research papers

The epithelial adhesin 1 (Epa1p) from the human-pathogenic yeast *Candida glabrata*: structural and functional study of the carbohydrate-binding domain

Francesco S. Ielasi, Klaas Decanniere and Ronnie G. Willaert*

Research Group Structural Biology Brussels (SBB), Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Correspondence e-mail:
ronnie.willaert@vub.ac.be

The yeast *Candida glabrata* represents the second major cause of clinical candidiasis cases in the world. The ability of this opportunistic pathogen to adhere to human epithelial and endothelial cells relies on the Epa adhesins, a large set of cell-wall proteins whose N-terminal domains are endowed with a calcium-dependent lectin activity. This feature allows the yeast cells to adhere to host cells by establishing multiple interactions with the glycans expressed on their cell membrane. The ligand-binding domain of the Epa1p adhesin, which is one of the best characterized in the Epa family, was expressed in *Escherichia coli*, purified and crystallized in complex with lactose. Sequence identity with the domain of another yeast adhesin, the Flo5p flocculin from *Saccharomyces cerevisiae*, was exploited for molecular replacement and the structure of the domain was solved at a resolution of 1.65 Å. The protein is a member of the PA14 superfamily. It has a β-sandwich core and a D*cis*D calcium-binding motif, which is also present in the binding site of Flo5p. However, Epa1p differs from this homologue by the lack of a Flo5-like subdomain and by a significantly decreased accessibility of the solvent to the binding site, in which a calcium ion still plays an active role in the interactions with carbohydrates. This structural insight, together with fluorescence-assay data, confirms and explains the higher specificity of Epa1p adhesin for glycan molecules compared with the *S. cerevisiae* flocculins.

1. Introduction

Cell-adhesion proteins are critical to fungal cell interactions in development, symbiosis and pathogenesis (Dranginis *et al.*, 2007; Nather & Munro, 2008). Yeast cells use these interactions either to form large aggregates in liquid environments or to adhere to mammalian cells and cause yeast infections. While the ability of yeasts belonging to the *Saccharomyces* genus to flocculate at the end of the beer-fermentation process is well known and appreciated by brewers (Jin & Speers, 1999; Verstrepen *et al.*, 2003; Van Mulders *et al.*, 2010; Soares, 2011), *Candida* species show a rather undesirable tendency to adhere to human epithelial and endothelial cells, which eventually leads to candidiasis (Cormack *et al.*, 1999; Sundstrom, 2002; Pfaller & Diekema, 2007; Richter *et al.*, 2005). In each of the abovementioned processes, adhesion events are mediated by several families of glycosyl phosphatidyl inositol-anchored cell-wall proteins (GPI-CWPs). Two families of yeast adhesion-mediating GPI-CWPs are known to possess evident analogies in structure and function. The *Saccharomyces cerevisiae* Flo (flocculation) proteins (Verstrepen *et al.*, 2004; Goossens & Willaert, 2010) and the Epa (epithelial adhesion) proteins

from *C. glabrata* (Frieman *et al.*, 2002) show a well defined modular structure and a lectin-like calcium-dependent activity which enables the specific and reversible binding of surface-expressed oligosaccharide and polysaccharide residues and results in cell–cell or cell–surface interactions (Frieman *et al.*, 2002; Goossens & Willaert, 2010). The primary structures of these GPI-CWPs are composed of a C-terminal signal sequence for the addition of the GPI anchor, a central heavily glycosylated Ser/Thr-rich region and an N-terminal sequence which contains the domain necessary for carbohydrate recognition. In Flo and Epa adhesins, this domain has been identified as a PA14 domain from its homology to a 14 kDa protein fragment found in the anthrax toxin protective antigen (PA; Rigden *et al.*, 2004). PA14 domains are usually involved in carbohydrate recognition or carbohydrate metabolism as part of a very large number of enzymes, adhesins and toxins from both prokaryotic and eukaryotic organisms. They share a β-sandwich structure topology, in which the antiparallel β-sheets are mostly conserved while the loops are variable in
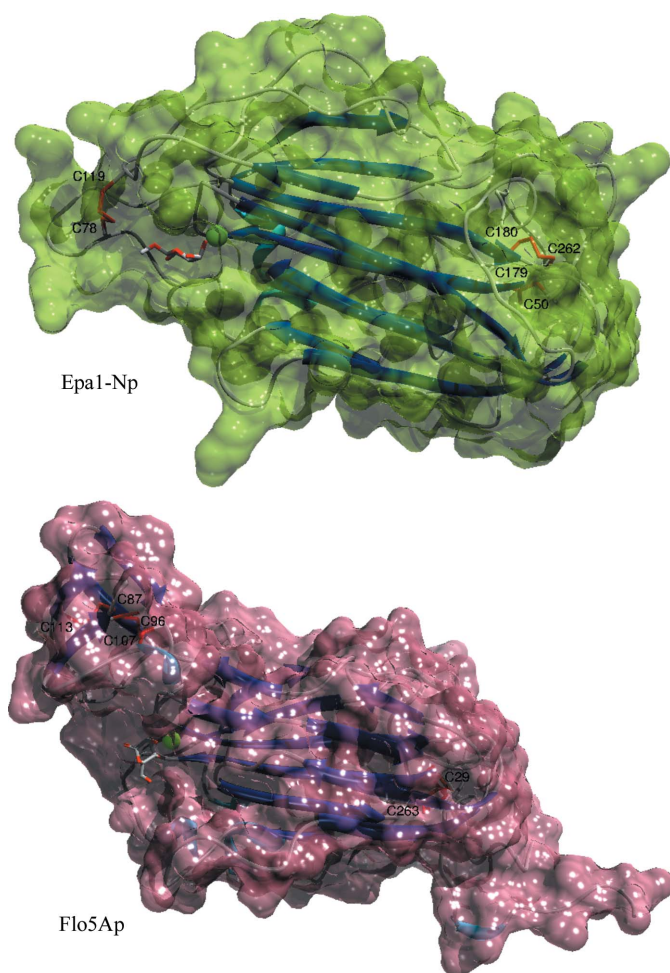

