

The 1.8 Å Resolution Structure of Hevamine, a Plant Chitinase/Lysozyme, and Analysis of the Conserved Sequence and Structure Motifs of Glycosyl Hydrolase Family 18

Anke C. Terwisscha van Scheltinga¹, Michael Hennig² and Bauke W. Dijkstra^{1*}

¹BIOSON Research Institute and Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4 9747 AG Groningen The Netherlands

²Biozentrum, University of Basel, 4056 Basel, Switzerland

The three-dimensional structure of hevamine, a plant enzyme with chitinase and lysozyme activity, has been refined at 1.8 Å resolution to an *R*-factor of 14.9% and a free *R*-factor of 19.6%. The final model consists of all 273 amino acid residues and 206 ordered water molecules. Two non-proline *cis*-peptides were identified, involving Phe32 and Trp255, both of which are implicated in substrate binding.

Other glycosyl hydrolase family 18 proteins with known three-dimensional structure are bacterial chitinase A, endo-β-*N*-acetylglucosaminidase F₁, endo-β-*N*-acetylglucosaminidase H, and the two plant proteins concanavalin B and narbonin, which have no known enzymatic activity. All these structures contain a (β α)₈ barrel fold, with the two family 18 consensus regions roughly corresponding to the third and fourth barrel strands. This confirms the grouping of these proteins into family 18, which was only based on weak and local sequence similarity. The substrate specificity of the enzymes is determined by the loops following the barrel strands that form the substrate binding site. All enzymes have an aspartic acid and a glutamic acid residue in positions identical with Asp 125 and the catalytic Glu127 of hevamine. The lack of chitinase activity of concanavalin B and narbonin can be explained by the absence of one of these carboxylate groups, and by differences in the loops that form the substrate-binding cleft in hevamine.

© 1996 Academic Press Limited

Keywords: TIM barrel; glycosyl hydrolase; X-ray structure comparison; *cis*-peptide; crystal contacts

*Corresponding author

Introduction

When infected by pathogens, plants may show pronounced defence responses. Cells surrounding the infected area are killed to prevent spreading of the pathogens (Ross, 1961), and a number of

so-called pathogenesis-related proteins (PR-proteins) are produced to attack and kill them (van Loon, 1985). The antifungal properties of these PR-proteins include 1,3-β-glucanase (Kauffmann *et al.*, 1987), lectin (Peumans & van Damme, 1995) and endochitinase (Legrand *et al.*, 1987) activity.

Hevamine is such a PR-protein, which has endochitinase as well as lysozyme activity. It can be isolated from the latex of the rubber tree (*Hevea brasiliensis*), where it is present in luteoids, vacuoles with a low internal pH (Tata *et al.*, 1983). The amino acid sequence of hevamine (Jekel *et al.*, 1991) shows significant similarity to those of other chitinases/lysozymes from plants and fungi, while there is a lower similarity to chitinases from bacteria, insects and viruses (Ayres *et al.*, 1994; Henrissat, 1990; Kramer *et al.*, 1993), and to human macrophage

Present address: M. Hennig, F. Hoffmann-La Roche AG, Pharma Division, 4070 Basel, Switzerland. Abbreviations used: PR, pathogenesis-related; chitotriose, *N*-,*N'*-,*N''*-triacetylchitotriose; chitohexaose, *N*-,*N'*-,*N''*-,*N'''*-,*N''''*-,*N'''''*-hexaacetylchitohexaose; endo F₁, endo-β-*N*-acetylglucosaminidase F₁; endo H, endo-β-*N*-acetylglucosaminidase H; *F*_o, observed structure factor; *F*_c, calculated structure factor; σ, standard deviation; RMS, root-mean-square; *V*_M, Matthew's volume; PEG, polyethylene glycol.

chitotriosidase (Boot *et al.*, 1995). All these chitinases have been classified as belonging to glycosyl hydrolase family 18 (Henrissat & Bairoch, 1993) on the basis of two consensus regions. The 2.2 Å X-ray structure of hevamine (Terwisscha van Scheltinga *et al.*, 1994) showed the enzyme to have a $(\beta\alpha)_8$ barrel fold with the two consensus regions (residues 73 to 80 and 119 to 127 in the hevamine sequence) approximately corresponding to the amino acid residues forming the third and fourth β -strand of the barrel. As most of these residues are buried, they are believed to be important for stabilization of the fold. In addition to the family 18 chitinases another, unrelated, family of chitinases exists, glycosyl hydrolase family 19 (Henrissat & Bairoch, 1993). As exemplified by the barley chitinase (Hart *et al.*, 1993), the fold of this family is totally different from that of the family 18 chitinases, and resembles that of the animal and phage lysozymes (Holm & Sander, 1994).

Binding studies with the chitin fragment chitotriose indicated that loops at the carboxy-terminal ends of the barrel β -strands form the substrate binding cleft (Terwisscha van Scheltinga *et al.*, 1994). This cleft contains six subsites, named from -4 to +2, and a conserved glutamic acid residue (Glu127), belonging to the second consensus sequence, is near subsites -1 and +1. A glutamic acid residue that is equivalent to Glu127 of hevamine was shown to be essential for catalytic activity of other family 18 chitinases (Tsujibo *et al.*, 1993; Watanabe *et al.*, 1993), suggesting that Glu127 in hevamine is a catalytic residue. This was confirmed by the X-ray structure of hevamine complexed with allosamidin, a bacterial family 18 chitinase inhibitor (Terwisscha van Scheltinga *et al.*, 1995).

Five other proteins with known X-ray structures contain the family 18 consensus regions, being chitinase A from *Serratia marcescens* (Perrakis *et al.*, 1994), endo- β -N-acetylglucosaminidase F₁ (endo F₁) from *Flavobacterium meningosepticum* (van Roey *et al.*, 1994), endo- β -N-acetylglucosaminidase H (endo H) from *Streptomyces plicatus* (Rao *et al.*, 1995), and the plant proteins concanavalin B from *Canavalia ensiformis* (Hennig *et al.*, 1995a) and narbonin from *Vicia narbonensis* (Hennig *et al.*, 1995b). Endo F₁ and endo H cleave the β (1,4)-glycosidic bond between the two core N-acetylglucosamine residues of asparagine-linked oligosaccharides, for narbonin and concanavalin B no enzymatic activity is known. Except for the 40% sequence identity between hevamine and concanavalin B (Hennig *et al.*, 1995a; Schlesier *et al.*, 1996) and the 32% sequence identity between endo F₁ and endo H, the overall similarity among the six proteins is low. Nevertheless, they all share the $(\beta\alpha)_8$ barrel fold of hevamine, with a similar position for the two consensus regions in all structures.

Here, we present the refined 1.8 Å resolution structure of hevamine and compare it with the other known family 18 structures. The increased resolution of the structure has enabled us to locate

with confidence water molecules in the active site, and to analyse the flexibility of the substrate-binding loops. These results are relevant for the further understanding of the catalysis by hevamine. From the comparison of the six family 18 three-dimensional structures, the role of the amino acid residues that form the conserved family 18 sequence motifs could be assessed.

Results and Discussion

Refinement

At the end of X-PLOR refinement, the *R*-factor and *R*_{free} had converged to 0.170 and 0.227, respectively. Subsequently, alternating cycles of manual rebuilding and TNT refinement were carried out to convergence. The final model contains all 273 amino acid residues and 206 water molecules. The *R*-factor (working set only) and *R*_{free} (test set only) are 0.148 and 0.196, respectively. The *R*-factor when using all data, i.e. including the test set, for refinement is 0.149.

Quality of the model

The hevamine model has good overall geometry (see Table 1), and the Ramachandran plot (Ramachandran & Sasisekharan, 1968) shows no residues in non-allowed regions (Figure 1). Ala47, the only residue found in the generously allowed region, is well-defined in electron density. It is fixed in its strained conformation through a hydrogen bond of the main-chain oxygen atom of residue 46 with the main-chain nitrogen atom of His49, and a hydrogen bond between its main-chain nitrogen atom and O^{δ1} of Asn45, of which N^{δ2} has been shown to hydrogen bond the carbohydrate substrate in subsite -3 (Terwisscha van Scheltinga *et al.*, 1994). Ala47 itself interacts with the bound substrate at subsite -4 through van der Waals contacts.

The side-chain of only one residue, Trp179, is close to an eclipsed conformation, with χ_1 being 6.9°. This residue is part of strand β_6 and is found in the core β region of the Ramachandran plot. Its side-chain is completely buried in the inside of the β -barrel, and its position is fixed by stacking on the aromatic side-chains of Tyr27, Tyr177 and Phe215, and by a hydrogen bond between N^{ε1} and the O^{δ1} atom of Asp123. Trp179 and Asp123 are strictly conserved in all family 18 plant chitinases.

The RMS coordinate error found from a Luzzati plot (Luzzati, 1952) on the basis of the test set is approximately 0.25 Å, the RMS shift of all atoms, compared with the 2.2 Å structure (Terwisscha van Scheltinga *et al.*, 1994), is 0.54 Å.

