Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structure of tomato wound-induced leucine aminopeptidase sheds light on substrate specificity

Kevin Duprez,[a] Melissa A. Scranton,[b]‡ Linda L. Walling[b] and Li Fan[a]*

[a]Department of Biochemistry, University of California-Riverside, Riverside, CA 92521, USA, and [b]Department of Botany and Plant Sciences and Center for Plant Cell Biology, University of California-Riverside, Riverside, CA 92521, USA

‡ Current address: Department of Food and Fuel for the 21st Century, University of California, San Diego, CA 92093, USA.

Correspondence e-mail: lifan@ucr.edu

Received 18 October 2013
Accepted 20 March 2014

PDB reference:
wound-induced leucine aminopeptidase, 4ksi

The acidic leucine aminopeptidase (LAP-A) from tomato is induced in response to wounding and insect feeding. Although LAP-A shows *in vitro* peptidase activity towards peptides and peptide analogs, it is not clear what kind of substrates LAP-A hydrolyzes *in vivo*. In the current study, the crystal structure of LAP-A was determined to 2.20 Å resolution. Like other LAPs in the M17 peptidase family, LAP-A is a dimer of trimers containing six monomers of bilobal structure. Each monomer contains two metal ions bridged by a water or a hydroxyl ion at the active site. Modeling of different peptides or peptide analogs in the active site of LAP-A reveals a spacious substrate-binding channel that can bind peptides of five or fewer residues with few geometric restrictions. The sequence specificity of the bound peptide is likely to be selected by the structural and chemical restrictions on the amino acid at the P1 and P1′ positions because these two amino acids have to bind perfectly at the active site for hydrolysis of the first peptide bond to occur. The hexameric assembly results in the merger of the open ends of the six substrate-binding channels from the LAP-A monomers to form a spacious central cavity allowing the hexameric LAP-A enzyme to simultaneously hydrolyze six peptides containing up to six amino acids each. The hexameric LAP-A enzyme may also hydrolyze long peptides or proteins if only one such substrate is bound to the hexamer because the substrate can extend through the central cavity and the two major solvent channels between the two LAP-A trimers.

1. Introduction

Leucine aminopeptidase (LAP; EC 3.4.11.1) is an aminoacyl-peptide hydrolase that catalyzes the removal of amino acids from the N-terminus of a peptide or protein (for reviews, see Henson & Frohne, 1976; Kim & Lipscomb, 1994; Matsui *et al.*, 2006). LAPs are widely distributed in both eukaryotes and prokaryotes from bacteria to humans and are of critical biological and medical importance. LAPs are known to play critical roles in many 'housekeeping' functions, such as the maturation of proteins, digestive and intracellular protein metabolism and hormone regulation (Taylor, 1993). Altered aminopeptidase activity has been associated with several pathological disorders such as cancer (Umezawa, 1980) and possibly eye-lens cataracts (Taylor *et al.*, 1982). LAPs belong to the M17 family of peptidases, which are highly conserved di-zinc metallopeptidases that are found in plants, animals and microbes (Matsui *et al.*, 2006). Animal LAPs may have a role in the turnover of oxidatively damaged proteins in the lens of the eye (Taylor *et al.*, 1982). Human LAP has been proposed to process peptides released from the 26S proteasome for use in MHC I presentation (Saveanu *et al.*, 2002). Interestingly,

© 2014 International Union of Crystallography

Escherichia coli LAP (also named PepA) is multifunctional as an aminopeptidase and as a DNA-binding protein that mediates site-specific recombination in *ColE1* plasmids (Stirling *et al.*, 1989) and acts as a transcription factor to modulate the *carAB* operon (Charlier *et al.*, 1995).

The complement of LAPs in plants is more complex and their roles are being elucidated (Fowler *et al.*, 2009). In plants, there are two classes of LAPs, which are 70–77% identical, and, with one exception (*Arabidopsis thaliana* LAP1), plant LAPs reside within the chloroplast stroma (Tu *et al.*, 2003; Narváez-Vásquez *et al.*, 2007; Walling, 2006). The LAPs with neutral pIs (LAP-N) are detected in all plants and are constitutively expressed (Bartling & Nosek, 1994; Bartling & Weiler, 1992; Gu, Pautot *et al.*, 1996; Chao *et al.*, 1999, 2000; Gu *et al.*, 1999). The LAPs with acidic pIs (LAP-A) are found only in a subset of the Solanaceae and are induced in response to both biotic and abiotic stresses (Chao *et al.*, 1999, 2000). LAP-A from tomato (*Solanum lycopersicum*) maintains the epitope conservation of LAP-N, although it exhibits variation in its gene regulation, being up-regulated in response to various environmental cues, such as wounding, water deficit, salinity and caterpillar feeding (Pautot *et al.*, 1993, 2001; Gu, Pautot *et al.*, 1996; Chao *et al.*, 1999). *In vitro* studies on peptides with different amino-acid sequences revealed that LAP-A preferably hydrolyzes substrates with basic (Arg) and nonpolar (Leu, Val, Ile and Ala) residues (Gu *et al.*, 1999; Gu & Walling, 2000), showing similar hydrolytic activity to the porcine and *E. coli* LAPs. The biological substrates of tomato LAP-A are unknown to date, but given the localization of LAP-A, its substrates must reside within the chloroplast stroma or in passage through the stroma in plants or reside within the insect gut. In plants LAP-A is involved in plastid–nucleus communication (retrograde signaling), as LAP-A is both a positive and a negative regulator of nuclear gene expression after injury (Fowler *et al.*, 2009; Scranton *et al.*, 2013).

Crystal structures of several LAP enzymes have been determined, including those from *Bos taurus* (Kraft *et al.*, 2006; Sträter, Sun *et al.*, 1999), *E. coli* (Sträter, Sherratt *et al.*, 1999), *Pseudomonas putida* (Kale *et al.*, 2010) and *Plasmodium falciparum* (McGowan *et al.*, 2010), all of which exhibit a hexameric structure with bilobal monomer subunits. From these structures, key residues were identified that facilitate metal-cofactor and substrate binding and catalysis (Gu & Walling, 2002; Kraft *et al.*, 2006; Sträter, Sherratt *et al.*, 1999). In particular, the crystal structures of bovine lens LAP bound with inhibitors and transition-state analogs such as bestatin and L-leucinephosphonic acid have provided important insights into the structure and catalytic mechanism of M17 LAPs (Burley *et al.*, 1991; Kim & Lipscomb, 1993; Sträter & Lipscomb, 1995*a*,*b*; Sträter, Sherratt *et al.*, 1999; Sträter, Sun *et al.*, 1999).

The hydrolysis of the peptide bond by LAP enzymes occurs at an active site consisting of two metals, a bridging water molecule or hydroxyl ion and a bicarbonate ion. The metal-bridging water molecule is believed to be the nucleophile attacking the scissile peptide bond of the substrate. Besides having a role in positioning and activating the nucleophile, the two metal ions at the active site are important for substrate binding and transition-state stabilization: one metal ion is crucial for binding the N-terminal amino group of the substrate, while the other metal ion binds the carbonyl O atom of the scissile amide bond and stabilizes the negative charge that develops on this O atom (the oxyanion) in the presumed tetrahedral *gem*-diolate transition state. The oxyanion is further stabilized by an interaction with the nearby Lys residue. The bicarbonate ion is believed to act as a general base, abstracting a proton from the nucleophilic water molecule and transferring it to the amino-terminal group of the P1′ product after cleavage of the peptide bond (Kale *et al.*, 2010).

