

JMB

The Refined Crystal Structure of an Endochitinase from *Hordeum vulgare* L. Seeds at 1.8 Å Resolution

P. John Hart, Heather D. Pflugger, Arthur F. Monzingo, Thomas Hollis and Jon D. Robertus*

Department of Chemistry and Biochemistry, University of Texas, Austin, TX 78712 U.S.A.

Class II chitinases (EC 3.2.1.14) are plant defense proteins. They hydrolyze chitin, an insoluble β -1,4-linked polymer of *N*-acetylglucosamine (NAG), which is a major cell-wall component of many fungal hyphae. We previously reported the three-dimensional structure of the 26 kDa class II endochitinase from barley seeds at 2.8 Å resolution, determined using multiple isomorphous replacement (MIR) methods. Here, we report the crystallographic refinement of this chitinase structure against data to 1.8 Å resolution using rounds of hand rebuilding coupled with molecular dynamics (X-PLOR). The final model has an *R*-value of 18.1% for the 5.0 to 1.8 Å data shell and 19.8% for the 10.0 to 1.8 Å shell, and root-mean-square deviations from standard bond lengths and angles of 0.017 Å and 2.88°, respectively. The 243 residue molecule has one β -sheet, ten α -helices and three disulfide bonds; 129 water molecules are included in the final model. We show structural comparisons confirming that chitinase secondary structure resembles lysozyme at the active site region. Based on substrate binding to lysozyme, we have built a hypothetical model for the binding of a hexasaccharide into the pronounced active site cleft of chitinase. This provides the first view of likely substrate interactions from this family of enzymes; the model is consistent with a lysozyme-like mechanism of action in which Glu67 acts as proton donor and Glu89 is likely to stabilize the transition state oxycarbonium ion. These binding site residues, and many hydrophobic residues are conserved in a range of plant chitinases. This endochitinase structure will serve as a model for other plant chitinases, and that catalytic models based on this structure will be applicable to the entire enzyme family.

Keywords: crystal structure; refinement; X-ray analysis; endochitinase; substrate binding

*Corresponding author

Introduction

Plants, in contrast to vertebrates, have no immune system and have developed other methods to defend themselves against pathogens such as viruses, bacteria and fungi. Plants under attack can induce genes that express a battery of defense proteins, known as pathogenesis-related (PR) proteins (Bowles *et al.*, 1991; Tasmussen *et al.*, 1992). These include: (1) enzymes that synthesize small molecular mass, antimicrobial compounds called phytoalexins

(Albersheim & Valent, 1978); (2) proteinase inhibitors (Walker-Simmons *et al.*, 1983); (3) ribosome inactivating proteins (RIPs: Roberts & Selitrennikoff, 1986); (4) non-catalytic, pore-forming polypeptides that insert themselves into fungal cell walls and dispel the cellular contents into the surrounding medium (Roberts & Selitrennikoff, 1990); and (5) lytic enzymes such as β -1,3 glucanases and chitinases (Abeles *et al.*, 1970; Boller *et al.*, 1983).

Chitin is a high tensile-strength, insoluble, linear β -1,4-linked polymer of *N*-acetylglucosamine (NAG). It is identical with cellulose except at the C-2 position, where it has an *N*-acetyl group instead of a hydroxyl group. Predominantly found in arthropod exoskeletons and fungal cell walls, it is one of the most abundant polysaccharides in nature. Plant lytic enzymes limit fungal growth by hydrolyzing β -glucan and chitin. Chitin comprises about 3 to 60% of cell-wall mass, depending on the type of fungus

Abbreviations used: PR, pathogenesis-related; RIP, ribosome-inactivating protein; NAG, *N*-acetylglucosamine; WGA, wheat germ agglutinin; MIR, multiple isomorphous replacement; HEWL and GEWL, hen and goose egg-white lysozyme; T4L, bacteriophage T4 lysozyme; PAP, pokeweed antiviral protein; NAM, *N*-acetyl muramic acid; r.m.s., root-mean-square.

(Wessels & Sietsma, 1981; Bartnicki-Garcia, 1968). The glycosidic (hydrolytic) fragments released by fungal cell-wall degradation act as elicitors of host stress metabolite (phytoalexin) biosynthesis (Walker-Simmons *et al.*, 1983).

Exochitinases, usually found in bacteria, hydrolyze chitin from the non-reducing end in a step-wise fashion (Roberts & Cabib, 1982). Endochitinases, found in a wide variety of plants, can hydrolyze the NAG polymer internally, releasing polysaccharides (Molano *et al.*, 1979; Boller *et al.*, 1983). Endochitinases, but not exochitinases, are effective in preventing the invasion of fungal mycelia (Roberts & Selitrennikoff, 1986, 1988; Leah *et al.*, 1991). Chitinases are most effective attacking the accessible nascent chitin fibers produced at the apex of growing hyphal tips of filamentous fungi (Mauch *et al.*, 1988; Roberts and Selitrennikoff, 1988). *In vivo* experiments demonstrated that transgenic tobacco plants constitutively expressing a bean endochitinase gene were better able to resist fungal infection than corresponding non-transformed plants (Broglie *et al.*, 1991).

Plant chitinases are monomeric proteins between 25 and 40 kDa. In a comprehensive classification of glycosyl hydrolases, Henrissat (1991) includes plant chitinases, along with those from fungal and bacterial sources, in groups 18 and 19. Others propose four classes of plant chitinases based on amino acid sequences, several of which may be present in the same plant (Shinshi *et al.*, 1990; Collinge *et al.*, 1993). Figure 1 shows the sequences of several representative chitinases and their domain structure. Class I, II and IV chitinases have homologous catalytic domains, of about 26 kDa. Class IV enzymes have three deletions within this main unit, but clearly belong to the same family (Collinge *et al.*, 1993). Class I and IV endochitinases also have an N-terminal cysteine-rich domain of 40 to 50 residues. This domain is homologous to wheat germ agglutinin (WGA), and presumably binds the enzyme to chitin. The three-dimensional structure of WGA has been determined (Wright, 1987). The N-terminal domain is linked to the chitinase domain by a glycine/proline-rich hinge segment. Class III chitinases show no sequence similarity to enzymes in class I, II or IV. They are relatively rare in higher plants, but common in fungi.

We previously reported the crystallization of the 26 kDa class II endochitinase from barley seeds (Hart *et al.*, 1992). A subsequent superior crystal form led to a preliminary structure at 2.8 Å resolution determined using multiple isomorphous replacement (MIR) methods (Hart *et al.*, 1993). The barley chitinase coordinates were deposited as entry 1BAA in the Protein Data Bank. Our analysis of the barley chitinase (Robertus *et al.*, 1995) shows it has elements of secondary structure that are conserved in all the known lysozyme structures, hen egg-white (HEWL), T4 phage (T4L), and goose egg-white (GEWL). This relationship was also detected by a computer search of the Protein Data Bank containing coordinates for barley chitinase (Holm & Sander, 1994). This finding

serves as a basis for the notion that barley chitinase may bind substrate and hydrolyze it by a mechanism similar to that of the well-studied lysozymes. Here, we report an overall description of the enzyme model based on refinement against 1.8 Å resolution data, discuss features of the model as the prototypical chitinase domain, and report the hypothetical binding and hydrolysis of a hexasaccharide based on lysozyme models. The refined chitinase coordinates have been deposited in the Protein Data Bank under access number 2BAA.

