
**Sequenzanalyse, Transkriptionsaktivität und Untersuchungen
zur Evolution des Plasmids pKAP298 aus dem intrazellulären
Parasiten *Caedibacter taeniospiralis***

Dem Fachbereich Biologie der Universität Kaiserslautern zur Erlangung des akademischen
Grades "Doktor der Naturwissenschaften" eingereichte

Dissertation

vorgelegt von
Jörn Jeblick

Vorsitzender: Prof. Dr. H.J. Schmidt
Erster Berichterstatter: PD Dr. J. Kusch
Zweiter Berichterstatter: Prof. Dr. J. Cullum

Lehrbereich Ökologie der Technischen Universität Kaiserslautern, November 2003

D 386

"Life is but a walking shadow, a poor player that struts and frets his hour upon the stage and then is heard no more. It is a tale told by an idiot, full of sound and fury signifying nothing"

William Shakespeare, Macbeth

... und für Optimisten:

"...die Chancen stehen eins zu einer Million - aber es könnte klappen..."

Terry Pratchett

gewidmet meinen Eltern, Großeltern und meiner wunderbaren Lebensgefährtin

1. EINLEITUNG.....	6
1.1 CAEDIBACTER-ARTEN UND IHRE VERWANDTSCHAFT ZUEINANDER	6
1.2 ART DER CAEDIBACTER-PARAMECIUM-INTERAKTIONEN	8
1.3 TÖTUNGSMECHANISMEN DER CAEDIBACTER-INFIZIERTEN PARAMECIEN	8
1.4 DIE pKAP-PLASMID-FAMILIE & GENETISCHE FAKTOREN DER R-KÖRPER SYNTHESE.....	10
1.4.1 Das Plasmid pKAP aus Caedibacter taeniospiralis.....	10
1.4.2 R-Körper-Gene	12
1.5 HYPOTHESEN ZUR NATUR DES TOXINS	12
1.6 ZIELE DER ARBEIT	13
2. MATERIAL UND METHODEN.....	15
2.1 MATERIALIEN.....	15
2.1.1 Geräte.....	15
2.1.2 Kunststoffmaterial	16
2.1.3 Glasmaterial	16
2.1.4 Optische Geräte.....	16
2.1.5 Fotomaterial	16
2.1.6 Chemikalien	17
2.1.7 Elektrophorese.....	17
2.1.7.1 Verwendete Agarose.....	17
2.1.7.2 DNA Molekulargewichtsmarker	18
2.1.7.3 Loading Dye	18
2.1.8 Enzyme.....	18
2.1.9 Transformationsvektoren.....	18
2.1.10 Wirtsorganismus für die Vektoren.....	19
2.1.11 Verwendete "Kits"	19
2.1.12 Oligonucleotide	20
2.1.13 Stammlösungen, Puffer und Kulturmedien	20
2.1.14 Biologisches Material.....	23
2.1.14.1 Paramecium Stamm.....	23
2.1.14.2 Endobionten.....	24
2.1.14.3 Futterbakterien.....	24
2.2 METHODEN	25
2.2.1 Aufzucht und Aufarbeitung von Paramecium/Caedibacter - Massenkulturen	25
2.2.1.1 Nachweis und Identifizierung des Caedibacter-Endobionten	25
2.2.1.2 Bereiten des Futter-/Kultur-Mediums.....	26
2.2.1.3 Aufzucht der Massenkulturen	27
2.2.1.4 Abernten der Massenkultur.....	28
2.2.1.5 Aufschluss der abgeernteten Paramecien / Isolation des Endobionten	29
2.2.2 DNA-Isolation.....	31
2.2.2.1 Plasmid-DNA-Isolation	31
2.2.2.2 DNA-Extraktion aus Agarose-Gelen	33
2.2.3 Isolation von Gesamt-RNA aus C. taeniospiralis.....	34
2.2.4 Umschreiben der Gesamt-RNA in cDNA.....	35
2.2.5 Modifikation von DNA.....	36
2.2.5.1 Restriktion	36
2.2.5.2 Dephosphorylierung von Vektor-DNA.....	37
2.2.5.3 Ligation.....	37
2.2.6 Transformation von Escherichia coli.....	38
2.2.6.1 Herstellung elektrokompetenter Zellen.....	38
2.2.6.2 Transformation durch Elektroporation	38
2.2.6.3 Identifizierung und Konservierung von Transformanten.....	39
2.2.7 DNA/RNA-Konzentrationsbestimmung über die optische Dichte.....	39
2.2.8 Agarose-Gel-Elektrophorese	40
2.2.9 Hybridisierungen	40
2.2.9.1 Southern Blots	40
2.2.9.2 Kolonie-Blots zum Transformanten-Screening	41
2.2.9.3 Nicht-radioaktive DNA-Markierung mit Digoxigenin-11-dUTP	41
2.2.9.4 Hybridisierung der Membranen mit der Sonden-DNA.....	42

2.2.9.5	Detektion der hybridisierten Membranen	43
2.2.10	Analyse der Southern Blots anhand der pKAP-Sequenz, Aktivitätsanalyse.....	44
2.2.10.1	Zuordnung der Blot-Gel-Banden zu den entsprechenden pKAP298-Fragmenten.....	44
2.2.10.2	Zuordnung der Signal-Banden zu den entsprechenden pKAP298-Fragmenten.....	44
2.2.10.3	Erstellung einer Gesamt-Transkriptions-Aktivitäts-Karte des Plasmids pKAP298.....	45
2.2.11	Sequenzierung des Plasmids pKAP298	45
2.2.11.1	Eigene Sequenzierung	45
2.2.11.2	Sequenzierung durch die Firma Seqlab	46
2.2.11.3	Zusammensetzen der Gesamtsequenz.....	47
2.2.12	Bioinformatische Analyse der pKAP298 - Sequenz	47
2.2.12.1	Erstellung von Restriktionskarten.....	47
2.2.12.2	Suche nach Nucleotid-Sequenzhomologien mittels BLAST	47
2.2.12.3	Suche nach potentiell kodierenden Open-Reading-Frames mittels "GeneMark.hmm"	47
2.2.12.4	Suche nach Protein-Homologien der potentiell codierenden ORFs.....	48
2.2.12.5	Suche nach potentiellen Transposons	48
3.	ERGEBNISSE.....	49
3.1	ISOLIERUNG VON CAEDIBACTER TAENIOSPIRALIS UND DES PLASMIDS pKAP298	50
3.1.1	Massenkulturen von Paramecium tetraurelia Stamm 298	50
3.1.2	PCR-Nachweis von Caedibacter taeniospiralis in Paramecien aus Massenkulturen.....	50
3.1.3	Abernten der Paramecium-Massenkulturen.....	51
3.1.3.1	Abernten zur pKAP298-Plasmid-Isolation	51
3.1.3.2	Abernten zur Gesamt-RNA-Isolation	52
3.1.4	Aufschluss der abgeernteten Paramecien und Isolation des Endobionten	52
3.1.4.1	Isolation des Endobionten zur pKAP298-Isolation	52
3.1.4.2	Isolation des Endobionten zur Gesamt-RNA-Isolation.....	53
3.1.5	Isolation des Plasmids pKAP298 aus dem aufgereinigten Endobionten	53
3.1.5.1	Plasmid-Isolation für Southern Blots / zur Subklonierung in E. Coli	53
3.1.5.2	Plasmid-Isolation für Sequenzierungsreaktionen.....	54
3.1.5.3	Kontrolle des Plasmid-Isolats auf Verunreinigung mit Nicht-Plasmid-DNA	54
3.2	SUBKLONIERUNG DER FEHLENDEN pKAP298-FRAGMENTE IN E. COLI.....	56
3.2.1	Subklonierung der mit Pst I und Sac I geschnittenen Fragmente.....	56
3.2.2	Klonierung von Subfragmenten des 10443 Bp Fragments in E. coli.....	60
3.2.3	Versuch die beiden fehlenden Subfragmente des 10443 Bp Fragments zu klonieren	61
3.2.4	Zusammenfassung.....	63
3.3	TRANSKRIPTIONS-AKTIVITÄTS-ANALYSE DES PLASMIDS pKAP298	63
3.3.1	Erstellen von Southern Blots	64
3.3.1.1	Southern Blots aus dem Gesamt-Plasmid pKAP298	64
3.3.1.2	Southern Blots aus den in E. coli subklonierten pKAP298 - Fragmenten	65
3.3.2	Isolation und Aufreinigung der Gesamt-RNA aus Caedibacter taeniospiralis 298	66
3.3.2.1	Isolation der Gesamt-RNA	66
3.3.2.2	Aufreinigung des Gesamt-RNA-Isolats von kontaminierender DNA	67
3.3.3	Umschreiben der Gesamt-RNA in cDNA.....	69
3.3.4	Markieren der cDNA mit digoxigenin-dUTP.....	70
3.3.5	Hybridisierung und Detektion der markierten cDNA-Sonden mit den Southern Blots.....	70
3.3.5.1	Hybridisierung der markierten cDNA-Sonden mit den Southern Blots.....	70
3.3.5.2	Detektion der hybridisierten Southern Blots.....	71
3.3.5.3	Dokumentation der hybridisierten Southern Blots.....	72
3.3.6	Identifizierung der Orte von Genaktivität.....	73
3.3.6.1	Auswertung der detektierten Blots	73
3.3.6.2	Erstellung einer Transkriptions-Aktivitäts-Karte des Plasmids pKAP298	93
3.4	ANALYSE DER pKAP298-SEQUENZ.....	94
3.4.1	Sequenzierung des Plasmids pKAP298	94
3.4.1.1	Eigene Sequenzierung	94
3.4.1.2	Sequenzierung durch Seqlab.....	96
3.4.1.3	Zusammensetzen der Sequenzen zum vollständigen Plasmid pKAP298	97
3.4.1.4	Restriktionskarte des Plasmids pKAP298	97
3.4.2	Bioinformatische Analyse der pKAP298-Sequenz.....	98
3.4.2.1	Potentiell codierende ORFs mit Nucleotid-Sequenzhomologien.....	99
3.4.2.2	Schon beschriebene translatierte ORFs von anderen pKAP-Plasmiden	99

3.4.2.3	Noch nicht beschriebene potentiell kodierende ORFs mit Sequenzhomologien	101
3.4.2.4	Potentiell kodierende ORFs ohne Sequenzhomologien (Orphans).....	113
3.4.2.5	Annotation des Plasmids pKAP298.....	114
3.4.2.6	Transkriptions-Aktivitäts-Karte mit allen wahrscheinlich kodierenden ORFs.....	115
3.4.2.7	Identifizierte mögliche Transposons auf dem Plasmid pKAP298	116
3.4.2.8	Mögliche Operons / polycistronische mRNA.....	121
4.	DISKUSSION	123
4.1	TRANSKRIPTIONS-AKTIVITÄT AUF DEM PLASMID pKAP298.....	123
4.1.1	<i>Analyse der Transkriptions-Aktivität</i>	<i>123</i>
4.1.1.1	ORFs mit Protein-Homologien auf Fragmenten mit nachgewiesener Gen-Aktivität ...	123
4.1.1.2	ORFs ohne Protein-Homologien auf Fragmenten mit nachgewiesener Gen-Aktivität.	126
4.1.2	<i>Weitere Forschungsansätze für zukünftige Untersuchungen am Plasmid pKAP298.....</i>	<i>127</i>
4.1.3	<i>Analyse möglicher Fehlerquellen</i>	<i>128</i>
4.2	EVOLUTION DES PLASMIDS pKAP298.....	131
4.2.1	<i>Untersuchung der Basensequenz nach Ähnlichkeiten zu Phagen-DNA-Sequenzen</i>	<i>132</i>
4.2.2	<i>Untersuchung der translatierten ORFs nach Ähnlichkeiten zu Proteinen viraler Herkunft</i>	<i>132</i>
4.2.3	<i>Taxonomie und Genom des putativen Bakteriophagen aus dem pKAP298 evolvierte</i>	<i>138</i>
4.2.4	<i>Transposons und das Gruppe-II-Intron.....</i>	<i>139</i>
4.2.5	<i>Verwandtschaft von pKAP298 zu anderen pKAP-Plasmiden.....</i>	<i>142</i>
4.2.6	<i>Zusammenfassung und Ausblick.....</i>	<i>143</i>
4.3	VERSUCH DER IDENTIFIZIERUNG DES PUTATIVEN TOXINS.....	145
5.	ZUSAMMENFASSUNG.....	150
6.	ABKÜRZUNGSVERZEICHNIS.....	152
7.	LITERATURVERZEICHNIS	154
8.	DANKSAGUNG	166
ANHANG	168	
	DAS PROJEKT "KONSTRUKTION EINES CRY4B-EXPRESSIONSVEKTORS FÜR <i>C.TAENIOSPIRALIS</i> "	168
	<i>Kurzer Überblick über das Projekt.....</i>	<i>168</i>
	<i>PCR-Longrun-Amplifikation des cry4B-Gens aus dem Plasmid pHT611</i>	<i>169</i>
	<i>Klonierung des amplifizierten cry4B-Gens in den TOPO-Vektor</i>	<i>171</i>
	<i>Umklonierung des cry4B-Gens aus dem TOPO-Vektor in den Shuttle-Vektor pNF2</i>	<i>173</i>
	<i>Der Versuch der "Endmontage" des Cry4B-Expressionsvektors in Saccharomyces cerevisiae... 176</i>	

1. Einleitung

Gegenstand dieser Arbeit ist das Plasmid pKAP298 aus *Caedibacter taeniospiralis* Stamm 298. *Caedibacter taeniospiralis* lebt als obligater Endobiont in *Paramecium tetraurelia*, einem Frischwasser-Ciliaten. In dem folgenden Text werde ich näher auf die Gattung *Caedibacter* eingehen (1.1) und die Art der *Paramecium/Caedibacter*-Interaktionen beschreiben (1.2). Da der *Caedibacter*-Endobiont dem Wirt *Paramecium* eine Killer-Eigenschaft gegenüber anderen, sensitiven Paramecien verleiht, gehe ich als nächstes auf die von *Caedibacter* vermittelten Tötungsmechanismen ein (1.3). Zur Ausprägung der Killer-Eigenschaft sind sogenannte R-Körper notwendig. Dies sind Proteinbänder, die in den Zellen zylindrisch aufgerollt vorliegen. Die für die R-Körper kodierenden Gene sind bei *Caedibacter taeniospiralis* auf einem Plasmid pKAP lokalisiert. Die pKAP-Plasmid-Familie und die genetischen Faktoren für die R-Körper-Synthese werden unter 1.4 beschrieben. Neben den R-Körpern, ist zur Vermittlung der Killer-Eigenschaft noch ein Toxin notwendig, welches wahrscheinlich ebenfalls von pKAP kodiert wird. Einige Hypothesen zur Natur dieses Toxins werden unter 1.5 dargelegt. In 1.6 definiere ich schließlich die Ziele dieser Arbeit. Das Nebenprojekt „Konstruktion eines Cry4B-Expressionsvektors für *Caedibacter taeniospiralis*“ wird im Anhang besprochen.

1.1 *Caedibacter*-Arten und ihre Verwandtschaft zueinander

Bisher konnten in *Paramecium* fünf verschiedene *Caedibacter*-Spezies entdeckt werden (PREER ET AL., 1974, QUACKENBUSH ET AL., 1978 und SCHMIDT ET AL., 1987). Ihr Habitat ist das Cytoplasma des infizierten *Parameciums*, *C. caryophilus* lebt sowohl im Cytoplasma, als auch im Macronucleus. Bakterien des Genus *Caedibacter* wurden bisher nur in Paramecien des Aurelia-Komplexes gefunden (siehe Tab. 1.1.1). Eine Ausnahme bildet hier wiederum *C. caryophilus*, der in *Paramecium caudatum* entdeckt und beschrieben wurde (SCHMIDT ET AL., 1987).

<i>Caedibacter</i> -Art	Wirt	Habitat im Wirt
<i>C. taeniospiralis</i>	<i>Paramecium tetraurelia</i>	Cytoplasma
<i>C. varicaedens</i>	<i>P. biaurelia</i>	Cytoplasma
<i>C. pseudomutans</i>	<i>P. tetraurelia</i>	Cytoplasma
<i>C. paraconjugatus</i>	<i>P. biaurelia</i>	Cytoplasma
<i>C. caryophilus</i>	<i>P. caudatum/P. novaurelia</i>	Macronucleus/Cytoplasma

Tab. 1.1.1: *Caedibacter*-Arten, ihr Wirt und ihr Habitat im Wirt (POND ET AL., 1989)

Die bisher klassifizierten *Caedibacter*-Arten leben in verschiedenen Wirten und bewohnen unterschiedliche Habitate im jeweiligen Wirt

Als *Caedibacter taeniospiralis* klassifizierte *Caedibacter* sind 0,4-1 µm breite, 1-4 µm lange, gram-negative, nicht sporenbildende, unbewegliche und fakultativ anaerobe Stäbchen. Heute werden sie taxonomisch den *Thiotrichales* zugeordnet (BEIR ET AL., 2002).

Taxonomie: Bacteria; Proteobacteria; Gammaproteobacteria; *Thiotrichales*; unklassifizierte *Thiotrichales*

DNA-DNA Hybridisierungen zeigten, dass sich die verschiedenen *Caedibacter*-Spezies genetisch relativ stark unterscheiden (QUACKENBUSH, 1977, 1978). So betrug die DNA/DNA Homologie zwischen *Caedibacter pseudomutans* 51m1 und *Caedibacter varicaedens* 562 nur 15%. Die Gemeinsamkeit der verschiedenen Spezies ist die Fähigkeit R-Körper auszubilden. Dies legt den Schluss nahe, dass der Genus *Caedibacter* eine künstliche Gruppierung aufgrund der Fähigkeit der R-Körper-Synthese ist. Dafür spricht auch die Untersuchung der 16S rRNA von *Caedibacter taeniospiralis* und *Caedibacter caryophilus* (BEIR ET AL. 2002), die zeigte, dass *C. taeniospiralis* eher zu den Gamma-Proteobakterien gehört und *C. caryophilus* größere Ähnlichkeiten mit Alpha-Proteobakterien hat. Dagegen scheinen die extrachromosomalen Elemente, die wahrscheinlich für die R-Körper-Synthese kodieren, relativ nahe verwandt zu sein. Die DNA-DNA Homologie der Bakteriophagen aus *Caedibacter pseudomutans* 51m1 und *Caedibacter varicaedens* 562 betrug zum Beispiel 73% (QUACKENBUSH ET AL., 1986A). Es ist also wahrscheinlich, dass Bakterien verschiedener Genera die gleichen - oder zumindest ähnliche - extrachromosomalen Elemente erworben haben.

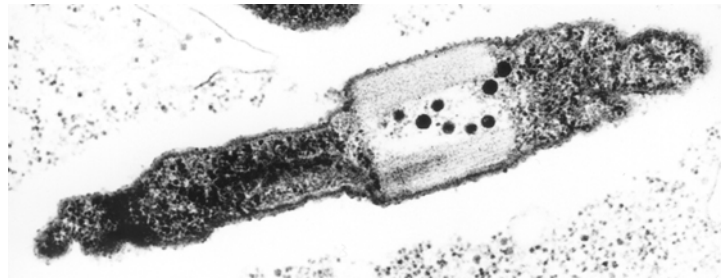


Abb. 1.2.1: „Bright“-Form von *Caedibacter caryophilus*

elektronenmikroskopische Aufnahme, 40000 X Vergrößerung, Foto: Stremmel. Der viereckige helle Bereich im Bakterium zeigt einen aufgerollten R-Körper

1.2 Art der *Caedibacter*-*Paramecium*-Interaktionen

Für *Caedibacter* ist die Endobiose in seinem Wirt obligat, da bisher keine *Caedibacter*-Art außerhalb ihres Wirtes kultiviert werden konnte. Für den Wirt ist der Endobiont nicht obligat. Es kann nicht jeder potentielle *Paramecium*-Wirt von *Caedibacter* infiziert werden. Es muss im Wirt ein Kerngen, genannt „K“, vorhanden sein, um eine Infektion durch *Caedibacter* zu ermöglichen (SONNEBORN, 1943 und 1959). Liegt dieses Kerngen heterozygot vor (Genotyp „kK“), findet man im Wirt nur maximal halb so viele Endobionten wie in Wirten mit homozygotem Genotyp („KK“) (BALSLEY, 1967; CHAO, 1953). Dies deutet auf eine Kodominanz der Allele hin. Erhält ein *Caedibacter* tragender Wirt durch Konjugation den Genotyp „kk“, so verliert er seine Endobionten. *Caedibacter* wird meist als Symbiont deklariert (HECKMANN UND GÖRTZ, 1991), kann allerdings auch als Parasit betrachtet werden (KUSCH ET AL., 2002), daher wird er in dieser Arbeit wertneutral als Endobiont bezeichnet. Der mögliche Vorteil der *Caedibacter*-Infektion für *Paramecien* liegt in dem Erwerb von Killer-Eigenschaften gegenüber sensitiven, nicht infizierten *Paramecien* (SONNEBORN, 1938; LANDIS, 1988; KUSCH ET AL., 2002). Andererseits überleben mit *Caedibacter* infizierte *Paramecien* im Labor, unter Hungerbedingungen, weniger lang als nicht infizierte *Paramecien* (ESTEVE, 1978). Ob dies unter natürlichen Bedingungen von Bedeutung ist, ist allerdings noch unklar.

1.3 Tötungsmechanismen der *Caedibacter*-infizierten *Paramecien*

Bakterien des Genus *Caedibacter* liegen in zwei morphologisch unterschiedlichen Formen vor, der sogenannten Nonbright-Form und der Bright-Form. Die Nonbright-Form ist die reproduktive Form, "Brights" können nicht mehr replizieren. Je nach Fütterungszustand der

Wirte kann der Anteil der "Brights" bis zu 50% der gesamten *Caedibacter*-Population betragen (PREER ET AL., 1974). Kennzeichen der "Brights" sind die refraktilen Körper (=R-Körper) - 20-30 μm lange, zwischen 0,4 μm und 0,8 μm breite Proteinbänder, die in den Zellen zylindrisch aufgerollt vorliegen (POND ET AL., 1989, SCHMIDT ET AL., 1987). Die verschiedenen *Caedibacter*-Arten besitzen spezifische R-Körper-Typen, *Caedibacter taeniospiralis* trägt Typ 51 R-Körper. Typ 7 R-Körper sind in *C. varicaedens*, *C. pseudomutans* und *C. paraconjugatus* zu finden, und *C. caryophilus* besitzt Typ Cc R-Körper.

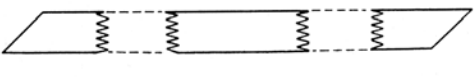
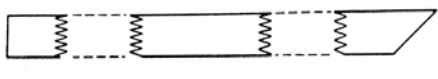
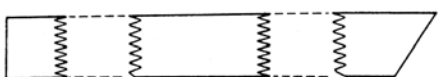
R body type	Ribbon dimensions		Ribbon Morphology		
	maximum length	width	Outer Terminus	Center	Inner Terminus
51	20 μm	0.4 μm			
7	20 μm	0.4 μm			
Cc	$\geq 20 \mu\text{m}$	0.8 μm			

Abb. 1.3.1.: Morphologie der in *Caedibacter* vorkommenden „R-Körper“ (POND ET AL., 1989). Die verschiedenen R-Körper-Typen besitzen unterschiedliche Größen und Morphologien

Es ist noch nicht in allen Einzelheiten geklärt, auf welche Weise sensitive Zellen abgetötet werden. Fest steht, dass die R-Körper an der Vermittlung der Killereigenschaft beteiligt sind. Die "Brights" werden über die Cytopyge in das umgebende Medium abgegeben (PREER ET AL., 1972) und können von sensitiven Paramecien durch Phagocytose aufgenommen werden (AUSTIN, 1948). Nachdem die "Brights" verdaut sind, lässt sich eine Konformationsänderung der R-Körper beobachten (PREER ET AL., 1974). Das Proteinband streckt sich durch Entrollen teleskopartig in die Länge, durchstößt die Membran der Nahrungsvakuole, und der Inhalt der Nahrungsvakuole gelangt in das Cytoplasma (JURAND ET AL., 1971, 1978). Der R-Körper transportiert somit wahrscheinlich ein noch nicht bekanntes Toxin zu seinem Wirkort (DILTS, 1986). Innerhalb weniger Stunden bis Tage sterben die sensitiven Zellen ab. Dabei zeigen die sensitiven Zellen vor ihrem Tod unterschiedliche prälethale Symptome, die charakteristisch für die verschiedenen *Caedibacter*-Arten sind (SONNEBORN, 1959).

Caedibacter-Art	“R-Körper”-Typ	Tötungsmechanismus
<i>C. taeniospiralis</i>	51	“hump”-killing
<i>C. varicaedens</i>	7	Verschiedene Typen
<i>C. pseudomutans</i>	7	“spin”-killing
<i>C. paraconjugatus</i>	7	“mate”-killing
<i>C. caryophilus</i>	Cc	“paralysis”-killing

Tab. 1.3.2: verschiedene Tötungsmechanismen unterschiedlicher *Caedibacter*-Arten

Tötungsmechanismus	Erklärung des Tötungsmechanismus
“hump”-killing	an den betroffenen Zellen bildet sich eine große Blase am aboralen Pol aus
“spin”-killing	“Sensitive” Zellen rotieren gegen den Uhrzeigersinn
“vacuolization”	Es bilden sich große Vakuolen in den betreffenden Zellen aus
“paralysis”	Betroffene Zellen schwimmen nicht weiter, ansonsten keine auffälligen Reaktionen
“mate”-killing	“sensitive” Zellen sterben nur dann ab, wenn sie in physischen Kontakt mit den “Killern” kommen, z.B. während der Konjugation

Tab. 1.3.3: Erklärung der Tötungsmechanismen

1.4 die pKAP-Plasmid-Familie und genetische Faktoren der R-Körper-Synthese in *Caedibacter taeniospiralis*

1.4.1 Das Plasmid pKAP aus *Caedibacter taeniospiralis*

Von den 5 bekannten *Caedibacter* Spezies ist *Caedibacter taeniospiralis* die einzige, die ein Plasmid als extrachromosomales Element beinhaltet (DILTS, 1976). DILTS entdeckte, dass die Kopienzahl des Plasmids in Bright-Partikeln höher ist als in der reproduktiven Form der *Caedibacter*. Das Plasmid pKAP ist in allen Stämmen von *Caedibacter taeniospiralis* vorhanden, hat aber in unterschiedlichen Stämmen verschiedene Größen. Die Plasmidgrößen bisher untersuchter pKAP-Plasmide liegen zwischen 41,5 und 50,5 kB. Durch DNA/DNA-Hybridisierungen und Restriktions-Vergleiche konnte gezeigt werden, dass sich die Plasmide sehr ähnlich sind (QUAKENBUSH ET AL., 1986). Die Plasmide unterscheiden sich nur durch

die Insertion von 1,5 kB und/oder 7,5 kB großen Stücken an unterschiedlichen Stellen der Plasmide. Diese Insertionen konnten als Transposons identifiziert und teilweise sequenziert werden (QUACKENBUSH ET AL., 1986, HERUTH ET AL., 1996). Tn4501 und Tn4501/2 sind jeweils 1,5 kB groß und sehr ähnlich. Auf manchen pKAP-Plasmiden, nicht jedoch auf pKAP298, ist zusätzlich noch das Transposon Tn4502 lokalisiert, das ebenfalls große Ähnlichkeit zu Tn4501 aufweist. Tn4503 umfasst 7,5 kB und besitzt eine Region, die zu den kleineren Transposons homolog ist. Es ist wahrscheinlich, dass Tn4503 aus einem 6 kB großen, nicht homologen Transposon besteht, in das Tn4501/2 insertiert worden ist. Wenn diese Transposons nicht vorhanden wären, hätte jedes pKAP-Plasmid die gleiche Restriktionskarte. Wie DNA-Hybridisierungen zeigten, sind die Transposons ebenfalls im Genom von *Caedibacter taeniospiralis* vorhanden. Bisher wurden auf dem Transposon (Tn4501 und Tn4502) nur die Gene der vermuteten Transposase charakterisiert, andere Funktionen der Transposons sind nicht bekannt. Wie in 1.4.2 beschrieben, befinden sich auf den pKAP-Plasmiden auch die R-Körper kodierenden Gene. Diese wurden bei den Plasmiden pKAP116 und pKAP47 sequenziert und wiesen keine Unterschiede auf (HERUTH ET AL., 1994). Von sonstigen Genen auf den Plasmiden ist nichts bekannt. Es wird jedoch vermutet, dass sich das/die Toxingen/e ebenfalls auf den Plasmiden befinden (siehe 1.5). Auch über den Ursprung der Plasmide ist ebenfalls nichts genaues bekannt, man vermutet aber, dass es sich bei den Plasmiden um Bakteriophagen-Derivate handelt, die zum Plasmid evolviert sind und die Fähigkeit zur Zell-Lyse verloren haben. Diese These unterstützt auch die Tatsache, dass, bei der Untersuchung von R-Körpern aus *Caedibacter varicaedens*, Phagen-Capsomer ähnliche Strukturen beobachtet werden konnten (PREER ET AL., 1972). Des weiteren konnten aus *Caedibacter varicaedens* Stamm 562 Bakteriophagen mit darin verpackter DNA isoliert werden (PREER ET AL., 1971). Ein weiterer Hinweis auf die Phagen-Herkunft ist die Tatsache, dass die Produktion von R-Körpern mit UV-Licht verstärkt werden kann (PREER ET AL., 1974). Es ist bekannt, dass UV-Strahlung einige Prophagen zur Aktivität induziert. Hungerbedingungen erhöhen ebenfalls die Produktion von R-Körpern. R-Körper aus *Caedibacter taeniospiralis* sind zwar nicht mit Capsomer-ähnlichen Strukturen assoziiert, es konnten allerdings kleine, helikale Strukturen beobachtet werden (~18 nm breit, unbestimmte Länge), die nur in der Bright-Form der *Caedibacter* auftauchen. Es wird vermutet, dass diese Strukturen inkomplette (oder falsch zusammengebaute) Bakteriophagen-(Schwanz)-Stücke sind (PREER ET AL., 1972). In dieser Arbeit wird das Plasmid pKAP298 aus *Caedibacter taeniospiralis* untersucht. Es ist 49112 Bp groß (laut der in dieser Arbeit durchgeführten Sequenzierung) und wurde bisher über eine Restriktionskarte (QUACKENBUSH ET AL., 1983A) charakterisiert. Diese Restriktionskarte diente in der Arbeit (bis die Basensequenz des Plasmids bekannt war) als Basis für die Auswahl geeigneter Restriktionsenzyme und zur Validierung der Ergebnisse.

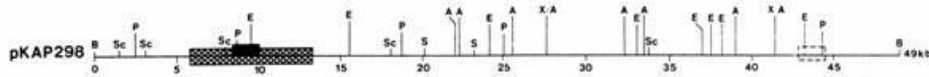


Abb. 1.4.1.1: Restriktionskarte des Plasmids pKAP298 (QUACKENBUSH ET AL., 1983)

Der umrahmte Bereich zwischen ~ 43 kb und 44,5 kb entspricht dem Transposon Tn4501. Der schraffierte Bereich zwischen ~ 5,5 kb und 13,5 kb entspricht dem Transposon Tn4503, in den ein zu Tn4501 homologes Transposon (Tn4501/2) inseriert ist.

1.4.2 R-Körper-Gene

Die für die R-Körper-Synthese codierenden Gene liegen auf extrachromosomalen Elementen. Bei R-Körpern Typ 7 und Cc wird die Synthese wahrscheinlich durch Bakteriophagen induziert (PREER ET AL., 1974, SCHMIDT ET AL., 1987). Bei R-Körpern des Typs 51 liegen die codierenden Gene jedoch auf einem Plasmid, das, je nach *Caedibacter*-Stamm, zwischen 41,5 bis 50,5 kb groß ist (siehe 1.4.1). Dies konnte durch Subklonierung einzelner Plasmid-Fragmente in *E. coli* nachgewiesen werden (QUACKENBUSH ET AL., 1983A,B). Die mit einem bestimmten Plasmid-Fragment transformierten Bakterien bildeten plötzlich R-Körper. Der kodierende Bereich ist 1,8 kb groß und wurde sequenziert (HERUTH ET AL., 1994). Er besteht aus drei Genen (*rebA*, *rebB*, *rebC*), die jeweils für ein 18 kDa, ein 13 kDa und ein 10 kDa großes Protein kodieren. Die R-Körper-Synthese erfolgt über zwei Polymerisationsschritte, bei denen einmal *rebA* und *rebB*, das andere Mal alle drei Proteine beteiligt sind. Über die Rolle eines putativen vierten Gens auf dem sequenzierten Fragment (*rebD*) ist nichts bekannt, es wird aber vermutet, dass es bei der Ausprägung des Toxins beteiligt ist (HERUTH ET AL., 1994, POND, 2002, unveröffentlicht). Die in *E. coli* exprimierten R-Körper wiesen keine Toxizität gegenüber sensitiven Paramecien-Stämmen auf. Die toxin-kodierende Sequenz ist noch nicht lokalisiert, sie wird jedoch auf den Plasmiden vermutet (POND ET AL., 1989), da mit der Synthese der R-Körper wahrscheinlich auch die Produktion des hypothetischen Toxins beginnt. Sicher ist dies jedoch nicht. Das Toxin könnte wohl auch in der Nonbright-Form der *Caedibacter* produziert werden, könnte dann allerdings, aufgrund des fehlenden R-Körpers, die Nahrungsvakuole im *Paramecium*-Wirt nicht verlassen.

1.5 Hypothesen zur Natur des Toxins

Es gibt Hinweise darauf, dass das Toxin ein Protein ist. So hat GIBSON (unpublizierte Daten), ein 20-25 kDa Protein aus einem Killer-*Paramecium*-Homogenat (*P. tetraurelia* Stamm 51) isoliert, das toxisch auf sensitive Paramecien wirkte. Das Protein wurde bisher nicht genauer charakterisiert. Immunologische Kreuz-Reaktionen mit den R-Körper-Proteinen wurden jedoch nicht beobachtet. Die Toxizität eines Killer-*Paramecium*-Homogenats nimmt zu, wenn

man das Homogenat vor dem Verfüttern an sensitive Paramecien 1-2 Stunden "reifen" lässt. Dies spricht dafür, dass das Toxin nicht immer in den *Caedibacter* vorhanden ist, sondern unter Stressbedingungen neu synthetisiert wird (BUTZEL ET AL., 1960). Wenn man das Homogenat mit verschiedenen biochemischen Inhibitoren behandelt, so kann die Toxizität verändert werden. Die Zugabe von Natrium-Azid und Dinitrophenol verringern die Toxizität, was darauf schließen lässt, dass die "Reifung" des Toxins vom oxidativen Stoffwechsel abhängt (BUTZEL ET AL., 1960). Der verringern Effekt von Dinitrophenol konnte durch Zugabe von ATP blockiert werden. Anscheinend ist die Produktion des Toxins nicht von der Proteinbiosynthese abhängig, da mit Chloramphenicol behandeltes Homogenat keine Verringerung der Toxizität zeigte. Die Behandlung von sensitiven Paramecien mit Chloramphenicol bot aber einen gewissen Schutz. Möglicherweise liegt das Toxin in den *Caedibacter* schon als Protoxin vor, welches, durch eine ATP-abhängige Reaktion, in das Toxin umgewandelt werden kann. Die Regulation dieser Modifikation könnte vom Krebs-Zyklus abhängig sein, denn die Behandlung mit Krebs-Zyklus-Inhibitoren wie Malonat oder Natrium-Fluorid erhöht die Toxizität. Erhitzung des Killer-Homogenats auf mindestens 55 °C, extreme pH-Werte und einige proteolytische Enzyme inaktivieren ebenfalls die Toxizität (SONNEBORN, 1959), was wiederum für ein Protein als Toxin spricht. Die Behandlung von sensitiven Zellen mit Lecithin bietet einen Schutz vor dem Toxin (BUTZEL ET AL., 1962). Dies könnte darauf hindeuten, dass der Wirkort des Toxins die Zellmembran ist. Das Toxin wirkt sich möglicherweise negativ auf die osmoregulatorischen Fähigkeiten der *Paramecium*-Zelle aus. Dazu passen auch die sichtbaren Effekte des Tötungsmechanismus, wie z.B. die Blasenbildung am aboralen Pol der sensitiven Zelle beim "hump-killing"-Mechanismus. N-acetyl-D-Galactosamin schützt sensitive Zellen ebenfalls vor dem Absterben. Es wird vermutet, dass der Zucker an das Toxin bindet und es somit inaktiviert.

1.6 Ziele der Arbeit

Ein Ziel dieser Arbeit war eine weitere Charakterisierung des Plasmids pKAP298, und zwar sowohl auf der Basensequenz-Ebene, als auch bei der Charakterisierung möglicher Gen-Produkte. Zum einen sollte das gesamte Plasmid sequenziert werden, um durch Sequenzvergleiche mögliche Homologien zu schon bekannten Sequenzen zu finden. Zum anderen sollten, durch Untersuchungen mit der transkribierten mRNA, Orte potentieller Genaktivität auf dem Plasmid lokalisiert werden (4.1). Dadurch sollten Rückschlüsse auf die evolutive Herkunft des Plasmids ermöglicht werden (4.2), und vom Plasmid kodierte Genprodukte sollten bestimmt werden. Ein weiteres Ziel war es, zur Identifizierung des wahrscheinlich auf dem Plasmid beheimateten Toxingens beizutragen (4.3). Diese Arbeit soll eine Basis für weitergehende und spezifischere Studien am Plasmid pKAP298 bilden. Darum

sollen auch Hinweise auf weitere Forschungsansätze gegeben werden. Bisher wurde noch nicht versucht, direkt über die mRNA *in vivo* Transkriptions-Aktivität auf dem Plasmid nachzuweisen. Diese Methode sollte mit dieser Arbeit für spätere, weiterführende Forschung etabliert werden. In meiner Diplomarbeit (JEBLICK, 2000) wurde versucht, das Plasmid pKAP298 in *E. coli* zu klonieren. Ein ca. 19,4 kB großes Stück konnte nicht subkloniert werden. In dieser Arbeit sollte nun erforscht werden, warum dies nicht gelang. Meine Vermutung war, dass auf dem nicht klonierbaren Stück ein für das Toxin kodierendes Gen lokalisiert ist. Durch die Sequenz-Analyse und Transkriptions-Aktivitäts-Analysen ist dies vielleicht nachweisbar. Des Weiteren sollte während dieser Arbeit versucht werden, das fehlende 19,4 kB große Fragment des Plasmids in kleineren Stücken in *E. coli* zu klonieren.

2. Material und Methoden

2.1 Materialien

2.1.1 Geräte

Neben den Standardlaborgeräten wurden folgende Geräte verwendet. Sie werden hier aufgrund ihrer besonderen Spezifität und Relevanz für die durchgeführten Versuche genannt:

Gerät	Hersteller
Autoklav Varioklav Typ E	H+P Labortechnik (Oberschleißheim)
Biofuge A (Rotor 1386)	Heraeus Sepatech (Osterode)
E. coli Pulser™ / Gene Pulser II	Bio-Rad
Feinwaage MC1 Analytic AC 210S	Sartorius (Göttingen)
Gene Quant II RNA/DNA Calculator	Pharmacia Biotech (Freiburg)
Kläreseparator KA 05-00-105 ("Durchlaufzentrifuge")	Westfalia Separator (Oelde)
Kühlzentrifuge Z233MK	Hermle Labortechnik (Wehingen)
Labor-pH-Meter 646	Knick (Berlin)
Laborwaage MC1 Laboratory LC 620P	Sartorius (Göttingen)
Magnetrührer (heizbar)	IKA Labortechnik (Staufen)
Midi-Gelkammer	Renner (Darmstadt)
Mikrowellengerät Alaska M1400	Beckmann Industr. (München)
Öltestzentrifuge HNS II mit ASTM-Zentrifugenglasrotor	Damon/IEC Division Needham, USA
Pipettierhilfen	Gilson (Villiers le bel, Frankreich)
Pipettierhilfen	Eppendorf (Hamburg)
Spannungsquelle	Elektrowerkstatt der Universität Kaiserslautern
Schüttler	IKA Labortechnik (Staufen)
Sicherheitsbrenner	Schütt Labortechnik (Göttingen)
Sorvall Kühlzentrifuge RC-5C Plus mit Rotoren SS-34 und GSA	Du Pont (Wilmington, USA)
Sterilbank Clean Air	Clean Air (Niederlande)
Programmierbarer Thermoblock Uno mit Heizdeckel	Biometra (Tampa, USA)

Gerät	Hersteller
Wärmeschrank GTRO 214	Memmert (Schwabach)
UV-Illuminator UV-FLX-35M	Serva (Heidelberg)

2.1.2 Kunststoffmaterial

Reaktionsgefäße und Pipettenspitzen wurden größtenteils von der Firma Eppendorf (Hamburg) bezogen. Sie wurden vor Gebrauch autoklaviert. Sterilfilter mit Porengröße 0,2 µm und 0,45 µm wurden von der Firma Schleicher & Schüll (Dassel) bezogen. Sterile Kunststoffröhrchen der Größen 15 ml und 50 ml wurden über die Firma Fischer Labortechnik (Saarbrücken) bezogen. 250 ml Zentrifugenbecher kamen von der Firma Du Pont (Wilmington, USA). Die 20 ml Einwegspritzen zur Sterilfiltration stammen von der Firma Braun (Melsungen). Petrischalen (Durchmesser 90 mm) wurden von der Firma Roth (Karlsruhe) bezogen.

2.1.3 Glasmaterial

Reagenzgläser, Erlenmeyerkolben, Bechergläser und Messzylinder stammen von der Firma Schott (Mainz) und wurden je nach Bedarf vor Gebrauch sterilisiert. Corex-Zentrifugenröhrchen der Größe 15 ml und 30 ml wurden von der Firma Du Pont (Wilmington, USA) bezogen. Pasteurpipetten kamen von der Firma Brand (Mainz) über die Firma Fischer Labortechnik (Frankfurt). Als Massenkulturgefäße wurden 20 Liter Kolben der Firma Schott (Mainz) benutzt.

2.1.4 Optische Geräte

Gerät	Hersteller
Binokular 1x-6.5x CO 11	Olympus (Hamburg)
Scanner SnapScan 1212P	Agfa (Leverkusen)
Sofortbildkamera MP4+	Polaroid (Offenbach)

2.1.5 Fotomaterial

Mit dem Polaroid-Kamerasystem MP4+ wurden Ethidiumbromid-gefärbte Agarosegele unter UV-Anregung fotografiert. Zur Dokumentation wurde der Sofortbild-Film Polaroid Nr. 667 (Polaroid, St. Albans, GB) benutzt. Alle Bilder wurden mit dem Scanner Snapscan 1212P

(Agfa, Leverkusen) digitalisiert und mit dem Programm Photoshop (Version 5.5) der Firma Adobe nachbearbeitet.

2.1.6 Chemikalien

Die verwendeten Chemikalien wurden, falls im Text nicht besonders vermerkt, in p.a. Qualität von folgenden Firmen bezogen:

Firma	Firmensitz
Baker	Deventer, NL
Biomol/Feinchemikalien	Hamburg
Biozym	Oldendorf
Boehringer Mannheim	Mannheim
Difco	Detroit, USA
Fluka Biochemika	Buchs, CH
Gibco BRL	über Life Technologies, Eggenstein
ICN Biochemicals	Aurora, USA
Kimberley Research	Atlanta, USA
Life Technologies	Eggenstein
Merck	Darmstadt
Pharmacia	Freiburg
Promega	Heidelberg
Riedel de Haen	Seelze
Roth	Karlsruhe
Serva Feinbiochemika	Heidelberg
Sigma	Deisenhofen

Im Falle von Fertigprodukten folgt in den Anleitungen die spezifische Produktbezeichnung.

2.1.7 Elektrophorese

2.1.7.1 *Verwendete Agarose*

Analytische DNA-Agarosegelelektrophoresen wurden mit "Ultra Pure" Agarose der Firma Life Technologies (Paisley, GB) hergestellt.

2.1.7.2 DNA Molekulargewichtsmarker

DNA Molekulargewichtsmarker wurden von den Firmen MBI-Fermentas (St. Leon Roth) und Peqlab (Erlangen) bezogen.

Marker	Bandengrößen
100 kB DNA-Ladder+	3 kB, 2 kB, 1.5 kB, 1.2 kB, 1031 Bp, 900 Bp ... bis 100 Bp
λ19Mix	48502 Bp, 38416 Bp, 33498 Bp, 29946 Bp, 24508 Bp, 23994 Bp, 19397 Bp, 17053 Bp, 15004 Bp, 12220 Bp, 10086 Bp, 8614 Bp, 8271 Bp, 1503 Bp
λEcoT13 Marker	19329 Bp, 7743 Bp, 6223 Bp, 4254 Bp, 3472 Bp, 2690 Bp, 1882 Bp, 1489 Bp, 925 Bp, 421 Bp, 74 Bp
1kb-Ladder	10000 Bp, 8000 Bp, 6000 Bp, 5000 Bp, 4000 Bp, 3500 Bp, 3000 Bp, 2500 Bp, 2000 Bp, 1500 Bp, 1000 Bp, 750 Bp, 500 Bp, 250 Bp

2.1.7.3 Loading Dye

Alle Ansätze, die auf ein Agarose-Gel aufgetragen werden sollten, wurden mit dem Beladungspuffer 6x Loading Dye Blue-Orange der Firma Promega (Heidelberg) versetzt. Die Farbstoffe Bromphenol-Blau, Orange G und Xylene C wandern ungefähr analog zu DNA-Fragmenten der Größen 300 Bp, 50 Bp bzw. 4 kB.

2.1.8 Enzyme

Die Enzyme T₄-DNA-Ligase und Taq-DNA-Polymerase, sowie sämtliche Restriktionsendonucleasen mit den entsprechenden Reaktionspuffern, wurden von der Firma Promega (Heidelberg) bezogen. Die im Anhang verwendete Pwo-Polymerase stammt von der Firma Peqlab (Erlangen).

2.1.9 Transformationsvektoren

Der Plasmid-Vektor pHSG298 wurde von der Firma Takara (Japan) bezogen. Er basiert auf dem pUC18-Vektor und besitzt ein Kanamycin-Resistenzgen.

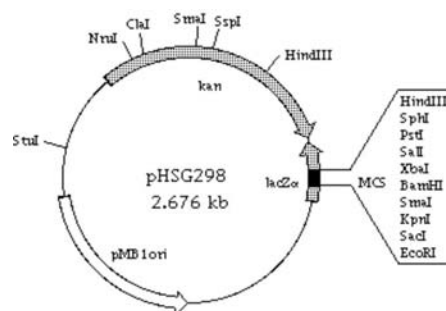


Abbildung 2.1.9.1: Plasmid-Vektor pHSG298 (nach Takara)

2.1.10 Wirtsorganismus für die Vektoren

Als Wirtsorganismen für die Klonierungen wurden Zellen der *E. coli* Stämme JM109 (Promega) und TOP10 (Invitrogen) verwendet. Diese wurden, wie im Methodenteil beschrieben, zur Elektroporation kompetent gemacht.

Der Genotyp des verwendeten Stammes JM109 lautet wie folgt: *endA1, recA1, gyrA96, thi, hsdR17, (r_K⁻, m_K⁺), relA1, supE44, λ⁻, Δ(lac-proAB), [F⁺, traD36, proAB, lacI^qZΔM15]*.

Der Genotyp des verwendeten Stammes TOP10 lautet wie folgt: F⁻ *mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 deoR araD139 Δ(ara-leu)7697 galU galK rpsL (Str^R) endA1 nupG*.

2.1.11 Verwendete “Kits”

Während dieser Arbeit wurden folgende “Kits” verwendet:

Kit	Hersteller
DIG High Prime DNA Labeling and Detection Starter Kit II	Roche (Mannheim)
NucleoSpin Extract 2 in 1	Machery Nagel (Düren)
Quiagen Plasmid MIDI Kit	Quiagen (Hilden)
RevertAid H Minus First Strand cDNA Synthesis Kit	MBI Fermentas (St. Leon Roth)
Thermo Sequenase Fluorescence Labeled Primer Cycle Sequencing Kit	Amersham Pharmacia Biotech Europe GmbH (Freiburg)
TOPO PCR XL Cloning Kit	Invitrogen (Frankfurt am M.)

2.1.12 Oligonucleotide

Für PCR-Reaktionen wurden folgende Oligonucleotide verwendet:

Name	Sequenz
Cry4BF	5'-ATAT GTCGAC <i>GGATCC</i> GAAATAGCGAGGTTTCCCATGTCTA-3'
Cry4bR	5'-ATAT GTCGACT CTAGATATCTATTTTTTATGTTTT-3'
pkap476F	5'-GCGTTGATAAAGCATAAGTATTA AAAA ATAG-3'
pkap1264R	5'-TACCTATAACATTCATCATTATTTGGAAGT-3'
YEP24F	5'-CTTTTCTAGGGGGTCTGACG-3'
YEP24R	5'-GGAGGACCGAAGGAGCTAC-3'

Die Abkürzungen R und F im Primernamen bedeuten “reverse” und “forward”. Sie besagen an welchen Strang der Template-DNA der jeweilige Primer bindet. Die in den Cry4BF und Cry4BR Primern **fett** gedruckten Bereiche bezeichnen eine *Sal* I-Schnittstelle. Diese Bereiche binden nicht an die Template-DNA, sondern dienen zur späteren Aufarbeitung des PCR-Produkts. Der kursive Bereich im Primer CryBF ist eine *Bam* HI-Schnittstelle. Sie wird ebenfalls später noch gebraucht (siehe Anhang). Die Primer wurden von der Firma MWG Biotech (Ebersberg) bezogen.

2.1.13 Stammlösungen, Puffer und Kulturmedien

Es werden hier selbst angesetzte Lösungen, Puffer und Kulturmedien aufgeführt. Lösungen aus verwendeten “Kits” werden nur genannt, wenn diese neu angesetzt werden mussten.

Stammlösung / Puffer / Medien	Herstellung
0,1 M IPTG	1,2 g IPTG werden in 50 ml Aqua _{bidest} gelöst, anschließend sterilfiltriert und bei 4°C gelagert
20 x Natriumphosphat Puffer für <i>Paramecium</i> -Kulturmedium	27,8 g NaH ₂ PO ₄ xH ₂ O 35,6 g Na ₂ HPO ₄ x2H ₂ O werden in 1 L Aqua _{bidest} gelöst und bei Raumtemperatur gelagert
20 x SSC	175.3 g NaCl 88.2 g Natrium-Citrat

Stammlösung / Puffer / Medien	Herstellung
	werden in 800 ml Aqua _{bidest} gelöst und der pH-Wert auf pH=7,0 eingestellt. Danach wurde das Volumen auf 1 L aufgefüllt.
50 x TAE-Puffer	<p>2,0 M Tris 1,0 M Essigsäure 0,1 M EDTA</p> <p>243 g Tris werden in 800 ml Aqua_{bidest} gelöst. Es werden 57,1 ml Eisessig und 37,2 g EDTA hinzugefügt sowie der pH-Wert mit konzentrierter HCl auf 7,5 eingestellt. Das Volumen der Lösung wird mit Aqua_{bidest} auf 1 L aufgefüllt. Die Lagerung erfolgt bei Raumtemperatur</p>
Antibody-Solution	Zu 10 ml Blocking-Solution wird 1,5 µl des Anti-Digoxigenin-Antikörpers aus dem Digoxigenin-Kit gegeben. Die Antikörper-Stammlösung muss vor der Zugabe für 5 min bei 10 000 g zentrifugiert werden
Blocking-Solution	10 ml 10 x Blocking-Solution aus dem Digoxigenin-Kit werden mit 90 ml Maleinsäurepuffer vermischt
DEPC- Aqua _{bidest}	<p>1 ml DEPC</p> <p>werden in 1 L Aqua_{bidest} gegeben und 1 h lang gerührt. Danach wird die Lösung autoklaviert</p>
Detektionspuffer	<p>0,1 M Tris 0,1 M Nach</p> <p>werden in 800 ml Aqua_{bidest} gelöst, der pH-Wert auf pH=9,5 eingestellt und mit Aqua_{bidest} auf 1 l aufgefüllt</p>
Kaliumphosphat-Puffer für <i>Caedibacter</i> -Isolation	<p>0,01 M KH₂PO₄ 0,01 M K₂HPO₄ 150 mM Sorbitol</p>

Stammlösung / Puffer / Medien	Herstellung
	werden in 1 L Aqua _{bidest} gelöst und bei Raumtemperatur gelagert
Kanamycin (50 mg/ml)	0,5 g Kanamycin wird mit Aqua _{bidest} auf 10 ml aufgefüllt, sterilfiltriert und bei -20°C gelagert
LB Medium	10 g Bacto [®] Tryptone (oder Peptone) 5 g Bacto [®] Yeast Extract 5 g NaCl werden in 1 L Aqua _{bidest} gelöst, der pH mit konzentrierter NaOH auf 7,0 eingestellt, in Glasflaschen abgefüllt und autoklaviert. Die Lagerung erfolgt bei Raumtemperatur
LB-Agarplatten mit Kanamycin, IPTG und X-Gal	3 g Agar werden mit 200 ml LB-Medium im Mikrowellenherd gekocht bis der Agar vollständig gelöst ist. Wasserverluste werden mit Aqua _{bidest} ergänzt. Nach Abkühlung des Mediums auf ca. 60°C wird 200 µl Kanamycin-Stammlösung, 2 ml IPTG-Stammlösung und 0,8 ml X-Gal-Stammlösung zugegeben. Danach wird der Agar bis zu einer Schichtdicke von ca. 3 mm in Petrischalen gegossen und aushärten gelassen. Die Lagerung der Platten erfolgt bei 4 °C.
Maleinsäurepuffer	0,1 M Maleinsäure 0,15 M NaCl werden in 800 ml Aqua _{bidest} gelöst, der pH-Wert mit NaOH auf pH=7,5 eingestellt und mit Aqua _{bidest} auf 1 l aufgefüllt
Post-Hybridisierungs-Lösung	100 ml 20 x SSC 1 g SDS

Stammlösung / Puffer / Medien	Herstellung
	werden mit 900 ml Aqua _{bidest} aufgefüllt und bei Raumtemperatur gelagert
Stammlösung für <i>Paramecium</i> -Kulturmedium	Die Grundlage des Mediums bildet Wiesenheu (Fa. Dehner) 400 g Wiesenheu werden mit 4 Litern Aqua _{bidest} für 1 Stunde gekocht und für 5 Stunden abkühlen gelassen. Das Heu wird danach gründlich ausgepresst und die Stammlösung über 8 Lagen Gaze filtriert. Anschließend wird die Stammlösung für 20 min bei 121 °C autoklaviert. Die Aufbewahrung der Lösung erfolgt bei Raumtemperatur
Waschpuffer zur Hybridisierung	0,1 M Maleinsäure 0,15 M NaCl 3 ml Tween 20 werden in 800 ml Aqua _{bidest} gelöst, der pH-Wert mit NaOH auf pH=7,5 eingestellt und mit Aqua _{bidest} auf 1 l aufgefüllt
X-Gal	100 mg 5-Chlor-4-Brom-3-Indolyl-β-D-Galaktosid werden in 2 ml N,N'-dimethylformid gelöst, das Reaktionsgefäß mit Aluminiumfolie umwickelt und bei -20°C gelagert

2.1.14 Biologisches Material

2.1.14.1 *Paramecium* Stamm

Zur Untersuchung wird ausschließlich der Stamm *Paramecium tetraurelia* 298K (PREER ET AL., 1972, gefunden in Empire Range/Panama) benutzt (dankend erhalten von Prof. Dr. Fokin, St. Petersburg).

2.1.14.2 Endobionten

Die Paramecien sind mit dem Endobionten *Caedibacter taeniospiralis* infiziert. Die Symbionten sind Träger des Plasmids pKAP298.

2.1.14.3 Futterbakterien

Paramecium-Futter-Medien wurden ausschließlich mit *Klebsiella minuta* als Futterorganismus beimpft. Aufgrund der unsterilen Haltung der *Paramecium*-Kulturen kam es nach und nach zur Besiedelung der Kulturen mit Hefen, anderen (nicht identifizierten) Bakterien, Flagellaten u.ä.

2.2 Methoden

2.2.1 Aufzucht und Aufarbeitung von *Paramecium/Caedibacter* - Massenkulturen

2.2.1.1 Nachweis und Identifizierung des *Caedibacter-Endobionten*

Da die Endobionten durch zu schnelles Wachstum des Wirtes ausdünnen können, bzw. der Wirt die *Caedibacter* vollständig verlieren kann (PREER ET AL., 1972), muss in regelmäßigen Abständen kontrolliert werden, ob noch Bakterien in den Paramecien vorhanden sind. Somit kann auch ausgeschlossen werden, dass versehentlich Kulturen vertauscht oder gemischt worden sind. Das Test-Verfahren beruht auf der PCR-Amplifizierung und elektrophoretischen Detektion eines ~800 Bp großen DNA-Fragments des Plasmids. Die PCR-Primer stammen aus einer schon bekannten Sequenz des Plasmids (R-Körper kodierender Bereich, HERUTH ET AL., 1994). Um eine hohe Spezifität zu gewährleisten, wird die Größe der Primer auf 30 Basen festgesetzt. Aus einem Aliquot der zu testenden Kultur werden 5-6 Zellen isoliert und 3 x in Aqua_{dest} gewaschen. Das Waschen hat schnell zu erfolgen, da die Zellen sonst in dem Aqua_{dest} platzen und so die Endobionten freisetzen. Die gewaschenen Zellen werden nun mit 25 µl Aqua_{dest} in ein 100 µl PCR-Tube überführt. Die Lyse der Zellen erfolgt über starkes Erhitzen der Zellen und anschließenden mehrmaligen Temperaturwechsel. Ein spezieller Lysis-Puffer ist in diesem Verfahren nicht nötig. Die Lyse findet in einem Thermocycler mit Heizdeckel statt.

Thermocycler-Programm:

- 10 min bei 99 °C
- 1 min bei 65 °C
- 1 min bei 99 °C
- die letzten beiden Schritte werden 10 x wiederholt

Nun wird aus dem Lyse-Produkt mittels PCR die Nachweisbande amplifiziert. Dabei werden folgende Primer verwendet: pkap476F und pkap1264R, Schmelztemperatur jeweils 58,6 °C.

Der PCR-Ansatz wird auf Eis zusammenpipettiert und gut durchmischt:

- 25 µl Lyse-Produkt
- 5 µl 10 x PCR-Puffer
- 5 µl MgCl₂-Puffer (Endkonzentration 2.5 mM MgCl₂)
- 4 µl dNTP-Mix (Konzentration 2 mM pro dNTP)
- 1 µl Primer pkap476F (100 pMol)
- 1 µl Primer pkap1264R (100 pMol)
- 0,5 µl *Taq*-DNA-Polymerase (2,5 units)
- 8,5 µl Aqua_{bidest}
- 50 µl Gesamtansatz

PCR-Programm für Thermocycler (105 °C Deckel-Temperatur):

- 30 s Denaturierung bei 94 °C
- 30 s Primer-Annealing bei 53 °C
- 1 min Primer Extension bei 72 °C
- die 3 Schritte werden 35 mal wiederholt

Die Dokumentation der PCR-Produkte erfolgt über Gel-Elektrophorese und anschließendes Fotografieren des mit Ethidiumbromid gefärbten Gels im UV-Licht.

2.2.1.2 Bereiten des Futter-/Kultur-Mediums

Als Basis des Kulturmediums wird Wiesenheu (Fa. Dehner Gartencenter) verwendet, das sich als gleichwertig zu dem bisher verwendeten "Wheat Gras Powder"-Futtergranulat erwiesen hat.

- 200 ml 20 x Natriumphosphat-Puffer werden mit 170 ml Kultur-Stammlösung in einen 5 l Erlenmeyerkolben gegeben und mit Aqua_{dest} auf 4,5 l aufgefüllt
- das Kulturmedium wird für 20 min bei 121 °C autoklaviert
- den Autoklav lässt man vollständig abkühlen, dabei verringert sich (durch Verdunstung im entstehenden Vakuum) das Volumen in den Kolben auf 4 Liter
- das autoklavierte Medium wird bei Raumtemperatur gelagert

Vor der Verwendung muss das Kulturmedium noch mit Futterbakterien beimpft werden. Dies geschieht durch Überimpfen von *Klebsiella minuta* aus einer Schräg-Agar-Kolonie in das Kulturmedium und anschließender Inkubation bei 33 °C über Nacht.

Futtermedien auf Basis von Hefe-Extrakt, frischer Bäckerhefe oder Pepton sind nicht empfehlenswert, da die Paramecien-Kulturen damit leicht überfüttert werden können. Die Futterbakterien vermehren sich in dem nährstoffreichen Medium sehr schnell und verbrauchen viel Sauerstoff, der dann den Paramecien fehlt. Die Paramecien-Dichte ist im überfütterten Medium dementsprechend viel geringer, im schlimmsten Fall kann die Kultur "umkippen" und alle Pantoffeltierchen sterben. Ungepuffertes Leitungswasser als Basis für das Kulturmedium lässt sich zwar auch benutzen, die maximale Dichte der Paramecien bleibt allerdings hinter der bei Verwendung von Aqua_{bidest} und Natriumphosphat-Puffer zurück. Zudem überleben hungernde Zellen im gepufferten Medium länger als im Medium mit Leitungswasser.

2.2.1.3 Aufzucht der Massenkulturen

Generell beruht die Aufzucht der Massenkulturen auf der Verdoppelung des Volumens alle 2 Tage. Da sich *Paramecium tetraurelia* schneller vermehrt als sein Endobiont *Caedibacter taeniospiralis* (PREER ET AL., 1972), hätte die Kultur nach 24 Stunden (bei 23 °C) zwar wieder ihre maximale Paramecien-Dichte, der Endobiont würde aber ausdünnen, was natürlich nicht erwünscht ist. Daher wird dem Endobionten ein weiterer Tag Zeit gegeben, seine ursprüngliche Dichte im Wirt zu erreichen. Wichtig ist das regelmäßige Filtern der Kulturen über Gaze, um grobe Verunreinigungen zu entfernen. Dadurch wird dem Medium organischer Detritus entzogen, der Schwefelwasserstoff erzeugenden Bakterien - und anderen Anaerobiern - als Substrat dienen kann.

Die Aufzucht der Kultur zum Zwecke der Isolation des Plasmids pKAP298 findet in 5 l Erlenmeyerkolben statt, die jeweils mit 4 Litern Kulturmedium befüllt werden.

- die 4 l *Paramecium*-Kultur wird kurz in kreisenden Bewegungen geschüttelt, damit sich die Kahmschicht auf der Kultur in kleinere Stücke zerteilt
- das Kulturmedium wird durch 8 Lagen Gaze filtriert und auf 2 Erlenmeyerkolben gleichmäßig aufgeteilt
- die Kolben werden mit angeimpften Futtermedium (2.2.2) auf 4 l aufgefüllt und 2 Tage bei Raumtemperatur inkubiert, danach wiederholt sich der Vorgang

Hat die Massenkultur ihr gewünschtes Volumen erreicht (z.B. 32 l), wird sie noch einmal filtriert und dann für 10 Tage bei 12 °C inkubiert. An Tag 5 der Inkubation erfolgt eine weitere Filtration. Falls der Endobiont während der Aufzucht doch ausgedünnt ist, kann er während dieser Zeit wieder seine maximale Konzentration erreichen. Außerdem soll die Kultur in einen Hungerzustand versetzt werden, da dadurch die Kopienzahl des gewünschten Plasmids pKAP 298 in den Endobionten steigt (DILTS, 1976). Aus einem Aliquot (z.B. 1 Liter) der Kultur wird eine neue Massenkultur angelegt.

Die Aufzucht einer Massenkultur zur Gesamt-RNA-Isolation ist identisch mit der zur Plasmid-Isolation, nur dass mehrere Ansätze parallel bereitet werden, die (um sie später qualitativ vergleichen zu können) die gleichen Ausgangsbedingungen haben müssen. Sollen also z.B. 2 Ansätze mit je 16 l gemacht werden, so werden 4 Erlenmeyerkolben Ausgangskultur in einem 20 l Kolben gründlich vermischt und gleichmäßig auf zwei 20 l Kolben aufgeteilt. Diese werden nun mit dem gleichen Volumen (ebenfalls gemischtem) Futtermedium aufgefüllt. Somit sind gleiche Ausgangsbedingungen gewährleistet. Die Inkubationsdauer und Inkubationstemperatur richten sich nach den zu untersuchenden Kulturparametern. Eine Filtration des Ansatzes findet nur statt, wenn die Kultur in Hungerzustand versetzt werden soll.

Folgende verschiedene Ansätze zur Gesamt-RNA-Isolation (mit je 16 l) wurden mehrfach angesetzt (jeweils gleichzeitig):

- Kultur gefüttert, Abernten nach 12 h , Temperatur 23 °C
- Kultur im Hungerzustand, Abernten nach 10 Tagen , Temperatur 23 °C

2.2.1.4 Abernten der Massenkultur

Bei dem Abernten der Massenkultur zur pKAP298-Plasmid-Isolation wird die Paramecien-Kultur zunächst über 8 Lagen Gaze gefiltert und dann mittels einer Durchlaufzentrifuge (Kläreseparator KA 05-00-105) bei niedriger Drehzahl (~ 3000 rpm) auf 400 ml eingeeengt. Nachdem die gesamte Kultur durch die Zentrifuge gelaufen ist, wird mit 2-3 l unbeimpften Kulturmedium nachgespült. Der Zentrifugenrotor, in dem sich die Paramecien-Suspension befindet, wird vorsichtig geschüttelt, um an der Wand haftende Paramecien in die Suspension zu bringen - wenn noch viele Schwebstoffe in der Kultur sind, bildet sich an der Rotorwand ein zäher Belag, in dem sich viele Paramecien verfangen. Die Paramecien-Suspension wird nun auf vier 100 ml Öltestbecher verteilt und kurz gewartet, bis die Paramecien mit den noch im Medium vorhandenen Schwebstoffen aggregieren und zu Boden sinken. Dadurch

verhindert man, dass allzu viele Paramecien und Schwebstoffe während der anschließenden Zentrifugation an der Öltestbecherwand hängen bleiben und kann die Schwebstoffe leicht absaugen, da sie nach der Zentrifugation ein Pellet über dem *Paramecium*-Pellet bilden. Die Öltestbecher werden danach bei 3000 rpm für 5 Minuten in der Öltestzentrifuge (HNS II mit ASTM-Zentrifugenglasrotor) zentrifugiert. Der Überstand wird (zusammen mit dem Schwebstoff-Pellet) mittels einer Wasserstrahlpumpe abgesaugt. Übrig bleibt das *Paramecium*-Pellet, das nun weiter aufgeschlossen werden kann.

Das Abernten der Massenkultur zur Gesamt-RNA-Isolation ist vom Ablauf identisch mit dem Abernten zur Plasmid-Isolation. Da jedoch eine "Momentaufnahme" des RNA-Zustandes des Endobionten *Caedibacter taeniospiralis* gemacht werden soll, und die zu isolierende mRNA bei sich ändernden Bedingungen schnell abgebaut wird (bzw. andere mRNA gebildet wird), muss sehr schnell gearbeitet werden. Ferner wird zum Nachspülen der Durchlaufzentrifuge eiskaltes, unbeimpftes Futtermedium benutzt, um die Zellen abzukühlen. Dadurch sollen die Stoffwechselfunktion verlangsamt werden. Die Öltestbecher werden nach dem Absaugen des Überstandes sofort auf Eis gestellt und weiter aufgeschlossen.

2.2.1.5 Aufschluss der abgeernteten Paramecien / Isolation des Endobionten

Caedibacter taeniospiralis 298

Die Isolierung des Endobionten *Caedibacter taeniospiralis* erfolgt mittels eines (von mir modifizierten) Verfahrens nach JO ANNE MUELLER (1963). MUELLER benutzt zur Isolation einen 0,01 M Kaliumphosphatpuffer zum Waschen der Bakterien. In diesem hypotonen Puffer platzen die Bakterienzellen allerdings beim Vortexen, was Totalverluste an DNA/RNA zur Folge hat. Dem in dieser Arbeit verwendeten Puffer wurde 150 mM Sorbitol hinzugefügt, um dies zu verhindern. Des weiteren benutzt MUELLER eine mit Ecteola 23 ((Epichlorohydrin Triethanolamin)-Zellulose) befüllte Säule zur Isolation der Bakterien. Diese verstopft jedoch sehr leicht und die Folge ist wiederum ein Totalverlust der DNA/RNA. In dieser Arbeit wird das Adsorbens Ecteola 23 direkt mit der Bakterien/Detritus-Suspension gemischt und vorsichtig durch einen Kaffeefilter gedrückt. Das folgende Protokoll bezieht sich auf eine Isolation aus 31 l ausgehungertes *Paramecium*-Kultur mit einer Zelldichte von ~1200 Zellen/ml. Der Zelltitert kann durch mehrmaliges Auszählen kleiner Aliquots bestimmt werden.

