

**Biosynthese kapsulärer Polysaccharide in**  
***Sinorhizobium meliloti***

**Vorgelegt von Diplom-Ingenieur**  
**Guido Franz Epple**  
**aus Paderborn**

**Von der Fakultät III - Prozesswissenschaften**  
**der Technischen Universität Berlin**  
**zur Erlangung des akademischen Grades**  
**Doktor der Ingenieurwissenschaften**  
**- Dr. -Ing. -**

**genehmigte Dissertation**

**Promotionsausschuss:**

**Vorsitzender: Prof. Dipl.-Ing. Dr. Ulf Stahl**  
**Berichter: Prof. Dr. Ulrich Szewzyk**  
**Berichter: Prof. Dr. Otto Geiger**

**Tag der wissenschaftlichen Aussprache: 20. Mai 2005**

**Berlin 2005**  
**D 83**

## Abkürzungen

AA	Acrylamid
Abb.	Abbildung
ACP	Acyl-Carrier-Protein(e)
AcpP	konstitutives ACP
Ap	Ampicillin
APS	Ammoniumpersulfat
AS	Aminosäure
bidest.	bidestilliert
bp	Basenpaar(e)
Bq	Becquerel
BSA	Rinderserumalbumin
C	Celsius
C%	Gewichtsprozentanteil von <i>N,N</i> -Methylenbisacrylamid an der Gesamtmonomermasse
Cb	Carbenicillin
Cm	Chloramphenicol
CoA	Coenzym A
Cpm	Impulse pro Minute (Counts per minute)
d	Tag
Da	Dalton
DC	Dünnschichtchromatografie
DEAE	Diethylaminoethyl
DNA	Desoxyribonukleinsäure
dNTP	Desoxynucleosidtriphosphat
DTE	1,4-Dithioerithritol
DTT	1,4-Dithiothreitol
EDTA	Ethylendiamintetraessigsäure

## Abkürzungen, Fortsetzung

EGTA	Ethylenglycol-bis-aminoethylether- <i>N,N,N,N</i> -tetraessigsäure
EPS	Exopolysaccharid
EtOH	Ethanol
g	Gramm
h	Stunde
HAc	Essigsäure
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranosid
k	Kilo ( $10^3$ )
<i>kan</i>	Kanamycin/Neomycin-vermittelnde Resistenzkassette
kb	Kilobase
KDO	3-Deoxy-D-manno-2-octulosonsäure
Km	Kanamycin
konz.	Konzentriert
KPS	Kapsuläre Polysaccharide
l	Liter
LPS	Lipopolysaccharid
M	Mega ( $10^6$ )
MeOH	Methanol
MG	Molekulargewicht
min	Minute
ml	Milliliter, $10^{-3}$ l
MM	Minimalmedium
NBT	Nitroblautetrazolium
<i>neo</i>	Kanamycin/Neomycin-vermittelnde Resistenzkassette
nm	Nanometer
Nod-Faktoren	Lipo-chitin-Oligosaccharid-Signalmoleküle
OD <sub>xxx</sub>	Optische Dichte bei xxx nm

## Abkürzungen, Fortsetzung

ORF	Offener Leserahmen (open reading frame)
p.A.	pro analysi
PAGE	Polyacrylamidgelelektrophorese
PCR	Polymerase-Kettenreaktion (polymerase chain reaction)
psi	pounds per square inch
RNA	Ribonukleinsäure
RNase	Ribonuklease
rpm	Umdrehungen pro Minute (revolutions per minute)
RT	Raumtemperatur
s	Sekunden
SDS	Natriumdodecylsulfat
Sm	Streptomycin
<i>spec</i>	Spectinomycin-vermittelnde Resistenzkassette
t	Zeit
T%	Gewichtsprozente von Acrylamid an der Gesamtmasse der Lösung
Tab.	Tabelle
Tc	Tetracyclin
TCA	Trichloressigsäure
TEMED	<i>N, N, N', N'</i> -Tetramethylethyldiamin
Tris	Tris(hydroxymethyl)aminomethan
TY	Trypton/Hefe-Extrakt (tryptone yeast extract)
U	Unit
UV	Ultraviolett
v/v	Volumen pro Volumen
Vol.	Volumen
WT	Wildtyp
w/v	Gewicht pro Volumen

## Abkürzungen, Fortsetzung

xg	Mehrfaches der Erdbeschleunigung g ( $g = 9,81 \text{ m/s}^2$ )
X-Gal	5-Brom-4-Chlor-3-indolyl- $\beta$ -D-galactopyranosid
$\mu\text{l}$	Mikroliter, $10^{-6} \text{ l}$

---

## INHALTSVERZEICHNIS

<b>1</b>	<b>EINLEITUNG.....</b>	<b>1</b>
1.1	DIE SYMBIONTISCHE STICKSTOFF-FIXIERUNG. ....	1
1.2	BEDEUTUNG VON POLYSACCHARIDEN FÜR DIE SYMBIONTISCHE STICKSTOFF-FIXIERUNG	2
1.2.1	<i>Cyclische Glucane</i> .....	2
1.2.2	<i>Lipopolysaccharide (LPS)</i> .....	2
1.2.3	<i>Exopolysaccharide (EPS)</i> .....	3
1.2.4	<i>Kapsuläre Polysaccharide (KPS)</i> .....	4
1.3	BIOSYNTHESE KAPSULÄRER POLYSACCHARIDE.....	5
1.3.1	<i>Escherichia coli</i> .....	5
1.3.2	<i>Sinorhizobium meliloti</i> .....	5
1.4	POLYKETIDSYNTHASEN .....	7
1.5	ZIELSETZUNG DIESER ARBEIT .....	8
<b>2</b>	<b>MATERIAL UND METHODEN .....</b>	<b>9</b>
2.1	BAKTERIENSTÄMME .....	9
2.2	PLASMIDE UND KOSMIDE .....	10
2.3	PUFFER UND LÖSUNGEN .....	14
2.4	KULTURMEDIEN UND ZELLZÜCHTUNG .....	16
2.4.1	<i>Medien und Anzucht</i> .....	16
2.4.2	<i>Medienzusätze</i> .....	17
2.5	LAGERUNG VON BAKTERIENKULTUREN .....	18
2.6	ENZYME UND MARKER .....	18
2.7	RADIOCHEMIKALIEN.....	18
2.8	REAKTIONSKITS .....	19
2.9	SYNTHETISCHE OLIGONUKLEOTIDE.....	19
2.10	DNA-ARBEITEN.....	21
2.10.1	<i>DNA-Isolierung</i> .....	21
2.10.1.1	Isolierung genomischer DNA .....	21
2.10.1.2	Plasmidpräparation aus <i>E. coli</i> und <i>S. meliloti</i> .....	21
2.10.2	<i>Reinigung und Fällung von DNA</i> .....	22
2.10.2.1	Reinigung von DNA mit einer Silikamatrix .....	22
2.10.3	<i>Quantifizierung und Reinheitsbestimmung von DNA</i> .....	22
2.10.4	<i>Enzymatische Modifizierung von DNA</i> .....	22
2.10.4.1	Restriktionshydrolyse .....	22
2.10.4.2	Dephosphorylierung mit alkalischer Phosphatase .....	23
2.10.4.3	Auffüllen überhängender Enden mit Klenow-Fragment.....	23
2.10.4.4	Entfernen überhängender Enden mit T4 DNA-Polymerase.....	23

2.10.4.5	Ligation.....	24
2.10.5	<i>Agarosegel-Elektrophorese von DNA</i> .....	24
2.10.6	<i>Transformation von E. coli</i> .....	24
2.10.6.1	Transformation nach der CaCl <sub>2</sub> -Methode .....	24
2.10.6.1.1	Herstellung kompetenter Zellen .....	24
2.10.6.1.2	Transformation kompetenter Zellen.....	25
2.10.6.2	Transformation durch Elektroporation .....	25
2.10.6.2.1	Herstellung kompetenter Zellen .....	25
2.10.6.2.2	Elektroporation .....	25
2.10.7	<i>Transfer mobilisierbarer Vektoren durch Konjugation</i> .....	26
2.10.7.1	Drei-Punkt-Kreuzung .....	26
2.10.7.2	"Knock-out-Mutagenese".....	27
2.10.8	<i>Amplifikation von DNA mittels der Polymerase-Kettenreaktion (PCR)</i> .....	28
2.10.9	<i>DNA-Sequenzierung</i> .....	29
2.10.9.1	Sequenzierungsreaktion.....	29
2.10.9.2	Elektrophorese .....	29
2.10.9.3	Entwicklung des Röntgenfilms .....	30
2.10.10	<i>DNA-Analyse Software und Internetdienste</i> .....	31
2.11	PROTEINCHEMISCHE ARBEITEN .....	31
2.11.1	<i>Native PAGE von Proteinen</i> .....	31
2.11.2	<i>SDS-PAGE</i> .....	32
2.11.3	<i>Tricin-Gele</i> .....	33
2.11.4	<i>Färben von Proteingelen</i> .....	34
2.11.4.1	Coomassiefärbung .....	34
2.11.4.2	Silberfärbung.....	34
2.11.5	<i>Blotten von Proteingelen</i> .....	35
2.11.6	<i>Blotten von Proteingelen nach radioaktiver Markierung von Proteinen.</i> .....	35
2.11.7	<i>Immudetektion von immobilisierten Proteinen (Western Blot)</i> .....	36
2.11.8	<i>Proteinbestimmung</i> .....	36
2.11.9	<i>Zellaufschluss</i> .....	37
2.11.9.1	French-Press-Aufschluss .....	37
2.11.9.2	Zellaufschluss mit Lysozym .....	37
2.11.10	<i>Aufreinigung des RkpF-Proteins</i> .....	37
2.11.10.1	Isopropanolfällung.....	37
2.11.10.2	Chromatografie an DEAE 52-Cellulose .....	38
2.11.10.3	Aufkonzentrieren von RkpF-Pools.....	38
2.11.10.4	Gelfiltration von RkpF.....	38
2.11.10.5	<i>In vitro</i> -Markierung von RkpF durch radioaktives CoA .....	39
2.11.10.6	Konversion von Apo- zu Holo-RkpF.....	39
2.11.10.7	Nachweis von RkpF in Mutanten von <i>Sinorhizobium meliloti</i> AK631 .....	40

2.11.11	<i>In vivo</i> -Markierung von Proteinen	40
2.11.11.1	Markierung mit $\beta$ -[ $^3\text{H}$ ]Alanin	40
2.11.11.2	Markierung mit [ $^{35}\text{S}$ ]Sulfat	41
2.11.12	Autoradiografie	41
2.12	UNTERSUCHUNG VON POLYSACCHARIDEN	41
2.12.1	Isolierung von Polysacchariden	41
2.12.2	Gelelektrophorese von Polysacchariden	42
2.12.3	Färben von Polysacchariden in DOC-PAGE-Gelen	43
2.13	ISOLIERUNG VON METABOLITEN	43
2.13.1	Extraktion von Metaboliten	43
2.13.1.1	Extraktion durch Umesterung	43
2.13.1.2	Extraktion mit Ethylacetat	44
2.13.2	Dünnschichtchromatografische Auftrennung von Extrakten	44
2.13.3	[ $^{14}\text{C}$ ]Acetat-Markierung	44
<b>3</b>	<b>ERGEBNISSE</b>	<b>46</b>
3.1	KLONIERUNGEN UND SEQUENZIERUNGEN	46
3.1.1	Klonierungen zur heterologen Expression	46
3.1.1.1	Klonierung von <i>rkpABCDEF</i> zur heterologen Expression	46
3.1.1.1.1	Klonierungsstrategie	46
3.1.1.1.2	Klonierung von <i>rkpA'</i>	47
3.1.1.1.3	Klonierung von <i>rkpE'F</i>	47
3.1.1.1.4	Klonierung von <i>rkpA<sub>367</sub>BCDEF</i> in pET9a	49
3.1.1.1.5	Klonierung von <i>rkpA<sub>262</sub>BCDEF</i> in pET9a	50
3.1.1.1.6	Klonierung von <i>rkpA<sub>114</sub>BCDEF</i> in pET9a	50
3.1.1.2	Klonierung einzelner ORFs für die heterologe Expression	50
3.1.1.2.1	Klonierung von <i>rkpA</i>	50
3.1.1.2.2	Klonierung von <i>rkpF<sub>L</sub></i>	51
3.1.2	Analyse der ORFs in der Komplementationsgruppe I	51
3.1.2.1	Sequenzabweichungen zum Genbankeintrag X64131	51
3.1.2.2	<i>In silico</i> -Analyse der Komplementationsgruppe I	52
3.1.2.2.1	Analyse der Genbanksequenz X64131	52
3.1.2.2.2	Analyse der korrigierten Sequenz	53
3.1.3	Analyse des komplementierenden Kosmides pPP428	54
3.1.3.1	Subklonierung und Sequenzierung von pPP428	54
3.1.3.2	Analyse der pPP428-Sequenz	55
3.1.4	"Knock-out-Mutagenese"	58
3.1.4.1	Klonierungsstrategie zur Inaktivierung von <i>rkpU</i>	58
3.1.4.2	Inaktivierung von <i>kpsF</i>	58
3.2	PROTEINCHEMISCHE ARBEITEN	59

3.2.1	<i>Untersuchungen zu RkpF</i> .....	59
3.2.1.1	Heterologe Expression von RkpF.....	59
3.2.1.1.1	Aufreinigung von RkpF.....	59
3.2.1.1.2	<i>In vivo</i> -Markierung von RkpF.....	61
3.2.1.1.3	<i>In vitro</i> -Konversion von Apo- zu Holo-RkpF.....	62
3.2.1.2	Nachweis von RkpF in <i>Sinorhizobium meliloti</i> .....	63
3.2.2	<i>Heterologe Expression von RkpA</i> .....	64
3.2.3	<i>Heterologe Expression von RkpABCDEF</i> .....	65
3.2.3.1	Nachweis eines RkpF-enthaltenden Proteinkomplexes.....	66
3.2.3.2	Biochemische Untersuchung des Proteinkomplexes .....	67
3.2.3.2.1	Einfluss der DTT-Konzentration auf den RkpF-enthaltenden Proteinkomplex .....	67
3.2.3.2.2	Einfluss von Harnstoff auf den RkpABCDEF-Proteinkomplex .....	68
3.2.3.3	Expression von RkpABCDEF bei verschiedenen Kultivierungsbedingungen .....	70
3.2.3.4	<i>In vivo</i> -Markierung des Proteinkomplexes.....	71
3.2.3.4.1	Markierung mit $\beta$ -[ <sup>3</sup> H]Alanin.....	71
3.2.3.4.2	Markierung mit [ <sup>35</sup> S]-Sulfat .....	72
3.2.3.5	Expression des RkpABCDEF-Proteinkomplexes in <i>Sinorhizobium meliloti</i> .....	73
3.3	PRODUKT DER PROTEINE RKPABCDEF.....	75
3.3.1	<i>Markierung mit [1-<sup>14</sup>C]Acetat</i> .....	75
3.3.2	<i>Untersuchung verschiedener Extraktionsmethoden</i> .....	75
3.3.3	<i>Dünnschichtchromatografische Auftrennung</i> .....	76
3.3.4	<i>Induktionsabhängigkeit der Produktbildung</i> .....	77
3.3.5	<i>Einfluss des RkpA-Startkodons auf die Produktbildung</i> .....	79
3.4	ANALYSE VON POLYSACCHARIDEN .....	80
3.4.1	<i>Analyse der kapsulären Polysaccharide</i> .....	80
3.4.1.1	Komplementation der Insertionsmutanten von <i>S. meliloti</i> AK631 .....	81
3.4.1.1.1	Mutante 649 ( <i>rkpU::Tn5</i> ).....	81
3.4.1.1.2	Mutante 644 ( <i>rkpC::Tn5</i> ) .....	82
3.4.1.1.3	Mutante 590 ( <i>rkpG::Tn5</i> ).....	83
3.4.1.1.4	Mutante 551 ( <i>rkpI::Tn5</i> ) .....	84
3.4.1.1.5	Mutante 633 ( <i>rkpJ::Tn5</i> ).....	85
<b>4</b>	<b>DISKUSSION</b> .....	<b>87</b>
4.1	UNTERSUCHUNGEN ZU RKP.....	87
4.1.1	<i>RkpF ist ein Acyl-Carrier-Protein</i> .....	87
4.1.2	<i>RkpF in Sinorhizobium meliloti AK631</i> .....	88
4.2	KLONIERUNG DER GENE RKPABCDEF .....	88
4.3	HETEROLOGE EXPRESSION DER GENE RKPABCDEF .....	91
4.4	NACHWEIS DES RKPABCDEF-PROTEINKOMPLEX IN <i>S. MELILOTI</i> .....	93

4.5	VERGLEICH VON RKPABCDEF AUS <i>SINORHIZOBIUM MELILOTI</i> AK631 MIT RkPA AUS DEM SEQUENZIERTEN GENOM VON <i>S. MELILOTI</i> 1021.....	93
4.6	PRODUKT DER PROTEINE RKPABCDEF.....	94
4.7	KOMPLEMENTATION VON RHIZOBIELLEN MUTANTEN.....	95
4.7.1	<i>Neueinteilung der Komplementationsgruppen der rkp-1-Region</i> .....	96
<b>5</b>	<b>ZUSAMMENFASSUNG</b> .....	<b>98</b>
<b>6</b>	<b>LITERATURVERZEICHNIS</b> .....	<b>100</b>
<b>7</b>	<b>ANHANG</b> .....	<b>111</b>
7.1	GENBANKEINTRAG X64131 .....	111
7.2	GENBANKEINTRAG RkPU .....	118
7.3	GENBANKEINTRAG KPSF .....	121
7.4	SEQUENZEN UND RESTRIKTIONSCHNITTSTELLEN VON PPP428.....	123
7.5	VERGLEICH VON RKPABCDEF VON <i>SINORHIZOBIUM MELILOTI</i> AK631 MIT RkPA AUS DEM SEQUENZIERTEN GENOM VON <i>S. MELILOTI</i> 1021.....	137
7.5.1	<i>Proteinalignment</i> .....	137
7.5.2	<i>Nucleotidalignment</i> .....	140
7.6	DANKSAGUNG .....	146
7.7	EIGENE VERÖFFENTLICHUNGEN .....	147
7.7.1	<i>Teile der Arbeit wurden bereits veröffentlicht</i> .....	147
7.7.2	<i>Sonstige Veröffentlichungen</i> .....	147
7.8	LEBENS LAUF .....	147

---

## Abbildungsverzeichnis

Abbildung 1: Konstruktion eines <i>rkpABCDEF</i> -enthaltenden Überexpressionsvektors.....	48
Abbildung 2: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1005.....	49
Abbildung 3: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1014.....	49
Abbildung 4: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1041.....	52
Abbildung 5: Ausschnitt eines BLASTx-Ergebnisses. ....	52
Abbildung 6: Mögliche Offene Leserahmen im <i>rkpABCDEF</i> -Gencluster.....	53
Abbildung 7: Analyse der Offenen Leserahmen und der bekannten Gene in pPP428.....	56
Abbildung 8: Überexpression und Reinigung des RkpF-Proteins.....	61
Abbildung 9: Zellextrakte von mit $\beta$ -[ <sup>3</sup> H]Alanin kultivierten Zellen.....	62
Abbildung 10: Zeitabhängige Umsetzung von Apo- zu Holo-RkpF durch die gereinigte Holo-ACP-Synthase AcpS aus <i>E. coli</i> . ....	63
Abbildung 11: Western-Blot zum Nachweis von RkpF in verschiedenen Mutanten von <i>Sinorhizobium meliloti</i> AK631.....	64
Abbildung 12: SDS-PAGE-Analyse der heterologen RkpA-Expression.....	65
Abbildung 13. Western-Blot Nachweis eines RkpF-enthaltenden Proteinkomplexes.....	67
Abbildung 14: DTT-abhängige Stabilität und Immunreaktivität des RkpF-enthaltenden Proteinkomplexes. ....	68
Abbildung 15: Untersuchung des Proteinkomplexes in Gegenwart von 5 M Harnstoff. ....	69
Abbildung 16: Expression des RkpABCDEF-Komplexes bei verschiedenen Kultivierungstemperaturen. ....	71

Abbildung 17: Western-Blot und Autoradiogramm von mit $\beta$ -[ $^3\text{H}$ ]Alanin <i>in vivo</i> -markierten Proteinen.....	72
Abbildung 18: Western-Blot-Analyse hochkonzentrierter Rohextrakte aus <i>S. meliloti</i> AK631. ....	74
Abbildung 19: Autoradiogramm der dünn-schichtchromatografisch aufgetrennten [1- $^{14}\text{C}$ ]Acetat-markierten Ethylacetatextrakte. ....	76
Abbildung 20: Untersuchung der induktionsabhängigen Bildung des RkpABCDEF-Proteinkomplexes und dessen vermutlichen Produktes.....	78
Abbildung 21: Autoradiogramm der dünn-schichtchromatografisch aufgetrennten [1- $^{14}\text{C}$ ]Acetat-markierten Ethylacetatextrakte. ....	80
Abbildung 22: Polysaccharide aus exponentiell wachsenden Kulturen von <i>S. meliloti</i> 649 ( <i>rkpU::Tn5</i> ).....	82
Abbildung 23: Polysaccharide aus exponentiell wachsenden Kulturen von <i>S. meliloti</i> 644 ( <i>rkpC::Tn5</i> ).....	83
Abbildung 24: Polysaccharide aus exponentiell wachsenden Kulturen von <i>S. meliloti</i> 590 ( <i>rkpU::Tn5</i> ).....	84
Abbildung 25: Polysaccharide aus exponentiell wachsenden Kulturen von <i>S. meliloti</i> 551 ( <i>rkpl::Tn5</i> ). ....	85
Abbildung 26: Klonierung des <i>rkpABCDEF</i> -Genklusters aus <i>Sinorhizobium meliloti</i> AK631 für die heterologe Expression in <i>E. coli</i> . ....	90
Abbildung 27: Organisation der <i>rkp-1</i> -Region. ....	97

---

# 1 Einleitung

## 1.1 Die symbiontische Stickstoff-Fixierung.

Rhizobien können mit Pflanzen der Leguminosenfamilie eine symbiontische Wechselbeziehung eingehen. Für die Ausbildung von neuen Pflanzenorganen, den Wurzelknöllchen, ist eine molekulare Kommunikation zwischen der Wirtspflanze und dem mikrosymbiontischen Bakterium notwendig. Von den Wurzeln der Wirtspflanze werden Flavonoid-Moleküle ausgeschieden, welche die Induktion einer Reihe von Genen im Bakterium induzieren (Long, 1989; Long, 1996). Diese Gene werden Knöllchenbildungs- oder *nod*-Gene genannt. Die *nod*-Genprodukte synthetisieren spezifische Lipochitin-Oligosaccharid-Signalmoleküle (Nod-Faktoren) (zusammengefasst in D'Haese & Holsters, 2002). Diese Signalmoleküle können bereits ohne direkten Kontakt zwischen den Symbiosepartnern Zellteilung in der Rindenschicht der Wirtspflanzenwurzel hervorrufen. Nach der erfolgten Anheftung der Bakterien wandern diese durch Infektionsschläuche in die Wurzeln ein und werden später von Pflanzenzellen durch Endocytose aufgenommen (Kijne, 1992). Im späten Stadium der Symbioseentwicklung differenzieren sich die Bakterien zu Bakteroiden, die dann mit Hilfe einer Nitrogenase im entwickelten Knöllchen Stickstoff fixieren (Fisher & Long, 1992; van Rhijn & Vanderleyden, 1995). Grundsätzlich unterscheidet man zwei Arten der Wurzelknöllchenbildung. Bei den undeterminierten Wurzelknöllchen findet eine fortschreitende Zellteilung statt und es finden sich alle Differenzierungsstadien in den einzelnen Wurzelknöllchen. Im Falle der determinierten Wurzelknöllchenbildung sind dagegen alle Zellen eines Knöllchens in einem ähnlichen Differenzierungszustand (Vasse *et al.*, 1990). Eine Besonderheit des Symbioseprozesses ist die so genannte Wirtsspezifität. Dies bedeutet, dass ein bestimmtes Bakterium der *Rhizobiaceae* nur mit einer oder wenigen Leguminosen eine Symbiose ausbilden kann. Diese Wirtsspezifität ist vor allem durch die Nod-Faktoren determiniert, jedoch spielen auch andere Moleküle eine wichtige Rolle (Broughton *et al.*, 2000).

## **1.2 Bedeutung von Polysacchariden für die symbiotische Stickstoff-Fixierung**

Während der mittleren Entwicklungsphase der Symbiose spielen verschiedene rhizobielle Polysaccharide eine wichtige Rolle.

### **1.2.1 Cyclische Glucane**

Die aus *Sinorhizobium meliloti* isolierten cyclischen Glucane bestehen aus ca. 20  $\beta$ -(1,2)-verknüpften Glucose-Bausteinen (Breedveld *et al.*, 1995). Die cyclischen Glucane sind hauptsächlich im Periplasma vorhanden, wo sie den Bakterien bei der Anpassung an veränderte osmotische Verhältnisse helfen. Für die cyclische Glucane von *Bradyrhizobium japonicum*, die  $\beta$ -(1,3)- und  $\beta$ -(1,6)- verknüpft sind, konnte gezeigt werden, dass sie in großen Mengen in Bakteroiden vorhanden sind. Eine Mutation in der Biosynthese der cyclischen Glucane führt bei der Symbiose von *Bradyrhizobium japonicum* mit seiner Wirtspflanze, der Sojabohne, zu ineffektiven Wurzelknöllchen (Dunlap *et al.*, 1996).

### **1.2.2 Lipopolysaccharide (LPS)**

Lipopolysaccharide (LPS) sind Polysaccharide, die über einen Lipidteil in der äusseren Membran der Gram-negativen Bakterien verankert sind. LPS haben einen charakteristischen Aufbau, der sich im Wesentlichen in drei Bestandteile gliedern lässt: 1. Lipid A, 2. Kernzone (engl. Core) welche in innere und äußere Kernzone unterteilt werden kann und 3. O-Antigen-Region. Der innere Teil der LPS, das Lipid A, besteht aus einer Glycolipidschicht, die auch die äußere Schicht der Zellmembran bildet. Diese Glycolipidschicht besteht aus einem Trisaccharid, das mit Hydroxyfettsäuren substituiert ist. In den LPS der Rhizobien wurden mehrere langkettige Fettsäuren entdeckt, die an der ( $\omega$ -1)-Position eine Hydroxy-funktion tragen (Bhat *et al.*, 1991b; Gil-Serrano *et al.*,

1994; Hollingsworth & Carlson, 1989). In Lipid A-Präparaten aller Vertreter der Familie *Rhizobiaceae* wurde 27-Hydroxyoctacosansäure nachgewiesen, deren Nachweis auch als taxonomischer Marker genutzt wird (Bhat *et al.*, 1991a). Ein Bestandteil der Core-Region ist 2-Keto-3-Deoxyoctulosonsäure (KDO). Das Vorhandensein von KDO wurde früher als Indikator für LPS genutzt, jedoch ist KDO auch in den kapsulären Polysacchariden (KPS) enthalten. Daher ist nicht eindeutig, ob Funktionen bei der Symbiose die früher den LPS zugeordnet worden sind, nicht in Wirklichkeit durch die KPS übernommen werden (Dazzo *et al.*, 1991). Jedoch konnte gezeigt werden, dass gereinigte LPS von *S. meliloti* die Abwehrreaktion von kultivierten *Medicago sativa*-Zellen modulieren können (Albus *et al.*, 2001). Daher wird vermutet, dass LPS ebenso, wie Exopolysaccharide (EPS) und KPS, die Abwehrreaktion der Pflanze gegen den Symbionten unterdrücken.

### 1.2.3 Exopolysaccharide (EPS).

Die Exopolysaccharide sind hochmolekulare Heterokomplexe aus Untereinheiten von sieben bis neun Hexosen. Es konnte gezeigt werden, dass Exopolysaccharide nur bei der Ausbildung von undeterminierten Wurzelknöllchen essenziell sind. Bei der Ausbildung einer Symbiose mit determinierten Wurzelknöllchen sind EPS-Mutanten nicht behindert (Hotter & Scott, 1991). Am besten untersucht ist die Rolle der EPS in der Wechselwirkung zwischen Alfalfa und seinem Mikrosymbionten *Sinorhizobium meliloti* Rm1021 (Reinhold *et al.*, 1994; Reuber & Walker, 1993). In diesem System konnte gezeigt werden, dass EPS I, welche auch Succinoglycan genannt werden, für die erfolgreiche Ausbildung des Infektionsschlauches und die Einwanderung der Bakterien in diesen benötigt werden. Die Funktion von *exo*<sup>-</sup>-Mutanten ließ sich durch die externe Zugabe von EPS I wieder herstellen (Battisti *et al.*, 1992). Jedoch waren *exo*<sup>-</sup>-Mutanten nicht komplett in ihrer Symbiosefähigkeit inhibiert und es konnte gezeigt werden, dass andere Polysaccharide wie EPS II oder die kapsulären Polysaccharide für die Funktion des EPS I substituieren können.

### 1.2.4 Kapsuläre Polysaccharide (KPS)

Viele Schritte der Symbiose zwischen den Leguminosen und den Bakterien der *Rhizobiaceae* sind mit pathogenen Interaktionen vergleichbar. Kapsuläre Polysaccharide (KPS) spielen bei Infektionen durch Gram-negative enteropathogene Bakterien eine wichtige Rolle. Sie verhindern, dass die Bakterien durch das Immunsystem des Wirtsorganismus erkannt werden. Im Gegensatz zu den Lipopolysacchariden wirken KPS nur schwach immunisierend. Die EPS-defiziente Mutante AK631 des Stammes *S. meliloti* 41, die einen Defekt im *exoB*-Lokus trägt, war überraschenderweise noch immer in der Lage Stickstoff-fixierende Knöllchen zu bilden (Putnoky *et al.*, 1990). Sie synthetisiert ein saures Polysaccharid, das der Gruppe II K-Antigene (kapsuläre Polysaccharide) aus *Escherichia coli* ähnelt (Reuhs *et al.*, 1993). Wenn AK631 jedoch zusätzliche Mutationen im *rkpZ*-Gen (Reuhs *et al.*, 1995) oder in Genen der *fix-23*-Region (Putnoky *et al.*, 1988; Putnoky *et al.*, 1990) aufweist, so wird das rhizobielle kapsuläre Polysaccharid nicht mehr gebildet; die entsprechenden *S. meliloti*-Mutanten zeigen einen *Fix<sup>-</sup>*-Phänotyp. Rhizobielle kapsuläre Polysaccharide können entweder direkt durch Analyse von Extrakten in Polyacrylamidgelen und anschließender Anfärbung mit Alcian Blau (Reuhs *et al.*, 1993) oder indirekt, über die Vermittlung einer Phagensensitivität nachgewiesen werden. Wenn rhizobielle kapsuläre Polysaccharide nicht mehr gebildet werden, kann auch keine Adsorption des Bakteriophagen 16-3 an den entsprechenden *S. meliloti*-Stamm erfolgen (Putnoky *et al.*, 1990). Des Weiteren scheinen die KPS eine Funktion als Signalsubstanz zu haben, da sie verschiedene Pflanzengene induzieren (Becquart-de Kozak *et al.*, 1997).

### **1.3 Biosynthese kapsulärer Polysaccharide.**

#### **1.3.1 *Escherichia coli***

In *Escherichia coli* gehören die Biosynthese und der Transport von kapsulären Polysacchariden zu den Pathogenitätsfaktoren. Es wird vermutet, dass es einen gemeinsamen genetischen Ursprung der KPS-Genloci von *Haemophilus influenzae*, *E. coli* und *Neisseria meningitis* gibt (Frosch *et al.*, 1991). Die Suche nach Genen, die an der Biosynthese der Gruppe II K-Antigene (kapsuläre Polysaccharide) beteiligt sind, hat bei *E. coli* zur Identifizierung, von 3 konservierten Genregionen geführt (Bliss & Silver, 1996; Boulnois *et al.*, 1987; Whitfield & Paiment, 2003). Eine "Region 2" ist an der Biosynthese des eigentlichen Polysaccharides beteiligt. Eine "Region 3" beinhaltet Gene, die für ABC-Transporter kodieren und vermutlich am Transport der synthetisierten kapsulären Polysaccharide über die zytoplasmatische Membran beteiligt sind. Eine "Region 1" schließlich ist am Transport der kapsulären Polysaccharide zur Zelloberfläche beteiligt. Gene, die an der Biosynthese einer Fettsäure oder eines  $\beta$ -Ketids beteiligt sein könnten, wurden bisher bei den Genen, die für die Biosynthese kapsulärer Polysaccharide in *E. coli* notwendig sind, nicht gefunden. Die *E. coli* Gruppe II K-Antigene, wie auch kapsuläre Polysaccharide aus Meningococcen und *Haemophilus*, sind häufig über eine Phosphodiesterbindung an ein Diacylglycerin geknüpft (Rick & Silver, 1996). Es wird diskutiert, dass diese Lipidsubstitution als Anker des kapsulären Polysaccharides an das Bakterium dienen könnte. Trotz intensiver Suche konnte jedoch bisher kein Lipidanteil, weder ein Phosphatidsäurerest noch ein Fettsäurerest, als Bestandteil der rhizobiellen kapsulären Polysaccharide identifiziert werden (Reuhs *et al.*, 1993).

#### **1.3.2 *Sinorhizobium meliloti***

Die Oberfläche der Bakterien hat einen bedeutenden Einfluss sowohl auf die Ausbildung der Symbiose als auch für die Anheftung von Bakteriophagen. Für den *Sinorhizobium meliloti*-Stamm 41 konnte gezeigt werden, dass in einem

*exoB*-defizienten Hintergrund die Bildung von kapsulären Polysacchariden (KR5-antigen) essenziell für die erfolgreiche Ausbildung einer funktionellen Symbiose ist (Petrovics *et al.*, 1993; Putnoky *et al.*, 1990). Bisher wurden drei Genregionen, die an der Biosynthese und dem Transport der rhizobiellen Kapsulären Polysaccharide beteiligt sind, identifiziert. Die *rkp-2*-Region enthält die beiden Gene *lpsL* und *rkpK*. Für *rkpK* konnte gezeigt werden, dass es für eine UDP-Glycose-Dehydrogenase kodiert, die an der Biosynthese von Glucuronsäure, einem der Bausteine des KR5-Antigenes, beteiligt ist (Kereszt *et al.*, 1998). Die *rkp-3*-Region enthält die sechs stammspezifischen Gene *rkpL*, *rkpM*, *rkpN*, *rkpO*, *rkpP* und *rkpQ*. Es wird vermutet, dass die Produkte dieser Gene an der Biosynthese spezifischer Zucker beteiligt sind (Kiss *et al.*, 2001). Für das Genprodukt von *rkpM* wurde gezeigt, dass es direkt mit dem Protein-H des Bakteriophagen 16-3 interagiert (Putnoky *et al.*, 2004). In der Nachbarschaft dieses Genclusters wurden fünf weitere Gene identifiziert, die an der Biosynthese oder dem Transport der KR5-Antigene beteiligt sein sollen. Während *rkpR*, *rkpT* und *rkpS* diese Funktion nur aufgrund ihrer Homologien zugeordnet wird, konnte für *rkpZ* und *rkpY* gezeigt werden, dass sie an der Biosynthese der KR5-Antigene beteiligt sind (Kiss *et al.*, 2001). Der Stamm *Sinorhizobium meliloti* 1021 besitzt kein *rkpZ*-Gen und im Gegensatz zu *Sinorhizobium meliloti* 41, ist eine *exoB*<sup>-</sup>-Mutante von *Sinorhizobium meliloti* 1021 nicht in der Lage symbiontisch Stickstoff zu fixieren. Wird jedoch das *rkpZ*-Gen aus AK631 in eine *exoB*<sup>-</sup>-Mutante von *Sinorhizobium meliloti* 1021 eingeführt, so kommt es zu einer Suppression des *Fix*<sup>-</sup>-Phänotyp (Reuhs *et al.*, 1995). Die *rkp-1*-Region enthält 10 Gene und diese scheinen an der Biosynthese eines spezifischen Lipid-Ankers des KR5-Antigens beteiligt zu sein (Epple, 1995; Kiss *et al.*, 1997; Petrovics *et al.*, 1993). Mutationen in der *rkp-1*-Region konnten in 4 Komplementationsgruppen (I-IV) eingeteilt werden (Putnoky *et al.*, 1990). Die DNA-Sequenzanalyse der Komplementationsgruppe I der *rkp-1*-Region erlaubte die Bestimmung von 6 offenen Leserahmen (ORFs 1-6) (Petrovics *et al.*, 1993), die inzwischen als *rkpABCDEF*-Gene bezeichnet werden (Reuhs, 1996). Für die davon abgeleiteten Genprodukte konnte ein

hoher Grad an Homologie mit der Fettsäuresynthetase aus der Ratte und eine ähnliche Organisation wie für deren multifunktionalen Enzymdomänen gefunden werden (Petrovics *et al.*, 1993).

Bei verschiedenen Bakterien wurden zwar Phospholipidsubstituenten in den KPS gefunden, spezielle Ketide oder Fettsäuren konnten jedoch bisher nicht nachgewiesen werden (Forsberg & Reuhs, 1997; Frosch *et al.*, 1992; Reuhs *et al.*, 1993); (Reuhs *et al.*, 1998). Deshalb wird vermutet, dass der von den Proteinen RkpABCDEF synthetisierte Acyl- oder  $\beta$ -Ketidrest nur für bestimmte Biosynthese- oder Exportschritte des rhizobiellen kapsulären Polysaccharides essenziell ist, aber von diesem wieder abgespalten wird (Reuhs *et al.*, 1998).

#### **1.4 Polyketidsynthasen**

Bei Polyketidsynthasen werden verschiedene Typen unterschieden. Beim Typ-I sind alle enzymatischen Aktivitäten, die für die Synthese des Polyketids benötigt werden, auf einem Polypeptid vereint. Beim Typ-II werden die einzelnen Reaktionen der Kettenverlängerung durch monofunktionale Proteine ausgeführt. Die wachsende Kohlenstoffkette wird dabei von einem Acyl-Carrier-Protein getragen. Obwohl die Proteine RkpABCDEF nach ihren Gensequenzen einem Typ-II-Polyketidsynthasesystem entsprechen, haben sie bei Homologievergleichen eine viel stärkere Ähnlichkeit zu Typ-I-Polyketidsynthasen. Die größte Ähnlichkeit besteht zu den Modulen 3 und 4 des Erythromycin-produzierenden Systems aus *Saccharopolyspora erythraea* (Bevitt *et al.*, 1992). Der modulare Aufbau der Typ-I-Polyketidsynthasen erlaubt die Kombination verschiedener Module in beliebiger Reihenfolgen und somit die Biosynthese von neuen Polyketiden (Bisang *et al.*, 1999; Eckermann *et al.*, 1998; Pieper *et al.*, 1995; Roberts *et al.*, 1993; Silakowski *et al.*, 1999; Xue *et al.*, 1999).

## **1.5 Zielsetzung dieser Arbeit**

Durch genetische Untersuchungen konnte gezeigt werden, dass die Gene *rkpABCDEF* an der Biosynthese der Kapsulären Polysaccharide (KR5-Antigene) in *Sinorhizobium meliloti* AK631 beteiligt sind (Petrovics *et al.*, 1993). Trotz der Homologien dieser für die KPS-Biosynthese essenziellen Gene zu Fettsäure- und Polyketidsynthasen konnte bisher kein Acyl- oder  $\beta$ -Ketiderest in den KR5-Antigenen nachgewiesen werden (Reuhs *et al.*, 1998). Da die Gene *rkpABCDEF* in *Sinorhizobium meliloti* nur sehr schwach exprimiert werden, soll zur Identifizierung des Produktes der Proteine RkpABCDEF eine heterologe Expression genutzt werden (Petrovics *et al.*, 1993). Für das Protein RkpF konnte bereits gezeigt werden, dass es sich um ein Acyl-Carrier-Protein handelt. Ebenfalls konnte gezeigt werden, dass RkpF in *E. coli* durch Übertragung der prosthetischen Gruppe 4'-Phosphopanthein posttranslational modifiziert wird (Epple, 1995). Diese Modifikation ist notwendig, um eine Synthese des vermuteten  $\beta$ -Ketides am Acyl-Carrier-Protein zu erlauben. Auf Basis dieser Erkenntnisse sollten die Gene der Komplementationsgruppe-I *rkpABCDEF* in einen Überexpressionsvektor kloniert werden und anschließend in *E. coli* heterolog exprimiert werden. Aus diesem heterologen System sollte ein spezifisches Produkt der Proteine RkpABCDEF isoliert und analysiert werden. Zur weiteren Überprüfung der Funktionalität der klonierten Gene sollten diese ebenfalls in ein Plasmid breiter Wirtsspezifität umkloniert werden. Mit diesem Plasmid sollten dann in verschiedenen Tn5-Insertionsmutanten von *Sinorhizobium meliloti* AK631 eine weitere Analyse der KR-5 Antigen-Biosynthese in *Sinorhizobium meliloti* durchgeführt werden.

