

*Structure – function analysis of TIM-barrel proteins: The  
catalytic mechanism of chitinolytic enzymes*

*Thesis Submitted for the Degree  
“Doctor of Philosophy”*

*By*  
**Gali Prag**

**This work was carried out under the  
supervision of Prof. Amos Oppenheim**

**Submitted to the Senate of the Hebrew University of Jerusalem**

**In July 2001**

## **Thanks**

To Prof. Amos Oppenheim, to whom I will be forever grateful, for his generous help and guidance and for the support he gave me through difficult times.

To my parents that seeded my curiosity and desire for knowledge.

To my girlfriend Dr. Dorit Raz, for her love, constant moral support, and for critical reading of this manuscript.

To my warm and caring family.

To Ms. Efrat Gabay who encouraged me throughout the past years.

### **I am also grateful to**

Dr. Paul Tucker, for his guidance and help with the crystallography studies at the Hamburg Synchrotron and for his warm hospitality.

Dr. Simon Greenberg, for his help with the Bioinformatic studies.

Prof. Constantinos Vorgias for his continuous help.

Dr. Kyriacos Petratos, and Dr. Yannis Papanikolau, for their help with the crystallography project in Heraklion Crete and for their warm hospitality.

Dr. Oded Livnah, for his advice and interest and for helping me with the crystallization in his lab.

Prof. Josef Shlomai, for his interest and advice.

Prof. Ilan Chet, for his scientific support.

Prof. Vincent Eijsink in Ås, for the wonderful conversation and ideas.

Ms. Dinah Teff, for her empathy and help with manuscript writing.

Ms. Simi Koby, for her help and involvement.

Mr. Giorgos Tavlas for his help and for his warm hospitality in Heraklion.

Dr. Yoram Shotland, for his scientific and personal support.

Mr. Oren Kobiler and Ms. Hagit Rudich and all my other friends at the Department of Molecular Genetics and Biotechnology.

# Content

<b>Summary</b>	Page 6
<b>Abbreviations</b>	Page 12
<b>Introduction</b>	Page 14
<b>Materials and Methods</b>	
Table lists of substrates, DNA oligonucleotides, bacterial strains and plasmids used in this work	Page 27
Construction and expression of chitobiase and chitinase A wild type and mutant proteins	Page 31
Screening of chitinase A mutants with low activity	Page 32
Cloning and knockout of the <i>E. coli celF / chbF</i> gene	Page 33
Random mutagenesis and creation of a combinatorial DNA library	Page 34
DNA sequencing and sequencing corrections	Page 35
Purification of chitobiase, chitinase A and ChbF	Page 36
Crystallization of chitobiase and chitinase A mutant proteins	Page 37
Kinetic analysis of chitobiase and chitinase A wild type and mutant proteins	Page 38
Synchrotron data collection under cryo-cooling conditions	Page 39
Molecular replacement and model refinement	Page 40
Homology search and sequence alignment	Page 41
Modeling of human Hexosaminidase A	Page 41

## **Results**

### **Chapter I. Genetic analysis of chitobiase catalytic TIM-barrel domain:**

Conservation of the catalytic domain Page 48

**Chapter II. Structures of chitobiase mutants with the substrate di-*N*-acetyl-D-glucosamine: The catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540** Page 58

**Chapter III. OligoNAG substrates bound to chitinase A** Page 69

**Chapter IV. Analysis of the catalytic mechanism of chitinase A** Page 78

**Chapter V. Trapping reaction intermediates in chitinase A D391A mutant** Page 85

**Discussion** Page 96

**Bibliography** Page 111

## **Appendixes**

**A1** Multiple alignment of glycosyl hydrolase protein family 18.

**A2** List of papers influenced by this thesis

**The following three manuscripts are a direct outcome of this thesis**

**Prag, G., Papanikolau, Y., Tavlas G., Vorgias, E.C., Petratos, K. and**

**Oppenheim, B.A.** Structures of chitobiase mutants with the substrate Di-*N*-acetyl-D-glucosamine: The catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J. Mol. Biol.* (2000) **300**, 611. Attached

**Papanikolau, Y., Prag, G., Tavlas G., Vorgias, E.C., Oppenheim, B.A. and Petratos, K.** High resolution structural analyses of mutant chitinase A complexes with substrates provide new insight into the mechanism of catalysis. *Biochemistry* (2001) **40**, 11338. Attached

**Prag, G., Vorgias, E.C. and Oppenheim, B.A.** Conservation of structural elements and catalytic mechanism in *Serratia marcescens* chitinolytic enzymes. *Chitin Enzymology* (2001) **3**, 351 Not attached

**The following two manuscripts are not related to the study of structure - function analysis of TIM-barrel proteins.**

**Prag, G., Greenberg, S. and Oppenheim, B.A.** Structural principles of prokaryotic gene regulatory proteins and the evolution of repressors and gene activators. *Mol. Microbiol.* (1997) **26**, 619. Attached.

**Giladi, H., Koby, S., Prag, G., Engelhorn, M., Geiselmann, J. and Oppenheim, B.A.** Participation of IHF and a distant UP element in the stimulation of the phage  $\lambda$  P<sub>L</sub> promoter. *Mol. Microbiol.* (1998) **30**, 433. Attached

**A3** CD-ROM and Internet links, see instructions on Page 155. Attached

**Hebrew summary**

## Summary

Chitin degrading enzymes are ubiquitous and are found in prokaryotic and in eukaryotic organisms as well as in archaea. *Serratia marcescens* has been used as a model system for the utilization of chitin as a carbon source. In the presence of chitin, *S. marcescens* expresses chitinase A, chitinase B, chitinase C and chitobiase that degrade chitin to mono *N*-acetylglucosamine (NAG). The structures of *S. marcescens* chitinolytic enzymes, chitinase A and chitobiase were solved in collaboration with the laboratory of K. Wilson at EMBL. The structure of chitinase B was recently solved by D. van Aalten and his coworkers. The catalytic domains of all three enzymes were found to possess a TIM-barrel structure. Chitinase A, B and C were assigned to glycosyl hydrolase protein family 18, while chitobiase belongs to protein family 20. Interestingly, chitobiase shows similarity to human hexosaminidase A (HexA). Mutation leading to a defect in HexA results in Tay Sachs disease.

This research is directed toward the understanding of the mode-of-action of chitobiase and chitinase A. In this study I applied the disciplines of bioinformatics, genetics, biochemistry and structural biology. The first goal was to further investigate the hypothesis that chitobiase and HexA are orthologs. These studies provide a structural rationalization for Tay-Sachs mutations. The second goal was to generate mutations into chitobiase and chitinase A and characterize these mutants. In this study I aimed to identify the importance of conserved residues in both enzymes. Another goal of this research was to characterize the mode of substrate binding of chitobiase

and chitinase A. Structural studies of enzyme-substrate interactions were made possible by the availability of mutant derivatives generated in this work. Furthermore, my biochemical and structural studies yielded a deeper understanding of the enzymatic mechanism.

Based on the structure of chitobiase I built a model of the catalytic domain of the human enzyme. This model helps to rationalize the severity of Tay Sachs mutations.

The active sites of chitinase A, chitinase B and chitobiase are located at the carboxy-terminal end of the  $\beta$ -strands of the TIM-barrel, suggesting that these proteins all have a common origin. Information on the co-crystal structure of chitinase-substrate complexes is scant. Earlier attempts to obtain the co-crystal structures of oligoNAG-chitinase complexes proved unsuccessful. I generated mutations changing the catalytic residues of chitinase A and chitobiase. Biochemical characterization of the mutant proteins allowed us to select a number of mutants for structural studies. I was able to obtain co-crystals with the native substrates and solved their structure. Our investigations led us to suggest that chitinase A and chitobiase employ similar catalytic mechanism.

Multiple alignments of glycosyl hydrolases protein families 18 and 20, show the conservation of residues in the  $\beta$ 4-loop #4; DXXDXDXE in family 18 and HXGGDE in family 20 (X represents any amino acid). Both signatures include Glu315 of chitinase A and Glu540 of chitobiase that act as proton donors. It was found that the relative positions of the catalytic residues DXE and DE are similar in

both enzymes. However, in chitinase A, Asp313 appears, in the wild-type, in two alternative conformations. Comparison of the structures of the enzyme-substrate complexes shows high similarity in the position of the -1 to +1 sugars. It was also observed that in both enzyme-substrate complexes, the planes of the sugars at -1 and +1 are tilted around the scissile bond in a similar manner. In both enzymes a 'chair' to 'boat' conformational change of the -1 sugar was observed. These energetically non-favored structures probably favor the hydrolysis reaction.

Comparison of the structures of the catalytic sites of chitinase A and chitobiase provides a structural explanation for their different functions. Chitinase A possesses a long groove in which an octaNAG substrate is firmly imbedded. This groove allows mainly exochitinase activity on long chitin chains, removing diNAG residues from the reducing end. Chitobiase possesses a tunnel shape active site that may accommodate tetraNAG substrate. This enzyme acts as exochitinase, removing one NAG residue at a time from the non-reducing end of short oligomers. We found that domain-I of chitobiase also participates in substrate binding. This domain is not found in human HexA.

In general, glycosidases degrade carbohydrates by a general acid-base catalysis that involves two amino acid residues, a proton donor and a nucleophile. Hydrolysis of the scissile bond results in either the retention or the inversion of the anomeric configuration of the remaining carbon. While the catalytic glutamate is highly conserved, no amino acid residue that could act as a nucleophile was identified in protein families 18 and 20. A different catalytic mechanism in which the substrate acts as a nucleophile (substrate-assisted mechanism) was proposed. In this mechanism

the glutamate residue acts as a proton donor, while the terminal oxygen (O7) of the acetamido group of the -1 NAG acts as a nucleophile. Nucleophilic attacks of the oxacarbonium C1 by the acetamido O7 result in the formation of oxazolinium ring intermediate. Therefore the acetamido group of the -1 NAG has to occupy a specific position with respect to C1. Our results lead us to propose that in order to act as a nucleophile the acetamido group must rotate around the C2-N2 bond. However, direct evidence for substrate-assisted catalysis is yet lacking.

Structural analysis of alanine replacement mutations in chitinase A and chitobiase provide insight into the function of the conserved Asp313 and Asp539 in chitinase A and chitobiase, respectively. Our results strongly suggest that one function of these Asp residues is to ensure the precise positioning of the acetamido group. In addition, these Asp residues provide additional negative charge at the active site and stabilize the partially positive charge that was distributed on the oxazolinium ring. We further suggest that these Asp residues alternate between two conformations. One conformation facilitates substrate binding while the alternative conformation locks the acetamido group of the -1 NAG in a position that favors the substrate-assisted reaction. I attempted to obtain crystals that would show a trapped intermediate. Utilizing the partially defective mutant D391A of chitinase A, a short soaking time combined with cryo-cooling conditions yielded significant data about possible intermediates of the reaction. These experiments allowed us to identify the trapped cleaved sugar. Furthermore, preliminary indications of the presence of the oxazolinium ring intermediates were obtained.

This study provides further information on the possible evolution of the pathway for chitin metabolism in which the catalytic mechanism and key catalytic residues are conserved. In contrast to previous suggestions, our analysis leads us to favor the hypothesis that the genes coding for protein families 18 and 20 diverged from a common ancestral gene coding for a TIM-barrel domain. These diverge to acquire different substrate specificity but retain the catalytic aspartate and glutamate residues in  $\beta$ 4-loop #4. Evolutionary tinkering led to the establishment of the signatures around the catalytic sites while conserving the Asp-Glu structure required for catalysis. Interestingly, family 19 chitinases, which do not possess a TIM-barrel structure, utilize an alternative, acid-base catalytic mechanism. In several members of family 18, such as the plant seed storage proteins concanavalin B and narbonin, which evolved to lose catalytic activity, the conserved DXXDXDXE motif has been modified. Our findings provide a molecular explanation for the lack of catalytic activity of these proteins.

The biochemical and structural information presented here elucidates how *S. marcescens* employs its enzymes in chitin degradation. Chitin degradation is initiated by the action of endochitinases, probably chitinase A and chitinase C. Chitinase A also acts as an exochitinase, cleaving diNAG dimers from the reduced end. Chitinase B acts as an exochitinase cleaving triNAG and diNAG from the non-reducing end of the chitin oligoNAG chains generated by the action of chitinase A and C. The oligomer, triNAG and diNAG products are subsequently degraded to metabolizable NAG monomers by chitobiase. This analysis clearly explains the synergistic activity of these enzymes that has previously been reported.

Solving the detailed structures of the chitinolytic enzyme-substrate complexes yields deeper understanding of the mechanism of action of these enzymes. These biochemical and structural analyses lead us to propose that both chitinase A and chitobiase share a similar catalytic mechanism.

## List of abbreviations

### Amino Acids one letter code:

**A**-Alanine; **C**-Cysteine; **D**-Aspartate; **E**-Glutamate; **F**-Phenylalanine; **G**-Glycine;

**H**-Histidine; **I**-Isoleucine; **K**-Lysine; **L**-Leucine; **M**-Methionine; **N**-Asparagine;

**P**-Proline; **Q**-Glutamine; **R**-Arginine; **S**-Serine; **T**-Threonine; **V**-Valine;

**W**-Tryptophan; **Y**-Tyrosine

**ARP/wARP** *weighted* Automated Refinement Procedure

**CCP4** Collaborative Computational Project, Number 4

**DE motif** Aspartate – Glutamate residues pair

**DESY** Deutcheland Electron Synchrotron

**EMBL** European Molecular Biology Laboratory

**GCG** Genetics Computer Group

**HexA** *Human*  $\beta$ -*N*-acetyl-hexosaminidase A

**IPTG** isopropyl- $\beta$ -D-thiogalactopyranoside

**FFT** Fast Fourier Transform

**Fc** Calculated Factors

**Fo** Observed Factors

**NAG** N-acetyl- $\beta$ -D-glucosamine

**MR** molecular Replacement

**4MU** 4-Metyl-Umbelliferyl

**PCR** Polymerase Chain reaction

**PDB** Protein Data Bank (of protein structures)

**pNp** para-Nitro-Phenyl

**RCSB**      Research Collaboratory for Structural Bioinformatics

**RMSD** Root Mean Square Deviation

**SDS-PAGE**   Sodium Dodecyl Sulfate Poly Acryl Amide Gel Electrophoresis

**TIM**          Trios IsoMerase phosphate

**X-NAG**      5-Bromo-4-Chloro3-Indolyl-N-acetyl- $\beta$ -D-glucosamin

# Introduction

## Proteins, enzymes and catalytic domains

Proteins have evolved during evolution, through selective pressure, to perform specific functions. It was estimated that about 1000 folds exist in nature (Chothia, 1992). The fact that out of 14,055 solved structures in the Protein Data Bank (PDB 11410 PDB Entries at the 1 July 2000), only 564 unique folds were found, supports this notion (SCOP: Structural Classification of Proteins. 1.53 release “<http://scop.mrc-lmb.cam.ac.uk/scop>”). Assembly of several domains (modules) with different functions into one protein chain yields proteins that can carry complex processes. For example, repressors such as the Lac repressor contain a DNA binding domain, an inducer-binding domain and a region responsible for tetramer formation (See Appendix A2). Thus, the concerted activities of all these domains are essential for Lac repressor activity.

Glycosyl hydrolases were classified into families on the basis of amino acid sequence similarities (Henrissat, 1991). The 3D-structures of proteins are more conserved than their amino acid sequences. Thus, it is expected that proteins in a given family would have a sufficiently similar fold to permit a detailed structure-function analysis based on protein homology modelling. Enzymes with different substrate specificity are sometimes classified in the same family, indicating an evolutionary divergence to acquire a new specificity. For example, chitobiase (*N*-acetyl glucosaminidase), an exochitinase which cleaves *N*-acetyl-Glucosamine residues from oligo-chitin chains, and human hexosaminidase A which cleaves *N*-acetyl-galactosamine from GM<sub>2</sub> molecules, are both members of protein family 20 (see

below). Interestingly, both enzymes can hydrolyze the same chromogenic substrate pNp-NAG. In contrast, members of structurally different protein families can hydrolyze the same substrate. For example, the glycosyl hydrolase families 18 and 19 are chitin-degrading enzymes. Not surprisingly, the underlying catalytic mechanisms of these two protein families are different.

### **TIM-Barrel Protein domain**

One of the most abundant domain structures in nature is TIM-barrel (Triose IsoMerase phosphate), also called  $(\beta/\alpha)_8$ -barrel (Alber *et al.*, 1981). It was estimated that about 10% of the proteins in nature contain a TIM-barrel domain (Hegyi & Gerstein, 1999; Wierenga, 2001). A large number of these proteins are glycosyl hydrolases. The TIM-barrel is made of eight  $\beta$ -strands tethered to eight  $\alpha$ -helices by loops. The loops at the amino termini of the barrel  $\beta$ -strands are usually short, while the loops at the carboxy termini are long and form a unique shape that composes the structure of the active site and determines substrate specificity (Figure I.1). It has been suggested that the TIM-barrel provides a framework which preserves the positions of these loops. It was found that in many TIM-barrel enzymes the catalytic proton donor is located at loop #4 of the barrel (Branden & Tooze, 1991). This observation suggests that TIM-barrel enzymes are evolutionary related and are derived from a common ancestor (divergent evolution – see discussion). Recently, based on sequence and structure similarities, it was suggested that two TIM-barrel proteins in the histidine biosynthesis pathway evolved by gene duplication and gene fusion from a common half-barrel  $(\beta/\alpha)_4$  ancestor (Lang *et al.*, 2000). In addition, it was shown that it is

possible to generate a circular permutation of the coding sequence of an active TIM-barrel enzyme (Branden & Tooze, 1991; Jia *et al.*, 1996).

### **Catalytic mechanism of glycosyl hydrolases**

Hydrolysis is one of the most common reactions in biological systems in which a covalent bond is cleaved. In a few cases, such as hydrolysis of ATP or NADH, this process is coupled with utilization of the chemical bond energy. However in most cases, for example in hydrolytic degradation of polymers such as proteins or polysaccharides, the reaction is not coupled with the specific utilization of the free energy released. However, although hydrolysis is an exothermic reaction, some of the scissile bond energy is transferred via the enzyme during the catalytic process to split a water molecule. Enzymatic hydrolysis of glycosidic bonds is executed via a general acid / base mechanism (Davies & Henrissat, 1995; McCarter & Withers, 1994; Sinnott, 1990). In this mechanism two amino acid residues are active, one as a proton donor and the other as a nucleophile. There are two alternative outcomes of the reaction: an overall retention (also named a double displacement mechanism) in which the configuration of the anomeric carbon is retained after the bond cleavage, and an inversion (also named a single displacement mechanism) in which the configuration of the anomeric carbon is inverted following the cleavage of the scissile bond. In both mechanisms the position of the proton donor is identical within a hydrogen bonding distance from the glycosidic oxygen. However in retaining enzymes the residue acting as a nucleophile is in closer vicinity to the sugar anomeric carbon than in inverting enzymes. Inverting enzymes contain a water molecule between the base and the sugar carbon. The distance between the two catalytic residues is characteristic of the mode

and stereochemistry of the reaction mechanism: 4.8-5.3 Å is typical for retaining enzymes, and 9.0-9.6 Å for inverting enzymes.

### **Chitin and its metabolism**

Chitin, the second most abundant biopolymer on earth (Huber *et al.*, 1995), is composed of units of *N*-acetyl-D-glucosamine (NAG) linked by  $\beta$ -1-4-glycosidic bonds (Figure I.2). This fibrous polysaccharide serves as the major structural component of the exoskeleton of arthropods and it occurs in the cell wall of many fungi. In yeast, chitin maintains the mother-bud junction structure. Chitinases (chitin degrading enzymes) are also widespread in nature and have been found in bacteria, fungi, plants, invertebrates and vertebrates (Perrakis *et al.*, 1994; Perrakis *et al.*, 1993). Chitinases participate in life cycle functions such as cell wall division and morphogenesis, as a defense system against fungi and as a source of NAG for growth. In recent years several genes coding for chitinase were found in humans. These include an acidic chitinase expressed in the stomach and in the lung, that may function as a part of our immune system (Boot *et al.*, 1995; Boot *et al.*, 1999), and chitotriosidase which is secreted by activated human macrophages. Chitotriosidase was found to be markedly elevated in plasma of Gaucher disease patients (Young *et al.*, 1997). Chitinase and hexosaminidase will be discussed below.

### **Chitin metabolism by *Serratia marcescens***

The soil Gram negative bacterium *Serratia marcescens* is one of the most effective bacteria for chitin degradation (Monreal & Reese, 1969). Thus several research groups employ *S. marcescens* as a model for a chitinolytic degradation system. Analysis of the proteins produced by chitin-induced *S. marcescens* revealed a

number of proteins, identifiable by their apparent molecular mass on SDS-PAGE. At least three chitinases, chitinase A, B and C, a chitobiase and a chitin binding protein (CBP21) were found to be involved in the chitin hydrolysis (Suzuki *et al.*, 1998; Warren, 1996; Watanabe *et al.*, 1997). The synergistic action of these enzymes is summarized in the discussion section (and also in our recent review (Prag *et al.*, 2001)). In my thesis I focused on two enzymes of the pathway chitinase A and chitobiase.

### **Chitinase A**

*S. marcescens* chitinase A belongs to the glycosyl hydrolase family 18 (Henrissat, 1991). Chitinase A was cloned and characterized in our laboratory in the late eighties (Shapira *et al.*, 1989). The crystal structure of the native enzyme was solved and refined to a 2.3 Å resolution; however, it was impossible to obtain co-crystals with the oligoNAG substrate (Perrakis *et al.*, 1994; Vorgias *et al.*, 1992). Chitinase A consists of three domains: *i.* an all- $\beta$ -strand amino terminal domain similar to fibronectin type III (FnIII) domain, *ii.* a catalytic TIM-barrel, and *iii.* a small  $\alpha$ + $\beta$ -fold domain. The function of the small  $\alpha$ + $\beta$ -domain is not known (Figure I.3A and I.3B). The FnIII like domain was found in several other glycosyl hydrolase protein families such as cellulases and amylases, and may play a role in polymer binding (Watanabe *et al.*, 1992). Interestingly, it was suggested that the FnIII like domain evolved by horizontal evolution and gene fusion (Bork & Doolittle, 1992; Prag *et al.*, 1997). It was proposed that a long groove, located at the carboxy-terminal end of the  $\beta$ -strands of the TIM-barrel, binds the substrate (Perrakis *et al.*, 1994; Tews *et al.*, 1997). A multiple alignment of family 18 of glycosyl hydrolases, shows the complete

conservation of both Asp313 and Glu315 in  $\beta$ -loop #4 (positions of the amino acid residues are those of *S. marcescens* chitinase A) which are part of a DXXDXDXE conserved motif (Figure I.4). The glutamate residue was proposed to function as the proton donor (Perrakis *et al.*, 1994; Schmidt *et al.*, 1994; Watanabe *et al.*, 1993). It was suggested that these signature motifs participate in the formation of a dipole down through the center of the barrel (van Aalten *et al.*, 2000). Biochemical studies with family 18 chitinases showed that these enzymes act as retaining enzymes (Armand *et al.*, 1994).

Two alternative models were proposed for the catalytic mechanism of family 18 chitinases (Perrakis *et al.*, 1994; Terwisscha van Scheltinga *et al.*, 1994). These are the acid-base catalysis, as was first suggested for hen egg white lysozyme, and the substrate-assisted model. In accordance with the first model, two acidic residues were identified as a proton donor (E315) and as a nucleophile (D391). The second model is based on the crystal structure complex of Hevamine (a family-18 plant chitinase) with the chitinase inhibitor allosamidin. According to the substrate-assisted mechanism E315 acts as a proton donor while the acetamido O7 of the -1 NAG acts as a nucleophile (Tews *et al.*, 1996; Tews *et al.* 1997).

### **Chitobiase**

*S. marcescens* chitobiase, also named  $\beta$ -glucosaminidase, is capable of degrading chitobiose (diNAG) (Kless *et al.* 1986). This enzyme can also cleave longer substrates of oligoNAG and artificial substrates such as pNp-NAG. Thus it should also be named exochitinase. The *chb* gene codes for chitobiase, composed of 885 amino acids from which a 28-residues leader peptide is removed upon transport to the

periplasm. The crystal structure of native chitobiase was solved and refined to 1.9 Å (Tews *et al.* 1996). The structure revealed four domains tethered by flexible linkers (Figure I.3). Domain I (residues 28-181) is very similar to the cellulose binding domain (CBD) of *Cellulomonas fimi* cellulase, and is comprised of two  $\beta$ -sheets with two and a half turn helices. Domain II (residues 214-335), has a topological similarity to a metalloproteinase (serralysine) from *S. marcescens*. This domain is probably conserved in all protein members of the family, however its function is yet unknown (Mark *et al.*, 2001). Domain III is the catalytic TIM ( $\alpha/\beta$ )-barrel (residues 336-765). Domain IV is relatively small (67 amino acids; residues 819-885) and is similar to the  $\beta$ -sandwich immunoglobulin proteins.

Similarly to chitinase A, it was suggested that hydrolysis by chitobiase proceeds via a substrate-assisted mechanism in which the configuration of the anomeric carbon is retained. In this reaction model glutamate 540, located in loop #4 of the catalytic domain, acts as a proton donor and the acetamido group acts as the nucleophile. (Koshland Jr., 1953; Tews *et al.*, 1996; Drouillard *et al.*, 1997).

Multiple alignment of family 20 of glycosyl hydrolases, based on 17 sequences extracted from the SwissProt bank, revealed a conserved signature of six residues HXGGDE (where X represents any amino acid) in  $\beta$ -loop #4 (see Figure I.4).

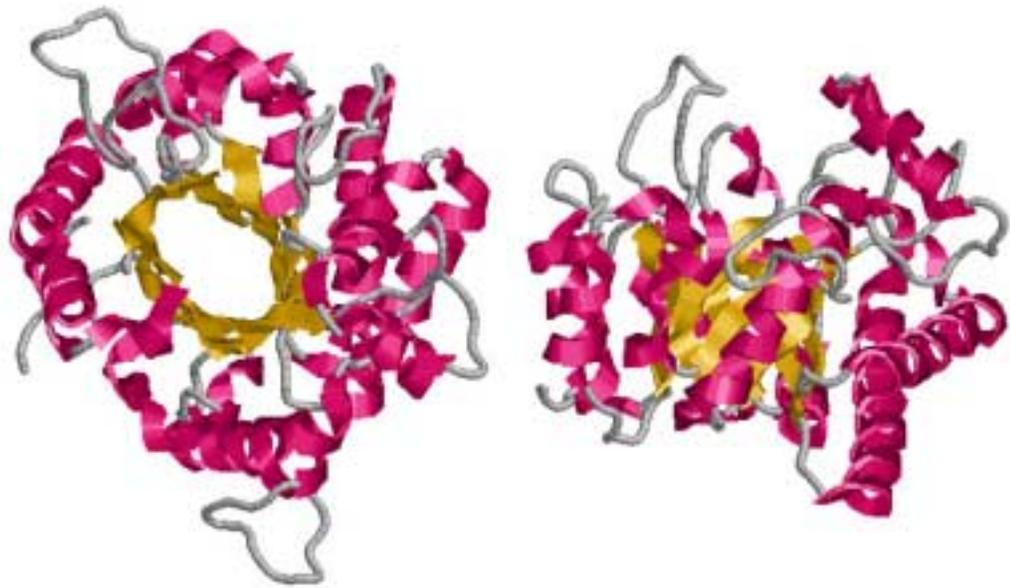
### **Conservation of chitobiase and human hexosaminidase A**

Human  $\beta$ -hexosaminidases, belonging to the glycosyl hydrolase protein family 20, catalyze the cleavage of terminal  $\beta$ -N-acetylglucosamine or  $\beta$ -N-acetylgalactosamine residues from glycoconjugates [reviewed in (Gravel *et al.*, 1995; Gravel *et al.*, 1991)].  $\beta$ -hexosaminidase A (HexA), is a heterodimer made of  $\alpha$  and  $\beta$

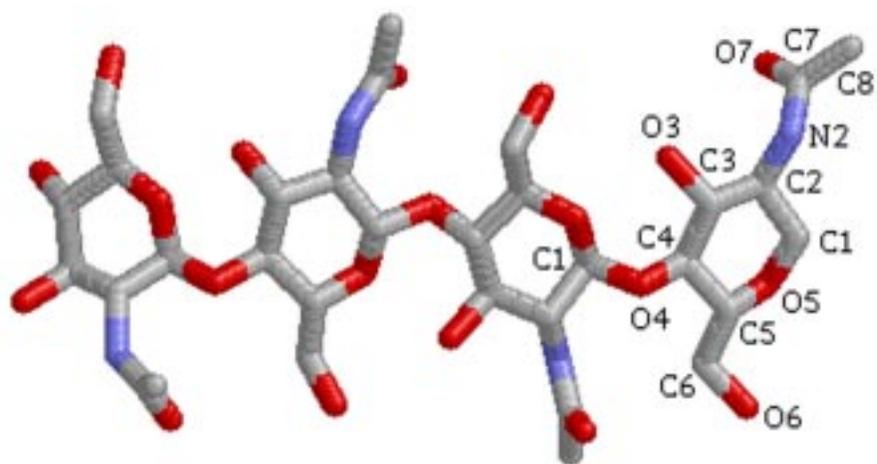
subunits coded by the *HEXA* and *HEXB* genes respectively. Hexosaminidase B (HexB) is a homodimer of the  $\beta$  subunit. The homodimer of the  $\alpha$  subunit is an inactive enzyme. An important substrate of HexA is GM<sub>2</sub>-ganglioside, a major constituent of neural membranes. Deficiencies in these enzymes lead to accumulation of gangliosides, resulting in gangliosidosis such as Tay-Sachs or Sandhoff diseases. Numerous mutations in human HexA have been identified.

Structural modeling of the HexA catalytic domain, based on the bacterial chitobiase, was constructed and the Tay-Sachs and Sandhoff mutations were analyzed (Tews *et al.*, 1996a). Point mutations located around the active site show severe phenotypes, while mutations located far from the active site show less severe phenotypes. In this study we tested further the hypothesis of structural homology by the introduction of specific Tay-Sachs mutations into the *chb* gene (see Chapter I). In parallel to our study, further experimental support for the homology modeling of the two enzymes was obtained by mutational replacement of conserved residues in HexA (Fernandes *et al.*, 1997).

In this thesis, I have focused on two enzymes, chitobiase and chitinase A. I have carried out bioinformatic, genetic, biochemical and structural studies with the aim of answering key questions about the mechanisms of action of these enzymes. Co-crystallization with natural substrates allowed us to elucidate the mode of substrate binding in both enzymes. Our findings support the hypothesis that both enzymes utilize substrate-assisted mechanisms.

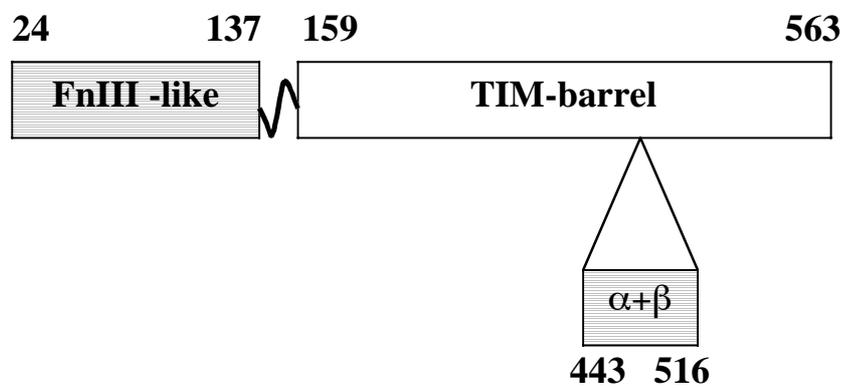


**Figure I.1. Scheme showing the architecture of ( $\alpha/\beta$ ) TIM-barrel domain.**  $\beta$ -strands and  $\alpha$ -helices are colored yellow and pink respectively, loops are colored gray. The shown structure is the catalytic domain of chitinase A. The right view shows a  $90^\circ$  axial rotation of the left view. Loops at the upper side of the right view were predicted to generate the active site formation (Perrakis *et al.* 1994)



**Figure I.2. Structure of chitin.** The structure shows four residues of *N'*-acetyl- $\beta$ -D-Glucosamine linked by  $\beta$  1- $\rightarrow$ 4 glycosidic bond. Carbon, oxygen and nitrogen atoms are indicated at the reduced end of the chain. Color code: carbon-gray; oxygen-red; nitrogen-light blue. The structure shown is from the chitinase A D313A mutant complex with octaNAG (see Chapter IV).

