8-2019

Structural characterization of a prolyl aminodipeptidase (PepX) from *Lactobacillus helveticus*

Deanna Dahlke Ojennus
Whitworth University, dojennus@whitworth.edu

Recommended Citation
Ojennus, D. D., Bratt, N. J., Jones, K. L. & Juers, D. H. (2019). Acta Cryst. F75, 625-633. DOI: 10.1107/S2053230X19011774.

electronic reprint

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Structural characterization of a prolyl aminodipeptidase (PepX) from *Lactobacillus helveticus*

Deanna Dahlke Ojennus, Nicholas J. Bratt, Kent L. Jones and Douglas H. Juers

Acta Cryst. (2019). F**75**, 625–633

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structural characterization of a prolyl aminodipeptidase (PepX) from *Lactobacillus helveticus*

Deanna Dahlke Ojennus,[a]* Nicholas J. Bratt,[a] Kent L. Jones[b] and Douglas H. Juers[c]

[a]Department of Chemistry, Whitworth University, 300 West Hawthorne Road, Spokane, WA 99251, USA, [b]Department of Mathematics and Computer Science, Whitworth University, 300 West Hawthorne Road, Spokane, WA 99251, USA, and [c]Department of Physics and Program in Biochemistry, Biophysics and Molecular Biology, Whitman College, 345 Boyer Avenue, Walla Walla, WA 99632, USA. *Correspondence e-mail: dojennus@whitworth.edu

Received 17 May 2019
Accepted 26 August 2019

Edited by S. Sheriff, Bristol-Myers Squibb, USA

Keywords: prolyl aminodipeptidase; PepX; α/β-hydrolases; protein structure; calcium binding; calcium-blade zone; gluten; *Lactobacillus helveticus*.

PDB reference: prolyl aminodipeptidase from *Lactobacillus helveticus*, 6nff

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

Prolyl aminodipeptidase (PepX) is an enzyme that hydrolyzes peptide bonds from the N-terminus of substrates when the penultimate amino-acid residue is a proline. Prolyl peptidases are of particular interest owing to their ability to hydrolyze food allergens that contain a high percentage of proline residues. PepX from *Lactobacillus helveticus* was cloned and expressed in *Escherichia coli* as an N-terminally His-tagged recombinant construct and was crystallized by hanging-drop vapor diffusion in a phosphate buffer using PEG 3350 as a precipitant. The structure was determined at 2.0 Å resolution by molecular replacement using the structure of PepX from *Lactococcus lactis* (PDB entry 1lns) as the starting model. Notable differences between the *L. helveticus* PepX structure and PDB entry 1lns include a cysteine instead of a phenylalanine at the substrate-binding site in the position which confers exopeptidase activity and the presence of a calcium ion coordinated by a calcium-binding motif with the consensus sequence D*X*(DN)*X*DG.

1. Introduction

Aminopeptidases are of interest in the food industry for the production of protein hydrolysates, which are more easily digested, have an altered texture or taste, display improved antioxidant properties and have a reduced presence of common food allergens. The X-prolyl dipeptidyl aminopeptidase (PepX) from *Lactobacillus helveticus* is a 92 kDa peptidase which specifically hydrolyzes a peptide bond C-terminal to a penultimate proline residue at the N-terminus of a peptide. To date, the only known structure of a bacterial PepX enzyme is that from *Lactococcus lactis* (Rigolet *et al.*, 2002). *L. lactis* PepX has 34% sequence identity to the *L. helveticus* enzyme and >50% sequence similarity (Altschul *et al.*, 1990). *L. helveticus* PepX is known to form dimeric assemblies through noncovalent interactions (Vesanto *et al.*, 1995) and the *L. lactis* structure was found to be a homodimer with twofold rotation symmetry, in which each monomer contains a single active site within an α/β-hydrolase fold (Rigolet *et al.*, 2005).

L. helveticus is a lactic acid-fermenting bacterium that is commonly used in the fermentation of milk for commercial cheese production. It is known to be effective in breaking down potentially allergenic milk proteins which contain a high content of proline residues. A recombinant *L. helveticus* PepX enzyme, both alone and in combination with PEPN (a general aminopeptidase), has been shown to improve the antioxidant effect of casein hydrosylates as well as to decrease the

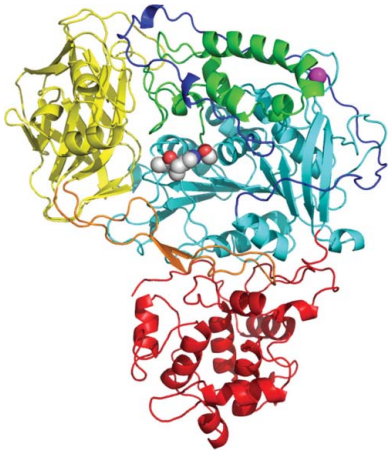

Table 1
Cloning and expression of PepX.

Source organism	*L. helveticus*
DNA source	Leeners helveticus cheese starter culture
Forward primer†	TCATTCACATCATATGAAATATAACCAATA TGCTTACG
Reverse primer‡	TCACATCTCGAGTTATTTTTCATAAAACTT GATTTCAGT
Cloning/expression vector	Novagen pET-14b
Expression host	*E. coli* BL21(DE3)pLysS
Complete amino-acid sequence of the construct produced§	MGSSHHHHHHSSGLVPRGSHMKYNQYAYVE TDFQQQVKELIDINFLPKNYQVWDFGSL LAKLVKNAIAEAKTDAAKNAKLAEFAVS DHQTLADFLKEKPTEIGTKQFYNVALQL LGYHVHYDYDFADPTGFMQRNALPFLQD ISDNQKLISAFYRLLNTRAKNGQILLDV MAGKGYFTQFWGQNKFKFFNGKSIPVFD TNKVIREVVYVETDLDTDHDGKSDLIQV TVFRPEETNKGLKVPALYTASPYFGGII ANEKRNHNVDENLSDSTEWNDPQYVHSP IVKAEKPDGSSRPATEEAVHKSSYPLNE YMLARGFASVFAGAIGTRGSDGVRITGA PEETESAAAVIEWLHGDRVAYTDRTRTV RTTADWCNGNIGMTGRSYLGTLQIAIAT TGVKGLKTVVSEAAISSWYDYYREHGSV IAPEACQGEDLDLLAETCQSNLWDAGSY LKIKPEYDKMQKQLREKEDRNTGQYSDF WEAGNYRHHADGIKCSWISVHGLNDWNV KPKNVYKIWQLVKKMPMKHHLFLHQGPH YNMNNLVSIDFTDLMNLWFVHELLGIEN NAYNQWPTVMIQDNLQADKWHEEPDWSN DLGQEKIYYPTDEGELFQDGNGKAQKSF TDVGGIEFKKAGISESDWQYKFICGDEK WAKPSLRFETDEFTHPTTIVGRPEVKVR VSASLPKGEISVALVELGERQRLTATPK FLMHGGQELGYRFGTDTLQEFVPDKKTK AKLITKAHMNLQNFKDMKKPEAIDADKF YDLDFLLQPTYYTIPSGSKLALIIYSTD QGMTKRPLEDETYTIDLANTEIKFYEK

† The forward primer contains an NdeI restriction site. ‡ The reverse primer contains an XhoI restriction site. § The amino-acid sequence includes a 20-residue N-terminal affinity tag prior to the first residue of the PepX sequence.

bitterness typically caused by large hydrophobic peptides that contain a high percentage of prolines (Stressler *et al.*, 2013, 2016).

