

**Metagenomanalyse eines hydrolytischen Konsortiums:
Identifizierung und biochemische Charakterisierung von
Polysaccharid-abbauenden Biokatalysatoren aus nicht
kultivierten Mikroorganismen**

Dissertation

zur Erlangung des Doktorgrades

der Mathematischen-Naturwissenschaftlichen Fakultäten

der Georg-August-Universität zu Göttingen

vorgelegt von

Sonja Voget

aus Hamburg

Göttingen 2006

D7

Referent: Prof. Dr. Wolfgang Streit

Korreferent: PD Dr. Rolf Daniel

Tag der mündlichen Prüfung:

Abkürzungen.....	vi
I. Einleitung.....	1
1. Das Potential der Metagenomik.....	1
2. Metagenomik und die weiße Biotechnologie.....	2
3. Cellulasen.....	4
II. Material und Methoden.....	7
1. Verwendete Organismen, Plasmide und Oligonukleotide.....	7
2. Nährmedien.....	9
2.1 LB-(Luria-Bertani) Medium (Sambrook et al., 1989).....	9
2.2 TB-Medium (Sambrook et al., 1989).....	9
2.3 CMC (Carboxymethylcellulose)- Medium (Teather, 1982).....	9
2.4 Mineralmedium.....	10
2.5 Antibiotika und sonstige Medienzusätze.....	11
3. Wachstumsbedingungen und Stammhaltung.....	11
3.1 Zellanzucht.....	11
3.2 Stammkulturen.....	12
3.3 Messung der optischen Dichte.....	12
4. Standardtechniken für das Arbeiten mit Nukleinsäuren.....	12
4.1 Behandlung von Geräten und Lösungen für das Arbeiten mit Nukleinsäuren.....	12
4.2 Reinigung und Konzentrierung von Nukleinsäuren.....	12
4.2.1 Phenol/Chloroform-Extraktion.....	12
4.2.2 Fällung von DNA mit Ethanol oder Isopropanol.....	13
4.2.3 Dialyse von Gesamt-DNA.....	13
4.2.4 Konzentrationsbestimmung von DNA.....	13
4.3 Methoden zur Charakterisierung von DNA-Fragmenten.....	14
4.3.1 Agarosegelelektrophorese und Ethidiumbromid-Färbung von DNA.....	14
4.4 Isolierung von DNA.....	15
4.4.1 Isolierung von Gesamt-DNA des bakteriellen Konsortiums.....	15
4.4.2 Isolierung von Plasmid-DNA durch alkalische Lyse.....	16
4.4.3 Isolierung von Plasmid-DNA mit Hilfe des „QIAprep Spin Miniprep“.....	17

4.4.4 Isolierung von Plasmid-DNA mit Hilfe des „QIAprep Spin Midiprep“	17
4.4.5 Elution von DNA-Fragmenten aus Agarosegelen	17
4.5 Enzymatische Modifikationen von DNA	17
4.5.1 DNA-Verdau durch Restriktionsendonukleasen.....	17
4.5.2 Partieller Verdau von Gesamt-DNA	18
4.5.3 Dephosphorylierung von DNA-Enden	18
4.5.4 Ligation von DNA-Fragmenten.....	19
4.6 Amplifikation von DNA durch Polymerase-Kettenreaktion (PCR)	19
4.6.1 Amplifikation von DNA-Fragmenten mit der <i>Taq</i> - o. <i>Pfu</i> -DNA Polymerase .	19
4.6.2 Amplifikation von DNA-Fragmenten mit der KOD Hifi DNA Polymerase	20
4.7 Ligation von PCR Produkten in den pDrive Klonierungs-Vektor	21
4.8 Transposonmutagenese	21
5. Transformationsverfahren	22
5.1 Hocheffiziente Transformation von <i>E. coli</i>	22
5.1.1 Herstellung von kompetenten Zellen mit Hilfe der CaCl ₂ Methode.....	22
5.1.2 Transformation von kompetenten Zellen.....	23
5.1.3 Selektion rekombinanter <i>E. coli</i> Klone durch X-Gal-Test	23
5.2 Elektroporation.....	23
5.2.1 Herstellung „kompetenter Zellen“ für die Elektroporation.....	23
5.2.2 Transformation mittels Elektroporation.....	24
5.4 Transduktion mit dem “Gigapack® III Gold Packaging Extract”	24
5.4.1 Herstellung phagenkompetenter Zellen	25
5.4.2 Titerbestimmung	25
6. Die Prozessierung von Rohsequenzen.....	26
6.1 Prozessierung von DNA-Sequenzen mit dem STADEN-Software-Paket....	26
6.1.1 „Pregap4“	26
6.1.2 „Gap4“	27
7. Überwindung von Sequenzlücken.....	28
7.1 Identifikation von Contigüberlappungen	28
7.2 „Primer-Walking“	28
7.3 Kombinatorische PCR.....	29
8. Sequenzanalyse	29
8.1 ORF -Vorhersage.....	29

8.2 Untersuchung der 16S rRNA-Gene auf chimere Sequenzen.....	30
9. Proteinbiochemische Methoden.....	31
9.1 Standardtechniken	31
9.1.2 Quantitative Proteinbestimmung (Bradfordtest)	31
9.1.3 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE).....	31
9.1.4 Molekularmassenbestimmung von Proteinen in SDS-Gelen	33
9.1.5 Coomassie-Färbung.....	34
9.1.6 Proteinfärbung mit Ponceau S Rot.....	34
9.1.7 Umpufferung durch Dialyse.....	35
9.1.8 Konzentration von Proteinlösungen	35
9.2 Zellernte und Rohextraktgewinnung	35
9.2.1 Zellernte	35
9.2.2 Zellaufschluß mit der „French Pressure Cell“	35
9.2.3 Zellaufschluß durch Ultraschall	36
9.3 Proteinreinigung durch FPLC.....	36
9.3.1 Ionenaustausch-Chromatographie (IEC).....	37
9.3.2 Hydrophobe-Interaktions-Chromatographie (HIC)	38
9.3.3 Gelfiltration (SEC)	38
9.4 Immunologische Methoden	39
9.4.1 Western-Blot (Towbin et al., 1992).....	39
9.4.2 Nachweis von Proteinen mit aufeinanderfolgenden Histidinen	40
9.4.3 Detektion	41
9.5 Nachweis von Enzymaktivitäten.....	41
9.5.1 Qualitativer Agarase-Aktivitätstest	41
9.5.2 DNSA-Test zur Bestimmung der Menge reduzierender Zucker	42
9.5.3 Produktanalyse durch Dünnschichtchromatographie (TLC).....	43
9.5.4 Messung der Pektat Lyase Enzymaktivität.....	45
9.6 Charakterisierung der Enzymaktivität	46
9.6.1 pH-Abhängigkeit der Enzymaktivität	46
9.6.2 Temperaturabhängigkeit der Enzymaktivität	47
9.6.3 Einfluß von Metallionen und Chelatoren auf die Enzymaktivität	47
III. Ergebnisse.....	48
1. Analyse der phylogenetischen Diversität mittels 16S rRNA-Genanalyse	48
1.1 Vereinzelung des agarolytischen Konsortiums	49

1.2 DNA-Isolierung und Klonierung der 16S rRNA-Gene	49
1.3 Sequenzanalyse	50
2. Durchmusterung der Metagenom-Bank nach biotechnologisch interessanten Enzymen	54
2.1 Durchmusterung der Metagenom-Bank nach cellulolytischer Aktivität.....	54
2.1.1 Subklonierung von Cosmiden mit cellulolytischer Aktivität.....	57
2.1.2 Identifizierung der cellulolytischen ORFs mittels Transposonmutagenese ..	57
2.1.3 Sequenzanalyse des Klons pSVCEL-1	58
2.2 Durchmusterung der Metagenom-Bank nach agarolytischer Aktivität.....	62
2.2.1 Sequenzierung von Cosmiden mit agarolytischer Aktivität	62
2.2.2 Analyse der Cosmide VII C10, XIV A3, III E5 und X B11	63
2.2.2.1 Analyse des Cosmids VII C10	64
2.2.2.2 Analyse des Cosmids XIV A3	66
2.2.2.3 Analyse des Cosmids III E5	68
2.2.2.4 Analyse des Cosmids X B11	69
2.3 Sequenzbasierende Durchmusterung der Cosmide III F2, VII H3, XIV B3, VII A1 und VII D9 nach Enzymen von industrieller Bedeutung	70
2.3.1 Analyse der Cosmide III F2, VII H3, XIV B3, VII A1 und VII D9	70
2.3.1.1 Analyse des Cosmids III F2	71
2.3.1.2 Analyse des Cosmids VII H3	72
2.3.1.3 Analyse des Cosmids XIV B3	73
2.3.1.4 Analyse der Cosmide VII A1 und VII D9	73
3. Nachweis der Funktionalität ausgewählter Polysaccharid-modifizierender Enzyme der Klone pSVCEL-1 und VII C10	76
3.1 Funktionalitätsnachweis der Pektat-Lyase PelB des Klons pSVCEL-1	76
3.2 Funktionalitätsnachweis und Expressionsstudien zu der Agarase AguB des Klons VII C10	78
4. Klonierung, heterologe Expression und Aufreinigung der Cellulase CelA.....	80
4.1 Aufreinigung von CelA mittels Anionenaustauschchromatographie	81
4.2 Aufreinigung von CelA mittels hydrophober Interaktionschromatographie ..	83
4.3 Aufreinigung von CelA mittels Gelfiltration	85
5. Biochemische Charakterisierung von CelA.....	87
5.1 Bestimmung des Temperaturspektrums	87
5.2 Bestimmung der Thermostabilität	87
5.3 Bestimmung des pH-Spektrums	89

5.4 Abhängigkeit der Cellulase von verschiedenen Additiven.....	90
5.4.1 Auswirkungen von einwertigen Metallionen auf die Enzymaktivität	90
5.4.2 Auswirkungen von mehrwertigen Metallionen auf die Enzymaktivität.....	92
5.4.3 Einfluss weiterer Zusätze auf die Aktivität.....	93
5.5 Substratspektrum der Cellulase CelA	94
5.5.1 Qualitative Untersuchung des Substratspektrums und des Spaltungsmechanismus mittels Dünnschichtchromatographie	95
IV. Diskussion	98
1. Phylogenetische Analyse des angereicherten Boden-Konsortiums	98
2. Anreicherungskulturen als verlässliche Quelle neuer Biokatalysatoren....	104
2.1 Genduplikationen durch horizontalen Gentransfer.....	111
3. Charakterisierung einer Cellulase des angereicherten Boden-Konsortiums	114
4. Ausblick	116
V. Zusammenfassung.....	117
VI. Literaturverzeichnis	119
VII. Anhang	129

Abkürzungen

Abb.	Abbildung
Amp	Ampicillin
AS	Aminosäure(n)
BCIP	5-Brom-4-chlor-3-indolyl-phosphat
bidest	bidestiliert
bp	Basenpaare
BSA	Rinderserumalbumin
bzw.	beziehungsweise
°C	Grad Celsius
CBD	Cellulosebindedomäne
CMC	Carboxymethylcellulose
Da	Dalton
DMF	Dimethylformamid
DMSO	Dimethylsulfoxide
DNA	Desoxyribunukleinsäure
DNase	Desoxyribunuklease
dNTP	Desoxyribonuklosid-5'-triphosphat
ds	doppelsträngig, Doppelstrang
E.	<i>Escherichia</i>
EDTA	Ethylendiamintetraacetat
E-Cup	Eppendorfreaktionsgefäß
et al.	et alii; und andere
FPLC	Fast-Protein-Liquid-Chromatography
g	Gramm
GHF	Glycosidhydrolase-Familie
h	Stunde(n)
IPTG	Isopropyl-β-D-thiogalactopyranosid

kb	Kilobasen
l	Liter
LB	Luria-Bertani
Lsg.	Lösung
m	milli-($1 \cdot 10^{-3}$), Meter
M	Molar (mol/l), molekulare Masse
MKS	Multi-Klonierungs-Stelle
μ	Mikro-($1 \cdot 10^{-6}$)
max.	maximal
min	Minute
NBT	Nitroblau-Tetrazolin
OD	Optische Dichte
ORF	offener Leserahmen (engl. <i>open reaging frame</i>)
P.	<i>Pseudomonas</i>
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase Kettenreaktion
Pfu	<i>Pyrococcus furiosus</i>
PEG	Polyethylenglycol
pH	negativer dekadischer Logarithmus der Protonenkonzentration
r	resistent/Resistenz
RNA	Ribonukleinsäure
RT	Raumtemperatur
sec.	Sekunde
SDS	Natrium-Dodecylsulfat (engl. <i>Sodium-Dodecylsulfate</i>)
SG	Sammelgel
Std.	Stunden
SV	Säulenvolumen
T	Temperatur

Tab.	Tabelle
TAE	Tris-Acetat-EDTA-Puffer
TE	Tris-EDTA-Puffer
TEMED	N,N,N',N'-tetramethylethyldiamin
TG	Trenngel
Tris	Tris (hydroxymethyl)-aminoethan
U	Unit
Upm	Umdrehungen pro minute
UV	Ultraviolett
usw.	und so weiter
Vol.	Volumen
v/v	Volumen pro Volumen
w/v	Masse pro Volumen
X-Gal	5-Brom-4-chlor-3-indolyl- β -D-galactopyranosid

I. Einleitung

1. Das Potential der Metagenomik

Obwohl mit bloßem Auge nicht sichtbar, stellen Prokaryoten zwei der drei Domänen des Lebens und nehmen eine wichtige Stellung innerhalb des Ökosystems der Erde ein. Sie katalysieren einzigartige Transformationen im biogeochemischen Stoffkreislauf der Erde, produzieren wichtige Bestandteile der Erdatmosphäre und stellen einen nicht zu unterschätzenden Anteil der genetischen Diversität der Erde dar. Diese Diversität wird vor allem durch die Anpassungsfähigkeit an unterschiedliche Habitate bedingt. Prokaryoten kommen in Sedimenten ebenso vor, wie in aquatischer Umgebung, in gemäßigten Habitaten und unter teilweise extremen Bedingungen, die Temperatur, Druck, pH-Wert, Nährstoffgehalt und Salz- sowie Schwermetall-Konzentrationen betreffen können.

Vor nicht einmal 20 Jahren erstellte Woese mit seinen Mitarbeitern auf Grundlage von 16S rRNA-Genanalysen kultivierter Mikroorganismen den ersten bakteriellen Stammbaum mit insgesamt 12 Phyla (Woese, 1987). Seitdem hat sich die Anzahl an 16S rDNA-Sequenzen vervielfacht. Die NCBI Datenbank GenBank enthielt im Jahr 2003 ca. 78000 16S rDNA Sequenzen und 50 bakterielle Phyla, von denen die Hälfte aus bislang nicht kultivierten Organismen besteht (Schloss and Handelsman, 2004; Wheeler et al., 2000). Schätzungen gehen davon aus, daß die Anzahl der Prokaryoten ca. $4 - 6 \times 10^{30}$ Zellen beträgt, wobei der größte Anteil in den oberen Schichten des Erdbodens vorkommt (Whitman et al., 1998). Ein Gramm Erde enthält bis zu 4×10^7 Zellen und schätzungsweise 2000 bis 18000 unterschiedliche Spezies (Torsvik and Ovreas, 2002). Zur Zeit (Stand 10/2005) sind jedoch nur 263 komplette und 520 partielle bakterielle und archaelle Genome öffentlich zugänglich (Genom Datenbank NCBI). Hinzu kommt, daß mit derzeitigen Kultivierungsmöglichkeiten nur ca. 1 % aller Mikroorganismen zugänglich gemacht werden können (Torsvik and Ovreas, 2002).

Diese Diskrepanz zu verringern und sich das genetische Potential der bakteriellen Diversität anzueignen ist einer der Kernpunkte der Metagenomik bzw. Umweltgenomik. Der Begriff Metagenomik beschreibt die funktionale und sequenz-basierende Analyse der Gesamtheit der genetischen Information einer Umweltprobe (Handelsman et al., 1998). Eine metagenomische Analyse beinhaltet die direkte Isolation der DNA aus einer Umweltprobe, Klonierung in geeignete Vektoren mit anschließender Transformation in ein Wirtsbakterium und die Durchmusterung der erhaltenen Transformanten (Abbildung 1). Die Metagenom-Bank kann nach phylogenetischen

Markern durchmustert, aber auch zur Identifizierung neuer bioaktiver molekularer Substanzen und Biokatalysatoren genutzt werden.

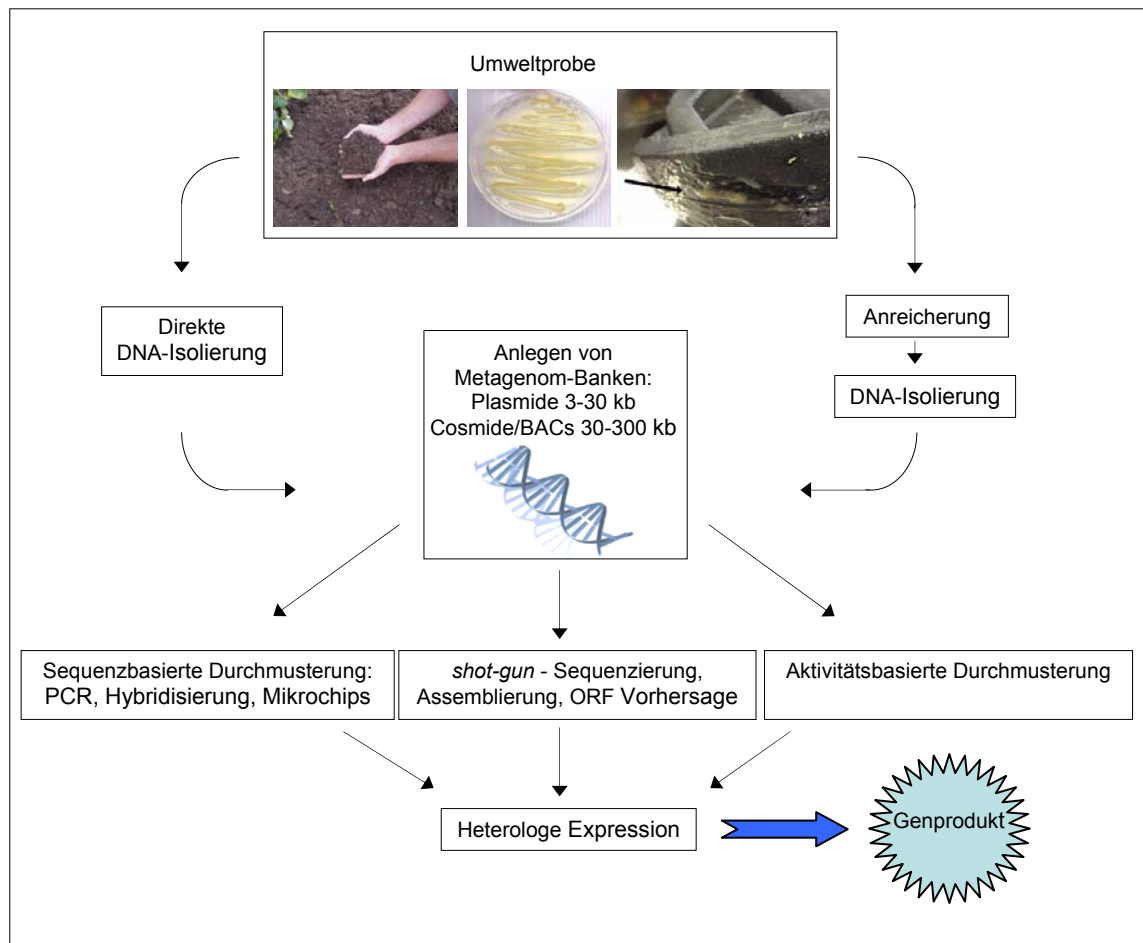


Abb. 1: Vorgehensweise der Metagenomik bei der Identifizierung von Biokatalysatoren aus Umweltproben.

2. Metagenomik und die weiße Biotechnologie

Unter weißer Biotechnologie, auch industrielle Biotechnologie genannt, versteht man den Einsatz biotechnischer Methoden innerhalb industrieller Prozesse. Der Begriff steht jedoch auch für den Gedanken einer neuen ökologischen und ökonomischen Nachhaltigkeit (Zinke, 2004). Der Vorteil gegenüber den etablierten chemischen Verfahren liegt auf der Hand: Biokatalyse funktioniert bei niedrigeren Temperaturen, produziert weniger giftige Abfälle, Emissionen und unerwünschte Nebenprodukte (Gavrilescu and Chisti, 2005). Obwohl das Potential von Enzymen zur Produktion von Fein-Chemikalien und pharmazeutischen Wirkstoffen immens ist, war es bisher durch die Anzahl geeigneter Biokatalysatoren limitiert (Burton et al., 2002). Die Erschließung der mikrobiellen Diversität durch die Fortschritte auf dem Gebiet der Metagenomik trug

zu einem Paradigmenwechsel bei: Anstatt die Prozesse auf die vorhandenen Enzyme anzupassen, ist es nun möglich durch geeignete Screening-Verfahren das passende Enzym für optimierte Prozesse zu finden (Bull et al., 2000; Cowan et al., 2004).

Möglich wurde dies vor allem durch die direkte Isolation von DNA aus einer Umweltprobe, was 1985 von Pace und seinen Mitarbeitern vorgestellt und mit der Erstellung einer λ Phagen Bank aus einer Seewasser Probe 1991 realisiert wurde (Pace et al., 1985; Schmidt et al., 1991). Die Entwicklung von Strategien zur Erstellung der ersten Umwelt-Banken aus dem Boden kam langsamer voran, da mit der DNA polyphenolischen Substanzen extrahiert werden, die schwer zu entfernen sind und Probleme bei der enzymatischen Modifikation der DNA, wie Restriktion, Ligation und PCR, bereiten (Rondon et al., 2000; Tsai and Olson, 1992).

Bei der Durchmusterung von Metagenom-Banken nach neuen Biokatalysatoren gibt es zwei klassische Ansätze (Abbildung 1). Zum einen die aktivitätsbasierende Durchmusterung, bei der z.B. durch Verwendung von Plattentests die Standort-Genbank nach der gewünschten Enzymaktivität durchsucht wird. Zum anderen die sequenzbasierende Durchmusterung, die das Design von degenerierten Primern auf der Grundlage konservierter DNA-Regionen zur Identifizierung neuartiger Gene durch direkte Amplifikation oder durch Kolonie-Hybridisierung beinhaltet (Bell et al., 2002). Mit sinkenden Kosten wird immer öfter auch die Sequenzierung großer Abschnitte, oder sogar ganzer Metagenom-Banken auf der Grundlage des „whole genome shotgun sequencing“ durchgeführt, einer Methode die sonst für die Sequenzierung von Einzelorganismen verwendet wird (Schmeisser et al., 2003; Venter et al., 2004). Ein Beispiel hierfür ist die Untersuchung eines mikrobiellen Konsortiums aus dem sauren Milieu eines Grubenabflusses, wo mittels „shot gun“-Sequenzierung der Gesamt-DNA zwei von fünf Genome der einzelnen Organismen komplettiert werden konnten (Tyson et al., 2004). Ein Beispiel für ein umfangreicheres Projekt stellt der Versuch dar, die mikrobielle Vielfalt der Sargasso See durch Sequenzierung zu erschließen, bei dem über 1,5 Gbp an DNA Information geschaffen wurden (Venter et al., 2004). Zum Vergleich: Bei der erstgenannten Studie wurden 76 Mbp an Sequenzinformation generiert, jedoch handelte es sich dabei auch um ein Habitat mit geringerer Diversität. Mit Hilfe dieser Methoden ist bereits eine große Anzahl neuer Biokatalysatoren und pharmazeutischer Wirkstoffe identifiziert worden. Hierzu zählen Esterasen/Lipasen (Ferrer et al., 2005; Rhee et al., 2005), Proteasen (Santosa, 2001), Amidasen (Gabor et al., 2004), Polyketidsynthasen (Ginolhac et al., 2004), und Antibiotika (Gillespie et al., 2002). Ebenso wurde eine große Anzahl Polysaccharid modifizierender Enzyme

isoliert, darunter Amylasen (Yun et al., 2004), Agarasen (Voget et al., 2003), Chitinasen (Hoster et al., 2005) und Pektat Lyasen (Solbak et al., 2005).

Ein Problem bei der erfolgreichen Identifizierung neuer Biokatalysatoren ist jedoch, neben dem Problem der Isolierung hochmolekularer reiner Gesamt-DNA, häufig die niedrige Frequenz an positiven Klonen in einer Metagenom-Bank (Daniel, 2005). Die größte Limitierung geht hierbei von der Expression in heterologen Wirtsstämmen, zumeist *E. coli*, aus. Es können die Transkription, Translation oder auch die Proteinfaltung der Fremd-Gene betroffen sein (Voget et al., 2005). Die Durchmusterung von mehreren Hunderttausend Klonen mit letzten Endes einem positivem Treffer ist keine Seltenheit (Henne et al., 2000; Riesenfeld et al., 2004). Um die Frequenz zu erhöhen bieten sich Anreicherungen vor der Erstellung der Genbanken an (Streit et al., 2004). Diese Vorgehensweise umgeht zudem die Problematik der DNA-Isolation, bedeutet aber auch einen Rückgang der Diversität in der untersuchten Umwelt-Probe, da schneller wachsende Organismen andere verdrängen. Mit Bedacht angewandt hat sich diese Methode jedoch als sehr erfolgreich bei der Isolierung neuartiger Biokatalysatoren erwiesen. Beispiele hierfür sind z.B. Amidasen (Gabor et al., 2004), Alkohol Oxidoreduktasen (Knietsch et al., 2003), Chitinasen (Hoster et al., 2005), B₁₂-abhängige Dehydratasen (Knietsch et al., 2003) sowie Biotinsynthese Operons (Entcheva et al., 2001).

Die meisten Anreicherungen basieren auf dem Einsatz selektiver Kohlenstoff- oder Stickstoffquellen, jedoch gibt es auch andere Ansätze. So können Mikroorganismen anhand der Verwertung von radioaktiv gelabelten Substraten identifiziert und die gelabelte DNA mittels Ultrazentrifugation selektiert werden (SIP, „Stable Isotope probing“) (Dumont and Murrell, 2005). Eine andere Methode ist die Anreicherung in Gegenwart von Bromodeoxy-Uridin (BrdU) unter Verwendung selektiver Kohlenstoffquellen (Yin et al., 2000). Die BrdU gelabelte DNA kann anschließend mit immunologischen Methoden separiert werden.

3. Cellulasen

Cellulose bildet mit 35 bis 50 % den Hauptbestandteil der Pflanzenzellwand und ist somit das am häufigsten vorkommende Biopolymer der Erde (Lynd et al., 2002). Aufgrund seines immensen Vorkommens und den damit verbundenen niedrigen Kosten bildet pflanzliche Biomasse einen geeigneten Rohstoff zur nachhaltigen Energiegewinnung (Demain et al., 2005). Die Herstellung von erneuerbaren Kraftstoffen (Bioethanol) gewinnt angesichts der steigenden CO₂-Emissionen und

schwindenden Erdölreserven an Aktualität und die Erforschung und Anwendung der Cellulose abbauenden Enzyme steht mehr und mehr im Interesse der Biotechnologie. Jedoch ist der Anwendungsbereich der cellulolytischen Enzyme damit nicht erschöpft. In der Nahrungsmittelindustrie finden sie Verwendung bei der Extraktion und Klärung von Fruchtsäften, der Herstellung von fermentierten Sojaprodukten wie Sojasauce und der Gewinnung von Agar aus Seegras. In der Textilindustrie entfernen Cellulasen den überschüssigen Farbstoff bei der Herstellung von Blue Jeans („stone wash“ Effekt) um nur einige Beispiele zu nennen (Bhat and Bhat, 1997).

Cellulose besteht aus linear β -1,4 verknüpften Glukose-Einheiten, wobei bis zu 15 000 Moleküle eine Cellulosekette bilden können. In dieser Kette ist ein Glukosemolekül um 180° zu seinem Nachbarmolekül gedreht und wird durch Wasserstoffbrückenbindungen in dieser Position gehalten. Ebenfalls mittels Wasserstoffbrückenbindungen sind die Glukosemoleküle eines solchen Stranges mit denen eines anderen verbunden. Mehrere solcher Mikrofibrillen bilden eine Cellulosefaser (Abbildung 2). Innerhalb der Struktur der Cellulose wechseln kristalline Bereiche mit nichtkristallinen, so genannten amorphen Bereichen ab. Der Grad der Kristallisation kann in der Natur zwischen 40 % und 90 % betragen (Hilden and Johansson, 2004).

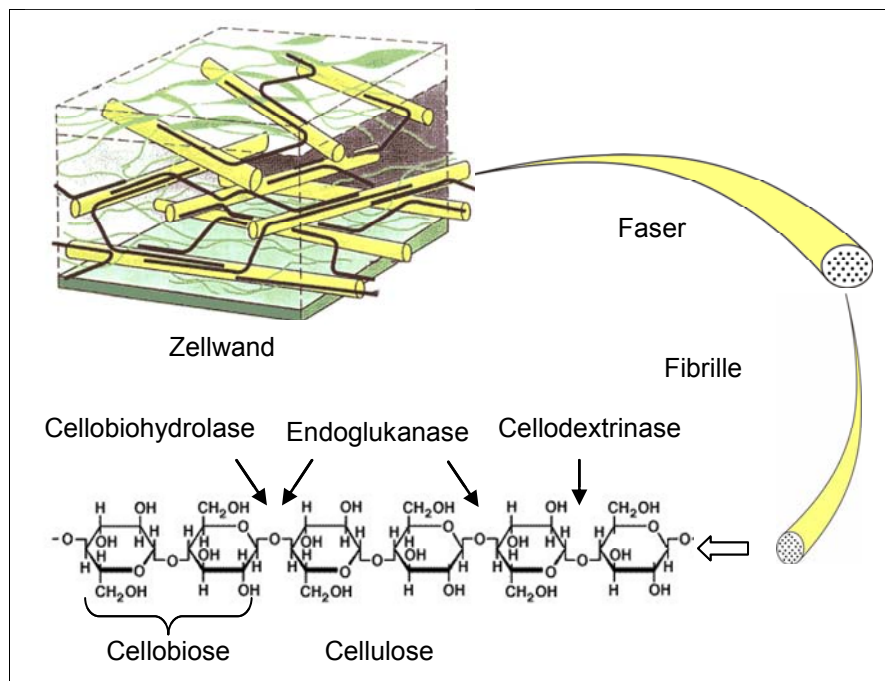


Abb. 2: Aufbau der Cellulosefaser und der pflanzlichen Zellwand.

Dargestellt ist die Primäre Zellwand mit den Bestandteilen Cellulose (gelb); Hemicellulose (schwarz) und Pektin (grün). Im Detail ist der Aufbau einer Cellulosefaser abgebildet. Sie besteht aus mehreren Cellulosefibrillen, die sich aus Celluloseketten zusammensetzen.

Aufgrund der komplexen Struktur der Cellulose, welche in der Zellwand von anderen Polymeren wie Hemicellulose und Pektin umgeben ist, entwickelten Cellulose-abbauende Organismen eine Anzahl von spezialisierten Enzymen. Cellulasen gehören zu der Familie der Glykosylhydrolasen (GHF) und hydrolisieren die β -1,4 glykosydischen Bindungen. Sie können in drei Gruppen eingeteilt werden: Endoglukanasen schneiden nach dem Zufallsprinzip innerhalb der amorphen Bereiche der Cellulosekette und setzen so Oligosaccharide unterschiedlicher Länge frei. Die Exoglukanasen hingegen wirken an den Enden der Cellulosekette und spalten Glukose (Cellodextrinasen) oder Cellobiose (Cellobiohydrolasen) ab. Die β -Glukosidasen hydrolisieren lösliche Cellodextrine und Cellobiose schließlich zu Glukose.

Bakterielle Cellulasen lassen sich in komplexe und nicht-komplexe Systeme unterteilen. Mit wenigen Ausnahmen verwerten anaerobe Bakterien Cellulose mittels komplexer Cellulasysteme, den so genannten Cellulosomen (Doi and Kosugi, 2004). Hierbei befinden sich die cellulolytischen Enzyme innerhalb eines nicht-katalytischen Proteins, welches für die Bindung an die Cellulose verantwortlich ist (Schwarz, 2001). Die von aeroben Bakterien produzierten Cellulasen befinden sich nicht in einem solchen Gerüst, sondern binden direkt an die Cellulose (Yi-Heng Percival Zhang, 2004). Diese Cellulasen besitzen meistens einen modularen Aufbau mit einer katalytischen Domäne, die über flexible Linker mit Cellulose-Binde-Domänen (CBD`s) verbunden sind. CBD`s vermitteln die Bindung an unlösliche Cellulose (Bolam et al., 1998). So zeigen denn auch die Cellulasen, die keine CBD`s besitzen nur eine eingeschränkte Fähigkeit, unlösliche Cellulose zu hydrolisieren (Coutinho et al., 1993; Fontes et al., 1997).

Viele der biotechnologisch interessanten Cellulasen stammen von Isolaten aus extremen Habitaten, deren Kultivierbarkeit aufgrund der speziellen Anforderungen schwierig ist (Gomes and Steiner, 2004). Die Umgehung dieser Problematik durch Methoden der Metagenomik eröffnet neue Möglichkeiten bei der Isolierung neuartiger Cellulasen von bisher unkultivierbaren Mikroorganismen (Healy et al., 1995; Rees et al., 2003; Voget et al., 2005).

Im Zuge der vorliegenden Arbeit erfolgte eine umfassende Analyse eines angereicherten Boden-Konsortiums. Mittels einer 16S rDNA-Sequenzanalyse sollte die Diversität des durch Anreicherung selektierten Konsortiums ermittelt werden. Anschließend sollte die erstellte Umwelt-Genbank mit aktivitäts- und sequenzbasierenden Techniken nach hydrolytischer Aktivität durchmustert, sowie eine aus dem Konsortium isolierte Cellulase einer ausführlichen biochemischen Charakterisierung unterzogen werden.

II. Material und Methoden

1. Verwendete Organismen, Plasmide und Oligonukleotide

Die im Rahmen dieser Arbeit verwendeten Organismen sind in Tabelle 1, die Plasmide in Tabelle 2 aufgeführt. Tabelle 3 gibt eine Übersicht über die verwendeten Oligonukleotide.

Tabelle 1: Verwendete Bakterienstämme

Bakterienstamm	Eigenschaften	Herkunft/Referenz
<i>Escherichia coli</i> XL1-blue	<i>thi1, endA1, gyrA96 (Nal^r), supE44, recA1, hsdR17, λ⁻, lac⁻, rpsL(Sm^r), [F⁺traD36, proAB, lacI^q, lacZΔM15, Tn10(Tc^r)]</i>	Bullock et al., 1987
<i>Escherichia coli</i> BL21 (DE3)	F ⁻ , <i>ompT, hsdS_B (r⁻_B m⁻_B) gal, dcm, λDE3</i>	Novagen
<i>Escherichia coli</i> VCS257	DP50 Derivat: <i>tonA53 dapD8 lacY1glnV44(supE44) (gal-uvrB)47 tyrT58 (supF58) gyrA29 (thyA57) hsdS3(rk-mk-) mcrA</i>	Sambrook et al., 1989

¹⁾ Geno- und Phänotypbezeichnung siehe (Bachmann, 1987)

Tabelle 2: Verwendete Vektoren und rekombinante Plasmide

Plasmid	Eigenschaften	Herkunft/Referenz
pBlueskript SK(+)	Amp ^r , <i>lacPOZ'</i>	Stratagene (Heidelberg)
pET19b	T7 Promotor, <i>lacI</i> Amp ^r	Stratagene (Heidelberg)
pTZ19	<i>lacZ, Cm^r</i>	MBI, Fermentas
pWE15	Amp ^r , Neo ^r , <i>cos</i>	Stratagene (Heidelberg)
SuperCos 1	Amp ^r ; Kan ^r ; <i>cos</i>	Stratagene (Heidelberg)
pDrive	<i>oriEc(pUC), P_{lac}/lacZ, Kan^r, Amp^r</i>	Qiagen (Hilden)
pSVCosCEL1-8	pWE15-Vektor mit Insert der Umwelt-DNA	diese Arbeit

pSVCEL-1	pSK(+) Vektor mit 9,7 kb <i>EcoRI</i> -Fragment aus pSVCosCEL1	diese Arbeit
----------	--	--------------

¹⁾ Phänotypbezeichnungen: Amp^r, Ampicillin-Resistenz; Kan^r, Kanamycin-Resistenz; Cm^r, Chloramphenicol-Resistenz, Neo^r, Neomycin-Resistenz;

Genotypbezeichnungen: *lacI*, Repressor des Lactose Operons; *lacPOZ'*, Promotor-Operator-Region des Lactose-Operons und α -Peptid der β -Galactosidase; *oriEc*, *origin of replication* in *E. coli*; *P_{lac}*, *lac* Promotor

Tabelle 3: Verwendete Oligonukleotide

Bezeichnung	Sequenz	Referenz
abi_forward	5'-ACGACGTTGTAAAACGACGGCCAG-3'	G ₂ L
abi_reverse	5'-TTCACACAGGAAACAGCTATGACC-3'	G ₂ L
T7-forward (pET)	5'-TAGGGGAATTGTGAGCGGAT-3'	diese Arbeit
T7-reverse (pET)	5'-ATCCGGATATAGTTCCTCC-3'	diese Arbeit
616V	5'-AGA GTT TGA TYM TGG CTC AG-3'	(Brosius et al., 1981)
1492R	5'-CGG YTA CCT TGT TAC GAC-3'	(Kane et al., 1993)
GM1F	5'-CCA GCA GCC GCG GTA AT-3'	(Muyzer et al., 1993)
celA-forward	5'- <u>CATATGAAATTTCCGCATTGCTTG</u> -3'	diese Arbeit
celA-reverse	5'- <u>GGATCCCTATTTCTTATTTTCACG</u> -3'	diese Arbeit
orf03-pSVCEL-1-forward	5'- <u>CATATGCGATTA</u> AAAATTCTATTG-3'	diese Arbeit
orf03-pSVCEL-1-reverse	5'- <u>GGATCCTTAGGGAATGGGGCAAGA</u> -3'	diese Arbeit
pelA-forward	5'- <u>GGCATATGAAATTCAGGACTATTTCCAAATGC</u> -3'	diese Arbeit
pelA-reverse	5'- <u>GGGGATCCTTACTGGCAAGACACCTTACCCAAA</u> -3'	diese Arbeit

Y= C oder T; M= A oder C; eingefügte Restriktionsstellen sind unterstrichen.

2. Nährmedien

Die angegebenen Mengen beziehen sich auf ein Endvolumen von 1 Liter in demineralisiertem Wasser. Zur Herstellung von Festmedien wurde jeweils 1,5 % (w/v) Agar hinzugegeben. Alle Lösungen und Medien wurden, sofern nicht anders angegeben, autoklaviert. Die Lagerung erfolgte bei 4 °C.

2.1 LB-(Luria-Bertani) Medium (Sambrook et al., 1989)

Trypton	10 g
Hefeextrakt	5 g
NaCl	10 g
H ₂ O _{dest.}	ad 1000 ml

2.2 TB-Medium (Sambrook et al., 1989)

Trypton	12 g
Hefeextrakt	24 g
Glycerin (87%, v/v)	4 ml
H ₂ O _{dest.}	ad 960 ml

5 x Phosphatpuffer

KH ₂ PO ₄	11,55 g
K ₂ HPO ₄	62,70 g
H ₂ O _{dest.}	ad 100 ml

Nach dem Autoklavieren wurden 20 ml 5 x Phosphatpuffer und 20 ml 10 %ige Glukoselösung zugegeben.

2.3 CMC (Carboxymethylcellulose)- Medium (Teather, 1982)

Dieses Medium diente der Durchmusterung der Genbank nach Klonen mit cellulytischer Aktivität.

Lösung A:

Trypton	10 g
Hefeextrakt	5 g

NaCl	10 g
H ₂ O _{dest.}	ad 500 ml

Lösung B:

CMC	x g
H ₂ O _{dest.}	ad 500 ml

Lösung B wurde in der Mikrowelle erhitzt bis die Carboxymethylcellulose vollständig gelöst war. Danach wurden beide Lösungen vereinigt und nach Hinzufügen von 1,5 % (w/v) Agar autoklaviert.

2.4 Mineralmedium

Die Bestandteile wurden als 100 x konzentrierte Stammlösungen angesetzt und getrennt autoklaviert. Nach dem Erkalten wurden die Lösungen zu sterilem H₂O_{dest.} (ad 1000 ml) gegeben. Zur Herstellung von Festmedium wurden Mineralmedium und Agar (1,5 % w/v) in H₂O_{dest.} jeweils doppelt konzentriert angesetzt und autoklaviert. Erst nachdem die Lösungen auf ca. 45 – 50 °C erkaltet waren, wurden sie vereinigt.

Zusammensetzung:

NaNO ₃ x 3 H ₂ O	1,0 g
K ₃ HPO ₄ x 3 H ₂ O	1,01 g
MgSO ₄ x 7 H ₂ O	0,25 g
CaCl ₂ x 7 H ₂ O	0,015 g
FeSO ₄ x 7 H ₂ O	0,14 g
EDTA	0,02 g
Spurenelementlösung	10 ml
H ₂ O _{dest.}	ad 1000 ml
keine pH-Korrektur.	

Spurenelementlösung (100 fach konzentriert):

Co(NO ₃) ₂ x 6 H ₂ O	15 mg
H ₃ BO ₄	61,8 mg

MnCl ₂ x 4 H ₂ O	198,0 mg
ZnSO ₄ x 7 H ₂ O	57,5 mg
NiCl ₂ x H ₂ O	47,5 mg
CuSO ₄ x 5 H ₂ O	2,5 mg
Na ₂ MoO ₄	2,4 mg
CaCl ₂ x 6 H ₂ O	4,0 mg
H ₂ O _{dest.}	ad 1000 ml

2.5 Antibiotika und sonstige Medienzusätze

Von den Medienzusätzen (Tab. 4) wurden Stammlösungen angesetzt, sterilfiltriert und bei -20°C gelagert. Die Zugabe der Antibiotika und anderer hitzelabiler Medienbestandteile erfolgte nach Autoklavieren und Abkühlung der Medien auf ca. 55°C.

Tabelle 4: Konzentrationen der verwendeten Antibiotika und anderer Medienzusätze

Medienzusatz	Lösungsmittel	Stammlösung (mg/ml)	Arbeitskonzentration (µg/ml)
Ampicillin	H ₂ O _{dest.}	100	100
Kanamycin	H ₂ O _{dest.}	20	20
Tetracyclin	Ethanol	10	10
Chloramphenicol	Ethanol	34	25
IPTG	H ₂ O _{dest.}	24	0,1-1mM
X-Gal	Dimethylformamid	20	0,002 % ig (v/v)

3. Wachstumsbedingungen und Stammhaltung

3.1 Zellanzucht

E. coli-Stämme wurden über Nacht in LB-Flüssigmedium in Reagenzgläsern (5 ml) oder Erlenmeyerkolben (30 ml bis 2 l Kulturen) angezogen. Plasmidhaltige Stämme

wurden durch Zugabe des entsprechenden Antibiotikums unter Selektivdruck gehalten. Hauptkulturen wurden mit 1-3 % (v/v) einer über Nacht gewachsenen Vorkultur angeimpft. Bei der Verwendung von Überexpressionsstämmen wurde die Expression der Proteine mit IPTG bei einer OD_{595} von 0,3 bis 0,6 induziert. Die Inkubation von *E. coli*-Kulturen erfolgte bei 30 °C bzw. 37 °C auf einem Schüttler.

3.2 Stammkulturen

E. coli konnte bis zu vier Wochen auf festem Komplexmedium mit dem entsprechenden Antibiotikum bei 4 °C gelagert werden. Für eine längere Lagerung wurden Glycerinkulturen angelegt. Frische Übernachtskulturen wurden mit sterilem Glycerin im Verhältnis 1:1 gemischt und bei -70 °C eingefroren.

3.3 Messung der optischen Dichte

Die optische Dichte wurde in Einmalküvetten mit einer Schichtdicke von 1 cm bei einer Wellenlänge von 595 nm gegen unbeimpftes Medium photometrisch bestimmt (Pharmacia Biotec; Ultraspec 3000). Um Meßungenauigkeiten zu vermeiden, wurden die Proben mit unbeimpftem Medium entsprechend verdünnt, sobald eine Extinktion von 0,3 überschritten wurde.

4. Standardtechniken für das Arbeiten mit Nukleinsäuren

4.1 Behandlung von Geräten und Lösungen für das Arbeiten mit Nukleinsäuren

Thermostabile Lösungen, Glaswaren und andere Geräte wurden zur Inaktivierung von Nukleasen für 30 min bei 121 °C autoklaviert. Hitzelabile Lösungen wurden sterilfiltriert. Nicht autoklavierbare Materialien wurden mit 70 % (v/v) Ethanol gespült oder abgeflammt.

4.2 Reinigung und Konzentrierung von Nukleinsäuren

4.2.1 Phenol/Chloroform-Extraktion

Zur Extraktion von Proteinen wurden die DNA-Lösungen im Verhältnis 1:1 mit einem Phenol/Chloroform/Isoamylalkohol-Gemisch (25:24:1 (v/v/v)) versetzt und gründlich durchmischt. Zur Phasentrennung wurden die Ansätze 20 min bei 13000 Upm und 4 °C zentrifugiert und die wässrige, DNA-haltige Oberphase vorsichtig

abgehoben. Zur Entfernung von Phenolresten wurde noch eine 1:1 Extraktion mit Chloroform/Isoamylalkohol (24:1 (v/v)) durchgeführt. Anschließend erfolgte eine Fällung der DNA mit Ethanol.

4.2.2 Fällung von DNA mit Ethanol oder Isopropanol

Zum Fällen von DNA wurde die DNA-Lösung mit 0,3 Volumen einer 7,5 M Ammoniumacetatlösung und 0,7 Volumen Isopropanol oder 2,5 Volumen eiskaltem unvergälltem Ethanol (96 % (v/v)) versetzt. Nach Durchmischung (bei Ethanol-Fällung Lagerung bei $-20\text{ }^{\circ}\text{C}$ für 0,5-12 h) konnte die DNA durch Zentrifugation (13000 Upm, $4\text{ }^{\circ}\text{C}$, 15-30 min) gefällt werden. Das Pellet wurde zweimal mit unvergälltem 70 %igen Ethanol gewaschen und dann in der Vakuumzentrifuge (Speed Vac[®]Plus SC110A Savant) für 5-10 min getrocknet. Die DNA wurde nun im gewünschten Volumen H₂O oder TE-Puffer resuspendiert.

TE-Puffer

Tris-Cl: 10 mM
EDTA: 1 mM
pH 7,4

4.2.3 Dialyse von Gesamt-DNA

Aus Umweltproben isolierte DNA (s. 2.4.4) enthält oftmals zu hohe Salzkonzentrationen, die bei weiteren Anwendungen stören könnten. Sie muß daher zunächst einer Dialyse unterzogen werden. Dazu wurden Dialyseschlauch-Stücke in TE-Puffer (s. 2.4.2.2) autoklaviert. Anschließend konnte die DNA-Lösung in die Schlauch-Stücke eingefüllt werden, die an beiden Enden mit Dialyseschlauch-Klammern verschlossen wurden. Die Dialyse der DNA erfolgte bei $4\text{ }^{\circ}\text{C}$ für 24 h gegen TE-Puffer, der nach etwa 4-5 h ausgewechselt wurde. Anschließend erfolgte die Fällung der DNA mit Ethanol (s. 2.4.2.2).

4.2.4 Konzentrationsbestimmung von DNA

Zur Ermittlung der Konzentration wurde die OD₂₆₀ einer DNA-Lösung in einer Quarzküvette gemessen. Bei dieser Wellenlänge entspricht eine OD₂₆₀ von 1 einer DNA-Konzentration von 50 µg/ml. Die Reinheit konnte durch die Bildung der Quotienten von Absorptionswerten bei verschiedenen Wellenlängen ermittelt werden. Für reine DNA-Lösungen gelten folgende Richtwerte: OD₂₆₀ : OD₂₈₀ $\geq 1,8$ (Sambrook et al., 1989), niedrigere Werte weisen auf Verunreinigungen mit Proteinen hin.

4.3 Methoden zur Charakterisierung von DNA-Fragmenten

4.3.1 Agarosegelelektrophorese und Ethidiumbromid-Färbung von DNA

Eine der am häufigsten angewandten Methoden zur Charakterisierung von DNA ist die Agarosegelelektrophorese. Mit Hilfe dieser Technik können DNA-Moleküle im Gel aufgetrennt werden. Dabei hängt die Laufgeschwindigkeit eines DNA-Fragments von mehreren Faktoren ab, wie z.B. der Größe des DNA-Moleküles, der Agarosekonzentration und der davon abhängigen Porengröße, der angelegten Spannung und den Salzkonzentrationen des Puffers. Es wurden 0,8 %ige Gele für die Analyse von Restriktionsfragmenten und 1 %ige oder 1,2 %ige Gele zur Auftrennung von PCR-Produkten in 1 x TAE-Puffer, der auch als Laufpuffer verwendet wurde, gegossen. Vor Beladen der Geltaschen wurde die DNA mit Ladepuffer im Verhältnis 10:1 vermischt, was der Beschwerung und der Markierung der Lauffront, durch den enthaltenen Farbstoff Bromphenolblau, bei ca. 500 bp diente. Die Auftrennung erfolgte im Minigel (Hoefer® HE33 Mini Horizontal Submarine Unit) bei einer Spannung von 60-100 V und einer Laufzeit von 45-90 min. Nach Beendigung der Gelelektrophorese erfolgte die Färbung der DNA durch Überführung des Gels in ein Ethidiumbromidbad für ca. 10 min und anschließender Entfärbung im Wasserbad für ebenfalls 10 min. Mit Hilfe der Gel-Dokumentationsanlage (BioRad® GelDoc 1000) wurde die DNA bei UV-Licht durch das eingelagerte Ethidiumbromid sichtbar gemacht, fotografisch dokumentiert und eventuell bearbeitet. Mittels eines Thermodruckers konnten diese Bilder ausgedruckt werden.

