Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

X-ray diffraction structure of a cell-wall invertase from *Arabidopsis thaliana*

Maureen Verhaest,[a] Willem Lammens,[a,b] Katrien Le Roy,[b] Barbara De Coninck,[b] Camiel J. De Ranter,[a] André Van Laere,[b] Wim Van den Ende[b] and Anja Rabijns[a]*

[a]Laboratorium voor Biokristallografie, Faculteit Farmaceutische Wetenschappen, K. U. Leuven, Herestraat 49, O&N II, Bus 822, B-3000 Leuven, Belgium, and [b]Laboratorium voor Moleculaire Plantenfysiologie, Faculteit Wetenschappen, Departement Biologie, K. U. Leuven, Kasteelpark Arenberg 31, B-3001 Heverlee, Belgium

Correspondence e-mail: anja.rabijns@pharm.kuleuven.be

Received 19 June 2006
Accepted 24 October 2006

PDB Reference: cell-wall invertase 1, 2ac1, r2ac1sf.

Cell-wall invertases play crucial roles during plant development. They hydrolyse sucrose into its fructose and glucose subunits by cleavage of the α1–β2 glycosidic bond. Here, the structure of the *Arabidopsis thaliana* cell-wall invertase 1 (AtcwINV1; gene accession code At3g13790) is described at a resolution of 2.15 Å. The structure comprises an N-terminal fivefold β-propeller domain followed by a C-terminal domain formed by two β-sheets. The active site is positioned in the fivefold β-propeller domain, containing the nucleophile Asp23 and the acid/base catalyst Glu203 of the double-displacement enzymatic reaction. The function of the C-terminal domain remains unknown. Unlike in other GH 32 family enzyme structures known to date, in AtcwINV1 the cleft formed between both domains is blocked by Asn299-linked carbohydrates. A preliminary site-directed mutagenesis experiment (Asn299Asp) removed the glycosyl chain but did not alter the activity profile of the enzyme.

1. Introduction

Sucrose (α-D-glucopyranosyl β-D-fructofuranoside) is one of the most widespread disaccharides in nature (Avigad & Dey, 1997) and is mainly limited to oxygenic photosynthetic organisms, including plants, unicellular algae and cyanobacteria (Salerno & Curatti, 2003). Sucrose is ubiquitous in higher plants and is the first free sugar resulting from photosynthesis (Koch, 2004). It is the major transport compound to bring energy and carbon skeletons from source to sink tissues. Some plant species use raffinose and/or stachyose for this purpose (Keller & Pharr, 1996). Plant growth and development is often accompanied by changes in source–sink relations (Sturm, 1999). Sessile life forms such as plants have apparently developed complex regulatory mechanisms enabling them to respond in a flexible way to different abiotic and biotic signals. Sucrose and hexoses play major roles as metabolic signals, regulating plant development by affecting different classes of genes (Smeekens, 2000; Gibson, 2005). Carbohydrate partitioning and sugar sensing are intimately connected to sucrose metabolism and these processes are vital throughout plant development (Koch, 2004).

The channelling of sucrose into sink metabolism often requires the cleavage of the α1–β2 glycosidic bond. In plants, this process can be catalyzed by sucrose synthase (SuSy; EC 2.4.1.13), a glycosyl transferase, and/or invertases (β-fructoside hydrolases splitting sucrose into glucose and fructose; EC 3.2.1.26). Alternatively, sucrose can be directly polymerized into fructans by the action of fructosyltransferases [*e.g.* 1-SST

(EC 2.4.1.99) and 1-FFT (EC 2.4.1.100)]. These enzymes do not use phosphate or nucleotide sugars, but use the energy in the Glc-Fru linkage to create novel Fru-Fru linkages (Vijn & Smeekens, 1999; Van den Ende *et al.*, 2004). To our knowledge, no α-glucosidases that break down sucrose have been reported from plants. Animals, in contrast, appear to have no β-fructosidases and only use α-glucosidases to break down the sucrose in their food.

Plant invertases can be classified into four categories: (i) soluble cytoplasmic invertases with a neutral to alkaline pH optimum, (ii) soluble vacuolar invertases, (iii) soluble apoplastic and (iv) insoluble cell-wall-bound invertases. The latter three have acidic pH optima (Tymowsk-Lalanne & Kreis, 1998; Kim *et al.*, 2000). Two invertase gene families can be discerned in *Arabidopsis thaliana*. The first family has eight members encoding six putative cell-wall invertases and two vacuolar invertases (Ji *et al.*, 2005). Two of the so-called cell-wall invertase genes do not encode a functional invertase but encode fructan exohydrolases (FEH; De Coninck *et al.*, 2005). The second family contains nine genes encoding cytosolic neutral/alkaline invertases; these enzymes belong to family 100 of the glycosyl hydrolases (GH 100; Ji *et al.*, 2005), a classification based on overall amino-acid sequence similarities (http://afmb.cnrs-mrs.fr/CAZY/; Henrissat, 1991). Vacuolar acid invertases may control sugar composition in fruits and storage organs (Scholles *et al.*, 1996). However, most research has been focused on cell-wall invertases since they play a crucial role in apoplastic phloem unloading and in doing so intervene in carbohydrate partitioning and the long-distance transport of sucrose. Cell-wall invertases have been implicated in defence responses (Roitsch *et al.*, 2003) and the regulation of seed (Miller & Chourey, 1992) and pollen development (Goetz *et al.*, 2001). Finally, because sucrose and hexoses also regulate gene expression, it is evident that invertase enzymes play a fundamental role in controlling gene expression, cell differentiation and development (Roitsch & Gonzalez, 2004). The clearest demonstrations of the importance of cell-wall invertase in plant development were provided by the antisense inhibition of carrot cell-wall invertase, which resulted in the prevention of tap-root formation (Tang *et al.*, 1999), and the aberrant pedicel and endosperm development in a maize mutant affected in a cell-wall invertase (Miller & Chourey, 1992).

Taking into account the enormous impact that cell-wall invertases can have on plant development, we set out to determine the three-dimensional structure of a cell-wall invertase from *A. thaliana*, a widely used model plant (AtcwINV1; gene accession code At3g13790). AtcwINV1 clearly shows the highest expression level of the six cell-wall-type hydrolases (Sherson *et al.*, 2003) and its expression level can be further induced after infection (Benhamou *et al.*, 1991; Fotopoulos *et al.*, 2003). Heterologous expression of AtcwINV1 in *Pichia pastoris* demonstrated that it is a typical invertase and not an FEH (De Coninck *et al.*, 2005).

