

Abstract

Title of Dissertation: COMPLEX POLYSACCHARIDE
DEGRADATION BY *MICROBULBIFER*
DEGRADANS STRAIN 2-40: STUDIES OF
THE CHITINOLYTIC SYSTEM AND
CARBOHYDRASE ARCHITECTURE

Michael B. Howard, Doctor of Philosophy, 2004

Dissertation Directed By: Professor Steven W. Hutcheson, Department of
Cell Biology and Molecular Genetics

Microbulbifer degradans strain 2-40 is a Gram negative marine bacterium that is able to depolymerize and metabolize a wide variety of complex polysaccharides, including chitin. Chitin is the second most abundant biopolymer in nature and is a widely available nutrient in many environments. The chitinolytic system of *M. degradans* consists of three chitin depolymerases (ChiA, ChiB, and ChiC), three N-acetylglucosaminidases (HexA, HexB, and HexC), a chitodextrinase (CdxA), a chitin-binding protein (CbpA), and a suite of enzymes involved in the transport and metabolism of GlcNAc and GlcNAc₂. ChiB, the largest eubacterial chitinase described, includes two complete Glycoside Hydrolase family 18 catalytic domains; one is exolytic and the other is endolytic. These catalytic domains share similar reaction optima, temperature, pH, and metal ion sensitivities, and are shown to function synergistically in the depolymerization of crystalline chitin. ChiA, ChiB, and 44 other *M. degradans* carbohydrases contain polyserine linker regions. These linkers are composed predominantly of serine (79%), have an average length of 39

residues, and are encoded by all six serine codons without any obvious bias or pattern. Polyserine domains are found only between functional groups (e.g., catalytic, binding, or anchoring domains) and are only found in putative secreted, carbohydrate depolymerases. The sequence of the *M. degradans* genome was determined during the course of this work and permitted the first genomic analysis of a chitinolytic organism to be performed. Further, these experiments and subsequent analyses have lead to a greater understanding of how *M. degradans* is able to metabolize such a diverse collection of biopolymers.

COMPLEX POLYSACCHARIDE DEGRADATION BY *MICROBULBIFER*
DEGRADANS STRAIN 2-40: STUDIES OF THE CHITINOLYTIC SYSTEM AND
CARBOHYDRASE ARCHITECTURE.

By

Michael B. Howard

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

Professor Steven W. Hutcheson, Chair

Professor Ronald M. Weiner

Associate Professor Richard C. Stewart

Assistant Professor Jocelyne DiRuggiero

Professor William E. Bentley (Dean's Representative)

© Copyright by
Michael B. Howard
2004

Doctor of Philosophy

There are some things which cannot be learned quickly,
and time, which is all we have,
must be paid heavily for their acquiring.

They are the very simplest things,
and because it takes a man's life to know them
the little new that each man gets from life is very costly
and the only heritage he has to leave.

-Ernest Hemingway

Acknowledgements

I would like to express my sincerest thanks to my wife, for without her limitless patience and willingness to tolerate long hours and an occasionally irritable husband, this work could not have been finished.

Table of Contents

Acknowledgements.....	iii
List of Tables	vi
List of Figures.....	vii
Abbreviations.....	x
Chapter 1. General Introduction	1
Characteristics and Physiology of <i>Microbulbifer degradans</i> strain 2-40.....	1
The Structure and Uses of Chitin and its Derivatives.....	3
Microbial Chitinolytic Systems	7
Statement of Purpose	16
Chapter 2: The Chitinolytic System of <i>Microbulbifer degradans</i>	17
Introduction.....	17
Materials and Methods.....	18
Results.....	23
Discussion.....	43
Chapter 3: Chitinase B of <i>Microbulbifer degradans</i> Contains Two Catalytic Sites with Different Activities on Chitin	52
Introduction.....	52
Materials and Methods.....	58
Results.....	61
Discussion	81
Chapter 4: Polyserine Linkers Among <i>Microbulbifer degradans</i> Carbohydrases and Other Prokaryotic Proteins.....	88
Introduction.....	88
Materials and Methods.....	89
Results.....	90
Discussion	102
Chapter 5: Other <i>Microbulbifer</i> species: Metabolic Profiles, Cloning of Select Genes, and Partial Enzyme Characterization.....	109
Introduction.....	109
Materials and Methods.....	111
Results.....	115
Conclusions.....	126
Chapter 6: General Conclusions and Future Directions	132

Appendix I: Growth of <i>Salmonella enterica</i> on Germinating Alfalfa Sprouts	137
Abstract	137
Introduction	139
Materials and Methods	141
Results	144
Discussion	160
Appendix II: Methods for the Detection and Characterization of Chitinases and Other Chitin Modifying Enzymes	171
Introduction	171
Assays to detect chitinolytic activity of culturable microorganisms	173
Degenerate PCR primers and probes for environmental DNA samples, culturable, and nonculturable microorganisms	182
Detection of specific chitooligosaccharides	183
Detection of chitin-binding proteins and chitin-binding assays	191
Conclusions	192
Bibliography	194

List of Tables

Table 2-1	Relevant physical and biochemical characteristics of the components of the <i>M. degradans</i> chitinolytic system.....	45
Table 3-1	Activity of polypeptides containing ChiB catalytic domains on different MUF analogs.....	69
Table 3-2	Activity of polypeptides containing ChiB catalytic domains on native chitin.....	80
Table 4-1	<i>M. degradans</i> proteins that contain polyserine linkers.....	91
Table 4-2	Additional PSL proteins among prokaryotes.....	100
Table 5-1	Complex polysaccharides metabolized by described <i>Microbulbifer</i> spp. and <i>T. turnerae</i>	116
Table 5-2	Activity of <i>M. hydrolyticus</i> chitinases on MUF-diNAG and MUF-triNAG.....	128
Table A2-1	Retention time of chitooligosaccharides using the Chang HPLC Method.....	189

List of Figures

Figure 1-1	Structures of the components of chitin.....	5
Figure 1-2	The GlcNAc metabolic pathway of <i>E. coli</i>	11
Figure 2-1	Zymogram of culture supernatants of chitin-grown <i>M. degradans</i>	24
Figure 2-2	Extracellular components of the <i>M. degradans</i> chitinolytic system.....	28
Figure 2-3	Periplasmic and cytoplasmic components of the <i>M. degradans</i> chitinolytic system.....	33
Figure 2-4	Comparison of the <i>E. coli</i> and <i>M. degradans</i> GlcNAc metabolism clusters.....	37
Figure 2-5	Proteases are secreted in response to chitin.....	41
Figure 2-6	A model of chitin degradation by <i>M. degradans</i>	47
Figure 3-1	Endo- vs. Exo-chitinolytic activity.....	54
Figure 3-2	Comparison of the domain architectures of <i>M. degradans</i> ChiB and <i>T. kodakaraensis</i> Tk-ChiA.....	56
Figure 3-3	Expression of ChiB is dependent on carbon source.....	62
Figure 3-4	Each GH18 domain of ChiB can independently degrade chitin.....	66
Figure 3-5	pH optimum of GH18 _N and GH18 _C	71
Figure 3-6	Temperature optimum of GH18 _N and GH18 _C	73
Figure 3-7	Effect of different ionic conditions on the activity of GH18 _N and GH18 _C	75

Figure 3-8	Reaction products released from native chitin by the activity of each catalytic domain.....	77
Figure 3-9	Model of chitin depolymerization by ChiB.....	86
Figure 4-1	Duplication of two <i>M. degradans</i> genes that encode polyserine linkers.....	98
Figure 4-2	Polyserine domains are flexible and may increase the range of motion of an enzyme.....	107
Figure 5-1	Architecture of <i>Microbulbifer hydrolyticus</i> enzymes identified in this work.....	118
Figure 5-2	ChiA, ChiB, and ChiC degrade chitin.....	124
Figure A1-1	Growth of <i>S. enterica</i> serovar Cubana 98E01362SH2 on germinating alfalfa sprouts.....	145
Figure A1-2	Correlation between <i>Salmonella</i> populations on alfalfa sprouts and in waste irrigation water.....	148
Figure A1-3	Populations of <i>S. enterica</i> serovar Cubana 98E01362SH2 in directly sampled waste irrigation water.....	150
Figure A1-4	Growth of <i>S. enterica</i> serovar Cubana 98E01362SH2 on alfalfa seeds and sprouts during different points in the germination process.....	153
Figure A1-5	Level of reducing sugars in spent irrigation water over time.....	155
Figure A1-6	Growth of <i>S. enterica</i> serovar Cubana 98E01362SH2 in LB, M63 minimal medium, and sterilized spent irrigation water.....	157

Figure A1-7	Growth of various serovars of <i>S. enterica</i> on alfalfa sprouts.....	161
Figure A1-8	Levels of detected populations of various serovars of <i>S. enterica</i> in spent irrigation water.....	163
Figure A1-9	Growth of <i>S.e. Typhimurium</i> wt and various mutants with impaired pathogenicity on germinating alfalfa sprouts.....	165
Figure A2-1	2-aminobenzoic acid labeled chitooligosaccharides.....	189

Abbreviations

aa	amino acid
AmbAll	pectate lyase catalytic domain
AZCL	azureine cross-linked
CBD	carbohydrate binding domain
CBM	carbohydrate binding module
Chbs	chitobiose
Chi	chitin
COS	chitooligosaccharides
CP	complex polysaccharide
diNAG	chitobiose
DNSA	3,5-dinitrosalicylic acid
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol bis(2-aminoethyl ether)-N,N,N',N'-tetraacetic acid
FN3	fibronectin type 3 domain
fru-6-P	fructose-6-phosphate
GC	guanosine + cytosine content
GH	glycoside hydrolase family catalytic domain
GlcNAc	N-acetyl-D-glucosamine
GlcNAc-6-P	N-acetyl-D-glucosamine-6-phosphate
Gln	glucosamine
Gln-6-P	glucosamine-6-phosphate
Glu	glucose
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	histidine
ICP	insoluble complex polysaccharide
IM	inner membrane
IPTG	Isopropyl-beta-D-thiogalactopyranoside
JGI	Joint Genome Institute
KB	kilobases
kDa	kilo-Daltons
LPB	Lipoprotein box (acylation site)
MUF	4'-methylumbelliferone
MW	molecular weight
NAG	N-acetyl-D-glucosamine
NTA	nitrilotriacetic acid
OM	outer membrane
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PKD	polycystic kidney disease domain
PSL	polyserine linker
PTS	phosphotransferase system
RepDom	repeat domain

SDS	sodium dodecylsulfate
SSS	polyserine linker
triNAG	chitotriose
Tris	tris (hydroxymethyl) aminomethane

Chapter 1. General Introduction

Characteristics and Physiology of Microbulbifer degradans strain 2-40

Microbulbifer degradans strain 2-40 was originally isolated from decaying *Spartina alterniflora*, a salt marsh cord grass, in Mathews County, Virginia in the mid-1980s (2). *Ab initio*, its agarolytic ability was easily observed and set it apart from other isolates. Subsequent analyses revealed that *M. degradans* is able to degrade agar, agarose, chitin, cellulose, pectin, pullulan, xylan, starch, sodium polypectate, alginate, and laminarin (2, 152). The ability of a single bacterium to degrade this array of complex polysaccharides has not been previously reported and appears to be a distinguishing feature of this strain.

General phenotypic characterization has revealed that *M. degradans* is a Gram negative bacterium that forms pleiomorphic rods and is a strict aerobe when grown on minimal media (41). It has an absolute requirement for sodium ions and grows optimally in 2.3% sea salt at 30°C and pH 7.5, though growth is observed from 4°C to 37°C and at pH 4.5 to 10. It is oxidase and catalase positive, beta hemolytic, and is motile by a single polar flagellum. The chromosome has a G+C content of ~46% and is roughly 5.1 MB in size; *M. degradans* does not contain plasmids or megaplasms (N.Ekberg, unpublished observations). The major fatty acids of *M. degradans* are 16:0 (37%), 14:0 (15%), and 12:1 3-OH (6%). Colonies are white to cream colored and sink into agar-based solid media due to agarolytic activity. Eumelanin is produced in late stationary/death phase in both solid and liquid cultures. Broth cultures incubated without shaking do not form a pellicle (M.Howard, unpublished

observations). Numerous attempts to transform *M. degradans*, identify a stably maintained plasmid or suicide plasmid, utilize known transposon mutagenesis strategies (both mating and phage-based), or produce gene knockouts have not been successful (M. Howard and N. Ekborg, unpublished observations).

The phylogeny of *M. degradans* has not been completely established. It was originally suggested that it be included in the genus *Alteromonas* (41). However, 16S rDNA sequence analysis has revealed that it is ~93% similar to *Teridinibacter turnerae*, a dinitrogen fixing, cellulolytic endosymbiont of shipworms (Teredinidae: Bivalvia) (28). Additionally, it shares ~92% similarity to *Microbulbifer hydrolyticus*, the Type strain of the genus *Microbulbifer*. *M. degradans* shares some similarity to *M. hydrolyticus* in terms of cell biology; each forms bleb like structures in the presence of complex polysaccharides (42). However, *Microbulbifer* spp. including *M. hydrolyticus*, *M. salipaludis*, *M. elongatum*, *M. arenaceous*, and *M. maritmus* are all less than 92% similar to *M. degradans* at the 16S level and have a G+C content above 51% (102, 136, 154, 155). The fatty acid profile of *M. degradans* is significantly different from *M. hydrolyticus* and other *Microbulbifer* spp. The taxonomy of this organism is therefore somewhat uncertain; the profile compared to the genus *Microbulbifer* is similar in some aspects, but the G+C content of the chromosomes differs by more than 5%. Conversely, *Teridinibacter* spp. are endosymbionts (a significant ecological feature) and are able to fix dinitrogen (a significant metabolic feature). Further, this genus is designated by a published name that suggests all members must be endosymbionts of shipworms, a misleading notion considering that fewer than ten strains were used to construct the genus (28). For the

purposes of this work, *M. degradans* will be considered as a proposed member of the genus *Microbulbifer* with the designation *Microbulbifer degradans* (in reference to the battery of complex polysaccharides it can degrade). This nomenclature is used by GenBank in reference to the shotgun genome sequence and has been used in numerous publications by our group.

Because of its extraordinary degradative abilities, *M. degradans* strain 2-40 was selected for sequencing by the United States Department of Energy Joint Genome Institute. This sequencing effort has provided a draft genome sequence and partial annotation; annotated portions of the genome have a minimum of 9X coverage. Sequence analysis has revealed more than 130 putative proteins involved in polysaccharide depolymerization, a large number of proteases, sugar transporters, carbohydrate binding proteins, and some genetic duplication in several of the complex polysaccharide degrading systems. Further, the genome sequence has confirmed the G+C content and genome size reported earlier. The large amount of data provided by the genome sequence has allowed a detailed and in depth analysis of the complex polysaccharide degrading systems of *M. degradans* to be undertaken; the chitinolytic system is of particular interest to this work.

The Structure and Uses of Chitin and its Derivatives

Chitin is a ubiquitous polysaccharide composed of repeating units of β -1,4 linked N-acetyl-D-glucosamine and occasional D-glucosamine residues (92) (Figure 1-1). It has been estimated that up to ten gigatons of chitin are synthesized annually, thus chitin is an abundant and inexpensive renewable resource. Chitin is set apart from other common polysaccharides (such as cellulose, xylan, pectin, and starch) by the

fact that it contains nitrogen. Chitin is usually at least 90% acylated, is completely insoluble in water, and can be modified to suit numerous industrial and medical applications.

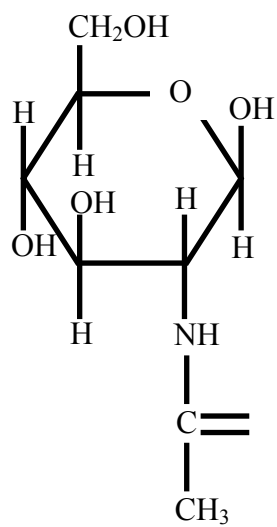
Chitin exists in a variety of crystalline states in the environment (92). Alpha chitin consists of antiparallel chains which permit extensive hydrogen bonding between acetamido groups. This results in a very durable, inflexible polymer. Alpha chitin is found in the spines of *Sagitta*, arthropod cuticles, and in the calyces of hydrozoa. Beta chitin consists of parallel chains of the polysaccharide, allowing for a more flexible compound. Beta chitin is found in squid pens, diatoms, and the shells of mollusks. Gamma chitin, by far the most abundant form of the polymer, is a mixture of α and β chitin. This form can be found in insect exoskeletons, the cell walls of fungi, the cocoons of insects, diplopods, arachnids, and in a variety of mollusks. Regardless of the crystalline structure of chitin, it is almost always complexed with salts, protein, and other carbohydrates.

The medical and industrial uses of chitin, chitin derivatives, and chitinases are expanding and have recently been reviewed in detail (93, 128). In industrial and applied fields, chitin and chitosan (partially deacylated chitin) are used in such areas as bioremediation of heavy metal contamination, water purification, paper and textile production, as additives in animal feed, and as a component of consumer products such as cosmetics (45, 99, 104, 151). Uses for chitin-derived products in the medical field include drug and vaccine delivery systems, wound and burn treatments, blood cholesterol lowering agents, anti-clotting agents, antibacterial compounds, and

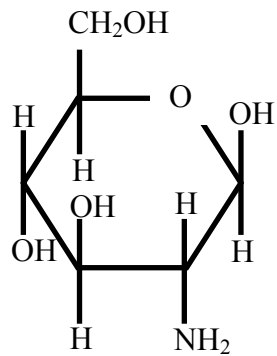
Figure 1-1

Structures of the components of chitin

Chitin is a polymer of N-acetylglucosamine and, to a lesser extent, glucosamine. The total glucosamine content is usually less than 30%. The linkages between the monomers are in a β conformation and link the molecules between the 1 carbon and 4 carbon. The acetamido group on the 2 carbon of N-acetylglucosamine is responsible for a large amount of the hydrogen bonding between strands and determines the crystalline properties of the polymer. The length of chitin strands is often many hundreds of residues.



N-acetylglucosamine (~70%)



Glucosamine (~30%)

enhanced dissolution of some drugs, such as ibuprofen (27, 57, 58, 93, 128).

Chitinases have been shown to act as biocontrol and anti-fungal agents, as well as having anti-biofouling applications (36). Chitinases can be used to create chitooligosaccharides of specific lengths, which have a variety of roles in the industrial and research applications mentioned above. Chitin-binding modules have been incorporated into protein expression vectors as fusion tags to facilitate the purification of recombinant proteins (3).

Microbial Chitinolytic Systems

Chitinases are produced by prokaryotes and eukaryotes for a number of purposes. Organisms that utilize chitin for form or defense modify their own structural chitin by the activity of chitinases. Plants and bacteria utilize these enzymes for defense against chitin-containing predators or parasites such as fungi, nematodes, and insect larva. Of primary interest for the purpose of this work, and by far the most common use of chitinolytic systems in the biome, is the acquisition of nutrients by the action of these enzymes. It has been demonstrated that many bacterial species are able to utilize chitin as a sole source of carbon and nitrogen.

Many bacteria, particularly from the marine environment, have developed highly efficient systems for the depolymerization, transport, and metabolism of chitin and its derivatives. Chitinovorous bacteria have been identified from marine sediments, soil samples, and desert sands. Within the oceans there is a constant “rain of chitin” as organisms that contain this polysaccharide die or molt (157). This results in an almost limitless supply of chitin for marine bacteria. Surprisingly, almost no free chitin can be found in marine sediments, thus demonstrating the

efficacy of these systems (31). Further, the transport and metabolism of N-acetylglucosamine is a common event in the oceans and the systems and pathways used to metabolize this substrate have also been described in detail (110, 120).

The bacterial enzymes involved in chitin degradation and metabolism are found in the extracellular environment, periplasm, and cytoplasm (31). Extracellular components include chitin depolymerases, chitin-binding proteins with no observed catalytic activity, and proteases (88, 138). Chitin depolymerases can act as endo- or exo-enzymes; endochitinases cleave the polymer at random glycosidic linkages, whereas exochitinases act at the non-reducing end of free chitin chains to release chitobiose (135). Secreted proteases have been identified that are upregulated by the presence of chitin. These enzymes appear to degrade the proteinaceous components of chitin to facilitate access to the polysaccharide polymer. A protease with a chitin binding domain has been identified from *Alteromonas* sp. strain O-7, though the lack of such a domain among other secreted proteases does not preclude their activity on chitin (89).

Non-catalytic chitin-binding proteins are often components of chitinolytic systems, though their function unclear (35, 124). There are some reports of these proteins interacting with the outer-membrane of gram negative organisms, in theory to anchor the cell to a nutrient source (114). It has also been proposed that the binding domains of these proteins may disorganize the crystalline nature of the polymer, rendering them more susceptible to the activity of depolymerases (61). However, it is widely accepted that carbohydrate binding modules bind their cognate substrate almost irreversibly, so it is unclear how such disorganized strands could

then become accessible to other enzymes. The conserved nature of this type of protein among many different microbial chitinolytic systems suggests that it has a significant role in the metabolism of chitin.

The result of the activity of extracellular components of most bacterial chitinase systems is a collection of short chain chitooligosaccharides (degree of polymerization ≤ 8) and chitobiose. Chitobiose is able to freely diffuse into the periplasm, whereas specific porins are thought to be present to transport longer oligosaccharides (65). Specific chitoporins have been identified in some *Vibrio* spp., though these proteins are either specific to *Vibronaceae* or are not similar to those in other known bacterial genera.

The periplasmic components of chitinolytic systems may include chitodextrinases and N-acetyl- β -D-glucosaminidases. Chitodextrinases degrade chitooligosaccharides to GlcNAc and chitobiose, whereas N-acetyl- β -D-glucosaminidases are able to split chitobiose into GlcNAc (66, 67). The resultant sugar pool is composed of a large amount of GlcNAc and a smaller amount of chitobiose. The Phosphotransferase Systems (PTS) of some bacteria have been shown to include a specific transporter subunit for GlcNAc; these systems are able to simultaneously phosphorylate and translocate GlcNAc to the cytoplasm (64). Some bacterial species are also able to phosphorylate and transport chitobiose. The ability of *S. marcescens* to upregulate its chitinolytic system has been shown to be dependent on the transport of chitobiose to the cytoplasm (146, 150). Other bacteria may use antiporters or other transport mechanisms to move GlcNAc to the cytoplasm, as is observed in lactose transport by *E.coli*.

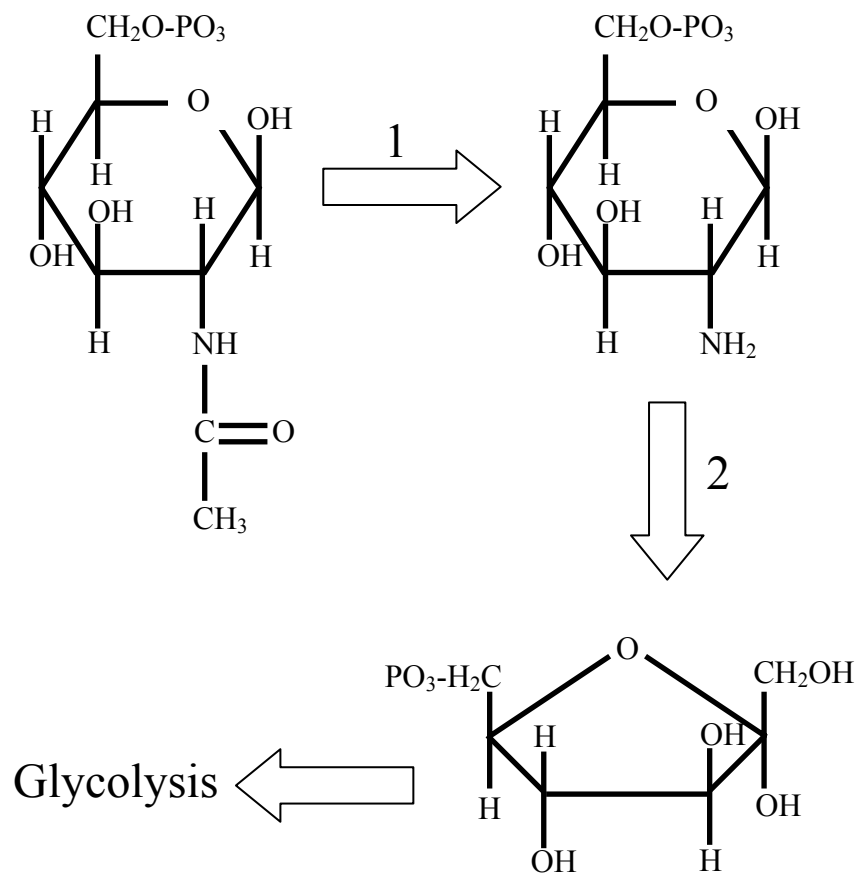
Cytoplasmic components of chitin-metabolizing systems are essentially the same as those used by bacteria that utilize GlcNAc as a sole carbon source, but cannot depolymerize chitin (*E.coli* for example) (109). GlcNAc is either phosphorylated by a GlcNAc phosphorylase or has already been modified by a PTS to form GlcNAc-6-P. This compound can be directly metabolized or isomerized to form GlcNAc-1-P and shunted to cell wall biogenesis. The elegant work of J. Plumbridge *et al* has shown that GlcNAc-6-P is first deacylated by NagA to form glucosamine-6-P (Gln-6-P) and then isomerized by NagB to form fructose-6-P which subsequently enters glycolysis (107-110) (Figure 1-2). The acetate formed by the deacylation of GlcNAc-6-P is also metabolized and the deamination of Gln-6-P releases ammonium, a usable nitrogen source. NagC is a regulator that represses the expression of *nag* genes in the absence of GlcNAc.

Chitin depolymerases and other enzymes involved in insoluble complex polysaccharide (ICP) degradation are typically modular and contain a number of different, though often conserved, domains (47, 48). Chitin depolymerase catalytic domains from both archaea and eubacteria are most often classified as Glycoside Hydrolase Family 18 (GH18) domains. These domains are identified by similarity to a consensus sequence (SIGGWT-X₂₅-DIDWE) that always includes the retention of the key glutamic acid residue. GH18 domains function via a retaining mechanism (71). Interestingly, GH18 domains also rely on substrate assistance for their function where the acetamido group at the 2 position forms an

Figure 1-2

The GlcNAc metabolic pathway of *E. coli*

GlcNAc metabolism requires a cluster of genes encoding proteins with a variety of functions. A PTS transporter subunit specific to GlcNAc simultaneously translocates and phosphorylates the molecule. In the cytoplasm, GlcNAc-6-P is then deacylated by NagA, a GlcNAc-6-P deacetylase (step 1). The resulting molecule, glucosamine-6-P, is then deaminated and isomerized by NagB (step 2) to form fructose-6-P. This sugar then enters glycolysis.



intermediate to permit the reaction to proceed. This fact may account for the observation that GH18 chitinases have decreased activity on highly deacetylated chitins. Glycoside Hydrolase Family 19 (GH19) catalytic domains are primarily found in the chitinases of plants, with the exception of several chitinases from *Streptomyces* spp., *Shewanella*, and *Salmonella* (144). The consensus sequence for this family of enzymes is much more complex and includes two key glutamic acid residues. GH19 domains have an inverting mechanism which results in an α -anomeric product from the β -linked substrate (71). Each family requires water as a component of the degradation reaction. Interestingly, chitiodextrinases also contain GH18 domains, complete with all requisite conserved residues, but are unable to depolymerize chitin (68). Further, both endo- and exo-acting chitinases contain the same catalytic modules despite obvious differences in their cognate substrates (71).

Chitin-binding domains are also found among chitinases, proteases, and non-catalytic chitin binding proteins, in theory to prevent diffusion of an enzyme or protein from its intended substrate (47). This may be of particular importance in marine and littoral environments where enzyme diffusion and dilution are major factors affecting complex polysaccharide degradation. Such diffusion could cause the release of nutrients away from the secreting cell, resulting in an ecological advantage for other bacterial residents of the same niche. Chitin-binding modules appear to bind chitin almost irreversibly (3). A chitin-binding domain from a *Bacillus circulans* chitinase has been used by New England Biolabs in a protein expression vector wherein the coding sequence for a binding domain is transcriptionally fused in frame to a gene of interest and expressed. When purified on chitin columns, the

recombinant protein must be proteolytically cleaved from the chitin binding tag, as it cannot be eluted from the column with SDS, concentrated NaCl washes, or by GlcNAc or chitooligosaccharide solutions. Similar irreversible binding has been observed among cellulose binding domains (127).

Domains of uncertain function are also frequently encountered within prokaryotic chitinases (47). Fibronectin type III domains (FN3) and Polycystic Kidney Disease (PKD) domains have also been described in chitinolytic, xylanolytic, and cellulolytic enzymes. FN3 domains were shown to disorganize cellulose strands of filter paper and increase their susceptibility to depolymerases, and therefore may have a binding/disorganizing function (61). PKD domains have no proposed role in ICP degradation at this time (47). Chitinase 18A of *Clostridium thermocellum* contains a dockerin domain and was shown to be associated with the cellulosome of this organism (159).

Repetitive sequences that link the domains of complex polysaccharide degrading enzymes, including chitinases, have also been described (127). These domains fall into two main classes: polyglycine, serine, or threonine linkers that have a presumed flexible structure (53), and proline/hydroxyl-amino acid linkers with a rigid, extended conformation (149). Polyserine linkers have not been observed in chitinases other than those of *M. degradans*, but are common among the cellulases, pectate lyases, and xylanases of *Cellvibrio japonicus* (53). Polyserine domains can be considered one of the distinguishing features of the complex polysaccharide degrading systems of *Cellvibrio* spp. and *Microbulbifer* spp. Several organisms also encode one to three enzymes with polyserine domains, but they are far less common

than proline/hydroxyamino acid linkers. These linkers have been described in a wide variety of bacterial complex polysaccharide depolymerases from both gram positive and negative organisms and are often shorter than polyserine domains. In *Cellulomonas*, these domains are glycosylated and loss of this modification leads to proteolytic cleavage by extracellular proteases (74). Regardless of the type of domain linker, these motifs have been observed in chitinases between discrete binding domains and between binding and catalytic domains.

The regulation of microbial chitinolytic systems has not been well described (31). Most are subject to repression in the presence of glucose and require chitobiose for induction of extracellular chitinolytic enzymes (146). Many chitinolytic organisms are able to move toward sources of chitin, suggesting that signal transduction plays a role in sensing and responding to this polymer (8, 15).

Streptomyces thermoviolaceus OPC-520 encodes a two component regulatory system that controls its chitinolytic system (72, 143). ChiS of this organism has similarity to a histidine kinase, whereas ChiR has similarity to a response regulator. However, similar proteins in *Streptomyces coelicolor* A3(2) were disrupted with no overall effect on chitinolytic activity (50). Among gram negative bacteria, little is known about the regulation of these systems. It has recently been proposed that a non-catalytic chitin-binding protein produced by *Vibrio furnissii* interacts with a cytoplasmic membrane-anchored regulator to repress chitinase expression (76). The presence and binding of chitooligosaccharides causes the repressor to change conformation, resulting in signal transduction via phosphorylation of a cytoplasmic response regulator. This model has yet to be demonstrated empirically.

Statement of Purpose

The degradation and recycling of carbon and nitrogen by microorganisms is a crucial event in our environment. In many cases, a consortia of organisms takes part in the synergistic metabolism of recalcitrant complex polysaccharides such as cellulose and chitin; *Microbulbifer degradans* strain 2-40, with the ability to metabolize a wide variety of polysaccharides and complex, heterogeneous materials in monoculture, may represent a novel portion of the cycle in which carbon is recycled in the biome. This work was undertaken to evaluate the chitinolytic system of *M. degradans* and apply that knowledge to a wider set of experiments in order to elucidate the mechanisms by which this bacterium is able to efficiently metabolize polysaccharide polymers.

Chapter 2: The Chitinolytic System of *Microbulbifer degradans*

The data in this section appeared in *J Bacteriol* 185(11): 3352-60.

Introduction

M. degradans is able to utilize chitin as a sole carbon source (2). Alpha and beta chitin upregulate chitinolytic enzymes and support robust growth in liquid culture. The growth rates of *M. degradans* on GlcNAc (the monomer of which chitin is composed) and glucose are similar, demonstrating the efficiency of this system. In the presence of polymeric chitin, it was previously shown that three bands of chitinolytic activity were observed in glycol chitin zymograms, though their size was uncertain (L. Whitehead and R. Weiner, unpublished observations). These data, taken together, indicated that *M. degradans* encoded a complete chitinolytic system.

The components of the *M. degradans* chitinolytic system were identified by a combination of screens of genomic libraries for chitinolytic clones and computational biology/bioinformatics. The initial genomic library screens were carried out during the preliminary phases of sequencing the *M. degradans* genome in the event that those efforts met with difficulty. Traditional methods were utilized until the genomic sequencing project generated sufficient data that components of the chitinolytic system could be identified by sequence similarity to other enzymes.

Materials and Methods

Strains and growth conditions

Microbulbifer degradans strain 2-40 was grown in half-strength Marine Medium 2216 (Difco) at 30°C with shaking at 225 rpm. Agar was added to a final concentration of 1.5% to prepare solid media. Alternatively, *M. degradans* was grown in minimal medium [per liter: 2.3% Instant Ocean (Aquarium Systems, Mentor OH), 0.5% ammonium chloride, 0.2% carbon source, 50 mM Tris HCl, pH 7.6] at 30°C. *Escherichia coli* EC100 (Epicentre Technologies, Madison WI), *E. coli* DH5 α (Invitrogen, Carlsbad CA) and *E. coli* Tuner™ (DE3)(pLacI) (Novagen, Madison WI) were grown in Luria-Bertani (LB) broth or agar supplemented with the appropriate antibiotics and incubated at 37°C. Kanamycin, 50 μ g/ml; Chloramphenicol, 30 μ g/ml; Ampicillin, 100 μ g/ml.

Materials and Reagents

Genes of interest were amplified using ProofPro DNA polymerase (Continental Lab Products, San Diego CA). Restriction enzymes, amylose resin, T4 DNA ligase, and the pMal-2pX expression vector were obtained from New England Biolabs (Beverly MA). Ni-nitrilotriacetic acid (Ni-NTA) agarose and the pETBlue2 expression vector were from Novagen (Madison WI). The pCC-1 fosmid vector was obtained from Epicentre Technologies (Madison WI). Centrifugal filters were from Millipore (Bedford MA). All other reagents and chemicals were obtained from Sigma-Aldrich (St. Louis MO).

Construction of the M. degradans genomic library

M. degradans chromosomal DNA was isolated using a Wizard™ Genomic DNA Preparation kit (Novagen) and sheared by passage through a Hamilton syringe. Genomic fragments of approximately 40 KB in size were gel extracted from a 1.0% low melting point agarose gel and end-repaired using the pCC1 fosmid library production kit. Genomic fragments were ligated into the linearized, dephosphorylated pCC-1 vector and packaged into T1 phage particles. Packaged phage were used to transfect the library into *E.coli* EC100 and transfectants were isolated by virtue of the *cat* gene carried by pCC-1.

Genomic library screens for chitinolytic clones

E. coli library clones were initially screened for chitin depolymerase activity by plating on LB agar supplemented with 0.1% colloidal chitin or 0.08% chitin azure. Clones were incubated at 37°C and periodically screened for zones of clearing indicative of chitin depolymerase activity. Alternatively, the chitin analogs 4'-methylumbelliferyl- β -D-N,N'-diacetylchitobiose (MUF-diNAG) and 4'-methylumbelliferyl- β -D-N,N',N''-triacetylchitotrioside (MUF-triNAG) were used to screen library clones for chitinolytic activity (121). Single clones were grown in 100 μ l of LB broth in microtiter plates with shaking at 25°C for 12 h. A MUF analog was added to a final concentration of 1.5 μ M and incubated for an additional 24 h. Degradation of the analog was determined by fluorescence under 366 nm UV light.

Fosmid mutagenesis and sequencing

Fosmids of interest were purified using a Midiprep Plasmid Purification kit (Epicentre). The purified fosmid was mutagenized in vitro using an EZ-Tn <Kan-2> Insertion kit (Epicentre) and electroporated into *E. coli* DH5 α . The transformants were screened for loss of activity and the library construct was then purified from loss of function clones. The DNA sequence of the genomic insert flanking the transposon was determined using sequencing primers specific to the ends of the transposon.