Currently, the *in vivo* substrates of the wound-induced LAP-A from tomato have yet to be discovered. To provide an accurate structural model for explaining the biological and catalytic properties of LAP-A, we have analyzed this enzyme by X-ray crystallography. Here, we report the crystal structure of unliganded LAP-A determined at 2.2 Å resolution. Analysis of this structure along with substrate modeling allowed us to provide new insights into the structural and functional features of the LAP-A enzyme.

2. Materials and methods

2.1. Protein purification

The pQLapA-M construct expresses the mature LAP-A1 protein (amino-acid residues 54–571) as a His_6-LAP-A protein in *E. coli* as described previously (Gu *et al.*, 1999). The recombinant LAP-A protein was produced in 500 ml LB medium supplemented with 50 mg l^{-1} ampicillin and inoculated with 20 ml freshly grown overnight culture. The cells were grown at 37°C until an OD_{600} of 0.6 was reached, at which point IPTG was added to a final concentration of 0.4 m*M*. 6 h after IPTG induction, the cells were harvested by centrifugation at 6000*g* for 15 min at 4°C. The pelleted cells were resuspended in 70 ml cold lysis buffer (50 m*M* NaH_2PO_4/NaOH pH 8.0, 300 m*M* NaCl). The cells were lysed by the addition of 35 mg lysozyme followed by incubation on ice for 30 min. The lysate was sonicated on ice for 1 min with a 10 s on/off cycle. The lysate was clarified by centrifugation at 10 000*g* for 30 min at 4°C. The clarified supernatant was filtered through a 0.2 µm filter and loaded onto a 1.5 ml Ni–NTA agarose column equilibrated with lysis buffer. After passing the supernatant, the column was washed with 6.67 column volumes (CV) of 20 m*M* imidazole followed by 6.67 CV of 40 m*M* imidazole and was then eluted with 6.67 CV of 250 m*M* imidazole in 1 ml fractions on ice. Protein concentrations were measured using the bicinchoninic acid (BCA) assay (Smith *et al.*, 1985). Samples were either diluted with an equal volume of 80% glycerol and stored at −20°C or frozen in liquid nitrogen and stored at −80°C.

2.2. Crystallization and data collection

All crystallization experiments were performed at 298 K using the hanging-drop vapour-diffusion technique on siliconized cover slips sealed on 24-well Linbro plates (Hampton

Research, USA). LAP-A crystals were grown using 500 µl reservoir solution and drops consisting of a 1:1 ratio of protein solution (13 mg ml^{-1}) to mother liquor consisting of 0.1 *M* 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.5, 0.1 *M* ammonium sulfate, 45%(*v*/*v*) methylpentanediol (MPD); rod-shaped crystals of varying lengths were harvested after 5 d. Single crystals were selected with a nylon CryoLoop (Hampton Research, USA) and freshly cooled in a 100 K nitrogen stream with the mother liquor serving as a cryoprotectant. X-ray diffraction data were collected using an in-house Rigaku MicroMax-007 HF X-ray generator (wavelength of 1.541 Å) and a Rigaku R-AXIS IV^{++} image-plate detector at UC Riverside. Diffraction images were initially indexed with *CrystalClear* v.2.0 (Rigaku, USA) to evaluate the space group and the diffraction limit. A data set of 440 images was collected from one single crystal exhibiting the highest resolution (2.20 Å) using a crystal-to-detector distance of 160 mm, an oscillation width of 0.5° and an exposure time of 2 min per image. X-ray diffraction intensities were indexed, merged and scaled with *HKL*-2000 (Otwinowski & Minor, 1997).

2.3. Structure determination and refinement

The LAP-A structure was solved *via* molecular replacement (MR) using the *CCP*4 program *MOLREP* (Vagin & Teplyakov, 2010; Winn *et al.*, 2011). The conserved C-terminal region (amino acids 249–603, equivalent to LAP-A residues 215–571) of the M17 aminopeptidase protein from *Plasmodium falciparum* (PDB entry 3kqx; McGowan *et al.*, 2010) was used as a model after the conversion of all nonconserved residues to alanines. After MR, the missing N-terminal residues were built into the model *via* the *Buccaneer* pipeline (Cowtan, 2006, 2008). Cycles of refinement in *REFMAC* (Murshudov *et al.*, 1997, 2011) and *PHENIX* (Adams *et al.*, 2010; Afonine *et al.*, 2005; Moriarty *et al.*, 2009) alternated with manual model building in *Coot* (Emsley *et al.*, 2010) using σ-weighted ($F_o - F_c$) and ($2F_o - F_c$) electron-density maps. The R_{free} calculation was based on 5% of the total reflections and was evaluated throughout refinement. Different divalent metal ions were incorporated into the active site to obtain the best fit to the ($2F_o - F_c$) map. *SFCHECK* (Vaguine *et al.*, 1999) was used to validate the final model geometry and data fit. The final deposited model (PDB entry 4ksi), with an *R* factor of 13.7% and an R_{free} of 17.3%, contains residues 55–571 of the LAP-A protein, two active-site magnesium ions, two sulfate ions, three chloride ions, five glycerol molecules, one imidazole molecule, one MPD molecule and 472 water molecules, while the N-terminal His-tag residues were not defined owing to a lack of electron density.

2.4. Fluorescence-scan experiments

Excitation scans were carried out on Advanced Light Source beamline 12.3.1 using the *Blu-Ice* software suite to identify the existence of the metallic elements Ca, Cu, Fe, Mn, Ni and Zn. The results indicated that none of these are present in the LAP-A crystal. The absorption edge (fluorescence emission at 1250 eV and excitation energy at 1303 eV) of Mg is out of the range of the beamline, so a scan for Mg was not carried out.

Table 1
Statistics of data collection and structure refinement.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P6_322$
Unit-cell parameters (Å, °)	$a = b = 160.6$, $c = 104.2$, $\alpha = \beta = 90.0$, $\gamma = 120.0$
Temperature (K)	100
Wavelength (Å)	1.5418
No. of reflections	728860
No. of unique reflections	38529 (2551)
Resolution range (Å)	30.0–2.20 (2.26–2.20)
Completeness (%)	99.90 (100)
R_{merge} (%)	13.6 (61.8)
$\langle I/\sigma(I)\rangle$	23.618 (5.256)
Multiplicity	17.9 (16.6)
Refinement	
R_{work} (%)	13.67 (15.4)
R_{free} (5% of data) (%)	17.35 (21.7)
Average *B* values (Å^2)	
Protein atoms	16.9
Magnesium ions	19.9
Sulfate ions	47.0
Chloride ions	62.7
Glycerol molecules	46.2
Imidazole molecules	41.6
MPD molecules	41.7
Water molecules	35.6
R.m.s. deviation from standard geometry	
Bond lengths (Å)	0.017
Bond angles (°)	1.751
Ramachandran plot statistics, residues in (%)	
Most favoured region	90.2
Additional allowed region	9.8
Generously allowed region	0

3. Results

3.1. Overall structure

The mature form of LAP-A consisting of amino acids 55–571 (Gu, Chao *et al.*, 1996) crystallized in space group $P6_322$ with one molecule in the asymmetric unit, and the structure was elucidated to 2.20 Å resolution (with an *R* factor of 13.7% and an R_{free} of 17.3%) by molecular replacement (see Table 1 for statistics of the structure determination). The LAP-A molecule exhibits a bilobal structure consisting of the conserved catalytic C-terminal domain (residues 269–571; red in Fig. 1*a*) and the N-terminal domain (residues 55–237; blue in Fig. 1*a*) connected by a long helix (residues 238–268; yellow in Fig. 1*a*). The N-terminal domain has no obvious enzymatic function and is not conserved in amino-acid sequence among LAP enzymes. LAP-A is a hexamer of two trimers (Fig. 1*b*). Each trimer is formed by the interactions of the C-terminal domains of three LAP-A monomers at the center with the N-termini stretching out to form a triangular shape. At the trimeric interface is a triangular cavity of ∼16 Å in side length and ∼10 Å in depth (yellow trangle in Fig. 1*c*), which is connected to the substrate-binding channel of each monomer (Fig. 1*c*). Overall, a spacious cavity is formed at the center of the trimer with a radius of ∼30 Å (yellow circle in Fig. 1*c*).