Fructosyl transferases, FEHs and cell-wall and vacuolar invertases from plants group together with β-fructosidases and FEHs from microbial origin within family 32 of the glycosyl hydrolases (GH 32). The related family 68, which also contains some invertases, bacterial levansucrases and inulosucrases, can be combined with family 32 into clan GH-J (β-fructosidase superfamily; Naumoff, 2001). Within this clan, four three-dimensional structures of microbial enzymes have been determined to date: the levansucrases (GH family 68) from *Bacillus subtilis* (Meng & Fütterer, 2003; PDB code 1oyg) and *Gluconacetobacter diazotrophicus* (Martinez-Fleites *et al.*, 2005; PDB code 1w18), a β-fructosidase (GH family 32) from *Thermotoga maritima* (Alberto *et al.*, 2004; PDB code 1uyp) and an inulinase (β-fructosidase, GH 32) from *Aspergillus awamori* (Nagem *et al.*, 2004; PDB code 1y4w). The first plant GH 32 enzyme structure, 1-FEH IIa from *Cichorium intybus*, has recently been obtained (Verhaest, Van den Ende *et al.*, 2005; PDB code 1st8). This enzyme cannot degrade sucrose, but typically catalyzes the breakdown of inulin-type fructan. A new EC number has recently been assigned to this enzyme (EC 3.2.1.153) in order to differentiate it from classic microbial inulinases or β-fructosidases (EC 3.2.1.80) that can also break down sucrose. To the best of our knowledge, no three-dimensional structure has been obtained for a typical plant invertase. Here, we describe the three-dimensional structure of *A. thaliana* cell-wall invertase 1 (AtcwINV1), a key metabolic enzyme involved in the regulation of overall plant growth and development.

2. Experimental

2.1. Cloning and mutagenesis

Cell-wall invertase 1 (AtcwINV1; gene accession code At3g13790) from *A. thaliana* was cloned into pPICZα vector and heterologously expressed in *P. pastoris* as described by De Coninck *et al.* (2005). The pPicZαA-AtcwINV1 construct was used as a template for site-directed mutagenesis (Asn299Asp), following the QuickChange protocol (Strategene) and using the primer 5′-GGG GTT GGA CT**G** ACG AGT CAT CG-3′ and the complementary reverse primer. Chemically competent *Escherichia coli* TOP10 cells were transformed with the mutated plasmid and the mutations were verified by sequencing.

2.2. Purification and enzymatic essay

The invertase was heterologously expressed in *P. pastoris* and purified from the yeast supernatant (De Coninck *et al.*, 2005). Enzyme assays were performed as described in Van Riet *et al.* (2005). The substrates used were 5 m*M* sucrose, 1-kestose, levan and inulin dissolved in 50 m*M* sodium acetate. Reaction products were analysed by anion-exchange chromatography with pulsed amperometric detection (AEC-PAD, Dionex, Sunnyvale, CA, USA) as described by Van den Ende & Van Laere (1996). Relative substrate specificities were determined by comparison of the peak areas.

2.3. Crystallization and data collection

Two types of crystals were obtained as described previously (Verhaest, Le Roy *et al.*, 2005). Diffraction data were collected

at the BM14 beamline at the ESRF synchrotron (Grenoble, France; crystal type I) and at the ID14-3 beamline at the ESRF synchrotron (crystal type II) using glycerol as a cryoprotectant. Type I crystals diffracted to a resolution of 2.4 Å, while type II crystals diffracted to a resolution of 2.15 Å. These data sets were processed using *DENZO* and *SCALEPACK* v.0.97.647d (crystal type I) or v.1.97.2 (crystal type II) (Otwinowski & Minor, 1997). Data-collection statistics are summarized in Table 1.

2.4. Structure determination

Since invertase shares 52% amino-acid sequence identity with 1-FEH IIa (Verhaest, Van den Ende *et al.*, 2005), the latter structure could be used to solve the phase problem of invertase *via* the molecular-replacement technique (Hoppe, 1957; Rossmann & Blow, 1962). This part of the structure determination of invertase was performed using the diffraction data obtained from the type I crystals (2.4 Å resolution). The molecular replacement and further structure refinement were performed with programs from the *CNS* package (v.1.1; Brünger *et al.*, 1998). The first refinement cycle was performed with 528 of a total of 541 amino acids. The initial invertase model obtained from the type I crystals had an R_{work} of 33.33% and an R_{free} of 37.23%. When the type II crystal diffraction data became available, a new molecular-replacement round was carried out using the initial invertase model and the new crystal type data. Further refinement cycles with intermittent manual rebuilding in *O* (Jones *et al.*, 1991) were then performed using the type II crystal data until values of 20.17% (R_{work}) and 24.38% (R_{free}) were obtained, with the first four amino-acid residues not included in the final model because of lack of electron density. During refinement, 371 water molecules were progressively added when they met the following requirements: (i) a minimum 3σ peak was present in the $|F_{obs}| - |F_{calc}|$ difference map, (ii) a peak was clearly visible in the $2F_{obs}| - |F_{calc}|$ map, (iii) the B value for the water molecule did not exceed 60 Å^2 during refinement and (iv) the water molecule was stabilized by hydrogen bonding. Additionally, four glycerols, five *N*-acetylglucosamines and five mannoses were added to the crystal-