## 2 Material und Methoden

### 2.1 Bakterienstämme

Die in der vorliegenden Arbeit verwendeten Bakterienstämme sind in Tabelle 1 aufgelistet. Wenn im Folgenden von *Sinorhizobium meliloti* die Rede ist, so handelt es sich dabei immer, falls nicht anders erwähnt, um den Stamm *S. meliloti* Stamm 41 AK631.

Tabelle 1: Bakterienstämme

Stamm	Relevante Eigenschaften	Referenz
<i>Escherichia coli</i>		
DH5 $\alpha$	<i>recA1</i> , $\phi$ 80 <i>lacZ</i> $\Delta$ M15	(Hanahan, 1983)
BL21(DE3)	F <sup>-</sup> , <i>ompT</i> , <i>lacUV5</i>	(Studier <i>et al.</i> , 1990)
OG7001	F <sup>-</sup> , <i>ompT</i> , <i>lacUV5</i> , <i>panD</i>	(Epple <i>et al.</i> , 1998)
J53 R751-pGM2		(Jacoby <i>et al.</i> , 1976)
JM 110		(Yanisch-Perron <i>et al.</i> , 1985)
<i>Sinorhizobium meliloti</i> Stamm 41		
AK631 Wildtyp	<i>exoB</i> , Fix <sup>+</sup>	(Putnoky <i>et al.</i> , 1988)
Sm AK631-Derivate		
649	Tn5-Insertion Upstream von <i>rkpA</i>	(Putnoky <i>et al.</i> , 1990)
644	Tn5-Insertion in <i>rkpC</i>	(Putnoky <i>et al.</i> , 1990)
590	Tn5-Insertion in <i>rkpG</i>	(Putnoky <i>et al.</i> , 1990)
551	Tn5-Insertion in <i>rkpI</i>	(Putnoky <i>et al.</i> , 1990)
633	Tn5-Insertion in <i>rkpJ</i>	(Putnoky <i>et al.</i> , 1990)
PP2756	<i>rkpF::kan</i>	Peter Putnoky
PP2757	<i>rkpF::kan</i>	Peter Putnoky
PP2915	AK631 mit <i>rkpF</i> auf pRK290	
	Orientierung 1,	Peter Putnoky
PP2916	AK631 mit <i>rkpF</i> auf pRK290	

PP2919	Orientierung 2, PP2756 mit <i>rkpF</i> auf pRK290	Peter Putnoky
PP2920	Orientierung 1, PP2756 mit <i>rkpF</i> auf pRK290	Peter Putnoky
	Orientierung 2,	Peter Putnoky

## 2.2 Plasmide und Kosmide

Die während der Doktorarbeit hergestellten und verwendeten Plasmide und Kosmide sind in den Tabellen 2 und 3 aufgelistet.

Tabelle 2: Plasmide, die zur Klonierung verwendet wurden

Plasmid	Relevante Eigenschaften	Referenz
pRK2013	Helfer-Plasmid, Kanamycin-resistent	(Figurski <i>et al.</i> , 1979)
pUC19	Klonierungsvektor, Carbenicillin-resistent	(Yanisch-Perron <i>et al.</i> , 1985)
pBK-CMV	Klonierungsvektor, Kanamycin-resistent	Stratagene, La Jolla, CA
pBSIIKS+	Klonierungsvektor, Carbenicillin-resistent	Stratagene, La Jolla, CA
pET9a	Expressionsvektor, Kanamycin-resistent	(Studier <i>et al.</i> , 1990)
pMP3510	IncP, Expressionsvektor der den Promotor von <i>nodABC</i> enthält, Tetracyclin-resistent	(Spaink <i>et al.</i> , 1995)
pTB3131	neo-Kassette von Tn5 in pUC19	(Kretzschmar <i>et al.</i> , 2001)
pPP428	pLAFR1 mit einem 30 kb großen sinorhizobiellen Fragment, enthält die <i>rkp-1</i> -Region	(Putnoky <i>et al.</i> , 1988)
pBBR1MCS	Klonierungsvektor, Chloramphenicol-resistent	(Kovach <i>et al.</i> , 1994)

pAT389	pBBR1MCS mit einem 10 kb <i>Hind</i> III-Fragment der <i>rkp</i> -1-Region, enthält die Komplementations- gruppe-I	Peter Putnoky
pTB1003	pET9a mit <i>rkpF</i> , kloniert mit <i>Nde</i> I/ <i>Bam</i> HI; <i>rkpF</i> unter der Kontrolle des T7-Promotors (Epple, 1995)	
pG-KJE8	Expression von GroEL, GroEL, DnaK und DnaJ	(Nishihara <i>et al.</i> , 1998)

Tabelle 3: Angefertigte Plasmide

Vektor	Relevante Eigenschaften
pTB1005	pUC19 mit dem 1043 bp großen PCR-Produkt <i>rkpA</i> ´, amplifiziert mit <i>Pfu</i> -Polymerase, kloniert mit <i>Nde</i> I/ <i>Bam</i> HI
pTB1007	pET9a mit dem 1043 bp großen <i>rkpA</i> -enthaltenden <i>Nde</i> I/ <i>Bam</i> HI-Fragment aus pTB1005
pTB1008	pTB1007 mit dem 441 bp großen <i>Kpn</i> I/ <i>Bam</i> HI-Fragment des PCR-Produktes <i>rkpE</i> ´ <i>F</i> ´, amplifiziert mit <i>Pfu</i> -Polymerase, kloniert mit <i>Kpn</i> I/ <i>Bam</i> HI
pTB1011	der mit <i>Sac</i> I/ <i>Xba</i> I aus pMP3510 entnommene <i>nodABC</i> -Promoter wurde in pBBR1MCS eingesetzt
pTB1014	pTB1008 mit dem 419 bp großen <i>Bam</i> HI-Fragment des PCR-Produkt <i>rkpE</i> ´ <i>F</i> ´, amplifiziert mit <i>Pfu</i> -Polymerase, kloniert mit <i>Bam</i> HI
pTB1017	pTB1007 mit dem 840 bp großen <i>Kpn</i> I/ <i>Bgl</i> II-Fragment des PCR-Produkt <i>rkpA</i> (oTBC7/oTBC100) amplifiziert mit <i>Pfu</i> -Polymerase in die <i>Kpn</i> I/ <i>Bam</i> HI-Schnittstellen kloniert; <i>rkpA</i> <sub>367</sub> unter der Kontrolle des T7-Promotors
pTB1019	pUC19 mit einem 915 bp <i>Sma</i> I/ <i>Kpn</i> I-Fragment von pAT389
pTB1025	pET9a mit dem 279 bp großen PCR-Produkt <i>rkpF</i> <sub>L</sub> amplifiziert mit <i>Pfu</i> -Polymerase, kloniert mit <i>Nde</i> I/ <i>Bam</i> HI

- pTB1030 pTB1017 mit dem 778 bp großen *NdeI/KpnI*-Fragment des PCR-Produkt *rkpA*<sub>262</sub>, amplifiziert mit *Pfu*-Polymerase ; *rkpA*<sub>262</sub> unter der Kontrolle des T7-Promotors
- pTB1040 pTB1014 mit dem 778 bp großen *NdeI/KpnI*-Fragment aus pTB1030
- pTB1041 pUC19 mit einem 10178 Kbp *HindIII*-Fragment aus pPP428
- pTB1044 pTB1014 mit dem 5,66 kb *KpnI/BclI*-Fragment aus pTB1041; *rkpA*<sub>367</sub> *BCDEF* unter der Kontrolle des T7-Promotors
- pTB1047 pTB1040 mit dem 5,66 kb *KpnI/BclI*-Fragment aus pTB1041; *rkpA*<sub>262</sub> *BCDEF* unter der Kontrolle des T7-Promotors
- pTB1056 ein *rkpA*<sub>367</sub> *BCDEF* enthaltendes *XbaI/BamHI*-Fragment (aufgefüllt mit Klenow) aus pTB1044 wurde mit *XbaI/HindIII* (aufgefüllt mit Klenow) in pMP3510 kloniert
- pTB1057 ein *rkpA*<sub>262</sub> *BCDEF* enthaltendes *XbaI/BamHI* (aufgefüllt mit Klenow) aus pTB1047 wurde mit *XbaI/HindIII* (aufgefüllt mit Klenow) in pMP3510 kloniert
- pTB1058 pMP3510 mit dem 10178 Kbp *HindIII*-Fragment aus pTB1041; *rkpUABCDEF* unter der Kontrolle des *nodABC*-Promotors
- pTB1063 pUC19 mit einem 462 bp *EcoRI*-Fragment aus pPP428
- pTB1064 pUC19 mit einem 698 bp *EcoRI*-Fragment ent aus pPP428
- pTB1069 pUC19 mit einem 1021 bp *EcoRI*-Fragment aus pPP428
- pTB1073 pUC19 mit einem 1428 bp *EcoRI*-Fragment aus pPP428
- pTB1074 pUC19 mit einem 1338 bp *EcoRI*-Fragment aus pPP428
- pTB1078 pUC19 mit einem 2836 bp *EcoRI*-Fragment aus pPP428
- pTB1081 pUC19 mit einem 7827 bp *EcoRI*-Fragment aus pPP428
- pTB1082 pTB1011 mit einem 7827 bp *EcoRI*-Fragment aus pTB1078 mit der anderen Orientierung als in pTB1081
- pTB1083 pUC19 mit dem 926 bp großen *NdeI/KpnI*-Fragment des PCR-Produkt *rkpA*<sub>114</sub> (oTBC132/oTBC133) amplifiziert mit *Pfu*-Polymerase

- pTB1084 pTB1082 mit *Pst*I linearisiert und in pTB1011 eingesetzt; *rkpGHIJ* unter der Kontrolle des *nodABC*-Promotors
- pTB1085 pTB1017 mit dem 926 bp großen *Nde*I/*Kpn*I-Fragment aus pTB1083; *rkpA*<sub>114</sub> unter der Kontrolle des T7-Promotors
- pTB1090 pTB1044 mit dem 1 kb *Xba*I/*Kpn*I-Fragment aus pTB1085; *rkpA*<sub>114</sub> *BCDEF* unter der Kontrolle des T7-Promotors
- pTB1094 ein *rkpU* enthaltendes 1,7 kb *Sph*I-Fragment aus pTB1041 wurde in pMP3510 eingesetzt, *rkpU* unter der Kontrolle des *nodABC*-Promotors.
- pTB1095 wie pTB1094 nur andere Orientierung des Inserts
- pTB1098 pUC19 mit einem 2874 bp *Eco*RI-Fragment aus pPP428
- pTB1100 pBS mit 9723 bp *Bgl*II/*Xba*I-Fragment aus pPP428
- pTB1103 pTB1100 mit *Xba*I linearisiert und in pTB1011 eingesetzt; *rkpGHJKpsF* unter der Kontrolle des *nodABC*-Promotors
- pTB1111 pUC19 mit einem 3338 bp *Hind*III/*Xba*I-Fragment aus pPP428
- pTB1112 ein *rkpA*<sub>114</sub> *BCDEF* enthaltendes *Xba*I/*Bam*HI (aufgefüllt mit Klenow) aus pTB1090 wurde mit *Xba*I/*Hind*III (aufgefüllt mit Klenow) in pMP3510 kloniert; *rkpABCDEF* unter der Kontrolle des *nodABC*-Promotors
- pTB1115 ein *rkpG* enthaltendes Fragment *Bam*HI/*Nco*I (aufgefüllt mit Klenow) aus pTB1078 wurde mit *Bam*HI/*Eco*RV in pTB1011 kloniert; *rkpG* unter der Kontrolle des *nodABC*-Promotors
- pTB1116 pTB1081 wurde mit *Kpn*I verdaut und relegiert; *rkpGH* unter der Kontrolle des *nodABC*-Promotors
- pTB1120 pTB1111 mit einer Kanamycin-Resistenz-Kassette, eingefügt in die *Agel*-Schnittstelle von *kpsF*, mit entgegengesetzten Transkriptionsrichtungen
- pTB1121 pTB1111 mit einer Kanamycin-Resistenz-Kassette, eingefügt in die *Agel*-Schnittstelle von *kpsF*, mit gleichen Transkriptionsrichtungen
- pTB1123 pTB1121 mit *Hind*III linearisiert und in pMP3510 eingesetzt

### 2.3 Puffer und Lösungen

TE-Puffer:	10 mM Tris/HCl 1 mM EDTA, pH 8
Resuspendierungspuffer: (DNA-Extraktion)	50 mM Tris/HCl 10 mM EDTA, pH 8 400 µg/ml RNAase A
Lysispuffer: (DNA-Extraktion)	0,2 N NaOH, 1% SDS
Neutralisationspuffer: (DNA-Extraktion)	3 M K-Acetat, pH 5,2
TAE-Puffer:	40 mM Tris/Acetat, 1 mM EDTA
TBE-Puffer:	2,5 mM EDTA, 90 mM Borsäure, 90 mM Tris
4 x Laufpuffer: (PAGE)	1,5 M Tris/HCl, pH 8,8
4 x Sammelpuffer: (PAGE)	500 mM Tris/HCl, pH 6,8
2 x Auftragspuffer: (native PAGE)	20% Glycerin (w/v) in 4 x Laufpuffer

---

	2% (v/v) einer gesättigten Bromphenolblaulösung in EtOH
2 x Auftragspuffer: (SDS-PAGE)	20% Glycerin (w/v) 8% (w/v) SDS 2% (v/v) Mercaptoethanol 2% (v/v) einer gesättigten Bromphenolblaulösung in EtOH in 4 x Laufpuffer
Tankpuffer: (PAGE)	25 mM Tris 192 mM Glycin
Tankpuffer: (SDS-PAGE)	25 mM Tris 192 mM Glycin 0,1% (w/v) SDS
Gelpuffer: (Tricin-PAGE)	3 M Tris 0,3% (w/v) SDS pH 8,45 mit HCl eingestellt
Anodenpuffer: (Tricin-PAGE)	0,2 M Tris pH 8,9 mit HCl eingestellt
Kathodenpuffer: (Tricin-PAGE)	0,1 M Tris 0,1 M Tricin 0,1% (w/v) SDS
2 x Auftragspuffer: (Tricin-PAGE)	8% (w/v) SDS 24 % (w/v) Glycerin 100 mM Tris/HCl, pH 6,8

	2% (v/v) einer gesättigten Bromphenolblaulösung in EtOH 2% (v/v) Mercaptoethanol
Blottingpuffer:	192 mM Glycin 20 % (v/v) MeOH 25 mM Tris, ergibt pH 8,8 und wurde nicht eingestellt
PBS:	137 mM NaCl; 2.7 mM KCl; 1.4 mM $\text{KH}_2\text{PO}_4$ ; 4.3 mM $\text{Na}_2\text{HPO}_4 \cdot 7 \text{H}_2\text{O}$ ; pH 7.4

## 2.4 Kulturmedien und Zellzüchtung

### 2.4.1 Medien und Anzucht

Alle Medien wurden direkt nach dem Ansetzen 20 Minuten bei 121°C autoklaviert. Für feste Medien wurde vor dem Autoklavieren 1,8 % Agar hinzugefügt.

*E. coli* wurde auf LB-Medium als Vollmedium oder M9-Medium als Minimalmedium bei 37°C angezogen (Miller, 1972). Als Regenerationsmedium nach der Elektroporation von *E. coli* wurde SOC-Medium benutzt.

*S. meliloti* wurde bei 29°C kultiviert. Als Vollmedium wurde TY-Medium (Beringer, 1974) verwendet.

LB-Medium (1 l)	10 g NaCl 10 g Bacto Trypton 5 g Hefeextrakt
M9-Medium (1 l)	7,5 g $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$

3 g  $\text{KH}_2\text{PO}_4$   
0,5 g NaCl  
1 g  $\text{NH}_4\text{Cl}$   
werden mit  $\text{H}_2\text{O}$  auf 987 ml aufgefüllt  
und autoklaviert, anschließend werden:  
1 ml einer 1 M  $\text{MgSO}_4$ -Lösung  
10 ml einer 1 M Glucose-Lösung  
1 ml einer 1%-igen Thiamin-Lösung  
1 ml einer 0,1 M  $\text{CaCl}_2$ -Lösung  
zugegeben

SOC-Medium (1 l)      20 g Bacto Trypton  
5 g Hefeextrakt  
580 mg NaCl  
190 mg KCl  
1,2 g  $\text{MgSO}_4$   
20 ml einer 1 M Glucose-Lösung  
(getrennt autoklaviert)

TY-Medium (1 l)      5 g Bacto Trypton  
3 g Hefeextrakt  
10 ml einer 0,45 M  $\text{CaCl}_2$ -Lösung  
(getrennt autoklaviert)

#### 2.4.2 Medienzusätze

Die Medienzusätze wurden, falls erforderlich, in den folgenden Konzentrationen zugegeben:

Carbenicillin	100 $\mu\text{g/ml}$
Chloramphenicol	20 $\mu\text{g/ml}$
Gentamicin	60 $\mu\text{g/ml}$

Kanamycin	50 µg/ml
Neomycin	100 µg/ml
Piperazillin	40 µg/ml
Rifampicin	5 µg/ml
Spectinomycin	300 µg/ml
Streptomycin	500 µg/ml
Tetracyclin	2 µg/ml ( <i>S. meliloti</i> ), 20 µg/ml ( <i>E. coli</i> )
X-Gal	40 µg/ml

## 2.5 Lagerung von Bakterienkulturen

*E. coli* wurde in LB mit 15% (v/v) Glycerin, *S. meliloti* in Pepton (6 g/l) mit 15 % Glycerin bei – 80°C gelagert.

## 2.6 Enzyme und Marker

Restriktionsenzyme	Roche Molecular Biochemicals, Mannheim Pharmacia, Lippsale MBI Fermentas, Vilnius
<i>Taq</i> -DNA-Polymerase	MBI Fermentas, Vilnius
<i>Pfu</i> -DNA-Polymerase	Promega, Madison, WI
Phosphatase (CIP)	Roche Molecular Biochemicals, Mannheim
T4 DNA-Ligase	Roche Molecular Biochemicals, Mannheim
λ-DNA	MBI Fermentas, Vilnius

## 2.7 Radiochemikalien

Chemikalie	Aktivität	spez. Radioaktivität
β-[3- <sup>3</sup> H]Alanin	37 MBq/ml	2,22 TBq/mMol
[1- <sup>14</sup> C]Acetat	7,4 MBq/ml	2,2 GBq/mMol
[ <sup>3</sup> H]CoA	35,15 MBq/ml	1,48 TBq/mMol

[ <sup>35</sup> S]Sulfat	37 MBq/ml	> 37 TBq/mMol (carrier-free)
--------------------------	-----------	---------------------------------

Alle Radiochemikalien wurden von Amersham Pharmacia Biotech, Uppsala, Schweden geliefert.

## 2.8 Reaktionskits

Gene DNA Purification System	BioRad, München
QIAGEN Plasmid Kit	Qiagen, Hilden
Sequenase 7-deaza dGTP Sequencing Kit	Amersham Buchler, Braunschweig

## 2.9 Synthetische Oligonukleotide

Alle synthetischen Oligonukleotide wurden von der Firma Metabion, München hergestellt. Die Primer sind jeweils in 5'-3'-Richtung angegeben.

oTBC2	AAA ggA TCC TCA Cgg CTT gCT TAC TTT TTC ATg g	Primer für die Amplifikation von <i>rkpE'</i>
oTBC3	ggA ATA CAT ATg gAC CAg CAg Cag Cgg GTT CTg	Primer für die Amplifikation von <i>rkpA</i> <sub>367</sub> '
oTBC4	CgC CTg TgC Cgg ATC Cgg ATC gCTg	Primer für die Amplifikation von <i>rkpA</i> <sub>367</sub> '
oTBC6	TCA ggC gAT gCg ggC gC	Sequenzierungsprimer für pTB1005
oTBC7	ATC Agg CgA TgC ggg CgC	Primer für die Amplifikation von <i>rkpA</i> <sub>367</sub> ''
oTBC8	ggA gCA ggT ACC gCT TCg TTC ggA AgC ACC	Primer für die Amplifikation von <i>rkpE'</i>
oTBC100	ATA AgA TCT TAT AgT CTC TCg AgC Agg	Primer für die Amplifikation von <i>rkpA</i>
oTBC109	AAC TAT gTC gCC gCA AAC gg	Sequenzierungsprimer für pTB1014

- oTBC111 Cag gAA TAC ATA Tgg gCA CgC CCg  
Primer für die Amplifikation von *rkpA*<sub>262</sub>'
- oTBC113 Cag gAA TAC ATA Tgg ATC Tcg Tcg CCA Tg  
Primer für die Amplifikation von *rkpF*<sub>L</sub>
- oTBC121 gTC gCC CAC ggT gCC gC  
Sequenzierungsprimer bindet an 2635-2651 in X64131
- oTBC131 TTA gAT CAT ATg Acg gTT gAA ATC ATA ggg  
Primer für die Amplifikation von *rkpA*<sub>114</sub>'
- oTBC132 gCg CCg CgA ACC gAT GG  
Primer für die Amplifikation von *rkpA*<sub>114</sub>
- oTBC149 TCT CAC ACC CAT CAT CGA CC  
Sequenzierungsprimer für pPP428
- oTBC150 CAA AAA ACg CCT CgT CAg C  
Sequenzierungsprimer für pPP428
- oTBC151 CCC TTT CTC TTT gCA TCC  
Sequenzierungsprimer für pPP428
- oTBC152 gTg CCA gAA CgA gTg CAT CC  
Sequenzierungsprimer für pPP428
- oTBC153 CTT CgC TCA TCg AAA TCC  
Sequenzierungsprimer für pPP428
- oTBC154 CCA CTC gCC TTC AAC TgA Cg  
Sequenzierungsprimer für pPP428
- oTBC155 ATA ggA gAT TgC ggCA gC  
Sequenzierungsprimer für pPP428
- oTBC156 TTA CgT CAC CgA CCT CAA CC  
Sequenzierungsprimer für pPP428
- oTBC157 gAC ATT gAC ggC CTT CTC g  
Sequenzierungsprimer für pPP428
- oTBC158 AAC CTC ACC ATg ACC AAC C  
Sequenzierungsprimer für pPP428
- oTBC159 gCg ATC AAg ACT CCT CTg C

Sequenzierungsprimer für pPP428  
oTBC160      gAg Agg ATC gAA AgC gCA gC  
Sequenzierungsprimer für pPP428  
oTBC161      CCT TAC CAA CAC gAT CAC C  
Sequenzierungsprimer für pPP428  
oTBC162      ATg ATg AAA gCC Agg ATg AAg C  
Sequenzierungsprimer für pPP428  
oTBC163      gAA ACT TTT Cgg CAA gCT Cg  
Sequenzierungsprimer für pPP428

## **2.10 DNA-Arbeiten**

### **2.10.1 DNA-Isolierung**

#### **2.10.1.1 Isolierung genomischer DNA**

Eine über Nacht gewachsene *S. meliloti*-Kultur (1 ml) wurde durch Zentrifugation geerntet und mit 1 ml TE-Puffer, pH 8.0 gewaschen. Die Bakterien wurden in 400 µl TE-Puffer, 50 µl 10%-iger SDS und 50 µl (5 mg/ml) Proteinase K (Sigma, Deisenhofen) resuspendiert und für 1 Stunde bei 37°C inkubiert. Anschließend wurde die Suspension mit einer Injektionskanüle G 20-11/2 geschert. Die wässrige Phase wurde nacheinander mit einem Volumen Phenol, einem Volumen Phenol/Chloroform (1:1 v/v) und einem Volumen Chloroform/Isoamylalkohol (24:1 v/v) extrahiert und die genomische DNA in der Wasserphase nachfolgend mit 2 Volumen Ethanol gefällt. Das Pellet wurde in 100 µl TE-Puffer gelöst und bei -20°C gelagert.

#### **2.10.1.2 Plasmidpräparation aus *E. coli* und *S. meliloti***

Plasmid-DNA wurde nach der Methode von Birnboim & Doly (Birnboim & Doly, 1979) isoliert. Für Sequenzierungsreaktionen wurde die Plasmid-DNA mit dem "QIAGEN Plasmid Kit" der Firma Qiagen (Hilden) gereinigt.

## **2.10.2 Reinigung und Fällung von DNA**

Die Reinigung der DNA durch Extraktion mit Phenol/Chloroform, sowie die Fällung der DNA mit Ethanol oder Isopropanol wurde, wie bei Sambrook *et al.* (Sambrook *et al.*, 1989) beschrieben, durchgeführt. Zur Reinigung der DNA für die Elektroporation wurde die Silikamatrix-Methode verwendet.

### **2.10.2.1 Reinigung von DNA mit einer Silikamatrix**

Aus Agarosegelen und für die Elektroporation wurde die DNA mit dem "Gene DNA Purification System" der Firma BioRad (München) gereinigt.

## **2.10.3 Quantifizierung und Reinheitsbestimmung von DNA**

Reinheit und Konzentrationen von DNA-Lösungen wurden spektrofotometrisch, wie in Sambrook *et al.* (Sambrook *et al.*, 1989) beschrieben, bestimmt. Geringe DNA-Mengen und kleine Volumina wurden durch Vergleich der Ethidiumbromidfluoreszenz mit einem DNA-Standard quantifiziert.

## **2.10.4 Enzymatische Modifizierung von DNA**

### **2.10.4.1 Restriktionshydrolyse**

Restriktionshydrolysen wurden nach Angaben des Herstellers (Pharmacia, Freiburg; Roche Molecular Biochemicals, Mannheim; Fermentas, Vilnius) durchgeführt. Üblicherweise wurde 1 h bei 37°C im Wasserbad in einem Mindestvolumen von 10 µl inkubiert.

#### **2.10.4.2 Dephosphorylierung mit alkalischer Phosphatase**

Zur Verhinderung von Selbstligationen der gespaltenen Vektoren bei der Klonierung wurden die überhängenden Enden dephosphoryliert. Dafür wurde dem Restriktionsansatz nach der Restriktion 1 U alkalische Phosphatase (Roche Molecular Biochemicals, Mannheim) zugegeben. Zuerst wurde für 30 min bei 37 °C und für weitere 30 min bei 52 °C inkubiert. Nach Zugabe von 1/5 Vol. EGTA (100 mM, pH 8,0) wurde 30 min bei 65°C inkubiert.

#### **2.10.4.3 Auffüllen überhängender Enden mit Klenow-Fragment**

Die 5´-überhängenden Enden geschnittener DNA wurden mit dem Klenow-Fragment nach der in Sambrook *et al.* (Sambrook *et al.*, 1989) beschriebenen Vorschrift aufgefüllt. Zum Restriktionsansatz von 20 µl wurden 3 µl dNTPs (10 mM), 3 µl Puffer und 1 U Klenow-Fragment gegeben und auf 50 µl mit Aqua bidest. aufgefüllt. Nach der Inkubation für 30 min bei 37°C wurde für 10 min bei 75°C eine Hitzedenaturierung durchgeführt.

#### **2.10.4.4 Entfernen überhängender Enden mit T4 DNA-Polymerase**

Die 3´-überhängenden Enden geschnittener DNA wurden mit T4 DNA-Polymerase nach der in Sambrook *et al.* (Sambrook *et al.*, 1989) beschriebenen Vorschrift entfernt. Zum Restriktionsansatz von 20 µl wurden 3 µl dNTPs (10 mM), 3 µl Puffer und 0,1 U T4 DNA-Polymerase gegeben und auf 50 µl mit Aqua bidest. aufgefüllt. Nach der Inkubation für 5 min bei 16°C wurde für 10 min bei 80°C eine Hitzedenaturierung durchgeführt.

### **2.10.4.5 Ligation**

Die Ligation wurde nach Vorschrift des Herstellers (Roche Molecular Biochemicals, Mannheim) durchgeführt. Das Reaktionsvolumen von 10 µl mit 1 U T4 Ligase wurde für 16 h bei 16°C inkubiert.

### **2.10.5 Agarosegel-Elektrophorese von DNA**

Die Agarosegel-Elektrophoresen wurden, wie bei Sambrook *et al.* (Sambrook *et al.*, 1989) beschrieben, für 1 h bei RT mit 100 V in TAE-Puffer durchgeführt. Die Agarose-Konzentration betrug je nach Größe der zu analysierenden DNA-Stücke 0,75-2 %. Vor der Elektrophorese wurde der DNA-Lösung 1/5 Vol. Auftragspuffer (Sambrook *et al.*, 1989) mit Bromphenolblau zugegeben. Nach Ende der Elektrophorese wurde das Gel 15-30 min in Ethidiumbromidlösung (0,1 µg/ml) inkubiert. Unter UV-Licht war die DNA durch die Fluoreszenz des interkalierten Ethidiumbromids sichtbar. Als Größenstandard wurde λ-DNA (Pharmacia, Freiburg), mit *Hind*III oder mit *Eco*RI und *Hind*III gespalten, eingesetzt.

### **2.10.6 Transformation von *E. coli***

#### **2.10.6.1 Transformation nach der CaCl<sub>2</sub>-Methode**

##### **2.10.6.1.1 Herstellung kompetenter Zellen**

LB-Medium (50 ml) wurde mit 0,5 ml einer über Nacht gewachsenen Kultur von *E. coli* beimpft und bei 37°C und 150 rpm bis zur OD<sub>600</sub>=0,4 gezüchtet. Anschließend wurde die Kultur 10 min auf Eis gestellt und danach 5 min mit 2000 rpm bei 4°C abzentrifugiert. Der Überstand wurde dekantiert und die Zellen vorsichtig in 30 ml eiskalter CaCl<sub>2</sub>-Lösung (60 mM) resuspendiert. Die Zellsuspension wurde über Nacht in einem Eisbad im Kühlschrank inkubiert. 27 ml der CaCl<sub>2</sub>-Lösung wurden vorsichtig abpipettiert, und die sedimentierten

Zellen wurden in den verbleibenden 3 ml resuspendiert und in Anteilen von je 200 µl bei -80°C eingefroren.

#### **2.10.6.1.2 Transformation kompetenter Zellen**

Die kompetenten Zellen wurden auf Eis aufgetaut und mit 1-10 µl Plasmid-DNA-Lösung vorsichtig vermischt. Nach 30 min Inkubation auf Eis wurden die Zellsuspensionen für 1 min einem Hitzeschock von 42°C ausgesetzt und anschließend sofort für 5 min auf Eis gestellt. Nach Zugabe von 1 ml LB-Medium wurden die Zellen 1-1,5 h bei 37°C und 150 rpm geschüttelt, danach auf LB-Platten mit dem entsprechendem Antibiotikum ausplattiert und über Nacht bei 37°C im Brutschrank inkubiert.

### **2.10.6.2 Transformation durch Elektroporation**

#### **2.10.6.2.1 Herstellung kompetenter Zellen**

1 Liter LB-Medium ohne NaCl wurde mit 1/100 Vol. einer Übernachtskultur angeimpft und bis zu einer  $OD_{600}=0,6$  gezüchtet. Die Kultur wurde dann 15-30 min auf Eis gekühlt und bei 3000 x g abzentrifugiert. Die Zellen wurden in 0,5 l kaltem Reinstwasser resuspendiert und erneut zentrifugiert. Anschließend wurden die Zellen in 20 ml einer kalten 10% Glycerin-Lösung resuspendiert, erneut zentrifugiert und noch einmal in kalter 10% Glycerin-Lösung resuspendiert. Die Zellsuspension wurde in Aliquots bei -80 °C eingefroren.

#### **2.10.6.2.2 Elektroporation**

Es wurden entweder 1-2 µl nicht entsalzter Ligationsansatz für die Elektroporation von 100 µl einer Zellsuspension verwendet, oder der gesamte Ligationsansatz wurde mit Silikamilch gereinigt, in H<sub>2</sub>O bidest. aufgenommen und komplett zur Elektroporation eingesetzt. Das Gemisch wurde in eine Elektroporationsküvette mit 0,2 cm Elektrodenabstand gefüllt. Die Zellen

wurden dann bei 2,5 kV, 200  $\mu$ s und 40  $\mu$ F in einer Gen Pulser II System Elektroporationsapparatur der Firma BioRad elektroporiert. Zur Regeneration wurde 1 ml eines 37°C warmen SOC-Mediums zugegeben. Die Zellen wurden 1 h bei 37°C inkubiert. Danach wurden Verdünnungen auf den entsprechenden Selektivmedien ausplattiert.

## 2.10.7 Transfer mobilisierbarer Vektoren durch Konjugation

### 2.10.7.1 Drei-Punkt-Kreuzung

Zur Einschleusung von Plasmiden in *S. meliloti*-Zellen wurde die Konjugation verwendet. Bei der Konjugation kommt es zum Austausch von genetischem Material zwischen einem Donorstamm und einem Empfängerstamm (Rezipient) über eine Brückenbildung zwischen beiden Stämmen. Normalerweise enthält der *E. coli*-Donorstamm das zu übertragende Plasmid, welches eine Mobilisierungsstelle (Mob-site) besitzt. Um den Transfer auszulösen, bedarf es zusätzlich eines Helfer-Stammes, der die *tra*-Gene enthält. Die drei verschiedenen Stämme Donor, Helfer und Rezipient werden in einem bestimmten Verhältnis miteinander vermischt und für einige Stunden miteinander inkubiert.

Für diese Drei-Punkt-Kreuzung wurde eine, über Nacht gewachsene, *S. meliloti*-Kultur (Rezipient) in frisches TY-Medium zu einer  $OD_{600}=0,15$  überimpft und für 3 h bei 30°C inkubiert. Die über Nacht auf LB-Platten mit entsprechendem Antibiotikum gewachsenen Helfer- und Donor-Stämme wurden 1 Stunde vor der Kreuzung in LB-Medium zu einer  $OD_{600}=0,15$  angeimpft. Die Kulturen wurden, falls Antibiotikareste entfernt werden mussten, mit TY-Medium gewaschen. Jeweils Suspensionen mit 100  $\mu$ l Rezipient, 40  $\mu$ l Helfer-Stamm und 20  $\mu$ l Donor wurden in einem Eppendorfgefäß miteinander vermischt und in 10  $\mu$ l-Portionen auf TY-Platten getropft. Nachdem die Platten getrocknet waren, wurden sie bei 29°C über Nacht inkubiert. Als Kontrollen dienten Kreuzungen ohne Helfer-Stamm oder ohne Donor-Stamm. Am nächsten Morgen wurden die Kulturen mit einer Impföse im 3-Strich-Verfahren

auf TY-Platten ausgestrichen. Als selektierende Antibiotika wurden Streptomycin (Selektion für *S. meliloti*), Piperazillin (Selektion gegen *E. coli*) und Tetracyclin (Selektion für das Plasmid) verwendet.

Nach etwa 5 Tagen waren Einzelkolonien auf den Platten sichtbar. Die Bakterien wurden in zwei weiteren Durchgängen vereinzelt. Die erfolgreich isolierten Transkonjuganten wurden bei  $-80^{\circ}\text{C}$  gelagert.

### 2.10.7.2 "Knock-out-Mutagenese"

Die "Knock-out-Mutagenese" wurde in zwei Schritten vollzogen. Im ersten Schritt wurden Transkonjuganten, wie unter 2.10.7.1 beschrieben, hergestellt. Diese Transkonjuganten enthielten auf dem Plasmid eine Resistenz-Kassette im Gen, welches unterbrochen werden sollte. Im zweiten Schritt wurde ein zweiter broad-host-range-Vektor der gleichen Inkompatibilitätsgruppe (IncP) in die Transkonjuganten eingekreuzt. Dafür wurde der *E. coli*-Stamm J53 R751-pGM2 (Jacoby *et al.*, 1976) verwendet. Er besitzt die *tra*-Gene in seinem Genom, weshalb kein zusätzlicher Helfer-Stamm notwendig war.

Eine, über Nacht in TY-Medium mit Tetracyclin gewachsene, *S. meliloti*-Transkonjuganten-Kultur (Rezipient) wurde 3 h vor der Kreuzung in frischem TY-Medium zu einer  $\text{OD}_{600}=0,15$  überimpft. Der Donor *E. coli*-Stamm J53 R751-pGM2, welcher über Nacht auf einer LB-Platte mit 30  $\mu\text{l/ml}$  Gentamicin gewachsen war, wurde 1 h vor der Kreuzung ebenfalls zu einer  $\text{OD}_{600}=0,15$  angeimpft. Für die Kreuzung wurde 1 ml der *S. meliloti*-Kultur in ein Eppendorfgefäß gegeben und anschließend bei 4000xg für 2 min zentrifugiert. Das Medium wurde entfernt und 1 ml der J53 R751-pGM2-Kultur wurde auf die Zellen pipettiert. Nach erneuter Zentrifugation wurden die Bakterien in 100  $\mu\text{l}$  TY-Medium resuspendiert und in 10  $\mu\text{l}$ -Portionen auf TY-Platten getropft. Nach dem Trocknen der Platten wurden sie bei  $29^{\circ}\text{C}$  inkubiert. Am nächsten Morgen wurden die Kulturen mit einer Impföse im 3-Strich-Verfahren auf TY-Platten ausgestrichen. Als selektierende Antibiotika wurden Piperazillin (Selektion gegen *E. coli*), Gentamicin (Selektion für pGM2) und Neomycin (Selektion für Resistenz-Kassette) verwendet.

Nach etwa 5 Tagen waren Einzelkolonien sichtbar, die auf frischen Platten weiter vereinzelt wurden. Bei einfachen Rekombinationen wäre das ganze Plasmid und damit auch die Tetracyclin-vermittelnde Resistenz des Vektors ins Genom integriert. Solche Tetracyclin-resistente Transkonjuganten wurden verworfen. Tetracyclin-sensitive, doppelt rekombinierte Transkonjuganten wurde in einem weiteren Durchgang vereinzelt und anschließend bei  $-80^{\circ}\text{C}$  gelagert.

### 2.10.8 Amplifikation von DNA mittels der Polymerase-Kettenreaktion (PCR)

Zur Vervielfältigung von DNA-Stücken und besonders zum Einfügen neuer Restriktionsschnittstellen für die Ligation wurde die Polymerase-Kettenreaktion (PCR) verwendet. Routinemäßig wurden 500 pg DNA und stöchiometrische Mengen der beiden Primer ( $5\ \mu\text{M}$ ) eingesetzt. Die Reaktion erfolgte im Reaktionspuffer des Herstellers und in Gegenwart einer Mischung aus dNTPs, wobei die Endkonzentration jedes einzelnen dNTP  $250\ \mu\text{M}$  betrug. Das Endvolumen der Ansätze umfasste  $100\ \mu\text{l}$  wenn *Pfu*-Polymerase oder  $25\ \mu\text{l}$  wenn *Taq*-Polymerase verwendet wurde. In einem programmierbaren Heizblock ("Cycler") durchlief die Probe dann ein Temperaturprogramm mit üblicherweise 30 Zyklen. Die Temperaturschritte betragen:

5 min bei  $95^{\circ}\text{C}$ ,

dann Zyklen von 45 sec bei  $95^{\circ}\text{C}$

1 min bei der für das Primerpaar optimalen  
Bindungstemperatur

Inkubation bei  $72^{\circ}\text{C}$  je nach Templategröße

( $2,5\ \text{min}/\text{kb}$  bei *Pfu*-Polymerase oder  $1\ \text{min}/\text{kb}$   
für *Taq*-Polymerase)

gefolgt von einem abschließenden Elongationsschritt von 10 min bei  $72^{\circ}\text{C}$ .

## 2.10.9 DNA-Sequenzierung

DNA wurde nach der Kettenabbruchmethode von Sanger (Sanger *et al.*, 1977) mit Didesoxynukleotiden und [<sup>35</sup>S]-dATP (Amersham Buchler, Braunschweig) sequenziert. Für die Sequenzierungsreaktion wurde der "Sequenase 7-deaza dGTP Sequencing Kit" der Firma Amersham Buchler (Braunschweig) verwendet. Die Elektrophorese wurde mit einem Hoefer SQ3 Sequencer (Pharmacia, Freiburg) durchgeführt. Ein Teil der DNA-Sequenzierung wurde von der Firma Meixner, Berlin durchgeführt

### 2.10.9.1 Sequenzierungsreaktion

Für eine Sequenzierungsreaktion wurden 5 µg denaturierte DNA eingesetzt. Dazu wurde 5 µg doppelsträngiger DNA in einem Volumen von 30 µl mit 7,5 µl Denaturierungslösung (2 M NaOH, 2 mM EDTA) vermischt und 30 min bei 37 °C inkubiert. Anschließend wurden 6,6 µl 3 M Natriumacetatlösung, pH 4,6 und 120 µl eiskalter Ethanol zugegeben und 20 min bei -80 °C inkubiert. Die denaturierte DNA wurde durch 15 min Zentrifugation bei 15000xg sedimentiert, mit 70 % Ethanol gewaschen, erneut zentrifugiert und getrocknet. Die Sequenzierungsreaktion wurde nach dem Protokoll des Herstellers (Amersham Buchler, Braunschweig) mit [<sup>35</sup>S]-dATP (Amersham Buchler, Braunschweig) durchgeführt. Die radioaktiv markierte DNA wurde bis zur Elektrophorese bei -20 °C aufbewahrt.