Results

High-resolution data collection and model refinement

The area detector was used to collect native data to a nominal resolution of 1.80 Å; data collection statistics are summarized in Table 1. The data set was 98% complete for the resolution shell from 3.88 to 1.80 Å, and 94% complete for all data between 50 and 1.80 Å, with an overall R_{merge} of 7.7%. These data were scaled to, and combined with, the original native data. The mean fractional isomorphous difference between the high-resolution data and the original native data was 5.8%. The merged native data set was 99% complete to 1.80 Å resolution.

Refinement of the original hand-built 2.8 Å resolution chitinase structure consisted of 20 rounds of manual model rebuilding coupled with energy minimization and simulated annealing. Table 2 summarizes the refinement process in terms of model statistics. The crystallographic R -value is calculated for all data in the resolution range of 5 Å to the resolution of the given model to facilitate comparisons. The final model (20) has an R -value of 0.18 for the 5.0 to 1.8 Å data shell, and r.m.s. deviations from standard bond lengths and angles of 0.017 Å and 2.88°. Also tabulated to monitor the progress of refinement are the r.m.s. differences in atomic positions between intermediate models and number 20, the final model for the refinement. Finally, the mean difference in phase angles between intermediates and model 20 are calculated for all reflections between 15 and 2.5 Å.

A plot of the ψ versus ϕ dihedral angles (Ramachandran & Sasikharam, 1968) is presented in Figure 2. The plot does not include values for proline or glycine residues. The distribution of ψ and ϕ values lies mostly within the energetically favorable right-handed α -helical region. Five non-glycine residues (Tyr84, Tyr96, Gln118, Gln162 and Asn233) are in the left-handed α -helical conformation and three other non-glycine residues (Asp77, Arg90 and His121) do not belong to any of the typical secondary structural conformations. In both cases, the residues participate in sharp turns that have local favorable interactions that could compensate for their otherwise less probable conformations.

Figure 3 depicts the progress made in map quality over the course of the refinement process. Figure 3(a) shows representative electron density in the original

Table 1

Data collection statistics for the 1.8 Å refinement of barley chitinase

Max shell res (Å)	Avg. res. (Å)	No. reflections collected	Fraction measured (%)	Total observations	R_{sym} (%)
3.88	4.38	1192	58	11,345	6.15
3.07	3.40	2024	100	23,894	5.98
2.69	2.86	2008	99	21,372	6.65
2.44	2.56	1998	98	19,482	7.97
2.27	2.35	1958	98	17,489	9.36
2.13	2.20	1973	97	16,641	11.35
2.03	2.08	1945	97	15,510	13.58
1.94	1.98	1973	97	14,861	15.65
1.86	1.90	1968	97	14,100	18.49
1.80	1.83	1936	98	13,313	20.91
Overall	2.29	18,975	94	168,007	7.69

2.8 Å MIR map. Figure 3(b) shows the corresponding region in the 1.8 Å resolution ($2F_o - F_c$) map. Note the improved electron density for C^β of Trp158 and the carbonyl group of Ala59, as well as the clear holes through the center of the Phe60 and Phe151 aromatic rings. These samples are representative of the quality of the initial and final electron density maps overall.

Model description

The class II barley endochitinase molecule is a compact globular structure approximately 40 Å × 45 Å × 42 Å. Figure 4 illustrates the overall fold of the enzyme in ribbon form. It has three disulfide bonds; the cysteine pairs are 23 to 85, 97 to 105 and 204 to 236. The protein has one antiparallel β-sheet (residues 112 to 114 and 118 to 120), and has ten helical segments labeled A to J. These comprise approximately 47% of the linear sequence. The sequence numbers of these helices,

together with data relevant to capping, termination, and other unusual properties, are presented in Table 3. One helix of note is helix D, which has a proline residue (127) in the center. This clearly disrupts the normal helical hydrogen-bonding pattern. The carbonyl group of Tyr123 rotates away from the helix axis. The ring of Pro127 stacks against that of Trp103, making a strong hydrophobic contact.

Six ion pairs (with lengths under 3.4 Å) are formed among the charged groups, primarily in solvent-exposed regions of constrained mobility. There is one case of stacking between aromatic residues (Phe191:Tyr219), and many examples of perpendicular ring interaction (Burley & Petsko, 1985). These occur as pairs (Tyr111:Tyr125, Phe80:Phe218, Trp72:Trp82), a three-ring cluster (Phe44:Phe60:Trp158), and an extended network (Phe11:Phe28:Tyr29:Tyr31:Phe34:Tyr84:Phe151).

An elongated cleft runs the length of the molecule, and is presumably responsible for substrate binding and catalysis. The base or “floor” of the cleft is

Table 2

Summary of model statistics for the 1.8 Å refinement of barley chitinase

Model	R-value	Δbond (Å)	Δangle (deg.)	r.m.s. difference from model 20 (Å)		⟨Δphase⟩ from model 20 (deg.)	Comment
				Main-chain	Side-chain		
1	41.0	0.026	5.51	1.05	1.90	40.4	
2	20.0	0.020	4.11	0.90	1.77	35.6	
3	22.4	0.022	4.20	0.86	1.76	34.7	Data to 2.5 Å
4	18.2	0.020	3.55	0.44	1.03	27.4	Isotropic Bs refined
5	17.9	0.021	3.51	0.42	0.98	26.4	
6	18.8	0.020	3.45	0.41	0.98	25.3	Data to 2.3 Å
7	20.1	0.020	3.36	0.38	0.95	24.2	
8	20.6	0.020	3.32	0.37	0.98	24.4	Data to 2.1 Å
9	20.3	0.020	3.30	0.33	0.91	24.0	
10	20.1	0.020	3.23	0.31	0.83	23.6	
11	20.5	0.019	3.19	0.30	0.81	22.3	Data to 2.0 Å
12	20.6	0.019	3.15	0.31	0.78	21.8	
13	20.8	0.020	3.19	0.32	0.79	21.8	Data to 1.9 Å
14	20.9	0.020	3.17	0.31	0.87	22.1	
15	21.1	0.019	3.13	0.30	0.87	21.8	Data to 1.8 Å
16	19.0	0.018	3.06	0.28	0.78	17.1	59 water molecules added
17	18.4	0.018	3.04	0.28	0.74	16.6	87 water molecules added
18	18.1	0.018	3.00	0.26	0.58	15.3	
19	18.1	0.017	2.89	0.11	0.35	10.6	
20	18.1	0.017	2.88				Refinement ended

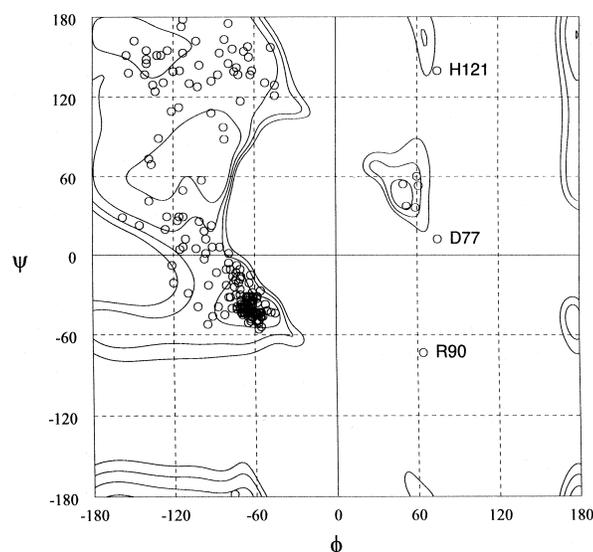


Figure 2. Ramachandran plot for the final model (20). The plot displays main-chain conformational angles. The contour lines represent an energy surface based on an alanine dipeptide model (Peters & Peters, 1981). Proline and glycine residues are not included.

flanked by two “wings”, which may serve to align the chitin polymer into position for hydrolysis.