- es werden 1,5 g Ecteola 23 abgewogen und in 100 ml Kaliumphosphatpuffer 30 min quellen und äquilibrieren gelassen. Danach wird der Puffer vorsichtig abgossen und durch 100 ml neuen Kaliumphosphatpuffer ersetzt. Erneut lässt man die

Suspension 30 min äquilibrieren. Danach wird der überschüssige Puffer vorsichtig entfernt. Der Ansatz des Adsorbens erfolgt am zweckmäßigsten vor dem Abernten der Kultur

- in jeden der 4 Ölttestbecher mit *Paramecium*-Pellet wird 2 ml Aqua_{bidest} pipettiert und das Pellet resuspendiert. In diesem Schritt ist es essentiell, Aqua_{bidest} zur Resuspension zu verwenden, da hierdurch die *Paramecium*-Zellen aufquellen und leicht aufgeschlossen werden können
- Die *Paramecium*-Suspension wird sofort in einen eiskalten 15 ml Potter-Homogenisator überführt und 1 min auf Eis homogenisiert. Es erfolgt eine Überprüfung eines Aliquots im Binokular, ob alle Zellen gründlich homogenisiert worden sind. Sollte dies nicht der Fall sein, wird so lange weiter homogenisiert, bis keine ganzen Zellen mehr zu erkennen sind
- Um zu verhindern, dass die Bakterien platzen, wird dem Homogenat 4 ml Kaliumphosphatpuffer zugegeben und erneut für 30 s homogenisiert
- Das Homogenat wird auf zwei 50 ml Corex-Zentrifugenröhren verteilt, jeweils mit Kaliumphosphatpuffer auf ~40 ml aufgefüllt und durch Schütteln der Röhren gründlich vermischt
- Es folgt eine Zentrifugation für 3 min in der Kühlzentrifuge bei 0 °C und 4500 x g. Der Überstand wird verworfen
- Die Pellets des Homogenats werden durch gründliches Vortexen in 2 ml Kaliumphosphatpuffer resuspendiert, die Röhre wiederum mit Puffer auf 40 ml aufgefüllt und die Suspension gut durchmischt
- Erneut wird wie oben zentrifugiert und der Überstand verworfen
- Diese Waschschrte werde so lange wiederholt, bis der Überstand annähernd klar ist (4 bis 5 mal)
- Das Pellet wird nun in 4 ml Kaliumphosphatpuffer gründlichst resuspendiert, zu dem vorbereiteten Adsorbens Ecteola 23 gegeben und mit ihm vermischt
- Die Ecteola/Homogenat-Suspension wird für 2 Minuten auf Eis inkubiert
- Nun wird die Suspension in einen (in einem Trichter befindlichen) Kaffeefilter gegeben (von TIP, Größe 2, auf Trichtergröße zurechtgeschnitten und mit Kaliumphosphatpuffer äquilibriert)
- Nachdem die gesamte Suspension durch den Filter geflossen ist, wird der Filter mit dem Ecteola 23 vorsichtig (damit der Filter nicht zerreißt) ausgepresst. Es wird noch zweimal je 3 ml Kaliumphosphatpuffer auf den Filter gegeben und danach vorsichtig herausgepresst. Zu groß darf der Druck nicht sein, da sich sonst schon adsorbierter Detritus vom Ecteola lösen kann und dann das Isolat kontaminiert

- Das gereinigte Isolat wird nochmals bei 0 °C und 4500 x g für 3 min zentrifugiert und der Überstand verworfen

Im Pellet befinden sich im Idealfall nur noch die Bright- und die Nonbright-Form der *Caedibacter* und während der Homogenisierung nicht aufgeschlossene Organellen aus *Paramecium tetraurelia* (z.B. Mitochondrien). Um das Isolat weiter zu fraktionieren und dadurch diese Organellen entfernen zu können, müsste eine Perkoll-Gradienten-Dichtezentrifugation durchgeführt werden. Da im weiteren Verlauf der Arbeit die übriggebliebenen Bestandteile nicht stören, kann auf eine solche Aufreinigung verzichtet werden.

Die Isolation des Endobionten zur Gesamt-RNA-Isolation läuft analog zur Isolation für die Plasmid-Gewinnung ab. Da es jedoch zu keiner Verfälschung des RNA-Zustandes in den zu isolierenden Bakterien kommen soll, muss auch hier wieder auf größte Geschwindigkeit und konstant niedrige Temperaturen geachtet werden. Alle Schritte sollten nahe 0 °C durchgeführt werden. Außerdem werden die jeweiligen, im obigen Protokoll angegebenen Zeiten, um die Hälfte verkürzt und das Pellet höchstens zweimal mit Kaliumphosphatpuffer gewaschen.

2.2.2 DNA-Isolation

2.2.2.1 Plasmid-DNA-Isolation

Alle Plasmid-Isolations-Protokolle in dieser Arbeit basieren auf dem Prinzip der "alkalischen Lyse" (BIRNBOIM UND DOLY, 1979). Dabei werden mittels NaOH und SDS die Zellwand, die Zellmembran und die restlichen Zellbestandteile in Lösung gebracht. Genomische DNA wird denaturiert und fragmentiert. Durch die Zugabe eines hochionalen Neutralisierung-Puffers und die dadurch induzierte pH-Wert-Änderung zum sauren Bereich, fallen die meisten Proteine, andere Zellbestandteile, sowie die genomische DNA, als unlöslicher Niederschlag aus. Die Plasmide renaturieren, bleiben aufgrund ihrer zirkulären Form in Lösung und können, entweder durch eine geeignete Trägermatrix (hier Silika) oder durch Phenol-Chloroform-Extraktion, weiter aufgereinigt werden. Als letzter Konzentrationsschritt erfolgt eine Alkohol-Präzipitation der Plasmid-Lösung. Für die pKAP298-Isolation aus den isolierten *Caedibacter* wird das Plasmid-Midi-Kit von Quiagen verwendet.

Sollen mit dem Isolat Southern Blots erstellt, oder das Isolat zur Subklonierung in *E. coli* verwendet werden, wird auf die Aufreinigung mittels einer Silika-Säule verzichtet, da die

Plasmid-Verluste, aufgrund der Plasmidgröße von 49 kB, zu groß sind (ungefähr 50 % Verlust). Die Isolation erfolgt nach dem Protokoll von Quiagen. Nach dem Neutralisations-Schritt wird das Isolat dann über Phenol-Chloroform-Extraktion und dreifaches Ausschütteln mit Ether von Proteinresten gereinigt. Um den restlichen Ether aus der Plasmid-Lösung zu entfernen, wird die Lösung 10 min auf 75 °C erhitzt. Da der Siedepunkt des Ethers bei 35 °C liegt, verdampft er völlig aus der Lösung. Diese Hitzebehandlung hat weiterhin den Vorteil, dass möglicherweise noch vorhandene DNAsen hitzeinaktiviert werden. Eine Aufkonzentrierung des Isolats erfolgt über Isopropanol-Fällung (gleiches Volumen, 20 min bei -20 °C, Zentrifugation 10 min bei mind. 13 000 x g) und zweifaches Waschen mit 70 % EtOH mit anschließendem Trocknen des Plasmid-Präzipitats. Das Plasmid-Pellet wird in einem geeigneten Volumen 10 mM Tris/HCl (pH 8,5) gelöst und bei -20 °C gelagert.

Die Isolierung des Plasmids pKAP 298 aus dem *Caedibacter*-Pellet zum Zwecke der Sequenzierung gestaltet sich sehr ähnlich. Da jedoch die mit der Sequenzierung beauftragte Firma Seqlab (Göttingen) mit Phenol-Chloroform extrahiertem Plasmid-Isolat bei der Sequenzierung keine Signale bekam, wird das als Template bestimmte Isolat über die im Kit enthaltene Silika-Matrix-Säule aufgereinigt und der damit einhergehende Verlust an Plasmid in Kauf genommen. Die Isolation erfolgt nach dem Quiagen Protokoll, das Plasmid-Eluat wird mittels Isopropanol-Fällung aufkonzentriert.

Isolationen rekombinanter Vektoren aus *E. coli* werden nach dem Quiagen-Protokoll durchgeführt und das Lysat, je nach Verwendungszweck, über eine Säule oder gar nicht (bei dem Screening nach Inserts) aufgereinigt.

Bei allen Isolaten werden möglicherweise noch vorhandene DNAsen für 10 min bei 75 °C inaktiviert.

Um zu kontrollieren ob die pKAP298-Isolate weitestgehend frei von unerwünschter Nicht-Plasmid-DNA ist, wird ein Aliquot (~500 ng) mit dem Restriktionsenzym *Pst* I verdaut. Der Ansatz wird in einem Agarose-Gel elektrophoretisch aufgetrennt und mit Ethidiumbromid gefärbt. Der Verdau des Plasmids ist nötig, da das Plasmid in seiner zirkulären Form teilweise nicht in das Gel hineinwandert, bzw. aufgrund seiner Größe in dem Bereich gescherter genomischer DNA liegt, und somit mögliche Verunreinigungen überdeckt. Außerdem bekommt man mit dem zirkulären Plasmid keine definierten Banden, sondern eher einen Schmier. Das Gel wird zur Dokumentation unter UV-Licht-Anregung fotografiert.

2.2.2.2 DNA-Extraktion aus Agarose-Gelen

Alle DNA-Extraktionen aus Agarose-Gelen in dieser Arbeit werden mit dem "NucleoSpin Extract 2 in 1"-Kit der Firma Machery&Nagel durchgeführt. Das Protokoll basiert auf dem originalen Protokoll der Firma M&N, wurde aber, aufgrund der Größe der zu extrahierenden Banden, etwas modifiziert.

- Auf dem UV-Tisch wird die zu extrahierende Bande möglichst exakt aus dem Gel ausgeschnitten und in ein 2 ml Tube überführt, dessen Gewicht vorher bestimmt worden ist
- Das Tube wird erneut gewogen und aus der Gewichts Differenz das Gewicht des Agarose-Stücks errechnet
- Pro 100 mg Agarose werden 300 µl Puffer NT1 hinzugegeben
- Das Tube wird für 10 min im Wasserbad bei 55 °C inkubiert und alle 2 min kurz gevortext. Die Inkubation ist zu verlängern, falls sich das Agarose-Stück nicht vollständig im Puffer gelöst hat
- Die Lösung wird in eine Aufreinigungs-Säule pipettiert und 1 min bei 4000 rpm in der Tischzentrifuge zentrifugiert
- Der Durchfluss wird erneut auf die Säule gegeben und wieder für 1 min, diesmal bei 12 000 rpm, zentrifugiert
- Der Durchfluss wird verworfen und 600 µl Waschpuffer NT3 in die Säule pipettiert
- Es folgt eine Zentrifugation für 1 min bei 12 000 rpm
- Nachdem der Durchfluss verworfen worden ist, wird der Waschschrift mit 300 µl Puffer NT3 wiederholt und für 2 min bei 12 000 rpm zentrifugiert. Der Puffer mit dem darin enthaltenen Ethanol muss vollständig aus der Säule entfernt sein, damit die Elution nicht inhibiert wird
- Die Säule wird auf ein 1,5 ml Tube gesteckt und die Hälfte des gewünschten Elutionsvolumens an Elutionspuffer hinzupipettiert (bei 30 µl Endvolumen dementsprechend 15 µl). Der Elutionspuffer ist 10 mM Tris-HCl (pH 8,5)
- Die Säule wird mit Parafilm verschlossen und 5 min bei 50 °C inkubiert
- Nun wird 1 min bei 12 000 rpm zentrifugiert und auf die Säule die restliche Hälfte des Elutions-Puffers pipettiert
- Erneut wird die mit Parafilm verschlossene Säule 5 min bei 50 °C inkubiert und danach 2 min bei 12 000 rpm zentrifugiert. Die lange Inkubation bei 50 °C soll die Elutions-Ausbeute steigern, da manche der zu extrahierenden Fragmente relativ groß sind (bis zu 10 kB), und diese daher stark an das Säulenmaterial binden, was die Elution bei Raumtemperatur ineffektiv macht.

- Das DNA-Eluat wird 10 min auf 75 °C erhitzt, um mögliche DNAsen zu inaktivieren

Als letzten Schritt der DNA-Extraktion wird noch die DNA-Konzentration des Eluats bestimmt und die Lösung bis zur weiteren Verwendung bei - 20 °C gelagert.

2.2.3 Isolation von Gesamt-RNA aus *C. taeniospiralis*

Alle in dieser Arbeit durchgeführten RNA-Isolationen werden nach der "optimierten Guanidinisothiocyanat/Phenol-Methode zur Extraktion von RNA", mittels des Reagenz "pegGOLD RNAPure™" der Firma Peqlab, durchgeführt. Es wird das von Peqlab mitgelieferte Protokoll verwendet. Was das Protokoll der Firma Peqlab verschweigt ist, dass sich im Gesamt-RNA-Isolat auch Plasmide befinden, die bei dem Verfahren nicht entfernt werden. In diesem Fall ist es das in dieser Arbeit zu untersuchende Plasmid pKAP298. Da aus dem Isolat markierte Sonden hergestellt werden, die dann auf das geblottete Plasmid pKAP298 hybridisiert werden sollen, ist es fatal, wenn bei der Sondenherstellung schon das Plasmid vorhanden ist - alle Banden des Southern Blots würden Signale ergeben. Die Plasmid-Kontamination des RNA-Isolats muss daher vollständig, durch intensiven DNaseI-Verdau im Anschluss an die Isolation, beseitigt werden.

Nach der Isolation wird die Konzentration des RNA-Isolats bestimmt. Ganz exakt ist die ermittelte Konzentration nicht, da sich noch doppelsträngige DNA-Kontaminationen (unter anderem das Plasmid pKAP298) in dem Isolat befinden, und diese bei 260 nm stärker absorbieren als einzelsträngige RNA. Der Wert ist aber eine gute Näherung und dient als Basis zur Berechnung der zum DNA-Verdau nötigen DNaseI. Um die DNA-Kontaminationen aus dem Isolat zu entfernen, werden ihm pro gemessenem µg RNA 10 Units DNaseI (RNasefrei) zugegeben und das Isolat für 30 Minuten bei 37 °C inkubiert. Es werden wiederum Aliquots zur Konzentrationsbestimmung und zur Dokumentation mittels Agarose-Gelelektrophorese abgezweigt und die RNA-Konzentration erneut gemessen.

Die für die Dokumentation vorgesehenen Aliquots werden auf ein 1,5 % Agarose-Gel aufgetragen und mit 7,5 V/cm Spannung elektrophoretisch getrennt. Die Spannung muss so hoch sein, damit sich die Auftrennung schnell vollzieht und so die RNA wenig abgebaut wird. Es werden dem Gel keine denaturierenden Agenzien beigemischt. Anschließend wird das Gel fotografisch dokumentiert. Das kontaminationsfreie RNA-Isolat kann nun bei - 70 °C aufbewahrt werden.

Auf eine Inaktivierung der DNaseI wird verzichtet, da - falls sich zeigen sollte, dass noch DNA-Kontaminationen vorhanden sind - das Isolat durch Auftauen und anschließendes Inkubieren bei 37 °C jederzeit "nachverdaut" werden kann.

2.2.4 Umschreiben der Gesamt-RNA in cDNA

Zum Umschreiben der isolierten Gesamt-RNA in cDNA wird ausschließlich das "RevertAid H Minus First Strand cDNA Synthesis"-Kit der Firma MBI Fermentas benutzt. Die in diesem Kit verwendete reverse Transkriptase „H Minus M-MuLV“ hat den Vorteil, dass sie den RNA-Strang während der Transkription nicht abbaut, so dass er als Template für weitere Transkriptionen dienen kann. Somit ist die Ausbeute an cDNA höher als bei der Verwendung einer Transkriptase mit ribonuclease-Aktivität. Ein selektives Umschreiben von mRNA ist natürlich nicht möglich, da prokaryotische mRNA keinen Poly-A-Schwanz hat, an dem Oligo-dT-Primer binden könnten. Die Transkription findet daher unselektiv mit Random-Hexamer-Primern statt.

In ein 100 µl PCR-Reaktionsgefäß wird so viel RNA-Isolat pipettiert, dass sich in dem Gefäß 1 µg RNA befindet (höchstens jedoch 10 µl) und mit Aqua_{bidest} auf 10 µl aufgefüllt. Das Tube wird im Thermocycler (mit Heizdeckel) für 5 min auf 95 °C erhitzt, um die sich noch im Isolat befindliche DNaseI zu inaktivieren. Das Reaktionsgefäß wird kurz auf Eis gestellt, 1 µl der Random-Hexamer-Lösung hinzugegeben und gründlich vermischt. Erneut wird 5 min bei 95 °C im Thermocycler inkubiert, um mögliche Sekundärstrukturen der RNA zu denaturieren und somit eine gleichmäßige Anlagerung der Random-Hexamere zu ermöglichen. Das Tube wird auf Eis gestellt und 4 µl des 5 x Reaktionspuffers, 2 µl 10 mM dNTP-Lösung und 1 µl RNase-Inhibitor hinzu pipettiert. Es folgt eine Inkubation für 5 min bei 25 °C. In die Reaktionslösung werden 1 µl der reversen Transkriptase „H Minus M-MuLV“ pipettiert und gründlich vermischt. Das Tube wird 10 min bei 25 °C gehalten, um die Anlagerung der Random-Hexamere zu gewährleisten. Die eigentliche reverse Transkription findet für 80 min bei 42 °C im Thermocycler mit Heizdeckel statt. Um die reverse Transkriptase zu inaktivieren, wird der Ansatz für 10 min auf 75 °C erhitzt.

Auf eine Aufreinigung der entstandenen cDNA (über eine Säule oder durch alkoholische Präzipitation) wird, wegen des möglichen Verlustes an cDNA, verzichtet. Das für die Markierung verwendete Klenow-Enzym arbeitet auch in dem für die reverse Transkription verwendeten Puffer. Die fehlende Aufreinigung hat allerdings den Nachteil, dass die Konzentrationsbestimmung des Transkripts durch optische Dichte-Messung unmöglich wird,

da nicht eingebaute dNTPs und Random-Hexamere, sowie die nicht abgebaute RNA, das Messergebnis verfälschen. Die Aufbewahrung des Transkripts erfolgt bei - 20 °C.

2.2.5 Modifikation von DNA

2.2.5.1 Restriktion

Alle DNA-Restriktionen mit einer einzigen Endonuclease werden in den geeigneten Puffersystemen durchgeführt. Wenn DNA einem Restriktionsverdau mit mehreren Enzymen unterzogen werden soll, so ist die unterschiedliche Aktivität der Restriktionsenzyme zu beachten. Der Puffer wird so gewählt, dass das langsamste Enzym optimal arbeitet, und die anderen Enzyme gleichzeitig nicht zu stark gehemmt werden. Auch die Anzahl der Units pro Enzym kann variiert werden, um einen möglichst schnellen Verdau zu erreichen. Des weiteren gilt, dass der Komplettdau mit steigender Anzahl von Schnittstellen länger dauert. Aus diesen Faktoren, und bei Kenntnis der Konzentration der zu schneidenden DNA, kann die für den Verdau nötige Zeit abgeschätzt werden. Bei einem Multi-Enzym-Verdau zur elektrophoretischen Auftrennung ist darauf zu achten, dass das Endvolumen des Ansatzes plus Loading Dye nicht das Volumen des Gel-Slots übersteigt.

Allgemeiner Ansatz eines 3 fach Multi-Enzymverdau für den Slot eines großen Gels (100 ml Gel, 40 µl Slot):

- 30 µl Substrat DNA
 - x µl Enzym 1
 - y µl Enzym 2
 - z µl Enzym 3
 - (6-x-y-z) µl Aqua_{bidest}
- 36 µl Gesamtansatz (+ 4 µl Loading-Dye)

Der Reaktionsansatz wird für die geschätzte, zum Komplettdau nötige Zeit bei 37 °C inkubiert. Ein Abstoppen der Reaktion erfolgt durch 10 min Erhitzen auf 75 °C. Sollte eine Aufreinigung des Ansatzes notwendig sein, so wird diese Aufreinigung mittels des "NucleoSpin Extract 2in1"-Kits der Firma Machery&Nagel durchgeführt. Um bei der Vektor-Restriktion einen hohen Transformanten-Hintergrund durch ungeschnittenen Vektor zu vermeiden, wird der vollständige Verdau durch Gel-Elektrophorese und durch Transformation des unligierten Vektors überprüft.

2.2.5.2 Dephosphorylierung von Vektor-DNA

Wenn ein Plasmid-Vektor nur mit einem Restriktionsenzym geschnitten wird, ist es ratsam, den Vektor mit CIAP zu dephosphorylieren, um eine Religierung des Vektors während der Ligation, und damit einen hohen Transformanten-Hintergrund, zu verhindern. Auf eine Aufreinigung des verdauten Vektors vor der Dephosphorylierung wird, wegen des möglichen DNA-Verlustes, verzichtet. Die alkalische Phosphatase arbeitet in allen zum Verdau verwendeten Puffern. Auf eine genaue Berechnung des pro Ansatz zu verwendenden Enzyms wird verzichtet. Da das Enzym später ohnehin durch einen Reinigungs-Schritt komplett entfernt wird, kann es im Überschuss zugegeben werden.

- pro Vektor-Ansatz werden 0,2 µl CIAP hinzupipettiert und gründlich vermischt
- das Tube wird 30 min bei 37 °C inkubiert, danach für 15 min bei 56 °C
- es werden dem Ansatz nochmals 0,2 µl CIAP hinzugegeben und 45 min bei 37 °C inkubiert

Der verdaute Vektor wird nun mit dem "NucleoSpin Extract 2in1"-Kit von den verwendeten Enzymen gereinigt. Nach der Elution mit dem geeigneten Volumen Elutions-Puffers wird das Eluat 15 min auf 75 °C erhitzt, um eventuell noch vorhandenes aktives Enzym sicher zu inaktivieren. Als letzten Schritt wird die DNA-Konzentration des Eluats bestimmt und der verdaute Vektor bis zur Weiterverwendung bei - 20 °C gelagert.

2.2.5.3 Ligation

Die Ligations-Reaktionen werden mit einem molaren Vektor/Insert Verhältnis von idealerweise 1:1 durchgeführt. Der Ligationsansatz wird nach folgender Formel berechnet:

$$\frac{Vektor(ng) \times Insert(kb)}{Vektor(kb)} \times \frac{Insert}{Vektor} = Insert(ng)$$

Standard-Ligations-Ansatz:

x µl	Plasmid-Vektor DNA Lösung
y µl	Insert DNA-Lösung
z µl	10x T ₄ Ligase Puffer
<u>0,5 µl</u>	<u>T₄ Ligase</u>
Σ=x+y+z+0,5 µl Gesamtansatz	

Die Ligations-Ansätze werden für 2 Stunden bei Raumtemperatur und danach über Nacht bei 4 °C inkubiert. Die Reaktion wird durch 10 min Erhitzen auf 75 °C abgestoppt. Anschließend werden die Ansätze entweder direkt zur Transformation eingesetzt oder bei - 20 °C gelagert. Auf eine Aufreinigung des Ligations-Ansatzes zur Entfernung der Puffer-Ionen wird (wegen möglichem DNA-Verlust) verzichtet. Wenn die kompetenten Zellen sorgfältig hergestellt worden sind, d.h. wenn sie eine geringe Leitfähigkeit aufweisen, kommt, trotz der im Ligations-Ansatz vorhandenen Ionen, eine für eine effiziente Transformation ausreichende Puls-Länge zustande.

2.2.6 Transformation von *Escherichia coli*

2.2.6.1 Herstellung elektrokompetenter Zellen

Die Herstellung kompetenter Zellen erfolgt nach dem Protokoll, das in der Bedienungsanleitung zum *E. coli* Pulser™ beschrieben ist. 1 Liter LB-Medium wird mit 1/100 Volumen einer *E. coli* Übernachtskultur (Stamm JM109 oder TOP10) beimpft. Unter Schütteln (ca. 200 rpm) werden die Zellen bis zu einer ABS_{600} von 0,7 wachsen gelassen. Die Zellen werden dann in einer Kühlzentrifuge bei 0 °C und 4000 g für 15 min zentrifugiert. Danach wird das Pellet viermal in 250 ml eiskaltem 10 % (v/v) Glycerol gewaschen und abzentrifugiert. Die Zellen werden schließlich in 4 ml 10 % Glycerol-Lösung resuspendiert, in 80 µl Aliquots aufgeteilt und im kalten Ethanolbad bei - 70 °C schockgefroren. Die Lagerung erfolgt ebenfalls bei - 70 °C.

2.2.6.2 Transformation durch Elektroporation

Die Transformation durch Elektroporation erfolgt mittels des *E. coli* Pulser™ (Bio-Rad, München). Ein 80 µl Aliquot kompetenter Zellen wird aufgetaut und auf Eis gestellt. Der Zellsuspension wird 1 µl Ligationsansatz zugegeben, gut vermischt und ca. 1 Minute auf Eis inkubiert. Die Zell/DNA-Suspension wird in eine sterile 80 µl Elektroporations-Küvette pipettiert und 5 min auf Eis inkubiert. Der Pulser wird auf 1800 Volt eingestellt und die Küvette gepulst. 10 s nach dem Puls wird 500 µl LB-Medium (alternativ auch SOC-Medium) in die Küvette pipettiert und gründlich mit der Zellsuspension vermischt. Der Ansatz wird in ein 2 ml Tube überführt und für 45 min bei 37 °C und Schütteln (~200 rpm) inkubiert. Zum Abschluss wird am Pulser noch die Pulsdauer abgelesen. Sie sollte bei etwa 3,5-5 ms liegen. Nach der Inkubation wird der Transformationsansatz in Aliquots von 100 µl bis 200 µl auf

selektive LB-Agarplatten ausplattiert und mindestens 20 h bei 37 °C inkubiert. Die LB-Platten enthalten 50 µg/ml Kanamycin, 80 µg/ml X-Gal und 0,5 mM IPTG.

2.2.6.3 Identifizierung und Konservierung von Transformanten

Rekombinante *E. coli*-Klone werden zunächst über Blue/White-Screening, dann über Kolonie-Blot-Hybridisierung und Restriktions-Verdau des isolierten rekombinanten Plasmids, nach den erwünschten Inserts gescreent. Klone mit erwünschten Inserts werden in Aliquots, als Suspension in 15 % (v/v) Glycerol, bei - 70 °C konserviert. Zusätzlich wird das dazugehörige, rekombinante Plasmid-Isolat ebenfalls bei - 70 °C konserviert.

2.2.7 DNA/RNA-Konzentrationsbestimmung über die optische Dichte

Die Reinheit und Konzentration der isolierten DNA und RNA wird durch eine spektralphotometrische Messung der optischen Dichte bestimmt. Hierfür wird ein Gene Quant II RNA/DNA Calculator verwendet. Eine O.D.₂₆₀ von 1 entspricht 50 µg dsDNA oder 40 µg ssDNA bzw. RNA pro ml (MANIATIS ET AL., 1989). Der Quotient aus der optischen Dichte der Probe bei 260 nm und bei 280 nm ist ein Maß für die Reinheit der DNA-Probe. Das Verhältnis (Ratio) der O.D.₂₆₀/O.D.₂₈₀ liegt bei reiner DNA bei 1,8, bei reiner RNA bei 2,0 (MANIATIS ET AL., 1989). Zur Messung der O.D. bei den oben angegebenen Wellenlängen wird eine Quarzküvette mit einer Schichtdicke von 5 mm verwendet. Alle DNA- bzw. RNA-Konzentrationsbestimmungen in dieser Arbeit werden nach diesem Verfahren durchgeführt. Als Referenz dient das jeweilige reine Lösungsmittel der DNA/RNA-Proben. Es ist genau darauf zu achten, dass die zu messenden Proben keine Restkonzentrationen an Ethanol oder Phenol enthalten, da diese Stoffe Licht in einem ähnlichen Wellenbereich wie Nucleinsäuren absorbieren, was die Messung in Richtung zu hoher Konzentration verfälscht. Des weiteren sind erst Messungen ab 20 ng/µl als relativ genau anzusehen, da das Gerät eine Messungsgenauigkeit von ~2 ng/µl hat (empirisch durch mehrmalige Referenzwert-Messung festgestellt).

Zur Bestimmung von RNA-Konzentrationen wird die Absorbtion der Lösung bei 260 nm gemessen und die Konzentration wie folgt berechnet:

$$\text{RNA-Konzentration} = \text{Absorbtion}_{260} \times 40 \times 2 \text{ (Schichtdicke der Küvette nur 5 mm)}$$

2.2.8 Agarose-Gel-Elektrophorese

Die Auftrennung von DNA-Fragmenten erfolgt in 1 x TAE als Laufpuffer. Die Konzentration der Gele beträgt, je nach Größe der zu trennenden Fragmente, zwischen 0,7 % und 1,5 %. Auch die angelegte Spannung (zwischen 2 und 7,5 V/cm) variiert je nach Fragmentgröße. Sollen große Fragmente getrennt werden, wird eine niedrige Agarose-Konzentration in Kombination mit einer niedrigen Spannung gewählt, bei kleinen Fragmenten umgekehrt. Bei Ansätzen mit sowohl großen als auch kleinen Fragmenten (z.B. geschnittenes pKAP298-Isolat), wird eine relativ hohe Gel-Konzentration (bis zu 1,4 %) und eine relativ niedrige Spannung gewählt. Die jeweiligen Proben werden mit 1/6 Volumen Loading-Dye versetzt und in die Gel-Slots pipettiert. Als Längenstandards dienen die in 2.1.7.2 genannten Molekulargewichtsmarker. Das Agarose-Gel wird nach der Elektrophorese für 20 Minuten in einer Lösung aus Aqua_{bidest} mit einigen Tropfen Ethidiumbromid gefärbt. Nach dem Färben wird das Gel für ca. 20 Minuten in Aqua_{bidest} entfärbt. Anschließend wird das entfärbte Gel auf dem UV-Lichtkasten betrachtet und gegebenenfalls fotografiert.

2.2.9 Hybridisierungen

2.2.9.1 Southern Blots

Der Transfer von DNA-Fragmenten (geschnittenes pKAP298 oder in *E. coli* subklonierte pKAP-Fragmente) erfolgt mittels des Kapillar-Verfahrens auf Hybond-Nylon-Membranen der Firma Amersham Pharmacia Biotech Europe GmbH (Freiburg). Das Agarose-Gel wird zunächst 20 min in 0,25 M HCl depuriniert, dann 20 min denaturiert (0,5 M NaOH, 3 M NaCl) und schließlich 30 min neutralisiert (0,5 M Tris-HCl, 1,5 M NaCl, pH 7). Als Laufpuffer dient 10 x SSC. Eine (Glas-)Platte wird auf ein Reservoir mit dem Laufpuffer platziert. Auf die Platte wird ein Stück 3 mm Whatman-Filterpapier (mit der doppelten Breite des zu blottenden Agarose-Geles) gelegt, so dass beide Enden des Filterpapiers in das Reservoir eintauchen. Das Agarose-Gel wird mit der Oberseite nach unten auf das Papier gelegt und eventuell vorhandene Luftblasen ausgestrichen. Auf das Gel kommt ein genauso großes Stück der Nylon-Membran. Luftblasen werden wieder durch Ausstreichen entfernt. Auf die Nylonmembran werden 2 Lagen 3 mm Whatman-Filterpapier gleicher Größe gelegt und ein wenig mit Laufpuffer befeuchtet, damit sie der Membran eng anliegen. Nun werden die aus dem Gel überstehenden Ränder des untersten Filterpapiers mit Parafilm bedeckt, so dass dort später kein Puffer durchfließen kann und der gesamte Fluss durch das Agarose-Gel geht. Auf den Blot werden mittig ~50 Lagen Papierhandtücher gelegt. Die Papierhandtücher werden mit einer Platte bedeckt und diese mit Bleigewichten beschwert. Zunächst sollte das Gewicht nicht zu schwer sein, da sonst das Agarose-Gel zerquetscht werden könnte. Nach 1 h

kann dann das endgültige Gewicht aufgelegt werden. Der Blot-Vorgang dauert mindestens 24 h. Die Membran mit der übertragenen DNA wird mit 200 mJ/cm² im UV-Licht fixiert, in 10 x SSC gewaschen und in 2 x SSC bis zur Weiterverwendung bei 4 °C aufbewahrt. Das Agarose-Gel kann durch Färbung mit Ethidiumbromid auf die Vollständigkeit der Übertragung kontrolliert werden

2.2.9.2 Kolonie-Blots zum Transformanten-Screening

Alle auf Inserts zu testende rekombinante Klone werden auf eine selektive LB-Agar-Platte überimpft, nummeriert und 24 h bei 37 °C inkubiert. Von dieser Platte werden Teile der gewachsenen Kolonien auf eine Nylonmembran (Hybond der Firma Amersham Pharmacia Biotech Europe GmbH) aufgetragen und mit der gleichen Nummerierung versehen. Mit der Kolonie-Seite nach oben wird die Membran auf ein 3 mm Whatman-Filterpapier gelegt, das vorher in Lysis-Puffer QP2 (aus dem Qiagen Plasmid Midi Kit) getränkt worden ist. Die Membran wird 10 min bei Raumtemperatur inkubiert. Danach wird die Membran mit den lysierten Bakterien auf ein weiteres Filterpapier transferiert, das vorher in Neutralisierungspuffer QP3 getränkt worden ist. Es folgt eine Inkubation für 10 min bei Raumtemperatur. Die Membran wird nun, unter fließendem Wasser und vorsichtigem Reiben, sorgfältig von Bakterien-Resten gesäubert. Fixiert wird die Plasmid-DNA im UV-Licht mit einer Strahlendosis von 200 mJ/cm². Die Membran wird in 10 x SSC gewaschen und in 2 x SSC bis zur Weiterverwendung bei 4 °C aufbewahrt.

2.2.9.3 Nicht-radioaktive DNA-Markierung mit Digoxigenin-11-dUTP

Für die Markierung von DNA wird das "DIG High Prime DNA Labeling and Detection Starter Kit II" der Firma Roche (Mannheim) verwendet. Im Vergleich zur radioaktiven Isotopen-Markierung hat das Kit den Vorteil, dass es auch außerhalb eines Isotopenlabors angewendet werden kann. Außerdem bedingt das Prinzip der Markierung mit Digoxigenin eine Verstärkung des Signals, da der an das Digoxigenin bindende Antikörper eine enzymatische Reaktion katalysiert und die Signalstärke somit vom während der Detektion im Überschuss vorhandenen Substrat abhängt. Da präparationsbedingt nur sehr wenig erwünschte mRNA (*Caedibacter taeniospiralis* ist Prokaryot) und schließlich markierte cDNA-Sonde zu erwarten ist, ist eine Verstärkung des Signals essentiell. Bei einer Isotopen-Markierung wäre eine Detektion wahrscheinlich aussichtslos.

2.2.9.3.1 Markieren der cDNA aus Gesamt-RNA-Isolaten

Alle cDNA-Sonden-Markierungen werden mit dem "DIG High Prime DNA Labeling and Detection Starter Kit II" der Firma Roche (Mannheim) durchgeführt. Weil das Kit extrem teuer ist, wird zunächst von jeder Markierungs-Reaktion nur ein viertel Ansatz gemacht, und erst wenn bei der späteren Detektion deutliche Banden zu sehen sind (also alle bisherigen Schritte, von der RNA-Isolation über den DNaseI-Verdau bis hin zur reversen Transkription, funktioniert haben), der gesamte cDNA-Ansatz markiert. Zunächst werden 4 µl des zu markierenden cDNA-Transkripts im Thermocycler (mit Heizdeckel) für 10 min auf 99 °C erhitzt, um die cDNA vollständig zu denaturieren. Das Tube wird sofort für 2 min in Eiswasser gestellt. Es wird nun 1 µl der "DIG-High-Prime"-Reaktions-Lösung hinzu pipettiert und gründlich vermischt. Das Reaktionsgefäß wird für 20 Stunden bei 37 °C inkubiert und anschließend für 10 min auf 75 °C erhitzt, um die Enzym-Reaktion zu stoppen. Mit diesem viertel Ansatz wird nun hybridisiert und detektiert. Wenn Banden zu sehen sind (allerdings nicht ALLE, siehe unten), wird der restliche cDNA-Ansatz markiert. Der Ablauf ist identisch, nur dass 16 µl cDNA-Lösung und 4 µl "DIG-High-Prime"-Reaktions-Lösung benutzt werden. Wenn die Detektion nicht das gewünschte Ergebnis bringt, muss wieder an dem Schritt angefangen werden, der vermutlich nicht funktioniert hat. Ergeben z.B. auf dem detektierten Blot alle pKAP298-Banden positive Signale, hat der DNaseI-Verdau des Gesamt-RNA-Isolats nicht funktioniert, und das RNA-Isolat (2.3.3) muss "nachverdaut" werden. Die markierte Sonde wird bis zur Weiterverwendung bei - 20 °C aufbewahrt.

2.2.9.3.2 Markieren von pKAP298-DNA zur Insert-Detektion

Die DigOx-Markierung von pKAP298 läuft analog zur Markierung der cDNA ab, nur dass gleich ein gesamter Ansatz markiert wird (4 µl "DIG-High-Prime"-Reaktions-Lösung und 16 µl pKAP298-Isolat \cong 1µg). Der Erfolg der Markierung wird durch Hybridisierung der Sonde an pKAP298-DNA überprüft.

2.2.9.4 Hybridisierung der Membranen mit der Sonden-DNA

Als Hybridisierungs-Temperatur wird durchgehend 39 °C gewählt. Sie erweist sich als guter Kompromiss zwischen geringer Signalstärke bei höheren Temperaturen und der vermehrt auftretenden, unspezifischen Blot-Sonden-Bindung bei niedrigeren Temperaturen. Der zu hybridisierende Southern/Kolonie Blot wird in ein 50 ml Falcon-Tube überführt, 10 ml der "DIG Easy Hyb"-Lösung hinzugegeben und das Tube für 30 min bei 39 °C unter leichtem Schütteln (~150 rpm) im Wärmeschrank prähybridisiert. Die markierte DNA-Sonde wird mit

Aqua_{bidest} auf 100 µl aufgefüllt und im Thermocycler für 10 min auf 99 °C erhitzt, um eine vollständige Denaturierung zu erreichen, und danach sofort auf Eis gestellt. 15 ml der "DIG Easy Hyb"-Lösung werden in ein 50 ml Falcon-Tube gegeben und das Tube im Wasserbad auf 68 °C erhitzt. Die denaturierte Sonde wird in das Falcon-Tube pipettiert und im Wasserbad 10 min bei 68 °C inkubiert. Somit bleibt die Sonde einzelsträngig. Die prähybridisierte Membran wird in das noch 68 °C warme Falcon-Tube mit der denaturierten Sonde gesteckt (DNA-Seite nach innen) und das Falcon-Tube geschlossen. Um den Deckel des Tubes werden noch 2 Lagen Textil-Klebeband gewickelt, um einen Verlust der Hybridisierungslösung während der Inkubation zu verhindern. Das Tube wird für 24 h unter leichtem Schütteln (~150 rpm) bei 39 °C im Wärmeschrank inkubiert.

Bei der Hybridisierung von Southern Blots wird die Membran nach 24 h aus der Hybridisierungslösung in 39 °C warme "DIG Easy Hyb"-Lösung überführt und 10 min bei 39 °C gehalten. Gleichzeitig wird die Hybridisierungslösung für 10 min auf 68 °C erhitzt, um die möglicherweise renaturierte Sonde wieder einzelsträngig zu machen. Die Southern Blot-Membran wird schließlich wieder in die Hybridisierungslösung gegeben und erneut unter leichtem Schütteln 24 h bei 39 °C inkubiert.

Nach der Hybridisierung wird die jeweilige Membran in ein 50 ml Tube mit 20 ml Post-Hybridisierungslösung überführt und 15 min bei Raumtemperatur auf einem Rolltisch inkubiert. Eine zweite Stringenzwaschung findet bei 65 °C für 20 min in 20 ml 0,5 x SSC mit 0,5 % (w/v) SDS statt.

2.2.9.5 Detektion der hybridisierten Membranen

Alle verwendeten Puffer-Lösungen müssen Raumtemperatur haben. Man kann die angesetzten Puffer für mehrere Detektionen verwenden, sie müssen dann allerdings im Kühlschrank gelagert und bei Bedarf im Wasserbad auf Raumtemperatur erhitzt werden. Die Detektion findet wieder in 50 ml Falcon-Tubes statt, die einzelnen Inkubationen der Tubes werden auf einem Rolltisch durchgeführt. Die hybridisierte und stringenz-gewaschene Membran wird für 15 min in 25 ml Waschpuffer gewaschen. Es folgt eine 30 min Inkubation in 40 ml Blocking-Solution. Nun wird die Membran für 30 min in 10 ml Antibody-Solution gegeben. Die Membran wird 30 min in 40 ml Waschpuffer gewaschen. Vor der Inkubation auf dem Rolltisch sollte das Tube geschüttelt werden, damit der Waschpuffer auch zwischen Tubewand und Membran dringen kann. Der Waschschrift wird wiederholt. Die gewaschene Membran wird für 5 min in 40 ml Detektionspuffer äquilibriert. Aus einem Plastik-Autoklavierbeutel wird eine Tasche geschnitten, in welche die Membran mit der DNA-Seite

nach oben gelegt wird. Auf die DNA-Seite der Membran wird gleichmäßig 500µl der "CSPD ready-to-use"-Lösung verteilt und die Oberseite der Plastik-Tasche auf den Blot zurückgelegt. Die Tasche mit dem Blot wird nun an 3 Seiten mit einem Folienschweißgerät verschweißt und 10 min bei Raumtemperatur inkubiert. Vorsichtig wird nun die überschüssige CSPD-Lösung aus der noch offenen Taschenseite gedrückt und diese Seite ebenfalls verschweißt. Es ist darauf zu achten, dass keine CSPD-Lösung auf die Außenseite der Blot-Tasche gelangt, da dies später zu einer Schwärzung des Filmes an der betreffenden Stelle führt. Der verschweißte Blot wird für 10 min bei 37 °C inkubiert. In der Dunkelkammer wird auf die DNA-Seite der Membran ein Stück zurechtgeschnittener Film gelegt (Kodak BioMAX Light, mit der beschichteten Seite in Richtung Membran) und für 10-12 h detektiert (Kolonie-Blots aufgrund der größeren Signalstärke nur 1-2 h). Falls die Banden (bei den Southern Blots) zu schwache Signale liefern, wird nochmals detektiert, diesmal dementsprechend länger (bis zu 24 h). Der Film wird 5 min im Entwicklungsbad entwickelt, 30 s gewässert und 10 min im Fixierbad fixiert. Danach wird er 2 min unter fließendem Wasser gespült und getrocknet. Nach der Detektion wird die Membran in einem 50 ml Tube mit 40 ml Stripping-Solution für 30 min bei 37 °C unter leichtem Schütteln inkubiert, um die Sonden-DNA vom Blot zu waschen und das Digoxigenin zu zerstören. Die "gestrippte" Membran wird zweimal 15 min in 2 x SSC gewaschen und kann dann mehrmals wiederverwendet werden.

2.2.10 Analyse der Southern Blots anhand der pKAP-Sequenz, Aktivitätsanalyse

2.2.10.1 Zuordnung der Blot-Gel-Banden zu den entsprechenden pKAP298-

Fragmenten

Die Vorgehensweise zur Banden-Zuordnung ist für alle Southern Blots identisch, unabhängig davon woher die geblottete DNA stammt (original pKAP298-DNA oder amplifizierte subklonierte Fragmente). Zunächst werden, anhand der durch Sequenzierung ermittelten pKAP298 Basensequenz und deren Restriktionskarte, die Banden der Fotos der zum Blotten verwendeten Gele den entsprechenden Fragmenten des Plasmids pKAP298 zugeordnet.

2.2.10.2 Zuordnung der positiven Signal-Banden zu den entsprechenden pKAP298-

Fragmenten

Danach werden die positiven Signal-Banden der detektierten Southern Blots den entsprechenden Banden der Gel-Fotos zugeordnet und mit Hilfe der Sequenzdaten in eine entsprechende Restriktionskarte des Plasmids pKAP298 eingetragen. Signale werden bei Southern Blots, die mit in *E. coli* amplifizierten pKAP-Fragmenten erstellt wurden, als positiv

gewertet, wenn die Bande in mehreren detektierten Blots Signale ergibt und die Signale signifikant stärker als der unspezifische Background sind. Bei Southern Blots, die mit dem verdauten Gesamt-Plasmid pKAP298 erstellt wurden, gelten die gleichen Kriterien, mit der Ergänzung, dass, wenn positive Signale auf sich überlagernden Fragmentbanden detektiert werden, alle potentiell für das Signal verantwortlichen Fragmentbanden als unsicher gewertet werden.

2.2.10.3 Erstellung einer Gesamt-Transkriptions-Aktivitäts-Karte des Plasmids pKAP298

Die verschiedenen Ergebnisse der Auswertung der Southern Blots werden zusammengefügt und eine Gesamt-Transkriptions-Aktivitäts-Karte des Plasmids pKAP298 erstellt. Dabei gehen von den mit dem ganzen Plasmid erstellten Blots nur die positiven Signal-Banden in die Gesamtauswertung mit ein, die auf pKAP298-Fragmenten lokalisiert sind, welche nicht subkloniert werden konnten. Der Rest der Gesamt-Transkriptions-Aktivitäts-Karte setzt sich aus den Ergebnissen der Auswertung der Southern Blots zusammen, die mit in *E. coli* subklonierten pKAP-Fragmenten erstellt wurden. Der Grund dafür ist, dass bei diesen Blots positive Signale eindeutig den entsprechenden Fragmenten auf dem Plasmid pKAP298 zugeordnet werden können, was bei den Gesamt-Plasmid-Blots, durch Überlagerung beim Verdau entstehender gleich großer pKAP298-Fragmente auf den Blots, oftmals nicht möglich ist.

2.2.11 Sequenzierung des Plasmids pKAP298

Nur ein kleiner Teil der Sequenzierung des Plasmids wurde in der Abteilung durchgeführt. Nachdem die Kosten und der Zeitaufwand ausgerechnet wurden, kamen wir zu dem Ergebnis, dass die Sequenzierung außerhalb erstens wesentlich günstiger und zweitens natürlich viel weniger zeitaufwändig ist. Deshalb wurde der Großteil der Sequenzierung von der Firma Seqlab (Göttingen) durchgeführt.

2.2.11.1 Eigene Sequenzierung

Die Sequenzierung in der Abteilung wird unter der tatkräftigen Mithilfe von Herrn H.W. Breiner durchgeführt. Da die zu sequenzierenden Fragmente sich in einen Vektor mit M13universal-Primer-Bindungsstellen befinden, werden zur Sequenzierung fluoreszenzmarkierte M13universal-Primer benutzt. Sequenziert wird mittels der Kettenabbruch-Methode nach SANGER ET AL. (1977). Für die Reaktion wird der "Thermo

Sequenase Fluorescence Labeled Primer Cycle Sequencing Kit" der Firma Amersham Pharmacia Biotech Europe GmbH (Freiburg) benutzt. Pro zu sequenzierendem Fragment werden 2 Forward- und 2 Reverse- Ansätze, nach dem Protokoll des Herstellers, erstellt und in einem Thermocycler mit Heizdeckel folgendem Programm unterzogen:

- 2 min bei 95 °C
- 15 s bei 95 °C
- 15 s bei 57 °C (Annealing-Temperatur der Primer)
- 15 s bei 70 °C (Elongation)

die Schritte 2 bis 4 werden 30 mal wiederholt.

Danach werden jedem Tube 6 µl Stop-Solution zugegeben und die Tubes bis zum Auftragen auf das Sequenziergel bei 4 °C aufbewahrt. 1 µl pro Ansatz wird auf ein 41 cm langes, 5 % Polyacrylamid-Gel mit 64-Haifischzahnkamm aufgetragen. Die Gel-Elektrophorese findet unter Standard-Bedingungen (nach Angaben des Herstellers) auf einem automatischen Sequenzierer Li-Cor 400 (MWG Biotech, Ebersberg) statt. Die gewonnenen Daten werden mit dem zum Sequenzierer gehörenden Softwarepaket Base ImagIR™ ausgewertet. Die Basensequenzen für die mit den reversen Primern durchgeführten Sequenzierungen werden mit dem Programm "Revseq" (<http://bioweb.pasteur.fr/seqanal/interfaces/revseq.html>) revers-komplementär gemacht. Nun können die 4 erhaltenen Sequenzen pro Fragment mit dem Programm ClustalX aligned werden. Das heißt, die einzelnen Sequenzen werden nach ihrer größten Übereinstimmung untereinander sortiert und daraus eine Konsensus-Sequenz gebildet.

2.2.11.2 Sequenzierung durch die Firma Seqlab

Der Firma Seqlab wurden rekombinante Vektoren mit pKAP298-Inserts und das Gesamt-Plasmid pKAP298 (zur Sequenzierung der nicht subklonierten pKAP298-Fragmente) als Templates zugeschickt und dort sequenziert. Die Plasmid-Vektoren mit pKAP298-Inserts mussten zuvor in *E. coli* amplifiziert und isoliert werden. Die Sequenzierung wurde in einem ABI 377 Sequenzierer der Firma Applied Biosystems mittels single-strand Primer-Walking durchgeführt. Die Start-Primer zur Sequenzierung nicht subklonierter pKAP298-Fragmente wurden von mir mitgeliefert. Die einzelnen Basensequenzen wurden mir per Email zugeschickt.

2.2.11.3 Zusammensetzen der Gesamtsequenz

Um die gesamte pKAP298 Sequenz zu ermitteln, müssen die Basensequenzen der einzelnen Fragmente aufbereitet und in der richtigen Reihenfolge zusammengesetzt werden. Mittels des Online-Programms Webcutter (<http://www.firstmarket.com/cutter/cut2.html>) und den geeigneten Restriktions-Enzymen werden die einzelnen Sequenzen zunächst "virtuell" geschnitten, um die Orientierung zu überprüfen (anhand der schon vorhandenen Restriktionskarte des Plasmids pKAP298 von QUACKENBUSH ET AL., 1983A). Gegebenenfalls wird die Sequenz mit dem Programm "Revseq" revers-komplementär gemacht. Nun werden bei den Sequenzen der subklonierten Fragmente die Vektoranteile entfernt. Dies geschieht wieder durch "virtuelles" Schneiden mit dem Webcutter. Als Restriktionsenzyme werden diejenigen benutzt, die schon bei der Subklonierung verwendet wurden. Alles vor der ersten und nach der letzten Schnittstelle stammt vom Vektor und wird gelöscht. Jetzt kann die Gesamtsequenz Fragment für Fragment zusammengesetzt werden. Doppelte Schnittstellen am Anfang und Ende der einzelnen Sequenzen werden entfernt.

2.2.12 Bioinformatische Analyse der pKAP298 - Sequenz

2.2.12.1 Erstellung von Restriktionskarten

Um die Identifizierung der DNA-Banden auf den Gelbildern der Southern Blots zu erleichtern, wird die pKAP298-Gesamtsequenz mittels der Programms Webcutter "virtuell" mit den gleichen Restriktionsenzymen verdaut, die auch schon zum DNA-Verdau für die Erstellung der Blots verwendet worden sind. Die Restriktionskarten sowie alle weiteren Grafiken dieser Arbeit werden mittels "Corel Draw 9" (Corel Corp., Unterschleißheim) erstellt.

2.2.12.2 Suche nach Nucleotid-Sequenzhomologien mittels BLAST

Die Gesamtsequenz wird mit dem Online-Programm "BLAST" (NCBI, <http://www.ncbi.nih.gov/BLAST/>) auf Sequenzhomologien zu schon bekannten Sequenzen durchsucht.

2.2.12.3 Suche nach potentiell kodierenden Open-Reading-Frames mittels

"GeneMark.hmm"

Die Gesamtsequenz wird mittels des Online-Programms "GeneMark.hmm" (<http://opal.biology.gatech.edu/GeneMark/>) nach potentiell Protein-kodierenden Open Reading Frames durchsucht. Der Suchalgorithmus des Programms GeneMark.hmm basiert auf der

Anwendung des "hidden Markov model". Empirische Beobachtungen an beschriebenen Protein-kodierenden Sequenzen zeigten, dass in einem solchen Bereich die Basensequenz nicht zufällig ist, sondern bestimmten Verteilungsmustern folgt. Das Programm sucht also in der Sequenz nach solchen Mustern und gibt den dazugehörigen Open-Reading-Frame an (LUKASHIN UND BORODOVSKI, 1998). Nach LUKASHIN UND BORODOVSKI ist das Programm in der Lage, aus 10 getesteten Genomen (*E. coli*, *B. subtilis*...) durchschnittlich 94,6 % der bisher schon identifizierten Gene wiederzufinden. Als durchschnittliche Rate von detektierten falschen positiven Genen (also nicht kodierende Bereiche) wird 10 % angegeben. Es besteht zwar bei Prokaryoten die Tendenz den größten potentiellen ORF zu transkribieren, die Initiation könnte aber auch an einem alternativen Startkodon weiter downstream im ORF stattfinden. Daher werden die als möglicherweise kodierend identifizierten ORFs (und ein Teil der upstream liegenden Region) nach potentiellen Ribosomen-Bindestellen untersucht. Dadurch lässt sich eine wahrscheinlichere Aussage darüber machen, an welchem Startkodon im ORF die Transkription (und später die Translation) beginnt. Laut LUKASHIN UND BORODOVSKI können so durchschnittlich bis zu 78 % der identifizierten ORFs in ihrer exakten Länge und mit ihrem wahrscheinlichsten Startkodon charakterisiert werden.

2.2.12.4 Suche nach Protein-Homologien der potentiell codierenden ORFs

Die von GeneMark.hmm identifizierten ORFs werden mittels Protein-BLAST (NCBI, <http://www.ncbi.nih.gov/BLAST/>) nach Homologien mit bekannten Enzymen untersucht. Besonderer Augenmerk wird auf "konservierte Domänen" gelegt. Als zufällig werden Übereinstimmungen gewertet, deren Erwartungswert (für einen Zufallsbefund) größer als 0,001 ist, es sei denn, dass die zu untersuchende Sequenz relativ klein und daher der Erwartungswert nicht sehr aussagekräftig ist. Zufällige Übereinstimmungen werden in dieser Arbeit nicht dargestellt.

2.2.12.5 Suche nach potentiellen Transposons

Die Basensequenz wird mit dem EMBOSS-Online-Programm „Einverted“ (<http://bioweb.pasteur.fr/seqanal/interfaces/einverted.html>) nach invertierten Wiederholungs-Sequenzen durchsucht, welche ein Kennzeichen für mögliche Transposons sind. Gleichzeitig wird an den Enden der putativen Transposons nach kurzen Wiederholungs-Sequenzen gesucht, die an beiden Enden in gleicher Richtung orientiert sind. Diese sind die Sequenzen, an denen die Insertion des Transposons stattgefunden hat. Schließlich wird auf dem putativen Transposon (anhand der Aminosäure-Sequenz) nach einer möglichen Transposase, dem die Transposition katalysierenden Enzym, und anderen putativ kodierenden Bereichen gesucht.

3. Ergebnisse

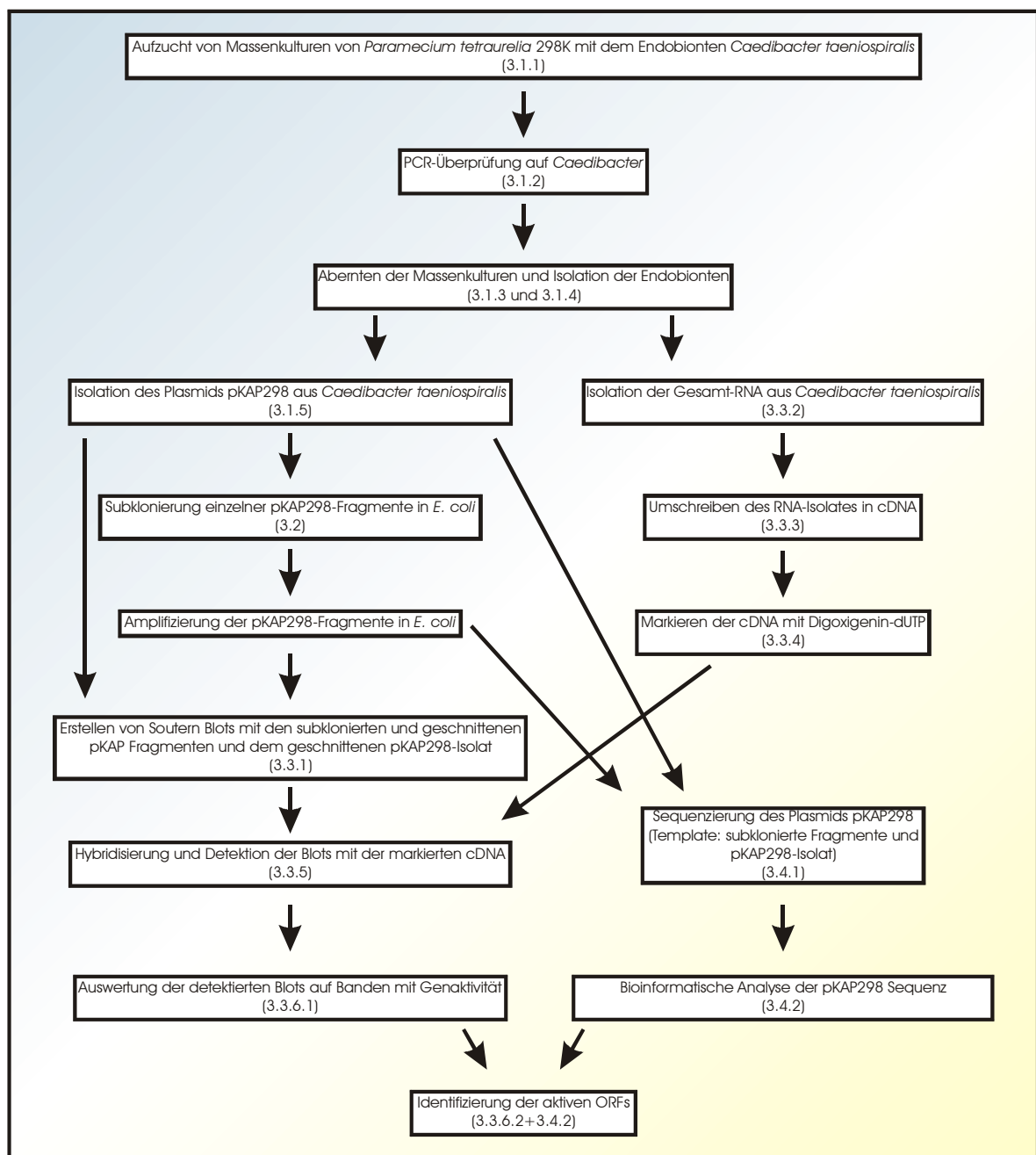


Abb. 3.1: Übersicht der durchgeführten Versuche und Untersuchungen

Die einzelnen Untersuchungen sind in ihrer chronologischen Reihenfolge dargestellt. Die Pfeile zeigen an, in welchen Folgeschritten die Ergebnisse (oder Daten) der einzelnen Versuche verwendet wurden. Aus den Ziffern in den Klammern lässt sich entnehmen, wo im Ergebnisteil die bei den jeweiligen Schritten erhaltenen Ergebnisse zu finden sind.

3.1 Isolierung von *Caedibacter taeniospiralis* und des Plasmids pKAP298

3.1.1 Massenkulturen von *Paramecium tetraurelia* Stamm 298

Für Plasmid- und Gesamt-RNA-Isolationen wurden Massenkulturen von *Paramecium tetraurelia* Stamm 298 angelegt. Da das untersuchte Bakterium *Caedibacter taeniospiralis* ein obligater Endobiont von *Paramecium tetraurelia* ist, also nicht ohne seinen Wirt kultiviert werden kann, mussten für die Untersuchungen an dem Bakterium große Mengen an *Paramecium*-Kulturen herangezogen werden. Für diese Arbeit wurden ca. 3000 l *Paramecium*-Kultur verwendet. Um möglichst ähnliche Kulturen zu erhalten, wurde das Kulturmedium wie unter 2.2.1.2 beschrieben bereitet, immer mit dem gleichen Futterbakterien beimpft, die Kulturen bei identischer Temperatur gehalten (ca. 23 °C, bzw. 12 °C) und die gleiche Zeit hungern gelassen. Somit konnte bei den Massenkulturen für die Plasmid-Isolation davon ausgegangen werden, nach dem Aufarbeiten der Kultur die ungefähr gleiche Menge an Plasmid-Isolat zu erhalten. Auf eine genaue Dokumentation der Kultur-Dichten wurde verzichtet, da es für die Arbeit lediglich relevant war, welche Mengen an Plasmid bzw. Gesamt-RNA isoliert werden konnten. Bei durchgeführten Auszählungen lag die durchschnittliche maximale Dichte der Kulturen bei 1200-1500 Zellen/ml. Nach der Verdoppelung des Kulturmediums hatte die Kultur-Dichte nach 24 h bei 23 °C wieder diesen Wert erreicht.

3.1.2 PCR-Nachweis von *Caedibacter taeniospiralis* in *Paramecium*

Im Intervall von 8 Wochen wurden die *Paramecium*-Massenkulturen mittels PCR auf das Vorhandensein der Endobionten überprüft. Spontan hatte keine der Kulturen ihre Endobionten verloren. Eine Verringerung der Zelldichte von *C. taeniospiralis* in ihrem Wirt *P. tetraurelia* 298 konnte nicht beobachtet werden. Die Menge des isolierten Plasmids pKAP298 aus den einzelnen Massenkulturen bewegte sich, bei identischer Kultivierung, in jeweils ähnlicher Größenordnung.

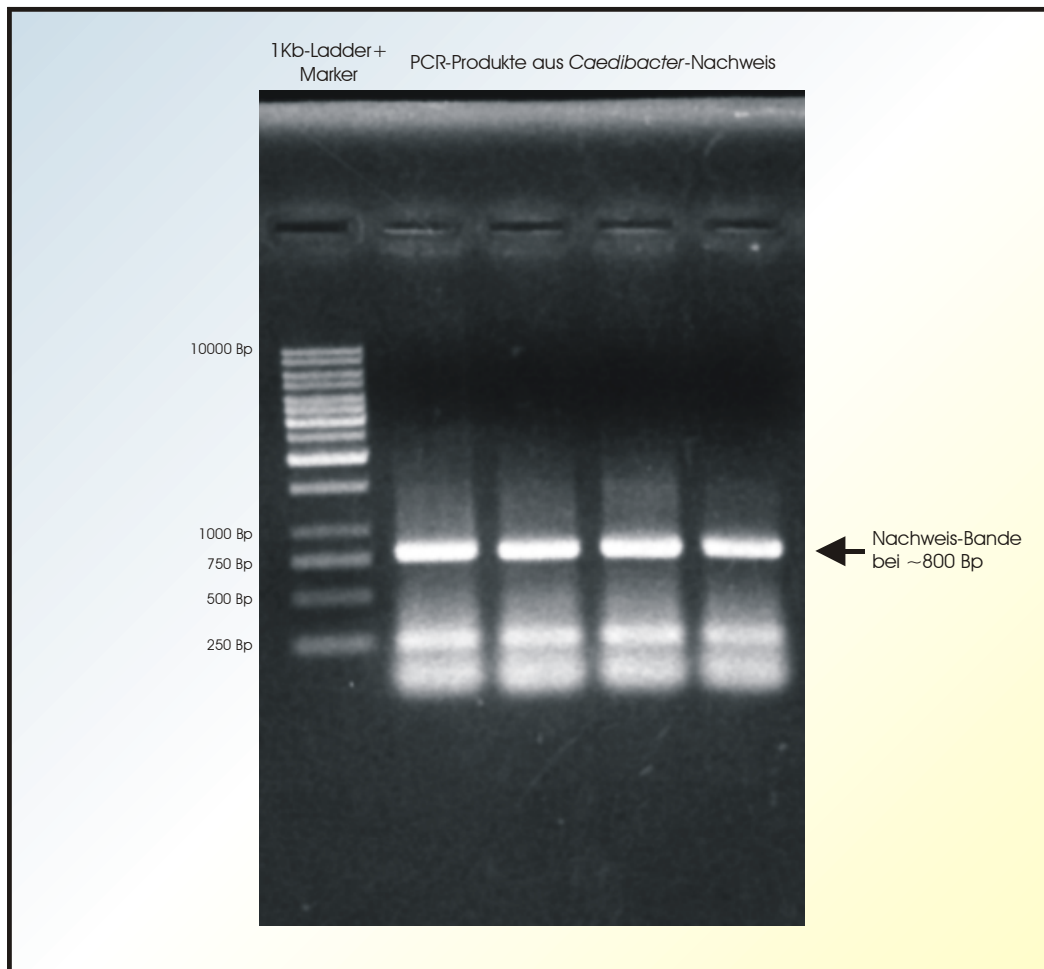


Abb. 3.1.21: PCR-Nachweis von *Caedibacter taeniospiralis* in der Massenkultur

Die Bande bei 800 Bp zeigt, dass sich *Caedibacter* in den Paramecien befanden. Gel 1,5 %, 5,6 V/cm, 45 min.

3.1.3 Abernten der Paramecium-Massenkulturen

3.1.3.1 Abernten zur *pKAP298*-Plasmid-Isolation

Nach einer Hunger-Inkubation von 10 Tagen bei 12 °C, wurden regelmäßig bis zu 31 Liter der Kultur abgeerntet. Bei einem durchschnittlichen Zelltitert von ungefähr 1200 Tieren/ml erhielt man so ein Paramecien-Pellet von ca. 37 200 000 Tieren. Ein Einengen der Paramecien mittels Durchlaufzentrifuge erwies sich als sehr effektiv. In dem ablaufenden Medium ließen sich, bei Betrachtung durch ein Binokular, keine Paramecien feststellen. Es blieben allerdings im eingengten Medium auch Detritus-Partikel, die durch die Gaze-Filtration nicht entfernt wurden, und kleinere Mikroorganismen (wie Flagellaten, Hefen und Bakterien) zurück. Diese ließen sich jedoch relativ effektiv durch das Zentrifugieren in der Öltestzentrifuge HNS II entfernen. Ein Teil setzte sie sich am Rand der Öltestbecher ab, der Rest bildete über dem

Paramecium-Pellet ein weiteres Pellet, das mittels einer Wasserstrahlpumpe und einer langen Pasteur-Pipette abgesaugt werden konnte. Das Paramecien-Pellet war, bei Betrachtung im Binokular, frei von Detritus und anderen Organismen. Als optimale Einstellung an der Öltestzentrifuge HNS II wurden 5 Minuten bei maximaler Geschwindigkeit ermittelt. Die Paramecien waren danach pelletiert und das Pellet ließ sich relativ leicht wieder resuspendieren.

3.1.3.2 Abernten zur Gesamt-RNA-Isolation

Nachdem die jeweiligen Massenkulturen angelegt und die betreffende Zeit (12 h oder 10 Tage) inkubiert worden waren, wurden sie abgeerntet. Die Prozedur war identisch mit dem Abernten zur Plasmid-Isolation, es musste aber, um die RNA-Zusammensetzung in den Bakterien möglichst wenig zu verfälschen, schneller gearbeitet und die Zellen in der Durchlaufzentrifuge mit eiskaltem Kulturmedium gewaschen werden. Nach einer nur 2 minütigen Zentrifugation in der Öltestzentrifuge wurden die Öltestbecher mit den Zellen auf Eis gestellt und sofort weiter aufgearbeitet. In *Paramecium*-Pellets zur Gesamt-RNA-Isolation befand sich, durch die verkürzte Zentrifugations-Zeit, mehr Detritus als in den Pellets aus 3.1.3.1.

3.1.4 Aufschluss der abgeernteten Paramecien und Isolation des Endobionten

Caedibacter taeniospiralis

3.1.4.1 Isolation des Endobionten zur pKAP298-Isolation

Die *Paramecium*-Pellets wurden, wie im Methodenteil beschrieben, aufgearbeitet. Die Resuspension der Pellets in je 2 ml Aqua_{bidest}, anstatt im Kaliumphosphat-Puffer, war zum Aufschluss essentiell. Wie sich in Vorversuchen gezeigt hatte, bewirkten die im Kaliumphosphat-Puffer vorhandenen 150 mM Sorbit ein Schrumpfen der Paramecien-Zellen, was ein Aufschließen der Zellen mittels eines Potter-Homogenisators unmöglich machte. Durch die Verwendung von Aqua_{bidest} statt des Kaliumphosphat-Puffers quollen die Zellen auf und konnten sehr leicht homogenisiert werden. Bei den weiteren Isolations-Schritten war nun wiederum die Verwendung des Kaliumphosphat-Puffers mit 150 mM Sorbit nötig, da dieser ein Platzen der Bakterien-Zellen verhinderte. Nach 4-5 Waschschrritten des Homogenats in dem Kaliumphosphat-Puffer war das *Caedibacter*/Zellorganellen-Pellet so weit von *Paramecium*-Zelltrümmern befreit, dass es, gründlich resuspendiert, zu der vorbereiteten Ecteola 23-Suspension gegeben werden konnte. Der Aufreinigungs-Schritt mittels Ecteola 23 erwies sich als extrem wichtig, da hierbei sehr effektiv die restlichen

Paramecium-Zelltrümmer und weiterer Detritus beseitigt werden konnten. Bei Plasmid-Isolationen, in denen *Caedibacter*-Isolate verwendet wurden, die nicht mittels Ecteola 23 aufgereinigt worden waren, ließ sich eine starke Kontamination des Plasmid-Isolats mit Fremd-DNA (wahrscheinlich gescherte genomische DNA von *Paramecium*) feststellen (siehe Test auf DNA-Kontamination des Plasmid-Isolats, 3.1.5.3). Die im Filter befindliche Ecteola/*Caedibacter*-Suspension konnte, um den letzten Rest der Bakterien aus der Suspension zu isolieren, mehrmals (durch nicht zu starkes Drücken) ausgepresst werden, ohne dass eine größere Fremd-DNA-Kontamination im späteren Plasmid-Isolat zu finden war.