### 2.10.9.2 Elektrophorese

Der Zusammenbau des Hoefer SQ3 Sequencers erfolgte nach Vorschrift des Herstellers. Damit sich das Gel nach der Elektrophorese leichter von der Glasplatte löst, wurde auf die Innenseite einer der Glasplatten Repel-Silane (Pharmacia, Freiburg) aufgetragen. Die filtrierte Gellösung wurde nach Zugabe von 1 ml 10 % APS und 25 µl TEMED zwischen die Glasplatten gegossen und über Nacht bei RT polymerisiert. Nach dem Einsetzen der Glasplatten in die

Apparatur wurde in einem Vorlauf das Gel bei 60 W auf 55 °C erwärmt. Vor dem Auftragen wurden die Proben 2 min bei 75 °C inkubiert. Die Elektrophorese lief 1,5-6 h bei 60 W. Nach der Elektrophorese wurde das Gel auf Whatman 3MM Papier in einem Hoefer SE1160 Gelrockner (Pharmacia, Freiburg) bei 80 °C im Vakuum getrocknet. Um die radioaktiv markierte DNA sichtbar zu machen, wurde das getrocknete Gel mindestens 2 Tage an einen Kodak BIOMAX MR1 Röntgenfilm (Sigma, Deisenhofen) gepresst und wie in 2.10.9.3 beschrieben entwickelt.

Gellösung:            30 ml Protogel (Biozym, Oldendorf)  
                          7,5 ml 20 x TBE-Puffer  
                          63 g Harnstoff (Roth, Karlsruhe)  
                          ad 150 ml H<sub>2</sub>O

Tankpuffer:           1x    TBE

20xTBE-Puffer       216 g Tris  
                          110 g H<sub>3</sub>BO<sub>3</sub>  
                          80 ml 0,5 M EDTA, pH 8,0  
                          ad 1 l H<sub>2</sub>O

### 2.10.9.3    Entwicklung des Röntgenfilms

Zur Entwicklung des Röntgenfilms wurde dieser nach der Exposition 3-5 min in Entwickler-Lösung (GBX Developer; Sigma, Deisenhofen) gelegt, anschließend kurz in H<sub>2</sub>O gewaschen und schließlich 1-3 min in Fixierungslösung (GBX Fixer; Sigma, Deisenhofen) geschwenkt. Danach wurde nochmals kurz mit H<sub>2</sub>O gewaschen und der entwickelte Film getrocknet.

## 2.10.10 DNA-Analyse Software und Internetdienste

DNA- und Aminosäuresequenzen wurden mit Hilfe verschiedener Versionen des BLAST-Algorithmus (Altschul *et al.*, 1990) mit den Sequenzen der Datenbanken am "National Center for Biotechnology Information NCBI", USA verglichen (<http://www.ncbi.nlm.nih.gov>). Das Suchergebnis lieferte Hinweise auf mögliche Strukturen und Funktionen ähnlicher Gene und Genprodukte. Zur Vorhersage von Offenen Leserahmen (ORFs) wurde das Programm ORF-Finder verwendet ([http://bioinformatics.org/sms/orf\\_find.html](http://bioinformatics.org/sms/orf_find.html)). Proteinalignments wurden mit dem Programm Clustal W (Thompson *et al.*, 1994) durchgeführt. Das Auffinden und Markieren von übereinstimmenden Aminosäuren wurde mit dem Programm BoxShade ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)) vorgenommen. Zur Suche nach Sequenzmotiven in Proteinsequenzen wurde die PROSITE Datenbank (Bairoch *et al.*, 1997) verwendet. Für allgemeine DNA-Analysen wurde das Programm OMIGA der Firma Oxford Molecular Ltd. verwendet.

## 2.11 Proteinchemische Arbeiten

### 2.11.1 Native PAGE von Proteinen

Zur Auftrennung von Proteinen im nativen Zustand wurde ein 20%T; 2,7%C-Trenngel, das mit einem 4%T; 2,7%C- Sammelgel überschichtet war, benutzt (Jackowski & Rock, 1983). Für das Trenngel wurden auf 10 ml einer 30%T; 2,7%C-Monomerstammlösung 1,2 ml H<sub>2</sub>O, 3,75 ml 4 x Laufpuffer, 60 µl 10% APS und 17,5 µl TEMED zugegeben. Für das Sammelgel wurden auf 1 ml einer 30%T; 2,7%C-Monomerstammlösung 4,55 ml H<sub>2</sub>O, 1,88 ml 4 x Sammelpuffer, 30 µl 10% APS und 10 µl TEMED zugegeben. Die Proben wurden mit dem 2 x Auftragspuffer gemischt und die Gele bei RT mit einer Spannung von 100 V für 2 h entwickelt.

4 x Laufpuffer:  
(PAGE)

1,5 M Tris/HCl, pH 8,8

4 x Sammelpuffer: (PAGE)	500 mM Tris/HCl, pH 6,8
2 x Auftragspuffer: (Native PAGE)	20% Glycerin (w/v) in 4 x Laufpuffer 2% (v/v) einer gesättigten Bromphenolblaulösung in EtOH

### 2.11.2 SDS-PAGE

Die Auftrennung von Proteingemischen aufgrund ihres relativen Molekulargewichts unter reduzierenden Bedingungen erfolgte mittels Polyacrylamid-Gelelektrophorese (PAGE) nach Laemmli (Laemmli, 1970). Die Auftrennung wurde mit Acrylamid/Bisacrylamid-Sammelgelen von 4% und Trenngelen von 5; 7,5; 12 oder 15% durchgeführt.

Pipettierschema für SDS -Polyacrylamid-Gele:

Lösungen	Sammelgel 4%	Trenngel 5%	Trenngel 7.5%	Trenngel 12%	Trenngel 15%
AA/BIS	3,2 ml	4,6 ml	6,9 ml	8,9 ml	10,85 ml
1 M TRIS/HCl (pH 8,7)	-	7,5 ml	7,5 ml	7,5 ml	7,5 ml
1 M TRIS/HCl (pH 6,8)	1,25 ml	-	-	-	-
SDS (10%)	0,1 ml	0,1 ml	0,1 ml	0,1 ml	0,1 ml
ddH <sub>2</sub> O	5,5 ml	7,7 ml	5,4 ml	3,4 ml	1,4 ml
TEMED	25 µl	70 µl	70 µl	70 µl	70 µl
APS (10%)	50 µl	17 µl	17 µl	17 µl	17 µl

Die Proteinlysate wurden mit 2 x Auftragspuffer versetzt und vor dem Auftrag auf das Gel 5 min auf 95°C erhitzt. Zur Analyse von Proteinlysaten unter nicht-reduzierenden Bedingungen wurden diese in Auftragspuffer ohne β-Mercaptoethanol aufgenommen und nicht aufgeköcht. Die Gele wurden bei RT mit einer Spannung von 100 V für 1-2 h entwickelt.

AA/BIS:	30% (w/v) Acrylamid, 0.8% (w/v) Bisacrylamid (Roth, Karlsruhe)
2 x Auftragspuffer: (SDS-PAGE)	20% Glycerin (w/v) 8% (w/v) SDS 2% (v/v) Mercaptoethanol 2% (v/v) einer gesättigten Bromphenolblaulösung in EtOH in 4 x Laufpuffer
Tankpuffer: (SDS-PAGE)	25 mM Tris 192 mM Glycin 0,1% (w/v) SDS

### 2.11.3 Tricin-Gele

Zur Auftrennung von kleinen Proteinen im denaturierten Zustand wurde eine SDS-PAGE nach Schägger und von Jagow (Schägger & von Jagow, 1987) durchgeführt. Dabei wurde ein 16,5%T; 6%C-Trenngel benutzt, das mit einem 4%T; 3%C-Sammelgel überschichtet war. Für das Trenngel wurden auf 5 ml einer 49,5%T; 6%C-Monomerstammlösung 5 ml H<sub>2</sub>O, 5 ml Gelpuffer, 4 g Glycerin (100%), 75 µl 10% APS und 7,5 µl TEMED zugegeben. Für das Sammelgel wurden auf 500 µl einer 49,5%T; 3%C-Monomerstammlösung 4,2 ml H<sub>2</sub>O, 1,55 ml Gelpuffer, 32 µl 10% APS und 3,2 µl TEMED zugegeben. Die Proben wurden mit 2 x Auftragspuffer gemischt und vor dem Lauf für 30 min bei 40°C inkubiert. Die Gele wurden bei einer Temperatur von 10°C zuerst für 80 min bei 20 V und dann, bis der Bromphenolblau-Marker die untere Gelkante erreichte, bei einer Spannung von 35 V entwickelt.

Gelpuffer: (Tricin-PAGE)	3 M Tris 0,3% (w/v) SDS pH 8,45 mit HCl eingestellt
-----------------------------	---

Anodenpuffer: (Tricin-PAGE)	0,2 M Tris pH 8,9 mit HCl eingestellt
Kathodenpuffer: (Tricin-PAGE)	0,1 M Tris 0,1 M Tricin 0,1% (w/v) SDS
2 x Auftragspuffer: (Tricin-PAGE)	8% (w/v) SDS 24 % (w/v) Glycerin 100 mM Tris/HCl, pH 6,8 2% (v/v) einer gesättigten Bromphenolblaulösung in EttOH 2% (v/v) Mercaptoethanol

## 2.11.4 Färben von Proteingelen

### 2.11.4.1 Coomassiefärbung

Die Polyacrylamidgele wurden mit Serva blue G nach der Vorschrift von Schägger und von Jagow (Schägger & von Jagow, 1987) gefärbt. Die Gele wurden zum Fixieren der Proteine für 30 min in einer Lösung aus 50% MeOH und 10% HAc geschüttelt. Gefärbt wurde dann für 1 h mit einer 10% HAc, die 0,025% (w/v) Serva blue G enthielt. Entfärbt wurden die Gele mit 10% HAc über 2 h, wobei alle 30 min die Lösung erneuert wurde.

### 2.11.4.2 Silberfärbung

Silberfärbung von Proteingelen wurde nach der Vorschrift von Nesterenko (Nesterenko *et al.*, 1994) nach dem folgenden Schema durchgeführt

Fixierung	1,5 ml TCA (50%), 25 µl 37% HCHO in 60 ml 50% Aceton	5 min
Spülen	Aqua bidest.	3 x 5 sec
Waschen	Aqua bidest.	5 min
Spülen	Aqua bidest.	3 x 5 sec

Vorbehandlung	60 ml 50% Aceton	5 min
Vorbehandlung	100 µl 10% Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> * 5 H <sub>2</sub> O in 60 ml Aqua bidest.	1 min
Spülen	Aqua bidest.	3 x 5 sec
Imprägnieren	800 µl 20% AgNO <sub>3</sub> , 600 µl 37% HCHO in 60 ml Aqua bidest.	8 min
Spülen	Aqua bidest.	3 x 5 sec
Entwickeln	1,2 g Na <sub>2</sub> CO <sub>3</sub> , 25 µl 37% HCHO, 25 µl 10% Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> * 5 H <sub>2</sub> O in 60 ml Aqua bidest.	10-20 sec
Stopp	1% HAc in Aqua bidest.	30 sec
Spülen	Aqua bidest.	3 x 15 sec

### 2.11.5 Blotten von Proteingelen

Zum Blotten der Proteine wurde die Semi-Dry-Blottingapparatur Multiphor der Firma Amersham Bioscience benutzt. Als Blottingpapier wurde 3MM der Firma Whatmann benutzt. Die Proteine wurden auf eine Nitrocellulosemembran BA83, 0,2 µm der Firma Schleicher und Schüll geblottet. Nachdem die Apparatur nach Herstellerangaben vorbereitet worden war, wurde mit dem Blottingpuffer bei 0,8 mA/cm<sup>2</sup> Blotfläche für 2 h geblottet

### 2.11.6 Blotten von Proteingelen nach radioaktiver Markierung von Proteinen.

Zum Blotten radioaktiv markierter Proteine wurde die Naßblotapparatur Mighty Small Transfer Electrophoresis Unit der Firma Hoefer benutzt. Als Blottingpapier wurde 3MM der Firma Whatmann benutzt. Die Proteine wurden auf eine Nitrocellulosemembran BA83, 0,2 µm der Firma Schleicher und Schüll geblottet. Nachdem die Apparatur nach Herstellerangaben vorbereitet worden war, wurde mit dem Blottingpuffer bei 250 mA für 1 h geblottet.

### 2.11.7 Immundetektion von immobilisierten Proteinen (Western Blot)

Nach dem Transfer der Proteine auf die Nitrocellulosemembran werden im Einzelnen folgende Schritte durchgeführt:

- Blocken der Membran mit einer 1% BSA-Lösung in PBS bei 4°C über Nacht
- 2 mal 1 Stunde waschen der Membran mit PBS-Tween (0,5% Tween 20)
- Inkubation mit primärem Antikörper (verdünnt in PBS-Tween) für 2 Stunden bei RT
- Die Membran kurz mit H<sub>2</sub>O waschen
- 2-mal 15 Minuten waschen der Membran mit PBS-Tween (0,5% Tween 20)
- Inkubation mit sekundärem Antikörper (verdünnt in PBS-Tween) für 1 Stunden bei RT
- Die Membran kurz mit H<sub>2</sub>O waschen
- 2-mal 15 Minuten waschen der Membran mit PBS-Tween (0,5% Tween 20)

Färben der Membran mit BCIP/NBT Färbetabletten der Firma Sigma nach Angaben des Herstellers.

### 2.11.8 Proteinbestimmung

Proteinbestimmungen wurden mit der Biuret-Methode nach Itzhaki und Gill durchgeführt (Itzhaki & Gill, 1964.). Für die Reaktionslösung wurden 315 mg CuSO<sub>4</sub> \* 5 H<sub>2</sub>O und 1,5 g Kalium-Natriumtartrat in 125 ml H<sub>2</sub>O gelöst. Anschließend wurden 75 ml 10% NaOH und 250 mg KJ zugegeben und die Lösung mit H<sub>2</sub>O auf 250 ml aufgefüllt. Für die Messung wurden 833,3 µl der Reaktionslösung mit 166,7 µl der Probe gemischt und für 20-30 min inkubiert. Anschließend wurde die Extinktion bei 540 nm gemessen. Die Eichgerade wurde mit BSA als Standardprotein erstellt.

## **2.11.9 Zellaufschluss**

### **2.11.9.1 French-Press-Aufschluss**

Die Zellen von *E. coli* oder *S. meliloti* wurden in 1/10 Vol. der ursprünglichen Kulturflüssigkeit mit Aufschlusspuffer (50 mM Tris/HCl pH 6,8, 100 mM KCl) aufgenommen. Die Zellsuspensionen wurden dann durch drei French-Press-Passagen bei 20000 psi aufgeschlossen. Anschließend wurden Zelltrümmer und nicht aufgeschlossene Zellen durch Zentrifugieren bei 6000xg und 4°C für 30 min abgetrennt. Der so erhaltene Rohextrakt wurde entweder bei -20 °C eingefroren oder sofort weiterverarbeitet.

### **2.11.9.2 Zellaufschluss mit Lysozym**

Die im Lysispuffer (50 mM Tris/HCl pH 6,8; 25% (w/v) Sucrose; 5 mM EDTA; 10 mM DTT) resuspendierten Zellen wurden mehrmals bei -20°C eingefroren und auf Eis aufgetaut. Nach dem letzten Auftauen wurde 1/5 Vol. einer Lysislösung (250 mM Tris/HCl pH 7,5; 10 mM DTT; 5 mg/ml Lysozym) zugegeben und der Ansatz für 15 min auf Eis inkubiert. Danach wurden 2/5 Vol. einer 100 mM EDTA-Lösung (pH 8,0), die 10 mM DTT enthielt, zugegeben und erneut für 15 min auf Eis inkubiert. Anschließend wurden die Proben bei 7000xg für 5 min zentrifugiert und die Überstände für die weiteren Untersuchungen verwendet.

## **2.11.10 Aufreinigung des RkpF-Proteins**

### **2.11.10.1 Isopropanolfällung**

Zu dem auf 4°C temperierten Rohextrakt von *E. coli* BL21(DE3)pTB1003 (Epple, 1995) wurde langsam, unter Rühren, die entsprechende Menge kalten Isopropanols (4°C) zugetropft. Nach der Isopropanolzugabe wurde noch kurz gerührt und die Lösung dann für 1 h bei 4°C inkubiert. Anschließend wurde bei 6000xg für 30 min zentrifugiert und der Überstand in einen Rundkolben

überführt. Bei einer Wasserbad-Temperatur von 40°C wurde die Lösung am Rotationsverdampfer eingengt, bis der Anteil an Isopropanol kleiner als 5% war.

### **2.11.10.2 Chromatografie an DEAE 52-Cellulose**

Für die weitere Isolierung des RkpF-Proteins wurde dem Überstand der Isopropanolfällung CHAPS (1 mM) zugegeben. Die Lösung wurde mit einer Flussrate von 10 ml/h auf eine 30 ml DE 52-Säule (Durchmesser 1,5 cm) aufgetragen. Die Säule wurde anschließend mit 90 ml Puffer (50 mM Tris/HCl pH 8,0; 1 mM CHAPS) gewaschen. Die Elution wurde mit einem linearen NaCl-Gradienten (90 ml) von 0-1 M bei einer Flussrate von 19 ml/h durchgeführt. Die Elution der Proteine wurde mit einem Durchflussfotometer bei 254 nm verfolgt. Der Gradient wurde mit einem Leitfähigkeitsdetektor überprüft. Es wurden Fraktionen mit der Größe von 1,5 ml gesammelt. Die Fraktionen in denen der Anteil von RkpF über 90% lag wurden vereinigt.

### **2.11.10.3 Aufkonzentrieren von RkpF-Pools**

Die RkpF enthaltenden Proben wurden mit Puffer (100 mM Tris/HCl, pH 8,0; 2 mM CHAPS) verdünnt bis die Konzentration an NaCl in der Probe unter 50 mM lag. Dann wurde die Probe mit einer Flussrate von 25 ml/h auf eine 10 ml DE52-Säule (Durchmesser 1 cm) aufgetragen. Nach dem Waschen der Säule mit 20 ml Puffer (100 mM Tris/HCl pH 8,0; 1 mM CHAPS) wurden die Proteine anschließend mit 800 mM NaCl (in 100 mM Tris/HCl pH 8,0; 1 mM CHAPS) eluiert. Die gesammelten Fraktionen (0,75 ml) wurden in der nativen PAGE analysiert.

### **2.11.10.4 Gelfiltration von RkpF**

Zur Gelfiltration von RkpF wurde eine Säule (Durchmesser 2,6 cm; Länge 90 cm) mit 475 ml (60 g) G50-Superfine (Pharmacia) in einem Puffer mit 25 mM

Tris und 192 mM Glycin (Laufpuffer der nativen PAGE) gefüllt. Die Proben wurden mit Saccharose gemischt und in einem minimalen Volumen auf die Säule aufgebracht. Die Chromatografie wurde mit einer Flussrate von 36 ml/h durchgeführt und es wurden Fraktionen zu 5,5 ml gesammelt.

#### **2.11.10.5 *In vitro*-Markierung von RkpF durch radioaktives CoA**

Für die *in vitro*-Markierung von Holo-RkpF wurde die prosthetische Gruppe von radioaktivem CoA mit der gereinigten Holo-ACP-Synthase (Lambalot & Walsh, 1995) auf Apo-RkpF übertragen. Verschiedene Proteinmengen von Apo-RkpF in 10 mM Tris/HCl pH 8,0 wurden mit 40 mM DTT, 25 mM MgCl<sub>2</sub>, 3 µM CoA, 500 pM [<sup>3</sup>H]-CoA und 100 µg Holo-ACP-Synthase in einem Volumen von 100 µl für 45 min bei 37°C inkubiert. Anschließend wurden 30 µl einer BSA-Lösung (12,5 mg/ml) zugegeben und die Proteine mit 800 µl kalter 10%-igen TCA gefällt. Die Proben wurden dann für 5 min bei 12000xg zentrifugiert und das Pellet noch zweimal mit jeweils 800 µl kalter TCA-Lösung (10%) gewaschen. Nachdem das TCA vollständig entfernt worden ist, wurden die Pellets in 100 µl 1 M-Tris-Lösung gelöst und die Menge an Radioaktivität in einem Szintillationszähler bestimmt.

#### **2.11.10.6 Konversion von Apo- zu Holo-RkpF**

Für die Konversion von der Apo- in die Holo-Form des RkpF-Proteins wurden 4 nmol Apoform mit 0,6 nmol der nach Lambalot und Walsh (Lambalot & Walsh, 1995) gereinigten Holo-ACP-Synthase, 10 mM DTT, 5 mM MgCl<sub>2</sub> und 300 µM CoA bei 37°C für 2 h inkubiert. Zum Abstoppen der Reaktion wurde EGTA (pH 8,0) bis zu einer Endkonzentration von 80 mM zugegeben.

### **2.11.10.7 Nachweis von RkpF in Mutanten von *Sinorhizobium meliloti* AK631**

Die zu untersuchenden Mutanten von *S. meliloti* AK631 wurden in 20 ml TY-Medium bis zu einer  $OD_{600}=1,2$  gezüchtet. Die Zellen wurden geerntet und in 1 ml Puffer (100 mM Tris/HCl pH 8,0) resuspendiert. Zur Induktion des plasmidkodierten RkpF wurde in den entsprechenden Kulturen 5 µg/ml Tetracyclin zum Medium zugegeben. Die Zellen wurden durch zwei French-Press-Passagen bei 20000 psi aufgeschlossen. Zelltrümmer wurden durch Zentrifugation abgetrennt und der Überstand bei  $-20^{\circ}\text{C}$  eingefroren. Zu je 150 µl dieser Rohextrakte wurden 225 µl Isopropanol zugegeben. Nach einer Inkubation von 60 min bei  $4^{\circ}\text{C}$  wurden die Proben 20 min bei 7000xg und  $4^{\circ}\text{C}$  zentrifugiert. Die Überstände wurden anschließend in der Speedvac getrocknet und in 15 µl Puffer resuspendiert. Die Proben wurden auf einer 20% nativen PAGE aufgetrennt und auf eine Nitrocellulosemembran geblottet. Der Blot wurde mit Anti-Rkp-F (1:5000) und Anti-Rabbit (1:10000)-Antikörpern entwickelt.

### **2.11.11 *In vivo*-Markierung von Proteinen**

#### **2.11.11.1 Markierung mit $\beta$ -[ $^3\text{H}$ ]Alanin**

Für die Markierung mit  $\beta$ -[ $^3\text{H}$ ]Alanin, das in die prosthetischen Gruppe 4`-Phosphopantethein von ACPs eingebaut wird, werden die Proteine in *E. coli* OG7001 exprimiert. OG7001 ist eine  $\beta$ -Alanin-auxothrophe Mutante von *E. coli* BL21(DE3) (Epple *et al.*, 1998). Aus einer Übernachtskultur in M9, die 1 µM  $\beta$ -Alanin enthielt, wurden zwei neue Kulturen in M9 mit einer  $OD_{600}=0,15$  angeimpft. Nur eine der Kulturen wurde mit  $\beta$ -Alanin supplementiert und das Wachstum der Kulturen wurde verfolgt, bis in der  $\beta$ -Alanin-freien Kultur eine Begrenzung des Wachstums durch den Mangel an  $\beta$ -Alanin auftrat. Ein Teil der  $\beta$ -Alanin-limitierten Kultur wurde geerntet und in frischem M9-Medium zu einer  $OD_{600}=0,15$  resuspendiert. Von dieser Kultur wurde 1 ml abgenommen und 1

$\mu\text{M}$   $\beta$ -[ $^3\text{H}$ ]Alanin zugegeben. Der restlichen Kultur wurde  $1 \mu\text{M}$   $\beta$ -Alanin zugegeben und die  $\text{OD}_{600}$  verfolgt. Bei einer  $\text{OD}_{600}=0,3$  wurde IPTG ( $100 \mu\text{M}$ ) zugegeben und die Zellen wurden nach weiteren 4 Stunden geerntet. Die Zellpellets der markierten Kultur wurden 3-mal mit  $1 \text{ ml}$  M9-Medium gewaschen, in  $100 \mu\text{l}$  Puffer ( $50 \text{ mM}$  Tris/HCl pH 6,8;  $25\%$  (w/v) Sucrose;  $5 \text{ mM}$  EDTA;  $10 \text{ mM}$  DTT) resuspendiert und bei  $-20^\circ\text{C}$  eingefroren.

### 2.11.11.2 Markierung mit [ $^{35}\text{S}$ ]Sulfat

Die Markierung von Proteinen mit [ $^{35}\text{S}$ ] wurde analog zu der Markierung mit [ $^3\text{H}$ ] durchgeführt. Das  $\text{MgSO}_4$  des M9-Mediums wurde durch ein [ $^{35}\text{S}$ ]-enthaltendes Sulfat ersetzt.

### 2.11.12 Autoradiografie

Zellulosenitratmembranen, auf die mit  $\beta$ -[ $^3\text{H}$ ]Alanin radioaktiv markierte Proteine geblottet worden waren, wurden zweimal mit Autofluor (National Diagnostics) besprüht. Die Membran wurde anschließend mit einem Röntgenfilm bei  $-80^\circ\text{C}$  inkubiert.

## 2.12 Untersuchung von Polysacchariden

### 2.12.1 Isolierung von Polysacchariden

Die Isolierung von zellassoziierten Polysacchariden wurde nach der Mini-Phenol-Wasser-Extraktionsmethode durchgeführt (Reuhs *et al.*, 1998). Es wurden die Zellen aus  $3 \text{ ml}$  *S. meliloti* TY-Kulturen mit einer  $\text{OD}_{600}=0,6$  geerntet. Zur Entfernung der Exopolysaccharide wurden die Zellen mit  $\text{H}_2\text{O}$  gewaschen. Die Zellen wurden dann in  $500 \mu\text{l}$  Puffer ( $50 \text{ mM}$   $\text{Na}_2\text{HPO}_4 \cdot 7 \text{ H}_2\text{O}$ ;  $5 \text{ mM}$  EDTA;  $0,05\%$   $\text{NaN}_3$ , pH 7) resuspendiert und intensiv gevortext. Anschließend wurden  $500 \mu\text{l}$  einer  $90\%$ -gen Phenol-Lösung zugegeben und die Proben für  $15 \text{ min}$  bei  $65^\circ\text{C}$  inkubiert. Während dieser Inkubation wurden die Proben alle  $3 \text{ Minuten}$

gevortext und anschließend 5 min im Eisbad abgekühlt. Nach einer Zentrifugation (5 min, 10000xg, 4°C) wurde die wässrige Phase abgenommen, gegen Aqua bidest. dialysiert (MWCO 12000-14000) und in der Speedvac getrocknet.

### 2.12.2 Gelelektrophorese von Polysacchariden

Die getrockneten Polysaccharide wurden in 20 µl Probenpuffer gelöst und auf 18%-DOC-PAGE-Gelen analysiert (Reuhs *et al.*, 1998). Die Gele wurden wie folgt hergestellt:

Lösung	Sammelgel	Trenngel
Monomer-Lösung (37,5:1) Roth	750 µl	7,68 ml
Laufpuffer	-	3,2
Sammelpuffer	1,59 ml	-
H <sub>2</sub> O	5,62 ml	5,12 ml
APS (10%)	40 µl	28,5 µl
TEMED	10,5 µl	6,25 µl
Gesamtvolumen	7 ml	16 ml

Laufpuffer                    22,71 g Tris / 100 ml H<sub>2</sub>O pH 8,8 (1,87 M)

Sammelpuffer                7,69 g Tris / 100 ml H<sub>2</sub>O pH 6,8 (0,63 M)

Tankpuffer                    2,7 g Glycin (290 mM), 4,5 g Tris (37 mM),  
2,5 g Natriumdeoxycholat (6 mM) /1000 ml H<sub>2</sub>O

Probenpuffer                800 µl Sammelpuffer mit 80 mg Glycerin und  
15 µl einer gesättigten Bromphenolblaulösung in  
EtOH

Nach einem Vorlauf mit 15 mA/Gel für 10 min werden 1,5 µl der Proben aufgetragen und die Gele bei konstanter Stromstärke für 70 min entwickelt.

### 2.12.3 Färben von Polysacchariden in DOC-PAGE-Gelen

Für die Alcian Blau-Behandlung (Reuhs *et al.*, 1998) wurden die Gele sofort nach dem Gellauf wie folgt behandelt:

- 30 min sanftes Schütteln in 0,005% Alcian Blau (in 40% ETOH und 5% HAc)
- erneuern der Alcian Blau-Lösung und über Nacht schütteln
- waschen in Aqua bidest für 1 min
- oxidieren der Polysaccharide mit 0,7% Natriummetapersulfat
- 5-mal für 5 min waschen in Aqua bidest

Anschließend wurde eine Silberfärbung durchgeführt:

Imprägnieren	800 µl 20% AgNO <sub>3</sub> , 600 µl 37% HCHO in 60 ml Aqua bidest.	8 min
Spülen	Aqua bidest.	3 x 5 sec
Entwickeln	1,2 g Na <sub>2</sub> CO <sub>3</sub> , 25 µl 37% HCHO, 25 µl 10% Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> * 5 H <sub>2</sub> O in 60 ml Aqua bidest.	10-20 sec
Stopp	1% HAc in Aqua bidest.	30 sec
Spülen	Aqua bidest.	3 x 15 sec

## 2.13 Isolierung von Metaboliten

### 2.13.1 Extraktion von Metaboliten

#### 2.13.1.1 Extraktion durch Umesterung

Wasserfreier saurer Methanol wurde durch das Einleiten von HCl-Gas in Methanol erzeugt. Zellpellet aus 1 ml M9-Kultur des zu analysierenden *E. coli* Stamms wurden mit 100 µl wasserfreiem saurem Methanol (1 N) versetzt und in einem verschlossenen Glasgefäß 20 min bei 80°C inkubiert. Dann wurden 100 µl 2 M NaCl zugegeben und die Probe 2-mal mit 600 µl CH<sub>2</sub>Cl<sub>2</sub> extrahiert. Die vereinigten organischen Phasen wurden am Stickstoffstrom getrocknet, die Rückstände in 100 µl CH<sub>2</sub>Cl<sub>2</sub> gelöst und bei -20°C eingefroren.

### 2.13.1.2 Extraktion mit Ethylacetat

Für die Extraktion wurde das Zellpellet aus 2 ml M9-Kultur des zu analysierenden *E. coli*-Stamm 2-mal mit 6 ml Ethylacetat (mit 1% HAc) extrahiert. Die vereinigten organischen Phasen wurden am Stickstoffstrom getrocknet, die Rückstände in 100  $\mu$ l  $\text{CH}_2\text{Cl}_2$  gelöst und bei  $-20^\circ\text{C}$  eingefroren. Für die Extraktion von größeren Kulturen wurde am Ende der Induktionsperiode der Kultur ein Vol. Ethylacetat (1% HAc) zugesetzt und über Nacht geschüttelt. Am nächsten Tag wurde die Kultur nochmal mit 2 Vol. Ethylacetat (1% HAc) extrahiert und die organischen Phasen vereint.

### 2.13.2 Dünnschichtchromatografische Auftrennung von Extrakten

Die Extrakte wurden mit Hilfe von Kristallspitzen auf den unteren Bereich einer Dünnschichtchromatografie-Platte (HPTLC silica gel 60 F254 von Merck, Darmstadt; oder Reverse Phase C18-Silica Chromatografie-Platten) aufgetragen. Als Laufmittel wurden verwendet:  $\text{CHCl}_3$ :MeOH:H<sub>2</sub>O (14:6:1);  $\text{CHCl}_3$ :MeOH:HAc (13:5:2); Hexan:Diethylether (85:15); Hexan:Diethylether:HAc (85:15:1); Acetonitril:H<sub>2</sub>O (55:45); die obere Phase des Gemisches aus Ethylacetat:EtOH: 15% Amoniumoacetat (pH 9,6) (9:4:8) oder Ethylacetat:EtOH:HAc (17:2:1). Die HPTLC-Platten wurden nach dünnschichtchromatografischen Auftrennung für 1 h getrocknet und dann autoradiografiert.

### 2.13.3 [ $1\text{-}^{14}\text{C}$ ]Acetat-Markierung

Zur Markierung von Metaboliten mit [ $1\text{-}^{14}\text{C}$ ]Acetat wurden die entsprechenden *E. coli*-Stämme in M9-Medium gezüchtet. 20 ml M9 wurden mit 200  $\mu$ l aus einer Überernachtkultur angeimpft und die  $\text{OD}_{600}$  verfolgt. Bei einer  $\text{OD}_{600}=0,28$  wurden die Kulturen mit IPTG (100  $\mu$ M) induziert. Teilen oder der gesamten Kulturen wurde dann 2  $\mu$ l [ $1\text{-}^{14}\text{C}$ ]Acetat (7,4 kBq/ $\mu$ l) pro ml

Kulturvolumen zugegeben. Nach 4 Stunden wurden die Zellen geerntet und das Pellet zweimal mit M9-Medium gewaschen. Die Pellets wurden dann bei  $-20^{\circ}\text{C}$  eingefroren.

## 3 Ergebnisse

### 3.1 Klonierungen und Sequenzierungen

Wenn keine anderen Angaben gemacht werden, so bezieht sich die Nummerierung von DNA-Bereichen auf den Genbankeintrag X64131 (Petrovics *et al.*, 1993)

#### 3.1.1 Klonierungen zur heterologen Expression

##### 3.1.1.1 Klonierung von *rkpABCDEF* zur heterologen Expression

###### 3.1.1.1.1 Klonierungsstrategie

Um die Gene *rkpABCDEF* unter der Kontrolle des T7-Promotors induzierbar in *E. coli* zu exprimieren, sollten diese in den Vektor pET9a kloniert werden. Da eine Klonierung des gesamten DNA-Bereichs aufgrund von internen Schnittstellen für *Bam*HI nicht möglich ist, waren drei aufeinander folgende Klonierungsschritte notwendig (Abb. 1). Für einen ersten Klonierungsschritt sollte mit Hilfe der Polymerasekettenreaktion (PCR) und entsprechender Oligonukleotide (oTB3 und oTB4) der Anfang des *rkpA*-Gens (367 - 1387) so amplifiziert werden, dass am Anfang des Segmentes eine *Nde*I-Schnittstelle und am Ende eine *Bam*HI-Schnittstelle entstand. Das PCR-Fragment wurde mit *Nde*I und *Bam*HI verdaut und in die *Nde*I/*Bam*HI-Schnittstellen des pET9a-Vektors kloniert um ein pET9a-Derivat, (Plasmid I) zu erhalten, welches den Anfangsteil des *rkpA*-Gens (*rkpA'*) enthält. Die durch PCR eingefügte *rkpA'*-Teilsequenz wurde durch DNA-Sequenzierung bis zur *Kpn*I-Schnittstelle (1044) auf Fehlerfreiheit überprüft. In einem zweiten Klonierungsschritt sollte mit Hilfe der PCR und entsprechender Oligonukleotide (oTB5 und oTB2) das Ende des *rkpE*-Gens und das *rkpF*-Gen (*rkpE'F*) so amplifiziert werden, dass am Anfang des Segmentes eine *Kpn*I-Schnittstelle

und am Ende eine *Bam*HI-Schnittstelle entstand. Das PCR-Fragment wurde mit *Kpn*I und partiell mit *Bam*HI verdaut und ein 0,82 kb-Fragment wurde in die *Kpn*I-*Bam*HI-Schnittstelle des Plasmides I kloniert, um ein *rkpA*'-, *rkpE*'F-enthaltendes pET9a-Derivat (Plasmid II) zu erhalten. Die durch PCR eingefügte *rkpE*'F-Teilsequenz wurde durch DNA-Sequenzierung bis zur *Bcl*I-Schnittstelle (6714) auf Fehlerfreiheit überprüft. In einem dritten Klonierungsschritt sollte das 5,7 kb große *Kpn*I/*Bcl*I-Fragment der Komplementationsgruppe-I der *rkp*-1-Region in die *Kpn*I/*Bcl*I-Schnittstellen des Plasmides II kloniert werden, um ein Konstrukt (Plasmid III) zu erhalten, bei dem die gesamte Komplementationseinheit-I der *rkp*-1-Region unter der Kontrolle des T7-Promotors steht.

#### **3.1.1.1.2 Klonierung von *rkpA*'**

Mit den Oligonukleotiden oTBC3 und oTBC4 und pPP428 (Putnoky *et al.*, 1988) als Template-DNA wurde ein *rkpA*-enthaltendes PCR-Fragment erzeugt, welches mit *Nde*I/*Bam*HI zuerst in pUC19 (pTB1005) kloniert wurde. Die Sequenzierung dieses Plasmides zeigte Abweichungen von der publizierten Sequenz (Abb. 2). Eine Überprüfung ergab, dass die Sequenz in pTB1005 mit der des komplementierenden Kosmides pPP428 übereinstimmte. Anschließend wurde *rkpA*' in pET9a umklontiert (pTB1007). Die veränderte Sequenz führt zu einem Frameshift in einem Bereich von 6 Aminosäuren, sowie zu einer zusätzlichen Aminosäure in RkpA (Abb. 2)

#### **3.1.1.1.3 Klonierung von *rkpE*'F**

Das mit den Oligonukleotiden oTBC 2 und oTBC 8 amplifizierte 867 bp PCR-Fragment wurde in zwei Schritten kloniert. Das PCR-Fragment wurde mit *Kpn*I/*Bam*HI geschnitten und das entstandene 441 bp-Fragment mit *Kpn*I/*Bam*HI in pTB1007 kloniert. Dieses Plasmid pTB1008 enthält *rkpA*'E''F. In dieses Plasmid wurde dann das 426 bp *Bam*HI-Fragment des PCR-Fragmentes kloniert. Die Orientierung des Inserts wurde durch Restriktionsverdaus

kontrolliert. Das korrekte Plasmid pTB1014 wurde durch Sequenzierung überprüft. Auch hier wurde eine Abweichung vom Genbankeintrag gefunden (Abb. 3). Die gefundene Sequenz konnte auch im komplementierenden Kosmid pPP428 bestätigt werden.

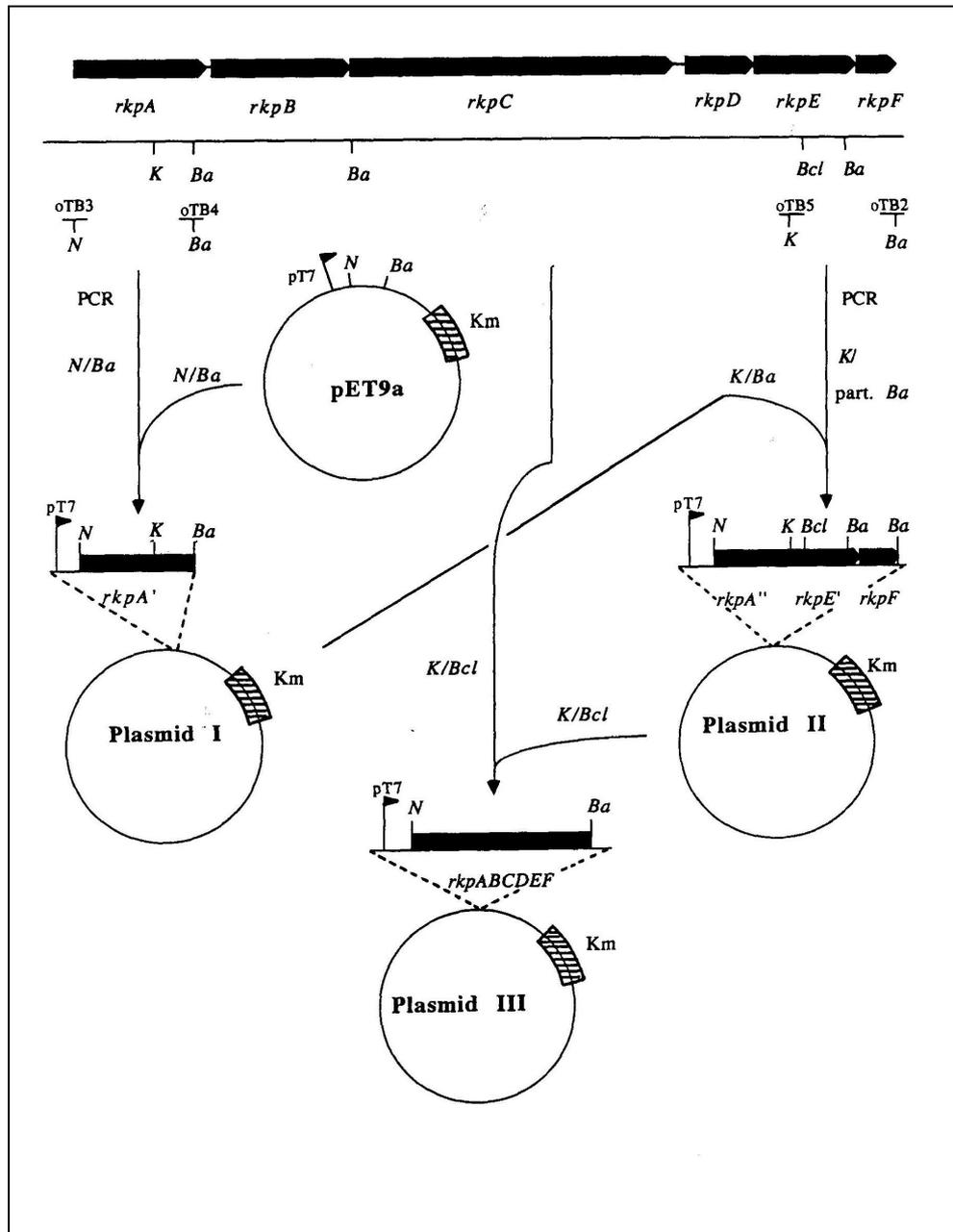


Abbildung 1: Konstruktion eines *rkpABCDEF*-enthaltenden Überexpressionsvektors. Die einzelnen Klonierungsschritte sind im Text beschrieben (3.1.1.1.1). Antibiotika-Resistenzen sind als schraffierte Segmente, *rkp*-Gensequenzen als geschwärzte Bereiche dargestellt. Fähnchen deuten Promotorsequenzen an. Abkürzungen: Km, Kanamycin; pT7, Promotor der von T7-Polymerase erkannt wird. Restriktionsstellen *KpnI* (*K*), *Bam*HI (*Ba*), *BclI* (*Bel*) und *NdeI* (*N*) sind angedeutet.

