**Identifizierung und Charakterisierung von Genen
und korrespondierenden Genprodukten aus
Metagenombanken, die Butanol-Dehydrogenase-
Aktivität oder proteolytische Aktivität vermitteln**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Tanja Waschkowitz
aus Kassel

Göttingen 2006

D 7

Referent: PD Dr. R. Daniel

Korreferent: Prof. Dr. W. Liebl

Tag der mündlichen Prüfung: 03.05.2006

*Meiner Mutter und im Andenken
an meinen Vater gewidmet*

INHALTSVERZEICHNIS

Inhaltsverzeichnis	I
Abkürzungsverzeichnis	VI
1 Einleitung	1
1.1 Erschließung und Analyse der mikrobiellen Diversität durch Metagenomik	1
1.2 Butanol, industrielle Bedeutung und Produktion	4
1.3 Klassifizierung, Funktionen und industrielle Anwendung von Proteasen	7
1.4 Zielsetzung der Arbeit	10
2 Material und Methoden	11
2.1 Verwendete Organismen, Plasmide und Oligonukleotide	11
2.2 Nährmedien	13
2.2.1 LB-(Luria-Bertani)-Medium (SAMBROOK et al., 1989)	13
2.2.2 SOB-Medium (HANAHAN, 1983)	14
2.2.3 Denitrifikanten-Anreicherungsmedium	14
2.2.4 Medium für Aldehyd-Indikator-Platten-Test (CONWAY et al., 1987, modifiziert)	14
2.2.5 M9-Mineralmedium (MILLER, 1972, modifiziert)	15
2.2.6 Protease-Expressionsmedium	15
2.2.7 Spurenelementelösung SL-4 (PFENNIG und LIPPERT, 1966)	16
2.2.8 Spurenelementelösung SL-10 (WIDDEL et al., 1983)	16
2.2.9 Vitaminlösung (WOLIN et al., 1964, modifiziert)	17
2.2.10 Antibiotika und sonstige Medienzusätze	17
2.3 Zellanzucht	18
2.3.1 Aerobe Anzucht	18
2.3.2 Anaerobe Anzucht	18
2.3.3 Stammhaltung	18
2.3.4 Anzucht von rekombinanten <i>E. coli</i> -Stämmen zur heterologen Expression	19
2.3.4.1 Expression mit dem „pET Directional TOPO Expression“-Kit	19
2.4 Bestimmung von Wachstumsparametern	21
2.4.1 Bestimmung der optischen Dichte	21
2.4.2 Bestimmung des pH-Wertes	21
2.5 Standardtechniken für das Arbeiten mit Nukleinsäuren	22
2.5.1 Behandlung von Geräten und Lösungen	22

2.5.2	Fällung, Reinigung und Konzentration von Nukleinsäuren	22
2.5.2.1	Fällung von Nukleinsäuren	22
2.5.2.2	DNA-Fällung nach Saccharose-Dichtegradienten-Zentrifugation	22
2.5.2.3	Mikrodialyse von DNA	23
2.5.2.4	Dialyse von Umwelt-DNA	23
2.5.2.5	Konzentrationsbestimmung von Nukleinsäuren	24
2.5.2.6	Isolierung von DNA-Fragmenten mittels „QIAquick Gel Extraction“-Kit	24
2.5.2.7	Reinigung von DNA-Fragmenten mittels „QIAquick PCR Purification“-Kit	25
2.6	Isolierung von Nukleinsäuren	25
2.6.1	Isolierung von DNA aus Umwelt-Proben (ZHOU et al., 1996)	25
2.6.2	Reinigung von Umwelt-DNA (HENNE et al., 1999)	26
2.6.3	Isolierung von Plasmid-DNA	27
2.6.3.1	Plasmidschnellpräparation (HOLMES und QUIGLEY, 1981)	27
2.6.3.2	Minipräparation von Plasmiden mittels „QIAprep Spin Miniprep“-Kit	28
2.6.3.3	Minipräparation von Plasmiden mittels „NucleoSpinPlasmid“-Kit	28
2.6.3.4	Präparative Plasmidisolierung über Anionenaustauschchromatographie	29
2.7	Gelelektrophorese von Nukleinsäuren	29
2.7.1	Standard-Agarose-Gelelektrophorese	29
2.7.2	Größenbestimmung von Nukleinsäuren	30
2.8	Enzymatische Modifikation von DNA	31
2.8.1	Schneiden von DNA durch Restriktionsendonukleasen	31
2.8.2	Dephosphorylierung von DNA-Fragmenten	32
2.8.3	Ligation von DNA-Fragmenten	32
2.8.3.1	Ligation mit der T4-DNA-Ligase	32
2.8.3.2	Ligation mit dem „TOPO-T/A-Cloning“-Kit	32
2.9	Saccharose-Dichtegradienten-Zentrifugation (SAMBROOK et al., 1989, modifiziert)	34
2.10	Übertragung von DNA in <i>E. coli</i> und Selektion rekombinanter Klone	34
2.10.1	Hocheffiziente Transformation in <i>E. coli</i> (INOUE et al., 1990)	34
2.10.2	Transformation in <i>E. coli</i> durch Elektroporation (DOWER et al., 1988)	35
2.10.3	Der X-Gal-Test zur Selektion rekombinanter Klone	36
2.10.4	Screening auf die Verwertung von Butanol	37
2.10.5	Plattentest zum Screening auf proteolytische Aktivität	38
2.11	Polymerase-Kettenreaktion (PCR)	38
2.12	Sequenzierung und Analyse von Sequenzdaten	40
2.12.1	Durchführung von DNA-Sequenzierungen	40
2.12.2	Sequenzierungsprojekte pTWP9 und pTWP13	40
2.12.3	Prozessierung von Rohsequenzen	40

2.12.4	„Pregap4“ und „Gap4“	41
2.12.5	Methoden zur Schließung von Sequenzlücken	42
2.12.5.1	Identifikation von „Contig“-Überlappungen mit „Gap4“ oder „Align two Sequences“ (NCBI)	42
2.12.5.2	„Primer-Walking“	43
2.12.6	Sequenzierungsprojekte pTWB3 und pTWB4	44
2.12.7	Sequenzanalyse	44
2.12.7.1	ORF („open reading frame“)-Vorhersage	45
2.12.7.2	Annotation	46
2.13	Standardtechniken für das Arbeiten mit Proteinen	46
2.13.1	Herstellung von Rohextrakten	46
2.13.2	Proteinbestimmung	47
2.13.3	SDS-Polyacrylamid-Gelelektrophorese (PAGE)	48
2.13.4	Proteinfärbung in SDS-Polyacrylamidgelen	50
2.13.4.1	Coomassie-Färbung	50
2.13.4.2	Silberfärbung (BLUM et al., 1987)	50
2.13.5	Bestimmung der Molekularmassen von Proteinen	52
2.13.6	Konzentrierung und Umpufferung von Proteinlösungen	53
2.13.6.1	Konzentrierung und Umpufferung mit Hilfe von Vivaspin Konzentratoren	53
2.13.6.2	Konzentrierung und Umpufferung mit dem Vivaflow 200 Komplettsystem	53
2.13.7	Entsalzen von Proteinlösungen	54
2.14	Western Blotting (TOWBIN et al., 1979, modifiziert)	54
2.14.1	Immunodetektion von Proteinen auf Blot-Membranen	55
2.15	Methoden zur Proteianreicherung bzw. Reinigung	57
2.15.1	Metall-Chelat-Affinitätschromatographie	57
2.15.1.1	Vorbereitung der Säule	58
2.15.1.2	Reinigung löslicher Proteine unter nativen Bedingungen	58
2.15.1.3	Regeneration der Säule	59
2.15.2	Hydrophobe Interaktionschromatographie (HIC)	59
2.16	Bestimmung von Enzymaktivitäten	61
2.16.1	Bestimmung der proteolytischen Aktivität	61
2.17	Biochemische Charakterisierung der Proteasen	64
2.17.1	Bestimmung der Substratspezifität	64
2.17.2	Bestimmung der proteolytischen Aktivität in Abhängigkeit von der Temperatur	67
2.17.3	Bestimmung der proteolytischen Aktivität in Abhängigkeit vom pH-Wert	67
2.17.4	Bestimmung der Thermostabilität	67
2.17.5	Einfluss von Inhibitoren auf die proteolytische Aktivität	68
2.18	Chemikalien, Enzyme und Gase	68
3	Experimente und Ergebnisse	70
3.1	Konstruktion von Metagenombanken in <i>Escherichia coli</i>	72

3.1.1	Isolierung und Reinigung von DNA aus Umwelt-Proben	72
3.1.2	Enzymatische Modifikationen der isolierten Umwelt-DNA	74
3.1.3	Klonierung der isolierten Umwelt-DNA in Vektoren und Transformation in <i>Escherichia coli</i>	76
3.1.4	Entwicklung einer Methode zur Klonierung von Umwelt-DNA mit Hilfe des TOPO-T/A-Systems	77
3.1.5	Konstruktion und Charakterisierung von Metagenombanken	80
3.2	Durchmusterung von Metagenombanken nach Genen, die Butanol-Dehydrogenase-Aktivität vermitteln	82
3.2.1	Screening-Strategie	83
3.2.2	Durchführung und Ergebnisse des Screenings zur Identifizierung potentieller Butanol-Dehydrogenasen	85
3.2.3	Molekulare Analyse der Plasmide pTWB3 und pTWB4	86
3.2.4	Charakterisierung der Butanol-Dehydrogenase-Aktivität von <i>E. coli</i> /WL3/pTWB3 und <i>E. coli</i> /WL3/pTWB4	94
3.3	Durchmusterung von Metagenombanken nach Genen, die proteolytische Aktivität vermitteln	98
3.3.1	Screening-Strategie	98
3.3.2	Ergebnisse des Screenings nach Genen die proteolytische Aktivität vermitteln	99
3.3.3	Molekulare Analyse der Plasmide pTWP9 und pTWP13	101
3.3.3.1	Sequenzierung der Inserts von pTWP9 und pTWP13	101
3.3.3.2	Festlegung von offenen Leserahmen und Annotation der Gene	102
3.3.4	Molekulare Charakterisierung von <i>mprA</i> und <i>mprB</i>	116
3.4	Biochemische Charakterisierung von MprA und MprB	122
3.4.1	Klonierung von <i>mprA</i> und <i>mprB</i> sowie unterschiedlicher Genfragmente in den Expressionsvektor pET101/D-TOPO	122
3.4.2	Überprüfung der Funktionalität der konstruierten <i>mprA</i> und <i>mprB</i> -Gene und Gen-Derivate	125
3.4.3	Heterologe Expression der <i>mprA</i> und <i>mprB</i> -Gene und deren Derivate in <i>E. coli</i> BL21	127
3.4.4	Reinigung von MprA, MprA-K1, MprB und MprB-K1 aus konzentrierten Kulturüberständen über Metall-Chelat-Affinitätschromatographie	131
3.4.5	Reinigung von MprA, MprA-K1, MprB und MprB-K1 aus konzentrierten Kulturüberständen über hydrophobe Interaktionschromatographie (HIC)	132
3.4.6	Bestimmung der proteolytischen Aktivität in Abhängigkeit von der Temperatur	137
3.4.7	Bestimmung der Thermostabilität	139
3.4.8	Bestimmung der proteolytischen Aktivität in Abhängigkeit vom pH-Wert	141
3.4.9	Einfluss von Inhibitoren auf die proteolytische Aktivität	142
4	Diskussion	145
4.1	Konstruktion von Metagenombanken	145

4.2	Analyse des Metagenoms	152
4.3	Identifizierung von Genen aus Metagenombanken die Butanol-Dehydrogenase-Aktivität vermitteln	161
4.4	Identifizierung von Genen aus Metagenombanken die proteolytische Aktivität vermitteln	168
4.4.1	Genetische Organisation der Inserts von pTWP9 und pTWP13 und Vergleich mit korrespondierenden Genomabschnitten von verschiedenen <i>Xanthomonas</i> -Arten	169
4.4.2	Charakterisierung der Proteasen MprA und MprB	178
5	Zusammenfassung	185
6	Literaturverzeichnis	188
7	Anhang	204
7.1	Angaben zur Insertsequenz von pTW9	204
7.1.1	Lage der offenen Leserahmen	204
7.1.2	Insertsequenz von pTWP9	205
7.2	Angaben zur Insertsequenz von pTWP13	214
7.2.1	Lage der offenen Leserahmen	214
7.2.2	Insertsequenz von pTWP13	214
7.3	Angaben zu den Genomabschnitten der 5 <i>Xanthomonas</i>-Vertreter	221
7.3.1	<i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	221
7.3.2	<i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	222
7.3.3	<i>X. campestris</i> pv. <i>campestris</i> str. 8004	223
7.3.4	<i>X. axanopodis</i> pv. <i>citri</i> str. 306	225
7.3.5	<i>X. oryzae</i> pv. <i>oryzae</i> KACC10331	226

ABKÜRZUNGSVERZEICHNIS

α	alpha
Abb.	Abbildung
AG	Aktiengesellschaft
Ap ^r	Ampicillin-Resistenz
APS	Ammoniumpersulfat
ATP	Adenosin-5'-triphosphat
β	beta
bp	Basenpaare
bzw.	beziehungsweise
°C	Grad Celsius
C.	Clostridium
ca.	circa
cm	Zentimeter
C-Quelle	Kohlenstoffquelle
CTAB	Cetyl-Trimethyl-Ammoniumbromid
C-terminal	Carboxyterminal
Da	Dalton
dATP	Desoxyadenosin-5'-triphosphat
dCTP	Desoxycytidin-5'-triphosphat
dest.	destilliert
dGTP	Desoxyguanosin-5'-triphosphat
d.h.	das heisst
DH	Dehydrogenase
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleotid-5'-triphosphat
<i>E.</i>	<i>Escherichia</i>
ε	Extinktionskoeffizient
EDTA	Ethylendiamintetraessigsäure
ed(s).	Herausgeber
et al.	et alteri (und andere)
Fa.	Firma

g	Gramm
xg	-fache Erdbeschleunigung (9,81 m/s ²)
γ	gamma
GmbH	Gesellschaft mit beschränkter Haftung
h	Stunde
IPTG	Isopropyl-β-thiogalaktosid
k	Kilo
K	Kalium
Kbp bzw. kb	Kilo Basenpaare
KP	Kaliumphosphat
l	Liter
LB	Luria Bertani
M	molar (Mol pro Liter); Mega-
m	Meter
Mbp	Mega Basenpaar
Mg	Magnesium
μ	Mikro-
min	Minute
N ₂	Stickstoff
n	Nano-
NAD ⁺	Nikotinamidadenindinukleotid
NADH	Nikotinamidadenindinukleotid, reduziert
Nr.	Nummer
NTA	Nitrilotriessigsäure
N-terminal	Aminoterminal
Ω	Ohm
OD	optische Dichte
<i>orf</i>	offener Leserahmen
PAGE	Polyacrylamid-Gelelektrophorese
PCR	„polymerase chain reaction“
PEG	Polyethylenglykol
pH	negativer dekadischer Logarithmus der Protonenkonzentration

PMSF	Phenylmethylsufonylfleurid
R _f	elektrophoretische Mobilität
RNA	Ribonukleinsäure
rRNA	ribosomale RNA
RT	Raumtemperatur
S	Svedberg-Einheit
s	Sekunde
s.	siehe
SDS	Natriumdodecylsulfat
s.o.	siehe oben
<i>sp.</i>	<i>species</i>
T	Temperatur
Tab.	Tabelle
<i>Taq</i>	<i>Thermus aquaticus</i>
TEMED	N, N, N', N'-Trimethylethylendiamid
T _m	Schmelztemperatur
Tricine	N-[Tris(hydroxymethyl)-methyl]-glycin
Tris	Tris(hydroxymethyl)-aminomethan
U	Unit (Einheit der Enzymaktivität)
Upm	Umdrehungen pro Minute
V	Volt
Vol.	Volumen
v/v	Volumen pro Volumen
W	Watt
w/v	Masse pro Volumen
<i>X.</i>	<i>Xanthomonas</i>
x g	-fache Erdbeschleunigung (9,81 m/s ²)
X-Gal	5-Brom-4-Chlor-3-Indolyl-D-Galaktosid
z.B.	zum Beispiel

Abkürzungen für Aminosäuren

A	Ala	Alanin	M	Met	Methionin
C	Cys	Cystein	N	Asn	Asparagin
D	Asp	Aspartat	P	Pro	Prolin
E	Glu	Glutamat	Q	Gln	Glutamin
F	Phe	Phenylalanin	R	Arg	Arginin
G	Gly	Glycin	S	Ser	Serin
H	His	Histidin	T	Thr	Threonin
I	Ile	Isoleucin	V	Val	Valin
K	Lys	Lysin	W	Trp	Tryptophan
L	Leu	Leucin	Y	Tyr	Tyrosin

1 Einleitung

1.1 Erschließung und Analyse der mikrobiellen Diversität durch Metagenomik

Der größte Anteil an der Biomasse der Erde wird von Mikroorganismen gestellt. Nach Schätzungen geht man davon aus, dass ca. 1 bis 100 Millionen unterschiedliche prokaryotischer Arten auf der Erde vorkommen (WILSON, 1992; STRICKEBERGER, 1996). Aufgrund ihrer vielfältigen physiologischen, genetischen und metabolischen Eigenschaften sind Mikroorganismen in der Lage unter extremsten Umweltbedingungen wie z.B. in heißen Quellen, arktischen Standorten, Tiefsee oder Salinen zu leben. Sie stellen somit ein enormes Potenzial an natürlich vorkommender Biodiversität dar, die als Quelle für neuartige Biokatalysatoren und Wirkstoffen genutzt werden kann (SEOW et al., 1997; DANIEL et al., 2002).

Jedoch kann bis heute nur ein geringer Anteil der auf der Erde vorkommenden Prokaryoten (weniger als 1 %) mit Standardtechniken kultiviert werden (AMMAN et al., 1995). Daher ist der größte Anteil der vorhandenen mikrobiellen Diversität bisher unentdeckt geblieben. Mit Hilfe von molekularbiologischen Untersuchungen, z.B. durch 16S rRNA-Genanalysen, ist es möglich, die vorkommende, phylogenetische Diversität zu erfassen und einen Einblick in die Vielfalt mikrobiellen Lebens zu bekommen. Dies gelang Pace und seinen Mitarbeitern im Jahr 1985 als ersten durch die Entwicklung von kultivierungsunabhängigen Techniken, bei denen die gesamte mikrobielle DNA aus einer Umweltprobe isoliert und direkt kloniert wurde. Durch Sequenz-Analysen der 5S und 16S rRNA-Gene war es ihnen möglich, die Vielfalt der Mikroorganismen aus den heißen Quellen des Yellow Stone Parks ohne vorherige Kultivierung zu beschreiben (PACE et al., 1985). Mit Hilfe von molekularen, phylogenetischen Analysen ist es gelungen, den ersten bakteriellen Stammbaum aus dem Jahr 1987 mit insgesamt 12 Phyla (WOESE, 1987) bis zum Jahr 2003 auf 52 Phyla zu erweitern (RAPPE und

GIOVANNONI, 2003). Etwa die Hälfte der Phyla enthält bislang nicht kultivierbare Organismen (RIESENFELD et al., 2004).

In der Vergangenheit war die Kultivierbarkeit von Mikroorganismen eine Grundvoraussetzung zum Auffinden neuer industriell relevanter Biomoleküle (DANIEL, 2002). Durch die Metagenomforschung wurde es möglich, diese weitgehend ungenutzte Quelle mikrobieller Diversität sowohl für die Anwendung als auch für die Forschung unabhängig von der Kultivierbarkeit der Mikroorganismen zu erschließen. Der Begriff Metagenom wurde im Jahr 1998 von Handelsman und Mitarbeitern eingeführt, und bezeichnet die Gesamtheit aller in einem Habitat vorhandenen, mikrobiellen Genome (HANDELSMAN et al., 1998; RONDON et al., 2000). Zur Nutzbarmachung des Metagenoms wurden diverse molekulare Methoden entwickelt, die auf direkter Isolierung und Klonierung der mikrobiellen DNA eines Standortes ohne vorherige Anreicherung und Isolierung von Mikroorganismen basieren (HENNE et al., 1999; RONDON et al., 2000). Die dabei in heterologen Wirten hergestellten Genbibliotheken (Metagenombanken) aus unterschiedlichen Umwelt-Standorten können als genetische Ressource beliebig vermehrt, analysiert und wirtschaftlich genutzt werden (Übersicht: DANIEL, 2002; HANDELSMAN, 2004).

Die Analyse von Metagenomen kann durch sequenzbasierende oder aktivitätsbasierende Verfahren erfolgen (DANIEL, 2002; STREIT et al, 2004; DANIEL, 2004; HANDELSMAN, 2004). Eine Möglichkeit ist eine im großen Maßstab durchgeführte Sequenzierung von mikrobieller Umwelt-DNA oder Metagenombanken, die auf der Grundlage des „whole genome shotgun sequencing“ basiert. Ein prominentes Beispiel dafür ist die von der Arbeitsgruppe um Craig Venter durchgeführte Sequenzierung des Metagenoms der Sargasso Sea, die sich vor der Küste Floridas befindet (VENTER et al., 2004). In dem Projekt wurden mehrere hundert Liter Meerwasser aus der nährstoffarmen Sargasso Sea vor den Bermuda Inseln gefiltert. Mit der von den Filtern isolierten DNA wurden Genbanken angelegt, mit denen „high-throughput“-Sequenzierungen durchgeführt wurden. Auf diese Weise konnten der öffentlichen NCBI-Datenbank mehr als 1,6 Gbp an DNA-Sequenzen, in

mehr als 1 Million individuellen Einträgen sowie in etwa die gleiche Anzahl an potentiell translatierten Proteinen zur Verfügung gestellt werden. Diese Datenflut führte dazu, dass die Größe der öffentlichen Datenbank auf einen Schlag verdoppelt wurde. Von den identifizierten potentiellen Genen zeigten etwa 70.000 keinerlei Ähnlichkeiten zu bekannten Genen. Besonders überraschte dabei der Fund von mehr als 800 bislang unbekannten Genen, die für Photorezeptoren kodieren (VENTER et al., 2004). Weitere Beispiele für Projekte dieser Art sind die metagenomische Analyse eines Trinkwasserbiofilms (SCHMEISSER et al., 2003) und eines „acid mine drainage“-Biofilms (TYSON et al., 2004). Ähnliche in Planung befindliche Projekte haben zum Beispiel zum Ziel, die Zusammensetzung mikrobieller Organismen in urbaner Luft (VENTER et al.) oder die Zusammensetzung der oralen mikrobiellen Lebensgemeinschaft zu analysieren (National Institute of Dental and Craniofacial Research, Bethesda, USA). Die „high-throughput“-Sequenzanalyse von Metagenomen aus verschiedenen Standorten ermöglicht auch die Erstellung von genetischen „Fingerprints“ von Habitaten und die komparative Metagenomik (TRINGE et al., 2005).

Die funktionelle Analyse von Metagenombanken bietet eine nahezu unerschöpfliche Quelle für biotechnologisch-relevante Biomoleküle (DANIEL, 2004). Die industrielle bzw. „weiße“ Biotechnologie hat sich ökologisch und ökonomisch als echte Alternative zu klassischen synthetischen Verfahren durchgesetzt. Es wird prognostiziert, dass bis zum Jahr 2010 zwischen 10 % bis 20 % aller chemischen Produkte durch biotechnologische Verfahren produziert werden (LORENZ und ECK, 2005). Um den industriellen Bedarf an neuen Biomolekülen decken zu können, ist es notwendig die mikrobielle Diversität für die Anwendung nutzbar zu machen. Mit verschiedenen Screeningverfahren können Metagenombanken z.B. auf eine gewünschte Enzymaktivität durchmustert werden (SCHLOSS und HANDELSMANN, 2003). Diese Techniken konnten in der Vergangenheit erfolgreich eingesetzt werden, um industriell-relevante neuartige Enzyme sowie pharmazeutisch wichtige Naturstoffe zu entdecken (LORENZ und ECK, 2005). Beispiele hierfür sind die Isolierung von neuartigen Amylasen (RONDON et al., 2000; YUN et al., 2004), Cellulasen (HEALY et al., 1995; VOGET et al., 2005), Oxidoreduktasen

(KNIETSCH et al., 2003a und 2003b), Glycerin/Diol-Dehydratasen (KNIETSCH et al., 2003c), Amidasen (GABOR et al., 2004) Lipasen/Esterasen (HENNE et al., 2000; RONDON et al., 2000; LEE et al., 2004) und Antibiotika (BRADY und CLARDY, 2000; WANG et al., 2000; GILLESPIE et al., 2002) aus Metagenombanken. Im Rahmen der Arbeit soll die funktionelle Analyse von Metagenombanken angewendet werden, um neuartige Gene zu identifizieren, die eine Butanol-Dehydrogease-Aktivität bzw. proteolytische Aktivität, vermitteln.

1.2 Butanol, industrielle Bedeutung und Produktion

Der Alkohol n-Butanol ist eine klare, mittelflüchtige und neutrale Flüssigkeit mit einem charakteristischen Geruch. Das Produkt mit der Summenformel $C_4H_{10}O$ findet z.B. als Lösungsmittel oder als Ausgangsmaterial für Synthesen vielseitige Verwendung in der Industrie (<http://enius.de/schadstoffe/1-butanol.html>). Im Nachfolgenden sind einige Anwendungsbeispiele von n-Butanol aufgeführt.

In der Lackindustrie findet Butanol Verwendung als Lösungsmittel bei Lackzubereitungen. Besonders vorteilhaft ist dabei, dass Butanol das Weißanlaufen beim Trocknen bestimmter Lacke in feuchter Atmosphäre verhindert. Er senkt in Alkydharzlacken bereits nach Zugabe geringer Mengen die Viskosität beträchtlich und verbessert damit Streicheigenschaften und Verlauf dieser Lacke. Es wird in geringen Konzentrationen auch Lacken auf der Basis alkohollöslicher Harze zugesetzt, um das Fadenziehen zu verhindern (Technisches Merkblatt für n-Butanol; Fa. BASF, Ludwigshafen). Ferner kann es verwendet werden als:

- Lösungsmittel für Farbstoffe (z.B. Druckfarben)
- Extraktionsmittel zur Gewinnung von Arzneimittelwirkstoffen und Naturstoffen (z.B. Antibiotika, Hormone, Vitamine, Alkaloide und Kampfer)

- Zusatz in Reinigungsmitteln (z.B. Fußbodenreiniger oder Fleckentferner)
- Zusatz in Kraftstoffen für Ottomotoren (verhindert eine Vereisung des Vergasers)
- Laufmittel bei der Papier- und Dünnschichtchromatographie
- Ausgangsmaterial zur Herstellung von z.B. Monocarbonsäurebutylestern und Flotationschemikalien (Technisches Merkblatt für n-Butanol; Fa. BASF, Ludwigshafen; DÜRRE, 2005).

Der weitaus größte Anteil an n-Butanol wird heutzutage chemisch in zwei Schritten hergestellt. Im ersten Schritt („Oxo-Synthese“) wird Propen mit Synthesegas zu Butyraldehyd umgesetzt, aus welchem dann im zweiten Schritt durch eine Hydrierungsreaktion Butanol entsteht (DÜRRE, 2005):

Schritt 1: „Oxo-Synthese“

Schritt 2: „Hydrierung“

Eine interessantere Alternative der Butanol-Herstellung ist die biotechnologische Aceton-Butanol-Fermentation durch einige Vertreter der Gattung *Clostridium* wie z.B. *C. acetobutylicum*, *C. beijerinickii*, *C. saccharobutylicum* und *C. saccharoperbutylacetonicum*. Die Aceton-Butanol-Fermentation durch *C. acetobutylicum* ist in Abbildung 1 dargestellt. *C. acetobutylicum* zeichnet sich dabei durch einen zweiphasigen Gärungsstoffwechsel aus (JONES und WOODS., 1986). In der logarithmischen Wachstumsphase (Säurephase) werden hauptsächlich die organischen Säuren Butyrat und Acetat gebildet. Dabei kommt es zu einem starken Abfall des pH-Werts, wodurch der Gärungsstoffwechsel auf die Bildung neutraler

Gärungsprodukte umgestellt wird. In der zweiten Phase (Lösungsmittelphase) werden nun Gärungsprodukte wie Ethanol, Isopropanol, Aceton und Butanol produziert und ausgeschieden. Wie aus Abbildung 1 zu entnehmen ist, sind an der direkten Bildung von Butanol aus Butyraldehyd drei Enzyme beteiligt. Dabei handelt es sich um die bifunktionale Butyraldehyd-/Butanol-Dehydrogenase E (AdhE) sowie um die beiden Butanol-Dehydrogenasen A und B (BdhA und BdhB).

Abbildung: 1 Schema, der Aceton-Butanol-Fermentation durch *Clostridium acetobutylicum*.

Jeder Pfeil kennzeichnet eine Enzymreaktion. Es sind nur die Enzyme, die direkt an der Butanol-Synthese beteiligt sind, aufgefhrt. AdhE: Butyraldehyd-/Butanol-Dehydrogenase E; BdhA: Butanol-Dehydrogenase A; BdhB: Butanol-Dehydrogenase B. Eingerahmt sind die mglichen Grungsprodukte.

Die biologische Produktion der chemisch wichtigen Lsungsmittel Aceton und Butanol hatte in der ersten Hlfte des 20-sten Jahrhunderts (bis etwa 1950) eine enorme Bedeutung und wurde nur durch die Ethanol-Fermentation der

Hefe übertroffen (JONES und WOODS, 1986; DÜRRE, 1998). Ab etwa 1950 verdrängten kostengünstigere Verfahren der petrochemischen Industrie die Butanol-Produktion via Fermentation. Auch erzielte das Fermentationsverfahren schlechtere Ausbeuten als die chemische Produktion. Spätestens die 1973 aufkommende Ölkrise führte zu einem erneuten, starken Interesse an der Aceton-Butanol-Fermentation (DÜRRE, 1998). Intensive Forschungen der letzten Jahre führten zum Auffinden fast aller am Fermentationsprozess von *Clostridium acetobutylicum* beteiligten Enzyme. Die meisten Gene wurden kloniert, sequenziert und die Enzyme charakterisiert (WELCH et al., 1989; YOUNGLESON et al., 1989; PETERSEN et al., 1991; WALTER et al., 1992; FISCHER et al., 1993; NAIR et al., 1994; SAUER und DÜRRE, 1995; FONTAINE et al., 2002). Auch die Genomsequenzierung von *C. acetobutylicum* konnte abgeschlossen werden (NÖLLING et al., 2001).

In Anbetracht der weltweit immer knapper werdenden Rohstoffe ist es von großem industriellem Interesse, die chemische Produktion von Butanol durch ein biotechnologisches Verfahren zu ersetzen. Das kann jedoch nur erreicht werden, wenn die Butanol-Ausbeuten im Fermentationsprozess durch gezielte Optimierung (z.B. „Enzyme Engineering“) gesteigert werden. Eine weitere Möglichkeit der Optimierung besteht in dem Auffinden von neuen Genen z.B. aus Metagenombanken die hohe Butanol-Dehydrogenase-Aktivität vermitteln. Mit Hilfe dieser neuen Gene könnte ein Produktionsstamm zur biotechnologischen Produktion von Butanol entwickelt werden.

1.3 Klassifizierung, Funktionen und industrielle Anwendung von Proteasen

Proteasen (Peptidasen) umfassen eine breite Enzymklasse, die die hydrolytische Spaltung von Peptidbindungen in Proteinen oder Peptiden katalysieren. Sie sind ubiquitär in allen Organismen verbreitet und kommen dabei in allen Geweben und Zellen vor. Entsprechend der Nomenklatur der

„International Union of Biochemistry and Molecular Biology“ gehören Proteasen der 3. Hauptklasse der Enzyme, den Hydrolasen, an und bilden dort die Unterklasse 3.4.

Zunächst erfolgt eine Unterteilung der Peptidasen in Endopeptidasen bzw. Exopeptidasen. Endopeptidasen spalten immer innerhalb einer Polypeptidkette, während Exopeptidasen ausschließlich am N-terminalen Ende der Polypeptidkette (Aminopeptidasen) bzw. C-terminalen Ende (Carboxypeptidasen) agieren. Aufgrund der Beschaffenheit ihrer katalytischen, aktiven Zentren können 4 unterschiedliche Typen (Serin-, Cystein-, Aspartat- und Metalloproteasen) klassifiziert werden. Neben diesen vier Hauptklassen existieren aber auch Proteasen, die einen unbekannten Reaktionsmechanismus aufweisen und deswegen in eine eigenständige Klasse eingestuft werden. Basierend auf Sequenzähnlichkeit und evolutionärer Verwandtschaft erfolgen weitere Untergliederungen (RAO et al., 1998).

Ein spezielleres Klassifikationssystem für proteolytische Enzyme, genannt MEROPS, wurde von Alan Barrett und seinen Mitarbeitern im Jahr 1993 entwickelt (RAWLINGS et al., 1993). In der MEROPS-Datenbank (www.merops.ac.uk), die seit 1996 existiert, werden Proteasen aufgrund von strukturellen Aspekten sowie evolutionären Verwandtschaftsbeziehungen auf der Basis der Aminosäuresequenz klassifiziert. Jede Protease ist in eine bestimmte Familie eingeordnet. Eine Familie wiederum beinhaltet Peptidasen, die signifikante Ähnlichkeiten in ihrer Aminosäuresequenz besitzen. Mehrere Familien werden aufgrund ihrer gemeinsamen evolutionären Verwandtschaftsbeziehungen, die auf Übereinstimmungen in der Tertiärstruktur basieren, in einem Clan zusammengefasst. Die Eingruppierung oberster Hierarchie erfolgt aufgrund der unterschiedlichen Katalysemechanismen, wobei jedem katalytischen Typ mehrere Clans untergeordnet werden (RAWLINGS et al., 1993 und 2006).

Proteasen sind für alle Organismen lebensnotwendig. Sie nehmen oft Schlüsselstellungen in einem breiten Spektrum von biologischen Prozessen ein und erfüllen eine große Zahl essentieller Funktionen in Lebewesen. Dabei

differenziert man zwischen intrazellulären Proteasen und extrazellulären Proteasen, die in das Medium sekretiert werden. Weiterhin muss zwischen limitierter Proteolyse und totaler Proteolyse unterschieden werden. Bei der totalen Proteolyse werden in der Regel nicht mehr benötigte Proteine vollständig abgebaut. Bei der limitierten Proteolyse wird dabei durch gezielte Abspaltung einer geringen Anzahl von Peptidbindungen die Aktivierung oder Inaktivierung von anderen Proteinen hervorgerufen. Nachfolgend sind einige Funktionen von Proteasen aufgeführt (Übersicht: RAO et al., 1998):

- Beteiligung an diversen biologischen Prozessen, die über Spaltung von Peptidbindungen verlaufen (z.B. Blutgerinnung, Fibrinolyse, Kontrolle des Blutdrucks, Zellteilung, Gewebeabbau und Aufbau, Sporulation, Tumorzellinvasion und Entzündungsreaktionen)
- Beteiligung an der Reifung neu synthetisierter Proteine, indem sie z.B. Signalsequenzen oder Zielsteuerungssequenzen abspalten
- Verdauung von Nahrungsproteinen
- Prozessierung von Vorläuferproteinen z.B. von Hormonen oder Organellproteinen
- Regulation des Proteingehalts innerhalb der Zelle

Aufgrund der vielfältigen Wirkungsweise finden Proteasen vielseitige Anwendungen in der Industrie. Sie bilden mit 60 % den weitaus größten Anteil an industriell genutzten Enzymen (RAO et al., 1998). Die größte industrielle Bedeutung haben dabei bakterielle Proteasen, die als Standard-Zusatz in Wasch- und Reinigungsmitteln Verwendung finden. Aber auch tierische Proteasen wie z.B. die Labenzyme Chymosin und Renin aus dem Kälbermagen werden industriell zur Käseherstellung eingesetzt. In der Lederindustrie spielen Proteasen eine entscheidende Rolle bei der Enthaarung, Beize und Weichmachung von Leder. Weitere industrielle Verwendungen von Proteasen bestehen im Brauereiwesen und der Backwarenindustrie (RUTTLOFF, 1994).

1.4 Zielsetzung der Arbeit

In dieser Arbeit sollten zunächst Metagenombanken aus verschiedenen Standorten konstruiert werden. Im nächsten Schritt sollte die in den Genbanken niedergelegte mikrobielle Diversität genutzt werden, um in aktivitätsbasierenden Screeningverfahren potentielle Gene zu identifizieren, deren Genprodukte Butanol-Dehydrogenase-Aktivität bzw. proteolytische Aktivität vermitteln. Weiterhin sollte eine molekulare und biochemische Charakterisierung von den identifizierten Genen bzw. Genprodukten erfolgen.

2 Material und Methoden

2.1 Verwendete Organismen, Plasmide und Oligonukleotide

Die im Rahmen dieser Arbeit verwendeten Organismen sind in Tabelle 1, die Vektoren und die konstruierten rekombinanten Plasmide in Tabellen 2 und 3 aufgeführt. Tabelle 4 zeigt die eingesetzten Oligonukleotide:

Tabelle 1: Verwendete Organismen

Stamm	relevanter Geno- oder Phänotyp	Herkunft/Referenz
<i>Escherichia coli</i> TOP10	$F^- mcrA \Delta(mrr-hsdRMS-mcrBC)$ $\Phi 80 lacZ \Delta M15 \Delta lacX74 recA1 deoR$ $araD139 \Delta(ara-leu)7697 galU galK$ $rpsL (Str^R) endA1 nupG$	Invitrogen, Carlsbad, USA
<i>Escherichia coli</i> DH5 α	$F^- endA1 \Phi 80 lacZ \Delta M15 \Delta(lacZYA-$ $argF)U169, endA1, recA1, hsdR17$ $(r_K-m_K+), deoR, thi-1, supE44, \lambda-,$ $gryA96, relA1$	HANAHAN, 1983
<i>Escherichia coli</i> BL21(DE3)	$F^- ompT hsdS_B (r_B-m_B-), gal dcm$ $rne131 (DE3)$	Invitrogen, Carlsbad, USA
<i>Escherichia coli</i> WL3	$adhC81 fadR adhE3$	LOROWITZ et al., 1982

Tabelle 2: Verwendete Vektoren

Plasmid	relevantes Merkmal	Herkunft/Referenz
pBluescript [®] SK+	$Ap^r, lacPOZ^-, pMB1$ -Replikon	Stratagene, San Diego, USA
pBluescript [®] IISK+	$Ap^r, lacPOZ^-, pMB1$ -Replikon	Stratagene, San Diego, USA
pCR2.1-TOPO	$Ap^r, Kan^r, pMB1$ -Replikon	INVITROGEN, CARLSBAD, USA
pET101/D-TOPO	$Ap^r His_6$ -tag C-terminal, V5-epitop-tag	Invitrogen, Carlsbad, USA

Tabelle 3: Übersicht über die im Rahmen der Arbeit konstruierten rekombinanten Plasmide

Plasmid	relevantes Merkmal	Herkunft/Referenz
pTWB3	pCR2.1-TOPO mit Insert aus Umwelt-DNA	diese Arbeit
pTWB4	pCR2.1-TOPO mit Insert aus Umwelt-DNA	diese Arbeit
pTWP9	pBluescript®SK mit Insert aus Umwelt-DNA	diese Arbeit
pTWP13	pBluescript®SK mit Insert aus Umwelt-DNA	diese Arbeit
pET-MprA	pET101/D-TOPO mit Insert <i>orf29</i> aus Insert pTWP9	diese Arbeit
pET-MprA-K1	pET101/D-TOPO mit Insert <i>orf29</i> (C-terminal verkürzt) aus Insert pTWP9	diese Arbeit
pET-MprB	pET101/D-TOPO mit Insert <i>orf8</i> aus Insert pTWP13	diese Arbeit
pET-MprB-K1	pET101/D-TOPO mit Insert <i>orf8</i> (C-terminal verkürzt) aus Insert pTWP13	diese Arbeit
pET-MprB-K2	pET101/D-TOPO mit Insert <i>orf8</i> (N-terminal verkürzt) aus Insert pTWP13	diese Arbeit
pET-MprB-K3	pET101/D-TOPO mit Insert <i>orf8</i> (C- und N-terminal verkürzt) aus Insert pTWP13	diese Arbeit

Tabelle 4: Verwendete Oligonukleotide

Bezeichnung	Sequenz
rev	5'-TTCACACAGGAAACAGCT-3'
for	5'-ACGACGTTGTAAAACGACGGCCAG-3'
pET101for	5'-CGATCCCGCGAAATTAATACG-3'
pET101rev	5'-GGGTTATGCTAGTTATTGCTC-3'
pET-MprA/K1-for	5'- <u>CACCAT</u> GCATTCCAGCAGTCG-3'
pET-MprA-K1-rev	5'-GAAGGTGACGCTCCAGCTG-3'
pET-MprB/K1-for	5'- <u>CACCAT</u> GCAGTCCAGCAGTCGT-3'
pET-MprA/MprB-rev	5'-GAAGGTGACGCTCCAGCTG-3'
pET-MprB-K2/K3-for	5'- <u>CACCAT</u> GGCCGATATCGGCACC-3'
pET-MprB-K1/K3-rev	5'-GCCGGTGCTGGCGGAGATG-3'

Die Erkennungssequenz für die gerichtete Klonierung in den Expressionsvektor pET101/D-TOPO ist unterstrichen.

2.2 Nährmedien

Alle aufgeführten Medien wurden durch 30-minütiges Autoklavieren bei 121°C sterilisiert. Die C- Quellen für Mineralmedien wurden separat autoklaviert und erst unmittelbar vor dem Animpfen zugegeben. Zur Herstellung von Agarplatten wurde dem Medium vor dem Autoklavieren 1,5 % (w/v) Agar zugefügt.

Anaerobe Medien wurden nach der von COSTILOW (1981) beschriebenen Methode unter Begasung mit sauerstofffreiem Stickstoff hergestellt

2.2.1 LB- (Luria-Bertani) Medium (SAMBROOK et al., 1989)

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

2.2.2 SOB-Medium (HANAHAHAN, 1983)

Trypton	20	g
Hefeextrakt	5	g
NaCl	0,58	g
KCl	0,18	g
MgCl ₂	2,0	g
MgSO ₄ x 7H ₂ O	2,5	g
H ₂ O _{dest.}	ad 1000	ml

2.2.3 Denitrifikanten-Anreicherungsmedium

KNO ₃	20	g
K ₂ HPO ₄	0,5	g
MgSO ₄ x 7H ₂ O	0,2	g
SL 10	1	ml
Vitaminlösung nach Wolin (s. 2.2.9)	1	ml
H ₂ O _{dest.}	ad 1000	ml

Dem Medium wurde als C-Quelle entweder Zimtsäure (2,5 mM) oder 2-Penten (2 mM) nach dem Autoklavieren zugesetzt.

2.2.4 Medium für Aldehyd-Indikator-Platten-Test (CONWAY et al., 1987, modifiziert)

KH ₂ PO ₄	7	g
K ₂ HPO ₄	3	g
MgSO ₄ x 7 H ₂ O	0,1	g
Proteose Pepton	2	g
Vitaminlösung nach Wolin (s. 2.2.9)	2	ml
SL-10	2	ml
Agar	15	g
H ₂ O _{dest.}	ad 1000	ml

Das Medium wurde autoklaviert und auf ca. 45 °C abgekühlt. Danach konnten die Indikatoren zugefügt werden:

Pararosanilin	20	mg
Na ₂ S ₂ O ₅	100	mg

Pararosanilin und Natriumbisulfit wurden vor der Zugabe im Dunkeln in 20 bis 30 ml H₂O gelöst und der pH-Wert auf 7,0 eingestellt. Als C-Quelle wurde dem Medium 2 ml Butanol zugesetzt.

2.2.5 M9-Mineralmedium (MILLER, 1972; modifiziert)

KH ₂ PO ₄	3,0	g
Na ₂ HPO ₄	6,0	g
NH ₄ Cl	3,0	g
NaCl	0,5	g
MgSO ₄ x 7 H ₂ O	0,05	g
CaCl ₂ x 12 H ₂ O	0,002	g
FeSO ₄	0,02	g
Hefeextrakt	0,2	g
SL-4	1	ml
Vitaminlösung nach Wolin (s. 2.2.9)	1	ml
H ₂ O _{dest.}	ad 1000	ml
pH-Wert einstellen auf	7,4	

Dem Medium wurde als C-Quelle Glucose in einer Endkonzentration von 20 mM zugesetzt.

2.2.6 Protease-Expressionsmedium

Für die Anzucht von rekombinanten *E. coli*-Stämmen diente als Grundlage modifiziertes M9-Mineralmedium (s. 2.2.5), welches zur Förderung der

Protease-Sekretion mit L-Arginin (0,3 M) versetzt wurde. Als C-Quelle wurde Glucose mit einer Endkonzentration von 20 mM eingesetzt.

2.2.7 Spurenelementelösung SL-4 (PFENNIG und LIPPERT, 1966)

Na ₂ -EDTA	5,0	g
FeSO ₄ x 7 H ₂ O	2,0	g
ZnSO ₄ x 7 H ₂ O	0,1	g
MnCl ₂ x 4 H ₂ O	0,03	g
H ₃ BO ₄	0,3	g
CoCl ₂ x 6 H ₂ O	0,2	g
CuCl ₂ x 2 H ₂ O	0,01	g
NiCl ₂ x 6 H ₂ O	0,02	g
NaMoO ₄ x 2 H ₂ O	0,03	g
H ₂ O _{dest.}	ad 1000	ml

Der pH-Wert wurde auf 6,7 eingestellt und die Lösung lichtgeschützt bei 4 °C gelagert.

2.2.8 Spurenelementelösung SL-10 (WIDDEL et al., 1983)

HCl (25 %)	10	ml
FeCl ₂ x 4 H ₂ O	1,5	g
ZnCl ₂	70	mg
MnCl ₂ x 4 H ₂ O	100	mg
H ₃ BO ₄	6	mg
CoCl ₂ x 6 H ₂ O	190	mg
CuCl ₂ x 2 H ₂ O	2	mg
NiCl ₂ x 6 H ₂ O	24	mg
Na ₂ MoO ₄ x 2 H ₂ O	36	mg
H ₂ O _{dest.}	ad 1000	ml

2.2.9 Vitaminlösung (WOLIN et al., 1964, modifiziert)

Biotin	2	mg
Folsäure	2	mg
Pyridoxin x HCl	10	mg
Thiamin x HCl	15	mg
Riboflavin	5	mg
Nikotinsäure	5	mg
Ca-Pantothenat	5	mg
Vitamin B ₁₂	10	µg
p-Aminobenzoessäure	5	mg
α-Liponsäure	1	mg
H ₂ O _{dest.}	ad 1000	ml

2.2.10 Antibiotika und sonstige Medienzusätze

Für die Anzucht rekombinanter *E. coli*-Stämme wurden die in Tabelle 5 aufgeführten Medienzusätze verwendet:

Tabelle 5: Antibiotika und sonstige Medienzusätze

Medienzusatz	Stammlösung	Arbeitskonzentration
Ampicillin	50 mg/ml in 50 % Ethanol	50-100 µg/ml
Kanamycin	50 mg/ml in H ₂ O _{des}	50-100 µg/ml
IPTG	25 mg/ml in H ₂ O _{dest.}	50 µg/ml
X-Gal	20 mg/ml in Dimethylformamid	40 µg/ml

Die IPTG-Stammlösung, die Kanamycin-Stammlösung und die Ampicillin-Stammlösung wurden sterilfiltriert. Die Zugabe der Komponenten erfolgte erst nach dem Abkühlen der autoklavierten Medien auf mindestens 50 °C. Die Lagerung erfolgte bei -20 °C.

2.3 Zellanzucht

2.3.1 Aerobe Anzucht

Die aerobe Anzucht von *E. coli* erfolgte in Reagenzgläsern oder Erlenmeyerkolben mit Schikanen. Das Kulturvolumen betrug ca. 20 % des Füllvolumens der Reagenzgläser bzw. Erlenmeyerkolben. Die Kulturen wurden unter Schütteln bei 30 ° oder 37 °C inkubiert.

2.3.2 Anaerobe Anzucht

Zur Herstellung von anaeroben Anreicherungskulturen aus diversen Umweltproben wurden 500 ml anaerobisiertes Dentrifikantenmedium (s. 2.2.3) mit 5-10 g Umweltprobe als Inokulum versetzt. Die Inkubation erfolgte für eine Woche bei 30 °C in 1 l-Serumflaschen (Fa. Müller und Krempel AG, Bülach, Schweiz). Die Kulturen wurden jeden Tag auf Gasbildung hin überprüft.

2.3.3 Stammhaltung

Agarplatten mit *E. coli* konnten maximal bis zu 8 Wochen bei 4 °C gelagert werden. Zur längerfristigen Konservierung wurden Stammkulturen angelegt. Dazu wurden diese bei 37 °C in LB-Medium (s. 2.2.1) angezogen. In der logarithmischen Wachstumsphase wurden dann 0,75 ml Zellsuspension entnommen, mit 0,3 ml Glycerin (87 %, v/v) versetzt und bei -70 °C gelagert.

2.3.4 Anzucht von rekombinanten *E. coli* Stämmen zur heterologen Genexpression

2.3.4.1 Expression mit dem „pET Directional TOPO Expression“-Kit

Der „pET Directional TOPO® Expression“-Kit (Fa. Invitrogen, Carlsbad, USA) nutzt die Klonierungsstrategie des „TOPO-Cloning“ (s. Abb.3 und detaillierte Erklärung des Grundprinzips unter 2.8.3.2), um blunt-end-PCR-Produkte gerichtet in einen „high-level“-Expressionsvektor zu klonieren. Außerdem besteht die Möglichkeit das Zielgen N- oder C-terminal mit einem Tag (z.B. His₆-Tag) zu versehen, der später eine einfache Detektion mit Antikörpern oder die Reinigung des gebildeten Proteins ermöglicht. Die physikalische Karte des verwendeten Vektors pET101/D-TOPO (mit C-terminalem His₆-Tag) ist in Abbildung 2 dargestellt.

Abbildung 2: Physikalische Karte des Expressionsvektors pET101/D-TOPO

lacO: lac-Operator; *lacI*: lac-Repressor; RBS: Ribosomenbindestelle; EK: Enterokinase-Erkennungsstelle; *rop*: erleichtert „low-copy“-Replikation in *E. coli*;

Die gerichtete TOPO-Klonierung (s. Abb. 2) wurde durch das Anhängen einer spezifischen Sequenz (5'-CACC) an das durch PCR amplifizierte Zielgen erreicht, die komplementär (3'-GTGG) als Einzelstrang-Überhang auf dem Vektor vorhanden war. Dieser Überhang drang in das 5'-Ende des PCR-Produkts ein und stabilisierte so die korrekte Orientierung des inserierten Gens. Das pET-Expressionssystem macht sich die hohe Aktivität und Spezifität der T7-RNA-Polymerase zu Nutze. Dies ermöglichte ausgehend vom T7-Promotor eine regulierte Expression von heterologen Genen in *E. coli* (ROSENBERG et al., 1987; STUDIER und MOFFATT, 1986; STUDIER et al. 1990). Für die Expression wird der *E. coli*-Stamm BL21 Star (DE3) herangezogen. Dieser Stamm trägt den defekten, lysogenen Bakteriophagen DE3 des λ -Phagen, der folgende Elemente chromosomal kodiert enthält:

- den Lac-Repressor *lacI*
- die T7-RNA-Polymerase-Gene unter der Kontrolle des *lacUV5*-Promotors
- einen Teil des *lacZ*-Gens

Abbildung 3: Schematische Darstellung der gerichteten TOPO-Klonierung.

Wenn dieser Vektor in einem DE3-lysogenen *E. coli*-Stamm etabliert wurde, wirkte der gebildete *lac*-Repressor zum einen an dem *lacUV5*-Promotor im Wirtschromosom und reprimierte dort die Expression der T7-Polymerase und zum anderen am T7-Promotor des Vektors. Letzteres verhinderte die Expression der inserierten Gene. Die Expression der T7-RNA-Polymerase und der inserierten Gene ließ sich somit durch die Zugabe von IPTG induzieren.

2.4 Bestimmung von Wachstumsparametern

2.4.1 Bestimmung der optischen Dichte

Die Bestimmung der optischen Dichte (OD) erfolgte anhand von Trübungsmessungen bei einer Wellenlänge von 600 nm gegen sterile Nährlösung in einem Spektralphotometer (UV-VIS Spektrophotometer V-550, Fa. Jasco, Jena) in Plastikküvetten mit einer Schichtdicke von 1 cm. Bei Zelldichten $E_{600} > 0,3$ wurden die Proben mit steriler Nährlösung entsprechend verdünnt. Die Messungen erfolgten gegen Medium als Leerwert.

2.4.2 Bestimmung des pH-Wertes

Der pH-Wert wurde mit einem WTW pH-Meter pH 526 (Fa. WTW GmbH, Weilheim) mit einer Glas-Calomel-Elektrode (Fa. Cahn-Ingold, Frankfurt) bestimmt. Vor den Messungen wurde das Gerät für den entsprechenden Messbereich mit Eichlösungen (pH 4,0; pH 7,0 und pH 10,0 der Fa. Riedel de Haen AG, Hannover) geeicht.

2.5 Standardtechniken für das Arbeiten mit Nukleinsäuren

2.5.1 Behandlung von Geräten und Lösungen

Alle hitzestabilen Lösungen und Geräte wurden zur Inaktivierung von Nukleasen für 20 min bei 121 °C autoklaviert. Hitzelabile Geräte wurden mit 70 % (v/v) Ethanol gewaschen und Metalloberflächen abgeflammt. Hitzelabile Biochemikalien wurden sterilfiltriert.

2.5.2 Fällung, Reinigung und Konzentration von Nukleinsäuren

2.5.2.1 Fällung von Nukleinsäuren

Die Konzentration von Nukleinsäuren erfolgte durch Fällung mit Ethanol oder Isopropanol. Durch Zugabe von 3 M Na-Acetat (pH 5,5) wurde, falls notwendig, die Ionenstärke erhöht (Endkonzentration 0,3 M). Zur Präzipitation der Nukleinsäuren erfolgte die Zugabe von 2 Vol. Ethanol (96 %) bzw. 1 Vol. Isopropanol (96 %). Die Ethanolfällung erfolgte durch Inkubation bei -70 °C für mindestens 20 min, bei Isopropanol wurde mindestens 5 min bei Raumtemperatur inkubiert. Die Nukleinsäuren wurden durch Zentrifugation (13000 bis 16000 x g, 4 °C oder RT, 30 min) pelletiert. Das Pellet wurde mit 1 ml eiskaltem Ethanol (70 %, v/v) gewaschen, getrocknet und in dem gewünschten Volumen sterilem H₂O_{dest.} aufgenommen.

2.5.2.2 DNA-Fällung nach Saccharose-Dichtegradienten-Zentrifugation

Ziel dieser Fällung war die Abtrennung von Saccharose nach der Saccharose-Dichtegradienten-Zentrifugation (s. 2.9). Dazu wurde jedes Aliquot aus dem Gradienten mit 0,5 Vol. PEG-NaCl-Lösung (20 % PEG 6000; 2,5 M NaCl) versetzt und über Nacht bei RT inkubiert. Anschließend erfolgte eine Zentrifugation der Lösungen (30 min, 13.000 rpm, RT). Der Überstand wurde

abgenommen, das Pellet mit 0,5 ml Ethanol (70 % (v/v), RT) versetzt und nochmals unter den gleichen Bedingungen zentrifugiert. Wiederum wurde der Überstand abgenommen, das Pellet getrocknet und in 20 bis 30 µl sterilem H₂O_{dest.} aufgenommen. Alle Schritte wurden bei RT durchgeführt, da bei Einsatz niedrigerer Temperaturen das Polyethylenglykol ausfällt.

2.5.2.3 Mikrodialyse von DNA

Um die Effizienz der Elektroporation (s. 2.10.2) zu erhöhen, wurden die Salze aus Ligationsansätzen durch eine Mikrodialyse entfernt. Die Dialyse wurde in einer kleinen Petrischale gegen steriles H₂O als Dialysepuffer durchgeführt. Zunächst wurden kleine Membranfilterscheiben (Porengröße 0,025 µm, Fa. Milipore, Eschborn) mit der glänzenden Seite nach oben auf die Wasseroberfläche gegeben. Dann wurden mit einer Pipette vorsichtig 5 bis 10 µl der Probe auf die Dialysemembran aufgetragen. Nach 30 min konnten die dialysierten Proben abgenommen werden.

2.5.2.4 Dialyse von Umwelt-DNA

Enthielt die aus den Standort-Proben isolierte DNA zu hohe Salzkonzentrationen, musste sie vor einer weiteren Reinigung zunächst einer Dialyse unterzogen werden. Zu diesem Zweck wurden 20 bis 30 cm lange Dialyseschlauchstücke in TE-Puffer autoklaviert. Anschließend wurde die DNA-Lösung in die Schläuche eingefüllt, die an beiden Enden mit Dialyseschlauch-Klammern verschlossen wurden. Die DNA-Lösung wurde dann 24 h gegen TE-Puffer (1x) dialysiert. Nach etwa 4 bis 5 h wurde der Puffer ausgewechselt. Anschließend erfolgte eine Fällung der DNA mit Ethanol (s. 2.5.2.1).

TE-Puffer (10x):

Tris	100	mM
EDTA	10	mM
H ₂ O _{dest.}	ad 1000	ml
pH 8,0		

2.5.2.5 Konzentrationsbestimmung von Nukleinsäuren

Die Bestimmung der Konzentration von Nukleinsäuren in wässriger Lösung erfolgte photometrisch bei einer Wellenlänge von 260 nm. Eine Extinktion von 1 entspricht etwa:

- 50 µg/ml doppelsträngiger DNA,
- 40 µg/ml einzelsträngiger DNA oder
- 30 µg/ml einzelsträngiger Oligonukleotide.

Die Reinheit von Nukleinsäure-Lösungen konnte durch das Verhältnis OD₂₆₀ zu OD₂₈₀ bestimmt werden. Bei reinen DNA-Lösungen sollte dieser Wert zwischen 1,8 und 2,0 liegen (SAMBROOK et al., 1989), niedrigere Werte weisen auf Verunreinigungen mit Proteinen hin.

2.5.2.6 Isolierung von DNA-Fragmenten mittels „QIAquick Gel Extraction“-Kit

Zur Isolierung von DNA-Fragmenten wurde der „QIAquick Gel Extraction“-Kit (Fa. Qiagen GmbH, Hilden) herangezogen. Die Methode beruht auf der Bindung der DNA an eine Silica-Membran in Gegenwart von hohen Salzkonzentrationen. Das Kit kann zur Extraktion von DNA-Fragmenten von einer Größe von 0,1 bis 10 kbp aus Agarosegelen (s. 2.7.1) genutzt werden. Die Extraktion erfolgte nach den im Herstellerprotokoll beschriebenen Arbeitsschritten. Die DNA wurde mit H₂O_{dest.} eluiert und bis zur weiteren Verwendung bei –20 °C aufbewahrt.

2.5.2.7 Reinigung von DNA-Fragmenten mittels „QIAquick PCR Purification“-Kit

Zur Reinigung von PCR-Produkten und anderen linearen DNA-Fragmenten wurde der „QIAquick PCR Purification“-Kit (Fa. Qiagen GmbH, Hilden) genutzt. Wie auch beim „QIAquick Gelextraction“-Kit (s. 2.5.2.6) wird die DNA bei hohen Salzkonzentrationen an eine Silica-Membran gebunden. Die Reinigung wurde nach den Anweisungen des Herstellers durchgeführt. Die Elution der DNA erfolgte mit $\text{H}_2\text{O}_{\text{dest.}}$. Die so gereinigten PCR-Produkte wurden bis zu ihrer weiteren Verwendung bei $-20\text{ }^{\circ}\text{C}$ gelagert.

2.6 Isolierung von Nukleinsäuren

2.6.1 Isolierung von DNA aus Umweltproben (ZHOU et al., 1996)

Diese Methode der DNA-Isolierung aus Bodenproben beruht auf der direkten Lyse der Zellen, so dass sie vor dem Aufschluss nicht von den Bodenpartikeln abgetrennt werden müssen. Weiterhin wird so die zusätzlich im Boden frei vorliegende DNA erhalten.

Zur DNA-Isolierung wurden 50 g Boden mit 135 ml DNA-Extraktionspuffer (DEP) und 1 ml Proteinase K (10 mg/ml) versetzt und für 30 min bei $37\text{ }^{\circ}\text{C}$ horizontal geschüttelt (225 Upm). Nach der Zugabe von 15 ml 20 % (w/v) SDS wurde das Gemisch für 2 h bei $65\text{ }^{\circ}\text{C}$ inkubiert und dabei alle 15 bis 20 min vorsichtig über Kopf geschwenkt. Zur Trennung der bereits gelösten DNA von der Bodenmatrix wurde anschließend zentrifugiert (10 min, 7000 Upm, Rotor GS3, Sorvall Zentrifuge, Fa. Du Pont Instruments, Newton, USA, RT) und der Überstand in einen GS3-Zentrifugenbecher überführt. Das Pellet wurde noch zweimal in je 45 ml DEP und 5 ml 20 % (w/v) SDS suspendiert und nach einer 10-minütigen Inkubation bei $65\text{ }^{\circ}\text{C}$ wie oben beschrieben zentrifugiert. Zur Extraktion der DNA wurde den gesammelten Überständen 1 Vol. Chloroform/Isoamylalkohol (24:1) zugesetzt und für 10 min bei 7000 Upm

(Rotor GS3, Sorvall Zentrifuge, Fa. Du Pont Instruments, Newton, USA) und RT zentrifugiert. Die Fällung der DNA erfolgte durch Zugabe von 0,6 Vol. Isopropanol für 1 h bei RT. Nach einer Zentrifugation bei 9000 Upm (Rotor GSA, Sorvall Zentrifuge, Fa. Du Pont Instruments, Newton, USA) für 20 min bei RT wurde das Pellet mit eiskaltem Ethanol (70 % v/v) gewaschen und in 4 bis 5 ml sterilem H₂O aufgenommen.

DNA-Extraktionspuffer (DEP):

Tris	100	mM
EDTA	100	mM
Na ₂ HPO ₄	100	mM
NaCl	1,5	M
CTAB	1 %	(w/v)
pH 8,0		

2.6.2 Reinigung von Umwelt-DNA (HENNE et al., 1999)

Bei der verwendeten Methode zur Isolierung von DNA aus Standortproben (s. 2.6.1) werden neben der DNA auch Huminsäuren gefällt, die in nachfolgenden Klonierungsschritten Restriktionsenzyme und Ligase inhibieren können. Durch Reinigung über das „WizardTM Plus Minipreps DNA Purification System“ (Fa. Promega Deutschland GmbH, Mannheim) und eine anschließende Anionenaustauschchromatographie (Qiagen-tip-100 Säulen, Fa. Qiagen GmbH, Hilden), konnten die Huminsäuren zum größten Teil abgetrennt werden. Jedoch ist diese Methode mit einem hohen Verlust an DNA verbunden.

Der erste Reinigungsschritt erfolgte nacheinander über zwei Minisäulen. Dazu wurden 150 µl DNA-Lösung mit 150 µl H₂O verdünnt, mit 1 ml „DNA Purification Resin“ versetzt und auf den Säulenkörper gegeben. Nach dem Waschen mit 2 ml „Wash“-Puffer wurde zum Trocknen des Säulenmaterials für 2 min zentrifugiert. Die Elution der DNA aus der Säulenmatrix erfolgte mit jeweils 50 µl

TE-Puffer (65 – 70 °C). Da bei dieser Methode die Bindekapazität des Säulenmaterials stark überschritten wird, wurde die Elution sechsmal wiederholt. Das vereinigte Eluat wurde wieder mit 1 ml Resin versetzt und auf die zweite Säule aufgetragen. Die Elution wurde viermal wiederholt. Sollte die Umwelt-DNA nach dieser Behandlung immer noch Verunreinigungen aufweisen, wurden Eluate aus mehreren vorher beschriebenen Reinigungs-Ansätzen vereinigt und mit 1 Vol. eines MOPS/NaCl-Puffers (1,5 M NaCl, 100 mM MOPS, pH 7,0) versetzt, um eine effiziente Bindung an die nachfolgende Anionenaustauschchromatographie-Säule zu gewährleisten. Die verwendeten Qiagen-tip-100-Säulen wurden mit 4 ml QBT-Puffer äquilibriert und die DNA-Lösung aufgetragen, und entsprechend der mitgelieferten Vorschrift mit den Puffern QC gewaschen, mit QF eluiert und mit Isopropanol gefällt. Die Puffer QBT, QC und QF wurden vom Hersteller Qiagen mitgeliefert.

2.6.3 Isolierung von Plasmid-DNA

2.6.3.1 Plasmidschnellpräparation (HOLMES und QUIGLEY, 1981)

Diese Methode diene vor allem zum schnellen Nachweis und zur Charakterisierung von Klonierungsergebnissen.

1,5 ml einer über Nacht gewachsenen Kultur wurden geerntet (5000 x g, 5 min, RT) und in 350 µl STET-Puffer suspendiert. Nach Zugabe von 25 µl Lysozym-Lösung (10 mg/ml 10 mM Tris-HCl, pH 8,0) und einer kurzen Inkubation des Ansatzes bei RT erfolgte anschließend eine Denaturierung für 1 min bei 100 °C. Zellreste und denaturierte Proteine wurden nach Zentrifugation (16000 x g, 10 min) mit einem sterilen Zahnstocher entfernt. Nach der Zugabe von 40 µl Na-Acetat (2,5 M, pH 5,2) und 420 µl Isopropanol wurde das Plasmid durch Zentrifugation (16000 x g, 5 min) pelletiert. Das enthaltene Pellet wurde mit Ethanol (70 %, (v/v)) gewaschen, getrocknet und in 50 µl sterilem H₂O_{dest.} suspendiert.

STET-Puffer:

NaCl	100	mM
Tris-HCl	10	mM
EDTA	1	mM
Triton X-100	5 %	(w/v)
pH 8,0		

2.6.3.2 Minipräparation von Plasmiden mittels „QIAprep Spin Miniprep“-Kit

Zum Erhalt von Plasmid-DNA mit einem sehr hohen Reinheitsgrad wurde der „QIAprep Spin Miniprep“- Kit (Fa. Qiagen, Hilden) genutzt. 1,5 ml einer 5-ml-Übernachtskultur wurden für 5 min bei 6000 x g zentrifugiert und das Zellpellet in 250 µl Puffer P1 suspendiert. Es folgte die Zugabe von 250 µl Puffer P2, Schwenken bis zum Aufklaren der Lösung und Neutralisation durch Zusatz von 350 µl Puffer N3. Durch Zentrifugation (10 min, 13000 x g) wurden die ausgefällten Proteine und Zelltrümmer sedimentiert. Der Überstand wurde anschließend auf eine QIAprep Säule aufgetragen, die eine Silica-Gel-Membran für die selektive Adsorption von Plasmid-DNA enthält. Durch Zentrifugation (1 min, 13000 x g) erfolgte die Bindung der DNA an die Membran, der Durchfluss wurde verworfen. Nach einem Waschschrift mit 750 µl Puffer PE und dem Trocknen durch Zentrifugation (1 min, 13000 x g) wurde die Plasmid-DNA mit 30 bzw. 50 µl H₂O_{dest.} eluiert und konnte bei -20 °C gelagert werden. Alle verwendeten Puffer wurden vom Hersteller mitgeliefert.

2.6.3.3 Minipräparation von Plasmiden mittels „NucleoSpinPlasmid“-Kit

Alternativ zum „QIAprep Spin Miniprep“- Kit der Firma Qiagen, kam auch der „NucleoSpinPlasmid“-Kit (Fa. Macherey und Nagel, Düren) zum Einsatz. Beide Kits basieren auf den gleichen Prinzipien und Vorgehen. Die Isolierung der Plasmide erfolgte nach den Angaben des Herstellers. Die Elution der Plasmid-

DNA wurde mit $\text{H}_2\text{O}_{\text{dest}}$ durchgeführt und anschließend bei $-20\text{ }^{\circ}\text{C}$ aufbewahrt. Alle Puffer wurden vom Hersteller mitgeliefert.

2.6.3.4 Präparative Plasmidisolierung über Anionenaustauschchromatographie

Die Isolierung größerer Mengen an Plasmid-DNA erfolgte nach alkalischer Lyse mit den Puffern P1, P2 und P3 über Anionenaustauschchromatographie (Qiagen Plasmid Midi-Kit; Fa. Qiagen, Hilden).

