Study of a Bacillus circulans chitin-binding domain by a green fluorescent protein binding assay and detection of lysozymes by improved zymograms

Markus Hardt
Louisiana State University and Agricultural and Mechanical College, mhardt@lsu.edu
STUDY OF A *Bacillus circulans* CHITIN-BINDING DOMAIN BY A GREEN FLUORESCENT PROTEIN BINDING ASSAY AND DETECTION OF LYSOZYMES BY IMPROVED ZYMOTRAGRAMS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University
and Agricultural and Mechanical College

in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biological Sciences

by

Markus Hardt
Vordiplom Eberhard-Karls Universität Tübingen (Germany), May 1996
December 2002
DEDICATION

“Life results from the non-random survival of randomly varying replicators.”

Richard Dawkins

This work is dedicated to my wife Erika, both of our families and to my friends who have encouraged me and supported my growth as a scientist and as a person.

Very special thanks go to Heinrich, Wolfgang, the Kupferhammer-Fußballer, Daphne Dog and Henry Jones for challenging me not only to think and not only to experience but most of all to enjoy living.
ACKNOWLEDGMENTS

I would like to thank first and foremost my major professor, Dr. Roger A. Laine, for his support, enthusiasm and infinite ideas. I especially wish to thank my research committee members, Drs. Sue G. Bartlett, Grover L. Waldrop, Ding S. Shih, K. Gus Kousoulas and Frank K. Cartledge, for their valuable insights and discussions, which enabled the completion of this work.

In addition, special thanks to Dr. Betty Zhu whose patience and wealth of knowledge were instrumental and more than I could have asked for. Thanks to Marianne McKee, Tiffany Jeanminette, Ryan Hay, Denny Dartez, Shannon O'Neal, and Bryan Perilloux for their valuable help in the lab. For their assistance with microscopic imaging, I wish to thank Cindy Henk and Hollie Hale-Donze of the LSU Socolofsky Microscopy Center.
# TABLE OF CONTENTS

Dedication................................................................................................................................... ii

Acknowledgments ................................................................................................................... iii

List of Tables .......................................................................................................................... viii

List of Figures .......................................................................................................................... ix

List of Equations .................................................................................................................... xii

List of Abbreviations .............................................................................................................. xiii

Abstract .................................................................................................................................... xvi

## Chapter 1: Introduction .........................................................................................................1

1.1 Objective............................................................................................................................1

1.2 Carbohydrates ....................................................................................................................1

1.2.1 Chitin .................................................................................................................................3

1.2.2 Chitosan ............................................................................................................................5

1.2.3 Other Polysaccharides .....................................................................................................5

1.2.4 Peptidoglycan (Murein) .................................................................................................6

1.3 Bacterial Polysaccharolytic Enzymes ..............................................................................6

## Chapter 2: Chitin-Binding Domain of *Bacillus circulans* Chitinase ChiA1 .............8

2.1 Chitinases ...........................................................................................................................8

2.2 *Bacillus circulans* WL-12 ...............................................................................................9

2.3 Modular Structure of ChiA1 .............................................................................................10

2.3.1 FNIIID Domain ..............................................................................................................11

2.3.2 Catalytic Domain ............................................................................................................11

2.4 Chitin-Binding Domain ChBD_{ChiA1} ...........................................................................14

2.4.1 Structure of Binding Domain .......................................................................................14

2.5 Functional Role of Carbohydrate-Binding Domains (CBDs) .........................................15

2.6 Sequence Analysis ............................................................................................................17

2.7 Classifications of Carbohydrate-Binding Domains .........................................................17

2.8 Structural Neighbors of ChBD_{ChiA1} .........................................................................20

2.9 Binding Mechanism ...........................................................................................................23

2.9.1 Interactions between Proteins and Carbohydrates .....................................................23

2.9.2 ChBD_{ChiB} ...................................................................................................................25

2.10 Structures of Other Microbial Carbohydrate-Binding Domains ....................................26

2.10.1 Thermodynamic Analysis of Carbohydrate-Binding .................................................26

2.11 Residues Proposed to Be Involved in Binding ..............................................................29
Chapter 3: Green Fluorescent Protein .................................................................31
  3.1 GFPuv and Other GFP Variants ........................................................................32
  3.2 Applications of GFP ........................................................................................33
  3.3 Monitoring of Protein Folding by GFP ...............................................................34
  3.4 Regulation of the araBAD Promoter .................................................................35
  3.5 Evaluation and Optimization of Recombinant Protein Production Using GFP ..................36
  3.6 GFP in Electrophoretic Methods ......................................................................36

Chapter 4: Cloning of p2x12 and Site-Directed Mutagenesis of ChBD-GFP Fusion Protein .................................................................38
  4.1 Reagents ............................................................................................................38
  4.2 Bacterial Strains, Plasmids, and Culture Medium ................................................38
  4.3 General DNA Techniques ...................................................................................38
  4.4 Synthesis of Oligonucleotides ..........................................................................38
  4.5 Transformation and Plasmid Isolation ...............................................................39
  4.6 Construction of pHis-ChBD-GFPuv .................................................................39
  4.7 Cloning of ChBDChla1 .....................................................................................39
  4.8 Site-Directed Mutagenesis by Overlap-Extension PCR ......................................40
  4.9 Results and Discussion ....................................................................................42

Chapter 5 Isolation and Purification of ChBD-GFP Fusion Proteins .................47
  5.1 Introduction .........................................................................................................47
  5.2 Methods and Materials .....................................................................................47
    5.2.1 Induction of ChBD-GFP Protein Expression by L-Arabinose .........................47
    5.2.2 Induction of Mutant-Proteins and Influence of Incubation Temperature on Solubility ....47
    5.2.3 Production of Recombinant Proteins .............................................................48
    5.2.4 Visualization and Quantitation of GFP in Polyacrylamide Gels ......................49
    5.2.5 Protein Assay ..............................................................................................51
    5.2.6 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) ..................................51
    5.2.7 Affinity Chromatography ............................................................................51
  5.3 Results and Discussion ....................................................................................52
    5.3.1 Protein Production and Purification ...............................................................52
    5.3.2 Induction of ChBD-GFP Protein Expression by L-Arabinose .........................52
    5.3.3 Induction of Mutant-Proteins and Influence of Incubation Temperature on Solubility ....52
    5.3.4 Purification of Fusion Proteins ....................................................................54
    5.3.5 Chitin Affinity Chromatography ................................................................56

Chapter 6: Adsorption of Carbohydrate-Binding Proteins to Polysaccharides ...............................................................................58
  6.1 Introduction .......................................................................................................58
  6.1.1 Mathematical Approach ...............................................................................58
  6.2 GFP-Based Binding Assay .................................................................................62
  6.3 Establishment of Assay Conditions ....................................................................63
    6.3.1 Fluorescent Spectra of GFP, Wild-Type ChBD-GFP Fusion Protein and its Mutants...64
LIST OF TABLES

Table 1: Proteins Used for Multiple Sequence Alignment ........................................................... 18
Table 2: Microbial Carbohydrate-Binding Domains Whose Structure Has Been Solved .......... 27
Table 3: Applications of Green Fluorescent Protein in Various Research Areas ..................... 34
Table 4: Oligonucleotides Used for Cloning and Site-Directed Mutagenesis ............................... 44
Table 5: List of Plasmid Constructs Used in this Study ............................................................... 45
Table 6: Commonly Coeluted Proteins from Ni-NTA Columns .................................................. 55
Table 7: Sensitivity of Commonly Used Methods to Determine Protein Concentration ............ 67
Table 8: Binding Parameters of ChBD to Insoluble Chitin Preparations ................................... 77
Table 9: Coefficients Of Determination (R^2) Obtained from Nonlinear Regression Analysis of ChBD-GFP Adsorption Isotherms by Different Binding Models ................................. 78
Table 10: Measured Adsorption Parameters for the Binding of ChBD-GFP and Mutant Proteins to Chitin Beads ............................................................................................................. 81
LIST OF FIGURES

Figure 1: Structure of Cellulose Tetramer ................................................................................................... 2
Figure 2: Structure of Chitin Tetramer ........................................................................................................ 4
Figure 3: Modular Architecture of Chitinases A1 (ChiA1) and D1 (ChiD1) from *B. circulans* WL-12 ................................................................................................................................. 10
Figure 4: Schematic Drawing of the Structure of the Catalytic Domain of ChiA1 ........................................ 13
Figure 5: Schematic Ribbon Drawing of the Tertiary Structure of ChBD_ChiA1 ........................................ 15
Figure 6: Nucleic Acid and Corresponding Amino Acid Sequence of ChBD_ChiA1 .................................... 16
Figure 7: Multiple Sequence Alignment of ChBD_ChiA1 with Related Proteins ........................................... 19
Figure 8: Amino Acid Sequence Alignment of *B. circulans* WL-12 ChBD_ChiA1 (ChiA1) with the ChBD of Chitinase B (ChiB) of *S. marcescens* and CBDEGZ (CBDEGZ) of *E. chrysanthemi* endoglucanase Z ........................................................................... 20
Figure 9: Superposition of the Structures of ChBD_ChiA1, ChBD_ChiB and CBDEGZ ................................. 21
Figure 10: Sequence Alignment between 1ed7 (ChBD_ChiA1), the Related 1aiw, 1e6z and the 1fc5 and 1e1t Proteins, Which Are Different in Sequence but Similar in Structure ................................................................ 22
Figure 11: Schematic Representation of the Cellulose-Binding Domain CBD_EGZ of *E. chrysanthemi* ........................................................................................................................................... 24
Figure 12: Structure of Chitinase B of *Serratia marcescens* ....................................................................... 25
Figure 13: Map of the Electrostatic Potential of the Solvent Accessible Surface of ChBD .......................... 28
Figure 14: Structure of ChBD_ChiA1. Residues H681, T682, W687, and E688 are Suspected to be Involved in Binding ......................................................................................................................... 29
Figure 15: Schematic Diagram of the NMR-Derived Structure of ChBD_ChiA1 ......................................... 30
Figure 16: Schematic Representation of the Crystal Structure of the Wild-Type Green-Fluorescent Protein ................................................................................................................................. 32
Figure 17: Schematic Map of the Plasmid p2x12 Containing the Nucleotide Sequence for the Chitin-Binding Domain-GFPuv Fusion Protein ..................................................................................... 41
Figure 18: Scheme of the Introduction of Site-Specific Mutations by Overlap Extension PCR ................. 43
Figure 19: Analysis of the Site-Directed Mutagenesis of H681A ................................................................. 46
Figure 41: Localization of Fungal Infection in Tissue with Calcofluor White............................. 95
Figure 42: Localization of Fungal Infection in Tissue with FITC-Wheat Germ Agglutinin....... 96
Figure 43: Localization of Fungal Infection in Tissue with ChBD-GFP......................................96
Figure 44: Zymogram of Hen-Egg White Lysozyme and Ostrich Lysozyme......................... 100
Figure 45: Zymogram of a Dilution Series of Hen-Egg White Lysozyme............................... 102
Figure 46: Plot of Logarithm of Applied Lysozyme Units vs. Areas of Clearing Zones in Zymogram................................................................................................................... 103
Figure 47: Excitation Spectrum of GFPuv, ChBD-GFP and Mutated ChBD-GFP Fusion Proteins .................................................................................................................................135
Figure 48: Emission Spectrum of GFPuv, ChBD-GFP and Mutated ChBD-GFP Fusion Proteins .................................................................................................................................136
LIST OF EQUATIONS

Equation 1: Equilibrium Equation for Adsorption................................................................. 59
Equation 2: The Concentration of Available Binding Sites \( n \).................................................. 59
Equation 3: Basic One-Site Binding Langmuir Equation...................................................... 59
Equation 4: Double-Reciprocal Form of Transformed Langmuir Equation.......................... 60
Equation 5: Two Binding-Site Langmuir Model Equation.................................................... 61
Equation 6: Empirical Freundlich Equation ......................................................................... 61
Equation 7: Combined Langmuir Freundlich Equation (Analogous to Hill Equation) ........ 61
Equation 8: Temkin Model .................................................................................................... 61
Equation 9: The Jovanovic Model for One Adsorption Site.................................................. 62
Equation 10: Depletion Isotherm .......................................................................................... 68
Equation 11: Modified Langmuir Equation ......................................................................... 79
Equation 12: Original Affinity Equation .............................................................................. 86
LIST OF ABBREVIATIONS

Amp: Ampicillin
ANDE: Nondenaturing gel affinity electrophoresis
bp: Base pairs
°C: Degree Celsius
CAZY: Carbohydrate-active enzyme database
CBD: Carbohydrate-binding domain
ChBD: Chitin-binding domain
ChBD_{ChiA1}: Chitin-binding domain of \textit{Bacillus circulans} Chitinase A1
ChBD-GFP: Chitin-binding domain-GFP fusion protein
ChiA1: Chitinase A1 of \textit{Bacillus circulans}
ChiB: Chitinase B of \textit{Serratia marcescens}
DEAE: Diethylaminoethyl
DALI: Distance matrix alignment
DNA: Deoxyribonucleic acid
EDTA: Ethylenediaminetetraacetic acid
FACS: Fluorescence-assisted cell sorting
FNIIID: Fibronectin type III domain
FRAP: Fluorescence-recovery after photobleaching
FSSP: Families of structurally similar proteins
GFP: Green fluorescent protein of \textit{Aequorea victoria}
GFPuv: GFP variant optimized for excitation with ultraviolet light
GlcNAc: 2-acetamido2-deoxy-glucopyranosid
GMS: Gomori’s methenamine silver
HEWL: Hen egg white lysozyme
IPTG: Isopropyl-1-thio-β-D-galactopyranoside
ITC: Isothermal titration microcalorimetry
kDa: Kilo Dalton
LB: Luria-Bertani broth
LB Amp: Luria-Bertani broth with ampicillin
M: Molar
μL: Microliter
mL: Milliliter
ML: Micrococcus lysodeikticus
μM: Micromolar
mM: Millimolar
MurNAc: 2-acetamido-2-deoxy-3-O-lactyl-glucopyranosid
NMR: Nuclear magnetic resonance
OEWL: Ostrich egg white lysozyme
PAGE: Polyacrylamide gel electrophoresis
PCR: Polymerase chain reaction
PDB: Protein data bank, a repository of three dimensional biological macromolecular structure data
PFAM: Protein Families database of Alignments and hidden Markov models
pmol: Picomoles
PMT: Photomultiplier tube
R.M.S.D.: The root mean square superposition residual in Angstroms
RNA: Ribonucleic acid
SDS: Sodium dodecylsulfate
SMART: Simple Modular Architecture Research Tool
TIM-barrel: The prototypical eight-stranded α/β barrel fold of triose phosphate isomerase, which the most commonly observed tertiary structure observed in protein crystal structures.

Tris: Tris (hydroxymethyl) aminomethane

UV: Ultraviolet

VAST: Vector alignment search tool
ABSTRACT

A fluorescent binding assay was developed to investigate the effects of site-directed mutagenesis on the binding affinity and binding specificity of the chitin-binding domain of chitinase A1 from Bacillus circulans WL-12. The chitin-binding domain (ChBD) was genetically fused to the N-terminus of the green fluorescent protein, GFP. The polyhistidine-tagged hybrid protein was expressed in Escherichia coli under the dose-dependent regulation of the araBAD promoter and purified using metal affinity-, chitin- or ion-exchange chromatography. Residues suggested to be involved in binding from previous three-dimensional studies were mutated and their contributions to binding and substrate specificity were evaluated by depletion assays. Purified fusion proteins were incubated with chitin beads, polysaccharide-protein complexes were removed by centrifugation and the free protein concentration was measured fluorometrically. The experimental binding isotherms were analyzed by non-linear regression using a modified Langmuir equation. Binding affinity and specificity were alternatively studied by affinity electrophoresis under non-denaturing conditions. Non-conservative substitution of tryptophan residue (W687) with alanine abolished chitin-binding affinity. Double mutation E668K/P689A also impaired binding significantly. Other substitutions in the binding site had little effect on overall affinity for chitin. Interestingly, mutation T682A led to a higher specificity towards chitinous substrates than observed for the wild-type. Furthermore, the ChBD-GFP hybrid protein proved to be useful for specifically labeling cell walls of fungi and yeast and for the detection of fungal infections in tissue samples.

Additionally, an improved method for detecting cell lytic activity by a color-based zymogram was developed. Proteins were separated by electrophoresis in SDS-polyacrylamide
gels copolymerized with Remazol-brilliant-blue labeled whole cells of Micrococcus lysodeikticus. After electrophoresis, the enzymes were allowed to refold and lyse the blue-labeled cells embedded in the gel, producing clearing zones in an otherwise bluish gel. This improved zymogram method allows the rapid, sensitive and simultaneous determination of cell lytic specificity, relative activity and molecular weight. This assay should be useful for many research disciplines investigating the role of lysozymes and other cell wall hydrolases capable of refolding after SDS treatment.
CHAPTER 1
INTRODUCTION

1.1 Objective
The goal of this study was (1) to better understand the general nature of interactions between carbohydrate-binding domains and polysaccharides and (2) to specifically identify the binding site residues of the Bacillus circulans ChiA1 chitin-binding domain (ChBD), their individual contribution to affinity and specificity towards chitin and other polymers and (3) to use ChBD as an affinity label to localize chitin in biological samples.

To achieve the above goals, the B. circulans ChBD$_{	ext{ChiA1}}$ was fused to the green-fluorescent protein (GFP). The integration of GFP promised advantages in several aspects of the study: (1) optimization of protein expression, (2) monitoring of global folding characteristics of fusion proteins, (3) monitoring of protein purification, (4) as a highly sensitive and specific signal in affinity electrophoresis, (5) display of mutations in native electrophoresis, (6) as a highly sensitive and specific signal for measuring fusion protein concentration in depletion binding assay, and (7) in fluorescence microscopy.

In this study, the value of the versatile GFP-fusion tag for mutagenic studies is shown as well as the potential for studies using combinatorial approaches (e.g. directed evolution). Furthermore, an improved zymogram method for fast screening of lysozyme activities was developed.

1.2 Carbohydrates
Carbohydrates are the primary products of carbon dioxide fixation by plants and comprise the bulk of biomass on earth in the form of mono-, di-, oligo- and polysaccharides. The multiple linkage types and branching of carbohydrates yield a remarkable diversity that is not
matched by nucleic acids or peptides. Oligosaccharides can vary in linear sequence and identity of the subunits, anomeric configuration (α, β), position of linkage on ring, ring size (furanose or pyranose), branching and possible additional substitutions (Laine, 1997). A variety of monosaccharides can be homo- or heteropolymerized. Oligo- and polysaccharides play important roles in biological processes and functions such as structure, energy storage, biological recognition and adhesion.

The most abundant polysaccharide in nature is cellulose, followed by hemicelluloses and chitin. Whereas cellulose is a linear homopolymer of β-1,4 linked glucose, chitin is a linear homopolymer of β-1,4 linked 2-deoxy-2-acetamido-D-glucose (N-acetyl-glucosamine, GlcNAc) and hemicelluloses are heteropolymers. The most abundant subunit in hemicellulose is D-xylose, followed by various other sugars, including D-glucose, D-glucuronic acid, D-mannose.

Hemicelluloses have an amorphous structure of little strength. In comparison, the two very

Figure 1: Structure of cellulose tetramer
similar structures of cellulose and chitin are crystalline, strong, and resistant to hydrolysis. The polymers can extend over a thousand residues, forming flat and rigid shapes stabilized by oxygen bridges. They can pack into parallel crystalline arrays stabilized by hydrogen bonds. This molecular strength makes cellulose in plants and chitin in fungal cell walls and arthropod cuticles an integral structural component that gives form, structural support and protection due to its resistance to chemical and even to enzymatic hydrolysis. Cello-oligosaccharides as well as chito-oligosaccharides become insoluble in water if the chain length exceeds about nine to ten units.

Since polysaccharides display the largest renewable biomass, the ecological role of polysaccharides in the global carbon cycle is of great environmental impact.

1.2.1 Chitin

Chitin can be found in a wide range of organisms. Chitin forms a polymer that allows it to function as a load-bearing component of the skeletal materials of many lower animals, for example the exoskeleton of arthropods (including insects and crustaceans). Chitin is also found in coelenterates, nematodes, protozoa, mollusks and the cell walls of many fungi. Chitin is almost always associated with other structural components like protein and glucans (Gooday, 1990). In the cuticle of crustaceans, chitin is also associated with calcium carbonate and pigments forming a complex armor-like matrix (Shimahara et al., 1984). There are varying degrees of acetylation of chitin occurring in nature, providing a continuum from fully acetylated to completely deacetylated. Deacetylation may be involved in interaction with proteins (Blackwell, 1988). Completely acetylated chitin can be found in the spines of certain marine diatoms. This highly crystalline material lacks a protein matrix. Chitin has an obvious structural similarity to cellulose: Chitin and cellulose chains have the same basic 2 helical conformation, in which two monomer residues repeat every 10.3-10.4 Å (Blackwell, 1988). Under the electron
Figure 2: Structure of chitin tetramer.

microscope, chitin has a fibrous, semi-crystalline morphology, with amorphous, less ordered edges (Blackwell, 1988). If more than one in six monomers in chitin is D-glucosamine, the polymer would be considered a form of chitosan (Blackwell, 1988).

Chitin occurs in two, possibly three, polymorphic forms based on crystalline packing: the α-form (crustaceans, insects, fungi) is more common than the β-form, which is found exclusively in structures of aquatic organisms (e.g. squid pens, spines of polychaete Aphrodite, Pogonophora tubes, protective tubes of deep sea vestimentiferan worms (Lamellibrachia satsuma) located on hydrothermal vents on the east Pacific ridge). The possible third form, γ-chitin, has been suggested to compose the stomach lining of Loligo (Blackwell, 1988). α-chitin has an antiparallel arrangement of polymer chains, while the configuration of β-chitin is parallel. This can be distinguished by infrared spectroscopy and X-ray diffraction spectra. α-chitin forms hydrogen bonds between sheets, β-chitin does not. β-chitin swells readily in water, where it
forms a series of crystalline hydrate structures. In the parallel chain structure, all chains have the same directionality. In the swollen state, water molecules are intercalated between the stacks of chains. β-chitin can be irreversibly converted to α-chitin by treatment with acid (Blackwell, 1988). In this case, bundles of chains oriented in one direction meld with bundles parallel in the other direction. There is a strong similarity of chitin to cellulose: Native cellulose (Cellulose I) has parallel chains, and does not swell in water, but in alkali. Swelling and a regeneration from solution leads to the more stable cellulose II form with antiparallel chain arrangement (Blackwell, 1988).

1.2.2 Chitosan

Some investigators consider chitosan to be defined as a linear chitin-polymer with an acetylation degree of less than 17%. This yields a positively charged polymer. Chitosan can be derived from chitin by alkaline or enzymatic deacetylation. Chitosan is soluble in weak acids. Partially deacetylated chitin has potent immunological activities (Tokura et al., 1999), such as the activation of peritoneal macrophages in vivo (Nishimura et al., 1984), the suppression of growth of Meth-A tumor cells in syngeneic mice and the stimulation of nonspecific host resistance against E. coli infections (Nishimura et al., 1984). Other reported functions of this aminopolysaccharide include coagulation of pollutants (No and Meyers, 2000), acceleration of burn healing (Nelson et al., 1994), lowering blood cholesterol levels (Ylitalo et al., 2002), improvements of crop yields (Bhaskara Reddy et al., 1999), enhancement of drug dissolution (Illum, 1998), and viscosity modification (Wang and Xu, 1994).

1.2.3 Other Polysaccharides

Mannans are β-1,4 linked homopolymers of mannose that are more flexible than cellulose or chitin and therefore pack less tightly. Xylans are β-1,4 linked polymers of
xylopyranose, which form twisted structures. Xylans can have various substitutions (acetyl, arabinofuranosyl and glucuronosyl residues), which also prevent tight packing. Polysaccharide components of the plant cell in addition to cellulose include xyloglucans, galactomannans, pectins and glucans. Starch is an accessible storage sugar that is composed of α-amylose, a helical linear chain containing a thousand or more α-1,4 linked glucopyranose residues, and amylopectin, a highly branched structure with an average of 20 α-1,4 linked glucopyranose units branched by α-1,6 bonds.