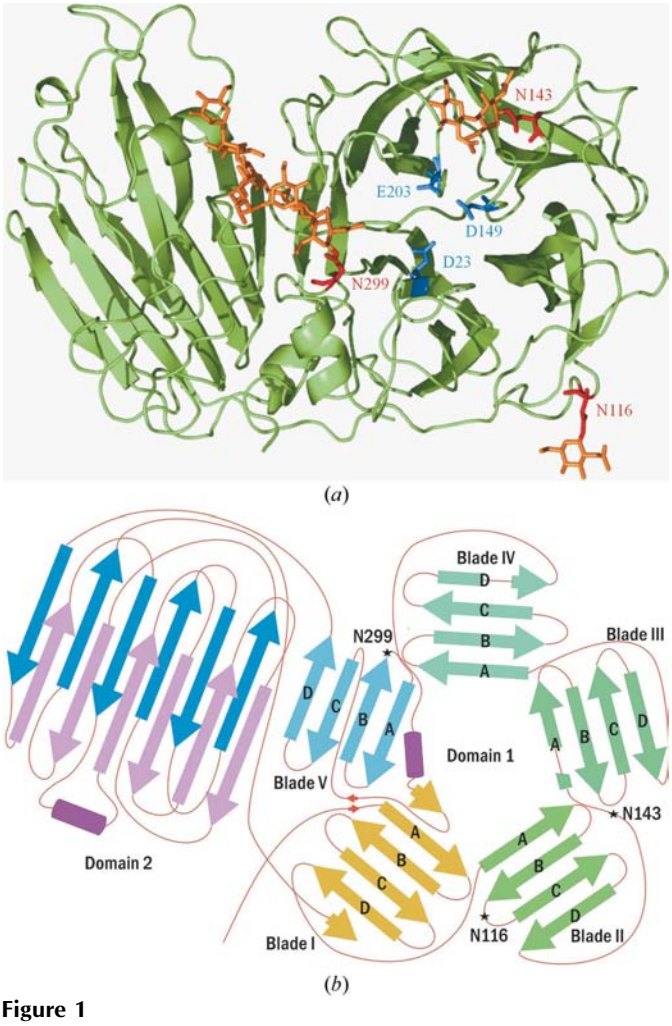

Figure 1
The three-dimensional structure of invertase. (*a*) The overall three-dimensional structure of invertase. Three glycosylation sites could be allocated to invertase on residues Asn116, Asn143 and Asn299 (indicated in red). The glycosylation chains are indicated in orange. The figure was prepared with *PyMOL* (DeLano, 2002). (*b*) Schematic diagram of the topology of invertase. β-Strands are depicted by arrows and α-helices by cylinders. Asterisks represent the glycosylation sites.

Table 1
Data-collection and reduction statistics for *A. thaliana* invertase.

Values in parentheses are for the highest resolution shell.

	Crystal type I	Crystal type II
Space group	$P3_2$	$C222_1$
Unit-cell parameters (Å)		
a	105.12	112.93
b	105.12	162.89
c	50.89	74.12
Wavelength used (Å)	0.984	0.931
Resolution limits (Å)	25.0–2.40 (2.49–2.40)	20.0–2.15 (2.19–2.15)
Total observations	72547	149921
Unique observations	23906 (1953)	38119 (1880)
Redundancy	3.0	3.9
Completeness (%)	97.4 (79.5)	99.9 (99.9)
Mean $I/\sigma(I)$	14.09 (1.9)	12.22 (2.48)
R_{sym} (%)	7.6 (33.2)	9.2 (50.3)

Table 2
Refinement statistics of the *A. thaliana* invertase structure.

Reflections (working/test)	63116/3279
Total No. of non-H atoms	4464
Protein atoms	4315
Water molecules	371
Glycerol molecules	4
N-Acetylglucosamine molecules	5
Mannose molecules	5
R_{work} (%)	20.17
R_{free} (%)	24.38
R.m.s.d. bond lengths (Å)	0.006
R.m.s.d. bond angles (°)	1.268
R.m.s.d. *B* factors (Å^2)	
Bonded main chain	1.199
Bonded side chain	2.006
Average *B* factors (Å^2)	
All protein atoms	29.06
Main-chain protein atoms	28.14
Side-chain protein atoms	29.96
Other entities	41.32

Table 3
The fivefold β-propeller structures.

The r.m.s. deviations from the cell-wall invertase from *A. thaliana* are shown. Numbers in parentheses indicate the numbers of residues used for r.m.s.d. calculations. (aa, amino acids). Amino-acid similarity was calculated using *ClustalW*.

Protein	Organism	Protein size (aa)	Fivefold β-propeller size (aa)	Total aa similarity with invertase (%)	Fivefold β-propeller aa similarity (%)	R.m.s. deviation from invertase (Å)	PDB code	Reference
AtcwINV1	*A. thaliana*	537	337	100	100	—	2ac1	This work
1-FEH IIa	*C. intybus*	543	335	52	54	0.84 (515 aa)	1st8	Verhaest, Van den Ende *et al.* (2005)
Invertase/β-fructosidase	*T. maritima*	432	295	22	28	1.7 (379 aa)	1uyp	Alberto *et al.* (2004)
Levansucrase	*B. subtilis*	447	447	11	9	2.0 (194 aa)	1oyg	Meng & Fütterer (2003)
Exo-inulinase	*A. awamori*	518	353	22	26	1.7 (396 aa)	1y4w	Nagem *et al.* (2004)
α-L-Arabinanase arb43A	*C. japonicus*	315	315	5	5	2.0 (202 aa)	1gyd	Nurizzo *et al.* (2002)
Tachylectin-2	*T. tridentatus*	236	236	8	8	2.3 (44 aa)	1tl2	Beisel *et al.* (1999)
Apyrase	*H. sapiens*	331	300	6	6	2.3 (66 aa)	1s1d	Dai *et al.* (2004)
Endo-1,5-α-L-arabinanase	*B. thermodenitrificans*	312	312	5	5	2.1 (192 aa)	1wl7	Yamaguchi *et al.* (2005)
Levansucrase	*G. diazothrophicus*	493	493	4	6	2.0 (204 aa)	1w18	Martinez-Fleites *et al.* (2005)

Table 4
Hydrogen bonds in the molecular velcro in the cell-wall invertase.

		Distance (Å)
Phe14 O	Gln316 N	3.05
Asn19 N	Gly311 O	2.87
Met21 N	Ser313 O	2.88
Trp63 N	Gln539 O	2.80
Trp63 O	Ser541 N	2.80

lographic model. The refinement statistics are summarized in Table 2. Ramachandran statistics showed that 84.2% of the residues reside in the most favoured region of the plot, 14.7% are in the additionally allowed regions and 0.6% are in the generously allowed regions. Two residues (His170 and His485) were found in a disallowed region of the Ramachandran plot. These unusual main-chain conformations could be corroborated in simulated-annealing electron-density maps and may be a consequence of the interaction of these main-chain atoms with surrounding residues.