Comparison between barley chitinase and HEWL

The polymers hydrolyzed by chitinase, cellulase, chitosanase, lysozyme and glucanase are very similar, and it might be expected that some of these enzymes would be related in an evolutionary sense. Amino acid sequence comparisons, using the suite of programs in FASTA (Pearson & Lipman, 1988), show no significant similarity and the enzymes often differ significantly in size. A computer analysis of C^α coordinates (Holm & Sander, 1994) and a structural analysis of molecular models in our laboratory revealed that barley chitinase shares several elements of secondary structure with lysozymes.

The structural similarity between barley chitinase and HEWL is illustrated in Figure 5. Chitinase, with 243 residues, is much larger than the 129 residue HEWL. However, key elements of secondary structure forming the substrate binding and active site areas are common. The 30 to 161 segment of the chitinase C^α backbone is shown as the heavy bonds, superimposed with the 8 to 102 segment of HEWL. The position of a HEWL hexasaccharide substrate is shown as light bonds. Note particularly the pattern of secondary structural elements corresponding to the chitinase C and F helices, and the β -sheet forming the left-hand wall of the substrate binding pocket. These substrate binding elements are very strongly conserved, and various insertions and deletions occur outside this crucial area. The key catalytic residue of HEWL, Glu35, superimposes with Glu67 of chitinase as shown with light bonds (r.m.s. distance of C^α positions is 0.5 Å). The r.m.s. deviation

of all matching C^α positions in the substrate binding elements is 1.7 Å. Helices B and J of chitinase also appear to superimpose with corresponding helices in HEWL. The r.m.s. deviation of matching C^α positions with these helices included is 4.0 Å.

Enzyme/substrate structural studies

Concurrent with crystallographic refinement, experiments were undertaken to help determine how substrate might bind to the enzyme. This would provide direct visualization of which amino acid side-chains bind substrate, and which are poised near the scissile bond. It might be possible, based on these studies, to postulate a catalytic mechanism that could be tested biochemically.

The endochitinases from maize and yam have been shown to bind tetraNAG fairly well, and to hydrolyze it into two diNAG molecules. TriNAG is also a substrate, but its hydrolysis rate is much slower than for tetraNAG, probably due to poorer binding (Koga *et al.*, 1989). When monoclinic barley chitinase crystals were soaked in tetra-NAG they cracked and dissolved completely within two hours. Twenty-four hours later they reformed, but a difference Fourier showed no sugar bound. Subsequent difference Fouriers of chitinase crystals soaked in 0.5 M diNAG also showed no sugar binding. These observations are consistent with the notion that tetraNAG substrate binding causes the molecule to undergo a conformational change incompatible with maintenance of crystal contacts; upon cleavage to disaccharide, the product is released and the enzyme reverts to its unliganded conformation. Cocrystallization experiments with tri-NAG were carried out but no binding was observed.

Hypothetical substrate binding

A hypothetical hexaNAG substrate model, based on that from lysozyme, was fitted and energy minimized as described in Materials and Methods. The r.m.s. difference of protein side-chain atomic positions between the starting and minimized models was 0.36 Å. The r.m.s. difference of sugar atomic positions was 0.32 Å. Protein backbone atoms were constrained to remain fixed. Figure 6(a) shows the region around the putative active site cleft of barley chitinase, as seen in native crystal structure. Several potential hydrogen bonds are formed between the protein and each sugar moiety of the hexasaccharide model. The sugars are labeled A to F, from the non-reducing end. As shown in Figure 6(b), the $N^{\epsilon 1}$ atom of Trp103 donates a hydrogen bond to the O-7 atom of the A sugar, while O-6 of the A sugar donates a hydrogen bond to the carbonyl oxygen atom of Pro163. The B sugar is involved in three hydrogen bonds: O-5 with the amide nitrogen atom of Asn124, O-6 with the amide group of Gln162, and O-7 with the N^{ϵ} atom of Lys165. The amide nitrogen atom of Asn124 also hydrogen bonds with O-3 of the

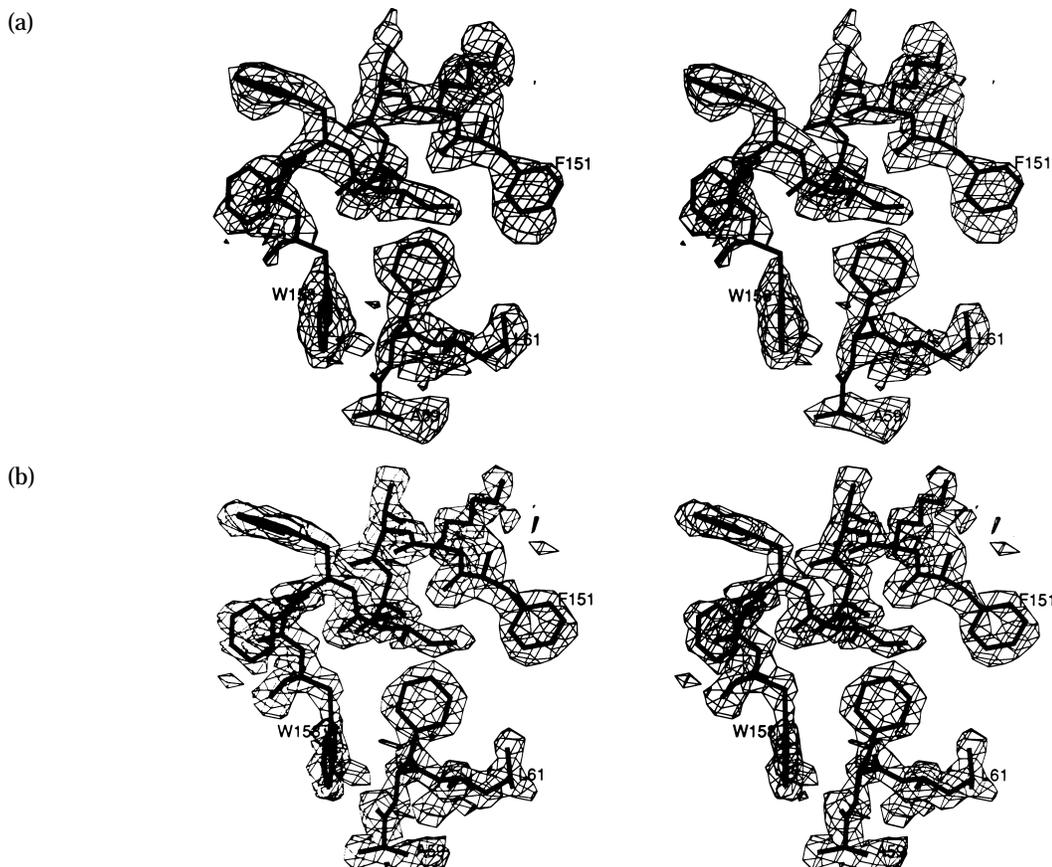


Figure 3. Electron density map quality at the beginning and end of refinement: (a) 2.8 Å resolution MIR map; (b) the $(2F_o - F_c)$ map superimposed on the final model (see the text). In both maps, 10% of the cell volume lies within the contours.