1.2.4 Peptidoglycan (Murein)

Peptidoglycan is composed of alternating 2-acetamido-2-deoxy-D-glucopyranoside (GlcNAc) and 2-acetamido-2-deoxy-3-O-lactyl-D-glucopyranoside (MurNAc) residues, cross-linked with short D-amino acid containing peptides. Peptidoglycan forms a rigid cell wall surrounding all bacteria and is the target of penicillin and lysozymes.

1.3 Bacterial Polysaccharolytic Enzymes

O-glycoside hydrolases (EC 3.2.1.-) are a widespread group of enzymes, which hydrolyze the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. A classification system based on amino-acid sequence comparison has been established (Henrissat, 1991; Henrissat and Davies, 1997). Currently 87 glycosyl hydrolase families are listed in the carbohydrate-active enzyme database (CAZY)¹ (Coutinho and Henrissat, 1999), with structural representatives for 31 families.

Microorganisms synthesize extracellular enzymes that convert the abundant insoluble carbohydrate polymers like cellulose, hemicellulose and chitin to yield soluble oligomers and monomers that can be taken up and used as carbon and energy source. Due to the complexity of

¹ Carbohydrate-Active Enzymes server (Coutinho, P and Henrissat, B (1999) at URL: http://afmb.cnrs.fr/~cazy/CAZY/index.html
the substrate and the inaccessibility of the often crystalline and insoluble polysaccharides, many organisms have developed sophisticated glycosyl hydrolase systems with different substrate and product specificity. Vibrios, for example, may have more than 15 proteins dedicated to chitin hydrolysis, transport and usage (Park et al., 2000). These microorganisms participate in the important ecological recycling of polysaccharides in nature. It has been estimated that in the aquatic biosphere alone about $10^{11}$ metric tons of chitin are produced annually, with the majority being salvaged by marine bacteria (Tracey, 1957). There are no reports of massive accumulations of chitin in nature, and the occurrence of chitin is limited in fossils, which leads to the conclusion that carbon and nitrogen fixed in chitin is recycled rapidly (Keyhani and Roseman, 1999).

So far over 400 genes of polysaccharolytic enzymes (amylases, cellulases, hemicellulases, chitinases, etc.) have been cloned, sequenced and grouped into families. Most of the hydrolytic enzymes have a modular structure with a distinct catalytic domain and one or more noncatalytic modules. These latter modules are predominantly substrate-binding domains that are independent of the active site. Some of the auxiliary domains function as linkers or in protein-protein interaction. Certain Clostridia species produce cellulolytic complexes (cellulosomes) made of multiple cell-wall degrading enzymes organized on a scaffolding protein by means of interacting dockerin and cohesin domains (Beguin and Lemaire, 1996). No such organized ‘chitinosomes’ have yet been described. Expression of polysaccharolytic enzymes can be constitutive or be induced by the presence of substrate or substrate fragments depending on the enzyme and organism (Jeuniaux, 1966).
CHAPTER 2
CHITIN-BINDING DOMAIN OF *BACILLUS CIRCULANS* CHITINASE CHIA1

2.1 Chitinases

Chitinases (EC 3.2.1.14) catalyze the degradation of the fibrous insoluble polysaccharide chitin by hydrolyzing the β-1,4 glycosidic linkages either proceeding from the nonreducing end (exo-) or randomly within (endo-) the chitin chain. Based on amino acid sequence and structural similarities, chitinases belong to families 18 and 19 of glycosyl hydrolases (Henrissat, 1991). The catalytic domains of family 18 chitinases have a characteristic \((\alpha/\beta)\)\textsubscript{8}-TIM-barrel fold. Class 19 chitinases have a high \(\alpha\)-helical content and their structure resembles hen egg-white lysozyme (Henrissat and Davies, 1997).

Chitinases are found in organisms that possess chitin as a constituent (fungi, yeast, crustaceans, insects) as well as in organisms that do not synthesize chitin (bacteria, plants, vertebrates) (Jeuniaux, 1966). There are, however, some exceptions. For example, rhizobia synthesize Nod-factors (chitin oligomers with various substituents, including an N-linked fatty acid) to signal the leguminous host plant in the nodulation process (John et al., 1993). In glycoproteins the core region of asparagine-linked oligosaccharides contains \(N,N'\)-diacetylchitobiose (Kornfeld and Kornfeld, 1985). And finally, the epidermal cuticle of the bony fish *Paralipophrys trigloides* is chitinous (Wagner et al., 1993).

Invertebrates require chitinases for partial degradation of their exoskeletons. Fungi and yeasts use these enzymes to modify chitin, which is used as an important cell wall component, and require them for cell separation during growth (Kuranda and Robbins, 1991). Chitinase production by higher plants is suggested to be a part of their defense mechanism against fungal pathogens (Benhamou, 1995). In vertebrates, chitinases are produced in insectivorous fishes,
amphibians, reptiles, birds and mammals and are likely to play a digestive and a defensive physiological role (Overdijk et al., 1996). Production of chitinases is widely distributed among bacteria. Bacteria produce chitinases to digest chitin and utilize it as carbon and energy sources. Chitinolytic bacteria, therefore, play an important role in the ecological recycling of chitin in nature (Gooday, 1990). Chitinases have been found in gliding bacteria, pseudomonads, vibrios, Photobacterium, enteric bacteria, actinomycetes, bacilli and clostridia (Jeuniaux, 1966; Gooday, 1979).

The three-dimensional structures of the following chitinases have been determined and their structures are deposited in the Brookhaven Protein Databank (PDB)² (Bernstein et al., 1977): (1) endochitinase from barley (Hordeum vulgare) seeds (Song and Suh, 1996) (PDB identifier 1cns), (2) hevamine with combined chitinase and lysozyme activities from the plant Hevea brasiliensis (2hvm) (Terwisscha van Scheltinga et al., 1996), (3) chitinase ChiA from Serratia marcescens (1edq) (Papanikolau et al., 2001) and (4) chitinase ChiB from Serratia marcescens (1e6z, see Figure 12) (van Aalten et al., 2000) and (5) a chitinase from the pathogen Coccidioides immitis (1d2k) (Hollis et al., 2000). Unlike the bacterial enzymes, the two plant chitinases and the fungal chitinase do not contain chitin-binding domain separate from their catalytic domains.

2.2 Bacillus circulans WL-12

As many chitinous invertebrates, protozoa and fungi live in the soil, or accumulate in it after death, it is not surprising that bacteria capable of chitin degradation can be readily isolated from soil. Bacillus circulans WL-12 is a gram-positive soil bacterium capable of lysing yeast and fungal cell walls (Tanaka and Phaff, 1965; Rombouts and Phaff, 1976). In the presence of chitin, B. circulans can secrete up to 10 different chitinases, derived by proteolytic modifications of the

² http://www.rcsb.org/pdb/index.html
primary gene products ChiA1, ChiC1 and ChiD1 (Alam et al., 1996), which all have multi-domain structures (Watanabe et al., 1990). Several other bacteria, including *Streptomyces lividans* (Miyashita et al., 1991), *Aeromonas* sp. (Shiro et al., 1996), *Serratia marcescens* (Harpster and Dunsmuir, 1989) are reported to secrete multiple, genetically different chitinases to form a synergistic chitinolytic enzyme system for the effective degradation of the polymer.

### 2.3 Modular Structure of ChiA1

Exo-chitinase ChiA1 of *Bacillus circulans* WL-12 is composed of 699 amino acids, has a molecular weight of 74 kDa and is an example of a family 18 modular bacterial glycosyl hydrolase. ChiA1 is composed of an N-terminal catalytic domain, Asp\(^{42}\)-Val\(^{460}\) with the characteristic (β/α)\(_8\)-barrel fold, two almost identical (74.4% sequence identity) fibronectin type III-like domains (FnIIIDs), Ala\(^{464}\) to Thr\(^{549}\) and Ala\(^{559}\) to Thr\(^{644}\), and the C-terminal chitin-binding domain ChBD\(_{ChiA1}\), Ala\(^{655}\)-Gln\(^{699}\), see Figure 3. The individual structures of the FnIII domains and the chitin-binding domain have been solved by NMR (Ikegami et al., 2000; Jee et al., 2002) and the tertiary structure of the catalytic domain was revealed by x-ray crystallography (Matsumoto et al., 1999).

![Figure 3: Modular architecture of chitinases A1 (ChiA1) and D1 (ChiD1) from *B. circulans* WL-12. SP, signal peptide (solid black bar); CatD, catalytic domain (gray bar); FnIIID, fibronectin type III domain (white bar); ChBD, chitin-binding domain (dark gray bar). Numbers in parentheses indicate the extend of the domains.](image-url)
2.3.1 FNIIID Domain

The fibronectin type III domain (FnIIID) is a seven-stranded \( \beta \)-sandwich fold that is very common in modular proteins and can be found in about 2% of animal proteins (Bork and Doolittle, 1992). The FnIIIDs of ChiA1 were the first fibronectin type III-like structures found in bacteria (Watanabe et al., 1990). In general, bacterial FnIIIDs have been identified exclusively in glycosyl hydrolases (chitinases, cellulases, amylases) of soil bacteria (Little et al., 1994). Although the function of bacterial FnIIIDs is still unknown, a role in spatial orientation of the domain in respect to the catalytic domain has been suggested (Watanabe et al., 1994). A deletion study of ChiA1 (lacking the second FnIID or both FnIIIDs) showed a reduction of enzymatic activity towards colloidal chitin, but did not impact chitin-binding activity (Watanabe et al., 1994). The structure of FnIID of chitinase ChiA1 has been solved by NMR (Jee et al., 2002) and shows significant similarity to the Greek key \( \beta \)-sandwich fold displayed by animal FnIIIDs.

The acquisition of the FNIII domain is regarded as an example of horizontal gene transfer from animal to bacteria (Little et al., 1994). Soil bacteria are thought to acquire significant proportions of their genome from other organisms (Little et al., 1994).

2.3.2 Catalytic Domain

The catalytic domain of ChiA1 has an \((\alpha/\beta)_8\)-TIM-barrel core structure with eight \( \beta \)-sheets forming the inside of the barrel surrounded by eight helices characteristic for family 18 glycosyl hydrolases (Matsumoto et al., 1999). The three-dimensional structures of several family 18 chitinases have been solved, including the previously mentioned hevamine (PDB entry: 2hvm) (Terwisscha van Scheltinga et al., 1996), chitinase ChiA (1edq) (Papanikolau et al., 2001) and ChiB from Serratia marcescens (1e6z, see Figure 12) (van Aalten et al., 2000) and the chitinase from Coccidioides immitis (1d2k (Hollis et al., 2000)).
The catalytic mechanisms of chitinases belonging to the glycosyl hydrolases families 18 and 19 differ. The mechanism of family 19 glycosyl hydrolases has been well characterized, as described below, for hen egg-white lysozyme (HEWL) (Blake et al., 1965). For a long time, the “Phillips” mechanism was believed to be the model for the catalytic mechanism of β-glycosidases with a net retention of the configuration at the anomeric carbon atom (Phillips, 1967). The protonated carboxylic acid (Glu35) of lysozyme was thought to act as a general acid catalyst protonating the glycosidic oxygen of the scissile bond, leading to bond cleavage and formation of a positively charged oxycarbonium ion. The deprotonated carboxylate (Asp52) was believed to stabilize the transient oxycarbonium intermediate via electrostatic interactions. Chemical modification (Lin and Koshland, 1969) and site-directed mutagenesis (Malcolm et al., 1989) confirmed the involvement of the proposed residues. However, new studies (Vocadlo et al., 2001) indicate that HEWL forms a covalent intermediate with the substrate as first proposed by Koshland (1953). Instead of stabilizing the long-lived oxycarbonium intermediate, Asp52 acts as a nucleophile forming the glycosyl-enzyme intermediate.

In contrast, the reaction mechanism of family 18 glycosyl hydrolases involves only a single glutamic acid, Glu144 in ChiB or Glu204 in ChiA1, residue as catalyst and the proton donors and acceptors are ‘substrate assisted’ (Watanabe et al., 1993)). The acetamide carbonyl oxygen of the neighboring substrate unit acts as nucleophile in a double displacement mechanism involving an oxazoline intermediate (Brameld et al., 1998). Two small inserted β-domains between the seventh and eighth strand of the barrel form in conjunction with the barrel a deep substrate-binding cleft. These insertions also classify ChiA1 of B. circulans to the subfamily A of family 18 glycosyl hydrolases (Matsumoto et al., 1999). Stacking interactions mediated by aromatic residues lining the binding cleft are believed to bend and twist the chitin oligomer chain
at the third sugar ring from the reducing end. Cleavage is expected to occur at the second linkage, consistent with the observation that $N,N'$-diacetylchitobiose was the limit digest product when ChiA1 hydrolyzed chitin oligomers ($N,N'$-diacetylchitobiose is not a substrate) (Watanabe et al., 1990).

Residues that are involved in catalytic activity of ChiA1 are Glu 204 and Asp 200, as shown by site-directed mutagenesis (Watanabe et al., 1993). In alignment with the binding cleft, the solvent-exposed tryptophan residues Trp122 and Trp134 are thought to be involved in guiding the chitin chain into the catalytic site (Watanabe et al., 2001) (Figure 4). The distances between the aromatic rings of W122, W134 and the Y56 at the entrance of the binding cleft correspond to the position of multiple GlcNAc units in a chitin oligomer.

![Figure 4](image.png)

**Figure 4**: Schematic drawing of the structure of the catalytic domain of ChiA1. β-sheets are represented as blue arrows and α-helices as red barrels. Aromatic residues in the binding cleft Tyr 56 and Trp 53, Trp433, Trp164 and Trp285 are shown. Green indicates the catalytic residues Glu 204 and Asp 200. Trp 134 and Trp122 on the surface are shown. The approximate alignment of the chitin chain is indicated by the yellow line.
Amino acid replacement has shown that W122 and W134 of the catalytic domain are essential for the hydrolysis of crystalline chitin and appear to be involved in the binding of ChiA1 to crystalline chitin (Watanabe et al., 2001).

The catalytic domain of ChiA1 is very similar in sequence and structure to chitinase A of Serratia marcescens. It has been shown that in chitinase A the additional linker domain ChiN contains exposed tryptophans that guide chitin chains into the catalytic site (Uchiyama et al., 2001), however, no solvent-exposed tryptophans can be found in the corresponding FnIII domains of ChiA1 of B. circulans.

2.4 Chitin-Binding Domain ChBD_{ChiA1}

2.4.1 Structure of Binding Domain

The three dimensional structure of the C-terminal chitin-binding domain ChBD_{ChiA1} encompassing the residues Ala^{655}-Gln^{699} of ChiA1 has been determined by NMR (Ikegami et al., 2000). The structure of ChBD_{ChiA1} is compact and globular. The tertiary structure is composed of two antiparallel β-sheets, a shorter one formed by the β-strands β_{1} (Thr^{660}-Tyr^{662}) and β_{4} (His^{681}-Ser^{683}), and a longer one composed of β-strands (β_{2} (Gln^{666}-Tyr^{670}), β_{3} (Lys^{673}-Cys^{677}) and β_{5} (Trp^{696}-Leu^{698}) as shown in Figure 5, which form a twisted β-sandwich with an angle of about 45° between the sheets. Hydrophobic and aromatic residues (Trp^{656}, Tyr^{662}, Val^{668}, Tyr^{670}, Tyr^{675}, Cys^{677}, Leu^{695}, Trp^{696}) with low solvent accessibility (Ikegami et al., 2000) form the core region of the domain, which, despite its lack of disulfide bonds, is very rigid and compact. The β sheet formed by β_{2}, β_{3} and β_{5} forms a flat surface on the molecule.

ChBD_{ChiA1} is required for the highly specific binding of ChiA1 to insoluble or crystalline chitin. ChBD_{ChiA1} has been reported not to bind to the more flexible conformations of chito-oligosaccharides, soluble derivatives of chitin, or other polysaccharides (Hashimoto et al., 2000).
Deletion studies showed that the C-terminal ChBD_{ChiA1} is required for the specific binding of ChiA1 to insoluble chitin, with ChiA1ΔChBD having only little affinity to chitin (Hashimoto et al., 2000).

![Figure 5: Schematic ribbon drawing of the tertiary structure of ChBD_{ChiA1}. The five numbered β-strands (with annotated endings) form two antiparallel β-sheets. The figure was drawn with ViewerPro (Accelrys).](image)

Interestingly, chitinase D1 of *Bacillus circulans* also binds to insoluble chitin and possesses an N-terminal ChBD that is almost identical to ChBD_{ChiA1} (Figure 3). The third chitinase of *B. circulans*, Chitinase C1, lacks a ChBD and significant chitin-binding activity.

### 2.5 Functional Role of Carbohydrate-Binding Domains (CBDs)

The separation of catalytic and non-catalytic substrate-binding into structurally and functionally independent domains has been observed in other hydrolases that degrade insoluble polysaccharide, including cellulases, hemicellulases, xylanases, amylases and β-1,3 glucanases (Reinikainen et al., 1992; Brun et al., 1997).
The CBD is thought to help deliver the catalytic site to the substrate and assist in maintaining the catalytic site on the substrate and therefore increase the avidity of the enzyme for the insoluble substrate. Deletion studies have shown that removal of the CBD reduced the rate of hydrolysis of insoluble substrate (Watanabe et al., 1994; Svitil and Kirchman, 1998). A second proposed function suggests that the CBD actively disrupts the tight crystal structure of the substrate. By interfering with the non-covalent interactions between adjacent substrate molecules, the glycosidic bond of closely packed carbohydrate chains would be rendered more accessible to hydrolysis by the enzyme (Din et al., 1994). Irreversible adsorption to polysaccharide substrate has been observed for many CBDs by several authors as reviewed by Warren (1996). Third, the CBD may be part of a guiding system that leads the polysaccharide chain into the active site of the catalytic module, providing a mechanism for processivity (Sakon et al., 1997), as will be described later for chitinase B of S. marcescens (van Aalten et al., 2000). In general, carbohydrate-binding domains possess a high specificity: some bind only to amorphous substrate, some only to crystalline substrate. For example, CBH1 of Streptomyces olivaceoviridis binds to α-chitin but not to β-chitin (Zeltins and Schrempf, 1997). However, broad specificity can be seen in some cellulose-binding domains that can also bind to chitin structures.
2.6 Sequence Analysis
To identify the homology of ChBD_{ChiA1} with other proteins, a FASTA3 search (Pearson and Lipman, 1988) was performed against Swissprot, Trembl and TremblNew protein databases in March 2002. The search revealed 20 different proteins that have significant similarity with ChBD_{ChiA1}. Interestingly, the similar sequences are found exclusively in chitinases and not in other polysaccharolytic enzymes (Figure 7), suggesting that they mediate specific binding to chitin and not to other polysaccharides like cellulose. However, the alignment also reveals a relationship with several proteases, among them protease C from *Streptomyces griseus*, for which an involvement in the degradation of chitin-linked proteins has been suggested, but has not yet been experimentally shown (Sidhu et al., 1994). Proteins are closely associated with all forms of chitin, excluding β-chitin of diatoms (Gooday, 1990).

An alternative pathway of chitin degradation involves the deacetylation of chitin to chitosan and a subsequent depolymerization of chitosan by chitosanases. Therefore, it is not surprising that deacetylase DA1 from *Vibrio alginolyticus* has a ChBD as well. However, DA1 has not been reported to have chitin-binding activity (Ohishi et al., 2000). There are several other proteins in the alignment for which the role of chitin binding is unclear, including a putative oxidoreductase from *Y. pestis* and a putative dioxygenase from *S. griseus*.

2.7. Classifications of Carbohydrate-Binding Domains
Based on amino acid sequence homologies, CBDs can be grouped into distinctive families. SMART (Schultz et al., 1998) (Simple Modular Architecture Research Tool) classifies ChBD_{ChiA1} into the chitin-binding domain ChBD3 family (Accession code SM0495), which encompasses 102 protein entries in non-redundant databases (March 2002). Chitin-binding domain type 3 is present mostly in bacterial chitinases, as well as in endoglucanases. Type 1 is a
Table 1: Proteins used for multiple sequence alignment. Identity of sequence with ChBDChiA1 is indicated.

<table>
<thead>
<tr>
<th>#</th>
<th>Identification</th>
<th>Protein</th>
<th>Organism</th>
<th>%</th>
<th>Gram</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CHI1_BACCI</td>
<td>ChBD chitinase A1 precursor</td>
<td><em>Bacillus circulans</em> WL-12</td>
<td>100</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>Q48494</td>
<td>Chitinase</td>
<td><em>Kurthia zopfii</em></td>
<td>86.7</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>Q9KHB3</td>
<td>Chi1</td>
<td><em>Bacillus circulans</em></td>
<td>82.2</td>
<td>+</td>
</tr>
<tr>
<td>4</td>
<td>CHID_BACCI</td>
<td>Chitinase D precursor</td>
<td><em>Bacillus circulans</em> WL-12</td>
<td>64.4</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>Q9XA22</td>
<td>Putative dioxygenase</td>
<td><em>Streptomyces coelicolor</em></td>
<td>57.8</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>Q9FC87</td>
<td>Putative bifunctional protein (frag.)</td>
<td><em>Streptomyces coelicolor</em></td>
<td>55.6</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>PRTC_STRGR</td>
<td>Serine protease C precursor</td>
<td><em>Streptomyces griseus</em></td>
<td>60.0</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>Q9L0J5</td>
<td>Putative serine protease</td>
<td><em>Streptomyces coelicolor</em></td>
<td>55.6</td>
<td>+</td>
</tr>
<tr>
<td>9</td>
<td>Q9WX19</td>
<td>Family 19 chitinase precursor</td>
<td><em>Aeromonas sp.</em> 10S-24</td>
<td>60.0</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>Q9WXD3</td>
<td>Chitinase C</td>
<td><em>Serratia marcescens</em> 2170</td>
<td>53.3</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>Q59143</td>
<td>Chitinase precursor</td>
<td><em>Aeromonas sp.</em></td>
<td>48.9</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Q59141</td>
<td>Chitinase precursor</td>
<td><em>Aeromonas sp.</em></td>
<td>46.7</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>Q59142</td>
<td>Chitinase precursor</td>
<td><em>Aeromonas sp.</em></td>
<td>44.4</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>Q9XA96</td>
<td>Putative serine protease</td>
<td><em>Streptomyces coelicolor</em></td>
<td>42.2</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>Q91I1H5</td>
<td>Chitinase ChiC</td>
<td><em>Pseudomonas aeruginosa</em></td>
<td>44.4</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>Q59145</td>
<td>Chitinase II precursor</td>
<td><em>Aeromonas sp.</em></td>
<td>37.8</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>Q9KSH6</td>
<td>Hypothetical protein VC1280</td>
<td><em>Vibrio cholerae</em></td>
<td>37.8</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>Q99PX1</td>
<td>Deacetylase DA1 precursor</td>
<td><em>Vibrio alginolyticus</em></td>
<td>35.6</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>Q48373</td>
<td>Chitinase precursor Chi 69</td>
<td><em>Janthinobacterium lividum</em></td>
<td>40.0</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>Q9RCG5</td>
<td>Chitinase Chi67</td>
<td><em>Doohwaniella chitinasi</em></td>
<td>33.3</td>
<td>-</td>
</tr>
</tbody>
</table>

chitin-binding domain family of fungal and plant ChBDs, whereas type 2 ChBD can be found mainly, but not exclusively, in insects and viruses.