25 ml Zellkultur wurden bei einer OD_{600} von 1 bis 1,5 durch Zentrifugation geerntet und in 4 ml P1-Puffer suspendiert. Es folgte die Zugabe von 4 ml P2-Puffer und eine Inkubation für 5 min bei RT. Anschließend wurden 4 ml P3-Puffer zugefügt und für 30 min auf Eis inkubiert. Nach Zentrifugation (30 min, $16000 \times g$, $4\text{ }^{\circ}\text{C}$) wurde der Überstand abgenommen und zur vollständigen Aufklärung erneut für 15 min zentrifugiert. Der dabei erhaltene Überstand konnte dann auf eine mit 4 ml QBT-Puffer äquilibrierte Quiagen-tip-100 Säule aufgetragen werden. Es folgte ein Waschschrift mit 20 ml QC-Puffer und die Elution der Plasmid-DNA durch Zugabe von 5 ml QF-Puffer. Die DNA wurde mit 0,7 Vol. Isopropanol gefällt und durch Zentrifugation (30 min, $16000 \times g$, $4\text{ }^{\circ}\text{C}$) pelletiert. Das Sediment wurde anschließend mit Ethanol (70 %, v/v) gewaschen und getrocknet und in 100 bis 200 μl sterilem $\text{H}_2\text{O}_{\text{dest}}$ aufgenommen. Alle verwendeten Puffer wurden vom Hersteller mitgeliefert

2.7 Gelelektrophorese von Nukleinsäuren

2.7.1 Standard-Agarose-Gelelektrophorese

Zur analytischen Auftrennung von DNA-Fragmenten diente die Standard-Agarose-Gelelektrophorese (SAMBROOK et al., 1989). Es wurden Gele mit

einem Volumen von ca. 25 ml gegossen. Die Agarosekonzentration variierte zwischen 0,8 und 1,2 % (w/v). Die Proben wurden vor dem Auftragen mit 0,2 Volumen Stopp-Mix versetzt, um sie zu beschweren und um die Lauffront zu markieren. Als Elektrophoresepuffer wurde 1-fach TAE-Puffer verwendet. Die Elektrophorese wurde bei einer konstanten Spannung zwischen 80 und 120 V durchgeführt, wodurch sich eine Laufzeit von 1 bis 1,5 h ergab. Nach Beendigung der Elektrophorese wurden die Agarosegele in einem Ethidiumbromidbad (1 µg/ml H₂O) für 10 bis 30 min angefärbt. Nach kurzem Entfärben im Wasserbad konnten die Gele auf der Gel Doc 2000-Anlage (Fa. Biorad, München) bei Licht der Wellenlänge 254 nm begutachtet werden. Zur Dokumentation diente ein Videoprinter (Intas Video Graphic Printer UP-890 CE, Fa. Sony Corporation, Japan), der eine Bildübertragung auf Thermopapier ermöglichte.

Stopp-Mix:

Saccharose	50	% (w/v)
Harnstoff	7	M
EDTA	1	mM
Bromphenolblau	0,1	% (w/v)

TAE-Puffer (50x):

Tris	2	M
Eisessig	1	M
EDTA	0,1	M
H ₂ O _{dest.}	ad 1000	ml
pH 8,0		

2.7.2 Größenbestimmung von Nukleinsäuren

Als Größenstandard für lineare DNA-Fragmente wurde der "Ladder-Mix" (Fa. MBI Fermentas, St. Leon-Rot) eingesetzt. Die Größen der Fragmente betrugen in kbp:

10,0	3,5	1,2	0,6
8,0	3,0	1,03	0,5
6,0	2,5	0,9	0,4
5,0	2,0	0,8	
4,0	1,5	0,7	

und 3 weitere kleinere Fragmente, die meist im Gel nicht mehr sichtbar waren.

2.8 Enzymatische Modifikation von DNA

2.8.1 Schneiden von DNA durch Restriktionsendonukleasen

Die Sequenz-spezifische Spaltung von DNA mit Restriktionsendonukleasen lieferte lineare DNA-Fragmente mit definierten Enden. Die Spaltung erfolgte in den vom Hersteller mitgelieferten Puffersystemen bei den empfohlenen Temperaturen. Die Ansätze wurden mit einem Gesamtvolumen von mindestens 15 µl für 2 bis 15 h durchgeführt. Zur Vermeidung einer zu hohen Glycerinkonzentration im Reaktionsansatz durfte das Volumen der zugegebenen Enzymlösung höchstens 1/10 des Gesamtvolumens betragen.

Die partielle Spaltung von chromosomaler Standort-DNA wurde zeitlimitiert durchgeführt. Es wurde routinemäßig mit 1 U Restriktionsenzym gearbeitet. Die Verdauung wurde in bestimmten Zeitabständen durch Zugabe von EDTA (20 mM) und Inkubation auf Eis gestoppt.

Die partielle Spaltung der Plasmide pTWP9 und pTWP13 zur Erstellung von „Shotgun“-Plasmidgenbanken erfolgte enzymlimitiert. Dafür wurde ausgehend von 0,5 U Enzym pro 40 µl Ansatz über eine Verdünnungsreihe in 1:2-Verdünnungsschritten die Enzymkonzentration bis zu einer Verdünnungsstufe von 1:512 verdünnt. Die Restriktion erfolgte für 20 min bei entsprechender Temperatur. Abgestoppt wurde die Reaktion durch Inkubation bei 65 °C für 15 min.

2.8.2 Dephosphorylierung von DNA-Fragmenten

Zur Erhöhung der Effizienz einer Ligationsreaktion zwischen Insert-DNA und Vektor-DNA wurde die Vektor-DNA an den 5'-Enden dephosphoryliert. Auf die Weise sollte eine Rezirkularisierung des Vektors verhindert werden. Es wurde die alkalische Phosphatase der Fa. MBI Fermentas GmbH eingesetzt. Nach vollständiger Spaltung des Vektors wurde 1 U alkalische Phosphatase pro μg DNA zugesetzt und für 30 min bei 37 °C inkubiert.

2.8.3 Ligation von DNA-Fragmenten

2.8.3.1 Ligation mit T4-DNA-Ligase

Zur Verknüpfung von linearisierter, dephosphorylierter Vektor- und Insert-DNA wurde T4-DNA-Ligase eingesetzt, die in Anwesenheit von ATP eine Phosphodiesterbindung zwischen 3'-OH und 5'-PO₄-Enden von Nukleinsäuren katalysiert. Die Ligationen wurden in einem Volumen von 20 μl durchgeführt. Es wurden jeweils 0,3 bis 1 μg Vektor- bzw. Insert-DNA eingesetzt. Der Ansatz enthielt weiterhin 1 U T4-DNA-Ligase und in 1-facher Konzentration den vom Hersteller mitgelieferten Reaktionspuffer. Die Ligationsansätze mit kohäsiven Enden wurden über Nacht bei 16 °C, für 2 bis 4 h bei Raumtemperatur oder für 48 h bei 4 °C inkubiert. Zur Ligation glatter Enden wurden 2 U Ligase eingesetzt und die Ansätze für 1 h bei 37 °C inkubiert.

2.8.3.2 Ligation mit dem „TOPO-T/A-Cloning“-Kit

Für eine besonders effektive Klonierung von PCR-Produkten, die mit der *Taq*-DNA-Polymerase in einer Polymerasekettenreaktion (PCR, s. 2.11) amplifiziert wurden, wurde der „TOPO-TA-Cloning®“-Kit (Fa. Invitrogen, Carlsbad, USA) verwendet. Für diese Art der Ligation wird keine Ligase benötigt. Alle Schritte wurden laut Hersteller-Angaben durchgeführt.

Die *Taq*-DNA-Polymerase besitzt eine von der Matrize unabhängige terminale Transferase-Aktivität und hängt an die 3'-Enden von PCR-Produkten einzelne Desoxyadenosine (A) an. Der im Kit vorhandene, bereits linearisierte Vektor pCR2.1-TOPO besitzt analog dazu an den 3'-Enden einzelne Desoxythymidin-Rest (T). Dies erlaubt eine Klonierung nach der so genannten TA-Methode (s. Abb. 4). Des Weiteren besitzt der Vektor die kovalent gebundene Topoisomerase I des *Vaccinia*-Virus. Die Topoisomerase I bindet an spezifische Sequenzen doppelsträngiger DNA und spaltet dort das Phosphodiester-Gerüst stromabwärts der Sequenz 5'-CCCTT in einem Strang (SHUMAN, 1991).

Abbildung 4: Schematische Darstellung der TOPO-TA-Klonierung.

Die Energie dieser Spaltungsreaktion wird durch Ausbildung einer kovalenten Bindung zwischen dem 3'-Phosphat des gespaltenen Stranges und dem Tyrosin-274 der Topoisomerase I konserviert. Die Reaktion zwischen DNA und Protein kann durch den Angriff des 5'-Hydroxyl-Restes umgekehrt werden, was dann zur Abspaltung der Topoisomerase I führt (SHUMAN, 1994). Dieses Prinzip wird zur effizienten Klonierung von PCR-Produkten genutzt.

2.9 Saccharose-Dichtegradienten-Zentrifugation (SAMBROOK et al., 1989; modifiziert)

Die Größenfraktionierung von Fragmenten partiell verdauter, chromosomaler DNA, die zur Klonierung eingesetzt werden sollte, erfolgte mittels einer Saccharose-Dichtegradientenzentrifugation. Bei dieser sehr schonenden Methode werden die verschiedenen Fragmentgrößen entsprechend ihrer Schwebedichte im Saccharosegradienten getrennt. Zunächst wurde ein Saccharose-Stufengradient (10 - 40 %) in 10 ml-Zentrifugenröhrchen (Sorvall Polyallomer 3699, Fa. Du Pont de Nemours Deutschland GmbH, Dreieich) hergestellt. Dazu wurden je 1,9 ml 40 %, 32,5 %, 25 %, 17,5 % und 10 % Saccharose (jeweils w/v in 1 M NaCl, 20 mM Tris-HCl, 5 mM EDTA, pH 8,0) beginnend mit der höchsten Konzentration übereinander geschichtet und nach jeder Stufe eingefroren. Auf den fertigen Gradienten wurde 100 - 200 µg partiell verdaute chromosomale DNA aufgetragen. Die Zentrifugation erfolgte in einem „swing-out“-Rotor (Sorvall TH641, Fa. Du Pont de Nemours) in einer Ultrazentrifuge (Sorvall OTD 55 B, 27000 Upm, 24 h, 20 °C). Zur Ernte des Gradienten wurden vorsichtig je 400 µl von oben beginnend abgenommen. Anschließend erfolgte eine DNA-Fällung mit PEG/NaCl aus Saccharose (s. 2.5.2.2). 10 µl jeder dritten Fraktion wurden durch Agarosegelelektrophorese (s. 2.7.1) auf Größe und Qualität hin überprüft.

2.10 Übertragung von DNA in *E. coli* und Selektion rekombinanter Klone

2.10.1 Hocheffiziente Transformation in *E. coli* (INOUE et al., 1990)

Zur Herstellung hocheffizienter Transformanten wurde *E. coli* aus einer Stammkultur auf eine LB-Agar-Platte ausgestrichen und über Nacht bei 37 °C inkubiert. Aus 10 bis 12 Einzelkolonien wurde eine 5 ml Vorkultur angeimpft. Diese diente als Inokulum für 250 ml SOB-Medium (s. 2.2.2). Diese Kultur wurde bei 16 °C schüttelnd bis zu einer OD₆₀₀ von 0,6 inkubiert. Nach einer

Inkubation für 10 min auf Eis wurden die Zellen zentrifugiert (10 min, 2500 x g, 4 °C), in 80 ml eiskalter TB-Lösung suspendiert und für 10 min auf Eis abgekühlt. Es folgte eine Zentrifugation der Zellen (10 min, 2500 x g, 4 °C) und Suspension in 20 ml TB-Lösung mit DMSO (Endkonzentration 7 %, v/v). Nach einer Inkubation für 10 min auf Eis wurden Aliquots von 200 µl in flüssigem N₂ eingefroren und direkt zur Transformation eingesetzt oder bei - 70 °C gelagert.

Zur Transformation wurden 200 µl Zellsuspension auf Eis aufgetaut und nach der Zugabe von 1 bis 20 µl DNA-Lösung (ca. 10 bis 500 ng) für 30 min auf Eis inkubiert. Es folgte ein Hitzeschock für 1 min bei 42 °C, eine kurze Inkubation auf Eis und die Zugabe von 0,8 ml LB-Medium (s. 2.2.1). Zur Ausprägung der plasmidkodierten Antibiotikaresistenz wurden die Ansätze für 60 min bei 37 °C inkubiert und anschließend die Suspension auf LB-Agarplatten mit entsprechendem Antibiotikum ausplattiert.

TB-Lösung:

Pipes	10	mM
MnCl ₂ *	55	mM
CaCl ₂	15	mM
KCl	250	mM
pH 6,7		

*Das MnCl₂ wurde ungelöst und separat autoklaviert!

2.10.2 Transformation in *E. coli* durch Elektroporation (DOWER et al., 1988)

Diese Methode beruht auf der Permeabilisierung der Zellmembran durch Elektroschock. Dieser verursacht einen kurzfristigen reversiblen lokalen Zusammenbruch des Membranpotentials bzw. eine Desorganisation der Membranstruktur (CHASSY et al., 1988) und erleichtert damit die Aufnahme von Makromolekülen wie z. B. Plasmid-DNA in die Zelle.

Zur Herstellung elektrokompetenter Zellen wurde eine 5-ml-Übernachtskultur in 250 ml LB-Medium (s. 2.2.1) überimpft und bei 30 °C unter Schütteln bis zu einer OD₆₀₀ von 0,5 bis 1,0 inkubiert. Die Kultur wurde für 15 bis 30 min auf Eis gehalten und danach für 10 min bei 5000 x g und 4 °C zentrifugiert. Das Pellet wurde dann zweimal mit 1 Volumen H₂O_{dest.} gewaschen und anschließend noch einmal mit 10 ml Glycerin (10 %, (v/v)). Abschließend erfolgte die Suspendierung in 0,5 ml Glycerin (10 %, v/v). Aliquots der Suspension wurden bei -70 °C gelagert. Pro Ansatz wurden 40 µl Zellsuspension auf Eis aufgetaut, in eine sterile, vorgekühlte Elektroporationsküvette (Elektrodenabstand 0,2 cm) überführt und mit 1 bis 5 µl dialysierter DNA-Lösung versetzt. Die Elektroporation erfolgte in einem Gene Pulser TM (Bio Rad Laboratories GmbH, München) bei 25 µF, 200 Ω und 2,5 kV, wodurch eine Feldstärke von 12,5 kV/cm und eine Zeitkonstante von 3,5 bis 4,5 ms erreicht wurde. Nach dem Stromimpuls wurden die Zellen sofort in 1 ml LB-Medium aufgenommen, 60 min bei 37 °C inkubiert und anschließend auf LB-Agarplatten mit den entsprechenden Antibiotika ausplattiert.

2.10.3 Der X-Gal-Test zur Selektion rekombinanter Klone

Bei Klonierungen ist es entscheidend, Klone, die nach Übertragung von DNA nur den Vektor enthalten, von denjenigen die ein Plasmid mit Insert tragen, zu unterscheiden. Aus diesem Grund wurde ein zusätzliches Screeningsystem neben der plasmidkodierten Antibiotikaresistenz benutzt. Plasmide der pBluescript-Reihe enthalten neben dem Promotor- und Operator-Bereich des *lac*-Operons auch das 5'-Ende des *lacZ*-Genes (VIEIRA und MESSING, 1982). Dieses kodiert für das so genannte α-Peptid, welches die inaktive β-Galaktosidase von Wirtsstämmen wie *E. coli* DH5α komplementieren kann. Da sich die multiple Klonierungsstelle der verwendeten Vektoren innerhalb des aminoterminalen Bereichs des *lacZ*-Gens befindet, ist eine α-Komplementation nur möglich, wenn das Plasmid kein Insert besitzt. Der Nachweis der aktiven β-Galaktosidase auf Agarplatten erfolgte durch Zugabe des Induktors IPTG (Isopropyl-β-thiogalaktopyranosid) und des Substrats X-Gal (5-Brom-4-chlor-

indoyl- β -D-galaktosid), welches durch das Enzym gespalten wurde. Dabei entstand in Anwesenheit von Sauerstoff der blaue Farbstoff 5-Brom-4-chlor-indigo, welcher der Indikator für rekombinante *E. coli*-Stämme ist, die ein Plasmid ohne Insert tragen.

2.10.4 Screening auf die Verwertung von Butanol

Die Überprüfung rekombinanter *E. coli*-Stämme auf die Verwertung von Butanol wurde auf Aldehyd-Indikator-Agar (s. 2.2.4) durchgeführt. Das Medium enthielt das zu testende Substrat Butanol (0,2 % v/v) sowie den Farbstoff Pararosanilin und Natriumbisulfit. Der Farbstoff bildete zusammen mit dem Natriumbisulfit ein Schiff'sches Reagenz, das farblos ist. Werden nun bei der Substrat-Verwertung Carbonyle (Aldehyde bzw. Ketone) gebildet, kommt es zur Ausbildung einer Schiff'schen Base, die intensiv rot gefärbt war. Daher wiesen Kolonien von zur Butanol-Verwertung befähigten *E. coli*-Klonen in diesem Medium eine rote Farbe auf und färbten auch den Agar rot an, alle anderen Klone blieben hell (s. Abb. 5). Für das Screening wurde als Wirtstamm die *E. coli*-Mutante WL3 verwendet. Nach der Transformation von rekombinanten Plasmiden wurde auf den Indikatorplatten ausplattiert und für 16 bis 72 h bei 30 oder 37 °C inkubiert.

Abbildung 5: Testprinzip der Aldehyd/Keton-Indikatorplatten.

2.10.5 Plattentest zum Screening auf proteolytische Aktivität

Der Nachweis proteolytischer Aktivität bei rekombinanten *E. coli*-Stämmen fand auf LB-Agar statt, der mit Skim-Milk (2 % w/v) versehen war. Die Skim-Milk wurde zuvor in etwas $\text{H}_2\text{O}_{\text{dest.}}$ gelöst, getrennt vom Medium autoklaviert und nach dem Autoklavieren zum Medium zugegeben. Proteolytische Aktivität führt zur Hydrolyse der Skim Milk, welche dadurch deutlich aufklart. Für das Screening wurden *E. coli*-Top10-Zellen transformiert und geeignete Verdünnungen auf Skim-Milk-Agar ausplattiert. Positive Klone konnten nach 2 bis 3 Tagen Wachstum bei 37 °C, durch eine deutliche Hofbildung (Aufklärung) um die Kolonie detektiert werden.

2.11 Polymerase-Kettenreaktion (PCR)

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) oder die *Pfu*-DNA-Polymerase (Fa. MBI Fermentas) eingesetzt. Die Reaktionen wurden entsprechend den Herstellerempfehlungen in einem Volumen von 50 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)	5	μl
MgCl_2 oder MgSO_4	1,5	mM
$\text{H}_2\text{O}_{\text{dest.}}$	ad 50	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
<u>30 Zyklen</u>		
Denaturierung:	95 °C	1 min
Hybridisierung:	x°C	1 min
Kettenverlängerung:	72 °C	y min
<u>Ende der Zyklen</u>		
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 Hybridisierung wurde entsprechend der zu erwartenden Fragmentlänge eine Zeit von ca. 1 min/1000 Bp für die *Taq*-Polymerase bzw. 2 min/1000 Bp für die *Pfu*-Polymerase 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. Die Temperaturzyklen wurden mit dem „Mastercycler Gradient“ (Fa. Eppendorf, Hamburg) durchgeführt. Die PCR-Produkte wurden durch Agarose-Gelelektrophorese (s. 2.7.1) analysiert.

2.12 Sequenzierung und Analyse von Sequenzdaten

2.12.1 Durchführung von DNA-Sequenzierungen

Alle Sequenzierungen von rekombinanten Plasmiden und PCR-Produkten wurden durch das „Göttingen Genomics Laboratory“ (Institut für Mikrobiologie und Genetik, Universität Göttingen) oder von der Firma SeqLab (Göttingen) mit den jeweiligen firmeninternen Verfahren durchgeführt.

2.12.2 Sequenzierungsprojekte pTWP9 und pTWP13

Zur Ermittlung der Insertsequenzen der Plasmide pTWP9 und pTWP13 wurden „Shotgun“-Genbanken von den jeweiligen Plasmiden erstellt und diese wie unter 2.12.1 beschrieben sequenziert. Nach Erhalt der Sequenzdaten erfolgte zunächst die Prozessierung der Rohsequenzen.

2.12.3 Prozessierung der Rohsequenzen

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. Hier soll nur ein Überblick über die Anwendung des Software-Packets im Rahmen der Sequenzierungsprojekte pTWP9 und pTWP13 gegeben werden. Eine ausführliche Dokumentation kann unter http://www.mrc-lmb.cam.ac.uk/pubseq/staden_home.html abgerufen werden. Im wesentlichen sind es zwei Programme, die zur Prozessierung der DNA-Sequenzen benutzt wurden: „Pregap4“ und „Gap4“.

2.12.4 „Pregap4“ und „Gap4“

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

- Screening auf Vektor-Kontaminationen: Sequenzen, die den Vektor-Sequenzen (pBluescript[®] IISK⁺ oder pCR2.1-TOPO) entsprechen, fallen aus der Datenbank heraus
- Screening auf Kontamination mit *E. coli* DNA: bei der Herstellung der Genbank kann es zu Kontaminationen mit der DNA des Wirtsstammes (*E. coli* TOP10) kommen. Diese Daten müssen ebenfalls entfernt werden
- Qualitätskontrolle: Sequenzen mit einer geringen Qualität fallen aus der Datenbank heraus
- „Vektor-Clipping“: Es werden die Basen einer Sequenz maskiert, die vor dem Beginn der Insertsequenz, der Vektor-Sequenz entsprechen
- Umformatierung in ein neues Dateiformat: Zur Assemblierung benötigt „Gap4“ ein anderes Dateiformat („experiment files“)

Zur Erleichterung der Prozessierung der Rohdaten wurden im Göttinger Genomanalyselabor Computerskripte entwickelt, die hier nicht näher erläutert werden sollen. 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.).

„**Gap4**“ ist ein Genomassemblierungsprogramm, es kann aber auch für kleinere DNA-Abschnitte wie Cosmide oder große Plasmide verwendet werden (BONFIELD et al., 1995). Neben der Assemblierung kann das Programm auch zur Editierung der Sequenzen benutzt werden. „Gap4“ ist des Weiteren unerlässlich beim so genannten „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 so genannten „Contigs“ zusammengefasst. Aus den „alignten“ 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 10.000 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.

2.12.5 Methoden zur Schließung von Sequenzlücken

2.12.5.1 Identifikation von „Conitig“-Überlappungen mit „Gap4“ oder „Align two Sequences“ (NCBI)

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 „Contig“-Überlappung bedurfte im „Join Contigs“ Editor einer genauen Überprüfung. Zum anderen wurde ‚extern‘ nach

überlappenden „Contigs“ gesucht, indem die Enden einiger „Contigs“ gegen andere „Contigs“ bei NCBI („National Center for Biotechnology; <http://www.ncbi.nlm.nih.gov>) mit dem BLAST-Algorithmus „Align two Sequences“ verglichen wurden. Contigs, die potentiell überlappten, wurden dann in Gap4 manuell über „Join Contigs“ zueinander ausgerichtet und gegebenenfalls miteinander verbunden.

2.12.5.2 „Primer-Walking“

Mit Hilfe der „Gap4“-Funktion „Template Display“ können einzelne „Contigs“ visuell dargestellt werden; dort sind alle im „Contig“ enthaltenen „Readings“ und deren „Templates“ aufgeführt. Ein „Template“ entspricht dabei einem Insert eines rekombinanten Plasmids der „Shotgun“-Genbank. Anzumerken ist, dass die Inserts der „Shotgun“-Plasmide von beiden Seiten her ansequenziert wurden („forward“ und „reverse“). Da die durchschnittliche Insertgröße bei etwa 2-4 kb lag und ein typischer Sequenzierlauf etwa 600-800 Basen lang war, bedeutete dies, dass ein Insert (=„Template“) nicht vollständig sequenziert wurde. Im „Template Display“ nun 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 (<http://www.ncbi.nlm.nih.gov>) mit „BLASTX“ identifiziert werden. Damit war es nun möglich, durch ein so genanntes „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 dem Ableiten eines weiteren Oligonukleotids schließen. Konnte auch

mit dieser Methode keine Lücken schließende Sequenz erhalten werden, wurde versucht mit einer PCR-Reaktion, in welcher das ursprüngliche Plasmid als „Template“ eingesetzt wurde und den die Lücke flankierenden Oligonukleotiden, das fehlende DNA-Fragment zu erhalten. Dieses PCR-Produkt konnte dann ebenfalls sequenziert werden.

Nach Beendigung der Sequenzierungsphase und dem Schluss der letzten Lücken in den Insert-Sequenzen von pTWP9 und pTWP13 konnte mit der Sequenzanalyse begonnen werden.

2.12.6 Sequenzierungsprojekte pTWB3 und pTWB4

Die Insertgrößen der Plasmide pTWB3 und pTWB4 waren deutlich kleiner als die von pTWP9 und pTWP13. Deswegen konnten die Insert-Sequenzen durch „Primer-Walking“ ermittelt werden. Dazu wurden so lang neue Oligonukleotide in „forward“- und in „reverse“-Leserichtung abgeleitet, bis die Insert-Sequenzen jeweils doppelt abgedeckt waren. Die Bearbeitung der Rohsequenzen (Prozessierung, Assemblierung und Editierung) wurde mit dem Programm „Sequence Manager“ auf einem Macintosh PowerPC (Fa. Apple Computer Inc., Cupertino, USA) durchgeführt. Im Anschluss konnte mit der Sequenzanalyse begonnen werden.

2.12.7 Sequenzanalyse

Nach Beendigung der Sequenzierungsphase, erfolgte die Sequenzanalyse, dazu zählte zunächst das Auffinden aller in der Sequenz vorkommenden Gene (ORF-Vohersage). Dieses wurde mit entsprechenden Computeralgorithmen durchgeführt. Anschließend erfolgte die Annotation der festgelegten Gene, das heißt jedem ORF wurde versucht eine Funktion zuzuordnen, die das aus dem Gen resultierende Protein beschreibt.

2.12.7.1 ORF („open reading frame“)-Vorhersage

Alle ORF-Vorhersageprogramme für Prokaryonten basieren auf dem Auffinden von Start- (ATG, GTG, TTG)- und Stopp-Codons (TAA, TGA, TAG) in allen sechs möglichen Leserahmen. Danach werden solche Sequenzbereiche festgelegt, die zwischen dem Start- und Stopp-Codon auf demselben Leserahmen eine bestimmte „Cutoff“-Größe erreichen (oft 30 Codons). Dieser Bereich wird als ORF bezeichnet, stellt also ein potentielles Gen dar. Die ORF-Vorhersage wurde manuell mit der frei im Internet erhältlichen Software „Artemis“ (<http://www.sanger.ac.uk/Software/Artemis>; RUTHERFORD et al., 2000) durchgeführt. Es ist möglich, sich mit dem Programm alle potentiellen ORFs mit einer minimalen Größe darstellen zu lassen. Man kann die Größe des ORFs aber auch definieren. In dieser Arbeit wurde eine ORF-Größe von mindestens 30 Codons festgelegt. Die daraus resultierenden ORFs in den 6 Leserahmen wurden nun mit „Artemis“ auf weitere Parameter untersucht.

- **Ribosomenbindestelle:** Bei jedem ORF wurde überprüft, ob 4-12 Basen stromaufwärts des Start-Codons eine Ribosomenbindestelle vorhanden war
- **ORF-Länge:** Bei der Festlegung der Länge eines ORFs war die Art des Vorhandenseins der Ribosomenbindestelle und die Art des Start-Codons entscheidend; ATG-Start-Codons wurden gegenüber GTG und TTG bevorzugt. Mit der Funktion „Trim selected features to next any“ konnte das nächste, stromabwärts liegende Start-Codon gefunden werden
- **GC-Frameplot:** Bei „Artemis“ ist es möglich, sich die G+C-Häufigkeit an den 3 Codonpositionen anzeigen zu lassen. Empirische Untersuchungen (persönliche Mitteilung Dr. Heiko Liesegang, Universität Göttingen) haben gezeigt, dass sich diese Häufigkeiten am Start und am Stopp eines Gens häufig umkehren. Dies kann eine Hilfe und ein Hinweis für das Vorliegen eines Gens sein, ist aber keine notwendige Bedingung.

Mit Artemis ist es auch möglich, sich die kodierende Aminosäuresequenz der potentiellen ORFs anzeigen zu lassen. Somit konnte im nächsten Schritt versucht werden durch Sequenzabgleich auf Proteinebene bei der NCBI-Sequenzdatenbank (<http://www.ncbi.nlm.nih.gov/BLAST/>), jedem potentiellen ORF eine Funktion zuzuweisen (Annotation).

2.12.7.2 Annotation

Die Annotation hat zum Ziel, jedem aus einem Gen resultierenden Protein eine Funktion zuzuordnen. Dazu erfolgte von jedem ORF ein manueller Sequenzabgleich auf Proteinebene bei der NCBI-Sequenzdatenbank (<http://www.ncbi.nlm.nih.gov/BLAST/>). Eine signifikante Homologie zwischen zwei Proteinsequenzen wird im allgemeinen definiert als eine Proteinidentität größer als 30 % bei einer Überlappung („Alignment“), von mehr als 60 % der Eingabe („query“)- und der Vergleichssequenz („subject“). Die Identifizierung von konservierten Domänen erfolgt durch Sequenzabgleiche mit den Domänen-Datenbanken „Pfam“ (<http://www.sanger.ac.uk/Software/Pfam>) und „ProSite“ (<http://www.expasy.org/prosite>). Wenn keine Ähnlichkeiten zu bereits bekannten Sequenzen vorhanden waren, wurde der entsprechende ORF als hypothetisch bezeichnet.

2.13 Standardtechniken für das Arbeiten mit Proteinen

2.13.1 Herstellung von Rohextrakten

Die Herstellung von Rohextrakten erfolgte durch Zellaufschluss mit der French-Press. Zunächst wurden die Zellen durch Zentrifugation (5000 Upm, 15 min, 4 °C, GS3-Rotor, Sorvall RC-5B-Zentrifuge, Fa. Du Pont de Nemours GmbH) geerntet und das Pellet in einem geeigneten Puffer aufgenommen. Die

Zellsuspension wurde dann in eine vorgekühlte French-Press-Apparatur (Fa. SLM Aminco, Urbana, USA) gefüllt. Der Aufschluß der Zellen fand bei einem Druck von 140 Mpa statt. Der Vorgang wurde zwei- bis dreimal wiederholt, wobei vor dem ersten Mal eine Spatelspitze DNase I (Fa. Boehringer, Mannheim) zugegeben wurde. Anschließend erfolgte eine Zentrifugation bei 15000 Upm und 4 °C für 30 min (SS34-Rotor, Sorvall RC-5B-Zentrifuge, Fa. Du Pont de Nemours GmbH). Der erhaltene Extrakt wurde als Rohextrakt bezeichnet und konnte bis zur weiteren Verwendung bei –20 °C gelagert werden.

2.13.2 Proteinbestimmung

Die Bestimmung der Proteinkonzentration erfolgte in modifizierter Form nach der Methode von BRADFORD (1976). Es wurden 20 µl einer Probe zu 1 ml Bradford-Reagenz gegeben. Wenn nötig, konnte die Probe mit H₂O_{dest.} verdünnt werden. Nach einer Inkubation für 10 min bei RT wurde die Absorption bei einer Wellenlänge von 580 nm gegen einen Leerwert gemessen. Zur Konzentrationsbestimmung wurde eine Eichgerade mit Rinderserumalbumin in einem Bereich von 0 bis 10 µg/ml aufgenommen.

Bradford-Reagenz:

Serva Blau G-250	70	mg
Ethanol (96 %, v/v)	50	ml
H ₃ PO ₄ (85 %, v/v)	100	ml
H ₂ O _{dest.}	ad 1000	ml

Nach vollständigem Lösen des Farbstoffs in Ethanol und Zugabe der übrigen Bestandteile erfolgte eine Filtration durch einen Faltenfilter. Das Reagenz wurde lichtgeschützt gelagert und war ca. 1 Jahr haltbar.

2.13.3 SDS-Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Die Auftrennung von Proteinen nach ihrem Molekulargewicht unter denaturierenden Bedingungen durch Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) erfolgte in Minigel-Elektrophorese-Kammern (Fa. Bio-Rad, München) mit 8,5 x 8,0 x 0,1 cm Glasplatten. Trenn- und Sammelgel hatten folgende Zusammensetzung (Mengen für 2 Gele):

<u>Trenngel:</u>		<u>10% iq</u>	<u>12% iq</u>
	Acrylamidlösung	3,2 ml	4 ml
	Trenngelpuffer, pH 8,8	2 ml	2 ml
	H ₂ O _{dest.}	4,54 ml	3,9 ml
	10 % SDS-Lösung	100 µl	100 µl
	APS-Lösung	60 µl	60 µl
	TEMED	8 µl	8 µl
<u>Sammelgel:</u>		<u>10% iq</u>	<u>12% iq</u>
	Acrylamidlösung	800 µl	375 µl
	Sammelgelpuffer, pH 6,8	500 µl	500 µl
	H ₂ O _{dest.}	3,6 ml	3 ml
	10 % SDS-Lösung	50 µl	50 µl
	APS-Lösung	20 µl	20 µl
	TEMED	6 µl	6 µl

Zunächst wurden Acrylamidlösung, Gelpuffer, SDS-Lösung. und H₂O_{dest.} gemischt. Nach der Zugabe von TEMED und APS-Lösung wurde das Gel gegossen und mit H₂O_{dest.} überschichtet, um eine ebene Fläche zu erhalten. Nach Polymerisation des Trenngels (ca. 30 min) konnte das H₂O_{dest.} abgenommen und das Sammelgel gegossen werden, in das zur Formung von Geltaschen ein Kamm luftblasenfrei eingesetzt wurde. Nach 30 min war das Gel vollständig polymerisiert und konnte verwendet werden. Die Proben wurden vor dem Auftragen im Verhältnis 2:1 mit Denaturierungspuffer gemischt und für 5

min bei 96 °C im Heizblock denaturiert. Anschließend wurden die Proben auf das mit Elektrophoresepuffer überschichtete Gel mit Hilfe einer Hamiltonspritze aufgetragen. Die Elektrophorese erfolgte zunächst bei einer Spannung von 100 V, bis die blaue Markerbande das Trenngel erreicht hatte, und wurde dann auf 130 V erhöht. Als Netzgerät diente ein Gene-Power-Supply GPS 200/400 (Fa. Amersham Biosciences). Die Elektrophorese wurde beendet wenn die Markerbande das Gelende erreicht hatte.

Acrylamidlösung

Acrylamid	29,1	g
Bisacrylamid	0,9	g
H ₂ O _{dest.}	ad 100	ml

Trenngelpuffer

Tris	56,8	g
H ₂ O _{dest.}	ad 250	ml
pH 8,8 (mit 5N HCl)		

Glycin	144,2	g
SDS	10	g
H ₂ O _{dest.}	ad 1000	ml
pH 8,3 (nicht einstellen!)		

Sammelgelpuffer

Tris	37,8	g
H ₂ O _{dest.}	ad 250	ml
pH 6,8		

Denaturierungspuffer

Tris	0,2	g
SDS	1,2	g
Serva-Blau G250	3	mg
Glycerin	3	ml
2-Mercaptoethanol	0,5	ml
H ₂ O _{dest.}	ad 10	ml
pH 6,8		

APS-Lösung

Ammoniumpersulfat	1	g
H ₂ O _{dest.}	ad 10	ml

SDS-Lösung

SDS	10	g
H ₂ O _{dest.}	ad 100	ml

2.13.4 Proteinfärbung in SDS-Polyacrylamidgelen

2.13.4.1 Coomassie-Färbung

Nach Beendigung der Elektrophorese wurden die Gele kurz mit $\text{H}_2\text{O}_{\text{dest.}}$ gespült und dann für 60 min in die Coomassie-Färbelösung gelegt. Nach Spülen mit $\text{H}_2\text{O}_{\text{dest.}}$ wurde das Gel entfärbt, wobei die Entfärbelösung mehrfach gewechselt wurde. Bei guter Transparenz des Hintergrundes konnten die Gele ausgewertet werden und in Fixierlösung aufbewahrt werden.

Coomassie-Färbelösung:

Coomassie-Blau G-250	0,08	% (w/v)
Coomassie-Blau R-250	0,08	% (w/v)
Methanol	45	% (v/v)
Eisessig	10	% (v/v)

Entfärbelösung:

Methanol	33	% (v/v)
Eisessig	10	% (v/v)

Fixierlösung:

Eisessig	10	% (v/v)
----------	----	---------

2.13.4.2 Silberfärbung (BLUM et al., 1987)

Die Silberfärbung ist eine sehr empfindliche Färbemethode, da selbst geringe Mengen an Protein nachgewiesen werden können. Das Gel wurde dazu nach der Elektrophorese in folgenden Schritten bearbeitet:

- mindestens 1 h (höchstens 24 h) im Fixierer inkubieren
- 3 x 5 min in 50 % (v/v) Ethanol waschen
- 1 min in der Thiosulfatlösung inkubieren
- 3 x 20 sec in $\text{H}_2\text{O}_{\text{bidest}}$ waschen
- 20 min imprägnieren
- in Entwicklerlösung inkubieren, bis die Banden sichtbar sind
- sofort kurz in $\text{H}_2\text{O}_{\text{bidest}}$ spülen
- 10-20 min in Stopplösung inkubieren (bis leichte Trübung entsteht)
- mindestens 15 min in $\text{H}_2\text{O}_{\text{bidest}}$ waschen

Die Gele wurden in Plastikfolie eingeschweißt und sofort mit Hilfe eines Scanners dokumentiert, da die Gefahr des Nachdunkelns bestand. Die Lagerung erfolgte bei 4 °C.

Fixierer:

Methanol	50	ml
Eisessig	12	ml
Formaldehyd	100	µl
H ₂ O _{bidest}	ad 100	ml

Die Methanol-Eisessiglösung wurde auf Vorrat angesetzt; nur Formaldehyd wurde unmittelbar vor Gebrauch zugesetzt.

Thiosulfatlösung:

Na ₂ S ₂ O ₃ x 5 H ₂ O	2	mg
H ₂ O _{bidest}	ad 100	ml

Die Lösung wurde stets frisch angesetzt.

Imprägnierlösung:

AgNO ₃	200	mg
Formaldehyd	75	µl
H ₂ O _{bidest}	ad 100	ml

Die Lösung wurde stets frisch angesetzt.

Entwicklerlösung:

Na ₂ CO ₃ x 10 H ₂ O	6	g
Thiosulfatlösung (s.o.)	2	ml
Formaldehyd	50	µl
H ₂ O _{bidest}	ad 100	ml

Die Lösung wurde stets frisch angesetzt.

Stopplösung:

EDTA	1,86	g
H ₂ O _{bidest}	ad 100	ml

Die Lösung wurde mit NaOH auf einen pH-Wert von 7,0 titriert und konnte mehrere Wochen gelagert werden.

2.13.5 Bestimmung der Molekularmassen von Proteinen

Um das Molekulargewicht von Proteinen in SDS-Gelen und auf Blotmembranen bestimmen zu können, wurden gleichzeitig auch entsprechende Standardproteine mit aufgetrennt. Die Zusammensetzung der eingesetzten Proteinmarker und die molekularen Massen nach Herstellerangaben sind nachfolgend aufgeführt.

Protein Molecular Weight Marker (Fa. MBI Fermentas):

Protein	Herkunft	kDa
β-Galaktosidase	<i>E. coli</i>	116
BSA	Rinderserum	66
Ovalbumin	Hühnereiweiß	45
Laktat-Dehydrogenase	Schweinemuskel	35
RE Bsp981	<i>E. coli</i>	25
β-Laktoglobulin	Kuhmilch	18
Lysozym	Hühnereiweiß	14

peqGOLD Prestained Protein Marker (Fa. PEQLAB Biotechnologie)*:

Protein	Herkunft	kDa
β-Galaktosidase	<i>E. coli</i>	118
BSA	Rinderserum	85
Ovalbumin	Hühnereiweiß	47
Laktat-Dehydrogenase	Schweinemuskel	36
β-Laktoglobulin	Kuhmilch	26
Lysozym	Hühnereiweiß	20

*Kovalent gebundene Farbstoffe beeinflussen das Laufverhalten von Proteinen, deswegen ist dieser Marker nur für annäherungsweise Molekulargewichtsbestimmungen geeignet und wurde deswegen nur für Gele verwendet, die im Anschluss geblottet wurden.

Nach der SDS-PAGE und Färbung der Proteine konnte die Molekularmasse eines Proteins ermittelt werden, indem dessen Wanderungsgeschwindigkeit mit der von Markerproteinen verglichen wurde. Die Auftragung des Logarithmus der

Molekularmassen der Markerproteine gegen den Rf-Wert (Wanderungsstrecke des Proteins bezogen auf die Wanderungsstrecke des Farbstoffmarkers) ergibt eine Eichgerade, anhand derer die Molekularmasse des jeweiligen Proteins berechnet werden konnte.

2.13.6 Konzentrierung und Umpufferung von Proteinlösungen

2.13.6.1 Konzentrierung und Umpufferung mit Hilfe von Vivaspin Konzentratoren

Um Proteinlösungen zu konzentrieren, wurden Mikrokonzentratoren (Vivaspin 6 ml; Fa. Sartorius, Göttingen) mit einer Ausschlussgröße von 10 und 30 kDa verwendet. Dabei konnten Proteinlösungen von 6 ml auf 0,5 ml eingeeengt und umgepuffert werden. Alle Schritte wurden nach Angaben des Herstellers durchgeführt. Die Zentrifugation erfolgte nach Herstellerangaben bis das gewünschte Volumen erreicht war.

2.13.6.2 Konzentrierung und Umpufferung mit dem Vivaflow 200 Komplettsystem

Um größere Volumen (z.B. die Kulturüberstände der Protease-Expressionskulturen) einzuengen, wurde das Vivaflow 200 Komplettsystem (Fa. Sartorius, Göttingen) komplementiert mit einer PES (Polyethersulfon)-Membran (Ausschlussgröße: 30 kDa; Fa. Sartorius, Göttingen) eingesetzt. Damit konnten einerseits innerhalb von ca. 1 bis 2 Stunden 2 l Kulturüberstand auf ca. 30 ml eingeeengt werden. Andererseits konnte das System auch zur Umpufferung der konzentrierten Proteinlösung benutzt werden. Alle Schritte wurden nach Angaben des Herstellers durchgeführt.

2.13.7 Entsalzen von Proteinlösungen

Für die Entsalzung von Proteinlösungen wurden PD10-Säulen (Fa. Amersham Biosciences) verwendet. Das Prinzip beruht auf einer Gelfiltration an Sephadex G-25. Das Säulenmaterial musste vor Benutzung mit dem Elutionspuffer äquilibriert werden. Es konnten pro Lauf maximal 2,5 ml Probe aufgetragen werden, anschließend wurde mit 3,5 ml eluiert. Die Handhabung erfolgte nach den Anweisungen des Herstellers.

2.14 Western Blotting (TOWBIN et al., 1979, modifiziert)

Die Methode des elektrophoretischen Proteintransfers auf Blot-Membranen (TOWBIN et al., 1979) wurde für den spezifischen Nachweis von Proteinen mit His₆-Tag bzw. V₅-Epitop angewandt. Der Western-Blot wurde in einer Mini-Blot-Apparatur („Mini Trans-Blot® Electrophoretic Transfer Cell“, Fa. Bio-Rad Laboratories GmbH, München) durchgeführt. Nach der SDS-PAGE wurde das Gel sowie auf Gelgröße zugeschnittene 2 Lagen Whatman Papier (Whatman 3 MM, Schleicher und Schuell GmbH, Dassel) und die Nitrocellulose-Membran (Hybond ECL Nitrocellulose Membrane; Fa. Amersham Biosciences, Freiburg) zusammen mit den beiden Glasfiberpads 10 min in gekühltem Transfer-Puffer getränkt. Anschließend wurde der Blot luftblasenfrei in der Gelhalterungskassette auf der Anodenseite aufgebaut und durch Andrücken der Kathodenseite auf die Anodenseite fest verschlossen.

Anode (+)

- Glasfiberpad
- 1 Lage Whatman
- Blot-Membran
- SDS-Polyacrylamidgel
- 1 Lage Whatman
- Glasfiberpad

Kathode (-)

Zusammen mit der Blot-Elektrodeneinheit wurde die Kassette in die mit Transfer-Puffer gefüllte Blot-Kammer eingesetzt. Der Transfer erfolgte bei einer Spannung von 120 V und 4 °C für 1 Stunde.

Transfer-Puffer:

Tris	25	mM
Glycin	192	mM
Methanol	200	ml
H ₂ O _{bidest.}	ad 1000	ml
pH 8,1 bis 8,4		

Der pH-Wert des Transfer-Puffers durfte nicht mit Säure oder Base eingestellt werden; lag der pH außerhalb dieses Bereichs, wurde der Puffer neu angesetzt. Bei der vor dem Western-Blot durchgeführten SDS-PAGE wurde ein so genannter „Prestained Marker“ (peqGOLD prestained Protein Marker; Fa. PeqLab, Erlangen, s. 2.13.5) verwendet. Anhand dieser bereits gefärbten Banden konnte der Proteintransfer überprüft werden. Dadurch entfiel ein Färbeschritt mit Ponceau S. Die Membran konnte direkt nach dem Western-Blot der gewünschten Färbung unterzogen werden. Als Alternative zur Benutzung von Nitrocellulose-Membranen wurden auch hydrophobe Polyvinylidendifluorid-Membranen (PVDF-Membranen, Fa. Amersham Biosciences, Freiburg) eingesetzt. Dabei musste darauf geachtet werden, dass diese zunächst in Methanol voräquilibriert wurden.

2.14.1 Immunodetektion von Proteinen auf Blot-Membranen

Geblottete Antigene bieten die Möglichkeit sie mit Konjugaten oder Antikörpern anfärben zu lassen. Bei der Verwendung von Antikörpern hängt die Spezifität der Immunfärbung von der Spezifität des primären und sekundären Antikörpers und deren eingesetzten Konzentrationen ab. In dieser Arbeit wurde die Immunodetektion eingesetzt, um carboxyterminale His₆- bzw. V5-Epitop

versehene Fusionsproteine nachzuweisen. Folgende primären- und sekundären Antikörper kamen dabei zum Einsatz.

Primär-Antikörper:

- Anti-His Antikörper (Fa. Amersham Biosciences, Freiburg); monoklonaler Antikörper aus Maus gerichtet gegen His₆-Polypeptid, Verdünnung: 1:3000 in TBS-T Puffer.
- Anti-V5-Ap Antikörper (Fa. Invitrogen, Carlsbad, USA); monoklonaler, mit alkalischer Phosphatase konjugierter Antikörper aus Maus gerichtet gegen das V5-Epitop aus *Paramyxovirus*, Verdünnung: 1:2000 in TBS-T Puffer.

Sekundär-Antikörper:

- Alkalische-Phosphatase-gekoppelter Anti Maus IgG + IgM (H+L) aus Ziege (Fa. Amersham Biosciences, Freiburg): Verdünnung: 1:10.000 in TBS-T Puffer.

Durchführung:

Alle Inkubationen erfolgten, wenn nicht anders beschrieben, bei Raumtemperatur auf einem Schüttler. Nach dem elektrophoretischen Transfer der Proteine von dem SDS- Polyacrylamidgel auf die NC-Membran wurde diese zur Absättigung unspezifischer Bindungsstellen 60 min in der Blocklösung inkubiert und anschließend dreimal für je 10 min in TBS-T gewaschen. Zum spezifischen Nachweis von Proteinen wurde die Membran für 1 Stunde mit dem Primärantikörper inkubiert (bei Anti-V5-Alkalische-Phosphatase Konjugat: 2 Stunden) und anschließend dreimal für je 15 min mit TBS-T gewaschen. Nach der darauf folgenden Inkubation mit dem Sekundärantikörper (entfällt für Anti-V5-Alkalische-Phosphatase Konjugat) für 1 Stunde wurde die NC-Membran erneut dreimal für je 15 min mit TBS-T gewaschen. Im Anschluss erfolgte die Detektion mit alkalischer Phosphatase.

Detektion:

Für die Immunodetektion der mit alkalischer Phosphatase gekoppelten Antikörper wurde das BCIP/NBT Color Development Substrat (Fa. Promega, Mannheim) eingesetzt. Die Substratfärbung wurde nach Angaben des Herstellers durchgeführt.

TBS Puffer (10x):

Tris-HCl	0,1	M
NaCl	1,4	M
pH 7,4		

TBS-T Puffer (10x):

Tris-HCl	0,1	M
NaCl	1,4	M
Tween-20	0,1	% (w/v)
pH 7,4		

Blocklösung:

Skim-Milk	5 % (w/v) in TBS-T Puffer (1x)
-----------	--------------------------------

2.15 Methoden zur Proteianreicherung bzw. Reinigung

2.15.1 Metall-Chelat-Affinitätschromatographie

Das Prinzip der Metall-Chelat-Affinitätschromatographie beruht auf der Fähigkeit bestimmter Aminosäuren z.B. Histidin als Elektronendonator, an Proteinoberflächen zu fungieren und so an Metallionen wie Ni^{2+} und Cu^{2+} zu binden (HOCHULI et al., 1988; YIP et al., 1989). Die Metallionen wiederum sind über Chelatgruppen an ein Trägermaterial (Agarose oder Sepharose) gebunden.

Die Metall-Chelat-Chromatographie besitzt im neutralen pH-Bereich ihre größte Spezifität und Effizienz. Die Elution der Proteine findet unter Verwendung eines Imidazolgradienten statt. Imidazol führt zu einer kompetitiven Verdrängung des Histidins. Alternativ kann eine Elution durch einen pH-Gradienten erfolgen.

2.15.1.1 Vorbereitung der Säule

Für die Metall-Chelat-Chromatographie wurde Ni^{2+} -Nitrilotriessigsäure-Agarose (NTA-Agarose; Fa. Qiagen GmbH) verwendet. Nach dem Einsetzen der unteren Fritte wurden 2 - 4 ml Ni^{2+} -NTA-Agarose in eine PD 10-Säule (Fa. Amersham Biosciences, Freiburg) eingefüllt. Nach der Sedimentation des Trägermaterials wurde die Säule mit dem entsprechenden Puffer gespült und bei 4 °C aufbewahrt.

2.15.1.2 Reinigung löslicher Proteine unter nativen Bedingungen

Zur Reinigung löslicher Proteine unter nativen Bedingungen wurde der konzentrierte Kulturüberstand mit der Ni^{2+} -NTA-Agarose, die mit Puffer I äquilibriert worden war, unter Rühren bei 4 °C für 1 h inkubiert. Nach dieser Inkubation wurde die Lösung in den Säulenkörper einer PD-10 Säule gefüllt. Nach dem Absetzen der Agarose wurde diese zunächst mit einem geeigneten Volumen an Puffer I gewaschen. Es folgte ein weiterer Waschschrift mit Puffer II. Die Elution der Proteine erfolgte in mehreren Schritten mit Puffer III. Vor einer erneuten Verwendung der Säule wurde das Chromatographiematerial mit Puffer III gespült und dann mit Puffer I äquilibriert.

Puffer I

KH_2PO_4	50	mM
KCl	300	mM
Imidazol	10	mM
pH 8,0		

Puffer II

KH_2PO_4	50	mM
KCl	300	mM
Imidazol	20	mM
pH 8,0		

Puffer III

KH_2PO_4	50	mM
KCl	300	mM
Imidazol	250	mM
pH 8,0		

2.15.1.3 Regeneration der Säule

Die Notwendigkeit einer Regeneration der Ni^{2+} -NTA-Agarose wurde durch eine Farbänderung von grün-blau nach bräunlich-grau angezeigt. Durch das chelatierende Agens EDTA erfolgte die Elution des Nickels von der Säule. Nach einer anschließenden Wiederbeladung mit Ni^{2+} -Ionen stand das Säulenmaterial erneut für mehrere Reinigungszyklen zur Verfügung. Die Regeneration des Säulenmaterials erfolgte anhand des Herstellerprotokolls.

2.15.2 Hydrophobe Interaktionschromatographie (HIC)

Die hydrophobe Interaktionschromatographie wurde mit einer ÄKTA-FPLC-Anlage bestehend aus folgenden Komponenten (alle Fa. Amersham Biosciences, Freiburg) durchgeführt.

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 Probenschleife

Säule:

- HIC Phenyl Sepharose HP XK 16/10

Das Trennprinzip der HIC basiert auf der unterschiedlichen Oberflächenhydrophobizität der einzelnen Proteine. Durch Auftragung 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 eine 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 Fertigsäule 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 0,5 MPa Rückdruck.

Allgemeine Regeln beim Umgang mit der Äkta-FPLC-Anlage:

- Proben vor dem Auftragen auf eine Säule immer durch Filtration (0,2 µm bzw. 0,45 µm Porenweite) oder durch hochtourige Zentrifugation partikelfrei machen.
- sämtliche Laufmittel und Puffer vor Gebrauch ebenfalls filtrieren.
- Säulen in 20 % (v/v) Ethanol lagern; direkte Umpufferung der Säulen von 20 % (v/v) Ethanol zu einem salzhaltigen Eluenten oder umgekehrt vermeiden, dazwischen mit H₂O_{dest.} spülen.
- alle Anschlüsse „drop to drop“ verbinden, um das Eindringen von Luft in das System zu vermeiden.

Durchführung:

Zur Auftragung der Probe wurde die Säule mit mindestens 2 Säulenvolumen auf Hochsalzbedingungen äquilibriert. Die Probe musste sich ebenfalls in einem Milieu hoher Ionenstärke befinden. Hierzu wurde sie vorsichtig auf 1 M NH₄SO₄ hochgesalzen. Die Elution erfolgte im absteigenden Salzgradienten, wobei evtl. mit H₂O_{dest.} nacheluiert werden musste, um stark hydrophobe Proteine wieder von der Säule zu lösen.

Die Parameter der Säulenläufe werden nachfolgend aufgelistet:

- Säulenvolumen: 20 ml
- Puffer A: 50 mM Tris-HCl, pH 8,0
- Puffer B: 50 mM Tris-HCl, pH 8,0 + 1 M NH_4SO_4
- Äquilibration: 5 Säulenvolumen Puffer B
- Gradient: 100 %- 0 % Puffer B (10 Säulenvolumen)
- Probevolumen: 35 ml bis 40 ml
- Fraktionen: je 6 ml
- Flussrate: 1,5 ml/min

2.15.3 Reinigung von Proteinen mit dem HiTrap IEX Selection Kit

Der Hi Trap IEX Selection Kit (Fa. Amersham Biosciences) beinhaltet 7 verschiedene Ionenaustauscher-Säulen (Füllvolumen je 1 ml). Das Kit wurde eingesetzt, um zu überprüfen ob eine Bindung des Zielproteins an Anionenaustauscher bzw. Kationenaustauscher stattfindet. Die Auftragung der Proteinlösung erfolgte über eine Peristaltische Pumpe mit einer Flussrate von 1ml/min. Alle Schritte wurden nach den Angaben des Herstellers durchgeführt.

2.16 Bestimmung von Enzymaktivitäten

2.16.1 Bestimmung der proteolytischen Aktivität

Die proteolytische Aktivität wurde colorimetrisch mit Casein (Fa. Sigma Chemie GmbH, Deisenhofen) als Substrat bestimmt. Die Messungen wurden in 2 ml Eppendorf Reaktionsgefäßen durchgeführt. Jede einzelne Enzymmessung hatte ihren eigenen Referenzwert. Folgendes Pipettierschema lag dem Enzymtest zugrunde.

	<u>Enzymmessung</u>	<u>Referenzwert</u>
Reagenz B: (Casein-Lsg.)	0,5 ml	0,5 ml
1-2 min bei 55 °C im Heizblock vorinkubieren, dann fortfahren:		
Enzymlösung:	0,1 ml	-
1 Stunde bei 55 °C im Heizblock inkubieren, dann fortfahren:		
Reagenz C: (TCA-Lsg.)	0,5 ml	0,5 ml
Enzymlösung:	-	0,1 ml
30 min bei 55 °C im Heizblock inkubieren		

Danach erfolgte ein Zentrifugationsschritt für 3 min, in einer Eppendorf-Tischzentrifuge. Dadurch wurde das durch die TCA-Lösung (Reagenz C) gefällte, nicht hydrolysierte Casein pelletiert. Der klare Überstand wurde für die Farbreaktion eingesetzt.

Durchführung der Farbreaktion:

Je 400 µl des klaren Überstandes von der Enzymmessung bzw. dem Referenzwert wurden mit 1 ml Reagenz E und 200 µl Folin & Ciocalteu's Phenol-Reagenz (F-C; Fa. Sigma Chemie GmbH, Deisenhofen) versetzt. Nach einer 30-minütigen Inkubation bei 37 °C und einer erneuten Zentrifugation (3 min, 13.000 rpm, RT) in einer Eppendorf-Tischzentrifuge, konnte die Extinktion des Überstandes bei 660 nm gegen die jeweiligen Referenzwerte ermittelt und die Aktivität anhand einer Eichkurve berechnet werden.

Erstellen einer Eichkurve:

Eine Eichkurve wurde mit einer 1,1 mM L-Tyrosin-Lösung nach folgendem Pipettierschema in einer Dreifachbestimmung erstellt:

	Std 0	Std 1	Std 2	Std 3	Std 4
Tyrosin-Lsg	-	10 µl	20 µl	40 µl	80 µl
H₂O_{dest.}	400 µl	390 µl	380 µl	360 µl	320 µl
Reagenz E	1 ml	1 ml	1 ml	1 ml	1 ml
F-C	200 µl	200 µl	200 µl	200 µl	200 µl

Nach einer 30-minütigen Inkubation bei 37 °C und einer Zentrifugation (3 min, 13.000 rpm, RT) in einer Eppendorf-Tischzentrifuge, konnten die Extinktionen des Überstandes bei 660 nm gemessen werden.

$$\Delta OD_{660} \text{ Standard} = OD_{660} \text{ Standard (1-4)} - OD_{660} \text{ Standard 0}$$

Zur Erstellung der Eichkurve wurden die ΔOD_{660} Standard gegen die µmol Tyrosin aufgetragen und die Steigung der Ausgleichsgeraden ermittelt.

Berechnung der proteolytischen Aktivität:

$$\Delta OD_{660} \text{ Probe} = OD_{660} \text{ Enzymmessung} - OD_{660} \text{ Referenzwert}$$

Die Ermittlung der aus der ΔOD_{660} Probe resultierenden µmol Tyrosin erfolgt über die Eichkurve.

$$U/ml = \frac{(\mu\text{mol Tyrosin}) \times (V_A)}{(V_E) \times (T) \times (V_F)}$$

V_A: Gesamtvolumen Enzym-Assay in ml

V_E: eingesetzte Enzymmenge in ml

T: Zeit des Enzym-Assays in min

V_F: eingesetztes Volumen in der Farbreaktion in ml

$$\text{U/mg} = \frac{\text{U/ml}}{\text{Proteinmenge in mg/ml}}$$

U (Unit)-Definition:

1 **U** ist definiert, als die Enzymmenge, die bei der Hydrolyse von Casein in 1 Minute 1 µmol Tyrosin freisetzt.

Reagenz A:

50 mM Kaliumphosphat-Puffer, pH 7,5

Reagenz B:

Caseinlösung (0,65 % w/v in Reagenz A)

Reagenz C:

110 mM Trichloressigsäure

Reagenz E:

500 mM Natriumcarbomat-Lösung

Standard-Lösung:

1,1 mM L-Tyrosin-Lösung

2.17 Biochemische Charakterisierung der Proteasen

2.17.1 Bestimmung der Substratspezifität

Zur Erstellung eines Substratspektrums, wurde die proteolytische Aktivität an unterschiedlichen Substraten untersucht. Die getesteten Substrate, das Testprinzip sowie die Vorgehensweise werden nachfolgend beschrieben.

Nachweis der proteolytischen Aktivität mit Skim Milk als Substrat:

Untersuchungen zur Hydrolyse von Skim-Milk wurde mit einem auf Aktivität basierendem Plattentest auf LB-Skim-Milk-Agar qualitativ durchgeführt. Dieses Testsystem wurde in der Arbeit auch schon als Screening-Methode angewandt und ist unter 2.10.4 beschrieben.

Nachweis der proteolytischen Aktivität mit Azoalbumin und Azocasein als Substrat:

Qualitative Untersuchungen zur Hydrolyse von Azoalbumin erfolgten auch im Aktivitäts-Plattentest. LB- bzw. M9-Mineralagar wurden nach dem Autoklavieren mit dem Substrat Azoalbumin (0,3 % w/v) und dem entsprechendem Antibiotikum komplementiert. Das Substrat wurde vorher in etwas sterilem H₂O_{dest.} gelöst, da es sich nicht sterilfiltrieren ließ. Der Testagar erhielt durch das Substrat Azoalbumin eine orange Farbe. Nach 2 bis 3 Tagen Inkubation (37 °C) der rekombinanten *E. coli*-Stämme auf den Testplatten, konnte die proteolytische Aktivität durch eine Hofbildung (Entfärbung: durch Hydrolyse des Substrates) um die Kolonie gezeigt werden.

Nachweis der elastinolytischen Aktivität mit Elastin Congo Rot als Substrat:

Da viele Proteasen befähigt sind Elastin zu hydrolysieren, erfolgten Untersuchungen auf elastinolytische Aktivität ebenfalls qualitativ mit einem Plattentest. Als Grundmedium wurde LB-Agar verwendet. Nach dem Autoklavieren erfolgten die Zugaben von Elastin Congo Rot (0,3 % w/v) und dem entsprechendem Antibiotikum. Elastin Congo Rot ist im Agar unlöslich und erscheint als „krümmelige“ Substanz innerhalb des Agars, daher ist beim Gießen der Platten, auf eine gleichmäßige Verteilung des Substrates zu achten. Durch elastinolytische Aktivität wird der an das Substrat gekoppelte rote Farbstoff abgespalten und geht in Lösung. Die Aktivität wird somit durch eine rote Hofbildung um die Kolonie angezeigt. Die Platten wurden bis zu einer Woche bei 37 °C inkubiert.

Nachweis der proteolytischen Aktivität mit Gelatine, BSA und Casein als**Substrat:**

Der qualitative proteolytische Aktivitätsnachweis für die Substrate Gelatine, BSA und Casein erfolgte über die Zymographie. Dabei handelt es sich um eine elektrophoretische Methode. Bei dieser Technik wird die Proteaseaktivität in einem mit dem Substrat copolymerisiertem SDS-Polyacrylamidgel detektiert. An der Stelle, wo das Substrat hydrolysiert wird, erscheint nach Anfärbung mit Coomassie Brilliantblau eine durchsichtige Bande.

Die Herstellung der Gele erfolgte, wie unter 2.13.3 beschrieben. Der einzige Unterschied war, dass die Hälfte des Volumens an $\text{H}_2\text{O}_{\text{dest}}$ im Trenngel, durch die entsprechende Substratlösung (0,65 %, w/v) ersetzt wurde. Zum Auftragen der Proben in die Geltaschen wurden sie mit Zymogram-Probenpuffer versetzt (Verhältnis Probe zu Puffer betrug 3:1). Die Elektrophorese erfolgte bei 120 V (s. 2.13.3). Zur Entfernung von SDS wurden die Gele nach der Elektrophorese für 30 min in Triton X-100-Lösung (2,5 % v/v) gewaschen. Nach einer Inkubation für 2 bis 8 h in KP-Puffer (5 mM, komplementiert mit 0,5 mM ZnCl_2) konnten die Gele mit Coomassie-Färbelösung (s. 2.13.4.1) für 30 bis 60 min angefärbt werden. Im Anschluss erfolgte die Entfärbung (s. 2.13.4.1), bis die Aktivitätsbande deutlich sichtbar war.

Zymogram-Probenpuffer:

Tris	62,5	mM
SDS	4 %	(w/v)
Glycerin	25 %	(v/v)
Bromphenolblau	0,01 %	(w/v)
pH 6,8		

2.17.2 Bestimmung der proteolytischen Aktivität in Abhängigkeit von der Temperatur

Zur Ermittlung der proteolytischen Aktivität in Abhängigkeit von der Temperatur, wurde die Enzymlösung bei unterschiedlichen Temperaturen (von 15 °C – 85 °C) für 1 h bei der jeweiligen Temperatur im Heizblock inkubiert. Im Anschluss erfolgte die Bestimmung der proteolytischen Aktivität wie unter 2.16.1 beschrieben.

2.17.3 Bestimmung der proteolytischen Aktivität in Abhängigkeit vom pH-Wert

Für die Bestimmung der proteolytischen Aktivität in Abhängigkeit vom pH-Wert, wurde das Casein in Puffern mit unterschiedlichen pH-Werten gelöst. Es wurden zwei Puffersysteme gewählt:

- pH 6 bis 8: Kaliumphosphat-Puffer (50 mM)
- pH 9 bis 11,5: Natriumcarbonat-Puffer (50 mM)

Die Messung der Protease-Aktivität erfolgte wie unter 2.16.1 beschrieben.

2.17.4 Bestimmung der Thermostabilität

Zur Ermittlung der Thermostabilität der Proteasen wurde die Enzymlösung für eine Dauer von 3 h bei drei unterschiedlichen Temperaturen (50 °C, 60 °C und 70 °C) im Heizblock inkubiert. Zu verschiedenen Zeitpunkten innerhalb der drei Stunden wurden Aliquots entnommen und bei 4 °C aufbewahrt. Im Anschluss wurde die Ausgangsaktivität, die proteolytische Restaktivität der entnommenen Aliquots und die Restaktivität der 3-Stunden-Probe bestimmt (s. 2.16.1).

2.17.5 Einfluss von Inhibitoren auf die proteolytische Aktivität

Um den Einfluss von Inhibitoren auf die Enzymaktivität zu untersuchen, erfolgte eine Inkubation (1 h, RT) der Enzymlösung mit unterschiedlichen Inhibitoren. Folgende Protease-Inhibitoren wurden eingesetzt:

- EDTA (10 mM)
- 1,10 Phenantrolin (10 mM)
- PMSF (10 mM)

Im Anschluss erfolgte die Bestimmung der proteolytischen Restaktivität, wie unter 2.16.1 beschrieben.

2.18 Chemikalien, Enzyme und Gase

Es wurden Chemikalien des handelsüblichen Reinheitsgrades der Firmen Boehringer (Mannheim), Merck (Darmstadt), Riedel de Haen (Hannover), Serva (Heidelberg) und Sigma Aldrich Chemie (Deisenhofen) verwendet. Des Weiteren wurde mit Biochemikalien, Enzymen etc. folgender Firmen gearbeitet:

Amersham Biosciences, Freiburg:

Nitocellulose-Membranen, PVDF-Membranen, PD10-Säulen, HiTrap IEX Selection Kit, Anti-His Antikörper, Alkalische-Phosphatase-gekoppelter Anti Maus IgG + IgM

Biozym GmbH, Hess. Oldendorf

Elektroporationsküvetten

Eurogentec, Seraing, Belgien:

Agarose

Gerbu Biotechnik GmbH, Gaiberg:

IPTG, X-Gal, Ampicillin, Kanamycin, Proteinase K

Invitrogen, Carlsbad, CA, USA:

pET Directional TOPO Expression Kit, TOPO TA Cloning Kit, Anti-V5-Ap Antikörper

Macherey und Nagel, Düren:

NucleoSpin Plasmid-Kit

MBI Fermentas GmbH, St. Leon-Rot:

Taq-DNA-Polymerase, *Pfu* DNA-Polymerase T4-DNA-Ligase, T4-DNA-Polymerase, Ladder-Mix, Nukleotide, Restriktionsendonukleasen, Protein Molecular Weight Marker

Messer Griesheim GmbH, Kassel:

Gase (N₂, H₂, N₂/CO₂)

Metabion, Martiensried:

Oligonukleotide

Millipore, Eschborn:

Mikrodialyse-Filter

PEQLAB Biotechnologie:

peqGOLD Prestained Protein Marker

Promega, Deutschland GmbH, Mannheim:

WizardTM *Plus* Minipreps DNA Purification System, BCIP/NBT Color Development Substrat

Qiagen, Hilden:

Plasmid Midi-Kit, QIAprep Spin Miniprep Kit, Qiaquick PCR-Purification-Kit; Qiaquick Gel Extraction Kit

Sartorius AG, Göttingen:

Sterilfilter, Vivaspin Konzentratoren

Whatman, Maidstone, England:

Whatman-Filterpapier 3MM

3 Experimente und Ergebnisse

Die Nachfrage nach biotechnologisch relevanten Biokatalysatoren und Wirkstoffen steigt ständig an. In der Vergangenheit wurden natürliche Biokatalysatoren hauptsächlich durch traditionelle Methoden identifiziert. Zunächst erfolgte ein groß angelegtes Screening nach Mikroorganismen mit einer erwünschten Enzymaktivität. Hierzu erfolgte eine Isolierung und Anreicherung des Mikroorganismus, und anschließend wurde das gewünschte Zielgen isoliert. Ein Nachteil dieses klassischen Verfahrens besteht jedoch darin, dass die Kultivierbarkeit der Mikroorganismen hierfür eine Voraussetzung ist. Da aber schätzungsweise weniger als 1 % der auf der Erde befindlichen Mikroorganismen mit Standardtechniken kultivierbar sind (AMMAN et al., 1995), kann ein Großteil der natürlich vorkommenden Biodiversität nicht genutzt werden.