3.1.4.2 Isolation des Endobionten zur Gesamt-RNA-Isolation

Die einzelnen Isolationen des Endobionten zur Gesamt-RNA-Isolation wurde, nach dem im Methodenteil beschriebenen, Verfahren durchgeführt. Alle Schritte wurden auf Eis (bzw. Zentrifugationen nahe 0 °C) durchgeführt, die jeweils angegebenen Zentrifugations-Zeiten wurden um die Hälfte verkürzt und das Pellet höchstens zweimal mit Kaliumphosphat-Puffer gewaschen. Die Ecteola/Bakterien-Suspension wurde dann sofort vorsichtig durch den Filter gedrückt. In diesem Fall war es besser, Kontaminationen durch die schnelle Aufarbeitung in Kauf zu nehmen, als möglicherweise das Endergebnis der RNA-Isolation zu verfälschen. Eine mögliche DNA- und Protein-Kontamination wurde, während der weiteren Aufarbeitung des RNA-Isolats, durch einen längeren DNaseI-Verdau und Phenol-Chloroform-Extraktion entfernt. Das Pellet wurde sofort weiter aufgearbeitet (siehe 3.3.2)

3.1.5 Isolation des Plasmids pKAP298 aus dem aufgereinigten Endobionten

3.1.5.1 Plasmid-Isolation für Southern Blots / zur Subklonierung in *E. Coli*

Es kam bei dieser Art der Plasmid-Isolation darauf an, eine möglichst große Konzentration von pKAP298-Plasmid in einem möglichst geringen Volumen von Lösungspuffer zu erreichen, ohne dass mögliche Kontaminationen (Proteine etc.) die nachfolgenden Enzymreaktionen inhibiert hätten. Um eine maximale Plasmid-Ausbeute aus dem *Caedibacter*-Isolat zu erreichen, wurde auf die Aufreinigung des Isolats über die dem Quiagen Plasmid Midi Kit mitgelieferten Silika-Matrix-Säulen verzichtet. Parallelversuche von Plasmid-Isolationen mit Verwendung der Säulen zeigten, dass die Ausbeute an Plasmid bis zu 50 % niedriger ausfiel als bei Isolationen ohne Säule. Der Grund dafür ist wohl die Größe des Plasmids, und die damit einhergehende stärkere Bindung an die Säule, was eine vollständige Elution erschwerte. Um mögliche Inhibierungen der nachfolgenden Enzymreaktionen zu verhindern, wurde das Isolat mit Phenol-Chloroform extrahiert. Um

Messfehler bei der späteren Plasmid-Konzentrationsbestimmung zu verhindern, war es wichtig, das restliche Phenol im extrahierten Isolat durch Ausschütteln mit Ether vollständig zu entfernen. Um bei der Präzipitation mit Isopropanol nicht zu viel Plasmid zu verlieren, mussten die angegebenen Inkubierungs- und Zentrifugations-Zeiten eingehalten werden. Des Weiteren war bei der Lösung der präzipitierten DNA zu beachten, dass sich das Plasmid während des Zentrifugations-Schritts nicht nur im sichtbaren Pellet, sondern auch an der umliegenden Wand des verwendeten Tubes, abgelagert hatte. Darum war es wichtig, das Lösungsmittel (10 mM Tris/HCl) auch über die das Pellet umgebende Tube-Wand gelangen zu lassen. Am effektivsten gelang dies durch "Anschneiden" des Tubes und die dadurch bewirkte gleichmäßige Verteilung der Tris-Lösung an der Tube-Wand. Um möglicherweise noch vorhandene Nucleasen zu inaktivieren, wurde das Plasmid-Isolat für 10 min in ein 75 °C warmes Wasserbad gestellt. Schließlich wurde die Plasmid-Konzentration des Isolats bestimmt. Aus 31 Liter Massenkultur, im Hungerzustand und mit einer Zelldichte von 1200 Zellen/ml, ließen sich typischerweise bis zu 15 µg pKAP298-Plasmid isolieren.

3.1.5.2 Plasmid-Isolation für Sequenzierungsreaktionen

Da die Firma SeqLab bei der Verwendung von pKAP298-Template, das mit Phenol-Chloroform-Extraktion aufgearbeitet war, bei der Sequenzierung keine Signale erhielt, wurde mit dem Plasmid-Isolat für die Sequenzierung anders verfahren. Bei der Aufreinigung des Isolats wurde die dem Quiagen-Kit mitgelieferte Säule benutzt. Dies bewirkte anscheinend eine gründlichere Reinigung der Plasmid-DNA von Kontaminationen. Die größere Reinheit des Plasmid-Isolats wurde mit einem hohen Verlust an Plasmid-DNA erkauft. So ließen sich, nach dieser Methode, aus 31 Litern Massenkultur höchstens 8 µg pKAP298-Plasmid isolieren.

3.1.5.3 Kontrolle des Plasmid-Isolats auf Verunreinigung mit Nicht-Plasmid-DNA

Vor der Weiterverwendung mussten alle pKAP298-Isolate auf eine Kontamination mit Nicht-Plasmid-DNA überprüft werden. Bei Isolaten, die für die Herstellung von Southern Blots verwendet werden sollten, durfte die Kontamination mit Fremd-DNA groß sein, damit, nach dem Restriktionsverdau und der elektrophoretischen Auftrennung des Plasmids, kleine Plasmid-Fragment-Banden nicht durch einen großen Hintergrund-DNA-Schmier überdeckt wurden. Eine genaue Zuordnung der Banden (und damit die Gen-Aktivitäts-Analysen) wären sonst unmöglich gewesen. Plasmid-Isolate für die Subklonierung in *E. coli* durften ebenfalls keinen großen Anteil an Fremd-DNA haben, da der DNA-Schmier bei den Ligations-Reaktionen mit der gewünschten Insert-DNA konkurriert, und die Transformations-Effizienz

drastisch gesenkt hätte. Dies gilt insbesondere für kleinere DNA-Fragmente, die der Wahrscheinlichkeit nach häufiger einligiert werden als die gewünschten großen Inserts. Bei Plasmid-Isolaten für Sequenzierungen musste schließlich der Anteil vom pKAP298-Template an der gemessenen Gesamt-DNA-Konzentration des Isolats abgeschätzt werden, um einen Ansatz der Sequenzier-Reaktion mit der richtigen Menge Template-DNA zu ermöglichen. Ein Aliquot eines jeden Plasmid-Isolats (~500 ng bis 1 µg) wurde mit dem Restriktionsenzym *Pst* I verdaut, einer Agarose-Gelelektrophorese unterzogen und fotografisch dokumentiert.

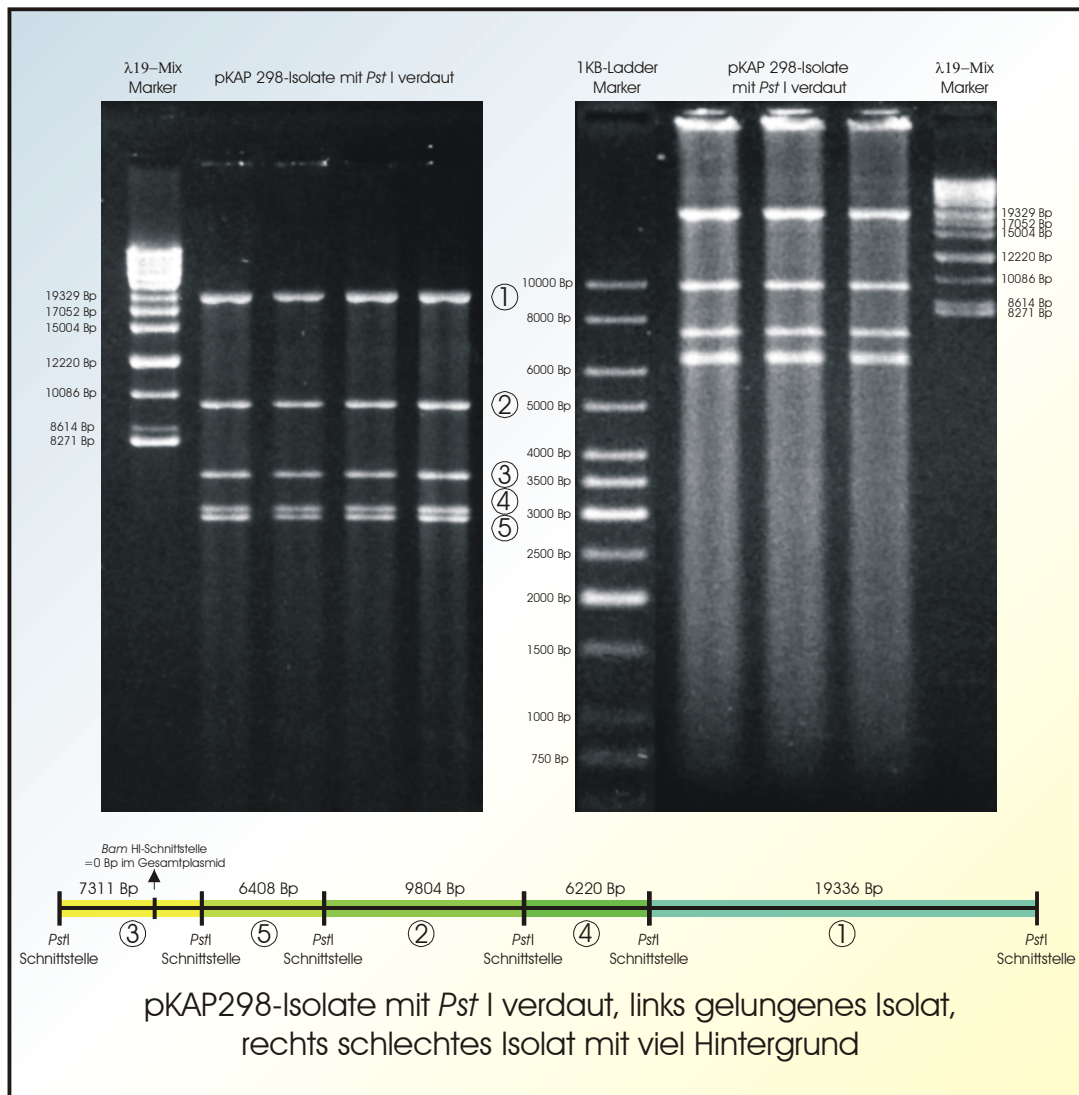


Abb. 3.1.5.3: Kontrolle des mit *Pst* I verdauten pKAP298-Plasmid-Isolats auf Verunreinigungen

linkes Gel ein Plasmid-Isolat mit wenig Fremd-DNA, rechts ein Isolat mit hohem DNA-Hintergrund (unbrauchbar). Agarose-Gele 0,8 %, 2,7 V/cm, 4 h.

3.2 Subklonierung der fehlenden pKAP298-Fragmente in *E. coli*

Um pKAP298-Fragmente für die Erstellung von Southern Blots zu amplifizieren und um die Sequenzierung des Plasmids zu erleichtern, wurde versucht, das komplette Plasmid pKAP298 in möglichst großen Fragmenten in *E. coli* zu klonieren. Während meiner Diplomarbeit (JEBLICK, 2000) wurden schon 60 % des Plasmids subkloniert. Ein ca. 19,4 kB großes Stück konnte nicht subkloniert werden. Dies wurde in dieser Arbeit nochmals versucht, indem das große Stück durch Multi-Enzym-Verdau in kleinere Fragmente geschnitten und diese dann in *E. coli* subkloniert wurden. Dies gelang nur zum Teil, zu den möglichen Gründen siehe 4.3. Es wird hier der genaue Ablauf der Klonierungs-Strategie dargestellt. Bei den Grafiken ist zu beachten, dass sie nicht an dem von QUACKENBUSH ET AL. (1983) definierten "Anfang" des Plasmids, d.h. an der *Bam* HI -Schnittstelle, beginnen, sondern an einer *Pst* I - Schnittstelle. Der Anfangspunkt des Plasmids ist als *Bam* HI - Schnittstelle eingetragen.

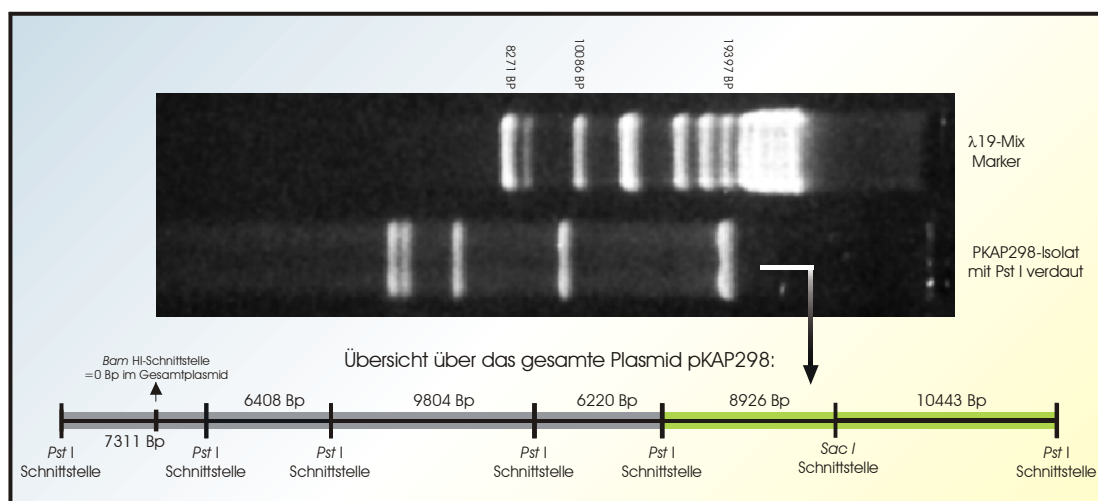


Abb. 3.2.1: Das noch nicht in *E. coli* subklonierte pKAP298-Fragment

Das noch nicht subklonierte pKAP298-Fragment (im Gel die rechte Bande) ist in grün dargestellt. Grau unterlegt sind die bereits während meiner Diplomarbeit subklonierten pKAP298-Fragmente.

3.2.1 Subklonierung der mit *Pst* I und *Sac* I geschnittenen Fragmente

Das bisher nicht subklonierte Fragment wird durch 2 *Pst* I - Schnittstellen flankiert. Etwa in der Mitte des Stückes befindet sich eine *Sac* I - Schnittstelle (siehe Abb. 3.2.1). Um die beiden nach einem Doppelverdau entstehenden Fragmente in *E. coli* zu klonieren, wurde das pKAP298-Plasmid-Isolat zunächst mit den Restriktionsenzymen *Pst* I und *Sac* I geschnitten,

einer Agarose-Gel-Elektrophorese unterzogen und die betreffenden Fragment-Banden aus dem Gel extrahiert und aufgereinigt. Danach wurde die DNA-Konzentration der Isolate gemessen.

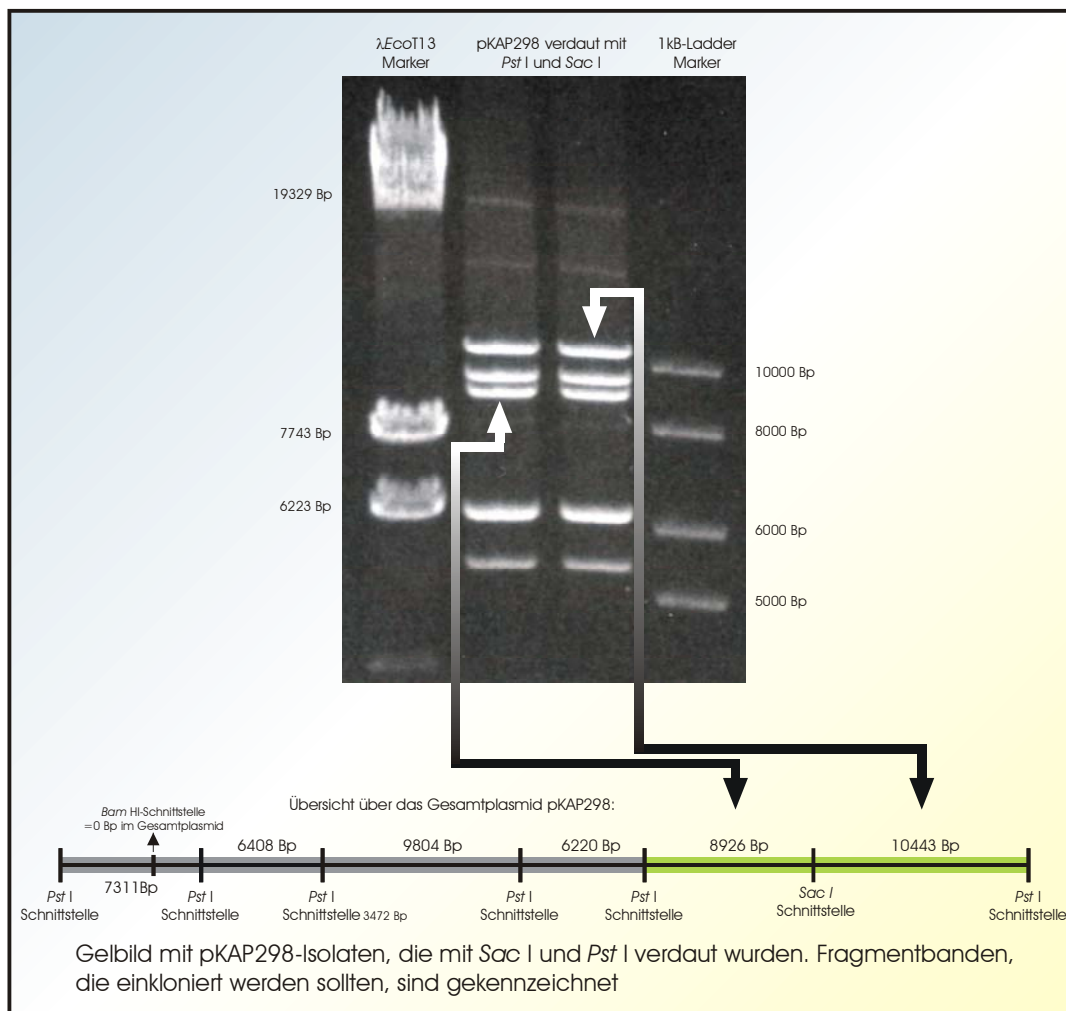


Abb. 3.2.1.1: Gelbild mit pKAP298-Fragmenten, die noch subkloniert werden sollten

Auf das Gel wurden pKAP298-Isolate aufgetragen, die mit *Pst* I und *Sac* I verdaut worden waren. Fragmente, die inkloniert werden sollten, sind mit Pfeilen gekennzeichnet. Auf der Plasmid-Übersicht unten ist nur eine *Sac* I Schnittstelle eingetragen. Gel 0,8 %, 2 V/cm, 5 h.

Der zur Transformation benutzte Vektor pHSG298 wurde zunächst mit den Restriktionsenzymen *Pst* I und *Sac* I geschnitten. Der Reaktionsansatz wurde aufgereinigt und die DNA-Konzentration des gereinigten Vektor-Isolats bestimmt. Auf eine Dephosphorylierung des Vektors wurde verzichtet, da er mit zwei verschiedenen Restriktionsenzymen geschnitten worden war, die jeweils unterschiedliche DNA-Überhänge produzieren, was eine Religierung des Vektors relativ effektiv verhinderte. Es wurden Ligations-Ansätze der beiden extrahierten pKAP298-Fragmente laut Methodenteil zusammenpipettiert und

inkubiert. Das Verhältnis von Vektor zu Insert betrug in beiden Fällen idealerweise 1:1. Jeweils 1 µl der Ligations-Ansätze wurde für die Transformation mittels Elektroporation benutzt. Danach wurden die Pulslängen abgelesen. Es ergab sich eine durchschnittliche Pulslänge von 3,8 ms. Die Pulslängen waren deshalb relativ kurz, weil auf die Präzipitation der Ligations-Ansätze verzichtet wurde, und deshalb in den Transformations-Ansätzen noch die Ionen des Ligations-Puffers vorhanden waren. Es erschien dennoch sinnvoller, eine etwas geringere Transformations-Effektivität in Kauf zu nehmen, anstatt durch eine misslungene Präzipitation den gesamten Ligations-Ansatz zu verlieren. Nach Inkubation der Transformations-Ansätze wurden diese auf selektive LB-Platten ausplattiert und inkubiert. Die gewachsenen Kolonien wurden nach den gewünschten Inserts gescreent. Das Screening fand zunächst über Blue-White-Screening statt. Da aber auch weiße Kolonien nicht unbedingt Inserts aus dem Plasmid pKAP298 hatten, wurden durch pKAP298-Sonden-Hybridisierung diejenigen weißen Kolonien herausgefiltert, die sicher ein Insert aus dem pKAP298-Plasmid trugen. Nachdem aus allen weißen Kolonien ein Kolonie-Blot erstellt worden war, wurde diese Membran mit einer pKAP298-Sonde hybridisiert und auf einem Kodak BioMAX Light Film dokumentiert. Alle Kolonien, die positive Signale gaben, hatten mit Sicherheit ein Fragment aus pKAP298 als Insert. Aus Kolonien mit pKAP298-Insert wurden dann Übernacht-Kulturen angelegt

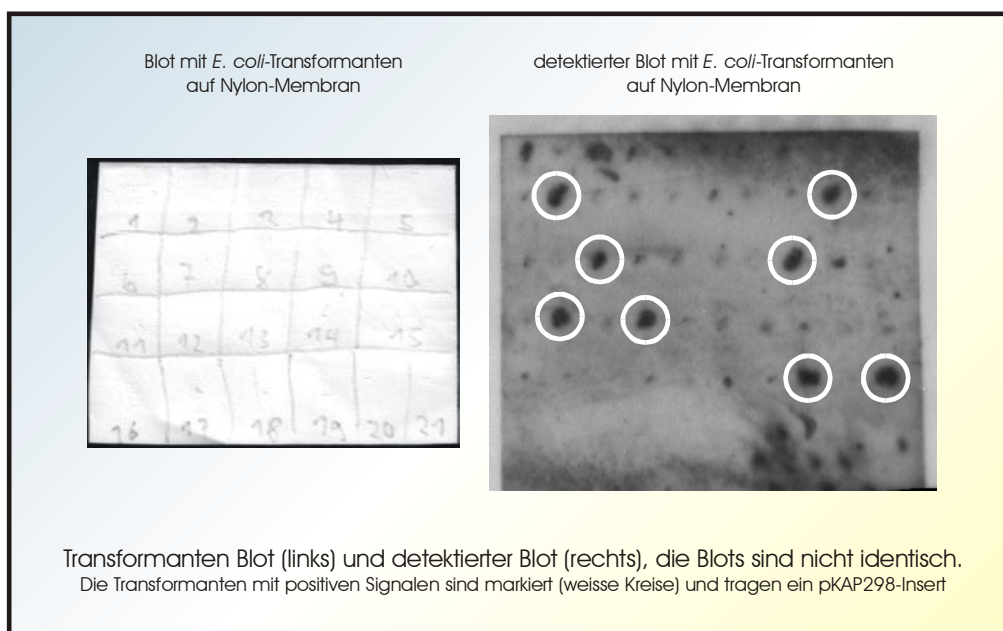


Abb. 3.2.1.2: Beispiel für Transformanten Screening mittels Hybridisierung

Links eine Nylon-Membran, auf die rekombinante *E. coli* geblottet wurden, rechts ein detektierter Blot (nicht identisch mit der links dargestellten Membran). Transformanten, die positive Signale ergaben (weisse Kreise), trugen sicher ein pKAP298-Insert.

Aus den Übernacht-Kulturen der Kolonien mit pKAP298-Insert wurden danach die rekombinanten Plasmide isoliert. Diese Plasmid-Isolate wurden mit *Pst* I und *Sac* I geschnitten, einer Agarose-Gel-Elektrophorese unterzogen und fotografisch dokumentiert. Kolonien mit dem richtigen Insert wurden als Glycerin-Kultur und Vektor-Isolat konserviert.

Das Ergebnis der Transformationen und der Auswertung einiger hundert Transformanten war, dass nur das 8926 Bp-Fragment erfolgreich in *E. coli* subkloniert werden konnte. Das 10443 Bp große Fragment ließ sich als Ganzes nicht in *E. coli* klonieren, obwohl die einzelnen Ligations-Reaktionen und Transformationen funktioniert hatten. Dies ließ sich daran erkennen, dass unter den mit den 10443 Bp-Ligations-Ansätzen transformierten Zellen durchaus viele Kolonien mit pKAP298-Inserts zu finden waren.

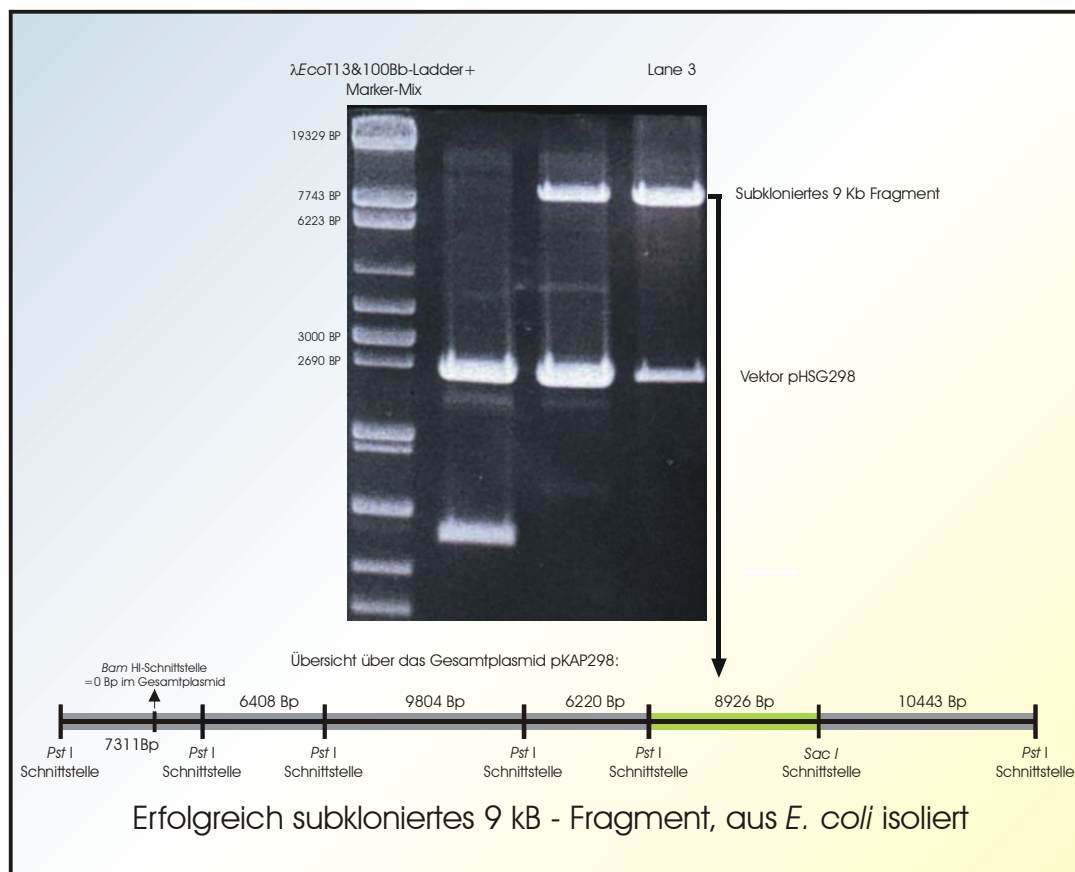


Abb. 3.2.1.3: Gelbild mit erfolgreich subkloniertem 9 kB-Fragment

Lane 3 zeigt ein mit *Pst* I und *Sac* I geschnittenes Plasmid-Isolat eines Transformanten mit gesuchtem 8926 Bp Insert aus dem Plasmid pKAP298. Die obere Bande entspricht dem gesuchten pKAP298-Insert, die untere ist der Vektor pHSG298. Gel 1 %, 3,7 V/cm, 2 h.

3.2.2 Klonierung von Subfragmenten des 10443 Bp Fragments in *E. coli*

Da es nicht gelungen war, das 10443 Bp-Fragment als Ganzes in *E. coli* zu klonieren, wurde nun versucht, es in kleineren Fragmenten zu subklonieren. Das fehlende Fragment weist einige *Eco* RI - Schnittstellen auf, die zur weiteren Zerkleinerung des Fragments benutzt werden konnten.

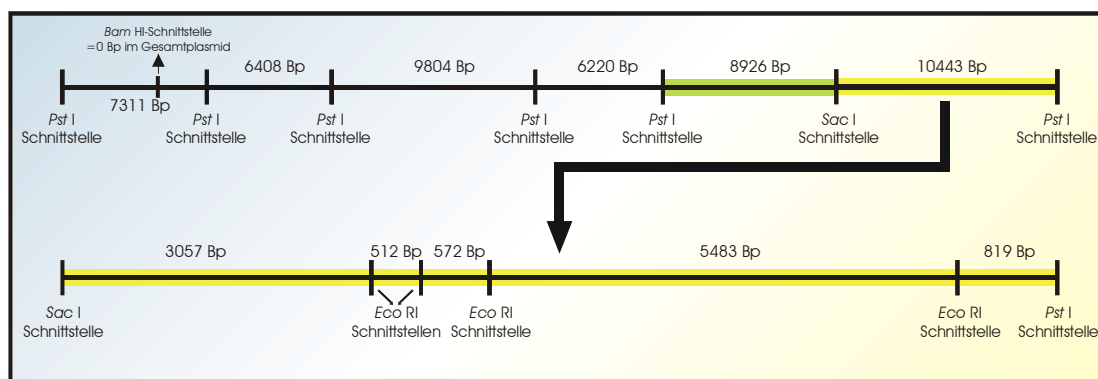


Abb. 3.2.2.1: Eco RI Schnittstellen im noch fehlenden pKAP298-Fragment

Es ist das noch nicht subklonierte pKAP298-Fragment mit seinen *Eco* RI-Schnittstellen, die zur weiteren Zerkleinerung benutzt werden konnten, dargestellt.

Ein pKAP298-Isolat wurde mit *Pst* I und *Sac* I verdaut, die Fragmente per Gel-Elektrophorese aufgetrennt und die 10,5 kB Bande aus dem Gel extrahiert. Das extrahierte 10,5 kB-Fragment wurde mit *Eco* RI verdaut, aufgereinigt und die DNA-Konzentration bestimmt. Da in der geschnittenen Insert-Lösung Fragmente mit unterschiedlichen Enden waren, musste auch der Vektor pHSG298 mit verschiedenen Enzymen verdaut werden. Es wurden 3 verschiedene Vektor-Verdaue angesetzt. Im ersten Ansatz wurde der Vektor mit *Sac* I und *Eco* RI verdaut, im zweiten mit *Pst* I und *Eco* RI und im dritten nur mit *Eco* RI. Der Vektor, der nur mit *Eco* RI verdaut worden war, wurde zusätzlich dephosphoryliert, um hohen Transformanten-Hintergrund nach der Transformation zu vermeiden. Alle 3 Vektor-Verdaue wurden aufgereinigt und die DNA-Konzentration bestimmt. Aus den 3 unterschiedlich geschnittenen Vektoren wurden mit dem verdauten 10,5 kB Fragment Ligations-Ansätze erstellt. Als durchschnittliche Insert-Größe wurde 2,7 kB angenommen, was der Größe des Vektors entspricht. Je 1 µl der drei Ligations-Ansätze wurden in *E. coli* transformiert. Zunächst wurde wieder mittels Sonden-Hybridisierung auf pKAP298-Inserts vorselektiert und dann die Inserts per Restriktionsverdau identifiziert. Auf diese Weise konnten 3 der fehlenden 5 kleineren Fragmente subkloniert werden. Die Klone mit den richtigen Inserts wurden als Glycerin-Kultur und Vektor-Isolat konserviert.

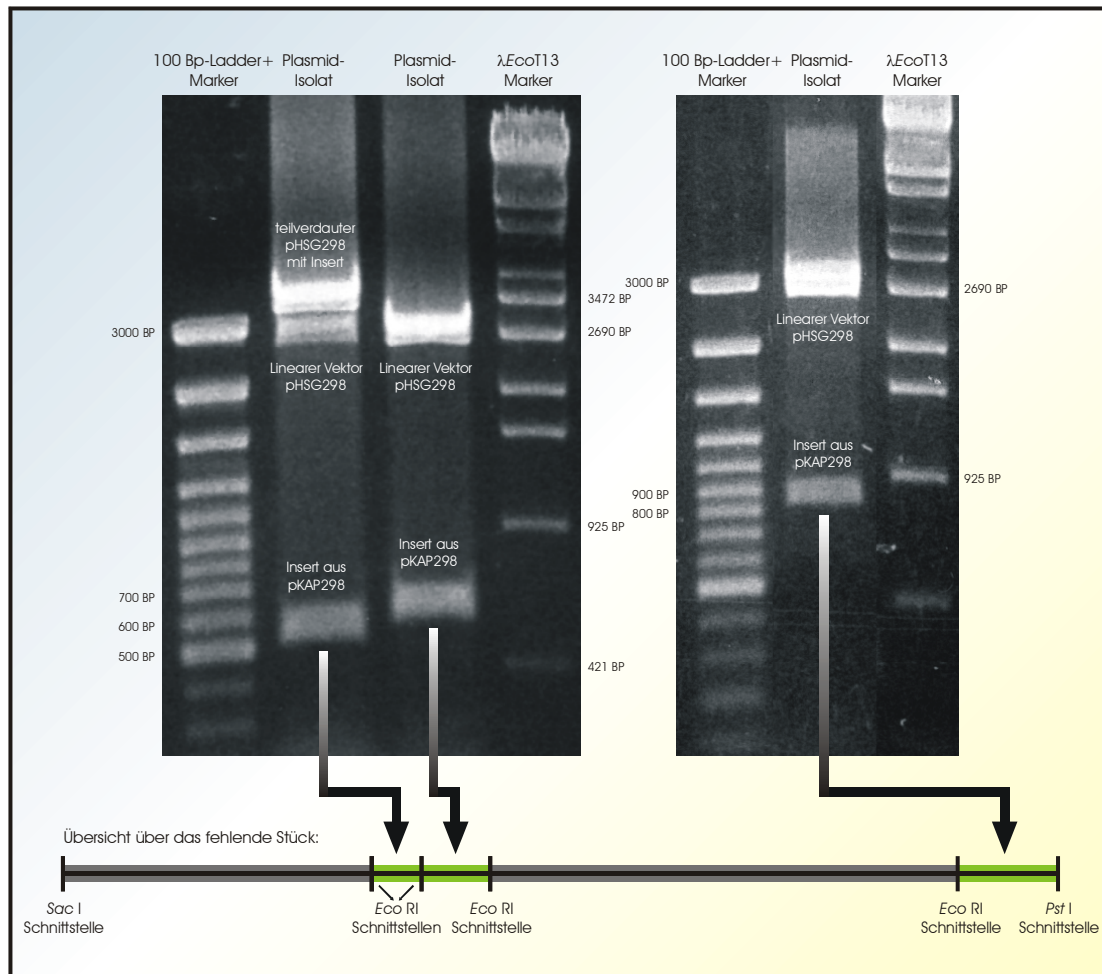


Abb. 3.2.2.2: Rekombinante Plasmide, die gesuchte pKAP298-Fragmente enthalten

Auf den Gelbildern sind mit *Eco* RI (bzw. im rechten Gelbild mit *Eco* RI und *Pst* I) geschnittene Plasmid-Isolate aus *E. coli*-Transformanten aufgetragen, die bisher fehlende pKAP298-Fragmente enthalten. Die Übersicht unten in der Abbildung zeigt, um welche der pKAP298-Fragmente aus dem fehlenden 10443 Bp Stücke es sich handelt. Gele 1,4%, 4,5 V/cm, 2 h.

3.2.3 Versuch die beiden fehlenden Subfragmente des 10443 Bp Fragments in *E. coli* zu klonieren

Es wurde pKAP298-Plasmid-Isolat mit *Sac* I und *Eco* RI verdaut, elektrophoretisch aufgetrennt, die gewünschten Fragment-Banden ausgeschnitten und die DNA daraus extrahiert. Die DNA-Konzentration der extrahierten Fragmente wurde gemessen.

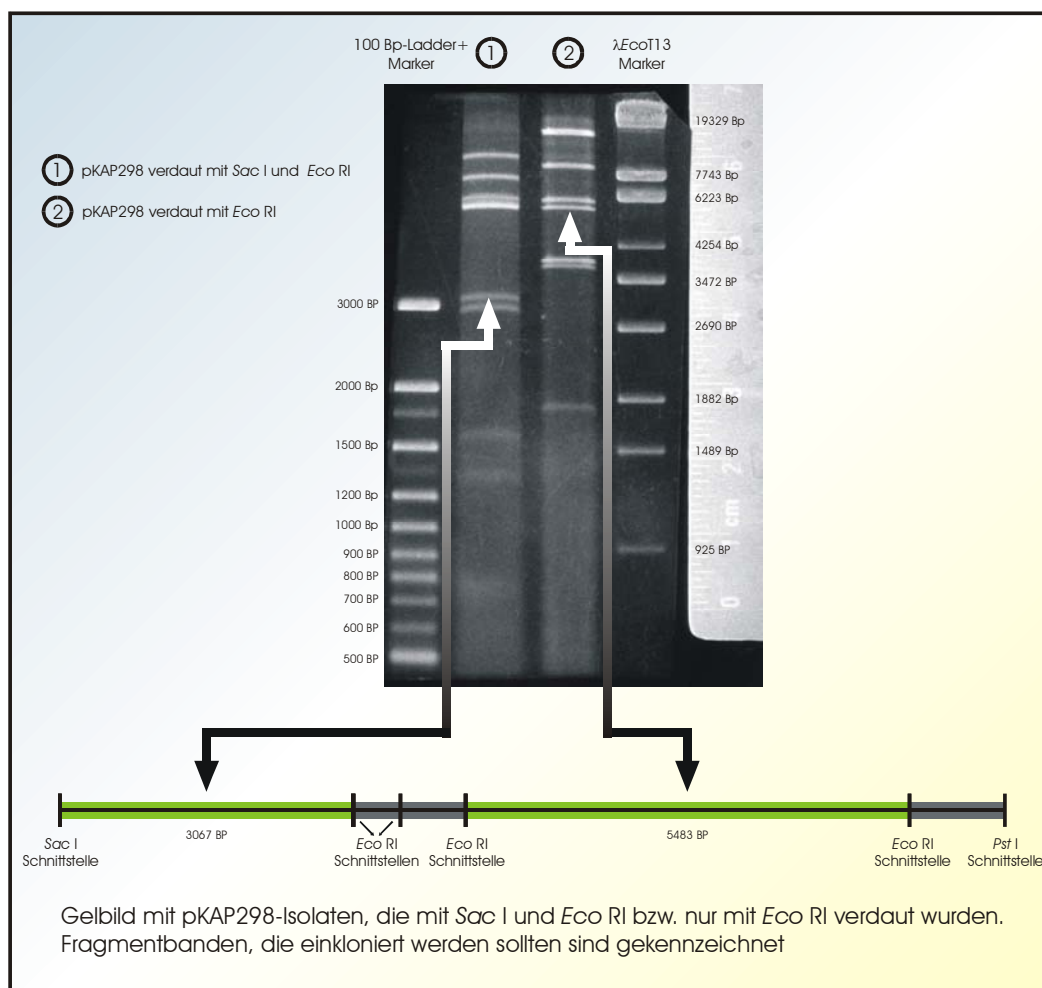


Abb. 3.2.3.1: Gelbild mit noch nicht subklonierten pKAP298-Fragmenten

In Lane 1 ist ein mit *Sac* I und *Eco* RI verdautes pKAP298-Isolat aufgetragen. In Lane 2 wurde das Plasmid-Isolat nur mit *Eco* RI verdaut. Die noch fehlenden und zu klonierenden Fragmentbanden sind mit Pfeilen markiert.

Der Vektor pHSG298 wurde durch Verdau mit *Sac* I + *Eco* RI, bzw. nur *Eco* RI, und anschließendem Aufreinigen für die Ligation vorbereitet. Der nur mit *Eco* RI verdaute Vektor wurde zusätzlich dephosphoryliert. Ligiert und transformiert wurde nach dem Standard-Protokoll. Es wurden sowohl JM109- als auch TOP10-Zellen zur Transformation verwendet. Gescreent wurde wiederum per pKAP298-Sonden-Hybridisierung und Restriktionsverdau der pKAP298-Insert enthaltenden Plasmide.

Nachdem viele positive Transformanten erfolglos nach den gewünschten Inserts gescreent worden waren, wurde die Subklonierung aufgegeben. In vielen Klonen waren pKAP298-Fragmente in den Vektor eingebaut worden, jedoch niemals die gesuchten. Die Größe der Fragmente kann nicht der Grund für den Misserfolg gewesen sein, da oftmals Fragmente von ähnlicher Größe wie die fehlenden Fragmente in den Vektor einligiert worden waren.

Mögliche Gründe, warum die Subklonierung der fehlenden Fragmente nicht funktioniert hat, werden in 4.3 diskutiert.

3.2.4 Zusammenfassung

Folgende Grafik zeigt, welche neuen Fragmente während dieser Arbeit erfolgreich in *E. coli* subkloniert werden konnten (grün). Die grau unterlegten Fragmente konnten nicht erfolgreich subkloniert werden.

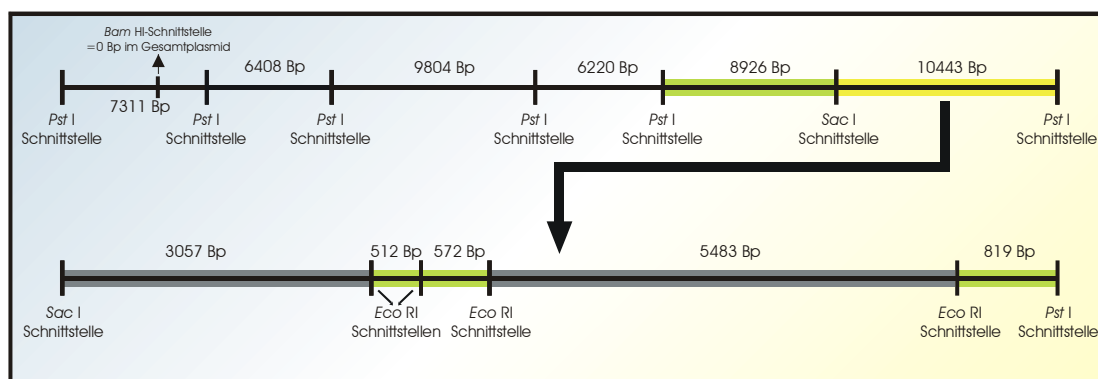


Abb. 3.2.4.1: Zusammenfassung der Subklonierungs-Versuche

Erfolgreich subklonierte Fragmente sind grün unterlegt, fehlende grau. Nicht markierte Fragmente wurden während meiner Diplomarbeit (JEBLICK, 2000) subkloniert.

3.3 Transkriptions-Aktivitäts-Analyse des Plasmids pKAP298

Das Prinzip der Gen-Aktivitäts-Analyse

Aus *Caedibacter taeniospiralis* wird die Gesamt-RNA isoliert. Das Isolat wird in cDNA umgeschrieben und die cDNA mit Digoxigenin markiert. Gleichzeitig wird aus *Caedibacter* das Plasmid pKAP298 isoliert, mit Restriktions-Endonucleasen geschnitten und als Southern Blot geblottet. Alternativ dazu wird der Großteil des Plasmids in *E. coli* subkloniert und amplifiziert. Aus der amplifizierten DNA werden ebenfalls Southern Blots erstellt. Diese Blots werden nun mit den markierten cDNA-Sonden hybridisiert und detektiert. Auf Fragmenten, die positive Signale ergeben, muss Transkriptions-Aktivität stattgefunden haben. Anhand der Basensequenz der Fragmente kann man nun nach potentiell kodierenden Open-Reading-Frames in der betreffenden Sequenz suchen, die möglicherweise für die Aktivität verantwortlich sind. Es ist zwar meist keine 100 % Identifizierung der aktiven ORFs möglich, aber es werden dadurch Hinweise auf weitere Forschungsansätze gegeben.

3.3.1 Erstellen von Southern Blots

3.3.1.1 Southern Blots aus dem Gesamt-Plasmid pKAP298

Die pKAP298-Plasmid-Isolate wurden zunächst mit den geeigneten Restriktionsenzymen geschnitten. Die Auswahl der Restriktions-Enzyme erfolgte nach 2 Kriterien. Zum einen sollte das Plasmid in möglichst viele Fragmente zerlegt werden, um Orte potentieller Genaktivität möglichst genau lokalisieren zu können. Zum anderen sollen die entstehenden Fragmente möglichst unterschiedlich groß sein, um eine deutliche Auftrennung der Fragmente in der Gel-Elektrophorese zu ermöglichen. Dazu wurden folgende Restriktionsenzyme in unterschiedlichen Kombinationen benutzt: *Ava* I, *Eco* RI, *Eco* RV, *Hind* III, *Kpn* I, *Pst* I, *Sac* I, *Sal* I und *Xho* I. Die geschnittenen pKAP298-Plasmid-Isolate wurden elektrophoretisch aufgetrennt und fotografisch dokumentiert.

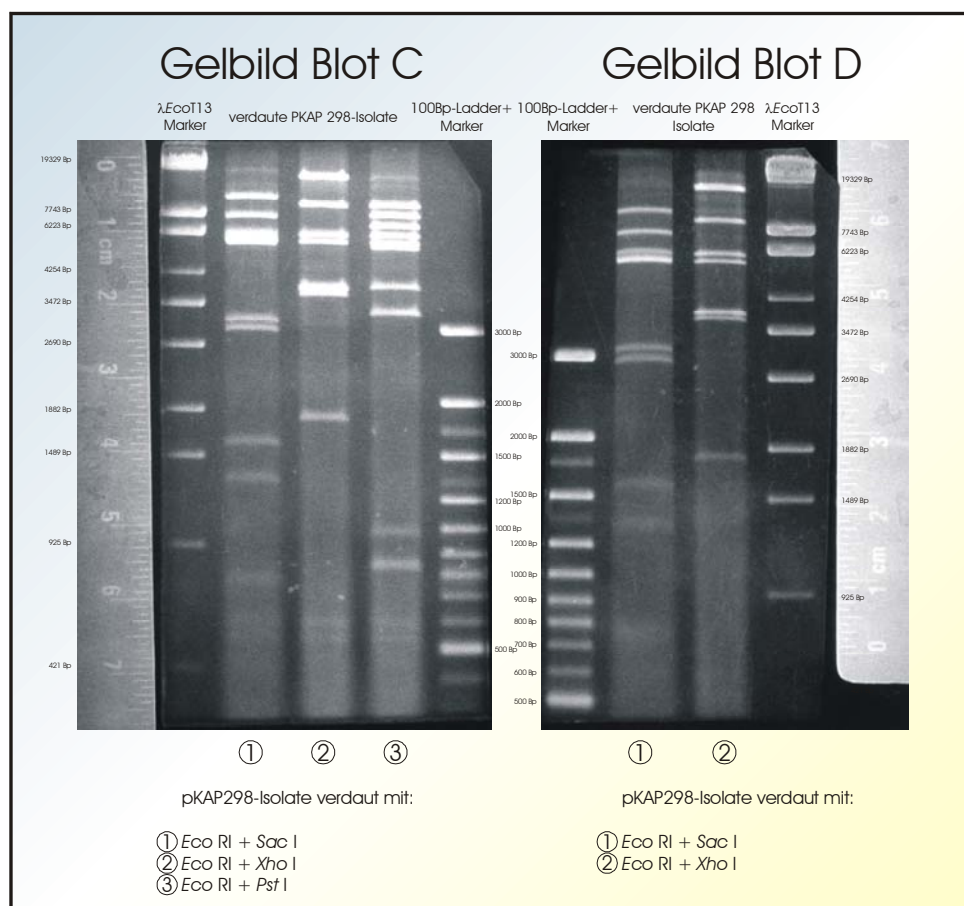


Abb. 3.3.1.1.1: Beispiel-Gelfotos

Es sind zwei Gelfotos abgebildet, aus denen Southern Blots hergestellt wurden. Auf die Gele wurden geschnittene pKAP298-Isolate aufgetragen (Bild links Lanes 1-3 , Bild rechts Lanes 1+2). Die benutzten Restriktionsenzyme werden darunter angegeben

Es werden in dieser Arbeit nur diejenigen Agarose-Gele abgebildet, die bei der Hybridisierung mit der markierten cDNA-Sonde und anschließender Detektion auswertbare Ergebnisse lieferten. Die Gelbilder werden bei der Auswertung der detektierten Blots (3.3.6.1) dargestellt. Außerdem sind sie als Corel-Draw-Grafiken auf der beiliegenden CD abzurufen (im Ordner "3_3_1_1_Corel_Gelbilder"). Das Problem bei der elektrophoretischen Auftrennung des geschnittenen Plasmids war, dass ein großes Spektrum an Bandengrößen aufgetrennt werden musste. Die kleinsten Fragmente lagen bei etwa 200 Bp, die größten bei etwa 19 kB. Da für die optimale Auftrennung von großen und kleinen Fragmenten jeweils andere Bedingungen gelten, musste zwischen den beiden Extrembedingungen ein Kompromiss gefunden werden, bei dem alle Fragmentgrößen einigermaßen scharf getrennt wurden. Um eine längere Auftrennungs-Strecke benutzen zu können, und damit eine bessere Trennung der Fragmente zu erreichen, wurden großes Agarose-Gele (100 ml, Länge 17 cm) hergestellt. Die Agarose-Konzentration wurde auf 1,4 % festgesetzt, um eine gute Auftrennung der kleineren DNA-Fragmente zu fördern. Die an den Gelen angelegte Spannung wurde im Gegenzug relativ niedrig (etwa 2,3 V/cm) gewählt, was eine bessere Trennung der großen Fragmente bewirkte. Die Agarose-Gele wurden nun, nach dem Protokoll im Methodenteil, auf den Blot-Vorgang vorbereitet und anschließend die DNA-Fragmente per Kapillar-Technik auf Nylonmembranen übertragen.

3.3.1.2 Southern Blots aus den in *E. coli* subklonierten pKAP298 - Fragmenten

Die Erstellung von Southern Blots mit den subklonierten pKAP298-Fragmenten lief analog zu der Erstellung der Southern Blots mit dem Gesamt-Plasmid ab. Zuvor mussten jedoch aus den rekombinanten *E. coli* die Plasmide mit den entsprechenden Inserts isoliert werden. Die Plasmide wurden mit den geeigneten Enzymen verdaut, elektrophoretisch aufgetrennt und die Agarose-Gele fotografisch dokumentiert. Die Agarose-Gele wurden nun, nach dem Protokoll im Methodenteil, auf den Blot-Vorgang vorbereitet und anschließend die DNA-Fragmente per Kapillar-Technik auf Nylonmembranen übertragen. Aufgrund des Umfangs der Arbeit, werden die Agarose-Gel-Fotos an dieser Stelle nicht abgebildet. Sie werden bei der Auswertung der detektierten Blots (3.3.6.1) dargestellt. Außerdem sind sie als Corel-Draw-Grafiken auf der beiliegenden CD abzurufen (im Ordner "3_3_1_2_Corel_Gelbilder").

3.3.2 Isolation und Aufreinigung der Gesamt-RNA aus *Caedibacter taeniospiralis*

3.3.2.1 Isolation der Gesamt-RNA

Aus den isolierten *Caedibacter* wurde, wie im Methodenteil beschrieben, die Gesamt-RNA isoliert. Um möglichen RNA-Abbau durch noch vorhandene oder während der Präparation eingeschleppte RNAsen zu verhindern, musste schnell und immer mit Handschuhen gearbeitet werden. Die Unversehrtheit der RNA ließ sich durch eine Agarose-Gel-Elektrophorese überprüfen. Wenn im Isolat noch Banden zu erkennen waren, und der RNA-Schmier über 2 kB hinaus ging, war das Isolat in Ordnung.

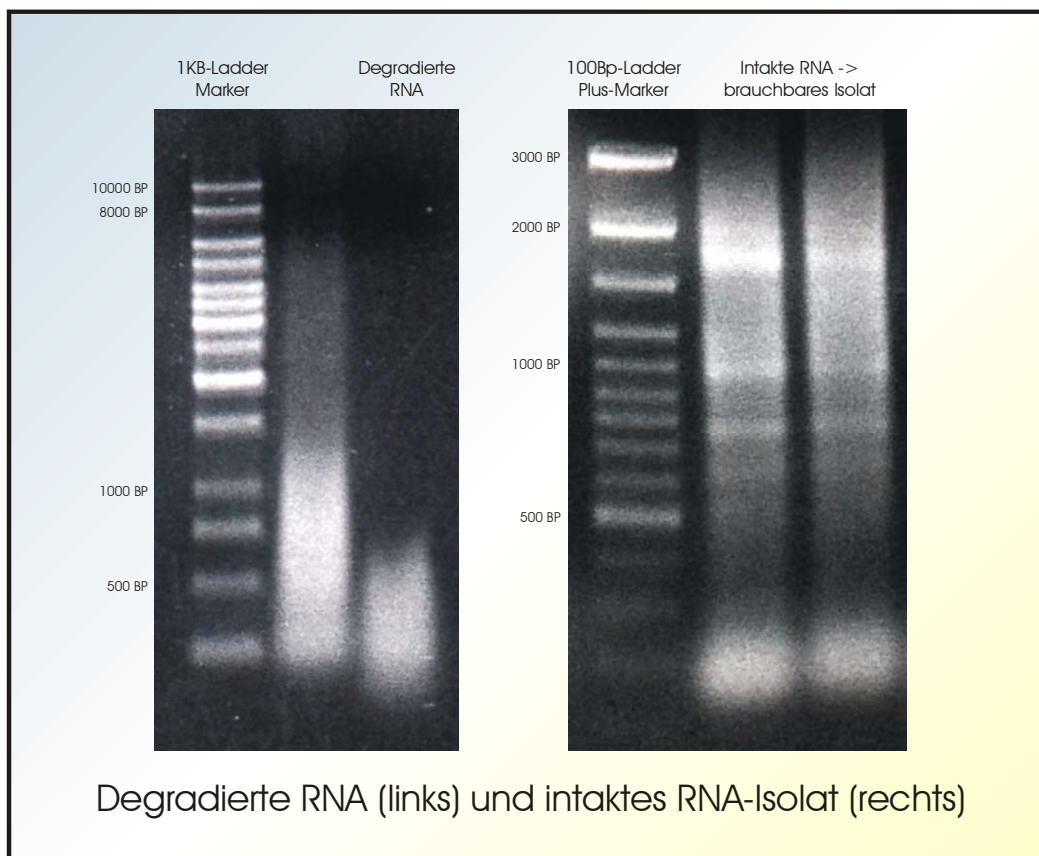


Abb. 3.3.2.1.1: Vergleich zwischen unbrauchbarem und gutem RNA-Isolat

Im linken Gelbild ist ein degradiertes RNA-Isolat dargestellt, im rechten Gelbild ein intaktes RNA-Isolat. Bei dem intakten RNA-Isolat sind die Banden ribosomaler RNA gut zu erkennen. Dieses Isolat konnte weiterverarbeitet werden. Gele: 1,5 %, 7,5 V/cm, 30 min.

Da die RNA-Isolationen für die entsprechenden Kultur-Zustände mehrmals wiederholt wurden, werden hier nur 2 Ergebnisse von typischen Isolationen dargestellt. Die Menge der isolierten Gesamt-RNA konnte (auch bei gleichen Ausgangsbedingungen) stark differieren,

was zum einen am Ernährungszustand der Zellen, zum anderen an der Präparation liegen konnte (z.B. Verluste durch die Ecteola-Säule bei der Bakterien-Isolation).

1. Ausgangs-Kultur: 16 Liter *Paramecium*-Massenkultur (23 °C), gefüttert, abgeerntet nach 12 h. Gemessene Konzentration des RNA-Isolats in 100 µl DEPC-Aqua_{bidest}: 192,8 ng/µl.
2. Ausgangs-Kultur: 16 Liter *Paramecium*-Massenkultur (23 °C), hungernd, abgeerntet nach 10 Tagen. Gemessene Konzentration des RNA-Isolats in 100 µl DEPC-Aqua_{bidest}: 142,4 ng/µl.

Diese Werte sind natürlich nicht exakt, da noch DNA (wie das Plasmid pKAP298) in den Isolaten vorhanden war. Möglicherweise verfälschten auch Phenol-Reste, die trotz mehrmaligem Waschens des Isolats mit 70% Ethanol nicht entfernt werden konnten, die Konzentrations-Messung. Die gemessenen Konzentrationen dienten jedoch zur Berechnung der einzusetzenden DNase im DNaseI-Verdau (3.3.2.2).

3.3.2.2 Aufreinigung des Gesamt-RNA-Isolats von kontaminierender DNA

Das Gesamt-RNA-Isolat musste, um die bei der weiteren Aufbereitung störende DNA-Kontamination (z.B. das Plasmid pKAP298) zu beseitigen, einem gründlichen DNaseI-Verdau unterzogen werden. Als Basis zur Berechnung der pro Verdau einzusetzenden DNaseI wurden die in 3.3.2.1 gemessenen RNA-Konzentrationen verwendet. Die Inkubierungs-Zeit durfte nicht zu lange sein, um mögliche RNA-Degradation durch noch vorhandene RNAsen so gering wie möglich zu halten. Ferner war darauf zu achten, dass RNasefreie DNaseI benutzt wurde, da sonst die RNA durch die in der Enzymlyse enthaltene RNase vollständig abgebaut worden wäre.

Nach dem DNaseI-Verdau wurden wieder die RNA-Konzentrationen der Isolate (aus 3.3.2.1) gemessen:

1. gemessene Absorbtion $1,71 * 80 = 136,8$ ng/µl
2. gemessene Absorbtion $1,32 * 80 = 105,6$ ng/µl

Die Messungen zeigten, dass bis zu 30% der Gesamt-Nucleinsäuren-Konzentration des Isolats DNA-Kontamination waren.

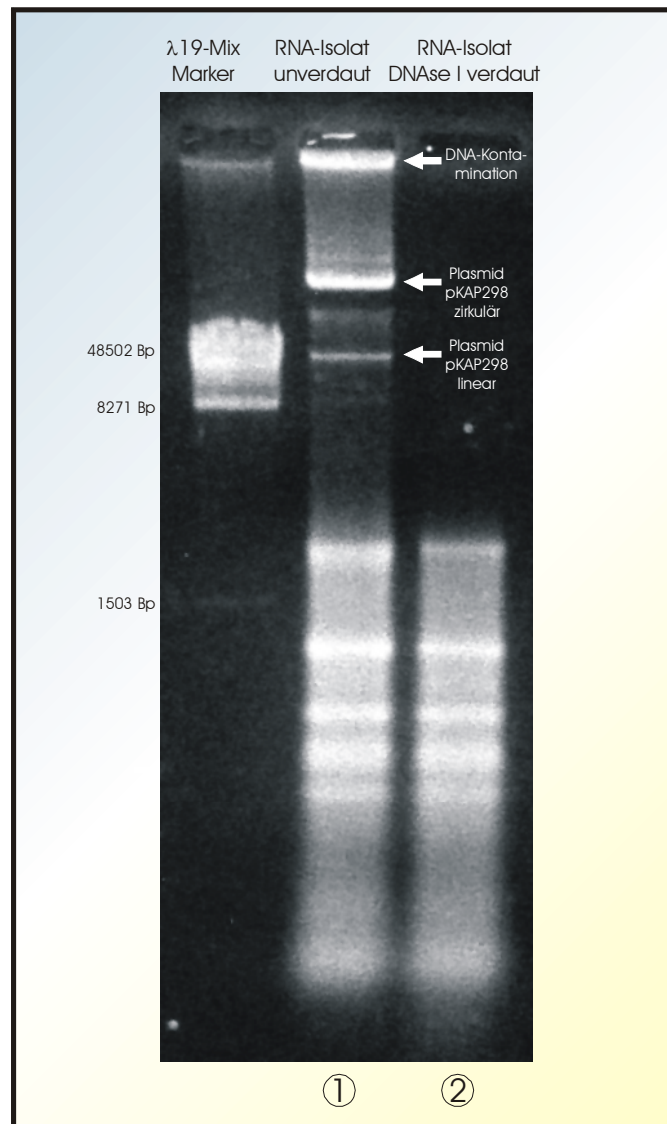


Abb. 3.3.2.2.1: RNA-Isolat vor und nach DNaseI-Verdau

Lane 1 zeigt ein RNA-Isolat vor, Lane 2 das gleiche Isolat nach einem 30 min DNaseI-Verdau. Es ist gut zu erkennen, dass die sichtbare DNA-Kontamination vollständig verdaut worden ist. Die RNA ist relativ intakt geblieben. Gel 1,5 %, 7,5 V/cm, 30 min.

Die verdauten RNA-Isolate wurden, zusammen mit den entsprechenden unverdauten Isolaten, einer Agarose-Gel-Elektrophorese unterzogen. Dadurch ließ sich erkennen, ob die RNA während des Verdau stark abgebaut worden war. Außerdem ließ sich dadurch abschätzen, ob der DNaseI-Verdau funktioniert hatte (siehe Abb. 3.3.2.2.1). Die letzte Gewissheit, ob das Isolat kontaminationsfrei war, erhielt man allerdings erst nach der Detektion mit der aus dem Isolat hergestellten und markierten Sonde. Das Gesamt-RNA-Isolat wurde bei -70 °C aufbewahrt.

3.3.3 Umschreiben der Gesamt-RNA in cDNA

Die Gesamt-RNA-Isolate aus 3.3.2 wurden mit dem "RevertAid H Minus First Strand cDNA Synthesis"-Kit der Firma MBI Fermentas in cDNA umgeschrieben. Als Template für die Transkription wurde jeweils 1 µg des DNase-verdauten RNA-Isolats benutzt. Die reverse Transkription wurde nach dem Standard-Protokoll aus dem Methodenteil durchgeführt. Da das Transkript wegen möglichen cDNA-Verlusten nicht aufgereinigt wurde, war eine exakte DNA-Konzentrationsbestimmung unmöglich. Die Transkripte lieferten bei den Konzentrationsmessungen immer sehr ähnliche Werte. Wurden nach der Markierung eines Aliquots der cDNA und anschließender Detektion keine Signale festgestellt, konnte ein weiteres Aliquot des jeweiligen cDNA-Transkripts mittels Agarose-Gel-Elektrophorese untersucht werden. War ein DNA-Schmier zu sehen, der über 3 kB hinausging, konnte davon ausgegangen werden, dass zumindest der Transkriptions-Schritt funktioniert hatte.

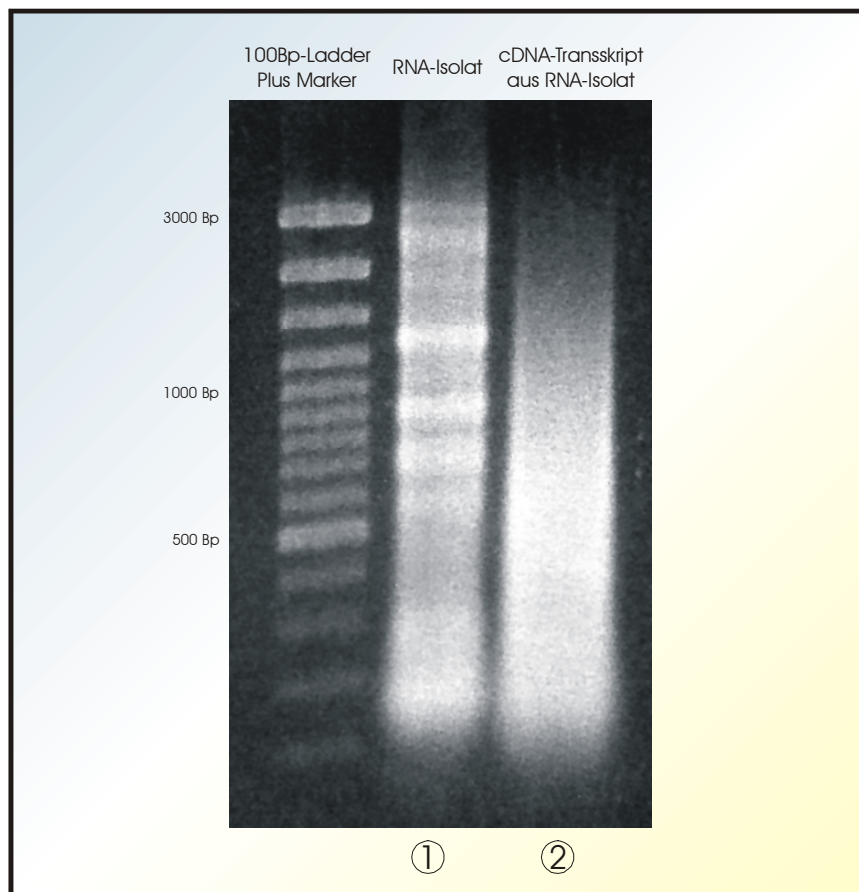


Abb. 3.3.3.1: gelungene cDNA-Transkription

Das cDNA-Transkript (Lane 2) wurde aus dem RNA-Isolat (Lane 1) hergestellt. Man kann erkennen, dass auch größere Fragmente transkribiert worden sind. Gel 1,5 %, 7,5 V/cm, 30 min.

3.3.4 Markieren der cDNA mit digoxigenin-dUTP

Von den cDNA-Transkripten wurden jeweils zunächst 4 µl (ein viertel Ansatz) mit Digoxigenin markiert, hybridisiert und detektiert (3.3.5). Dies geschah, um zu überprüfen, ob die bisherigen Schritte (3.3.2-3.3.3) funktioniert hatten. War die Überprüfung positiv verlaufen, wurden die restlichen 16 µl der jeweiligen cDNA markiert. Wichtig war eine lange Denaturierung des cDNA-Transkripts (10 min bei 99 °C), um eine gleichmäßige Anlagerung der Random-Hexamer-Primer an die cDNA zu ermöglichen. Um optimale Erträge an markierter cDNA zu erhalten, fand die Markierungs-Reaktion für 20 h im Thermocycler mit Heizdeckel statt. Die Effektivität der Markierung konnte durch den Vergleich einer Verdünnungsreihe des markierten Produktes mit einer Verdünnungsreihe einer markierten Standard-Lösung getestet werden. Aliquots der Verdünnungsreihen wurden auf eine Nylon-Membran pipettiert, im UV-Licht fixiert und detektiert. Waren nach der Detektion bei der cDNA-Reihe Signale zu sehen, hatte die Markierungs-Reaktion funktioniert. Dies sagt nur aus, dass und wie viel der cDNA markiert worden ist, nicht jedoch welche cDNA markiert wurde. Dieser Schritt wurde als Test, ob die Markierung funktioniert hat, nur dann durchgeführt, wenn bei der Hybridisierung und Detektion mit dem markierten Produkt keine Banden zu sehen waren. Ergab der Test, dass die Markierungs-Reaktion erfolgreich gewesen war, musste der Fehler in einem früheren Schritt gelegen haben (3.3.2-3.3.3).

3.3.5 Hybridisierung und Detektion der markierten cDNA-Sonden mit den Southern Blots

Da insgesamt etwa 150 Hybridisierungen und Detektionen durchgeführt wurden, und die Dokumentation den Rahmen der Arbeit sprengen würde, werden hier nur diejenigen dokumentiert, die gelungen, und daher relativ gut auszuwerten waren (siehe 3.3.6.1). Es konnten nicht alle erstellten Southern Blots mit allen verschiedenen cDNA-Sonden hybridisiert werden, da die entsprechenden Hybridisierungs-Lösungen mit den Sonden nur für wenige Hybridisierungen verwendet werden konnten, bevor die Signalstärke (aufgrund der quantitativen Entfernung der Sonde durch Hybridisierung an die Blots) zu schwach wurde.

3.3.5.1 Hybridisierung der markierten cDNA-Sonden mit den Southern Blots

Die Hybridisierungsbedingungen wurden für alle Blots gleich gewählt. Die jeweilige Hybridisierung fand für 48 h bei 39 °C unter leichtem Schütteln statt (~150 rpm). Wichtig war eine gründliche Denaturierung der jeweiligen Sonde (zunächst 10 min bei 99 °C und dann in der Hybridisierungs-Lösung nochmals 10 min bei 68 °C), um deren Anlagerung an die geblottete DNA zu ermöglichen. Die Hybridisierungs-Temperatur durfte keinen

Schwankungen unterliegen - besonders nicht Schwankungen nach unten, da dies zu einer massiven, unspezifischen Sonden-Anlagerung geführt hätte. Bei einem Ausfall des Wärmeschrankes und einer Rest-Temperatur von 32 °C, konnte auf dem betreffenden Blot sogar die DNA-Markerbanden erkannt werden, die normalerweise nicht detektiert werden dürften.

Hybridisierungen wurden in folgenden Kombinationen durchgeführt:

- Southern Blots aus verdautem pKAP298-Plasmid mit markiertem cDNA-Transkript aus RNA-Isolaten, jeweils von gefütterten und hungernden Massenkulturen
- Southern Blots aus in *E. coli* subklonierten pKAP298-Fragmenten mit markiertem cDNA-Transkript aus RNA-Isolaten, jeweils von gefütterten und hungernden Massenkulturen

3.3.5.2 Detektion der hybridisierten Southern Blots

Die Detektion aller hybridisierten Blots wurde auf die gleiche Weise durchgeführt. Die Membranen wurden einer gründlichen Stringenz-Waschung unterzogen (siehe Methodenteil), um unspezifisch gebundene Sonden-DNA abzuwaschen. Dies gelang jedoch nicht immer vollständig. Bei Southern Blots mit subklonierten pKAP298-Fragmenten war, aufgrund der großen Menge an geblotteter DNA, fast immer ein schwacher Background an unspezifisch gebundener Sonde zu sehen. Bei der Auswertung der detektierten Filme (siehe 3.3.6.1) wurden daher nur Signale gewertet, die sich deutlich vom DNA-Banden-Hintergrund unterschieden. Wichtig war ein gründliches Abwaschen der unspezifisch auf die Membran gebundenen Digoxigenin-Antikörper durch zweimaliges 30 min Inkubieren des Blots im Waschpuffer, da nicht abgewaschene Antikörpern sich bei der Detektion als schwarze Flecken - bzw. hohen Background - auf dem Film bemerkbar machten. Waren auf dem detektierten Film die Banden nur schwach zu sehen, und gleichzeitig der Background ebenfalls nur schwach, wurde die Inkubierung mit dem BioMax Film verlängert.