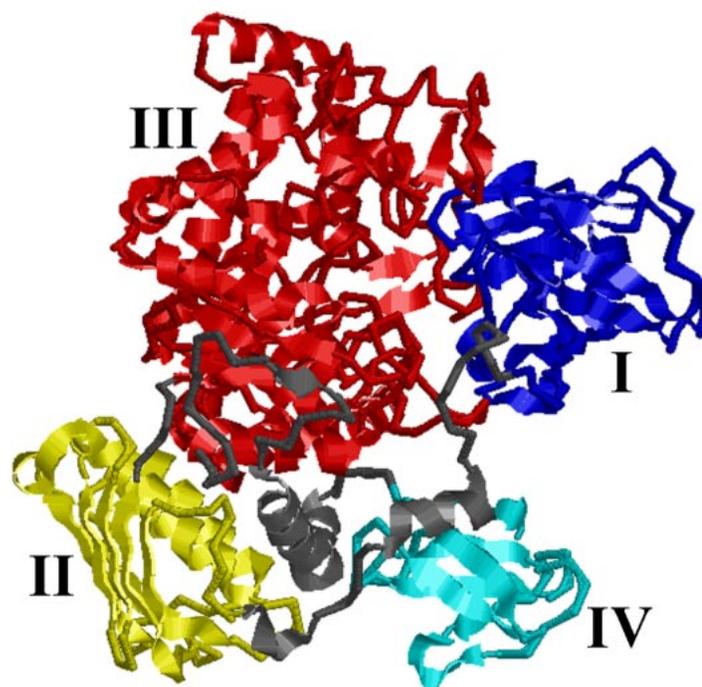
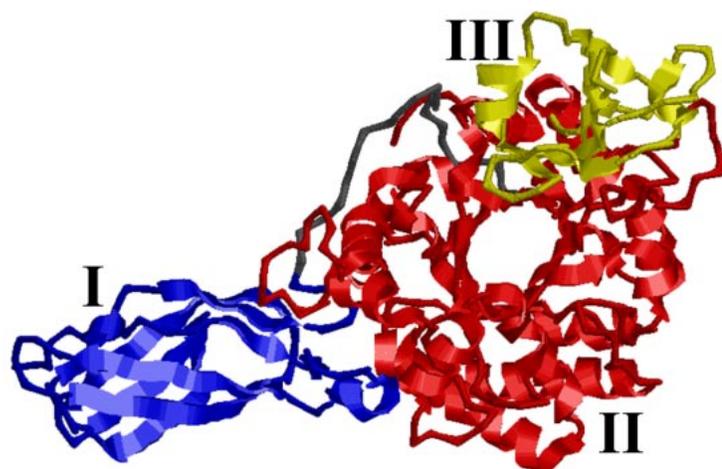
## Chitinase A



## Chitobiase



Figure I.3A. Domains of chitinase A and chitobiase



**Figure I.3B. Domain structures of chitinase A and chitobiase.** A ribbon scheme showing the crystal structures of chitinase A and chitobiase.

Family 20			Family 18								
*	CHB_SERMA	529	QPIKTWHFGG	DEAK		*	CHIA_SERMA	307	FDGVD	IDWE	FPP
*	HEX_STRPL	303	TPGRYLHIGG	DEAH		*	EBAG_STRPL	124	LDGVDF	DD	EYA
	HEXA_HUMAN	312	FPDFYLHLGG	DEV		*	EBA1_FLAME	120	LDGVFF	DD	EYS
	HEXB_HUMAN	344	FPDQFIHLGG	DEVE		*	CHLY_HEVBR	119	LDGIDF	DI	EHG
	HEX_VIBVU	508	QPLTDYHIG	DETA		*	CHIB_SERMA	136	FDGVD	IDWE	FPP
	CHB_VIBHA	526	APLTTWHFGG	DEAK			CHIC_SERMA	133	FDGL	DIDL	EQA
	HEXB_MOUSE	323	FPDQFIHLGG	DEVE			CHID_BACCI	295	FNGL	DIDL	EGS
	HEXA_MOUSE	312	FPDFYLHLGG	DEV			CHIX_STROI	376	LDGL	DIDF	EGH
	HEXB_ALT	520	VPLNTYHIG	DETA			CHI1_APHAL	163	FDGID	IDWE	FPP
	HEXB_FELCA	286	FPDHFVHLGG	DEVE			CHI1_BACCI	196	FDGV	DL	WE
	HEX1_CANAL	313	FIDVVFHVGN	DELQ			CHI1_CANAL	139	VDGF	DF	ENK
	HEX1_ENTHI	290	FGTDYVHVGG	DEW			CHI1_COCIM	163	FDGID	IDWE	FPP
	HEXA_DICDI	297	FIDNYFHTGG	DELV			CHI1_RHIOL	158	IDGV	DL	EGG
	HEXC_BOMMO	359	ESTDMFHMGG	DEVS			CHI2_COCIM	167	VDGF	DF	IEKG
	HEXA_PORGI	324	FPGTYFHIGG	DECP			CHI3_CANAL	149	LDGF	DF	ENN
	STRH_STRPN	349	KKTEIFNIGL	DEYA			CHI4_TRIHA	163	FDGID	IDWE	FPP
	STRH_STRPN	793	GKTKIFNFGT	DEYA			CHIA_ALT	305	YDGV	IDWE	FPP
							CHIA_ARATH	148	LDGID	FN	ELG
							CHIA_CICAR	142	LDGID	FD	ESG
							CHIA_CUCCA	144	LDGV	DF	ESG
							CHIA_PHAAN	148	LDGV	DF	ESG
							CHIA_TOBAC	141	LDGID	FD	EGG
							CHIA_VITVI	144	LDGID	FD	ELG
							CHIB_TOBAC	143	LDGID	FD	ELG
							CHIE_BETVU	143	LDGID	FD	ESG
							CHIT_BRUMA	140	FDGF	DL	WE
							CHIT_CAEEL	171	FDGID	IDWE	FPP
							CHIT_MANSE	138	FDGL	DL	WE
							CHIT_STRLI	374	FDGID	LD	WE
							DIAC_HUMAN	135	MDGIN	ID	EQE
							DIAC_RAT	120	MDGIN	ID	EQE
							CHI1_RHINI	156	IDGID	LD	EGG
							CHI2_CANAL	145	VDGF	DF	IEHG
							CHI2_RHIOL	158	IDGV	DL	EGG
							CHIT_NPVAC	297	FDGV	IDWE	FPP
							CHIT_STRPL	375	FDGID	LD	WE
							CHIT_YEAST	149	VDGF	DF	ENN
							EBA3_FLAME	159	LDGIS	LD	EHS
							EBAG_FLASP	120	LDGV	DL	DEYS
						**	C3L1_HUMAN	131	FDGL	DL	LAWLYP
						**	OGP_HUMAN	133	FDGL	DL	FFLYP
						** *	NARBONIN	124	IDGID	DIHY	EHI
						**	*CONB_CANEN	123	LDGI	HF	DIQKP

**Figure I.4 Sequence alignment of protein families 18 and 20 showing the conserved signatures.** Multiple alignment was performed as described in the Materials and Methods section. \* Marks proteins whose structure was solved; \*\* Proteins with no chitinase activity. The catalytic glutamic residues are shown in red and the conserved aspartic residues in blue.

## Materials and Methods

**Table M.1** List of substrates for enzymatic and crystallographic studies

Short name	Chemical name
<b>Chitin</b>	poly- <i>N</i> -acetyl- $\beta$ -D-glucosamine (from crab shells)
<b>diNAG</b>	$N^I, N^{II}$ -Diacetylchitobiose
<b>tetraNAG</b>	$N^I, N^{II}, N^{III}, N^{IV}$ -Tetraacetylchitotetraose
<b>hexaNAG</b>	$N^I, N^{II}, N^{III}, N^{IV}, N^V, N^{VI}$ -Hexaacetylchitohexaose
<b>octaNAG</b>	$N^I, N^{II}, N^{III}, N^{IV}, N^V, N^{VI}, N^{VII}, N^{VIII}$ -Octaacetylchitooctaose
<b>4MU-NAG</b>	4-Methylumbelliferyl- $\beta$ -D- $N^I$ -acetylglucosamine
<b>4MU-diNAG</b>	4-Methylumbelliferyl- $\beta$ -D- $N^I, N^{II}$ -Diacetylchitobiose
<b>4MU-triNAG</b>	4-Methylumbelliferyl- $\beta$ -D- $N^I, N^{II}, N^{III}$ -Triacetylchitotriose
<b>4MU-tetraNAG</b>	4-Methylumbelliferyl- $\beta$ -D- $N^I, N^{II}, N^{III}, N^{IV}$ -Tetraacetylchitotetraose
<b>pNp-NAG</b>	p-Nitrophenyl- $\beta$ -D- $N^I$ -acetylglucosamine
<b>pNp-diNAG</b>	p-Nitrophenyl - $\beta$ -D- $N^I, N^{II}$ -Diacetylchitobiose
<b>pNp-triNAG</b>	p-Nitrophenyl - $\beta$ -D- $N^I, N^{II}, N^{III}$ -Triacetylchitotriose
<b>X-NAG</b>	5-Bromo-4-Chloro-3-Indolyl- $N^I$ -acetyl- $\beta$ -D-glucosamine

All the above substrates were purchased from Sigma.

**Table M.2 List of DNA oligonucleotides**

Nucleotides sequence	<sup>a</sup> Description	<sup>b</sup> Lab. #
<b>1. <i>chb</i></b>		
5' TCGCGATAACAATAATTAAGAAGGATG	cloning	A28
5' GAAGATCTCGGCGACTCCACGTACTGCTCTC	cloning	A29
5' CTGGCACTTTGGCGGCGCGGAGGCGAAAAACATC	D539A	A25*
5' GCACTTTGGCGGCCAGGAGGCGAAAAACATCCGC	D539N	A52*
5' GCACTTTGGCGGCGATGATGCGAAAAACATCCGC	E540D	A1 *
5' GCACTTTGGCGGCGATGCGGCGAAAAACATCCGC	E540A	A14*
5' GCACTTTGGCGGCGCGGCGGCGAAAAACATCCGC	D539A-E540A	A41*
5' CAACCCGGACTACGTGCAATGGACTTCCCTTACG	Y669E	A90*
5' CAACCCGGACTACGTCTTCATGGACTTCCCTTAC	Y669F	A159
5' CCTCGACGTGGCGGCCAATTCCATAAGAAGGACG	R349A	A10*
5' GGAAGTTGTGCGCCACGTCG	R349H	A25
5' GGTGATCCCGGAGATCCACATGCCGGCGCACGC	D448H	A3 *
5' AAGCCGCCATCAGATCCAGC	Q363L	A26
5' GCCGGTTCAGACCAACG	S&RM	1794
5' GGATCACCTCAATCTGG	S&RM	1795
5' CGGCGACTCCACGTACTGCTCTC	S&RM	A12
5' GCGCGCTTGCCCGCTTGACGGC	S&RM	A15
5' CCGCGCTACCGCCCATAGCCCGC	S&RM	A16
5' CGGCGGACGTCGTCAGTCAGCG	S&RM	A17
<b>2. <i>chiA</i></b>		
5' GACCTGGAAGTTCTTCGCCGGCGTGGATATCGAC	D308A	A146
5' GTTCTTCGACGGCGTGGCTATCGACTGGGAGTTC	D311A	A147
5' GCCCGGGAAGTCCCACGCGATATCCACGCCGTC	D313A	A70
5' TTGCCGCCCCGGAACGCCAGTCGATATCCACGC	E315A	A66
5' TTGCCGCCCCGGAAGTCCAGTCGATATCCACGC	E315Q	A64
5' CCGGCCGCAAGTATGCGCTGACCTCCGCCATCAG	E358A	A149
5' CTTCTGATGAGCTTCGACTTCTATGGCGC	Y390F	A158
5' CTTCTGATGAGCTACGCCTTCTATGGCGC	D391A	A148
5' CGAAGGCGCCATACAAGTGGTAGCTCATCAGGAA	D391N	A74
5' CGAAGGCGCCATAGAAGTTCGTAGCTCATCAGGAA	D391E	A68
5' GTGTGGAATTGTGAGCGG	S&RM.	A123
5' CACCgACggCAgCCATTTg	S&RM.	A124
5' GAAGCAGGCGCATCCTGAC	S&RM	A125
5' CATCAGACCGCGCTGAATGC	S&RM	A126
5' CAGCGCGGTCTGATGCCCCAGG	S&RM	A86
5' CGTCGGTTCGGTGAAAGAGTTCC	S&RM	A87
<b>3. <i>celF /chbF</i></b>		
5' CGCGGATCCAGCCAGAAATTAAGATCGTCAC	cloning	A76
5' CGTCTCGAATTTTTTCGTGTAATTCATGGGG	cloning	A77

The above oligonucleotides were used for cloning, site directed and random mutagenesis and sequencing. <sup>a</sup> Description of the procedure; <sup>b</sup> the serial number of the oligonucleotides in our laboratory; S&RM is Sequencing and Random Mutagenesis; \* Complementary primers were used for mutagenesis (not shown). All these oligonucleotides were purchased from Biotechnology General Israel.

**Table M.3 List of bacterial strains**

<b>Strain</b>	<b>Genotype</b>	<b>Source</b>
XL1B	(company catalog)	Stratagene
XL2B	(company catalog)	Stratagene
JM109	(company catalog)	NEB
A5039	K37 lacZ::Tn5 lacI <sup>q</sup>	Our collection
A2097	C600 r <sup>-</sup> m <sup>+</sup> <i>pro thr leu</i> B [1857 $\Delta$ BamHI $\Delta$ HI] lacZ::Tn10	Our collection
A9262	JM109/pCBII	Our collection
A9263	XL1B/pGPchbQ363L	This work
A9264	XL2B/pGP2	This work
A9266	XL1B/pCB-E540D	This work
A9267	JM109/ptacChiA	Our collection
A9268	A5039/ pCBII	This work
A9269	A5039/ pGP3	This work
A9270	A5039/ pCB-E540D	This work
A9279	JM109/pGPchb	This work
A9287	XL2B/pGPchb	This work
A9288	XL2B/pGPchbKM	This work
A9300	XL1B/pGPchbKM - D448H	This work
A9301	XL1B/pGPchbKM – D539A	This work
A9302	XL1B/pGPchbKM – E540A	This work
A9445	XL2B/pGP3	This work
A9475	XL1B/pGPchb-D539A,E540A	This work
A9477	A2097/pPLchiA-D313A	This work
A9478	XL1B/pCB-R349H	This work
A9492	XL1B/pGPchiA-E315Q	This work
A9493	XL1B/ pGPchiA-D391E	This work
A9499	XL1B/ pGPchiA- D313A ,D391E	This work
A9586	XL2B/pGPchiA-D308A	This work
A9587	XL2B/ pGPchiA-E356A	This work
A9588	XL2B/ pGPchiA-D391A	This work
A9590	XL2B/ pGP-His-chbF	This work
A9591	XL2B/ pGP-His-chbF:Kan <sup>R</sup>	This work
A9592	XL2B/pUC4K	This work
A9610	DY329 (Yu <i>et al.</i> 2000)	Court D.L.
A9611	A9610 celF/chbF::Kan <sup>R</sup>	This work
A9612	A9611/ pGP-His-chbF	This work
A9614	A9611/ pGPchiA	This work
A9615	A9610/pUC4K	This work
A9616	A9611/ pGPchbF	This work
A9617	A9610/pGPchiA	This work
A9618	A9610/pGPchiA-E315Q	This work
A9619	A9610/pKK223-3	This work
A9621	XL2B/pGPchiA-Y390F	This work
A9622	XL2B/pGPchiA-D391N	This work

**Table M.4 Plasmid constructions**

<b>Plasmid</b>	<b>Derived from</b>	<b>Source</b>
pKK177-3	pKK233-3	Altuvia S.
pUC18	pBR322 and M13mp18	NEB
pUC4K	pUC19 and <i>Kan<sup>R</sup></i> cassette	NEB
pKK-223-3	pBR322 and tac promoter	Pharmacia
pQE30	pBR322	QIAGEN
pCBII	pEMBL18	Our collection
pGPchb	pCBII	This work
pGPchbQ363L	pGPchb	This work
pGPchbR349H	pGPchb	This work
pGPchbR349L	pGPchb	This work
pGPchbD448H	pGPchb	This work
pGPchbD539A	pGPchb	This work
pGPchbE540A	pGPchb	This work
pGPchbE540D	pCBII	This work
pGPchbD539A/E540D	pGPchbD539A	This work
pGPchbY669E	pGPchb	This work
pGPchbD539A/Y669E	pGPchbD539A	This work
pGPchbKM	pUC4K and pGPchb	This work
pGPchbKM-D539A	pUC4K and pGPchbE540D	This work
pLchiA	pSA100 and <i>chiA</i> gene	Our collection
pGPchiA	pKK-233-3 and pLchiA	This work
pGPchiAD308A	pGPchiA	This work
pGPchiAD313A	pGPchiA	This work
pGPchiAE315A	pGPchiA	This work
pGPchiAE315Q	pGPchiA	This work
pGPchiAY390F	pGPchiA	This work
pGPchiAD391A	pGPchiA	This work
pGPchiAD391N	pGPchiA	This work
pGPchiAD391E	pGPchiA	This work
pGPchiAF396A	pGPchiA	This work
pGPchiAD391E/D313A	pGPchiA D313A	This work
pGP-His-chbF	pQE30 and <i>celF/chbF</i> gene	This work
pGP-His-chbF- <i>Kan<sup>R</sup></i>	<i>Kan<sup>R</sup></i> insertion into <i>chbF</i>	This work
pGPchbF	pKK177-3 and <i>celF/chbF</i> gene	This work

Plasmid constructions performed in this work are described in the Methods section.

### **Construction and expression of chitobiase mutants**

The *chb* gene was subcloned into the expression vector pKK177-3 (Amann & Brosius, 1985), by PCR, using *Pow* DNA polymerase (Boehringer) and plasmid pCBII as a template (Kless *et al.*, 1989), to yield plasmid pGPchb. Substitutions were introduced into the *chb* gene using a modified protocol of QuickChange site directed mutagenesis kit (Stratagene). The PCR step was performed with 0.5  $\mu$ M of each primer (33-36 bp), 0.2 mM dNTP mix, 0.2  $\mu$ g pGPchb DNA template, 0.1 U of *Pfu*-DNA polymerase in 1 X reaction buffer at total volume of 25  $\mu$ l. The duration of synthesis was 14 min at 68° C followed by the digestion of the parental DNA for 2 hours at 37° C with *DpnI* enzyme. DNA was introduced into *E.coli* XL1B or XL2B cells by electroporation and colonies were screened with 5-Bromo-4-Chloro-3-Indolyl N-Acetyl- $\beta$ -D-Glucosaminide (X-NAG) by pouring 2-3 ml of 1.25 mg x ml<sup>-1</sup> X-NAG in soft agar (0.6% agar dissolved in LB medium). Mutant clones that were obtained at high frequency (more than 85%) were purified and confirmed by DNA sequencing.

To test the importance of the conserved Asp539 - Glu540 pair for the catalytic activity of chitobiase, we subcloned the *chb* gene and generated a number of mutations by site directed mutagenesis. Mutants were recognized by their inability to convert X-NAG into its insoluble dye and were further confirmed by DNA sequencing. This system provides an easy tool for genetic analysis of  $\beta$ -N-acetylhexosaminidases.

Preliminary screening showed that alanine replacement mutations D539A and E540A led to a drastic reduction in enzymatic activity (approximately 3% of wild-type activity). Similarly, E540D was found to be defective in chitobiase activity

(0.5% of wild-type activity). Almost no residual activity was observed in the double mutant E539A/E540A (0.1% of wild-type activity).

### **Construction and expression of chitinase A mutants**

Site-specific mutations were introduced by PCR. The full-length PCR products of wild type and mutated genes were digested by *EcoRI* and *HindIII* restriction endonucleases and cloned into these sites of the pKK223-3 plasmid to give pGPchiA wild type and mutant clones. Transformation of the ligated DNA was performed by electroporation into *E. coli* XL2B competent cells. Mutant clones were grown on LB plates supplemented with 50 g/ml ampicillin, colloidal chitin, IPTG (Isopropyl- $\beta$ -D-Thio-Galactopyranoside) and X-NAG at 37 ° C. Mutant colonies were recognized by their inability to degrade chitin and by the slow conversion of X-NAG (Figure M.1). We found that these growth conditions lead to the activation of the endogenous *E. coli* phospho-chitobiase. Candidate clones were grown overnight in LB containing 50  $\mu$ g X ml<sup>-1</sup> ampicillin and assayed for chitinase expression by using the substrate pNp-diNAG and by running crude protein extracts of each clone on SDS-PAGE. Wild type and mutant clones were confirmed by DNA sequencing.

### **Screening of chitinase A mutants with low activity**

It was previously shown that *E. coli* colonies that express chitinase A on colloidal chitin plates result in a clear halo around the colonies (Reid & Ogrzydziak, 1981). The highly defective chitinase A cannot degrade the colloidal chitin, thus the media around the colony remains turbid. We found that chitinase A activity can be monitored indirectly by adding X-NAG at a final concentration of 0.005 % to the media. In this screening system very low activity of chitinase A is sufficient to

produce diNAG and tri-NAG oligos that act as an inducer for the *E. coli cel* operon and expression of the *celF/chbF* gene (Keyhani & Roseman, 1997). The ChbF (an enzyme product of the *chbF* gene), in turn cleaves the X-NAG into NAG and an insoluble blue dye (see Figure M.1). While wild type chitinase resulted in colour development within 16 hours after plating, mutants of chitinase with less than 1 % activity developed the colour within 24-28 hours.

### **Cloning and knockout of the *E.coli celF/chbF* gene**

The *E.coli celF/chbF* gene, proposed to have a phospho chitobiase activity (Keyhani & Roseman, 1997), was cloned as a His-tag fused protein in the pQE system (Qiagen). Two oligonucleotide DNA primers were designed to clone the gene (see Table M.2). PCR was performed using a standard procedure of Dynazyme DNA polymerase. A genomic DNA template was prepared from *E. coli* XL2B strain for the PCR amplification. The PCR product and the pQE30 vector were digested with the *Bam*HI and *Kpn*I restriction endonucleases, and subsequently the fragments were ligated to form the pGPchbF-His plasmid. The ligated DNA plasmid was transformed into competent *E. coli*-XL2B cells by electroporation and several clones were picked for further study. One clone, showing protein production at the proposed size of 51.3 KDa as judged by 10 % SDS-PAGE, was sequenced to ensure wild type genotype cloning.

Construction of a knockout *chbF* strain was done with the recently developed recombination system, based on bacteriophage lambda RED proteins (Yu *et al.*, 2000). pGP-His-chbF plasmid DNA was digested with the restriction endonuclease *Nsi*I that cut the plasmid only at the nucleotide number 429 of the *chbF* gene. A kanamycin

resistant cassette was prepared by digesting the pUC4K plasmid with *Pst*I, and purification of the DNA fragment of the cassette region from agarose gel. Ligation of both DNA fragments results in an insertion of the kanamycin resistant cassette into the *chbF* gene. By performing a PCR with the primers we used for the gene cloning, a linear DNA fragment was obtained. This DNA contained an inserted kanamycin resistant cassette flanked by large regions of the *chbF* gene (pGP-His-*chbF*-*Kan*<sup>R</sup>). The linear DNA was transformed by electroporation into the DY329 lambda induced competent cells (Yu et al., 2000). Several hundred clones were found on LB-kanamycin plates. Ten colonies were tested for *in vivo* ChbF activity and determined to be unable to convert the X-NAG into insoluble blue dye following the induction of the *cel / chb* operon, as we showed for several *E. coli* strains (Figure M.1).

### **Random mutagenesis and creation of combinatorial DNA library**

We searched for mutants with alter specificity, in which chitinase A would acquire a chitobiase activity (i.e. the mutant enzyme will be able to cleave diNAG or X-NAG into monomers). I used a PCR technique to randomly introduce mutations into the *chiA* gene (Fromant *et al.*, 1995). The PCR was performed with unbalanced concentrations of nucleotides in the presence of manganese (final concentrations of 1 mM dATP, 0.02 mM dGTP, 0.2 mM dCTP and 0.2 mM dTTP) with 0.1mM MnCl<sub>2</sub> and 6 mM MgCl<sub>2</sub>. Moreover, the DNA polymerase enzyme that I used was TaqPolymerase, which has no proof editing system. The combination of unbalanced concentrations of nucleotides with the addition of manganese to the reaction mixture, promotes the incorporation of mistakes during elongation. It is impossible to use this mutagenesis technique on full length DNA fragment coding for chitinase A (1.7 Kbp)

(Fromant et al., 1995). Thus I performed the PCR on small fragments (about 200-400 bp each) and used these PCR products to construct a combinatorial mutated library of the *chiA* gene. The library was constructed by modifying the DNA shuffling technique (Zhang *et al.*, 1997). Series of overlapping primers that result in DNA fragments of about 200-400 bp with overlapping regions of about 20-30 bp were synthesized and subjected to random PCR as described above. The PCR products were these purified from agarose gel and DNA fragments were used for a second PCR without primers and templates, allowing the re-assembly of the fragments into full length DNA encoding for the *chiA* gene. A third PCR amplification was performed with external primers on the final full length reassembled DNA products. The DNA product coding for chitinase A was digested with restriction endonucleases and ligated into the pKK223-3 vector, as described for the construction of pGPchiA. A similar technique was employed with chitobiase.

### **DNA sequencing and sequence corrections**

All mutated clones were checked by DNA sequencing with the help of Dr. Korner M., as part of the service of the DNA analysis unit of the Hebrew University. Our electron density maps, of both mutants, suggested that residues 484 and 566 were incorrectly assigned as Pro and Gly respectively (Tews *et al.*, 1996a). Resequencing the appropriate coding regions showed that these positions code for Gln and Ser respectively, a conclusion supported by the electron density maps. Similarly, sequence mistakes of five amino acid residues were corrected in chitinase A.

### **Purification of chitobiase**

Induced culture (5 liters) was harvested using a low speed centrifuge. Cells were resuspended in 20 ml of 0.5 M Sucrose 20mM Tris-HCl pH 8.0 and 0.2 mg x ml<sup>-1</sup> lysozyme, and incubated for 1 hour at 4°C. After centrifugation, NH<sub>4</sub>SO<sub>4</sub> was brought to 2 M and the periplasmic proteins were applied on a Phenyl-Sepharose column, pre-equilibrated with 2 M ammonium sulfate and 20 mM Tris-HCl buffer pH 8.0. After washing with the same buffer, the protein was eluted by a descending gradient of ammonium sulfate (1.2 to 0 M) and Tris-HCl pH 8.0 buffer. The volume of the solution containing the protein was reduced using an Amicon filter. After dialysis against a large volume of 10 mM sodium phosphate buffer pH 8.0, the protein was applied on SP-Sepharose column. Bound protein was eluted with 0-0.4 M NaCl, 10 mM sodium phosphate buffer pH 8.0 gradient. The location of the fraction containing the enzyme was determined by assaying for the hydrolysis of pNp-NAG. The fractions with highest activity were checked by 12.5% SDS-PAGE (Laemmli, 1970). The protein was concentrated by ultra filtration and the protein concentration was determined by the Bradford reagent.

### **Purification of chitinase A**

Cell cultures (5 L) were induced by IPTG, harvested by low speed centrifugation (3000 g), and chitinase A was purified from the periplasmic fraction. About 8 grams of cells were resuspended in 20 ml of 0.5 M sucrose 20 mM Tris-HCl pH 8.0 and 0.2 mg X ml<sup>-1</sup> lysozyme, and incubated for 1 hour at 4°C. After centrifugation, the periplasmic proteins were applied on a Sepharose CL-6B column, eluted by a descending gradient of ammonium sulfate (1.2 to 0 M) in Tris-HCl pH 8.0

buffer, concentrated by ultrafiltration and applied on a SP-Sepharose column pre-equilibrated with 20 mM acetic acid buffer pH 4.8 and 15% PEG 6000. Bound proteins were eluted by increasing the NaCl gradient and decreasing the PEG concentration. Finally the protein was resuspended in a large volume of water and applied on a hydroxyl-apatite column, extensively washed with 0.5 mM NaCl and eluted with 1.2 M ammonium sulfate. The pure proteins were concentrated to a final concentration of about 40 mg X ml<sup>-1</sup> as measured by using Bradford reagent.

### **Purification of ChbF**

An induced culture (3 liters) of A9612 cells was harvested using a low speed centrifuge. Cells were resuspended in 5 ml of 50 mM MnCl<sub>2</sub>, 20mM Tris-HCl pH 8.0 and 0.2 mg x ml<sup>-1</sup> lysozyme, and incubated for 1 hour at 4°C. The suspension volume was enlarged to 50 ml with 50 mM MnCl<sub>2</sub>, 10mM imidazole and 20mM Tris-HCl pH 8.0 and applied on highly cross-linked Ni<sup>2+</sup>-column using FPLC. The protein was eluted by imidazole gradient (0 –200mM). The volume of the solution containing the protein was reduced using an Amicon filter. After dialysis against a large volume of 50 mM MnCl<sub>2</sub>, 1.5 M ammonium sulfate and 20mM Tris-HCl pH 8.0, the protein was applied to a Phenyl-Sepharose column, pre-equilibrated with the same buffer. After washing, the protein was eluted by a descending gradient of ammonium sulfate (1.2 to 0 M) with the same buffer. The volume of the solution containing the protein was reduced using an Amicon filter and Milipore for final concentration.

### **Crystallization of chitobiase mutants**

We have improved the crystallization protocol, (Tews *et al.*, 1992) as follows: Co-crystals / crystals were grown by the hanging-drop vapor diffusion method

(Figure M.2A). Reservoir buffer contained 2.3 M ammonium sulfate and 100 mM sodium acetate buffer pH 4.8. The aqueous protein solution 40 mg X ml<sup>-1</sup> was mixed with an equal volume of reservoir buffer containing 10 mM diNAG (5 mM for tetra and hexaNAG). Crystals of approximately 0.5 X 0.3 X 0.2 mm<sup>3</sup> were formed within 2-3 days.

### **Crystallization of chitinase A mutants**

Co-crystals / crystals were grown by hanging-drop vapor diffusion. Drops of 6 µl containing 4 µl of 0.5 mM (30 mg X ml<sup>-1</sup>) protein and 2 µl of reservoir solution were equilibrated against 1 ml of reservoir solution consisting of 0.75 M sodium citrate (pH 7.2) and 20% (v/v) methanol (Vorgias *et al.*, 1992). In the co-crystallization experiments, the substrates were added (to the drop) to a final concentration of 5 mM. Co-crystals of approximately 0.6 X 0.2 X 0.2 mm<sup>3</sup> were formed within a period of 3-4 days at 18°C (see Figure M.2B).

### **Kinetic analysis**

The kinetic constants,  $K_M$  and  $k_{cat}$ , of wild-type and mutant-enzymes were determined with the substrate analogue p-nitro-phenyl-NAG (pNp-NAG for chitobiase; and pNp-diNAG for chitinase A) at concentrations ranging from 0.5 µM to 5 mM for chitobiase and 10 µM to 1.5 mM for chitinase A. The reactions were performed in 0.1 M of potassium phosphate buffer pH 7.9 at 42° C and monitored for the accumulation of p-Nitrophenol at OD<sub>405</sub> (930 Uvicon spectrophotometer, Kontron Instruments). The kinetic constants were obtained by fitting the

measurements, (average of 3 experiments) of the initial rates of the reactions to the Michaelis-Menten equation using Prism 2.0 software (GraphsPad).