3. Results and discussion

3.1. Overall fold of the invertase structure

The overall three-dimensional structure of the *A. thaliana* cell-wall invertase consists of an N-terminal fivefold β-propeller domain followed by a C-terminal domain formed by two β-sheets (Fig. 1*a*). Only two α-helices and four 3_{10}-helices are present (*DSSP*; Kabsch & Sander, 1983).

The propeller domain is based on a fivefold repeat of blades (numbered I–V), each containing four antiparallel β-strands, around a central axis (Fig. 1*b*). Strands are labelled *A*, *B*, *C* and *D* from the inside of the propeller outwards. A loop is inserted in two blades, creating a disruption in strand *A* of blade III between residues 149 and 153 and in strand *D* of blade IV between residues 262 and 270. A long loop between strands *B* and *C* of blade V is interrupted by an α-helix (residues 303–311) followed by a short β-strand that is hydrogen bonded to blade I.

The β-propeller fold has been found in many protein structures in fourfold to eightfold variants. While the first (sixfold) β-propeller fold was described in 1983 (influenza virus neuraminidase; Varghese *et al.*, 1983), the first fivefold β-propeller structure was only published in 1999 (tachylectin-2; Beisel *et al.*, 1999). Since then, the reported number of fivefold β-propeller structures has gradually increased. To date, ten fivefold β-propeller structures have been published (Table 3), including (i) an invertase/β-fructosidase from *T. maritima* (Alberto *et al.*, 2004), (ii) a levansucrase from *B. subtilis* (Meng & Fütterer, 2003), (iii) α-L-arabinanase Arb43A from *Cellvibrio japonicus* (Nurizzo *et al.*, 2002), (iv) a levansucrase from *G. diazothrophicus* (Martinez-Fleites *et al.*, 2005), (v) fructan-1 exohydrolase IIa (1-FEH IIa) from *Cichorium intybus* (Verhaest, Van den Ende *et al.*, 2005), (vi) an exo-inulinase from *Aspergillus awamori* (Nagem *et al.*, 2004), (vii) tachylectin-2 from *Tachypleus tridentatus* (Beisel *et al.*, 1999), (viii) a human apyrase (Dai *et al.*, 2004), (ix) an endo-1,5-α-L-arabinanase from *B. thermodenitrificans* (Yamaguchi *et al.*, 2005) and (x) the cell-wall invertase from *Arabidopsis thaliana* (this work). The structures of AtcwINV1 (GH 32), 1-FEH IIa (GH 32), invertase/β-fructosidase (GH 32), exo-inulinase (GH 32) and a levansucrase (GH 68), all members of the GH-J clan, are presented in Fig. 2. Most β-propeller structures are stabilized by closure of the ring of blades by a combination of β-strands from the N- and C-termini in the last blade, called 'molecular velcro' (Fülöp & Jones, 1999). The closure of the invertase propeller, however, is characterized by an atypical molecular velcro similar to that present in the 1-FEH IIa propeller (Verhaest, Van den Ende *et al.*, 2005). In invertase, a β-strand from blade V (amino acids 311–313, strand *C*) forms two hydrogen bonds with blade I (amino acids 19–21, strand *A*) (see Table 4). Similar hydrogen bonds exist in 1-FEH (Verhaest, Van den Ende *et al.*, 2005), invertase/β-fructosidase (Alberto *et al.*, 2004) and exo-inulinase (Nagem *et al.*, 2004). In invertase, a second closure is made by a β-strand coming from the second domain (amino acids 539–541) and forming two hydrogen bonds with the *D* β-strand of blade I at the N-terminus (amino acid 63). Consequently, blade I has two extra small β-strands (one at

the beginning and one at the end) forming a kind of double closure consisting of six β-strands in a 1 + 4 + 1 arrangement (Fig. 1*b*). This double closure is absent in invertase/β-fructosidase and exo-inulinase, where a 1 + 4 arrangement is observed. Furthermore, like invertase/β-fructosidase (Alberto *et al.*, 2004), exo-inulinase (Nagem *et al.*, 2004) and 1-FEH IIa (Verhaest, Van den Ende *et al.*, 2005), invertase shows an extra hydrogen bond between the loop at the N-terminus and the third strand of blade V (amino acids 14 and 316), again forming an extra closing contact.

The second domain of invertase starts near Leu342 and consists of two six-stranded β-sheets. The two β-sheets are composed of antiparallel β-strands and form a sandwich-like fold. This domain displays one disulfide bridge between Cys399 and Cys448. In invertase, 1-FEH IIa and exo-inulinase, almost all β-strands of this domain are longer than in invertase/β-fructosidase. The invertase loop from amino acids Val379–Pro412 contains a small α-helix; an equal arrangement is present in the 1-FEH IIa loop. In contrast, the equivalent loops in invertase/β-fructosidase or exo-inulinase are very small and lack this α-helix. In levansucrase, this second domain is completely missing.

Although to date nothing is known about the glycosylation of AtcwINV1 in *Arabidopsis*, in the invertase expressed in *P. pastoris* three glycosylation sites could be allocated in the electron-density maps, specifically on residues Asn116, Asn143 and Asn299 (Fig. 1*a*). Furthermore, the electron-density maps displayed at least one *N*-acetylglucosamine at each residue. One extra *N*-acetylglucosamine could be assigned to Asn143 and a more complex carbohydrate branching could be assigned to Asn299, which carried a total of two *N*-acetylglucosamines and five mannoses. The binding