Nach der Detektion ließ sich zum ersten Mal definitiv erkennen, ob die bisherigen Schritte, von der RNA-Isolation bis zur Hybridisierung und Detektion, funktioniert hatten. War dies nicht der Fall, wurden die einzelnen Schritte überprüft. Da die Unversehrtheit der RNA schon in 3.3.2 durch Gel-Elektrophorese überprüft worden war, konnte der Fehler nur im Transkriptions-Schritt (3.3.3) oder der Markierung (3.3.4) gelegen haben. Ob die jeweiligen Schritte funktioniert hatten, wurde nach den bei den einzelnen Punkten angegebenen Verfahren geprüft.

3.3.5.3 Dokumentation der hybridisierten Southern Blots

Aufgrund des Umfanges der Arbeit werden die detektierten Filme bis auf ein Beispiel-Film an dieser Stelle nicht abgebildet. Sie werden jedoch später, bei der Auswertung der detektierten Blots, dargestellt (siehe 3.3.6.1). Außerdem sind sie als Corel-Draw-Grafiken auf der beiliegenden CD abzurufen (in den Ordnern "3_3_5_3_Corel_detektierte_Blots_pKAP298" und "3_3_5_3_Corel_detektierte_Blots_subklonierte_Fragmente").

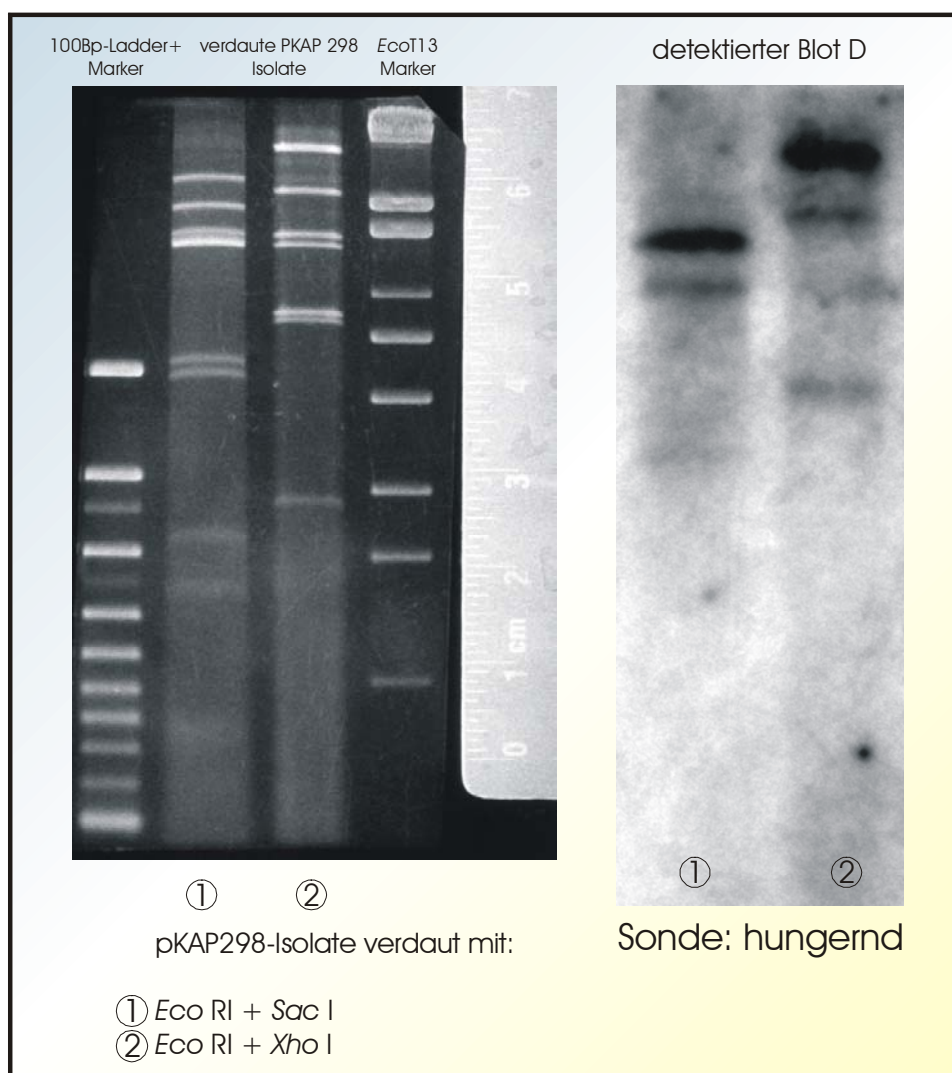


Abb. 3.3.5.3.1: Beispiel für detektierten Southern Blot-Film

Aus dem Gelbild (links) wurde ein Southern Blot erstellt und mit der Sonde "hungernd" hybridisiert. Danach wurde detektiert. Der entwickelte Film ist im Bild rechts dargestellt

3.3.6 Identifizierung der Orte von Genaktivität

(Dieser Teil der Arbeit konnte chronologisch erst stattfinden, nachdem die Basensequenz des Plasmids pKAP298 bekannt war)

Die Detektion der hybridisierten Blots mit RNA-Isolaten aus gefütterten und hungernden Kulturen ergab keine einheitlichen, für die einzelnen Zustände der Massenkulturen eindeutig reproduzierbare, Ergebnisse (Gründe siehe Diskussion 4.1, wahrscheinlich ist der RNA-Zustand durch die lange Isolierungsdauer identisch geworden). Auf die getrennte Auswertung von gefütterten und hungernden Kulturen wird daher verzichtet. Die Bezeichnungen "gefüttert" oder "hungernd" werden nur noch zur Unterscheidung der einzelnen detektierten Blots benutzt.

3.3.6.1 Auswertung der detektierten Blots

Die einzelnen detektierten Blots wurden, wie im Methodenteil beschrieben, ausgewertet. Auf den nächsten Seiten wird die Auswertung der einzelnen detektierten Southern Blots in Bezug auf Orte mit Transkriptions-Aktivität dargestellt. Da aufgrund der Größe der Grafiken auf den Seiten kein Platz mehr für Bildunterschriften ist, wird an dieser Stelle eine Legende für die einzelnen Grafiken geliefert.

Detektierte Blots, die aus in *E. coli* subklonierten pKAP298-Fragmenten erstellt wurden, sind "SubA" und "SubB" benannt, gefolgt von der Bezeichnung der jeweiligen verwendeten Sonde (z.B. "gefüttert 1"). Links oben ist das Agarose-Gel abgebildet, aus dem der Southern Blot hergestellt wurde (in der Legende "A"). Die einzelnen Banden sind nummeriert, damit ihre Position auf den Fragmentbanden identifiziert werden kann. Links unten ist der Film des mit der entsprechenden Sonde hybridisierten und detektierten Southern Blots abgebildet (aus in der Legende "B"). In der Mitte der Grafiken sind die 5 (bzw. bei den Blots "SubB" nur 4) pKAP-Fragmentbanden mit den entsprechenden Restriktions-Schnittstellen dargestellt (in der Legende "C"). Die Zahlen in den Kreisen geben an, welches Fragment welcher Agarose-Gel-Bande entspricht. In diese pKAP-Fragment-Skizzen sind die Banden, auf denen sichere Transkriptions-Aktivität nachgewiesen werden konnte, grün markiert. Unsichere Bande sind hellgrün. Die unterste Grafik zeigt die Anordnung der einzelnen Fragmentbanden auf dem Plasmid pKAP298, der Startpunkt des linearisierten Plasmids ist die eingetragene *Bam* HI-Schnittstelle im 7311 Bp-Fragment.

Detektierte Blots, die aus dem verdauten Plasmid pKAP298 erstellt wurden, sind als Blot "A" bis "E" bezeichnet, gefolgt von der jeweils zur Hybridisierung verwendeten Sonde ("gefüttert", "hungernd"). Oben links ist das Agarose-Gel abgebildet, aus dem der Blot hergestellt wurde (in der Legende "A"), die einzelnen Banden sind wiederum nummeriert.

Rechts oben ist der Film des mit der entsprechenden Sonde hybridisierten und detektierten Southern Blots abgebildet (in der Legende "B"). Unten ist schematisch das linearisierte Plasmid pKAP298 dargestellt, jeweils mit den Schnittstellen der zum Verdau verwendeten Restriktionsenzyme (in der Legende "C"). Auch hier sind die Fragmente nummeriert, um eine Identifikation der positiven Signalbanden zu ermöglichen. In diesen pKAP-Skizzen sind die Banden, auf denen sichere Transkriptions-Aktivität nachgewiesen werden konnte, grün markiert, unsichere Bande sind hellgrün (z.B. sich überlagernde Banden). Alle Abbildungen sind auch auf der beigefügten CD als Corel-Draw-Grafiken abzurufen (im Ordner "3.3.6.1_Auswertung_der_detektierten_Blots").

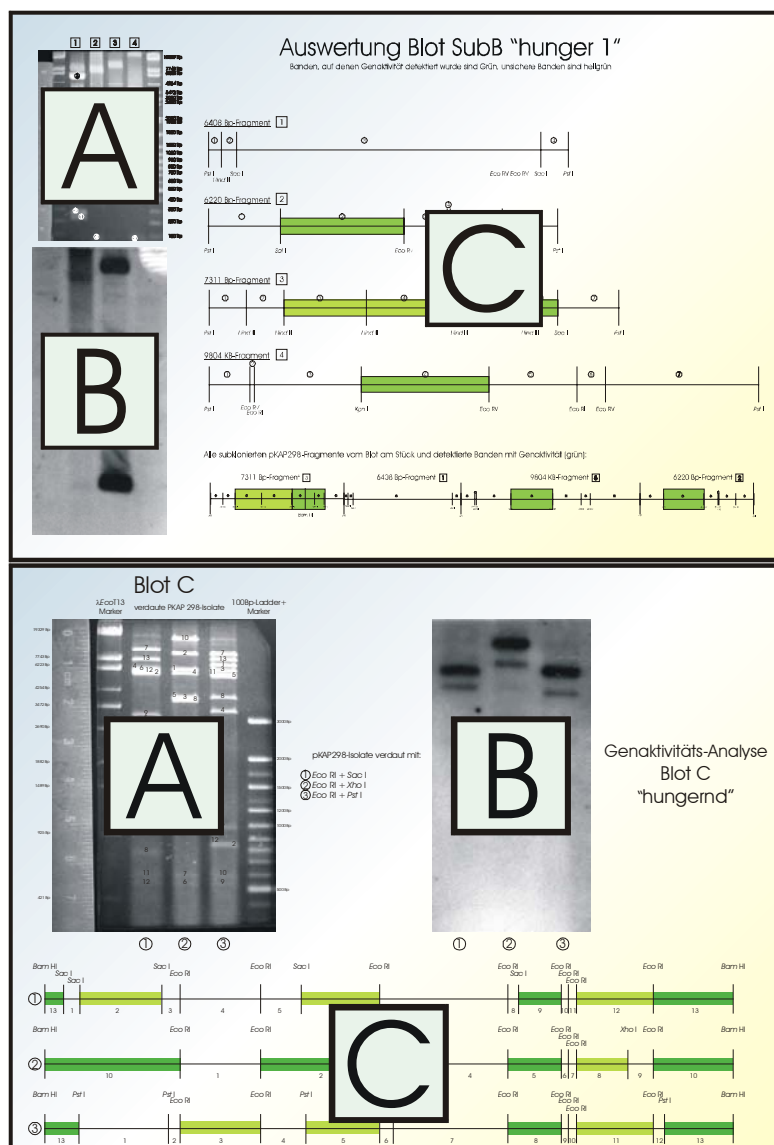
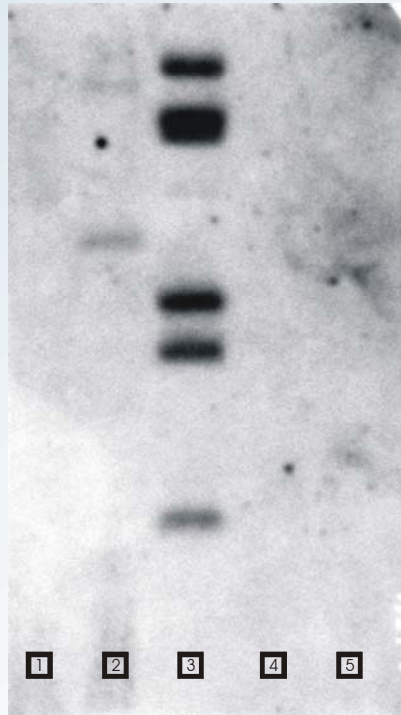
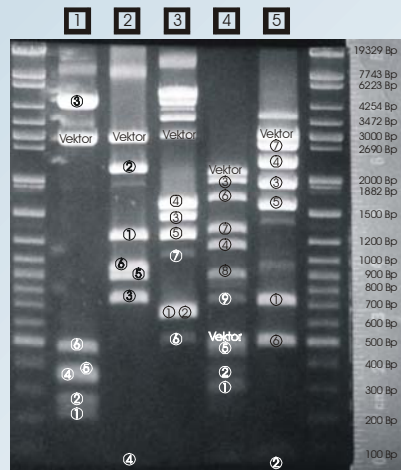


Abb. 3.3.6.1.1: Legende der auf den nächsten Seiten folgenden Grafiken

Die jeweiligen Beschreibungen von "A", "B" und "C" sind dem obigen Text zu entnehmen.



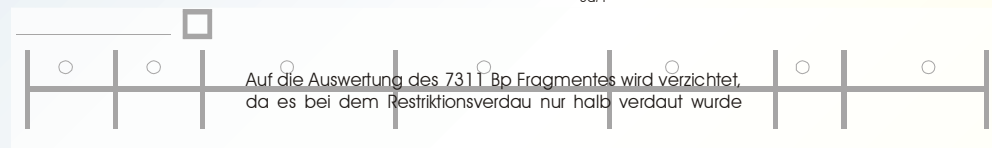
Auswertung Blot SubA "gefüttert 1"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün

6408 Bp-Fragment **1**

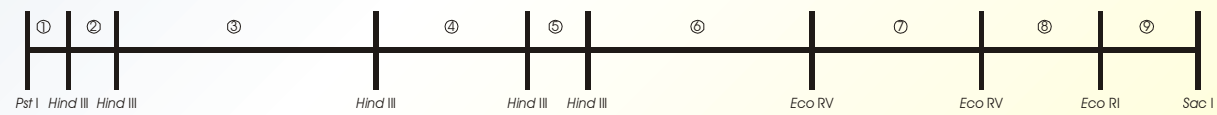


6220 Bp-Fragment **2**

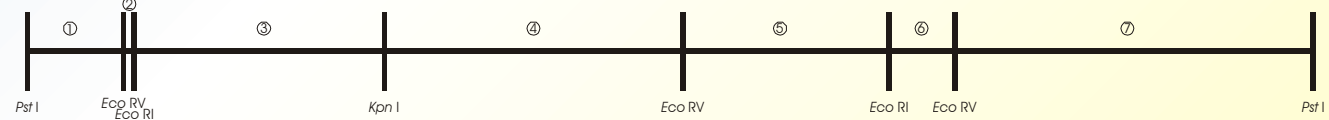


Auf die Auswertung des 7311 Bp Fragmentes wird verzichtet, da es bei dem Restriktionsverdau nur halb verdaut wurde

8926 Bp-Fragment **4**

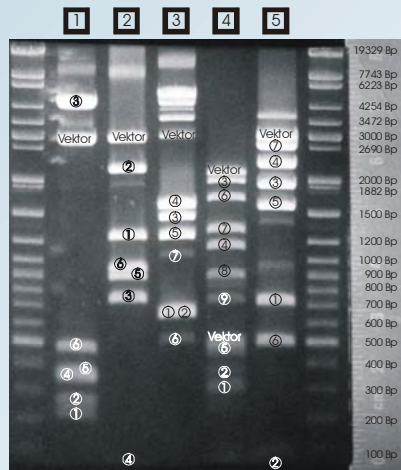


9804 KB-Fragment **5**



Alle subklonierten pKAP298-Fragmente vom Blot am Stück + detektierte Banden mit Genaktivität (grün):





Auswertung Blot SubA "gefüttert 2"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün

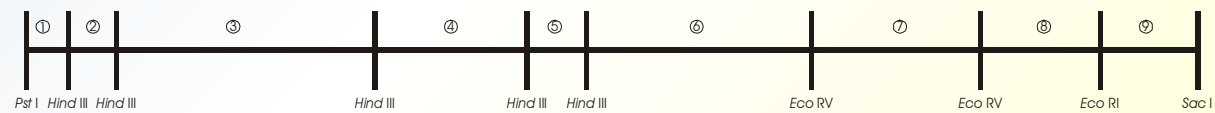
6408 Bp-Fragment **1**



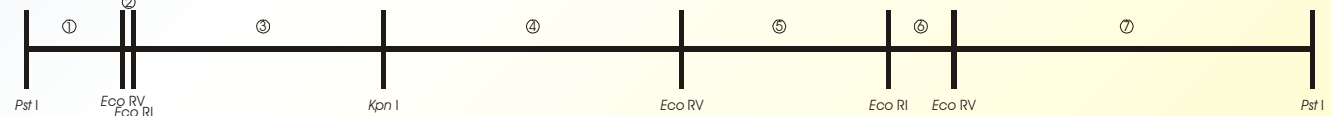
6220 Bp-Fragment **2**



8926 Bp-Fragment **4**

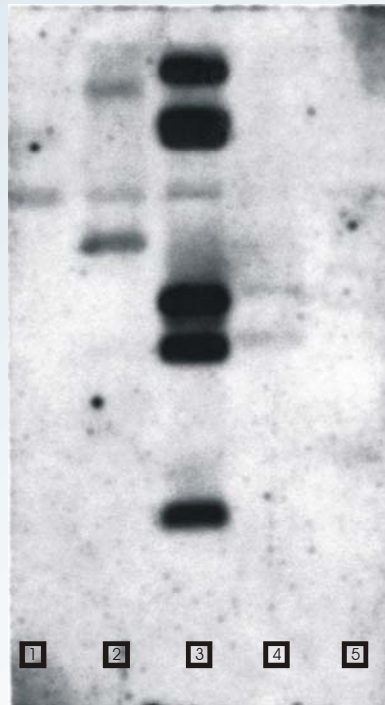
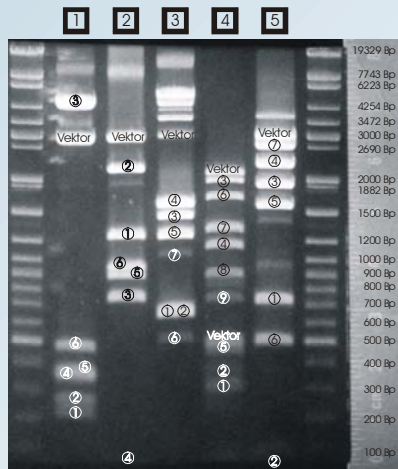


9804 KB-Fragment **5**



Alle subklonierten pKAP298-Fragmente vom Blot am Stück + detektierte Banden mit Genaktivität (grün):





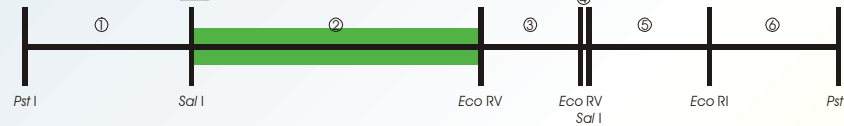
Auswertung Blot SubA "gefüttert 3"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün

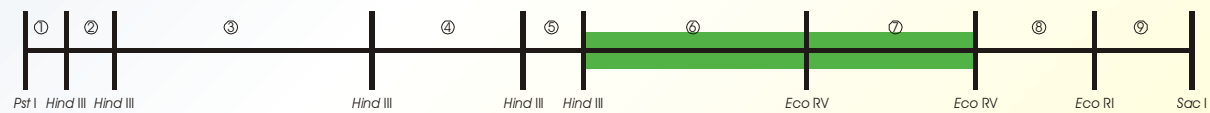
6408 Bp-Fragment **1**



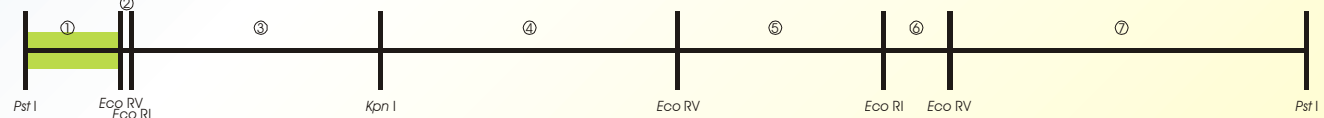
6220 Bp-Fragment **2**



8926 Bp-Fragment **4**

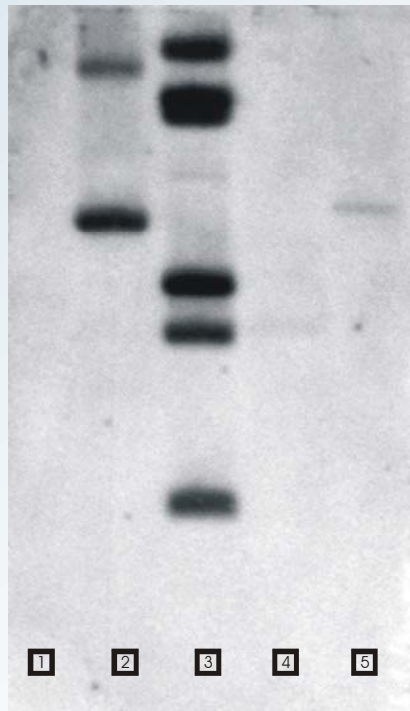
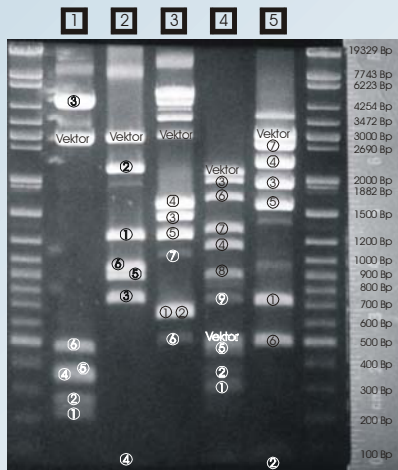


9804 KB-Fragment **5**



Alle subklonierten pKAP298-Fragmente vom Blot am Stück + detektierte Banden mit Genaktivität (grün):





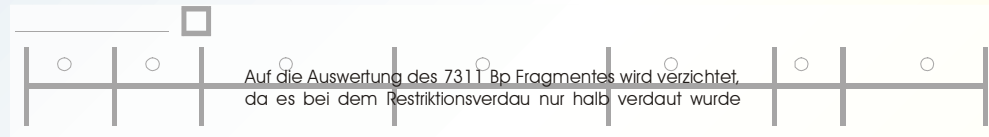
Auswertung Blot SubA "hunger 1"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün

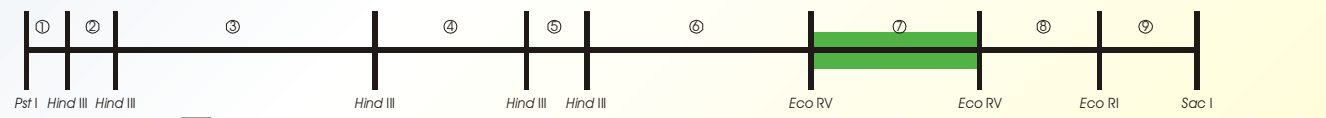
6408 Bp-Fragment **1**



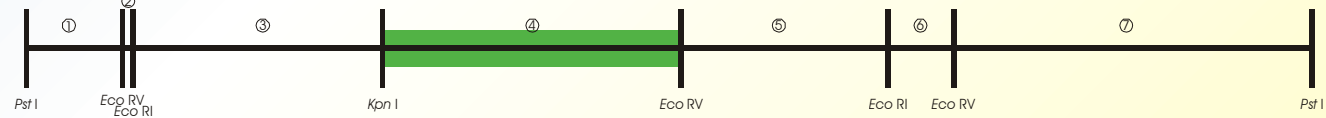
6220 Bp-Fragment **2**



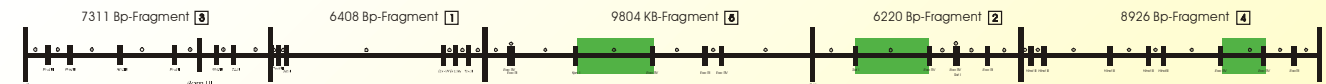
8926 Bp-Fragment **4**

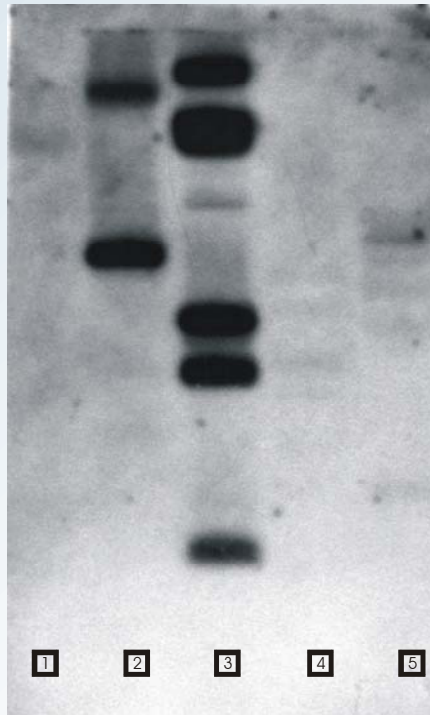
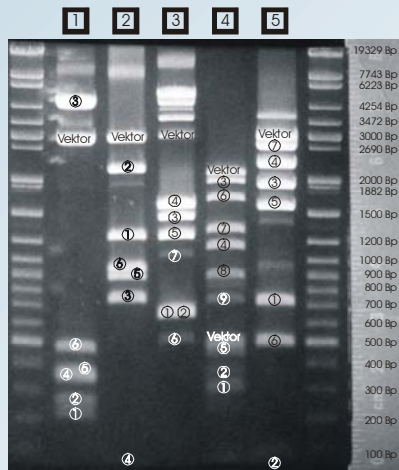


9804 KB-Fragment **5**



Alle subklonierten pKAP298-Fragmente vom Blot am Stück + detektierte Banden mit Genaktivität (grün):

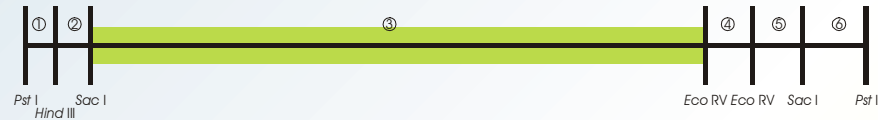




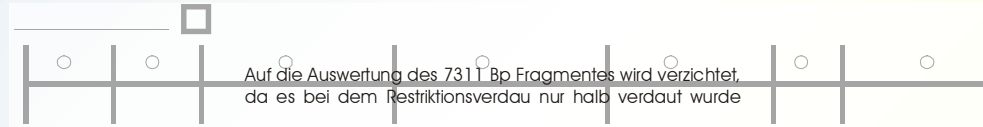
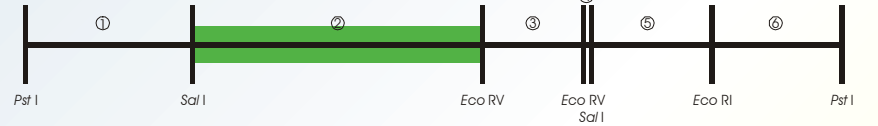
Auswertung Blot SubA "hunger 2"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün

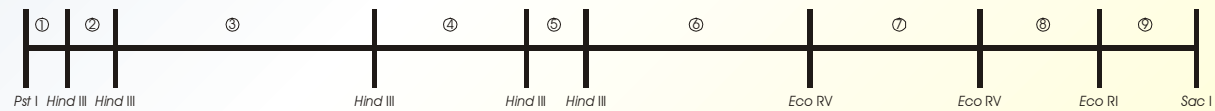
6408 Bp-Fragment **1**



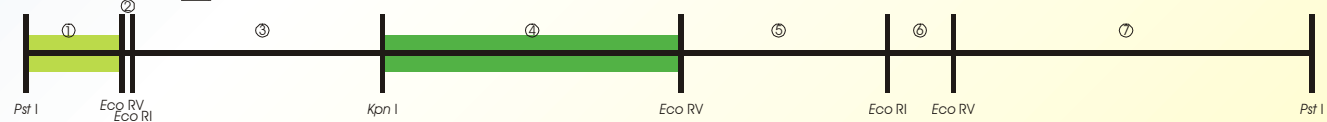
6220 Bp-Fragment **2**



8926 Bp-Fragment **4**



9804 KB-Fragment **5**

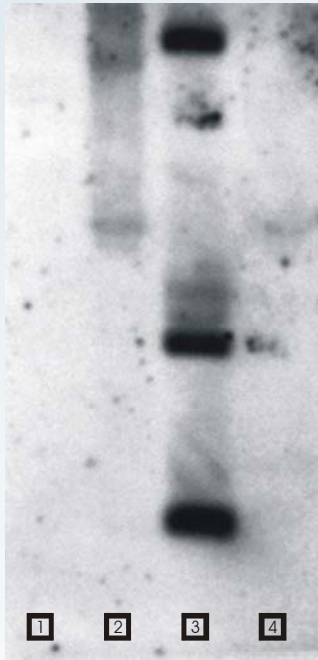
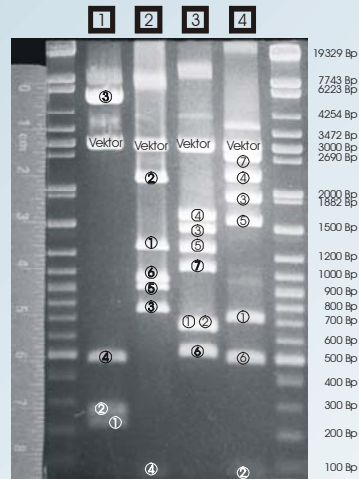


Alle subklonierten pKAP298-Fragmente vom Blot am Stück + detektierte Banden mit Genaktivität (grün):



Auswertung Blot SubB "gefüttert 1"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün



6408 Bp-Fragment **1**



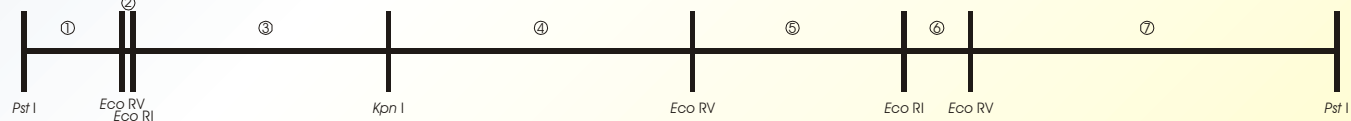
6220 Bp-Fragment **2**



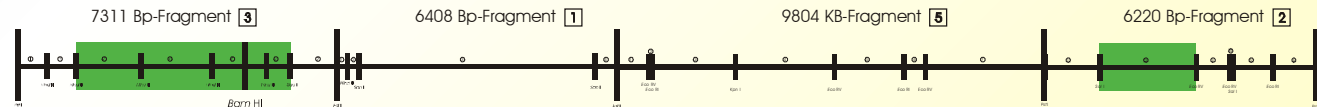
7311 Bp-Fragment **3**



9804 KB-Fragment **4**

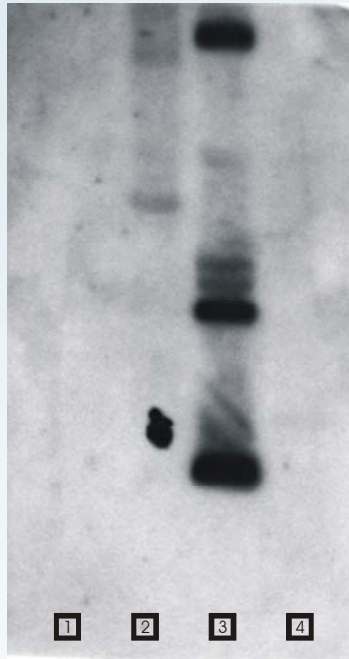
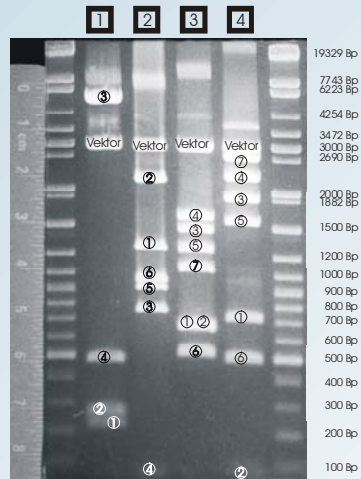


Alle subklonierten pKAP298-Fragmente vom Blot am Stück und detektierte Banden mit Genaktivität (grün):



Auswertung Blot SubB "gefüttert 2"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün



6408 Bp-Fragment **1**



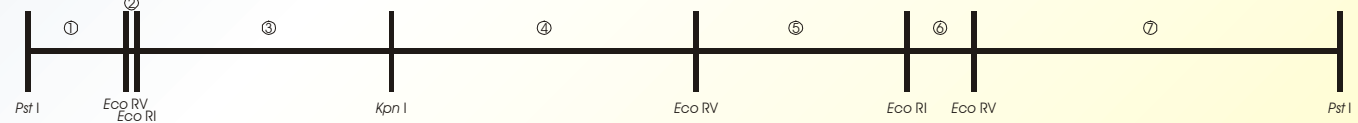
6220 Bp-Fragment **2**



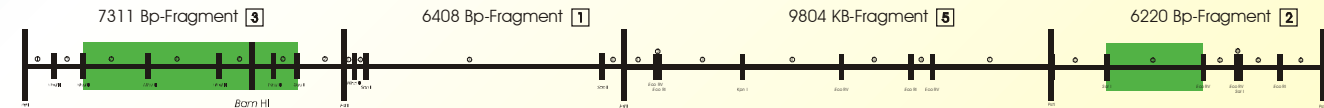
7311 Bp-Fragment **3**



9804 KB-Fragment **4**

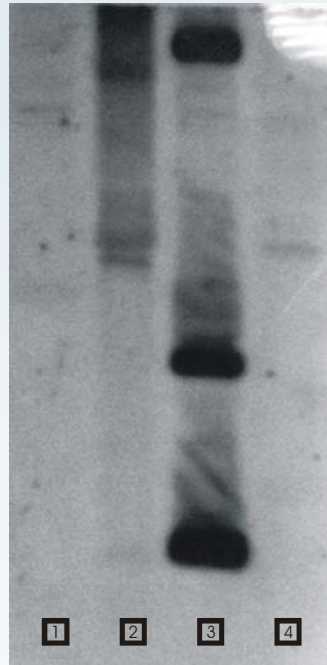
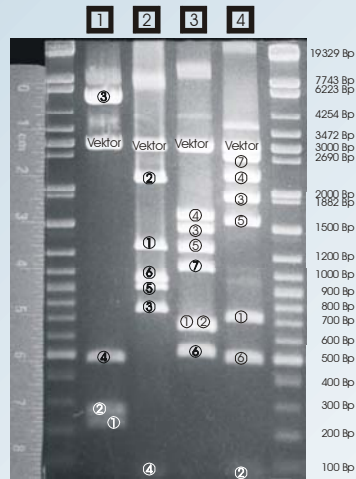


Alle subklonierten pKAP298-Fragmente vom Blot am Stück und detektierte Banden mit Genaktivität (grün):



Auswertung Blot SubB "hunger 1"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün



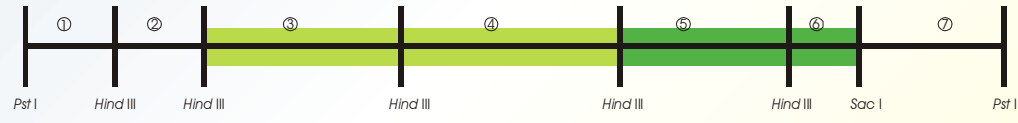
6408 Bp-Fragment **1**



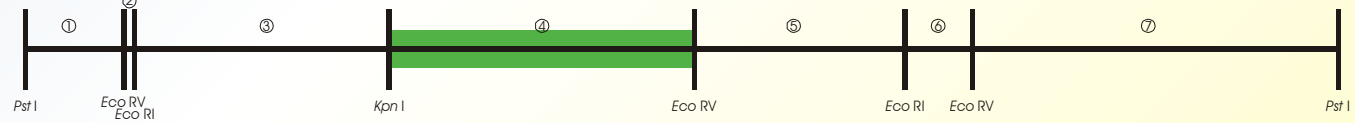
6220 Bp-Fragment **2**



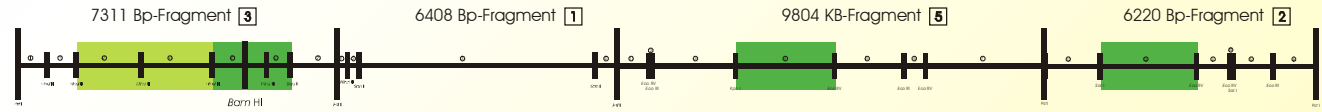
7311 Bp-Fragment **3**



9804 KB-Fragment **4**

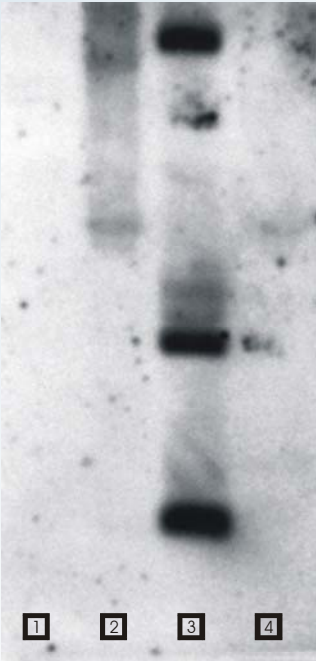
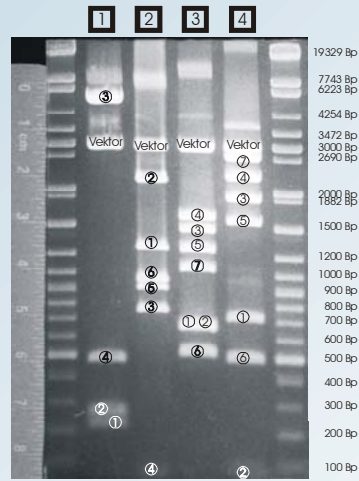


Alle subklonierten pKAP298-Fragmente vom Blot am Stück und detektierte Banden mit Genaktivität (grün):



Auswertung Blot SubB "hunger 2"

Banden, auf denen Genaktivität detektiert wurde sind grün, unsichere Banden sind hellgrün



6408 Bp-Fragment 1



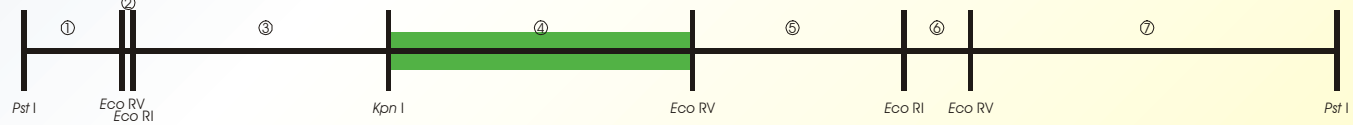
6220 Bp-Fragment 2



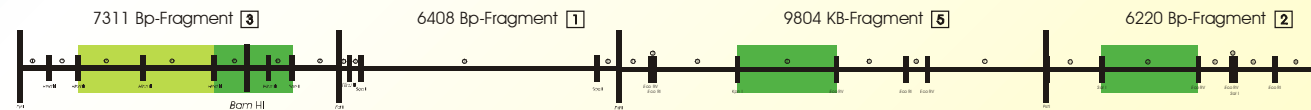
7311 Bp-Fragment 3

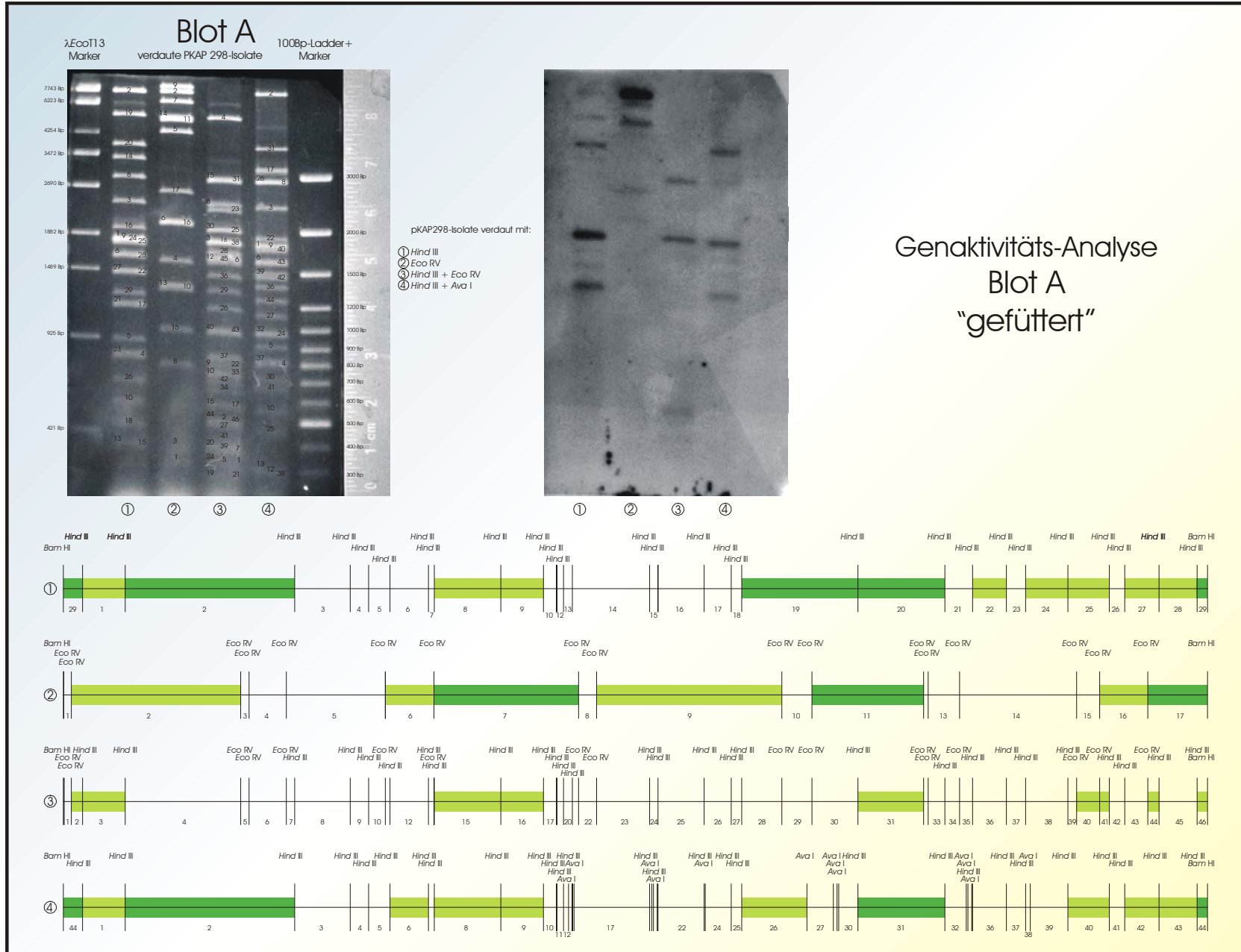


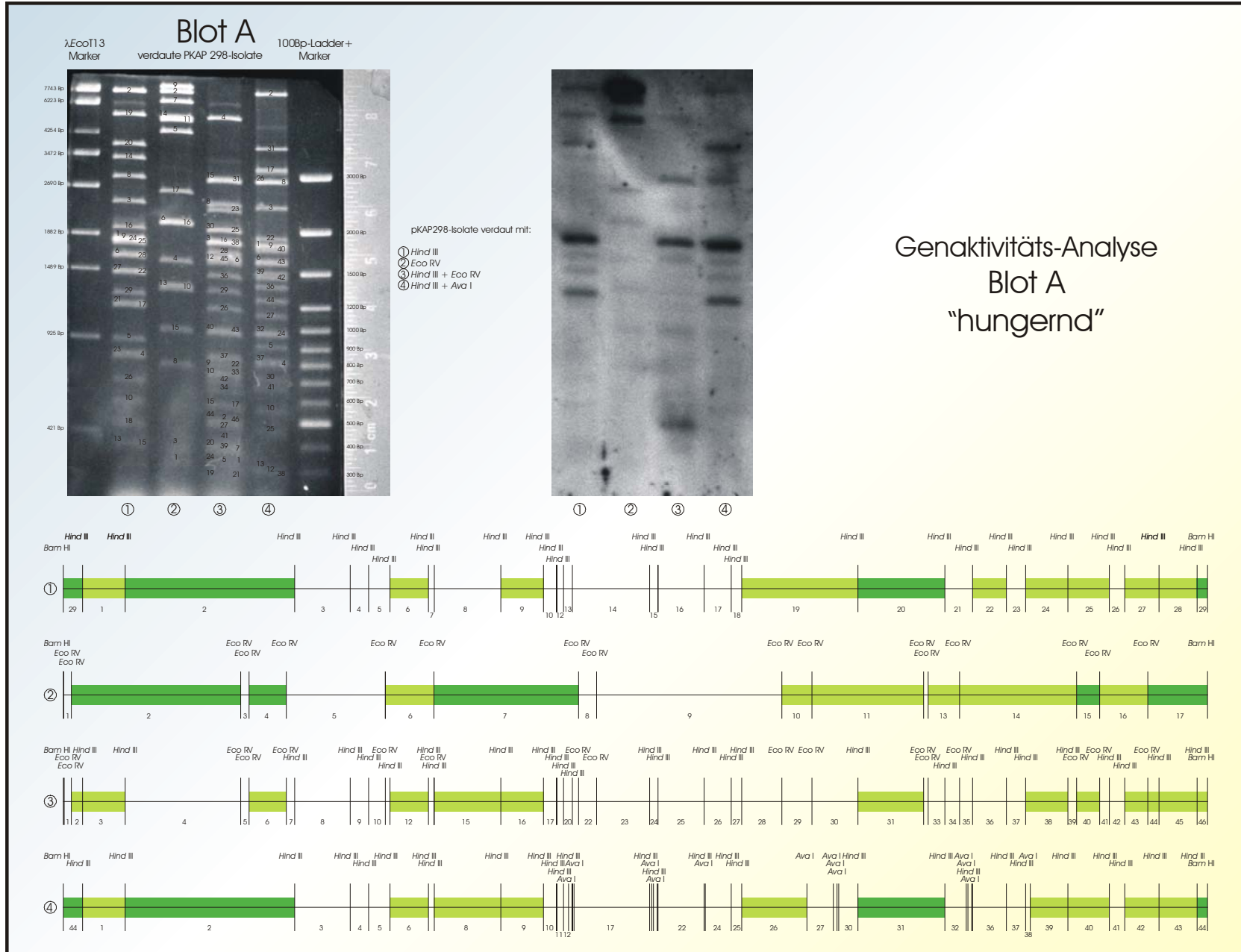
9804 KB-Fragment 4

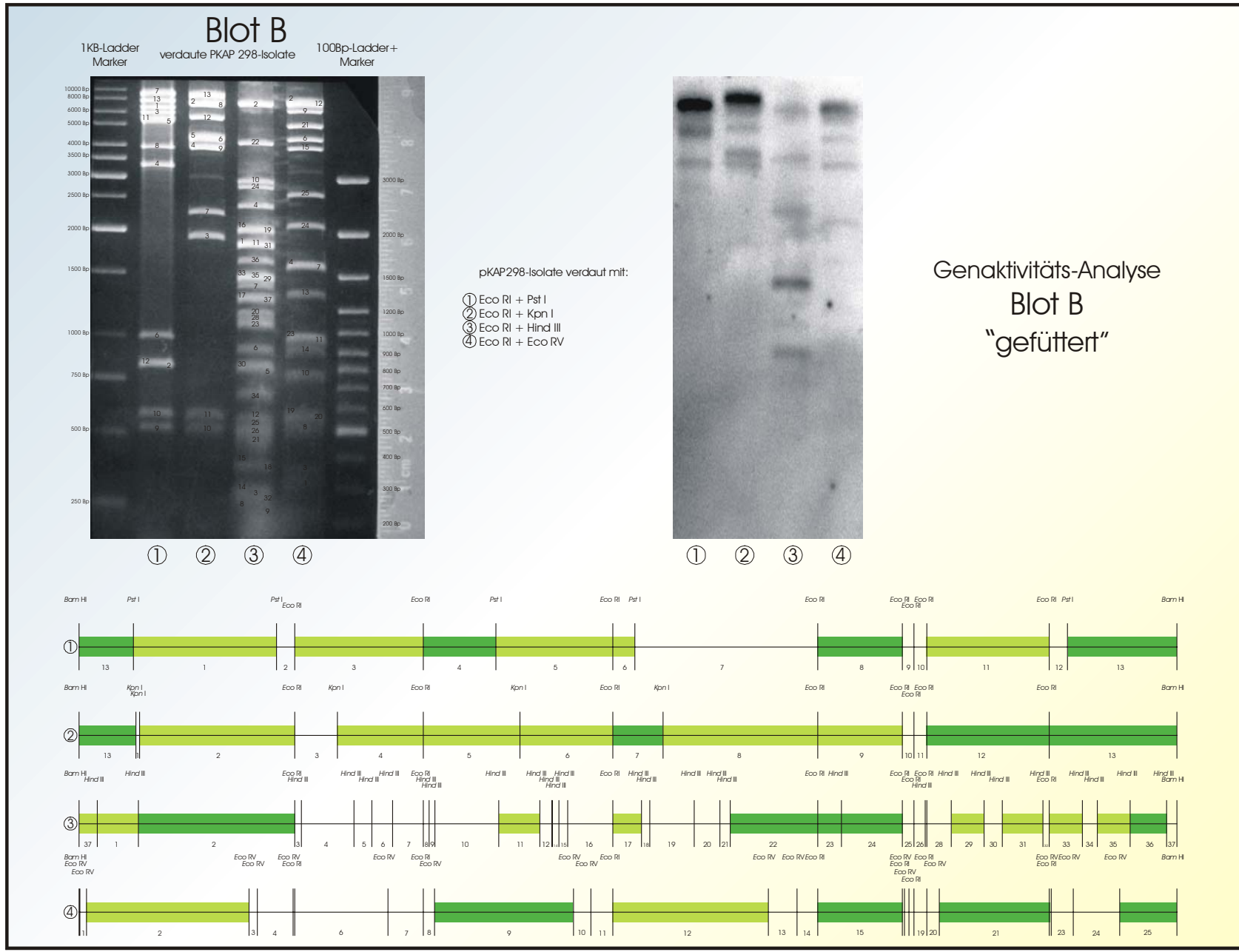


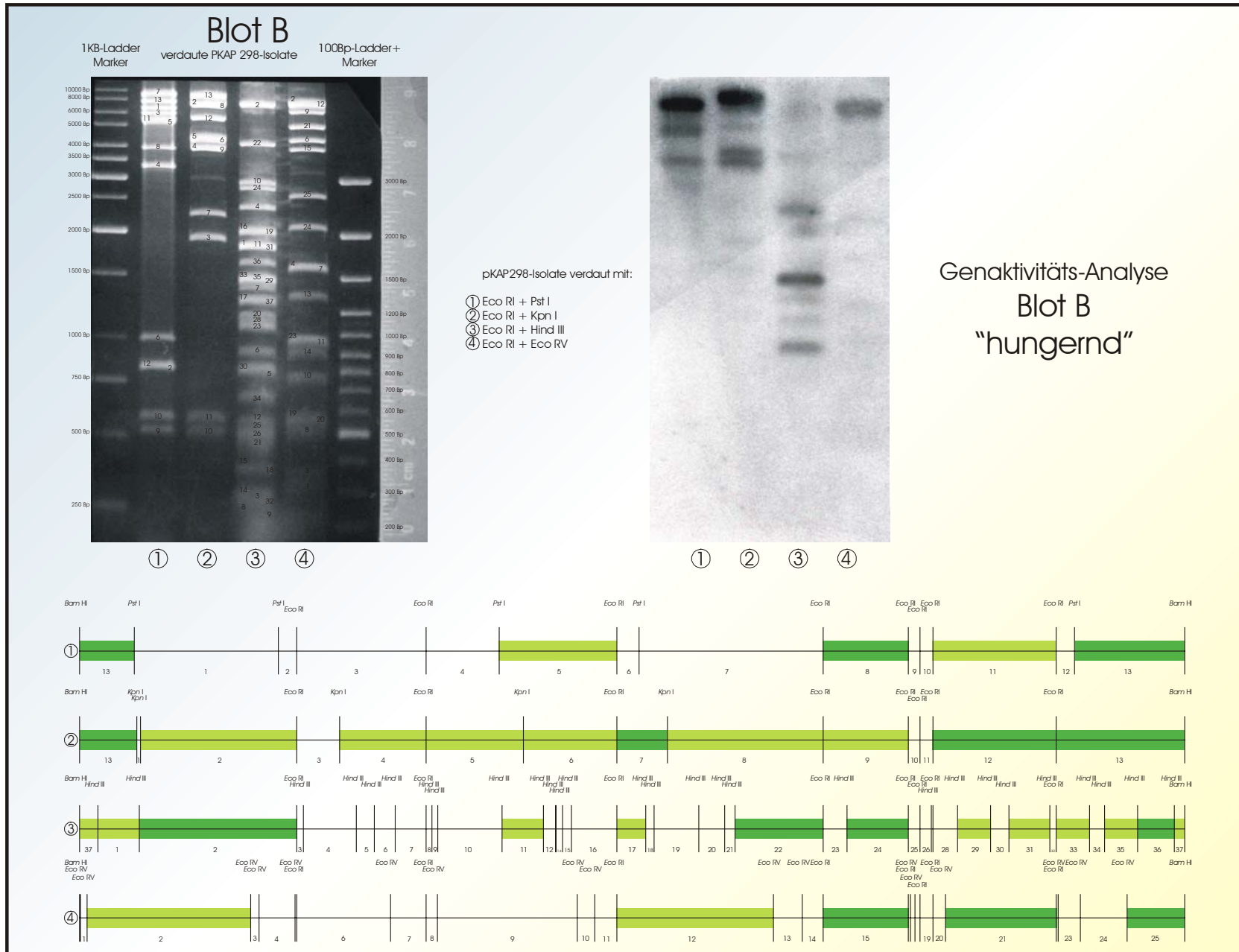
Alle subklonierten pKAP298-Fragmente vom Blot am Stück und detektierte Banden mit Genaktivität (grün):

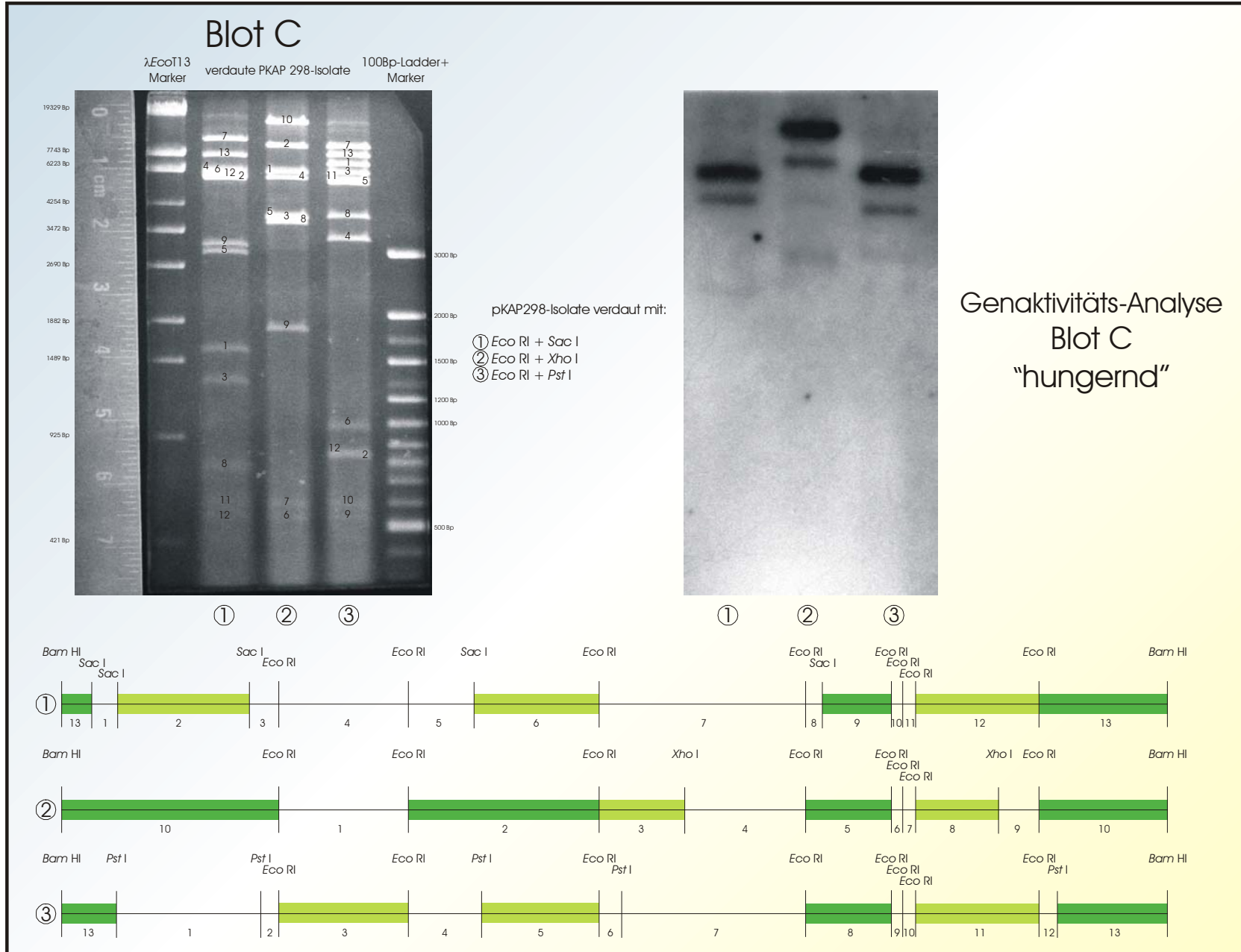




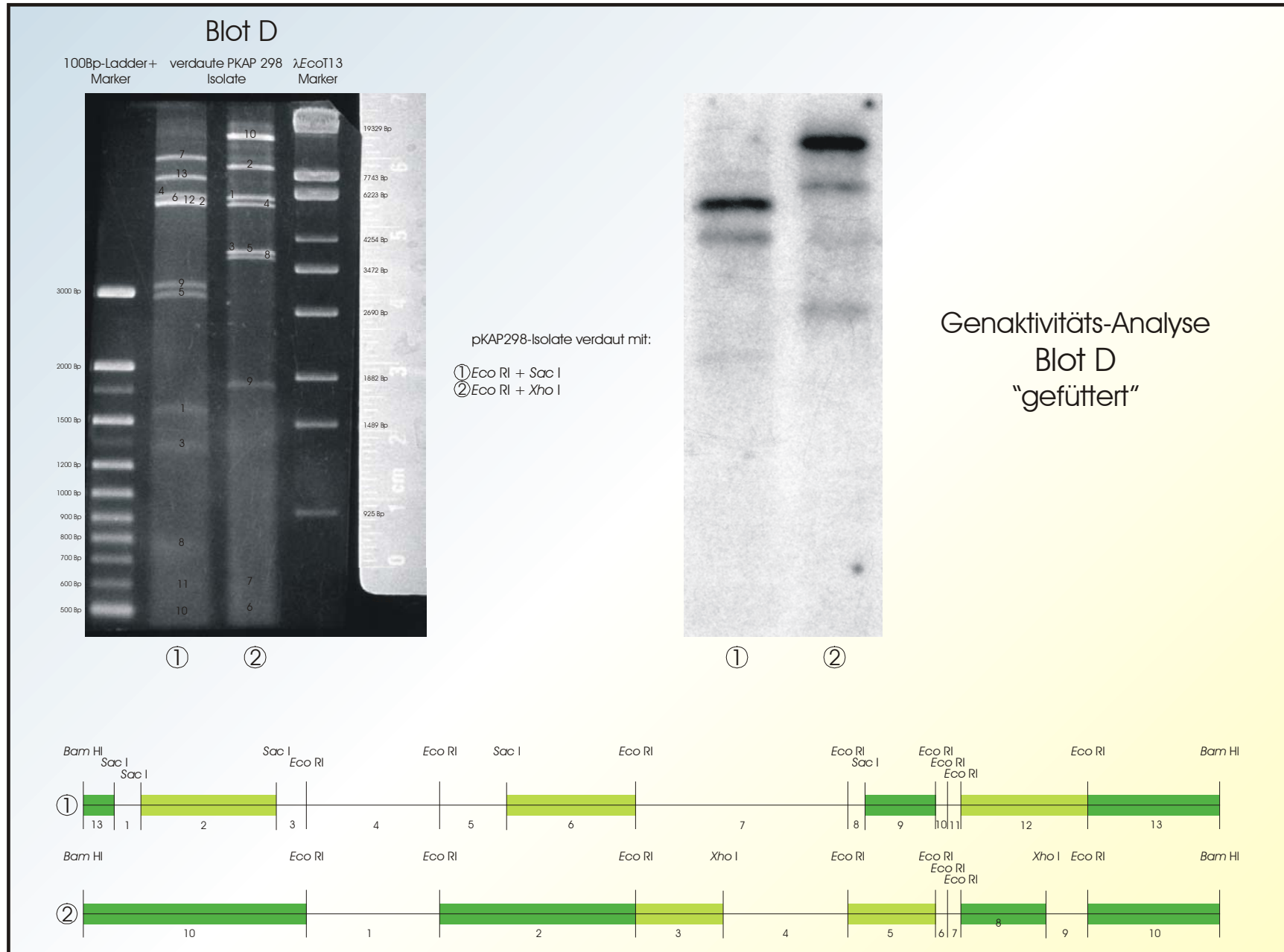


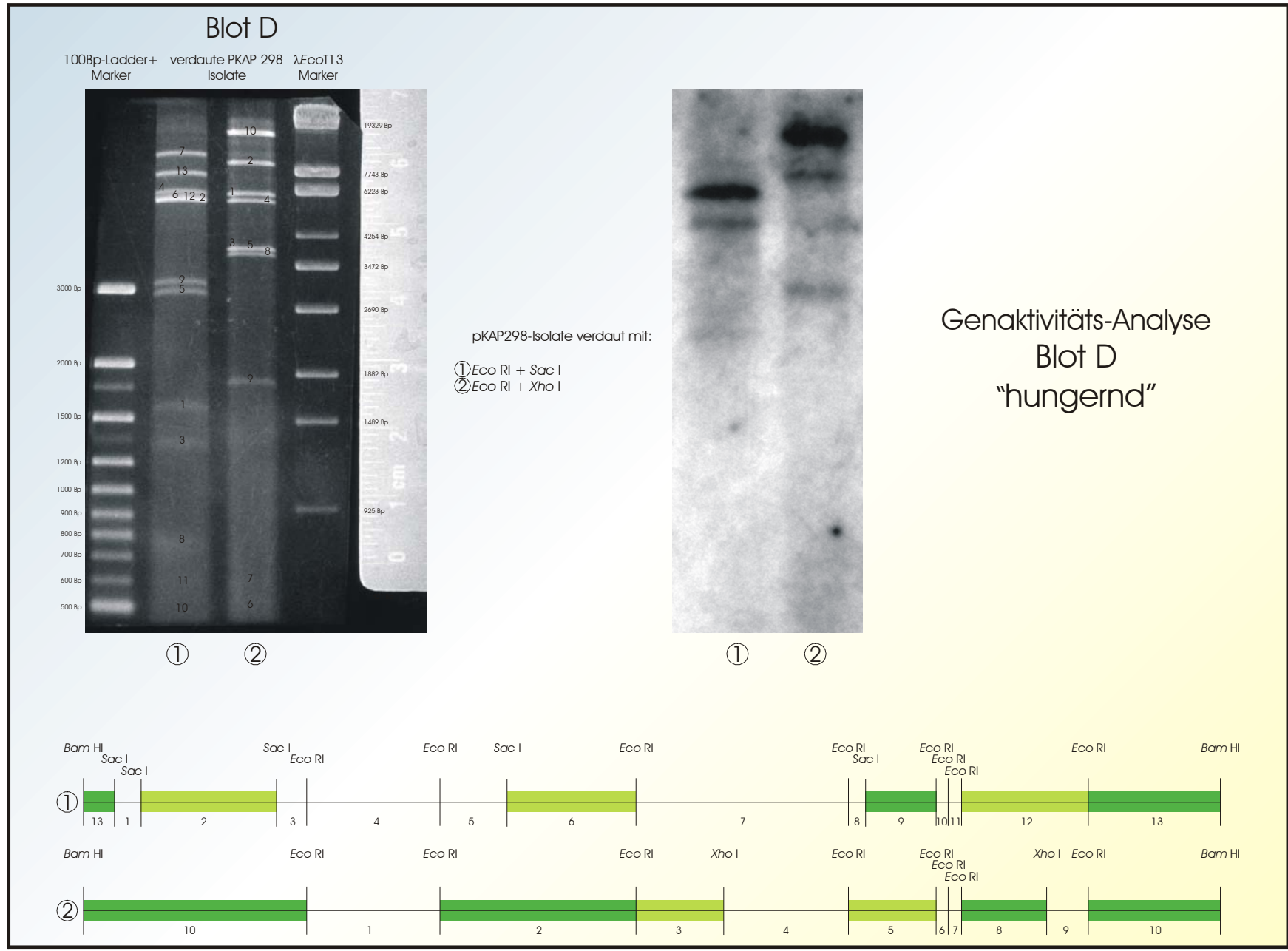


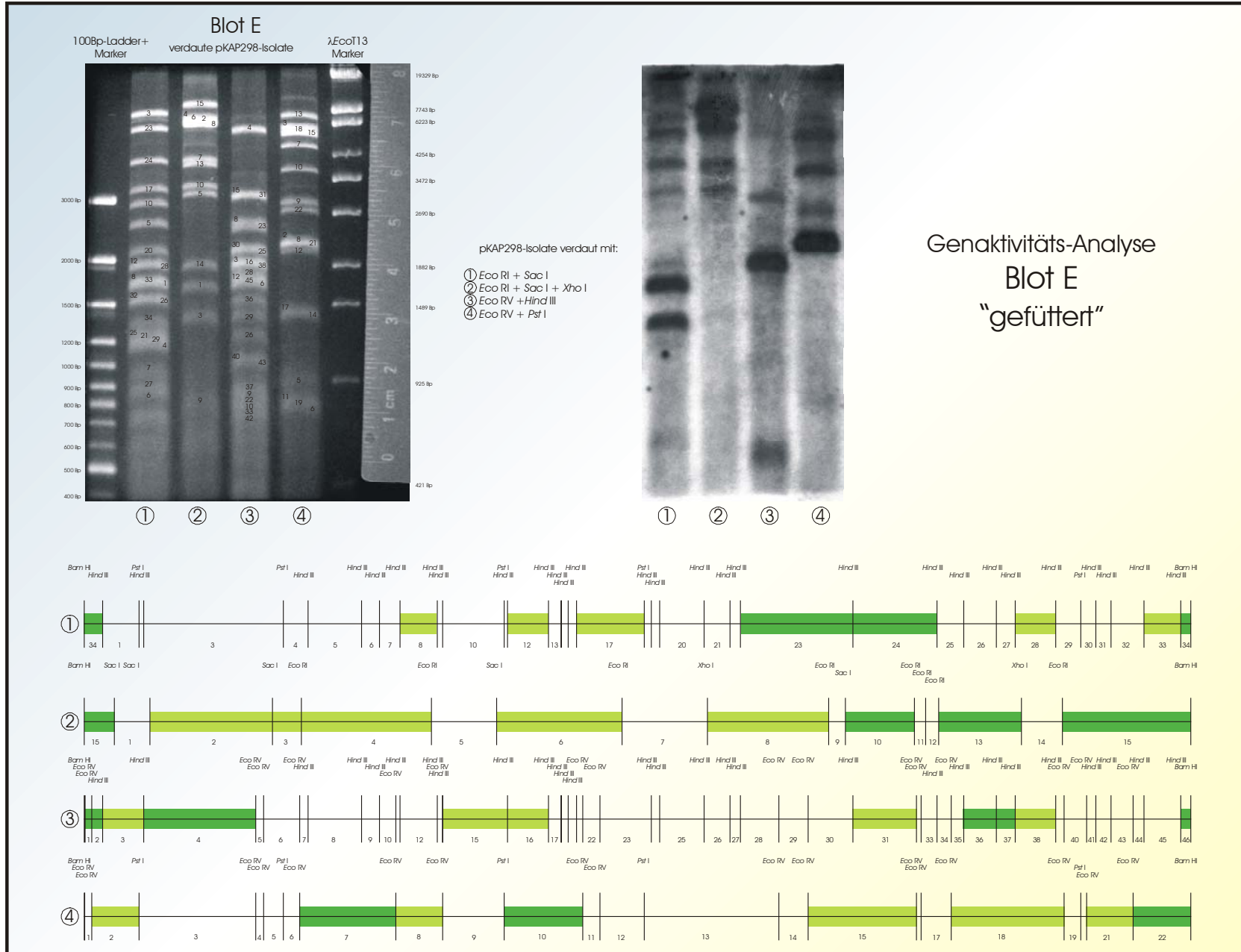


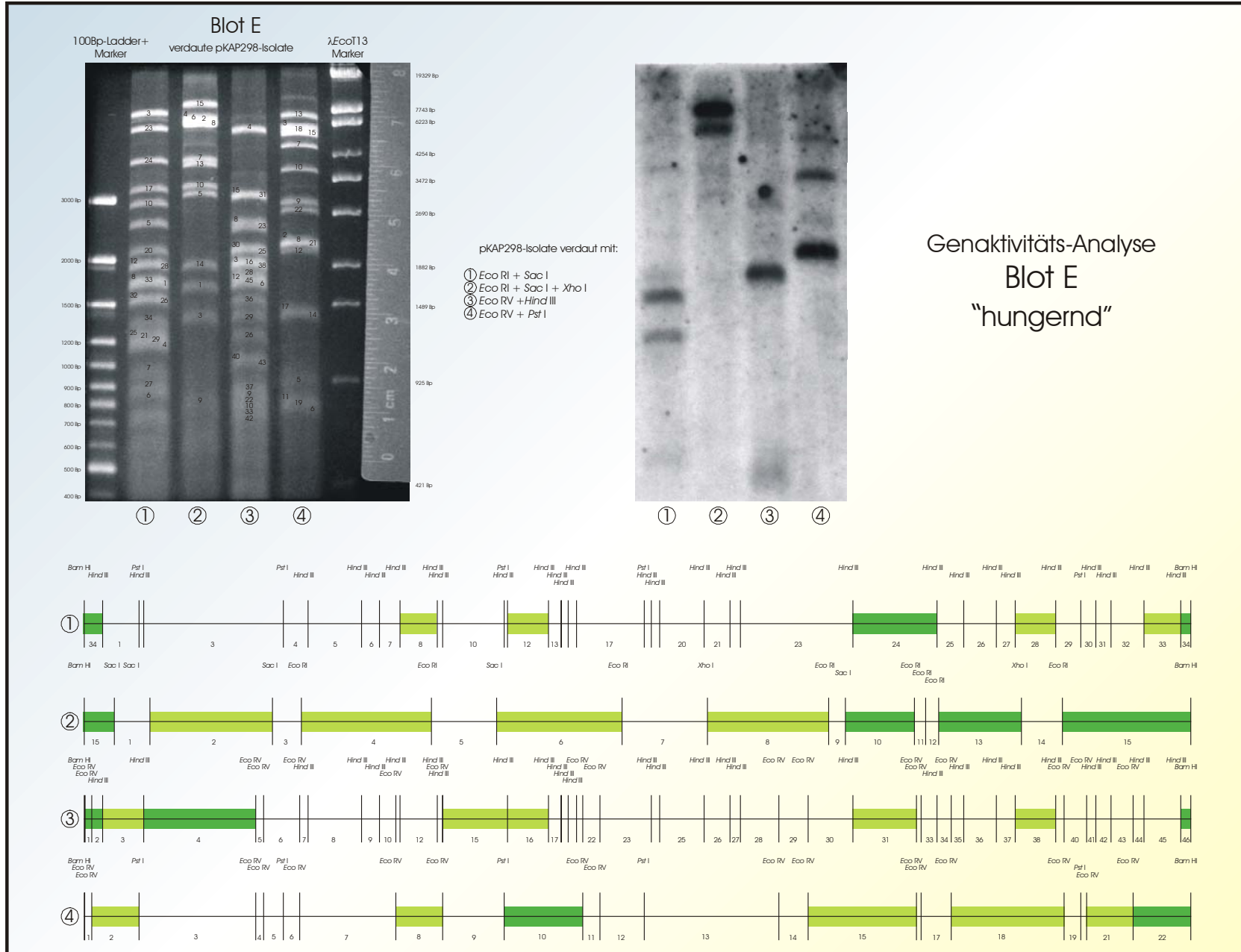


Genaktivitäts-Analyse
Blot C
"hungernd"









3.3.6.2 Erstellung einer Transkriptions-Aktivitäts-Karte des Plasmids pKAP298

Nach den im Methodenteil beschriebenen Kriterien wurde eine "Gesamt-Transkriptions-Aktivitäts-Karte" des Plasmids pKAP298 erstellt.