Two trimers come together to form a hexamer mainly through pairwise interactions between the N-terminal domains (Fig. 1*b*), resulting in a central hollow chamber that hosts the six active sites connected by open substrate-binding channels (Fig. 1).

Each active site contains two metal ions (Mg^{2+}) bridged by a water or hydroxyl ion as observed in other LAP structures (Fig. 1*d*). While many published M17 peptidase structures, including the bovine lens LAP structures, have two Zn^{2+} ions at the active site, Mg^{2+} ions fitted best to the electron density at the active site of the LAP-A enzyme during structural refinement. Fluorescence-scan experiments at a synchrotron beamline excluded the presence of the metallic elements Ca, Cu, Fe, Mn, Ni and Zn. This is consistent with the observation that LAP-A has a better *in vitro* activity in the presence of Mg^{2+} (Gu *et al.*, 1999). The distance between the two metal ions is 3.0 Å. The two Mg^{2+} ions are mainly coordinated by carboxylate O atoms from the side chains of three aspartates and one glutamate (Asp347, Asp367, Asp427 and Glu429). In addition, metal M1 (Mg601) is coordinated by the main-chain carbonyl O atom of residue Asp427, and metal M2 (Mg602) makes a bond to the ε-amino group of a lysine residue (Lys342). All of the amino-acid residues that coordinate the two metal ions are conserved among the LAP enzymes, and the coordinating bond distances and metal–metal distance in the LAP-A structure are very similar to those reported for homologous LAP structures (Kale *et al.*, 2010; Kraft *et al.*, 2006; Sträter, Sherratt *et al.*, 1999; Sträter, Sun *et al.*, 1999). One sulfate ion was also observed near the metal-binding site, possibly playing the same role as a previously observed bicarbonate ion in bovine lens LAP and *E. coli* PepA (Sträter, Sun *et al.*, 1999). The sulfate ion is bound to Arg431 and makes hydrogen bonds to the metal-bridging water molecule

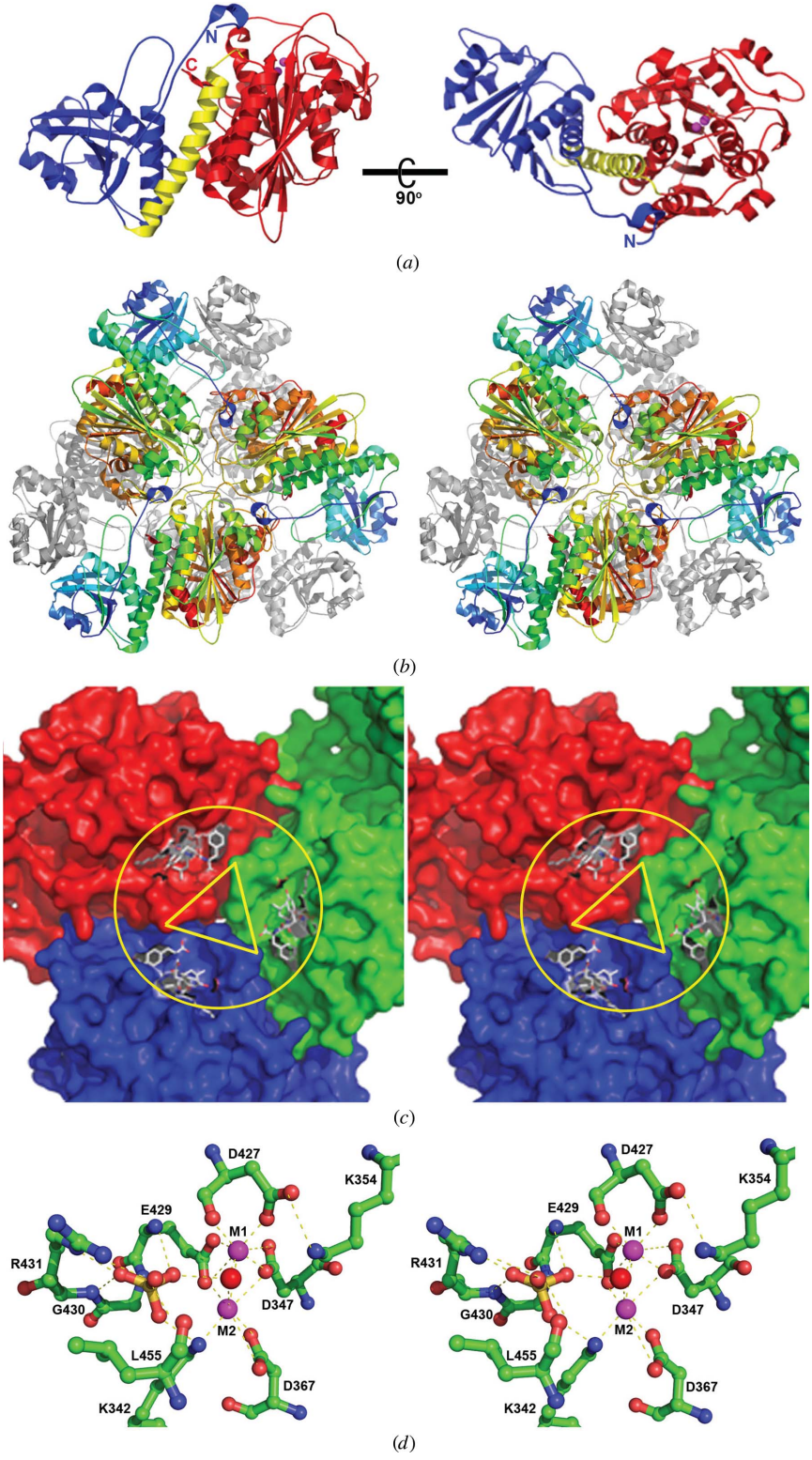

Figure 1
Structure of LAP-A. (*a*) Monomer structure shown as ribbons. The N-terminal and C-terminal domains are colored blue and red, respectively. The linking helix is shown in yellow. (*b*) Stereoview of the LAP-A hexameric structure shown as ribbons. The bottom trimer is colored grey while the upper trimer is colored in rainbow for each monomer, with the N-terminus in blue and the C-terminus in red. (*c*) Structure of the LAP-A trimer shown as molecular surfaces. The three monomers are colored red, blue and green, respectively. The central cavity at the trimeric interface is indicated by a yellow triangle, while the space forming the spacious central hollow chamber of the hexamer is indicated by a yellow circle. The substrate-binding channels at three active sites are indicated by microginin (shown in stick representation). Microginin was docked into the active sites by superposing the structure of bovine lens LAP bound with microginin (PDB entry 2j9a) onto the structure of LAP-A. (*d*) Stereoview of the conserved residues at the active site. Amino-acid residues are shown in ball-and-stick representation. The atoms are colored as follows: N, blue; O, orange; C, green; two metal (Mg^{2+}) ions (M1 and M2), magenta; bridging water, red; S atom of the sulfate ion, yellow.

and the amino groups of residues Glu429 and Gly430, as well as the ε-amino group of residue Lys342 (Fig. 1*d*).

