Two high-resolution crystal structures of potato 1,3-β-glucanase reveal subdomain flexibility with implications for substrate binding

Agnieszka Wojtkowiak[a], Kamil Witek[b], Jacek Hennig[b] and Mariusz Jaskolski[a,c,*]

[a]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznan, Poland; [b]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw and [c]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Correspondence e-mail: mariuszj@amu.edu.pl

PDB References: higher-density crystal structure of potato endo-1,3-β-glucanase, 3ur7; lower-density crystal structure of potato endo-1,3-β-glucanase, 3ur8.

Running title: 1,3-β-Glucanase from potato

Synopsis The 1.40 and 1.26 Å resolution crystal structures of plant endo-1,3-β-glucanase, a member of the glycoside hydrolase family 17, reveal high flexibility of a subdomain that forms part of the active-site cleft, and an unusual crystal packing mode, characterized by infinite chains of protein molecules linked via His-tag docking in the next active site.

For submission to: *Acta Crystallographica* Section D

Abstract Endo-1,3-β-glucanases are widely distributed among bacteria, fungi and higher plants. They are responsible for the hydrolysis of the glycosidic bond in specific polysaccharides with tracts of unsubsituted β-1,3-linked glucosyl residues. The plant enzymes belong to the glycoside hydrolase family 17 (GH17) and are also members of class 2 of pathogenesis-related (PR) proteins. For endo-1,3-β-glucanase from *Solanum tuberosum* (potato, cultivar Désirée), X-ray diffraction data were collected to 1.40 and 1.26 Å resolution for two crystals which, despite a similar packing framework, represent two separate crystal forms. In particular, they differ in the Matthews coefficient and are consequently referred to as higher density (HD, 1.40 Å) and lower density (LD, 1.26 Å) forms. The general fold of the protein resembles that of other known plant endo-1,3-β-glucanases and is defined by a $(\beta/\alpha)_8$ barrel with an additional subdomain built around the C-terminal half of the barrel. The structures reveal high flexibility of the subdomain, which forms part of the catalytic cleft. Comparison with other GH17 endo-1,3-β-glucanase structures reveals differences in the arrangement of the secondary structure elements in this region, which can be correlated with sequence variability and may suggest distinct substrate binding patterns. The crystal structures reveal an unusual packing mode, clearly visible in the LD structure, caused by the presence of the C-terminal His_6-tag, which extends from the compact fold of the enzyme molecule and docks in the catalytic cleft of a neighboring molecule. In this way, an infinite chain of His-tag-linked protein molecules is formed along the **c** direction.

Keywords: glucoside hydrolase; GH17; subdomain; pathogenesis-related protein; His-tag; crystal packing;

1. Introduction

Endo-1,3-β-glucanases (EC 3.2.1.39) are members of the ubiquitous group of glycosidases, i.e. enzymes capable of hydrolysing the glycosidic bond. The cleavage reaction of this specific type of glycosidases is limited to β-1,3-glucosidic linkages present in unbranched segments consisting of several β-1,3-linked glucosyl residues (Witek *et al.*, 2008). The natural substrates of endo-1,3-β-glucanases have a complex structural form of a triple helix, e.g. curdlan (Chuah *et al.*, 1983), or/and are often branched, e.g. callose with (1→6)-β- branching, or/and are composed of mixed glycosidic links, e.g. (1→3), (1→4)-β-glucans. The products of the hydrolysis reaction are (1→3)-β-D-oligoglucosides ranging in length from two to nine glucose moieties, among which the vast majority are tri- and tetrasaccharides (Moore &

Stone, 1972; Keen & Yoshikawa, 1983; Hrmova & Fincher, 1993). Endo-1,3-β-glucanases are implicated in various physiological roles. For instance, in viruses, they are predicted to be involved in the degradation of the host cell wall during virus egress or/and entry (Sun *et al.*, 2001). Endo-1,3-β-glucanases present in Archaea play a role during the fermentation process (Gueguen *et al.*, 1997), while the bacterial enzymes have been shown to have lytic activity against fungi (Fiske *et al*., 1990) and metabolic function (Fuchs *et al*., 2003). In the animal kingdom, endo-1,3-β-glucanases are restricted only to some invertebrates; for example, the enzymes expressed in nematodes allow them to feed on fungi (Kikuchi *et al.*, 2005), while in algae they are involved in the digestion of their storage polysaccharides. (1→3)-β-D-Glucans are a major component of fungal cell wall. The endo-1,3-β-glucanases present in these organisms are involved in cell wall modification during growth, morphogenesis, budding, sporulation and conjugation (Bielecki & Galas, 1991). (1→3)-β-D-Glucans are also components of cell walls in plants, but are restricted in these organisms to more specialized functions. However, plant endo-1,3-β-glucanases play a role in several physiological and developmental processes, e.g. in cell division, microsporogenesis and pollen development, seed germination, and flowering. Plant endo-1,3-β-glucanases are classified as pathogenesis-related class-2 (PR-2) proteins (van Loon *et al*., 1994) because they are expressed in the plant tissue in response to attack by pathogenic microorganisms as well as to wounding or abiotic stress. In particular, they participate in the defense reaction against fungi by their ability to hydrolyze the fungal cell walls. Endo-1,3-β-glucanases have allergenic properties and can be found in pollen grains (Huecas *et al*., 2001). They have been identified among the most allergenic components of natural rubber latex proteins (Sunderasan *et al*., 1995) and as cross-reactive allergens in the latex-fruit syndrome (Wagner *et al*., 2004).

According to amino-acid-sequence-based classification of glycoside hydrolases (GH) (Henrissat, 1991), endo-1,3-β-glucanases are grouped into five families with the following numbers: 16, 17, 55, 64 and 81 [The Carbohydrate-Active EnZymes database (CAZy)-http://www.cazy.org/; Cantarel *et al*., 2009]. So far, crystallographic studies have been presented for all families except GH81. Although proteins from these families act on similar substrates, they have evolutionarily distinct folds. GH16 endo-1,3-β-glucanases are bacterial (Fibriansah *et al*., 2007; Hong *et al*., 2008) and archaeal (Ilari *et al*., 2009) proteins with a β-sandwich jelly-roll folding motif. The plant proteins, on the other hand, are members of family GH17 and exhibit a $(\beta/\alpha)_8$ TIM barrel fold (Varghese *et al*., 1993; Receveur-Brechot *et al*., 2006; Fuentes-Silva, *et al*., unpublished results). Fungal endo-1,3-β-glucanases are

representatives of family GH55 and consist of two domains with a right-handed parallel β-helix fold, forming a rib-cage-like overall shape (Ishida *et al*., 2009). The fold of bacterial enzymes GH64 is distinct from that of bacterial GH16 glycosidases and consists of two domains, a β-barrel domain and a mixed α/β domain (Wu *et al*., 2009).

The GH17 family is classified within the GH-A clan. Clans GH-A through GH-N have been established based on tertiary structure similarity and conservation of the catalytic residues and mechanism. All GH-A clan members possess the $(\beta/\alpha)_8$ barrel fold and two catalytic glutamate residues: a proton donor and a nucleophile, located near the C-terminal ends of β-strands 4 and 7, respectively. The hydrolysis of the glycosidic bond catalyzed by GH-A glycosidases is characterized by retention of the stereochemistry of the anomeric carbon at the cleavage point (Jenkins *et al.*, 1995). Due to the characteristic location of the catalytic residues, the GH-A clan is also referred to as the 4/7 superfamily. The hydrolysis reaction proceeds through a double-displacement mechanism. The nucleophile and proton-donor carboxylic groups are located on opposite sides of the hydrolyzed glycosidic bond and are separated by a distance of approximately 5.5 Å, with the proton donor situated within hydrogen-bonding distance of the glycosidic oxygen. After protonation of the glycosidic oxygen by the proton donor, the nucleophile attacks the sugar ring from the opposite side relative to the leaving group to form a covalent glycosyl-enzyme intermediate, which is subsequently hydrolyzed by a water molecule in the next step of the reaction.