An indication of the quality of the structure on a per residue basis can be obtained from the real space correlation coefficient of the fit of the final model to the final σ_A weighted electron density map. This correlation coefficient is never lower than

Table 1. Statistics of data collection and refinement

	High-resolution data set		Low-resolution data set
Number of measurements	107,989		19,569
Number of unique reflections	23,108		4420
R_{merge} (outer resolution shell)	0.048 (0.149)		0.045 (0.061)
Resolution (Å)	1.80		3.20
Merged data set			
Number of unique reflections	23,188		
Completeness (%)			
∞ –1.80 Å	96.8		
1.85–1.80 Å	80.8		
Refinement			
Resolution range used (Å)	8.00–1.80		
Completeness of working set (%)	88		
Completeness of test set (%)	10		
R -factor, all data (%)	14.9		
R_{free} , test set (%)	19.6		
Number of protein atoms	2087		
Number of water molecules	206		
RMS deviations from ideal values			
Bond lengths (Å)	0.010		
Bond angles (deg.)	1.29		
Dihedrals (deg.)	22.9		
<i>B</i> -value correlation between bonded atoms (Å ²)	2.7		
Temperature factors	B_{min} (Å ²)	B_{max} (Å ²)	$\langle B \rangle$ (Å ²)
All protein atoms	2.0	94.7	14.5
Main-chain atoms	2.0	33.9	11.4
Side-chain atoms	2.8	94.7	17.9
Water molecules	8.8	76.9	33.6
$R_{\text{merge}} = \frac{\sum_{hkl} \sum_j I(hkl, j) - \langle I(hkl) \rangle }{\sum_{hkl} \sum_j \langle I(hkl) \rangle}$ $R\text{-factor} = \frac{\sum_{hkl} \ F_o(hkl) - F_c(hkl) \ }{\sum_{hkl} F_o(hkl) }$ $R_{\text{free}} = \frac{\sum_{hkl} \ F_o(hkl, \text{test}) - F_c(hkl, \text{test}) \ }{\sum_{hkl} F_o(hkl, \text{test}) }$			

0.70 (Figure 2), suggesting that there is no severe error in the structure. A similar result was obtained for the correlation of the final model to an OMIT (Bhat, 1988) map (Figure 2).

Several residues have an average temperature factor for side-chain atoms higher than 50 Å². (Table 2). For all these residues the main-chain atoms are well defined in electron density with an average temperature factor lower than 25 Å², suggesting that the residues are not part of disordered regions. The side-chains are polar or charged, and all point into the solvent; none of them hydrogen bonds to a protein atom. Therefore, it appears that these side-chains are not well defined because they are flexible.

Overall structure

Hevamine has a ($\beta\alpha$)₈ barrel fold (Terwisscha van Scheltinga *et al.*, 1994). The ($\beta\alpha$)₈ barrel β -strands and α -helices are numbered 1 to 8, the names of the loops are defined by the secondary structure elements they connect. Although most secondary structure elements are regular, the 1.8 Å refined structure shows a clear distortion of the central region of α 7, the helix formed by residues 231 to 244 (Figure 3). This is caused by Pro241, which cannot form a hydrogen bond with its main-chain nitrogen atom to the carbonyl oxygen atom of Ser237. At one side of this proline residue, the main-chain nitrogen

atoms of Ile239 and Leu240 form π -helix hydrogen-bonding contacts with the carbonyl oxygen atoms of Val234 and Leu235, respectively. At the other side, the main-chain nitrogen atom of Glu242 forms a 3₁₀ hydrogen bond with the carbonyl oxygen atom of Ile239. As a result, all main-chain nitrogen atoms of the helix (except for the amino-terminal three residues and Pro241) are involved in hydrogen-bonding main-chain oxygen atoms inside the helix, but the carbonyl oxygen atoms of Ile236, Ser237 and Arg238 do not make any hydrogen bonds with protein atoms (Figure 3).

The hevamine structure contains three *cis*-peptides. A *cis*-proline, Pro162, is found in the rigid β 5 α 5 loop, which forms one end of the substrate-binding cleft. Two other *cis*-peptides were originally modelled in the 2.2 Å electron density as *trans*-peptides, but with the 1.8 Å data it immediately became clear that this was incorrect. They are found between Ala31 and Phe32, and between Trp255 and Ser256 (Figure 4). The former is at the position where strand β 2 shifts its hydrogen-bonding interactions from β 1 to strand β 2' (Figure 4(a)), the latter is between β 8 and α 8' (Figure 4(b)). Phe32 and Trp255 are located at the bottom of the substrate-binding cleft, have a solvent-accessible surface of 13 and 49 Å², respectively, and are conserved in all known family 18 plant chitinase sequences. The complex of hevamine with chitotriose shows that these residues are involved in

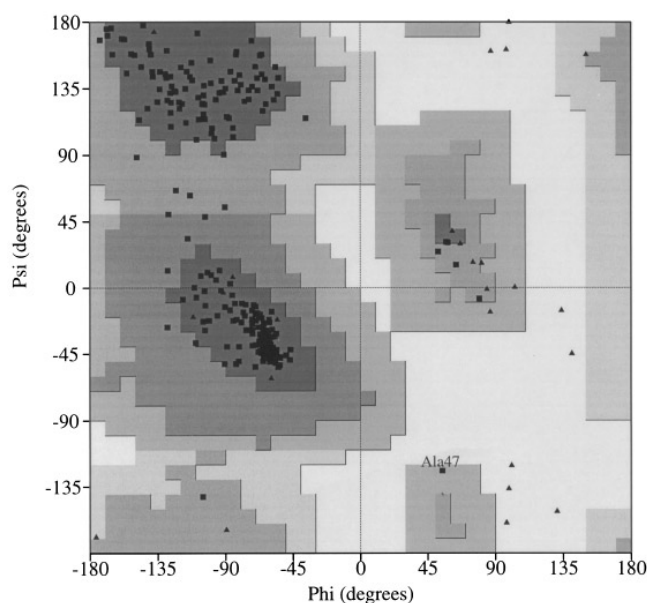


Figure 1. Ramachandran plot for hevamine, created using PROCHECK (Laskowski *et al.*, 1993). Glycine residues are shown as black triangles, non-glycine residues as black squares. Grey scale corresponds to the energetic preference of non-glycine residues: dark grey marks the most favored, grey the additional allowed, light grey the generously allowed and white the disallowed regions. Ala47, the one non-glycine residue found in the generously allowed regions, is labelled.

substrate binding (Terwisscha van Scheltinga *et al.*, 1994).

The active-site cleft

The substrate-binding cleft of hevamine is formed by the carboxy-terminal residues of the barrel β -strands and the subsequent loops. Although these $\beta\alpha\alpha$ loops on average are almost twice as long as the loops at the amino-terminal end of the β -strands, the average temperature factor is lower ($15.3 \text{ \AA}^2$ compared with $19.0 \text{ \AA}^2$). Indeed, the residues in the loops have well-defined conformations. Disulphide bonds within loop $\beta 2'\alpha 2$ and between the loops $\beta 5\alpha 5$ and $\beta 6\alpha 6$ restrict their flexibility. Loop $\beta 5\alpha 5$ has a very rigid structure, as it contains a half-cystine and three proline residues. The average temperature factor for its main-chain atoms is $8.7 \text{ \AA}^2$, which is significantly lower than the overall average temperature factor for main-chain atoms of $11.7 \text{ \AA}^2$.

The active-site cleft is lined with bound solvent molecules. Several of these solvent molecules are at positions where polar groups of the substrate bind, as identified in the complexes of hevamine with chitotriose or allosamidin (Terwisscha van Scheltinga *et al.*, 1994, 1995). Of the oxygen and nitrogen atoms of the chitotriose, bound at subsites -4 to -2 , two atoms are in a position similar to a water molecule in the native structure (Table 3). Of the allosamizoline moiety of allosamidin, which mimics an *N*-acetylglucosamine residue and which

binds at subsite -1 , this is the case for three atoms. For the sugar residues that were modelled at subsites $+1$ and $+2$ (Terwisscha van Scheltinga *et al.*, 1994), three atoms are in positions similar to water molecules, one atom of the sugar at subsite $+1$ and two atoms of the residue at subsite $+2$ (Table 3). These three water molecules are hydrogen bonded to residues in the loops $\beta 5\alpha 5$ and $\beta 6\alpha 6$, that are strictly conserved in family 18 plant and fungal chitinase sequences.

Crystal contacts

The packing of the hevamine molecules in the crystal lattice is tight, with a V_M of $2.13 \text{ \AA}^3/\text{Da}$. The crystals are very stable in a solution containing NaCl and in ammonium sulphate or PEG 6000 solutions. Moreover, the pH of the solutions can be varied at least from 2.0 to 9.5.

The interactions of molecules in the crystal are extensive, with 24 hydrogen bonds (distances between proton donor and acceptor shorter than $3.5 \text{ \AA}$) between a hevamine molecule and its neighbors (Table 4). Furthermore, there are 46 water molecules that form hydrogen bonds to two different protein molecules. The total solvent-accessible surface that is buried by crystal contacts is approximately $2600 \text{ \AA}^2$ per molecule, which is 23% of the total hevamine surface area. This tight packing may explain the stability of hevamine crystals, even when the precipitant in the stabilising solution is changed. The contacts involve no salt-bridges, but the hydrogen bonds do involve charged residues. As these charged residues are largely buried when making the crystal contacts, their degree of protonation probably is quite independent of the pH of the solvent, explaining the relative insensitivity of the hevamine crystals regarding changes in pH.

Comparison with homologous structures

At present, six high-resolution three-dimensional structures are known that contain the two family 18 consensus motifs. Four of them, hevamine, chitinase A, endo F_1 and endo H, cleave the glycosidic bond between two *N*-acetylglucosamine residues. The other two, narbonin and concanavalin B, have no known enzymatic activity. As all six structures comprise a very similar fold (Figure 5), it is of interest to try to establish the evolutionary relationships among the six proteins from their structures.