Mit der Metagenomforschung ist man in der Lage dieses Problem zu umgehen. Unabhängig von der Kultivierbarkeit der Mikroorganismen wird die gesamte genomische DNA aus Umweltproben direkt oder nach vorheriger Anreicherung der Mikroorganismen in komplexen Anreicherungskulturen isoliert und zur Erstellung von Umweltgenbanken (Metagenombanken) verwendet. Die Genbanken können dann in geeigneten Screening-Verfahren nach gewünschten Enzymaktivitäten durchmustert werden. Mit Hilfe der Metagenomik konnten bisher sowohl biotechnologisch interessante Enzyme wie z.B. Amylasen (RONDON et al., 2000; YUN et al., 2004), Glycerin/Diol-Dehydratasen (KNIETSCH et al. 2003c), Oxidoreduktasen (KNIETSCH et al., 2003a und 2003b), Lipasen/Esterasen (HENNE et al., 2000) und Cellulasen (HEALY et al. 1995; VOGET et al., 2005) als auch Gene zur Synthese von Biotin (ENTCHEVA et al. 2001) und Antibiotika (GILLESPIE et al., 2002) entdeckt werden.

Im Rahmen dieser Arbeit sollten verschiedene Metagenombanken erstellt werden (s. Abb. 6). Weiterhin sollten die selbst konstruierten Genbanken und auch schon vorhandene Metagenombanken auf Gene, die Butanol-

Dehydrogenase-Aktivität oder proteolytische Aktivität vermitteln, mit auf Enzymaktivität basierenden Screeningverfahren durchmustert werden. Im Anschluss wurde versucht Zielgene zu identifizieren und die daraus resultierenden Genprodukte zu charakterisieren.

Abbildung 6: Schema der im Rahmen der Arbeit durchgeführten Schritte.

3.1 Konstruktion von Metagenombanken in *Escherichia coli*

Zur Konstruktion einer Metagenombank kann man zwei Ansätze verfolgen. Einerseits kann die dafür benötigte Umwelt-DNA direkt, ohne vorherige Kultivierung oder Anreicherung von Mikroorganismen aus einer Standortprobe isoliert werden. Damit bleibt die Biodiversität der Umweltprobe erhalten. Andererseits kann vorher ein Anreicherungsschritt z.B. in komplexen Anreicherungsmedien mit selektiver C- oder N-Quelle mit der Standortprobe als Inokulum vorgenommen werden. Durch einen gezielten Anreicherungsschritt ist es möglich, die Trefferquote an positiven Klonen in dem später durchgeführten Screening-Verfahren zu erhöhen (KNIETSCH et al., 2003a und 2003b; ENTCHEVA et al., 2001; GABOR et al., 2004). Allerdings bedeutet ein vorheriger Anreicherungsschritt eine Reduzierung der Diversität der Probe, da durch die gezielte Anreicherung einige Mikroorganismen schneller wachsen können und damit andere verdrängen.

3.1.1 Isolierung und Reinigung von DNA aus Umwelt-Proben

Zur Herstellung von Metagenombanken wurden Sediment- und Bodenproben folgender Umwelt-Standorte verwendet:

- Flusssediment des Flusses Leine (bei Göttingen)
- Wattsediment aus der Nordsee (in der Nähe von Niebüll)
- Sediment aus dem Golf von Eilat (Israel)
- Sediment aus einem Salzbecken (Israel)
- Bodenprobe aus einer Erzhalde (Harz)
- Wiesenboden (bei Göttingen)
- Zuckerrübenboden (bei Göttingen)
- Wüstensand (Ägypten)
- Klärschlamm (Faulturm, Kläranlage Göttingen)
- Komposterde (bei Leverkusen)

Die mikrobielle DNA wurde entweder direkt aus den Proben isoliert, oder es erfolgte ein vorheriger Anreicherungsschritt in Mineralmedium mit Zimtsäure (2,5 mM) bzw. 2-Penten (2 mM) als alleiniger C-Quelle und Energiequelle unter anaeroben oder aeroben Bedingungen (s. 2.2.3). Als Inokulum wurden bei den Anreicherungskulturen (500 ml) 10 bis 15 g eines Standortprobengemisches aus mehreren Umweltproben eingesetzt. Nach 2 bis 5 Tagen Inkubation wurden die Kulturen durch Zentrifugation geerntet. In dem daraus resultierenden Pellet befanden sich sowohl die angereicherten Mikroorganismen als auch die Matrix-Bestandteile der verwendeten Umweltproben. Die Pellets dienten als Ausgangsmaterial zur Isolierung von mikrobieller DNA.

Im Rahmen dieser Arbeit wurde die DNA nach der Methode von Zhou und seinen Mitarbeitern isoliert (s. 2.6.1). Bei dieser Vorgehensweise werden die Mikroorganismen ohne vorherige Abtrennung von Umweltpartikeln einer direkten Lyse unterzogen. Somit ist gewährleistet, dass auch die frei in der Probe vorliegende DNA mit erfasst wird (ZHOU et al., 1996; SANTOSA, 2001). Dabei gelang der Zellaufschluss durch eine anhaltende Hitzeinkubation in einem Hochsalzextraktionspufferpuffer mit Hilfe von Detergenzien (SDS, CTAB), EDTA und Proteinase K. Es folgte eine Abtrennung aller Zell- und Umweltproben-Partikel durch Zentrifugation. Nach durchgeführter Chloroform/Isoamylalkohol-Extraktion konnte die DNA mit Isopropanol gefällt werden. Da oftmals bei der Isopropanolfällung auch sehr viel Salz mit auskristallisierte, erfolgte im Anschluss eine Dialyse der DNA über Nacht gegen TE-Puffer (s. 2.5.2.4). Danach wurde die DNA erneut gefällt (s. 2.5.2.1).

Ein Nachteil dieser Methode waren die häufig auftretenden Koextraktionen von Matrixbestandteilen wie z.B. Huminsäuren, die der isolierten DNA eine bräunliche Farbe verliehen. Eine Entfernung von Huminsäuren war notwendig, da sie Restriktionsenzyme, Ligasen und Polymerasen inhibieren (TEBBE und VAHJEN, 1993) und diese Enzyme für die nachfolgende Klonierung der DNA eingesetzt wurden. Es erfolgte eine Reinigung der Umwelt-DNA über das „Wizard DNA Purification System“ und falls notwendig, zusätzlich eine weitere über Anionenaustauschchromatographie an „Qiagen-Tip-100“-Säulen (s. 2.6.2).

Die auf diesem Wege erhaltene hochmolekulare Umwelt-DNA ist in Abbildung 7 dargestellt und konnte für weitere Modifikationen im Zuge der Konstruktion einer Metagenombank eingesetzt werden.

Abbildung 7: Darstellung isolierter und gereinigter Umwelt-DNA im Agarosegel.

Die Umwelt-DNA wurde, wie unter 3.1.1 erläutert, isoliert und gereinigt. Aufgetragen wurden jeweils 2 µl. M: Längenstandard; 1: Isolierte chromosomale Umwelt-DNA vor Reinigung; 2: Isolierte chromosomale Umwelt-DNA nach Reinigung über „Wizard DNA Purification System“ und „Qiagen Tip-100“-Säule.

3.1.2 Enzymatische Modifikationen der isolierten Umwelt-DNA

Um später Umwelt-DNA-Fragmente von geeigneter Größe klonieren zu können, erfolgte eine Größenfragmentierung der gereinigten DNA. Dazu wurde die DNA einem partiellen zeitlimitierten Restriktionsverdau mit dem Restriktionsenzym *Bsp143I* (einem Isoschizomer von *Sau3AI*) unterzogen (s. 2.8.1). Dieses Restriktionsenzym besitzt den Vorteil, dass es nur eine 4 bp-lange Erkennungssequenz hat (5'...GATC...3') und demnach häufig schneidet. Das Ziel ist es, im Idealfall eine gleichmäßige Verteilung aller vorkommenden DNA-Fragmente zu erhalten. Der partielle Restriktionsverdau wurde dabei über einen Zeitraum von 10 Minuten durchgeführt. Zu unterschiedlichen Zeitpunkten wurden Aliquots entnommen und die Enzym-Reaktion mit EDTA auf Eis abgestoppt. Im Anschluss erfolgte die Überprüfung der partiellen

Größenfragmentierung durch Agarosegelelektrophorese (s. 2.7.1). Um eine Klonierung von sehr kleinen DNA-Fragmenten zu verhindern, erfolgte eine Trennung der partiell geschnittenen DNA nach Fragmentgröße in einem Saccharose-Dichtegradienten durch Zentrifugation (s. 2.9). In Abbildung 8 sind der typische Verlauf eines partiellen Restriktionsverdaus und einer Größenfraktionierung durch Saccharose-Dichtegradienten-Zentrifugation dargestellt.

Abbildung 8: Elektrophoretische Trennung partiell verdauter und durch Saccharose-Dichtegradienten-Zentrifugation größenfraktionierter Umwelt-DNA.

A) Partieller Verdau der Umwelt-DNA mit *Bsp143I*. M: Längenstandard; Spur 1: Umwelt-DNA 15 s verdaut; Spur 2: Umwelt-DNA 30 s verdaut; Spur 3: Umwelt-DNA 1 min verdaut; Spur 4: Umwelt-DNA 1,5 min verdaut; Spur 5: Umwelt-DNA 2 min verdaut; Spur 6: Umwelt-DNA 5 min verdaut; Spur 7: Umwelt-DNA 10 min verdaut. Gut zu erkennen ist die Abnahme an hochmolekularer, genomischer Umwelt-DNA (durch einen roten Pfeil markiert).

B) Saccharose-Dichtegradienten-Zentrifugation. Spuren 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21 und 23: Zeigen je 3 µl jeder zweiten erhaltenen DNA-Fraktion aus dem Gradienten. M: Längenstandard.

3.1.3 Klonierung der Umwelt-DNA in Vektoren und Transformation in *Escherichia coli*

Die Auswahl eines geeigneten Vektorsystems zur Konstruktion von Metagenombanken hängt von unterschiedlichen Faktoren ab. Sie wird bestimmt von der Qualität der isolierten DNA, der gewünschten Insertgröße, von dem Wirtsorganismus, in dem die Genbank hergestellt werden soll, und von der erwünschten Kopienzahl der rekombinanten Plasmide. Letzteres ist entscheidend für das spätere Screening nach Enzymaktivitäten (DANIEL, 2002 und 2005). Im Rahmen der Arbeit wurden ausschließlich Plasmidgenbanken in dem Wirt *Escherichia coli* konstruiert. Als geeignetes Vektorsystem wurden die pUC19-Derivate pBluescript® SK(+) bzw. pBluescript® SKII(+) ausgewählt. Der Vorteil von Plasmidgenbanken gegenüber z.B. Cosmidgenbanken besteht darin, dass die höhere Kopienzahl der Vektoren eine Detektion eines nur schwach exprimierten Fremdgenes ermöglicht. Weiterhin ist auch eine Klonierung von noch geringfügig verunreinigter DNA möglich. Besonders hervorzuheben sind allerdings die einfachen technischen Verfahren der Klonierung. Nachteile von Plasmidgenbanken gegenüber Cosmidgenbanken sind dagegen, dass man bei einem Screening eine größere Anzahl an Klonen durchmustern muss, um potentielle Kandidaten zu finden (DANIEL, 2005).

Es wurde angestrebt Plasmidgenbanken mit einer Insertgröße zwischen 2 bis 8 Kbp zu erstellen. Dazu wurden geeignete DNA-Fractionen aus der Saccharose-Dichtegradienten-Zentrifugation ausgewählt und mit Hilfe der T4-Ligase in die entsprechenden Vektoren kloniert (s. 2.8.3.1). Die Vektoren wurden vorher mit dem Restriktionsenzym *Bam*HI (Erkennungssequenz: 5'...G ↓GATC↑C...3';) linearisiert (s.2.8.1), damit eine Klonierung der mit *Bsp*143I verdauten Umwelt-DNA gewährleistet war. Weiterhin erfolgte eine Behandlung mit alkalischer Phosphatase (s. 2.8.2), um eine hohe Religationsrate des Vektors auszuschließen. Nach der Klonierung erfolgte die Transformation der rekombinanten Plasmide via Hitzeschock in *E. coli* DH5α-Zellen oder *E. coli* Top10-Zellen (s. 2.10.1). Eine Selektion der Transformanten erfolgte mit Hilfe der Antibiotikaressistenz des Vektors und des X-Gal-Tests (s. 2.10.3).

3.1.4 Entwicklung einer Methode zur Klonierung von Umwelt-DNA mit Hilfe des TOPO-T/A-Systems

Die Klonierung der Umwelt-DNA einiger Standorte (z.B. Komposterde und Bodenprobe Erzabbau) gestaltete sich äußerst schwierig und die Anzahl an erhaltenen rekombinanten Transformanten pro Ligationsansatz war sehr gering (oft weniger als 500 rekombinante Klone pro Ligationsansatz). Dieses Phänomen könnte damit begründet werden, dass es trotz intensiver Reinigungsversuche der Roh-DNA aus diesen Standorten nicht möglich war, die T4-Ligase inhibierenden Matrixsubstanzen, vollständig zu entfernen. Deswegen wurde im Rahmen dieser Arbeit eine alternative Klonierungsmethode von Umwelt-DNA entwickelt, bei welcher der Ligationsschritt mit der T4-Ligase umgangen wurde.

Die neu ausgearbeitete Methode basiert auf der TOPO-T/A-Klonierungsmethode der Firma Invitrogen (s. 2.8.3.2). Normalerweise wurde diese Technik dafür entwickelt, um mit *Taq*-Polymerase amplifizierte PCR-Produkte direkt ohne Zusatz von Ligase in den mitgelieferten Vektor zu klonieren. Die *Taq*-DNA-Polymerase zeichnet sich dadurch aus, dass sie unabhängig von der eingesetzten Template-DNA eine terminale Transferase-Aktivität besitzt und somit an die 3'-Enden von PCR-Produkten einzelne Desoxyadenosine (A) anhängt. Der vom Hersteller mitgelieferte Klonierungsvektor pCR2.1-TOPO liegt linearisiert vor. Er besitzt komplementär zu den Überhängen der mit *Taq*-DNA-Polymerase amplifizierten PCR-Produkte, an den 3'-Enden einzelne Desoxythymidine (T). Weiterhin ist an den Vektor kovalent die Topoisomerase I des *Vaccinia*-Virus gebunden. Diese bindet an eine spezifische doppelsträngige Sequenz im Vektor (5'...CCCTT...3') und spaltet dort das Phosphodiester-Gerüst stromabwärts dieser Sequenz in einem Strang (SHUMAN, 1991). Die Energie dieser Spaltungsreaktion wird durch Ausbildung einer kovalenten Bindung zwischen dem 3'-Phosphat des gespaltenen Stranges und dem Tyrosin₂₇₄ der Topoisomerase I konserviert. Durch einen Angriff der 5'-Hydroxylgruppe des PCR-Produktes auf die kovalente Phospho-Tyrosyl-Bindung zwischen Vektor und Enzym kann die Reaktion umgekehrt werden. Die Topoisomerase wird somit entfernt und das

PCR-Produkt in den Vektor kloniert (SHUMAN, 1994). In Abbildung 9 ist die TOPO-T/A-Klonierung schematisch dargestellt.

Abbildung 9: Schematische Darstellung der TOPO-T/A-Klonierungsmethode. (aus Handbuch für „TOPO TA Cloning®“ Version R, Fa. Invitrogen)

Um diese Methode für die Klonierung von Umwelt-DNA anwenden zu können, musste die gereinigte und partiell-verdaute Umwelt-DNA speziell aufgearbeitet werden. Die Modifikationsschritte umfassen dabei die drei nachfolgend beschriebenen Reaktionen:

1. Erzeugung von „glatten Enden“ mit Hilfe der T4-DNA-Polymerase:

Zunächst wurde die Umwelt-DNA mit Hilfe des Enzyms T4-DNA-Polymerase an ihrem 3'-Ende aufgefüllt. Dieser Schritt war unumgänglich, da der Verdau der Umwelt-DNA mit dem unsymmetrisch schneidenden Restriktionsenzym *Bsp143I* „klebrige Enden“ erzeugt hatte. Dazu wurden etwa 5 µg größenfraktionierte Umwelt-DNA mit 10 mM dNTP-Mix, 10 µl Reaktionspuffer und 5 U T4-DNA-Polymerase versetzt. Mit sterilem H₂O_{dest.} wurde der Reaktionsansatz auf ein Gesamtvolumen von 50 µl gebracht und für 1 h bei

Raumtemperatur inkubiert. Ein zusätzlicher Vorteil der Auffüllreaktion war, dass das Enzym auch „Proof reading“-Aktivität besitzt und somit geschädigte DNA-Abschnitte repariert werden konnte. Im Anschluss an die Inkubation wurde der Reaktionsansatz mit Hilfe des „QIAquickPCR-Purification-Kits“ der Firma Qiagen (s. 2.5.2.7) gereinigt und die DNA mit 50 µl sterilem H₂O_{dest.} von der Säule eluiert.

2. Erzeugung von 3`-adenylierter (3`-A)-DNA:

Das Anfügen von Desoxyadenosin an das zuvor geglättete 3`-Ende der Umwelt-DNA erfolgte mit dem Enzym *Taq*-DNA-Polymerase. Dazu wurden die 50 µl aufgefüllte und gereinigte DNA aus Schritt 1 mit 6 µl dATP (2 mM), 6 µl MgCl₂ (25 mM), 8 µl Reaktionspuffer und 2 U *Taq*-DNA-Polymerase versetzt und mit sterilem H₂O_{dest.} auf ein Gesamtvolumen von 80 µl gebracht. Der Reaktionsansatz wurde für 20 Minuten bei 72 °C inkubiert und danach analog zu Schritt 1 gereinigt.

3. Dephosphorylierung der DNA:

Der letzte Schritt der Modifizierung der Umwelt-DNA für die TOPO-T/A-Klonierung umfasste die Dephosphorylierung der DNA an ihrem 5`-Ende mit Hilfe der alkalischen Phosphatase (CIAP). Dazu wurde die aufgearbeitete DNA aus Schritt 2 mit 2 µl der CIAP, 10 µl Reaktionspuffer und sterilem H₂O_{dest.} auf ein Gesamtvolumen von 100 µl gebracht. Nach einer Inkubation für 30 Minuten bei 37 °C konnte die DNA über das „QIAquick PCR-Purification-Kit“ gereinigt werden. Diesmal erfolgte die Elution von der Säule nur mit 30 µl sterilem H₂O_{dest.} Die DNA konnte hiernach für die TOPO-T/A-Klonierung in den Vektor pCR®2.1-TOPO eingesetzt werden.

Alle Schritte der Klonierungsreaktion und der anschließenden Transformation der rekombinanten Plasmide in *E. coli* Top10-Zellen wurden nach Hersteller-Angaben durchgeführt. Eine Selektion der Transformanten erfolgte mit Hilfe der Antibiotikaresistenz des Vektors und des X-Gal-Testes (s. 2.10.3) auf LB-Ampicillin-Agar, der zusätzlich X-Gal und IPTG enthielt. Mit dieser Methode war es im Vergleich zur Klonierung mit der T4-Ligase zum einen möglich, die

Anzahl an rekombinanten Plasmiden pro Ligationsansatz um das 2 bis 5-fache zu steigern, und zum anderen den prozentualen Anteil an rekombinanten Plasmiden auf 6 % (Vergleich der Inserthäufigkeit zwischen Genbanken GB11 und GB7, s. Tab. 6) bis 19 % (Vergleich der Inserthäufigkeit zwischen Genbanken GB 1 und GB5, s. Tab, 6) zu erhöhen.

3.1.5 Konstruktion und Charakterisierung von Metagenombankenbanken

Wenn eine ausreichend große Anzahl an rekombinanten Klonen auf den Agar-Platten vorhanden war, wurden zur Konservierung der rekombinanten Plasmide alle Klone von den Platten abgeschwemmt und die Plasmide über eine Anionenaustauschchromatographie mit Hilfe des Qiagen Plasmid Midi-Kit isoliert (s. 2.6.3.4). Die Aufbewahrung der daraus resultierenden Plasmidgenbanken erfolgte bei -20 °C.

Zur Charakterisierung der Plasmidgenbanken wurden Parameter wie die durchschnittliche Insertgröße, die durchschnittliche Inserthäufigkeit und die Anzahl an rekombinanten Plasmiden bestimmt. Die Ermittlung der Insertgrößen erfolgte, indem 1 µl der Plasmid-DNA in *E. coli* Top10-Zellen retransformiert und 50 rekombinante Retransformanten in LB-Ampicillin-Medium angezogen wurden. Nach der Isolierung ihrer Plasmide (s. 2.6.3.1) konnte die Insertgröße via Restriktionsanalyse (s. 2.8.1) ermittelt werden. Der prozentuale Anteil an Plasmiden mit Insert (durchschnittliche Inserthäufigkeit) konnte über den X-Gal-Test (s. 2.10.3) bestimmt werden. Die Anzahl an rekombinanten Klonen erhielt man durch Auszählung der abgeschwemmten Klone von den Platten. In Tabelle 6 sind alle im Rahmen der Arbeit konstruierten Genbanken mit ihren Parametern aufgeführt.

Tabelle 6: Aufstellung aller konstruierten Umweltgenbanken und ihrer Charakteristika.

Standort, Behandlung	Genbank- Bezeichnung	Klonierungs- Vektor	Anzahl rekombinanter Plasmide	Ø Insert- größe	Insert- Häufigkeit	Gesamt- DNA [Mbp]
Standortprobengemisch: Wüstensand, Leinesediment, Faulturm, Sediment Golf von Eilat; Vorinkubation: 2-Penten, anaerob	GB1	pBluescriptSK II	52.000	4,2 kb	62 %	218
	GB5	pCR2.1-TOPO	78.000	5,3 kb	74 %	413
Standortprobengemisch: Salzbeckensediment, Bodenprobe, Sediment Golf von Eilat Wattsediment Vorinkubation: Zimtsäure, anaerob	GB11	pBluescriptSK I	46.000	3,1 kb	58 %	143
	GB2	pBluescriptSK II	55.000	3,9 kb	59 %	215
	GB7	pCR2.1-TOPO	65.000	4,0 kb	62 %	260
	GB6	pCR2.1-TOPO	65.000	2,8 kb	69 %	182
Standortprobengemisch: Faulturm, Bodenprobe, Wattsediment, Salzbeckensediment, Leinesediment; Vorinkubation: 2-Penten, aerob	GB3	pBluescriptSK II	70.000	5,4 kb	63 %	378
Wattsediment Vorinkubation: Zimtsäure, anaerob	GB9	pCR2.1-TOPO	28.000	2,6 kb	55 %	73
Bodenprobe: Erzabbaubereich Schacht Fortuna, Harz; Vorinkubation: Zimtsäure, anaerob	GB4	pBluescriptSK I	10.000	3,9 kb	71 %	39
	GB8	pCR2.1-TOPO	37.000	4,6 kb	76 %	170
Komposterde keine Vorinkubation	GB12	pBluescriptSK II	32.000	3,5 kb	68 %	76
	GB10	pCR2.1-TOPO	60.000	2,9 kb	72 %	210
Zusammenfassung aller 12 Genbanken:				598.000	66 %	2,2 Gbp

Insgesamt wurden 12 Metagenombanken in drei unterschiedlichen Klonierungsvektoren konstruiert, deren durchschnittlichen Insertgrößen zwischen 2,8 Kbp und 5,4 Kbp liegen. Die Anzahl an rekombinanten Plasmiden pro Genbank variiert von 10.000 bei Genbank GB4 bis 78.000 bei Genbank GB5. Die niedrigste Inserthäufigkeit mit 55 % weist die Genbank GB9 auf. Die höchste Inserthäufigkeit mit 76 % besitzt Genbank GB8. Fasst man die Werte aller Genbanken zusammen, konnten insgesamt 2,2 Gbp Umwelt-DNA kloniert werden. Deren genetische Information ist auf 598.000 rekombinanten Plasmiden mit einer durchschnittlichen Insertgröße von 3,9 kbp und einer Inserthäufigkeit von 66 % enthalten.

Diese konstruierten Metagenombanken sowie schon in der Arbeitsgruppe vorhandene Metagenombanken wurden als Ausgangsmaterial für die Suche von neuartigen Genen verwendet, die Butanol-Dehydrogenase- oder proteolytische Aktivität vermitteln.

3.2 Durchmusterung von Metagenombanken nach Genen, die Butanol-Dehydrogenase-Aktivität vermitteln

Ein Ziel der Industrie ist es, ein wirtschaftliches Fermentationsverfahren zur Butanol-Produktion zu etablieren als Alternative zur chemischen Produktion, da das industriell häufig verwendete Lösungsmittel Butanol heutzutage chemisch hergestellt wird, die chemische Produktion von den knapper werdenden Erdölvorkommen abhängig ist. Wie aus Abbildung 10 zu entnehmen, sind im Fermentationsprozess von *Clostridium acetobutylicum* Butanol-Dehydrogenasen im letzten Schritt der Fermentation direkt an der Herstellung von Butanol beteiligt. Eine Möglichkeit zur Optimierung der Butanol-Fermentation wäre, mit Hilfe der Metagenomik neuartige Enzyme, die Butanol-Dehydrogenase-Aktivität vermitteln, zu entdecken. Durch gezieltes „metabolic engineering“ könnte ein Produktionsstamm entwickelt werden, der sich in einem

späteren Produktionsprozess durch hohe Butanol-Ausbeuten auszeichnet und damit eine echte Alternative zur chemischen Produktion wäre.

Abbildung 10: Schematische Darstellung der fermentativen Butanol-Produktion durch *Clostridium acetobutylicum*.

Dargestellt ist nur skizzenhaft der Reaktionsweg bis zum Butanol mit den daran entscheidend beteiligten Enzymen. Alle anderen von *Clostridium acetobutylicum* gebildeten Fermentationsprodukte wurden vernachlässigt.

AdhE: Butyraldehyd-/Butanol-Dehydrogenase E; BdhA: Butanol-Dehydrogenase A; BdhB: Butanol-Dehydrogenase B.

3.2.1 Screening-Strategie

Für die Identifizierung neuartiger Butanol-Dehydrogenasen aus Metagenombanken wurde ein auf Enzymaktivität basierendes Screening-Verfahren angewendet. Dazu wurde die Fähigkeit der Butanol-Dehydrogenasen ausgenutzt, dass sie nicht nur an der Bildung von Butanol (Reduktionsrichtung) beteiligt sind, sondern dass sie auch die Rückreaktion vom Butanol zum Butyraldehyd (Oxidationsrichtung) katalysieren.

Das Screening erfolgte auf Aldehyd-Indikatorplatten nach einer modifizierten Methode von Conway (CONWAY et al., 1987). Dieses Nachweisverfahren konnte schon erfolgreich zur Auffindung neuartiger Alkohol-Oxidoreduktasen eingesetzt werden (KNIETSCH et al.; 2003). Der verwendete Indikatoragar (s. 2.2.4) enthielt als Grundbetsandteil Mineralmedium mit einem gewissen Anteil an Komplexbestandteilen (Pepton) und Ampicillin. Daher waren zunächst alle rekombinanten *E. coli*-Stämme zum Wachstum befähigt. Durch Zugabe des

Testsubstrates Butanol wurden jedoch die Klone im Wachstum bevorzugt, die Butanol als zusätzliche C-Quelle nutzen konnten. Die Umsetzung des Substrats Butanol zu Butyraldehyd, konnte durch eine Farbreaktion angezeigt werden. Der Indikatoragar enthielt zusätzlich die Bestandteile Natriumdisulfit und Pararosanilin. Diese beiden Substanzen bilden zusammen ein Schiff'sches Reagenz (ein primäres Amin), welches fast farblos ist. Kommt es nun in einer Additionsreaktion mit einer Carbonylverbindung (z.B. Butyraldehyd) zur Ausbildung einer Schiff'schen Base (Azomethin), wird dies durch eine intensive rote Farbe angezeigt. Eine Detektion von potentiellen positiven Klonen konnte, wie in Abbildung 11 dargestellt, durch die rote Koloniefarbe dieser Klone erfolgen. Eine hohe Umsetzung des Substrats führte zum Anfärben des Agars und zur Bildung von roten Höfen um die Kolonien. Eine Schwierigkeit des Screenings bestand zunächst in der Auswahl eines geeigneten Wirtsstammes, denn sowohl *E. coli* DH5 α als auch *E. coli* Top10 zeigten aufgrund ihrer wirtseigenen Alkohol-Dehydrogenasen eine starke Hintergrundaktivität auf den Indikatorplatten. Beide Stämme konnten das Testsubstrat Butanol umsetzen und wiesen demnach eine rote Koloniefarbe auf.

Abbildung 11: Phänotyp von positiven Klonen auf Aldehyd-Indikatorplatten.
Der Pfeil zeigt einen Butanol-verwertenden rekombinanten *E. coli*-Stamm.

Als geeigneter Wirtsstamm erwies sich die *E. coli*-Mutante WL3 (LOROWITZ und CLARK, 1982). Dieser Stamm ist dadurch charakterisiert, dass er eine Mutation in dem *adhE*-Gen trägt und somit keine wirtseigene AdhE Alkohol-

Dehydrogenase-Aktivität aufweist. Auch besitzt er keine Gene für spezifische Butanol-Dehydrogenasen. Der Stamm wurde auf den Indikatorplatten getestet und zeigte dort keine Hintergrundfärbung.

3.2.2 Durchführung und Ergebnisse des Screenings zur Identifizierung potentieller Butanol-Dehydrogenasen

Zur Identifizierung neuartiger Butanol-Dehydrogenasen wurden rekombinante Plasmide von vier der in Tabelle 6 aufgeführten Umweltgenbanken (GB1, GB2, GB8 und GB10) in *E. coli* WL3 transformiert (s. 2.10.1 und 2.10.2) und geeignete Verdünnungen auf den Aldehyd-Indikatorplatten mit 1-Butanol als Testsubstrat ausplattiert. Die Inkubation der Platten erfolgte bei 30 °C oder 37 °C für 16 bis 72 Stunden (s. 2.10.4). Positive Klone mit roter Koloniefarbe wurden auf eine neue Indikatorplatte ausgestrichen. Zeigten die Klone danach immer noch einen roten Phänotyp wurden sie in einer 5 ml Kultur angezogen und die Plasmide isoliert (s. 2.6.3.2 und 2.6.3.3). Es erfolgte eine Retransformation der isolierten Plasmide in die *E. coli*-Mutante und eine Überprüfung des Phänotyps auf Indikatorplatten. Tabelle 7 gibt eine Zusammenfassung der bei der Durchmusterung der Genbanken erhaltenen Ergebnisse.

Tabelle 7: Zusammenfassung der Ergebnisse des Screenings nach neuartigen Butanol-Dehydrogenasen.

Genbank	Anzahl getesteter Klone	Getestete Umwelt-DNA [Mbp]	Positive Klone nach initialen Screening	Klone mit stabilem Phänotyp	Bezeichnung der Plasmide
GB1	12.000	31,3	1	0	-
GB2	15.000	34,5	0	0	-
GB8	11.000	38,5	2	0	-
GB10	15.000	31,3	3	2	pTWB3; pTWB4
Summe	53.000	130,1	5	2	pTWB3; pTWB4

Im Rahmen des initialen Screenings konnten 6 Klone mit rotem Phänotyp aus drei unterschiedlichen Genbanken (GB1, GB8 und GB10) identifiziert werden. Nach erfolgter Retransformation der isolierten Plasmide zeigten nur noch zwei Klone, die beide aus der Genbank GB10 stammten, einen stabilen roten Phänotyp auf den Indikatorplatten. Die Genbank GB10 wurde aus Komposterde ohne vorherigen Inkubationsschritt hergestellt (s. Tab. 6). Die Bezeichnung der Plasmide dieser Klone lautet pTWB3 und pTWB4.

Insgesamt wurden im Rahmen des Screenings 130,1 Mbp Umwelt-DNA nach Genen die Butanol-Dehydrogenase-Aktivität vermitteln durchmustert, dabei konnten zwei Plasmide identifiziert werden, die einen stabilen Phänotyp aufweisen. Das ergibt einen Wert von einem positiven Klon auf 65 Mbp durchmusterter Umwelt-DNA.

3.2.3 Molekulare Analyse der Plasmide pTWB3 und pTWB4

Die Sequenzierung der aus dem Screening nach Genen für neuartige Butanol-Dehydrogenasen hervorgegangenen Plasmide pTWB3 und pTWB4 wurde durch das „Göttingen Genomics Laboratory“ durchgeführt (s. 2.12.1 und 2.12.2.) Die Prozessierung, Assemblierung und Editierung der erhaltenen Sequenzen erfolgte manuell mit dem Programm „Sequence Manager“ auf einem Macintosh PowerPC (s. 2.12.6). Die Sequenzierungen ergaben Insertgrößen von 1723 bp für pTWB3 und 2346 bp für pTWB4. Die G+C-Gehalte der klonierten Umwelt-DNA-Fragmente lagen bei 69,6 % für pTWB3 und 55,6 % für pTWB4. Als nächstes erfolgte die Suche nach möglichen offenen Leserahmen auf den beiden Insertsequenzen. Dazu wurde die Software „Artemis“ verwendet (RUTHERFORD et al., 2000). Die ORF-Festlegung erfolgte dabei manuell mit den unter 2.12.7.1 beschriebenen Kriterien. Die genetischen Organisationen der Inserts von pTWB3 und pTWB4 sind in den Abbildung 12 bzw. 14 dargestellt und die DNA-Sequenzen mit den festgelegten offenen Leserahmen in Abbildung 13 bzw. 15 gezeigt.

Abbildung 12: Genetische Organisation des Inserts von pTWB3.

In schwarz dargestellt sind die beiden offenen Leserahmen, *orf31* und *orf32*, des Inserts von Plasmid pTWB3. Da es sich bei beiden Leserahmen um keine vollständigen Gene handelt, mussten sie jeweils durch die Vektorsequenz des Klonierungsvektors pCR2.1-TOPO komplementiert werden. Dies ist in grau dargestellt und durch Sternchen markiert.

Für das Inserts von pTWB3 konnten zwei offene Leserahmen festgelegt werden, die als *orf31* und *orf32* bezeichnet wurden. Beide *orfs* liegen auf dem Plusstrang der Insert-DNA und bilden keine vollständigen Leserahmen. Um komplette Gene bilden zu können, müssen sie durch Vektorsequenzen komplementiert werden (s. Abb. 12).

Der *orf31* (Insertposition: 1 - 714) mit einer Gesamtgröße von 786 bp wird am Anfang durch 72 bp Vektorsequenz des Klonierungsvektors pCR2.1-TOPO komplementiert (s. Abb. 13). Ohne Vektorsequenz liegt der G+C-Gehalt bei 69 %, mit der Vektorsequenz bei 68 %. Das aus *orf31* abgeleitete Genprodukt besitzt eine Größe von 261 Aminosäuren (24 Aminosäuren aus der Vektorsequenz plus 237 aus der Insertsequenz). Ein Sequenzabgleich auf Proteinebene bei der NCBI-Datenbank zeigte für das *orf31*-Genprodukt (ohne die Aminosäuren, die aus der Vektorsequenz resultieren) die größten Übereinstimmungen (43 % Identität bzw. 58 % Ähnlichkeit) zu einer Acyltransferase aus *Anaeromyxobacter dehalogenans* (s. Tab. 8). Weiterhin konnten Übereinstimmungen zu den konservierten Domänen von Acyltransferasen (pfam01553 und COG0204) gefunden werden (s. Tab. 8).

Das *orf32*-Gen (Insertposition: 802 – 1723) hat komplementiert eine Gesamtgröße von 999 bp, die sich aus 922 bp Insertsequenz zusammensetzt,

die am Ende von 77 bp Vektorsequenz ergänzt wird (s. Abb. 13). Der G+C-Gehalt liegt ohne Vektorsequenz bei 71 % und mit bei 70 %. Das abgeleitete Genprodukt mit einer Gesamtgröße von 332 Aminosäuren (307 Aminosäuren aus der Insertsequenz plus 25 aus der Vektorsequenz) zeigte die größte Ähnlichkeit zu einer potentiellen Dihydroflavonol 4-Reduktase aus *Anaeromyxobacter dehalogenans*. Die Identität betrug 53 % und die Ähnlichkeit 68 %. Zusätzlich lagen Übereinstimmungen zu der konservierten Domäne der 3- β -Hydroxysteroid-Dehydrogenase/Isomerase-Familie (pfam01073) und der konservierten Domäne von Epimerasen (COG0451) vor (s. Tab. 8).

pTWB3:

orf31

V ACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGGTAC

V CGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTT

1 CGCGCGCTCTACGAGAACGTTCTGCACGCGCACGTCGAGGGGCAGTCGAA

51 CATCCCGACGCACACCAACTTCATCGCGGCGACCAACCACACGTCGCACC

101 TCGACACCGGGCTCATCAAGTACGCGCTCGGCGAGTACGGCCGCGACATG

151 GTGGCGCTGGCGGCGGCCGACTACTTCTTCGACAACCGCTACAAGCGCGC

201 CTACATGAACAACCTTCACGAACATCGTGCCCATCGAGCGTTCGGGGAGTC

251 TGCGCCAGTCGCTCCGCCACGCGCGCTCGTTCCTTGAGCGCGGCTACGCG

301 GCGCTCATCTTCCCCGAGGGGACGCGCTCGATGTCGGGGCAGATGGCGGA

351 GTTCAAGTCGGGCTTCGCCTACCTCGCCTTGAACACGCGGACCGGCATCC

401 TGCCCGCGTACCTGTGGGGCACGTACGAGGCGTTCCCCAAGGGCTCGACC

451 TTCCTCAAGGGGCGCGAGGTCGGCGCGCGCATCGGCCGCTTCCTCGACAT

501 CGAGGAGCTGGAAGATTTGACCAAGGGGATGCCGCGCGCCGAGGCCTACC

551 GGCTCGTCGCCGCGCTCGTCCGCCACGAGATAGAGAACCTGCGCGACGGC

601 ACGCGCCACGCCTTCGACGGGGCCGCGCTGCGCCGCCGCTGGCGCGCCGA

651 GCGCCGCCACGAGACCCCGAGGCCGCGCACGACGAGGGCGAGCTGGCGA

701 CCACCGGCGATTGAAAGTTTGAACGCTGAACGCCGGACGATGACCTGAAG

751 GCTTCCGGCTTTCACCTTCACCGTCCGGCGTTCAGCGTCCATCGTTCCGAA

orf32

801 CATGGCGACGACTAAGAAAACCAGCACGAAGAAGGCCGCGACGAAGGCGA

851 AGAAGGCCGCGGCCAACCCGCCGAGGCTGCCGCCGGCGCGTCAAGACG


```

901  ACCTCGCTCGCCAGGCGCGCGAAGCCGGTCACGCTCGTCACGGGCGGGAC
951  GGGCTTCCTCGGCTCGCACCTCGTCAGACTTCTCGTCGAGGAGGGCGCGG
1001 CGACCCTGCGCGTGCTCTCGACCTCGAAGCCCGCGTGGCTCGAAGGGCTG
1051 GGGGTGGAGACGGTGGTCGGGTCGGTCAACGAGGCCGCGGACGTGCGGCG
1101 CGCGCTCGAAGGCGTGGCCGAGGTCTACCACCTCGCGGGGCGCGTCTCGC
1151 GCGAGCCGGAAGACGCGCACATGATGTACAAGCTGCACGTCGACGGCACG
1201 CGCGTGCTGTGCGAGGCGGCGCGCGAGGCGAATGTGAAGTCCATCGTCAT
1251 GGCTTCGACGAGCGGGACGGTCGCGGTACCGAGACGGGCGACCTCGCGC
1301 CCGACGAGGAGTGGCCGACGCCGCTCGACATCGTCTCGCGGTGGCCCTAC
1351 TACGCGAGCAAGGTCTACCAGGAGCGCGTCGCGCTGGAGTACTTTTCGGG
1401 GAAGGGGCGGCGGCTCGTCCTGATGAACCCGAGCCTGCTGCTCGGGCCGG
1451 GCGACGAGCGCCTGAGCTCGACCAAGGTCGTGCTCGACTTCCTCGCGCGC
1501 AAGATCCCGACCGTCCCGAAGGGGGGCCTGAGCTTCGTGGACGCGCGCGA
1551 CGCGGCCCGCGCCTTCCGCGCGGCGATGGAGCGCGGCGCGCACGGCGAGC
1601 GCTACCTGCTGGGCTCGGCCAATTGGACGTTTCGACAAGTTCTTCGGGCGG
1651 CTGGAGCGCATCACGAAGGTGCCCGCCCCCTTCCTCTCGCTCCCCTCGAA
1701 GCTCTCGCTACCGGCGCGCGCGAAGGGCGAATTCTGCAGATATCCATCA
V    CACTGGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTTCGCCTATAG

```

Abbildung 13: DNA-Sequenz des pTWB3-Inserts.

Die Zahlen geben die Position der Basen innerhalb des Inserts an.
V: Vektorsequenz des Vektor pCR2.1-TOPO, welche auch an der schwarzen kursiven Schrift zu erkennen ist. Farblich unterlegt sind die zwei mit „Artemis“ ermittelten, offenen Leserahmen. Türkis: *orf31*; lila: *orf32*

Da die Leserahmen innerhalb des Inserts keine vollständigen Gene bilden, müssen sie jeweils durch Vektorsequenzen (gezeigt durch Transparenz der farblichen Unterlegung) komplementiert werden. Die Pfeile geben die Leserichtung der Gene an. Unterstrichen dargestellt ist die für das Gen jeweilige potentielle „Shine-Dalgarno-Sequenz.“

Genau wie bei dem Insert von pTWB3 konnten bei pTWB4 zwei offene Leserahmen (*orf41* und *orf42*) festgelegt werden, die aber auch hier keine vollständigen Gene bilden und somit durch die Vektorsequenz des Klonierungsvektors pCR2.1-TOPO ergänzt werden mussten (s. Abb. 13). Weiterhin haben die beiden *orfs* unterschiedliche Leserichtungen.

Abbildung 14: Genetische Organisation des Inserts von pTWB4.

In schwarz dargestellt sind die beiden offenen Leserahmen *orf41* und *orf42*, des Inserts von pTWB4. Da es sich bei beiden Leserahmen um keine vollständigen Gene handelt, mussten sie jeweils durch die Vektorsequenz des Klonierungsvektors pCR2.1-TOPO komplementiert werden. Diese ist in grau dargestellt und durch Sternchen markiert.

Der offene Leserahmen *orf41* hat eine Gesamtgröße von 699 bp, die sich aus 554 bp Insertsequenz (Position. 1 - 548) und 151 bp Vektorsequenz zusammensetzen (s. Abb. 15 und 16). Der G+C-Gehalt des komplementierten *orfs* liegt bei 56 %, ohne die Vektorsequenz beträgt dieser 58 %. Das von *orf41* abgeleitete Genprodukt hat eine Größe von 232 Aminosäuren (182 Aminosäuren aus der Insertsequenz und 50 aus der Vektorsequenz) und zeigt die größten Übereinstimmungen (59 % Identität bzw. 78 % Ähnlichkeit) zu einer Aminopeptidase (Accession-Nr. AAP07400.1) aus *Bacillus cereus*. Weiterhin konnten hohe Ähnlichkeiten zu der konservierten Domäne der Familie der Peptidasen M29 (pfam02073) mit einem „e-Value“ von $3e^{-63}$ und der konservierten Domäne der Leucyl-Aminopeptidasen (COG2309) gefunden werden (s. auch Tab. 8).

Der *orf42* mit einer Gesamtgröße von 1617 bp setzt sich aus 1561 bp Insertsequenz (Position: 785 - 2346) und 56 bp Vektorsequenz zusammen (s. Abb. 14 und 15). Der G+C-Gehalt beträgt sowohl für den vollständigen offenen Leserahmen als auch ohne die komplementierende Vektorsequenz 58 %. Das vollständige Gen kodiert für 538 Aminosäuren (520 Aminosäuren aus der Insertsequenz und 18 aus der Vektorsequenz). Ein Sequenzabgleich auf Proteinebene bei der NCBI-Datenbank zeigt große Ähnlichkeiten zu einer neutralen Zink-Metallopeptidase aus *Bacillus licheniformis* ATCC 14580

(Identität: 57 %; Ähnlichkeit: 72 %), welches auch durch die Übereinstimmungen zu den konservierten Domänen der Peptidase-Familie M3 (pfam01432) und der Oligoendopeptidase F (COG1164; „e-Value“: $4e^{-98}$) bestätigt wurde.

pTWB4:

```

V   TTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAA
V   CAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGGTAC
V   CGAGCTCGGATCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTT
1   GGATCGTCCGTGTCCACGCGCGTCGCCCGGAAAATCGCGTCCCACAGGGC
51  AGGCACCTGTTTCGTGGCGGGCAAGTCAGGAAACACCATGGCCGCCCATT
101 CCGGCGACGGGACAGCAATTACGCTCCAGCTTACCTTGTCCGACATGGTG
151 TAGTTGCGGAACGTCCGCAACGCTTGTCCCGCTGCCTTGTTCGAGTCGC
201 GATTCGCTCCAGCTTGACACCTTTGAGCAAATCCGGGTGGTTCGAGTTGA
251 TGTACAAAACGCCGCTCCGTTTTTCGGCCATCTCCTCCAGCCCTTTTGCC
301 TTCCACTGCGGATAATCGGCAAAAGCTTCATCCGGCGCGAGATCGTATTT
351 CATGCGGGTCAGCTCGTCATCGGTCCAGTCGATATGTACGTGTTTGGCTC
401 CTGCTTCATACGCTTTGCGCGCCACCTTGCGCACGAACGGAGCAGCGGAG
451 AGCGGAGCAGTCACAACCAGTGTTTGCATCGGCTGTACATTGACGCCGAC
501 TTTTACGGCGAGATCGGCATAGCGGTCAAACAAAAGCTCGAAAGACATAG
551 GAAAGCCTCCTTTTCACTTGCCTTCAGCACCTCTCTATTATACCTGCACA
601 ACGGGAGTTGCTCGCTAGTCATTTTGAAAAAATTCCTTGGAAGCGCAG
651 AAACAGACAGATTACGGAAGTTTTTCGCGACATTTTTTCGTTTATTTATTTT
701 TTGGAATGGGAGGAAAATTCAAACGAGAATCAAACCTGAATAATAGAAAA
751 CATTTTATTGTAGGAATCATAAGGAGGCAGACACAGTGAAGACGAGACAG
801 ACGCGTGACCAGGTAGCCGTGGAATATACATGGAACCTGGATGACTTATT
851 TGCCAGTCAAGCAGCGTGGGAGCAGGAGCTGGACGAGATTATCAGCCAAC
901 TGCCGGCCGTAACGGCTTTCAAAGGGCAACTGCACACAGGCGCAGAAACA
951 TTGCTGGCCTGCCTGGAAGCAAAGAAGCGATCGAGGTTTCGCGCCAATCT
1001 TGCGGCTACCTATGCCCCTCTGCGCTCATCCGAGGACGGAACCAATCCGC
1051 AAAATCAGGCAAATTCGCGCGGTGTCGGTGACGTCGTCTCCAGGCTGAAC
1101 GCAGAGCTATCGTTTGTGCCGTGCGAAATTTTGGCCTTGCCGGACGGCAC
1151 GATCGAGAAGTTCCTGCAAGAAGAGCCGGGTCTTGCGCCTTTCCGCAAAA
1201 GCCTTTTCGATCTGCTGGAAACCAAGTCGTATCGGCTGTCTGCGGAAACC

```

orf41

orf42

```

1251 GAAGCGTCGCTTGCCGCACTCGGCGAAGTATTCGGCGCTCCGTACACGAT
1301 CTACCAGCGCGGGAAGCTGTCCGATATGTCATTTGCGCCCGTGCAGGACG
1351 GGGCCGGAGCAGAGCGCCCGGTGTCTTTTGCCCTGTTTGAAGCCGACTAC
1401 GAAATGTCTGCGGACACGACGCTGCGCCGGGCTGCCTACGAATCGTTTTTC
1451 AAACACGCTTGCTGCCTACCAAACTCGTTTTGCCGCTGTGTACGCCACTG
1501 AGGTGAAAAAGCAAACCGTGCTGTCCAGGCTGCGCGGCTTCGAGACAGTG
1551 ACTGACATGCTCTTGACACGACGCAAGTAACACTGGACATGTACCACAA
1601 CATTCTCGATATCATCGGCACAGAGCTGGCGCCGCACATGCGCCGCTACG
1651 CCAAGCTGAAGGCGCGCGAGCTGGGGCTTGAAAAGCTGAAATTCTGCGAC
1701 CTGAAGGCGCCGCTTGATCCGGAATTCAGCCCGCCGATTACGATTGAGGA
1751 AGCGGGCCGCCTGATTAGTCTGCTTTGCAAGTAATGGGACCGGAATACT
1801 CCGCAATCATGGAAAAGGCGTTGAACGAACGCTGGATCGACTACGTGGAC
1851 AACGTGGGCAAATCGACAGGAGCGTTTTGCACGACACCGTACGGCAACCA
1901 CTCCTACATCCTCATTACGTGGTGGGCAACATGCGCAGCGCCTTCACGC
1951 TGGCTCACGAGCTGGGGCACTGCGGCCATTTACGCTGGCGACCCGCGAG
2001 CAAAGCTATACCAATTCGCGTCCGTCTTGTACTTTATCGAAGCGCCGTC
2051 GACGATCAACGAGCTGCTTCTGGGGCAGCACATCATGGCGCAGTCCGACG
2101 ATCCGCGCCTCAAGCGCTGGGTGATTTTGCAACTGATGAATACGTACTAC
2151 CATAACTTCGTGACGCACCTGCTGGAAGGCCGCATGCAGCGCAAAGTGTA
2201 CGAGGCGGCGCAAACGACGTGCCGCTGACGGCAAAAAAGCTGTCGGAAT
2251 GGAAAGGGCAAGTGTTGGCGGAGTTCTGGGGAGACGCCGTGGAGCTTGAC
2301 GAGGCAGCCAGCTTGACGTGGATGCGCCAGCCGCACTACTACATGGAAGG
V   GCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCT
V   A

```

Abbildung 15: DNA-Sequenz des pTWB4-Inserts.

Die Zahlen geben die Position der Basen innerhalb des Inserts an. V: Vektorsequenz des Vektors pCR2.1-TOPO, welche auch an der schwarzen kursiven Schrift zu erkennen ist. Farblich unterlegt sind die zwei mit „Artemis“ ermittelten offenen Leserahmen. Gelb: *orf41*; grün: *orf42*. Da die Leserahmen innerhalb des Inserts keine vollständigen Gene bilden, mussten sie jeweils durch Vektorsequenzen (gezeigt durch Transparenz der farblichen Unterlegung) komplementiert werden. Die Pfeile geben die Leserichtung der Gene an. Unterstrichen dargestellt ist die für das Gen jeweilige potentielle „Shine-Dalgarno-Sequenz“.

Tabelle 8: Aufstellung der identifizierten ORFs der Inserts von pTWB3 und pTWB4 mit ihren nächsten verwandten Proteinen und konservierten Domänen.

Die Identifizierung der verwandten Proteine, der konservierten Domänen sowie die Sequenzähnlichkeitswerte erfolgte mit Hilfe der BLAST Programme der NCBI. Die Sterne zeigen unvollständige ORFs an. Alle Angaben basieren nur auf der Insertsequenz ohne komplementierende Vektorsequenz.

ORF	Größe	Nächstes verwandtes Protein und Accession Nummer	Organismus	E-value	Identität / Ähnlichkeit (Homologer AS-Bereich)	Konservierte Domäne
<i>orf31</i> *	237 AS	Phospholipid/Glycerin-Acetyltransferase; ZP_00402075.1	<i>Anaeromyxobacter dehalogenas</i> 2CP-C	1e ⁻⁴³	43 % / 58 % (1 bis 208)	COG0204; PlsC; 1-Acyl-sn-Glycerin-3-Phosphat-Acetyltransferase; (4e ⁻²¹) pfam01553; Acyltransferase; (1e ⁻¹⁶)
<i>orf32</i> *	307 AS	Putative Dihydroflavonol 4-Reduktase ZP_00402082.1	<i>Anaeromyxobacter dehalogenas</i> 2CP-C	3e ⁻⁶⁸	53 % / 68 % (45 bis 307)	COG0451; WcaG; Nukleosid-Di-Phosphat-Zucker Epimerase; (7e ⁻¹⁸) Pfam01073; 3-Beta-Hydroxysteroid-Dehydrogenase/Isomerase; (6e ⁻¹³)
<i>orf41</i> *	182 AS	Aminopeptidase; AAP07400.1	<i>Bacillus cereus</i> ATCC 14579	7e ⁻⁶¹	59 % / 78 % (1 bis 182)	pfam02073; Peptidase_M29; Thermophile Metalloprotease (29); (3e ⁻⁶³) COG2309; AmpS; Leucyl-Aminopeptidase; (1e ⁻³⁹)
<i>orf42</i> *	520 AS	Peptidase M, neutrale Zink-Metallopeptidase; AAU21932.1	<i>Bacillus licheniformis</i> ATCC 14580	1e ⁻¹⁷³	57 % / 72 % (4 bis 520)	COG1164; Oligoendopeptidase F.; (4e ⁻⁹⁸) pfam01432; Peptidase Familie M3; (5e ⁻²³)

Keines der vier identifizierten Gene (*orf31*, *orf32*, *orf41* und *orf42*) der Plasmide pTWB3 und pTWB4, die aus dem Screening nach potentiellen Butanol-Dehydrogenase-Aktivitäten hervorgegangen waren, zeigte signifikante Sequenzähnlichkeiten zu bekannten Alkohol-Dehydrogenasen oder Oxidoreduktasen. Deswegen sollten Enzymaktivitätsmessungen mit den Genprodukten der Plasmide Aufschluss darüber geben, ob die Genprodukte der Plasmide pTWB3 und pTWB4 tatsächlich Butanol-Dehydrogenase-Aktivitäten vermitteln.

3.2.4 Charakterisierung der Butanol-Dehydrogenase-Aktivität von *E. coli*/WL3/pTWB3 und *E. coli*/WL3/pTWB4

Die Bestimmung verschiedener Alkohol- und Aldehyd-Dehydrogenase-Aktivitäten, insbesondere die Butanol-Dehydrogenase-Aktivität der aus den Inserts von pTWP9 und pTWP13 abgeleiteten Genprodukte, wurde im Rahmen einer Kooperation in der Abteilung für Mikrobiologie und Biotechnologie an der Universität Ulm im Rahmen einer Diplomarbeit durchgeführt und ist zu diesem Zeitpunkt noch nicht ganz abgeschlossen. Alle Messungen fanden in der *E. coli*-Mutante WL3 statt, welche auch beim Platten-Screening als Wirtstamm eingesetzt wurde. Diese Mutante zeigte im Vergleich zu anderen *E. coli*-Stämmen (DH5 α , JM109 und SURE[®]) mit Abstand die geringsten Hintergrundaktivitäten (Daten nicht gezeigt).

Zunächst wurden die Plasmide pTWB3 und pTWB4 in *E. coli* WL3 transformiert. Nach erfolgreicher Transformation erfolgte die Anzucht der rekombinanten *E. coli*-Stämme in 500 ml Kulturen in SOC-Medium unter anaeroben Bedingungen. Bei einer OD von 1 bis 1,5 wurden die Zellen geerntet und Rohextrakte hergestellt. Als Negativkontrolle wurde der Klonierungsvektor pCR2.1-TOPO in *E. coli* WL3 transformiert, analog der rekombinanten Stämme angezogen, geerntet und der Rohextrakt hergestellt. Die Bestimmung verschiedener Alkohol- und Aldehyd-Dehydrogenase-Aktivitäten der Rohextrakte erfolgte mit unterschiedlichen Cofaktoren im optisch-enzymatischen Test. Aus Tabelle 9 lassen sich die ermittelten Enzymaktivitäten der Rohextrakte der rekombinanten

E. coli-Stämme sowie die Enzymaktivitäten des Rohextraktes der Negativkontrolle entnehmen.

Tabelle 9: Ergebnisse der in den Rohextrakten von *E. coli* WL3 mit den Plasmiden pCR2.1-TOPO, pTWB3 und pTWB4 ermittelten Alkohol- und Aldehyd-Dehydrogenase-Aktivitäten.

Getestete Reaktion Substrat → Produkt		Spezifische Aktivitäten [U/mg] von rekombinanten <i>E. coli</i> -Stämmen mit		
		pTWB3	pTWB4	pCR2.1-TOPO
Cosubstrat: NADPH + H ⁺	Formaldehyd → Methanol	0,01	0,01	0
	Acetaldehyd → Ethanol	0,03	0,02	0,01
	Propionaldehyd → Propanol	0,03	0,06	0,01
	Butyraldehyd → Butanol	0,08	0,18	0
	Valeraldehyd → Pentanol	0,06	0,15	0
	Hexaldehyd → Hexanol	0,07	0,10	0
Cosubstrat: NADH + H ⁺	Formaldehyd → Methanol	0	0	0
	Acetaldehyd → Ethanol	0	0	0
	Propionaldehyd → Propanol	0	0,01	0
	Butyraldehyd → Butanol	0,01	0,03	0
	Valeraldehyd → Pentanol	0	0,02	0,01
	Hexaldehyd → Hexanol	0,01	0,01	0
Cosubstrat: NAD ⁺	Methanol → Formaldehyd	0	0	0
	Ethanol → Acetaldehyd	0	0	0
	Propanol → Propionaldehyd	0	0	0
	Butanol → Butyraldehyd	0	0,02	0
	Pentanol → Valeraldehyd	0	0	0
	Hexanol → Hexaldehyd	0	0	0
Cosubstrat: NADP ⁺	Methanol → Formaldehyd	0	0	0
	Ethanol → Acetaldehyd	0	0	0
	Propanol → Propionaldehyd	0	0	0
	Butanol → Butyraldehyd	0	0,01	0
	Pentanol → Valeraldehyd	0	0	0
	Hexanol → Hexaldehyd	0	0	0

Wie erwartet, konnte im Rohextrakt der Negativkontrolle *E. coli*/WL3/pCR2.1-TOPO bei allen durchgeführten Testansätzen nahezu keine Butanol-Dehydrogenase-Aktivität ermittelt werden. Die rekombinanten *E. coli* WL3-Stämme mit den Plasmiden pTWB3 und pTWB4 zeigten dagegen deutliche Alkohol-Dehydrogenase-Aktivitäten in Reduktionsrichtung mit den in der Tabelle 9 aufgeführten Aldehyden als Substrat und NADPH+H⁺ als Cosubstrat. Die höchste spezifische Aktivität konnte dabei für den Rohextrakt von *E. coli*/WL3/pTWB4 mit 0,18 U/mg bei der Umsetzung von Butyraldehyd zu Butanol gemessen werden. Außerdem konnte für diesen Rohextrakt auch eine starke Umsetzung von Valeraldehyd zu Pentanol und von Hexaldehyd zu Hexanol gezeigt werden. Der Rohextrakt von *E. coli*/WL3/pTWB3 zeigte ebenfalls die höchste spezifische Enzymaktivität bei der Reaktion von Butyraldehyd zu Butanol. Diese lag jedoch mit 0,08 U/mg deutlich unter der von *E. coli*/WL3/pTWB4. Weiterhin konnten auch mit dem Rohextrakt von *E. coli*/WL3/pTWB3 die Umsetzungen der längerkettigen Aldehyde zu den entsprechenden Alkoholen gezeigt werden. Wurde dagegen das Cosubstrat NADH+H⁺ eingesetzt, so waren in Reduktionsrichtung die Enzymaktivitäten der Rohextrakte deutlich geringer als mit dem Cosubstrat NADPH+H⁺ oder teilweise auch gar nicht mehr messbar. Beim Einsatz der entsprechenden Alkohole als Substrat (Oxidationsrichtung), unabhängig vom verwendeten Cosubstrat, wurde nahezu keine Enzymaktivität gemessen werden. Die einzige Ausnahme zeigte der Rohextrakt von *E. coli*/WL3/pTWB4, bei dem eine sehr geringe spezifische Enzymaktivität bei der Reaktion von Butanol zu Butyraldehyd gemessen werden konnten (s. Tab. 9).

Da die auf pTWB3 und pTWB4 lokalisierten Gene den entsprechenden rekombinanten *E. coli*-Klonen Butanol-Dehydrogenase-Aktivitäten vermitteln konnten, sollte als nächstes untersucht werden, welches der jeweils 2 vorhandenen Gene (Abb. 12 und Abb. 14) für die Aktivität verantwortlich ist. Es erfolgte eine Subklonierung jedes einzelnen *orfs* mit potentieller natürlicher Promotorregion getrennt in einen pUC18-Klonierungsvektor, so dass die von jedem ORF vermittelte Aktivität einzeln in der *E. coli*-Mutante WL3 charakterisiert werden konnte. Die durchgeführten Enzymmessungen ergaben, dass bei dem Plasmid pTWB3 sowohl der *orf31* als auch der *orf32* Genprodukte

bilden, die Butanol-Dehydrogenase-Aktivität vermittelten. Die Rohextrakte der Einzelklone mit *orf31* oder *orf32* zeigten beide Butanol-Dehydrogenase-Aktivität, die jeweils fast die gleiche spezifische Aktivität von 0,08 U/mg bzw. 0,07 U/mg erreichten wie der Rohextrakt von pTWB3 (0,08 U/mg).

Bei dem Plasmid pTWB4 dagegen wurde eindeutig das Genprodukt von *orf42* identifiziert, welches die Butanol-Dehydrogenase-Aktivität vermittelt. Es konnte für das Genprodukt von *orf42* die gleiche spezifische Aktivität von 0,18 U/mg ermittelt werden, die auch schon für den Rohextrakt von pTWB4 gezeigt worden ist. Der erreichte Wert von 0,18 U/mg ist genauso hoch, wie die Gesamtaktivität der Enzyme AdhE, BdhA und BdhB aus *Clostridium acetobutylicum*, die von Dürre und seinen Mitarbeitern (DÜRRE et al. 1987) unter denselben Testbedingungen (allerdings NADH-abhängig) bestimmt werden konnte. Für das Genprodukt von *orf41* konnte dagegen keine signifikante spezifische Butanol-Dehydrogenase-Aktivität (0,01 U/mg) festgestellt werden. Interessanterweise ergaben die molekularbiologischen Untersuchungen keinerlei Hinweise darauf, dass *orf42* für eine Butanol-Dehydrogenase oder andere Oxidoreduktase kodiert. Das abgeleitet Genprodukt zeigte starke Ähnlichkeiten zu einer neutralen Zink-Metalloproteinase aus *Bacillus licheniformis* (s. Tab. 8).

Die Ergebnisse deuten darauf hin, dass die Butanol-Dehydrogenase-Aktivität von *E. coli* WL3-pTWB4 nicht durch eine Alkohol-Dehydrogenase-Reaktion des Genprodukts von *orf42* vermittelt wurde. Es kann vermutet werden, dass die Reaktion durch einen indirekten Weg geschieht, indem die Aktivität des *orf42*-Genprodukts einem wirtseigenen, bisher unbekannten *E. coli*-Enzym die Substratumsetzung ermöglicht.

3.3 Durchmusterung von Metagenombanken nach Genen, die proteolytische Aktivität vermitteln

3.3.1 Screening-Strategie

Die Durchmusterung von Metagenombanken nach Genen mit proteolytischer Aktivität wurde in einem aktivitätsbasierenden Screeningverfahren auf LB-Skim-Milk-Agarplatten durchgeführt (s. 2.10.5). Die Testagarplatten sind durch eine starke Trübung charakterisiert, welche durch das emulgierte Substrat Skim-Milk hervorgerufen wird. Weist ein rekombinanter *E. coli*-Stamm proteolytische Aktivität auf den Testagarplatten auf, führt dies zur hydrolytischen Spaltung der Peptidbindungen des Magermilchproteins, welches durch eine Aufklärung des Mediums angezeigt wird. Positive rekombinante *E. coli*-Stämme wurden daher anhand einer Aufklarungszone um die Kolonie detektiert (s. Abb. 16). Als Wirtsstamm konnte *E. coli* Top10 eingesetzt werden, da dieser Stamm keine unspezifischen Hintergrundaktivitäten auf den Testplatten zeigte.

Abbildung 16: Phänotyp positiver *E. coli*-Klone auf LB-Skim-Milk-Agarplatten.
Der Pfeil zeigt einen rekombinanten *E. coli*-Stamm, der proteolytische Aktivität aufweist.

3.3.2 Ergebnisse des Screenings nach Genen, die proteolytische Aktivität vermitteln

Zur Identifizierung von Genen für neuartige Proteasen wurden rekombinante Plasmide von allen im Rahmen der Arbeit konstruierten 12 Metagenombanken sowie von schon in der Arbeitsgruppe existierenden 17 weiteren metagenomischen Plasmidgenbanken in den Wirtstamm *E. coli* Top10 transformiert (s. 2.10.1). Geeignete Verdünnungen der Transformationsansätze wurden auf den LB-Skim-Milk-Agarplatten ausplattiert und für 2 bis 4 Tage bei 37 °C inkubiert. Weiterhin wurden für das Screening auch zwei „large-insert“ Metagenombanken in Cosmiden eingesetzt. Die Genbanken, die aus heimischen Flusssedimenten konstruiert wurden, waren als rekombinante *E. coli*-Stämme in Mikrotiterplatten hinterlegt und konnten somit auf die Testplatten überstempelt werden. Auch diese Platten wurden für 2 bis 4 Tage bei 37 °C inkubiert. Es erfolgte eine tägliche Kontrolle aller Testplatten. Positive Klone mit proteolytischer Aktivität, welche durch Aufklärung des Mediums um die Kolonie zu erkennen waren, wurden im Vereinzelungsausstrich auf neue Testplatten ausgestrichen. Zeigten die rekombinanten *E. coli*-Stämme dann erneut proteolytische Aktivität, wurden sie in einer 5 ml Kultur angezogen und die Plasmide bzw. Cosmide isoliert. Zur Überprüfung des Phänotyps erfolgte eine Retransformation (s. 2.10.1 für Plasmide bzw. 2.10.2 für Cosmide) der isolierten Plasmide oder Cosmide in *E. coli* Top10-Zellen und ein erneutes Screening auf LB-Skim-Milk-Agarplatten.

Im Rahmen des initialen Screenings konnten 19 Klone mit potentieller proteolytischer Aktivität aus 9 unterschiedlichen Metagenombanken identifiziert werden (s. Tab. 10). Dabei stammten 4 der 19 positiven Klone aus „large-insert“ Metagenombanken. Nach erfolgter Retransformation der isolierten Plasmide/Cosmide zeigten nur noch 5 der 19 identifizierten *E. coli*-Klone proteolytische Aktivität und somit einen stabilen Phänotyp. Die 5 positiven Klone verteilen sich auf drei Plasmidgenbanken (je ein Klon von GB5 und GB8 sowie 2 Klone aus GB12; s. Tab. 6) und auf eine Cosmidgenbank (1 Klon).

Tabelle 10: Zusammenfassung der Ergebnisse des Screenings nach Genen, die proteolytische Aktivität vermitteln.

Genbank	Anzahl getesteter Klone	Getestete Umwelt-DNA [Mbp]	Positive Klone nach initialem Screening	Klone mit stabilem Phänotyp	Bezeichnung der Plasmide bzw. Cosmide
GB1	12.000	31	2	-	-
GB2	12.000	28	3	2	pTWP9; pTWP13
GB3	17.000	58	-	-	-
GB4	13.000	36	2	-	-
GB5	18.000	71	2	1	pTWP3
GB6	20.000	39	-	-	-
GB7	25.000	62	-	-	-
GB8	15.000	52	2	1	pTWP8
GB9	13.000	19	-	-	-
GB10	20.000	42		-	-
GB11	13.000	23	1	-	-
GB12	15.000	36	-	-	-
17 weitere Genbanken	330.000	860	3	-	-
2 Cosmid-Genbanken	10.000	350	4	1	pC-GII(A)-7m
Summe:	533.000	1,7 Gbp	19	5	

Insgesamt wurden 1,7 Gbp Umwelt-DNA nach Genen, die proteolytische Aktivität vermitteln, auf LB-Skim-Milk-Agarplatten durchmustert. Dabei konnten 4 Plasmide und 1 Cosmid identifiziert werden, die den rekombinanten *E. coli*-Stämmen eine stabile proteolytische Aktivität vermittelten. Das ergibt einen Wert von 1 positivem Klon auf ca. 340 Mbp durchmusterter Umwelt-DNA. Die Bezeichnungen der erhaltenen Plasmide bzw. Cosmide der positiven Klone sind Tabelle 10 zu entnehmen. Hinzuzufügen ist, dass die rekombinanten *E. coli*-Stämme mit den Plasmiden pTWP9 und pTWP13 die deutlichste Hofbildung auf den Testplatten zeigten. Beide Plasmide stammen aus der

Genbank GB2, die aus einem Standortprobengemisch mit einem vorgezogenen Inkubationsschritt mit Zimtsäure unter anaeroben Bedingungen konstruiert wurde. Aufgrund des deutlich ausgeprägten Phänotyps wurden diese beiden Plasmide zur weiteren Charakterisierung ausgewählt.