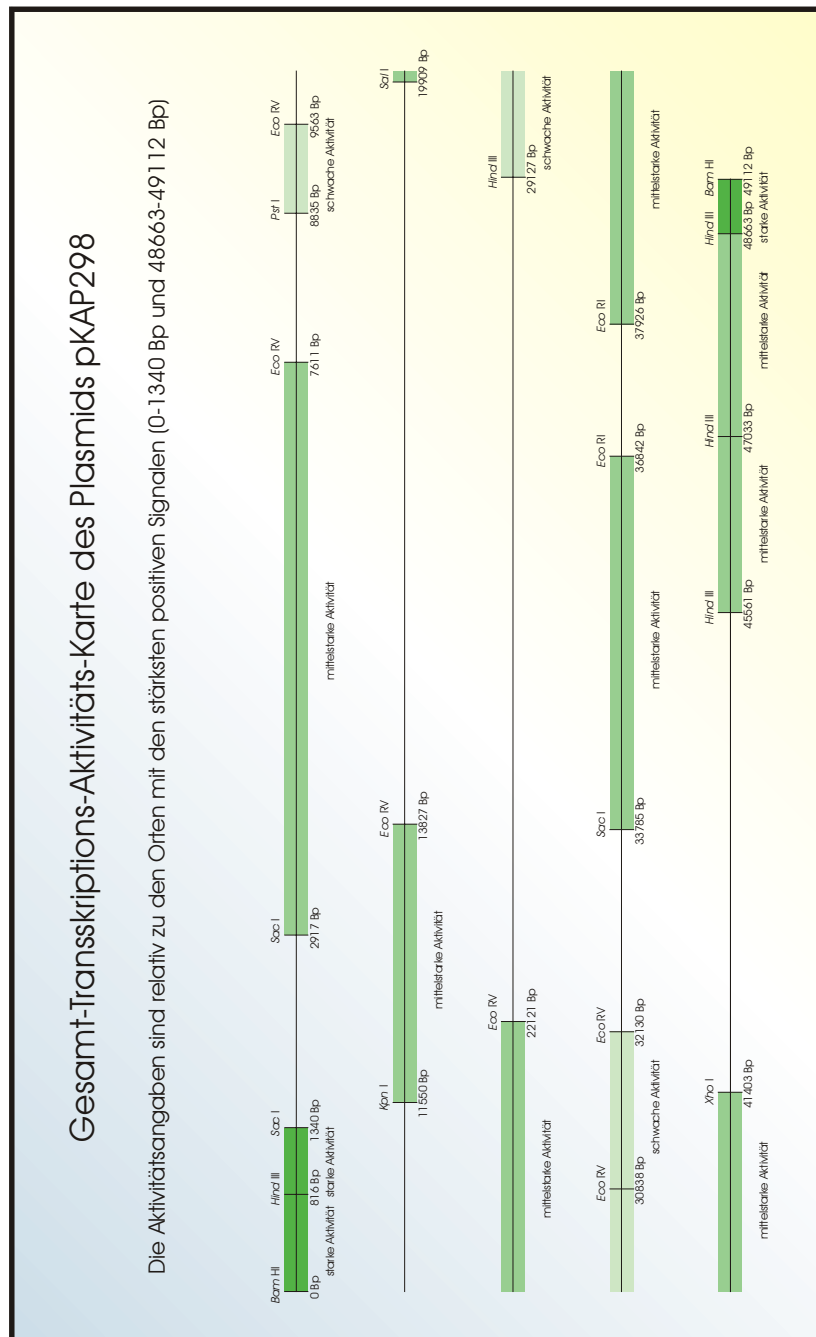


Abb. 3.3.6.2.1: Gesamt-Transkriptions-Aktivitätskarte des Plasmids pKAP298

Die vier oberen Linien entsprechen jeweils 10 kB des Gesamtplasmids, die untere nur 9 kB. In grün eingetragen sind die pKAP298-Regionen, auf denen Transkriptions-Aktivität detektiert werden konnte.

3.4 Analyse der pKAP298-Sequenz

3.4.1 Sequenzierung des Plasmids pKAP298

3.4.1.1 Eigene Sequenzierung

Da die Sequenzierung außer Haus viel kostengünstiger und zeitsparender war, wurden nur die 2 kleinsten (512 Bp und 572 Bp) der subklonierten pKAP298-Fragmente in der Abteilung Ökologie der TU Kaiserslautern sequenziert.

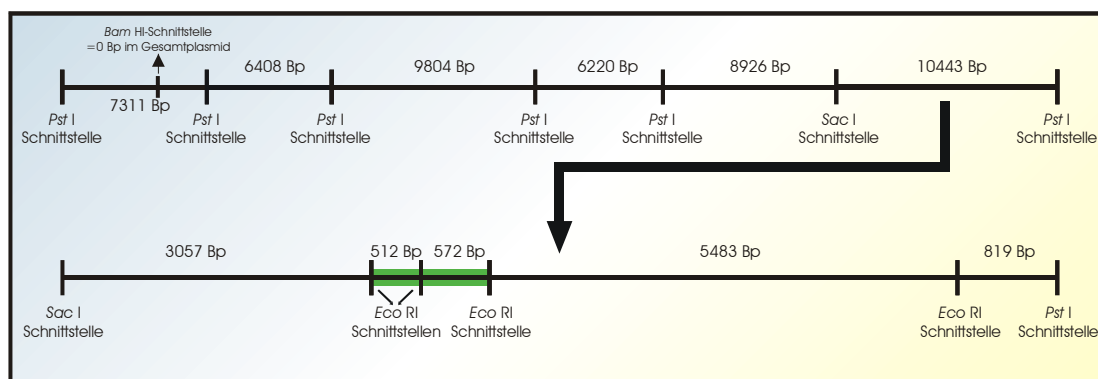


Abb. 3.4.1.1.1: selbst sequenzierte pKAP298-Fragmente (grün dargestellt)

Zunächst wurde durch Plasmid-Isolation aus den rekombinanten *E. coli* die Template-DNA gewonnen. Die Konzentration der Plasmid-DNA betrug:

- 156 ng/µl für das Fragment mit 512 Bp
- 129 ng/µl für das Fragment mit 572 Bp

Mit diesen Konzentrationen wurden jeweils 4 Template/Primer-Premixe pro Fragment angesetzt - 2 mit M13universal-forward-Primern, 2 mit M13universal-reverse-Primern. Je 4,5 µl dieser Mixe wurden mit 1,5 µl der einzelnen Basen-Reaktionslösungen gemischt und dem Sequenzier-Programm im Thermocycler unterzogen. Danach wurde ein Sequenzier-Gel bereitet, die 32 fertigen Proben aufgetragen, die Elektrophorese gestartet und die Bandenmuster am Computer detektiert. Aus dem Bild mit dem Bandenmuster wurde mit dem Programm "ImageAnalysis" die Basensequenz der einzelnen Ansätze extrahiert. Pro Fragment erhielt man so theoretisch 4 Sequenzen (2 Forward und 2 Reverse). Bei dem größeren der beiden Fragmente waren die Banden einer der 4 aufgetragenen Proben nicht auszuwerten (im SDS-Gel schief gelaufen), daher wurden für dieses Fragment statt 4 nur 3 Sequenzen ermittelt. Um die Sequenzen miteinander vergleichen zu können, mussten die

reversen Sequenzen noch revers-komplementär gemacht werden. Dies geschah mit dem Programm "Revseq". Die jeweiligen 4 bzw. 3 Sequenzen pro Fragment wurden nun in dem Programm ClustalX aligned und die Konsensus-Sequenz ermittelt. Dabei wurden eindeutige Alignment-Fehler, wie falsche Gaps, korrigiert. Es folgen die Alignments der zwei sequenzierten Fragmente. "K...." steht für das 512 Bp-Fragment, "G...." steht für das 572 Bp-Fragment.

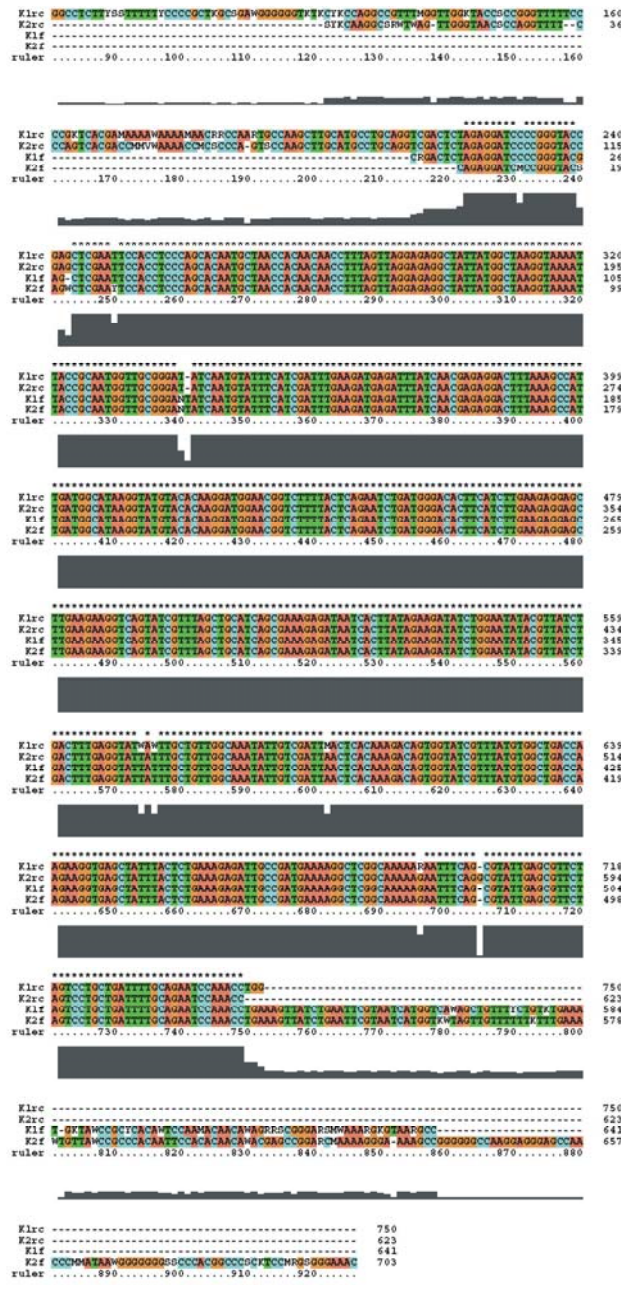


Abb. 3.4.1.1.2: ClustalX Alignment der Sequenzierung des 512 Bp Fragments

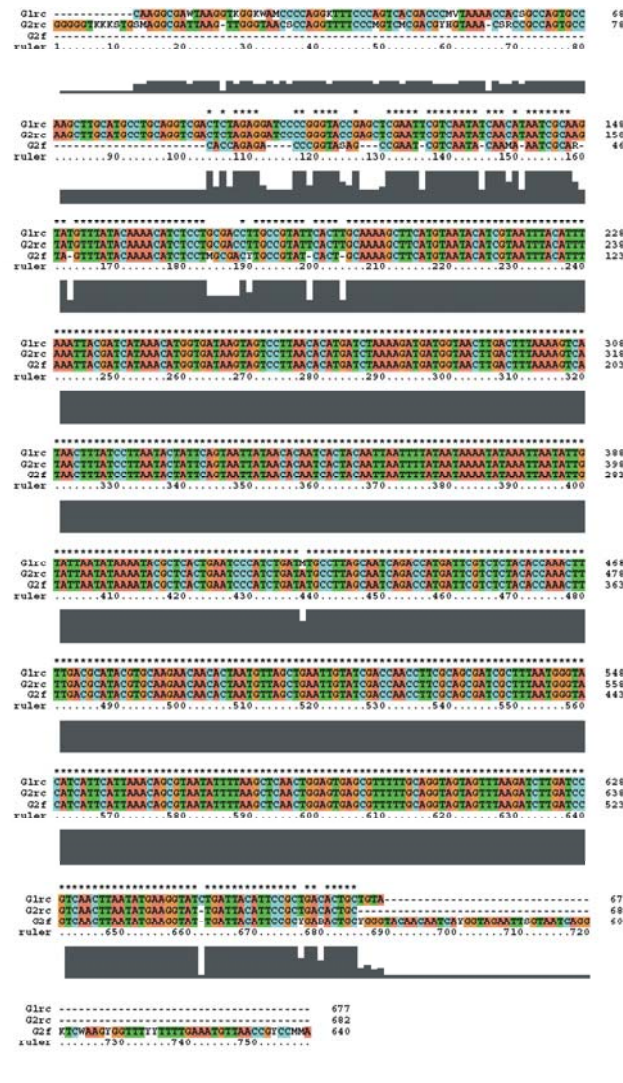


Abb. 3.4.1.1.3: ClustalX Alignment der Sequenzierung des 572 Bp Fragments

Auf der Gesamtsequenz des Plasmids pKAP298 ist das 512 Bp Stück (ausgehend von der *Bam* HI - Schnittstelle als Startpunkt) zwischen den Basen 36844 bis 37356 lokalisiert, das 572 Bp Stück zwischen den Basen 37357 und 37928.

3.4.1.2 Sequenzierung durch Seqlab

Der Firma Seqlab wurden als Template Plasmid-Isolate der subklonierten pKAP298-Fragmente zugeschickt. Zur Sequenzierung der pKAP298-Stücke, die nicht subkloniert werden konnten, wurde der Firma das ganze Plasmid pKAP298 zugeschickt und die entsprechenden flankierenden Start-Primer aus den benachbarten, schon bekannten Sequenzen geliefert. Die Sequenzierung wurde mittels single-strand Primer-Walking durchgeführt, mit einer von Seqlab angegebenen Fehler-Rate von 0,1 %. Auf 49 kB kann also

mit ca. 49 falschen Basen gerechnet werden. In der Sequenz fand sich 16 mal "N", die dazugehörige Base kann bei Bedarf mittels Doppelstrang-Sequenzierung identifiziert werden, was während dieser Arbeit nicht mehr möglich war.

3.4.1.3 Zusammensetzen der Sequenzen zum vollständigen Plasmid pKAP298

Die bei der eigenen Sequenzierung gewonnenen Sequenzen und die Basensequenzen von SeqLab wurden, wie im Methodenteil beschrieben, zusammengesetzt. Die Gesamtsequenz wird hier, aufgrund ihrer Größe (49,1 kB), nicht abgedruckt. Sie ist auf der beigefügten CD als Datei zu finden (das File "pKAP298_komplette_Sequenz.doc" oder als Textfile gleichen Namens)

3.4.1.4 Restriktionskarte des Plasmids pKAP298

Die Gesamt-Sequenz wurde mit dem Programm Webcutter virtuell (mit den in der Arbeit verwendeten Enzymen) verdaut:

pKAP298, 49112 base pairs

Table by Enzyme Name

Enzyme name	No. cuts	Positions of sites	Recognition sequence
Ava I	15	21789 21954 22027 25365 25438 25603 27667 32022 33158 33321 33394 38864 38937 39102 41605	c/ycgrg
<i>Bam</i> HI	1	3	g/gatcc
<i>Eco</i> RI	8	9646 15404 23874 33042 36844 37356 37928 43411	g/aattc
<i>Eco</i> RV	17	25 339 7613 7971 9565 13829 15907 22123 22891 30840 32132 36936 37135 38480 43497 44490 46553	gat/atc
<i>Hind</i> III	29	818 2650 9936 12309 13107 14019 15668 15917 18786 20608 21161 21182 21478 21856 25169 25536 27513 28670 29129 34112 37851 39035 40490 41316 43121 44895 45563 47035 48665	a/agctt
<i>Kpn</i> I	5	2532 2703 11552 19717 26120	ggtac/c
<i>Pst</i> I	5	2429 8837 18641 24861 44230	ctgca/g
<i>Sac</i> I	5	1342 2919 8350 18306 33787	gagct/c
<i>Sal</i> I	2	19911 22951	g/tcgac
<i>Xho</i> I	2	27667 41605	c/tcgag

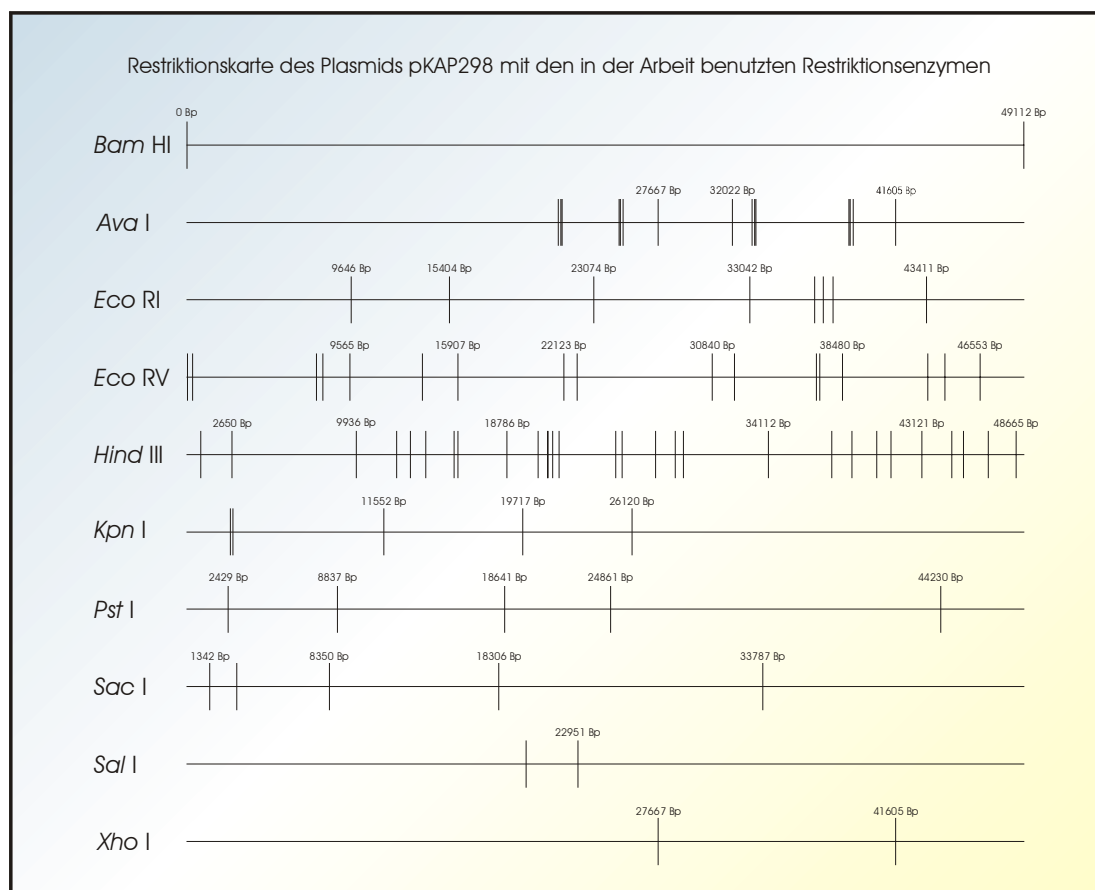


Abb. 3.4.1.4.1: Grafische Darstellung der Restriktions-Schnittstellen auf pKAP298

Einzelne Schnittstellen sind zur Orientierung beschriftet, die genaue Lage der unbeschrifteten Schnittstellen kann der Tabelle darüber entnommen werden (direkt aus Webcutter)

3.4.2 Bioinformatische Analyse der pKAP298-Sequenz

Mittels des Programms GeneMark.hmm konnten 63 potentiell codierende ORFs auf dem Plasmid identifiziert werden. Laut den Programmierern (LUKASHIN UND BORODOVSKI, 1998) ist das Programm in der Lage, durchschnittlich 94,6 % der potentiell kodierenden Bereiche auf einer beliebigen prokaryotischen DNA-Sequenz zu identifizieren. Die durchschnittliche Rate von nicht identifizierten ORFs beträgt 5,4 %, was im Fall des bearbeiteten Plasmids zwischen 3 und 4 nicht identifizierte, potentiell codierende ORFs wären. Als Prozentsatz für falsche positive ORFs werden ungefähr 10 % angegeben. Auf die 63 identifizierten Bereiche gerechnet ergibt das zwischen 6 und 7 falsche positiven ORFs. Eine Überprüfung der Ergebnisse über GetORF und BLAST ergab, dass GeneMark.hmm alle ORFs mit signifikanten Sequenzhomologien identifiziert hat. Für diese Überprüfung wurden aus der pKAP-Sequenz mittels GetORF (<http://bioweb.pasteur.fr/seqanal/interfaces/getorf.html>) alle ORFs ab einer Größe von 99 Bp translatiert extrahiert (bacterial code) und mittels BLAST auf Ähnlichkeiten zu schon

bekannten Proteinen untersucht. Diese Ergebnisse wurden dann mit den aus GeneMark.hmm erhaltenen Daten verglichen. Ob die putativen Gene wirklich mit dem von GeneMark.hmm angegebenen Startkodon beginnen, müsste experimentell geklärt werden (sofern sie überhaupt exprimiert werden). Die von GeneMark.hmm ermittelten Daten und die Grafische Auswertung der Sequenz sind als Dateien auf der beigefügten CD abrufbar ("pKAP_GeneMark_hmm_Ergebnisse.htm" und "pKAP_GeneMark_hmm_grafische_Auswertung.pdf" im Ordner "Genemark_Ergebnisse")

Von den 63 identifizierten, möglicherweise codierenden ORFs beginnen 56 mit dem Startkodon ATG, 4 mit GTG und 3 mit TTG.

Der Gesamt-GC-Gehalt des Plasmids pKAP298 beträgt 37 % (ermittelt mit dem Programm GeeCee aus dem EMBOSS-Programmpaket, <http://bioweb.pasteur.fr/seqanal/interfaces/geecce.html>).

3.4.2.1 Potentiell codierende ORFs mit Nucleotid-Sequenzhomologien

Die Nucleotid-Sequenz der identifizierten ORFs wurde mit BLAST auf Ähnlichkeiten zu schon bekannten Sequenzen durchsucht. Signifikante Sequenzhomologien konnten nur mit schon sequenzierten ORFs von anderen Plasmiden der pKAP-Familie nachgewiesen werden. Diese ORFs werden (translatiert) in 3.4.2.2 genauer besprochen. Die erhaltenen BLAST-Daten sind auf der beigefügten CD zu finden (im Ordner "Genemark_Ergebnisse\GeneMark_ORFs_Nucleotid-Blast_Ergebnisse", als HTML-Files).

3.4.2.2 Schon beschriebene translatierte ORFs von anderen pKAP-Plasmiden

Auf den Plasmiden pKAP116 und pKAP47 wurden schon Gene beschrieben (HERUTH ET AL., 1994, 1996). Hierbei handelt es sich um eine putative Transposase (von dem Transposon Tn4501) und vier bei der R-Körper-Synthese beteiligte Proteine. Auf dem hier untersuchten Plasmid pKAP298 wurden die entsprechenden ORFs gesucht und mit den schon von den anderen pKAP bekannten Sequenzen verglichen.

3.4.2.2.1 ORF 53 [von 42889 Bp- 44173 Bp] (REVERSE SENSE)

Hierbei handelt es sich um eine schon von HERUTH ET AL. (1996) beschriebene Transposase vom Transposon Tn4501 (siehe 3.4.2.7.1).

BLAST-Ähnlichkeiten:

```

gi|1438744|gb|AAB04101.1| transposase Tn4501
Length = 240

Score = 489 bits (1259), Expect = e-137
Identities = 240/240 (100%), Positives = 240/240 (100%)

Query: 185 MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAELLRDIKRRGLVAGPELATGDGA 244
           MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAELLRDIKRRGLVAGPELATGDGA
Sbjct: 1    MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAELLRDIKRRGLVAGPELATGDGA 60

Query: 245 MGFWSALEKEFPKTCQRCWCWVHKTANILNKLPKAMQSRKSKLHNMYPGKEQAEAEYK 304
           MGFWSALEKEFPKTCQRCWCWVHKTANILNKLPKAMQSRKSKLHNMYPGKEQAEAEYK
Sbjct: 61  MGFWSALEKEFPKTCQRCWCWVHKTANILNKLPKAMQSRKSKLHNMYPGKEQAEAEYK 120

Query: 305 EFMVATYARYPKAVECLTKDKERMFTFFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG 364
           EFMVATYARYPKAVECLTKDKERMFTFFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG
Sbjct: 121 EFMVATYARYPKAVECLTKDKERMFTFFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG 180

Query: 365 CYSRETLCAFFKRLARQAETWSRLRGHQRLAEVINMVKFIDGISEHEVDKIKQDKQNA 424
           CYSRETLCAFFKRLARQAETWSRLRGHQRLAEVINMVKFIDGISEHEVDKIKQDKQNA
Sbjct: 181 CYSRETLCAFFKRLARQAETWSRLRGHQRLAEVINMVKFIDGISEHEVDKIKQDKQNA 240

```

Beide Enzyme sind in ihrer Sequenz identisch, wobei allerdings ORF 53 (laut Genemark.hmm) an einem anderen Startkodon beginnt, und damit größer als die Transposase des Transposons Tn4501 aus pKAP116 und pKAP47 ist.

3.4.2.2.2 ORFs 1-3 und 63

Bei diesen ORFs handelt es sich um die bei der R-Körper-Synthese beteiligten Proteine RebA, RebB, RebC und RebD (HERUTH ET AL., 1994).

ORF 1 [von 62 Bp- 307 Bp] (REVERSE SENSE) = RebD

```

gi|436873|gb|AAA73411.1| RebD
Length = 82

Score = 105 bits (262), Expect = 2e-22
Identities = 76/82 (92%), Positives = 76/82 (92%), Gaps = 1/82 (1%)

Query: 1  MADNAITPNLYVNPQVTDVSTSN-MMNVI GMAPAQAMGVLYQGVASTANLAIQNAQSSSQ 59
           MADNAITPNLYVNPQVTD      MMNVI GMAPAQAMGVLYQGVASTANLAIQNAQSSSQ
Sbjct: 1  MADNAITPNLYVNPQVTDVDFCHFQIMMNVI GMAPAQAMGVLYQGVASTANLAIQNAQSSSQ 60

Query: 60  QLNQIGQAVTSVACERIMKMMG 81
           QLNQIGQAVTSVACERIMKMMG
Sbjct: 61  QLNQIGQAVTSVACERIMKMMG 82

```

ORF 2 [von 590 Bp- 907 Bp] (REVERSE SENSE) = RebB

```

gi|436871|gb|AAA73409.1| RebB
Length = 105
Score = 117 bits (293), Expect = 4e-26
Identities = 105/105 (100%), Positives = 105/105 (100%)

Query: 1  MSNVNSQITDSVTQTNTKILGEMPAFTTGSMLQMATQAAGLSIQNSVTNQQSNMLHQAS 60
           MSNVNSQITDSVTQTNTKILGEMPAFTTGSMLQMATQAAGLSIQNSVTNQQSNMLHQAS
Sbjct: 1  MSNVNSQITDSVTQTNTKILGEMPAFTTGSMLQMATQAAGLSIQNSVTNQQSNMLHQAS 60

Query: 61  TTQGMSILYSVDTAANAQAIGSVNRSNDTSRLTDALAVIKAAKNG 105
           TTQGMSILYSVDTAANAQAIGSVNRSNDTSRLTDALAVIKAAKNG
Sbjct: 61  TTQGMSILYSVDTAANAQAIGSVNRSNDTSRLTDALAVIKAAKNG 105

```

ORF 3 [von 988 Bp- 1332 Bp] (REVERSE SENSE) = RebA

```

gi|436870|gb|AAA73408.1| RebA
Length = 114
Score = 119 bits (299), Expect = 8e-27
Identities = 113/114 (99%), Positives = 113/114 (99%)

Query: 1  MATTTSTADVGNTTDIVNSQITDSVTQTNTMVLGSTPAQIMTNLMQMSAQANGLAMQNAV 60
           MATTTSTADVGNTTDIVNSQITDSVTQTNTMVLGSTPAQIMTNLMQMSAQAN LAMQNAV
Sbjct: 1  MATTTSTADVGNTTDIVNSQITDSVTQTNTMVLGSTPAQIMTNLMQMSAQANALAMQNAV 60

Query: 61  VGQKQTNMLSDAATTQGLSIMYTLAASDGGQAVSTVDNSTDMAGLIDALGVDKA 114
           VGQKQTNMLSDAATTQGLSIMYTLAASDGGQAVSTVDNSTDMAGLIDALGVDKA
Sbjct: 61  VGQKQTNMLSDAATTQGLSIMYTLAASDGGQAVSTVDNSTDMAGLIDALGVDKA 114

```

ORF 63 [von 48821 Bp- 49165 Bp] (REVERSE SENSE) = RebC

```
gi|436872|gb|AAA73410.1|   RebC
Length = 56
Score = 95.5 bits (236), Expect = 2e-19
Identities = 55/55 (100%), Positives = 55/55 (100%)

Query: 1  MADIEHFSEDIQRMMDPKNNVASQGAYDIVSMAAASAVQNELAMHSGFETTSNT 55
          MADIEHFSEDIQRMMDPKNNVASQGAYDIVSMAAASAVQNELAMHSGFETTSNT
Sbjct: 1  MADIEHFSEDIQRMMDPKNNVASQGAYDIVSMAAASAVQNELAMHSGFETTSNT 55
```

Bis auf RebD, das sich von seinen Pendants aus pKAP116 und pKAP47 in 6 Aminosäuren unterscheidet, sind die R-Körper kodierenden Gene nahezu identisch - nur bei RebA findet man einen Austausch einer einzigen Aminosäure. Dieser Fakt könnte wichtig sein, da dem *RebD*-Gen eine wichtige Rolle bei der Toxinbildung zugeschrieben wird (POND, 2002, unveröffentlicht, siehe 4.3).

3.4.2.3 Noch nicht beschriebene potentiell kodierende ORFs mit Aminosäure-Sequenzhomologien

Es folgt die Auflistung von translatierten, potentiell kodierenden ORFs mit signifikanten Ähnlichkeiten zu schon bekannten Proteinen. Es wird das translatierte hypothetische Protein und die von Protein-BLAST entdeckten Ähnlichkeiten aufgelistet. Dazu werden noch die vom Programm GeneMark errechneten Wahrscheinlichkeiten, dass der jeweils beschriebene ORF für eine Protein kodiert, wiedergeben (zwischen 0,5 und 1, der Treshold betrug 0,5). Die von Protein-BLAST erhaltenen Treffer werden, aufgrund des Umfangs, nur verkürzt wiedergegeben. Die kompletten BLAST-Ergebnisse sind auf der beiliegenden CD als HTML-Files gespeichert (im Ordner "Genemark_Ergebnisse\GeneMark_ORFs_Protein-Blast_Ergebnisse").

3.4.2.3.1 ORF 7 [von 5654 Bp- 6235 Bp] (REVERSE SENSE)

Translation:

```
>ORF7_Translation: 5654..6235 (reverse), 194 amino acids
LPSNKIDVKLIKDHWDVILRLIASIKSRDFRSTILKRLSSYDKEHALQEAIQEFGKIIK
SIFILNYIDSVELRQSIKQLNKGELSNKFSKAVSFAPNQEINEAEREDQKETAACKMII
QNAILWNYIELTMMILNEKDEDKRKEMLANIKESSILTWRHVNMHGTYDFSKLIASNEP
KMTAQEILDAEVA*
```

Konservierte Domänen:

[gnl|CDD|6485](#), pfam01526, Transposase 7, Transposase. This family includes transposases of Tn3, Tn21, Tn1721, Tn2501, Tn3926 transposons from E-coli. The specific binding of the Tn3 transposase to DNA has been demonstrated. Sequence analysis has suggested that the invariant triad of Asp689, Asp765, Glu895 (numbering as in Tn3) may correspond to the D-D-35-E motif previously implicated in the catalysis of numerous transposases.
CD-Length = 968 residues, **only 17.4% aligned**
Score = 139 bits (353), Expect = 1e-34

[gnl|CDD|13788](#), COG4644, COG4644, Transposase and inactivated derivatives, TnpA family [DNA replication, recombination, and repair]
CD-Length = 323 residues, **only 44.9% aligned**
Score = 96.9 bits (241), Expect = 1e-21

Protein-Sequenz-Ähnlichkeiten:

[gi|2117143|emb|CAA71190.1](#) site-specific recombinase [*Escherichia coli*]
 Length = 232
 Score = 114 bits (285), Expect = 9e-25
 Identities = 61/171 (35%), Positives = 99/171 (57%), Gaps = 5/171 (2%)

[gi|24376247|ref|NP_720355.1](#) ISSod9, transposase [*Shewanella oneidensis* MR-1]
 Length = 988
 Score = 114 bits (284), Expect = 1e-24
 Identities = 64/166 (38%), Positives = 97/166 (58%), Gaps = 5/166 (3%)

[gi|14279038|dbj|BAB59051.1](#) transposase [*Pseudomonas putida*]
 Length = 988
 Score = 112 bits (281), Expect = 3e-24
 Identities = 61/166 (36%), Positives = 98/166 (59%), Gaps = 5/166 (3%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,66

Mögliche Funktion des Genprodukts: Als Fragment einer putativen Transposase/Rekombinase könnte das Protein an der Transposition des Transposons Tn4503 (siehe 3.4.2.7.3) beteiligt sein. Da das Protein (laut Ergebnis bei der Suche nach konservierten Domänen) inkomplett ist, ist es wahrscheinlich nicht mehr zur Katalyse eines Transpositions-Ereignisses fähig.

3.4.2.3.2 ORF 8 [von 6205 Bp- 8712 Bp] (REVERSE SENSE)**Translation:**

```
>ORF8_Translation: 6205..8712 (reverse), 836 amino acids
MSRFQVLPPEYAKQFDEPPKFNRTEREYFTVDLMLDHSVTRGVNDESKVGAIIQLGYFR
ASGKFYATESFRKNDIRYVCSLLKIPYKSNLKNYPSTSKSNHKKRLILDASGFEPFNKN
KELFDETIKQMIENHMMPRKIVFSVIDLNFNQKKIETPSYDAFARSITKLYHAFESDLTTQ
LAKVIGDEHKQALESLLEKDSKGGKALISRLKHINHSVKPGQTKNSIARFDELDAIYQNIY
GLLSNIKISSEAISYQKLTIKYQLAQINNLVDDNQRYLYLLAFVHVHQQKWCQDVFLQIV
KKSVRATDNKASKSIDATYAHNAPTKNELTDSVADGYDYRNLVTVQIEKKAQDQNLNNDQ
FRCDVQDILNKNKSDHLNQSDQLQTLQDQIMDKESHQGYFDQHDQLSRKLQNKVAKIILK
RLHFTFEPVAEDLKNIAEYYQSGASITKTLAPDAFLNDEQHAQIYQDQGFNISLYKSMY
VNLFDAFSCGAASLAISYKYLPIASAYMDESSWKKNNKLLKRLGLEEFENIDEVLVKLQ
TLLDQKYDVNNAKIHSGENFFIRFNAKNDLSLIIRTPAVDKPNYESVAQMIGEDRNIPIIL
EMLTSLNKLFLRFTTPLTHFKVSKVKIPDDELFIAGIFALGSDIGIQLSKSSIGINANK
LNNAVKWFYFAVENLHAINLLTDFMNKMSLPEQFKKEKDFLHTSSDGRKRSVSIESLVSN
YSYKYFGHGKGINIYTFVDERGILFHSTVFSSSDRDAPYVIDGLLHNEGKSDMHSTDTH
GYTEAIFAISHLGITFAPRIAKLAKQNYLPSEASKKDWLKKVTRYCLAIKLMSS*
```

Konservierte Domänen:

[gnl|CDD|6485](#), pfam01526, Transposase_7, Transposase. This family includes transposases of Tn3, Tn21, Tn1721, Tn2501, Tn3926 transposons from *E. coli*. The specific binding of the Tn3 transposase to DNA has been demonstrated. Sequence analysis has suggested that the invariant triad of Asp689, Asp765, Glu895 (numbering as in Tn3) may correspond to the D-D-35-E motif previously implicated in the catalysis of numerous transposases.
 CD-Length = 968 residues, 83.3.1% aligned
 Score = 213 bits (543), Expect = 8e-56

Protein-Sequenz-Ähnlichkeiten:

[gi|15076640|dbj|BAB62426.1](#) transposase [*Bacillus cereus*]
 Length = 983
 Score = 138 bits (347), Expect = 5e-31
 Identities = 177/801 (22%), Positives = 331/801 (41%), Gaps = 52/801 (6%)

[gi|5360567|dbj|BAA82058.1](#) TnpA [*Bacillus megaterium*]
 Length = 983
 Score = 137 bits (346), Expect = 6e-31
 Identities = 177/801 (22%), Positives = 331/801 (41%), Gaps = 52/801 (6%)

[gi|482355|pir|A42707](#) type II transposase homolog - *Bacillus firmus* (fragment)
 Length = 840
 Score = 131 bits (330), Expect = 4e-29

Identities = 180/784 (22%), Positives = 328/784 (41%), Gaps = 61/784 (7%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,72

Mögliche Funktion des Genprodukts: Als putative Transposase könnte das Protein für die Transposition des auf dem Plasmid lokalisierten Transposon Tn4503 verantwortlich sein.

3.4.2.3.3 ORFs 9,10 und 11

ORF 9 erstreckt sich von 8887 Bp - 9429 Bp

ORF 10 erstreckt sich von 9438 Bp - 9599 Bp

ORF 11 erstreckt sich von 9617 Bp - 10159 Bp

Alle 3 ORFs zusammengenommen ergeben die durch Insertionen, bzw. Punktmutationen oder Deletionen unterbrochene Transposase aus dem Transposon Tn4501 (HERUTH ET AL., 1996). Die ursprüngliche Transposase war (laut Genemark.hmm) wiederum größer als seine Pendants aus pKAP116 und pKAP47 (siehe ORF 53, 3.4.2.2.1). Alle 3 ORFs sind auf dem Transposon Tn4501/2 (siehe 3.4.2.7.2) lokalisiert, welches wiederum auf Tn4503 liegt.

ORF 9:

<gi|1407813|gb|AAB03866.1> transposase
Length = 312
Score = 105 bits (261), Expect = 4e-22
Identities = 47/53 (88%), Positives = 49/53 (92%)
Query: 127 FQQALAPILGENAKNVSAQVICRLKETWHAELALWRSRSLSDKHVYVWGDGI 179
FQQALAP ++AKNVSAQVICRLKETWHAELALWRSRSLSDKHVYVW DGI
Sbjct: 15 FQQALAPYWVKSANKVSAQVICRLKETWHAELALWRSRSLSDKHVYVWVDGI 67

ORF 10:

<gi|1407813|gb|AAB03866.1> transposase
Length = 312
Score = 104 bits (260), Expect = 3e-22
Identities = 51/51 (100%), Positives = 51/51 (100%)
Query: 1 MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAE LLRDIKRRGLVAG 51
MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAE LLRDIKRRGLVAG
Sbjct: 73 MESEKTCMLVIGATVDGKKELVAFDDGFRESAESWAE LLRDIKRRGLVAG 123

ORF11:

<gi|1407813|gb|AAB03866.1> transposase
Length = 312
Score = 355 bits (911), Expect = 2e-97
Identities = 180/180 (100%), Positives = 180/180 (100%)
Query: 1 MGFWSALEKEFPKTCQRCWVHKTANILNKLPKAMQSRASKSLHNMYMAPGKEQAE EAYK 60
MGFWSALEKEFPKTCQRCWVHKTANILNKLPKAMQSRASKSLHNMYMAPGKEQAE EAYK
Sbjct: 133 MGFWSALEKEFPKTCQRCWVHKTANILNKLPKAMQSRASKSLHNMYMAPGKEQAE EAYK 192
Query: 61 EFVATYEARYPKAVECLTKDKERMFTFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG 120
EFVATYEARYPKAVECLTKDKERMFTFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG
Sbjct: 193 EFVATYEARYPKAVECLTKDKERMFTFYNFPAEHWCHIRSTNPIESSFATIRHRTRQARG 252
Query: 121 CYSRETL LCAFFKLARQAEKTSRRLRGHQLAEVINMVKFIDGISEHEVDKIKQDKQNA A 180
CYSRETL LCAFFKLARQAEKTSRRLRGHQLAEVINMVKFIDGISEHEVDKIKQDKQNA A
Sbjct: 253 CYSRETL LCAFFKLARQAEKTSRRLRGHQLAEVINMVKFIDGISEHEVDKIKQDKQNA A 312

3.4.2.3.4 ORF 12 [von 10168 Bp- 11016 Bp] (REVERSE SENSE)

Translation:

```
>ORF12_Translation: 10168..11016 (reverse), 283 amino acids
MENDAIERFIKYLESTEKSPATLINYRRDLNAPAKWFEDNNGDAFLPAKITTDLRQYKS
MLIDHPFKPKTINRRIGSLRSPINLWLDIGEINNKFPLPKLVKENQPTPMWLDKNQHAL
MRHLERYGNERDASIIISIFMNTGMRVQEFVNLKWDVTLSDKKGIIMIRHRSKANKYREIP
LNKDARYAFVKLGFKTHAGSDEYVLQGGQGNITTRGIQMMIKRRVAYTDLDYLSPHMLRH
TFCKNLVNAVGSLEKVAVLGAGHETLETTKIYCHPSMESYCYI*
```

Konservierte Domänen:

[gnl|CDD|14105](#), COG4974, XerD, Site-specific recombinase XerD [DNA replication, recombination, and repair]
CD-Length = 300 residues, 94.7% aligned
Score = 155 bits (394), Expect = 5e-39

[gnl|CDD|14104](#), COG4973, XerC, Site-specific recombinase XerC [DNA replication, recombination, and repair]
CD-Length = 299 residues, 90.3% aligned
Score = 119 bits (300), Expect = 3e-28

[gnl|CDD|7681](#), pfam00589, Phage_integrase, Phage integrase family. Members of this family cleave DNA substrates by a series of staggered cuts, during which the protein becomes covalently linked to the DNA through a catalytic tyrosine residue at the carboxy end of the alignment. The catalytic site residues in CRE recombinase are Arg-173, His-289, Arg-292 and Tyr-324.
CD-Length = 175 residues, 97.7% aligned
Score = 93.5 bits (232), Expect = 2e-20

[gnl|CDD|8420](#), pfam02899, Phage_integr_N, Phage integrase, N-terminal SAM-like domain.
CD-Length = 87 residues, 92.0% aligned
Score = 66.5 bits (162), Expect = 3e-12

[gnl|CDD|10452](#), COG0582, XerC, Integrase [DNA replication, recombination, and repair]
CD-Length = 309 residues, 87.1% aligned
Score = 50.4 bits (117), Expect = 2e-07

Protein-Sequenz-Ähnlichkeiten:

[gi|2127287|pir|I40587](#) site-specific recombinase (Tn5401) - *Bacillus thuringiensis*
Length = 306
Score = 118 bits (296), Expect = 9e-26
Identities = 78/265 (29%), Positives = 131/265 (49%), Gaps = 11/265 (4%)

[gi|23099003|ref|NP_692469.1](#) integrase:recombinase [*Oceanobacillus ihyensensis* HTE831]
Length = 305
Score = 114 bits (285), Expect = 2e-24
Identities = 87/281 (30%), Positives = 145/281 (51%), Gaps = 18/281 (6%)

[gi|21401814|ref|NP_657799.1](#) Phage_integrase, Phage integrase family [*Bacillus anthracis* A2012]
Length = 297
Score = 113 bits (282), Expect = 4e-24
Identities = 86/285 (30%), Positives = 138/285 (48%), Gaps = 18/285 (6%)

[gi|30263835|ref|NP_846212.1](#) site-specific recombinase, phage integrase family [*Bacillus anthracis* str. Ames]
Length = 299
Score = 112 bits (281), Expect = 5e-24
Identities = 86/285 (30%), Positives = 138/285 (48%), Gaps = 18/285 (6%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,64

Mögliche Funktion des Genprodukts: Der ORF ist auf dem Transposon Tn4503 lokalisiert. Als putative site-spezifische Rekombinase ist das Protein möglicherweise an der replikativen Transposition des Transposons Tn4503 beteiligt (siehe 4.2.4).

3.4.2.3.5 ORF 24 [von 22017 Bp- 23696 Bp] (REVERSE SENSE)

Translation:

```
>ORF24_Translation: 22017..23696 (reverse), 560 amino acids
LEYLTLEQKLTVKQKEALDLAQYFRYAATSQVKNKAYRKAIEDYGFYDGTGQWDKSQLH
KLSANGYPPEVLNIVKPIINSLTGVDVQTRFRVAFRSSSGNVKDDRLAKAITHLGYHIQE
NQDMTHKATQKFMMLITGIGWSNLFKQPSTSEIIYEHVHPLDVLDFDDDLSSQLNNSF
VIRIRWISLAEMKRLWPKHSSYFDSRFSVKNVREAKEWKYQGEVSPFIQQRAEIELDT
```


FYLGEGYGRVRVIEVQYRQEMMSFEGMDINGHYFATFNEKEALSLVEKMTDLERKSSKA
 IMRVLFTEDRLLLEYAPLKNVNDLDFDFTYIPCLWSRRYS DGVDPDWLSSMKPVQRRYNSL
 QTRLSYLLSSTRAIYDEGAFSNKEPHEIEAMLSDPGAARIEISPGSRFELHANNQISGGYF
 TLLKKTVEDIQRVSGYDEALGKETNAQSGVAIRARQVNSVRGQTIAFDNLKMKKREAR
 MMLNLIQSGSGAFLESQILTEDEKREIVILNCVHDVDRKRVVNDIRTLPLSIHIEEVPD
 FESPTQEQQARLDRVPRAS*

Konservierte Domänen:

Keine

Protein-Sequenz-Ähnlichkeiten:

[gi|18071218|ref|NP_542303.1|](#) unknown [*Sinorhizobium meliloti* phage PBC5]
 Length = 749
 Score = 107 bits (268), Expect = 4e-22
 Identities = 112/473 (23%), Positives = 206/473 (43%), Gaps = 30/473 (6%)

[gi|15800971|ref|NP_286987.1|](#) putative portal protein of bacteriophage BP-933W [*Escherichia coli*
 O157:H7 EDL933]
 Length = 714
 Score = 79.7 bits (195), Expect = 1e-13
 Identities = 98/444 (22%), Positives = 183/444 (41%), Gaps = 47/444 (10%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,71

Mögliche Funktion des Genprodukts: Das putative Protein könnte, als Phagen-Portal-Protein, an der Beladung eines hypothetischen Phagen-Capsids mit dem Phagen-Genom beteiligt sein.

3.4.2.3.6 ORF 25 [von 23710 Bp- 25134 Bp] (REVERSE SENSE)

Translation:

>ORF25_Translation: 23710..25134 (reverse), 475 amino acids
 MNIDSFVVDAAFNHVQKHFLLYTPHKKRQFHAAGASAKERLFLAGNRTGKYCGVAESV
 MHLTGYPQWIGKRFRTRPIKAWAASVTALTAEVLEKAYLEMAEDLVIGVDRLRHSYK
 IDYKTTGGYSELTFKSYEQGRKKFQAAKLDLVHLDEEPPRDIYVESLMRTMSTEVDFNEGIV
 LLTMTPLGLTDLILEFQETTIEREVINGSGVSEMSEMTEEVIKVDEGSIVNNRFYIQAS
 WDDNPHLLDSAKETLSKSLKPEKEARKHGIPSLGSGLVYPVSEVAVVNVNPFVLPKHWR
 VFGLDFWGNPTAALFAVIDRDNVYMLTGEYYSERTPQQHVYELKKGADKINGVYDP
 AGEQSSQRDGGDLAQLYRDSGLRYLYKADNAKEEGIMKVLQRFQNGKLIKIFNLNNTLKE
 FRMYSRDEKGVKKGNDHMLDCLRYIVMSGERLAKSDIQRNSTYHNRPSLKGIF*

Konservierte Domänen:

[gnl|CDD|14649](#), COG5565, COG5565, Bacteriophage terminase large (ATPase) subunit and inactivated derivatives
 [General function prediction only]
 CD-Length = 79 residues, 98.7% aligned
 Score = 66.7 bits (162), Expect = 6e-12

Protein-Sequenz-Ähnlichkeiten:

[gi|15965769|ref|NP_386122.1|](#) PROBABLE DNA PACKAGING PROTEIN GP2 [*Sinorhizobium meliloti*]
 Length = 477
 Score = 273 bits (699), Expect = 3e-72
 Identities = 171/469 (36%), Positives = 240/469 (51%), Gaps = 65/469 (13%)

[gi|27476053|ref|NP_775255.1|](#) terminase [*Pseudomonas aeruginosa* phage PaP3]
 Length = 482
 Score = 250 bits (638), Expect = 5e-65
 Identities = 158/483 (32%), Positives = 231/483 (47%), Gaps = 67/483 (13%)

[gi|24371583|ref|NP_720326.1|](#) gp2 [*Salmonella typhimurium* bacteriophage ST64T]
 Length = 517
 Score = 205 bits (522), Expect = 1e-51
 Identities = 147/475 (30%), Positives = 207/475 (43%), Gaps = 77/475 (16%)

[gi|9635535|ref|NP_059627.1](#) terminase large subunit [*Enterobacteria* phage P22]
 Length = 499
 Score = 204 bits (520), Expect = 2e-51
 Identities = 153/476 (32%), Positives = 212/476 (44%), Gaps = 79/476 (16%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,77

Mögliche Funktion des Genprodukts: Das putative Protein könnte, als Phagen-Packaging-Protein, an der Beladung eines hypothetischen Phagen-Capsids mit dem Phagen-Genom beteiligt sein. Als putative große Untereinheit einer Terminase könnte es ATP- und Capsid-bindende, sowie DNA schneidende Aktivität aufweisen (FUJISAWA ET AL., 1997).

3.4.2.3.7 ORF 32 [von 28353 Bp- 29252 Bp] (REVERSE SENSE)

Translation:

```
>ORF32_Translation: 28353..29252 (reverse), 300 amino acids
MKNVHDSVFKDLIANRDFAVSFLMTYLPKELVELVDWQTVKLESANVEHVRQQQKDNQKQ
KEQSDLTFLFKFDGKNGAVFVHIESQTGDDGTILIRTRHYQTSYLLDYIKRHKTVKGLP
LVVSIYYANQKPFSSHLSNIHDYFANTELAKKYAFTTQFIDLNRYSDEEILEHGFIAGYE
LILKAIREKNI DGKGLDIAINQIEAYDHIARQVLI RYMSQYS DMETKDFHDKLIYSKPDLR
GDVMTVAEQWEQKGIQKGIQT TARNFLMGLSAEQVVKGTGLDQDTVLK LKKEVEQTQH*
```

Konservierte Domänen:

[gnl|CDD|16051](#), pfam04754, Transposase_31, Putative transposase, YhgA-like. This family of putative transposases includes the YhgA sequence from *Escherichia coli* and several prokaryotic homologues.
 CD-Length = 288 residues, 92.4% aligned
 Score = 115 bits (290), Expect = 5e-27

[gnl|CDD|14548](#), COG5464, COG5464, Uncharacterized conserved protein [Function unknown]
 CD-Length = 289 residues, 97.2% aligned
 Score = 77.7 bits (191), Expect = 1e-15

Protein-Sequenz-Ähnlichkeiten:

[gi|26251226|ref|NP_757266.1](#) Hypothetical protein yfcI [*Escherichia coli* CFT073]
 Length = 306
 Score = 110 bits (276), Expect = 2e-23
 Identities = 88/301 (29%), Positives = 154/301 (51%), Gaps = 18/301 (5%)

[gi|9507490|ref|NP_052497.1](#) 55 pct identical (4 gaps) to sp:YTL2_SALTY[hypothetical protein near tlpA operon of *S. typhimurium*] [Plasmid ColIb-P9]
 Length = 313
 Score = 102 bits (253), Expect = 1e-20
 Identities = 75/250 (30%), Positives = 124/250 (49%), Gaps = 17/250 (6%)

[gi|16763859|ref|NP_459474.1](#) putative transposase [*Salmonella typhimurium* LT2]
 Length = 311
 Score = 101 bits (252), Expect = 1e-20
 Identities = 82/281 (29%), Positives = 136/281 (48%), Gaps = 23/281 (8%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,72

Mögliche Funktion des Genprodukts: Als putative Transposase könnte das Protein für die Transposition des auf dem Plasmid lokalisierten Transposon Tn4504 (siehe 3.4.2.7.4) verantwortlich sein.

3.4.2.3.8 ORF 39 [von 32377 Bp- 32929 Bp] (REVERSE SENSE)

Translation:

>ORF39_Translation: 32377..32928 (reverse), 184 amino acids
 MINAEKKLNQENDKQEVLRGRGRPCCKDNRSMSRNELIKHLKSLKTKFSREDCNHHFFYEL
 MNLLREQLLNRIILLEGGRFETRVSVFRKNAHNPRTEKIKPREKRVVLFKFAFNEFKAL
 ISDEAHTPNPNDLLDYVDNLDDKKPRGRPKKIVQKSSIEDKPKRPRGRPKIFFKDTPKNDKE
 LES*

Konservierte Domänen:

[gnl|CDD|10644](#), COG0776, HimA, Bacterial nucleoid DNA-binding protein [DNA replication, recombination, and repair
 CD-Length = 94 residues, 94.7% aligned
 Score = 52.8 bits (127), Expect = 2e-08

[gnl|CDD|7512](#), pfam00216, Bac_DNA binding, Bacterial DNA-binding protein.
 CD-Length = 90 residues, 98.9% aligned
 Score = 51.7 bits (124), Expect = 5e-08

[gnl|CDD|367](#), smart00411, BHL, bacterial (prokaryotic) histone like domain
 CD-Length = 90 residues, 98.9% aligned
 Score = 48.2 bits (115), Expect = 5e-07

Protein-Sequenz-Ähnlichkeiten:

[gi|11467715|ref|NP_050767.1](#) DNA-binding protein hu homolog [*Guillardia theta*]
 Length = 93
 Score = 43.5 bits (101), Expect = 0.002
 Identities = 23/86 (26%), Positives = 42/86 (48%)

[gi|16125619|ref|NP_420183.1](#) integration host factor alpha-subunit [*Caulobacter crescentus* CB15]
 Length = 100
 Score = 43.5 bits (101), Expect = 0.002
 Identities = 23/90 (25%), Positives = 44/90 (48%)

[gi|118264|sp|P05385|DBH_CLOPA](#) DNA-BINDING PROTEIN HU
 Length = 91
 Score = 42.7 bits (99), Expect = 0.003
 Identities = 29/86 (33%), Positives = 42/86 (48%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,65

Mögliche Funktion des Genprodukts: Als Integration Host Factor könnte das putative Protein an der Replikation des Plasmids beteiligt sein. Es ist ebenso ist eine transkriptions-regulatorische Aufgabe dieses Proteins möglich (wie in *E. coli*, OBERTO ET AL., 1994).

3.4.2.3.9 ORF 42 [von 34060 Bp- 35679 Bp] (REVERSE SENSE)**Translation:**

>ORF42_Translation: 34060..35679 (reverse), 540 amino acids
 MDNPFINNHGKTEDMSQETGDKTGDMSQETGDKTGDMSQETGDKTEDMSQEIGDKTGDMS
 QETGDKTGDMSQETGDMSQETGDKTGDMSQETGDKTGDMSQETGDKTGDMSQETGDKTGD
 MSQETGDKTEDMSQETGDKTGDMSQETGDKTEDMSQETGDKTEDMSQEIGDKTGDMSQET
 GDKTEDMSQEIGDKTEDMSQETGNKIQLNGCQFIVIRIMISSLENLCLLITCHINTNNII
 KQIKSNKDIVKAIYRLKKNYISVYERKTRNGWVKYQIRPDVFEYVKSLSQSTQGDFEK
 NTPQKNESDSHIYNIDKIEHSMCLNMGQDEKTDSSKYSQSNVVEKLTTEELEFIKEQYEL
 LVKENHDLRKITLNNSEATKTNETTRDELNKTDSLTLSHVNLSDVNLDLTYETIDLTKIN
 DLLKSMGKAPFFYSSDINEIRINAEENNIEMNPLQLQIFLRHFEHQMTYQWEDISKRWDEK
 AINNPKHVFIGTFKKRHVALVDGYSCQEQQATSHIEAQKSIEALKQKQREKEKEIRKNE*

Konservierte Domänen:

[gnl|CDD|14402](#), COG5280, COG5280, Phage-related minor tail protein [Function unknown]
 CD-Length = 634 residues, only 32.3% aligned
 Score = 45.8 bits (108), Expect = 1e-05

[gnl|CDD|8456](#), pfam02987, LEA, Late embryogenesis abundant protein. Different types of LEA proteins are expressed at different stages of late embryogenesis in higher plant seed embryos and under conditions of dehydration stress. The function of these proteins is unknown.

CD-Length = 70 residues, 95.7% aligned
Score = 45.2 bits (107), Expect = 2e-05

Protein-Sequenz-Ähnlichkeiten:

[gi|27712910|ref|XP_232290.1](#) similar to *Drosophila melanogaster* CG15040 gene product [*Plasmodium yoelii yoelii*] [*Rattus norvegicus*]
Length = 355

Score = 92.0 bits (227), Expect = 2e-17
Identities = 77/221 (34%), Positives = 113/221 (51%), Gaps = 26/221 (11%)

[gi|11358481|pir|T47561](#) late embryogenesis abundant protein-like - *Arabidopsis thaliana*
Length = 479

Score = 82.4 bits (202), Expect = 2e-14
Identities = 64/200 (32%), Positives = 104/200 (52%), Gaps = 5/200 (2%)

[gi|421875|pir|S33520](#) Lea protein - soybean
Length = 458

Score = 75.5 bits (184), Expect = 2e-12
Identities = 57/194 (29%), Positives = 99/194 (51%), Gaps = 3/194 (1%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,80

Mögliche Funktion des Genprodukts: Das putative Gen könnte für ein Phagen-Schwanz-Protein kodieren (siehe konservierte Domänen). Es könnte daher mitunter für die, mit den R-Körpern assoziierten, helikalen Proteinstrukturen kodieren (siehe Einleitung 1.4.1).

3.4.2.3.10 ORF 43 [von 35749 Bp- 36408 Bp] (REVERSE SENSE)

Translation:

>ORF43_Translation: 35749..36408 (reverse), 220 amino acids
VALIELDQQANLTRTLGVDARKKPVLDIYINDKLEISDLLINIFDGLDFIPSRVDNALLD
NSLMLGKHLPLDKVFNKPFSELKSKYDIIVIDCPSPISGSAVSAATLASDEIIMFITPTDYA
LAGLELTHHEILNLCEKYDKDIVLKLILFNKFSRSTNLSHEVMKYVLKKNQCNSILLRSYI
RSLQSIENCIAASISVFDSPFKKTPEKEDFSLTSEILGI*

Konservierte Domänen:

[gnl|CDD|10910](#), COG1192, Soj, ATPases involved in chromosome partitioning [Cell division and chromosome partitioning]

CD-Length = 259 residues, 83.3.2% aligned
Score = 99.5 bits (247), Expect = 3e-22

[gnl|CDD|4405](#), pfam00991, ParA, ParA family ATPase
CD-Length = 110 residues, **only 71.8% aligned**
Score = 55.7 bits (134), Expect = 4e-09

[gnl|CDD|10328](#), COG0455, COG0455, ATPases involved in chromosome partitioning [Cell division and chromosome partitioning]

CD-Length = 262 residues, 78.6% aligned
Score = 54.1 bits (130), Expect = 1e-08

[gnl|CDD|10868](#), COG1149, COG1149, MinD superfamily P-loop ATPase containing an inserted ferredoxin domain [Energy production and conversion]

CD-Length = 284 residues, **only 22.9% aligned**
Score = 36.4 bits (84), Expect = 0.003

Protein-Sequenz-Ähnlichkeiten:

[gi|32470489|ref|NP_863174.1](#) putative Soj [*Enterococcus faecium*]
Length = 267

Score = 87.4 bits (215), Expect = 1e-16
Identities = 61/224 (27%), Positives = 119/224 (53%), Gaps = 13/224 (5%)

[gi|29377956|ref|NP_817082.1](#) ParA family protein [*Enterococcus faecalis* V583]
Length = 260

Score = 85.5 bits (210), Expect = 6e-16
Identities = 65/229 (28%), Positives = 117/229 (51%), Gaps = 23/229 (10%)

[gi|15081480|ref|NP_149993.1](#) Soj protein [*Clostridium perfringens*]
Length = 250

Score = 72.4 bits (176), Expect = 5e-12
Identities = 62/219 (28%), Positives = 110/219 (50%), Gaps = 5/219 (2%)

[gi|29611531|ref|NP_818984.1](#) chromosome-partitioning ATPase [*Bacteroides thetaiotaomicron*]
 Length = 252
 Score = 71.6 bits (174), Expect = 7e-12
 Identities = 54/218 (24%), Positives = 105/218 (48%), Gaps = 6/218 (2%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,62

Mögliche Funktion des Genprodukts: Das putative Protein könnte bei der Replikation des Plasmids pKAP298 eine Rolle spielen. Für ähnliche Soj/ParA-Proteine wurde eine wichtige Funktion bei der Segregation der replizierten (Plasmid- oder Chromosomen-) DNA sich teilender Zellen festgestellt (MOTALLEBI-VESHAREH ET AL., 1990, siehe 4.3). Zur ParA-Familie gehörende Proteine kommen auch auf low-copy-number Plasmiden, wie dem Prophagen P1 aus *E. coli*, vor (DAVEY ET AL., 1994).

3.4.2.3.11 ORF 44 [von 36899 Bp- 37660 Bp]

Translation:

```
>ORF44_Translation: 36899..37660 (direct), 254 amino acids
MAKVKLPQWLRDINVFHREFEYIYQREDFKAIDGIRYVHKDGTVFYSESDDGTLHLEEELE
EGQYRLAASAKEIITYRRYLEYTLSDFEVLFVAVGKYCRLTHKDSGIVYVADQEGELFTLK
EIADEKARQKQFQRIERSSPADFAESKPESYLNSTVIVVPAVSAECNQYLHIKLTDDQDLK
LLPAKTLTPVELKILRCLMNDVPIKAIKAAKVGGRYNSANISVVLARMRQKFGVETNHGLIA
KAYQMGFSEIRILY*
```

Konservierte Domänen:

[gnl|CDD|9095](#), pfam00196, GerE, Bacterial regulatory proteins, luxR family
 CD-Length = 66 residues, 97.0% aligned
 Score = 36.7 bits (85), Expect = 0.003
 Query: 185 **KTLTPVELKILRCLMNDVPIKAIKAAKVGGRYNSANISVVLARMRQKFGVETNHGLIAKAYQ** 244
 Sbjct: 2 **DSLSPREREVLRLLAAGKSNKEIADILGISEK-TVKVHRSNIMRKLNVHRSVELIRLAI** 60
 Query: 245 **MGFSE** 249
 Sbjct: 61 **LGLIA** 65

Protein-Sequenz-Ähnlichkeiten:

[gi|10640351|emb|CAC12165.1](#) DNA repair protein RAD51 related protein [*Thermoplasma acidophilum*]
 Length = 245
 Score = 36.6 bits (83), Expect = 0.42
 Identities = 24/70 (34%), Positives = 37/70 (52%), Gaps = 3/70 (4%)
[gi|13878692|sp|Q9HJD3|RADB_THEAC](#) DNA repair and recombination protein radB
 Length = 229
 Score = 36.2 bits (82), Expect = 0.51
 Identities = 24/70 (34%), Positives = 37/70 (52%), Gaps = 3/70 (4%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,77

Mögliche Funktion des Genprodukts: Das putative Protein weist in seiner Aminosäuresequenz Ähnlichkeiten zu der Proteinfamilie der luxR-Proteine auf, genauer zu GerE. GerE wird als Transkriptions-Aktivator klassifiziert (CUTTING ET AL., 1986). Möglicherweise ist ORF 44 an der Regulation bestimmter Gene auf pKAP298 beteiligt.