### **Synchrotron data collection under cryo-cooling conditions**

In the soaking experiments we searched for flash cryo-cooling conditions prior to data collection, in order to prevent the development of ice water around the protein crystal. We found that using 5-10 % higher concentrations of the hanging drop solution with 18 % of glycerol protects the protein crystals and forms a glass liquid-like crystal in the mounting loops.

With the co-crystallized complexes, crystals were frozen by immersion into liquid N<sub>2</sub>. Data collections were carried out at 100°K. All crystals of chitinase A belonged to the face-centered orthorhombic space group  $C222_1$  and exhibited only minor variations in their unit cell dimensions. With chitobiase, all crystals studied belonged to the primitive space group  $P2_12_12_1$ , and exhibited only minor variations in their unit cell dimensions as well.

Diffraction data were collected from a single crystal of chitobiase D539A mutant complex under cryo-cooling conditions (100°K) at the EMBL X11 synchrotron beamline at the DORISIII storage ring of DESY, Hamburg. The resolution range of the data was 10-1.8 Å (Table 2.2). The data of a single crystal of the mutant E540D complex were collected at the home facility (at the laboratory of Dr. Petratos K. IMBB FORTH Heraklion) using a Rigaku rotating anode (Cu-K $\alpha$ ) X-ray generator (T=100°K) within resolution limits of 10-1.9 Å (Table 2.2). Data processing was performed using *DENZO* and equivalent Bragg reflections were

merged using SCALEPACK from the HKL package (Otwinowski & Minor, 1997). Figure M.3 shows a representative diffraction image.

All data of chitinase A co-crystal mutants were collected at the EMBL synchrotron, except one data set of the D391A-hexaNAG complex that was done at the laboratory of Dr. Petratos K. in Heraklion. In the soaking experiments we used the EMBL X11 / X13 and X31 synchrotron beamlines. Data processing was done with the XDS program (Kabsch, 1993).

### **Molecular replacement and model refinement**

Molecular replacement (MR) was carried out with AmoRe (Navaza, 1994) using the native chitobiase structure (PDB code: 1qba) as a model; or chitinase A coordinates (PDB code: 1ctn). Prior to refinement, 5% of the data were randomly flagged for cross validation ( $R_{free}$ ). Refinement was performed under restrained conditions using the programs REFMAC (Murshudov *et al.*, 1997) and ARP / wARP (Lamzin *et al.*, 1993) from the CCP4 suite (Bailey, 1994) until convergence of the indices  $R_{factor}$  and  $R_{free}$  was reached (Table M.5 and Table 2.2 in the Results section). Fourier maps with coefficients of  $2mF_{obs}-DF_{calc}$  and  $mF_{obs}-DF_{calc}$  were calculated, where  $m$  is the figure of merit and  $D$  is the error distribution derived from the  $\sigma_A$  function (Read, 1986). Manual interventions were carried out using both electron density maps that were inspected with 'O' (Jones *et al.*, 1991) in order to check the agreement of the model with the X-ray data. Finally, the stereochemistry of the model (structural validation) was analyzed using PROCHECK (Laskowski *et al.*, 1993) and WHAT\_CHECK (Vriend, 1990).

## **Sequence alignment**

A homology search was carried out with blast against the SwissProt protein bank. The first 17 homologous sequences were extracted. Because most of the family 20 glycosyl hydrolases carry several domains besides the catalytic domain, each one of these sequences was aligned with the sequence coding for the TIM-barrel domain of *S.marcescens* chitobiase using the Bestfit program of the GCG [(not shown; (Wisconsin, Genetics Computer Groups)]. In this step I reduced the interference of other domains existing in these sequences to the multiple alignment. Further, the segments that probably code for the TIM-barrel domain were aligned with the clustalW program (Thompson *et al.*, 1994). Multiple alignment was carried out by assignment of secondary structure elements from chitobiase to the penalty options of clustalW. Finally the alignment was inspected by comparing the predicted secondary structure elements of human hexosaminidase A with those of chitobiase (using the PHD server <http://dodo.cpmc.columbia.edu/predictprotein/>; (Rost *et al.*, 1994)), and taking into account known Tay-Sachs mutations and active site residues of chitobiase. The final alignment of the glycosyl hydrolases family 20 TIM-barrel domain is presented in Appendix A1.

## **Modeling of human Hexosaminidase A**

The crystal structure coordinates of chitobiase and the multiple alignment were the initial input used to generate a 3D model of the  $\alpha$ -chain of human hexosaminidase A, with the homology procedure of the *Insight II* package programs (*Insight II*, 1997). The predicted secondary structure of hexosaminidase and the secondary structure elements of the known chitobiase structure allowed me to assign

these elements to the sequence of hexosaminidase. Tay-Sachs mutations that were performed in this work and candidate active site residues were the key points for the structural alignment. The structural alignment was adjusted manually by choosing each of the residues that were assigned, considering the criteria of mutations and Dayhoff evolutionary mutation matrix, minimal interference of secondary structure elements and active site formation residues. Loops were generated according to the loop generating procedures of the program. Those with the lowest RMS value were selected from the 10 different options given by the generator.

**Table M.5 Data collection and refinement statistics of chitinase A co-crystal of mutant proteins with octaNAG**

<b>Data set</b>	<b>E315Q-octaNAG</b>	<b>D313A-octaNAG</b>
<b>A. Diffraction data</b>		
Wavelength (Å)	0.8469	0.8469
Space group symmetry	$C222_1$	$C222_1$
Resolution limit (Å)	10.0 - 1.9	10.0 - 1.8
Data Redundancy <sup>1</sup>	4.1(4.1)	3.7(3.4)
Completeness (%) <sup>1</sup>	99.0(100.0)	97.7(99.0)
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	16.5(5.9)	24.4(7.7)
$R_{\text{merge}}$ (%) <sup>1</sup>	5.0(24.2)	3.4(14.7)
<b>B. Refinement</b>		
$R_{\text{factor}}$ (%)	17.3	17.8
$R_{\text{free}}^2$ (%)	21.7	21.7
No. of protein atoms	4190	4183
No. of hetero-atoms	113	113
No. of ordered waters	806	863
<b>C. R.M.S. deviations from ideal</b>		
Bond length (Å)	0.012	0.012
Bond angle (Å)	0.028	0.026
<b>D. Average B-factor values (<math>\langle \text{Å}^2 \rangle</math>)</b>		
Protein atoms (Å <sup>2</sup> )	22.6	21.8
Substrate (Å <sup>2</sup> )	41.4	43.6

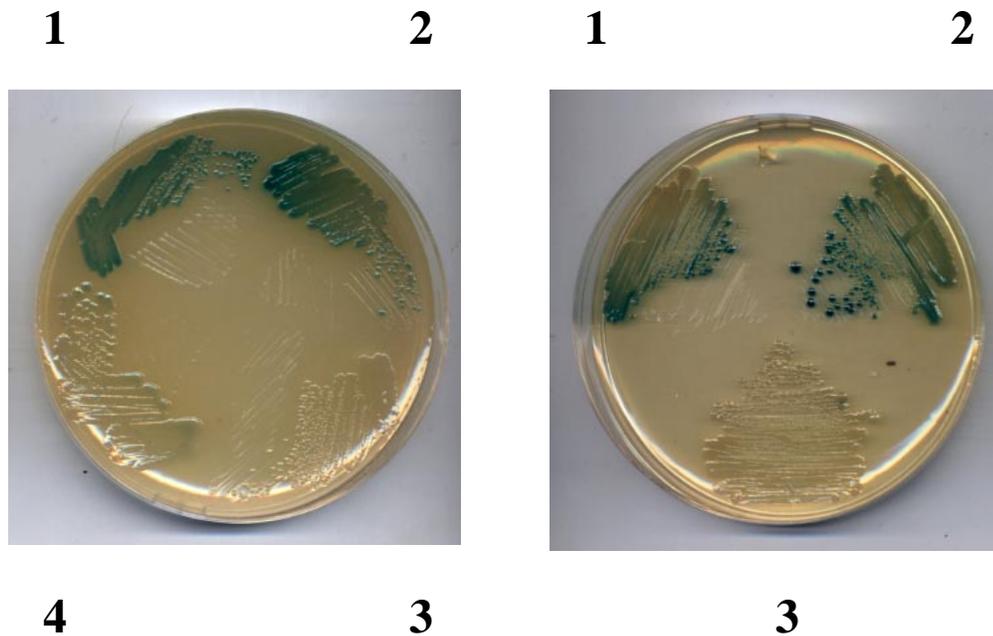
<sup>1</sup>The numbers in parentheses are statistics from the highest resolution shell as reported by Scalepack HKL.

<sup>2</sup> $R_{\text{free}}$  was calculated with a test set of a random 5% of the diffraction data.

**Table M.6 PDB accessions for chitinase A and chitobiase complexes**

<b>Complex structure</b>	<b>Accession</b>	
	<b>PDB code</b>	<b>Structure Factors</b>
<b>Chitinase A</b>		
D313A – octaNAG	1EIB	r1eibsf
E315Q – octaNAG	1EHN	r1ehnsf
Y390F – hexaNAG	1FFR	r1ffrsf
Wild Type-allosamidin	1FFQ	r1ffqsf
D391A - hexaNAG	To be deposited	
D391A - tetraNAG	1K9T	
<b>Chitobiase</b>		
D539A - diNAG	1C7S	RCSB001441
E540D - diNAG	1C7T	RCSB001442

The coordinates and the structure factors of the complexes were deposited in the Protein Data Bank at the RCSB with the above accessions



**Figure M.1. Screening of chitinase A activity.** Indirect monitoring of the activity of chitinase A by probing the activity of *Escherichia coli* ChbF (see text).

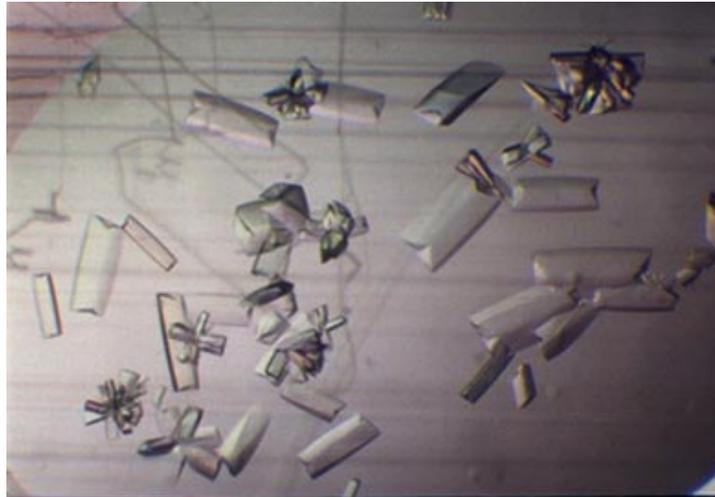
(A) LB-chitin plate with X-NAG: shows colonies are:

1. DY329 / pGPchiA;
2. DY329::*chbF*<sup>-</sup> / pGPHis-*chbF* + pGPchiA;
3. DY329::*chbF*<sup>-</sup> / pGPchiA ;
4. DY329 / pKK223-3.

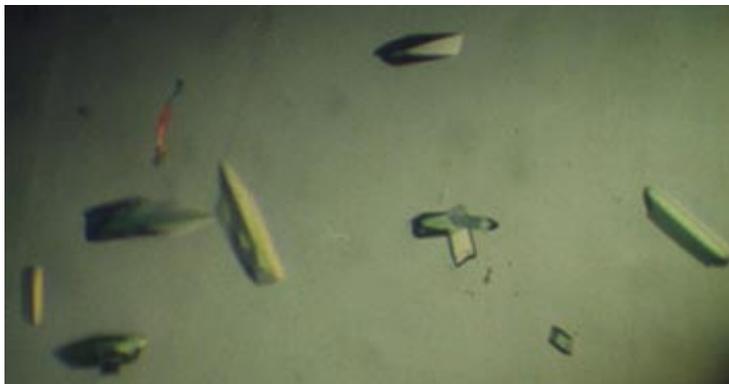
(B) LB plate with diNAG and X-NAG: shows colonies are:

1. DY329 / pKK223-3;
2. DY329::*chbF*<sup>-</sup> / pGPHis-*chbF*;
3. DY329::*chbF*<sup>-</sup> / pKK223-3.

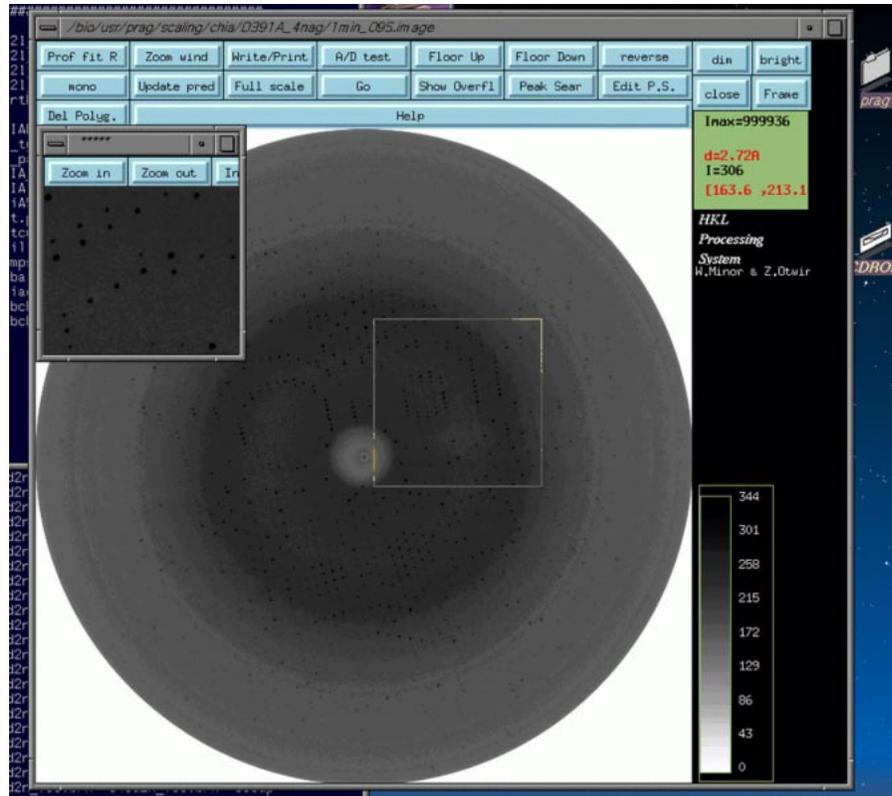
## Chitinase A



## Chitobiase



**Figure M.2. Co-crystals of chitinase A and chitobiase.** Co-crystals of mutant proteins with native substrates (oligoNAG) were grown as described in the Methods section.



**Figure M.3. Diffraction pattern of chitinase A D391A-tetraNAG crystal.** Data was collected under cryo-cooling conditions at the EMBL, DESY in Hamburg as described in the Methods section. The image is one of 200 images that were collected. Inset shows zoom into the low resolution area (about 30-3 Å). This picture has been performed by using the program *HKL-Denzo*.

## Results

### Chapter I. Genetic analysis of the conserved catalytic TIM-barrel domain of chitobiase

*S. marcescens* chitobiase is the first member of the glycosyl-hydrolase family 20 for which the structure was solved (Tews *et al.*, 1996a). It serves as a representative structural model for other family members (Fernandes *et al.*, 1997; Tews *et al.*, 1996a). Alignment of *S. marcescens* chitobiase with other family 20 protein members revealed sequence identities ranging between 15-55 % (not shown). Multiple alignment analysis of the catalytic TIM-barrel domain of family 20 hydrolases and distance calculations revealed a phylogenetic tree with two branches (see Figures 1.1 and A.1 in the appendix). One branch includes the bacterial enzymes, which are clustered with about 50 % sequence identity. The second branch includes the mammalian enzymes, which are clustered with about 80 % amino acids identity. Interestingly, the  $\alpha$  chain of human hexasaminidase A (HexA), a heterodimer made of  $\alpha$  and  $\beta$  homologous chains) shows a higher similarity to its ortholog, mouse  $\alpha$  chain, than to the paralog human  $\beta$  chain. A higher similarity between the human and mouse  $\beta$  chains is found as well. This suggests that gene duplication in an ancestral organism (paralog) occurred before human and mouse were evolutionary diverged.

Structural analysis shows that many of the conserved residues are located at the proposed active site. This finding indicates that all family 20 members utilize the same catalytic mechanism. The mammalian enzymes are acting as a hetero or homodimers. However the crystal structure of chitobiase and the recently solved structure of  $\beta$ -N-

acetyl-hexosaminidase from *Streptomyces plicatus* suggest that these bacterial enzymes are monomers (Mark *et al.*, 2001; Mark *et al.*, 1998). It is assumed that dimerization generates a specific surface for substrate binding; however, no structural prediction concerning dimer formation can be derived from the analysis of the bacterial enzymes.

### **Bacterial chitobiase and human hexosaminidase A**

Modeling of human hexosaminidase is important because many point mutations in these enzymes lead to GM<sub>2</sub>-gangliosidosis (including Tay-Sachs and Sandhoff diseases (reviewed in (Gravel *et al.*, 1995); See Table 1.1). Thus this investigation helped to rationalize Tay-Sachs and Sandhoff diseases.

Improved alignment [in which the known secondary structure elements of the chitobiase TIM-barrel are assigned to the hexosaminidase A sequences, (using clustalW)] revealed about 25 % amino acid identity between the human and the bacterial proteins. We used the improved alignment to build a model of the  $\alpha$  chain human hexosaminidase A (see Methods), a model structure very similar to that proposed by Tews *et al.*, (1996). Based on the structural model many Tay-Sachs mutations can be assigned to the enzyme's active site. Moreover, in some cases mutations located far from the active site shows somewhat less severe phenotypes. However, a cluster of mutations located at the outer side of the TIM-barrel far from the active site shows severe phenotypes as well. It was suggested that these mutants are located at the heterodimer interface and are involved in the dimer formation.

## **Introduction of point mutations into *chb***

We further tested the hypothesis of structural homology by the introduction of specific Tay-Sachs mutations into the *chb* gene. Mutational sites were chosen based on the sequence conservation and structural analysis of the wild type chitobiase (Figures 1.2 and A.3 in the Appendix). Table 1.1 summarizes the known mutations in glycosyl-hydrolase family 20 members and includes the mutational sites analyzed in this work. More than 20 mutations involved in Tay-Sachs and Sandhoff diseases in the human hexosaminidase A are known (see below).

Site directed mutagenesis was performed as described in the Methods section on a subcloned *chb* gene. We set up a new highly efficient screening system to isolate chitobiase mutants by using the X-NAG substrate (see Methods). Mutant clones unable to develop a blue color on X-NAG plates were confirmed by DNA sequencing and further investigated for chitobiase activity. To quantify the activity of the mutant clones we assayed crude extracts of induced cultures (Table 1.2). We introduced the mutations R349A, R349L and D448H (R178 and D258 in HexA, Table 1.1). We found that chitobiase mutant proteins R349A, R349L and D448H produce defective enzymes (Table 1.2). Understanding of the function of these residues can be obtained by inspection of the catalytic site. Zooming into the active site (Figure 1.3) shows that R349 forms two critical hydrogen bonds with the OH3 and OH4 hydroxyl groups of the -1 NAG, thus anchoring the substrate into the base of the catalytic pocket. D448 forms a hydrogen bond with the key residue D539 and a second hydrogen bond via a water molecule to OH3 hydroxyl of the -1 NAG. Thus both R349 and D448 appear to

participate in substrate binding, suggesting that a similar defect is found in the Tay-Sachs R178 and D258 mutations.

We also introduced the mutation Q363L (Tay-Sachs V192L) and found that the enzyme retained about 40 % of the enzymatic activity (Table 1.2). A control mutant was constructed as well. The Tay-Sachs mutational change V192L affects a non-conserved residue in chitobiase, and was therefore predicted to have no direct effect on the active site (see Figure 2-1). Inspection of the HexA model showed that V192 is located on the surface of  $\alpha$ -helix # 1 of the TIM-barrel. This residue is part of a cluster of Tay-Sachs mutations located on the surface of the TIM-barrel, which includes W420C, W474C, E482K and W485R. It probably forms an interface binding area for interacting with the subunits (Tews *et al.*, 1996a), or with the substrate-specific protein cofactor, the GM<sub>2</sub>-activator protein (Burg *et al.*, 1985; Wright *et al.*, 2000).

It was previously suggested that the conserved E540 of chitobiase acts as the proton donor (HexA E323). We have generated the mutants E540D and E540A that were found to have very low activity (Table 1.2). Similar findings were obtained with the analogous mutations E323D in HexA and E314Q mutants from *S. plicatus* (Mark *et al.*, 2001; Mark *et al.*, 1998). Two additional conserved sites, D539 and Y669, were also studied. We found that mutations in both residues decrease the enzymatic activity (Table 1.2). Inspection of chitobiase structure suggested that both residues are in contact with the acetamido group of the -1 NAG substrate. Biochemical and structural studies of mutants in the proposed catalytic residues D539 and E540 are described in Chapter II and in Prag *et al.*, (2000). In conclusion, the analysis of the mutants in

chitobiase provides further support for the structural and functional homology between the human and the bacterial enzymes.

**Table 1.1 List of the available mutants in the glycosyl hydrolase family 20**

Human TSD HexA		<sup>a</sup> Human HexA	<sup>b</sup> <i>S.plicatus</i> hex	<sup>c</sup> <i>S.marcescens</i> chitobiase
$\alpha$ -chain	$\beta$ -chain	$\alpha$ -chain		
R170Q/W				R341
	I207V			L345
<u>R178C/H/L</u>			R162H	<b>R349H/A</b>
V192L				<b>Q363L</b>
V200M				K371
<u>S210F</u>				G381
R247W				Q437
R249W				R439
G250D				Q440
<u>D258H</u>			D246N	<b>D448H</b>
G269S				S459
S279P				T487
F300L				F512
		E307D		E519
DF305				I517
				<b>D539A</b>
		E323D	E314Q	<b>E540A/D</b>
<u>W420C</u>				V668
				<b>Y669E</b>
	Y456S			E675
W458R				W763
		E462D		E739
W474C				F752
	R505Q			R754
E482K				E760
R499C/H				R776
R504C/H				K779

List of known mutations in the glycosyl hydrolase family 20. The first two columns show Tay-Sachs disease (TSD) mutants in the human hexosaminidase  $\alpha$  and  $\beta$  chains. Severe Tay-Sachs B1 phenotypes characterized by normal amounts but catalytically inactive enzymes are underlined.

<sup>a</sup> Site directed mutations introduced into hexA. (Fernandes *et al.*, 1997).

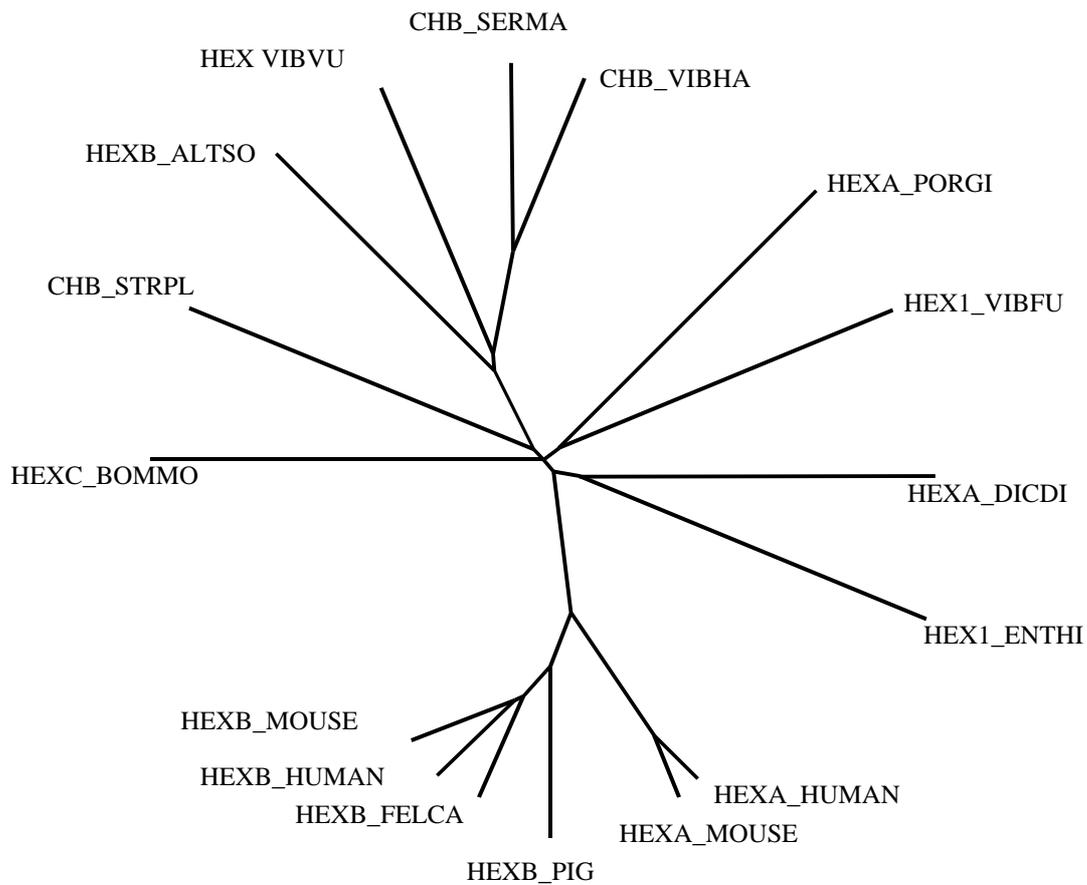
<sup>b</sup> Analysis of conserved residues in the *S. plicatus*  $\beta$ -*N*-acetylhexosaminidase (see (Mark *et al.*, 2001)).

<sup>c</sup> List of residues in the TIM-barrel domain of *S. marcescens* chitobiase corresponding to those in HexA. Mutations introduced in the *S. marcescens* chitobiase gene are shown in bold letters (see text and multiple alignment in Appendix A1).

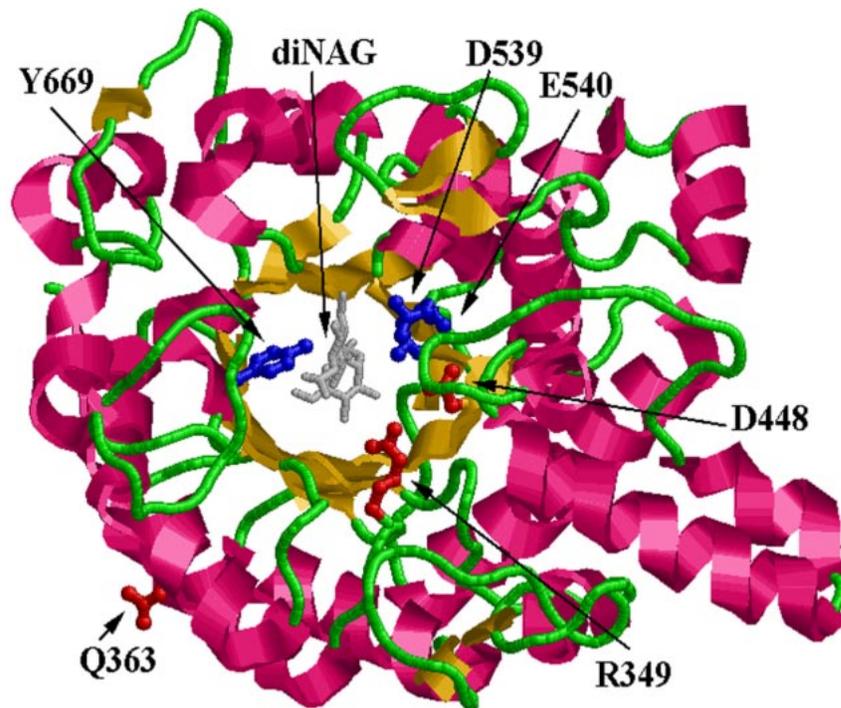
**Table 1.2**      **Relative activities of wild-type and mutant chitobiase proteins**

<b>Protein</b>	<b>Relative activity</b>
Wild type	1
R349L	0.035
R349H	0.028
Q363L	0.385
D448H	0.012
D539A	0.008
E540A	0.032
E540D	0.005
D539A / E540D	> 0.001
Y669E	0.075
D539A / Y669E	0.008

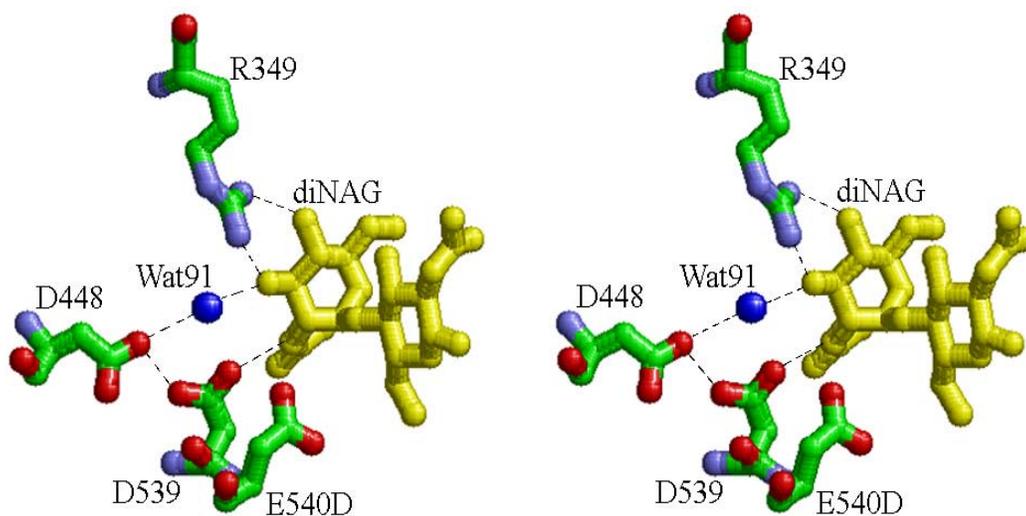
Crude extracts of overnight induced cultures were assayed for activity using p-Nitro-Phenyl-NAG as a substrate (see the Methods section). The results above are averages of at least three independent assays with standard deviations of less than 5 %.



**Figure 1.1. Phylogenetic tree of the glycosyl-hydrolase family 20 TIM Barrel domain.** Multiple alignment of glycosyl-hydrolase family 20 was performed using the clustalW program as described in the Methods section. Based on the structure of chitobiase, the region coding for the TIM-barrel domain of all other family members was realigned. Neighbor-Joining distances were calculated to generate the above unrooted tree. Codes: SERMA, *Serratia marcescens*; VIBHA, *Vibrio harveyi*; PORGI, *Porphyromonas gingivalis*; VIBFU, *Vibrio furnissii*; DICDI, *Dictyostelium discoideum*; ENTHI, *Entamoeba histolytica*; HUMAN, *Homo sapiens*; MOUSE, *Mus musculus*; PIG, *Sus scrofa*; FELCA, *Felis silvestris catus*; BOMMO, *Bombyx mori*; STRPL, *Streptomyces plicatus*; ALTSO, *Alteromonas* sp. (strain O-7); VIBVU, *Vibrio vulnificus*.



**Figure 1.2. Projection of Tay-Sachs mutations on the catalytic domain of chitobiase.** The positions of Tay-Sachs mutations (red) and other mutations (blue) that were generated in this study are shown. Also shown is the location of the diNAG substrate (gray) at the active site. The structure shown is of the TIM- barrel domain of chitobiase E540D - diNAG mutant complex (see Chapter II)



**Figure 1.3. A stereo view showing the active site of chitobiase - diNAG complex.** The figure is focused on the enzyme-substrate interactions with four active site residues that are described in the text. The coordinates of chitobiase E540D - diNAG mutant complex (see Chapter II) were used to generate this figure. Dashed lines represent putative hydrogen bonds. Color code: diNAG, yellow; atoms of amino acid residues: carbon, green; nitrogen, light blue and oxygen, red. The oxygen atom of a water molecule is rendered as a blue ball.