Figure 1
Comparison between the structures and surface shapes of Epa1-Np and Flo5Ap. In both panels, the PA14-like β-sandwich core of the protein is indicated in dark blue, the ligand structure included in the electron density is shown in the binding pocket together with calcium (green sphere), cysteines are labelled and disulfide bridges are indicated as red connections.

their sequences and lengths (Petosa *et al.*, 1997; Rigden *et al.*, 2004). The atomic structure of the N-terminal part of the Flo5 flocculin (Flo5Ap), including the PA14 domain, from *S. cerevisiae* has recently been determined by X-ray crystallography (Veelders *et al.*, 2010). Additionally, active involvement of calcium in the interaction with mannose residues was shown. A Ca^{2+} ion is bound to the lateral chains of two highly conserved aspartates which form a D*cis*D peptide in a surface loop. This rare structural motif and the divalent metal ion both directly contribute to recognition of mannose and mannose-containing glycans and are a distinctive feature of the Flo5p active site, which exhibits ligand affinity in the millimolar range.

Up to 23 different sequences are present in the *EPA* gene family (Kaur *et al.*, 2005). Each of them is predicted to encode a protein that is directly or indirectly involved in the adhesion of *Candida* to host tissues. However, the only proof that Epa proteins are responsible for adhesive events on epithelial and endothelial cells comes from studies of the *EPA1*, *EPA6* and *EPA7* genes (Cormack *et al.*, 1999; Castaño *et al.*, 2005; Zupancic *et al.*, 2008). The *EPA1* gene encodes the Epa1p adhesin, a 112 kDa glycoprotein which is responsible for at least 95% of wild-type *C. glabrata* adhesion to epithelial cells *in vivo*. Epa1p-mediated adhesion has been shown to be significantly decreased in the presence of lactose and *N*-acetyllactosamine, but not in sialyl derivatives of these disaccharides (Cormack *et al.*, 1999). A more recent study indicated that the N-terminal sequence of Epa1p preferentially binds glycans with a terminal galactose residue and a β1–3 or β1–4 glycosidic bond to galactose, glucose, *N*-acetylgalactosamine or *N*-acetylglucosamine (Zupancic *et al.*, 2008). The binding was significantly enhanced for branched glycans with two terminal β-galactose residues or when the penultimate residue was sulfated, while it was dramatically decreased by a terminal α-galactose, by sulfation of the terminal β-Gal, by fucosylation or sialylation of the penultimate residue or when GlcNAc was the penultimate residue (Zupancic *et al.*, 2008). Interestingly, it was found that of Galβ1–3Gal and Galβ1–3GalNAc, Epa1p exhibited a higher specificity for the latter. This disaccharide constitutes one of the core structures in the O-glycans of human mucins and is also known as Thomsen–Friedenreich antigen (TF or T antigen), a tumour-associated molecule (Brockhausen, 1999; Yu, 2007; Rhodes *et al.*, 2008).

In this work, we report the purification of the N-terminal region of Epa1p (Epa1-Np; Fig. 1), its crystal structure cocrystallized with β-lactose (Galβ1–4Glc) and an interaction study with three different saccharides (β-galactose, lactose and the T antigen) by spectrofluorimetry.

2. Materials and methods

2.1. Expression and purification

The *EPA1* gene (UniProt entry Q6VBJ0) from *C. glabrata* (ATCC 2001), spanning residues 31–278, was cloned into the pET-21b(+) expression vector (EMD Chemicals) containing a C-terminal His$_6$ tag and was overexpressed in the *Escherichia*