The structures of three plant endo-1,3-β-glucanases have been reported, for *Hordeum vulgare* (barley) at 2.2 Å (PDB code 1GHS; Varghese *et al.*, 1994), *Musa acuminata* (banana) at 1.45 Å (2CYG; Receveur-Brechot *et al*., 2006), and *Hevea brasiliensis* (rubber tree) at 2.5 Å (3EM5, 3F55; Fuentes-Silva *et al*., unpublished results). They share similar overall fold and active-site topology with endo-1,3-1,4-β-glucanase from *H. vulgare* (1GHR;Varghese *et al.*, 1994; 1AQ0; Müller *et al.*, 1998), which represents the same GH17 family. The 3D structures of members of this family exhibit the characteristic TIM-barrel fold, defined by eight parallel β-strands in the interior of the protein, surrounded by a ring of helices. Typically, there is a single α-helix crossover between each pair of adjacent β-strands. The endo-1,3-β-glucanases from *H. vulgare*, *M. acuminata* and *H. brasiliensis* have additional structural elements in the C-terminal half of the barrel. They include two pairs of short antiparallel β-strands which, together with additional neighboring short helices and loops, form a small subdomain built around α-helix 6. A deep catalytic cleft, approximately 40 Å long, runs along the upper part of the entire molecule. The length of the cleft suggests that it can accommodate up to eight

glucosyl residues of the (1→3)-β-D-glucan substrate, as confirmed by kinetic and thermodynamic studies (Hrmova *et al.*, 1995).

Endo-1,3-β-glucanase from *S. tuberosum* (potato) is synthesized as a 338-residue precursor protein (NCBI GenBank accession No. AJ586575, UniProt ID Q70C53) with a 23-residue signal peptide at the N-terminus. Secretion of the enzyme to the extracellular space is connected with a removal of this peptide. The enzyme functions as a monomer and its kinetic parameters have been determined by Witek *et al.* (2008). The present paper describes two crystal structures of recombinant mature endo-1,3-β-glucanase from *S. tuberosum* with an additional His-tag octapeptide (-LNHHHHHH) at the C-terminus. A survey of the impact of His-tag on the structure of the tagged proteins shows that it has very little effect on refinement statistics and no significant effect on the structure of the native protein (Carson *et al.*, 2007). Here, we demonstrate that the presence of the His-tag facilitates crystallization and has no apparent influence on the overall fold of the protein. The tag residues form intermolecular contact between monomers within the asymmetric unit and are involved in crystal packing contacts. Moreover, the histidine tag protrudes into the catalytic cleft and interacts with highly conserved residues in the substrate-binding regions, mimicking substrate recognition. The paper compares the structure of endo-1,3-β-glucanase from *S. tuberosum* with the previously reported structures of GH17 proteins from plants and discusses the consequences of the observed differences in the subdomain structure, a region that is postulated to take part in substrate binding.

2. Materials and methods

2.1. Cloning and expression

cDNA coding for the endo-1,3-β-glucanase from *S. tuberosum* was amplified by PCR using the *gluB20-2* ORF cloned into pTOPO vector [pTOPOgluB(20-2)ORF; Barabasz, 2005] as a template. The forward (**CATATG**CAGCCTATCGGAGTATGCTAT) and the reverse **(CTCGAG**ATTAAAATTGAGTTGATACTT) primers introduced, respectively, NdeI and XhoI restriction sites (bold). The PCR product was cloned into pGEM-T Easy vector (Promega). Following nucleotide sequence confirmation, the gene was subcloned into pET-30a(+) vector (Novagen), which added a hexahistidine tag at the C-terminal end of the expressed protein. The histidine tag consists of eight residues, with the following sequence -LNHHHHHH.

The recombinant protein was expressed in *E. coli* BL21 strain (Studier & Moffatt, 1986), using TB medium (Sambrook *et al*., 1989). The bacterial culture was incubated at 310 K to A_{600}=1. Protein overexpression was induced with the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.25 mM, and incubation was continued at 291 K for 24 h. Finally, the culture was centrifuged at 8,000×g at room temperature for 10 min. The bacterial cells were collected and kept overnight at 253 K.

2.2. Purification

The bacterial pellet was suspended and lysed in 50 ml of buffer Z [50 mM sodium phosphate buffer, pH 7.8, 1 mM PMSF, 10% (w/v) glycerol] per 1 l of the source culture. The suspension was supplemented with $MgSO_4$ to a final concentration of 10 mM and centrifuged at 23,500×g, 277 K, for 30 min. The solution was saturated with $(NH_4)_2SO_4$ to 80% and left overnight at 277 K with continuous stirring. The pellet was dialyzed against three changes of buffer Z with 10× sample volume. After centrifugation, the soluble recombinant protein was initially purified on a DEAE-cellulose column. The column was washed with buffer Z, and the column flow-through was supplemented with NaCl and imidazole to the final concentrations of 300 mM and 20 mM, respectively. Finally, the protein suspension was applied to an Ni-NTA agarose column. The column was washed with washing buffer (50 mM sodium phosphate, pH 7.8, 300 mM NaCl, 20 mM imidazole) and the purified protein was eluted from the column with elution buffer (50 mM sodium phosphate, pH 7.8, 300 mM NaCl, 200 mM imidazole). The homogenous protein fractions were pooled together, dialyzed against 25 mM sodium phosphate buffer, pH 7.8, with addition of 10% (w/v) glycerol, and kept as 1 mg aliquots at 253 K.

For crystallization experiments, the protein was concentrated to 8 $mg.ml^{-1}$ and the buffer was exchanged to 20 mM Tris-HCl, pH 8.0, using Millipore Centricon 10 filters.

2.3. Crystallization

Crystals were grown by vapor diffusion at 292 K in hanging drops mixed from 1.5 μl protein solution and 1.5 μl reservoir solution. The starting condition was obtained from Structure Screen 1 (Molecular Dimensions, Ltd) with the reservoir solution containing 0.1 M sodium acetate, pH 4.6, 0.2 M ammonium acetate, 30% PEG 4000. The crystallization experiments suffered from nucleation problems, leading to amorphous precipitates and only sporadic measurable crystals, which were often twinned. These problems were overcome by lowering

the PEG 4000 concentration to 25% and using streak seeding. Crystals of two different forms denoted higher density (HD) and lower density (LD) appeared after two days, often in the same drop. There was no correlation between crystallization conditions, crystal morphology, cryoprotection or crystal handling, and the appearance of the specific crystal form. Differentiation between the two crystal forms was based on the results of X-ray diffraction data processing.

2.4. Data collection and processing

For data collection, 1:4 mixture of PEG 400 and the reservoir solution was used as a cryoprotectant. X-Ray diffraction data for the HD crystals were collected at MAX-lab in Lund, using beamline I911-2, in two passes: a medium-resolution pass (30-2.15 Å, oscillation 1.2º), and a high-resolution pass (30-1.40 Å, oscillation 0.75º). Data for the LD crystals were collected at EMBL Hamburg at beamline X11 in two passes, at medium resolution (40-1.85 Å, oscillation 1º), and at high resolution (40-1.26 Å, oscillation 0.5º). Both data sets were indexed, integrated and scaled with *HKL2000* (Otwinowski & Minor, 1997). In both cases the space group is $P2_1$. The asymmetric units of both crystal forms contain two protein molecules with Matthews coefficients (Matthews, 1968) of 1.96 and 2.03 $Å^3\ Da^{-1}$ for the HD and LD crystals, respectively. A summary of the data collection and processing statistics is given in Table 1.

2.5. Structure determination and refinement

The structure of the HD crystal was determined by molecular replacement with the *MOLREP* program (Vagin & Teplyakov, 1997), using directly the coordinates of endo-1,3-β-glucanase from *H. vulgare* (PDB code 1GHS, molecule A) as a search model. The solution was characterized by a correlation coefficient of 0.406 and an R-factor of 0.544. Structural refinement was performed using *REFMAC5* (Murshudov *et al*., 1997). After each refinement step, the *XtalView* program (McRee, 1999) was used for viewing electron-density maps and manual rebuilding of the model. The refined model at 1.40 Å resolution was subsequently used for the determination of the structure of the LD crystal. Structure determination and refinement was carried out as for the HD crystal. Anisotropic modeling of the atomic displacement parameters was used in each case, permitted by the high resolution of the diffraction data (1.40 Å and 1.26 Å, respectively). In both structures, there are two protein molecules in the asymmetric unit, labeled A and B. A summary of the refinement statistics is given in Table 1. For most calculations, the *CCP*4 suite of programs was used (Collaborative

Computational Project, Number 4, 1994). The structures were validated using *MolProbity* (Chen *et al.*, 2010). Molecular and electron density illustrations were prepared in *PyMol* (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC).