PFAM (Protein Families Database of Alignments and hidden Markov Models) is a database of multiple alignments and hidden Markov models of protein domains or conserved
Figure 7: Multiple sequence alignment of ChBDChiA1 with related proteins. Sequence alignment was carried out by FASTA3. Conserved amino acid residues are indicated by different shades of background.

protein regions. The PFAM database (Version 7.2) classifies ChBDChiA1 into the carbohydrate-binding domain group CBM_5_12, Accession code: PF02839 (Bateman et al., 2002). The group CBM_5_12, which has 114 members, emerged from the previous separate carbohydrate-binding groups CBM_5 and CBM_12, which were based on a classification by CAZY, the Carbohydrate-active enzyme database. ChBDChiA1 belongs to family 12 carbohydrate-binding modules (mostly chitinases), whereas the structurally related binding domains, ChBDChiB of Serratia marcescens and CBD_EGZ of Cel5 cellulase (formerly known as endoglucanase EGZ) of the gram-negative bacterium plant pathogen Erwinia chrysanthemi have been classified based on sequence conservation to the distantly related family 5 carbohydrate-binding modules. The two databank entries, SM0495 and PF02839, are cross-referenced in entry IPR003610 of the InterPRO database (Apweiler et al., 2001).
Figure 8: Amino acid sequence alignment of *B. circulans* WL-12 ChBDChiA1 (ChiA1) with the ChBD of Chitinase B (ChiB) of *S. marcescens* and CBD_{EGZ} (CBDEGZ) of *E. chrysanthemi* endoglucanase Z. Conserved amino acid residues are indicated by different shades of background. The numbers at the right of each sequence represent the last residue position in the sequence. The number at the top and bottom represent the sequence numbers of ChBDChiA1 and CBDEGZ, respectively. The boxes represent β-strand regions of the protein structures, black boxes indicate a conserved β-sheet three-dimensional structure. Large letters in the sequence indicate 3D-conserved residues. Residues that have been proposed to be involved in binding are indicated by the symbol ● below the letter.

**2.8 Structural Neighbors of ChBD_{ChiA1}**

The solution of the three-dimensional structures of ChBD_{ChiA1} (PDB entry: 1ed7), ChBD_{ChiB} (1e6z) and CBD_{EGZ} (1aiw) allowed establishment of a relation between family 12 and family 5 binding families. According to the FSSP (Families of Structurally Similar Proteins) (Holm and Sander, 1994) database, which uses the DALI (Distance Matrix Alignment) (Holm and Sander, 1993) algorithm, ChBD_{ChiA1} shares a structurally conserved N-terminal sequence with CBD_{EGZ} as has been previously observed by Hashimoto *et al.* (2000).

The VAST algorithm (Vector Alignment Search Tool)\(^3\) (Madej *et al.*, 1995) allows automatic comparison of the 3D protein structures deposited in the Molecular Modeling Database (MMDB\(^4\)), a database of macromolecular 3D structures derived from the protein databank. A VAST search revealed, in agreement with the latest results in the FSSP database, that the C-alpha-atoms of 31 residues of ChBD_{ChiA1} (1ed7) can be superimposed on


Figure 9: Superposition of the structures of ChBD\textsubscript{ChiA1}, ChBD\textsubscript{ChiB} and CBD\textsubscript{EGZ}. The red tubes represent the C-alpha backbone of the equivalent conserved regions of each protein. The individual segments are shown in pink, brown and blue for ChBD\textsubscript{ChiA1}, ChBD\textsubscript{ChiB} and CBD\textsubscript{EGZ} respectively. The residues involved in binding are shown in dark brown for ChBD\textsubscript{ChiB} and blue for ChBD\textsubscript{ChiB}. ChBD\textsubscript{ChiA} residues, which have been proposed to be involved in binding, are shown in green. Superposition was made with the Cn3D program\textsuperscript{5} from the National Center for Biotechnology Information (NCBI) (Wang \textit{et al.}, 2000).

\textsuperscript{5} http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml
equivalent pairs of CBDEGZ (1aiw) (0.9 Å r.m.s.d. (the root mean square superposition residual in Ångstroms) and 31.8% identity over 22 residues), as previously noted by Ikegami et al. (2000) and with the newly solved structure of ChBDChiB (1e6z) of S. marcescens (1.1 Å r.m.s.d. and 29% identity over 31 residues, Figure 9). The sequences have high similarity in the hydrophobic β-sheet core structure. The binding site residues are similar between CBDEGZ and ChBDChiB, but the sequence of ChBDChiA1 in this region is very different (Figure 8). Interestingly, 1ed7 also shows marginal similarity to MoeA, a protein from the molybdopterin synthesis pathway of E. coli (PDB: 1fc5) (Schrag et al., 2001) and LysU, a lysyl-tRNA synthetase from E. coli (1el1t) (Desogus et al., 2000). However, sequence disparity between the proteins (see Figure 10) and their unrelated functionality make their structural similarity seem coincidental.

Figure 10: Sequence alignment between 1ed7 (ChBDChiA1), the related 1aiw, 1e6z and the 1fc5 and 1el1t proteins, which are different in sequence but similar in structure.

Despite the structural similarity, the C-terminal portion of the three carbohydrate-binding domains differs significantly. CBDEGZ exhibits the highly conserved stWWst (small-turn-aromatic-aromatic-small-turn) motif among chitinases and endoglucanases of the former family V CBDs: Ala-Asn-Trp⁴³-Tyr⁴⁴-Thr-Ala, with W43 and Y44 shown to be involved in stacking interaction with glucose. ChBDChiB exhibits a slight variation of the theme: TKWGYYI, with the two (boldface) aromatic residues shown to be involved in binding. ChBDChiA1, however does not share this functional ‘aromatic duo’ binding motif (see Figure 8) and may not be considered integral to the ChBD3 family, as suggested by Brun et al. (2000). Therefore, it is proposed that
ChBD_{\text{ChiA}1} along with the proteins that share the alternate loop sequence (Table 1, Figure 7), form a subfamily of ChBD3 with a different chitin-binding mechanism, as originally suggested by Ikegami et al. (2000).

2.9 Binding Mechanism
2.9.1 Interactions between Proteins and Carbohydrates

The suggested major types of interactions between proteins and their carbohydrate ligands are (1) hydrogen–bonding, (2) direct metal coordination using divalent cations such as Ca^{2+}, Mn^{2+} (C-type lectins) or indirect (legume lectins), (3) van der Waals interactions and (4) hydrophobic interactions (Elgavish and Shaanan, 1997). Carbohydrates display both polar and apolar characteristics. Hydroxyl or other polar groups dominate the periphery of the carbohydrate monomer, whereas the pyranose ring is hydrophobic. The large number of hydroxyl groups of the various sugar molecules offers the most obvious partners for a complex hydrogen bond network with a corresponding array of polar residues of the protein either directly with oxygen or hydroxyl groups of the carbohydrate or indirectly via water molecules (Quiocho, 1989). Hydrogen bonding can help stabilize the appropriate orientation of binding residues or it may contribute to disruptive forces (Din et al., 1994). Hydrophobic interactions are associated with the face-to-face stacking of aromatic side chains, primarily tryptophan and tyrosine along the aliphatic axial face of pyranose rings (Quiocho, 1986; Spurlino et al., 1992). Tyrosine and tryptophan rings are thought to stack against pyranose rings of cellulose providing a hydrophobic driving force for binding and serving as additional hydrogen-bond donors and acceptors (Tormo et al., 1996; Brun et al., 1997; Mattinen et al., 1997). Therefore, it is not surprising that in carbohydrate-binding proteins, aromatic residues (especially tryptophan) are often highly conserved. Involvement of hydrophobic interactions in CBD binding mechanisms has been observed by NMR (Mattinen et al., 1997; Mattinen et al., 1997), microcalorimetry, site-directed
mutagenesis (Din et al., 1994; Mattinen et al., 1997; Simpson and Barras, 1999) and by chemical modification (Bray et al., 1996).

Figure 11: Schematic representation of the cellulose-binding domain CBDEGZ of E. chrysanthemi. CBDEGZ has a ski boot shape, formed by a triple antiparallel $\beta$-sheet perpendicular to a less-ordered loop. The three solvent-exposed aromatic residues (W18, W43, Y44), that have been proven to be involved in binding, are shown (Brun et al., 1997). The aromatic triad could cover about five to six glucose moieties by stacking interaction. The distance between equivalent atoms within three glucose moieties in cellulose is about 10Å.

However, aromatic residues may also play a role in the folding of small domains. The cellulose-binding domain CBD$_{EGZ}$ of E. chrysanthemi, which is structurally related to ChBD$_{ChiA1}$, is a well-studied example of stacking interaction with the crystalline substrate via a hydrophobic face. Three solvent-exposed aromatic rings Trp$^{18}$, Trp$^{43}$ and Tyr$^{44}$ are linearly arranged to form a flat hydrophobic surface. Each aromatic ring is believed to stack on every other pyranose ring of cellulose (see Figure 11). The role of the aromatic residues was verified by amino-acid replacement (Simpson and Barras, 1999).
2.9.2 ChBD$_{\text{ChB}}$

The complete structure of chitinase B, a family 18 glycosyl hydrolase of the soil bacterium *Serratia marcescens* has been recently solved by X-ray crystallography (Figure 12) (van Aalten *et al.*, 2000). ChiB is only the third solved structure of a cellulase or chitinase that includes both the catalytic and carbohydrate-binding domain. A linker region connects the chitin-binding domain stiffly to the catalytic domain. A 55 Å aromatic stretch continues from the

![Figure 12: Structure of chitinase B of *Serratia marcescens*. ChBD attached to the catalytic domain. The side chains of the two aromatic residues involved in the binding of chitin (Trp479, Tyr481) are shown in green. They form, together with two exposed aromatic residues of the linker region (Trp252 and Tyr 240, shown in blue), an aromatic guiding stretch for the chitin substrate towards the catalytic residues (shown in pink). The approximate alignment of the chitin chain is indicated by the yellow line.](image-url)
binding domain to the catalytic domain, a distance corresponding to a binding site of a chitin chain with 10 subsites. This alignment has been proposed to support a mechanism for processivity, allowing the polymer to slide into the active site while retaining its interaction in the binding domain (van Aalten et al., 2000).

2.10 Structures of Other Microbial Carbohydrate-Binding Domains
Carbohydrate-binding domains that exhibit affinity to crystalline substrates contain solvent exposed aromatic residues that form a planar hydrophobic surface implicated in binding to polysaccharides. Examples include Trichoderma reesei CBD_{CBHI} (Kraulis et al., 1989), Cellulomonas fimi CBD_{CEX} (Xu et al., 1995), E. chrysanthemi CBD_{EGZ} (Brun et al., 1997), Thermomonospora fusca CBD_{E4} (Sakon et al., 1997) and Clostridium thermocellum CBD_{CIP} (Tormo et al., 1996). In contrast, domains that bind single polysaccharide chains display an open cleft as seen in C. fimi CBD_{N1} (Johnson et al., 1996) and C. fimi xylan-binding domain XBD_{D1} (Simpson et al., 1999) (Table 2). The following CBDs have been characterized by site-directed mutagenesis: CBD_{N1} (Kormos et al., 2000), CBD_{CBHI} (Reinikainen et al., 1992; Linder et al., 1995; Mattinen et al., 1997), CBD_{CenA} (Din et al., 1994), CBD_{XylA} (Poole et al., 1993; Nagy et al., 1998), CBD_{EGZ} (Simpson and Barras, 1999) and CBD_{CBHI} (Koivula et al., 1998).

2.10.1 Thermodynamic Analysis of Carbohydrate Binding

The thermodynamics of CBDs’ interactions with soluble carbohydrates and crystalline substrates has been investigated by isothermal titration microcalorimetry (ITC) and important thermodynamical differences in the two binding processes have been observed. The binding to oligomers is believed to be driven by enthalpy against unfavorable entropic contribution due to conformational restriction of the soluble ligands upon binding. This reflects
Table 2: Microbial carbohydrate-binding domains whose structure has been solved

<table>
<thead>
<tr>
<th>CBD</th>
<th>Enzyme (Organism)</th>
<th>3D Structure</th>
<th>PFAM Family</th>
<th>Overall Shape</th>
<th>Binding Site Residues</th>
<th>Binding Substrate</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBD_{CBH1} (Cel7A)</td>
<td>Celllobiohydrolase I (Trichoderma reesei)</td>
<td>1cbh (NMR)</td>
<td>CBM_1</td>
<td>Wedge shaped irregular β-sheet</td>
<td>3 conserved tyrosines on hydrophobic planar surface Y31, Y32</td>
<td>Crystalline cellulose, not to chitin, amorphous cellulose</td>
<td>(Kraulis et al., 1989)</td>
</tr>
<tr>
<td>CBD_{EG1}</td>
<td>Endoglucanase I (T. reesei)</td>
<td>NMR</td>
<td>CBM_1</td>
<td>Wedge shaped irregular β-sheet</td>
<td>Three conserved aromatic residues</td>
<td>Crystalline and amorphous cellulose, cello-oligosaccharides</td>
<td>(Mattinen et al., 1998)</td>
</tr>
<tr>
<td>CBD_{EX} (CBM2a)</td>
<td>β1,4 glucanase (Cellulomonas fimi)</td>
<td>1exg (NMR)</td>
<td>CBM_2</td>
<td>Elongated, 9 stranded β-barrel</td>
<td>3 solvent exposed Trp., W17, W54, W72</td>
<td>Crystalline cellulose, chitin</td>
<td>(Xu et al., 1995)</td>
</tr>
<tr>
<td>CBD_{Cp}</td>
<td>Cellulosomal scaffolding subunit (Clostridium thermocellum)</td>
<td>1nbc (X-ray)</td>
<td>CBM_3</td>
<td>9-stranded, jelly roll β sandwich</td>
<td>Planar linear strip of surface exposed aromatic residues.</td>
<td>Crystalline cellulose, chitin</td>
<td>(Tormo et al., 1996)</td>
</tr>
<tr>
<td>CBD_{E4}</td>
<td>Endo/Exocellulase E4 (Thermomonospora fusca)</td>
<td>1tf4 (X-ray)</td>
<td>CBM_3</td>
<td>Antiparallel β-sandwich with 10 β-strands</td>
<td>Strip of conserved polar residues</td>
<td>Crystalline cellulose</td>
<td>(Sakon et al., 1997)</td>
</tr>
<tr>
<td>CBD_{N1} (CBM4-1)</td>
<td>CenC cellulase (Cellulomonas fimi)</td>
<td>1ulo (NMR)</td>
<td>CBM_4_9</td>
<td>10 β strands, in 2 antiparallel sheets forming a jelly roll β sandwich.</td>
<td>Binding cleft with hydrophobic central slip Y19, Y85 flanked by polar groups.</td>
<td>Soluble cello-oligosaccharides, amorphous, not crystalline cellulose</td>
<td>(Johnson et al., 1996)</td>
</tr>
<tr>
<td>CBD_{N2}</td>
<td>CenC cellulase (C. fimi)</td>
<td>1cxl (NMR)</td>
<td>CBM_4_9</td>
<td>11 β-strands forming a jelly roll β-sandwich</td>
<td>Binding cleft</td>
<td>Soluble cello-oligosaccharides, amorphous, not crystalline cellulose</td>
<td>(Brun et al., 2000)</td>
</tr>
<tr>
<td>ChBD_{ChiB}</td>
<td>Chitinase B (S. marcescens)</td>
<td>1e6z (X-ray)</td>
<td>CBM_5_1</td>
<td>3 antiparallel β strands connected by loops</td>
<td>Y479, Y481</td>
<td>Crystalline cellulose, chitin</td>
<td>(van Aalten et al., 2000)</td>
</tr>
<tr>
<td>CBD_{EGZ}</td>
<td>Endoglucanase Z (Cel5 cellulase) (E. chrysanthemi)</td>
<td>1aiw (NMR)</td>
<td>CBM_5_1</td>
<td>Ski boot</td>
<td>3 conserved aromatic residues: W18, W43 and Y44</td>
<td>Crystalline cellulose, chitin</td>
<td>(Brun et al., 1997)</td>
</tr>
<tr>
<td>ChBD_{ChaA1}</td>
<td>Chitinase A1 (B. circulans)</td>
<td>1ed7 (NMR)</td>
<td>CBM_5_1</td>
<td>Twisted β-sandwich</td>
<td>This study, W687</td>
<td>Chitin, not cellulose</td>
<td>(Ikegami et al., 2000)</td>
</tr>
</tbody>
</table>
the dominance of intermolecular hydrogen bonds and/or van der Waals forces in the interaction, as shown for CBD_N1 (Tomme et al., 1996). In contrast, binding to crystalline substrate is primarily driven by entropy, e.g. CBD_Cex binding to crystalline cellulose (Creagh et al., 1996). The hydrophobic stacking interactions result in dehydration of the peptide-carbohydrate binding interface formed by the aromatic side chains and pyranose rings of the carbohydrate and the subsequent return of ordered water molecules to the bulk solution. These processes are associated with large negative heat capacity values.

Figure 13: Map of the electrostatic potential of the solvent accessible surface of ChBD. Blue indicates a positive, red a negative potential. The side-chains of the charged residues K673 and E688 and W687 are shown.

These thermodynamical observations suggest that there may be an important difference between domains that bind crystalline substrates and others that bind
conformational mobile ligands. CBDs binding to crystalline polysaccharides (members of families CBM_3 and CBM_5_12) have solvent exposed aromatic residues on flat or convex binding sites, whereas domains binding to single-stranded polysaccharides (CBM_4_9 family members) have more restricted binding sites, i.e. a groove lined with aromatic residues and flanked by polar amino acid residues.

2.11 Residues Proposed to Be Involved in Binding

ChBD$_{ChiA1}$ contains only three charged residues (Lys$^{673}$, Lys$^{676}$ and Glu$^{688}$). Accordingly, the surface is dominated by non-charged residues, as shown in the map of the electrostatic potential at the protein surface (Figure 13).

Figure 14: Structure of ChBDChiA1. Residues H681, T682, W687, and E688 (side-chains are shown and labeled) are suspected to be involved in binding. Y670, Y675, W696 (green side-chains) are highly conserved aromatic residues at the core of the fold that is shared by other binding domains. An electrostatic map of the surface overlays the structure.
Residues of ChBD_{ChiA} that may be involved in binding were suggested by Ikegami et al. (2000) according to the following criteria: sequence conservation among the group, surface accessibility, and hydrophobic or aromatic characteristics for interaction with chitin. Proposed residues are His^{681} (27.1), Thr^{682} (34.4), Trp^{687} (22.5), Pro^{689} (38.6), Pro^{693} (45.5), values in parenthesis indicate solvent accessibility. All potential binding site residues are located on the same face of the domain (Figure 14). It has been noted that the proposed aromatic binding site residues in ChBD_{ChiA}, His^{681} and Trp^{687}, are significantly less solvent accessible and Thr^{682} and Pro^{693} are relatively more solvent accessible than comparable residues in CBD_{EGZ} (Ikegami et al., 2000). In the present work, a series of ChBD_{ChiA1} mutants was constructed in which the residues that potentially interact with chitin (Figure 15) were independently replaced by a alanine residues or other amino acids.

Figure 15: Schematic diagram of the NMR-derived structure of ChBD_{ChiA1} (Ikegami et al., 2000), highlighting the potential binding site residues investigated in this mutagenic study.
CHAPTER 3
GREEN FLUORESCENT PROTEIN

The jellyfish *Aequorea victoria* found in the Pacific Northwest displays bioluminescence that is emitted from the green fluorescent protein (GFP) (Johnson et al., 1962). The light organs are located on the rim of the hemispherical umbrella of the jellyfish. *In vivo*, GFP receives the energy for the fluorescence by radiationless energy transfer from the aequorin photoprotein, whose own blue light emission is triggered by calcium (Morise et al., 1974). *In vitro*, however, GFP can be excited by UV-light in the absence of any co-factors. Wild-type GFP has a major excitation peak at 395 nm and a minor peak at 475 nm and emits bright green light with an emission peak at 509 nm with a shoulder at 540 nm. GFP is a relatively small protein of 238 amino acids. The gene for GFP has been cloned and sequenced (Prasher et al., 1992) and several variants of GFP have been engineered. The GFP of the sea pansy *Renilla* shares the same chromophore as *Aequorea* (Ward et al., 1980). The *p*-hydroxybenzylideneimidaazolidinone chromophore is formed post-translationally from the serine$^{65}$-tyrosine$^{66}$-glycine$^{67}$ tripeptide sequence. After the auto-cyclization initiated by a nucleophilic attack of the amide of G$^{67}$ on the carbonyl group of Y$^{66}$ and subsequent dehydration, the α-β bond of tyrosine$^{66}$ is oxidized by molecular oxygen (Heim et al., 1994). The time required for the spontaneous chromophore formation has been reported to be from 30 minutes (Waldo et al., 1999) to 4 hours. The crystal structures of GFP (PDB entry: 1ema (Ormo et al., 1996) (Figure 16) and 1gfl (Yang et al., 1996)), revealed that the fluorophore is located on a single central α-helix surrounded by a tightly packed, 11-stranded β-can structure. Short sequence segments cap the top and bottom end of the barrel structure. Overall, the closed cylindrical structure shields the fluorophore effectively from the solvent and gives the protein its reported high stability against proteases, chaotropic
agents, temperature, salt, moderate concentrations of organic solvents (i.e. acetonitrile, imidazole), and detergents (1% SDS) as well as alkaline pH. However, pH values below 5.0 destroy the fluorescence. Denatured GFP has reportedly been able to refold and the fluorescence properties can be restored (Bokman and Ward, 1981; Ward and Bokman, 1982; Surpin and Ward, 1989).

3.1 GFPuv and other GFP Variants

To optimize the characteristics of GFP for various applications an ever-expanding series of GFP-variants has been created and is commercially available. The variants include changes in: excitation/emission spectra, temporal stability, fluorescence intensity and codon usage for
optimal expression in specific organisms (Heim and Tsien, 1996). GFPuv is a variant of GFP optimized for maximal fluorescence upon UV light excitation (360-400 nm) and for high expression in bacteria (Crameri et al., 1996).

GFPuv was developed by in vitro DNA shuffling and displays the ‘cycle 3’ mutations F99S, M153T and V163A (Crameri et al., 1996). The mutations effect protein folding and fluorophore formation, thereby increasing the fluorescence 18-fold compared to the wild-type GFP. Furthermore, Arg codons rarely used in E.coli have been removed from the GFPuv DNA sequence and replaced by codons that are more favorable. GFPuv can be expressed in E.coli as soluble fluorescent protein at higher levels than the wild-type GFP. GFPuv has excitation maxima at 397 nm and 475 nm, with an absorbance extinction coefficient of 30 and 6.5-8.5 \([10^3 \text{ M}^{-1} \text{ cm}^{-1}]\), respectively. The emission spectrum peaks at 506 nm according to literature values with a quantum yield of 0.79 (Crameri et al., 1996).

### 3.2 Applications of GFP

The green fluorescent protein has been used in a broad range of applications (Conn, 1999), listed in Table 3. The versatility of the jellyfish green fluorescent protein (GFP) is mainly due to its ability to function as an independent domain in a fusion-protein construct without altering the properties of the protein of interest. GFP’s stability, fairly small size and lack of requirement for cofactors and substrates make it an ideal reporter protein. GFP can be expressed in many different cell types and organisms and is usually non-toxic for the host organism. GFP can be visualized by fluorescence microscopy and quantitative studies can be done using fluorescence-activated cell sorting (FACS) instruments, fluorimeters, fluorescence-plate readers, or fluorimagers.

---

Table 3: Applications of Green Fluorescent Protein in various research areas.