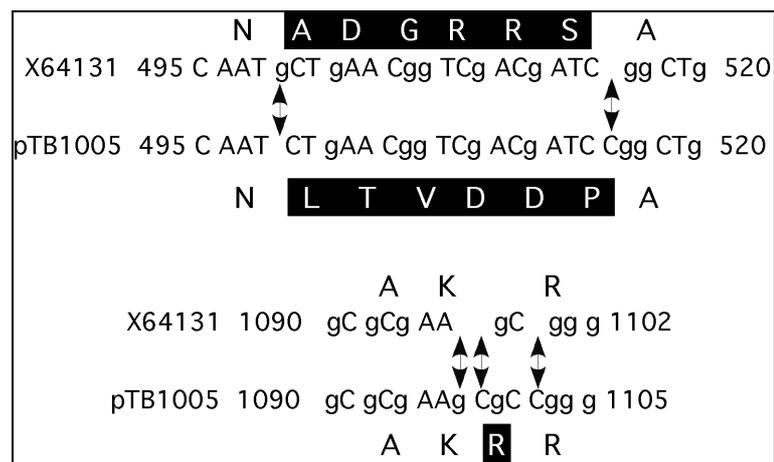


Abbildung 2: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1005. Veränderte Nukleotide sind durch Pfeile gekennzeichnet. Die veränderten Aminosäuren sind schwarz hinterlegt.



Abbildung 3: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1014. Veränderte Nukleotide sind durch Pfeile gekennzeichnet. Die veränderte Aminosäure ist schwarz hinterlegt.

#### 3.1.1.1.4 Klonierung von *rkpA*<sub>367</sub>*BCDEF* in *pET9a*

Da *BclI* ein methylierungssensitives Restriktionsenzym ist, wurde die DNA für diese Klonierung aus dem *E. coli*-Stamm JM110 isoliert. Das 5,66 kb *KpnI/BclI* Fragment aus pTB1041 wurde in pTB1014 eingesetzt. Das erhaltene

Plasmid pTB1044 enthält die Gene  $rkpA_{367}BCDEF$  unter der Kontrolle des T7-Promotors.

#### **3.1.1.1.5 Klonierung von $rkpA_{262}BCDEF$ in pET9a**

Da eine erneute Analyse der ORFs in der Komplementationsgruppe-I der  $rkp$ -1-Region (3.1.2.2) auch einen alternativen Translationsstartpunkt für RkpA bei 262 vermuten ließ, wurde auch diese Variante zur heterologen Expression kloniert. Das 778 bp *NdeI/KpnI*-Fragment aus pTB1030 wurde in pTB1014 eingesetzt. Das resultierende Plasmid pTB 1040 enthält die Gene  $rkpA_{262}E'F$ . In pTB1040 wurde dann das 5,66 kb *KpnI/BclI* Fragment aus pTB1041 eingesetzt. Das erhaltene Plasmid pTB1047 enthält die Gene  $rkpA_{262}-F$  unter der Kontrolle des T7-Promotors.

#### **3.1.1.1.6 Klonierung von $rkpA_{114}BCDEF$ in pET9a**

Ein weiterer alternativer Translationsstartpunkt für RkpA bei 114 (siehe 3.1.2.2) wurde ebenfalls zur heterologen Expression kloniert. Das 1 kb *XbaI/KpnI*-Fragment aus pTB1085 wurde in pTB1044 eingesetzt. Das resultierende Plasmid pTB1090 enthält die Gene  $rkpA_{114}BCDEF$  unter der Kontrolle des T7-Promotors.

### **3.1.1.2 Klonierung einzelner ORFs für die heterologe Expression**

#### **3.1.1.2.1 Klonierung von $rkpA$**

Ein 840 bp großes *KpnI/BglI*-Fragment des PCR-Produktes der Oligonukleotide oTBC7 und oTBC100 wurde in die *KpnI/BamHI*-Schnittstellen von pTB1007 kloniert. Das resultierende Plasmid pTB1017 enthält das  $rkpA_{367}$ -Gen unter der Kontrolle des T7-Promotors. In diesem Plasmid wurden dann mit

*NdeI/KpnI* die alternativen Startpunkte von *rkpA* ausgetauscht. So enthält pTB1030 das *rkpA*<sub>262</sub>-Gen und pTB1085 das *rkpA*<sub>114</sub>-Gen

### 3.1.1.2.2 Klonierung von *rkpF<sub>L</sub>*

Die Analyse der ORFs in der Komplementationsgruppe I der *rkp-1*-Region (3.1.2.2) zeigte, dass auch eine um 4 Aminosäuren größere Variante des RkpF-Proteins möglich ist. Wenn RkpF in den Konstrukten mit *rkpA-F* exprimiert wird, sollte sich RkpF in seiner natürlichen Form bilden. Um zu sehen ob die 4 zusätzlichen Aminosäuren das Laufverhalten in der Nativen-PAGE beeinflussen wurde die *rkpF<sub>L</sub>*-Variante ebenfalls kloniert. Dafür wurde das PCR-Produkt der Oligonukleotide oTBC2 und oTBC113 in den pET9a-Vektor kloniert. Die Sequenz im Plasmid pTB1025 wurde durch Sequenzierung überprüft.

## 3.1.2 Analyse der ORFs in der Komplementationsgruppe I

### 3.1.2.1 Sequenzabweichungen zum Genbankeintrag X64131

Neben den in den Abbildungen 2 und 3 bereits gezeigten Sequenzabweichungen traten weitere auf. Bei der Kontrolle des Plasmides pTB1041 mit verschiedenen Restriktionsenzymen fiel auf, dass das Bandenmuster des *Bam*HI-Verdau nicht den Erwartungen entsprach. Die Sequenzierung ergab die in Abb. 4 dargestellten Abweichungen vom Genbankeintrag. Da eine der Veränderungen im mutmaßlichen Stoppcodon von RkpB auftrat, wurde die gesamte Komplementationsgruppe-I einer erneuten *in silico*-Analyse unterzogen. Bei einer BLAST-Analyse fiel auf, dass in einem andern Leseraster auch „Upstream“ des publizierten Startpunktes von RkpA noch eine Homologie zu Polyketidsynthasen bestand (Abb. 5). Eine Sequenzierung des entsprechenden Bereiches ergab, dass durch eine zusätzliche Base (G210) im Genbankeintrag eine Verschiebung des Leserasters vorlag.

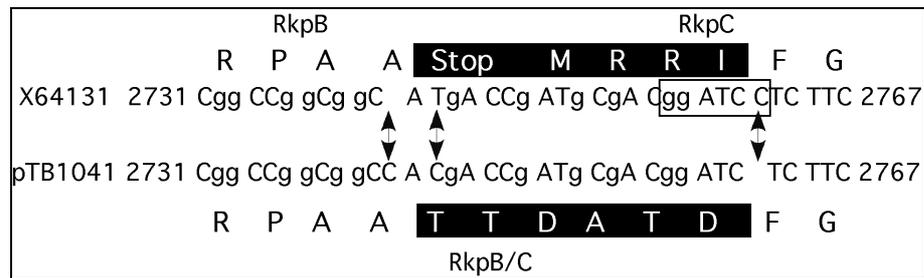


Abbildung 4: Sequenzunterschiede zwischen dem Genbankeintrag (X64131) und dem Plasmid pTB1041. Veränderte Nukleotide sind durch Pfeile gekennzeichnet. Die veränderten Aminosäuren sind schwarz hinterlegt. Die *Bam*HI-Schnittstelle ist eingerahmt.

```

>gi|31044143|gb|AAP42855.1| NanA1 [Streptomyces nanchangensis]
Length = 2902

Score = 32.7 bits (73), Expect = 2.6
Identities = 13/27 (48%), Positives = 19/27 (70%)
Frame = +3

Query: 120 VEIIGRACLAPGAKSPQALFKILRQGK 200
      V I+G AC PGA P+A +++LR G+
Sbjct: 24  VAIVGLACRLPGADPEAFWRLLRAGE 50

>gi|49146124|ref|YP_025562.1| Type I modular polyketide synthase [Mycobacterium ulcerans]
gi|42414758|emb|CAE46851.1| Type I modular polyketide synthase [Mycobacterium ulcerans]
Length = 16990

Score = 32.0 bits (71), Expect = 4.4
Identities = 12/26 (46%), Positives = 17/26 (65%)
Frame = +3

Query: 120 VEIIGRACLAPGAKSPQALFKILRQG 197
      V ++G +C PGA P AL+ +LR G
Sbjct: 20  VAVVGMSCRVPGASDPDALWALLRDG 45

```

Abbildung 5: Ausschnitt eines BLASTx-Ergebnisses (Altschul *et al.*, 1990) mit dem Bereich 1-200 des Genbankeintrags (X64131) als Vergleichssequenz. Gezeigt sind die Alignments der abgeleiteten Proteinsequenz mit NanA1 (Sun *et al.*, 2003) und einer modularen Typ-I-Polyketidsynthase aus *Mycobacterium ulcerans* (Stinear *et al.*, 2004).

### 3.1.2.2 *In silico*-Analyse der Komplementationsgruppe-I

#### 3.1.2.2.1 Analyse der Genbanksequenz X64131

Die Analyse mit verschiedenen Programmen zum Auffinden von ORFs lieferten übereinstimmende Ergebnisse. Diese übereinstimmenden Ergebnisse wichen jedoch von den veröffentlichten Daten ab (Petrovics *et al.*, 1993). Im

Einzelnen wurde für RkpA ein Startcodon bei 262, für RkpD bei 5580 und für RkpF bei 7221 vorhergesagt.

### 3.1.2.2 Analyse der korrigierten Sequenz

Die Computeranalyse der korrigierten Sequenz weicht nur an Stellen von der Analyse der Originalsequenz ab, wo es zu Verschiebungen im Leseraster kam. So beginnt der ORF für RkpA bei 114, und durch die Sequenzänderung im vorhergesagten Stoppkodon von *rkpB* entsteht ein neuer großer ORF für RkpB/C. Die Unterschiede zwischen der veröffentlichten und der neu vorgeschlagenen Organisation der Komplementationsgruppe-I sind in Abb. 6 dargestellt und in Tabelle 4 zusammengefasst.

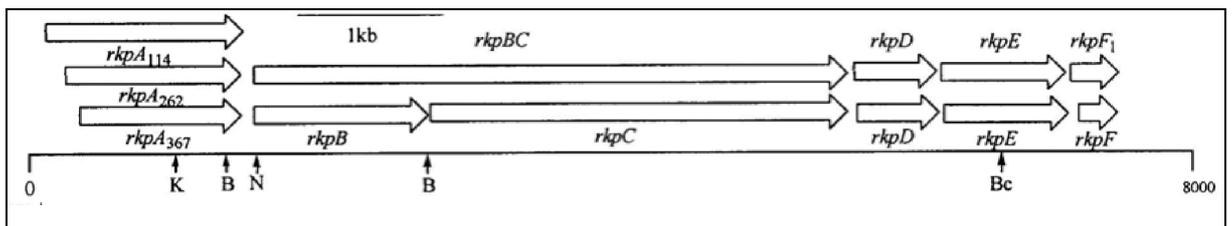


Abbildung 6: Mögliche Offene Leseraster im *rkpABCDEF*-Gencluster. Die unterste der drei abgebildeten Möglichkeiten entspricht dem Genbankeintrag X64131 (Petrovics *et al.*, 1993). Abkürzungen: K, *KpnI*; B, *Bam*HI; N, *Nde*I; Bc, *Bcl*I

Tabelle 4: Gegenüberstellung der veröffentlichten und der durch erneute Sequenzierung und *in silico*-Analyse bestimmten ORFs in der Komplementationsgruppe-I der rkp-1-Region

ORF	Genbankeintrag X64131				Korrigierte Sequenz			
	Start	Stopp	AA	MW in Da	Start	Stopp	AA	MW in Da
<i>rkpA</i>	367	1485	372	39061	114	1487	457	48575
<i>rkpB</i>	1558	2745	395	42073	-	-	-	-
<i>rkpB/C</i>	-	-	-	-	1560	5648	1363	145620
<i>rkpC</i>	2749	5649	966	103517	-	-	-	-
<i>rkpD</i>	5730	6296	188	20324	-	-	-	-
<i>rkpD<sub>i</sub></i>	-	-	-	-	5582	6298	238	25152
<i>rkpE</i>	6319	7173	284	29999	6321	7175	284	29973
<i>rkpF</i>	7236	7505	89	9587	-	-	-	-
<i>rkpF<sub>i</sub></i>	-	-	-	-	7223	7507	94	10116

### 3.1.3 Analyse des komplementierenden Kosmides pPP428

Die Untersuchungen zur Biosynthese kapsulärer Polysaccharide verschiedener Mutanten von AK631 (siehe 3.4.1.1) lieferte ungewöhnliche Ergebnisse. Die beschriebenen Mutanten konnten nicht durch die individuellen Gene der vorhergesagten Komplementationsgruppen trans-komplementiert werden (Kiss *et al.*, 1997; Petrovics *et al.*, 1993). Peter Putnoky machte eine ähnliche Beobachtung. Zum Beispiel konnte er die Phagensensitivität der Mutante PP2756 (*rkpF::kan*) nicht durch *rkpF*-Komplementierung in trans wieder herstellen (persönliche Mitteilung). Dass RkpF jedoch in PP2915 exprimiert wird, konnte gezeigt werden (siehe 3.2.1.2). Diese Ergebnisse zusammen mit der in der Literatur beschriebenen Mutante 649 (Putnoky *et al.*, 1990) ließen vermuten, dass sich auf dem Kosmid pPP428 (Putnoky *et al.*, 1988) noch weitere Gene befinden, die an der Biosynthese der kapsulären Polysaccharide beteiligt sind.

#### 3.1.3.1 Subklonierung und Sequenzierung von pPP428

Das Kosmid pPP428 wurde durch einen partiellen *EcoRI*-Verdau genomischer DNA von AK631 erzeugt (Putnoky *et al.*, 1988). Um die gesamte Sequenz des Inserts von pPP428 zu erhalten, wurde ein kompletter *EcoRI*-Verdau durchgeführt und alle Fragmente in pUC19 kloniert. In Tabelle 5 sind

die zur Sequenzierung genutzten Plasmide, die Größen der in ihnen enthaltenen Fragmente und deren Position in der Gesamtsequenz zusammengefasst. Die Übergänge der Plasmide wurden direkt im Kosmid sequenziert. Das Kosmid pPP428 hat ein Insert von 29875 bp. Der Genbankeintrag X64131 umfasst den Bereich 7788-20755 der pPP428-Sequenz. Die Bereiche 1-4241 und 22080-29875 wurden einzelsträngig sequenziert. Die doppelsträngig sequenzierten Bereiche 4241-7788 (AF247710) und 20755-22080 (AF247711) wurden in der NCBI-Genbank eingetragen (siehe 7.2 und 7.3).

Tabelle 5: Subklone von pPP428 und ihre Position in der Gesamtsequenz.

Plasmid	Klonierung	Bereich in pPP428	Insertgröße in bp
pTB1074	<i>EcoRI</i>	1-1338	1338
pTB1073	<i>EcoRI</i>	1333-2758	1428
pTB1077	<i>EcoRI</i>	2752-5587	2836
pTB1041	<i>HindIII</i>	5155-15332	10178
pTB1078	<i>EcoRI</i>	12923-20750	7827
pTB1063	<i>EcoRI</i>	20751-21215	462
pTB1064	<i>EcoRI</i>	21209-21905	698
pTB1069	<i>EcoRI</i>	21898-22918	1021
pTB1098	<i>EcoRI</i>	229822-25857	2874
pTB1065	<i>EcoRI</i>	25852-26407	557
pTB1072	<i>EcoRI</i>	26421-19332	2901

### 3.1.3.2 Analyse der pPP428-Sequenz

Bei der Computeranalyse der kompletten sinorhizobiellen Sequenz des Kosmides pPP428 konnten zwei weitere ORFs mit Homologien zu Genen, die an der Biosynthese von kapsulären Polysacchariden beteiligt sind, identifiziert werden. Der erste ORF (*rkpU*) liegt mit divergenter Transkriptionsrichtung upstream von *rkpA*. Dieser ORF zeigt Homologien zu einer Familie von Lipoproteinen (PROSITE: PDOC00013). Für zwei Mitglieder der Lipoproteinfamilie, *ctrl* aus *Neisseria meningitidis* (Frosch *et al.*, 1992) und *wza* (K30) aus *E. coli* (Drummelsmith & Whitfield, 2000) konnte gezeigt werden, dass sie am Transport von KPS über die äußere Membran beteiligt sind. Im Gegensatz zur bisherigen Vorstellung, dass das von den Proteinen RkpABCDEF gebildete Ketid als Anker für das zu synthetisierende KPS

fungiert, könnte alternativ dieses Ketid auch der Anker des sinorhizobiellen Lipoprotein RkpU sein. Der zweite ORF, der in divergenter Transkriptionsrichtung downstream von *rkpJ* liegt, hat Homologien zum *kpsF*-Gen aus *E. coli*. Die genaue Funktion von *kpsF* ist bisher unklar. Es konnte jedoch in *E. coli* gezeigt werden, dass *kpsF* essenziell für den Transport der KPS ist (Cieslewicz & Vimr, 1996; Cieslewicz & Vimr, 1997; Simpson *et al.*, 1996). Die Ergebnisse der Sequenzanalyse des komplementierenden Kosmides pPP428 sind in Abb. 7 und Tabelle 6 zusammengefasst.

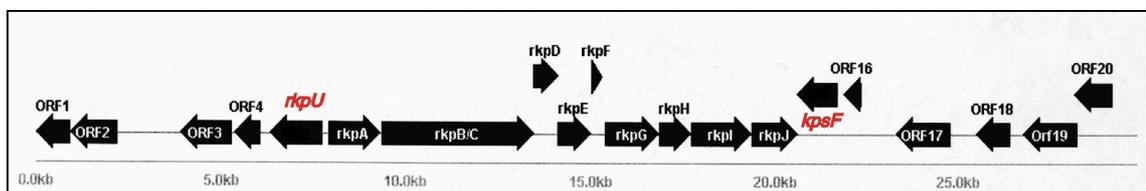


Abbildung 7: Analyse der Offenen Leserahmen und der bekannten Gene in pPP428. Alle ORFs sind in Tabelle 6 beschrieben.

Tabelle 6: Offene Leserahmen auf dem Kosmid pPP428

Gen	Länge [bp]/[AS]	mögliche Funktion	Identität der Aminosäuresequenz [%] zu:	Referenz
ORF 1	913/304 unvollständig	Kupfer-Bindeprotein	43 % <i>pcoA</i> aus <i>Escherichia coli</i> ; Acc.-Nr.: S52253 (Brown <i>et al.</i> , 1995)	Sequenzvergleich
ORF 2	1242/414	Nickel- und Kobalt-Resistenzprotein	21 % <i>CnrC</i> aus <i>Alcaligenes eutrophus</i> Acc.-Nr.: P37974 (Liesegang <i>et al.</i> , 1993)	Sequenzvergleich
ORF 3	1353/451	Nicotinat-Phosphoribosyltransferase	33 % <i>pncB</i> aus <i>Salmonella typhimurium</i> , Acc.-Nr.: P22253 (Vinitsky <i>et al.</i> , 1991)	Sequenzvergleich
ORF 4	660/220	Pyrazinamidase/Nicotinamidase	45 % <i>pncA</i> aus <i>Escherichia coli</i> ; Acc.-Nr.: P21369 (Jerstrom <i>et al.</i> , 1989)	Sequenzvergleich
<i>rkpU</i>	1416/472	Capsular Polysaccharide Export Outer Membrane Protein	32 % <i>ctrl</i> aus <i>Neisseria meningitidis</i> , Acc.-Nr.: P32013 (Frosch <i>et al.</i> , 1992)	diese Arbeit
<i>rkpA</i>	1374/ 458	Beta-Ketoacyl-Synthase	42 % zum Modul 3 der Erythronolid-Synthase aus <i>Saccharopolyspora erythraea</i> Acc.-Nr.: Q03132 (Bevitt <i>et al.</i> , 1992)	(Petrovics <i>et al.</i> , 1993)
<i>rkpB/C</i>	4092/1364	Acyltransferase/Dehydratase	37 % zum Modul 3 der Erythronolid-Synthase aus <i>Saccharopolyspora erythraea</i> , Acc.-Nr.: Q03132 (Bevitt <i>et al.</i> ,	(Petrovics <i>et al.</i> , 1993)

			1992)	
<i>rkpD</i>	564/188	Enoylreductase	50 % Polyketid-Synthase aus <i>Sorangium cellulosum</i> , Acc.-Nr.: AAF26921 (Molnar <i>et al.</i> , 2000)	(Petrovics <i>et al.</i> , 1993)
<i>rkpE</i>	852/284	Ketoreductase	38 % <i>mtaB</i> aus <i>Stigmatella aurantiaca</i> , Acc.-Nr.: AAF19810 (Silakowski <i>et al.</i> , 1999)	(Petrovics <i>et al.</i> , 1993)
<i>rkpF</i>	267/89	Acyl-Carrier-Protein	33% zum Modul 3 der Erythronolid-Synthase aus <i>Saccharopolyspora erythraea</i> , Acc.-Nr.: Q03132 (Bevitt <i>et al.</i> , 1992)	(Petrovics <i>et al.</i> , 1993), (Epple <i>et al.</i> , 1998)
<i>rkpG</i>	1338/446	Glycin-Acetyltransferase	31 % 2-Amino-3-Ketobutyrate Coenzyme A Ligase aus <i>Bacillus subtilis</i> , Acc.-Nr.: O31777	(Kiss <i>et al.</i> , 1997)
<i>rkpH</i>	837/279	Ribitol Type-Dehydrogenase	39% zur 17-beta-Hydroxysteroid-Dehydrogenase aus <i>Gallus gallus</i> , Acc.-Nr.: BAA19567 (Wajima <i>et al.</i> , 1999)	(Kiss <i>et al.</i> , 1997)
<i>rkpI</i>	1611/537	?	23% zu CgmB aus <i>S. meliloti</i> welches Phosphoglycerinreste von Phosphatidylglycerin auf Cyclische Beta-1,2-Glucane überträgt, Acc.-Nr.: AAB41532 (Wang <i>et al.</i> , 1999)	(Kiss <i>et al.</i> , 1997)
<i>rkpJ</i>	1236/412	?	31 % zu <i>kpsD</i> aus <i>Escherichia coli</i> , Acc.-Nr.: P42218 (Pazzani <i>et al.</i> , 1993)	(Kiss <i>et al.</i> , 1997)
<i>kpsF</i>	1077/359	?	37 % zu <i>kpsF</i> aus <i>Escherichia coli</i> , Acc.-Nr.: CAA64561 (Simpson <i>et al.</i> , 1996)	diese Arbeit
ORF 16	453/151	?	67 % zu hypothetischem 25.8 kD Protein Y4IC aus <i>Rhizobium sp.</i> NGR234, Acc.-Nr.: P55486 (Freiberg <i>et al.</i> , 1997)	Sequenzvergleich
ORF 17	321/107	Protein-Export-Protein	31 % zu <i>secDF</i> aus <i>Streptomyces coelicolor</i> A3(2), Acc.-Nr.: CAA22394 (Redenbach <i>et al.</i> , 1996)	Sequenzvergleich
ORF 18	879/293	Protein-Export-Protein	35 % zu <i>secD</i> aus <i>Sinorhizobium meliloti</i> , Acc.-Nr.: AAF34182 (Streit & Phillips, 1997)	Sequenzvergleich
ORF 19	1443/481	Acetolactate-Synthase	37 % zu <i>ilvB</i> aus <i>Escherichia coli</i> , Acc.-Nr.: P08142 (Wek <i>et al.</i> , 1985)	Sequenzvergleich
ORF 20	1014/338	Aspartat-Aminotransferase A (TransaminaseA)	32% zu <i>aatA</i> aus <i>Sinorhizobium meliloti</i> , Acc.-Nr.: Q02635 (Rastogi & Watson, 1991)	Sequenzvergleich

### 3.1.4 "Knock-out-Mutagenese"

Um die direkte Beteilligung der neu gefundenen ORFs an der Biosynthese der kapsulären Polysaccharide nachzuweisen, sollen Mutanten erzeugt werden, in denen diese durch die Insertion von Antibiotikaresistenzgenen unterbrochen sind. Die in diesem Absatz benutzten Angaben zu Positionen von Restriktionsschnittstellen beziehen sich auf die Sequenz des Inserts von pPP428 (siehe Anhang).

#### 3.1.4.1 Klonierungsstrategie zur Inaktivierung von *rkpU*

Ein 3427 bp *FspI*(5597)/*KpnI*(8824)-Fragment aus pTB1041 soll mit *KpnI*/*EcoRV* in pBS kloniert werden. In diesem Konstrukt soll ein 981 bp (6559-7540) großes *EcoRV*-Fragment aus *rkpU* (6300-7659) durch die Kanamycin-Resistenz-Kassette aus pTB3131 ersetzt werden. Die anschließende Umklonierung mit *BamHI*/*KpnI* in pMP3510 würde ein Plasmid ergeben, welches dann zur Inaktivierung von *rkpU* in *S. meliloti* AK631 genutzt werden könnte.

#### 3.1.4.2 Inaktivierung von *kpsF*

Das 3338 bp große *HindIII*(18837)/*XbaI*(22175)-Fragment von pPP428 wurde in pUC19 kloniert. Das resultierende Plasmid pTB1111 wurde dann mit *AgeI* (20820) verdaut. Die entstandenen Überhänge wurden mit der Klenow-Polymerase aufgefüllt und das ca 1kb große Kanamycin-Resistenz-vermittelnde Gen, welches mit *SmaI* aus pTB3131 entnommen worden war, eingesetzt. In dem entstehenden Plasmid ist der ORF *kpsF* (21657-20578) durch das Kanamycin-Resistenz-Gen unterbrochen. In pTB1120 ist die Richtung des Kanamycin-Resistenz-vermittelnden Gens entgegengesetzt der ursprünglichen Richtung des ersetzten *kpsF*-Genes, wogegen in pTB1121 beide ORFs in die gleiche Richtung transkribiert werden. pTB1120 wurde anschließend mit *HindIII* geschnitten und mit dem ebenfalls mit *HindIII* geschnittenen pMP3510

fusioniert. Das entstandene Fusionsplasmid pTB1123 wurde dann durch eine Drei-Punkt-Kreuzung in *S. meliloti* AK631 transkonjugiert. Leider konnten in den anschließenden Experimenten keine Bakterien isoliert werden, bei denen eine Rekombination stattgefunden hatte.

## **3.2 Proteinchemische Arbeiten**

### **3.2.1 Untersuchungen zu RkpF**

#### **3.2.1.1 Heterologe Expression von RkpF**

##### **3.2.1.1.1 Aufreinigung von RkpF**

Die bereits zuvor erarbeitete Aufreinigungsvorschrift (Epple, 1995) für das RkpF-Protein wurde verbessert, um eine homogene Präparation von RkpF zu erhalten. Bei einer  $OD_{600}=0,6$  wurde einer Kultur von *E. coli* BL21(DE3) pTB1003, IPTG (0,1 mM) zugegeben und die Zellen nach 4 Stunden weiterer Inkubation bei 37°C geerntet. Das Zellpellet wurde bei -20°C eingefroren. 1,1 g Zellen (Nassgewicht) wurden in 30 ml Aufschlusspuffer (50 mM Tris/HCl pH 6,8, 100 mM KCl) resuspendiert und durch drei French-Press-Passagen (20000 psi) aufgeschlossen. Nach einer Zentrifugation bei 6800xg wurde der Überstand abgenommen. Diesem Rohextrakt wurde dann unter Rühren bei 4°C tropfenweise Isopropanol, bis zu einer Endkonzentration von 60% zugegeben. Nach einer 1-stündigen Inkubation bei 4°C wurden die ausgefallenen Bestandteile durch Zentrifugation bei 6800xg abgetrennt und der Überstand in einen Rundkolben überführt. Anschließend wurde der Gehalt an Isopropanol am Rotationsverdampfer mit einer Wasserbadtemperatur von 40°C und einem Druck von 200 mbar auf unter 5% verringert. Der Lösung wurde dann CHAPS (1 mM) zugegeben und eine DE-52-Säulenchromatographie durchgeführt. RkpF enthaltende Fraktionen wurden gepoolt und aufkonzentriert. Es folgte eine Größenausschlusschromatografie mit einer G50-Superfine-Säule. Die RkpF enthaltenden Fraktionen wurden erneut gepoolt und aufkonzentriert. Die

Ausbeute einer solchen Aufreinigung waren 5,4 mg homogenes RkpF-Protein (Tab. 7). Abb. 8 zeigt eine 20%-Native-PAGE der verschiedenen Aufreinigungsschritte.

Tabelle 7: Ausbeute der Aufreinigung von RkpF

Aufreinigungsstufe	Proteinkonzentration in mg/ml	Volumen in ml	Gesamtprotein in mg
Rohextrakt	5,6	29	162
Isopropanolfällung	1,2	31	37
DE-52	1,2	6	7,2
G50	1,8	3	5,4

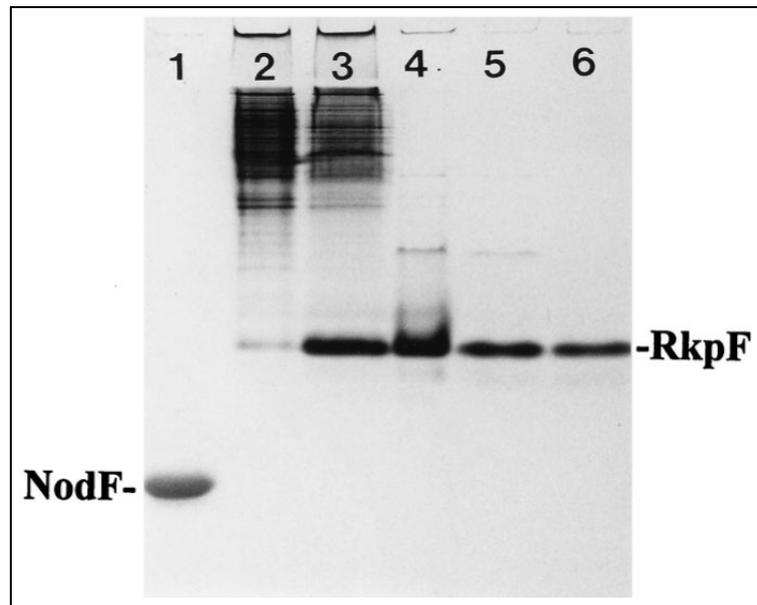


Abbildung 8: Überexpression und Reinigung des RkpF-Proteins. 20 % Native PAGE mit anschließender Coomassiefärbung. Analysiert wurden gereinigtes NodF-Protein (Ghose *et al.*, 1996)(Spur 1), BL21 (DE3) pTB1003 ohne (Spur 2) und mit (Spur 3) Induktion durch IPTG, sowie Proteinpools nach der Isopropanolfällung (Spur 4), der DE-52- (Spur 5) und der G50-Säulenchromatographie (Spur 6).

### 3.2.1.1.2 *In vivo*-Markierung von RkpF

Ein häufig auftretendes Problem bei der heterologen Expression von ACPs in *E. coli* ist, dass die posttranslationale Modifikation von der Apo- in die Holoform durch Übertragung der prosthetischen Gruppe 4'-Phosphopantetein nicht stattfindet. Um zu überprüfen ob RkpF vergleichbar zum NodF-Protein (Geiger *et al.*, 1991) in *E. coli* modifiziert wird, wurden die Zellen mit radioaktiv markiertem  $\beta$ -Alanin, das spezifisch in 4'-Phosphopantethein eingebaut wird, gezüchtet. Wie unter 2.11.11.1 beschrieben, wurden OG7001 pMP2301, OG7001 pET9a und OG7001 pTB1003 gezüchtet und mit Lysozym aufgeschlossen. Die Proteinextrakte wurden mit einer 20% nativen PAGE aufgetrennt und auf eine Nitrozellulosemembran geblottet. Die mit Autofluor (National Diagnostics) besprühte Membran wurde anschließend für 5 Wochen mit einem Röntgenfilm bei  $-80^{\circ}\text{C}$  inkubiert. Wie in Abb. 9 zu sehen ist wurde

RkpF *in vivo* mit  $\beta$ -[ $^3$ H]Alanin markiert. RkpF wird also von einer Phosphopantetheinyltransferase in *E. coli* posttranslational modifiziert.

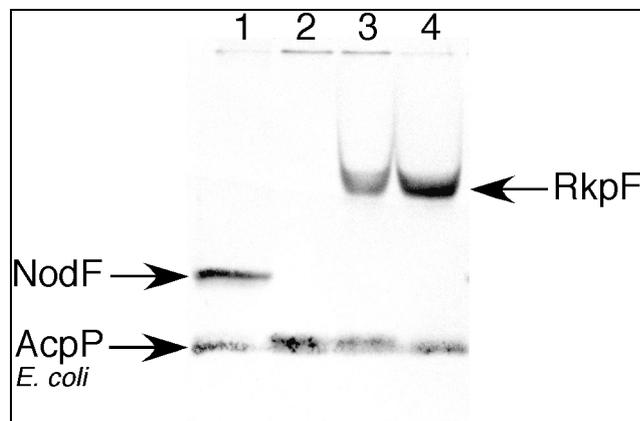


Abbildung 9: Zellextrakte von in Gegenwart von  $\beta$ -[ $^3$ H]Alanin kultivierten Zellen wurden mit einer 20% nativen PAGE aufgetrennt und anschließend autoradiografiert. Analysiert wurden Extrakte von OG7001 pMP2301 (Spur 1), OG7001 pET9a (Spur 2) und OG7001 pTB1003 (Spur 4), jeweils mit IPTG induziert und OG7001 pTB1003 ohne Induktion (Spur 3)

### 3.2.1.1.3 *In vitro*-Konversion von Apo- zu Holo-RkpF

Nachdem gezeigt werden konnte, dass RkpF in *E. coli* posttranslational modifiziert wird, sollte überprüft werden, ob diese Modifikation auch *in vitro* durch die gereinigte Holo-ACP-Synthase (AcpS) aus *E. coli* durchgeführt wird (Lambalot & Walsh, 1995). Dazu wurde gereinigtes RkpF wie in 2.11.10.6 beschrieben mit AcpS und CoA inkubiert. In einer zeitabhängigen Untersuchung konnte die quantitative Umsetzung von Apo-RkpF zu Holo-RkpF durch AcpS gezeigt werden (Abb. 10). Dass es sich bei den gezeigten Banden tatsächlich um die Apo- und Holo-Formen von RkpF handelt, konnte durch massenspektrometrische Untersuchungen, die von Jane Thomas-Oates, (Michael Barber Center for Mass Spectrometry, Department of Chemistry, UMIST, Manchester, UK) durchgeführt worden sind, bestätigt werden (Epple *et al.*, 1998).

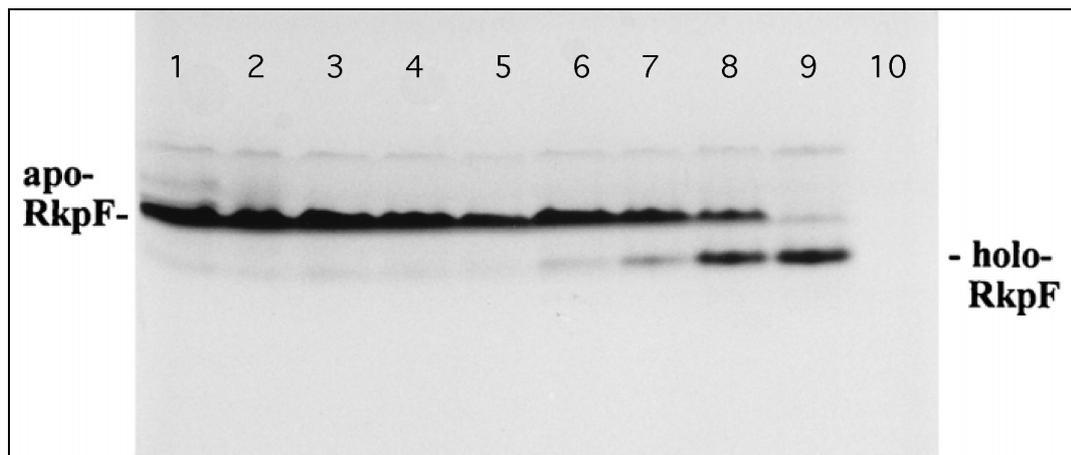


Abbildung 10: Zeitabhängige Umsetzung von Apo- zu Holo-RkpF durch die gereinigte Holo-ACP-Synthase AcpS aus *E. coli*. Die Reaktion wurde zu verschiedenen Zeitpunkten gestoppt und die Proben mit einer 20% nativen PAGE mit anschließender Coomassiefärbung analysiert. Apo-RkpF mit CoA und AcpS für 0 min (Spur 4), 8 min (Spur 5), 16 min (Spur 6), 32 min (Spur 7), 64 min (Spur 8) oder 128 min (Spur 9) inkubiert. Die anderen Proben wurden ebenfalls jeweils 128 min inkubiert, jedoch ohne CoA (Spur 1) oder ohne AcpS (Spur 2), in Gegenwart von EGTA (Spur 3) oder ohne Apo-RkpF (Spur 10)

### 3.2.1.2 Nachweis von RkpF in *Sinorhizobium meliloti*

In *Sinorhizobium meliloti* AK631 konnte kein RkpF-Protein nachgewiesen werden. Das liegt wahrscheinlich am sehr niedrigen Transkriptionsniveau des *rkpABCDEF*-Genklusters (Petrovics *et al.*, 1993). Bei Genen mit vergleichbar niedrigem Transkriptionsniveau in Reporter-Genstudien konnte das entsprechende Protein erst nach dem Einführen zusätzlicher Genkopien nachgewiesen werden (Herman Spaink, persönliche Mitteilung). Peter Putnoky (Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary) hat das *rkpF*-Gen in AK631 inaktiviert. Die erhaltene Mutante PP2756 war nicht in der Lage KPS zu produzieren, konnte jedoch nicht durch das *rkpF*-Gen trans-komplementiert werden (Peter Putnoky, persönliche Mitteilung). Die Mutanten mit dem *rkpF*-Gen in trans (PP2919 und PP2920) und der Wildtyp mit dem *rkpF*-Plasmid (PP2915 und PP2916) wurden auf die Expression von RkpF überprüft.

Abb. 11 zeigt, dass in AK631 und der *rkpF*-Mutante kein RkpF nachgewiesen werden kann. Jedoch konnte in allen Stämmen mit plasmidkodiertem RkpF das Protein nach Anreicherung durch Isopropanolfällung (siehe 2.11.10.7) detektiert werden. Die stärkste Expression zeigt PP2919 in dem sogar die Apo- und die Holo-Form nachzuweisen sind.