3. Results and discussion

3.1. Model quality and structure overview

Mature endo-1,3-β-glucanase from *S. tuberosum* consists of 315 amino acid residues (residues 24-338). The recombinant protein used in this study has an additional eight residues at the C-terminus, encoding a histidine tag (residues 339-346), which is comprised of two linker residues (LeuAsn) and six histidines. The model of the HD (higher density) crystal structure contains all the residues of the protein sequence (24-338) and one residue, His344, from the affinity tag (in both molecules). Because of lack of contiguous electron density connecting the protein molecules and the affinity-tag residues, the numbers of the His residues were assigned by analogy to the LD (lower density) crystal structure. In the LD structure, five disordered residues (Gln223-Asp227) of molecule B could not be modeled because of poor electron density. The histidine tag residues are present in both molecules, with very clear contiguous electron density extending from the protein C-terminus, except for the last two histidine residues in each molecule (residue numbers 345-346). The final models have very good overall geometry (Table 1), and the Ramachandran plot statistics (Ramachandran *et al*., 1963) indicate over 98% of the main-chain dihedral angles in the most favored regions with no residues in disallowed regions (Chen *et al.*, 2010).

As in other $(\beta/\alpha)_8$ (or TIM) barrels, the overall fold of the protein (Fig. 1) consists of eight parallel β strands β1-β8 (for the naming convention see Varghese *et al.*, 1994) forming the interior of the structure, with connections provided by eight external helices α1-α8 and loops. Ideally, between each pair of adjacent β-strands in a TIM barrel there is a single α-helix. In the present case, helix α8 is reduced to a short 3_{10} helix (Fig. 1). Another distinguishing feature of the present fold is the existence of two helices between strands β3 and β4, with an extra 3_{10} helix A3 present in addition to the typical α-helix, α3. Moreover, there are other arrangements of the secondary structure elements, built around the C-terminal half of the barrel, that are characteristic for this protein. Two short antiparallel β-strands (B5a and B5b) are located in the β5-α5 loop, and two short α-helices (A6a and A6b) are located in loop β6-α6. These extra structural elements, together with the neighboring loops, create a subdomain that is situated around helix α6. Helix α6 is perpendicular to the β-strands and other α-helices

defining the barrel fold. The N- and C-termini of the molecule (disregarding the affinity tag) are separated by a distance of about 5.5 Å and lie at the bottom surface of the barrel.

The protein possesses an elongated ellipsoidal shape (Fig. 2) with overall dimensions of ~32×40×50 Å. Parallel to the longest axis of the ellipsoid, a catalytic cleft approximately 40 Å long runs along the upper surface of the molecule. This shape of the catalytic cleft is typical for endo-glycosidases and allows for binding of several sugar units (Davies *et al*., 1995). The cleft runs toward the β2-α2 loop and strand β6, and is extended beyond β6 by the presence of the transverse helix α6 and the A6b-α6 loop, which is part of the subdomain (Fig. 2). The side walls of the catalytic cleft are formed by loops connecting the β-strands with the helices, and by the helices themselves.

3.2. Two crystal forms

The two crystal forms have the same space group ($P2_1$) and similar lattice parameters (Table 1). However, while the **b** axis is identical, the **a** and **c** parameters are systematically longer for the LD form (by 1.8 and 2.6%, respectively). The monoclinic angle is wider by 1.2° in the LD form. The R_{merge} value for scaling the two data sets together is as high as 0.490. In both crystals, the asymmetric unit contains two protein molecules in quite similar packing arrangement. Although in some cases a clear differentiation between protein crystal polymorphs is problematic and there may be a "continuum of polymorphic modifications" (Michalska *et al*., 2008), in the present case there is no doubt about the existence of two distinct crystal forms for the following reasons. (i) The aggregated change of the unit cell volume is significantly higher than the experimental error; (ii) the two diffraction data sets cannot be scaled together; (iii) the protein molecules have visibly different orientations with respect to the crystallographic directions and with respect to each other; (iv) there is a visible change in the translation of the molecules along their main packing direction (**c**); (v) the intermolecular interactions leading to this packing arrangement, i.e. docking of the His-tag tail in the active-site cleft of the adjacent molecule, have perfect definition in the electron density in one of the crystal forms (LD) but not in the other; (vi) there is a visible conformational change in one of the structural elements (loop A6b-α6) upon transition from form LD to HD; (vii) the lattice contacts in the two structures are not the same. The structural aspects of these arguments (iii-vii) will be discussed in the following sections.

3.3. The asymmetric unit and impact of the His-tag on crystal packing

The crystallographic asymmetric units of the HD and LD crystals contain two protein molecules, A and B, in each case. The choice of the unit-cell origin is consistent. In both structures, the monomers are arranged in a similar manner and are related by a translation of about 40.5 Å and 40.2 Å in the HD crystal structure, and 41.6 Å and 41.1 Å in the LD crystal structure, approximately along the **c** direction with a concomitant rotation of about 1.2° and 5.7°, respectively. Contacts between the monomers within the asymmetric unit (disregarding the affinity tag) of the HD crystal structure are formed by loops α2-β3, α3-β4, and the N-terminal fragment of β4 from molecule A, and by loop β4-α4 and the N-terminal fragment of helix α4 from molecule B. These interactions are preserved in the case of the LD crystal structure, with the exception of loop α3-β4 from molecule A. Almost identical contacts in two different crystal forms can be surprising, but are correlated with the flexibility of this part of the protein (see Structural comparisons). For the purpose of comparisons, models of the HD and LD structures were superposed using only the contact residues as targets, to illustrate that although these regions superpose well and interactions are preserved in both structures, the orientations of the entire molecules are different (Fig. 3).

Direct contacts between molecules A and B within the asymmetric unit in both structures are also formed by the His-tag peptides. The histidine tag introduced at the C-terminus of the protein extends from the bottom side of the barrel (relative to the upper side where the catalytic cleft is located) of molecule A into the catalytic cleft on the upper side of molecule B. Conversely, the His-tag of molecule B packs into the catalytic cleft of a **c**-translated copy of molecule A (Fig. 4). The interactions of the His-tags in the catalytic clefts are different for molecules A and B, and they also differ between the two structures. The His-tags in the LD crystal structure are much better defined in electron density, with six residues (Leu339-His344) modeled in each molecule with low B-factors, than in the HD crystal, where only one residue (His344) is visible in each molecule. In both structures, the side chain of His344 forms the same intermolecular hydrogen bonds with Tyr201 and Glu319 of the complementary protein molecule (Fig. 5). The missing residues of the His-tag in the HD structure cannot be modeled by analogy to the LD structure not only because of very poor electron density, but also because of the altered orientation and distance between molecules A and B in the two structures.

The impact of His-tag peptides on protein crystal structures has been investigated by Carson *et al*. (2007) who showed that almost all of the resolved tags are involved in crystal packing contacts. Insertion of a histidine-tag in the catalytic cleft of an enzyme has been observed previously for a glycoside hydrolase - lichenase (Taylor *et al*., 2005). This mode of

interaction reflects the electrostatic affinity between the negatively charged active site of the enzyme (Glu residues) and the positively chareged His residues, a situation that is favored in mildly acidic buffers. The present case, however, is unusual as the protein molecules are placed in the crystal structures on top of each other, forming an infinite, straight chain of molecules linked by their histidine tags (Fig. 4).

3.4. Crystal contacts

Analysis of the molecular surface area buried on crystal packing (Table S1) shows, as expected, that the total solvent-accessible area is larger for the LD crystal form molecules. Omitting the His-tag residues from the models results in an increase of this area in both crystal forms. It is interesting to note that there is a disparity between the contributions of molecules A and B to the lattice contacts in crystal forms HD and LD. The disparity arises from differences in packing arrangement and remains even after the elimination of potential structural reasons, i.e. removal of the His-tag residues and of the Gln223-Asp227 fragment (part of loop A6b-α6) from molecule B of the HD crystal form, which is absent in the LD model. In all other molecules, loop A6b-α6, which forms part of the subdomain, participates in extensive lattice contacts. Mapping the residues involved in intermolecular interactions onto the molecular surface of the protein (Fig. 6) illustrates the disparity of the different molecules in lattice contacts. In the above calculations, a contact between a pair of atoms sitting in different asymmetric units was detected only if the distance between their van der Waals spheres is less than 0.5 Å (Vriend, 1990). Additionally, contacts between pairs of atoms from different monomers in the same asymmetric unit are presented.