Concanavalin B, narbonin, endo F_1 and endo H have a chain length similar to that of hevamine (299, 290, 289 and 271 amino acid residues, respectively), and consist also of only one domain. Chitinase A, however, consists of 563 amino acid residues, and has two extra domains, one at the amino terminus of the $(\beta\alpha)_8$ barrel, and one as an insertion between $\beta 7$ and $\alpha 7$. The function

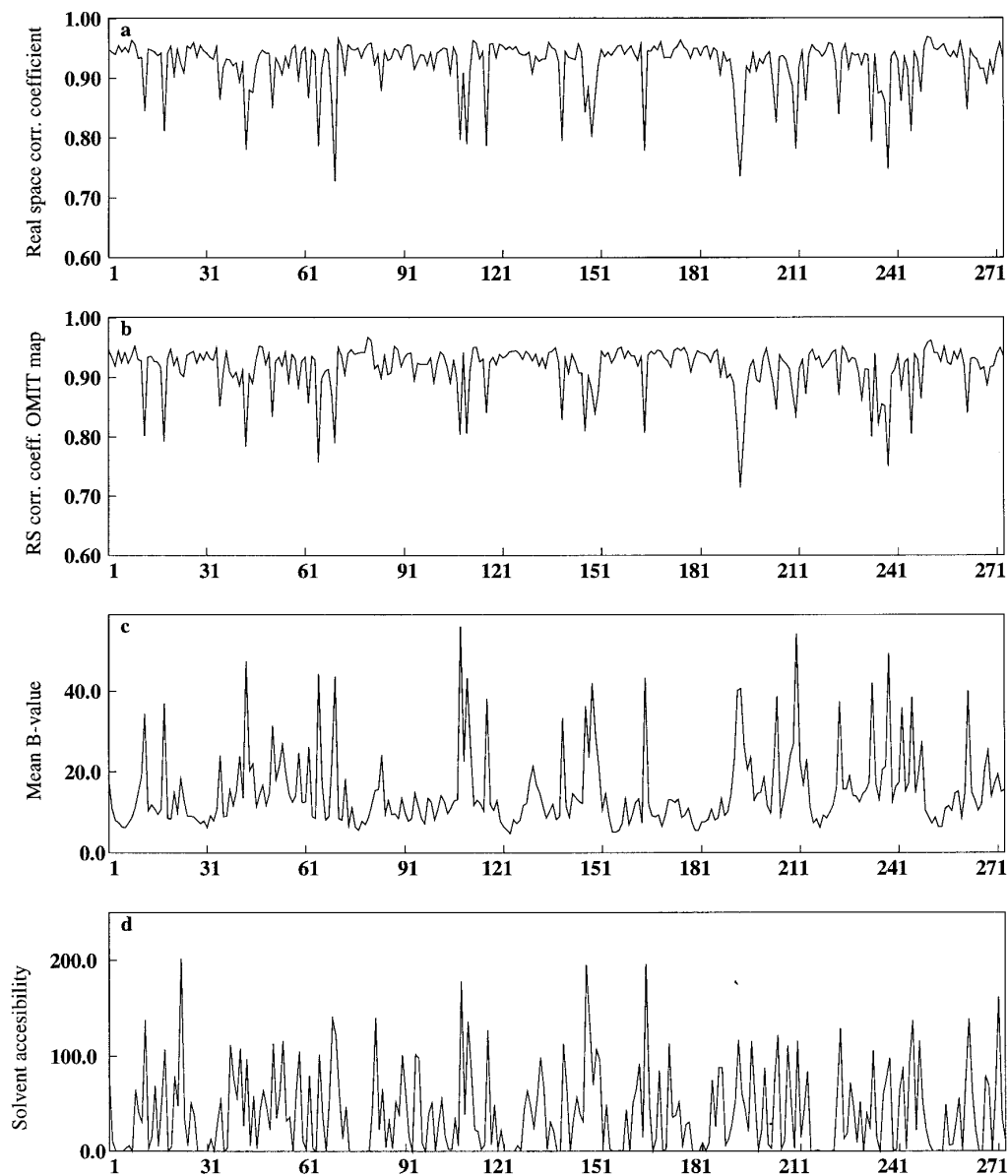


Figure 2. Quality of the structure, averaged over all atoms per residue. (a) Real space correlation coefficients for the final σ_A -weighted $2F_o - F_c$ electron density map. (b) Real space correlation coefficients for the final OMIT map. (c) Mean temperature factor ($\text{\AA}^2$). (d) Solvent accessibility ($\text{\AA}^2$).

of the amino-terminal domain, which has a fold similar to a fibronectin III module, is not known. The insertion domain, comprising a five-stranded antiparallel β -sheet and two helices,

makes up one of the sides of the substrate binding cleft.

The structure of hevamine is most similar to that of concanavalin B, with an RMS difference of 1.3 $\text{\AA}$

Table 2. Amino acid residues with the average temperature factor of the side-chain atoms higher than 50 $\text{\AA}^2$

Residue	$\langle B \rangle$ side-chain ($\text{\AA}^2$)	$\langle B \rangle$ main-chain ($\text{\AA}^2$)	Solvent-accessible surface ($\text{\AA}^2$)	Position in structure
Gln43	55.6	13.4	97	$\beta 2'$
Arg65	55.3	9.0	102	$\alpha 2$
Gln70	59.2	12.7	122	$\alpha 2\beta 2$
Lys108	64.4	20.9	178	$\alpha 3\beta 3$
Arg164	51.2	10.6	196	$\beta 5\alpha 5$
Asn210	58.9	23.8	116	$\beta 6\alpha 6$
Arg238	59.5	17.7	89	$\alpha 7$

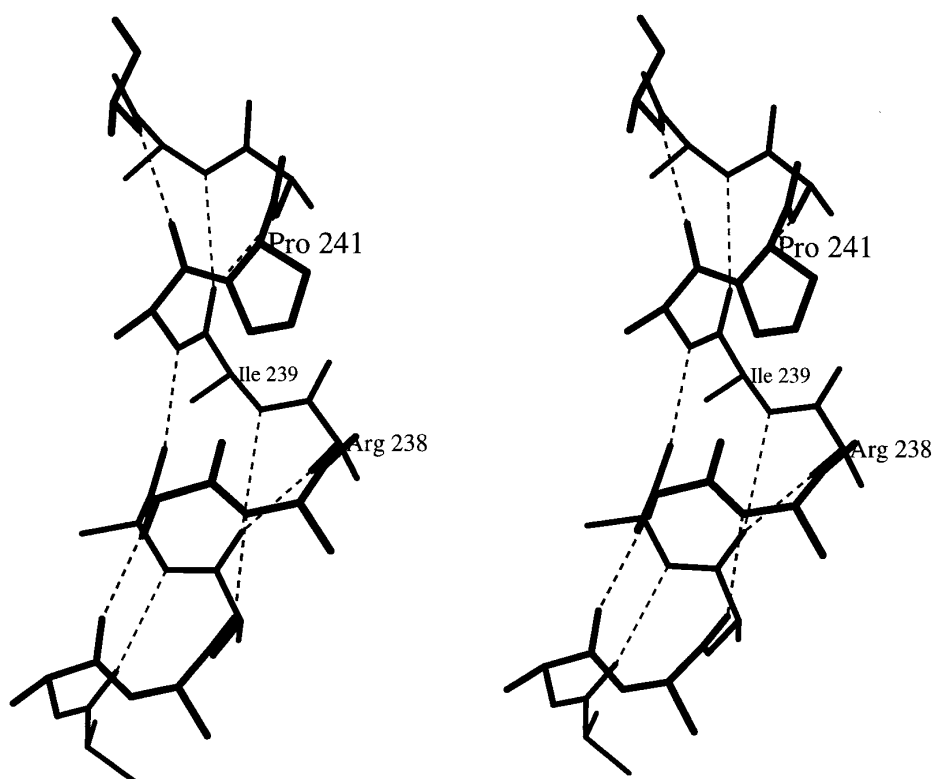


Figure 3. Main-chain hydrogen bonds in the distorted part of helix $\alpha 7$.

for 255 equivalent C^α pairs. Despite the low level of sequence identity, the structures of narbonin, chitinase A, endo F₁ and endo H superimposed well on hevamine, with RMS differences of 2.1 Å (185 C^α pairs), 2.0 Å (164 C^α pairs), 2.1 Å (161 C^α pairs) and 1.9 Å (153 atom pairs), respectively. An alignment of the sequences, based on the superposition of their three-dimensional structures, is shown in Figure 6.

On the basis of their three-dimensional structures, Farber & Petsko (1990) defined four families of $(\beta\alpha)_8$ barrel proteins. Narbonin (Reardon & Farber, 1995), concanavalin B and hevamine can be grouped in family C, whose members share a single domain structure with an extra helix $\alpha 8'$ and have the major barrel axis near $\beta 3$. Of the members of family C, hevamine superimposes best on the α -subunit of tryptophan synthase (Hyde *et al.*, 1988), with an RMS difference of 2.4 Å for 130 equivalent C^α pairs. However, the other family 18 structures do not fit in this family: chitinase A because it has additional domains, and the endo-*N*-acetylglucosaminidases because they lack helix $\alpha 8'$. Moreover, there is no functional relationship between the family 18 enzymes and the members of the Farber & Petsko family C, which use the helix dipole of $\alpha 8'$ to bind phosphate, and do not have any glycosyl hydrolase activity. The family 18 proteins should all belong to the same structural $(\beta\alpha)_8$ barrel family, which means that some of the criteria used by Farber & Petsko (1990), such as the number and location of extra domains, helices or

β -sheets, are not universally valid for classifying protein structures.

