range of temperatures and is designed by nature to function at high pH values (see Section 4.1). Although the proteome of *R. etli* is rather poorly characterized, it has already been reported that *R. etli* possesses some other thermostable and thermolabile enzymes, for example a thermostable xylitol dehydrogenase (optimal temperature 70°C, optimal pH 9.5; Tiwari *et al.*, 2010) and a thermolabile glutaminase encoded by the *glsA* gene (Calderón *et al.*, 1999). As thermal resistance is an important feature of enzymes with medicinal applications, our results

suggest that ReAIV might be superior to ReAV as a starting point for enzyme engineering towards potential therapeutic applications.

5. Conclusions and outlook

In addition to the well characterized inducible and thermolabile class 3 L-asparaginase ReAV (Loch *et al.*, 2021), the genome of the bacterium *R. etli* encodes another enzyme with the same activity, named ReAIV, which has a different expression profile (constitutive) and biophysical properties (thermostable). Its sequence identity and similarity to ReAV are low (31.1% and 44.6%, respectively) and its sequence is significantly shorter, particularly at the C-terminus. Based on five accurate crystal structures, some determined at very high resolution, we show that despite these differences ReAIV is structurally very similar to ReAV and has a nearly identical active site. We also report the catalytic parameters of the ReAIV enzyme.

With the discovery of the new class 3 L-asparaginases (originally termed *R. etli*-type), interesting questions arise concerning the origin of these enzymes and their evolutionary fate. Since class 3 L-asparaginases are quite similar in fold to penicillin-binding proteins and the related serine β-lactamases, it is quite possible that they have evolved from these cell-wall-metabolism-related bacterial proteins. Their evolution would, however, have generated important differences, such as the presence of a zinc cation in the active site or a specific manner of dimerization. From a pan-genomic analysis of the bacterial kingdom (Zielezinski *et al.*, 2022) it is quite obvious that ReAIV-type enzymes are indeed characteristic of rhizobia and may have evolved in this group of symbiotic, nitrogen-fixing bacteria. ReAV-type enzymes are, however, rarely found in rhizobia and the presence of ReAV in *R. etli* seems to be the result of a horizontal gene-transfer event from *Burkholderia*. It is interesting to ask why rhizobia would acquire another copy of a similar enzyme via a roundabout route instead of duplicating an existing gene. Is ReAIV perhaps in fact the result of gene duplication that has spread more easily in rhizobia? However the question would still remain: why? We are not yet at a point to answer these and similar evolutionary questions. More research is necessary, but we are certainly on the threshold of a fascinating field.

6. Data availability

Atomic coordinates and structure factors corresponding to the final crystallographic models of ReAIV generated in this study have been deposited in the Protein Data Bank (PDB) as entries 8osw (R4mC-1), 8clz (R4mC-2), 8cly (R4tP), 8ori (R4oP-1) and 8col (R4oP-2). The corresponding raw diffraction images have been deposited in the Macromolecular Xtallography Raw Data Repository (https://mxrdr.icm.edu.pl) at https://doi.org/10.18150/NOBOWP (R4mC-1), https://doi.org/10.18150/WFJHW8 (R4mC-2), https://doi.org/10.18150/SWISNO (R4tP), https://doi.org/10.18150/PLH5BS (R4oP-1) and https://doi.org/10.18150/7RCKMY (R4oP-2).

Acknowledgements

We thank the EMBL/DESY Outstation in Hamburg for the allocation of synchrotron beamtime at PETRA III and the beamline scientists for help with data collection. We thank Dr Ewa Sitkiewicz at the Mass Spectrometry Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw for carrying out the MS measurements. We also thank Dr Milosz Ruszkowski for helping with X-ray data collection. Author contributions were as follows. JIL refined the R4mC-2, R4oP-2 and R4tP structures, analyzed the results, performed nanoDSF and nanoDLS experiments, prepared the illustrations and wrote the manuscript. PW produced and purified the protein, crystallized the R4mC-1, R4oP-1 and R4oP-2 crystal forms, conducted enzyme kinetic assays, analyzed the results and wrote the manuscript. JS carried out all ITC experiments, analyzed the data and wrote the manuscript. BI worked out the protein expression and purification protocol, crystallized the R4mC-2 and R4tP crystal forms and solved their structures. KP prepared and purified the proteins. MC carried out the analysis of protein cavities. MaG supervised bench work and refined the R4oP-1 structure. MiG collected and processed X-ray diffraction data, supervised all crystallographic calculations and analyzed the crystal structures. MJ conceived and coordinated the project, designed the experiments, analyzed the results and wrote the manuscript. The authors declare no competing interests.