3.5. Structural comparisons

A superposition of all the protein chains (except the His-tag residues) in all pairwise combinations shows that the secondary structures of the monomers are practically identical. The r.m.s.d. value calculated in *ALIGN* (Cohen, 1997) for 297 C^{α} pairs of the HD monomers is 0.19 Å and 0.45 Å for 301 C^{α} pairs of the LD structure. The r.m.s.d. values indicate that the protein molecules in the HD crystal are less divergent than in the LD structure. Moreover, molecule A from the LD crystal has a better superposition with both molecules of the HD crystal than with molecule B from the same structure (Fig. 7, Table 2). The highest atomic deviations are observed in the loop regions (Fig. 8), especially for one of the loops forming the subdomain, namely loop A6b-α6 (Phe220-Asn232), where the deviations exceed 5 Å for the C^{α} atoms of Arg224. The mobility of this loop is also indicated by its partial disorder in

molecule B of the LD crystal. The conformational changes of this loop are in correlation with crystal contacts. The other components of the subdomain, namely helices A6a and A6b, and loops A6a-A6b and B5a-B5b, also display visible flexibility. Another part of the protein with apparent flexibility is loop β4-α4 together with the N-terminal fragment of helix α4. However, both molecules from the HD crystal and molecule A from the LD crystal superpose very well in this region, while for molecule B(LD) the C^{α} deviations at Glu124 are as high as 1.7 Å. Interestingly, this region is responsible for molecular contacts between the monomers within the asymmetric unit.

So far, four structures of endo-1,3-β-glucanase have been deposited in the PDB (Berman *et al*., 2000) for the following plants: *H. vulgare* (1GHS), *M. acuminata* (2CYG) and *H. brasiliensis* (3EM5, 3F55). Structural comparisons of the present models of endo-1,3-β-glucanase from *S. tuberosum* with those structures show that the fold is essentially the same in all cases (Fig. 9). The C^{α} r.m.s.d. values for superpositions using molecule A of the LD crystal as the target are about 0.8 Å in all cases. A detailed analysis shows that the core β-sheet of the barrel is highly conserved. The main differences are found in the loops and in the helical regions forming the outer shell of the protein, and are correlated with sequence insertions and deletions (Fig. 10). The differences in the amino acid sequence within loops β4-α4 and β5-α5 result in shifts of helices α4 and α5, respectively. The loop within the subdomain, located between helices A6b and α6, is shorter in the *S. tuberosum* protein and the antiparallel β-strand present in this region in the enzymes from *H. vulgare*, *M. acuminata* and *H. brasiliensis* is absent in the potato protein altogether. The level of sequence identity between the present enzyme from *S. tuberosum* and those from *H. vulgare*, *M. acuminata* and *H. brasiliensis* is 47%, 50% and 55%, respectively.

3.6. The catalytic cleft and the active site

The active site of plant endo-1,3-β-glucanases is located in a 8-9 Å deep cleft running along the upper part of the molecule. Analogously to all other members of the GH-A clan, two strictly conserved glutamate residues located near the C-terminal ends of the β strands 4 and 7 are predicted to act as the proton donor and nucleophile, respectively (Jenkins *et al.*, 1995). In the enzyme from *S. tuberosum*, these residues correspond to Glu118 (proton donor) and Glu259 (nucleophile). Both catalytic residues are situated in the canyon, about one-third of the distance from the β2-α2 loop to the opposite end of the cleft formed by the subdomain. Inside the catalytic cleft, there are a number of aromatic residues (Tyr58, Tyr201, Phe204, Phe305, Phe322) showing strict conservation in all plant endo-1,3-β-glucanases, which may

be involved in stacking interactions with the rings of the glucosyl residues of the substrate (Varghese *et al.*, 1994). The importance of the strictly conserved hydrophilic residues Glu259, Glu310, Lys313 and Glu319 has been investigated in site-directed mutagenesis studies (Chen *et al.*, 1995). Substitution of each of these residues resulted in a reduction of the enzymatic activity.

In all four protein molecules of the two structures presented in this study, histidine residue 344 from the His-tag forms hydrogen bonds with Glu319 and Tyr201. The interactions of the histidine tag within the catalytic cleft may provide a hint about the substrate binding mode and about the residues involved (Fig. 5), although it is obvious that binding of an oligohistidine peptide by an oligosaccharide-processing enzyme may not reflect all of the specific interactions responsible for substrate recognition. Therefore, a crystal structure of an enzyme-oligosaccharide complex will be necessary to verify the speculations.

The length of the catalytic cleft in all known plant endo-1,3-β-glucanase structures is approximately the same, 40 Å, and is compatible with accommodation of up to eight glucosyl residues of a (1→3)-β-D-glucan substrate. Kinetic and thermodynamic studies together with product analysis of oligosaccharide hydrolysis indicate the existence of eight subsites of the binding cleft, numbered –3, –2, –1, +1, +2, +3, +4, +5, with the scissile bond located between units –1 and +1 (for subsite nomenclature see Davies *et al.*, 1997), and the oligosaccharide substrate orientated with its non-reducing end (–3) over the β2-α2 loop, and the reducing end (+5) over the subdomain (Hrmova *et al.*, 1995). The binding energies calculated at the individual subsites have the highest values at subsites –2, +4 and +5. Based on the location of the catalytic residues and the length of an octasaccharide, it has been suggested for the *H. vulgare* enzyme that subsite +5 is located over the antiparallel β-strand present in the A6b-α6 loop, which is part of the subdomain (Chen *et al.*, 1995; Hrmova *et al.*, 1995). The high mobility of this loop in the present structure, its partial disorder (Fig. 8), and the absence of an antiparallel β-strand in the case of *S. tuberosum* endo-1,3-β-glucanase (Fig. 9), may all indicate an alternative pattern of oligosaccharide binding affinities, different from that used by other plant enzymes in this family.

4. Conclusions

We have presented two crystal structures (HD and LD) of endo-1,3-β-glucanase from *S. tuberosum* (potato, cultivar Désirée) determined to resolutions of 1.40 and 1.26 Å. The enzyme has the TIM-barrel $(\beta/\alpha)_8$ folding pattern, also found in three other plant endo-1,3-β-

glucanases of the glycoside hydrolase family GH17. Differences between these structures are found mostly in the loops and in the helical regions forming the outer shell of the protein fold, and in the extra structural elements forming its additional subdomain. In the present endo-1,3-β-glucanase from *S. tuberosum*, the subdomain is composed of (i) two short additional helices A6a and A6b located in the loop between β6 and α6, (ii) two short antiparallel β-strands B5a and B5b located in the loop between β5 and α5, and (iii) additional loops situated around helix α6. The catalytic cleft, which runs along the upper part of the entire molecule over the core β-sheet together with the subdomain, has been suggested as the binding site for a long, octameric segment of the substrate (1→3)-β-D-glucan polysaccharides. The findings of the present study, demonstrating that the A6b-α6 loop that forms part of the subdomain, is characterized by high mobility and lacks the antiparallel β-strand present in other structures of similar plant enzymes, may indicate that the substrate binding pattern is diffent in various plant endo-1,3-β-glucanases. The tight packing of the protein molecules in the two crystal structures [Matthews coefficients 1.96 $Å^3.Da^{-1}$ (HD crystal) and 2.03 $Å^3.Da^{-1}$ (LD crystal)] is a consequence of the presence of the histidine tag attached to the C-terminus of the protein, which protrudes from one molecule and docks in the catalytic cleft of an adjacent molecule.

Acknowledgement

This work was supported in part by the Polish Node of the Potato Genome Sequencing Consortium contract no. 47/PGS/2006/01. Some of the calculations were carried out in the Poznan Metropolitan Supercomputing and Networking Center.

References

Barabasz, A. (2005). PhD Thesis, Institute of Biochemistry and Biophysics, Pol. Acad. Sci., Warsaw.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000). *Nucleic Acids Res.* **28**, 235-242.

Bielecki, S. & Galas, E. (1991). *Crit. Rev. Biotechnol.* **10,** 275-304.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233-D238.

Carson, M., Johnson, D. H., McDonald, H., Brouillette, C. & DeLucas, L. J. (2007). *Acta Cryst.* D**63**, 295-301.

Chen, L., Garrett, P.T. J., Fincher, G. B. & Høj, P. B. (1995). *J. Biol. Chem.* **270,** 8093-8101.

Chen, V.B., Arendall III, 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.

Chuah, C.T., Sarko, A., Deslandes, Y. & Marchessault, R.H. (1983). *Macromolecules* **16**, 1375-1382.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760-763.

Cohen, G. R. (1997). *J. Appl. Cryst.* **30**, 1160-1161.

Davies, G. J. & Henrissat B. (1995). *Structure* **3**, 853-859.

Davies, G. J., Wilson, K. S. & Henrissat, B. (1997). *Biochem. J.* **321**, 557-559.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392-400.

Fibriansah, G., Masuda, S., Naoya, K., Nakamura, S. & Kumasaka, T. (2007). *Proteins* **69**, 683-690.

Fiske, M. J., Tobey-Fincher, K. L. & Fuchs, R. L. (1990). *J. Gen. Microbiol.* **136,** 2377-2383.