50 x TAE-Puffer (pH 8,5)

Tris	242 g
Eisessig	57,1 ml
Na ₂ EDTA x 2 H ₂ O	32,1 g
H ₂ O	ad 1000 ml

10 x DNA-Ladepuffer

Ficoll 400	20 g
Bromphenolblau	0,25 g
SDS	1,0 g

EDTA	0,1 M, pH 8
H ₂ O	ad 100 ml

Ethidiumbromid-Färbebad:

Zwei Tropfen einer 2 mg/ml Ethidiumbromidlösung wurden mit 500 ml H₂O vermischt.

4.4 Isolierung von DNA

4.4.1 Isolierung von Gesamt-DNA des bakteriellen Konsortiums

Kulturen wurden bei 4 °C und 5000 Upm pelletiert und in 1-2 ml 1M NaCl resuspendiert und eine Stunde auf Eis inkubiert. Nach erneuter Zentrifugation (20 min bei 13.000 Upm und 4 °C) wurde das Pellet in 1,5 ml TE-Saccharose-Lösung (20 %, w/v) aufgenommen und die Zellen in 1,2 ml DEP mit 1 % SDS für 2-4 h bei 37 °C lysiert. Anschließend wurde der DNA-Extrakt mit 600 µl Sarkosyl-Proteinase K-Lösung versetzt und über Nacht bei 37 °C inkubiert. Die Gesamt-DNA wurde mit einer Phenol/Chloroform-Lösung (1:1, v/v) extrahiert. Diese Extraktion wurde 2-3 mal wiederholt, bis eine klare, viskose DNA-Lösung in der wässrigen Phase vorhanden war. Anschließend wurde eine Chloroform-Extraktion durchgeführt und die DNA gegen 2 l TE-Puffer bei 4 °C über Nacht dialysiert.

DNA-Extraktionspuffer (DEP)

Tris	100 mM
Na ₂ -EDTA	100 mM
Na ₂ -HPO ₄	100 mM
NaCl	1,5 M
CTAB	1 %
pH 8	

TE-Saccharose-Puffer

20 % Saccharose in TE-Puffer

Sarkosyl-Proteinase K-Lösung

25 µl einer 25 %igen Proteinase K-Lösung (w/v) wurden auf 1 ml 5 %ige Sarkosyl-Lösung gegeben.

4.4.2 Isolierung von Plasmid-DNA durch alkalische Lyse

Diese Methode wurde eingesetzt, um eine große Anzahl von Klonen/Transformanten auf das Vorhandensein positiver rekombinanter Plasmide zu überprüfen.

Durchführung:

1,5 ml einer Übernachtskultur wurden 1 min im Eppendorfgefäß bei 13.000 Upm abzentrifugiert. Das Pellet wurde in 200 µl Puffer P1 resuspendiert. Anschließend erfolgte die alkalische Lyse der Zellen durch Zugabe von 200 µl Puffer P2, dreimaligem Invertieren des Eppendorfgefäßes und 5 minütiger Inkubation auf Eis. Durch Zugabe von 200 µl Puffer P3 erfolgte die Fällung der Proteine (2-3 sec vortexen und 5 min auf Eis inkubieren). Nach 5 minütigem Zentrifugieren bei 13.000 Upm wurde der Überstand in ein neues Eppendorfgefäß überführt. Bei Bedarf an sehr reiner Plasmid-DNA wurde eine Phenol/Chloroform Extraktion zur Verminderung von Proteinkontaminationen durchgeführt; dazu wurde jeder Ansatz mit 200 µl eines Phenol/Chloroform-Gemisches versetzt, gevortext und nach 5 minütigem Zentrifugieren der Überstand in ein neues Gefäß überführt. Durch Zugabe von 1 ml Isopropanol und 5 minütiger Inkubation auf Eis wurde die DNA gefällt und nach 30 minütigem Zentrifugieren bei 13.000 Upm zweimal mit 1 ml 70 %igem Ethanol gewaschen. Das Pellet wurde bei 37 °C getrocknet und in 50-100 µl H₂O_{dest.} resuspendiert.

Puffer P1

Tris-HCl 50 mM

EDTA 10 mM

pH 8

RNA konnte durch Zusatz von RNaseA (1 mg/ml) zu Puffer 1 entfernt werden.

Puffer P2

NaOH 0,2 M

SDS 1 % (w/v)

Puffer P3

Eisessig 29,5 ml

KOH-Plätzchen ad pH 4,8

H₂O_{dest.} ad 100 ml

4.4.3 Isolierung von Plasmid-DNA mit Hilfe des „QIAprep Spin Miniprep“

Das Plasmidisolierungskit von Qiagen (Hilden) ermöglichte die Präparation kleinerer Mengen hochreiner DNA. Dabei wurden bis zu 25 µg Plasmid-DNA (bei hoher Kopienzahl) aus 3-5 ml *E. coli*-Kulturen gewonnen. Das Prinzip beruht auf alkalischer Lyse der Zellen mit nachfolgender Adsorption der DNA an eine Silicagelmatrix in Gegenwart hoher Salzkonzentrationen. Die in H₂O_{bidest.} eluierte DNA war für die PCR und Sequenzierungen verwendbar.

4.4.4 Isolierung von Plasmid-DNA mit Hilfe des „QIAprep Spin Midiprep“

Das Plasmidisolierungskit der Firma Qiagen (Hilden) ermöglicht die Aufreinigung größerer Plasmidmengen. Abhängig von der Kopienzahl wurde eine 25-500 ml Kultur über Nacht bei 37 °C inkubiert. Nach der Zellernte wurde das Pellet in einem vom Hersteller vorgeschriebenen Volumen P1-Puffer gelöst. Nach der Zugabe der gleichen Menge P2-Puffer durfte der Ansatz nicht länger als 5 min bei RT inkubiert werden. Anschließend wurde P3-Puffer hinzugegeben und das Gemisch für 30 min auf Eis inkubiert. Nach Zentrifugation (4 °C, 30 min, 13.000 Upm) wurde der Überstand abgenommen und zur vollständigen Entfernung der Proteine erneut für 15 min zentrifugiert. Es folgte eine Reinigung über eine Qiagen Tip 100-Säule (Anionenaustauschchromatographie). Der Überstand aus der letzten Zentrifugation wurde auf die mit Niedrigsalzpuffer (QBT) äquilibrierte Säule aufgetragen und nach einem Waschschrift (QC-Puffer) mit salzhaltigem Tris-Puffer (QF) unter schwach alkalischen Bedingungen eluiert. Die DNA wurde mit 0,7 Vol. Isopropanol gefällt, gewaschen (70 % w/v Ethanol), getrocknet und in TE-Puffer oder H₂O_{bidest.} aufgenommen.

4.4.5 Elution von DNA-Fragmenten aus Agarosegelen

Nach PCR oder Restriktionsverdau und Überprüfung der Ergebnisse auf einem Agarosegel konnten DNA-Fragmente, die für eine Ligation benötigt wurden, unter Verwendung des „QiaQuick Gel Extraktion Kits“ (Qiagen, Hilden) aus dem Gel eluiert werden.

4.5 Enzymatische Modifikationen von DNA

4.5.1 DNA-Verdau durch Restriktionsendonukleasen

Diese Methode beruht auf der hohen Spezifität der aus verschiedenen Prokaryonten isolierten Restriktionsendonukleasen (TypII) zu ihrer (meist) palindromischen Erkennungssequenz in einem DNA-Doppelstrang. Nach Bindung an

diese palindromische Sequenz hydrolysieren die Enzyme an definierter Stelle die Phosphodiesterbindung zwischen je zwei Basen in beiden Strängen, wobei überhängende oder glatte Enden entstehen können. Es wurden überwiegend Restriktionsenzyme und die mitgelieferten 10fachen Inkubationspuffer der Firma MBI Fermentas, verwendet. In der Regel hatten die Restriktionssansätze ein Gesamtvolumen von 20 µl.

Restriktionsansatz

DNA	x µg
10 x Spaltungspuffer	2 µl
Restriktionsenzym	0,5 µl (1-5 U/µg DNA)
H ₂ O	ad 20 µl

Die Inkubation erfolgte 2-3 h, oder auch über Nacht, bei der für das Restriktionsenzym optimalen Temperatur von meist 37 °C. Zur Inaktivierung der Restriktionsenzyme wurden die Ansätze 10 min bei 75 °C inkubiert oder mit Chloroform extrahiert.

4.5.2 Partieller Verdau von Gesamt-DNA

Der partielle Verdau von Gesamt-DNA aus Umweltproben wurde in einer Verdünnungsreihe durch das Enzym *Bsp143I* (ein Isoschizomer von *Sau3AI*) durchgeführt. Es wurden 10 U Enzym 8-10 fach verdünnt. Die eingesetzte DNA-Menge sollte nicht mehr als 250 ng DNA betragen. Nach 20 min Inkubation bei 37 °C wurde der Verdau gestoppt und auf einem 0,8 %igem Agarosegel überprüft.

4.5.3 Dephosphorylierung von DNA-Enden

Zur Vermeidung von Selbstligationen konnten die 5'-Phosphatgruppen restriktionsverdauter DNA enzymatisch durch alkalische Phosphatase abgespalten werden.

Dieses fand mittels der „Antarctic Phosphatase“ (Fa. NEB) statt. Es wurden 1/10 Volumen des 10x Puffers zu 1 µg DNA gegeben. Nach hinzufügen von 1 µl Antarctic Phosphatase wurde der Ansatz gemixt. Handelte es sich um 5'-Überhänge oder gerade Enden wurde für 15 min bei 37 °C inkubiert, bei 3'-Überhängen für 60 min. Die Inaktivierung erfolgte durch fünf minütiges Erhitzen bei 65 °C.

4.5.4 Ligation von DNA-Fragmenten

DNA-Ligasen katalysieren die Bildung von Phosphodiester-Bindungen zwischen einem 5'-Phosphat- und einem 3'-OH-Ende in doppelsträngiger DNA. Mit Hilfe der T4-DNA-Ligase (Promega, Deutschland GmbH, Mannheim) wurden DNA-Fragmente ligiert. Die Ligationen wurden in möglichst wenig Volumen durchgeführt, mit einem Überschuß an Insert-DNA. Ligationen von DNA-Fragmenten mit kohäsiven Enden erfolgten bei 16 °C und solche mit glatten Enden (Zugabe von 0,1 Vol. 50 %igem (v/v) Polyethylenglykols (PEG) zur Erhöhung der Viskosität) bei Raumtemperatur, jeweils über Nacht. Die Ligationsansätze konnten ohne weitere Behandlung zur Transformation eingesetzt werden.

4.6 Amplifikation von DNA durch Polymerase-Kettenreaktion (PCR)

4.6.1 Amplifikation von DNA-Fragmenten mit der *Taq*- oder *Pfu*-DNA-Polymerase

Die Polymerase-Kettenreaktion (PCR) wurde zur Herstellung spezifischer DNA-Fragmente für Klonierungen oder Sequenzierungen verwendet. Als Enzyme wurden die *Taq*-DNA-Polymerase (Fa. MBI Fermentas) und die *Pfu*-DNA-Polymerase (Fa. MBI Fermentas) eingesetzt. Die Reaktionen wurden entsprechend den Herstellerempfehlungen in einem Volumen von 50 bzw. 100 µl durchgeführt und enthielten folgende Bestandteile:

DNA	0,5 - 1	µg
Oligonukleotide	je 100	pmol
dNTP-Gemisch	200	µM
DNA Polymerase	1	U
Reaktionspuffer (10x)	10	µl
MgCl ₂ oder MgSO ₄	1,5	mM
H ₂ O _{dest.}	ad 100	µl

Die Temperaturzyklen wurden den jeweiligen Anforderungen in Bezug auf die Hybridisierungstemperatur (x) und die Dauer der Kettenverlängerung (y) angepasst. Sie wurden nach folgendem Schema durchgeführt:

Denaturierung	95 °C	2 min
30 Zyklen		
Denaturierung	95 °C	1 min

Hybridisierung	x °C	1 min
Kettenverlängerung	72 °C	y min
Ende der Zyklen		
Kettenverlängerung	72 °C	10 min

Für die Einstellung der Parameter waren die zu erwartende Fragmentlänge und die Schmelztemperatur T_m der verwendeten Oligonukleotide entscheidend. Für die Elongation wurde entsprechend der zu erwartenden Fragmentlänge eine Zeit von ca. 1 min/1000 Bp gewählt. Die optimale Hybridisierungstemperatur wurde nach der Schmelztemperatur T_m der Oligonukleotide festgelegt, die über folgende Formel errechnet wurde:

$$T_m = 69,3 + 0,41 (\% \text{ GC-Gehalt}) - 650/L$$

L = Länge des Oligonukleotids; % GC-Gehalt = prozentualer GC-Gehalt des Oligonukleotids

Durch Verkürzung der Hybridisierungsdauer, Verringerung der Oligonukleotidkonzentration und Variation der $MgCl_2$ -Konzentration konnte die Spezifität der PCR optimiert werden.

Bei der 16S rRNA-Gen spezifischen PCR wurden stark verlängerte Zeiten (ca. 3 x) für die Kettenverlängerung verwendet, um die Entstehung von chimären Sequenzen zu minimieren (Meyerhans et al., 1990).

Die Temperaturzyklen wurden mit dem "Thermocycler Primus" (Fa. MWG Biotech, Ebersberg) durchgeführt. Die PCR-Produkte wurden durch Agarose-Gelelektrophorese (2.4.3.1) analysiert.

4.6.2 Amplifikation von DNA-Fragmenten mit der KOD Hifi DNA Polymerase

Diese Polymerase (Fa. Novagen) hat eine „Proofreading“-Aktivität und extrem kurze Hybridisierungs- und Verlängerungszeiten. Dadurch wichen das Pipettierschema und die Temperaturzyklen extrem von den der anderen verwendeten Polymerasen ab.

DNA	0,5 - 1	µg
Oligonukleotide	0,4	µM
dNTP-Gemisch	200	µM

DNA Polymerase	1	U
Reaktionspuffer (10x)	5	μ l
MgCl ₂	1	mM
H ₂ O _{dest.}	ad 50	μ l

Fragmentgröße 1-2 kb

Denaturierung	98 °C	5 min
Zyklen	25	
Denaturierung	98 °C	15 s
Hybridisierung	T _m - 5 °C	2 s
Kettenverlängerung	72	20 s

Fragmentgröße 3-4 kb

Denaturierung	98 °C	5 min
Zyklen	25 - 30	
Denaturierung	98 °C	15 s
Hybridisierung	T _m - 5 °C	5 s
Kettenverlängerung	72	40 s

4.7 Ligation von PCR Produkten in den pDrive Klonierungs-Vektor

Der pDrive Klonierungs Vektor (Qiagen, Hilden) wird in einer linearen Form geliefert und enthält an jedem Ende einen U-Überhang, der eine spezifische Bindung mit dem PCR-Produkt möglich macht. Ebenfalls besitzt der Vektor einen T7 und einen SP6 Promotor an jeder Seite der Klonierungsstelle, so dass eine *in vitro* Transkription des PCR-Produktes möglich wird. Für die PCR wurde die *Taq*-Polymerase verwendet, da diese A-Überhänge an den Enden des PCR-Produkts erzeugt, die für eine anschließende Ligation in den pDrive Vektor benötigt werden.

4.8 Transposonmutagenese

Um funktionelle Gene zu identifizieren, wurde das EZ::TN™ <KAN-2> Transposon der Firma Epicentre (Madison, Wisconsin) verwendet. Das Transposon

insetierte zufällig in die Ziel-DNA (*in vitro*). Mittels einer Kanamycinresistenz, die durch das Transposon vermittelt wurde, konnte die mutagenisierte DNA selektioniert werden. Die Mutagenese erfolgte nach Angaben des Herstellers. Das Transposon enthielt ebenfalls Primerbindestellen für die Primer KAN-2 FP-1 *Forward* und KAN-2 RP-1 *Reverse* (Tab. 6. Diese Primer ermöglichten eine bidirektionale Sequenzierung des DNA-Abschnitts, in den das Transposon inseriert hatte.

Tabelle 6: EZ::TN™ <KAN-2> Transposon Primer für eine bidirektionale Sequenzierung

Primer	Sequenz	Tm [°C]
KAN-2 FP-1 <i>Forward</i> Primer	5' - ACCTACAACAAAGCTCTCATCAACC - 3'	55
KAN-2 RP-1 <i>Reverse</i> Primer	5' - GCAATGTAACATCAGAGATTTTGAG - 3'	55

5. Transformationsverfahren

5.1 Hocheffiziente Transformation von *E. coli*

5.1.1 Herstellung von kompetenten Zellen mit Hilfe der CaCl₂ Methode

200 ml LB-Medium wurden 1 %ig (v/v) mit einer Übernachtskultur angeimpft und bis zu einer OD₅₉₅ von ca. 0,4 angezogen. Nach 10 min Inkubation auf Eis erfolgte die Pelletierung der Zellen bei 4 °C und 4.000 Upm für 7 min. Das Pellet wurde in 10 ml eiskalter CaCl₂ -Lösung resuspendiert und wurde anschließend bei 4 000 Upm für 5 min bei 4 °C abzentrifugiert. Die Pellets wurden in 2 ml eiskalter CaCl₂ Lösung resuspendiert und in 200 µl Aliquots in vorgekühlte Eppendorfgefäße portioniert. Die Zellen konnten direkt für eine Transformation verwendet werden oder wurden sofort bei -70 °C schockgefroren.

Calciumchloridlösung

CaCl ₂	60 mM
Glycerol	15 % (w/v)
PIPES	10 mM
pH 7	

5.1.2 Transformation von kompetenten Zellen

Eingefrorene kompetente Zellen (200µl) wurden 10 min auf Eis aufgetaut. Nach Zugabe von 10-150 ng DNA und vorsichtigem Mischen erfolgte eine Adsorption der DNA an die Zellen durch 30 minütige Inkubation auf Eis. Durch einen Hitzeschock für 90 sec bei 42 °C sollte die DNA in die Zellen aufgenommen werden. Anschließend wurden die Zellen für 3-5 min auf Eis inkubiert. Nach Zugabe von 800 µl LB-Medium wurden die Zellen zur Ausbildung der Antibiotikaresistenz 45-60 min bei 37 °C inkubiert. 100 µl des Transformationsansatzes wurden auf vorgewärmte Agarplatten mit entsprechendem Antibiotikazusatz zur Selektion plattiert und über Nacht bei 37 °C inkubiert.

5.1.3 Selektion rekombinanter *E. coli* Klone durch X-Gal-Test

Der verwendete *E. coli* Stamm XL1-Blue ist durch eine Deletion im *lacZ*-Gen gekennzeichnet und kann daher keine aktive β-Galaktosidase mehr bilden. Durch Transformation mit Plasmiden wie pBlueskript SK+, die über die α- Untereinheit der β-Galaktosidase verfügen, in der sich wiederum die Multiklonierungsstelle befindet, kommt es zur Bildung eines aktiven Enzyms. Dies kann dadurch auf X-Gal- und IPTG-haltigen Platten sichtbar gemacht werden, dass das *lacZ*-Gen durch IPTG induziert wird und die β-Galaktosidase das Glucoseanalogon X-Gal spaltet, wonach durch Luftoxidation ein blauer Farbstoff entsteht. Soll in die Multiklonierungsstelle des Plasmids ein *Insert* ligiert werden, kann der Erfolg der Insertion nach Transformation anhand der Kolonienfarbe sichtbar gemacht werden. Plasmide mit einer Insertion in der Multiklonierungsstelle können keine aktive β-Galaktosidase bilden und erscheinen weiß (Blau/Weiß-Selektion).

5.2 Elektroporation

Durch kurze elektrische Impulse wurde die Zellmembran für DNA passierbar gemacht. Die Transformationseffizienz bei dieser Methode lag bei ca. 10^9 - 10^{10} Transformanten pro µg eingesetzter DNA. Entscheidende Parameter sind die Stärke des elektrischen Feldes, die Dauer des Pulses und die Konzentration an DNA. Dabei mußte ein Kompromiss zwischen der Erhöhung der Transformationseffizienz und dem Absinken der Lebensfähigkeit der Zellen durch höhere Feldstärken und längere Pulszeiten gefunden werden.

5.2.1 Herstellung „kompetenter Zellen“ für die Elektroporation

Alle Lösungen und Gefäße mußten eisgekühlt verwendet werden. Vornehmliche Aufgabe der Zellpräparation war die Beseitigung von Salzen, um die

lonenstärke der Suspension zu verringern. Die Zellpellets sollten nur äußerst vorsichtig resuspendiert werden.

500 ml LB-Medium wurden mit 2,5 ml Übernachtskultur angeimpft und bis zu einer OD_{595} von 0,5-0,7 unter starkem Schütteln bei 37 °C inkubiert. Anschließend wurden die Zellen 10–15 min auf Eis gelagert. Die Ernte der Zellen erfolgte durch Zentrifugation bei 5.000 Upm für 20 min bei 4 °C. Das Pellet wurde in der verbleibenden Flüssigkeit resuspendiert. Die Zellen wurden mit 500 ml sterilem $H_2O_{bidest.}$ gewaschen und erneut bei 5.000 Upm für 20 min bei 4 °C abzentrifugiert. Das Pellet wurde mit 40 ml 10 % Glycerin gut gemischt und abermals abzentrifugiert (4.000 Upm, 10 min, 4 °C). In ca. 1 Volumen 10 %igem Glycerin wurden die Zellen resuspendiert und aliquotiert (60µl). Eine Lagerung erfolgte bei -70 °C.

5.2.2 Transformation mittels Elektroporation

In einer Elektroporationsküvette wurden 0,1-1 µg DNA mit 60 µl Zellsuspension gemischt und auf Eis inkubiert. Die Elektroporation erfolgte bei 2.500 V, 25 µF und 200 Ω. Die Entladungszeit sollte zwischen 3 und 5 ms liegen. Die Kontakte sollten trocken sein. Nach der Elektroporation erfolgte die Zugabe von 1 ml LB-Medium und sofortiges Umfüllen des Ansatzes in E-Cups. Die Zellen wurden für ca. 1 h bei 37 °C im Schüttler inkubiert und anschließend 25 µl auf Selektionsmedium ausplattiert. Eine Inkubation erfolgte bei 37 °C über Nacht.

5.4 Transduktion mit dem “Gigapack® III Gold Packaging Extract”

Um Genbanken mit relativ großen Fragmenten (~30 kb) herzustellen, wurden Cosmidvektoren verwendet. Diese konnten, falls sie dementsprechend große Inserts trugen, *in vitro* in λ-Phagen Köpfe verpackt und in dafür geeignete *E. coli*- Stämme transduziert werden.

Zunächst wurden die vom Hersteller (Fa. Stratagene, San Diego, USA) gelieferten, bei -70 °C gelagerten Lysate durch Reiben zwischen den Fingern aufgetaut. Dann wurden sofort 0,1 - 1,0 µg (max. 4 µl) ligierte DNA hinzugefügt und vorsichtig mit der Pipettenspitze gemischt. Anschließend wurde das Gemisch kurz (3 - 5 s) zentrifugiert und für 2 h bei RT inkubiert. Anschließend wurden 500 µl SM-Puffer und 20 µl Chloroform zugefügt, vorsichtig gemischt und zum Sedimentieren kurz (5 - 10 s) zentrifugiert. Die in Phagenköpfe verpackte DNA befand sich nun im Überstand und konnte bei 4 °C bis zu 4 Wochen aufgehoben werden.

Für die Infektion mit den durch den “Gigapack® III Gold Packaging Extract” (Fa. Stratagene) *in vitro* verpackten Ligationsansätzen wurde *E. coli* VCS257 eingesetzt.

Zur Ausbildung des für die Transduktion notwendigen λ -Rezeptors wurden die Stämme in LB-Medium (s. 2.2.1) mit 0,2 % (w/v) Maltose und 10 mM MgSO_4 für 6 bis 8 h bei 37 °C oder für 12 h bei 30 °C angezogen. Das Pellet wurde in MgSO_4 -Lösung (10 mM) suspendiert und auf eine OD_{600} von 0,5 - 1,0 eingestellt. Diese Suspension wurde bei 4 °C gelagert und noch am gleichen Tag zur Transduktion eingesetzt.

Vom Verpackungsansatz wurden verschiedene Verdünnungen (z.B. 1:10, 1:50) in SM-Puffer hergestellt. Jeweils 25 μl dieser Verdünnungen wurden mit 25 μl Empfängerzellen vermischt und für 30 min bei RT inkubiert. Anschließend wurden jeweils 200 μl LB-Medium hinzugefügt. Zur Ausbildung der entsprechenden Antibiotikaresistenz erfolgte eine Inkubation für 1 h bei 37 °C. Zum Ausplattieren auf Agarplatten mit entsprechendem Antibiotikum wurden die Ansätze 1 min zentrifugiert, der Überstand abgenommen und die sedimentierten Zellen in 50 μl LB-Medium resuspendiert.

SM-Puffer

NaCl	5,8 g
$\text{MgSO}_4 \times 7 \text{ H}_2\text{O}$	2,0 g
Tris (1 M)	50,0 ml
Gelatine (2%, (w/v))	5,0 ml
$\text{H}_2\text{O}_{\text{dest.}}$	ad 1000 ml
pH 7,5	

5.4.1 Herstellung phagenkompetenter Zellen

Als Wirtsstamm diente *E. coli* VCS 257. Von einer frisch ausgestrichenen LB-Platte wurden die Zellen in 10 ml LB-Medium mit 0,2 % (w/v) Maltose und 10 mM MgSO_4 bis zu einer OD_{595} von 1 angezogen. Die Kultur wurde auf Eis abgekühlt und abzentrifugiert (4 °C, 10 min, 2.000 Upm). Das Pellet wurde zur Hälfte des ursprünglichen Volumens in 10 mM MgSO_4 aufgenommen. Die Suspension wurde weiter bis zu einer OD_{595} von 0,5 verdünnt, um die Infektion durchzuführen.

5.4.2 Titerbestimmung

Der Verpackungs-Mix wurde 1:5, 1:10 und 1:50 mit SM-Puffer (s.o.) verdünnt und davon jeweils 25 μl mit 25 μl kompetenten Zellen ($\text{OD}_{595} = 0,5$) vermischt. Nach einer Inkubation von 30 min bei 22 °C wurden je 200 μl LB-Medium hinzugegeben. Die Proben wurden für eine weitere Stunde bei 37 °C inkubiert. Anschließend wurden die

Zellsuspensionen auf LB-Agarplatten unter Selektionsdruck ausplattiert. Um die Bank anzulegen, wurde die effektivste Verdünnung verwendet.

6. Die Prozessierung von Rohsequenzen

Nach einem Sequenzierlauf werden die Rohdaten, die Fluoreszenzsignale und -intensitäten automatisch analysiert und in DNA-Sequenzen umgerechnet, was als *Base calling* bezeichnet wird. Benutzt wurde der „Basecaller Phred“ (Ewing and Green, 1998). Dieser liest die Rohdaten von den DNA Sequenziermaschinen (*trace files*), rechnet sie in DNA-Basen um und teilt diesen zusätzlich einen Qualitätswert zu. „Phred“ kann sowohl die Rohdaten vom „ABI 377 DNA-Gerät“ als auch vom „Megabace 1000“ lesen. Die Sequenzen und die dazugehörigen Qualitätswerte können dann zur Assemblierung mit Hilfe von Sequenzassemblierungsprogrammen benutzt werden (Bonfield and Staden, 1996).

6.1 Prozessierung von DNA-Sequenzen mit dem STADEN-Software-Paket

Das STADEN-Software-Paket beinhaltet verschiedene Programme zum Selektieren, Aufbereiten, Assemblieren und Editieren von DNA-Sequenzen (Staden et al., 2000). Es würde den Rahmen dieser Arbeit sprengen, auf die Struktur und Funktionsweise der einzelnen Programme einzugehen. Eine ausführliche Dokumentation kann unter http://www.mrc-lmb.cam.ac.uk/pubseq/staden_home.html abgerufen werden. Im Wesentlichen wurden zur Prozessierung der DNA-Sequenzen die Programme „Pregap4“ und „Gap4“ genutzt.

6.1.1 „Pregap4“

Bevor DNA-Sequenzen in das Genomassemblierungsprogramm „Gap4“ eingespeist werden können, müssen diese durch diverse Prozesse geleitet werden, welche das Programm „Pregap4“ leistet (Bonfield and Staden, 1996):

- a. Screening auf Vektor-Kontaminationen: Sequenzen, die den Vektor-Sequenzen (pBluescriptIISK⁺ oder pWE15) entsprechen, fallen aus der Datenbank heraus.
- b. Screening auf Kontamination mit *E. coli* DNA: Bei der Herstellung der Genbank kann es zu Kontaminationen mit der DNA des Wirtsstammes (*E. coli* XL1blue) kommen. Diese Daten müssen ebenfalls entfernt werden.
- c. Qualitätskontrolle: Sequenzen mit einer geringen Qualität fallen aus der Datenbank heraus.

- d. „Vektor-Clipping“: Es werden die Basen einer Sequenz maskiert, die im Übergang zum Insert der Vektor-Sequenz entsprechen.
- e. Umformatierung in ein neues Dateiformat: Zur Assemblierung benötigt „Gap4“ ein anderes Dateiformat (*experimental files*).

Zur Erleichterung der Prozessierung der Rohdaten wurden im Göttinger Genomanalyselabor Computerskripte entwickelt, die hier nicht näher erläutert werden. Prinzipiell bewerkstelligen solche Skripte das Umbenennen von Dateinamen, das Einleiten der Rohsequenzen in „Pregap4“ und das Sortieren der Ausgabe von „Pregap4“ in mehrere Listen („Failed, Passed, ScreenVector“, etc.).

6.1.2 „Gap4“

Gap4 ist ein Genomassemblierungsprogramm, das auch für kleinere DNA-Abschnitte wie Cosmide oder große Plasmide verwendet werden kann (Bonfield et al., 1995). Neben der Assemblierung kann das Programm auch zur Editierung der Sequenzen benutzt werden. Eine weitere Nutzung von Gap4 liegt in dem sogenannten *Finishing*, dem Schließen von Sequenzlücken in einem Sequenzierungsprojekt.

Assemblierung: Mit Hilfe von Algorithmen („PHRAP“, <http://bozeman.mbt.washington.edu/phrap.docs/phrap.html>) werden Sequenzüberlappungen ausfindig gemacht. Sequenzläufe (*readings*), die solche Sequenzüberlappungen aufweisen, werden dann passend zueinander ausgerichtet (*Alignment*) und zu zusammenhängenden DNA-Abschnitten, den sogenannten Contigs, zusammengefasst. Aus den *alignen* Sequenzen innerhalb eines Contigs bildet das Programm dann eine Consensus-Sequenz.

Editierung: Ziel ist es hierbei, die Qualität der Consensus-Sequenz eines Contigs zu verbessern und Fehler in der Assemblierung zu detektieren und aufzulösen. Dabei wurden entsprechende Einzelsequenzen markiert, aus dem Contig extrahiert und dann manuell in den Datensatz zu passenden Contigs zurücktransferiert. Die Fehlerrate sollte nicht höher als 1 in 10000 Basen betragen. Dabei lässt sich die Qualität jeder einzelnen Base in der Consensus-Sequenz über deren *Base Confidence*-Wert überprüfen.

Finishing: In „Gap4“ können zwei (oder mehrere) Contigs anhand der Lage und Orientierung der enthaltenden *readings* zueinander geordnet werden, was eine Hilfe zur Überwindung von Sequenzlücken zwischen Contigs ist. Gegebenenfalls können Contigs dann semi-manuell mit Hilfe der Funktion *Join Contigs* miteinander verbunden werden.

7. Überwindung von Sequenzlücken

7.1 Identifikation von Contigüberlappungen

Nach erfolgter Assemblierung, deren Überprüfung und der ersten Editierung sämtlicher Contigs wurde versucht, überlappende Contigs zu identifizieren. Dies geschah zum einen automatisch über die Funktion *Find Internal Joins* in „Gap4“. Jede gefundene potentielle Contigüberlappung bedurfte im *Join Contigs Editor* einer genauen Überprüfung. Zum anderen wurde ‚extern‘ nach überlappenden Contigs gesucht, indem die Enden einiger Contigs gegen die anderer bei NCBI (National Center for Biotechnology; www.ncbi.nlm.nih.gov) mit dem Algorithmus *Align two Sequence* verglichen wurden. Contigs, die potentiell überlappten, wurden dann in „Gap4“ manuell über *Join Contigs* zueinander ausgerichtet und gegebenenfalls miteinander verbunden.

7.2 „Primer-Walking“

Mit Hilfe der „Gap4“-Funktion „Template Display“ können einzelne Contigs visuell dargestellt werden: dort sind alle im Contig enthaltenden „Readings“ und deren Templates aufgeführt. Ein Template entspricht dabei einem Insert eines rekombinanten Plasmids der Cosmid-Genbank. Die Insert der Plasmide wurden von beiden Seiten ansequenziert (*forward* und *reverse*). Da die durchschnittliche Insertgröße zwischen 2 und 4 kb lag und ein typischer Sequenzierlauf etwa 600-700 Basen lang war, bedeutet dies, dass ein Insert (= Template) nicht vollständig sequenziert wurde. Im *Template Display* sind die beiden *readings* desselben Templates zueinander angeordnet. Mit Hilfe der Funktion *Order Contigs* ließen sich dann solche Contigs aneinander reihen, die an den Rändern *readings* beinhalteten, die zu demselben Template gehörten. Beispielsweise begann oder endete der Contig A mit dem *forward reading* des Templates X, während Contig B nun mit dem *reverse reading* des Templates X endete oder begann. Teilweise konnten benachbarte Contigs durch gemeinsame BLAST-Hits bei NCBI (www.ncbi.nlm.nih.gov) mit „blastX“ identifiziert werden. Damit war es nun möglich, durch ein sogenanntes *Primer-Walking* die Lücke zwischen diesen beiden Contigs zu schließen: es wurde ein Sequenzierprimer vom Ende eines der beiden *readings* abgeleitet (mit der Orientierung zum Insertinneren), der dann zur Sequenzierung des entsprechenden *Templates* eingesetzt wurde. So konnte man die überbrückende Sequenz zwischen den Contigs ermitteln und die Lücke sofort oder nach einem weiteren Ableiten eines Oligonukleotids schließen. Waren keine Lücken überspannenden Plasmide vorhanden, wurde das Cosmid selbst als *Template*

eingesetzt. Konnte auch mit dieser Methode keine Lücken schließende Sequenz erhalten werden, wurde versucht mit einer PCR-Reaktion mit dem Cosmid als *Template* und den die Lücke flankierenden Oligonukleotiden das fehlende DNA-Fragment zu amplifizieren.

7.3 Kombinatorische PCR

Des Weiteren wurde die Methode der kombinatorischen PCR eingesetzt. Hierbei wurden zuerst Oligonukleotide von allen Contigrändern abgeleitet mit der Orientierung zum Contigende hin. Nun wurden diese in allen denkbaren Kombinationen miteinander eingesetzt, wobei eine durchschnittliche Lückengröße von nicht mehr als 2 kb angenommen wurde. Falls die Anzahl an möglichen Kombinationen zu groß erschien, wurden auch Subsätze von Oligonukleotidkombinationen angesetzt. Jedes erhaltene PCR-Produkt wurde sofort sequenziert und die Sequenz zum Lückenschluss eingesetzt.

8. Sequenzanalyse

Nach Beendigung der Sequenzierungsphase und dem Schluss der letzten Lücken in der Cosmid-Sequenz wurde versucht, alle in der Sequenz vorkommenden Gene zu identifizieren. Dafür wurden zunächst mögliche Gene vorhergesagt (ORF-Vorhersage) und diesen anschließend mögliche Funktionen zugewiesen.

8.1 ORF -Vorhersage

Alle ORF-Vorhersageprogramme für Prokaryoten basieren auf dem Auffinden von Start- (ATG, GTG, TTG) und Stopcodons (TAA, TGA, TAG) auf allen sechs möglichen Leserahmen. Danach werden solche Sequenzbereiche festgelegt, die zwischen dem Start- und Stopcodon auf demselben Leserahmen eine bestimmte *Cutoff*-Größe erreichen (oft 30 Codons). Dieser Bereich wird als ORF bezeichnet, stellt also ein potentiell Gen dar. Die ORF-Vorhersage wurde manuell mit dem Programm „Artemis“ (www.sanger.ac.uk/Software/Artemis) durchgeführt. Dieses Programm zeigt alle möglichen ORFs mit einer minimalen Größe an. Die daraus resultierenden ORFs in den sechs Leserahmen wurden auf folgende Parameter hin untersucht:

Homologien zu bekannten Proteinen: Die angezeigten potentiellen ORFs werden in „Artemis“ auch in die Aminosäuresequenz übersetzt. Damit war es möglich, in der NCBI-Datenbank (www.ncbi.nlm.nih.gov) mittels „blastP“ nach ähnlichen Proteinen zu

suchen. Gab es solche mit einer guten Übereinstimmung, war dies das stärkste Anzeichen dafür, dass der vorliegende ORF tatsächlich ein Gen darstellte.

Die Länge: Bei der Festlegung der Länge eines ORFs war die Art des Startcodons entscheidend: ATG-Startcodons wurden gegenüber den anderen bevorzugt, TGA-Stopcodons den TAG-Stopcodons vorgezogen.

Der *GC-Frameplot*: In „Artemis“ ist es möglich sich die G+C-Häufigkeit an den 3 möglichen Codonpositionen graphisch anzeigen zu lassen. Empirische Forschungen haben gezeigt, dass sich diese Häufigkeiten am Start und Stop eines Gens häufig umkehren. Dies kann eine Hilfe und ein Hinweis für das Vorliegen eines Gens sein, ist aber keine notwendige Bedingung.

Ribosomenbindestelle: Bei jedem ORF wurde überprüft, ob 4-12 Basen stromaufwärts eine Ribosomenbindestelle vorhanden war.

Mit diesen Kriterien wurden die wahrscheinlichsten ORFs übernommen. Dabei wurden soweit es möglich war Überlappungen von ORFs in unterschiedlichen Leserahmen ausgeschlossen.

8.2 Untersuchung der 16S rRNA-Gene auf chimere Sequenzen

Die Sequenzen der 16SrRNA-Gene wurden mit den Programmen „CHIMERA_CHECK version 2.7“ des „Ribosomal Database Project“ (RDP; (Larsen et al., 1993) unter der Internetadresse <http://rdp.cme.msu.edu/cgi/chimera.cgi> und dem Programm „Bellerophon chimera detection programm“ unter der Internetadresse <http://foo.math.uq.au/~huber/bellerophon.pl> auf das Vorhandensein von Chimeren überprüft.

Das Programm „CHIMERA_CHECK version 2.7“ überprüft, ob es einen möglichen Schnittpunkt in der zu untersuchenden Sequenz gibt, der die Sequenz in zwei Teile teilt, die jeweils zu einer anderen Sequenz aus der RDP-Datenbank die größte Übereinstimmung haben. Dabei macht das Programm „CHIMERA_CHECK version 2.7“ keine definitive Aussage darüber, ob eine Sequenz ein Chimere ist, so dass bei jeder Sequenz im Einzelnen auf das Vorhandensein von Artefakten entschieden werden muss.

Das Programm „Bellerophon chimera detection programm“ berechnet dagegen 2 mögliche Stammbäume mit dem Anfang und dem Ende aller Sequenzen. Ergeben sich dabei unterschiedliche Verzweigungen, ist dies ein Anzeichen dafür, dass es sich bei den entsprechenden Sequenzen um Chimere handelt.

9. Proteinbiochemische Methoden

9.1 Standardtechniken

9.1.2 Quantitative Proteinbestimmung (Bradfordtest)

Für eine quantitative Proteinbestimmung wurde der „Protein Assay Kit“ von der Firma BioRad (München) verwendet, der auf der colorimetrischen Methode von Bradford (1976, mod.) basiert. Der anionisch vorliegende Farbstoff Coomassie Brilliant Blue G250 bindet dabei an positiv geladene Aminosäuren der Proteine. Maximal können Proteinmengen von 10 µg bestimmt werden. Die untere Nachweisgrenze liegt bei 1 µg.

Durchführung:

Die Proteinlösung wurde mit 1 ml der nach Angaben des Herstellers (Konzentrat-Verdünnung 1:5 mit H₂O) angesetzten Färbelösung versetzt, der Ansatz gut gemischt und für 10 min bei RT im Dunkeln inkubiert. Nach nochmaligem Durchmischen erfolgte die Messung der OD bei 595 nm gegen eine Nullprobe ohne Protein. Die Quantifizierung erfolgte mit Hilfe einer Eichgeraden, die mit BSA im Bereich von 0-20 µg Protein/Ansatz erstellt wurde.

9.1.3 SDS-Polyacrylamidgelelektrophorese (SDS-PAGE)

Bei der SDS-Polyacrylamidgel-Elektrophorese (Laemmli, 1970) werden alle Proteine durch Einlagerung des amphipatischen, negativ geladenen SDS (Natriumdodecylsulfat) denaturiert und mit einer identischen Ladungsdichte versehen. Dies folgt aus der Einlagerung eines SDS-Moleküls pro 2 Aminosäuren, so dass die negative Gesamtladung der Größe des Proteins proportional ist. Durch die identische Ladungsdichte ist nun eine Auftrennung der Proteine nach der Masse in einem Polyacrylamid-Netzwerk möglich. Für die Elektrophorese wurde eine Minigel-Elektrophorese-Kammer verwendet. Es wurden ausschließlich diskontinuierliche Gele gegossen, die aus einem Trenngel und einem darüber liegenden Sammelgel bestanden, was zu einer besseren Auftrennung der Proben führte. Für die Trennleistung in einem bestimmten Molekularmassenbereich ist der Vernetzungsgrad des SDS-Gels entscheidend. Proteine mit einer Masse von 10.000-180.000 Da werden in einem 10 %igen Gel am effektivsten getrennt. In einem Massenbereich von 30.000 bis 250.000 empfiehlt sich ein 7,5 %iges Gel.

Acrylamid- Stammlösung (AA): lag als 40 %ige Fertiglösung vor (Fa. Biomol).

Trenngel- Stammlösung

Tris	18,2 g
SDS	0,4 g
H ₂ O _{bidest}	ad 100 ml
pH 8,8	

Sammelgel- Stammlösung

Tris	6,1 g
SDS	0,4 g
H ₂ O _{bidest}	ad 100 ml

10 x Elektrophorese- Puffer

Tris	30,3 g
Glycin	144,1 g
SDS	10,0 g
H ₂ O _{bidest}	ad 1000 ml
pH 8,4	

4 x Probenpuffer

Glycerin	7,5 ml
β- Mercaptoethanol	2,5 ml
SDS	1,2 g
Bromphenolblau (2%)	0,5 ml
Tris	0,4 g
H ₂ O _{bidest}	ad 50 ml

Ammoniumpersulfat (APS)

10 % in H₂O_{bidest}

N,N,N',N'-Tetramethyldiamin (TEMED)

Gelherstellung:

Für ein Gel wurden je eine große und eine kleine Glasplatte mit 70 %igem Ethanol gereinigt und mit dazwischenliegenden Abstandhaltern in eine Gelapparatur eingesetzt. Danach wurde das Trenngel zwischen die senkrecht stehenden Glasplatten gegossen und vorsichtig mit H₂O überschichtet, um eine gerade Geloberfläche zu erhalten. Nach mindestens 45 minütiger Polymerisation des Trenngels wurde das Wasser mit saugfähigem Papier entfernt und das Sammelgel darüber gegossen. Sofort danach wurde der Kamm luftblasenfrei eingesetzt und das Gel konnte nach dem Auspolymerisieren und Ziehen des Kammes für die Gelelektrophorese verwendet werden. Die Elektrophorese wurde bei 30 mA durchgeführt.

Tabelle 7: Pipettierschema (ausreichend für 2 Gele):

Lösung	Trenngel 10 %	Sammelgel 4 %
AA	2,0 ml	0,4 ml
TG	2,0 ml	--
SG	--	0,96 ml
H ₂ O _{bidest.}	4,0 ml	2,64 ml
TEMED	6 µl	4 µl
APS max.	60 µl	30 µl

Probenvorbereitung:

Die Proben wurden, sofern nicht anders angegeben, 1:4 mit 4 x SDS-Probenpuffer versetzt und vor dem Auftragen mindestens 5 min bei 95 °C erhitzt und anschließend kurz abzentrifugiert.

9.1.4 Molekularmassenbestimmung von Proteinen in SDS-Gelen

Bei der Elektrophorese von Proteinen wurden Marker mit definierten Molekularmassen auf ein SDS-Gel mit aufgetragen. Die Wanderungstrecken der verschiedenen Proteine konnten gegen den dekadischen Logarithmus ihrer Molekülmassen aufgetragen werden. Aus der so erhaltenen Eichgerade wurde die Molekularmasse eines mit aufgetragenen unbekanntes Proteins ermittelt. Verwendet wurden der „High Molecular Weight Standard Mixture“ Stock No. SDS-6H (Sigma)

„High Molecular Weight Standard Mixture“:

Markerproteine:	Molekularmasse [Da]
Myosin, Kaninchenmuske	205.000
β -Galatosidase, <i>E. coli</i>	116.000
Phosphorylase b, Kaninchenmuskel	97.400
Albumin, Rind	66.000
Albumin, Hühnerei	45.000
Carboanhydrase, Rinder- Erythrocyten	29.000

Für SDS-Polyacrylamidgele, die anschließend zum Protein-Transfer im Western-Blot eingesetzt wurden, wurde der Proteinstandard „BenchMark™ Prestained Protein Ladder“ der Firma GibcoBRL eingesetzt.

9.1.5 Coomassie-Färbung

Nach Beendigung der Elektrophorese wurden die SDS-Gele mit Coomassie-Blue-Färbelösung versetzt und ca. 30 min lang auf einer Laborwippe leicht geschwenkt. Entfärbt wurden die Gele mit einer Entfärberlösung, die solange gewechselt wurde, bis der Hintergrund der Gele farblos war. Die Gele wurden bis zum Trocknen mit dem BIORAD Gelrockner in H₂O_{dest.} aufbewahrt.

Färbelösung

Coomassie Blue R-250	0,1	% (w/v)
Methanol 100 % (v/v)	400	ml
Essigsäure 100 % (v/v)	100	ml
H ₂ O _{dest}	500	ml

Entfärberlösung

Methanol 100 % (v/v)	400	ml
Essigsäure 100 % (v/v)	100	ml
H ₂ O _{dest}	500	ml

9.1.6 Proteinfärbung mit Ponceau S Rot

Diese Methode konnte zur schnellen, reversiblen Anfärbung von Proteinen auf Nitrocellulose-Blotmembranen genutzt werden. Hierbei wurde die geblottete Membran (2.9.4.1) 10 min mit Ponceau S gefärbt (0,1 % (w/v) Ponceau S in 5 %iger (w/v)

Trichloressigsäure). Nach dem Färbeschritt konnte die Membran durch Spülen unter Wasser langsam entfärbt werden.

9.1.7 Umpufferung durch Dialyse

Proteinlösungen wurden durch Dialyse in Dialyseschläuchen entsalzt und umpuffert (VISKING® Dialysis Tubing, Serva, Heidelberg). Zunächst wurde ein angemessen großes Schlauchstück abgeschnitten, wobei 1 cm Dialyseschlauch ca. 2,5 ml Flüssigkeit aufnehmen kann. Der Schlauch wurde in $H_2O_{\text{bidest.}}$ für 15 min ausgekocht, anschließend mit der Proteinlösung befüllt und an beiden Enden mit Schlauchklammern abgedichtet. Die Dialyse erfolgte über Nacht unter leichtem Rühren gegen 2-5 l Puffer bei 4 °C.

9.1.8 Konzentration von Proteinlösungen

Um Proteinlösungen einzuengen, wurden Mikrokonzentratoren (Amicon Ultra, Millipore, Cork, Irland) mit einem Ausschlußgröße von 10 kDa bzw. 30 kDa verwendet. Dabei konnten Proteinlösungen von 14 ml auf 200 µl eingengt werden. Die Zentrifugation erfolgte in einer Kühlzentrifuge (Centrifuge 5804, Eppendorf, Hamburg) bei 3000 Upm.

9.2 Zellernte und Rohextraktgewinnung

9.2.1 Zellernte

Zur Gewinnung von Rohextrakten wurden *E. coli*-Zellen in Flüssigkultur angezogen (2.3.1) und geerntet. Das Pellet wurde entsprechend den weiteren Erfordernissen in Puffer aufgenommen. Dieser Puffer diente auch als Waschpuffer. Nach gegebenenfalls mehrmaligem Waschen durch Resuspendieren und Abzentrifugieren (15 min, 7000 Upm, 4 °C) wurde das Naßgewicht des Pellets bestimmt und die Zellen in 2-3 ml Puffer pro g Naßgewicht gründlich resuspendiert.

9.2.2 Zellaufschluß mit der „French Pressure Cell“

Der Aufschluß von Zellen in der „French Pressure Cell“ (American Instrument Company, Silver Spring, USA) erfolgt dadurch, daß die Zellsuspension zunächst einem Druck von 6,9 MPa ausgesetzt wird, der absinkt, wenn ein kleines Ventil geöffnet wird, und die Zellen hindurchgepreßt werden. Der plötzliche Druckabfall bewirkt das Platzen und somit die Lyse der Zellen. Das zähflüssige Lysat wird aufgefangen und zentrifugiert, wobei Zelltrümmer sedimentieren und von dem cytoplasmatischen Überstand getrennt werden. Der Überstand mit löslichen, intrazellulären Proteinen wird

als Rohextrakt bezeichnet. Diese Methode wurde ausschließlich bei größeren Kulturvolumina eingesetzt.