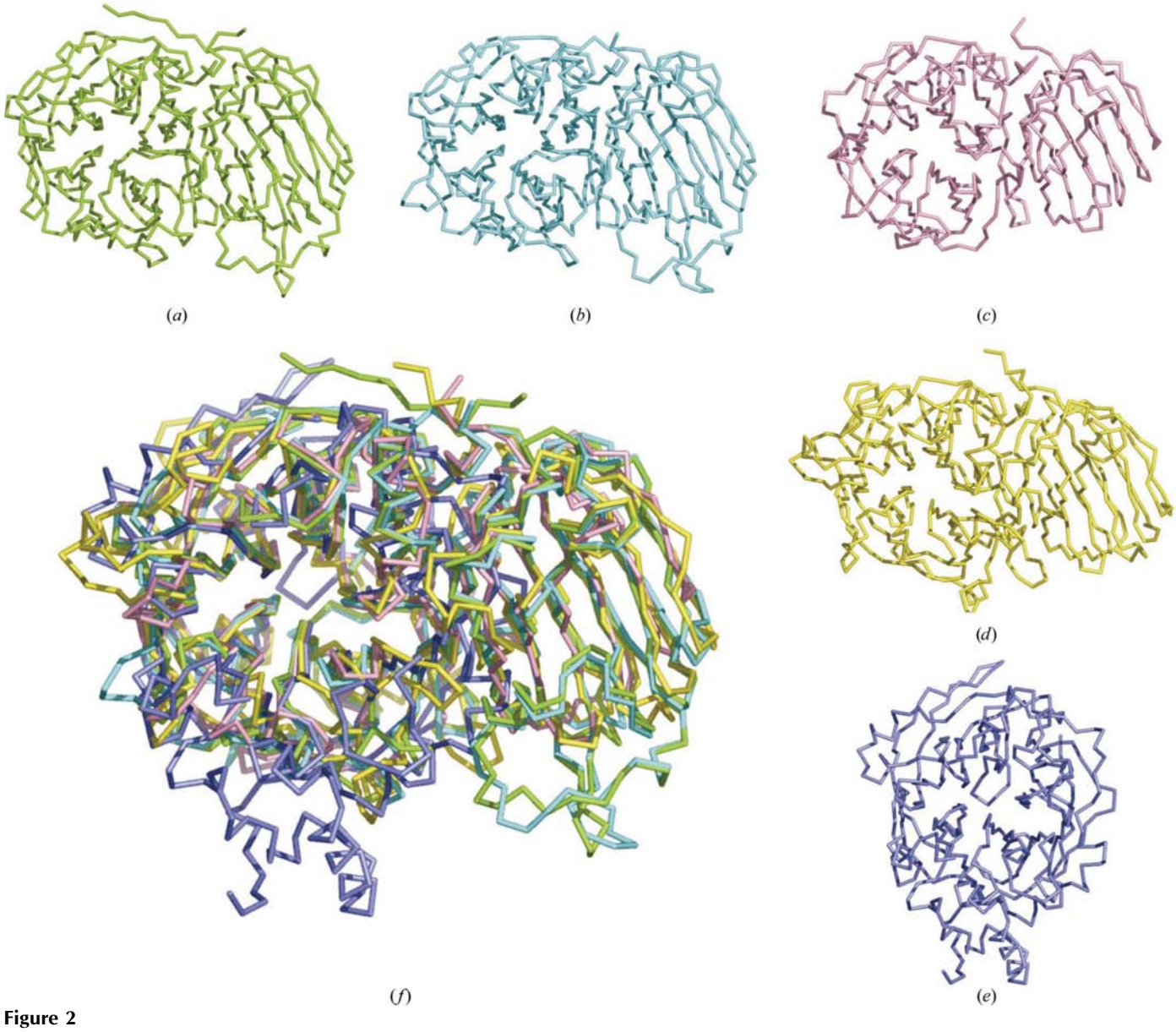

Figure 2
The three-dimensional structures of 1-FEH IIa (*a*), invertase (*b*), invertase/β-fructosidase (*c*), exo-inulinase (*d*) and levansucrase from *B. subtilis* (*e*). The superposition of these five proteins is also shown (*f*). Figures were prepared with *PyMOL* (DeLano, 2002).

Table 5
Relative substrate specificities of wild-type and mutant (Asn299Asp) AtcwINV1.

	Wild-type (%)	Asn299Asp (%)
Sucrose, 5 m*M*	100	10
1-Kestose, 5 m*M*	35	36
Levan, 5 m*M*	9	10
Inulin, 2.5% (5 m*M*)	3	4

between these various sugars on Asn299 is listed in Fig. 3. The glycosylations at Asn116 and Asn143 are both at the surface of the protein, far from the active site. Therefore, it is expected that the Asn116 and Asn143 glycosylations are not functionally important in the enzymatic reaction. However, the glycosylation chain at Asn299 blocks the cleft that emerges from the pocket-shaped active site and continues at the interface between the two domains. Plant invertases preferentially degrade sucrose (Table 5), but many of them are also able to degrade the kestose trisaccharides and raffinose to a certain extent (De Coninck *et al.*, 2005 and references therein). Longer fructan chains are poor substrates for these invertases (Table 5), but residual activity can still be detected, especially after long-term incubations. It can be speculated that the presence of the glycosylation sugars in the cleft prevents the binding of longer fructan substrates (see §4).