## **Chapter II. Structures of chitobiase mutants complexed with the substrate di-*N*-acetyl-*D*-glucosamine: The catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540**

*S. marcescens* chitobiase belongs to glycosyl hydrolases family 20 (Henrissat, 1991). In this family the catalytic domain is an  $\alpha/\beta$  TIM-barrel, and the active site lies at the center of the barrel convex side made of the loops at the c-terminus end of the barrel  $\beta$ -strands. It was suggested that hydrolysis by chitobiase proceeds via a substrate-assisted mechanism in which the configuration of the anomeric carbon is retained. In this reaction glutamate 540, located in loop #4 of the catalytic domain, acts as a proton donor and the acetamido group acts as the nucleophile (Drouillard *et al.*, 1997; Tews *et al.*, 1996b).

### **Conservation of catalytic residues**

Multiple alignment of family 20 of glycosyl hydrolases, based on 17 sequences extracted from the SwissProt bank, showed the complete conservation of both Asp539 and Glu540 in loop #4 (positions of the amino acid residues are those of *S. marcescens* chitobiase). A signature of six residues HXGGDE (where X represents any amino acid) can be easily recognized (see Figure I.4 in the Introduction). It was previously shown that changing the catalytic glutamate of chitobiase in *Streptomyces plicatus* and in human  $\beta$ -hexosaminidase A inactivated the enzyme (Fernandes *et al.*, 1997; Mark *et al.*, 1998). The two conserved glycine residues are probably essential in forming the structure of the end of  $\beta$ -strands #4 and loop #4, and in facilitating its conformational flexibility. The conserved His535, which is located at the core of the barrel, makes a number of electrostatic and aromatic interactions with residues at the bottom of the

catalytic pocket. These are probably essential for maintaining the structure of the catalytic site. In this study we chose to change Asp539 and Glu540 of *S. marcescens* chitobiase, the only member for which the structure of the enzyme was solved at the time the experiments were carried out. The availability of inactive mutant proteins allowed us to achieve co-crystallization with the native substrate. The structures of the enzyme-substrate co-crystals of the mutant proteins allowed us to gain insight into the function of these conserved residues in the active site.

### **Biochemical characterization of *chb* mutants**

The kinetic parameters of the purified wild-type and mutant proteins D539A, E540A and E540D were determined (Table 2.1). The residual activity of the E539A/E540A protein was too low to allow kinetic analysis. Mutation D539A was found to increase  $K_M$  and to decrease  $K_{cat}$ , suggesting that D539 is involved in both substrate binding and catalytic activity. This mutational change caused a decrease of over 1000 fold in the enzymatic efficiency ( $K_{cat}/K_M$ ). Mutations E540A and E540D were found to decrease  $K_M$  by about 5 fold and to decrease  $K_{cat}$ , suggesting that E540 is involved in catalysis. These mutations led to a decrease in enzymatic efficiency of about 30 and 120 fold respectively (Table 2.1). It is possible that the somewhat higher activity of the E540A mutant over that of E540D is due to the participation of a water molecule accommodated between the scissile bond and the C $\beta$  group of the mutated Ala540 residue. Thus we demonstrated that the conserved Asp539 - Glu540 (DE) pair plays an important role in the catalytic site.

## Co-crystal structures of chitobiase mutants with diNAG

To investigate the structural consequences of the amino acid substitutions, the structures of D539A and E540D mutants complexed with diNAG were determined. This work was done as a collaboration with Dr. K. Petratos. We co-crystallized the mutant proteins with diNAG and solved their structures (Table 2.2). Diffraction data were collected to 1.9 Å and 1.8 Å resolutions. The root-mean-square-deviations (RMSD) of the backbone coordinates between the wild-type enzyme and the D539A and E540D mutants, were calculated to be 0.31 Å and 0.37 Å respectively. These findings show that the structure of the mutant proteins did not change. The backbone structure of the TIM-barrel domains of mutant proteins and the position of the substrates are shown in Figure 2.1. The conformation of the diNAG bound to the enzyme is similar to that described previously (Tews *et al.*, 1996b). However, the two mutants revealed subtle changes in both amino acid residues and in the substrate.

Surprisingly, the analysis of the D539A-diNAG complex reveals that although the position of the substrate is maintained, the acetamido group is flipped by 175° degrees (dihedral angle of the N2-C7 bond) with respect to the wild-type and E540D complexes (Figures 2.2 and 2.3). It appears that the flipped acetamido group occupies part of the space that is filled by the side chain of D539 in the wild-type. Thus, it seems that one function of D539 is to restrain the acetamido group in a specific orientation by forming a hydrogen bond with N2 of the -1 sugar. The loss of this hydrogen bond is probably responsible for the reduced affinity of the mutant protein to the substrate. The altered conformation of the acetamido group in the D539A-diNAG complex may be favored by its hydrogen bond to Glu540 (Figure 2.3). In addition,

residues Trp616 and Trp639 form hydrophobic interactions with the acetamido group of the -1 sugar in its two alternative conformations. We speculate that these interactions limit the position of the acetamido group to two distinct conformations.

The mutant protein D539A caused additional changes at the active site. The electron density of the side chain of residue E540 is poor ( $\rho = 0.65 \times \sigma$ ). This is reflected by increased atomic temperature factors ( $B_{factor}$ ). The average  $B_{factor}$  for the E540 side chain was refined to  $41.1 \text{ \AA}^2$ , as compared to the total averaged  $B_{factor}$  of  $23.6 \text{ \AA}^2$  for all side chains in the molecule ( $\sigma = 8.4 \text{ \AA}^2$ , Table 2.2). Similar analysis of the wild-type complex (PDB code: 1qbb) yielded a  $B_{factor}$  for the E540 side chain of  $26.4 \text{ \AA}^2$  and total averaged  $B_{factor}$  of  $16.4 \text{ \AA}^2$  ( $\sigma = 2.8 \text{ \AA}^2$ ). These results suggest a greater flexibility of E540 in the mutant. Furthermore, E540 assumes an altered conformation in which the distance between the O $\epsilon$ 1-carboxylic group of E540 and the glycosidic oxygen increases from  $2.8 \text{ \AA}$  to  $5.3 \text{ \AA}$  (Figures 2.2 and 2.3). These results suggest that D539 acts to restrain the movement of E540 and could account for the decreased activity of the D539A protein.

The conserved D539 residue may have additional functions. It possibly aids in the contact between the acetamido O7 of the -1 sugar and the anomeric carbon C1 of the substrate, and in stabilizing the partial positive charge of the acetamido group while forming the proposed oxazolinium ring (Brameld *et al.*, 1998; Terwisscha-van-Scheltinga *et al.*, 1995; Terwisscha-van-Scheltinga *et al.*, 1994). D539 may also improve proton donation by forming an electrostatic interaction with E540. Thus it is presently difficult to distinguish the various contributions of the D539 residue to the

enzymatic activity. Attempts to get the D539N mutant, which could have helped in the above analysis, failed.

The analysis of the E540D-diNAG co-crystal is simpler since the structural difference from the wild-type enzyme is confined to the absence of one CH<sub>2</sub> group. In the E540D-diNAG complex the distance between the carboxylic end and the glycosidic oxygen was found to be 4.3 Å (Figure 2.3). Thus, increasing the distance between the carboxylic end and the glycosidic oxygen from 2.8 Å (as found in the wild-type enzyme) to 4.3 Å is sufficient to reduce the catalytic activity of the mutant (Table 2.1). These findings support the hypothesis that E540 acts as the proton donor. It is not clear why the mutations changing the E540 residue appear to increase the affinity of the enzyme to the substrate.

### **The catalytic mechanism of chitobiase**

Our detailed structural analysis of two mutants at the catalytic site allowed a more comprehensive analysis of the chitobiase mechanism of action. It is clear from our results that in the complex the structure of the substrate is distorted. As was previously observed (Tews *et al.*, 1996), the planes of the two sugars are tilted around the glycosidic linkage by about 90<sup>0</sup> with respect to one another. This distortion is stabilized by hydrophobic and polar interactions. A similar distortion was found in chitinase A-octaNAG complexes (see Chapter IV and Discussion). It was previously shown that the cleavage of diNAG proceeds via a substrate-assisted mechanism, in which the configuration of the anomeric carbon is retained (Drouillard *et al.*, 1997). Our results support the proposal that Glu540 donates a proton to the glycosidic bond. This results in the cleavage of diNAG and in the release of the +1 sugar from the

active site. Our results show the importance of the conserved Asp539 residue. As discussed above, this residue performs multiple functions.

After breakage of the scissile bond, the positive charge of the C1 of the -1 sugar is stabilized by a nucleophilic attack from the O7 of the acetamido group, resulting in the formation of an oxazolinium ring. We suggest that the oxazolinium ring is stabilized by a hydrogen bond of Asp539 with the N2. Interestingly, we found that one function of Asp539 is to keep both the Glu540 and the *N*-acetyl group in the proper position essential for catalysis. Finally, hydrolysis of a water molecule completes the reaction by the hydroxyl attack at the C1 of the -1 sugar and the reprotonation of Glu540.

The results presented in this chapter provide an explanation for the importance and conservation of the Asp -Glu in glycosyl hydrolase family 20. It is interesting that chitinases belonging to family 18, of which there are now over 50 sequenced representatives, all possessing an Asp-X-Glu motive at the catalytic site (see Chapters III - V). In most cases, X is found to be a hydrophobic residue. Inspection of the determined 3D structures of the chitinases suggests that the aspartate and glutamate residues assume a similar configuration in both chitobiase and chitinases. The function of the Asp-X-Glu motive in family 18 will be described in the next chapters.

**Table 2.1** Kinetic constants of wild-type and mutant chitobiase proteins

<b>Enzyme</b>	<b><math>K_M</math></b> <b>(<math>\mu\text{M}</math>)</b>	<b><math>K_{cat}</math></b> <b>(<math>\text{sec}^{-1}</math>)</b>	<b><math>K_{cat}/K_M</math></b> <b>(<math>\text{sec}^{-1} \times \mu\text{M}^{-1}</math>)</b>
WT	0.063	827.0	$1.3 \times 10^4$
D539A	1.991	17.0	$0.8 \times 10^1$
E540A	0.014	6.0	$4.2 \times 10^2$
E540D	0.010	1.4	$1.1 \times 10^2$

Kinetic assays were performed with pNp-NAG as described in the Materials and Methods section. Standard errors were less than  $\pm 15\%$  for  $K_M$  and less than  $\pm 18\%$  for  $K_{cat}$  determinations.

**Table 2.2** Diffraction data processing and refinement statistics

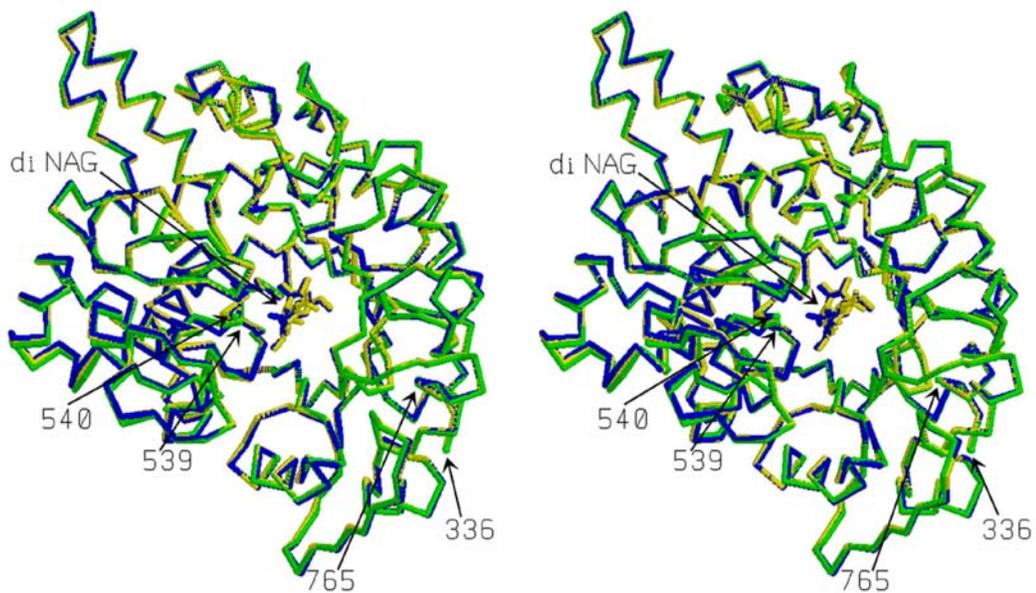
	D539A	E540D
<b>A. Diffraction Data</b>		
Unit cell		
a (Å)	109.8	109.2
b (Å)	100.0	99.4
c (Å)	86.3	86.6
Space group symmetry	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2
Diffraction limit (Å)	1.75	1.85
Observations	937195	1001880
Unique reflections	76047	71983
<sup>4</sup> Completeness (%)	86.9 (90.8)	96.9 (94.0)
<sup>1,4</sup> R <sub>merge</sub> (%)	2.4 (9.5)	5.5 (21.6)
Overall I/σ	30.8	27.8
<b>B. Refinement</b>		
Resolution limits	15-1.8	10-1.9
<sup>2</sup> R <sub>factor</sub> (%)	17.1	19.1
<sup>3</sup> R <sub>free</sub> (%)	22.3	24.6
Number of residues	857	857
Number of water molecules	823	834
<b>C. RMS deviations from ideality</b>		
Bond length (Å)	0.022	0.029
Bond angle (°)	2.29	2.89
Aromatic planar groups (Å)	0.012	0.015
<b>D. Average of B<sub>factor</sub> Values (&lt;Å<sup>2</sup>&gt;)</b>		
All atoms	21.8	25.7
Main chain	19.6	24.7
Side chain	23.6	26.8
Substrate (diNAG)	15.5	24.9
Solvent (H <sub>2</sub> O, SO <sub>4</sub> <sup>2-</sup> )	31.0	34.9

<sup>1</sup>R<sub>merge</sub> =  $\sum |I_i - \langle I \rangle| / \sum I_i$  where  $I_i$  is an individual observed intensity measurement and  $\langle I \rangle$  is the average intensity for this reflection.

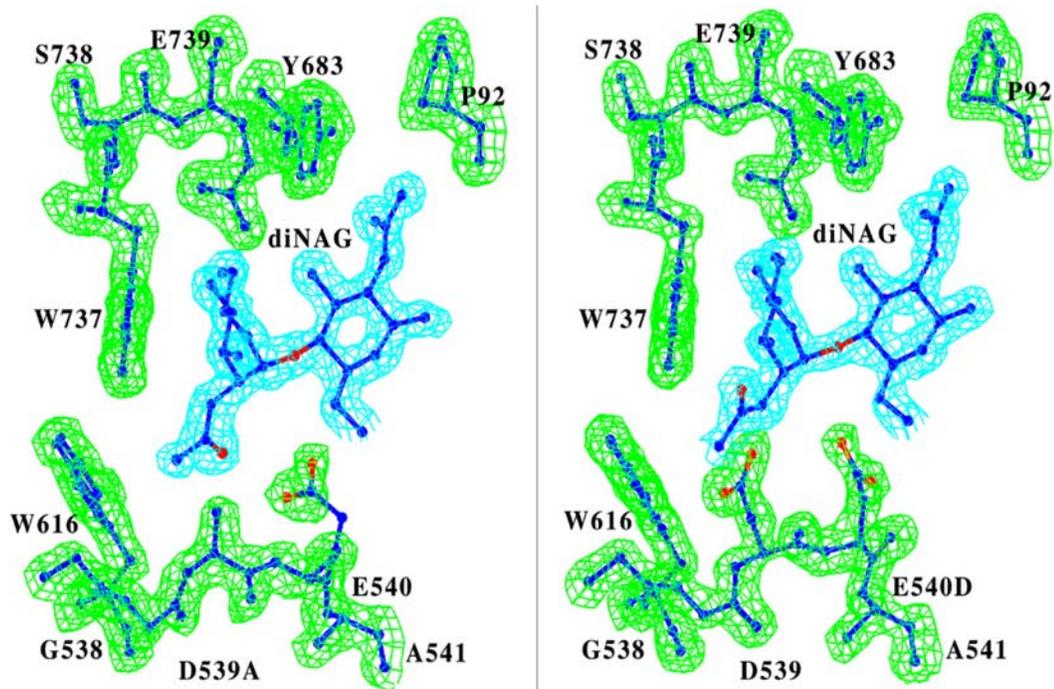
<sup>2</sup>R<sub>factor</sub> =  $\sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$  where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively.

<sup>3</sup>R<sub>free</sub> =  $\sum |F_{\text{obs}} - |F_{\text{calc}}|| / \sum F_{\text{obs}}$  calculated from 5% of the reflections selected randomly and omitted from the refinement process.

<sup>4</sup> Values in parentheses refer to the corresponding values of the highest resolution shell (from scalepack output. 1.95-1.8 Angstrom).

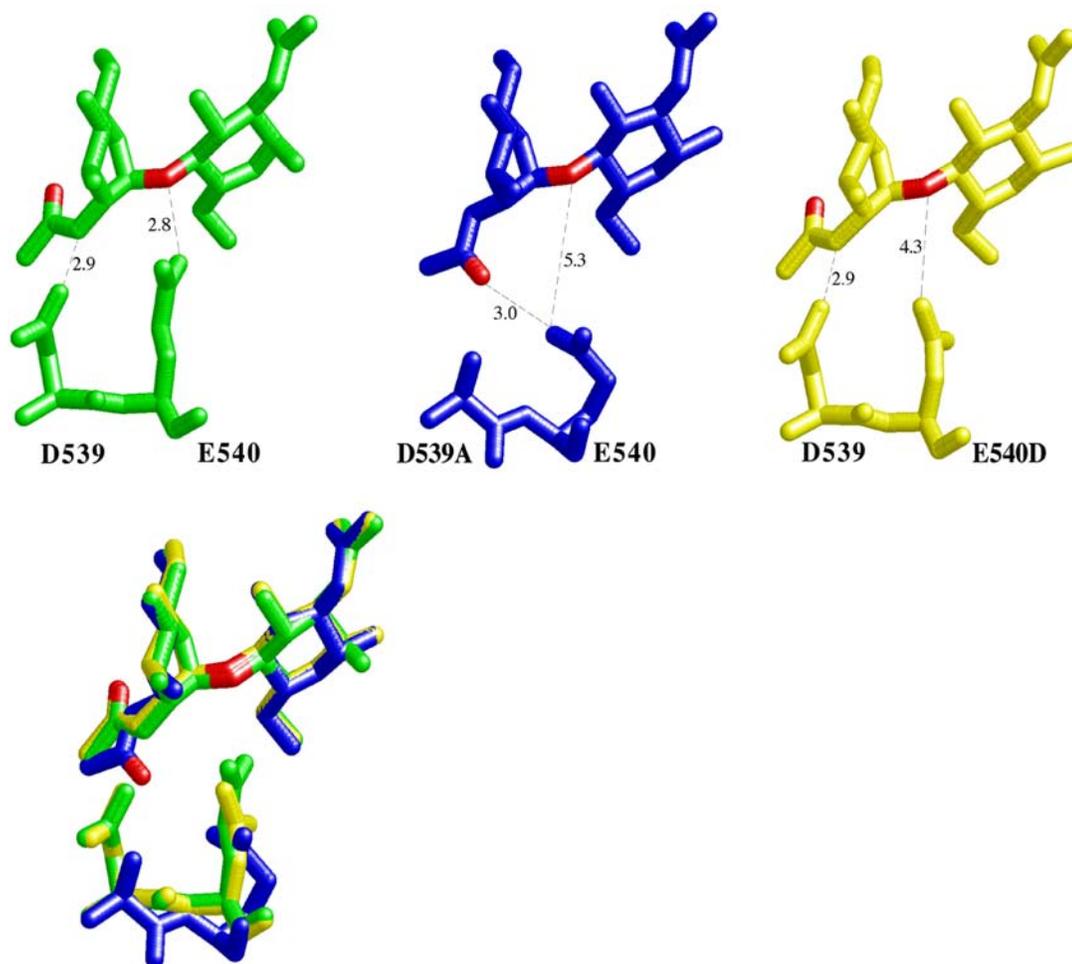


**Figure 2.1. Comparison of the TIM-barrel catalytic domains of wild-type and mutant chitobiase structures.** Stereo views of wild-type (green) mutant D539A (blue) and mutant E540D (yellow) of the backbone coordinates and of the diNAG atoms superimposition are shown. Location of residues at the amino (Pro336) and the carboxy (Arg765) ends of the TIM-barrel domain, and of the catalytic Asp539 and Glu540 residues are indicated. The wild-type data were taken from Tews *et al*, 1996 (1qba). The program *insight II* was used to generate this figure.



**Figure 2.2. Models of the complexes of *S. marcescens* chitobiase mutants.**

The figure focuses on the active sites occupied by the substrate. Part of the calculated electron density maps  $2mF_{\text{obs}} - DF_{\text{calc}}$  (see Methods) are shown from the final refined models of the mutants D539A at 1.8 Å (left) and E540D at 1.9 Å (right), contoured at  $1.5 \times \sigma$  level. The sugar and amino acid residues are shown in blue. The glycosidic oxygen O7A of the acetamido group and the carboxylic oxygen atoms of residues 539 and 540 are colored red. The program O was used to generate this figure.



**Figure 2.3. Structural comparisons of the catalytic sites of wild-type and mutant complexes.** Wild-type (green), D539A (blue) and E540D (yellow) are shown. Superimposition of the three complexes is shown below. The glycosidic oxygen and the proposed acetamido O7 nucleophile are colored in red. Distances are indicated in Angstrom units. The wild-type data were taken from Tews *et al*, 1996 (1qbb). The program RasMol was used to generate this figure.

### **Chapter III. Structural analysis of oligoNAG substrates bound to chitinase A**

The crystal structure of native chitinase A was refined at a 2.3 Å resolution (Perrakis *et al.* 1994). Based on this structure it was proposed that the substrate-binding site forms along the groove that is located at the carboxy-termini of the  $\beta$ -strands of the TIM-barrel domain. However, attempts to obtain the co-crystal structures of chitinase-oligoNAG complexes in which the non-cleaved substrate occupies the catalytic site proved unsuccessful (Perrakis *et al.*, 1994; Terwisscha-van-Scheltinga *et al.*, 1996). Only two NAG residues were found at the active site of the enzyme, followed by a soaking experiment with tetraNAG. This is possibly due to the enzymatic cleavage of the substrate within the crystal lattice (Tews *et al.*, 1997). To understand how chitinase A binds to the chitin oligomer, we generated site-specific mutants at the catalytic site and obtained chitinase A- oligoNAG co-crystals that enabled a comprehensive structure-function analysis.

#### **Identification of mutants**

Mutants were generated by site directed mutagenesis. Mutants that showed a decreased turnover number ( $K_{cat}$ ) but were able to bind the substrate ( $K_M$ ), were chosen for further crystallographic studies (Table 3.1). The biochemical properties of these mutants are described in Chapter IV. The mutations D313A, E315Q, Y390F and D391A were found to be suitable for further study. Co-crystals of these mutant proteins with oligoNAG were grown to the face-centered orthorhombic space group  $C222_1$  and data were collected under cryogenic conditions as described in the Methods. The final refined complex resolutions were less than 2 Å, sufficient for a

detailed analysis of the enzyme-substrate interactions (for a table describing the crystallographic data and refinement statistics see the Material and Methods section). The structures of the wild type and mutant proteins are highly similar. However, specific, subtle, structural changes in residues 315 and 313 were observed. This finding allowed us to investigate how the substrate binds to chitinase A.

### **Analysis of chitinase A binding groove in a co-crystal structure with oligoNAG**

The complex of chitinase A E315Q with octaNAG was analyzed in detail. It allowed us to identify eight subsites located at the long groove across the barrel. The groove appeared to be open at the side facing the FnIII-like domain. This opening probably accounts for the ability of the enzyme to bind longer oligoNAG chains and chitin. At the other end of the groove the presence of Tyr418, which contacts the +2 NAG, may reduce the ability of a longer substrate extended beyond this tyrosine. Thus the +3 NAG will have to bypass the block posed by residue Tyr418 (see Chapter V). The function of Tyr418 may explain the propensity of diNAG to be a major product produced by the enzyme.

The overall structure of the complex E315Q-octaNAG is shown in Figure 3.1. The substrate can clearly be seen firmly embedded within the deep tunnel (about 8-11 Å in width and about 15 Å deep). Close inspection of the structures indicated that a large number of amino acid residues contribute to the binding of octaNAG (Table 3.1). Direct contacts are mainly mediated by aromatic and charged residues. In addition, indirect contacts via water molecules are also present (see Figures 3.1 A and B respectively and Table 3.1). Weaker interactions extend from -2 to -6 subsites. Finally, a small number of the interactions of the -5 and -6 sugar residues are probably

due to the packing in the crystal lattice and may not reflect any biological function. Many, but not all, of the residues that participate in substrate binding are highly conserved (Table 3.1). All acetamido groups of the -4 NAG to +2 NAG residues were found to form specific interactions with the enzyme. We identified at least one hydrogen bond between the enzyme and each of the acetamido groups of the -4 NAG to +2 NAG residues, providing for oligoNAG specificity. We suspect that the less conserved residues that we identified as participating in substrate binding may represent different evolutionary solutions for substrate binding. Alternatively, these non-conserved residues reflect minor differences of substrate specificity among glycosyl hydrolase family 18.

### **The catalytic pocket and substrate distortion**

The -1 and +1 sugars and the scissile bond were identified as previously suggested (Perrakis *et al.*, 1994; Tews *et al.*, 1997). These two NAG residues are tilted around the glycosidic linkage by about  $90^{\circ}$  with respect to one another in a non-favored conformation that may reduce the activation energy (Figure 3.2). A number of interactions of the -1 NAG take place with residues Tyr163, Phe191, Trp275, Asp313, Glu315, Met388, Tyr390, Asp391, Tyr444, Arg446, and Trp539 (see also Table 3.1). In addition, the +1 NAG interacts with residues Trp275, Phe316, Met388, Asp391, Phe396 and Arg446. We assume that Glu315, mutated in the E315Q protein, also interacts with the +1 NAG. These interactions are probably responsible for forcing the substrate into the non-favored conformation for catalysis (see Chapter IV). Moreover, we found that the sugar residue bound at subsite -1 (-1 NAG) adopts a 'boat'  ${}^{1,4}B$  conformation. Similarly, a 4-sofa conformation was reported for the structure of

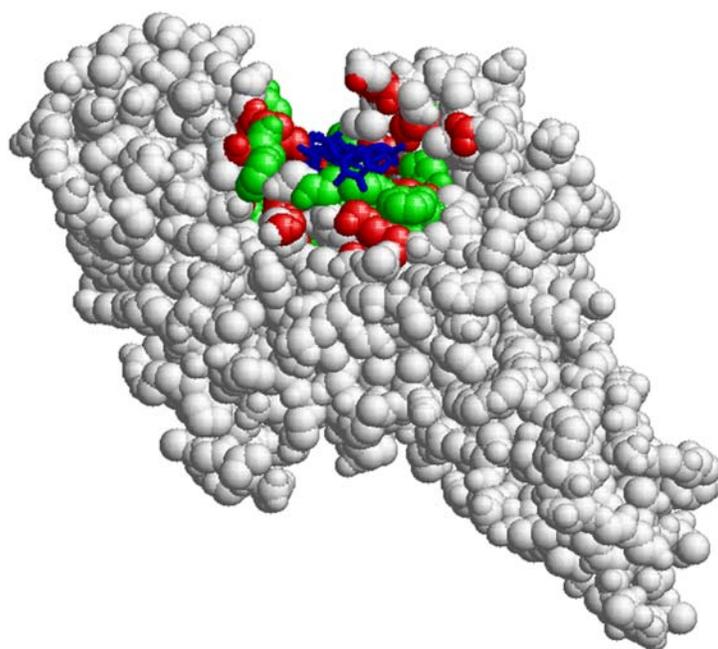
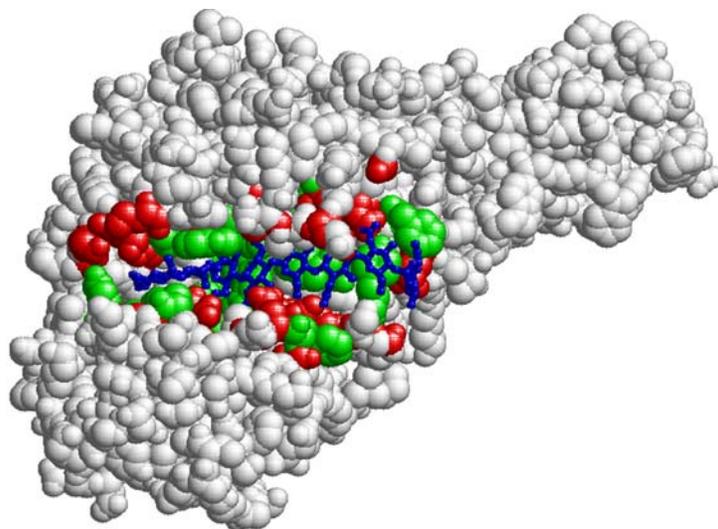
chitobiase-diNAG complex, see Chapter II and in Prag *et al.*, (2000), Tews *et al.*, (1996a), and in modeled structure of the chitinase A complex with hexaNAG (Tews *et al.*, 1997). Such a conformation allows the C1, C2, C5 and O5 of the -1 pyranose ring to adopt the coplanar configuration required for catalysis (see chapter IV). The acetamido group of the -1 NAG, proposed to be participating in the substrate-assisted catalysis, is embedded in a pocket that extends from the substrate-binding groove. In the E315Q complex the conformation of the acetamido group of the -1 NAG is energetically favored (as are all other acetamido groups of the octaNAG substrate). We propose that this structure describes an early binding step in which a pre-Michaelis complex is formed. Interestingly, in complexes of D391A mutant protein, this acetamido group was found in a position rotated by about 180° with respect to that found with E315Q (see Chapter V for the significance of this finding). The configuration that is found in the E315Q and D313A complexes may favor entry of the -1 sugar to the catalytic site whereas the configuration that is found in Y390F and D391A complexes facilitates the interaction of Asp313 with N2 of the *N*-acetyl group. This causes the *N*-acetyl group to occupy the energetically non-favored conformation that is required for the substrate-assisted model of catalysis (see Chapters IV and V). This well-defined pocket is located at the center of the barrel and is surrounded by the conserved residues Tyr163, Phe191, Asp313, Met388, Tyr390 and Trp539 (Figure 3.3 and Table 3.1). All these residues are conserved in family 18 chitinases. Thus we suggest that this pocket is structurally conserved and plays an important role in the catalytic mechanism of family 18 chitinases.

The findings described in this section provide an explanation as to how the enzyme binds long oligoNAG substrates, and suggest an explanation for why the main products of the enzyme are units of diNAG. Our results lead us to suggest that chitinase A acts mainly as an exoenzyme cleaving two residues at a time from the reducing end.