coli Origami 2 strain. The protein sequence for the construct was chosen by comparison with other proteins containing the PA14 domain (Veelders *et al.*, 2010). Cells were grown at 310 K until an $OD_{600 nm}$ of 0.5 was reached. The temperature was then decreased to 285 K and protein production was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 300 μ*M* followed by incubation for 2 d. The cells were harvested by centrifugation for 30 min at 4000*g* and 277 K and were resuspended in buffer 1 (20 m*M* NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.4, 1 *M* NaCl, 60 m*M* imidazole, 2.5 m*M* β-mercaptoethanol). The cells were lysed using a French press and the lysate was centrifuged for 30 min at 40 000*g* and 277 K. The supernatant was applied onto an Ni-affinity column (Ni Sepharose 6 Fast Flow, GE Healthcare) pre-equilibrated with buffer 1. Before elution, buffer 1 was changed directly on the column to buffer 2 (50 m*M* Tris–HCl pH 7.4, 200 m*M* NaCl, 60 m*M* imidazole, 2.5 m*M* β-mercaptoethanol). An isocratic elution of the protein was performed with buffer 2 containing 500 m*M* imidazole. Fractions containing Epa1-Np were pooled and CaCl$_2$ was added to a final concentration of 10 m*M*. The protein was subjected to a second affinity-purification step by applying it onto a column packed with lactose–agarose beads (Sigma–Aldrich) which was pre-equilibrated with buffer 3 (50 m*M* Tris–HCl pH 7.4, 200 m*M* NaCl, 10 m*M* CaCl$_2$, 2.5 m*M* β-mercaptoethanol). After binding, Epa1-Np was derivatized on the column with iodoacetamide under mild reaction conditions in order to block one of the free cysteine residues (Cys272). The reaction buffer (10 m*M* iodoacetamide, 50 m*M* Tris–HCl pH 8.5, 150 m*M* NaCl, 5 m*M* EDTA) was injected into the column and the reaction was carried out for 30 min. The protein was eluted with buffer 4 (50 m*M* Tris pH 7.4, 200 m*M* NaCl, 10 m*M* CaCl$_2$, 200 m*M* lactose). Fractions were eluted and dialysed against 50 m*M* Tris, 200 m*M* NaCl. The purified and modified protein was concentrated using a 3 ml Amicon ultrafiltration stirred cell (Millipore). The Epa1-Np concentration (molecular weight of 29 733 Da) was determined spectroscopically by measuring the absorbance of the sample at 280 nm.

2.2. Crystallization and data collection

The JBScreen 5–8 (Jena Bioscience) crystallization screens were chosen for initial screening of crystallization conditions. Screen plates for the sitting-drop method were prepared in 96-well Intelli-Plates using a Phoenix Liquid Handling System robot (Art Robbins Instruments). Crystallization was tested using Epa1-Np at concentrations of 1.5 and 3 mg ml^{-1} in 50 m*M* Tris, 200 m*M* NaCl, 10 m*M* CaCl$_2$, 20 m*M* lactose at 293 K. The hit conditions (100 m*M* HEPES pH 7.5, 20% PEG 8000) were optimized by varying the PEG 8000 concentration (from 15 to 25%), the pH (6.5 and 7.5) and the protein concentration (1.5, 3.0, 4.5 and 6.0 mg ml^{-1}). For the optimization screens, 24-well plates (Becton Dickinson) and the hanging-drop method were used. Drops were prepared by mixing 2 μl protein solution with 2 μl screen solution taken from the 500 μl in the reservoir. Crystallization was successfully achieved in about three weeks with an initial protein

Table 1
Data-collection, refinement and *MolProbity* structure-validation statistics for the Epa1-Np–lactose complex.

Values in parentheses are for the last shell.

Data-collection statistics	
Wavelength (Å)	1.000
Space group	$C222_1$
Unit-cell parameters (Å)	$a = 75.48$, $b = 105.45$, $c = 69.57$
Resolution range (Å)	42.00–1.65 (1.69–1.65)
No. of unique reflections	33944 (2464)
No. of observed reflections	33625
Completeness (%)	99.1 (87.7)
Multiplicity	6.9 (5.5)
$\langle I/\sigma(I) \rangle$	20 (3.2)
R_{merge}	0.063 (0.46)
Refinement statistics	
No. of reflections used in refinement	31929 (2046)
No. of reflections above σ cutoff in final cycle	31929 (2046)
Final overall R factor	0.174 (0.200)
Final R_{work}	0.172 (0.204)
Final R_{free}	0.210 (0.250)
Structure validation	
All-atom clashscore	5.81
Poor rotamers (%)	0.51
Ramachandran outliers (%)	0.00
Ramachandran favoured (%)	96.00

concentration of 6 mg ml^{-1} in 100 m*M* HEPES pH 7.5, 22.5% PEG 8000 at a temperature of 293 K. Diffraction data were collected on beamline PXIII at the Swiss Light Source.

2.3. Structure determination and refinement

The structure of Epa1-Np was solved by molecular replacement using *BALBES* including the *warpNtrace* option (Long *et al.*, 2008). The structure of Flo5Ap from *S. cerevisiae* in complex with calcium and mannose (PDB entry 2xjp; Veelders *et al.*, 2010) was used to generate a homology model for the adhesin. Refinement with *REFMAC*5 (Murshudov *et al.*, 2011) was alternated with model inspection and correction using *Coot* (Emsley *et al.*, 2010) until convergence. Structure figures were generated with *UCSF Chimera* (Pettersen *et al.*, 2004). Data-collection and refinement statistics are given in Table 1.

2.4. Spectrofluorimetric binding assay

The binding of β-D-galactose, Galβ1–4Glc (Sigma–Aldrich) and Galβ1–3GalNAc (Dextra) was studied by fluorescence spectroscopy. Measurements were performed using an LS55 fluorescence spectrometer (PerkinElmer) at a temperature of 298 K. Fluorescence spectra were acquired in the range between 300 and 400 nm using an excitation wavelength of 295 nm (excitation of Trp residues), an excitation bandwidth of 5 nm, an emission bandwidth of 8 nm and a scan speed of 300 nm min^{-1}. Each spectrum was acquired as an average of three consecutive scans over the indicated emission range. A decrease in fluorescence intensity was observed upon carbohydrate binding at 350 nm. The initial sample volume was 500 μl, consisting of 1 μ*M* Epa1-Np, 50 m*M* Tris, 200 m*M* NaCl and 10 m*M* CaCl$_2$. The sample was titrated with appropriate aliquots of 0.25, 2.5, 5 and 50 m*M* ligand solutions, which were prepared by dissolving the carbohydrates in the same buffer

solution as used for the protein. After each addition of carbohydrate, the sample was gently mixed with a pipette. Percentages of fluorescence quenching obtained from each titration step were plotted *versus* the carbohydrate concentration and the resulting curves were fitted using the *Prism*5 (GraphPad) software. A single-site ligand-binding model was used for all ligands.