Fuchs, K. P., Zverlov, V. V., Velikodvorskaya, G. A., Lottspeich, F. & Schwarz, W. H. (2005). *Microbiology* **149,** 1021-1031.

Gille, C. & Frömmel, C. (2001). *Bioinformatics* **17**, 377-378.

Gueguen, Y., Voorhorst, W. G. B., van der Oost, J. & de Vos, W. M. (1997). *J. Biol. Chem.* **272,** 31258-31264.

Henrissat, B. (1991). *Biochem. J.* **280**, 309-316.

Hong, T.-Y., Hsiao, Y.-Y., Meng, M. & Li, T. T. (2008). *Acta Cryst.* D**64**, 964-970.

Hrmova, M., & Fincher, G.B. (1993). *Biochem. J.* **289**, 453–461.

Hrmova, M., Garrett, T. P. J. & Fincher, G. B. (1995). *J. Biol. Chem.* **270**, 14556-14563.

Huecas, S., Villalba, M. & Rodriguez, R. (2001). *J. Biol. Chem.* **276**, 27959-27966.

Ilari, A., Fiorillo, A., Angelaccio, S., Flori, R., Chiaraluce, R., van der Oost, J. & Consalvi, V. (2009). *FEBS J.* **276**, 1048-1058.

Ishida, T., Fushinobu, S., Kawai, R., Kitaoka, M., Igarashi, K. & Samejima, M. (2009). *J. Biol. Chem.* **284**, 10100-10109.

Jenkins, J., Leggio, L. L., Harris, G. & Pickersgill, R. (1995). *FEBS Lett.* **362**, 281-285.

Jones, T. A. & Kleywegt, G. J. (1999). *Proteins: Struct. Funct. Genet. Suppl.* **3**, 30-46.

Keen, N. T. & Yoshikawa, M. (1983). *Plant Physiol*. **71**, 460-465.

Kikuchi, T., Shibuya, H. & Jones, J. T. (2005). *Biochem. J.* **389**, 117-125.

van Loon, L. C., Pierpoint, W. S., Boller, T. & Conejero, V. (1994). *Plant Mol. Biol. Rep.* **12**, 245-264.

Matthews, B. W. (1968). *J. Mol. Biol.* **33,** 491-497.

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156-165.

Michalska, K., Borek, D., Hernandez-Santoyo, A. & Jaskolski, M. (2008). *Acta Cryst.* D**64**, 309-320.

Moore, A. E. & Stone, B. A. (1972). *Biochim Biophys Acta.* **258**, 248-264.

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

Müller, J. J., Thomsen, K. K. & Heinemann, U. (1998). *J. Biol. Chem.* **273**, 3438-3446.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 22756-22763.

Painter, J. & Merritt, E. A. (2006). *J. Appl. Cryst.* **39**, 109-111.

Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95-99.

Receveur-Bréchot, V., Czjzek, M., Barre, A., Roussel, A., Peumans, W. J., Van Damme, E. J. M. & Rogué, P. (2006). *Proteins* **63**, 235-242.

Sambrook, J.E., Fritsch, F. & T. Maniatis (1989). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

Studier, F.W. & Moffatt, B.A. (1986). *J. Mol. Biol.* **189**, 113-130.

Sun, L., Gurnon, J. R., Adams, B. J., Graves, M. V. & Van Etten, J. L. (2001). *Virology* **276**, 27-36.

Sunderasan, E., Hamzah, S., Hamid, S., Ward, M. A., Yeang, H. Y. & Cardosa, M. J. (1995). *J. Nat. Rubb. Res.* **10**, 82-99.

Taylor, E. J., Goyal, A., Guerreiro, C. I., Prates, J. A., Money, V. A., Ferry, N., Morland, C., Planas, A., Macdonald, J. A., Stick, R. V., Gilbert, H. J., Fontes, C. M. & Davies, G. J. (2005). *J. Biol. Chem.* **280**, 32761-32767.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022-1025.

Varghese, J. N., Garrett, T. P. J., Colman, P. M., Chen, L., Høj, P. B. & Fincher, G. B. (1994). *Proc. Natl. Acad. Sci. USA* **91**, 2785-2789.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52-56.

Wagner, S., Radauer, C., Hafner, C., Fuchs, H., Jensen-Jarolim, E., Wüthrich, B., Scheiner, O. & Breiteneder, H. (2004). *Clin. Exp. Allergy* **34,** 1739-1746.

Witek, A. I., Witek, K. & Hennig, J. (2008). *Acta Biochim Pol.* **55**, 791-797.

Wu, H. M., Liu, S.W., Hsu, M. T., Hung, C. L., Lai, C. C., Cheng, W. C., Wang, H. J., Li, Y. K. & Wang, W. C. (2009). *J. Biol. Chem.* **284,** 26708-26715.

Figure captions

Fig. 1. Topology diagram of endo-1,3-β-glucanase from *S. tuberosum*. Color code: red – β-strand, green – α-helix, blue – 3_{10} helix.

Fig. 2. Overall fold of endo-1,3-β-glucanase from *S. tuberosum* represented by molecule A of the lower density (LD) crystal structure. The strands of the inner β-barrel (red) are surrounded by α-helices (green), and by additional helices and a β-sheet form the subdomain (blue). (A) Top view down the TIM-barrel axis. (B) Side view of the molecule, with the subdomain facing the viewer.

Fig. 3. Superposition of the asymmetric units of the HD (molecule A – blue, molecule B – green) and LD (molecule A – red, molecule B – orange) crystals, using only the contact residues (shown in detail on the right, and indicated by an arrow) as the superposition target.

Fig. 4. Crystal packing of endo-1,3-β-glucanase from *S. tuberosum* in the LD crystal. Molecule A is marked in blue, molecule B in green.

Fig. 5. The histidine tag residues (339-344) from molecule A of the LD crystal structure docked in the catalytic cleft of molecule B. The C-terminal tail of molecule A is shown in F_o-F_c *OMIT* map (green) contoured at 2.5σ. Selected side chains in the catalytic cleft of molecule B are shown in $2F_o$-F_c electron density contoured at the 1.0σ level.

Fig. 6. Intermolecular contacts mapped onto the surface of endo-1,3-β-glucanase from *S. tuberosum*. For each protein molecule, viewed down the TIM-barrel axis, two views are shown, with a rotation of 180° around the vertical axis. Green color represents residues participating in protein-protein interactions within the asymmetric unit, orange color represents residues participating in protein-protein crystal contacts (disregarding histidine tag and residues interacting with histidine tag in both cases). Magenta color marks residues interacting with the histidine tag, and blue color marks the histidine tag residues. The left view is for the molecule facing the viewer with the catalytic canyon.

Fig. 7. Superposition of the present four models of endo-1,3-β-glucanase from *S. tuberosum*. Color code: higher density (HD) crystal structure molecule A – blue, molecule B – green; lower density (LD) crystal structure molecule A – red, molecule B – orange. The arrow indicates loop A6b-α6.

Fig. 8. Superposition of endo-1,3-β-glucanases from *S. tuberosum* (represented by molecule A of the LD crystal structure) (orange), *M. acuminata* (yellow), *H. vulgare* (green) and *H. brasiliensis* (blue).

Fig. 9. Structural sequence alignment (calculated with the program *STRAP*; Gille & Frömmel, 2001) comparing endo-1,3-β-glucanases from *S. tuberosum, H. brasiliensis, M. acuminata* and *H. vulgare*. The secondary-structure elements above the *S. tuberosum* sequence are marked as α-helices (green), 3_{10}-helices (blue) and β-strands (red); the subdomain β-strands not present in the structure of the *S. tuberosum* protein are indicated by transparent arrows. Residues forming the secondary structures are highlighted in corresponding colors. The catalytic residues are marked with an empty (proton donor) and filled (nucleophile) star.