Peptidases from lactobacilli, as well as fungal or other bacteria-derived prolyl endopeptidases (PEPs), have been suggested for use in the enzymatic detoxification of gluten, a food allergen found in wheat, rye and barley that is toxic to people with celiac disease (CD; Caputo *et al.*, 2010; Wieser & Koehler, 2012). Reports of CD patients being able to tolerate sourdough breads better than other gluten-containing breads could possibly be owing to the activity of lactobacilli peptidases (Gerez *et al.*, 2008). The gliadin and glutenin proteins found in gluten, like milk proteins, also contain a high percentage of proline residues that are not easily digested by our own gastric and pancreatic enzymes. For example, a CD immunogenic 33-mer peptide from α2-gliadin has been shown to persist in both *in vivo* and *in vitro* assays against intestinal peptidases (Shan *et al.*, 2002). Lactobacilli-extracted peptidases, including PepX, have shown promise in detoxifying gluten peptides (Gerez *et al.*, 2008; De Angelis *et al.*, 2010; Francavilla *et al.*, 2017; Gallo *et al.*, 2005); however, they are not particularly active or stable at low pH and are also susceptible to other digestive proteases. Regardless of these known limitations, it has been suggested that PepX in combination with PEPN and possibly PEPO (a neutral endopeptidase) would be sufficient to completely detoxify gluten (De Angelis *et al.*, 2010; Gallo *et al.*, 2005).

Here, we report the crystal structure of the *L. helveticus* PepX enzyme at 2.0 Å resolution as well as the activity of the recombinant enzyme on a gluten peptide.

2. Materials and methods

2.1. Cloning, expression and purification

The *pepx* gene was PCR-amplified from an *L. helveticus* cheese starter culture (Leeners). The identity of the bacterial isolates grown on MRS plates was confirmed at the species level by 16S rRNA gene sequencing using universal 8F and 1492R primers (SeqWright). The gene was amplified with forward and reverse primers (Invitrogen) containing NdeI and XhoI restriction sites (Table 1). The *pepx* gene was ligated into a pET-14b expression vector to attach an N-terminal histidine tag and placed into an *Escherichia coli* BL21(DE3)pLysS cell line (Novagen). The *pepx* gene was sequenced (SeqWright) and the translated protein sequence was found to be a 99% match to NCBI reference sequence WP_003629197.1, with the following four amino-acid variations: S36G, Q347R, L401S and R462G. *E. coli* cultures were grown at 37°C in LB-Miller broth until they reached an OD_{600} of between 0.5 and 1.0, and protein expression was induced by the addition of 0.8 µ*M* IPTG. The cultures were grown for an additional 4 h at room temperature and the cells were harvested by centrifugation.

The PepX protein was purified using procedures that were slightly modified from those reported previously (Vesanto *et al.*, 1995; Stressler *et al.*, 2013). The cells were resuspended in 20 m*M* Tris buffer pH 8.0, 1 m*M* EDTA pH 8.0, 0.5 mg ml^{-1} lysozyme and were allowed to incubate on ice with stirring for 1 h. The lysate was doubled in volume with water and brought to 0.25 U ml^{-1} DNAse I (New England Biolabs), 5 µg ml^{-1} pancreatic RNAse A (Sigma), 1 m*M* $MgCl_2$, 0.1 m*M* $CaCl_2$. The lysate was stirred at room temperature for 30 min and then sonicated. The cell debris was removed by centrifugation and the soluble lysate was fractionated by ammonium sulfate precipitation (30 and 80% saturation). PepX found in the 80% fractionated pellet was resuspended and dialyzed in 50 m*M* potassium phosphate buffer pH 6.5. PepX was purified by chromatography using a Nickel HiTrap IMAC Sepharose FF column (GE Healthcare Life Sciences) with a running buffer consisting of 50 m*M* potassium phosphate pH 6.5, 20 m*M* imidazole, 500 m*M* NaCl, eluting with a linear gradient to 500 m*M* imidazole. Fractions containing PepX as determined by SDS–PAGE were combined, dialyzed into 50 m*M* potassium phosphate buffer pH 6.5 and further purified by HiTrap Q FF anion-exchange chromatography (GE Healthcare Life Sciences) using a running buffer consisiting of 50 m*M* Tris pH 7.2 and eluting with a linear gradient to 1.0 *M* NaCl. Fractions containing PepX were combined and dialyzed into 50 m*M* potassium phosphate buffer pH 6.5.

Table 2
Crystallization.

Method	Hanging-drop vapor diffusion
Plate type	VDXm plate with siliconized glass cover slides (Hampton Research)
Temperature (K)	296
Protein concentration (mg ml^{-1})	4–7
Buffer composition of protein solution	50 m*M* potassium phosphate buffer pH 6.5
Composition of reservoir solution	75 m*M* potassium phosphate buffer pH 6.0, 1 m*M* DTT, 6.8–7.2% PEG 3350
Volume and ratio of drop	5 µl; 1:1 and 1:5
Volume of reservoir (ml)	1

2.2. Crystallization

The PepX used for crystallization was concentrated to between 4 and 7 mg ml^{-1} using Corning Spin-X UF Concentrators (10 000 molecular-weight cutoff). Rhomboid-shaped crystals formed after 1–2 weeks of growth, with the largest crystals ranging between 200 and 300 µm in size (Table 2).

2.3. Crystal mounting, X-ray data collection, processing and refinement

The crystals were manipulated under a humid air flow (∼95% relative humidity; Farley *et al.*, 2014). The crystals were first mounted at room temperature (295 K) in MicroRT tubes (MiTeGen, Ithaca, New York, USA) using nylon loops (Hampton Research, Aliso Viejo, California, USA) or micromeshes (MiTeGen). A small amount (∼20 µl) of the crystal well solution was added to the MicroRT tube before mounting. The PepX crystals diffracted with visible spots to 4 Å resolution at room temperature and unit-cell parameters could be determined, but the diffraction did not last very long (∼1 h). Low-temperature diffraction was therefore pursued. Crystals were mounted in a cold nitrogen-gas stream (Cryojet, Oxford Instruments, Abingdon, England; 100 K sample/shield flows at 6/4 l min^{-1}, respectively) using the vial-mounting approach (Farley *et al.*, 2014). Crystals were tested with (listed with decreasing levels of thermal contraction) methanol, 2-methyl-2,4-pentanediol, ethylene glycol, glycerol and xylose as cryoprotective agents (Alcorn & Juers, 2010). Of these, the best results were observed with 30–40% ethylene glycol; however, the diffraction was weak. Consequently, a highly redundant data set was collected at 100 K to improve the accuracy of the high-resolution data. For the data set reported here, a cover slip containing PepX crystals was transferred to a humidity-controlled stage for crystal manipulation. A PepX crystal (180 × 200 × 300 µm) was transferred from the cover slip to an ∼20 µl drop of solution consisting of synthetic mother liquor [7%(*w*/*v*) PEG 3350, 75 m*M* potassium phosphate pH 6.0] supplemented with 40%(*v*/*v*) ethylene glycol for 30 s. The crystal was mounted with a nylon loop and transferred to a vial containing 500 µl of the cryosolution for transfer to the cold gas stream. The crystal was simultaneously mounted in the gas stream as the vial was removed.