3. Results and discussion

3.1. Overall structure of Epa1-Np and comparison with the PA14 and Flo5Ap structures

The Epa1-Np fold is mainly established by a β-sandwich which is made up of two antiparallel β-sheets containing six and four strands. As in Flo5Ap (Fig. 1), this constitutes the kernel of the protein, which belongs to the PF07691 PFAM family represented by the PA14 domain from the anthrax
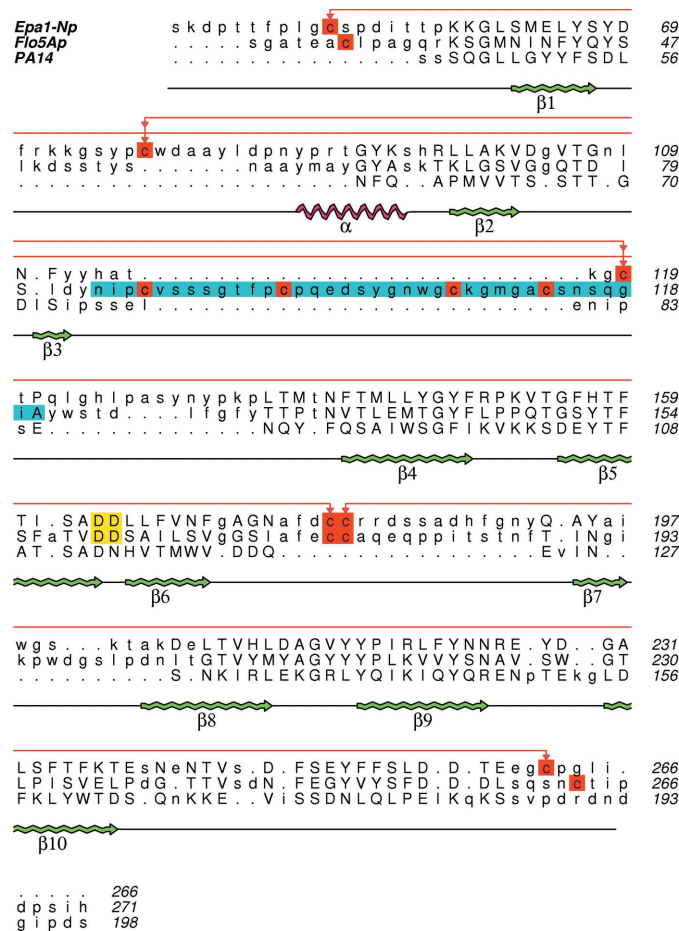
toxin protective antigen (UniProt entry P13423; PDB entry 1acc; Petosa *et al.*, 1997; Fig. 2). Homologues of PA14 are present in most Flo (Goossens & Willaert, 2010) and Epa adhesins (Rigden *et al.*, 2004; Zupancic *et al.*, 2008), as well as in the recently identified Pwp7 (PA14 domain-containing wall protein) adhesin from *C. glabrata* (Desai *et al.*, 2011). A carbohydrate-binding function rather than a catalytic role has been proposed for this domain. However, it can also reside in the catalytic domain of some enzymes, but always interacting with sugar moieties, such as in the β-glucosidase from *Kluyveromyces marxianus* (Yoshida *et al.*, 2010). In general, several lectins share an anti-β-sandwich topology, especially