Zymography

M. degradans was grown for 48 h in minimal media containing 0.2% colloidal chitin at 30°C. The following steps were performed at 4°C: Cultures were centrifuged at 10,000 X g for 20 minutes. The supernatant was passed through a 0.22 μ m syringe filter and concentrated 50 fold using a centrifugal concentrator with a 10 kDa cutoff pore size. Proteins within these concentrated supernatants were fractionated by SDS-PAGE with a stacking gel by the method of Laemmli in an 8% acrylamide separating gel containing a final concentration of 0.01% ethylene-glycol chitin or 0.05% gelatin (142). Gels were then incubated in refolding buffer (50 mM Tris, 1 mM EDTA, 5 mM 2-mercaptoethanol [pH 7.5]) at 4°C for 24 h. Gels were washed for one hour in 100 mM sodium phosphate buffer (pH 7.0) at 25°C with frequent changes of buffer followed by incubation in the same buffer at 37°C for 16 h. Gels containing glycol chitin were rinsed and washed in developing buffer (0.5 M Tris, 0.01% Calcofluor [pH 7.5]) for 5 min and then rinsed with distilled water for 2 h. Zones of chitin depolymerase activity appeared as dark bands when viewed under

UV light. Gels containing gelatin were rinsed, stained with 0.1% Coomassie Brilliant Blue R-250, and destained. Bands of protease activity appeared as unstained bands against a blue background.

Protein Expression and Purification

Genes of interest were amplified using PCR and primers tailed with appropriate restriction sites. Each PCR product was then incubated with the appropriate restriction enzymes and ligated into pETBlue2 or pMal-2pX expression vectors and transformed into *E.coli* Tuner™ or *E.coli* DH5α for protein expression. A 50 ml culture of an expression culture was grown at 37°C to an optical density at 600 nm of 0.5, induced with isopropyl-β-D-thiogalactopyranoside (IPTG), and grown for an additional 3 h. Cells were harvested and resuspended in lysis buffer and clarified lysates were prepared as recommended by the manufacturer of the affinity resin. pETBlue-2 His tag fusions were purified using Ni-NTA agarose and pMal-2pX maltose-binding protein (MBP) fusions were purified using amylose resin. Purified recombinant proteins were collected, concentrated using centrifugal filters, buffer exchanged into storage buffer (50 mM Tris, 50 mM NaCl, 10% glycerol, pH 7.5), and stored at -80°C.

Chitinase activity assays using chitin analogs and chitooligosaccharides

To determine the activity of the chitin depolymerases against the chitin analogs MUF-diNAG and MUF-triNAG, 900 μl of a 50 μM solution of each analog was added to 100 μl of purified enzyme and incubated at 37°C for 30 min. The

fluorescence of the reaction (excitation wavelength, 365 nm; emitted wavelength, 460 nm) was determined using a Hoefer TKO-100 fluorometer and compared to a standard curve prepared with 4'-methylumbelliferone. One unit of enzyme activity is defined as the number of μ moles of 4'-methylumbelliferone released per min per mg of enzyme (121). To determine the activity of enzymes of interest on native chitooligosaccharides, degradation products of reactions were separated by thin layer chromatography. Enzyme reactions contained 0.45 μ mol of chitooligosaccharide substrate in 10 mM Tris at pH 7.5. After 1 h at 30°C, reactions were stopped by boiling for 10 min. Degradation products were fractionated on silica gel plates by developing in 5:2:1 2-propanol-ethanol-water for 1 h. The plate was air dried, sprayed with 10% sulfuric acid in ethanol, and baked at 120°C for 20 min. Chitooligosaccharide spots appeared brown and were compared to standards.

Complementation of a nagA mutant

The *M. degradans nagA* gene was amplified using PCR and ligated into pBluescript SK+ to create pNagA. *E.coli* K-12 strain IBPC531 (*nagA::cm*) was transformed with pNagA, plated on minimal medium (M63 minimal salts, 0.2% GlcNAc, thiamine, 0.05% IPTG), and incubated at 37°C (110). IBPC531 and IBPC531 (pBluescript SK+) were plated on the same medium as controls.

Sequence manipulations and analysis

Protein domains and modules were identified using the Simple Modular Architecture Tool (SMART) and the PFAM database (126). Similarity searches were

performed using the BLAST algorithm at the NCBI server (1). Type II secretion signals were identified using the iPSORT program and the SignalP version 1.1 program (96). Multiple sequence alignments were performed using the ClustalW program. Estimated protein molecular weights were calculated at the ExPASy server of the Swiss Institute of Bioinformatics.

Results

M. degradans culture supernatants contain multiple chitinases

To estimate the number of chitinases secreted by *M. degradans* in response to chitin, concentrated culture supernatants were analyzed using zymograms. Late-log phase cultures grown in α -chitin were concentrated and separated in a 0.01% glycol chitin zymogram. After renaturation and incubation, zones of chitin depolymerization were evident (Figure 2-1). Three chitin depolymerases were observed using this method with apparent sizes of 140, 90, and 60 kDa. These three bands were consistently observed at the same sizes and were found to be in agreement with the subsequently obtained genomic data.

Cloning of a chitodextrinase

E. coli clones containing the genomic library of *M. degradans* were screened for activity against MUF-diNAG and MUF-triNAG. Six clones were identified that could degrade both of these substrates. The clone that appeared to cleave the substrates the most rapidly, 3-42, was selected for further analysis. The library

Figure 2-1

Zymogram of culture supernatant of chitin-grown *M. degradans*

The supernatant of an early stationary phase culture of *M. degradans* grown with chitin as the sole carbon source was sterilized by filtration and concentrated 50 fold. The concentrated sample was separated in a 0.01% glycol chitin zymogram and the resident proteins refolded. After incubation in activity buffer and staining with Calcofluor, dark bands indicative of chitin depolymerase activity are evident. The observed sizes are in good agreement with the predicted weights of ChiA, ChiB, and ChiC obtained from the genomic sequence.

140 kDa ➡

90 kDa ➡

60 kDa ➡



construct carried by 3-42 was purified, mutagenized using an EZ-Tn *in vitro* transposon mutagenesis kit, and re-introduced into *E. coli*. Transformants that were no longer active against the MUF analogs were identified and the genomic DNA sequence flanking the transposon was determined. A complete open reading frame was identified by successive rounds of primer walking. The gene responsible for the MUF degrading activity encodes a 1088 aa protein with a calculated molecular mass of 115 kDa. It includes a type II secretion signal, two Polycystic Kidney Disease (PKD) domains, a 403 aa Glycoside Hydrolase family 18 (GH18) domain, and a 41 aa chitin-binding domain. The GH18 domain is most similar to the catalytic site of the chitodextrinase ChiD from *Alteromonas* sp. strain O-7 (72% identity, 83% similarity). The chitin-binding domain was most similar to a chitin-binding domain within the ChiA protein of *Pseudoalteromonas* sp. strain S91 (64% identity, 71% similarity). A PCR screen of the six clones able to degrade the MUF analogs showed that three encoded this gene.

Sequence features of three chitin depolymerases identified in the genomic sequence of M. degradans

A draft version of the *M. degradans* genomic sequence became available before additional analysis of the six analog degrading library clones could be undertaken. To identify candidate genes involved in the chitinolytic pathway of *M. degradans*, the sequence of bacterial enzymes with demonstrated chitinolytic activity were obtained from GenBank and used in BLAST searches of the *M. degradans* genomic sequence. These searches were initially performed using the full length

sequences of enzymes produced by closely related *Alteromonas* spp. and later with the sequence of the catalytic and binding domains of enzymes from marine and other sources. This led to the identification of three genes encoding candidate chitin depolymerases and confirmed the sequence of the chitodextrinase gene obtained from previous experiments.

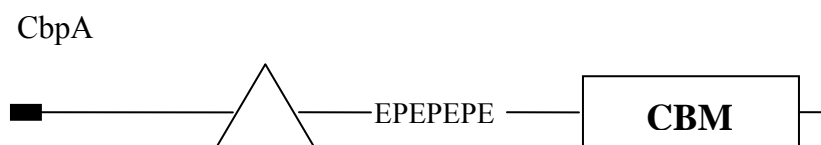
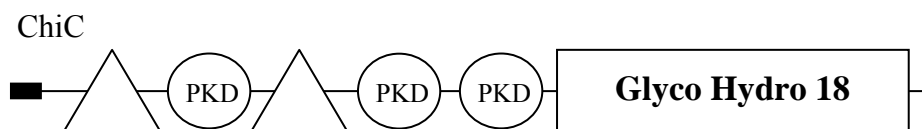
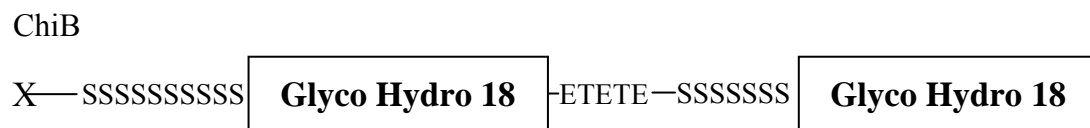
The first chitin depolymerase, ChiA, is a 543 aa protein with a calculated mass of 57.0 kDa. ChiA includes two chitin binding domains and a GH18 catalytic domain (Figure 2-2). The first binding domain is most similar to a chitin-binding domain of ChiA from *Pseudoalteromonas* sp. strain S91 (60% identity, 75% similarity), while the second binding domain is most similar to a chitin-binding domain from ChiA of *Vibrio cholerae* (44% identical, 57% similar). The 299 aa GH18 domain is most similar to the GH18 domain of ChiA from *V.cholerae* (65% identity, 77% similarity). The GH18 domain includes the consensus sequence SLGG-X₃₂-DIDLE, typical of family 18 chitinases. Two polyserine linkers are also found in ChiA; the first is a 46 residue linker located between the two chitin-binding domains, while the second is 36 residues in length and separates the second chitin-binding domain from the catalytic domain.

The second chitin depolymerase, ChiB, is a 1271 aa enzyme with a calculated mass of 136.1 kDa. ChiB has two discrete GH18 domains, but has no predicted chitin-binding or carbohydrate binding domains (Figure 2-2). The first GH18 domain is 384 aa in length and is most similar to the catalytic site of ChiB from *Serratia marcescens* strain BJL200 (55% identity, 69% similarity). The second GH18 domain is most similar to a chitinase from *Vibrio* sp. strain 5SM-1 (49% identity, 66%

Figure 2-2

Extracellular components of the *M. degradans* chitinolytic system

From top to bottom: ChiA, ChiB, ChiC, CbpA. The conserved domains of the extracellular components of the chitinolytic system were determined using the Simple Modular Architecture Tool. The localization of these enzymes is based on the presence of a Type II secretion signal and the known localization of similar microbial enzymes. Triangles, chitin-binding domains; Glyco Hydro 18, Glycoside Hydrolase family 18 catalytic domain; SSS, polyserine linker domain; PKD, polycystic kidney disease domain; EP, Glutamic acid-Proline linker; CBM, carbohydrate binding domain of uncertain specificity; black box, Type II secretion signal; X, putative lipoprotein acylation site.



similarity) and is 394 aa long. The amino-terminal and carboxy-terminal GH18 domains include the consensus sequence SVGG-X₃₈-DIDWE and SIGG-X₃₇-DIDWE, respectively. Polyserine linkers are also present in ChiB; the first begins immediately after a predicted lipoprotein acylation site at the extreme amino terminus and continues for 28 residues. After a 12 aa domain, a 109 residue serine-rich linker is present. A third polyserine linker separates the two catalytic domains and contains 51 residues. An additional domain linker composed of (Glu-Thr)₁₀ is located between the third polyserine domain and the second GH18 catalytic domain.

The third chitin depolymerase, ChiC is 792 aa in length and has a calculated mass of 87.1 kDa. ChiC includes two chitin-binding domains; the first is a 46 aa domain that is most similar to a chitin-binding domain from ChiB of *Vibrio harveyi* (57% identity, 63% similarity), while the second is a 49 aa domain and is most similar to a chitin-binding domain from ChiA of *V. cholerae* (57% identity, 75% similarity) (Figure 2-2). ChiC also contains three PKD domains. The 350 aa GH18 domain is most similar to the GH18 domain of ChiC from *Streptomyces peucetius* (61% similarity, 71% identity) and has the consensus sequence SFGG-X₃₇-DIDWE. No repetitive linker domains are found within ChiC.

Sequence features of the periplasmic components of the M. degradans chitinolytic system

Chitooligosaccharides and chitobiose are produced from the activity of chitin depolymerases. These sugars are then taken up and modified in the periplasm before transport to the cytoplasm. Several enzymes thought to localize to the periplasm were

identified in the genomic sequence in addition to the putative chitodextrinase, CdxA, which had been cloned during initial genomic library screens.

Two predicted N-acetyl-glucosaminidases were identified in the genomic sequence. The first, HexA, is a 795 aa enzyme with a predicted mass of 88.5 kDa. HexA includes a GH20b domain that is most similar to the active site of the *Alteromonas* sp. strain O-7 N-acetylhexosaminidase (32% identity, 57% similar) and a 348 aa GH20 domain related to the active site of a *Pseudoalteromonas* strain S91 N-acetylglucosaminidase (58% identity, 73% similar) (Figure 2-3).

The second N-acetyl-glucosaminidase, HexB, is an 889 aa enzyme with a predicted mass of 98.4 kDa that contains a putative carbohydrate binding domain, a GH20b domain, and a 406 aa GH20 domain that is most similar to the active site of an N-acetylhexosaminidase from *Vibrio vulnificus* (55% identity, 70% similarity) (Figure 2-3).

Biochemical activities of CdxA, HexA, and HexB

To confirm the function of each of the periplasmic chitooligosaccharide modifying enzymes, each was purified by virtue of a HexaHis tag and incubated with chitobiose, chitooligosaccharides, MUF-diNAG, and MUF-triNAG. Each was also analyzed in a glycol chitin zymogram. Consistent with N-acetyl-glucosaminidase activity, HexA and HexB could degrade MUF-diNAG, MUF-triNAG, and chitobiose, but not polymeric chitin or chitooligosaccharides. CdxA, however, was active on MUF-diNAG, MUF-triNAG, chitooligosaccharides, but not chitobiose or polymeric chitin. These data indicate that the chitobiose and chitooligosaccharides that enter the *M.*

degradans periplasm are acted upon by different enzymes. The chitooligosaccharides are likely degraded by the activity of CdxA to release chitobiose and some GlcNAc (if the chitooligosaccharide chain contains an odd number of residues). HexA and HexB then act on the majority of the chitobiose to form GlcNAc. This large pool of GlcNAc and some chitobiose is then taken up by the cell for metabolism.

Figure 2-3

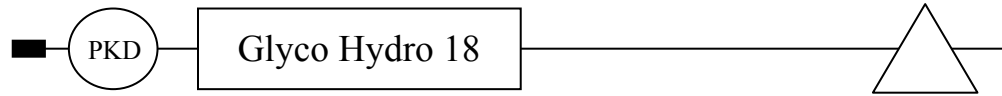
Periplasmic and cytoplasmic components of the *M. degradans* chitinolytic system

From top to bottom: CdxA, HexA, HexB, HexC. The periplasmic and cytoplasmic components of the chitinolytic system were designated as such by the presence of a Type II secretion signal and the known localization of similar microbial enzymes.

The presence of a putative lipoprotein acylation site at the amino terminus of HexA suggests that HexA may be anchored to the outer membrane. The conserved domains of these enzymes were identified using the SMART protein architecture tool.

Triangle, chitin-binding domain; PKD, polycystic kidney disease domain; Glyco Hydro 18, Glycoside Hydrolase family 18 domain; GH20b, Glycoside Hydrolase family 20 ancillary domain, Glyco Hydro 20, Glycoside Hydrolase family 20 domain, SSBD, soluble sugar binding domain; Glyco Hydro 3, Glycoside Hydrolase family 3 domain; black box, Type II secretion signal; X, putative lipoprotein acylation site.

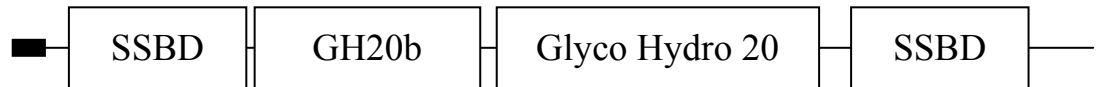
CdxA



HexA



HexB



HexC



Cytoplasmic components and inner membrane transporters of the M. degradans chitinolytic system

The combined activity of the *M. degradans* extracellular and periplasmic chitinolytic enzymes theoretically results in a pool of GlcNAc and some chitobiose. Chitobiose is known to be a critical signaling molecule in *S. marcescens* and upregulation of chitin depolymerase genes does not occur until some chitobiose reaches the periplasm. This is most likely also the case for *M. degradans*; Western blots using a polyclonal antibody raised against ChiB show that it is not expressed in the presence of GlcNAc (see Chapter 3). After serving its signaling function, chitobiose is presumably cleaved into two molecules of GlcNAc by cytoplasmic N-acetylglucosaminidases. Consistent with this theory, the gene for HexC was identified in the genomic sequence of *M. degradans*. This enzyme contains 345 aa and has a predicted mass of 37.4 kDa. The amino terminus has a negative hydropathy and is not consistent with Type II secretion signals of gram negative bacteria, suggesting it remains in the cytoplasm. HexC has a GH3 catalytic domain that is most similar to a *Pseudomonas aeruginosa* N-acetylhexosaminidase (50% identity, 65% similarity) (Figure 2-3).

Other components of the *M. degradans* cytoplasmic chitinolytic system include genes for proteins with similarity to NagA, NagB, and NagC of *E. coli* K-12. These genes are part of an apparent operon (*nagC-nagB-nagA*) (Figure 2-4). The predicted NagC_{mdg} is a member of the LacI family of transcriptional regulators and was most similar to LacI of *Xyella fastidiosa* strain 9a5c (47% identity, 65%

similarity). As in the NagC proteins of other bacteria, NagC_{mdeg} contained a helix-turn-helix motif.

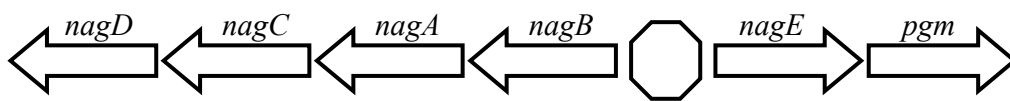
NagA_{mdeg} contains a putative N-acetylglucosamine-6-P deacetylase domain and was overall most similar to the *Caulobacter crescentus* N-acetylglucosamine-6-P deacetylase (55% identity, 70% similarity). Expression of NagA_{mdeg} in *E.coli* K-12 strain IBPC531, a mutant lacking a functional *nagA* gene, restored growth of this strain on minimal medium containing GlcNAc as the sole carbon source, thus confirming the function of NagA_{mdeg}. NagB_{mdeg} includes two sugar isomerase domains and is most similar to NagB of *C. crescentus* (56% identity, 69% similarity). A homologue to the *E.coli* sugar transporter NagE was also identified in the *M. degradans* genome that is expressed from an apparent *nagF-hexA-nagE* operon. NagE_{mdeg} contains a typical secretion signal and 11 transmembrane domains. It is overall most similar to a Glu/Gal transporter from *X. fastidiosa* strain 9a5c (43% identity, 61% similarity). NagF is most similar a GlcNAc kinase from *Brucella melitensis* (32% identity, 48% similarity). The orientation of these operons is different from their putative counterparts in *E. coli* (Figure 2-4).

Figure 2-4

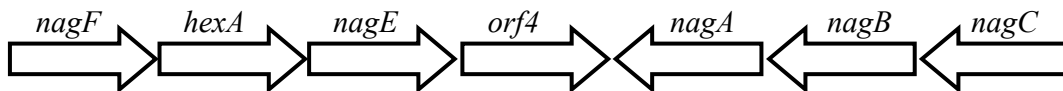
Comparison of the *E. coli* and *M. degradans* GlcNAc metabolism clusters

E. coli is able to metabolize GlcNAc as a sole carbon source by virtue of the *nag* operon. The *E. coli* GlcNAc cluster includes 6 genes (*nagA*, *nagB*, *nagC*, *nagD*, *nagE*, and *pgm*). *pgm* is a phosphoglucomutase. These are regulated from a common locus between the operons where the promoters for each are located. Here, NagC binds to repress the transcription of the operons in the absence of GlcNAc. In *M. degradans*, the GlcNAc cluster is composed of 7 genes (*nagA*, *nagB*, *nagC*, *nagE*, *hexA*, and *orf4*). There is no homologue of the *E. coli nagD* gene in the *M. degradans* genome, and *orf4* has no predicted function. The method of repression of these operons by NagC has not been established. Octagon, NagC binding site.

E. coli NAG Cluster



M. degradans NAG Cluster



Two genes encoding components of a possible chitobiose transport system were identified in the *M. degradans* genome. The first is similar to a diacetylchitobiose phosphorylase from *Vibrio furnissii* (33% identity, 49% similarity), while the second encodes a protein with similarity to a sodium/galactoside permease from *Salmonella enterica* serovar Typhimurium LT2 (36% identity, 55% similarity). The presence of this putative phosphorylase suggests that chitobiose may not enter *M. degradans* in a PTS-dependant manner.

Accessory enzymes of the M. degradans chitinolytic pathway

Additional enzymes were identified that may be involved in the chitinolytic pathway of *M. degradans*. The first of these is a potential chitin-binding protein, CbpA. CbpA contains two carbohydrate binding domains separated by a proline/threonine rich linker, but does not contain a putative catalytic site (Figure 2-2). The first binding domain consists of 220 aa and is most similar to a chitin-binding region of CbpD of *P. aeruginosa* (32% identity, 45% similarity). The second binding domain of CbpA is a 95 aa type 2 carbohydrate binding domain similar to a CBM2 domain within a rhamnogalacturonan lyase of *Cellvibrio japonicus* (64% identity, 74% similarity).

To determine if specific proteases are upregulated in response to chitin, a gelatin zymogram was employed. Concentrated culture supernatants of *M. degradans* utilizing different carbon sources were analyzed in this gel to test for protease activity. In the case supernatants from chitin-grown cells, two bands of protease

activity were evident at approximately 100 kDa and 30 kDa (Figure 2-5). These bands were not present in glucose grown cells. Attempts were made to isolate this band and analyze it using MALDI-TOF mass spectroscopy but were not successful, possibly due to the large amount of protein present in the zymogram gel.

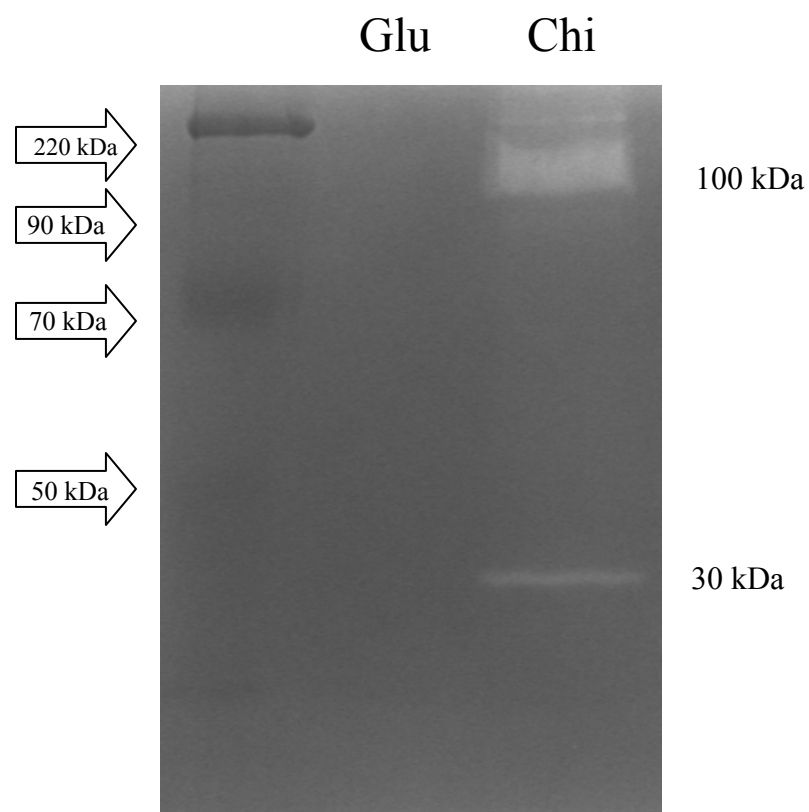
Genetic organization of the M. degradans chitinolytic system

Many marine bacteria, such as *Alteromonas* sp. strain O-7 and *Pseudoalteromonas* sp. strain S9, have clusters of genes involved in chitin degradation. However, in the case of *M. degradans* the major components of the system are dispersed throughout the genome, the only exception being the *nag* utilization cluster. Further, none of the chitinases were clustered with genes encoding other putative degradative enzymes. Because of the unusually large number of CP degrading enzymes encoded within the genome, it is thought that lateral gene transfer may have been involved in the acquisition of these genes. Because it is unlikely that clusters of chitinase genes would be taken up and incorporated at the same location, the dispersed nature of the system lends support to the horizontal acquisition of these genes.

Figure 2-5

Proteases are secreted in response to chitin

Supernatants of early stationary phase glucose and chitin grown cultures were collected and sterilized by filtration. These supernatants were concentrated 50 fold and separated in a zymogram containing 0.08% gelatin. After refolding and incubation in activity buffer, the zymogram was stained with Coomassie Blue. After destaining, clear bands typical of protease activity were apparent. The larger protease band is approximately 100 kDa in size, while the smaller is approximately 30 kDa. Glu, concentrated glucose culture supernatants; Chi, concentrated chitin culture supernatants. Protein markers were Kalidescope Markers (Biorad).



M. degradans is sensitive to Streptozotocin

Streptozotocin is a GlcNAc analog; when taken up by a cell and phosphorylated, it forms the strong alkylating agent diazomethane that interferes with DNA synthesis (120). Sensitivity to streptozotocin has been suggested as indicator of PTS activity. *M. degradans* was grown in minimal media with chitin or GlcNAc as the sole carbon source both with and without streptozotocin added to a final concentration of 20 µg/ml. Growth was completely inhibited in the cultures in which streptozotocin was present. This data, taken with the fact that some elements of a PTS system are evident in the genome (e.g., genes for PTS enzyme I and HPr), suggests that *M. degradans* may transport GlcNAc in a PTS-dependent manner. However, PTS systems are rare among strict aerobes. Further, streptozotocin may be taken up by *M. degradans* in a non-PTS dependent manner and phosphorylated by the activity of NagF, which theoretically would also produce diazomethane. This would explain non-PTS-dependent sensitivity to streptozotocin. In addition, genes for many of the specific PTS enzyme II subunits for some of the sugars utilized by *M. degradans* are missing from the genome. Thus, the exact mechanisms by which GlcNAc and chitobiose are transported are unknown. The lack of a workable genetic system for *M. degradans* has made this a particularly difficult question to address.

Discussion

The chitinolytic system of *M. degradans* is composed of a typical complement of enzymes, however some of these enzymes have novel and interesting features (Table 2-1). The presence of two catalytic domains within a single Eubacterial

chitinase has not been described, and the presence of polyserine domains within the *M. degradans* chitinases is also novel. Taken together, the data obtained during the preceding experiments allow a model pathway of chitin metabolism by this organism to be proposed (Figure 2-6). The activity of the three chitin depolymerases, which are shown here to degrade polymeric chitin, release chitobiose and longer, soluble chitooligosaccharides from chitin strands. In *Vibrio harveyi*, chitooligosaccharides are transported to the periplasm via a specific outer membrane porin; no homologues of this protein were identified in the *M. degradans* genome (63). Chitotriose, chitobiose, and GlcNAc likely enter the periplasm by diffusion. In the periplasm, CdxA degrades chitooligosaccharides to chitobiose, while the growing pool of chitobiose is cleaved to form GlcNAc by the activity of HexA and HexB, both N-acetylglucosaminidases.

The mechanism by which chitobiose and GlcNAc are transported to the cytoplasm of *M. degradans* is not clear. A homologue of the *E. coli* NagE protein is clustered with genes for other GlcNAc modifying enzymes in the *M. degradans* genome. NagE is the GlcNAc specific subunit of the *E. coli* PTS system, but it is unclear if *M. degradans* encodes a functional PTS system (109). Therefore, two models of GlcNAc transport and modification are possible. First, NagE_{mdg} may not be involved in PTS transport and another unknown enzyme is involved

Table 2-1

Relevant physical and biochemical characteristics of the components of the *M. degradans* chitinolytic system¹

Enzyme	Amino Acids	Estimated MW	GC Content²	N-term Hydropathy	Active on Chitobiose	Active on C.O.S.	Active on Chitin	Glycoside Hydrolase Catalytic Domain	Cellular Location	Accession Number
ChiA	543	59.7	49.7	1.573	-	N/D	+	GH18	Extracellular	DAA01335
ChiB	1271	139.8	46.9	2.167	-	+	+	GH18 (2)	Extracellular	DAA01334
ChiC	792	87.1	54.2	1.000	-	N/D	+	GH18	Extracellular	DAA01336
CdxA	1088	119.7	52.7	1.547	-	+	-	GH18	Periplasmic	AAO47607
HexA	795	87.5	46.0	1.527	+	N/D	-	GH20	Periplasmic or Extracellular	DAA01338
HexB	889	97.8	47.2	0.993	+	N/D	-	GH20	Periplasmic	DAA01339

¹Each enzyme was tested in a battery of biochemical experiments to determine its substrate specificities. Protein domains and

modules were identified using the Simple Modular Architecture Tool (SMART) and the PFAM database. Type II secretion signals

were identified using the iPSORT program and the SignalP version 1.1 program. Estimated protein molecular weights were calculated

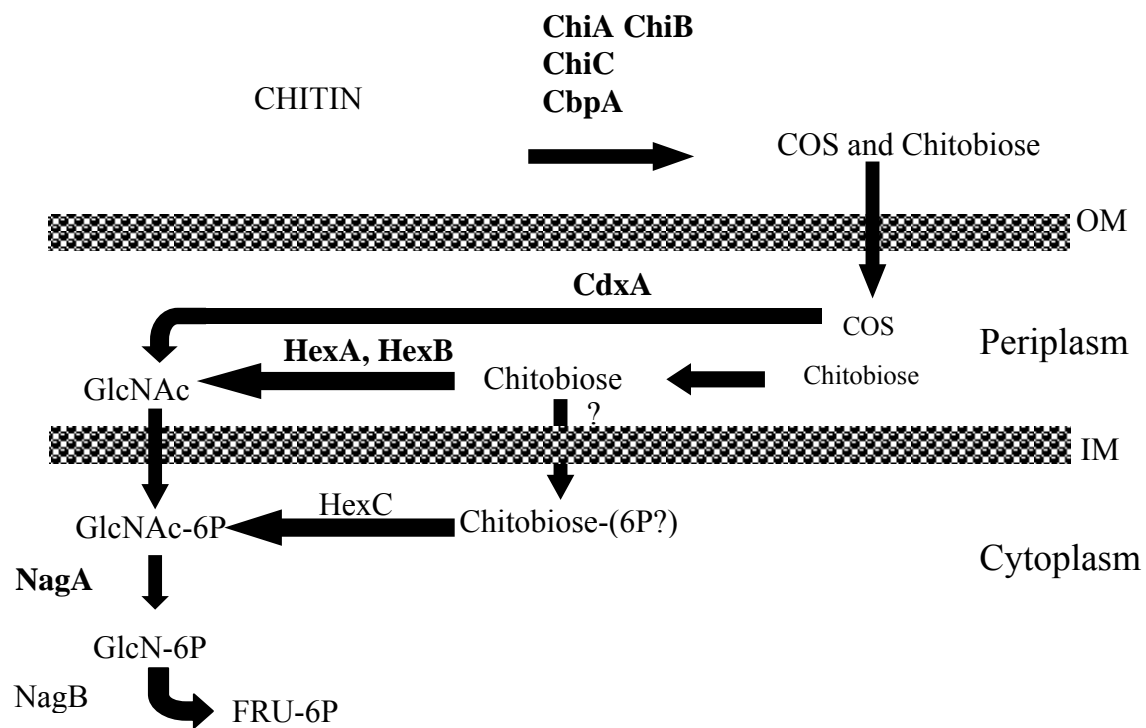
at the ExPASy server of the Swiss Institute of Bioinformatics. N/D, Not determined; GH18, Glycoside Hydrolase family 18 domain; COS, chitooligosaccharides.

²GC content of the *M. degradans* genome is approximately 44.6%

Figure 2-6

A model of chitin degradation by *M. degradans*

Chitin is likely degraded by the activity of the three chitin depolymerases (ChiA, ChiB, ChiC) to form chitooligosaccharides and chitobiose, which enter the periplasm. There, CdxA (a chitodextrinase) degrades chitooligosaccharides to N-acetylglucosamine and chitobiose, and most chitobiose is cleaved to form additional N-acetylglucosamine by HexA and HexB. Chitobiose and N-acetylglucosamine are then transported, perhaps via PTS transporters, to the cytoplasm. GlcNAc-6-P is then deacylated by NagA to form glucosamine-6-P, which may then be isomerized by NagB and metabolized. Chitobiose may function as a signaling molecule before it is cleaved by HexC and similarly metabolized. OM, outer membrane; IM, inner membrane, GlcNAc, N-acetyl-glucosamine; GlcN, glucosamine; COS, chitooligosaccharides; Fru-6-P, fructose-6-phosphate.



in the PTS dependent transport of GlcNAc. As GlcNAc is transported to the periplasm by this unknown PTS transporter, it is simultaneously phosphorylated at the C6 position. The second possibility is that NagE_{mdg} or some other transporter brings GlcNAc to the cytoplasm in a non-PTS manner and GlcNAc must be phosphorylated there by the activity of NagF, a predicted GlcNAc kinase. In either case, GlcNAc-6-P is likely then deacylated, isomerized, and metabolized via glycolysis.

Chitobiose transport is also an important and somewhat unclear process in *M. degradans*. GlcNAc is not sufficient to upregulate the expression of chitin depolymerases in this organism, therefore another soluble sugar must reach the cytoplasm to function as a signaling molecule. Many chitinolytic systems are fully upregulated by the presence of chitobiose, and a *S. marcescens* mutant unable to transport chitobiose was unable to depolymerize or metabolize chitin (146). Therefore, it is likely that some chitobiose reaches the *M. degradans* cytoplasm. A pair of genes with some similarity to a chitobiose kinase and a sugar transporter was identified in tandem in the genome, though their function could not be confirmed. In the event that chitobiose is taken up and phosphorylated, it is likely cleaved after a period of time by HexC, a cytoplasmic N-acetylglucosaminidase. This would result in the release of a GlcNAc-6-P molecule and a GlcNAc molecule. Because GlcNAc must be phosphorylated to be utilized, this GlcNAc may be the intended substrate of NagF as opposed to the above model wherein all GlcNAc available to the cell reaches the cytoplasm in an unphosphorylated state.

Another intriguing feature of the *M. degradans* chitinolytic system is found within CdxA. A chitin-binding domain is located at the extreme carboxy terminus of this protein. CdxA is predicted to act within the periplasm of *M. degradans*, so in theory the chitin-binding domain would never encounter its intended substrate. However, it may be that certain chitin-binding domains have adapted to bind short chain chitooligosaccharides. This would be similar to the catalytic GH18 domains of chitodextrinases that have evolved to degrade chitooligosaccharides but have no activity on polymeric chitin. This chitin-binding domain may be an evolutionary artifact, or it may represent an opportunity to analyze the structure of a chitooligosaccharide binding domain. A third possibility that must be considered is that CdxA is extracellular, though without a specific antibody against this enzyme this notion would be difficult to demonstrate empirically. Further, it seems unlikely that an enzyme that releases small, soluble sugars would act outside of the cell where the liberated sugars are susceptible to diffusion and loss.

The polyserine linkers found in ChiA and ChiB are also a novel feature of this enzyme system. In ChiA, one of these linkers separates two chitin-binding domains, while the other separates a binding domain from the catalytic domain. In ChiB, they separate a putative lipoprotein acylation site from the remainder of the protein and also two catalytic domains. The polyserine domains average 39 residues in length and are predominantly composed of serine, though a few threonine and glycine residues are also present. All six codons for serine are used to encode these domains with no detectable bias and in no obvious pattern or repeat; in no case are three identical serine codons found in tandem (See Chapter 4 for a more detailed analysis of these

domains). ChiA and ChiB are the first chitinases reported to contain polyserine linkers of this type.

During this work, concurrent efforts were made to create a workable genetic system for *M. degradans*. Clearly the ability to perform site directed mutagenesis, complementation, and expression of proteins with epitope tags within *M. degradans* would have greatly strengthened the preceding work. Many attempts were made to mutagenize and transform *M. degradans*, but were not met with success. Because of this difficulty, much of the subsequent work performed has been based on recombinant protein expression in *E. coli* and characterization of those enzymes. However, the use of classical genetics is still not a feasible option to study *M. degradans*.

Chapter 3: Chitinase B of *Microbulbifer degradans* Contains Two Catalytic Sites with Different Activities on Chitin

The data in this section appeared in *J Bacteriol.* 186(5):1297-303.

Introduction

Chitin is a difficult substrate for microbial degradation because it is usually crystalline and complexed with protein, salts, and other carbohydrates. However, many microorganisms have developed efficient strategies for the depolymerization, transport, and metabolism of chitin and its derivatives. These systems involve multiple enzyme activities, usually encoded on separate polypeptides.