C sugar, and O-6 of the C sugar forms a hydrogen bond with Oⁿ of Tyr123. The N^{o2} atom of Asn199 donates a hydrogen bond to O-5 of the D sugar. The carboxylate group of Glu67, although in position to receive a hydrogen bond from O-6 of sugar D, is

likely the catalytic proton donor and in the desolvated substrate binding complex is probably donating a hydrogen bond to O-4 of the leaving group E sugar. The amide nitrogen atom of Gln118 donates a hydrogen bond to O-7 of sugar E. Alternatively, the

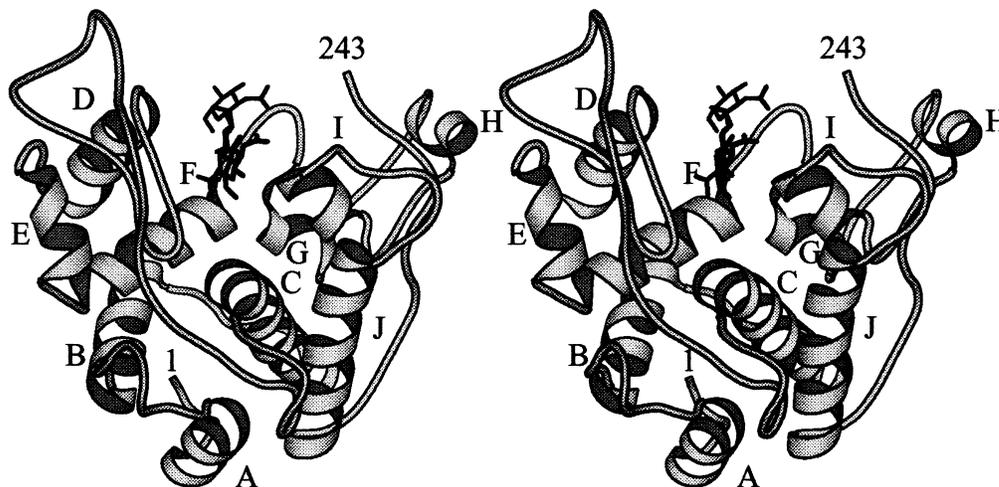


Figure 4. Ribbon diagram of the barley endochitinase molecule. Helices are labeled A to J. This Figure was created using the program MOLSCRIPT (Kraulis, 1991).

Table 3

Description of α -helices		
Helix	Residues	Comments
A	7–15	N cap: Ser7 to N-10
B	30–38	N cap: Thr30 to N-33; C-terminal 2 residues in 3_{10} helix
C	49–68	N cap: Ser49 to N-52; C cap: Thr68 to O-64
D	121–133	Pro127 stacks on Trp103
E	139–146	N cap: Asn139 to N-141
F	146–160	N cap: Asp146 to N-149
G	167–174	N cap: Ser167 to N-170
H	179–186	N cap: Ser179 to N-182
I	191–199	C terminus becomes 3_{10} helix
J	211–226	N cap Asp209 to N-211

side-chain of Glu89 may be fully extended, displacing the side-chain of 118, and forming a hydrogen bond with O-7. As described in the Discussion, Glu89 is also possibly involved in catalysis. A guanidinium nitrogen atom of Arg211 donates hydrogen bonds to O-3 and O-7 of the F sugar.

Discussion

Model refinement

The initial hand-built model was basically correct as evidenced by the drop in *R*-value from 41% to 20% in a single round of automated refinement at 2.8 Å resolution (Hart *et al.*, 1993). Subsequent phase extension and refinement were facilitated by the fact that few changes were needed from the original hand-built model.

Omit maps were calculated for several other exterior regions of the molecule that seemed ambiguous. However, only minor changes were made in side-chain conformation. The density for the solvent-exposed region spanning residues 89 to 96 was weaker than average due to the relatively high (mid to upper 30s) temperature factors these

residues exhibited. This somewhat disordered loop was difficult to model and contains one of the residues (Arg90) that falls in an energetically unfavorable position (Figure 2). Another residue with an unfavorable conformation, Asp77, lies at the end of a β -loop adjacent to Pro76. The conformational restrictions of the proline residue may contribute to the high energy conformation of residue 77. Another residue with an unfavorable conformation, His121, lies at the end of a β -strand (residues 118 to 120) and at the beginning of an α -helix (residues 122 to 133) and has hydrogen bonds involving both. The peptide nitrogen atom of 121 donates a hydrogen bond to the carbonyl oxygen atom of 111 as part of a β -sheet while the carbonyl oxygen atom of 121 receives a hydrogen bond from the nitrogen atom of 125.

Fifty-nine water molecules were added to model 15, which used data to 1.8 Å resolution. Their inclusion caused a surprisingly large drop in the *R*-value of 2.1%. The final model (20) contains 129 water molecules.

It appears that further refinement at 1.8 Å resolution would be superfluous, as evidenced by the fact that *R*-values had stopped changing in any meaningful way and the average phase shifted only

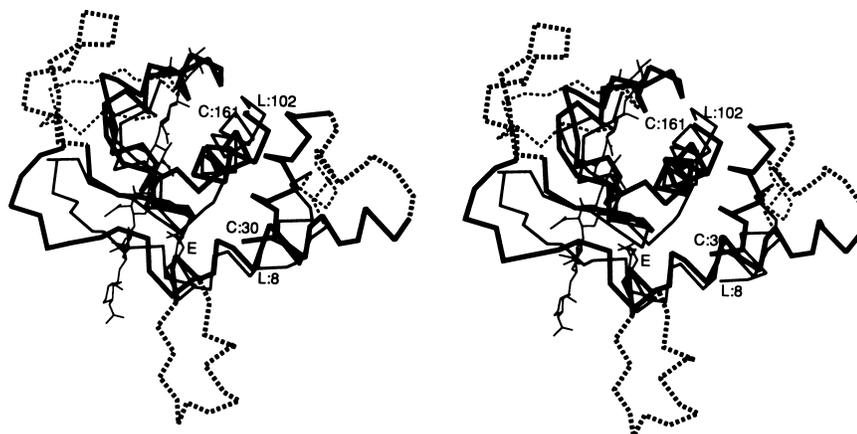


Figure 5. Comparison of structural similarities between barley chitinase and HEWL. Chitinase, with 243 residues, is much larger than the 129 residue HEWL. However, key elements of secondary structure forming the substrate binding and active site areas are common. The 30 to 161 segment of the chitinase C $^{\alpha}$ backbone is shown as the heavy bonds, superimposed with the 8 to 101 segment of HEWL. The binding of a hexasaccharide is shown as light bonds. The crucial catalytic glutamate residues, 67 in chitinase and 35 in HEWL, are also shown by light bonds.