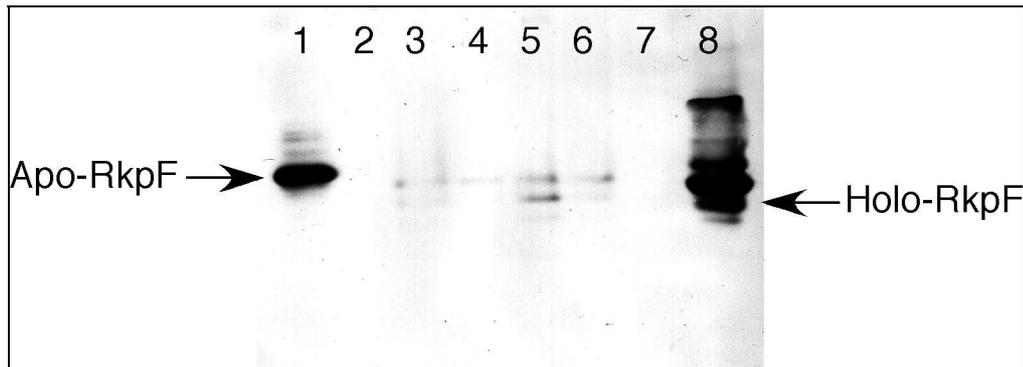


Abbildung 11: Western-Blot zum Nachweis von RkpF in verschiedenen Mutanten von *Sinorhizobium meliloti* AK631. Als Kontrollen wurden gereinigtes Apo-RkpF (Spur 1) und Holo-RkpF (Spur 8) aufgetragen. Die sonstigen Proben sind AK631 (Spur 2), PP2915 (Spur 3), PP2916 (Spur 4), PP2919 (Spur 5), PP2920 (Spur 6) und PP2756 (Spur 7).

### 3.2.2 Heterologe Expression von RkpA

Die Analyse der Offenen Leserahmen des *rkpABCDEF*-Genklusters (siehe 3.1.2.2.) zeigte drei mögliche Startpunkte für RkpA. Alle drei Varianten wurden in den pET9a-Vektor kloniert und in BL21(DE3) exprimiert. Bei der Expression von RkpA<sub>367</sub>, welches der publizierten Form (Petrovics *et al.*, 1993) entspricht, konnte keine zusätzliche Bande in der erwarteten Größe von 39061 Da gefunden werden (Abb. 12A). Im Gegensatz dazu wurden bei der Expression von RkpA<sub>262</sub> (Abb. 12B) und RkpA<sub>114</sub> (Abb. 12C) zusätzliche Protein-Banden mit den entsprechenden Massen von 42917 Da und 48575 Da nachgewiesen. Dabei fällt auf, dass die längste Form, deutlich stärker exprimiert wird als die mittlere Form. Dieses Verhältnis findet sich auch bei der heterologen Expression des gesamten Genklusters wieder. Auch der RkpABCDEF-

Proteinkomplex wird bei Verwendung der Sequenz ab Codon 114 stärker exprimiert als wenn mit Codon 262 begonnen wird (siehe 3.2.3.1).

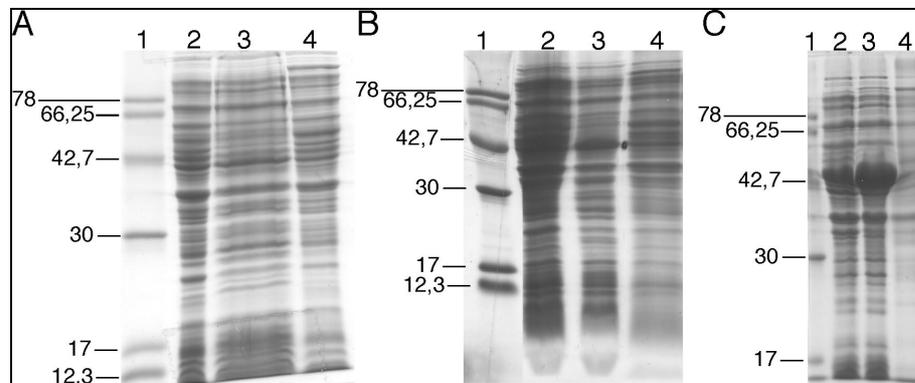


Abbildung 12: SDS-PAGE-Analyse der heterologen RkpA-Expression. A) BL21(DE3) pTB1017 (RkpA<sub>367</sub>) ohne (Spur 2) und mit Induktion durch IPTG (Spur 3), aufgetrennt auf einem 15% SDS-PAGE-Gel mit anschließender Coomassiefärbung. B) BL21(DE3) pTB1030 (RkpA<sub>262</sub>) ohne (Spur 2) und mit Induktion durch IPTG (Spur 3), aufgetrennt auf einem 15% SDS-PAGE-Gel mit anschließender Coomassiefärbung. C) BL21(DE3) pTB1085 (RkpA<sub>114</sub>) ohne (Spur 2) und mit Induktion durch IPTG (Spur 3), aufgetrennt auf einem 7,5% SDS-PAGE-Gel mit anschließender Coomassiefärbung. Als Kontrollen wurden jeweils Extrakte von BL21(DE3), die mit IPTG induziert worden waren, aufgetragen (Spuren 4). Die Größen der Molekulargewichtsstandards in kDa sind angegeben (Spuren 1).

### 3.2.3 Heterologe Expression von RkpABCDEF

In *Sinorhizobium meliloti* AK631 konnte gezeigt werden, dass von den Genen *rkpABCDEF* nur eine mRNA gebildet wird (Petrovics *et al.*, 1993). Somit kann die Bildung von RkpF als Indikator für die Translation des gesamten Genklusters genommen werden. Jedoch konnte in Rohextrakten von BL21(DE3) pTB1044 (*rkpA<sub>367</sub>BCDEF*) und BL21(DE3) pTB1047 (*rkpA<sub>262</sub>BCDEF*) keine RkpF-Bande in 20% nativen PAGE-Gelen oder Tricin-enthaltenden SDS-PAGE-Gelen detektiert werden.

### 3.2.3.1 Nachweis eines RkpF-enthaltenden Proteinkomplexes.

Falls es auf Grund eines frühzeitigen Abbruchs der Translation der singulären mRNA zu keiner RkpF-Synthese käme, sollten zumindest die anderen Rkp-Proteine gebildet werden. Bei der Analyse von Rohextrakten fiel auf, dass bei der Expression von *rkp<sub>262</sub>ABCDEF* von Plasmid pTB1047 und *rkp<sub>114</sub>ABCDEF* von Plasmid pTB1090 nach Induktion mit IPTG eine zusätzliche Proteinbande mit einem scheinbaren Molekulargewicht von ca. 250 kDa entsteht. Diese Masse könnte einem RkpABCDEF-Proteinkomplex (259,4 kDa) entsprechen. Diese große Proteinbande reagierte mit dem Antikörper gegen RkpF (Abb. 13). Vergleichbar zu den Untersuchungen zur heterologen Expression von RkpA wird auch der Proteinkomplex in der Form mit dem Startcodon bei 114 viel stärker gebildet (3.2.2). Bei der Expression von *rkp<sub>114</sub>ABCDEF* werden große Mengen des Proteinkomplexes in „Inclusion-Bodies“ gebildet.

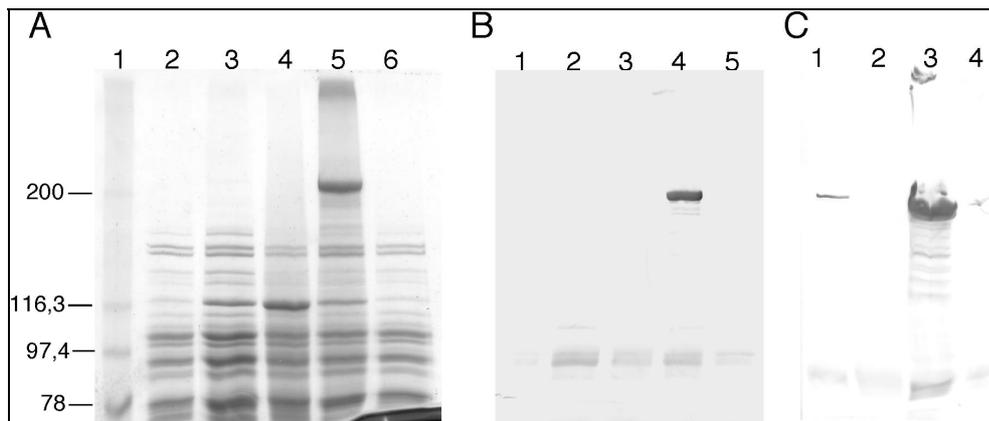


Abbildung 13. Western-Blot-Nachweis eines RkpF-enthaltenden Proteinkomplexes. Rohextrakte von *E. coli* mit verschiedenen Expressionplasmiden wurden auf 7,5% SDS-PAGE-Gelen aufgetrennt. Die Proteine wurden durch Coomassiefärbung (A) oder durch einen Western-Blot mit Anti-RkpF (1:10000) detektiert (B, C). Die Proben sind BL21(DE3) pTB1044 ( $rkpA_{367}BCDEF$ ) ohne (A Spur 2, B Spur 1) und mit IPTG-Induktion (A Spur 3, B Spur 2), BL21(DE3) pTB1047 ( $rkpA_{262}BCDEF$ ) ohne (A Spur 4, B Spur 3, C Spur 2) und mit IPTG-Induktion (A Spur 5, B Spur 4, C Spur 1), BL21(DE3) pTB1090 ( $rkpA_{114}BCDEF$ ) ohne (C Spur 4) und mit IPTG-Induktion (C Spur 3) sowie als Kontrolle BL21(DE3) mit IPTG-Induktion (A Spur 6, B Spur 5). Die Größen der Molekulargewichtsstandards in kDa sind angegeben (A Spur 1)

### 3.2.3.2 Biochemische Untersuchung des Proteinkomplexes

Unter den Bedingungen der SDS-PAGE sollten nur Proteinkomplexe mit kovalenten Bindungen erhalten bleiben. Als erstes wurde der Einfluss von DTT untersucht.

#### 3.2.3.2.1 Einfluss der DTT-Konzentration auf den RkpF-enthaltenden Proteinkomplex

RkpF könnte in seiner Holo-Form über die SH-Gruppe der prosthetischen Gruppe an die Cysteine eines anderen Proteins binden. Deshalb soll die Stabilität des Proteinkomplexes und seine Immunreaktivität mit dem Anti-RkpF-Antikörper bei verschiedenen DTT-Konzentrationen untersucht werden. Dafür wurden Rohextrakte von BL21(DE3) pTB1090 ( $rkpA_{114}BCDEF$ ) mit

verschiedenen DTT-Konzentrationen für 30 min bei RT inkubiert und die Proben auf 7,5% SDS-PAGE-Gelen analysiert. Abbildung 14A zeigt, dass nach Coomassiefärbung kein Intensitätsunterschied für die 250 kDa-Bande bei unterschiedlichen DTT-Behandlungen zu sehen ist. Abbildung 14B zeigt, dass die Immunreaktivität des Komplexes mit steigender DTT-Konzentration geringer wird. Entweder vermindert DTT die Immunreaktivität von RkpF oder RkpF ist durch einen von DTT beeinflussbaren Mechanismus an den Proteinkomplex gebunden.

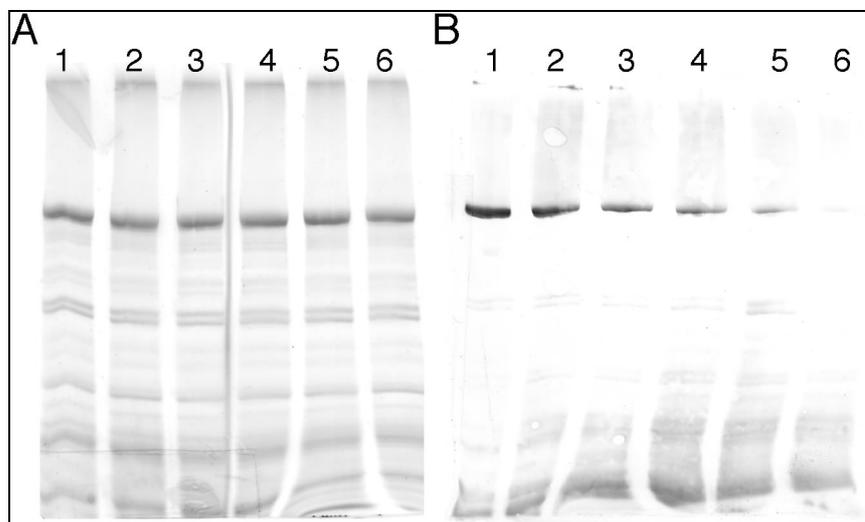


Abbildung 14: DTT-abhängige Stabilität und Immunreaktivität des RkpF-enthaltenden Proteinkomplexes. Rohextrakte von BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) wurden mit DTT inkubiert und auf 7,5% SDS-PAGE-Gelen analysiert. A) Coomassie-gefärbtes Gel. B) Western-Blot mit Anti-RkpF-Antikörper (1:10000). Die DTT-Konzentration waren 0 mM (Spur 1), 10 mM (Spur 2), 20 mM (Spur 3), 40 mM (Spur 4), 80 mM (Spur 5) und 160 mM (Spur 6).

### 3.2.3.2.2 Einfluss von Harnstoff auf den RkpABCDEF-Proteinkomplex

Der Proteinkomplex bildet bei der heterologen Expression „Inclusion Bodies“. Es besteht also die Möglichkeit, dass es sich nicht um einen Proteinkomplex handelt, sondern um miteinander interagierende Proteine, die zusammen ausgefallen sind. Möglicherweise reichte die denaturierende

Wirkung des SDS nicht aus, um diese Proteine voneinander zu trennen. Um dieser Frage nachzugehen, wurde eine SDS-PAGE mit 5 M Harnstoff im Gel durchgeführt. Proben von BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) wurden auf einem 7,5% SDS-Urea-PAGE-Gel aufgetrennt. Eine Hälfte des Gels wurde mit Coomassie gefärbt und die andere Hälfte geblottet und mit Anti-RkpF-Antikörper entwickelt. In Abb. 15 ist zu erkennen, dass auch die Gegenwart von 5 M Harnstoff den Proteinkomplex nicht auflöst, und auch RkpF unter diesen Bedingungen nicht herausgelöst wird.

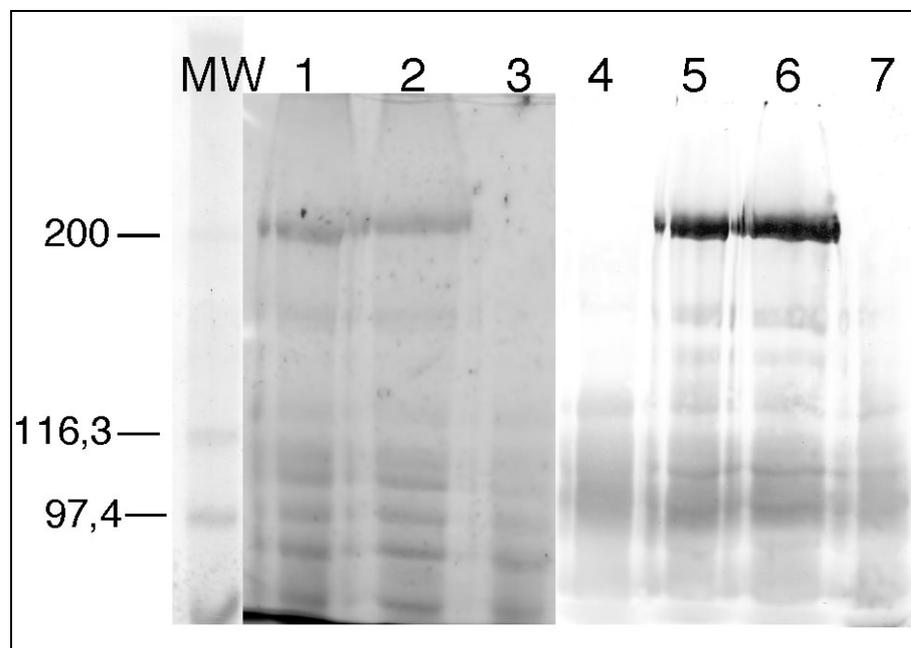
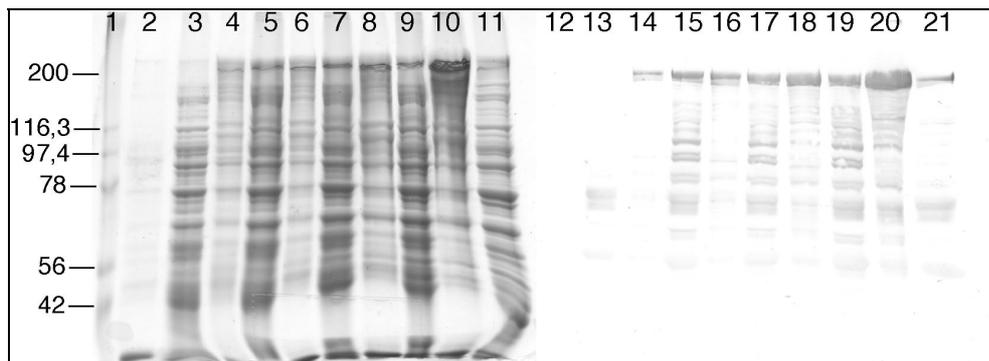


Abbildung 15: Untersuchung des Proteinkomplexes in Gegenwart von 5 M Harnstoff. Ein 7,5% SDS-Urea-PAGE-Gel wurde zum Teil mit Coomassie gefärbt und die andere Hälfte durch einen Western-Blot mit Anti-RkpF (1:10000) detektiert. BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) wurde ohne (Spur 1 und 5) und mit 160 mM DTT untersucht (Spur 2 und 6). Als Kontrollen dienten BL21 (DE3) (Spur 4) und BL21(DE3) pTB1003 (*rkpF*) (Spur 3 und 7). Die Größen der Molekulargewichtsstandards in kDa sind angegeben.

### 3.2.3.3 Expression von RkpABCDEF bei verschiedenen Kultivierungsbedingungen

Da es das Ziel war, das Produkt der heterolog exprimierten Proteine RkpABCDEF zu identifizieren, sollten die Proteine in ihrer natürlichen Faltung vorliegen, damit die Biosynthese des entsprechenden Produkts möglich ist. Die Ablagerung von Proteinen in „Inclusion Bodies „ ist häufig eine Folge einer unzureichenden Faltung der Proteine. Deshalb wurden verschiedene Faktoren untersucht, mit denen die Anreicherung des Proteinkomplexes in der löslichen Form erzielt werden sollte. Die Variation der IPTG-Konzentration hatte keinen Einfluss auf die Verteilung des Proteinkomplexes. Ebenso hatte die Koexpression der Chaparon-Komplexe GroEL/GroES und DnaK/DnaJ von einem induzierbaren Plasmid (Nishihara *et al.*, 1998) keinen Einfluss. Eine weitere häufig verwendete Möglichkeit zur Vermeidung von „inclusion bodies“ ist die Kultivierung der Zellen bei niedrigeren Temperaturen. Um den Einfluss der Kultivierungstemperatur zu überprüfen, wurden parallele Kulturen von BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) bei 37°C angezchtet. Bei einer  $OD_{600}=0,5$  wurden die Kulturen auf 15°C, 20°C, 25°C, 30°C und 37°C temperiert und bei einer  $OD_{600}=0,6$  mit 100 µM IPTG induziert. Nach 4 Stunden Inkubation bei den entsprechenden Kultivierungstemperaturen wurden die Zellen geerntet. Die Zellen wurden mit der French-Press aufgeschlossen und anschließend bei 40000xg zentrifugiert. Die Überstände und die resuspendierten Pellets wurden auf 10% SDS-PAGE-Gelen getrennt und anschließend entweder durch Coomassiefärbung oder Western Blot mit Anti-RkpF analysiert (Abb. 16). Es ist deutlich zu sehen, dass die Proteinmenge des Komplexes von der Kultivierungstemperatur abhängig ist. Bei 25°C und 30°C Kultivierungstemperatur ist ein gutes Expressionsniveau vorhanden und der Anteil des Proteinkomplexes im Überstand ist deutlich besser als bei 37°C. Bei 15°C wird nur sehr wenig Komplex gebildet, der anscheinend auch kein RkpF enthält.



Temperatur	15		20		25		30		37	
Fraktion	P	Ü	P	Ü	P	Ü	P	Ü	P	Ü
Spuren	2,12	3,13	4,14	5,15	6,16	7,17	8,18	9,19	10,20	11,21

Abbildung 16: Expression des RkpABCDEF-Komplexes bei verschiedenen Kultivierungstemperaturen. Rohextrakte von BL21(DE3) pTB1090 (*rkpA<sub>1-14</sub>BCDEF*), die bei den angegebenen Temperaturen induziert worden waren, wurden mit 40000xg zentrifugiert und die Überstände (Ü) und resuspendierten Pellets (P) auf 10% SDS-Gelen aufgetrennt und mit Coomassiefärbung und Western-Blot mit Anti-RkpF (1:10000) analysiert. Die Größen der Molekulargewichtsstandards in kDa sind angegeben (Spur 1)

### 3.2.3.4 *In vivo*-Markierung des Proteinkomplexes

#### 3.2.3.4.1 Markierung mit $\beta$ -[ $^3$ H]Alanin

Eine Voraussetzung für die Biosynthese des vermuteten  $\beta$ -Ketides durch den RkpABCDEF-Proteinkomplex ist die posttranslationale Modifikation von RkpF durch die Phosphopantetheinyltransferase AcpS. Für das einzeln heterolog exprimierte RkpF konnte diese Modifikation bereits nachgewiesen werden (siehe 3.2.1.1.2) (Epple *et al.*, 1998). Nun sollte untersucht werden, ob auch der Proteinkomplex *in vivo* mit  $\beta$ -[ $^3$ H]Alanin markiert werden kann. Dazu wurden Kulturen (1 ml) von OG7001 pTB1003 (*rkpF*) und OG7001 pTB1090 (*rkpA<sub>1-14</sub>BCDEF*) mit 23,5  $\mu$ l  $\beta$ -[ $^3$ H]Alanin (37 MBq/ml; 2,22 TBq/mMol) wie unter 2.11.11.1 beschrieben markiert. Die Rohextrakte des Lysozymbaufschlusses wurden auf einem 10% SDS-Gel aufgetrennt und auf Nitrocellulose geblottet. Der mit Autofluor behandelte Teil der Membran wurde mit einem Röntgenfilm

für 3 Tage (Abb.17A) oder 4 Wochen (Abb.17B) inkubiert. Der Proteinkomplex wurde auch durch  $\beta$ - $^3\text{H}$ Alanin markiert, woraus man schließen kann, dass RkpF im Komplex in seiner Holo-Form vorliegt. Die starke Überexposition (Abb. 17B) zeigte weitere markierte Banden bei OG7001 pTB1090 (Spur 6), die in der Kontrolle OG7001 pTB1003 (Spur 5) nicht entstanden. Hierbei könnte es sich um Fragmente oder Vorstufen des Proteinkomplexes handeln. Die Konzentration dieser Proteine reichte jedoch nicht aus, um RkpF unter den verwendeten Bedingungen mit dem Anti-RkpF-Antikörper zu detektieren.

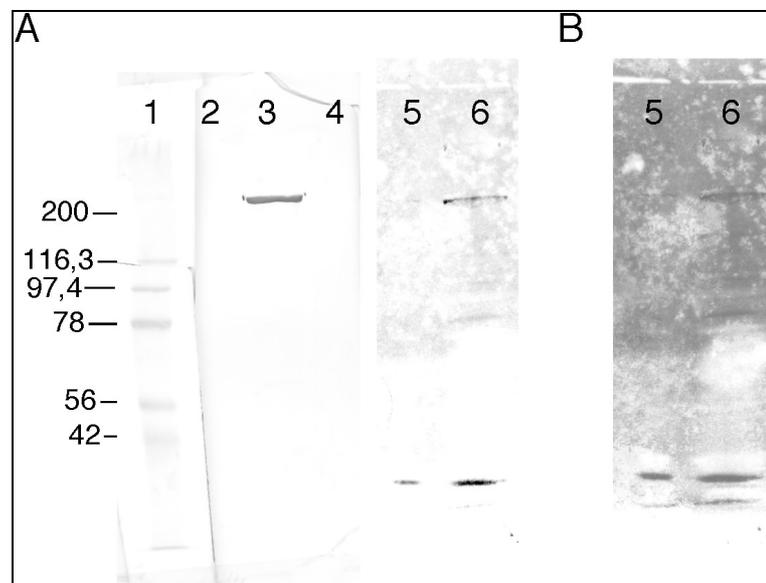


Abbildung 17: Western-Blot und Autoradiogramm von mit  $\beta$ - $^3\text{H}$ Alanin *in vivo* markierten Proteinen. Teile des Blots eines 10% SDS-Gels wurden mit Amidoschwarz gefärbt (MGM Spur 1), mit Anti-RkpF entwickelt (Spur 3, unmarkierter Rohextrakt von OG7001 pTB1090 (*rkpA<sub>114</sub>BCDEF*) oder autoradiografiert (OG7001 pTB1003 (*rkpF*), Spur 5 und OG7001 pTB1090, Spur 6). Die Spuren 2 und 4 waren leer. B) Überexposition der Spuren 5 und 6 aus A. Die Größen der Molekulargewichtsstandards in kDa sind angegeben (Spur 1)

### 3.2.3.4.2 Markierung mit [ $^{35}\text{S}$ ]Sulfat

Um abschätzen zu können in welchem Ausmasse der Proteinkomplex RkpABCDEF posttranslational modifiziert worden war, wurden Markierungen mit  $\beta$ - $^3\text{H}$ Alanin und [ $^{35}\text{S}$ ]Sulfat durchgeführt. Der Proteinkomplex enthält 29

Cysteine und 49 Methionine. Damit ergibt sich ein Verhältnis von 79 S-Atomen pro  $\beta$ -Alanin. Bei den Markierungen mit  $\beta$ - $^3\text{H}$ Alanin wurde außer dem heterologen ACP auch das AcpP aus *E. coli* markiert. Aus dem Verhältnis der markierten Banden und den Proteinmengen zwischen AcpP und den überexprimierten ACPs lässt sich abschätzen, dass nicht mehr als ca. 5-10% des heterolog exprimierten Proteins durch  $\beta$ - $^3\text{H}$ Alanin markiert werden. Da jedoch der Einbau des  $^{35}\text{S}$ -Isotops gleichmäßig erfolgt, ergibt sich daraus ein Verhältnis von 790 S-Atomen pro  $\beta$ -Alanin. Um eine vergleichbare Markierung der Proteine zu erzielen, wurde daher das  $^{35}\text{S}$ Sulfat mit einer um den Faktor 1000 niedrigeren spezifischen Radioaktivität eingesetzt. Zu parallel gezüchteten Kulturen wurden wie unter 2.11.11.2 beschrieben 60 MBq/mM  $\beta$ - $^3\text{H}$ Alanin und 23 kBq/mM  $^{35}\text{S}$ Sulfat zugegeben. Mit den verwendeten Konzentrationen von 1  $\mu\text{M}$   $\beta$ -Alanin und 1 mM Sulfat ergaben sich Gesamtaktivitäten von 60 Bq/ml  $\beta$ - $^3\text{H}$ Alanin und 23 Bq/ml  $^{35}\text{S}$ Sulfat. Die geernteten Zellen wurden in SDS-Auftragspuffer resuspendiert und die extrahierten Proteine auf einem 10% SDS-Gel aufgetrennt. Die Banden des RkpABCDEF-Proteinkomplexes wurden nach Coomassiefärbung aus dem Gel ausgeschnitten, zerkleinert und die Radioaktivität wurde im Szintillationszähler quantifiziert. Es wurden 340 cpm für die mit  $\beta$ - $^3\text{H}$ Alanin markierten Proben gemessen. Daraus errechnet sich eine Stoffmenge von  $9,4 \cdot 10^{-11}$  Mol  $\beta$ - $^3\text{H}$ Alanin in der Gelbande. Für das  $^{35}\text{S}$ -Isotop ergaben sich  $2,6 \cdot 10^{-9}$  Mol  $^{35}\text{S}$  in der Gelbande. Durch das molare Verhältnis von 79:1 ergibt sich daraus, dass ca. 3% des in der Gel-Bande enthaltenden Proteins mit  $\beta$ -Alanin und somit mit 4'-Phosphopantethein modifiziert sind.

### **3.2.3.5 Expression des RkpABCDEF-Proteinkomplexes in *Sinorhizobium meliloti***

Bisher ist noch nicht klar, ob der Proteinkomplex ein Artefakt der heterologen Expression des *rkpABCDEF*-Genklusters ist. Trotz vielfacher Versuche konnte bisher keine RkpF-Bande in *Sinorhizobium meliloti* AK631

nachgewiesen werden. Deshalb wurde ein hochkonzentrierter Rohextrakt von *S. meliloti* AK631 hergestellt und dieser dann größenspezifisch angereichert. Dazu wurden die Zellen aus 5 Liter einer TY-Kultur bei einer  $OD_{600}=1,0$  geerntet. Die Zellen wurden gewaschen und in 50 ml Puffer (50 mM Tris/HCl pH 8,0; 10 mM DTT; 1 mM PMSF) resuspendiert und durch 3 French-Press-Passagen bei 20000 psi aufgeschlossen. Nach einer Zentrifugation für 30 min bei 4°C und 10000xg wurde der Überstand abgenommen und die Proteinkonzentration des Überstandes zu 48 mg/ml bestimmt. Um große Proteine anzureichern, wurde dieser Rohextrakt mit Ultrafree15 Zentrifugationsfiltern (MWCO 100000) um ca. Faktor 10 aufkonzentriert. Auf Grund des hohen Lipidgehaltes war dieses Konzentrat hoch viskos und die Proteinkonzentration konnte nicht bestimmt werden. Anschließend wurden der Rohextrakt, das Konzentrat und der Durchlauf der Aufkonzentrierung in einem Western-Blot analysiert. Abb.18 zeigt, dass sowohl im Rohextrakt als auch in der aufkonzentrierten Probe eine hochmolekulare Bande mit dem Anti-RkpF-Antikörper reagiert. Damit konnte ein Proteinkomplex von ca. 200 kDa auch in *S. meliloti* nachgewiesen werden.

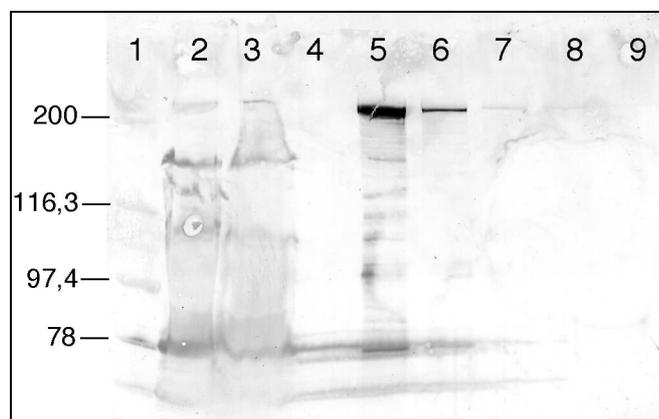


Abbildung 18: Western-Blot-Analyse hochkonzentrierter Rohextrakte aus *S. meliloti* AK631. Ca. 500 µg Gesamtprotein eines Rohextraktes von *S. meliloti* AK631 (Spur 2), sowie der Rückstand (Spur 3) und der Durchlauf (Spur 4) der größenselektiven Aufkonzentrierung wurden auf einem 7,5% SDS-PAGE Gel aufgetrennt. Als Kontrolle wurden 15 µg (Spur 5), 5 µg (Spur 6), 1,6 µg (Spur 7), 0,5 µg (Spur 8) und 0,18 µg (Spur 9) eines Rohextraktes von BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) aufgetragen. Die Größen der Molekulargewichtsstandards in kDa sind angegeben (Spur 1)

### **3.3 Produkt der Proteine RkpABCDEF**

Durch den Nachweis von RkpF konnte gezeigt werden, dass der gesamte klonierte Gencluster *rkpABCDEF* in *E. coli* exprimiert wird. Da sich der entstandene Proteinkomplex durch  $\beta$ -[<sup>3</sup>H]Alanin markieren lässt, ist auch die für eine Aktivität notwendige posttranslationale Modifikation von RkpF erfolgt. Damit sind die Voraussetzungen gegeben, um in *E. coli* nach einem Produkt zu suchen, das von den Proteinen RkpABCDEF gebildet wird.

#### **3.3.1 Markierung mit [1-<sup>14</sup>C]Acetat**

Um alle synthetisierten Fettsäuren oder  $\beta$ -Ketide detektieren zu können, wurden diese mit [1-<sup>14</sup>C]-Acetat markiert. Dazu wurden die Zellen wie unter 2.13.3 beschrieben gezüchtet. Ein Aliquot der gesamten Kultivierungslösung sowie ein Aliquot des gewaschenen und des resuspendierten Zellpellet wurden im Szintillationszähler gemessen. Die Daten ergaben, dass unter diesen Bedingungen ca. 30% der eingesetzten Radioaktivität zellgebunden war.

#### **3.3.2 Untersuchung verschiedener Extraktionsmethoden**

Das Produkt der Proteine RkpABCDEF könnte sowohl frei, als auch gebunden in den Zellen vorliegen. Deshalb wurden verschiedene Extraktionsmethoden mit den markierten Zellen durchgeführt. Neben der Extraktion von Phospholipiden (Bligh & Dyer, 1959), wurde auch eine Umesterung mit saurer Methanollösung und eine bei der Analyse von Polyketiden gebräuchliche Ethylacetatextraktion durchgeführt. Die so erhaltenen Extrakte wurden anschließend in verschiedenen dünnschichtchromatografischen Systemen untersucht. Als Kontrolle wurde jeweils der entsprechende Extrakt von BL21(DE) pTB1003 (*rkpF*) analysiert.

### 3.3.3 Dünnschichtchromatografische Auftrennung

Die Lipidextrakte wurden auf HPTLC-Dünnschichtplatten mit verschiedenen Laufmitteln (siehe 2.13.2) aufgetrennt. Alle Dünnschichtplatten wurden getrocknet und autoradiografiert. Nur nach der Extraktion mit Ethylacetat und der Auftrennung mit Ethylacetat:EtOH:HAc (17:2:1) als Laufmittel, konnte eine für die Expression der Gene *rkpABCDEF* spezifischer Substanz mit einem Rf-Wert von 0,58 gefunden werden (Abb. 19).

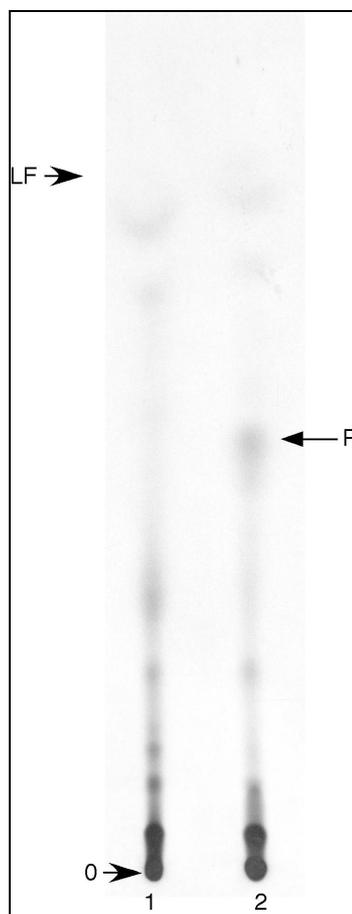
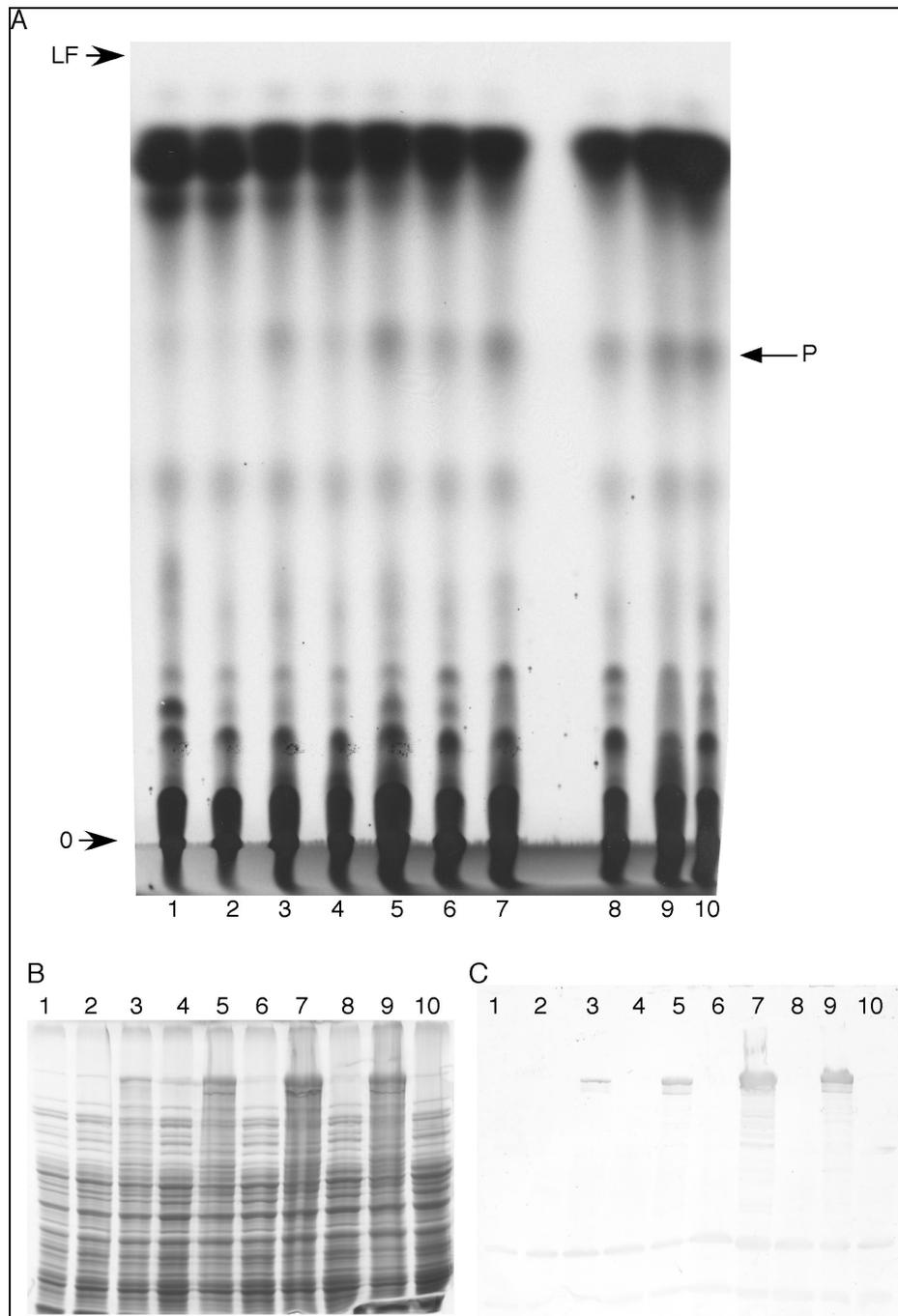


Abbildung 19: Autoradiogramm der mit Ethylacetat:EtOH:HAc (17:2:1) als Laufmittel dünn-schichtchromatografisch aufgetrennten [ $^{14}\text{C}$ ]Acetatmarkierten Ethylacetatextrakten aus *E. coli* BL21(DE3) pTB1003 (*rkpF*) (Spur 1) und BL21(DE3) pTB1090 (*rkpA<sub>1,14</sub>BCDEF*) (Spur 2). O, Startpunkt, LF, Laufmittelfront; P, RkpABCDEF-abhängige Substanz.

### 3.3.4 Induktionsabhängigkeit der Produktbildung

Um zu untersuchen ob das Auftreten der zusätzliche Substanz in den Ethylacetatextrakten aus BL21(DE3) pTB1090 (*rkpA<sub>1,4</sub>BCDEF*) mit der heterologen Expression der Gene *rkpABCDEF* einhergeht, wurde folgender Versuch durchgeführt. Einer M9-Kultur von BL21(DE3) pTB1090 wurde bei einer  $OD_{600}=0,18$ , [ $1-^{14}C$ ]Acetat zugegeben. Bei einer  $OD_{600}=0,28$  wurde die Kultur geteilt und eine Hälfte mit IPTG induziert. Aus diesen Kulturen und unmarkierten Parallelkulturen wurden zu den Zeitpunkten 0, 1, 2, 3 und 4 Stunden Proben entnommen. Um die Proteinexpression zu verfolgen, wurden die Proben mit SDS-Auftragspuffer versetzt und auf 10% SDS-PAGE-Gelen aufgetrennt. Die Proteine wurden mit Coomassiefärbung und Western-Blot detektiert (Abb. 20 B, C). Die Ethylacetatextrakte der markierten Proben wurden dünnschichtchromatografisch aufgetrennt und autoradiografiert (Abb. 20 A). Bereits zum Zeitpunkt 0 Stunden ist die zusätzliche Bande leicht zu sehen. Die Induktion und die längeren Inkubationszeiten führen zu einer Verstärkung des Spots. Jedoch verstärkt sich der Spot nicht im gleichen Maße wie die Proteinmenge. Wie bereits zuvor gezeigt wurde (3.2.3.4.2) ist nur ein kleiner Prozentsatz des Proteinkomplexes posttranslational modifiziert und deshalb potenziell metabolisch aktiv. So wird auch in Proben, in denen im Western-Blot kein Proteinkomplex synthetisiert wird, eine vergleichbare Menge des Produktes der Proteine RkpABCDEF gebildet.