Pseudoalteromonas sp. strain S91 (138), *Serratia marcescens* (132), and *Streptomyces colicolor* A3(2) (123), for example, secrete several chitin-depolymerizing enzymes in the presence of chitin. Endo- and exo-chitinases have been described that function cooperatively to depolymerize chitin (17, 37). Endochitinases randomly cleave glycosidic linkages, generating free ends and long chitooligosaccharides (Figure 3-1). These are then acted upon by exo-chitinases that release chitobiose from the non-reducing ends of each. While exo- and endo-chitinases are able to depolymerize chitin alone, the presence of both activities significantly increases the efficiency of chitinolytic systems.

The Glycoside Hydrolase Family 18 (GH18) domain is the most common catalytic domain of microbial chitin depolymerases (47, 48). Despite sharing a consensus sequence and a conserved catalytic glutamic acid residue, GH18 domains may differ in their activity toward polymeric chitin and chitooligosaccharides (i.e.,

endo- vs. exo- activity). Chitodextrinases, which depolymerize chitooligosaccharides but not chitin, also contain GH18 domains (47). Chitinolytic enzymes with GH18 domains have been isolated from organisms as diverse as psychrophilic eubacteria and hyperthermophilic archaea, demonstrating the wide range of conditions to which these domains have adapted. Because conserved residues are found in GH18 domains with divergent optima and substrate specificities, sequence analysis is insufficient to determine the enzymatic specificities of newly discovered chitinases.

ChiB of *M. degradans* is a modular 1271 amino acid enzyme with a calculated weight of 136.1 kDa (54). The amino terminus is predicted to contain a secretion signal that is separated from the remainder of the protein by a polyserine domain of 148 amino acids, 99 of which are serine residues (Figure 3-2). ChiB is predicted to include two complete GH18 domains separated by a 180 amino acid linker domain which includes an acidic region consisting of TE-(ET)₁₀ and another polyserine domain containing 39 serine residues. Only *Thermococcus kodakaraensis* KOD1 encodes a similar chitinase (Tk-ChiA) with two catalytic domains, though it differs from ChiB of *M. degradans* in that it has three chitin-binding domains and proline rich linker sequences (134) (Figure 3-2).

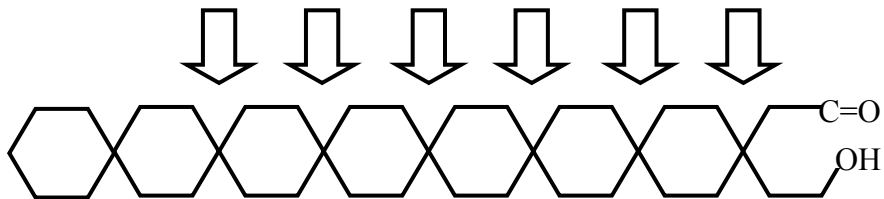
Figure 3-1

Endo- vs. Exo-chitinolytic activity

Endochitinases are able to cleave a chitin polymer at any glycosidic linkage other than the first at the non-reducing end. This activity usually results in the formation of long chitooligosaccharides from the chitin polymer. Exochitinases processively cleave chitobiose from the non-reducing end of free chitin chains or chitooligosaccharides.

Endo-Chitinase

Random cutting at any linkage



1st

2nd

3rd

Exo-Chitinase

Processive cutting of chitobiose from free ends

Figure 3-2

Comparison of the domain architectures of *M. degradans* ChiB and *T.*

kodakaraensis Tk-ChiA

ChiB of *M. degradans* contains two complete GH18 catalytic domains separated by a polyserine linker and a threonine/glutamic acid linker. A putative lipoprotein acylation site is present at the amino terminus and is separated from the first catalytic domain by a long polyserine linker. *Tk*-ChiA contains a type II secretion signal followed by a chitin-binding domain, a GH18 domain, two more chitin-binding domains, and a second GH18 domain. In *Tk*-ChiA, the first GH18 domain is exolytic, releasing chitobiose and some chitotriose from free chitin strands, while the other catalytic domain is endochitinolytic. Triangle, chitin-binding domain; X, putative lipoprotein acylation site; GH18, Glycoside Hydrolase family 18 catalytic domain.

T. kodakaraensis Tk-ChiA



M. degradans ChiB



Materials and Methods

Chemicals and Reagents

Standard reagents, chitooligosaccharides, methylumbelliferone substrates, and chitin were obtained from Sigma (St. Louis, MO). Ethylene glycol chitin was purchased from Fisher Scientific (Pittsburgh, PA). Ni-NTA agarose was obtained from Qiagen (Valencia, CA). Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Beverly, MA). Bugbuster™ NT and pETBlue2 were obtained from Novagen (Madison, WI).

Cloning and Expression of GH18_N and GH18_C

Oligonucleotide primers were designed to amplify the nucleotide sequence corresponding to each catalytic domain by PCR using purified *M.degradans* genomic DNA as a template. Primer sequences: GH18_N-F (468), CTTGGCGCGCCATGGTGTAGATGCCGAATTG; GH18_N-R (1924), CCGGGTACCGTTGTCTTCGTAATTGCCTTC; GH18_C-F (2512), CTTGGCGCGCCATGGCGAAACAGATTTAG; GH18_C-R (3800), CCGGGTACCCTGCTTTTCGTTGCCGAA; restriction sites underlined, relative position of the 5' nucleotide start of each primer within the *chiB* sequence is shown in parentheses. GH18_{N+C} was created using primers GH18_N-F (468) and GH18_C-R (3800). Each amplified fragment was then digested with the appropriate restriction enzymes and ligated into the protein expression vector pETBlue2 using T4 DNA ligase. Expression constructs were verified by sequencing and transformed into *E.coli* Tuner™ DE3(pLacI) cells. Protein expression was performed according to the

manufacturer's protocol. Cells were lysed with BugBuster™ NT lysis buffer, centrifuged, and the supernatant collected. Supernatants containing recombinant enzymes were applied to a Ni-NTA agarose column and purified according to the manufacturer's protocol for native protein purification. Purified enzyme samples were quantified using a BSA protein quantification kit (Pierce, Rockford IL).

Glycol Chitin Zymography

Ethylene glycol chitin was incorporated into the separating portion of an SDS-PAGE gel to a final concentration of 0.01%. After fractionation of the proteins, the zymogram was incubated in refolding buffer (50 mM Tris-Cl, 1mM EDTA, 5 mM 2-mercaptoethanol, pH 7.5) overnight at 4°C and subsequently analyzed for chitin depolymerase activity as described in the previous chapter.

Enzyme assays using chitin analogs

Solutions of 4'-methylumbelliferyl-N-N'-diacetylchitobiose [MUF-diNAG] and 4'-methylumbelliferyl-N-N'-N''-triacetylchitotriose [MUF-triNAG] were prepared in 50 mM sodium phosphate buffer (pH 7.0). Reaction mixtures contained 2 µg of purified enzyme and 30 µM analog solution. After incubation for 5-10 minutes at 37°C for GH18_N or 5-20 minutes at 30°C for GH18_C reactions were stopped by submersion in an ice water bath. Liberated methylumbelliferone was detected using a Hoefer TKO-100 fluorometer. The reaction was measured at multiple time points between 5 and 20 minutes and was found to be linear, with less than 10% of the substrate being degraded.

Oligosaccharide electrophoresis

Labeling reactions were incubated with 2 volumes of labeling solution (1.0 M sodium cyanoborohydride, 0.2 M 2-aminobenzoic acid) at 65°C with frequent shaking for 2 hours and then dried under vacuum. Each sample was resuspended in standard 2X SDS-PAGE loading buffer and fractionated in a 15% polyacrylamide gel at 45 mA constant current. Labeled oligosaccharides were visualized under UV light.

Determination of Reaction Optima for each Domain

MUF-diNAG or MUF-triNAG was added to 20 μ g of purified enzyme and incubated at a given pH or temperature and activity detected as described above. The buffers used were: sodium acetate (pH 4.0-5.5), MES (5.5-6.5), PIPES (6.5-7.0), HEPES (7.0-8.0), and Tris base (8.0-9.5). For a given enzyme, reaction conditions that permitted maximum activity were assigned a value of 100%. Where indicated, EDTA, EGTA, KCl, NiCl₂, SrCl₂, MgCl₂, MnCl₂, CuCl₂, CaCl₂, or HgCl₂ were added to reaction mixtures to a final concentration of 10 mM; NaCl was added at concentrations up to 1.0 M. Reactions containing metal ions contained 200 pmol enzyme and were incubated for ten minutes at 37°C for GH18_N or twenty minutes at 30°C for GH18_C.

Enzyme assays using chitin and chitin derivatives

Purified enzyme and substrate (2mg chitin or 10 nmol chitooligosaccharide) were added to 50 mM HEPES, pH 7.5 and incubated at 30°C. The amount of reducing

sugar generated was determined by the DNSA assay as described. Specific enzyme activity was estimated by comparison to a standard curve.

Protein sequence analysis

Analysis of protein domains was performed using the Simple Modular Architecture Research Tool. Similarity between proteins and protein domains was determined by the BLAST algorithm. The lipoprotein anchoring site within ChiB was identified using the Database of Bacterial Lipoproteins. The nucleotide and protein sequence of ChiB have been placed in GenBank under the accession BK001042.

Results

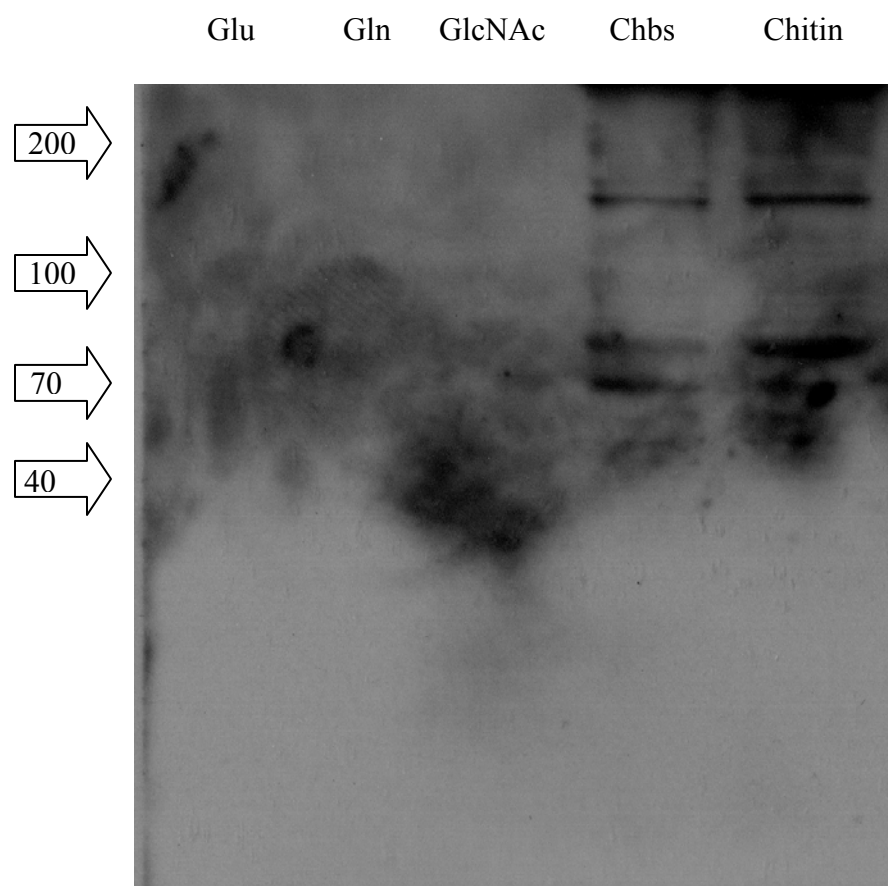
ChiB is expressed in response to different chitins but not GlcNAc

To confirm that ChiB is expressed and to determine what stimuli upregulate the expression of this enzyme, a polyclonal antibody prepared against ChiB was used in immuno-blots. *M. degradans* was grown in minimal media containing various chitins or chitin derivatives as the sole carbon source. Chitin and chitobiose induced the expression of ChiB; glucose, glucosamine, and GlcNAc did not function as inducers (Figure 3-3). Similar to other bacterial chitinolytic systems, the presence of glucose in a culture containing chitin repressed the expression of ChiB.

Figure 3-3

Expression of ChiB is dependent on carbon source

Chitin or chitooligosaccharides often function as a signalling molecule to upregulate the expression of genes for chitinolytic enzymes. It was also feasible that some of the chitin depolymerases of *M. degradans* might only be upregulated under certain conditions. The expression of ChiB was monitored using different carbon sources supplementing minimal media to a final concentration of 0.2%. Mid log phase cells were collected, lysed in loading buffer, and separated in an SDS-PAGE gel. A Western-blot was performed using a polyclonal antibody raised against ChiB. The estimated weight of the highest band in the following blot is in good agreement with the predicted weight of ChiB. Sole carbon sources used to culture cells: Lane 1, Glucose (Glu); Lane 2, Glucosamine (Gln); Lane 3, GlcNAc; Lane 4, Chitobiose (Chbs); Lane 5, Chitin. Bands between 70 and 100 kDa are likely degradation products of ChiB.



ChiB is predicted to contain two catalytic domains and a lipoprotein acylation site

ChiB was previously predicted to contain two catalytic domains based on conserved sequence motifs (54). The first catalytic site, GH18_N, was identified in the amino-terminal region of ChiB (residues 221-605). It consists of 385 amino acids and is most similar to the GH18 domain of the exo-chitinase ChiB of *Serratia marcescens* (S52422) (55% I, 69% S). A second predicted catalytic site, GH18_C, is present in the carboxy-terminal domain of ChiB (residues 860-1254). This domain is composed of 395 amino acids and is most similar to a chitinase from *Vibrio* sp. strain 5SM-1 (AAL46648) (49% I, 66% S). The two GH18 domains of ChiB only share 29% identity and 42% similarity when aligned at the amino acid level. GH18_N and GH18_C include the motifs SVGGWAESN-X₃₃-FDGIDIDWEYP and SIGGWTMSTPF-X₂₆-FDGVDIDWEYP, respectively. These sequences are nearly identical to the consensus sequence that characterizes a GH18 domain and each also includes the key catalytic Glu residue (underlined).

ChiB contains a predicted lipobox within amino acid residues 16-19 composed of L-S-A-C. In addition, two positively charged residues are found within the first 5 amino acids (N at position 2 and K at position 5) and are separated from the lipobox by a hydrophobic stretch of ten amino acids. These characteristics satisfy the major criteria required for a lipoprotein secretion signal and acylation site (60, 81).

GH18_N and GH18_C independently depolymerize chitin

To determine if the GH18 domains of ChiB are catalytically active against chitin, the sequence corresponding to each domain (GH18_N, codons 156-641; GH18_C, codons 837-1266) was amplified by PCR and ligated into pETBlue2 to create

carboxy-terminal hexaHis fusions. The polypeptides were expressed in *E.coli* and purified on Ni-NTA agarose columns. The chitinolytic activity of each GH18 domain was tested using a glycol chitin zymogram. Consistent with their conserved sequence features, the ability of each catalytic domain to independently depolymerize chitin was observed in zymograms (Figure 3-4). Clear zones indicative of depolymerization were observed and corresponded to the predicted masses of the recombinant polypeptides (50.5 kDa, GH18_N; 47.7 kDa, GH18_C).

GH18_N and GH18_C differentially degrade chitin analogs

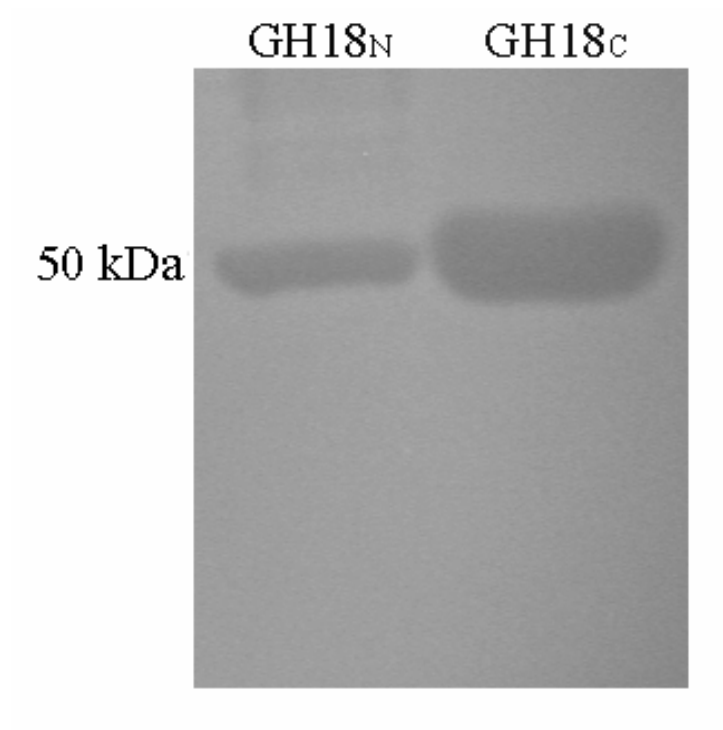
One possible explanation for the presence of two catalytic domains within ChiB is that each has a different role in the degradation of chitin, as was observed in an archaeal chitinase from *T. kodakaraensis* KOD1 (135). The chitin analogs MUF-diNAG and MUF-triNAG consist of chitobiose or chitotriose linked to a methylumbelliferone moiety at the reducing end that fluoresces under UV light only when cleaved from the saccharide (100). In theory, both exochitinases and endochitinases will hydrolyze the second glycosidic linkage from the non-reducing end of MUF-diNAG, thus releasing fluorescent MUF. Exochitinase activity on MUF-triNAG will result in the formation of chitobiose and non-fluorescent MUF-GlcNAc, while endochitinolytic activity can hydrolyze both the second and third glycosidic linkages of MUF-triNAG, thus releasing MUF.

Purified enzyme samples were added to solutions of either analog and the release of methylumbelliferone was monitored fluorometrically during the period of linear accumulation of product. When incubated with MUF-diNAG, the rate of MUF

Figure 3-4

Each GH18 domain of ChiB can independently degrade chitin

The nucleotide sequence corresponding to each GH18 domain was amplified using PCR and *M. degradans* genomic DNA as template. These DNA fragments were ligated into the vector pETBlue2 and expressed. Purified samples of peptides that included each domain were separated in a 0.01% glycol chitin zymogram, refolded, and incubated in activity buffer. After staining with Calcofluor, dark zones indicative of chitin depolymerase activity were observed.



release by GH18_N was 13.6-fold higher than that observed when GH18_C was utilized. However, when GH18_C was incubated with MUF-triNAG, the rate of MUF release was 2.7-fold higher than when incubated with GH18_N (Table 3-1). These results suggest that GH18_N may have exo-chitinase activity whereas GH18_C may have endo-chitinase activity.

The GH18 domains have similar reaction optima

The presence of two catalytic domains for the same substrate within a single enzyme is rare. If the dual domains of ChiB act together to degrade chitin, it would follow that these domains are most active under similar physical conditions. Ionic, pH, and temperature optima were determined for each domain. Purified samples of each enzyme were incubated with the optimal MUF substrate as identified above. GH18_N had a pH optimum between 7.2 and 8.0, while GH18_C was most active from 7.2 to 7.8 (Figure 3-5). The temperature optimum of GH18_N was determined to be 37°C, with retention of 80% activity at 30°C. GH18_C was most active at 30°C and retained only 67% of its activity at 37°C (Figure 3-6). A significant loss of activity was observed for each domain at temperatures above 40°C.

To examine the effect of ionic conditions on each domain, various chloride salts were added to reaction mixtures to a final concentration of 10 mM. The addition of Mg²⁺, Mn²⁺, Ca²⁺, K⁺, EDTA, and EGTA to 10 mM, and NaCl up to 1.0 M had no effect on the activity of either domain. The activity of GH18_N was reduced 36% by Ni²⁺, 8% by Sr²⁺, and 41% by Cu²⁺, while the activity of GH18_C was reduced 14% by

Table 3-1

Activity of polypeptides containing ChiB catalytic domains on different MUF analogs¹

	Rate of MUF release (nmol/min/mg enzyme)		Ratio of Rates²
	MUF-diNAG	MUF-triNAG	
GH18 _N	80 ± 1.04	5.9 ± 1.77	13.6
GH18 _C	6.9 ± 0.311	19 ± 0.429	0.363

¹ Polypeptides containing each active domain were purified as described in Materials and Methods and 2 µg incubated with the chitin analogs MUF-diNAG and MUF-triNAG. Reactions were stopped after 5-20 minutes and the fluorescence of the reaction measured. The rate of reaction was linear for at least 20 minutes. The data is the mean of three replicates and each experiment was repeated three times with similar results.

² Ratio of Rate of MUF release from MUF-diNAG / rate of release of MUF from MUF-triNAG.

Ni^{2+} , 5% by Sr^{2+} , and 53% by Cu^{2+} . Hg^{2+} completely inhibited the activity of both domains (Figure 3-7).

The GH18 domains have different activities on chitooligosaccharides

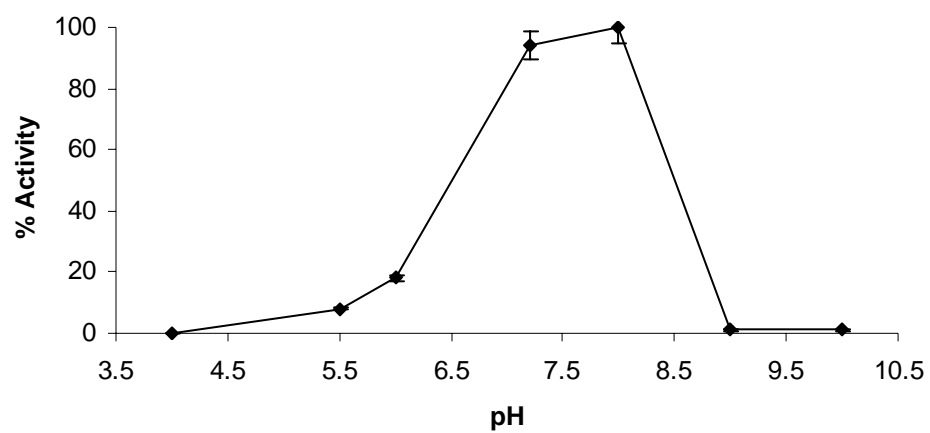
The products formed from the activity of GH18_N and GH18_C on native chitooligosaccharides was determined. Native chitooligosaccharides (GlcNAc₄₋₆) were incubated with purified samples of each polypeptide and degradation products were labeled with 2-aminobenzoic acid and fractionated by oligosaccharide gel electrophoresis. Consistent with exo-chitinase activity, the sole degradation product of GH18_N activity on GlcNAc₄ was chitobiose (Figure 3-8). Further, GH18_N primarily released chitobiose from GlcNAc₅ and GlcNAc₄ was not observed. When incubated with GlcNAc₆, GH18_N produced chitobiose and GlcNAc₄, but did not produce GlcNAc₃ or GlcNAc₅. In contrast, GH18_C produced a mixture of chitooligosaccharides when acting on GlcNAc₄, GlcNAc₅, and GlcNAc₆, consistent with the ability of an endochitinase to cleave a chitooligosaccharide at any glycosidic linkage after the first bond at the non-reducing end. When incubated with 2-aminobenzoic-labeled chitohexose, GH18_N produced a small amount of labeled GlcNAc₄ and an increasing amount of labeled chitobiose over time, consistent with processive degradation from the non-reducing end. GH18_C activity on pre-labeled chitohexose produced labeled GlcNAc₂, GlcNAc₃, and GlcNAc₄. The absence of labeled GlcNAc₅ suggests that the first glycosidic bond at the non-reducing end cannot be cleaved by this enzyme.

Figure 3-5

pH optimum of GH18_N and GH18_C

Purified samples of each recombinant protein were incubated with their preferred MUF substrate under different pH conditions. The fluorescence of the reaction was monitored and the condition which permitted maximal activity was assigned a value of 100%. While each domain had a slightly different optimum, each was maximally active over a range of pH values from 7.2 to 8.0, regardless of the buffer system used.

pH Optimum of GH18-N



pH Optimum of GH18-C

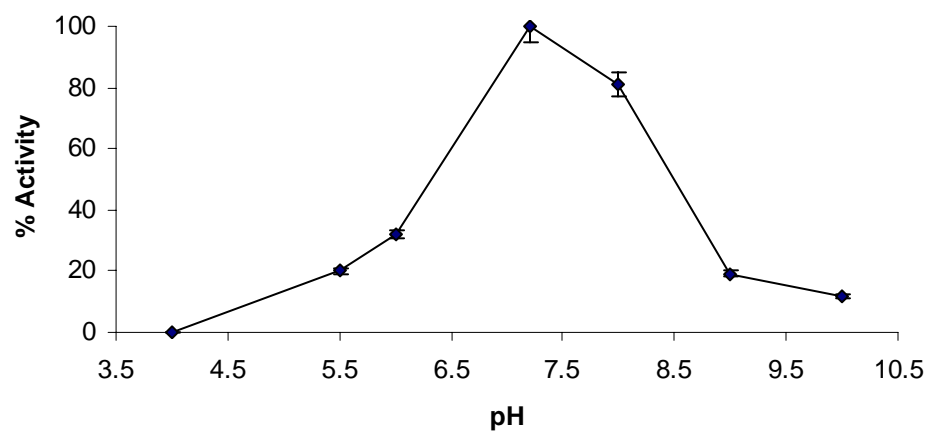
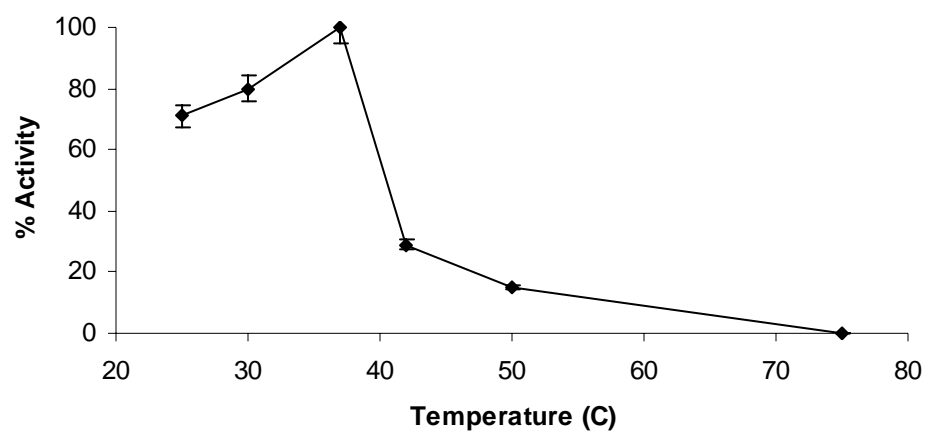


Figure 3-6

Temperature optimum of GH18_N and GH18_C

Purified samples of each recombinant protein were incubated with their preferred MUF substrate under different temperature conditions. The fluorescence of the reaction was monitored and the condition which permitted maximal activity was assigned a value of 100%. While each domain had a slightly different optimum, each was maximally active over a range of temperatures from 30°C to 37°C. Neither domain was thermostable.

Temperature Optimum of GH18-N



Temperature Optimum of GH18-C

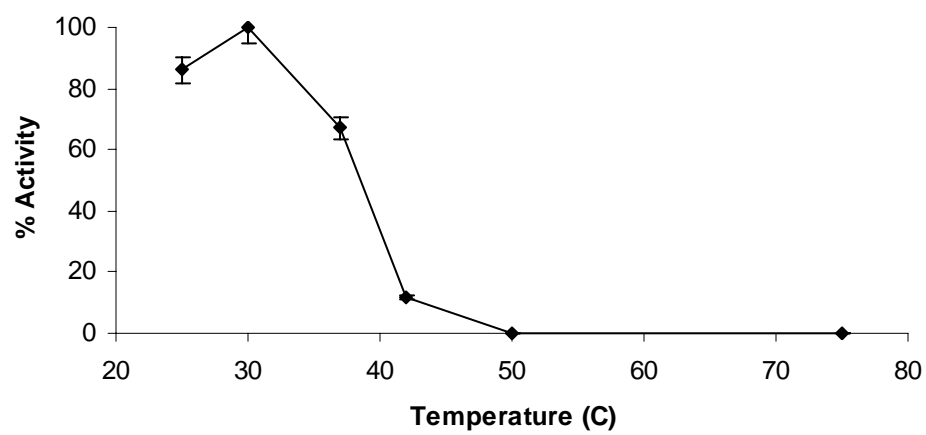


Figure 3-7

Effect of different ionic conditions on the activity of GH18_N and GH18_C

The effect of various metal ions, EDTA and EGTA up to 10 mM, and NaCl up to 1.0 M, on the activity of each catalytic domain was determined. Reactions consisted of recombinant samples of each protein at their optimal pH and temperature with their preferred MUF analog. As a control, a sample to which no metal ions were added was measured and assigned a relative value of 100%. Grey bars, GH18-N; White bars, GH18-C.

Effect of Ionic Conditions on GH18-N and GH18-C

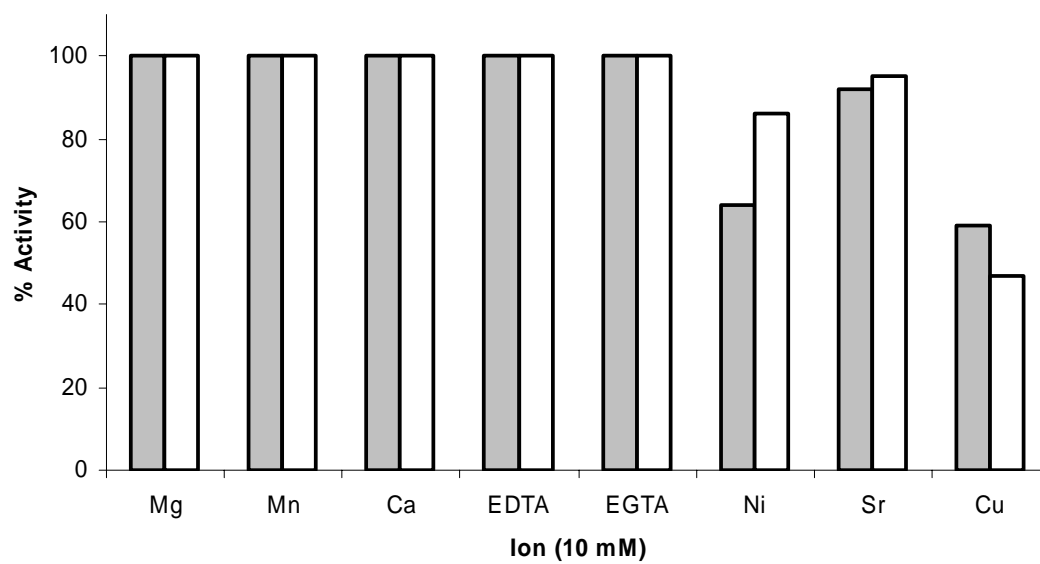
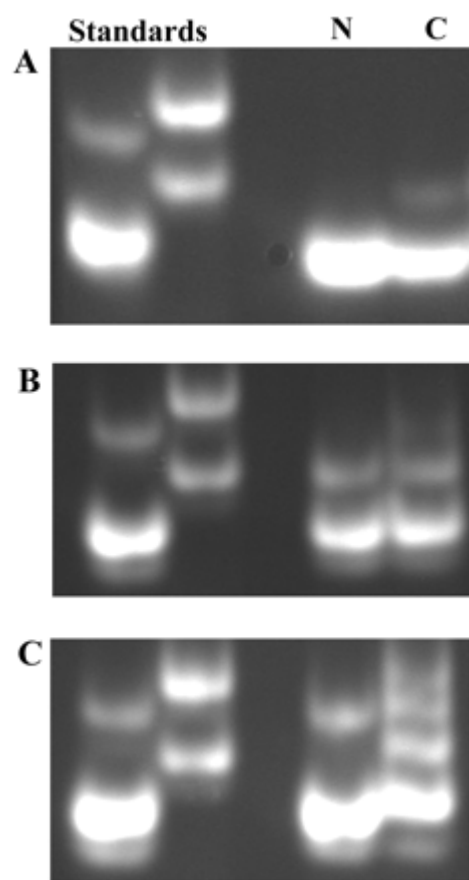


Figure 3-8

Reaction products released from native chitin by the activity of each catalytic domain

Purified samples of each enzyme were incubated with chitotetraose (A), chitopentose (B), or chitohexose (C). To determine the chitooligosaccharides released from the substrates by each enzyme, the reactions were stopped by boiling, dried under vacuum, and the reducing ends of the degradation products were labeled with para-aminobenzoic acid. The labeled end products were separated in a 15% acrylamide gel and visualized under UV light. Standards in each gel, from top to bottom (lanes staggered for clarity): Chitopentose, chitotetraose, chitotriose, chitobiose. N, GH18_N; C, GH18_C.



GH18_N and GH18_C function cooperatively to degrade native chitin

The impact of the differential activities in a single reaction mixture, both when the catalytic domains are linked on a single polypeptide and when expressed as separate enzymes, was examined. Equivalent amounts (250 pmol) of GH18_N or GH18_C were added individually to native chitin to determine the rate at which each could release reducing sugars, an indication of depolymerization. GH18_N released 0.0158 μ mol reducing sugar/min when added to native chitin, whereas GH18_C released 0.0340 μ mol/min (Table 3-2). A similar rate was measured at multiple time points during the initial 30 minutes of each reaction.

To determine if the active domains function cooperatively to degrade chitin, equivalent amounts of each polypeptide were added to native chitin in a single reaction. If the domains act independently of each other, the theoretical combined rate of degradation should be greater than the sum of the two independent activities calculated above, i.e., 0.0498 μ mol reducing sugar/min/500 pmol total protein. Consistent with the proposed endo- and exo-activity of each domain, the actual rate was 140% higher than the theoretical rate (Table 3-2).

Because in their native state the domains are linked on a single polypeptide, the rate of depolymerization was also measured when both catalytic domains were present and fused by their native linkage. Full-length enzyme could not be used in these experiments because of difficulties in expressing the complete protein, possibly due to the serine-rich, 150-residue linker region at the amino terminus. The truncated form of ChiB lacking the postulated lipoprotein-anchoring site and linker region, GH18_{N+C}, was used instead (residues 156-1266). GH18_{N+C} released 0.0645 μ mol

Table 3-2

Activity of polypeptides containing ChiB catalytic domains on native chitin

Polypeptide components of reaction (pmol)			Rate of Depolymerization^a
GH18_N	GH18_C	GH18_{N+C}	
250			0.0158 ± 0.0015
	250		0.0340 ± 0.0017
250	250		0.1190 ± 0.0081 ^b
		250 ^c	0.0645 ± 0.0031

^a Rate given as □mol reducing sugar released/min/ 250 or 500 pmol catalytic domains depending on the reaction components.

^b The theoretical calculated rate (0.0498 μmol/min/500 pmol catalytic domains) was determined by adding the observed rate for each domain acting alone and assuming no cooperative interaction.

^c Each molecule of GH18_{N+C} contains two catalytic domains, therefore 500 pmoles of catalytic domains were present.

reducing sugar/min when 250 pmol of polypeptide were added (and therefore 250 pmol of each active domain), an increase of 23% over the theoretical combined rate (Table 3-2).

Discussion

Analysis of the chitinolytic system of *M. degradans* revealed an unusual chitin depolymerase, ChiB, which appeared to include two catalytic domains. One of the catalytic domains of ChiB is shown here to function as an endochitinase while the other functions as an exochitinase. ChiB is the first Eubacterial chitinase demonstrated to contain two functional GH18 catalytic domains. The lack of carbohydrate binding domains and typical accessory domains (e.g., Fibronectin Type III domains, Polycystic Kidney Disease domains) coupled with the discrete activities of each catalytic domain emphasize the novelty of this enzyme.

When expressed as separate polypeptides, each GH18 domain of ChiB was able to depolymerize chitin in zymograms and was most active under similar temperature, pH, and ionic conditions. GH18_N was more active on MUF-diNAG than MUF-triNAG and displayed a pattern of activity typical of an exochitinase on chitooligosaccharides. Chitobiose was released from the non-reducing end of GlcNAc₄ - GlcNAc₆. Conversely, GH18_C released MUF most rapidly from MUF-triNAG and was able to cleave chitooligosaccharides at multiple linkages, demonstrating endochitinase activity. GH18_C was more than twice as active on native chitin as GH18_N; because native chitin has a paucity of free, exposed ends,

exochitinases have far fewer sites at which they can act as compared to random cutting endochitinases that can cleave virtually any glycosidic linkage in the polymer. The synergistic degradation of chitin observed when both domains were present further supports their proposed function. The presence of both domains on separate polypeptides increased the release of reducing sugars 140% (2.4 fold) over the theoretical combined rate calculated if the domains were only to act additively. This synergism would not be observed if both domains had the same activity.