Figure 2
Structure-based multiple alignment of the sequences of Epa1-Np (PDB entry 4a3x), Flo5Ap (PDB entry 2xjp) and PA14 domain from anthrax toxin protective antigen (PDB entry 1acc). The alignment was generated with *PDBeFold* (http://pdbe.org/fold) and edited with *ALINE* (Bond & Schüttelkopf, 2009). Residues indicated by capital letters are aligned in all three sequences. The following are highlighted: the Flo5 subdomain (light blue), cysteine residues (red) and the D*cis*D peptides (yellow). Disulfides bridges in Epa1-Np are indicated with red arrows and Epa1-Np structural motif positions are indicated below the sequence alignment.
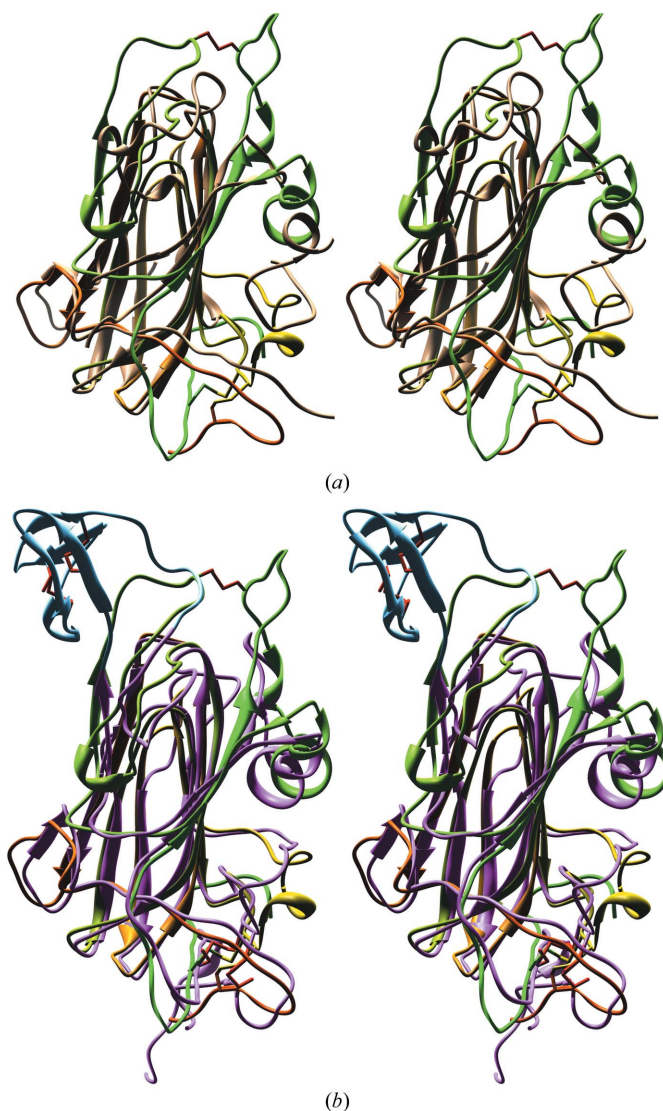

Figure 3
Superposition of the Epa1-Np structure on the two main PA14 homologues, represented as stereoviews. In both cases the structure of the *Candida* adhesin is indicated as a green–yellow–orange spectrum (moving from the N-terminus to the C-terminus); disulfide bonds are indicated as red connections. (*a*) Superposition on the PA14 domain structure (light brown) from the *Bacillus anthracis* protective antigen (r.m.s.d. of 2.25 Å for C$^\alpha$ atoms of 115 residues, 14% sequence identity). (*b*) Superposition on the Flo5Ap structure from *S. cerevisiae* (r.m.s.d. of 1.77 Å for C$^\alpha$ atoms of 191 residues, 32% sequence identity); the central body of the protein is represented in purple and the Flo5 subdomain is represented in light blue.

C-lectins, which also rely on the complexation of one or more Ca^{2+} ions for their carbohydrate-binding properties (Gabius *et al.*, 2011). Flo5Ap has indeed been shown to be a C-lectin with a calcium ion in its binding pocket and to be actively involved in interactions with mannosides (Veelders *et al.*, 2010). Here, it is demonstrated that Epa1-Np is also endowed with a calcium-dependent activity and possesses a similar binding site, although there are some structural differences and profound divergences in ligand specificity from the *S. cerevisiae* flocculins.

The C- and N-termini are both located opposite the binding pocket in the protein. They are tethered to the central body of the protein by two disulfide bridges (Fig. 1), one between Cys50 and Cys179 and the second between Cys180 and Cys262. This fixation of the two termini to the loop between strands $\beta6$ and $\beta7$ gives the protein a more compact shape and blocks access to the space between the two central sheets, thereby decreasing the surface of the β-sandwich exposed to the solvent. The location of the two juxtaposed extremities of the protein compared with the adhesive pocket confirms the N-terminal domain as the apex of the full Epa1p, which is used by the yeast cells as a 'hand' to reach for host glycans that are far away from the cell wall in the extracellular space. It was not possible to include fragments of the N-terminal (residues 31–39) and C-terminal (residues 267–278 plus linker-His$_6$ tag) parts in the final structure since no electron density was found for these fragments, although the indicated residues were included in the construct. In any case, these amino acids do not seem to play an important role in the structure and function of Epa1-Np, including residue Cys272, since this cysteine is not involved in any disulfide bond in Epa1-Np folding and could be blocked with iodoacetamide without affecting the ligand-binding activity. Problems were also found in fitting the side chain of Met63 in the electron density. A point mutation is unlikely, since the presence of the methionine codon and its correct translation were confirmed by plasmid DNA sequencing and direct MS analysis of the protein (data not shown). The difference found between the experimental electron density and the hypothetical density expected for a methionine at position 63 is therefore likely to be a result of radiation damage.

The superposition of Epa1-Np onto the PA14 peptide of the anthrax protective antigen (PDB entry 1acc) and Flo5Ap (PDB entry 2xjp) shows similarities and differences between the pathogenic yeast adhesin and its two homologues. In the first case (Epa1-Np/PA14; r.m.s.d. of 2.25 Å determined for the C^α atoms of 115 residues; 14% sequence identity; Fig. 3*a*), there is no similarity between the two domains apart from the central antiparallel β-sandwich. One of the most important features of Epa1-Np, *i.e.* the *cis*-peptide formed by Asp164 and Asp165, is absent in PA14, in which the *trans* Asp113–Asn114 gives a totally different conformation to the β-turn which is incompatible with calcium binding.