Conserved residues in the first family 18 consensus motif

In each of the six structures the two family 18 consensus regions (residues 73 to 80 and 119 to 127 in hevamine numbering) roughly correspond to strands $\beta 3$ and $\beta 4$. Although these regions are highly conserved, not one residue is strictly conserved in all known family 18 sequences, and only a few in those of the six proteins with known structures (Figure 6). These are Lys73, Gly80, Asp120 and Gly121 (hevamine numbering). Hydrogen bonds formed by the consensus region residues, except for the β -barrel hydrogen bonds, are shown in Table 5. The hydrophobic residues in the consensus region are not conserved, although they remain hydrophobic in all structures, which reflects the aspecificity of hydrophobic interactions.

The first consensus region begins in hevamine with Lys73. Why Lys at this position is highly conserved is not clear. In all six structures the lysine is the first residue of $\beta 3$. The side-chain has an extended conformation and points towards the solvent. In all cases the N^ϵ atom makes no hydrogen bonds with protein atoms, and the van der Waals interactions differ in the various structures. In all structures the N^ϵ atom is between 4.5 and 7.0 Å from the carboxyl group of the conserved aspartate residue (Asp120 for hevamine), and it might be that

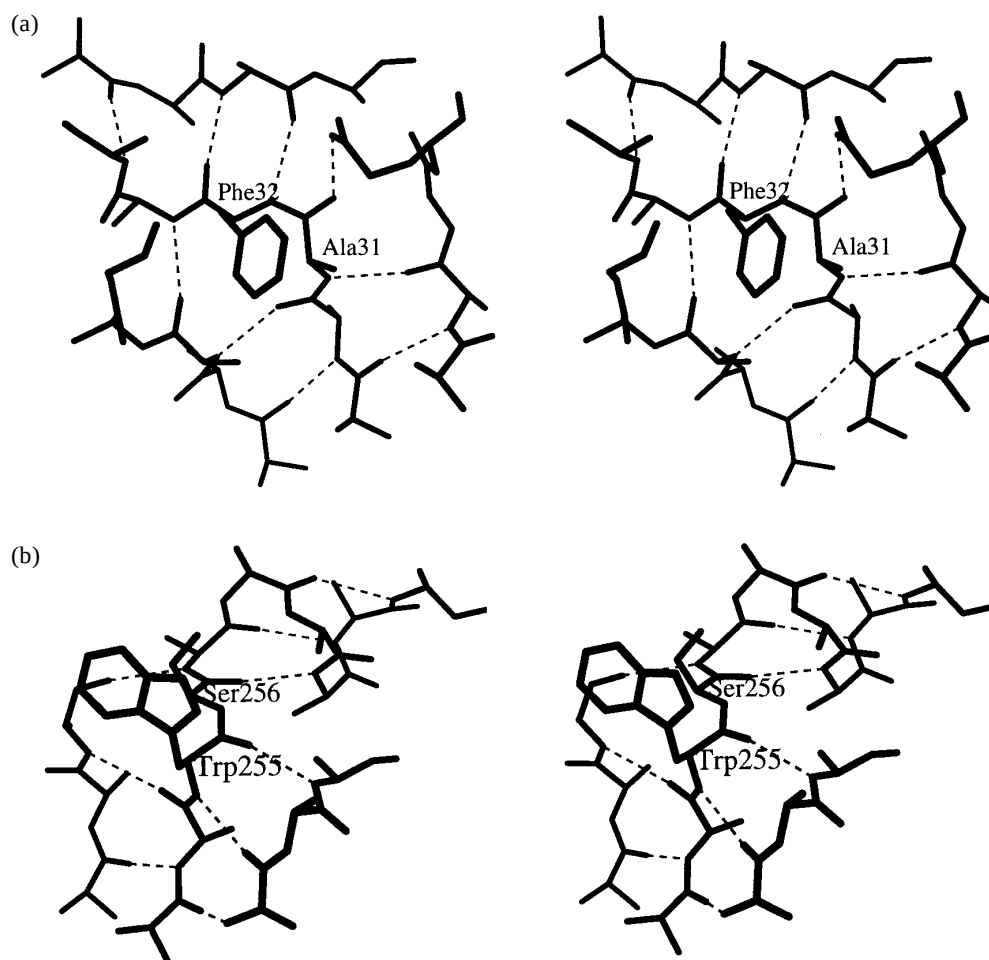


Figure 4. The *cis*-peptides found in the hevimine structure. (a) Ala31 and Phe32 are located in $\beta 2$. The main-chain nitrogen atom of Ala31 forms hydrogen-bonding contacts with the main-chain oxygen of Gln9, that is located in $\beta 1$, whereas the main-chain nitrogen atom of residue 32 forms hydrogen-bonding contacts with the main-chain oxygen atom of Asn45, in strand $\beta 2'$. (b) Trp255 is the last residue of $\beta 8$, and its main-chain oxygen atom hydrogen bonds the main-chain nitrogen atom of Ala220. Ser256 is the first residue of $\alpha 8'$.

these two charges interact during the folding process of these proteins.

Gly80 is the last residue of $\beta 3$. In hevimine, concanavalin B, narbonin and chitinase A, Gly79 is also conserved. Both glycine residues have ϕ, ψ angles that are not allowed for non-glycine residues, although they do form β -sheet hydrogen bonds with the neighbouring strands. In the four structures, Gly80 is close to a phenylalanine residue at position 32. A residue other than glycine at position 80 would not fit without major rearrangements in the β -barrel. In contrast, in both endo F_1 and endo H the residues equivalent to Gly79 and Gly80 have ϕ, ψ angles in the most favoured regions for β -strands, and the residue equivalent to Gly79 is leucine (Leu92 and Leu93, respectively). The leucine side-chains point in the direction where in the other structures the main-chain is positioned, causing a drastic change in the conformation of loop $\beta 3\alpha 3$. In the hevimine structure, this loop forms part of one side of the substrate-binding cleft, with residues Gly79, Gly81 and Ile82 interacting with the substrate at subsite -2. The conformation

of this loop is similar in the structures of concanavalin B, narbonin and chitinase A. In the structure of endo F_1 and endo H, this loop is replaced by a much shorter loop formed by residues Leu92, Gly93 and Asn94, and Leu91, Gly92 and Asn93, respectively. In endo F_1 and endo H, in contrast to the chitinases, the substrate does not consist of linear $\beta(1-4)$ linked sugar residues; the minimum required substrate is the branched $\text{Man}\alpha(1-3)\text{Man}\alpha(1-6)[\text{Man}\alpha(1-3)]\text{Man}\beta(1-4)\text{GlcNAc}\beta(1-4)\text{GlcNAc}$ (van Roey *et al.*, 1994). The $\beta(1,4)$ linkages are rigid, with hydrogen-bonding interactions between the O-5 atom of one and the O-3 atom of the following sugar residue, whereas the three α -linked mannose residues are more flexible. This means that the β -linked mannose and two *N*-acetylglucosamine residues are in a relative conformation that is similar to the conformation of *N*-acetylglucosamine residues in chitin, and could bind at subsites in a similar way as in hevimine. Cleavage takes place between the two *N*-acetylglucosamine residues, and as Glu132, which is equivalent to the catalytic Glu of hevimine, is

Table 3. Atoms of the chitotriose, chitohexaose and allosamidin models closer than 1.0 Å to a water molecule of the 1.8 Å structure

	Distance to water molecule (Å)	Hevamine hydrogen-bonding partners of water molecule
A. Chitotriose atom		
−4:O-7	0.4	Gly11 N Gly48 N
−2:O-7	0.8	Trp255 N ^{ε1}
B. Allosamidin atom		
−1:N-2	0.3	Asp125 O ^{δ2}
−1:O-3	0.7	Gly81 N
−1:O-7	0.4	Tyr183 O ⁿ
C. Chitohexaose atom		
+1:O-6	0.80	Gln181 O ^{ε1} Tyr183 O
+1:O-4	0.28	Gln158 O ^{ε1} Asn184 O
+2:N-7	0.53	Asn185 O ^{δ1}

Chitotriose and allosamidin coordinates were obtained from crystallographic complexes (Terwisscha van Scheltinga *et al.*, 1994, 1995), the chitohexaose was modelled on the basis of the chitotriose coordinates (Terwisscha van Scheltinga *et al.*, 1994).

proposed to be essential for catalytic activity (van Roey *et al.*, 1994), it is reasonable to assume that these two sugars bind at positions similar to subsites −1 and +1 of hevamine. The mannose residue would then bind at a position similar to subsite −2 of hevamine, and as the substrate is branched at this mannose moiety, the shape of the subsite should be very different from that of hevamine. Indeed, in the endo F₁ structure 3β α 3, which forms part of subsite −2 of hevamine, has a conformation that is very different from the conformations in the other structures, allowing accommodation of the carbohydrate chains that are branched at the central mannose residue.

Conserved residues in the second family 18 consensus motif

In the second consensus region, Asp120 is the first residue of strand β4. It makes two hydrogen

bonds that are conserved in all six structures; one hydrogen bond is between its main-chain nitrogen atom and the main-chain carbonyl oxygen atom of residue 74, the other hydrogen bond is between O^{δ2} and the main-chain nitrogen atom of residue 74. The latter prevents a regular parallel β-barrel hydrogen-bonding pattern, in which there would have been a hydrogen bond between N-74 and the carbonyl oxygen atom of residue 119. This irregularity is compensated for by Gly121, which has ϕ, ψ -angles mostly outside the regions allowed for non-glycine residues, and allows the β-barrel to continue. As both Asp120 and Gly121 are fully conserved in all family 18 sequences, the irregularity of the β-barrel in this region is probably characteristic for the whole family.