Table 1
Data collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	HD crystal form	**LD crystal form**
Data collection statistics		
Radiation source	I911-2, MAX-lab, Lund	X11, EMBL Hamburg
Wavelength [Å]	1.0430	0.8148
Temperature of measurements [K]	100	100
Space group	$P2_1$	$P2_1$
Unit-cell parameters [Å,°]	a=74.1, b=49.1, c=80.5, β=102.4	a=75.4, b=49.1, c=82.6, β=103.6
Mosaicity [°]	0.52	0.95
Molecules in ASU	2	2
Solvent content [%]	37	40
Resolution range [Å]	30.0-1.40 (1.45-1.40)	40.0-1.26 (1.31-1.26)
Reflections, collected / unique	496236 / 110982	696322 / 151662
R_{merge} [1]	0.071 (0.490)	0.059 (0.554)
Completeness [%]	100 (100)	95.4 (93.8)
$\langle I/\sigma(I)\rangle$	20.1 (2.7)	19.2 (2.0)
Average redundancy	4.5 (3.1)	4.6 (3.3)
B-value from Wilson plot [Å^2]	27.1	12.9
Refinement statistics		
Resolution [Å]	20.0-1.40	19.5-1.26
No. of reflections	109834	149903
No. of reflections in test set	1117	1654
R_{work} / R_{free}	0.161 / 0.186	0.142 / 0.182
No. of residues	634	637
Water molecules	506	763
Na^+	1	–
R.m.s.d. from ideal [2]		
bond lengths [Å]	0.019	0.018
bond angles [°]	1.73	1.74
Average B-factors [Å^2]	16.8	9.3
Ramachandran statistics [%]		
most favored regions	98.6	98.1
additionelly allowed regions	1.4	1.9
Clashscore from Molprobity	0.40	0.98
Poor rotamers from Molprobity	0.74	0.91
PDB code	3ur7	3ur8

[1] $R_{merge} = \Sigma_h \Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h.

[2]Engh & Huber (1991).

Table 2
Superposition statistics for plant endo-1,3-β-glucanases.

Calculations were carried out in *ALIGN* (Cohen, 1997) for C^{α} atoms (auto mode). A_{HD}, B_{HD} – protein chains of the present higher density crystals structure; A_{LD}, B_{LD} – protein chains of the present lower density crystals structure; A_{1GHS}, A_{2CYG} and A_{3EM5} – protein chains of the PDB models 1GHS, 2CYG and 3EM5, respectively.

Chains fitted	R.m.s.d. (Å) / No. pairs	Max. distance (Å)
A_{HD} onto B_{HD}	0.19 / 297	2.32
A_{HD} onto A_{LD}	0.26 / 296	4.93
A_{HD} onto B_{LD}	0.46 / 302	1.66
B_{HD} onto A_{LD}	0.26 / 301	5.05
B_{HD} onto B_{LD}	0.44 / 303	1.71
A_{LD} onto B_{LD}	0.45 / 301	2.04
A_{LD} onto A_{1GHS}	0.84 / 285	5.10
A_{LD} onto A_{2CYG}	0.83 / 287	5.82
A_{LD} onto A_{3EM5}	0.83 / 289	3.08
A_{1GHS} onto A_{2CYG}	0.62 / 280	6.68
A_{1GHS} onto A_{3EM5}	0.85 / 296	4.32
A_{2CYG} onto A_{3EM5}	0.78 / 304	4.77