3.4.2.3.12 ORF 47 [von 39601 Bp- 40626 Bp]**Translation:**

```
>ORF47_Translation: 39601..40626 (direct), 342 amino acids
MSNAELNKTLKLLSSIAECARGNRSFEFTSLAHLLDAAEFLSYCYGLDRNKAVGIDKVS
WQEYGVVDVADKIEENLVMLRKRKTFKMPSPRRVYIPKNGESRPLGISAIENKIVESGIMLI
LQSIYEQDFLECSYGFRRPGRNTHQALNEVDKAIMTQPVNHLVEADIKGFFDNVSHEKLD
FKVIRVKDTSLHLIDCFLLRAGYIDKGVLDIDTEKGTQPGSILSPMLANIFLHYVLDKWF
E DTVKQHVVEGFCRLVRYADDFVCLIQYQEDARKIERALANRFNKHELQLHPEKSRNISFGR
FEKLNALNAGRKANTFDFLGFTHFCDKTRKGYFQTSENER*
```

Konservierte Domänen:

[gnl|CDD|12678](#), COG3344, COG3344, Retron-type reverse transcriptase [DNA replication, recombination, and repair]
 CD-Length = 328 residues, 86.9% aligned
 Score = 113 bits (284), Expect = 3e-26

[gnl|CDD|16610](#), pfam00078, rvt, Reverse transcriptase (RNA-dependent DNA polymerase). A reverse transcriptase gene is usually indicative of a mobile element such as a retrotransposon or retrovirus. Reverse transcriptases occur in a variety of mobile elements, including retrotransposons, retroviruses, group II introns, bacterial msDNAs, hepadnaviruses, and caulimoviruses.
 CD-Length = 200 residues, 93.0% aligned
 Score = 105 bits (263), Expect = 7e-24

Protein-Sequenz-Ähnlichkeiten:

[gi|27377026|ref|NP_768555.1](#) blr1915 [*Bradyrhizobium japonicum*]
 Length = 491
 Score = 270 bits (690), Expect = 3e-71
 Identities = 135/323 (41%), Positives = 202/323 (62%), Gaps = 1/323 (0%)

[gi|12620616|gb|AAG60892.1|AF322013.11](#) ID459 [*Bradyrhizobium japonicum*]
 Length = 491
 Score = 268 bits (685), Expect = 9e-71
 Identities = 134/323 (41%), Positives = 201/323 (62%), Gaps = 1/323 (0%)

[gi|2688957|gb|AAC82519.1](#) RetA [IncL/M plasmid R471a]
 Length = 495
 Score = 267 bits (683), Expect = 1e-70
 Identities = 136/323 (42%), Positives = 199/323 (61%), Gaps = 1/323 (0%)

[gi|26991132|ref|NP_746557.1](#) group II intron-encoding maturase, putative [*Pseudomonas putida* KT2440]
 Length = 373
 Score = 229 bits (585), Expect = 4e-59
 Identities = 117/253 (46%), Positives = 161/253 (63%), Gaps = 1/253 (0%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,62

Mögliche Funktion des Genprodukts: ORF 47 ist wahrscheinlich Teil eines Gruppe-II-Introns. Das putative Protein schreibt möglicherweise das als ssRNA vorliegende Gruppe-II-Intron in doppelsträngige DNA um, die dann (vielleicht vom gleichen Enzym) an irgendeiner Stelle im Genom, oder sonstiger dsDNA, integriert werden kann (LEHMANN ET AL., 2003).

3.4.2.3.13 ORF 48 [von 40642 Bp- 40935 Bp]**Translation:**

```
>ORF48_Translation: 40642..40935 (direct), 98 amino acids
VQDLNLWLKSIIRNVLPTKEWWKILKAKLRGHFEYYGVSGNYPSIAKYYSIAMRQTLKWL
NRRSQKRKMNWDFNKYVERHPLPRPRINHIFYYTLVAL*
```

Konservierte Domänen:

Keine

Protein-Sequenz-Ähnlichkeiten:

[gi|12620547|gb|AAG60823.1|AF322012_128](#) ID272 [*Bradyrhizobium japonicum*]
 Length = 193
 Score = 58.5 bits (140), Expect = 2e-08
 Identities = 31/63 (49%), Positives = 39/63 (61%), Gaps = 1/63 (1%)

[gi|27350092|dbj|BAC47104.1](#) blr1839 [*Bradyrhizobium japonicum* USDA 110]
 Length = 172
 Score = 57.0 bits (136), Expect = 6e-08
 Identities = 31/63 (49%), Positives = 39/63 (61%), Gaps = 1/63 (1%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: >0,50

Mögliche Funktion des Genprodukts: Unbekannt. Dieser ORF ist allerdings direkt neben der putativen reversen Transkriptase eines Gruppe-II-Introns lokalisiert (ORF 47) und könnte daher Teil dieses Introns sein. Dafür spricht auch, dass sowohl ORF 48 als auch ORF 47 große Ähnlichkeiten zu putativen Proteinen aus *Bradyrhizobium japonicum* aufweisen.

3.4.2.3.14 ORF 49 [von 41075 Bp- 41308 Bp]**Translation:**

>ORF49_Translation: 41075..41308 (direct), 78 amino acids
 MKATIEKTGKIGKINKMTHDAYLAHEALSASQVKLLLDNPHLFLRKEPMVQTFDMLLGSV
 IHSVLVLEPEECLILSMR*

Protein-Sequenz-Ähnlichkeiten:

[gi|6103651|gb|AAF03706.1](#) unknown [lambdoid phage HB-4]
 Length = 309
 Score = 35.0 bits (79), Expect = 0.29
 Identities = 22/58 (37%), Positives = 33/58 (56%), Gaps = 4/58 (6%)

[gi|2507105|sp|P15032|RECE_ECOLI](#) Exodeoxyribonuclease VIII (EXO VIII)
 Length = 866
 Score = 34.7 bits (78), Expect = 0.36
 Identities = 22/58 (37%), Positives = 33/58 (56%), Gaps = 4/58 (6%)

[gi|15831188|ref|NP_309961.1](#) exonuclease VIII RecE [*Escherichia coli* O157:H7]
 Length = 866
 Score = 34.3 bits (77), Expect = 0.46
 Identities = 22/58 (37%), Positives = 33/58 (56%), Gaps = 4/58 (6%)

[gi|147536|gb|AAA03212.1](#) recE
 Length = 661
 Score = 34.3 bits (77), Expect = 0.51
 Identities = 22/58 (37%), Positives = 33/58 (56%), Gaps = 4/58 (6%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,53

Mögliche Funktion des Genprodukts: siehe ORF 54

3.4.2.3.15 ORF 50 [von 41308 Bp- 41937 Bp]**Translation:**

>ORF50_Translation: 41308..41937 (direct), 210 amino acids
 MPKLDARTKAGKEEKQAFEEAHEGKVIITDALFDTAQSCAEAITSSSANALLQNGVAEAS
 FFGQLEGYDAKCRPDYCRERSLIIDIKKTTDASPEGFTRADANFKYYIQAAWYLDVLKS
 IGEVPVKDFLFIAVEATAPFMVGIYRLADSAIEFGRLKYREALKTYRNINDFAKPIYRDT
 TRKYIQTIDLPAVVYAQNKNSNLMEVYHG*

Protein-Sequenz-Ähnlichkeiten:

[gi|13186143|emb|CAC33454.1](#) ORF B-Homologous to RecE (Exodeoxyribonuclease VIII, Exo VIII) of *E. coli* [*Legionella pneumophila*]
 Length = 280
 Score = 84.3 bits (207), Expect = 1e-15
 Identities = 63/198 (31%), Positives = 98/198 (49%), Gaps = 6/198 (3%)

[gi|29566166|ref|NP_817737.1](#) gp60 [Mycobacteriophage Che9c]
 Length = 314
 Score = 60.5 bits (145), Expect = 2e-08
 Identities = 57/193 (29%), Positives = 87/193 (45%), Gaps = 15/193 (7%)

[gi|15801815|ref|NP_287833.1](#) putative exodeoxyribonuclease VIII of prophage CP-933R [*Escherichia coli* O157:H7 EDL933]
 Length = 866
 Score = 55.5 bits (132), Expect = 5e-07
 Identities = 53/201 (26%), Positives = 96/201 (47%), Gaps = 19/201 (9%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,75

Mögliche Funktion des Genprodukts: siehe ORF 54

3.4.2.3.16 ORF 54 [von 44336 Bp- 44800 Bp]**Translation:**

>ORF54_Translation: 44336..44800 (direct), 155 amino acids
 MRYNLVKNELVKTLLSDGRAEQSIFWIDEGTDILCKARADYINVEQGYIVDIKTTSNLAD
 DSSFGFTLKQYHYDLSAGFYVDGFKKAIGRDVFYFIVIESNPPHNYSIFKSSIDLLAQQ
 RNKYKALDTFFTRANQSGNFAKPYNNGNLIELTA*

Protein-Sequenz-Ähnlichkeiten:

[gi|1742216|dbj|BAA14950.1](#) Exodeoxyribonuclease VIII (EC 3.1.11.-) (Exo VIII). [*Escherichia coli*]
 Length = 797
 Score = 62.8 bits (151), Expect = 2e-09
 Identities = 40/120 (33%), Positives = 56/120 (46%), Gaps = 6/120 (5%)

[gi|6103651|gb|AAF03706.1](#) unknown [lambdoid phage HB-4]
 Length = 309
 Score = 61.2 bits (147), Expect = 4e-09
 Identities = 34/83 (40%), Positives = 43/83 (51%), Gaps = 3/83 (3%)

[gi|15801815|ref|NP_287833.1](#) putative exodeoxyribonuclease VIII of prophage CP-933R [*Escherichia coli* O157:H7 EDL933]
 Length = 866
 Score = 61.2 bits (147), Expect = 5e-09
 Identities = 34/83 (40%), Positives = 43/83 (51%), Gaps = 3/83 (3%)

[gi|13186143|emb|CAC33454.1](#) ORF B-Homologous to RecE (Exodeoxyribonuclease VIII, Exo VIII) of *E. coli* [*Legionella pneumophila*]
 Length = 280
 Score = 52.0 bits (123), Expect = 3e-06
 Identities = 35/107 (32%), Positives = 61/107 (57%), Gaps = 5/107 (4%)

Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt: 0,77

Mögliche Funktion des Genprodukts: Die ORFs 49,50 und 54 waren sehr wahrscheinlich einmal ein einziges Protein, bevor die Insertion des Transposons Tn4501 und einer anderen Sequenz (ORFs 51 und 52, Funktion unbekannt) dieses putative Gen unterbrochen hat - auch scheinen an diesem Ort (wie auch immer geartete) Mutationen stattgefunden haben, die aus ursprünglich einem ORF mindestens 3 gemacht haben. Die Aminosäuresequenz weist große Ähnlichkeit mit dem RecE-Protein aus *E. coli* und zu diversen Phagen-Proteinen auf. RecE

besitzt eine Exonuclease-Aktivität und spielt wahrscheinlich eine wichtige Rolle bei der Rekombination des Phagen-Genoms (FTOUHI ET AL., 1990; ZHANG ET AL., 2003). Da das ursprüngliche Genprodukt (der ganze zu *RecE* homologe ORF) nicht mehr existiert, und die einzelnen jetzt vorhandenen ORFs möglicherweise nicht mehr für funktionelle Proteine kodieren, ist es denkbar, dass das Plasmid pKAP298, als putatives Bakteriophagen-Derivat, die Fähigkeit zur selbstkodierte Rekombination verloren hat.

3.4.2.4 Potentiell kodierende ORFs ohne Sequenzhomologien (Orphans)

Die von GeneMark.hmm identifizierten ORFs ohne signifikante Aminosäure-Sequenzhomologien werden hier als Tabelle aufgeführt. Es wird die Nummer des ORFs, seine Lage auf dem Plasmid, die Orientierung und die durchschnittliche Wahrscheinlichkeit, dass es ein codierender ORF ist, wiedergegeben.

Nummer	Ort auf dem Plasmid (in Bp)	Länge des ORF (in Bp)	Orientierung	Wahrscheinlichkeit
4	1447-2199	753	-	0,55
5	2229-4211	1983	-	0,85
6	4208-5584	1376	-	0,83
13	11218-12369	1152	+	0,68
14	12423-13325	903	+	0,62
15	13358-14494	1137	-	0,86
16	14550-16070	1521	-	0,71
17	16082-18187	2106	-	0,64
18	18198-18827	630	-	>0,50
19	18827-19174	348	-	0,74
20	19222-20343	1122	-	0,75
21	20592-21572	981	-	0,81
22	21572-21721	150	-	0,50
23	21758-21874	117	+	>0,50
26	25138-25503	355	-	0,61
27	25519-25659	141	-	>0,50
28	25851-26144	294	-	0,56
29	26474-27247	774	+	0,88
30	27376-27684	309	+	0,65
31	27978-28214	237	-	0,67

Nummer	Ort auf dem Plasmid (in Bp)	Länge des ORF (in Bp)	Orientierung	Wahrscheinlichkeit
33	29813-29887	75	-	>0,50
34	30144-30224	81	-	>0,50
35	30457-30576	120	-	>0,50
36	30584-31894	1311	-	0,80
37	31891-32031	141	-	>0,50
38	32139-32315	177	-	0,52
40	33127-33243	117	+	>0,50
41	33657-34067	411	-	0,59
45	37968-38264	297	+	0,60
46	38505-38789	285	+	0,66
51	41930-42553	624	+	0,53
52	42596-42769	174	+	0,55
55	44890-45357	468	+	0,65
56	45369-45704	336	+	>0,50
57	45713-46018	306	+	0,62
58	46181-46912	732	-	0,61
59	46890-47228	339	-	0,66
60	47257-48015	759	-	0,73
61	48211-48477	267	+	0,57
62	48524-48811	288	-	0,69

Tab. 3.4.2.4.1: von GeneMark.hmm identifizierte putativ kodierende ORFs ohne Sequenz-Homologien zu schon bekannten Genen/Proteinen. Es ist die Lage auf dem Plasmid (relativ zur Bam HI - Schnittstelle), die Länge des ORFs, die Orientierung und die durchschnittliche Wahrscheinlichkeit, dass es sich um einen kodierenden ORF handelt, angegeben

3.4.2.5 Annotation des Plasmids pKAP298

Das Plasmid pKAP298, die potentiell kodierenden ORFs, sowie die Transposons wurden mit dem NCBI-Programm Sequin annotiert und werden demnächst veröffentlicht. Die Annotations-Datei ("pkap298.sqn") befindet sich auf der beiliegenden CD im Ordner "Annotation_pKAP298", das Programm Sequin im Ordner "Programme".

3.4.2.6 Transkriptions-Aktivitäts-Karte mit allen wahrscheinlich kodierenden ORFs

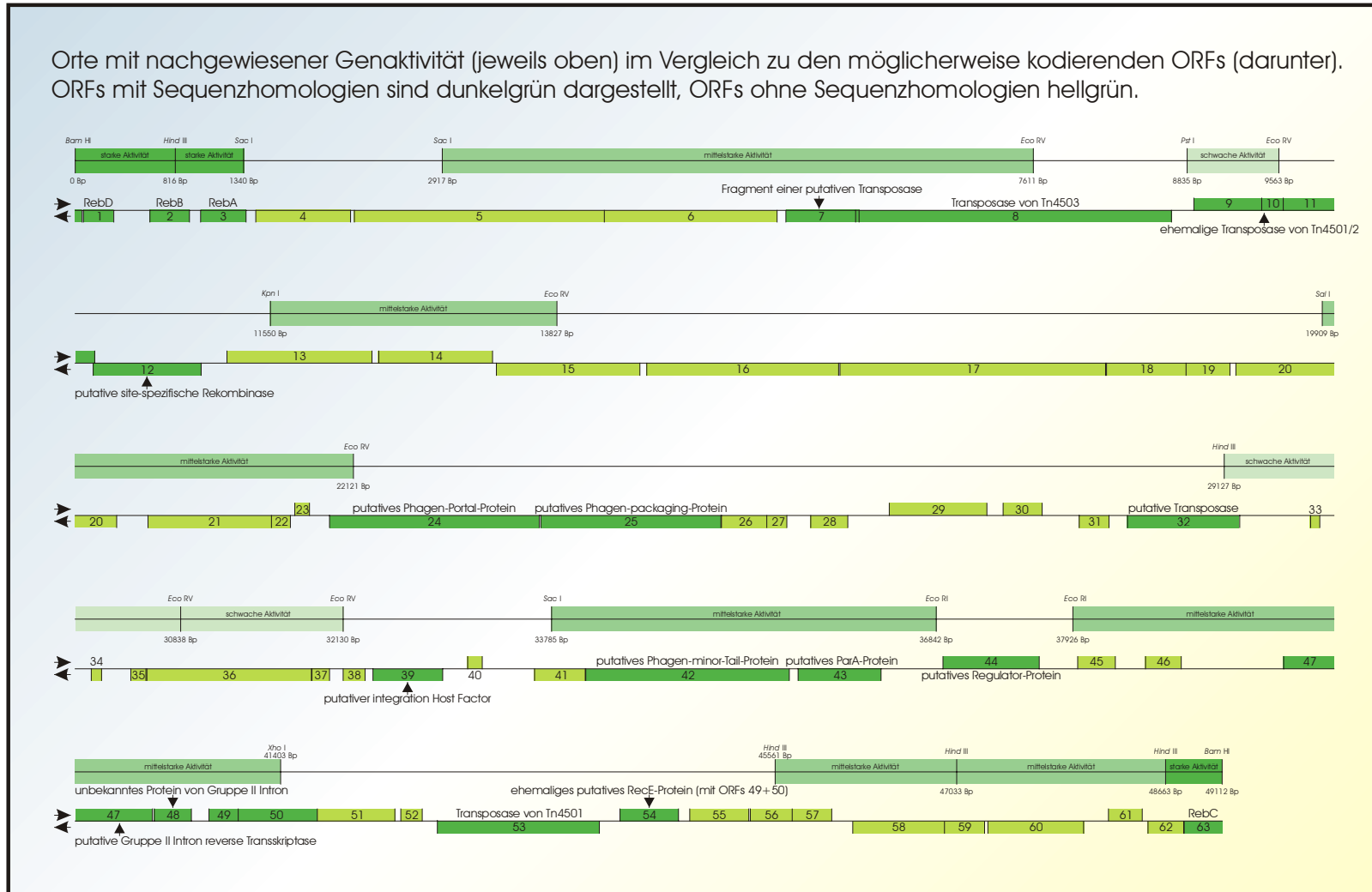


Abb. 3.4.2.6.1: von GeneMark.hmm ermittelten möglicherweise kodierenden ORFs (jeweils untere Linie der 5 Linienpaare) im Vergleich zu den Orten mit detektierte Transkriptions-Aktivität (jeweils obere Linie der Linienpaare). ORFs mit Sequenzhomologien sind dunkelgrün dargestellt, ORFs ohne Homologien hellgrün. ORFs mit positiver Orientierung stehen über der Linie, revers orientierte unterhalb. Die oberen vier Linienpaare entsprechen jeweils 10 kB des Gesamtplasmids, das untere Linienpaar nur 9 kB

3.4.2.7 Identifizierte mögliche Transposons auf dem Plasmid pKAP298

Auf dem Plasmid pKAP298 konnten 3 Transposons identifiziert werden, die von QUACKENBUSH ET AL. (1986A,B) schon für die Plasmide pKAP116 und pKAP47 teilweise beschrieben wurden. Des weiteren wurde ein neues mögliches Transposon entdeckt.

3.4.2.7.1 Transposon Tn4501

Das Transposon ist lokalisiert zwischen den Basen 42862 und 44301 (=1439 Bp). Es konnte in dieser Arbeit erstmals in voller Länge und mit seinen Wiederholungs-Sequenzen charakterisiert werden.

Invertierte Wiederholungs-Sequenz (Ergebnis vom Programm Einvert):

```
Score 50: 30/34 ( 88%) matches, 2 gaps
  42862 aatattaatcgagttattgtcaaatttagtgatgg 42897
        |||| |||| ||| ||| || |||||
  44301 ttat-atta-ctctatatgagcttaaatcacatacc 44268
```

Direkte Wiederholung der Insertions-Sequenz vor und nach dem Transposon:

```
Linker Terminus      AAACAACAGGAGATAA
Rechter Terminus     AATCAGCAGGACTGAA
```

Mögliche Transposase: ORF 53, auf pKAP47 und pKAP116 beschrieben und sequenziert von HERUTH ET AL. (1996).

3.4.2.7.2 Zu Tn4501 homologes Transposon Tn4501/2

Das Transposon ist lokalisiert zwischen den Basen 8756 und 10196 (=1440 Bp) und weist große Ähnlichkeiten zu Tn4501 auf. Das Transposon wurde in dieser Arbeit erstmals sequenziert.

Invertierte Wiederholungs-Sequenz (Ergebnis vom Programm Einvert):

```
Score 54: 30/36 ( 83%) matches, 1 gaps
  8756 catccatcgatgagatataactcgaatttagtgatgg 8792
        |||| || || ||| ||| || |||||
  10196 gtagctacct-ctcaataacagtttaaatcacatacc 10161
```

Direkte Wiederholung der Insertions-Sequenz vor und nach dem Transposon konnten nicht identifiziert werden.

Mögliche Transposase: ORFs 9,10,11. Die Transposase ist unterbrochen und daher wohl nicht mehr funktionsfähig.

Die Namensgebung Tn4501 und Tn4501/2 (QUACKENBUSH ET AL., 1986A,B) ist etwas unglücklich gewählt, da es sich bei den beiden Transposons um "Insertions-Sequenzen" handelt (die nur für eine Transposase kodieren) und nicht um zusammengesetzte Transposons des Typs Tn.

Die fehlenden direkten Wiederholungssequenzen und die fragmentierte Transposase lassen darauf schließen, dass Tn4501/2 die Fähigkeit zur Transposition verloren hat. Das Transposon scheint ein Überrest eines relativ lange zurückliegenden Transpositions-Ereignisses zu sein. Dieses nichtfunktionelle Transposon ist eingebettet in ein weiteres Transposon:

3.4.2.7.3 Transposon Tn4503

Das Transposon ist lokalisiert zwischen den Basen 5591 und 13333 (=7742 Bp). Es wurde bis zu dieser Arbeit noch nicht sequenziert.

Invertierte Wiederholungs-Sequenz (Ergebnis vom Programm Einvert):

```
Score 80: 32/36 ( 88%) matches, 0 gaps
      5591 ggggtctccataaagatttgggccaaaattcccactt 5626
          |||| ||| ||||||||||||| ||||||||||| |||
      13333 cccataggaatttctaaaccagttttaaggcgaa 13298
```

Direkte Wiederholung der Insertions-Sequenz vor und nach dem Transposon:

Linker Terminus	GTCATT
Rechter Terminus	GTCATT

Mögliche Transposase: ORF 8, zwischen den Basen 6205 und 8712 (reverse), Ähnlichkeiten zu bekannten konservierten Protein-Domänen:

[gn1|CDD|6485](#), pfam01526, Transposase_7, Transposase. This family includes transposases of Tn3, Tn21, Tn1721, Tn2501, Tn3926 transposons from *E. coli*. The specific binding of the Tn3 transposase to DNA has been demonstrated. Sequence analysis has suggested that the invariant triad of Asp689, Asp765, Glu895 (numbering as in Tn3) may correspond to the D-D-35-E motif previously implicated in the catalysis of numerous transposases.
CD-Length = 968 residues, 83.3.1% aligned
Score = 213 bits (543), Expect = 8e-56

Andere auf dem Transposon lokalisierte putative Gene mit Aminosäure-Sequenzähnlichkeiten zu schon bekannten Proteinen:

ORF 12 kodiert für ein putatives Protein, das Ähnlichkeiten mit einer site-spezifischen Rekombinase hat. Dieses putative Protein könnte ebenfalls an der Transposition des Transposons Tn4503 beteiligt sein. Des weiteren ist mit dem ORF 7 eine weitere (wahrscheinlich inaktive, da nur fragmentarische) Transposase (oder Rekombinase) auf dem Transposon lokalisiert.

Andere auf dem Transposon lokalisierte putative Gene ohne Aminosäure-Sequenzähnlichkeiten zu schon bekannten Proteinen:

ORF 13 und ORF 14 liegen ebenfalls auf dem Transposon Tn4503.

3.4.2.7.4 Putatives Transposon Tn4504

Das putative Transposon befindet sich zwischen den Basen 25381 und 33379 (=7998 Bp). Es wurde bisher noch nicht beschrieben.

Invertierte Wiederholungs-Sequenz (Ergebnis vom Programm Einvert):

```
Score 706: 258/269 ( 95%) matches, 2 gaps
25381 ttttgatatccatacacaataatcgaatcggctatcctcggcgggtataaattttgcccgggccgctccgcct
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
33379 aaaacatataggtatgtgttttagcttagccgataggagcccatattttaaacgggcccggcagggcgga
```

Die invertierte Wiederholungs-Sequenz ist nur teilweise wiedergegeben. Die Länge der invertierten Wiederholungs-Sequenz deutet auf ein zusammengesetztes Transposon des Typs Tn5 (nicht replikativ) hin.

Direkte Wiederholung der Insertions-Sequenz vor und nach dem Transposon:

Linker Terminus	TCATGATAGA
Rechter Terminus	TCTATCATGA

Die beiden möglichen direkten Wiederholungen weisen nur eine relativ geringe Ähnlichkeit zueinander auf. Es ist also davon auszugehen, dass - falls es sich hierbei um ein Transposon handelt - das Transpositionsereignis evolutionär schon länger zurückliegt.

Mögliche Transposase: ORF 32 besitzt Ähnlichkeiten zu bekannten konservierten Protein-Domänen:

```
gn1|CDD|16051, pfam04754, Transposase_31, Putative transposase, YhgA-like. This family of putative transposases includes the YhgA sequence from Escherichia coli  
CD-Length = 288 residues, 92.4% aligned  
Score = 115 bits (289), Expect = 6e-27
```

Andere auf dem Transposon lokalisierte putative Gene mit Aminosäure-Sequenzähnlichkeiten zu schon bekannten Proteinen:

ORF 39 kodiert für ein putatives bakterielles DNA-Binde-Protein. In seiner Aminosäure-Sequenz weist es Ähnlichkeiten zu einem Integration Host Factor auf.

Andere auf dem Transposon lokalisierte putative Gene ohne Aminosäure-Sequenzähnlichkeiten zu schon bekannten Proteinen:

Die ORFs 26-31, 33-38 und 40 sind ebenfalls auf dem putativen Transposon Tn4504 gelegen.

Die nächste Seite zeigt die Grafik von 3.4.2.6 mit den zusätzlich eingetragenen identifizierten Transposons.

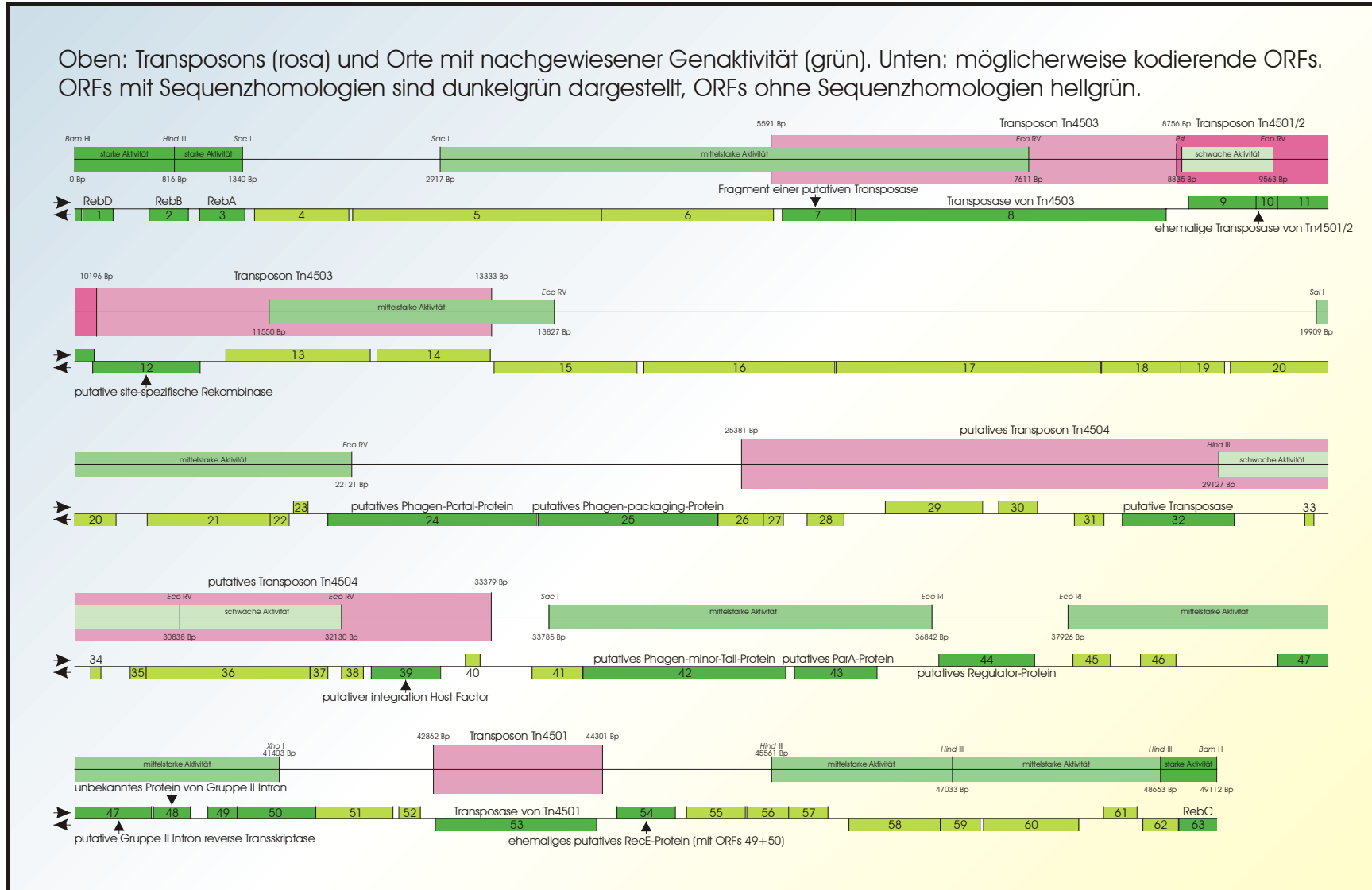


Abb. 3.4.2.7.1: Identifizierte Transposons (rosa) und Orte mit detektierter Transkriptions-Aktivität (jeweils obere Linie der Linienpaare) im Vergleich zu den von GeneMark.hmm ermittelten möglicherweise kodierenden ORFs (jeweils untere Linie). ORFs mit Sequenzhomologien sind dunkelgrün dargestellt, ORFs ohne Homologien hellgrün. Die oberen vier Linienpaare entsprechen jeweils 10 kB des Gesamtplasmids, das untere Linienpaar nur 9 kB

3.4.2.8 Mögliche Operons / polycistronische mRNA

In Bakterien wird mRNA häufig polycistronisch transkribiert, d.h. mehrere Gene, die an einem ähnlichen Stoffwechselfvorgang beteiligt sind und als Cluster dicht beieinander liegen, werden als eine zusammenhängende mRNA transkribiert (CHERRY, 2003). Ein Beispiel dafür ist das Lac-Operon aus *E. coli*, dessen 3 Strukturgene (*lacZ*, *lacY* und *lacA*) als polycistronische mRNA transkribiert werden. Die mRNA wird dann geschnitten und die einzelnen Gene translatiert (MURAKAWA ET AL., 1991). Auf vielen Bakteriophagen sind Gene, die zum gleichen Phagen-Zyklus gehören, nebeneinander angeordnet und unterliegen der regulativen Kontrolle eines einzigen Operons. Auch sie werden polycistronisch transkribiert (siehe Bakteriophage Lambda, SEYFFERT, 1998).

Auch auf pKAP298 waren solche möglicherweise polycistronisch transkribierten Regionen zu finden. Um eine solche Region identifizieren zu können, wurden folgende Kriterien aufgestellt:

1. die möglicherweise polycistronisch transkribierten Gene mussten nahe beieinander liegen und die gleiche Orientierung haben
2. die Gene durften nicht auf einem Transposon oder einem Gruppe-II-Intron liegen, sondern die Einheit musste in Gänze auf dem "ursprünglichen" pKAP298 lokalisiert sein
3. das putative Cluster musste entweder ganz auf einer transkriptionsaktiven pKAP-Region liegen, oder es durfte - falls die Einheit bei der mRNA-Isolation nicht transkribiert wurde - auf keinem Gen des Clusters Aktivität nachgewiesen werden. Falls eine putative Einheit zum Teil auf einem als aktiv erkannten pKAP298-Fragment lokalisiert war, musste auf diesem Fragment noch mindestens ein weiterer, nicht zu diesem Cluster gehörender, ORF (bzw. ein anderes Cluster) lokalisiert sein, der möglicherweise für die detektierte Aktivität verantwortlich sein konnte.

Anhand dieser Kriterien konnten 8 möglicherweise polycistronisch transkribierte Cluster auf dem Plasmid pKAP298 identifiziert werden. Diese Regionen wurden in die aus 3.4.2.6 und 3.4.2.7 bekannte pKAP298-Karte eingetragen (siehe nächste Seite, gelbe Regionen). Ob diese Regionen wirklich polycistronisch transkribiert werden, ist noch experimentell zu überprüfen (z.B. über RT-PCR). Auch über mögliche Regulations-Mechanismen lässt sich noch nichts aussagen.

Oben: Transposons (rosa) und Orte mit nachgewiesener Genaktivität (grün). Unten: mögliche polycistronische Cluster (gelb), möglicherweise kodierende ORFs. ORFs mit Sequenzhomologien sind dunkelgrün dargestellt, ORFs ohne Sequenz-Homologien hellgrün.

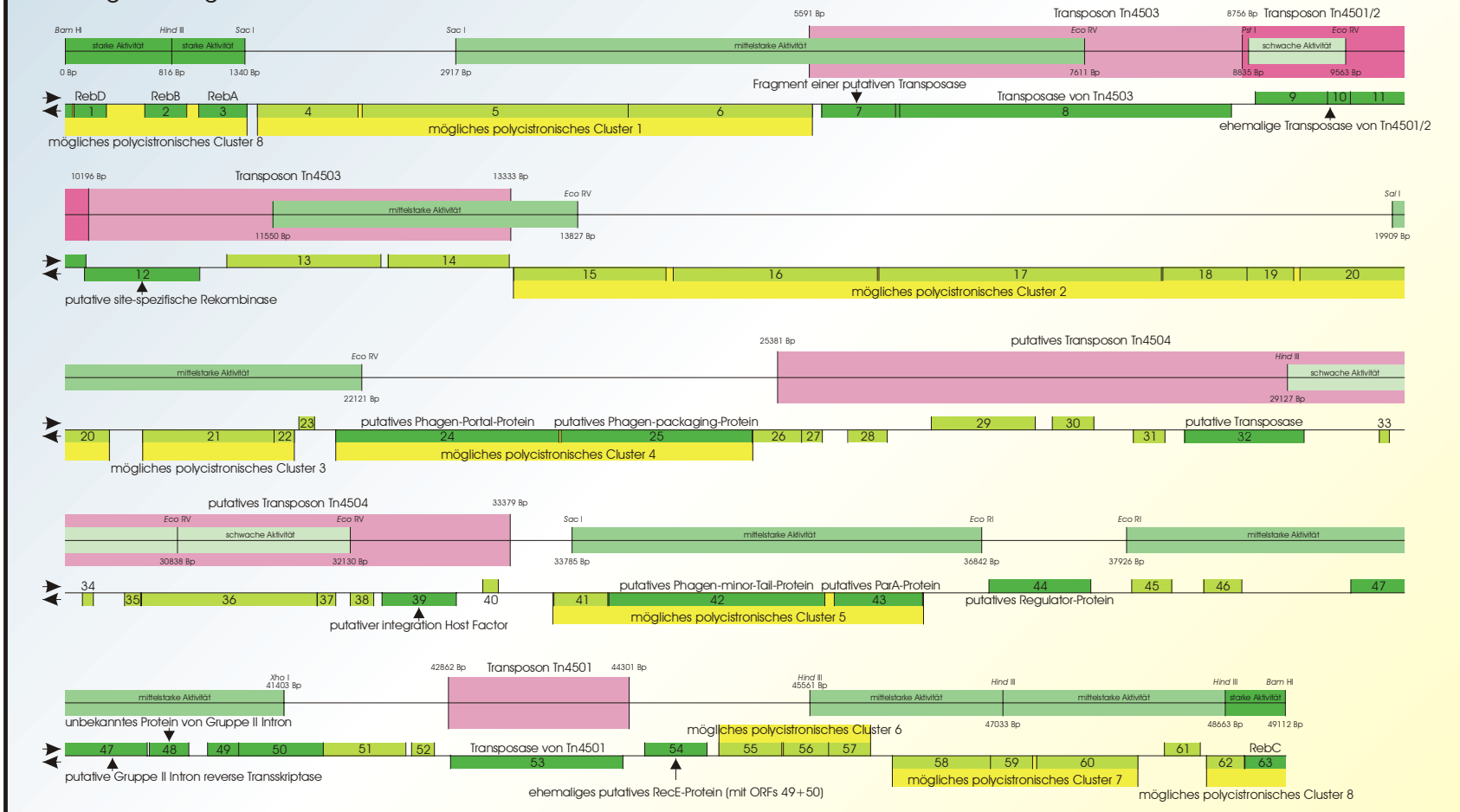


Abb. 3.4.2.8.1: putative polycistronische Cluster auf dem Plasmid pKAP298 (jeweils untere Linie der Linienpaare, gelb). Identifizierte Transposons (rosa) und Orte mit detektierte Transkriptions-Aktivität (grün) sind jeweils auf der oberen Linie eingetragen, von GeneMark.hmm ermittelten möglicherweise kodierenden ORFs jeweils auf der unteren Linie. ORFs mit Sequenzhomologien sind dunkelgrün dargestellt, ORFs ohne Homologien hellgrün. Die oberen vier Linienpaare entsprechen jeweils 10 kB des Gesamtplasmids, das untere Linienpaar nur 9,1 kB.

4. Diskussion

4.1 Transkriptions-Aktivität auf dem Plasmid pKAP298

Bisher wurden noch keine Untersuchungen zur *in vivo* Transkriptions-Aktivität auf dem Plasmid pKAP298 durchgeführt. Die für die R-Körper kodierenden Gene wurden zum Beispiel durch Subklonierung einzelner pKAP-Fragmente (aus pKAP116 und pKAP47) in *E. coli*, Analyse der translatierten Proteine und anschließende Sequenzierung identifiziert (HERUTH ET AL., 1994). Da auf einem 49112 Bp großen Plasmid mehr potentielle Gene zu erwarten sind, die exprimiert werden, galt es eine Methode zu entwickeln, die eine Lokalisation solcher transkriptionsaktiven Orte auf dem Plasmid ermöglicht. Der Ansatz war, diese Orte über die dort transkribierte mRNA zu lokalisieren. Dazu wurde die RNA isoliert, in cDNA transkribiert, markiert und mit Southern Blots, auf denen sich verdaute pKAP298-DNA befand, hybridisiert und detektiert. Anhand von positiven Signalen aus der Detektion konnten dann Rückschlüsse auf Orte mit Gen-Aktivität gezogen werden.

4.1.1 Analyse der Transkriptions-Aktivität

4.1.1.1 ORFs mit Protein-Homologien auf Fragmenten mit nachgewiesener Gen-Aktivität

Die stärkste Transkriptions-Aktivität konnte auf den Plasmid-Fragmenten nachgewiesen werden, auf denen die R-Körper kodierenden Gene lokalisiert sind. Laut POND ET AL. (1989) werden die R-Körper verstärkt unter Stressbedingungen (Hunger, UV-Bestrahlung) exprimiert. Unter den bei der RNA-Isolation herrschenden Bedingungen (Zellen im Hungerzustand) konnte also mit einer Transkriptions-Aktivität der R-Körper Gene gerechnet werden. Diese R-Körper-Gene sind bei jedem Mitglied der pKAP-Familie um die einzige *Bam* HI - Schnittstelle gruppiert (HERUTH ET AL., 1994). Auf dem bearbeiteten Plasmid pKAP298 liegen sie in den Bereichen 1-1340 Bp und 48663-49112 Bp. Das RebD-Gen (GeneMark ORF 1) erstreckt sich von 62-307 Bp, RebB (ORF 2) von 590-907, RebA zwischen 988-1332 Bp und RebC (ORF 63) von 48821-53 Bp (über die *Bam* HI - Schnittstelle als Plasmid-Startpunkt). Die Tatsache, dass auf den Fragmenten mit den R-Körper-Genen Transkriptions-Aktivität nachgewiesen werden konnte, zeigt, dass die Methode der Aktivitätsbestimmung über RNA-Isolation und die Hybridisierung des markierten cDNA-Transkripts funktionierte. Ob die R-Körper-Gene polycistronisch oder monocistronisch transkribiert werden, und wie die Transkription/Translation reguliert wird, ist noch unbekannt.

Alle anderen Fragmentbanden, die positive Signale lieferten, wiesen durchgehend eine geringere Signalstärke auf. Dies kann mehrere Gründe haben. Es könnten auf diesen Fragmenten weniger oder kleinere Gene transkribiert worden sein, oder die potentiellen Gene (bzw. polycistronischen Cluster) besitzen schwächere Promotoren, die nur eine kleinere Anzahl von mRNA-Transkripten zulassen. Ein weiterer Grund für die geringere Signalstärke könnte sein, dass die Expression bestimmter Gene (bzw. Operons) (in dem Zellstadium, in dem die RNA-Isolation stattfand) durch Regulatoren teilweise inhibiert worden war. Gleiches gilt für Bereiche, auf denen keine Transkriptions-Aktivität nachgewiesen werden konnte. Dort können in anderen Zellstadien sehr wohl Gene (Cluster) exprimiert werden, diese waren aber bei der RNA-Isolation nicht aktiv. Als mögliche Regulatoren kämen unter anderem die ORFs 39 und 44 in Betracht. Für Proteine, die zur Familie der Integration Host Factor gehören (zu denen wohl auch ORF 39 zählt), konnten in diversen Bakterien transkriptions-regulatorische Funktionen nachgewiesen werden (KIRNDORFER ET AL., 1998, OBERTO ET AL., 1994). ORF 44 weist Ähnlichkeit zu der GerE Proteinfamilie auf, die als Transkriptions-Aktivator klassifiziert wird (CUTTING ET AL., 1986). In *Bacillus subtilis* zum Beispiel, spielt GerE bei der Regulation, und damit bei der Expression, bestimmter Gene während der Sporulation eine Rolle (ZHENG ET AL., 1992).

Ursprünglich sollte anhand der auf dem Plasmid nachgewiesenen Transkriptions-Aktivität versucht werden, die transkribierten ORFs eindeutig zu identifizieren. Aufgrund der Größe der verdauten und dann geblotteten pKAP-Fragmente, und der daraus folgenden großen Zahl an potentiellen ORFs pro Fragment, gelang dies jedoch nicht. Daher wurde darauf verzichtet die einzelnen putativ transkriptionsaktiven ORFs (bzw. Gen-Cluster) auf mögliche Promotoren und andere, die Wahrscheinlichkeit der Transkription betreffende, Merkmale zu untersuchen. Die erhaltenen Daten wären rein hypothetisch gewesen. Mit den durch diese Arbeit erhaltenen Daten ist es in Zukunft möglich, die betreffenden ORFs auf ihre tatsächliche Transkriptions-Aktivität zu untersuchen.

Von den identifizierten ORFs mit Protein-Sequenz-Homologien zu schon bekannten Proteinen lagen viele nicht (oder nur teilweise) in Bereichen mit nachgewiesener Gen-Aktivität. Dies bedeutet allerdings nicht, dass die potentiellen Gene gar nicht exprimiert werden, sondern zunächst einmal nur, dass sie in dem Zellstadium, in dem sich die *Caedibacter* bei der Isolation befanden, nicht exprimiert wurden (siehe oben) oder die produzierte mRNA unter der Nachweisgrenze dieses Verfahrens lag. Um ORFs als potentiell aktiv werten zu können, mussten diese ganz (oder zumindest zum größten Teil) in Regionen mit nachgewiesener Genaktivität liegen. Eine nur partielle Lokalisierung eines aktiven ORFs

im Bereich positiver Signale hätte logischerweise auch im angrenzenden, negativen pKAP-Fragment zu Signalen geführt. Konnten diese nicht nachgewiesen werden, war der betreffende ORF auch nicht aktiv. Diese Bedingung erfüllten nur wenige der ORFs mit Protein-Sequenz-Homologien. Dies waren unter anderem die oben beschriebenen R-Körper-Gene (ORFs 1-3 und 63).

Ein durchschnittlich mittelstarkes Signal lieferte der Bereich 2917 Bp bis 7611 Bp. Auf ihm sind 2 ORFs mit Protein-Ähnlichkeiten zu finden, erstens der ORF 7, als Fragment einer möglichen Transposase/Rekombinase, und zweitens der ORF 8, der ebenfalls Homologien zu putativen Transposasen aufweist. ORF 8 befindet sich nur zu ca. 60 % auf dem positiven Bereich. Da auf dem benachbarten Fragment keine Aktivität nachgewiesen werden konnte, scheidet ORF8 wahrscheinlich als Verursacher des Signals aus. Auf einem Bereich mit schwachem Signal (8835-9563 Bp) liegt der ORF 9, ein Fragment einer vormals größeren Transposase vom Transposon Tn4501/2. Der Bereich zwischen 33785 Bp und 36842 Bp wies im Durchschnitt ein mittelstarkes Signal auf. Dies ist besonders interessant, da dort zwei putative Gene mit Protein-Sequenz-Homologien lokalisiert sind. Zum Einen ist dies der ORF 42. Laut BLAST weist er Ähnlichkeiten zu Phagen-Schwanz-Proteinen auf - er könnte für mitunter die mit den R-Körper assoziierten helikalen Protein-Strukturen kodieren (siehe Einleitung 1.4.1, PREER ET AL., 1972). Zum anderen liegt in dem Bereich der ORF 43, der möglicherweise (als putatives Soj/ParA-Enzym) an der Replikation des Plasmids beteiligt ist (DAVEY ET AL., 1994). Beide Gene könnten möglicherweise polycistronisch transkribiert werden (siehe 3.4.2.8, Cluster 5). Geht man von der These aus, dass das Plasmid pKAP298 von einem Bakteriophagen abgeleitet ist (PREER ET AL., 1974), macht es durchaus Sinn, dass beide ORFs zum Zeitpunkt der Isolation aktiv sind. Da die *Caedibacter* bei der Isolation wahrscheinlich in einen Stresszustand versetzt wurden, könnten auf dem Plasmid die für einen lytischen Zyklus kodierenden späten Phagen-Gene induziert worden sein (KLAUS ET AL., 1992). Dies wären die R-Körper-Gene (als möglicherweise aberrante Capsid-Proteine), ein putatives Phagen-Schwanz-Gen ORF 42 und die an der Replikation des Phagen-Genoms beteiligten Enzyme, zu denen wahrscheinlich ORF 43 gehört. Auf dem Bereich zwischen 37926 Bp und 41403 Bp, der durchschnittlich mittelstarke Signale ergab, liegen 4 putative Gene mit Protein-Ähnlichkeiten, davon einer (ORF 50) nur zu zwei Dritteln - er scheidet daher wahrscheinlich als Verursacher der Signale aus. Bei den anderen handelt es sich um die ORFs 47, 48 und 49. Die ORFs 47 und 48 sind wahrscheinlich Teil eines Gruppe-II-Introns (siehe 4.2.2), wobei ORF 47 für eine reverse Transkriptase und ORF 48 für ein unbekanntes Protein kodiert. ORF 49 hat (wie ORF 50 und 54) Ähnlichkeiten mit dem aus *E. coli* bekannten Protein RecE (FOUTS ET AL., 1983).

4.1.1.2 ORFs ohne Protein-Homologien auf Fragmenten mit nachgewiesener Gen-Aktivität

Die Eigenschaften (Länge, Lokalisierung und Orientierung) der hier aufgeführten Open Reading Frames können im Abschnitt 3.4.2.4 nachgelesen werden. In dem mittelstarke Aktivität aufweisenden Bereich zwischen 2917 Bp und 7611 Bp sind (neben ORF 7 und 8, siehe oben) zwei ORFs ohne bekannte Protein-Homologien lokalisiert. ORF 5 liegt nur zu 65 % im Signal-Bereich und scheidet daher als Signalgeber aus. Der ORF 6 dagegen liegt ganz auf dem Fragment und könnte daher für die Transkriptions-Aktivität verantwortlich sein. Geht man jedoch davon aus, dass die ORFs 4 bis 6 polycistronisch transkribiert werden (siehe 3.4.2.8, Cluster 1), ist ORF 6 als Verursacher des Signals unwahrscheinlich, da für ORF 4 keine Signale detektiert werden konnten. Es bliebe also nur noch ORF 7. Zwischen 11550 Bp und 13827 Bp wurden ebenfalls durchschnittlich mittelstarke Signale detektiert. Auf diesem Fragment liegen 3 potentiell kodierende ORFs. Die ORFs 13 und 15 liegen nur zum Teil auf dem Fragment (ORF 13 zu 71 % und ORF 15 zu 41 %) und sind daher vermutlich nicht für die positiven Signale verantwortlich. Das putative Gen ORF 14 liegt dagegen ganz auf dem Fragment und liefert demzufolge wahrscheinlich die Signale. Die ORFs 21 bis 23 sind auf einem Bereich mittelstarker Signale zwischen 19909 Bp und 22121 Bp lokalisiert, der ORF 20 nur zu 38 %. Er wird daher ausgeschlossen. Die ORFs 22 und 23 sind sehr klein (150 Bp und 117 Bp). Die durchschnittliche Gengröße der bisher sequenzierten Bakteriengenome beträgt zwischen 900 und 1000 Bp (CASJENS, 1998). Daraus könnte man folgern, dass es eher größere Gene sind, die exprimiert werden, und dass daher die ORFs 22 und 23 wahrscheinlich nicht für die positiven Signale dieses Fragments verantwortlich sind. Es bleibt also noch der ORF 21 als möglicherweise transkribiertes Gen. Andererseits ist ORF 22 direkt neben ORF 21 gelegen und könnte zusammen mit diesem transkribiert werden (siehe 3.4.2.8, Cluster 3). Der Bereich von 29127 Bp bis 32130 Bp weist schwache Transkriptions-Aktivität auf. Auf ihm liegen die ORFs 33 bis 37. Der wahrscheinlichste Verursacher des positiven Transkriptions-Signals ist ORF 36 (1311 Bp), da die anderen nur sehr kurz sind (zwischen 75 Bp und 114 Bp). ORF 45 und ORF 46 liegen auf dem mittelstark-aktiven Bereich zwischen 37926 Bp und 41403 Bp, wo auch die ORFs mit Protein-Homologien 47 bis 50 zu finden sind. Welche dieser Gene am wahrscheinlichsten aktiv sind lässt sich nicht sagen. Das letzte Fragment mit nachgewiesener Transkriptions-Aktivität liegt zwischen 45561 Bp und 48663 Bp. Dort wurden durchschnittlich mittelstarke Signale detektiert. Auf diesem Bereich befinden sich die ORFs 56 bis 62, wobei ORF 56 und ORF 62 nur zu 42 % bzw. 41 % auf dem Fragment liegen. ORF62 befindet sich zu 59 % auf einem Bereich starker Aktivität, wo auch die R-Körper Gene lokalisiert sind (48663 Bp bis 49112 Bp). Im Gegensatz zu ORF 56 kann ORF 62 daher nicht als Signalgeber ausgeschlossen werden. Von den ORFs 57 bis 61

sind ORF 58 und ORF 60 die größten (732 Bp und 759 Bp). Eine tendenzielle Aussage, welcher ORF die Signale liefert, ist aber nicht möglich. Die ORFs 58 bis 60 könnten als Cluster transkribiert werden (siehe 3.4.2.8, Cluster 7) - dann wären alle 3 ORFs potentiell für die Signale verantwortlich. Die ORFs 55 bis 57 könnten ebenfalls polycistronisch transkribiert werden (siehe 3.4.2.8, Cluster 6), was die ORFs 56 und 57 als Signal-Produzenten ausschließen würde, da für ORF 55 keine Aktivität nachgewiesen werden konnte.

4.1.2 Weitere Forschungsansätze für zukünftige Untersuchungen am Plasmid pKAP298

Eine genaue Aussage darüber, ob es die oben genannten potentiellen Gene sind, die auf den entsprechenden positiven Fragmenten für Transkriptions-Aktivität gesorgt haben, ist derzeit nicht möglich. Diese Arbeit bildet jedoch die Basis für weitere Untersuchungen an dem Plasmid pKAP298 und dessen potentiellen Genen. Da nun die Basensequenz des Plasmids bekannt ist, und gleichzeitig Bereiche auf dem Plasmid mit sicherer Transkriptions-Aktivität beschrieben worden sind, ist es in Zukunft möglich, die betreffenden aktiven ORFs und deren Produkte zu identifizieren. Aus den betreffenden Bereichen des Plasmids, die Gen-Aktivität aufwiesen, könnte man die potentiellen codierenden ORFs heraussuchen und aus deren Sequenz 100-200 Bp lange Stücke auswählen, die für den jeweiligen ORF typisch sind, also nicht mit ORFs in anderen Leserastern überlappen. Am besten wählt man eine Region nahe des Stopkodons, da das genaue Startkodon, an dem die Transkription beginnt, mit GeneMark.hmm nicht genau vorhergesagt werden kann (LUKASHIN UND BORODOVSKI, 1998). Diese Stücke könnten (komplementär) als Oligonucleotide synthetisiert und in gleicher Konzentration nebeneinander auf eine Nylonmembran fixiert werden. Diese Membran lässt man wiederum mit markiertem RNA/cDNA-Transkript hybridisieren und kann dann, im Idealfall, die zur Zeit der RNA-Isolation aktiven ORFs identifizieren. Die dadurch erhaltenen Ergebnisse liefern dann möglicherweise auch Hinweise auf Operons und polycistronisch transkribierte Gen-Cluster.

Die während der Arbeit hergestellten *E. coli*-Klone mit pKAP298-Fragmenten können auch weiterhin für Studien zur Genaktivität benutzt werden. Die zu untersuchenden pKAP-Fragmente können in beliebiger Menge hergestellt werden. Die Fragment-DNA könnte dann mit anderen, als den in dieser Arbeit benutzten, Restriktionsenzymen geschnitten und geblottet werden. Dadurch lassen sich vielleicht die Orte mit Transkriptions-Aktivität weiter eingrenzen. Ein Problem ist jedoch, dass je kleiner und öfter die Fragmente geschnitten werden, desto näher liegen auch die entstehenden Banden auf dem Blot beieinander und desto

schwieriger wird die genaue Auswertung. Hinsichtlich der Identifizierung von transkriptionsaktiven ORFs auf dem Plasmid, wäre deshalb die oben beschriebene Methode über Oligonucleotide die bessere Wahl.

Von den R-Körper-Genen ist bekannt, dass sie in *E. coli* ohne einen im Vektor vorhandenen Promotor ausgeprägt werden (QUACKENBUSH ET AL., 1983B). Somit werden also zumindest die Promotoren der R-Körper-Gene von *E. coli* erkannt. Vermutlich befinden sich auf den subklonierten pKAP-Fragmenten noch weitere ORFs mit Promotoren, die von *E. coli* erkannt werden. Man könnte also aus den entsprechenden Bakterien mit pKAP-Fragmenten die Gesamt-RNA isolieren und in cDNA umschreiben. Diese wird wieder markiert und mit den oben beschriebenen Oligonucleotiden aus den zu untersuchenden ORFs hybridisiert. So könnte man auf die aufwändige Gesamt-RNA-Isolation aus den originalen *Caedibacter*-Kulturen verzichten und trotzdem mögliche Genaktivität nachweisen. Die Sicherheit, dass wirkliche alle in *Caedibacter* aktiven ORFs erfasst werden, ist allerdings mit dieser Methode nicht gegeben, da zu den ORFs gehörende Promotoren von *E. coli* vielleicht nicht erkannt werden, bzw. dass z.B. ein Regulator-Protein aus *Caedibacter* fehlt (oder dass das Regulator-Gen auf einem anderen pKAP298-Fragment liegt, z.B. ORF 39 und 44, siehe 4.1.1.1). Auch mögliche Induktoren (z.B. Stoffwechselprodukte aus *Caedibacter*) von Operons könnten fehlen. Andererseits könnten in *E. coli* auch mehr potentielle Gene exprimiert werden, da möglicherweise ein (oder mehrere), von einem anderen pKAP-Fragment oder aus *Caedibacter* selbst stammender, Inhibitor fehlt. Mittels des Gesamt-RNA-Isolationsverfahren aus rekombinanten *E. coli* könnte man allerdings noch weitere Studien zur Gen-Expression auf dem Plasmid pKAP298 betreiben, z.B. ob zu verschiedenen Zellstadien andere ORFs transkribiert werden. Aus den rekombinanten *E. coli* mit pKAP-Fragmenten ließe sich die Gesamt-RNA-Isolation zu verschiedenen Zeitpunkten (oder Kultur-Bedingungen) sehr einfach und schnell durchführen, ohne dass das Ergebnis durch eine lange Isolationszeit verfälscht werden würde. Möglicherweise könnten Gene aus dem vegetativen Zyklus und Stress-Gene (z.B. Hunger-, Hitze-, oder UV-Stress) identifiziert und unterschieden werden.

4.1.3 Analyse möglicher Fehlerquellen

Die Methode, über mRNA Orte mit Gen-Aktivität zu identifizieren, stellte sich in diesem Fall als sehr schwierig durchzuführen dar. Das erste Problem war, dass der Wirt des zu untersuchenden Plasmids, *Caedibacter taeniospiralis*, ein obligater Endobiont von *Paramecium tetraurelia* ist, also nicht außerhalb von *Paramecium tetraurelia* kultiviert werden kann (POND ET AL., 1989). Zur Gewinnung der RNA mussten also immer Massenkulturen von *Paramecium tetraurelia* angelegt werden, aus denen dann, in einem

mindestens 30 minütigem Verfahren, zunächst die *Caedibacter*-Bakterien isoliert werden mussten. Da mRNA in Bakterien eine sehr geringe Lebensdauer von durchschnittlich nur einer bis maximal 30 Minuten hat (GRUNBERG-MANAGO, 1999), die Isolation aber mindestens 30 min dauerte, war es nicht möglich eine "Momentaufnahme" des mRNA-Zustands der Zellkultur beim Abernten zu erhalten. Ein "Einfrieren" des momentanen mRNA-Zustands der Bakterien ist direkt in den Massenkulturen nicht möglich, da die Kulturen zu groß sind (16 Liter), als dass man sie z.B. innerhalb von 1-2 Minuten auf 0°C abzukühlen und damit die Neubildung (bzw. den Abbau) von mRNA verhindern könnte. Ein (wahrscheinlich) verfälschender Isolations-Schritt ist also bisher unumgänglich. Es wurde RNA aus gefütterten und hungernden Zellkulturen isoliert und hybridisiert - die entstehenden Blots wiesen ein sehr ähnliches Bandenmuster auf, wobei keine einheitlichen Unterschiede zwischen RNA-Isolaten aus gefütterten und ungefütteten Kulturen gefunden werden konnten. Die RNA-Isolate aus gefütterten Zellkulturen sollten den mRNA-Zustand von *Caedibacter taeniospiralis* in der Wachstumsphase dokumentieren, RNA-Isolate aus hungernden Kulturen den mRNA-Zustand im (Hunger-)Stress. Laut DILTS (1986) sollen die für die R-Körper kodierenden Gene auf dem Plasmid vorwiegend bei Zellen im Stress aktiv sein. Die ausgewerteten Blots von RNA-Isolaten aus Hungerkulturen zeigten dies auch sehr deutlich anhand der starken positiven Signale auf den Plasmid-Fragmenten, auf denen die R-Körper-Gene lokalisiert sind. Allerdings zeigten RNA-Isolate aus gefütterten Kulturen ebenfalls starke Aktivität auf den entsprechenden Plasmid-Bereichen. Dies könnte darauf hindeuten, dass die mindestens 30 Minuten dauernde Isolation der *Caedibacter* aus den Paramecien die Bakterien aus den gefütterten Kulturen ebenfalls in einen Stress-Zustand versetzt hat. Wahrscheinlich hat sich die Transkriptions-Aktivität der gefütterten Kulturen der Transkriptions-Aktivität von Kulturen im Hungerzustand angeglichen. Mit dem hier beschriebenen Verfahren ist also "nur" eine Analyse der mRNA aus *Caedibacter* im (Hunger-)Stress-Zustand möglich. Deshalb wurde auch auf eine getrennte Auswertung von detektierten Blots aus gefütterten und hungernden Kulturen verzichtet und die Ergebnisse aus beiden Untersuchungen zusammen ausgewertet. Die entstandene Karte des Plasmids pKAP298, auf der Orte mit detektierter Transkriptions-Aktivität eingetragen sind, stellt also eine Synthese der ausgewerteten Blots aller RNA-Isolate dar.

Eine weitere Schwierigkeit war, dass der Wirt des Plasmids ein Prokaryot ist, die mRNA also nicht selektiv anhand des Poly-A-Schwanzes isoliert werden konnte. Es musste daher immer die Gesamt-RNA für die Weiterverwendung isoliert werden. Da der durchschnittliche Anteil von mRNA an der Gesamt-RNA in einer Bakterien-Zelle normalerweise nur ungefähr 5 % beträgt (der Rest ist zum größten Teil ribosomale RNA, LEWIN 2002), befand sich folglich sehr viel störende andere RNA in den Isolaten. Außerdem ist davon auszugehen, dass der

Anteil der vom Plasmid stammenden mRNA wiederum nur einen kleinen Anteil an der gesamten mRNA ausmacht. Die große Kontamination mit nicht-mRNA könnte sich in dem (unselektiven) Umschreiben in cDNA und anschließendem (ebenfalls unselektiven) Markieren der cDNA bemerkbar machen. Da viel mehr der kontaminierenden (m)RNA vorhanden ist als erwünschte mRNA, wird auch mehr davon umgeschrieben und markiert, was die Menge an potentiell geeigneten Sonden weiter senkt.

Die aus der Gesamt-RNA hergestellten Sonden sollten auf Fragmente des Plasmids pKAP298 hybridisiert werden. Verfahrensbedingt befand sich im RNA-Isolat auch das Plasmid pKAP298. Würde ein solches kontaminiertes Isolat markiert und hybridisiert werden, würden alle pKAP-Fragmente auf dem Blot starke Signale geben, da das bei der Markierungsreaktion ebenfalls markierte pKAP-Kontaminat an seine Fragment-Pendants auf dem Blot hybridisiert. Dies kam während dieser Arbeit sehr häufig vor, was mehrere Tage Arbeit zunichte machte, da das Endergebnis der ganzen Schritte erst nach einigen Tagen sichtbar war. Der einzige Ausweg war ein langer DNaseI-Verdau der Gesamt-RNA nach der Isolation. Dieser durfte allerdings auch nicht zu lange dauern, da möglicherweise noch im Isolat vorhandene RNAsen die RNA (zumindest teilweise), trotz RNase-Inhibitor, abgebaut hätten. Auch dies konnte während der Arbeit oftmals beobachtet werden.

Bei der Hybridisierung der markierten Sonden auf Southern Blots mit verdauten pKAP-Fragmenten, die in *E. coli* amplifiziert worden waren, kam es (aufgrund der hohen DNA-Konzentration auf dem Blot) häufig zu unspezifischen Hybridisierungen der Sonde, die dann, auch mit verlängerter Stringenz-Waschung, nicht vollständig entfernt werden konnten. So waren zum Beispiel häufig positive Signale auf den Vektorbanden zu detektieren, die eigentlich keine Signale hätten ergeben sollten, da sie nicht von pKAP298 stammten. Auch eine höhere Hybridisierungs-Temperatur brachte nicht den gewünschten Erfolg. Von einem solchen detektierten Blot konnten nur pKAP-Fragmente als positiv gewertet werden, deren Signalstärke (bei mindestens gleicher Länge) deutlich über der Signalstärke der Vektorbanden lag. Zumindest konnte man bei den Southern Blots, die mit in *E. coli* amplifizierten pKAP-Fragmenten hergestellt wurden, die einzelnen (positiven) Banden eindeutig unterscheiden und den entsprechenden Orten auf dem Plasmid pKAP298 zuordnen. Bei Southern Blots, die aus dem verdauten Gesamt-Plasmid hergestellt wurden, konnten - aufgrund der Größe des Plasmids und der Vielzahl der beim Verdau entstehenden (und sich überlappenden) Banden - positive Signale häufig nicht eindeutig einem bestimmten Fragment, und damit dem entsprechenden Ort auf dem Plasmid, zugeordnet werden. Der Großteil der Aktivitätsanalyse wurde daher, der Eindeutigkeit wegen, an den vorher in *E. coli* subklonierten und amplifizierten pKAP298-Fragmenten durchgeführt. Die Aktivitätsanalyse der Fragmente, die

nicht subkloniert werden konnten, erfolgte dann folgerichtig durch Sonden-Hybridisierung mit Southern Blots, auf denen das verdaute Gesamt-Plasmid aufgetragen war. Diese detektierten Blots wurden dann gezielt auf mögliche Transkriptions-Aktivität auf den nicht subklonierten Fragmenten analysiert. Andere Fragmente mit positiven Signalen wurden für die Erstellung einer Aktivitäts-Karte des Plasmids nicht gewertet. So ließ sich dann eine Transkriptions-Aktivitäts-Karte des Plasmids pKAP298 erstellen, die zwar relativ grob ist, aber dafür mit hoher Wahrscheinlichkeit die tatsächlichen Orte mit Genaktivität auf dem Plasmid wiedergibt. Dabei ist wiederum zu beachten, dass die Karte nur für *Caedibacter*-Isolate gilt, die nach dem in dieser Arbeit beschriebenen Verfahren isoliert wurden, da die Bakterien während der Isolation wohl in einen Stress-Zustand kommen (siehe oben). Wäre es möglich die Bakterien in ihrer Wachstums-Phase auf die Genaktivität zu untersuchen, käme vielleicht eine ganz andere Transkriptions-Aktivitäts-Karte zustande. Leider ist das mit diesem Verfahren nicht möglich, da die Bakterien zunächst einmal aus ihrem Wirt isoliert werden müssen (siehe oben). Ein Ausweg wäre der Nachweis von Transkriptions-Aktivität mittels RT-PCR zu verschiedenen Zellzyklen, wobei die Primer so gewählt werden könnten, dass sie gezielt die mRNA der zu testenden ORFs amplifizieren. Dies bedeutet allerdings für 63 zu testende ORFs einen hohen Material- und Zeitaufwand. Als Template für die Reaktionen müssten wenige (ganze) infizierte Paramecien-Zellen dienen, da eine langwierige Isolation und Aufreinigung von *Caedibacter* wiederum das Ergebnis verfälschen würde. Ob die Template-Menge von zu untersuchender mRNA dann für eine RT-PCR ausreicht, und ob die unerwünschte Gesamt-RNA aus *Paramecium* und *Caedibacter* (sowie sonstige im Ansatz vorhandene Kontaminationen) nicht die Reaktion be- oder verhindern, müsste experimentell geklärt werden. Über RT-PCR könnten auch definitive Aussagen über die mögliche polycistronische Transkription von auf dem Plasmid identifizierten, hypothetischen Gen-Clustern getroffen werden (siehe 3.4.2.8). Die Primer wären in diesem Fall beispielsweise so zu wählen, dass sich das Amplifikat sukzessive über mehrere putativ zu einem Cluster gehörende ORFs erstreckt.

4.2 Evolution des Plasmids pKAP298

Wie in der Einleitung (1.4.1) beschrieben wird vermutet, dass das Plasmid ursprünglich von Bakteriophagen-DNA oder RNA abstammt (PREER ET AL., 1974) und dann zu einem reinen Plasmid evolvierte, bzw. dass die ursprüngliche Phagen-DNA schon als Plasmid vorlag. Diese These kann durch meine Arbeit gestützt und erweitert werden.

4.2.1 Untersuchung der Basensequenz nach Ähnlichkeiten zu Phagen-DNA-Sequenzen

Zunächst wurde die Basensequenz des Plasmids pKAP298 nach Sequenzhomologien zu schon bekannten Phagen-Sequenzen untersucht. Hierbei konnten keine Sequenzhomologien zu den Basensequenzen schon bekannter Virengenome gefunden werden.

4.2.2 Untersuchung der translatierten identifizierten ORFs nach Ähnlichkeiten mit Proteinen viraler Herkunft

Als zweiter Schritt wurden die auf dem Plasmid identifizierten und translatierten, potentiell kodierenden ORFs nach Aminosäure-Sequenzhomologien zu schon bekannten Proteinen viraler Herkunft untersucht. Hierbei wurden einige Übereinstimmungen gefunden, die darauf hindeuten, dass das Plasmid tatsächlich ursprünglich von einem Bakteriophagen abstammt.

Der ORF 25 weist in seiner Aminosäure-Sequenz die größte Ähnlichkeit zu einem hypothetischen Protein aus *Sinorhizobium meliloti* auf, das GALIBERT ET AL. (2001) als putatives DNA packaging Protein GP2 charakterisieren:

```
Score = 273 bits (699), Expect = 3e-72  
Identities = 171/469 (36%), Positives = 240/469 (51%), Gaps = 65/469 (13%)
```

Gleichzeitig bestehen große Ähnlichkeiten zu Terminasen, bzw. DNA-Packaging-Proteinen, von anderen Phagen. Eine 32 % Übereinstimmung (47 % positive Aminosäuren) besteht mit einer putativen Terminase aus dem *Pseudomonas aeruginosa* Phagen PaP3, die von HU ET AL. (2002) beschrieben worden ist. Zu einem hypothetischen Gp2-Protein aus der *Salmonella typhimurium* Bacteriophage ST64T (MMOLAWA ET AL. 2003) weist der ORF 25 eine Sequenz-Übereinstimmung von 30% auf (43 % positive AS). Mit einer von VANDER BYL ET AL. (2000) beschriebenen Terminase (große Untereinheit) aus der Enterobacteria Phage P22 lässt sich eine 32 % Übereinstimmung (44 % positive AS) feststellen. Terminasen, bzw. DNA-Packaging-Proteine, sind an der Beladung des Phagen-Capsids mit dem viralen Genom (DNA oder RNA) beteiligt (FUJISAWA ET AL., 1997). Dies ist, in Verbindung mit dem Auftreten von ikosaedrischen Capsiden in manchen *Caedibacter*-Arten, die teilweise mit DNA beladen sind (PREER ET AL., 1972 und 1974), ein starker Hinweis auf die Phagen-Herkunft des Plasmids pKAP298.