3.3.3 Molekulare Analyse der Plasmide pTWP9 und pTWP13

3.3.3.1 Sequenzierung der Inserts von pTWP9 und pTWP13

Die Sequenzierung der Inserts der Plasmide pTWP9 und pTWP13 konnte nicht durch „Primer-walking“ wie bei den Plasmiden pTWB3 und pTWB4 erfolgen, da sich nach Restriktionsanalyse der isolierten Plasmide herausstellte, dass die Größe der Insertsequenzen in dem Klonierungsvektor pBluescript® SKII (+) bei beiden Plasmiden jeweils mehr als 15 Kbp betrug. Somit wurden „Shotgun“-Genbanken der Plasmide hergestellt. Dazu erfolgte ein partieller enzymlimitierter Verdau der Plasmide pTWP9 und pTWP13 mit den Restriktionsenzymen *Bsp143I* und *AluI* (s. 2.8.1). Die daraus resultierenden DNA-Fragmente wurden für die Topo-T/A-Klonierung aufgearbeitet (s. 3.1.4) und anschließend in den Vektor pCR2.1-TOPO kloniert. Nach erfolgter Transformation der „Shotgun“-Plasmide in *E. coli* Top10-Zellen erfolgte eine Selektion von rekombinanten Plasmiden über den X-Gal-Test auf Agarplatten (s. 2.10.3). Alle weißen insert-tragenden Klone wurden angezogen, die Plasmide isoliert und zur Sequenzierung verwendet. Die Sequenzierungen wurden vom „Göttingen Genomics Laboratory“ und von der Firma SEQLAB (Göttingen) durchgeführt. Die Prozessierung der Rohsequenzen erfolgte mit dem STADEN Software Paket (s. 2.12.3 und 2.12.4). Innerhalb dieses Pakets wurden die Sequenzen letztendlich mit dem Programm „Gap4“ bearbeitet. Mit Hilfe dieses Programms waren die Assemblierung, die Editierung und der Lückenschluss zur Erhaltung der kompletten Insertsequenz möglich.

Die auf diesem Weg erhaltenen Insertsequenzen der Plasmide pTWP9 und pTWP13 zeigten Insertgrößen von 28.113 bp (pTWP9) und 20.028 bp (pTWP13). Der G+C-Gehalt lag bei 68,3 % für das Insert von pTWP9 und bei 69,5 % für die Insertsequenz von pTWP13.

3.3.3.2 Festlegung von offenen Leserahmen und Annotation der Gene

Die Festlegung von offenen Leserahmen auf den Insertsequenzen der Plasmide pTWP9 und pTWP13 erfolgte manuell mit dem Programm „Artemis“ (RUTHERFORD et al., 2000) mit den unter 2.12.7.1 beschriebenen Kriterien. Mit dem Programm Artemis konnten ebenfalls für alle festgelegten offenen Leserahmen die zu den DNA-Sequenzen korrespondierenden Aminosäure-Sequenzen abgeleitet werden. Die vollständigen DNA-Sequenzen der Inserts der Plasmide pTWP9 und pTWP13 sowie die Positionen der festgelegten offenen Leserahmen und deren abgeleiteten Proteinsequenzen befinden sich im Anhang der Arbeit (s. 7.1 und 7.2). Eine Annotation (s. 2.12.7.2) der Gene erfolgte durch einen Sequenzabgleich der abgeleiteten Aminosäuresequenzen mit gespeicherten Proteinsequenzen der NCBI-Datenbank. Weiterhin konnten auf diese Weise auch mögliche konservierte Domänen identifiziert werden. Eine Funktion wurde einem Genprodukt zugewiesen, wenn mindestens ein E-value von $1e^{-10}$ vorlag. Konnten keine signifikanten Sequenzähnlichkeiten gefunden werden, wurde die Bezeichnung „hypothetisches Protein“ gegeben. Konnten dagegen signifikante Sequenzähnlichkeiten zu anderen konservierten, hypothetischen Proteinen gefunden werden, so erfolgte die Bezeichnung „konserviertes hypothetisches Protein“. Die genetische Organisation des Inserts von pTWP9 ist in Abbildung 17 dargestellt. Die Annotationen der korrespondierenden Genprodukte sowie potentielle konservierte Domänen und die nächsten verwandten Proteine mit Angabe von Sequenzähnlichkeiten können der Tabelle 11 entnommen werden.

Abbildung 17: Genetische Organisation des Inserts von pTWP9.

Die Pfeile markieren Länge und Orientierung von offenen Leserahmen. Unvollständige Leserahmen sind durch Sterne gekennzeichnet. Farbige Pfeile repräsentieren Gene, deren Genprodukt eine Funktion zugewiesen werden konnte. Genprodukte von dunkelgrauen Genen zeigen konserviert hypothetische Proteine, während weiße Pfeile Gene darstellen, deren Genprodukte als hypothetische Proteine bezeichnet wurden. Die gepunkteten Linien markieren Genbereiche, die auch im Insert von pTWP13 vorkommen.

Tabelle 11: Annotationen der ORFs des Inserts pTWP9 und Darstellung des nächsten verwandten Proteins und von konservierten Domänen.

Die Identifizierung der verwandten Proteine, der konservierten Domänen sowie die Ermittlung der Sequenzähnlichkeitswerte erfolgten mit Hilfe der BLAST Programme der NCBI. Unvollständige ORFs sind durch Sterne markiert. Farblich unterlegt ist der ORF, der für die proteolytische Aktivität verantwortlich ist.

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession –Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
1*	230	<i>rplA</i>	50S ribosomales Protein L1	50S ribosomales Protein L1 (AAM40195.1); <i>Xanthomonas campestris</i> pv. <i>campestris</i> str. ATCC 33913	83 % / 92 % (AS: 3-230) 1e ⁻¹⁰²	pfam00687; ribosomales Protein der Familie L1p/L10e 1e ⁻⁵⁰
2	142	<i>rplK</i>	50S ribosomales Protein L11	50S ribosomales Protein L11 (AAM35844.1); <i>X. axanopodis</i> pv. <i>citri</i> str. 306	88 % / 96 % (AS: 1-142) 1e ⁻⁶⁶	smart00649; ribosomales Protein L11/112; 3e ⁻⁴¹
3	200	<i>nusG</i>	Transkriptions-Antiterminationsfaktor	Transkriptions-Antiterminationsfaktor (AAM40193.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	91 % / 95 % (AS: 17-199) 8e ⁻⁹³	COG0250; NusG, Transkriptions-Antiterminator; 7e ⁻⁵²
4	127	<i>secE</i>	Präprotein-Translokase UE	Präprotein -Translokase (UE) (AAM35842.1); <i>X. axanopodis</i> pv. <i>citri</i> str. 306	57 % / 74 % (AS: 1-127) 6e ⁻³⁸	-
5	396	<i>tufA</i>	Elongationsfaktor Tu	Elongationsfaktor Tu (CAJ2267.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	92 % / 96 % (AS: 1-396) 0	pfam00009; GTP EFTU, Elongationsfaktor Tu GTP-Bindedomäne; 3e ⁻⁵⁵

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession -Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
6	363	<i>ychF</i>	GTP-Bindeprotein	GTP-Bindeprotein (BAE70157.1); <i>X. oryzae</i> pv. <i>oryzae</i> MAFF 311018	88 % / 94 % (AS: 1-363) 0	pfam06071; DUF933, Protein mit unbekannter Funktion; 3e ⁻³⁶
7	193	<i>pth</i>	Peptidyl-tRNA-Hydrolase	Peptidyl-tRNA Hydrolase (AAM40189.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	77 % / 89 % (AS: 1-193) 6e ⁻⁸⁷	pfam01195; Peptidyl-tRNA-Hydrolase; 2e ⁻⁴⁸
8	226	<i>rpY</i>	50S ribosomales Protein L25	50S ribosomales Protein L25 (AAM40188.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	63 % / 79 % (AS: 4-206) 8e ⁻⁷²	COG1825; RplY, ribosomales Protein L25; 1e ⁻²²
9	319	<i>prsA</i>	Ribose-Phosphat-Pyrophosphokinase	Ribose-Phosphat-Pyrophosphokinase (CAJ22611.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	84 % / 91 % (AS: 1-319) 1e ⁻¹⁴⁸	COG0462; PrsA, Phosphoribosylpyrophosphat-Synthetase; 8e ⁻⁹⁷
10	302	<i>ipK</i>	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase (YP_202243.1); <i>X. oryzae</i> pv. <i>oryzae</i> KACC10331	70 % / 77 % (AS: 9-295) 3e ⁻¹⁰⁷	COG1947; IspE, 4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase; 3e ⁻⁴⁵
11	212	<i>loIB</i>	Lipoprotein-Precursor der äußeren Membran	Lipoprotein-Precursor der äußeren Membran (AAM40185.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	50 % / 61 % (AS: 3-206) 4e ⁻⁴⁹	pfam03550; LOIB, outer membrane Lipoprotein; 2e ⁻²³
12	557	<i>orf12</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40184.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	63 % / 77 % (AS: 24-556) 0	-

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
13	436	<i>hemA</i>	Glutaryl-tRNA-Reduktase	Glutaryl-tRNA-Reduktase (BAE70164.1); <i>X. oryzae</i> pv. <i>oryzae</i> MAFF 311018	68 % / 79 % (AS: 1-424) 5e ⁻¹⁵⁴	COG0373; HemA, Glutaryl-tRNA Reduktase; 3e ⁻¹¹⁶
14	360	<i>prfA</i>	Peptidketten-Release-Faktor	Peptidketten-Release-Faktor (BAE70165.1); <i>X. oryzae</i> pv. <i>oryzae</i> MAFF 311018	78 % / 88 % (AS: 1-357) 9e ⁻¹⁵⁵	COG0216; PrfA, Peptidketten-Release-Faktor A; 5e ⁻¹²⁴⁴
15	148	<i>orf15</i>	GCN5-ähnliche N-Acetyltransferase	GCN5-ähnliche N-Acetyltransferase (ABA75299.1); <i>Pseudomonas fluorescens</i> PfO-1	45 % / 62 % (AS: 7-148) 6e ⁻²⁷	-
16	172	<i>moaB</i>	Molypdopterin Biosynthese-Protein	Molypdopterin Biosynthese-Protein (AAM40180.1); <i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913	77 % / 86 % (AS: 1-172) 3e ⁻⁷¹	pfam00994; Mögliche Molypdopterin Bindedomäne; 3e ⁻¹⁵
17	258	<i>orf17</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (CAJ22603.19); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	70 % / 80 % (AS: 1-256) 1e ⁻¹⁰⁴	pfam03976; DUF344, Domäne mit unbekannter Funktion; 2e ⁻⁷⁷
18	90	<i>acyP</i>	Acylphosphatase	Acylphosphatase (ZP_00473130.1); <i>Chromohalobacter salexigens</i> DSM 3043	54 % / 69 % (AS: 9-90) 6e ⁻¹⁷	COG1254; AcyP, Acylphosphatase; 5e ⁻²²
19	194	<i>yneN</i>	Mögliches Thioredoxin	Mögliches Thioredoxin (CAJ22600.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	57 % / 71 % (AS: 3-189) 1e ⁻⁵⁹	COG0526; TrxA, Thiol-Disulfid-Isomerase und Thioredoxin; 1e ⁻¹⁰

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
20	452	<i>rbn</i>	Ribonuklease BN	Ribonuklease BN (AAM35825.1); X. <i>axanopodis</i> pv. <i>citri</i> str. 306	59 % / 75 % (AS: 31-448) 4e ⁻¹⁴³	pfam03631; Ribonuklease BN-ähnliche Familie; 4e ⁻³²
21	197	<i>trp</i>	Tryptophan-Repressor-Bindeprotein	Tryptophan-Repressor-Bindeprotein (BAE70173.1); X. <i>oryzae</i> pv. <i>oryzae</i> MAFF 311018	79 % / 90 % (AS: 1-196) 1e ⁻⁸⁷	COG0655; WrbA, Multimeres Flavodoxin WrbA; 3e ⁻²³
22	113	<i>orf22</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40173.1); X. <i>campestris</i> pv. <i>campestris</i> str. ATCC 33913	53 % / 66 % (AS: 7-109) 8e ⁻²⁰	-
23	177	<i>orf23</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
24	217	<i>orf24</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
25	217	<i>orf25</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
26	162	<i>aspG</i>	Asparaginase	Asparaginase (AAW76873.1); X. <i>oryzae</i> pv. <i>oryzae</i> KACC10331	81 % / 91 % (AS: 1-161) 8e ⁻⁷¹	cd00411; Asparaginase; 5e ⁻¹¹

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
27	113	<i>orf27</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAR36832.1) <i>Geobacter metallireducens</i> GS-15	57 % / 74 % (AS: 5-112) $2e^{-29}$	COG3189; Uncharakterisiertes konserviertes Protein; $2e^{-33}$
28	49	<i>orf28</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
29	862	<i>mprA</i>	Metalloprotease	Metalloprotease (BAB85124.1); <i>Pseudomonas sp. A28</i>	49 % / 65 % (AS: 10-737) 0	Ausführliche Beschreibung s. Kap. 3.3.4
30	183	<i>orf30</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM35814.1); <i>X. axanopodis</i> pv. <i>citri</i> str. 200	84 % / 91 % (AS: 1-183) $6e^{-85}$	pfam01883; Domäne mit unbekannter Funktion DUF59; $1e^{-12}$
31	142	<i>orf31</i>	Konserviertes hypothetisches Protein	Hypothetisches Protein (YP_362687.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	49 % / 62 % (AS: 19-142) $7e^{-23}$	-
32*	259	<i>ptnb</i>	Mögliche NAD(P)-Transhydrogenase β -UE	Mögliche NAD(P)-Transhydrogenase β -UE (CAJ22586.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	92 % / 94 % (AS: 1-258) $5e^{-130}$	pfam02233; PNTB, NAD(P)-Transhydrogenase β UE; $9e^{-88}$

Für das Insert von pTWP9, welches eine Größe von 28.113 bp hatte, konnten 32 offene Leserahmen festgelegt werden. Vier der aus den ORFs abgeleiteten Genprodukte zeigten keine signifikanten Ähnlichkeiten zu Proteinsequenzen der NCBI-Datenbank und wurden deswegen als „Hypothetische Proteine“ bezeichnet. Fünf der Genprodukte wiesen Übereinstimmungen zu konservierten hypothetischen Proteinen ohne bekannte Funktion auf. Dagegen konnte 23 abgeleiteten Genprodukten eine Funktion zugewiesen werden. Dabei war auffällig, dass bei der Mehrzahl (24 Genprodukte) das nächste verwandte Protein aus einem Organismus der Gattung *Xanthomonas* stammte (s. Tab. 11). Weiterhin zeigten die abgeleiteten Sequenzen der Genprodukte häufig sehr hohe Identitäten (zwischen 70 % bis 92 %) zu ihren nächsten verwandten Proteinen. Funktionell am häufigsten vertreten sind dabei Genprodukte, die an der bakteriellen Translation beteiligt sind. Die Gene *rplA* und *rplK* kodieren beide für ribosomale Proteine der großen Untereinheit von bakteriellen Ribosomen. Das Genprodukt von *tufA*, der Elongationsfaktor Tu, ist an der Elongation der wachsenden Polypeptidkette beteiligt, indem er die freien Aminoacyl-tRNAs an die ribosomale Aminoacyl-A-Bindungsstelle befördert. Die anderen drei an der bakteriellen Translation beteiligten Genprodukte werden durch die Gene *ychF*, *pth* und *rplY* kodiert. Das Gen *prfA* kodiert ebenfalls für ein Protein des bakteriellen Translationsvorgangs.

Der Orf29 mit dem Gen-Namen *mprA* kodiert für eine Metalloprotease mit einer Größe von 862 Aminosäuren. Diese zeigt auf Proteinebene 49 % Identität und 65 % Ähnlichkeit zu einer charakterisierten Metalloprotease Empl aus *Pseudoalteromonas sp.* A28 (LEE et al, 2002). Da *mprA* das einzige Gen auf pTWP9 darstellt, welches für eine potentielle Protease kodiert, wurde davon ausgegangen, dass es für die proteolytische Aktivität von *E. coli*/pTWP9 verantwortlich war. Dies konnte in biochemischen Untersuchungen bestätigt werden (s. 3.4). Die Protease-Sequenz von MprA wurde deswegen einer genaueren molekularen Charakterisierung unterzogen (s. 3.3.4).

Abbildung 18: Genetische Organisation des Inserts von pTWP13.

Die Pfeile markieren Länge und Orientierung von offenen Leserahmen. Unvollständige Leserahmen sind durch Sterne gekennzeichnet. Farbige Pfeile repräsentieren Gene, deren Genprodukten eine Funktion zugewiesen werden konnte. Genprodukte von dunkelgrauen Genen zeigen konservierte, hypothetische Proteine, während weiße Pfeile Gene darstellen, deren Genprodukte als hypothetische Proteine bezeichnet wurden. Die gepunkteten Linien markieren Genbereiche, die auch im Insert von pTWP9 vorkommen.

Tabelle 12: Annotationen der ORFs des Inserts pTWP13 sowie Darstellung der nächsten verwandten Proteine und von konservierten Domänen.

Die Identifizierung der verwandten Proteine, der konservierten Domänen sowie der Sequenzähnlichkeitswerte erfolgte mit Hilfe der BLAST Programme der NCBI. Unvollständige ORFs sind durch Sterne markiert. Farbig unterlegt ist der ORF, der für die proteolytische Aktivität verantwortlich ist.

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession -Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
1*	210	<i>orf1</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
2	93	<i>acyP</i>	Acylphosphatase	Acylphosphatase (ZP_00600900.1); <i>Rubrobacter xylanophilus</i> DSM 9941	52 % / 63 % (AS: 6-92) 3e ⁻¹⁸	COG1254; AcyP, -14 Acylphosphatasen; 1e
3	192	<i>yneN</i>	Mögliches Thioredoxin	Putatives Thioredoxin (CAJ22600.1); <i>Xanthomonas campestris</i> pv. <i>vesicatoria</i> str. 85-10	61 % / 71 % (AS: 3-187) 2e ⁻⁶¹	COG0526; TrxA, Thiol-Disulfid-Isomerase und Thioredoxine; 1e ⁻¹¹
4	425	<i>rbn</i>	Ribonuklease BN	Ribonuklease BN (AAM35825.1); <i>X. axonopodis</i> pv. <i>citri</i> str. 306	56 % / 73 % (AS: 4-425) 6e ⁻¹³⁶	pfam03631; Ribonuklease BN-ähnliche Familie; 5e ⁻²⁴
5	197	<i>trp</i>	Tryptophan-Repressor-Bindeprotein	Tryptophan-Repressor-Bindeprotein (BAE70173.1); <i>X. oryzae</i> pv. <i>oryzae</i> MAFF 311018	79 % / 89 % (AS: 1-196) 1e ⁻⁸⁶	COG0655; WrbA, Multimeres Flavodoxin WrbA; 7e ⁻²²

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
6	110	<i>orf6</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40173.1); <i>X. campestris</i> pv. <i>campestris</i> str. 8004	51 % / 65 % (AS: 1-110) 3e ⁻²⁰	-
7	162	<i>aspG</i>	Asparaginase	Asparaginase (AAM35820.1); <i>X. axonopodis</i> pv. <i>citri</i> str. 306	83 % / 92 % (AS: 1-161) 4e ⁻⁷³	cd00411; Asparaginase; 9e ⁻¹¹
8	862	<i>mprB</i>	Metalloprotease	Metalloprotease (BAB85124.1); <i>Pseudomonas</i> sp. A28	50 % / 66 % (AS: 17-737) 0	Ausführliche Beschreibung s. Kap. 3.3.4
9	51	<i>orf9</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
10	184	<i>orf10</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM35814.1); <i>X. axonopodis</i> pv. <i>citri</i> str. 306	76 % / 87 % (AS: 3-183) 9e ⁻⁷⁶	pfam01883; DUF59, Domäne mit unbekannter Funktion; 1e ⁻¹³
11	147	<i>orf11</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (CAJ22587.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	51 % 62 % (AS: 10-140) 6e ⁻²⁶	-
12	464	<i>pntB</i>	Mögliche β-UE der NAD(P)-Transhydrogenase	Mögliche β-UE der NAD(P)-Transhydrogenase (CAJ22586.1); <i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10	79 % / 85 % (AS: 1-463) 0	pfam02233; PNTB, NAD(P)-Transhydrogenase β-UE; 2e ⁻¹⁰⁵

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
13	101	<i>pntA</i>	α -UE der NAD(P) - Transhydrogenase	α -UE der NAD(P) - Transhydrogenase (AAM35811.1); X. <i>axonopodis</i> pv. <i>citri</i> str. 306	90 % / 95 % (AS: 1-100) $9e^{-45}$	COG3288; PntA, NAD(P)-Transhydrogenase α -UE; $4e^{-7}$
14	195	<i>orf14</i>	RNA-Polymerase ECF-Sigma-Faktor	RNA Polymerase ECF-Sigma-Faktor (CAJ22585.1); X. <i>campestris</i> pv. <i>vesicatoria</i> str. 85-10	68 % / 75 % (AS: 14-194) $1e^{-61}$	COG1595; RpoE, RNA-Polymerase Sigma UE; $3e^{-13}$
15	131	<i>orf15</i>	Hypothetisches Protein	Keine signifikanten Übereinstimmungen		
16	152	<i>orf16</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40160.1); X. <i>campestris</i> pv. <i>campestris</i> str. 8004	43 % / 57 % (AS: 1-151) $6e^{-31}$	-
17	377	<i>pnt</i>	Pyridin-Nukleotid-Transhydrogenase	Pyridin-Nukleotid-Transhydrogenase (YP_202272.1); X. <i>oryzae</i> pv. <i>oryzae</i> KACC10331	72 % / 81 % (AS: 1-367) $3e^{-140}$	COG3288; PntA, NAD(P)-Transhydrogenase α -UE; $2e^{-74}$
18	744	<i>orf18</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40157.1); X. <i>campestris</i> pv. <i>campestris</i> str. ATCC 33913	32 % / 48 % (AS: 18-732) $8e^{-81}$	-
19	199	<i>orf19</i>	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM40155.1); X. <i>campestris</i> pv. <i>campestris</i> str. ATCC 33913	66 % / 75 % (AS: 4-189) $3e^{-66}$	-

ORF	Größe in AS	Gen-name	Annotation	Nächstes verwandtes Protein (Accession-Nr.) und Organismus	Identität / Ähnlichkeit (Homologer AS-Bereich) E-value	Konservierte Domänen mit E-value
20	306	exo	Exodeoxyribonuklease	Exodeoxyribonuklease (AAM40154.1); X. campestris pv. campestris str. ATCC 33913	77 % / 84 % (AS: 10-299) 5e ⁻¹²⁷	smart00475; 53EXOc, 5'-3'-Exonuklease; 5e ⁻⁵⁷
21	186	orf21	MutT-ähnliches Protein	MutT-ähnliches Protein (AAM5999.1); X. axonopodis pv. citri str. 306	61 % / 78 % (AS: 10-184) 1e ⁻⁵⁸	COG0494; MutT, NTP Pyrophosphohydrolasen inklusive oxidativer Reparatur-Enzyme; 7e ⁻⁷
22	325	orf22	Konserviertes hypothetisches Protein	Konserviertes hypothetisches Protein (AAM2671.1); Chloobium tepidum TLS	89 % / 94 % (AS: 1-325) 1e ⁻¹⁶⁸	-

Für das 20.028 bp-Insert von pTWP13 mit einem G+C-Gehalt von 69 % konnten 22 offene Leserahmen festgelegt werden (s. Abb. 18). Die von *orf1*, *orf9* und *orf15* abgeleiteten Genprodukte zeigten keine signifikanten Sequenzübereinstimmungen zu Proteinsequenzen in der NCBI-Datenbank. Signifikante Ähnlichkeiten zu Proteinen mit unbekannter Funktion konnten für die Genprodukte von *orf6*, *orf10*, *orf11*, *orf16*, *orf18*, *orf19* und *orf22* festgestellt werden. Besonders auffällig war die hohe Übereinstimmung (89 % Identität und 95 % Ähnlichkeit) des hypothetischen Genprodukts von *orf22* zu einem konservierten hypothetischen Protein aus dem Organismus *Chlorobium tepedium*. Eine Funktion konnte 12 der 22 abgeleiteten Genprodukte zugewiesen werden. Dabei stammt wie bei den abgeleiteten Genprodukten der Insertsequenz von pTWP9 die nächste verwandte Proteinsequenz häufig aus Organismen der Gattung *Xanthomonas*. Vergleicht man die genetischen Organisationen der beiden Insertsequenzen so ist zu erkennen, dass die Aufeinanderfolge der Orfs 2-14 von Insert pTWP13 (s. Abb.18) fast identisch zu der genetischen Organisation der Orfs 18-31 mit Ausnahmen der Orfs 23 bis 25 der Insertsequenz von pTWP9 ist (s. Abb. 17). Es kann deswegen angenommen werden, dass es sich bei den beiden Insertsequenzen um zwei überlappende Genomabschnitte aus eng verwandten Organismen handelt. Diese Vermutung wird durch die Identifizierung des Genprodukts, welches vermutlich für die proteolytische Aktivität verantwortlich ist, bestärkt. Es handelt sich hierbei um das Genprodukt von Orf8, dessen Gen als *mprB* bezeichnet wurde. Ebenso wie das von Gen *mprA* (ORF29, pTWP9; s. Tab.11) abgeleitete Genprodukt kodiert *mprB* für eine Metalloprotease mit einer Größe von 862 Aminosäuren. Als nächstes verwandtes Protein konnte wie bei MprA, die Metalloprotease Empl aus dem Organismus *Pseudoalteromonas sp.* A28 (LEE et al., 2002), identifiziert werden. Allerdings sind die Übereinstimmungen von MprB zu Empl mit 50 % Identität und 66 % Ähnlichkeit etwas größer als die Übereinstimmungen von MprA zu Empl (49 % Identität und 65 % Ähnlichkeit). Diese Unterschiede beruhen darauf, dass die Gensequenzen von *mprA* und *mprB* und die daraus abgeleiteten Proteinsequenzen nicht identisch sind (s. 3.3.4).

3.3.4 Molekulare Charakterisierung von *mprA* und *mprB*

Zunächst erfolgte eine Eingruppierung der potentiellen Proteasen MprA und MprB. Dazu wurden die Aminosäuresequenzen von MprA und MprB in die speziell für Proteasen bestehende Datenbank MEROPS (<http://merpos.sanger.ac.uk/>) geladen und einer Ähnlichkeitssuche mit anderen Proteasen unterzogen. Auch bei dieser Datenbank wurde als nächstes verwandtes Protein die Metalloprotease Empl aus *Pseudoalteromonas sp.* A28 (LEE et al., 2002) identifiziert. Zusätzlich wurden MprA und MprB aufgrund von Sequenzhomologien in die Familie M4 der Metalloproteasen eingestuft, in der auch die Protease Empl vertreten ist. Diese auch als „Thermolysin-ähnliche Proteasen“ bezeichnete Familie umfasst ausschließlich sekretierte Endopeptidasen aus Bakterien. Einzige Ausnahmen bilden zwei archaeele M4-Metalloproteasen aus *Methanosarcina acetivorans* sowie zwei Vertreter aus Pilzen. Eine weitere Gemeinsamkeit der Vertreter der M4-Metalloproteasen ist, dass alle als inaktives „Prä-Proenzym“ (Vorläufer-Protein) gebildet werden. Sie besitzen alle N-terminale Propeptide und einige auch C-terminale Prosequenzen, denen unterschiedliche Funktionen bei der Reifung der Proteasen zugesprochen werden. Durch Abspaltung der Prosequenzen in unterschiedlichen Mechanismen entstehen schließlich die reifen, aktiven Enzyme (Übersicht: RAWLINGS und Barrett, 2004).

Bei dem Aminosäureabgleich der Proteine MprA und MprB mit der NCBI-Datenbank konnten für beide Proteine durch das Tool „Conserved Domain Search“ mehrere konservierte Domänen, die häufig in Metalloproteasen vorkommen, gefunden werden. Die Tabelle 13 gibt einen Überblick über die identifizierten konservierten Domänen sowie deren Lokalisation innerhalb der Proteinsequenzen. Für beide Proteine konnten jeweils sechs konservierte Domänen festgelegt werden. Die Verteilung der konservierten Domänen dehnen sich fast über die kompletten Aminosäuresequenzbereiche von MprA und MprB aus. Im N-terminalen Bereich konnte für beide Proteine jeweils ein Propeptidsequenz, wie sie für M4-Metalloproteasen beschrieben ist, identifiziert werden. Dem Propetid wird eine Funktion als intramolekulares Chaperon zugesprochen. Danach folgten bei beiden Proteasen konservierte

Sequenzbereiche, die typisch für katalytische Domänen von Thermolysin-ähnlichen Metallopeptidasen sind (JONGENEEL et al., 1989). Nach einer strukturellen Domäne, der alpha-helikalen Domäne, konnten je zwei C-terminale Propeptid-Domänen (PPC-Domänen) erkannt werden. Diese Domänen werden nur am C-terminalen Ende sekretierter Proteasen gefunden. Normalerweise sind sie im reifen Enzym nicht anwesend. Die genaue Funktion von bakteriellen, C-terminalen Propeptiden ist noch nicht geklärt. Es wird aber angenommen, dass sie eventuell bei der Lokalisation und Sekretion des Vorläufer-Proteins behilflich sind. PPC-Domänen wurden bisher ausschließlich in Bakterien und Archaeen gefunden, nicht aber in Eukaryoten (YEATS et al., 2003). Am C-Terminus konnte bei MprA und MprB ein Sequenzbereich für eine regulatorische Domäne (P-Domäne) identifiziert werden. Diese Domäne kommt häufiger in Subtilisin-ähnlichen Proprotein-Konvertasen und anderen eher eukaryotischen Proteasen vor. Im Vergleich zu diesen Enzymen sind nur etwa 60 % der Domäne in MprA und MprB vorhanden (s. Tab. 13).

Da es sich bei allen Vertretern der M4-Metalloproteasen ausschließlich um sekretierte Enzyme handelt, benötigen sie für den Export aus der Zelle eine aminoterminal Signalsequenz. Mit Hilfe des Programms „SignalP 3.0“ (<http://www.cbs.dtu.dk/services/SignalP/>) konnte für beide Protease-Sequenzen ein Signalpeptid identifiziert werden. Bei MprA und MprB umfasste das Signalpeptid die Aminosäuren 1 bis 27 (s. Tab. 13).

Tabelle 13: Aufstellung der identifizierten konservierten Domänen für MprA und MprB.

Die Suche nach konservierten Domänen erfolgte bei der NCBI mit Hilfe des Tools „Conserved Domain Search“. Aminoterminalen Signalpeptide wurden mit dem Programm SignalP 3.0 identifiziert.

Protein	Größe in AS	Konservierte Domänen mit E-Value	Aminosäure-Sequenzbereich
MprA	862	Signalpeptid	1 bis 27
		pfam03413, Pep_M4_propep, Thermolysin Metallopeptidase, Propeptid; $2e^{-24}$	70 bis 214
		pfam01447, Peptidase_M4, Thermolysin Metallopeptidase, katalytische Domäne; $5e^{-29}$	226 bis 371
		pfam02868, Peptidase_M4_C, Thermolysin Metallopeptidase, alpha-helikale-Domäne; $3e^{-30}$	373 bis 512
		pfam04151, PPC, bakterielle Propeptid C-terminale Domäne; $1e^{-21}$	546 bis 630
		pfam04151, PPC, bakterielle Propeptid C-terminale Domäne; $5e^{-21}$	656 bis 740
		COG4935; regulatorische P-Domäne von Subtilisin-ähnlichen Proprotein-Konvertasen und anderen Proteasen; $9e^{-11}$	757 bis 862
MprB	862	Signalpeptid	1 bis 27
		pfam03413, Pep_M4_propep, Thermolysin Metallopeptidase, Propeptid; $1e^{-21}$	67 bis 214
		pfam01447, Peptidase_M4, Thermolysin Metallopeptidase, katalytische-Domäne; $1e^{-26}$	226 bis 371
		pfam02868, Peptidase_M4_C, Thermolysin Metallopeptidase, alpha-helikale-Domäne; $4e^{-30}$	373 bis 503
		pfam04151, PPC, bakterielle Propeptid C-terminale Domäne; $2e^{-21}$	545 bis 626
		pfam04151, PPC, bakterielle Propeptid C-terminale Domäne; $4e^{-21}$	656 bis 740
		COG4935; regulatorische P-Domäne von Subtilisin-ähnlichen Proprotein-Konvertasen und anderen Proteasen; $1e^{-10}$	757 bis 862

Aufgrund der erhaltenen Ergebnisse konnten für die Proteasen MprA und MprB die in Abbildung 19 dargestellten Domänenstrukturen festgelegt werden. Ein Vergleich mit der Domänenstruktur der M4-Metalloprotease Empl aus *Pseudoalteromonas sp. A28* (LEE et al., 2002) zeigte, dass mit Ausnahme der regulatorischen P-Domäne, die sich jeweils am C-terminalen Ende von MprA und MprB befindet, die gleiche Domänenstruktur vorlag.

Abbildung 19: Vergleich der Domänenstrukturen der drei Metalloproteasen MprA, MprB (beide diese Arbeit) und Empl (LEE et al., 2002).

Eine weitere Gemeinsamkeit aller M4-Metalloproteasen ist ein innerhalb der katalytischen Domäne existierendes konserviertes Sequenz-Motiv. Dieses Motiv, welches auch als „Jongeneel Konsensus Motiv“ (JONGENEEL et al., 1989) bekannt ist, umfasst die Aminosäure-Signatur [GSTALIVN]-x(2)-H-E-[LIVMFYW]-{DEHRKP}-H-x-[LIVMFYWGSPQ]. Dabei fungieren die beiden Histidine als Zink-Liganden und das Glutamat bildet den katalytischen aktiven Rest. Metalloproteasen, die diese Signatur in ihrer katalytischen Domäne aufweisen, werden deswegen als Zink-abhängige Metalloproteasen bezeichnet (RAWLINGS und BARRETT, 1995). Die Aminosäuresequenzbereiche von MprA (AS 359 bis 368), MprB (AS 359 bis 368) und Empl (AS 344 bis 352) weisen diese Signatur mit einer Ausnahme vollständig auf (s. Abb. 20).

MprA: 359 V A G H E V S H G F 368

MprB: 359 V A G H E V S H G F 368

Empl: 344 V S A H E V S H G F 353

Signatur: [GSTALIVN]-x(2)-H-E-[LIVMFYW]-{DEHRKP}-H-x-[LIVMFYWGSPQ]

Abbildung 20: Darstellung der Aminosäuresequenzen von MprA, MprB (beide diese Arbeit) und Empl (LEE et al., 2002), die das Konsensus-Motiv für Zink-abhängige Metalloproteasen beinhalten.
Schwarz: variable und beliebige AS; Rot: festgelegte AS; Blau: AS, die nicht der Signatur entsprechen.

Die Aminosäuresequenzen von MprA, MprB und Empl entsprechen jeweils in der siebten Position (S₃₆₅, S₃₆₅ und S₃₅₀) der Signatur nicht dem Konsensus-Motiv. Es befindet sich an dieser Position bei allen drei Proteasen ein Serinrest.

Analysen nachgewiesen werden. Zusätzlich zur Klonierung der vollständigen *mprA*- und *mprB*-Gene erfolgte eine Klonierung von Teilfragmenten, die unterschiedliche Domänen der Enzyme umfassten, ebenfalls in den Expressionsvektor pET101/D-TOPO (s. Abb. 22). Auch hierbei wurde die gerichtete Klonierung so gewählt, dass alle Konstrukte mit C-terminalem V5-Epitop und His₆-Tag versehen waren. Alle für die Klonierung vorgesehenen DNA-Fragmente wurden mit Hilfe der *Pfu*-DNA-Polymerase amplifiziert (s. 2.11). In Tabelle 14 sind die zur Amplifizierung des jeweiligen Fragments eingesetzten Oligonukleotide, die verwendete Template-DNA und die Größen der amplifizierten PCR-Produkte aufgeführt. Alle DNA-Fragmente konnten mit der erwarteten Größe durch PCR amplifiziert werden. Die jeweiligen PCR-Produkte wurden isoliert, gereinigt (s. 2.5.2.6) und in den Expressionsvektor pET101/D-TOPO kloniert (2.3.4.1). Die dabei hergestellten Plasmide wurden zunächst in *E. coli* Top10-Zellen transformiert (s. 2.10.1) und die rekombinanten *E. coli*-Stämme auf LB-Ampicillin-Agar-Platten (s. 2.2.1) ausplattiert. Zur Analyse wurden pro Ansatz jeweils 10 Kolonien in 5 ml Kulturen (LB-Medium mit Ampicillin) angezogen und die Plasmide isoliert (s. 2.6.3.3). Die korrekte Orientierung der Inserts wurde durch Restriktionsanalysen (s. 2.8.1) in allen Fällen bestätigt.

Tabelle 14: Übersicht der für die Klonierung in pET101/D-TOPO mittels PCR amplifizierten *mprA*- und *mprB*-Genfragmente.

Die Sequenzen der verwendeten Oligonukleotide sind in Tabelle 4 aufgeführt. Eine genaue Darstellung der klonierten *mprA*- und *mprB*-Genfragmente ist Abbildung 23 zu entnehmen.

DNA-Fragment	Größe des Fragments	Verwendete Oligonukletide	Template-DNA
<i>mprA</i>	2593 bp	pET-MprA/K1-for; pET-MprA/MprB-rev	pTWP9
<i>mprA-K1</i>	1638 bp	pET-MprA/K1-for; pET-MprA-K1-rev	pTWP9
<i>mprB</i>	2593 bp	pET-MprB/K1-for; pET-MprA/MprB-rev	pTWP13
<i>mprB-K1</i>	1638 bp	pET-MprB/K1-for; pET-MprB-K1/K3-rev	pTWP13
<i>mprB-K2</i>	1941 bp	pET-MprB-K2/K3-for; pET-MprA/MprB-rev	pTWP13
<i>mprB-K3</i>	940 bp	pET-MprB-K2/K3-for; pET-MprB-K1/K3-rev	pTWP13

Anschließend wurden alle Plasmide zur Überprüfung der Gensequenzen und der korrekten Insertion der klonierten DNA-Fragmente durch das „Göttingen Genomics Laboratory“ sequenziert. Bei keinem Derivat lagen Abweichungen von der Originalsequenz vor und alle Insertionspositionen waren korrekt. In Abbildung 22 sind die in dem Expressionsvektor pET101/D-TOPO erzeugten Protease-Derivate schematisch dargestellt.

Abbildung 22: Übersicht der in pET101/D-TOPO erzeugten Protease-Gen-Derivate.

Die Derivate *mprA* und *mprB* enthalten jeweils den vollständigen Genbereich. Bei *mprA-K1* und *mprB-K1* wurden jeweils nur die Genregionen der Signalsequenz, des N-terminalen Propetids, der katalytischen Domäne und die Genregionen der alpha-helikalen Domäne kloniert. Der gesamte C-terminale Genbereich (C-terminales Propeptid und regulatorische P-Domäne) wurde deletiert. Dem Konstrukt *mprB-K2* fehlen die Genbereiche für die Signalsequenz und das N-terminale Propeptid. Eine sowohl N- als auch C-terminale Deletion ist bei dem Konstrukt *mprB-K3* zu verzeichnen, hier wurden nur die Genbereiche für die katalytische Domäne und die alpha-helikale Domäne kloniert.

3.4.2 Überprüfung der Funktionalität der hergestellten *mprA* und *mprB*-Gene und Gen-Derivate

Zur Überprüfung welche der konstruierten *mprA* und *mprB*-Derivate proteolytische Aktivität vermitteln wurden zunächst alle Plasmide in *E. coli* BL21-Zellen transformiert (s. 2.10.1). Von den Transformationsansätzen, erfolgte ein direktes Ausstreichen der rekombinanten *E. coli*-Stämme auf Skim-Milk-Agarplatten (s. 2.10.5), die zusätzlich IPTG für die Induktion der Genexpression beinhalten. Als Negativkontrolle wurde der pET101/D-TOPO Expressionsvektor ohne Insert in *E. coli* BL21-Zellen transformiert und ebenfalls auf die Indikatorplatten ausgestrichen. Als Positivkontrollen dienten die aus dem Screening hervorgegangen Klone *E. coli*/Top10/pTWP9 und *E. coli*/Top10/pTWP13. Eine deutliche Hofbildung durch hydrolytische Spaltung des Substrats Skim-Milk war bei den Positivkontrollen und den rekombinanten *E. coli*-Stämmen mit den Derivaten *mprA*, *mprB*, *mprA-K1* und *mprB-K1* zu verzeichnen, während die Derivate *mprB-K2* und *mprB-K3* und die Negativkontrolle, keine proteolytische Aktivität vermitteln konnten (s. Abb. 23).

Die Überprüfung der Fähigkeit zur hydrolytischen Spaltung der Substrate Azoalbumin, Azocasein und Elastin-Congo-Rot erfolgte ebenfalls auf Indikatoragar, der anstelle von Skim-Milk die jeweiligen Substrate enthielt (s. 2.17.1). Die Negativkontrolle sowie rekombinante *E. coli*-Stämme mit den Derivaten *mprB-K2* und *mprB-K3* konnten keines der getesteten Substrate hydrolytisch spalten. Die rekombinanten *E. coli*-Stämme mit den Derivaten *mprA*, *mprB*, *mprA-K1* und *mprB-K1* sowie rekombinante *E. coli*-Stämme mit den beiden Ursprungs-Plasmiden pTWP9 und pTWP13 zeigten eine deutliche Hofbildung auf Testplatten mit den Substraten Azoalbumin und Azocasein aber keine elastinolytische Aktivität auf Testplatten mit dem Substrat Elastin-Congo-Rot (s. Tab. 15).

Abbildung 23: Funktionalitätsüberprüfung von MprA, MprA-K1, MprB, MprB-K1, MprB-K2 und MprB-K3 auf Testagarplatten.

Die Plasmide pET-MprA, pET-MprA-K1, pET-MprB, pET-MprB-K1, pET-MprB-K2 und pET-MprB-K3 wurden wie unter 2.10.1 beschrieben in *E. coli* BL21-Zellen transformiert und auf Skim-Milk-Agarplatten ausgestrichen. Proteolytische Aktivität wird durch Hofbildung (Aufklärung) um die Kolonie angezeigt. Als Negativkontrollen wurden *E. coli*/BL21/pET101/D-TOPO und *E. coli*/Top10/pBluescript SK II und als Positivkontrollen die beiden rekombinanten Stämme aus dem Screening (*E. coli*/Top10/pTWP9 und *E. coli*/Top10/pTWP13) ausgestrichen. Die Platten wurden für 2 bis 4 Tage bei 37 °C inkubiert.

Tabelle 15: Nachweis der Funktionalität der unterschiedlichen Protease-Konstrukte mit verschiedenen Testsubstraten.

Alle Plasmide wurden in *E. coli* BL21-Zellen transformiert (s. 2.10.1) und auf Testagarplatten mit den unterschiedlichen Substraten ausgestrichen (s. 2.17.1). Proteolytische Aktivität konnte durch Aufklärung des Agars um den Impfausstrich detektiert werden.

Stamm	Getestetes Substrat			
	Skim-Milk	Azoalbumin	Azocasein	Elastin-Congo-Rot
<i>E. coli</i> /BL21/MprA	+	+	+	-
<i>E. coli</i> /BL21/MprA-K1	+	+	+	-
<i>E. coli</i> /BL21/MprB	+	+	+	-
<i>E. coli</i> /BL21/MprB-K1	+	+	+	-
<i>E. coli</i> /BL21/MprB-K2	-	-	-	-
<i>E. coli</i> /BL21/MprB-K3	-	-	-	-
<i>E. coli</i> /Top10/pTWP9	+	+	+	-
<i>E. coli</i> /Top10/pTWP13	+	+	+	-
Negativkontrolle	-	-	-	-

3.4.3 Heterologe Expression der *mprA* und *mprB*-Gene und deren Derivate in *E. coli* BL21

Für die heterologe Expression von *E. coli*/BL21/MprA, *E. coli*/BL21/MprA-K1, *E. coli*/BL21/MprA-K2, *E. coli*/BL21/MprA-K3, *E. coli*/BL21/MprB und *E. coli*/BL21/MprB-K1 mittels des T₇-Polymerase Promotor-Systems wurden zunächst alle Plasmide in *E. coli* BL21-Zellen transformiert (s. 2.10.1). Mit jedem Transformationsansatz wurden 25 ml Protease-Expressionsmedium mit Ampicillin (s. 2.2.6) angeimpft und über Nacht inkubiert. Die Anzucht der Hauptkulturen (4 x 500 ml) erfolgte aus den Vorkulturen (4 x 25 ml) bis zu einer OD von ca. 0,8 bei 37 °C auf einem Schüttler. Danach wurde die Expression der unterschiedlichen Protease-Derivate durch Zugabe von 0,5 mM IPTG induziert und die Expressionskulturen für 12 bis 15 h bei 30 °C weiter inkubiert. Als Negativkontrolle wurde *E. coli*/BL21/pET101/D-TOPO verwendet. Die

Zellernte erfolgte durch Zentrifugation. Die Zellpellets wurden in 50 mM Tris-Puffer (pH 8,0), der mit 0,5 mM ZnCl_2 versetzt war, suspendiert und die Zellen mit der French-Presse (s. 2.13.1) aufgeschlossen. Alle Kulturüberstände der Expressionskulturen wurden zunächst gefiltert, auf ein Volumen von 30 bis 50 ml eingeeengt und mit 50 mM Tris-Puffer (pH 8,0), der mit 0,5 mM ZnCl_2 versetzt war, umgepuffert (s. 2.13.6.2). Im Anschluss erfolgte die Bestimmung der spezifischen proteolytischen Aktivität der Rohextrakte sowie der konzentrierten und umgepufferten Kulturüberstände aus den verschiedenen Expressionskulturen (s. 2.16.1).

In den Rohextrakten von *E. coli*/BL21/MprA, *E. coli*/BL21/MprA-K1, *E. coli*/BL21/MprA-K2, *E. coli*/BL21/MprA-K3, *E. coli*/BL21/MprB und *E. coli*/BL21/MprB-K1 sowie im Rohextrakt der Negativkontrolle konnten fast die gleichen spezifischen proteolytischen Aktivitäten (0,01 bis 0,015 U/mg) ermittelt werden. Diese proteolytischen Aktivitäten sind auf die wirtseigenen intrazellulären Proteasen von *E. coli* BL21 zurückzuführen, da keiner der rekombinanten *E. coli*-Stämme im Vergleich zur Negativkontrolle eine gesteigerte proteolytische Aktivität gezeigt hat. Ein anderes Ergebnis zeigten allerdings die spezifischen Aktivitäten der konzentrierten Kulturüberstände der rekombinanten *E. coli*-Stämme, die in Tabelle 16 aufgeführt sind. In den Kulturüberständen von *E. coli*/BL21/MprA-K2, *E. coli*/BL21/MprA-K3, beide mit Deletion im Genbereich der für die Signalsequenz und das aminoterminalen Propeptid kodiert, konnten im Vergleich zum Kulturüberstand der Negativkontrolle keine gesteigerten spezifischen proteolytischen Aktivitäten ermittelt werden. Dagegen zeigten die Kulturüberstände von *E. coli* BL21/MprA und *E. coli*/BL21/MprB mit vollständigem *mprA* bzw. *mprB*-Gen sowie *E. coli*/BL21/MprA-K1 und *E. coli*/BL21/MprB-K1 mit Deletionen im Genbereich für die C-terminalen Propeptide und die regulatorische P-Domäne eine gegenüber der Negativkontrolle um das 55 bis 80 fache gesteigerte spezifische proteolytische Aktivität. Die höchsten spezifischen Aktivitäten konnten für *E. coli*/BL21/MprA (0,16 U/mg) und *E. coli*/BL21/MprB (0,14 U/mg) ermittelt werden. *E. coli*/BL21/MprA-K1 und *E. coli*/BL21/MprB-K1 zeigten etwas geringere spezifische Aktivitäten (0,11 U/mg bzw. 0,12 U/mg).

Tabelle 16: Spezifische proteolytische Aktivitäten der Proetase-Derivate im konzentrierten Kulturüberstand.

Die aus den Expressionskulturen durch Abzentrifugation der Zellen erhaltenen Kulturüberstände wurden konzentriert, umgepuffert (s. 2.13.6.2) und die spezifischen proteolytischen Aktivitäten ermittelt (s. 2.16.1)

Stamm	Spezifische proteolytische Aktivität [U/mg]
<i>E. coli</i> /BL21/MprA	0,16
<i>E. coli</i> /BL21/MprA-K1	0,11
<i>E. coli</i> /BL21/MprB	0,14
<i>E. coli</i> /BL21/MprB-K1	0,12
<i>E. coli</i> /BL21/MprB-K2	0,001
<i>E. coli</i> /BL21/MprB-K3	0,002
Negativkontrolle	0,002

Die erhaltenen Ergebnisse belegten, dass die *mprA* und *mprB*-Gene für extrazelluläre Proteasen kodieren, die auch vom Wirtstamm *E. coli* BL21 bei heterologer Produktion in das Medium sekretiert werden. Ferner ist das aminoterminal Propeptid, das meistens autokatalytisch abgespalten wird, essentiell für die Entstehung des reifen und aktiven Proteins, da alle rekombinanten *E. coli*-Stämme mit einem Plasmid ohne den Genbereich, der für das aminoterminal Propeptid kodiert, keine proteolytische Aktivität zeigten. Dies stimmt auch mit der Eingruppierung von MprA und MprB in die Metalloprotease-Familie M4 überein, da das aminoterminal Propeptid bei Mitgliedern dieser Familie als intramolekulares Chaperon fungiert und somit bei der Faltung der aktiven Protease behilflich ist (O' DONOHUE und BEAUMONT, 1996; SHINDE und INHOUYE, 2000). Bei der extrazellulären Metalloprotease Pseudolysin aus *Pseudomonas aeruginosa*, wird angenommen, dass das aminoterminal Propeptid, nachdem im Periplasma die Abspaltung erfolgt ist mit dem restlichen Peptid assoziiert („Pro-Mature Complex“). Somit wird eine Inhibition der proteolytischen Aktivität erreicht, bis die Protease in das Medium sekretiert wird. Dort dissoziiert das aminoterminal Propeptid von der nun vollständig gereiften Protease ab, die somit aktiv wird (KESSLER und SAFRIN, 1994; BRAUN et al., 1998; KESSLER et al., 1998).

Durch Western-Blot-Analyse und nachfolgend durchgeführte Immunodetektion mit einem spezifischen Antikörper gegen das in allen MprA und MprB-Derivaten angehängte C-terminale V5-Epitop wurde überprüft, ob alle rekombinanten Proteine von dem Wirtstamm *E. coli* BL21 produziert werden. Der Rohextrakt aus jeder Expressionskultur sowie der konzentrierte Kulturüberstand wurden elektrophoretisch im SDS-Polyacrylamidgel aufgetrennt (s. 2.13.3), die Gele anschließend auf Nitrocellulose-Membranen geblottet (s. 2.14) und die Immunodetektion mit dem Anti-V5-AP-Antikörper, wie unter 2.14.1 beschrieben, durchgeführt (s. Abb. 24).

Abbildung 24: Nachweis von MprA, MprA-K1, MprB, MprB-K1, MprB-K2 und MprB-K3 aus Rohextrakten durch Immunodetektion.
Die Rohextrakte wurden durch SDS-PAGE augetrennt (s. 2.13.3), auf Nitrocellulose-Membranen geblottet (s. 2.14) und eine Immunodetektion mit dem Anti-V5-AP-Antikörper durchgeführt (s. 2.14.1).

Ein Nachweis für die heterologe Produktion der unterschiedlichen Protease-Derivate MprA, MprA-K1, MprB, MprB-K1, MprB-K2 und MprB-K3 mit dem Anti-V5-AP-Antikörper konnte nur in den entsprechenden Rohextrakten erfolgen (s. Abb. 24). In den konzentrierten Kulturüberständen war eine Detektion nicht möglich. Das lässt darauf schließen, dass die verschiedenen Konstrukte am C-terminalen Ende weiteren Prozessierungen unterliegen und die abgespaltenen Fragmente im Anschluss degradiert werden. Im Weiteren wurden in das Medium sekretierte Protease-Derivate aus dem konzentrierten Kulturüberstand gereinigt und biochemisch charakterisiert.

3.4.4 Reinigung von MprA, MprA-K1, MprB und MprB-K1 aus konzentrierten Kulturüberständen über Metall-Chelat-Affinitätschromatographie

Die zuvor dargestellte Western-Blot-Analyse (s. Abb 24, 3.4.3) hatte gezeigt, dass die Protease-Derivate MprA, MprA-K1, MprB und MprB-K1 in den konzentrierten Kulturüberständen nicht nachzuweisen waren. Im C-terminalen Bereich der Protease-Derivate befand sich ebenfalls ein His₆-Tag (s. Abb. 22), der eine Reinigung über Metall-Chelat-Affinitätschromatographie an Ni-NTA-Agarose ermöglichen sollte (s. 2.15.1). Im Anschluss an die Chromatographie erfolgte die Bestimmung der proteolytischen Aktivität von allen erhaltenen Fraktionen (Durchfluss, Waschen 1, Waschen 2 und Elution) sowie von den jeweiligen Ausgangs-Proteinproben über die Zymographie (s. 2.171). In Abbildung 25 wird stellvertretend für alle Reinigungsversuche das Ergebnis der Affinitätschromatographie für MprA gezeigt. Deutlich zu erkennen ist, dass eine Reinigung von MprA über Metall-Chelat-Affinitätschromatographie nicht erreicht werden konnte, da sich der Hauptanteil des Proteins schon im Durchfluss der Säule befand und somit nicht an die Ni-NTA-Agarose gebunden hatte. Weitere proteolytische Aktivitäten, die immer schwächer wurden, waren in den beiden Waschfraktionen zu verzeichnen. Keine proteolytische Aktivität konnte dagegen in der Elutionsfraktion nachgewiesen werden. Somit konnte eine Reinigung von MprA, MprA-K1, MprB und MprB-K1 über Metall-Chelat-Affinitätschromatographie ebenfalls nicht erfolgreich durchgeführt werden. Diese Ergebnisse untermauern die Vermutung, dass die verschiedenen Protease-Derivate an ihren C-terminalen Enden prozessiert werden.

Abbildung 25: Zymogramm des Reinigungsverlaufs von MprA über Metall-Chelat-Affinitätschromatographie.

Der Reinigungsversuch von MprA erfolgte über Metall-Chelat-Affinitätschromatographie an Ni-NTA-Agarose (s. 2.15.1). Im Anschluss an die Reinigung wurde die proteolytische Aktivität der gesammelten Fraktionen über die Zymographie in einem SDS-Casein-Gel nachgewiesen (s. 2.17.1). P: Proteolytische Aktivität im konzentrierten Kulturüberstand; D: Proteolytische Aktivität in der Durchfluss-Fraktion; W1: Proteolytische Aktivität der ersten Waschfraktion mit Puffer P1; W2: Proteolytische Aktivität der zweiten Waschfraktion mit Puffer P2; E: Proteolytische Aktivität der Elutionsfraktion mit Puffer P3.

3.4.5 Reinigung von MprA, MprA-K1, MprB und MprB-K1 aus konzentrierten Kulturüberständen über hydrophobe Interaktionschromatographie (HIC)

Das Trennprinzip der HIC basiert auf der unterschiedlichen Oberflächenhydrophobizität der einzelnen Proteine. Zunächst wurden alle Proben auf ein Milieu von hoher Ionenstärke eingestellt, indem sie vorsichtig auf 1 M Ammoniumsulfat hochgesalzen wurden. Danach erfolgte die Reinigung der Protease-Derivate mit einer Äkta-FPLC-Anlage über eine Phenyl Sepharose HP XK 16/10 Fertigsäule (s. 2.15.2). Bei allen Säulenläufen wurden je 40 Fraktionen gesammelt. Nach der Chromatographie wurde von jeder Fraktion die proteolytische Aktivität ermittelt (s. 2.16.1). Die Chromatogramme der verschiedenen Säulenläufe zur Reinigung der Protease-Derivate sind in Abbildung 26 gezeigt.

Abbildung 26: Chromatogramme der HIC-Säulenläufe der Proben MprA, MprA-K1, MprB und MprB-K1. Von jeder Fraktion wurde nach den Säulenläufen die proteolytische Aktivität bestimmt (s. 2.16.1).

Proteolytische Aktivitäten konnten bei den Proben MprA in den Fraktionen 30 bis 35, bei MprA-K1 in den Fraktionen 31 bis 35, bei MprB in den Fraktionen 30 bis 34 und bei MprB-K1 in den Fraktionen 30 bis 35 ermittelt werden (s. Abb. 26). Die Fraktionen mit den höchsten proteolytischen Aktivitäten wurden von jeder Probe gepoolt und die spezifischen Aktivitäten der gepoolten Fraktionen bestimmt (s. Tab. 17). Die höchste spezifische Aktivität zeigte die gepoolte Fraktion der Probe MprA mit 1,51 U/mg. Etwas geringere spezifische Aktivitäten, wiesen die gepoolten Fraktionen der Probe MprB (1,28 U/mg) und MprA-K1 (1,19 U/mg) auf. Die niedrigste spezifische Aktivität wurde für die Probe MprB-K1 mit 0,85 U/mg ermittelt (s. Tab. 17).

Tabelle 17: Spezifische proteolytische Aktivitäten der gepoolten Fraktionen aus den verschiedenen Säulenläufen der HIC.

Die aus der HIC hervorgegangenen Fraktionen mit den jeweils höchsten proteolytischen Aktivitäten wurden von jeder Probe gepoolt und die spezifische Aktivität erneut ermittelt.

Probe	Fraktionen	Proteolytische Aktivität [U/mg]
MprA	31 bis 33	1,51
MprA-K1	32 bis 34	1,19
MprB	31 bis 34	1,28
MprB-K1	30 bis 33	0,85

Die spezifischen proteolytischen Aktivitäten der verschiedenen Proben waren nach den Reinigungen über hydrophobe Interaktionschromatographie (s. Tab. 17) deutlich höher als in konzentrierten Kulturüberständen (s. Tab. 16). Das bedeutet, dass alle sekretierten Protease-Derivate über die HIC angereichert werden konnten. Für jede Probe wurden die Anreicherungsfaktoren und die Ausbeuten berechnet (s. Tab. 18).

Tabelle 18: Anreicherung von MprA, MprA-K1, MprB und MprB-K1 aus konzentrierten Kulturüberständen über hydrophobe Interaktionschromatographie (HIC).

Die Kulturüberstände der verschiedenen Expressionskulturen wurden konzentriert und die proteolytischen Aktivitäten bestimmt. Nach Reinigung über HIC wurde von jeder erhaltenen Fraktion die proteolytische Aktivität ermittelt. Fraktionen mit hohen proteolytischen Aktivitäten wurden gepoolt und erneut die spezifische Aktivität bestimmt.

Probe/ Anreicherungs- schritt	Gesamt- Volumen [ml]	Gesamt- Protein [mg]	Gesamt- Aktivität [U]	spez. Aktivität [U/mg]	Ausbeute [%]	Anreicherungs- faktor
<u>MprA:</u> konz. Kulturüberstand HIC	35 18	13,65 0,72	2,18 1,09	0,16 1,51	100 50	1 9,4
<u>MprA-K1:</u> konz. Kulturüberstand HIC	38 18	14,44 1,08	1,59 1,29	0,11 1,19	100 81	1 8,5
<u>MprB:</u> konz. Kulturüberstand HIC	50 24	16,5 1,44	2,31 1,84	0,14 1,28	100 80	1 9,1
<u>MprB-K1:</u> konz. Kulturüberstand HIC	40 24	15,6 2,4	3,12 2,04	0,12 0,85	100 65	1 7,1

Der größte Anreicherungsfaktor konnte für die Probe MprA mit 9,4 erzielt werden. Die anderen proteolytisch aktiven Derivate wurden durch hydrophobe Interaktionschromatographie 7,1-fach (MprB-K1) bis 9,1-fach (MprB) angereichert. Die Ausbeuten der angereicherten Proben lagen zwischen 50 % bis 81 %. Durch SDS-Page (s. 2.13.3) wurde überprüft, ob die unterschiedlichen Konstrukte gereinigt werden konnten. In Abbildung 27 wird am Beispiel von der Probe MprB die erzielte Reinigung über HIC dargestellt.

Abbildung 27: Darstellung der Reinigung von MprB durch SDS-PAGE.

Spur 1: konzentrierter Kulturüberstand der Probe MprB. Spur 2: gepoolte Fraktionen des Säulenlaufs der Probe MprB. M: Molekularmassenmarker in kDa. Das 12 %ige SDS-Gel wurde nach der Elektrophorese einer Silberfärbung (s. 2.13.4.2) unterzogen. Der Pfeil kennzeichnet eine angereicherte Proteinbande.

Eine vollständige Reinigung der Probe MprB über hydrophobe Interaktionschromatographie konnte nicht erreicht werden, da in Spur 2 (s. Abb. 27) noch sehr viele Proteinbanden zu erkennen sind. Ein Vergleich der Spuren 1 und 2 zeigte jedoch, dass eine partielle Reinigung von MprB vorlag, da im konzentrierten Kulturüberstand (Spur 1) deutlich mehr Proteinbanden als nach der HIC-Reinigung (Spur 2) zu erkennen sind. Weiterhin ist in Spur 2 eine angereicherte Proteinbande mit einer Molekularmasse von ca. 33 kDa zu erkennen. Diese entspricht der aus der Gensequenz von *mprB* abgeleiteten Molekularmasse der Protease nach Prozessierung der Signalsequenz, des N-terminalen und des C-terminalen Propeptids (34 kDa). Ähnliche Ergebnisse wurden bei der SDS-PAGE von MprA, MPrA-K1 und MPrB-K1 erzielt. Es lag in allen Fällen eine partielle Reinigung vor.

Die durch HIC partiell gereinigten Protease-Derivate wurden ebenfalls durch Zymographie (s. 2.16.1) analysiert (s. Abb. 28).

Abbildung 28: Zymographie der partiell gereinigten Proteine MprA, MprA-K1, MprB und MprB-K1

Die Proben wurden in einem 12 %igen SDS-Casein-Gel elektrophoretisch aufgetrennt. Zur Entfernung des SDS erfolgte eine Inkubation in Triton X 100-Lösung für 30 min. Nach Inkubation für 2 bis 8 h in 50 mM KP-Puffer (mit 0,5 mM ZnCl_2) wurde das Gel mit Coomassie-Färbelösung angefärbt und anschließend so lange entfärbt, bis die Aktivitätsbanden deutlich sichtbar waren. Als Vergleich wurde auch die Negativkontrolle mit aufgetrennt.

Bei allen Protease-Derivaten sind aufgehellte Banden im SDS-Casein-Gel zu erkennen. Die Negativkontrolle zeigte keine solche Bande (s. Abb. 28). An den aufgehellten Bereichen wurde durch die proteolytische Aktivität der jeweiligen Proben die Spaltung des copolymerisierten Substrats Casein hervorgerufen. Das Auftreten von mehreren Banden pro Spur ist nicht untypisch und wurde schon oft für andere Proteasen beschrieben (KEIL-DLOUHA, 1976; LACKS und SPRINGHORN, 1980). Es wird angenommen, dass das Auftreten von

mehreren Banden in einer Spur entweder durch Aggregatbildungen der Protease mit dem copolymerisierten Substrat hervorgerufen wird oder durch Selbstverdau der Proteasen aktive Teilpeptide mit unterschiedlicher Molekularmasse entstehen können. Durch Zymographie konnte bei allen Protease-Derivaten auch die Fähigkeit zur hydrolytischen Spaltung der Substrate Gelatine und Rinderserumalbumin (BSA) gezeigt werden (Daten nicht gezeigt).

Für weitere Reinigungsschritte wurde die Bindung der partiell gereinigten Protease-Derivate an verschiedene Anionen- und Kationenaustauscher- Säulen mit Hilfe des „Hi Trap IX Selection Kit“ getestet (s. 2.15.3). Mit einer peristaltischen Pumpe wurde je 1 ml der gepoolten Fraktionen von jeder Probe aus den HIC-Läufen auf die Säulen aufgetragen. Die Wasch- und Elutionsschritte erfolgten nach den Angaben des Herstellers. Von allen gesammelten Fraktionen (Durchfluss, Waschen und Elution) wurde die proteolytische Aktivität bestimmt (s. 2.16.1). Nur in den Durchfluss-Fraktionen konnte proteolytische Aktivität ermittelt werden. Demnach erfolgte keine Bindung der Proteasen an die Säulenmaterialien. Somit war es nicht möglich die Protease-Derivate im Rahmen dieser Arbeit über Anionen- bzw. Kationenaustausch-Chromatographie weiter zu reinigen. Für die im Folgenden beschriebene Charakterisierung wurden daher die gepoolten Fraktionen mit den partiell gereinigten Proteasen verwendet.

3.4.6 Bestimmung der proteolytischen Aktivität in Abhängigkeit von der Temperatur

Für die Bestimmung der Temperaturoptima von MprA, MprA-K1, MprB und MprB-K1 wurden die Enzymlösungen für 1 h bei Temperaturen von 15 °C bis 85 °C in Heizblöcken inkubiert (s. 2.17.2). Im Anschluss wurden die proteolytischen Aktivitäten der verschiedenen Ansätze bestimmt. Die für die jeweiligen Proteasen ermittelten Temperaturkurven sind in Abbildung 29

dargestellt. Es wurden jeweils die relativen Aktivitäten gegen die entsprechenden Temperaturen aufgetragen. Die höchste gemessene proteolytische Aktivität für jedes Protease-Derivat wurde als 100 % festgesetzt.

Abbildung 29: Abhängigkeit der proteolytischen Aktivität von der Temperatur bei MprA, MprA-K1, MprB und MprB-K1.

Die Enzymlösungen wurden für 1 h bei unterschiedlichen Temperaturen inkubiert und im Anschluss die proteolytischen Aktivitäten ermittelt. 100 % relative Aktivität entsprechen dabei nachfolgend aufgeführten proteolytischen Aktivitäten. MprA: 1,48 U/mg; MprA-K1: 1,23 U/mg; MprB: 1,21 U/mg; MprB-K1: 1,01 U/mg.

Maximale spezifische Aktivitäten konnten für MprA (1,48 U/mg) und MprA-K1 (1,23 U/mg) bei 65 °C ermittelt werden. MprB (1,21 U/mg) und MprB-K1 (1,01 U/mg) haben dagegen das Temperaturoptimum ihrer Proteolyse-Aktivität bei 60 °C. Auffällig ist, dass MprB im Temperaturbereich zwischen 45 °C und 65 °C nicht unter 94 % im Vergleich zur höchsten Aktivität fällt. MprA zeigt eine ähnlich hohe relative Aktivität (mindestens 86 %; 1,27 U/mg) zwischen 45 °C und 65 °C. Die Protease-Derivate MprA-K1 und MprB-K1, beide mit C-terminalen Deletionen, weisen einen schmaleren Temperaturbereich auf, in dem die relativen Aktivitäten nicht unter 82 % fallen. Das war zwischen 55 °C

und 65 °C der Fall. Bei 75 °C zeigen alle Proben einen deutlichen Abfall und erreichen maximal 15 % ihrer höchsten Aktivitäten. Im Temperaturbereich von 15 °C bis 35 °C konnte für MprA die höchste proteolytische Aktivität im Vergleich zu den anderen getesteten Proteasen ermittelt werden.