X-ray data were collected using an Oxford Diffraction Nova X-ray source (50 kV/0.8 mA) and an Oxford Diffraction Onyx CCD detector. Bragg diffraction spots were visible to 3.0 Å resolution, so a data-collection strategy was created with 2.0 Å as a target resolution using *CrysAlis^{Pro}* (Oxford Diffraction). Data were integrated and scaled in *CrysAlis^{Pro}* and were merged with *AIMLESS* (Evans, 2006) and *CTRUNCATE* (French & Wilson, 1978; Winn *et al.*, 2011). Using the criterion of $CC_{1/2}$ = 0.25, a resolution limit of 2.0 Å was chosen (Table 3; Karplus & Diederichs, 2015).

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Oxford Diffraction Nova sealed-tube micro-focus
Wavelength (Å)	1.54056
Temperature (K)	100
Detector	Oxford Diffraction Onyx CCD
Crystal-to-detector distance (mm)	85.0
Rotation range per image (°)	0.3
Total rotation range (°)	604.8
Exposure time per image (s)	60
Space group	$P4_32_12$
a, *b*, *c* (Å)	154.63, 154.63, 106.59
α, β, γ (°)	90, 90, 90
Mosaicity (°)	1.02
Resolution range (Å)	30.33–2.00 (2.04–2.00)
Total No. of reflections	1599555 (45905)
No. of unique reflections	87344 (4411)
Completeness (%)	100.0 (100.0)
Multiplicity	18.3 (10.4)
$\langle I/\sigma(I)\rangle$	8.1 (0.5)†
R_{meas}	0.40 (4.95)
$CC_{1/2}$	0.99 (0.25)
Overall *B* factor from Wilson plot (Å^2)	21.4

† $CC_{1/2}$ was used to determine the resolution cutoff. $\langle I/\sigma(I)\rangle$ falls below 2.0 at 2.57 Å resolution.

Although the data set was highly redundant, the anomalous signal was weak. Consequently, the structure was solved using molecular replacement with *Phaser* in *PHENIX* (McCoy *et al.*, 2007; Adams *et al.*, 2010) with PDB entry 1lns as a starting model (Rigolet *et al.*, 2002). This is an X-prolyl dipeptidyl aminopeptidase from *L. lactis* with 34% amino-acid identity to PepX. The molecular-replacement solution was then used as input to *AutoSol* in *PHENIX*. In this case, because the sequence identity between the search model and the target structure was low, rather than rebuilding the search model, *AutoBuild* in *PHENIX* built the target structure from scratch using as input the amino-acid sequence (NCBI reference amino-acid sequence WP_003629197.1; PepX from *L. helveticus*) and the output map from *Phaser* density-modified with *RESOLVE* (Terwilliger, 2003). The autobuilt structure was then inspected residue by residue using $F_o - F_c$, $2F_o - F_c$ and, in places of ambiguity, OMIT electron density. After a few cycles of refinement in *PHENIX* and model building in *Coot* (Emsley *et al.*, 2010), the four amino-acid changes identified above were confirmed in the electron-density maps. Sample electron density is shown in Supplementary Fig. S1. Additional rounds of refinement and model building, including the addition of solvent molecules, resulted in a final model containing 794 amino-acid residues, water molecules, ethylene glycol molecules, a calcium ion and four phosphate ions. The protein chain includes all 793 residues of the cloned sequence

as well as one of 20 extra N-terminal residues (which include a six-His tag for purification). The remaining 19 residues are not visible in the electron-density map, and appear to occupy a relatively large solvent channel in the crystal (Table 4). The coordinates and structure factors have been deposited in the Protein Data Bank with code 6nff.

2.4. Activity assays

The Michaelis constant (K_m) for purified PepX at 1 n*M* was measured using varying concentrations of the dipeptide substrate mimic Gly-Pro-*p*-nitroanilide hydrochloride (GPpNA; Sigma) in 50 m*M* MES buffer pH 6.5 and monitoring the change in absorption at 410 nm over time at 40°C. Successive dipeptidase activity was monitored by digestion of the octamer peptide FPQPQLPY (Biomatik), with PepX at 100 n*M* and octamer peptide at 1 mg ml^{-1} in 0.1 *M* Tris pH 8.0 at room temperature (23 ± 1°C). Time points were taken at 0, 10, 45 and 120 min and were analyzed by RP-HPLC (Restek, C18, 5 µm, 100 × 4.6 mm). The peptides eluted during a linear gradient from 5 to 30% acetonitrile (in water and 0.1% TFA) at a flow rate of 1 ml min^{-1} over 20 min.

Table 4
Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	30.33–2.00 (2.02–2.00)
Completeness (%)	97.6 (94.0)
No. of reflections, working set	80812 (2541)
No. of reflections, test set	4332 (130)
Final R_{cryst}	0.236 (0.476)
Final R_{free}	0.273 (0.447)
Estimated coordinate error (Å)	0.35
No. of non-H atoms	
Total	7050
Protein	6400
Ion (Ca^{2+}, phosphate)	21
Ligand (ethylene glycol)	68
Water	561
R.m.s. deviations	
Bonds (Å)	0.008
Angles (°)	1.2
Average *B* factors (Å^2)	
Overall	34
Protein	34
Ion (Ca^{2+}, phosphate)	43
Ligand (ethylene glycol)	38
Water	36
Ramachandran plot†	
Most favored (%)	96.0
Allowed (%)	3.8

† The Ramachandran plot was produced with *MolProbity* (Chen *et al.*, 2010)

3. Results and discussion

3.1. Overall structure of PepX from *L. helveticus*

The structure of PepX from *L. lactis* (PDB entry 1lns) is the prototype of the S15 peptidase family of the SC clan (Rigolet *et al.*, 2002). The structure of PepX from *L. helveticus* (PDB entry 6nff) is the second structure determined for this clan and family of serine peptidases. The crystal structure was determined at 2.0 Å resolution (Tables 3 and 4) and PDB entry 6nff exhibits a four-domain structure with an overall shape that resembles the continent of Africa, as initially pointed out by Rigolet *et al.* (2002) (Fig. 1). For the structure with PDB code 1lns, the crystals were also grown with PEG (4000) as precipitant at pH 5.2 and the X-ray data were collected at 298 K.