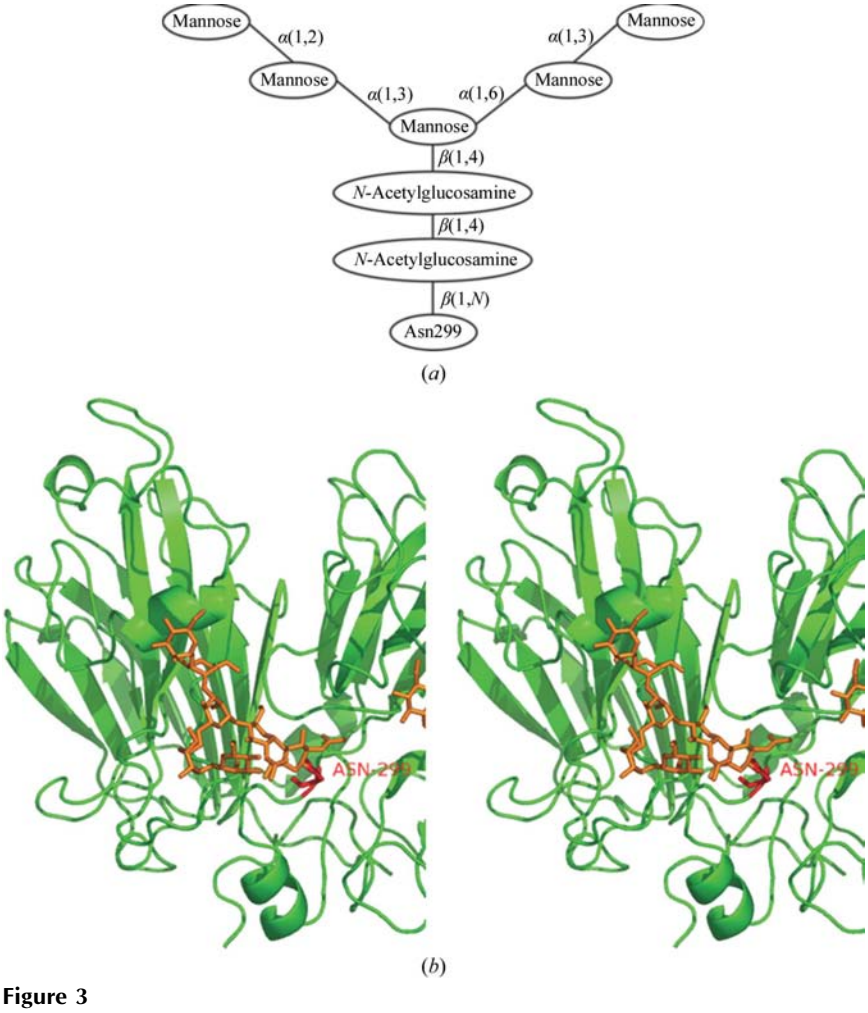

Figure 3
The different bonds between the Asn299 N-linked carbohydrates in invertase (*a*). The carbohydrate configuration was obtained through electron-density map interpretation. The stereo figure (*b*) shows the glycosylation chain (orange) on Asn299 (red). This figure was prepared with *PyMOL* (DeLano, 2002).

3.2. Active site and reaction mechanism

The structure of invertase shows a common fold (Fig. 2) with 1-FEH IIa, invertase/β-fructosidase and exo-inulinase, three other members of the GH family 32 proteins (Verhaest, Van den Ende *et al.*, 2005; Alberto *et al.*, 2004; Nagem *et al.*, 2004), as well as with the two GH family 68 levansucrase proteins (Meng & Fütterer, 2003; Martinez-Fleites *et al.*, 2005). Through superposition of these structures, it can be seen that the active sites of both GH families are located in the β-propeller domain and show extensive overlap with one another (Fig. 2). Equivalent residues of these five proteins are shown in Fig. 4. The implications of the structural similarities and differences will be discussed below.

A glycerol molecule resulting from the soak of the crystal in the cryoprotectant solution can be observed in the active site (Fig. 5) making the following interactions: O1 of glycerol interacts with Gln39 OE1 (3.09 Å) and Asn22 ND2 (3.45 Å), O2 interacts with Ser83 N (3.06 Å), Ser83 OG (3.39 Å) and Asp149 OD1 (3.34 Å) and O3 interacts with Glu203 OE1 (2.86 Å) and Asp149 OD2 (3.41 Å). The active site of invertase is composed of amino acids that are conserved within the GH family 32: Asp23, Asp149 and Glu203 in AtcwINV1 (Fig. 5), each spaced by 4.73–6.45 Å from each other, which is consistent with the double-displacement enzymatic reaction mechanism catalyzed by this enzyme. These residues are 'members' of the conserved regions NDPNG, FRDP and WECPD and play a crucial role in the catalytic mechanism of hydrolysis of the glycosidic bond (Reddy & Maley, 1990, 1996; Batista *et al.*, 1999; McCarter & Withers, 1994).

The general glycosyl hydrolase reaction mechanism involves the protonation of the glycosidic oxygen by an acid/base catalyst followed by a nucleophilic attack on the anomeric carbon of the sugar substrate by a carboxylate group (Koshland & Stein, 1954). Consistent with this general mechanism and by analogy with a study of Reddy & Maley (1996) on yeast invertase, the equivalent amino acid Asp23 of invertase was identified as the nucleophile and Glu203 as the acid/base catalyst. Consequently, the reaction scheme for invertase can be summarized as follows. Sucrose binds to the active site, where its glycosidic oxygen is protonated by Glu203. Subsequently, a nucleophilic attack

is performed by the carboxylate of Asp23, forming a covalent fructose–enzyme intermediate. Finally, this intermediate is hydrolyzed, releasing fructose and the free enzyme.

Structural analysis does not suggest a direct role for invertase Asp149 in the catalytic process, in contrast to mutation experiments on the equivalent levansucrase amino-acid residue (Batista *et al.*, 1999). Superposition of invertase with the sucrose present in levansucrase reveals that Asp149 is too far from the C2 hydroxyl group or the glycosidic oxygen to act as a catalytic residue. However, Nagem *et al.* (2004) suggest that Asp149 affects the nucleophilicity of Asp23. As is observed in all β-glycosyl hydrolases, the acid catalyst Glu203 of invertase needs to undergo a pK_a change of approximately two to three pH units before and after the formation of the covalent intermediate. In analogy with the 1-FEH IIa structure (Verhaest, Van den Ende *et al.*, 2005), we propose that the pK_a of the side chain of Glu203 in invertase is modulated by the proximity of Tyr279 (2.52 Å), which is strongly conserved in many GH 32 enzymes.