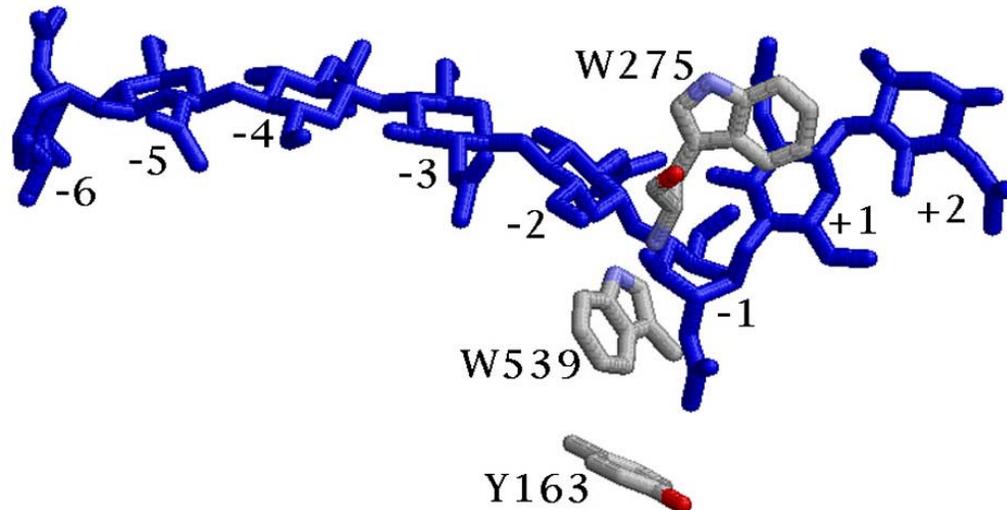
**Table 3.1** Contacts between chitinase A E315Q and octaNAG

Subsite	Residue	Distance (Å)	Surface contact (Å <sup>2</sup> )	Hydrogen Bond	Hydrophobic interaction	H <sub>2</sub> O	
+2 NAG	<u>Trp275</u>	3.1	27.0	-	+		
	Phe316	4.2	6.1	-	+		
	Lys320*	3.9	10.4	+	-	804	
	Ser364*	4.2	3.4	+	-	802	
	Gly366	3.9	13.7	-	+		
	Lys369	2.6	46.1	+	-		
	<u>Tyr390</u>	4.9	1.3	+	-		
	<u>Asp391</u>	3.2	23.6	+	-		
	Phe392	4.1	17.3	-	+		
	Phe396	3.4	47.9	-	+		
	Tyr418*	3.3	59.2	+	+	606	
	+1 NAG	<u>Trp275</u>	3.4	62.5	+	+	
		<u>Gln315</u>	4.9	1.5	+	-	
		Phe316	3.4	26.1	-	+	
Ser364		4.5	5.7	+	-		
Met388		3.9	21.4	+	-		
<u>Tyr390</u>		4.5	5.9	+	-		
<u>Asp391</u>		3.2	23.1	+	-		
Phe396		4.1	12.3	-	+		
Arg446*		3.4	26.2	+	-	797	
-1 NAG		<u>Tyr163</u>	3.4	32.1	-	+	
		<u>Phe191</u>	3.7	13.7	-	+	
		<u>Trp275</u>	3.0	15.3	+	-	
		<u>Asp313</u>	3.3	30.1	+	-	
		<u>Gln315</u>	3.0	21.8	+	-	
	Ala362	3.5	12.8	-	+		
	Met388	3.5	18.4	-	+		
	<u>Tyr390</u>	3.6	15.8	+	-		
	<u>Asp391</u>	3.1	11.0	+	-		
	Tyr444	3.6	13.0	-	+		
	Arg446	3.2	13.9	+	-		
	Trp539	3.4	36.3	+	+		
	-2 NAG	<u>Trp167</u>	4.0	8.5	+	-	
		Arg172	4.4	8.5	+	-	
<u>Phe191</u>		3.7	13.5	-	+		
Ile207		4.6	4.0	-	+		
<u>Trp275</u>		3.0	24.5	+	-		
Thr276		3.4	14.6	+	-		
Tyr444		4.0	20.8	-	+		
Arg446*		3.9	13.9	+	-	800	
Glu473		2.7	31.8	+	-		
Ile476		4.0	29.7	-	+		
Trp539		3.0	12.2	+	+		
Glu540		2.9	24.2	+	-		
-3 NAG		<u>Trp167</u>	3.6	52.2	-	+	
		Arg172	2.6	14.4	+	-	
	<u>Phe191</u>	4.5	2.7	-	+		
	Ser210*	5.3	1.0	+	-	329	
	His229*	4.3	15.9	+	+	789	
	Thr276	2.7	37.7	+	-		
	Leu277	4.1	21.3	-	+		
	Trp472	4.3	1.0	-	+		
	Glu473	2.6	34.2	+	-		
	-4 NAG	<u>Trp167</u>	3.7	24.7	-	+	
		<u>Tyr170</u>	4.0	13.2	-	+	
		<u>Gly171</u>	4.9	7.6	+	-	
		Arg172	2.6	35.2	+	-	
		Ile207	4.5	6.6	-	+	
Glu208*		5.1	1.8	+	-	790	
Ser210		4.3	20.6	+	-		
His229		5.1	5.7	+	-		
Asp230*		5.7	11.4	+	-	784	
Trp472		4.1	32.1	-	+		
-5 NAG		<u>Tyr170</u>	3.0	42.1	+	+	
		<u>Gly171</u> *	4.1	8.0	+	-	788
		Glu208*	3.9	25.8	+	+	790
		Asp230*	5.3	4.1	+	-	784
-6 NAG	<u>Tyr170</u>	3.4	31.5	-	+		
	Phe232	3.3	32.7	-	+		
	Lys237*	2.7	34.8	+	-	373	

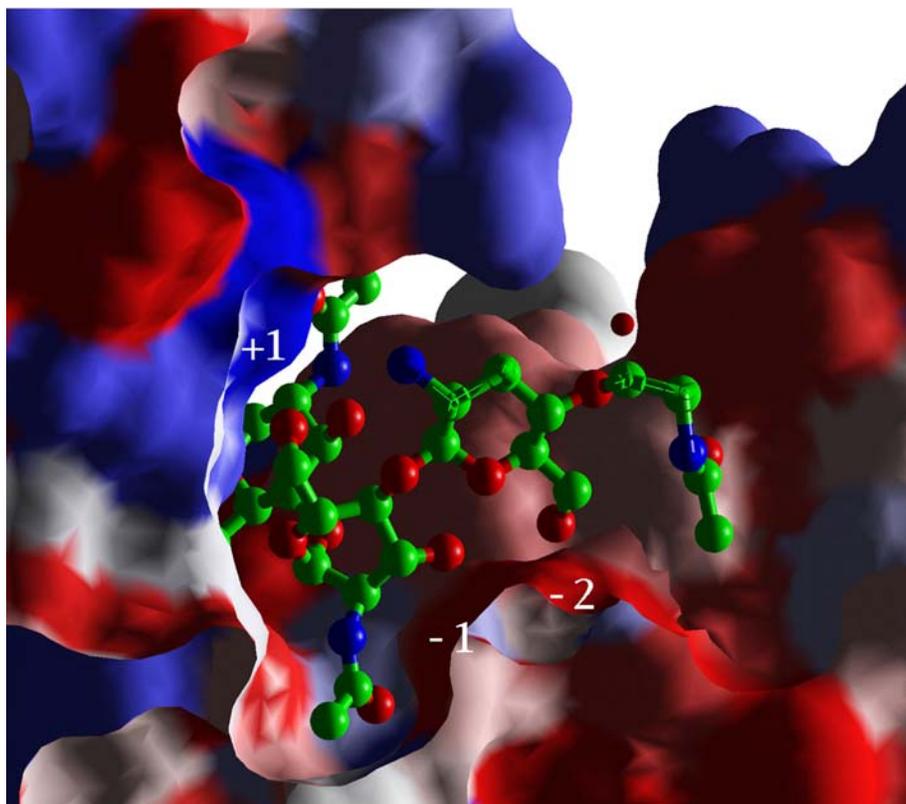
Enzyme-substrate interactions with each NAG group is given. The amino acid residues define the subsites on chitinase A. The minimal distance and surface contact between the amino acid residue and the NAG residue are given. (\*) Interactions mediated by water molecules are indicated. Underline marks highly conserved residues. This analysis was carried out by the LPC server (<http://bioinfo.weizmann.ac.il:8500/oca-bin/lpccsu/>; Sobolev *et al.*, 1999) and refined by inspection.



**Figure 3.1. Chitinase A – octaNAG complex.** OctaNAG bound to the catalytic groove of chitinase A E315Q mutant is shown. Aromatic and hydrophilic residues that interact with the substrate are colored in green and red respectively; The octaNAG substrate is colored blue. Two orientations were chosen to show the length and the depth of the groove.



**Figure 3.2. Substrate torsion in the complex of chitinase A with octaNAG.** The octaNAG substrate and several aromatic residues that probably induce the torsion at the catalytic center are shown. NAG residues (blue) are indicated (see text). Color code of amino acid atoms: carbon-gray; nitrogen-light blue and oxygen-red.



**Figure 3.3. The catalytic pocket of chitinase A.** The acetamido group of the -1 NAG is imbedded in the pocket that is located at the center of the binding groove. Protein surface was calculated and colored as a hydrophobic gradient (red represents hydrophobic residues and blue hydrophilic residues).

## **Chapter IV. Analysis of the catalytic mechanism of chitinase A**

Two models were suggested for the catalytic mechanism for chitinase A. The first model suggests a double displacement (retaining) mechanism in which Glu315 acts as a proton donor and Asp391 acts as a nucleophile (Perrakis *et al.*, 1994). In the alternative substrate-assisted model Glu315 is proposed to act as a proton donor while the nucleophilic action is carried out by the terminal oxygen (O7) of the acetamido group of the -1 NAG (Tews *et al.*, 1997). In this second model the double bond of O7 is opened and its unpaired electron attacks the oxacarbonium, yielding an oxazolinium ring intermediate (see Figure D4 in the Discussion sections). The high-resolution structures of chitinase A oligoNAG-complexes allowed us to carry out a detailed analysis of the enzymatic mechanism. In this chapter and in Chapter V, there is a focus on the catalytic mechanism of chitinase A.

### **The catalytic center of wild type chitinase A**

In a collaboration with Dr. K. Petratos an improved crystal structure of wild-type chitinase A with an increased resolution (1.55 Å) was obtained [phases were taken from the native structure, PDB code: 1ctn; (Perrakis *et al.*, 1994) to be published]. This structure was found to be highly similar to the previously published structure. However, residue Asp313 was found in two conformations (Figure 4.1A). In one conformation Asp313 interacts with Glu315. This conformation was also found in other chitinases (Hollis *et al.*, 2000; Rao *et al.*, 1995; Terwisscha-van-Scheltinga *et al.*, 1994; van Aalten *et al.*, 2000; Van-Roey *et al.*, 1994) and in chitobiase (residue Asp539, see Chapter II and Discussion section). In the alternative conformation, Asp313 interacts with residue Asp311. This conformation is identical to that found

previously in the structure of chitinase A (Perrakis *et al.*, 1994), and to the conformation found in the E315Q-octaNAG complex (see below). As will be discussed below, we propose that the ability of Asp313 to alternate between two conformations plays a key role in the catalytic reaction of chitinase A.

### **Biochemical analysis of active site mutants**

During the construction of mutants by site directed mutagenesis we, developed a new genetic screen for chitinase mutants (see Materials and Methods). The principle of the screen is based on the activation of the endogenous *E. coli* gene *celF* / *chbF* as a reporter gene (Keyhani & Roseman, 1997; Parker & Hall, 1990) in the presence of diNAG (the latter is a product of chitinase A when colonies expressing the enzyme are growing on chitin plates). The CelF / ChbF enzyme, a product of the *celF* gene, cleaves X-NAG into mono NAG and an insoluble blue dye. To test the importance of specific residues for the enzymatic activity we determined the kinetic parameters of the purified wild type and mutant proteins with the substrate pNp-diNAG (Table 4.1). Similar results were obtained with the substrates 4MU-diNAG and pNp-diNAG. The mutations D313A and E315Q, both highly conserved residues, caused about a 100 - 200 fold reduction in  $K_{cat}$ . We mutated residue Asp391 that was proposed to participate in chitinase A activity (Perrakis *et al.*, 1994). This residue is conserved as either Asp or Asn. We found that D391E mutant showed a wild-type phenotype, and D391N mutant resulted in only a minor reduction in chitinase A activity. However the D391A mutant reduced the turnover number by about 250-fold. Residue Y390 is also conserved among all family 18 members, and faces the catalytic center. A mutant with a minimal change of substituting this tyrosine with phenylalanine was constructed.

Y390F showed a 200-fold reduction in  $K_{cat}$ . Thin layer chromatography assays of the hydrolysis of hexaNAG by the wild type and mutant enzymes showed that mutants Y390F and D391A retain significant activity, while the mutants D313A and E315Q were inactive (data not shown). Our results demonstrate that all mutated residues probably play an important role in the catalytic activity of chitinase A.

### **Structural analysis of the catalytic mechanism**

We have co-crystallized and solved the complex structures of all four mutants with oligoNAG as described in the Materials and Methods section and in Papanikolaou *et al.*, (2001). This work was done as a collaboration with the laboratory of Dr. Kyriacos Petratos in IMBB FORTH, Heraklion. The structures of wild-type and mutant proteins are highly similar (see Chapter III). However, specific, subtle, structural changes in residues 315 and 313 were observed. The -1 and +1 sugars and the scissile bond were identified as suggested previously (Perrakis *et al.*, 1994; Tews *et al.*, 1997). These two NAG residues are found in a stressed configuration as described in Chapter III. The acetamido group of the -1 NAG, proposed to participate in the substrate-assisted catalysis, is embedded in the pocket extending from the substrate-binding groove (see Chapter III).

The D313A mutant co-crystallized with octaNAG showed that the position of Glu315 is identical to that found in the native wild-type enzyme, while the Asp313 side chain is absent (Figure 4.1B). The sugar is present in an almost identical configuration to that present in the E315Q co-crystal (see Figure 4.1C). The distance between the carboxylic end of Glu315 and the glycosidic oxygen is 2.7 Å, a distance that is optimal for its proton donor activity. Therefore, the inactivity of this mutant

must be caused by the absence of the aspartic side chain. In addition, the distance between O7 of the *N*-acetyl group and C1 is 3.7 Å. This distance is too large for nucleophilic attack. It is possible that an interaction of Asp313 with the acetamido group is required for placing it in the energetically non-favored conformation, permitting the substrate-assisted reaction to occur (see below). The inactivity of the D313A protein could therefore also be due to the inability of O7 of the *N*-acetyl group to act as a nucleophile. Attempts to get the D313N mutant, which I predicted would have milder effects on the enzymatic activity and could have helped in the above analysis, failed.

In the E315Q-octaNAG complex the Glu to Gln change is clearly observed. In the complex, the side-chain of Gln315 is found in two conformations (Figure 4.1C). One of these conformations is identical to that found in the wild-type enzyme in the absence of a substrate, and in the Y390F and D391A complexes (see Chapter V). The simplest explanation for the inactivity of the E315Q mutant is the inability of Gln315 to act as a proton donor. In addition, the O7 of the *N*-acetyl group is positioned 3.9 Å from the C1 carbon, probably too far and at an improper angle to carry out the nucleophilic attack efficiently.

We have co-crystallized the D391A and Y390F mutants with hexaNAG. The overall structure of these mutants is very similar to the wild-type and to the other mutants, and the mutational change could be clearly identified. However, in these complexes Asp313 occupies only the conformation in which it is next to Glu315. In this conformation Asp313 interacts with the N2 of the acetamido of the -1 NAG. This interaction forces the rotation of the acetamido group into a non-favored conformation

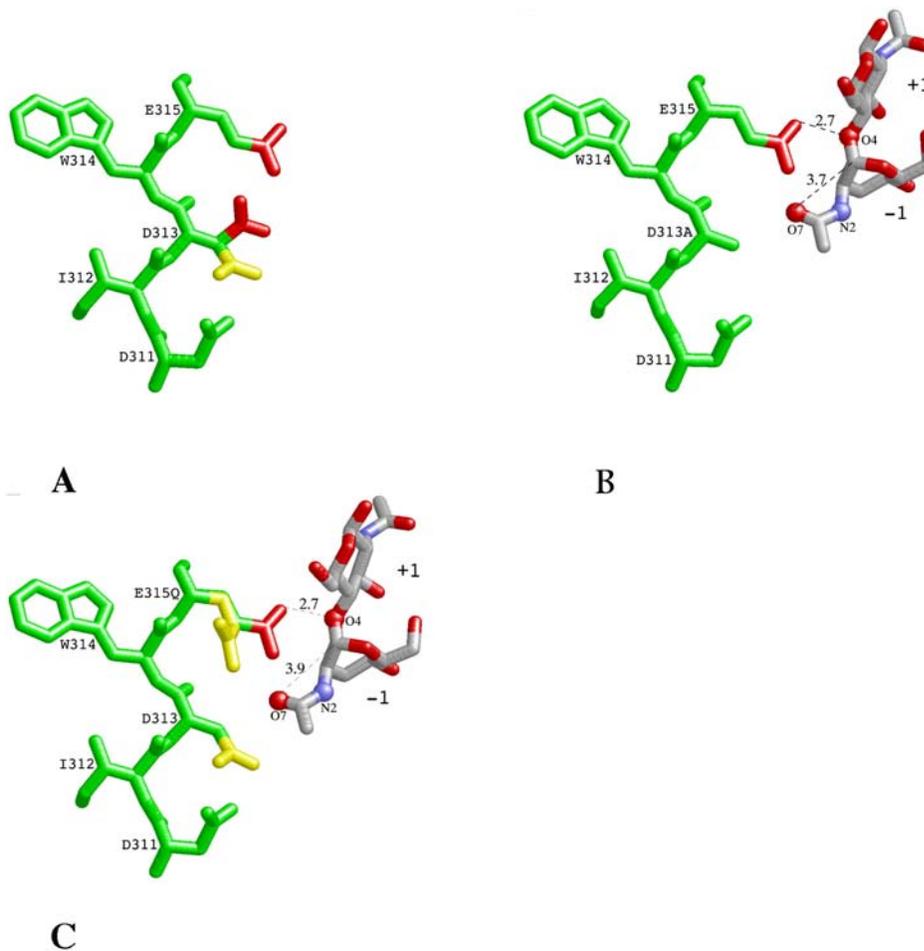
required for the substrate-assisted reaction. The importance of this rotation is discussed in Chapter V.

In conclusion, the highly conserved Asp313 and Glu315 residues were shown to play an important role in the catalytic activity of family 18 chitinases. Glu315 acts as a proton donor. Residue Asp313 probably reduces the activation energy of the reaction by rotating the acetamido of the -1 sugar, a rotation required for the proposed substrate-assisted reaction. The partial negative charge of Asp313 is postulated to further stabilize the intermediate by reducing the positive charge of the N2 in the oxazolinium ring formed during the nucleophilic attack by the acetamido O7 atom of the -1 sugar.

**Table 4.1 Kinetic parameters of wild-type and mutant chitinase A proteins**

<b>Protein type</b>	<b><math>K_{cat}</math> (Sec<sup>-1</sup>)</b>	<b><math>K_M</math> (<math>\mu</math>M)</b>	<b>Relative <math>K_{cat}/K_M</math></b>
Wild type	331 $\pm$ 6.20	52 $\pm$ 2	1.0000
D313A	1.6 $\pm$ 0.09	296 $\pm$ 5	0.0008
E315Q	4.9 $\pm$ 0.12	530 $\pm$ 9	0.0014
Y390F	1.4 $\pm$ 0.10	87 $\pm$ 2	0.0025
D391A	1.2 $\pm$ 0.11	260 $\pm$ 8	0.0007

Kinetic assays were performed with pNp-diNAG as described in the Methods section. The results represent averages of 3 independent experiments with standard errors. These mutants were unable to degrade colloidal chitin.



**Figure 4.1. Structure of chitinase A-oligoNAG complexes.** The position of -1 and +1 sugar residues, the catalytic residues D313 and E315 are shown. (A) Wild type structure. (B) D313A – octaNAG complex. and (C) E315Q – octaNAG complex. Amino acid residues are colored in green with the exceptions of one of the two conformations of Asp313 and Q315 which are colored in red and yellow. Sugar atoms are colored as follows: carbon gray; nitrogen light blue and oxygen red. Distances are indicated in Angstrom.

## **Chapter V. Trapping the reaction intermediates in chitinase A D391A mutant**

The substrate-assisted model for the catalytic activity of chitinase A implies that a terminal oxygen atom (O7) of the -1 NAG acetamido group acts as a nucleophile. Thus, an oxazolonium ring is predicted to exist as an intermediate (see scheme 1 in the Discussion section). The complexes of E315Q and D313A mutants with octaNAG revealed pre-Michaelis structures in which catalysis was not initiated. Thus it was not surprising that no intermediate was observed (see Chapter IV). This chapter describes our attempt to trap reaction intermediate structures of chitinase A. Shortening the reaction time by soaking the enzyme with the substrate, followed by a rapid quenching of the reaction at a low temperature (under cryogenic conditions) is now a common technique. However, in many cases the enzymatic turnover is considerably higher than the diffusion of the substrate into the active sites at the crystal lattice. Thus a combination of substrate, intermediates and product structures are trapped, which leads to fuzzy electron density maps. We therefore selected to analyze a mutant protein with a reduced turnover rate and to trap possible intermediates by observing the crystals under cryogenic conditions.

### **Trapping of reaction intermediates**

Chitinase A D391A mutant showed low but significant activity against oligoNAG, as observed by analyzing the hydrolysis of native substrates (not shown). To trap intermediate complexes, crystals of the D391A mutant were soaked with hexaNAG or with tetraNAG for 1, 10, 30, 60 and 180 minutes, and flash cooled (100° K) upon mounting for data collection (Mesecar *et al.*, 1997) and see the Material and

Methods section. Diffraction data were collected using synchrotron radiation at the EMBL DESY, Hamburg with the kind help of Dr. Paul Tucker. A full data set from a single crystal that was soaked with tetraNAG for one minute was collected to 1.8 Å resolution. A longer soaking time yielded poor diffraction due to increased crystal mosaicity. The aspartic to alanine substitution and the tetraNAG were clearly observed in the crystal. Surprisingly, the tetraNAG was found to be bound at the -1, +1, +2 and +3 subsites (Figure 5.1). The +3 NAG skips over the Tyr418 residue at the end of the substrate-binding groove. In contrast to the structures of E315Q complex, Asp313 was found to be interacting with Glu315 and the acetamido N2 atom of the -1 NAG. This conformation was found in other chitinases, including the complex of hevamine with allosamidin (Terwisscha-van-Scheltinga *et al.*, 1995) and the improved resolution structure that was recently deposited in the Protein Data Bank (1edq).

Inspection of the electron density maps revealed that most of the substrate is intact, however, some of it is probably partially cleaved. All the sugar residues have a relaxed 'chair' conformation. A critical conformational change from 'chair' to 'boat' is found at the -1 NAG. This high-energy conformation is a result of the binding forces at the -1 subsite, and probably plays a critical role in the enzymatic activity. It seems that the -1 NAG may assume more than one conformation. However the electron density of the -1 NAG is somewhat less clear, thus we could not model more than one conformation. We speculate that several conformations of -1 NAG residue may represent the movement from 'chair' to 'boat' structures upon substrate binding. The reverse process from 'boat' to 'chair' probably takes place after cleavage of the glycosidic bond. In the intact structure of the oligoNAG, the -1 NAG to +1 NAG

residues are distorted and tilted by about 90° around the glycosidic bond (Figure 5.2). This distortion is very similar to the structures found in the chitinase A D313A-octaNAG and E315Q-octaNAG complexes and also in chitobiase-diNAG complexes.

Strikingly, we found that the position of the acetamido group of the -1 NAG is in an energetically non-favored conformation in which the terminal O7 faces the C1 oxacarbonium (Figures 5.1 and 5.2). This conformation appears to be stabilized by the interactions of the acetamido N2 with O $\delta$ 2 of Asp313 and of acetamido O7 with the hydroxyl group of Tyr390 (Figure 5.2). The acetamido group is in a position which could allow a nucleophilic attack of the oxacarbonium, as would be expected from the substrate-assisted model. Note that in the complexes of D313A-octaNAG and E315Q-octaNAG, the position of the acetamido group of the -1 NAG is rotated by about 180°. No density extending between O7 and C1 of the -1 NAG was observed, suggesting that the nucleophilic attack did not take place. Moreover, the distance between these two atoms is 2.9 Å, which is too large for a covalent bond formation. Such a structure was also observed in the complexes of chitobiase (Prag *et al.*, 2000; Prag *et al.*, 2001; Tews *et al.*, 1996a), and see also Chapter II and Discussion). The lack of a covalent bond between O7 and C1 is probably due to the fact that most of the substrate is not yet cleaved.

A structural comparison of the complexes of D391A-teraNAG with D313A-octaNAG and E315Q-octaNAG showed that the acetamido group of the -1 NAG is rotated by about 180°. We suggest that this complex represents a trapped reaction intermediate.

### **Analysis of the co-crystal of D391A mutant with hexaNAG**

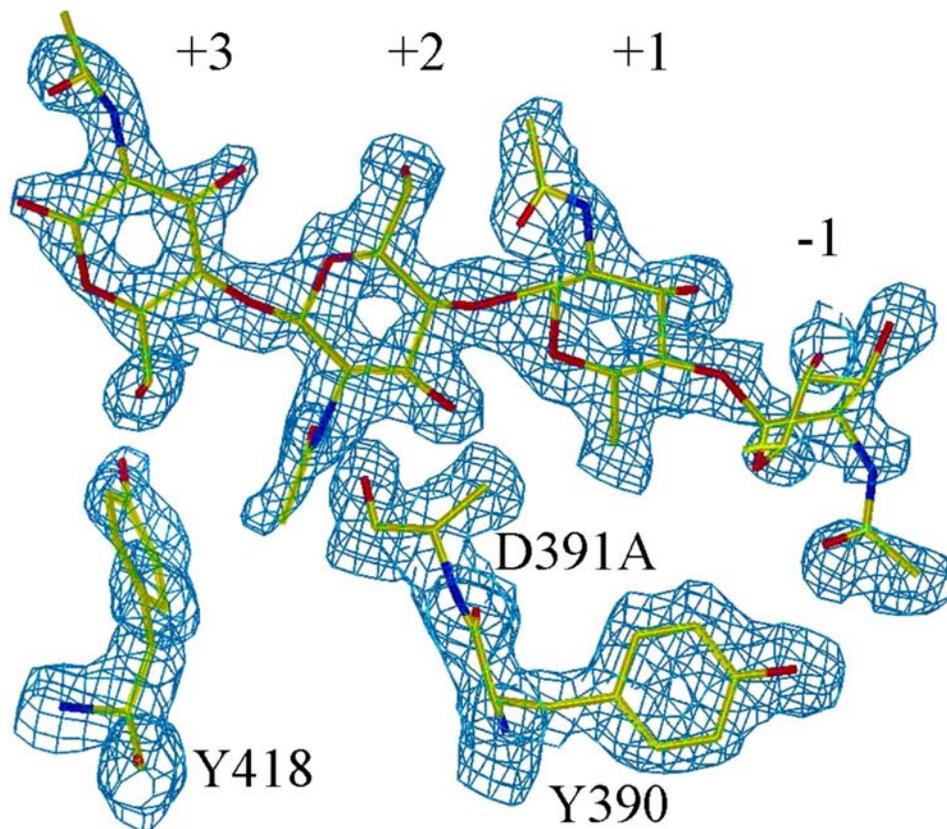
We have co-crystallized the D391A mutant with hexaNAG, and solved the structure under cryogenic conditions using Cu-K $\alpha$  X-ray generator at a 1.9 Å resolution (see Methods). The structure revealed a complex in which the hexaNAG was cleaved between the -1 and the +1 subsites (Figure 5.3A). These results indicate, for the first time, where cleavage takes place. The alanine substitution at residue 391 is clearly seen (Figure 5.3B). In addition, Glu315 assumes an identical conformation as we found in the soaking experiment. These results provide further support for the suggestion that Glu315 acts as the proton donor. All other critical residues are found to be identical in structure to those observed in the soaking experiment. We could easily model the structure of NAG residues at the -4, -3, +1 and +2 subsites based on our electron density maps. No binding to the +3 subsite was observed, suggesting that this subsite has low affinity to the substrate. However the electron density of the -2 NAG residues was found to be broken and incomplete (Figure 5.4). Using the coordinates of the sugar residues from our previous complex D313A-octaNAG allowed us to model the location of the NAG residue (Figure 5.4). However, the electron density surrounding the -1 NAG showed that the structure of the NAG residue is different from that found with D313A-octaNAG and D391A-tetraNAG complexes (Figure 5.4). The disagreement between these models and the fact that the oligoNAG was found to be cleaved, led us to try to model the -1 NAG residue with an oxazolonium-intermediate structure. Based on recently solved complex of chitinase A with allosamidin (Petratos K. unpublished data) and the D391A-tetraNAG complex,

we have built a model of the intermediate containing an oxazolonium ring at the -1 subsite. The complexes of D391A-hexaNAG and hevamine-allosamidin were superimposed, revealing the superposition of the -1 NAG with the cyclopenta-allosamizoline. The coordinates of the -1 NAG were then manually fitted to assume a structure similar to the cyclopenta-allosamizoline moiety (using the program 'O') in order to generate an oxazolonium ring with the pyranoside in a 'chair' conformation. Finally, the coordinates of the intermediate were energetically minimized using a consistent valence force field procedure (cvff of *Insight-II Discover*), and added to the CCP4 dictionary for further refinement. This procedure improved the electron density maps as well as the final refinement statistic parameters. In an attempt to improve our model, omitted maps were further calculated. Comparison of the  $2mFo-DFc$  with the  $mFo-DFc$  ( $\sigma$ -weighted) maps (Read, 1986), with the coordinates of the -1 and 2 NAG residues omitted from the refinement process, showed the electron density around the -1 NAG with the oxazolonium ring structure (Figure 5.5).