Durchführung:

Die Zellsuspension wurde durch dreimalige Passage durch die „French Pressure Cell“ aufgeschlossen. Das zähflüssige Lysat wurde in SS34-Röhrchen aufgefangen und 20 min zentrifugiert (15000 Upm, 4 °C).

9.2.3 Zellaufschluß durch Ultraschall

Kulturvolumina von ≤ 5 ml wurden durch Ultraschallaufschluß unter Verwendung einer Ultraschall-Sonde (Ultraschallprozessor UP 200S, 24 kHz, 200 W, Fa. Dr. Hielscher GmbH, Teltow; Sonotrode S2, 2 mm) lysiert. Die Zellsuspension wurde in einem Eppendorf-Cup 5 min beschallt (35 %ige Amplitude, Cycle 0,5). Das Abzentrifugieren der Zelltrümmer erfolgte bei 13000 Upm für 15 min in einer Tischzentrifuge (Biofuge fresco, Kendro, Osterode). Der Überstand wurde in ein neues Eppendorf-Cup überführt.

9.3 Proteinreinigung durch FPLC

Bei einer säulenchromatographischen Proteinreinigung mittels FPLC (*Fast Protein Liquid Chromatography*) wird das zu trennende Substanzgemisch in Lösung (mobile Phase) über eine Säule aus einer festen, porösen Matrix (stationäre Phase) geschickt. Die Wechselwirkungen der gelösten Bestandteile mit der stationären Phase verzögern die Wanderungsgeschwindigkeit durch die Matrix in substanzspezifischer Weise.

Durchführung:

Alle Puffer und Proben wurden vor dem Auftragen filtriert, um eine Verunreinigung des Säulenmaterials mit Staubpartikeln zu vermeiden. Die Filtration der Proben erfolgte durch Einweg-Sterilfilter (Minisart 0,2 μm , Sartorius). Die Puffer wurden mit $\text{H}_2\text{O}_{\text{bidest}}$ hergestellt und anschließend mit einer Vakuumpumpe (Savant) durch Cellulosenitrat-Filter (0,45 μm , Schleicher & Schuell, Dassel) filtriert. Um das Eindringen von Luftblasen in das Säulenmaterial zu verhindern, wurden alle Anschlüsse *drop to drop* verbunden. Die Säulen werden in 20 %igem (v/v) Ethanol gelagert. Direkte Umpufferung der Säulen von Ethanol zu einem salzhaltigen Puffer oder umgekehrt sollte durch Spülen mit MilliQ-Wasser vermieden werden.

Die verwendete ÄKTA-FPLC-Anlage besteht aus folgenden Komponenten (alle GE Healthcare, Freiburg):

Geräte:

Monitor UPC-900

Pumpe P-920

In-line Mixer M-925

Motor Valve INV-907

Fraction Collector Frac-900

10 ml und 50 ml Superloop, 1 ml und 250 µl Probenschleifen

Säulen:

SOURCE 30Q HR10/10

HIC PhenylSephrose HP

Superdex 200 prep grade XK 16/60

9.3.1 Ionenaustausch-Chromatographie (IEC)

Das Trennprinzip der Ionenaustausch-Chromatographie beruht auf elektrostatischen Wechselwirkungen zwischen den Proteinen und den funktionellen Gruppen des Gelmaterials. Dabei treten geladene Gruppen auf der Oberfläche der Proteine mit entgegengesetzt geladenen Gruppen auf der Austauschermatrix in Interaktion und führen so zu einer Bindung an die Säule, wobei die Stärke dieser Bindung von den individuellen Ladungseigenschaften der einzelnen Polypeptide abhängt. Entscheidende Parameter sind die Ionenstärke (Salzkonzentration) und der pH-Wert. Die Elution erfolgt durch eine Erhöhung der Ionenstärke, wodurch die elektrostatische Wechselwirkung zwischen Protein und Gelmaterial geschwächt wird.

Für die IEC wurde in dieser Arbeit der starke Anionenaustauscher SOURCE 30Q mit einem Säulenvolumen von 8,25 ml verwendet. Das Material besteht aus monodispersen Partikeln mit 30 µm Durchmesser und einer kontrollierten Porenstruktur, die hohe Stabilität und geringen Rückdruck gewährleisten. Die funktionelle Gruppe der SOURCE 30Q ist ein quartäres Amin, das über einen hydrophilen Spacer an die Säulenmatrix gebunden ist. Die Proteinbindungskapazität beträgt bis zu 25 mg Protein pro ml Gelbett. Die empfohlene Flußrate wird mit 300-990 cm/h (4-13 ml/min) bei einem maximalen Rückdruck von 5 MPa angegeben.

Durchführung:

Die Säule wurde zur Vorbereitung zunächst mit 1 Säulenvolumen (SV) Puffer A (= Niedrigsalzpuffer) und anschließend mit 2 Säulenvolumina Puffer B (= Hochsalzpuffer) gespült, um die funktionellen Gruppen mit Gegenionen abzusättigen. Dann wurde das Medium mit mindestens 8 SV Auftragungspuffer A auf die

Startbedingungen äquilibriert. Die Probe (in Puffer A) wurde unter Niedersalzbbedingungen auf die Säule aufgetragen und in einem ansteigenden Salzgradienten eluiert.

9.3.2 Hydrophobe-Interaktions-Chromatographie (HIC)

Diese Methode basiert auf der unterschiedlichen Hydrophobizität der einzelnen Proteinoberflächen. Durch Auftragen der Proteine in einem Puffer mit hoher Ionenstärke entstehen starke Wechselwirkungen zwischen hydrophoben Oberflächenbereichen der Polypeptide und den hydrophoben Liganden der Säulenmatrix. Treibende Kraft für diese Wechselwirkung ist die Entropie, die durch die Aggregation unpolarer Gruppen in polarer Umgebung zunimmt. Die Elution erfolgt durch eine Verringerung der Polarität des Puffers. Dies kann durch Reduzierung der Ionenstärke, den Austausch des Salzes gegen eines mit geringerem Aussalzeffekt (Hofmeister-Serie) oder durch die Zugabe von unpolaren Agenzien (Harnstoff, Detergenzien, organische Lösungsmittel) erfolgen.

Es wurde mit einer Phenyl Sepharose HP XK 16/10 Säule mit 20 ml Säulenvolumen gearbeitet, bei der die Gelmatrix aus stark quervernetzten Agarose-Partikeln besteht. Die empfohlene Flußrate beträgt 50-150 cm/h bei maximal 0,3 MPa Rückdruck.

Durchführung:

Zum Auftragen der Probe wurde die Säule mit mindestens 2 SV auf Hochsalzbedingungen äquilibriert. Die Probe musste sich ebenfalls in einem Milieu hoher Ionenstärke befinden. Hierzu wurde sie durch Dialyse über Nacht auf 1 M NH_4SO_4 hochgesalzen. Die Elution erfolgte mit einem absteigendem Salzgradienten, wobei evtl. mit $\text{H}_2\text{O}_{\text{bidest.}}$ nacheluiert werden musste, um stark hydrophobe Proteine wieder von der Säule zu lösen.

9.3.3 Gelfiltration (SEC)

Bei dieser nicht adsorptiven Methode erfolgt die Trennung der Proteine aufgrund ihrer Größe. Größeren Molekülen steht ein vergleichsweise kleines Verteilungsvolumen in der jeweiligen Gelmatrix zur Verfügung, so dass sie früher als kleinere Proteine eluieren. Für die Gelfiltration wurde eine Superdex 200 High Load 16/60 Säule verwendet. Empfohlen wird für diese Säule eine Flußrate von 10-50 cm/h bei maximal 0,3 MPa Rückdruck.

Molekularmassenbestimmung eines Proteins: Da das Elutionsvolumen mit der Größe des jeweiligen Proteins korreliert, kann die Gelfiltration nicht nur als

Reinigungsschritt, sondern auch zur Molekülmassenbestimmung herangezogen werden. Hierzu wurde der Verteilungskoeffizient K_{av} eines Proteins ermittelt

$$K_{av} = (V_e - V_o) / (V_z - V_o)$$

V_t = gesamtes Säulenvolumen

V_o = Ausschlußvolumen der Säule

V_e = Elutionsvolumen des Proteins

Durch Kalibrierung der Gelfiltrationssäule mit Eichproteinen, Ermittlung des jeweiligen Verteilungskoeffizienten und Auftragung des dekadischen Logarithmus der Molekülmasse gegen den K_{av} -Wert, konnte eine Eichgerade erstellt werden, aus der die Molekülmasse eines beliebigen Proteins anhand seines Verteilungskoeffizienten errechnet wurde.

9.4 Immunologische Methoden

9.4.1 Western-Blot (Towbin et al., 1992)

Nach der Auftrennung durch SDS-PAGE konnten Proteine aus dem Gel auf eine Nitrocellulosemembran übertragen werden. Die Übertragung erfolgt dabei elektrophoretisch unter Verwendung einer Western-Blot-Kammer nach dem Semi-Dry-Verfahren. 6 Lagen Whatmanpapier und eine Nitrocellulosemembran (Fa. Sartorius AG, Göttingen) wurden, nachdem sie auf Gelgröße zurechtgeschnitten worden waren, zusammen mit dem SDS-Gel in Transferpuffer gelegt. Anschließend wurden auf die Anodenplatte in folgender Reihenfolge drei Lagen Whatmanpapier, die Nitrocellulosemembran, das Gel und wiederum drei Lagen Whatmanpapier, jeweils unter Vermeidung von Luftblasen aufgelegt. Danach rollte man mit einer Glaspipette mehrmals über den Stapel, um überschüssigen Transferpuffer mit saugfähigem Papier zu entfernen, legte die Kathodenplatte auf und beschwerte die Apparatur mit einem Gewicht (etwa 1 kg). Nach Anschluß an ein Netzgerät (Gene Power Supply GPS 200/400, Pharmacia, Freiburg) wurden die Proteine bei konstanter Spannung und beginnend mit 5 mA/cm² Membranfläche solange übertragen, bis die Stromstärke auf weniger als die Hälfte des Ausgangswertes abgesunken war (nach ungefähr 30 min).

Transferpuffer

Tris	15,14 g
Glycin	14,40 g
Methanol	200 ml

H₂O_{dest} ad 1000 ml

pH 8,6

Lagerung bei 4 °C

9.4.2 Nachweis von Proteinen mit aufeinanderfolgenden Histidinen

Proteine mit mindestens fünf aufeinanderfolgenden Histidinresten in ihrer Aminosäuresequenz können durch spezifische Antikörper (Polyklonaler Anti-His-Tag Kaninchen-Antikörper der Firma MoBiTec, Göttingen) nachgewiesen werden. Dazu wurde die Western-Blot-Membran zunächst 2 x 10 min in 1 x TBS-Puffer gewaschen und dann für 1h bis über Nacht in Blockierungs-Puffer (TBST mit 3 % (w/v) BSA) inkubiert, um alle unspezifischen Bindestellen abzusättigen. Anschließend wurde 3 x 10 min in 1 x TBST-Puffer gewaschen, um die Membran danach 1 h lang mit dem primären Penta-His-Antikörper (~1:5000 in TBST-Puffer) inkubieren lassen zu können. Zur Entfernung ungebundener Antikörper wurde 3 x 10 Minuten lang mit TBST-Puffer gewaschen. Der sekundäre, mit alkalischer Phosphatase gekoppelte Antikörper (Anti-Rabbit-IGg-AP der Firma SIGMA) wurde 1:2.500 in TBST-Puffer aufgenommen und konnte während einer 30-minütigen Inkubation an die bereits spezifisch gebundenen primären Antikörper binden. Nach erneutem Waschen mit TBS-Puffer (3 x 10 min) erfolgte schließlich die Detektion.

TBS-Puffer

Tris 100 mM

NaCl 0,9 % (w/v)

pH 7,5

TBST-Puffer

Tris 100 mM

NaCl 0,9 % (w/v)

Tween-20 0,1 % (v/v)

pH 7,5

Blockierungs-Puffer

3 % BSA in TBS-Puffer

9.4.3 Detektion

Proteine mit konjugierter alkalischer Phosphatase (AP), wie zum Beispiel der während dieser Arbeit als sekundäre Antikörper eingesetzte Anti-Kaninchen-IgG-AP, konnten durch Zugabe einer Färbelösung sichtbar gemacht werden. Dazu wurden die Membranen, auf der sich die indirekt gebundenen und AP gekoppelten Proteine befanden, 3 min in Detektionspuffer äquilibriert. Nach Zugabe der BCIP/NBT Färbelösung zur Membran wurde solange im Dunkeln ohne Schütteln inkubiert, bis Signale in Form von dunkelbraunen präzipitiertem Substrat zu erkennen waren. Die Reaktion wurde durch Waschen der Membran mit Leitungswasser gestoppt.

Detektionspuffer

Tris-HCl	100	mM
NaCl	100	mM
pH 9,5		

Färbelösung

NBT-Lösung	66	µl
BCIP-Lösung	33	µl
Detektionspuffer	ad 10	ml

Die Färbelösung wurde vor Gebrauch frisch angesetzt, wobei 10 ml für eine Membran eingesetzt wurden.

NBT-Lösung: 75 mg/ml NBT in 70%-igem Dimethylformamid

BCIP-Lösung: 50 mg/ml BCIP in 70%-igem Dimethylformamid

9.5 Nachweis von Enzymaktivitäten

9.5.1 Qualitativer Agarase-Aktivitätstest

Um feststellen zu können, ob eine Proteinprobe Agarase-Aktivität aufwies, wurde folgendermaßen verfahren:

100 µl der zu untersuchenden Probe wurden mit 1 ml 0,2 %iger (w/v), in Tris-Puffer gelöster, flüssigen *Low Melt Agarose* (Sea Plaque® GTC® Agarose von BMA, Rockland, ME USA) in ein Eppendorfreaktionsgefäß pipettiert und bei 37 °C über Nacht auf einem Schüttler inkubiert. Als Negativkontrolle wurden 1 ml der 0,2 %igen *Low Melt Agarose* mit 100 µl Tris-Puffer vermischt und ebenfalls über Nacht inkubiert.

In Ansätzen mit Agarase-positiven Proben blieb die *Low Melt Agarose* nach ca. 12-stündiger Inkubation auch bei RT flüssig, während die Negativkontrolle nach ca. 30 min bei RT auspolymerisierte und fest wurde.

Um anschließend einen DNSA-Test durchführen zu können, wurden je 40 µl eines Agarase-positiven Zell- oder Proteinüberstandes zu 710 µl der 0,2 %igen (w/v) *Low Melt Agarose* in einem Eppendorfreaktionsgefäß gemischt und bei 37 °C über Nacht auf einem Schüttler inkubiert.

9.5.2 DNSA-Test zur Bestimmung der Menge reduzierender Zucker (Standard-Test)

Bei der Hydrolyse von polymeren Zuckern werden Oligomere bzw. Monomere gebildet, wodurch sich die Zahl der reduzierenden Enden vervielfacht. Eine Quantifizierung der reduzierenden Aldehydgruppen, bzw. α -Hydroxysubstituenten bei Ketosen, wird durch den Einsatz von 3,5-Dinitrosalicylsäure ermöglicht. In Gegenwart von Phenol und Natriumsulfit wird 3,5-Dinitrosalicylsäure durch die reduzierenden Enden zu 3-Amino-5-nitrosalicylsäure reduziert. Diese Reaktion verursacht bei 100 °C einen Farbumschlag von gelb nach braun. Die bräunliche Färbung des Testansatzes, die bei einer Wellenlänge von 575 nm photometrisch quantifiziert werden kann, ist proportional zur Bildung von 3-Amino-5-nitrosalicylsäure, deren Menge wiederum äquimolar zu der Anzahl der reduzierenden Enden ist.

Durchführung:

Der Reaktionsansatz wurde ohne Enzym bei der gewählten Inkubationstemperatur 10 min im Heizblock präinkubiert. Im Abstand von 20 s wurde dann das Enzym zugegeben, der Ansatz kurz gevortext und so schnell wie möglich wieder in den Heizblock zurückgestellt. Die Inkubation erfolgte für eine enzymabhängige Inkubationstemperatur und -zeit. Die Reaktion wurde durch Zugabe von 750 µl DNSA-Reagenz gestoppt, wobei wieder 20 s Abstand zwischen den E-Cups eingehalten wurden, um eine gleich lange Inkubation zu gewährleisten. Nach kurzem Mischen wurden die Deckel der Eppendorf-Cups mit einer Nadel durchstoßen. Die Ansätze wurden dann 15 min bei 98 °C im Heizblock inkubiert und anschließend auf Eis abgekühlt. Durch Messung der OD bei 575 nm (gegen einen Ansatz ohne Enzym) konnte mit einer aus definierten Glukosekonzentrationen hergestellten Eichgerade die Anzahl der reduzierenden Enden und somit auch die hydrolytische Aktivität des Enzyms bestimmt werden. Dabei entsprach 1 Unit der Enzymaktivität, die in 1 min zur Freisetzung von 1 µmol reduzierenden Enden führte. Alle Messungen erfolgten als Doppelbestimmungen.

Enzymassay für die Cellulase CelA:

250 mM K ₂ HPO ₄ , pH 6,5	100	μl
Substrat (2%)	250	μl
H ₂ O _{bidest}	150 – x	μl
Enzym	x	μl (2 μg)
	Σ = 500	μl

Inkubation für 15 min bei 40°C

Lösungen:

DNSA-Reagenz:

Dinitrosalicylsäure	10	g
Phenol	2	ml
Na ₂ SO ₃	0,5	g
K-Na-Tartrat	200	g
NaOH	10	g
H ₂ O _{bidest}	ad 1000	ml

Vor Licht schützen und bei 4°C lagern.

Substrat:

Carboxymethylcellulose in H ₂ O _{bid}	2	%
(medium viscosity, Sigma)		

250 mM K₂HPO₄-Puffer:

	K ₂ HPO ₄
K ₂ HPO ₄	4,35 g
H ₂ O _{bidest}	ad 100 ml

pH 6,5

9.5.3 Produktanalyse durch Dünnschichtchromatographie (TLC)

Die Dünnschichtchromatographie ist eine Form der Verteilungschromatographie zwischen einer hydrophilen (stationären) und einer hydrophoben (mobilen) Phase und dient zur Analyse von Stoffgemischen verschiedener organischer Verbindungen wie Zuckern, Aminosäuren und Nukleotiden. Auf einer mit Sorptionsmittel (Kieselgel,

Cellulose) beschichteten Dünnschichtfolie werden wenige Mikroliter des zu analysierenden Stoffgemisches aufgetragen. In einer geschlossenen Kammer bewirken Kapillarkräfte das Aufsteigen eines Lösungsmittelgemisches aus einer wäßrigen und einer organischen Komponenten (aufsteigende Chromatographie). Die wäßrige Komponente bildet mit dem Sorptionsmittel eine gelartige stationäre Phase, während die organische Komponente als mobile Phase eine Auftrennung der Moleküle im Substanzgemisch gemäß ihrer Polarität bewirkt. Die Wanderungsgeschwindigkeiten der zu trennenden Substanzen erfolgt in Abhängigkeit von ihrer relativen Löslichkeit in der stationären und der mobilen Phase, wobei unpolare Moleküle schneller wandern als polare. Aufgetrennte Substanzen werden durch Besprühen des Chromatogramms mit Reagenzlösungen detektiert, welche mit den Molekülen gefärbte Produkte bilden. Im Rahmen dieser Arbeit wurde die Dünnschichtchromatographie zur Auftrennung von Mono- und Oligosacchariden genutzt, wie sie bei hydrolytischer Spaltung auftreten. Anhand eines Cellooligosaccharid-Standards (G1-C5) wurden die Produkte identifiziert. Als stationäre Phase dienten mit Kieselgel beschichtete Aluminium-Folien (Kieselgel 60, F₂₅₄, Merck, Darmstadt). Als mobile Phase wurden folgende Laufmittel verwendet:

Laufmittel 1:		Laufmittel 2:		Laufmittel 3.	
1- Propanol	60 %	1- Propanol	50 %	Eisessig	25 %
Ethylacetat	10 %	Nitromethan	30 %	Ethylacetat	50 %
H ₂ O _{bidest}	30 %	H ₂ O _{bidest}	20 %	H ₂ O _{bidest}	25 %
Laufzeit 2x 3h		Laufzeit 1x 2 h		Laufzeit 1x3 h	

Weitere Lösungen:

Substratlösungen:	1	% (w/v)
Cellooligosaccharid-Standard:	0,1	% (w/v)

Stammlösung der Entwicklungsreagenz

Anilin	1 ml
Diphenylamin	1 g
Aceton	100 ml

Lagerung lichtgeschützt bei 4°C.

Entwicklungsreagenz

Stammlösung der Entwicklungsreagenz	5 ml
85 % (v/v) Ortho-Phosphorsäure	0,5 ml

Die Herstellung erfolgt kurz vor der Anwendung. Das Volumen ist ausreichend für die Entwicklung einer DC-Platte.

Durchführung:

Der Testansatz wurde bei 40°C für eine angemessene Zeit inkubiert und anschließend auf Eis gestellt. In der Zwischenzeit wurden die Dünnschichtplatten und die Kammer für den Lauf vorbereitet. Aliquots der Testansätze wurden entlang einer Linie 1 cm vom unteren Rand der DC-Platte strichförmig aufgetragen (Abstand der einzelnen Proben mindestens 0,7 cm; Abstand zum Seitenrand der DC-Platte mindestens 1,5 cm). Wurden Proben in einem Umfang von mehr als 2 µl aufgetragen, mußten die Probenauftrags-striche zwischen den einzelnen Aufträgen 10 min bei RT oder mit einem Fön gut getrocknet werden.

Die Chromatographiekammer wurde mit Whatmanpapier ausgekleidet und die Schliffänder von Deckel und Kammer mit Vaseline gefettet. Anschließend wurden 100 ml des jeweiligen Laufmittels über das Whatmanpapier in die Kammer gefüllt. Die trockene DC-Platte wurde in die gut äquilibrierte Dünnschicht-Kammer eingesetzt, wobei sie, um eine gerade Lauffront zu bekommen, die Seitenstreben nicht berühren durfte. Nachdem die Lauffront 2 cm unter dem oberen Plattenrand angekommen war, wurde die DC-Platte getrocknet (Lagerung bei -20 °C in Alufolie möglich) und mit frisch hergestellter Sprühlösung (Entwicklungsreagenz +1/10 Vol. Phosphorsäure) besprüht. Die besprühte Platte wurde in Alufolie eingewickelt und für 12 min bei 140 °C inkubiert. Zur Aufbewahrung wurde die Platte anschließend in Schlauchfolie eingeschweißt. Eine eventuelle Dokumentation durch Einscannen der DC-Platten mußte sofort erfolgen, da diese nach einigen Wochen einen blauen Hintergrund entwickeln.

9.5.4 Messung der Pektat Lyase Enzymaktivität

Die Bestimmung der Enzymaktivität erfolgte durch Messung des linearen Anstiegs der Absorption bei 235 nm für einen definierten Zeitraum. Die Messungen wurden in einem Zweistrahlphotometer (Varian, Darmstadt, CARY 100 Conc) im Abstand von 0,1 s durchgeführt. Dabei wurden Quarzküvetten mit einer Schichtdicke von 1 cm und einem Volumen von 100 µl verwendet. Zur Berechnung der Aktivität (U) bzw. der spez. Aktivität (U/mg) wurde das Lambert-Beersche Gesetz verwendet. Es beschreibt den Zusammenhang zwischen der

Konzentration eines absorbierenden Inhaltstoffs und der Absorption von Licht in Flüssigkeiten. Das Gesetz bildet die Grundlage der modernen Spektroskopie als analytische Methode.

$$E_{\lambda} = -\log\left(\frac{I}{I_0}\right) = c \cdot \epsilon_{\lambda} \cdot l$$

E_{λ} = Extinktion (Absorption an der Wellenlänge λ), I = Intensität des transmittierten Lichtes, I_0 Intensität des eingefallenen (eingestrahlenen) Lichtes, c = Konzentration der absorbierenden Substanz in der Flüssigkeit, ϵ_{λ} = dekadischer molarer Extinktionskoeffizient bei der Wellenlänge λ ,

l = Weglänge des Lichtes in der Flüssigkeit (auch d = Schichtdicke des Behältnisses)

Eine Unit des Enzyms ist definiert als die Menge des Enzyms welche 1 μmol eines Produkts pro Minute bildet; der molare Extinktions Koeffizient (ϵ) beträgt $4075 \text{ mol}^{-1} \text{ cm}^{-1}$ (Lieber et al., 1993).

Enzymassay für die Aktivitätsmessung:

10 mM CaCl_2	100 μl
0,5 M Glycin NaOH Puffer, pH 9,0	100 μl
2 % Polygalakturonsäure (Fa.Sigma)	125 μl
Enzym	675 –x μl
H_2O	ad 1000 μl

Die Zugabe des Enzyms erfolgte nach einer Vorinkubation des Reaktionsansatzes von 10 min bei 40°C . Als Negativkontrolle fungierte ein Testansatz ohne Enzym.

9.6 Charakterisierung der Enzymaktivität

9.6.1 pH-Abhängigkeit der Enzymaktivität

Zur Ermittlung der cellulolytischen Aktivität von CelA in Abhängigkeit vom pH-Wert wurde der Standard-DNSA-Test (2.9.5.1) mit variablem pH-Wert bei 40°C durchgeführt. Hierzu wurden drei Puffersysteme verwendet. Für pH-Werte im Bereich 4,0 bis 6,0 wurde Acetat-Puffer verwendet, im pH-Bereich von 6,0 bis 7,5 MacIlvaine-Puffer und N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) wurde für den pH-Bereich von 9,7 bis 10,5 benutzt. Alle drei Puffer wurden in einer Endkonzentration von

50 mM verwendet. Sämtliche berechneten Aktivitäten wurden als relative Aktivität bezogen auf die höchste gemessene Aktivität angegeben.

9.6.2 Temperaturabhängigkeit der Enzymaktivität

Zur Ermittlung der cellulolytischen Aktivität von CelA in Abhängigkeit von der Temperatur wurde der Standard-DNSA-Test (2.9.5.1) bei Inkubationstemperaturen von 23°C bis 95°C durchgeführt. Der K_2HPO_4 -Puffer wurde hierzu bei den jeweiligen Meßtemperaturen auf pH 6,5 titriert. Das Temperatur-Optimum wird durch zwei gegenläufige Faktoren bestimmt. Einerseits wird die Reaktion durch die steigende Temperatur thermodynamisch beschleunigt, andererseits wird die Umsetzung von Substraten durch die thermische Inaktivierung des Enzyms ab einer gewissen Temperatur begrenzt. Da die thermische Inaktivierung und somit die Lage des Temperatur-Optimums zeitabhängig ist, wurde die Inkubationszeit von 20 min beibehalten. Auch hier erfolgte die Angabe der Aktivität in relativen Enzymaktivitäten

9.6.3 Einfluß von Metallionen und Chelatoren auf die Enzymaktivität

Der Einfluß ein- und zweiwertiger Kationen sowie von Chelatoren und potentiellen Regulatoren auf die Aktivität der rekombinanten Cellulase wurde durch den DNSA-Test (2.9.5.1) unter den als Standard festgelegten Bedingungen ermittelt. Aufgrund der Färbung einiger Additivilösungen bzw. deren Einfluß auf die Farbreaktion des DNSA-Tests wurde die Enzymaktivität jedes Ansatzes gegen eine eigene Nullprobe mit den gleichen Zusätzen gemessen.

III. Ergebnisse

Der globale Umsatz durch biotechnologisch genutzte Enzyme stieg in den vergangenen Jahren von 1 Milliarde Euro in 1996 (Godfrey and West, 1996) auf 2 Milliarden Euro in 2001 (Godfrey, 2003). Es ist absehbar, daß die Nachfrage nach neuen Enzymen weiterhin im Zuge der Entwicklung neuer und spezialisierter Anwendungen steigen wird, jedoch war bisher die Anzahl an geeigneten Biokatalysatoren beschränkt (Burton et al., 2002). Die direkte Isolation von Umwelt-DNA umgeht die Notwendigkeit der Kultivierung auf der Suche nach neuen Enzymen und die damit verbundene Problematik. Auf diese Art wird eine Durchmusterung von Genbanken nach den gewünschten Enzymaktivitäten ermöglicht. Mittels dieser Methodik konnte bereits eine große Anzahl neuer Enzyme identifiziert werden (Streit et al., 2004). Hierzu gehören u.a. Esterasen und Lipasen (Ferrer et al., 2005; Rhee et al., 2005), Oxidoreduktasen (Knietsch et al., 2003), Antibiotika (Gillespie et al., 2002) und polysaccharidmodifizierende Enzyme, wie Amylasen (Yun et al., 2004), Chitinasen (Hoster et al., 2005) und Cellulasen (Grant et al., 2004; Rees et al., 2003).

Im Rahmen dieser Arbeit wurde eine Metagenom-Bank eines bakteriellen Konsortiums sowohl aktivitätsbasierend als auch sequenzbasierend nach neuen Biokatalysatoren durchmustert und darüber hinaus ein Einblick in die phylogenetische Diversität dieses angereicherten Konsortiums gewonnen. Eine im Zuge dieser Arbeit identifizierte Cellulase wurde zudem einer biochemischen Charakterisierung unterzogen.

1. Analyse der phylogenetischen Diversität mittels 16S rRNA-Genanalyse

Bei dem in dieser Arbeit genutzten bakteriellen Konsortium handelt es sich um eine Anreicherungskultur auf Minimalmedium mit Agar als einziger Kohlenstoffquelle. Für die Anreicherung wurde eine Bodenprobe eines landwirtschaftlich genutzten Feldes in unmittelbarer Nähe des Instituts für Mikrobiologie und Genetik in Göttingen verwendet. Um einen Einblick in die Diversität dieses bakteriellen Konsortiums zu erhalten, sollte zunächst eine 16S rRNA-Genanalyse durchgeführt werden.

1.1 Vereinzelnung des agarolytischen Konsortiums

Es wurde eine Vereinzelnung des Konsortiums durchgeführt, um neben der Gesamt-DNA auch die DNA der Einzelisolate für die phylogenetische Analyse nutzen zu können. Die Vereinzelnung fand auf Vollmedium (LB-Medium) statt.

Da das Konsortium in der Lage war Agar als einzige Kohlenstoffquelle zu nutzen wurde zusätzlich für die Einzelisolate das Wachstum auf Minimal-Medium getestet, dem als C-Quelle nur Agar hinzugefügt wurde (II.2.4).

Insgesamt konnten fünf Einzelisolate mit unterschiedlichem Phänotyp isoliert werden (Tabelle 1). Darunter waren einerseits eher große und schleimige Isolate wie die Klone 1 und 3, wobei das Isolat 3 bei Wachstum auf Minimal-Agar sporulierte. Andererseits gab es auch Isolate, welche kleinere Kolonien bildeten. Hierzu gehörten die Isolate 2, dessen weißliche Kolonien matt waren, und 4, dessen Kolonien gelb waren und glänzten. Das Isolat 5 zeigte einen gelblich-weißen Kolonie-Typ, der leicht glänzte.

Tabelle 1: phänotypische Beschreibungen der Einzelisolate

Bezeichnung des Isolats	Phänotyp	Wachstum auf Minimalmedium
1	große, schleimige, gelb-weiße Kolonien; glänzend	ja
2	kleine, weißliche Kolonien; matt	ja
3	mittelgroße, schleimige, beige Kolonien; glänzend	sporuliert
4	kleine, gelbe Kolonien; glänzend	nein
5	gelblich-weiße Kolonien; leicht glänzend	ja

1.2 DNA-Isolierung und Klonierung der 16S rRNA-Gene

Die chromosomale DNA sowohl des Konsortiums als auch der fünf Einzelisolate wurde wie unter II.4.4.1 beschrieben isoliert. Die variable V3-Region des 16S rRNA-Gens (Neefs et al., 1990) wurde per PCR amplifiziert (II.4.6.1). Zur Amplifizierung des ca. 1,5 kb großen Fragmentes wurden die Primer 616V und 1492R verwendet, welche spezifisch für *Bacteria* sind (Tabelle 2).

Tabelle 2: Verwendete Oligonukleotide zur Amplifizierung der 16S rRNA-Gensequenzen

Bezeichnung	Primersequenz	T _m [°C]	Referenz
616V	5'-AGA GTT TGA TYM TGG CTC AG-3'	55	(Brosius et al., 1981)
1492R	5'-CGG YTA CCT TGT TAC GAC-3'	55	(Kane et al., 1993)
GM1F	5'-CCA GCA GCC GCG GTA AT-3'	58	(Muyzer et al., 1993)

Y = C oder T; M = A oder C

Bei der PCR ist zu beachten, dass durch unvollständige Strangsynthese während der Elongations-Phase artifizielle PCR-Produkte entstehen können, die aus der rDNA zweier Organismen zusammengesetzt sind. Dies wird durch fragmentierte Template-DNA gefördert (Liesack et al., 1991; Paabo et al., 1990). Aus diesem Grund wurde eine 2-fach verlängerte Elongationszeit gewählt, sowie die Template-DNA mittels Agarose-Gelelektrophorese (II.4.3.1) auf hochmolekulare DNA kontrolliert. Die erhaltenen PCR-Produkte wurden anschließend wie in II.4.7 beschrieben in den pDrive-Vektor (Fa. Qiagen, Hilden) kloniert. Positive Transformanten wurden mittels Restriktionsanalyse überprüft (II.4.5.1). Anschließend erfolgte die Sequenzierung der hochvariablen Bereiche mit Hilfe des GM1F-Primers.

Die 16S rRNA-Gensequenzen für die fünf Einzelisolate wurden von eins bis fünf durchnummeriert. Die Sequenzierung der 16S rDNA-Klone der Gesamt-DNA des Konsortiums ergab zusätzlich fünf weitere, unterschiedliche 16S rRNA-Gensequenzen, die im weiteren mit AK für „agarolytisches Konsortium“ abgekürzt werden. Bei der Auswahl wurden Sequenzen mit über 97 % Identität als gleich angesehen.

1.3 Sequenzanalyse

Die erhaltenen Sequenzen wurden mit den Datensätzen des ribosomalen Datenbank Projektes (RDP II; (Larsen et al., 1993)) verglichen.

Die RDP II- Datenbank beinhaltet zum jetzigen Zeitpunkt (09/2005) über 177 880 bakterielle rDNA Sequenzen. Mit dem integrierten Programm „Sequence Match“ wurden die gewonnenen 16S rRNA-Gensequenzen mit der gesamten Datenbank verglichen, die sowohl Isolate als auch unkultivierte Organismen beinhaltet. Es wurden bei der Analyse alle Sequenzen der Datenbank miteinbezogen, vollständige (> 1200 bp) als auch unvollständige (< 1200 bp) 16S rRNA-Gensequenzen.

70 % der erhaltenen 16S rRNA-Gensequenzen lassen sich dem Phylum der *Proteobacteria* zuordnen (Tabelle 3). Innerhalb der γ -*Proteobacteria* zeigte das Isolat Nummer 1 und drei der 16S rDNA-Klone des Konsortiums (AK8, AK12 und AK13) die höchsten Übereinstimmungen zu Vertretern der *Pseudomonadaceae*, während das Isolat Nummer 4 die größten Übereinstimmungen zu einem unkultivierten Vertreter der Spezies *Xanthomonas* zeigte. Die α - und β -*Proteobacteria* sind mit je einer 16S rDNA-Sequenz des Konsortiums (AK14 und AK17) vertreten, wobei der Klon AK14 die größte Ähnlichkeit zu der Spezies *Agrobacterium* zeigt und der Klon AK17 zu einem unkultivierten Bakterium der Spezies *Janthinobacterium*. Die nächsten Vertreter der Isolate Nummer 3 und 5 gehören innerhalb des Phylums *Firmicutes* zu der Klasse *Bacillus*. Eines der Isolate zeigte zudem die größte Übereinstimmung zu der Spezies *Micrococcus*, welche zu dem Phylum der Aktinobakterien gehört.

Tabelle 3: Phylogenetische Einordnung der 16SrRNA-Gensequenzen anhand der RDP II-Ergebnisse

Klon	Länge (bp)	Ähnlichste Sequenz	GenBank Nummer	Phylogenetische Klassifikation	Herkunft	Ähnlichkeitswert*
1	859	<i>Pseudomonas</i> sp. BE08	AY45670 1	γ -Proteobacteria; Pseudomonadaceae Pseudomonas	NN	0,549
2	830	<i>Micrococcus</i> sp. Ellin149	AF40899 1	Actinobacteria; Micrococcaceae; Micrococcus	Bodenprobe Australien	0,984
3	831	<i>Brevibacillus</i> sp. YT-658	AB09080 3	Firmicutes; Bacilli; Brevibacillus	Bodenprobe Japan	0,970
4	833	unkultiviertes Bakterium SHA-39	AJ306770	γ -Proteobacteria Xanthomonadaceae unklassifiziert	x ¹	0,962
5	832	<i>Bacillus</i> sp. MK03	AB06267 8	Firmicutes; Bacilli; Bacillus	NN	0,967
AK8	800	unkultiviertes Bakterium DSSD95	AY32879 2	γ -Proteobacteria Pseudomonadaceae Cellvibrio	Trinkwasser verteilungs- system Simulator, USA	0,934
AK12	849	unkultiviertes Bakterium DSSD95	AY32879 2	γ -Proteobacteria Pseudomonadaceae Cellvibrio	s.o.	0,817
AK13	833	<i>Pseudomonas</i> sp. MFY30	AY33134 1	γ -Proteobacteria; Pseudomonadaceae	Bodenprobe Frankreich	0,931

Pseudomonas						
AK14	831	<i>Agrobacterium tumefaciens</i> UP-3	AY36432 9	α -Proteobacteria; Rhizobiaceae; Agrobacterium	NN	0,986
AK17	830	unkultiviertes Bakterium 300C-G04	AY66200 7	β -Proteobacteria; Oxalobacteraceae; Janthinobacterium	Grundwasserprobe, USA	0,972

* der Ähnlichkeitswert gibt die Anzahl der gemeinsamen (einzigartigen) Oligos zwischen zwei Sequenzen an, dividiert durch die niedrigste Anzahl an einzigartigen Oligos einer der beiden Sequenzen.

x¹ 1,2-dichlorpropan dechlorierende Mischkultur

NN nicht angegeben

Mit Hilfe des Programms „clustalW“ wurde unter Verwendung der eigenen und der Sequenzen der zwei nächsten Nachbarn ein Multi-Alignment erstellt, welches mit dem Programm „Mega2“ zu einem Stammbaum auf der Grundlage der „Maximum Likelihood-Methode“ verrechnet wurde (Abb. 1). Die Wahrscheinlichkeit der Verzweigungen wurde mittels der „Bootstrap“-Analyse überprüft.

Der erstellte Stammbaum spiegelt die Diversität des Bodenkonsortiums wieder und verdeutlicht die Einteilung in die einzelnen phylogenetischen Klassen.

Mit 50 % bilden die γ -Proteobacteria den größten Anteil an der Zusammensetzung des Konsortiums. Es folgen mit 20 % Vertreter der Firmicutes. Mit jeweils 10 % sind die α - und β -Proteobacteria sowie die Actinobacteria vertreten.

Von besonderem Interesse ist der hohe Anteil von vier 16S rRNA-Gensequenzen mit großen Ähnlichkeiten zu bisher unkultivierbaren Bakterien. Hierzu zählen drei der fünf Klone, welche nicht isoliert werden konnten und nur durch 16S rDNA-Klone des Konsortiums vertreten sind. Jedoch zeigt auch ein Einzelisolat die größten Übereinstimmungen zu einem bisher nicht kultivierbaren Organismus.

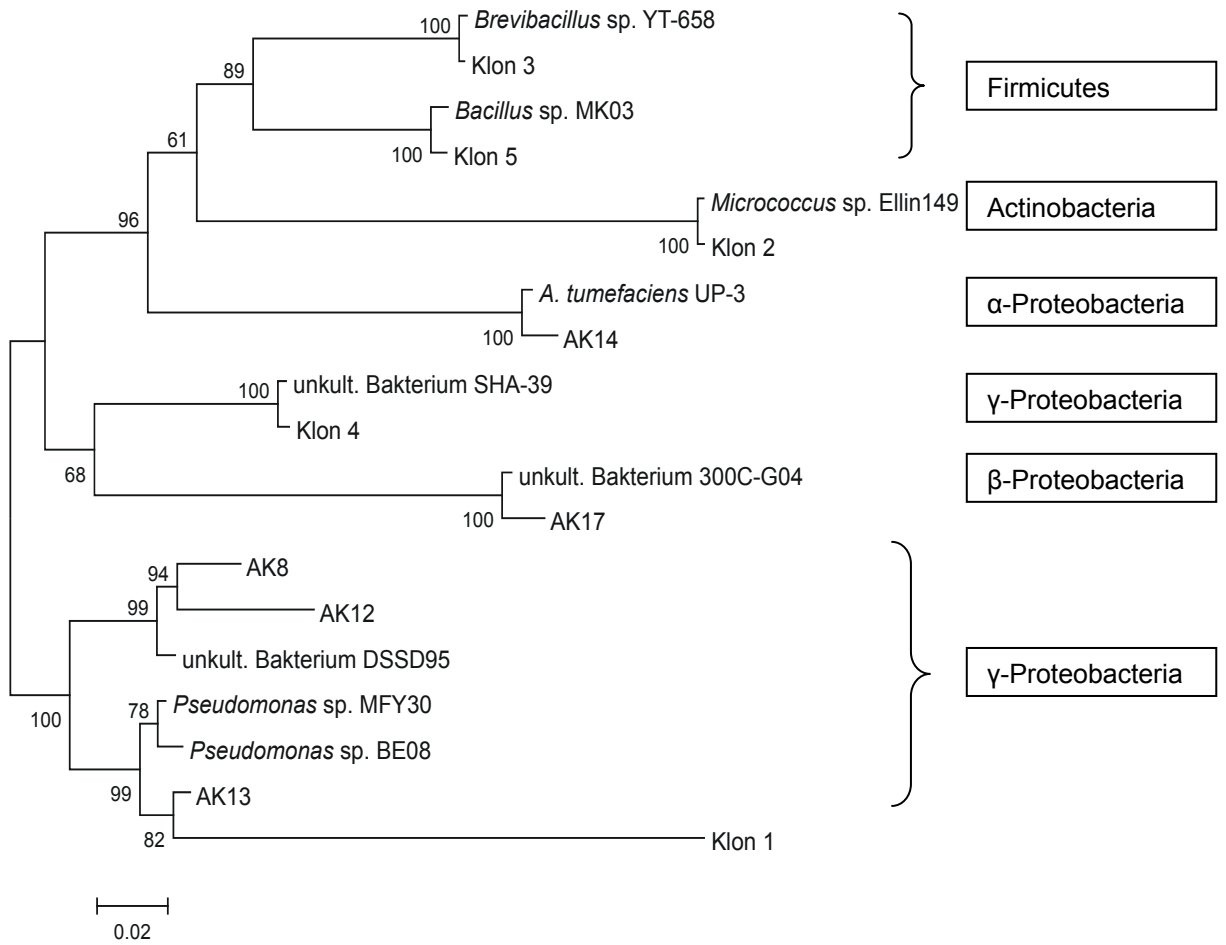


Abb. 1: Phylogenetischer Stammbaum des agarolytischen Bodenkonsortiums

2. Durchmusterung der Metagenom-Bank nach biotechnologisch interessanten Enzymen

Die Durchmusterung von Metagenom-Banken nach interessanten Biokatalysatoren beinhaltet immer auch die Frage nach der Wahl der Screening-Methode. Aktivitätsbasierende Durchmusterung bringt schnelle Ergebnisse bei der Suche nach neuen Genen mit gewünschten Eigenschaften, jedoch ist die dafür notwendige heterologe Expression im Wirtstamm ein Nachteil. Sequenzbasierende Durchmusterung von Genbanken umgeht dieses Problem, jedoch werden hierbei meist Gene bekannter Enzymklassen identifiziert, während neuartige Gene aufgrund fehlender Ähnlichkeiten zumeist als hypothetisch klassifiziert werden.

Aus diesen Gründen sollten beide Methoden kombiniert werden, um neue Biokatalysatoren zu identifizieren. Hierbei lag der Schwerpunkt auf den polysaccharidabbauenden Enzymen, insbesondere den Cellulasen und Agarasen.

2.1 Durchmusterung der Metagenom-Bank nach cellulolytischer Aktivität

Die vorhandene Metagenom-Bank, bestehend aus 1700 Cosmid-Klonen, des Bodenkonsortiums wurde mittels eines kolorimetrischen Tests auf cellulolytische Aktivität getestet. Hierzu wurden die Klone aus den Mikrotiterplatten auf LB-CMC-Agar gestempelt (II.2.3). Es erfolgte eine Inkubation der Agar-Platten von 16 h bei 37°C und eine weitere Inkubation von 24-48 h bei Raumtemperatur. Anschließend wurden die Platten für 20 Minuten mit 0,2 % Kongo-Rot überschichtet und mehrfach mit 1 M NaCl gewaschen (Teather, 1982). Kolonien mit cellulolytischer Aktivität waren durch gelbliche Höfe gegen einen roten Hintergrund gekennzeichnet (Abbildung 2). Insgesamt wurden fast 1700 Klone durchmustert, wobei elf Klone mit reproduzierbarer Aktivität identifiziert werden konnten.

Um das Vorhandensein von identischen Klonen auszuschließen, wurde eine Restriktionsanalyse durchgeführt. Von den elf Klonen mit cellulolytischer Aktivität blieben acht Klone mit unterschiedlichem Restriktionsmuster übrig (Tabelle 4). Abbildungen aller positiven Klone sind auf der beigefügten CD-Rom zu finden.

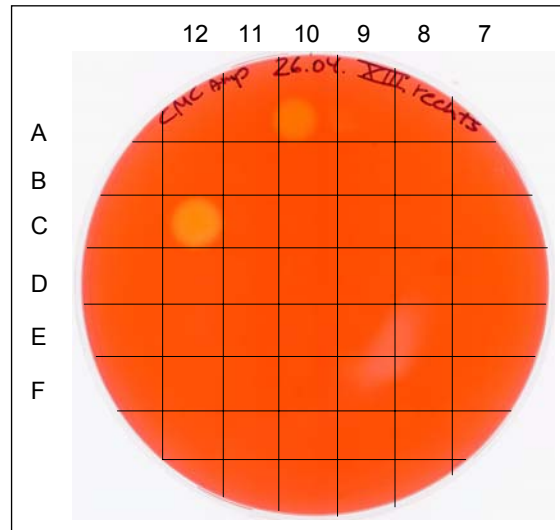


Abb. 2: Ergebnis des Cellulase Screenings auf Carboxymethylcellulose
Dargestellt sind die Cosmid-Klone XIII A10 und XIII C12.

Tabelle 4: Auflistung der Cosmid-Klone mit cellulolytischer Aktivität

Klonbezeichnung	Zuordnung zur Metagenom-Bank		
	Plattenummer	Position	Insertgröße [kb] *
pSVCosCEL-1	XIII	C12	22,0
pSVCosCEL-2	XI	A6	16,0
pSVCosCEL-3	XIII	A10	34,5
pSVCosCEL-4	XII	G2	32,5
pSVCosCEL-5	XI	A5	31,0
pSVCosCEL-6	XIV	D9	31,0
pSVCosCEL-7	XVIII	D9	16,0
pSVCosCEL-8	XVIII	D5	35,0

* Die Insertgröße wurde mittels Restriktionsanalyse und anschließender Agarose-Gelelektrophorese abgeschätzt und stellt einen Näherungswert dar.

Um einen Überblick über die Substratspezifitäten der acht Cosmid-Klone zu erhalten, wurden Aktivitätstests mit verschiedenen Oligosacchariden durchgeführt. Es wurden sowohl β -1,4-verknüpfte Glukane (CMC, Lichenan), β -1,3-verknüpftes Glukan (Laminarin) als auch mikrokristalline Cellulose (Avicel) und β -1,4-verknüpfte Xylose (Hafer-Spelzen-Xylan) eingesetzt. Die Enzymaktivität wurde mittels des unter II.9.5.2

beschriebenen DNSA-Tests zur Bestimmung der Menge an reduzierenden Enden bestimmt. Für die Messung wurden die zellfreien Rohextrakte eingesetzt. Als Negativ-Kontrolle diente der Rohextrakt des Wirtsstammes *E. coli* VCS257. In Tabelle 5 sind die errechneten Aktivitäten in mU pro mg Protein angegeben.

Tabelle 5: Aktivitäten der positiven Cosmid-Klone gegenüber verschiedenen Substraten

Klon	CMC	Lichenan	Laminarin	Avicel	Xylan
	1 %	0,5 %	0,5 %	1 %	1 %
Spezifische Aktivität [mU/mg]					
pSVCosCEL-1	-	-	-	-	0,2
pSVCosCEL-2	2,9	10,5	0,9	-	1,1
pSVCosCEL-3	1,0	3,6	0,2	-	0,2
pSVCosCEL-4	0,4	2,7	0,3	-	-
pSVCosCEL-5	0,4	2,2	0,6	-	-
pSVCosCEL-6	0,3	1,2	0,3	-	-
pSVCosCEL-7	0,2	1,3	0,2	-	-
pSVCosCEL-8	-	-	-	-	-

Alle ermittelten Aktivitäten fielen gering aus. Dies wird darauf zurückgeführt, daß sich die verwendeten Klone nicht in speziellen Überexpressionsvektoren, sondern im Cosmid-Vektor pWE15 befanden.