3.4.7 Bestimmung der Thermostabilität

Zur Ermittlung der Temperaturstabilität von MprA, MprA-K1, MprB und MprB-K1 wurden alle Proben bei drei unterschiedlichen Temperaturen (50, 60 und 70 °C) für eine Zeitdauer von 3 h inkubiert. Die Bestimmung der proteolytischen Aktivität aller Proben erfolgte beim Startpunkt, nach 30, 60, 120 und 180 min Inkubation (s. 2.17.4). Die Diagramme zur Thermostabilität sind in Abbildung 30 gezeigt. Die Protease-Aktivitäten werden als relative Aktivitäten dargestellt, wobei die Ausgangsaktivitäten als 100 % festgesetzt wurden. Alle Proteasen zeigten bei 70 °C eine schnelle Inaktivierung, da dort schon nach 30 min fast keine proteolytischen Aktivität mehr messbar war. Bei 60 °C nahmen die Protease-Aktivitäten im Verlauf von 3 h langsam ab. Nach 180 min zeigten die Proben MprA und MprA-K1 noch 54 % bzw. 51 % der Ausgangsaktivität, während bei MprB und MprB-K1 nur noch 31 % bzw. 37 % der Ausgangsaktivität vorlagen. Bei 50 °C wiesen die C-terminal deletierten Proben MprA-K1 und MprB-K1 die größte Thermostabilität auf, da nach 180 min immer noch 81 % bzw. 86 % der Ausgangsaktivität erreicht wurden. Die Proteasen ohne C-terminale-Deletionen MprA und MprB zeigten nach 3 h bei 50 °C noch 67 % bzw. 60 % der Ausgangsaktivität. Erwähnenswert ist, dass bei allen Protease-Derivaten bei Lagerung für mehrere Wochen bei 4°C keine Aktivitätsverluste zu verzeichnen waren. Somit zeichnen sich die untersuchten Enzyme durch eine hohe Stabilität in einem weiten Temperaturbereich (4 °C bis 60 °C) aus.

Abbildung 30: Bestimmung der Thermostabilität von MprA, MprA-K1, MprB und MprB-K1 bei 50, 60 und 70 °C.

Die Enzymlösungen wurde jeweils für 180 min bei den drei unterschiedlichen Temperaturen inkubiert. Die proteolytische Aktivität bzw. Restaktivität wurden jeweils zum Zeitpunkt 0, nach 30, 60, 120 und 180 min für alle Temperaturen ermittelt. 100 % relative Aktivität (Zeitpunkt: 0) entsprechen dabei nachfolgend aufgeführten spezifischen Aktivitäten. MprA: 1,56 U/mg; MprA-K1: 1,14 U/mg; MprB: 1,32 U/mg; MprB-K1: 0,89 U/mg.

3.4.8 Bestimmung der proteolytischen Aktivität in Abhängigkeit vom pH-Wert

Die Bestimmung der proteolytischen Aktivitäten aller Protease-Derivate in Abhängigkeit vom pH-Wert erfolgte wie unter 2.17.3 beschrieben. Die pH-Spektren von MprA, MprA-K1, MprB und MprB-K1 im Bereich von pH 6 bis pH 11 sind in Abbildung 31 gezeigt. Die aufgeführten Aktivitäten werden erneut als relative Aktivitäten wiedergegeben. Die jeweils höchste gemessene Aktivität jeder Probe wurde als 100 % festgesetzt.

Abbildung 31: Abhängigkeit der proteolytischen Aktivität vom pH-Wert.

Die Enzymlösungen wurden mit Casein-Lösungen, die in Puffern mit unterschiedlichen pH-Werten angesetzt waren, für 1 h inkubiert und anschließend die proteolytischen Aktivitäten ermittelt. 100 % relative Aktivität entsprechen dabei nachfolgend aufgeführten spezifischen Aktivitäten. MprA: 1,58 U/mg; MprA-K1: 1,19 U/mg; MprB: 1,36 U/mg; MprB-K1: 0,88 U/mg.

Alle Proben zeigten ihre höchsten proteolytischen Aktivitäten bei pH 8. Es konnte über den gesamten pH-Bereich Aktivität festgestellt werden, jedoch betrugen die relativen Aktivitäten von MprA, MprA-K1 und MprB-K1 bei pH 6 nur 6 % (MprA) bzw. 3 % (MprA-K1 und MprB-K1) der höchsten Aktivität bei pH 8. Dagegen zeigte die Probe MprB bei pH 6 eine spezifische Aktivität von 0,3 U/mg, welches 22 % der höchsten spezifischen Aktivität bei pH 8 ausmachte. Von allen Protease-Derivaten konnte für MprB das weiteste pH-Spektrum festgestellt werden. Zwischen pH 6,5 und pH 10 fiel die relative Aktivität nicht unter 65 % (pH 10). Die Probe MprA zeigte einen ähnlichen Verlauf wie MprB, jedoch sank hier die relative Aktivität bei pH 10 auf 55 % ab. Erstaunlicherweise betrug die relative Aktivität von MprA bei pH 11 sogar noch 30 % (0,47 U/mg). Jeweils über 50 % der relativen Aktivität erreichten die Proben MprA-K1 und MprB-K1 im pH-Bereich von 6,5 bis 9. Bei pH 11 zeigten beide nur noch 18 % bzw. 12 % ihrer höchsten Aktivität. Alle Proben weisen somit ein breites pH-Spektrum auf, wobei jedoch die höchsten Aktivitäten im alkalischen Bereich festgestellt werden konnten.

3.4.9 Einfluss von Inhibitoren auf die proteolytische Aktivität

Zur Untersuchung der Wirkung von Inhibitoren auf die Protease-Aktivitäten von MprA, MprA-K1, MprB und MprB-K1 wurden die Enzymlösungen für 1 h mit den Protease-Inhibitoren EDTA, 1,10-Phenanthrolin und PMSF inkubiert (s. 2.17.5) und anschließend die Restaktivität ermittelt (s. 2.17.1). In Abbildung 32 werden die Ergebnisse dieser Untersuchungen dargestellt.

Abbildung 32: Einfluss von Inhibitoren auf die Protease-Aktivität von MprA, MprA-K1, MprB und MprB-K1.

Die Enzymlösungen wurden für 1 h mit unterschiedlichen Inhibitoren inkubiert und anschließend die Restaktivität ermittelt. Als Vergleich dienten Enzymlösungen ohne Behandlung. Die Werte ohne Inhibitoren wurden als 100 % (a) festgesetzt. 100 % relative Aktivität entsprechen dabei nachfolgend aufgeführten spezifischen Aktivitäten. MprA: 1,61 U/mg; MprA-K1: 1,35 U/mg; MprB: 1,48 U/mg; MprB-K1: 1,08 U/mg. Folgende Inhibitoren wurden eingesetzt. (b): 10 mM EDTA, (c): 10mM 1,10-Phenantrolin und (d): 10 mM PMSF.

Die Zugabe von 10 mM EDTA oder 1,10-Phenantrolin zu den Enzymlösungen führte zur vollständigen Inaktivierung der proteolytischen Aktivität bei allen Enzymen. Sowohl EDTA als auch 1,10-Phenantrolin sind aufgrund ihrer chelatierenden Eigenschaften typische Inhibitoren von Metalloproteasen (WINDLE und KELLEHER, 1997; NEGRETE-ABASCAL et al., 2003; KOOL et al., 2005). Die Zugabe von PMSF zu den Enzymlösungen hatte dagegen keinen signifikanten Einfluss auf die Aktivität der Proteasen. Die Restaktivitäten nach 1-stündiger Inkubation mit PMSF betrugen bei den Proben MprA, MprA-K1 und MprB noch zwischen 82 % und 84 % der Vergleichsaktivitäten. Bei MprB-K1 konnte sogar noch 100 % der Ausgangsaktivität verzeichnet werden. Die erzielten Ergebnisse bestätigen die Eingruppierung von MprA und MprB in die

Familie der Zink-abhängigen Metalloproteasen, welches auch mit den in der Literatur beschriebenen Eigenschaften zu Metalloproteasen übereinstimmt. Dort konnte auch festgestellt werden, dass der Serin-Protease-Inhibitor PMSF fast keinen Einfluss auf die Aktivität von Metalloproteasen ausübt, während die Inkubation mit chelatierenden Agenzien wie EDTA oder 1,10-Phenantrolin zur vollständigen Inaktivierung führt (MIYAMOTO et al., 2002; LEE et al., 2002).

4 Diskussion

In den vergangenen Jahren wurde es durch kultivierungsunabhängige metagenomische Analysen ermöglicht, die natürlich vorkommende mikrobielle Diversität als nahezu unerschöpfliche Ressource für neuartige Biokatalysatoren und Wirkstoffe nutzbar zu machen. Dazu wird die gesamte mikrobielle Umwelt-DNA aus einer Standortprobe isoliert und direkt kloniert. Die Hinterlegung der auf diesem Weg erhaltenen genetischen Informationen erfolgt in komplexen Genbanken so genannten Metagenombanken. Diese können in heterologen Wirten (meistens *Escherichia coli*) beliebig oft vermehrt und mit sequenzabhängigen oder auf biologischer Aktivität basierenden Verfahren analysiert und sowohl für die Anwendung als auch für die Forschung genutzt werden (STREIT et al., 2004; DANIEL, 2004; HANDELSMAN, 2004 SCHMITZ und STREIT 2004). Im Rahmen dieser Arbeit wurden verschiedene Metagenombanken konstruiert und zusammen mit bereits vorhandenen aktivitätsbasierend auf Gene durchmustert, die Butanol-Dehydrogenase-Aktivität oder proteolytische Aktivität vermitteln. Die für die entsprechenden Aktivitäten verantwortlichen Gene wurden anschließend identifiziert und charakterisiert.

4.1 Konstruktion von Metagenombanken

Eine entscheidende Voraussetzung für die Konstruktion von komplexen Metagenombanken ist die Isolierung von intakter, hochmolekularer und klonierbarer DNA. Inzwischen gibt es eine große Anzahl an Publikationen zur Isolierung von DNA aus Umweltproben und auch kommerzielle Isolierungs-Kits sind erhältlich (DANIEL 2005). Grundsätzlich unterscheidet man zwischen zwei verschiedenen Strategien der DNA-Isolierung. Die eine Variante basiert auf einer vorherigen Abtrennung der Mikroorganismen von umgebenden Matrixpartikeln (Zellseparation) und anschließender Lyse der isolierten Zellen (TORSVIK, 1980; JAKOBSEN und RASMUSSEN, 1992; BAKKEN und

LINDAHL, 1995; CURTOIS et al., 2001; GABOR et al., 2003). Die Trennung der Zellen von Matrixbestandteilen kann auf unterschiedliche Weise erfolgen. Häufig werden diverse Filtersysteme mit definierter Porengröße dafür verwendet (SOMMERVILLE et al., 1989) oder die Probe wird zunächst in einem Puffer homogenisiert. Die gelösten Matrixbestandteile können im Anschluss durch einen niedrigtourigen Zentrifugationsschritt entfernt werden (HOLBEN, 1996). Nach Entfernung der Matrixsubstanzen werden die nicht an diese Substanzen gebundenen Zellen durch hochtourige Zentrifugation sedimentiert und lysiert. Anschließend erfolgt die Isolierung der DNA (HOLBEN, 1996). Die Methode der vorherigen Zellseparation bietet sich vor allem bei Umweltproben mit geringer Biomasse an (z.B. Wasserproben), weil dadurch eine Konzentrierung der Probe erreicht wird (GIOVANNONI et al., 1990). Ferner ist diese Strategie vorteilhaft bei Proben, die mit Fremdsubstanzen (z.B. Klärwerksproben) verunreinigt sind, da diese für die nachfolgende Klonierung der DNA störend sind (DANIEL, 2002). Ein Vorteil bei auf Filtration basierender Zellseparation besteht darin, dass über die Porengröße der verwendeten Filter eine Vorselektion in Bezug auf den Ursprung der zu isolierenden DNA (eukaryotisch, prokaryotisch oder viral), vorgenommen werden kann (BREITBART et al., 2002; VENTER et al., 2004).

Die zweite prinzipielle Methode der DNA-Isolierung aus Umweltproben beruht auf der direkten Lyse der Mikroorganismen in der Probe. Dabei wird auch die frei in der Probe vorliegende DNA miterfasst (TEBBE und VAHJEN, 1993; WIKSTROM et al., 1996; ZHOU et al., 1996; SANTOSA, 2001; KNIETSCH et al., 2003a und 2003b, diese Arbeit). Bei Umweltproben, die reich an organischen Stoffen sind (z.B. Bodenproben, Schlamm), erfolgt aber dabei auch die Koextraktion von Humin- und Fulvinsäuren, die für nachfolgende molekularbiologische Arbeitsschritte sind (TSAI und OLSON, 1992; SMALLA et al., 1993; TEBBE und VAHJEN, 1993; PURDY, 2005). Daher müssen diese Substanzen nach der Isolierung der störend DNA durch gezielte Reinigungsschritte wie z.B. über Chromatographie an Silikagelmateriale entfernt werden. Es gibt aber auch Möglichkeiten zur Reduzierung der Menge an koextrahierten Humin- und Fulvinsäuren. Zum Beispiel können Huminsäurekontaminationen erfolgreich durch chemische Koagulation mit $\text{Al}(\text{NH}_4)(\text{SO}_4)_2$ (BRAID et al., 2003), durch Vorbehandlung mit Triton X-100 in Gegenwart von

Lösungsmitteln (PUROHIT et al., 2003) oder durch einen Waschschrift der Probe in Triton X-100 und EDTA (WATSON und BLACKWELL, 2000; FORTIN et al., 2004) reduziert werden. Allerdings haben sich diese Methoden bis heute noch nicht durchgesetzt und werden selten angewendet (PURDY, 2005).

Die verschiedenen Methoden für den eigentlichen Zellaufschluss sind bei der direkte Lyse der Zellen oder der vorherigen Zellseparation mit anschließender Zelllyse nahezu identisch. In dieser Arbeit wurde die DNA-Isolierung nach der Methode der direkten Lyse der Zellen ohne vorherige Zellseparation angewendet (s. 3.1.1). Die Zelllyse wird dabei durch Kombination eines chemischen- und enzymatischen Verfahrens erzielt. Die Zellen werden durch Verwendung eines Hochsalz-Extraktionspuffers (1,5 M NaCl) in einer ausgedehnten Hitzeinkubation in Gegenwart von EDTA, Proteinase K sowie der Detergenzien SDS und CTAB lysiert. Nach Abtrennung der Matrixbestandteile und der Zelltrümmer wird die DNA durch eine Chloroform/Isoamylalkohol-Behandlung extrahiert und in Gegenwart von Isopropanol gefällt. Die Reinheit der daraus resultierenden hochmolekularen Roh-DNA war abhängig von dem für die DNA-Isolierung verwendeten Umwelt-Standort oder Proben-Gemisch. Mit viel koextrahierten Huminsäuren verunreinigt waren die isolierte Roh-DNA der Standortortproben-Gemische, die Zuckerrübenboden oder Wiesenboden enthielten (s. Tab. 6; 3.1.5). Eine stark bräunliche Färbung, die indikativ für Kontaminationen mit Huminsäuren ist, zeigte auch die Roh-DNA aus der Kompostprobe und der Bodenprobe aus dem Erzabbau-Gebiet. Eine Entfernung der störenden Huminsäuren konnte, wie auch in anderen Studien gezeigt wurde, durch Chromatographie an Silikagelmaterialein und durch Anionenaustausch-Chromatographie fast vollständig erreicht werden (ZHOU et al., 1996; HENNE et al., 1999; DANIEL, 2002; PURDY, 2005). Nur die gereinigte DNA der Kompostprobe und die DNA der Probe aus dem Erzabbaugebiet wiesen nach diesen Reinigungsschritten immer noch eine leicht bräunliche Färbung auf. Dies hatte negative Auswirkungen auf die Klonierung dieser DNA (s. 3.1.4).

Außer der chemischen und der enzymatischen Zelllyse sind auch Methoden für den mechanischen Zellaufschluss wie z.B. durch Einsatz eines „Bead-Beaters“

oder durch Zelllyse mit der French-Press in der Literatur beschrieben (MORE et al., 1994; DE LIPTHAY et al., 2004). Der Nachteil des mechanischen Zellaufschlusses liegt in der Qualität der isolierten Roh-DNA, da dies häufig zu einer starken Scherung der DNA führt und somit keine hochmolekulare DNA erhalten wird. Ein Vorteil ist jedoch, dass sich Zellen von Gram-positiven Bakterien besser aufschließen lassen als durch chemische oder enzymatische Behandlung (KAUFMANN et al., 2004). Weitere Methoden zur Zelllyse sind die Ultraschall-Behandlung und der Zellaufschluss über das so genannte „Freeze/Thaw“-Verfahren (PICARD et al., 1992; TSAI und OLSON, 1992). Auch für die Reinigung der Roh-DNA gibt es außer den erwähnten Chromatographie-Verfahren an Silikagelmateriale oder Anionenaustauschern weitere alternative Methoden. Zum Beispiel kann eine effiziente Reinigung der isolierten DNA durch Agarosegelelektrophorese (STEIN et al., 1996), durch CsCl-Dichtegradientenzentrifugation und/oder Hydroxyapatit-Chromatographie erreicht werden (OGRAM, 1987; SELENSKA und KLINGMÜLLER, 1991; BRUCE et al., 1992; LLOYD-JONES und HUNTER, 2001). Diese Methoden sind allerdings sehr zeitaufwendig und führen oft zu hohen DNA-Verlusten.

Anhand des vorgestellten breiten Spektrums der Möglichkeiten zur DNA-Isolierung und DNA-Reinigung ist es ersichtlich, dass es kein universelles Protokoll gibt, welches für alle Standortproben angewendet werden kann. Daher ist es empfehlenswert mehrere Protokolle zu testen und danach zu entscheiden, welche Methoden sich für den jeweiligen Standort eignen (DANIEL, 2002; PURDY, 2005). Entscheidend ist nur die Qualität der isolierten DNA. Diese sollte hochmolekular und zur Gewährleistung einer effizienten Klonierung möglichst frei von Huminsäuren sein, die für Restriktionsenzyme, DNA-Polymerasen und DNA-Ligasen inhibierend sind.

Für die Konstruktion von Metagenombanken aus der isolierten DNA können als Vektorsysteme Plasmide (HENNE et al., 1999; KNIETSCH et al., 2003a und 2003b; GABOR et al., 2004), Cosmide bzw. Fosmide (STEIN et al., 1996; ENTCHEVA et al., 2001; SCHMITZ, 2005) und „bacterial artificial chromosomes“ (BACs) (HANDELSMAN et al., 1998; RONDON et al., 2000; GILLESPIE et al., 2002) verwendet werden. Die Auswahl des geeigneten

Vektorsystems ist dabei abhängig von der gewünschten durchschnittlichen Insertgröße, von der Kopienzahl und von dem Wirtsorganismus, der für Genbankherstellung verwendet werden soll. Ein weiteres Kriterium ist die Qualität der isolierten DNA (DANIEL, 2005). Es konnte in einigen Fällen Umwelt-DNA, die noch Verunreinigungen mit Huminsäuren aufwies, nur in Plasmide kloniert werden (DANIEL, 2002). Man unterscheidet so genannte „small-Insert“-Metagenombanken (Insertgrößen <15 kb), die in Plasmiden hergestellt werden, und „large-insert“-Metagenombanken, die in Cosmiden bzw. Fosmiden (Insertgrößen bis 40 kb) und BACs (Insertgrößen >40 kb) erstellt werden. In „high-copy“-Plasmiden hergestellte „small-Insert“-Metagenombanken haben den Vorteil, dass sie in hoher Kopienzahl im Wirtsorganismus repliziert werden. Dies kann die Detektion eines nur schwach exprimierten Fremd-Gens bei Verwendung von auf biologischer Aktivität beruhenden Screening-Verfahren ermöglichen (DANIEL, 2002). Weiterhin ist die Konstruktion von „small-insert“-Metagenombanken weniger aufwendig als die Konstruktion von „large-insert“-Metagenombanken. Ein Nachteil von „small-insert“-Metagenombanken im Vergleich zu „large-insert“-Metagenombanken ist dagegen die geringere Insertgröße. Daher muss für das Auffinden der gewünschten Zielgene eine größere Anzahl an Klonen durchmustert werden. Ferner liegen bei „small-insert“-Metagenombanken größere Gencluster nur unvollständig kloniert vor und werden daher bei aktivitätsbasierenden Screening-Verfahren nicht erfasst. Ein Beispiel hierfür sind Gencluster für die Synthese von Polyketiden, die mehr als 100 kb umfassen können (DANIEL, 2005).

Im Rahmen dieser Arbeit wurden ausschließlich Metagenombanken in „high-copy“-Plasmiden hergestellt und mit auf aktivitätsbasierenden Screening-Verfahren anschließend auf Gene durchmustert, die Butanol-Dehydrogenase-Aktivität oder proteolytische Aktivität vermitteln. Es wurden nur „small-insert“-Metagenombanken konstruiert, da die Zielaktivitäten häufig nur von einzelnen Genen vermittelt werden und daher eine Konstruktion von „large-insert“-Metagenombanken nicht nötig war. Weiterhin erhoffte man sich durch die in dem Wirtstamm *E. coli* hervorgerufene hohe Kopienzahl, eine Detektion von schwach exprimierten Zielgenen zu erzielen. Auch ermöglichen die kleineren Insertgrößen eine schnellere Identifizierung der Zielgene nach Isolierung der

rekombinanten Plasmide aus positiven Klonen. Die für die Klonierung verwendete Umwelt-DNA wurde nach ihrer Isolierung und Reinigung, verdaut, größenfragmentiert und mit Hilfe der T4-DNA-Ligase in „high-copy“-Vektoren kloniert. Dabei traten, wie vorher schon erwähnt, Schwierigkeiten mit der DNA aus der Komposterde und der Bodenprobe aus dem Erzabbaugebiet auf, die noch geringe Verunreinigungen mit Huminsäuren aufwiesen. Die Klonierbarkeit dieser DNA war sehr gering (oft weniger als 500 rekombinante Klone pro Ligationsansatz). Daher wurde im Rahmen der Arbeit eine Klonierungsmethode entwickelt, die auf der Methode der TOPO-T/A-Klonierung basiert. Diese zeichnet sich dadurch aus, dass sie unabhängig von der Ligation mit DNA-Ligasen ist, die durch Huminsäuren inhibiert werden (s. 3.1.4). Mit dieser neu ausgearbeiteten Methode, die ähnlich der von Wilkinson und Mitarbeitern entwickelten Methode zur Klonierung von DNA aus geothermalen Sedimentproben (WILKINSON et al., 2002) war, konnte die DNA der beiden Habitate effizient kloniert werden. Die Klonierungsmethode wurde auch für die Klonierung der DNA aus anderen Standorten bzw. für Standortprobengemische angewendet. Im Vergleich zur Klonierung mit T4-DNA-Ligase konnte bei Verwendung der TOPO-T/A-Klonierungsmethode die Anzahl an rekombinanten Klonen um das 2 bis 5-fache gesteigert und der prozentuale Anteil von Plasmiden mit Inserts um 6 % bis 19 % erhöht werden (s. 3.1.4).

Insgesamt wurden im Rahmen der Arbeit 12 unterschiedliche „small-insert“-Metagenombanken hergestellt, die zusammengefasst 598.000 rekombinante Klone mit einer durchschnittlichen Insertgröße von 3,9 kb und einer durchschnittlichen Inserthäufigkeit von 66 % aufwiesen (s. Tab. 6; 3.1.5). Die hergestellten Metagenombanken repräsentierten insgesamt 2,2 Gbp an klonierter Umwelt-DNA. In Tabelle 19 werden die im Rahmen der Arbeit konstruierten Genbanken mit anderen in der Literatur beschriebenen „small-insert“-Metagenombanken verglichen. Aus dieser Gegenüberstellung ist ersichtlich, dass die in der Arbeit konstruierten und die publizierten „small-insert“-Metagenombanken eine ähnliche Anzahl an rekombinanten Plasmiden aufweisen und auch die durchschnittlichen Insertgrößen miteinander vergleichbar sind.

Tabelle 19: Vergleich der im Rahmen dieser Arbeit konstruierten mit anderen „small-insert“-Metagenombanken

Bei allen aufgeführten Genbanken handelt es sich um „small-insert“-Genbanken, die in Plasmiden hergestellt wurden. Eine umfassende Charakterisierung, der im Rahmen dieser Arbeit konstruierten Genbanken, ist Tab. 6 zu entnehmen.

Habitat (Bezeichnung)	Anzahl rekombinanter Plasmide	Ø Insertgröße [Kbp]	Klonierte DNA [Mbp]	Referenz
^a Standort-Gemisch A (GB1)	52.000	4,2	218	diese Arbeit
^a Standort-Gemisch A (GB2)	78.000	5,3	413	diese Arbeit
^b Standort-Gemisch B (GB3)	46.000	3,1	143	diese Arbeit
^b Standort-Gemisch B (GB4)	55.000	3,9	215	diese Arbeit
^b Standort-Gemisch B (GB5)	65.000	4,0	260	diese Arbeit
^c Standort-Gemisch C (GB6)	65.000	2,8	182	diese Arbeit
^a Standort-Gemisch A (GB7)	70.000	5,4	378	diese Arbeit
Wattsediment (GB8)	28.000	2,6	73	diese Arbeit
Bodenprobe Erzabbau (GB9)	10.000	3,9	39	diese Arbeit
Bodenprobe Erzabbau (GB10)	37.000	4,6	170	diese Arbeit
Komposterde (GB11)	60.000	3,5	210	diese Arbeit
Bodenprobe	30.000	3,5	110	YUN et al., 2004
Sediment Golf von Eilat	63.200	5,6	212	KNIETSCH et al., 2003c
Sediment Solar Lake	37.300	4,0	98	KNIETSCH et al., 2003b
Zuckerrübenboden	71.050	5,4	272	KNIETSCH et al., 2003c
Flusssediment	31.800	4,2	83,3	SCHMITZ, 2005

- (a): Standort-Gemisch A, verwendet wurden Wüstensand, Leinesediment, Faulturm, Sediment Golf von Eilat.
 (b): Standort-Gemisch B, verwendet wurden Salzbeckensediment, Boden, Sediment Golf von Eilat.
 (c): Standort-Gemisch C, verwendet wurden Faulturm, Boden, Wattsediment, Salzbeckensediment.

4.2 Analyse des Metagenoms

Die in der isolierten Umwelt-DNA bzw. in den daraus hergestellten Metagenombanken gespeicherte genetische Information kann auf unterschiedliche Weise genutzt werden. Sie z.B. kann für phylogenetische Analysen der in einem Standort vorhandenen mikrobiellen Diversität verwendet werden. Dies geschieht durch Sequenzanalyse von phylogenetischen Markergenen. Die meisten dieser Untersuchungen basieren auf der 16S rRNA-Genanalyse, die von Carl Woese eingeführt wurde (WOESE, 1987; WOESE et al., 1990). Das 16S-rRNA-Gen kodiert für die ribosomale 16S rRNA der kleinen 30S-Untereinheit von prokaryotischen Ribosomen. Das Gen besitzt sowohl stark konservierte Genregionen, die den Einsatz von universellen Oligonukleotiden zur PCR-Amplifikation ermöglichen, als auch variable Bereiche, die als phylogenetische Chronometer genutzt werden können (GONZALEZ und SAIZ-JIMENEZ, 2005). Woese nahm 16S rRNA-Genanalysen an kultivierten Mikroorganismen vor und konnte anhand der erhaltenen 16S rRNA-Gensequenzen im Jahr 1987 den ersten bakteriellen Stammbaum erstellen, der 12 unterschiedliche Phyla von kultivierten Organismen enthielt (WOESE, 1987). Die Einführung von kultivierungsunabhängigen Techniken, die erstmals von Pace und Mitarbeitern im Jahr 1985 beschrieben wurde (PACE et al., 1985) haben dazu geführt, dass bis heute der bakterielle Stammbaum auf 52 Phyla erweitert werden konnte (RAPPE und GIOVANNONI, 2003, s. Abb. 33). Dabei enthalten 26 Phyla keine kultivierten Organismen und werden nur durch 16S rRNA-Gensequenzen repräsentiert. Vierzehn Phyla, die kultivierte Vertreter enthalten, sind seit 1987 dazugekommen (RIESENFELD et al., 2004). Von diesen Phyla weisen aber 13 weniger als 5 kultivierte Spezies auf (GONZALEZ und SAIZ-JIMENEZ, 2005).

Abbildung 33: Phylogenetischer Stammbaum der Bakterien
 Der Stammbaum wurde aus RAPPE und GIOVANNONI (2003) entnommen. Die schwarzen Keile repräsentieren Phyla, die 1987 von Woese beschrieben wurden. Weiße Keile zeigen die 14 Phyla, mit kultivierten Vertretern, die seit 1987 dazugekommen sind. In grau dargestellt sind die 26 Phyla, die durch keine kultivierten Organismen repräsentiert werden. Die Länge des Maßstabs entspricht 0,05 Änderungen pro Basen-Position.

Durch 16S rRNA-Genanalysen von Metagenombanken oder von der isolierten Umwelt-DNA konnten bisher mikrobielle Lebensgemeinschaften vieler verschiedener Habitate beschrieben werden. Beispiele dafür sind heiße Quellen (HUGENHOLTZ et al., 1998), Flüsse (BÖCKELMANN et al., 2000; SCHMITZ,

2005), Trinkwasserbiofilm (SCHMEISSER et al., 2003), Sargasso Sea (VENTER et al., 2004) und Böden (RONDON et al., 2000). Durch weiterentwickelte Methoden konnten genetische „Fingerprints“ von Habitaten erstellt und miteinander verglichen werden (GONZALES und SAIZ-JIMENEZ, 2004; TRINGE et al., 2005). Im Rahmen der Arbeit wurden keine 16S rRNA-Genanalysen oder andere phylogenetische Untersuchungen durchgeführt. Jedoch konnten bei zwei rekombinanten Plasmiden (pTWP9 und pTWP13), die aus dem Screening der Metagenombanken nach Genen für Proteasen hervorgegangen waren, durch Sequenzanalyse der Inserts Rückschlüsse auf die Herkunft der klonierten Umwelt-DNA gezogen werden. Bei der Annotation der festgelegten offenen Leserahmen der Inserts von pTWP9 und pTWP13 war auffällig, dass das nächste verwandte Protein meistens aus der Gattung *Xanthomonas* stammte. Hierzu zählten auch Gene, die für stark konservierte ribosomale Proteine kodieren (s. Tab. 11 und 12; 3.3.3). Daher ist es wahrscheinlich, dass die in pTWP9 und pTWP13 inserierten DNA-Fragmente von Vertretern aus der Gattung *Xanthomonas* stammen, die zu den γ -*Proteobacteria* gehören.

Eine Möglichkeit zur anwendungsorientierten Nutzung der in Metagenombanken niedergelegten mikrobiellen Diversität ist die Identifizierung von neuartigen biotechnologisch relevanten Biomolekülen im Rahmen von Screening-Verfahren (DANIEL, 2004; HANDELSMAN, 2004). Die industrielle Biotechnologie, die auch als weiße Biotechnologie bezeichnet wird, gewinnt immer mehr an Bedeutung in Bezug auf die Herstellung neuer Produkte oder die Optimierung bestehender Produktionsverfahren (LORENZ und ZINKE, 2005). Der Vorteil von biokatalytischen Verfahren gegenüber konventionellen chemischen Prozessen besteht darin, dass Biokatalyse im Vergleich zu vielen chemischen Verfahren unter milderer Bedingungen durchgeführt werden kann. Weiterhin fallen weniger giftige Abfälle oder unerwünschte Nebenprodukte an und oftmals sind Mikroorganismen oder deren Enzyme in der Lage, Produkte in einem ein- oder wenigstufigen Prozess zu produzieren. Ein weiterer Vorteil liegt in dem Einsatz von nachwachsenden Rohstoffen wie z.B. Stärke oder Cellulose als Ausgangsmaterialien für biotechnologisch hergestellte Produkte, welches zur Schonung der immer knapper werdenden fossilen Ressourcen beiträgt

(WILLKE und VORLOP, 2004; GAVRILESCU und CHISTI, 2005). In einem Bericht von McKinsey gehen Experten davon aus, dass bis zum Jahr 2010 schon 20 Prozent aller verkauften chemischen Substanzen, welches einem Jahresumsatz von rund 160 Milliarden Dollar entspräche, aus biotechnologischen Produktionsprozessen stammen könnten (LORENZ und ECK, 2005).

Die Durchmusterung von Metagenombanken nach neuen industriell interessanten Enzymen kann prinzipiell mit zwei unterschiedlichen Strategien erfolgen. Zum einen kann das Screening auf der Basis von Sequenzähnlichkeiten oder konservierten Sequenzmotiven erfolgen. Dazu können Sonden aus konservierten DNA-Regionen konstruiert werden. Anschließend können diese beispielsweise zur Identifizierung der Klone mit den gewünschten Genen durch Koloniehybridisierung eingesetzt werden (KNIETSCH et al., 2003c). Eine weitere Möglichkeit ist eine direkte Amplifikation der Zielgene durch PCR mit degenerierten Primerpaaren (ESCHENFELD et al., 2001). Ein Nachteil von allen auf Sequenzähnlichkeiten beruhenden Screeningverfahren ist, dass die Funktionalität der von den identifizierten Genen kodierten Genprodukte nicht gewährleistet ist, da auch unvollständige Gene erhalten werden können. Ferner lassen sich mit diesen Verfahren nur neue Vertreter schon bekannter Gen- bzw. Proteinfamilien finden. Die andere Strategie zur Durchmusterung von Metagenombanken basiert auf der biologischen Aktivität von heterolog produzierten Genprodukten. Hierfür können z.B. spezielle Plattentestverfahren eingesetzt werden, bei denen die Klone mit der gewünschten Aktivität durch Umsatz von mit Farbstoffen gekoppelten Substraten oder Reaktionen mit anderen Indikatoren identifiziert werden. Ein Vorteil der aktivitätsbasierenden Durchmusterung ist, dass man strukturell vollkommen neuartige Gene bzw. Genprodukte für die gewünschte Aktivität entdecken kann. Ferner ist bei dieser Screening-Strategie gewährleistet, dass man nur funktionelle Zielgene bzw. Genprodukte detektiert (DANIEL, 2004; STREIT et al., 2004; DANIEL, 2005).

Beide prinzipiell unterschiedlichen Screening-Strategien von Metagenombanken konnten in der Vergangenheit erfolgreich zur Identifizierung

neuer industriell relevanter Enzyme und von pharmazeutisch wichtigen Naturstoffen beitragen (DANIEL, 2005; LORENZ und ECK, 2005). Beispiele dafür sind die Identifizierung neuer Amidasen (GABOR et al., 2004), Amylasen (RONDON et al., 2000), Cellulasen (REES et al., 2003), Dehydratasen (KNIETSCH et al., 2003c), Oxidoreduktasen (KNIETSCH et al., 2003b), Lipasen /Esterasen (Henne et al., 2000; LEE et al., 2004), Proteasen (GUPTA et al., 2002; diese Arbeit) und Antibiotika (GILLESPIE et al., 2002).

In der vorliegenden Arbeit wurden die hergestellten Metagenombanken auf neuartige Gene durchmustert, die Butanol-Dehydrogenase-Aktivität oder proteolytische Aktivität vermitteln. Für beide Zielmoleküle wurden aktivitätsbasierende Screening-Verfahren auf Indikatoragarplatten angewendet. Der für das Butanol-Dehydrogenase-Screening eingesetzte Indikatoragar konnte auch schon erfolgreich zur Identifizierung neuartiger Oxidoreduktasen, welche zur Oxidation von kurzkettigen Polyolen (C2-C4) mit benachbarten Hydroxylgruppen befähigt waren, angewendet werden (KNIETSCH et al., 2003a und 2003b). Anstelle von Glycerin/1,2-Propandiol, 1,2-Ethandiol und 2,3-Butandiol wurde in der vorliegenden Arbeit 1-Butanol als Testsubstrat verwendet. Der Test wurde in einer abgewandelten Form zuerst von Conway und Mitarbeitern publiziert (CONWAY et al., 1987). Der Indikatoragar beinhaltete den Farbstoff Pararosanilin, welcher mit dem ebenfalls im Agar befindlichen Natriumbisulfit ein farbloses Schiff'sches Reagenz bildete. Kam es durch eine Aktivität des *E. coli*-Klons zur Oxidation von Butanol und zur Bildung des entsprechenden Carbonyls ging dieses eine Additionsverbindung mit dem Schiff'schen Reagenz ein. Daraus resultierte ein Additionsprodukt, eine Schiff'sche Base (Azomethin), welche sich durch eine rote Farbe auszeichnete. Daher konnten positive Klone aufgrund ihrer Rotfärbung und der Ausbildung eines roten Hofes um die Kolonie leicht auf den Indikatoragarplatten identifiziert werden (s. Abb. 11; 3.2.1). Insgesamt wurden 130 Mbp Umwelt-DNA aus vier unterschiedlichen Metagenombanken (GB1, GB2, GB8 und GB10; s. Tab. 6; 3.1.5) auf das Vorhandensein der Zielgene durchmustert. Es konnten dabei zwei rekombinante *E. coli*-Klone identifiziert werden, die einen stabilen roten Phänotyp aufwiesen. Daraus resultierte ein Verhältnis von 1 positiver Klon mit Butanol-Dehydrogenase-Aktivität auf 65 Mbp durchmusterter Umwelt-DNA (s.

3.2.2). In der oben schon erwähnten mit dem gleichen Testprinzip durchgeführten Durchmusterung von Metagenombanken nach Oxidoreduktasen konnte in zwei unterschiedlichen Studien Trefferquoten von 1 positiver Klon pro 213 Mbp und von 1 positiver Klon pro 77 Mbp durchmusterter Umwelt-DNA erreicht werden (KNIETSCH, 2002; KNIETSCH et al., 2003a und 2003b). Im Vergleich dazu wurde im Rahmen dieser Arbeit eine höhere Trefferquote erzielt.

Bei der Durchmusterung von Metagenombanken nach Genen, die proteolytische Aktivität vermitteln, wurden die eigenen und auch schon in der Arbeitsgruppe vorhandene Metagenombanken eingesetzt. Für das auf Enzymaktivität basierende Screeningverfahren wurde ein Testagar verwendet, der schon mehrfach in der Literatur für die Identifizierung von Protease-Genen beschrieben wurde (RAJAMANI und HILDA, 1987; RONDON et al., 2000). Das im Agar emulgierte Testsubstrat Magermilchprotein (Skim-Milk) verlieh dem Agar eine starke Trübung. Proteolytische Aktivität führte zur hydrolytischen Spaltung der Peptidbindungen des Testsubstrats, welches durch eine Aufklärung des Mediums um die Kolonie angezeigt wurde (s. Abb. 16, 3.3.1). Das Screening erfolgte im größeren Umfang als das zuvor beschriebene. Insgesamt wurden 533.000 *E. coli*-Klone, die 1,7 Gbp Umwelt-DNA enthielten, auf proteolytische Aktivität untersucht (s. Tab. 10; 3.3.2). Es konnten 5 *E. coli*-Klone mit stabiler proteolytischer Aktivität identifiziert werden, welches einem Verhältnis von 1 positiven Klon auf 340 Mbp gescreenter Umwelt-DNA entspricht. In der Literatur sind drei weitere Beispiele für die Durchmusterung von Metagenombanken nach Genen für Proteasen beschrieben. Rondon und Mitarbeiter erstellten zwei „large-insert“-Metagenombanken durch direkte Klonierung von aus Bodenproben isolierter DNA in BAC-Vektoren (RONDON et al., 2000). Zusammen enthielten die beiden Metagenombanken ca. 1,1 Gbp an klonierter Umwelt-DNA. Für die auf Enzymaktivität beruhende Durchmusterung nach Protease-Genen wurde nur eine Genbank, die 100 Mbp klonierter Umwelt-DNA enthielt, verwendet. Dabei konnte kein Klon mit proteolytischer Aktivität identifiziert werden (s. Tab. 20). In einer Untersuchung von Lämmle konnte ebenfalls bei der Durchmusterung einer Metagenombank, die 107 Mbp Boden-DNA enthielt, kein Klon mit proteolytischer Aktivität gefunden werden (LÄMMLE, 2004; s. Tab. 20). Diese Ergebnisse stimmen gut mit den in dieser

Arbeit erhaltenen überein. Es wurde ein Verhältnis von einem positiven Klon auf 340 Mbp durchmusterter Umwelt-DNA ermittelt. Daher lagen Rondon und Mitarbeiter sowie Lämmle mit 100 bzw. 107 Mbp durchmusterter Umwelt-DNA deutlich unter der im Rahmen dieser Arbeit ermittelten Menge an klonierter Umwelt-DNA, die zur Identifizierung eines positiven Klons durchmustert werden musste. In der Studie von Gupta und Mitarbeitern wurden eine aus Boden-DNA erstellte „small-insert“-Metagenombank mit Hilfe von Skim-Milk-Indikatoragar auf das Vorhandensein von Protease-Genen getestet (GUPTA et al., 2002). Die Arbeitsgruppe durchmusterte 1 Gbp Umwelt-DNA, die in 100.000 rekombinanten Plasmiden niedergelegt war. Dabei konnte ein positiver Klon identifiziert werden. Somit wurde hier eine Trefferquote von einem positiven Klon auf 1 Gbp durchmusterter Umwelt-DNA erzielt (s. Tab. 20). Sequenzanalysen des Protease-Gens ergaben, dass es sich um eine Metalloprotease handelte, die 30 % Identität zu einer Protease aus einem Seestern aufwies (GUPTA et al., 2002). Zusammenfassend lässt sich feststellen, dass die in dieser Arbeit erreichten Trefferquoten deutlich über denen liegen, die in vergleichbaren Durchmusterungen von anderen Autoren erzielt wurden (s. Tab. 20).

Aus den erhaltenen Ergebnissen ist ersichtlich, dass eine hohe Anzahl an klonierter Umwelt-DNA bzw. Klonen getestet werden muss, um einen Klon mit gewünschter Aktivität zu erhalten. Die Anzahl an zu durchmusternden Klonen pro Treffer ist abhängig von der Frequenz der Organismen, die das gesuchte Gen tragen (GABOR et al., 2003; DANIEL, 2005). Es konnte in unterschiedlichen Studien gezeigt werden, dass eine Erhöhung der Frequenz der Organismen mit den Zielgenen in der zur Genbankerstellung verwendeten DNA, auch zu einer Erhöhung der Trefferzahl an positiven Klonen bei der Durchmusterung von Metagenombanken führt (HEALY et al., 1995; ENTCHEVA et al., 2001, KNIETSCH et al., 2003b und 2003c; GABOR et al., 2004). Eine Erhöhung der Trefferzahl kann durch Nutzung von Probenmaterial erreicht werden, welches eine natürliche oder im Labor erzeugte Anreicherung von Organismen mit der gewünschten Eigenschaft aufweist (DANIEL, 2004). Zur Detektion von z.B. thermostabilen Enzymen aus Metagenombanken empfiehlt es sich als Ausgangsmaterial zur Konstruktion der Genbanken

Proben aus heißen Quellen zu verwenden, da dort thermophile Organismen natürlich angereichert sind (VIEILLE und ZEIKUS, 2001). Im Labor erzeugte Anreicherungen können durch eine Inkubation der Umwelt-Probe in Gegenwart von einer selektiver C- oder N-Quelle erreicht werden (Gabor, 2004; DANIEL, 2005). Ein Beispiel dafür ist die schon vorher beschriebene Durchmusterung von Metagenombanken nach Genen für Oxidoreduktasen (KNIETSCH, 2002; KNIETSCH et al., 2003a und 2003b). Es wurden zum einen Genbanken bei denen kein Anreicherungsschritt vor der DNA-Isolierung vorgenommen wurde und zum anderen Anreicherungs-genbanken für die Durchmusterung verwendet. Das Testprinzip war bei beiden Untersuchungen dasselbe. Durch den Anreicherungsschritt konnte die Trefferquote um das 2,7-fache gesteigert werden (s. Tab. 20). Ein Nachteil von Anreicherungsschritten ist der Verlust an mikrobieller Diversität in der Probe, da durch die Anreicherung eine Selektion vorgenommen wird bei der schneller wachsende Organismen andere verdrängen (CALDWELL et al., 1997). Eine weitere Möglichkeit, um eine gezielte Anreicherung vorzunehmen, ist die Methode des „stable-isotope probing“. Bei diesem Verfahren werden mit dem in der Natur seltenen ^{13}C -Isotop markierte Substrate eingesetzt, die selektiv für das Vorhandensein der Zielgene sind. Die Mikroorganismen, die diese Substrate verwerten können und somit die Zielgene besitzen, bauen das ^{13}C in ihrer DNA ein. Die DNA der Zielorganismen wird daraufhin schwerer als die ^{12}C -DNA von Mikroorganismen, die die Zielgene nicht besitzen und daher das verwendete Substrat nicht verwerten können. Nach durchgeführter DNA-Isolierung kann durch CsCl-Dichtegradienten-Zentrifugation eine Trennung zwischen der „leichten“ ^{12}C -DNA und der „schweren“ ^{13}C -DNA erzielt werden. Für die Konstruktion der Genbank wird nur die ^{13}C -DNA verwendet, die nur aus Zielgen-enthaltenden Mikroorganismen stammt (WELLINGTON et al., 2003; DUMONT und MURRELL, 2005; SCHWARZ et al., 2005). Auch in der vorliegenden Arbeit wurden bei einigen der konstruierten Metagenombanken Anreicherungsschritte mit selektiver C-Quelle (Zimtsäure bzw. 2-Penten) vor der DNA-Isolierung durchgeführt. Da diese Anreicherungsschritte jedoch für andere als die im Rahmen dieser Arbeit gesuchten Gene gedacht waren, können aus den erhaltenen Ergebnissen keine Schlüsse auf Auswirkungen des Anreicherungsschrittes auf die Trefferquote abgeleitet werden.

Tabelle 20: Vergleich der Ergebnisse, der in dieser Arbeit durchgeführten aktivitätsbasierenden Screenings mit in der Literatur beschriebenen ähnlichen Untersuchungen.
 Farblich unterlegt, sind die Ergebnisse dieser Arbeit. Grau hinterlegt ist das aktivitätsbasierende Screening nach Oxidoreduktasen, welches mit Genbanken durchgeführt wurde, bei denen vorher ein Anreicherungs-schritt durchgeführt wurde.

Screening-Ziel	Genbank-Typ	Anzahl getesteter Klone	Durchmusterung Umwelt-DNA [Gbp]	Anzahl an Treffern	Mbp durchmusterter DNA pro Treffer	Referenz
♦Proteasen	Plasmid/ Cosmid	533.000	1,7	5	340	diese Arbeit
♦Proteasen	BAC	3648	0,1	0	-	RONDON et al., 2000
♦Proteasen	Plasmid	100.000	1	1	1000	GUPTA et al., 2002
♦Proteasen	Plasmid	32.000	0,1	0	-	LÄMMLE, 2004
*Butanol-Dehydrogenasen	Plasmid	53.000	0,13	2	65	diese Arbeit
*Oxidoreduktasen	Plasmid	1.000.000	3,2	15	213	KNIETSCH, 2002; KNIETSCH et al., 2003a
*Oxidoreduktasen	Plasmid	400.000	1,23	16	77	KNIETSCH, 2002; KNIETSCH et al., 2003b

(♦) Screeningverfahren: Plattentest mit Skim-Milk als Substrat; (*) Screeningverfahren: Indikatoragar mit Pararosanilin und Natriumbisulfit als Indikatoren aber mit unterschiedlichen Testsubstraten.

4.3 Identifizierung von Genen, die Butanol-Dehydrogenase-Aktivität vermitteln

Das Lösungsmittel Butanol ist von großer industrieller Bedeutung, da es als eine der wichtigsten Grundchemikalien vielseitige Verwendung findet (s. 1.2). Kommerziell wird es heutzutage überwiegend in einem zweistufigen chemischen Prozess hergestellt. Einige Mikroorganismen sind ebenfalls in der Lage, Butanol zu produzieren. Hauptproduzenten sind Vertreter der Gattung *Clostridium* z.B. *C. acetobutylicum*, *C. beijerinckii*, *C. saccharobutylicum* und *C. saccharoperbutylacetonicum* (DÜRRE, 2005). Neben Vertretern der Gattung *Clostridium* sind aber auch andere Mikroorganismen zur Butanolproduktion befähigt. Bei dem obligat anaeroben Organismus *Butyrivacterium methylotrophicum*, der C₁-Substrate wie z.B. Kohlenmonoxid, Kohlendioxid, Methanol und Formiat verwerten kann, konnte eine NADH- und eine NADPH-abhängige Butanol-Dehydrogenase-Aktivität von 0,68 U/mg bzw. 0,13 U/mg ermittelt werden (SHEN et al., 1999). *Pseudomonas butanovora* ein Gram-negatives aerobes Proteobakterium, welches von einer Ölfirma aus Belebtschlamm isoliert wurde (TAKAHASHI, 1980; TAKAHASHI et al., 1980), besitzt zwei Butanol-Dehydrogenasen (ARP, 1999; VANGNAI et al., 2002a). Die beiden Enzyme, die als BDH und BOH bezeichnet werden, zeigen allerdings eine PQQ (Pyrrolochinolin-chinon)-abhängige und NAD(P)H-unabhängige Dehydrogenase-Aktivität (VANGNAI et al., 2002a). Dieser Organismus ist in der Lage mit n-Alkanen (C₂ bis C₉) bzw. deren Oxidationsprodukten zu wachsen (TAKAHASHI, 1980; TAKAHASHI et al., 1980). So wie es für andere n-Alkan-verwertende Organismen beschrieben ist, wird die Oxidation im ersten Schritt durch eine Monooxygenase katalysiert (HAMAMURA et al., 1997). Allerdings ist Butanol in dem Stoffwechsel von *P. butanovora* nicht das Endprodukt der Alkanoxidation sondern es wird weiter über Butyraldehyd zu Butyrat umgesetzt. Letzteres geht dann möglicherweise in die β -Oxidation ein (ARP, 1999). Weitere Studien haben gezeigt, dass BDH und BOH in der Lage sind auch die Rückreaktion zu katalysieren. Daher können diese Enzyme auch Butyraldehyd zu Butanol umsetzen (VANGNAI et al., 2002b). Außerdem zur Butanol-Produktion befähigt ist das Schwefelproduzierende, extrem thermophile Archaeon *Hyperthermus butylicus*. Die

daran beteiligten Enzyme sowie deren Gene sind allerdings noch nicht identifiziert (ZILLIG et al., 1990).

Am besten untersucht ist die Butanol-Produktion von *C. acetobutylicum* (s. Abb. 1; 1.2). Dabei ist erwähnenswert, dass *C. acetobutylicum* in der Lage ist, mit nachwachsenden Ressourcen wie z.B. Stärke, Xylan und Cellobiose als C-Quelle zu wachsen und daraus Butanol zu bilden (LEE et al., 1985; MITCHELL, 1998). In der ersten Phase (Säurephase), die während der exponentiellen Wachstumsphase stattfindet, werden dabei überwiegend die organischen Säuren Acetat und Butyrat sowie H_2 und CO_2 gebildet. Dadurch kommt es zu einer starken Abnahme des pH-Werts im Medium auf 4 bis 4,5. Dies ist eine Bedrohung für die Zellen, da sie durch den Zusammenbruch des Protonengradienten nicht mehr in der Lage sind ihren internen pH-Wert aufrecht zu halten, was zum Zelltod führt (GOTTWALD und GOTTSCHALK, 1985; DÜRRE et al., 1988). Um das Absinken des pH-Werts zu verhindern, kommt es zu einer drastischen Umstellung („Shift“) im Gärungsstoffwechsel von *C. acetobutylicum*. Es werden nun überwiegend die Lösungsmittel Aceton und Butanol und in einigen anderen Stämmen Isopropanol anstelle von Aceton produziert und in das Medium ausgeschieden (Lösungsmittelphase). Die produzierten Lösungsmittel sind ebenfalls toxisch für die Zellen, da die Membranintegrität zerstört wird. Am gefährlichsten ist dabei Butanol, welches in hohen Konzentrationen z.B. die ATPase und die Zucker-Aufnahme inhibiert. Jedoch führt die Lösungsmittel-Produktion zum Ansteigen des pH-Werts, und die Zellen gewinnen genügend Zeit, um die Sporenbildung einzuleiten, welche ein Überleben gewährleistet (Übersicht: DÜRRE, 2005). Die Gene und Genprodukte, die direkt an der Bildung von Butanol beteiligt sind, wurden identifiziert, gereinigt und charakterisiert (DÜRRE, 2005). Es handelt dabei sich um drei Enzyme, eine bifunktionale Butyraldehyd-/Butanol-Dehydrogenase (AdhE) sowie die Butanol-Dehydrogenasen A und B (BdhA und BdhB). Die tatsächliche Beteiligung von AdhE an der Butanol-Produktion konnte in mehreren Studien eindeutig gezeigt werden. So wird AdhE erst zu Beginn der Lösungsmittelphase exprimiert (SAUER und DÜRRE, 1995). Weiterhin erfolgte eine Wiederherstellung der Butanol-Bildung in einer Lösungsmittel-negativen Mutante nach Transformation des *adhE*-Gens (NAIR et al., 1994; NAIR und

PAPOUTSAKIS, 1994). In einer anderen Studie konnte eine drastische Abnahme der Butanol-Produktion nach Inaktivierung des *adhE*-Gens durch Insertion beobachtet werden (GREEN und BENNETT, 1996). AdhE ist aufgrund hoher Homologien den multifunktionalen Alkohol-Dehydrogenasen zuzuordnen und zeigt große Sequenzähnlichkeiten zu Genen bzw. Genprodukten aus *E. coli* (GOODLOVE et al., 1989), *Salmonella typhimurium* (DAILLY et al., 2000), *Lactococcus lactis* (ARNAU et al., 1998) und *Streptococcus bovis* (ASANUMA et al., 2004) sowie zu eukaryotischen Genen bzw. Genprodukten aus *Entamoeba histolytica* (YANG et al., 1994) und *Giardia lamblia* (SANCHEZ, 1998). Weiterhin konnte gezeigt werden, dass das *adhE*-Gen mit anderen Genen in dem so genannten *sol*-Operon organisiert ist, welches auf dem Megaplasmid pSOL1 von *C. acetobutylicum* lokalisiert ist (NÖLLING et al., 2001). Die spezifische Butanol-Dehydrogenase-Aktivität liegt für AdhE mit $\text{NADH} + \text{H}^+$ als Cofaktor bei 0,16 U/mg und für die Rückreaktion mit NAD^+ bei 0,03 U/mg (NAIR et al., 1994). Das *adhE*-Gen konnte bisher außer in *C. acetobutylicum* bei keinem anderen Lösungsmittel-produzierenden *Clostridium*-Vertreter nachgewiesen werden (DÜRRE, 2005). Die anderen besitzen statt eines *adhE*-Gens ein Aldehyd-Dehydrogenase-Gen (*ald*) (DÜRRE, 2005). Vor nicht allzu langer Zeit konnte ein zweites *adhE*-Gen (*adhE2*) bei *C. acetobutylicum* identifiziert werden (FONTAINE et al., 2002). Es zeigt 66 % Identität zu *adhE*. Es konnte gezeigt werden, dass AdhE2 sowohl eine NADH-abhängige Butanol-Dehydrogenase-Aktivität (0,12 U/mg) als auch eine Butyraldehyd-Dehydrogenase-Aktivität (0,02 U/mg) aufweist (FONTAINE et al., 2002). Die beiden Butanol-Dehydrogenasen BdhA und BdhB zeigen sowohl NADH-abhängige als auch NADPH-abhängige Enzymaktivitäten (PETERSEN, 1991). Die Proteinsequenzen beide Enzyme weisen 76 % Identität und 86 % Ähnlichkeit zueinander auf und sind durch Homologien zu Alkohol-Dehydrogenasen charakterisiert (WALTER et al., 1992). Auch bei anderen Clostridien konnten Butanol-Dehydrogenasen identifiziert werden (PALOSAARI und ROGERS., 1988; YAN und CHEN., 1990; WOODS und REID., 1995).

Trotz vorhandener Butanol-produzierender Mikroorganismen ist es bis heute nicht gelungen, die chemische Herstellung von Butanol durch ein ökonomisch konkurrenzfähiges Fermentationsverfahren zu ersetzen. Die Ausbeuten der

mikrobiellen Produktion sind niedriger und das biotechnologische Verfahren ist mit mehr Kosten verbunden (DÜRRE, 1998). Im Zuge der immer knapper werdenden Erdölressourcen stößt jedoch die mikrobielle Butanolproduktion auf zunehmendes Interesse der chemischen Industrie. Ein zusätzlicher Vorteil der biotechnologischen Produktion von Butanol wäre die damit gekoppelte Möglichkeit der Entsorgung von kohlenhydrathaltigen Abfällen, die z.B. bei der Getreideproduktion anfallen und als Ausgangsstoffe der Fermentation genutzt werden können. Derzeit wird in unterschiedlichen Studien versucht, ein wirtschaftliches biotechnologisches Verfahren zur Herstellung von Butanol zu entwickeln. Eine Arbeitsgruppe arbeitet derzeit an einem biotechnologischem Verfahren zur Produktion von Butanol aus Synthesegas (CO und H_2), welches aus Steinkohle oder beim Thermorecycling von Abfällen entsteht. Ein Organismus der Synthesegas als Substrat gut verwertet wurde schon gefunden (DÜRRE, 2005b). Ferner konnten die Mitarbeiter des Fraunhofer Instituts zeigen, dass mit n-Butan als alleiniger C-Quelle angereicherte Bakterienstämme überwiegend n-Butan zu 1-Butanol oxidieren (www.igb.fraunhofer.de/WWW/GF/Biokatalyse/dt/).

Ein weiterer Ansatz zur Entwicklung eines biotechnologisch einsetzbaren Produktionsstammes für Butanol wurde in dieser Arbeit verfolgt. Mit Hilfe der Metagenomik sollten neuartige Gene, die Butanol-Dehydrogenase-Aktivität vermitteln, identifiziert werden. Durch „enzyme engineering“ bzw. „pathway engineering“ könnte es hinterher möglich sein, solche Gene zur Optimierung eines Produktionsstammes zu verwenden. Mit einem auf Enzymaktivität basierenden Screening-Verfahren (s. 3.2.1) erfolgte die Durchmusterung von insgesamt 53.000 rekombinanten *E. coli*-Stämmen (0,13 Gbp Umwelt-DNA; s. Tab 6, 3.2.2). Als Wirt für die Durchmusterung wurde die *E. coli*-Mutante WL3 eingesetzt, da sie durch eine Mutation des wirtseigenen *adhE*-Gens nicht in der Lage war das Testsubstrat Butanol umzusetzen. Diese Mutante zeigte im Vergleich zu anderen *E. coli*-Stämmen keine Hintergrundaktivitäten auf den Indikatorplatten. Es konnten zwei positive *E. coli*-Klone (*E. coli*/WL3/pTWB3 und *E. coli*/WL3/pTWB4) auf den Indikatoragarplatten identifiziert werden, die das Testsubstrat Butanol oxidieren konnten. Die aus den positiven Klonen isolierten rekombinanten Plasmide pTWB3 bzw. pTWB4 wurden isoliert und die

Insertsequenzen analysiert. Auf beiden Inserts konnten je zwei offene Leserahmen (*orf31* und *orf32* für pTWB3; *orf41* und *orf42* für pTWB4) festgelegt werden, die aber jeweils keine vollständigen Gene bildeten und durch Vektorsequenzen komplementiert wurden (s. Abb. 12 und Abb. 14; 3.2.3). Interessanterweise zeigten die abgeleiteten Genprodukte der vier Leserahmen keinerlei Ähnlichkeiten zu bekannten Alkohol-Dehydrogenasen oder Oxidoreduktasen. Beide Klone wurden daraufhin einer biochemischen Charakterisierung unterzogen. In Rohextrakten von *E. coli*/WL3/pTWB3 und *E. coli*/WL3/pTWB4 wurde die spezifische Alkohol-Dehydrogenase-Aktivität sowohl in Oxidationsrichtung mit primären Alkoholen (C1 bis C6) als Substraten als auch in Reduktionsrichtung mit den korrespondierenden Aldehyden als Substraten bestimmt. Die höchste spezifische Aktivität zeigte der Rohextrakt von *E. coli*/WL3/pTWB4 mit 0,18 U/mg bei der NADPH-abhängigen Umsetzung von Butyraldehyd zu Butanol. Auch im Rohextrakt von *E. coli*/WL3/pTWB3 konnte eine spezifische Aktivität von 0,08 U/mg für die gleiche Reaktion ermittelt werden. Weiterhin waren beide Extrakte in der Lage, die längerkettigen Aldehyde (C5 und C6) jedoch mit niedrigerer Aktivität NADPH-abhängig umzusetzen. In Oxidationsrichtung fielen die NADP-abhängigen spezifischen Aktivitäten deutlich geringer aus. Mit NAD^+ bzw. $\text{NADH}^+ + \text{H}^+$ als Cosubstrat waren bis auf eine Ausnahme keine Aktivitäten messbar (s. Tab. 9; 3.2.4). Subklonierungen der einzelnen Gene von pTWB3 und pTWB4 und anschließende Bestimmung der Enzymaktivität in den entsprechenden Rohextrakten der Subklone sollten Aufschluss darüber geben, welches Gen jeweils für die Butanol-Dehydrogenase-Aktivität verantwortlich war. Für das Plasmid pTWB3 konnte eine solche Zuordnung nicht erfolgen. Bei dem Plasmid pTWB4 konnte *orf42* als das Gen identifiziert werden, welches die Butanol-Dehydrogenase-Aktivität vermittelt. Dabei war die spezifische Aktivität in dem Rohextrakt des entsprechenden Subklons in Reduktionsrichtung mit Butyraldehyd von 0,18 U/mg genau so hoch, wie die spezifische Aktivität im Rohextrakt von *E. coli*/pTWB4 (s. 3.2.4 und Tab. 9). Wie vorher schon erwähnt, zeigte die Sequenzanalyse des *orf42*-Genprodukts keinerlei Ähnlichkeiten zu in Datenbanken vorhandenen Sequenzen von Butanol-Dehydrogenasen oder anderen Oxidoreduktasen. Die Proteinsequenz des abgeleiteten Genprodukts wies große Übereinstimmungen (57 % Identität und 72 % Ähnlichkeit) zu einer

neutralen Zink-abhängigen Metallopeptidase (Oligoendopeptidase) aus *Bacillus licheniformis* ATCC 14580 (Accession-Nr.: AAU21932.1) auf (s. Tab. 8; 3.2.3). Zusätzlich zeigten sich signifikante Ähnlichkeiten zu der konservierten Domäne COG1146 der Oligoendopeptidase F. Diese Peptidasen werden in der MEROPS-Datenbank der Familie M3 zugeordnet. Alle Vertreter dieser Familie zeichnen sich durch die Gemeinsamkeit eines konservierten Sequenzmotivs (Zink-Bindemotiv) aus (RAWLINGS und BARRETT, 2004). Dieses Sequenzmotiv umfasst die Aminosäuren HEXXH. Dabei fungieren die beiden Histidin-Reste des konservierten Motivs zusammen mit einem weiteren ca. 15 Aminosäuren C-terminal befindlichen Glutamatrest als Zink-Liganden. In Abbildung 34 ist ein Aminosäure-Alignment des abgeleiteten *orf42*-Genprodukts mit der konservierten Domäne COG1146 der Oligoendopeptidase F aufgeführt. Aus diesem Alignment geht hervor, dass das *orf42*-Genprodukt das konservierte Zink-Bindemotiv und den Glutamatrest vollständig beinhaltet. Zur M3-Familie zugehörige Peptidasen werden unterschiedliche intrazelluläre Funktionen zugesprochen. Eine Gemeinsamkeit aller M3-Peptidasen ist, dass sie alle Peptide von geringer Größe degradieren. Die Spaltung kann dabei N-terminal oder auch C-terminal erfolgen (<http://merops.sanger.ac.uk/>; RAWLINGS und BARRETT, 2006).

Aufgrund der erhaltenen Ergebnisse kann man darauf schließen, dass *orf42* für eine Peptidase der M3-Familie kodiert. Somit können in Bezug auf die vom *orf42*-Genprodukt vermittelte Butanol-Dehydrogenase-Aktivität nur Vermutungen angestellt werden. Am wahrscheinlichsten scheint dabei, dass das *orf42*-Genprodukt dem Klon *E. coli*/WL3/pTWB4 die Substratumsetzung von Butyraldehyd zu Butanol auf einem indirekten Weg vermittelt. Möglich wäre, dass die vermutete Peptidase-Aktivität des Genprodukts ein *E. coli*-Regulatorprotein oder ein anderes *E. coli*-Protein verändert und eine bisher in *E. coli* unbekannte NADPH-abhängige Butanol-Dehydrogenase-Aktivität produziert bzw. aktiviert wird. Im Rahmen der Durchmusterung von Metagenombanken aufgetretene Umsetzungen von Testsubstraten auf indirektem Weg unter Beteiligung von Wirtsproteinen sind in der Literatur beschrieben. Bei der Durchmusterung von Umweltgenbanken nach Genen für Oxidoreduktasen, welche mit dem in dieser Arbeit verwendeten Testprinzip

erfolgte, konnte auch in einem Fall eine Substratumsetzung auf die Beteiligung von einem Wirtsprotein zurückgeführt werden. Bei einem der identifizierten Klone beruhte die Entstehung der aus Glycerin gebildeten Carbonylverbindung auf der kombinierten Aktivität von einer Glycerin-Kinase, die von der in das Plasmid inserierten Umwelt-DNA kodiert wurde und der wirtseigenen Glycerin-3-Phosphat-Dehydrogenase (KNIETSCH, 2002; KNIETSCH et al., 2003b). Diese Ergebnisse von auf biologischer Aktivität basierenden Screeningverfahren zeigen, dass es zwar möglich ist die Produktbildung nachzuweisen aber auf welchem Weg diese erfolgt, ist jedoch zunächst nicht immer sofort feststellbar.

<i>orf42</i>	6	TRDQVAV ^Y ETW ^N LD ^D LFASQA ^A WE ^Q ELDE ^I IISQLPAVTA ^F KGQ ^L H ^T GAETLLAC ^L EAK ^E A	65
COG1164	3	NRSEVPE ^K YT ^W D ^L SD ^L YPGDELWEA ^F LEESK ^E LLKALE ^F YEL ^I LS-AED ^L LEALEEY ^E K	61
<i>orf42</i>	66	IEVRANLAAT ^Y AR ^L RSSE ^D GTNPQ ^N ANSARVGDV ^V SRLNAE ^L SFVPSE ^I LALPDGT ^I EK	125
COG1164	62	LEELLGR ^L YAYAS ^M K ^L STD ^T DEEA ^Q K ^L YQ ^K LEEF ^Y ADFSS ^A LSFFEP ^E LE ^L DEE ^V IQ ^S	121
<i>orf42</i>	126	FLQEE ^P GLAP ^F RK ^S L ^F DLLET ^K SY ^R LSA ^E TEAS ^L AALGEVFGAP ^Y TIYQ ^R GK ^L SD ^M SFAP	185
COG1164	122	LLTSG ^P ELAD ^Y RFY ^L EELFRR ^K PH ^Q LSA ^E EK ^L LA ^Q LSPVGN ^S ASN ^I FFDLTNSD ^L T ^F PD	181
<i>orf42</i>	186	VQDGAGA ^R FPVS ^F ALFEADYEMSAD ^T TLRRA ^A YESFSNTLAAY ^Q NSFAA ^V YATE ^V K ^Q TV	245
COG1164	182	IGDAK ^G K ^L TLE ^Q L ^N ---LLED ^P DREV ^R KA ^A YEALYKAYE ^K HRNT ^L AAL ^N TL ^V K ^V LAF	238
<i>orf42</i>	246	LSRLRG ^F ETV ^T DML ^L HTQ ^Q VTLD ^M YHN ^I LDIIGTELAPH ^M RRYA ^K LKARE ^L GLEK ^L K ^F CD	305
COG1164	239	LARARN ^D DVIDSALLRNE ^V DREV ^D ND ^L IESVKEA ^F LPL ^L HR ^Y YKLAK ^V LGLEK ^L RP ^Y D	298
<i>orf42</i>	306	LKAP-LD ^P EFSP ^P ITIEEAG ^R L ^I Q ^S ALQ ^V MGPEY ^S SAIME ^K ALNERWIDYVDNVG ^K STG ^A F	364
COG1164	299	LYAP ^L LDK ^D PSPEY ^S EEAKE ^L VLKALAP ^L GPEY ^A KIAR ^R AFDERWIDVYPRK ^G K ^R SGAY	358
<i>orf42</i>	365	CTTPY-GN ^H SY ^I LIT ^W SGN ^M RS ^A FTLA ^H ELGH ^C GHFT ^L ATRE ^Q SYTNSR ^P S ^L YFI ^I AP ^S T	423
COG1164	359	SIGF ^Y K ^G D ^H PFIL ^M NYD ^G TLRD ^V FTLA ^H ELGH ^S VHSY ^F SR ^K HQ ^P YLYAD ^S IFLA ^I AST	418
<i>orf42</i>	424	INELL ^L GQHIMA ^Q SD ^D PRLKRW ^V ILQLMNT ^Y YHN ^F VTH ^L LEGRM ^Q RKV ^Y EAAQNDV ^P LTA	483
COG1164	419	FNE ^M LLFDYLLER ^F KDPEER ^L AILE ^E KLEGF ^F FATLFRQ ^T LFAEF ^E HRV ^H ELIEEGEELTA	478
<i>orf42</i>	484	KKLSEWKG ^Q VLA ^E FWGDAVELDEA ^A SLTW ^M RQPH ^Y YM	520
COG1164	479	EELNELY ^L ELQ ^E Y ^Y GDAVK ^L DEL ^S GLEWARIP ^H F ^Y H	515

Abbildung 34: Aminosäurealignment des *orf42*-Genprodukts mit der konservierten Domäne COG1164 der Oligoendopeptidase F.