Zeitpunkt in h		0		1		2		3		4	
Induktion mit IPTG		-	+	-	+	-	+	-	+	-	+
Spur	A	1	2	3	4	5	6	7	8	9	10
	B	2	1	4	3	6	5	8	7	10	9
	C	2	1	4	3	6	5	8	7	10	9

Abbildung 20: Untersuchung der induktionsabhängigen Bildung des RkpABCDEF-Proteinkomplexes und dessen vermutlichen Produktes. A) Autoradiogramm der mit Ethylacetat:EtOH:HAc (17:2:1) als Laufmittel dünnenschichtchromatografisch aufgetrennten [ $^{14}\text{C}$ ]Acetat-markierten Ethylacetatextrakten aus BL21(DE3) pTB1090 ( $rkpA_{114}BCDEF$ ). B) Coomassie-gefärbtes 10 % SDS-PAGE-Gel von BL21(DE3) pTB1090. C) Western-Blot mit Anti-RkpF (1:10000) analysiert. O, Startpunkt, LF, Laufmittelfront; P, spezifische Substanz.

### 3.3.5 Einfluss des RkpA-Startkodons auf die Produktbildung

Der RkpF-enhaltende Proteinkomplex konnte in *Sinorhizobium meliloti* AK631 nachgewiesen werden. Dadurch wird auf einen Translationsstartpunkt von RkpA hingewiesen, der zu einer Komplexbildung führt. Jedoch konnte noch kein funktioneller Nachweis für den korrekten Translationstart von *rkpA* gefunden werden. Als weiterer Ansatz sollte die Produktbildung der verschiedenen Varianten des klonierten Genklusters untersucht werden. Dafür wurden Kulturen von BL21(DE3) pTB1003 (*rkpF*), BL21(DE3) pTB1044 ( $rkpA_{367}BCDEF$ ), BL21(DE3) pTB1047 ( $rkpA_{262}BCDEF$ ) und BL21(DE3) pTB1090 ( $rkpA_{114}BCDEF$ ) mit 1- $^{14}\text{C}$ Acetat markiert und die Ethylacetatextrakte dünnenschichtchromatografisch aufgetrennt. Die anschließende Autoradiografie zeigte überraschenderweise, dass auch mit dem ursprünglich vorgeschlagenen Translationsstartpunkt bei 367 eine Produktbildung stattfindet (Abb. 21).

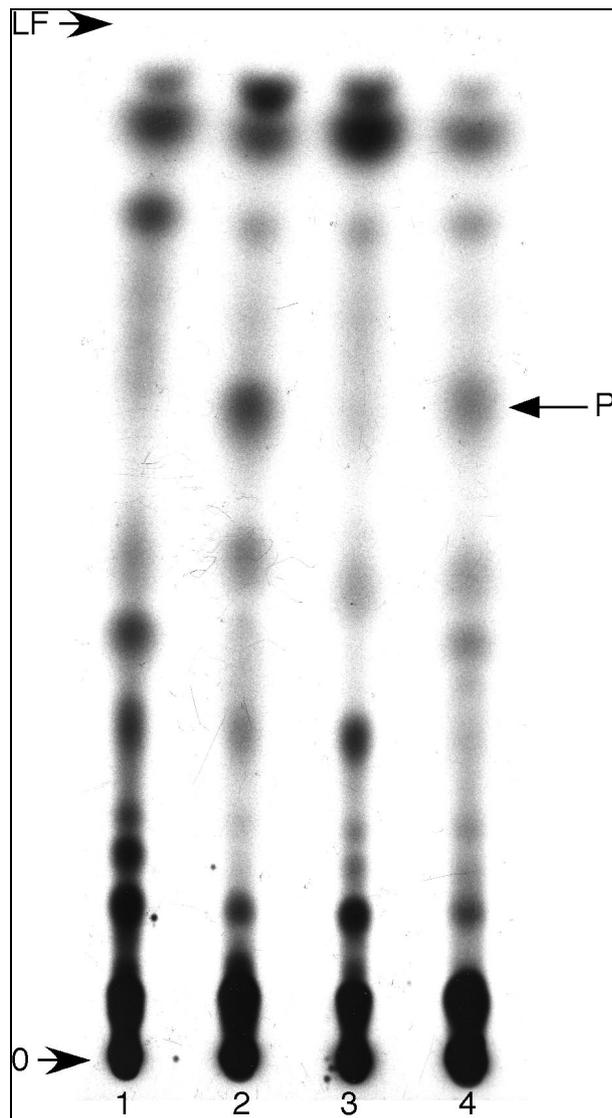


Abbildung 21: Autoradiogramm der mit Ethylacetat:EtOH:HAc (17:2:1) als Laufmittel dünn-schichtchromatografisch aufgetrennten  $[1-^{14}\text{C}]$ Acetatmarkierten Ethylacetatextrakten aus *E. coli* BL21(DE3) pTB1003 (*rkpF*) (Spur 1), BL21(DE3) pTB1044 (*rkpA*<sub>367</sub>*BCDEF*) (Spur 2), BL21(DE3) pTB1047 (*rkpA*<sub>262</sub>*BCDEF*) (Spur 3) und BL21(DE3) pTB1090 (*rkpA*<sub>114</sub>*BCDEF*) (Spur 4). O, Startpunkt, LF, Laufmittelfront; P, spezifische Substanz.

### 3.4 Analyse von Polysacchariden

#### 3.4.1 Analyse der kapsulären Polysaccharide.

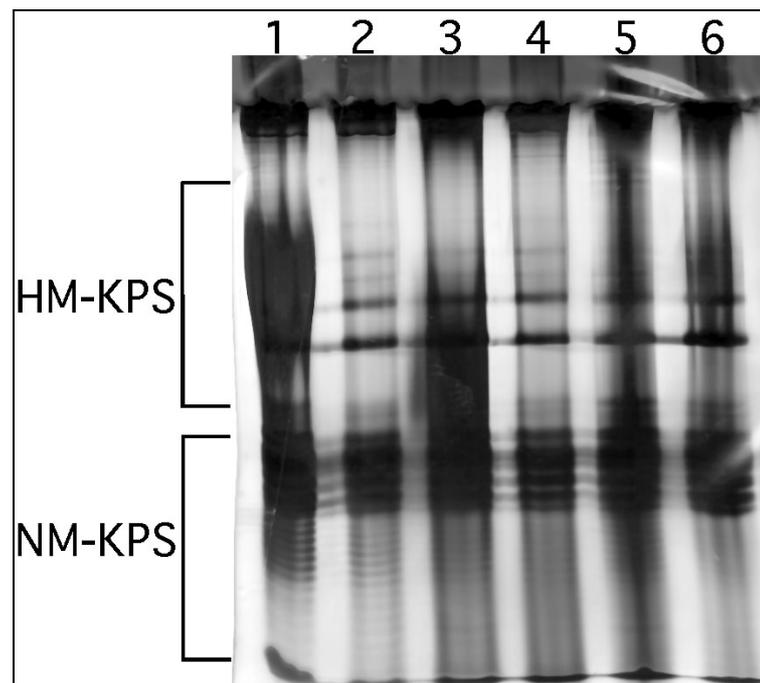
Der *rkpABCDEF*-Gencluster wurde als Komplementationgruppe-I der *rkp*-1-Region beschrieben. Wie unter 3.1.2.2 beschrieben, wurden während der

Klonierung zur heterologen Expression dieses Genklusters Fehler in der publizierten Sequenz gefunden (Petrovics *et al.*, 1993). Um zwischen den drei möglichen Startpunkten von RkpA zu differenzieren, wurde das Genkluster in drei Varianten in ein Plasmid mit breitem Wirtsspektrum kloniert. Durch die Wiederherstellung der Biosynthese der KPS sollte der korrekte Startpunkt bestimmt werden. Da keines der Plasmide die erwartete Komplementation herbeiführte, wurde die Sequenz des Kosmides pPP428 wie unter 3.1.3 beschrieben analysiert. Die Analysen ergaben, dass auch die bisherige Einteilung der Komplementationsgruppen in der *rkp-1*-Region nicht vollständig war.

### **3.4.1.1 Komplementation der Insertionsmutanten von *S. meliloti* AK631**

#### **3.4.1.1.1 Mutante 649 (*rkpU::Tn5*)**

Das neu gefundene Gen *rkpU* wird in entgegengesetzter Transkriptionsrichtung zum *rkpABCDEF*-Genkluster abgelesen. Jedoch kann diese Mutante nicht durch *rkpU trans*-komplementiert werden. Die Bildung von KPS wird durch die Gene *rkpUABCDEF* wieder hergestellt und durch *rkpG* nochmal gesteigert (Abb. 22 und Tab. 8)

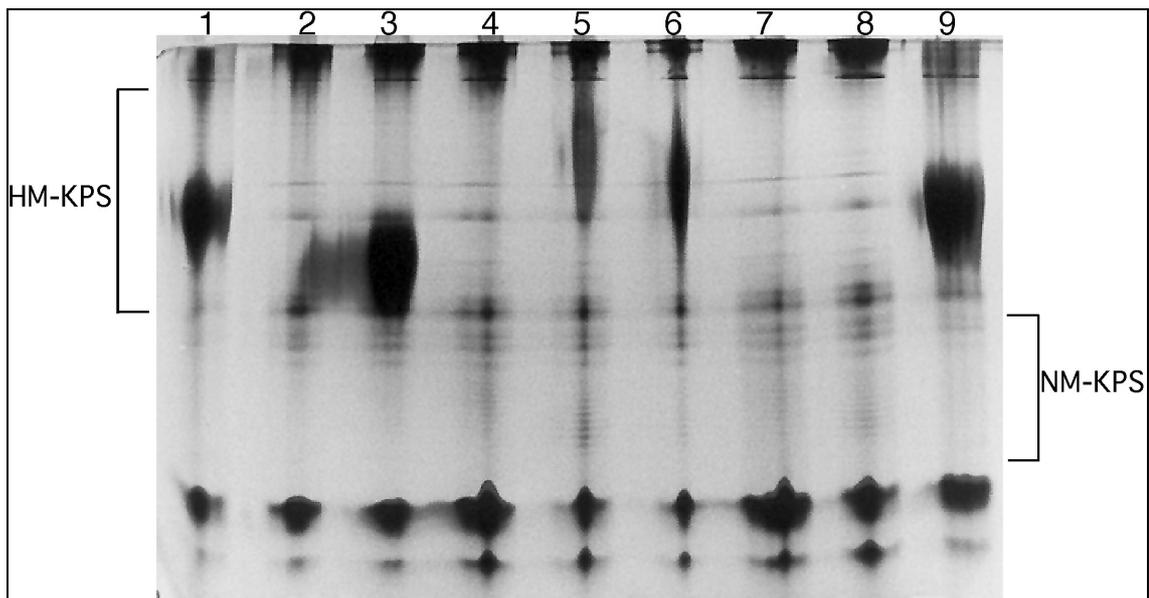


Spur	Stamm	Gene <i>in trans</i>	KPS
1	AK631	Wildtyp	+
2	649	-	-
3	649	Komplementierendes Kosmid	+
4	649	<i>rkpU</i>	-
5	649	<i>rkpUABCDEFG</i>	+
6	649	<i>rkpUABCDEFG</i>	(+)

Abbildung 22: Polysaccharide aus exponentiell wachsenden Kulturen von *S. meliloti* 649 (*rkpU::Tn5*), die verschiedene Plasmide enthielten, wurden extrahiert und auf einem 18% DOC-PAGE-Gel aufgetrennt. KPS wurden durch Alcian-Blau-Behandlung und Silberfärbung detektiert. Die Symbole bedeuten: - kein KPS; (+) wenig KPS; + dem Wildtyp entsprechende KPS-Menge, HM-KPS hochmolekulare KPS, NM-KPS niedermolekulare KPS.

#### 3.4.1.1.2 Mutante 644 (*rkpC::Tn5*)

Obwohl das Gen *rkpU* in dieser Mutante intakt sein sollte, wird die KPS-Biosynthese nur durch die *in trans* vorliegenden Gene *rkpUABCDEFG* komplementiert (Abb. 23 und Tab. 8).

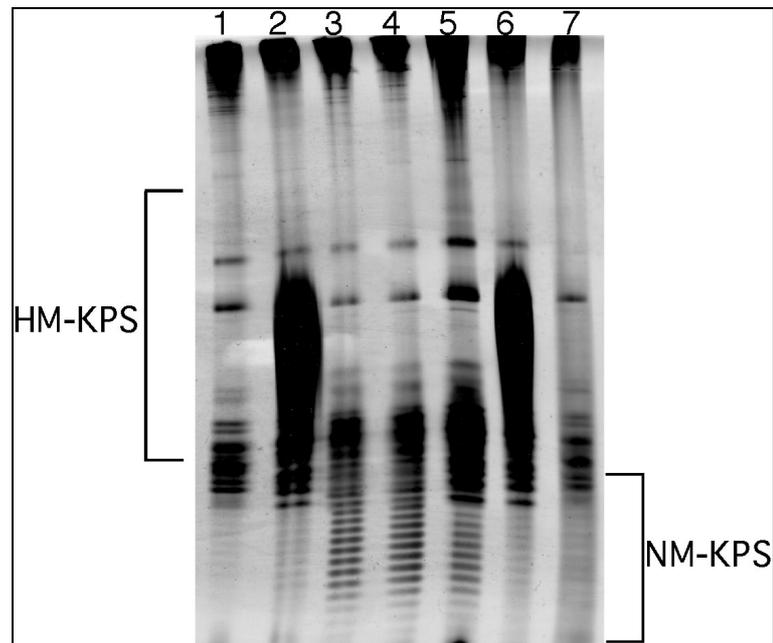


Spur	Stamm	Gene <i>in trans</i>	KPS
1,9	AK631	Wildtyp	+
2	644	-	-
3	644	Komplementierendes Kosmid	+
4	644	<i>rkpUABCDEF</i>	-
5	644	<i>rkpUABCDEFGF</i>	+
6	644	<i>rkpUABCDEFGH</i>	+
7	644	<i>rkpABCDEF</i>	-
8	644	<i>rkpABCDEFGHIJ</i>	-

Abbildung 23: Polysaccharide aus exponentiell wachsenden Kulturen von *S. meliloti* 644 (*rkpC::Tn5*), die verschiedene Plasmide enthielten, wurden extrahiert und auf einem 18% DOC-PAGE-Gel aufgetrennt. KPS wurden durch Alcian-Blau-Behandlung und Silberfärbung detektiert. Die Symbole bedeuten: - kein KPS; (+) wenig KPS; + dem Wildtyp entsprechende KPS-Menge, HM-KPS hochmolekulare KPS, NM-KPS niedermolekulare KPS.

#### 3.4.1.1.3 Mutante 590 (*rkpG::Tn5*)

Die Komplementation der Mutante 590 zeigt ein von den anderen Mutanten leicht abweichendes Bild. Nach den publizierten Daten sollten die Gene *rkpGH* die Komplementationsgruppe-II bilden. Auffällig ist bei dieser Mutante, dass durch die *in trans*-befindlichen Gene *rkpG* und *rkpGHIJ* es zu einer verstärkten Synthese von niedermolekularem KPS kommt. Eine teilweise Komplementation der Biosynthese der hochmolekularen KPS findet durch die Gene *rkpABCDEFGF* statt. Eine vollständige Wiederherstellung der KPS-Biosynthese konnte nur durch pPP428 erzielt werden (Abb. 24 und Tab. 8).

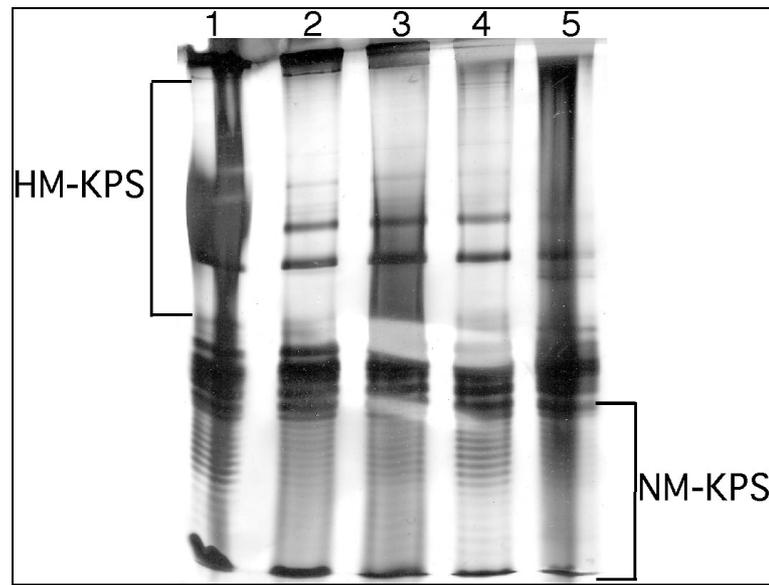


Spur	Stamm	Gene <i>in trans</i>	KPS
1,7	590	-	-
2	590	Komplementierendes Kosmid	+
3	590	<i>rkpGHJ</i>	L
4	590	<i>rkpG</i>	L
5	590	<i>rkpABCDEFG</i>	L, (+)
6	AK631	Wildtyp	+

Abbildung 24: Polysaccharide aus exponentiell wachsenden Kulturen von *S. meliloti* 590 (*rkpU::Tn5*), die verschiedene Plasmide enthielten, wurden extrahiert und auf einem 18% DOC-PAGE-Gel aufgetrennt. KPS wurden durch Alcian-Blau-Behandlung und Silberfärbung detektiert. Die Symbole bedeuten: - kein KPS; (+) wenig KPS; + dem Wildtyp entsprechende KPS-Menge; L erhöhte Menge an niedermolekularen KPS, HM-KPS hochmolekulare KPS, NM-KPS niedermolekulare KPS.

#### 3.4.1.1.4 Mutante 551 (*rkpI::Tn5*)

Das Gen *rkpI* wurde als die Komplementationseinheit-III der *rkp-1*-Region bestimmt (Kiss *et al.*, 1997). Eine Komplementation der Mutante 551 (*rkpI::Tn5*) war jedoch nur mit den Genen *rkpGHJkpsF* möglich. Wurden nur die Gene *rkpGHJ* eingebracht, konnte die Biosynthese der KPS nicht wiederhergestellt werden (Abb. 25). Damit konnte gezeigt werden, dass auch das neu entdeckte Gen *kpsF* an der KPS-Biosynthese beteiligt ist. Ob bereits die Gene *rkpIjksF* ausreichend sind, wurde nicht untersucht (Tab. 8).



Spur	Stamm	Gene <i>in trans</i>	KPS
1	AK631	Wildtyp	+
2	551	-	-
3	551	Komplementierendes Kosmid	+
4	551	<i>rkpGHIJ</i>	-
5	551	<i>rkpGHIJkpsF</i>	+

Abbildung 25: Polysaccharide aus exponentiell wachsenden Kulturen von *S. meliloti* 551 (*rkpl::Tn5*), die verschiedene Plasmide enthielten, wurden extrahiert und auf einem 18% DOC-PAGE-Gel aufgetrennt. KPS wurden durch Alcian-Blau-Behandlung und Silberfärbung detektiert. Die Symbole bedeuten: - kein KPS; (+) wenig KPS; + dem Wildtyp entsprechende KPS-Menge, ,HM-KPS hochmolekulare KPS, NM-KPS niedermolekulare KPS.

#### 3.4.1.1.5 Mutante 633 (*rkpJ::Tn5*)

Das Gen *rkpJ* wurde als die Komplementationseinheit-IV der *rkp-1*-Region bestimmt (Kiss *et al.*, 1997). Jedoch wurde, wie bei der Mutante 551, das neu entdeckte Gen *kpsF* für eine Komplementation benötigt (Tab. 8).

Tabelle 8: Komplementation der verschiedenen Insertionsmutanten von *S. meliloti* AK631

Stamm	<i>Tn5</i> -Insertion	<i>In trans</i> -eingebrachte Gene	Plasmide	KPS
649	<i>rkpU</i>	-	-	-
"	"	<i>rkpU</i>	pTB1094	-
"	"	<i>rkpUABCDEF</i>	pTB1058	(+)
"	"	<i>rkpUABCDEFGF</i>	pTB1058,pTB1115	+
644	<i>rkpC</i>	-	-	-
"	"	<i>rkpABCDEF</i>	pTB1112	-
"	"	<i>rkpUABCDEF</i>	pTB1058	-
"	"	<i>rkpUABCDEFGHJkpsF</i>	pTB1058, pTB1103	+
"	"	<i>rkpUABCDEFGHIJ</i>	pTB1058, pTB1084	+
"	"	<i>rkpABCDEFGHIJ</i>	pTB1112, pTB1084	-
"	"	<i>rkpUABCDEFGF</i>	pTB1058,pTB1115	+
"	"	<i>rkpUABCDEFGH</i>	pTB1058,pTB1116	+
590	<i>rkpG</i>	-	-	-
"	"	<i>rkpUABCDEF</i>	pTB1058	-
"	"	<i>rkpG</i>	pTB1115	L
"	"	<i>rkpABCDEFGF</i>	pTB1112,pTB1115	L, (+)
"	"	<i>rkpGHIJ</i>	pTB1084	L
551	<i>rkpl</i>	-	-	-
"	"	<i>rkpGHIJ</i>	pTB1084	-
"	"	<i>rkpGHIJkpsF</i>	pTB1103	+
"	"	<i>rkplJkpsF</i>		n. B.
"	"	<i>rkpJkpsF</i>		n. B.
633	<i>rkpJ</i>	-	-	-
"	"	<i>rkpGHIJkpsF</i>	pTB1103	+
"	"	<i>rkpGHIJ</i>	pTB1084	-
"	"	<i>rkplJkpsF</i>		n.B.
"	"	<i>rkpJkpsF</i>		n. B.

Die Symbole bedeuten: - kein KPS; (+) wenig KPS; + dem Wildtyp entsprechende KPS-Menge; L erhöhte Menge an niedermolekularen KPS; n. B. nicht Bestimmt

---

## 4 Diskussion

### 4.1 Untersuchungen zu RkpF

#### 4.1.1 RkpF ist ein Acyl-Carrier-Protein.

Der sechste ORF (jetzt *rkpF*) der Komplementationseinheit-I der *rkp-1*-Region ist zu Acyl-Carrier-Proteinen (ACPs) homolog. Nach Klonierung von *rkpF* in einen Überexpressionvektor wurde das heterolog in *E. coli* überexprimierte Protein gereinigt. Das Protein RkpF konnte *in vivo* durch  $\beta$ - $^3\text{H}$ Alanin markiert werden. Da  $\beta$ -Alanin ein spezifischer Baustein während der Biosynthese von 4`-Phosphopantethein ist, bedeutet der Einbau von  $\beta$ - $^3\text{H}$ Alanin in RkpF, dass RkpF posttranslational mit der prosthetischen Gruppe 4`-Phosphopantethein modifiziert wird. Es konnte weiterhin gezeigt werden, dass die Modifikation durch die gereinigte Phosphopantetheinyltransferase AcpS *in vitro* durchgeführt werden kann (Epple *et al.*, 1998). Die posttranslationale Modifikation führt zu einem veränderten Laufverhalten in der nativen PAGE. Diese Veränderung konnte auch für das konstitutive AcpP von *E. coli* gezeigt werden, und wurde dort als Stabilisierung der Helix 2 durch die prosthetische Gruppe interpretiert (Jackowski & Rock, 1983). Bei RkpF ist diese Veränderung noch deutlicher und stabiler als bei dem AcpP aus *E. coli* (Epple *et al.*, 1998). Neben RkpF konnten bereits drei weitere ACPs aus Rhizobien identifiziert und überexprimiert werden. Dabei handelt es sich um das konstitutive AcpP (Lopez-Lara & Geiger, 2000; Platt *et al.*, 1990), NodF (Geiger *et al.*, 1991; Ghose *et al.*, 1996; Ritsema *et al.*, 1994) sowie AcpXL (Brozek *et al.*, 1996; Lopez-Lara & Geiger, 2000). Neben diesen bereits untersuchten ACPs lässt eine Analyse des Genoms von *S. meliloti* noch auf das Vorhandensein weiterer ACPs schließen (Geiger & Lopez-Lara, 2002).

Das ACP NodF ist an der Biosynthese und Übertragung einer ungewöhnlichen mehrfach ungesättigten Fettsäure beteiligt, die in die Nod-Faktoren eingebaut wird (Geiger *et al.*, 1991; Geiger *et al.*, 1994; Geiger *et al.*, 1998; Yang *et al.*, 1999). AcpXL ist an der Übertragung der ungewöhnlich

langen Fettsäure 27-Hydroxyoctacosansäure auf das Zuckergerüst des Lipid A beteiligt (Brozek *et al.*, 1996). Trotz ihrer niedrigen Sequenzidentität von nur 26-32% (Brozek *et al.*, 1996) zeigen initiale Analysen der Sekundärstruktur und der generellen drei-dimensionalen Faltung, dass die Gesamtstruktur der ACPs sehr stark konserviert ist (Ghose *et al.*, 1996). Trotz ihrer strukturellen Ähnlichkeit können jedoch die ACPs sich nicht funktionell ersetzen. Für NodF wurde gezeigt, dass die Spezifität seiner Funktion in der C-terminalen Hälfte des Proteins kodiert ist (Ritsema *et al.*, 1998). Für die spezialisierten ACPs NodF, AcpXL und RkpF konnte gezeigt werden, dass Sie durch die gereinigte Phosphopantetheinyltransferase AcpS aus *E. coli* (Lambalot & Walsh, 1995) posttranslational modifiziert werden können (Epple *et al.*, 1998; Lopez-Lara & Geiger, 2000). Für ACPs aus der Biosynthese aromatischer Polyketide in *Streptomyces coelicolor* konnte gezeigt werden, dass AcpS ebenfalls acetyliertes Phosphopantetheine von Acetyl-CoA auf Fren- oder Gra-Apo-ACP übertragen kann (Carreras *et al.*, 1997). Von RkpF konnten auf diese Weise jedoch keine acetylierten Formen erzeugt werden.

#### 4.1.2 RkpF in *Sinorhizobium meliloti* AK631

Peter Putnoky vom Institute of Genetics, Biological Research Center, Hungarian Academy of Science hat eine Insertionsmutante in *rkpF* hergestellt und gezeigt, dass eine derartige Mutante keine KPS bildet.

Überraschenderweise gelang es jedoch nicht, diese Mutante mit einem funktionellen *rkpF*-Gen *in-trans* zu komplementieren. Es konnte in dieser Arbeit gezeigt werden, dass mit *rkpF in trans* komplementierte Stämme sowohl die Apo- als auch die Holoform von RkpF exprimierten. Jedoch scheint RkpF in *S. meliloti* AK631 so nicht zu existieren, sondern nur als Modul in einem größeren Protein vorzukommen.

#### 4.2 Klonierung der Gene *rkpABCDEF*

Die durchgeführten Arbeiten zu den Klonierungen der Gene *rkpABCDEF* in den Überexpressionsvektor pET9a für die heterologe Expression in *E. coli*

sind in Abb. 26B zusammengefasst. Die mit PCR amplifizierten Genbereiche wurden zur Kontrolle nachsequenziert. Bei diesen Sequenzierungen ergaben sich einige Abweichungen von der in der Genbank vorhandenen Sequenz (Gene bank accession number: X64131). Die fraglichen Bereiche wurden mit Hilfe von Subklonen des komplementierenden Kosmides erneut sequenziert. Dabei ergab sich, dass der Genbankeintrag einige Fehler enthielt. Die meisten Fehler hatten nur einen Aminosäureaustausch oder Leserasterverschiebungen über kleine Bereiche zur Folge. Drei fehlerhafte Basen jedoch hatten größere Auswirkungen. So ergab sich aus einem Fehler im vermuteten Stoppkodon des *rkpB*-Gens und einem Fehler in der vermuteten Intergenregion zwischen den Genen *rkpB* und *rkpC*, dass es nur einen großen ORF *rkpB/C* gibt (Abb. 26A). Ein weiterer folgenreicher Fehler war eine zusätzliche Base an Position 210 des Genbankeintrags. Durch diese zusätzliche Base veränderte sich der Leserahmen von *rkpA*, so dass neben dem vermuteten Startpunkt bei Base 367 und einem weniger wahrscheinlichen Startpunkt bei 262 auch noch ein Startpunkt bei 114 möglich war. Der mögliche Startpunkt bei 114 ist vor allem deshalb interessant, da die Homologien zu den modularen Polyketidsynthesen bei der Base 120 beginnen. Aus diesem Grunde wurden die Gene *rkpABCDEF* mit allen drei möglichen Startpunkten in den pET9a-Vektor kloniert.

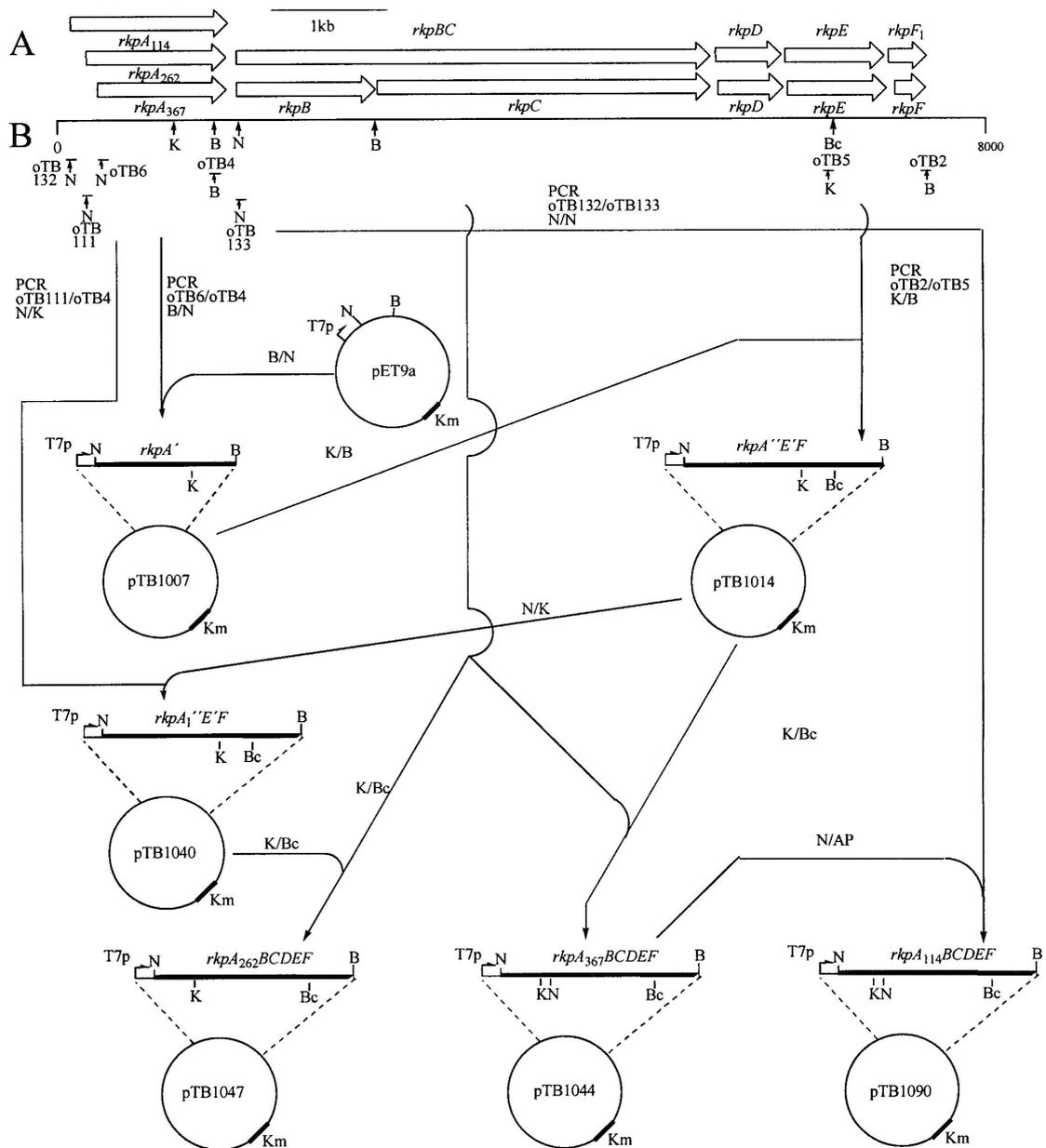


Abbildung 26: Klonierung des *rkpABCDEF*-Genklusters aus *Sinorhizobium meliloti* AK631 für die heterologe Expression in *E. coli*. A) Mögliche Offene Leserahmen im Gencluster *rkpABCDEF*. Die unterste der drei abgebildeten Möglichkeiten entspricht der publizierten Form (Petrovics *et al.*, 1993). B) Klonierungsschema für die Klonierung des *rkpABCDEF*-Genklusters in den Überexpressionsvektor pET9a mit den drei verschiedenen Startpunkten für RkpA. Abkürzungen: oTB, PCR-Primer; K, *KpnI*; B, *BamHI*; N, *NdeI*; Bc, *BclI*; AP, Alkalische Phosphatase; T7p, Promotor für die Expression mit der T7 Polymerase; Km, Kanamycin-Resistenzgen.

### 4.3 Heterologe Expression der Gene *rkpABCDEF*

Die heterologe Expression der Gene *rkpABCDEF* mit den verschiedenen Startpunkten für *rkpA* lieferte überraschende Ergebnisse. Das Expressionsverhalten und die Proteinanalyse der drei Stämme *E. coli* BL21(DE3) pTB1044 (*rkpA*<sub>367</sub>*BCDEF*), *E. coli* BL21(DE3) pTB1047 (*rkpA*<sub>262</sub>*BCDEF*) und *E. coli* BL21(DE3) pTB1090 (*rkpA*<sub>114</sub>*BCDEF*) zeigte große Unterschiede. In Extrakten aller drei Stämme konnte eine zusätzliche Proteinbande in der SDS-PAGE von ca. 20 kDa detektiert werden. Eine Bande identischer Größe entsteht bei der heterologen Expression des RkpD-Proteins. Eine weitere zusätzliche Proteinbande hat die scheinbare Größe von 150 kDa in der SDS-PAGE und könnte der Größe des RkpB/C Proteins entsprechen. Dieser Befund scheint die Daten der Sequenzierung zu bestätigen, dass es zwischen den ursprünglich vermuteten Genen *rkpB* und *rkpC* kein Stoppkodon gibt. In keinem der drei Extrakte konnte die monomere Form von RkpF nachgewiesen werden. In den Rohextrakten der Stämme mit den früheren Startpunkten für *rkpA* wurde abhängig von der Induktion der Gene *rkpABCDEF* ein Protein von ca. 250 kDa gebildet. Bei Verwendung des Transkriptionsstarts bei 114 wurde dieses Protein in so großen Mengen synthetisiert, dass sich "Inclusion-Bodies" bildeten. Dieses Expressionsverhalten deckt sich mit den Ergebnissen, die aus der Klonierung der einzelnen *rkpA*-Gene erhalten wurden. Bei Expression mit dem Startcodon bei 367 konnte keine zusätzliche Proteinbande entdeckt werden. Bei der Verwendung der anderen beiden möglichen Startpunkte entstand jeweils ein Protein mit der für RkpA erwarteten scheinbaren Masse in der SDS-PAGE. Bei Verwendung des Startpunktes bei 114 wurde dieses Protein in sehr großen Mengen und in Form von "Inclusion-Bodies" gebildet. Ein Protein, das sich unter denaturierenden Bedingungen entsprechend einem Molekulargewicht von 250 kDa in der SDS-PAGE bewegte, ließ vermuten, dass es sich dabei möglicherweise um einen stabilen Komplex der RkpABCDEF Proteine handelt. Dieser Befund ist überraschend, da der Komplex aus 5 Einzelproteinen besteht, die eigentlich in SDS-haltigen Lösungen voneinander dissoziieren müssten. Der Proteinkomplex mit einem

Molekulargewicht von 250 kDa reagierte auch mit dem gegen RkpF gerichteten Antikörper. Dieser Befund zeigt, dass RkpF ein Bestandteil des Proteinkomplexes ist. Weder die Verwendung reduzierender Agentien (Mercaptoethanol, DTT) noch hohe Konzentrationen an Harnstoff oder Kombinationen dieser Substanzen führten dazu, dass dieser Komplex dissoziierte. Durch radioaktive Markierung des Proteinkomplexes mit  $\beta$ - $^3\text{H}$ Alanin konnte gezeigt werden, dass RkpF zumindest teilweise in seiner Holoform im Komplex vorliegt. Durch Doppelmarkierung mit  $\beta$ - $^3\text{H}$ Alanin und  $^{35}\text{S}$ Sulfat konnte ermittelt werden, dass ca. 3% des in "Inclusion-Bodies" exprimierten Proteinkomplexes die Holoform von RkpF enthalten. Nach ihrer genetischen Organisation entsprechen die Gene *rkpABCDEF* einer Typ-II-Polyketidesynthase. Bei Typ-II-Polyketidesynthasen werden die einzelnen Reaktionen der Kettenverlängerung durch monofunktionale Proteine ausgeführt. Bei einem Homologievergleich ähneln die Proteine RkpABCDEF jedoch stärker den Typ-I-Polyketidesynthasen, bei denen alle enzymatischen Aktivitäten, die für die Synthese des Polyketids benötigt werden, auf einem Polypeptid vereint sind. Die Typ-I-Polyketidesynthasen haben einen modularen Aufbau und jedes Modul führt einen Schritt der Kettenverlängerung des wachsenden Polyketides aus. Jedoch konnte auch gezeigt werden, dass Module wiederholt durchlaufen werden (Gaitatzis *et al.*, 2002; He & Hertweck, 2003; Wilkinson *et al.*, 2000). Inzwischen sind viele dieser iterativen Typ-I-Polyketidesynthasen bekannt (reviewed in Shen, 2003). Der RkpABCDEF-Proteinkomplex könnte eine Zwischenstufe der bisher bekannten Typ-I- und Typ-II-Polyketidsynthasen sein. Interessanterweise ist im sequenzierten Genom von *S. meliloti* 1021 nur ein ORF annotiert, der den gesamten Genen *rkpABCDEF* aus *S. meliloti* Stamm41 AK631 entspricht und der somit zumindest in *S. meliloti* 1021 für ein Protein von 250 kDa kodieren würde.

#### **4.4 Nachweis des RkpABCDEF-Proteinkomplex in *S. meliloti***

Es war unklar, ob der ungewöhnliche 250 kDa-große *rkpABCDEF*-abhängige Proteinkomplex ein Artefakt der Expression in *E. coli* war oder ob er auch in *S. meliloti* auftritt. Das niedrige Expressionsniveau der *rkpABCDEF*-Gene (Petrovics *et al.*, 1993) machte es schwierig diesen Komplex in Wildtyp-Stämmen nachzuweisen. In einem hoch aufkonzentrierten Rohextrakt von *S. meliloti* konnte der Proteinkomplex jedoch nachgewiesen werden. Eine monomere Form von RkpF konnte im Wildtypstamm nicht nachgewiesen werden. Um eine effektive Biosynthese eines Produktes von 5 verschiedenen Proteinen, die in niedrigen Konzentrationen vorhanden sind, zu gewährleisten, wäre die Ausbildung eines stabilen Komplexes der beteiligten Proteine von großem Vorteil.