A comparison of Epa1-Np with yeast flocculin (Flo5Ap; r.m.s.d. of 1.77 Å determined for the C^α atoms of 191 residues; 32% sequence identity; Fig. 3*b*) shows the absence of the Flo5Ap subdomain (Veelders *et al.*, 2010), which is substituted

by a loop between strands $\beta3$ and $\beta4$, while the turns between $\beta1$ and the α-helix and between $\beta6$ and $\beta7$ are longer. These two loop fragments seem to shield the underlying structure of the Epa1-Np active site from the external environment and the disulfide bridge connecting them (between Cys78 and Cys119) may be very important for efficient and specific binding of the carbohydrate ligands.

3.2. Features of the carbohydrate-binding site

Interactions of the N-terminal domain of Epa1p with carbohydrates have already been extensively investigated in a qualitative and semi-quantitative manner (Zupancic *et al.*, 2008). Glycan microarray analysis was employed to determine the binding specificities of a purified N-terminal Epa1p fragment and of whole *S. cerevisiae* cells expressing Epa-Cwp2 fusion proteins for surface-immobilized glycans. Furthermore, adherence experiments to epithelial cells in the absence and presence of different inhibiting carbohydrates were carried out. In particular, the adhesion of *S. cerevisiae* cells expressing on the surface the N-termini of the closely related Epa6p and Epa7p and several chimeric constructs that were generated by swapping defined regions between Epa6-Np and Epa7-Np was evaluated. Two binding models were proposed. In both, the pentapeptide EYDGA (residues 227–231, as well as the corresponding EYDAE peptide in Epa7p) and two different binding regions were involved in the interaction mechanism. In the first model the pentapeptide was suggested to be a 'low-affinity' β-galactose-binding sequence compared with the DNDGA peptide of Epa6p, which was able to bind both galactose epimers and thus was defined as a 'high-affinity' binder. The presence of an additional site for interaction with a second sugar moiety was hypothesized. In contrast, the second model depicted the same pentapeptide as interacts with the second carbohydrate residue, placing the galactose-binding site elsewhere in the protein. A mutation in the first two amino acids would thus disrupt any contact with the second carbohydrate ring and would decrease the galactose concentration needed to inhibit the binding of the protein to glycans.

By cocrystallizing Epa1-Np with lactose in the presence of calcium, it was possible to discover the network of interactions in the binding site (Fig. 4). The divalent metal ion is required for substrate binding (as is also the case for Flo5Ap). Its role is not merely structural: the structure confirms direct contacts not only with the protein but also with hydroxyl groups of the carbohydrate ligands. This allows the adhesin to select specific carbohydrate conformations, *i.e.* only those that are able to enter into the coordination sphere of the calcium ion and to efficiently interact with the surrounding protein environment. As previously mentioned, the principal calcium-binding motif is the $D_{164}cisD_{165}$ peptide, which is located in the $\beta5$–$\beta6$ turn. The interaction between the lateral chains of the aspartates and the cation occurs in the range 2.3–2.5 Å. A second β-turn (between strands $\beta9$ and $\beta10$) is involved in the coordination of calcium. In this case, the backbone contributes directly to the formation of the complex with the carbonyls of Glu227
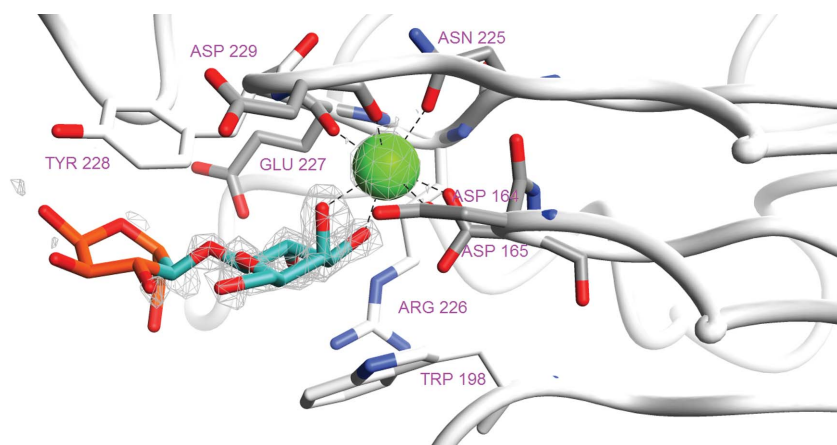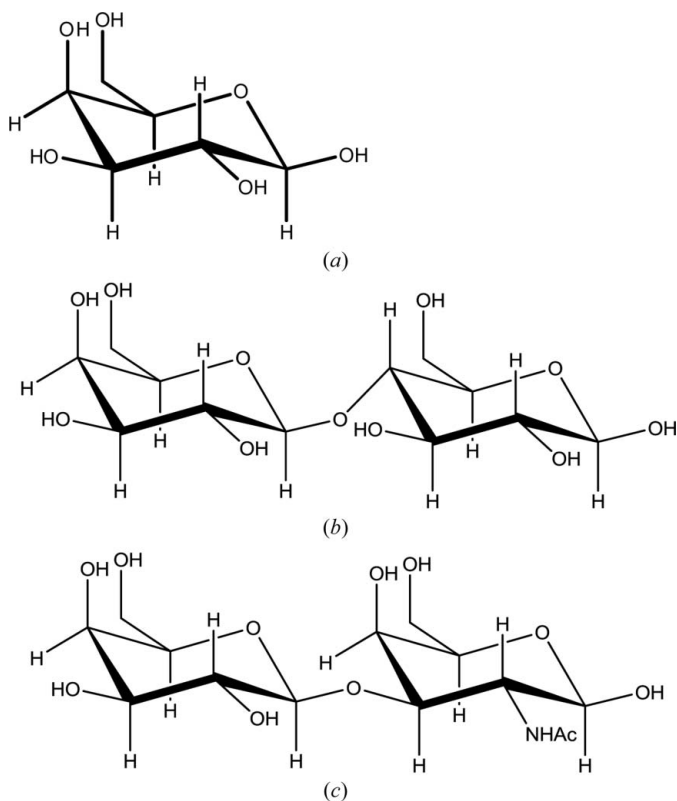
Figure 4
View of the Epa1-Np carbohydrate-binding site. All amino acids involved in the coordination of calcium and interactions with lactose are labelled. Electron density representing a composite OMIT map in proximity to the Ca atom and the carbohydrate is shown. The part of the lactose molecule which could be fitted to the electron density is indicated in light blue, while the excluded part of the ligand is indicated in orange.