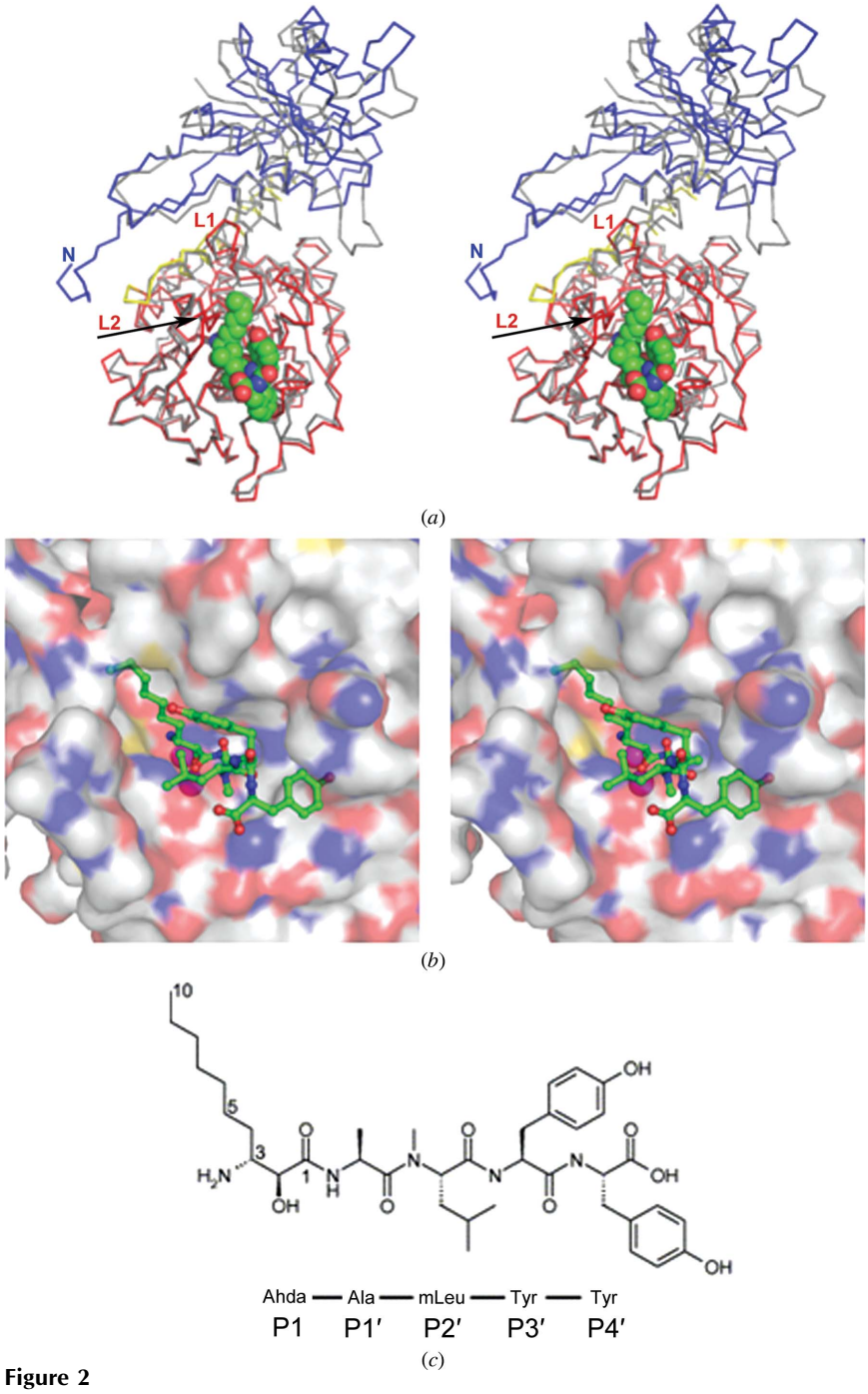

Figure 2
Structural comparison of LAP-A with bovine lens LAP bound to FR1. (*a*) The structure of monomeric LAP-A was superimposed with that of bovine lens LAP bound to microginin FR1 (PDB entry 2j9a). The backbone structures are presented as wires with FR1 in spheres. LAP-A domains are colored as in Fig. 1(*a*). The two loops differ in the catalytic domains of the two structures, indicated by L1 and L2. Bovine lens LAP is in grey with FR1 in green for C atoms, orange for O atoms and blue for N atoms. (*b*) Stereoview of FR1 modeled at the substrate-binding channel of the LAP-A monomer. The surface of the LAP-A structure is shown with C atoms in white, O atoms in orange, N atoms in blue and S atoms in yellow. The two metal ions are shown as magenta spheres. FR1 is shown in ball-and-stick representation with C atoms in green, O atoms in red and N atoms in blue. (*c*) Covalent structural diagram of microginin FR1. Top, chemical diagram of FR1; middle, amino-acid sequence; bottom, the residue designations at the active site of LAP-A.

Both the LAP-A monomer and hexamer are structurally similar, as expected from sequence conservation, to other LAPs of the M17 peptidase family, including the bovine lens LAP. When the three-dimensional structure of LAP-A is superimposed on that of bovine lens LAP bound to the inhibitor microginin (FR1; PDB entry 2j9a; Kraft *et al.*, 2006), the catalytic domains align very well (Fig. 2*a*) except for two loops, residues 355–362 (L2 in Fig. 2*a*) and 549–556 (L1 in Fig. 2*a*) in LAP-A, which move closer at the top of the substrate-binding channel in the LAP-A structure (Fig. 2*a*). Therefore, more structural restrictions on substrate selection are expected for LAP-A than for bovine lens LAP.

3.2. Modeling of different substrates at the active site of LAP-A

Microginin FR1 is a pentapeptide analog (Ahda-Ala-mLeu-Tyr-Tyr; Fig. 2*c*) and is an inhibitor of many LAP enzymes (Kraft *et al.*, 2006; Okino *et al.*, 1993; Ishida *et al.*, 1997). Overall, FR1 fits quite well into the substrate-binding pocket of LAP-A (Fig. 2*b*). Therefore, each LAP-A monomer should have room to hold a pentapeptide at the active site.

The substrate-binding pocket of the LAP-A monomer is a spacious channel whose entrance is wide open to allow peptides to enter the channel and reach the metal-binding active site located at the end of the channel (Fig. 3*a*). The active site contains the sulfate ion (sitting on the positive blue surface in Fig. 3*a*) and two metal ions coordinated mostly by carboxylate O atoms (Fig. 1*d* and the negative red spotted surface in Fig. 3*a*) and the bridging water. The channel itself is so large that the docked FR1, a pentapeptide analog, looks like it is floating in space with few contacts with the walls of the channel (Fig. 2*b*), suggesting that there are few geometric restrictions on peptide binding by LAP-A. However, when a natural substrate of LAP-A binds, in order for the hydrolysis of the first peptide bond to occur the first (P1 position) and the second (P1′ position) amino acids must be positioned perfectly against the two metal ions, the bridging water and the sulfate ion so that they can facilitate the reaction (Fig. 3). The binding pockets for the P1 and P1′ amino acids are oriented like two 'rabbit ears' in front of the active site (Fig. 3*a*). The P1 pocket is long and wide enough to accom-