Mit Carboxymethylcellulose (CMC) als Substrat zeigten die Klone pSVCosCEL-2 und pSVCosCEL-3 die höchsten Aktivitäten mit 2,9 mU/mg bzw. 1,0 mU/mg. Diese beiden Klone zeigten auch mit dem linearen Glukan Lichenan die höchsten Aktivitäten, die 10,5 mU/mg für pSVCosCEL-2 und 3,6 mU/mg für pSVCosCEL-3 betragen. Mit dem β -1,3-verknüpften Laminarin betrug die höchste gemessene Aktivität 0,9 mU/mg, die auf den Klon pSVCosCEL-2 entfiel. Mikrokristalline Cellulose konnte hingegen keiner der getesteten Klone umsetzen. Zur Umsetzung des ebenfalls getesteten Xylans waren nur drei der Klone in der Lage, wobei der Klon pSVCosCEL-2 die höchste Aktivität mit 1,1 mU/mg zeigte.

2.1.1 Subklonierung von Cosmiden mit cellulolytischer Aktivität

Die Identifizierung der Cellulase-positiven Cosmid-Klone mittels Restriktionsanalyse zeigte, daß die Größe der Inserts zwischen ca. 16 und ca. 35 kb betrug. Um eine zügige Identifizierung der zu der cellulolytischen Aktivität korrespondierenden ORFs zu gewährleisten, fand zunächst eine Subklonierung der Cosmid-Klone statt. Hierzu wurde die Cosmid-DNA der Klone extrahiert (II.4.4.4) und mit dem Restriktionsenzym *EcoRI* geschnitten (II.4.5.1). Da auf beiden Seiten eines Inserts jeweils eine *EcoRI* Schnittstelle in der Multiklonierungsstelle des Vektors pWE15 vorhanden ist, wurde auf diese Weise das Insert aus dem Vektor herausgeschnitten. Somit wurde gewährleistet, daß später identifizierte Klone nur DNA des Metagenoms enthielten. Die erhaltenen Restriktionsfragmente wurden nach erfolgter Reinigung in den ebenfalls mit *EcoRI* geschnittenen und dephosphorelierten (II.4.5.3) Vektor pTZ19R ligiert (II.4.5.4). Inserttragende Transformanten wurden parallel auf LB-Agarplatten mit Chloramphenicol zur Sicherung der Klone und auf LB-Agarplatten mit 0,2 % Carboxymethylcellulose zur Identifizierung Cellulase-positiver Sub-Klone gepickt. Auf diese Art konnten *EcoRI*-Klone mit cellulolytischer Aktivität von den Cosmid-Klonen pSVCosCEL-1 und pSVCosCEL-3 isoliert werden (Tabelle 6). Mittels Restriktionsanalyse und anschließender Agarose-Gelelektrophorese wurden die beiden *EcoRI*-Klone auf ihre Insertgröße überprüft. Für den Klon pSVCEL-1 konnte eine Insertgröße von ca. 7 kb und für den Klon pSVCEL-3 eine Insertgröße von ca. 15 kb ermittelt werden.

Tabelle 6: Ergebnis der Subklonierung der Cellulase-positiven Cosmide

Cosmidklon	Bezeichnung des <i>EcoRI</i> -Klons	Insertgröße
pSVCosCEL-1	pSVCEL-1	~ 7 kb
pSVCosCEL-3	pSVCEL-3	~ 15 kb

2.1.2 Identifizierung der cellulolytischen ORFs mittels Transposonmutagenese

Aufgrund der kleineren Insertgröße wurde der Klon pSVCEL-1 für die weitere Bearbeitung ausgewählt. Die Plasmid-DNA des Klons wurde wie unter II.4.4.3 beschrieben isoliert und zur *in vitro* Transposonmutagenese eingesetzt. Die Mutagenese erfolgte nach Angaben des Herstellers. Das Transposon EZ::TN™<KAN-2> inserierte zufällig in die Ziel-DNA. Nach erfolgter Mutagenisierung wurde die DNA in *E. coli* XL1-blue transformiert (II.5.1.2). Mittels einer Kanamycinresistenz, die durch das

Transposon vermittelt wurde, fand eine Selektion auf Transposon tragende Klone statt. Das Transposon enthielt Primerbindestellen für die Primer KAN-2 FP-1 *forward* und KAN-2 RP-1 *reverse* (II.4.8, Tab. 6). Diese Primer ermöglichten eine bidirektionale Sequenzierung des DNA-Abschnitts, in den das Transposon inseriert hatte. Von den so erhaltenen Sequenzen ließen sich erneut Oligonukleotide zur weiteren Sequenzierung ableiten. Auf diese Art wurde der Klon pSVCEL-1 komplett sequenziert.

2.1.3 Sequenzanalyse des Klons pSVCEL-1

Es erfolgte eine Sequenzierung über die eigentliche Insertgröße des *EcoRI*-Klons pSVCEL-1 hinaus, um auch die angrenzenden ORFs zu komplettieren. Für die Assemblierung der Sequenzen wurde das Programm „Gap4“ des STADEN Software Pakets benutzt (Staden et al., 2000) (II.6 ff.). In der Abbildung 3 ist die für die Klonierung verwendete *EcoRI*-Schnittstelle eingetragen. Die Insertgröße des Klons pSVCEL-1 beträgt demnach 6,7 kb, was in guter Übereinstimmung mit der durch Restriktionsanalyse und Agarose-Gelelektrophorese angenommenen Größe von ca. 7 kb liegt. Insgesamt wurden 9,7 kb des ca. 22 kb großen Cosmid-Klons pSVCosCEL-1 sequenziert. Mit der erhaltenen Sequenz wurde eine manuelle ORF - Vorhersage unter Verwendung des Programms „Artemis“ durchgeführt (II.8.1). Es konnten fünf ORFs identifiziert werden (Abbildung 3), deren Aminosäuresequenz mittels „blastp“ mit der NCBI-Datenbank verglichen wurde.

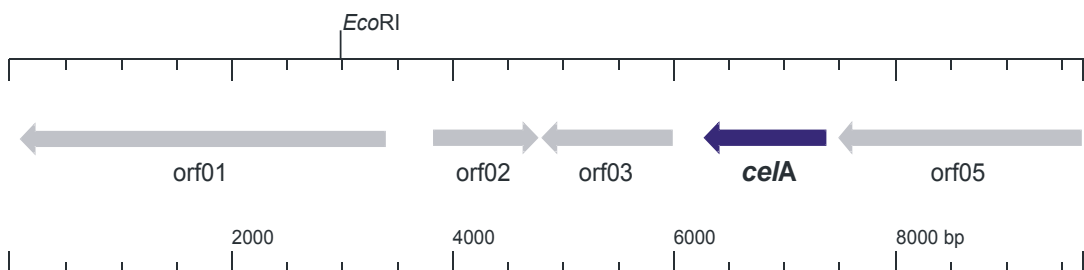


Abb. 3: Darstellung der Anordnung der offenen Leserahmen (ORFs) auf einem 9,7 kb großen Teilstück von pSVCosCEL-1

Tabelle 7: Annotation der identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Alignment Region	Homologe Alignment Region	Identität (%)
01	1095	β-1,3-Endoglukanase – <i>S. degradans</i> 2-40	ZP_0031583 9	20-813	159-953/1238	62
02	310	konserviertes hypothetisches Protein - <i>Chromobacterium violaceum</i>	AAQ60986	49-299	9-261/278	33
03	386	Pektat Lyase - <i>Frankia</i> sp.	ZP_0057293 4	148-334	143-334/365	45
<i>celA</i>	363	Cellulase A - <i>Cellvibrio mixtus</i>	AAB61461	26-362	27-363/363	77
04	729	hypothetisches Protein plu4553 - <i>P. luminescens</i>	NP_931717	145-728	678-1256/1257	27

Die DNA-Sequenz wurde bei NCBI-GenBank mit der *Accession*-Nummer AAB61461 eingetragen. Identifizierten ORFs wurde ein Name zugeteilt, wenn der ermittelte E-Wert über 10^{-80} lag.

Tabelle 7 stellt die fünf identifizierten ORFs den Proteinen mit den jeweils größten Ähnlichkeiten gegenüber. Der durch Transposonmutagenese identifizierte, für die cellulolytische Aktivität verantwortliche ORF wurde als *celA* bezeichnet. Die Aminosäuresequenz von *celA* zeigte die größte Homologie (86 % Ähnlichkeit und 77 % Übereinstimmung) zu einer bereits untersuchten Cellulase aus *Cellvibrio mixtus* (Fontes et al., 1998). Durch die Sequenzierung konnte eine Größe von 1095 bp für *celA* ermittelt werden, was einer Aminosäuresequenz von 363 AS entspricht (Abbildung 4). Unter Verwendung des online zugänglichen Programms „SignalP 3.0“ des Zentrums für biologische Sequenzanalysen (CBS) der technischen Universität von Dänemark (<http://www.cbs.dtu.dk/services/SignalP/>) wurde ein Signal-Peptid von 23 AS Länge identifiziert. Im Zuge dieser Arbeit wurde das korrespondierende Enzym CeiA heterolog in *E. coli* expremiert, aufgereinigt und anschließend charakterisiert (siehe Kapitel 3 ff.).

Interessanterweise fanden sich auf dem Klon pSVCEL-1 weitere ORFs mit hohen Ähnlichkeiten zu Proteinen, die am Polysaccharid-Abbau beteiligt sind. Hierzu zählen der ORF 01, der große Ähnlichkeit (76 % und 62 % Übereinstimmung) zu einer möglichen Glykosyl Hydrolase (EC 3.2.1.39) der Familie 81 aus *Saccharophagus*

degradans aufwies. Der ORF 03 zeigte die größte Ähnlichkeit (59 % und 45 % Übereinstimmung) zu einer wahrscheinlichen Pektat Lyase aus *Frankia* sp.. Der ORF 03 hat eine Größe von 1160 bp, was einer Aminosäuresequenz von 386 AS entspricht (siehe Anhang, Abbildung 1).

```

                                     -35
                                     |
1  attgtatntt  tatttgtgtt  atttctagcg  cctctaatac  ttatatccaa  taagaacca
    -10
    |
61  ctgaggcgct  gctatgaaat  ttccgcattg  cttgaaaacta  tttttatttc  cgttgtttat
    m k f p h c l k l f l f p l f
121  ttttatcagt  gcaaatgcgc  tggccgatga  tgcctggcaa  aatgccagtg  gctggtgga
    i f i s a n a l a d d a w q n a s g w w
181  tgcaaccgat  attcccgcgt  ttgataaaaa  caaaattacc  cggcagcttt  cattggtaaa
    n a t d i p a f d k n k i t r q l s l v
241  agttcaaggc  aataaatttg  tcgatgccaa  tggcaagggtg  gtggtatttc  gcggtgtgaa
    k v q g n k f v d a n g k v v v f r g v
301  tatttctgat  ccggataaaa  tcgtcaacga  caaacacttt  aacaaaaaac  attttgaagt
    n i s d p d k i v n d k h f n k k h f e
361  cattaagcc  tggggtgcca  atgtggtgcg  cgtgccgggtg  catccgctcg  cctggcgcaa
    v i k a w g a n v v r v p v h p s a w r
421  gcagggtaaa  aaagcctact  tggccctgct  cgaccaaaaca  gttatctggg  caaatgaatt
    k q g k k a y l a l l d q t v i w a n e
481  gggatgtat  gtgattatcg  attggcactc  gattggtaat  ctcaaactcg  aaatgttcca
    l g m y v i i d w h s i g n l k s q m f
541  aaacaattcc  tattacaccg  ataagggtga  aactaatgat  ttttggcgca  cggtatccga
    q n n s y y t d k g e t n d f w r t v s
601  gcgctatgcc  ggtataaacg  cagtggcttt  ttacgaaatt  ttaaatgagc  cgactgtttt
    e r y a g i n a v a f y e i f n e p t v
661  taatggccgt  ttaggtatgg  tgagttgggc  ggagtggaag  gcgatcaacg  aggaaattat
    f n g r l g m v s w a e w k a i n e e i
721  caccgtaatt  caagcgcata  accctaatac  aatttcactc  gtggccggtt  ttaattgggc
    i t v i q a h n p n a i s l v a g f n w
781  ttacgattta  acaccggtag  caacagcacc  tatcgatcgt  accaatgtgg  cttatgtcag
    a y d l t p v a t a p i d r t n v a y v
841  ccacccttat  ccaatgaaag  tgggcgcacc  ctatgaaaaa  aattgggagc  gcgattttgg
    s h p y p m k v g a p y e k n w e r d f
901  ttttatcgcc  gataaatatc  cagtgtttgc  cactgaaatt  ggctaccagt  tagcgaccga
    g f i a d k y p v f a t e i g y q l a t

```

```
961  taaagggggcc catattccgg tgattgacga cggtaaatac ggcaagcgca tcaccgatta
      d k g a h i p v i d d g k y g k r i t d
1021 ttttggcgat aaaggcatta gttgggtggc gtgggtgttt gaccctgatt ggtcgccgca
      y f g d k g i s w v a w v f d p d w s p
1081 gatgtttacc gactacaaaa cctacgagcc aactatgcag ggcaaacatt tccgcaagt
      q m f t d y k t y e p t m q g k h f r e
1141 tatgctgcgt gaaaataaga aatagttttt atataaccct ctctgacaa tctatgaaag
      v m l r e n k k -
```

Abb. 4: Nukleotidsequenz und die abgeleitete Aminosäuresequenz von CelA

Die dem Signal Peptid entsprechende Sequenz ist unterstrichen, die wahrscheinliche Shine-Dalgarno Sequenz ist fett und die Terminator Region ist kursiv gedruckt. Die entsprechende Aminosäuresequenz ist im Ein-Buchstabencode in grün dargestellt.

2.2 Durchmusterung der Metagenom-Bank nach agarolytischer Aktivität

Das in dieser Arbeit untersuchte Boden-Konsortium besitzt die Fähigkeit Agar zu hydrolisieren. Bei Vorarbeiten wurde die Metagenom-Bank nach den, die agarolytische Aktivität vermittelnden, Genen durchmustert (Uesbeck, 2002; Voget et al., 2003). Hierzu wurde ein Test entwickelt, der auf der Hydrolyse von Low-Melt-Agarose basiert. Die Klone der Cosmid-Bank wurden in 96 *well* Mikrotiterplatten in LB Medium angezogen. Nach erfolgter Anzucht bei 37 °C wurde der zellfreie Rohextrakt gewonnen (II.9.2.2). Zur Vereinfachung wurden die 96 Klone eines Blockes hierbei vereinigt. Die erhaltenen Rohextrakte wurden dann in einen qualitativen Agarase-Aktivitätstest eingesetzt (II.9.5.1). Positive Blöcke wurden anhand der verflüssigten Low-Melt-Agarose identifiziert, wohingegen die Agarose der negativen Blöcke sich verfestigte. Zeigten die vereinigten Rohextrakte eines Blockes agarolytische Aktivität, so wurde der Test für die 96 Einzelklone eines Blockes wiederholt.

Auf diese Weise wurden in der 1700 Klone umfassenden Metagenom-Bank vier Cosmid-Klone mit agarolytischer Aktivität identifiziert. Es handelte sich um die Klone III E5, VII C10, X B11 und XIV A3. Die römischen Ziffern geben die Nummer der Mikrotiterplatte der Metagenom-Bank an, während die Position innerhalb der Platte durch eine Buchstaben-Ziffer-Kombination gekennzeichnet ist. Die durch den qualitativen Agarase-Aktivitätstest identifizierten vier Cosmide wurden im Zuge dieser Arbeit komplett sequenziert.

2.2.1 Sequenzierung von Cosmiden mit agarolytischer Aktivität

Für die Sequenzierung der vier Cosmide wurden so genannte *shotgun*-Genbanken für jedes einzelne Cosmid angelegt. Bei genomischen *shotgun*-Genbanken wird versucht, durch mechanische oder enzymatische Behandlung der eingesetzten DNA unterschiedlich große Fragmente zu erhalten, so daß eine statistische Verteilung aller Fragmente über die gesamte DNA gewährleistet ist. Für die Erstellung von *shotgun*-Genbanken der vier agarolytischen Cosmide fand ein partieller Restriktionsverdau mit dem Enzym *Bsp143I* statt (II.4.5.2). Es folgte eine präparative Agarose-Gelelektrophorese und anschließende Extraktion der DNA aus dem Gel (II.4.4.5), wobei DNA-Fragmente von 2,5 bis 4 kb Größe ausgewählt wurden. Die erhaltenen DNA-Fragmente wurden in den zuvor mit dem Restriktionsenzym *BamHI* geschnittenen Vektor pTZ19R kloniert und in *E. coli* XL1-blue transformiert. Die anschließende Sequenzierung mit Standard-Primern (II.1; Tab. 3) erfolgte im Göttinger Labor für Genomanalyse (G₂L). Es folgte das *base calling* der Rohdaten und die

Prozessierung der Sequenzen (II.6, II.6.1.1). Eine Assemblierung der Sequenzen fand mit dem Programm „Gap4“ des STADEN Software Pakets statt (II.6.1.2), bestehende Sequenzlücken wurden geschlossen bzw. Sequenzbereiche mit schlechter Qualität oder Abdeckung überarbeitet (II.7 ff.). Nach Abschluss der Sequenzierarbeiten fand eine Festlegung der ORFs mit dem Programm „Artemis“ statt.

2.2.2 Analyse der Cosmide VII C10, XIV A3, III E5 und X B11

Die bei der Sequenzierung gewonnenen Sequenzdaten wurden für eine Annotation mittels „blastp“ bei der NCBI Datenbank eingesetzt. Die Tabelle 8 stellt die identifizierten ORFs den Proteinen mit den jeweils größten Ähnlichkeiten gegenüber. Die ausführlichen Tabellen zu jedem Cosmid-Klon befinden sich im Anhang.

Tabelle 8: Zusammenfassung der biotechnologisch interessanten ORFs der Cosmide VII C10, XIV A3, III E5 und X B11.

Cosmid-Klon / ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
VII C10						
34	260	Dehydrogenase mit unters. Spezifitäten – <i>Nostoc punctiforme</i>	ZP_00109053.1	21-258	9-246/248	65 6e-83
<i>amyA</i>	608	α -Amylase – <i>Anaeromyxobacter dehalogenans</i>	ZP_00400688.1	82-531	47-463/545	38 9e-81
<i>gnuB</i>	525	β -1,4-Endoglukanase – <i>P. fluorescens</i>	P18126	1-524	1-510/511	60 0.0
48	99	Endoglukanase – <i>S. degradans 2-40</i>	ZP_00317934.1	1-98	305-402/403	69 9e-40
<i>aguB</i>	449	β -Agarase I – <i>S. degradans 2-40</i>	AAT67062.1	15-445	9-445/593	53 e-140
<i>aguA</i>	772	hypoth. Protein – (β -Galaktosidase) <i>S. degradans 2-40</i>	ZP_00315251.1	10-772	1-769/769	53 0.0
XIV A3						
<i>aguC</i>	451	hypoth. Protein – <i>S. degradans 2-40</i>	ZP_00315360.1	1-448	294-743/747	65 0.0
<i>aguD</i>	449	beta-Agarase I – <i>S. degradans 2-40</i>	AAT67062.1	15-445	9-445/593	54

						1e-139
<i>aguE</i>	772	hypoth. Protein (β -Galaktosidase) – <i>S. degradans</i> 2-40	ZP_00315251.1	10-772	1-769/769	53 0.0
78	392	β -Xylosidase – <i>S. degradans</i> 2-40	ZP_00315659.1	64-336	48-336/368	36 3e-46
80	409	Laminarinase – <i>C. thermocellum</i>	CAC27412.2	20-177	424-596/1321	32
<i>aguF</i>	580	Agarase – <i>Pseudomonas</i> sp. ND137	BAD88713.1	31-580	34-590/591	38 1e-92

III E5

<i>aguG</i>	457	hypoth. Protein (β -Galaktosidase) – <i>S. degradans</i> 2-40	ZP_00315251.1	1-457	313-769/769	63 0.0
<i>aguH</i>	449	beta Agarase I. – <i>S. degradans</i> 2-40	AAT67062.1	15-445	9-445/593	54 1e-139
<i>agul</i>	451	hypoth. Protein. – <i>S. degradans</i> 2-40	ZP_00315360	1-448	294-743/747	65 0.0
<i>amyB</i>	634	1,4-alpha-Glukan spaltendes Enzym – <i>Azoarcus</i> sp.	YP_161008.1	28-627	23-625/636	59 0.0

X B11

83	299 Fragment	mögliche Esterase – <i>S. degradans</i> 2-40	ZP_00316903.1	1-281	8-289/380	60 1e-89
<i>aguJ</i>	772	hypoth. Protein (β -Galaktosidase) – <i>S. degradans</i> 2-40	ZP_00315251.1	10-772	1-769/769	53 0.0
<i>aguK</i>	449	β -Agarase.– <i>S. degradans</i> 2-40	AAT67062.1	15-445	9-445/593	54 1e-139
<i>aguL</i>	451	hypoth. Protein – <i>S. degradans</i> 2-40	ZP_00315360.1	1-448	294-743/747	64 0.0

War ein Nachweis für das homologe Enzym vorhanden, so wurde der entsprechende ORF in der Tabelle dunkelgrau hinterlegt. Handelte es sich um eine automatische Annotation ohne funktionellen Nachweis wurde der jeweilige ORF hellgrau hinterlegt.

2.2.2.1 Analyse des Cosmids VII C10

Für das Cosmid VII C10 wurden 26 ORFs auf 34,8 kb Länge ermittelt (siehe Abb. 6). Der GC-Gehalt betrug 47,5 %.

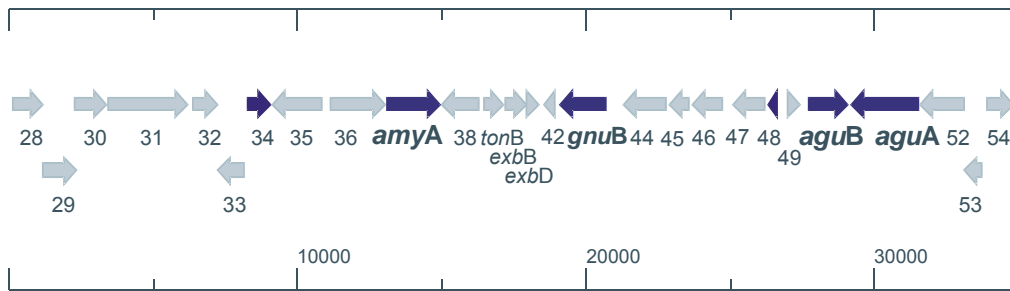


Abb. 6: Darstellung der offenen Leserahmen auf dem Cosmid VII C10

ORFs von biotechnologischem Interesse sind blau markiert.

Es konnten zwei ORFs identifiziert werden, deren korrespondierende Enzyme für die agarolytische Aktivität verantwortlich sein können. Sie wurden mit *aguA* und *aguB* bezeichnet. Durch die Sequenzierung konnte eine Größe von 2318 bp für *aguA* ermittelt werden, was einer Aminosäuresequenz von 772 AS entspricht. Ein erster Datenbank-Abgleich im Jahr 2003 zeigte die größte Übereinstimmung (41 %) zu der β -Agarase AgaB aus *Vibrio* sp. JT0107 (Sugano et al., 1994). Der jüngste Abgleich (August 2005) zeigte ein verändertes Bild. Die größte Ähnlichkeit (53 % Identität) zeigte sich jetzt zu einer möglichen β -Galaktosidase aus *Saccharophagus degradans* 2-40. Dieser Organismus ist in der Lage ein großes Spektrum an komplexen Polysacchariden, u.a. auch Agar, abzubauen (Ekborg et al., 2005).

Der ORF *aguB* ist 1349 bp groß und kodiert für ein 449 AS langes Protein. Er zeigte die größte Ähnlichkeit (53 % Identität) zu der β -Agarase AgaB, ebenfalls aus *S. degradans* 2-40 (Ekborg et al., 2005).

Weitere potentiell biotechnologisch relevante ORFs konnten identifiziert werden:

ORF 34 zeigte große Ähnlichkeiten (65 % Identität) zu einer Dehydrogenase aus *Nostoc punctiforme*, die zu der Familie der kurzkettigen Alkohol-Dehydrogenasen zu zählen ist.

Der ORF *amyA* zeigte die größte Übereinstimmung (34 % Identität) zu einer α -Amylase aus *Dictyoglomus thermophilum*. Die Funktionalität von AmyA konnte im Rahmen einer Dissertation nachgewiesen werden (Stoekigt, 2005).

Gleich zwei ORFs wiesen Ähnlichkeiten zu der Enzymgruppe der Endoglukanasen auf, die beim Abbau von Cellulose eine wichtige Rolle spielen. ORF *gnuB* zeigte die größte Ähnlichkeit zu einer β -Endoglukanase aus *Pseudomonas fluorescens* sp. *cellulosa*. Für

diese Endoglukanase konnte Aktivität gegen β -1,4-verknüpfte Glukane sowie gegen kristalline Cellulose nachgewiesen werden (Gilbert et al., 1990). Eine nicht weiter charakterisierte Endoglukanase aus *S. degradans* 2-40 zeigte 69 % Übereinstimmung zu ORF 48.

2.2.2.2 Analyse des Cosmids XIV A3

Für das Cosmid XIV A3 konnten 21 ORFs auf 38 kb Länge ermittelt werden (Abb. 8). Der GC-Gehalt betrug 44,3 %. In der Tabelle 8 sind die Ergebnisse des Abgleichs mit der NCBI-Datenbank mittels „blastp“ aufgelistet.

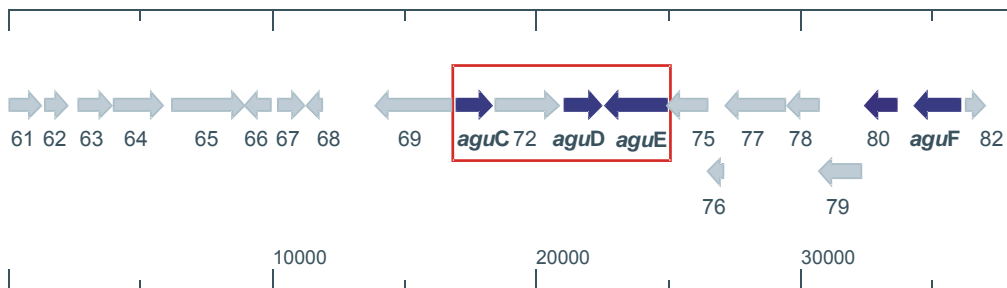


Abb. 8: Darstellung der offenen Leserahmen auf dem Cosmid XIV A3

ORFs von biotechnologischem Interesse sind blau markiert; das Cluster aus Agarase-Genen ist rot umrandet.

Auf diesem Cosmid-Klon konnten insgesamt vier Gene identifiziert werden, deren korrespondierende Enzyme für die agarolytische Aktivität verantwortlich sein könnten. Sie wurden mit *aguC*, *aguD*, *aguE* und *aguF* gekennzeichnet. Die drei erstgenannten ORFs lagen innerhalb eines 7,1 kb umfassenden Clusters (in der Abbildung rot umrandet) und zeigten, wie schon beim Cosmid VII C10, die größte Übereinstimmung zu Proteinen aus dem Organismus *S. degradans* 2-40. Darüber hinaus stimmen die ORFs *aguD* und *aguE* mit den ORFs *aguB* bzw. *aguA* überein. Dementsprechend fielen auch die Übereinstimmungen zu der NCBI-Datenbank identisch aus.

Der ORF *aguC* zeigte bei dem Datenbank-Abgleich von 2003 die größte Übereinstimmung (50 %) zu der β -Agarase AgaB aus *Vibrio* sp. JT0107. Der erneute Abgleich mit der ständig aktualisierten NCBI-Datenbank zeigte die größte Übereinstimmung (65 %) zu einem bislang als hypothetisch annotierten Protein aus *S. degradans* 2-40.

Der ORF *aguF* besitzt eine Größe von 1742 bp (580 AS) und zeigte die höchste Übereinstimmung zu einer beschriebenen Agarase aus *Pseudomonas* sp. ND137 (Matsumoto et al., 2005).

Im Zuge der Sequenzierung konnten zwei weitere ORFs von potentiell biotechnologischem Interesse ermittelt werden. Hierbei handelt es sich zum einen um den ORF 78, der eine Übereinstimmung von 36 % zu einer annotierten β -Xylosidase von *S. degradans* 2-40 aufwies. Beim Abbau des Xylans, einen wichtigen Bestandteil der Pflanzenzellwand, hydrolysieren die Endo- β -1,4-Xylanasen das Xylan-Rückgrat zu Xylobiose und kurzen Xylose-Oligomeren, die anschließend von den β -D-Xylosidasen zu Xylose-Monomeren abgebaut werden. Diese Enzyme finden vielfältige Anwendungsmöglichkeiten z.B. in der Nahrungsmittelindustrie oder der Papierherstellung (Collins et al., 2005).

Zum anderen zeigte der ORF 80 Ähnlichkeit (32 % Übereinstimmung) zu der bereits charakterisierten Endoglukanase Lic16A aus *Clostridium thermocellum*, die zu der Glykosyl Hydrolase Familie 16 (EC 3.1.2.6) gehört (Fuchs et al., 2003). Für Lic16A konnten sehr hohe Aktivitäten gegen lösliche β -1,3-verknüpfte Glukane festgestellt werden.

2.2.2.3 Analyse des Cosmids III E5

Für den dritten sequenzierten Cosmid-Klon III E5 konnten 15 ORFs auf zwei Contigs von insgesamt 21,7 kb Länge ermittelt werden (Abbildung 9). Der GC-Gehalt für Contig eins betrug 50 %, für Contig zwei 63,3 %.

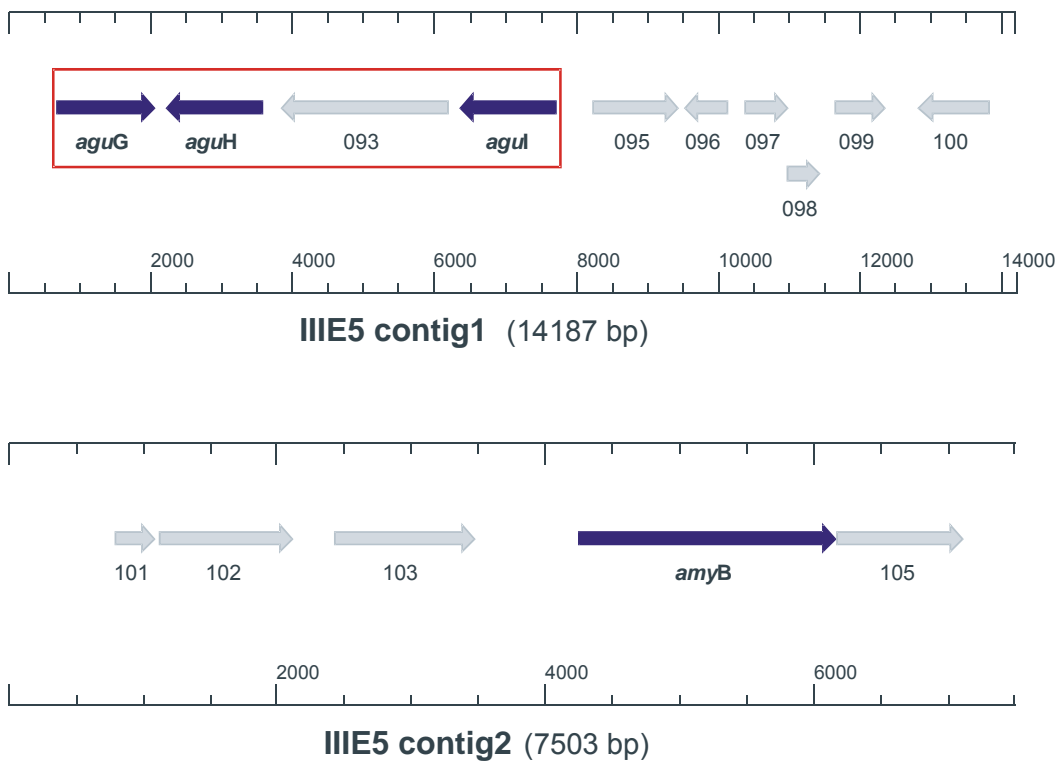


Abb. 9: Darstellung der offenen Leserahmen auf dem Cosmid III E5

ORFs von biotechnologischem Interesse sind blau markiert; das Cluster aus Agarase-Genen ist rot umrandet.

Auf diesem Cosmid wurde ebenfalls ein Gen-Cluster von drei ORFs gefunden, welche für die agarolytische Aktivität des Klons verantwortlich sein könnten. Die fortlaufend bezeichneten ORFs *aguG*, *aguH* und *agul* zeigen wie die entsprechenden ORFs der Cosmid-Klone VII C10 und XIV A3 die größten Übereinstimmungen zu Proteinen aus *S. degradans* 2-40 (Tabelle 8). Darüber hinaus stimmte der ORF *aguH* mit *aguB* und *aguD* überein, so wie der ORF *agul* identisch mit *aguC* ist. Der ORF *aguG* stimmt mit den ORFs *aguA* und *aguE* überein. Er ist jedoch mit 1373 bp rund 40 % kürzer als *aguA* und *aguE*, die beide 2318 bp lang sind.

Darüber hinaus konnte auf diesem Cosmid-Klon noch ein ORF (*amyB*) mit großer Ähnlichkeit zu einer Amylase aus *Azoarcus* sp. EbN1 identifiziert werden. Für das mit AmyB bezeichnete Enzym wurde, wie für AmyA, im Zuge einer Dissertation die Funktionalität nachgewiesen (Stoeckigt, 2005).

2.2.2.4 Analyse des Cosmids X B11

Für den Cosmid-Klon X B11 wurden fünf ORFs auf 9,8 kb Länge festgestellt (Abbildung 10). Der GC-Gehalt betrug 44,6 %. Auch hier konnten mittels „blastp“ gleich mehrere ORFs identifiziert werden, die Ähnlichkeiten zu agarolytischen Enzymen aufwiesen (Tabelle 8).

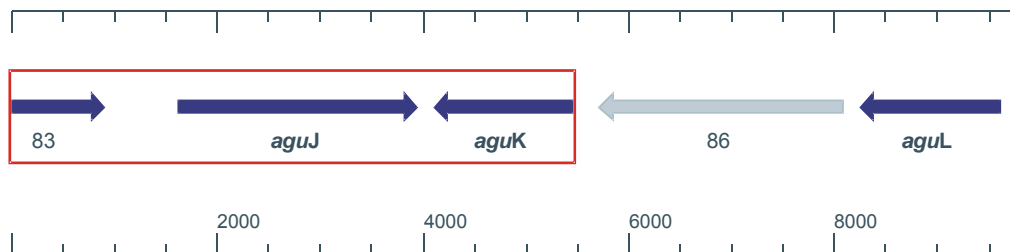


Abb. 10: Darstellung der offenen Leserahmen auf dem Cosmid X B11

ORFs von biotechnologischem Interesse sind blau markiert; das Cluster aus Agarase-Genen ist rot umrandet.

Die entsprechenden ORFs wurden mit *aguJ*, *aguK* und *aguL* bezeichnet. Auch hier lagen die drei ORFs in einem Gencluster zusammen. Der ORF *aguJ* entsprach den ORFs *aguE* und *aguA*; der ORF *aguK* den ORFs *aguH*, *aguD* und *aguB*. Der ORF *aguL* den ORFs *aguI* und *aguC*.

Der ORF 83 zeigte große Ähnlichkeit (60 % Identität) zu einer annotierten möglichen Esterase der α/β -Hydrolase Superfamilie.

2.3 Sequenzbasierende Durchmusterung der Cosmide III F2, VII H3, XIV B3, VII A1 und VII D9 nach Enzymen von industrieller Bedeutung

Die sequenzbasierte Durchmusterung von Metagenom-Banken beinhaltet ursprünglich das Design von degenerierten Primern auf der Grundlage konservierter DNA-Regionen zur Identifizierung neuartiger Gene durch direkte Amplifikation oder durch Kolonie-Hybridisierung (Bell et al., 2002; Knietsch et al., 2003).

Die Weiterentwicklung der Technologien zum automatisierten Sequenzieren und zum Assemblieren großer Datenmengen sowie die damit einhergehende Kostensenkung hat neue Möglichkeiten zur Identifizierung bisher unbekannter Biokatalysatoren eröffnet. Neue Ansätze beinhalten daher die Sequenzierung großer Abschnitte, oder sogar ganzer Metagenom-Banken auf der Grundlage des *whole genome shotgun sequencing*, einer Methode die sonst für die Sequenzierung von Einzelorganismen verwendet wird (Schmeisser et al., 2003; Tyson et al., 2004; Venter et al., 2004).

Aus diesem Grund wurden fünf Cosmide der Metagenom-Bank nach dem Zufallsprinzip ausgewählt und komplett sequenziert.

2.3.1 Analyse der Cosmide III F2, VII H3, XIV B3, VII A1 und VII D9

Die Sequenzierung erfolgte wie unter 2.2.1 beschrieben. Die bei der Sequenzierung gewonnenen Sequenzdaten wurden für eine Annotation mittels „blastp“ bei der NCBI Datenbank eingesetzt.

Die Tabelle 9 stellt die identifizierten ORFs den Proteinen mit den jeweils größten Ähnlichkeiten gegenüber. Die detaillierten Tabellen zu jedem Cosmid-Klon befinden sich im Anhang.

Tabelle 9: Zusammenfassung der biotechnologisch interessanten ORFs der Cosmide III F2, VII H3, VII A1 und VII D9

Cosmid-Klon / ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
III F2						
120	697	Pektat Lyase – S. <i>degradans</i> 2-40	ZP_00314795.1	85-695	72-680/680	46 4e-136
<i>pelA</i>	1085	Pektat Lyase – S.	ZP_00314538	488-	725-	46

<i>degradans 2-40</i>				1085	1306/1306	1e-130
125	239	metallo beta-Lactamase – <i>P. putida</i>	CAD12765.1	1-239	1-239/246	59 2e-77
VI H3						
199	416	Aminopeptidase – <i>A. tumefaciens</i>	NP_530980.1	10-416	6-412/412	85 0.0
<i>bacA</i>	268	bacitracin Resistenz Protein – <i>A. tumefaciens</i>	NP_531000.1	1-268	1-268/268	78 2e-110
XIV B3						
174	155	mögliche Metall-abhängige Hydrolase – <i>S. degradans 2-40</i>	ZP_00317817.1	1-149	2-149	58 2e-44

War ein Nachweis für das homologe Enzym vorhanden, so wurde der entsprechende ORF in der Tabelle dunkelgrau hinterlegt. Handelte es sich um eine automatische Annotation ohne funktionellen Nachweis wurde der jeweilige ORF hellgrau hinterlegt.

2.3.1.1 Analyse des Cosmids III F2

Auf dem Cosmid III F2 konnten 21 ORFs auf 28 kb Länge identifiziert werden (Abbildung 11). Der GC-Gehalt betrug 51 %.

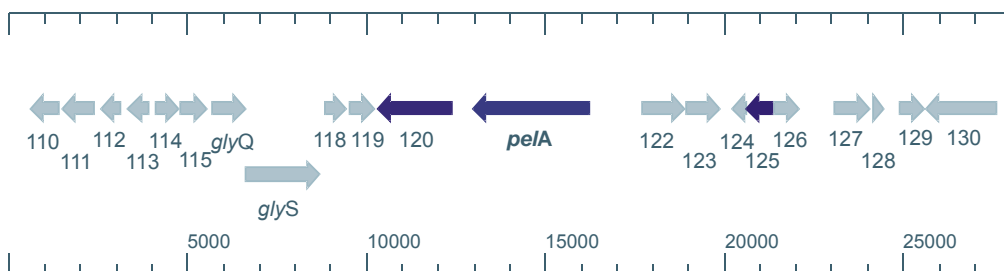


Abb. 11: Darstellung der offenen Leserahmen auf dem Cosmid III F2

ORFs von biotechnologischem Interesse sind blau markiert.

Interessanterweise zeigten 50 % der ORFs auf diesem Cosmid-Klon die größten Übereinstimmungen zu dem bereits erwähnten Organismus *S. degradans 2-40* (Anhang, Tabelle 5). Die „blastp“-Ergebnisse der restlichen ORFs zeigten keine

besondere Präferenz zu einem bestimmten Organismus, jedoch gehören alle, wie auch *S. degradans*, zu den Proteobakterien.

Es konnten drei ORFs mit Ähnlichkeiten zu Proteinen identifiziert werden, die von biotechnologischem Interesse sein könnten:

Die ORFs 120 und *peIA* zeigten die größten Übereinstimmungen (jeweils 46 %) zu zwei Pektat Lyasen aus *S. degradans* 2-40. Pektinolytische Enzyme werden von pflanzenpathogenen Organismen zum Abbau der Pflanzenzellwand produziert. Industrielle Anwendungen findet diese Enzymgruppe mit dem zumeist alkalischen pH-Optimum vor allem in der Textilindustrie (Solbak et al., 2005).

Des Weiteren zeigte der ORF 125 eine Übereinstimmung von 59 % zu der neuartigen Metallo- β -Laktamase IMP-12 aus einem klinischen *Pseudomonas putida* Isolat (Docquier et al., 2003).

2.3.1.2 Analyse des CosmidsVII H3

Nach erfolgter Sequenzierung konnten 22 ORFs auf dem 21 kb großen Cosmid-Klon VII H3 festgestellt werden (Abbildung 12). Der GC-Gehalt betrug 60,3 %. Auch hier war es möglich, ORFs von potentielltem Interesse für die Biotechnologie mittels Sequenzabgleichung mit der NCBI-Datenbank zu identifizieren.

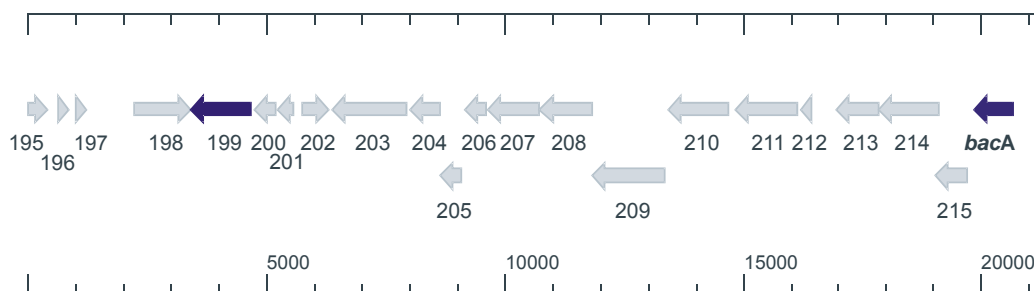


Abb. 12: Darstellung der offenen Leserahmen auf dem Cosmid VII H3

ORFs von biotechnologischem Interesse sind blau markiert.

Der ORF 199 zeigte eine hohe Übereinstimmung von 85 % zu einer annotierten Aminopeptidase aus *Agrobacterium tumefaciens* C58.

Darüber hinaus konnte ein ORF mit hoher Übereinstimmung (78 %) zu einem Protein gefunden werden, welches Resistenz gegen das Antibiotikum Bacitracin vermittelt. Dieser ORF wurde mit *bacA* bezeichnet.

90 % der ORFs konnten den Organismen *Agrobacterium tumefaciens* und *Sinorhizobium meliloti* zugeordnet werden (Anhang, Tabelle 6). Aufgrund der bei der Genom-Sequenzierung von *A. tumefaciens* gefundenen hohen Ähnlichkeiten der Beiden wird davon ausgegangen, daß sie einen gemeinsamen Vorfahren haben (Wood et al., 2001).

2.3.1.3 Analyse des Cosmids XIV B3

Auf dem Cosmid-Klon XIV B3 konnten 33 ORFs auf einer Länge von 38,5 kb ermittelt werden (Abbildung 13). Der GC-Gehalt betrug 49,5 %.

Teilt man die ORFs anhand des Datenbank-Abgleichs in verschiedene Bakteriengruppen ein, ergibt sich folgendes Bild (Anhang, Tabelle 7):

45 % der ORFs zeigen die höchste Übereinstimmung zu dem γ -Proteobakterium *S. degradans* 2-40. Weitere 21 % zu Vertretern der Pseudomonaden. Die restlichen 33 % lassen sich, mit einer Ausnahme, ebenfalls den Proteobakterien zuordnen. Insgesamt dominieren die γ -Proteobacteria mit 76 %, gefolgt von den β -Proteobacteria mit 9 %, und den α - bzw. δ -Proteobacteria mit jeweils 6 %.

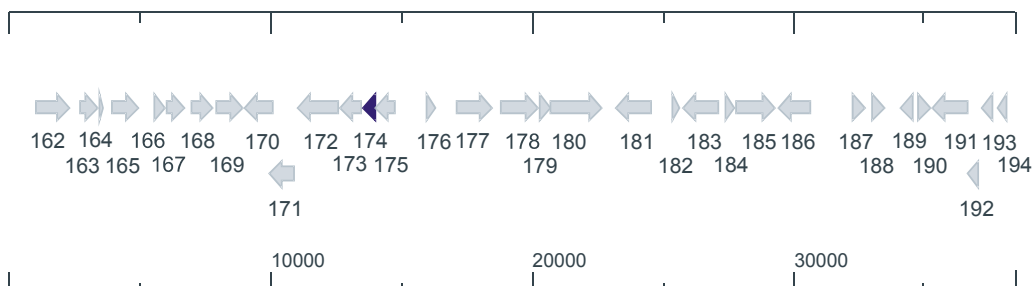


Abb. 13: Darstellung der offenen Leserahmen auf dem Cosmid XIV B3

ORFs von biotechnologischem Interesse sind blau markiert.

Auch auf diesem Cosmid konnte ein ORF von möglichem biotechnologischem Interesse identifiziert werden:

Der ORF 174 zeigte eine Übereinstimmung von 85 % zu einer potentiellen Metall-abhängigen Hydrolase aus *S. degradans* 2-40.

2.3.1.4 Analyse der Cosmide VII A1 und VII D9

Auf den beiden folgenden sequenzierten Cosmid-Klonen VII A1 und VII D9 konnten keine ORFs von möglichem biotechnologischem Interesse identifiziert werden.

Der Klon VII A1 ist 30 kb groß und enthält 23 ORFs (Abbildung 14), bei einem GC-Gehalt von 52,6 %. Der Klon VII D9 ist mit 11,6 kb Insertgröße wesentlich kleiner. Auf ihm konnten acht ORFs zugewiesen werden. Der GC-Gehalt betrug 61 %.

Betrachtet man die Verteilung der ORFs anhand der Zugehörigkeit zu einzelnen Organismen, so ergibt sich folgendes Bild:

Bei dem Cosmid VII A1 zeigten 48 % der ORFs die größten Übereinstimmungen zu *Pseudomonas fluorescens* Pf0-1 (Anhang, Tabelle 8). Darüber hinaus bilden diese ORFs ein größeres, zusammenhängendes Cluster, und, wie man anhand der GenPept Nummern erkennen kann, handelt es sich um einen zusammenhängenden Genabschnitt aus *P. fluorescens* Pf0-1. Dieser Bereich ist in der Abbildung 12 durch eine rote Umrandung gekennzeichnet. Einzig der ORF 135 passt nicht in dieses Cluster und zeigt die größte Übereinstimmung zu einem hypothetischen Protein aus *Chromobacterium violaceum*.

Weitere 18 % der ORFs sind *S. degradans* 2-40 zuzuordnen, jedoch ist keine derartige Cluster-Bildung zu beobachten.

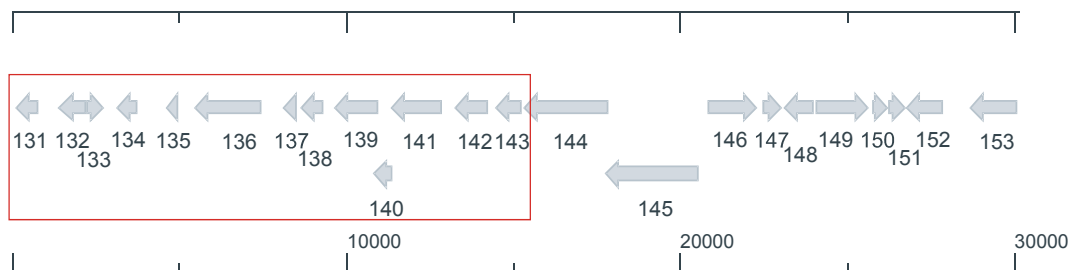


Abb. 14: Darstellung der offenen Leserahmen auf dem Cosmid VII A1

Die rote Box kennzeichnet die ORFs mit der höchsten Homologie zu einem zusammenhängenden Genabschnitt aus *P. aeruginosa*.

Bei dem Cosmid-Klon VII D9 lassen sich zwar keine Gen-Cluster feststellen, jedoch lassen sich anhand des Datenbank-Abgleichs sämtliche ORFs der *Rhizobium-/Agrobacterium*-Gruppe zuordnen (Anhang, Tabelle 9).

62,5 % der ORFs zeigen hierbei die höchste Übereinstimmung mit Proteinen aus *Agrobacterium tumefaciens* C58.

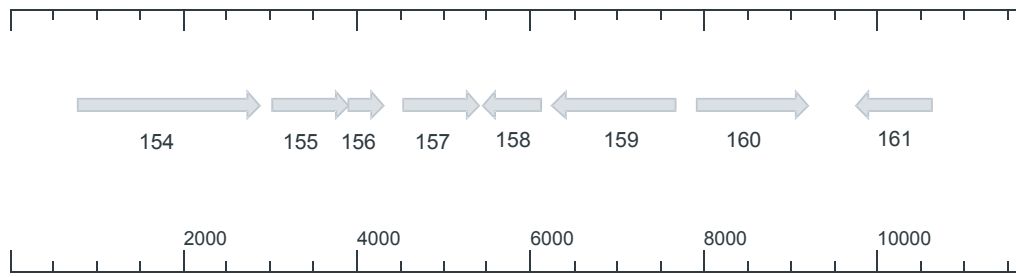


Abb. 15: Darstellung der offenen Leserahmen auf dem Cosmid VII D9

3. Nachweis der Funktionalität ausgewählter Polysaccharid-modifizierender Enzyme der Klone pSVCEL-1 und VII C10

3.1 Funktionalitätsnachweis der Pektat-Lyase PelB des Klons pSVCEL-1

Um die Funktionalität des zu dem ORF 03 korrespondierenden Enzyms zu bestätigen, fand eine Überprüfung auf pektinolytische Aktivität statt. Hierzu wurde das Gen zunächst mittels der Primer ORF03-pSVCEL-1-*forward* und -*reverse* (Tab. 3; II.1) unter Verwendung der KOD-Polymerase amplifiziert und in den Expressionsvektor pET19b kloniert. Der erhaltene Klon wurde pSVpETPEL-1 genannt. Obwohl durch Sequenzierung und Western-Blot Analyse (II.9.4.2) der angehängte N-terminale His₆-*Tag* nachgewiesen werden konnte, war eine Aufreinigung mittels Affinitätschromatographie über Ni-NTA-Säulen nicht möglich. Für eine Aufreinigung mittels Ionenaustausch-Chromatographie wurde der starke Anionenaustauscher Source 30 Q gewählt. Die Tabelle 10 gibt die Parameter des Säulenlaufes wieder.