Nur 14 Basen in gleicher Richtung (reverse) auf dem Plasmid pKAP298 entfernt liegt der ORF 24, welcher eine hohe Ähnlichkeit zu einem hypothetischen Protein des *Sinorhizobium meliloti* Phagen PBC5 aufweist (SCHULMEISTER ET AL., 2001):

```
gi|18071218|ref|NP_542303.1| unknown [Sinorhizobium meliloti phage PBC5]
Length = 749
Score = 107 bits (268), Expect = 4e-22
Identities = 112/473 (23%), Positives = 206/473 (43%), Gaps = 30/473 (6%)
```

SCHULMEISTER ET AL. (2001) verweisen in der Annotation zu dem hypothetischen Protein auf eine Ähnlichkeit zu einem putativen Portal-Protein aus dem Bakteriophagen 933W (PLUNKETT ET AL., 1999). Auch der ORF 24 weist Homologien zu diesem hypothetischen Portal-Protein auf:

```
gi|9632518|ref|NP_049512.1| putative portal protein [Bacteriophage 933W]
Length = 714
Score = 79.7 bits (195), Expect = 1e-13
Identities = 98/444 (22%), Positives = 183/444 (41%), Gaps = 47/444 (10%)
```

Portal-Proteine sind bei Phagen ebenfalls an der Beladung des Capsids mit dem Viren-Genom beteiligt (FUJISAWA ET AL., 1997). Sie sind Teil eines Konnektors, der die Verbindung zwischen Kopf und Schwanz eines Bakteriophagen darstellt (BLACK, 1989, KLAUS ET AL., 1992). Die beiden ORFs (24 und 25) liegen auf dem Plasmid relativ weit von möglichen Transposasen, Gruppe-II-Introns und den schon beschriebenen Transposons entfernt, so dass davon ausgegangen werden kann, dass sie vom originalen Phagen-Genom stammen, und nicht durch sekundäre Transpositions-Ereignisse vom Wirtsgenom auf das Plasmid gelangt sind. Laut der durchgeführten Transkriptions-Aktivitäts-Analyse findet an den Stellen, an denen die ORFs 24 und 25 auf dem Plasmid lokalisiert sind, keine Transkriptions-Aktivität statt. Die beiden putativen Gene gehören wahrscheinlich zu den "späten" Phagen-Proteinen, d.h. sie werden exprimiert, wenn im lytischen Zyklus der Phage zusammengebaut wird (LEWIN, 2002). Dies kann z.B. durch UV-Strahlung oder möglicherweise, wie in diesem Fall, durch Stress induziert werden. Da diese Gene, trotz der induzierten Stressbedingungen (ausgehungerte Zellen + lange *Caedibacter* Isolationsdauer), nicht transkribiert werden, könnte bedeuten, dass die Gene inaktiviert sind (dass z.B. die potentiellen Promotoren mutiert sind oder die Gen-Regulation nicht mehr funktioniert). Dies wäre ein Hinweis darauf, dass die putativen Gene evolutionär alt und nicht mehr essentiell sind. Oftmals werden die späten Phagen-Gene nur von einer modifizierten RNA-Polymerase transkribiert, für die ein Gen auf dem Phagen-Genom kodiert (GEIDUSCHEK ET AL., 1988). Ein zu einer RNA-Polymerase homologes Protein konnte auf pKAP298 nicht entdeckt werden. Möglicherweise sind für die Expression der beiden Gene auch bestimmte Transkriptions-Regulatoren/-Aktivatoren notwendig (KLAUS ET AL., 1992). Für Regulatoren/Aktivatoren könnten auf pKAP298 die ORFs 39 und 44 kodieren, die in keiner Detektion Transkriptions-Aktivität aufwiesen (siehe 4.1.1.1). Eine weitere Möglichkeit für eine Inaktivierung wäre, dass die ORFs 24 und 25 mit den ORFs 41 bis 43 ein Gen-Cluster bildeten (siehe 3.4.2.8, Cluster 4 und 5) und früher polycistronisch, unter Kontrolle des gleichen Regulationsmechanismus, transkribiert wurden,

bevor das Cluster durch die Insertion des Transposons Tn4504 unterbrochen und damit die Transkription unterbunden wurde. Auf dem putativen Cluster 5 (ORFs 41 bis 43) lässt sich noch Transkriptions-Aktivität nachweisen.

Ein hypothetisches Protein, das (schwache) Ähnlichkeit zu einem hypothetischen Phagen-Schwanz-Protein aufweist, konnte im ORF 42 identifiziert werden:

```
gnl|CDD|14402, COG5280, COG5280, Phage-related minor tail protein [Function unknown]
CD-Length = 634 residues, only 32.3% aligned
Score = 45.8 bits (108), Expect = 1e-05
```

Der ORF könnte mitunter für die mit den R-Körpern assoziierten helikalen Proteinstrukturen kodieren (PREER ET AL., 1974).

Die ORFs 49, 50 und 54 scheinen auf dem hypothetischen ursprünglichen Phagen-Genom direkt nebeneinander gelegen zu haben (bzw. sie waren eine zusammenhängende Sequenz), bevor sie, durch die Insertion des Transposons Tn4501 (QUACKENBUSH ET AL., 1983A, 1986A,B) und einer weiteren Sequenz, von einander getrennt wurden. Alle 3 ORFs weisen Ähnlichkeit zu einem RecE-Protein der Prophage VT1-Sakai aus *Escherichia coli* Stamm O157:H7 auf (YOKOYAMA ET AL., 2000):

ORF 49:

```
gi|13361399|dbj|BAB35357.1| exonuclease VIII RecE [Escherichia coli O157:H7]
Length = 866
Score = 34.3 bits (77), Expect = 0.46
Identities = 22/58 (37%), Positives = 33/58 (56%), Gaps = 4/58 (6%)
```

ORF 50:

```
Score = 55.5 bits (132), Expect = 5e-07
Identities = 53/201 (26%), Positives = 96/201 (47%), Gaps = 19/201 (9%)
```

ORF 54:

```
Score = 61.2 bits (147), Expect = 5e-09
Identities = 34/83 (40%), Positives = 43/83 (51%), Gaps = 3/83 (3%)
```

Der ORF 50 stimmt zu 31 % (49 % positive AS) mit einem hypothetischen Protein aus *Legionella pneumophila* überein. In der Annotation diese Proteins (LUNEBERG ET AL., 2001) wird auf eine Homologie zu dem RecE-Protein aus *E. coli* verwiesen, welches auf der im Genom von *E. coli* integrierten, defekten Prophage Rac lokalisiert ist (FOUTS ET AL., 1983). Alle 3 ORFs besitzen Ähnlichkeiten zu einem hypothetischen Protein aus der Lambdoid Phage HB-4 (KIM ET AL., 1999). In der Annotation dieses Proteins weisen die Autoren wiederum auf Homologien zu dem RecE-Protein aus *E. coli* hin. ORF 50 und 54 ähneln

einem hypothetischen Protein der Mycobacteriophage Che9c (PEDULLA ET AL., 2003). Man kann durch den relativ hohen BLAST-"Score" (und das obwohl die ORFs teilweise wesentlich kleiner waren als die gefunden ähnlichen Sequenzen) davon ausgehen, dass diese Sequenzen ihren Ursprung auf einem Phagen-Genom hatten. Das ähnliche Enzym RecE aus diversen Phagen spielt eine Rolle bei der Phagen-Rekombination (SHIRAIISHI ET AL., 2002, ZHANG ET AL., 2003). Da die putativen Enzyme aus den ORFs 49,50 und 54 als Fragmente wohl nicht funktionell sind, ist es wahrscheinlich, dass der ursprüngliche Bakteriophage, aus dem pKAP298 evolvierte, die Fähigkeit zur Phagen-kodierten Rekombination verloren hat. Die (homologe, site-spezifische oder illegitime) Rekombination des Plasmids könnte allerdings weiterhin stattfinden, möglicherweise katalysiert durch Enzyme des Wirts *Caedibacter*. Falls das putative RecE-Protein an der Integration des Phagen in das Wirtsgenom beteiligt war (z.B. durch Induktion von DNA-Doppelstrangbrüchen, ZHANG ET AL., 2003), könnte der Ur-Phage die Fähigkeit zur Integration in das Wirtsgenom verloren haben. Das Plasmid pKAP298 könnte möglicherweise trotzdem, zumindest partiell, in das *Caedibacter*-Genom integriert werden. Auf dem Plasmid und auf dem Wirtsgenom sind die gleichen Transposons lokalisiert. Diese Transposons könnten als Ausgangspunkte für eine homologe Rekombination zwischen dem Plasmid und dem Wirts-Genom dienen. Ob dies bereits geschehen ist, oder wie hoch die Wahrscheinlichkeit für ein solches Ereignis ist, müsste noch näher erforscht werden. Aus dem *Caedibacter taeniospiralis*-Genom sind bisher allerdings keine Sequenzen bekannt.

Die Suche nach konservierten Protein-Domänen ergab für ORF 47 eine große Übereinstimmung mit einer möglichen reversen Transkriptase:

```
gnl|CDD|16610, pfam00078, rvt, Reverse transcriptase (RNA-dependent DNA polymerase). A reverse transcriptase gene is usually indicative of a mobile element such as a retrotransposon or retrovirus. Reverse transcriptases occur in a variety of mobile elements, including retrotransposons, retroviruses, group II introns, bacterial msDNAs, hepadnaviruses, and caulimoviruses.  
CD-Length = 200 residues, 93.0% aligned  
Score = 105 bits (264), Expect = 7e-24
```

Die größte Sequenz-Ähnlichkeit besteht zu einem putativen Protein aus *Bradyrhizobium japonicum*, das KANEKO ET AL. (2002) als wahrscheinliche reverse Transkriptase/Maturase klassifizierten:

```
gi|27377026|ref|NP_768555.1| b1r1915 [Bradyrhizobium japonicum]  
Length = 491  
Score = 270 bits (690), Expect = 2e-71  
Identities = 135/323 (41%), Positives = 202/323 (62%), Gaps = 1/323 (0%)
```

Weitere Sequenz-Homologien bestehen zu einem putativen Protein aus *Sinorhizobium meliloti* (beschrieben von GALIBERT ET AL., 2001):

```
gi|16263485|ref|NP_436278.1| Putative reverse transcriptase [Sinorhizobium meliloti]
Length = 505
Score = 151 bits (381), Expect = 2e-35
Identities = 99/283 (34%), Positives = 145/283 (51%), Gaps = 53/283 (18%)
```

Dieser Befund könnte dafür sprechen, dass es sich bei dem potentiellen Bakteriophagen, aus dem pKAP298 hervorging, um einen Phagen mit einzelsträngiger RNA als Genom gehandelt hat. Phagen, deren Genom aus RNA besteht, benötigen eine reverse Transkriptase, um das Genom zunächst in doppelsträngige DNA umzuschreiben (LEWIN, 2002). Andererseits weist die Aminosäure-Sequenz eine große Übereinstimmung zu dem Protein RetA auf, das auf dem broad-host-range Plasmid IncL/M R471a (aus *Serratia marcescens*) lokalisiert ist (KULAEVA ET AL., 1998):

```
gi|2688957|gb|AAC82519.1| RetA [IncL/M plasmid R471a]
Length = 495
Score = 267 bits (683), Expect = 1e-70
Identities = 136/323 (42%), Positives = 199/323 (61%), Gaps = 1/323 (0%)
```

Das *retA*-Gen wurde von DAI ET AL. (2002) als Gruppe-II-Intron charakterisiert. Auch die Ähnlichkeit des ORFs 47 zu Proteinen anderer Gruppe-II-Introns spricht gegen den ORF 47 als reverse Transkriptase eines RNA-Phagen. Mit einem putativen Protein aus *Pseudomonas putida* Stamm KT2 440 stimmt ORF 47 zu 46 % überein (63 % positive AS). Dieses Protein wurde von NELSON ET AL. (2002) als putative Maturase eines Gruppe-II-Introns klassifiziert.

```
gi|24986172|gb|AAN70021.1|AE016640_9 group II intron-encoding maturase, putative [Pseudomonas putida
KT2440]
Length = 373
Score = 229 bits (585), Expect = 4e-59
Identities = 117/253 (46%), Positives = 161/253 (63%), Gaps = 1/253 (0%)
```

Ein Protein des Plasmids pNL1 aus *Novosphingobium aromaticivorans* stimmt mit ORF 47 zu 41 % überein (59 % positive AS). Dieses Protein wurde von ROMINE ET AL. (1999) als putative Maturase eines Gruppe-II-Introns klassifiziert:

```
gi|11362250|pir|T31160 maturase-related protein (group II intron) - Sphingomonas aromaticivorans plasmid pNL1
Length = 633
Score = 157 bits (396), Expect = 3e-37
Identities = 84/203 (41%), Positives = 121/203 (59%), Gaps = 3/203 (1%)
```

Der ORF 48 besitzt ebenfalls Ähnlichkeiten zu den oben beschriebenen Proteinen. Er ist daher wahrscheinlich Teil des putativen Gruppe-II-Introns. Gruppe-II-Introns sind mobile Elemente, deren Mobilität nicht, wie bei Transposons, mit doppelsträngiger DNA beginnt, sondern als einzelsträngige RNA. Die reverse Transkriptase/Maturase sorgt für das Umschreiben der RNA in DNA und oftmals auch für die Integration an einer bestimmten Zielstelle im Genom (KLEIN ET AL., 2002, LEHMANN ET AL., 2003). Ob das putative Gruppe-II-Intron noch die Fähigkeit zur Transposition besitzt, konnte in dieser Arbeit nicht geklärt

werden. Es müsste untersucht werden, ob sich das putative Gruppe-II-Intron auch im Genom von *Caedibacter taeniospiralis* wiederfindet (obwohl Gruppe-II-Introns vorwiegend zwischen mobilen Elementen transponieren, KLEIN ET AL., 2002). Auf dem pKAP298-Fragment, auf dem das putative Gruppe-II-Intron lokalisiert ist, konnte Gen-Aktivität nachgewiesen werden (siehe 4.1.1.1), ob es allerdings die zum Gruppe-II-Intron gehörenden ORFs sind, die transkribiert wurden, blieb unklar.

Für den ORF 12 fand BLAST große Homologien zu konservierten Proteindomänen aus der Phagen-Integrasen Familie:

```
gnl|CDD|7681, pfam00589, Phage_integrase, Phage integrase family.
CD-Length = 175 residues, 97.7% aligned
Score = 93.5 bits (232), Expect = 2e-20
gnl|CDD|8420, pfam02899, Phage_integr_N, Phage integrase, N-terminal SAM-like domain.
CD-Length = 87 residues, 92.0% aligned
Score = 66.5 bits (162), Expect = 3e-12
```

Auch in der Aminosäure-Sequenz ließen sich Ähnlichkeiten zu einer putativen Phagen-Integrase aus *Bacillus anthracis* (READ ET AL., 2002) feststellen:

```
gi|21401814|ref|NP_657799.1| Phage_integrase, Phage integrase family [Bacillus anthracis A2012]
Length = 297
Score = 113 bits (282), Expect = 4e-24
Identities = 86/285 (30%), Positives = 138/285 (48%), Gaps = 18/285 (6%)
```

Phagen-Integrasen sind an dem Einbau des Phagen-Genoms in das Wirtsgenom beteiligt, so dass der temperente Phage als Prophage in die lysogene Phase gelangt (KLAUS ET AL., 1992). Bei dem ORF 12 ist es allerdings unwahrscheinlich, dass er wirklich vom originären Phagen kommt, da er auf dem Transposon Tn4503 lokalisiert ist. Er stammt also höchstwahrscheinlich aus dem *Caedibacter taeniospiralis* Genom (siehe 4.2.4). Des weiteren weist die Aminosäuresequenz starke Ähnlichkeit zu einer site-spezifischen Rekombinase vom Transposon Tn5401 aus *Bacillus thuringiensis* auf (BAUM, 1994):

```
gi|2127287|pir|I40587 site-specific recombinase (Tn5401) - Bacillus thuringiensis
Length = 306
Score = 118 bits (296), Expect = 9e-26
Identities = 78/265 (29%), Positives = 131/265 (49%), Gaps = 11/265 (4%)
```

Als relativ sicher vom ursprünglichen Phagen stammend können nur die ORFs 24,25,42,49,50 und 54 angesehen werden. Alle anderen in 4.2.2 beschriebenen ORFs mit Ähnlichkeiten zu Phagen-Proteinen kommen wahrscheinlich nicht vom originären Phagen, sondern sind durch Transpositions-Ereignisse auf pKAP298 gelangt. Auf dem Plasmid konnten keine ORFs gefunden werden, die Ähnlichkeiten zu bekannten Proteinen aufweisen, die an der Phagen-induzierten Lyse der infizierten Zellen beteiligt sind - vielleicht sind diese Gene verlorengegangen (siehe 4.2.6), oder aber ORFs ohne Sequenzhomologien kodieren für solche Proteine.

4.2.3 Taxonomie und Genom des putativen Bakteriophagen, aus dem pKAP298 evolvierte

In der nachfolgenden Tabelle sind diejenigen Bakteriophagen näher beschrieben, mit denen die als sicher vom putativen pKAP-Phagen stammenden ORFs (24,25,49,50 und 54, siehe 4.2.2) Übereinstimmungen aufweisen. Aus diesen Daten lässt sich eine hypothetische Taxonomie des Ur-Phagen erstellen.

Name	Taxonomie	Genom-Typ	Genomgröße	beschrieben von
Bacteriophage 933W	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae; Lambda-like viruses.	dsDNA	61670 Bp	PLUNKETT ET AL., 1999
Enterobacteria phage P22	Viruses; dsDNA Viruses; no RNA stage; Caudovirales; Podoviridae; P22-like viruses	dsDNA	41724 Bp	VANDER BYL ET AL., 2000
Lambdoid phage HB-4	Viruses, unclassified bacteriophages	?	?	KIM ET AL., 1999
Mycobacteriophage Che9c	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.	dsDNA	57050 Bp	PEDULLA ET AL., 2003
<i>Pseudomonas aeruginosa</i> phage PaP3	Viruses; dsDNA Viruses; no RNA stage; Caudovirales; Podoviridae; T7-like Viruses	dsDNA	44249 Bp	HU ET AL., 2002
<i>Salmonella typhimurium</i> bacteriophage ST64T	Viruses; dsDNA Viruses; no RNA stage; Caudovirales; Podoviridae; unclassified Podoviridae	dsDNA	40679 Bp	MMOLAWA ET AL., 2001
<i>Sinorhizobium meliloti</i> phage PBC5	Viruses, unclassified bacteriophages	?	57416 Bp	SCHULMEISTER ET AL., 2001
<i>Staphylococcus aureus</i> phage phi 11	Viruses, unclassified bacteriophages	?	43604 Bp	IANDOLO ET AL., 2002

Tab. 4.2.3.1: Bakteriophagen, zu denen pKAP298 Sequenz-Homologien besitzt

Es werden jeweils der Name, die Taxonomie, der Genomtyp und die Genom-Größe dargestellt

Lässt man die unklassifizierten Bakteriophagen außer Betracht, so kann man feststellen, dass alle Bakteriophagen, die Sequenz-Ähnlichkeiten mit pKAP298 aufweisen, aus der Ordnung

der *Caudovirales* stammen. Gemeinsame Merkmale dieser Ordnung sind, dass die zu ihr gehörenden Bakteriophagen Schwänze besitzen und ihr Genom aus linearer doppelsträngiger DNA besteht (MANILOFF UND ACKERMANN., 1998). Die in *Caedibacter taeniospiralis* mit den R-Körpern assoziierten langen helikalen Proteinstrukturen könnten daher möglicherweise wirklich als Produkte einer Phagenschwanz-Synthese (PREER ET AL., 1974) gedeutet werden. MANILOFF UND ACKERMANN (1998) unterteilen diese Ordnung in 3 Familien, die *Myoviridae* (kontraktile Schwänze), die *Podoviridae* (kurze, nicht kontraktile Schwänze) und die *Siphoviridae* (lange, nicht kontraktile Schwänze). Vergleicht man die durchschnittliche Genomgröße der zu pKAP298 ähnlichen Bakteriophagen (aus der Tabelle 4.2.3.1), so ergibt sich für die zur Familie der *Siphoviridae* gehörenden Phagen eine Größe von ca. 59 kB. Die Vertreter aus der Familie der *Podoviridae* besitzen eine durchschnittliche Genomgröße von ca. 42 kB. Das Plasmid pKAP298 ist, abzüglich der aus *Caedibacter taeniospiralis* stammenden Transposons (QUACKENBUSH ET AL., 1986A,B, HERUTH ET AL., 1996 und diese Arbeit), ca. 38,5 kB groß. Die Genomgröße würde also eher für eine Einordnung bei den *Podoviridae* sprechen. Es ist allerdings noch nicht bekannt, ob die putative pKAP298-Phage während ihrer Evolution Genom-Fragmente verloren (z.B. durch ein Rekombinations-Ereignis) oder (außer den bereits identifizierten Transposons und des Gruppe-II-Introns) neue DNA integriert hat. Daher ist die Klassifizierung über das Genom in diesem Fall nicht sehr aussagekräftig.

4.2.4 Transposons und das Gruppe-II-Intron

Die auf dem Plasmid pKAP298 identifizierten (und zu den Transposons von pKAP116 und pKAP47 homologen) Transposons Tn4501, Tn4501/2 und Tn4503 stammen wahrscheinlich aus dem Genom von *Caedibacter taeniospiralis*. Für ihre Pendanten auf pKAP169 und pKAP30 konnte dies im Labor nachgewiesen werden, da auf diesen Plasmiden spontane Insertionen von Tn4501 und Tn4503 beobachtet werden konnten (DILTS ET AL., 1986A,B; QUACKENBUSH ET AL., 1986B). Überdies ergaben DNA-DNA Hybridisierungen der Transposon-DNA mit genomischer DNA große Übereinstimmungen. Bisher wurde nur das Transposon Tn4501 sequenziert (HERUTH ET AL., 1996). Für die Aufklärung der Herkunft des Plasmids pKAP298 können diese Transposons nicht dienen, da sie alle sekundär in *Caedibacter taeniospiralis* auf das Plasmid gelangt sind.

Das Transposon Tn4501/2 sollte starke Ähnlichkeiten zum Transposon Tn4501 aufweisen (QUACKENBUSH ET AL., 1986A,B). Dies konnte in dieser Arbeit durch den Sequenz-Vergleich beider Transposasen bestätigt werden:

```

gi|1438743|gb|U61289.1|CTU61289 Caedibacter taeniospiralis plasmid pKAP transposon TN4501 transposase
gene, complete cds
Length = 1350

Score = 2425 bits (1261), Expect = 0.0
Identities = 1343/1359 (98%), Gaps = 12/1359 (0%)
Strand = Plus / Plus

```

Auch die große Übereinstimmung der invertierten Wiederholungs-Sequenzen spricht dafür, dass Tn4501 und Tn4501/2 homolog sind:

```

10196 tagctacctctcaataacagtttaaatacacatacc 10161 Tn4501
      | | ||| ||| || |||||
44301 ttat-attactctatatgagcttaaatacacatacc 44268 Tn4501/2

```

Dies entspricht einer Übereinstimmung von ca. 70 %.

Die auf dem Transposon TN4501/2 entdeckte putative Transposase (ORFs 9,10 und 11) wurde durch Deletionen, Punktmutationen und/oder Insertionen mehrmals unterbrochen, so dass dieses Transposon wahrscheinlich nicht mehr zur Transposition fähig ist. Auch zu dem von HERUTH ET AL. (1996) sequenzierten Transposon Tn4502 (nicht auf pKAP298 zu finden) besteht eine sehr große Übereinstimmung:

```

gi|1407812|gb|U60645.1|CTU60645 Caedibacter taeniospiralis plasmid pKAP transposon TN4502 transposase
gene, complete cds
Length = 1887

Score = 2517 bits (1309), Expect = 0.0
Identities = 1405/1423 (98%), Gaps = 14/1423 (0%)
Strand = Plus / Plus

```

Im Rahmen dieser Arbeit konnten keine weiteren, auf Tn4501 und Tn4501/2 lokalisierten, putativen Gene entdeckt werden. Somit sollten Tn4501 und Tn4501/2 als "Insertions-Sequenzen" bezeichnet werden. Die Transposition der Transposons verläuft wahrscheinlich nicht replikativ.

Das Transposon Tn4503 (auf dem Tn4501/2 lokalisiert ist) konnte in dieser Arbeit näher beschrieben werden. Es ist 7724 Bp groß, hat eine direkt wiederholte Insertions-Sequenz von 6 Basen und eine invertierte Wiederholungs-Sequenz von 36 Basen, von denen 32 genau übereinstimmen. Ein auf dem Transposon Tn4503 lokalisiertes putatives Gen (ORF 12) weist Homologien zu einer site-spezifischen Rekombinase von dem Transposon Tn5401 aus *Bacillus thuringiensis* auf (BAUM, 1994, siehe auch 4.2.2). Als auf einem Transposon lokalisierte site-spezifische Rekombinase ist ORF 12 möglicherweise an der Transposition des Transposons beteiligt. Es könnte die Aufgabe haben, (als Resolvase) das bei der Transposition entstehende Kointegrat aus Donor- und Rezipienten-DNA durch eine site-

spezifische Rekombination zu trennen. Dies deutet auf eine replikative Transposition des Transposons Tn4503 hin. Für das Transposon Tn5401 aus *Bacillus thuringiensis* (wo eine zu ORF 12 ähnliche Rekombinase lokalisiert ist, s.o.) konnte eine replikative Transposition experimentell bestätigt werden (BAUM, 1994). In dem von BAUM beschriebenen Transposon Tn5401 ist das Resolvase-Gen (*tnpI*) direkt neben dem Transposase-Gen (*tnpA*) lokalisiert. Auf Tn4503 ist diese Anordnung durch die Insertion des Transposons Tn4501/2 zerstört. Laut BAUM ET AL. (1999) hat das Protein TnpI eine Repressorfunktion für die Transkription von *tnpA*. Möglicherweise wirkt ORF 12 ebenfalls als Repressor und hemmt die Bildung der Transposase (ORF 8), und damit die Transpositionshäufigkeit. Die Transkriptions-Aktivitäts-Untersuchungen ergaben für beide ORFs (8 und 12) keine Aktivität (bei den bei der RNA-Isolation herrschenden Bedingungen). Der ORF 7 hat seine größte Übereinstimmung mit einer site-spezifischen Rekombinase aus einem Quecksilber-Resistenz vermittelnden Transposon, das in *E. coli* subkloniert wurde (YURIEVA ET AL., 1997, 35 % identische AS und 57 % positive AS). Ob das putative Protein funktionell ist konnte nicht geklärt werden. ORF 8 ähnelt einer von WANG ET AL. (1989) entdeckten Transposase aus *Bacillus cereus* (22 % identische AS, 41 % positive AS). Zudem besitzt ORF8 Ähnlichkeit zu einer von HUANG ET AL. (1999) beschriebenen Transposase, die auf einem Quecksilber-Resistenz vermittelnden Transposon aus *Bacillus megaterium* liegt (22 % identische AS und 41 % positive AS). Quecksilber-Resistenz-Gene (oder Gene zur Vermittlung anderer Resistenzen) konnten auf dem Transposon Tn4503 nicht identifiziert werden. Es befinden sich jedoch noch zwei weitere putative Gene auf Tn4503, die allerdings keine Ähnlichkeiten zu schon bekannten Protein-Sequenzen aufweisen (ORF 13 und ORF 14). Daher kann hier auch keine Aussage zu deren wahrscheinlicher Funktion gemacht werden. Einer der beiden ORFs (oder möglicherweise auch beide) ergab(en) bei der Transkriptions-Aktivitäts-Analyse positive Signale (siehe 4.1.2). Welches der 3 putativen Transpositions-Enzyme (ORFs 7, 8, 12) letztlich für die Mobilität des Transposons Tn4503 verantwortlich ist, oder ob mehrere Enzyme zusammen wirken, konnte durch diese Arbeit nicht sicher aufgeklärt werden. Tatsache ist jedoch, dass Tn4503 noch die Fähigkeit zur Transposition besitzt, wie spontane Neu-Insertionen in die Plasmide pKAP169 und pKAP30 bewiesen haben (DILTS ET AL., 1986A,B, QUACKENBUSH ET AL., 1986B).

Das neu entdeckte putative Transposon Tn4504 (3.4.2.7.4) scheint ursprünglich nicht aus *Caedibacter taeniospiralis* zu stammen. Eine Region mit gleichen Restriktions-Schnittstellen ist auf allen bisher beschriebenen pKAP-Plasmiden zu finden (QUACKENBUSH ET AL., 1983A). Die relativ geringe Ähnlichkeit der direkten Wiederholungs-Sequenzen spricht dafür, dass das putative Transpositions-Ereignis evolutionär längere Zeit zurückliegt, die sehr große Übereinstimmung der invertierten Wiederholungen (95% bei einer Länge von 269 Basen)

spricht dagegen. Der ORF 32 könnte, laut des Suchergebnisses nach konservierten Protein-Domänen ([pfam04754.2](#), [Transposase 31](#)), für eine Transposase kodieren. In seiner Aminosäure-Sequenz weist er ebenfalls Homologien zu einer von MCCLELLAND ET AL. (2001) beschriebenen, putativen Transposase aus *Salmonella typhimurium* LT2 auf (29 % identische AS, 48 % positive AS). Auf dem Transposon wurde ein weiteres hypothetisches Gen mit Sequenz-Homologien entdeckt, das möglicherweise für ein DNA bindendes Protein kodiert (ORF 39). Es könnte eine Rolle in der Replikation des Plasmids pKAP298 oder bei der Gen-Regulation spielen (siehe 4.1.1.1).

Ein Gruppe-II-Intron konnte bisher auf den pKAP-Plasmiden nicht identifiziert werden. Das neu entdeckte putative Gruppe-II-Intron (~39601 Bp - 40935 Bp, ORFs 47 und 48, siehe 4.2.2) stammt höchstwahrscheinlich nicht aus dem Genom von *Caedibacter taeniospiralis*, da es (laut Restriktionskarte von QUACKENBUSH ET AL., 1983A) auf allen bekannten pKAP-Plasmiden vorhanden ist und spontane Neu-Insertionen eines solchen Introns auf pKAP bisher nicht beobachtet werden konnten. Aus der Sequenz-Ähnlichkeit der zum Gruppe-II-Intron gehörenden ORFs lassen sich Rückschlüsse auf mögliche vorherige Wirte der pKAP-Plasmid-Familie ziehen. Die größte Sequenz-Ähnlichkeit der beider ORFs bestehen zu putativen Proteinen aus *Bradyrhizobium japonicum* (KANEKO ET AL., 2002, siehe 4.2.2).

4.2.5 Verwandtschaft von pKAP298 zu anderen pKAP-Plasmiden

Die Beziehung von pKAP298 zu anderen pKAP-Plasmiden lässt sich aufgrund des Vergleichs der Aminosäuresequenz bereits bekannter Gene dieser Plasmide untersuchen. Von pKAP47 und pKAP116 ist die Sequenz der R-Körper-Gene (HERUTH ET AL., 1994) und der Transposase des Transposons Tn4501 (HERUTH ET AL., 1996) bekannt.

Das *rebB*-Gen von pKAP298 (ORF 2) stimmt genau mit seinen Pendanten auf pKAP47 und pKAP116 überein, ebenso *rebC* (ORF 63). Das *rebA*-Gen (ORF 3) unterscheidet sich in der Aminosäure 53 von *rebA* aus pKAP47 und pKAP116. Statt Alanin ist in der pKAP298-Sequenz Glycin zu finden. Interessanterweise zeigt sich der gleiche Unterschied auch in der Sequenz von *rebA* aus *Caedibacter caryophilus* (WEISEL, 2001, Diplomarbeit). Auch dort findet sich im *rebA*-Gen Glycin statt Alanin. In *C. caryophilus* sind die R-Körper Gene nicht auf einem Plasmid, sondern wahrscheinlich auf einem Bakteriophagen lokalisiert (SCHMIDT ET AL., 1987). Sollte dieser Unterschied kein Artefakt durch die Sequenzierung sein, könnte man daraus schließen, dass pKAP298 größere Übereinstimmung mit dem postulierten ursprünglichen Bakteriophagen besitzt als pKAP47 und pKAP116 (sofern pKAP298 und der *C. caryophilus*-Phage homolog sind). Dies wäre möglicherweise ein Hinweis darauf, dass sich

pKAP47 und pKAP116 aus pKAP298 entwickelt haben. Das *rebD*-Gen aus pKAP298 unterscheidet sich (translatiert) in den Aminosäuren 19 bis 24 von seinen Homologen aus pKAP47 und pKAP116. Der Grund dafür sind 2 Insertionen in der Nucleotidsequenz (auf pKAP47 die Stellen 54 und 67), die in diesem Bereich von 54 Bp bis 72 das Leseraster verschieben. Diese Verschiebung wird durch eine weitere Insertion an Stelle 73 wieder ausgeglichen, so dass die restliche Aminosäuresequenz von *rebD* auf allen pKAP-Plasmiden übereinstimmt. Vergleicht man dagegen wiederum das *rebD*-Gen aus pKAP298 mit seinem Pendant aus *C. caryophilus* (WEISEL, 2001, Diplomarbeit), so sind beide völlig identisch - hier fehlen die Insertionen aus pKAP47 und pKAP116. Dies könnte die Hypothese stützen, dass pKAP298 dem ursprünglichen Bakteriophagen am nächsten kommt, und dass sich (zumindest) pKAP47 und pKAP116 aus pKAP298 entwickelt haben. Dass das Ergebnis ein Artefakt durch Fehler bei der Sequenzierung ist, kann als unwahrscheinlich angesehen werden, da die Sequenzen von *rebD* aus pKAP47 und pKAP116 genau übereinstimmen und 3 identische Sequenzierungs-Fehler innerhalb einer 19 Bp Sequenz relativ unwahrscheinlich sind.

Ein Sequenzvergleich der Transposase des Transposons Tn4501, das auf allen pKAP-Plasmiden zu finden ist (QUACKENBUSH ET AL., 1983A) und aus dem Genom von *Caedibacter taeniospiralis* stammt (DILTS ET AL., 1986A,B, QUACKENBUSH ET AL., 1986B), zeigt, dass die Basen- und Aminosäure-Sequenz des Proteins auf allen bisher untersuchten pKAP-Plasmiden identisch sind. Dies könnte damit begründet werden, dass sich das Transposon evolutionär noch nicht so lange auf den pKAP-Plasmiden befindet wie die R-Körper-Gene.

4.2.6 Zusammenfassung und Ausblick

Die Hypothese, dass das Plasmid pKAP298 von einem Bakteriophagen abstammt (PREER ET AL., 1974), konnte durch diese Arbeit gestützt werden. Dafür sprechen die identifizierten ORFs mit Sequenz-Ähnlichkeiten zu Phagen-Genen (siehe 4.2.2). Wahrscheinlich evolvierte pKAP298 aus einem Bakteriophagen mit doppelsträngiger DNA (siehe 4.2.3). Ob der putative Ursprungs-Phage als Plasmid (wie der *E. coli* Prophage P1, COOPER ET AL., 2002) oder als lineares DNA-Molekül (wie der *E. coli* Bakteriophage N15, RAVIN ET AL., 2000) existierte und sich erst sekundär zum Plasmid wandelte, ist bisher nicht eindeutig geklärt. Der hypothetische Phage wurde wahrscheinlich nicht in das Genom seines Wirtes integriert, sondern autonom von dessen Genom repliziert. Dafür spricht die Entdeckung von putativen Replikations-Proteinen auf pKAP298: ORF 39 als möglicher Integration Host Factor und ORF 43 als putatives ParA-Protein. ParA Homologe sind auf vielen Plasmiden mit niedriger

Kopienzahl (BIGNELL ET AL., 2001), sowie auf autonom replizierenden Phagen zu finden (DAVEY ET AL., 1994). Als Beispiele seien hier wiederum die Phagen P1 und N15 genannt. ParA ist als ATPase bei der aktiven Segregation von replizierter DNA auf die Tochterzellen während der Zellteilung beteiligt (MOTALLEBI-VESHAREH ET AL., 1990, siehe 4.3).

Die für die R-Körper kodierenden Gene von pKAP298 und *Caedibacter caryophilus* sind identisch (siehe 4.2.5, WEISEL, 2001). In *C. caryophilus* konnte durch CsCl-Gradienten-Zentrifugation Bakteriophagen-DNA nachgewiesen werden (SCHMIDT ET AL., 1987). Wenn die R-Körper kodierenden Gene bei *Caedibacter caryophilus* wirklich auf diesem Bakteriophagen lokalisiert sind, spricht dies für die Hypothese, dass der Vorläufer von pKAP298 linear war, denn SCHMIDT ET AL. (1987) und WEISEL (2001) konnten keine Plasmid-DNA nachweisen. Der Bakteriophage aus *C. caryophilus* müsste sequenziert werden, um herauszufinden in wie weit er mit der Gesamtsequenz von pKAP298 übereinstimmt. Als forward- und reverse-Sequenzier-Primer könnten Oligonucleotide aus der R-Körper kodierenden Region dienen. Sollten sich weitere große Ähnlichkeiten ergeben, käme der Bakteriophage als Vorgänger von pKAP298 in Betracht. In mit *C. caryophilus* infizierten Paramecien konnten SCHMIDT ET AL. (1987) im Makronucleus freie R-Körper entdecken und vermuteten, dass diese durch die phagen-vermittelte Lyse der *Caedibacter* freigesetzt wurden. Daraus kann gefolgert werden, dass auf dem Bakteriophagen aus *C. caryophilus* aktive, für die Zell-Lyse kodierende, Gene lokalisiert sind. Bei *Caedibacter taeniospiralis* konnte bisher keine Zell-Lyse festgestellt werden (ausgenommen, dass die Bright-Form sensitiver auf Tenside reagieren, POND ET AL., 1989), und auch in der pKAP298-Sequenz konnten keine ORFs mit Homologien zu bekannten Lyse-Genen entdeckt werden. Falls der Bakteriophage aus *C. caryophilus* wirklich homolog zu pKAP298 sein sollte, so könnten diese Gene als Folge des Ringschlusses des linearen Bakteriophagen (vielleicht aufgrund eines wirts- oder phagen-induzierten Rekombinations-Ereignisses) verloren gegangen sein. Die entsprechenden Lyse-Gene müssten daher auf einem der Enden des *C. caryophilus*-Phagen lokalisiert sein - dies wäre zu überprüfen.

Das Plasmid pKAP298 könnte also durch Rekombination vom Bakteriophagen zu einem "broad host range"-Plasmid evolviert sein. Für pKAP als "broad-host-range"-Plasmid spricht auch die Tatsache, dass manche als *Caedibacter taeniospiralis* klassifizierte Bakterienstämme (A30 und 51) nur 41 % DNA-Homologie zueinander aufweisen (QUACKENBUSH ET AL., 1977, 1978), während für andere *Caedibacter taeniospiralis*-Stämme eine Übereinstimmung von immerhin 80 % nachgewiesen werden konnte. Der verbindende Faktor scheint das Plasmid pKAP zu sein. Einige identifizierte ORFs mit potentieller Phagen-Herkunft und die des putativen Gruppe-II-Introns (ORFs 24, 25, 47, 48) weisen Ähnlichkeiten zu Proteinen auf, die in der Ordnung *Rhizobiales* identifiziert und beschrieben worden sind (siehe 4.2.2)

(KANEKO ET AL., 2002 und GALIBERT ET AL., 2001). Vielleicht war ein zu dieser Ordnung gehörender Organismus einmal Wirt des putativen Ur-Bakteriophagen während seiner Evolution.

4.3 Versuch der Identifizierung des putativen Toxins

Die Suche nach einem möglichen Toxin gleicht einem Indizienprozess. Geht man davon aus, dass das Toxin vom Plasmid kodiert wird (POND ET AL., 1989) und dass die Expression mit Beginn der R-Körper-Synthese einsetzt (da "Nonbright"-*Caedibacter* nicht toxisch sind, DILTS ET AL., 1986), muss das Toxin auf einem pKAP298 Fragment liegen, das bei meinen Untersuchungen Transkriptions-Aktivität aufwies. Bei allen detektierten RNA/cDNA-Sonden konnte Aktivität auf den R-Körper kodierenden Regionen des Plasmids nachgewiesen werden, daher sollte auch das Toxin produziert worden sein. GIBSON UND WELLS (unveröffentlicht) fanden durch Protein-Chromatographie an *P. tetraurelia* 51 heraus, dass ein 20-25 kDa Protein toxisch auf sensitive Paramecien wirkte und diese innerhalb von 3 Tagen starben. Nimmt man nun die Kriterien, dass das Toxin auf einem transkriptionsaktiven Fragment liegen muss und eine Masse zwischen 20 und 25 kDa hat, bleiben genau 3 als möglicherweise kodierend identifizierte ORFs übrig: ORF 7 mit ca. 22,4 kDa, ORF 9 mit ca. 20,3 kDa und ORF 43 mit ca. 24,6 kDa. Die errechnete Masse der putativen Proteine ist möglicherweise nicht exakt, da der jeweilige von GeneMark.hmm vorausgesagte Transkriptions-Startpunkt nicht sicher ist. LUKASHIN UND BORODOVSKI (1998) geben aber eine durchschnittliche Wahrscheinlichkeit von 78 % an, dass ein von GeneMark.hmm identifizierter kodierender ORF an dem angegebenen Startkodon transkribiert wird. Ein weiterer, die reale Masse der putativen Proteine verändernder Faktor, könnten post-translationale Modifikationen der Proteine sein. Die ORFs 7 und 9 stammen nicht vom originären pKAP-Plasmid, sondern sind durch ein Transpositions-Ereignis als Transposon aus dem *Caedibacter*-Genom auf das Plasmid gelangt (DILTS ET AL., 1986A,B, QUACKENBUSH ET AL., 1986B). Auch die Tatsache, dass *Caedibacter taeniospiralis*-Stämme, die ein pKAP-Plasmid ohne dieses Transposon (Tn4503) besitzen, ebenfalls toxisch auf sensitive Paramecien wirken (POND ET AL., 1989), spricht gegen die von den ORFs 7 und 9 kodierten Proteine als mögliche Toxine. Es bleibt also noch der ORF 43. Interessanterweise liegt dieser ORF auf einem pKAP298-Fragment (33785 Bp - 36842 Bp), das sich nicht in *E. coli* subklonieren ließ (siehe 3.2.4). Ein 155 Bp kleineres, mit den gleichen Restriktionsenzymen verdautes Fragment (15402 Bp - 18304 Bp, siehe Abb. 3.2.3.1) wurde jedoch häufig in den untersuchten *E. coli* - Klonen gefunden - obwohl definitiv die richtige Bande aus dem Gel ausgeschnitten und zur Klonierung aufbereitet worden war. Dies zeigt, dass zumindest die Methodik funktioniert hatte - aber auf dem Fragment könnten möglicherweise für *E. coli*

toxische Sequenzen sitzen. Eine Analyse der Sequenz ergab keine Palindrome, die durch Formation von Hairpins die teilweise Deletion des Inserts hätten induzieren können (PINDER ET AL., 1998). Allerdings ist auf dem Fragment eine Region mit 12 direkten, 33 Basen langen, Repeats lokalisiert, die durch Formation stabiler Sekundärstrukturen möglicherweise die Replikation des rekombinanten Vektors inhibiert haben könnten - der genaue Mechanismus solcher Blockierungen *in vivo* ist noch unklar (KRASILNIKOVA ET AL., 1998):

Start	End	Score	Size	Count	Identity	Consensus
35072	35467	333	33	12	96.2	cctgtttcttgagacatgtccccggtttgtcc

(ermittelt mit Etandem, EMBOSS, <http://bioweb.pasteur.fr/seqanal/interfaces/etandem.html>)

Es ist für dieses Fragment bisher nicht sicher geklärt, warum es nicht subkloniert werden konnte. Geht man von der Toxin-Theorie aus, könnte das putative, von ORF 43 kodierte, Protein die Subklonierung verhindert haben. Als zur Soj-/ParA-Familie homologes Protein spielt es (als ATPase) wahrscheinlich bei der Aufteilung des Plasmids pKAP298 auf die zwei Tochterzellen bei der Zellteilung eine Rolle (MOTALLEBI-VESHAREH, 1990). Auch bei der Segregation des Genoms bei *E. coli* und *B. subtilis* sind zur Soj-/ParA-Familie gehörende Proteine an der gleichmäßigen Aufteilung des Genoms in die 2 Tochterzellen beteiligt (BIGNELL ET AL., 2002). Ein weiteres Protein, das an diesem Vorgang beteiligt ist, ist ParB. EASTER ET AL. (2002) fanden an *Caulobacter crescentus* durch Klonierungs- und Induzierungs-Versuche heraus, dass, wenn das zu ParA homologe Protein in höherer Konzentration als ParB vorkommt, es zu schweren Zellteilungs- und Chromosomen-Verteilungs-Defekten kommt. Auf dem Plasmid pKAP298 konnte bisher kein zu ParB homologes Protein entdeckt werden. Wenn das putative ParA Protein in *E. coli* exprimiert würde, könnte es aufgrund des Missverhältnisses von (chromosomalem) ParB und (chromosomalem+Plasmid-kodierten) ParA zu Zellteilungs-Defekten kommen. Falls diese hypothetischen Defekte nicht immer auftreten, also nicht völlig letal sind, würde es sich wahrscheinlich zumindest in einer verlangsamten Vermehrung der Bakterien bemerkbar machen. *Caedibacter taeniospiralis* besitzt eine sehr geringe Vermehrungsrate (PREER ET AL., 1972) - vielleicht ist gerade dieses putative Gen dafür verantwortlich, indem es die Zellteilung stört. Möglicherweise ist genau dieses Gen daran schuld, dass *Caedibacter taeniospiralis* nicht außerhalb seines Wirtes *Paramecium tetraurelia* kultiviert werden kann. Dies wäre ein ökologisch besonders interessanter Aspekt. Man könnte das Plasmid pKAP als Parasit ansehen, dessen einziger Zweck seine Vermehrung ist (denn Plasmid-vermittelte Vorteile für *Caedibacter* konnten bisher nicht entdeckt werden) und dabei den Wirt *Caedibacter* schädigt (durch die mögliche Verringerung der Vermehrungsrate und den erhöhten Ressourcenverbrauch). *Caedibacter* kann man ebenfalls als Parasit ansehen, da er sich wohl von *Paramecium*-Stoffwechselprodukten ernährt und seinen Wirt damit schädigt (infizierte

Paramecien sterben unter Hungerbedingungen schneller als nicht-infizierte, ESTEVE, 1978). Gleichzeitig scheint das Plasmid, durch die in *Caedibacter* stattfindende Expression der R-Körper und eines Toxins, dem *Caedibacter*-Wirt *Paramecium* einen Selektions-Vorteil gegenüber nicht infizierten Paramecien zu verschaffen (KUSCH ET AL., 2002). Der Parasit eines Parasits hätte also einen positiven Effekt auf den Wirt seines Wirts, während er seinen direkten Wirt schädigt. Dieses Beziehungsgeflecht sollte auf molekularbiologischer Ebene noch eingehender untersucht werden. In wie weit das hypothetische von ORF 43 kodierte Protein einen toxischen Effekt auf sensitive Paramecien haben könnte, muss noch erforscht werden. Ein interessanter Ansatz ist, dass BLAST für dieses putative Protein Ähnlichkeiten zu einer Membran-assoziierten ATPase fand, die auch bei Eukaryoten eine Rolle bei ATP-abhängigen Ionenpumpen spielt (ZHOU ET AL., 2000):

```
gnl|CDD|16569, cd00550, ArsA_ATPase, Oxyanion-translocating ATPase (ArsA). This ATPase is involved in transport of arsenite, antimonite or other oxyanions across biological membranes in all three kingdoms of life. ArsA contains a highly conserved AAA motif present in the AAA+ ATPase superfamily associated with a variety of cellular activities. To form a functional ATP-driven pump, ArsA interacts with the permease ArsB, which is a channel-forming integral membrane protein. One of the most interesting features of ArsA is the allosteric activation by its transport substrates. A divalent cation, typically Mg2+, is required for its enzymatic activity.
```

```
CD-Length = 294 residues, only 43.5% aligned  
Score = 43.0 bits (101), Expect = 4e-05
```

Die Übereinstimmungen sind zwar nicht groß, es könnte allerdings sein, dass eben dieses hypothetische Protein in *Paramecium* ebenfalls als ATPase - und somit als Triebfeder einer solchen Ionenpumpe - wirkt. Diese Theorie würde auch sehr gut zu der Vermutung passen, dass der Wirkort des putativen Toxins die Zellemembran ist und dort die osmoregulatorischen Fähigkeiten der Paramecien stört (BUTZEL ET AL., 1962). Auch die Killer-Mechanismen "hump-killing" (Tab. 1.3.3, PREER ET AL., 1974) und "vacuolization" wären ein Hinweis auf eine vom putativen Protein vermittelte Störung der Osmoregulation. Für zu ParA homologe Proteine aus Prokaryoten ist eine Membran-Assoziation ebenfalls nachgewiesen (BIGNELL ET AL., 2001).

Ein weiteres pKAP298-Fragment konnte, trotz wiederholter und gründlicher Vorgehensweise, ebenfalls nicht in *E. coli* subkloniert werden (37926 Bp - 41403 Bp). Auf diesem Fragment sind zwei ORFs ohne Homologien zu schon bekannten Proteinen lokalisiert (ORF 45 und ORF 46). Des weiteren befinden sich 4 ORFs mit Ähnlichkeiten zu bekannten Protein auf dem Fragment. Dabei handelt es sich um eine putative reverse Transkriptase eines Gruppe-II-Introns (ORF 47), ein hypothetisches Protein, dessen Funktion unbekannt ist (ORF 48), und Teile eines ursprünglichen, putativen, RecE-Proteins (ORF 49 und 50). Inwiefern dieses nicht subklonierbare pKAP298 Fragment für *E. coli* und vielleicht für sensitive Paramecien toxische Sequenzen enthält, konnte nicht geklärt werden. Auf dem Fragment wurde jedoch in

jedem detektierten Southern Blot Transkriptions-Aktivität nachgewiesen, so dass man davon ausgehen kann, dass dort wirklich Gene exprimiert werden. Inverted Repeats und direkte Repeats, die zu ungewöhnlichen Sekundärstrukturen der DNA führen, und somit die Replikation des rekombinanten Vektors hätten verhindern können (PINDER ET AL., 1998 und KRASILNIKOVA ET AL., 1998), konnten auf diesem Fragment nicht nachgewiesen werden.

Stabile DNA-Sekundärstrukturen auf diesen beiden nicht klonierbaren Fragmenten (oder anderswo auf dem Plasmid) könnten auch dafür verantwortlich sein, dass das Projekt "Konstruktion eines Cry4B-Expressionsvektors mit pKAP298" nicht mit Erfolg beendet werden konnte (siehe Anhang). Es wäre ebenfalls möglich, dass auf den Fragmenten (oder anderswo auf dem Plasmid) Gene lokalisiert sind, die auch von Eukaryoten exprimiert werden und auf RNA-Ebene (durch das Silencen essentieller Gene) oder auf Enzym-Ebene für toxische Produkte kodieren (vielleicht ein ParA-Pendant). Dieser Sachverhalt müsste ebenfalls näher untersucht werden.

Generell könnte man, um die auf dem Plasmid identifizierten putativ kodierenden ORFs auf ihre Toxizität zu untersuchen, einen Vektor mit einem Operon benutzen, das sich (z.B. durch IPTG) zur Protein-Expression eines inklonierten putativen ORFs induzieren lässt. Der zu untersuchende ORF könnte mittels PCR (mit einer Polymerase, die eine Proofreading-Eigenschaft besitzt, wie *Pwo*-Polymerase) aus dem Plasmid amplifiziert und in einen solchen Vektor umkloniert werden. Selbst wenn der ORF (wie vermutlich ORF 43) für ein Protein kodiert, das für *E. coli* toxisch ist, sollte er doch, im nicht induzierten Zustand, das Überleben der Bakterien ermöglichen. Für die Toxizität der *Caedibacter* sind die R-Körper-Proteine ein zwingend notwendiger Faktor - wahrscheinlich wird das hypothetische Toxin durch die Zerstörung der Nahrungsvakuolen (aufgrund des "Aufrollens" der R-Körper) in das Cytoplasma der sensitiven Paramecien freigesetzt (JURAND ET AL., 1978). Während meiner Diplomarbeit habe ich das pKAP298-Fragment, das die R-Körper kodierenden Gene trägt, in *E. coli* subkloniert - es ist also ein Klon vorhanden, der R-Körper exprimiert. Man könnte also die R-Körper produzierenden Bakterien zusammen mit den induzierten *E. coli*, die den potentiell toxischen ORF inkloniert haben, an sensitive Paramecien verfüttern. Treten toxische Effekte auf (in diesem Fall hump-killing), hätte man einen Beweis dafür, dass der untersuchte ORF für das Toxin kodiert. Möglich ist allerdings auch, dass mehrere Proteine für die Toxizität verantwortlich sind.

Ignoriert man das Kriterium, dass das hypothetische Toxin möglicherweise zwischen 20 kDa und 25 kDa groß ist, und nimmt nur die Lage eines ORFs auf einem Fragment mit nachgewiesener Transkriptions-Aktivität als hinreichendes Merkmal, kommt man auf eine

Reihe weiterer potentiell für das Toxin kodierender ORFs. Diese sind schon in 4.1.1.1 und 4.1.1.2 aufgeführt. Theoretisch könnte das putative Toxin auch auf pKAP-Fragmenten lokalisiert sein, die keine Transkriptions-Aktivität aufwiesen. Es könnte auch aus dem Genom von *Caedibacter taeniospiralis* stammen, was jedoch unwahrscheinlich ist, da manche *C. taeniospiralis*-Stämme wenig DNA-Verwandtschaft zu anderen Stämmen ihrer Gattung besitzen (QUACKENBUSH, 1977, 1978, siehe 4.2.6) und trotzdem eine gleich geartete Toxizität aufweisen. Als wahrscheinlichster Sitz des Toxins wird daher das Plasmid pKAP298 angenommen, dabei präferentiell die Fragmente, auf denen während dieser Arbeit Transkriptions-Aktivität nachgewiesen werden konnte. Unter diesen Fragmenten bieten sich die nicht in *E. coli* subklonierbaren Fragmente (33785 Bp - 36842 Bp und 37926 Bp - 41403 Bp) für die weitere Forschung nach dem Toxin an.

Die Vermutung, dass das *rebD*-Gen (ORF 1) möglicherweise für das Toxin kodiert (POND, unveröffentlicht, 2002), kann durch diese Arbeit nicht gestützt werden. Die Aminosäuresequenz des putativen Proteins auf pKAP298 unterscheidet sich in 6 hintereinanderliegenden Aminosäuren von denen auf pKAP47 und pKAP116 (siehe 4.2.5). Dieser Unterschied könnte ausreichen, die Funktionalität des putativen Proteins zu zerstören (je nachdem ob ein enzymatisch aktives Zentrum des Proteins zerstört wird). Bisher weiß man nicht, ob das *rebD*-Gen überhaupt exprimiert wird (HERUTH ET AL., 1996, POND, unveröffentlicht, 2002). Dies wäre mittels der in 4.1.2 beschriebenen Oligonucleotid-RNA/cDNA-Sonden-Hybridisierungsmethode oder über RT-PCR zu überprüfen.

5. Zusammenfassung

- aus Kulturen von *Paramecium tetraurelia* Stamm 298 wurde der Endobiont *Caedibacter taeniospiralis* isoliert
- aus *Caedibacter taeniospiralis* wurde das Plasmid pKAP298 isoliert
- 60 % der fehlenden pKAP298-Sequenz wurde in *E. coli* subkloniert, so dass nun (mit den in meiner Diplomarbeit subklonierten Stücken) insgesamt 84,2 % der Gesamtsequenz von pKAP298 in *E. coli* subkloniert sind
- mit dem geschnittenen Plasmid pKAP298 und den amplifizierten pKAP-Fragmenten aus *E. coli* wurden Southern Blots hergestellt
- aus *Caedibacter taeniospiralis* wurde die Gesamt-RNA isoliert
- die Gesamt-RNA wurde in cDNA umgeschrieben und mit Digoxigenin markiert
- die markierten cDNA-Sonden wurden mit den Southern Blots aus pKAP298-Fragmenten hybridisiert und per Chemolumineszenz detektiert
- aus den detektierten Blots wurde eine Karte des Plasmids pKAP298 erstellt, auf der die Orte mit nachgewiesener Transkriptions-Aktivität (zum Zeitpunkt der RNA-Isolation) eingetragen sind
- das Plasmid pKAP298 (49,1 kB) wurde sequenziert
- die Basensequenz des Plasmids wurde mit GeneMark.hmm nach potentiell kodierenden Open Reading Frames durchsucht
- die potentiell kodierenden Bereiche wurden mit (Nucleotid- und Protein-) BLAST auf Ähnlichkeiten zu schon bekannten Sequenzen untersucht
- die pKAP298-Sequenz wurde nach Transposons durchsucht. Es wurde ein neues Transposon auf pKAP298 entdeckt, schon beschriebene Transposons konnten näher charakterisiert werden
- anhand der potentiell kodierenden ORFs, deren Sequenz-Ähnlichkeiten und den gefundenen Transposons, wurde das Plasmid annotiert, sowie mögliche vom Plasmid kodierte Genprodukte charakterisiert
- es wurde ein Gruppe-II-Intron auf pKAP298 identifiziert und beschrieben
- es wurde eine Karte des Plasmids pKAP298 erstellt, auf der die kodierenden ORFs und die gefundenen Transposons, sowie die Orte mit nachgewiesener Transkriptions-Aktivität, eingetragen wurden
- den Orten mit Transkriptions-Aktivität wurden die entsprechenden potentiell kodierenden ORFs zugeordnet und diskutiert (4.1)
- aufgrund der (Nuclein- und Aminosäure-)Sequenz-Homologien zu schon bekannten Sequenzen wurden Überlegungen zur Evolution des Plasmids pKAP298 angestellt (4.2). Dabei konnte die Hypothese, dass pKAP298 aus einem Bakteriophagen

evolvierte, gestützt werden, da putative Phagen-Gene auf pKAP298 entdeckt wurden. Der putative Ur-Phage wurde anhand der Homologien näher charakterisiert und den *Caudovirales* zugeordnet. Durch den Vergleich der pKAP298-Sequenz mit schon bekannten Sequenzen der Plasmide pKAP116 und pKAP47 und einer (R-Körper-kodierenden) Sequenz aus *Caedibacter caryophilus* (die Sequenz ist wahrscheinlich auf einem Bakteriophagen lokalisiert) wurde postuliert, dass pKAP298 evolutionär älter als pKAP116 und pKAP47 ist, und diese sich aus pKAP298 entwickelt haben.

- mit Hilfe der nachgewiesenen Transkriptions-Aktivität und den potentiell kodierenden Regionen wurde versucht das putative Toxin zu identifizieren (4.3).
- es wurde versucht einen Cry4B-Expressionsvektor für *Caedibacter taeniospiralis* zu konstruieren (Anhang)

6. Abkürzungsverzeichnis

Abkürzung	Bedeutung
Abb.	Abbildung
abs.	absolut
Aqua _{bidest}	Aqua bidestilliert
AS	Aminosäure
ATP	Adenosin-5'-triphosphat
Bp / kB	Basenpaare / Kilo-Basenpaare
BLAST	"basic local alignment search tool"
bzw.	beziehungsweise
<i>C. xxx</i>	<i>Caedibacter xxx</i>
cm ²	Quadratzentimeter
d.h.	das heißt
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytidintriphosphat
dGTP	Desoxyguanosintriphosphat
DigOx	Digoxigenin-11-uridintriphosphat
DNA	Desoxyribonucleinsäure
dNTP	Mix aus dATP, dCTP, dGTP, dTTP
dTTP	Desoxythymidintriphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
ECTEOLA	Epichlorhydrin Triethanolamin Cellulose
EDTA	Ethylendiamintetraessigsäure
EtOH	Ethanol
(μ)F	(Mikro)Farad
g	je nach Kontext Gramm oder Erdbeschleunigung $\approx 9,81 \text{ m/s}^2$
h	Stunde
IPTG	Isopropyl-β-D-thiogalactosid
kB	Kilobasenpaare
konz.	konzentriert
l / μl	Liter / Mikroliter
LB-Medium	Luria-Bertani Medium
m / mm / μm	Meter / Millimeter / Mikrometer

Abkürzung	Bedeutung
M / mM	Mol / Millimol
min	Minute
NaCl	Natriumchlorid
OD	optische Dichte
ORF(s)	open reading frame(s)
p. a.	per analysi (zur Analyse)
<i>P. xxx</i>	<i>Paramecium xxx</i>
PCR	Polymerase Chain Reaction
pH	potentia hydrogenii
R-Körper	refraktiler Körper
RNA	Ribonucleinsäure
rpm	rounds per minute
RT	Raumtemperatur
s	Sekunde
SDS	Sodium Dodecyl Sulphate
SSC	Standard-Salzcitratlösung (Natriumchlorid/Natriumcitrat)
Tab.	Tabelle
TAE	Tris-Acetat-EDTA
TE	Tris-EDTA
Tris	Tris(hydroxymethyl)-aminomethan
u	unit
UV	Ultra violett
V	Volt
v/v	Volumenprozent
Vol.	Volumen
w/v	Gewichtsprozent
WGP	Wheat Gras Powder
X-Gal	5-Chlor-4-Brom-3-Indonyl- β -D-Galactopyranosid
z.B.	zum Beispiel
°C	Grad Celsius
Ω	Ohm

7. Literaturverzeichnis

- AUSTIN, M.L. (1948): The killing substance, paramecin: activity of single particles. *Am. Nat.* **82**:51-59
- BALSLEY, M. (1967): Dependence of the *kappa* particles of stock 7 of *Paramecium aurelia* on a single gene. *Genetics* **56**: 125-131
- BAUM, J.A. (1994): Tn5401, a new class II transposable element from *Bacillus thuringiensis*. *J. Bacteriol.* **176**: 2835-2845
- BAUM, J.A., GILMER, A.J., LIGHT METTUS, A.M. (1999): Multiple roles for TnpI recombinase in regulation of Tn5401 transposition in *Bacillus thuringiensis*. *J. Bacteriol.* **181**:6271-7
- BEIER, C.L., HORN, M., MICHEL, R., SCHWEIKERT, M., GORTZ, H.-D., AND WAGNER, M. (2002) : The genus *Caedibacter* comprises Endobionts of *Paramecium* spp. related to the *Rickettsiales* (*Alphaproteobacteria*) and to *Francisella tularensis* (*Gammaproteobacteria*). *Appl. Environ. Microbiol.* **68**:6043-6050
- BEGON, M., HARPER, J.L., TOWNSEND, C.R. (1991): Ökologie, 2.Auflage. *Birkhäuser, Basel*
- BERNARD, P., GABANT, P., BAHASSI, E.M., COUTURIER, M. (1994): Positive-selection vectors using the F plasmid *ccdB* killer gene. *Gene.* **148**:71-4
- BIGNELL C., THOMAS, C.M. (2001): The bacterial ParA-ParB partitioning proteins. *J. Biotechnol.* **91**(1):1-34
- BIRNBOIM, H.C., DOLY, J. (1979). *Nucleic Acids Res.* **7**, 1513-1523
- BLACK, L. W. (1989): DNA packaging in dsDNA Bakteriophages. *Annu Rev. Microbiol.* **43**: 267-292
- BUTZEL, H.M., BROWN, L.H., MARTIN, W.B. (1960): Effects of detergents upon killer-sensitive reactions in *Paramecium aurelia*. *Physiol. Zool.* **33**: 213-224

- BUTZEL, H.M, PAGLIARA, A. (1962): The effect of biochemical inhibitors upon the killer sensitive system in *Paramecium aurelia*. *Exp Cell Res.* **27**: 382-395
- CHANG, H.W., JULIN, D.A. (2001): Structure and function of *Escherichia coli* RecE Protein, a Member of the RecB Nuclease Domain Family. *J. of Biol. Chemistry* **276**: 46004-46010
- CHAO, P.K. (1953): Kappa concentration per cell in relation to the life cycle, genotype and mating type in *Paramecium aurelia*, variety 4. *Proc Natl. Acad. Sci. USA* **39**: 103-113
- CERRY, J.L. (2003): Genome size and operon content. *J Theor Biol.* **221**:401-10
- COOPER, S., KEASLING, J.D. (2002): Experimental and Theoretical Considerations of P1-plasmid Replication and Segregation during the E. coli Cell Cycle. <http://www-personal.umich.edu/~cooper/P1.pdf>
- CUTTING, S., MANDELSTAM, J. (1986). The nucleotide sequence and the transcription during sporulation of the *gerE* gene of *Bacillus subtilis*. *J. Gen. Microbiol.* **132**:3013–3024
- DAVEY, M.J., FUNNELL, B.E. (1994): The P1 plasmid partition protein ParA. A role for ATP in site-specific DNA binding. *J Biol Chem.* **269**:29908-13
- DELECLUSE, A., PONCHET, S., KLIER, A., RAPOPORT, G. (1993): Expression of *cry4A* and *cry4B* Genes, Independently or in Combination, in a Crystal-Negative Strain of *Bacillus thuringiensis israelensis*. *Appl. and envir. Microb.* **59**: 3922-3927
- DILTS, J.A. (1976): Covalently closed, circular DNA in kappa Endobionts of *Paramecium*, *Genet. Res.* **27**: 161-170
- DILTS J.A. (1986A): the importance of refractile body in expression of the killer trait in *Paramecium*. In: R.B. Wickner, A. Hinnebusch, A.M. Lambowitz, C. Gunsalus and A. Hollaender (ed.), *Extrachromosomal elements in lower eukaryotes*. Plenum Press, New York
- DILTS, J.A., QUAKENBUSH, R.L. (1986B): A mutation in the R-body-coding sequences destroys expression of the killer trait in *P. tetraurelia*. *Science* **232**: 641-643

- EASTER, J., GOBER, J.W. (2002): ParB-Stimulated Nucleotide Exchange Regulates a Switch in Functionally Distinct ParA Activities. *Mol. Cell* **10**: 427-434
- ERDMANN, N., PETROFF, T., FUNNELL, B.E. (1999): Intracellular localization of P1 ParB protein depends on ParA and *parS*. *Proc. Natl. Acad. Sci. USA* **96**: 14905-14910
- ESTÈVE, J.C. (1978): Une population de type "killer" chez *P. caudatum* (Ehrenberg), *Protistologica*, **14**: 201-207
- FOKIN, S., KARPOV, S. (1995): Bacterial endocytobionts inhabiting the pernicular space of protista. *Endocyt. Cell Res.*, **11**: 81-94
- FOUTS, K.E., WASIE-GILBERT, T., WILLIS, D.K., CLARK, A.J., BARBOUR, S.D. (1983): Genetic analysis of transposon-induced mutations of the *rac* prophage in *Escherichia coli* K-12 which affect expression and function of *recE*. *J.Bacteriol.* **156**:718-726
- FTOUHI, N., GUILLEN, N. (1990): Genetic Analysis of Fusion Recombinants in *Bacillus subtilis*: Function of the *recE* Gene. *Genetics*, **126**: 487-496
- FUJISAWA H, MORITA M. (1997): Phage DNA packaging. *Genes Cells.* **2(9)**:537-45
- GEIDUSCHEK, E.P., KASSAVETIS, G.A. (1988): Changes in RNA polymerase. *The Bacteriophages* **1**: 93-115
- GRUNBERG-MANAGO, M. (1999): Messenger RNA Stability and its Role in Control of Gene Expression in Bacteria and Phages. *Ann. Rev. of Gen.* **33**: 193-227
- HECKMANN, K. AND GÖRTZ, H.D. (1991): Procaryotic symbionts of ciliates. In: Balows, A., Truper, H.G., Dworkin, M., Schleifer, K.H. (eds.): *The Prokaryotes*, 2d edn, Springer Verlag, Berlin: 3865-3890
- HERUTH, D.C., POND, F.R., DILTS, J.A., QUAKENBUSH, R.L. (1994). Characterization of Genetic Determinants for R Body Synthesis and Assembly in *Caedibacter taeniospiralis* 47 and 166. *Journal of Bacteriology.* , **176**: 3559-3567

- HERUTH, D.C., CHRISTIAN, S.L., HUCKABY, J.J., QUACKENBUSH, R.L. AND POND, F.R. (1996): Characterization of transposable genetic elements in *Caedibacter taeniospiralis* plasmid DNA, NCBI
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=nucleotide&list_uids=1438743&dopt=GenBank
- HUANG, C.C., NARITA, M., YAMAGATA, T., ITOH, Y. AND ENDO, G. (1999): Structure analysis of a class II transposon encoding the mercury resistance of the Gram-positive Bacterium *bacillus megaterium* MB1, a strain isolated from minamata bay, Japan. *Gene* **234**: 361-369
- HU, F., ZHANG, K., JIN, X., RAO, X., TAN, Y., ZHU, J., CHEN, Z. (2002): Complete genome sequence of *Pseudomonas aeruginosa* phage PaP3. unpubliziert, direct submission:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Protein&list_uids=27476053&dopt=GenPept
- IANDOLO, J.J., WORRELL, V., GROICHER, K.H., QIAN, Y., TIAN, R., KENTON, S., DORMAN, A., JI, H., LIN, S., LOH, P., QI, S., ZHU, H., ROE, B.A. (2002): Comparative analysis of the genomes of the temperate bacteriophages phi11, phi12 and phi13 of *Staphylococcus aureus* 8325. *Gene* **289**: 109-118
- JEBLICK (2000): Subklonierung des Plasmids pKAP298 aus *Caedibacter taeniospiralis* in *E. coli* und die Lokalisierung eines Replikations-Startpunktes auf dem Plasmid pKAP298. Diplomarbeit der Abteilung Ökologie des Fachbereiches Biologie der Universität Kaiserslautern
- JURAND, A., PREER, J.R., JR. AND RUDMAN, B.M. (1978): Further investigations on the prelethal effects of killing action of *kappa*-containing killer stocks of *Paramecium aurelia*. *J. Exp. Zool.* **206**: 25-48
- JURAND, A., RUDMANN, B.M. AND PREER J.R., JR. (1971): Prelethal effects of killing action by stock 7 of *Paramecium aurelia*. *J. Exp. Zool.* **177**: 365-388
- KANABROCKI, J.A., QUACKENBUSH R.L., POND F.R. (1986): Organization and Expression of Genetic determinants for Synthesis and Assembly of type 51 R-bodies. *J. of Bacteriology*, **168**: 40-48
- KANEKO, T., NAKAMURA, Y., SATO, S., MINAMISAWA, K., UCHIUMI, T., SASAMOTO, S., WATANABE, A., IDESAWA, K., IRIGUCHI, M., KAWASHIMA, K., KOHARA, M., MATSUMOTO, M., SHIMPO, S., TSURUOKA, H., WADA, T., YAMADA, M., TABATA,