Carbohydrases with two catalytic domains are rare among prokaryotes. Only a small number have been characterized, mostly from ruminants and thermophiles. For example, *Ruminococcus flavefaciens* 17 (34) and *Fibrobacter succinogenes* S85 (105), produce xylanases with two catalytic domains, though the latter appears to encode a xylanase with two domains of the same function. Two extreme thermophiles, *Anaerocellum thermophilum* (a \square -subgroup proteobacterium) and *Thermococcus kodakaraensis* KOD1 (an archaeon), produce enzymes with two catalytic domains (134, 158). *A. thermophilum* produces a cellulase with separate GH9 and GH48 domains that encode for endo- and exo-glucanase activity, respectively. A chitinase from *T. kodakaraensis*, Tk-ChiA, was shown to have an amino-terminal exochitinase domain, while the carboxy-terminus contains an endochitinase domain. Unlike ChiB of *M. degradans*, this enzyme also contains chitin-binding domains and is not predicted to anchor to the cell surface. Further, the exolytic domain of Tk-ChiA is able to weakly cleave the third glycosidic linkage from the non-reducing end of free chitin chains, an activity not observed in experiments with GH18_N.

The dual catalytic domains of ChiB function cooperatively to degrade chitin to chitobiose. Though maximal depolymerization was achieved when the catalytic domains of ChiB were on separate polypeptides, there are clear benefits to their presence as a single unit. First, a single promoter region is able to regulate the expression of two enzymatic activities. This permits two essential components of the chitinolytic system to be simultaneously regulated from a single locus, much like an operon regulating genes encoding a polycistronic mRNA. However, unlike an operon where several individual proteins are produced, a single enzyme is encoded. The amount of energy and secretion machinery needed to deliver two enzymatic functions to the exterior of the cell is therefore decreased. Second, encoding both activities on a single polypeptide ensures the proximity of the two domains during the *in situ* depolymerization of chitin. This allows for a synergistic and focused degradation of the polymer. In the environment, secreted enzymes may diffuse away from their intended targets and not be available to assist other components of a degradative system. This is partially solved by the presence of carbohydrate binding domains (which are not found within ChiB), but there is no assurance that both endo- and exo-acting enzymes will bind to the same location and have the opportunity to act in concert to achieve the full potential of the system unless linked on a single polypeptide.

When both domains were present on the same polypeptide, the synergism between the domains was less obvious. The activity detected when the domains are joined was only a modest increase (23%) over the theoretical activity when compared to the activity of the two catalytic domains as separate entities. The decreased activity

of the domains when linked may be the result of the domains then moving as a single protein as each encounters substrate. For example, as the exolytic domain is cleaving soluble chitooligosaccharides, perhaps away from the insoluble polymer, the endolytic domain is unable to contact, and therefore degrade, its primary substrate. One can envision that the amount of reducing sugars released would increase if the domains were free to act at different locations. However, such an arrangement may not be of benefit in nature where substrate is much more limited and less often encountered than in a laboratory reaction.

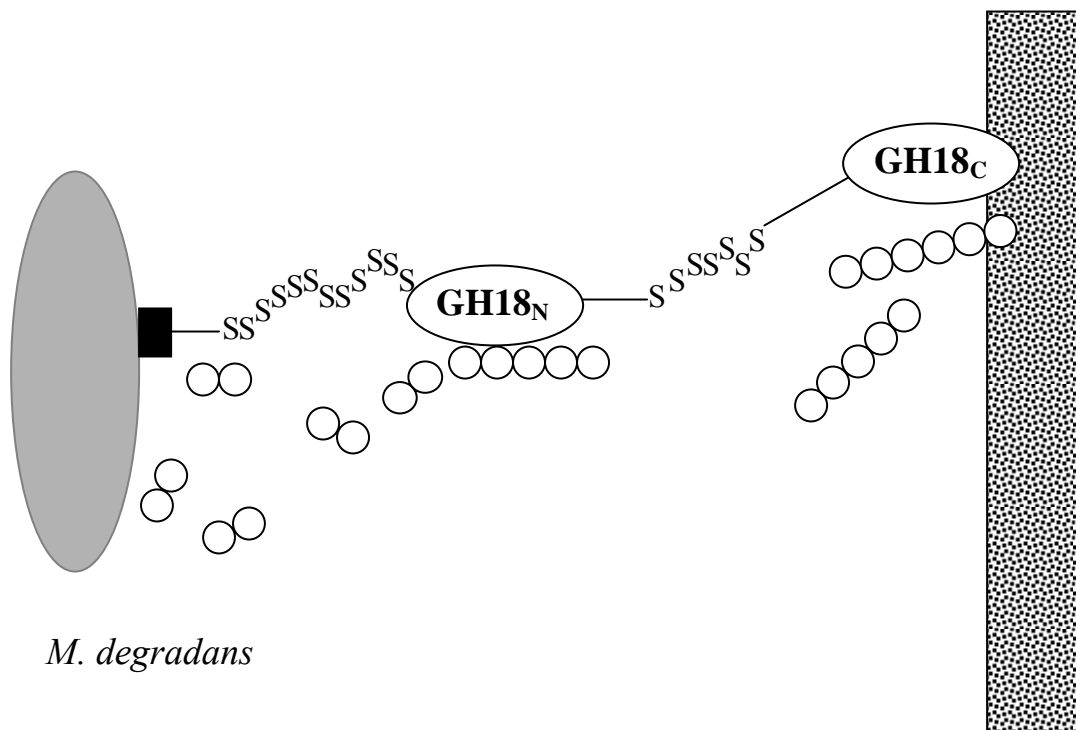
Based upon the data presented in this work and on the known properties of chitinases, a model of ChiB activity can be proposed (Figure 3-9). Each catalytic site has been shown to be independently active, so the linkage between the domains may prevent interference between them during the degradation of chitin. The significance of the repetitive sequence in this region is unclear. The processive cutting nature of exochitinases and random cutting behavior of endochitinases has been described (122) and can be applied to the activity model of ChiB. As GH18_C releases chitooligosaccharides from the polymer, they can be immediately acted upon by GH18_N which processively cleaves chitobiose from the non-reducing end. The lipoprotein acylation site present at the amino terminus of ChiB likely functions to anchor the enzyme to the outer membrane. This notion is strengthened by the observation that chitinase activity has been associated with outer membrane preparations of *M. degradans* (L. Whitehead and R. Weiner, unpublished observations). The membrane anchorage would keep two critical enzymatic activities in close proximity to the cell and perhaps forgo the necessity of chitin-binding

domains. If this is the case, the importance of the catalytic domain arrangement within ChiB becomes apparent; chitooligosaccharides released by the activity of the distal GH18_C can be transferred to the exo-acting domain, which is in close proximity to the outer membrane where newly formed chitobiose can be taken up by the cell.

Figure 3-9

Model of chitin depolymerization by ChiB

ChiB is likely to attach to the surface of the cell via a lipoprotein anchor (black box). Activity of the endochitinolytic GH18_C releases chitooligosaccharides from polymeric chitin (speckled box). Free chitooligosaccharides (small circles) are then acted upon by the exochitinolytic GH18_N that processively releases chitobiose from the non-reducing end. Free chitobiose would then be taken up by the cell and metabolized. The polyserine linkers (SSS) may provide flexibility to the enzyme and optimize interaction with substrates. Note that cartoon is not drawn to scale.



Chapter 4: Polyserine Linkers Among *Microbulbifer degradans* Carbohydrases and Other Prokaryotic Proteins

The data in this section appeared in *Prot Sci* 13(5): 1422-25.

Introduction

Functional domains (e.g., catalytic and binding domains) within some prokaryotic carbohydrases are separated by linker regions consisting of simple or repetitive sequence rich in proline, threonine, serine, and/or glycine (10, 40, 70, 127, 140). These linker domains have been identified in a diverse range of microbes, including gram negative, gram positive, terrestrial, and marine bacteria, as well as archaea. Domain linkers have been hypothesized to provide proteins with a flexible hinge or to increase the distance between active domains, presumably to optimize interaction with substrates (12, 19, 32, 117). Threonine/proline rich domains have been found in prokaryotic carbohydrases and shown to function as sites of posttranslational modification, namely glycosylation. The addition of sugars to these sites appears to protect secreted carbohydrases from protease activity. Polyserine linker domains [PSLs] are thought to be extremely flexible, in stark contrast to the more common proline rich linker sequences which are predicted to have extended, rigid conformations (127).

While linker domains located between active sites of carbohydrases are not uncommon among prokaryotes, linkers composed predominantly of serine are rare. The soil bacterium *Cellvibrio japonicus* (formerly *Pseudomonas fluorescens* subsp.

cellulosa) had been the only organism known to encode multiple carbohydrases containing PSLs, as extensively characterized by Gilbert *et al* (16, 43, 62, 84, 87). *C. japonicus* enzymes that contain PSLs include cellulases, xylanases, a rhamnogalacturonan lyase, pectate lyases, and mannanases. In these proteins, PSLs are only located between functional domains. The deletion of the PSLs in XylA and XylC of *C. japonicus* decreases activity on xylan/cellulose complexes, but on not soluble substrates.

C. japonicus was the only prokaryote known to have more than two PSL proteins (with a current total of 13 sequenced and characterized PSL enzymes) until the genome sequence of *M. degradans* became available. Forty six secreted carbohydrases and related proteins in *M. degradans* contain PSLs that separate functional domains. This presents an ideal opportunity to examine the complete distribution of PSLs within a given prokaryote because the draft genomic sequence is available and its complement of carbohydrases is unparalleled.

Materials and Methods

PSL containing proteins were identified using protein sequences based upon the translated nucleotide sequences of 140 completed microbial genomes and, where possible, the 125 unfinished microbial genomes found at the NCBI microbial genome homepage (<http://www.ncbi.nlm.nih.gov/genomes/MICROBES/Complete.html>). Non-redundant, annotated protein sequence databases were searched for PSL proteins using the PIR pattern/peptide match program at the Protein Information Resource server (<http://pir.georgetown.edu/>). The domain architecture of each PSL protein was

analyzed using the Simple Modular Architecture Research Tool (<http://smart.embl-heidelberg.de>). Type II secretion signals were identified using the iPSORT program (<http://www.hypothesiscreator.net/iPSORT>) and lipoprotein acylation sites were identified at the DOLOP website (<http://www.mrc-lmb.cam.ac.uk/genomes/dolop>).

Results

Forty six M. degradans proteins contain PSLs

The *M. degradans* genome was searched for open reading frames encoding five or more serine residues in tandem. Forty six genes were identified that encoded proteins with PSLs; 18 contain a single PSL, while 28 have two or more (Table 4-1). These domains have an average length of 39 residues and an average composition of 79% serine, 11% glycine, 7% threonine, and 3% alanine; glycine residues are predominantly located in the sequence immediately flanking tracts of polyserine sequence. Several of the PSLs also contain a single Asp or Cys residue. Though serine is the predominant residue within each domain, none are identical in terms of exact residue composition or sequence. The nucleotide sequences corresponding to each of the domains are even less similar and each is unique within the genome. Each of the six codons for serine are used to encode the PSLs and none are used preferentially, nor are they arranged in any obvious pattern or repeat. No other sequences encoding domains consisting of predominantly one amino acid are found in the *M. degradans* genome.

Table 4-1

***M. degradans* proteins that contain polyserine linkers**

<i>M. degradans</i> PSL Protein Accession	Mw	PSL start residue (length)^a	Domains separated by PSL^b	Most Similar Protein	Source Organism of Most Similar Protein	Accession of Most Similar Protein
Cellulases						
ZP_00064659	81.9	144 (52) 286 (93)	CBD2 - CBD2 CBD2 - GH6	beta-1,4-exocellulase (e ⁻¹³¹)	<i>Thermobifida fusca</i>	AAA62211
ZP_00066178	67.9	343 (21) 511 (33)	GH5 - CBDIV CBDIV - [92]	Cellulase (e ⁻¹¹⁸)	<i>Pseudomonas sp.</i> <i>ND137</i>	BAB79288
ZP_00064986	65.9	127 (62) 240 (40)	CBD2 - CBD10 CBD10 - GH5	Cellulase (e ⁻⁷⁰)	<i>Bacillus sp. BP-23</i>	CAA73113
ZP_00068260	66.9	338 (33)	GH5 - CBD6	Cellulase (e ⁻¹³⁴)	<i>Pseudomonas sp.</i> <i>ND137</i>	BAB79288
ZP_00065776	89.5	608 (58) 717 (52)	GH9 - CBD10 CBD10 - CBD2	Carboxymethyl - cellulase (0.0)	<i>Cellvibrio japonicus</i>	CAA31082
ZP_00064853	60.7	36 (58)	LPB - CBDIV	Cellulase H (6e ⁻⁷⁰)	<i>Clostridium</i> <i>thermocellum</i>	JH0157
ZP_00067454	49.1	31 (47)	LPB - GH5	Cellulase G (7e ⁻⁵⁹)	<i>Fibrobacter</i> <i>succinogenes</i>	AAB38548
ZP_00067367	77.2	122 (40) 214 (58)	CBD2 - CBM10 CBD10 - GH5	Cellodextrinase CelC (0.0)	<i>Cellvibrio japonicus</i>	CAA43597
Pectate Lyases						
ZP_00067834	78.9	358 (36) 511 (30)	AmbAll - [96] [96] - FN3	Pectate Lyase (e ⁻¹⁶⁹)	<i>Thermobifida fusca</i>	ZP_00056995

ZP_00067832	136.4	363 (31)	AmbAll - [96]	Pectate Lyase	<i>Thermobifida fusca</i>	ZP_00056995
		520 (38)	[96] - FN3	(e^{-169})		
		782 (63)	CBDIV - AmbAll			
ZP_00068232	73.3	151 (46)	CBD2 - CBD6	Pectate lyase 10A	<i>Cellvibrio japonicus</i>	AAG29353
		342 (33)	CBD6 - RepDom	(e^{-175})		
ZP_00067017	42.9	89 (30)	[86] - FN3	Pel-28K ($3e^{-38}$)	<i>Bacillus</i> sp. P-2850	BAC11009
		204 (19)	FN3 - PecLyase			
ZP_00065374	n/a ^c	1 (13) ^e	n/a ^c	Pectate Lyase (e^{-151})	Uncultured bacterium	AAP70368
		160 (50)	CBDIV - AmbAll			
ZP_00064694	52.3	117 (37)	CBDII - CBDIV	Esterase D (e^{-65})	<i>Cellvibrio japonicus</i>	CAA41727
		291 (19)	CBDIV - PecLyase			
ZP_00064697	81.5	134 (51)	CBDII - CBD6	Pectate Lyase (e^{-138})	<i>Pirellula</i> sp.	CAD74167
		321 (34)	CBM6 - AmbAll			
ZP_00066059	60.8	32 (50)	LPB - CBM6	Pectate Lyase 10A	<i>Cellvibrio japonicus</i>	AAG29353
				(e^{-90})		
ZP_00066061	75.6	89 (59)	[77] - PecLyase	Pectate Lyase	<i>Erwinia chrysanthemi</i>	CAA57041
				(e^{-125})		
ZP_00065735	46.1	89 (33)	[67] - FN3	Pectate Lyase	<i>Erwinia carotovorum</i>	CAA55814
		206 (21)	FN3 - PecLyase	($6e^{-89}$)		

Xylanases

ZP_00067027	77.7	645 (42)	CBM_4_9 - Transmembrane domain	1,3-glucanase (e^{-19})	<i>Lysobacter enzymogenes</i>	AAN77505
ZP_00067229	129.6	611 (27)	CBDIV - CBDII	Beta-xylosidase	<i>Clostridium thermocellum</i>	ZP_00060110
		724 (25)	CBDII - CBD_4_9	(e^{-73})		
ZP_00067071	80.8	224 (22)	GH11 - RepDom	1,4-xylanase (0.0)	<i>Cellvibrio japonicus</i>	CAA88763
		360 (15)	RepDom - RepDom			
ZP_00066073	65.0	131 (47)	CBD2 - CBD10	Xylanase A (e^{-108})	<i>Cellvibrio japonicus</i>	P14768
		231 (50)	CBD10 - GH10			

ZP_00066395	61.7	310	(28)	GH10 - CBDIV	Xylanase B (e^{-78})	<i>Cellvibrio japonicus</i>	CAA38389
		482	(29)	CBDIV - CBD_5_12			

Other Carbohydrases

ZP_00064900	65.4	305	(28)	CBDIV - [307]	Alginate lyase (e^{-71})	<i>Klebsiella pneumoniae</i>	AAA25049
DAA01335	56.3	77	(46)	ChtBD3 - ChtBD3	Chitinase A (e^{-117})	<i>Vibrio cholerae</i>	ACC72236
		181	(36)	ChtBD3 - GH18			
DAA01334	135.2	39	(28)	LPB - [15]	Chitinase B (e^{-113})	<i>Serratia marcescens</i>	CAA85292
		81	(18)	[15] - [12]			
		108	(80)	[12] - GH18			
		786	(53)	GH18 - GH18			
ZP_00065784	52.6	323	(25)	GH5 - CBD10	β -1,4-Mannanase (e^{-148})	<i>Pseudomonas sp. ND137</i>	BAB79290
		441	(32)	CBD10 - CBD10			
ZP_00065857	n/a ^c	1	(14) ^e	n/a ^c	endo-1,4-Mannanase 26B (e^{-125})	<i>Cellvibrio japonicus</i>	AAO31762
ZP_00065570	57.6	337	(37)	GH5 - CBD10	1,4-Mannanase 5B (e^{-144})	<i>Cellvibrio japonicus</i>	AAO31760
		418	(50)	CBD10 - CBD2			
ZP_00066475	93.6	778	(39)	CBD6 - CBDIV	Rhamnogalacturonan lyase (0.0)	<i>Cellvibrio japonicus</i>	AAK20911

PSL Proteins of Uncertain Function

ZP_00066179	67.7	351	(32)	PecLyase - CBDIV	Cellulase (e^{-20})	<i>Pseudomonas ND137</i>	BAB79288
		523	(35)	CBDIV - [92]			
ZP_00068161	37.2	51	(20)	[31] - [207]	No Significant Similarity	n/a	n/a
		298	(37)	[207] - [32]			
ZP_00065412	44.2	284	(34)	CBD2 - ChtBD2	Hypothetical Prot. ($2e^{-18}$)	<i>Cytophaga hutchinsonii</i>	ZP_00119937
ZP_00066911	73.7	553	(33)	GH32 - CBD2	Hypothetical Prot. (e^{-63})	<i>Streptomyces avermitilis</i>	BAC70108

ZP_00066983	79.4	428 (23) 643 (17)	GH-like - [187] [187] - [95]	Hypothetical Prot. (4e ⁻⁷⁸)	<i>Bacillus cereus</i> <i>ATCC 14579</i>	AAP07701
ZP_00065369	103.7	30 (26)	LPB - [892]	Hypothetical Prot. (e ⁻¹¹²)	<i>Clostridium</i> <i>thermocellum</i>	ZP_00060315
ZP_00066022	71.4	379 (52)	PecLyase-like - CBD2	No Significant Similarity	n/a	n/a
ZP_00066068	84.8	145 (33) 220 (22)	CBD2 - [42] [42] - [546]	Hypothetical Prot. (7e ⁻²⁹)	<i>Clostridium</i> <i>thermocellum</i>	ZP_00060970
ZP_00066137	92.7	604 (37) 741 (40)	AmbAll - RepDom RepDom - RepDom	Alginate synthesis related protein (9e ⁻³⁰)	<i>Aquifex aeolicus</i> VF5	AAC07558
ZP_00065695	85.0	126 (23) 290 (22)	CBD2 - [132] [132] - [406]	Cellodextrinase C (4e ⁻²⁴) ^d	<i>Cellvibrio japonicus</i>	CAA43597
ZP_00065777	56.3	127 (50) 225 (106)	CBD2 - CBD10 CBD10 - [231]	Cellodextrinase C (2e ⁻³⁵) ^d	<i>Cellvibrio japonicus</i>	CAA43597
Not in database	17.1	19 (34)	LPB - Bacterial Ig-like dom.	No Significant Similarity	n/a	n/a
ZP_00065481	64.4	23 (26)	LPB - [540]	No Significant Similarity	n/a	n/a
ZP_00065540	80.3	285 (57)	FA58C - PecLyase	Putative Lyase precursor (2e ⁻²¹)	<i>S. coelicolor</i> A3(2)	CAB61820
ZP_00066396	97.5	125 (62) 336 (22)	CBD2 - [149] [149] - [575]	Hypothetical Prot. (e ⁻¹⁴³)	<i>Pirellula</i> sp.	CAD73232
ZP_00068207	72.5	23 (33)	LPB - CBDIV	Hypothetical Prot. (5e ⁻²³)	<i>Streptomyces</i> <i>avermitilis</i>	BAC71716

^a The borders of each PSL were defined as the beginning and end of consecutive serine, alanine, or glycine residues. Several linkers contain a single aspartic acid or cystine residue. The first amino acid in the PSL is defined as the start and its position in the peptide is given, as well as its length in residues.

^b CBD, Carbohydrate Binding Domain; ChtBD, Chitin Binding Domain; CBDIV, Cellulose Binding Domain, type IV; GH, Glycoside Hydrolase Domain; FN3, Fibronectin Type III Domain; LPB, lipoprotein box; AmbAll, Pectate Lyase catalytic domain; PecLyase, Pectate Lyase catalytic domain; RepDom, repeat domain; FA58C, discoidan domain. The length of domains of unknown function is given in brackets.

^c Sequence is incomplete at the N-terminus.

^d Similarity to these proteins is limited to a predicted carbohydrate binding domain.

M. degradans PSL proteins are modular, secreted carbohydrases

All 46 *M. degradans* PSL proteins are carbohydrate depolymerizing enzymes, carbohydrate binding proteins, or proteins with similarity to known proteins involved in carbohydrate degradation. These include 2 chitinases, 8 cellulases, 10 pectate lyases, 5 xylanases, 3 mannanases, a rhamnogalacturonan lyase, an alginate lyase, and 16 proteins of uncertain function. Among the 16 proteins for which no activity could be predicted, each had weak similarity to a known degradative enzyme or contained sequence similarity to known carbohydrate binding module [CBM] or catalytic domain. In cases where no sequence similarity was identified, the PSLs separate the proteins into segments large enough to contain presently unconfirmed catalytic sites or CBMs. Each of the 46 PSL containing proteins is predicted to contain a Sec-dependant, Type II secretion signal at the amino terminus.

In *M. degradans*, PSLs always separate predicted binding and/or catalytic domains. In no case did a PSL appear at the extreme carboxy-terminus of a protein, nor did two PSLs separate short segments of sequence with no predicted function. Interestingly, in nine proteins a PSL is located at the amino terminus following the secretion signal. Because in other cases PSLs are only located between functional groups, the amino terminus of each of these proteins was further analyzed. All nine of these proteins contain an apparent lipoprotein acylation site; *i.e.* each had at least one positively charged residue within the first five amino acids, a hydrophobic stretch of 8-10 residues, and a lipobox containing the appropriately conserved amino acids, including the requisite cysteine residue. In gram-negative bacteria, when the cysteine

residue within a lipobox is acylated the protein becomes anchored to the inner or outer membrane (81).

Analysis of M. degradans PSLs at the nucleotide level

Forty two of the 46 genes encoding PSL proteins are unique within the *M. degradans* genome sequence. The remaining four genes include two pairs of paralogs (Fig 4-1). The genes for two predicted pectate lyases (ZP00067834 and ZP00067832) exhibit greater than 75% identity among accessory domains (i.e., a carbohydrate binding domain and a Fibronectin Type III domain) and more than 80% identity between sequences corresponding to catalytic domains. However, the nucleotide sequence corresponding to the similarly located PSLs is less than 20% identical. Likewise, two cellulases (ZP00066178 and ZP00068260) also appear to have significant similarity at the nucleotide level except for their PSLs. The nucleotide sequences for the amino-terminal Glycoside Hydrolase family 5 domains and the predicted carbohydrate binding domains are more than 75% identical between genes, while the corresponding sequence for the PSL was dissimilar.

PSL proteins among other prokaryotes

Beyond the PSL proteins of *M. degradans* and *C. japonicus*, a small number of PSL proteins were identified during searches of the nonredundant database as well as complete and incomplete microbial genome sequences (Table 4-2). A cellulase from *Pseudomonas* sp. ND137 that contains a single PSL was identified. The only other sequences available for this organism are two xylanases, an agarase, and a

Figure 4-1

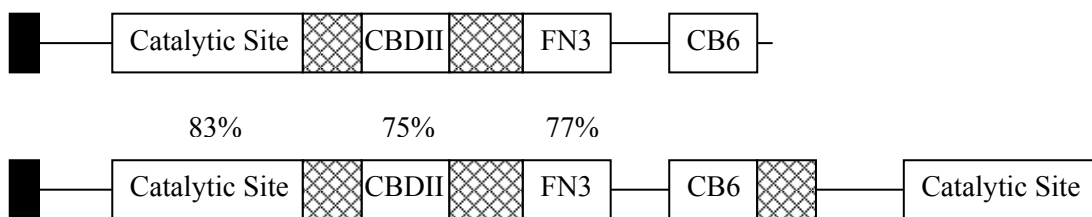
Similarity among two pairs of *M. degradans* genes that encode polyserine linkers

Similar sets of pectate lyases (ZP00067834, top; ZP00067832, bottom) (**A**) and cellulases (ZP00066178, top; ZP00068260, bottom) (**B**) are represented. The percentages indicate the degree of identity between the corresponding nucleotide sequences of each conserved domain. Regions not included within the boundaries of an indicated conserved domain share less than 10% identity at the nucleotide level.

Note that similarly located polyserine domains do not share nucleotide identity.

Black boxes, predicted secretion signals; CBDII, Carbohydrate Binding Domain, type 2; FN3, Fibronectin type III domain; CB6, Carbohydrate Binding Domain, type 6; hatched boxes, polyserine domains.

A



B

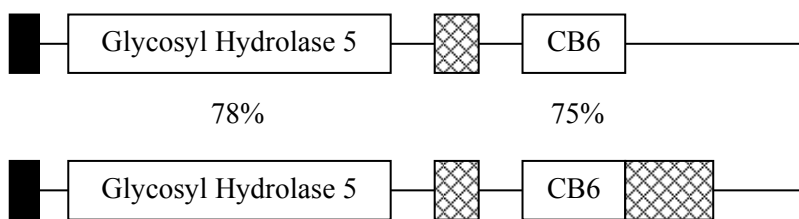


Table 4-2

Additional PSL proteins among prokaryotes¹

Source Organism	PSL Protein Annotation	PSL Separated Domains	Accession Number
<i>Proteins with demonstrated activity</i>			
<i>Bartonella quintana</i>	Hemin binding protein	-	AAM68130
<i>Cellvibrio japonicus</i>	Cellulase A	GH9 – CBM10 PKD – CBDII	P10476
	Cellulase B	CBDII – CBM10 CBM10 – GH45	P18126
	Cellulase E	GH5 – CBM10 CBM10 – CBMII	S56132
	Cellodextrinase C	CBDII – CBM10 CBM10 – GH5	CAA43597
	Esterase D	CBDII – CBM6 CBM6 – GH62	CAA41727
	Mannanase 26B	LPB – CBM10	AAO31762
	Mannanase 5B	GH5 – CBM10 CBM10 – CBDII	AAO31760
	Mannanase 5C	CBM10 – GH5	AAO31761
	Pectate Lyase 10A	CBDII – CBM6 CBM6 – Active Site	AAG29353
	Rhamnogalacturonan Lyase	Active site – FN3 FN3 – CBDII	AAK20911
	Xylanase A	CBM10 – GH10	CAA33469
	Xylanase B	CBDII – CBM6 CBM6 – GH10	CAA38389
	Xylanase C	CBDII – CBM6 CBM6 – GH62	P23031
<i>Cellvibrio mixtus</i>	Xylanase A	PDeac – CBM10	S59631
<i>Erwinia chrysanthemi</i>	OutD	-	B47021
<i>Pseudomonas</i> sp. ND137	Cellulase	CBDII - ?	BAB79288
<i>Ruminococcus albus</i>	1,4-glucanase	Dockerin – CBM_4_9	BAA92430
<i>Vibrio cholerae</i>	Hemagglutinin	-	AAA65694

<i>Hypothetical Proteins or Proteins Identified by Similarity</i>			
<i>Bacillus halodurans</i>	HP	-	BAB03774
<i>Bradyrhizobium japonicum</i>	Efflux protein (by similarity)	-	BAC48973
<i>Campylobacter jejuni</i>	HP	-	CAB75242
<i>Corynebacterium glutamicum</i>	HP	-	BAC00299
<i>Lactobacillus plantarum</i> WCFS1	Hypothetical Cell Surface Protein	-	NP_784025
	HP (lipoprotein)	-	CAD63604
	HP (lipoprotein)	-	AAK05819
<i>Pseudomonas fluorescens</i> PfO-1	Outer membrane protein (by similarity)	-	ZP_00088263
<i>Streptomyces avermitilis</i>	HP (lipoprotein)	-	BAC67844
<i>Xyella fastidiosa</i> strain <i>Temecula1</i>	Cellobiosidase	GH6 – CBDII	NP_778753
<i>Xylella fastidiosa</i> strain <i>9a5c</i>	Cellobiosidase	GH6 - CBDII	NP_298556

¹ PSL containing proteins were identified using protein sequences based upon the translated nucleotide sequences of 140 completed microbial genomes and, where possible, the 125 unfinished microbial genomes found at the NCBI microbial genome homepage (<http://www.ncbi.nlm.nih.gov/genomes/MICROBES/Complete.html>). Non-redundant, annotated protein sequence databases were searched for PSL proteins using the PIR pattern/peptide match program at the Protein Information Resource server (<http://pir.georgetown.edu/>). Abbreviations: CBM, Carbohydrate binding module; CBD, Carbohydrate binding domain; LPB, Lipoprotein box; PDeac, polysaccharide deacetylase domain; PKD, Polycystic kidney disease domain; GH, Glycoside hydrolase domain; HP, hypothetical protein; FN3, Fibronectin Type III domain.

mannanase, none of which contains domain linkers of any type. A cellobiosidase from *Xyella fastidiosa* strain *Temecula1* and another from *Xylella fastidiosa* strain *9a5c* were discovered, each of which contains a PSL separating a GH6 domain from a CBMII. The corresponding genes are greater than 86% identical over their entire length, with 87% identity at the amino acid level. No other proteins from these organisms contain PSLs. *Ruminococcus albus* encodes a 1,4-glucanase with a PSL between a catalytic and a binding domain. *Erwinia chrysanthemi* OutD, a pectic enzyme secretion protein, also contains a PSL, however it does not contain currently characterized domains. Several other hypothetical proteins or proteins identified only by sequence similarity contain polyserine domains (Table 4-2), though it could not be determined if they function as linkers because these proteins have no conserved domains. Interestingly, no proteins with PSLs were identified among archaea. Finally, several proteins from various sources were found to contain serine rich stretches at their extreme carboxy termini, but these domains are generally more heterogeneous in their residue composition than the PSLs defined in this work.

Discussion

Polyserine domains and serine rich proteins are much more common among eukaryotes than prokaryotes. The genomic sequence of *Homo sapiens*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, and *Plasmodium falciparum* reveal open reading frames encoding repetitive single amino acid sequences, including polyserine regions, in a diverse range of proteins. However, the function of PSLs in eukaryotes has not been established.

Although heterogeneous linker domains within prokaryotic carbohydrases have been recognized for several decades, reports of PSLs have been limited to one or two proteins from any single organism with the notable exception of *C. japonicus*. Our current searches of complete microbial genomes failed to identify a significant number of additional proteins that include PSLs. Even the available complete genome sequence of *Shewanella oneidensis* MR-1, a phylogenetic relative of *M. degradans* within Alteromonadaceae, does not reveal open reading frames encoding PSL proteins. Based upon thorough searches of existing prokaryotic databases, the known enzymes of *C. japonicus*, searches of the nonredundant database, and the considerable data afforded by analysis of the *M. degradans* genome, it is now possible to posit that in prokaryotes, PSLs are almost always found within secreted, complex polysaccharide depolymerizing enzymes or proteins involved in carbohydrate binding or metabolism. The function of PSLs as linkers is further supported by the observation that these domains are only located between predicted functional domains.

Genetic duplication appears to have resulted in genes for two additional PSL proteins within the *M. degradans* genome (Figure 4-1), but all other genes for PSL containing proteins are unique within the known sequence. In *C. japonicus*, the genes for XylB and XylC are located in tandem in the genome and appear to contain duplicate sequence in their amino-terminal region (43). This sequence encodes two polyserine linkers; the first is identical at the nucleotide level and amino acid level, while the second is partially duplicated. In XylC, the second polyserine linker becomes divergent after 15 codons, thus ending the identical sequence between the

proteins. In neither organism were there duplicate genes in which one contained polyserine linker sequence and the other did not. Thus, it does not appear that a known method of transposition or a recent, repetitive duplication event generated PSLs. Further, it does not appear that PSLs are extensively duplicated, as most organisms encoding PSLs do so only within a single protein.

Interestingly, eight of the *M. degradans* PSL proteins are most similar to *C. japonicus* enzymes wherein sequence, overall domain architecture, and PSL location are conserved. Horizontal transfer is known to play a role in the acquisition of new genetic material by bacteria, though it often occurs in specific eco-niches, such as the rumen (39, 94). It is unlikely that *C. japonicus*, a soil bacterium, and *M. degradans*, a marine bacterium, have recently shared a common environment. Thus, these genes may have been exchanged before each evolved to different habitats. Alternatively, these carbohydrases may have been inherited from a common ancestor. In either case, these domain arrangements have been conserved for a long period of time, suggesting that the placement of the domains and PSLs within each enzyme is functionally significant.

While *M. degradans* encodes 46 proteins with PSLs involved in complex carbohydrate degradation, it is postulated to contain nearly twice that number of extracellular carbohydrases wherein the domains are not separated by repetitive linking sequence. Similarly, *C. japonicus* also encodes carbohydrases that do not contain PSLs. The deletion of polyserine linkers from two *C. japonicus* xylanases decreased their activity on insoluble substrates, but does not altogether abolish their activity (13). Possibly relevant as well, threonine/proline rich linkers have been

shown to be dispensable with only moderate loss of activity (149). These observations suggest that PSLs are not absolutely required for carbohydrase function, but have evolved to enhance the activity of certain enzyme configurations, particularly during *in situ* degradation of ICP. Furthermore, among the cellulase and pectate lyase paralogs within *M. degradans*, the nucleotide sequence corresponding to PSLs are variable while the surrounding domains are not. Though the nucleotide sequences encoding PSLs are dynamic, their amino acid sequences are static; this suggests a specific, though currently unknown, role for serine within these linkers.

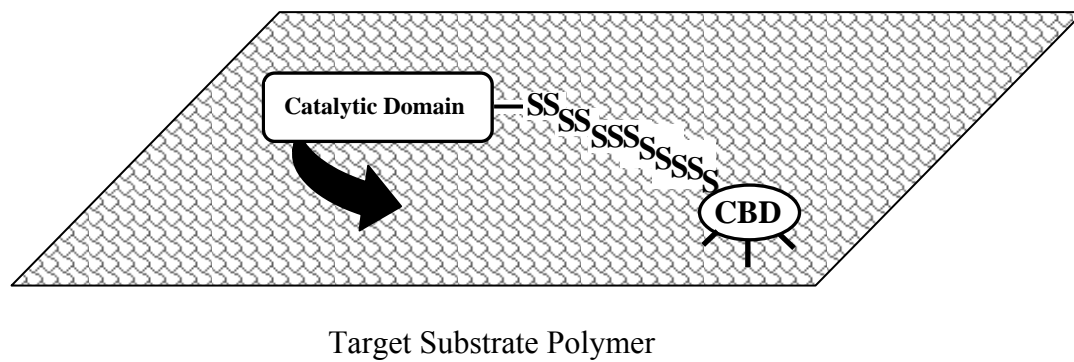
The degradation of ICP is not governed by conventional enzyme/substrate interactions. The direct contact of a carbohydrase with an ICP for a relatively extended period is required for the efficient degradation of these complex substrates. Therefore, a flexible linker region (such as a PSL) coupling a catalytic and a binding domain could expand the potential substrate target area available to the enzyme after the CBM makes contact with the polymer (Figure 4-2). Similarly, PSLs would enhance substrate availability to an enzyme anchored to a bacterial outer membrane, a potential survival advantage in the marine environment where diffusion and dilution are major factors affecting extracellular enzymes. This report is the first observation of PSLs separating an anchoring domain from the remainder of a protein. In nine *M. degradans* enzymes and in several hypothetical proteins from other organisms PSLs are located immediately after an amino-terminal lipobox, suggesting that PSLs can function to extend the catalytic and/or binding domains of a surface associated enzyme from the outer membrane.

M. degradans is unique among marine bacteria in its ability to degrade at least 9 ICP. Moreover, the draft genome sequence reveals over 130 putative carbohydrases involved in the degradation of these ICP. That 46 of these proteins contain PSLs and that they are limited to secreted enzymes involved in ICP degradation is an extremely interesting finding that underlies the importance of the PSL motif in carbohydrate catalysis in nature.