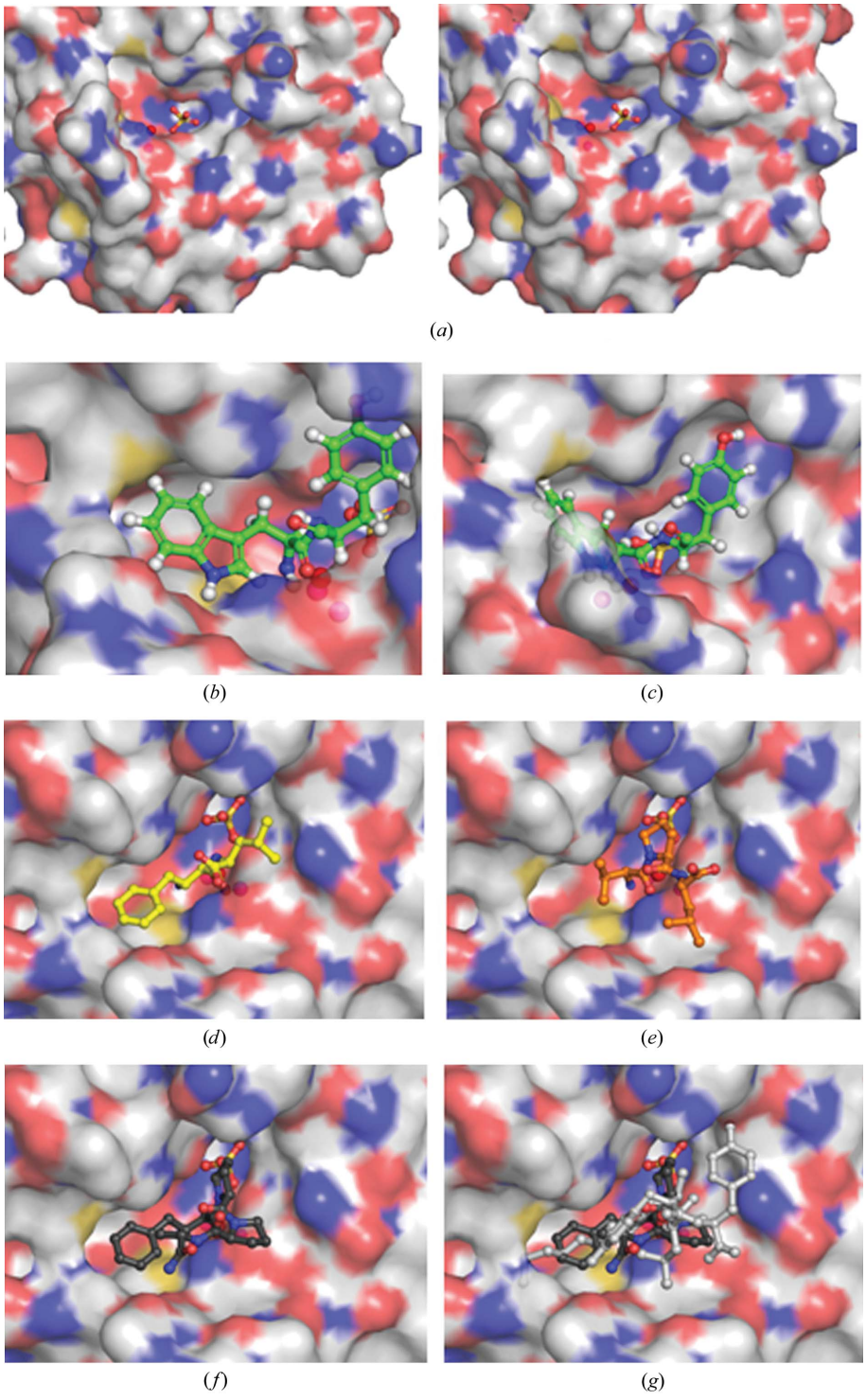

Figure 3
Different substrates fit into the substrate-binding channel of LAP-A. (*a*) Stereoview of the substrate-binding channel at the active site. (*b*) The P1 site occupied by Trp (green). The Trp-Tyr dipeptide is modeled by mutating FR1 in Fig. 2(*b*). (*c*) The P1′ site occupied by Tyr (green). (*d*) Bestatin at the active site. Bestatin (yellow) is modeled by superimposing the LAP-A monomer with *Pseudomonas putida* LAP (ppLAP) in complex with bestatin (PDB entry 3h8g; Kale *et al.*, 2010). (*e*) Val-Pro-Leu tripeptide at the active site. The tripeptide is modeled by superimposing the Val-Pro-Leu tripeptide (brown) bound to *E. coli* aminopeptidase P (PDB entry 2v3x; Graham & Guss, 2008) with apstatin in (*f*). (*f*) Apstatin at the active site. Apstatin (black) is modeled by superimposing apstatin bound to *E. coli* aminopeptidase P (PDB entry 1n51; Graham *et al.*, 2004) with bestatin in (*d*). (*g*) Comparison of apstatin (black) and FR1 (silver) at the active site. H atoms are shown as white balls in (*b*) and (*c*). LAP-A is presented as partially transparent surfaces colored as in Fig. 2(*b*).

modate all of the natural amino acids including tryptophan (Fig. 3*b*). The P1′ pocket is also large enough for many amino acids including Tyr (Fig. 3*c*). This suggests that the LAP-A enzyme can hydrolyze many peptides of varied sequences and lengths. Shown in Fig. 3 are different peptides or peptide analogs modeled at the substrate-binding channel of the active site starting from the dipeptide Trp-Tyr (Figs. 3*b* and 3*c*), the dipeptide analog bestatin (Phe-Leu analog; Fig. 3*d*), the tripeptide Val-Pro-Leu (Fig. 3*e*), the tetrapeptide analog apstatin (Phe-Pro-Pro-Ala analog; Fig. 3*f*) and the pentapeptide analog FR1 (Ahda-Ala-mLeu-Tyr-Tyr; Fig. 3*g*). The substrate-binding channel is long and large enough to hold pentapeptides at the active site of each LAP-A monomer.

In theory, longer peptides or proteins could bind to the LAP-A monomer since the substrate-binding channel is open to the solvent. However, the length of a peptide substrate appears to be limited by the hexameric assembly of the LAP-A enzyme; these substrate limitations are biologically relevant since LAP-A hexamers, but not monomers, are catalytically active (Gu *et al.*, 1999). Inside the LAP-A hexamer, the central cavity (yellow triangle in Fig. 1*c*) formed by the substrate-binding channels has room to allow each monomer to bind a substrate which may have at least one more amino acid than FR1. Therefore, if only geometry limitations are considered, the hexameric LAP-A enzyme may engage in the catalysis of six hexapeptides at the same time. However, a substrate can extend from its substrate-binding channel into the central cavity and outside the hexamer through the solvent channels between the two trimers if the rest of the active sites are either empty or occupied by substrates consisting of five amino acids or fewer (Figs. 1*c* and 4). Therefore, the hexameric LAP-A enzyme may bind one substrate with more than six amino-acid residues at a time. This structural feature is consistent with an early observation that *in vitro* LAP-A can process a putative 19-residue precursor of systemin (Leu-systemin) to its 18-amino-acid form (Gu & Walling, 2000).