#### **4.5 Vergleich von RkpABCDEF aus *Sinorhizobium meliloti* AK631 mit RkpA aus dem sequenzierten Genom von *S. meliloti* 1021**

Nach Beendigung der praktischen Arbeiten wurde das Genom des eng verwandten *S. meliloti* 1021 publiziert (Barnett *et al.*, 2001; Capela *et al.*, 2001; Finan *et al.*, 2001; Galibert *et al.*, 2001). Im sequenzierten Genom von *S. meliloti* 1021 wird anstatt der Gene *rkpABCDEF* nur noch ein langer ORF als *rkpA* annotiert. Da der Genbankeintrag für *rkpABCDEF* von *S. meliloti* Stamm 41 AK631 (X64131) Fehler aufweist, konnte nicht ausgeschlossen werden, dass auch in *S. meliloti* Stamm 41 AK631 nur ein offener Leserahmen für *rkpABCDEF* vorhanden war. Jedoch ergab die Überprüfung der eigenen Sequenzierungsdaten, dass die Stoppkodons von *rkpAEF* vorhanden sind. Im Nukleotidalignment sind die Sequenzen von *S. meliloti* 1021 und Stamm 41 AK631 zu 98,75% identisch. Insgesamt sind 62 Positionen verändert. Bei den Veränderungen sind der Austausch, Insertion oder Deletion von einzelnen Nukleotiden, jedoch auch Insertionen und Deletionen von Triplets zu finden (siehe Anhang). Für ein ClustalW Alignment der Proteinsequenzen wurde aus

den einzelnen Proteinen aus AK631 ein Fusionsprotein durch jeweiliges Einfügen von 3 Alaninen an den Verbindungsstellen hergestellt. Abweichungen zwischen den Sequenzen sind in den Bereichen aller Stoppkodons (RkpA 458, RkpB/C 1852, RkpD 2092, RkpE 2392, RkpF 2500) zu finden, aber auch innerhalb der einzelnen Proteine. Die Verteilung der Abweichungen und die Tatsache, dass in weiten Teilen dieser Abweichungen die eigenen Sequenzdaten mit dem entsprechenden Genbankeintrag übereinstimmen, lässt Sequenzierungsfehler als Ursache der Unterschiede ausscheiden. Bei der Evolution der Polyketidsynthasen haben sich daher vielleicht erst aus Type-II-Polyketidsynthasen durch posttranslationale Modifikation stabile Proteinkomplexe gebildet, wie im *S. meliloti* Stamm 41 AK631. Aus dieser Vorstufe haben sich dann die iterativen Typ-I-Polyketidesynthasen entwickelt, wie eine in *S. meliloti* Stamm 1021 zu finden ist. Im benachbarten Protein RkpU sind die Proteinsequenzen der beiden Stämme identisch.

#### **4.6 Produkt der Proteine RkpABCDEF**

Um das Produkt der Proteine RkpABCDEF zu identifizieren, wurden diese Gene in *E. coli* nach IPTG-Induktion überexprimiert. Zellen von *E. coli* BL21(DE3) pTB1090 (*rkpA<sub>114</sub>BCDEF*) und als Kontrolle BL21(DE3) pTB1003 (*rkpF*) wurden in Minimalmedium unter Zusatz von <sup>14</sup>C-Acetat gezüchtet. Diese Kulturen wurden mit verschiedenen Lösungsmitteln extrahiert und die Extrakte mit Dünnschichtchromatografie untersucht. Wenn die Kulturen mit Ethylacetat (1% Essigsäure) extrahiert wurden und diese Extrakte auf kieselgelhaltigen Dünnschichtplatten mit dem Laufmittel Ethylacetat:Ethanol:Essigsäure (17:2:1; v/v) entwickelt wurden, konnte auf den Autoradiogrammen der Dünnschichtplatten eine zusätzliche Substanz nachgewiesen werden. Das Auftreten dieser neuen Substanz geht mit der Induktion der *rkpA<sub>114</sub>BCDEF*-Gene und der Bildung des 250 kDa-großen Proteinkomplexes einher. In zeitabhängigen Experimenten konnte die von RkpABCDEF synthetisierte Substanz, mittels der Dünnschichtchromatografie, bereits nachgewiesen

werden, bevor der Proteinkomplex in einem Western-Blot detektiert werden konnte. Die Bildung des Produktes der Proteine RkpABCDEF konnte nicht mit einem bestimmten Startpunkt von RkpA korreliert werden. Leider konnte im Rahmen der vorliegenden Arbeit die Substanz nicht weiter angereichert werden. Der ursprüngliche Ansatz (Spaink *et al.*, 1991) konnte aus experimentellen Gründen nicht weiterverfolgt werden. Die alternativen Versuche zur Etablierung eines HPLC-Protokolles waren leider nicht erfolgreich.

#### **4.7 Komplementation von rhizobiellen Mutanten**

Es wurde versucht, den KPS-Phenotyp vorhandener rhizobieller Mutanten in der *rkp-1*-Region mit den klonierten *rkpABCDEF*-Genen zu komplementieren. Mit keiner der drei *rkpABCDEF*-enthaltenden Varianten konnten rhizobielle Mutanten in der *rkp-1*-Region komplementiert werden. Die folgende Sequenzierung der bisher unbekannt Bereiche des komplementierenden Kosmides pPP428 führte zur Identifizierung zweier neuer ORFs, die an der Biosynthese der KPS beteiligt sind. Ein ORF (*rkpU*) liegt mit divergenter Transkriptionsrichtung vor *rkpA*. Dieser ORF zeigt Homologien zu einer Familie von Lipoproteinen (PROSITE: PDOC00013). Für zwei Mitglieder dieser Lipoproteinfamilie, *ctrl* aus *Neisseria meningitidis* (Frosch *et al.*, 1992) und *wza* (K30) aus *E. coli* (Drummel-Smith & Whitfield, 2000) konnte gezeigt werden, dass sie am Transport von KPS über die äußere Membran beteiligt sind. Im Gegensatz zur bisherigen Vorstellung, dass das von den Proteinen RkpABCDEF gebildete Ketid als Anker für das zu synthetisierende KPS fungiert, könnte alternativ dieses Ketid auch der Anker des sinorhizobiellen Lipoproteins RkpU sein. Die *S. meliloti* Mutante 644 (*rkpC::Tn5*), die eine Insertionsmutation in *rkpC* hat, ließ sich jedoch auch durch die Gene *rkpUABCDEF* nicht vollständig komplementieren. Für eine vollständige Komplementation dieser Mutante wurde auch noch *rkpG* benötigt. Der andere relevante offene Leserahmen, der in divergenter Transkriptionsrichtung Downstream von *rkpJ* liegt, zeigt Homologien zum *kpsF*-Gen aus *E. coli*. Die

genaue Funktion von *kpsF* ist bisher noch unklar, jedoch konnte in *E. coli* gezeigt werden, dass es essenziell für den Transport der KPS ist (Cieslewicz & Vimr, 1996; Cieslewicz & Vimr, 1997; Simpson *et al.*, 1996). Für eine Komplementation der Mutanten 551 (*rkpI::Tn5*) und 633 (*rkpJ::Tn5*) war das Gen *kpsF* essenziell.

#### 4.7.1 Neueinteilung der Komplementationsgruppen der *rkp-1*-Region

Die bisherige Einteilung der *rkp-1*-Region in die Komplementationsgruppen: I mit den Genen *rkpABCDEF*, II mit den Genen *rkpGH*, III mit dem Gen *rkpI* und IV mit dem Gen *rkpJ* (Kiss *et al.*, 1997; Petrovics *et al.*, 1993) konnte nicht bestätigt werden. Gegen diese Einteilung sprachen auch die Daten von Peter Putnoky zur Trans-Komplementation der *rkpF*-defizienten Mutante PP2756 (siehe 3.2.1.2) sowie die Beschreibung der Mutante 649 bei der die *Tn5*-Insertion außerhalb der 4 Komplementationsgruppen liegt. Die Ergebnisse der Komplementationsexperimente lassen vermuten, dass die *rkp-1*-Region nur zwei statt der vier beschriebenen Komplementationsgruppen umfasst. Die erste Gruppe besteht aus den Genen *rkpUABCDEFGH* und die zweite aus den Genen *rkpIJKpsF*. Jedoch ist diese Einteilung noch nicht eindeutig, da keine Mutanten in *rkpU*, *rkpH* und *kpsF* zu Verfügung standen. Aus den Komplementationsexperimenten geht jedoch hervor, dass die neu entdeckten Gene *rkpU* und *kpsF* an der Biosynthese bzw. dem Transport der KPS beteiligt sind. Das neu identifizierte Lipoprotein gehört zur gleichen Komplementationsgruppe wie die Gene *rkpABCDEF* und die vermutliche Acyltransferase RkpG. Das lässt vermuten, dass dieses Lipoprotein der Akzeptor für das durch RkpABCDEF gebildete und von RkpG übertragene  $\beta$ -Ketid sein könnte. Die Insertionsmutante in *rkpU* wäre damit ein idealer Stamm, um nach dem Produkt der Proteine RkpABCDEF in *S. meliloti* zu suchen. Die wahrscheinliche Organisation der *rkp-1*-Region ist in Abb. 27 dargestellt.

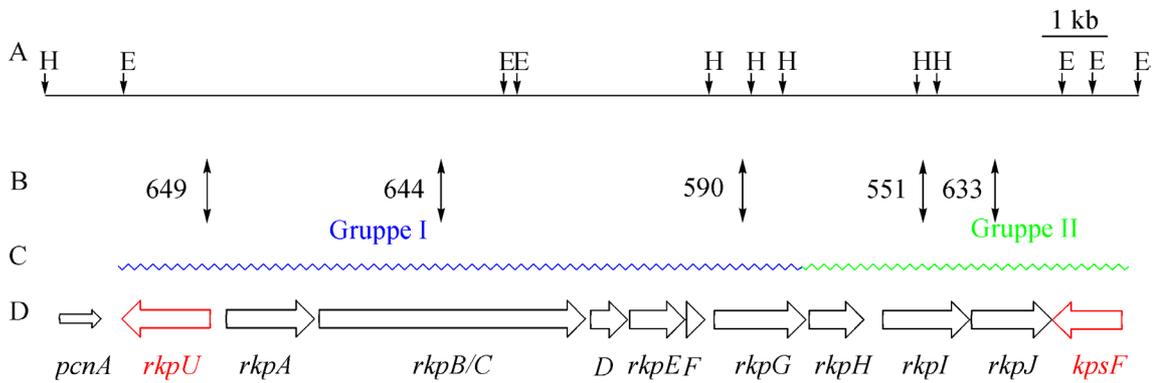


Abbildung 27: Organisation der *rkp-1*-Region. A) Schnittstellen der Restriktionendonukleasen *Hind*III (H) und *Eco*RI (E). B) Die Pfeile zeigen die Positionen der *Tn5*-Insertionen in den Mutanten von *S. meliloti* AK631 (Kiss *et al.*, 2001; Petrovics *et al.*, 1993). C) Komplementationsgruppen in der *rkp-1*-Region. D) Gene und ihre Orientierung in der *rkp-1*-Region

---

## 5 Zusammenfassung

Bei vielen Interaktionen pathogener Bakterien und ihrer Wirte spielen kapsuläre Polysaccharide (KPS) eine wichtige Rolle. Bei der Ausbildung einer funktionellen Symbiose des Bakteriums *Sinorhizobium meliloti* Stamm 41 und der Wirtspflanze *Medicago sativa* sind kapsuläre Polysaccharide essentiell. Mutationen in den Genen *rkpABCDEFGHIJ* verhindern die Synthese der Polysaccharide. Solche Mutanten sind nicht mehr in der Lage Stickstoff-fixierende Wurzelknöllchen zu bilden (Kiss *et al.*, 1997; Petrovics *et al.*, 1993; Reuhs *et al.*, 1993). Die Gene *rkpABCDEF* zeigen starke Ähnlichkeiten zu Genen von modularen Polyketidsynthasen und eukaryotischen Fettsäuresynthethasen. Bisher konnte jedoch in den KPS weder ein  $\beta$ -Ketid noch ein Fettsäurerest entdeckt werden. Im Rahmen dieser Arbeit konnte gezeigt werden, dass es sich beim Protein RkpF tatsächlich um ein Acyl-Carrier-Protein handelt (Epple *et al.*, 1998). Für die heterologe Überexpression der Gene *rkpABCDEF* in *E. coli* wurden diese in den pET9a-Vektor kloniert. Um die korrekte Klonierung der Gene *rkpABCDEF* zu überprüfen, sollte die rhizobielle *rkpC*-Mutante 644 komplementiert werden. Diese Komplementation war jedoch mit den Genen *rkpABCDEF* nicht möglich. Anschliessende Untersuchungen ergaben, dass weitere bisher unbekannte Gene, für die KPS-Bildung benötigt werden und auf dem komplementierenden Kosmid pPP428 lokalisiert sein mussten. Die Sequenzierung der verbleibenden 17 kb des Kosmides führte zur Entdeckung zweier neuer Gene, die an der Biosynthese bzw. dem Transport der kapsulären Polysaccharide beteiligt sind. Zusätzlich wurden die Komplementationsgruppen der *rkp*-1-Region neu bestimmt. Eines der neuen Gene zeigt Homologien zu einer Familie von Lipoproteinen, die am Export von Polysacchariden beteiligt sind (Drummelsmith & Whitfield, 2000; Frosch *et al.*, 1992). Dieses Lipoprotein könnte einen möglichen Akzeptor für das  $\beta$ -Ketid darstellen, welches durch die Proteine RkpABCDEF synthetisiert wird. Das zweite neue Gen zeigt Homologien zum *kpsF*-Gen aus *E. coli*, dessen genaue Rolle beim Transport der kapsulären Polysaccharide noch nicht

geklärt ist (Cieslewicz & Vimr, 1997). Die heterologe Expression der Gene *rkpABCDEF* in *E. coli* führte zur Bildung eines stabilen Proteinkomplexes mit einer Masse von ca. 250 kDa. Überraschenderweise ist dieser Proteinkomplex auch in einem denaturierenden SDS-Gel stabil. Mit Western-Blots konnte gezeigt werden, dass RkpF ein Bestandteil dieses Komplexes ist. Der Komplex konnte auch mit [<sup>3</sup>H]-β-Alanin markiert werden. Dies ist ein Hinweis darauf, dass RkpF zumindest teilweise mit 4'-Phosphopantethein substituiert ist und entsprechend in seiner Holoform vorliegt. Der ungewöhnliche Proteinkomplex konnte ausserdem in *S. meliloti* Stamm 41 nachgewiesen werden. Wenn die Gene *rkpABCDEF* in Anwesenheit von <sup>14</sup>C-markiertem Acetat in *E. coli* überexprimiert wurden, erfolgte die Bildung einer neuen Substanz. Diese neue Substanz konnte mittels Dünnschichtchromatografie nachgewiesen werden. Das Auftreten dieser neuen Substanz ist von der Induktion der Gene *rkpABCDEF* abhängig.

## 6 Literaturverzeichnis

**Albus, U., Baier, R., Holst, O., Puhler, A. & Niehaus, K. (2001).** Suppression of an elicitor-induced oxidative burst reaction in *Medicago sativa* cell cultures by *Sinorhizobium meliloti* lipopolysaccharides. *New Phytol* **151**, 597-606.

**Altschul, S. F., Gish, W., Miller, W., Meyers, E. W. & Lipman, D. J. (1990).** Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215**, 403-410.

**Bairoch, A., Bucher, P. & Hofmann, K. (1997).** The PROSITE database, its status in 1997. *Nucl Acids Res* **25**, 217-221.

**Barnett, M. J., Fisher, R. F., Jones, T. & other authors (2001).** From the Cover: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. *Proc Natl Acad Sci U S A* **98**, 9883-9888.

**Battisti, L., Lara, J. C. & Leigh, J. A. (1992).** Specific oligosaccharide form of the *Rhizobium meliloti* exopolysaccharide promotes nodule invasion in alfalfa. *Proc Natl Acad Sci U S A* **89**, 5625-5629.

**Becquart-de Kozak, I., Reuhs, B. L., Buffard, D., Breda, C., Kim, J. S., Esnault, R. & Kondorosi, A. (1997).** Role of the K-Antigen Subgroup of Capsular Polysaccharides in the Early Recognition Process Between *Rhizobium meliloti* and Alfalfa Leaves. *Mol Plant-Microbe Interact* **10**, 114-123.

**Beringer, J. E. (1974).** R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84**, 188-198.

**Bevitt, D. J., Cortes, J., Haydock, S. F. & Leadlay, P. F. (1992).** 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur J Biochem* **204**, 39-49.

**Bhat, U. R., Carlson, R. W., Busch, M. & Mayer, H. (1991a).** Distribution and phylogenetic significance of 27-hydroxy-octacosanoic acid in lipopolysaccharides from bacteria belonging to the alpha-2 subgroup of Proteobacteria. *Int J Syst Bacteriol* **41**, 213-217.

**Bhat, U. R., Mayer, H., Yokota, A., Hollingsworth, R. I. & Carlson, R. W. (1991b).** Occurrence of lipid A variants with 27-hydroxyoctacosanoic acid in lipopolysaccharides from members of the family *Rhizobiaceae*. *J Bacteriol* **173**, 2155-2159.

- Birnboim, H. C. & Doly, J. (1979).** A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**, 1513-1523.
- Bisang, C., Long, P. F., Cortes, J. & other authors (1999).** A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* **401**, 502-505.
- Bligh, E. G. & Dyer, W. J. (1959).** A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**, 911-917.
- Bliss, J. M. & Silver, R. P. (1996).** Coating the surface: a model for expression of capsular polysialic acid in *Escherichia coli* K1. *Mol Microbiol* **21**, 221-231.
- Boulnois, G. J., Roberts, I. S., Hodge, R., Hardy, K. R., Jann, K. B. & Timmis, K. N. (1987).** Analysis of the K1 capsule biosynthesis genes of *Escherichia coli*: definition of three functional regions for capsule production. *Mol Gen Genet* **208**, 242-246.
- Breedveld, M. W., Hadley, J. A. & Miller, K. J. (1995).** A novel cyclic beta-1,2-glucan mutant of *Rhizobium meliloti*. *J Bacteriol* **177**, 6346-6351.
- Broughton, W. J., Jabbouri, S. & Perret, X. (2000).** Keys to Symbiotic Harmony. *J Bacteriol* **182**, 5641-5652.
- Brown, N. L., Barrett, S. R., Camakaris, J., Lee, B. T. & Rouch, D. A. (1995).** Molecular genetics and transport analysis of the copper-resistance determinant (*pco*) from *Escherichia coli* plasmid pRJ1004. *Mol Microbiol* **17**, 1153-1166.
- Brozek, K. A., Carlson, R. W. & Raetz, C. R. H. (1996).** A Special Acyl Carrier Protein for Transferring Long Hydroxylated Fatty Acids to Lipid A in *Rhizobium*. *J Biol Chem* **271**, 32126-32136.
- Capela, D., Barloy-Hubler, F., Gouzy, J. & other authors (2001).** From the Cover: Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc Natl Acad Sci U S A* **98**, 9877-9882.
- Carreras, C. W., Gehring, A. M., Walsh, C. T. & Khosla, C. (1997).** Utilization of enzymatically phosphopantetheinylated acyl carrier proteins and acetyl-acyl carrier proteins by the actinorhodin polyketide synthase. *Biochemistry* **36**, 11757-11761.
- Cieslewicz, M. & Vimr, E. (1996).** Thermoregulation of *kpsF*, the first region 1 gene in the *kps* locus for polysialic acid biosynthesis in *Escherichia coli* K1. *J Bacteriol* **178**, 3212-3220.

- Cieslewicz, M. & Vimr, E. (1997).** Reduced polysialic acid capsule expression in *Escherichia coli* K1 mutants with chromosomal defects in *kpsF*. *Mol Microbiol* **26**, 237-249.
- D'Haese, W. & Holsters, M. (2002).** Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* **12**, 79R-105.
- Dazzo, F. B., Truchet, G. L., Hollingsworth, R. I. & other authors (1991).** Rhizobium lipopolysaccharide modulates infection thread development in white clover root hairs. *J Bacteriol* **173**, 5371-5384.
- Drummelsmith, J. & Whitfield, C. (2000).** Translocation of group 1 capsular polysaccharide to the surface of *Escherichia coli* requires a multimeric complex in the outer membrane. *EMBO J* **19**, 57-66.
- Dunlap, J., Minami, E., Bhagwat, A. A., Keister, D. L. & Stacey, G. (1996).** Nodule development induced by mutants of *Bradyrhizobium japonicum* defective in cyclic B-glucan synthesis. *Mol Plant-Microbe Interact* **9**, 546-555.
- Eckermann, S., Schroder, G., Schmidt, J. & other authors (1998).** New pathway to polyketides in plants. *Nature* **396**, 387-390.
- Epple, G. (1995).** Klonierung und Expression eines neuartigen Acyl-Carrier-Proteins aus dem *fix-23*-Locus von *Rhizobium meliloti*. In *Institut für Biotechnologie; Fachgebiet Technische Biochemie*, pp. 53. Berlin; Germany: Technische Universität Berlin.
- Epple, G., van der Drift, K. M., Thomas-Oates, J. E. & Geiger, O. (1998).** Characterization of a novel acyl carrier protein, RkpF, encoded by an operon involved in capsular polysaccharide biosynthesis in *Sinorhizobium meliloti*. *J Bacteriol* **180**, 4950-4954.
- Finan, T. M., Weidner, S., Wong, K. & other authors (2001).** From the Cover: The complete sequence of the 1,683-kb pSymB megaplasmid from the N<sub>2</sub>-fixing endosymbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A* **98**, 9889-9894.
- Fisher, R. F. & Long, S. R. (1992).** Rhizobium--plant signal exchange. *Nature* **357**, 655-660.
- Forsberg, L. S. & Reuhs, B. L. (1997).** Structural characterization of the K antigens from *Rhizobium fredii* USDA257: evidence for a common structural motif, with strain-specific variation, in the capsular polysaccharides of Rhizobium spp. *J Bacteriol* **179**, 5366-5371.

- Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A. & Perret, X. (1997).** Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387**, 394-401.
- Frosch, M., Edwards, U., Bousset, K., Krausse, B. & Weisgerber, C. (1991).** Evidence for a common molecular origin of the capsule gene loci in gram-negative bacteria expressing group II capsular polysaccharides. *Mol Microbiol* **5**, 1251-1263.
- Frosch, M., Muller, D., Bousset, K. & Muller, A. (1992).** Conserved outer membrane protein of *Neisseria meningitidis* involved in capsule expression. *Infect Immun* **60**, 798-803.
- Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G. & Muller, R. (2002).** The Biosynthesis of the Aromatic Myxobacterial Electron Transport Inhibitor Stigmatellin Is Directed by a Novel Type of Modular Polyketide Synthase. *J Biol Chem* **277**, 13082-13090.
- Galibert, F., Finan, T. M., Long, S. R. & other authors (2001).** The Composite Genome of the Legume Symbiont *Sinorhizobium meliloti*. *Science* **293**, 668-672.
- Geiger, O., Spaink, H. P. & Kennedy, E. P. (1991).** Isolation of the *Rhizobium leguminosarum* NodF nodulation protein: NodF carries a 4'-phosphopantetheine prosthetic group. *J Bacteriol* **173**, 2872-2878.
- Geiger, O., Thomas-Oates, J., Glushka, J., Spaink, H. & Lugtenberg, B. (1994).** Phospholipids of *Rhizobium* contain nodE-determined highly unsaturated fatty acid moieties. *J Biol Chem* **269**, 11090-11097.
- Geiger, O., Glushka, J., Lugtenberg, B. J., Spaink, H. P. & Thomas-Oates, J. E. (1998).** NodFE-dependent fatty acids that lack an alpha-beta unsaturation are subject to differential transfer, leading to novel phospholipids. *Mol Plant-Microbe Interact* **11**, 33-44.
- Geiger, O. & Lopez-Lara, I. M. (2002).** Rhizobial acyl carrier proteins and their roles in the formation of bacterial cell-surface components that are required for the development of nitrogen-fixing root nodules on legume hosts. *FEMS Microbiol Lett* **208**, 153-162.
- Ghose, R., Geiger, O. & Prestegard, J. H. (1996).** NMR investigations of the structural properties of the nodulation protein, NodF, from *Rhizobium leguminosarum* and its homology with *Escherichia coli* acyl carrier protein. *FEBS Lett* **388**, 66-72.

- Gil-Serrano, A. M., Gonzalez-Jimenez, I., Tejero-Mateo, P., Megias, M. & Romero-Vazquez, M. J. (1994).** Analysis of the lipid moiety of lipopolysaccharide from *Rhizobium tropici* CIAT899: identification of 29-hydroxytriacontanoic acid. *J Bacteriol* **176**, 2454-2457.
- Hanahan, D. (1983).** Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.
- He, J. & Hertweck, C. (2003).** Iteration as Programmed Event during Polyketide Assembly; Molecular Analysis of the Aureothin Biosynthesis Gene Cluster. *Chem Biol* **10**, 1225-1232.
- Hollingsworth, R. & Carlson, R. (1989).** 27-Hydroxyoctacosanoic acid is a major structural fatty acyl component of the lipopolysaccharide of *Rhizobium trifolii* ANU 843. *J Biol Chem* **264**, 9300-9303.
- Hotter, G. S. & Scott, D. B. (1991).** Exopolysaccharide mutants of *Rhizobium loti* are fully effective on a determinate nodulating host but are ineffective on an indeterminate nodulating host. *J Bacteriol* **173**, 851-859.
- Itzhaki, R. F. & Gill, G. M. (1964.).** A micro-biuret method for estimating protein. *Anal Biochem* **9**, 401-410.
- Jackowski, S. & Rock, C. (1983).** Ratio of active to inactive forms of acyl carrier protein in *Escherichia coli*. *J Biol Chem* **258**, 15186-15191.
- Jacoby, G. A., Jacob, A. E. & Hedges, R. W. (1976).** Recombination between plasmids of incompatibility groups P-1 and P-2. *J Bacteriol* **127**, 1278-1285.
- Jerlstrom, P. G., Bezjak, D. A., Jennings, M. P. & Beacham, I. R. (1989).** Structure and expression in *Escherichia coli* K-12 of the L-asparaginase I-encoding ansA gene and its flanking regions. *Gene* **78**, 37-46.
- Kereszt, A., Kiss, E., Reuhs, B. L., Carlson, R. W., Kondorosi, A. & Putnoky, P. (1998).** Novel rkp gene clusters of *Sinorhizobium meliloti* involved in capsular polysaccharide production and invasion of the symbiotic nodule: the rkpK gene encodes a UDP-glucose dehydrogenase. *J Bacteriol* **180**, 5426-5431.
- Kijne, J. W. (1992).** The Rhizobium infection process. In *Biological Nitrogen Fixation*, pp. 293-398. Edited by G. B. Stacey, R.H. and Evans, H.J. New York: Chapman.
- Kiss, E., Reuhs, B. L., Kim, J. S., Kereszt, A., Petrovics, G., Putnoky, P., Dusha, I., Carlson, R. W. & Kondorosi, A. (1997).** The rkpGHI and -J genes

are involved in capsular polysaccharide production by *Rhizobium meliloti*. *J Bacteriol* **179**, 2132-2140.

**Kiss, E., Kereszt, A., Barta, F., Stephens, S., Reuhs, B. L., Kondorosi, A. & Putnoky, P. (2001).** The rkp-3 gene region of *Sinorhizobium meliloti* Rm41 contains strain-specific genes that determine K antigen structure. *Mol Plant-Microbe Interact* **14**, 1395-1403.

**Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., 2nd & Peterson, K. M. (1994).** pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16**, 800-802.

**Kretzschmar, U., Schobert, M. & Gorisch, H. (2001).** The *Pseudomonas aeruginosa* acsA gene, encoding an acetyl-CoA synthetase, is essential for growth on ethanol. *Microbiology* **147**, 2671-2677.

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

**Lambalot, R. H. & Walsh, C. T. (1995).** Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *J Biol Chem* **270**, 24658-24661.

**Liesegang, H., Lemke, K., Siddiqui, R. A. & Schlegel, H. G. (1993).** Characterization of the inducible nickel and cobalt resistance determinant cnr from pMOL28 of *Alcaligenes eutrophus* CH34. *J Bacteriol* **175**, 767-778.

**Long, S. R. (1989).** Rhizobium-legume nodulation: life together in the underground. *Cell* **56**, 203-214.

**Long, S. R. (1996).** Rhizobium symbiosis: nod factors in perspective. *Plant Cell* **8**, 1885-1898.

**Lopez-Lara, I. M. & Geiger, O. (2000).** Expression and purification of four different rhizobial acyl carrier proteins. *Microbiology* **146**, 839-849.

**Miller, J. K. (1972).** Experiments in Molecular Genetics. pp. 431-432. Cold Spring Harbor, NY: Cold Spring Harbour Laboratory.

**Molnar, I., Schupp, T., Ono, M. & other authors (2000).** The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chem Biol* **7**, 97-109.

**Nesterenko, M. V., Tilley, M. & Upton, S. J. (1994).** A simple modification of Blum's silver stain method allows for 30 minute detection of proteins in

polyacrylamide gels. *Journal of Biochemical and Biophysical Methods* **28**, 239-242.

**Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H. & Yura, T. (1998).** Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*. *Appl Environ Microbiol* **64**, 1694-1699.

**Pazzani, C., Rosenow, C., Boulnois, G. J., Bronner, D., Jann, K. & Roberts, I. S. (1993).** Molecular analysis of region 1 of the *Escherichia coli* K5 antigen gene cluster: a region encoding proteins involved in cell surface expression of capsular polysaccharide. *J Bacteriol* **175**, 5978-5983.

**Petrovics, G., Putnoky, P., Reuhs, B., Kim, J., Thorp, T. A., Noel, K. D., Carlson, R. W. & Kondorosi, A. (1993).** The presence of a novel type of surface polysaccharide in *Rhizobium meliloti* requires a new fatty acid synthase-like gene cluster involved in symbiotic nodule development. *Mol Microbiol* **8**, 1083-1094.

**Pieper, R., Luo, G., Cane, D. E. & Khosla, C. (1995).** Cell-free synthesis of polyketides by recombinant erythromycin polyketide synthases. *Nature* **378**, 263-266.

**Platt, M. W., Miller, K. J., Lane, W. S. & Kennedy, E. P. (1990).** Isolation and characterization of the constitutive acyl carrier protein from *Rhizobium meliloti*. *J Bacteriol* **172**, 5440-5444.

**Putnoky, P., Grosskopf, E., Ha, D. T., Kiss, G. B. & Kondorosi, A. (1988).** *Rhizobium* fix genes mediate at least two communication steps in symbiotic nodule development. *J Cell Biol* **106**, 597-607.

**Putnoky, P., Petrovics, G., Kereszt, A., Grosskopf, E., Ha, D. T., Banfalvi, Z. & Kondorosi, A. (1990).** *Rhizobium meliloti* lipopolysaccharide and exopolysaccharide can have the same function in the plant-bacterium interaction. *J Bacteriol* **172**, 5450-5458.

**Putnoky, P., Deak, V., Bekasi, K., Palvolgyi, A., Maasz, A., Palagyi, Z., Hoffmann, G. & Kerepesi, I. (2004).** H Protein of Bacteriophage 16-3 and RkpM Protein of *Sinorhizobium meliloti* 41 Are Involved in Phage Adsorption. *J Bacteriol* **186**, 1591-1597.

**Rastogi, V. K. & Watson, R. J. (1991).** Aspartate aminotransferase activity is required for aspartate catabolism and symbiotic nitrogen fixation in *Rhizobium meliloti*. *J Bacteriol* **173**, 2879-2887.

**Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H. & Hopwood, D. A. (1996).** A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. *Mol Microbiol* **21**, 77-96.

**Reinhold, B. B., Chan, S. Y., Reuber, T. L., Marra, A., Walker, G. C. & Reinhold, V. N. (1994).** Detailed structural characterization of succinoglycan, the major exopolysaccharide of *Rhizobium meliloti* Rm1021. *J Bacteriol* **176**, 1997-2002.

**Reuber, T. L. & Walker, G. C. (1993).** Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* **74**, 269-280.

**Reuhs, B. L., Carlson, R. W. & Kim, J. S. (1993).** *Rhizobium fredii* and *Rhizobium meliloti* produce 3-deoxy-D-manno-2-octulosonic acid-containing polysaccharides that are structurally analogous to group II K antigens (capsular polysaccharides) found in *Escherichia coli*. *J Bacteriol* **175**, 3570-3580.

**Reuhs, B. L., Williams, M. N., Kim, J. S., Carlson, R. W. & Cote, F. (1995).** Suppression of the Fix- phenotype of *Rhizobium meliloti* *exoB* mutants by *lpsZ* is correlated to a modified expression of the K polysaccharide. *J Bacteriol* **177**, 4289-4296.

**Reuhs, B. L. (1996).** Acidic capsular polysaccharides (K antigens) of *Rhizobium*. In *Biology of Plant-Microbe Interactions*, pp. 331-336. Edited by G. Stacey, B. Mullin & P. M. Gresshoff. St. Paul.: International Society for Molecular Plant-Microbe Interactions,.

**Reuhs, B. L., Geller, D. P., Kim, J. S., Fox, J. E., Kolli, V. S. & Pueppke, S. G. (1998).** *Sinorhizobium fredii* and *Sinorhizobium meliloti* produce structurally conserved lipopolysaccharides and strain-specific K antigens. *Appl Environ Microbiol* **64**, 4930-4938.

**Rick, P. D. & Silver, R. P. (1996).** Enterobacterial common antigen and capsular polysaccharides. In *Escherichia coli and Salmonella*, pp. 104-122. Edited by F. C. Neidhard. Washington D. C.: ASM Press.

**Ritsema, T., Geiger, O., van Dillewijn, P., Lugtenberg, B. J. & Spaink, H. P. (1994).** Serine residue 45 of nodulation protein NodF from *Rhizobium leguminosarum* bv. *viciae* is essential for its biological function. *J Bacteriol* **176**, 7740-7743.

**Ritsema, T., Gehring, A. M., Stuitje, A. R. & other authors (1998).** Functional analysis of an interspecies chimera of acyl carrier proteins indicates a specialized domain for protein recognition. *Mol Gen Genet* **257**, 641-648.

**Roberts, G. A., Staunton, J. & Leadlay, P. F. (1993).** Heterologous expression in *Escherichia coli* of an intact multienzyme component of the erythromycin-producing polyketide synthase. *Eur J Biochem* **214**, 305-311.

**Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

**Sanger, F., Nicklen, S. & Coulson, A. R. (1977).** DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467.

**Schägger, H. & von Jagow, G. (1987).** Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368-379.

**Shen, B. (2003).** Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **7**, 285-295.

**Silakowski, B., Schairer, H. U., Ehret, H. & other authors (1999).** New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J Biol Chem* **274**, 37391-37399.

**Simpson, D. A., Hammarton, T. C. & Roberts, I. S. (1996).** Transcriptional organization and regulation of expression of region 1 of the *Escherichia coli* K5 capsule gene cluster. *J Bacteriol* **178**, 6466-6474.

**Spaink, H. P., Sheeley, D. M., van Brussel, A. A. & other authors (1991).** A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature* **354**, 125-130.

**Spaink, H. P., Wijffjes, A. H. & Lugtenberg, B. J. (1995).** Rhizobium NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides. *J Bacteriol* **177**, 6276-6281.

**Stinear, T. P., Mve-Obiang, A., Small, P. L. C. & other authors (2004).** From the Cover: Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc Natl Acad Sci U S A* **101**, 1345-1349.

**Streit, W. R. & Phillips, D. A. (1997).** A biotin-regulated locus, bioS, in a possible survival operon of *Rhizobium meliloti*. *Mol Plant-Microbe Interact* **10**, 933-937.

- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990).** Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89.
- Sun, Y., Zhou, X., Dong, H., Tu, G., Wang, M., Wang, B. & Deng, Z. (2003).** A Complete Gene Cluster from *Streptomyces nanchangensis* NS3226 Encoding Biosynthesis of the Polyether Ionophore Nanchangmycin. *Chem Biol* **10**, 431-441.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994).** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.
- van Rhijn, P. & Vanderleyden, J. (1995).** The Rhizobium-plant symbiosis. *Microbiol Rev* **59**, 124-142.
- Vasse, J., de Billy, F., Camut, S. & Truchet, G. (1990).** Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J Bacteriol* **172**, 4295-4306.
- Vinitsky, A., Teng, H. & Grubmeyer, C. T. (1991).** Cloning and nucleic acid sequence of the *Salmonella typhimurium* pncB gene and structure of nicotinate phosphoribosyltransferase. *J Bacteriol* **173**, 536-540.
- Wajima, Y., Furusawa, T., Kawauchi, S., Wakabayashi, N., Nakabayashi, O., Nishimori, K. & Mizuno, S. (1999).** The cDNA cloning and transient expression of an ovary-specific 17beta-hydroxysteroid dehydrogenase of chickens. *Gene* **233**, 75-82.
- Wang, P., Ingram-Smith, C., Hadley, J. A. & Miller, K. J. (1999).** Cloning, sequencing, and characterization of the *cgmB* gene of *Sinorhizobium meliloti* involved in cyclic beta-glucan biosynthesis. *J Bacteriol* **181**, 4576-4583.
- Wek, R. C., Hauser, C. A. & Hatfield, G. W. (1985).** The nucleotide sequence of the *ilvBN* operon of *Escherichia coli*: sequence homologies of the acetohydroxy acid synthase isozymes. *Nucleic Acids Res* **13**, 3995-4010.
- Whitfield, C. & Paiment, A. (2003).** Biosynthesis and assembly of Group 1 capsular polysaccharides in *Escherichia coli* and related extracellular polysaccharides in other bacteria. *Carbohydr Res* **338**, 2491-2502.
- Wilkinson, B., Foster, G., Rudd, B. A. & other authors (2000).** Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase. *Chem Biol* **7**, 111-117.

---

**Xue, Q., Ashley, G., Hutchinson, C. R. & Santi, D. V. (1999).** A multiplasmid approach to preparing large libraries of polyketides. *Proc Natl Acad Sci U S A* **96**, 11740-11745.

**Yang, G.-P., Debelle, F., Savagnac, A. & other authors (1999).** Structure of the *Mesorhizobium huakuii* and *Rhizobium galegae* Nod factors: a cluster of phylogenetically related legumes are nodulated by rhizobia producing Nod factors with alpha,beta-unsaturated - *N*-acyl substitutions. *Mol Microbiol* **34**, 227-237.

**Yanisch-Perron, C., Vieira, J. & Messing, J. (1985).** Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

## 7 Anhang

### 7.1 Genbankeintrag X64131

LOCUS RMXFIX23 12965 bp DNA linear BCT 14-DEC-2001

DEFINITION R.meliloti DNA for fix23 ORF1, ORF2, ORF3, ORF4, ORF5 and ORF6.

ACCESSION X64131

VERSION X64131.1 GI:1235585

KEYWORDS Fatty acid synthase; fix23 gene; nodule development.

SOURCE Sinorhizobium meliloti (Rhizobium meliloti)

ORGANISM Sinorhizobium meliloti  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

REFERENCE 1 (bases 1 to 7542)

AUTHORS Petrovics,G., Putnoky,P., Reuhs,B., Kim,J., Thorp,T.A., Noel,K.D., Carlson,R.W. and Kondorosi,A.

TITLE The presence of a novel type of surface polysaccharide in Rhizobium meliloti requires a new fatty acid synthase-like gene cluster involved in symbiotic nodule development

JOURNAL Mol. Microbiol. 8 (6), 1083-1094 (1993)

MEDLINE 93368424

PUBMED 8361353

REFERENCE 2

AUTHORS Kiss,E., Reuhs,B.L., Kim,J.S., Kereszt,A., Petrovics,G., Putnoky,P., Dusha,I., Carlson,R.W. and Kondorosi,A.

TITLE The rkpGHI and -J genes are involved in capsular polysaccharide production by Rhizobium meliloti

JOURNAL J. Bacteriol. 179 (7), 2132-2140 (1997)

MEDLINE 97234622

PUBMED 9079896

REFERENCE 3

AUTHORS Petrovics,G.

TITLE Direct Submission

JOURNAL Submitted (13-JAN-1992) G. Petrovics, B.R.C. Institute of Genetics, Hungarian Academy of Sciences, H-6701 Szeged, Temesvari krt 62, P

O

REMARK Box 521, HUNGARY Revised by [3]

REFERENCE 4 (bases 1 to 12965)

AUTHORS Kiss,E.

TITLE Direct Submission

JOURNAL Submitted (18-MAR-1996) Kiss E., B.R.C. Institute of Genetics, Hungarian Academy of Sciences, H-6701 Szeged, Temesvari krt 62, P

O

COMMENT Box 521, HUNGARY e-mail: rhizo@everx.szbk.u-szeged.hu  
On Mar 21, 1996 this sequence version replaced gi:46270.  
Genes in the fix23 region are clustered on the chromosome close to the pyr-2 and pyr-29 markers. The predicted sequence of the fix23 ORF's are homologous to different fatty acid synthase and polyketide synthase enzymes, especially around the active sites.