The catalytic domain of PDB entry 6nff (residues 183–388 and 463–573) contains a canonical α/β-hydrolase fold (Carr & Ollis, 2009). It consists of a ten-stranded β-sheet (1, 2, 4, 3, 5, 6, 7, 8, 9 and 10) in which strands 2 and 10 are antiparallel to the rest (Fig. 2). Helices intervene between all strands, except between strands 1 and 2 and between strands 9 and 10, to give a total of ten helical structures (A, A′, A″, B, C, D, E, F, G and G′). The conserved catalytic triad (Ser363, Asp483 and His514) is located N-terminal to helix D, within the loop between β-strand 7 and helix F and the loop between β-strand 8 and helix G. The catalytic domain also contains a long loop that is described as a 'lasso' loop by Rigolet *et al.* (2002) in the structure with PDB code 1lns and is unique among α/β-hydrolase folds. This loop (residues 228–286) is located between β-strand 3 and helix B and surrounds the smaller helical domain while also making contacts with the C-terminal domain. Although PDB entries 6nff and 1lns both exhibit the same ten-stranded β-sheet pattern, there are several differences in the secondary-structural elements that connect the strands. For example, strands 2 and 3 are connected by a helix in PDB entry 6nff but not in PDB entry 1lns. A connecting helix between strands 2 and 3 is not part of the typical eight-stranded canonical α/β-hydrolase fold, but is occasionally observed in α/β-hydrolases (Carr & Ollis, 2009) such as plant carboxylesterases (Ileperuma *et al.*, 2007). Also, a small additional β-sheet structure, which also deviates from the canonical fold, is observed in the regions between strands 4 and 5 of PDB entry 1lns but is absent in PDB entry 6nff.

The two amino-acid Ramachandran outliers reported in Table 4 (using *MolProbity*) are located at positions Ala304 and Pro513. Two additional outliers, as reported by *DeepView*/*Swiss-PdbViewer* (Guex & Peitsch, 1997), are His100 and Ser363. The electron density for these four residues is unambiguous. His100 is rather far from the active site, but appears to be involved in the dimer contact. Ala304, Ser363 and Pro513 are all in the catalytic domain. Ser363 is the catalytic nucleophile, while Pro513 immediately precedes the catalytic histidine in the amino-acid sequence. Steric strain is frequently observed at the catalytic serine in α/β-hydrolases to form the 'nucleophile elbow' and oxyanion hole (Carr & Ollis, 2009). Here, we see that steric strain may also be required to correctly position the residue neighboring the catalytic histidine. Dimitriou and coworkers described a 'catalytic acid zone' for α/β-hydrolases (Dimitriou *et al.*, 2017, 2019) that is necessary for proper coordination of the catalytic histidine within the catalytic triad. Of the seven representatives of group A catalytic acid zone structures that they analyzed across six protein families, the equivalent histidine neighboring position is found to have higher backbone energy strain than the catalytic histidine in all but one structure (Supplementary Table S1). The histidine neighbor in three of the seven structures falls into an allowed, although not favored,

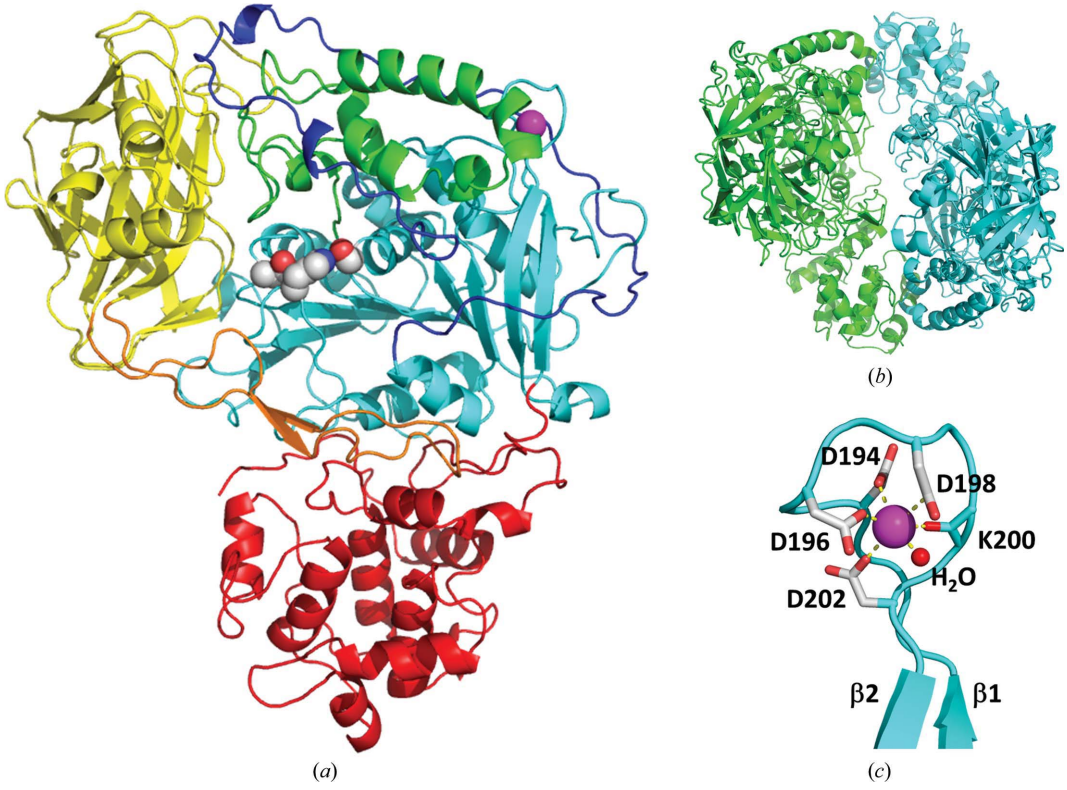

Figure 1
Three-dimensional structure (PDB entry 6nff) of *L. helveticus* PepX (*a*) shown with the N-terminal domain in red, the C-terminal domain in yellow, the helical domain in green and the α/β-hydrolase domain in cyan. The active-site residues located in the catalytic domain are shown as spheres in CPK colors and a bound calcium ion, also located within the catalytic domain, is shown as a magenta sphere. The 'lasso loop' (residues 228–286) of the catalytic domain is shown in dark blue and the extended loop of the C-terminal domain (672–710) is shown in orange. (*b*) Structure of the dimer observed in the crystal looking down a twofold rotational crystal symmetry axis. The dimer shown is supported as biologically relevant by calculations with *PISA*, suggesting the burial of 4250 Å^2 at the shown interface and a dissociation free energy of 9.1 kcal mol^{-1} (Krissinel & Henrick, 2007). (*c*) The calcium-binding motif located between β1 and β2 of the catalytic domain. The calcium ion exhibits octahedral coordination with Asp194 (2.4 Å), Asp196 (2.3 Å), Asp198 (2.5 Å), Asp202 (2.4 Å), the carbonyl of Lys200 (2.3 Å) and a water molecule, HOH1137 (2.3 Å). The water molecule also appears to be within hydrogen-bonding distance of Asp198 and the backbone carbonyl of Ser201 (not shown). This figure was prepared with *PyMOL* (Schrödinger).

region of the Ramachandran plot (using *DeepView*): Asp234 in hydroxynitrilase (PDB entry 3c6x), Asp237 in salicylic acid-binding protein 2 (PDB entry 1xkl) and Gly555 in acylamino-acid-releasing enzyme (PDB entry 2hu5) (Dimitriou *et al.*, 2017; Schmidt *et al.*, 2008; Forouhar *et al.*, 2005; Kiss *et al.*, 2007).

The helical domain (residues 389–462) is between strand 6 and helix E of the catalytic domain and makes substantial contacts with it. Cap domains involved in substrate binding are not uncommon in α/β-hydrolase folds (Carr & Ollis, 2009). It consists of five helices (three α-helices and two 3_{10}-helices) and provides important contacts at the substrate-binding site that are proposed to position the substrate and confer exopeptidase activity (Supplementary Table S2; Rigolet *et al.*, 2002, 2005). The helical domain also makes contact with the C-terminal domain and is surrounded by the 'lasso' loop of the catalytic domain.