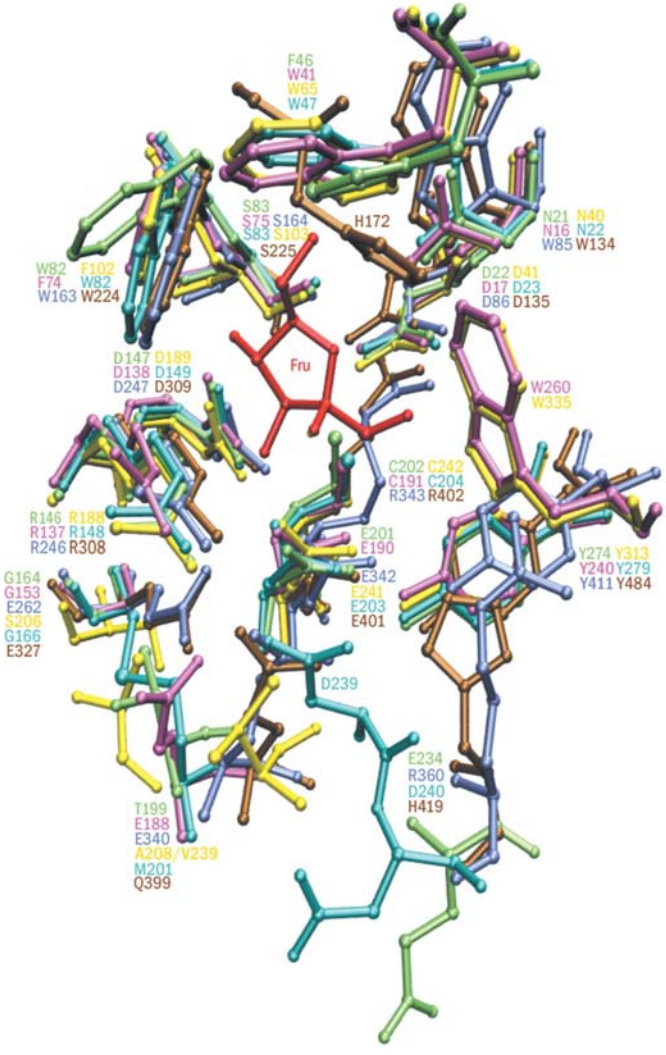

Figure 4
The three-dimensional structure of the active site. Superposition of the active site of invertase (light blue) with 1-FEH IIa (green), invertase/β-fructosidase (pink), exo-inulinase (yellow), levansucrase from *B. subtilis* (dark blue) and levansucrase from *G. diazothrophicus* (brown) showing similarities and variations in the catalytic region. The fructose (red) is oriented as in exo-inulinase to indicate the active site. This figure was prepared with *VMD* (Humphrey *et al.*, 1996) and *POV-Ray* (http://www.povray.org).

3.3. Plant invertases and FEHs show different substrate specificities

Essentially, plant invertases preferentially degrade sucrose (Table 5), while plant FEHs are unable to degrade sucrose but instead degrade fructans (Van den Ende *et al.*, 2004). As mentioned previously, invertases of microbial origin degrade fructans with a higher degree of polymerization (DP) at a much higher rate than plant invertases. Carefully inspecting the structures of all known GH 32 proteins immediately shows the presence of a blocking glycosylation chain in the cleft of the plant invertases. Plant 1-FEHs and microbial invertases do not have this glycosylation and the cleft is in a more open conformation. For 1-FEH IIa it was previously proposed that this cleft could play a role in binding the higher DP inulins (Verhaest, Van den Ende *et al.*, 2005). Hence, the hypothesis that the presence of a glycosyl moiety in the cleft of AtcwINV1 causes the reduced activity towards higher DP inulins was checked by mutational analysis. For this purpose, an Asn299Asp mutant was created by site-directed mutagenesis and heterologously expressed in *P. pastoris*. However, the recombinant protein did not show any difference in substrate

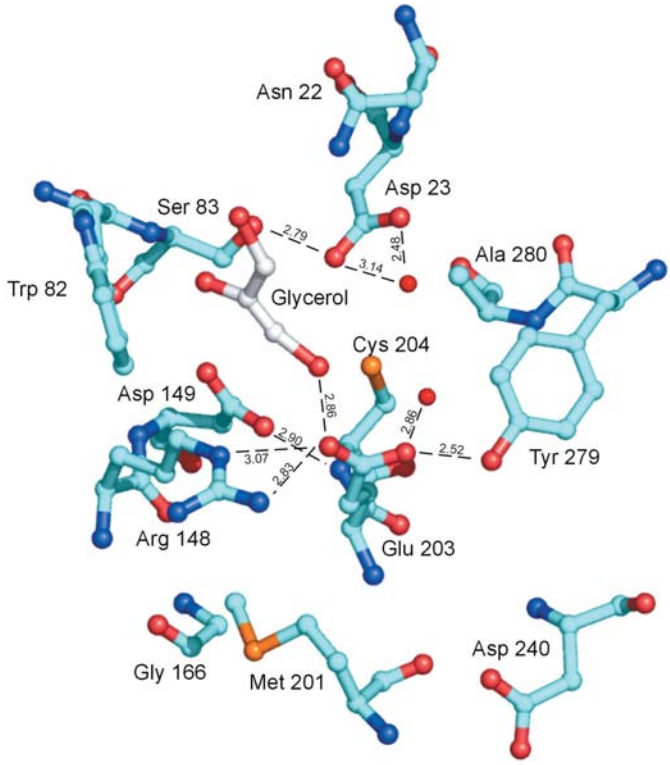

Figure 5
Active site of invertase. The key residues of the active site and one glycerol molecule are displayed. Bonding interactions are shown as dashed lines, while the spheres represent water molecules. The corresponding distances are given in Å. This figure was prepared with *PyMOL* (DeLano, 2002).

specificity in comparison with the wild-type enzyme (Table 5). For this reason, it can be concluded that the sugar chains of the glycosylation at Asn299 are not directly responsible for the inefficient degradation of longer inulins by plant invertases. Further mutational studies are needed to verify the role of this glycosylation at Asn299 and to resolve the functional differences between plant FEHs and invertases.

4. Conclusions

Plant invertases are key metabolic enzymes involved in the regulation of overall plant growth and development. For the first time a three-dimensional structure has been obtained for a typical plant invertase, a cell-wall invertase from *A. thaliana* (AtcwINV1). The structure comprises an N-terminal fivefold β-propeller domain followed by a C-terminal domain formed by two β-sheets. The active site is positioned in the fivefold β-propeller domain containing the nucleophile Asp23 and the acid/base catalyst Glu203 of the double-displacement enzymatic reaction. The glycosylation chain occluding the cleft does not seem to play an important role in substrate specificity. This new structure of a GH family 32 member could be the starting point for further mutation studies to test the functional hypotheses of the different amino acids suggested in this study. The determination of the AtcwINV1 structure allows further exploration of the biotechnological potential of interfering with invertase function in plants. Manipulated cell-wall invertases with lower K_m values might be used in the future to increase yield, fruit quality or secondary product formation in crop plants (Sonnewald *et al.*, 1997).

AR and WVdE are postdoctoral Research Fellows of the Fund for Scientific Research-Flanders (Belgium) (FWO-Vlaanderen). We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron-radiation facilities and we would like to thank the beamline scientists for assistance in using beamlines BM14 and ID14-3. BDC received a grant from IWT-Vlaanderen.