Rot dargestellt sind identische Aminosäuren, blau dargestellt sind ähnliche Aminosäuren, schwarze Aminosäuren zeigen keinerlei Übereinstimmungen. Mit gelber Farbe unterlegt ist das konservierte Zink-Bindemotiv HEXXH sowie der dritte Zink-Ligand, der in beiden Sequenzen ein Glutamat darstellt.

Um den Nachweis zu erbringen, ob die Reduktion von Butyraldehyd zu Butanol tatsächlich von einem wirtseigenen Enzym vollzogen wird, sind weitere Untersuchungen notwendig. Eine Möglichkeit ist, die Proteine im Rohextrakt von *E. coli*/WL3/pTWB4 durch PAGE unter nativen Bedingungen elektrophoretisch aufzutrennen. Anschließend kann durch eine Aktivitätsfärbung mit Butanol als Substrat, das für die Produktbildung verantwortliche Protein lokalisiert werden. Eine Ansequenzierung der entsprechenden Proteinbande und ein Vergleich der dabei erhaltenen Aminosäuresequenz mit der Proteindatenbank des *E. coli*-Genoms könnte zeigen, ob es sich um ein *E. coli*-Protein handelt und um welches. Entsprechende Untersuchungen werden momentan in Zusammenarbeit mit der Arbeitsgruppe von Prof. P. Dürre (Universität Ulm) durchgeführt.

4.4 Identifizierung von Genen aus Metagenombanken, die proteolytische Aktivität vermitteln

Proteasen (Peptidasen) sind Enzyme die Peptidbindungen hydrolysieren. Sie übernehmen zahlreiche essentielle Funktionen, die für das Überleben aller Organismen notwendig sind und kommen ubiquitär vor. Es wird angenommen, dass mindestens 2 % aller in Organismen vorkommenden Gene für Proteasen kodieren (<http://merops.sanger.ac.uk>). Sie spielen eine entscheidende Rolle in vielen physiologischen und pathologischen Prozessen. Proteasen sind z.B. am Protein-Katabolismus, an der Blutgerinnung, am Zellwachstum, an der Zellmigration, an Entzündungsprozessen und Tumorbildung sowie an der Prozessierung von Vorläufer-Proteinen beteiligt. Letzteres führt zur Freisetzung von z.B. aktiven Hormonen und Toxinen (RAO et al., 1998). Dabei sind intrazelluläre Proteasen eher an intrazellulären Prozessen beteiligt, die extrazellulären Proteasen dagegen überwiegend an der Protein-Hydrolyse in zellfreien Umgebungen. Dies ermöglicht der Zelle die Verwertung von hydrolytisch gespaltenen Proteinbestandteilen oder dient zur Zerstörung von Gewebebestandteilen (GUPTA et al., 2002a). Aufgrund des breiten

Wirkungsspektrums sind vor allem extrazelluläre Proteasen kommerziell von Interesse und finden vielfältige Anwendungen in industriellen Prozessen (s. 1.3). Die größte Bedeutung haben dabei die mikrobiellen Proteasen (GUPTA et al., 2002b). In der Waschmittelindustrie werden Proteasen als Standardzusatz zur Entfernung von Eiweißflecken verwendet. Dabei umfasst allein der Einsatz der Proteasen in der Waschmittelindustrie 25 % der weltweit industriell genutzten Enzyme. Fast alle dieser Proteasen werden von verschiedenen Vertretern der Gattung *Bacillus* produziert (RAO et al., 1998). Auch in der Pharmaindustrie gewinnen Proteasen aufgrund von vielen nachgewiesenen Beteiligungen an pathologischen Prozessen zunehmend an Bedeutung. Die bakteriellen Toxine wie z.B. das Cholera toxin (*Vibrio cholerae*), Tetanustoxin (*Clostridium tetani*), Botulinumtoxin (*Clostridium botulinum*), Diphtherietoxin (*Corynebacterium diphtheriae*) und Anthraxtoxine (*Bacillus anthracis*) werden als inaktive Vorläuferproteine gebildet, die erst durch die von verschiedenen Proteasen katalysierte limitierte Proteolyse in ihren aktiven Zustand überführt werden (GORDON und LEPLA, 1994). Daher sind diese Proteasen Zielmoleküle in der medizinischen Forschung für die Entwicklung neuer Therapeutika (RAO et al., 1998). In der vorliegenden Arbeit wurden insgesamt 1,7 Gbp an Umwelt-DNA in einem auf Enzymaktivität basierenden Screening-Verfahren auf proteolytische Aktivität durchmustert. Es konnten 5 *E. coli*-Klone identifiziert werden, von denen zwei Klone *E. coli*/Top10/pTWP9 und *E. coli*/Top10/pTWP13 und die daraus isolierten rekombinanten Plasmide pTWP9 bzw. pTWP13 näher analysiert wurden (s. 3.3).

4.4.1 Genetische Organisation der Inserts von pTWP9 und pTWP13 und Vergleich mit korrespondierenden Genomabschnitten von verschiedenen *Xanthomonas*-Arten

Von den aus der Durchmusterung der Metagenombanken erhaltenen 5 *E. coli*-Klonen mit stabilem proteolytischen Phänotyp zeigten *E. coli*/Top10/pTWP9 und *E. coli*/Top10/pTWP13 die stärkste proteolytische Aktivität auf den Indikatoragarplatten und wurden daher weiter analysiert. Die Insertsequenzen der isolierten Plasmide pTWP9 (28.113 bp) und pTWP13 (20.028 bp) wurden

bestimmt. (s. 3.3.3). Die Analyse der genetischen Organisationen der beiden Inserts zeigte, dass diese in einem Teilbereich sehr ähnlich waren (s. Abb. 35A und 35B). Bei der Annotation der aus den offenen Leserahmen der Inserts von pTWP9 und pTWP13 abgeleiteten Genprodukte war auffällig, dass das nächste verwandte Protein oftmals aus einem Vertreter der Gattung *Xanthomonas* stammte (s. Tab. 11 und 12; 3.3.3.2). Daher wurde vermutet, dass die in pTWP9 und pTWP13 klonierte Umwelt-DNA zwei überlappende Genomabschnitte desselben Vertreters oder von eng verwandten Arten der Gattung *Xanthomonas* repräsentiert. Zur Überprüfung dieser Hypothese erfolgte ein Vergleich der genetischen Organisation des pTWP9- und pTWP13-Inserts mit den entsprechenden Genomabschnitten von fünf unterschiedlichen *Xanthomonas*-Spezies (s. Abb. 35), deren vollständigen Genomdaten in der NCBI-Datenbank hinterlegt sind. Es handelte sich dabei um die Genomabschnitte von *X. campestris* pv. *vesicatoria* str. 85-10 (Accession-Nr.: AM039952), *X. campestris* pv. *campestris* str. ATCC 33913 (Accession-Nr.: AE008922), *X. campestris* pv. *campestris* str. 8004 (Accession-Nr.: CP000050), *X. axanopodis* pv. *citri* str. 306 (Accession-Nr.: AE008923) und *X. oryzae* pv. *oryzae* KACC10331 (Accession-Nr.: AE013598). Bei allen fünf *Xanthomonas*-Arten handelt es sich um phytopathogene Organismen mit unterschiedlich breiten Wirtsspektren. Die Wirtsspezifität spiegelt sich dabei in den Namen der einzelnen Arten oder Pathovaren wider. *Xanthomonas* gelangt über Wassertropfen durch natürliche Öffnungen oder Verletzungen in den Interzellularraum der Pflanze. Dort kommt es zur Vermehrung und Ausbildung von typischen Krankheitssymptomen, den wässrigen Läsionen, die im weiteren Verlauf nekrotisch wirken und zum Welken der Pflanze führen (LEYNS et al., 1984). Vertreter der Gattung *Xanthomonas* zeichnen sich im Vergleich zu anderen phytopathogenen Bakterien wie z.B. *Xylella fastidiosa* oder *Pseudomonas syringae* durch einen hohen G+C-Gehalt der DNA aus. Bei den oben erwähnten *Xanthomonas*-Vertretern variiert der G+C-Gehalt des Genoms zwischen 63 und 67 % (LEE et al, 2005). Die G+C-Gehalte der in pTWP9 und pTWP13 inserierten metagenomischen DNA lagen mit 68,3 bzw. 69,5 % somit im Bereich der G+C-Gehalte, die für die Genome der fünf *Xanthomonas*-Arten ermittelt wurden.

Abbildung 35: (folgende Seite)
Vergleich der genetischen Organisation der Inserts von pTWP9 und pTWP13 mit korrespondierenden Genomabschnitten von fünf unterschiedlichen *Xanthomonas*-Vertretern.

Die Genomdaten von (C) *X. campestris* pv. *vesicatoria* str. 85-10 (Accession-Nr.: AM039952; Genomabschnitt: 1118626 bis 1074579), (D) *X. campestris* pv. *campestris* str. ATCC 33913 (Accession-Nr.: AE008922; Genomabschnitt: 1040326 bis 995188), (E) *X. campestris* pv. *campestris* str. 8004 (Accession-Nr.: CP000050; Genomabschnitt: 4000216 bis 4045359), (F) *X. axanopodis* pv. *citri* str. 306 (Accession-Nr.: AE008923; Genomabschnitt: 1119791 bis 1071729) und (G) *X. oryzae* pv. *oryzae* KACC10331 (Accession-Nr.: AE013598; Genomabschnitt: 3859493 bis 3902413) wurden der NCBI-Datenbank entnommen. Die Zahlen unter den Genomabschnitten der *Xanthomonas*-Vertreter repräsentieren die Gen-Position. Die Namen und Funktionen der abgeleiteten Genprodukte sind dem Anhang (s. 7.3) zu entnehmen. Die Annotation der aus den Inserts von (A) pTWP9 und (B) pTWP13 abgeleiteten Genprodukte ist in Tab. 11 und 12 dargestellt. Identische Farben kennzeichnen Gene mit derselben annotierten Funktion. Sterne symbolisieren unvollständige Gene. Rote Dreiecke zeigen Gene, die nur in den Inserts von pTWP9 und pTWP13 vorkommen. Blaue Dreiecke kennzeichnen Gene, die nur bei *Xanthomonas* auftreten. Die Gene, die für Proteasen kodieren, sind gelb hinterlegt.

Der Vergleich der genetischen Organisation der Inserts von pTWP9 und pTWP13 mit der der Genomabschnitte der fünf *Xanthomonas*-Arten zeigt, dass sehr starke Ähnlichkeiten in der Genorganisation von pTWP9 und pTWP13 zu der Genabfolge bei verschiedenen *Xanthomonas*-Vertretern vorliegen (s. Abb. 35). Das wird besonders deutlich beim Vergleich der Gene bzw. der korrespondierenden Genprodukte 1a bis 14a von pTWP9 mit den ersten 14 Genen bzw. Genprodukten von allen *Xanthomonas*-Vertretern. Die den aus diesem Bereich abgeleiteten Genprodukten zugewiesenen Funktionen stimmten ohne Ausnahmen überein. Ferner zeigten die Genprodukte große Ähnlichkeiten in der Aminosäuresequenz zu dem jeweils korrespondierenden Genprodukt. Allerdings sind auch Unterschiede festzustellen. Es gibt Gene, die nur auf den beiden Inserts von und pTWP9 und pTWP13 vorkommen wie z.B. Gen 15a bei pTWP9 oder Gen 9b bei pTWP13. Von den insgesamt 54 potentiellen Genen, die auf den Inserts von pTWP9 und pTWP13 lokalisiert sind, zeigten 9 keine Ähnlichkeit zu Genen, die auf den fünf untersuchten *Xanthomonas*-Genomabschnitten vorhanden sind. Andere Gene wiederum treten nur in den Genomabschnitten der *Xanthomonas*-Arten auf. Ein Beispiel dafür stellt ein Gen dar, welches für eine Xylanase kodiert. Dieses Gen liegt in allen fünf *Xanthomonas*-Genomabschnitten (Gen 24 bzw. bei *X. oryzae* Gen 23) und nicht in den Inserts von pTWP9 und pTWP13 vor. Insgesamt kommen bei allen *Xanthomonas*-Genomabschnitten 52 der 219 *Xanthomonas*-Gene weder bei pTWP9 noch bei pTWP13 vor (s. Abb. 35). Beim Vergleich der Genomabschnitte der unterschiedlichen *Xanthomonas*-Vertreter untereinander sind ebenfalls Unterschiede in der Genorganisation festzustellen (s. Abb. 35 D-G). So besitzen z.B. *X. campestris* pv. *campestris* str. ATCC 33913 (Gen 36 und 37), *X. campestris* pv. *campestris* str. 8004 (Gen 36 und 37) und *X. axanopodis* pv. *citri* str. 306 (Gen 41 und 42) Gene, die für einen Transkriptions-Regulator bzw. eine Hydrolase kodieren, die in den untersuchten Genregionen von *X. campestris* pv. *vesicatoria* str. 85-10 und *X. oryzae* pv. *oryzae* KACC10331 nicht vorkommen. Insgesamt kommen 6 der 219 *Xanthomonas*-Gene nur bei einer *Xanthomonas*-Art vor (s. Abb. 35). Für die klonierten Genregionen der Inserts von pTWP9 und pTWP13 wurde jeweils ein Einzelvergleich mit jedem Genomabschnitt der fünf *Xanthomonas*-Vertreter vorgenommen, indem die prozentualen Anteile an abgeleiteten Genprodukten

ohne signifikante Identität (Identitäten unter 25 %), zwischen 25 % und 49 % Identität, zwischen 50 % und 69 % Identität und ≥ 70 % Identität in der Aminosäuresequenz berechnet wurden (s. Abb.36).

Abbildung 36: Vergleiche der Übereinstimmung der aus den Inserts von pTWP9 und pTWP13 abgeleiteten Genprodukte mit den aus den Genregionen von 5 *Xanthomonas*-Arten abgeleiteten Genprodukten
 Bei pTWP9 entsprechen 100% 32 Genprodukte, bei pTWP13 entsprechen 100% 22 Genprodukte.

Für die Genprodukte des Inserts von pTWP9 konnte die größte Ähnlichkeit zu den Genprodukten des Genombereichs von *X. campestris* pv. *vesicatoria* str. 85-10 festgestellt werden. Von den 32 Genprodukten bei pTWP9 zeigen 26 (81 %) signifikante Übereinstimmungen (Identität $\geq 25\%$) zu entsprechenden Genprodukten von *X. campestris* pv. *vesicatoria* str. 85-10. Von denen weisen 14 (44 %) mehr als 70 % Identität, 11 (34 %) zwischen 50 % und 69 % Identität und 1 (3 %) zwischen 25 % und 49 % Identität auf (s. Abb. 36). Die geringste Übereinstimmung konnte zu den Genprodukten von *X. campestris* pv. *campestris* str. 8004 ermittelt werden, da nur 37 % der Genprodukte mehr als 70 % Identität aufweisen (s. Abb. 36). Die Genprodukte des Inserts von pTWP13 zeigen die größten Übereinstimmungen dagegen zu den Genprodukten von *X. axanopodis* pv. *citri* str. 30 auf. Bei 18 von 22 Genprodukten (82 %) konnten signifikante Ähnlichkeiten identifiziert werden. Davon weisen 36 % (8 Genprodukte) mehr als 70 % Identität, 14 % (3 Genprodukte) zwischen 50 % und 69 % Identität und 32 % (7 Genprodukte) zwischen 25 % und 49 % Identität auf. Keine signifikante Identität konnte bei 4 Genprodukten (18 %) festgestellt werden (s. Abb. 36). Auch für die Genprodukte des Inserts von pTWP13 konnten die geringsten Übereinstimmungen zu den entsprechenden Genprodukten von *X. campestris* pv. *campestris* str. 8004 identifiziert werden.

Die größten Abweichungen der genetischen Organisationen der Inserts von pTWP9 und pTWP13 im Vergleich zu den Genomabschnitten aller 5 *Xanthomonas*-Vertreter betrifft die Genregionen, die die Protease-Gene *mprA* und *mprB* flankieren (s. Abb. 35). Dieser Sequenzbereich ist auch der überlappenden Bereich der beiden Inserts aus pTWP9 und pTWP13 (Gene 18a bis 32a bei pTWP9 und Gene 1b bis 12b bei pTWP13). Bei vier der fünf *Xanthomonas*-Vertreter liegen drei Protease-Gene hintereinander vor. Nur in dem Genomabschnitt von *X. oryzae* pv. *oryzae* KACC10331 ist wie bei pTWP9 und pTWP13 nur ein Protease-Gen lokalisiert. Die aus *mprA* und *mprB* abgeleitete Proteasesequenzen zeigen die größten Übereinstimmungen (49 % bzw. 50 % Identität) zu einer Zink-Metalloprotease (Empl) aus *Pseudoalteromonas* sp. (s. 3.3.3 und 3.3.4). Bei Sequenzvergleichen von MprA und MprB mit den entsprechenden Genprodukten der Protease-Gene der fünf

Xanthomonas-Arten (s. Abb.35) konnten ebenfalls geringe Identitäten festgestellt werden. Bei den drei hintereinander geschalteten Protease-Genen der *Xanthomonas*-Spezies, konnten signifikante Identitäten jedoch nur zur jeweils dritten der aus diesen Genen abgeleiteten Proteinsequenzen identifiziert werden. Diese Identitäten liegen allerdings nur 25,4 bis 30,3 % (s. Tab. 21).

Tabelle 21: Vergleich der MprA- und MprB-Proteinsequenz mit den Sequenzen der entsprechenden Genprodukte aus 5 verschiedenen *Xanthomonas*-Arten.

Die Accession-Nummern der *Xanthomonas*-Genomabschnitte sind der Legende aus Abbildung 35 zu entnehmen. MprA und MprB stammen aus dieser Arbeit. Die Identitäten in der Aminosäuresequenzen wurden mit dem Programm „Emboss Pairwise Alignment“ (www.ebi.ac.uk/emboss/align) ermittelt.

<i>Xanthomonas</i>-Art und Position des Genprodukte	Identität zu MprA [%]	Identität zu MprB [%]
<i>X. campestris</i> pv. <i>vesicatoria</i> str. 85-10, <i>X.</i> Produkt Gen 29 (s. Abb. 35C)	27,7	29,3
<i>X. campestris</i> pv. <i>campestris</i> str. ATCC 33913 Produkt Gen 29,(s. Abb. 35D)	29,4	29,3
<i>X. campestris</i> pv. <i>campestris</i> str. 8004 Produkt Gen 29 (s. Abb. 35E)	30,5	30,3
<i>X. axanopodis</i> pv. <i>citri</i> str. 306 Produkt Gen 30, (s. Abb. 35F)	27,6	28,3
<i>X. oryzae</i> pv. <i>oryzae</i> KACC10331 Produkt Gen 26 (s. Abb. 35G)	25,7	25,4

Die vorgenommenen Vergleiche der genetischen Organisationen der Inserts aus pTWP9 und pTWP13 und der darauf lokalisierten Genprodukte mit den entsprechenden Genregionen und Genprodukten von *X. campestris* pv. *vesicatoria* str. 85-10, *X. campestris* pv. *campestris* str. ATCC 33913, *X. campestris* pv. *campestris* str. 8004, *X. axanopodis* pv. *citri* str. 306 und *X. oryzae* pv. *oryzae* KACC10331 zeigten große Übereinstimmungen aber auch einige Unterschiede der Inserts von pTWP9 und pTWP13 zu den untersuchten Genomabschnitten der *Xanthomonas*-Arten. Die ermittelten Unterschiede waren jedoch nicht größer als die Unterschiede, die in den untersuchten Genomabschnitten der verschiedenen *Xanthomonas*-Arten untereinander

vorlagen. Ferner lagen die G+C-Gehalte der Inserts von pTWP9 und pTWP13 im Bereich der G+C-Gehalte der *Xanthomonas*-Genome. Aus den Vergleichen kann man daher schließen, dass sowohl die in pTWP9 als auch die in pTWP13 inserierte Umwelt-DNA Genomabschnitte von Mitgliedern der Gattung *Xanthomonas* repräsentieren.

In der Genorganisation der überlappenden Bereiche der Inserts von pTWP9 und pTWP13 waren ebenfalls Unterschiede festzustellen (s. Abb. 35). Ein Alignment der Proteinsequenzen von gleichartigen Genprodukten aus dem überlappenden Bereich gab Aufschluss darüber, wie hoch die Identitäten bzw. Ähnlichkeiten dieser Genprodukte auf Aminosäuresequenzebene sind (s. Tab 22).

Tabelle 22: Vergleich der aus dem überlappenden Sequenzbereich des pTWP9- und pTWP13-Inserts abgeleiteten Genprodukte. Die Identitäten bzw. Ähnlichkeiten der Aminosäuresequenzen der Genprodukte wurden mit dem Programm „Emboss Pairwise Alignment“ (www.ebi.ac.uk/emboss/align) ermittelt. Sterne symbolisieren unvollständige Genprodukte, bei denen sich die Identitäten und Ähnlichkeiten nur auf den vorhandenen Bereich beziehen.

pTWP9			Identität/ Ähnlichkeit [%]	pTWP13		
Gen- Position	Gen- name	Annotation		Gen- Position	Gen- name	Annotation
-	-	-	-	1	<i>orf1*</i>	Hypothetisches Protein
18	<i>acyP</i>	Acylphosphatase	72 / 80	2	<i>acyP</i>	Acylphosphatase
19	<i>yneN</i>	Mögliches Thioredoxin	78 / 86	3	<i>yneN</i>	Mögliches Thioredoxin
20	<i>rbn</i>	Ribonuklease BN	75 / 83	4	<i>rbn</i>	Ribonuklease BN
21	<i>trp</i>	Tryptophan-Repressor-Bindeprotein	91 / 95	5	<i>trb</i>	Tryptophan-Repressor-Bindeprotein
22	<i>orf22</i>	Konserviertes hypothetisches Protein	74 / 83	6	<i>orf6</i>	Konserviertes hypothetisches Protein
23	<i>orf23</i>	Hypothetisches Protein	-	-	-	-
24	<i>orf24</i>	Hypothetisches Protein	-	-	-	-
25	<i>orf25</i>	Hypothetisches Protein	-	-	-	-
26	<i>orf26</i>	Asparaginase	94 / 98	7	<i>orf7</i>	Asparaginase
27	<i>orf27</i>	Konserviertes hypothetisches Protein	-	-	-	-
28	<i>orf28</i>	Hypothetisches Protein	-	-	-	-
29	<i>mprA</i>	Metalloprotease	92 / 96	8	<i>mprB</i>	Metalloprotease
-	-	-	-	9	<i>orf9</i>	Hypothetisches Protein
30	<i>orf30</i>	Konserviertes hypothetisches Protein	83 / 91	10	<i>orf10</i>	Konserviertes hypothetisches Protein
31	<i>orf31</i>	Konserviertes hypothetisches Protein	68 / 77	11	<i>orf11</i>	Konserviertes hypothetisches Protein
32	<i>orf32</i>	Mögliche NAD(P)-Transhydrogenase, β -UE	97 / 98	12	<i>orf12*</i>	Mögliche NAD(P)-Transhydrogenase, β -UE

Der Vergleich belegt, dass es im überlappenden Bereich 9 Genprodukte gibt, die sehr hohe Übereinstimmungen zueinander aufweisen (68 bis 97 % Identität bzw. 77 bis 98 % Ähnlichkeit). Hierzu gehören z.B. die Metalloproteasen (92 % Identität), die Asparaginasen (94 % Identität) oder die potentiellen β -Untereinheiten der NAD(P)-Transhydrogenasen (97 % Identität). Allerdings gibt es auch Genprodukte, die entweder nur bei pTWP9 (Genprodukte von *orf23* bis *orf25*, *orf27* und *orf28*) oder nur bei pTWP13 (Genprodukte von *orf1* und *orf9*) vorhanden sind. Dies zeigt an, dass die in pTWP9 und pTWP13 inserierte Umwelt-DNA nicht aus identischen Organismen sondern wahrscheinlich aus verschiedenen Stämmen der gleichen Art oder aus unterschiedlichen Arten der gleichen Gattung stammt. Bei der Gattung handelt es sich wie oben beschrieben wahrscheinlich um *Xanthomonas*.

4.4.2 Charakterisierung der Proteasen MprA und MprB

Die Aminosäuresequenzen von MprA und MprB, die beide 862 Aminosäuren umfassen, zeigen die größten Identitäten (49 % bzw. 50 %) zu der Proteinsequenz der Zink-Metalloprotease Empl aus *Pseudoalteromonas* sp. A28 (s. Tab 11 und 12 in 3.3.3.2). Der Stamm wurde aus einer Meerwasserprobe aus der Ariake Sea vor der Küste Japans isoliert. Dort trägt die Kieselalge *Skeletonema costatum* zu einer unerwünschten Blütenbildung im Wasser bei. Durch Untersuchungen konnte gezeigt werden, dass Empl algizide Wirkung besitzt und das Absterben der Kieselalge hervorruft (LEE et al., 2000 und 2002). Bei vielen Proteasen mit pathologischer Wirkung handelt es sich um Zink-abhängige Metalloproteasen (MIYOSHI und SHINODA, 2000). Ein Beispiel für solche Zink-abhängige Metalloproteasen ist das von dem humanpathogenen Bakterium *Vibrio vulnificus* produzierte Vibriolysin, welches zu hämorrhagischer Gewebeerstörung durch Auflösung der zellulären Basalmembran führt (MIYOSHI und SHINODA, 2000). Für MprA und MprB konnte aufgrund spezifischer Motive in der Aminosäuresequenz eine Zuordnung zu den Zink-abhängigen Metalloproteasen erfolgen (s. Tab 13; 3.3.4). Mit Hilfe der

MEROPS-Proteasedatenbank wurden beide Enzyme in die Familie M4 der Zink-abhängigen Metalloproteasen eingestuft, zu der auch Empl gehört. Alle Mitglieder dieser Familie zeichnen sich durch das Vorkommen von zwei Konsensus-Motiven in ihrer Aminosäuresequenz aus. Beide Motive sind in der MprA- und MprB-Sequenz vollständig vorhanden (s. Abb. 37). Bei einem der Motive handelt es sich um das so genannte Zink-Bindemotiv, welches die Aminosäuren HEXXH beinhaltet. Das zweite Motiv umfasst die Aminosäuren GXXNEXXSD und befindet sich ca. 15 bis 20 Aminosäuren vom Zink-Bindemotiv in Richtung des C-terminalen Endes entfernt (MOYOSHI und SHINODA, 2000). Dabei fungieren die beiden Histidin-Reste des Zink-Bindemotivs als Zink-Liganden. Als dritter Zink-Ligand tritt das Glutamat aus dem zweiten Konsensus-Motiv auf und als vierter Zink-Ligand fungiert ein Wassermolekül (RAWLINGS und BARRETT, 1995).

Konsensus	HExxH	GxxNExxSD
MprA	348 STFYPLVS-ADVAGHEVSHGFTEQHSNLTYSGQSGGINEAFSDMGGEATEHYWKGSNDFL	407
MprB	348 STFYPLVS-SDVAGHEVSHGFTEQHSNLTYSGQSGGINEAFSDMGGEATEHYWKGSNDFL	407
MprI	270 NTFYPLVS-LDVSAHEVSHGFTEQNSGLVYRYKSGGLNEAFSDMAGEAAEFYMKGSNDWL	329
EmpI	273 NTFYPLVS-LDVSAHEVSHGFTEQNSGLIYSGKSGGLNEAFSDMAGEAAEFYMKGSNDWL	332
Vvp	269 STFYPLVD-INVSAHEVSHGFTEQNSGLIYSNMSGGMNEAFSDIAGEAAEFYMKGSVDWI	328
Thermolysin	299 QTFIPLSGGIDVVAHELTHAVTDYTAGLIYQESGAINEAISDFIGFTLVEFYANKNPDWE	359
	. ** * . : * . ** : : . * : : . * * * . : : : : : * . * : : . * :	

Abbildung 37: Aminosäurealignment im Bereich der für M4-Metalloproteasen typischen Konsensus-Motive.

In dem multiplen Alignment werden die Aminosäuresequenzbereiche von Mitgliedern der Familie M4 der Zink-abhängigen Metalloproteasen gezeigt, die das konservierte Zink-Bindemotiv (HEXXH) und ein zweites Konsensus-Motiv (GXXNEXXSD) beinhalten. MprA und MprB: in dieser Arbeit identifizierte Metalloproteasen; MprI: Metalloprotease aus *Alteromonas* sp., Accession-Nr.: BAB79615.1; Empl: Metalloprotease aus *Pseudoalteromonas* sp., Accession-Nr.: BAB85124.1; Vvp: Metalloprotease aus *Vibrio vulnificus*, Accession-Nr.: BAC22940.1; Thermolysin: Metalloprotease aus *Bacillus thermoproteolyticus*, Accession-Nr.: CAA01492.1. Das Alignment wurde mit dem Programm ClustalW erstellt. (*) symbolisiert identische Aminosäuren; (:) symbolisiert konservierte Substituenten; (.) symbolisiert semi-konservierte Substituenten.

Alle Mitglieder der Familie M4 der Zink-abhängigen Metalloproteasen haben neben den zwei konservierten Aminosäuremotiven weitere Gemeinsamkeiten. Sie werden alle als inaktive Vorläufer-Proteine gebildet und es handelt sich fast ausschließlich um sekretierte Proteine (RAWLINGS und BARRETT, 1995 und 2004). Diese erlangen ihre Aktivität durch mehrere Prozessierungsschritte, in denen N-terminale und, wenn vorhanden, auch C-terminale Prosequenzen meistens autokatalytisch abgespalten werden. Daher beinhalten die M4-Metalloproteasen unterschiedliche konservierte Domänen in ihrer Aminosäuresequenz und benötigen zusätzlich für den Export aus der Zelle eine N-terminale Signalsequenz. Auch für MprA und MprB konnten mehrere für M4-Metalloproteasen typische konservierte Domänen sowie eine N-terminale Signalsequenz identifiziert werden. Die Domänenstrukturen der beiden Metalloproteasen sind in Abbildung 19 dargestellt (s. 3.3.4). Die Aminosäuresequenzen von MprA und MprB beinhalten jeweils eine N-terminale Signalsequenz, ein N-terminales Propeptid, eine katalytische Domäne, eine strukturelle Domäne (α -helikal), je zwei C-terminale Propeptide und eine für M4-Metalloproteasen untypische regulatorische P-Domäne. Den einzelnen Domänen werden dabei unterschiedliche Funktionen zugesprochen. Für verschieden N-terminale Propeptide von M4-Metalloproteasen wie z.B. bei denen von Pseudolysin und Thermolysin konnte gezeigt werden, dass die Prosequenzen als intramolekulare Chaperone fungieren und somit wichtige Funktionen bei der Faltung der Proteine übernehmen (O'DONOHUE und BEAUMONT, 1996; MARIE-CLAIRE et al., 1999). Zusätzlich werden ihnen inhibitorische Wirkungen eingeräumt, die eine vorzeitige Ausprägung der Protease-Aktivität im Cytoplasma verhindern (KESSLER und SAFRIN, 1994). In der katalytischen Domäne ist das Zink-Bindemotiv lokalisiert (RAWLINGS und BARRETT, 2004). C-terminale Propeptid-Domänen sind nicht bei allen M4-Metalloproteasen verbreitet. Thermolysin und Pseudolysin verfügen z.B. über keine C-terminalen Propeptide (KUBO und IMANAKA, 1988; BEVER und IGLEWSKI, 1988). Wenn diese jedoch auftreten, können sie einzeln oder wie bei MprA (diese Arbeit), MprB (diese Arbeit), MprI aus *Alteromonas* sp. (MIYAMOTO et al., 2002) und Empl aus *Pseudoalteromonas* sp. Str. A28 (LEE et al., 2002) auch mehrfach vorkommen. Die C-terminalen bakteriellen Propeptide werden erst autokatalytisch prozessiert, wenn die noch inaktiven

Vorläufer-Proteine aus der Zelle exportiert wurden. Für die Pro-Aminopeptidase prozessierende Protease (PA-Protease) konnte gezeigt werden, dass im Periplasma zuerst die N-terminale Prosequenz abgespalten wird. Das C-terminale Propeptid wird nach Translokation des Vorläufer-Proteins durch die äußere Membran prozessiert (NIRASAWA et al., 2002). Es wird vermutet, dass C-terminale Prosequenzen Funktionen bei der Translokation der Proteine durch die äußere Membran übernehmen (KIM und MATSUZAWA, 2000). Die regulatorische P-Domäne, die allerdings nur zu ca. 60 % in MprA und MprB vorhanden ist, kommt meistens in Subtilisin-ähnlichen Proprotein-Konvertasen und anderen eher eukaryotischen Proteasen vor. Es wird angenommen, dass regulatorische P-Domänen Chaperon-Funktionen ausüben oder in posttranslationalen Modifikationen involviert sind (GLUSCHANKOF und FULLER, 1994).

In der vorliegenden Arbeit wurde überprüft, welche Domänen bei MprA und MprB zur Bildung einer aktiven Protease erforderlich sind. Aus diesem Grund wurden sowohl MprA und MprB als auch verschiedene Derivate von beiden Proteasen konstruiert, in *E. coli* BL21 heterolog produziert und auf Protease-Aktivität getestet (s. 3.4.1 und 3.4.2). Die Derivate zeichneten sich durch unterschiedliche Deletionen von einzelnen Domänen aus. MprA-K1 und MprB-K1 wiesen C-terminale Deletionen auf (s. Abb. 22; 3.4.1) und verfügten somit über keine C-terminalen Propeptide und keine regulatorische P-Domäne (s. Abb. 22; 3.4.1). Das Derivat MprB-K2 wurde N-terminal deletiert und besaß kein Signalpeptid und kein N-terminales Propeptid. Bei MprB-K3 erfolgten Deletionen sowohl am N- als auch am C-Terminus. Dieses Derivat umfasste nur noch die katalytische Domäne und die α -helikale Domäne (s. Abb. 22; 3.4.1). Durch Western Blot-Analyse konnte gezeigt werden, dass MprA und MprB sowie alle Derivate in *E. coli*-BL21-Zellen heterolog produziert wurden (s. Abb. 24; 3.4.3). Weiterhin konnte nachgewiesen werden, dass das N-terminale Propeptid und die Signalsequenz essentiell für die Ausprägung der proteolytischen Aktivität sind, da nur rekombinante *E. coli*-Stämme, die MprA, MprB, MprA-K1 und MprB-K1 produzierten, proteolytische Aktivität aufwiesen (s. Abb. 23; 3.4.2). Die Messungen der Protease-Aktivität von allen konstruierten Derivaten und den unveränderten Enzymen im Kulturüberstand im

Vergleich zu der im Zellextrakt belegten, dass die proteolytische Aktivität erst im Kulturüberstand und somit nach Export aus den *E. coli*-Zellen nachweisbar ist (s. Tab. 16; 3.4.2). Diese Ergebnisse bestätigten, dass MprA und MprB in einer für M4-Metalloproteasen typischen Weise vom Wirtstamm *E. coli* BL21 in das Medium sekretiert werden und dort erst ihre proteolytische Aktivität erlangen. Hierfür sind die Signalsequenz und das N-terminale Propeptid erforderlich. Das meistens bei anderen M4-Metalloproteasen autokatalytisch abgespaltenen C-terminalen Propeptid und die regulatorische P-Domäne scheinen dagegen nicht erforderlich für die proteolytische Aktivität von MprA und MprB zu sein. Alle Derivate mit C-terminaler Deletion (MprA-K1 und MprB-K1) zeigten nur geringfügig niedrigere spezifische Aktivitäten wie MprA und MprB (s. Tab. 16; 3.4.2). Die nicht-essentielle Funktion der C-terminalen Propeptide bei M4-Metalloproteasen war angezeigt, da es auch wie zuvor beschrieben Mitglieder dieser Familie gibt, die kein C-terminales Propeptid besitzen. Ferner stimmen die mit den Derivaten von MprA und MprB erhaltenen Ergebnisse gut mit denen überein, die bei einer Studie mit der PA-Protease aus *Aeromonas caviae* T-64 erzielt wurden. Es konnte dabei ebenfalls durch Deletions-Experimente belegt werden, dass das N-terminale Propeptid essentiell für die Aktivität der M4-Metalloprotease ist und das C-terminale Propeptid dafür nicht benötigt wird (NIRASAWA et al., 1999).

Des Weiteren erfolgte innerhalb der vorliegenden Arbeit eine biochemische Charakterisierung von MprA, MprB, MprA-K1 und MprB-K1. Hierzu wurden partiell aus dem Kulturüberstand der entsprechenden *E. coli*-Klone gereinigte Proteine eingesetzt, die ca. 8,5-fach angereichert waren (s. Tab. 18; 3.4.5). Die Proteasen MprA, MprB und die C-terminal deletierten Derivate MprA-K1 und MprB-K1 wiesen Aktivität in einem Temperatur-Bereich von 15 bis 75 °C auf (s. Abb. 29 und 3.4.6). Maximale spezifische Protease-Aktivitäten konnten für MprA und MprA-K1 bei 65 °C und für MprB und MprB-K1 bei 60 °C ermittelt werden. Im Temperaturbereich zwischen 45 °C und 65 °C sanken die relativen Aktivitäten im Vergleich zur maximalen Aktivität nicht unter 55 % ab. Ein ähnliches Temperaturoptimum von 60 °C konnte für die M4-Metalloprotease MprI aus *Alteromonas* sp. Str. 0-7 bestimmt werden (MIYAMOTO et al., 2002). Untersuchungen zur Temperaturstabilität bei 50, 60 und 70 °C von MprA, MprB,

MPrA-K1 und MPrB-K1 zeigten, dass nach drei Stunden Inkubation bei 50 und 60 °C die spezifischen Aktivitäten bei allen Enzymen mindestens noch 31 % im Vergleich zur Ausgangsaktivität betrugen. Im Gegensatz dazu führte schon eine 30 minütige Inkubation bei 70 °C nahezu zur vollständigen Inaktivierung aller Enzyme. (s. Abb. 30 und Kap. 3.4.7). Ähnliche Untersuchungen wurden für die Metalloprotease CP1 aus *Pseudoalteromonas* sp. str. CP76, welche ein Temperaturoptimum bei 55 °C aufweist, durchgeführt. Eine Hitzeinkubation bei 55 °C führte nach 180 min zur vollständigen Inaktivierung, während bei 50 °C nach 3 h noch ca. 30 % der Ausgangsaktivität vorhanden waren (SANCHEZ-PORRO et al., 2003). Weiterhin zeigten MPrA, MPrB, MPrA-K1 und MPrB-K1 über einen weiten pH-Bereich von 6 bis 11 proteolytische Aktivität. Das Optimum lag in allen Fällen bei pH 8 (s. Abb. 31 und 3.4.8). Zwischen pH 7 und 9 betrug die spezifische Aktivität bei allen Enzymen mindestens 75 % der bei pH 8 gemessenen Maximalaktivität. Erstaunlicherweise konnte bei pH 11 noch proteolytische Aktivität detektiert werden. Somit sind MPrA und MPrB als alkalische Metalloproteasen einzustufen, wobei die meisten der alkalischen Metalloproteasen nur proteolytische Aktivität im pH-Bereich von 7 bis 9 zeigen (MIYAMOTO et al., 2002).

Experimente mit typischen Protease-Inhibitoren bestätigten ebenfalls die Zugehörigkeit von MPrA und MPrB zu den Zink-abhängigen Metalloproteasen (s. Abb. 32 und 3.4.9). Die Inkubation von MPrA, MPrB, MPrA-K1 und MPrB-K1 in Gegenwart der Metallchelatoren EDTA bzw. 1,10-Phenantrolin führte zur vollständigen Inaktivierung bei allen Enzymen. Die Sensitivität gegenüber diesen Verbindungen ist charakteristisch für Metalloproteasen (NEGRETE-ABASCAL et al., 2003; KOOI et al., 2005). Die Inkubation der Enzyme mit dem für Serin-Proteasen typischen Inhibitor PMSF hatte nahezu keinen Effekt auf die Protease-Aktivität von MPrA, MPrB, MPrA-K1 und MPrB-K1. Vergleichbare Ergebnisse wurden auch mit den Metalloproteasen MprI aus *Alteromonas* sp. und Empl aus *Pseudoalteromonas* sp. erzielt (MIYAMOTO et al., 2002; LEE et al., 2002).

Die im Rahmen der biochemischen Untersuchungen erhaltenen Ergebnisse bestätigten die Eingruppierung von MPrA und MPrB in die Metalloprotease-

Familie M4. Für beide Enzyme gilt es, in zukünftigen Arbeiten die Reinigung zu optimieren, da diese im Rahmen der vorliegenden Arbeit jeweils nur partiell durchgeführt werden konnte. Eine Möglichkeit dafür besteht in der Auswahl von anderen Reinigungsverfahren. Voraussetzung hierfür ist jedoch eine Verbesserung der heterologen Enzymproduktion, da nur sehr geringe Mengen der Proteasen von *E. coli* produziert und in den Kulturüberstand sekretiert wurden. Dementsprechend stand nur wenig aktive und prozessierte Protease als Ausgangsmaterial für die Reinigung zur Verfügung. Denkbar wäre zur Steigerung der Protease-Produktion eine heterologe Expression von *mprA* und *mprB* in alternativen Wirten. Erste Versuche wurden dafür schon durchgeführt. Es konnte gezeigt werden, dass die Protease-Produktion in den Wirten *Bacillus subtilis*, *Streptomyces coelicolor* und *Streptomyces lividans* zu keiner Steigerung im Vergleich zur Produktion in *E. coli* führte. Alternativ wird zurzeit versucht, MprA und MprB in *X. campestris* zu produzieren. Die Verwendung dieses Wirtes ist angezeigt, da die in pTWP9 und pTWP13 inserierte Umwelt-DNA und somit auch *mprA* und *mprB* wahrscheinlich aus Organismen der Gattung *Xanthomonas* stammen.

5 Zusammenfassung

1. Zur Herstellung von Metagenombanken wurde die mikrobielle DNA aus verschiedenen Boden- und Sedimentproben nach der Methode der direkten Lyse isoliert, gereinigt, partiell verdaut und anschließend in die Vektoren pBluescript SK⁺ und pBluescript SK II⁺ inseriert. Im Rahmen der vorliegenden Arbeit wurde eine auf T/A-Klonierung basierende Methode zur effizienten Klonierung von Umwelt-DNA, die noch mit Matrixbestandteilen wie z.B. Huminsäuren verunreinigt ist, entwickelt. Es konnte mit diesem Verfahren die Klonierungseffizienz 2- bis 5-fach und die Inserthäufigkeit um 6 % bis 19 % gesteigert werden. Insgesamt wurden 12 Metagenombanken konstruiert, die zusammengefasst 598.000 rekombinante Plasmide beinhalten. Diese repräsentieren 2,2 Gbp an klonierter Umwelt-DNA.

2. Zur Identifizierung von Genen, die Butanol-Dehydrogenase-Aktivität vermitteln, wurden in einem auf Enzymaktivität basierenden Screening-Verfahren 53.000 *E. coli*-Klone (130 Mbp Umwelt-DNA) durchmustert. Als Wirtstamm diente dabei die *adhE*-negative *E. coli*-Mutante WL3. Es konnten zwei positive *E. coli*-Klone bei der Durchmusterung identifiziert werden. Die daraus isolierten rekombinanten Plasmide wurden als pTWB3 und pTWB4 bezeichnet.

3. In den Rohextrakten von *E. coli* WL3/pTWB3 und *E. coli* WL3/pTWB4 konnte eine NADPH-abhängige Butanol-Dehydrogenase-Aktivität nachgewiesen werden. Die spezifische Aktivität für die Bildung von Butanol aus Butyraldehyd im Rohextrakt von *E. coli* WL3/pTWB4 war mit 0,18 U/mg höher als die im Rohextrakt von *E. coli* WL3/pTWB3 gemessene von 0,08 U/mg. Im Weiteren wurde die Inserts von pTWB3 und pTWB4 sequenziert und analysiert. Dabei konnten jeweils zwei offene Leserahmen festgelegt werden, die als *orf31* und *orf32* (pTWB3) sowie als *orf41* und *orf42* (pTWB4) bezeichnet wurden. Die aus den vier Genen abgeleiteten Aminosäuresequenzen zeigten keinerlei Ähnlichkeiten zu Proteinsequenzen von bekannten Alkohol-Dehydrogenasen oder Oxidoreduktasen. Durch Subklonierungen und anschließenden Nachweis

der Butanol-Dehydrogenase-Aktivität in Rohextrakten der korrespondierenden *E. coli*-Klone konnte bei pTWB4 *orf42* als das Gen identifiziert werden, welches verantwortlich für die Butanol-Dehydrogenase-Aktivität von *E. coli*/WL3/pTWB4 ist. Die Aminosäuresequenz des *orf42*-Genprodukts zeigte große Übereinstimmung zu der Proteinsequenz einer Oligoendopeptidase aus *Bacillus licheniformis*. Es ist daher wahrscheinlich, dass die Substratumsetzung von Butyraldehyd zu Butanol durch *E. coli*/WL3/pTWB4 auf einem indirekten Weg erfolgt, wobei das *orf42*-Genprodukt eine Umsetzung des Substrats durch wirtseigene Proteine auf bisher unbekannte Weise ermöglicht.

4. Die Identifizierung von Genen aus Metagenombanken, die proteolytische Aktivität vermitteln, erfolgte aktivitätsbasierend auf Skim-Milk-Indikatoragar. Für das Screening wurden die im Rahmen der Arbeit konstruierten und weitere, in der Arbeitsgruppe vorhandene, Metagenombanken eingesetzt. Insgesamt wurden 533.000 *E. coli*-Klone (1,7 Gbp Umwelt-DNA) auf proteolytische Aktivität durchmustert. Dabei konnten fünf *E. coli*-Klone mit stabiler proteolytischer Aktivität identifiziert werden. Zwei der fünf Klone wurden für eine weitere Charakterisierung aufgrund ihrer starken proteolytischen Aktivität ausgewählt. Die aus den beiden Klonen isolierten rekombinanten Plasmide wurden als pTWP9 und pTWP13 bezeichnet und weisen eine Insertgröße von 28.113 bp bzw. 20.018 bp auf.

5. Sequenzanalysen zeigten, dass auf dem pTWP9-Insert 32 und auf dem pTWP13-Insert 22 potentielle Gene vorhanden sind. Die beiden Inserts weisen in einem kleinen Bereich, in dem auch die Protease-Gene *mprA* (pTWP9) und *mprB* (pTWP13) lokalisiert sind, Überlappungen auf. Die überlappenden Genregionen waren jedoch nicht vollkommen identisch und auch *mprA* und *mprB* Gene zeigen auf DNA-Ebene nur 91,5 % Identität. Weiterhin wiesen die genetischen Organisationen der Inserts aus pTWP9 und pTWP13 signifikante Ähnlichkeiten zu Genomabschnitten von fünf unterschiedlichen *Xanthomonas*-Arten auf. Daher ist es wahrscheinlich, dass die in pTWP9 und pTWP13 inserierte Umwelt-DNA aus verschiedenen Vertretern der Gattung *Xanthomonas* stammt.

6. Die Aminosäuresequenzen der aus *mprA* und *mprB* abgeleiteten Genprodukte besitzen 49 % bzw. 50 % Identität und 65 % bzw. 66 % Ähnlichkeit zu einer Zink-abhängigen Metalloprotease aus *Pseudoalteromonas* sp. str. A28. Durch weitere Sequenzanalysen konnten MprA und MprB in die Familie M4 der Zink-abhängigen Metalloproteasen eingestuft werden. Diese Familie zeichnet sich dadurch aus, dass alle Proteasen als inaktive Vorläufer-Proteine gebildet werden und es sich fast ausschließlich um Exoenzyme handelt. Die Proteinsequenzen von MprA und MprB verfügen jeweils über für Proteasen der M4-Familie typische Domänen. Diese umfassen ein N-terminales Propeptid, eine katalytische Domäne mit dem konservierten Aminosäuremotiv HEXXH, eine α -helikale-Domäne sowie zwei C-terminale Prosequenzen. Ferner befinden sich am N-Terminus eine Signalsequenz für den Proteinexport und am C-Terminus eine regulatorische P-Domäne, die für andere Mitglieder der M4-Metalloproteasen bisher nicht beschrieben ist.

7. Die Gene *mprA* und *mprB* sowie unterschiedliche Derivate mit Deletionen in Genbereichen, die für die Signalsequenz, das N-terminale Propeptid oder die C-terminalen Prosequenzen kodieren, wurden in den Expressionsvektor pET101/D-TOPO kloniert und die Genprodukte heterolog in *E. coli* BL21-Zellen produziert. Analysen der proteolytischen Aktivität von entsprechenden Klonen und Proteinpräparationen belegten, dass sowohl das N-terminale Propeptid als auch die Signalsequenz von essentieller Bedeutung für die Protease-Aktivität von MprA bzw. MprB sind. Deletionen der C-terminalen Domänen hatten dagegen keinen signifikanten Einfluss auf die proteolytische Aktivität. Ferner zeigten nur in den Kulturüberstand exportierte Enzyme Aktivität. Im Weiteren wurden die durch hydrophobe Interaktionschromatographie partiell gereinigten Enzyme MprA und MprB sowie die im Bereich der C-terminalen Prosequenzen deletierten Derivate MprA-K1 und MprB-K1 biochemisch charakterisiert. Die Enzyme konnten die Proteasesubstrate Casein, Skim-Milk, Gelatine, Azoalbumin, Azocasein und Rinderserumalbumin umsetzen. Die maximalen proteolytischen Aktivitäten von MprA und MprA-K1 wurden bei 65 °C und von MprB und MprB-K1 bei 60 °C erreicht. Das pH-Optimum der Protease-Aktivität war bei allen Enzymen pH 8.

6 Literaturverzeichnis

AMANN, R.I., LUDWIG, W., SCHLEIFER, K.H., (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59: 143-169.

ARNAU, J., JORGENSEN, F., MADSEN, S.M., VRANG, A., ISRAELSEN, H., (1998). Cloning of the *Lactococcus lactis adhE* gene, encoding a multifunctional alcohol dehydrogenase, by complementation of a fermentative mutant of *Escherichia coli*. *J. Bacteriol.* 180: 143-169.

ARP, D.J., (1999). Butane metabolism by butane-grown "*Pseudomonas butanovora*". *Microbiology* 145: 1173-1180.

ASANUMA, N., YOSHII, T., HINO, T., (2004). Molecular characteristics and transcription of the gene encoding a multifunctional alcohol dehydrogenase in relation to the deactivation of pyruvate formate-lyase in the ruminal bacterium *Streptococcus bovis*. *Arch. Microbiol.* 181: 122-128.

BAKKEN, L.R., LINDAHL, V., (1995). Separation and purification of bacteria from soil. In „Nucleic Acids in the Environment, Methods and Applications“ (J.T. Trevors and J.D. van Elsas, eds.), 8-27, Springer-Verlag, Berlin.

BEVER, R.A., IGLEWSKI, B.H., (1988). Molecular characterization and nucleotide sequence of the *Pseudomonas aeruginosa* elastase structural gene. *J. Bacteriol.* 170: 4309-4314.

BLUM, H., BEIER, H., GROSS, H. J., (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8: 93- 99.

BÖCKELMANN, U., MANZ, W., NEU, T.R., SZEWZYK, U., (2000). Characterization of the microbial community of lotic organic aggregates ('river snow') in the Elbe River of Germany by cultivation and molecular methods. *FEMS Microbiol. Ecol.* 33: 157-170.

BONFIELD, J.K., SMITH, K., STADEN, R., (1995). A new DNA sequence assembly program. *Nucleic Acids Res.* 24: 4992-4999.

BONFIELD, J.K., STADEN, R., (1996). Experiment files and their application during large-scale sequencing projects. *DNA Sequence* 6: 109-117.

BRADFORD, M.M., (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* 72: 248-254.

BRADY, S.F., CLARDY, J., (2000). Long-chain N-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. *J. Am. Chem. Soc.* 122: 12903-12904.

BRAID, M.D., DANIELS, L.M., KITTS, C.L., (2003). Removal of PCR inhibitors from soil DNA by chemical flocculation. *J. Microbiol. Methods* 52: 389-393.

BRAUN, P., de GROOT, A., BITTER, W., TOMMASSEN, J., (1998). Secretion of elastinolytic enzymes and their propeptides by *Pseudomonas aeruginosa*. *J. Bacteriol.* 180: 3467-3469.

BREITBART, M., SALAMON, P., ANDRESEN, B., MAHAFFY, J.M., SEGALL, A.M., MEAD, D., AZAM, F., ROHWER, F., (2002). Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. Usa* 99: 14250-14255.

BRUCE, K.D., HORNS, W.D., HOBMAN, J.L., OSBORN, A.M., STRIKE, P., RITCHIE, D.A., (1992). Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. *Appl. Environ. Microbiol.* 58: 3413-3416.

CALDWELL, D.E., WOLFAARDTS, M., KOBER, D.R., LAWRENCE, J.R., (1997). Cultivation of microbial consortia and communities, p.79-90. In Hurst, J.K., Knudsen, G.R., Mc Inerney, M.J., Stetzenbach, L.D., Walter, M.V., (ed.), *Manual of environmental microbiology*. ASM Press, Washington, USA

CONWAY, T., SEWELL, G.W., OSMAN, Y.A., INGRAM, L.O., (1987). Cloning and sequencing of the alcohol dehydrogenase II gene from *Zymomonas mobilis*. *J. Bacteriol.* 169: 2591-2597.

COSTILOW, R.N., (1981). Biophysical factors in growth. In: Gerhardt P., Murray G. E., Costilow N., Nester E. W., Wood W. A., Krieg R., Philips G. B. (Hrsg.) *Manual of methods for general bacteriology*, 66-78. American Society for Microbiology, Washington D.C., USA.

COURTOIS, S., FROSTEGARD, A., GORANSSON, P., DEPRET, G., JEANIN, P., SIMONET, P., (2001). Quantification of bacterial subgroups in soil: Comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation. *Environ. Microbiol.* 3: 431-439.

DAILLY, Y.P., BUNCH, P., CLARK, D.P., (2000). Comparison of the fermentative alcohol dehydrogenase of *Salmonella typhimurium* and *Escherichia coli*. *Microbios* 103: 179-196.

DANIEL, R., (2002). Construction of environmental libraries for functional screening of enzyme activity, p. 63-78. In S. Brakmann and K. Johnson (Ed.), *In: Direct molecular evolution of proteins..* Wiley-VCH Verlag GmbH, Weinheim, Germany.

DANIEL, R., (2004). The soil metagenome – a rich resource for the discovery of novel natural products. *Current Opinion In Biotechnology* 15: 199-204.

DANIEL, R., (2005). The metagenomics of soil. *Nature Rev. Microbiol.* 3 (6): 470-47.

DE LIPTHAY, J.R., ENZINGER, C., JOHNSEN, K., AMAND, J., SØRENSEN, S.J., (2004). Impact of DNA extraction method on bacterial community composition measured by denaturing gradient gel electrophoresis. *Soil Biol. Biochem.* 36: 1607-1614.

DOWER, W.J., MILLER, J.F., RAGSDALE, C.W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* 16: 6127- 6145.

DUMONT, M.G., MURRELL, J.C., (2005). Stable isotope probing – linking microbial identity to function. *Nature Rev. Microbiol.* 3 (6): 499-504.

DÜRRE, P., KUHN, A., GOTTWALD, M., GOTTSCHALK, G., (1987). Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol.* 26: 268-272

DÜRRE, P., BAHL, H., GOTTSCHALK, G., (1988). Membrane processes and product formation in anaerobes. In: Erickson, L.E., und Fung. D.Y.-C. (ed) *Handbook of Anaerobic Fermentations*. Marcel Dekker Inc., New York. USA: 187-206.

DÜRRE, P., (1998). New insights and novel developments in clostridial acetone/butanol/isopropanol fermentation. *Appl. Microbiol. Biotechnol.* 49: 639-648.

DÜRRE, P., (2005). Formation of solvents in clostridia: In: Dürre P (ed) *Handbook on Clostridia*, CRC Press-Taylor and Francis Group, Boca Raton, USA, pp 671-693.

DÜRRE, P., (2005b). Abfallstoffe als Nahrung-Entwicklung eines biotechnologischen Prozesses zur Herstellung von Butanol aus billigem Synthesegas. *GenomXPress*, Sonderausgabe September 2005. Dürre, P (ed.). 28.

ENTCHEVA, P., LIEBL, W., JOHANN, A., HARTSCH, T., 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.

ESCHENFELD, W.H., STOLZ, L., ROSENBAUM, H., KHAMBATTA, Z.S., QUAITE-RANDALL, E., WU, S., KILGORE, D.C., TRENT, J.D., DONELLY, M.I., (2001). DNA from uncultured organisms as a source of 2,5-diketo-D-gluconic acid reductases. *Appl. Environ. Microbiol.* 67: 4206-4214.

FISCHER, R.J., HELMS, J., DÜRRE, P., (1993). Cloning, sequencing and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. *J. Bacteriol.* 175: 6959-6969.

FONTAINE, L., MEYNIAL-SALLES, I., GIRBAL, L., YANG, X., CROUX, C., SOUCAILLE, P., (2002). Molecular characterization and transcriptional analysis of *adhE2*, the gene encoding NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. 184: 821-830.

FORTIN, N., BEAUMIER, D., LEE, K., GREER, C.W., (2004). Soil washing improves the recovery of total community DNA from polluted and high organic content sediments. J. Microbiol. Methods 56: 181-191.

GABOR, E.M., de VRIES, E.J., JANSSEN D.B., (2003). Efficient recovery of environmental DNA for expression cloning by indirect extraction methods. FEMS Microbiol. Ecol. 44: 153-163.

GABOR, E.M., de VRIES, E.J., 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 (9): 948-58.

GAVRILESCU, M., CHISTI, Y., (2005). Biotechnology- a sustainable alternative for chemical industry. Biotechnol. Adv. 23 (7-8): 471-479.

GILLESPIE, D.E., BRADY, S.F., BETTERMANN, A.D., CIANCOTTO, N.P., LILES, M.R., RONDON, M.R., CLARDY, J., GOODMAN, R.M., HANDELSMAN, J. (2002). Isolation of antibiotics turbomycin A and B from metagenomic library of soil microbial DNA. Appl. Environ. Microbiol. 68: 4301-4306.

GIOVANNONI, S.J., DE LONG, E.F., SCHMIDT, T.M., PACE, N.R., (1990). Tangential flow filtration and preliminary phylogentic analysis of marine picoplankton. Appl. Environ. Microbiol. 56: 2572-2575.

GLUSCHANKOF, P., FULLER, R.S., (1994). A C-terminal domain conserved in precursor processing proteases is required for intramolecular N-terminal maturation of pro-Kex2 protease. Embo J. 13 (10): 2280-2288.

GODFREY, T., WEST, S., (1996). Industrial enzymology, 2nd ed., p. 3. Macmillan Publishers inc., New York, N.Y.

GONZALEZ, J.M., SAIJ-JIMENEZ, C., (2004). Microbial activity in biodeteriorated monuments as studied by denaturing gradient gel electrophoresis. J. Separ. Sci. 27: 1024-1028.

GONZALEZ, J.M., SAIJ-JIMENEZ, C., (2005). Application of molecular nucleic acid-based techniques for the study of microbial communities in monuments and artworks. Intern. Microbiol. 8: 189-194.

GOODLOVE, P.E., CUNNINGHAM, P.R., PARKER, J., CLARK, D.P., (1989). Cloning and sequence analysis of the fermentative alcohol-dehydrogenase-encoding gene of *Escherichia coli*. *Gene* 85 (1): 209-214.

GORDON, V.M., LEPPLA, S.H., (1994). Proteolytic Activation of Bacterial Toxins: Role of Bacterial and Host Cell Proteases. *Infect. And Immunity* 62 (2): 333-340

GOTTWALD, M.A., GOTTSCHALK, G., (1984). The internal pH of *Clostridium acetobutylicum* and its effect on the shift from acid to solvent formation. *Arch. Microbiol.* 143: 42-46.

GREEN, E.M., BENNETT, G.N., (1996). Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. *Appl. Biochem. Biotechnol.* 57/58, 213.

GUPTA, R., BEG, Q.K., LORENZ, P., (2002a). Bacterial alkaline proteases: molecular approaches and industrial applications. *Appl. Microbiol. Biotechnol.* 59: 15-32.

GUPTA, R., BEG, Q.K., KHAN, S., CHAUHAN, B., (2002b). An overview on fermentation, downstream processing and properties of microbial alkaline proteases. *Appl. Microbiol. Biotechnol.* 60: 381-395.

HAMAMURA, N., PAGE, C., LONG, T., SEMPRINI, L., ARP, D.J., (1997). Chloroform cometabolism by butane-grown CF8, *Pseudomonas butanovora*, and *Mycobacterium vaccae* JOB 5 and methane-grown *Metylosinus trichosporium* OB3b. *Appl. Environ. Microbiol.* 63: 3607-3613.

HANDELSMAN, J., RONDON, M.R., BRADY, S.F., CLARDY, J., 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.

HANDELSMAN, J., (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* 68: 669-685.

HANAHAN, D., (1983). Studies on transformation of *Escherichia coli* with plasmid. *J. Mol. Biol.* 166: 557-580.

HEALY, F.G., RAY, R.M., ALDICH, A.C., WILKIE, L.O., INGRAM, L.O., SHANMUGAM, K.T., (1995). Direct isolation of functional genes encoding cellulases from the microbial consortia in thermophilic, anaerobic digester maintained on lignocellulose. *Appl. Microbiol. Biotechnol.* 43: 667-674.

HENNE, A., DANIEL, R., SCHMITZ, R.A., GOTTSCHALK, G., (1999). Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Appl. Environ. Microbiol.* 65: 3901-3907.

HENNE, A., SCHMITZ, R.A., BÖMEKE, M., GOTTSCHALK, G, DANIEL, R. (2000). Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. Appl. Environm. Microbiol. 66: 3113-3116.

HOSTER, F., SCHMITZ, J.E., DANIEL, R., (2005). Enrichment of chitinolytic microorganisms: isolation and characterization of a chitinase exhibiting antifungal activity against phytopathogenic fungi from a novel *Streptomyces* strain. Appl. Microbiol. Biotechnol. 66 (4): 434-442.

HOLBEN, W.E., (1996). Isolation and purification of bacterial community DNA from environmental samples. In Manual of environmental microbiology p. 431-436. Hurst, C.J., Knudsen, G.R., McInerney, M.J., Stetzenbach, L.D., Walter, M.V., (ed.), ASM Press, Washington, USA.

HOLMES, D.S, QUIGLEY, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114: 193-197.

HUGENHOLTZ, P., GOEBBEL, B.M., PACE, N.R., (1998). Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. 180: 4765-4774.

INOUE, H., NOJIMA, H., OKAYAMA, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. Gene 96: 23-28.

JACOBSEN, C.S., RASMUSSEN, O.F., (1992). Development and application of a new method to extract bacterial DNA from soil based on separation of bacteria from soil with cation-exchange resin. Appl. Environ. Microbiol. 58:2458-2462

JONES, D.T., WOODS, D.R., (1986). Acetone-butanol fermentation revisited. Microbiol. Rev. 50: 484-524.

JONGENEEL, C.V., BOUVIER, J., BAIROCH, A., (1998). A unique signature identifies a family of zinc-dependent metallopeptidases. FEBS Lett. 242: 211-214.

KAUFFMANN, I.M., SCHMITT, J., SCHMID, R.D., (2004). DNA isolation from soil samples for cloning in different hosts. Appl. Microbiol. Biotechnol. 64: 665-670.

KEIL-DLOUHA, V., (1976). Chemical characterization and study of the autodigestion of pure collagenase from *Achromobacter iophagus*. Biochim. Biophys. Acta 429 (1): 239-251.

KESSLER, E., SAFRIN M., (1994). The propeptide of *Pseudomonas aeruginosa* elastase acts as an elastase inhibitor. J. Biol. Chem. 269: 22726-22731.

KESSLER, E., SAFRIN, M., GUSTIN, J.K., OHMAN, D.E., (1998). Elastase and the LasB protease of *Pseudomonas aeruginosa* are secreted with their propeptides. J. Biol. Chem. 273: 30225-30231.

KIM, D.W., MATSUZAWA, H., (2000). Requirement for the COOH-terminal pro-sequence in the translocation of aqualisin I across the cytoplasmic membrane in *Escherichia coli*. Biochem. Biophys. Res. Commun. 277: 216-220.

KNIETSCH, A., (2002). Konstruktion und Durchmusterung von Metagenombanken: Identifizierung und Charakterisierung von Genen und Genprodukten von am Polyol-stoffwechsel beteiligten Oxidoreduktasen und Coenzym B₁₂-abhängigen Dehydratasen. Dissertation der Georg-August-Universität zu Göttingen.

KNIETSCH, A., WASCHKOWITZ, T., BOWIEN, S., HENNE, A., DANIEL, R. (2003b). Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli*. J. Mol. Microbiol. Biotechnol. 5 (1): 46-56.

KNIETSCH, A., WASCHKOWITZ, T., BOWIEN, S., HENNE, A., DANIEL, R. (2003b). 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., BOWIEN, S., WHITED, G., GOTTSCHALK, G., DANIEL, R. (2003c). 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.

KOOI, C., CORBETT, C.R., SOKOL, P.A., (2005). Functional Analysis of the *Burkholderia cenocepacia* ZmpA Metalloprotease. J. Bacteriol. 187 (13): 4421-4429.

KUBO, M., IMANAKA, T., (1988). Cloning and nucleotide sequence of the highly thermostable neutral protease gene from *Bacillus stearothermophilus*. J. Gen. Microbiol. 134: 1883-1892.

LACKS, S.A., SPRINGHORN, S.S., (1980). Renaturation of enzymes after polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. J. Biol. Chem. 255 (15): 7467-7473.

LAEMMLI, U.K., (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

LÄMMLE, K., (2004). Neue Enzyme für industrielle anwendungen aus Boden-Genbanken. Dissertation der Universität Stuttgart.

LEE, S.F., FORSBERG, C.W., GIBBINS, L.N., (1985). Xylanolytic activity of *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 50: 1068-1076.

LEE, S.-O., KATO, J., TAKIGUCHI, N., KURODA, A., IKEDA, T., MITSUTANI, A., OHTAKE, H., (2000). Involvement of an extracellular protease in algicidal activity of the marine bacterium *Pseudoalteromonas* sp. Strain A28. Appl. Environ. Microbiol. 66: 4334-4339.

LEE, S.-O., KATO, J., NAKASHIMA, K., KURODA, A., IKEDA, T., TAKIGUCHI, N., OHTAKE, H., (2002). Cloning and Characterization of Extracellular Metal Protease Gene of the Algicidal Marine Bacterium *Pseudoalteromonas* sp. Strain A28. Biosci. Biotechnol. Biochem. 66 (6): 1366-1369.

LEE, S.-W., WON, K., LIM, H.K., KIM, J.C., CHOI, G.J., CHO, K.Y., (2004). Screening for novel lipolytic enzymes from uncultured soil microorganisms. Appl Microbiol Biotechnol. 65 (6): 720-726.

LEE, B.-M., PARK, Y.-J., PARK, D.-S., KANG, H.-W., KIM, J.-G., SONG, E.-S., PARK, I.-C., YOON, U.-H., HAHN, J.-H., KOO, B.-S., LEE, G.-B., KIM, H., PARK, H.-S., YOON, K.-O., KIM, J.-H., JUNG, C., KOH, N.-H., SEO, J.-S., GO, S.-J., (2005). The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res. 33 (2): 577-586.

LEYNS, F., DE CLEENE, M., SWINGS, J., DE LEY, J., (1984). The host range of the genus *Xanthomonas*. Bot. Rev. 50: 305-355.

LORENZ, P., ECK, J., (2005). Metagenomics and industrial applications. Nature Rev. Microbiol. 3 (6): 510-516.

LLOYD-JONES, G., HUNTER, D.W.F., (2001). Comparison of rapid DNA extraction methods applied to contrasting New Zealand soils. Soil Biol. Biochem. 33: 2053-2059.

LORENZ, P., ZINKE, H., (2005). White biotechnology: differences in US and EU approaches? Trends Biotechnol. 23 (12): 570-574.