Funding information

This work was supported by National Science Centre (NCN, Poland) grants 2020/37/B/NZ1/03250 and 2018/31/B/NZ1/00047.

References

Addlagatta, A., Krzywda, S., Czapinska, H., Otlewski, J. & Jaskolski, M. (2001). *Acta Cryst.* D**57**, 649–663.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Alexandre, A. & Oliveira, S. (2013). *Crit. Rev. Microbiol.* **39**, 219–228.
Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G. & Nielsen, H. (2019). *Nat. Biotechnol.* **37**, 420–423.
Borek, D. & Jaskólski, M. (2001). *Acta Biochim. Pol.* **48**, 893–902.
Calderón, J., Huerta-Saquero, A., Du Pont, G. & Durán, S. (1999). *Biochim. Biophys. Acta*, **1444**, 451–456.
Chakraborty, D., Taly, A. & Sterpone, F. (2015). *J. Phys. Chem. B*, **119**, 12760–12770.
Chwastyk, M. & Cieplak, M. (2021). *Front. Mol. Biosci.* **8**, 692230.
Chwastyk, M., Galera-Prat, A., Sikora, M., Gómez-Sicilia, À, Carrión-Vázquez, M. & Cieplak, M. (2014). *Proteins*, **82**, 717–726.
Chwastyk, M., Jaskolski, M. & Cieplak, M. (2014). *FEBS J.* **281**, 416–429.
Chwastyk, M., Jaskolski, M. & Cieplak, M. (2016). *Proteins*, **84**, 1275–1286.
Chwastyk, M., Panek, E. A., Malinowski, J., Jaskólski, M. & Cieplak, M. (2020). *Front. Mol. Biosci.* **7**, 591381.
Czapinska, H., Otlewski, J., Krzywda, S., Sheldrick, G. M. & Jaskólski, M. (2000). *J. Mol. Biol.* **295**, 1237–1249.
DeLano, W. (2002). *CCP4 Newsl. Protein Crystallogr.* **40**, 82–92.
Dumina, M. & Zhgun, A. (2023). *Int. J. Mol. Sci.* **24**, 2674.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Farias, S. T. & Bonato, M. C. M. (2003). *Genet. Mol. Res.* **2**, 383–393.
Graham, P. H., Draeger, K. J., Ferrey, M. L., Conroy, M. J., Hammer, B. E., Martinez, E., Aarons, S. R. & Quinto, C. (1994). *Can. J. Microbiol.* **40**, 198–207.
Gribenko, A. V., Patel, M. M., Liu, J., McCallum, S. A., Wang, C. & Makhatadze, G. I. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 2601–2606.
Howieson, J. G. & Dilworth, M. J. (2016). Editors. *Working with Rhizobia*. Canberra: Australian Centre for International Agricultural Research.
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.
Ishizuka, H., Yamamoto, T., Arata, Y. & Fujiwara, S. (1973). *Bull. Chem. Soc. Jpn*, **46**, 468–471.
Isupov, M. N., Obmolova, G., Butterworth, S., Badet-Denisot, M. A., Badet, B., Polikarpov, I., Littlechild, J. A. & Teplyakov, A. (1996). *Structure*, **4**, 801–810.
Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head–Gordon, T., Geng, W., Krasny, R., Wei, G.-W., Holst, M. J., McCammon, J. A. & Baker, N. A. (2018). *Protein Sci.* **27**, 112–128.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Kumar, S., Tsai, C. J., Ma, B. & Nussinov, R. (2000). *J. Biomol. Struct. Dyn.* **17**, Suppl. 1, 79–85.
Lin, M.-H., Gresshoff, P. M. & Ferguson, B. J. (2012). *Plant Physiol.* **160**, 2028–2039.
Loch, J. I., Imiolczyk, B., Sliwiak, J., Wantuch, A., Bejger, M., Gilski, M. & Jaskolski, M. (2021). *Nat. Commun.* **12**, 6717.
Loch, J. I. & Jaskolski, M. (2021). *IUCrJ*, **8**, 514–531.
Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D. & Lopez, R. (2019). *Nucleic Acids Res.* **47**, W636–W641.