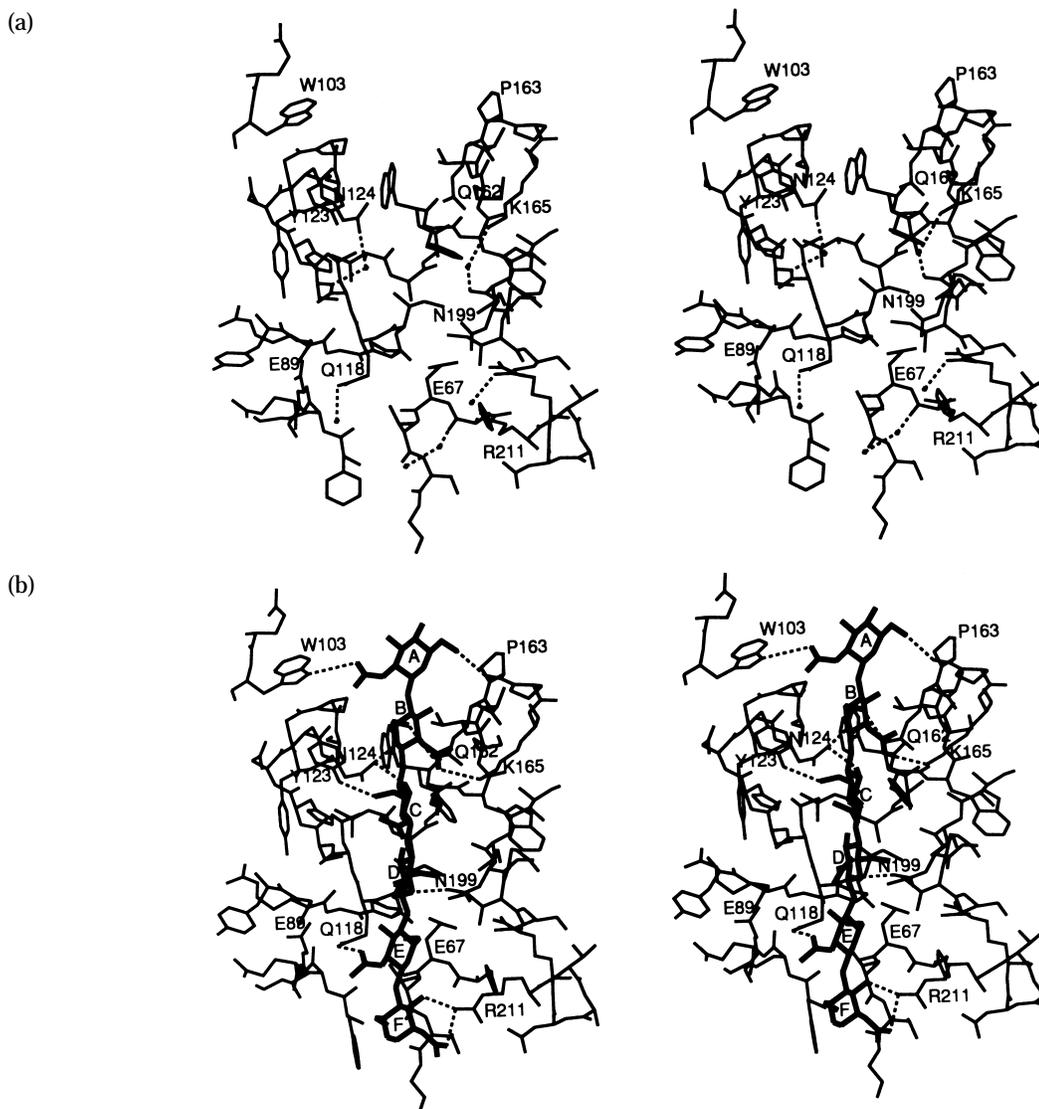


Figure 6. Hypothetical model of hexaNAG binding to barley chitinase. (a) The active site region of barley chitinase, including solvent molecules is shown. The molecule contains a very prominent and elongated cleft. (b) A hexaNAG model has been fit to the active site cleft, based largely on the known binding of substrate to lysozyme. The substrate model is shown in dark bonds, while crucial amino acids from the enzyme are shown in lighter bonds. Hydrogen bonds are shown as broken lines. The sugar moieties are labeled A to F from the non-reducing end.

10.6° from model 19 to model 20. It must be realized that there is no single best model for a protein and that there exists a family of equivalent models with essentially identical *R*-factors and distortions from standard bond lengths and angles. We compared two equivalent models separated by 0.05 ps along the simulated annealing path for the X-PLOR refinement of model 20; their average phase difference was 10.8°. It appears that the phase shift among the equivalent "best" models of this protein is about 10°.

The class II (barley) chitinase fold compared with other chitinases

Within the plant chitinase family, the class II barley enzyme serves as a prototypical model for class I, II and IV chitinases. This is clear from the sequence

comparisons and alignments shown in Figure 1. Of the 243 amino acid residues in the barley chitinase molecule, 49 are invariant in the chitinase catalytic domains shown in Figure 1 and many more residues are highly conserved. Pairwise sequence comparisons range from 73% identity with the arabidopsis (class I) enzyme to a still significant 38% with corn (class IV).

The invariant non-polar residues tend to fall into the hydrophobic core of the molecule, as expected for proteins with similar folding patterns. For example, Figure 1 shows many residues are conserved in a patch around the invariant Phe60 of helix C. Others are conserved around the invariant Trp158 of helix F. As seen in Figure 5, the C and F helices form the core of the substrate binding cleft. The conserved hydrophobic residues interact to secure that vital

active site geometry. The side-chain of the invariant Phe60 directly contacts the side-chains of three invariant residues on helix F, Phe151, Ala154 and Trp158.

The conserved polar residues tend to lie along the prominent cleft, presumed to be the substrate binding and catalytic site. For example, three invariant residues appear to bind the substrate, Gln118, Tyr123 and Asn124 (Figure 6(b)). This will be expanded upon below, in the discussion of substrate modeling.

As shown in Figure 1, the key difference between the class IV chitinase, represented by corn, and the class I and II molecules is two large deletions. The corn deletion between barley residues 71 and 83 occur downstream from the catalytic Glu67 at the end of the C helix. Notice in Figure 5 that this corresponds to a deletion (broken line) seen also in HEWL. This loop can clearly be dispensed with and the corn enzyme resembles HEWL in this area. The deletion corresponding to barley residues 167 to 188 does not appear on Figure 5, but it too is in a region of the molecule not involved in the active site. HEWL shows a corresponding deletion in this variable part of the molecule.