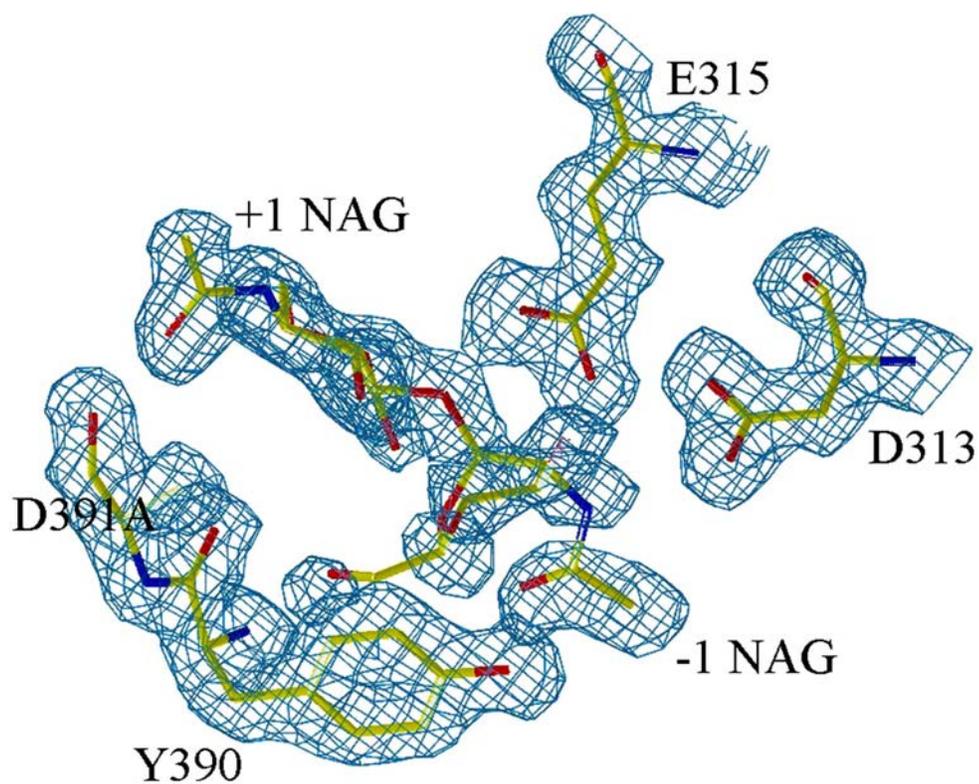
### **The possible function of Asp391**

Based on the results presented above, it is difficult to clearly define the function of Asp391. A number of hypotheses explaining why the D391A mutant protein has low activity can be considered. According to the simplest hypothesis, the mutant is defective in the proper binding of the substrate. As shown in Table 3.1, Asp391 forms interactions with three NAG residues at the -1, +1 and +2 subsites. At an early stage of this study we substituted Asp391 with asperagine and glutamic acids. We found that both D391N and D391E mutants had wild-type phenotypes and these mutant proteins were not further studied. Inspection of the interactions that Asp391

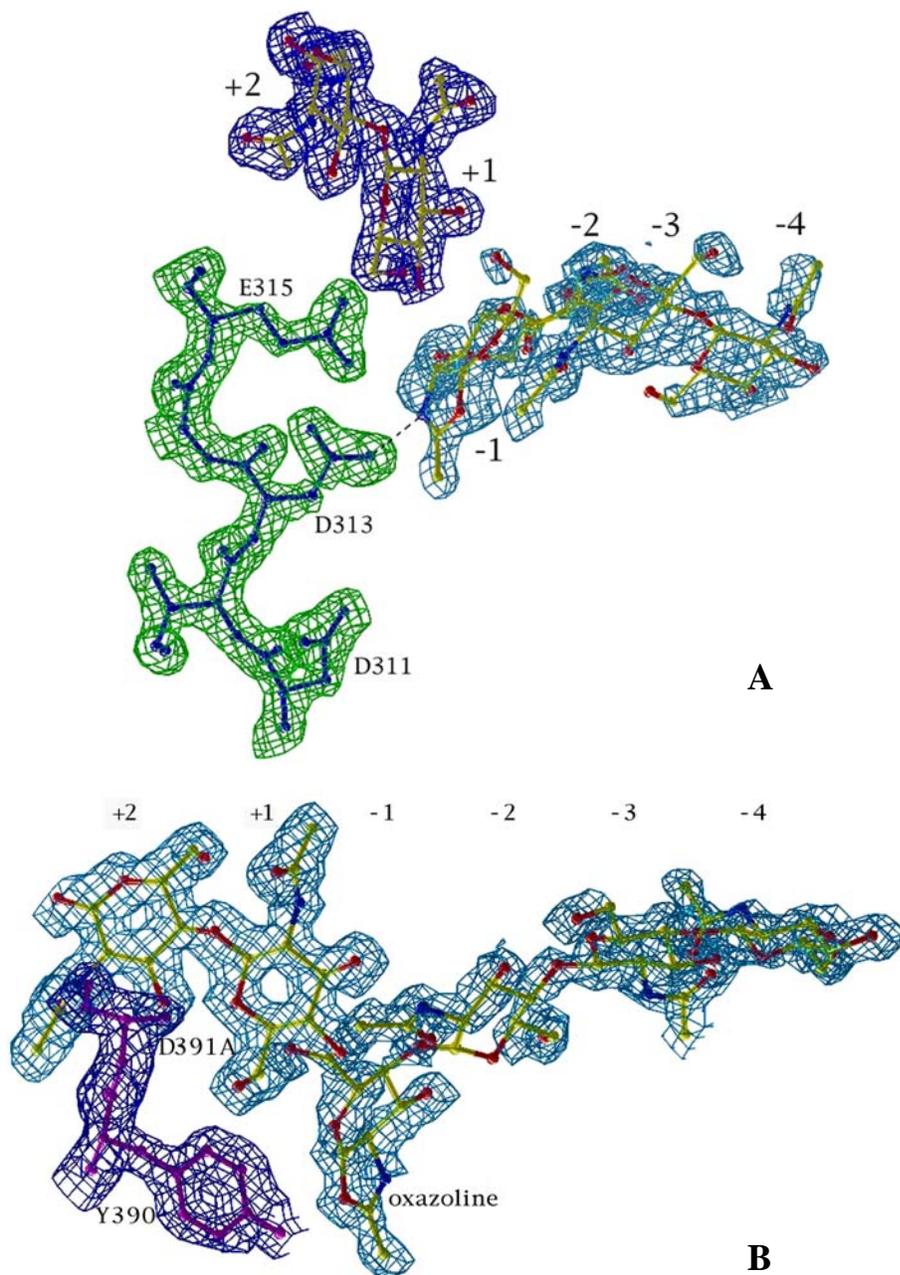
makes with the substrate (in the E315Q or D313A mutant complexes), suggests that both D391N and D391E substitutions can accommodate all these interactions. Alternatively, D391 could act as a nucleophile, as was shown in a number of glycosyl hydrolase protein families (Davies & Henrissat, 1995; Rye & Withers, 2000).



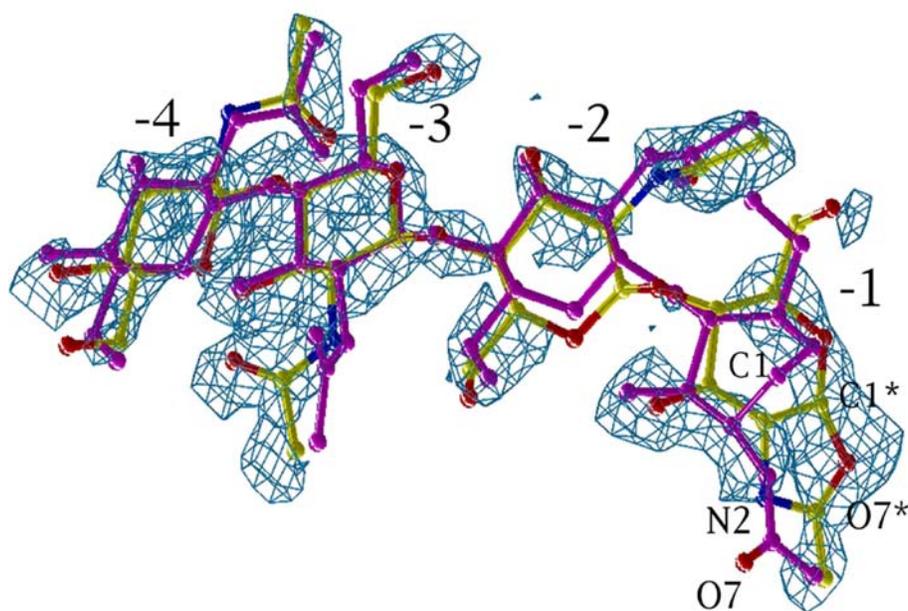
**Figure 5.1. A +3 NAG residue at the D391A-tetraNAG complex.** The electron density map shows the D391A-tetraNAG complex. The positions of sugar residues along the binding groove are indicated. The +3 NAG residue is clearly seen above Tyr418. The location of the +2 NAG and +1 NAG residues are identical to those we found in all other complexes. A  $2mF_{obs}-DF_{cal}$  map was calculated from the last refined model and contoured at  $1.2 \sigma$ .



**Figure 5.2. The acetamido group conformation of the -1 NAG favors substrate-assisted mechanism.** The electron density map focuses on the catalytic center of the complex of D391A-tetraNAG. Interactions between Y390 and O7 of the acetamido group of the -1 NAG, and between the glycosidic oxygen and E315, are shown. A  $2mF_{obs}-DF_{cal}$  map was calculated from the last refined model and contoured at  $1.2 \sigma$ .



**Figure 5.3. Cleaved substrate trapped at the active site of chitinase A mutant.** The electron density maps show the co-crystal of D391A with hexaNAG. **(A)** The map clearly shows that the substrate is cleaved between the +1 and the -1 subsites. Interaction between D313 and N2 of the proposed oxazolinium ring is also shown. **(B)** From a different angle (to the above there are). Shown a proposed formation of oxazolinium ring, interaction between Y390 and O7 of the -1 NAG, and the alanine substitution at residue 391. These  $2mF_{obs}-DF_{cal}$  maps are contoured at  $1.0 \sigma$ .



**Figure 5.4. Structural comparison of the substrates from D391A-hexaNAG and E315Q-octaNAG complexes.** The figure focuses on the -1 to -4 NAG residues. Coordinates of the two complexes were superimposed based on C $\alpha$  atoms of the TIM-barrel domains. The shown electron density map was calculated from the D391A-hexaNAG co-crystal as described in the text ( $2mF_{obs}-DF_{cal}$  contoured at  $1.2 \sigma$ ). While the position of the N2 atoms remained in the same position, the positions of C1, C7, C8 and O7 atoms were changed. \* Indicate the proposed positions of atoms of the D391A complex.

## Discussion

*Serratia marcescens* has been used as a model system for the utilization of chitin as a carbon source. In the presence of chitin, the bacterium expresses chitinase A, chitinase B, chitinase C and chitobiase that degrade chitin to mono *N*-acetylglucosamine (NAG) (Warren, 1996). In addition, these cells express a chitin binding protein (CBP21) with an unknown function (Chu *et al.*, 2001; Folders *et al.*, 2000). Although intensive studies of a system for chitin degradation have been studied for over a decade it is yet not fully understood why *S. marcescens* requires the activity of four enzymes. The structures of *S. marcescens* chitinolytic enzymes, chitinase A, chitinase B and chitobiase were solved and their catalytic domains were found to possess a TIM ( $\alpha/\beta$ )-barrel structure (Perrakis *et al.*, 1994; Tews *et al.*, 1996a; van Aalten *et al.*, 2000). In contrast, very little is known about the function and structure of chitinase C (Suzuki *et al.*, 1999). Chitinase A, B and C were assigned to the glycosyl hydrolase protein family 18, while chitobiase belongs to the protein family 20 (Henrissat, 1991). In these enzymes the active site is located at the carboxy-terminal end of the  $\beta$ -strands of the TIM-barrel, suggesting that all these proteins, although differing greatly in their primary sequences, have a common origin.

Information on the co-crystal structure of chitinase-substrate complexes was scant. Earlier attempts to obtain the co-crystal structures of oligoNAG-chitinase complexes in which the non-cleaved substrate occupies the catalytic site proved unsuccessful (Perrakis *et al.*, 1994; Terwisscha-van-Scheltinga *et al.*, 1996). Creating mutations of the catalytic residues of *S. marcescens* chitinase A and

chitobiase allowed us to obtain co-crystals with the native (Prag *et al.*, 2000; Prag *et al.*, 2001). Our investigations led us to suggest that *S. marcescens* enzymes participating in chitin metabolism are all employed a similar catalytic mechanism.

### **Analysis of Chitinase A, chitinase B and chitobiase substrate binding**

Chitinase A possesses a groove with at least eight subsites in which OctaNAG faces the FnIII-like domain on one side and is bound by residue Tyr418 at the other. Inspection of the FnIII-like domain suggests that longer chitin chains could interact with a number of aromatic residues located on its surface (Figure D.1) with the most distant tryptophan residue 45 Å from subsite -6. This arrangement could allow the interaction of about 11 additional NAG residues over and above the 8 residues bound to the groove. The FnIII-like domain is tethered to the TIM-barrel domain by a flexible hinge (Perrakis *et al.*, 1994), which could permit a different special arrangement between these two domains. We suggest that this permits chitinase A to cleave diNAG without dissociation from the substrate and that it may act processively. A similar mechanism was previously suggested for cellulases, in which the cellulose binding module binds to the cellulose via aromatic residues (Sakon *et al.*, 1997; Irwin *et al.*, 1998). We speculate that a mutation in the Tyr418 residue may alter the specificity of chitinase A to act preferentially as an endochitinase. In contrast to the TIM-barrel, the FnIII-like domain may bind oligoNAG at its surface or within the domain, following unfolding and refolding of the FnIII-like domain, as shown for the FnIII domain of the muscle protein titin (Klimov & Thirumalai, 2000; Zhang *et al.*, 1999). The inner conserved aromatic residues in the FnIII-like domain may participate in such a mechanism (Figure D.1). Chitobiase binds the oligomers with

only five subsites or less, requiring a free non-reducing end. Moreover, instead of an open groove, the active site of chitobiase has a boot shape tunnel.

### **Atomic level conservation at the active site**

Multiple alignments of glycosyl hydrolases protein families 18 and 20, show the conservation of residues in the  $\beta$ 4-loop #4; DXXDXDXE in family 18 and HXGGDE in family 20. Both signatures include Glu315 of chitinase A and Glu540 of chitobiase that act as the proton donor [(Perrakis *et al.*, 1994; Prag *et al.*, 2001; Tews *et al.*, 1996a) and this thesis] (Figure I.4 in the introduction). It is interesting to examine the function of these conserved residues. In chitinase A the catalytic Glu315 is located at the beginning of loop#4 of the TIM-barrel. A string of conserved Asp residues of the DXXDXDXE motive transverse the TIM-barrel forming a dipole together with additional residues. A similar set of acidic residues could be also identified in chitobiase including residues Glu540, Asp539, Asp448 and Glu446. It is possible these essential Asp residues provide an additional negative charge that somehow affects the reaction. For example, it is possible that once Glu donates its proton by attacking the glycosidic bond, the acidic residues generate a negative charge gradient that favors the forward reaction.

As discussed in the Introduction, the TIM-barrel structure is found in a large number of proteins whose structure was solved. One may wonder why chitinase A and chitobiase utilize for their catalytic domain the TIM-barrel structural motive. We speculate that this structure provides a flexible framework to which the catalytic residues are attached. Evolutionary changes of the TIM-barrel can thus be translated

to changes in the position of residues involved in substrate binding and catalysis leading to improved enzymatic activity.

We have mutated the conserved catalytic Asp and Glu residues mentioned above of both chitinase A and chitobiase and found that both residues are essential for the catalytic activity. The structure of the catalytic residues Asp and Glu are similar in both enzymes (Figure D.2). However, in chitinase A Asp313 appears, in the wild-type, in two alternative conformations. Comparison of the structures of the enzyme-substrate complexes showed a high similarity in the position of the -1 to +1 sugars. It was also observed that in both enzyme-substrate complexes the planes of the sugars at -1 and +1 are tilted around the scissile bond in a similar manner. In both enzymes a critical chair to boat sugar conformational change is involved in the bending and rotation of the substrate upon binding (Figure D.2). These energetically non-favored structures may favor the hydrolysis reaction. Furthermore, the distance between the proton donor and the glycosidic oxygen is conserved. These findings suggest that chitinase A, chitobiase and probably chitinase B possess similar catalytic sites. The information obtained in this work raises several puzzling questions: How do these enzymes recognize the synthetic chromogenic substrates and why is chitinase unable to degrade the chitobiase substrate pNpNAG? The answers to these questions awaits further crystallographic studies.

The sequences of proteins belonging to the glycosyl hydrolase families 18 and 20 are too different to be aligned on the basis of their primary sequences only. However, the structure of the catalytic TIM-barrel domain is highly conserved. We found that the structure of residues in the conserved motives (DE of chitobiase and

DWE of chitinase A) together with the structure of the -1/+1 diNAG are sufficient to obtain atomic structural alignment of C $\alpha$  of chitobiase with chitinase A. Chitinase B could be aligned with chitinase A, as these proteins are highly similar (Figure D.3). This analysis revealed, for example, a conserved tyrosine residue (Tyr669 of chitobiase and Tyr390 of chitinase A). Mutations Y669E and Y390F showed reduced activity.

### **A model of the catalytic mechanism**

In general, glycosidases degrade carbohydrates by a general acid-base catalysis that involves two amino acid residues, a proton donor and a nucleophile. Hydrolysis of the scissile bond results in either the retention or the inversion of the C1-carbon anomeric configuration. While the catalytic glutamate is highly conserved, no amino acid residue that could act as a nucleophile was identified in protein families 18 and 20 (see below). A different catalytic mechanism was therefore proposed. Based on the complex of Hevamine (a family 18 plant chitinase) with allosamidin and on the complex of chitobiase with diNAG, it was suggested that catalysis takes place via a substrate-assisted mechanism (Figure D.4), (Brameld *et al.*, 1998; Prag *et al.*, 2000; Koshland Jr., 1953; Papanikolau *et al.*, 2001; Terwisscha-van-Scheltinga *et al.*, 1994; Tews *et al.*, 1996a). Our data lead us to favor the substrate-assisted model. In this proposed mechanism the glutamate residue acts as a proton donor, while the terminal oxygen (O7) of the acetamido group of the -1 NAG acts as a nucleophile. In order to act as a nucleophile the acetamido group has to rotate around the C2-N2 bond as is suggested by the studies on both enzymes. The precise positioning of the O7 atom is achieved by its

interaction with a tyrosine residue (Tyr390 in chitinase A, Tyr669 in chitobiase). In addition, the N2 atom of the acetamido interacts with an aspartic residue (Asp313 in chitinase A, Asp539 in chitobiase). Following the nucleophilic attack an oxazolinium ring is formed which stabilizes the intermediate. This stabilization is probably achieved by formation of a hydrogen bond between N2 of the acetamido and Asp313 or Asp539 of chitinase A and chitobiase respectively. To complete the reaction a water molecule is split. The proton is thus trapped by the glutamate residue and the OH<sup>-</sup> completes the hydrolysis by attacking the C1 atom of the intermediate. Direct evidence for substrate-assisted catalysis is not yet available (i.e. clear demonstration of the oxazolinium intermediate), probably due to the short half life time of the intermediate. The following evidence supports the substrate-assisted model:

1. The removal of the acetyl group from the -1 NAG inhibits the activity of chitobiase (Drouillard *et al.*, 1997).
2. A perfect inhibitor of these enzymes should mimic an intermediate of the reaction. The known inhibitors, allosamidin and thiozolinium, indeed contain an oxazolinium ring (Knapp *et al.*, 1996).
3. Complexes of hevamin and chitinase A with allosamidin and the complex of  $\beta$ -Hexosaminidase from *S. plicatus* with NAG-thiozoline demonstrate the presence of the oxazolinium rings at the exact position of the proposed intermediate (Mark *et al.*, 2001; and see Chapter V).

4. Indirect support for the formation of a covalent oxazolinium ring intermediate is suggested by the ability of chitinase to produce high molecular weight chitin from a derivative of diNAG-oxazoline (Kobayashi *et al.*, 1996).

The following points doubt the alternative acid-base mechanism of hydrolysis:

1. Asp391, which was suggested to act as a nucleophile in chitinase A, is not conserved and is found as Asp or Asn. Asperagine is known to be a poorer nucleophile. Mutations D391N and D391E retain wild-type activity.
2. No candidate residue that could act as a nucleophile in chitinase could be identified.
3. In retaining enzymes the distance between the carboxylic oxygen atoms of the proton donor and the nucleophile is 4.5 to 5.5 Å. The closest aspartate residue that could have acted as a nucleophile is 6.6 Å and 7.7 Å away from the proton donor in chitinase A and chitinase B, respectively.

### **The function of the essential aspartic residue**

We have shown that one function of chitinase Asp539 and chitinase A Asp313, is to ensure the placement of the acetamido group in the conformation favoring catalysis via the substrate assisted catalysis. In addition, Asp539 appears to be required for restraining the movement of Glu540. In both enzymes, Asp313 and Asp539 may provide an additional negative charge at the active site and stabilize the partial positive charge of the acetamido group while forming the oxazolinium ring.

The improved crystal structure of wild-type chitinase A demonstrates that Asp313 is found in two alternative conformations. This finding suggests that Asp313

is flexible and that this flexibility is an essential part of the catalytic reaction. We propose that when Asp313 faces away from Glu315 it improves the binding of the substrate. Furthermore, when the hydrolytic reaction is initiated 313 helps to move the acetamido group to the proper position for its action as a nucleophile. Asp539 of chitobiase may have a similar function.

### **Evolutionary considerations**

This study provides further information on the possible evolution of the pathway for chitin metabolism. The large differences in amino acid sequences and the presence of additional domains led to the suggestion that proteins possessing a TIM-barrel domain evolved by convergent evolution (Branden & Tooze, 1991; Perrakis *et al.*, 1996). Our analysis leads us to favor the hypothesis that the genes coding for protein families 18 and 20 diverged from a common ancestral gene coding for a TIM-barrel domain. These enzymes diverge to acquire different substrate specificity. However, evolutionary forces conserved  $\beta$ 4-loop #4 with its catalytic aspartate and glutamate residues. Further tinkering led to the establishment of the signatures around the catalytic sites. Interestingly, family 19 chitinases that do not possess a TIM-barrel structure, utilize an alternative, acid-base catalytic mechanism (Brameld & Goddard III, 1998; Hart *et al.*, 1995). We anticipate that it will be possible to change the specificity of families 18 and 20 enzymes by modifying their substrate binding properties without altering the catalytic residues (Altamirano *et al.*, 2000).

We searched whether the unique structure of the DXE and DE motives within the DXXDXDXE and HXGGDE signatures (see Figure I.4 in the

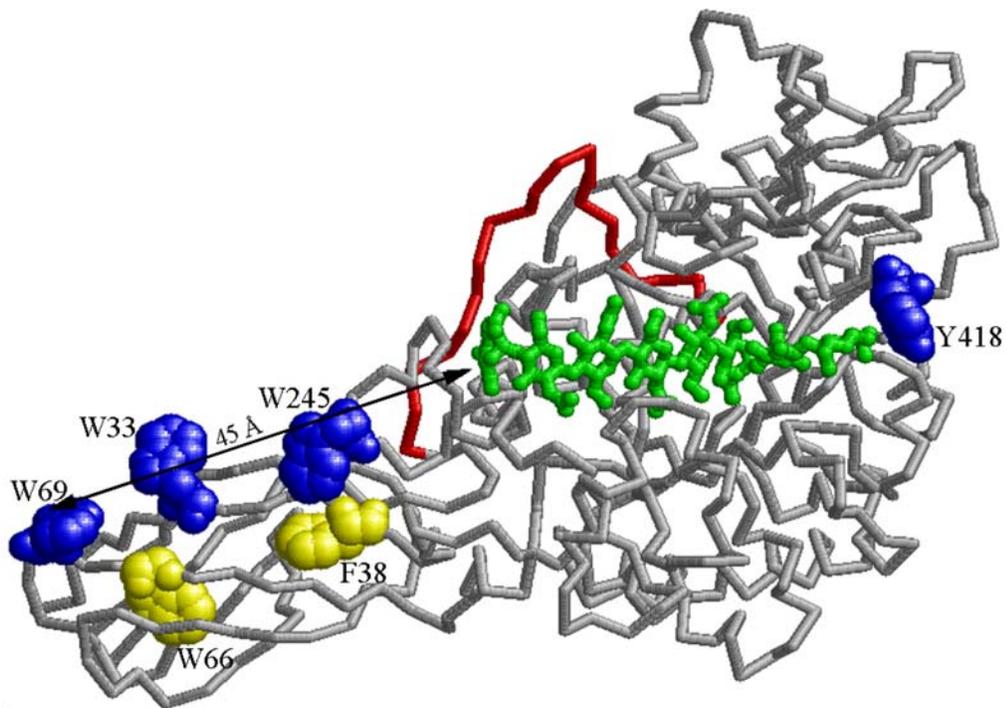
Introduction section) is conserved in other glycosyl hydrolase protein families. We found similar structures in at least 9 glycosyl hydrolase families in which the residue next to E were either D, E, N or Q. The function of these structurally conserved auxiliary residues merits further investigation.

Several members of family 18, such as the plant seed storage proteins concanavalin B and narbonin, are known to be devoid of catalytic activity (Figure I.4; Hennig *et al.*, 1995; Hennig *et al.*, 1992; Sun *et al.*, 2001). The structure of the inactive catalytic domain of these proteins is similar to chitinase A and to other chitinases. The conserved motif is modified to DXXDXDXQ in concanavalin B and to DXXDXHXE in narbonin (Terwisscha-van-Scheltinga *et al.*, 1996). Similarly, two human proteins, C3L1 and OGP39, (carrying the DXXDXAXL and DXXDXFXL respectively) were recently identified as inactive members of family 18 proteins (Malette *et al.*, 1995). Our model for the function of Asp313 and Glu315 of chitinase A provides a molecular explanation for the lack of catalytic activity of these proteins.

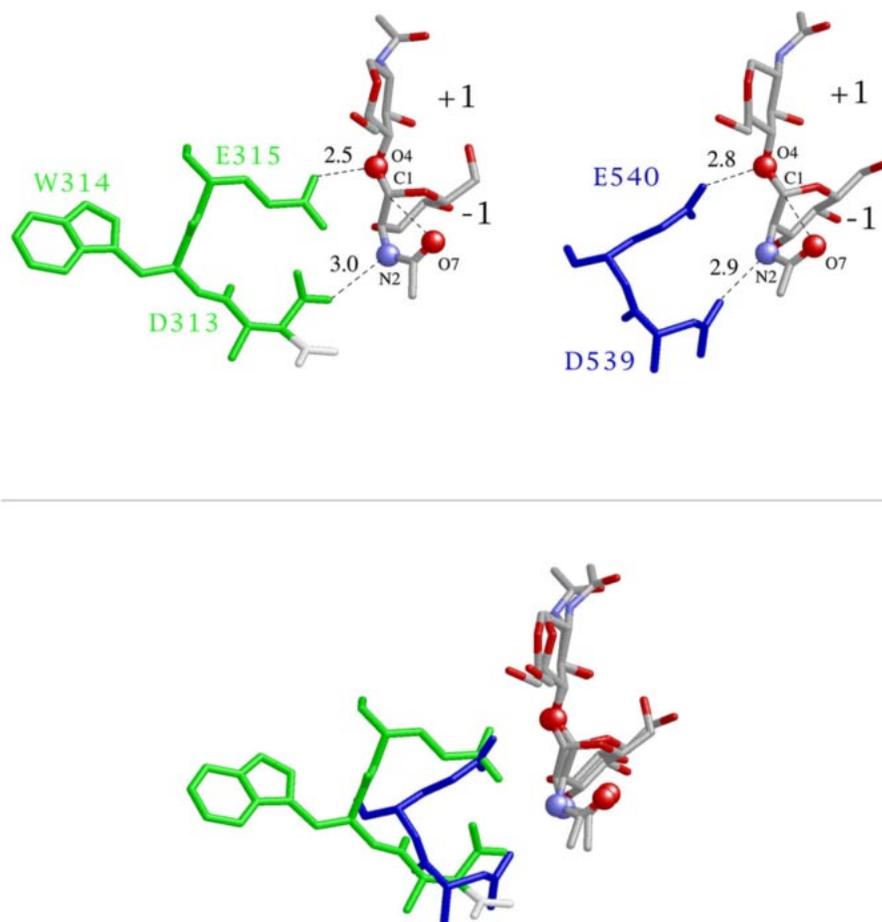
### **The cooperative action in chitin degradation**

The biochemical and structural information presented here provides a more complete picture to explain how *S. marcescens* employs its enzymes in chitin degrading (Figure D.5). Chitin degradation is probably initiated by the action of endochitinases, chitinase A and chitinase C (Watanabe *et al.*, 1997). Chitinase A also acts as an exochitinase, cleaving mainly diNAG from the reduced end of the chitin. Chitinase B acts as an exochitinase cleaving triNAG and diNAG from the non-reducing end of the chitin oligoNAG chains generated by the action of chitinase A

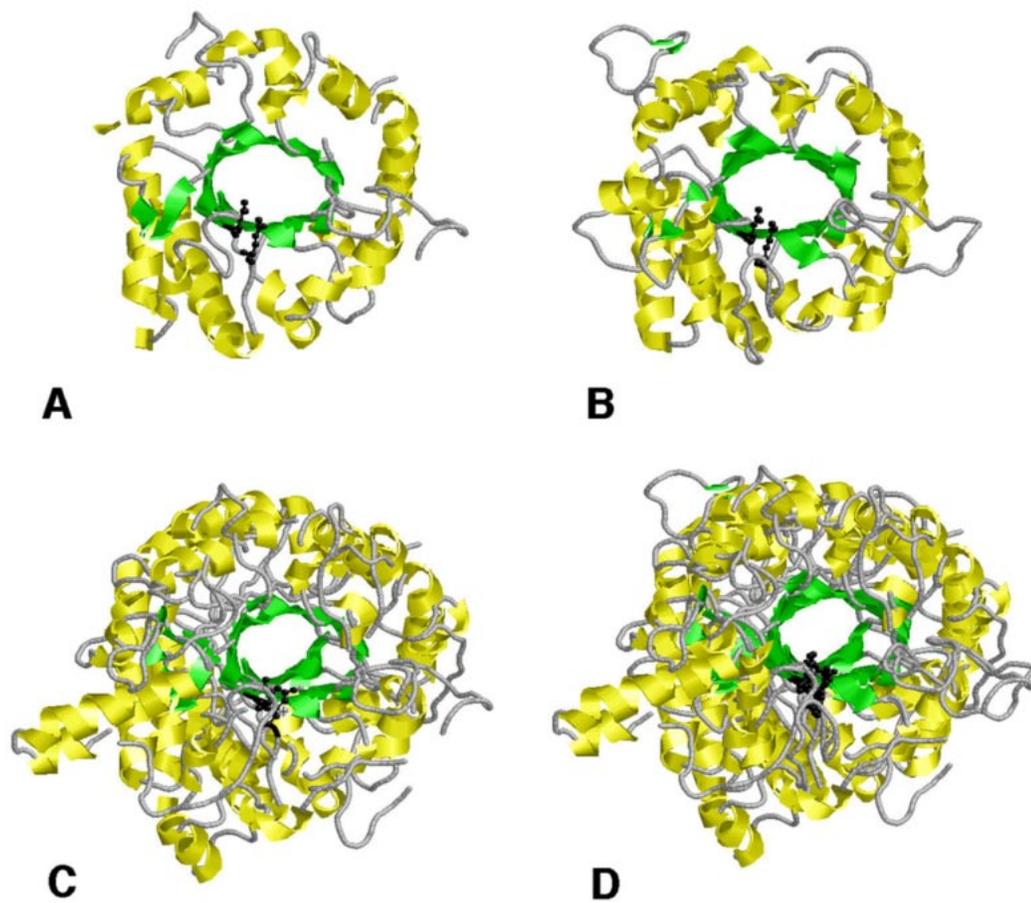
and C (Brurberg *et al.*, 1996; van Aalten *et al.*, 2000). Short oligomers, triNAG and diNAG products, are subsequently degraded to metabolizable NAG monomers by the action of chitobiase. This analysis explains the synergistic activity of these enzymes, which has been reported previously (Brurberg *et al.*, 1996; Suzuki *et al.*, 1999). It would be interesting to examine whether the cooperative action of these proteins is reflected in protein-protein interactions generating a higher order chitin-degrading complex as found in the cellulose degradation (Mechaly *et al.*, 2001; Mechaly *et al.*, 2000).