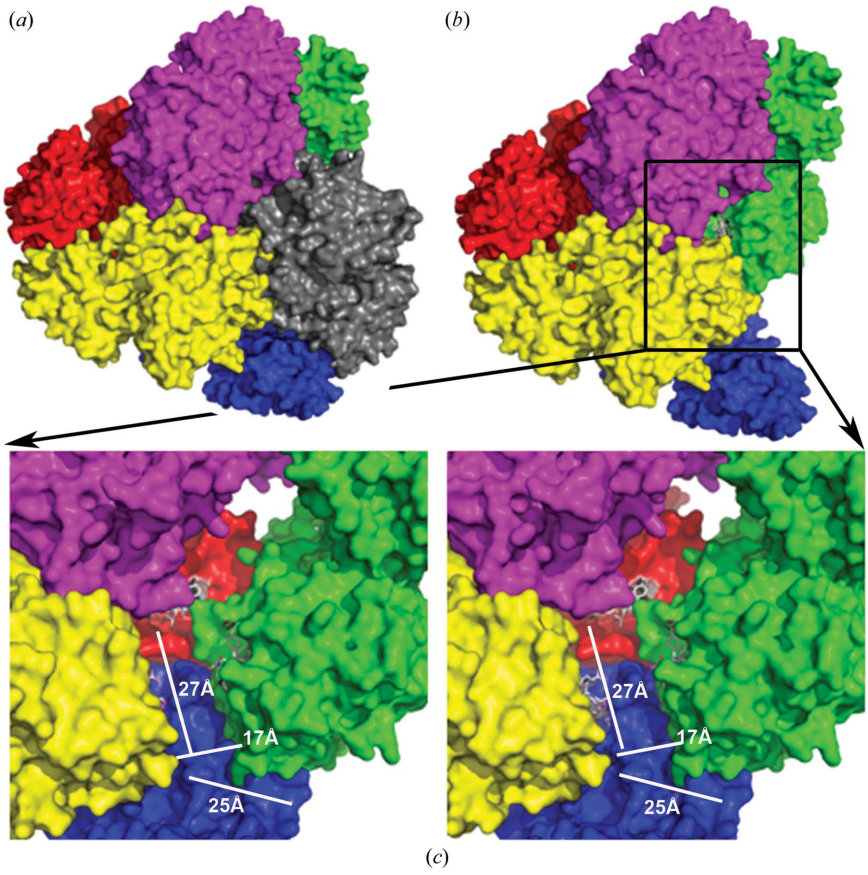

Figure 4
Substrate-binding channels of the hexameric LAP-A enzyme. (*a*) The hexameric LAP-A enzyme. LAP-A monomer surfaces are displayed in different colors to distinguish the monomers. (*b*) Hexameric LAP-A missing one monomer. (*c*) Stereoview of the possible substrate-binding channels for a long peptide in the hexameric LAP-A enzyme. FR1 molecules are shown in sticks with C atoms in white, O atoms in red and N atoms in blue. The dimensions of the channels are indicated by white lines.

Interestingly, the two possible substrate-binding channels allowing a longer peptide to bind are about 27 and 25 Å in length (Fig. 4*c*), respectively. Considering that the distance between two adjacent C^{α} atoms in a peptide is about 3.8 Å, these two channels together have the length to cover 12 amino-acid residues. Together with the six amino-acid residues bound at the substrate-binding channel of the active site, a long peptide of 18 residues fits nicely from the active site to the edge of the LAP-A hexamer.

The geometrical shape of the substrate-binding channel at the active site and the chemical requirements for breaking the first peptide bond suggest that the substrate specificity of LAP-A is mainly decided by the P1 and P1′ residue binding pockets (Fig. 3). The P1 pocket surface is a mixture of hydrophobic and negatively charged patches (white and red, respectively), while the P1′ pocket surface is a mixture of hydrophobic and positively charged patches (white and blue, respectively). This difference in chemical properties between the P1 and P1′ pockets is likely to provide a preference for different amino acids at the P1 and P1′ positions since these two spacious pockets provide few geometry selections. Based on the structural analysis, the P1 position would favor Leu, Met and Arg, as observed by early *in vitro* assays, while the P1′ position would be appropriate for Leu and Phe but not Arg (Gu & Walling, 2000). One special case is proline. Proline at P1′ is allowed in space and fits chemically (Figs. 3*e* and 3*f*), but its unusual structure places it too close to the sulfate ion and may disrupt the active site (Fig. 5); therefore, peptides with proline at the P1′ position may not be good substrates of LAP-A. The substrate-binding channel becomes much wider after the P1′ position, so fewer structural restrictions are placed on amino acids at the P2′ position and beyond. In addition, these amino acids are away from the reaction center and have more freedom to position themselves to fit the channel, so that they are expected to make lower or no contributions to the substrate specificity. For example, Ahda (P1) and Ala (P1′) of FR1 align well with Phe (P1) and Leu (P1′) of bestatin as well as Val (P1) and Pro (P1′) of the Val-Pro-Leu tripeptide (Fig. 5*a*). However, mLeu (P2′) of FR1 has a different orientation to Leu (P2′) of the tripeptide (Fig. 5*a*). More dramatically, Leu (P2′) of the tripeptide is positioned differently to Pro (P2′) of apstatin (Phe-HCOH-Pro-Pro-Ala) in Fig. 5(*b*).

4. Discussion

The tomato LAP-A enzyme exhibits the same hexameric oligomerization and conserved monomeric structure as observed in other published LAP enzyme structures (Kale *et al.*, 2010; Kraft *et al.*, 2006; Sträter, Sherratt *et al.*, 1999). However, each LAP enzyme is unique in both structure and substrate specificity. For example, the two loops forming the top part of the P1 pocket are closer to each other in the LAP-A structure than in the bovine lens LAP structure (Fig. 2*a*). Therefore, it is not surprising that the plant, animal and prokaryotic LAP enzymes displayed significant differences in their abilities to hydrolyze selected dipeptide and tripeptide substrates (Gu & Walling, 2000). Kinetic analysis (Gu & Walling, 2000, 2002) of the activity of tomato LAP-A towards 60 dipeptides and seven tripeptides indicated that the P1, P1′ and P2′ residues of the peptides influenced both the substrate affinity (K_m) and the catalytic ability (V_{max} and k_{cat}) of the enzyme. Amino acids at the P1′ position have the highest impact on the catalytic ability of the LAP-A enzyme. Results from the hydrolysis of nine Leu-Xaa dipeptides indicated 44-fold variations in V_{max} when different amino acids occupy the P1′ position, ranging from 440 µmol min^{-1} per milligram of LAP-A protein for the Leu-Phe dipeptide to 9.8 µmol min^{-1} per milligram of protein for the Leu-Asp dipeptide (see Table 3 in Gu & Walling, 2002). There is about a tenfold difference in V_{max} when

different amino acids occupy the P1 position based on the results of assays on ten Xaa-Leu dipeptides, ranging from 419 μmol min^{-1} per milligram of LAP-A protein for the Leu-Leu dipeptide to 45 μmol min^{-1} per milligram of LAP-A protein for the Tyr-Leu dipeptide (see Table 2 in Gu & Walling, 2002).

However, among the six tested Arg-Gly-Xaa substrates with amino acids of different sizes and properties at the P2′ position, the differences in V_{max} are limited to a sixfold range, with the lowest being 13 μmol min^{-1} per milligram of LAP-A protein for the Arg-Gly-Arg tripeptide and the highest 77 μmol min^{-1} per milligram of LAP-A protein for the Arg-Gly-Gly tripeptide (see Table 1 in Gu & Walling, 2000). These biochemical results support the prediction from the above structural analysis that the peptide-sequence specificity is predominately defined by the amino acids at the P1 and P1′ positions, and the amino acids at the P2′ position and beyond contributes less and less to sequence selection of the substrate as their positions are farther away from the first peptide bond.