<table>
<thead>
<tr>
<th>Applications of GFP</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(in vivo) Protein localization</td>
<td>reviewed in (Tsien, 1998) and (Kain et al., 1995)</td>
</tr>
<tr>
<td>Gene expression (reporter of transcriptional activity)</td>
<td>(Chalfie et al., 1994; Li et al., 1998)</td>
</tr>
<tr>
<td>Detection of mutations</td>
<td>(Aoki et al., 2002)</td>
</tr>
<tr>
<td>Protein folding and solubility monitoring</td>
<td>(Waldo et al., 1999)</td>
</tr>
<tr>
<td>Binding assays (small peptide, cortisol, biotin)</td>
<td>(Hernandez and Daunert, 1998; Lewis et al., 1999; Lewis and Daunert, 2000; Deo and Daunert, 2001)</td>
</tr>
<tr>
<td>Process optimization for recombinant protein production</td>
<td>(Rucker et al., 2001; Albano et al., 1998; Zhang et al., 2001)</td>
</tr>
<tr>
<td>Membrane trafficking</td>
<td>(Lippincott-Schwartz et al., 1998)</td>
</tr>
<tr>
<td>Transformation marker</td>
<td>(Lukacova et al., 1999)</td>
</tr>
<tr>
<td>FRET (Fluorescence Resonance Energy Transfer) Experiments</td>
<td>(Heim and Tsien, 1996)</td>
</tr>
<tr>
<td>PCR cloning</td>
<td>(Dabrowski et al., 2000)</td>
</tr>
<tr>
<td>Drug screening</td>
<td>(Valdez et al., 1998)</td>
</tr>
<tr>
<td>DNA affinity chromatography</td>
<td>(Jarrett and Taylor, 1998)</td>
</tr>
<tr>
<td>Ca^{2+}-biosensor</td>
<td>(Miyawaki et al., 1997; Romoser et al., 1997)</td>
</tr>
<tr>
<td>Exploration of protein interaction</td>
<td>(Miyawaki and Tsien, 2000; Mahajan et al., 1998; Park and Raines, 1997; Kiessig et al., 2001)</td>
</tr>
<tr>
<td>FRET based protease assay</td>
<td>(Heim and Tsien, 1996; Mitra et al., 1996)</td>
</tr>
<tr>
<td>Cytoplasmic pH measurement</td>
<td>(Elsliger et al., 1999)</td>
</tr>
</tbody>
</table>

3.3 Monitoring of Protein Folding by GFP

As a conceivable result of a site-directed mutagenesis, the folding of the whole protein can be disrupted. The misfolding can cause hydrophobic residues that normally are buried in the interior of the protein to become exposed to the solvent and form interfaces that can lead to aggregation of insoluble protein. Overexpression of a protein by itself can lead to accumulation and formation of inclusion bodies. Although inclusion bodies contain high amounts of pure protein that can be isolated and purified (Georgiou and Valax, 1999), the protein is often
misfolded and inactive. Insoluble GFP is reportedly non-fluorescent and correct native folding is considered to be a prerequisite for the formation of the chromophore. The solubility of overexpressed heterogeneous proteins in *E. coli* can be improved experimentally by the following methods: fusion to a more soluble protein, coexpression of chaperones and folding catalysts, expression at lower temperature, modified growth media and by regulation of protein production (Hannig and Makrides, 1998). However, most of these approaches are largely empirical and the theoretical background remains unknown.

Recently it has been reported that the folding of overexpressed heterogeneous proteins in *E. coli* can be monitored by a C-terminal GFP tag to the protein of interest (Waldo *et al.*, 1999). It has been shown that low fluorescence values for cells expressing GFP-fusion proteins correlate to low solubility of the target protein when expressed without a GFP-tag. The folding proceeds from N-terminus to the C-terminus and correct folding of the whole protein (fusion partner and GFP) can be inferred from the development of GFP fluorescence.

Controlling protein expression aids in maximizing soluble protein yield. Examples of currently available dose-dependent protein expression induction systems are: (1) pBAD expression vectors using the *araBAD* promoter and induction by *L*-arabinose (Guzman *et al.*, 1995), (2) BL21 SI cells (Invitrogen) which use the T7 RNA polymerase under control of osmolarity using the salt inducible *proU* promoter (Donahue and Bebee, 1999), and (3) Tuner cells (Novagen), which are lacYZ deletion mutants of BL21. The lac permease allows uniform entry of IPTG into all the cells allowing a concentration-dependent level of expression.

### 3.4 Regulation of the *araBAD* Promoter

The *araBAD* promoter effects positive and negative regulation by the product of the *araC* gene. *AraC* is a transcriptional regulator that forms a complex with *L*-arabinose. In absence of *L*-
arabinose, AraC dimer bind the O₂ and I₁ half sites of the araBAD operon which form a 210 bp loop that blocks transcription. L-arabinose binds to araC and releases the araC dimer from the O₂ site and the dimer binds to the I₂ site, adjacent to the I₁ site. The DNA loop is released and transcription can begin. The cAMP activator protein (CAP)-cAMP complex binds to DNA and stimulates the binding of AraC to I₁ and I₂. Basal expression can be suppressed by introducing glucose into the growth medium, which lowers cAMP levels (Guzman et al., 1995).

3.5 Evaluation and Optimization of Recombinant Protein Production using GFP

The value of GFP-tags in rapid analysis of a variety of conditions essential for expression of recombinant proteins has been recently recognized (Rucker et al., 2001). The fluorescence of GFP allows the initial identification of bacterial cells that have been transformed with the plasmid of interest without the need of laborious screening methods like colony PCR or restriction analysis of plasmid DNA. Growth conditions for protein expression (e.g. induction time, growth temperature, concentration of inducer) can easily be established by measuring the whole cell fluorescence of standardized bacterial conditions. Cell lysis, protein solubilization, chromatographical purification steps and protein storage conditions can be directly monitored by the specific GFP-fluorescence eliminating the need for immunological detection methods.

3.6 GFP in Electrophoretic Methods

In electrophoretic applications, GFP and GFP-fusion proteins can be detected by their fluorescence with high selectivity and sensitivity even in complex mixtures. The method does not result in difficulties that may arise with radioisotopic or chemical fluorescence labeling as mentioned, nor are traditional staining steps like Coomassie-blue required. To study protein-protein interactions, GFP has been previously used to detect mobility-changes in affinity capillary electrophoresis and in gel based band-shift assays (Park and Raines, 1997; Kiessig et
Here the use of GFP in affinity electrophoresis under non-denaturing conditions to study protein-polysaccharide binding is reported the first time.
CHAPTER 4
CLONING OF P2X12 AND SITE-DIRECTED MUTAGENESIS OF CHBD-GFP FUSION PROTEIN

4.1 Reagents

*Vent* polymerase, DNA ligase, restriction enzymes, deoxynucleoside triphosphates used in PCR reactions, molecular mass markers and chitin beads were obtained from New England Biolabs. Ampicillin, L-(-)-arabinose, chemicals for electrophoresis and all other chemicals were purchased from Sigma unless noted otherwise.

4.2 Bacterial Strains, Plasmids, and Culture Medium

Plasmid pGEM-T easy (Promega, Madison, Wisconsin) was used for cloning of PCR products. pBAD His A (Guzman *et al.*, 1995) (Invitrogen, Carlsbad, California) was used for expression of His6-tagged fusion proteins. *Escherichia coli* DH5α (*F* endA1 hsdR17(rk mk*) supE44 thi-1 recA1 gyrA96 relA1 (argF-lacZYA)U169 ΦlacZ M15) was the host strain used to prepare DNA for cloning and sequencing. *E. coli* strain Top10 (*F* endA1 hsdR17(rk mk*) supE44 thi-1 recA1 gyrA96 relA1 (argF-lacZYA)U169 ΦlacZ M15) (Invitrogen) was used for the expression of the proteins. Gene expression was induced with L-arabinose. Strains were grown in Luria-Bertani broth (LB Amp) medium containing 100 µg of ampicillin per mL. Bacto Tryptone and Bacto Yeast Extract (Difco, Detroit, Michigan) was used to prepare LB medium.

4.3 General DNA Techniques

Various molecular cloning techniques were performed with appropriate enzymes as directed by the manufacturers. Restriction analysis of DNA fragments on agarose gels were performed as described in Sambrook *et al.* (1989).

4.4 Synthesis of Oligonucleotides

Oligonucleotides were synthesized at the LSU GeneLab core facility.
4.5 Transformation and Plasmid Isolation

Plasmid DNA was isolated by the modified alkaline method (Sambrook et al., 1989) or using the Qiaprep plasmid miniprep kit (Qiagen, Valencia, CA). Competent cells for electroporation were prepared by a method adapted from Tung (1995). Ligations were desalted prior to electroporation according to Atrazhev (1996). Electroporation was performed using a Gene-Pulser II (Biorad) with prechilled 0.1 mL cuvettes, under conditions suggested by the manufacturer (25 µF, 200 Ω and 1.8 kV). After electroporation Top10 cells were rescued immediately with 200 µl SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 10 mM MgSO₄, and 10mM MgCl₂), spread on selective agar plates and incubated at 37°C overnight. Positive transformants were identified by restriction analysis.

4.6 Construction of pHis-ChBD-GFPuv

The coding sequence for the GFPuv protein (GFP variant optimized for maximal UV excitation and efficient bacterial expression) was excised from vector pGFPuv (BD Biosciences, Clontech) (Genbank accession # U62636) using the restriction enzymes PstI and EcoRI. The fragment was ligated into the appropriately digested multiple cloning site of the expression vector pBAD His A (Invitrogen) (Guzman et al., 1995) to create the plasmid pB33 that allows expression of N-terminally (His)₆-tagged GFPuv protein. In pBAD, gene expression is under control of the arabinose promoter/repressor araBAD, which is inducible with L-arabinose.

4.7 Cloning of ChBD_{ChIA1}

The DNA region encoding the ChBD of chitinase A1 was cloned into the above described expression vector pB33 downstream of the polyhistidine sequence by polymerase chain reaction (PCR) based methods (Mullis and Faloona, 1987). The DNA for ChBD was amplified from the plasmid pTYB1 (New England Biolabs; Chong et al. (1997)) using the synthetic oligonucleotides CBDF5 and CBDR3 as primers (Table 4). The PCR mixture (100 µL)
contained 5 ng of pTYB1 DNA, 200 µM of each dNTP, 20 pmol of each primer and 3 Units of Vent Polymerase in its recommended buffer. Twenty-five thermal cycles were performed as follows: denaturing at 94°C for 1 min, annealing at 54°C for 1 min, and primer extension at 72°C for 2 min.

The primary PCR product was purified using the PCR Wizard kit (Promega) and the DNA was extended by PCR under the above conditions except that primers, CBDF5PST and CBD3KPN, were used to introduce PstI and KpnI (underlined in Table 4) as flanking restriction enzyme sites. The final product was purified by electrophoresis, digested with PstI and KpnI, and ligated into the PstI/KpnI digested plasmid pB33 to yield plasmid p2x12 (see Figure 17). Positive clones were identified by restriction analysis and verified by DNA sequencing.

To facilitate purification of the recombinant proteins, vectors p2x12 and pB33 carry the nucleotide sequence for an N-terminal hexahistidine tag, which is separated from the inserted ChBD-GFPuv nucleotide sequence by a small linker sequence. Top10 cells were transformed with the newly constructed plasmids by electroporation as described above. After electroporation, cells were spread on selective agar plates soaked with 40 µL L-arabinose solution and incubated at 37°C overnight. Positive transformants were identified by green fluorescence when illuminated with UV-light and by restriction analysis.

### 4.8 Site-Directed Mutagenesis by Overlap-Extension PCR

To replace the residues implicated in substrate binding, the ChBD sequence was subjected to overlap extension PCR mutagenesis (Ho et al., 1989). A scheme for the overlap- extension PCR method is shown in Figure 18. This method requires the use of a total of three separate PCR reactions and the use of four primers. The mutagenic primer F and primer CBD3KPN are used to amplify the temporary product F-C. Primers BadF5 and the mutation site-specific primer R
Figure 17: Schematic map of the plasmid p2x12 containing the nucleotide sequence for the chitin-binding domain-GFPuv fusion protein.

are used to produce B-R. The primers used for each site-specific mutation are listed in Table 4, for example to introduce the mutations Trp687 → Ala, the primer W687A-F is used as primer F. The reverse complement sequence of W687A-F corresponds to primer W687A-B and is used as primer B. Mutagenic primers did not include rare codons for *E. coli* (Codon usage database, [http://www.kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/) (Nakamura *et al.*, 2000)). A typical reaction contained 50 ng of template DNA, 5 pmol of each flanking primer and mutagenic primer, 200 µM of each dNTP and 2 Units of Taq Polymerase (Life Technologies (Invitrogen, Carlsbad CA) in its recommended buffer. Amplification was obtained after 25 cycles of 45 sec at 94°C, 45 sec at 54°C and 70 sec at 72°C. Each of the temporary products F-C and B-R contains the sequence alteration in complementary ends. The PCR products were separated by gel electrophoresis, bands excised and purified with a Qiaquick Gel Extraction Kit (Qiagen) and recombined in an
assembly PCR. The assembly PCR was performed under conditions listed above, except that PfuTurbo Hotstart DNA Polymerase (Stratagene, La Jolla CA) and flanking BadF5 and cbdr3 primers were used exclusively for amplification. After the denaturing step, the strands having matching sequences at their 3’-end overlapped and act as primers for each other. This final assembly PCR step spliced the two temporary products together resulting in the original full-length product now carrying the mutations encoded by the primers R and F. The mutated DNA product was digested with restriction enzymes BgIII and Kpn1 and re-ligated into the appropriately digested original template plasmid pB33. To easily identify successful mutations additional silent mutations were introduced to add an additional restriction site (Nhe1 or Pst1) into the sequence. All mutated vectors were identified by restriction analysis and confirmed by DNA sequencing of the whole gene.

4.9 Results and Discussion
Plasmids pB33 and p2x12 were successfully constructed, harboring the gene for polyhistidine-tagged GFP and ChBD-GFP fusion protein, respectively. The NMR structure of ChBDChiA1 leads to the prediction that seven amino acid residues (Pro680, His681, Thr682, Trp687, Glu688, Pro689 and Pro693) potentially interact with the N-acetyl-glucosamine chain of chitin and contribute to binding. The following site-directed mutants were successfully prepared and purified: P680A, H681W, T682A, W687A, E688Q, P689A, P693A and the double mutant E688K/P689A. The progress and verification of the PCR-based mutagenesis is shown for the mutation H681A in Figure 19. In two separate PCR reactions, the preliminary products F-C and B-R were formed and purified (lane 1 and 2). In a third PCR-reaction, the two fragments were assembled to form the slightly larger product B-R (lane 3). To verify the introduction of the mutation, the PCR product was digested with Nhe1 (lane 4). Although the Nhe1 digest did not reach completion, the presence of the 340 bp fragment indicates the introduction of an additional
Figure 18: Scheme of the introduction of site-specific mutations by overlap extension PCR.
Table 4: Oligonucleotides used for cloning and site-directed mutagenesis. Restriction-sites are underlined. Inverted letters indicate mutations.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Codon Change</th>
<th>Sequence</th>
<th>RE-site</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Flanking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBDF5</td>
<td></td>
<td>5'-ACAAATCCTGGGTATGCCCTGCT-3'</td>
<td>Forw.</td>
<td></td>
</tr>
<tr>
<td>CBDR3</td>
<td></td>
<td>5'-TTGAAGCTCGCACAAAGGC-3'</td>
<td>Rev.</td>
<td></td>
</tr>
<tr>
<td>CBDF5PST</td>
<td></td>
<td>5'-CCTCTGCAGCACAATCTCTGGGTATCCCGCT-3'</td>
<td>PstI</td>
<td>Forw.</td>
</tr>
<tr>
<td>CBD3KPN</td>
<td></td>
<td>5'-CCTCGGTACCATTGAAGCTGCCACAAGGC-3'</td>
<td>KpnI</td>
<td>Rev</td>
</tr>
<tr>
<td>BADF5</td>
<td></td>
<td>5'-ATGCCATAGCATTTTTATCC-3'</td>
<td></td>
<td>Forw.</td>
</tr>
<tr>
<td>Mutagenesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P680A-F</td>
<td>CCC → GCC</td>
<td>5'-CAGGCCCACACACCTGCTGCAGGAGG-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>P680A-B</td>
<td></td>
<td>5'-CCTGCTAGCGAGGTGTGGGCTG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>H681A-F</td>
<td>CAC → GCA</td>
<td>5'-CCGCAACCTGCTGAGAGGATGG-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>H681A-B</td>
<td></td>
<td>5'-CCATCCTGCTAGCGAGGTGCTG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>H681W-F</td>
<td>CAC → TGG</td>
<td>5'-TAATGCTAGGCACCTGACCTCC-3'</td>
<td>PstI</td>
<td>Forw.</td>
</tr>
<tr>
<td>H681W-B</td>
<td></td>
<td>5'-GGAGGTCCAGGGAGGATGACCTTTA-3'</td>
<td>PstI</td>
<td>Rev</td>
</tr>
<tr>
<td>T682A-F</td>
<td>ACC → ACG</td>
<td>5'-CCACAACCTGCTGCAGGATGAGG-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>T682A-B</td>
<td></td>
<td>5'-CCATCCTGCTGCTAGCGAGGATG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>W687A-F</td>
<td>TGG → GCG</td>
<td>5'-CCTGCTAGCGAGGAGGGAACCC-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>W687A-B</td>
<td></td>
<td>5'-TGTTTCGCTCTGCTGCTGAGG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>E688Q-F</td>
<td>GAA → CAA</td>
<td>5'-ACCTGCTGACAGGATGCAACCA-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>E688Q-B</td>
<td></td>
<td>5'-TGTTTCGCTCTGCTGCTGAGG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>E668K-P689A-F</td>
<td>GAA → AAA</td>
<td>5'-CCTGCTGAGGATGGCAAACCA-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>E668K-P689A-B</td>
<td>CCA → GCA</td>
<td>5'-CCTGCTGAGGATGGCAAACCA-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>P689A-F</td>
<td>CCA → GCA</td>
<td>5'-CCTGCTGAGGATGGCAAACCA-3'</td>
<td>NheI</td>
<td>Forw.</td>
</tr>
<tr>
<td>P689A-B</td>
<td></td>
<td>5'-GATGCTTACCACCCTGCAGGAGG-3'</td>
<td>NheI</td>
<td>Rev</td>
</tr>
<tr>
<td>P693F-F</td>
<td>CCT → GCT</td>
<td>5'-TGCTGCTTGGGAGGGATGGCAAACCA-3'</td>
<td>PstI</td>
<td>Forw.</td>
</tr>
<tr>
<td>P693F-B</td>
<td></td>
<td>5'-ATGCCAGCTGCCACAAGGCAGGCAA-3'</td>
<td>PstI</td>
<td>Rev</td>
</tr>
</tbody>
</table>
Table 5: List of plasmid constructs used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBAD His A</td>
<td>Invitrogen (Guzman et al., 1995)</td>
</tr>
<tr>
<td>pGEM-T easy</td>
<td>(Mezei and Storts, 1994) Promega, Madison WI</td>
</tr>
<tr>
<td>pTYB1</td>
<td>New England Biolabs, Beverly MA (Chong et al., 1997))</td>
</tr>
<tr>
<td>pGFPuv</td>
<td>BD Biosciences Clontech, Palo Alto CA (Chalfie et al., 1994; Crameri et al., 1996)</td>
</tr>
<tr>
<td>pB33</td>
<td>761-bp PstI-EcoR1 fragment containing GFPuv gene in pBAD His A</td>
</tr>
<tr>
<td>p2x12</td>
<td>162-bp PstI-KpnI fragment containing ChBD\textsubscript{ChIA1} cloned into pB33</td>
</tr>
<tr>
<td>pP680A</td>
<td>p2x12 in which the Pro680 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
<tr>
<td>pH681A</td>
<td>p2x12 in which the His681 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
<tr>
<td>pH681W</td>
<td>p2x12 in which the His681 codon of ChBD\textsubscript{ChIA1} is substituted by a Trp codon</td>
</tr>
<tr>
<td>pT682A</td>
<td>p2x12 in which the Thr682 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
<tr>
<td>pW687A</td>
<td>p2x12 in which the Trp687 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
<tr>
<td>pE688Q</td>
<td>p2x12 in which the Glu688 codon of ChBD\textsubscript{ChIA1} is substituted by a Gln codon</td>
</tr>
<tr>
<td>pE688K/P689A</td>
<td>p2x12 in which the Glu688 codon is substituted by a Lys codon and the</td>
</tr>
<tr>
<td>pP689A</td>
<td>p2x12 in which the Pro689 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
<tr>
<td>pP693A</td>
<td>p2x12 in which the Pro693 codon of ChBD\textsubscript{ChIA1} is substituted by a Ala codon</td>
</tr>
</tbody>
</table>

of an additional \textit{Nhe1} site. Therefore, the amino acid replacement succeeded. For additional verification and to exclude any PCR errors, the final plasmid was sequenced (Data not shown).
Figure 19: Analysis of the site-directed mutagenesis of H681A. M is the molecular weight standard (100 bp-ladder, New England Biolabs). Lane 1 and 2 are the preliminary products F-C (58 bp) and B-R (353 bp), respectively. Lane 3 is the assembly product B-C (389 bp). Lane 4 is a partial Nhe1 digest of B-C showing the undigested band at 389 bp and the major digestion products at 340 bp, 243 bp, and 146 bp.
CHAPTER 5

ISOLATION AND PURIFICATION OF CHBD-GFP FUSION PROTEINS

5.1 Introduction

Chitin-binding Domain-GFP fusion protein and its mutants were expressed in *E. coli* and purified by immobilized metal-chelate affinity chromatography (IMAC) and chitin affinity or ion exchange chromatography.

5.2 Methods and Materials

5.2.1 Induction of ChBD-GFP Protein Expression by *L*-Arabinose

*E. coli* Top10 cells carrying plasmid p2x12 were grown in 50 mL LB Amp medium in 250 mL Erlenmeyer flasks at 37°C with vigorous shaking. The medium was supplemented with *L*-arabinose in a concentration range from 0.0002 to 0.2 % (w/v). Top10 cells carrying no plasmid were grown as negative control. At various points of time over 24 h, 200 µl samples of the growth culture were taken. The cells were diluted to an OD(600) reading of 0.150 with 10 mM Tris-HCl, 150 mM NaCl pH 7.5 and their whole-cell fluorescence (394 nm excitation and 511 nm emission) was measured using a fluorescence plate reader. The results are shown in Figure 22.

5.2.2 Induction of Mutant-Proteins and Influence of Incubation Temperature on Solubility

*E. coli* Top10 cells carrying plasmids encoding mutated ChBD-GFP sequences, were incubated in culture tubes containing 5 mL of LB Amp at either 25°C or 37°C under vigorous shaking. After a total expression time of 8 hours, cells were harvested. Whole-cell fluorescence was measured as above. Three mL of the culture were pelleted and washed twice with 1 mL of buffer. Pellets were resuspended in 150 µl of buffer and sonicated for cell lysis. The sonicant was centrifuged for 10 minutes at 14000 × g. The pellet was resuspended in 150 µl of buffer. 50
µl of the soluble fraction, and 50 µl of the resuspended insoluble fraction were each diluted with
50 µl of buffer and the fluorescence was measured using 394/511 nm.

5.2.3 Production of Recombinant Proteins

To express GFPuv, ChBDChiA1-GFP and mutants under the araBAD promoter, E. coli
strain Top10 cells were transformed with the vectors pB33, p2x12, or its derivatives by
electroporation. E. coli cells carrying p2x12 were grown at 37°C in 2 liters of SOB medium (20
g Bacto tryptone, 5 g Bacto yeast extract, 0.5 g NaCl per liter, 2.5 mM KCl, 10 mM MgCl₂)
supplemented with 100 µg/mL ampicillin in baffled flasks to assure sufficient aeration. In the
exponential growth phase, protein expression was induced with L-arabinose to a final
concentration of 0.002% (w/v). Following further cultivation for 8 hours, the cells were
harvested by centrifugation. Proteins were produced in soluble form in E. coli and purified by
immobilized metal-chelate affinity chromatography (IMAC). The cells were harvested by
centrifugation (4500×g) for 20 min at 4°C, and the supernatant was discarded. The cell pellet
was resuspended in 1/40 volume 50 mM NaH₂PO₄, pH 8.0, containing 300 mM NaCl (Buffer
A). Cells were lysed by incubation with lysozyme and sonication. Cellular debris was removed
by centrifugation at 8000 × g for 20 min at 4°C. The supernatant fraction was applied to a 5 mL
bed volume Ni-NTA agarose column (Qiagen), which was subsequently washed with 50 mL
buffer A with 20 mM imidazole. Proteins were eluted with buffer A containing 150 mM
imidazole. The proteins were dialyzed twice against 50 mM NaH₂PO₄, pH 8.0, containing 150
mM NaCl (Buffer B) and stored at 4°C with the addition of 0.02% sodium azide to prevent
microbial contamination.

Native ChBD-GFP and mutants with high affinity to chitin were further purified by chitin
affinity chromatography. A chitin column was prepared from chitin beads (New England
Biolabs). Protein solutions were applied to the chitin column, previously equilibrated with buffer B. The column with bound proteins was washed extensively with buffer B and proteins were eluted in two steps with buffer B containing 30% and 40% acetonitrile.

Mutants W687A and double mutant E688K/P689A exhibited low chitin-binding affinity and were purified by ion-exchange chromatography as follows: Proteins were dialyzed against 50 mM NaH₂PO₄, pH 8.0, loaded onto an equilibrated DEAE-Biogel column and washed with the same buffer. Bound proteins were eluted with a linear gradient of NaCl (0-300 mM). Protein concentrations were determined by the method of Bradford (1976) or by absorption at 280 nm. SDS-polyacrylamide gel electrophoresis of the purified fusion proteins in 12.5% polyacrylamide was conducted by the method of Laemmli (1970). Proteins were detected by fluorescence imaging and/or by staining with Coomassie blue solution as described (Laemmli, 1970) (Figure 20, Figure 21). To verify the homogeneity of the protein preparations, an aliquot of each sample was subjected to affinity electrophoresis, as described below.