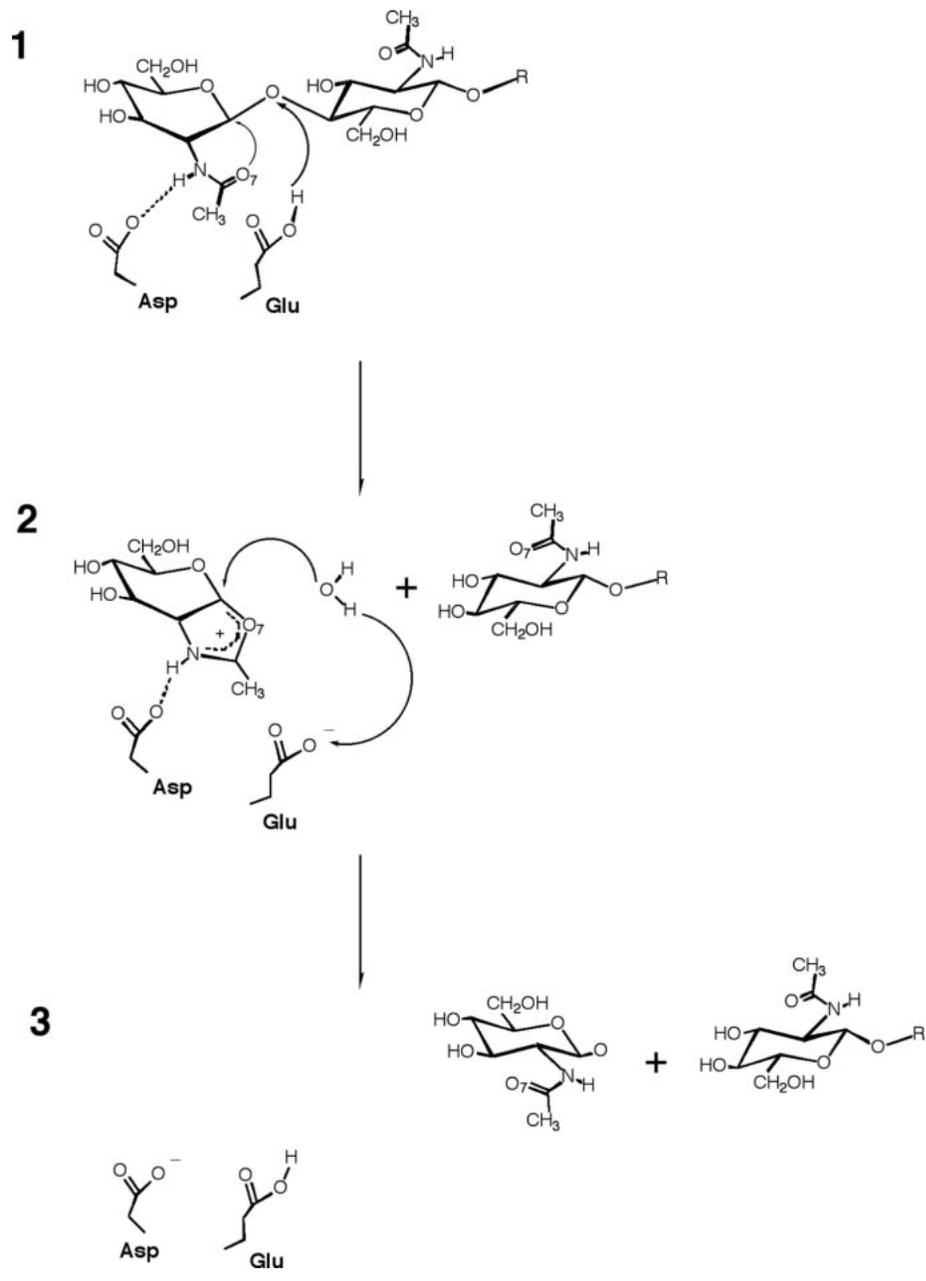
**Figure D.1. Proposed action of the FnIII-like domain of chitinase A.** Aromatic residues on the surface of the FnIII-like domain and Tyr418 are colored blue. Two aromatic residues that are partially buried within the FnIII-like domain are colored yellow. The flexible hinge between the FnIII-like and the TIM-barrel domains is colored red. The arrow indicates the distance from the -6 NAG to the distal tryptophan.



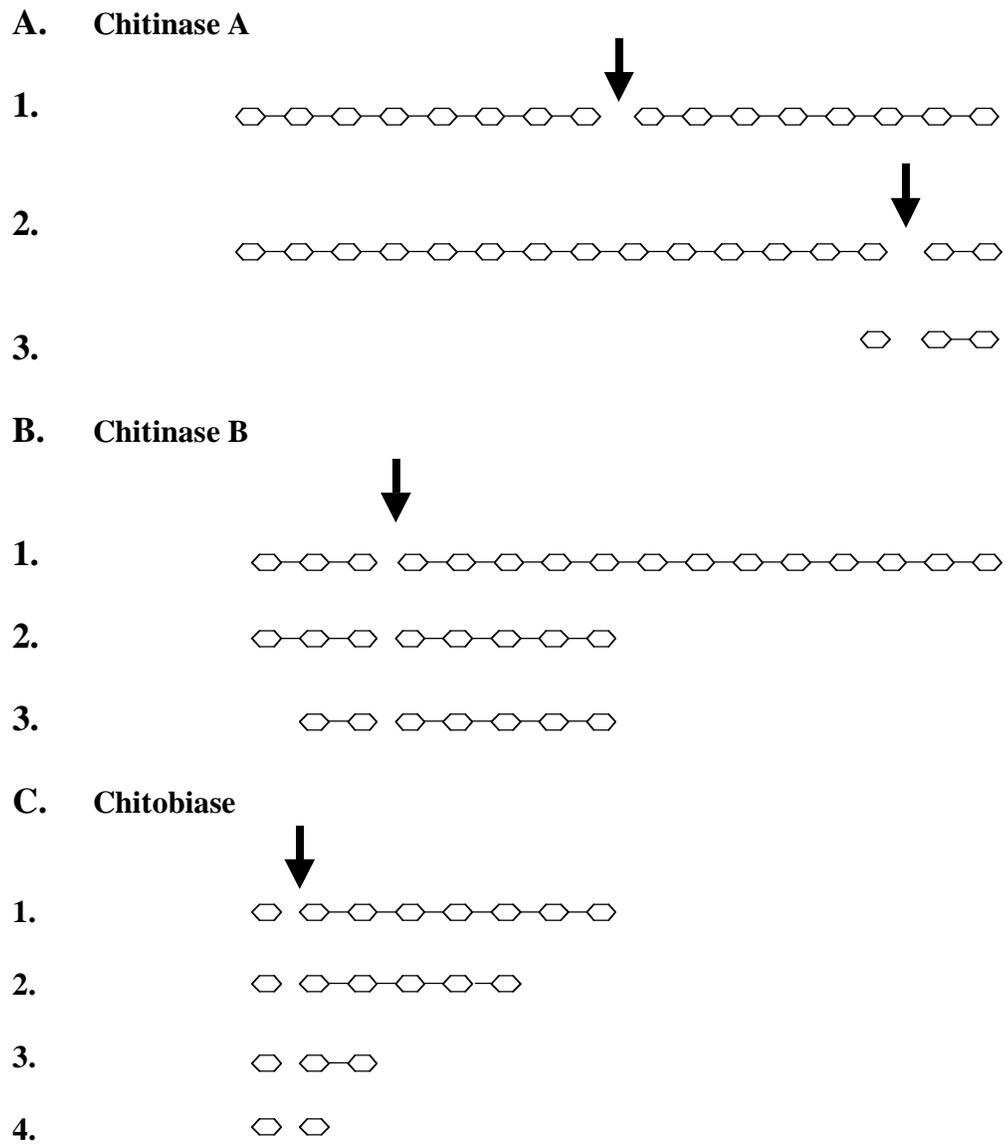
**Figure D.2. Conservation of chitinase A and chitobiase catalytic sites.** The catalytic residues of chitinase A (left) and chitobiase (right) are shown, as well as the overlap of the two enzymes (below). The -1 and +1 sugars are shown. Distances are given in Å. The complex of wild-type chitinase A with the substrate was modeled based on mutant complexes.



**Figure D.3. Structural comparison of the TIM-barrel domains** (A) chitinase A; (B) chitinase B; (C) chitobiase and (D) A superposition of all three enzymes. Alpha-helices are colored yellow and beta-strands are colored green. The catalytic Asp-Glu residues are colored black.



**Figure D.4. Model for the catalytic mechanism of chitinase A and chitobiase.** The model focuses on the catalytic Asp and Glu residues at the active site.



**Figure D.5. The cooperative action of *Serratia marcescens* chitinases and chitobiase in chitin degradation.**

## References

**Alber, T., Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D., Rivers, P. S. & Wilson, I. A.** (1981). On the three-dimensional structure and catalytic mechanism of triose phosphate isomerase. *Philos Trans R Soc Lond B Biol Sci* **293**, 159.

**Altamirano, M. M., Blackburn, J. M., Aguayo, C. & Fersht, A. R.** (2000). Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold. *Nature* **403**, 617.

**Amann, E. & Brosius, J.** (1985). 'ATG vectors' for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* **40**, 183.

**Armand, S., Tomita, H., Heyraud, A., Gey, C., Watanabe, T. & Henrissat, B.** (1994). Stereochemical course of the hydrolysis reaction catalyzed by chitinases A1 and D from *Bacillus circulans* WL-12. *FEBS Lett* **343**, 177.

**Bailey, S.** (1994). The ccp4 suite - programs for protein crystallography. *Acta Crystallographica Section D Biological Crystallography* **50**, 760.

**Boot, R. G., Renkema, G. H., Strijland, A., van Zonneveld, A. J. & Aerts, J. M.** (1995). Cloning of a cDNA encoding chitotriosidase, a human chitinase produced by macrophages. *J Biol Chem* **270**, 26252.

**Boot, R. G., van Achterberg, T. A., van Aken, B. E., Renkema, G. H., Jacobs, M. J., Aerts, J. M. & de Vries, C. J.** (1999). Strong induction of members of the chitinase family of proteins in atherosclerosis: chitotriosidase and human cartilage gp-39 expressed in lesion macrophages. *Arterioscler Thromb Vasc Biol* **19**, 687.

**Bork, P. & Doolittle, R. F.** (1992). Proposed acquisition of an animal protein domain by bacteria. *Proc Natl Acad Sci U S A* **89**, 8990.

**Brameld, K. A. & Goddard III, W. A.** (1998). The role of enzyme distortion in the single displacement mechanism of family 19 chitinases. *Proc. Natl. Acad. Sci. USA* **95**, 4276.

**Brameld, K. A., Shrader, W. D., Imperiali, B. & Goddard, W. A., 3rd.** (1998). Substrate assistance in the mechanism of family 18 chitinases: theoretical studies of potential intermediates and inhibitors. *J Mol Biol* **280**, 913.

**Branden, C. & Tooze, J.** (1991). *Introduction to protein structure*, Garland Publishing, Inc., London.

**Brurberg, M. B., Nes, I. F. & Eijsink, V. G.** (1996). Comparative studies of chitinases A and B from *Serratia marcescens*. *Microbiology* **142**, 1581.

**Burg, J., Banerjee, A. & Sandhoff, K.** (1985). Molecular forms of GM2-activator protein. A study on its biosynthesis in human skin fibroblasts. *Biol Chem Hoppe Seyler* **366**, 887.

**Chothia, C.** (1992). Proteins. One thousand families for the molecular biologist. *Nature* **357**, 543.

**Chu, H. H., Hoang, V., Hofemeister, J. & Schrempf, H.** (2001). A *Bacillus amyloliquefaciens* ChbB protein binds beta- and alpha-chitin and has homologues in related strains. *Microbiology* **147**, 1793.

**Davies, G. & Henrissat, B.** (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853.

**Drouillard, S., Armand, S., Davies, G. J., Vorgias, C. E. & Henrissat, B.** (1997). Serratia marcescens chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. *Biochem J* **328**, 945.

**Fernandes, M. J., Yew, S., Leclerc, D., Henrissat, B., Vorgias, C. E., Gravel, R. A., Hechtman, P. & Kaplan, F.** (1997). Identification of candidate active site residues in lysosomal beta-hexosaminidase A. *J Biol Chem* **272**, 814.

**Folders, J., Tommassen, J., van Loon, L. C. & Bitter, W.** (2000). Identification of a chitin-binding protein secreted by Pseudomonas aeruginosa. *J Bacteriol* **182**, 1257.

**Fromant, M., Blanquet, S. & Plateau, P.** (1995). Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* **224**, 347.

**Gravel, R. A., Clarck, J. T. R., Kaback, M. M., Mahuran, D. D., Sandhoff, K. & Suzuki, K.** (1995). The GM2 gangliosidosis. In *The metabolic and molecular basis of inherited disease* 7 edit. (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds.), pp. 2839-2879. McGraw-Hill, New York.

**Gravel, R. A., Triggs-Raine, B. L. & Mahuran, D. J.** (1991). Biochemistry and genetics of Tay-Sachs disease. *Can J Neurol Sci* **18**, 419.

**Hart, P. J., Pfluger, H. D., Monzingo, A. F., Hollis, T. & Robertus, J. D.** (1995). The refined crystal structure of an endochitinase from Hordeum vulgare L. seeds at 1.8 Å resolution. *J Mol Biol* **248**, 402.

**Hegy, H. & Gerstein, M.** (1999). The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. *J Mol Biol* **288**, 147.

**Hennig, M., Schlesier, B., Dauter, Z., Pfeffer, S., Betzel, C., Hohne, W. E. & Wilson, K. S.** (1992). A TIM barrel protein without enzymatic activity? Crystal-structure of narbonin at 1.8 Å resolution. *Febs Lett* **306**, 80.

**Hennig, M., Jansonius, J. N., Terwisscha-van-Scheltinga, A. C., Dijkstra, B. W. & Schlesier, B.** (1995). Crystal structure of concanavalin B at 1.65 Å resolution. An "inactivated" chitinase from seeds of *Canavalia ensiformis*. *J Mol Biol* **254**, 237.

**Henrissat, B.** (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* **280**, 309.

**Hollis, T., Monzingo, A. F., Bortone, K., Ernst, S., Cox, R. & Robertus, J. D.** (2000). The X-ray structure of a chitinase from the pathogenic fungus *Coccidioides immitis*. *Protein Sci* **9**, 544.

**Huber, R., Stohr, J., Hohenhaus, S., Rachel, R., Burggarf, S., Jannasch, H. W. & Stetter, K. O.** (1995). *Thermococcus chitinophagus* sp. nov., novel chitin-degrading, hyperthermophilic archaeum from a deep-sea hydrothermal vent environment. *Arch Microbiol* **164**, 255.

**Insight II.** (1997). User guide October 1997 edit. Biosym/MSI, Sun Diego.

**Irwin, D., Shin, D. H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A. & Wilson, D. B.** (1998). Roles of the catalytic domain and two cellulose binding

domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. *J Bacteriol* **180**, 1709-1714.

**Jia, J., Huang, W., Schorken, U., Sahm, H., Sprenger, G. A., Lindqvist, Y. & Schneider, G.** (1996). Crystal structure of transaldolase B from *Escherichia coli* suggests a circular permutation of the alpha/beta barrel within the class I aldolase family. *Structure* **4**, 715.

**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.

**Kabsch, W.** (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795.

**Keyhani, N. O., & Roseman, S.** (1997). Wild-type *Escherichia coli* grows on the chitin disaccharide, *N,N'*-diacetylchitobiose, by expressing the *cel* operon. *Proc. Natl. Acad. Sci. USA* **94**, 5.

**Kless, H., Sitrit, Y., Chet, I. & Oppenheim, A. B.** (1989). Cloning of gene coding for chitobiase of *Serratia marcescens*. *Mol. Gen. Genet* **217**, 471.

**Klimov, D. K. & Thirumalai, D.** (2000). Native topology determines force-induced unfolding pathways in globular proteins. *Proc Natl Acad Sci U S A* **97**, 7254.

**Knapp, S., Vocadlo, D., Gao, Z., Krik, B., Lou, J. & Withers, S. G.** (1996). NAG-thiazoline, an N-acetyl-b-hexosaminidase inhibitor that implicates acetamido participation. *J. Am. Chem. Soc.* **118**, 6804.

**Kobayashi, S., Toshitsugu, K. & Shoda, S.-i.** (1996). Synthesis of artificial chitin: irreversible catalytic behavior of a glycosyl hydrolase through a transition state analogue substrate. *J. Am. Chem. Soc.* **118**, 13113.

**Koshland Jr., D. E.** (1953). Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev* **28**, 416.

**Laemmli, U. K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680.

**Lamzin, Vs, W. & Ks.** (1993). Automated refinement of protein models. *Acta Crystallographica Section D Biological Crystallography* **49**, 19.

**Lang, D., Thoma, R., Henn-Sax, M., Sterner, R. & Wilmanns, M.** (2000). Structural evidence for evolution of the beta/alpha barrel scaffold by gene duplication and fusion. *Science* **289**, 1546.

**Laskowski, Ra, M., Mw, M., Ds, T. & Jm.** (1993). Procheck - a program to check the stereochemical quality of protein structures. *Journal Of Applied Crystallography* **26**, 9.

**Malette, B., Paquette, Y., Merlen, Y. & Bleau, G.** (1995). Oviductins possess chitinase- and mucin-like domains: a lead in the search for the biological function of these oviduct-specific ZP-associating glycoproteins. *Mol Reprod Dev* **41**, 384.

**Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G. & James, M. N.** (2001). Crystallographic evidence for substrate-assisted catalysis in a bacterial beta-hexosaminidase. *J Biol Chem* **276**, 10330.

**Mark, B. L., Wasney, G. A., Salo, T. J., Khan, A. R., Cao, Z., Robbins, P. W., James, M. N. & Triggs-Raine, B. L.** (1998). Structural and functional characterization of *Streptomyces plicatus*  $\beta$ -*N*-acetylhexosaminidase by comparative molecular modeling and site-directed mutagenesis. *J Biol Chem* **273**, 19618.

**McCarter, J. D. & Withers, S. G.** (1994). Mechanism of enzymatic glycoside hydrolysis. *Curr. Op. Struct. Biol.* **4**, 885.

**Mechaly, A., Fierobe, H. P., Belaich, A., Belaich, J. P., Lamed, R., Shoham, Y. & Bayer, E. A.** (2001). Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition. *J Biol Chem* **276**, 9883.

**Mechaly, A., Yaron, S., Lamed, R., Fierobe, H. P., Belaich, A., Belaich, J. P., Shoham, Y. & Bayer, E. A.** (2000). Cohesin-dockerin recognition in cellulosome assembly: experiment versus hypothesis. *Proteins* **39**, 170.

**Mesecar, A. D., Stoddard, B. L. & Koshland, D. E., Jr.** (1997). Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences. *Science* **277**, 202.

**Monreal, J. & Reese, E. T.** (1969). The chitinase of *Serratia marcescens*. *Can J Microbiol* **15**, 689.

**Murshudov, G., Vagin, A. & Dodson, E. J.** (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallog. sect. D* **53**, 240.

**Navaza, J.** (1994). Amore - an automated package for molecular replacement. *Acta Crystallographica Section A* **50**, 7.

**Otwinowski, Z. & Minor, W.** (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* **276**, 20.

**Papanikolau, Y., Prag, G., Tavlas, G., Vorgias, C. E., Oppenheim, B. A. & Petratos, K.** (2001). High Resolution Structural Analyses of Mutant Chitinase A Complexes with Substrates Provide New Insight into the Mechanism of Catalysis. *Biochemistry*, **40**, 11338.

**Parker, L. L. & Hall, B. G.** (1990). Mechanisms of activation of the cryptic cel operon of Escherichia coli K12. *Genetics* **124**, 473.

**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.

**Perrakis, A., Tews, I., Wilson, K. S. & Vorgias, C. E.** (1996). Structural aspects on the catalytic mechanism of chitinases, hevamine, and chitobiasis "far away and yet so close?". In *Chitin Enzymology* (**Muzzarelli, R. A. A.**, ed.), Vol. 2, pp. 109-122. European Chitin Society, Ancona.

**Perrakis, A., Wilson, K. S., Oppenheim, A. B., Chet, I. & Vorgias, C. E.** (1993). Phylonetic relationships of chitinases. In *Chitin Enzymology* (**Muzzarelli, R. A. A.**, ed.), Vol. 1, pp. 217-232. European Chitin Society, Ancona.

**Prag, G., Greenberg, S. & Oppenheim, A. B.** (1997). Structural principles of prokaryotic gene regulatory proteins and the evolution of repressors and gene activators. *Mol Microbiol* **26**, 619.

**Prag, G., Papanikolau, Y., Tavlas, G., Vorgias, C. E., Petratos, K. & Oppenheim, A. B.** (2000). Structures of chitobiasis mutants complexed with the

substrate Di-N-acetyl-d-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J Mol Biol* **300**, 611.

**Prag, G., Vorgias, C. E. & Oppenheim, A. B.** (2001). Conservation of structural elements and catalytic mechanism in *Serratia marcescens* chitinolytic enzymes. In *Chitin Enzymology* (**Muzzarelli, R. A. A.**, ed.), Vol. **3**, pp. 351. European Chitin Society, Ancona.

**Rao, V., Guan, C. & Van-Roey, P.** (1995). Crystal structure of endo-beta-N-acetylglucosaminidase H at 1.9 Å resolution: active-site geometry and substrate recognition. *Structure* **3**, 449.

**Read, R. J.** (1986). Improved fourier coefficient for maps using phases from partial structures with errors. *Acta. Crystallog. sect. A* **42**, 144.

**Reid, J. D. & Ogrydziak, D. M.** (1981). Chitinase-overproducing mutant of *Serratia marcescens*. *Appl Environ Microbiol* **41**, 664.

**Rost, B., Sander, C. & Schneider, R.** (1994). PHD-an automatic mail server for protein secondary structure prediction. *Comput Appl Biosci* **10**, 53.

**Rye, C. S. & Withers, S. G.** (2000). Glycosidase mechanisms. *Curr Opin Chem Biol* **4**, 573.

**Sakon, J., Irwin, D., Wilson, D. B. & Karplus, P. A.** (1997). Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. *Nat Struct Biol* **4**, 810-818.

**Sayle, R. A. & Milner-White, E. J.** (1995). Rasmol - biomolecular graphics for all. *Trends In Biochemical Sciences* **20**, 374.

**Schmidt, B. F., Ashizawa, E., Jarnagin, A. S., Lynn, S., Noto, G., Woodhouse, L., Estell, D. A. & Lad, P.** (1994). Identification of two aspartates and a glutamate essential for the activity of endo-beta-N-acetylglucosaminidase H from *Streptomyces plicatus*. *Arch Biochem Biophys* **311**, 350.

**Shapira, R., Ordentlich, A., Chet, I. & Oppenheim, A. B.** (1989). Control of plant diseases by chitinase expressed from cloned DNA in *Escherichia coli*. *Phytopathology* **79**, 1246.

**Sinnot, M. L.** (1990). Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* **90**, 1171.

**Sobolev, V., Sorokine, A., Prilusky, J., Abola, E. E. & Edelman, M.** (1999). Automated analysis of interatomic contacts in proteins. *Bioinformatics* **15**, 327.

**Sun, Y. J., Chang, N. C., Hung, S. I., Chang, A. C., Chou, C. C. & Hsiao, C. D.** (2001). The crystal structure of a novel mammalian lectin, Ym1, suggests a saccharide binding site. *J Biol Chem* **276**, 17507.

**Suzuki, K., Suzuki, M., Taiyoji, M., Nikaidou, N. & Watanabe, T.** (1998). Chitin binding protein (CBP21) in the culture supernatant of *Serratia marcescens* 2170. *Biosci Biotechnol Biochem* **62**, 128.

**Suzuki, K., Taiyoji, M., Sugawara, N., Nikaidou, N., Henrissat, B. & Watanabe, T.** (1999). The third chitinase gene (*chiC*) of *Serratia marcescens* 2170 and the relationship of its product to other bacterial chitinases. *Biochem J* **343**, 587.

**Terwisscha-van-Scheltinga, A. C., Hennig, M. & Dijkstra, B. W.** (1996). 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. *J Mol Biol* **262**, 243.

**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.

**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.

**Tews, I., Dauter, Z., Oppenheim, A. B. & Vorgias, C. E.** (1992). Crystallization of recombinant chitobiase from *Serratia marcescens*. *J Mol Biol* **228**, 696.

**Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S. & Vorgias, C. E.** (1996a). Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. *Nat Struct Biol* **3**, 638.

**Tews, I., Terwisscha-van-Scheltinga, A. C., Perrakis, A., Wilson, K. S. & Dijkstra, B. W.** (1997). Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *J. Am. Chem. Soc.* **119**, 7954.

**Tews, I., Vincentelli, R. & Vorgias, C. E.** (1996b). *N*-Acetylglucosaminidase (chitobiase) from *Serratia marcescens*: gene sequence, and protein production and purification in *Escherichia coli*. *Gene* **170**, 63.

**Thompson, J. D., Higgins, D. G. & Gibson, T. J.** (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through

sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673.

**van Aalten, D. M. F., Synstad, B., Brurberg, M. B., Hough, E., Riise, B. W., Eijsink, V. G. H. & Wierenga, R. K.** (2000). Structure of a two-domain chitotriosidase from *Serratia marcescens* at 1.9-Å resolution. *PNAS* **97**, 5842.

**Van-Roey, P., Rao, V., Plummer, T., Jr. & Tarentino, A. L.** (1994). Crystal structure of endo-beta-*N*-acetylglucosaminidase F1, an alpha/beta-barrel enzyme adapted for a complex substrate. *Biochemistry* **33**, 13989.

**Vorgias, C. E., Kingswell, A. J., Dauter, Z. & Oppenheim, A. B.** (1992). Crystallization of recombinant chitinase from the cloned *chiA* gene of *Serratia marcescens*. *J. Mol. Biol.* **226**, 897.

**Vriend, G.** (1990). WHAT IF: a molecular modeling and drug design program. *Journal Of Molecular Graphics* **8**, 52.

**Warren, R. A. J.** (1996). Microbial hydrolysis of polysaccharides. *Annu. Rev. Microbiol.* **50**, 183.

**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.

**Watanabe, T., Kimura, K., Sumiya, T., Nikaidou, N., Suzuki, K., Suzuki, M., Taiyoji, M., Ferrer, S. & Regue, M.** (1997). Genetic analysis of the chitinase system of *Serratia marcescens* 2170. *J Bacteriol* **179**, 7111.

**Watanabe, T., Oyanagi, W., Suzuki, K., Ohnishi, K. & Tanaka, H.** (1992). Structure of the gene encoding chitinase D of *Bacillus circulans* WL-12 and possible homology of the enzyme to other prokaryotic chitinases and class III plant chitinases. *J Bacteriol* **174**, 408.

**Wierenga, R. K.** (2001). The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Lett* **492**, 193.

**Wisconsin Package Version 10, (GCG)** Madison, Wisc. .

**Wright, C. S., Li, S. C. & Rastinejad, F.** (2000). Crystal structure of human GM2-activator protein with a novel beta-cup topology. *J Mol Biol* **304**, 411.

**Young, E., Chatterton, C., Vellodi, A. & Winchester, B.** (1997). Plasma chitotriosidase activity in Gaucher disease patients who have been treated either by bone marrow transplantation or by enzyme replacement therapy with alglucerase. *J Inherit Metab Dis* **20**, 595.

**Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G. & Court, D. L.** (2000). An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* **97**, 5978.

**Zhang, B., Xu, G. & Evans, J. S.** (1999). A kinetic molecular model of the reversible unfolding and refolding of titin under force extension. *Biophys J* **77**, 1306.

**Zhang, J. H., Dawes, G. & Stemmer, W. P.** (1997). Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci USA* **94**, 4504.

## Appendix A1 Multiple alignment of the glycosyl hydrolase family 20 proposed TIM-barrel domain

	β-1	α-1	β-2	
CHB_SERMA	DAPRF <b>YRGIFLD</b> DVARNFHK	<b>KDAVLRLLDQMAA</b> YKLN <b>KHFHLS</b>	DDEGWRIEIPGLPELT	393
HEXA_HUMAN	DFPRFPHRGLLLDTSRHYLPLSSILD	TLTLDVMAYNKLN	VFWHLVDDPSFPYESFTFPELM	222
HEXA_MOUSE	DFPRFPHRGVLLDTSRHYLPLSSILD	TLTLDVMAYNKFN	VFWHLVDDSSFPYESFTFPELT	222
HEXB_HUMAN	DSPRFSHRGILIDTSRHYLPVKIILK	TLDAMAFNKFNVL	HWHIVDDQSFPYQSITFPELS	255
HEXB_MOUSE	DSPRFPHRGILIDTSRHFLPVKTI	LKTLDAMAFNKFNVL	HWHIVDDQSFPYQSTTFPELS	234
HEXB_FELCA	DSPRFPHRGILIDTARHFLPVKSIL	KTLDAMAFNKFNVL	HWHIVDDQSFPYQSVTFPELS	197
HEXB_PIG	DFPRFPHRGILIDTGRHF	LSVKTIKTLDAMAFNKFNVL	HWHIVDDQSFPYQSINFGVLS	229
CHB_VIBHA	DAPRFDYRGVMVDVARNFH	SKDAILATLDQMAAYKM	NKLHLHLTDDEGWRIEIPGLPELT	387
HEX_VIBVU	DEPRLDYRGMHMDVSRNFH	SKELVFRFLDQMAAYKM	NKFHFLADDEGWRIEINGLPELT	370
HEXB_ALTSO	DTPRYDFRGLHVDVARNF	RSKAFILQTI EQMAAYKLN	NKLHLHLADDEGWRLAIDGLDEL	382
CHB_STRPL	DTPRYAWRSAMLDVSRH	FFGVDEVKRYIDRVARY	KYNKLHLHLSDDQGWRIAIDSWPRLA	206
HEXA_PORGI	DEPAFGYRGFMLDVCRH	FLSVEDIKKHIDIMAM	FKINRFHWHLTEDQAWRIEIKKYPRLT	225
HEX1_ENTHI	DAPRFKWRGLMVDPSRN	PLSPLMFKRIIDTLAS	VKANVLHIHLSDAQTFVFESKKYPLLH	196
HEX1_VIBFU	DAPRFKYRGMMLDCAR	HFHPLERVKRLINQLA	HYKFNTFWHLTDDEGWRIEIKSLPQLT	315
HEXA_DICDI	DSPRYPWRGFMVDSAR	HYIPKNMILHMIDSL	GFSKFNTHLWHMVDAAVFPVESTTYPDLT	210
HEXC_BOMMO	DRPVYPYRGILLDTARN	FYSIDSIKRTIDAMA	AVKLNTHFWHITDSQSFPLVLQKRPNLS	266
	* * * . : * * :	. : : . * * : * * :	: :	*

		α-2	β-3	
CHB_SERMA	EVGGQRCHDLSETTCLLPQY	QGP--DVYG-GFFS	<b>RQDYIDI IKYAQA</b> RQI <b>EV</b> IP <b>EIDM</b> P	450
HEXA_HUMAN	RKGSYNP-----	VTHIYTAQDVKEVIEYAR	LRGIRVLAEFDTP	260
HEXA_MOUSE	RKGSFNP-----	VTHIYTAQDVKEVIEYAR	LRGIRVLAEFDTP	260
HEXB_HUMAN	NKGSYS-----	LSHVYTPNDVRMVIEYAR	LRGIRVLP EFDTP	292
HEXB_MOUSE	NKGSYS-----	LSHVYTPNDVRMVLE	YARLRGIRVIPEFDTP	271
HEXB_FELCA	NKGSYS-----	LSHVYTPNDVHTVIE	YARLRGIRVIPEFDSP	234
HEXB_PIG	SKGSYS-----	LSHVYTPNDVRMVIE	YARIRGIRVMPEFDTP	266
CHB_VIBHA	EVGANRCFDTQEK	SCLLPQLGSGPTTDN	FGSGYFSKADYVEILKYAKARNIEV	447
HEX_VIBVU	QVGAHRCHDVEQNK	CMPQLGSGAELP	NNSGYYTREDYKEILAYASARNIQV	430
HEXB_ALTSO	SVGAYRCFDLTET	RCLLPQLGAGNDK	NAQVNGFYSAEDYIEILRYAKAHHIE	442
CHB_STRPL	TYGG-----S-----	TEVGGGP-----	GGYYTKAEYKEIVRYAASRHLEVVPE	248
HEXA_PORGI	EVGS-----T-----	RTEGDGTQ----	YSGFYTQEQVRDIVQYASDHFITV	269
HEX1_ENTHI	QKGMVD-----	ESFVLTQSFLRELA	QYGANRGVIVYGEIDTP	233
HEX1_VIBFU	DIGAWRG-----	VDEVLEPQYSLLTE	KHGGFYTQEEIREVIAAYAAERGIT	368
HEXA_DICDI	-KGAFS-----	PSATFSHDDIQEV	VAYAKTYGIRVIPEFDIP	246
HEXC_BOMMO	KLGAYS-----	PTKVYTKQDIREV	VEYGLERGVRLPEFDAP	303
	*	:	: * . : * : :	*