The LAP-A enzyme is most efficient in the hydrolysis of dipeptides and their coversion to free amino acids. Among all of the dipeptide substrates tested, the LAP-A enzyme displayed the lowest activity for the hydrolysis of the Leu-Asp dipeptide, at 9.8 μmol hydrolyzed dipeptide per minute per milligram of LAP-A protein (Gu & Walling, 2000); that is, each LAP-A hexamer hydrolyzes 50 Leu-Asp dipeptides per second, a catalytic rate comparable to those of many proteases in the cell. Interestingly, the activity of the LAP-A enzyme for hydrolysis of tripeptides is usually 2–10-fold lower than for the hydrolysis of dipeptides according to results on the hydrolysis of the Arg-Gly dipeptide (139 μmol hydrolyzed dipeptides per minute per milligram of LAP-A protein) and Arg-Gly-Xaa tripeptides (77–13 μmol hydrolyzed tripeptides per minute per milligram of LAP-A protein) by the LAP-A enzyme (Gu & Walling, 2000). Considering these biochemical results together with the structure of the substrate-binding channel described in the report, we expect that the LAP-A enzyme hydrolyzes substrates of more than three amino acids with an efficiency that is the same as or lower than that for tripeptides. Although the LAP-A hexameric enzyme is able to hydrolyze peptides of more than six amino acids, it has the lowest efficiency in the hydrolysis of such long peptides. This observation can be explained well by the LAP-A structure. As described above, the six substrate-binding channels in the LAP-A hexameric enzyme can each hold a peptide of five or fewer amino acids and hydrolyze them simultaneously, but can only bind a single substrate with six or more amino acids at a time in order to avoid substrate collision at the central cavity if more long peptides are present.

LAP-A has only been found in a subset of the Solanaceae and is induced in response to both biotic and abiotic stresses (Tu *et al.*, 2003; Chao *et al.*, 1999, 2000). Currently, the biological substrates of LAP-A have yet to be discovered; however, the LAP-A substrates are likely to be found within chloroplast stroma or in passage through the stroma in plants or within the insect midgut (Narváez-Vásquez *et al.*, 2007). Several possible roles of LAP-A in the tomato defense response have been proposed (Pautot *et al.*, 1993; Chao *et al.*, 1999; Gu, Pautot *et al.*, 1996; Fowler *et al.*, 2009). (i) LAP-A may be involved in the rapid turnover of proteins to salvage C and N from cells committed to death owing to wounding or pathogen attack. The N-terminal residue of a protein strongly influences its rate of turnover and LAP-A may speed the turnover or extend the life of some proteins by removing their N-terminal residues (Varshavsky, 1996; Bradshaw *et al.*, 1998). (ii) LAP-A may process the N-termini of selected peptides or proteins, which may influence the magnitude or effectiveness of the plant wound/defense or water-deficit response. LAP-A modulates a plastid-derived retrograde signal to positively and negatively regulate a subset of nuclear genes after injury (Fowler *et al.*, 2009; Scranton *et al.*, 2012). (iii) LAP-A may be important for the degradation of stress-induced proteins that accumulate as a response to wounding or pathogen attack; decreasing levels of proteins that reduce plant cell viability may also be an important component in recovery

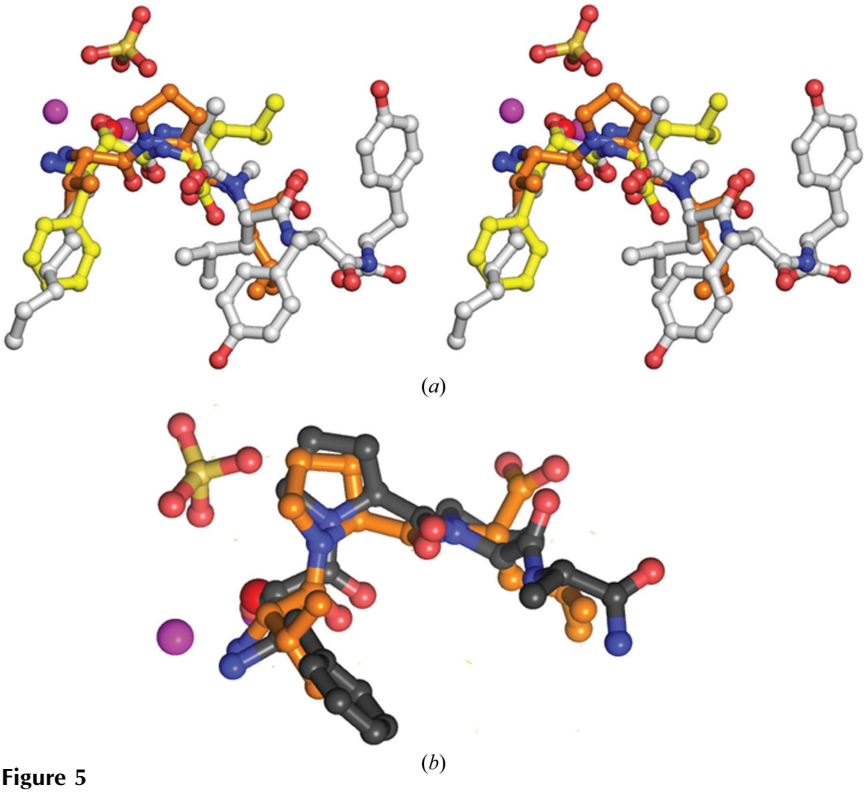

Figure 5
Comparison of different substrates at the active site of LAP-A. (*a*) Stereoview of FR1 (silver), bestatin (yellow) and Val-Pro-Leu (brown) in ball-and-stick representation. (*b*) The Val-Pro-Leu tripeptide is superimposed with apstatin; both are shown in ball-and-stick representation. The two metal ions (magenta) and bridging water (red) are shown together with the sulfate ion (yellow and brown) at the active site of LAP-A.

from biotic and/or abiotic stress. (iv) LAP-A may degrade peptides within the insect midgut (Gu *et al.*, 1999; Pautot *et al.*, 2001). While the delivery of LAP-A in artificial diets does not impact insect growth and development, the fact that LAP-A is one of the most stable proteins in the insect suggests that LAP-A may work in cooperation with other anti-nutritive enzymes such as wound-induced arginase to deplete essential arginine from the insect diet and thereby impact insect life-history parameters (Chen *et al.*, 2007; Fowler *et al.*, 2009).

The LAP-A structure along with the substrate modeling presented in this report indicates that the LAP-A enzyme is naturally built for these biological functions. The spacious substrate-binding channel at the active site places few restrictions on substrates, so that the LAP-A enzyme can hydrolyze many peptides or proteins with different sequences even if certain amino acids are preferred at the P1 or P1′ positions. In addition, we propose that the LAP-A enzyme makes use of its different efficiencies in hydrolyzing proteins, peptides and short peptides to fulfill its various biological functions. LAP-A has the lowest efficiency in the hydrolysis of proteins; however, this activity is likely to be biologically relevant since LAP-A only needs to remove the N-terminal amino acid of a protein to influence protein stability. While no N-end rule has been rigorously established in chloroplasts, stability determinants of stromal proteins in tobacco and the role of ClpS (a putative N-end rule adaptor) in substrate recognition by the ClpP protease in *Arabidopsis* have recently been described (Apel *et al.*, 2010; Nishimura *et al.*, 2013).

The fact that LAP-A has moderate activity towards peptides suggests that LAP-A may cleave peptides *in planta*. LAP-A may be involved in the turnover of ClpP protease-generated peptides, thereby facilitating the catabolism of peptides to their constituent amino-acid residues. A similar role for the processing of 26S proteasome-generated peptides has been proposed for mammalian LAP after interferon treatment and for the cytosolic *Arabidopsis* LAP1 after cadmium stress (Beninga *et al.*, 1998; Polge *et al.*, 2009; Towne *et al.*, 2005). In contrast, the role of LAP-A in catabolizing biologically active peptides in the stroma is not known, since only a single active peptide (glutathione) has been identified in plastids to date. Glutathione is an abundant tripeptide that is detected in all plant subcellular compartments and is a key regulator of redox homeostasis (Noctor *et al.*, 2012). However, its N-terminal γ conjugation (γ-Glu-Cys-Gly) makes it an unlikely substrate for LAP-A. Finally, the highest activity of LAP-A is towards dipeptides, suggesting that LAP-A may quickly hydrolyze dipeptides into free amino acids; this is a key role in the salvaging of C and N from cells committed to death owing to wounding or pathogen attack.