Tabelle 10: Parameter des Source 30 Q Säulenlaufes.

Medium:	Source 30 Q
Säule:	HR 10/10
Säulenvolumen:	8,25 ml
Puffer A:	50 mM Tris-HCl pH 8.0
Puffer B:	50 mM Tris-HCl pH 8.0 + 1 M NaCl
Flussrate:	1,5 ml/min
Fraktionen	5 ml
Äquilibrierung:	2 SV Puffer B, 6 SV Puffer A
Gradient:	0 – 100 % Puffer B (5 SV)

Es wurden alle Fraktionen gesammelt und anschließend der Proteingehalt bestimmt (II.9.1.2). Wie anhand des Chromatogramms in Abbildung 16 zu sehen ist, befand sich ein Großteil des aufgetragenen Proteins in der Auftragsfraktion zwei und hatte nicht an das Säulenmaterial gebunden.

Eine Analyse mittels SDS-Gelelektrophorese zeigte in den Durchlaufaktionen 1-3 Proteinbanden in der erwarteten Größe von 39 kDa (Abbildung 17).

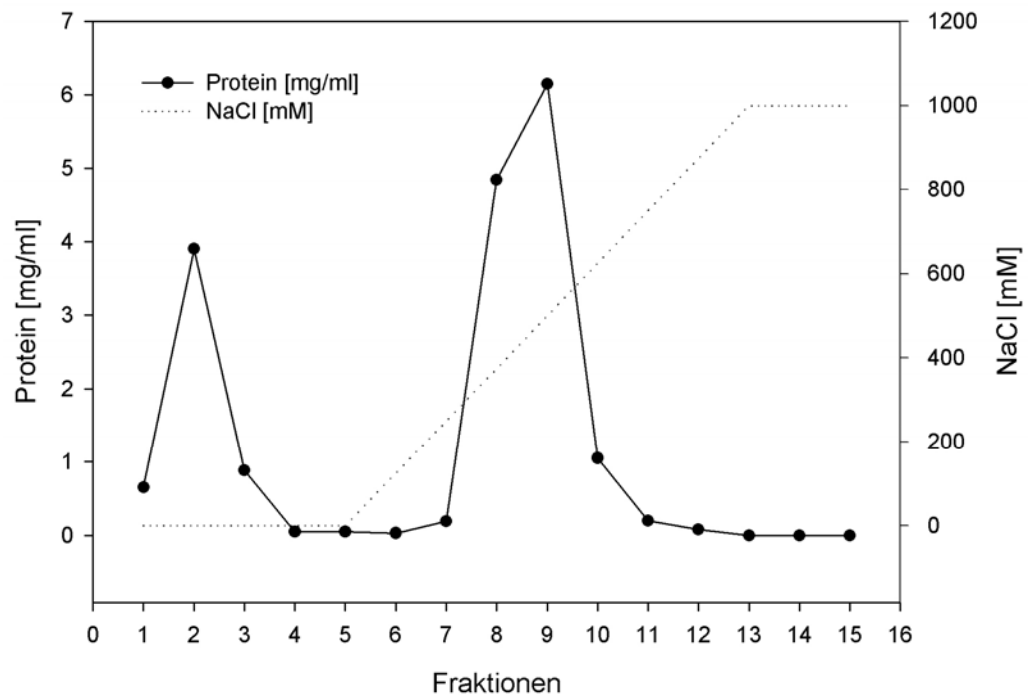


Abb. 16: Chromatogramm des Source30Q Säulenlaufs zur Reinigung der Pektat Lyase

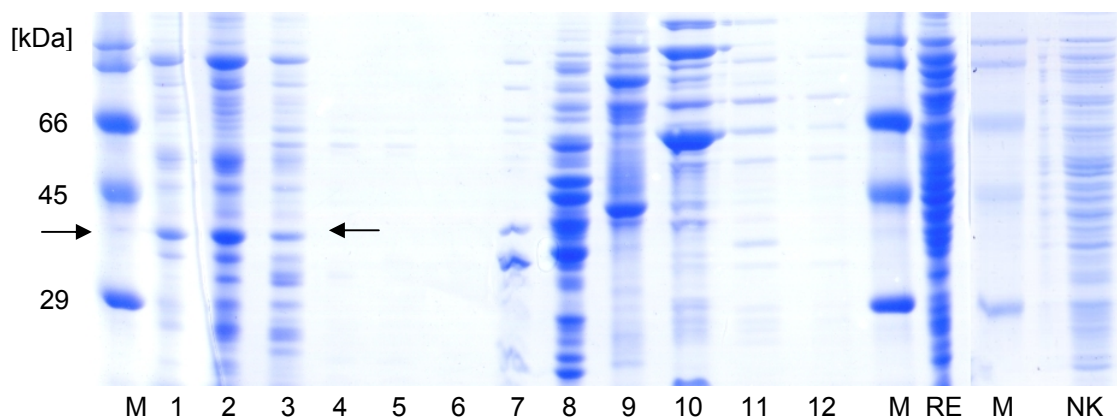


Abb. 17: SDS-Gel aller Fraktionen des Source30Q Säulenlaufs zur Reinigung der Pektat-Lyase

M: *high molecular weight marker* (Sigma); Spuren 1-12: Fraktionen 1-12; RE: Rohextrakt pSVPEL-1. NK: Rohextrakt pET19b ohne Insert. Die Pfeile markieren die erwartete Laufhöhe der Pektat Lyase bei ca. 40 kDa.

Es erfolgte eine Identifizierung durch Aktivitätsmessung der Fraktionen eins bis drei sowie acht und neun. Pektat-Lyasen katalysieren die Hydrolyse der α -1,4-Verknüpfungen der Polygalakturonsäure (PGA), dem Rückrat des Pektins. Dies macht man sich bei der Bestimmung der Enzymaktivität zunutze, indem man den linearen Anstieg der Absorption bei 235 nm verursacht durch die Bildung von ungesättigten Produkten, mittels eines Zweistrahlphotometers misst. Nach Ermittlung der Steigung kann dann durch Anwendung des Lambert-Beerschen Gesetzes die spezifische Aktivität berechnet werden (siehe II.9.5.4).

Da Pektat-Lyasen Ca^{2+} -Ionen für ihre enzymatische Aktivität benötigen und die Effektivität bei alkalischen pH-Werten zwischen 8 und 10 am größten ist (Sakai et al., 1993), wurde der Testansatz auf diese Bedürfnisse abgestimmt. Der Standard-Reaktionsansatz enthielt 0,25 % Polygalakturonsäure als Substrat in 50 mM Glycin/NaOH-Puffer (pH 9.0) und dem Zusatz von 1 mM CaCl_2 . Die Enzymzugabe erfolgte nach einer Vorinkubation von 10 min bei 40 °C. Als Negativkontrolle fungierte ein Testansatz ohne Enzym.

Nur in Fraktion zwei des Source Q30-Säulenlaufes konnte Enzymaktivität festgestellt werden. Es wurde eine Aktivität von 9 U/mg berechnet.

3.2 Funktionalitätsnachweis und Expressionsstudien zu der Agarase AguB des Klons VII C10

Im Zuge dieser Arbeit wurde das Gen *aguB* in den Überexpressionsvektor pET21a kloniert und das korrespondierende Enzym AguB heterolog in *E. coli* BL21 expremiert. Bei einer Länge von 449 AS wurde ein Molekulargewicht von 49 kDa erwartet. Wie in Abb. 18 zu sehen ist, lag das Protein als Einschlußkörper (*inclusion bodies*) vor. Diese treten bei der Akkumulation hoher Mengen von inaktivem Fremdprotein in der Wirtszelle auf und sind auf eine unzureichende Proteinfaltung zurückzuführen. Um den Anteil an löslichem Protein zu erhöhen, wurden Versuche mit verringerter Wachstumstemperatur, -dauer sowie veränderter IPTG-Zugabe durchgeführt. Diese ergaben jedoch keine besseren Ergebnisse. Eine Aufreinigung mittels des angefügten C-terminalen His-Tags unter nativen Bedingungen war nicht möglich. Der His-Tag konnte jedoch durch eine Aufreinigung unter denaturierenden Bedingungen bestätigt werden. Die Funktionalität des Enzyms konnte mittels des qualitativen Agarase-Aktivitätstests (II.9.5.1) unter Verwendung des Rohextrakts nachgewiesen werden.

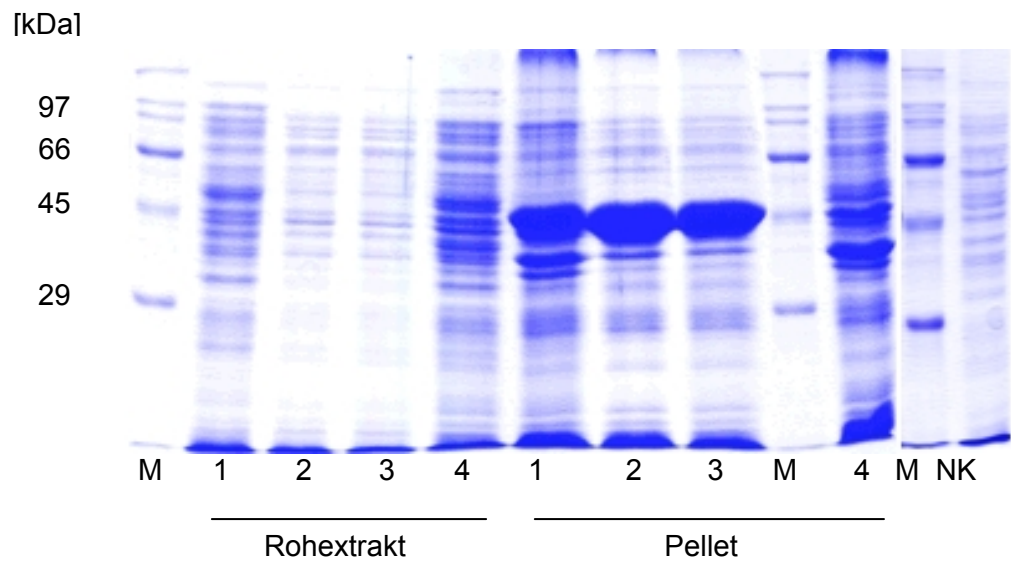


Abb. 18: SDS-Gel der Expressionsstudien mit der Agarase AguB

Spur M: *high molecular weight marker* (Fa. Sigma); Probe 1: Induktion mit 0,1 mM IPTG und Inkubation über Nacht; Probe 2: 0,1 mM IPTG, vier Std. Inkubation, Probe 3: 0,5 mM IPTG, vier Std. Inkubation; Probe 4: ohne IPTG, vier Std. Inkubation. NK: Rohextrakt pET21a ohne Insert.

4. Klonierung, heterologe Expression und Aufreinigung der Cellulase CelA

Der mittels Transposonmutagenese auf pSVCEL-1 identifizierte ORF *celA* wurde mit den Primern *celA-forward* und *celA-reverse* amplifiziert (II.1, Tab. 3; II.4.6.2). Zur Amplifikation wurde die KOD-Polymerase (Fa. Novagen/ Merck KGaA, Darmstadt, Deutschland) gewählt, da diese Polymerase eine *Proofreading*-Aktivität bei extrem kurzen Hybridisierungs- und Verlängerungszeiten hat. Die Primer waren so gewählt, daß eine 5'-*NdeI* sowie eine 3'-*BamHI* Schnittstelle eingefügt wurden. Nach Überprüfung des PCR-Produktes durch Agarose-Gelelektrophorese und anschließende Gel-Extraktion (II.4.4.5) folgte die Klonierung in den Vektor pBSK+ (Fa. Stratagene, San Diego, USA). Positive Transformanten konnten durch Blau-Weiß-Selektion (II.5.1.3) identifiziert werden. Es folgte eine Restriktion mit den Enzymen *BamHI* und *NdeI* (2.4.5.1) und Ligation (II.4.5.4) in den auf gleiche Weise geschnittenen Expressions-Vektor pET19b (Fa. Novagen).

Das pET-Expressionssystem macht sich die hohe Aktivität und Spezifität der T7-RNA-Polymerase zu Nutze. Dies ermöglicht ausgehend vom T7-Promotor eine regulierte Expression von heterologen Genen in *E. coli*. Außerdem besteht die Möglichkeit, das Zielgen N- oder C-terminal mit einem *Tag* (u.a. His₆-*Tag*) zu versehen, der später eine einfache Detektion mit Antikörpern oder die Reinigung des gebildeten Proteins ermöglicht.

Die erhaltenen rekombinanten Klone wurden mit T7-Primern sequenziert (II.1; Tab. 3) und der Klon pSVpETCEL-1 für die weiteren Arbeiten ausgewählt. Die Anzucht erfolgte in 2 l LB Medium mit Ampicillin als Selektivmarker bei 30 °C und 90 Upm. Die Überexpression wurde durch die Zugabe von 0,1 mM IPTG induziert und über Nacht inkubiert.

Die Zellen wurden geerntet (II.9.2.1) und in 50 mM Tris-HCl Puffer pH 8.0 resuspendiert. Der Zellaufschluß erfolgte mit einer *French Pressure Cell* (II.9.2.2) (American Instrument Company, Silver Spring, USA). Zur Gewinnung der Rohextrakts wurden die Zelltrümmer durch Zentrifugation pelletiert. Aus 2 l Kulturvolumen wurden 16 g Zellen (Nassgewicht) geerntet, aus denen 21 ml Rohextrakt mit einer Proteinkonzentration von 46,1 mg/ml gewonnen wurden.

Obwohl durch Sequenzierung und Western-Blot Analyse (II.9.4.2) der angehängte N-terminale His₆-*Tag* nachgewiesen werden konnte, war eine Aufreinigung mittels Affinitätschromatographie über Ni-NTA-Säulen nicht möglich. Aus diesem Grund

erfolgte eine dreistufige Aufreinigung durch das ÄKTA FPLC System (GE Healthcare) und geeigneten Säulen, die in den folgenden Kapiteln genauer erläutert wird.

4.1 Aufreinigung von C_{el}A mittels Anionenaustauschchromatographie

Das Trennprinzip der Ionenaustausch-Chromatographie beruht auf elektrostatischen Wechselwirkungen zwischen den Proteinen und den funktionellen Gruppen des Gelmaterials (II.9.3.1).

Da der errechnete Isoelektrische Punkt von C_{el}A bei 6,1 lag, wurde bei der Verwendung eines Puffers mit höherem pH, in diesem Fall pH 8,0, von einer negativen Ladung des Proteins ausgegangen und der Anionenaustauscher gewählt.

Für die Aufreinigung war zunächst eine Entsalzung des Rohextraktes nötig. Dies erfolgte mittels des *Desalting Kits* der Firma GE Healthcare und Vorgehensweise nach Protokoll. Anschließend erfolgte der Auftrag des gesamten Rohextraktes auf den verwendeten, starken Anionenaustauscher Source 30 Q. Tabelle 11 listet die Parameter des Säulenlaufes auf.

Tabelle 11: Parameter des Source 30 Q Säulenlaufes.

Medium:	Source 30 Q
Säule:	HR 10/10
Säulenvolumen:	8,25 ml
Probe:	968 mg
Puffer A:	50 mM Tris-HCl pH 8.0
Puffer B:	50 mM Tris-HCl pH 8.0 + 1 M NaCl
Flussrate:	1,5 ml/min
Fraktionen	5 ml
Äquilibration:	2 SV Puffer B, 6 SV Puffer A
Gradient:	0 – 100 % Puffer B (5 SV)

Die Elution erfolgt bei der Ionenaustauschchromatographie durch eine Erhöhung der Ionenstärke, wodurch die elektrostatische Wechselwirkung zwischen Protein und Gelmaterial geschwächt wird. Für die Aufreinigung von C_{el}A wurde hierbei NaCl unter Nutzung eines linearen Gradienten verwendet. Alle Fraktionen wurden gesammelt und

für jede sowohl der Protein-Gehalt (II.9.2.1) als auch die Cellulase Aktivität bestimmt. Die Enzymaktivität wurde hierbei unter Verwendung des DNSA Standardtests zur Bestimmung der Menge an reduzierenden Enden gemessen (II.9.5.2). Als Substrat diente das β -Glukan der Gerste. Parallel fand eine Überprüfung durch SDS-Gelelektrophorese statt (II.9.1.3 ff.).

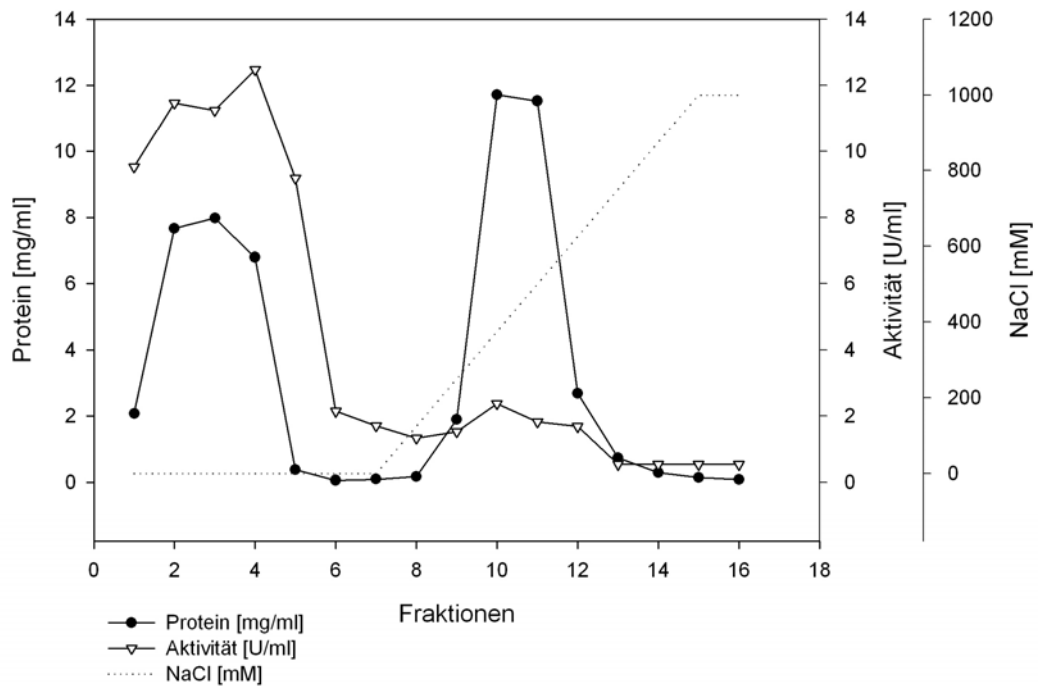


Abb. 19: Chromatogramm des Source 30 Q Säulenlaufs zur Aufreinigung des rekombinant hergestellten CelA Proteins.

Wie anhand des Chromatogramms in Abbildung 19 zu sehen ist, hatte ein Großteil des Proteins nicht an das Säulenmaterial gebunden und befand sich in den Durchlauf fraktionen 1-4. Die Aktivitätsmessungen und die SDS-Gelelektrophorese bestätigten den Verdacht, das CelA nicht an die Säulenmatrix gebunden hatte. Trotzdem wurde dieser Aufreinigungsschritt beibehalten, da durch die Bindung von Wirtsproteinen an die Matrix des Anionenaustauschers ein Aufreinigungseffekt erzielt werden konnte. Dies ist auch in der Abbildung 20 zu erkennen, die eine Analyse der Eluate der einzelnen Aufreinigungsschritte darstellt. Die durch die SDS-Gelelektrophorese ermittelte Größe von CelA beträgt 42 kDa, was in guter Übereinstimmung mit der vorhergesagten Größe von 42,1 kDa liegt.

Die entsprechenden Fraktionen wurden vereinigt und es wurde ein Proteingehalt von 2,39 mg/ml und eine Enzymaktivität von 31,3 U/mg festgestellt (S. 34, Tabelle 12).

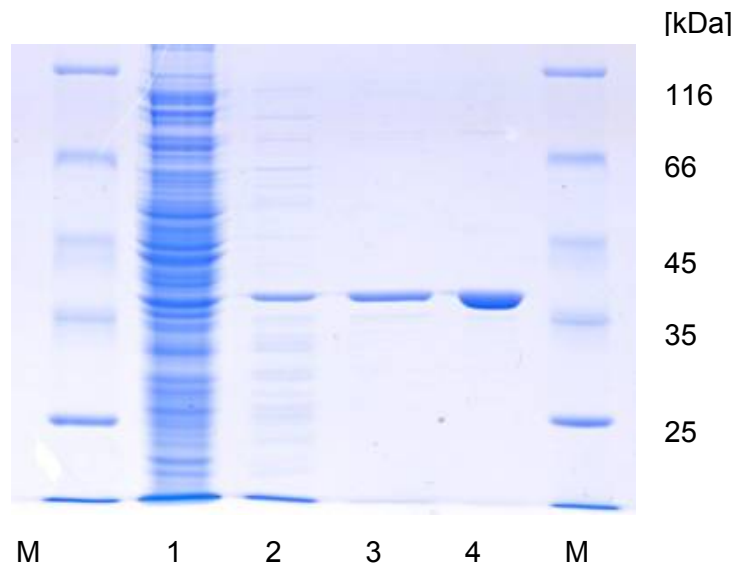


Abb. 20: SDS-Gel des Aufreinigungsverlaufs der Cellulase CelA

Spuren: M *Protein molecular weight marker* (Fermentas), 1 Rohextrakt, 2 Pool Source Q30 (3 μg), 2 Pool HIC Phenyl-Sepharose (2 μg), 4 Pool Superdex 200 (2 μg)

4.2 Aufreinigung von CelA mittels hydrophober Interaktionschromatographie (HIC)

Bei diesem Aufreinigungsprinzip wird sich die Bindung von hydrophoben Resten an die Säulenmatrix zunutze gemacht. Die gelösten Salze verringern die Löslichkeit des Proteins und verstärken so die hydrophoben Interaktionen des Proteins mit dem Säulenmaterial.

Zur Vorbereitung wurden die vereinigten Fraktionen des vorangegangenen Säulenlaufs gegen Hochsalzpuffer dialysiert (50 mM K_2HPO_4 , 1 M NH_4SO_4 pH 6.5). Tabelle 12 listet die Einzelheiten des HIC-Phenyl-Sepharose Laufes auf.

Tabelle 12: Parameter des HIC Phenyl-Sepharose Laufes

Medium:	HIC Phenyl Sepharose
Säule:	XK 16/10
Säulenvolumen:	20,11 ml
Probe:	47,8 mg
Puffer A:	50 mM K_2HPO_4 , 1M NH_4SO_4 pH 6.5

Puffer B:	50 mM K_2HPO_4
Flussrate:	1,0 ml/min
Fraktionen	4 ml
Äquilibration:	2 SV Puffer B, 6 SV Puffer A
Gradient:	0 – 100 % Puffer B (5 SV)

Die Elution erfolgte mittels eines linear, absteigenden Salzgradienten, wobei ein Protein umso eher eluiert je weniger hydrophob es wäre. Wieder wurden alle Fraktionen gesammelt und für jede sowohl der Protein-Gehalt als auch die Cellulase Aktivität bestimmt.

Wie das Chromatogramm des Säulenlaufs zeigt, stimmten die Fraktionen mit der höchsten Enzymaktivität mit denen des größten Proteingehaltes überein (Abbildung 21). Die Fraktionen 30 bis 34 wurden vereinigt und ein Proteingehalt von 0,35 mg/ml und eine spez. Aktivität von 102,1 U/mg ermittelt (S. 36, Tabelle 13).

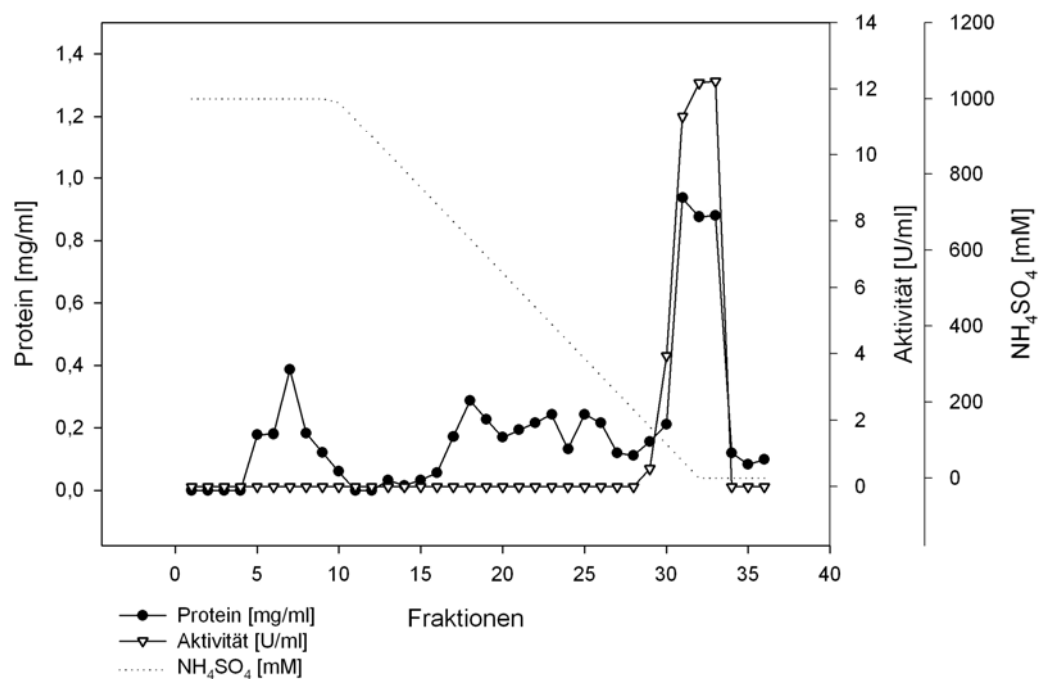


Abb. 21: Chromatogramm des HIC-Phenyl-Sepharose Säulenlaufs zur Aufreinigung des rekombinant hergestellten Celia Proteins.

4.3 Aufreinigung von CclA mittels Gelfiltration

Bei dieser nicht adsorptiven Methode erfolgt die Trennung der Proteine aufgrund ihrer Größe. Größeren Molekülen steht ein vergleichsweise kleines Verteilungsvolumen in der jeweiligen Gelmatrix zur Verfügung, so dass sie früher als kleinere Proteine eluieren. Für die Gelfiltration wurde eine Superdex 200 High Load – Säule verwendet (II.9.3.3).

Um eine möglichst gute Auftrennung zu gewährleisten, wurde das aus der hydrophoben Affinitätschromatographie erhaltene Eluat von zuvor 22 ml auf 2 ml eingengt (II.9.1.8).

Die eigentliche Auftrennung erfolgte unter Verwendung von 50 mM K_2HPO_4 Puffer pH 6,5 und einer konstanten Flußrate von 1 ml/min. Wieder wurden alle Fraktionen auf ihren Proteingehalt und Enzymaktivität überprüft und mittels SDS-Gelelektrophorese kontrolliert. Die Fraktionen 13-15 zeigten die höchste Aktivität und den höchsten Proteingehalt. Sie wurden vereinigt und bei 4 °C bis zur weiteren Verwendung gelagert.

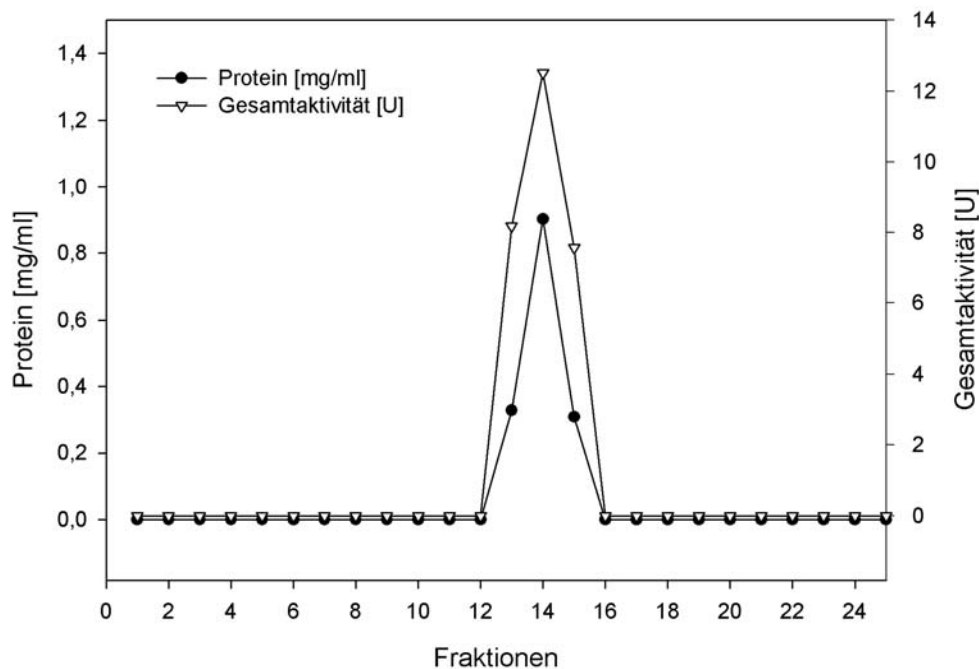


Abb. 22: Chromatogramm des Superdex 200 Säulenlaufs zur Aufreinigung des rekombinant hergestellten CclA Proteins.

Tabelle 13 gibt eine Zusammenfassung der einzelnen Aufreinigungsschritte wieder. Durch die drei Stufen konnte insgesamt eine 25-fache Aufreinigung erzielt werden. Der

Proteingehalt der Elutionsfraktionen des Superdex 200 Säulenlaufes betrug 0,4 mg/ml bei einer spezifischen Enzymaktivität von 388,3 U/mg.

Tabelle 13: Aufreinigungsprotokoll von CelA aus *E. coli* pSVpETCEL-1

Reinigungsschritt	Volumen [ml]	Protein [mg]	Aktivität [U]	Ausbeute [%]	Faktor
Rohextrakt	21	968	14952	100	1
Source Q30	20	47,8	5580	37	2
Phenylsepharose	22	7,7	770	5	6,6
Superdex 200	4	1,6	624	4	25

5. Biochemische Charakterisierung von CelA

5.1 Bestimmung des Temperaturspektrums

Zur Ermittlung der cellulolytischen Aktivität von CelA in Abhängigkeit von der Temperatur wurde der DNSA Standardtest (II.9.5.1) mit CMC als Substrat bei Inkubationstemperaturen von 23 °C bis 95 °C durchgeführt (II.9.6.2).

Die Auftragung erfolgte als relative Aktivität, wobei der höchste Messwert gleich 100 % gesetzt wurde.

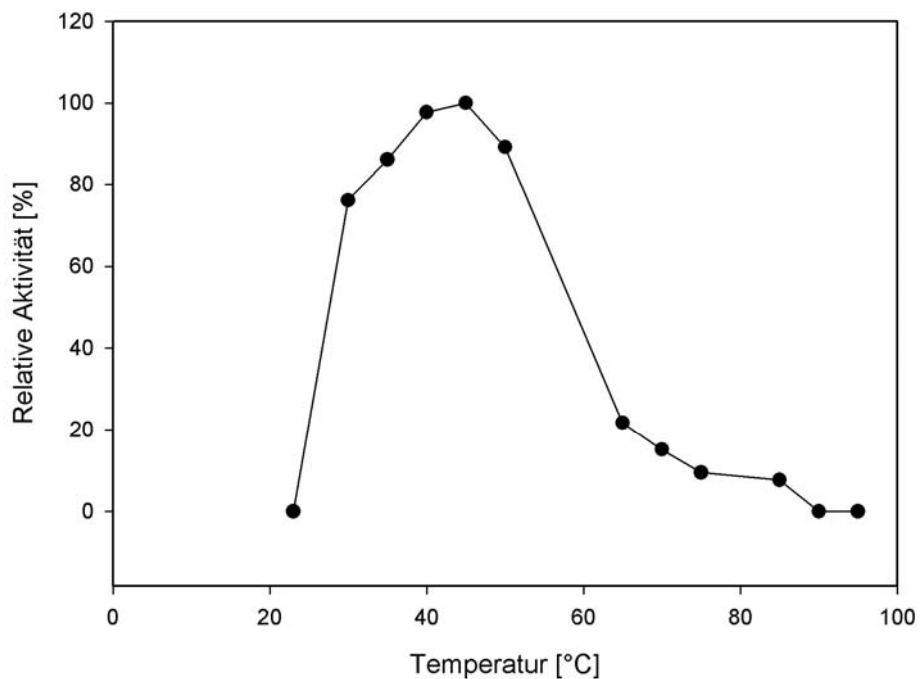


Abb. 23: Bestimmung des Temperaturspektrums der Cellulase CelA

Die maximale Aktivität wurde bei 45 °C verzeichnet. CelA besitzt jedoch ein breites Temperatur Spektrum, wobei 76 % der höchsten Aktivität bei 30 °C und 89 % bei 50 °C gemessenen wurden.

5.2 Bestimmung der Thermostabilität

Die Thermostabilität von CelA wurde für den Temperaturbereich von 40 °C bis 60 °C bestimmt. Es erfolgte eine Inkubation des Enzyms bei den angegebenen Temperaturen. Nach der Probennahme fand die Ermittlung der Enzymaktivität, wie

schon bei der Messung zum Temperaturoptimum, mittels des DNSA-Standardtests (II.9.5.1) statt.

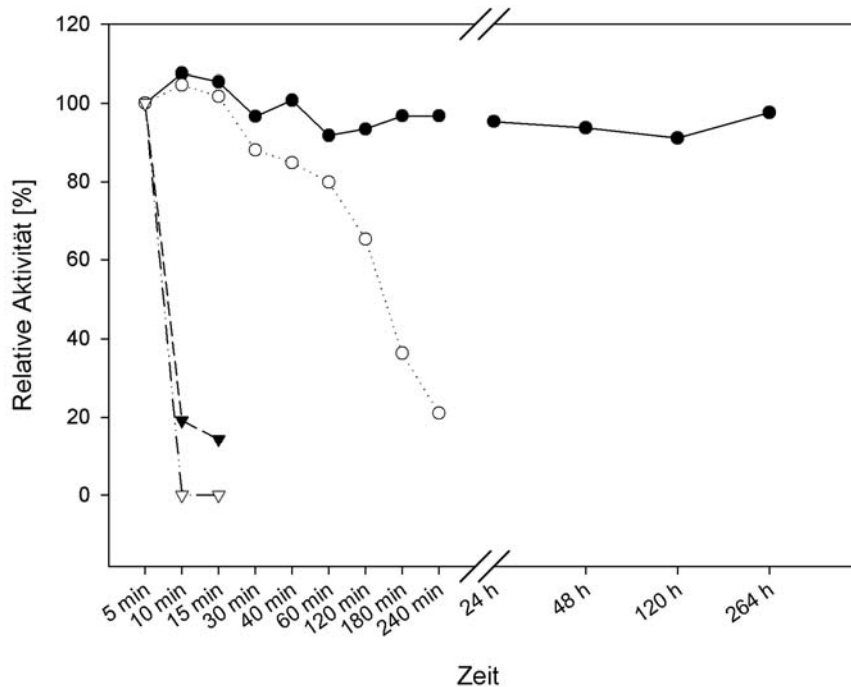


Abb. 24: Bestimmung der Temperaturstabilität von CelA

Die Aktivität wurde über einen Zeitraum von bis zu 264 Std. (=11 Tagen) mittels DNSA Standardtest gemessen. Die Temperaturen betragen 40°C (●), 50°C (○), 55°C (▼) und 60°C(▽).

Bei einer Vorinkubation des Enzyms bei 60 °C zeigte sich bereits nach 10 min eine komplette Inaktivierung von CelA (Abbildung 24). Wurde das Enzym bei 55 °C inkubiert, war nach 10 min nur noch eine Aktivität von 20 % messbar.

Die Inkubation bei 50 °C zeigte hingegen ein ganz anderes Bild: Nach 10 min war sogar ein leichter Anstieg der Aktivität zu verzeichnen. Ein Abfall konnte erst nach 30 min Inkubation festgestellt werden, jedoch lag die Aktivität hier immer noch bei 88 % der maximalen Aktivität. Noch nach 180 min (= 3 h) waren 36 % der Aktivität feststellbar.

Es stellte sich heraus, daß das Enzym bei 40 °C äußerst stabil war. Nach 4 h Vorinkubation bei dieser Temperatur zeigte CelA immer noch eine Aktivität von 96,7 % der maximalen Aktivität. Die Messungen wurden über einen Zeitraum von 11 Tagen (= 264 h) durchgeführt. Während dieser Zeit zeigte CelA eine mehr oder minder gleich bleibende Aktivität, die nach 11 Tagen noch 97 % betrug.

5.3 Bestimmung des pH-Spektrums

Zur Ermittlung der cellulolytischen Aktivität von CelA in Abhängigkeit vom pH-Wert wurde der DNSA-Standardtest (II.9.5.1) mit variablem pH-Wert bei 40°C durchgeführt. Die verwendeten Puffer sind in Kapitel II.9.6.1 aufgeführt.

Bei der Auftragung wurden alle Messpunkte prozentual in Relation zur höchsten Aktivität gesetzt.

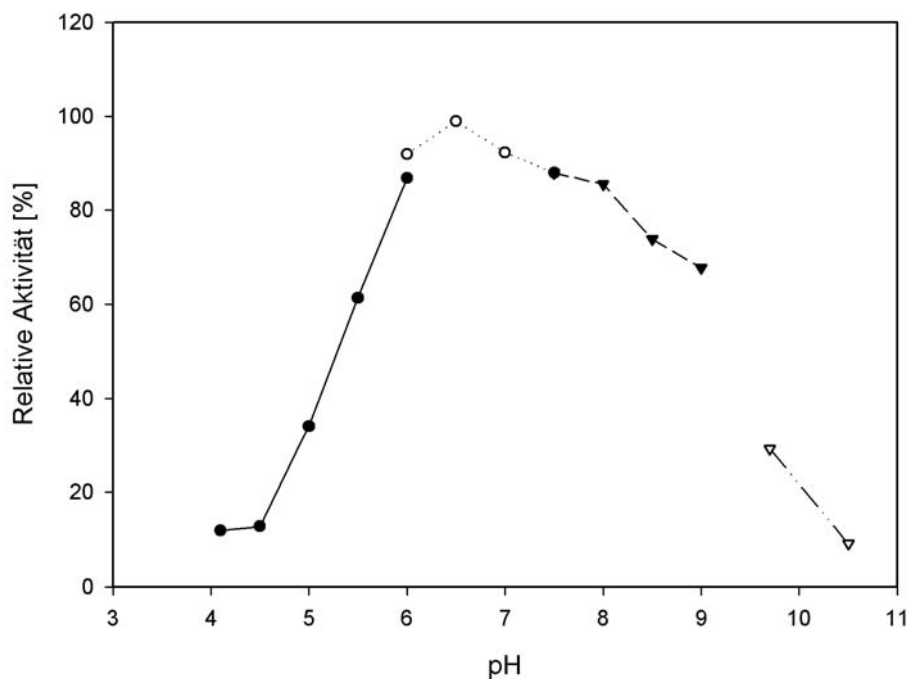


Abb.25: Bestimmung des pH-Spektrums der Cellulase CelA.

Die Enzymaktivität wurde mittels DNSA Standardtest bei pH-Werten von pH 4,1 bis pH 10,5 gemessen. Es wurden folgende Puffer verwendet: Acetat (●), MacIlvaine (○), Tris-HCl (▼) und CAPS (▽).

Es zeigte sich, daß das Enzym über ein breites pH-Spektrum hinweg aktiv war mit einem Aktivitätsmaximum bei pH 6,5. Das Enzym zeigte noch 60 % seiner Maximalaktivität bei einem pH-Wert von 5,5 und 68 % bei pH 9,0. Unterhalb von pH 5,0 und oberhalb von pH 9,7 ist nur noch eine geringe Aktivität zu verzeichnen.

5.4 Abhängigkeit der Cellulase von verschiedenen Additiven

5.4.1 Auswirkungen von einwertigen Metallionen auf die Enzymaktivität

Es wurden die Chlorid-Salze von Na^+ , K^+ , Rb^+ , Li^+ und Ni^+ verwendet. Die Messung der Enzymaktivität fand wieder mit dem DNSA Standardtest unter Verwendung von CMC als Substrat statt (II.9.6.3).

In einem ersten Durchgang wurden die Salze dem Inkubationsansatz direkt hinzugegeben. Die Zugabe des Enzyms erfolgte nach einer Präinkubation des Reaktionsansatzes bei $40\text{ }^\circ\text{C}$ von 10 min wie unter II.9.5.2 beschrieben. Es folgte eine Inkubation von 20 min.

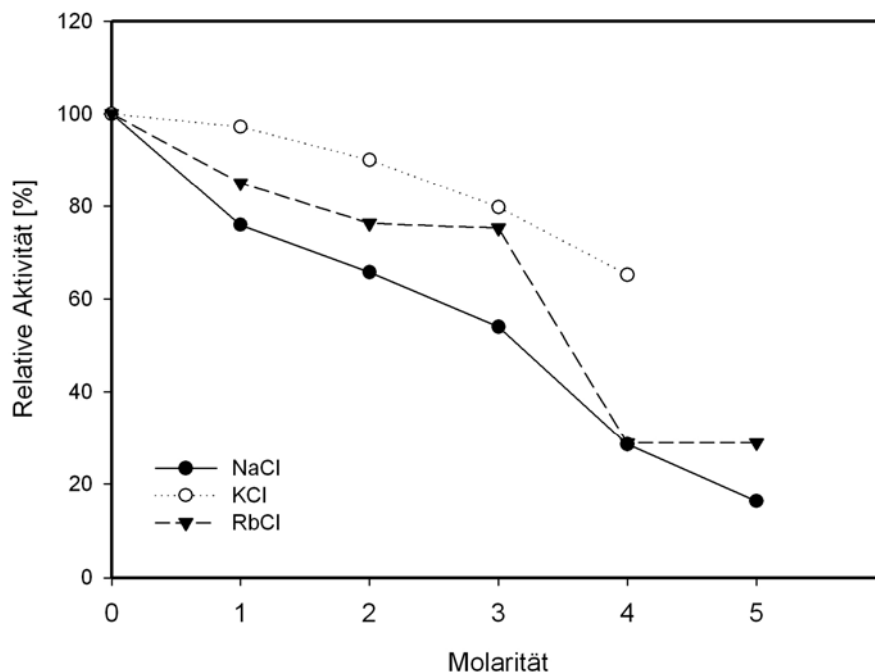


Abb. 26 Auswirkungen einwertiger Salze auf die Enzymaktivität von CelA

CelA zeigte bei Inkubation in 3 M KCl eine relative Aktivität von 79,7 % und bei der Gegenwart von 4 M KCl noch 65,2 % der maximal gemessenen Aktivität (Abbildung 26). Fand der Test zur Messung der Enzymaktivität in der Gegenwart von RbCl statt, so war bei 3 M eine Aktivität von 75,3 % zu verzeichnen und bei 5 M RbCl noch eine relative Aktivität von 28,9 %. Bei Gegenwart von 3 M NaCl zeigte CelA immer noch eine relative Aktivität von 54 %. Erst bei Inkubation in 5 M NaCl sank diese auf 16,4% ab.

Ebenfalls getestet wurde die Auswirkung von LiCl und NiCl auf die Enzymaktivität. In Gegenwart von 1 mM LiCl zeigte CelA noch eine relative Aktivität von 75,4 %. Bei 1 mM NiCl sogar von 95,6 %.

Aufgrund dieser Ergebnisse fand eine Untersuchung der Auswirkungen der einzelnen Salze über einen längeren Zeitraum statt. Hierzu wurde das Enzym, wie bei der Messung zur Temperaturstabilität, bei hohen Salzkonzentrationen in einem Zeitraum von bis zu 20 h inkubiert. Die eingesetzten Salzkonzentrationen betragen 3 M für NaCl und RbCl, sowie 4 M für KCl. Zu definierten Zeitpunkten fand eine Probenentnahme statt, und die Enzymaktivität wurde durch den DNSA-Standardtest, ohne die Gegenwart von Salz, bestimmt.

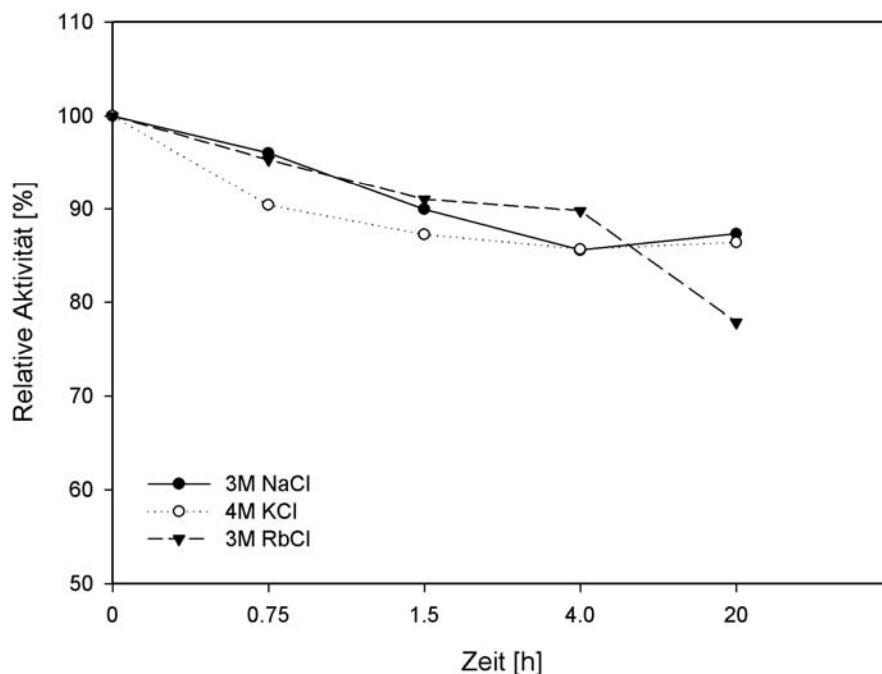


Abb. 27 Bestimmung der Salztoleranz von CelA nach Vorinkubation bei verschiedenen Salzkonzentrationen.

Es zeigte sich, daß CelA immer noch eine hohe Aktivität nach bis zu 20 stündiger Inkubation bei hohen Salzkonzentrationen besaß.

Nach 20 h Vorinkubation bei 3 M NaCl zeigte das Enzym immer noch eine relative Aktivität von 87 %. Die Vorinkubation bei 3 M RbCl für 20 h resultierte in einer relativen Aktivität von immer noch 77 %. Die Auswirkungen von einer Präinkubation in Gegenwart von 4 M KCl auf die Enzymaktivität war ebenfalls gemäßigt. CelA zeigte nach 20 h unter diesen Bedingungen noch eine relative Aktivität von 86 %.

5.4.2 Auswirkungen von mehrwertigen Metallionen auf die Enzymaktivität

Auch bei dieser Untersuchung wurden die Chlorid-Salze für die Aktivitätsbestimmung verwendet. Alle Metallionen wurden in einer Endkonzentration von 1 mM eingesetzt. Nur bei Manganchlorid wurden zusätzlich höhere Endkonzentrationen von 2,5 und 5 mM verwendet. Die Salze wurden dem Inkubationsansatz direkt hinzugegeben. Die Zugabe des Enzyms erfolgte nach einer Präinkubation des Reaktionsansatzes bei 40 °C von 10 min (II.9.5.2). Es folgte eine Inkubation von 20 min.

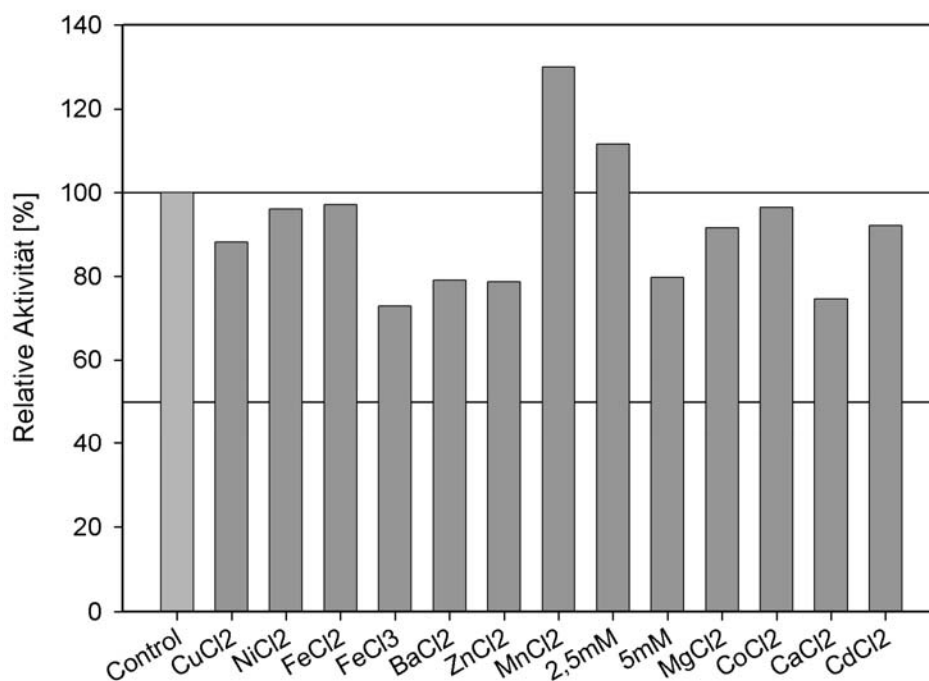


Abb. 28 Auswirkungen mehrwertiger Metallionen auf die Aktivität von CεLA.