Hypothetical substrate binding and mechanism of action

Attempts to soak substrate analogs into chitinase crystals have so far been unsuccessful, but a model for substrate binding based on data for HEWL has been made. Computer comparisons (Holm & Sander, 1994) and the superposition shown in Figure 5 show that barley chitinase strongly resembles HEWL. The substrates for these enzymes are similar and it is not unreasonable to suppose that the mechanisms of action of the hydrolases are similar. The notion that the chitinase mechanism of action is similar to that for lysozyme is supported by the observation that molecular superposition matches Glu67 of chitinase with Glu35 of HEWL and, by extension, with Glu11 of T4L and with Glu73 of GEWL (Rossmann & Argos, 1976; Matthews *et al.*, 1981; Grütter *et al.*, 1983).

In the HEWL mechanism, Asp52 is generally considered to stabilize oxycarbonium ion build-up on the D sugar during hydrolysis. The most likely analog in chitinase is Glu89, although it does not match well in space with Asp52 in the least-squares superposition; the C² atoms differ by 5.3 Å. However, the C² position of this second active site residue varies considerably even among the lysozymes; for example, Asp52 of HEWL differs from Asp86 of goose lysozyme by 5.7 Å (Grütter *et al.*, 1983). The carboxylate groups are believed to make electrostatic interaction with the transition state and such interactions are not sensitive to angular displacement. As long as the charges are close, they provide stabilization.

Binding of tetrasaccharide to HEWL has been observed crystallographically and a hexasaccharide

model was proposed in sites labeled A through F (Ford *et al.*, 1974; Blake *et al.*, 1967). Similarly, a hexasaccharide was positioned in chitinase using the least-squares operator that relates chitinase to HEWL and the resulting model was subjected to energy minimization. Figure 6(b) shows the postulated interactions between barley chitinase and the hexasaccharide.

In support of the model, alkylation of a tyrosine residue homologous to Tyr123 greatly reduces activity in the chitinase from *Zea mays* (Verburg *et al.*, 1992). As seen in Figure 6(b), this residue is probably involved in substrate binding, making a strong bond to the C sugar. Site-directed mutagenesis of the corresponding Tyr to Phe in *Arabidopsis* decreases activity about 40% (Verburg *et al.*, 1993). Figure 1 shows that, in addition to the catalytic residues Glu67 and Glu89, the putative substrate binding residues of barley chitinase are conserved in plant chitinases. Some, like asparagine 124 and 199 are invariant, while others like Trp103 and Tyr123 are conserved in all but one species.

The hypothetical substrate model does not incorporate any major conformational change in chitinase, but it is reasonable to believe that some change must occur upon substrate binding. This would account for the observation that binding of tetrasaccharide substrates causes the crystals to dissolve. It may be that our nearly static model ignores some large structural rearrangement. However, since our model is consistent with the limited chemical modification data, and rationalizes the conservation of residues in the active site, it seems more likely that the model is basically correct and that only minor conformational changes occur upon substrate binding, but these are sufficient to disrupt the weak crystal packing forces.

Comparison with HEWL and the substrate model shows that Glu67 of chitinase is likely to be the proton donor in the catalytic mechanism. Glu67 is invariant in the plant chitinases (Figure 1). It is the homolog of Glu35 in HEWL and is responsible for protonation of the E sugar leaving group. In the desolvated, substrate-bound enzyme, Glu67 is likely to be protonated. One side-chain oxygen atom of Glu67 lies 4.4 Å from O-4 of the D site sugar; this is comparable to the distances observed for Glu35 of HEWL in the binding of NAM-NAG-NAM and a tetrasaccharide lactone (Ford *et al.*, 1974; Kelly *et al.*, 1979).

It is known that the HEWL mechanism is a double displacement type, proceeding with retention of anomeric configuration (Dahlquist *et al.*, 1969). As a result of site-directed mutagenesis studies on phage T4 lysozyme, a single displacement mechanism has been suggested for that enzyme, although there is no supporting biochemical evidence (Kuroki *et al.*, 1993). It is not known if the barley chitinase mechanism proceeds with retention of anomeric configuration. In the absence of this information, the most likely mechanism is a double displacement similar to that for HEWL and implied in the discussions above. However, the substrate model

above and, in particular, the position of Glu89, allows an alternative. As presently built, the model has room for a water molecule to bind between Glu89 and the α face of the D sugar; the OG atom of Ser120 could hydrogen bond to the water molecule. In a single displacement mechanism, Glu67 would protonate O-4 of the leaving E sugar, while Glu89 activates the water molecule for a nucleophilic attack on C-1 of the D sugar. There is no water molecule observed at this position in the 1.8 Å crystal structure, but water 277 is only 5.0 Å away and could be displaced to the catalytic position upon binding of the substrate.

Materials and Methods

High-resolution data collection

Crystallization and structural determination of the barley endochitinase to 2.8 Å resolution have been reported (Hart *et al.*, 1993). The original native data were 96.3% complete to 2.2 Å resolution. For higher-resolution data collection, a single large chitinase crystal of dimensions 0.43 mm \times 0.86 mm \times 0.96 mm, was mounted in an appropriate quartz capillary. Three-dimensional diffraction data were collected at 25°C on a SDMS multi-wire area detector using the method of Xuong *et al.* (1985). Crystal alignment was facilitated by the program INDEX, written by Dr Stephen Ernst in our laboratory, which allows rapid orientation angle determination from a random crystal mounting. The crystal was rotated around ω , collecting 0.12° frames for 60 seconds each. The X-ray source was an Elliot GX-20 rotating anode generator operating at 40 kV, 40 mA with a graphite monochromator. Data were collected between 50.3 and 1.80 Å resolution and evaluated using the University of California, San Diego (UCSD) software system (Howard *et al.*, 1985). The high-resolution diffraction data collected from this crystal are summarized in Table 1. The data were scaled to, and combined with, the original native diffraction data, keeping the new intensities where the data sets overlapped.

Refinement of the model

The crystallographic refinement of the barley endochitinase used energy minimization and simulated annealing found in the X-PLOR package (Brünger, 1988). The strategy of refinement was similar to that reported for the pokeweed antiviral protein (PAP; Monzingo *et al.*, 1993). A "round" of refinement is defined as a hand rebuilding of the model on an Evans and Sutherland PS390 graphics system running the program FRODO (Jones, 1982), followed by a variable number of cycles of automated refinement using the molecular dynamics option of X-PLOR on a CRAY Y-MP8/864. The automated refinement used terms from 5 Å to the nominal resolution limit of the run, ranging from 2.8 to 1.8 Å. Fourier maps were calculated using all terms between 15 Å and the upper resolution limit for the data appropriate to that round of refinement.

Data were extended conservatively in the refinement process, with reflections added to 2.5 Å during round 2, 2.3 Å during round 5, 2.1 Å during round 7, 2.0 Å during round 10, 1.9 Å during round 12, and 1.8 Å during round 14. Individual isotropic temperature factor refinement was begun during round 3: 59 bound water molecules were

included in round 15, and 87 in round 16. Each putative water molecule was examined individually and added to the model if: (1) it was within hydrogen-bonding distance (3.4 Å) of an appropriate atom of the protein structure; (2) it had good (3σ) difference density; and (3) it had good density in a ($2F_o - F_c$) Fourier map. The refinement was considered complete when the average phase shift between model 19 and model 20 leveled off at 10.6°, 3σ difference maps were flat, and the *R*-value remained essentially constant.