Table 2
Carbohydrate-binding affinities obtained by fluorescence spectroscopy.

Substrate	Structure	Dissociation constant (μM)
Galβ (galactose)	Fig. 5(a)	115 ± 11
Galβ1–4Glc (lactose)	Fig. 5(b)	31.9 ± 4
Galβ1–3GalNAc (T antigen)	Fig. 5(c)	3.5 ± 0.4

Figure 5
Carbohydrate structures (see Table 2).

and Asp229. This loop corresponds to a variable amino-acid sequence, which reflects the low importance in this region of side-chain chemistry for calcium binding (Rigden *et al.*, 2004). An exception is Asn225 (which is also conserved in Flo5Ap), the side chain of which also contributes to the complexation of Ca^{2+}. The metal ion, which is locked in the binding pocket, can still accept two ligand groups from the 3′- and 4′-hydroxyl O atoms of the galactose ring. Consequently, the calcium coordination shell shows a distorted pentagonal bipyramidal geometry, which is quite common for binding sites of non-EF-hand proteins (Yang *et al.*, 2002).

During structure refinement, we were able to model the galactose residue and the β1–4 glycosidic linkage of lactose into the electron density, but we were not able to completely model the glucose ring. Nonetheless, we could visualize the interactions of the 3′- and 4′-hydroxyl groups with the metal ion and with Asp165 and Asp164, respectively. This once again confirms the importance of the *cis*-dipeptide. Based on this observation, the lack of binding found for glycans sulfated at the 3′ and 4′ positions of their terminal galactose residue can easily be explained (Zupancic *et al.*, 2008). The 2′-hydroxyl position is compatible with a role of hydrogen-bond acceptor from the Arg226 guanidinium group and donor to the Glu227 carboxylate group. Hydrophobic interactions also seem to be quite important for sugar binding (Gabius *et al.*, 2011), especially if we consider the residue Trp198 on the β7–β8 loop. The π-electrons delocalized on the tryptophan ring are involved in a stacking interaction with the aliphatic C—H patches on the galactose moiety. This may also lock the whole loop in a rigid conformation, which can consequently be seen as a 'sealing lid' for the ligand pocket. This stacking interaction would be incompatible with the axial 1′-hydroxyl of the α-galactose anomer, as well as with an α-glycosidic linkage connecting a terminal galactose to any glycosidic chain.

Since the Glc unit of lactose is missing from the electron density, only a few assumptions can be made about the interactions of Epa1-Np with the penultimate carbohydrate ring in the ligand molecule. The positions of the Glu227 and Tyr228 side chains suggest a possible binding interface in this region with a second sugar moiety in the pocket. In particular, the tyrosine could establish a stacking hydrophobic interaction with the latter, thereby increasing the binding interactions of the ligand in the adhesive domain. The higher glycan affinity of the Epa1 adhesin compared with Flo5Ap and its relatively more 'open' adhesive region can be explained by the presence of this rather 'closed' binding compartment. The fixed β3–β4 and β1–α loops and the more flexible β7–β8 turn seem to allow simultaneous access of both small carbohydrates and more complex glycan architectures and guarantee an efficient shielding of this compartment from solvent molecules upon binding.

3.3. Determination of the carbohydrate-binding affinity constants by fluorescence spectroscopy

As previously discussed, a tryptophan residue (Trp198) is involved in the carbohydrate-binding mechanism. The properties of this amino acid as a natural environment-sensitive fluorophore were exploited in order to probe the interactions of the Epa1-Np active site with different saccharides in a quantitative manner. β-Galactose, lactose and the Thomsen–Friedenreich antigen (Galβ1–3GalNAc) were chosen for fluorescent titration assays (Table 2).

For all three titration experiments, the best nonlinear regression fitting of the fluorescence quenching data was performed using a 'single binding site' model. This is in agreement with our structure, which shows no evidence for a second lactose-binding site. Among the three selected saccharides, β-galactose was the ligand with the lowest affinity. This affinity is comparable with the affinity of D-mannose for the high-affinity binding site in Flo1-Np (Goossens *et al.*, 2011), but is significantly higher than that of mannose for Flo5Ap, for which the K_d was determined to have a value in the millimolar range (Veelders *et al.*, 2010). Lactose showed an intermediate affinity, with a K_d almost four times lower than that of galactose. This confirms once again how a second carbohydrate residue can affect the interactions of the ligand molecule with the binding pocket. The T antigen, a disaccharide with a β1–3 instead of a β1–4 glycosidic linkage and GalNAc in place of Glu, is an even better binder, with an affinity that is tenfold higher.