5.2.4 Visualization and Quantitation of GFP in Polyacrylamide Gels

The Storm 860 system (Molecular Dynamics (Amersham Biosciences), Sunnyvale CA) has previously been used for the visualization and quantification of GFP fluorescence in gels and agar plates (Kondepudi, 1999) in expression analysis and in the study of protein-protein interactions (Park and Raines, 1997). The blue-fluorescence mode (450+30 nm) of the Storm scanner provides efficient excitation of GFP and GFP variants to visualize GFP directly. According to literature values, the detection limits for purified GFP-S65T and EGFP variants in gel electrophoresis polyacrylamide gels is about 8 ng and 15 ng for wild-type GFP. The linear range of detection for each GFP is between 1.5 and 3 orders of magnitude (Kondepudi, 1999). For gel scanning, a PMT setting of 1000 volts was used at a resolution of 100 nm. Harsh
Figure 20: SDS-PAGE of purified ChBD-GFP fusion protein and mutants. Proteins were incubated at 37°C prior to electrophoresis. Gels were stained with Coomassie blue. M: Protein Marker (New England Biolabs) 83 kDa MBP-paramyosin, 62 kDa glutamic dehydrogenase, 47.5 kDa aldolase, 32.5 kDa triosephosphate isomerase, 25 kDa lactoglobulin A. Lanes 1 – 6 are: ChBD-GFP wild type, P680A mutant, H681A, T682A, W687A, E688Q.

Figure 21: Fluorescence Image of SDS-PAGE of purified fusion protein (identical to Figure 20). After electrophoresis, gel was scanned with STORM scanner using the blue-fluorescence mode.
denaturing conditions, like boiling of the sample in the presence of reducing agents, resulted in a loss of fluorescence and were avoided.

### 5.2.5 Protein Assay

Protein concentrations were estimated from absorbance at 280 nm using the molar absorption coefficients \( \varepsilon (280\text{nm}) \) (ChBD-GFPuv) = 55.8 / (mM*cm), which were calculated from amino acid compositions of each protein (Pace et al., 1995) using the WinGene/WinPep software (Hennig, 1999). Alternatively the Bradford assay (Bradford, 1976) was used to determine protein concentrations.

### 5.2.6 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel electrophoresis of the purified fusion proteins in 12.5% polyacrylamide was conducted by the method of Laemmli (1970). GFP maintains its fluorescence under regular SDS-PAGE conditions. After electrophoresis, fluorescent images of the gels were taken with STORM scanner as described above.

### 5.2.7 Affinity Chromatography

To test the binding specificity of ChBD-GFP, the protein was subjected to affinity chromatography. A peptidoglycan affinity column was prepared as previously described (Yoshimoto and Tsuru, 1974). N-acetylglucosamine (GlcNAc)-agarose was purchased from Sigma and chitin beads (Chong et al., 1997) were acquired from New England Biolabs. After initial protein purification using IMAC, proteins were dialyzed and aliquots were applied to the affinity columns and washed with PBS. Protein adsorption to the affinity bed was detected by a handheld UV-light and fluorescence of the affinity bed. Protein elution was monitored by simultaneously measuring the fluorescence of tryptophans present in proteins and GFP-fluorescence. ChBD-GFP failed to bind to GlcNAc-agarose or peptidoglycan-agarose, but bound
strongly to a column with chitin beads. Bound proteins were washed with PBS and eluted with a step-gradient of acetonitrile-PBS buffer. Chitin-affinity chromatography was used to purify the wild type and mutated proteins that demonstrated sufficient affinity to chitin. Low affinity mutants were purified by ion-exchange chromatography.

5.3 Results and Discussion

5.3.1 Protein Production and Purification

The green fluorescent protein was genetically fused to the C-terminus of the chitin-binding domain of \( B. \) circulans chitinase ChiA1. The fusion protein was successfully overexpressed in \( E. \) coli and purified using metal affinity chromatography and ion-exchange chromatography. The typical yield was about 3 mg·L\(^{-1}\). In the absence of a reducing agent, GFP fusion proteins formed dimers in SDS-PAGE analysis. If samples were not boiled for the analysis, fluorescent bands for the monomer, as well as for the dimer, could be detected in the gel after electrophoresis (see Figures 20 and 21).

5.3.2 Induction of ChBD-GFP Protein Expression by \( L \)-Arabinose

The expression of ChBD-GFP protein was regulated by the concentration of \( L \)-arabinose in the growth medium (Figure 22Figure). Approximately 7 h after inoculation, fluorescence began to develop significantly and leveled off after 12 hours. The highest values for whole-cell fluorescence were reached at a final concentration of 0.0002 % \( L \)-arabinose, followed closely by the 0.002 % induction level. Higher \( L \)-arabinose concentrations led to a significant reduction of whole-cell fluorescence. This likely is due to protein-overproduction, which results in accumulation of the GFP-fusion protein in an insoluble and non-fluorescent form.

5.3.3 Induction of Mutant-Proteins and Influence of Incubation Temperature on Solubility

As has been discussed above, the introduction of mutations can alter folding of the binding-domain, which can lead to global misfolding and insoluble proteins. The C-terminal
Figure 22: Concentration Relationship of soluble (fluorescent) ChBD-GFP-production to L-arabinose. The temporal development of whole cell fluorescence (394nm excitation/ 511 nm emission) was measured for *E. coli* Top10 cultures harboring plasmid p2x12 induced with total concentrations of 0.0002 % (w/v) to 0.2 % L-arabinose. Top10 cells grown in LB medium were used as a control.

Fused GFP allows monitoring of folding problems. To optimize conditions for expressions for mutant proteins, the incubation temperature and amount of L-arabinose for each mutant-conjugate were varied. The fluorescence values for the soluble and insoluble fraction are shown for the mutants T682A and W687A in Figure 23. Interestingly, it was found that for mutant proteins optimal expression conditions are at a significantly higher induction level (0.2% L-arabinose) than observed for the wild type. In addition, the fluorescence in the soluble fraction was higher when the mutant proteins were expressed at reduced temperatures. As has been previously reported (Waldo *et al.*, 1999), the fluorescence values in the soluble fraction correlate with high whole-cell fluorescence values. From the time-course of development of GFP-
fluorescence, the effect of point mutations on the stability of individual mutants was observed and with the exception of the H681W mutant, the substitutions only slightly affected the stability. In the case of H681W mutant protein, colonies harboring the plasmid encoding the mutant showed very faint fluorescence. However, neither the variation of temperature or \( \ell \)-arabinose concentrations, or the addition of stabilizing factors to the growth medium allowed expression of soluble and fluorescent H681W mutant protein. Therefore, the introduction of the tryptophan residue at the site led to a major change in the folding of the chitin-binding domain, which must have led to overall misfolding of the fusion protein and prevented GFP-fluorescence. The substitution with alanine (H681A) did not interfere with the fluorescence formation, and fluorescent H681A mutant protein was produced.

5.3.4 Purification of Fusion Proteins

Chitin-binding Domain-GFP fusion protein and its mutants were successfully purified by immobilized metal-chelate affinity chromatography (IMAC) and subsequent chitin affinity chromatography or ion-exchange chromatography for mutant proteins W687A and E688K/P689A with low avidity for chitin. The progress of the purification of mutant protein W687A is analyzed by SDS-PAGE, as shown in Figure 24.

Fractions eluted from the Ni-NTA column contained contaminating proteins, a common observation in metal affinity chromatography. A list of commonly coeluted proteins in IMAC is shown in Table 6. To remove the contaminating proteins, mutant proteins were further purified either by chitin affinity chromatography or, in the case of low chitin affinity mutants like W687A, by ion-exchange chromatography. Occasionally, dimerization of GFP-fusion protein was observed by SDS-PAGE. Dimerization of GFP occurs typically at high protein concentrations (5-10 mg/mL) and/or at high salt concentrations via hydrophobic interactions.
Figure 23: Influence of incubation-temperature and \( L \)-arabinose induction on the expression of soluble mutant fusion protein. Data is shown for mutants T682A (black) and W687A (red) at 25\(^\circ\)C (round) and 37\(^\circ\)C (square data symbols). Measurements for soluble fractions have the top half of the symbol filled, whereas insoluble have the bottom part filled with the respective color.

Dimerization does not inhibit fluorescence, however the excitation peak at 470 nm diminishes and the 395 nm peak increases. Dimerization can be reversed by exposure of the protein to reducing agents.

Table 6: Commonly coeluted proteins from Ni-NTA columns. Source: Newsgroup Discussion.

<table>
<thead>
<tr>
<th>MW (Da)</th>
<th>Protein Name</th>
<th>Protein ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>74283</td>
<td>Hypothetical Methyltransferase</td>
<td>P77398</td>
</tr>
<tr>
<td>69110</td>
<td>HSP70</td>
<td>P04475</td>
</tr>
<tr>
<td>66890</td>
<td>Glucosamine-Fructose-6-Phosphate Aminotransferase</td>
<td>P17169</td>
</tr>
<tr>
<td>23639</td>
<td>cAMP Receptor Protein</td>
<td>P03020</td>
</tr>
<tr>
<td>20851</td>
<td>Peptidyl-Prolyl Cis-Trans Isomerase</td>
<td>P30856</td>
</tr>
</tbody>
</table>
Figure 24: SDS-PAGE analysis of purification of W687A mutant. (A) lane M: protein marker, lane 1: 250 imidazole eluate from Ni-NTA column; Lanes 2-4: fractions from NaCl gradient elution from DEAE-Biogel column. Lane 2 shows the monomer and dimer of the mutant protein. Gels were stained with Coomassie blue. (B) Unstained gels analyzed by the fluorescence scanner. Lanes correspond to those in (A).

5.3.5 Chitin Affinity Chromatography

Wild type ChBD-GFP fusion protein and mutant proteins with significant chitin-affinity were purified by chitin affinity chromatography after IMAC. Elution from chitin beads required 35-40 % acetonitrile indicating strong hydrophobic interactions between the protein and chitin (Figure 25). The fluorescence of GFP decreased at high acetonitrile concentrations, but could be completely recovered upon removal of acetonitrile by dialysis. Once removed from acetonitrile, the purified ChBD-GFP protein displayed homogeneous chitin-binding activity as judged from affinity gel electrophoresis (data not shown).
Figure 25: Elution profile of chitin affinity chromatography of wild-type ChBD-GFP fusion protein. The fluorescence emission of GFP at 511 nm upon excitation at 394 nm and the fluorescence emission of protein at 342 nm upon excitation at 280 nm were measured for each fraction eluted from the column.
CHAPTER 6

ADSORPTION OF CARBOHYDRATE-BINDING PROTEINS TO POLYSACCHARIDES

6.1 Introduction

The adsorption of carbohydrate-binding proteins, -domains and glycosyl hydrolases to polysaccharides is a complex process, in part, because most polysaccharides are insoluble (e.g. chitin, cellulose) and they present a substrate with multiple binding sites that can be considered overlapping (Gilkes et al., 1992). The insolubility leads to the dependence of the concentration of accessible binding sites on the location and solvent accessibility of the sites on the surface of the polysaccharide rather than on the molecular weight of the polysaccharide. Adsorption of a protein ligand to this array of overlapping binding sites leads to another complication, the exclusion of binding sites by bound protein ligand. The number of accessible binding sites is limited due to the physical dimensions of a bound protein, which will cover an area that exceeds more than one binding site. Furthermore, around the bound protein, additional sites are excluded, although not covered, due to steric hindrance. Considering non-symmetrical shapes of the protein ligand, the effective size of the exclusion area can vary. At high protein concentration, slow rearrangements of protein on the surface can minimize the effective exclusion areas and more protein can be accommodated. This could partially explain the kinetic behavior of the adsorption, which consists of a rapid first stage and a slow second step. In many studies of polysaccharide adsorption, saturation of the binding could not be achieved even at the highest protein concentrations employed, which also may be linked to the rearrangement of adsorbed proteins.

6.1.1 Mathematical Approach

Under the assumption that the binding of ChBD to chitin follows classical Michaelis-Menton kinetics, the adsorption equilibrium can be described by Equation 1.
Equation 1: Equilibrium equation for adsorption, where \([B]\) is the concentration of protein-chitin complex in \(\mu\text{mol}\) bound protein/\(\text{g}\) substrate; \([F]\) is the concentration of free protein in solution in \(\mu\text{mol}/\text{L}\); \(n\) is the concentration of available binding sites in \(\mu\text{mol}/\text{g}\) substrate, and \(K_A\) is the association constant of the protein-substrate complex in \(\mu\text{mol}/\text{L}\).

Further assuming that (1) a single protein-ligand only interacts with a single binding site, (2) binding sites are uniform and (3) adsorbing protein molecules do not show positive or negative cooperative effects, the concentration of available binding sites can be expressed as shown in Equation 2.

\[
n = n_0 - [B]
\]

Equation 2: The concentration of available binding sites \(n\) is equal to the concentration of binding sites in the absence of protein ligand \(n_0\) [\(\mu\text{mol}/\text{g}\)] minus the concentration of protein-ligand complexes.

Equation 2 was substituted into Equation 1 and rearranged, resulting in the Langmuir equation (Langmuir, 1918) (Equation 3) that describes basic one-binding site adsorption isotherms.

\[
[B] = \frac{n_0 \cdot K_A \cdot [F]}{1 + K_A \cdot [F]}
\]

Equation 3: Basic one-site binding Langmuir equation.

The Langmuir equation has been used in many studies to describe the adsorption of proteins to polysaccharides. However, the Langmuir equation is often regarded as an oversimplified model, and the key assumptions listed above are not met (Bothwell and Walker, 1995). As previously noted, polysaccharides like chitin and cellulose can be considered arrays of overlapping potential binding sites. The number of available binding sites, therefore, becomes a
probability function depending on the number of bound proteins, their individual configuration
and their packing on the binding site array. The statistical exclusion of adjacent binding sites was
first studied for the one-dimensional case of DNA-protein interactions (McGhee and von Hippel,
1974). The two-dimensional case has been addressed for cellulose adsorption by a simulation
approach (Sild et al., 1996) and by approximation at low occupancy of binding sites, when the
increased spacing of ligands allowed the expected crowding factor to be small (Gilkes et al.,

Gilkes et al. (1992) introduced the distribution coefficient, \( \alpha \), which describes the number
of binding sites occupied by a single ligand molecule, into the original Langmuir equation and
transformed it into the double-reciprocal form to emphasize data at a lower concentration range
(Equation 4).

\[
\frac{1}{[B]} = \frac{1}{K_A \cdot n_0} \cdot \frac{1}{[F]} + \frac{\alpha}{n_0}
\]

Equation 4: Double-reciprocal form of transformed Langmuir equation according to Gilkes et al.
(1992)

Plotting of data in double reciprocal form (1/[B] vs. 1/[F]) allows a linear analysis according to
Equation 4. However, deviations from the model (overlapping binding sites, heterogeneity of
polysaccharide substrate (e.g. crystallinity, acetylation degree of chitin), non-specific binding,
presence of multiple binding sites with different affinities (e.g. crystal faces)) require linear
analysis from the limiting slope, neglecting values at higher protein concentrations. Due to the
inversion of the variables in Equation 4, small values of the dependent variable, which are more
prone to be affected by non-specific binding, will be emphasized in the placement of the fitted
line (Bothwell and Walker, 1995). When limiting slopes are fitted through the low binding data,
the resulting line is often not representative of the overall trend of the data (Gilkes et al., 1992;
Hashimoto et al., 2000). A deviation of linearity can be interpreted as negative or positive cooperativity, presence of multiple binding sites, or as result of non-specific binding. This approach can lead to rather arbitrary slopes assigned to the data.

The adsorption of CBDs and enzymes to polysaccharides has alternatively been described by the Langmuir equation assuming two binding sites (Equation 5) (Medve et al., 1997), and by the empirical Freundlich isotherm for the adsorption on a heterogeneous surface (Equation 6).

Equation 5: Two binding-site Langmuir model equation.

\[ [B] = \frac{n_1 \cdot [F]}{K_{D1} + [F]} + \frac{n_2 \cdot [F]}{K_{D2} + [F]} \]

Equation 6: Empirical Freundlich equation, where \( K \) is the Freundlich equilibrium constant, and \( m \) the power term of the isotherm (\( m > 1 \)). Adsorption energy decreases with the logarithm of saturation.

Other adsorption models include the combined Langmuir Freundlich Model (analogous to Hill equation) (Equation 7) and the Temkin model, which describes adsorption in the case of linear decrease in adsorption energy with increasing saturation (Equation 8).

Equation 7: Combined Langmuir Freundlich Equation (analogous to Hill equation)

\[ [B] = \frac{1}{n[F]^m} \]

Equation 8: Temkin model, where \( k \) is equal \( RT/q_0 \alpha \) and \( b \) is equal to \( k \ln(\alpha_0) \), where \( R \) is the gas constant, \( T \) the absolute temperature, \( q_0 \) the heat at adsorption at zero saturation and \( \alpha \) and \( \alpha_0 \) are constants.
The Jovanovic model (Equation 9) was originally developed to describe the adsorption of gases, but later was also used to describe adsorption of peptides on ion-exchange adsorbents.

\[ [B] = n[1 - e^{-k_F} \] \]

Equation 9: The Jovanovic model for one adsorption site uses the same parameters as the basic Langmuir equation (Equation 3).

6.2 GFP-Based Binding Assay

In this study, a fluorescence binding assay based on GFP fusion-protein was used for the first time to characterize the carbohydrate adsorption of a mutated carbohydrate-binding domain.

GFP has previously been successfully used to study the binding of small peptides (Lewis et al., 1999), cortisol (Deo and Daunert, 2001) and biotin (Hernandez and Daunert, 1998; Deo and Daunert, 2001). GFP offers high sensitivity compared to other methods previously used in carbohydrate adsorption studies, i.e. intrinsic tryptophan fluorescence (Hashimoto et al., 2000), SDS-PAGE (Poole et al., 1993), protein concentration assays (Lowry, Bradford, OD280) (Gilkes et al., 1992). High sensitivity has been alternatively achieved by using radioisotopic labeling (H3-labeled cellulose-binding domain of the major cellobiohydrolase of Trichoderma reesei (Linder and Teeri, 1996)), chemical fluorescent labeling (Creagh et al., 1996) or through amplification of an enzymatic reaction by fused enzymes like β-galactosidase, luciferase, alkaline phosphatase (Poole et al., 1993). However, radioisotopic labeling is associated with potential health hazards and special handling needs (storage, licensing, disposal) and enzyme-based assays require the addition of substrate. A GFP-based assay combines ease of use and sensitivity and has several advantages over the employment of a chemical fluorescent label. For example, traditional chemical fluorescence labeling requires the purification of protein and the covalent conjugation with reactive derivatives of organic fluorophores and a purification of the labeled protein from unreacted dye. The reaction itself can involve substitution of multiple sites, unfavorable side-
reactions, or the involvement or blocking of functional groups, which might interfere with bind-
ing activity. However, this could be prevented by peptide synthesis or by the use of modified
amino acids for expression of protein of interest with special aminoacyltransferases.

Construction of GFP-fusion proteins yields homogeneous samples with exactly one
fluorescent label at a predetermined position either N- or C-terminally to the protein sequence of
interest. Fusion proteins can be cheaply and conveniently produced by overexpression in a host
organism like *E. coli* and the construction of equally labeled mutated proteins is facilitated. The
GFP-tag also makes the purification of the analyte readily monitorable or renders purification
obsolete in some instances due to its specific signal.

Here the construction of a chitin-binding domain GFP fusion protein and the generation
of mutant ChBD genes is reported, in which putative binding-site residues were individually
substituted. The corresponding mutants were purified and their affinity to chitin and other
insoluble polysaccharides were analyzed by means of a fluorescence microplate assay using
GFP-fused polypeptides. The insolubility of the chitin and other polysaccharide substrate is an
inherent difficulty in the measurement of the adsorption. The excitation photons and the
fluorescence photons can be scattered by particles and interfere with the measurement of
fluorescence (Gabel *et al.*, 1971). To circumvent the problem, adsorbed protein concentrations
were calculated from the difference of total protein added and the concentration of unabsorbed
(free) protein in solution.

### 6.3 Establishment of Assay Conditions

Fluorescence measurements were performed on a SPECTRAmax Gemini XS microplate
spectrofluorometer (Molecular Devices, Sunnyvale, CA) using untreated black FluoroNunc
polystyrene plates with 400 µL wells (Nalge Nunc International, Naperville IL). All reported
fluorescence intensities are the average of a minimum of three replicates. To achieve maximal sensitivity for the assay, it was necessary to use an assay volume of at least 200 µL. The same amount of fluorophores (0.3 nmol and 0.06 nmol) recorded about 150 % more fluorescence in a volume of 200 or 300 µL compared to a volume of 100 µL (Figure 26).

**6.3.1 Fluorescent Spectra of GFP, Wild-Type ChBD-GFP Fusion Protein and its Mutants**

To maximize sensitivity, the optimal combination of excitation and emission wavelengths was determined for GFP and ChBD-GFP. The GFP used for this study is the UV-optimized green fluorescent protein (GFPuv) variant with an N-terminal hexahistidine purification tag. ChBD-GFP is identical to the GFP with the exception of the insertion of the ChBDChiA1 chitin-binding domain between the N-terminal His6-tag and the GFP. The literature values for the
emission maximum are 509 nm and 395 nm for the excitation maximum. GFPuv and ChBD-GFPuv in Buffer A (50 mM NaH2PO4, 150 mM NaCl pH 8.0) were scanned with fixed emission

Figure 27: Combined excitation and emission spectrum of ChBD-GFP fusion protein and GFP (511 nm was the fixed emission wavelength, while the excitation was scanned from 350nm to 485 nm in 1nm steps; Emission scan: Excitation fixed at 394 nm and emission measured for wavelengths 465 to 580).

at 511 nm. The excitation maximum was determined to be 394 nm for GFPuv and 396 nm for ChBD-GFP, which is in close agreement to the literature value. No significant difference in the spectra has been determined between GFP and ChBD-GFP suggesting that the additional ChBD domain does not interfere with the fluorescence properties of the green fluorescent protein. The fluorescence emission and excitation spectra of native GFP are retained (Figure 27). The spectra for the mutated ChBD-GFP proteins are shown in Appendix B. The fluorescence properties of


GFP were not altered by the fusion or the mutation in the ChBD domain. The excitation spectrum showed an excitation maximum at 394 nm and a minor peak at 478 nm. The emission spectrum showed a maximum at 511 nm with a shoulder at 545 nm.

### 6.3.2 Calibration Plot for ChBD-GFP Fusion Protein

Stock solutions of ChBD-GFP, GFP and mutant proteins were serially diluted with Buffer A. Calibration curves were constructed separately for each protein by measuring the fluorescence intensity at 511 nm using an excitation wavelength of 394 nm. Buffer A was used as blank for measurements.

Figure 28: Calibration plot of the native GFP and ChBD-GFP fusion protein. Data are average ± one standard deviation (n=3). Detection limits were calculated using S/N=3. Detection limit for ChBD-GFP was $1.32 \cdot 10^{-9}$M and $1.51 \cdot 10^{-9}$ M for GFP.
6.3.3 Determination of Detection Limit

The detection limit determined for ChBD-GFPuv was $1.3 \times 10^{-9}$ M using a signal to noise ratio of 3. A calibration plot for GFPuv revealed that the fluorescence intensity was not changed by the protein fusion (Figure 28). In general, the sensitivity of GFP-based protein quantification is about nanomolar. With a molecular weight of 37750 Da for ChBD-GFP, the sensitivity equals 37.75 ng/mL. The sensitivity is not matched by other commonly employed methods (Table 7).

6.3.4 Time-Association Curve

An association study was performed to demonstrate the binding of ChBD-GFP to chitin beads and to determine the effect of incubation time on the binding. A suspension of chitin beads (9 µg) was incubated with $2.16 \times 10^{-6}$ M ChBD-GFP WT fusion protein for increasing amounts of time. For each measurement, the solution was centrifuged at 10,000 × g and 100 µL of the supernatant was measured for fluorescence (Figure 29). After 90 minutes, no significant increase in the binding can be observed.