The N-terminal domain of PDB entry 6nff (residues 1–182) consists of a bundle of eight α-helices very similar to the domain found in PDB entry 1lns. At the time that the structure with PDB code 1lns was published (Rigolet *et al.*, 2002), no structural homology was observed to anything else in the PDB for the N-terminal domain, and this continues to be the case for PDB entry 6nff. SCOPe lists the N-terminal PepX domain as its own family and superfamily (a.40.2 and a.40.2.1; Murzin *et al.*, 1995; Fox *et al.*, 2014). This superfamily is one of three identified as containing a CH-like domain protein fold among the all-α-helical protein domain class. The N-terminal domain was proposed to be responsible for the dimerization of PDB entry 1lns owing to extensive contacts between monomers, and this is also observed in our structure (Fig. 1; Rigolet *et al.*, 2002). The crystal dimer is likely to be physiologically relevant, as dimerization of *L. helveticus* PepX in solution has been observed and reported previously (Vesanto *et al.*, 1995).

The C-terminal domain (residues 574–793) highly resembles the SCOPe galactose-binding domain-like superfamily, with a twisted β-sandwich structure (Murzin *et al.*, 1995; Fox *et al.*, 2014). Three structurally homologous domains are known in the PepX C-terminal domain-like family (b.18.13): α-amino-acid ester hydrolase from *Acetobacter pasteurianus* and *Xanthomonas citri*, the C-terminal domain of bacterial cocaine esterase from *Rhodococcus* sp. MB1 and the PepX C-terminal domain from *L. lactis*. The C-terminal domain of PDB entry 6nff also contains an unusual long loop (residues 672–710) that wraps around the catalytic domain, as was also observed for PDB entry 1lns (Fig. 1; Rigolet *et al.*, 2002).

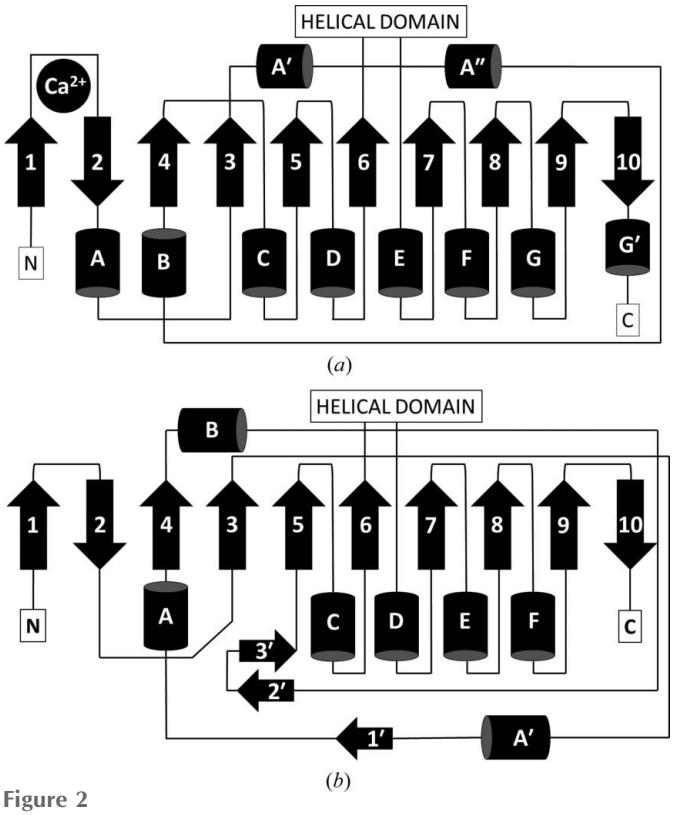

Figure 2
Secondary-structural map as assigned by *STRIDE* (Heinig & Frishman, 2004) for the α/β-hydrolase domain of *L. helveticus* PepX (*a*) compared with *L. lactis* PepX (*b*).

TM-align (Zhang & Skolnick, 2005) was used to align the structures with PDB codes 1lns and 6nff and reported an aligned length of 755 residues with an overall r.m.s.d. of 2.4 Å for C^{α} atoms. The normalized TM-score was 0.94, which strongly indicates high structural similarity. Regions of divergence in the structures of the PDB entries 6nff and 1lns primarily occur within the loops between secondary structures. However, three major regions of divergence, in which more than ten sequential amino acids diverge by 2 Å or more, are observed within the 'lasso loop', the second and third helices of the N-terminal domain and the first two strands and helices of the C-terminal domain (Supplementary Fig. S2).

3.2. Substrate-binding site

The substrate-binding site is found at residues surrounding the catalytic triad (Ser363, Asp483 and His514) within the catalytic domain as well as neighboring residues in the helical domain (Supplementary Table S2). Although no structure is available of any PepX bound to substrate, the binding site is nearly identical to human dipeptidyl peptidase-4 (DPP-IV), a functional homolog for which a structure bound to substrate is known (PDB entry 1r9n; Rigolet *et al.*, 2005; Aertgeerts *et al.*, 2004). Thus, parallels can be drawn to describe the PepX binding site (Rigolet *et al.*, 2005).

Few differences exist between the catalytic and helical domains of PDB entries 6nff and 1lns. The number of strands in the central β-sheet is the same, but PDB entry 6nff has a helix between strands 2 and 3 while PDB entry 1lns does not. PDB entry 1lns also has an additional small β-sheet structure between strands 4 and 5, while PDB entry 6nff does not (Fig. 2 and Supplementary Fig. S3). Despite these differences, the residues at the substrate-binding site are strictly conserved and superimpose well (Fig. 3 and Supplementary Table S2).

The only significant difference between the two structures in the active-site region is the residue corresponding to Cys408, which is a phenylalanine (Phe393) in PDB entry 1lns (and a glutamate, Glu205, in PDB entry 1r9n). This residue was postulated by Rigolet *et al.* (2005) to be responsible for exopeptidase activity in PepX by interacting with the N-terminus of the substrate, similar to Glu205 in DPP-IV, where the glutamate can hydrogen-bond to the N-terminus and, along with a second glutamate, block the entry of the substrate any further into the interior of the protein. As a consequence, only the second peptide bond of the substrate is positioned for nucleophilic attack by the catalytic serine, conferring exopeptidase activity rather than endopeptidase activity. This amino-acid difference may explain the differences in specificities for dipeptide substrate mimics that have been reported based on the identity of the N-terminal amino acid (Stressler *et al.*, 2013; Degraeve & Martial-Gros, 2003; Zevaco *et al.*, 1990). PepX from *L. lactis* hydrolyzes the dipeptide substrate mimic RPpNA (116% relative activity; Zevaco *et al.*, 1990) faster than APpNA (100% relative activity), while PepX from *L. helveticus* has been reported to hydrolyze APpNA faster (100% relative activity) than RPpNA (70.7 and 59% relative activity; Stressler *et al.*, 2013; Degraeve & Martial-Gros, 2003). Although the structural reason for this selectivity is not clear, the difference in relative activities does not appear to be a question of steric hindrance when considering that arginine paired with phenylalanine is more favorable than alanine as the N-terminal residue in the substrate. This suggests a favorable cation–π interaction in PDB entry 1lns that is absent in PDB entry 6nff. However, when looking at the interactions between binding-site residues and the YPSK peptide in the DPP-IV structure (PDB entry 1r9n; Aertgeerts *et al.*, 2004), this position (Glu205 in PDB entry 1r9n, Phe393 in PDB entry 1lns and Cys408 in PDB entry 6nff) appears to interact solely with the amide backbone of the N-terminal residue and not the side chain (Fig. 3).