This work was partially supported by a National Science Foundation grant (IOS 0725093) to LLW.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.

Afonine, P. V., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl. Protein Crystallogr.* **42**, contribution 8.

Apel, W., Schulze, W. X. & Bock, R. (2010). *Plant J.* **63**, 636–650.

Bartling, D. & Nosek, J. (1994). *Plant Sci.* **99**, 199–209.

Bartling, D. & Weiler, E. W. (1992). *Eur. J. Biochem.* **205**, 425–431.

Beninga, J., Rock, K. L. & Goldberg, A. L. (1998). *J. Biol. Chem.* **273**, 18734–18742.

Bradshaw, R. A., Brickey, W. W. & Walker, K. W. (1998). *Trends Biochem. Sci.* **23**, 263–267.

Burley, S. K., David, P. R. & Lipscomb, W. N. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 6916–6920.

Chao, W. S., Gu, Y.-Q., Pautot, V., Bray, E. A. & Walling, L. L. (1999). *Plant Physiol.* **120**, 979–992.

Chao, W. S., Pautot, V., Holzer, F. M. & Walling, L. L. (2000). *Planta*, **210**, 563–573.

Charlier, D., Hassanzadeh, G., Kholti, A., Gigot, D., Piérard, A. & Glansdorff, N. (1995). *J. Mol. Biol.* **250**, 392–406.

Chen, H., Gonzales-Vigil, E., Wilkerson, C. G. & Howe, G. A. (2007). *Plant Physiol.* **143**, 1954–1967.

Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.

Cowtan, K. (2008). *Acta Cryst.* D**64**, 83–89.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fowler, J. H., Narváez-Vásquez, J., Aromdee, D. N., Pautot, V., Holzer, F. M. & Walling, L. L. (2009). *Plant Cell*, **21**, 1239–1251.

Graham, S. C. & Guss, J. M. (2008). *Arch. Biochem. Biophys.* **469**, 200–208.

Graham, S. C., Maher, M. J., Simmons, W. H., Freeman, H. C. & Guss, J. M. (2004). *Acta Cryst.* D**60**, 1770–1779.

Gu, Y.-Q., Chao, W. S. & Walling, L. L. (1996). *J. Biol. Chem.* **271**, 25880–25887.

Gu, Y.-Q., Holzer, F. M. & Walling, L. L. (1999). *Eur. J. Biochem.* **263**, 726–735.

Gu, Y.-Q., Pautot, V., Holzer, F. M. & Walling, L. L. (1996). *Plant Physiol.* **110**, 1257–1266.

Gu, Y.-Q. & Walling, L. L. (2000). *Eur. J. Biochem.* **267**, 1178–1187.

Gu, Y.-Q. & Walling, L. L. (2002). *Eur. J. Biochem.* **269**, 1630–1640.

Henson, H. & Frohne, M. (1976). *Methods Enzymol.* **45**, 504–520.

Ishida, K., Matsuda, H., Murakami, M. & Yamaguchi, K. (1997). *Tetrahedron*, **53**, 10281–10288.

Kale, A., Pijning, T., Sonke, T., Dijkstra, B. W. & Thunnissen, A.-M. W. H. (2010). *J. Mol. Biol.* **398**, 703–714.

Kim, H. & Lipscomb, W. N. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 5006–5010.

Kim, H. & Lipscomb, W. N. (1994). *Adv. Enzymol. Relat. Areas Mol. Biol.* **68**, 153–213.

Kraft, M., Schleberger, C., Weckesser, J. & Schulz, G. E. (2006). *FEBS Lett.* **580**, 6943–6947.

Matsui, M., Fowler, J. H. & Walling, L. L. (2006). *Biol. Chem.* **387**, 1535–1544.

McGowan, S., Oellig, C. A., Birru, W. A., Caradoc-Davies, T. T., Stack, C. M., Lowther, J., Skinner-Adams, T., Mucha, A., Kafarski, P., Grembecka, J., Trenholme, K. R., Buckle, A. M., Gardiner, D. L., Dalton, J. P. & Whisstock, J. C. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 2449–2454.

Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. (2009). *Acta Cryst.* D**65**, 1074–1080.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Narváez-Vásquez, J., Tu, C.-J., Park, S.-Y. & Walling, L. L. (2007). *Planta*, **227**, 341–351.

Nishimura, K., Asakura, Y., Friso, G., Kim, J., Oh, S., Rutschow, H., Ponnala, L. & van Wijk, K. J. (2013). *Plant Cell*, **25**, 2276–2301.

Noctor, G., Mhamdi, A., Chaouch, S., Han, Y., Neukermans, J., Marquez-Garcia, B., Queval, G. & Foyer, C. H. (2012). *Plant Cell Environ.* **35**, 454–484.

Okino, T., Matsuda, H., Murakami, M. & Yamaguchi, K. (1993). *Tetrahedron Lett.* **34**, 501–504.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pautot, V., Holzer, F. M., Chaufaux, J. & Walling, L. L. (2001). *Mol. Plant Microbe Interact.* **14**, 214–224.
Pautot, V., Holzer, F. M., Reisch, B. & Walling, L. L. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9906–9910.
Polge, C., Jaquinod, M., Holzer, F., Bourguignon, J., Walling, L. & Brouquisse, R. (2009). *J. Biol. Chem.* **284**, 35412–35424.
Saveanu, L., Fruci, D. & van Endert, P. (2002). *Mol. Immunol.* **39**, 203–215.
Scranton, M. A., Fowler, J. H., Girke, T. & Walling, L. L. (2013). *PLoS One*, **8**, e77889.
Scranton, M. A., Yee, A., Park, S.-Y. & Walling, L. L. (2012). *J. Biol. Chem.* **287**, 18408–18417.
Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985). *Anal. Biochem.* **150**, 76–85.
Stirling, C. J., Colloms, S. D., Collins, J. F., Szatmari, G. & Sherratt, D. J. (1989). *EMBO J.* **8**, 1623–1627.
Sträter, N. & Lipscomb, W. N. (1995*a*). *Biochemistry*, **34**, 9200–9210.
Sträter, N. & Lipscomb, W. N. (1995*b*). *Biochemistry*, **34**, 14792–14800.
Sträter, N., Sherratt, D. J. & Colloms, S. D. (1999). *EMBO J.* **18**, 4513–4522.
Sträter, N., Sun, L., Kantrowitz, E. R. & Lipscomb, W. N. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 11151–11155.
Taylor, A. (1993). *FASEB J.* **7**, 290–298.
Taylor, A., Daims, M., Lee, J. & Surgenor, T. (1982). *Curr. Eye Res.* **2**, 47–56.
Towne, C. F., York, I. A., Neijssen, J., Karow, M. L., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Neefjes, J. J. & Rock, K. L. (2005). *J. Immunol.* **175**, 6605–6614.
Tu, C.-J., Park, S.-Y. & Walling, L. L. (2003). *Plant Physiol.* **132**, 243–255.
Umezawa, H. (1980). *Recent Results Cancer Res.* **75**, 115–125.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.
Varshavsky, A. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 12142–12149.
Walling, L. L. (2006). *Curr. Opin. Plant Biol.* **9**, 227–233.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.