Table 7: Sensitivity of commonly used methods to determine protein concentration.

<table>
<thead>
<tr>
<th>Protein Assay</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>NanoOrange</td>
<td>100 ng/mL -10 µg/mL</td>
</tr>
<tr>
<td>BCA Method</td>
<td>0.5 µg/mL – 1.5 mg/mL</td>
</tr>
<tr>
<td>Bradford</td>
<td>1 µg/mL –1.5 mg/mL</td>
</tr>
<tr>
<td>Lowry Assay</td>
<td>1 µg/mL – 1.5 mg/mL</td>
</tr>
<tr>
<td>280nm Absorbance</td>
<td>50 µg/mL -2 mg/mL</td>
</tr>
<tr>
<td>GFP-based</td>
<td>~ 27ng/mL</td>
</tr>
</tbody>
</table>
6.3.5 Ligand Dilution Study

To determine minimum substrate concentrations for the assay, the protein amount was fixed at 2.16×10⁻⁶ M solution of ChBD-GFP and the amount of chitin bead substrate varied (Figure 30). The fluorescence of the supernatant decreased with increasing substrate concentrations until it reached a plateau, indicating that ChBD-GFP binds to the chitin beads until all of it is absorbed.

Figure 29: Effect of the incubation time on the binding of a 2.16×10⁻⁶ M solution of ChBD-GFP to chitin beads.

6.4 Fluorescent Binding Assay

A depletion isotherm assay was used to generate binding data of ChBD-GFP and mutants on chitin. Each isotherm was constructed from the measured concentration of free protein and

\[
[B] = [T] - [F]
\]

Equation 10: Depletion Isotherm. [B], [T] and [F] are the concentration of bound protein, total protein and unbound (free) protein, respectively.
the calculated concentration of bound protein from a binding reaction between ChBD-GFP and chitin (Equation 10). Binding equilibria were measured in triplicates for each protein concentration.

Figure 30: Binder Dilution Study: Adsorption of $2.16 \times 10^{-6}$ M solutions of ChBD-GFP to increasing amounts of chitin beads was measured after incubation for 90 minutes at room temperature.

To exclude the possibility of lateral protein-protein interactions affecting the binding, low protein concentrations were employed. Chitin beads (New England Biolabs) were used as substrate. Various concentrations of protein were incubated in 1.5-mL microcentrifuge tubes containing 120 µg insoluble polysaccharide substrate in a final volume of 120 µL buffer PBS for
3h at room temperature. Control tubes contained protein without chitin. Tube contents were constantly mixed on a rocking platform. The samples were centrifuged twice at 10 000 × g for 3 min to separate free protein from bound protein. To avoid possible scattering problems, 100 µL of the supernatant was collected, diluted to 300 µL with PBS and the fluorescence emission was measured to determine the concentration of free protein using standard curves prepared for each mutant protein separately. The amount of bound protein was calculated from the initial protein concentration. It could be argued that the GFP-tag as well as the His-tag might contribute to polysaccharide binding and could be a factor in non-specific interaction. Therefore, a His-tagged GFPuv was constructed and used it as a negative control to the binding assay.

6.4.1 Effect of pH on Binding

The influence of pH value on binding of ChBD-ChiA1-GFP wild type to chitin was investigated. The following buffers were used at a final concentration of 25 mM: Sodium Citrate pH 5.0 and pH 6.0, sodium phosphate pH 7.0 and pH 8.0, Tris-HCl pH 9.0 and glycine-NaOH pH 10.5. Two hundredth of a mole of protein were incubated with 240 µg of chitin beads in a total volume of 240 µL for 3 h. Fluorescence measurements were taken as described above. The sensitivity of GFP to pH values below pH 5 limited the testable pH-range.

6.4.2 Substrate Specificity Assay

The binding specificity of wild-type and mutated ChBD-GFP proteins for insoluble polysaccharides was determined as follows: Purified proteins (0.1 µM final) were mixed with 240 µg polysaccharide in a total volume of 240 µl PBS and incubated for three hours while shaking.

The polysaccharides used for this experiment were powdered chitin from shrimp and crab shells (Sigma), colloidal chitin (Hirano and Nagao, 1988) (Sigma), chitin beads (Chong et al.,
(NEB), regenerated chitin (Molano et al., 1977), β-chitin from squid pen (Industry Research Limited, New Zealand), Curdlan (β-1,3 D-glucan) from Alcaligenes faecalis (Sigma), lichenan (β-1,3 and 1,4 linked glucan) from Cetraria islandica (Sigma), xylan (β-1,4 linked xylan) from beechwood (Sigma), acid-washed cellulose (Baker), chitosan (Funakoshi Co. Ltd Japan), Micrococcus lysodeikticus lyophilized whole cells (as a source of peptidoglycan, Sigma) and agarose I (Amresco Inc., Solon Ohio). As a negative control, polysaccharides were also incubated with GFP. The percentage of adsorbed protein rate was calculated from the ratio of supernatant fluorescence of ChBD-GFP in the presence and in the absence of polysaccharides. Adsorption percentages were corrected for the basal adsorption of GFP to the polysaccharides.

6.4.3 Influence of NaCl on Binding

In a total volume of 240 µl, 240 µg chitin beads were incubated for 3 hours with 0.85 µM ChBD-GFP and NaCl concentrations varying from 0 to 2 M. Samples were centrifuged and the fluorescence of the supernatant was measured as described above. The percentage of adsorbed protein rate was calculated from the ratio of supernatant fluorescence of ChBD-GFP in the presence and in the absence of chitin. The results are shown in Figure 32.

6.5 Results and Discussion

It has been demonstrated that a genetic fusion of GFP can be successfully used to study the binding properties of a carbohydrate-binding domain. Every aspect of the presented work can potentially be automated, which would allow high-throughput screening of binding domains that could be evolved by combinatorial approaches.

6.5.1 Effect of pH on Binding

The influence of pH value on binding of ChBD_{ChiA1}-GFP wild type and mutants is shown in Figure 31. ChBD_{ChiA1}-GFP wild type shows a maximal binding of 94% at pH 6.0. The binding
percentage drops slightly at higher pH values to 80 % at pH 10.5. The mutant E688Q also has maximum of pH 6.0, however, its binding is less influenced by pH. In comparison to the wild type, E688Q binds more weakly at pH 6.0, about the same at pH 7 and better at all other tested pH values. H681A mutant shows pH sensitivity similar to the wild type with overall lower binding-affinity. T682A mutant has maximal binding at pH 7 and shows a drop at pH 9, its overall affinity is lower than the wild type. The weakly binding double-mutant E688K/P689A shows increased binding at pH 6.0, with a drop at pH 7.0. Binding under acidic conditions (below pH 5.0) was not possible with the fluorescent binding assay, as GFP loses its fluorescence under these conditions.

Figure 31: Effect of pH on binding of ChBD-GFP wild type and mutants to regenerated chitin.
A previous study revealed a binding maximum of tag-free ChBD$_{\text{ChiA1}}$ at pH 9.0 near its isoelectric point and a substantial decrease of binding at pH values below pH 4.0. Acetic acid (20 mM, pH 3.0) was used to elute the ChBD from a chitin affinity column (Hashimoto et al., 2000). Similar to this study, the highest binding ratios for ChBD-GFP fusion protein and mutants are observed under pH conditions near their isoelectric points, namely pI 6.27 for the wild-type, when the ratio of charges on the protein are minimized.

### 6.5.2 Influence of NaCl on Binding

The binding of ChBD-GFP to regenerated chitin was significantly affected by ionic strength (Figure 32). In the presence of higher concentrations of NaCl, more ChBD-GFP

![Figure 32: Influence of NaCl on binding of ChBD-GFP to regenerated chitin.](image-url)
adsorbed to chitin. The results are in agreement with a previous study, which also showed that the binding of ChBD_{ChiA1} is higher at 0.5 M NaCl than in the absence of salt (Hashimoto et al., 2000). Electrostatic interactions between the polysaccharide surface and the protein are expected to be weakened by increased salt concentrations. However, the amount of adsorbed protein increased. Therefore, nonelectrostatic (hydrophobic) interactions are considered to play a dominant role in adsorption ChBD-GFP to chitin. This conclusion is supported by the elution of ChBD-GFP from chitin with acetonitrile, which would disrupt hydrophobic interactions, and the pH insensitivity of the binding.

6.5.3 Substrate Specificity

To determine the structural specificity of ChBD-GFP and how the point mutations affect substrate specificity, various insoluble polysaccharides were incubated with the recombinant proteins. After the bound protein-polysaccharide complexes were removed by centrifugation, the concentration of unbound protein in the supernatant was measured by fluorescence (Figure 33).

With the exception of mutants W687A and E688K/P689A, wild-type ChBD-GFP and its mutants showed very high affinity to chitin, especially chitin beads. High affinity substrates were also colloidal and regenerated chitin. The affinity towards β-chitin was significantly lower and powdered chitin showed only limited adsorption. Interestingly, the wild type and most mutant proteins showed low but significant binding to lichenan, xylan, cellulose, chitosan and agarose. None of the proteins was able to discriminate between chitin substrates with antiparallel arrangement of the polymers (regenerated chitin, colloidal chitin) and β-chitin, which has a parallel arrangement. Similar to the wild type, none of the generated mutant proteins interacted with peptidoglycan (whole bacterial cells) or curdlan. The specificity of mutant proteins towards the different insoluble polysaccharides shifted. [W687A]-ChBD binds only to a small extent to
Figure 33: Binding specificity of ChBD-GFP and its mutants to insoluble polysaccharides. The percentage of absorbed protein (± standard deviation) is shown for each protein and the according substrate. Mutant T682A and P693F showed higher specificity towards chitinous substrates. W687A is the only studied protein, which has residual affinity to murein. Mutants E688Q and P689A have higher affinity to chitosan than the wild type, P689A also binds more to cellulose. E688K/P689A mutant protein showed avidity to chitin beads, but lost avidity to regenerated and powdered chitin. Interestingly, the relative avidity of [E688K/P689A]-ChBD to β-chitin and cellulose compared to other substrates is higher than that of the wild type. [E688K/P689A]-ChBD, however, did not interact with chitosan at all. This suggests that the positive charge of the introduced lysine interferes with the binding to the
positively charged chitosan, since the sole P689A mutation increased the affinity to chitosan. The regenerated and powdered chitin can be partially deacetylated and be positively, resulting in electrostatic repulsion of a positive binding protein, which can explain why the double-mutant [E688K/P689A]-ChBD has lost avidity to these specific chitin preparations. Overall, mutant protein T682A (followed by P693F) displays a more selective binding to chitinous substrates than wild type and other mutant proteins. The E688Q mutant protein shows a very similar binding pattern to the wild type. The results for the binding specificity are in good agreement with previous findings, which found minimal binding of wild-type ChBD towards avicel, soluble starch and chitosan, but significant binding to powdered, colloidal, regenerated and β-chitin (Hashimoto et al., 2000). The low affinity to powdered chitin and other chitin substrates can be explained by the likely heterogeneity of samples. Compared to the reported results for pure ChBD, ChBD-GFP has a somewhat higher affinity to chitin and a broader specificity to polysaccharides, which actually corresponds more to the characteristics of chitinase A1 as a complete protein.

6.5.4 Binding Parameters of ChBD-GFP to Insoluble Chitin Preparations

The capacity and binding affinity of wild-type ChBD-GFP for different preparations of insoluble chitin was determined (Figure 34). ChBD has the highest affinity for β-chitin \( (K_A \ 2.00 \mu M^{-1}) \), followed by regenerated chitin, chitin beads (spherical regenerated chitin), and colloidal chitin (Table 8) The observed capacity was highest for chitin beads with 3.25 µmol/g chitin, followed by colloidal chitin, β-chitin and regenerated chitin. The results illustrate the overall affinity of ChBD to different chitin preparations and the way in which the number of available binding sites on chitin can be influenced by substrate heterogeneity (e.g. differences in degrees of acetylation, crystallinity and crystalline forms) and by substrate accessibility (porosity and the macroscopic size of chitin microfibrils).
As previously mentioned ChBD_{ChiA1} reportedly does not bind to the more flexible confirmations of chito-oligosaccharides or soluble derivatives of chitin (hexa-N-acetylchito-hexaose, 40% deacetylated chitin, carboxymethyl chitin, ethylene glycol chitin) as shown by NMR chemical shift mapping experiments and differential scanning calorimetry measurements (Hashimoto et al., 2000). The presence of soluble chitin derivatives does not affect the binding of ChBD_{ChiA1} to regenerated chitin (Hashimoto et al., 2000).

Table 8: Measured binding parameters of ChBD-GFP to insoluble chitin preparations. Parameters (± standard error) were calculated from adsorption data by non-linear regression using the modified Langmuir equation (Equation 11). $K_A$ is the association constant, $n_0$ is the total concentration of binding sites on the substrate. $R^2$ is the coefficient of determination.

<table>
<thead>
<tr>
<th></th>
<th>$K_A$ [µM$^{-1}$]</th>
<th>$n_0$ [µmol/g]</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chitin beads</td>
<td>0.95 ± 0.13</td>
<td>3.25 ± 0.06</td>
<td>0.987</td>
</tr>
<tr>
<td>Colloidal chitin</td>
<td>0.59 ± 0.10</td>
<td>1.98 ± 0.13</td>
<td>0.987</td>
</tr>
<tr>
<td>Regenerated chitin</td>
<td>1.62 ± 0.33</td>
<td>1.23 ± 0.06</td>
<td>0.967</td>
</tr>
<tr>
<td>β-chitin</td>
<td>2.00 ± 0.83</td>
<td>1.38 ± 0.15</td>
<td>0.846</td>
</tr>
</tbody>
</table>

### 6.5.5 Evaluation of Binding Models

The experimental binding isotherm of ChBD-GFP to chitin was analyzed by non-linear regression using several adsorption models (Figure 35). To facilitate the evaluation of the fitted curves, especially in the low binding concentration range, the experimental data was represented in a semilogarithmic plot.

The basic Langmuir equation $[B] = n_0 [F] / (1 + K_A[F])$ has been widely used to describe the adsorption of carbohydrate-binding domains to insoluble polysaccharides (Creagh et al., 1996; Bolam et al., 1998; Boraston et al., 2001). However, two key assumptions of the model require a uniform binding site and no interaction between binding proteins. Since polysaccharides like chitin can be described as an array of potential overlapping binding sites, the key postulations
Figure 34: Binding isotherms of wt-ChBD-GFP to different insoluble chitin substrates, colloidal chitin, regenerated chitin, β-chitin and chitin beads.

of the model can only be satisfied at low occupation of binding sites. When the Langmuir model along with the Jovanovic model were applied to the data, the fitted curve correlated very well (Figure 35, Table 9).

Table 9: Coefficients Of Determination ($R^2$) obtained from nonlinear regression analysis of ChBD-GFP adsorption isotherms by different binding models.

<table>
<thead>
<tr>
<th>Binding Model</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langmuir</td>
<td>0.991</td>
</tr>
<tr>
<td>Freundlich</td>
<td>0.934</td>
</tr>
<tr>
<td>Hill</td>
<td>0.991</td>
</tr>
<tr>
<td>Temkin</td>
<td>0.932</td>
</tr>
<tr>
<td>Jovanovic</td>
<td>0.996</td>
</tr>
</tbody>
</table>
Figure 35: Curve fitting of the experimental binding isotherm of ChBD-GFP by nonlinear regression analysis. Five different adsorptions models were used: Freundlich model, Langmuir model, Temkin model and the Jovanovic model. The combined Langmuir-Freundlich (Hill) model converged to the basic Langmuir model. Coefficients of determination for the fitted curves are shown in Table 9.

The Jovanovic curve reaches and levels off to asymptotic saturation levels rapidly, which does not correlate to the non-saturable binding behavior typically observed for carbohydrate-binding (Gilkes et al., 1992; Linder et al., 1995; Tsujibo et al., 1998) including ChBD_{ChiA1}. Therefore for the further investigation of the role of individual residues in the adsorption of the chitin-binding domain, binding isotherms for the generated ChBD-mutants were measured and binding parameters were evaluated by the modified Langmuir equation (Kormos et al., 2000) (Equation 11).

\[
[B] = \frac{n_0 \cdot K_A \cdot ([F] - G)}{1 + K_A \cdot ([F] - G)} - G
\]

Equation 11: Modified Langmuir equation. \(N_0\) is the concentration of total available binding sites in \(\mu\text{mol/ g substrate}\), \(K_A\) is the association constant of the protein-substrate complex in \(\mu\text{mol/L}\), 

\[\text{Bound Protein [\(\mu\text{mol/g}\)]} \]

\[\text{Free Protein [\(\mu\text{M}\)]} \]
[F] is the concentration of free protein in µmol/L, [B] is the concentration of bound protein in µmol/g chitin and G is a correction factor for optical effects caused by the presence of fine particles of chitin.

Values for $n_0$ and $K_A$ were obtained by least squares non-linear regression of the isotherm data with Origin software (OriginLab Corporation, Northampton, MA). The isotherm data was also analyzed by a double reciprocal plot to diagnose possible deviations from the Langmuir equation. The data in the double reciprocal plot was linear indicating the absence of negative or positive cooperativity in the binding, as well as the presence of a single uniform class of binding sites.

6.5.6 Adsorption Isotherms

The chitin-binding affinities of wildtype and mutant forms of ChBD$_{ChiA1}$-GFP were determined using a solution depletion method to generate binding isotherms. Concentrations of free ChBD$_{ChiA1}$-GFP were measured directly by the fluorescence assay, and the bound ChBD$_{ChiA1}$-GFP was calculated by subtracting the free protein from the total ChBD$_{ChiA1}$-GFP used for the binding reaction. Binding isotherms of [Bound] (µnmol/g of chitin) vs. [Free] (µM) were generated for ChBD$_{ChiA1}$-GFP wild type, P680A, H681A, T682A, W687A, E688Q, E688K/P689A, and P693F (Figure 36). Binding levels of mutants E688K/P689A and W687A were particularly low and could simply represent the level of nonspecific binding.

6.5.7 Adsorption Parameters

In mutagenic studies, amino acid substitution can lead to global and/or fine conformational changes in the protein. Larger structural changes are expected not to have occurred in the mutations described here, since the C-terminal GFP serves as a folding-monitor and a misfolded ChBD is expected to affect correct folding of GFP, as mentioned above. Only mutation H681W led to the production of non-fluorescent protein, whereas H681A protein was fluorescent.
Table 10: Measured adsorption parameters for the binding of ChBD-GFP and mutant proteins to chitin beads. Parameters (± standard error) were calculated from adsorption data by non-linear regression using the modified Langmuir equation (Equation 11). $K_A$ is the association constant, $n_0$ is the total concentration of binding sites on the substrate. The binding of W687A and E688K/P689A was too low to be measured accurately by the depletion isotherm method and could not be quantified (n.q.). $R^2$ is the coefficient of determination.

<table>
<thead>
<tr>
<th>Protein</th>
<th>$K_A$ [µM$^{-1}$]</th>
<th>$n_0$ [µmol/g]</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.95 ± 0.13</td>
<td>3.25 ± 0.06</td>
<td>0.987</td>
</tr>
<tr>
<td>P680A</td>
<td>3.44 ± 0.59</td>
<td>2.33 ± 0.12</td>
<td>0.965</td>
</tr>
<tr>
<td>H681A</td>
<td>1.08 ± 0.20</td>
<td>1.88 ± 0.07</td>
<td>0.971</td>
</tr>
<tr>
<td>T682A</td>
<td>3.15 ± 0.62</td>
<td>1.44 ± 0.06</td>
<td>0.968</td>
</tr>
<tr>
<td>W687A</td>
<td>n.q.</td>
<td>n.q.</td>
<td>n.q.</td>
</tr>
<tr>
<td>E688Q</td>
<td>1.26 ± 0.64</td>
<td>2.99 ± 0.60</td>
<td>0.990</td>
</tr>
<tr>
<td>E688K/P689A</td>
<td>n.q.</td>
<td>n.q.</td>
<td>n.q.</td>
</tr>
<tr>
<td>P689A</td>
<td>0.95 ± 0.18</td>
<td>2.09 ± 0.16</td>
<td>0.963</td>
</tr>
<tr>
<td>P693F</td>
<td>3.02 ± 0.51</td>
<td>1.31 ± 0.07</td>
<td>0.984</td>
</tr>
</tbody>
</table>

6.5.8 Individual Mutations

Residues on the binding face with hydrogen bonding potential (Thr682, His681) were substituted with alanine. The only charged residue in the binding site, Glu688, was substituted by glutamine in E688Q mutant and by lysine in the double mutations E688K/P689A to conserve the polar characteristic. Solvent exposed residues suspected to be able to establish close van der Waals contacts with the substrate surface (Pro680, Pro689, and Trp687) were replaced by alanine or with tryptophan for His681 and phenylalanine for Pro693.

Alanine replacement at W687 and the double mutation E688K/P689A had the largest affect on ChBD binding. The affinity for chitin was so low that affinity constants could not be reliably estimated by nonlinear regression analysis. The low affinity also complicated the
extrapolation of an accurate saturation value for the number of binding site residues for these two mutants on chitin. The data sets for all other mutations displayed convergence on a saturation value, providing good estimates for binding parameters.

6.5.9 W687A

W687 had greatly reduced the binding, which was too low to be quantified by the binding isotherm method. W687 seems to be crucial for the interaction with chitin. The pyrrolic amine has potential to form hydrogen bonds.

6.5.10 H681

His681 is the only other residue in this mutation series that has a ring structure with double bonds. H681 does not seem to contribute to binding, as the [H681A]-ChBD has a slightly increased affinity for chitin. A substitution by tryptophan was not acceptable and resulted in apparently misfolded non-fluorescent protein. Subtle differences in the π-electron system of histidine may render it less suitable for stacking interactions with the substrate compared to the common aromatic binding site residues phenylalanine, tryptophan and tyrosine. In addition, the pK of the imidazole group allows for a protonation of the ring under physiological conditions near pH 7, which could interfere with any stacking interactions. Linder et al. (1999) successfully used this particular characteristic to substitute binding-site tyrosines with histidine to design a pH-dependent CBD.

Considering the pH-independency of wild-type ChBD_{ChiA1} and of the H681A mutant, it is not surprising that H681 is not involved in a stacking interaction.

6.5.11 Other Mutations

The E688Q mutation did not affect binding to chitin (Figure 36, Table 10) suggesting that a charge is not required in the binding site. E688K/P689A led to strongly reduced binding and
Figure 36 Adsorption isotherms of ChBD-GFP and mutant proteins to chitin beads with non-linear regression using a modified Langmuir equation. Wild type and H681A mutant in graph (A), T682A, E688K/P689A, and P689A in (B) and mutants P680A, E688Q and P693F in graph (C).
completely abolished binding to chitosan. A positive charge leads to electrostatic repulsion of the positively charged chitosan polymer, indicating that E688 must be involved to some degree in binding. The P689A mutation by itself did not change the binding affinity to chitin at all.

The T682A mutation increased binding by three times compared of the wild type, suggesting that threonine contributes little to the binding by hydrogen bond formation. P693F substitution increased binding by a third of native affinity, suggesting that phenylalanine can substitute for the solvent exposed (45.5%) Pro693. The P680A also increased affinity by a third.

The capacity of the chitin beads substrate, \( n_0 \), determined for each mutant protein ranges from 3.25 to 1.31 \( \mu \text{mol/g} \) chitin beads and remained fairly unchanged. As has been pointed out previously for cellulose (Henrissat et al., 1988; Creagh et al., 1996), it is conceivable that chitin is a similar heterogeneous substrate containing different classes of binding sites. Individual mutations may exclude ChBD binding to one class of binding sites, reducing the saturation values.

In a previous study tag-free ChBD\(_{\text{ChiA1}}\) had a relative equilibrium association constant (\( K_r \)) of 14.9 liters/g (Chitinase A1 has a \( K_r \) of 21.1 l/g) towards regenerated chitin and a \( \alpha/[N_0] \) (an indicator of the number of lattice units occupied by a single ligand) of 1.3 g/\( \mu \text{mol} \) (ChiA1: 10.1 g/\( \mu \text{mol} \)) at pH 6.0 (Hashimoto et al., 2000). The data presented here cannot be directly compared to these results, since substrate preparation, protein concentration measurements and methods for binding analysis are different. The authors also reported difficulties measuring low protein concentrations with fluorescence readings at 280 nm excitation and 342 nm emission. The studied protein was purified by
chitin affinity chromatography using low pH conditions (pH 3.0), which conceivably could have led to irreversible protein denaturation and inhomogeneous binding activity of the samples. The authors did not provide evidence of a homogeneous binding activity of purified protein.