LOROWITZ, W., CLARK, D., (1982). *Escherichia coli* Mutants with a Temperature-Sensitive Alcohol Dehydrogenase. J. Bacteriol. 152 (2): 935-938

MARIE-CLAIRE, C., RUFFET, E., BEAUMONT, A., ROQUES, B.P., (1999). The prosequence of thermolysin acts as an intramolecular chaperone when expressed in trans with the mature sequence in *Escherichia coli*. J. Mol. Biol. 285: 1911-1915.

MILLER, J.H. (1972). Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MITCHELL, W.J., (1998). Physiology of carbohydrate to solven conversion by clostridia. *Adv. Microbiol. Physiol.* 39: 31-130.

MIYAMOTO, K., NUKUI, E., HIROSE, M., NAGAI, F., SATO, T., INAMORI, Y., TSUJIBO, H., (2002). A Metalloprotease (MprIII) Involved in the Chitinolytic System of a Marine Bacterium *Alteromonas* sp. Strain 0-7. *Appl. Environ. Microbiol.* 68 (11). 5563-5570.

MIYOSHI, S.-I., SHINODA, S., (2000). Micribial metalloproteases and pathogenesis. *Microbes and Infec.* 2: 91-98.

MORE, M.I., HERRICK, J.B., SILVA, M.C., GHIOSE, W.C., MADSEN, E.L., (1994). Quantitative cell lysis of indigenous microorganisms and rapid extraction of DNA from sediment. *Appl. Environ. Microbiol.* 60: 1572-1580.

MUYZER, G., DE WAAL, E.C., 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.

NAIR, R.V., BENNETT, G.N., PAPOUTSAKIS, E.T., (1994). Molecular characterization of an aldehyde/alcohol dehydrogenase gene from *clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* 871-885.

NAIR, R.V., PAPOUTSAKIS, E.T., (1994). Expression of plasmid-encoded *aad* in *Clostridium acetobutylicum* M5 restores vigorous butanol production. *J. Bacteriol.* 176: 5843-5846.

NEGRETE-ABASCAL, E., VACA PACHECO, S., PANIAGUA, G.L., PEREZ MENDEZ, A., IBARRA CABALLERO, J., PEREZ MARQUEZ, V.M., TENORIO, V.R., (2004). Metalloproteases secreted by *Actinobacillus sius*. *Curr. Microbiol.* 49: 55-58

NIRASAWA, S., NAKAJIMA, Y., ZHANG, Z.Z., YOSHIDA, M., HAYASHI, K., (1999). Intramolecular chaperone and inhibitor activities of a propeptide from a bacterial zinc aminopeptidase. *Biochem. J.* 341: 25-31.

NÖLLING, J., BRETON, G., OMELCHENKO, K.S., MAKAROVA, K.S., ZENG, Q., GIBSON, R., LEE, H.M., DUBOIS, J., QIU, D., HITTI, J., GTC SEQUENCING CENTER PRODUCTION, FINISHING AND BIOINFORMATICS TEAMS, WOLF, Y.I., TATUSOV, R.L., SABATHE, F., DOUCETTE_STAMM, L., SOUCAILLE, P., DALY, M.J., BENNETT, G.N., KOONIN, E.V., SMITH, D.R., (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* 183: 4823-4838.

O'DONOHUE, M.J., BEAUMONT, A., (1996). The roles of the prosequence of thermolysin in enzyme inhibition and folding in vitro. *J. Biol. Chem.* 271: 26477-26481.

OGRAM, A., SAYLER, G.S., BARKAY, T., (1987). The extraction and purification of microbial DNA from sediments. *J. Microbiol. Methods* 7: 57-66.

PACE, N.R., STAHL, D.A., LANE, D.J., OLSEN, G.J., (1985). Analysing natural microbial populations by rRNA sequencing. *ASM News* 51, 4-12.

PALOSAARI, N., ROGERS, P., (1988). Purification and properties of the inducible coenzyme A-linked butyraldehyde dehydrogenase from *Clostridium acetobutylicum*. *J. Bacteriol.* 170 (7): 2971-2976.

PETERSEN, D.J., WELCH, R.W., RUDOLPH, F.B., BENNETT, G.N., (1991). Molecular cloning of an alcohol (butanol) dehydrogenase gene cluster from *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol.* 173: 1831-1834.

PFENNIG, N., LIPPERT, K.D. (1966). Über das Vitamin B12-Bedürfnis phototropher Schwefelbakterien. *Arch. Microbiol.* 55: 245-256.

PICARD, C., PONSONNET, C., PAGET, E., NESME, X., SIMONET, P., (1992). Detection and enumeration of bacteria in soils by direct DNA extraction and polymerase chain reaction. *Appl. Environ. Microbiol.* 58: 2717-2722.

PURDY, K.J. (2005). Nucleic acid recovery from complex environmental samples. *Meth. Enzymol.* 397: 271-292.

PUROHIT, H.J., KAPLEY, A., MOHARIKAR, A.A., NARDE, G., (2003). A novel approach for extraction of PCR-compatible DNA from activated sludge samples collected from different biological effluent treatment plants. *J. Microbiol. Methods* 52: 315-323.

RAJAMANI, S., HILDA, A., (1987). Plate assay to screen fungi for proteolytic activity. *Curr. Sci.* 56: 1179-1181.

RAO, M.B., TANKSALE, A.M., GHATGE, M.S., DESHPANDE, V.V., (1998). Molecular and Biotechnological Aspects of Microbial Proteases. *Microbiol. Mol. Biol. Rev.* 62 (3): 597-635.

RAPPE, M.S., GIOVANNONI, S.J., (2003). The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* 57: 369-394.

RAWLINGS, N.D., BARRETT, A.J., (1993). Evolutionary families of peptidases. *Biochem. J.* 290: 205-218.

RAWLINGS, N.D., BARRETT, A.J., (1995). Evolutionary Families of Metalloproteases. *Methods Enzymol.* 248:183-228.

RAWLINGS, N.D., BARRETT, A.J., (2004). Introduction: Metallopeptidases and their clans. In *Handbook of Proteolytic Enzymes*, 2 edn (Barrett,A.J., Rawlings,N.D. & Woessner,J.F. eds) p.231-267, Elsevier, London (2004)

RAWLINGS, N.D., MORTON, F.R., BARRETT, A.J., (2006). MEROPS: the peptidase database. *Nucleic Acids Res.* 34, D270-D272.

REES, H., GRANT, S., JONES, B., GRANT, W., HEAPHY, S., (2003). Detecting cellulase and esterase enzymes activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* 7: 415-421.

RIESENFELD, C.S., SCHLOSS, P.D., HANDELSMA, J., (2004). METAGENOMICS: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet.* 38: 525-552.

RILEY, M., TAKASHI, A., ARNAUD, M.B., BERLYN, M. K.B., BLATTNER, F.R., CHAUDHURI, R.R., GLASNER, J.D., HORIUCHI, T., KESELER, I.M., KOSUGE, T., MORI, H., PERNA, N.T., PLUNKETT, G., RUDD, K.E., SERRES, M.H., THOMAS, G.H., THOMSON, N.R., WISHART, D., WANNER, B.L., (2006). *Escherichia coli* K-12: a cooperatively developed annotation snapshot-2005. *Nucleic Acids Res.* 34 (1): 1-9.

RONDON, M.R., GOODMAN, R.M., HANDELSMAN, J., (1999). The earth's bounty: assessing and accessing soil microbial diversity. *TIBTECH* 17: 403-409.

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., TIONG, C.L., GILMAN, M., OSBURNE, M.S., CLARDY, J. HANDELSMAN, J. GOODMAN, R.M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66: 2541-2547.

RUTHERFORD, K., PARKHILL, J., CROOK, J., HORSNELL, T., RICE, P., RAJANDREAM, M.A., BARRELL, B., (2000). Artemis: Sequence visualization and annotation. *Bioinformatics* 16 (10): 944-945

RAPPE, M.S., GIOVANNONI, S.J., (2003). The uncultured microbial majority. *Annu. Rev. Microbiol.* 57: 369-94.

RUTTLOFF, H. (ed.) (1994). Industrielle Enzyme (2. Auflage). Behr's Verlag, Hamburg.

SAMBROOK, J., FRITSCH, E., MANIATIS, T., (1989). Molecular cloning: a laboratory manual (2.Aufl.). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA

SANCHEZ, L.B., (1998). Aldehyde dehydrogenase (CoA-acetylating) and the mechanism of ethanol formation in the amitochondriate protist, *Giardia lamblia*. *Arch. Biochem. Biophys.* 354: 57-64.

SANCHEZ-PORRO, C., MELLADO, E., BERTOLDO, C., ANTRANIKIAN, G., VENTOSA, A., (2003). Screening and characterization of the protease CP1 produced by the moderately halophilic bacterium *Pseudoalteromonas* sp. strain CP76. *Extremophiles* 7: 221-228.

SANTOSA, D.A., (2001). Rapid extraction and purification of environmental DNA for molecular cloning applications and molecular diversity studies. *Mol. Biotechnol.* 17: 59-64.

SAUER, U., DÜRRE, P., (1995). Differential induction of genes related to solvent formation during shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. *FEMS Microbiol. Lett.* 125: 115-120.

SCHLOSS, P.D., HANDELSMAN, J., (2004). Status of the microbial Census. *Mol. Biol. Rev.* 68: 686-691

SCHMEISSER, C., STÖCKIGT, C., RAASCH, C., WINGENDER, J., TIMMIS, K.N., WENDEROTH, D.F., FLEMMING, H.C., LIESEGANG, H., SCHMITZ R.A., JAEGER, K.E., STREIT, W.R., (2003). Metagenome survey of biofilms in drinking-water networks. *Appl. Environ. Microbiol.* 69 (12): 7298-309.

SCHMITZ, J.E., (2005). Isolierung von DNA und Konstruktion einer Metagenombank aus dem sediment des Flusses Leine: partielle sequenzierung und Annotation des Metagenoms sowie Analyse der mikrobiellen Diversität. Dissertation Georg-August-Universität zu Göttingen.

SCHWARZ, S., WASCHKOWITZ, T., DANIEL, R., (2006). Enhancement of gene detection frequencies by combining DNA-based stable-isotope probing with the construction of metagenomic DNA libraries. *World J. Microbiol. Biotechnol.* Im Druck. Online erhältlich unter DOI 10.1007/s11274-005-9042-z

SELENSKA, S., KLINGMÜLLER, W., (1991). DNA recovery and direct detection of *Tn5* sequences from soil. *Lett. Appl. Microbiol.* 13: 21-24.

SHEN, G.J., SHIE, J.S., GRETHLEIN, A.J., JAIN, M.K., (1999). Biochemical basis for carbon monoxide tolerance and butanol production by *Butyribacterium methylotrophicum*. *Appl. Microbiol. Biotechnol.* 51: 827-832.

SHINDE, U., INOUE, M., (2000). Intramolecular chaperones: polypeptide extensions that modulate protein folding. *Semin. Cell Dev. Biol.* 11 (1): 35-44.

SHUMAN, S., (1991). Recombination Mediated by Vaccinia Virus DNA Topoisomerase I in *Escherichia coli* is Sequence Specific. *Proc. Natl. Acad. Sci. USA* 88: 10104-10108.

SHUMAN, S., (1994). Novel Approach to Molecular Cloning and Polynucleotide Synthesis Using Vaccinia DNA Topoisomerase. *J. biol. Chem.* 269: 32678-32684.

SEOW, K.T., MEURER, G., GERLITZ, M., WENDT-PIENKOWSKI, E., HUTCHINSON, R., DAVIES, J., (1997). A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* 179: 7360-7368.

SMALLA, K., CRESSWELL, N., MENDONCA-HAGLER, L.C., WOLTER, A., VAN ELSAS, J.D., (1993). Rapid DNA extraction protocol from soil for polymerase chain reaction-mediated amplification. *J. Appl. Bacteriol.* 74: 78-85.

SOMMERVILLE, C.C., KNIGHT, I.T., STREUBE, W.L., COLWELL, R.R., (1989). Simple, rapid method for direct isolation of nucleic acids from aquatic environments. *Appl. Environ. Microbiol.* 55: 548-554.

STADEN, R., BEAL, K.F., BONFIELD, J.K., (2000). The Staden package, 1998. *Methods Mol. Biol.* 132: 115-130.

STEIN, J.L., MARSH, T.L., WU, K.Y., SHIZUYA, H., DELONG, E.F., (1996). Characterization of uncultivated prokaryotes: isolation and analysis of a 4-kilobase-pair genome fragment from a planktonic marine archaeon. *J. Bacteriol.* 178: 591-599.

STREIT, W.R., DANIEL, R., 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., SCHMITZ, R.A., (2004). Metagenomics- the key to the uncultured microbes. *Curr. Opin. Microbiol.* 7: 492-498.

STRICKEBERGER, M.W., (1996). In: *Evolution*. Johnes & Bartlett Publishers, Boston, USA.

TAKAHASHI, J., (1980). Production of intracellular and extracellular protein from n-butane by *Pseudomonas butanovora* sp. nov. *Appl. Microbiol.* 26: 117-127.

TAKAHASHI, J., ICHIKAWA, Y., SAGAE, H., KOMURA, I., KANOU, H., YAMADA, K., (1980). Isolation and identification of n-butane- assimilating bacterium. *Agric. Biol. Chem.* 44: 1835-1840.

TANG, B., NIRASAWA, S., KITAOKA, M., HAYASHI, K., (2002). In vitro stepwise autoprocessing of the proform of pro-aminopeptidase processing protease from *Aeromonas caviae* T-64. *Bioch. et Biophys. Acta* 1596: 16-27.

TEBBE, C.C., VAHJEN, W., (1993). Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and yeast. *Appl. Environ. Microbiol.* 59: 2657-2665.

TORSVIK, V.L., (1980). Isolation of bacterial DNA from soil. *Soil Biol. Biochem.* 12: 15-21.

TOWBIN, H., STAHELIN, T., GORDON, J. (1979). Elektrophoretic transfer of proteins from polyacrylamid gels to nitrocellulose sheets: procedure and some applications. *Proc Nat Acad Sci. USA.* 76: 4350-4354.

TRINGE, S.G., von MERING, C., KOBAYASHI, A., SALAMOV, A.A., CHEN, K., CHANG, H.W., PODAR, M., SHORT, J.M., MATHUR, E.J., DETTER, J.C., BORK, P., HUGENHOLTZ, P., RUBIN, E.M., (2005). Comparative Metagenomics of Microbial Communities. *Science* 308: 554-557.

TSAI, Y.L., 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.

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., BANFIELD, J.F., (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428: 37-43.

VANGNAI, A.S., ARP, D.J., SAYAVEDRA-SOTO, L.A., (2002a). Two distinct alcohol dehydrogenases participate in butane metabolism by *Pseudomonas butanovora*. *J. Bacteriol.* 184: 1916-1924.

VANGNAI, A.S., SAYAVEDRA-SOTO, L.A., ARP, D.J., (2002b). Roles for the two 1-butanol dehydrogenases of *Pseudomonas butanovora* in butane and 1-butanol metabolism. *J. Bacteriol.* 184: 4343-4350.

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., FOUTS, D. E., LEVY, S., KNAP, A.H., LOMAS, M W., NEALSON, K., WHITE, O., PETERSON, J., HOFFMANN, J., PARSON, R., BADEN-TILLSON, H., PFANNKOCH, C., ROGERS, Y.H., SMITH, H.O., (2004). Environmental genome shotgun sequencing of the Sargosso Sea. *Science* 304: 66-74.

VIEILLE, C., ZEIKUS, G.J., (2001). Hyperthermophilic enzymes: sources, uses and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65: 1-43.

VIEIRA, J., MESSING, J. (1982). The pUC plasmids, a M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19: 259- 268.

VOGET, S., STEELE, H.L., STREIT, W.R., (2005). Metagenomes– an unlimited resource for novel genes, biocatalysts and metabolites. *Minerva Biotec* 17: 47-53.

WANG, G.Y., GRAZANI, E., WATERS, B., PAN, W., LI, X., MC DERMOTT, J., MEURER, G., SAXENA, G., ANDERSEN, R.J., DAVIES, J., (2000). Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* 2: 2401-2404.

WALTER, K.A., BENNETT, G.N., PAPOUTSAKIS, E.T., (1992). Molecular characterization of two *Clostridium acetobutylicum* ATCC 824 butanol dehydrogenase isoenzyme genes. *J. Bacteriol.* 174: 7149-7158.

- WATSON, R.J., BLACKWELL, B., (2000).** Purification and characterization of a common soil component which inhibits the polymerase chain-reaction. *Can. J. Microbiol.* 46: 633-642.
- WELCH, R.W., RUDOLPH, F.B., PAPOUTSAKIS, E.T., (1989).** Purification and characterization of the NADH-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (ATCC 824). *Arch. Biochem. Biophys.* 273: 309-318.
- WELLINGTON, E.M.H., BERRY, A., KRSEK, M., (2003).** Resolving functional diversity in relation to microbial community structure in soil. *Curr. Opin. Microbiol.* 6: 295-301.
- WIKSTROM, P., WIKLUND, A., ANDERSSON, A.C., FORSMAN, M., (1996).** DNA recovery and PCR quantification of catechol 2,3-dioxygenase genes from different soil types. *J. Biotechnol.* 52: 107-120.
- WILKINSON, D.E., JEANICKE, T., COWAN, D.A., (2002).** Efficient molecular cloning of environmental DNA from geothermal sediments. *Biotechnol. Lett.* 24: 155-161.
- WILLKE, T., VORLOP, K.D., (2004).** Industrial bioconversion of renewable resources as an alternative to conventional chemistry. *Appl. Microbiol. Biotechnol.* 66: 131-142.
- WILSON, E.O., (1992).** In: *The diversity of life*, pp. 131-162: W.W. Norton & Co. New York, USA.
- WINDLE, H.J.P., KELLEHER, D., (1997).** Identification and Characterization of a Metalloprotease Activity from *Helicobacter pylori*. *Infect. Immun.* 65 (8): 3132-3137.
- WOESE, C.R., (1987).** Bacterial evolution. *Microbiol. Rev.* 51: 221-271.
- WOESE, C.R., KANDLER, O., WHEELIS, M.L., (1990).** Towards a natural system of organisms: proposal for the domains Archaea, Bacteria ad Eucarya. *Proc. Natl. Acad. Sci. USA* 87: 4576-4579.
- WOODS, D.R., REID, S.J., (1995).** Regulation of nitrogen metabolism, starch utilization and the β -hbd-adh1 gene cluster in *Clostridium acetobutylicum*. *FEMS Microbiol. Rev.* 17: 299-306.
- YAN, R.-T., CHEN, J.-S., (1990).** Coenzyme A-acylating aldehyde dehydrogenase from *Clostridium beijerinckii* NRRL B592. *Appl. Environ. Microbiol.* 56 (9): 2591-2596.
- YANG, W., LI, E., KAIRONG, T., STANLEY JR., S.L., (1994).** *Entamoeba histolytica* has an alcohol dehydrogenase homologues to the multifunctional adhE gene product of *Escherichia coli*. *Mol. Biochem. Parasitol.* 64: 253-260.

YEATS, C., RAWLINGS, N.D., BATEMAN, A., (2004). The PepSY domain: a regulator of peptidase activity in the microbial environment? Trends Biochem. Sci. 29: 169-172.

YOUNGLESON, J.S., JONES, W.A., JONES, D.T., WOODS, D.R., (1989). Molecular analysis and nucleotide sequence of the *adh1* gene encoding an NADH-dependent butanol dehydrogenase in the Gram-positive anaerobe *Clostridium acetobutylicum*. Gene 78: 355-364.

YUN, J., KANG, S., PARK, S., YOON, H., KIM, M.J., HEU, S., RYU, S., (2004). Characterization of a Novel Amylolytic Enzyme Encoded by a Gene from a Soil-Derived Metagenomics Library. Appl. Environ. Microbiol. 70: 7229-7235.

ZHOU, J. M., BRUNS, M.A., TIEDJE, J.M., (1996). DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. 62: 316-322.

ZILLIG, W., HOLZ, I., JANEKOVIC, D., KLENK, H.P., IMSEL, E., TRENT, J., WUNDERL, S., FORJAZ, V.H., COUNTINHO, R., FERREIRA, T., (1990). *Hyperthermus butylicus*, a hyperthermophilic sulfur-reducing archaebacterium that ferments peptides. J. Bacteriol. 172: 3959-3965.

7 Anhang

7.1 Angaben zur Insertsequenz von pTWP9

7.1.1 Lage der offenen Leserahmen

ORF	Gen-name	Annotation	Sequenz-Lokalisation	Orien-tierung
1	<i>rplA</i>	50S ribosomales Protein L1	70-762	-
2	<i>rplK</i>	50S ribosomales Protein L11	767-1195	-
3	<i>nusG</i>	Transkriptions-Antiterminationsfaktor	1360-1962	-
4	<i>secE</i>	Präprotein Translokase UE	1994-2377	-
5	<i>tufA</i>	Elongationsfaktor Tu	2722-3912	-
6	<i>ychF</i>	GTP-Bindeprotein	4459-5550	-
7	<i>Pth</i>	Peptidyl-tRNA Hydrolase	5599-6180	-
8	<i>rplY</i>	50S ribosomales Protein L25	6232-6912	-
9	<i>prsA</i>	Ribose-Phosphat-Pyrophosphokinase	7166-8125	-
10	<i>ipK</i>	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase	8331-9293	-
11	<i>lolB</i>	Lipoprotein Precursor der äußeren Membran	9236-9874	-
12	<i>orf12</i>	Konserviertes hypothetisches Protein	9871-11544	-
13	<i>hemA</i>	Glutamyl-tRNA-Reduktase	11624-12934	+
14	<i>prtA</i>	Peptidketten-Release-Faktor	13123-14214	+
15	<i>orf15</i>	GCN5 ähnliche N-Acetyltransferase	14354-14800	+
16	<i>moaB</i>	Molypdopterin Biosynthese Protein	14874-15392	+
17	<i>orf17</i>	Konserviertes hypothetisches Protein	15493-16269	+
18	<i>acyP</i>	Acyolphosphatase	16302-16574	-
19	<i>yneN</i>	Mögliches Thioredoxin	16588-17172	-
20	<i>rbn</i>	Ribonuklease BN	17169-18527	-
21	<i>trp</i>	Tryptophan Repressor Bindeprotein	18585-19178	+
22	<i>orf22</i>	Konserviertes hypothetisches Protein	19175-19516	+
23	<i>orf23</i>	Hypothetisches Protein	19722-20255	-
24	<i>orf24</i>	Hypothetisches Protein	20439-21092	-
25	<i>orf25</i>	Hypothetisches Protein	21213-21746	-
26	<i>aspG</i>	Asparaginase	21894-22382	+
27	<i>orf27</i>	Konserviertes hypothetisches Protein	22384-22725	+
28	<i>orf28</i>	Hypothetisches Protein	23039-23188	+
29	<i>mprA</i>	Metalloprotease	23201-25789	-
30	<i>orf23</i>	Konserviertes hypothetisches Protein	26218-26769	-
31	<i>orf23</i>	Konserviertes hypothetisches Protein	26871-27299	+
32	<i>ptnB*</i>	Mögliche NAD(P)-Transhydrogenase β UE	27333-28112	-

7.1.2 Insertsequenz von pTWP9

```

1 gatccttcgc tgcgtcagg atgacgcaat ggcaaaggcg tttgcgtac gcgcaaacgc
61 gccggtgcgt cacttcagcg tcagcgctga ctggctgacg atcacgcccc ggcccatggg
121 tgagctcagc gagatcttct gcaggtactg gcccttgagg ctgcagcggt tggccttgat
181 cagggtccagc agcagcgccg gcaggttgct cttcagcttg ccggcctcga agtcggcctt
241 gccgatggcg cagtggatga tgccggcctt gtcggtgcgg tagcgacact ggccgcccct
301 ggcggttctc accgcctcgg ccgggttcgg cgacacggcg ccgaccttcg gggtgggcat
361 caggccgcgc ggaccgagca gctggccgag cttgccgacc acgcgcacgc cgtccggcgt
421 ggcgatgacc acgtcgtagt tgaggtcgcc ggccctgcac ttctcggcca ggtcgtccat
481 gccacaggct tcggcaccgg cggccaggcg ctcgtcgcc ttggcgcccc cggcgccgaa
541 caccgccacg cgacgcgtct tgccggtgccc ggccggcagc accggtcgaac cgcgacactg
601 ctggctcgac ttcttgcat ccacgccaag gcgcacggcg acgtcgacgg actcgacgaa
661 cttcaccttg ctgttgccct ggacgatctt caacgcctca tcgatggcgt aggccttgcc
721 cggtcgcac ctggcctggg ccttgattcg cttggtcgct atgtcctcag cctccaccg
781 tcaaacccat ggagcgggcc gagcccgcca tcgtcttcac cgccgcgtcc aggtcggccg
841 cggtcaggct ggcttccttc gccctggcga tgcctcgag ctgcttgagg gtgacctgac
901 ccaccttctc ggtgttcggg cgcttgagc cggactggac gcccgccgcc ttcttcaaca
961 gcaccgtggc cggcgtgctc ttggtgatga aggtgaaggc accggtccgaa tacgcgggtg
1021 tgatcaccgg agtcggcagg ccgggtccca gcttctgcgt ggccgacatt aacgccttgc
1081 agaattccat gatgttgagg ccgcgtggc ccaacgcagg acccaccggc ggcgacggat
1141 tcgcctgccc ggcttgacc tgacgctga tgtaaccgac gactttcttt gccattgctc
1201 actgctcctg cgagtgcgaa cgcttgagg ctcctcgctt gcttgccgct gcgaaccact
1261 cgacggcgcg aattttgttg ctcccgagg agcggttca gccgcgagct ctccggccct
1321 tggcaagct cgcggtgaa gccgctccta caacaacggc caggccttct cgacctggcc
1381 gaactccagt tcaaccggcg tcgaacgccc gaagatcagg accgccacc gcagcgggct
1441 cttgtcgtag ttgacttcct cgaccacgccc gttgaaatcg ttgaacgggc cgtcgaccac
1501 gcgcaccatc tggccggcct cgaacagcac cttcgagcgc ggcttctcga cgccctcctg
1561 gaccgggtcc aggatcgagg ccgcctcgct gtcgcgaatc ggcaagggct tgcggcggtg
1621 gccgccgatg aagcccatga cttcggtgtt ttcttgatc aggtgccagc tctcgtgtgc
1681 gatgcggggg atgccgcctt cgtcgtgggt gacgatctgc accagcacgt agccggggaa
1741 gaacttgctc tcggaacggc gcttctgccc tgcgcgcac tcgatcactt cctcggctcg
1801 caccatcacg tcgccgaaac gatcttccat gccggtgcgc acgatgcggc cgcgacgggc
1861 ctgggccacc gacttctcga aaccggaata ggcatgcacg acataccagc gcttggtcgc
1921 gtccggcggt gtggtgcccg tattgtgtgc ttgcgtttcc atcaacctgc tgctccagct
1981 gtcagtgcac ccgtcaacgt gccaggaaac acttcaccgc ggtctgcacg atcacgtcga
2041 agaaggccag caccaggctg atcaggatca ccaccgcgat caccaccag gtcgaacgcg
2101 tcgcttctct gcgcgtcggc cacaccactt tgcgcagctc gaaacgagct tccgagagga
2161 actcgggggt ctgcacgccc ttgtggctgg tcaggaacac gaccgcggcg ccgaccacgc
2221 cggcgatcac cgacatcacc cgggcgacgc tcggccactg cccgtcgaac cagtagaaga
2281 cgaacaggcc cgccgccacc agcgcgatcg ccagcgcgta cttcatcagg tcgcccggag
2341 tcgcccgggc tttggattgt tcgaccttgc tgttcacgtg atgcgaatcc ctcgcaggag
2401 cccgcgcac cagcgcgga cttcttcaat gctttgtccg gttggacgac gatccggccg
2461 tccggcttgg tttgtggcac gccaggaggg actcgaaccc ccaacctgcg gttttggaga
2521 ccgctgctct gccaatgag ctactggcgt atctaaatct tctgtcttcc cgaaaatggc
2581 aacggcggtc cgcccgcccc tgccgtcctt ctgccttcaa ttcctcttca tgggtcaaga
2641 aaacttgaag ctggttcccc gcctgaacgg ggccggcgcc cccggcaagg gcgtccgcgc
2701 gagcgcggac gcgtcgtcgc gtcacttgat gatcttgccc accacgcccg cgccgacggg
2761 ccggccgccc tcgcggtatc cgaaccgcag gccctcgctc atcgccaccg gggtgatcag
2821 cgtcaccttc atcttcacgt tgcctccggc catcaccatc tccacgccct ccggcagctc
2881 gcaggcgccg gtgatgtccg tgggtgcgaa gtagaactgc ggacggtagc ccttgaagaa
2941 cggcgtgtgg cgcccgccct cgtccttgct caggacatag acctcggcct cgaactcggg
3001 gtgcgggggt atcgaaccgg gcttgggcag cacctggcca cgctccacgt cgtcacgctt

```

3061 ggtgccgcgc agcagcaggc cggcggtgtc gcccgcttgg ccctgggtcca gcagcttgccg
 3121 gaacatctcc acgccggtga ccgtggtctt gaccgtcggg cggatgccca cgatctcgat
 3181 ctcgtcgccc accttgatga tcccgcgctc gatgcgaccg gtcaccacgg tgccgcgccc
 3241 ggaaatcgag aacacgtctt ccaccggcat caggaacgcc cggtcgatgt cgcgctcggg
 3301 ctgcgggatc caggtgtcca gcgcttccac cagcgcgatg atcgacggca cgccgatctc
 3361 ggactgggtc ccttcacgcg ccttgagcgc cgagcccttg atgatcgggg tgtcgtcgcc
 3421 cgggaactcg tacttgaca gcagctcgcg cacctccatc tccaccagct ccatcagctc
 3481 ggcgtcgtcc accatgtccg ccttggtcag gtagaccacg atgtacggca cgcccacctg
 3541 gcgcgccagc aggatgtgct cgcgggtctg cggcatcggg ccgtcggccg ccgaacacac
 3601 caggatcgcg ccgtccatct gcgccgctcc ggtgatcatg ttcttcacgt agtcccgctg
 3661 ccccgggcag tccacgtgcg cgtagtgcg gttcgggctc tcgtattcca cgtgcgcggg
 3721 cgagatcgtg atcccgcgcg ccttctcctc cggcgccgcg tcgatcgcgt cgtaggcctt
 3781 gaactcgcgg ccaaacgct ccgcgccac cttcgtcagc gcccgctca gcgtcgtctt
 3841 gccgtgggtc acgtgaccga tcgtgcccac gttcacgtgg ggcttggtgc gctcgaattt
 3901 accctttgcc atggctcgt ctcgggtgtt tgggtggaact ggcggtttgg actgcttgtt
 3961 tgtcttcgtt gcgaaccagc acatgcaagt actggctttt gcggatggat ccgcttgttg
 4021 gtgctcacga atggaatcga accatcgacc tcctccttac caaggagggt ctctaccgac
 4081 tgagctacgt gagcgctggg ttgtgggcct ggtgcagggt gttcaatgga gcgggagacg
 4141 ggaatcgaa cgcaccatc agcttggaag gctgaggttc taccgttgaa ctactccgc
 4201 gccggaaccg ctgaagcttt tgccgcgtgg gtacaactgg tggagggagg tggattcgaa
 4261 ccaccgaagg cgtaagccag cagatttaca gtctgcccc gttggccgct tgggtatccc
 4321 tccggctcgtt caaatcagc caatcgtgaa acagcccgcg attttgacgt gatggccggg
 4381 gagtgtcaac ggccaggctt gccgcggacg gttgcgcgca gcgcaatcaa gccggtccgg
 4441 ccgcgcggc ggcgcggttc agacattgaa gcggaagtgc agcacgtcgc cttcctggac
 4501 ccggtattcc ttgccctcca gccgcagccg gccggcatcg cgggcgcgg actcgcgccg
 4561 gtacttgatg aagtcacaa aaccgatggt ctgcggcggg atgaagccct tctcgaagtc
 4621 ggtatggatc acggcggcgg cctgcggcgc ggtcgaaccg ccttcacgg tccaggcgcg
 4681 gacttccttg accccggcgg tgaagtaggt ctgcaggccc agcagcttgt aggcggcgcg
 4741 gatgacccgg ttcaggcccg gctcgtccag gccaggtcg gcgaggaagg cgtcgcggtc
 4801 ctcgtcctcg agctggctca gtcctcctc gatggcggcc gagaccggga ccacctcggc
 4861 accctcgcgc agcgcgccg cccgcaccgc gtccagggtg gggttgtcgc tgaagccgtc
 4921 ctcgagcacg ttggcgacgt acatcaccgg cttgagcgtc agcaggaaca ggtcgcggat
 4981 cgcggcgcgg tcctcgtccg acagccccag cgcgcgcgcc ggcttgccg cgtccagccc
 5041 ggcgcgtaac cgcgcagca cctcgaccg cgccctggcg tccttgctgc cggctctggc
 5101 cgtgcgctcg gcgcgctgca gcgcctgtc gaccgattcc aggtccgcca gcgccagctc
 5161 ggtgtcgatc gtctcgatgt cggcgatcgg gtcgatccgg ttggtgacgt ggatcacgtc
 5221 gtcgttctcg aagcagcgca ccacgtgggt gatcgcgtcg acttcgcgga tgtgcgccag
 5281 gaacttggtg cccaggccct cgcgcgtggc ggcgccggcg accaggccgg cgatgtcgac
 5341 gaactcgacc gcggtcggga tcacctctg cggcttgacg atttcggcca gcgcgttcag
 5401 gcgcggatcc ggacccggca ccacgccgac gttgggttcg atcgtgcaga aggggaagtt
 5461 ggccgcggcg atgccggcct tggtcagcgc gttgaacagg gtcgacttgc cgacgttcgg
 5521 caggccgacg atgccgatt ggatgccat cgtgtactac tccgtgcgcg gggcgccggc
 5581 gccggcgcgg tgatgccgct attcgcgcgg cgtgtgcagc cgcttcattg cttcgttgaa
 5641 gtcgcggcg accgccagcg gcagcacgtc gagcgattg ccgatcgcgt gcaggatcgc
 5701 ggctcgtcgc tcgcggccgg gcttgcccag caccacggc gtgaccgggt ccttggtccc
 5761 gggatggccg atgccgatcc gcaggcgggt gaaacgccc tggccgagca gttgcattgt
 5821 gtcgcgcagg ccgttctggc cgcgtggcc gccgtcgaac ttgagccgcg ccgcaccggg
 5881 cggcagggtc agttcgtcgt gcgccagcag cgcttcccc ggctcgatct tccagtagcg
 5941 cagggccgcg gtgaccgact tgcccagacg gttcatgaag gtcgcgggct tcagcagcca
 6001 caccggctgc ccggcgatct cgaccttgca ggtctcgcgg aacagcttcg attccagccc
 6061 gaagcgcgcc ccgcactgct gcgcgagggc gtccacaaac cagaaccggg cgttggtgccg
 6121 ggtcctgaga tgctcgggac cggggttgcc cagcccgacg atgagtcgca gaccgccat
 6181 cgcgaaagcg cgcagcgggt gcgccggacc gggcgactg ctgcgcgacg attactcgcc

6241 cttgccttcg tcggcggcta cgtcggtgcc ttccggcggt accggagcgg ccgcggcggc
 6301 gccctcggcg gcttcctgtg tctcctcgct ttccggcggt gcaaagcggc ccaccaccac
 6361 cgcgacgtcg tggtcggcgc ccaggcgag ctccggcact tccacgccct tcggcaactg
 6421 gacctgcgac aggtgcatgg tgtcgccggc ctgcaggttc gacaggtcga tgtcgatgaa
 6481 ctccggcagg tcgcccgga ggcagctcac ttcatcttgg gtcagctcgt gcatgaccac
 6541 gaccccgcg gtcttgccgg ccggcgactt gtcgacgttg aggaagtga gcggcaccga
 6601 cagcgggatg gcttcgttag cgctcacgag ctggaagtcc atgtgcatga tccgctgctt
 6661 gtacggatgg cgctgcatgt cagcagcag cactttctcg gacttgccgt cgatttcaag
 6721 gtccaggatc gacgcgtaga accactcgct ctggctgacc agccaggtgc gctcgtgctc
 6781 gagctggatg ctgaccggat cggcggtggc gccgtagacg atggccggga cctggccggc
 6841 atgacgcagg cggcggctcg cacccttccc ctcgctcctt cggctggtgg ccgtgatctt
 6901 gtgttctgac attttcgttt actcgctttg cttgctattg gaccgcctcg cggcggcgat
 6961 ggcactcacc cgcgaccagg tgagtgtggg tgccggccgc ggccggcgcc cgtggccggg
 7021 catgagcccg atccacggtt atgacgccc gcacagggat gtgccggcgc ggccggcgcc
 7081 agcccgctgc ggcgacgtcg cccggctttg ccgggagtcg tcgcttgcgg gtcgggcagg
 7141 agcccgatcc cgcaatacaa caccctcagt ccacgtacag cgaactgacg gattccccga
 7201 aggcgatgcy gcggatggtt tccgccagca gttcggcgac gctgagctgg cggatcgtgg
 7261 cgcagcctg cgcgccttcc gacagcggga tgggtgcggt caccaccagc tcgtcgagtt
 7321 gcgacttgcc caggttgtcc agcgcgtcgc cggacagcac cggatgcgtg cagtacgcga
 7381 ccaccttggt cgcgcctgc gccttcaacg ccgcgcgcgc cgcgcacagg gtaccggcgg
 7441 tgtcgacgat gtcgtcgacc agcacgcaga ccttgccggc gacatcgccg atgatgttca
 7501 tcaccgtggc gacgttggcc ttggggcggc gcttgcgat gatggccagg tcggcgtcgt
 7561 ccaggcgctt ggcgatcgcc ctggcccgca ccacgcgcgc gacgtcgggc gataccacga
 7621 tcacgttctg ggtcccgtag tggcgccaga tgtcgccag cagcagcggc gacgagtaca
 7681 cgttgcgac cgggatgtcg aagaaacctt ggcgtcggc ggcatgcagg tccaccgtga
 7741 gcacccggtc ggtgtgcagc gtggagaaca tctgcgcgc gacctggcg gtgatcggca
 7801 cgcgcgcgga cctgggacgg cggctcctgc gggcgtagcc gaagtacggc accaccggg
 7861 tcacgctggc caccgatgcy cgcttgagcy cgtcgatcag caccagcagc tccaccaggt
 7921 tctccgcgct cgggtgcgac gtgggctgca gcacgaacac gtcctggccg cggacgttct
 7981 cctcgatctc gacctgcact tcgccatcgg aaaagcggcc gaccagggcc ttgccgggac
 8041 gcacgcgag ctcttgcac acggcatagg ccacggccgc gttggcggtt ccggtgaaca
 8101 ccagcaggtt gcatcgccct tgcacgtca tgtccgcttc gctctgaccg tggcggtggg
 8161 gggtcgaagc catgcgcggg gcaaggcctt ccgttgtttc aatccccgcg cgcaggcacg
 8221 ggttgcgcy gtgtaacctg tccgcgcgg cccgtaccgc gaggcggggg accttgcgag
 8281 ccgatcgcaa ttggcagggg cggcaggatt cgaacctgcy aatgccagga tcaaaacctg
 8341 gtgccttggg cctcttggcy acgcccctgc agacgttctg catccagcgc caggtgcagc
 8401 ggcgaccgcy tcgcgccttc cgcgatccac gcccgagac cggccggcaa gctcgccagc
 8461 ccggcccgty cgaagcagg atcgtcgaat tcgacgaagc agccactgcc cgaccgggtc
 8521 aatcgcggcc ggcgatccg cgccaacgcy gcgaacgcy cctccacggc cggttcgcg
 8581 cgacgcagca ccggctcgaa ggcattgcc agcggagcac ccgaaacgaa gtcggatatt
 8641 gtcgcgggcy ccgcatccg cgtcaattca ggggcttgg acaattccgc ggtcgaagcy
 8701 tggacgcgcy ggtcggcgag caggtaccag gccggcgga gttccagcgc gaccaggtcc
 8761 tcgcccagcy ctctggccca ggcattctcg ccgcgcagca acaccggcac gtcggcgccc
 8821 agctgcaggc cgagcgcggc gagccggtcg aggtccaggc ccgcgtccca aagctggttg
 8881 agcggccacca gcacggtggc ggcacggac gaaccggcg cgaagccggc accggatgga
 8941 atgcgctttt cgacgcagat gtcagcgctt tgcgcgatgt tcgcggtctt tgccagcaat
 9001 ttccggcat gacccagag gtcgtcgcc gcacccacc cgggtcgga gctgccgtgg
 9061 cggacgatgc ggcgtcgct ccgggtgcgc aggcggatcg tgcgcccc gtccagcagg
 9121 cgaatacgg tctgcagcag gtggtagccg tcggcgcggc ggccggggat gcgcaggaac
 9181 aggttgagct tggccggggc tggccagtcc gaccagggca tggctaccgt cgcgttcacg
 9241 ggctggcgtt gcccgactc cagtcgtcga gcaccagcg caccggggtc tcgtcgcgct
 9301 cggcggtgac ccgcatcgcc aactccagc ctgcccggc ttgaaggcgc caatcgccgt
 9361 agtccagggt ccagccgtcc tgcgtcagcc gcgcagccg gccgtcgccc gcgaattgca

9421 gttgcgccgg gccgccctgc gtggcgcggg cgccgcgcac ccacgaggcc agcgctgtga
 9481 ccgggatttc ccagccggtg gcctcgcgca gcagctgcgt cgctcgcgca ccgctgcgcy
 9541 ggccgcccgtc gaggccttcc aggcgcgcgt tgccggcatc cccctcgatc cgccagctct
 9601 ggcgcggtgac cggcgcactg agcgcgaccc gatacgcgtc gccgtcctgc tgccagtcga
 9661 tgcggccgct gccgccgcgc tggccgttgg acagggccac ccggccctgc aggttccatc
 9721 gcgctgttcg tgccagcacc gcttcgcgcy cctgctggcg ggccacggcg gcggtgtgtt
 9781 cggccggcgcy caacggcgcy cgccggcgcy tggtagcgca accggcgagc aggcggcgcy
 9841 cgagaatggc cagcggcagt ccgcgcacgc tcatgcgcy gtcttcgcca gcgcygctt
 9901 gagtgcgcy ttgtcgggat cgagcttgc gccctgctcy aagaacctgc gcgcttcgcy
 9961 cttgtcgctt ttacccaca gtacttcgcy caggtgcgc gcgatttcgc gatccttctg
 10021 cagcgcaac gcccgccgca actccaccag cgcttcgcy ttgcgccga ggcgatacag
 10081 caccagccg tagctgtcga tgatcgcgcy attgtcaggc tcggcgaccc gggcgcgctc
 10141 gatcagctgc agcgcttctg tatagcgcy ggtgcggtcg gcaagcgtgt agccgagcgc
 10201 gttgagcgcy gcgacgttgt ccggtcggcy gaccaggatc cggcggaagt cggcctcgcy
 10261 gcggcgatg tcgtcgcygc gctccacc cagcgcgcy gcatacagca gcgcygcatc
 10321 gtccgggaac gcggccaggc cacggcgcta ggctcgagt tcggcgcygc cgttgcgctc
 10381 ctgctgccgc agttccgct cgaacagga ggcttgcg cycgctcct catcgcgcygc
 10441 cgcatcgcc tgcaagtggc gcaggtcgcy ataggcctcg tcggcgcygc cgagcttctg
 10501 cagcacgcy gcggcgcyca accgtgcctg ccagcgctgg tcgccaccg gcagctgcg
 10561 gtaccactgc aacgcaccc gataacgctc gaggaattcc gcggtctggc cgagcagcag
 10621 gcgccgttgc ggatcggtt tggcggaatc gcgttcaat gcttcataaa gcttgcgag
 10681 ggcgcccttgc tctcggcct tggccagcag cgacgcagc aggtgtagc tctggtcgtc
 10741 ctgcccccg cgcgcgagta ccgccccgc ggcgaccggg tcggcgaggg cgtcgtattc
 10801 ggagccacc gacagccgca ggtcttcgct gatccgggtg gcgtcgtcga gcccggcgag
 10861 gatccggcgcy gcgtcgtcgc cttgccccgc ttcgcycaac tggctggcat gcagcaacgc
 10921 caccgtggc tcgccccgga agcgcgcygc cacggcctcg acgatgcgcy cggtcagtgc
 10981 cggtgtctc agcgtcgtcg ccaggcctcc gaatgcgagc caggcctgca actggttcgcy
 11041 gatcaggttgc ccgtccacca tctcgcccaa caatttcgcy gagagcccg gatgcttgc
 11101 gctgcccagc gccgcccagc catggcgcca gcctccgccc tcagggtgc gcacaggac
 11161 ttccagtccg cgacggcgcy cgcgctcatc tcgcygcgc agggccaggg tgacttcagc
 11221 cgcgcgagc gccaacgatt gcggcgcgcy cgcgcgccat agttccagc cctcgccgcygc
 11281 acggcgctca tcgtcgccga gcagtgcgat ccgcgagcgc cgttcggcca gtccgcatc
 11341 gccgtcgga gcctttgcgc catcgaggta ccagcgcgcy gcctcgtcca gcttccccgc
 11401 ctgcaacgcy aactcgccgc cgatggtggc ctcgacggtg tcgtcgtcgc ggcccgacc
 11461 ccgacgggc gtggcggaac aggcagcag cccgagcatg gcgacgggca gccagaagc
 11521 gggtaaatac agcgaacgcy gcacggatg cgaccgggcy gtaaatggc ggccctgctt
 11581 cgccgaagc ttagcgcaag cgctgaata gacacgccac cggatgaccc tgctcgccat
 11641 cggcatcaac caccagaccg cgccggtcgc cttgcgcgaa cgggtggcgt ttgcccacca
 11701 cagcctgtcg gccgctgcy acaccttgc gaagctgcy tcggtgcagc aggtcgcgct
 11761 gctgtccacc tgcaaccgca ccgagctcta tgcggttgc gaggacgagc gccggcgct
 11821 gggcgactgg ctcgcccagc atcccagca cagcacggcy ttcggcgggc acgacctgc
 11881 cgctatctc tacgggaccc aggacgagc cgccgtgcgc cactgttcc gggtcgccc
 11941 cgggttgga tcgtgtgtgc tcggcgagc gcagatcctc ggccaggtca agcacgctt
 12001 ggcggtcgc cgagcgcgcy gcacactcgc cagccagctc gaccgctgt tccagcacgc
 12061 cttcgtcgc gccaaagcgc cgcgaccca taccgcatc ggtgccaacc cgttctcgtt
 12121 cgctcgacc gcggtgcgcy tggcgagc ttcgttcgcy cgtccgtcc actcgcggtt
 12181 actgctgtc ggcgccgcy aaaccatcga gctggtcgc cgccacctc cgagctccag
 12241 ggtcaagcgc ctgctgttcg ccaaccgtac cctggcccat gccaggacc tcgcccagc
 12301 gcatggcgcy gtcgctgcy cgctggacga gatcgaaac cacttgggcy aagtgcacat
 12361 cgtgttctc gccaccgcca gcccggaacc ggtgctgtc cgcgcgagc tcgcccggc
 12421 gctggcgag cgcaagcacc ggccgatgt gctgctcgc ctcgccgtgc cgcgcgacat
 12481 cgagccgagc gtggcgagc tgcaggagc gttcctctac accatcgagc acctggagc
 12541 cgcgatcgag gacaaccgcy gcagccgcy cgaagctgc gacggcgcy aggcgatcgt

12601 cgggctgcag gtctcgcgct tcatggaac gctggccgcc agcactcgcg tggaaaccgct
 12661 caagcggctg cgcgcgcacg gcgacagcgc ccgcgccgaa gtgctggcca agggccggca
 12721 gcagctcgcc gccggcaacg atccccaggc ggtgctcgag ttctctcgcc acacctgac
 12781 caaccgcctg ctgcacgcgc cgacggtcgc gctgcgccag gccgcgctgg aaggcgacaa
 12841 cgaactggca cgcgccagcg agaagctgtt cctggaagcg ccggggggtg agaaccggaca
 12901 acctggaacc ggtgaaacgg gtcccggag ctagggaagg cctgaaccat cggatgctgt
 12961 catcgacaat tggagctgt catcccagc gtagcgagg atgacaaatc agtaccccg
 13021 ggggtgacaga ccgatacgtt cgggatgaca accccggcac gttgtcgccg ttcccggtt
 13081 tcccagtgct cgtgtcgcc cgttctgtct tcatccccg ccgggacccc catgctcccc
 13141 agcctgcgcc gcaaacctga agccctggcc gaacgccgcg aggaactcga acgctgctc
 13201 gccgatccgg gcacgatcgc cgacgccgag cgcttccgcg cgttgtcgcg cgagtctcg
 13261 cagctggaac cggtcgcgtt ggcgctggcc gccgagaacc gggcgcaacg cgacctcgcc
 13321 gctgcgcagg cgatgcgttc ggacccggag ctggcgggc tggccgaaga ggaaatcgcg
 13381 gccgcgaacg cgcgcctcca ggaactggaa tccgagctga tggcgagct ggtgcggagg
 13441 gatccgcgcg acgacggcaa cctgtacctc gaagtgcgcg ccggcaccgg cggcgacgag
 13501 gccgcgatct tcgccggaga cctgttccgc atgtacgcgc gctacgccga gcgccagggc
 13561 tggagctgg aggtggaatc ggcgagcgtc ggcgagcacg gcggctacaa ggaaatcgtc
 13621 gcccgcgctc agggccgcgg cgcctattcg aaactgaagt tcgaatccgg caccaccgc
 13681 gtgcaacgcg tgcgcgacac cgaatcccag ggcgcgcatc acacatcggc ggcgaccgtg
 13741 gcgatcatcg ccgaggacag cagcgatgtc gcgatcgaga tcaaccggc cgacctcagg
 13801 atcgacacct tccgttcttc cggcgccggc ggccagcacg tcaacaagac cgactcggcg
 13861 atccgcatca cccacctgcc cagcggcgtg gtggtggaga accagaccga gcgcagccag
 13921 caccgcaacc gcgacaaggc gatgaagagg ctcaggggcg tgctggtcga ggccgaaacc
 13981 agccggcgcg aagccgcgca ggcgcaggac cgcaagctgc aggtcggcag cggcgaccgc
 14041 agccagcgca tccgcaccta caactaccg cagggccgga tcaccgacca ccgcgtcgaa
 14101 ggccctcacg tgtacgacct gcccacgtg ctggaaggca acctcgacgc gctcatcgaa
 14161 cgccctggcc gcgaacagca ggcgcgacga ctggcccgcc tgggcaggga atgaatgcct
 14221 ttcttctct ccgactgggg ggagaagagc ctgccctgag cctgtcgaag ggtgccgaag
 14281 ccgatgagg ggcggcgga gccggcgaca cccggatagc acgcatcagc cgcgagccga
 14341 ataaactcgg cccatggaca cgaccatccg ccgcgcggga ccgcaggacc tcgacgcgct
 14401 cgccgtgctg ttcgacgcct atcgccgctt ctacgggcaa ccggccgacc tggcgcgggc
 14461 ccgcgactgg ctgcgcacgc gcatgcgctt cggcgaatcg gtcgtctga tcgccgaacg
 14521 cgacggcatg gccgcggcct tcaccagct gtaccgatg cattcctcg tccgcaccgc
 14581 gcgcacctgg atactcaacg acctgtacgt cgcggagcag gcgcggcgct cgcggctggc
 14641 ccgggcattg ctcgatgccg ccgcacggtc cgcgcgcgag gacggcgccg ccggcatcag
 14701 cctggagacc acccgcgaca acgcggcgcc gcgggcgctg taccgcgccg cgggctggca
 14761 cgaggaccag acccagtggt attcgcttcc gctggaatga tcgcccgat gccgcgtgcg
 14821 gtcggttgaa accctcccta aactgcgccc aacctcgac cagtccggac accatgggca
 14881 ccgatcgctt ttcatcccc ctgtcccttt gcgtgctgac cgtttcggac tcgcgcacgc
 14941 tggccgagga cagctccggc gactgcctcg tggccgggct caccgcggcc ggccaccgcc
 15001 tgcacgctcg cgcgctcctg cccgacgacc gttacaagct gcgcgcgctg gtatcgcaat
 15061 ggatcgccga cgcggcgctc gacggcatcc tggtcacgg cggcaccggt ttcaccggcc
 15121 gtgattccac gccggaagcc ctgcttccgc tgctcgacaa gcagatgccg ggcttcggcg
 15181 agctgttccg ggcgctgagc ttcgaggaga tcggcacctc gtcgctgcag tcgcgcgctt
 15241 tcgcggcct cgccaacggc accttcctgt tcgccctgcc cggttcgacc tcggcttgcc
 15301 gcacggcatg ggacaggatc atccgcacgc agctggacgc gcgcaccggt ccctgcaacc
 15361 tcgcgcgctt cgttccccg ctcaggaggt gaccatgtcg atcgatgcaa cctgcgcct
 15421 gctcgcgcgc gcccgcgctt tgtacggcat ggcccgccg gcggcgcgca aaagccgcca
 15481 ggccgcaagc cgatgaagaa gctcaagcgc aaggattatg aaaaggcgct ggagccgatg
 15541 caggaggaaac tggcgcgctt ggcgcgctgg gcgcggcatg gcggcgccg catcgtggtc
 15601 gtgttcgaag gccgcgacac cgcgggcaag ggcggcgcca tcgatgcat ccgcgcgcac
 15661 ctcaaccgcg gccagtgcg catcgccgca ttgccgaag cgaccgagcg cgaagcctcg
 15721 caatggtatt tccagcgcta cgtgccctac ctgcccgcg ccggcgagat cgtgctgttc

15781 gaccgcagct ggtacaaccg cgccggcgctc gagaaggtga tgggcttcgc cagcgacgcc
 15841 caggtcaagg ccttcctgca gcaggcgccg caattcgaac ggcaactggg cgacgacggc
 15901 atcctgctgt tcaagtactg gctgtgctgc gaccaggcgc aacaggagca gcgtttcgcc
 15961 gaacgcctgg acgatccgct caagcgctgg aagctctcgc cgatcgacct gcaggcgcg
 16021 gcgaagtacg ccgactacac ccgcgcccg gaagccatgc tcaaggccac gcacaccgac
 16081 ttcgcgccat ggacgctggg cgacttcaac gaccagaagc gggggccgct gaccctgctg
 16141 cgcgattttt tggaccgcct gcccgaacac cgcggtgccc cgacacggg cgatttcgcy
 16201 ccgttgcccg gcaaaccgct gaaggaacgc tttggcgcg cgaatccgat cgccacgtac
 16261 cacggctgaa cgcgcgagcg gccgccacgc ggcgactgga ttcaaccgat cgaaaaatcc
 16321 gcgccggcct cgcttgggtc ggccggtgag cgcgcgacgc gatcgacccg cgcatgcyggg
 16381 gggccgtgcc ggagccatgc ggcgagtccg tcgatcaacg cgctgtcgcc ggcggccagc
 16441 acttctacgc ggccgtccgg caggttcctg gcatggcgcc gcaagcccag gcgcagcgcc
 16501 tcgtcgcggg tcgcgcgcg gaagcacacg ccctgcacct tgccgctgac caggaaagcg
 16561 gtgtgcccgt ccatgcgtgc gggcgcgta cccgcctgc gggccaccgg cggcgcgcat
 16621 cgcgcgctcg atgtccttcg ccgtgatcgg accgatgaag tgcttggcga ccttgccgtc
 16681 gggggcgatc aggtaggtcg tcggcaggcc gaccggggtc gcgaaatccc tggggcggtc
 16741 gtacacgtcg acgatggcga tcggatacac caccgggtgc ttctgcagga aggccttcag
 16801 ctcgcccggt tcgatgtcct cgtaggccag gccgatcacg tcgatgtgtt cgcgcatcgc
 16861 gtccagcgcc gacaggtcgg gcatttcctt caggcaaggg ccgcaccagg tcgccagaa
 16921 gttcaccacc acccagttgc cgcgatgggt ggcgaggtcg tagggcttgc cgtcgagcgt
 16981 ggtgactcgc agcgtgggct gttccggtt cggcgcgcc gcttcggtcg gttgcgccg
 17041 cgcggtatgt tcgcggcgcg tcgcggccgg cggctgcgcc gcaggctgcy gttggggcg
 17101 ctcgttcgcc ggccgcttgc agggggccag cgccagcaag gccaggcacg acatggtcaa
 17161 caggcggttc atgcggatcc tttcgggagg tcgtcgaaga cattggagac gcggcgctt
 17221 agcagctcgc gcaacggcag gcgcaactgg tcgagtctct ccatcgcgcc gacggccgag
 17281 gcgtcgtcgc ggcacggcag gtgcgcctcg gcgatcggga tgcgcagctc gcaggcttcg
 17341 tagagttcgg ccatggtgag gtgcgtcagg tcgcgggcaa gcagccattc gcccgaatcg
 17401 gcgcggctga ccagtccgat cgcgctcaat tccccgagca gttcctgcac cagcgcgctc
 17461 gtgagcatcg gctcgagctg ctgcaattcg tcgctgtgca ggccgcgccc cttgtgccgt
 17521 tgctcggcaa agcggccgag catccgcagc aggcctgaca actcgtagcc gttgggcagg
 17581 cgcatcgccg cgggctggta gcggaacgcc gacatcgatg cggcgcaacga agcgccgagc
 17641 aggatcgaca cccagccgaa gtagatccac aacagggaaga ccggcacgaa gcgcagcgcc
 17701 ccgtagatct tctggtaggc gccgaagctg cccaggtaaca gcgcgatgcc ccatttcacc
 17761 acttcgaaca gcagaccga cagcagcgcg ccgaccatcg cgtggcgcca cttgatcgtg
 17821 cgatgcggca ccccccggaa gatcgccgcy aacgcggcca gctcgatcgt catcgccgcy
 17881 agccgcagca tcaccgcttc cagcgcgcg cggggtcgg tctcgaacac cgccagcgcy
 17941 aagaaccgcy tcgacaacgc caggctgggt gccgcgacca gcgcgcccag cgtcagtacc
 18001 gtccagtaga cgaggaagcg cgtgaattc gggcgcgcg tcctgacccg ccagatgcyg
 18061 ttgaagggtg cctcgacgcc ggtcagggtg atcagcagcy aggccaccag cgcgatcacg
 18121 ccggcggtg tcagctgccc ggcatgccc gagatctgcy gcaggtagcc ctcgaccctg
 18181 cgcgccgagc tggggacgaa attcgaaaag atgtagtgc tgagttgccc gctccagtcg
 18241 ttgaaggccg gaaacgcaga cagcaccgcy aacaccacca tcgacagcg caccagcgcy
 18301 aacaccgtgg ttaggacag cgcgcggcg gcctcgaaca ggcggtcgtc gaggaaaccg
 18361 cgccacagga aggcgaagaa cgagcgcgcc cgcgcgggt cgcgcaggcg gtcgctccag
 18421 cggttgatcg agtccagcg ctccatcgcy gcagggtaac ccatgcccg tggggcgccg
 18481 agccggttgc cgccgcagc gcaggcgag gaccggccc cggcatgtg cgcggcgctg
 18541 gcgcgagcc atccccatc cttgcgcctt cattggagca cgtgatgccc gagatcctg
 18601 tgttgattta cagccgcyg ggttcggtg cgaaactgg acggcaggtc gcgcgcygca
 18661 tcggcgaggt cgatggcatg gccgcgcyt tgcgcacggg accgcgggtg gccgcgacca
 18721 cgcatgcccg gaggccccg gtgcccagc acggcgcgcc ctacgtcgag cgcccgatc
 18781 tcgccaatg cgccggcctg ctgctcgga gcccgaccg tttcggaac atggccgcyg
 18841 cgctcaagca cttcatcgac gggctcgcyg ccgaatggg cagcggcacc ctggtcgga
 18901 agccggccgg cgtgttcacc tccaccagca cgatgcagc cggccaggag tccacgctg

18961 tgtcgaatgct gctgccgctg ctgcaccatg gctgctgat cgcgggcatc ccgttcaccg
 19021 agcccgcgct caacagcacg cgcaccggcg gcacgccgta cggcgccagc catgtagccg
 19081 ggctgcagga cgaccgcaa cccagcgacg acgaagccgt gctggcgcg cgcctcggcc
 19141 gccggaatgc gcagctcgcc gcgaggctgg cgccatgact cgcacccgcg gcatcctctg
 19201 cgccgccctg ctgcgctcg ccgctgtgta cgcgtggtgg ttcgcgcgcg acgggcagtg
 19261 ggtggcggtc gtcgtgttcg cgctgccacc gctgctgctg gcgctctcgc tgcctcgcg
 19321 cggcgccgccc cgacccggct tctgggcccgg ggtattggcg ctgctgtggt tctcgcacgg
 19381 catcatggtg gcctggacgc gcgcccggga gcgcggtctg gcgctggcgg ccgtcgcgct
 19441 ggcgggtggcg atcgtgcttg ccgccagcct gcccgggctg cgcgcgcgct tcggcgcgcg
 19501 ttcgcggtct tcctgaacaa tctcggttg accgggctg ttacagcgac tccaggggat
 19561 cgcgcgacg tcccagccgc tgccagacac tataaggaca cctaggggctc gctcgtgccc
 19621 gccatcgcca gcatcagggc tagcgcttcc agggccgggc aggcgtgta tctgtcgcgg
 19681 cagcaatgta gtgcttgctt ggatcgatga ttgtccgacg attaggtgcc aacgtcggtc
 19741 gtactgtcca cgtaccccaa tgtaatcgcg gtctggggtg gcagcaggca agtcacgctg
 19801 aactggccat ggtcctggat cgtcccccag tcgccaagt cgtccggctc ccaaattagg
 19861 gcggcgccgg tcgatgtggg cggggctacg cccgttttgc tcacataagt caggttgaag
 19921 tcggcgctcg acgccacgcc ggtacaggac atatcgtgca aagcgccgctc aaggctggcg
 19981 acatacagcc atatcgccgt gtttttgcta ccacctgaaa cagccccctc tcccgcaggg
 20041 tagacgcagg tgacgaaagt gctggctcag ccttcgttgc gcacgcctgt cgcccggttg
 20101 cggatggcg aggacgtcgt gggcaatgtc aatcggcaga aattggagcc caccaaggcg
 20161 atctggcgct cctgtgtcgc ggccgtggca tgggaattca tcaactcccc gcaggcaagc
 20221 accgccaacg cgatgatgtt ggtgtaaagc tgcattgtga tatccatggc tgggatgttt
 20281 gctgcgcgtc accatcgccg cgcaaggtag aagtgtcccc atgggtcccc atggcggtc
 20341 gggacagcgt cgactccaga gtgggcccgt gttcgcagtc ggagggggat ctgcggtaca
 20401 acgtttacgg gtagcgccg caggtagcat ccgccttctt accggttcgt ggcgtttccg
 20461 ccgcgagccg cgtagatgcg ggccgtgggtg ctgccatcgg cattgcgtac gtattctgcc
 20521 cgcgcccgtg ccagcttctg gccgctgggtg gccaggtagg tcaggaccca gtcttcggcg
 20581 ttcgccgatg ccgcaaaagt gccaacgat gcgcagctgc ggcggcgcca atccacatcc
 20641 ggggtgcgcg gaaccaagcc ggtactgttc atcgccctgc cgaccatcgt ctccagcaac
 20701 tggctctcga cgagggtgc ggccgggtcc ctgggatcgc tggaaaattg cgcttcgagc
 20761 gcctgcaatc gtgcgcgatg ttgttgctcg gcctgttccg gagtcaagtc gggcaatgga
 20821 ggtgcgggac tcctggaggc gccgtctcgt gcaccgaacg catcccgga aacattggtg
 20881 gcgacacctg cgggataacc cgtgggagct tgctcgcgct gcgcttccag cgctgcagc
 20941 cgcgccagga tcctggccat ctgcagctgg tcgcccgcg gtcgttctct gtgggtccag
 21001 accaggggcg ctgcgaccag caatgcaacc gcgagcactg tcccgcgat ccatggcagt
 21061 tggttcctgc gttgcgaatg atgtgcggc attgcctctc cgtaagtcac ggaacaaata
 21121 cctgcgcaac tcgaatcaa cccctcggtg ctgcatgagt cccgtgcgccc tggattgcc
 21181 ggcgcacggg cttgttgcca accatgaccg gatcaagtac ccacgttggc cgtactggtc
 21241 acgtatccac tggatgatcg ggtctggggc ggcagcaggc aggttacgct gaactggccg
 21301 taatcgggaa tggctcgtct gccgcccga tccccggct cccagattcc gtagccaccc
 21361 gacttggtgg tcaggggcgt gatcgtcttg ctcacgtaag tcagggtgaa gaagccgtcc
 21421 gaggccacgc cggtagacga cacatcgtgc gacacgcat ccaggctggc gacatacagc
 21481 cagatctcgt gtctcgagct tccgcctgca gtgcgcccct ctcgccaggg gaaggcgag
 21541 gtgacgtacg tgctggtcgt gccttcattg cgtacgcccg tcgcgcggtt gcgaatgggc
 21601 gaggacgtcg tcggcaaggc cagctggcag aaattcgggg caatcagggc gatctggcgg
 21661 ctctgcgtcg cggcctgcgc ctggaagccc atcagcccc caccagcaac cagcaccacg
 21721 ccgacgaact tgctgaaaag ttgcattcct acttctccac tgatggggcg gacgcccggg
 21781 catccataca aacggcgat gcggcgccca acggacagcc gagccagcca gatgggtccg
 21841 ccagggaac ccggatacaa tgcccgaac cggcacagcg ctctggacgg atcatggagc
 21901 agctcctcgt cgtcaccacc ggcggcacca tcgacaaggc gtacttcgac gacaagtcgg
 21961 actaccaggt cggcgaaccg cagatcgggc gcatcctgca ggaactgggg gtgggcttcc
 22021 gcttcagcgt ggtcccgatc ctgcgcaagg actcgtgta catcgacgcc gccgaccgcg
 22081 aactgctcgc cgcgaccatc gccgcgacg atgcgcgcca cgtgctgggtg acccagggca