Figure 4-2

Polyserine linkers are flexible and may increase the range of motion of an enzyme

Carbohydrate binding domains are known to bind complex polysaccharides almost irreversibly. After a single binding event, the enzyme has a limited circular area on which it can act. Polyserine domains may increase this area of activity and also improve the interaction of an enzyme with a difficult substrate (such as complex plant matter) by enabling the enzyme to take on a favorable conformation.



Chapter 5: Other *Microbulbifer* species: Metabolic Profiles, Cloning of Select Genes, and Partial Enzyme Characterization

Introduction

The genus *Microbulbifer* was proposed by Gonzales *et al* in 1997 as a new genus within the gamma-3 subgroup of proteobacteria (42). The type strain, *Microbulbifer hydrolyticus*, has been joined in recent years by 14 additional species. *M. hydrolyticus* was isolated from coastal Georgia in an enrichment culture in which the predominant carbon source was lignin and other complex polysaccharides such as cellulose and hemicellulose. It was subsequently shown to metabolize cellulose, chitin, starch, and xylan. *M. hydrolyticus* forms surface protuberances and blebs in the presence of complex polysaccharides; these structures have been estimated to double the surface area of the outer membrane.

The genome of another related bacterium, *Microbulbifer degradans*, has been sequenced by the Joint Genome Institute in conjunction with the US Department of Energy. *M. degradans* is able to degrade and metabolize at least nine complex polysaccharides (CP) including agar, alginate, cellulose, chitin, pectin, xylan, and others, and similarly forms surface protuberances, blebs, and fibrils in response to CP (2, 151). Forty-six *M. degradans* enzymes involved in CP metabolism contain

flexible polyserine linkers that separate functional groups, a significant structural feature that is thought to optimize substrate/enzyme interactions. Linkers of this type are rare among prokaryotic proteins; only *M. degradans* and *Cellvibrio japonicus* are known to encode multiple enzymes with polyserine domains.

Beyond the genomic sequence of *M. degradans*, there is a paucity of research and sequence data on other members of this genus. An agarase from *Microbulbifer* sp. JAMB-A7 was recently characterized and the corresponding gene sequence submitted to GenBank, and the sequence of a DNA gyrase from *Microbulbifer cystodytense* has also been provided (101, 102). However, the predominant data available on this group of bacteria are derived from phylogenetic analyses and rDNA sequencing. Because these species appear to share the ability to degrade numerous CP (Table 1), they may play a significant role in carbon cycling within marine ecosystems. Clearly studies of the CP metabolizing systems of this genus could lead to the discovery of novel enzymes, new metabolic pathways used for CP metabolism, and novel methodologies utilized by marine bacteria for the efficient degradation of recalcitrant CP waste.

This work was undertaken to partially characterize selected CP degrading enzymes of *M. hydrolyticus* and to compare them to known degradative enzymes; three chitinases, a xylanase, and an 1,4- β -D-glucan glucohydrolase were identified. *M. hydrolyticus*, *M. elongatus*, and *M. salipaludis* were tested for the ability to degrade additional complex polysaccharides. The overall similarity of *M. hydrolyticus* enzymes to *M. degradans* proteins, the presence of polyserine linker

domains in some *M. hydrolyticus* enzymes, and the role of this genus in CP degradation are discussed.

Materials and Methods

Strains and growth conditions *M. hydrolyticus* (ATCC 700072^T) and *M. salipaludis* (JCM 11542) were grown with shaking at 200 rpm in Difco 2216 Marine Broth at 37°C. *M. elongatus* (ATCC 10144^T) was grown similarly at 25°C. *E.coli* DH5α (Invitrogen, Carlsbad CA), *E.coli* EC100 (Epicentre Technologies, Madison WI) and *E.coli* Tuner™ (DE3)(pLacI) (Novagen, Madison WI) were grown in LB broth or on LB agar at 37°C supplemented with the appropriate antibiotics when necessary (Kanamycin, 50 µg/ml; Chloramphenicol, 30 µg/ml; Ampicillin, 100 µg/ml).

DNA manipulations and Library Construction Genomic DNA was extracted from *M. hydrolyticus* using the Wizard™ Genomic DNA purification kit (Promega, Madison WI) according to the manufacturer's protocol. A genomic library was constructed using the EpiFos Fosmid Library Kit (Epicentre Technologies) according to the manufacturer's protocol. Briefly, genomic DNA was sheared by passage through a Hamilton syringe, end repaired, and then separated in an agarose gel. Fragments of approximately 40 KB in size were isolated and ligated into pEpiFos-1. After packaging into T1 phage particles, the library was transfected into *E.coli* EC100.

Chitinase, cellulase, and xylanase library screens LB agar was supplemented with the dye linked carbohydrates AZCL-Xylan or AZCL-HE-Cellulose (Megazyme, Inc., Bray, Ireland) to a final concentration of 0.08%. Library clones were plated on this media and incubated at 37°C. Clones producing xylanase or cellulase activity were surrounded by a bright blue halo. Alternatively, library clones were grown on LB agar for 24 h and overlaid with a 0.5% agarose solution containing 20 µM 4'-methylumbelliferyl-N-N'-diacetylchitobiose (MUF-diNAG), 4'-methylumbelliferyl-N-N'-N''-triacetylchitotriose (MUF-triNAG), or 4'-methylumbelliferyl-cellobioside (Sigma, St. Louis MO). Clones were monitored periodically over 24 h for activity; positive clones were surrounded by a bright blue halo under 366 nm UV light.

Fosmid purification and mutagenesis Fosmid library constructs were purified from clones of interest using a Plasmid Midiprep Kit (Biorad, Hercules CA). Fosmids were then mutagenized using a EZ::TN <Kan-2> *in vitro* insertion kit (Epicentre Technologies), electroporated into *E.coli* DH5α, and assayed for loss of function using the described assays.

Gene expression and protein purification Genes of interest were amplified using PCR with *M. hydrolyticus* genomic DNA as template. The sequence of the oligonucleotide primers were as follows (restriction sites underlined, position of 5' nucleotide within the gene sequence in parentheses): ChiA-F(34), TAGGCGCGCCATGCATTGCTAGCGCTCGGC; ChiA-R(2025), CCATCGATGTCTGCACCCGCATTGGC; ChiB-F(19),

TAGGCGCGCCATAGCGTGATCGCTCTCGGGC; ChiB-R(1575),
CCATCGATCGGCAGGTTATGCACATGG; ChiC-F(80),
TAGGCGCGCCATCTGCTCGGCTTATTGGCCG, ChiC-R(1620),
GGGGTACCAACGCATAGGCAGTACCCG. After digestion with the appropriate restriction enzymes, amplified DNA was ligated into pETBlue2 (Novagen) using T4 DNA ligase (New England Biolabs, Beverly MA). Constructs were introduced into *E.coli* Tuner™ cells and protein expression performed according to the manufacturer's protocol. Cells were lysed using BugBuster™ Protein Extraction Reagent (Novagen) and recombinant, HexaHis tagged protein purified using Ni-NTA columns (Qiagen, Valencia CA) according to the manufacturer's protocol for native protein purification.

Zymography and Enzyme Reactions *E. coli* cells expressing ChiA, ChiB, or ChiC were lysed in 2X SDS-PAGE loading dye. Samples were separated in a glycol chitin zymogram as described previously. Proteins were refolded overnight at 4°C in refolding buffer [50 mM Tris, 10 mM EDTA, 5 mM 2-mercaptoethanol, pH 7.5], allowed to act on the incorporated substrate in activity buffer [50 mM Tris, 200 mM NaCl, pH 7.2], and stained and visualized as described. Enzyme reactions consisted of purified enzyme added to reaction buffer [50 mM Tris, 25 mM NaCl, 50 µM MUF-diNAG or MUF-triNAG, pH 7.5] in a final volume of 1 ml. Reactions were incubated at 30°C and sampled at various time points. Fluorescence units were measured using a Hoefer TK-100 fluorometer (excitation wavelength, 365 nm; emitted wavelength, 460 nm) and compared to a standard curve prepared with 4'-

methylumbelliferone. The reaction rates represent the initial linear rate and in each case less than 30% of the substrate was degraded (thus an excess of substrate was present).

Complex polysaccharide utilization assays To screen other *Microbulbifer* spp. for the ability to use alginate or pectin as a sole carbon source, strains were grown in minimal medium (2.5% sea salt, 0.05% yeast extract) with the polysaccharide of interest (0.2%). Assays for the utilization of cellulose, pullulan, and xylan employed AZCL-linked carbohydrates (Megazyme) incorporated into marine agar prepared with 1.5% agar and 1/3rd strength Difco 2216 Marine Medium. *M. hydrolyticus* was a positive control for xylan and cellulose utilization and *M. degradans* was utilized as a positive control for pectin and alginate utilization. Negative controls included uninoculated media and media prepared without the addition of a complex polysaccharide.

Protein sequence analysis Identification of protein domains was performed using the Simple Modular Architecture Research Tool. Similarity between proteins and protein domains was determined by the BLAST algorithm. The lipoprotein anchoring site within CelA was identified using the online Database of Bacterial Lipoproteins.

Chemicals and Reagents Restriction enzymes were from New England Biolabs. ProofPro DNA polymerase was from Continental Lab Products (San Diego CA).

Glycol Chitin was obtained from Fisher Scientific (Pittsburgh PA). All other chemicals were obtained from Sigma-Aldrich (St. Louis MO) unless otherwise noted.

Accession numbers Genes and the corresponding proteins that were identified in this work were submitted to GenBank and assigned the following accession numbers: ChiA, AY646086; ChiB, AY646087; ChiC, AY646088; XylA, AY646089; CelA, AY646090.

Results

Described Microbulbifer spp. degrade many complex polysaccharides

Microbulbifer spp. degrade a variety of complex polysaccharides. Because this activity appears to be a defining feature of this genus and its close relatives, experiments were performed to evaluate the ability of *M. hydrolyticus*, *M. salipaludis*, and *M. elongatus* to depolymerize and metabolize complex carbohydrates not addressed in their published characterizations. Each of these organisms was shown to degrade pullulan, and *M. hydrolyticus* and *M. salipaludis* were also shown to metabolize alginate (Table 5-1). Pectin was metabolized by *M. hydrolyticus* and *M. elongatus*. *M. salipaludis* was not able to degrade cellulose or pectin, though it was previously shown to degrade xylan. Finally, *M. elongatus* was unable to degrade xylan under the conditions used. It appears that the enzyme systems of these organisms are not repressed by the presence of a non-cognate complex

Table 5-1

Complex polysaccharides metabolized by described *Microbulbifer* spp. and *T. turnerae*

Substrate	<i>M.</i> <i>hydrolyticus</i>	<i>M.</i> <i>degradans</i>	<i>M.</i> <i>salipaludis</i>	<i>M.</i> <i>elongatus</i>	<i>M.</i> <i>arenaceus</i> [†]	<i>M. sp.</i> <i>JAMB-A7</i> [†]	<i>Teredinibacter</i> <i>turnerae</i> [†]
Alginate	+*	+	+*	+	ND	ND	ND
Agar	-	+	w	+	-	+	-
Cellulose	+	+	-*	+	ND	ND	+
Chitin	+	+	-	+	+	ND	ND
Pectin	+*	+	-*	+*	ND	ND	+
Starch	+	+	+	+	+	+	ND
Xylan	+	+	+	-*	ND	ND	+
Pullulan	+*	+	+*	+*	ND	ND	ND
Reference	(42)	(41)	(155)	(154)	(136)	(101)	(28)

* as determined in this work; +, substrate depolymerized; -, substrate not depolymerized; w, weakly positive; ND, not determined;

[†] strain not available

polysaccharide and that these strains utilize multiple sugars simultaneously. The agarolytic activities of *M. elongatus* and *M. salipaludis* did not interfere with the metabolism of pullulan when incubated on solid medium, though agarase activity was obvious from the pitting phenotype of the colonies.

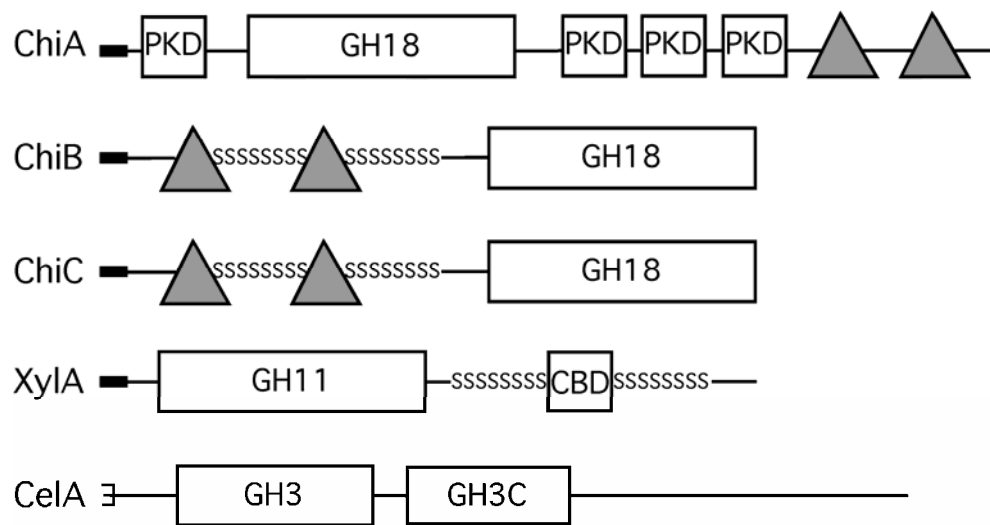
Identification of a M. hydrolyticus xylanase and 1,4-β-D-glucan glucohydrolase

To identify clones encoding enzymes active on cellulose and xylan, *E. coli* carrying the *M. hydrolyticus* genomic library were screened on media containing the dye-linked substrates AZCL-Xylan and AZCL-HE-Cellulose. The fosmid construct of a library clone found to degrade AZCL-Xylan was purified, mutagenized *in vitro* using a transposon mutagenesis kit, and transformed into *E. coli* DH5α. Loss-of-function mutants were isolated and the nucleotide sequence of the genomic insert surrounding the transposon was determined using sequencing primers specific to the ends of the transposon. The mutagenized open reading frame was designated *xylA* and was 59% identical at the nucleotide level to ZP_00067071, a predicted *M. degradans* xylanase. Xylanase A is a 417 aa protein that overall is most similar to a xylanase from *Pseudomonas* sp. ND137 (BAB79287). It contains a Glycoside Hydrolase family 11 (GH11) catalytic domain (aa 71-241) that is most similar to the GH11 domain of ZP_00067071 (85% identity, 93% similarity) (Figure 5-1). A possible carbohydrate binding domain (aa 262-385) was identified, though this region has only weak similarity to known binding motifs. XylA also contains two polyserine linkers. The first separates the GH11 domain from the potential binding domain and

Figure 5-1

Architecture of *Microbulbifer hydrolyticus* enzymes identified in this work

The sequence features of the identified degradative enzymes were determined using the programs listed in Materials and Methods. GH, Glycoside Hydrolase family; PKD, polycystic kidney disease domain; CBD, carbohydrate binding domain; SSSS, polyserine linker; E, lipoprotein acylation site; black box, secretion signal; grey triangle, chitin binding domain.



consists of GTS-[GGSS]₄-SGS. The second linker separates the putative binding region from the final 13 residues of the enzyme and consists of [GGSS]₅. As this linker does not appear to separate two functional domains, its significance not clear. A homologue of this protein found in *M. degradans* (ZP_00067071) contains a polysaccharide deacetylase domain after a similarly located polyserine linker domain, suggesting that XylA is a truncated derivative of this *M. degradans* enzyme.

A similar screen was undertaken to identify cellulase producing library clones using AZCL-HE-Cellulose, a strategy that has been used successfully to identify cellulase producing clones in a similarly prepared *M. degradans* genomic library (M. Howard, unpublished results). Surprisingly, no clones with activity on this substrate were identified, though *M. hydrolyticus* is known to degrade cellulose and weakly degraded AZCL-HE-Cellulose. As an alternative, MUF-cellobiose was utilized in agarose overlays. While this substrate is not likely to identify major endo-cellulases, exo-cellulase and 1,4- β -D-glucan glucohydrolase activity are detected and it was hypothesized that some of the genes for endo-cellulases could be clustered with genes for exo-acting enzymes. Several clones with activity on MUF-cellobiose were identified; mutational and sequence analysis indicated all of these clones encoded the same gene, designated *celA*. At the nucleotide level, *celA* was only 29% identical to a *Caulobacter crescentus* CB15 xylanase (AAK24025) and was not similar to any *M. degradans* gene.

CelA is an 882 aa protein that overall it is most similar to CelD, a 1,4- β -D-glucan glucohydrolase from *Cellvibrio japonicus* (CAA46499). It includes a Glycoside Hydrolase family 3 (GH3) domain (aa 78-397) and a GH3C domain (451-

673) (Figure 5-1). The GH3 domain is most similar (58% identity, 77% similarity) to the GH3 domain of CAA46499. The amino terminus of CelA contains a predicted lipoprotein acylation site with a conserved lipobox consisting of the residues L-S-A-C (aa 26-29). Additionally, two positively charged residues and a hydrophobic stretch were found in the residues preceding the lipobox, consistent with lipoprotein sequence features. CelA was active on MUF-cellobiose but not AZCL-Xylan or AZCL-HE-Cellulose, suggesting that it is a 1,4- β -D-glucan glucohydrolase and not an endo- or exo-cellulase. The enzyme from *C. japonicus* to which CelA is most similar released glucose from the non-reducing end of cello-oligomers, but had no activity on cellulose.

Sequence features of three Microbulbifer hydrolyticus chitinases

The genomic library of *M. hydrolyticus* was also screened for clones expressing chitinases by overlaying colonies with an agarose solution containing MUF-triNAG or MUF-diNAG. Clones were screened for activity against these analogs over 24 h. Positive clones were analyzed as described above. Three putative chitinases were identified and designated *chiA*, *chiB*, and *chiC*. *chiA* and *chiB* appear to be transcribed as an operon. They are in tandem in the genome and are separated by 66 nucleotides and no obvious promoter sequence could be found upstream of *chiB* (within the coding sequence of *chiA*). In addition, a transposon insertion in the sequence upstream of *chiA* prevented the expression of *chiB*. The position of *chiC* in the genome in relation to these genes is unknown, but was not within the flanking 10 KB of genomic sequence contained in the fosmid clone carrying *chiA* and *chiB*. *chiA*

and *chiB* had no significant identity to any known genes, while *chiC* shared a small region of identity (< 10% overall) with *chiC* of *M. degradans*.

Chitinase A is a 957 aa enzyme that contains 4 Polycystic Kidney Disease (PKD) domains, a Glycoside Hydrolase family 18 (GH18) domain (aa 162-544) and two chitin-binding domains (Figure 5-1). The GH18 domain is most similar to an endo-chitinase from *Serratia marcescens* (AAL57854) (70% identity, 81% similarity) and has a consensus sequence of SIGGW-X₃₅-DIDWEY. The overall domain architecture is most similar to an *Aeromonas hydrophila* chitinase (AAF70180) with the only major difference being the presence of an additional PKD domain in ChiA.

Chitinase B is a 525 aa enzyme that contains two chitin-binding domains and a GH18 domain (aa 219-506). The GH18 domain was most similar to a *Chromobacterium violaceum* chitinase (AAQ60603) (44% identity, 54% similarity) and has a consensus sequence of SYGGE-X₃₁-DIDLES. The chitin-binding domains and the catalytic site are separated by polyserine linkers of 40 and 36 residues, respectively (Figure 5-1). Each is composed of serine and some glycine (32 Ser of 40 total residues and 26 Ser of 36 total residues, respectively). The domain architecture was identical to ChiA of *M. degradans* and these proteins were overall 46% identical and 58% similar.

Chitinase C is a 548 aa enzyme with two chitin-binding domains, a GH18 catalytic domain (aa 220-544), and two polyserine linkers in the same arrangement as Chitinase B (Figure 5-1). The GH18 domain has a consensus sequence of SGSSE-X₂₄-DLDWEY and is most similar to a chitinase from *Streptomyces peucetius* (AAF43629) (68% identity, 73% similarity). The first polyserine linker is 32 residues

in length (27 Ser, 4 Gly, 1 Thr) and the second is 38 residues in length (27 Ser, 9 Gly, 1 Ala, 1 Cys). Although ChiC has the same domain arrangement as *M. degradans* ChiA, the catalytic site is most similar to that of *M. degradans* ChiC. Though the domain arrangement and amino acid sequence of these enzymes are similar, the genes are dissimilar at the nucleotide level. This suggests that while the domain architecture is similar, the presence of two similar enzymes is not the result of genetic duplication. The polyserine linkers of ChiB and ChiC are dissimilar at the nucleotide level and all codons for serine are utilized with no detectable pattern or codon bias.

ChiA, ChiB, and ChiC are endo-chitinases

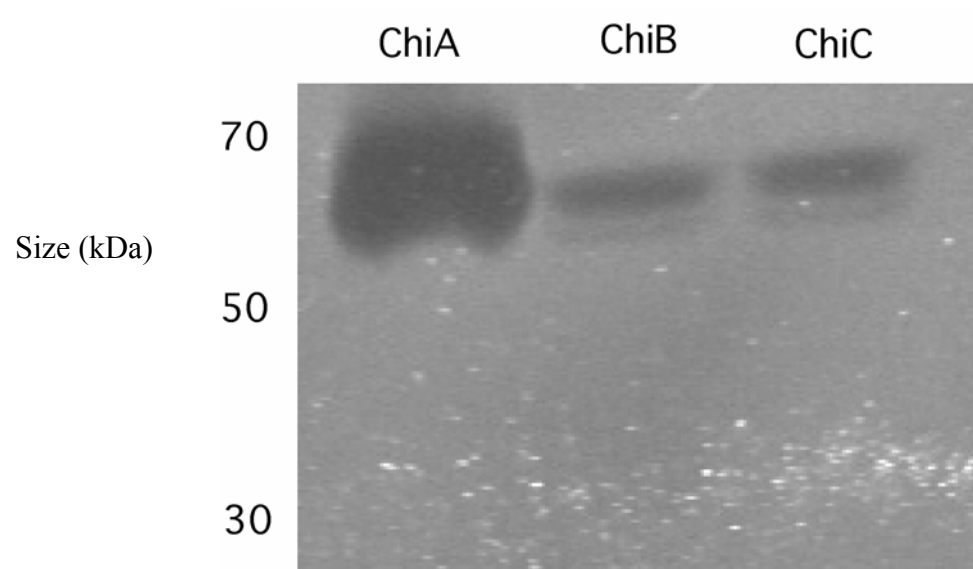
Because chitin analogs prepared with short chitooligosaccharides were used to identify genomic library clones encoding these putative chitinases, activity on polymeric chitin was confirmed using a glycol chitin zymogram. ChiA was truncated by removing the chitin-binding domains to facilitate expression, while ChiB and ChiC lacking secretion signals were expressed as full length proteins. Whole cell lysates of *E. coli* cultures expressing each enzyme were analyzed in glycol chitin zymogram. Dark bands indicative of chitin depolymerase activity were observed for each enzyme preparation (Figure 5-2).

To determine if the identified chitinases were endo- or exo-acting enzymes, the chitin analogs MUF-diNAG and MUF-triNAG were employed. Experiments have demonstrated that exo-chitinases have a higher activity on MUF-diNAG, while endo-chitinases are more active on MUF-triNAG (121). Each enzyme was purified

Figure 5-2

ChiA, ChiB, and ChiC degrade chitin

The activity of recombinant ChiA, ChiB, and ChiC from *M. hydrolyticus* was analyzed in a glycol chitin zymogram. After staining with Calcofluor, dark bands typical of chitin depolymerase activity were observed in each lane. The observed weights are in agreement with the theoretical weights of the expressed enzymes (ChiA, 71 kDa; ChiB, 60 kDa; ChiC, 60 kDa).



and incubated with MUF-diNAG or MUF-triNAG. The amount of liberated 4'-methylumbelliferone was quantified and compared to a standard curve. ChiA and ChiC were 4.6 and 4.4 fold more active on MUF-triNAG than MUF-diNAG, respectively, and ChiB was 8.3 fold more active on MUF-triNAG (Table 5-2). These data suggest that the *M. hydrolytica*s chitinases act as endo-cutting enzymes.

Conclusions

The genus *Microbulbifer* is an emerging group of organisms that express highly modular enzymes that function in the degradation of a wide variety of complex polysaccharides (CP). Like *M. degradans*, the other *Microbulbifer* spp. tested in this work (*M. elongatus*, *M. salipaludis*, and *M. hydrolyticus*) were found to degrade a broad spectrum of CP. Each is able to degrade pullulan and alginate. In addition, *M. hydrolyticus* was able to degrade cellulose, xylan, and pectin. *M. salipaludis* was able to degrade xylan, but not cellulose or pectin, while *M. elongatus* was not able to degrade xylan. Though *M. degradans* appears to have the ability to degrade the most diverse range of polysaccharides, the other *Microbulbifer* spp. examined encode enzyme systems for the degradation of many algal- and plant-derived CP. *Microbulbifer* spp. have been isolated from coastal salt marshes where CP components of higher plant cell walls (cellulose, pectin, xylan) and of algal origin (cellulose, alginate, agar) are abundant (2, 42). Therefore, these enzymes systems are likely to provide a metabolic advantage in a salt marsh environment. Further, these species have been isolated in North America, Europe, and Asia, and the genome of a

putative *Microbulbifer* species was partially sequenced during recent efforts to analyze the bacterial population of the Sargasso Sea (147). This suggests that these species are widely distributed and may play a role in global carbon cycling. Because the ability to degrade a wide variety of CP appears to be a defining feature of the genus *Microbulbifer*, it would be prudent that a battery of tests to examine CP metabolism be included in all future characterizations of novel species assigned to this genus. This will avoid gaps in the known metabolic abilities of described *Microbulbifer* spp. (e.g., Table 5-1).

Another feature shared by *Microbulbifer* spp. and close relatives appears to be the presence of polyserine linkers within secreted degradative enzymes. Polyserine linkers are rare among prokaryotic proteins and are almost exclusively found among secreted CP depolymerizing enzymes (53). Only two genera, *Microbulbifer* and *Cellvibrio*, are known to include two or more species that encode CP depolymerases with polyserine linkers. Among the organisms that are now known to encode multiple enzymes with polyserine linkers (*M. degradans*, *M. hydrolyticus*, *C. japonicus*, and *C. mixtus*), the linkers separate functional domains (surface anchoring, binding, and/or catalytic motifs), in theory to enhance the flexibility and range of motion of an enzyme. Consistent with this theory, the deletion of polyserine linkers from a *C. japonicus* xylanase resulted in a decrease in activity on polymeric xylan but not soluble substrates (13). Additionally, the polyserine linkers of these species are encoded by all six codons for serine without any detectable bias, nor in any obvious pattern. Interestingly, polyserine linkers are also found in some degradative enzymes

Table 5-2

Activity of *M. hydrolyticus* chitinases on MUF-diNAG and MUF-triNAG

<i>Enzyme</i>	<i>MUF-diNAG</i> ($\mu\text{mol/min/mg}$) ^b	<i>MUF-triNAG</i> ($\mu\text{mol/min/mg}$) ^b	<i>Ratio of Rates</i> <i>MUF-triNAG / MUF-</i> <i>diNAG</i>
ChiA ^a	0.090 \pm 0.0072	0.415 \pm 0.0398	4.6
ChiB	0.145 \pm 0.0157	1.200 \pm 0.0096	8.3
ChiC	0.200 \pm 0.0164	0.880 \pm 0.0114	4.4

^a Truncated as described in Materials and Methods

^b Rates expressed as μmoles of 4'-methylumbelliferone liberated from the indicated substrate per minute per mg of enzyme. Rates estimated by comparison to a standard curve. Data are the average of at least three experiments.

of *Teredinibacter tunerae* (Personal communication, D. Distel to R. Weiner), a close phylogenetic relative of *M. degradans*.

Thirty two of the more than 100 postulated degradative enzymes encoded by *M. degradans* have features typical of lipoproteins, and several other degradative enzymes from *C. japonicus* and *C. mixtus* have apparent lipoprotein acylation sites as well (e.g., CAA88764, AAO31762, AAS19695). The attachment of secreted proteins to the surface of a bacterial cell can increase the amount of soluble sugars available for metabolism and prevent diffusion of secreted enzymes, a major factor in marine and littoral environments. CelA from *M. hydrolyticus* is also a predicted lipoprotein, suggesting that it may localize to the cell surface. CelD of *C. japonicus*, the enzyme to which CelA is most similar, also has sequence features typical of a lipoprotein and has been found to associate with the cell membrane of *C. japonicus*. Therefore, it seems that this group of organisms may utilize surface attachment of degradative enzymes as a survival strategy in the environments they inhabit.

Because these species share some common metabolic activities, the chitinolytic system of *M. hydrolyticus* was examined and compared to that of *M. degradans* to determine if they share common genes and proteins as well. The chitinolytic system of *M. hydrolyticus* includes at least three chitin depolymerases. ChiB and ChiC have identical domain arrangements consisting of two chitin-binding domains followed by a GH18 catalytic domain, each of which is separated from the others by a polyserine linker (Fig 5-1). Interestingly, this is the same arrangement of domains and linkers found in ChiA of *M. degradans* (54). This domain arrangement may be particularly efficient at facilitating substrate/enzyme interactions and has thus

been retained by both species. The genes for these chitinases are dissimilar at the nucleotide level, leading to interesting questions concerning convergent evolution in the creation of this enzyme configuration. Further, the genes surrounding the chitinases in each species are different, and none of the *M. degradans* chitinases are clustered together. While these enzyme systems may share proteins with similar features and domain arrangements, it appears that the sequences of the genes for these enzymes are dissimilar.

All of the identified *M. hydrolyticus* chitinases were at least 4 times more active on MUF-triNAG than MUF-diNAG, suggesting that each is an endo-chitinase. Endo-chitinases randomly cleave the chitin polymer at any linkage other than the first at a non-reducing end. The result of endo-chitinolytic activity on a chitin strand is a pool of chitobiose and long chain, soluble chitooligosaccharides that are further processed by periplasmic enzymes or exo-chitinases. Therefore, it is likely that additional enzymes of the *M. hydrolyticus* chitinolytic pathway have yet to be discovered.

M. hydrolyticus has now been shown to encode enzymes with polyserine linkers and also a 1,4- β -D-glucan glucohydrolase with a putative lipoprotein acylation site; these are two enzyme features that are extensively observed among *M. degradans* CP depolymerizing enzymes. Despite several phylogenetic differences between known and described *Microbulbifer* spp., *Cellvibrio* spp., and *Teredinibacter turnerae*, all appear to have several common features among their CP metabolizing systems that can be better addressed using the genomic sequence of *M. degradans*. This sequence affords the invaluable opportunity to evaluate complete enzyme

systems and related proteins from a genomic perspective and apply that data to other similar organisms.

Chapter 6: General Conclusions and Future Directions

The observation that *Microbulbifer degradans* is able to depolymerize and metabolize a wide variety of complex polysaccharides was made nearly two decades ago (2). With recent experimental data in hand and the impending completion of the genomic sequence of this organism, it is now possible to elaborate on *how* this remarkable metabolic feat is achieved by *M. degradans*. As is always the case, answering one question leads to a host of others; this work was undertaken to create a framework and knowledge base for future studies. This was achieved by first examining the chitinolytic system and its enzymes and later branched into the architecture and organization of other carbohydrases.

Studies of the chitinolytic system revealed a complement of enzymes that was similar to that found in other described microbial systems (54). Chitin depolymerases, soluble sugar degrading enzymes, sugar binding proteins, and GlcNAc metabolizing enzymes were all present and had some similarity to other known enzymes. Upon closer inspection, however, the novel aspects of the system became apparent. ChiA and ChiB each contained polyserine domains that joined functional domains by what is presumed to be a flexible linker. ChiB was found to contain two discrete catalytic domains with different activities. These domains were examined and shown to synergistically degrade polymeric chitin. Finally, ChiB was found to contain a putative lipoprotein acylation site at its amino terminus, thus providing evidence for a possible mechanism of surface localization. This is a

particularly interesting observation, as this entire project was initiated, in part, to determine how the observed surface attachment of some carbohydrases was facilitated. The novel characteristics of some of the enzymes of the chitinolytic system were found to be conserved in most of the other complex polysaccharide degrading systems as well. Thus, *M. degradans* seems to have several tools to efficiently degrade complex polysaccharides: a multitude of enzymes for each substrate, a set of enzymes with two catalytic domains, polyserine linkers that facilitate substrate/enzyme interactions, and the potential surface anchoring of degradative enzymes.

The first strategy utilized by *M. degradans* for the efficient degradation of such a wide variety of complex polysaccharides is the production of an overwhelming number of different enzymes targeted to the same substrate. This feature is less obvious in the chitinase system (3 depolymerases) and the agarase system (5 depolymerases) than in the plant cell wall degrading enzyme suites including cellulases/glucanases (>20 depolymerases), pectate lyases (>10 enzymes), and xylanases (>7 enzymes). Because *M. degradans* can break down plant matter in monoculture and plant tissue is often complex and varies from species to species, each of the individual enzymes may have subtle differences in substrate preference. For instance, highly decorated xylans and celluloses may require different enzymes for complete depolymerization to occur. Clearly the evolutionary and energetic pressures needed to maintain such a large complement of seemingly redundant genes speaks to their importance in some metabolic function or process. Within this section, future work could be aimed at determining what specific enzymes are

upregulated in response to different carbon sources; the impending arrival of a microarray should assist in collecting the data needed to answer these questions.

The second strategy for efficient degradation of CP, enzymes with two catalytic domains, is a more limited effort on the part of *M. degradans*. ChiB was analyzed in detail in this work and represents, *inter alia*, how two different enzymatic activities linked on a single polypeptide can be of great value to a bacterium (52). The genomic sequence reveals a cellulase, a pectate lyase, and an alginate lyase that may also have two distinct catalytic domains. The putative domains within this cellulase have been tentatively shown to have activity on different glucans (L. Taylor, unpublished observations). Though these four enzymes represent only a small percentage of the carbohydrases encoded by *M. degradans*, it is exceedingly rare to find an organism that encodes a single degradative enzyme with two catalytic domains, to say nothing of an organism that may encode as many as four. Clearly future work on this subset of enzymes may lead to interesting and valuable answers about synergistic degradation and enzyme organization.

Polyserine linkers represent the third strategy for efficient depolymerization employed by *M. degradans*. These linkers, based on the known properties of serine and the crystal structure of a eukaryotic protein, are hypothesized to be highly flexible (53). This flexibility is most likely not *sine qua non* for enzymatic activity, but instead serves to optimize substrate/enzyme interactions and increase the range of activity of an enzyme bound to a polymer. Polyserine linkers are found only within carbohydrases and only between functional groups; this suggests a specific and conserved role for these domains as linkers of carbohydrate depolymerizing and

binding domains found within secreted enzymes. Though not found in all carbohydrases encoded by *M. degradans*, they are present in more than one third of these enzymes. These domain linkers were also found in some of the degradative enzymes encoded by *M. hydrolyticus* and *M. elongatus*. This leaves little doubt that they play a specific role in CP degradation. Future directions here include NMR analysis of these domains to confirm that they are flexible and perhaps assign some quantitative values to this feature. In addition, shortening and lengthening these domains and evaluating the effect this has on the biochemical activity of an enzyme would also be informative.

The final degradation strategy utilized by *M. degradans* is the potential surface localization of carbohydrases via a lipoprotein anchor or another mechanism. The advantages of surface localized enzymes are plentiful and clear: diffusion and loss of high value enzymes is mitigated, loss of soluble sugars released by these enzymes is minimized, and multiple enzymatic activities can be kept in proximity to each other so that the synergistic potential of the systems can be fully achieved. This theory of surface attachment is supported by previous work on *M. degradans* wherein surface association of carbohydrase activity was observed until stationary phase when the activity was released into the supernatant (R. Weiner, unpublished observations). In the case of PulA, a lipid anchored carbohydrase from *Klebsiella pneumoniae*, this same pattern of attachment and release was observed (115). While this release is most likely the result of artificial conditions created in a laboratory flask, it is clear that lipoprotein modifications can be utilized to anchor carbohydrases to the cell surface and that this anchoring is growth-phase dependent. Further, the observation

that polyserine linkers separate the putative lipoprotein acylation sites from the rest of the protein in nine enzymes supports the notion that these anchoring motifs are in some way active, as polyserine domains are only located between functional domains.

Enzyme architecture refers not only to the type of functional domains found within a protein, but also to their organization, how they are linked, and how the domains interact with other proteins, the cell, and their cognate substrate. Evolution is thought to change enzymes over time to better facilitate their intended activity; in the marine environment enzymatic activity must take place on complex, heterogenous matter where diffusion and dilution are major factors. The complex polysaccharide degrading enzymes of *M. degradans* that have adapted to these stresses have an architecture composed of numerous catalytic and binding domains, often joined by flexible linkers. Because the genes for most of these enzymes appear to be unique within the genome, it does not appear that these proteins arose via a “cut and paste” mechanism wherein random functional domains were fused to create novel enzymes. It is possible to speculate that the configurations of many of these proteins have developed slowly over time and are may be examples of optimized domain arrangement and linkage.