These results, in combination with the structural findings, corroborate the previous results regarding the carbohydrate specificity of Epa1p (Zupancic *et al.*, 2008) and quantitatively define the influence of the individual ligand structure elements on carbohydrate–protein affinity.

4. Conclusions

In this work, we were able to solve the structure of the N-terminal domain of Epa1p, an adhesin from *C. glabrata* which mediates adhesion to epithelial and endothelial human cells. This domain shows a predominant PA14-like β-sandwich topology and a Ca^{2+}-dependent carbohydrate-binding site. The metal ion is heptacoordinated, involving not only the residues belonging to the binding pocket but also the ligand molecule itself. This is proof of an active role of calcium in the interaction network which is generated upon the adhesion event. Similarities to Flo5Ap are evident on a structural and functional basis, but in the case of Epa1-Np the specificity for carbohydrate ligands with a well determined stereochemistry is higher owing to a binding site with a decreased exposure to the external environment. A functional study performed by spectrofluorimetry confirmed this higher specificity towards β-galactose and, more remarkably, β-galactose-containing disaccharides.

References

Bond, C. S. & Schüttelkopf, A. W. (2009). *Acta Cryst.* D**65**, 510–512.
Brockhausen, I. (1999). *Biochim. Biophys. Acta*, **1473**, 67–95.
Castaño, I., Pan, S.-J., Zupancic, M., Hennequin, C., Dujon, B. & Cormack, B. P. (2005). *Mol. Microbiol.* **55**, 1246–1258.
Cormack, B. P., Ghori, N. & Falkow, S. (1999). *Science*, **285**, 578–582.
Desai, C., Mavrianos, J. & Chauhan, N. (2011). *FEMS Yeast Res.* **11**, 595–601.
Draginis, A. M., Rauceo, J. M., Coronado, J. E. & Lipke, P. N. (2007). *Microbiol. Mol. Biol. Rev.* **71**, 282–294.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Frieman, M. B., McCaffery, J. M. & Cormack, B. P. (2002). *Mol. Microbiol.* **46**, 479–492.
Gabius, H.-J., André, S., Jiménez-Barbero, J., Romero, A. & Solís, D. (2011). *Trends Biochem. Sci.* **36**, 298–313.
Goossens, K. V., Stassen, C., Stals, I., Donohue, D. S., Devreese, B., De Greve, H. & Willaert, R. G. (2011). *Eukaryot. Cell*, **10**, 110–117.
Goossens, K. & Willaert, R. (2010). *Biotechnol. Lett.* **32**, 1571–1585.
Jin, Y.-L. & Speers, R. A. (1999). *Food Res. Int.* **31**, 421–440.
Kaur, R., Domergue, R., Zupancic, M. L. & Cormack, B. P. (2005). *Curr. Opin. Microbiol.* **8**, 378–384.
Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nather, K. & Munro, C. A. (2008). *FEMS Microbiol. Lett.* **285**, 137–145.
Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H. & Liddington, R. C. (1997). *Nature (London)*, **385**, 833–838.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
Pfaller, M. A. & Diekema, D. J. (2007). *Clin. Microbiol. Rev.* **20**, 133–163.
Rhodes, J. M., Campbell, B. J. & Yu, L.-G. (2008). *Biochem. Soc. Trans.* **36**, 1482–1486.
Richter, S. S., Galask, R. P., Messer, S. A., Hollis, R. J., Diekema, D. J. & Pfaller, M. A. (2005). *J. Clin. Microbiol.* **43**, 2155–2162.
Rigden, D. J., Mello, L. V. & Galperin, M. Y. (2004). *Trends Biochem. Sci.* **29**, 335–339.
Soares, E. V. (2011). *J. Appl. Microbiol.* **110**, 1–18.
Sundstrom, P. (2002). *Cell. Microbiol.* **4**, 461–469.
Van Mulders, S. E., Ghequire, M., Daenen, L., Verbelen, P. J., Verstrepen, K. J. & Delvaux, F. R. (2010). *Appl. Microbiol. Biotechnol.* **88**, 1321–1331.
Veelders, M., Brückner, S., Ott, D., Unverzagt, C., Mösch, H. U. & Essen, L. O. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 22511–22516.

Verstrepen, K. J., Derdelinckx, G., Verachtert, H. & Delvaux, F. R. (2003). *Appl. Microbiol. Biotechnol.* **61**, 197–205.

Verstrepen, K. J., Reynolds, T. B. & Fink, G. R. (2004). *Nature Rev. Microbiol.* **2**, 533–540.

Yang, W., Lee, H.-W., Hellinga, H. & Yang, J. J. (2002). *Proteins*, **47**, 344–356.

Yoshida, E., Hidaka, M., Fushinobu, S., Koyanagi, T., Minami, H., Tamaki, H., Kitaoka, M., Katayama, T. & Kumagai, H. (2010). *Biochem. J.* **431**, 39–49.

Yu, L.-G. (2007). *Glycoconj. J.* **24**, 411–420.

Zupancic, M. L., Frieman, M., Smith, D., Alvarez, R. A., Cummings, R. D. & Cormack, B. P. (2008). *Mol. Microbiol.* **68**, 547–559.