3.3. Calcium-binding site

Another significant difference between the structures with PDB codes 6nff and 1lns is the presence of a calcium-binding site (Figs. 1 and 2). The amino-acid sequence of *L. helveticus* contains a D*X*(DN)*X*DG consensus sequence for a calcium-binding motif (Supplementary Fig. S3; Rigden *et al.*, 2011). Unlike the canonical EF-hand structure, in which this consensus sequence occurs in a loop between two α-helices (Kawasaki *et al.*, 1998; Gifford *et al.*, 2007), the structure with PDB code 6nff shows this sequence occurring in a loop between two β-strands (strands 1 and 2; see Figs. 1 and 2) of the catalytic domain. This structure most closely resembles a calcium blade, a term coined by Rigden and coworkers, who

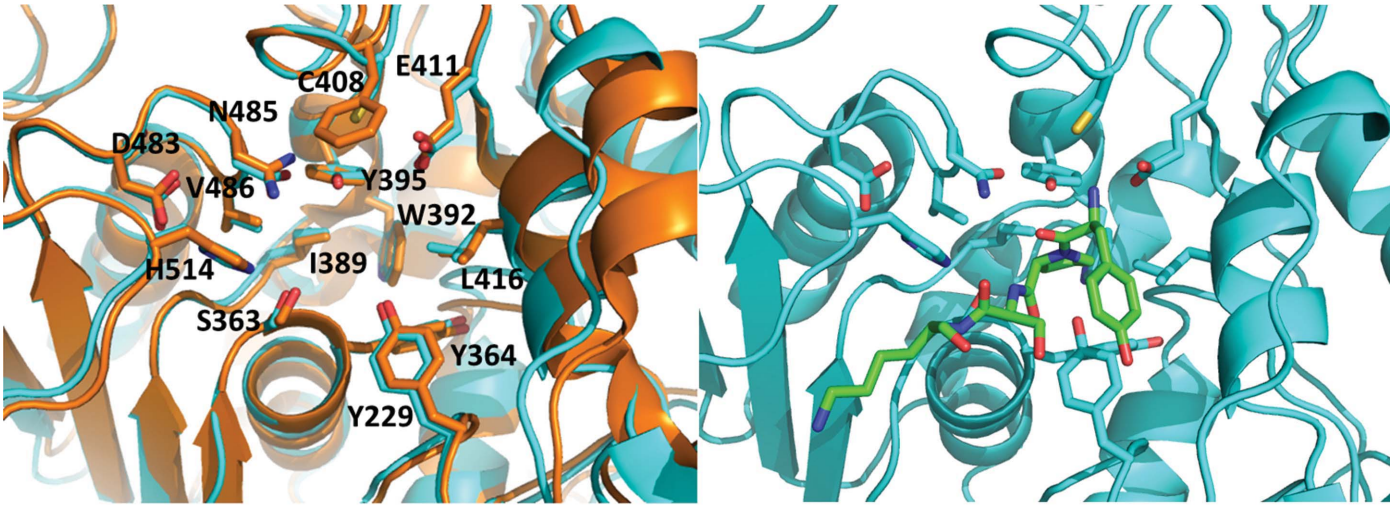

Figure 3
(*a*) Superposition of the active/binding-site residues between *L. helveticus* PepX (PDB entry 6nff, cyan) and PepX from *L. lactis* (PDB entry 1lns, orange). Binding-site residues are conserved, with the exception of Cys408, which is a phenylalanine (Phe393) in PDB entry 1lns. Labels are for the residues in PDB entry 6nff. The alignment was performed in *PyMOL* (Schrödinger) using the *super* utility, which reported an r.m.s.d. of 1.4 Å for 3848 aligned atoms after five cycles of refinement. (*b*) Model of *L. helveticus* PepX (cyan) bound to the YPSK peptide (green). The model was produced by aligning the active/binding sites of PDB entries 1r9n and 6nff in *PyMOL* (Schrödinger) and then transferring the bioactive peptide from the structure with PDB code 1r9n to that with PDB code 6nff. The resulting complex was solvated in an octahedral water box using *CHARMM-GUI* (Jo *et al.*, 2008; Brooks *et al.*, 2009; Lee *et al.*, 2016), and *NAMD* (Phillips *et al.*, 2005) scripts were generated. The model was energy-equilibrated using *NAMD* with the CHARMM36m force field (Huang & MacKerell, 2013) at 310.15 K. The resulting model was obtained after step four of the *CHARMM-GUI*-generated *NAMD* equilibration script, which ran 25 000 energy-minimization steps.

identified several examples of calcium-binding motifs across many different fold families all containing the D*X*(DN)*X*DG consensus sequence in a strand–loop–strand configuration, of which several occurred in β-propeller proteins (Rigden *et al.*, 2011; Rigden & Galperin, 2004). Denesyuk *et al.* (2017) later defined a structural motif called the calcium-blade zone which describes the necessary interactions between the motif and the calcium ion beyond the three aspartic acid residues of the consensus sequence and identified how a calcium-blade zone differs from an EF-hand zone even though they may both contain the D*X*(DN)*X*DG consensus sequence.

The calcium-binding blade-like site identified in PDB entry 6nff matches the consensus sequence of a calcium-blade zone with ten amino acids that form the closed loop: LDTDH-DGKSD (residues 193–202). The calcium is coordinated in a distorted octahedral shape by the side chains of Asp194, Asp196 and Asp198 in the first, second and third coordination positions. The fourth coordination position is occupied by the backbone carbonyl of Lys200 and the fifth position by the side chain of Asp202. The sixth coordination position is occupied by a water molecule that is also within hydrogen-bonding distance of the second O atom in the side chain of Asp198 (2.0 Å) and the backbone carbonyl of Ser201 (2.3 Å). This calcium-blade zone most closely resembles those identified in the seven- and eight-stranded β-propeller folds of integrin α-IIb (PDB entry 3t3p) and rhamnogalacturonan lyase (PDB entry 2zux) (Denesyuk *et al.*, 2017; Zhu *et al.*, 2012; Ochiai *et al.*, 2009). To our knowledge, this is the first calcium-binding protein from the SCOP superfamily and with the α/β-hydrolase fold (SCOPe c.69 and c.69.1) that contains a calcium-blade zone sequence and a calcium-binding blade-like site, increasing the number of identified distinct folds containing this motif from 12 to 13 (Denesyuk *et al.*, 2017).