Appendix I: Growth of *Salmonella enterica* on Germinating Alfalfa Sprouts

This work was undertaken as a rotation project and does not relate to the core thesis. The data in this section appeared in *Appl Environ Microbiol* 69(1): 548-53.

Abstract

Alfalfa sprouts and other seed sprouts have been implicated in numerous outbreaks of salmonellosis. The source of these epidemics appears to have been low-level contamination of seeds by *Salmonella* that developed into clinically significant populations during the seed germination process. To test the possibility that *Salmonella enterica* strains carry host range determinants that allow them to grow on alfalfa, strains isolated from alfalfa or other sources were surveyed for their ability to grow on germinating alfalfa seeds. An *S. enterica* serovar Cubana strain originally isolated from contaminated alfalfa sprouts multiplied most rapidly during the initial 24 h of the seed germination process. Germinating alfalfa seeds supported the multiplication of *S. enterica* cells prior to the emergence of the root radicle at 72 h. Thereafter, lower rates of multiplication were observed, concurrent with a decrease in reducing sugars released from plant tissue. The ability of *S. enterica* to grow on germinating alfalfa seeds was independent of the serovar, isolation source, or virulence of the strain. Isolates obtained from alfalfa grew to similar levels as those isolated from contaminated meat products or stools. Each of the strains could be detected in the waste irrigation water, with populations being strongly correlated with

those detected on the germinating alfalfa seeds. The *S. enterica* strains were able to use waste irrigation water as a sole carbon and nitrogen source, and mutants lacking known virulence traits did not demonstrate impaired growth. Therefore, *S. enterica* strains appear to grow saprophytically on soluble organics released from seeds during early phases of germination. The ability to detect *S. enterica* in the waste irrigation water early in the germination process indicates that this method may be used as a simple way to monitor the contamination of sprouts during commercial operations.

Introduction

In recent years, numerous outbreaks of salmonellosis have been associated with the consumption of contaminated alfalfa sprouts (131). A California study implicated the consumption of contaminated alfalfa sprouts in nearly 50% of the documented cases of salmonellosis prior to the institution of U.S. Food and Drug Administration guidelines for the production of alfalfa sprouts in 1998 (90). These outbreaks have involved a variety of *Salmonella enterica* serotypes, including strains of Bovismorbificans, Stanley, Newport, Montevideo, Meleagridis, Infantalis, Anatum, Senftenberg, Havana, Cubana, Tennessee, Saint-Paul (United Kingdom), Gold-Coast (United Kingdom), Mbandaka, and Enteritidis (7, 82, 116). Related outbreaks involving *Escherichia coli* O157:H7 (23, 90) and *Bacillus cereus* (111) have also been reported.

In each of the aforementioned sprout-associated outbreaks, low level seed contamination appears to have been the source of the *Salmonella* (82). Contaminating populations were usually undetectable in seed lots prior to germination. In cases where clinically significant populations developed, *S. enterica* was found growing epiphytically on sprout surfaces without producing obvious signs of contamination or plant disease (11, 113). Alfalfa sprouts and other seed sprouts are able to support significant naturally occurring microbial populations. The microbial bioburden of sprouts, which are primarily root tissue, form biofilms that can reach 10^8 CFU/gram fresh weight (gfw) (33). These biofilms can be observed on sprout hypocotyls as

early as 2 days post-germination, and by day 4 they can be found on all parts of the plant. Several *S. enterica* strains have been shown to reach clinically significant levels (10^6 to 10^7 CFU/gfw) 2 days after germination (137). Little is known about how any of these organisms grow in or on developing alfalfa sprouts.

Bacterial colonization of roots is a complex process (80). Type III protein secretion systems (TTSS) have recently been implicated in the growth and survival of *Pseudomonas fluorescens*, a rhizosphere-associated organism (112). It is thought that *P. fluorescens* utilizes a TTSS to achieve an ecological advantage over other rhizosphere-associated bacteria by translocating effector molecules into the cytoplasm of host cells. There, the effectors promote nutrient efflux into the rhizosphere. TTSS are required for the growth of a number of plant-associated bacteria, such as *Pseudomonas syringae*. Effector proteins translocated by the TTSS appear to interfere with host cell signaling pathways in order to promote nutrient efflux and suppress the host's cellular-defense responses. Specific strains of *P. syringae* have limited host ranges that appear to be linked to the effectors that they produce (22, 56). The asymptomatic epiphytic growth of a *P. syringae* strain under field conditions has also been linked to the activities of a TTSS (49).

S. enterica strains have been shown to express two distinct TTSS that are encoded by pathogenicity islands known as SPI1 and SPI2 (44). The SPI1-associated TTSS functions in the invasion of host cells, whereas the SPI2-linked TTSS is expressed in the post-invasion environment (79). Interestingly, at least one of the secreted effector proteins translocated by the SPI1 TTSS shares homology with a

secreted effector protein found in many plant pathogens that is required for plant-associated growth (130). There is also evidence that some virulence factors of mammalian pathogens are able to affect both animal and plant hosts. For example, a clinical isolate of *Pseudomonas aeruginosa*, which can cause opportunistic infections in humans, has been shown to produce symptoms in *Arabidopsis thaliana*. At least 17 *P. aeruginosa* genes have been associated with symptom production in both plants and mice (106).

In order to determine the nature of *S. enterica* growth on alfalfa sprouts, survival and growth of strains isolated from contaminated alfalfa was compared to that of strains from nonplant sources and distinct serotypes. To evaluate the role of known pathogenicity determinants in the colonization of alfalfa by *S. enterica*, strains with lacking genes for such determinants were obtained and studied.

Materials and Methods

Strains

Strains used in the experiments are shown in Table x. Cultures were routinely grown in Trypticase soy agar (TSA) medium, Luria broth (LB), or M63 minimal salts medium supplemented with 0.2% glucose, with shaking at 37°C.

Surface sterilization of alfalfa seeds

Alfalfa seeds were surface sterilized by using a modification of the procedures recommended by the U.S. Food and Drug Administration (4-6). Seeds were soaked in

a 10% sodium hypochlorite solution for 15 min with shaking, rinsed in distilled H₂O, and allowed to air dry at room temperature. Seeds were stored at 4°C until use.

Treated seed batches were free of detectable *S. enterica*.

Infestation of alfalfa seeds with S. enterica.

Seeds were inoculated with *S. enterica* by vacuum infiltration by adding sterilized seeds and 1 drop of Tween 20 to an overnight broth culture of an *S. enterica* strain. This mixture was placed under a vacuum for 5 min. The seeds were then collected, rinsed with distilled water three times to remove surface-associated *S. enterica*, and allowed to air dry.

Seed germination

Surface-sterilized alfalfa seeds were soaked in tap water for 30 min with shaking, rinsed with tap water, and spread evenly in a 5- by 32-cm sprouting tray. Fifteen grams of seeds was placed in each of the seed trays, which were incubated in an Easy Green sprouting unit (Seed and Grain Technologies, Albuquerque, N.Mex.) maintained at 25°C. This sprouting apparatus provided intermittent spray irrigation of developing sprouts and single-pass, flowthrough irrigation similar to that used by large-scale commercial operations. Sprouts were misted for 30 min every 3 h at an average irrigation rate of 500 µl/cm² of tray/day. To reproduce the low level contamination of sprout batches from infected seed, six inoculated seeds were randomly adding to seed trays containing surface-sterilized seeds of the same age and stage of germination. This method led to an appropriate ratio of 1 infested seed per 1,000 sterile seeds.

Quantification of S. enterica on germinating alfalfa seeds

Approximately 0.5-g samples of germinating seeds/sprouts were removed in triplicate from random positions within the seed batch and placed into 2 ml of sterile saline.

Adherent *S. enterica* cells were released by vortex treatment for 2 min. Washes were serially diluted and plated onto XLT4 agar (86). After incubation at 37°C, black colonies exhibiting typical of *S. enterica* on this medium were enumerated.

Quantification of reducing sugars in waste irrigation water

Reducing sugars were quantified by the method of Garcia et al (38). Fresh bicinchoninic acid (BCA) reagent was prepared by mixing equal parts of BCA solution A (per liter, 54.28 g of Na_2CO_3 , 24.2 g of NaHCO_3 , 1.942 g of disodium 2,2'-bicinchoninate) and BCA solution B (per liter, 1.248 g of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1.262 g of L-serine). Five hundred-microliter samples of waste irrigation water were mixed with an equal volume of BCA reagent and incubated at 80°C for 30 min. After the samples cooled to room temperature, the optical densities at 560 nm were recorded and compared to a standard curve prepared with a diluted glucose solution.

Sterilization of waste irrigation water

Waste irrigation water was collected in batches for 24 h from surface-sterilized seeds. The waste irrigation water was centrifuged for 10 min at 8,000 x g and then filter sterilized with 0.2 μm syringe filters. The doubling times of broth cultures were calculated during logarithmic growth by using 4- and 6-h time points and the standard equation.

Results

Growth of S. enterica serovar Cubana on germinating alfalfa sprouts

In order to define the basic growth characteristics of *S. enterica* on germinating alfalfa seeds under the experimental conditions, populations of a *S. enterica* strain previously isolated from contaminated alfalfa sprouts, *S. enterica* serovar Cubana 98E01362SH2, were monitored on germinating alfalfa seeds during the 5-day period typically used for commercial production of alfalfa sprouts. Initial populations of *S. enterica* in the inoculated seeds used to infest seed batches were below the sensitivity of the plate assays used for determining populations on homogenized seeds. Populations of 98E01362SH2 could be detected on germinating seeds approximately 6 h after the addition of the inoculated seeds and reached a maximum by 48 h (Figure A1-1). The average maximum number of CFU detected on sprouts after 48 h was $(5.0 \pm 2.5) \times 10^5$ /gfw. The doubling time for 98E01362SH2 during the initial phases of growth was determined to be approximately 47 min. *S. enterica* was not detected in germinating, uninoculated, surface-sterilized alfalfa seeds during the 5-day monitoring period. None of the samples obtained from infected batches had statistically larger populations of *S. enterica* cells than those of other samples collected at the same time from the same tray. This indicates that the initially inoculated seeds did not create limited 'zones of contamination'.

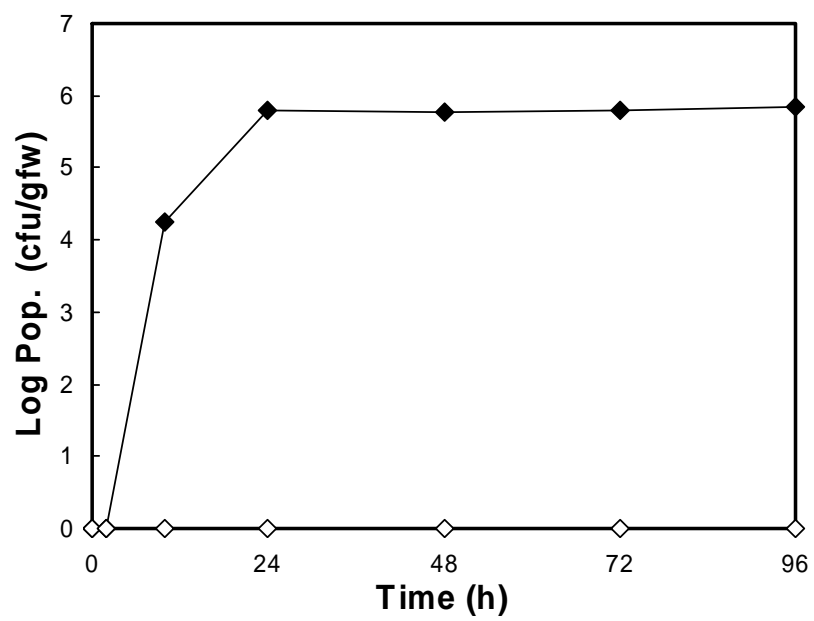
Detection of S. enterica in waste irrigation water

S. enterica could be detected in the waste irrigation water when *S. enterica* populations on developing sprouts were larger than 10^2 CFU/gfw. A strong linear

Figure A1-1

Growth of *S. enterica* serovar Cubana 98E01362SH2 on germinating alfalfa sprouts

The growth of a serovar isolated from contaminated alfalfa sprouts, *S. enterica* serovar Cubana 98E01362SH2, was determined on alfalfa sprouts. Samples of seed/sprout tissue were collected, vortexed in sterile saline, serial diluted, and plated onto XLT4 agar. Colonies displaying the black phenotype of *Salmonella* on this medium were enumerated. Closed diamonds, *S. enterica* serovar Cubana 98E01362SH2; open diamonds, uninfected negative control batch of alfalfa seed.



correlation ($R^2 = 0.9136$) was noted between *S. enterica* populations growing on sprouts and the populations detected in the waste irrigation water (Figure A1-2). Average-size populations detected in the irrigation water were $33\% \pm 14\%$ of the population sizes detected on sprouts, but by day 5, the population levels detected in the waste irrigation water decreased to $17\% \pm 2\%$ of the population levels detected on germinating seeds

To determine how early in the germination process *S. enterica* appears in waste irrigation water, populations of 98E01362SH2 in irrigation water collected from contaminated seed batches were monitored during the initial 48 h after inoculation. 98E01362SH2 could be detected in the waste irrigation water as early as 6 h after the initiation of germination (Figure A1-3), coincident with the appearance of detectable *Salmonella* on alfalfa sprouts.

Evaluation of the ability of germinating alfalfa seeds to support S. enterica growth

Because the largest population increase occurred within the first 24 h of germination ($>10^4$ -fold increase), the ability of *S. enterica* to colonize growing sprouts later in development was evaluated. A seed batch was inoculated as described above, and six germinating seeds/sprouts from the inoculated seed batch were transferred at 24, 48, 72, and 96 h to batches of previously uncontaminated germinating seeds. Populations that developed on germinating seeds and in the waste irrigation water were monitored as before. Seed batches inoculated at 0, 24, and 48 h after the initiation of germination

Figure A1-2

Correlation between *Salmonella* populations on alfalfa sprouts and in waste irrigation water.

The growth of different serovars on alfalfa sprouts and in irrigation water were monitored over time. A correlation between the populations was determined ($R^2 = 0.9136$) indicating that sampling of irrigation water may give a reasonable estimate of populations growing on the plant matter.

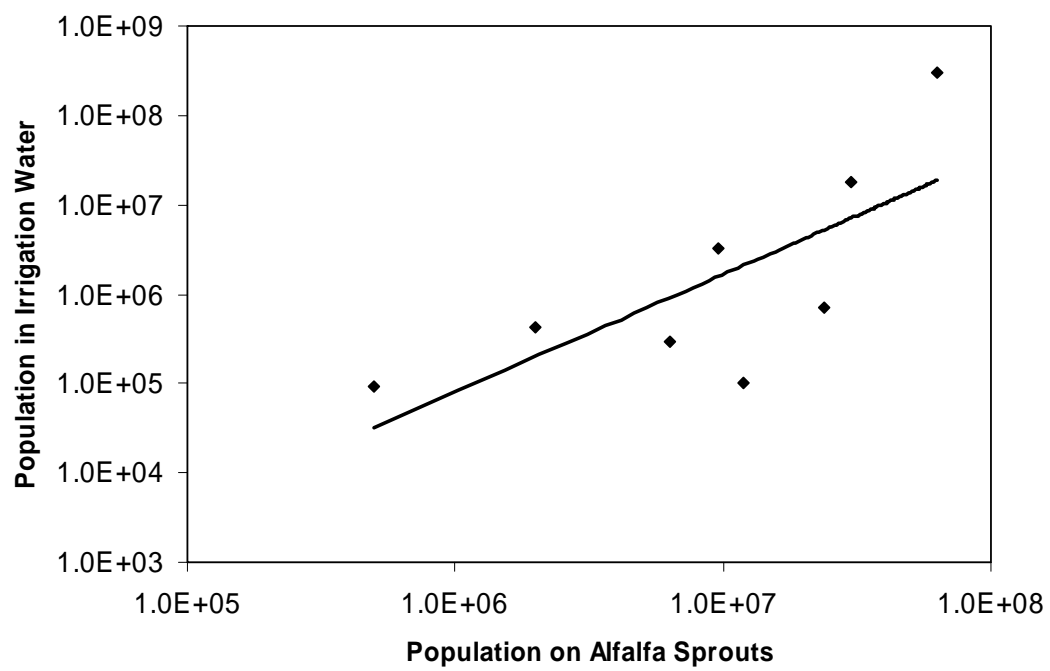
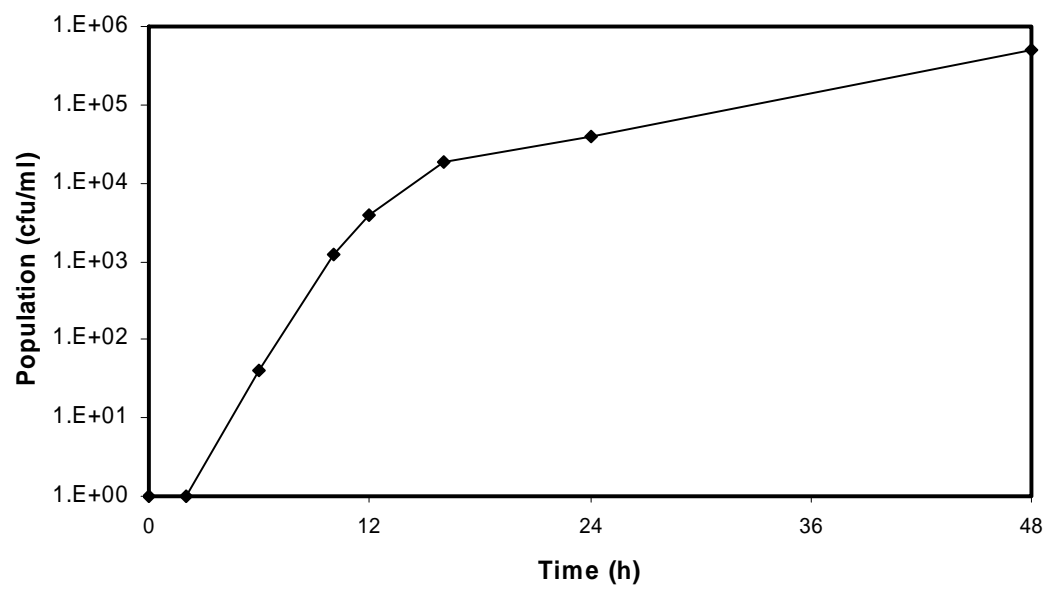


Figure A1-3

Populations of *S. enterica* serovar Cubana 98E01362SH2 in directly sampled waste irrigation water

Samples of irrigation water were collected as soon as the water had moved over the bed of seeds and before entering the collection vessel. Therefore, these populations are based on what can be detected after irrigation water has left the sprouting tray; the bacteria have not yet had the opportunity to grow on sugars or proteins in the collected waste irrigation water.

Salmonella Populations in Directly Sampled Irrigation Water



all established roughly equivalent populations of 98E01362SH2 within 48 h after inoculation of the seed batch (Figure A1-4). A 10-fold reduction in the size of maximum population was detected in seeds inoculated at 72 h, and little growth was detected in seed lots inoculated at 96 h.

Growth of S. enterica in waste irrigation water

Rapid multiplication of 98E01362SH2 during the initial phase of seed germination coincided with a predicted period of nutrient efflux from germinating seeds. To estimate the rate of nutrient release from germinating seeds, levels of reducing sugars in collected waste irrigation water were monitored. The initial concentration of reducing sugars in the effluent was 0.9 mM and declined to less than 0.1 mM by 96 h after the initiation of germination (Figure A1-5).

To evaluate whether *S. enterica* could grow on the nutrients released from germinating seeds, we monitored the growth of 98E01362SH2 populations in filter-sterilized waste irrigation water collected from germinating surface-sterilized seeds during the initial 24 h of the germination process. After a lag of 2 h, the growth of 98E01362SH2 was observed, with a doubling time of 83 min (Figure A1-6). This rate was similar to the growth rate of 98E01362SH2 in minimal salts medium observed after a lag time of 5 h. For comparison, the doubling time of 98E01362SH2 in a rich medium (LB) was approximately 54 min at this temperature.

Figure A1-4

Growth of *S. enterica* serovar Cubana 98E01362SH2 on alfalfa seeds and sprouts during different points in the germination process.

A seed batch was inoculated as described, and six germinating seeds/sprouts from the inoculated seed batch were transferred at 24, 48, 72, and 96 h to batches of previously uncontaminated germinating seeds. Populations that developed on germinating seeds and in the waste irrigation water were monitored as before.

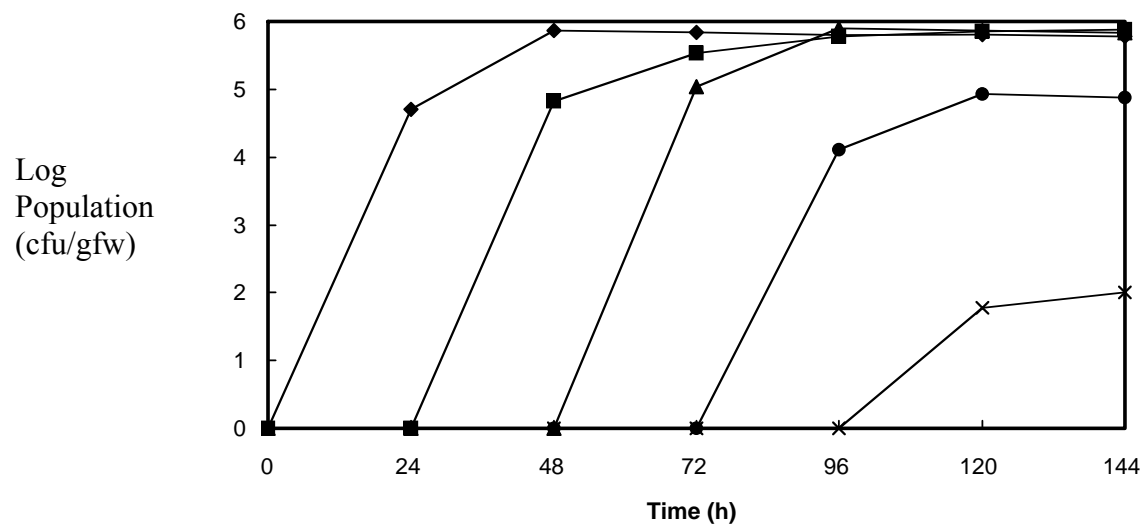


Figure A1-5

Level of reducing sugars in spent irrigation water over time.

Spent irrigation water was collected throughout the germination process. The level of reducing sugars in each sample was determined by the method of Garcia and plotted as a function of time.

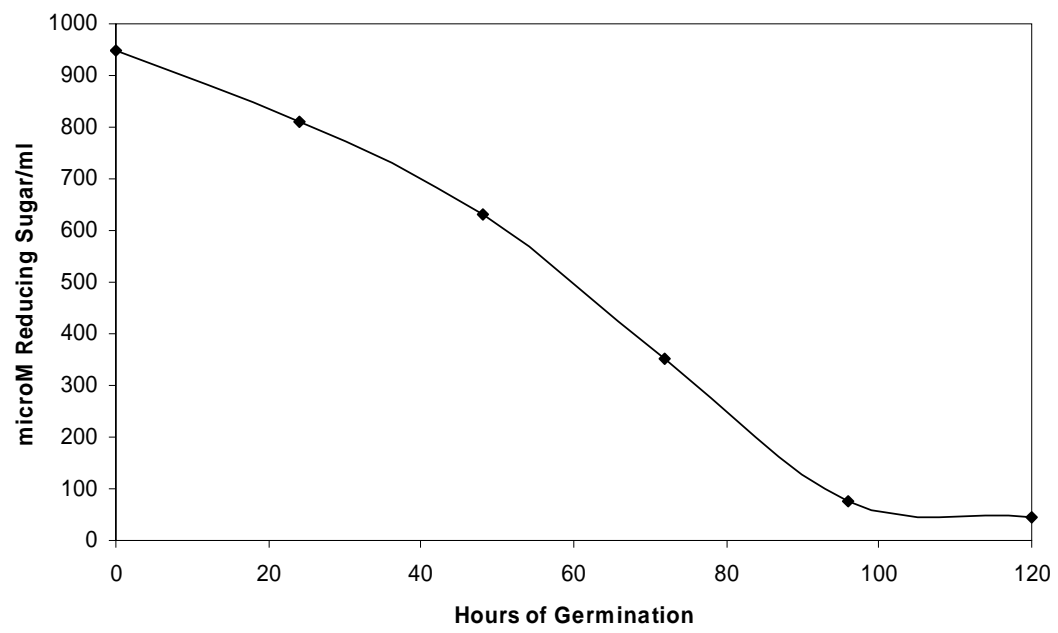
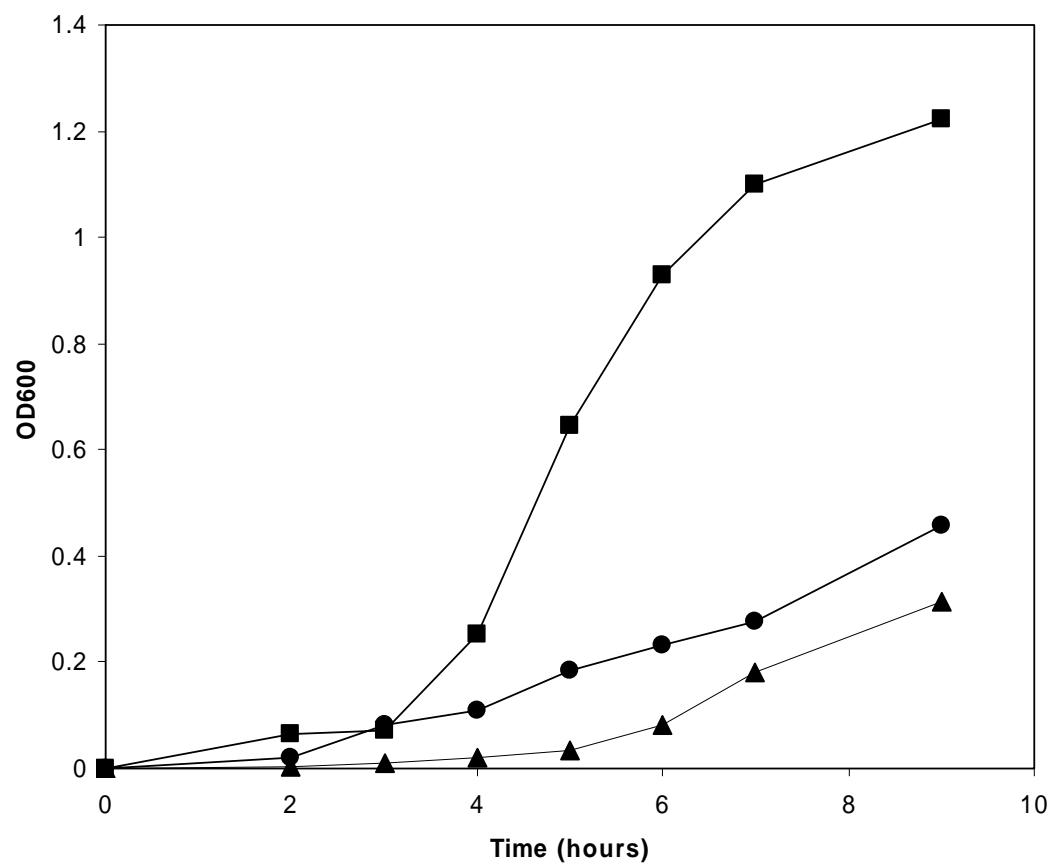


Figure A1-6

Growth of *S. enterica* serovar Cubana 98E01362SH2 in LB, M63 minimal medium, and sterilized spent irrigation water.

Spent irrigation water was collected in batch and filter sterilized. The ability of *S. enterica* serovar Cubana 98E01362SH2 to use the spent irrigation water as a sole carbon and nitrogen source was compared to growth of the same strain in LB or M63. Squares, LB; circles, irrigation water; triangles, M63 Glucose.



Evaluation of the ability of other S. enterica strains to grow on alfalfa sprouts

To determine if the ability of *S. enterica* strains to grow on germinating alfalfa seeds depended on the strain, the growth of isolates obtained from contaminated alfalfa sprouts (serovars Cubana and Tennessee) was compared to the growth of strains unlikely to have been in recent close association with plant tissue. Nonplant associated strains were isolated from contaminated meat by-products or from stools of patients infected after consuming contaminated foodstuffs other than alfalfa sprouts. All nine strains tested were able to grow on germinating alfalfa sprouts (Figure A1-7). After 24 h of growth, the average population size for alfalfa-isolated strains was 1.2×10^5 CFU/gfw whereas the average population size for the other strains was 7.2×10^5 CFU/gfw. Based on a two-sided Student *t* test, however, there was no significant difference in the means at the 95% confidence level. Strains of both groups exhibited up to an additional 100-fold increase in population sizes after 48 h of growth on germinating alfalfa seeds. As before, the populations detected in the irrigation water mirrored the populations detected on the developing sprouts and the *Salmonella* populations were not less than 10^4 /ml for any of the strains tested (Figure A1-8). These results indicate that the *S. enterica* strains tested share the ability to grow on alfalfa sprouts, regardless of source, serotype, or previous history.

Role of S. enterica pathogenicity determinants in colonization of alfalfa sprouts

To determine if there was any requirement for pathogenicity determinants during growth on alfalfa, *S. enterica* serovar Typhimurium SPI1 deletion mutants RM11 and SD69, *prgH*::Tn mutant EE656 (which lacks a structural element of the SPI1-

associated TTSS) and mutant 336 carrying a transposon mutation in a key transcriptional regulator of virulence gene expression (*hilA*) were tested for their ability to grow on germinating alfalfa. All strains were able to grow on the germinating seeds during the initial 48 h (Figure A1-9). The population levels of the mutants observed on the sprouts were similar to those of the wild-type strain at all sampling times. Again, no statistically significant differences in the means were detected at the 95% confidence level. These results are consistent with the conclusion that early growth of *S. enterica* on sprouts appears to be saprophytic and utilizes nutrients released from seeds early in the germination process.

Discussion

The contamination of alfalfa sprouts by *S. enterica* strains and other pathogenic bacteria has become a significant public health concern and has resulted in multiple warnings by regulatory agencies (e.g.,(4-6)). Comparatively little was known about how *S. enterica* strains grow on germinating alfalfa sprouts and what aspects of that growth can be exploited to reduce the human health risk. We have shown here that the ability of *S. enterica* strains to grow on germinating alfalfa sprouts is unrelated to their pathogenicity. All of the *S. enterica* strains tested were able to grow to clinically significant populations, irrespective of the isolation source, serovar, or virulence. The highest rates of multiplication were correlated with peak nutrient efflux from germinating seeds, suggesting that these organisms grow saprophytically on alfalfa. The presence of *S. enterica* in the waste irrigation water was shown to be a key indicator of contamination of alfalfa sprouts.

Figure A1-7

Growth of various serovars of *S. enterica* on alfalfa sprouts

Serovars from different isolation sources were tested for survival and growth on germinating alfalfa sprouts. Open bars, levels at 24h; black bars, levels at 48h.

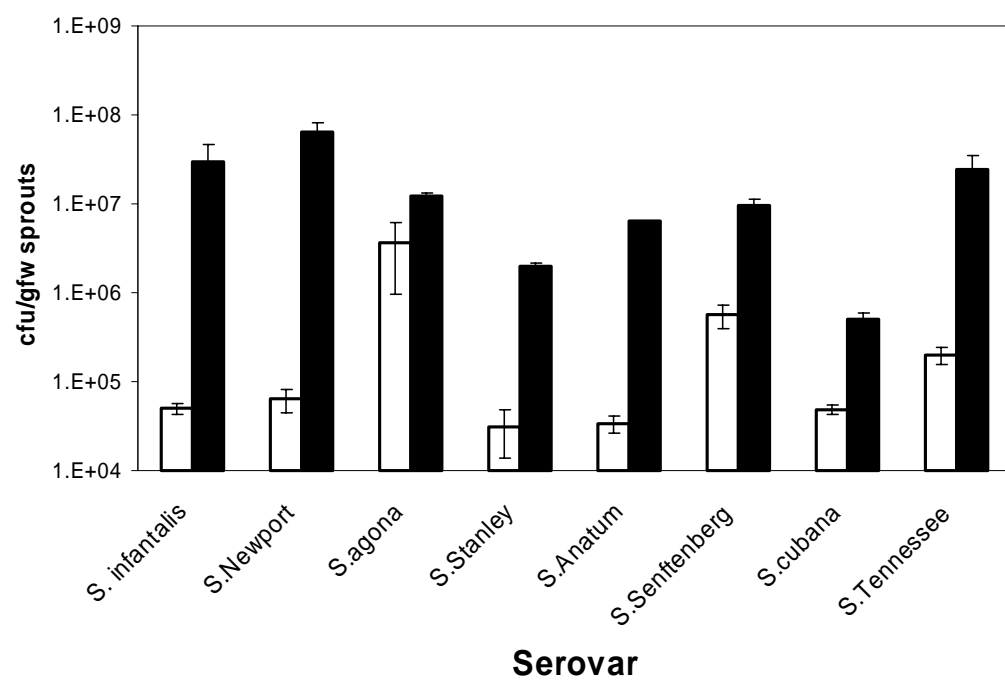


Figure A1-8

Levels of detected populations of various serovars of *S. enterica* in spent irrigation water

Serovars from different isolation sources were tested for survival and population sizes detected in waste irrigation water collected in batch. Open bars, levels at 24h; black bars, levels at 48h.

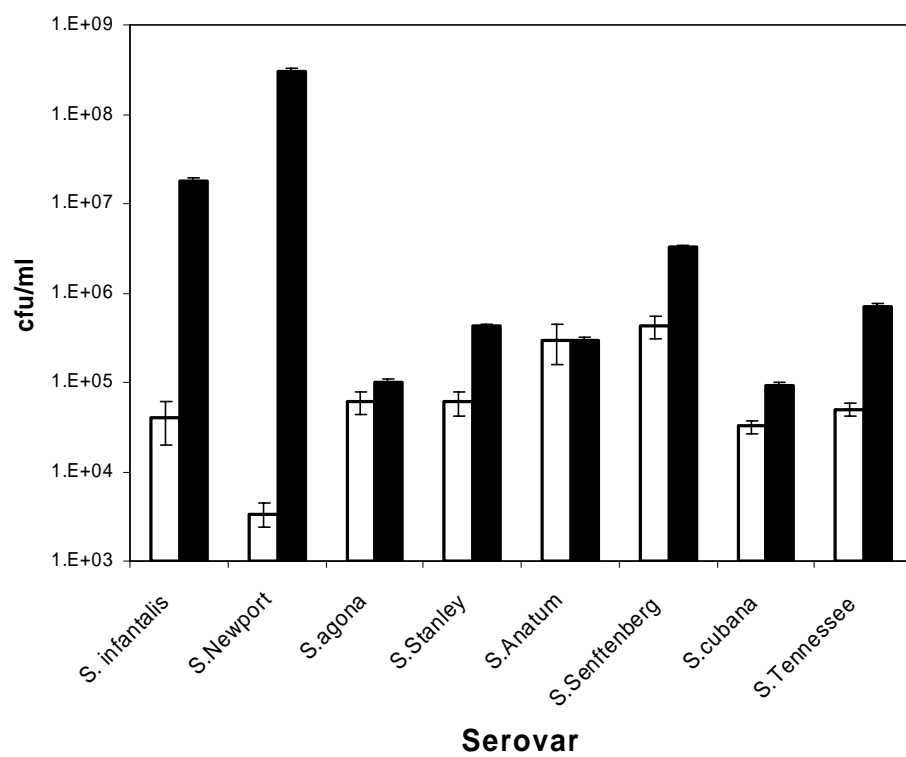
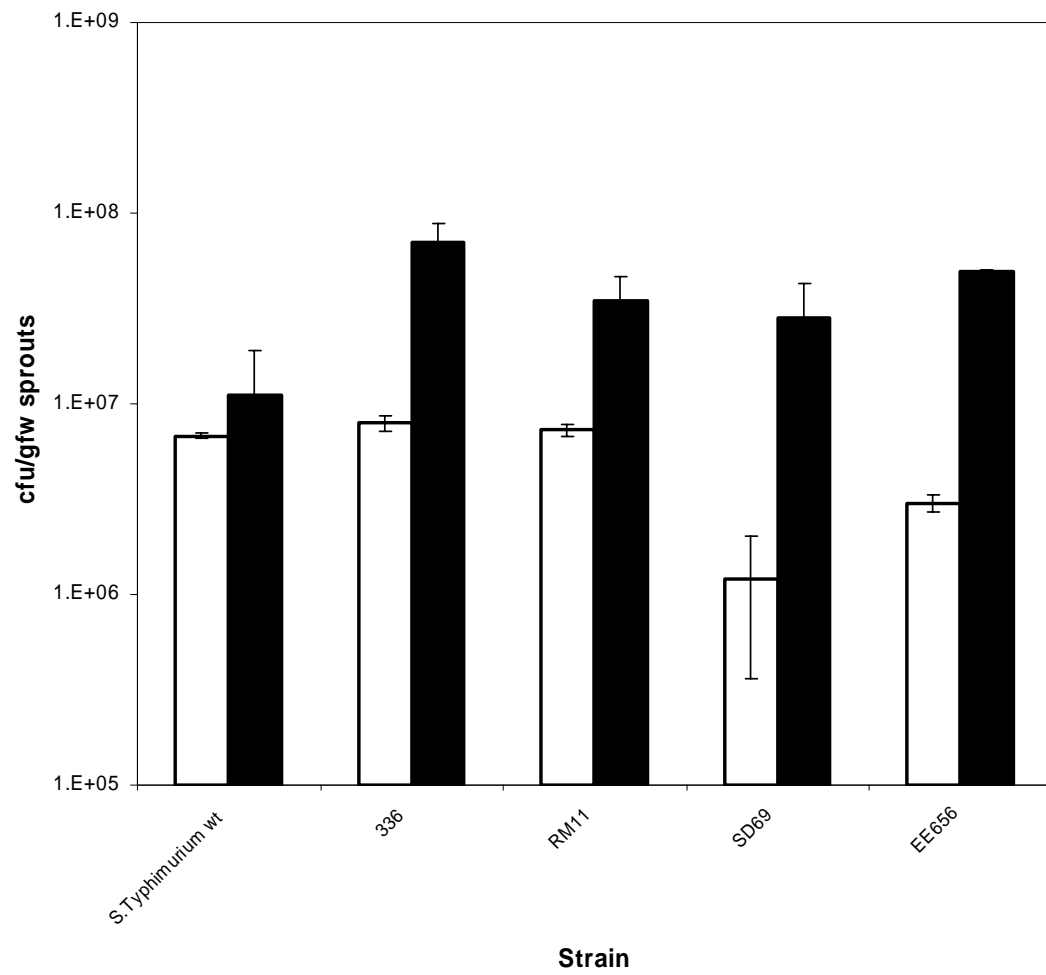


Figure A1-9

Growth of *S.e. Typhimurium* wt and various mutants with impaired pathogenicity on germinating alfalfa sprouts

S. enterica serovar *Typhimurium* SPI1 deletion mutants RM11 and SD69, *prgH*::Tn mutant EE656 (which lacks a structural element of the SPI1-associated TTSS) and mutant 336 carrying a transposon mutation in a key transcriptional regulator of virulence gene expression (*hilA*) were tested for their ability to grow on germinating alfalfa. Open bars, levels at 24h; black bars, levels at 48h.