22141 ccgacagcat ggtcgaaacg gcgaaggtgc tggcgagcat ccccgcaag accatcgctc
 22201 tgaccgggtgc gctgaaccg gcgcgtttcc gcggctccga cgccgagttc aacatcggta
 22261 ccgcgggtcgg cgtgtgcag tcgctgccgc caggggttta tatcgccatg aacggccgca
 22321 tctgggatcc ggcgaaggtc cgcaagaacg tcgcccga cgcgttcgag gcgctcgact
 22381 gaaatgccgg tgcacagggt caagcgggtc tacgagccgc ccgaggagga cgacggcctg
 22441 cgctgtctgg tcgatcgctt gtggccgcgc gccctgtcca aggaccgggc ccgggtggac
 22501 gagtggctta aggagatcgc gcccagtcgc gggctgcgca agcagttcca cgcacgcccg
 22561 gagcgtctgg acgacttcgc cgacgcctac cttgccgagg cgccgcccga ccccgcatg
 22621 cccggcctgc gcagcctggc gcgacggcag cgcacacccc tgctttacgc cttccacgac
 22681 gagggccaca accacgcgct gctgctgcgc gagggtgctgc ggtagcgag gcgctgtgtg
 22741 gcgcccctgc cccagtgccg cggggacatg gaaggggtat ggaaggggag cttctccccg
 22801 ggtgggcacc tctgcggaga aaccgacggc gcaaccttca gcggacacgc cgtgaatatc
 22861 ccctgttagg ccctgcggag gcttccatgc cgccgacggg tccgtgggga gctcggacca
 22921 ccgggtccta caagtgggcc ggcccctgtg gaaagaagcc tgccgaagcg acgagctcca
 22981 agctcgcgca atccacggtc gctgcccggc gcaccggggg acggagagtg gaccagggat
 23041 ggcccgcgct ccgactgtcg tcaggcggac gcagtacccc ggcccacggg ggcagcgact
 23101 ccatccgaaa ccgcgcccc atcgaaaaga gaagcaccat aaaaagaccc cggcgcaagg
 23161 ccggggtcgc aagtccatca gggggtaatg gtcgaggcga tcagaagggt acgctccagc
 23221 tgcgatgta gccggtgtcg ccggagtaca cgtccttgac ccgacgcttc caccgcccgt
 23281 tcaacgcctc ggtggagagg ttgacgggtg aggtggcaac cacgtgtgcg gcgctgtcgc
 23341 taccgctgga gttcttcagg cggtaggtcg agccgtccgg cgcgaccagg tcgatctgca
 23401 ggtcgccgcy gtaggtgtgg acgatgttga ccgacaccga ggcgttgctc ggagcgttgc
 23461 cgctacggcc ggagaccgtg atcgggctgt cgatcgtgcc ggggctgggg atgtttagt
 23521 tgctcccgtt gctgtaggtc tgccgcccac caccgcccgc ggtgctgtag tcgcccacca
 23581 ggctgacgcc cgagaacgcc gaatacgctt tcaagcgcac gtagtaggtg cccgcccgcg
 23641 gcgctcgcaa ggtgcagctc tccgcgttgc cgctcttgta cggacggcaa tcgtacgacg
 23701 tgctggctcg cgcgctgcgc aacttcacgt acatgtcggc atcgccgggt ccgcccggaca
 23761 tggatgaacg caggttgagc gcgcccgtcg gcacgctcag cgtgtagtcc accgcaactgc
 23821 cggtcgccgc cgacaggccg gtcgcccgca ctcccttggt cagcacgcct ccgcccgcgc
 23881 caccaccacc ggtggtgtag tcgcccacca ggctgacgcc ggagaacgcg gagtacgcct
 23941 tcagccgcac gtagtaggta cccgcccagc gcgcccga cgtgcaggtc tcggcattgc
 24001 cggacttgta cggacggcag tcgtacgagg tgcggtcgg cgcgctgccg aacttgacgt
 24061 acatgtcggc gtcaccggtg ccgcccggaca tggatgaacg caggttgctc gcaccagccg
 24121 ggacgaccag cgtgtagtgc accgaactgc cggctgatgc cgactggttg gtcacgggca
 24181 cgcccttggt caacggaccg ccggtgtcgc caccgcccc gcccccgccg ctgcagctca
 24241 cgccgaccgt ggtgaaggcg gcggtgacgt tggccacgct gtaaccagg tcgcccggccg
 24301 ccgtttcaac gccacaggcg gcggtgttgt aggtgctgct cgcggtccag tacagggcat
 24361 tggcgccggc gaagacctg aacgccttcg gcgtatccca gccggcggtc ttggccagggt
 24421 tgagaacgc cttgtgttac acgccagagg agtagtgac gtccatgccg ctggtgtagt
 24481 cgccggcgct gtcgatcgag ccgcccctct gggtcggtat gcacatgtag cgcagcgaac
 24541 cgctgccctt gaagatctcg ctgccgacga ggaagtcgtt gctgcccttc cagtagtgct
 24601 cgggtggctt gccgcccag tcggagaagg cttcgttgat cccgcccagc tggccggaat
 24661 aggtcagggt ggagtgtgc tcggtgaagc cgtgcgagac ctcgtggccg gcgacgtcgg
 24721 cgctgaccag cgggtagaac gtgctggcgc cgtcgcgcaa gctcatggcc gagccgtccc
 24781 agaacgcgtt ctcgtagtgt ctgctgtagt gcacgcgcat caccagctgg aagctgagcg
 24841 cgtgttagcc ggtgtaggcc gggatcatct tgggtgacac gccgcccgaag aagtgcgcat
 24901 cgttgatcgg cgagtaccg ccattgatgg tctttagtgt gttgcgccc caggcgtagg
 24961 aaaaacgcgt ggtgccgctg gtgccgccc tgaggttcac cgacttcacg ttggcggtgt
 25021 tcatggtgca ggtggagccg ctctgcgcca cgtcgaggta gccgtagatg ccgcccggagc
 25081 cccattcgta ctggccggtc ttctggttgc gcgcccggcc ggtgccgac tcgcccgggtg
 25141 tcaggccgtc ccagtgtctg agcaccttgc cgttgccgac gtcgacgac acgaacgggc
 25201 gggccgggct gccgccttgc gcggcgctcg cgaacatctc caccagccat gccaggtgcg
 25261 cattgtcgcg ttcgtcgacg aagatcatct gctgcgcctt ctcgcccga atcttcaacg

25321 cccccgagcg attgccgaag gccgcgtgcc gggcgagggt caacgcctgc gccctggcga
 25381 ccttggcggc cacggtcggc aggtcgcccg ccaggccttc cagcttgccg ccgaacagggt
 25441 tgcgcgcggc gccgtccttg cgttcgctga cgatcacctg ctgcgcgaac accgggatgc
 25501 cgcggaagggt ctgctggtag cgatagtgcc ggctaccgtc gcggtcgacc gcgctgttga
 25561 gcagcttcaa cccgaatcg gcatccaggg acagcaattc ggcatggcgg tcggccgatt
 25621 gcgcccggcac gccgaggcgc tgggtggcgg ccttgactg cctgtaaacc tgggcgacgt
 25681 cctgcttggt caggtcgacg cgatgggccc cctgcgcgga catggcggcg aaggcgaacg
 25741 ccaggggcga gagcacggca aggtaacgaa cagcactgct ggaatgcac acgggtgcct
 25801 tagggttgac ggttacggtt gatggcgcgg ccaccggcg gcgcaccggc agcgactccc
 25861 ctcatctgct gaacgcgacg atggcgctgc gctctcccc caccgggtgc cgggcgcctg
 25921 tcctggcggc ctgctgtcgt ccggcatcca tgcgcctccc cgactgtctc gaatgggccc
 25981 ccaatgcggc aatggccctt cccctgcccg cgggcggtag cccgggaagc gaggtcgaag
 26041 ccaggaggcg acgcgatcga acctgcgccg cgaaagtac gctatgcgac agtcagctgg
 26101 cgttcaagct gcaggcgacg cttttccggg gctgaaaaag tgacatggat cacatttcga
 26161 taccgtaacc ggttgcatat ttgctgtaac cgtttaccgc cggcgtgacg cggggtgtca
 26221 gagcatcccg gtttccaggc gcgcggcctc ggacatcatg ttgcggttcc acggtgggtc
 26281 gaacaccagt tcgacgtcgg cctcgccac ggctcgggac agttccagct tgttgccac
 26341 gtcgtcgacc aggatctcgc ccatgccgca gcccggcgcg gtcagggtca tcttgacctc
 26401 gaccaggcgt tgcctgctct cgcatgctgt caccaccgac tcgtacacca ggcccaggtc
 26461 gacgacgttg atcgggatct ccgggtcgaa gcaggtagc agctgttcc atacgagttg
 26521 ctgcaccgtc tcgtcggagg ccccgaggag caattccagg ggctcgacgg gttccttgcc
 26581 gatggcgctc gcacctctgc cggcgatccg gaacagggtt ccttcgacga acaccgtgta
 26641 gctgccgccc agggcctggg tgatgtagcc gacgctgccg gccggcaggg tgaccgcgtc
 26701 gccctgcggc accagcacgg cggcacagtc gcgctcgaag cgcacgggtt cgctgctgcg
 26761 ggaatacatg ttcatggcca agtgggcttg cagcatccac gctgcgcgaa gggtcattc
 26821 taccgcgcg cgagcgtat gctgtggtgc ttgccgacac gaagcccgcc atgaccgacc
 26881 ccaccgcgcg cccggcgac cggcgttgc aggccttgtt cgcgccgga ctggtggcgc
 26941 tgggcagctg cgggttcgct gccgcctgga tcctgctcgc gttcgcgcgc gaccgccaat
 27001 gcagctggat ggcggtttcc gccgcgatcg acgcggtgtt gttgctgcgg ctggcacgga
 27061 tgccggcggg cgtgccgcgc atgctcgccc ccatgctcgc cactccgtg tccatcgtgc
 27121 tcgccaactg ggggatcgcc gcggcccaga tcggcaagct gatgggcctg ctgccctggg
 27181 agtcggcttt caagctgggc ccgcggtacg catggacctt ggccgggctc gccaaacagc
 27241 cggccgacct ggcctggctg gcggtcgctt tgctgctggc ggcatcgct tcgcgctgag
 27301 tccgggtcag gccatggccc gagcgggcgc ggtcaatgcc cgccgtccag cgccttcagc
 27361 tcgctgacca gcgcaccggc catctccgcg ccgtcgccgt acagcatcct ggtgtgtgcg
 27421 ccgtagaaca gcgcgttctc gatcccggcg aagccggtgc ccttgccgcg cttgatgacc
 27481 acggtgttct tcgactgcga gacgtcgagg atcggcatgc cgtagatcgg ggacgcggga
 27541 tcggtcttcg ccaccgggtt gaccacgtcg ttggcgccga tcaccagcga gacatcggtg
 27601 ttggggaact cggggttgat gtcgtccatg tcggcgatca ggtcgtaggg caccggggcc
 27661 tcggccagca gcacgttcat gtgcccgggc atgcgcccgg cgaccgggtg gatcgcgaac
 27721 ttgaccttga cccgcgcctc ctgaaggcgc tgcgccagct cccagatctt gtgctgcgcc
 27781 tgcgccaccg ccatgccgta accgggcacg atcaccacgc gctgcgcgta ggccatcatc
 27841 gcggcgacgt cgcgggcctc gatcggcttc tgcgcccccg cgatctcctg cgcctgcccg
 27901 ctgctgccgc cgaaattgga gaacagcacg ttgcggatcg gccggttcat cgccttggcc
 27961 atcaaccgcg tcagcagcat gccggccgcg ccgaccatca tgccggcgat gatcagcgcc
 28021 tcgttgccga gcacgtaacc ctggaacgcc accgccaggc cgggtcaaggc gttgtacagc
 28081 gagatcacca ccggcatgtc ggcgcgcgcg atc

7.2 Angaben zur Insertsequenz von pTWP13

7.2.1 Lage der offenen Leserahmen

ORF	Gen-name	Annotation	Sequenz-Lokalisation	Orientierung
1	<i>orf1</i>	Hypothetisches Protein	1-631	-
2	<i>acyP</i>	Acyolphosphatase	870-1151	-
3	<i>yneN</i>	Mögliches Thioredoxin	1177-1755	-
4	<i>rbn</i>	Ribonuklease BN	1752-3029	-
5	<i>trp</i>	Tryptophan Repressor Bindepotein	3101-3694	+
6	<i>orf6</i>	Konserviertes hypothetisches Protein	3691-4023	+
7	<i>aspG</i>	Asparaginase	4132-4620	+
8	<i>mprB</i>	Metalloprotease	5197-7782	-
9	<i>orf9</i>	Hypothetisches Protein	7798-7953	-
10	<i>orf10</i>	Konserviertes hypothetisches Protein	8208-8762	-
11	<i>orf11</i>	Konserviertes hypothetisches Protein	8838-9281	+
12	<i>orf22</i>	Mögliche NAD(P)-Transhydrogenase β UE	9256-10650	-
13	<i>orf23</i>	Mögliche NAD(P)-Transhydrogenase α UE	10647-10952	-
14	<i>orf24</i>	RNA-Polymerase ECF-Sigma Faktor	11124-11711	+
15	<i>orf25</i>	Hypothetisches Protein	11708-12103	+
16	<i>aspG</i>	Konserviertes hypothetisches Protein	12108-12566	+
17	<i>orf27</i>	Pyridin-Nukleotid-Transhydrogenase	12846-13979	-
18	<i>orf28</i>	Konserviertes hypothetisches Protein	14039-16273	-
19	<i>mprA</i>	Konserviertes hypothetisches Protein	16400-16999	+
20	<i>orf23</i>	Exodeoxyribonuklease	16996-17916	+
21	<i>orf23</i>	MutT-ähnliches Protein	18205-18765	+
22	<i>ptnB*</i>	Konserviertes hypothetisches Protein	18983-19960	+

7.2.2 Insertsequenz von pTWP13

```

1 gatccgcaag ctttcagca aggaagaagt gttgctgcc tagtggtaga tcgctgcat
61 gaattgcaa gttgctcagg tccaggctctg ccgctgcaat gccagttgc ttccctaggt
121 gttcccgaag ctgcgtctgc tggcttctgc ccaagtcgga caattggcac tgattgacaa
181 aaacattgtc tgccctcttc agcgagacct tcaactccgac attcgagaca aaccggagct
241 tagtggggtg gcccgagaag tcaactcctgc gcttgtatag tttcgccaga ggcgacgcat
301 ctgcgctctt gtccttacgc tgcggcctt ccgttcggag gtcttttagc ttccaccac
361 ctgtctgctc gttctttttt atctggcaga actcgaccgc tgtgggcgaa tcggcagagt
421 ccagaagcgc tacgtcttct ttgacttcca taccactgc aaagtcttcg tcgcttgca
481 gcagttctgc tactgcaaaa gctccccacc attgttggtg ggtaaaccgg agagtggcat
541 ttgcaccggt cttctcgtcg gtcggcttgt cccgaagtag cagctcaaga gcatccggtt
601 tgtgctctcg gtcacctaac aggtccgtca tgcgctttcc cttgctcccc tattcgagc
661 gtacgcgcat gcagggcgaa tgaaattcct gaggccagaa aaagcaaagg acgaccccca
721 ggacgccctt tgggtgacta gagacatcta ctcgtagaca cagccctta atctcaatta
781 acccattgat ctaaatgatc ttttgtctgg cggtagcgga agtgggatcc tgtgccatca
841 gtggcagcac aggtcgaaa gccacccac tagccaccg aaaaccccg catcgaccg
901 ccccatccg tctcatgatg ctcaaccgca tccaccgcg ccagcgcg ggcgtgtgc
961 agccattgag ccagggcgtc gagcgcgcg gcgttgccg tggccagcac ttcgacgcg
1021 ccgtcgggca ggttgcgggc gtggccgcgc aggccaggc gcagtgcctc atcgcggtg

```

1081 ctggcgcgga agcacacgcc ttgcaccttg ccgctgatca gaaaacgcgc gctgcggtcc
 1141 atggccggca tgcccgtgtg gaagggaaac gcgcgctcag cccgcctgcy ggccgcgggc
 1201 ggcagcgatc gccgcgtcga tgtcggcggc ggtcaccgga ccgaggaagt gcttggcgag
 1261 cttgccgtcg ggcgcgatca ggtaggtcgt cggcaggccg cgcggcggtg cgaaatccc
 1321 cggcgggtcg tagacatcga cgatcgcgat cgggtagacc accggtatgct cgcgcaggaa
 1381 ggctgcagc ggcggcggtc cgatctcctc gtaggccagg ccgatcacct cgacgtgctc
 1441 gcgcatcgcy tccagcgccg acagttccgg catctccttc aggcacggcc cgcaccaggt
 1501 cgcgcagaag ttgaccacca cccagtggcc gcggtgcgcy gccaggtcgt agggcttgcc
 1561 gtcgagcgty gtgacctgca gcgtcggctg ttgcggcaac gccggctccg ctggcgccctg
 1621 cgcggcgggc ggcgcggccg cgggcgcggg ggtcgcggcc ggctggtcgy cggcggggtg
 1681 ggcagcatcg ccggcggtt tgcaggcgcc caggccagc ggcggccaggc aggcaggca
 1741 caacagacgy ttcattggga ttccttcgac agggcttcgt agacgttga cagcgcgcy
 1801 ttgagcagtt cgcgcagcgy caggcgagc tggtcagtt cttcgcgcgc ggcgatgcc
 1861 agcgcgctgt cgtggcacgy cagccgcgc tccggcgatc ggatgcgcaa ctggcaggcc
 1921 tcgtacaggt cggccatgct gaggtcgtcc aggtcgcgcy ccagcagcca ctgcgggat
 1981 tcggcgcgcy tgaccacgcc gatggccgcy agctgccga gcgactcctg gatcatggcy
 2041 tcggtgagga tcggctccag ctgctggatc tggtcgctgt gcaggccgcy gccctggcy
 2101 cggctactcg caaagcggcc cagcaggcgc agcaggccgt agagctcgta gccctcgggc
 2161 aggcgcatcg ccgccgctg tagcggaa cccgacatcg acgcccaaa cgaggcgccc
 2221 agcaggatcg agaccagcc gaagtagatc cacagcagca ggatcggcy agcccgagc
 2281 ttgccgtaga tcttctggtg ggagccgaaa ctgccagggt acagccccag cccccagcy
 2341 agcagttcga acagcgccac cgacagcacc gcgcggccca ccgcgtggcy ccagtgcacg
 2401 gtgcggtgcy gcacctgcy gaagatcgtc gcgaacccg ccagctcgat cgtcatcgcy
 2461 gccagccgca gcatcaggcy tccagcgcy cggcccggt cggcttgga caccgccagc
 2521 gcgaagaact gggtcgacag cggcaggctg gccgcgcac ccagcgcy cagcgtcagc
 2581 accgtccagt acaccaggaa gcggccaaac ctggggcgcy tggctctgac ccgcaaatg
 2641 cgggtgaagg cggcctcgac gctggtcagc gtgatcagca ccgagatcac cagcgcgatc
 2701 acgccgaccg cggtcagctg cccgggtgtg gcgaggaact cggccaggta tcgtccagc
 2761 gcgcgcgcy agctgggcac gaagttgtc aacacgtagc gggtagcgc gtcgccccac
 2821 ttctcgaa caagggaatcc cgacagcacc ccgaacgcca ccatcgacag cggcaccagc
 2881 gcgaacacgy tgggtgtagga cagcgcgccg gcggcctcga acaggcggtc gtcgcggaag
 2941 cgcgcgcaca ggaaccgga gaacgcactg gcgcgcgcgc ggtcgcgcag gcgatcgctc
 3001 cagcggttga tcgagtccag cggttccatc gggcgaaggg taccgcatgc cccgcgcgc
 3061 cggcgccatc gcctatgctt gccggcgaca ggaggacgc atgcccgaga ttctggtgct
 3121 gtactacagc cgcggcggtt cggtgcgcy gctggcgcy caggctcgcy cgcggctcgy
 3181 cgaggctcgy ggcattggcy cgcgcctgcy cagggtgccg ccggtggccg cgaccacgca
 3241 ggtcgcgagc gcaccggtgc ccgagcagcy cgcgcctac gtcgagcgcc gggacttgga
 3301 ggaatgcgcc ggctgtgctc tcggcagccc gaccgcttc ggcaacatgg ccgcgccgt
 3361 gaaacacttc atcgacgggc tcggcgccga atgggcccag ggcaccctgg tcggcaagcc
 3421 ggcggcgcy ttacactcca ccgccacct gcacggcgcy caggaaagcca cgctgctgtc
 3481 gatgctgctg ccgtgctgc accacggctg cgtgatcgcc ggcatcctgt tcaccgaggc
 3541 ggcgctcaac accaccgca gcggcgccac gccctacggc gccagccagc tcgcccgcgc
 3601 gcaggacgac ccgcagccca gcgacgagga ggcgctgctg gcgcgcgcgc tcggccgcgc
 3661 catcgcgag cttggcgcy ggctggcgcc gtgagccgcy cgcgtggcgt gctcggcgcy
 3721 gccctggtcg cactggccgc gctgtatgcy tgggtgttcg ggcgcgagcy gcagtgggtc
 3781 gccgtggccg tgttcgcgt gccgcgcgt ctgttgccgc tggcggtggc gcgcggccgy
 3841 atccgcgcy gcttctgggc cgggtgctg gcgtggcgt ggttcgcgca cggcgtcatg
 3901 gcggcggtga cgcggccgcc cgaacgcgcy ccggcgctgy ccgcgaccgc gctggcgctg
 3961 gtgatcgtgt tcgccgccag cctgcccggg ctgcgcgcgc gcttcgccc caagcgcgcc
 4021 tgagcgccgy cgcggccgca tctttgccc gcccgatgcy cccccgtgc gcgcaggcgc
 4081 gcctgggcy ggcggtataa tctgcgccc gcaacgcgac ggaacgaggc gatggacgag
 4141 ttactggtcg tcaccaccgy cggcaccatc gacaaggtgt atttcgacga caagtcggac
 4201 taccagggtg gcgagccgca gatcgccac atcctgcaggy agctgggggt ggcgttccgc

4261 ttcaacgtga tcccgatcct gcgcaaggac tcgctgcaca tcgacgccgc cgaccgcgag
 4321 ctgctgcgcg ccaccatcgc cgcgcaggac gcgcgtcatg tgctggtgac ccatggcacc
 4381 gacagcatgg tggaaacggc caaggtactg gcgagcatcc ccggcaagac catcgtcctc
 4441 accggggcgc tgaatccggc gcgctttcgc ggctccgacg cggagttaa catcggcacc
 4501 gcggtcggcg cggtgcaatc gctgccggcg ggcgtgtaca tcgccatgaa cggccgcac
 4561 tggaaacccg ccaaagtgcg caagaacgtc gccgccaacc gcttcgaagc gatcgactga
 4621 aagaagcggc accgcaccgc aggcacaccg cccggcacgc cccctcccc acgggaccaa
 4681 acggctcatc tccaagggc gggaacctct gtgggaagaa gccggcgggc cgacctgtgc
 4741 gggacacgcc gtgaatccgt ccctggaggc tcggcgggcgt ctccaaagtg cgggggcctc
 4801 tgcgagggaa gccggtagcc cggcctgtgc gggacacgcc gtgaatacgt ccatgtaggc
 4861 tcctcggcgg catccatgcc gccgacggtc ccgcacaggc cggggccaccg gcttcggggc
 4921 aagttagccg gcgctttag agaaaccagc cgaagcaacg atcaccaatc tcgcgaaggc
 4981 gccgcccggc gggctggggg tgcggagagc aggccatgga tggcctgcgc cccgacttaa
 5041 agacaggacg tcgtcaggag ggcgtagcac cccagcccg ccggcgggcg actcagtcgc
 5101 aagccgggtc cggcagtcgg gagaaaggcc ataaaaagac cccggcgcaa ggccggggtc
 5161 gcaagtcctg caggggtaac actcgcaaa tcagaagggtg acgctccagc tgtcgatata
 5221 gccggtgtcg ccggagtaca cgtccttgac ccgcagcttc cagcgccgt tcaacgcctc
 5281 ggtggagagg ttaacggtgt aggtggcaac caggttgcg gcgctgtcgc tgccgctgga
 5341 gttcttcagg cgttaggtcg agccgtccgg cgcgaccagg tcgatctgca ggtcgcgcg
 5401 gtaggtgtgg acgatgtga ccgacaccga ggcgttgcgc ggagcgttcc cgctacggcc
 5461 ggagaccgtg atcgggctgt cgatcgtgcc ggggctgggg atgtttagt tgctccggtt
 5521 gctgtaggtc tcgcgcgcac caccgcgcgc ggtgctgtag tcgccgacca ggctgacccc
 5581 ggagaacgcc gattacgcct tcaaccgcac gtagtaggtg cccgccgacg gcgtggcgaa
 5641 ggtgcagctc tccgcgttgc cgtcttcta cggacggcag tcgtaggagc tgtcggtcgg
 5701 agcgtgcgcg aacttgacgt acatgtccgc atcgccgggtg ccgccggaca tgggtgaagg
 5761 caggttgag gcaccgctgg gcacgctcag ggtgtagttg accgagctgc cggtgccgc
 5821 ggacaggccg gtggccggca cgccttgggt cagcaccgcg tcgccaccgc caccaccgc
 5881 gccggtgtcg tagtcgcga ttaggtgac gccggagaac gccgagtag ccttcagccg
 5941 cagtagtag gtgcccgcg acggcgccgc gaaggtgcag ctctccgagt tgccgctctt
 6001 gtacggacgg cagtcgtagg agctgtcggg cggagcgtg ccgaacttga cgtacatgtc
 6061 ggcatcgccg gtaccgcgg acatggtgaa ggtcaggtt ctggcgccg ccgggacggc
 6121 caggtgttag gtcaccgagc tgccggtgct ggccgagatg ttggtgaccg ccacgccctt
 6181 ggtcagcgga ccgccggtgc tgccgccacc gccgccgcg ctgcagctca cgccgaccgc
 6241 ggtgaacgcc gcggtgacat cggccacggt gtagccagg tcggtcgctg cggctgtcac
 6301 gccacaggcg ccttcattga aggtgctgct ggcgtccag tacaggcggt tggcgccggc
 6361 gaagacctg aacgccttcg gtgtgtccca gccggcggtc ttggccagg tgcagaacgc
 6421 cttgtgttac acgcccagg agtagtgac gtccatgcc ctggttagt cggcgcggtt
 6481 gtcgacgag ccgccgtcct gggctcgatt gcacatgtag cgcagcgaac cgctgccctt
 6541 gaagatctcg ctgccgacga ggaagtcgtt gctgcccttc cagtagtgct cgggtgcctc
 6601 gccgcccatt tcggagaagg cttcgttgat gccgcccgac tggccggaat aggtcagggt
 6661 ggagtgtgc tcggtgaagc cgtgcgagac ctcatggcg gcgacgtcgg agctgaccag
 6721 cgggtagaac gtgctggcg cgtcgccgaa gctcatggc gagccgtccc agaaccggtt
 6781 ctgtagttg ctgctgtagt gcacgcgat caccagctgg aagctgagcg cgtttagacc
 6841 ggtgtaggcc gggtagatct tggtagtacc gccgccgaag aagtgcgct cgttgatcgg
 6901 cgagtagccg ccgttgatgg tctttaggt gttgcgcggg caggttagg cgtacgcggt
 6961 ggtgccgctg gtgccgcgt tgaggttagc cgacttcacg ttggcggtt tcatggtgca
 7021 ggtcgagccg ctctgcgcca cgtccaggta gccgtagatg ccgccggagc cccactcgta
 7081 ctggccggtc tcttggttgc cgcgggggc ggtgccgata tcggcggtg tcaggccgtc
 7141 ccagtgtctg agcaccttgc cgttgccgcg gtcgacgatc acgaacgggc gcgtcggcgc
 7201 gccgccgtt gcgctgtcgg cgaacatctc caccaccag gccaggtgcg cgcggttgtt
 7261 gtcgtcgacg aagatcatct gctgcgcctt ctgcgcgag atcttcatcg ccgccacgcg
 7321 gttgccgaag gccgcatgcc gggccagcgt caacgcattg gccctggcga cggcgccggc
 7381 caccgcggc agttcggccg cgagcccttc cacttgcgc ccgaacaggc tgcgcacggc

7441 gccgtccttg ccttcgctga cgatcacctg ctgcgccgaac accgggatgc cacggaaggt
 7501 ctgctggttag cggtagtgcc ggggtgccgtc gcggtcggtg gcgctggtca gcagccgcag
 7561 cccggaatcg gcgtccagcg ccagcatttc ggcatggcgg tcgccggcct tggccggcat
 7621 cccgagctgc gcactgacgg ccttgtagtg cttgtagagc tgggtcgcat cctgcttggtg
 7681 caggctcgacc cgggtggccg cttgcgcgga cacggcggcc agcgcaaacg ccaatggcga
 7741 aagcacggca agacaacgaa cacgactgct ggactgcatc atggtgtcct ctgacggtta
 7801 cggttggtgg cgcggtcgcc cggcgggcga tcggtcgagc tccccacatt cgcgcggcc
 7861 ggacgacagg acgtcgtcgc cgaccgcctc cacacgggtg cctgtgaccg tgatcgcccc
 7921 cgcgaccggc ctgcgtaacc ttccagcgtc catgcttccc ccgaagcttc ggaatggggc
 7981 gcggcagcgg caaggggccc gcacggcccc gcaacgatgc cgggtggtgc aaatcggacc
 8041 ctggggcgcc cccgaatctg gccgcaaac gacgctaagc gacaccccg gccaaactcaa
 8101 gctgcgggcg cagcaattca ggcgtcataa acgtgatcca catcacgtat tgagaacgta
 8161 accgattcca attttcatgg aatcgtttac aggcggcgag gcggcgctca gtacatcccg
 8221 gtttccaggc gcgcggcctc ggacatcatg ctggccgtcc acggcgggtc gaacaccagc
 8281 tggatgtcgg ttctggccac cgtcgggatc agctccagct tgctgcgcgc gtcctcgatc
 8341 aggatgtcgc ccatgccgca gccgggcgcg gtgaggggtca tccgcacctt gaccaggcgc
 8401 tggccgtcgt cccggtgcag gaccgcggcc tcgtagacca ggcccaggtc gaccacgttg
 8461 atcgggatct cggggtcgta gcaggtcgc agctggcgcc agaccagctc ctccacggcg
 8521 tcgtcgggtg cgtcttcag tgccagcggc gcgggtgggt ccttgccgat gggtcggcg
 8581 tcctcgccgg cgatccgga caggttgccc tccacgaaca ccgtgtagct gccgccagc
 8641 gcctgggtga tatagccgac gctgcggcc gccaggggtca cggcgtcgcc ctgcggcacc
 8701 agcacggcgg cgcagtcgcg ttcaagcgc acgggttcgc tgctgcggga atacatgttc
 8761 atcgtggccg gggaaatggg gcggatggcg tcattctacc gccgcggga cgctatgctg
 8821 tggcgctttg cccctgcatg agccttgcca tgatcgacat caaggcgcgt ccatcgcgct
 8881 ggctggcgcc cgccctgctg gtgctgggca gcagcggtt cgccggcgcc tggatcctgc
 8941 tcgccctgat ccgcgaccgg caatgcagct ggatggcggg gctcgcggcg atcgacgcgg
 9001 cgctgctgct gcggctggcg cgcacgccgg cgggctggcc gcgctgctg gcggccgtgg
 9061 ccgccaccgc cgccaccatc gccctggcca actgggggat cgccggcgcg caggtcggcc
 9121 ggctgctggg cctgctgccc tgggaatcgg cgtgcgggt cgggcccgcg tacgcctgga
 9181 ccctggccgg gctcgccaac cgccggccg acctggcctg gctggcggcg gcgctgctgc
 9241 tggcggcgct tgccgctagt ggccgcgctc cagcgcttg agctcgctga ccagcgcgcc
 9301 ggccatctcc gcgccgtcgc cgtagagcat gcgggtgttg tcggcgtaga acagcgctt
 9361 ctcgatgccg gcaaagccgg tgcccttgcc gcgcttgatg accaccgtgt tcttcgactg
 9421 cgagacgtcc aggatcgga tgccgtagat cggcgaggcc ggatcggtcc tggccaccgg
 9481 gttgaccacg tcgttggcac cgatcaccag cgagacatcg gtattgggga actcgggggt
 9541 gatgtcgtcc atgtcggcga tcaggctgta gggcacgccg gcctcgcca acagacgtt
 9601 catgtgcccg ggcatgcgtc cggcgaccgg atggatcgcg aacttgacct tgacccgcg
 9661 ctctgcaat cgctgcgcca gctcccagat cttgtgctgc gcctgcgcca ccgccatgcc
 9721 gtagccgggc acgatacca cgcgctcgcc ataagccatc atcgcgcgga cgtcgccggc
 9781 ctcgatcggc ttctgcgccc ccgcatcgcc ctgcgcccgg ccgctgctgc cgccgaagtt
 9841 ggagaacagc acgcccgaga tcggccggtt catcgccctg gccatcaacc gcgtcagcag
 9901 catgccggcc gcgccgacca tcatgccggc gatgatcagc gcctcgctgc cgagcacgta
 9961 gccctcgaac gccaccgcca ggccggtcaa ggcttggtac agcgagatca ccaccggcat
 10021 gtcggcgccg ccgatcgga gcgtcatcag cacgccagc gccagcgcg cgacgaagaa
 10081 cgcgatcacc agcggcacgc cgagcgactg caccaccatc gcccagcg ccaccgccac
 10141 cagcgcgacc agcgcgttga accactgctg gccggggaag gtgtaacgct tgtccatgcg
 10201 cccgtcgagc ttggcccagg cgatcaccga acccgacagc gagaccgcac cgatcagtcg
 10261 gccgaccacg cccagcacca gcgtcagccg cggcgcgggc gcggtgcccc cgagcgagaa
 10321 gcgcagcagc tccaccgcgc cgatcgccgc ggctgaaccg ccgcccatgc cgttgaaacag
 10381 cgccaccatc tgccggcatgt cggatgatcg cacgcgcttg cccagaccc aggcgaccgc
 10441 ggtaccacgc gccagcgcga ccacgatcag cgtagggttg tgcagcccc gcagcaggaa
 10501 cgtcgcgaaa gtggccagca gcatgccac gccggccac tggatgcgc tgcgcgcggt
 10561 cttcgggctg gccatgcgct gcaggccgag caaaaacagg gtggcgcgga tgaagtagct

10621 ggcttcacc agcagcgga gccagctcat ggccgttccc cgccgtcggg cttgcggctg
 10681 gtcttgaaca tctccagcat ccgctcggtg accacgtagc cgccggcgac gttgccggcg
 10741 cccagcagta ccgcgacaaa gccgatcgcc ttctccagcg tgggtgctggc gtggcccgac
 10801 accaccatcg cgccgatcag gacgatgccg tggatgaagt tggagcccg catcagcggc
 10861 gtatgcagga tcaccggcac ccgcgagatg atcacgtggc cgccgatcgc cgccagcatg
 10921 aacacgtaca gcgccacgaa tccgtcgcgc atcgggtgcct ccccgcgcg tggtcgccc
 10981 atcataaccg cgccgggtggc ggccagggcg cgacgcgggc cgccgatccg gcggccaccg
 11041 cgggtcaacc cgggcggtcg ggccggtttt ccagtctgca gcacgccgct gcggccgat
 11101 acccggtcg agccgcccgc cctatgctgc tgcgcgtgaa ccccccgcc gccgaggacc
 11161 gccacccga tgcgcgcgt gcggccgccc cgccgcccga cccggcgccc acgtggaggc
 11221 ggttcctcgc gcagatcgcc ccgcgcgcgt tccgctcgc cgaactcggc ctgcgccagc
 11281 gcgaggacgc cctggacgcg gtccaggacg cgatgctgaa gatgctgcc taccacgcgc
 11341 gcccgccggc ggagtgccg ccgctgttct ggagcatcct gcgcagccgc atcgtcgacc
 11401 tgcagcgcg ccgcaacttc cgcctgcgt ggctgctgcc ggccggcggc gcggacagcg
 11461 gcgggatcga ctggatcgac ccgggccccg acccgcgcg cgccacgac ggccgcaag
 11521 gctacgcgcg gctggccacg gcgctggcg cactgccgcg gcgcagcgc gaggcgttca
 11581 cgctgcgggt gctggaggaa ctgcagtcg cgaccacgc gcgcgtcatg ggctgcagcg
 11641 agggttcggg gaagacgcac ctgtcgcgcg cgccgaccgc gctgcagcag caactggagg
 11701 actggcgatg aacgcaccgc ccccccgca gcccgggcg gatcccgacg cgcacggcat
 11761 cgaccgcgc gcacgcgcgc actaccggca ggcgctggcc gcgctgctgc cgcgcaccgc
 11821 cagcaactg cacaaccgcc tgcaggcggc gctcgcggc cagggccggc cgcagccgcc
 11881 acggctgccc gcctggctct gggcgcgccc cgccgcgctg gtgctggccc tggccttcgg
 11941 gctgcgggga ccggcgcgcc cgagccggc gccccgcgcc actgcaccgc tggccacgca
 12001 aaccccgccc gccaccgaca acgccccggc cagctcgcac caggatccc acttctacct
 12061 gtggctggca tcgcgcgat ccgtcacgct cgcctcggag taagcccatg cacaccccc
 12121 acccgacgtt ccgcctggcc cgcctgggc tgcgtgctggc gctggccgcc ggccacgcc
 12181 tggccggcga ccggaccccg cctccgcagg ctgcggcgcc gccacccctg ccggcctggg
 12241 agcagctgtc cccggcgag cgcgagtgat tgatcgcgcc gctgcgcgag cgctggaacg
 12301 ccaaccccg cgagcgtcg cggaacctac cgcacgcgca gggttggcag tcgatgacct
 12361 cggaacagcg caccgcgcgc gccacggcc tgcaccgctg ggagcgcatg gatccggagc
 12421 gccgcacgca gatgcgcgc ctgttcgggc agatgcgcag catgaccccc gaacagcgca
 12481 aggccttcgc cgagcgtgg cagcgatga cggccgagca gcgcgcgcgc tgggtgcagg
 12541 cccacccgcc cgccgcggac gaggtagcgc ggcagcttgc cagccgcgca cccccccgt
 12601 catccccggc tccatcgtag gaggcagctg agtcgcgatt tcgcaccaag cgggcatcgc
 12661 gactcacgtc gctcctacat cgggtggccc gccatccggg cagctctgt catgccggc
 12721 ccgccacctg gacacgcctc gtcacgccc tccccggctc accgtaggag cgacgtaagt
 12781 cgcgatttcg caccggcggc gcacgcgcac ttgcgtcgt cctgcaccgg tattgcgagc
 12841 gcggctcagg tcgcgggaagc gaggggcgcg gcggcgggg tcgcggctgc gggcggtgc
 12901 gccagaccg tcttcgccag cagttcgtcg tcccagtcga acgccagtgc gccgtccttg
 12961 agcagcagac tggcgaagt gaacacgtt cgcgctgaca tctcgtggc gtgcagcgcg
 13021 ccgctgctgg ccaggtccag cgcccgggc atgacgacgc cgccgcgctc gatggtctcg
 13081 ccgggcccgg tcagctcgca gttgccgcg gtttccggc ccaggtcgac gatcacgctg
 13141 cccggctgca tgcctccac catcgccgcg gacacgatct tcggcgccg ccgccccggc
 13201 acccgggcg tgcagacgac gacatcgatg gtcttcagg gctcggccag gcggcgctgc
 13261 tgctcggcgc gctcctcgtc ggtgagcgca cgccatagc caccctcgcc tgcggcgctg
 13321 acgcccaggc cgaggaaact gccgcccagc gactggatct gctcgcgcgt ttccgggcgc
 13381 acgtcgaaac cctcgacct cgctccaagc cgcttggcg tggcgatcgc ctgcaggccg
 13441 gcgaccccg cgccgaccac cagcaccttc gacggcgga tgggtccggc ggcggtggtc
 13501 agcatcgga agaagcgcg cgccagctgc gcggcgatca gcacgcctt gtagccggcc
 13561 atgcccggct gcgaactgag cacgtccatc gcctgcgcac ggggtggtgc gcgcaaccgc
 13621 tccagcgga aggcgacgat cgccagcgcc tgcagcgcg ccgcgcgcgc cggatcgccc
 13681 tgcggctgca gggtgccgac cagcaccgc ccgtgctgca ggcgctcgat cgccggcgcc
 13741 ggccggcgct gcacgcacag caccaggtcg gcgctgcca gcgcgcgctc ggctcgggc

13801 accagctgcg cgccggcgcc ggcgtaggcc gggtcgggga agtgcgcgcc gcgaccgagg
 13861 ccgcgctcga cccgcacgct ggcgcggcg gcgacgaatt tcttgcaagg ttccggggtc
 13921 agcgccacgc ggcgctcgcc ggcgcgggt tcccgcgcca ctccaatgat caccgccatg
 13981 cccgactccg tcgcatcgat gcgcgcatcc tagcgaatgc gcgcaggccg cgcgcatatc
 14041 agggcgcgcg cgtggcgggc tggcgagcg tctgccgcac ctggcgcatc gcctggtcga
 14101 acggcaactc gtcggcgctcc accgacaggc gcccgctctt gaccagaccg gccagctgct
 14161 ccggcgaggc caccagcttg cgcacgccgc ggcgggtgac cagcagcagg cgcgagggtca
 14221 acgggctgat ccaggcgacc ttgaccgagc cctcgacgcc gtcctcgtgc agccgcagcc
 14281 attcgcccgg gcgcagcgcg cgcacgcgcg cggcgacttc cgggtcgaag tcgagcgcg
 14341 cagtgcgcc caccaccgc aacagccggg tatcgtcgct gtcacggcc agttcgggca
 14401 tctccgggtg cggcgctgc tcgcgcgcgg tgcgggggaa cgccagcgtc cgcgccagcc
 14461 cggccatcca ttcgctggcg acctggtcgt ccaggccgga actggacagg cattccacca
 14521 ggcccgcata cagcgccagc acgcgctcgg cgacgttgct gccggcgccg cgggcagcg
 14581 cctcgtcgac cgcgatcaac tcgtcgccca gtgccatcac ctggcgctgg cgctcgagg
 14641 cctggccttc gcgcaacagc acctgcacca ggtggtgctg ccagtggcg tccaggaaca
 14701 agccgatggt cggcgctcagc gcgcgccggg acaggcgctg cgccagcacc ccggcgccct
 14761 gcagccgcgc ctgcagcagg cgctcgcgcc cgtgcacggc ttccgcccag cgccgctcca
 14821 ccacctcggc gcggcgcgcg tgctgctcca gcaggtcgag cagctcgttg gcggtgagtt
 14881 cgaacaccgc caggtcctcg ttgtattcgc agaccacgcg ctcgaccagc gcgctggcg
 14941 gctgcagcag ctcgcggtcc tcggcgagg cgccgtcgtt gctttcgag ctcagcgcca
 15001 gcgcgctcag caggcgctgc gccgggtgat ccgggcgcac gaacagggtc tcgtcggtca
 15061 gcgcgacctt gacgtaggac atcaccagcc gcgcgaacag ccggcgcgac tcgtgcacca
 15121 gcgcgtggct cttgagcagc gactcgaaca gcaggccggc caggtcgatc gcgtcctcct
 15181 gcgcgcgctc caggcagggt tggctcgggt ccagcccgag ccggcgcgcg ccttcagca
 15241 gtcctcgcg gatcgctcg tcgagccgc acctgcccgt cagtgtgctc tcgaactggc
 15301 gcgcaggtgc gcgctgcagc agcgaggcca cgctgaccag ttcctgcacc cgcaactcgc
 15361 ggcgtccggg caggcgcgcg tccggggcgt gtcgctcggg tacggtctgg tcgcccagca
 15421 tgccctcgcg cactggtgc agcaggtcgc gcaggcgctg gtgctgcacc cgcgcccg
 15481 cggagggtcg gaacaggtcg gtgcccgcgc tggccgccgc gtcgcgccg gccgcccgt
 15541 ccggcgcgcg ggcgcgccacc ccgggcacga agccggtcgc gttcgcgcc gccaccggcg
 15601 ccggctccgc aggcgtggcc gccgggtcgc ggcgcaggtc ggcagggaaa cccttcgccc
 15661 ccagttcggg gttgaggcg ccgtagagct cgccagcacc cttggccagt tctgttcgt
 15721 actggcgga cagcagcgcg tcgagggtct cggagatgct ggcgtcgcg aaggctcgca
 15781 ggaacgcccc ggccaggcg gccgcgcga ccggattggc ggccggccgg gtccccagg
 15841 cgttggccag cgactcgaag cgcaggcca gcatttccag cggcgctgg tagcgcgcc
 15901 cgatcgatc gccagcgcg tggccggcga ggtggaagtc gagctcgttc tcgcgacca
 15961 gcccattgg catcgccggc ttgcccagga aaccgcgacc gcggaaatcg tcgaaattgc
 16021 gcgcgaccag ttcgcggaac cgcacacgt aagcgcgct gcgctggcg agggccagca
 16081 ccgtcatcag gtccttcgcg tggccgcga cgtgcccac gcgcagactg tcgtgcagtt
 16141 ggtcctcgat cgctgtgac aggcgcggc gcaccgcgc gagttgctcc acggcgatac
 16201 gcttgagttc atcgagcagg cgcgcggat cggatgccga cacgctggcc agaccggtag
 16261 gatgagtgga caccgctaatt tcccttgcg tcgacgttca acggcaaaaa atgacgcaaa
 16321 cgtgaacaaa ttgtcgtaac gatcctcgcc cacatgcagt aatcgaatcc gtgatgacca
 16381 ttccagtacc ggccgaacca tgagcgtagc tctcgactgc ctgaatacgc gccgctcgg
 16441 ccccgccaag caactcgcg ccccggtcc cgacgacgct accctgcaac gcatgctgcg
 16501 cgccgcagtg cgcgtcccg accatggtaa acgcatgccg tggcgattcc tgcgattgc
 16561 gggtagcgcc cgccaggccc ttggcgagc gttggtcgcc cgctcactgg aacgcgatcc
 16621 ggccatcgcc ccggtcgatc tggacaagga ccgccaccg ttcgcgacg cgccgctgg
 16681 gatcgcggtg atcgcccgcc tcgggcccga cccgaagatc cccgaatcc aacgcttctc
 16741 ctgggctcgt tcgctgtgct tcgcgctgtt cgaggcgcg caggcctgag gcttcggcg
 16801 gcaatggctg accggctggc cggcctacga cccggcgatc ctgcgctgc tcggcctgga
 16861 cgccgacgaa cgcacgcgc gttcatcca catcggcag ccgggatgg acatccccga
 16921 gcgcacccgc cccgaccgc tgcaactgct cagcgactgg accccgcgc cgcgccggga

16981 tgcggggcgg cccgcgtgag cgacggtcgc gtgcgcccga cgctgtacct ggtcgacgcc
 17041 agcctgtacg tgttccgcgc ctggcactcg atgcccgcag agttccacga cgccgacggc
 17101 tggccgacca atgccgtgca cggcttcgcc cgcttcctgc tggagctgct ggagcgcgaa
 17161 cgccccaggc atatcgccgt cgccttcgac gaggcgctgg acagctgctt ccgcaacgcg
 17221 ctgtacccgg cctacaaggc caaccgcgag cgggcgcccg aagcgctcaa gcgccagttc
 17281 gggcagtgca aggcgctgtg cggcgcgctc gggctggtga cgctggcgca ccgcgaatac
 17341 gaggcgcagc acctgatcgg cagcgccctg catcacgcgc gcacgcacgg ccatcgccggc
 17401 gtgatcgtgt cggcggacaa ggacctgtcg caactgctcg gcctgcacga cgagcaatgg
 17461 gacttcgccc gcggccagcg ctggggcgcg gccggggtca aggcgcggca cggggtcgat
 17521 gcgcaccaga tcgccgacta tctcgccctg tgcggcgatg cggtcgacaa catccccggc
 17581 gtgcccggca tcggcgccaa gaccgctgcg gtgctgctgg cgcacttcgg cagcctcgac
 17641 gcgctgctcg cgcgcacga cgagatcccg ttctcgctt tccgcggcgc cgcggccgcg
 17701 gcggccaagc tgcgcgcgca ccgcgagcag gcgctgctgt gccggcagct gagcaccatc
 17761 gcgctggacg cgcgctgga ccacgcccac gacttcgccc gcgccgccgc cgatcgacac
 17821 ggctcgcgga cgctgtgcga cgcactcagc ttcgggcccgc tgaccgcggc ccgcctgcac
 17881 gccgccggcg gcctggacta cgccacggcg cagtagcgcg acggccagcc gttcctttct
 17941 ctcgggcccgc gggcgtccc ttctccgccg cgggaccgcg ttctcttctc tccgccacg
 18001 gaccgcttct cttctcccc cgcgcggcgc ttctcttctc cccgcccgca cgccgcttct
 18061 cttctcccca cccgcgagcc gcgttcctcc tccccgccg cgagccgctt tccttctccc
 18121 cgcccgcggg gagaaggcgc ccgaaggcg gatgaggggc cgtccccgcg tacgccagcc
 18181 acggccgcgc ggtagcatcg gcgcattggc caccgcgacg aactcaagct cccgtccgaa
 18241 accctctacg aaggcaaatg gctgcgcctg cggcgcgcg gccgctggga atacgccgag
 18301 cgcaacctat ccgacggcat ggcggtgatc atcatcgccg tcaccccgga cgaccgctg
 18361 ctgttcgtcg agcagaccgc catcccgtg ggcggcaaga ccatcgagat gccggccggc
 18421 ctggtcggcg accagcccgc caggacacg ctggaggacg ccgcgcggcg cgagctggtc
 18481 gaggaaccgc gctggcacc cgcgggggtc gaagtgtgct tggctggccc gacctcctc
 18541 gggatgagca acgagcgcat cgcctttgtc cgcggccgcg gcctgaccgc cgtcgccccc
 18601 ggcggcgggc acgccaccga ggacatcgct gtgcacgcaa tcccgcgcg cgacgcggcg
 18661 gcctggctga tgcaaaagcg cagcgaaggc tacgagctcg acctcaagct gtgggcccgc
 18721 ctgtggatgc tggagcgcaa cccggacggc acgccgatcg aatgacggcc ggatgatggc
 18781 ccggccggaa ggcagccgtg cgtatcgctt gaaacgcggc gttatttctt gtcgcccggc
 18841 ttctgaccgc ttgcacacc ccaccgacgc cgcgattccg cgcataaatc attgaataag
 18901 cgcaattgac agggcaatcg agcgagctga atatcgggac cgttatggct ccgagctgga
 18961 gcaaatcagg tgttagccct taatgaaagc acggctcgaa gaactggcgg acaagagcat
 19021 cgccgcgatg gtgtcgcgca tcgaggtgta caacaagccg gactttaagt accgggagga
 19081 aaccttctcc attcttgcca tcaacgcttg ggaactgctc ttaaaggcaa aatggctcaa
 19141 agacaacaac aacaaagttc gctcgctata cgtaactgag aagcgaatgc ggccaaatgg
 19201 caagccctac aagcacgcca aggtaaagat gacgggatca ggcaaccctt ttacgcacag
 19261 ccttgactac ttggccaaga gaatggccga gaagaagacg ttggctgatg gggcgcataa
 19321 gaacattatt gcgctgtgcg agatcagaga ctcgctcagt catttctaca acaagagtgc
 19381 tctgttcgca gtcgggctac aggaggtcgg aagtgcattc gtcaggaaat atgcaaaagc
 19441 cgcgcagtca tggttcagca ttgacttcgc tcagtacaac ttctacctaa tgccacttgc
 19501 attcgtttaag aatcagcaac ccgctgaggg cttgctcctg agcaaggagc agaagaatgt
 19561 tgcgtacttc atttctctcc ttgaggtgct aaacgaccgc aaggcggact atgcggtgtc
 19621 ggtgaacgta gaactcaagt ttctcaagtc aaaagctgat gatgcaatca aggttcaggt
 19681 cagcaatgac cctggcgcac caaaagtcca gctcacagaa gagcacttga aagataagta
 19741 tcctctcaac tatgcggcac tgaacaaggc gtgtgctgct cgttactgcg atttctctca
 19801 gaacaaaaag taccatgatc tgcgcaaac cctgaaagcc gataagcgct attgtcatgt
 19861 caagaaactc gacccggaca acccaaagag tgccaaacaa gagtggatca gccaaagctgt
 19921 gttcaatgtt ctggataagt actacacggt taaggcgctaa caattcgctt aagccgatgc
 19981 cgcttcgcgg caccgcttaa ctcaggcggt agcttgctat ggaagatc

7.3 Angaben zu den Genomabschnitten der 5 *Xanthomonas*-Vertreter

7.3.1 *X. campestris* pv. *vesicatoria* str. 85-10

Die Genomdaten von *X. campestris* pv. *vesicatoria* str. 85-10 sind bei der NCBI-Datenbank unter der Accession-Nr. AM039952 veröffentlicht. Der in Abbildung 35 (C) dargestellte Genomausschnitt repräsentiert den Genombereich von 1118626 bis 1074579. Die Annotationen der Gene können der nachfolgend aufgeführten Tabelle entnommen werden.

Gen-Position in Abbildung 35 (C)	Gen-Synonym bei NCBI	Annotation
1	XCV0988	50S ribosomales Protein L1
2	XCV0987	50S ribosomales Protein L11
3	XCV0986	Transkriptions-Antiterminationsfaktor
4	XCV0985	Präprotein Translokase UE
5	XCV0984	Elongationsfaktor Tu
6	XCV0983	GTP-Bindeprotein
7	XCV0982	Peptidyl-tRNA Hydrolase
8	XCV0981	50S ribosomales Protein L25
9	XCV0980	Ribose-Phosphat-Pyrophosphokinase
10	XCV0979	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase
11	XCV0978	Lipoprotein Precursor der äußeren Membran
12	XCV0977	Konserviertes hypothetisches Protein
13	XCV0976	Glutamyl-tRNA-Reduktase
14	XCV0975	Peptidketten-Release-Faktor
15	XCV0974	Hypothetisches Protein
16	XCV0973	Molypdopterin Biosynthese Protein
17	XCV0972	Konserviertes hypothetisches Protein
18	XCV0970	Acylphosphatase
19	XCV0969	Mögliches Thioredoxin
20	XCV0968	Ribonuklease BN
21	XCV0967	Tryptophan Repressor Bindeprotein
22	XCV0966	Konserviertes hypothetisches Protein
23	XCV0965	Xylanase Vorläuferprotein
24	XCV0964	Konserviertes hypothetisches Protein
25	XCV0963	Asparaginase
26	XCV0962	Konserviertes hypothetisches Protein
27	XCV0961	Extrazelluläre Protease
28	XCV0960	Extrazelluläre Protease

29	XCV0959	Extrazelluläre Protease
30	XCV0958	„branched-chain“ Aminosäure-Aminotransferase
31	XCV0957	Konserviertes hypothetisches Protein
32	XCV0956	Konserviertes hypothetisches Protein
33	XCV0955	Mögliche NAD(P)-Transhydrogenase β UE
34	XCV0954	RNA-Polymerase ECF-Sigma Faktor
35	XCV0953	Hypothetisches Protein
36	XCV0952	Konserviertes hypothetisches Protein
37	XCV0951	Pyridin-Nukleotid-Transhydrogenase
38	XCV0950	Konserviertes hypothetisches Protein
39	XCV0949	Konserviertes hypothetisches Protein
40	XCV0948	Exodeoxyribonuklease
41	XCV0947	MutT-ähnliches Protein
42	XCV0946	Konserviertes hypothetisches Protein

7.3.2 *X. campestris* pv. *campestris* str. ATCC 33913

Die Genomdaten von *X. campestris* pv. *campestris* str. ATCC 33913 sind bei der NCBI-Datenbank unter der Accession-Nr. AE008922 veröffentlicht. Der in Abbildung 35 (D) dargestellte Genomausschnitt repräsentiert den Genombereich von 1040326 bis 995188. Die Annotationen der Gene können der nachfolgend aufgeführten Tabelle entnommen werden.

Gen-Position in Abbildung 35 (D)	Gen-Synonym bei NCBI	Annotation
1	XCC0885	50S ribosomales Protein L1
2	XCC0884	50S ribosomales Protein L11
3	XCC0883	Transkriptions-Antiterminationsfaktor
4	XCC0882	Präprotein Translokase UE
5	XCC0880	Elongationsfaktor Tu
6	XCC0876	GTP-Bindeprotein
7	XCC0875	Peptidyl-tRNA Hydrolase
8	XCC0874	50S ribosomales Protein L25
9	XCC0873	Ribose-Phosphat-Pyrophosphokinase
10	XCC0871	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase
11	XCC0870	Lipoprotein Precursor der äußeren Membran
12	XCC0869	Konserviertes hypothetisches Protein
13	XCC0868	Glutamyl-tRNA-Reduktase
14	XCC0867	Peptidketten-Release-Faktor
15	XCC0866	Hypothetisches Protein
16	XCC0865	Molypdopterin Biosynthese Protein

17	XCC0864	Transkriptionsregulator
18	XCC0863	Konserviertes hypothetisches Protein
19	XCC0862	Acylphosphatase
20	XCC0861	Mögliches Thioredoxin
21	XCC0860	Ribonuklease BN
22	XCC0859	Tryptophan Repressor Bindeprotein
23	XCC0858	Konserviertes hypothetisches Protein
24	XCC0857	Xylanase
25	XCC0856	Asparaginase
26	XCC0855	Transkriptionsregulator
27	XCC0854	Extrazelluläre Protease
28	XCC0852	Extrazelluläre Protease
29	XCC0851	Extrazelluläre Protease
30	XCC0850	„branched-chain“ Aminosäure-Aminotransferase
31	XCC0849	Konserviertes hypothetisches Protein
32	XCC0848	Konserviertes hypothetisches Protein
33	XCC0847	RNA-Polymerase ECF-Sigma Faktor
34	XCC0846	Konserviertes hypothetisches Protein
35	XCC0845	Konserviertes hypothetisches Protein
36	XCC0844	Transkriptionsregulator
37	XCC0843	Hydrolase
38	XCC0842	Konserviertes hypothetisches Protein
39	XCC0841	Konserviertes hypothetisches Protein
40	XCC0840	Konserviertes hypothetisches Protein
41	XCC0839	Exodeoxyribonuklease
42	XCC0838	MutT-ähnliches Protein
43	XCC0837	Konserviertes hypothetisches Protein

7.3.3 *X. campestris* pv. *campestris* str. 8004

Die Genomdaten von *X. campestris* pv. *campestris* str. 8004 sind bei der NCBI-Datenbank unter der Accession-Nr. CP000050 veröffentlicht. Der in Abbildung 35 (E) dargestellte Genomausschnitt repräsentiert den Genombereich von 4000216 bis 4045359. Die Annotationen der Gene können der nachfolgend aufgeführten Tabelle entnommen werden.

Gen-Position in Abbildung 35 (E)	Gen-Synonym bei NCBI	Annotation
1	XC_3350	50S ribosomales Protein L1
2	XC_3351	50S ribosomales Protein L11
3	XC_3352	Transkriptions-Antiterminationsfaktor
4	XC_3353	Präprotein Translokase UE
5	XC_3354	Elongationsfaktor Tu
6	XC_3355	GTP-Bindeprotein
7	XC_3356	Peptidyl-tRNA Hydrolase
8	XC_3357	50S ribosomales Protein L25
9	XC_3358	Ribose-Phosphat-Pyrophosphokinase
10	XC_3359	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase
11	XC_3360	Lipoprotein Precursor der äußeren Membran
12	XC_3361	Konserviertes hypothetisches Protein
13	XC_3362	Glutamyl-tRNA-Reduktase
14	XC_3363	Peptidketten-Release-Faktor
15	XC_3364	Hypothetisches Protein
16	XC_3365	Molypdopterin Biosynthese Protein
17	XC_3366	Transkriptionsregulator
18	XC_3367	Konserviertes hypothetisches Protein
19	XC_3368	Acylphosphatase
20	XC_3369	Mögliches Thioredoxin
21	XC_3370	Ribonuklease BN
22	XC_3371	Tryptophan Repressor Bindeprotein
23	XC_3372	Konserviertes hypothetisches Protein
24	XC_3373	Xylanase
25	XC_3374	Asparaginase
26	XC_3375	Transkriptionsregulator
27	XC_3376	Extrazelluläre Protease
28	XC_3377	Extrazelluläre Protease
29	XC_3378	Extrazelluläre Protease
30	XC_3379	„branched-chain“ Aminosäure-Aminotransferase
31	XC_3380	Konserviertes hypothetisches Protein
32	XC_3381	Konserviertes hypothetisches Protein
33	XC_3382	RNA-Polymerase ECF-Sigma Faktor
34	XC_3383	Konserviertes hypothetisches Protein
35	XC_3384	Konserviertes hypothetisches Protein
36	XC_3385	Transkriptionsregulator
37	XC_3386	Hydrolase
38	XC_3387	Konserviertes hypothetisches Protein
39	XC_3388	Konserviertes hypothetisches Protein
40	XC_3389	Konserviertes hypothetisches Protein
41	XC_3390	Exodeoxyribonuklease
42	XC_3391	MutT-ähnliches Protein
43	XC_3392	Konserviertes hypothetisches Protein

7.3.4 *X. axanopodis* pv. *citri* str. 306

Die Genomdaten von *X. axanopodis* pv. *citri* str. 306 sind bei der NCBI-Datenbank unter der Accession-Nr. AE008923 veröffentlicht. Der in Abbildung 35 (F) dargestellte Genomausschnitt repräsentiert den Genombereich von 1119791 bis 1071729. Die Annotationen der Gene können der nachfolgend aufgeführten Tabelle entnommen werden.

Gen-Position in Abbildung 35 (F)	Gen-Synonym bei NCBI	Annotation
1	XAC0962	50S ribosomales Protein L1
2	XAC0961	50S ribosomales Protein L11
3	XAC0960	Transkriptions-Antiterminationsfaktor
4	XAC0959	Präprotein Translokase UE
5	XAC0957	Elongationsfaktor Tu
6	XAC0953	GTP-Bindeprotein
7	XAC0952	Peptidyl-tRNA Hydrolase
8	XAC0951	50S ribosomales Protein L25
9	XAC0950	Ribose-Phosphat-Pyrophosphokinase
10	XAC0948	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase
11	XAC0947	Lipoprotein Precursor der äußeren Membran
12	XAC0946	Konserviertes hypothetisches Protein
13	XAC0945	Glutamyl-tRNA-Reduktase
14	XAC0944	Peptidketten-Release-Faktor
15	XAC0943	Hypothetisches Protein
16	XAC0942	Molypdopterin Biosynthese Protein
17	XAC0941	Transkriptionsregulator
18	XAC0940	Konserviertes hypothetisches Protein
19	XAC0939	Acylphosphatase
20	XAC0938	Mögliches Thioredoxin
21	XAC0937	Ribonuklease BN
22	XAC0936	Tryptophan Repressor Bindeprotein
23	XAC0935	Konserviertes hypothetisches Protein
24	XAC0934	verkürzte Xylanase
25	XAC0933	verkürzte Xylanase
26	XAC0932	Asparaginase
27	XAC0931	Transkriptionsregulator
28	XAC0930	Extrazelluläre Protease
29	XAC0929	Extrazelluläre Protease
30	XAC0928	Extrazelluläre Protease
31	XAC0927	„branched-chain“ Aminosäure-Aminotransferase
32	XAC0926	Konserviertes hypothetisches Protein
33	XAC0925	Konserviertes hypothetisches Protein
34	XAC0924	Mögliche NAD(P)-Transhydrogenase β UE

35	XAC0923	Mögliche NAD(P)-Transhydrogenase α UE
36	XAC0922	RNA-Polymerase ECF-Sigma Faktor
37	XAC0921	Konserviertes hypothetisches Protein
38	XAC0920	Konserviertes hypothetisches Protein
39	XAC0919	Pyridin-Nukleotid-Transhydrogenase
40	XAC0918	Pyridin-Nukleotid-Transhydrogenase
41	XAC0917	Transkriptionsregulator
42	XAC0916	Hydrolase
43	XAC0915	Konserviertes hypothetisches Protein
44	XAC0914	Konserviertes hypothetisches Protein
45	XAC0913	Konserviertes hypothetisches Protein
46	XAC0912	Exodeoxyribonuklease
47	XAC0911	MutT-ähnliches Protein
48	XAC0910	Konserviertes hypothetisches Protein

7.3.5 *X. oryzae* pv. *oryzae* KACC10331

Die Genomdaten von *X. oryzae* pv. *oryzae* KACC10331 sind bei der NCBI-Datenbank unter der Accession-Nr. AE013598 veröffentlicht. Der in Abbildung 35 (G) dargestellt Genomausschnitt repräsentiert den Genombereich von 3859493 bis 3902413. Die Annotationen der Gene können der nachfolgend aufgeführten Tabelle entnommen werden.

Gen-Position in Abbildung 35 (G)	Gen-Synonym bei NCBI	Annotation
1	XOO3594	50S ribosomales Protein L1
2	XOO3596	50S ribosomales Protein L11
3	XOO3597	Transkriptions-Antiterminationsfaktor
4	XOO3598	Präprotein Translokase UE
5	XOO3599	Elongationsfaktor Tu
6	XOO3600	GTP-Bindeprotein
7	XOO3601	Peptidyl-tRNA Hydrolase
8	XOO3602	50S ribosomales Protein L25
9	XOO3603	Ribose-Phosphat-Pyrophosphokinase
10	XOO3604	4-Diphosphocytidyl-2-C-Methyl-D-Erythritol-Kinase
11	XOO3605	Lipoprotein Precursor der äußeren Membran
12	XOO3606	Konserviertes hypothetisches Protein
13	XOO3607	Glutamyl-tRNA-Reduktase
14	XOO3608	Peptidketten-Release-Faktor
15	XOO3609	Hypothetisches Protein

16	XOO3610	Molypdopterin Biosynthese Protein
17	XOO3611	Transkriptionsregulator
18	XOO3612	Konserviertes hypothetisches Protein
19	XOO3613	Acylphosphatase
20	XOO3614	Mögliches Thioredoxin
21	XOO3615	Ribonuklease BN
22	XOO3616	Tryptophan Repressor Bindeprotein
23	XOO3618	Xylanase
24	XOO3619	Asparaginase
25	XOO3620	Transkriptionsregulator
26	XOO3621	Extrazelluläre Protease
27	XOO3624	„branched-chain“ Aminosäure- Aminotransferase
28	XOO3625	Konserviertes hypothetisches Protein
29	XOO3626	Konserviertes hypothetisches Protein
30	XOO3627	Mögliche Transposase
31	XOO3628	Mögliche Transposase
32	XOO3629	Mögliche NAD(P)-Transhydrogenase β UE
33	XOO3630	Mögliche NAD(P)-Transhydrogenase α UE
34	XOO3631	RNA-Polymerase ECF-Sigma Faktor
35	XOO3632	Konserviertes hypothetisches Protein
36	XOO3633	Pyridin-Nukleotid-Transhydrogenase
37	XOO3634	Mögliche Transposase
38	XOO3635	Mögliche Transposase
39	XOO3636	Konserviertes hypothetisches Protein
40	XOO3637	Konserviertes hypothetisches Protein
41	XOO3638	Exodeoxyribonuklease
42	XOO3639	MutT-ähnliches Protein
43	XOO 3640	Konserviertes hypothetisches Protein

Danksagung

Mein besonderer Dank gilt Herrn PD Dr. Rolf Daniel, der mir die Arbeit an diesem überaus interessanten Thema ermöglichte. Vielen Dank für die Betreuung und die stetige Diskussionsbereitschaft, die zum Gelingen der Arbeit beigetragen haben.

Herrn Prof. W. Liebl danke ich für die Übernahme des Korreferats.

Ein besonderer Dank für die mir entgegengebrachte Hilfsbereitschaft, die nette Arbeitsatmosphäre und den enormen Spaß bei diversen Tagungen oder bei unserer Exkursion zur Zugspitze geht an alle jetzigen und ehemaligen Labor-Mitglieder Anja, Baris, Caro, Carola, Christiane, Claudia, Frank, Hanna, Jessi, Jörg, Sebastian, Sönke, Steffi, Susanne und Tini.

Noch ein zusätzlicher Dank an Carola, Christiane und Jessi für ihr Engagement bei der gewissenhaften Durchsicht des Manuskripts.

Eine angenehme Atmosphäre fand ich auch immer in unserer neuen Abteilung vor. Dort bedanke ich mich vor allem bei Angel, Jörn und Maike für eure Geduld mit mir an der FPLC. Mechthild danke ich ganz besonders für die vielen Handballgespräche, auch wenn wir uns manchmal auf dem Spielfeld als Gegner gegenüberstanden.

Ein Dankeschön geht auch an das G2L für die unendlichen Sequenzierungen meiner Plasmide, ohne die diese Arbeit nicht möglich gewesen wäre.

Ein ganz großes Lob an unser Werkstatt-Team Gerd Birke, Patrick Regin und Olaf Waase. Ich danke euch für die Unterstützung und die schnelle Hilfsbereitschaft, wenn es mal wieder darum ging etwas möglichst schnell zu reparieren.

Mein ganz besonderer Dank geht an meinen Lebensgefährten und an meine Eltern für die liebevolle Unterstützung und das aufgebrachte Verständnis schon während meiner gesamten Studienzeit. Gut das es euch gibt !!!

Und natürlich darf einer nicht vergessen werden. Mein treuester Freund, mein Hund Dario, der mich jetzt schon seit elf Jahren begleitet und immer wieder für reichlich Abwechslung und manchmal auch Aufregung in meinem Alltag gesorgt hat. Dario du bist wirklich einmalig.

Lebenslauf

14.07.1969	Geburt in Kassel
1976 - 1980	Besuch der Grundschule in Uschlag
1980 - 1984	Besuch der Gesamtschule in Niestetal/Heiligenrode
1984 - 1987	Besuch der Kooperativen Gesamtschule in Göttingen
1987- 1990	Besuch des Hainberg-Gymnasium in Göttingen
05/1990	Abitur
07/1990 – 06/1992	Ausbildung zur „Staatlich geprüften Biologisch-technischen Assistentin“ an der Chemieschule in Hann. Münden
09/1992	Immatrikulation an der Georg-August-Universität zu Göttingen im Studienfach Biologie (Diplom)
10/1994	Vordiplom in den Fächern Mikrobiologie, Botanik, Chemie und Physikalische Chemie
10/1999 – 09/2000	Anfertigung der experimentellen Diplomarbeit unter Anleitung von Prof. D.G. Robinson am Albrecht-von-Haller-Institut für Pflanzenwissenschaften mit dem Titel: Untersuchungen zur Identifizierung der Fracht pflanzlicher endozytotischer Clathrin-coated Vesikel
02/2001	Diplomprüfung in den Fächern Mikrobiologie, Botanik und Chemie
04/2001	Beschäftigung als wissenschaftl. Hilfskraft am Institut für Mikrobiologie und Genetik
07/2001	Beschäftigung als wissenschaftl. Mitarbeiterin am Institut für Mikrobiologie und Genetik
04/2002	Beginn der experimentellen Arbeiten zur vorliegenden Dissertation