Die Enzymaktivität wurde mittels DNSA-Standardtest mit CMC als Substrat gemessen. Die Metalle wurden, wenn nicht anders vermerkt, in 1mM Endkonz. zugegeben.

Die meisten der eingesetzten Salze zeigten keine nennenswerte Auswirkung auf die Enzymaktivität (Abbildung 28. So wurde die Aktivität nur leicht gesenkt durch die Zugabe von Cu²⁺ (relative Aktivität von 88 %), Mg²⁺ (92 %), Cd²⁺ (92 %), Ni²⁺ (96 %), Co²⁺ (96 %) und Fe²⁺ (97 %). Unter Werte von 80 % sank die relative Aktivität nur bei Zugabe von Fe³⁺ (73 %), Ba²⁺ (79 %), Zn²⁺ (78 %) und Ca²⁺ (74 %).

Die Zugabe von 1 mM Mn²⁺ bewirkte hingegen eine Aktivitätszunahme um 30 %. Wurde die Konzentration auf 2,5 mM angehoben zeigte sich immer noch eine

Zunahme von 12 %. Bei einer Endkonzentration von 5 mM Mn^{2+} war ein Rückgang der Enzymaktivität auf 80 % der maximalen Aktivität zu beobachten.

5.4.3 Einfluss weiterer Zusätze auf die Aktivität

Es wurden die Auswirkungen verschiedener Lösungsmittel sowie EDTA auf die Enzymaktivität von C_{el}A getestet. Der Test wurde wie unter 3.4.4.1 mit CMC als Substrat durchgeführt. Die eingesetzten Endkonzentrationen sind der Abb. 29 zu entnehmen.

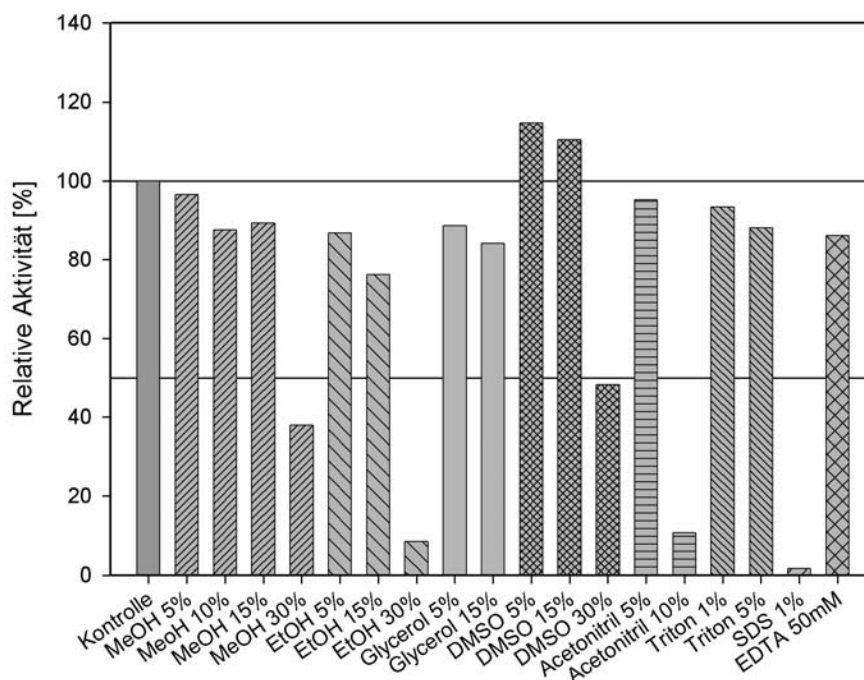


Abb. 29: Auswirkungen von Lösungsmitteln und EDTA auf C_{el}A.

C_{el}A zeigte sich äußerst stabil gegenüber den Lösungsmitteln Methanol und Ethanol (Abbildung 29). Diese zeigten erst bei höheren Konzentrationen signifikante Auswirkungen auf die Aktivität von C_{el}A. Die Gegenwart von 30 % Methanol resultierte in einer Aktivitätsabnahme auf 38 %, die von 30 % Ethanol auf 8 %. Glycerin zeigte keine großen Auswirkungen wohin gegen die Gegenwart von 5 % bzw. 15 % DMSO leicht stimulierend auf die Enzymaktivität wirkten (115 % bzw. 110 %). Eine Erhöhung der DMSO Konzentration auf 30 % resultierte hingegen in einen Abfall auf 48 % der Kontroll-Aktivität.

C_{el}A wurde stark inhibiert durch die Zugabe von 10 % Acetonitril und zeigte nur noch 11 % seiner normalen Aktivität.

Die Zugabe von bis zu 5 % Triton hatte wiederum keine signifikanten Auswirkungen auf die Aktivität (89 %), jedoch führte die Anwesenheit von 1 % SDS zu einem kompletten Verlust der Enzymaktivität.

Zusätzlich zeigte sich CelA relativ stabil gegenüber dem Zusatz von EDTA, zeigte es doch bei einer Konzentration von 50 mM noch eine relative Aktivität von 86 %.

5.5 Substratspektrum der Cellulase CelA

In den Aktivitätstest zur Identifizierung von CelA auf Agarplatten wurde Carboxymethylcellulose (CMC), ein Cellulosederivat, eingesetzt. Um einen Überblick über die Substratspezifitäten von CelA zu erhalten, wurden Aktivitätstests mit verschiedenen Oligosacchariden durchgeführt. Es wurde, neben CMC, zwei weitere β -1,4-verknüpfte Glukane (β -Glukan (Gerste), Lichenan) eingesetzt. Darüber hinaus wurden sowohl β -1,3-verknüpftes Glukan (Laminarin) als auch mikrokristalline Cellulose (Avicel) und β -1,4-verknüpfte Xylose (Hafer-Spelzen-Xylan) getestet. Die Enzymaktivität wurde mittels des unter II.9.5.2 beschriebenen DNSA-Tests zur Bestimmung der Menge an reduzierenden Enden bestimmt.

Tabelle 14: Substratspezifitäten von CelA.

Substrat	*spez. Aktivität [U/mg]
β -Glukan (Gerste)	710
CMC	68
Lichenan	401
Laminarin	0
Hafer-Spelzen-Xylan	0
Avicel	0

* 1 U entspricht der Enzymaktivität, die in 1 min zur Freisetzung von 1 μ mol reduzierenden Enden führt.

Die Tests zeigten, daß CelA eine Spezifität für lösliche β -1,4-verknüpfte Glukane besitzt (Tabelle 14). Die höchsten spezifischen Aktivitäten wurden hierbei mit dem linearen und unverzweigten β -Glukan der Gerste gemessen (710 U/mg). Es folgte das ebenfalls unverzweigte und lineare Glukan Lichenan mit 401 U/mg. Es besitzt gegenüber dem β -Glukan der Gerste einen höheren Anteil an β -1,3- zu β -1,4-Verknüpfungen (2:1). Mit Carboxymethylcellulose (CMC), ebenfalls β -1,4-verknüpft,

wurde eine spezifische Aktivität von 68 U/mg gemessen. Mit dem β -1,3-verknüpften Glukan Laminarin konnte hingegen keine Enzymaktivität festgestellt werden. Ebenso wenig mit β -1,4-verknüpfter Xylose, dem Xylan, oder mikrokristalliner Cellulose (Avicel).

5.5.1 Qualitative Untersuchung des Substratspektrums und des Spaltungsmechanismus mittels Dünnschichtchromatographie

Eine Untersuchung der Spaltungsprodukte mittels Dünnschichtchromatographie sollte klären, ob es sich bei CelA um eine Endo- oder Exo-Cellulase handelt. Endoglukanasen schneiden nach dem Zufallsprinzip innerhalb der amorphen Bereiche der Cellulosekette und setzen so Oligosaccharide unterschiedlicher Länge frei. Die Exoglukanasen hingegen wirken an den Enden der Cellulosekette und spalten Glukose (Cellodextrinasen) oder Cellobiose (Cellobiohydrolasen) ab. Die verwendeten Laufmittel und die genaue Vorgehensweise sind unter II.9.5.3 erläutert.

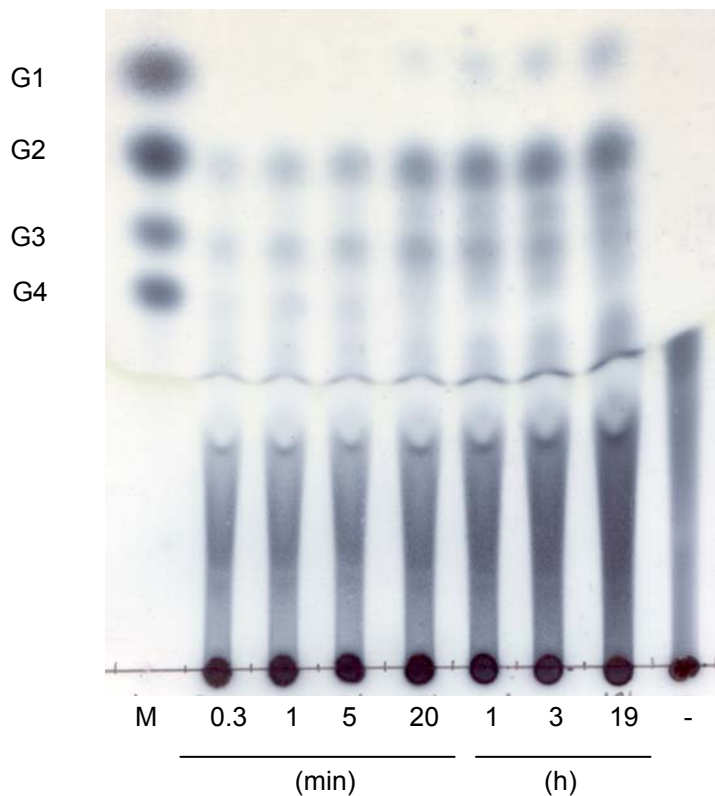


Abb. 30: Dünnschichtchromatographische Analyse des zeitlichen Abbaus von CMC durch CelA

Unter den Spuren ist der Zeitpunkt der Probennahme vermerkt. Spur (-): Kontrolle ohne Enzym; Spur (M): Marker bestehend aus je 0,1 % Glukose (G1), Cellobiose (G2), Cellotriose (G3) u. Cellotetraose (G4). Als Laufmittel diente ein Gemisch aus 1-Propanol, Ethylacetat und H₂O (6:1:3).

Der zeitliche Abbau von CMC zeigte, daß in der Anfangsphase nur Produkte mit höheren Polymerisationsgraden entstanden (Abbildung 30). Dies weist bereits auf einen Endo-Mechanismus des Enzyms hin. Das Hauptprodukt dieses Verdaus ist Cellobiose und zu kleineren Teilen auch Glukose.

Die Analyse der Spaltungsprodukte eines zeitlichen Abbaus von Lichenan unterstützten diese Vermutung, da auch hier am Anfang nur größere Produkte zu sehen waren (Abbildung 31). Das Hauptprodukt dieses Verdaus hatte einen Polymerisationsgrad (DP, *degree of polymerisation*) von drei, es entstanden jedoch auch kleinere Mengen an Produkten mit niedrigerem (DP 2) und höherem (DP 4) Polymerisationsgrad.

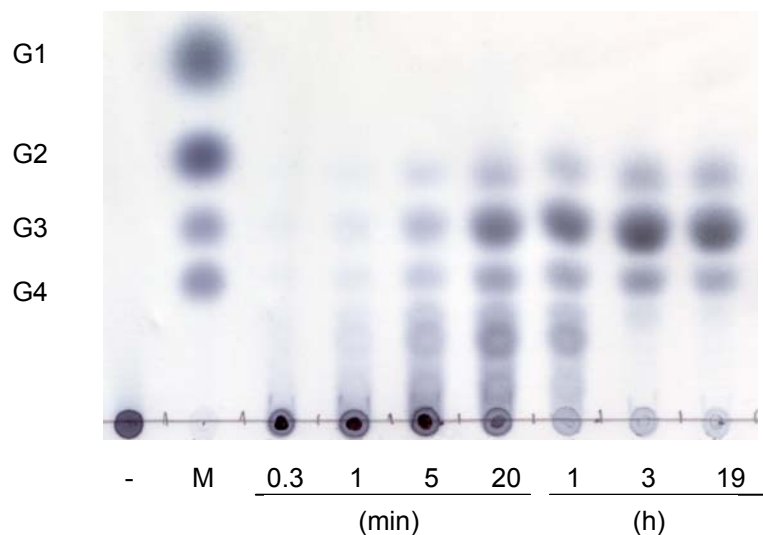


Abb. 31: Dünnschichtchromatographische Analyse des zeitlichen Abbaus von Lichenan durch CelA

Unter den Spuren ist der Zeitpunkt der Probennahme vermerkt. Spur (-): Kontrolle ohne Enzym; Spur (M): Marker bestehend aus je 0,1 % Glukose (G1), Cellobiose (G2), Cellotriose (G3) u. Cellotetraose (G4). Als Laufmittel diente Ethylacetat, Eisessig, H₂O (2:1:1).

Um den Endo-Mechanismus endgültig zu bestätigen, wurde als letztes ein Verdaus mit verschiedenen Cellooligosacchariden durchgeführt und die Spaltungsprodukte analysiert (Abbildung 32). Für CelA konnte die Aktivität einer Endoglukanase bestätigt werden, zeigte doch die dünnschichtchromatographische Analyse, daß aus Cellopentaose sowohl Cellobiose als auch Glukose entstehen. Als Konsequenz ist Cellobiose auch das Abbauprodukt aus der Spaltung von Cellotetraose. Die Analyse zeigte auch, daß CelA in der Lage ist, Cellotriose zu verwerten, jedoch nicht Cellobiose.

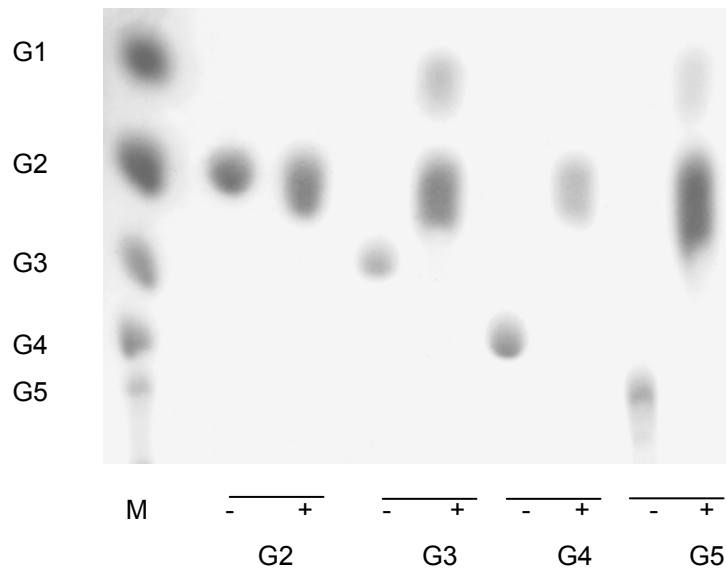


Abb. 32: Dünnschichtchromatographische Analyse der vollständigen Hydrolyse von Cellooligosacchariden durch CelA

(-): Kontrolle ohne Enzym; (+): Enzymzugabe; (M): Marker bestehend aus je 0,1 % Glukose (G1), Cellobiose (G2), Cellotriose (G3), Cellotetraose (G4) u. Cellopentaose (G5). Als Laufmittel diente 1-Propanol, Nitromethan, H₂O (5:3:2).

IV. Diskussion

Mit der ständigen Verbesserung der molekularen Methoden entwickelte sich auch ein neues Verständnis für die mikrobielle Diversität und deren Nutzung für biotechnologische Anwendungen. Standardmethoden, die auf das Screening von Einzelisolaten nach interessanten Enzymen basierten, unterlagen der Problematik, dass nur ca. 1 % aller Mikroorganismen mit bekannten Methoden kultivierbar waren (Torsvik and Ovreas, 2002). Umweltgenbanken hingegen boten die Möglichkeit sich unabhängig von der Kultivierbarkeit, das genetische Potential der bakteriellen Diversität eines Standortes anzueignen (Handelsman, 2004). Diese Gesamtheit der Genome aller Mikroorganismen wird als Metagenom bezeichnet (Handelsman et al., 1998) und gilt als nahezu unerschöpfliche Quelle neuartiger Enzyme (Streit and Schmitz, 2004).

Im Zuge der vorliegenden Arbeit erfolgte eine umfassende Analyse eines angereicherten Boden-Konsortiums. Mit aktivitäts- und sequenzbasierenden Techniken sollte die erstellte Umwelt-Genbank nach agarolytischer und cellulolytischer Aktivität durchmustert, sowie anschließend eine aus dem Konsortium isolierte Cellulase einer ausführlichen biochemischen Charakterisierung unterzogen werden. Zudem sollte die Diversität des angereicherten Konsortiums durch eine 16S rRNA-Genanalyse ermittelt werden.

1. Phylogenetische Analyse des angereicherten Boden-Konsortiums

Ausschlaggebend für die rasche Entwicklung der Metagenomik war die Idee mit Hilfe von molekularbiologischen Methoden phylogenetische Untersuchungen durchzuführen (Pace et al., 1985). Die meisten der Untersuchungen nutzen die kleine Untereinheit (16S) der ribosomalen RNA (rRNA). Dieses Molekül besitzt sowohl variable, als auch konservierte Bereiche, die die Konstruktion universeller Oligonukleotide erlauben. Bereits die erste Publikation 1987, die sich auf dieses Verfahren stützte sorgte für Aufsehen, da sich phänotypisch vollkommen unterschiedliche Organismen aufgrund der Analyse ihrer 16S rRNA-Gene als eng verwandt erwiesen oder sogar zu einem Phylum zusammengefasst wurden (Woese, 1987). Seit dieser Veröffentlichung, die 12 unterschiedliche bakterielle Phyla von kultivierten Organismen beinhaltete, hat sich die Anzahl der 16S rDNA-Sequenzen vervielfacht und neuere Studien gehen mittlerweile von mindestens 50 Phyla aus, von denen die Hälfte als Phyla-Kandidaten bezeichnet werden müssen, da es bislang keine kultivierten Vertreter gibt (Hugenholtz et al., 1998; Rappe and Giovannoni, 2003).

Die Datenbank GenBank des „National Center for Biotechnology Information“ (NCBI) listete im Oktober 2005 insgesamt 33 138 16S rDNA-Sequenzen von kultivierten Mikroorganismen und 132 320 Sequenzen von Mikroorganismen, die als unkultiviert, unidentifiziert oder unbekannt bezeichnet werden (Abbildung 1). Das sind 4,5 mal so viele 16S rDNA-Sequenzen wie noch im Jahr 2002 (Rappe and Giovannoni, 2003).

Betrachtet man die Herkunft der 16S rDNA-Sequenzen von bislang unkultivierten Organismen, gab es bis zum Jahr 2000 mehr Studien zu marinen Habitaten oder Süßwasserhabitaten als zu Bodenhabitaten (Rappe and Giovannoni, 2003). Dies lässt sich u.a. auf Probleme bei der DNA-Extraktion aus Bodenproben durch Verunreinigung mit polyphenolischen Substanzen zurückführen, was die Arbeit mit dem Habitat Boden bis Ende der 1990er Jahre erschwerte. Im Jahr 2005 waren die verfügbaren 16S rDNA-Sequenzen aus Sedimenten oder Böden um das 4-fache angestiegen, während die Anzahl der Sequenzen aus aquatischen Habitaten nur um das 2-fache gestiegen waren (Abbildung 1). Der Anstieg der 16S rDNA-Sequenzen aus Bodenhabitaten lässt sich aber auch mit einem gesteigerten Interesse an diesem Habitat erklären. Zum einen wegen seiner landwirtschaftlichen Bedeutung, zum anderen aufgrund der erhöhten mikrobiellen Diversität im Boden. So beträgt die Anzahl an prokaryotischen Zellen im offenen Meer 5×10^5 Zellen pro ml, in einem Gramm Ackerboden jedoch 2×10^9 Zellen (Whitman et al., 1998).

Bei dem hier untersuchten Konsortium sollten gezielt Bakterien mit hydrolytischen und vor allem agarolytischen Fähigkeiten durch die Anreicherung auf Minimalmedium mit Agar als einziger Kohlenstoffquelle selektiert werden. Agar hydrolysierende Bakterien kommen sowohl in marinen und Süßwasserhabitaten, als auch im Boden vor (Zhong et al., 2001). Die meisten kultivierten Vertreter stammen aus der Klasse der *γ-Proteobacteria* mit den Gattungen *Pseudomonas*, *Xanthomonas*, *Microscilla*, *Microbulbifer*, *Vibrio*, *Pseudoalteromonas* und *Aeromonas*. Andere Klassen mit agarolytischen Bakterien sind die *Bacteria* mit der Gattung *Bacillus*, und die Klassen der *Actinobacteria* und *Sphingobacteria*.

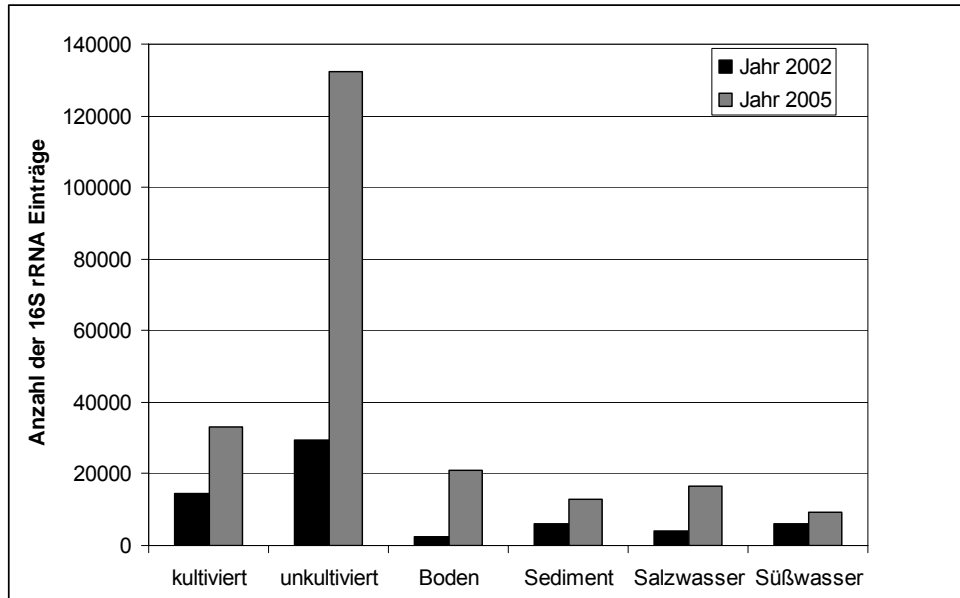


Abb. 1: Anzahl der in GenBank veröffentlichten mikrobiellen 16S rDNA Sequenzen in den Jahren 2002 und 2005.

Die Strategien zur Durchsuchung der GenBank Datenbank sowie die Daten für das Jahr 2002 wurden aus Rappe et al, 2003 entnommen. Wiedergegeben sind die Anzahl an 16S rDNA Sequenzen für als kultiviert oder als unkultiviert, unidentifiziert oder unbekannt gelistete Mikroorganismen. Zusätzlich ist die Anzahl an 16S rDNA Sequenzen der unkultivierten Mikroorganismen dargestellt, die einen Eintrag über die Herkunft der Organismen enthielten.

Teil der Arbeiten war eine Vereinzelung des Konsortiums und ein Vergleich der Diversität der kultivierbaren Organismen mit den 16S rDNA-Sequenzen des gesamten angereicherten Boden-Konsortiums. Es konnten fünf Einzelisolate durch Kultivierung gewonnen werden. Die Analyse der Gesamt-DNA erbrachte zusätzlich fünf weitere 16S rDNA-Sequenzen. Insgesamt stammten sieben Sequenzen aus dem Phylum der *Proteobacteria*. Hierbei dominierten die γ -*Proteobacteria* mit fünf Sequenzen gefolgt von den α - und β -*Proteobacteria* mit jeweils einer Sequenz. Zwei Sequenzen lassen sich den *Firmicutes* und eine den *Actinobacteria* zuordnen.

Bei der Analyse 16S rDNA-Sequenzen der kultivierbaren fünf Organismen zeigen die Klone Nummer 1 und 4 die höchste Ähnlichkeit zu Vertretern der *Proteobacteria*, die Klone Nummer 3 und 5 lassen sich den *Firmicutes* zuordnen und Klon Nummer 2 den *Actinobacteria*. Dieses Ergebnis deckt sich mit Beobachtungen, wonach die meisten der bisher kultivierten und damit am besten studierten Organismen aus nur vier Phyla stammen: den *Proteobacteria*, den *Actinobacteria*, den *Firmicutes* und den *Cytophagales*, auch *Cytophaga-Bacteroides-Flexibacter*-Gruppe (CFB) genannt (Hugenholtz, 2002).

Es konnten Ähnlichkeiten zu den Spezies *Pseudomonas*, *Xanthomonas*, *Bacillus*, und *Agrobacterium* festgestellt werden. Vertreter dieser Spezies sind für ihr Vorkommen im Boden bekannt (Ovreas and Torsvik, 1998). Ein Vertreter der *Bacilli* zeigte die größte Ähnlichkeit zu dem kultivierten Organismus *Bacillus* sp. MK03. Aus diesem Organismus konnten bereits eine Hydrolase und eine Agarase isoliert und charakterisiert werden (Suzuki et al., 2002; Suzuki et al., 2003). Drei der 16S rDNA-Sequenzen der Gesamt-DNA und das Isolat Nummer 4 zeigten die größten Übereinstimmungen zu bisher unkultivierten Bakterien.

Das Habitat Boden ist äußerst komplex und die unterschiedliche Bodenbeschaffenheit, -nutzung, Umwelteinflüsse oder die Jahreszeit der Probennahme haben große Auswirkung auf die mikrobielle Zusammensetzung (Smit et al., 2001). Ein direkter Vergleich der in dieser Arbeit vorkommenden bakteriellen Phyla und Klassen mit anderen Studien ist deshalb schwierig. Betrachtet man aber die Ergebnisse der Studien von vergleichbaren Bodenstandorten, sprich landwirtschaftlich genutzten Flächen, so fallen Unterschiede in der Diversität auf: In bisher fast allen auf molekularen Techniken basierenden Studien von Bodenhabitaten und Sedimenten konnten vermehrt Acidobakterien identifiziert werden (Barns et al., 1999; Liles et al., 2003). Ebenso gelten Vertreter des noch jungen Phylums *Verrucomicrobia* als typische Bodenorganismen (Hugenholtz et al., 1998). Die im Zuge dieser Arbeit erhaltenen 16S rDNA-Sequenzen konnten keinem dieser ubiquitär vorkommenden Phyla zugeordnet werden. Beide Phyla haben gemeinsam, dass sie bisher fast ausschließlich aus nicht kultivierten Organismen bestehen und als schwer kultivierbar gelten (Riesenfeld et al., 2004). Die Abwesenheit von Vertretern dieser Phyla in dem hier analysierten Konsortium lässt sich darauf zurück führen, dass es sich um eine Anreicherungskultur handelt, und die Diversität aufgrund der stattfindenden Selektion geringer ausfällt als bei der direkten Isolation der DNA aus einer Bodenprobe. Der für das Phylum namensgebende Organismus *Acidobacterium* ist moderat acidophil, aerob und heterotroph (Hugenholtz et al., 1998). Dementsprechend konnten bisher Vertreter der *Acidobacteria* durch Verwendung eines Mediums isoliert werden, welches neben Xylan als Kohlenstoffquelle einen pH-Wert von 5,5 besaß (Sait et al., 2002). Das in dieser Arbeit verwendete Minimalmedium besaß hingegen einen neutralen pH-Wert und Agar als einzige Kohlenstoffquelle und war somit für die Kultivierung dieser Organismen nicht geeignet.

Wie bereits erwähnt konnten nicht alle Organismen des Konsortiums isoliert werden, es zeigten sich Unterschiede bei dem Vergleich mit den gewonnenen phylogenetischen Daten der Gesamt-DNA. Diese Unterschiede in der Diversität von

Isolaten und der Gesamtheit der bakteriellen Population einer Bodenprobe wurde auch in anderen Studien beobachtet. So wurden bei der Untersuchung zweier Bodenproben von landwirtschaftlich genutzten Flächen die kultivierbaren Bakterien mit der Gesamtheit der bakteriellen Population verglichen ((Ovreas and Torsvik , 1998)). Es zeigte sich, dass die Diversität der Gesamtheit der Bakterien höher war, als innerhalb der kultivierbaren Organismen. Die Mehrheit der über 200 Isolate stammte aus den Phyla der *Proteobacteria*, *Cytophagales* und den *Actinobacteria*. Wobei die Reihenfolge die Gewichtung der unterschiedlichen Phyla wiedergibt.

Bei einer weiteren Studie (McCaig et al., 1999; McCaig et al., 2001), in der ebenfalls die Daten der phylogenetischen Analyse der Gesamt-DNA mit denen der kultivierbaren Organismen verglichen wurde, zeigte sich, dass die Ergebnisse hinsichtlich Diversität und Gewichtung der einzelnen Phyla völlig unterschiedlich waren. Als Isolationsmedium wurde ein Vollmedium mit Trypton verwendet. Es gehörten 40 % der Klone der Analyse der Gesamt-DNA zu den α -*Proteobacteria*, jedoch nur 9,5 % der Isolate stammten aus diesem Phylum. Die γ -*Proteobacteria* waren in der molekularen Studie mit nur 3,7 % vertreten, machten jedoch 24 % der Isolate aus. Diese Beobachtung konnte auch in dieser Arbeit gemacht werden: Keines der fünf Einzelisolate gehörte zu den α -*Proteobacteria* aber 25 % der Einzelisolate gehörten zu den γ -*Proteobacteria*.

Mitsui und Kollegen analysierten die Inkubationszeit und die benötigten Medien-Bestandteile für unterschiedliche phylogenetische Gruppen (Mitsui et al., 1997). Sie zeigten, dass Bakterien-Stämme der α -*Proteobacteria* oligotroph sind, d.h. sie sind in der Lage auf nährstoffarmen Medien (0,5 bis 15 mg/l Kohlenstoff) zu wachsen. Zudem wachsen sie sehr langsam. Demnach scheint das im Zuge dieser Arbeit für die Vereinzelung des Konsortiums verwendete nährstoffreiche LB-Medium nicht für das Wachstum und die Isolation von Vertretern der α -*Proteobacteria* geeignet zu sein. Obwohl es kein selektives Medium ist, scheinen schnell wachsende, aerobe Bakterien mit neutralem pH-Optima bevorzugt zu werden. Diese Vermutung wird auch gestützt durch die phylogenetische Untersuchung der Gesamt-DNA des Konsortiums. Hier konnte eine der 16S rDNA Sequenzen den α -*Proteobacteria* zugeordnet werden. Diese Beobachtung konnte auch in einer anderen Studie gemacht werden, bei der neue Organismen isoliert wurden (Hamaki et al., 2005). Es wurde ein Minimal-Medium verwendet, welches durch aufschwemmen und Zentrifugation gewonnenen Boden-Extrakt zusätzlich nur Agar enthielt. Die daraus isolierten Organismen gehörten zu den α -*Proteobacteria* und zeigten kein Wachstum auf verschiedenen Vollmedien.

In den vergangenen Jahren schien die einzige Möglichkeit Zugang zu bisher unbekanntem bakteriellen Spezies zu erhalten auf molekularen Techniken zu basieren. Klassische Anreicherungen und Kultivierungen galten aufgrund der künstlichen und sehr restriktiven Bedingungen als inadäquat und wurden zugunsten molekularer Techniken, wie die 16S rDNA-Sequenzanalyse von Gesamt-DNA vernachlässigt. Zwar revolutionierten diese Techniken unser Verständnis von der bakteriellen Diversität und erweiterten den bisher bekannten Stammbaum in ungeahntem Ausmaß, jedoch repräsentiert das 16S rRNA Gen nur 0,05 % der durchschnittlichen Genomgröße einer prokaryotischen Zelle (Rodriguez-Valera, 2002). Es ist damit nicht geeignet die neu auftretenden Fragen hinsichtlich Physiologie, der ökologischen Funktion und des biotechnologischen Potentials einer neuen Spezies zu beantworten. Viele Studien zeigten, dass Spezies mit über 97 % Übereinstimmung auf Ebene der 16S rRNA-Gene unterschiedliche physiologische und ökologische Eigenschaften aufwiesen (Achenbach and Coates, 2000). Zusätzlich besteht die Hälfte der bisher bekannten Phyla nur aus unkultivierten Vertretern und bei neun weiteren Phyla mit kultivierten Vertretern sind keine Genomdaten verfügbar (Hugenholtz, 2002). Im Anbetracht dieser Tatsachen erscheinen klassische Kultivierungsmethoden in einem neuen Licht. Diverse Studien zeigten, dass eine Vielzahl von bisher unkultivierten Bakterien mit einfachen Methoden zu isolieren sind (Joseph et al., 2003; McCaig et al., 2001; Sait et al., 2002). Diese Beobachtungen decken sich mit den Ergebnissen dieser Arbeit, da 50 % der Einzelisolate signifikante Ähnlichkeiten von 99 % zu bisher nicht kultivierten Organismen zeigten.

Mittels der phylogenetischen Analyse unter Kombination klassischer Kultivierungsmethoden und molekularer Techniken konnte ein Einblick in die Diversität eines angereicherten Bodenkonsortiums und der für die agarolytische Fähigkeit verantwortlichen Organismen gewonnen und der sinnvolle Einsatz dieser Methoden zur Isolation bisher unkultivierter Organismen bestätigt werden.

2. Anreicherungskulturen als verlässliche Quelle neuer Biokatalysatoren

Die am häufigsten angewendete Technik zur Identifizierung neuer Biokatalysatoren und bioaktiver Substanzen ist die aktivitätsbasierende Durchmusterung, bei der z.B. durch Verwendung von Plattentests die Standort-Genbank nach der gewünschten Enzymaktivität durchsucht wird. Im Gegensatz zu sequenzbasierenden Verfahren bietet diese Methode die Möglichkeit Gene neuer, unbekannter Enzymklassen zu identifizieren. Mit fortschreitender Etablierung und Anwendung dieser Methoden wurde aber auch die damit einhergehende Problematik sichtbar: Da ein Gramm Erde schätzungsweise zwischen 2 000 bis 18 000 unterschiedliche Spezies enthält, wird angenommen, dass eine Umweltbank, die dieses Metagenom repräsentieren soll aus über 10^7 Plasmid-Klonen (mit 5 kb Insert) oder 10^6 BAC-Klonen (mit 100 kb Insert) bestehen müsste (Handelsman et al., 1998). Da dies ein aussichtsloses Unterfangen darstellt scheint es sinnvoller zu sein die Anzahl der Spezies auf diejenigen mit der gewünschten Aktivität einzuschränken, um so mit einer Genbank von geringerem Umfang das betreffende Metagenom abzudecken. Dies würde auch ein weiteres Problem der Umweltgenomik umgehen: Die niedrige Frequenz an positiven Klonen innerhalb einer Genbank, deren DNA durch direkte Isolation aus der Umweltprobe gewonnen wurde. So konnte in einer Studie bei der Durchmusterung von 1 186 200 Klonen nur zehn Antibiotikaresistenz vermittelnde Klone identifiziert werden (Riesenfeld et al., 2004). Bei einer anderen Studie konnte aus einer Umweltbank von 730 000 Klonen bei der Durchmusterung nach Lipaseaktivität nur ein positiver Klon isoliert werden (Henne et al., 2000).

Geht man von einer idealen Genbank aus, bei der eine gleichmäßige Verteilung aller Spezies herrscht und keine inaktiven Klone vorhanden sind, so hängt die Anzahl der für einen positiven Treffer zu durchmusternden Klone von der Frequenz der Organismen ab, die das gesuchte Gen tragen (Gabor et al., 2004). Diese Frequenz kann durch einen Anreicherungsschritt vor der DNA-Isolation erhöht werden, da unter dem Selektionsdruck Organismen mit den gewünschten Eigenschaften bevorzugt werden (Streit et al., 2004). Natürlich beinhaltet die Anreicherung auch immer einen Abfall der Diversität in der untersuchten Probe, da nicht alle Organismen unter den gegebenen Laborbedingungen kultivierbar sind. Jedoch besteht bei einer gemischten Anreicherungskultur die Möglichkeit Organismen zu kultivieren, die einzeln niemals isolierbar wären, da sie in der Natur als Symbionten oder Teil eines Konsortiums vorkommen (Schink, 2002). Am weitesten verbreitet ist der Einsatz unterschiedlicher

Stickstoff- und Kohlenstoffquellen im Anreicherungsmedium, um einen Selektionsdruck aufzubauen.

Für die Anreicherung des in dieser Arbeit untersuchten Konsortiums wurde ein Minimalmedium mit Agar als einziger Kohlenstoffquelle verwendet, um agarolytische Organismen zu selektieren. Bei der anschließenden Durchmusterung der 1 700 Klone umfassenden Cosmid-Genbank auf agarolytische Aktivität konnten vier positive Klone isoliert werden, auf denen sich insgesamt zwölf Gene befanden, die für Agarasen kodieren (III. Ergebnisse, Tabelle 8). Eine Durchmusterung auf cellulolytische Aktivität der Genbank ergab elf positive Klone, von denen nach einer Restriktionsanalyse acht unterschiedliche übrig blieben (III. Ergebnisse, Tabelle 4). Dieses erhöhte Aufkommen an positiven Klonen bei einem aktivitätsbasierenden Screening ist selbst im Vergleich zu ähnlichen Studien bemerkenswert.

So wurde innerhalb einer Studie zur Identifizierung neuartiger Cellulasen und Esterasen Genbanken aus unterschiedlichen Anreicherungen und eine durch direkte DNA-Isolation erstellte Umwelt-Genbank hinsichtlich ihrer Trefferquote verglichen (Rees et al., 2003). Die Quote der Cellulase-positiven Klone betrug in der Genbank, welche aus einer Anreicherungskultur auf Carboxymethylcellulose (CMC) erstellt wurde, 1:15 000. Im Gegensatz hierzu betrug die Frequenz einer Genbank, welche aus einer Anreicherung auf Minimal-Medium erstellt wurde, nur noch 1:60 000 und die Umwelt-Genbank ergab keinen einzigen Treffer bei 100 000 durchmusterten Klonen.

Bei diesen Angaben ist aber zu beachten, dass es sich bei den verwendeten Genbanken dieser Studie um Plasmid-Genbanken handelt, bei der die Größe des klonierten DNA-Fragmentes bei durchschnittlich 5 kb lag, während die in dieser Arbeit verwendete Cosmid-Genbank eine sechs mal so große Insertgröße von 30 kb hatte. Ein direkter Vergleich ist schwierig, da beide Vektorsysteme sich nicht nur hinsichtlich der möglichen Insertgröße unterscheiden, sondern auch in anderen Punkten wie z.B. der Kopienzahl. Vor dem Anlegen einer Umwelt-Genbank muss dementsprechend das am besten geeignete System ausgewählt werden.

Diese Wahl hängt von mehreren Faktoren ab, wie z.B. der Qualität der isolierten DNA. Direkt aus einer Bodenprobe isolierte DNA ist häufig mit Huminsäuren und Mineralien kontaminiert, was zusätzliche Reinigungsschritte nötig macht. Dies führt vermehrt zu einem Fragmentieren der chromosomalen DNA, die dadurch nicht mehr für das Anlegen von Genbanken mit großen Inserts in z.B. Cosmid- oder BAC (Bacterial Artificial Chromosome)-Vektoren geeignet ist. Der Anreicherungs-schritt vor der Isolation umgeht diese Problematik und die Wahl des Vektorsystems hängt nun vor

allem von der angestrebten Kopienzahl, dem gewählten Wirtsorganismus und natürlich der gesuchten Enzymaktivität ab.

Der Vorteil von *small insert*-Genbanken liegt zum einen in der hohen Kopienzahl der Konstrukte im Wirtsorganismus. Zum anderen wird durch die geringe Länge des Inserts und dem damit verbundenen geringen Abstand zum vektoreigenen Promotor die Wahrscheinlichkeit erhöht, dass das Gen von Interesse transkribiert wird.

Tab. 1: Vergleich der Ergebnisse dieser Arbeit mit aktivitätsbasierenden Studien, die auch Anreicherungskulturen zur Grundlage haben.

Quelle	Vektor	Anzahl der Klone	Ø Insertgröße [kb]	Anzahl positiver Klone	Referenz
Ackerboden	Cosmid	1 700	30	8 Cellulasen, 12 Agarasen, 2 Amylasen, 14 weitere Biokatalysatoren	diese Arbeit
Ackerboden, Waldboden, Pferdedung	Cosmid	6 000 30 000 35 000	35	Biotinsynthese- Operon: 2 Klone/ 3 Klone/ 2 Klone	(Entcheva et al., 2001)
Ackerboden, Fluss- o. marine Sedimente, Salzsee	Plasmid	je 100 000 (400 000)	3,0 – 5,6	Alkohol-Oxidoreduktasen: 16 Klone	(Knietsch et al., 2003)
Ackerboden, Salzsee- und Fluss-Sediment	Plasmid	305 000 300 000 112 000	3,0 – 5,6	Dehydratasen	(Knietsch et al., 2003)
Meer-Sediment/ sandiger B./ Ackerb./ Uferboden	Plasmid	23 000 30 000 35 000 25 000	5,2	Amidasen: 1 Klon/ 2 Klone/ 0 Klone/ 2 Klone	(Gabor et al., 2004)
Seewasser	Plasmid	15 000 60 000	2 -10	Cellulasen: 1 Klon, Esterasen: 2 Klone	(Rees et al., 2003)
Sediment u. Boden eines Natronsees	Plasmid	35 000 37 000	2 - 5,5	Cellulasen: je 1 Klon pro Bank	(Grant et al., 2004)

Diese Vektoren eignen sich dementsprechend vor allem für die Durchmusterung nach einzelnen Genen. Voraussetzung ist jedoch, dass dieses Gen in der richtigen Orientierung zum Promotor kloniert und nicht durch eine transkriptionale Terminationssequenz von diesem getrennt ist (Gabor et al., 2004). Vorteile bringen hier Vektoren wie Cosmide, Fosmide und BACs, welche das Anlegen von *large insert*-Genbanken mit Insertgrößen über 30 kb erlauben. Zwar ist die Kopienzahl gegenüber den Plasmiden stark reduziert und aufgrund der Größe der Inserts ist eine Transkription durch den Vektor eigenen Promotor nicht wahrscheinlich. Jedoch zeigen diverse Studien, wie auch die Ergebnisse dieser Arbeit, dass die geringe Kopienzahl für eine erfolgreiche Durchmusterung ausreicht und die Transkription durch die auf der Umwelt-DNA vorhandenen Promotoren initiiert werden kann (Brady et al., 2004; Courtois et al., 2003; Entcheva et al., 2001; Gillespie et al., 2002; Rondon et al., 2000).

Mikroorganismen produzieren jedoch auch wertvolle Sekundärmetabolite, wie Antibiotika und Antitumorstoffe, deren Gene zusammen mit den zugehörigen Resistenz und regulatorischen Genen häufig geclustert vorkommen (Vining, 1992). Für das Auffinden dieser großen Operons eignen sich daher eher Vektoren, die größere Inserts zulassen. Auf diese Weise konnte in einem aktivitätsbasierten Screening unter Verwendung eines Cosmid Schaukel-Vektors und den heterologen Wirtsorganismen *E. coli* und *S. lividans* acht positive Klone für Polyketidsynthasen des Typs I aus einer 5 000 Klone umfassenden Genbank isoliert werden (Courtois et al., 2003). Ein weiteres erfolgreiches Beispiel ist die Isolierung sieben positiver Klone mit kompletten Biotinsynthese-Operons aus mehreren Cosmid-Genbanken unterschiedlicher Herkunft (Entcheva et al., 2001).

Eine zweite Möglichkeit sich das genetische Potential eines Standortes zugänglich zu machen bietet die sequenzbasierende Analyse. In Kombination mit der Sequenzierung einer *small insert*-Genbank wurde diese Methode für das Studium eines Modell-Biofilms aus Trinkwassersystemen verwendet (Schmeisser et al., 2003). Durch Analyse dieser nach dem Zufallsprinzip ausgewählten Cosmid-Klone konnten fünf ORFs mit potentiell biotechnologischen Nutzen identifiziert werden. Zwei der ORFs kodierten für putative Lipasen und drei für Amylasen.

Im Zuge dieser Arbeit erfolgte eine komplette Sequenzierung der vier Agarase-positiven und fünf weiterer, nach dem Zufallsprinzip ausgewählten, Cosmide. Ebenfalls wurde ein 9,7 kb großes Fragment (Klon pSVCEL-1) eines der acht Cellulase-positiven Cosmide sequenziert. Auf diese Weise wurden insgesamt 243 kb an Sequenzinformation generiert. Durch die Kombination dieser beiden Techniken

konnten insgesamt 36 Gene identifiziert werden, die für neuartige, potentielle Biokatalysatoren kodieren (siehe Tabelle 2).

Der hohe Anteil von 86 % der putativen und funktionalen polysaccharid-modifizierenden Enzymen an den in dieser Arbeit identifizierten Biokatalysatoren lässt sich auf die Anreicherung mittels Agar als einziger Kohlenstoffquelle zurückführen, wodurch Organismen mit hydrolytischen Fähigkeiten selektiert wurden. Ordnet man die ORFs anhand ihrer besten blastp-Treffer bestimmten Spezies zu, so fällt auf, dass 17 der 24 annotierten ORFs die höchsten Übereinstimmungen zu dem γ -Proteobakterium *Saccharophagus degradans* 2-40 zeigen. Dieser 1986 isolierte marine Organismus wurde zunächst anhand erster 16S rDNA Analysen der Spezies *Microbulbifer* innerhalb der Klasse der *Alteromonadacea* zugeordnet (Andrykovich and Marx, 1988; Gonzalez and Weiner, 2000). Aktuelle Studien ergaben jedoch, dass es sich um eine neue Spezies handelt (Ekborg et al., 2005). *S. degradans* 2-40 zeichnet sich durch die Fähigkeit aus eine Vielzahl von komplexen Polysacchariden abzubauen. Hierzu gehören Agar, Alginat, Chitin, Cellulose, Laminarin, Pektin, Pullulan, Stärke und Xylan (Ensor et al., 1999; Howard et al., 2004). Die nahe verwandten Vertreter der Spezies *Microbulbifer* sind ebenfalls bekannt für ihre hydrolytischen Fähigkeiten, wenn auch für keinen Organismus eine derartige hohe Anzahl an polysaccharidmodifizierenden Enzymen bekannt ist (Gonzalez et al., 1997; Tanaka et al., 2003; Yoon et al., 2004; Yoon et al., 2003).

Die sequenzbasierende Identifizierung von vier putativen Endoglukanasen zeigt, dass diese Methode eine sinnvolle Erweiterung zu der funktionellen Durchmusterung einer Umwelt-Genbank darstellt. Denn obwohl zwei der vier ORFs (*gnuB* und ORF 48 des Cosmids VII C10) die höchste Übereinstimmung zu β -1,4-Endoglukanasen (Cellulasen) zeigten, konnten sie nicht bei der Durchmusterung nach cellulolytischer Aktivität auf dem β -1,4-verknüpften Glukan CMC identifiziert werden. Gründe hierfür könnten in den bereits angesprochenen Problemen der Transkription bei „large insert“-Genbanken liegen. Die zwei anderen ORFs (ORF 80 des Cosmids XIV A3, ORF 01 des Subklons pSVCEL-1) zeigten außerdem die höchste Ähnlichkeit zu einer Laminarinase (β -1,3(4)-Endoglukanase) bzw. einer reinen β -1,3-Endoglukanase und hätten somit mit dem für die Durchmusterung verwendeten Substrat nicht detektiert werden können.

Neben Cellulose stellen die Hemicellulosen einen wichtigen Bestandteil der Pflanzenzellwand dar. Zu ihnen gehört neben Mannan, Galaktan und Arabinan auch Xylan, welches auf einem homopolymeren Rückgrat aus β -1,4-verknüpften Xylose-Einheiten basiert (Kulkarni et al., 1999). Im Zuge der Sequenzierung konnte ein ORF identifiziert werden, der große Übereinstimmungen zu einer β -Xylosidase zeigte. Beim

Abbau des Xylans hydrolysieren die Endo- β -1,4-Xylanasen das Xylan-Rückgrat zu Xylobiose und kurzen Xylose-Oligomeren, die anschließend von den β -D-Xylosidasen zu Xylose-Monomeren abgebaut werden (Beg et al., 2001). Xylolytische Enzyme finden vielfache industrielle Anwendungen: Sie werden zumeist in Kombination mit anderen Enzymen, wie Cellulasen, Amylasen und Pektinasen in der Zellstoff- und Papierindustrie eingesetzt. Sie finden aber auch Verwendung in der Lebensmittelindustrie bei der Klärung von Fruchtsäften, als Backenzym zur Verbesserung der Teigeigenschaften (Maschinengängigkeit, Stabilität) und zur Produktoptimierung (Krustenstabilität, Volumen), sowie in der Alkoholproduktion (Beg et al., 2001).

Tab. 2: Übersicht über die in dieser Arbeit identifizierten biotechnologisch interessanten ORFs.