Prior association with alfalfa did not appear to be a prerequisite for the growth of *S. enterica* strains on germinating alfalfa. All of the *S. enterica* strains tested were able to grow on alfalfa; these strains included representatives of nine serovars that had been collected from diverse sources. The two strains that had been isolated from contaminated alfalfa sprouts grew on germinating alfalfa seeds to levels similar to those of clinical isolates isolated from contaminated fish products and to those of other strains obtained from stools of patients not known to have consumed alfalfa sprouts. In view of the large number of distinct serovars that have been linked to alfalfa sprout-associated salmonellosis outbreaks, it appears that most, if not all, *S. enterica* strains are able to grow on germinating alfalfa seeds.

With other plant-associated bacteria, effectors secreted by the TTSS control the host range and individual strains are capable of colonizing only specific genotypes of the host. In these cases, population levels of the favored strain can be 1,000-fold higher than those of a nonfavored strain. No such specificity was observed among the tested *S. enterica* strains. In addition, SPI1 deletion mutants that lack the genes to form the SPI1-encoded TTSS, a *hila*::Tn mutant (with a mutation in a transcriptional regulator of the TTSS), and mutants lacking a critical gene (*prgH*) required for the assembly of the TTSS grew on germinating alfalfa. These results indicate that the SPI1-encoded TTSS of *S. enterica* is not required for growth on alfalfa. Therefore, the growth of *S. enterica* strains appears to be independent of known pathogenicity determinants.

The *S. enterica* strains tested grow saprophytically on nutrients released by the germinating seeds. The observation that *S. enterica* strains utilize waste irrigation

water as a growth medium indicates that these strains can metabolize the organic compounds released from germinating seeds. Seeds are known to release sugars and other organic molecules into the medium during the breakdown of the endosperm (18). The highest rates of *S. enterica* growth correlated with the release of reducing sugars into waste irrigation water. Although the measured levels of reducing sugars in the waste irrigation water were relatively low, the irrigation process caused substantial dilution of the released nutrients. Concentrations at the seed surface were thus likely to be much higher. Consistent with this conclusion, the doubling time for an *S. enterica* strain growing on germinating alfalfa seeds was similar to the rate observed in the rich LB medium at this temperature.

Populations in waste irrigation water were strongly correlated with populations present on developing alfalfa sprouts. *S. enterica* could be detected in waste irrigation water from contaminated seed batches within 12 h after the initiation of the germination process. Each *S. enterica* strain tested could be detected in the waste irrigation water collected from germinating seeds inoculated with that strain. *S. enterica* populations in waste irrigation water from inoculated seed batches reached a maximum by 48 h, irrespective of the strain, but germinating alfalfa seeds could support the growth of *S. enterica* strains when they were inoculated at any time during the initial 72 h of the process. Thereafter, the rate of growth appeared to be significantly lower. Because waste irrigation water still contained comparatively high numbers of *S. enterica*, the reduced ability of *S. enterica* to grow may have been related to the diminished release of sugars from the germinating seed, active defenses by the emerging root radicle, or a reduction in the ability of *S. enterica* strains to

adhere to and colonize plant tissue. Interestingly, the relative numbers populations of *S. enterica* strains in waste irrigation water were reduced late in the seed germination process. The reduced levels may have been due to the formation of biofilms on the emerging root radicles, as reported previously by other groups (21, 33).

Our data lend support to the U.S. Food and Drug Administration's recommended testing methods to ensure that commercial sprout batches are free of *S. enterica* contamination. The testing of waste irrigation water 48 h after the initiation of seed germination was recommended, but supporting evidence was not provided. In our experiments, we attempted to mimic conditions that would be present in a commercial sprouting operation. Separately infested seeds were used to inoculate germinating seed batches to recreate the field conditions where only a few contaminated seeds are detected within a contaminated seed lot (106). The irrigation conditions for germinating seeds were also similar to those used by commercial operations. The finding that *S. enterica* was present in the waste irrigation water under these conditions indicates that our observations are applicable to commercial operations utilizing similar growth conditions. The detection of *S. enterica* strains in the waste irrigation water at significant levels after 12 h suggests that the testing of wastewater can be performed very early in the sprouting process. Because these levels of *S. enterica* are detectable by several commercial diagnostic systems relatively early in the seed germination process, monitoring *S. enterica* in the waste irrigation water provides a simple and comparatively inexpensive method for detecting the contamination of alfalfa sprouts during commercial production operations. The early

identification of *S. enterica* contamination of alfalfa sprouts may reduce the chance that contaminated products from reaching the consumer.

Appendix II: Methods for the Detection and Characterization of Chitinases and Other Chitin Modifying Enzymes

This work was undertaken to summarize current and relevant methods from the literature. The data and material in this section appeared as a review article in *J Indust Microbiol Biotechnol* 30(11): 627-35.

Introduction

Chitin is composed of repeating N-acetyl-D-glucosamine (GlcNAc) residues and is a component of crustacean exoskeletons, diatoms, fungal cell walls, and squid pens. Several gigatons of chitin are thought to be produced annually in the biome, which makes it an abundant renewable resource (92). Chitin is difficult to purify and modify chemically, so identification of microbial chitin-modifying enzymes and elucidation of their activities could facilitate the efficient production of specific chitin products. Chitin is a versatile and promising biopolymer with numerous industrial, medical, and commercial uses.

Bacteria and fungi have developed systems for the depolymerization, transport, and metabolism of chitin and chitooligosaccharides (31, 66-68). Although chitin is ubiquitous in the marine environment, almost none can be found in marine sediments, demonstrating the efficiency of microbial chitin-degrading systems (157). In general, microbial degradation of chitin involves the activity of secreted chitin

depolymerases (EC 3.2.1.14; poly[1,4-(N-acetyl- β -D-glucosaminide)] glycanohydrolase) that release GlcNAc, chitobiose, and chitooligosaccharides from the polymer. These compounds then enter the periplasm where chitodextrinases (also EC 3.2.1.14) and N-acetylglucosaminidases (EC 3.2.1.52; β -N-acetyl-D-hexosaminide N-acetylhexosaminohydroase) act to form a pool of GlcNAc and, to a lesser extent, chitobiose (9). When transported into the cytoplasm, GlcNAc and chitobiose are metabolized or modified for use in cell wall biogenesis. The activity of each of these enzymes has the potential to be exploited to produce chitin-derived compounds of commercial interest. Some microorganisms, mostly fungi, are able to deacylate chitin by the activity of chitin deacetylases (EC 3.5.1.41; chitin amidohydrolase) to form chitosan. Chitosan can then be degraded by chitosanase (EC 3.2.1.132; chitosan N-acetylglucosaminohydrolase) and metabolized.

The diverse uses of chitin-derived products demand an equally diverse complement of enzymes that can be used to tailor them to specific needs. Chitinases and chitooligosaccharide-modifying enzymes have been isolated from a wide variety of organisms and characterized. Consequently, the known complement of these enzymes includes proteins with a range of pH and temperature optima, substrate specificities, and reaction end products. Because a large number of unculturable and/or unidentified bacteria, especially from the marine environment, are hypothesized to produce chitinases (25, 26), it is feasible that a large pool of uncharacterized chitin-degrading and -modifying enzymes has yet to be discovered. This review summarizes many of the current methods used to assay environmental isolates for chitinolytic activity, screen genomic libraries for genes encoding chitin-

modifying proteins, and characterize the enzymatic properties of purified chitinolytic enzymes. While bacterial chitinases are the subject of most of the assays described here, some protocols can be easily modified for use in fungal and other eukaryotic systems.

Assays to detect chitinolytic activity of culturable microorganisms

In many chitinolytic systems, chitin is hydrolyzed by secreted chitin depolymerases. Because these enzymes are able to diffuse through agar, assays to identify chitinolytic bacteria or genomic clones encoding chitin-degrading enzymes can be performed by monitoring the degradation of polymeric chitin incorporated into an agar medium. Though these assays have a limited sensitivity, they represent a simple and inexpensive method to identify chitinolytic microorganisms. Initial screens for chitinase activity should include both alpha and beta chitin, as different forms of chitin may induce the expression of specific chitinases (133). Alpha and beta chitin are incorporated into media at a concentration of 0.1% prior to heat sterilization. Alpha chitin isolated from shrimp or crab shells can be purchased from Sigma (C8908 and C9752, respectively), while beta chitin from squid pens can be obtained from Industrial Research Chemicals (Lower Hutt, New Zealand). Squid pen chitin has a large particle size and a plastic-like consistency, but can be macerated before addition to agar or broth cultures to increase the available surface area and facilitate degradation.

When screening environmental isolates for chitinolytic activity, assay media should not contain glucose, as catabolite repression of microbial chitin depolymerase activity is well documented (31, 59, 95). In the case of *E.coli* transformants

expressing a genomic library, LB agar supplemented with chitin is suitable. Halos or zones of clearing, usually 1-2 mm in diameter, will appear around colonies producing chitin depolymerase(s). Colonies should not be scored as negative until they have been incubated for several weeks; depolymerization of particulate chitin may take days or weeks to become visible, especially in the case of *E.coli* transformants expressing genomic libraries. Assay plates should be incubated in plastic bags with a damp paper towel to prevent drying during extended incubations.

Alternative substrates may be included in agar media to facilitate the visualization of chitin depolymerase activity. Chitin Azure (Sigma, C 3020) is prepared by covalently linking a soluble dye to colloidal chitin (129). When Chitin Azure is depolymerized, clear zones appear around colonies that are easier to visualize than the halos formed on chitin plates. Chitin Azure should be incorporated in media at a final concentration of 0.08%, autoclaved, and incubated as described above for chitin plates. Alternatively, ethylene glycol chitin [EGC] (Fisher, ICN157983-80) can be added to agar plates to screen for chitin depolymerase production (24). EGC is a soluble form of chitin that can be prepared by the method of Yamada *et al.* or purchased commercially (153) (Fisher Scientific). EGC has the consistency of Styrofoam™ and must be macerated before it will dissolve in water. Grinding EGC with a mortar and pestle facilitates suspension in water. The suspension will appear somewhat cloudy. EGC agar contains 0.01% trypan blue and 0.4% ethylene glycol chitin. Solutions of trypan blue and ethylene glycol chitin should be filter sterilized prior to addition to sterile, molten agar. After culturing, clear halos will appear around colonies of chitinolytic bacteria. Note that ethylene

glycol chitosan (Sigma, P 7364) can be utilized to assay for chitosanase production in the same manner.

Use of analogs for screening culturable bacteria or genomic libraries for chitinolytic activity

A variety of chitin analogs can be used to screen for the production of chitin-degrading enzymes. These analogs are especially effective for screening genomic libraries expressed in *E.coli*. Chitin depolymerase activity can be difficult to detect on chitin/agar plates because the production and secretion of foreign proteins may not be efficient in an *E.coli* host. In addition, chitin/agar screens identify only chitin depolymerase enzymes and not other chitin or chitooligosaccharide-modifying proteins. The *E.coli* strains EC300 (Epicentre Technologies), DH5 α , and DH5 α E (Invitrogen) are suitable hosts for genomic library screens because they do not cleave the analogs described below (M.Howard, unpublished observations).

Chitin analogs have been well documented as a valuable tool for initial identification of chitinolytic enzyme producing strains and clones (77, 91). 4-Methylumbelliferyl β -D-N,N'-diacetylchitobioside (MUF-diNAG) (Sigma, M 9763) and 4-Methylumbelliferyl- β -D-N,N',N''-triacetylchitotrioside (MUF-triNAG) (Sigma, M 5639) are prepared by linking a 4-methylumbelliferyl moiety to the reducing end of chitobiose or chitotriose, respectively (100). When cleaved by a chitin-degrading enzyme, methylumbelliferone is released which can be detected by its bright blue fluorescence under UV light. Quantification of this activity can be performed using a fluorometer (355 nm excitation, 465 nm emission) and a standard curve prepared with 4-methylumbelliferone (Sigma, M 1381). A variety of enzymes

can be detected with these analogs, including some chitin depolymerases, chitodextrinases, and N-acetylglucosaminidases (chitobiases). Chitinolytic enzymes can be tentatively classified (i.e., endo- vs. exo-acting) based on their activity toward analogs, but results should be confirmed by testing activity against unmodified chitooligosaccharides. MUF analog solutions should be prepared in 100 mM sodium phosphate buffer and stored at -20°C until use. These analogs can be used to screen outgrown colonies on agar by using a 0.7% agarose solution supplemented to a final concentration of 5 μM MUF-diNAG or MUF-triNAG (121) as a top agar overlay. After incubation of overlay plates at the desired temperature, colonies or transformants producing chitinolytic enzymes will be surrounded by a blue halo when visualized with UV light.

While the above plate screen is effective, diffusion of liberated 4-methylumbelliferone through the agarose may be problematic. An assay utilizing individual broth cultures of genomic clones is an alternative to agar media. Individual *E.coli* transformants are grown in 96 well microtiter plates at 25°C overnight and 1.5 nanomoles of a MUF analog is then added to each well and further incubated for 24 to 48 hours with gentle shaking (54). UV light can be used to visualize fluorescing cultures.

Several derivatives of MUF analogs have been prepared that can be used to specifically screen for chitodextrinase production. 4-methylumbelliferyl-N-N'-diacetyl-4-thio- β -chitobioside (Mu-TCB) and 4-methylumbelliferyl-N-N'-N''-triacetyl-4-thio- β -chitotrioside (Mu-TCT) contain thio-glycosidic linkages between GlcNAc residues (148). This alternative linkage prevents N-acetyl-glucosaminidases

from degrading the non-reducing end of these substrates. Mu-TCT has been demonstrated to be the optimal analog for chitodextrinase identification in enzymatic screens. To identify chitin depolymerase producing clones, which may also cleave Mu-TCB and Mu-TCT, all clones producing a positive reaction should be subsequently screened for chitin depolymerase activity using one of the methods described above.

The paranitrophenol [PNP] chitin analogs 4-nitrophenyl N-acetyl- β -D-glucosaminide (Sigma, N 9376), 4-nitrophenyl N,N'-diacetyl- β -D-chitobioside (Sigma, N 6133), and 4-nitrophenyl N,N',N''-triacetylchitotrioside (Sigma, N 8638) are well-established compounds for the analysis of chitinolytic enzymes, particularly N-acetyl- β -D-glucosaminidases (83, 125). PNP-analogs can be used much like the MUF analogs described above. When paranitrophenol is liberated from the carbohydrate polymer, it can be detected spectrophotometrically by monitoring the A_{410} . Svital *et al.* describe a method for analyzing crude soluble protein extracts from *E.coli* expressing chitinolytic enzymes cloned from *Vibrio harveyi* (133). Reaction buffer (10 mM Tris-Cl, 666 μ M PNP-chitobiose, pH 7.5) is combined with an enzyme preparation and incubated at a desired temperature. Stival *et al.* report success with this procedure after a 30 min incubation. Three volumes of a 1 M Tris base solution are added to stop the reaction. The amount of liberated PNP can then be quantified using a standard curve prepared with 4-nitrophenol (Sigma, 104-8).

Analogs can also be used to screen for the presence of chitosanases. Honda *et al.* utilized a chitin deacetylase purified from *C. lindemuthianum* to deacylate the GlcNAc residues of MUF-triNAG. MUF-triNAG was dissolved in 800 μ l reaction

buffer (10 mM sodium tetraborate-HCl, pH 8.4) and combined with 0.46 U of chitin deacetylase (51). The reaction was incubated at 37°C for 24 hours and the enzyme was then separated from the analog using a centrifugal filter. Deacylated MUF-triNAG (MUF-triGluN) can then be used to assay for chitosanase activity as described above. Though chitosanases have been shown to cleave all of the bonds in MUF-triGluN, cleavage of the methylumbelliferyl moiety is sufficient for detection of activity using assays similar to those described for MUF-triNAG/chitinase assay.

Detection of chitin modifying enzymes in culture supernatants and whole cell lysates using zymograms

Chitin depolymerases produced by an organism or a genomic clone can be identified and enumerated by using zymograms. Native or denaturing PAGE (73) can be employed in this type of zymography. Native gels do not require an overnight refolding step to restore enzymatic activity, but estimating the molecular weight of separated proteins requires additional steps (for method, see (46, 98)). Denaturing gels provide better separation of proteins and permit an estimation of molecular weight to be made directly, but are dependent upon successful refolding of the enzymes. A number of methods for detection of chitin depolymerase activity have been described that are amenable to both native and denaturing gels. Zymograms can be performed with whole cell lysates, membrane fractions, or concentrated culture supernatants. To ensure that supernatants are free of cells, they should be sterilized

with a 0.22 μm filter, and concentrated using a centrifugal filter device prior to analysis. A 50 to 100 fold concentration is recommended.

Because of its solubility, ethylene glycol chitin should be used as a substrate in zymograms when evaluating chitin depolymerase activity (142). EGC (0.01%) can be incorporated into an otherwise standard SDS-PAGE separating gel. Protein samples can then be fractionated by denaturing PAGE and incubated in refolding buffer (50 mM Tris base, 1 mM EDTA, 5 mM β -mercaptoethanol, pH 7.5) for 16 h at 4°C. Gels containing fractionated proteins should be washed in 10 mM Na phosphate buffer (pH 7.0) for one hour with several changes of wash buffer to remove trace amounts of EDTA and β -mercaptoethanol. Gels are then incubated in 10 mM sodium phosphate buffer (pH 7.0) for 12-36 hours at a desired temperature to allow refolded, active enzymes to degrade the ethylene glycol chitin incorporated into the separating gel. After incubation, zymograms are washed in 10 mM Tris-HCl (pH 7.4) containing 0.01% Calcofluor (Sigma, F 3543) for 5 minutes, taking care to fully dissolve the Calcofluor. The gel is then rinsed with dH₂O and periodically evaluated under UV light. Calcofluor will bind the glycol chitin and fluoresce brightly under UV light, while zones of depolymerase activity will appear dark.

An alternative method for producing zymograms has been described that detects not only the presence of some chitin depolymerases, but also chitodextrinase and N-acetylglucosaminidase activity (75). Proteins are fractionated by standard denaturing SDS-PAGE and incubated in refolding buffer as described in the previous screen. Gels are then placed in activity buffer (10 mM Na phosphate, 5 μM MUF analog of choice, pH 7.0) for 30 seconds. After incubation at the desired temperature,

gels are observed under UV light for activity. It is important that gels not be submerged during incubation in order to minimize diffusion of the released methylumbelliferone. Li *et al.* report detection of chitinase activity with this method after a five minute incubation.

Chitin deacetylase activity can also be detected in zymograms by incorporating EGC into the separating gel as in the above chitin depolymerase screen (141). Gels should be run as described above, stained with Calcofluor, and rinsed. Calcofluor has a higher binding affinity for chitosan than chitin, thus zones of chitin deacetylase activity will appear as bands of bright blue fluorescence against a blue fluorescent background.

Detection of chitinases by the release of reducing sugars

Chitooligomers, chitobiose, and GlcNAc, all of which have a reducing end, are released as chitin is degraded. This increase in reducing activity can be detected and quantified using a variety of methods. Commercially available kits for monitoring reducing sugar levels can be used, but an easy and inexpensive method described by Garcia *et al.* makes use of bicinchoninic acid (38) (Sigma, D 8284). Culture supernatants obtained from minimal media/chitin grown cultures can be tested for chitinolytic activity with this method. Additionally, mixtures containing purified chitinolytic enzymes and a solution of chitin or chitooligosaccharides can be analyzed to determine reaction rates (e.g., μM reducing sugar produced/min/mg enzyme) or substrate specificities.

To perform the Garcia protocol, prepare Solution A (per L: 54.28 g Na_2CO_3 , 24.0 g NaHCO_3 , 1.942 g 4,4'-dicarboxy-2,2'-biquinoline [bicinchoninic acid, disodium salt]) and Solution B (per L: 1.248 g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1.262 g L-serine). Equal amounts of Solutions A and B are mixed immediately before use to prepare the Reducing Sugar Assay Reagent. Reactions should be prepared by mixing equal volumes of Reducing Sugar Reagent and substrate (supernatant, enzyme reaction, or standard) and incubated at 80°C for 30 minutes. The samples are cooled to room temperature and the A_{560} measured. Reducing sugars are quantified by comparison to a standard curve prepared with GlcNAc. Note that the presence of protein will also produce a positive reaction, and therefore appropriate controls must be included.

An alternative to the Garcia protocol is the dinitrosalicylic acid reducing sugar test optimized by Miller (85). This assay makes use of a reagent that reacts only with reducing sugars and is not affected by the presence of protein or amino acids. After an enzyme/substrate reaction has been incubated for a desired time, add an equal volume of DNSA reagent (1% dinitrosalicylic acid, 0.2% phenol, 0.05% sodium sulfite, 1% sodium hydroxide) and boil for five minutes. The A_{575} is measured after boiling and levels of reducing sugars estimated by comparison to a standard curve prepared with the reducing sugar of interest (e.g., GlcNAc).

A method for the detection of reducing sugars by fluorescent labeling and subsequent HPLC of the products has been described (69). A fluorescent compound, 1-amino-1,3-naphthalene disulfonic acid, reacts with the reducing end of a sugar. The products of the reaction can then be separated by HPLC and the amount of labeled sugars quantified. Though more costly, this method has the advantage of

allowing the researcher to determine the abundance of specific chitooligosaccharides produced by a specific enzyme.

Degenerate PCR primers and probes for environmental DNA samples, culturable, and nonculturable microorganisms

Many microorganisms are not culturable on known laboratory media (29). The conserved nature of certain domains found within bacterial chitin-modifying proteins makes it possible to identify the cognate genes for chitin-processing enzymes from environmental samples by using molecular techniques. Many known microbial chitinases contain a GH18 domain and/or chitin-binding domains; GH18 domains are the predominant catalytic domain in microbial chitin depolymerases, and chitin-binding domains are common among bacterial chitinolytic enzymes (47). The nucleotide sequence corresponding to these conserved domains can be amplified from a known species, radiolabeled or labeled with a colorimetric detection kit, and used to probe a set of environmental isolates or a genomic library. Probes and degenerate primers can be used to identify chitinase genes from naked DNA in the environment or from a collection of microbes in an environmental sample that cannot be cultured.

Ramaiah *et al* (119) prepared a probe for chitinase genes derived from the *chiA* gene of *Vibrio harveyi*. The primers 5'-GATATCGACTGGGAGTTCCC-3' (forward) and 5'-CATAGAAGTCGTAGGTCATC-3' (reverse) were designed using the *chiA* sequences of *Serratia marcescens* (X03657, Z36294), *Alteromonas* sp. (D13762), *Bacillus circulans* (M57601), and *Aeromonas caviae* (U09139). These primers were then used to amplify a 225 bp fragment of *chiA* from *V.harveyi* (ATCC

14216) genomic DNA. This amplified fragment was radiolabeled and used to probe environmental isolates from the Chesapeake Bay using a standard hybridization. In some cases, a chitinase gene fragment could be directly amplified from colonies that hybridized the probe using the above primers.

Cottrell *et al.* (25) have described another set of primers that amplify chitinase genes, predominantly from α - and γ -proteobacteria isolated from environmental samples of marine origin. These primers are derived from a highly conserved region found within the *chiA* genes of *Alteromonas* sp. strain O-7, *Aeromonas caviae*, *Serratia marcescens*, and *Enterobacter agglomerans* and compensate for the variation in the third nucleotide position of each codon. The forward primer sequence is 5'-WSIGTIGGIGGITGGCANYT-3' and the reverse primer sequence is 5'-ATRTCICCRTTRTCIGCRTC-3' where W = A or T, S = C or G, Y = C or T, R = A or G, N = any nucleotide, and I = deoxyinosine. PCR conditions used to screen environmental isolates with these primers employed 35 cycles of: 1 min denaturation at 94°C, 1 min annealing at 50°C, and 3 min extension. The expected product size of a positive result is 900 bp. The amplified fragment can then be sequenced or ligated into an AT vector (Invitrogen) for further analysis. The nucleotide sequence of any amplified fragment must be determined to confirm the presence of a GH18-like sequence.

Detection of specific chitooligosaccharides

Biochemical assays are necessary to further classify chitinolytic enzymes. True rates of reaction and native substrate specificities are difficult to extrapolate from

enzyme/analog reactions. Therefore, it is important that methods for detection and quantification of native chitooligosaccharides be presented here.

Chitooligosaccharides are commercially available with degrees of polymerization ranging from 2 (chitobiose) to 8 (chitooctose) (Sigma). Degradative activity can be evaluated using chitin or various chitooligosaccharides and purified enzyme preparations followed by analysis to determine the products of each reaction.

To separate the reaction products for qualitative analysis, a relatively simple TLC procedure can be performed. Samples of reaction mixtures can be spotted on a silica gel plate and dried. Chromatograms are then developed using 2-propanol-ethanol-water at a ratio of 5:2:1 (v/v) (145). Plates should be thoroughly air-dried and then sprayed with 10% sulfuric acid in ethanol. The plates are again air-dried and baked at 120°C for 5-20 minutes. Chitooligosaccharides will appear as dark spots against a white background. Standards containing a mixture of chitooligosaccharides of various lengths are commercially available, or can be prepared from individual solutions. Chromatograms should be photographed soon after developing because spots will fade over time and are easily smeared.

An alternative method for fractionating chitin degradation products by TLC has been described. The chromatogram is developed with *n*-butanol-methanol-28% ammonia-water (5-4-2-1) (v/v). After drying, plates are sprayed with diphenylamine-aniline-phosphate reagent (0.4 g diphenylamine, 0.4 ml aniline, 3 ml 85% phosphoric acid, 20 ml acetone) and incubated at 80°C for 20 minutes. Chitooligosaccharides will appear as dark spots on a white background.

While TLC analysis allows for a qualitative assessment of product formation, it is relatively insensitive and not quantitative. More advanced and much more sensitive analysis can be performed to better characterize the products released by chitin-degrading enzymes using analytical instrumentation. Several methods have been described for the separation and quantification of chitooligosaccharides that employ HPLC. Difficulties involving limits of detection are reduced as compared to TLC, and quantification of end products is possible. HPLC has been used to determine the products formed by the action of chitodextrinases and N-acetylglucosaminidases, and also to determine the length of chitooligosaccharides released by the degradation of chitin (20).

Chang *et al.* found that a LiChrospher™ 100 NH₂ column (Fisher Scientific) with 5 µm packing coupled with an injection volume of 20 µl and a mobile phase composed of 75/25 (v/v) acetonitrile/water at a flow rate of 1ml/min resulted in optimum separation of chitooligosaccharides (20). Detection of chitooligosaccharides was performed by monitoring A₂₀₅ and peaks corresponding to specific products eluted as in Table A2-1. Other protocols that make use of different column packing and/or mobile phase components can also be explored (14, 30, 103).

Mass spectroscopy and NMR can also be utilized to determine the size of chitooligosaccharides produced by the activity of chitinases or chemical depolymerization of chitin. These methods can be modified for use with most mass spectrometers (78, 97, 118, 139, 156). Lopatin *et al.* report a method that allows for the identification of specific chitooligosaccharides within of a mixture containing chitooligosaccharides of varying lengths. A standard mass spectroscopic protocol

involves electrospray injection of a sample to determine the molecular weight of reaction products, though care should be taken to limit the salt and buffer content of the reaction.

Finally, oligosaccharide gel electrophoresis can be performed to determine chitooligosaccharide production or degradation in a reaction using a 2-aminobenzoic acid tag. A modified version of the Huang *et al.* protocol (55) has been used in our lab to label and separate chitooligosaccharides between 2 and 7 residues in length (Figure A2-1). To perform this procedure, chitooligosaccharide degradation products are dried under vacuum. Each reaction is then mixed with 100 μ l of labeling solution (1.0 M sodium cyanoborohydride, 0.2 M 2-aminobenzoic acid, prewarmed to 65°C). Standards are prepared by adding 10 μ l of a 1 mM solution of known chitooligosaccharides to 100 μ l labeling solution. Each labeling reaction is incubated at 65°C for two hours with periodic mixing. After cooling, 6 volumes of acetonitrile are added and samples are mixed briefly by vortexing. Products are then collected by centrifugation at 16,000 g for two minutes. Most of the supernatant (at least 500 μ l) should be discarded, taking care not to disturb or remove the gelatinous clusters at the bottom of the tube. Three hundred μ l of water is then added to each tube, thoroughly mixed, and dried under vacuum. Dried, labeled chitooligosaccharides should then be resuspended in 25 μ l SDS-PAGE loading buffer.

Table A2-1**Retention time of chitooligosaccharides using the Chang HPLC Method**

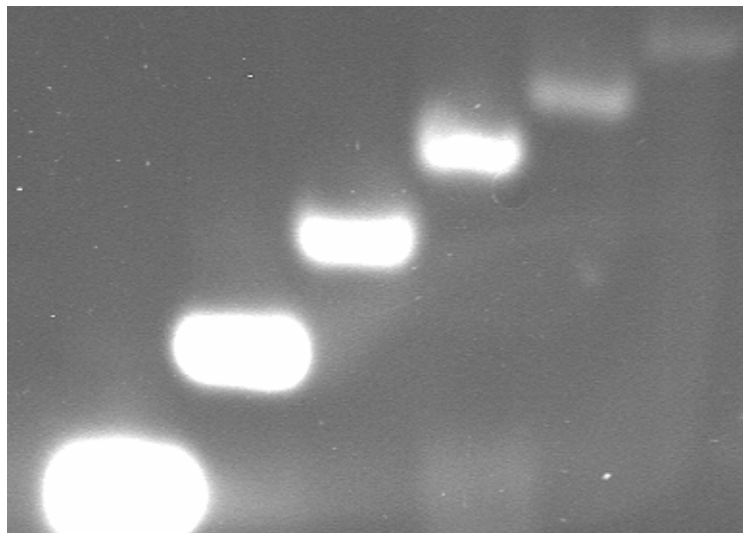
Retention time of chitooligosaccharides in a LiChrospher 100 NH₂ column with 5 micrometer packing chromatographed with 75/25 (v/v) acetonitrile/water at 1 ml/min.

Chitooligosaccharide	Retention time (min)
GlcNAc	5.73
(GlcNAc) ₂	8.40
(GlcNAc) ₃	12.00
(GlcNAc) ₄	17.87
(GlcNAc) ₅	26.67
(GlcNAc) ₆	40.07

Figure A2-1

2-aminobenzoic acid labeled chitooligosaccharides The reducing end of chitooligosaccharides can be labeled with 2-aminobenzoic acid and separated on a standard 15% SDS-PAGE gel. Labeled COS are then visualized under UV light. The numbers above the gel lanes indicate the degree of polymerization of the labeled chitooligosaccharides.

2 3 4 5 6 7



To separate the products of each labeling reaction, prepare a standard 1.5 mm, 10cm X 7cm SDS-PAGE gel (15% acrylamide, Tris-HCl buffer, per Sambrook) using a standard Tris-glycine buffer system. No stacking gel is necessary. Samples (2-10 μ l) are separated using constant current (45 mA) for 45 minutes. Labeled chitooligosaccharides can be visualized under UV light. The integrating function of gel visualization software may allow detection of 2-aminobenzoic acid- labeled saccharides that cannot be observed with the naked eye (52).

Detection of chitin-binding proteins and chitin-binding assays

A significant number of proteins associated with chitin degradation contain chitin-binding domains (47). These binding domains could conceivably be used to purify chitinolytic enzymes from protein preparations. In fact, a chitin-binding domain from a chitinase of *Bacillus circulans* is employed as a purification tag in the Impact™ protein expression systems (New England Biolabs, Beverly, MA).

Secreted, chitin-binding proteins or enzymes can be easily isolated from bacterial cultures by their affinity for chitin. This is especially useful in the event that promoters from an organism of interest are not active in *E.coli* or screening of genomic libraries is not possible due to gene toxicity. Chitin incorporated into liquid cultures can be isolated and washed to remove non-specifically adsorbed proteins. Many chitin-binding domains appear to bind the chitin polymer so tightly that conventional methods for eluting proteins from a matrix are ineffective. New England Biolabs reports that the *B. circulans* chitin-binding domain is not eluted from chitin

when treated with N-acetyl-glucosamine, chitooligosaccharides, or high concentrations of sodium chloride. SDS-PAGE loading dye can be used to elute bound proteins for use in zymograms. Alternatively, addition of 1% SDS or 6M guanidinium hydrochloride followed by incubation at 50-80°C can be used to denature and remove bound proteins (www.neb.com).

To determine if a purified protein of interest binds chitin, combine 1 mg binding substrate (α - or β -chitin, chitosan, etc.) with binding buffer (50 mM Tris-HCl, 0.1 M NaCl, pH 7.0) (145). After addition of a known amount of enzyme, reactions are incubated at a desired temperature for 1 hour with agitation. Reactions are then centrifuged at 25, 000 X g for ten minutes and the supernatant collected. The amount of protein in the supernatant is determined using a standard protein quantification protocol (e.g., a BSA protein quantification kit). The amount of bound protein is calculated by subtracting the amount of protein detected in the supernatant from the initial amount of protein added to the reaction.

Conclusions

The methods presented here can be used to identify and characterize new chitin degrading organisms and enzymes, and therefore increase our ability to modify chitin for applied uses. The non-toxic nature and absorbent properties of chitin make it environmentally safe for bioremediation and an excellent pharmaceutical tool for drug discovery and delivery. In addition, chitin is abundant and inexpensive, making it an attractive material for industrial processes and large-scale applications.

The conserved nature of several domains commonly found within chitin-modifying enzymes allows conserved and semi-conserved chitinases to be detected from raw nucleotide sequence. Extensive sequence analysis of the expanding list of microbial genome sequences is facilitated by the large number of web-based analysis tools, many of which are available at no cost to the researcher. Analysis of genomic sequence can lead to promising discoveries that must be subsequently demonstrated and characterized biochemically. Many of the methods presented here can be combined with bioinformatics to discover new and valuable chitin-modifying enzymes.