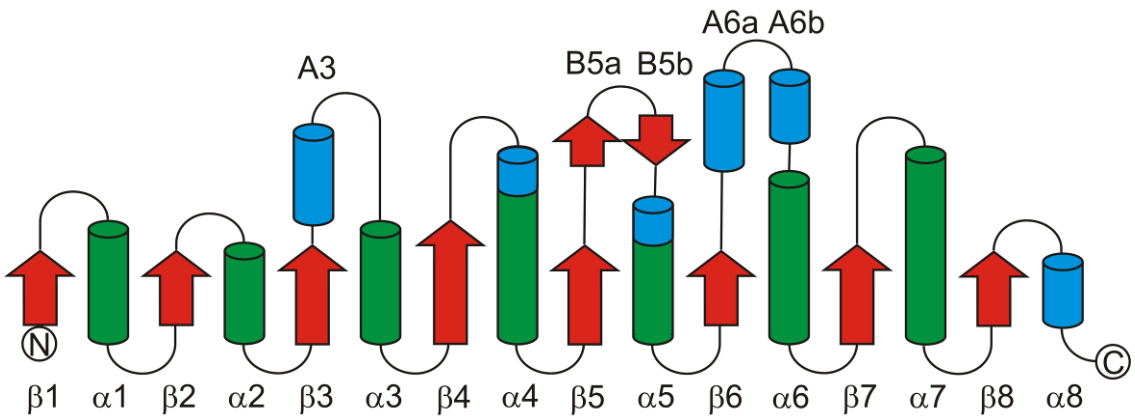

Fig. 1

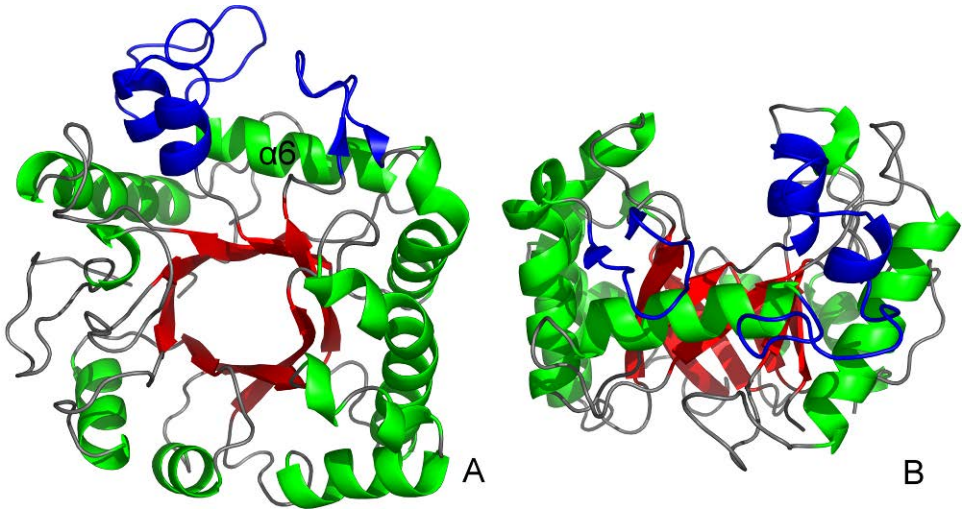

Fig. 2

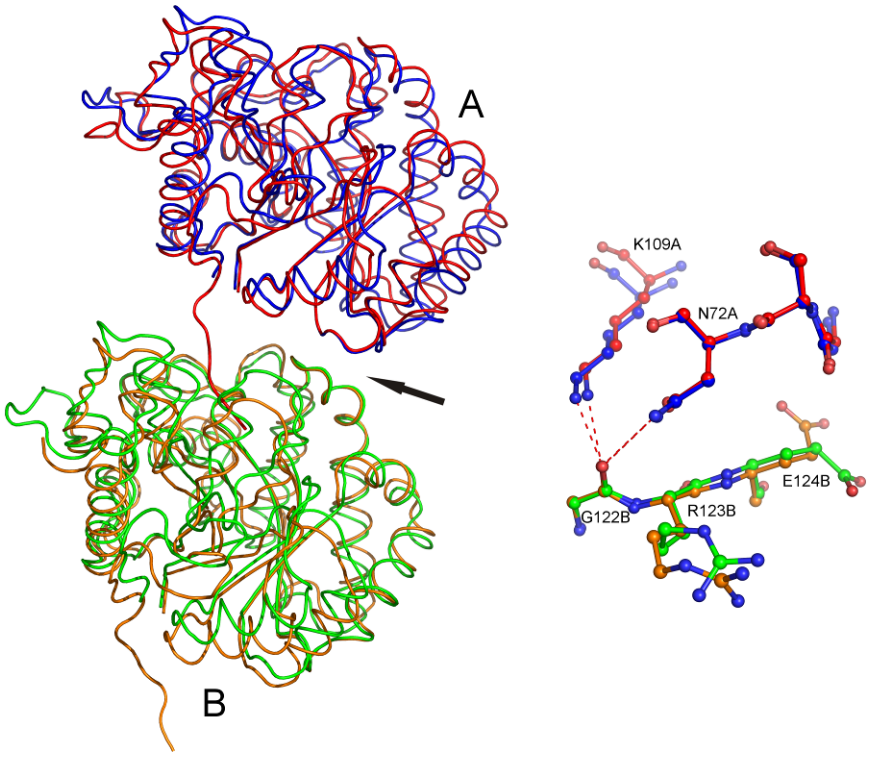

Fig. 3

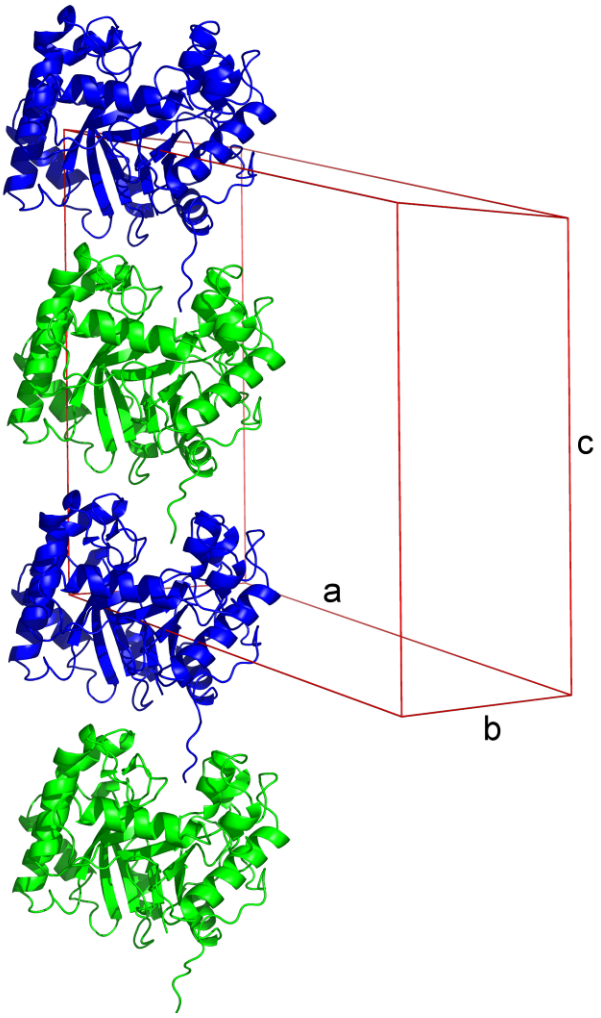

Fig. 4

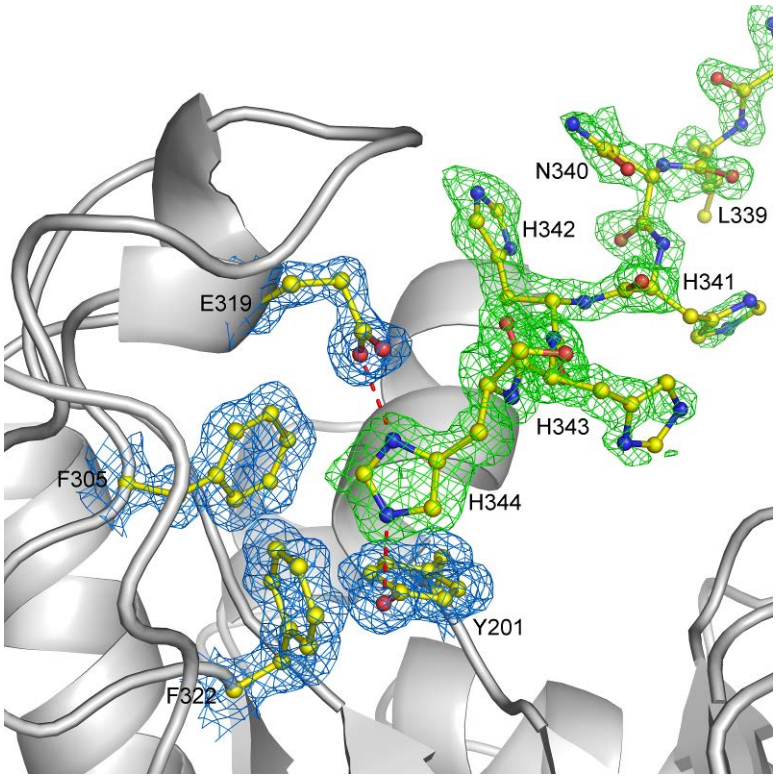

Fig. 5

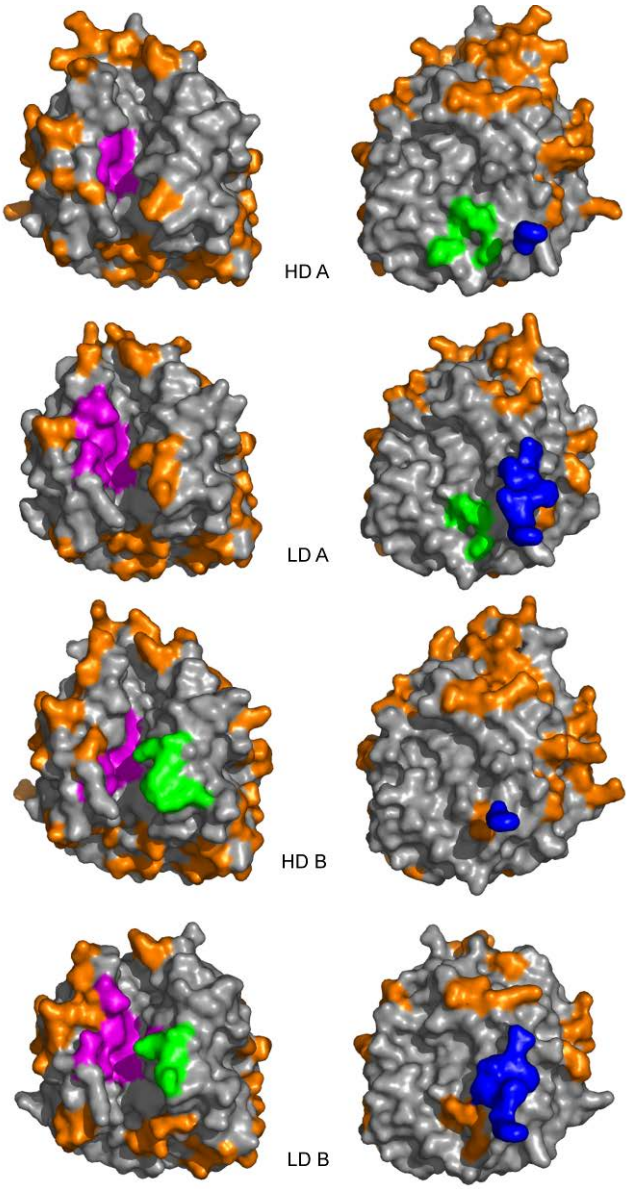

Fig. 6

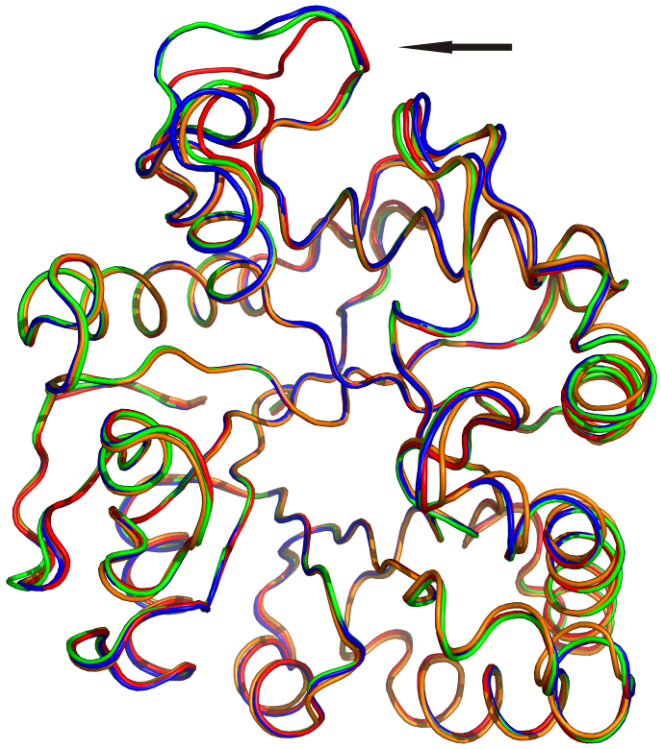

Fig. 7

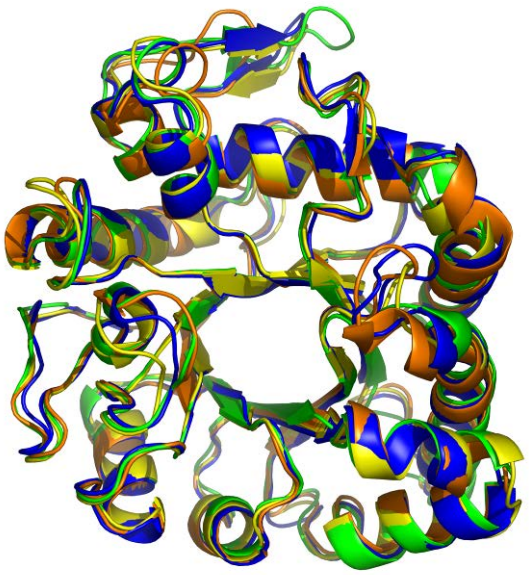

Fig. 8

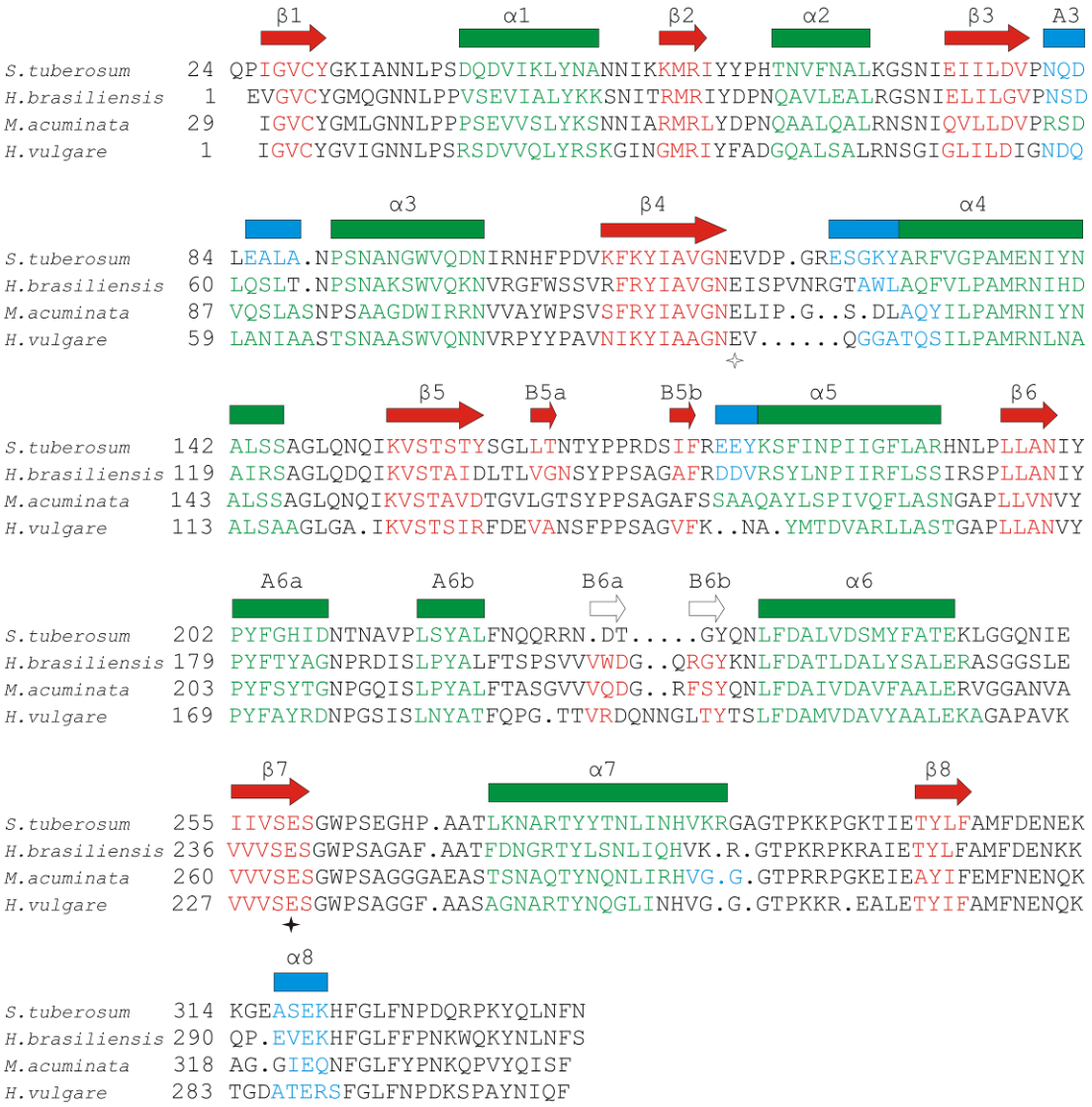
β1 α1 β2 α2 β3 A3
S.tuberosum 24 QPIGVCYGKIANNLPSDQDVIKLYNANNIKKMRIYYPHTNVFNALKGSNIEIILDVPNQD
H.brasiliensis 1 EVGVCYGMQGNNLPPVSEVIALYKKSNITRMRIYDPNQAVLEALRGSNIELILGVPNSD
M.acuminata 29 IGVCYGMLGNNLPPPSEVVSLYKSNNIARMRLYDPNQAALQALRNSNIQVLLDVPRSD
H.vulgare 1 IGVCYGVIGNNLPSRSDVVQLYRSKGINGMRIYFADGQALSALRNSGIGLILDIGNDQ
α3 β4 α4
S.tuberosum 84 LEALA.NPSNANGWVQDNIRNHFPDVKFKYIAVGNEVDP.GRESGKYARFVGPAMENIYN
H.brasiliensis 60 LQSLT.NPSNAKSWVQKNVRGFWSSVRFRYIAVGNEISPVNRGTAWLAQFVLPAMRNIHD
M.acuminata 87 VQSLASNPSAAGDWIRRNVVAYWPSVSFRYIAVGNELIP.G..S.DLAQYILPAMRNIYN
H.vulgare 59 LANIAASTSNAASWVQNNVRPYYPAVNIKYIAAGNEV......QGGATQSILPAMRNLNA
β5 B5a B5b α5 β6
S.tuberosum 142 ALSSAGLQNQIKVSTSTYSGLLTNTYPPRDSIFREEYKSFINPIIGFLARHNLPLLANIY
H.brasiliensis 119 AIRSAGLQDQIKVSTAIDLTLVGNSYPPSAGAFRDDVRSYLNPIIRFLSSIRSPLLANIY
M.acuminata 143 ALSSAGLQNQIKVSTAVDTGVLGTSYPPSAGAFSSAAQAYLSPIVQFLASNGAPLLVNVY
H.vulgare 113 ALSAAGLGA.IKVSTSIRFDEVANSFPPSAGVFK..NA.YMTDVARLLASTGAPLLANVY
A6a A6b B6a B6b α6
S.tuberosum 202 PYFGHIDNTNAVPLSYALFNQQRRN.DT.....GYQNLFDALVDSMYFATEKLGGQNIE