The side-chain of Asp123 is at the centre of the (β α)₈ barrel. For hevamine, this residue appears to be crucial for maintaining the fold, as its carboxylate group hydrogen bonds to side-chains of residues from four different β-strands: Tyr6 (β1), Ser77 (β3), Thr154 (β5) and Trp179 (β6). These interactions, however, are not fully conserved in all structures. Although Asp123 is conserved in all known family 18 chitinase sequences, in concanavalin B and endo F₁ it is replaced by histidine and phenylalanine, respectively. In concanavalin B, the histidine side-chain makes two hydrogen bonds, equivalent to those of Asp123 of hevamine with Tyr6 and Thr154, so that it is fully engaged in hydrogen-bonding. The residues at positions equivalent to Ser77 and Trp179, both fully conserved in all family 18 plant chitinase sequences, are alanine and phenylalanine, respectively. For endo F₁, the Phe128 side-chain stacks on the aromatic side-chains of Phe15, which is equivalent to Tyr6, and Tyr171. Surprisingly, endo H, which shares 32% sequence identity with endo F₁, has aspartate instead of phenylalanine at position 128, which makes interactions identical with Asp123 in hevamine.

Glu127 is essential for catalytic activity of family 18 chitinases. In the hevamine structure its carboxylate group is at hydrogen-bonding distance from the carboxylate group of Asp125, which is also

Table 4. Hydrogen bonds involved in crystal contacts

Molecule 1	Solvent-accessible surface (Å ²) of the amino acid residue		Molecule 2	Solvent-accessible surface (Å ²) of the amino acid residue		Distance (Å)
	Accessible in the crystal	Buried by crystal contacts		Accessible in the crystal	Buried by crystal contacts	
Asn12 O	121	17	Lys146 N ^ε	84	111	3.5
Ser21 O	39	40	Tyr140 O ⁿ	25	41	2.8
Arg23 N ⁿ¹	58	144	Gln90 O ^{ε1}	37	64	2.9
Gly39 O	11	65	Thr168 O ^{γ1}	35	50	2.9
Gly39 O	11	65	Asn171 O ^{δ1}	62	51	3.3
Thr41 O ^{γ1}	53	55	Thr168 O ^{γ1}	35	50	3.5
Ile69 O	104	35	Asn197 N ^{δ2}	20	35	3.0
Gly81 O	24	23	Lys244 N ^ε	39	45	2.6
Ala88 O	24	28	Arg204 N ⁿ¹	69	53	3.2
Ser89 O ^γ	4	35	Phe161 O	9	56	2.6
Tyr165 O ⁿ	21	27	Asp271 O ^{δ2}	63	99	2.7
Lys213 N ^ε	20	64	Glu223 O	105	24	3.5

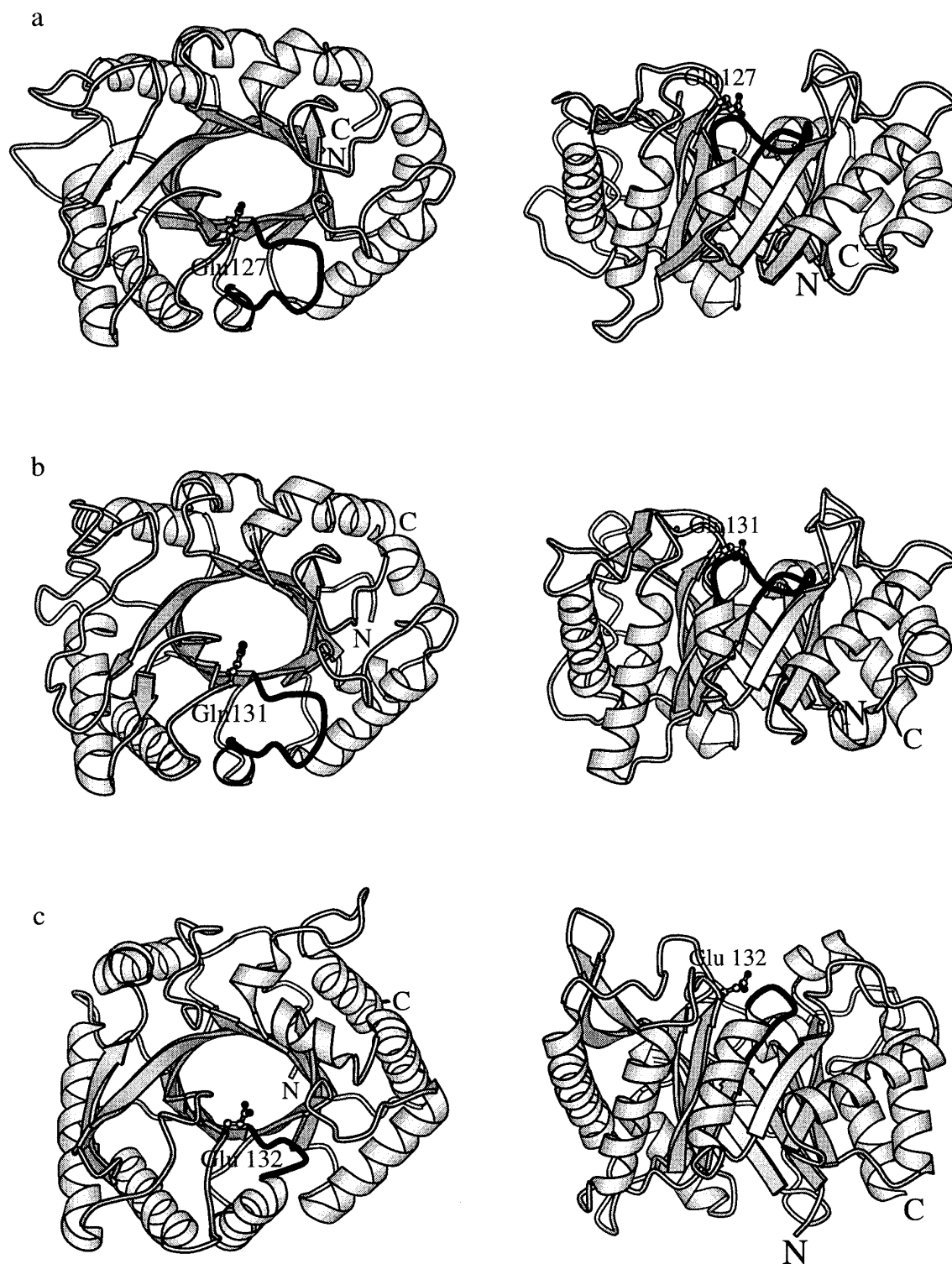


Figure 5. (a-c)

the case for the endo-*N*-acetylglucosaminidases. In chitinase A, however, the corresponding side-chain of Asp313, instead of pointing up to contact the catalytic glutamic acid residue, points down to form a hydrogen bond with Asp311, as well as with Oⁿ of the equivalent of Tyr6 in hevamine. In

narbonin, His130 is found at the position equivalent to Asp125, and its N^{e2} atom contacts O^{e1} of Glu132 through a water molecule. In concanavalin B, the glutamic acid is replaced by a glutamine residue, in accordance with the absence of any chitinase activity for concanavalin B (Hennig *et al.*, 1995a).

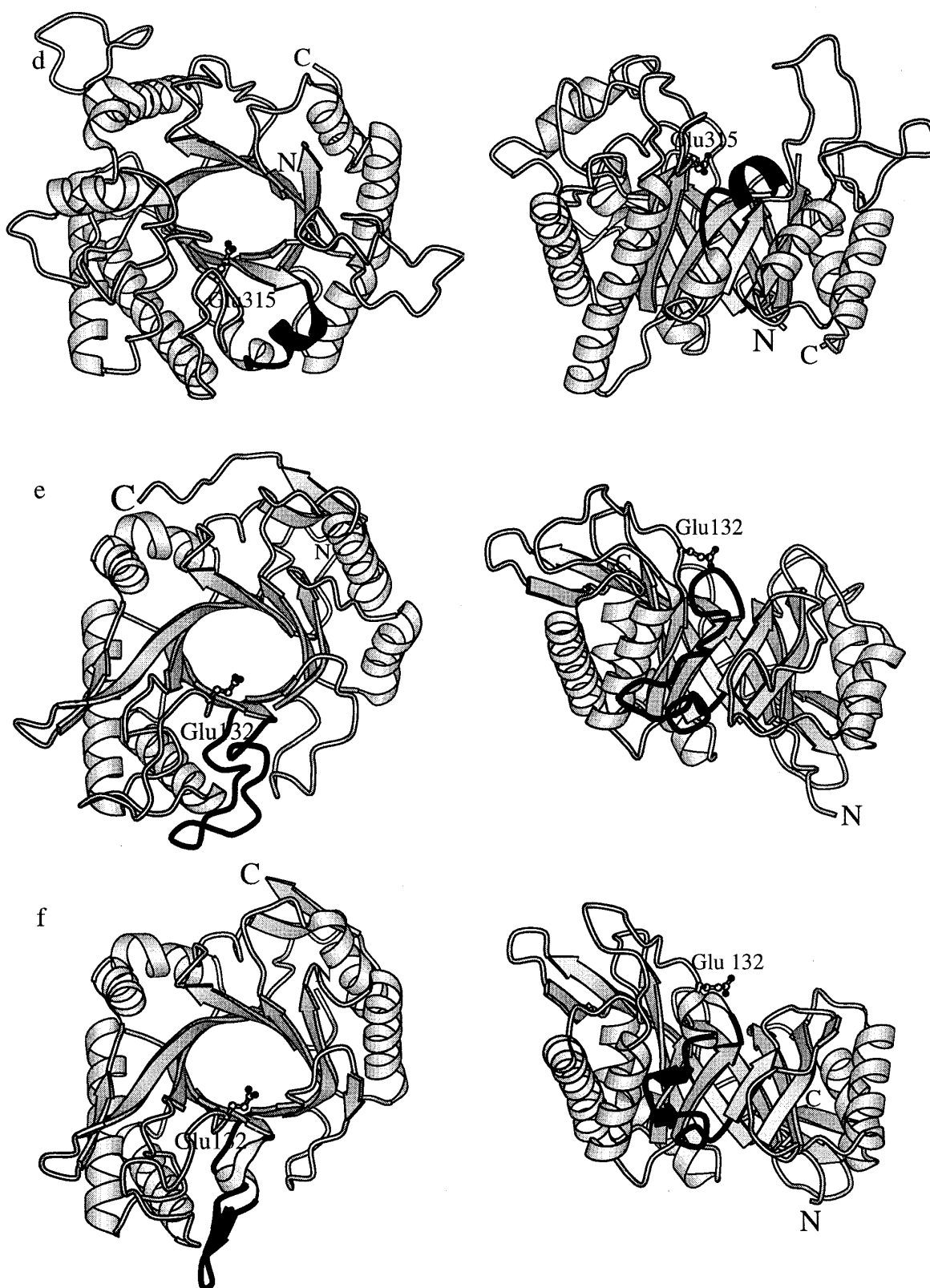


Figure 5 (d-f).