Bibliography

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**:3389-402.
2. **Andrykovich, G., and I. Marx.** 1988. Isolation of a new polysaccharide digesting bacterium from a salt marsh. *Appl Microbiol Biotechnol* **54**:1061-2.
3. **Anonymous** 10/12/2000 2000, posting date. IMPACT Protein Purification System FAQ. [Online.]
4. **Anonymous.** 1999. Guidance for Industry: Reducing microbial food safety hazards for sprouted seeds and guidance for industry: Sampling and microbial testing of spent irrigation water during sprout production. *Fed Regist.* **64**:57893-57902.
5. **Anonymous.** 1999. Microbiological Safety Evaluation and Recommendations on Sprouted Seeds. FDA/CFSAN.
6. **Anonymous.** 1997. Sanitary Guidelines for the growing and packing for sale of fresh sprouts. International Sprout Grower's Association.
7. **Backer, H., J. Mohle-Boetani, S. Werner, S. Abbott, J. Farrar, and D. Vugia.** 2000. High incidence of extra-intestinal infections in a *Salmonella* Havana outbreak associated with alfalfa sprouts. *Public Health Rep.* **115**:339-345.
8. **Bassler, B., P. Gibbons, and S. Roseman.** 1989. Chemotaxis to chitin oligosaccharides by *Vibrio furnissii*, a chitinivorous marine bacterium. *Biochem Biophys Res Commun* **161**:1172-6.
9. **Bassler, B. L., C. Yu, Y. C. Lee, and S. Roseman.** 1991. Chitin utilization by marine bacteria. Degradation and catabolism of chitin oligosaccharides by *Vibrio furnissii*. *J Biol Chem* **266**:24276-86.

10. **Beguin, P., and J. P. Aubert.** 1994. The biological degradation of cellulose. *FEMS Microbiol Rev* **13**:25-58.
11. **Beuchat, L., and J. Ryu.** 1997. Produce handling and processing practices. *Emerg. Infect. Dis.* **3**:459-465.
12. **Bhandari, D. G., B. A. Levine, I. P. Trayer, and M. E. Yeadon.** 1986. ¹H-NMR study of mobility and conformational constraints within the proline-rich N-terminal of the LC1 alkali light chain of skeletal myosin. Correlation with similar segments in other protein systems. *Eur J Biochem* **160**:349-56.
13. **Black, G. W., J. E. Rixon, J. H. Clarke, G. P. Hazlewood, M. K. Theodorou, P. Morris, and H. J. Gilbert.** 1996. Evidence that linker sequences and cellulose-binding domains enhance the activity of hemicellulases against complex substrates. *Biochem J* **319 (Pt 2)**:515-20.
14. **Blumberg, K., F. Liniere, L. Pustilnik, and C. A. Bush.** 1982. Fractionation of oligosaccharides containing N-acetyl amino sugars by reverse-phase high-pressure liquid chromatography. *Anal Biochem* **119**:407-12.
15. **Bouma, C. L., and S. Roseman.** 1996. Sugar transport by the marine chitinolytic bacterium *Vibrio furnissii*. Molecular cloning and analysis of the glucose and N-acetylglucosamine permeases. *J Biol Chem* **271**:33457-67.
16. **Brown, I. E., M. H. Mallen, S. J. Charnock, G. J. Davies, and G. W. Black.** 2001. Pectate lyase 10A from *Pseudomonas cellulosa* is a modular enzyme containing a family 2a carbohydrate-binding module. *Biochem J* **355**:155-65.
17. **Brurberg, M. B., I. F. Nes, and V. G. Eijsink.** 1996. Comparative studies of chitinases A and B from *Serratia marcescens*. *Microbiology* **142 (Pt 7)**:1581-9.
18. **Buchanan, B., W. Gruissem, and R. Jones.** 2000. *Biochemistry and Molecular Biology of Plants*. American Society of Plant Physiologists.
19. **Burton, J., S. G. Wood, A. Pedyczak, and I. Z. Siemion.** 1989. Conformational preferences of sequential fragments of the hinge region of human IgA1 immunoglobulin molecule: II. *Biophys Chem* **33**:39-45.
20. **Chang, K., J. Y. Lee, and W. Fu.** 2000. HPLC analysis of N-acetyl-chitooligosaccharides during the acid hydrolysis of chitin. *J Food Drug Anal* **8**:75-83.

21. **Charkowski, A., J. Barak, C. Sarreal, and R. Mandrell.** 2002. Differences in growth of *Salmonella enterica* and *E. coli* O157:H7 on alfalfa sprouts. *Appl. Environ. Microbiol.* **68**:3114-70.
22. **Collmer, A., J. Badel, A. Charkowski, W. Deng, D. Fouts, A. Ramos, A. Rehm, D. Anderson, O. Schneewind, K. van Dijk, and J. Alfano.** 2000. *Pseudomonas syringae* Hrp type III secretion system and effector proteins. *Proc. Natl. Acad. Sci (USA)* **97**:8770-7.
23. **Como-Sabetti, K.** 1997. Outbreaks of *E. coli* O157:H7 infection associated with eating alfalfa sprouts--Michigan and Virginia, June-July 1997. *Morb. Mortal. Wkly. Rep.* **46**:741-45.
24. **Connell, T. D., D. J. Metzger, J. Lynch, and J. P. Folster.** 1998. Endochitinase is transported to the extracellular milieu by the eps-encoded general secretory pathway of *Vibrio cholerae*. *J Bacteriol* **180**:5591-600.
25. **Cottrell, M. T., J. A. Moore, and D. L. Kirchman.** 1999. Chitinases from uncultured marine microorganisms. *Appl Environ Microbiol* **65**:2553-7.
26. **Cottrell, M. T., D. N. Wood, L. Yu, and D. L. Kirchman.** 2000. Selected chitinase genes in cultured and uncultured marine bacteria in the alpha- and gamma-subclasses of the proteobacteria. *Appl Environ Microbiol* **66**:1195-201.
27. **Cuero, R. G.** 1999. Antimicrobial action of exogenous chitosan. *Exs* **87**:315-33.
28. **Distel, D. L., W. Morrill, N. MacLaren-Toussaint, D. Franks, and J. Waterbury.** 2002. *Teredinibacter turnerae* gen. nov., sp. nov., a dinitrogen-fixing, cellulolytic, endosymbiotic gamma-proteobacterium isolated from the gills of wood-boring molluscs (Bivalvia: Teredinidae). *Int J Syst Evol Microbiol* **52**:2261-9.
29. **Edwards, C.** 2000. Problems posed by natural environments for monitoring microorganisms. *Mol Biotechnol* **15**:211-23.
30. **Fan, J. Q., A. Kondo, I. Kato, and Y. C. Lee.** 1994. High-performance liquid chromatography of glycopeptides and oligosaccharides on graphitized carbon columns. *Anal Biochem* **219**:224-9.
31. **Felse, P. A., and T. Panda.** 1999. Regulation and cloning of microbial chitinase genes. *Appl Microbiol Biotechnol* **51**:141-51.

32. **Ferreira, L. M., A. J. Durrant, J. Hall, G. P. Hazlewood, and H. J. Gilbert.** 1990. Spatial separation of protein domains is not necessary for catalytic activity or substrate binding in a xylanase. *Biochem J* **269**:261-4.
33. **Fett, W.** 2000. Naturally occurring biofilms on alfalfa and other types of sprouts. *J. Food Prot.* **63**:625-32.
34. **Flint, H. J., J. Martin, C. A. McPherson, A. S. Daniel, and J. X. Zhang.** 1993. A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domains, encoded by the *xynD* gene of *Ruminococcus flavefaciens*. *J Bacteriol* **175**:2943-51.
35. **Folders, J., J. Tommassen, L. C. van Loon, and W. Bitter.** 2000. Identification of a chitin-binding protein secreted by *Pseudomonas aeruginosa*. *J Bacteriol* **182**:1257-63.
36. **Fung, K. L., K. J. Zhao, Z. M. He, and M. L. Chye.** 2002. Tobacco-expressed *Brassica juncea* chitinase BjCHI1 shows antifungal activity in vitro. *Plant Mol Biol* **50**:283-94.
37. **Gao, J., M. W. Bauer, K. R. Shockley, M. A. Pysz, and R. M. Kelly.** 2003. Growth of Hyperthermophilic Archaeon *Pyrococcus furiosus* on Chitin Involves Two Family 18 Chitinases. *Appl Environ Microbiol* **69**:3119-28.
38. **Garcia.** 1993. Assessment of endo-1,4-beta-D-Glucanase activity by a rapid colorimetric assay using disodium 2,2'-bichinchoninate. *J Food Biochem* **17**:135-145.
39. **Garcia-Vallve, S., A. Romeu, and J. Palau.** 2000. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol Biol Evol* **17**:352-61.
40. **Gilkes, N. R., R. A. Warren, R. C. Miller, Jr., and D. G. Kilburn.** 1988. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *J Biol Chem* **263**:10401-7.
41. **Gonzalez, J., and R. Weiner.** 2000. Phylogenetic characterization of a marine bacterium strain 2-40, a degrader of complex polysaccharides. *Int J Syst Bacteriol* **8**:831-834.
42. **Gonzalez, J. M., F. Mayer, M. A. Moran, R. E. Hodson, and W. B. Whitman.** 1997. *Microbulbifer hydrolyticus* gen. nov., sp. nov., and *Marinobacterium georgiense* gen. nov., sp. nov., two marine bacteria from a lignin-rich pulp mill waste enrichment community. *Int J Syst Bacteriol* **47**:369-76.

43. **Hall, J., G. P. Hazlewood, N. S. Huskisson, A. J. Durrant, and H. J. Gilbert.** 1989. Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subspecies *cellulosa*: internal signal sequence and unusual protein processing. *Mol Microbiol* **3**:1211-9.
44. **Hansen-Wester, I., and M. Hensel.** 2001. *Salmonella* pathogenicity islands encoding type III secretion systems. *Microbes Infect* **3**:549-59.
45. **Hayatsu, H.** 2000. [Pigment molecules linked to polymer support: blue rayon, blue chitin, and green chitosan-synthesis and applications]. *Yakugaku Zasshi* **120**:534-47.
46. **Hedrick, J. L., and A. J. Smith.** 1968. Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch Biochem Biophys* **126**:155-64.
47. **Henrissat, B.** 1999. Classification of chitinases modules. *Exs* **87**:137-56.
48. **Henrissat, B., and A. Romeu.** 1995. Families, superfamilies and subfamilies of glycosyl hydrolases. *Biochem J* **311** (Pt 1):350-1.
49. **Hirano, S., and C. Upper.** 2000. Bacteria in the leaf ecosystem with emphasis on *Pseudomonas syringae*--a plant pathogen, ice nucleus, and epiphyte. *Microbiol. Mol. Bio. Rev.* **64**:624-53.
50. **Homerova, D., R. Knirschova, and J. Kormanec.** 2002. Response regulator ChiR regulates expression of chitinase gene, *chiC*, in *Streptomyces coelicolor*. *Folia Microbiol (Praha)* **47**:499-505.
51. **Honda, Y., M. Kiriata, T. Fukamizo, S. Kaneko, K. Tokuyasu, and R. Brzezinski.** 1999. Chitosanase-catalyzed hydrolysis of 4-methylumbelliferyl beta-chitotrioside. *J Biochem (Tokyo)* **126**:470-4.
52. **Howard, M. B., N. A. Ekborg, L. E. Taylor, 2nd, R. M. Weiner, and S. W. Hutcheson.** 2004. Chitinase B of "*Microbulbifer degradans*" 2-40 contains two catalytic domains with different chitinolytic activities. *J Bacteriol* **186**:1297-303.
53. **Howard, M. B., N. A. Ekborg, L. E. Taylor, S. W. Hutcheson, and R. M. Weiner.** 2004. Identification and analysis of polyserine linker domains in prokaryotic proteins with emphasis on the marine bacterium *Microbulbifer degradans*. *Protein Sci* **13**:1422-5.
54. **Howard, M. B., N. A. Ekborg, L. E. Taylor, R. M. Weiner, and S. W. Hutcheson.** 2003. Genomic Analysis and Initial Characterization of the

- Chitinolytic System of *Microbulbifer degradans* Strain 2-40. J Bacteriol **185**:3352-60.
55. **Huang, Z., T. Prickett, M. Potts, and R. F. Helm.** 2000. The use of the 2-aminobenzoic acid tag for oligosaccharide gel electrophoresis. Carbohydr Res **328**:77-83.
 56. **Hutcheson, S.** 2001. The molecular biology of hypersensitivity to plant pathogenic bacteria. J. Pl. Pathol. **83**:151-72.
 57. **Ilango, R., S. Kavimani, B. Jaykar, and G. Umamaheshwari.** 1999. Dissolution studies on tablets of ibuprofen using chitosan. Indian J Exp Biol **37**:505-8.
 58. **Illum, L.** 1998. Chitosan and its use as a pharmaceutical excipient. Pharm Res **15**:1326-31.
 59. **Ingram, C., and J. Westpheling.** 1995. The glucose kinase gene of *Streptomyces coelicolor* is not required for glucose repression of the *chi63* promoter. J Bacteriol **177**:3587-8.
 60. **Inouye, S., S. Wang, J. Sekizawa, S. Halegoua, and M. Inouye.** 1977. Amino acid sequence for the peptide extension on the prolipoprotein of the *Escherichia coli* outer membrane. Proc Natl Acad Sci U S A **74**:1004-8.
 61. **Kataeva, I. A., R. D. Seidel, 3rd, A. Shah, L. T. West, X. L. Li, and L. G. Ljungdahl.** 2002. The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. Appl Environ Microbiol **68**:4292-300.
 62. **Kellett, L. E., D. M. Poole, L. M. Ferreira, A. J. Durrant, G. P. Hazlewood, and H. J. Gilbert.** 1990. Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes. Biochem J **272**:369-76.
 63. **Keyhani, N., X. LI, and S. Roseman.** 2000. Identification and molecular cloning of a chitoporin. J Biol Chem **275**:33068-33076.
 64. **Keyhani, N., L. WAng, Y. Lee, and S. Roseman.** 2000. The chitin disaccharide, N,N'-diacetylchitobiose, is catabolized by *E.coli* and is transported/phosphorylated by the phosphoenolpyruvate:glycose phosphotransferase system. J Biol Chem **275**:33084-33090.
 65. **Keyhani, N. O., X. B. Li, and S. Roseman.** 2000. Chitin catabolism in the marine bacterium *Vibrio furnissii*. Identification and molecular cloning of a chitoporin. J Biol Chem **275**:33068-76.

66. **Keyhani, N. O., and S. Roseman.** 1996. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Molecular cloning, isolation, and characterization of a periplasmic beta-N-acetylglucosaminidase. *J Biol Chem* **271**:33425-32.
67. **Keyhani, N. O., and S. Roseman.** 1996. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Molecular cloning, isolation, and characterization of a periplasmic chitodextrinase. *J Biol Chem* **271**:33414-24.
68. **Keyhani, N. O., L. X. Wang, Y. C. Lee, and S. Roseman.** 1996. The chitin catabolic cascade in the marine bacterium *Vibrio furnissii*. Characterization of an N,N'-diacetyl-chitobiose transport system. *J Biol Chem* **271**:33409-13.
69. **Kim, Y. S., J. Liu, X. J. Han, A. Pervin, and R. J. Linhardt.** 1995. Analysis of fluorescently labeled sugars by reversed-phase ion-pairing high-performance liquid chromatography. *J Chromatogr Sci* **33**:162-7.
70. **Knowles, J., P. Lehtovaara, M. Penttila, T. Teeri, A. Harkki, and I. Salovuori.** 1987. The cellulase genes of *Trichoderma*. *Antonie Van Leeuwenhoek* **53**:335-41.
71. **Koga, D., M. Mitsutomi, M. Kono, and M. Matsumiya.** 1999. Biochemistry of chitinases. *Exs* **87**:111-23.
72. **Kormanec, J., B. Sevcikova, and D. Homerova.** 2000. Cloning of a two-component regulatory system probably involved in the regulation of chitinase in *Streptomyces coelicolor* A3(2). *Folia Microbiol (Praha)* **45**:397-406.
73. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-5.
74. **Langsford, M. L., N. R. Gilkes, B. Singh, B. Moser, R. C. Miller, Jr., R. A. Warren, and D. G. Kilburn.** 1987. Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. *FEBS Lett* **225**:163-7.
75. **Li, H., K. Morimoto, N. Katagiri, T. Kimura, K. Sakka, S. Lun, and K. Ohmiya.** 2002. A novel beta-N-acetylglucosaminidase of *Clostridium paraputrificum* M-21 with high activity on chitobiose. *Appl Microbiol Biotechnol* **60**:420-7.
76. **Li, X., and S. Roseman.** 2004. The chitinolytic cascade in *Vibrios* is regulated by chitin oligosaccharides and a two-component chitin catabolic sensor/kinase. *Proc Natl Acad Sci U S A* **101**:627-31.

77. **Lonhienne, T., K. Mavromatis, C. E. Vorgias, L. Buchon, C. Gerday, and V. Bouriotis.** 2001. Cloning, sequences, and characterization of two chitinase genes from the Antarctic *Arthrobacter* sp. strain TAD20: isolation and partial characterization of the enzymes. *J Bacteriol* **183**:1773-9.
78. **Lopatin, S. A., M. M. Ilyin, V. N. Pustobaev, Z. A. Bezchetnikova, V. P. Varlamov, and V. A. Davankov.** 1995. Mass-spectrometric analysis of N-acetylchitooligosaccharides prepared through enzymatic hydrolysis of chitosan. *Anal Biochem* **227**:285-8.
79. **Lucas, R., C. Lostroh, C. DiRusso, M. Spector, B. Wanner, and C. Lee.** 2000. Multiple factors independently regulate *hilA* and invasion gene expression in *Salmonella enterica* serovar Typhimurium. *J Bacteriol* **182**:1872-1882.
80. **Lustenberg, B., L. Dekkers, and G. Bloemberg.** 2001. Molecular determinants of rhizosphere colonization by Pseudomonads. *Annu. Rev. Phytopathol.* **39**:461-90.
81. **Madan Babu, M., and K. Sankaran.** 2002. DOLOP--database of bacterial lipoproteins. *Bioinformatics* **18**:641-643.
82. **Mahon, B.** 1997. An international outbreak of *Salmonella* infections caused by alfalfa sprouts grown from contaminated seed. *J. Infect. Dis.* **175**:876-82.
83. **Matsuo, Y., M. Kurita, J. K. Park, K. Tanaka, T. Nakagawa, M. Kawamukai, and H. Matsuda.** 1999. Purification, characterization and gene analysis of N-acetylglucosaminidase from *Enterobacter* sp. G-1. *Biosci Biotechnol Biochem* **63**:1261-8.
84. **McKie, V. A., J. P. Vincken, A. G. Voragen, L. A. van den Broek, E. Stimson, and H. J. Gilbert.** 2001. A new family of rhamnogalacturonan lyases contains an enzyme that binds to cellulose. *Biochem J* **355**:167-77.
85. **Miller, G.** 1959. Use of dinitrosalicylic acid reagent for detection of reducing sugar. *Anal Chem* **31**:426-428.
86. **Miller, R., C. Tate, E. Mallinson, and J. Scherrer.** 1991. Xylose-lysine-tergitol 4: an improved selective agar medium for the isolation of *Salmonella*. *Poult Sci* **70**:2429-32.
87. **Millward-Sadler, S. J., D. M. Poole, B. Henrissat, G. P. Hazlewood, J. H. Clarke, and H. J. Gilbert.** 1994. Evidence for a general role for high-affinity non-catalytic cellulose binding domains in microbial plant cell wall hydrolases. *Mol Microbiol* **11**:375-82.

88. **Miyamoto, K., E. Nukui, M. Hirose, F. Nagai, T. Sato, Y. Inamori, and H. Tsujibo.** 2002. A metalloprotease (MprIII) involved in the chitinolytic system of a marine bacterium, *Alteromonas* sp. strain O-7. *Appl Environ Microbiol* **68**:5563-70.
89. **Miyamoto, K., E. Nukui, H. Itoh, T. Sato, T. Kobayashi, C. Imada, E. Watanabe, Y. Inamori, and H. Tsujibo.** 2002. Molecular analysis of the gene encoding a novel chitin-binding protease from *Alteromonas* sp. strain O-7 and its role in the chitinolytic system. *J Bacteriol* **184**:1865-72.
90. **Mohle-Boetani, J.** 2001. *E. coli* O157:H7 and *Salmonella* infections associated with sprouts in California 1996-98. *Ann. Intern. Med.* **135**:239-247.
91. **Morimoto, K., S. Karita, T. Kimura, K. Sakka, and K. Ohmiya.** 1997. Cloning, sequencing, and expression of the gene encoding *Clostridium paraputrificum* chitinase ChiB and analysis of the functions of a novel cadherin-like domains and a chitin-binding domain. *J Bacteriol* **179**:7306-14.
92. **Muzzarelli, R.** 1999. Native, industrial, and fossil chitins. *In* M. RA (ed.), *Chitin and Chitinases*. Birkhauser, Basel.
93. **Muzzarelli, R. A., M. Mattioli-Belmonte, A. Pugnali, and G. Biagini.** 1999. Biochemistry, histology and clinical uses of chitins and chitosans in wound healing. *Exs* **87**:251-64.
94. **Netherwood, T., R. Bowden, P. Harrison, A. G. O'Donnell, D. S. Parker, and H. J. Gilbert.** 1999. Gene transfer in the gastrointestinal tract. *Appl Environ Microbiol* **65**:5139-41.
95. **Nguyen, J., F. Francou, M. J. Virolle, and M. Guerin.** 1997. Amylase and chitinase genes in *Streptomyces lividans* are regulated by reg1, a pleiotropic regulatory gene. *J Bacteriol* **179**:6383-90.
96. **Nielsen, H., J. Englebrecht, S. Brunak, and G. von Heijne.** 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10**:1-6.
97. **Nielsen, K. K., K. Bojsen, P. Roepstorff, and J. D. Mikkelsen.** 1994. A hydroxyproline-containing class IV chitinase of sugar beet is glycosylated with xylose. *Plant Mol Biol* **25**:241-57.
98. **Nishizawa, H., N. Kita, S. Okimura, E. Takao, and Y. Abe.** 1988. Determination of molecular weight of native proteins by polyacrylamide gradient gel electrophoresis. *Electrophoresis* **9**:803-6.

99. **No, H. K., and S. P. Meyers.** 2000. Application of chitosan for treatment of wastewaters. *Rev Environ Contam Toxicol* **163**:1-27.
100. **O'Brien, M., and R. R. Colwell.** 1987. A rapid test for chitinase activity that uses 4-methylumbelliferyl-N-acetyl-beta-D-glucosaminide. *Appl Environ Microbiol* **53**:1718-20.
101. **Ohta, Y., Y. Hatada, Y. Nogi, M. Miyazaki, Z. Li, M. Akita, Y. Hidaka, S. Goda, S. Ito, and K. Horikoshi.** 2004. Enzymatic properties and nucleotide and amino acid sequences of a thermostable beta-agarase from a novel species of deep-sea *Microbulbifer*. *Appl Microbiol Biotechnol* **64**:505-14.
102. **Ohta, Y., Y. Nogi, M. Miyazaki, Z. Li, Y. Hatada, S. Ito, and K. Horikoshi.** 2004. Enzymatic Properties and Nucleotide and Amino Acid Sequences of a Thermostable beta-Agarase from the Novel Marine Isolate, JAMB-A94. *Biosci Biotechnol Biochem* **68**:1073-81.
103. **Ohtakara, A., H. Matsunaga, and M. Mitsutomi.** 1990. Action pattern of *Streptomyces griseus* chitinase on partially N-acetylated chitosan. *Agric Biol Chem* **54**:3191-9.
104. **Onsoyen, E., and O. Skaugrud.** 1990. Metal recovery using chitosan. *J Chem Technol Biotechnol* **49**:395-404.
105. **Paradis, F. W., H. Zhu, P. J. Krell, J. P. Phillips, and C. W. Forsberg.** 1993. The *xynC* gene from *Fibrobacter succinogenes* S85 codes for a xylanase with two similar catalytic domains. *J Bacteriol* **175**:7666-72.
106. **Plotnikova, J., L. Rahme, and F. Ausubel.** 2000. Pathogenesis of the human opportunistic pathogen *P. aeruginosa* PA14 in *Arabidopsis*. *Plant Physiol* **124**:1766-1774.
107. **Plumbridge, J.** 2001. Regulation of PTS gene expression by the homologous transcriptional regulators, Mlc and NagC, in *Escherichia coli* (or how two similar repressors can behave differently). *J Mol Microbiol Biotechnol* **3**:371-80.
108. **Plumbridge, J., and A. Kolb.** 1991. CAP and Nag repressor binding to the regulatory regions of the *nagE-B* and *manX* genes of *Escherichia coli*. *J Mol Biol* **217**:661-79.
109. **Plumbridge, J. A.** 1990. Induction of the nag regulon of *Escherichia coli* by N-acetylglucosamine and glucosamine: role of the cyclic AMP-catabolite activator protein complex in expression of the regulon. *J Bacteriol* **172**:2728-35.

110. **Plumbridge, J. A.** 1989. Sequence of the *nagBACD* operon in *Escherichia coli* K12 and pattern of transcription within the nag regulon. *Mol Microbiol* **3**:505-15.
111. **Portnoy, B., J. Goepfert, and S. Harmon.** 1976. An outbreak of *B. cereus* food poisoning resulting from contaminated vegetable sprouts. *Am. J. Epidemiol.* **103**:589-94.
112. **Preston, G., N. Bertrand, and P. Rainey.** 2001. Type III secretion in plant growth-promoting *P. fluorescens* SBW25. *Molecular Microbiol.* **41**:999-1014.
113. **Proctor, M., M. Hamacher, M. Tortorello, J. Archer, and J. Davis.** 2001. Multistate outbreak of *Salmonella* Serovar Muenchen infections associated with alfalfa sprouts grown from seeds pretreated with calcium hypochlorite. *J. Clin. Microbiol.* **39**:3461-3465.
114. **Pruzzo, C., A. Crippa, S. Bertone, L. Pane, and A. Carli.** 1996. Attachment of *Vibrio alginolyticus* to chitin mediated by chitin-binding proteins. *Microbiology* **142** (Pt 8):2181-6.
115. **Pugsley, A. P., and M. G. Kornacker.** 1991. Secretion of the cell surface lipoprotein pullulanase in *Escherichia coli*. Cooperation or competition between the specific secretion pathway and the lipoprotein sorting pathway. *J Biol Chem* **266**:13640-5.
116. **Puohiniemi, R., T. Heiskanen, and A. Siitonen.** 1997. Molecular epidemiology of two international sprout-borne *Salmonella* outbreaks. *J. Clin. Microbiol.* **35**:2487-91.
117. **Radford, S. E., E. D. Laue, R. N. Perham, S. R. Martin, and E. Appella.** 1989. Conformational flexibility and folding of synthetic peptides representing an interdomain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Biol Chem* **264**:767-75.
118. **Rajamohanam, P. R., S. Ganapathy, P. R. Vyas, A. Ravikumar, and M. V. Deshpande.** 1996. Solid-state CP/MAS ¹³C-NMR spectroscopy: a sensitive method to monitor enzymatic hydrolysis of chitin. *J Biochem Biophys Methods* **31**:151-63.
119. **Ramaiah, N., R. T. Hill, J. Chun, J. Ravel, M. H. Matte, W. L. Straube, and R. R. Colwell.** 2000. Use of a *chiA* probe for detection of chitinase genes in bacteria from the Chesapeake Bay(1). *FEMS Microbiol Ecol* **34**:63-71.

120. **Riemann, L., and F. Azam.** 2002. Widespread N-Acetyl-D-Glucosamine Uptake among Pelagic Marine Bacteria and Its Ecological Implications. *Appl Environ Microbiol* **68**:5554-62.
121. **Robbins, P. W., C. Albright, and B. Benfield.** 1988. Cloning and expression of a *Streptomyces plicatus* chitinase (chitinase-63) in *Escherichia coli*. *J Biol Chem* **263**:443-7.
122. **Robbins, P. W., K. Overbye, C. Albright, B. Benfield, and J. Pero.** 1992. Cloning and high-level expression of chitinase-encoding gene of *Streptomyces plicatus*. *Gene* **111**:69-76.
123. **Saito, A., M. Ishizaka, P. B. Francisco, Jr., T. Fujii, and K. Miyashita.** 2000. Transcriptional co-regulation of five chitinase genes scattered on the *Streptomyces coelicolor* A3(2) chromosome. *Microbiology* **146 (Pt 11)**:2937-46.
124. **Saito, A., K. Miyashita, G. Biukovic, and H. Schrempf.** 2001. Characteristics of a *Streptomyces coelicolor* A3(2) extracellular protein targeting chitin and chitosan. *Appl Environ Microbiol* **67**:1268-73.
125. **Sakai, K., M. Narihara, Y. Kasama, M. Wakayama, and M. Moriguchi.** 1994. Purification and characterization of thermostable beta-N-acetylhexosaminidase of *Bacillus stearotheophilus* CH-4 isolated from chitin-containing compost. *Appl Environ Microbiol* **60**:2911-5.
126. **Schultz, J., R. R. Copley, T. Doerks, C. P. Ponting, and P. Bork.** 2000. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* **28**:231-4.
127. **Shen, H., M. Schmuck, I. Pilz, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. Warren.** 1991. Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. *J Biol Chem* **266**:11335-40.
128. **Singla, A. K., and M. Chawla.** 2001. Chitosan: some pharmaceutical and biological aspects--an update. *J Pharm Pharmacol* **53**:1047-67.
129. **Somers, P. J., R. C. Yao, L. E. Doolin, M. J. McGowan, D. S. Fukuda, and J. S. Mynderse.** 1987. Method for the detection and quantitation of chitinase inhibitors in fermentation broths; isolation and insect life cycle effect of A82516. *J Antibiot (Tokyo)* **40**:1751-6.
130. **Staskawicz, B., M. Mudgett, J. Dangl, and J. Galan.** 2001. Common and contrasting themes of plant and animal diseases. *Science* **292**:2285-9.

131. **Stewart, D., K. Reineke, J. Ulaszek, and M. Tortorello.** 2001. Growth of *Salmonella* during sprouting of alfalfa seeds associated with salmonellosis outbreaks. *J. Food Protec.* **64**:618-22.
132. **Suzuki, K., N. Sugawara, M. Suzuki, T. Uchiyama, F. Katouno, N. Nikaidou, and T. Watanabe.** 2002. Chitinases A, B, and C1 of *Serratia marcescens* 2170 produced by recombinant *Escherichia coli*: enzymatic properties and synergism on chitin degradation. *Biosci Biotechnol Biochem* **66**:1075-83.
133. **Svitil, A., S. Chadhain, J. Moore, and D. Kirchman.** 1997. Chitin degradation proteins produced by the marine bacterium *Vibrio harveyi* growing on different forms of chitin. *Appl Environ Microbiol* **63**:408-13.
134. **Tanaka, T., S. Fujiwara, S. Nishikori, T. Fukui, M. Takagi, and T. Imanaka.** 1999. A unique chitinase with dual active sites and triple substrate binding sites from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. *Appl Environ Microbiol* **65**:5338-44.
135. **Tanaka, T., T. Fukui, and T. Imanaka.** 2001. Different cleavage specificities of the dual catalytic domains in chitinase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J Biol Chem* **276**:35629-35.
136. **Tanaka, T., L. Yan, and J. G. Burgess.** 2003. *Microbulbifer arenaceus* sp. nov., a new endolithic bacterium isolated from the inside of red sandstone. *Curr Microbiol* **47**:412-6.
137. **Taormina, P., C. Beuchat, and L. Slutsker.** 1999. Infections associated with eating seed sprouts: An international concern. *Emerging Infec. Dis.* **5**:626-634.
138. **Techkarnjanaruk, S., and A. E. Goodman.** 1999. Multiple genes involved in chitin degradation from the marine bacterium *Pseudoalteromonas* sp. strain S91. *Microbiology* **145 (Pt 4)**:925-34.
139. **Tokuyasu, K., M. Mitsutomi, I. Yamaguchi, K. Hayashi, and Y. Mori.** 2000. Recognition of chitooligosaccharides and their N-acetyl groups by putative subsites of chitin deacetylase from a deuteromycete, *Colletotrichum lindemuthianum*. *Biochemistry* **39**:8837-43.
140. **Tomme, P., H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens.** 1988. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem* **170**:575-81.

141. **Trudel, J., and A. A.** 1990. Detection of chitin deacetylase activity after polyacrylamide gel electrophoresis. *Anal Biochem* **189**:249-253.
142. **Trudel, J., and A. Asselin.** 1989. Detection of chitinase activity after polyacrylamide gel electrophoresis. *Anal Biochem* **178**:362-6.
143. **Tsujibo, H., N. Hatano, T. Okamoto, H. Endo, K. Miyamoto, and Y. Inamori.** 1999. Synthesis of chitinase in *Streptomyces thermoviolaceus* is regulated by a two-component sensor-regulator system. *FEMS Microbiol Lett* **181**:83-90.
144. **Tsujibo, H., T. Okamoto, N. Hatano, K. Miyamoto, T. Watanabe, M. Mitsutomi, and Y. Inamori.** 2000. Family 19 chitinases from *Streptomyces thermoviolaceus* OPC-520: molecular cloning and characterization. *Biosci Biotechnol Biochem* **64**:2445-53.
145. **Tsujibo, H., H. Orikoshi, N. Baba, M. Miyahara, K. Miyamoto, M. Yasuda, and Y. Inamori.** 2002. Identification and characterization of the gene cluster involved in chitin degradation in a marine bacterium, *Alteromonas* sp. strain O-7. *Appl Environ Microbiol* **68**:263-70.
146. **Uchiyama, T., R. Kaneko, J. Yamaguchi, A. Inoue, T. Yanagida, N. Nikaidou, M. Regue, and T. Watanabe.** 2003. Uptake of N,N'-diacetylchitobiose [(GlcNAc)₂] via the phosphotransferase system is essential for chitinase production by *Serratia marcescens* 2170. *J Bacteriol* **185**:1776-82.
147. **Venter, J. C., K. Remington, J. F. Heidelberg, A. L. Halpern, D. Rusch, J. A. Eisen, D. Wu, I. Paulsen, K. E. Nelson, W. Nelson, D. E. Fouts, S. Levy, A. H. Knap, M. W. Lomas, K. Nealson, O. White, J. Peterson, J. Hoffman, R. Parsons, H. Baden-Tillson, C. Pfannkoch, Y. H. Rogers, and H. O. Smith.** 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**:66-74.
148. **Wang, L., N. Keyhani, S. Roseman, and Y. Lee.** 1997. 4-Methylumbelliferyl glycosides of N-acetyl 4-thiochito-oligosaccharides as fluorogenic substrates for chitodextrinase from *Vibrio furnissii*. *Glycobiology* **7**:855-60.
149. **Warren, R. A., C. F. Beck, N. R. Gilkes, D. G. Kilburn, M. L. Langsford, R. C. Miller, Jr., G. P. O'Neill, M. Scheufens, and W. K. Wong.** 1986. Sequence conservation and region shuffling in an endoglucanase and an exoglucanase from *Cellulomonas fimi*. *Proteins* **1**:335-41.

150. **Watanabe, T., K. Kimura, T. Sumiya, N. K Nikaidou, K. Suzuki, M. Suzuki, M. Taiyoji, S. Ferrer, and M. Regue.** 1997. Genetic analysis of the chitinase system of *Serratia marcescens* 2170. *J Bacteriol* **179**:7111-17.
151. **Weiner, R., L. Taylor, N. Ekborg, and L. Whitehead.** 2001. Degradosomes: Potential importance in the ocean's carbon cycle and in aquaculture and algalculture. *In* N. Seaxena (ed.), *Recent Advances in Marine Science and Technology*, 2000. Pacon International, New York.
152. **Whitehead, L. A., S. K. Stosz, and R. M. Weiner.** 2001. Characterization of the agarase system of a multiple carbohydrate degrading marine bacterium. *Cytobios* **106 Suppl 1**:99-117.
153. **Yamada, H., and T. Imoto.** 1981. A convenient synthesis of glycolchitin, a substrate of lysozyme. *Carbohydr Res* **92**:160-2.
154. **Yoon, J. H., H. Kim, K. H. Kang, T. K. Oh, and Y. H. Park.** 2003. Transfer of *Pseudomonas elongata* Humm 1946 to the genus *Microbulbifer* as *Microbulbifer elongatus* *comb. nov.* *Int J Syst Evol Microbiol* **53**:1357-61.
155. **Yoon, J. H., I. G. Kim, D. Y. Shin, K. H. Kang, and Y. H. Park.** 2003. *Microbulbifer salipaludis* *sp. nov.*, a moderate halophile isolated from a Korean salt marsh. *Int J Syst Evol Microbiol* **53**:53-7.
156. **Zhang, H., Y. Du, X. Yu, M. Mitsutomi, and S. Aiba.** 1999. Preparation of chitooligosaccharides from chitosan by a complex enzyme. *Carbohydr Res* **320**:257-60.
157. **Zobell, C., and S. Rittenberg.** 1937. The occurrence and characteristics of chitinoclastic bacteria in the sea. *J Bacteriol* **35**:275-287.
158. **Zverlov, V., S. Mahr, K. Riedel, and K. Bronnenmeier.** 1998. Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum thermophilum*' with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology* **144 (Pt 2)**:457-65.
159. **Zverlov, V. V., K. P. Fuchs, and W. H. Schwarz.** 2002. Chi18A, the endochitinase in the cellulosome of the thermophilic, cellulolytic bacterium *Clostridium thermocellum*. *Appl Environ Microbiol* **68**:3176-9.