Enzymklassen	Anzahl ORFs	ORF Bezeichnung und Herkunft
Polysaccharid-mod. Enzyme		
Amylasen ¹	2	<i>amyA</i> (VII C10), <i>amyB</i> (III E5)
Cellulasen ¹	8	<i>celA</i> (pSVCosCEL-1), pSVCosCEL-2 bis -8
Endoglukanasen	4	<i>gnuB</i> , ORF 48 (VII C10), ORF 80 (XIV A3), ORF 01 (pSVCEL-1)
Pektat Lyasen	3	ORF 03 ¹ (pSVCEL-1), <i>pelA</i> ¹ , ORF 120 (III F2)
Agarasen ¹	12	<i>aguA</i> , B (VII C10), <i>aguC</i> , D, E, F (XIV A3), <i>aguG</i> , H, I (III E5), <i>aguJ</i> , K, L (X B11)
Xylanasen	1	ORF 78 (XIV A3)
metallabhängige Hydrolase	1	ORF 174 (XIV B3)
Esterasen	1	ORF 83 (X B11)
Antibiotika Resistenz		
Metallo- β -Laktamase	1	ORF 125 (III F2)
Bacitracin Resistenz	1	<i>bacA</i> (VII H3)
Aminopeptidasen	1	ORF 199 (VII H3)

Alkohol Dehydrogenasen	1	ORF 34 (VII C10)
Σ 36		

Legende: Enzyme, die aktivitätsbasierend isoliert wurden sind grau hinterlegt;

¹ korrespondierendes Enzym ist aktiv.

Das Pektin stellt, nach Cellulose und Hemicellulose, den drittichtigsten Bestandteil der Pflanzenzellwand. Sein Abbau geschieht durch Zusammenarbeit mehrerer Enzyme wie den Polygalakturonasen, den Pektinesterasen, und den Pektin- und Pektatlyasen, die alle unter dem Begriff Pektinasen zusammengefasst werden (Solbak et al., 2005). Anwendungsgebiete umfassen alle Bereiche, in denen mit Pflanzenbestandteilen oder deren Früchten gearbeitet wird (Kashyap et al., 2001): in der Nahrungsmittelindustrie bei der Herstellung von Frucht- und Gemüsesäften, um die Saftausbeute zu erhöhen oder bei der Herstellung von Konzentraten aus Obst oder Gemüse (z.B. Tomatenmark). Auch bei der Weinherstellung finden Pektinasen Anwendung in der Klärung des frisch gepressten Mosts, wodurch die Konsistenz verbessert und der Wein dünnflüssiger wird. Insgesamt konnten drei ORFs für putative Pektat Lyasen identifiziert werden. Für das zum ORF 03 des Klons pSVCEL-1 korrespondierende Enzym PelB konnte im Zuge dieser Arbeit die Funktionalität nachgewiesen werden.

Durch Sequenzierung und Annotation konnten ebenfalls zwei ORFs mit hohen Ähnlichkeiten zu einer α -Amylase bzw. einem α -1,4-Glukan spaltenden Enzym ermittelt werden (siehe Tabelle 8, III. Ergebnisse). Amylytische Enzyme finden in vielen industriellen Bereichen Anwendung, vor allem in der Alkoholproduktion, der Stärkeverarbeitung, sowie der Papier- und Textilindustrie (Sorensen et al., 2004). Jedoch wäre auch in diesem Fall eine funktionsbasierende Durchmusterung der Genbank negativ verlaufen, da beide Cosmid-Klone in einem anschließend durchgeführten Plattentest auf Stärkehaltigem Medium keine Enzymaktivität zeigten. Erst nach Klonierung in entsprechende Expressionsvektoren verlief der Plattentest positiv (Stoekigt, 2005).

Die hohe Trefferquote, die mittels der aktivitätsbasierenden Durchmusterung erreicht werden konnte, bestätigt den Nutzen einer sinnvoll eingesetzten Anreicherung vor der DNA-Isolation. Zudem zeigte sich die Sequenzierung und Annotation großer Abschnitte von Umwelt-DNA als wertvolle Erweiterung, um die Anzahl der potentiellen Biokatalysatoren zu erhöhen. Im diesem Zusammenhang bieten „large insert“-Genbanken zudem den Vorteil Zugriff auf große, zusammenhängende Abschnitte der Umwelt-DNA möglich zu machen. Die Verteilung von Funktionen innerhalb eines

Konsortiums ebenso wie die genetische Organisation und horizontaler Gentransfer können so erschlossen werden (Handelsman, 2004).

2.1 Genduplikationen durch horizontalen Gentransfer

Agar ist ein komplexes Polysaccharid der Zellwand von Rotalgen und besteht aus zwei Komponenten, Agarose und Agaropektin. Den Hauptbestandteil bildet die Agarose, welche aus alternierenden Galaktose und 3,6-Anhydrogalaktose Einheiten aufgebaut ist, die zusammen als Agarobiose bezeichnet werden. Verknüpft sind die einzelnen Agarobioseeinheiten über β -1,4-glykosidische Bindungen während innerhalb der Agarobiose die Galaktose und 3,6-Anhydrogalaktose α -1,3-verknüpft sind. Bisher sind zwei Arten Agar hydrolysierender Enzyme bekannt, die β - und α -Agarasen. β -Agarasen hydrolysieren die β -1,4-glykosidische Bindung zwischen den Agarobioseeinheiten während α -Agarasen die α -1,3-Verknüpfungen zwischen der Galaktose und 3,6-Anhydrogalaktose spalten.

Die Sequenzierung und anschließende Annotation der vier Agarase-positiven Cosmid-Klone erbrachte zwölf ORFs, deren korrespondierenden Enzyme für die agarolytische Aktivität verantwortlich sein können (Abbildung 2).

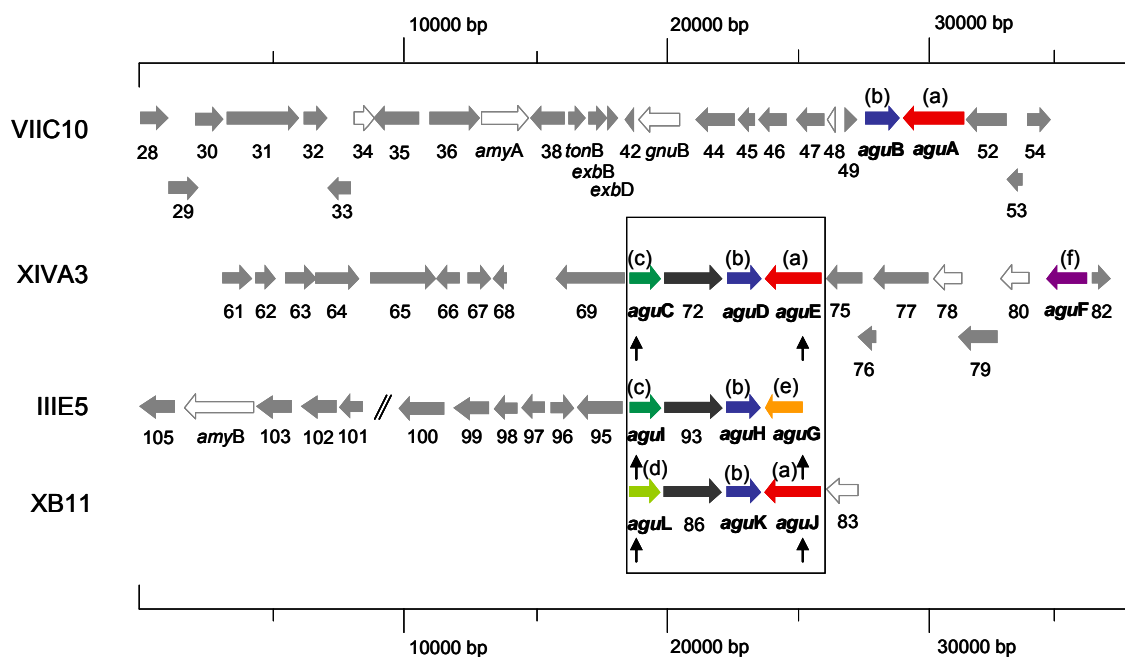


Abb. 2: Übersicht über die Anordnung der ORFs auf den Agarase-positiven Cosmiden.

Die für Agarasen kodierenden ORFs sind farbig hervorgehoben, wobei identische Gene mit gleichen Farben und Buchstaben markiert sind. Die ORFs weiterer Biokatalysatoren sind grau umrandet. Pfeile geben die Position eines 31 bp langen *inverted repeat* an.

Interessanterweise lag ein Großteil dieser ORFs in fast identischen Genclustern auf drei der vier Cosmide vor. Durch ein Alignment auf Aminosäureebene wurden die zwölf Gene und ihre Genprodukte in sechs verschiedene Gruppen von (a) bis (f) eingeteilt (Tabelle 3; Alignments befinden sich im Anhang). Allerdings beträgt der Unterschied zwischen den Proteinen der Gruppe (c) und (d) nur 1 %, die Länge ist mit je 451 AS identisch. Anders bei den Gruppen (a) und (e). Diese sind auf Sequenzebene 100 %ig identisch jedoch ist AguG um 315 AS kürzer als die Agarasen der Gruppe (a).

Die gefundenen Gencluster auf den Cosmiden XIV A3, III E5 und X B11 werden links und rechts von identischen je 31 bp langen *inverted repeats* flankiert und haben eine Größe von 6,2 kb. Sie enthalten neben den für die putativen Agarasen kodierenden ORFs noch einen ORF unbekannter Funktion von 2,36 kb Länge. *Inverted repeats* sind Teil von transponierbaren Elementen und können einen Hinweis auf die Entstehung dieser Gencluster geben.

Tab. 3: Einteilung in Gruppen und Funktion des nächsten Treffers der identifizierten ORFs.

Gruppe	ORFs	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	Identität (%)	Nummer der Glykosyl-Hydrolase-Familie
(a)	aguA, E, J	772	hypoth. Protein – (β -Galaktosidase)	ZP_00315251	53	50
(e)	aguG	457	<i>S. degradans</i>		63	
(b)	aguB, D, H, K	449	β -Agarase I (AgaB) – <i>S. degradans</i>	AAT67062	53	16
(c)	aguC, I	451	hypoth. Protein – <i>S. degradans</i>	ZP_00315360	65	50
(d)	aguL	451			64	
(f)	aguF	580	β -Agarase – <i>Pseudomonas</i> sp. ND137	BAD88713	38	16

Transponierbare Elemente spielen eine zentrale Rolle beim Austausch von DNA während des horizontalen Gentransfers. Diese Übertragung von Genen zwischen zwei Organismen wird durch Plasmide, Bakteriophagen und Transposons vermittelt (Frost et al., 2005). Diese besondere Art von mobilen DNA Elementen umfasst katabolische Transposons, die auch als „genomische Inseln“ bezeichnet werden, da sie bis zu 500 kb groß sein können (Tsuda et al., 1999; van der Meer and Sentchilo, 2003). Hierbei

unterscheidet man zwei Arten: Die Klasse-I-Transposons sind von IS-Elementen eingerahmt, welche die Funktionen für die Transposition (Transposase) des zentralen Abschnitts zur Verfügung stellen. Klasse-II-Transposons enthalten kurze *inverted repeats* an den Enden. Dementsprechend kodieren Abschnitte im zentralen Bereich die Transposition. Es sind eine Vielzahl von katabolischen Transposons bekannt, die für den bakteriellen Abbau von xenobiotischen Chemikalien verantwortlich sind (Springael and Top, 2004; Tan, 1999). Man nimmt sogar an, dass viele der für den Abbau dieser Substanzen verantwortlichen Stoffwechselwege durch horizontalen Gentransfer in einer Art Patchwork-Mechanismus entstanden sind. Andere katabolische Transposons enthalten jedoch auch Gene, die für den Abbau von Zuckern oder Zitronensäure kodieren, wie z.B. das Laktose Operon von *Yersinia enterocolitica* (Cornelis, 1981).

Das die Anpassung von Bakterien an die vorherrschenden Umweltbedingungen häufig mit horizontalem Gentransfer einhergeht zeigen auch die Pathogenitätsinseln, welche Virulenzfaktoren vermitteln, oder die symbiotische Insel aus *Mesorhizobium loti*, deren Übertragung nicht-symbiotische *Mesorhizobia* die Fähigkeit der Stickstofffixierung in einem Pflanzen-Wirt verleiht (Hacker and Kaper, 2000; Sullivan et al., 2002).

Ein Nachweis für die Existenz solcher genomischer Inseln in Agar abbauenden Organismen ist möglicherweise die Veröffentlichung eines 101 kb großen Plasmids aus *Microscilla* sp., deren Genprodukte in den Abbau von Polysacchariden sowie am Transport und Modifikation von Zuckern beteiligt sind (Zhong et al., 2001). Insgesamt konnten fünf unterschiedliche ORFs identifiziert werden, die für β -Agarasen kodieren. Darüber hinaus befinden sich auf diesem Plasmid sechs ORFs mit Homologien zu Transposasen, die von *inverted repeats* flankiert werden. Zusätzlich konnten zwei große *inverted repeats* von jeweils 1,4 kb Länge identifiziert werden. Sie flankieren einen Bereich, der ca. 2/3 des gesamten Plasmids umfasst und in dem sich sowohl die ORFs für die Agarasen als auch der Transposasen befinden.

Homologien zu einer Transposase wies nur der ORF 188 des Cosmids XIV B3 auf. Die auffällige Häufung von am Abbau von komplexen Polysacchariden aus Pflanzen beteiligten Enzymen innerhalb der sequenzierten Genabschnitte macht einen horizontalen Gentransfer aber durchaus wahrscheinlich. Eventuell handelt es sich bei den sequenzierten Cosmiden auch um Teile einer größeren genomischen Insel. Dies müsste im Zuge weiterer Arbeiten geklärt werden.

3. Charakterisierung einer Cellulase des angereicherten Bodenkonsortiums

Durch die Anreicherung gelang es ein Konsortium mit vielfältigen hydrolytischen Eigenschaften zu selektionieren. Neben weiteren am Abbau von komplexen Polysacchariden beteiligten Enzymen konnten insgesamt acht cellulolytische Klone einer 1700 Cosmid-Klone umfassenden Umwelt-Genbank durch aktivitätsbasierende Durchmusterung identifiziert werden.

Die Sequenzierung des Abschnitts eines dieser Cosmide ergab einen ORF von 1095 bp Länge, welcher für ein 363 AS großes Protein, CelA, kodiert. Hierbei handelt es sich um ein *single domain* Enzym, da es neben der katalytischen Domäne der Glykosylhydrolase-Familie 5 (GHF 5) keine Cellulosebindedomäne(n) besitzt. Ein Aminosäureabgleich mit der NCBI-Datenbank zeigte die höchsten Übereinstimmungen zu einer Cellulase aus *Cellvibrio mixtus*, welche ebenfalls keine Cellulosebindedomäne besitzt (Fontes et al., 1997). Diese Domänen spielen eine aktive Rolle bei der Hydrolyse von Cellulose, entweder durch Zerstörung der kristallinen Struktur der Cellulose (Din et al., 1993) oder indem sie das Enzym näher zu den amorphen Bereichen bringen (Hall et al., 1995). Cellulasen, denen diese Domäne fehlt zeigen verringerte bis gar keine Enzymaktivität gegen kristalline Cellulose während die Fähigkeit lösliche Cellulose zu hydrolysieren nicht betroffen ist (Bolam et al., 1998; Coutinho et al., 1993).

Auch CelA zeigte die höchsten Enzymaktivitäten gegen lösliche Formen der Cellulose wie β -Glukan und Carboxymethylcellulose, während gegen mikrokristalline Formen wie Avicel oder Filterpapier keine Aktivität festzustellen war. Insgesamt zeigte CelA eine Spezifität für lösliche β -1,4-verknüpfte Glukane und war nicht in der Lage reine β -1,3-verknüpfte Glukane wie das Laminarin oder β -1,4-verknüpfte Xylose (Xylan) abzubauen. Somit zeigte CelA die typische Enzymaktivität einer β -1,4-Glukanase, jedoch mit höheren Enzymaktivitäten als bei der Cellulase aus *Cellvibrio mixtus* festgestellt wurden. Durch anschließende dünnschichtchromatographische Untersuchungen der Spaltungsprodukte konnte nachgewiesen werden, dass es sich bei CelA um eine Endoglukanase handelt. Im Gegensatz zu den Exoglukanasen entstehen hierbei in der Anfangsphase des Verdaus nur Produkte mit höheren Polymerisationsgraden.

Weitere biochemischen Tests zeigten, dass das Enzym über ein breites pH-Spektrum hinweg aktiv war mit einem Aktivitätsmaximum bei pH 6,5. Das Enzym zeigte noch 60 % seiner Maximalaktivität bei einem pH-Wert von 5,5 und 68 % bei pH 9,0. Einen

ähnlich breites Spektrum ist sonst nur bei Vertretern von extremen Standorten anzutreffen. So besitzt eine Endoglukanase aus dem thermoacidophilen Bakterium *Alicyclobacillus acidocaldarius* ein pH-Spektrum von pH 5 bis 9 (Eckert et al., 2002). Für zwei andere Cellulasen aus Anreicherungskulturen wurde ebenfalls ein pH-Bereich von pH 5 bis 9 ermittelt (Healy et al., 1995). Es herrschten dort aber anaerobe und thermophile Anreicherungsbedingungen, welche die Selektion von extremophilen Organismen begünstigt haben dürfte.

Zusätzlich zu dem breiten pH-Spektrum zeigte sich CelA über einen weiten Temperatur-Bereich von 30 °C bis 50 °C aktiv, mit einem Maximum bei 45 °C. Darüber hinaus zeigte es über 11 Tage bei 40 °C eine gleichmäßig hohe Enzymaktivität. Der Grund für diese außergewöhnliche Thermostabilität kann wiederum an der fehlenden Cellulosebindedomäne (CBD) liegen. Es wird angenommen, dass Enzyme mit fehlender CBD langsamer mit dem Substrat interagieren und eine erhöhte Stabilität für diese Enzyme von Vorteil ist (Fontes et al., 1997). Ähnliche Beobachtungen hinsichtlich der Thermostabilität wurden dementsprechend auch für die Cellulase aus *Cellvibrio mixtus* gemacht.

Interessanterweise zeigte sich CelA äußerst stabil in der Gegenwart von Salzen. So führten Konzentrationen von 3M Natrium- oder Rubidiumsalm und sogar 4M Kaliumsalz im Reaktionsansatz immer noch zu relativen Enzymaktivitäten von 54 % bis 75 %. Bisher ist eine derartige Toleranz gegen hohe Salzkonzentrationen eher von Enzymen von marinen Organismen bekannt. So konnte bei zwei Cellulasen aus dem thermophilen Bakterium *Thermotoga maritima* die Thermostabilität durch die Zugabe von bis zu 5 M NaCl erhöht werden (Liebl et al., 1996).

Aufgrund seines breiten pH-Spektrums, seiner Temperatur-Stabilität und der Toleranz gegen erhöhte Salzkonzentrationen ist CelA ideal vorbereitet für die vorherrschenden Bedingungen in der Papier und Textilindustrie, die zu den wichtigsten Einsatzbereichen von Cellulasen zählen.

4. Ausblick

Die Ergebnisse dieser Arbeit zeigen, dass Umweltbanken das Potential bieten sich die genetische Diversität eines Standortes anzueignen, ohne auf die Kultivierbarkeit der Organismen angewiesen zu sein. Die hohe Anzahl an identifizierten funktionellen und putativen Biokatalysatoren bestätigt darüber hinaus den Nutzen den Anreicherungskulturen für die Erhöhung der Trefferquote besitzen.

Da die im Zuge dieser Arbeit identifizierten Gene bzw. deren korrespondierende Enzyme nicht alle charakterisiert werden konnten, bieten sich auch hier weiterführende Arbeiten an. Diese umfassen vor allem erste Tests hinsichtlich der Höhe der enzymatischen Aktivität sowie der Substratspezifität, des pH-Spektrums und des Temperatur-Optimums.

Obwohl aus einem mesophilen Habitat stammend, erwies sich die aus dem hydrolytischen Boden-Konsortium identifizierte Cellulase CelA als äußerst stabil, tolerant gegen hohe Salzkonzentrationen und besitzt ein breites pH-Spektrum. Es sind jedoch weitere Arbeiten sinnvoll, die den Einsatz dieses Enzyms unter industriellen Bedingungen untersuchen. Solche *scale up*-Prozesse sind unter anderem aus ökonomischer Sicht notwendig, um die Eignung eines Enzyms für industrielle Großdimensionen zu testen.

Weitere Arbeiten beinhalten auch Kristallisationsexperimente und eine Röntgenstrukturanalyse zur Aufklärung des aktiven Zentrums. Damit würden auch die Voraussetzungen geschaffen, dass Enzym mit Hilfe von *Protein Engineering* für spezifische industrielle Anwendungen zu optimieren.

Die Sequenzierung und anschließende Annotation der vier Agarase-positiven Cosmid-Klone erbrachte zwölf ORFs, deren korrespondierenden Enzyme für die agarolytische Aktivität verantwortlich sein können. Die Entstehung dieser Cluster durch horizontalen Gentransfer ist wahrscheinlich, jedoch konnte nicht abschliessend geklärt werden, ob die Cosmide Teil einer genomischen Insel sind. Hier erscheinen weitere Arbeiten sinnvoll, in denen eventuell überlappende Cosmide identifiziert und sequenziert werden, um so einen größeren, zusammenhängenden Genabschnitt zu generieren.

V. Zusammenfassung

Die Durchmusterung von Umweltgenbanken hat sich in den letzten Jahren als effektive Technik erwiesen, auf den stetig wachsenden Bedarf der Industrie an neuen Biokatalysatoren zu reagieren (Voget et al., 2005; Streit et al., 2004). Im Zuge dieser Arbeit sollte eine Umwelt-Genbank eines Boden-Konsortiums durch die Kombination unterschiedlicher metagenomischer Methoden analysiert und der Nutzen von Anreicherungen zur Erhöhung der Trefferquote untersucht werden. Insgesamt wurden 36 biotechnologisch interessante Gene bzw. deren Produkte aus unterschiedlichen Enzymklassen gefunden (Voget et al., 2003; Voget et al., 2006). Eine neuartige, äußerst stabile und halotolerante Cellulase wurde zudem einer genauen biochemischen Charakterisierung unterzogen (Voget et al., 2006).

Um diese Ziele zu erreichen wurden folgende Strategien verfolgt:

- I. Eine phylogenetische Analyse der Diversität des angereicherten Boden-Konsortiums.
 - II. Durchmusterung von Genbanken nach biotechnologisch interessanten Enzymen unter Verwendung unterschiedlicher Durchmusterungsmethoden inklusive der Sequenzierung großer, zusammenhängender DNA-Abschnitte.
 - III. Heterologe Expression und biochemische Charakterisierung der Cellulase CelA aus metagenomischer DNA
- I. Für eine phylogenetische Analyse wurde sowohl die Gesamt-DNA des Konsortiums verwendet, als auch die DNA der Einzelisolate. Die Analyse der 16S rRNA-Gene ergab, dass das Konsortium aus mindestens zehn Organismen besteht. Es konnten Sequenzähnlichkeiten zu Vertretern typischer Bodenbewohner wie *Agrobacterium*, *Bacillus*, *Brevibacillus*, *Micrococcus*, und *Pseudomonas* ermittelt werden. Vier 16S rRNA-Gensequenzen zeigten die größten Homologien zu bisher unkultivierten Bakterien der Spezies *Cellvibrio*, *Janthinobacterium* und *Xanthomonas*.
- II. Es wurden zwei Ansätze verfolgt, um die Wahrscheinlichkeit zu erhöhen biotechnologisch interessante Enzyme zu finden. Es wurde zum einen das klassische, aktivitätsbasierende, zum anderen das auf Sequenzähnlichkeiten basierende Durchmusterung genutzt. Die Durchmusterung nach agarolytischer Aktivität führte zur Identifizierung von vier Cosmid-Klonen aus der 1700 Klone umfassenden Metagenom-Genbank. Mittels Sequenzanalysen dieser Cosmide konnten 12 ORFs identifiziert werden, die für mögliche, funktionale Agarasen kodieren. Interessanterweise lag ein

Großteil dieser ORFs in fast identischen Genclustern auf drei der vier Cosmide vor. Zudem wurden diese Gencluster von *inverted repeats* flankiert. Dies deutete darauf hin, dass horizontaler Gentransfer an der Entstehung dieser Cluster beteiligt war. Darüber hinaus wurde die Genbank auf cellulolytische Aktivität durchmustert, wobei acht positive Cosmid-Klone identifiziert und bei drei Klonen die cellulolytischen ORFs mittels Transposonmutagenese identifiziert werden konnten. Bei Sequenzanalyse eines dieser Klone, pSVCEL-1, konnten zusätzlich eine Glykosylhydrolase und eine Pektat Lyase ermittelt werden. Insgesamt wurden 243 kb an metagenomischer Sequenzinformation gewonnen, die zur Identifizierung elf weiterer ORFs führten, deren Genprodukte am Abbau komplexer Polysaccharide beteiligt sein können.

III. Der 1092 bp große ORF *celA* wurde aus dem Cosmid pSVCosCEL-1 kloniert, in *E. coli* überexprimiert und das korrespondierende 42,1 kDa große Enzym (CelA) gereinigt und charakterisiert (Voget et al., 2006). Sequenzanalysen zeigten eine Ähnlichkeit von 86 % zu einer Cellulase aus *Cellvibrio mixtus* und führten zu einer Einordnung in die Glykosylhydrolase-Familie 5. Das rekombinante Protein zeigte sehr hohe Aktivitäten gegen β -1,4-verknüpfte Glukane wie CMC (68 U mg^{-1}), Lichenan (401 U mg^{-1}) und dem β -Glukan der Gerste (710 U mg^{-1}). Bei β -1,3-verknüpften Glukanen und Xylan (β -1,4-verknüpfte Xylose) konnte keine Aktivität festgestellt werden. Ebenso konnten unlösliche Substrate wie kristalline Cellulose (Avicel) nicht verwertet werden. Dies hängt sehr wahrscheinlich mit der fehlenden Cellulosebindedomäne zusammen, welche für den Abbau unlöslicher Substrate unerlässlich ist. Dünnschichtchromatographische Analysen zeigten, daß CelA einen für Endoglukanasen typischen Reaktionsmechanismus zeigte. CelA zeigte sich über einen großen pH-Bereich von pH 5,5 bis pH 9,0 stabil, mit einem Optimum bei pH 6,5. Damit besitzt das Enzym eines der größten bisher für Cellulasen veröffentlichten pH-Bereiche. Dies ist besonders bemerkenswert, da es sich bei dem Probenursprung nicht um ein extremes Habitat handelt. Ebenso zeigte CelA gute Aktivitäten über einen weiten Temperatur-Bereich. Es zeigte eine sehr hohe Stabilität bei 40°C über 11 Tage. Dies steht in Übereinstimmung mit früheren Studien, die Thermostabilität bei mesophilen Enzymen der GHF 5 als generelle Eigenschaft von Cellulasen ohne Cellulose-Bindedomänen nachwies. Interessanterweise zeigte sich CelA äußerst stabil in der Gegenwart von Salzen. So führten Konzentrationen von 3M Natrium- oder Rubidiumsalm und sogar 4M Kaliumsalm im Reaktionsansatz immer noch zu relativen Enzymaktivitäten von 54 % bis 75 %. Bisher ist eine derartige Toleranz gegen hohe Salzkonzentrationen eher von Enzymen aus marinen Organismen bekannt (Voget et al., 2006).

VI. Literaturverzeichnis

- Achenbach, L. A., and Coates, J. D. (2000).** Disparity between bacterial phylogeny and physiology. *Am Soc Microbiol News* 66, 714-715.
- Andrykovich, G., and Marx, I. (1988).** Isolation of a new polysaccharide digesting bacterium from a salt marsh. *Appl Microbiol Biotechnol* 54, 1061-1062.
- Bachmann, B. J. (1987).** Linkage map of *Escherichia coli* K12., Vol 2 (Washington D.C., USA: American Society for Microbiology (ASM)).
- Barns, S. M., Takala, S. L., and Kuske, C. R. (1999).** Wide Distribution and Diversity of Members of the Bacterial Kingdom Acidobacterium in the Environment. *Appl Environ Microbiol* 65, 1731-1737.
- Beg, Q. K., Kapoor, M., Mahajan, L., and Hoondal, G. S. (2001).** Microbial xylanases and their industrial applications: a review. *Applied Microbiology and Biotechnology* 56, 326.
- Bell, P. J., Sunna, A., Gibbs, M. D., Curach, N. C., Nevalainen, H., and Bergquist, P. L. (2002).** Prospecting for novel lipase genes using PCR. *Microbiology* 148, 2283-2291.
- Bhat, M. K., and Bhat, S. (1997).** Cellulose degrading enzymes and their potential industrial applications. 15, 583.
- Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998).** *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochem J* 331, 775-781.
- Bonfield, J. K., Smith, K., and Staden, R. (1995).** A new DNA sequence assembly program. *Nucleic Acids Res* 23, 4992-4999.
- Bonfield, J. K., and Staden, R. (1996).** Experiment files and their application during large-scale sequencing projects. *DNA Sequence* 6, 109-117.
- Brady, S. F., Chao, C. J., and Clardy, J. (2004).** Long-Chain N-Acyltyrosine Synthases from Environmental DNA. *Appl Environ Microbiol* 70, 6865-6870.
- Brosius, J., Dull, T. J., Sleeter, D. D., and Noller, H. F. (1981).** Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. *J Mol Biol* 148, 107-127.
- Bull, A. T., Ward, A. C., and Goodfellow, M. (2000).** Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol Mol Biol Rev* 64, 573-606.
- Burton, S. G., Cowan, D. A., and Woodley, J. M. (2002).** The search for the ideal biocatalyst. *Nature Biotechnology* 20, 37-45.
- Collins, T., Gerday, C., and Feller, G. (2005).** Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol Rev* 29, 3-23.
- Cornelis, G. (1981).** Sequence relationships between plasmids carrying genes for lactose utilization. *J Gen Microbiol* 124, 91-97.

Courtois, S., Cappellano, C. M., Ball, M., Francou, F. X., Normand, P., Helynck, G., Martinez, A., Kolvek, S. J., Hopke, J., Osburne, M. S., et al. (2003). Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* 69, 49-55.

Coutinho, J. B., Gilkes, N. R., Kilburn, D. G., Warren, R. A. J., and Miller, J. R. C. (1993). The nature of the cellulose-binding domain effects the activities of a bacterial endoglucanase on different forms of cellulose. *113*, 211.

Cowan, D. A., Arslanoglu, A., Burton, S. G., Baker, G. C., Cameron, R. A., Smith, J. J., and Meyer, Q. (2004). Metagenomics, gene discovery and the ideal biocatalyst. *Biochem Soc Trans* 32, 298-230.

Daniel, R. (2005). The metagenomics of soil. *Nat Rev Microbiol* 3, 470-478.

Demain, A. L., Newcomb, M., and Wu, J. H. (2005). Cellulase, clostridia, and ethanol. *Microbiol Mol Biol Rev* 69, 124-154.

Din, N., Gilkes, N. R., Tekant, B., Miller, R. C., Warren, R. A. J., and Kilburn, D. G. (1993). Non-hydrolytic disruption of cellulose fibers by the binding domain of bacterial cellulase. *Biotechnology* 9, 1096-1099.

Docquier, J.-D., Riccio, M. L., Mugnaioli, C., Luzzaro, F., Endimiani, A., Toniolo, A., Amicosante, G., and Rossolini, G. M. (2003). IMP-12, a New Plasmid-Encoded Metallo-beta-Lactamase from a *Pseudomonas putida* Clinical Isolate. *Antimicrob Agents Chemother* 47, 1522-1528.

Doi, R. H., and Kosugi, A. (2004). Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol* 2, 541-551.

Dumont, M. G., and Murrell, J. C. (2005). Stable isotope probing: Linking microbial identity to function. *Nature Reviews Microbiology* 3, 499.

Eckert, K., Zielinski, F., Lo Leggio, L., and Schneider, E. (2002). Gene cloning, sequencing, and characterization of a family 9 endoglucanase (CelA) with an unusual pattern of activity from the thermoacidophile *Alicyclobacillus acidocaldarius* ATCC27009. *Appl Microbiol Biotechnol* 60, 428-436.

Ekborg, N. A., Gonzalez, J. M., Howard, M. B., Taylor, L. E., Hutcheson, S. W., and Weiner, R. M. (2005). *Saccharophagus degradans* gen. nov., sp. nov., a versatile marine degrader of complex polysaccharides. *Int J Syst Evol Microbiol* 55, 1545-1549.

Ekborg, N. A., Howard, M. B., Taylor, L. E., Weiner, R. M., and Hutcheson, S. W. (2005). Genomic Analysis of the Agar Degrading System of *Microbulbifer degradans* strain 2-40 Reveals Two Beta-agarases with Unusual Protein Architecture and Properties (NCBI GenPept).

Ensor, L. A., Stosz, S. K., and Weiner, R. M. (1999). Expression of multiple complex polysaccharide-degrading enzyme systems by marine bacterium strain 2-40. *Journal of Industrial Microbiology and Biotechnology* 23, 123.

Entcheva, P., Liebl, W., Johann, A., Hartsch, T., and Streit, W. R. (2001). Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Appl Environ Microbiol* 67, 89-99.

Ewing, B., and Green, P. (1998). Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities. *Genome Res* 8, 186-194.

Ferrer, M., Golyshina, O. V., Chernikova, T. N., Khachane, A. N., Martins Dos Santos, V. A., Yakimov, M. M., Timmis, K. N., and Golyshin, P. N. (2005). Microbial enzymes mined from *Urania* deep-sea hypersaline anoxic basin. *Chem Biol* 8, 895-904.

Fontes, C. M., Clarke, J. H., Hazlewood, G. P., Fernandes, T. H., Gilbert, H. J., and Ferreira, L. M. (1997). Possible roles for a non-modular, thermostable and proteinase-resistant cellulase from the mesophilic aerobic soil bacterium *Cellvibrio mixtus*. *Appl Microbiol Biotechnol* 48, 473-479.

Fontes, C. M., Clarke, J. H., Hazlewood, G. P., Fernandes, T. H., Gilbert, H. J., and Ferreira, L. M. (1998). Identification of tandemly repeated type VI cellulose-binding domains in an endoglucanase from the aerobic soil bacterium *Cellvibrio mixtus*. *Appl Microbiol Biotechnol* 49, 552-559.

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* 3, 722.

Fuchs, K.-P., Zverlov, V. V., Velikodvorskaya, G. A., Lottspeich, F., and Schwarz, W. H. (2003). Lic16A of *Clostridium thermocellum*, a non-cellulosomal, highly complex endo- β -1,3-glucanase bound to the outer cell surface. *Microbiology* 149, 1021-1031.

Gabor, E. M., Alkema, W. B. L., and Janssen, D. B. (2004). Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* 6, 879-886.

Gabor, E. M., de Vries, E. J., and Janssen, D. B. (2004). Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ Microbiol* 6, 948-958.

Gavrilescu, M., and Chisti, Y. (2005). Biotechnology--a sustainable alternative for chemical industry. 23, 471.

Gilbert, H. J., Hall, J., Hazlewood, G. P., and Ferreira, L. M. (1990). The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies *cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. *Mol Microbiol* 4, 759-767.

Gillespie, D. E., Brady, S. F., Bettermann, A. D., Cianciotto, N. P., Liles, M. R., Rondon, M. R., Clardy, J., Goodman, R. M., and Handelsman, J. (2002). Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* 68, 4301-4306.

Ginolhac, A., Jarrin, C., Gillet, B., Robe, P., Pujic, P., Tuphile, K., Bertrand, H., Vogel, T. M., Perriere, G., Simonet, P., and Nalin, R. (2004). Phylogenetic Analysis of Polyketide Synthase I Domains from Soil Metagenomic Libraries Allows Selection of Promising Clones. *Appl Environ Microbiol* 70, 5522-5527.

Godfrey, T. (2003). The enzymes market for grain processing. In *Recent advances in enzymes in grain processing*. 401-406

Godfrey, T., and West, S. (1996). *Industrial Enzymology* (London: Maximillian Press Ltd.).

- Gomes, J., and Steiner, W. (2004).** The biocatalytic potential of extremophiles and extremozymes. *Food Technol Biotechnol* 42, 223-235.
- Gonzalez, J., and Weiner, R. (2000).** Phylogenetic characterization of a marine bacterium strain 2-40, a degrader of complex polysaccharides. *Int J Syst Bacteriol* 8, 831-834.
- Gonzalez, J. M., Mayer, F., Moran, M. A., Hodson, R. E., and Whitman, W. B. (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, 369-376.
- Grant, S., Sorokin, D. Y., Grant, W. D., Jones, B. E., and Heaphy, S. (2004).** A phylogenetic analysis of Wadi el Natrun soda lake cellulase enrichment cultures and identification of cellulase genes from these cultures. *Extremophiles* 8, 421.
- Hacker, J., and Kaper, J. B. (2000).** Pathogenicity islands and the evolution of microbes. *Annual Review of Microbiology* 54, 641-679.
- Hall, J., Black, G. W., Ferreira, L. M., Millward-Sadler, S. J., Ali, B. R., Hazlewood, G. P., and Gilbert, H. J. (1995).** The non-catalytic cellulose-binding domain of a novel cellulase from *Pseudomonas fluorescens* subsp. *cellulosa* is important for the efficient hydrolysis of Avicel. *Biochem J* 309, 749-756.
- Hamaki, T., Suzuki, M., FUDOU, R., JOJIMA, Y., KAJIURA, T., TABUCHI, A., SEN, K., and SHIBAI, H. (2005).** Isolation of Novel Bacteria and Actinomycetes Using Soil-Extract Agar Medium. *J BIOSCI BIOENG* 99, 485-492.
- Handelsman, J. (2004).** Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 68, 669-685.
- Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998).** Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5, R245-249.
- Healy, F. G., Ray, R. M., Aldrich, H. C., Wilkie, A. C., Ingram, L. O., and Shanmugam, K. T. (1995).** Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. *Applied Microbiology and Biotechnology* 43, 667.
- Henne, A., Schmitz, R. A., Bomeke, M., Gottschalk, G., and Daniel, R. (2000).** Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl Environ Microbiol* 66, 3113-3116.
- Hilden, L., and Johansson, G. (2004).** Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnol Lett* 26, 1683-1693.
- Hoster, F., Schmitz, J. E., and Rolf, D. (2005).** Enrichment of chitinolytic microorganisms: isolation and characterization of a chitinase exhibiting antifungal activity against phytopathogenic fungi from a novel *Streptomyces* strain. *Applied Microbiology and Biotechnology* 66, 434.
- Howard, M. B., Ekborg, N. A., Taylor, L. E., II, Weiner, R. M., and Hutcheson, S. W. (2004).** Chitinase B of "Microbulbifer degradans" 2-40 Contains Two Catalytic Domains with Different Chitinolytic Activities. *J Bacteriol* 186, 1297-1303.

- Hugenholtz, P. (2002).** Exploring prokaryotic diversity in the genomic era. *Genome Biology* 3, reviews0003.0001 - reviews0003.0008.
- Hugenholtz, P., Goebel, B. M., and Pace, N. R. (1998).** Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *J Bacteriol* 180, 4765-4774.
- Joseph, S. J., Hugenholtz, P., Sangwan, P., Osborne, C. A., and Janssen, P. H. (2003).** Laboratory Cultivation of Widespread and Previously Uncultured Soil Bacteria. *Appl Environ Microbiol* 69, 7210-7215.
- Kane, M. D., Poulsen, L. K., and Stahl, D. A. (1993).** Monitoring the enrichment and isolation of sulfate-reducing bacteria by using oligonucleotide hybridization probes designed from environmentally derived 16S rRNA sequences. *Appl Environ Microbiol* 59, 682-686.
- Kashyap, D. R., Vohra, P. K., Chopra, S., and Tewari, R. (2001).** Applications of pectinases in the commercial sector: a review. *Bioresource Technology* 77, 215-227.
- Knietsch, A., Bowien, S., Whited, G., Gottschalk, G., and Daniel, R. (2003).** Identification and Characterization of Coenzyme B12-Dependent Glycerol Dehydratase- and Diol Dehydratase-Encoding Genes from Metagenomic DNA Libraries Derived from Enrichment Cultures. *Appl Environ Microbiol* 69, 3048-3060.
- Knietsch, A., Bowien, S., Whited, G., Gottschalk, G., and Daniel, R. (2003).** Identification and characterization of coenzyme B12-dependent glycerol dehydratase- and diol dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures. *Appl Environ Microbiol* 69, 3048-3060.
- Knietsch, A., Waschowitz, T., Bowien, S., Henne, A., and Daniel, R. (2003).** Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli*. *Appl Environ Microbiol* 69, 1408-1416.
- Knietsch, A., Waschowitz, T., Bowien, S., Henne, A., and Daniel, R. (2003).** Construction and Screening of Metagenomic Libraries Derived from Enrichment Cultures: Generation of a Gene Bank for Genes Conferring Alcohol Oxidoreductase Activity on *Escherichia coli*. *Appl Environ Microbiol* 69, 1408-1416.
- Kulkarni, N., Shendye, A., and Rao, M. (1999).** Molecular and biotechnological aspects of xylanases. *FEMS Microbiol Rev* 23, 411-456.
- Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Larsen, N., Olsen, G. J., Maidak, B. L., McCaughey, M. J., Overbeek, R., Macke, T. J., Marsh, T. L., and Woese, C. R. (1993).** The ribosomal database project. *Nucleic Acids Res* 21, 3021-3023.
- Liebl, W., Ruile, P., Bronnenmeier, K., Riedel, K., Lottspeich, F., and Greif, I. (1996).** Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. *Microbiology* 142, 2533-2542.

- Lieber, H.-P., Thielecke, K., Buchholz, K., and Reilly, P. J. (1993).** High-performance anion-exchange chromatography of saturated and unsaturated oligogalacturonic acids. *238*, 307.
- Liesack, W., Sela, S., Bercovier, H., Pitulle, C., and Stackebrandt, E. (1991).** Complete nucleotide sequence of the *Mycobacterium leprae* 23 S and 5 S rRNA genes plus flanking regions and their potential in designing diagnostic oligonucleotide probes. *FEBS Letters* *281*, 114-118.
- Liles, M. R., Manske, B. F., Bintrim, S. B., Handelsman, J., and Goodman, R. M. (2003).** A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl Environ Microbiol* *69*, 2684-2691.
- Lorenz, P., and Eck, J. (2005).** Metagenomics and industrial applications. *Nature Reviews Microbiology* *3*, 510.
- Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S. (2002).** Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev* *66*, 506-577.
- Matsumoto, M., Aoki, Y., and Kamei, Y. (2005).** Molecular cloning of the agarase gene from *Pseudomonas* sp. ND137 in *Escherichia coli* (NCBI GenPept).
- McCaig, A. E., Glover, L. A., and Prosser, J. I. (1999).** Molecular Analysis of Bacterial Community Structure and Diversity in Unimproved and Improved Upland Grass Pastures. *Appl Environ Microbiol* *65*, 1721-1730.
- McCaig, A. E., Grayston, S. J., Prosser, J. I., and Glover, L. A. (2001).** Impact of cultivation on characterisation of species composition of soil bacterial communities. *35*, 37.
- Meyerhans, A., Vartanian, J. P., and Wain-Hobson, S. (1990).** DNA recombination during PCR. *Nucleic Acids Res* *18*, 1687-1691.
- Mitsui, H., Grolach, K., Lee, H.-j., Hattori, R., and Hattori, T. (1997).** Incubation time and media requirements of culturable bacteria from different phylogenetic groups. *30*, 103-110.
- Muyzer, G., de Waal, E. C., and Uitterlinden, A. G. (1993).** Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* *59*, 695-700.
- Neefs, J. M., Van de Peer, L. H., and de Wachter, R. (1990).** Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res* *18*, 2237-2317.
- Ovreas, L., and Torsvik, V. V. (1998).** Microbial Diversity and Community Structure in Two Different Agricultural Soil Communities. *Microb Ecol* *36*, 303-315.
- Paabo, S., Irwin, D. M., and Wilson, A. C. (1990).** DNA damage promotes jumping between templates during enzymatic amplification. *J Biol Chem* *265*, 4718-4721.
- Pace, N. R., Stahl, D. A., Lane, D. J., and Olsen, G. J. (1985).** Analyzing natural microbial populations by rRNA sequencing. *ASM News* *51*, 4-12.

- Rappe, M. S., and Giovannoni, S. J. (2003).** The uncultured microbial majority. *Annu Rev Microbiol* 57, 369-394.
- Rees, H., Grant, S., Jones, B., Grant, W., and Heaphy, S. (2003).** Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* 7, 415.
- Rhee, J. K., Ahn, D. G., Kim, Y. G., and Oh, J. W. (2005).** New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library. *Appl Environ Microbiol* 71, 817-825.
- Riesenfeld, C. S., Goodman, R. M., and Handelsman, J. (2004).** Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol* 6, 981-989.
- Riesenfeld, C. S., Schloss, P. D., and Handelsman, J. (2004).** METAGENOMICS: Genomic Analysis of Microbial Communities. *Annu Rev Genet* 38, 525-552.
- Rodriguez-Valera, F. (2002).** Approaches to prokaryotic biodiversity: a population genetics perspective. *Environmental Microbiology* 4, 628-633.
- Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., et al. (2000).** Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* 66, 2541-2547.
- Sait, M., Hugenholtz, P., and Janssen, P. H. (2002).** Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environmental Microbiology* 4, 654-666.
- Sakai, T., Sakamoto, T., Hallaert, J., and Vandamme, E. J. (1993).** Pectin, pectinase and protopectinase: production, properties, and applications. *Advances in applied microbiology* 39, 213-294.
- Sambrook, J., Fritsch, E., and Maniatis, T. (1989).** *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Santosa, D. A. (2001).** Rapid extraction and purification of environmental DNA for molecular cloning applications and molecular diversity studies. *Mol Biotechnol* 17, 59-64.
- Schink, B. (2002).** Synergistic interactions in the microbial world. *Antonie van Leeuwenhoek* 81, 257.
- Schloss, P. D., and Handelsman, J. (2004).** Status of the Microbial Census. *Microbiol Mol Biol Rev* 68, 686-691.
- Schmeisser, C., Stockigt, C., Raasch, C., Wingender, J., Timmis, K. N., Wenderoth, D. F., Flemming, H. C., Liesegang, H., Schmitz, R. A., Jaeger, K. E., and Streit, W. R. (2003).** Metagenome survey of biofilms in drinking-water networks. *Appl Environ Microbiol* 69, 7298-7309.
- Schmidt, T. M., DeLong, E. F., and Pace, N. R. (1991).** Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol* 173, 4371-4378.

Schwarz, W. H. (2001). The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology* 56, 634.

Smit, E., Leeflang, P., Gommans, S., van den Broek, J., van Mil, S., and Wernars, K. (2001). Diversity and Seasonal Fluctuations of the Dominant Members of the Bacterial Soil Community in a Wheat Field as Determined by Cultivation and Molecular Methods. *Appl Environ Microbiol* 67, 2284-2291.

Solbak, A. I., Richardson, T. H., McCann, R. T., Kline, K. A., Bartnek, F., Tomlinson, G., Tan, X., Parra-Gessert, L., Frey, G. J., Podar, M., et al. (2005). Discovery of Pectin-degrading Enzymes and Directed Evolution of a Novel Pectate Lyase for Processing Cotton Fabric. *J Biol Chem* 280, 9431-9438.

Sorensen, J. F., Kragh, K. M., Sibbesen, O., Delcour, J., Goesart, H., Svensson, B., Tahir, T. A., Brufau, J., Perez-Vendrell, A. M., Bellincampi, D., et al. (2004). Potential role of glycosidase inhibitors in industrial applicatons. *Biochim Biophys Acta* 1696, 275-287.

Springael, D., and Top, E. M. (2004). Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. *Trends Microbiol* 12, 53-58.

Staden, R., Beal, K. F., and Bonfield, J. K. (2000). The Staden package, 1998. *Methods Mol Biol* 132, 115-130.

Steele, H. L., and Streit, W. R. (2005). Metagenomics: Advances in ecology and biotechnology. *FEMS Microbiol Lett* 247, 105-111.

Stoeckigt, C. (2005) Konstruktion und Durchmusterung von Metagenombanken: Identifizierung und Charakterisierung von neuartigen amylolytischen Enzymen, **Dissertation**, Georg-August-Universitaet, Goettingen.

Streit, W. R., Daniel, R., and Jaeger, K. E. (2004). Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Curr Opin Biotechnol* 15, 285-290.

Streit, W. R., and Schmitz, R. A. (2004). Metagenomics - the key to the uncultured microbes. *Curr Opin Microbiol* 7, 492-498.

Sugano, Y., Matsumoto, T., and Noma, M. (1994). Sequence analysis of the *agaB* gene encoding a new beta-agarase from *Vibrio* sp. strain JT0107. *Biochim Biophys Acta* 1218, 105-108.

Sullivan, J. T., Trzebiatowski, J. R., Cruickshank, R. W., Gouzy, J., Brown, S. D., Elliot, R. M., Fleetwood, D. J., McCallum, N. G., Rossbach, U., Stuart, G. S., et al. (2002). Comparative Sequence Analysis of the Symbiosis Island of *Mesorhizobium loti* Strain R7A. *J Bacteriol* 184, 3086-3095.

Suzuki, H., Sawai, Y., Suzuki, T., and Kawai, K. (2002). Purification and characterization of an extracellular alpha-neoagarooligosaccharide hydrolase from *Bacillus* sp. MK03. *J Biosci Bioeng* 93, 456-463.

Suzuki, H., Sawai, Y., Suzuki, T., and Kawai, K. (2003). Purification and characterization of an extracellular β -agarase from *Bacillus* sp. MK03. *Journal of Bioscience and Bioengineering* 95, 328-334.