Oligo-NAG soaking experiments

Chitinase crystals were transferred to an artificial mother liquor (100 mM Tris-HCl (pH 8.5), 1.25 M sodium phosphate, 125 mM NaCl) containing 5 mM tetraNAG, 20 mM triNAG, or 30 to 500 mM diNAG, and observed over a period of approximately 30 hours. In co-crystallization experiments, 0.5 mM barley chitinase was incubated in solution with 20 mM triNAG. This solution was then used in micro-seeded hanging drop vapor diffusion experiments as described (Hart *et al.*, 1993). Crystals appeared after two days, and grew to usable size within five days. Three-dimensional diffraction data were collected and reduced as described (Hart *et al.*, 1993). Difference Fourier maps were calculated using in-house programs and inspected on the graphics system.

Comparison with HEWL

Coordinates of HEWL were taken from the Brookhaven Protein Data Bank (accession number 6LYZ). Model comparisons and manipulations were done on an Evans and Sutherland PS390 graphics system using the program FRODO (Jones, 1982). Least-squares superpositions were done using FRODO and the program X-PLOR (Brünger, 1988).

Substrate model building

Coordinates of hen egg-white lysozyme (HEWL) and the trisaccharide NAM-NAG-NAM bound to HEWL were taken from the Protein Data Bank, entries 6LYZ and 9LYZ (Blake *et al.*, 1965; Kelly *et al.*, 1979). Based on a superposition of HEWL and barley chitinase coordinates using the α -carbon atoms of a β antiparallel sheet and two α -helices (HEWL residues 24 to 36, 51 to 59 and 89 to 101; chitinase residues 56 to 68, 112 to 120 and 148 to 160), the trisaccharide was transformed to a putative binding site in the chitinase model. The NAM residues were converted to NAG. The transformed trisaccharide provided the basis for the B, C and D sugars of a hexa-NAG substrate model. The A site sugar moiety was based on coordinates from a tetrasaccharide lactone (Ford *et al.*, 1974). NAG residues were added at the reducing end of the trisaccharide model to fill the E and F sites. Side-chains of the protein model were moved to maximize potential hydrogen bonds with the model substrate. A model of the protein-hexasaccharide complex was energy minimized using X-PLOR. The protein backbone positions were constrained to remain fixed during the minimization.

Amino acid sequence analysis

The amino acid sequences for various plant chitinases were taken from Genbank files. The following sequences were analyzed: barley (Leah *et al.*, 1991); arabidopsis (Sumac *et al.*, 1990); corn (Huynh *et al.*, 1992); pea (Vad *et al.*,

1995); poplar (Parson *et al.*, 1989); potato (Gaynor & Unkenholz, 1989); rice (Huang *et al.*, 1991), and tobacco (van Buren *et al.*, 1992). The sequences were aligned using the program CLUSTAL (Higgins & Sharp, 1989). All the protein except barley have an amino terminal chitin binding domain and a flexible linker to the catalytic domain. To analyze the similarity of the catalytic domains, the ALIGN routine of the FASTA program suite was used (Pearson & Lipman, 1988).

Acknowledgements

We thank Dr Michael Ready and Edward Marcotte for help with sequence alignment and preparation of the ribbon drawing. This work was supported by grants GM 30048 and GM35989 from the National Institutes of Health and by grants from the Foundation for Research and the Welch Foundation.

References

- Abeles, F. B., Bosshart, R. P., Forrence, L. E. & Habig, W. H. (1970). Preparation and purification of glucanase and chitinase from bean leaves. *Plant Physiol.* **47**, 129–134.
- Albersheim, P. & Valebt, B. S. (1978). Host-pathogen interactions in plants. *J. Cell Biol.* **78**, 627–643.
- Bartnicki-Garcia, S. (1968). Cell wall chemistry, morphogenesis, and taxonomy of fungi. *Annu. Rev. Microbiol.* **22**, 87–102.
- Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). Structure of hen egg-white lysozyme. *Nature (London)*, **206**, 757–761.
- Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, C. D. & Sarma, V. R. (1967). Crystallographic studies of the activity of hen egg-white lysozyme. *Proc. Roy. Soc. ser. B*, **167**, 378–388.
- Boller, T., Gehri, A., Mauch, F. & Vogeli, U. (1983). Chitinase in bean leaves: induction by ethylene, purification, properties, and possible function. *Planta*, **157**, 22–31.
- Bowles, D. J., Gurr, S. J., Scollan, C., Atkinson, H. J. & Hammond-Kosack, K. E. (1991). Local and systemic changes in plant gene expression following root infection by cyst nematodes. In *Biochemistry and Molecular Biology of Plant-Pathogen Interactions* (Smith, C. J., ed.), pp. 225–236, Clarendon Press, Oxford.
- Brogliè, K. E., Chet, I., Holliday, M., Cressman, R., Biddle, P., Knowlton, S., Mauvais, C. J. & Brogliè, R. (1991). Transgenic plants with enhanced resistance to the fungal pathogen *Rhizoctonia solani*. *Science*, **254**, 1194–1197.
- Brünger, A. T. (1988). Crystallographic refinement by simulated annealing. In *Crystallographic Computing 4: Techniques and New Technologies* (Isaacs, N. W. & Taylor, R. R., eds), pp. 126–140, Clarendon Press, Oxford.
- Burley, S. K. & Petsko, G. A. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.
- Collinge, D. B., Kragh, K. M., Mikkelsen, J. D., Nielsen, K. K., Rasmussen, U. & Vad, K. (1993). Plant chitinases (mini-review). *Plant J.* **3**, 31–40.
- Dahlquist, F. W., Borders, C. L., Jacobson, G. & Raftery, M. A. (1969). The stereospecificity of human, hen and papaya lysozymes. *Biochemistry*, **8**, 694–700.
- Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tijan, R. (1974). Crystal structure of a lysozyme-tetrasaccharide lactone complex. *J. Mol. Biol.* **88**, 349–371.
- Gaynor, J. J. & Unkenholz, K. M. (1989). Sequence analysis of a genomic clone encoding an endochitinase from *Solanum tuberosum*. *Nucl. Acids. Res.* **17**, 5855–5856.
- Grütter, M. G., Weaver, L. H. & Matthews, B. W. (1983). Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? *Nature (London)* **303**, 828–831.
- Hart, P. J., Ready, M. P. & Robertus, J. D. (1992). Crystallization of an endochitinase from *Hordeum vulgare* L. seeds. *J. Mol. Biol.* **225**, 565–567.
- 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.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309–316.
- Higgins, D. G. & Sharp, P. M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS*, **5**, 152–253.
- Holm, L. & Sander, C. (1994). Structural similarity of plant chitinase and lysozyme from animals and phage. *FEBS Letters*, **340**, 129–132.
- Howard, A. J., Nielsen, C. & Xuong, N. H. (1985). Software for a diffractometer with multiwire area detector. *Methods Enzymol.* **114**, 452–472.
- Huang, J. K., Wen, L., Swegle, M., Tran, H. C., Thin, T. H., Noyler, H. N., Muthukrishnan, S. & Reeck, G. R. (1991). Nucleotide sequence of a rice genomic clone that encodes a class I endochitinase. *Plant Mol. Biol.* **61**, 479–480.
- Huynh, Q. K., Hironaka, C. M., Levine, E. B., Smith, C. E., Borgmeyer, J. R. & Shah, D. M. (1992). Antifungal proteins from plant purification, molecular cloning and antifungal properties of chitinases from maize seed. *J. Biol. Chem.* **267**, 6635–6640.
- Jones, T. A. (1982). FRODO: a graphics fitting program for macromolecules. In *Computational Crystallography* (Sayre, D., ed.), pp. 303–317, Oxford University Press, Oxford.
- Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G. & Phillips, D. C. (1979). X-ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme. *Nature (London)*, **282**, 875–878.
- Koga, D., Tsukamoto, T., Sueshige, N., Utsumi, T. & Ide, A. (1989). Kinetics of chitinase from yam, *Dioscorea opposita* THUNB. *Agric. Biol. Chem.* **53**, 3121–3126.
- Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
- Kuroki, R., Weaver, L. H. & Matthews, B. W. (1993). A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science*, **262**, 2030–2033.
- Leah, R., Tommerup, H., Svendsen, I. & Mundy, J. (1991). Biochemical and molecular characterization of three barley seed proteins with antifungal properties. *J. Biol. Chem.* **266**, 1564–1573.
- Matthews, B. W., Grütter, M. G., Anderson, W. F. & Remington, S. J. (1981). Common precursor of lysozymes of hen egg-white and bacteriophage T4. *Nature (London)*, **290**, 334–335.
- Mauch, F., Mauch-Mani, B. & Boller, T. (1988). Antifungal hydrolases in pea tissue. II. Inhibition of fungal growth by combinations of chitinase and β -1,3-glucanase. *Plant Physiol.* **88**, 936–942.