In the here presented study, an initial orthogonal purification scheme based on a polyhistidine tag was employed with a subsequent chitin affinity step. The effects of acetonitrile elution were studied by non-denaturing gel affinity electrophoresis (ANDE) and showed homogeneous binding activity of the samples.

6.5.12 Counter Argument

The 27-kDa GFP-tag significantly increases the size of the hybrid protein, and can shield the binding site of the ChBD or it can alter the flexibility of the polypeptide sequence. Experimental data does not demonstrate that GFP-tags interfere in general with biological functions of the target protein (Tsien, 1998). In nature, ChBDs are expressed as part of a larger protein. Therefore, in the ChBD-GFP fusion, GFP could emulate other domains (catalytic and FNIII domains) that are bound to the binding domain in native chitinases.
CHAPTER 7
AFFINITY NONDENATURING GEL ELECTROPHORESIS

7.1 Introduction

Affinity electrophoresis comprises all electrophoretic techniques in which specific interactions of proteins with ligands occur and the mobility of the protein changes in comparison to a reference substance. The principle of affinity electrophoresis in polyacrylamide gels was first used by Takeo to explore the interaction of phosphorylase with glycogen (Takeo and Nakamura, 1972). The approach and its various methodologies have been successfully used for the study of various classes of proteins, including glycosyl hydrolases, as reviewed by Takeo (1984; 1987) and Hořejši and Ticha (1986). The theory of affinity electrophoresis assumes (1) an equilibrium between proteins, ligands and protein-ligand complexes is established within the gel, (2) the protein-ligand complex has zero mobility and (3) the protein concentration is insignificant compared to the ligand concentration. Under these assumptions, the dependence of the change of protein mobility on the concentration of ligand in the gel can be used to calculate apparent dissociation constants (Equation 12) (Horeši et al., 1977) (Horeši, 1979).

\[
\frac{1}{r} = \frac{1}{R_0} \left(1 + \frac{c}{K_D}\right)
\]

Equation 12: Original affinity equation. \(r\) is the relative (to reference protein) migration distance of protein in presence of substrate; \(R_0\) is the relative migration distance of protein in absence of substrate; \(c\) is the concentration of substrate and \(K_D\) is the dissociation constant.

Nondenaturing affinity electrophoresis has been used to separate native proteins based on intrinsic charge and their different affinities to insoluble polysaccharide ligands immobilized in the gel by physical entrapment into the gel network. The high molecular weights of polysaccharides and of the protein-carbohydrate complexes prevent them from moving in the polyacryla-
mide gel because of molecular sieving. Affinity electrophoresis has been used for studying CBD-polysaccharide interactions semi-quantitatively (Simpson et al., 1999; Charnock et al., 2000) and quantitatively for affinity of CBD$_{N1}$ to the soluble barley β-glucan (Kormos et al., 2000) and the affinities of CBD modules of the fungus Piromyces equi (Freelove et al., 2001).

7.2 Methods and Materials
The capacity of ChBD-GFP and mutants to bind different insoluble polysaccharides was evaluated by nondenaturing gel affinity electrophoresis (ANDE). ANDE was carried out by preparing continuous polyacrylamide gels consisting of 7.5 \% (w/v) acrylamide in 25 mM Tris-HCl, 250 mM glycine buffer pH 8.3. Appropriate polysaccharides (0.1\%) were copolymerized in one of the gels. Samples of 20 µL of 2 µM solution of target proteins and non-interacting control GFP were mixed with Bromophenol Blue and glycerol and subjected to electrophoresis at room temperature with a current of 10 mA/gel for 1-2 h. Fluorescence images of the protein gels were acquired with the STORM fluorimager as described above and images were analyzed with TotalLab software (Phoretix, Newcastle upon Tyne UK).

7.3 Results and Discussion
In native polyacrylamide gel electrophoresis with a single pH electrolyte, proteins are separated based on their intrinsic charge (isoelectric point, pI) and molecular weight. The native state of the proteins is maintained. Only the pI of mutants E688Q (pI 6.37) and E688K/P689A (pI 6.46) differs from the wild type (pI 6.27). Their decreased mobility is observable (Figure 37) verifying the presence of the mutations on the protein level. The so called “GFP-display” method has been previously used to detect single amino acid changes in target peptide sequence with the use of native and SDS-urea gels (Aoki et al., 2000; 2002). This method allows monitoring of mutation efficiencies on a protein level, which can prove to be useful in directed evolution studies.
Figure 37: Affinity non-denaturing gel electrophoresis of GFP and ChBD-GFP wildtype and mutants in the presence (B) and absence (A) of swollen chitin. Proteins were electrophoresed in 7.5% non-denaturing polyacrylamide and visualized by fluorescence scanning of the gel. GFP (lane 1), ChBD-GFP (lane 2), P680A (lane 3), H681A (lane 4), T682A (lane 5), W687A (lane 6), E688Q (lane 7), E688K/P689A (lane 8), P689A (lane 9) and P693F (lane 10).

The capacity of ChBD-GFP and mutants to bind to a series of different insoluble polysaccharides (swollen chitin, β-chitin, curdlan, chitosan and cellulose) was evaluated by ANDE. The fluorescence image of ANDE-gels with swollen chitin is shown in Figure 37. The migration of the wild type and most mutants was almost completely retarded by the inclusion of chitin in the gel, even when lower concentrations of chitin were used. The strong protein adsorption and the difficulties preparing affinity gels on a quantitative basis due to the insolubility of the substrate and heterogeneous distribution of chitin in the gel made it unfeasible
to determine dissociation constants based on the affinity equation (Equation 12) for affinity electrophoresis. Individual affinity constants for mutant proteins were instead measured using the fluorescent binding assay. However, qualitative results can be gained from the direct comparison of the migration shifts of each mutant within the same affinity gel. The electrophoretic mobility of the GFP control and the W687A mutant was not affected by the presence of chitin, while the E688K/P689A mutant showed an intermediate retardation. The data indicates that E688K/P689A interacts weakly and W687A does not significantly bind chitin. The same migration pattern resulted for ANDE with β-chitin. No significant changes in the migration patterns were observed for the other polysaccharides studied (curdlan, chitosan and cellulose).

Affinity electrophoresis was also used for the testing of purity and binding homogeneity of purified protein. Although the fluorescence was maintained, 8M urea affected the ligand-binding activity of ChBD-GFP resulting in two fluorescent bands in ANDE: one with binding affinity and one without.
CHAPTER 8
CHBD-GFP FUSION PROTEIN AS A CHITIN SPECIFIC STAIN FOR FLUORESCENCE MICROSCOPY

8.1 Introduction

Fungal infections are a major medical problem for immunocompromised people (e.g. patients with acquired immune deficiency syndrome (AIDS), leukemia, transplant recipients and patients under cancer therapy). The most common systemic mycoses are comprised of: (1) Aspergillosis (infection or colonization of tissue and cavities, commonly bronchopulmonary system, by fungi in the genus Aspergillus, especially Aspergillus niger), (2) Blastomycosis of lung, skin and bone caused by Blastomyces dermatitidis, (3) Candidosis, which is among the most common opportunistic fungous infections (bronchopulmonary system, digestive tract, skin, vagina) caused by Candida albicans or other members of the Candida genus, (4) Coccidioidomycosis of upper respiratory tract and lungs, caused by Coccidioides immitis, (5) Cryptococcosis caused by Cryptococcus neoformans, (6) Histoplasmosis caused by Histoplasma capsulatum. Less common is Penicillosis, an infection caused by fungi of the genus Penicillium. Recently it has been noted by Ponikau et al. (in press) at the Mayo Clinic that over 95% of chronic sinusitis may have a primary cause from mycoses.

Infections can lead to pathological damage and suppress the besieged immune system. Mycoses of deep tissues, in particular, require early diagnosis. A diagnostic method targeting a broad spectrum of fungi and yeast is desirable. Chitin is common to most fungi, as it is a major constituent of the cell wall. Traditionally, histochemical stains have been used for the detection of fungi (Okudaira, 1985). Gomori’s methenamine silver (GMS) (Grocott, 1955) is the most commonly used stain in pathology to create contrasting images between fungus and host tissue,
however it is non-discriminative with respect to connective tissue polysaccharides (e.g. glycosaminoglycans, mucins). Other chemical stains commonly employed include: India ink, for the staining of capsulated organisms such as *Cryptococci*; Mag fura, believed to respond to elevated concentrations of divalent cations at chitin sites within the fungal cell wall and has been utilized for localizing chitin in a range of fungi and yeasts (Cox and Thomas, 1999). The fluorescent stilbene derivatives Uvitex 2B (Fungiqual A) (Wachsmuth, 1988), blankophor (Ruchel and Schaffrinski, 1999) and calcofluor white, a cotton whitener, have been used to detect fungal infections in tissue (Monheit *et al.*, 1984; Wachsmuth, 1988) and pathogenic fungi *in vitro* (Nicholas *et al.*, 1994). All three dyes fluoresce upon excitation with UV-light and offer a simple, rapid and inexpensive method to stain fungal cells. The significant background staining can be reduced by counterstaining with haemalum and eosin or Evans blue (Wachsmuth, 1988). Collagen, elastin, keratin fluoresce strongly. Even bacteria can be lightly stained with calcofluor, whereas Uvitex 2B shows slightly greater selectivity (Wachsmuth, 1988). Calcofluor white is believed to bind to $\beta$-1,4 glucans such as chitin, chitosan and cellulose (Wood and Fulcher, 1983).

Higher specificity can be achieved by the use of protein-based stains, especially using the fluorescent antibody technique developed by Coons *et al.* (1942; 1950). Antibodies specific to fungal strains have been used previously (Kaplan and Kaufman, 1961). Other than antibodies, lectins (including carbohydrate-binding domains) and enzymes have been used for revealing fungal infections (Benjaminson, 1969; Chamberland *et al.*, 1985; Laine and Lo, 2000). Proteins with high binding specificity can be labeled with fluorescent dye for fluorescence microscopy or with electron-dense substances like ferritin (Benjaminson *et al.*, 1966) or colloidal gold for electron microscopy (Chamberland *et al.*, 1985). Later fluorescein was conjugated to wheat germ

91
agglutinin to localize chitin in ascomycetous and basidomycetous yeasts (Simmons, 1989). The size of colloidal gold prevents this method from detecting chitin present in the more restricted inner cell wall layers (Molano et al., 1980). In *S. cerevisiae*, chitin is concentrated at the bud scar region (Molano et al., 1980). Polysaccharolytic enzymes conjugated with fluorescent dyes (e.g. fluorescein, rhodamine) have been used as specific stains (fluorescent enzyme technique) for the detection of polysaccharides: e.g. cellulase for staining cellulose (Seibert et al., 1978), chitinase (Benjaminson, 1969; Laine and Lo, 2000), lysozyme (Gould et al., 1963) and β-glucosidase for gram-positive and negative bacteria (Pital et al., 1967). Interestingly, in the time frame of the assay these enzymes should ideally be catalytically inactive (Laine and Lo, 2000) since cell lysis prevents a lasting stain, however, a low turnover number is characteristic of cell wall lytic enzymes.

8.2 Methods and Materials

A colony of *Saccharomyces cerevisiae* strain GRF180 was inoculated into 3 mL of yeast extract/peptone (YEP) medium containing 8 % glucose and grown at 30°C for 18 h. Yeast cells were pelleted by centrifugation, washed with PBS and resuspended to high turbidity in PBS. As negative control, suspension of *E. coli* (strain DH5α) and *Micrococcus lysodeikticus* (Sigma) were also prepared. To detect chitin, 30 µL of washed cells were mixed with ChBD-GFP (2 µM), calcofluor white (Sigma), 2 µg/mL of wheat-germ agglutinin-FITC (Sigma) or no fluorophore. All reactions were incubated at 25 °C for 1 h on a rocking platform. Following incubation, cells were pelleted, washed twice with PBS and resuspended in 10 % glycerol. A drop of suspended cells was applied to a glass slide and overlaid with a coverslip. Human tissue sections containing fungi (*Cryptococci, Histoplasmosis, Blastomycosis, Aspergillus* or *Candida albicans*) were obtained as FungiTissue-trol control slides from Sigma. To demonstrate that any labeling was due to the chitin-binding domain, polyhistidine tagged GFPuv was included in a control
experiment. Staining patterns were observed with fluorescence microscopy on a Microphot microscope (Nikon Instruments Inc., Melville NY) and images were taken with a SPOT RT digital camera (Diagnostic Instruments Inc., Sterling Heights, MI) at the Socolofsky Microscopy Center at Louisiana State University.

8.3 Results and Discussion

The specificity of ChBD-GFP hybrid protein to bind chitin present in organisms in situ was tested against *S. cerevisiae* cells and sections of human tissue with various fungal infections. The fungal labeling pattern was compared with results from common fluorescent chitin labeling methods involving calcofluor white and FITC-conjugated wheat germ agglutinin (WGA-FITC). Microscopic images of the specimen labeled by the three reagents are shown in Figure 38 to Figure 43.

![Figure 38: Localization of cell wall chitin in *S. cerevisiae*. Cells incubated with calcofluor white (80× magnification) fluorescent (A) and differential interference contrast (DIC) image (B). DIC images are shown for comparison.](image-url)
In *S. cerevisiae*, the four structural components of the cell wall, β-1,3 glucan, chitin, mannoprotein and β-1,6 glucan are to some extent crosslinked by the latter (Kollar *et al.*, 1997). Chitin is largely confined to the formation of buds and is retained in bud scars and septa (90 % total) with scattered occurrence in the rest of the cell wall (Molano *et al.*, 1980). WGA-FITC and ChBD-GFP stained the chitin rich bud scars and the whole contour of the yeast cells, while calcofluor stained entire *S. cerevisiae* cells without differentiation.

Figure 39: *S. cerevisiae* incubated with FITC-wheat germ agglutinin (40×) Fluorescent (A) and differential interference contrast (DIC) image (B).

Figure 40: *S. cerevisiae* incubated with ChBD-GFP (80×). Fluorescent (A) and differential interference contrast (DIC) image (B)
In human tissue sections, all three reagents generally detected about the same structures, with significantly higher background staining for calcofluor and WGA-FITC. The results demonstrate the ability of ChBD-GFP as a chitin specific stain even in the context of fungal and yeast cell wall matrices. The specificity was further verified by the lack of labeling of non-chitin containing negative controls, *E. coli*, oral epithelial cells and *M. lysodeikticus* (Data not shown). ChBD-GFP appears to be well suited to identify specifically chitin *in situ* by fluorescence microscopy.

![Figure 41: Localization of fungal infection in tissue with calcofluor white (40×).](image)

The value of ChBD-GFP to identify human tissue infested with fungus would have to be investigated in a broader spectrum of clinical samples. Of great medical interest would be whether ChBD-GFP can also be used to detect fungal antigens in human plasma, urine and other complex biological fluids by flow cytometry, providing a fast diagnosis and limiting the use of biopsy.
Figure 42: Localization of fungal infection in tissue with FITC-wheat germ agglutinin (40×).

Figure 43: Localization of fungal infection in tissue with ChBD-GFP (40×).
CHAPTER 9

ZYMOGRAM WITH REMAZOL-BRILLIANT-BLUE LABELED MICROCOCCUS LYSODEIKTICUS CELLS FOR THE DETECTION OF LYSOZYMES

9.1 Introduction

The zymogram method is based on the separation of proteins by gel electrophoresis and a subsequent *in situ* enzymatic step, in which the proteins are allowed to renature and relevant enzymes are allowed to convert catalytically substrates either that have been incorporated into the separating gel, sprayed onto the gel or overlaid within an agarose gel. Zymogram assays have been developed for various enzymes (Gabriel and Gersten, 1992) and have been extensively used in the detection of lysozymes and other cell wall hydrolases in previous studies (Audy *et al.*, 1989; Potvin *et al.*, 1988; Leclerc and Asselin, 1989).

Generally, inactivated *Micrococcus lysodeikticus* (*M. luteus* is a synonym) cells are trapped in the polymerized polyacrylamide gel as a substrate. Enzymes are separated by electrophoresis under denaturing conditions, which prevents premature interaction of enzyme with substrate. In the following refolding step, enzymes are allowed to renature and hydrolyze the embedded substrate, resulting in clearing zones in otherwise opaque gels. This method permits the simultaneous detection of enzyme activity and estimation of the molecular weight, aiding in the initial identification of the enzyme.

Inactivated *Micrococcus lysodeikticus* (ML) cells have been traditionally used as a substrate for measuring bacteriolytic activity with a spectrophotometer (Shugar, 1952) or in radial diffusion assays (lyso-plate assays) (Osserman and Lawlor, 1966). The common basic principle of these methods lies in the fact that the thick murein layer of the bacterial cells forms a turbid suspension, whose density will be reduced by presence of cell wall hydrolases. However,
sensitivity to experimental conditions (heterogeneity of the substrate preparation, buffer composition, etc.) compromises the reliability of the assay. For gel-based assays (lyso-plate, zymogram) the limitation lies mainly in the low contrast between clearing zones and background, often requiring additional time consuming staining/destaining steps, for example with methylene blue (Bernadsky et al., 1994; Watt and Clarke, 1994), to identify clearing zones.

The photometric suspension assay was improved by using Micrococcus lysodeikticus cells stained with the reactive blue dye Remazol-brilliant blue R (RBB) (Ito et al., 1992). RBB has been shown to bind to the hydroxyl groups of sugars (Stamm, 1963), which are present in peptidoglycan, the major constituent of bacterial cell walls. The hydrolysis of RBB-dye labeled ML-cells leads to the release of soluble blue products and clearing of a zone.

Here data is presented showing that whole cells of Micrococcus lysodeikticus prestained with RBB are an improved substrate for “Lyso”-zymograms, allowing a fast and real-time detection of lysozymes without the requirement for staining and destaining.

9.2 Materials and Methods

Materials required for gel preparation, lyophilized cells of Micrococcus lysodeikticus, chicken egg white lysozyme (14.3 kDa), bovine, serum albumin, Remazol Brilliant Blue R and all other reagents were purchased from Sigma Chemical Company (St. Louis, MO). Chemical reagents were of reagent grade or better.

9.2.1 Preparation of Ostrich lysozyme

Ostrich eggs were obtained from Three Fork Creek Ostrich Ranch (Pike Corinth, KY). Ostrich lysozyme (20.5 kDa), a goose-type lysozyme, was purified as described (Jolles et al., 1977).
9.2.2 Labeling of *Micrococcus lysodeikticus* Whole Cells with Remazol Brilliant Blue R

*M. lysodeikticus* whole cells were labeled with the vinyl-sulfone reactive dye Remazol brilliant blue R (RBB) according to the procedure for the synthesis of RBB-labeled starch (Rinderknecht *et al.*, 1967) with modifications suggested by Ito *et al.* (1992).

9.2.3 Polyacrylamide Gel Electrophoresis (PAGE)

PAGE (Laemmli, 1970) was performed in 12.5% (w/v) polyacrylamide gel (2.5 cm × 8.0 cm × 0.75 cm 5% stacking gel and a 5.5 cm × 8.0 cm × 0.75 cm separating gel), that contained 0.1% (w/v) blue ML cells. Polyacrylamide gels and buffers contained 0.1% (w/v) SDS. Samples were boiled for 2 minutes in 2 × sample buffer without reducing agent (62.5 mM Tris pH 6.8, 0.006% (w/v) bromophenol blue as tracking dye, 20% (v/v) glycerol, 2% (w/v) SDS).

Electrophoresis was performed using a Mini Protean II cell (Bio-Rad, Hercules CA) at a constant 120 V for two hours until the dye front approached the bottom of the resolving gel. The gel was washed twice with distilled water for 30 minutes to remove SDS, and incubated in a covered tray with 300 mL of renaturing buffer (50 mM NaPO₄, pH 7.0, containing 1% (v/v) Triton X-100) and gently shaken at 37°C.

Standard solutions of chicken egg white lysozyme (HEWL) were made in 50 mM potassium phosphate buffer (pH 7.0) by serial dilution.

A prestained protein marker (New England Biolabs, Beverly MA) was used as a molecular weight standard. The electrophoretic behavior of the standards is slightly affected by the prestaining and the apparent molecular weight is indicated: 175 kDa maltose-binding protein (MBP)-β-galactosidase, 83 kDa MBP-paramyosin, 62 kDa glutamic dehydrogenase, 47.5 kDa aldolase, 32.5 kDa triosephosphate isomerase, 25 kDa β-lactoglobulin A, 16.5 kDa lysozyme, 6.5 kDa aprotinin.
Digital images of the gels were taken with the NucleoVision 760 Imaging Workstation (Nucleotech Corporation, San Carlos, CA) using a UV transilluminator with a white light converter as light source. The presence of a cell wall hydrolase is detected by formation of transparent lytic zones in the otherwise opaque bluish gel. The areas of the lytic zones were measured using the ImageQuant software program [Molecular Dynamics (Amersham Biosciences), Sunnyvale CA].

9.3 Results and Discussion

Figure 44 shows zymograms of hen egg white and ostrich egg white lysozyme. Both enzymes were successfully renatured and both solubilized the blue ML substrate to produce clearing zones. As a negative control bovine serum albumin failed to yield clear spots, indicating that the clearing spots are not artifacts of electrophoresis or refolding.

Figure 44: Zymogram of hen-egg white lysozyme and ostrich lysozyme.
At high enzyme concentrations, clearing zones become visible after as little as one hour. As little as 0.1 unit of lysozyme (1.4 ng) could be detected after 12 h of incubation. Treatment of the protein samples with a reducing agent prior to electrophoresis had an expected inhibitory effect on renaturation. After electrophoresis, gels were incubated in refolding buffer to renature proteins. During renaturing, enzymes will diffuse in the gel and collide with embedded substrates to form immobilized enzyme-substrate complexes. Hydrolysis of blue-ML cells leads to release of soluble blue-labeled products, and the enzyme will be free to bind to additional substrate sites. Sites of hydrolysis form a readily observable clear reaction zone in a blue-stained gel.

The zymogram method can be considered as SDS-PAGE with a subsequent quasi-radial diffusion assay, however the diffusion is initiated from a band rather than from a point as in radial diffusion. By analogy with radial diffusion assays, the size of the clearing zone will, in part, be determined by the substrate affinity of the enzyme, its catalytic activity and its diffusion rate. In radial diffusion assays the diameter of the clearing zone is proportional to the logarithm of enzyme concentration applied, a method which has been applied to lysozyme quantification (Osserman and Lawlor, 1966). Figure 45 shows a zymogram of a dilution series of HEWL. The areas of clearing were measured and a linear relationship to the logarithm of enzyme concentration was found (Figure 46). An internal standard with known activity is required.

The presented zymogram method described here allows simultaneous detection and quantitation of multiple distinct enzymes in the same sample, whereas the radial diffusion assay only quantifies total enzyme activity.

The appearance and detection of clearing zones is improved by the use of dye-labeled ML cells, leading to increased contrast compared to the reduction of opacity with unlabeled cells. Visual monitoring of easily detectable clearing zones in real-time allows effortless and constant
Figure 45: Zymogram of a dilution series of hen-egg white lysozyme
following of the reaction process, without the requirement for additional time consuming
staining steps that could disrupt the catalytic reaction. This technique can be used as a fast and
convenient method for the discovery of new lysozymes. This approach may find use in
investigations of other cell wall hydrolases, which could include N-acetylmuramidases, N-
acetylglucosaminidase, N-acetylmuramyl-L-alanine amidases, endopeptidases and
transglycosylases (Ghuysen et al., 1968).
Figure 46: Plot of logarithm of applied lysozyme units vs. areas of clearing zones in zymogram.
CHAPTER 10
SUMMARY AND FUTURE DIRECTIONS

10.1 Summary
In this work, the binding affinity and substrate specificity of the chitin-binding domain (ChBD) of chitinase ChiA1 from *B. circulans* was studied by site-directed mutagenesis and the individual contributions of W687 and neighboring residues were determined by Langmuir-type adsorption isotherm analysis and affinity electrophoresis.