H.brasiliensis 179 PYFTYAGNPRDISLPYALFTSPSVVVWDG..QRGYKNLFDATLDALYSALERASGGSLE
M.acuminata 203 PYFSYTGNPGQISLPYALFTASGVVVQDG..RFSYQNLFDAIVDAVFAALERVGGANVA
H.vulgare 169 PYFAYRDNPGSISLNYATFQPG.TTVRDQNNGLTYTSLFDAMVDAVYAALEKAGAPAVK
β7 α7 β8
S.tuberosum 255 IIVSESGWPSEGHP.AATLKNARTYYTNLINHVKRGAGTPKKPGKTIETYLFAMFDENEK
H.brasiliensis 236 VVVSESGWPSAGAF.AATFDNGRTYLSNLIQHVK.R.GTPKRPKRAIETYLFAMFDENKK
M.acuminata 260 VVVSESGWPSAGGGAEASTSNAQTYNQNLIRHVG.G.GTPRRPGKEIEAYIFEMFNENQK
H.vulgare 227 VVVSESGWPSAGGF.AASAGNARTYNQGLINHVG.G.GTPKKR.EALETYIFAMFNENQK
α8
S.tuberosum 314 KGEASEKHFGLFNPDQRPKYQLNFN
H.brasiliensis 290 QP.EVEKHFGLFFPNKWQKYNLNFS
M.acuminata 318 AG.GIEQNFGLFYPNKQPVYQISF
H.vulgare 283 TGDATERSFGLFNPDKSPAYNIQF

Fig. 9

5. Supplementary materials

Table S1
Analysis of surface area [Å^2] buried on packing interactions.

	HD	HD with His-tag omitted	HD with His-tag & Q223B-D227B omitted	LD	LD with His-tag omitted
Total solvent-accessible area with generated symmetry-related atoms [Å^2]	19390	19480	19810	20110	20360
Total solvent-accessible area without considering symmetry-related atoms [Å^2]	27460	27200	27510	26460	25950
Total area difference owing to presence of symmetry-related atoms [Å^2]	-8060	-7710	-7690	-6350	-5590
Chain A area difference [Å^2] / % of total contact surface	-3930/48.8	-3800/49.3	-3800/49.4	-3600/56.7	-3210/57.4
Chain B area difference [Å^2] / % of total contact surface	-4130/51.2	-3910/50.7	-3890/50.6	-2750/43.3	-2380/42.6