- Molano, J., Polacheck, I., Duran, A. & Cabib, E. (1979). An endochitinase from wheat germ. *J. Biol. Chem.* **254**, 4901–4907.
- Monzingo, A. F., Collins, E. J., Ernst, S. R., Irvin, J. D. & Robertus, J. D. (1993). The 2.5 Å structure of pokeweed antiviral protein. *J. Mol. Biol.* **233**, 705–715.
- Parson, T. J., Bradshaw, H. D. & Gordon, M. P. (1989). Systemic accumulation of specific mRNAs in response to wounding poplar trees. *Proc. Nat. Acad. Sci., U.S.A.* **86**, 7895–7899.
- Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 2444–2448.
- Peters, D. & Peters, J. (1981). Quantum theory of the structure and bonding in proteins. Part 8. The alanine dipeptide. *J. Mol. Struct.* **85**, 107–123.
- Ramachandran, G. N. & Sasikharam, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–437.
- Rasmussen, U., Giese, H. & Mikkelsen, J. D. (1992). Induction and purification of chitinase in *Brassica napus* L. spp. *oleifera* infected with *Phoma lingam*. *Planta*, **187**, 328–334.
- Roberts, R. L. & Cabib, E. (1982). *Serratia marcescens* chitinase: one-step purification and use for the determination of chitin. *Anal. Biochem.* **127**, 402–412.
- Roberts, W. K. & Selitrennikoff, C. P. (1986). Isolation and partial characterization of two antifungal proteins from barley. *Biochim. Biophys. Acta*, **880**, 161–170.
- Roberts, W. K. & Selitrennikoff, C. P. (1988). Plant and bacterial chitinases differ in antifungal activity. *J. Gen. Microbiol.* **134**, 169–176.
- Roberts, W. K. & Selitrennikoff, C. P. (1990). Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. *J. Gen. Microbiol.* **136**, 1771–1778.
- Robertus, J. D., Hart, P. J., Monzingo, A. F., Marcotte, E. & Hollis, T. (1995). The structure of chitinases and prospects for structure-based drug design. *Can. J. Bot.* In the press.
- Rossmann, M. G. & Argos, P. (1976). Exploring structural homology of proteins. *J. Mol. Biol.* **105**, 75–96.
- Shinshi, H., Neuhaus, J. M., Ryals, J. & Meins, F. (1990). Structure of a tobacco endochitinase gene: evidence that different chitinase genes can arise by transposition of sequences encoding a cysteine-rich domain. *Plant Mol. Biol.* **14**, 357–368.
- Sumac, D. A., Hironaka, C. M., Yallaly, P. E. & Shah, D. M. (1990). Isolation and characterization of genes encoding basic and acidic chitinases in *Arabidopsis thaliana*. *Plant Physiol.* **93**, 907–914.
- Vad, K., deNeergard, E., Madriz-Ordenana, K., Mikkelsen, J. D. & Collinge, D. B. (1995). Accumulation of defense related transcripts and cloning of a chitinase mRNA from pea leaves (*Pisum sativum* L.) inoculated with *Ascochyta pisi* Lib. *Plant Sci.* In the press.
- van Buren, M., Neuhaus, J. M., Shinshi, G., Ryals, J. & Meins, F. (1992). The structure and regulation of homologous tobacco endochitinase genes of *Nicotiana sylvestris* and *N. tomentosiformis* origin. *Mol. Gen. Genet.* **232**, 460–469.
- Verburg, J. G., Smith, C. E., Lisek, C. A. & Khai Huynh, Q. (1992). Identification of an essential tyrosine residue in the catalytic site of a chitinase isolated from *Zea mays* that is selectively modified during inactivation with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide. *J. Biol. Chem.* **267**, 3886–3893.
- Verburg, J. G., Rangwala, S. H., Samac, D. A., Luckow, V. A. & Khai Huynh, Q. (1993). Examination of the role of tyrosine-174 in the catalytic mechanism of the *Arabidopsis thaliana* chitinase: comparison of variant chitinases generated by site-directed mutagenesis and expressed in insect cells using baculovirus vectors. *Arch. Biochem. Biophys.* **300**, 223–230.
- Walker-Simmons, M., Hadwider, L. & Ryan, C. A. (1983). Chitosans and pectic polysaccharides both induce the accumulation of the antifungal phytoalexin pisatin in pea pods and antinutrient proteinase inhibitors in tomato leaves. *Biochem. Biophys. Res. Commun.* **110**, 194–199.
- Wessels, J. G. H. & Sietsma, J. H. (1981). Fungal cell walls: a survey. In *Encyclopedia of Plant Physiology, New Series* (Tanner, W. & Loewus, F. A., eds), Vol. 13B, pp. 352–394, Springer Verlag, Berlin.
- Wright, C. S. (1987). Refinement of the crystal structure of wheat germ agglutinin isolectin 2 at 1.8 Å resolution. *J. Mol. Biol.* **194**, 501–529.
- Xuong, N. H., Nielsen, C., Hamlin, R. & Anderson, D. (1985). Strategy for data collection from protein crystals using a multiwire counter area detector diffractometer. *J. Appl. Crystallogr.* **18**, 342–350.

Edited by I. A. Wilson

(Received 14 November 1994; accepted 3 February 1995)