Although EF-hands and their roles in calcium regulation in eukaryotes have been well documented, calcium-binding proteins (CaBPs) and their roles in bacteria have yet to be fully explored (Michiels *et al.*, 2002; Domínguez *et al.*, 2015). The role of calcium in the *L. helveticus* PepX structure, if one exists, is not known. A possibility is that it exists only to confer stability to the α/β-hydrolase fold, as has been suggested for some other bacterial proteins that contain the D*X*(DN)*X*DG loop sequence (Rigden *et al.*, 2011; Suzuki *et al.*, 2009; Li *et al.*, 2010; Cioci *et al.*, 2006). Alternatively, it could be involved in the regulation of PepX activity. It has previously been documented that the activity of the *L. helveticus* PepX enzyme displays metal dependence, although not necessarily with respect to calcium. The enzyme is still fully active at EDTA concentrations of 1–10 m*M*, but is strongly inhibited in the presence of other divalent metals such as copper, cadmium, zinc, iron and mercury (Vesanto *et al.*, 1995; Stressler *et al.*, 2013). If these metals also interact with the calcium-binding blade-like site, inhibition is likely to result owing to distortion of the active/binding site (calcium is coordinated by residues found in the catalytic domain and is located 27 Å from the catalytic serine). Since PepX activity does not appear to rely on the presence of calcium, it is worth investigating whether calcium confers any structural stability or enables proper folding.

3.4. Activity of PepX Against an immunogenic octamer

The activity of the recombinant construct of *L. helveticus* PepX used for crystallization was tested against a GPpNA

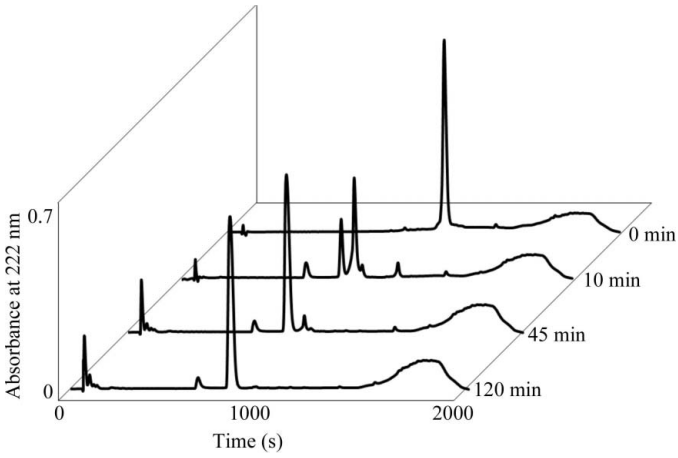

Figure 4
Hydrolysis of a gluten octamer peptide by *L. helveticus* PepX monitored over time by RP-HPLC. The octamer (FPQPQLPY) elutes at 1022 s, the hexamer (QPQLPY) at 820 s and the tetramer (QLPY) at 760 s. The identities of the eluted peptides were confirmed by LC-MS (data not shown).

substrate mimic. The K_m was determined to be between 250 and 310 μ*M* (data not shown), which is in good agreement with the K_m of 240 μ*M* reported by Degraeve & Martial-Gros (2003) for the same substrate. To examine the effectiveness of PepX against gluten-derived peptides with a high proline content, the activity of the enzyme was monitored over time with an octamer peptide substrate (FPQPQLPY) that is found within a celiac disease immunogenic 33-mer α-gliadin peptide (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF) identified as persisting under digestive conditions (Shan *et al.*, 2002). The sequence within the 33-mer was selected to include two *X*P*X*P repeats as well as an *XX*P motif which should not be recognized by PepX. A digest with 100 n*M* PepX and 1 mg ml^{-1} substrate showed nearly complete hydrolysis to liberate the first FP dipeptide within 10 min (Fig. 4). The second QP dipeptide was mostly liberated by 45 min, and after 2 h only the tetramer (QLPY) remained. Although hydrolysis of the second dipeptide appeared to lag behind that of the first dipeptide, this provides good evidence that PepX is capable of hydrolyzing peptides with *X*P repeats, which are frequently found within gluten protein sequences. PepX, when used in combination with one or two additional lactobacilli digestive enzymes, may be sufficient to completely detoxify gluten peptides, as suggested previously (De Angelis *et al.*, 2010; Gallo *et al.*, 2005).

4. Conclusion

The structure of *L. helveticus* PepX provides a second example from the S15 peptidase family of the SC clan. Although the structures are very similar, a single amino-acid difference at the substrate-binding site may explain the differences in substrate specificity that have been observed for *L. helveticus* PepX compared with *L. lactis* PepX. Also, the D*X*(DN)*X*DG calcium-binding site that is present in *L. helveticus* PepX but is absent in *L. lactis* PepX suggests a unique structural or regulatory role for metal binding that needs further investigation. Finally, it is worth noting that many commercially available digestive aids, such as the over-the-counter supplements Gluten-Zyme and GlutenEase, frequently contain probiotic blends that include *L. helveticus* or the DPP-IV enzyme. Although the effectiveness of these digestive aids has not been fully evaluated for the treatment of gluten intolerance and celiac disease, lactobacilli peptidases such as *L. helveticus* PepX continue to be an area of interest for enzyme-detoxification strategies.

5. Related literature

The following references are cited in the supporting information for this article: Ghosh *et al.* (2001), Grant *et al.* (2006), Ohara *et al.* (2014), Pletnev *et al.* (2003) and Zhu *et al.* (2003).

Acknowledgements

The authors would like to thank the many past undergraduate students who made significant contributions to the cloning, expression, purification and crystallization of PepX, especially Stacey Kobes, Andrew Bloom, Naji Saker, Anita (Rachel) Weeks and David Starkovich. Our thanks go to Whitworth student Weston Kroes and instrument specialist Scott Economu at Gonzaga University for analysis of peptides by LC-MS. We would also like to thank Tersa Almaw for helping to prepare the PepX model bound to the YPSK ligand. *NAMD* was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign.

Funding information

This work was funded by the generosity of the Empire Health Foundation and the Whitworth University summer STEM fellowship program, and was supported by the National Science Foundation by award 0723283 to Whitman College.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.

Aertgeerts, K., Ye, S., Tennant, M. G., Kraus, M. L., Rogers, J., Sang, B.-C., Skene, R. J., Webb, D. R. & Prasad, G. S. (2004). *Protein Sci.* **13**, 412–421.

Alcorn, T. & Juers, D. H. (2010). *Acta Cryst.* D**66**, 366–373.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.

Brooks, B. R., Brooks, C. L., MacKerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M. & Karplus, M. (2009). *J. Comput. Chem.* **30**, 1545–1614.

Caputo, I., Lepretti, M., Martucciello, S. & Esposito, C. (2010). *Enzym. Res.* **2010**, 1–9.

Carr, P. D. & Ollis, D. L. (2009). *Protein Pept. Lett.* **16**, 1137–1148.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cioci, G., Mitchell, E. P., Chazalet, V., Debray, H., Oscarson, S., Lahmann, M., Gautier, C., Breton, C., Perez, S. & Imberty, A. (2006). *J. Mol. Biol.* **357**, 1575–1591.

De Angelis, M., Cassone, A., Rizzello, C. G., Gagliardi, F., Minervini, F., Calasso, M., Di Cagno, R., Francavilla, R. & Gobbetti, M. (2010). *Appl. Environ. Microbiol.* **76**, 508–518.

Degraeve, P. & Martial-Gros, A. (2003). *Int. Dairy J.* **13**, 497–507.

Denesyuk, A. I., Permyakov, S. E., Johnson, M. S., Permyakov, E. A. & Denessiouk, K. (2017). *Biochem. Biophys. Res. Commun.* **483**, 958–963.