CHB_SERMA	AHARAAVVSMEARYK	KLHAAGKEQEANE	FRLVDPTDTSNTTSVQFFNRQS---	YLNPCLD	507
HEXA_HUMAN	GHTLSWGPG-----	IPG-----	LLTPCYSGSEPSGTFG-----	PVNPSLN	295
HEXA_MOUSE	GHTLSWGPG-----	APG-----	LLTPCYSGSHLSGTFG-----	PVNPSLN	295
HEXB_HUMAN	GHTLSWGKG-----	QKD-----	LLTPCYSRQNKLDSFG-----	PINPTLN	327
HEXB_MOUSE	GHTQSWGKG-----	QKN-----	LLTPCYNQKTKTQVFG-----	PVDPTVN	306
HEXB_FELCA	GHTQSWGKG-----	QKD-----	LLTPCYNEHKQSGTFG-----	PINPILN	269
HEXB_PIG	GHSRSWGKG-----	QKD-----	LLTPCYRKQVLSGTFG-----	PINPILN	301
CHB_VIBHA	AHARAAVVSMEARY	DRLMEEGKEAEANE	YRLMDPQDTSNVTTVQFYNKQS---	FINPCME	504
HEX_VIBVU	GHSLAAVKSM	EARYRKFMAEGDVV	KAEMYLLSDPNDDTQYYSIQHYQDN---	TINPCME	486
HEXB_ALTSO	GHSRAAI	IAMEARYKKLMA	QKPEDAQKYRLVETADKTRYSSI	QHYNDN---	498
CHB_STRPL	GHTNAALAS----	YAE LNCDG-----	VAPPLYTGTKVGFSS-----	SLCVDKD	286
HEXA_PORGI	GHAMAALAA-----	YP-----	QFRCFPPREFKP--RIIWGVEQD---	VYCAGKD	307
HEX1_ENTHI	AHTASWNLG-----	YPG-----	VVANCWYIVVSTSMRYGEN--	VLSLNPANP	273
HEX1_VIBFU	GHSRAAI	KA-----	LP-----	EWLFD	407
HEXA_DICDI	GHAAAWGIG-----	YPE-----	LVATCPDYAANVNNI-----	PLDISNP	280
HEXC_BOMMO	AHVGEGWQD-----	T-----	GLTVCFKAEPWTKFCVEPPCG--	QLNPTKE	341
	.	*			*

	$\alpha$ -3	$\beta$ -4	
CHB_SERMA	<u>SSQRFVDKVI</u> GEIAQMHKEA	GQPIKT <u>WHFCG</u>	GDEAKNIRLGAGYTDKAKPEPG--KGIIDQ 565
HEXA_HUMAN	NTYEFMSTFFLEVSSVFPDF	-----YLHLGGDEVDF	----- 327
HEXA_MOUSE	STYDFMSTLFLEISSVFPDF	-----YLHLGGDEVDF	----- 327
HEXB_HUMAN	TTYDFMSTLTFKEISEVFPDQ	-----FIHLGGDEVDF	----- 359
HEXB_MOUSE	TTYAFFNTTFKEISSVFPDQ	-----FIHLGGDEVDF	----- 338
HEXB_FELCA	STYNFLSQFFKEVSMVFPDH	-----FVHLGGDEVDF	----- 301
HEXB_PIG	TTYNFLSKFFKEISTVFPDE	-----FIHIGGDEVDF	----- 333
CHB_VIBHA	SSTRFVDKVI	SEVAAMHQEAGAPLTTWHFGGDEAKNIKLGAGFQDVNAEDKVS	SWKGTIDL 564
HEX_VIBVU	SSFVFM	KVIDEINKLHKEGGQPLTDYHIGADETA	-----G----- 522
HEXB_ALTSO	NTYTFIDKVLSEVKVLH	DRAGVPLNTYHIGADETA	----- 533
CHB_STRPL	VTYDFVDDVIGELAALTP	G---RYLHIGGDEAHS	----- 317
HEXA_PORGI	SVFRFISDVIDEVAPLFP	G---TYFHIGGDECPKD	----- 339
HEX1_ENTHI	NTFPIIDALMKELSDTF	GTD---YVHVGDEVWTS	----- 305
HEX1_VIBFU	GTFRFLDCVLEEVAALF	PSH---FIHIGADEVPDG	----- 439
HEXA_DICDI	ATFTFIQNLFTETIAPL	FIDN---YFHTGGDELVTG	----- 312
HEXC_BOMMO	ELYDYLEDIYVEMAEAF	EST---DMFHMGGDEV	SER----- 374
	. . * :	* * . **	

	$\alpha$ -4	$\beta$ -5	
CHB_SERMA	SNEDKPWAKSQ--VCQTM	IK--EGKV--AD	<u>MEHLPSYFGQEVSKLVKA</u> HIGID-- <u>RMQAW</u> 616
HEXA_HUMAN	-----CWKSNP---EIQDF	MR--KKGFG-EDFKQLESFYI	QTLLEDIVSSYGK---GYV
HEXA_MOUSE	-----CWKSNP---NIQAF	MK--KKGFG--TDFKQLESFYI	QTLLEDIVSDYDK---GYV
HEXB_HUMAN	-----CWESNP---KIQDF	MR--QKGGF-TDFKKLESFYI	QKVLDIATINK---GSIV
HEXB_MOUSE	-----CWASNP---NIQGF	MK--RKGFG-SDFRRLESFYI	KKILEIISLKK---NSIV
HEXB_FELCA	-----CWESNP---EIQGF	MK--QKGGF-KDFRRLESFY	LQKLLGIVSTVKK---GSIV
HEXB_PIG	-----CWASNS---EILQF	MQ--EKGFS-QISLNSNLCT	VFKISNMISAMKK---RPIV
CHB_VIBHA	SKQDKPFAQSP---QCQTL	LIT--DGTV--SDFAHLP	SHFAEEVSKIVA
HEX_VIBVU	-----AWGDSP---ECRKM	FVAPESGV--KNAKDINGY	FINRISHILDAKGL---TLG
HEXB_ALTSO	----VLWLESP---ACKKL	QAS---V--KDFTNFNGY	FIERVAKLLDKKGI---QV
CHB_STRPL	-----TPKADFVAFM	KRVQP-IVAKY	GK---TVV
HEXA_PORGI	-----RWKACS---LCQKR	MR--DNGL--KDEHELQSY	FIKQAEKVLQKHGK---RLI
HEX1_ENTHI	-----GWSKSKEYS	DIQKFMK--SKGL--NSL	TELEGYFNKYAQEQVIHNGK---HPV
HEX1_VIBFU	-----VWVNSP---KCQAL	MA--EEGY--TDAKELQ	GHLRLRYAEKKLKSLGK---RMV
HEXA_DICDI	-----CWLEDP---AIAN	WMT--KMGF---STTDAFQ	YFENNLDVTMKSINR---TKIT
HEXC_BOMMO	-----CWNSSE---EIQN	FMIQNRWNL	DKSSFLKLWNYFQKNAQDRAYKAF
			* :

	$\alpha$ -5	$\beta$ -6	$\alpha$ -6	$\beta$ -7			
CHB_SERMA	QDGLKDAES	<u>SKA</u> F	----ATSRVGVN	FWD	TLYWG	<u>GFDSVNDWANK</u> GYEVVV	SNPD-YVYMD 671
HEXA_HUMAN	QEVFDNK-VKI	QP-----DTIIQV	WREDIPV-NYMKELE	LVTKAGFR	ALLSAP---WYLN 423		
HEXA_MOUSE	QEVFDNK-VK	VRP-----DTIIQV	WREEMPV-EYMLEM	QDITRAGFR	ALLSAP---WYLN 422		
HEXB_HUMAN	QEVFDDK-AKL	AP-----GTIVEV	WKDSA----YPEEL	SRVTASGFP	VILSAP---WYLD 452		
HEXB_MOUSE	QEVFDDK-VEL	QP-----GTVVEV	WKSEH----YSYEL	KQVTGSGFP	PILSAP---WYLD 431		
HEXB_FELCA	QEVFDDH-VK	LLP-----GTIVQV	WKNQV----YTEEL	REVTAAGFP	VILSAP---WYLD 394		
HEXB_PIG	QEAFDGR-DK	FMP-----GTVVQV	WKIED----YKWEQ	SLITKAGFP	VILSAP---WYLD 426		
CHB_VIBHA	QDGLKYS	DGEKAF----ATENTR	VNFWDVLYWGGT	SSVYEW	SKKGYDVI	VSNPD-YVYMD 670	
HEX_VIBVU	NDGLSHKAL	DASSL---AGNPP	KAWVWGTMF	WGGVDQYNS	FANKGYD	VVVT	
HEXB_ALTSO	SDGLGDV	RAANM-----PANI	QSNGLGDI	KRKR	APVAHR	FANQGWV	
CHB_STRPL	HQLAGAEP	VEGA-----LVQY	WGLDRTGDAE	KA	EA	EAARNGTGLIL	
HEXA_PORGI	DEILEGGL	APSA-----TVMS	W---RG---ED	-GGIAA	ANM	NHDVIMTPG	
HEX1_ENTHI	EEVF	KKGNADKN-----TIIQV	WDDIR-----LLQ	QV	VNSGYKA	IF	
HEX1_VIBFU	EEAQHG	DKVSKD-----TVI	YSW---LS---EQ	-AALN	CARQ	GFDVILQ	
HEXA_DICDI	NDPIDYG-V	QLNP-----ETL	VQVSSGS-----DLQ	GIVNSGY	KALV	SFA---WYLD 400	
HEXC_BOMMO	TSTLT	TDY-THVEK	FLDK	DEYIIQV	WTTGADP-----QIQ	GLLQKGYRLIMS	
						* : :	

CHB_SERMA	FPYE-----VNPDERGYYWGTRFSDERKVFVSPAPDNMPQNAETSVDRDGNHFNA	720
HEXA_HUMAN	RIS-----YGPDWKDFYVVEPLAFEGTPEQKAL-----	451
HEXA_MOUSE	RVK-----YGPDWKDMYKVEPLAFHGTPEQKAL-----	450
HEXB_HUMAN	LIS-----YGQDWRKYYKVEPLDFGGTQKQKQL-----	480
HEXB_MOUSE	LIS-----YGQDWKNYYKVEPLNFEGSEKQKQL-----	459
HEXB_FELCA	WIS-----YGQDWRNYYKVDPLHFDGSEQKQL-----	422
HEXB_PIG	LIS-----YGQDWKNYYEVEPQDFPGSDKERKR-----	454
CHB_VIBHA	MPYE-----VDPKERGYYWATRATDTRKMFGFAPENMPQNAETSVDRDGNNGFTG	719
HEX_VIBVU	MPYE-----NDPEERGYYWATRFNDTKKVFVSMFPENVPANVEWMTDRMGAKISA	674
HEXB_ALTSO	FPYQ-----SHPEERGNHWASRAIESKKMFEFMPDNLPAHAEIWKNTNNHAYIA	679
CHB_STRPL	MKY-----TKDTPGLGSLWAG-YVEVQRSYDWDPA-----YLPG	428
HEXA_PORGI	HYQ-----GDPTVEPVVAIGG-YAPLEQVYAYNPLPK-----ELPA	463
HEX1_ENTHI	KQMPCLNSYDSSTCVNTHSMWVWTRNDMYDNDPVKSLSSSEKEN-----	440
HEX1_VIBFU	IAQD-----YAPEEPGVWDAG-VTPLERAYRYEPLVE-----VPE	563
HEXA_DICDI	KQNPDN-----NIHYEQDQTFYAADPTNNISTN-AEN-----	434
HEXC_BOMMO	CGFG-----AWVSGNWNWCSPYIGGQKVYGNPSPAVMALS-----	513

	$\beta$ -8	$\alpha$ -8	
CHB_SERMA	KSDK----PWPGAYGLSAQLWSETQRTDPQMEYMIFPRALSVAERSWIRAGWEQDYRAGR		776
HEXA_HUMAN	-----VIGGEACMWGEYVDNTN-LVPRLWPRAGAVAERLWSNKLTSDLTFAYE		498
HEXA_MOUSE	-----VIGGEACMWGEYVDSTN-LVPRLWPRAGAVAERLWSSNLTTNIDFAFK		497
HEXB_HUMAN	-----FIGGEACLWGEYVDATN-LTPRLWPRASAVGERLWSSKDVRDMDDAYD		527
HEXB_MOUSE	-----VIGGEACLWGEFVDATN-LTPRLWPRASAVGERLWSPKTVTDLENAYK		506
HEXB_FELCA	-----VIGGEACLWGEFVDATN-LTPRLWPRASAVGERLWSPEDITSVGNAYN		469
HEXB_PIG	-----VLGGEACLWGEYVDATN-LTPRLWPRASAVGERLWSHKDVRDIHDAYS		501
CHB_VIBHA	KGEI----EAKPFYGLSAQLWSETVRNDEQY EYMVFPRLVAAAQRAWHRADWENDYKGVV		775
HEX_VIBVU	TTGE----KTHDFLGVQALWSETIRTDQVEYMVLPRLMIAVAERGWHKASWEEEHKEGI		730
HEXB_ALTSO	NDSDSLNGVQFAGLQGLWSEMLRSDAQAEYMLYPRLLALAERAWHHAEWELPYQAGR		739
CHB_STRPL	APAD-----AVRGVEAPLWTETLSDPDQLDYMAFPRLPGVAELGWSPASTHDWDTYKV		481
HEXA_PORGI	DKH----R--YVLGAQANLWAEYLYTSERYDYQAYPRLLAVALTWTPLAKKDFADFCR		516
HEX1_ENTHI	-----VLGGEGCSWGESTDEQN-FFDRVFRQRYSAIAERLWSKESVVDKESHEV		487
HEX1_VIBFU	HDP----LR-KRILGIQCALWCELVNNQDRMDYMIYPRLLTALAGSGLDTPKIPA-----		611
HEXA_DICDI	-----IIGGEATMWAEQINQVN-WDVRVWPRAIGIAERLWSAQSVNSVSLALP		481
HEXC_BOMMO	YR-----DQILGGEVALWSEQSDPAT-LDGRLWPRAAFAERMWAEPSTAWQDAEHR		564
	* . * *	* . .	

## **Appendix A2. List of papers influenced by this thesis**

**The following three manuscripts are a direct outcome of this thesis**

**Prag, G., Papanikolau, Y., Tavlas G., Vorgias, E.C., Petratos, K. and Oppenheim, B.A.** Structures of chitobiase mutants with the substrate Di-*N*-acetyl-Dglucosamine: The catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J. Mol. Biol.* (2000) **300**, 611. Attached as a PDF file ([chitobiase.pdf](#)).

**Papanikolau, Y., Prag, G., Tavlas G., Vorgias, E.C., Oppenheim, B.A. and Petratos, K.** High resolution structural analyses of mutant chitinase A complexes with substrates provide new insight into the mechanism of catalysis. *Biochemistry* (2001) **40**, 11338. Attached as a PDF file ([chitinaseA.pdf](#)).

**Prag, G., Vorgias, E.C. and Oppenheim, B.A.** Conservation of structural elements and catalytic mechanism in *Serratia marcescens* chitinolytic enzymes. *Chitin Enzymology* (2001) **3**, 351. Not attached.

**The following two manuscripts are not related to the study of structure-function analysis of TIM-barrel proteins.**

**Prag, G., Greenberg, S. and Oppenheim, B.A.** Structural principles of prokaryotic gene regulatory proteins and the evolution of repressors and gene activators. *Mol. Microbiol.* (1997) **26**, 619. Attached as a PDF file ([StructPric.pdf](#)).

**Giladi, H., Koby, S., Prag, G., Engelhorn, M., Geiselman, J. and Oppenheim, B.A.** Participation of IHF and a distant UP element in the stimulation of the phage  $\lambda$  PL promoter. *Mol. Microbiol.* (1998) **30**, 433. Attached as a PDF file ([ihf.pdf](#)).

### **Appendix A3. CD-ROM and Internet links**

*i.* A CD-ROM that contains the Ph.D. thesis. In addition it contains the structures (see below for a list) that allow interactive inspection of animation and 3D structures.

The animation is a model based on the structures of the wild type enzyme and mutant complexes with substrates, and shows several steps along the enzymatic reaction. The CD-ROM also contains the published coordinates of chitinase A and chitobiase complexes.

*ii.* Viewing structures and animation from the web.

To view, download the *Chime* program from the *Chime home page* ( <http://www.mdlchime.com/chime/> ) and open one of the specific locations:

**Structure of chitinase A D313A – octaNAG complex :**

<http://md.huji.ac.il/~prag/thesis/1eib.html>

**Structure of chitinase A E315Q – octaNAG complex :**

<http://md.huji.ac.il/~prag/thesis/1ehn.html>

**Structure of chitobiase D539A – diNAG complex :**

<http://md.huji.ac.il/~prag/thesis/1c7s.html>

**Structure of chitinase A E540D – diNAG complex :**

<http://md.huji.ac.il/~prag/thesis/1c7t.html>

**Animation of the catalytic mechanism of chitinase A:**

<http://md.huji.ac.il/~prag/thesis/mechanism.html>

יחסי מבנה תפקיד בחלבונים דמויי חבית טים:  
המכניזם הקטליטי של אנזימים כיתינוליטיים

חיבור לשם קבלת תואר דוקטור לפילוסופיה

מאת גלי פרג

עבודה זו נעשתה בהדרכתו של

פרופסור עמוס אופנהיים

הוגש לסינט האוניברסיטה העברית בשנת 2001

## תוכו עניינים

6	תקציר
11	רשימת קיצורים
13	מבוא
	חומרים ושיטות
26	רשימת טבלאות של אליגונוקלאוטידים דנ " א, זני חיידקים ופלסמידים
30	בניית מוטנטים וביטוי חלבוני כיתינאז A וכיתוביאז
31	סריקה ואיבחון מוטנטים עם פעילות נמוכה של כיתינאז A
32	שיבוט ונוקאאוט של הגן <i>celF / chbF</i> מחיידק <i>Escherichia coli</i>
33	מוטגנזה אקראית ויצירת ספריות קומבנטוריות של דנ " א
34	קביעת רצף דנ " א ותיקון שגיאות
35	הפקה וניקוי החלבונים כיתוביאז, כיתינאז A ו ChbF
36	גיבוש מוטנטים של כיתוביאז וכיתינאז A
37	גאנליזה של הקנטיה הביוכימית של חלבונים מוטנטים וזן הבר מכיתוביאז וכיתינאז A
38	איסוף תוצאות פיזור קרינת X של גבישי חלבון ממאיץ חלקיקים בתנאי קור
39	החלפה מולקולרית ועידון המבנה
40	חיפוש וסידור מתואם של רצפים הומולוגים
40	בניית מודל של האנזים הקסוזאמינידאז מאדם על סמך המבנה של כיתוביאז
	תוצאות
47	<b>פרק א</b> אנליזה גנטית של תת מבנה דמוי טים של כיתוביאז ושימור המבנה הקטליטי.
	<b>פרק ב</b> מבנה מרוכב של מוטנטים בכיתוביאז עם סבסטרט טבעי: התפקיד הקטליטי של חומצות
57	אמינו השמורות אספרטאט 539 וגלוטמאט 540.
68	<b>פרק ג</b> קשירת הסובסטרט oligoNAG על ידי כיתינאז

77	<b>פרק ד</b> אנליזה של המכניזם הקטליטי של האנזים כיתנאז
48	<b>פרק ה</b> קיבוע ותפיסת מצבי בינים בראקציה אנזימטית בעזרת המוטנט D391A של כיתנאז
95	<b>דיון</b>
109	<b>רשימת ספרות</b>
122	<b>נספחים</b>
	<b>תקציר בעברית</b>

## תקציר

האנזימים המפרקים כיתין נפוצים ביותר ונמצאים באורגניזמים פרוקריוטיים, אאוקריוטיים וארכאה. החיידק *Serratia marcescens* משמש כמערכת מודל לחקר ניצול הכיתין כמקור פחמן. בנוכחות כיתין החיידק מבטא את האנזימים הכימיקליים כיתינאז A, כיתינאז B, כיתינאז C, ו-כיתוביאז המפרקים את הכיתין ליחידות N-acetylglucosamine (NAG). המבנה הגבישי של האנזימים כיתינאז A, כיתינאז B ו-כיתוביאז פוענח. בשלושת האנזימים תת-המבנה (domain) הקטליטי הינו דמוי חבית (טים). כיתינאז A, כיתינאז B וכיתינאז C שייכים למשפחת האנזימים ההידרוליטיים מספר 18, וכיתוביאז שייך למשפחה 20. מעניין כי כיתוביאז דומה לאנזים הקסוזאמינידאז באדם

(human hexosaminidase A - HexA). מוטציות המשבשות את פעילות האנזים באדם גורמות למחלת הטיי-זאקס (Tay-Sachs). בהתבסס על תת המבנה הקטליטי-חבית הטים של כיתוביאז, בנינו מודל של האנזים האנושי HexA. מודל זה מאפשר לנו להסביר מבנית את חומרת חלק מהמוטציות הגורמות למחלה. החדרה מוטציות גורמות המחלה בבני אדם, לגן המקודד לכיתוביאז איפשרו לנו לחקור ולתקן את המודל המבני.

האתר הפעיל של שני האנזימים, כיתינאז A וכיתוביאז בנוי מלולאות הנמצאות בקצה הקרבוקסילי של גדילי ה- $\beta$  של חבית הטים. לפיכך אנו מציעים כי לחלבונים אלו מקור אבולוציוני משותף. מעט מאוד ידוע על המבנה המרוכב (complex) של כיתינאז עם הסובסטרט הטבעי שלו. מאמצים לפענח מבנים כאלו נכשלו כנראה בגלל פירוק הסובסטרט על ידי האנזים. בעבודה זו החדרתי מוטציות המשנות את חומצות האמינו הקטליטיות. אפיון ביוכימי של חלבונים מוטנטים אלו איפשר לי בהמשך לבחור מוטנטים המתאימים למחקר מיבני. בעזרת המוטנטים הנבחרים הצלחתי לגדל גבישים מרוכבים של האנזים עם הסובסטרט הטבעי שלו. באמצעות קרינת X פיענחתי את המבנים המרוכבים של מוטנטים אלו. בהתבסס על

מחקר זה, אנו מציעים כי קיים דמיון רב במנגנון הפעולה של שני האנזימים, כיתינאז A וכיתוביאז.

השוואת רצפי חלבונים ממשפחות 18 ו-20, מראה שימור (חותמת) של שיירים ב גדיל  $\beta$ 4 ולולאה מספר-4, (DXXDXDXE במשפחה 18, ו- HXGGDE במשפחה 20). הוצע שבשתי חותמות אלו, תפקידה של החומצ האמינית גלוטמאט, (Glu315 של כיתינאז A, ו Glu540 של כיתוביאז) הינו למסור פרוטון במהלך הקטליזה. מצאנו שהמיקום המרחבי של זוג השיירים הקטליטיים DXE של כיתינאז A ו- DE של כיתוביאז דומים מאוד. אך בכיתינאז A, ח' אספארטית 313, (Asp313) נמצא במבנה החלבון הטבעי בשתי קונפורמציות אלטרנאטיביות. על ידי השוואת המבנים המרוכבים של כיטינאז A וכיתוביאז מצאנו דמיון רב במיקום הסוכרים -1 NAG ו +1 NAG ביחס לאתר הפעיל. כמו כן ראינו כי בשני המבנים המרוכבים, מישורי הסוכרים -1 NAG ו +1 NAG מסובבים בכ-  $90^\circ$  סביב הקשר שעתיד להיחתך. יתרה מזו, חל שינוי קונפורמציה בסוכר שבעמדה 1- ממבנה דמוי כיסא למבנה דמוי סירה. מבנה דמוי הסירה הוא בעל אנרגיה גבוהה שאינה מועדפת ולפיכך מקדם את התהליך ההידרוליטי של חיתוך הסוכר.

השוואת מבני האתר הפעיל של כיתינאז A וכיתוביאז מספקת הסבר מבני להבדלי התיפקוד של שני החלבונים. בכיתינאז A מבנה תעלה מאורך מאפשר קישור איתן של סובסטרט בעל שמונה סוכרים (octaNAG) המשובץ בעומק התעלה. תעלה זו מאפשרת את פעילותו העיקרית של האנזים כאקסוכיטינאז (exochitinase) המסיר דו-סוכר מהקצה המחוזר של שרשרת הכיטין וכן גם את פעולתו כאנדוכיטינאז (endochitinase) החותך את שרשרת הכיטין באופן אקראי לשתי חתיכות גדולות יחסית. לעומת זאת, האתר הפעיל של כיתוביאז בנוי כמנהרה היכולה להכיל כארבעה סוכרים (tetraNAG). אנזים זה פועל כאקסוכיטינאז המסיר

יחידת סוכר אחת מהקצה הבלתי מחוזר של שרשרת כיתין קצרה. מצאנו שתת המבנה מספר-I, (domain-I) של כיתוביאז משתתף אף הוא בקשירת הסוכר tetraNAG. תת מבנה זה לא נמצא באנזים HexA באדם.

המנגנון הקטליטי של רוב האנזימים ההידרוליטיים מפרקי הסוכרים (גליקוזידאזות), נקרא קטליזה של חומצה-בסיס (acid-base catalysis). במנגנון זה משתתפים שני שיירים של חומצות אמינו. האחד, בדרך כלל גלוטמאט, מתפקד כמוסר פרוטון (proton donor), השני, בדרך כלל אספרטאט, מתפקד כנוקליאופיל (nucleophile). בהידרוליזה של סוכרים הקונפיגורציה האנומרית של אטום הפחמן לאחר החיתוך האנזימאטי יכולה להיוותר בשני מצבים, האחד זהה לזה שהיה לפני החיתוך (retention), ובשניה, חל היפוך של הקונפיגורציה האנומרית (inversion). בשעה שהשייר הגלוטמאט - מוסר הפרוטון שמור מאוד, לא נמצא שייר המתאים לפעול כנוקליאופיל בשתי משפחות האנזימים 18 ו-20. לפיכך הוצע מודל הידרוליטי שונה, בו הגלוטמאט פועל כמוסר פרוטון אך הנוקליאופיל הינו הסובסטרט עצמו (substrate-assisted). אטום חמצן O7, הקשור לקבוצת האצטאמידו (קרבוניל) בסוכר NAG-1, מתקיף את פחמן C1 במצב הביניים של החיתוך (oxacarbenium). כתוצאה מהתקפה נוקלאופילית זו נוצר מצב מעבר שבו מבנה טבעת אוקסאזוליניום (oxazolinium ring) מיוצב זמנית. על פי תוצאות המחקר אנו מציעים שקבוצת האצטאמידו מסתובבת סביב הקשר C2-N2 על מנת שאטום החמצן O7 יימצא בעמדה מתאימה להתקפה נוקלאופילית. עדות ישירה למנגנון זה (לפיה נוצר קשר קו-וולנטי בין O7 ו-C1 טרם נצפתה).

אנליזה של מבני המוטנטים המרוכבים של שני האנזימים (D313A בכיתנאז A ו D539A בכיתוביאז), בהם הוחלף שייר האספארטאט השמור באלאנין, מספקת הסבר מעמיק על תפקידו של האספארטאט בקטליזה. על סמך התוצאות אנו מציעים שאספארטאט זה מבטיח

את מיקומו המדוייק של הנוקלאופיל. כמון כן אנו מציעים שהאספרטאט מספק מטען שלילי נוסף באתר הפעיל המייצב את המטען החלקי החיובי המתחלק על טבעת האוקסאזוליניום. בנוסף אנו מציעים ששייר האספרטאט נע בין שני מצבים מבניים חילופיים (קונפורמציות). קונפורמציה אחת מאפשרת את קשירת הסובסטרטאט בייתר קלות והקונפורמציה החלופית "נועלת" את קבוצת האצטאמידו ב- NAG-1 באורינטציה שאינה יציבה מבחינה אנרגטית התומכת בקטליזה מתווכת על ידי סובסטרטאט (substrate-assisted).

ניסיתי לתפוס ולפתור את המבנים המרוכבים של מצבי המעבר בראקציה (intermediate). לשם כך אפיינתי מוטנט D391A של כיתינאז A בו האנזים מראה פעילות שארית נמוכה אך מובהקת. שימוש בטכניקות של השרייה קצרת מועד של הגביש בתמיסה האם המכילה סובסטרטאט, קירור מיידי של הגביש (flash cryo-cooling) ואיסוף נתוני דיפרקציה של קרינת X במאיץ חלקיקים (synchrotron), לאסוף מידע על מבנים מרוכבים בהם מצבי ביניים של הראקציה נתפסו וקובעו. ניסויים אלה איפשרו לי לאפיין סוכר קשור וחתוך. מידע ראשוני מראה את הנוכחות של מצב המעבר הטבעתי - אוקסזוליניום באתר הפעיל.

מחקר זה מרמז כי אנזימים ממסלול מטבולי משותף המפרקים כיתין, הם ממקור אבולוציוני משותף שבו מנגנון הפעולה וחומצות אמינו קטליטיות נשמרים. הבדלים ניכרים ברצף השיירים האמיניים ונוכחותם של תת מבנים (domains) נוספים באנזימים המפרקים כיתין, הולידו את הרעיון שחלבונים אלו, בעלי תת המבנה טים, נוצרו בתהליך של אבולוציה מתכנסת (convergent evolution). הממצאים שלנו לעומת זאת מובילים אותנו להציע היפותזה לפיה לגנים המקודדים לחלבונים ממשפחות 18 ו 20, נגזרו מגן- "אב קדמון" משותף המקודד למבנה חבית טים (divergent evolution). חלבונים אלו התבדרו על מנת לרכוש מבנה המאפשר פעולה על סובסטרטאט ייחודי (substrate specificity). מגבלות אבולוציוניות קרוב לוודאי שימרו את גדיל ה-  $\beta$  ולולאה # 4 והמבנה המיוחד של הגלוטמאט והאספרטאט

הקטליטיים. מעניין כי מנגנון הפעולה של אנזימי כיתינאז ממשפחה 19, שאין בהם מבנה חבית טים הוא של חומצה-בסיס מטיפוס היפוך (invertin). מספר חלבונים ממשפחה 18, כגון חלבוני שימור מזרעים, קונקנבלין B ונרבונין הם ללא פעילות קטליטית. לחלבונים אלה תת מבנה חבית טים הדומה לזה של כיתינאזות. בשני החלבונים הלא פעילים האלה החותמת השמורה של משפחה 18, DXXDXDXE שונתה ל DXXDXDXQ ול DXXDXHXE בקונקנבלין ובנרבונין בהתאמה. באופן דומה גם בחלבונים מאדם C3L1 ו- OGP39 שלאחרונה אופיינו כחלבונים לא פעילים ממשפחה 18, חלו שינויים (מוטציות) בחותמת השמורה אבולוציונית. המודל שלנו לגבי פעילותם של Asp313 ו-Glu315 מכיתינאז A, וכן Asp539 ו-Glu540 מכיתוביאז מספקים הסבר מולקולרי לחוסר הפעילות הקטליטית של חלבונים אלו. המידע הביוכימי והמבני המוצג בעבודה מספק תמונה שלמה יותר המסבירה כיצד החיידק *S. marcescens* מפעיל את האנזימים הכיטונוליטיים בפירוק הכיתין. פירוק הכיתין, קרוב לוודאי, מתחיל בפעולתם של האנדוכיטונואזות, כיתינאז A וכיתינאז C. כיתינאז A גם פועל כאקסוכיטונואז המסיר דימרים (diNAG) מהקצה המחוזר של הכיתין. כיתינאז B פועל כאקסוכיטונואז המסיר טדימרים (triNAG) ודימרים (diNAG) מהקצה הלא מחוזר של הכיתין ושל אוליגומרים של כיתין הנוצרים על ידי הפעולה של כיתינאז A וכיתינאז C. התוצרים האוליגומרים, טרימרים ודימרים, לבסוף נחתכים למונומרים המטבוליטיים (NAG), על ידי כיתוביאז. אנליזה זו מסבירה בבירור את הפעילות הסינרגיסטית של אנזימים אלו. פיענוח מיבנים מרוכבים ברזולוציה גבוהה של אנזימים כיתונוליטיים עם סובסטרטים, איפשר לנו להבין באופן מעמיק את מנגנון ההכרה, הקישור והפעולה של האנזימים האלה. לאחר אנליזה פרטנית של הביוכימיה והמבנה של חלבונים אלה אנו מציעים שכיתינאז A וכיתוביאז הם בעלי מנגנון פעולה דומה.