- S.: Complete genomic sequence of nitrogen-fixing symbiotic *bacterium Bradyrhizobium japonicum* USDA110. *DNA Res.* **9**: 189-197
- KIM, J., NIETFELDT, J. AND BENSON, A. K. (1999): Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. *Proc. Natl. Acad. Sci. U.S.A.* **96** : 13288-13293
- KIRNDORFER, M., JAGER, A., KLUG, G. (1998): Integration host factor affects the oxygen-regulated expression of photosynthesis genes in *Rhodobacter capsulatus*. *Mol Gen Genet.* **258**(3):297-305
- KLAUS, S., KRÜGER, D. H., MEYER, J. (1992): Bakterienviren. *Gustav Fischer Verlag Jena*
- KLEIN, J. R., DUNNY, G. M. (2002): Bacterial group II introns and their association with mobile genetic elements. *Front Biosci.* **Aug 1**: 1843-56
- KRASILNIKOVA, M. M., SAMADASHWILY, G. M., KRASILNIKOV, A. S., MIRKIN, S. M. (1998): Transcription through a simple DNA repeat blocks replication elongation. *EMBO J.* **17**:5095-102
- KULAEVA, O. I., KOONIN, E. V., WOOTTON, J. C., LEVINE, A. S. AND WOODGATE, R. (1998): Unusual insertion element polymorphisms in the promoter and terminator regions of the *mucAB*-like genes of R471a and R446b. *Mutat. Res.* **397**: 247-262
- KUSCH, J. (1999): Bildung von Anopheles-spezifischen Toxinen durch *Paramecium*-Symbionten, und ihre ökologischen Effekte. *DFG-Forschungsantrag*
- KUSCH, J., CZUBATINSKI, L., WEGMANN, S., HÜBNER, M., ALTER, M., ALBRECHT, P. (2002): Competitive Advantages of *Caedibacter*-Infected *Paramecia*. *Protist*, Vol. **153**: 47-58
- LANDIS, W. G. (1986): Factors Determining the Frequency of the Killer Trait within Populations of *Paramecium aurelia* Complex. *Genetics*, **115**: 197-205
- LANDIS, W. G. (1988): Ecology, In: Görtz, H.-D. (1988): *Paramecium*. Springer, Berlin 419-436

- LAPAGE, S.P., SNEATH, P.H.A., LESSEL, E.F., SKERMAN, V.B.D., SEELINGER, H.P.R. AND CLARK, W.A. (1975): International code of nomenclature of bacteria-bacteriological code. *Am. Soc. Microbiol.*, Washington, DC.
- LEHMANN, K., SCHMIDT, U. (2003): Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit Rev Biochem Mol Biol.* **38**:249-303
- LEWIN, B. (2002): Molekularbiologie der Gene, 6. Auflage. *Spektrum Lehrbuch*
- LUKASHIN, A.V., BORODOVSKI, M. (1998): GeneMark.hmm: new solutions for gene finding. *Nucl. Acids Res.* Vol. **26**, No.4: 1107-1115
- LUNEBERG, E., MAYER, B., DARYAB, N., KOOISTRA, O., ZHRINGER, U., ROHDE, M., SWANSON, J. AND FROSCH, M. (2001): Chromosomal insertion and excision of a 30 kB unstable genetic element is responsible for phase variation of lipopolysaccharide and other virulence determinants in *Legionella pneumophila*. *Mol. Microbiol.* **39**, 1259-1271
- MANIATIS, T., FRITSCH, E.R., SAMBROOK, J.: Molecular cloning: a laboratory manual. *Second edition, Cold Spring Harbor Laboratory Press* 1989
- MANILOFF, J., ACKERMANN, H.W. (1998): Taxonomy of bacterial viruses: establishment of tailed virus genera and the order *Caudovirales*. *Arch Virol* **143**: 2051-2063
- MCCLELLAND, M., SANDERSON, K.E., SPIETH, J., CLIFTON, S.W., LATREILLE, P., COURTNEY, L., PORWOLLIK, S., ALI, J., DANTE, M., DU, F., HOU, S., LAYMAN, D., LEONARD, S., NGUYEN, C., SCOTT, K., HOLMES, A., GREWAL, N., MULVANEY, E., RYAN, E., SUN, H., FLOREA, L., MILLER, W., STONEKING, T., NHAN, M., WATERSTON, R. AND WILSON, R.K. (2001): Complete genome sequence of *Salmonella enterica* serovar *Typhimurium* LT2. *Nature* **413**: 852-856
- MMOLAWA, P.T., THOMAS, C.J. AND HEUZENROEDER, M.W. (2003): Genomic Structure of the *Salmonella enterica* Serovar *Typhimurium* DT64 Bacteriophage ST64T: Evidence for Modular Genetic Architecture. *J. Bacteriol.* **185**, 3473-3475

- MOTALLEBI-VESHAREH, M., ROUCH, D.A., THOMAS, C.M. (1990): A family of ATPases involved in active partitioning of diverse bacterial plasmids. *Mol. Microbiol.* **4**: 1455-63
- MUELLER, J.A. (1963): Separation of kappa particles with infective activity from those with killing activity and identification of the infective particles in *Paramecium aurelia*. *Experimental Cell Research* **30**: 492-508
- MURAKAWA, G.J., KWAN, C., YAMASHITA, J., NIERLICH, D.P. (1991): Transcription and decay of the lac messenger: role of an intergenic terminator. *J Bacteriol.* **173**:28-36
- NELSON, K., PAULSEN, J., WEINEL, C., DODSON, R., HILBERT, H., FOUTS, D., GILL, S., POP, M., MARTINS DOS SANTOS, V., HOLMES, M., BRINKAC, L., BEANAN, M., DEBOY, R., DAUGHERTY, S., KOLONAY, J., MADUPU, R., NELSON, W., WHITE, O., PETERSON, J., KHOURI, H., HANCE, I., LEE, P., HOLTZAPPLE, E., SCANLAN, D., TRAN, K., MOAZZEZ, A., UTTERBACK, T., RIZZO, M., LEE, K., KOSACK, D., MOESTL, D., WEDLER, H., LAUBER, J., HOHEISEL, J., STRAETZ, M., HEIM, S., KIEWITZ, C., EISEN, J., TIMMIS, K., DUESTERHOFT, A., TUMMLER, B., FRASER, C. (2002): Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* **4**: 799-808
- NICOLL, D.S.T. (1995): Labor im Fokus: Gentechnische Methoden. *Spektrum Akademischer Verlag Heidelberg*
- OBERTO, J., DRLICA, K., ROUVIERE-YANIV, J. (1994): Histones, HMG, HU, IHF: Meme combat. *Biochimie* **76**: 901-8
- PEDULLA, M. L., FORD, M. E., HOUTZ, J. M., KARTHIKEYAN, T., WADSWORTH, C., LEWIS, J. A., JACOBS-SERA, D., FALBO, J., GROSS, J., PANNUNZIO, N. R., BRUCKER, W., KUMAR, V., KANDASAMY, J., KEENAN, L., BARDAROV, S., KRIAKOV, J., LAWRENCE, J. G., JACOBS, W. R. JR., HENDRIX, R. W. AND HATFULL, G. F. (2003): Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**: 171-182
- PINDER, D.J., BLAKE, C.E., LINDSEY, J.C., LEACH, D.R. (1998): Replication strand preference for deletions associated with DNA palindromes. *Mol Microbiol.* **28**:719-27

- PLUNKETT, G., ROSE, D.J., DURFEE, T.J. AND BLATTNER, F.R. (1999): Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* **181**, 1767-1778
- POND, F., GIBSON, I., LALUCAT, J., QUAKENBUSH, R.L. (1989): R-Body-Producing Bacteria. *Bacteriological Review*, **53**: 25-67
- PREER, J.R., JR. (1950): Mikroskopically visible bodies in the cytoplasm of the killer strains of *Paramecium aurelia*. *Genetics*, **35**: 344-362
- PREER, L.B., PREER, J.R., JR (1964): Killing activity from lysed particles of *Paramecium*. *Genet. Res.*, **5**: 230-239
- PREER, J.R., JR, HUFNAGEL, L.A., PREER, L.B. (1966): Structure and behavior of "R"-bodies from killer paramecia. *J. Ultrastr. Research*, **15**: 131-143
- PREER, J.R, PREER, L.B., RUDMAN, B.M., JURAND, A. (1971): Isolation and composition of bacteriophage-like particles from kappa of killer paramecia. *Mol. Gen. Genet.* **111**: 202-208
- PREER, L.B., JURAND, A., PREER, J.R., RUDMAN, B.M. (1972): The classes of kappa in *Paramecium aurelia*. *J. Cell Sci.* **11**: 581-600
- PREER, J.R., JR., PREER, L.B., JURAND, A. (1974): Kappa and Other Endobionts in *Paramecium aurelia*. *Bacteriological Review*, **38**: 113-163
- QUAKENBUSH, R.L. (1977): Phylogenetic relationships of bacterial Endobionts of *Paramecium aurelia*: deoxyribonucleotide sequence relationship of 51 kappa and its mutants. *J. Bacteriol.* **129**: 895-900
- QUAKENBUSH, R.L. (1978): Genetic relationships among bacterial Endobionts of *Paramecium aurelia*: polynucleotide sequence relationships among members of *Caedibacter*. *J. Gen. Microbiol.* **108**:181-187
- QUAKENBUSH, R.L. (1983A): Plasmids from bacterial Endobionts of hump-killer paramecia. *Plasmid* **9**: 298-306

- QUAKENBUSH, R.L., BURBACH, J.A. (1983B): Cloning and expression of DNA sequences associated with the killer trait of *Paramecium tetraurelia* stock 47. *Proc. Natl. Acad. Sci. USA*, **80**: 250-254
- QUAKENBUSH, R.L., COX, B. J., KANABROCKI, J.A. (1986A): Extrachromosomal elements of extrachromosomal elements of paramecia, and their extrachromosomal elements, p. 265-278. In A. Hollaender et al. (ed.), Extrachromosomal elements of lower eukaryotes. *Plenum Publishing Corp.*, New York
- QUAKENBUSH, R.L., DILTS, J.A., COX, B.J. (1986B): Transposonlike elements in *Caedibacter taeniospiralis*. *J. Bacteriol.* **166**: 349-352
- QUAKENBUSH, R.L. (1988): Endobionts of killer Paramecia. In: Görtz, H.-D. (1988). *Paramecium*. Springer, Berlin 406-418
- RAVIN, V., RAVIN, N., CASJENS, S., FORD, M.E., HATFULL, G.F., HENDRIX, R.W. (2000): Genomic sequence and analysis of the atypical temperate bacteriophage N15. *J Mol Biol.* **299**: 53-73
- READ, T.D., SALZBERG, S.L., POP, M., SHUMWAY, M., UMayAM, L., JIANG, L., HOLTZAPPLE, E., BUSCH, J.D., SMITH, K.L., SCHUPP, J.M., SOLOMON, D., KEIM, P., FRASER, C.M. (2002): Comparative Genome Sequencing for Discovery of Novel Polymorphisms in *Bacillus anthracis*. *Science* (2002)
- ROMINE, M.F., STILLWELL, L.C., WONG, K.-K., THURSTON, S.J., SISK, E.C., SENSEN, C., GAASTERLAND, T., FREDRICKSON, J.K. AND SAFFER, J.D. (1999): Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. *J. Bacteriol.* **181**: 1585-1602
- RUPRECHT, V. (1999): Konstruktion eines Expressionsvektors für *Caedibacter taeniospiralis*. Diplomarbeit der Abteilung Ökologie des Fachbereiches Biologie der Universität Kaiserslautern
- SAIKI, R.K., GELFAND, D.H., STOFFEL, S., SCHARF, S.J., HIGUCHI, R., HORN, G.T., MILLIS, K.B. AND EHRLICH, H.A. (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA Polymerase. *Science* **293**: 487-491

- SANGER, F., NICKLEN, S., COULSON, A. (1977): DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467
- SCHLEE, D., KLEBER, H.-P. (Hrsg.) (1991): Biotechnologie I+II, 1.Auflage. *Gustav Fischer, Jena*
- SCHMIDT, H.J., GÖRTZ, H.-D., POND, F.R., QUAKENBUSH, R.L. (1987): Characterization of *Caedibacter* endonucleosymbionts from the macronucleus of *Paramecium caudatum* and the identification of a mutant with blocked R body synthesis. *Exp. Cell Res.* **174**: 49-57
- SCHMIDT, H.J., GÖRTZ, H.-D. AND QUAKENBUSH, R.L. (1987): *Caedibacter caryophilus* sp. nov., a killer symbiont inhabiting the macronucleus of *Paramecium caudatum*. *Int. J. Syst. Bacteriol.* **37**: 459-462
- SCHUHMANN, W. (1990): Biologie bakterieller Plasmide. *Friedr. Vieweg & Sohn Verlagsgesellschaft mbh, Braunschweig*
- SCHULMEISTER, S.A., KROL, J.E., VORHOELTER, F.-J., SKORUPSKA, A.M. AND LOTZ, W. (2001): Sequence of the genome of *Sinorhizobium meliloti* bacteriophage PBC5, NCBI direct submission, Nr. NC_003324.1
- SHIRAISHI K, HANADA K, IWAKURA Y, IKEDA H. (2002): Roles of RecJ, RecO, and RecR in RecET-mediated illegitimate recombination in *Escherichia coli*. *J Bacteriol.* **184**: 4715-21
- SHUMAN, S. (1991): DNA strand transfer reactions catalyzed by *vaccinia* topoisomerase I. *J Biol Chem.* **267**:8620-7
- SMITH-SONNEBORN, J.E., VAN WAGTENDONK, W.J. (1964): Purification and chemical characterization of kappa of stock 51, *Paramecium aurelia*. *Exp. Cell. Res.* **33**: 50-59
- SONNEBORN, T.M. (1938): Mating types, toxin interactions and heredity in *Paramecium aurelia*. *Science* **88**: 503
- SONNEBORN, T.M. (1943): Gene and Cytoplasm. I. The determination and inheritance of the killer character in variety 4 of *P. aurelia*. II. the bearing of determination and

- inheritance of characters in *P. aurelia* on problems of cytoplasmic inheritance, pneumococcus transformation, mutations and development. *Proc. Natl. Acad. Sci. USA*, **29**: 329-343
- SONNEBORN, T.M. (1959): *Kappa* and related particles in *Paramecium*. *Adv. Virus Res.*, **6**: 229-359
- STAHL, M.M., THOMASON, L., POTEETE, A.R., TARKOWSKI, T., KUZMINOW, A., STAHL, F. (1997): Annealing vs. Invasion in Phage λ Recombination. *Gen. Soc. Of America* **147**: 961-977
- STREMMEL, M. (1995): Experimente zur vergleichenden Untersuchung von "Killer-Paramecien" der Arten *Paramecium caudatum* und *Paramecium aurelia*. Diplomarbeit des Institutes für Allgemeine Zoologie und Genetik der Westfälischen Wilhelms-Universität Münster
- STREMMEL, M. (1999): Phylogenie und Artidentifizierung bei *Euplotes*-Arten und ihren Endobionten. Dissertation der Abteilung Ökologie des Fachbereiches Biologie der Universität Kaiserslautern
- STUTTARD, C., ROZEE, K.R. (1980): Plasmids and Transposons. *Academic Press, New York*
- VANDER BYL, C., KROPINSKI, A.M. (2000): Sequence of the genome of *Salmonella* bacteriophage P22. *J. Bacteriol.* **182**: 6472-6481
- WANG, Y., MOORE, M., LEVINSON, H.S., SILVER, S., WALSH, C., MAHLER, I. (1989): Nucleotide sequence of a chromosomal mercury resistance determinant from a *Bacillus sp.* with broad-spectrum mercury resistance. *J. Bacteriol.* **171**: 83-92
- WEBSEITE <http://info.bio.cmu.edu/Courses/03441/TermPapers/99TermPapers/Caulo/parab.html> : Plasmid Partitioning Genes and Their Homologs in Bacterial Cells
- WEISEL, T. (2001): Charakterisierung der R-Körper-Gene von *Caedibacter caryophilus* BGD19, Diplomarbeit Fachbereich Biologie, Abteilung Ökologie, Universität Kaiserslautern

- YAMAMOTO, T., WATKINSON, I.A., KIM, K., SAGE, M.V., STRATTON, R., AKANDE, Y.L., MA, D.P, ROE, B.A (1988): Nucleotide sequence of the gene coding for a 130-kDa mosquitocidal protein of *Bacillus thuringiensis israelensis*. *Gene* **66**: 107-120
- YOKOYAMA, K., MAKINO, K., KUBOTA, Y., WATANABE, M., KIMURA, S., YUTSUDO, C.H., KUROKAWA, K., ISHII, K., HATTORI, M., ABE, H., IIDA, T., YAMAMOTO, K., HAYASHI, T., YASUNAGA, T., HONDA, T., SASAKAWA, C. AND SHINAGAWA, H. (2000): Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic *Escherichia coli* O157:H7 strain derived from the Sakai outbreak. *Gene* **258**, 127-139
- YURIEVA, O., KHOLODII, G., MINAKHIN, L., GORLENKO, Z., KALYAEVA, E., MINDLIN, S., NIKIFOROV, V. (1997): Intercontinental spread of promiscuous mercury-resistance transposons in environmental bacteria. *Mol. Microbiol.* **24**: 321-329
- ZHANG, Y., MUYRERS, J.P.P, RIJNTJES, J., STEWART, A.F. (2003): Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. BMC Molecular Biology, <http://www.biomedcentral.com/1471-2199/4/1>
- ZHENG, L., HALBERG, R., ROELS, S., ICHIKAWA, H., KROOS, L., LOSICK, R. (1992). Sporulation regulatory protein GerE from *Bacillus subtilis* binds to and can activate or repress transcription from promoters for mother-cell-specific genes. *J. Mol. Biol.* **226**:1037–1050
- ZHOU, T., RADAEV, S., ROSEN, B.P., GATTI, D.L. (2000): Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. *EMBO J.* **19**(17):4838-45

Schließlich möchte ich mich bei allen Leuten bedanken, die durch ihre Mithilfe oder moralische Unterstützung am Gelingen dieser Arbeit beteiligt waren. Ein herzliches Danke geht daher an meine Eltern, meine Lebensgefährtin Anouk Weber und an alle meine Freunde.

Anhang

Das Projekt "Konstruktion eines Cry4B-Expressionsvektors für Caedibacter taeniospiralis 298"

Zusätzlich zum Hauptthema der Arbeit im Bereich der Grundlagenforschung, wurde noch ein anwendungsorientiertes Thema bearbeitet. Dies war die Konstruktion eines Cry4B-Expressionsvektors für *Caedibacter taeniospiralis*. Da der letzte, entscheidende Schritt, die "Endmontage" des Vektors in *Saccharomyces cerevisiae*, ohne Ergebnis blieb, wird dieses Projekt meiner Arbeit nur verkürzt und nicht in der üblichen Gliederung dargestellt. Auf in dem Hauptteil beschriebene Methoden wird nicht näher eingegangen, nur neu verwendete Methoden werden ausführlicher beschrieben.

Kurzer Überblick über das Projekt

Das *cry4B*-Gen aus einem 72 mDa Plasmid des *Bacillus thuringiensis israelensis* kodiert für eine Untereinheit eines Kristallproteins, das in der Lage ist, selektiv die Larven einiger Stechmücken-Arten zu töten (YAMAMOTO ET AL., 1988). Dieses Gen sollte in einen Vektor integriert werden, der die Fähigkeit besitzt, in *Caedibacter taeniospiralis* zu replizieren und das Cry4B-Kristallprotein zu exprimieren. Weil *Caedibacter taeniospiralis* als obligater Endobiont in Pantoffeltierchen lebt, und diese Paramecien ein Teil der Nahrung der zu tötenden Stechmückenlarven sind, könnte vielleicht ein biologisches Insektizid entwickelt werden, das über längere Zeit im Medium wirkt. Da *Caedibacter taeniospiralis* an sich schon ein Plasmid besitzt, lag es nahe, dieses als Teil des zu konstruierenden Vektors zu benutzen, weil dadurch die Replikation des Konstrukts in dem Bakterium gewährleistet wäre. Leider ist dieses Plasmid pKAP298 49 kBp groß und besitzt wahrscheinlich Sequenzen, die exprimiert toxisch für *E. coli* sind (bzw. es kann als Ganzes aufgrund irgendwelcher Sekundärstrukturen nicht repliziert werden, siehe Hauptteil der Arbeit, 4.3). Daher musste die "Endmontage" des Vektors in *Saccharomyces cerevisiae* stattfinden. Die Hefe ist in der Lage, größere Plasmide aufzunehmen. Außerdem würde sie (als Eukaryot) vielleicht die auf dem Plasmid vermuteten toxischen Sequenzen nicht exprimieren.

Konkreter Ablauf:

Das *cry4B*-Gen wurde mit seiner Promotorregion (das Gen wird mit seinem "original"-Promotor auch in *E. coli* exprimiert, YAMAMOTO ET AL., 1988) per Long-Run-PCR aus dem Vektor pHT611 (DELECLUSE ET AL., 1993) amplifiziert. Die Primer für die Amplifikation

enthielten beide eine *Sal* I-Schnittstelle, einer der beiden zusätzlich noch eine *Bam* HI-Schnittstelle. Das amplifizierte Gen wurde, per Topo-XL-PCR-Kloning-Kit (Invitrogen), in einen TOPO-Vektor kloniert und amplifiziert. Mit dem Restriktionsenzym *Sal* I wurde das Gen aus dem Vektor geschnitten und in den *E. coli/Saccharomyces cerevisiae* Shuttle Vektor pNF2 einkloniert. Der Schritt über den TOPO-Vektor war nötig, da sich das mit *Sal* I geschnittene PCR-Amplifikat nicht direkt in den Shuttle-Vektor klonieren ließ (wahrscheinlich waren die Überhänge zu kurz, als dass *Sal* I effektiv hätte schneiden können). Der Shuttle-Vektor und das Plasmid pKAP298 enthielten nun jeweils eine *Bam* HI-Schnittstelle. Nun wurde versucht die beiden, mit *Bam* HI geschnittenen, Teile des zukünftigen Expressionsvektors zu ligieren und in *Saccharomyces* zu klonieren. Dies misslang leider (mögliche Gründe siehe Hauptteil der Arbeit 4.3). Wäre die Konstruktion des Expressionsvektors gelungen, hätte man versuchen können, das fertige Konstrukt in *Caedibacter taeniospiralis* zu klonieren (z.B. mittels Partikel-Gun). Anschließend hätte man überprüfen können, ob das Bakterium das Cry4B-Kristallprotein exprimiert.

PCR-Longrun-Amplifikation des *cry4B*-Gens aus dem Plasmid pHT611

Das Plasmid pHT611 enthält das *cry4B*-Gen mit seinen regulatorischen Sequenzen aus einem Plasmid von *Bacillus thuringiensis israelensis* (DELECLUSE ET AL., 1993). Dieses Gen sollte, mit seinem Promotor, aus dem Plasmid amplifiziert werden.

Als Primer wurden die Oligonucleotide Cry4BF und Cry4BR verwendet (siehe Hauptteil der Arbeit 2.1.12). Beide Oligonucleotide enthalten als Überhänge *Sal* I - Schnittstellen, der Primer Cry4BF zusätzlich noch eine *Bam* HI - Schnittstelle. Als thermostabile DNA-Polymerase wurde die *Pwo*-Polymerase der Firma Peqlab verwendet, da sie bei hohen Temperaturen sehr stabil ist und außerdem noch eine proof-reading Eigenschaft besitzt, was die Fehlerrate bei der Amplifikation deutlich herabsetzt. Bei einer Größe des zu amplifizierenden Produkts von 3700 Bp ist dies dringend nötig. Allerdings arbeitet diese Polymerase viel langsamer als die übliche *Taq*-Polymerase.

Der PCR-Longrun-Ansatz wurde folgendermaßen angesetzt:

- 1 µl Plasmid pHT611-Lösung (100 ng)
- 5 µl 10 x PCR-Puffer (Promega)
- 8 µl MgSO₄-Puffer (von Peqlab, Endkonzentration 4 mM Mg SO₄)
- 4 µl dNTP-Mix (Konzentration 2 mM pro dNTP)
- 1 µl Primer Cry4BF (100 pMol)

- 1 µl Primer Cry4BR (100 pMol)
- 0,5 µl *Pwo*-DNA-Polymerase (0,5 units)
- 29,5 µl Aqua_{bidest}
50 µl Gesamtansatz

Die Longrun-PCR wurde mit folgendem Programm durchgeführt:

- Denaturierung bei 94 °C für 2 min
- Denaturierung bei 94 °C für 30 s
- Annealing bei 48 °C für 30 s
- Elongation bei 72 °C für 7 min 30 s

Die letzten 3 Schritte wurden 32 mal wiederholt.

Danach wurde pro Ansatz 1 Unit *Taq*-Polymerase hinzugegeben und für 5 min bei 72 °C inkubiert, um zu gewährleisten, dass an die 3'-Enden aller amplifizierten Fragmente ein deoxyAdenosin-Überhang hinzugefügt wurde. Dies war für die spätere Klonierung in den TOPO-Vektor wichtig. Die PCR-Amplifikation des *cry4B*-Gens verlief problemlos, es waren allerdings einige Ansätze zur Optimierung der MgSO₄-Konzentration nötig. Das PCR-Produkt wurde auf ein 1 % Agarose-Gel aufgetragen und einer Elektrophorese unterzogen (3,7 V/cm, bis untere Markerbande aus dem Gel herausgelaufen war).

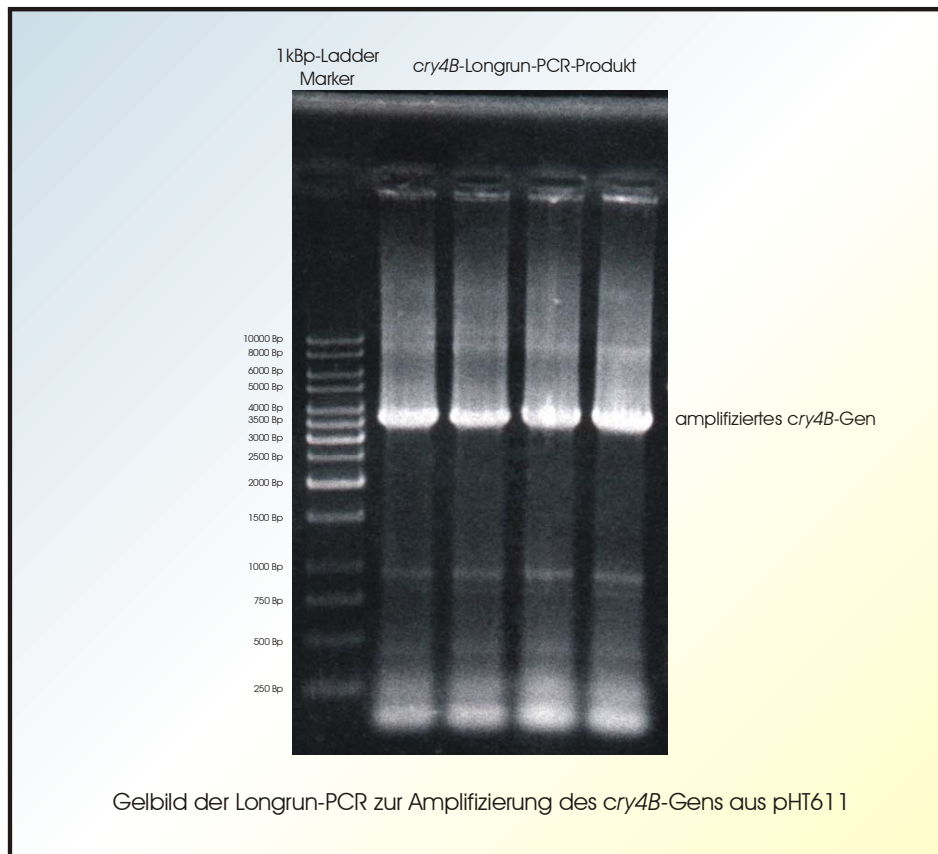


Abb. Anhang 1: Gelbild der Longrun-PCR mit amplifiziertem *cry4B*-Gen

Die Bande bei ca. 3,7 kB entspricht dem *cry4B*-Gen. Gel 1 %, 3,7 V/cm, 2 h.

Aus dem Gel wurde das amplifizierte Fragment mittels des "Nucleospin 2 in 1 -Kits" von Machery&Nagel extrahiert und aufgereinigt. So konnten im besten Fall 35 ng/ μ l des Amplifikats in 30 μ l 10mM Tris/HCl (=1050 ng) isoliert werden

Klonierung des amplifizierten *cry4B*-Gens in den TOPO-Vektor

Das Prinzip der Klonierung mit TOPO XL PCR Kloning Kit der Firma Invitrogen ist folgendes (aus der Anleitung zum Kit):

Der mit dem Kit gelieferte, linearisierte, Vektor pCR-XL-TOPO besitzt an seinen Enden deoxyThymidin-Überhänge, die einem PCR-Produkt mit deoxyAdenosin-Überhängen eine Anlagerung ermöglichen. Gleichzeitig ist das Enzym Topoisomerase I, aus dem Virus *Vaccinia*, kovalent mit dem Vektor verbunden. Die Topoisomerase schneidet das Phosphodiester-Backbone eines DNA-Stranges nach der spezifischen Sequenz 5'-CCCTT (Shuman, 1991). Die Energie des aufgebrochenen Backbones wird, durch die kovalente Bindung des 3' Phosphat-Restes des geschnittenen Stranges, an den Tyrosyl-Rest (Tyr-274)

der Topoisomerase I konserviert. Die Phospho-Tyrosyl-Bindung zwischen Enzym und Vektor kann dann durch das 5' Hydroxyl-Ende des PCR-Produktes angegriffen werden, wodurch sich die Reaktion umkehrt und die Topoisomerase freigesetzt wird. Der PCR-Strang wird dadurch kovalent an den TOPO-Vektor gebunden.

Das Ligations-Produkt kann dann per Elektroporation in *E. coli* subkloniert werden. Die Selektion erfolgt über die Unterbrechung des für *E. coli* tödlichen Gens *ccdB* (BERNARD ET AL., 1994) durch das Insert. Ohne Insert würde das Gen im Vektor exprimiert werden und so für den Tod der *E. coli*-Zellen sorgen. Alle lebensfähigen Transformanten enthalten also mit großer Wahrscheinlichkeit ein Insert im Vektor. Außerdem enthält der Vektor noch ein Kanamycin-Resistenzgen, das zur Selektion zwischen transformierten und nicht transformierten Zellen benutzt wird.

Die Durchführung im Einzelnen:

- 4 µl des PCR-Amplifikats des *cry4B*-Gens
- 1 µl des pCR-XL-TOPO Vektors

wurden bei 25 °C für 5 min inkubiert und anschließend sofort 1 µl 6X TOPO Cloning Stop Solution hinzugegeben.

0,5 µl des Ligations-Ansatzes wurden zur Elektroporation eines Ansatzes der mitgelieferten elektrokompenten TOP10 - *E. coli* Zellen verwendet (Genotyp siehe Hauptteil der Arbeit 2.1.10). Die Puls-Stärke wurde auf 1,6 kV eingestellt, die Pulsdauer betrug 3,7 ms. Nach dem Puls wurde sofort 900 µl LB-Medium hinzugegeben. Der transformierte Ansatz wurde 45 min bei 37 °C inkubiert und danach in Aliquots à 100 µl auf selektive LB-Agar-Platten mit Kanamycin (50 µg/µl) ausplattiert. Nach 24-stündiger Inkubation waren pro Platte mindestens 100-200 Kolonien gewachsen. Die Ligations- und Transformationseffektivität war also sehr hoch. Von 10 der vielen hundert Kolonien wurden Übernacht-Kulturen angelegt, aus diesen dann das rekombinante Plasmid isoliert und per Restriktionsverdau mit *Sal* I + *Xho* I und anschließender Agarose-Gel-Elektrophorese auf das richtige Insert überprüft. In 9 von 10 Fällen war das richtige Insert eingebaut worden. Von den Transformanten mit richtigem Insert wurden Glycerinkulturen angelegt.

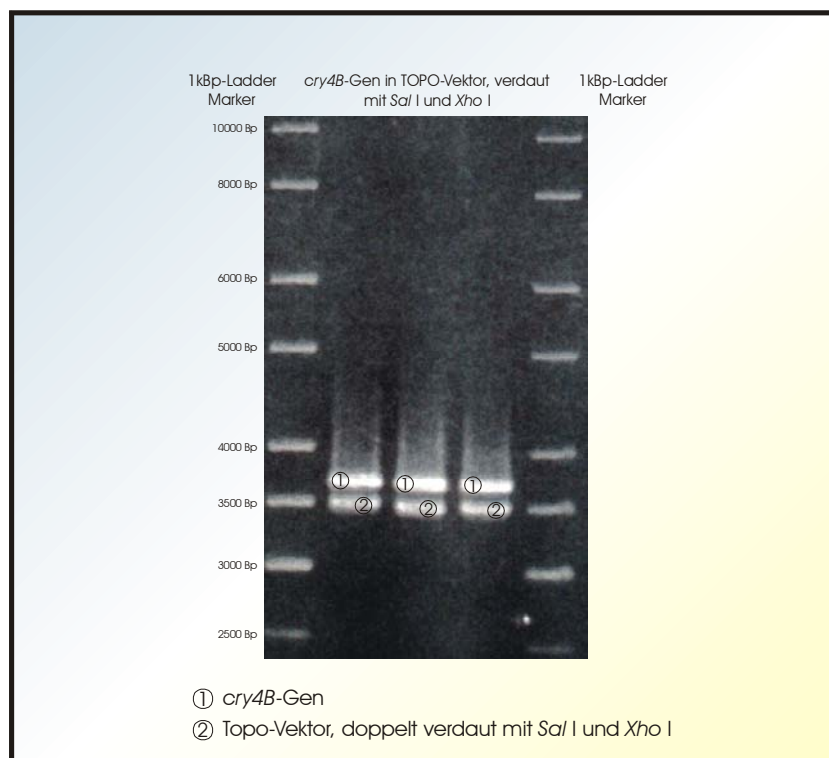


Abb. Anhang 2: Das *cry4B*-Gen im TOPO-Vektor, verdaut mit *Sal* I und *Xho* I

Die mit 1 bezeichneten Banden sind das amplifizierte *cry4B*-Gen, die mit 2 bezeichneten Banden entsprechen dem verdauten TOPO-Vektor. Gel 0,7 %, 3 V/cm, 5 h.

Umklonierung des *cry4B*-Gens aus dem TOPO-Vektor in den Shuttle-Vektor pNF2

Von einer *E. coli* Kolonie, die den rekombinanten Vektor mit *cry4B*-Insert enthielt, wurde eine 40 ml LB-Medium Kultur angelegt und das Plasmid isoliert. Dieses wurde dann mit *Sal* I und *Xho* I geschnitten, um die ähnlich großen Fragmente besser trennen zu können. Der Ansatz wurde einer Gel-Elektrophorese unterzogen und die *cry4B*-Bande aus dem Gel extrahiert. Gleichzeitig wurde der Shuttle-Vektor pNF2 in *E. coli* amplifiziert und isoliert. Das Isolat wurde ebenfalls mit *Sal* I geschnitten und mit CIAP dephosphoryliert. Der verwendete pNF2-Vektor (ATCC Nr. 37187) basiert auf dem Shuttle-Vektor Yep24. Er besitzt ein 2 μ -Replikon zur Replikation in der Hefe und ein pBR322 Replikations-Ursprung für die Replikation in *E. coli*. Als Selektionsmarker in *E. coli* trägt er Ampicillin- und Kanamycin-Resistenzgene, zur Selektion in *S. cerevisiae* kann ein *Ura*⁺-Gen und eine Resistenzgen gegen das Antibiotikum G418 benutzt werden. Der Vektor ist 10,3 kB groß und besitzt eine *Bam* HI-Schnittstelle, die von zwei *Sal* I-Schnittstellen flankiert wird. Er wurde von der Deutschen Sammlung von Mikroorganismen und Zellkulturen (Braunschweig) bezogen (Nr. 4862).

Aus beiden geschnittenen DNA-Fragmenten wurde ein Ligationsansatz zusammenpipettiert:

Das Verhältnis Vektor/Insert sollte ~1:3 betragen. Der Vektor war 10,3 kB groß, das Insert 3700 Bp

- 16 µl *cry4B*-Gen-Isolat (~160 ng)
- 1 µl pNF2-Vektor (~172 ng)
- 2 µl 10xT4-Ligase-Puffer
- 1 µl T4-Ligase (5 units)
- 20 µl Gesamtansatz

Der Ligationsansatz wurde für 16 h bei 4 °C inkubiert und dann jeweils 1 µl zur Elektroporation benutzt. Als Puls-Stärke wurde 1,8 kV gewählt, die durchschnittliche Pulsdauer mit selbst hergestellten elektrokompetenten TOP10 *E. coli* Zellen betrug etwa 4,7 ms. Die Transformationsansätze wurden in 500 µl LB-Medium für 45 min bei 37 °C inkubiert und auf selektive LB-Agarplatten (50 µg/ml Kanamycin) ausplattiert. Nach 24 h waren Kolonien gewachsen, von denen dann 5 ml Übernacht-Kulturen in LB-Medium angelegt wurden. Aus diesen wurde dann das pNF2-Konstrukt isoliert und mittels Restriktionsverdau mit *Sal* I und anschließender Elektrophorese auf das richtige Insert kontrolliert. Um den richtigen Klon zu finden, mussten über 150 Kolonien gescreent werden.

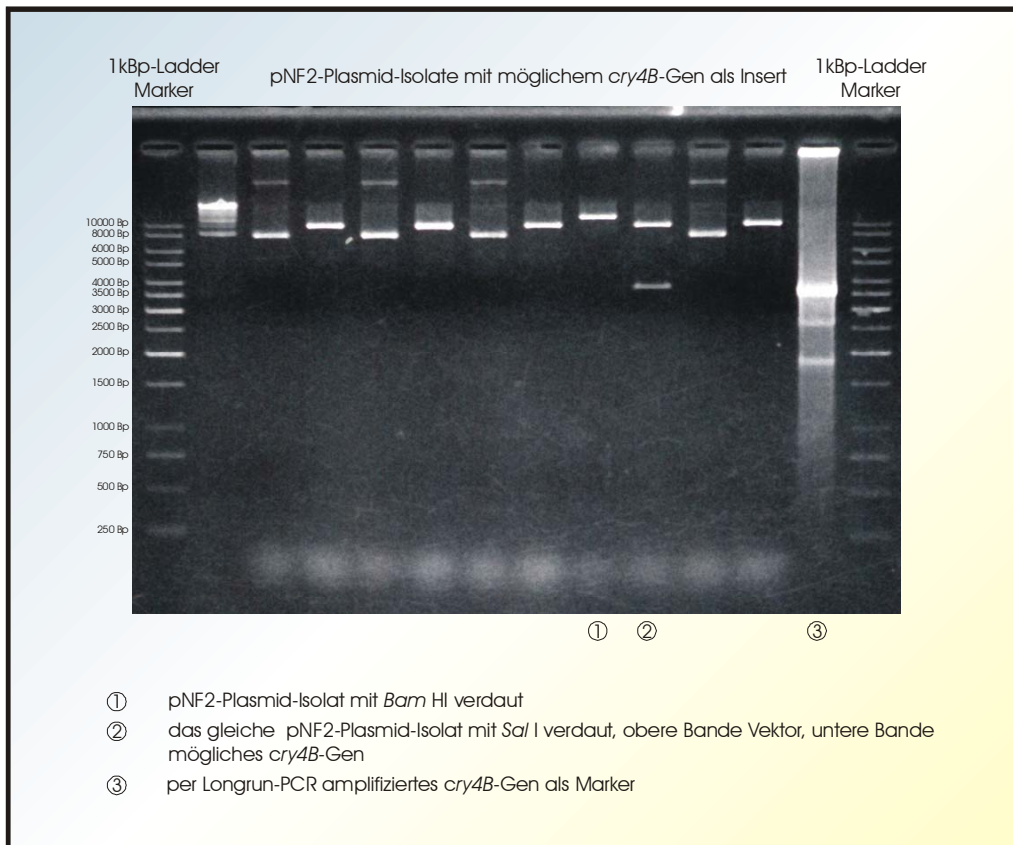


Abb. Anhang 3: Das *cry4B*-Gen im pNF2-Vektor, verdaut mit *Bam* HI und *Sal* I

In der mit 1 bezeichneten Lane ist ein pNF2-Isolat aufgetragen, das mit *Bam* HI verdaut wurde. In Lane 2 wurde das gleiche Isolat mit *Sal* I verdaut. Die obere Bande entspricht der pNF2-Vektorbande, die untere dem Insert. Das Insert weist die gleiche Größe auf, wie das in Lane 3 aufgetragene, per PCR amplifizierte, *cry4B*-Gen. Gel 1 %, 3,5 V/cm, 3 h.

Um letzte Gewissheit zu haben, ob wirklich das erwünschte Insert (*cry4B*-Gen) in den Vektor eingebaut worden ist, wurde mit dem Plasmid-Isolat des potentiell richtigen Klon eine Long-Run-PCR angesetzt (gleiche Bedingungen wie bei der PCR-Amplifikation des *cry4B*-Gens aus dem Vektor pHT611). Nach einer Gel-Elektrophorese wurden die Bandengrößen der Produkte zwischen Kontrolle (PCR mit pHT611) und zu testendem Plasmid (pNF2+X) verglichen. Danach konnte man sicher sein, den Klon mit dem richtigen Insert zu haben.

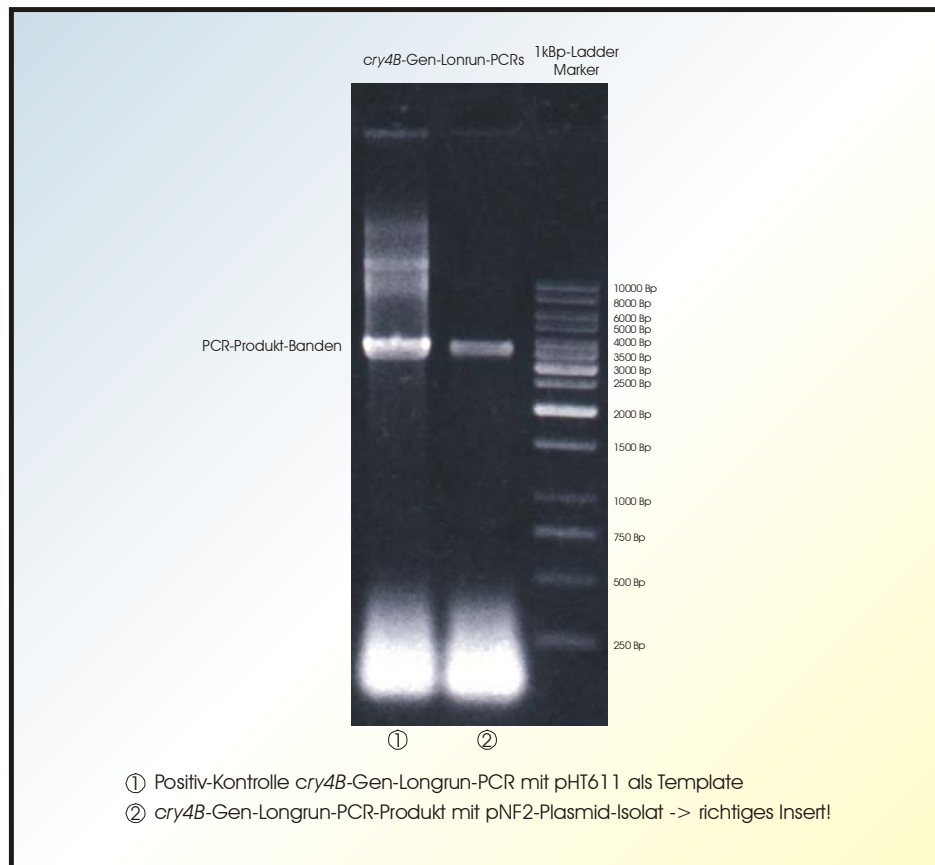


Abb. Anhang 4: PCR-Kontrolle, ob der Vektor pNF2 das erwünschte Insert (*cry4B*-Gen) trägt

Lane 1 enthält die Positivkontrolle mit pHT611 als PCR-Template, Lane 2 das PCR-Produkt mit dem rekombinanten pNF2-Vektor als Template. Beide Produktbanden sind identisch, also hatte pNF2 mit Sicherheit das *cry4B*-Gen als Insert eingebaut. Gel 1 %, 4,5 V/cm, 2 h.

Man hatte nun einen pNF2-Shuttle-Vektor mit *cry4B*-Insert, der genau eine *Bam* HI - Schnittstelle aufwies (durch die Primerwahl bei der PCR-Amplifikation des *cry4B*-Gens). Nun konnte versucht werden, das Plasmid pKAP298 (das ebenfalls genau eine *Bam* HI - Schnittstelle besitzt) in den Vektor zu ligieren und das Konstrukt in *Saccharomyces cerevisiae* zu klonieren.

Der Versuch der "Endmontage" des *Cry4B*-Expressionsvektors in *Saccharomyces cerevisiae*

Der zur Transformation verwendete Hefe-Stamm war *S. cerevisiae* GG100-14D (dankend erhalten von Dr. K. Eisleben, Universität Saarbrücken). Er zeichnet sich durch seine sehr gute Eignung für die Transformation per Elektroporation aus. Sein Genotyp ist: *Mato*, *ura3-52*, *his3*, *trp1*, *pho3*, *pho5*

Das pNF2/*cry4B*-Vektor-Konstrukt wurde zunächst in *E. coli* amplifiziert, isoliert, danach mit *Bam* HI verdaut und dephosphoryliert. Das zu ligierende Plasmid pKAP298 wurde, wie im Hauptteil der Arbeit beschrieben, aus *Caedibacter taeniospiralis* isoliert und mit *Bam* HI verdaut. Die Reinheit des Plasmid-Isolats wurde mittels Gel-Elektrophorese überprüft. Es konnten nur Isolate höchster Reinheit verwendet werden, die möglichst keinen "Schmier" von kleineren DNA-Fragmenten aufwiesen, da diese Fragmente, aufgrund ihrer geringeren Länge, mit höherer Wahrscheinlichkeit in den Vektor eingebaut worden wären.

Mit dem Vektor und dem Plasmid pKAP298 wurde dann ein Ligations-Ansatz erstellt. Idealerweise sollte das Verhältnis Vektor/Insert 1:1 betragen. Der Vektor hatte eine Größe von 10,3 kB, das Insert von ~49 kBp.

Typischer Ligations-Ansatz:

- 10 µl pKAP298-Isolat (~950 ng)
- 4 µl pNF2-Vektor (~260 ng)
- 2 µl 10xT4-Ligase-Puffer
- 3 µl Aqua_{bidest}
- 1 µl T4-Ligase (5 units)
- 20 µl Gesamtansatz

Der Ligations-Ansatz wurde 16 h bei 4 °C inkubiert.

Nun wurden elektrokompente Zellen des Hefestammes GG100-14D hergestellt. Dazu ließ man die Zellen in LB-Medium mit 2,5% Glucose bei 25 °C so lange wachsen, bis die Kultur eine OD₆₀₀ von ~0,8 erreicht hatte. Pro Elektroporations-Ansatz wurden 2 ml der Hefesuspension abzentrifugiert, 3 mal gründlich mit 1 M Sorbitol-Lösung gewaschen und in 80 µl 1 M Sorbitol resuspendiert. Den kompetenten Zellen wurde pro Ansatz 1 µl des Ligations-Ansatzes hinzugegeben und der Ansatz in eine Elektroporations-Küvette überführt. Diese wurde für 5 min auf Eis inkubiert. Die Elektroporation fand in dem Gene-Pulser II der Firma Bio-Rad statt. Die ermittelten, optimalen Einstellungen waren 1,25 kV, Widerstand 400 Ω und eine Kapazität von 25 µF. Der Spannungspuls dauerte zwischen 7 und 8,5 ms. Die transformierten Zellen eines Ansatzes wurden 1 h in 1 ml LB-Medium bei Raumtemperatur inkubiert, abzentrifugiert, in 1 M Sorbitol resuspendiert und dann auf einer Drop-Out-Agar-Platte mit 2,5% Glucose ausplattiert. Hefekolonien erschienen bei Raumtemperatur nach 3-5 Tagen.

Das verwendete Drop-Out-Medium ist eine synthetische Nährlösung ohne Uracil. Da der verwendete Hefe-Stamm Uracil aufgrund einer Mutation nicht selbst produzieren kann, würde er in diesem Medium nicht wachsen. Der Transformations-Vektor pNF2 kodiert aber für ein Protein, das der transformierten Hefe-Zelle wieder erlaubt, Uracil selbst zu produzieren. Die Zelle kann also in dem Uracil-freien Medium wachsen. Zusätzlich wurde dem Medium noch 50 µg/ml Kanamycin zugegeben, um das Überwachsen der Hefe-Zellen mit Prokaryoten zu verhindern.

Die gewachsenen Hefe-Kolonien wurden einzeln in Plastikröhrchen mit je 2 ml Drop-Out-Medium mit 2,5% Glucose überführt und 24 h, unter starkem Schütteln, bei Raumtemperatur inkubiert.

Die Überprüfung der einzelnen Transformanten auf das richtige Insert erfolgte auf 2 Arten:

Zunächst wurde versucht mittels Multiplex-PCR das richtige Insert zu identifizieren. Es wurde mit den isolierten rekombinanten Plasmiden eine PCR durchgeführt, deren Ansatz 2 Primerpaare enthielt. Das erste Primerpaar passte auf die R-Körper-kodierende Region des Plasmids pKAP298 (siehe Hauptteil der Arbeit 2.2.1) und hätte, bei einem Einbau des Plasmids als Insert, eine Bande bei ~800 Bp ergeben. Das andere Primerpaar amplifizierte eine ~600 Bp große Sequenz aus dem Vektor pNF2. Auf diese Weise konnte kontrolliert werden ob die PCR funktioniert hatte, da die pNF2-Bande in allen PCR-Produkten vorkommen musste.

Nun zu den einzelnen Schritten:

Um Ressourcen (und damit Geld zu sparen), wurden immer "(5 bis 10) in one"-Ansätze gemacht. Wurde ein Multi-Ansatz positiv getestet, konnten dann die einzelnen Transformanten des Gemisches untersucht werden.

5-10 der transformierten Hefe-Kulturen in Drop-Out-Medium wurden in einem 50 ml Falcon-Tube vereinigt und gründlich geschüttelt. 2 ml des Gemisches wurde mit 1 unit Zymolyase (ICN, Ohio, USA) vermischt und 30-60 min bei 33 °C inkubiert, bis sich Spheroblasten gebildet hatten. Wichtig war, dass sich die ursprünglichen Hefekulturen noch in der Phase exponentiellen Wachstums befunden hatten, da dann die größte Menge des Plasmids isoliert werden konnte. Wenn die Zellen der Suspension begonnen hatten, sich fädig/zäh zu aggregieren, wurde das Gemisch abzentrifugiert und das Zell-Pellet in 200 ml Resuspensions-Puffer QP1 (Quiagen Plasmid Midi Kit) resuspendiert. Die Zellen wurden dann mit 200 µl Lysis-Puffer QP2 versetzt und 5 min bei Raumtemperatur inkubiert. Danach wurden 200 ml

Neutralisierungs-Puffer QP3 hinzugegeben, gemischt und das Reaktionsgefäß für 15 min auf Höchstgeschwindigkeit zentrifugiert. Der Überstand wurde mit 1 Volumen Isopropanol präzipitiert (30 min / -20 °C) und mit 70% Ethanol gewaschen. Die DNA wurde dann in 20 µl Aqua_{bidest} gelöst und konnte als Template für die nachfolgende Multiplex-PCR benutzt werden.

Multiplex-PCR-Ansatz:

Für die Amplifizierung des pNF2-Stückes wurden die Primer YEP24F und YEP24R benutzt (siehe Hauptteil 2.1.12). Die Konzentration dieser Primer im PCR-Ansatz wurde um den Faktor 10 niedriger gewählt als die Konzentration der pKAP-Primer. Die entstehende Bande sollte nur dokumentieren, ob die PCR überhaupt funktioniert hatte. Die zur Überprüfung auf pKAP298 als Insert verwendeten Primer waren die schon im Hauptteil der Arbeit benutzen pkap476F und pkap1264R.

- 1 µl Hefe-Plasmid-Template
- 5 µl 10 x PCR-Puffer (Promega)
- 5 µl MgCl₂-Puffer (Promega, Endkonzentration 2,5 mM MgCl₂)
- 4 µl dNTP-Mix (Konzentration 2 mM pro dNTP)
- 1 µl Primer pkap476F (=100 pMol)
- 1 µl Primer pkap1264R (=100 pMol)
- 0,1 µl Primer YEP24F (=10 pMol)
- 0,1 µl Primer YEP24R (=10 pMol)
- 0,25 µl *Taq*-DNA-Polymerase (~1 unit)
- 32,5 µl Aqua_{bidest}
- 50 µl Gesamtansatz

Die PCR wurde im Thermocycler nach folgendem Programm durchgeführt:

- Denaturierung 30 s bei 94 °C
- Primer-Annealing 30 s bei 50 °C
- Primer-Extension 1 min bei 72 °C

Diese 3 Schritte wurden 35 mal wiederholt. Das PCR-Produkt wurde auf einem Agarose-Gel elektrophoretisch getrennt, gefärbt und ausgewertet. War nur die pNF2-Bande zu sehen, konnte man davon ausgehen, dass die PCR funktioniert hatte, aber kein pKAP298 als Insert

eingebaut worden war. Konnte man die pKAP298 Bande identifizieren trug möglicherweise einer der Transformanten das Plasmid als Insert.

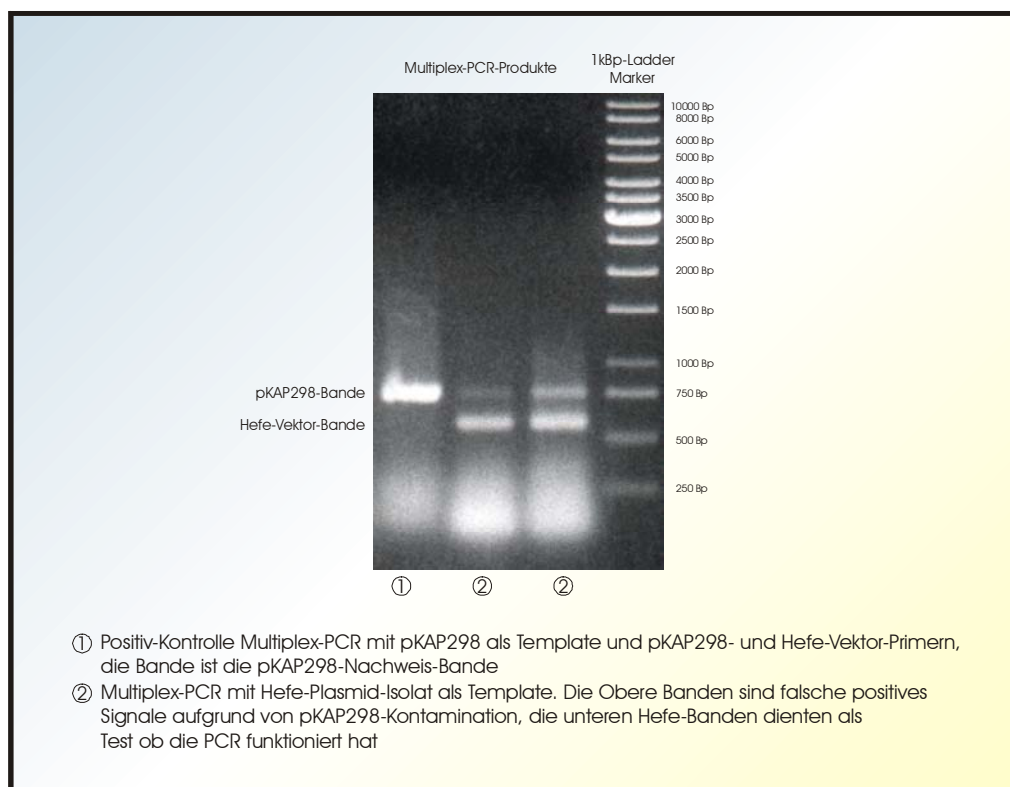


Abb. Anhang 5: Multiplex-PCR zur Überprüfung der Hefe-Klone auf pKAP298-Insert

Lane 1 ist die Positivkontrolle mit pKAP298 als PCR-Template. Die mit 2 bezeichneten Lanes sind die Multiplex-PCR Produkte mit Hefe-Plasmid-Isolat als Template. Die unteren Banden zeigen jeweils an, dass die PCRs funktioniert haben, die oberen, auf pKAP298 hinweisenden Banden, sind falsche positive Signale aufgrund von pKAP298-Kontamination. Gel 1 %, 3,5 V/cm, 3 h.

Ob ein Hefe-Transformant wirklich das Plasmid pKAP298 als Insert trug, wurde durch eine Plasmid-Isolation aus dem entsprechenden Transformanten mit anschließendem Restriktionsverdau und gelelektrophoretischer Analyse überprüft. Um das Ergebnis vorweg zu nehmen: kein Transformant hatte pKAP298 als Insert. Die positiven Signale kamen wohl aufgrund einer Kontamination mit pKAP298 zustande, die auch durch gründliches Waschen der Gerätschaften mit DNaseI-haltigem Wasser und Neuansatz sämtlicher Medien nicht entfernt werden konnte. Deshalb wurde zur weiteren Insert-Kontrolle ein anderes Verfahren gewählt.

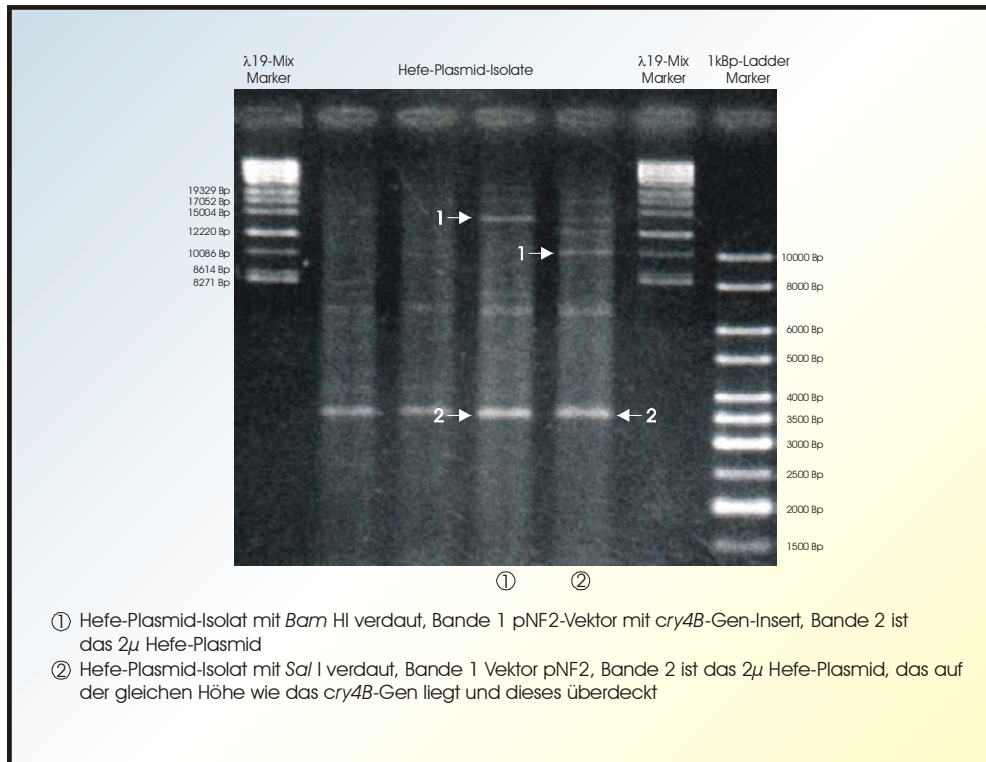


Abb. Anhang 6: Hefe-Plasmid-Isolate zur Überprüfung auf pKAP298-Insert

In Lane 1 ist ein mit *Bam* HI geschnittenes Hefe-Plasmid-Isolat aufgetragen. Bande 1 entspricht dem pNF2-Vektor mit *cry4B*-Insert (ca. 14 kB), Bande 2 ist das 2 μ Hefe-Plasmid. In Lane 2 ist das gleiche Hefe-Plasmid-Isolat aufgetragen, diesmal mit *Sal* I verdaut. Bande 1 entspricht dem Vektor pNF2 (ca. 10,3 kB), Bande 2 ist das *cry4B*-Gen (ca. 3,7 kB), das auf gleicher Höhe wie das 2 μ Hefe-Plasmid liegt und von diesem überdeckt wird. Gel 1 %, 3 V/cm, 3 h.

Die Hefe-Plasmid-Isolation lief ähnlich ab, wie die Template-Gewinnung für die Multiplex-PCR. Die zu überprüfenden Transformanten wurden in 10 ml Drop-Out-Medium mit 2,5% Glucose bis zu einer OD600 von 0,8 wachsen gelassen. Dann wurden 50 units Zymolyase hinzugegeben und 30-60 min bei 33 °C inkubiert. Nachdem sich Spheroblasten gebildet hatten und die Hefe-Zellen begannen fädig zu aggregieren, wurde der Ansatz abzentrifugiert und in 3 ml Resuspensions-Puffer QP1 resuspendiert. Danach wurden 3 ml Lysis-Puffer QP2 zugegeben und das Gemisch für 5 min lysiert. Anschließend wurde der Ansatz mit 3 ml Neutralisierungs-Puffer QP3 neutralisiert und für 30 min bei Höchstgeschwindigkeit abzentrifugiert. Der Überstand wurde wieder mit Isopropanol gefällt und mit 70% Ethanol gewaschen. Die DNA wurde in 100 μ l Aqua_{bidest} gelöst. Das Plasmid-Isolat wurde danach zum einen mit *Bam* HI, zum anderen mit *Sal* I verdaut, elektrophoretisch getrennt und das Gel auf das mögliche pKAP298-Insert untersucht.

Da sich, wie gesagt, kein Transformant mit dem richtigen Insert ermitteln ließ, also die positiven Signale in der PCR auf pKAP298-Kontamination zurückzuführen sein mussten,

wurde als alternative Kontrollmethode die Hybridisierung des Hefe-Plasmid-Isolats mit markierten Sonden gewählt. Zunächst wurden (wie im Hauptteil der Arbeit beschrieben) 2 Digoxigenin-markierte Sonden hergestellt. Zum einen wurde ein PCR-Produkt, mit pNF2 als Template und den Primern YEP24F und YEP24R, markiert (PCR-Bedingungen waren die gleichen wie bei der Multiplex-PCR). Zum anderen wurde das Plasmid pKAP298 selbst markiert. Es wurden aus den 2 ml Transformanten-Kulturen "(5 bis 10) in one" Hefe-Plasmid-Isolate nach der oben beschriebenen Methode hergestellt. Die DNA wurde jedoch in 100 µl Aqua_{bidest} gelöst. Je 20 µl der Isolate wurden nebeneinander auf eine Nylon-Membran pipettiert. Als Positivkontrolle kamen noch 5 µl pKAP298-Isolat (~300ng) hinzu, als Negativkontrolle 5 µl pNF2-Lösung (~300ng). Die DNA wurde anschließend unter UV-Licht mit 200 mJ/cm² auf der Membran fixiert. Danach wurde die Membran (wie im Hauptteil der Arbeit beschrieben) zunächst mit der markierten pKAP298-Sonde hybridisiert und detektiert. Lieferten die Hefe-Plasmid-Isolate keine positiven Signale (also IMMER), wurde die Membran gestrippt und mit der pNF2-Sonde hybridisiert und detektiert, um zu überprüfen, ob überhaupt Hefe-Plasmid in den aufgetragenen Isolaten vorhanden war.

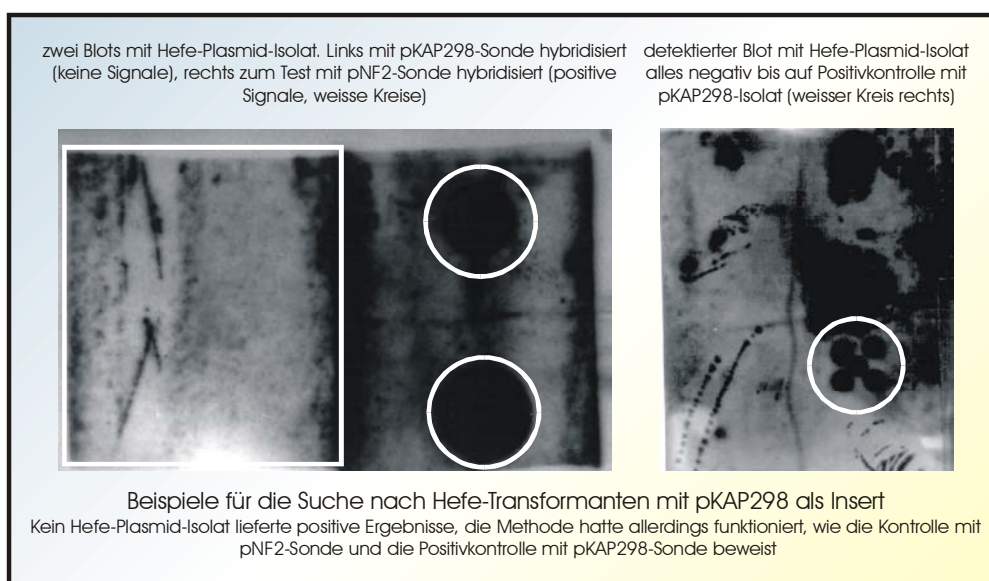


Abb. Anhang 7: detektierte Blots mit Hefe-Plasmid-Isolaten

Mit pKAP298-Sonden ergaben sich (außer bei der Positivkontrolle) keine positiven Signale. Die positiven Signale bei der Hybridisierung mit pNF2-Sonde zeigt allerdings, dass zumindest die Methode funktioniert hat.

Obwohl insgesamt über 2500 Klone mittels der beiden Test-Verfahren auf das richtige Insert überprüft wurden, konnte kein positiver Klon gefunden werden. Dies könnte zunächst daran liegen, dass die Wahrscheinlichkeit gegen mich war: ein 49 kB großes DNA-Fragment in einen 14 kB großen Vektor zu ligieren ist nun einmal relativ unwahrscheinlich - wenngleich

auch nicht unmöglich. Alle methodischen Schritte wurden mit größtmöglicher Sorgfalt durchgeführt. Besonderen Wert wurde auf die gründliche Dephosphorylierung des Vektors pNF2 gelegt, um mögliche Religierung des Vektor während der Ligation, und damit einen hohen Transformanten-Hintergrund, zu verhindern. Ebenfalls möglich ist, dass durch das Plasmid pKAP298, wie in *E. coli*, auch in der Hefe (obwohl Eukaryot) "toxische" Gene exprimiert werden oder stabile Sekundärstrukturen die Replikation verhindern (siehe Hauptteil der Arbeit 4.3), so dass eine Klonierung des Plasmids *per se* unmöglich ist.

Lebenslauf

Persönliche Angaben

Name	Jörn Jeblick
Geburtsdatum	14.10.1972
Geburtsort	Enkenbach-Alsenborn
Familienstand	ledig
Staatsangehörigkeit	deutsch

Bildungsgang

1978-1982	Grundschule in Morlautern
1982-1991	Albert-Schweitzer-Gymnasium in Kaiserslautern Abschluss: Allgemeine Hochschulreife
1991-1992	Studium der Chemie an der TU Kaiserslautern
1992-2000	Studium der Biologie an der TU Kaiserslautern Diplomarbeit in der Abteilung Ökologie Abschluss: Diplom-Biologe
2000-2001	Landesgraduiertenstipendium zur Promotion in der Abteilung Ökologie der TU Kaiserslautern
2001-2003	wissenschaftlicher Mitarbeiter zur Promotion in der Abteilung Ökologie der TU Kaiserslautern

Hiermit versichere ich, die vorliegende Arbeit in der Abteilung Ökologie der Technischen Universität Kaiserslautern selbständig durchgeführt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben.

Kaiserslautern, den _____

(Jörn Jeblick)

**Sequenzanalyse, Transkriptionsaktivität und Untersuchungen
zur Evolution des Plasmids pKAP298 aus dem intrazellulären
Parasiten *Caedibacter taeniospiralis***

Dem Fachbereich Biologie der Universität Kaiserslautern zur Erlangung des akademischen
Grades "Doktor der Naturwissenschaften" eingereichte

Dissertation

vorgelegt von
Jörn Jeblick

Vorsitzender: Prof. Dr. H.J. Schmidt
Erster Berichterstatter: PD Dr. J. Kusch
Zweiter Berichterstatter: Prof. Dr. J. Cullum

Lehrbereich Ökologie der Technischen Universität Kaiserslautern, November 2003