Dimitriou, P. S., Denesyuk, A., Takahashi, S., Yamashita, S., Johnson, M. S., Nakayama, T. & Denessiouk, K. (2017). *Proteins*, **85**, 1845–1855.

Dimitriou, P. S., Denesyuk, A. I., Nakayama, T., Johnson, M. S. & Denessiouk, K. (2019). *Protein Sci.* **28**, 344–364.

Domínguez, D. C., Guragain, M. & Patrauchan, M. (2015). *Cell Calcium*, **57**, 151–165.

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

Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.

Farley, C., Burks, G., Siegert, T. & Juers, D. H. (2014). *Acta Cryst.* D**70**, 2111–2124.

Forouhar, F., Yang, Y., Kumar, D., Chen, Y., Fridman, E., Park, S. W., Chiang, Y., Acton, T. B., Montelione, G. T., Pichersky, E., Klessig, D. F. & Tong, L. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 1773–1778.

Fox, N. K., Brenner, S. E. & Chandonia, J.-M. (2014). *Nucleic Acids Res.* **42**, D304–D309.

Francavilla, R., De Angelis, M., Rizzello, C. G., Cavallo, N., Dal Bello, F. & Gobbetti, M. (2017). *Appl. Environ. Microbiol.* **83**, e00376-17.

French, S. & Wilson, K. (1978). *Acta Cryst.* A**34**, 517–525.

Gallo, G., Angelis, M. D., McSweeney, P. L. H., Corbo, M. R. & Gobbetti, M. (2005). *Food Chem.* **91**, 535–544.

Gerez, C. L., Font de Valdez, G. & Rollán, G. C. (2008). *Lett. Appl. Microbiol.* **47**, 427–432.

Ghosh, D., Sawicki, M., Lala, P., Erman, M., Pangborn, W., Eyzaguirre, J., Gutiérrez, R., Jörnvall, H. & Thiel, D. J. (2001). *J. Biol. Chem.* **276**, 11159–11166.

Gifford, J. L., Walsh, M. P. & Vogel, H. J. (2007). *Biochem. J.* **405**, 199–221.

Grant, B. J., Rodrigues, A. P. C., ElSawy, K. M., McCammon, J. A. & Caves, L. S. D. (2006). *Bioinformatics*, **22**, 2695–2696.

Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.

Heinig, M. & Frishman, D. (2004). *Nucleic Acids Res.* **32**, W500–W502.

Huang, J. & MacKerell, A. D. (2013). *J. Comput. Chem.* **34**, 2135–2145.

Ileperuma, N. R., Marshall, S. D. G., Squire, C. J., Baker, H. M., Oakeshott, J. G., Russell, R. J., Plummer, K. M., Newcomb, R. D. & Baker, E. N. (2007). *Biochemistry*, **46**, 1851–1859.

Jo, S., Kim, T., Iyer, V. G. & Im, W. (2008). *J. Comput. Chem.* **29**, 1859–1865.

Karplus, P. A. & Diederichs, K. (2015). *Curr. Opin. Struct. Biol.* **34**, 60–68.

Kawasaki, H., Nakayama, S. & Kretsinger, R. H. (1998). *Biometals*, **11**, 277–295.

Kiss, A. L., Hornung, B., Rádi, K., Gengeliczki, Z., Sztáray, B., Juhász, T., Szeltner, Z., Harmat, V. & Polgár, L. (2007). *J. Mol. Biol.* **368**, 509–520.

Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.

Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., MacKerell, A. D., Klauda, J. B. & Im, W. (2016). *J. Chem. Theory Comput.* **12**, 405–413.

Li, N., Yun, P., Nadkarni, M. A., Ghadikolaee, N. B., Nguyen, K.-A., Lee, M., Hunter, N. & Collyer, C. A. (2010). *Mol. Microbiol.* **76**, 861–873.

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.

Michiels, J., Xi, C., Verhaert, J. & Vanderleyden, J. (2002). *Trends Microbiol.* **10**, 87–93.

Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). *J. Mol. Biol.* **247**, 536–540.

Ochiai, A., Itoh, T., Mikami, B., Hashimoto, W. & Murata, K. (2009). *J. Biol. Chem.* **284**, 10181–10189.

Ohara, K., Unno, H., Oshima, Y., Hosoya, M., Fujino, N., Hirooka, K., Takahashi, S., Yamashita, S., Kusunoki, M. & Nakayama, T. (2014). *J. Biol. Chem.* **289**, 24499–24510.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L. & Schulten, K. (2005). *J. Comput. Chem.* **26**, 1781–1802.

Pletnev, V., Addlagatta, A., Wawrzak, Z. & Duax, W. (2003). *Acta Cryst.* D**59**, 50–56.

Rigden, D. J. & Galperin, M. Y. (2004). *J. Mol. Biol.* **343**, 971–984.

Rigden, D. J., Woodhead, D. D., Wong, P. W. H. & Galperin, M. Y. (2011). *PLoS One*, **6**, e21507.

Rigolet, P., Mechin, I., Delage, M.-M. & Chich, J.-F. (2002). *Structure*, **10**, 1383–1394.

Rigolet, P., Xi, X. G., Rety, S. & Chich, J.-F. (2005). *FEBS J.* **272**, 2050–2059.

Schmidt, A., Gruber, K., Kratky, C. & Lamzin, V. S. (2008). *J. Biol. Chem.* **283**, 21827–21836.

Shan, L., Molberg, Ø., Parrot, I., Hausch, F., Filiz, F., Gray, G. M., Sollid, L. M. & Khosla, C. (2002). *Science*, **297**, 2275–2279.

Stressler, T., Eisele, T., Schlayer, M., Lutz-Wahl, S. & Fischer, L. (2013). *PLoS One*, **8**, e70055.

Stressler, T., Pfahler, N., Merz, M., Hubschneider, L., Lutz-Wahl, S., Claassen, W. & Fischer, L. (2016). *Appl. Microbiol. Biotechnol.* **100**, 7499–7515.

Suzuki, R., Katayama, T., Kitaoka, M., Kumagai, H., Wakagi, T., Shoun, H., Ashida, H., Yamamoto, K. & Fushinobu, S. (2009). *J. Biochem.* **146**, 389–398.

Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 38–44.

Vesanto, E., Savijoki, K., Rantanen, T., Steele, J. L. & Palva, A. (1995). *Microbiology*, **141**, 3067–3075.

Wieser, H. & Koehler, P. (2012). *J. AOAC Int.* **95**, 356–363.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* D**67**, 235–242.

Zevaco, C., Monnet, V. & Gripon, J.-C. (1990). *J. Appl. Bacteriol.* **68**, 357–366.

Zhang, Y. & Skolnick, J. (2005). *Nucleic Acids Res.* **33**, 2302–2309.

Zhu, J., Choi, W.-S., McCoy, J. G., Negri, A., Zhu, J., Naini, S., Li, J., Shen, M., Huang, W., Bougie, D., Rasmussen, M., Aster, R., Thomas, C. J., Filizola, M., Springer, T. A. & Coller, B. S. (2012). *Sci. Transl. Med.* **4**, 125ra32.

Zhu, X., Larsen, N. A., Basran, A., Bruce, N. C. & Wilson, I. A. (2003). *J. Biol. Chem.* **278**, 2008–2014.