References

Alberto, F., Bignon, C., Sulzenbacher, G., Henrissat, B. & Czjzek, M. (2004). *J. Biol. Chem.* **279**, 18903–18910.
Avigad, G. & Dey, P. M. (1997). *Plant Biochemistry*, edited by P. M. Dey & J. B. Harborne, pp. 143–204. London: Academic Press.
Batista, F. R., Hernandez, L., Fernandez, J. R., Arrieta, J., Menendez, C., Gomez, R., Tambara, Y. & Pons, T. (1999). *Biochem. J.* **337**, 503–506.
Beisel, H. G., Kawabata, S., Iwanaga, S., Huber, R. & Bode, W. (1999). *EMBO J.* **18**, 2313–2322.
Benhamou, N., Grenier, J. & Chrispeels, M. J. (1991). *Plant Physiol.* **97**, 739–750.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Dai, J., Liu, J., Deng, Y., Smith, T. M. & Lu, M. (2004). *Cell*, **116**, 649–659.
De Coninck, B., Le Roy, K., Francis, I., Clerens, S., Vergauwen, R., Halliday, A. M., Smith, S. M., Van Laere, A. & Van den Ende, W. (2005). *Plant Cell Environ.* **28**, 432–443.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, USA. http://www.pymol.org.
Fotopoulos, V., Gilbert, M. J., Pittman, J. K., Marvier, A. C., Buchanan, A. J., Sauer, N., Hall, J. L. & Williams, L. E. (2003). *Plant Physiol.* **132**, 821–829.
Fülöp, V. & Jones, D. T. (1999). *Curr. Opin. Struct. Biol.* **9**, 715–721.
Gibson, S. I. (2005). *Curr. Opin. Plant Biol.* **8**, 93–102.
Goetz, M., Godt, D. E., Guivarc'h, A., Kahmann, U., Chriqui, D. & Roitsch, T. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6522–6527.
Henrissat, B. (1991). *Biochem. J.* **280**, 309–316.
Hoppe, W. (1957). *Acta Cryst.* **10**, 750–751.
Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.
Ji, X., Van den Ende, W., Van Laere, A., Cheng, S. & Bennett, J. (2005). *J. Mol. Evol.* **60**, 615–634.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Keller, F. & Pharr, D. M. (1996). *Photoassimilate Distribution in Plants and Crops. Source–Sink Relationships*, edited by E. Zamski & A. A. Schaffer, pp. 157–184. New York: Marcel Dekker.
Kim, J. Y., Mahe, A., Guy, S., Brangeon, J., Roche, O., Chourey, P. S. & Prioul, J. L. (2000). *Gene*, **245**, 89–102.
Koch, K. (2004). *Curr. Opin. Plant Biol.* **7**, 235–246.
Koshland, D. E. Jr & Stein, S. S. (1954). *J. Biol. Chem.* **208**, 139–148.
McCarter, J. D. & Withers, S. G. (1994). *Curr. Opin. Struct. Biol.* **4**, 885–892.
Martinez-Fleites, C., Ortiz-Lombardia, M., Pons, T., Tarbouriech, N., Taylor, E. J., Arrieta, J. G., Hernandez, L. & Davies, G. J. (2005). *Biochem. J.* **390**, 19–27.
Meng, G. & Fütterer, K. (2003). *Nature Struct. Biol.* **10**, 935–941.
Miller, M. E. & Chourey, P. S. (1992). *Plant Cell*, **4**, 297–305.
Nagem, R. A. P., Rojas, A. L., Golubev, A. M., Korneeva, O. S., Eneyskaya, E. V., Kulminskaya, A. A., Neustroev, K. N. & Polikarpov, I. (2004). *J. Mol. Biol.* **344**, 471–480.
Naumoff, D. G. (2001). *Proteins*, **42**, 66–76.
Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J. & Davies, G. J. (2002). *Nature Struct. Biol.* **9**, 665–668.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reddy, A. & Maley, F. (1996). *J. Biol. Chem.* **271**, 13953–13957.
Reddy, V. A. & Maley, F. (1990). *J. Biol. Chem.* **265**, 10817–10820.
Roitsch, T., Balibrea, M. E., Hofmann, M., Proels, R. & Sinha, A. K. (2003). *J. Exp. Bot.* **54**, 513–524.
Roitsch, T. & Gonzalez, M. C. (2004). *Trends Plant Sci.* **9**, 606–613.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Salerno, G. L. & Curatti, L. (2003). *Trends Plant Sci.* **8**, 63–69.
Scholles, J., Bundock, N., Wilde, R. & Rolfe, S. (1996). *Planta*, **200**, 265–272.
Sherson, S. M., Alford, H. L., Forbes, S. M., Wallace, G. & Smith, S. M. (2003). *J. Exp. Bot.* **54**, 525–531.
Smeekens, S. (2000). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **51**, 49–81.
Sonnewald, U., Hajirezaei, M. R., Kossmann, J., Heyer, A., Trethewey, R. N. & Willmitzer, L. (1997). *Nature Biotechnol.* **15**, 794–797.
Sturm, A. (1999). *Plant Physiol.* **121**, 1–8.
Tang, G. Q., Lüscher, M. & Sturm, A. (1999). *Plant Cell*, **11**, 177–189.
Tymowsk-Lalanne, Z. & Kreis, M. (1998). *Adv. Bot. Res.* **28**, 71–117.
Van den Ende, W., De Coninck, B. & Van Laere, A. (2004). *Trends Plant Sci.* **9**, 523–528.

Van den Ende, W. & Van Laere, A. (1996). *Plant Physiol.* **149**, 43–50.
Van Riet, L., Nagaraj, V., Van den Ende, W., Clerens, S., Wiemken, A. & Van Laere, A. (2005). *J. Exp. Bot.* **57**, 213–223.
Varghese, J. N., Laver, W. G. & Colman, P. M. (1983). *Nature (London)*, **303**, 35–40.
Verhaest, M., Le Roy, K., Sansen, S., De Coninck, B., Lammens, W., De Ranter, C. J., Van Laere, A., Van den Ende, W. & Rabijns, A. (2005). *Acta Cryst.* F**61**, 766–768.
Verhaest, M., Van den Ende, W., Le Roy, K., De Ranter, C. J., Van Laere, A. & Rabijns, A. (2005). *Plant J.* **41**, 400–411.
Vijn, I. & Smeekens, S. (1999). *Plant Physiol.* **120**, 351–360.
Yamaguchi, A., Tada, T., Wada, K., Nakaniwa, T., Kitatani, T., Sogabe, Y., Takao, M., Sakai, T. & Nishimura, K. (2005). *J. Biochem.* **137**, 587–592.