- Tan, H. M. (1999).** Bacterial catabolic transposons. *Applied Microbiology and Biotechnology* 51, 1.
- Tanaka, T., Yan, L., and Burgess, J. G. (2003).** *Microbulbifer arenaceous* sp. nov., a new endolithic bacterium isolated from the inside of red sandstone. *Curr Microbiol* 47, 412-416.
- Teather, R. M. W., P.J. (1982).** Use of Congo red-poly saccharide interactions in enumeration of cellulolytic bacteria from the bovine rumen. *Appl Environ Microbiol* 43, 777-780.
- Torsvik, V., and Ovreas, L. (2002).** Microbial diversity and function in soil: from genes to ecosystems. *Curr Opin Microbiol* 5, 240-245.
- Towbin, H., Staehelin, T., and Gordon, J. (1992).** Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. *Biotechnology* 24, 145-149.
- Tsai, Y. L., and Olson, B. H. (1992).** Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Appl Environ Microbiol* 58, 2292-2295.
- Tsuda, M., Tan, H. M., Nishi, A., and Furukawa, K. (1999).** Mobile catabolic genes in bacteria. *Journal of Biocscience and Bioengineering* 87, 401-410.
- Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., Solovyev, V. V., Rubin, E. M., Rokhsar, D. S., and Banfield, J. F. (2004).** Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature*.
- Uesbeck, A. (2002)** Isolierung und Charakterisierung von agarolytischen Mikroorganismen, **Diplomarbeit**, Georg-August-Universität, Göttingen.
- van der Meer, J. R., and Sentchilo, V. (2003).** Genomic islands and the evolution of catabolic pathways in bacteria. *Curr Opin Biotechnol* 14, 248-254.
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D., Paulsen, I., Nelson, K. E., Nelson, W., et al. (2004).** Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66-74.
- Vining, L. C. (1992).** Secondary metabolism, inventive evolution and biochemical diversity-a review. *Gene* 115, 135-140.
- Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K. E., and Streit, W. R. (2003).** Prospecting for novel biocatalysts in a soil metagenome. *Appl Environ Microbiol* 69, 6235-6242.
- Voget, S., Steele, H. L., and Streit, W. R. (2005).** Metagenomes – an unlimited resource for novel genes, biocatalysts and metabolites. *Minerva Biotec* 17, 47-53.
- Voget, S., Steele, H. L., and Streit, W. R. (2006).** Characterization of a metagenome derived halotolerant cellulase. *Journal of Biotechnology* *in press*.
- Wheeler, D. L., Chappey, C., Lash, A. E., Leipe, D. D., Madden, T. L., Schuler, G. D., Tatusova, T. A., and Rapp, B. A. (2000).** Database resources of the National Center for Biotechnology Information. *Nucl Acids Res* 28, 10-14.

Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *PNAS* 95, 6578-6583.

Woese, C. R. (1987). Bacterial evolution. *Microbiol Rev* 51, 221-271.

Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F., Jr., et al. (2001). The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58. *Science* 294, 2317-2323.

Yi-Heng Percival Zhang, L. R. L. (2004). Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnology and Bioengineering* 88, 797-824.

Yin, B., Crowley, D., Sparovek, G., De Melo, W. J., and Borneman, J. (2000). Bacterial Functional Redundancy along a Soil Reclamation Gradient. *Appl Environ Microbiol* 66, 4361-4365.

Yoon, J.-H., Kim, I.-G., Oh, T.-K., and Park, Y.-H. (2004). *Microbulbifer maritimus* sp. nov., isolated from an intertidal sediment from the Yellow Sea, Korea. *Int J Syst Evol Microbiol* 54, 1111-1116.

Yoon, J.-H., Kim, I.-G., Shin, D.-Y., Kang, K. H., and Park, Y.-H. (2003). *Microbulbifer salipaludis* sp. nov., a moderate halophile isolated from a Korean salt marsh. *Int J Syst Evol Microbiol* 53, 53-57.

Yun, J., Kang, S., Park, S., Yoon, H., Kim, M.-J., Heu, S., and Ryu, S. (2004). Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomic Library. *Applied and Environmental Microbiology* 70, 7229.

Zhong, Z., Toukdarian, A., Helinski, D., Knauf, V., Sykes, S., Wilkinson, J. E., O'Bryne, C., Shea, T., DeLoughery, C., and Caspi, R. (2001). Sequence Analysis of a 101-Kilobase Plasmid Required for Agar Degradation by a *Microscilla* Isolate. *Appl Environ Microbiol* 67, 5771-5779.

Zinke, H. (2004). Weiße Biotechnologie: Neue Produkte, gesellschaftlicher Nutzen und Wertschöpfungspotentiale. *Zeitschrift für Biopolitik* 2, 31-41.

VII. Anhang

-35 -10

ttgcgg cagtcgtttt cgttgaaagc cttctgccta ctaaatttca tt**gagggg**atttctt

1 atgcgattaa aaattctatt gtccgggctg gtgctcgcca ctgttgcatc caatgcattg
 m r l k i l l s g l v l a t v a s n a l

61 gcagctaadc gtctgcggg ctatgtgacc atttgttcag aaggtaaaaag ctgttctgtc
 a a n r p a g y v t i c s e g k s c s v

121 gccagcgcta caacagttgc ttttgggcgc gctgataaat tcacttacia attgttgtcg
 a s a t t v a f g r a d k f t y k l l s

181 ggcagctttg tgtgttctga agcggatatt gggacgcgta ttgctggtgg cgtgaatgag
 g s f v c s e a v f g t r i a g g v n e

241 tgttctatcg cgcgcaaac taactcaagc agttctgtac ctgcgcagtc gtcatctagc
 c s i a r t t n s s s s v p a q s s s s

301 acatcgggtg tgagcagctc gaagagctcg gtggccagtt cggttgtaag tagcataaaa
 t s v v s s s k s s v a s s v v s s i k

361 agttcagtg caagttcatc atccagctca gtgactaatc ctctgttgg tgggttttat
 s s v a s s s s s v t n p p v g g f y

421 cccggctgtg aaaaaccaac tccgaccgaa accgttcaat ttaccgagac acgcgtcgtt
 p g c e k p t p t e t v q f t e t r v v

481 gcagctggtg aaatatttga tggtttaaat aaacgttaca actttaacgg cggcggtcag
 a a g e i f d g l n k r y n f n g g g q

541 gcggaagggc agccaccagt gttccgtgta ttggaaggcg gtgcggtaaa aaatgtgatt
 a e g q p p v f r v l e g g a v k n v i

601 atcggttcac tggcggctga tggatccac tgtttgggca actgctcgct tgatcacgta
 i g s l a a d g i h c l g n c s l d h v

661 tgggtgggaag atattggtga agacgcggcg acggcgatgg gaccagcggg caccattatg
 w w e d i g e d a a t a m g p a g t i m

721 agcatcactt gcggctcggc acttaacggg tcggataaaa ctttcaatt taatggccgt
 s i t c g s a l n g s d k t f q f n g r

781 ggtgaattac acatcagtca attttacgtg cagagttcag gcaaattggc acgtacctgt
 g e l h i s q f y v q s s g k l a r t c

841 ggcgattgca cgggcaacgg tggcccgcgc aaattagtg tcaacgatgt gatcacgcgc
 g d c t g n g g p r k l v i n d v i t r

901 gatgtcagca ccacgtcgg aatcaattcc aactttggtg atgtagcgac gatccgcaat
 d v s t i v g i n s n f g d v a t i r n

961 ctcaccctca ataatagcgg tacatccaaa accaaaatct gccaggtgta caaaggtgta
 l t l n n s g t s k t k i c q v y k g v

1021 gtgaaagggt cagggagcag ctccggcgtt ggtgtggagt ttgatacacc gaattgcgat

```
      v k g s g s s s a l g v e f d t p n c d
1081 gttcgctga acgatgtgac cttgctcggc gtaagccaaa tgaatacgtc tgcttgtgcg
      v r l n d v t l l g v s q m n t s a c a
1141 ggttcttgcc ccattccctaa
      g s c p i p -
```

Abb. 1: Nukleotidsequenz und die abgeleitete Aminosäuresequenz der Pektat-Lyase des Klons pSVCEL-1

Die dem Signal Peptid entsprechende Sequenz ist unterstrichen, die wahrscheinliche Shine-Dalgarno Sequenz ist fett und die Terminator Region ist kursiv gedruckt. Die entsprechende Aminosäuresequenz ist im Ein-Buchstabencode in grün dargestellt.

Tabelle 1: Annotation der auf VII C10 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
28	341	OM-Protein – <i>P. fluorescens</i>	ZP_0026 5447.1	1-315	184- 596/517	50 2e-82
29	390	Resistenz Efflux Pumpe (HlyD-Protein) – <i>A. vinelandii</i>	ZP_0009 0420.1	16-390	2-377/377	51 1e-99
30	348	hypoth. Protein – <i>P. fluorescens</i>	ZP_0026 5449.1	12-335	2-325/341	40 8e-61
31	925	hypoth. Protein – <i>N. europaea</i>	NP_8420 43.1	7-903	2-886/898	50 0.0
32	285	hypoth. Protein – <i>N. europaea</i>	NP_8422 48.1	55-283	5-232/241	34 2e-30
33	307	Transkriptions Regulator – <i>P. aeruginosa</i>	NP_2488 71.1	1-307	1-307/310	58 e-100
34	260	Dehydrogenase mit unters. Spezifitäten – <i>N. punctiforme</i>	ZP_0010 9053.1	21-258	9-246/248	65 6e-83
35	566	Gamma-glutamyltransferase – <i>S. degradans</i>	ZP_0031 6827.1	26-558	20-555/556	55 e-165
36	637	hypoth Protein – <i>G. violaceus</i>	NP_9255 74.1	3-637	68-699/706	47 e-161
<i>amyA</i>	608	α -Amylase – <i>A. dehalogenans</i>	ZP_0040 0688.1	82-531	47-463/545	38 9e-81
38	425	CBS Domänen Protein – <i>S. oneidensis</i>	NP_7183 90.1	1-425	1-427/439	65 e-154
<i>tonB</i>	209	TonB Protein – <i>P. syringae</i>	AAO5598 8.1	124-209	188- 273/273	51 8e-17
<i>exbB</i>	238	Biopolymer Transport Protein – <i>R. metallidurans</i>	ZP_0027 2239.1	6-221	5-221/243	62 1e-73
<i>exbD</i>	137	Biopolymer Transport Protein – <i>R. eutropha</i>	ZP_0016 8083.1	1-135	1-135/137	60 6e-42
42	107	hypoth Protein – <i>S. degradans</i>	ZP_0031 7246.1	1-104	1-104/109	69 7e-35
<i>gnuB</i>	525	Endoglukanase – <i>P. fluorescens</i>	P18126	1-524	1-510/511	60

						0.0
44	490	möglicher Transkriptions- regulator – <i>R. solanacearum</i>	NP_5212 40	1-487	1-484/488	49 e-121
45	219	konserv. hypoth. Protein – <i>R. solanacearum</i>	NP_5212 41.1	1-216	1-214/227	61 4e-69
46	349	integrales Membran-Protein – <i>C. violaceum</i>	NP_9017 49.1	1-341	1-341/341	50 4e-86
47	354	Ca ²⁺ /Na ⁺ Antiporter – <i>Crocospaera watsonii</i>	ZP_0017 7068.1	5-353	9-360/377	49 4e-86
48	99	Endoglukanase – <i>S. degradans</i>	ZP_0031 7934.1	1-98	305- 402/403	69 9e-40
49	155	hypoth. Protein – <i>Pirellula sp. 1</i>	NP_8646 04.1	1-148	390- 530/611	35 1e-20
<i>aguB</i>	449	beta-Agarase I – <i>S. degradans</i>	AAT6706 2.1	15-445	9-445/593	53 e-140
<i>aguA</i>	772	hypoth Protein – (β- Galaktosidase) <i>S. degradans</i>	ZP_0031 5251.1	10-772	1-769/769	53 0.0
52	355	COG2931: RTX Toxin u. verwandtes Ca ²⁺ - Bindeprot. – <i>Rubrivivax gelatinosus</i>	ZP_0024 5546.1	118-344	1167- 1388/1699	25 3e-05
53	208	Ähnlich zu RNA Polymerase Sigma Faktor <i>ncch</i> – <i>Pirellula sp. 1</i>	NP_8651 74.1	43-207	6-170/172	30 5e-12
54	179 Frag- ment	ABC-Typ Protease/Lipase Transport System – <i>Mesorhizobium sp. BNC1</i>	ZP_0019 5266.2	35-179	165- 309/742	85 4e-62

Die DNA-Sequenz wurde bei GenBank unter der Accession-Nummer AY212801 eingetragen. Identifizierten ORFs wurde ein Gen-Name zugeteilt, wenn der ermittelte E-Wert unter 10⁻⁸⁰ lag.

War ein Nachweis für das homologe Enzym vorhanden, so wurde der entsprechende ORF in der Tabelle dunkelgrau hinterlegt. Handelte es sich um eine automatische Annotation ohne funktionellen Nachweis wurde der jeweilige ORF hellgrau hinterlegt.

Tabelle 2: Annotation der auf XIV A3 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
61	398	Rekombinase, DNA-Invertase Pin Homologon – <i>C. hutchisonii</i>	ZP_00308018 .1	11-395	130-512/513	33 1e-47
62	304	mögliches GTP-Binde Protein – <i>Bacteroides fragilis</i>	CAH09580.1	1-302	1-302/305	59 5e-59
63	435	GTPase – <i>Bacteroides fragilis</i>	CAH09579.1	1-435	1-437/437	63 3e-160
64	619	hypoth Protein – <i>Nocardioides</i> sp. JS614	ZP_00657087 .1	1-539	52-572/649	26 2e-38
65	919	Rezeptor Protein (äußere Membran) – <i>Cytophaga hutchinsonii</i>	ZP_00309692 .1	8-905	7-928/929	26 1e-77
66	338	hypoth Protein DUF11 – <i>Deinococcus geothermalis</i>	ZP_00397476 .1	23-260	475-700/762	21 0.002
67	336	hypoth Protein	-	-	-	-
68	199	hypoth Protein	-	-	-	-
69	234	hypoth Protein	-	-	-	-
<i>actP</i>	962	ATPase – <i>A. tumefaciens</i>	NP_531635.1	20-914	11-854/861	51 0.0
<i>aguC</i>	451	hypoth. Protein – <i>S. degradans</i>	ZP_00315360 .1	1-448	294-743/747	65 0.0
72	785	hypoth. Protein – <i>S. degradans</i>	ZP_00315656	7-770	3-780/788	26 2e-67
<i>aguD</i>	449	beta-Agarase I – <i>S. degradans</i>	AAT67062.1	15-445	9-445/593	54 1e-139
<i>aguE</i>	772	hypoth. Protein (β -Galaktosidase) – <i>S. degradans</i>	ZP_00315251 .1	10-772	1-769/769	53 0.0
75	522	hypoth Protein – <i>C. crescentus</i>	NP422120.1	131-242	125-248/457	31
76	199	RNA Polymerase, σ^H – <i>X. axonopodis</i>	NP641654.1	23-158	15-184/206	31
77	760	hypoth. Protein – <i>Gibberella zeae</i>	NP625193.1	18-749	28-743/755	47

						0.0
78	392	Xylanase – <i>S. degradans</i>	ZP_00315659 .1	64-336	48-336/368	36 3e-46
79	558	Na ⁺ /solute Symporter – <i>Trichodesmium erythraeum</i>	ZP_00675040 .1	1-430	1-471/603	25 6e-37
80	409	Endoglukanase – <i>C. thermocellum</i>	CAC27412.2	20-177	424- 596/1321	32
<i>aguF</i>	580	Agarase – <i>Pseudomonas</i> sp. ND137	BAD88713.1	31-580	34-590/591	38 1e-92
82	236	Transkriptionsregula- tor – <i>X. campestris</i>	NP636966.1	12-207	115- 310/339	83 4e-93

Die DNA-Sequenz wurde bei GenBank unter der Accession-Nummer AY212800 eingetragen.

Tabelle 3: Annotation der auf III E5 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
<i>aguG</i>	457	hypoth. Protein (β -Galaktosidase) – <i>S. degradans</i>	ZP_003152 51.1	1-457	313-769/769	63 0.0
<i>aguH</i>	449	beta Agarase I. – <i>S. degradans</i>	AAT67062. 1	15-445	9-445/593	54 1e-139
093	785	hypoth. Protein – <i>S. degradans</i>	ZP_003156 56	7-770	3-780/788	26 2e-67
<i>agul</i>	451	hypoth. Protein. – <i>S. degradans</i>	ZP_003153 60	1-448	294-743/747	65 0.0
095	393	möglicher hydroxymethylpyrimidin Transporter – <i>P. fluorescens</i>	YP_257683 .1	41-392	78-429/430	91 1e-125
096	204	hypoth. Protein – <i>P. fluorescens</i>	ZP_002658 79.1	1-200	1-200/248	90 2e-102
097	205	NTP pyrophosphohydrolase – <i>P. fluorescens</i>	ZP_002658 80.1	1-205	3-207/207	91 1e-98
098	150	hypoth. Protein – <i>P. fluorescens</i>	ZP_002658 8.1	1-150	1-150/150	90 6e-76
099	238	mögliche Phosphohydrolase – <i>P. fluorescens</i>	ZP_002658 82.1	1-238	28-265/265	89 1e-124
100	325	hypoth. Protein – <i>S. degradans</i>	ZP_003152 42.1	4-277	5-245/412	30 5e-19
101	97	electron transport flavoprot, alpha SU – <i>Polaromonas sp.</i>	ZP_005095 44.1	1-97	217-313/313	96 1e-45
102	330	hypoth. Protein UPF0065 – <i>Polaromonas sp.</i>	ZP_005073 49.1	7-330	14-336/336	68 1e-107
103	347	Acyl-CoA DH verwandtes Prot. – <i>P. putida</i>	AAN68866. 1	30-347	31-347/347	46 9e-65
<i>amyB</i>	634	1,4-alpha-Glukan spaltendes Enzym – <i>Azoarcus sp.</i>	YP_161008 .1	28-627	23-625/636	59 0.0
106	311	Glucose-1-phosphat Adenylyltransferase – <i>Azoarcus sp.</i>	YP_160971 .1	5-297	5-297/404	67 2e-119

Die DNA-Sequenz wurde bei GenBank unter den Accession-Nummern AY236255 bzw. AY236226 eingetragen.

Tabelle 4: Annotation der auf X B11 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Aligne-ment region	Homologe Aligne-ment Region	Identität (%) / E-value
83	299 Fragment	mögliche Esterase – <i>S. degradans</i>	ZP_00316903. 1	1-281	8-289/380	60 1e-89
<i>aguJ</i>	772	hypoth. Protein (β- Galaktosidase) – <i>S. degradans</i>	ZP_00315251. 1	10-772	1-769/769	53 0.0
<i>aguK</i>	449	β-Agarase – <i>S. degradans</i>	AAT67062.1	15-445	9-445/593	54 1e-139
86	785	hypoth. Protein – <i>S. degradans</i>	ZP_00315656	7-770	3-780/788	26 2e-67
<i>aguL</i>	451	hypoth. Protein – <i>S. degradans</i>	ZP_00315360. 1	1-448	294- 743/747	64 0.0

Die DNA-Sequenz wurde bei GenBank unter der Accession-Nummer AY236223 eingetragen.

Tabelle 5: Annotation der auf III F2 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
110	277	hypoth. Protein. – <i>V. cholerae</i>	AAF94321.1	18-253	4-236/245	31
111	295	Lauroyl/myristoyl Acyltransferase – <i>S. degradans</i>	ZP_00318446.1	31-292	34-293/295	48 2e-65
112	188	COG0517: FOG: CBS-Domäne – <i>S. degradans</i>	ZP_00316481.1	1-178	1-178/187	32 7e-18
113	186	Cytochrom B561 – <i>S. degradans</i>	ZP_00318472.1	6-181	5-181/185	32 7e-20
114	207	hypoth. Protein – <i>S. degradans</i>	ZP00065671	17-198	48-201/209	24
115	241	hypoth. Protein – <i>S. degradans</i>	ZP00064837	25-240	33-241/243	21
<i>glyQ</i>	316	Glycyl-tRNA Synthetase, α UE – <i>S. degradans</i>	ZP_00318447.1	7-297	5-295/318	85 e-152
<i>glyS</i>	692	Glycyl-tRNA Synthetase, β UE – <i>S. degradans</i>	ZP_00318448.1	1-687	1-686/690	61 0.0
118	191	Histidinol-phosphatase – <i>Neisseria meningitidis MC58</i>	NP_275024.1	1-187	1-187/187	48 5e-45
119	234	mögliche Acyltransferase – <i>S. degradans</i>	ZP00065517.1	11-216	1-208/212	48
120	697	Pektat Lyase – <i>S. degradans</i>	ZP_00314795.1	85-695	72-680/680	46 4e-136
<i>peIA</i>	1085	Pektat Lyase – <i>S. degradans</i>	ZP_00314538	488-1085	725-1306/1306	46 1e-130
122	402	RtcB: konserviertes Protein – <i>S. degradans</i>	ZP_00318183.1	18-402	17-398/398	59 e-133
123	313	hypoth. Protein	-	-	-	-
124	119	konserviertes hypoth. Protein – <i>Vibrio parahaemolyticus</i>	NP_798177.1	1-118	1-118/119	75 9e-47
125	239	metallo beta-Lactamase – <i>P. putida</i>	CAD12765.1	1-239	1-239/246	59 2e-77

126	235	hypoth. Protein – <i>Burkholderia cepacia</i>	ZP_00214545.1	1-202	296-495/527	60 2e-69
127	325	Transmembranes Protein – <i>X. campestris</i>	NP_635882.1	48-324	15-296/297	43 1-46
128	97	hypoth. Protein – <i>S. meliloti</i>	NP436779	13-90	111-181/367	30
129	232	hypoth. Protein – <i>R. solanacearum</i>	NP518360	18-97	418-491/612	32
130	659	2-acyl- glycerophospho- ethanolamine Acyltransferase. – <i>X. campestris</i>	NP635885.1	29-659	1-640/640	73

Die DNA-Sequenz wurde bei GenBank unter der Accession-Nummer AY236224 eingetragen.

Tabelle 6: Annotation der auf VII H3 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	*ORF Align- ment Region	Homo- loge Alignment Region	Identität (%) / E-value
195	140 Frag- ment	hypoth. Protein – <i>S. meliloti</i>	CAC41738. 1	17-137	160- 280/290	65 8e-23
196	74	hypoth. Prot – <i>A. tumefaciens</i>	NP_39644 87.	1-74	1-74/74	87 1e-32
197	73	hypoth. Prot – <i>A. tumefaciens</i>	NP_39644 7.1	1-72	1-72/74	52 3e-08
198	392	Endonuklease/Exo- nuklease/Phosphata- se – <i>Frankia</i> sp.	ZP_005491 98.1	21-392	1-385/385	44 3e-74
199	416	Aminopeptidase – <i>A. tumefaciens</i>	NP_53098 0.1	10-416	6-412/412	85 0.0
200	157	hypoth. Prot – <i>A. tumefaciens</i>	NP_53098 1.1	1-157	1-157/158	78 3e-62
201	116	hypoth. Prot – <i>A. tumefaciens</i>	NP_53098 2.1	1-114	1-114/115	66 1e-38
202	186	Methyltransferase – <i>S. meliloti</i>	CAC49815. 1	4-186	2-183/183	53 6e-45
203	520	mögliche GMP	CAC41751.	1-520	1-520/520	91

		Synthase – <i>S. meliloti</i>	1			0.0
204	212	mögliche MTA/SAH Nukleosidase- <i>S. meliloti</i>	CAC41752. 1	1-208	1-208/212	69 4e-78
205	149	kons. hypoth. Protein – <i>S. meliloti</i>	CAC41753. 1	1-149	1-150/150	72 8e-55
206	150	Sensor Protein – <i>A. tumefaciens</i>	NP_53099 0.1	1-150	1-150/150	59 4e-45
207	358	hypoth. Protein – <i>S. meliloti</i>	CAC41755. 1	1-358	1-359/377	60 5e-124
208	368	hypoth. Protein – <i>A. tumefaciens</i>	NP_53099 2.1	31-333	36- 338/373	56 5e-83
209	508	hypoth. Protein – <i>M. loti</i>	NP_10617 7.1	32-508	57- 533/533	0.0
210	426	hypoth. Protein – <i>A. tumefaciens</i>	NP_53099 4.1	1-426	1-425/425	77 4e-174
211	432	Sun Protein – <i>A. tumefaciens</i>	NP_53099 5.1	3-430	1-428/429	76 0.0
212	72	Lipoprotein – <i>A. tumefaciens</i>	NP_53099 6.1	1-71	49- 119/122	78 3e-25
213	239	hypo Prot – <i>A. tumefaciens</i>	NP_53099 7.1	1-289	1-288/290	65 3e-96
214	422	hypoth. 4FE-4S Ferredoxin – <i>S. meliloti</i>	CAC41763. 1	40-394	4-357/366	75 5e-159
215	230	Glutathion S- Transferase – <i>S. meliloti</i>	CAC41764. 1	1-230	1-230/230	78 4e-97
<i>bacA</i>	268	Bacitracin Resistenz Protein – <i>A. tumefaciens</i>	NP_53100 0.1	1-268	1-268/268	78 2e-110

Tabelle 7: Annotation der auf XIV B3 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Align- ment region	Homo- loge Alignment Region	Identität (%) / E-value
162	429	Enolase – <i>P. aeruginosa</i>	NP_252325 .1	1-428	1-728/429	81 1e-177
163	238	4-diphosphocytidyl- 2C-methyl-D- erythritol Synthase – <i>Azotobacter vinelandii</i>	ZP_004179 82.1	5-225	2-224/236	51 3e-58

164	159	2C-methyl-D-erythritol 2,4-cyclodiphosphat Synthase – <i>P. aeruginosa</i>	ZP_002050 48.1	3-156	1-154/157	72 1e-49
165	326	hypoth. Protein – <i>Desulfovibrio vulgaris</i>	YP_011446 .1	1-325	1-324/325	36 5e-49
166	131	konserviertes Protein – <i>S. degradans</i>	ZP_003151 77.1	1-131	198-329/329	49 4e-26
167	223	Protein-L-isoaspartat carboxymethyltransferase – <i>S. degradans</i>	ZP_003151 76.1	11-223	1-213/213	69 4e-82
168	274	Lipoprotein – <i>P. chloramphis</i>	AAT46074. 1	17-274	39-294/294	50 3e-64
169	333	RNA Polymerase – <i>S. degradans</i>	ZP_003151 73.1	1-331	1-330/333	74 1e-132
170	359	DNA-3-methyladenin Glycosidase II – <i>Azoarcus</i> sp.	YP_161104 .1	50-253	1-294/300	51 4e-82
171	321	bifunktionales regulator. Protein – <i>Y. pestis</i>	NP_667414 .1	7-315	2-345/354	53 8e-84
172	516	Apolipoprotein acetyltransferase – <i>S. degradans</i>	ZP_003178 19.1	25-491	25-496/521	42 9e-108
173	285	möglicher Mg ²⁺ /Co ²⁺ Transporter CorC – <i>S. degradans</i>	ZP_003178 18.1	1-281	1-280/292	63 3e-92
174	155	mögliche Metall-abhängige Hydrolase – <i>S. degradans</i>	ZP_003178 17.1	1-149	2-149	58 2e-44
175	251	phosphate starvation-inducible protein PhoH – <i>S. degradans</i>	ZP_003178 16.1	8-249	97-338/364	77 9e-103
176	97	konserviertes Protein – <i>S. degradans</i>	ZP_003181 86.1	26-94	15-83/83	34 9e-06
177	446	Antwort Regulator beinhaltet CheY-artigen Receiver – <i>S. degradans</i>	ZP_003181 95.1	24-444	28-449/452	59 2e-138
178	483	Signal Transduktion Histidine Kinase – <i>S. degradans</i>	ZP_003181 93.1	18-473	36-490/505	39 1e-81
179	147	CheY-artiger	ZP_003181	1-147	1-147/152	71

		Rezeptor – <i>S. degradans</i>	92.1			6e-50
180	647	Signal Transduktion Histidine Kinase – <i>S. degradans</i>	ZP_003181 91.1	45-647	46- 658/666	44 5e-108
181	457	2-methylthioadenine Synthetase – <i>S. degradans</i>	ZP_003178 15.1	2-452	46- 495/496	80 0.0
182	109	konserviertes hypoth. Protein – <i>Polaromonas</i> sp.	ZP_005054 07.1	1-94	1-94/105	50 9e-14
183	466	Helikase – <i>Shewanella</i> <i>denitificans</i>	ZP_006348 42.1	1-390	1-392/433	53 6e-108
184	114	konserviertes hypoth. Protein – <i>S. degradans</i>	ZP_003178 14.1	1-110	1-197/109	68 2e-29
185	487	Glycerol Kinase – <i>S. degradans</i>	ZP_003166 47.1	1-452	1-457/491	31 2e-40
186	401	hypoth. Protein – <i>B. japonicum</i>	NP_768931 .1	67-383	144- 479/532	30 4e-28
187	148	Nuklease – <i>Crocospaera</i> <i>watsonii</i>	ZP_005168 82.1	11-133	13- 154/319	31 1e-09
188	159	mögliche Transposase – <i>Magnetospirillum</i> <i>gryphswaldense</i>	AAX11192. 1	50-124	8-82/124	44 1e-10
189		Adenin- phosphoribosyltrans- ferase – <i>X. anoxopodis</i>	AAM37243. 1	4-172	18-186	62 2e-50
190	172	konserviertes hypoth. Protein – <i>B. cenocepacia</i>	ZP_004632 83.1	5-169	3-166/172	59 3e-48
191	454	Sigma 70 Region 2 – <i>Anaeromyxobacter</i> <i>dehalogenans</i>	ZP_003998 38.1	40-417	8-385	72 2e-110
192	137	hypoth. Protein – <i>P. aeruginosa</i>	NP_25004. 1	1-136	1-136/137	55 4e-42
193	128	Protein unbek. Funktion – <i>P. aeruginosa</i>	AAV38040. 1	6-117	14- 121/122	32 2e-07
194	93	mögliche nukleosid- diphosphat Zucker Epimerase – <i>P. fluorescens</i>	ZP_002638 59.1	1-90	207- 296/300	57 4e-22

Tabelle 8: Annotation der auf VII A1 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Alignment Region	Homologe Alignment Region	Identität (%) / E-value
131	223	ATPase (DNA Reparatur) – <i>P. fluorescens</i> Pf0-1	ZP_00267005.1	26-223	2-199/199	85 6e-79
132	248	hypoth. Protein – <i>P. fluorescens</i> Pf0-1	ZP_00267004.1	1-248	1-248/248	85 3e-122
133	164	DNA Reparatur – <i>P. fluorescens</i> Pf0-1	ZP_00267003.1	1-164	1-164/164	91 2e-80
134	193	hypoth. Protein – <i>P. fluorescens</i> Pf0-1	ZP_00267002.1	1-197	1-197/197	83 1e-83
135	100	hypoth. Protein – <i>Chromobacterium violaceum</i>	AAQ58888.1	2-98	17-113/115	43 2e-17
136	655	Topoisomerase A – <i>P. fluorescens</i>	ZP_00267001.1	1-654	1-654/655	88 0.0
137	113	hypoth. Protein – <i>P. fluorescens</i> Pf0-1	ZP_00266999.1	1-112	1-112/113	89 9e-45
138	210	hypoth. Protein – <i>P. fluorescens</i> Pf0-1	ZP_00266998.1	1-210	42-251/251	89 6e-92
139	438	hypoth. Protein – <i>P. fluorescens</i>	ZP_00266997.1	43-437	1-395/396	82 e-176
140	182	hypoth. Protein – <i>P. fluorescens</i> Pf0-1	ZP_00266996.1	1-182	1-182/182	90 1e-78
141	493	Transkriptionaler Regulator – <i>P. fluorescens</i> Pf0-1	ZP_00266995.1	1-493	1-493/493	76 0.0
142	311	ATPases (Chromosom Teilung) – <i>P. fluorescens</i> Pf0-1	ZP_00266994	127-311	1-185/185	81 1e-76
143	255	hypoth. Protein – <i>Nitrosomonas europaea</i>	ZP_00669664	14-252	6-245/249	41 4e-42
144	837	UvrD/REP Helikase – <i>Magnetococcus</i> sp. MC-1	ZP_00605915	32-806	12-761/1155	41 6e-155
145	906	hypoth. protein – <i>N. europaea</i>	NP_842182.1	7-896	17-907/909	30 1e-97
146	471	Hexuronat Transporter – <i>Bacteroides fragilis</i>	YP_098170	3-468	21-494/497	52 2e-116

147	170	COG0517: FOG: CBS Domäne – <i>S. degradans</i>	ZP_00315283. 1	14-159	23- 169/171	41 2e-24
148	275	Glutamat Razemase – <i>Salmonella</i> <i>tyohimurium</i>	AAL22969	11-266	22- 274/283	51 5e-65
149	520	Protein konserviert in Bakterien – <i>S. degradans</i>	ZP_00314612. 1	84-490	50- 453/494	50 e-100
150	157	hypoth. Protein – <i>S.</i> <i>degradans</i>	ZP_00314611. 1	6-156	8-154/161	43 1e-16
151	172	Acetyltransferase, GNAT Familie – <i>Shewanella</i> <i>oneidensis</i>	NP_716994.1	8-170	9-166/171	28 9e-11
152	366	tRNA (uracil-5-)- Methyltransferase – <i>Vibrio</i> <i>parahaemolyticus</i>	NP_799318.1	5-366	1-368/369	53 e-112
153	468	Signal transd. Histidin Kinase – <i>S. degradans</i>	ZP_00316014. 1	67-467	46- 460/462	42 3e-84

Tabelle 9: Annotation der auf VII D9 identifizierten ORFs

ORF	Größe des Proteins (*AS)	Funktion, nächster Treffer	GenPept Nummer des nächsten Treffers	+ORF Aligment region	Homologe Aligment Region	Identität(%) / E-value
154	696	Kälte Schock <i>dead-box</i> Protein A – <i>A. tumefaciens</i>	AAK89264.1	1-601	1-601/615	67 0.0
155	289	Membran Protein – <i>Mesorhizobium</i> sp. BNC1	ZP_00194018.2	11-289	9-281/281	52 5e-77
156	136	hypoth. Protein – <i>A. tumefaciens</i>	NP_534623.1	1-136	1-136/136	36 3e-19
157	292	hypoth. Protein – <i>Rhizobium leguminosarum</i>	AAL83832.1	83-270	67-255/287	46 1e-38
158	222	ABC Transporter – <i>A. tumefaciens</i>	NP_531310.1	1-195	1-194/365	74 3e-80
159	475	Transkriptionaler Regulator (XRE-Familie) – <i>A. tumefaciens</i>	NP_531309.1	1-475	1-471/471	72 0.0
160	429	Isocitrat Lyase – <i>A. tumefaciens</i>	NP_531308.1	1-429	6-434/434	91 0.0
161	291	Transkriptions Regulator) – <i>S. meliloti</i>	NP_384816.1	1-287	1-286/291	60 2e-97

		10	20	30	40	50			
AguA	MKKSQYLNHI	LLCAISLFLI	AACSEQK---	---HNNNKNE	RANLNEIKES	ILLDFENEND			
AguE	MKKSQYLNHI	LLCAISLFLI	AACSEQK---	---HNNNKNE	RANLNEIKES	ILLDFENEND			
AguG	-----	-----	-----	-----	-----	-----			
AguJ	MKKSQYLNHI	LLCAISLFLI	AACSEQK---	---HNNNKNE	RANLNEIKES	ILLDFENEND			
Konsensus									

		70	80	90	100	110			
AguA	IKVVELKNSI	SQIIVNEGNH	SLEINFFSLN	NHDSSFEIKP	SNAWNFSDYE	TVALALDIEN			
AguE	IKVVELKNSI	SQIIVNEGNH	SLEINFFSLN	NHDSSFEIKP	SNAWNFSDYE	TVALALDIEN			
AguG	-----	-----	-----	-----	-----	-----			
AguJ	IKVVELKNSI	SQIIVNEGNH	SLEINFFSLN	NHDSSFEIKP	SNAWNFSDYE	TVALALDIEN			
Konsensus									

		130	140	150	160	170			
AguA	PTTSSAHIYV	YTYDKNGAFQ	LSNVAVPAES	THSYLIELKV	PSLLLNTGIR	NDPPSWPHNY			
AguE	PTTSSAHIYV	YTYDKNGAFQ	LSNVAVPAES	THSYLIELKV	PSLLLNTGIR	NDPPSWPHNY			
AguG	-----	-----	-----	-----	-----	-----			
AguJ	PTTSSAHIYV	YTYDKNGAFQ	LSNVAVPAES	THSYLIELKV	PSLLLNTGIR	NDPPSWPHNY			
Konsensus									

		190	200	210	220	230			
AguA	ISTIWRGGTK	TLDTSAITSV	RL LISGVLEN	KQLIIDNLRL	VEPDNFDSAY	LVGLVDEFGQ			
AguE	ISTIWRGGTK	TLDTSAITSV	RL LISGVLEN	KQLIIDNLRL	VEPDNFDSAY	LVGLVDEFGQ			
AguG	-----	-----	-----	-----	-----	-----			
AguJ	ISTIWRGGTK	TLDTSAITSV	RL LISGVLEN	KQLIIDNLRL	VEPDNFDSAY	LVGLVDEFGQ			
Konsensus									

		250	260	270	280	290			
AguA	NAKQDFVGKV	HSTEQRLRSL	RQEQKNLTDQ	QFPARSTFNG	WINGPKLTAT	GYFRTEKYDG			
AguE	NAKQDFVGKV	HSTEQRLRSL	RQEQKNLTDQ	QFPARSTFNG	WINGPKLTAT	GYFRTEKYDG			
AguG	-----	-----	-----	-----	-----	-----			
AguJ	NAKQDFVGKV	HSTEQRLRSL	RQEQKNLTDQ	QFPARSTFNG	WINGPKLTAT	GYFRTEKYDG			
Konsensus									

		310	320	330	340	350			
AguA	KWTLVDPEGY	LFFSNGLANI	RMANTSTITG	YDFDAALIKQ	RDADDYTPED	SIGLNTAPKS			

	670	680	690	700	710				
AguA	GIDQPYWDFL	AELDKPSIIG	EFHNGAIDSG	LLNPGLIHAE	SQFDRGEKYK	SYLNSVIDNP			
AguE	GIDQPYWDFL	AELDKPSIIG	EFHNGAIDSG	LLNPGLIHAE	SQFDRGEKYK	SYLNSVIDNP			
AguG	GIDQPYWDFL	AELDKPSIIG	EFHNGAIDSG	LLNPGLIHAE	SQFDRGEKYK	SYLNSVIDNP			
AguJ	GIDQPYWDFL	AELDKPSIIG	EFHNGAIDSG	LLNPGLIHAE	SQFDRGEKYK	SYLNSVIDNP			
Konsensus	*****	*****	*****	*****	*****	*****			

	730	740	750	760	770				
AguA	YFVGAHWFQY	IDSPLTGRAY	DGENYNVGFV	SVADIPYPPL	VKAAQEVNRN	LYQKRFGK			
AguE	YFVGAHWFQY	IDSPLTGRAY	DGENYNVGFV	SVADIPYPPL	VKAAQEVNRN	LYQKRFGK			
AguG	YFVGAHWFQY	IDSPLTGRAY	DGENYNVGFV	SVADIPYPPL	VKAAQEVNRN	LYQKRFGK			
AguJ	YFVGAHWFQY	IDSPLTGRAY	DGENYNVGFV	SVADIPYPPL	VKAAQEVNRN	LYQKRFGK			
Konsensus	*****	*****	*****	*****	*****	*****			

Abb. 2: Aminosäuresequenz-Alignment der Agarasen AguA, AguE, AguJ und AguG.

		10	20	30	40	50			
AguB	MKKHISCCVA	LMLTSVSVFA	ADWDSVPIPA	PAGAGKTWQL	QTISDEFNYN	ASPTVKPAAF			
AguD	MKKHISCCVA	LMLTSVSVFA	ADWDSVPIPA	PAGAGKTWQL	QTISDEFNYN	ASPTVKPAAF			
AguK	MKKHISCCVA	LMLTSVSVFA	ADWDSVPIPA	PAGAGKTWQL	QTISDEFNYN	ASPTVKPAAF			
AguH	MKKHISCCVA	LMLTSVSVFA	ADWDSVPIPA	PAGAGKTWQL	QTISDEFNYN	ASPTVKPAAF			
Konsensus	*****	*****	*****	*****	*****	*****			

		70	80	90	100	110			
AguB	NNRWNASYIN	AWTGPGDTEF	SAGHSYTTGG	SLALQAAAKP	GTNKIFAGII	SSKQTFITYPL			
AguD	NNRWNASYIN	AWTGPGDTEF	SAGHSYTTGG	SLALQAAAKP	GTNKIFAGII	SSKQTFITYPL			
AguK	NNRWNASYIN	AWTGPGDTEF	SAGHSYTTGG	SLALQAAAKP	GTNKIFAGII	SSKQTFITYPL			
AguH	NNRWNASYIN	AWTGPGDTEF	SAGHSYTTGG	SLALQAAAKP	GTNKIFAGII	SSKQTFITYPL			
Konsensus	*****	*****	*****	*****	*****	*****			

		130	140	150	160	170			
AguB	YLEARVKHTN	NTMANAVWML	SADSTQEIDA	MESYGSDRAG	QEWFDQRMHV	SHHLFIREPF			
AguD	YLEARVKHTN	NTMANAVWML	SADSTQEIDA	MESYGSDRAG	QEWFDQRMHV	SHHLFIREPF			
AguK	YLEARVKHTN	NTMANAVWML	SADSTQEIDA	MESYGSDRAG	QEWFDQRMHV	SHHLFIREPF			
AguH	YLEARVKHTN	NTMANAVWML	SADSTQEIDA	MESYGSDRAG	QEWFDQRMHV	SHHLFIREPF			
Konsensus	*****	*****	*****	*****	*****	*****			

		190	200	210	220	230			
AguB	QDYQPKDEGS	WIFNPHGTWR	GAMHTYGVHW	KDPWNLEYYI	DGVKVRTVSG	PAMIDPNNFT			
AguD	QDYQPKDEGS	WIFNPHGTWR	GAMHTYGVHW	KDPWNLEYYI	DGVKVRTVSG	PAMIDPNNFT			
AguK	QDYQPKDEGS	WIFNPHGTWR	GAMHTYGVHW	KDPWNLEYYI	DGVKVRTVSG	PAMIDPNNFT			
AguH	QDYQPKDEGS	WIFNPHGTWR	GAMHTYGVHW	KDPWNLEYYI	DGVKVRTVSG	PAMIDPNNFT			
Konsensus	*****	*****	*****	*****	*****	*****			

		250	260	270	280	290			
AguB	NGTGINKPLH	IIIDVEHQDW	RDVRPTAAEL	ADTNKSIMFV	DWIRVYKPA	ATTNSSASNS			
AguD	NGTGINKPLH	IIIDVEHQDW	RDVRPTAAEL	ADTNKSIMFV	DWIRVYKPA	ATTNSSASNS			
AguK	NGTGINKPLH	IIIDVEHQDW	RDVRPTAAEL	ADTNKSIMFV	DWIRVYKPA	ATTNSSASNS			
AguH	NGTGINKPLH	IIIDVEHQDW	RDVRPTAAEL	ADTNKSIMFV	DWIRVYKPA	ATTNSSASNS			
Konsensus	*****	*****	*****	*****	*****	*****			

		310	320	330	340	350			
AguB	STASSTPSNT	VQLIDFANYF	DTGKSTASVA	GDTYVGFNKS	GNGNINYNTA	GDWGDYLVTL			

```

AguD      STASSTPSNT VQLIDFANYF DTGKSTASVA GDTYVGFNKS GNGNINYNTA GDWGDYLVTL
AguK      STASSTPSNT VQLIDFANYF DTGKSTASVA GDTYVGFNKS GNGNINYNTA GDWGDYLVTL
AguH      STASSTPSNT VQLIDFANYF DTGKSTASVA GDTYVGFNKS GNGNINYNTA GDWGDYLVTL
Konsensus *****
                                     .....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
                                     370      380      390      400      410

AguB      PSDGKYKFEI ITASPMATNG IGAKLIIDGI YVGTINVGST GGWEVYAPFA LENNISIGAG
AguD      PSDGKYKFEI ITASPMATNG IGAKLIIDGI YVGTINVGST GGWEVYAPFA LENNISIGAG
AguK      PSDGKYKFEI ITASPMATNG IGAKLIIDGI YVGTINVGST GGWEVYAPFA LENNISIGAG
AguH      PSDGKYKFEI ITASPMATNG IGAKLIIDGI YVGTINVGST GGWEVYAPFA LENNISIGAG
Konsensus *****
                                     .....|.....| .....|.....| .....|.....
                                     430      440

AguB      THTIRIESVG SSAWQWNGDQ IRVTRVGSL
AguD      THTIRIESVG SSAWQWNGDQ IRVTRVGSL
AguK      THTIRIESVG SSAWQWNGDQ IRVTRVGSL
AguH      THTIRIESVG SSAWQWNGDQ IRVTRVGSL
Konsensus *****

```

Abb. 3: Aminosäure-Alignment der Agarasen AguB, AguD, AguH und AguK.

	370	380	390	400	410	
AguC	ASGLFHPGLV	HAANQEDRAK	MYKEYMQTVI	DNPYFIGAHW	FQYMDSPVTG	RSYDGENYNV
AguI	ASGLFHPGLV	HAANQEDRAK	MYKEYMQTVI	DNPYFIGAHW	FQYMDSPVTG	RSYDGENYNV
AguL	ASGLFHPGLV	HAANQEDRAK	MYKEYMQTVI	DNPYFIGAHW	FQYMDSPVTG	RSYDGENYNV
Konsensus	*****	*****	*****	*****	*****	*****

	430	440	450			
AguC	GFVSVTDTPY	APMVKAAKEL	HREMYTRRAK	K		
AguI	GFVSVTDTPY	APMVKAAKEL	HREMYTRRAK	K		
AguL	GFVSVTDTPY	APMVKAAKEL	HREMYTRRAK	K		
Konsensus	*****	*****	*****	*		

Abb. 4: Aminosäure-Alignment der Agarasen AguC, Agul und AguL.

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Wolfgang Streit für die, aller Unkenrufe zum Trotz, Aufnahme in seine Arbeitsgruppe und die Ermöglichung der selbständigen wissenschaftlichen Arbeit an einem interessanten Thema.

Bei Herrn Prof. Dr. Liebl möchte ich mich für die Aufnahme in seine Abteilung bedanken.

Herrn PD Dr. Rolf Daniel danke ich für die unkomplizierte Übernahme des Korreferats meiner Arbeit.

Ein großes Dankeschön gilt selbstverständlich allen Mitgliedern der Labore 37/42, die mich so freundlich in ihren Reihen aufgenommen haben und ohne die die Zeit nicht halb so spannend, diskussionsfreudig, ereignisreich und unterhaltsam gewesen wäre. Hierzu zählen besonders Alex Uesbeck, Anika Appelles, Ute Ludwig, Ole Fütterer, Mladen Tzvetkov, Mechthild Bömeke, Susanne Keese, Christina Schipper, Sabrina Koßmehl, Jessica Jabin, Steve Klopffleisch, Inga Schulte, Angel Angelov, Jörn Voss, Volker Hoffmann, Andrea Türmer, Meike Ballschmiter, Jörg Kleine, Judith Herath, und Silja Wittenberg.

Jetzt schon Vergangenheit und schon schmerzlich vermisst ist das „Dancing Lab“ mit seinen langjährigen Bewohnern Christel Schmeißer, Christian Elend und Martin Armbrrecht. Wenn lachen lebensverlängernd wirkt werde ich bestimmt steinalt!

Ein extra großes Dankeschön geht an Jörn Voss, der mir immer wieder geduldig die Mysterien der Protein-Aufreinigung und der ÄKTA erklärt hat und ohne dessen Hilfe ein Großteil meiner Arbeit nicht möglich gewesen wäre. Danke auch an Meike, Jörg, Angel und Volker, die ebenso geduldig meine Fragen zu der Welt im allgemeinen und Proteinen im besonderen beantwortet haben.

Bedanken möchte ich auch bei Marc Hoffmeister und Anne Krappmann, die mich schon mein ganzen Studium begleiten und immer zu Seite standen, besonders in den dunklen Zeiten.

Ohne die „unsichtbare“ Hilfe von vielen fleißigen Händen wäre diese Arbeit so nicht möglich gewesen. Hierzu zählen das Team vom G₂L, die die Sequenzierung der Plasmid-Banken durchgeführt haben, Herr Dr. Heiko Liesegang, der auch Windows-Usern die Welt von unix erklären kann, und natürlich die Werkstatt mit Gerd Birke, Olaf Waase und Patrick Regin, die immer zur Stelle sind wenns brennt.

Das Schönste was mir während meines Studiums „passiert“ ist, ist allerdings mein Freund Christian. Für keinen Titel der Welt würde ich ihn eintauschen! Danke dafür!

Lebenslauf

23.02.1976	Geburt in Hamburg
1982-1986	Besuch der Grundschule Quellmoor, Hamburg
1986-1990	Besuch des Friedrich-Ebert-Gymnasiums, Hamburg
1990-1995	Besuchs des Gymnasiums Hochrad, Hamburg
06/1995	Abitur
09/1996	Immatrikulation an der Georg-August-Universität zu Göttingen im Studienfach Biologie (Diplom)
04/1999	Vordiplom in den Fächern Mikrobiologie, Zoologie, Chemie und Physikalische Chemie
02/2001	Diplomprüfung in den Fächern Mikrobiologie, Immunologie und Genetik
04/2001-12/2001	Anfertigung der experimentellen Diplomarbeit unter Anleitung von PD Dr. Uwe Deppenmeier mit dem Titel: Untersuchungen zum DNA-Transfer und zur Biosynthese von Zellbausteinen in <i>Methanosarcina mazei</i> Gö1
04/2002	Beginn der experimentellen Arbeiten zur vorliegenden Dissertation