Maret, W. (2004). *Biochemistry*, **43**, 3301–3309.
Maret, W. & Li, Y. (2009). *Chem. Rev.* **109**, 4682–4707.
McCall, K. A., Huang, C. & Fierke, C. A. (2000). *J. Nutr.* **130**, 1437S–1446S.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Michalska, K. & Jaskolski, M. (2006). *Acta Biochim. Pol.* **53**, 627–640.
Michiels, J., Verreth, C. & Vanderleyden, J. (1994). *Appl. Environ. Microbiol.* **60**, 1206–1212.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Needleman, S. B. & Wunsch, C. D. (1970). *J. Mol. Biol.* **48**, 443–453.
Nguyen, H. A., Durden, D. L. & Lavie, A. (2017). *Sci. Rep.* **7**, 41643.
Nguyen, H. A., Su, Y. & Lavie, A. (2016). *J. Biol. Chem.* **291**, 17664–17676.
Ortuño-Olea, L. & Durán-Vargas, S. (2000). *FEMS Microbiol. Lett.* **189**, 177–182.
Pauling, L. (1960). *The Nature of the Chemical Bond*. Ithaca: Cornell University Press.
Pechkova, E., Sivozhelezov, V. & Nicolini, C. (2007). *Arch. Biochem. Biophys.* **466**, 40–48.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Rahaman, O., Kalimeri, M., Melchionna, S., Hénin, J. & Sterpone, F. (2015). *J. Phys. Chem. B*, **119**, 8939–8949.
Riikonen, A., Rouvinen, J., Tikkanen, R., Julkunen, I., Peltonen, L. & Jalanko, A. (1996). *J. Biol. Chem.* **271**, 21340–21344.
Ronning, D. R., Iacopelli, N. M. & Mishra, V. (2010). *Protein Sci.* **19**, 2498–2510.
Sanchez-Ruiz, J. M. & Makhatadze, G. I. (2001). *Trends Biotechnol.* **19**, 132–135.
Schalk, A. M., Nguyen, H.-A., Rigouin, C. & Lavie, A. (2014). *J. Biol. Chem.* **289**, 33175–33186.
Simas, R. G., Krebs Kleingesinds, E., Pessoa Junior, A. & Long, P. F. (2021). *J. Chem. Technol. Biotechnol.* **96**, 1326–1332.
Takano, K., Funahashi, J., Yamagata, Y., Fujii, S. & Yutani, K. (1997). *J. Mol. Biol.* **274**, 132–142.
Tiwari, M. K., Moon, H.-J., Jeya, M. & Lee, J.-K. (2010). *Appl. Microbiol. Biotechnol.* **87**, 571–581.
Verástegui-Valdés, M. M., Zhang, Y. J., Rivera-Orduña, F. N., Cheng, H.-P., Sui, X. H. & Wang, E. T. (2014). *Syst. Appl. Microbiol.* **37**, 605–612.
Vogt, G. & Argos, P. (1997). *Fold. Des.* **2**, S40–S46.
Vogt, G., Woell, S. & Argos, P. (1997). *J. Mol. Biol.* **269**, 631–643.
Wang, Y., Wang, G., Moitessier, N. & Mittermaier, A. K. (2020). *Front. Mol. Biosci.* **7**, 583826.
Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., Foster, L. J. & Brinkman, F. S. L. (2010). *Bioinformatics*, **26**, 1608–1615.
Zastrow, M. L. & Pecoraro, V. L. (2014). *Biochemistry*, **53**, 957–978.
Zhao, Y., Chwastyk, M. & Cieplak, M. (2017). *Sci. Rep.* **7**, 39851.
Zhou, X.-X., Wang, Y.-B., Pan, Y.-J. & Li, W.-F. (2008). *Amino Acids*, **34**, 25–33.
Zielezinski, A., Loch, J. I., Karlowski, W. M. & Jaskolski, M. (2022). *Sci. Rep.* **12**, 15797.

STRUCTURAL BIOLOGY

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

Rhizobium etli has two L-asparaginases with low sequence identity but similar structure and catalytic center

Joanna I. Loch, Paulina Worsztynowicz, Joanna Sliwiak, Marta Grzechowiak, Barbara Imiolczyk, Kinga Pokrywka, Mateusz Chwastyk, Miroslaw Gilski and Mariusz Jaskolski

Acta Cryst. (2023). D**79**, 775–791