Figure 5. Ribbon plots of the six family 18 structures, created using MOLSCRIPT (Kraulis, 1991). The right-hand panels are rotated 90° with respect to the left-hand panels. (a) Hevamine, with Glu127 in ball-and-stick representation, and loop $\beta 5\alpha 5$ colored black. (b) Concanavalin B, with Gln131 in ball-and-stick representation, and loop $\beta 5\alpha 5$ colored black. (c) Narbonin, with Glu132 in ball-and-stick representation, and loop $\beta 5\alpha 5$ colored black. (d) The catalytic domain of chitinase A, with Glu315 in ball-and-stick representation, and loop $\beta 5\alpha 5$ colored black. (e) Endo F₁, with Glu132 in ball-and-stick representation, and loop $\beta 5\beta 6$ colored black. (f) Endo H, with Glu132 in ball-and-stick representation, and loop $\beta 5\beta 6$ colored black.

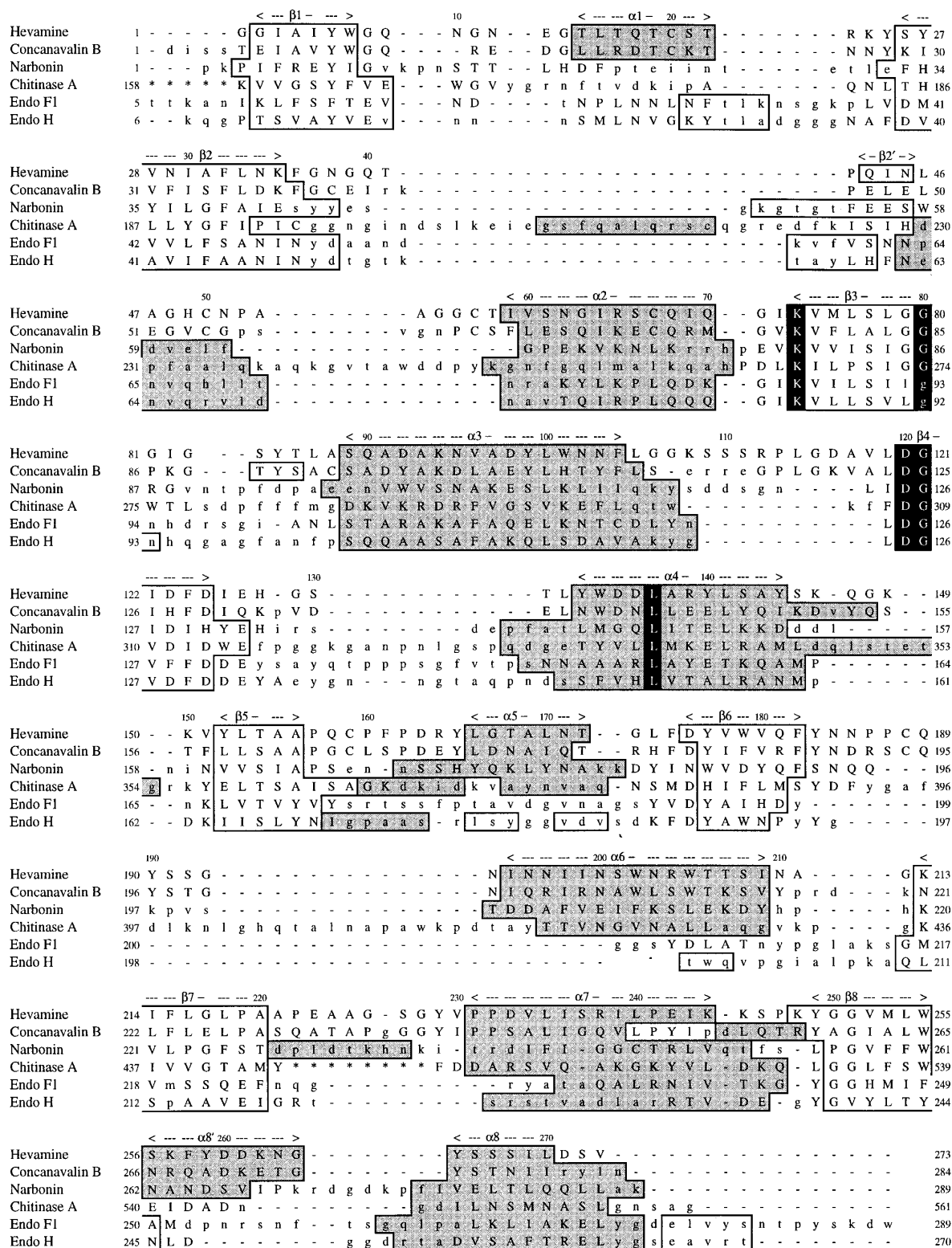


Figure 6. Structural alignment to the sequence of hevamine, based on the automated alignment procedure in O, and improved by hand. When the pairs of C α atoms fit better than 3.8 Å, the residue names are shown in capitals. The two additional domains of chitinase A are omitted from the comparison, and are indicated by asterisks. Shaded boxes indicate α -helices, non-shaded boxes β -strands. The residues that are fully conserved for all six proteins are shown in white, with black shading. The first and last residues on each line are labelled, as well as every tenth residue of the hevamine sequence. The amino-terminal 23 residues of chitinase A form a leader peptide, and are not part of the mature protein. The carboxy-terminal 15 residues of concanavalin B, the carboxy-terminal residue of narbonin, the amino-terminal four residues of endo F1, the amino-terminal five residues and carboxy-terminal residue of endo H and the carboxy-terminal two residues of chitinase A are not part of the models, as they are not visible in electron density. These residues are not included in the alignment.

Table 5. Non- β -structure hydrogen bonds of the consensus regions

Protein	Residue	Atom	Contact
Hevamine	Asp120	N O ^{$\delta 2$} O ^{$\delta 2$}	Val74 O Val74 N Gln68 N ^{$\epsilon 2$}
Concanavalin B	Asp124	N O ^{$\delta 2$} O ^{$\delta 2$}	Val79 O Val79 N Gln73 N ^{$\epsilon 2$}
Narbonin	Asp125	N O ^{$\delta 2$}	Val80 O Val80 N
Chitinase A	Asp308	N O ^{$\delta 2$} O ^{$\delta 2$} O ^{$\delta 1$}	Ile268 O Ile268 N Lys260 N ^{ζ} Lys260 N ^{ζ}
Endo F ₁	Asp125	O ^{$\delta 1$} N O ^{$\delta 2$} O ^{$\delta 2$}	Arg355 N ^{$\eta 2$} Val87 O Val87 N Gln81 N ^{$\epsilon 2$} /O ^{$\epsilon 1$}
Endo H	Asp125	N O ^{$\delta 2$} O ^{$\delta 2$}	Val86 O Val86 N Gln80 N ^{$\epsilon 2$}
Hevamine	Asp123	O ^{$\delta 1$} O ^{$\delta 1$} O ^{$\delta 2$} O ^{$\delta 2$}	Thr154 O ^{$\gamma 1$} Trp179 N ^{$\epsilon 1$} Tyr6 O ^{η} Ser77 O ^{γ}
Concanavalin B	His127	N ^{$\delta 1$} N ^{$\epsilon 2$}	Ser160 O ^{γ} Tyr10 O ^{η}
Narbonin	Asp218	O ^{$\delta 1$} O ^{$\delta 2$} O ^{$\delta 2$}	Ser163 O ^{γ} Tyr8 O ^{η} Ser83 O ^{γ}
Chitinase A	Asp311	O ^{$\delta 1$} O ^{$\delta 2$} O ^{$\delta 2$}	Thr360 O ^{$\gamma 1$} Ser271 O ^{γ} Asp313 O ^{$\delta 1$}
Endo F ₁	Phe128	—	—
Endo H	Asp128	O ^{$\delta 1$} O ^{$\delta 1$} O ^{$\delta 2$} O ^{$\delta 2$}	Ser166 O ^{γ} Trp192 N ^{$\epsilon 1$} Tyr14 O ^{η} Ser89 O ^{γ}
Hevamine	Asp125	O ^{$\delta 1$} O ^{$\delta 1$}	Glu127 O ^{$\epsilon 2$} Gln181 N ^{$\epsilon 2$}
Concanavalin B	Asp129	O ^{$\delta 1$} O ^{$\delta 2$}	Gln131 N ^{$\epsilon 2$} Tyr10 O ^{η}
Narbonin	His130	N ^{$\delta 1$}	Tyr8 O ^{η}
Chitinase A	Asp313	O ^{$\delta 1$} O ^{$\delta 2$}	Asp311 O ^{$\delta 2$} Tyr163 O ^{η}
Endo F ₁	Asp130	O ^{$\delta 1$}	Glu132 O ^{$\epsilon 2$}
Endo H	Asp130	O ^{$\delta 1$}	Glu132 O ^{$\epsilon 1$}
Hevamine	Glu127	O ^{$\epsilon 2$} O ^{$\epsilon 2$}	Asp125 O ^{$\delta 1$} Gln181 N ^{$\epsilon 2$}
Concanavalin B	Gln131	N ^{$\epsilon 2$}	Asp129 O ^{$\delta 1$}
Narbonin	Glu132	O ^{$\epsilon 1$}	Arg87 N ^{ϵ}
Chitinase A	Glu315	—	—
Endo F ₁	Glu132	O ^{$\epsilon 2$}	Asp130 O ^{$\delta 1$}
Endo H	Glu132	O ^{$\epsilon 1$}	Asp130 O ^{$\delta 1$}