The research results suggest that ChBD_{ChiA1} and the binding domains listed in the sequence alignment (Figure 7) form a subfamily of the CBD_5_12 family based on the following reasoning: The members of the subfamily do not share the conserved stWWst motif found in other CBDs of the 5_12 family. Instead, they share a highly conserved tryptophan residue, which has been shown to be critical for chitin binding of ChBD_{ChiA1}. Ala substitution of Trp687 affected the binding most severely.

The importance of hydrophobic interactions in the binding was indicated by: (1) elution of the ChBD-GFP fusion protein from chitin with acetonitrile, (2) increased binding correlated with ionic strength of buffer and (3) maximal binding near the pI. Hydrophobic interactions involving a dehydration of the binding interface have been shown to be the driving forces for other carbohydrate-binding domains (Creagh *et al.*, 1996) and for protein adsorption processes in general (Haynes and Norde, 1994). Mutations of other residues adjacent to W687 showed little change for overall binding affinity to chitin other than the E688K/P689A mutation. In addition, adjacent residues seem to affect substrate specificity. Interestingly, mutation T682A shifts the specificity of the binding domain towards chitinous substrates in reference to the wild-type. Double-mutation E688K/P689A decreased avidity towards chitosan and chitin preparations with partial deacetylation (e.g. powdered chitin), but did not reduce the avidity towards β-chitin.
Overall similar observations have been described for other mutagenic studies of CBDs (Linder et al., 1995; Creagh et al., 1996; Simpson and Barras, 1999; McLean et al., 2000). It is also conceivable that the multivalency of the polysaccharide substrate limits discrimination between residues responsible for high-affinity and those responsible for modest contributions.

Compared to extensive studies on microbial cellulose-binding domains, knowledge about the binding mechanisms of bacterial chitin-binding domains is still very limited, aside from a mutagenic study on the Streptomyces chitin-binding protein ChB1 (Zeltins and Schrempf, 1997). In ChB1, tryptophan residues were shown to play an important role in the interaction with chitin, however, ChB1 does not share sequence homology with ChBD\textsubscript{ChiA1}. In the current study, Trp687 has been identified as a crucial binding site residue of ChBD\textsubscript{ChiA1}. Additional amino acids equally as significant could not be identified and the binding site should be considered more complex than has been predicted. From comparison with other carbohydrate-binding domains, it seems likely that at least one additional aromatic amino acid is involved in binding, since there has been no report of just a single stacking site.

The complexity and precision in the formation of the protein-carbohydrate complex can limit the success achievable with single point mutations of binding-site residues. Nevertheless, the current GFP-based approach in conjunction with affinity electrophoresis allows efficient screening of a library of ChBD mutants created by combinatorial methods. A combinatorial approach may prove helpful in elucidating the complex protein-sugar interaction in a way that can only be achieved by empirical means.

10.2 Future Directions

10.2.1 Surface Diffusion

Since the insoluble polysaccharides discussed here comprise a two-dimensional array of binding sites for the carbohydrate-binding domain, it is conceivable, that the protein does not
only form a binding equilibrium between one binding site and the bulk solution, but that the protein remains on or at least close to the surface of the polysaccharide and slides horizontally to the next binding site. In other words, once the protein is initially adsorbed on the substrate, the local concentration of alternative binding sites is very high compared to the bulk solution; therefore, a complete dissociation from the substrate should be limited. Theoretical aspects of diffusion of enzymes on insoluble substrates have been analyzed previously (Katchalski-Katzir et al., 1985) and have been applied for a β-amylase-starch system (Henis et al., 1988) and for cellulases of *C. fimi* (Cex and CenA) and their respective cellulose-binding domains on cellulose (Jervis et al., 1997). The theory is based on the reduction of dimensionality in the diffusion of proteins introduced by Adam and Delbrück (1968). Instead of motion in the three-dimensional solution and fairly limited numbers of protein-substrate collisions, transport of proteins to their substrates can be subdivided. First, the protein would reach a surface by diffusion or convection in the three-dimensional solution. Once on the surface interface, the protein would diffuse one- or two-dimensionally to its substrate. The function of the chitin-binding domain may lie in the initial binding of the modular protein to the chitin surface. The binding is specific but does not have to be very strong, allowing for the surface diffusion. The catalytic domains can therefore reach their substrates, i.e. exo- or endochitinase, deacetylases or even proteases, which degrade chitin-associated proteins.

To investigate further the complex binding behavior of ChBD, it would be important to understand the mobility of the binding domain on the polysaccharide by measuring surface diffusion. Similar to an approach used in membrane studies, the ChBD-GFP fusion protein could be used in the determination of the surface diffusion rate by fluorescence recovery after photobleaching (FRAP) analysis. FRAP has been previously used to describe the lateral mobility of β-amylase on starch (Henis et al., 1988).
10.2.2 Protein Immobilization

Carbohydrate-binding proteins and -domains (cellulose-binding domains, maltose-binding protein) have been found useful as affinity tags for protein purification in heterologous protein expression systems (Maina et al., 1988; Shpigel et al., 1998). ChBDChia has been used as a self-removable affinity tag in conjunction with a protein-splicing element (intein) for purification of target proteins (Chong et al., 1997). This approach circumvents the use of site-specific proteases to remove affinity tags from proteins. Recently, the inclusion of a C-terminal GFP in the self-cleavable intein-ChBD tag has been reported. As mentioned, the GFP acts as a reporter of fusion protein solubility and folding status, allowing optimization of protein expression and the prediction of protein yields. Finally, the fluorescence can indicate the removal efficiency of the self-splicing intein-ChBD-GFP tag (Zhang et al., 2001). The intein-GFP system can be useful in automation of protein expression processes.

The usage of CBDs in general does not have to be limited to purification tools. CBDs could serve as an anchor for solid-phase chemistry, i.e. involving protein modifications or synthesis, or in protein arrays. To combine multiple enzymatic steps, for example, sugar nucleotide regeneration enzymes have been co-immobilized on so called superbeads (Chen et al., 2001). The dual tagging (His)6 and ChBD) approach used for the ChBD-GFP presented here, allows for the orthogonal removal of the immobilized protein (imidazole, acetonitrile). Overall, immobilization techniques will be important in the creation of cell free systems, which by themselves have the advantages of improved sample separation, higher stability, reusability and improved kinetics of enzymatic reactions.

10.2.3 Fluorescence Polarization

Small molecules move and rotate freely and very rapidly in solution, fluorescent labels attached to the molecules are therefore free of directionality (isotropic). Upon binding to a larger
molecule with slower mobility, the small molecule can gain directionality (anisotropy) and is said to be polarized, resulting in higher rotational correlation time. The fluorescence polarization can be used to assess ChBD-GFP-polysaccharide complex formation.

10.2.4 Fluorescence-Resonance Energy Transfer (FRET)

When two fluorophores are in sufficient proximity and in appropriate relative orientation to each other, an excited fluorescent donor molecule can transfer energy to the second, lower energy, fluorescent acceptor molecule in non-radiative manner ( Förster, 1948; Stryer, 1978). The fluorescence emission of the acceptor molecule is enhanced, with attendant loss of emission of the donor molecule. Fluorescence-resonance energy transfer (FRET) can be used to study macromolecular interactions in vivo an in vitro systems. FRET can be measured in spectrofluorometers, laser scanning confocal microscopy and conventional fluorescence microscopy (Xia and Liu, 2001). Interaction of ChBD with chitin could be investigated by measuring FRET using the GFP fusion protein and fluorescently labeled chitin. Fluorescent chitin can be prepared from dilute fluorescently labeled chitosan (Tommeraas et al., 2001) with subsequent reacetylation of the free amino groups.

10.2.5 Directed Evolution Approach

Even with knowledge of the detailed structure of both the binding domain and the chitin substrate there is no reliable prediction about the binding interactions available and the effect of any modification has to be empirically tested. The practical limitation in conventional site-directed mutagenesis became clear in this study: Only a small number of sites can be analyzed, otherwise the conventional approach becomes impractical due to the combinatorial effects of possible mutations. A ‘directed evolution’ approach that mimics the natural design process (iterative cycles of generation of diversity, selection and recombination) has become increasingly popular and seems very promising in the study of binding domains.
• Generation of Variation

To introduce variation the following methods have been employed: Conventional site-directed mutagenesis, mutagenic PCR (error prone PCR in presence of Mn^{2+} ions, use of mutation inducible nucleosides or randomized primers) and elongation mutagenesis (Matsuura et al., 1999). The region of randomization can be limited and individual amino acid positions can be randomized by using oligonucleotides with NNG/C or NNG/T codons that are able to code for all 20 amino acids and one stop codon.

• Recombination

A powerful mechanism in natural evolution is recombination, which allows demonstration of the synergistic effect of two beneficial mutations for the phenotype. High recombination efficiency can be obtained by the in vitro DNA-shuffling method (Stemmer, 1994; Stemmer, 1994), which uses a set of parental genes, digested with DNAse I, to create a pool of short DNA fragments that are reassembled to full-length in the presence of DNA polymerase. Alternative methods include ‘staggered extension process’ (Zhao et al., 1998) and ‘incremental truncation’ (Ostermeier et al., 1999).

• Screening

A classical method for selection, which links phenotype to the encoding genotype is the phage display library. Combinatorial libraries of peptides and proteins are subcloned most commonly into an M13 bacteriophage vector encoding coat protein III of the filamentous phage f1 and the library members are displayed as fusion to the phage capsid (Petrenko and Smith, 1997). Bacterial cell surface display can be achieved by the fusion of the target protein to outer membrane proteins (e.g. ice nucleation protein from Pseudomonas syringae (Jung et al., 1998)), lipoprotein-outer membrane protein chimera, secreted proteins or subunits of cellular appendages
(reviewed in (Georgiou et al., 1997)). Ribosomal display (Hoffmuller et al., 1998) allows for selection in a cell-free system. Successful mutations can be enriched by affinity purification.

### 10.3 ChBD-GFP-Based Directed Evolution Approach

The previously described fusion of the chitin-binding domain to GFP provides an elegant system for a directed evolution study with the goal of identifying residues required for binding to chitin, altering the affinity and/or altering the specificity to other polysaccharides. After randomization of the ChBD-sequence with error-prone PCR and ligation of the randomized sequences into plasmids, transformed cells can be screened for fluorescence on plates with inducer. GFP fluorescence indicates successful transformation, protein production and correct folding of the target protein, as mentioned. Cells that carry undesirable mutations that cause frame-shifts, changes in global-folding and premature stops will lack GFP fluorescence and are omitted from the functional screening. To test the efficiency of randomization on the amino acid level, selected colonies can be screened by SDS-PAGE with urea by the previously mentioned ‘GFP-display’ method (Aoki et al., 2002). Individual colonies can be numbered and used for mini protein expressions. Cell lysates acquired by sonification will be analyzed simultaneously by previously described affinity electrophoresis in the presence of co-solidified chitin or other substrates. Changes in affinity can be detected by the highly sensitive GFP fluorescence detectable by fluorescence scanner even in cell lysates, requiring no additional purification steps. Plasmid DNA of samples that resulted in shifts will then be isolated and the DNA sequenced, informing about the mutations.
BIBLIOGRAPHY


## APPENDIX A

MULTIPLE SEQUENCE ALIGNMENT OF PFAM CHBD3 FAMILY
<table>
<thead>
<tr>
<th>Accession</th>
<th>Organism Name</th>
<th>Enzyme Name</th>
<th>Sequence (Peptide)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2D1X2/828-867</td>
<td>Pseudalteromonas sp. s9</td>
<td>Chitinase c</td>
<td>RRAKRAIDWYTR ...</td>
</tr>
<tr>
<td>O9ZN1.0/788-828</td>
<td>Clostridium perfringens</td>
<td>Chitinase a</td>
<td>YPTKDESNTDTR ...</td>
</tr>
<tr>
<td>O50976/786-827</td>
<td>Clostridium perfringens</td>
<td>Chitinase b (chitodextrinase)</td>
<td>YPTKDESNTDTR ...</td>
</tr>
<tr>
<td>GUN1_BACS4/357-397</td>
<td>Bacillus sp. (strain n-4 / jcm 9156)</td>
<td>Endoglucanase a</td>
<td>YEPEDENTDTR ...</td>
</tr>
<tr>
<td>GUN5_BACCA1/357-396</td>
<td>Bacillus agaradhaerens</td>
<td>Endoglucanase 5a</td>
<td>YEPEDENTDTR ...</td>
</tr>
<tr>
<td>Q59222/346-385</td>
<td>Bacillus sp.</td>
<td>Endo-beta-1,4-glucanase precursor</td>
<td>YEPEDENTDTR ...</td>
</tr>
<tr>
<td>GUN1_BACS4/445-484</td>
<td>Bacillus sp. (strain n-4 / jcm 9156)</td>
<td>Endoglucanase b</td>
<td>YEPEDENTDTR ...</td>
</tr>
<tr>
<td>GUN2_BACS4/366-405</td>
<td>Bacillus sp. (strain n-4 / jcm 9156)</td>
<td>Endoglucanase a</td>
<td>YEPEDENTDTR ...</td>
</tr>
<tr>
<td>Q9RQ17/1137-1177</td>
<td>Cellulomonas pachnodae</td>
<td>Endo-1,4-beta-xylanase</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9HR23/32-73</td>
<td>Halobacterium sp. (strain nrc-1)</td>
<td>Vng0818c</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9HR43/32-73</td>
<td>Halobacterium sp. (strain nrc-1)</td>
<td>Chitinase</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9HR23/31-72</td>
<td>Halobacterium sp. (strain nrc-1)</td>
<td>Vng0819c</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>O85501/547-589</td>
<td>Bacillus subtilis</td>
<td>Chitinase</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>P94339/549-593</td>
<td>Bacillus licheniformis</td>
<td>Chitinase (chitodextrinase)</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>O9C19/447-489</td>
<td>Lactococcus lactis (subsp. lactis)</td>
<td>Chitinase</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9ZJX2/779-818</td>
<td>Pseudalteromonas sp. s9</td>
<td>Chitinase c</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9ZN1/108-151</td>
<td>Aeromonas sp. 10a-24</td>
<td>Chitinase</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q59141/122-165</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q59145/392-435</td>
<td>Aeromonas sp</td>
<td>Chitinase ii precursor</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q59144/144-187</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor fragment</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q59143/120-163</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q59142/139-182</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Q9RC05/122-170</td>
<td>Doohwaniella chitinasiens</td>
<td>Chitinase ch167</td>
<td>YPQRTRSSWYTR ...</td>
</tr>
<tr>
<td>Accession</td>
<td>Organism</td>
<td>Chitinase Type</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------</td>
<td>--------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Q9WX9/142-185</td>
<td>Aeromonas sp. 10s-24</td>
<td>Chitinase (pyral 1 ort) precursor</td>
<td></td>
</tr>
<tr>
<td>O48373/151-197</td>
<td>Janninobacterium lividum</td>
<td>Chitinase precursor</td>
<td></td>
</tr>
<tr>
<td>Q9KTW/1445-484</td>
<td>Vibrio cholerae</td>
<td>Chitinase, putative</td>
<td></td>
</tr>
<tr>
<td>Q9ZIX3/436-475</td>
<td>Pseudomonas sp. s9</td>
<td>Chitinase b</td>
<td></td>
</tr>
<tr>
<td>O5015/33-72</td>
<td>Streptomyces griseus</td>
<td>Chitinase e</td>
<td></td>
</tr>
<tr>
<td>Q9RJF3/4-25</td>
<td>Streptomyces</td>
<td>Chitinase (fragment)</td>
<td></td>
</tr>
<tr>
<td>Q9RQQ3/767-807</td>
<td>Aeromonas hydrophila</td>
<td>Chitinase</td>
<td></td>
</tr>
<tr>
<td>Q9ZIX3/771-711</td>
<td>Aeromonas hydrophila</td>
<td>extracellular chitinase chiA</td>
<td></td>
</tr>
<tr>
<td>Q9P9Q8/770-810</td>
<td>Aeromonas hydrophila</td>
<td>Chitinase 92</td>
<td></td>
</tr>
<tr>
<td>O43919/770-810</td>
<td>Aeromonas hydrophila</td>
<td>Chitinase a precursor</td>
<td></td>
</tr>
<tr>
<td>Q9RQX/1080-1122</td>
<td>Cellulomonas pachonodi</td>
<td>Endo-1,4-beta-xylanase</td>
<td></td>
</tr>
<tr>
<td>CH1A_FECOL/1459-508</td>
<td>Escherichia coli</td>
<td>probable bifunctional chitinase/lysozyme precursor</td>
<td></td>
</tr>
<tr>
<td>Q9ZIX3/483-522</td>
<td>Pseudomonas sp. s9</td>
<td>Chitinase b</td>
<td></td>
</tr>
<tr>
<td>Q9RQQ3/814-853</td>
<td>Aeromonas hydrophila</td>
<td>Chitinase</td>
<td></td>
</tr>
<tr>
<td>Q9LSD/818-857</td>
<td>Aeromonas hydrophila</td>
<td>extracellular chitinase chiA</td>
<td></td>
</tr>
<tr>
<td>Q9P9Q8/817-856</td>
<td>Aeromonas hydrophila</td>
<td>Chitinase 92</td>
<td></td>
</tr>
<tr>
<td>Q43919/817-856</td>
<td>Aeromonas caviae</td>
<td>Chitinase a precursor</td>
<td></td>
</tr>
<tr>
<td>Q9KT3/1048-1082</td>
<td>Vibrio cholerae</td>
<td>Chitinase, putative</td>
<td></td>
</tr>
<tr>
<td>O5263/635-699</td>
<td>Sodalis glossinidius</td>
<td>Exochitinase</td>
<td></td>
</tr>
<tr>
<td>QN9XD6/35-81</td>
<td>Burkholderiada gladioli</td>
<td>Chitosanase a</td>
<td></td>
</tr>
<tr>
<td>Q9RMB8/43-86</td>
<td>Arthrobacter sp</td>
<td>Chitinase precursor</td>
<td></td>
</tr>
<tr>
<td>Q5994/494-530</td>
<td>Aeromonas sp</td>
<td>Chitinase ii precursor</td>
<td></td>
</tr>
<tr>
<td>Q9RCG5/33-71</td>
<td>Deodhwiella chitinasiagens</td>
<td>Chitinase chi67</td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Species</td>
<td>Protein Type</td>
<td>Sequence Information</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------</td>
<td>----------------------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>O9N422:247-281</td>
<td>Streptomyces coelicolor</td>
<td>putative dioxygenase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O91115:436-470</td>
<td>Pseudomonas aeruginosa</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O924IP2:467-75</td>
<td>Streptomyces olivaceovirdis</td>
<td>chitinase precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9S6T9:467-75</td>
<td>Streptomyces coelicolor</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9239M6:467-75</td>
<td>Streptomyces coelicolor</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O971V7:44-83</td>
<td>Clostridium acetobutylicum</td>
<td>hypothetical protein cac1533</td>
<td>sequence details</td>
</tr>
<tr>
<td>CHID_VIBFU:998-1030</td>
<td>Vibrio furnissii</td>
<td>Chitodextrinase precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9KIP3:1001-1033</td>
<td>Vibrio cholerae</td>
<td>Chitodextrinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9UWR7:54-122</td>
<td>Pyrococcus kodakarenseis</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O94256:453-492</td>
<td>Serratia marcescens</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9529T2:475-515</td>
<td>Serratia marcescens</td>
<td>Chitinase b precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>CHIB_SERMA:453-492</td>
<td>Serratia marcescens</td>
<td>Chitinase b precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9239M6:385-419</td>
<td>Vibrio cholerae</td>
<td>hypothetical protein ve1280</td>
<td>sequence details</td>
</tr>
<tr>
<td>O99P71:3-415</td>
<td>Vibrio alginolyticus</td>
<td>Deacetylase da1 precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>GUNZ_EWICH:386-419</td>
<td>Erwinia chrysanthemi</td>
<td>Endoglucanase z precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9REW0:386-419</td>
<td>Erwinia chrysanthemi</td>
<td>Endo-1,4-beta-glucanase precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O96699:442-487</td>
<td>Alteromonas haloplanktos</td>
<td>Cellulase precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9Z2439:32-61</td>
<td>Xanthomonas sp. akj</td>
<td>Chitinase a precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O92667:47-76</td>
<td>Xanthomonas maltophilia</td>
<td>Chitinase a</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9Z27F2:276-312</td>
<td>Vibrio cholerae</td>
<td>Endochitinase chia precursor</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9KQ06:276-312</td>
<td>Vibrio cholerae</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>P95675:675-708</td>
<td>Vibrio harveyi</td>
<td>Chitinase a</td>
<td>sequence details</td>
</tr>
<tr>
<td>O9REW0:675-708</td>
<td>Vibrio sp. fl7</td>
<td>Chitinase</td>
<td>sequence details</td>
</tr>
<tr>
<td>Accession</td>
<td>Organism</td>
<td>Description</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>-----------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Q9WX923-59</td>
<td>Aeromonas sp. 10s-24</td>
<td>Chitinase (pra1 orf) precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9R4E72-19</td>
<td>Aeromonas</td>
<td>Chitinase vi (fragment)</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9R4F52-18</td>
<td>Aeromonas</td>
<td>Chitinase viii (fragment)</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q5914328-64</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9R4G50-1-19</td>
<td>Aeromonas</td>
<td>Chitinase iii (fragment)</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q5914125-61</td>
<td>Aeromonas</td>
<td>Chitinase precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q5914226-62</td>
<td>Aeromonas sp</td>
<td>Chitinase precursor (fragment)</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q5914429-65</td>
<td>Aeromonas</td>
<td>Chitinase precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q4837352-88</td>
<td>Janthinobacterium lividum</td>
<td>Chitinase precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9ZNS226-62</td>
<td>Aeromonas sp. 10s-24</td>
<td>Chitinase</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9FCC87667-701</td>
<td>Streptomyces coelicolor</td>
<td>putative bifunctional protein (fragment)</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9WX34433-667</td>
<td>Serratia marcescens</td>
<td>Chitinase c1</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q54442333-467</td>
<td>Serratia marcescens</td>
<td>Chitinase</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q999Y11332-396</td>
<td>Vibrio alginolyticus</td>
<td>Deacetylase da1 precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q999K66336-379</td>
<td>Vibrio cholerae</td>
<td>hypothetical protein vc1280</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9X9586412-455</td>
<td>Streptomyces coelicolor</td>
<td>putative serine protease</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9X8994653-687</td>
<td>Kurthia zopfi</td>
<td>Chitinase</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q999H3653-687</td>
<td>Bacillus circulans</td>
<td>Chitinase</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>CH1L_BACCI653-687</td>
<td>Bacillus circulans</td>
<td>Chitinase a1 precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>CHID_BACCI631-65</td>
<td>Bacillus circulans</td>
<td>Chitinase d precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>PRTC_STRGR411-445</td>
<td>Streptomyces griseus</td>
<td>Serine protease c precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
<tr>
<td>Q9L015417-451</td>
<td>Streptomyces coelicolor</td>
<td>putative serine protease precursor</td>
<td>YPRQGEETGTA...TFRKVSNDQTVLQ-RHTAYT...-</td>
</tr>
</tbody>
</table>
Figure 47: Excitation spectrum of GFPuv, ChBD-GFP and mutated ChBD-GFP fusion proteins (394 nm excitation wavelength, emission scanned from 420-550 nm).
Figure 48: Emission spectrum of GFPuv, ChBD-GFP and mutated ChBD-GFP fusion proteins (511 nm emission wavelength, excitation scanned from 350-485 nm).
Markus Hardt was born in Langenfeld (Rheinland), Germany. He graduated from Marianum in Leverkusen-Opladen (Germany) with an Abitur (High School Diploma) in May 1993. Markus then attended Eberhard-Karls Universität in Tübingen (Germany) where he obtained a Vordiplom degree in biochemistry in May 1996. From August 1996 until May 1997, he participated as a graduate student in an exchange program with Louisiana State University in Baton Rouge, Louisiana. After returning to his graduate studies at Eberhard-Karls Universität in Tübingen for one semester, he entered the Biochemistry Department at Louisiana State University in January 1998 and began working toward the doctor of philosophy degree in biochemistry under the direction of Dr. Roger A. Laine.