Additional conserved features

Apart from the family 18 consensus motifs, a few other features are conserved. In all six structures, non-proline *cis*-peptides occur. A *cis*-peptide similar to that found between Ala31 and Phe32 in hevamine is present in all structures. For concanavalin B, narbonin and chitinase A, there is a residue with a small side-chain (Gly, Ala or Ser) at position 31, and Phe at position 32. Although for endo F₁ and endo H the positions of the Phe residue and the small residue have interchanged, the Phe side-chain occupies approximately the same space in all structures. The *cis*-peptide

between Trp255 and Ser256 is conserved in concanavalin B, narbonin and chitinase A. Trp at position 255 forms the end of $\beta 8$, with the next residue starting the inserted helix $\alpha 8'$. In contrast to the other proteins, the endo-*N*-acetylglucosaminidases do not have the extra helix $\alpha 8'$, and the *cis*-peptide at the end of $\beta 8$ is not conserved. However, both enzymes do have an aromatic residue at position 255, with the side-chain oriented in a way similar to the tryptophan side-chains in the other structures. As the two aromatic residues at positions 32 and 255 participate in substrate binding for hevamine (Terwisscha van Scheltinga *et al.*, 1994), it seems that both concanavalin B and narbonin contain reminiscences of a chitin binding site.

Chitinase A has a third non-proline *cis*-peptide, between the catalytic Glu315 and Phe316. The nitrogen and oxygen atoms of the peptide are not involved in hydrogen bonds, and the peptide might be forced in a *cis*-conformation by the buried phenylalanine side-chain.

Functional importance of structural differences

The loops connecting $\beta 5$ with $\alpha 5$ (Figure 5) reflect the different functions of the four enzymes. For hevamine, which shows endochitinase activity, $\beta 5\alpha 5$ is a rigid, highly structured loop. It consists of nine residues, of which one is a cysteine residue involved in a disulphide bridge, and three are proline, among which is *cis*-Pro162. The loop is at one end of the substrate-binding cleft, forming part of subsite +2. As the loop is compact, leaving the substrate-binding cleft with an open end, the bound substrate would be allowed to extend into the solvent. In the structure of chitinase A, which mainly produces chitobiose (Roberts & Cabib, 1982), $\beta 5\alpha 5$ does not contain proline or cysteine residues and accommodates a small α -helix. The loop is more extended than in the hevamine structure and could obstruct substrates that extend beyond subsite +2. As cleavage takes place between site -1 and +1, this would explain the formation of mainly chitobiose as the product. For endo F₁, the loop has quite a different orientation, such that there is no subsite +2, and a sugar residue at this position would stick into the solvent. Endo F₁ cleaves between two *N*-acetylglucosamine residues, of which one is linked to a protein through an asparagine side-chain. As cleavage occurs between the subsites -1 and +1, the two *N*-acetylglucosamine residues should bind at these subsites, without further need for a subsite +2.

The lack of chitinase activity of concanavalin B and narbonin might be explained by substitutions of residues that are located in the active site of hevamine. In concanavalin B the position equivalent to Glu127, the putative proton donor of hevamine, is occupied by Gln131. In narbonin, the position equivalent to Asp125, proposed to stabilize

Table 6. Possible hydrogen-bonding interactions with chitinase inhibitors

Inhibitor	Hevamine	Concanavalin B	Narbonin
–4:O-3		Arg14 N ^{η1}	
–4:O-8	Gly11 N Gly48 N Gln9 O ^{ε1}	Arg14 N	
–3:N-2			Asn14 O ^{δ1}
–3:O-6			
–3:O-7	Asn34 N ^{δ2}	Asp37 O ^{δ1} (?)	
–2:O-6	Ile82 N	Lys87	Arg87 N
–2:O-7	Trp255 N ^{ε1}	Trp265 N ^{ε1}	Lys12 N ^ζ
–1:N-2	Asp125 O ^{δ2}		
–1:O-3	Gly81 N		
–1:O-6	Ala224 O	Thr232 O	
–1:O-8	Tyr183 O ^η	Tyr189 O ^η	

The inhibitor coordinates at subsite –4 are taken from the hevamine-chitotriose complex, at the subsites –3, –2 and –1 from the hevamine-allosamidin complex.

the positively charged oxazoline reaction intermediate (Terwisscha van Scheltinga *et al.*, 1995), is occupied by His130. Furthermore, several residues in concanavalin B and narbonin would clash with a chitin-like substrate. To analyse this in more detail, the protein structures of concanavalin B and narbonin were superimposed on hevamine complexed with chitotriose (Terwisscha van Scheltinga *et al.*, 1994) or allosamidin (Terwisscha van Scheltinga *et al.*, 1995). Atoms that are, after superimposing, within hydrogen-bonding distance from the inhibitors are given in Table 6. Large parts of the hevamine chitin-binding site are distorted in the two other proteins.

At subsite –4 Ala47 of hevamine is replaced in concanavalin B by Glu51, which would collide with the sugar residue. At subsite –1 Gly81 of hevamine, of which the main-chain nitrogen atom hydrogen bonds O-3 of the *N*-acetylglucosamine analogue of allosamidin, is replaced by Pro86. This rules out the possibility of concanavalin B to hydrogen bond to O-3. Moreover, the side-chain of the proline residue fills up part of the binding site.

For narbonin, a large part of subsite –4 is taken by the side-chain and main-chain of Asn14. The most striking disruption of the cleft, however, is caused by the replacement of Gly81 and Ala225 of hevamine by Arg87 and Asp231 of narbonin. These two residues, located in loop $\beta 3\alpha 3$ and loop $\beta 7\alpha 7$, respectively, are at opposite sides of the cleft and form a salt bridge, resulting in a complete obstruction of subsite –1. Moreover, the $\beta 5\alpha 5$ loop of narbonin differs drastically from those of hevamine and concanavalin B. For hevamine, this loop consists of nine residues, including one half-cystine and three proline residues, one of which is *cis*-Pro162. The loop is part of subsite +2, and forms one of the open ends of the substrate-binding cleft. The loop of concanavalin B is very similar (Figure 5). The $\beta 5\alpha 5$ loop of narbonin, however, consists of only four residues, and does not contain either the half-cystine or the *cis*-Pro observed in hevamine and concanavalin B. Instead of forming an open end of the cleft, it continues

upwards and, together with the first residues of helix $\alpha 5$, is likely to block subsite +2.

These results suggest that concanavalin B and narbonin cannot bind chitin in a fashion similar to hevamine. This is confirmed by the soaking of concanavalin B crystals in solution containing chitotriose, which does not yield electron density for a ligand (unpublished results) and unsuccessful search for chitinase activity under a wide range of conditions (Schlesier *et al.*, 1996).

Conclusions

Many interactions of the family 18 consensus regions are highly conserved, but only a few are fully conserved. Most important, Asp120 and Gly121 allow the irregular contact between the amino-terminal ends of strands $\beta 3$ and $\beta 4$. Asp125 and Glu127 are important for catalytic activity and, as such, conserved in the enzymatically active proteins.

Although the sequences of the six proteins compared generally have a very low identity level, the cores of the structures superimpose very well. The two conserved regions have similar positions in all structures, and appear to be important both for stability of the fold and for catalytic activity. The shape of the substrate-binding cleft of hevamine has been drastically altered in the concanavalin B and narbonin structures, preventing tight binding of chitin as observed for hevamine. The $\beta x\alpha x$ loops of the other enzymes, chitinase A, endo F₁ and endo H, differ to accommodate different substrates and to perform different reactions. Although the functions of concanavalin B and narbonin are not known, all six structures show a nice example of very distantly related proteins, which have evolved quite differently to adapt to their particular functions.

Methods

Crystallization and data collection

Crystals of hevamine were grown by the method of vapour diffusion (Rozeboom *et al.*, 1990). They grow in space group $P2_12_12_1$, with cell dimensions $a = 52.30$ Å, $b = 57.72$ Å and $c = 82.05$ Å, and with one molecule per asymmetric unit. The crystals are very stable, and can be transferred directly to synthetic mother liquors containing 20% (w/v) NaCl, 23% (w/w) PEG 6000 or 50% (w/w) ammonium sulphate at a pH ranging from 2.0 to 9.5, without causing any change in cell dimensions or loss in stability or resolution.

Two diffraction data sets were collected from a single crystal on a MAR Research imaging plate at room temperature at the synchrotron beam line X11 at DESY (Hamburg), using a wavelength of 0.92 Å: a high-resolution data set (1.8 Å, 97 frames with a 1.25 oscillation) and a low-resolution data set (3.2 Å, 57 frames with a 2.5 oscillation). Data were processed using the program MOSFLM (A.G.W. Leslie, MRC, Cambridge, UK) and CCP4 software, and the two data sets were merged using the Groningen BIOMOL software. Statistics of the data collection can be found in Table 1.

Refinement

Refinement at 1.8 Å was started with the 2.2 Å hevamine model (Terwisscha van Scheltinga *et al.*, 1994), leaving out the water molecules to prevent any bias in assigning their positions. Positional and isotropic *B*-factor refinement, using all data between 8.0 and 1.8 Å with $F > 0$, was initially carried out with the program X-PLOR (Brünger *et al.*, 1987), and in a later stage with the program TNT (Tronrud *et al.*, 1987). For both programs the stereochemical restraints derived by Engh and Huber (1991) were used: 10% of the diffraction data, chosen randomly as a test set, were omitted from the refinement, and were used to calculate a free *R*-factor (R_{free} ; Brünger, 1992). Refinement weights were varied to yield a low R_{free} and good geometry. Parts of the structure with bad geometry were identified using the program PROCHECK (Laskowski *et al.*, 1993), after which the model was rebuilt in σ_A -weighted (Read, 1986) $F_o - F_c$ and $2F_o - F_c$ electron density maps with O (Jones *et al.*, 1991). Regions with ill-defined electron density were excluded from the model, and rebuilt after additional positional refinement using OMIT maps. Water molecules were fitted in peaks of $F_o - F_c$ electron density larger than 3.0σ , under the conditions that at least one hydrogen bond should be formed with the protein or another water molecule, and that non-hydrogen-bonding atoms should not be closer than 2.8 Å. After a refinement cycle water molecules with *B*-values higher than five times the average *B*-value of the protein atoms, or with $2F_o - F_c$ electron density below 1σ , were deleted. When refinement and rebuilding had converged, a final TNT cycle using all data was carried out. Crystallographic structure factors and atomic coordinates are available with ID code 2HVM at the Brookhaven protein data bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973.

Structure analysis

Secondary structure assignments and solvent-accessible surfaces were obtained from the program DSSP (Kabsch & Sander, 1983). A residue is defined as being part of an α -helix or β -strand if it is involved with at least one hydrogen bond in the hydrogen-bonding pattern of the secondary structure element.

The structures of concanavalin B (Hennig *et al.*, 1995a; PDB entry 1CNV), narbonin (Hennig *et al.*, 1995b; PDB entry 1NAR), chitinase A (Perrakis *et al.*, 1994; PDB entry 1CTN), endo F_1 (van Roey *et al.*, 1994; PDB entry 2EBN), endo H (Rao *et al.*, 1995; PDB entry 1EDT) and the α -subunit of tryptophan synthase (Hyde *et al.*, 1988; PDB entry 1WSY) were superimposed on the hevamine model using the option of the program O (Jones *et al.*, 1991) to locate the longest matching fragments between two molecules, where each pair of C^α atoms in a fragment should fit better than 3.8 Å. The assignment of equivalent C^α pairs was checked and improved by hand, using the information of the superimposed three-dimensional structures.

Acknowledgements

We thank the staff of the EMBL outstation at DESY, Hamburg, for their help during data collection.

References

- Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M. & Possee, R. D. (1994). The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology*, **202**, 586–605.
- Bhat, T. N. (1988). Calculation of an OMIT map. *J. Appl. Crystallog.* **21**, 279–281.
- Boot, R. G., Renkema, G. H., Strijland, A., van Zonneveld, A. J. & Aerts, J. M. F. G. (1995). Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. *J. Biol. Chem.* **270**, 26252–26256.
- Brünger, A. T. (1992). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic *R* factor refinement by molecular dynamics. *Science*, **235**, 458–460.
- Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallog. sect. A*, **47**, 392–400.
- Farber, G. K. & Petsko, G. A. (1990). The evolution of α/β barrel enzymes. *Trends Biochem. Sci.* **15**, 228–234.
- Hart, P. J., Monzingo, A. F., Ready, M. P., Ernst, S. R. & Robertus, J. D. (1993). Crystal structure of an endochitinase from *Hordeum vulgare* L. seeds. *J. Mol. Biol.* **229**, 189–193.
- Hennig, M., Jansonius, J. N., Terwisscha van Scheltinga, A. C., Dijkstra, B. W. & Schlesier, B. (1995a). Crystal structure of concanavalin B at 1.65 Å resolution. An “inactivated” chitinase from seeds of *Canavalia ensiformis*. *J. Mol. Biol.* **254**, 237–246.
- Hennig, M., Schlesier, B., Pfeffer-Hennig, S., Dauter, Z., Nong, V. H. & Wilson, K. S. (1995b). Crystal structure of narbonin at 1.8 Å resolution. *Acta Crystallog. sect. D*, **51**, 177–189.
- Henrissat, B. (1990). Weak sequence homologies among chitinases detected by clustering analysis. *Protein Seq. Data Anal.* **3**, 523–526.
- Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293**, 781–788.
- Holm, L. & Sander, C. (1994). Structural similarity of plant chitinase and lysozymes from animals and phage—an evolutionary connection. *FEBS Letters*, **340**, 129–132.
- Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W. & Davies, D. R. (1988). Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* **263**, 17857–17871.
- Jekel, P. A., Hartmann, J. B. H. & Beintema, J. J. (1991). The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. *Eur. J. Biochem.* **200**, 123–130.
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
- Kauffmann, S., Legrand, M., Geoffroy, P. & Fritig, B. (1987). Biological function of ‘pathogenesis-related’ proteins: four PR proteins of tobacco have 1,3- β -glucanase activity. *EMBO J.* **6**, 3209–3212.

- Kramer, K. J., Corpuz, L., Choi, H. K. & Muthukrishnan, S. (1993). Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of *Manduca sexta*. *Insect Biochem. Mol. Biol.* **23**, 691–701.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
- Legrand, M., Kauffmann, S., Geoffroy, P. & Fritig, B. (1987). Biological function of pathogenesis-related proteins: four tobacco pathogenesis-related proteins are chitinases. *Proc. Natl Acad. Sci. USA*, **84**, 6750–6754.
- Luzzati, V. (1952). Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallog.* **5**, 802–810.
- Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S. & Vorgias, C. E. (1994). Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure*, **2**, 1169–1180.
- Peumans, W. J. & van Damme, E. J. M. (1995). Lectins as plant defense proteins. *Plant Physiol.* **109**, 347–352.
- Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–438.
- Rao, V. H., Guan, C. & van Roey, P. (1995). Crystal structure of endo- β -N-acetylglucosaminidase H at 1.9 Å resolution: active-site geometry and substrate recognition. *Structure*, **3**, 449–457.
- Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A*, **42**, 140–149.
- Reardon, D. & Farber, G. K. (1995). Protein motifs. 4. The structure and evolution of α/β barrel proteins. *FASEB J.* **9**, 497–503.
- Roberts, R. L. & Cabib, E. (1982). *Serratia marcescens* chitinase: one-step purification and use for the determination of chitin. *Anal. Biochem.* **127**, 402–412.
- Ross, A. F. (1961). Localized acquired resistance to plant virus infection in hypersensitive hosts. *Virology*, **14**, 329–339.
- Rozeboom, H. J., Budiani, A., Beintema, J. J. & Dijkstra, B. W. (1990). Crystallization of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. *J. Mol. Biol.* **212**, 441–443.
- Schlesier, B., Nong, V. H., Horstmann, C. & Hennig, M. (1996). Sequence analysis of concanavalin B from *Canavalia ensiformis* reveals homology to chitinases. *J. Plant Physiol.* **147**, 665–674.
- Tata, S. J., Beintema, J. J. & Balabaskaran, S. (1983). The lysozyme of *Hevea brasiliensis* latex: isolation, purification, enzyme kinetics and a partial amino-acid sequence. *J. Rubb. Res. Inst. Malaysia*, **31**, 35–48.
- Terwisscha van Scheltinga, A. C., Kalk, K. H., Beintema, J. J. & Dijkstra, B. W. (1994). Crystal structures of hevamine, a plant defence protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure*, **2**, 1181–1189.
- Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B. & Dijkstra, B. W. (1995). Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and X-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. *Biochemistry*, **34**, 15619–15623.
- Tronrud, D. E., Ten Eyck, L. & Matthews, B. W. (1987). An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallog. sect. A*, **43**, 489–501.
- Tsujibo, H., Orikoshi, H., Imada, C., Okami, Y., Miyamoto, K. & Inamori, Y. (1993). Site-directed mutagenesis of chitinase from *Alteromonas sp.* strain O-7. *Biosci. Biotechnol. Biochem.* **57**, 1396–1397.
- van Loon, L. C. (1985). Pathogenesis-related proteins. *Plant Mol. Biol.* **4**, 111–116.
- van Roey, P., Rao, V., Plummer, T. H., Jr & Tarentino, A. L. (1994). Crystal structure of endo- β -N-acetylglucosaminidase F₁, an α/β -barrel enzyme adapted for a complex substrate. *Biochemistry*, **33**, 13989–13996.
- Watanabe, T., Kobori, K., Miyashita, K., Fujii, T., Sakai, H., Uchida, M. & Tanaka, H. (1993). Identification of glutamic acid-204 and aspartic acid-200 in chitinase-A1 of *Bacillus circulans* WL-12 as essential residues for chitinase activity. *J. Biol. Chem.* **268**, 18567–18572.

Edited by R. Huber

(Received 3 May 1996; received in revised form 2 July 1996; accepted 9 July 1996)