FEATURES Location/Qualifiers

source 1..12965  
/organism="Sinorhizobium meliloti"  
/mol\_type="genomic DNA"  
/strain="41"  
/db\_xref="taxon:382"  
/map="between the met-5 and phe-15 auxotrophic markers"  
/cell\_line="AK631"  
/clone\_lib="pLAFR1"  
/note="RKP-I. (former fix-23) chromosomal region"

-35\_signal 290..295

```

-10_signal      312..314
RBS             355..358
gene           367..1485
              /gene="rkpA"
CDS            367..1485
              /gene="rkpA"
              /codon_start=1
              /transl_table=11
              /product="Fix23-1"
              /protein_id="CAA45483.1"
              /db_xref="GI:46271"
              /db_xref="GOA:Q52930"
              /db_xref="TrEMBL:Q52930"

/translation="MDPQQRVLLQLAWRALEDANISVASLHGENVG VYVGASSLDHAN
ADGRRSAAAGPYFMTGNTLSIVSNRISHVFGLSGPSMTVDTACSSSLVALDQAMRALN
AGEIDTAIVGGVNI LAHPLPFVGFQAARMLSP EGLCRAYDNDGAGYVRAEGGVVFLR
RTDRAQRERDRSYARIVASGVNSAGRTNGISLPSREAQANLLRMIYEGNGIDANQVAF
VEGHGTGTKVGDPAE VWSIGTVIGAKRAPVPIGSIKSNIGHTEPASGLFGMMKAVLAL
ENNYLPASLHFETPNEQIDFDGLNVRVTANPIELLK GKRRARLAGINSFGFGGANAHVV
ISDPDPAQA EKAATSPAGHVFLASAHTASSLENLLERL"
RBS            1550..1552
gene           1558..2745
              /gene="rkpB"
CDS            1558..2745
              /gene="rkpB"
              /codon_start=1
              /transl_table=11
              /product="Fix23-2"
              /protein_id="CAA45484.1"
              /db_xref="GI:46272"
              /db_xref="GOA:Q52931"
              /db_xref="TrEMBL:Q52931"

/translation="MRHRFAARSDHPEDIVRAIANHLEKPGSDIGETGEAAAKEAKVA
FVFSGNGSQWAGMGVEAFRENLFHRQSFTSVSALFRFHSDIVLTDLLTDPGLDRKLD
TKVAQPLLFAIQ AALSDSL VAMGIKPTAVFGHSVGEIAAAYAAGALSLVDAVSIVAKR
SLHQDLLAGQGTMAAVMLGEEAAKAF AAERGLDDVCVAAINAHNSVTISGPANEISAF
RDAARKAKI PVQIILDINYPFHHP I IDRAKEAFLSDIPDIAPRRTELAYLSTVTGAALD
GTALDPDYWWKNVREPVRFQAAA E AALDLGCMLFIEISPRPILGSYVKETIKQAAVPA
SVVATLLRDAGANGHDPISASMARAVAHGA AVDRSRVYGKRDAFIELPALPFEPVELR
PAA"
RBS            2724..2727
              /gene="rkpB"
gene           2749..5649
              /gene="rkpC"
CDS            2749..5649
              /gene="rkpC"
              /codon_start=1
              /transl_table=11
              /product="Fix23-3"
              /protein_id="CAA45485.1"
              /db_xref="GI:46273"
              /db_xref="GOA:Q52932"
              /db_xref="TrEMBL:Q52932"

/translation="MRRILFGRSAKPYRLNGWRGDPNAGSWKNHVDAHLFPDLAEHVV

```

```

DGKAILPGSGFIEIAVSAAQQYHGSDEVEITNLEIVRPLELSDSRIMELSTILSPETG
DIEIRSRERLSEDDWTVHAVARSRKPLPTTSNDPCSTKGLAKTATVTAAKAYETAKQF
GLDYGPQFQLLSKAVAYGDRLIDVELKHPQASGHPVTYALHPI SVDATFHGLVALFD
RFTGDRGGAYIPVRF GSVRVVNAGRPIARATIEIERVSANSIKARFHF FDETGVIAIAH
FEDCRFRRTYL RKHKTLEGLSFHYEAVLSDVVMPTKVATAMPATLAEPI DDNGVDNA
TLLEFNAAIYRACQEI ALKLGRTATVVRGDALPGDFAFQCFLTSCLLTLE DAGLCEHQN
GNWKVAREFSLPTVPEILGELYGDRPDRAVEAVLVNAYAEALSRLDALLGPVRAGDD
AASFNAGFISEATLDHQAVHSVASRSRMDQVLHTVEHVLAAQPAGACGSSNSAASPAS
AAGWPIWQRAMVQHFPFSSRATMHNGIWKSPSRMMNVRVLKKSEFADIGPFDLAVSAS
DNLYQLIEEESGVRTALRSMPGSGLVAAVSAPSI FADFALGLSGGWFFERSQTAEFPI
GRIAAVPHWQKCLADLDLGDVVVTDRECPHGNTILIEARGTAAAGAGAEPGETGTPYL
LIETATAQKPVAAGGRSAAPVVRVTGDLAADMATLRAALEAMGDRPLRAVFMSEAERAV
GDGSALLQTRVLALSFAFAEALKQHLAGVEPRATTGAPCARGPGRRLITSAGSGLNSG
LWAFARVLQNEYEFIDVHTLDLAGDADIGRKDMLAAALPLLTVSGQNREWLLDRKTGL
LSELRAVPGATNKVDGKTNAAATIRQRVSSQVASIAWEEASVPEPGPNEVVAVAA
TGLNFRDVMWAMGLLP EEAELEDGSPGPSAWNC PATWSRS AKRSTI SRSAMR "
RBS 5717..5719
gene 5730..6296
    /gene="rkpD"
CDS 5730..6296
    /gene="rkpD"
    /codon_start=1
    /transl_table=11
    /product="Fix23-4"
    /protein_id="CAA45486.1"
    /db_xref="GI:581519"
    /db_xref="GOA:Q52933"
    /db_xref="TrEMBL:Q52933"

/translation="MAAATLPVAFLTAYYAIVELGRIQAGETILIHGGSGVGLAALQ
VAKLKGAKVIATAGSAEKRRFLTMLGADHVFD SRSLSGFVEDVRSVTTGGEGVDLVLNSL
FAEAMEQSLSLVKPFGRFLELGKRDYYADSKIGLRPFRRNVSYFGIDADQLLVNAPDL
TRRIFMQIGALFEEGNVTPLPYRAF'DYG"
RBS 6295..6297
gene 6319..7173
    /gene="rkpE"
CDS 6319..7173
    /gene="rkpE"
    /codon_start=1
    /transl_table=11
    /product="Fix23-5"
    /protein_id="CAA45487.1"
    /db_xref="GI:46275"
    /db_xref="TrEMBL:Q52934"

/translation="MQNAGHIGKIVVLPVAGRHEVTAKAVRGMKVDPDGVHLVVGGI
GGFGLVAANWLVEKGARRIALCSRRGQPD AETRAMIELRAEWRRCGFGPCLRHHGCRG
RGDASGDASFGSTVRSVVHAAMVLDDALI SNLSREERNRPVIETKAKGAAILDRLTRGD
RLDNF ILFSSATTLVGNPGQANYVAANGYLEGLARARRQ ELAGLAIGFGAIADAGYL

```

```

TQNADVNDLLAKRIGKTALKAQVALDMVENHVAADPGTVDAAVVMI SEIDWTAARNLP
      VAATRSSR "
      RBS      7223..7225
      gene     7236..7505
              /gene="rkpF"
      CDS      7236..7505
              /gene="rkpF"
              /codon_start=1
              /transl_table=11
              /product="Fix23-6"
              /protein_id="CAA45488.1"
              /db_xref="GI:46276"
              /db_xref="TrEMBL:Q52935"

/translation="MIEGKSPQEAEDILFDLVAGEIAAILRVSKDVTTRGKILKEIGL
      DSLMAVELGMSFQQNTGFDMP LSGVADNTTVGDIARKLHEKVS KP "
      gene     7671..9011
              /gene="rkpG"
      CDS      7671..9011
              /gene="rkpG"
              /codon_start=1
              /transl_table=11
              /product="acyl-transferase"
              /protein_id="CAA45489.1"
              /db_xref="GI:1235586"
              /db_xref="GOA:Q52936"
              /db_xref="TrEMBL:Q52936"

/translation="MRNTHQSSERNRMAR SERELPPAPRRQQARFEDLPEYKQVLTQK
IASEQLGIANPFYRPHQTAAGAT TMIDGRKLINFASYDYLG LNRHAHVLERARETIAG
FGISASASRLVAGERPQHVELEEKIAQFYGVDAAVCFVSGYLTNVAAI SCLMGPKDLV
IHDEFIHNSALAGIKLSGATR RFFKHNDTADLEHVLRTVAGDYRRILVIVEGIYSMDG
DVANLPALLKLR AEYGFWLMVDEAHS LGVLGRHGRGLAEHFGADPHEVDIWMG TLSKT
TSSCGGYIAGSAAALAAVLKASAGGFVYSVGLAPVLAASAVASLDILASEPERTAAVRR
NGSLFLKLAKEAGLDTGLSGGFSVVPVIVGDSLRAVQLSNDLLAAGNVLPIIHPAVPE
      GQARLRVLHHLRPHRGTDPPHRHIGAERLKDLTDRNFGLGRVDIEKMLSAL SAC "
      gene     9048..9887
              /gene="rkpH"
      CDS      9048..9887
              /gene="rkpH"
              /codon_start=1
              /transl_table=11
              /product="ribitol type dehydrogenase"
              /protein_id="CAA45490.1"
              /db_xref="GI:1235587"
              /db_xref="GOA:Q52937"
              /db_xref="TrEMBL:Q52937"

/translation="MTHVIITGGSSGIGLAVASIYAARSRVSLVAR SADLLEKAAQKL
VAEHGLEAGAVRIEAADVSKGEEIEAAVFR CVDAFGPCDVLVTSAGVVEPAPFEAMQG
SAFHRQMETNFS GTVHAVRAVYPDMKNRPRHILMVSSGAGLIGIYGTAYCASKFALN
GFARRCAANARPQCRHFDLLSPDTETPQFKRELAARPVEASVIMGTVRPWTAEAVARK
IVKIDRRRFEIYFGAALYLLPLRAGGQAFPQLVVRPGDRAQQWTVTAAPAAMLVPAN
      IRL "
      gene     9922..11535
              /gene="rkpI"
      CDS      9922..11535
              /gene="rkpI"

```

```

/codon_start=1
/transl_table=11
/protein_id="CAA45491.1"
/db_xref="GI:1235588"
/db_xref="GOA:Q52938"
/db_xref="Swiss-Prot:Q52938"

/translation="MKWLKGIQLALIPFNLHDYPIALTLGCYLLSCAVIFYTDRFALP
ARERRKNAPKYGRNRDRIDHLARLPVIALVFAGFFAISWRPLYAAAGTMSFFIIIFTGI
SRAKFKFIREPLVFSIDIALVADVFKYKTI FYASSLNIVFWIVAFLYVFGVSALYMYFE
PTILPERSRLFVWLVMVGIAAGPWGLLFYGPVNRPTAALVQRLVKA INVKMNTVRFGT
FASVVFHFIIWLGVKRDKIVAELSEILRAAVHDLIGHEEAPLIIVWQSESFIDMRHF
VDSIKLPTVDRLRKQAVQWGRLSNVFEGGYTLRTEFAVL SGLVPDDIHVDASYPYLR
AHYADIVWPGKLRGWRTHFIHPYDRTFFLRHKAMPLLGFEKLTMLDAFDHKPERDGL
YVSDATLAARVLA EVEKLP EEESGFFFVASMANGHPWEPGRVGTLTNPVDIYLAILEQ
SDAALKQLIDGLNKLDRPVWFVFGDHAPLLKSFADFPDPRTDYFIVPLAKARAAAH
SPKRRKDEDPWNLLGSMLKXANLHKDALQ"
gene 11560..12798
      /gene="rkJ"
CDS 11560..12798
     /gene="rkJ"
     /codon_start=1
     /transl_table=11
     /protein_id="CAA45492.1"
     /db_xref="GI:1235589"
     /db_xref="TrEMBL:Q52939"

/translation="MTAVTTIVPAPRPTFLFLQGPASRVFAKVAGHLEVLGHCCLRIN
LNPGDQIFWRRRAAITTVDSLARWPAYVEAFIRRHAVTDLVLLGEERPYPHRIATAAAR
KAGVRVVFVEMGYLRPDWLTLERGMSNSHFPVDPAQILRAAMGLPEPDWQRLYRQS
FLAEAAAYDLLYNLNVFFCFLFRYPRHGI FHPLAEYAGWIRIRRLAAGKREQRADAL
IRLSAKNAPYFVYPLQLETDYQLRAHSPFGSQREAI DGI LASFAHRA PAGTRLAVKV
HPLDNSLISWRRIVADKAASLVGVDRLVYLDGGNLDLSL TEHSCGM/TVNSTAGLHALK
QKPKVKVLGKAVFDIAGLTDQQPLDAFWTAPEADPVL SGAMFRLMAASI QVRGNFY S
VRVRMPAARPSPQAAPQRRQ"

ORIGIN
  1 ggggataaaa cgggggatac cgcagagca ggaatgtcgc cgattgcaat ttacgttaga
  61 ccttgcaatg ggctaagaca attttaacgg cagttcagag cgttcgatct tacatgacgg
 121 ttgaaatcat agggcgtgca tgccttgccc cgggagcgaa atctccgcag gcgctgttca
 181 aaatccttcg tcagggcaag tgcaccgtgg acccgcgtcc cttcggaccg ctggggacctt
 241 gcgcgtttct gccacccggg catgggcacg cccggcaaga cctattcggt tgccgcccgc
 301 gtcctcgacc atatctacga ttctgatccc gccgttttcg ggatgtcgca gcgcgaggcc
 361 atgtatatgg acccgagca gcgggttctg ttgcagctcg cctggcgagc gctcgaggac
 421 gccaatattt ccgtggcgtc gctgcacggt gaaaatgtcg gcgtctatgt cggcgctcc
 481 agcctcgacc acgccaatgc tgacggctga cgatcggctg ccgcccggcc ctacttcatg
 541 acgggcaata cgctgtgat cgtctccaac cgtatttccc acgtcttcgg cctgagcggc
 601 ccgagcatga ccgtcgacac ggcctgttcc tcttcgctcg tggccctcga tcaggcgatg
 661 cgggcgctga acgcccgaga gatcgacacg gccatcgctg gcggcgtaa tatcctcggc
 721 caccgccttc ccttcgctcg cttcgcccag gcgcgcatgc tctcgcccga agggctctgc
 781 cgcgcctatg acaatgacgg cgccggctat gttcgcgccg agggcggcgt cgtcttcgctc
 841 ctgcgcccga cggaccgggc gcaacgcgag cgcgaccgga gctatgccc gacgctgcgc
 901 tccggcgtca attcggccgg gcgcaaccaac ggcattctcg tgccctcgg cgaggcgcag
 961 gccaatctgc tgaggatgat ctacgaaggc aacggcatcg acgccaacca ggttgccctc
1021 gtagaaggcc atgggtaccg cacgaaggtc ggcgaccggc cggaggtctg gtccatcggc
1081 acggtgatcg gcgcgaagcg ggctccgggt ccgatcggct cgataaagtc caatatcggc

```

1141	cataccgagc	cggcctccgg	cctgttcggg	atgatgaagg	cggttctggc	tctggagaac
1201	aattatctgc	cggcctcgct	gcatttcgag	acgccgaacg	agcagatcga	cttcgacggc
1261	ctcaatgtcc	gtgtgactgc	caatccgatc	gagcttctga	agggcaagcg	ggcgcggctc
1321	gccggcatta	attccttcgg	cttcggcggc	gccaacgcc	atgtcgtcat	cagcgatccg
1381	gatccggcac	aggcggaaaa	ggctgcgacc	tcgcccggcg	gccacgtgtt	tctgggaagc
1441	gccataaccg	cgtcgagcct	cgaaaacctg	ctcgagagac	tataaggccg	cctttgcggc
1501	gcgacgaagg	aagagacgcg	cgcgatcgtc	gccgcctccg	gcgccaatcg	gacgcataatg
1561	cgcatcgggt	tcgcccggcg	cagcgatcat	cccgaagata	tcgtccgggc	gattgccaac
1621	cacctggaaa	agcccgggctc	cgatatcggg	gagacggggc	aagcggccgc	gaaagaagca
1681	aaggctgcct	tcgtcttctc	cggcaacggc	tcgcaatggg	ccggcatggg	tgctgaagcc
1741	ttccgcgaga	atctgcaactt	ccgtcagtcg	ttcacttcgg	tcagcgcctt	tttcagggttc
1801	cattcggata	tcgtcctcac	cgatctcctg	accgaccccc	gactcgcacg	gaagctcgcg
1861	gatacgaagg	tggcccagcc	gctgcttttc	gcgatacagg	ccgcgctctc	cgactcgtc
1921	gtggccatgg	gcatcaagcc	gaccgcgctc	ttcggctcatt	ccgtcggcga	aatcgtcgcg
1981	gcttatgcag	cggggcgctc	gtcgtcgtc	gacgcgtgtt	cgatcgtcgc	caaacgatcg
2041	ctgcaccagg	accttctggc	cggccagggc	acgatggcgg	ccgttatgct	gggcgaggag
2101	gcagccaagg	ccttcgcggc	ggagcgcggg	ctcgacgatg	tctgctcgcg	agcaatcaac
2161	gctcacaatt	cggtgacgat	ctccggcccc	gcgaacgaga	tttcggcggtt	ccgcgacgct
2221	gctgcgaagg	cgaaaatccc	ggttcagatc	ctcgacatca	attatccggt	ccacattccg
2281	atcatcgacc	ggcggaaaaga	ggcgttctc	tcggatatcc	ctgatatcgc	ccccgcggg
2341	acggagctgg	cctatctgtc	gaccgtcacg	ggtgcgcgcg	tcgacgggac	tgctctcgat
2401	cccgactact	ggtggaagaa	cgttcgcgag	cccgtgcggtt	tcaggccgcg	tgccgaggcc
2461	gcgctcgatc	tcggctgcac	gctgttcac	gagatctcgc	cgcgtccgat	ccttggtccc
2521	tacgtcaagg	agacgatcaa	gcaggcggca	gtgcccggctt	ccgtggtggc	gacattgctg
2581	cgcgacgccc	gcgcgaacgg	ccacgatccg	atctccgctg	cgatggcgcg	cgcagtcgcc
2641	cacggtgccg	ccgtcgaccg	gtcgcgtgtc	tatggcaagc	gggacgcatt	catcgaactc
2701	cccgcgctgc	ccttcgagcc	ggtggaactg	cggcccggcg	catgaccgat	gcgacggatc
2761	ctcttcggac	gttccgcaaa	gcctatcgg	ctcaacggct	ggcgcggcga	ccccaatgcg
2821	ggaagctgga	aaaacctatgt	cgacgcgat	ctcttcccgg	atctcgccga	acacgtggtc
2881	gacggcaagg	cgatcctccc	gggcagcggg	ttcatcgaaa	tcgcccgtttc	cgccgcccag
2941	caatatcacg	gcagcgacga	ggtcgagatc	accaatctgg	agatcgtccc	cccgcctcga
3001	ctgagcgaca	gccgcatcat	ggaactctcg	accatcctgt	cgcccggagac	cggcgacatc
3061	gagatccgat	cgcgcgaacg	tctctcggaa	gacgattgga	ccgtgcacgc	cgctcgcgca
3121	agccgcaagc	ccttgccgac	gacgagcaac	gactgtccgt	ccaccaaagg	cttgccgaag
3181	acggcgaccg	tgacggcggc	caaggcctat	gaaacggcga	agcagttcgg	gctggactac
3241	ggcccgcaat	tcagcttct	ctccaaagcc	gtggcctatg	gcgaccggct	tatcgacgct
3301	gagttgaagc	acccccaggc	gtcggggcgc	ccccacgtca	cctacgcctt	tcattccgatc
3361	tcggtcgacg	cgactctcca	cggcctggtc	gcactcttcg	accgcttcac	cggcgaccag
3421	ggcggcgctt	atattcctgt	ccgtttcggc	tcctgctcgc	tcgtcaatgc	cggcccggccg
3481	atcggccgcg	cgacgatcga	gatcgcgagg	gtcagcgcca	actcgatcaa	ggcacgggtt
3541	catttctctc	acgagaccgg	cgtcgcgatc	gcacacttcg	aggattgccc	cttccgcctg
3601	acctatctgc	gcaagcacia	gacgcttgaa	ggcctttcct	tcactatgca	ggcagttctg
3661	tcgcagctgg	tcatgcccgg	cacgaaagtt	gcgacggcca	tgctcgcgac	cctggccgag
3721	ccgatcgatg	acaatggcgt	cgacaacgcg	accctgctct	tcaacgcgcg	catctaccgc
3781	gctgcccagg	aaattgccct	gaagctcggc	agggggaccg	ccacagtgcg	aggcgaccgcg
3841	ctgcccgggg	acttcgcggt	ccagtgtctc	ctgacgagct	gctgctgac	gctcggagat
3901	cccggccttt	gcgagcacca	gaacggcaac	tggaaggtcg	cacgggagtt	ctcgtttccg
3961	accgttcgg	agatccttgg	cgagctctat	ggcgcaccgc	ccgatcgtgc	ggtggaagcg
4021	gtgctcgta	acaacgccta	tgccggaagca	ctcagccgctc	tcgacgcgct	gcttggggccg
4081	gtgcgtgcag	gcgacgacgc	ggcttcggtc	aatgcccgggtt	tcacagcga	agcgcagctc
4141	gaccatcagg	ccgtgcattc	cgtggcgagc	ccgtcgcgca	tgaccaggtt	gcttcatacg
4201	gtcgaacacg	ttctcgtcgc	gcagccggca	ggcgcctgcg	gctcatcgaa	ctcggcagcg
4261	ttccgggctt	cagccgcggg	ctggcgcgatc	tggcagcgcg	caatgggtgca	gcactttcca
4321	ttttcgagcc	gcgcgacaat	gcacaacgga	atctggaaat	cgcttcgag	gaggatgaat
4381	gtgcgcgttc	tgaagaagag	cgaattcgcc	gatatcgggc	cgttcgacct	cgccgtcagt
4441	gcttcggaca	atctctacca	gctgatcgaa	gaggagagcg	gcgtccgcac	ggcccttcgc
4501	ttccatgccc	ccggaagcgg	gctcgtcgcg	gccgtcagcg	cgccgtcgat	cttcgcccag
4561	ttcccccctg	gtctctcggg	cggatgggtt	gagcgcagcc	agacggcaga	attcccgatc
4621	ggcaggatcg	cggccgtacc	ccattggcag	aagtgtctcg	cagatctcga	cctcggcgac
4681	gtggctcgtg	cggatcgcga	atgcccgcac	gggaaacacca	tctgatcga	ggcccggggc
4741	acagcggcgg	caggcgcggg	cgccgagcgc	ggcgaaaccg	gaacaccata	tctgctgatc
4801	gagactgcca	ctgcccagaa	accggtcgcg	gccggtggaa	ggagtgcgcg	tccgctcgt
4861	gtgacaggcg	accttgccgc	agatatggcg	acgctccgcg	cagccctcga	agccatgggg
4921	gacagaccgt	tgccgcgggt	tttcatgtcc	gaggcggagc	gcgcccgcg	cgacgggttcg
4981	gcctgctgc	agacgcgtgt	atggcgctc	agcgcctttg	ccgaggcatt	gaaacagcat
5041	ctcgcggcgg	tcgagccgag	ggcgacgacc	ggcgcgcctt	gtgctcgtgg	ccccggggcg
5101	cgctcccctta	tcacctcggc	aggctccggg	ctgaattccg	gattgtgggg	cttcgcccgc
5161	gtccttcaga	acgagtatga	gtttatcgac	gttcacacgc	tcgatcttgc	cggtgatgcc
5221	gacatcggcc	gcaaggatat	gcttgccgca	gcgcttcccc	tgctgacggg	ttcgggacag

5281 aaccgggaat ggctgctcga ccggaagacg ggccctgctgt cggagcttcg ggccgtaccc  
5341 ggcgcaacga acaaggttga cgggaaaacc aacgccttcg cggccgcgac cattcgcag  
5401 cgcgtcagct cgcaggttgc aagcatcgcc tgggaagagg cgcagctacc ggagccgggc  
5461 cccaacgagg tcgtcgttgc cgtcgcggca accggcttga atttccgcga cgtgatgtgg  
5521 gccatgggcc tgttgccgga agaggcgttg gaagacggtt cgcgggggca accatcggca  
5581 tggaaactgtc cggccacgtg gtcgcggctc gcgaagagat cgacgatctc gcggtcggcg  
5641 atgcggtgat ggccatcgcc tccgcgcct tctcgacca tgcggtcgtc gcgcgcgcc  
5701 catcgccaag tgcgcgagc gttgagcccg tggccgcggc cacattgccc gtcgccttc  
5761 ttaccgccta ctacgccatt gtcgagctcg gccgcattca ggctggagag acgattttga  
5821 tccacggtgg ctccggcggc gtcggccttg ccgcgctcca ggtcgcaaag ctcaagggcg  
5881 ccaaggtcat cgcacggcg ggaagcgccg aaaagcggcg cttcctgacg atgctcggcg  
5941 ccgaccatgt cttcgattcg cgttcgctcg gcttcgctga ggatgtccgt tcggtcaccg  
6001 gcggggaagg cgtcgacctc gtgctcaatt cgtttttcgc cgaagcgatg gacgagacc  
6061 tctcgttggg gaaacccttt ggccggttcc tggaaactcg caagcgcgac tattacgcc  
6121 acagcaagat cgggctcagg ccgttcgccc gcaatgtcag ctatttcggc atcgacgcc  
6181 atcagttgct ggtgaatgcg ccggatctga ccaggcgcat cttcatgcag atcggagcgc  
6241 tcttcgagga aggcaacgct acgccgcttc cctatcgcc ctctcgattac ggatgaaatc  
6301 gggcaacttg tccggctgat gcagaatgcc gggcatatcg gcaagatcgt ggtgctgcc  
6361 cccgttgccg gcaggcacga ggtgacggcg aaggctgtga ggggatgaa ggtcgaccg  
6421 gacggtcgtc atctcgtcgt cggcggcctc gcgggcttcg gctcgtcgc gcccaactgg  
6481 ctcgctcgaga aaggcgcacg ccgtattgct ctttgcctcc gccgaggcca gccggatgct  
6541 gaaacgcgcg cgtgatcga gctcggggca gaatggcggc gttgcggctt cgttccatgc  
6601 ttgcgacatc acggatgccc cggccgtgga gacgcttctg gcgacgcttc gttcgggaagc  
6661 accgttcgct ccgtcgttca ccggcgatg gcggcgatg gtgctcgac atgctgctg  
6721 agccgcgaac gcaaccggcc ggtcatcgaa acgaaagcca agggcggcggc aattctcgat  
6781 cgccttacc gtggtgacag gctcgacaac ttcacctgt tctcctcggc gacgacgctg  
6841 gtcggcaatc ccggccaggc gaactatgct gccgcaaacg gctatctcga aggccttgc  
6901 cgtgcacggc gccaggaggg cctggcaggg ctcgccatcg gcttcggcgc gattgcccga  
6961 gccggctacc tgacgcagaa ccggcgatgc cccgcaagcg tcgccaagcg catcggcaaa  
7021 acggcgctga aggccaggt ggcgctcgac atggtggaaa accacgctcg cgcggatccg  
7081 ggcacggctc acgcggccgt cgtgatgatt tcggagatcg actggacggc tgcgcggaac  
7141 ctgcccgtcg ccgcaacgcg ctcttcgagg tgatcctgcg cagcgcgat cagcatgccc  
7201 caggcgcgca gggcacgacg atggatctcg tcgccatgat cgaaggcaaa tcgcccagc  
7261 aggccgagga catcctcttc gatctggtgg ccggcgagat cgtcgccatc cttcgtgct  
7321 cgaaggatac ggtcactcgc gaaagatat tgaaggagat cggcctcgac agcctcatgg  
7381 ccgttgaact cggcatgagt tccagcaga acaccggctt cgacatgcct ttgagcggcg  
7441 tcgcggaaca taccacggct ggcgacatcg cgcgcaagct ccatgaaaa gtaagcaagc  
7501 cgtgaccagg gcaacgagaa cgaggccggt gcgacaagc ttgtaacaga actcaagcag  
7561 cgtcatgtcg gagcggacaa ggaaaaagca ctgagtcaat gagcaacgat ggaacggcc  
7621 agaccttccc gaaaatgaac agcagtctca aggagagcct gttggacagg atgcggaaca  
7681 cgcacatcgc gtcggagcgc aaccggatgg ccgctcgga gcgggagttg ccgcccgtc  
7741 cgcgtcgcca gcaggcgcgt ttcgaggatc tgcccagata taagcaggct cttacgcgca  
7801 agatcgccag cgagcagctc ggcacogcca acccgttcta ccggccgat cagacggcag  
7861 ccggtgcgac gacgatgac gatggcgcga agctgatcaa cttcgcctcc tatgactatc  
7921 tcgggctcaa ccggcatgct cacgtcctcg agcgggcgcg cgagacgatc gccggcttcg  
7981 ggatttccgc gtcggcgagc cgcctcgttg ccggcgaaac gccgcagcat gttggagctc  
8041 aagaaaagat cgcgcaatc tacggcgttg atgcccgcgt ctgtttcgta agcggctacc  
8101 tgacgaaatg cgcgcgacat agctgcctga tggggccgaa ggacctcgtc atccacgacg  
8161 agttcatcca caacagcgc ctcgcgggga tcaagcttcc aggtgcgaca gcgcccttct  
8221 tcaagcacia cgacaccgcg gatctcgagc atgtcctgcg caccgtcgcc ggtgactacc  
8281 ggcgcatcct ggtgatcgtc gagggcacat actcgatgga cggcgacgct gccaatctgc  
8341 cggcgttctt gaagctcagg gccgaatacg gcttctggct gatggtcgac gaggcgcatt  
8401 cgtcgggggt cctcggccgg caccggcagg gcctcgcgca acatttcggg gccgatccgc  
8461 acgaagtcca catctggatg ggaacattgt cgaagacgac gtcgagctgc ggcggctaca  
8521 tagcaggaag cgcgcctcgt gcggccgtgc taaaggcatc tgcggcggc ttcgctctata  
8581 gcgtcggcct cgcgcccgtg cttgcccgat ccgcccgtcg cagcctcgac atcctggcaa  
8641 gcgaaccgga acgcacggcg gcggtgagac gcaacggcag cctgttctcg aagcttgcca  
8701 aggaggccgg cctcgatacc gggctgagcg gcggcttttc ggtcgtgctt gtcacgtcg  
8761 gggattcgct gcgcgcctc cagcttctga acgacctgct gcgcccggc aacgtcctgc  
8821 cgtcatcca cccggcgggt ccggaaggac aggcgcgct tcgcgttctt catcaactgc  
8881 gaccacaccg aggaacagat ccgcccaccc gtcacattgg agctgaacgc ctgaaggatc  
8941 tcaccgacc caatttcggc ctcgggcgcg ctcgggcgcg tcgatatcga gaagatgct  
9001 cggcatgctg accaccagcc gccgcacggc gcgcaacggg ttcgtccatg acacatgtga  
9061 tcatcaccgg cggctcgagc ggaattggcc tcgcggtcgc ctcgatctac gcggcccgca  
9121 gccgctctc tcttgctcga cgtcggcccg atcttctgga gaaggcagca cagaagctgg  
9181 tagccgaaca cgttctcgag gccggagcgg tccgcataga ggcgcggat gtttccaaag  
9241 gagaggagat cagggcgcg gcttctcgt gtgtggatgc gtcggacc tgcgagctt  
9301 tcgtgacctc ggcaggcgtc gtcgaaccg ccgcatctga ggcgatgcag ggatcggcct  
9361 tccaccggca gatggaaac aacttttcgg gcacgggtgca tgcctgccc

9421 cggacatgaa gaacaggcgg cggcatattc tgatggtttc ttcgggtgcc ggcctgatcg  
 9481 gcatctacgg atataccgcc tactgcccct cgaagtccgc gcttaacggc ttcgccaggc  
 9541 gctgcccagc gaacgcacgc ccacaatgtc ggcatttcga tctgctttca cccgatacgg  
 9601 agacgcccga attcaagagg gagcttgccc cccggccagt ggaggcaagc gtaatcatgg  
 9661 gaaccgtccg cccatggacg gccgaggcgg tcgcccgaag gatcgtcaag ggcacgcacc  
 9721 ggcggcgctt cgagatttat tttggcgcgg ccttgatct tctgcccgtt cgggcccggc  
 9781 gtcaggcctt tcctcaattg gtggttcgac cgggcgatcg cgcgcaacag tggactgtga  
 9841 cggcggcgcc cgcggccatg cttgtgcccag ccaatattag gctttaagtg tcggggagaag  
 9901 ccttctgccc aggcggcgac ggtgaaatgg ctcaaaggta tcggcttggc gtcattccc  
 9961 tccaacctgc atgactacc gatagccttg accctcgggt gctacctgct gtccctgcgc  
 10021 gtgatcttct atacggaccg gttcgccctt cccggcgcgg aacgcaggaa gaacgcaccg  
 10081 aaatacgggc gcaatcggga caggatcgac catctcgccc ggctgcccgt catcgcgctc  
 10141 gtctttgccc gcttcttcgc gatctcgtgg cggcccctct atgcccgcgc cggcacgatg  
 10201 agtttcttca tcattttcac cggcatttcc cgcgcgaagt tcaagttcat ccgtgagccg  
 10261 ctggctctttt ccgatatcgc cctggtcgcg gatgtcttca agtacaagac gatcttctat  
 10321 gcgagttcgc tgaacatcgt cttctggatc gtcgccttcc tttacgtggt cggcgtatcg  
 10381 gcaactttaca tgtatttcga gccgacgatc ctgcccgaag gaagcaggct ctctctgggt  
 10441 ctgggtgatgg tggggatcgc cgcgggacc tggggactgt tgttttacgg gccctggaac  
 10501 gcggcgacgg ccgctctcgt gcagaggctg gtgaaggcga tcaacgtcaa gatgaacacg  
 10561 gtcgggttccg gtaccttcgc ttcgctcgtc ttccatttca tcatctgggt cgggtgcaag  
 10621 cgcgacaaga tcgctgcccga attgtcggaa atactgcccg ccgcccgtgca cgacctcatc  
 10681 ggcacagagg agcggcccgt gatcatcgta tggcagtcgg agtcccttcat cgacatgcgg  
 10741 catttcgggtg tcgattcgat caagcttccg acggctgacc ggttgccgaa gcaggcgggtg  
 10801 caatggggggc gcttgagcaa tgtcttcgag gccggttaca cgctgcccac cgaatttgcc  
 10861 gtcctgagcg gccttggtgc cgacgacatc cacgtcgatg ccagctatcc ttatctcgt  
 10921 gccgcccact atgcccgat atgctggccc gaaagctca agcggcgtg cgcacgcgat  
 10981 ttcatctatc cctatgaccg gacgttcttc ctgcccaca aggcgatgcc cctgctcggc  
 11041 ttcgagaagc ttaccatgct cgatgccttc gaccataaac cggagcggga cgggctttat  
 11101 gtctccgacg ccacgctggc agcccgcgtg ctgcccgaag tcgagaagct cccggaagag  
 11161 gagagcggct ttttcttcgt cgcctcgatg gcaaatcacg gccctggga accggggcgc  
 11221 gtcggtacgc tcaccaatcc cgtcgatata tacctggcca tctcagagca gtcgacgct  
 11281 gcgttgaagc agctgatcga cggcctcaac aagctcgacc ggcgggtctg gttcgtcttt  
 11341 tacggcgacc acgcggccct tctgaaatct ttcgggatc ccttcccgga tcccgcacg  
 11401 gattatttca tcgtgcccgt cgcgaaggcg cgcgcccggg ccacagccc gaagcggcgg  
 11461 aaggacgagg atccttggaa cctgctcggc tccatgctga agcacgcaa tctgcacaag  
 11521 gacgcgctgc aatagcgtg acttcccggaa atcccggca tgacagcagt aacgaccata  
 11581 gtaccgcgcg cgcgaagaac cttcctgttc ttgcagggac cggcgtcgcg cgtcttcgcc  
 11641 aaggtgcgcg ggcacctcga agtgcctggc cattgctgtc tgcgcatcaa cctcaatccc  
 11701 ggcgaccaga tattctggcg gcgcccggcg gccataacta ccgtggactc gttggcgcg  
 11761 tggccggcct atgtggaggc ctttatccgc cgcacgccc tcaccgacct cgtcctgctc  
 11821 ggcgaagagc gccctatca ccggatcggc acagcggccc cgcgaaaagc cggcgtccgc  
 11881 gtctttgtgg tcgaaatggg ctatctcagg cgggactggt tgacgctcga gcgcccggc  
 11941 atgctgtcca actcgcactt tccggtcgat cccgcgcaa tctcgcgagc gccaatggga  
 12001 ctgcccgagc ccgactggca acggtcttac cggcagagct ttcttgcgga ggcggcatac  
 12061 gacctcctct acaacctgcc gaacgtggtt ttctgcttcc tcttctgcta cccgcccac  
 12121 ggcaccttct atccgctcgc cgaatatgcc ggtggtatcc ggtaccggcg gctggcggcc  
 12181 ggcaaacgcg agcagcggcg cgcggatgcg ttgatccgct cgctttcggc gaaaaatgcg  
 12241 cegtacttct tctatccgct tcaacttgaa acggactatc aattgctgtc gattcacc  
 12301 ttcggcagcc agcgggaagc gatcgacggc atccttgctt ctttgcgca ccgtgccc  
 12361 gcaggaacca ggctcggcgt caaggtgcac ccgctcgaca acagcctgat ctctggcgc  
 12421 aggatcgtcg cggacaaggc cgcctcactc ggtgtcggcg acagggctct ctatctcgac  
 12481 ggcggcaatc tcgattcgtc cacggagcac agttgcccga ttgtaacggt caactcgacc  
 12541 gccggactgc atgctctgaa gcagggcaag cccgtgaagg tctgggaaa ggcgctctt  
 12601 gacattgccg gtctcaccga ccagcagccc ctcgacgctg tctggacggc gcccgaggcc  
 12661 ccggatccgg tgctgagcgg gccaatgttc aggtgatgg ccgcttcgat ccaggtgccc  
 12721 gggaaacttct attcctgctg ggtacggatg cggcccgcaa ggccatcgcc gcaggctgca  
 12781 ccgcaacgac gtcaatgaac cggggccctt cgtcagccc cgcggcggc caaagcggc  
 12841 tgaactcag cgaagatcag gcgacgccc cgcgacgag atcgtgaatg tgcaggatgc  
 12901 cgaccggggc gccggcctcg tcgacgagga acagcaggtt ggacctgtg gtcctgcatg  
 12961 aattc

//

## 7.2 Genbankeintrag rkpU

LOCUS AF247710 3547 bp DNA linear BCT 08-AUG-2000  
 DEFINITION Sinorhizobium meliloti capsule polysaccharide exporter protein RkpU

(rkpU), putative pyrazinamidase/nicotinamidase PncA (pncA), and putative nicotinate phosphoribosyltransferase PncB (pncB) genes, complete cds.

ACCESSION AF247710  
VERSION AF247710.1 GI:8571419  
KEYWORDS .  
SOURCE Sinorhizobium meliloti (Rhizobium meliloti)  
ORGANISM Sinorhizobium meliloti  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

REFERENCE 1 (bases 1 to 3547)  
AUTHORS Epple,G. and Geiger,O.  
TITLE New genes in the chromosomal rkpI-region involved in the biosynthesis of capsular polysaccharide  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3547)  
AUTHORS Epple,G. and Geiger,O.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAR-2000) Biotechnology, TU Berlin, Seestrasse 13, Berlin 13353, Germany

FEATURES Location/Qualifiers  
source 1..3547  
/organism="Sinorhizobium meliloti"  
/mol\_type="genomic DNA"  
/strain="41"  
/db\_xref="taxon:382"  
gene 136..1488  
/gene="rkpU"  
CDS 136..1488  
/gene="rkpU"  
/note="putative pore forming lipoprotein; belongs to the BexD/CtrA/VexA family (PROSITE PS00013)"  
/codon\_start=1  
/transl\_table=11  
/product="capsule polysaccharide exporter protein RkpU"  
/protein\_id="AAF76876.1"  
/db\_xref="GI:8571420"

/translation="MAGRAQSRLFCRLSRGGLLWDGLCAFQRPADSLPGKCGADICM  
ESLLCRTRSVVVKRV/TALLLCTALAGCQAVPGEGLPTTDIVSDAGQSGSEIGRRNATV  
FDIVDVGQSARLVSEYVSTTLSRRFGIGGVVGVVIGIGDQLKVTIFEAGSDGLFST  
TESKQTSIDLVVQPDGKAAIPYVGSVRFAGLTLEQARQAILDALKQKAVEPDVIVTST  
STASRIVTVSGAVGRPSVVPLNLVSETINEVIAKAGGPSAQPYETYVTLVRGKKTGTV  
LLKSIIESPSENIHVKPGDQIFVTRDPRTFTVLGQVRANQRVEFGANDLNLLEAVALA  
GGGSDRTVDAKGYFVFRYEEDIVMSLLGQERFHKLLSKGMKADRVGRYPVVYRFDMT  
KPDSLIVGQTFPVKNRDVIYASRHPSVDITKFLDFVARPIGIVNSGVNVADNLNDFNN  
"  
misc\_feature 1673..1771  
/note="similar to Sinorhizobium meliloti sequences downstream of the cycL gene, between the chvI and chvG genes (partial repeat element RIME2), downstream of the phoT gene, and between the fliG and fliN genes"  
gene 1798..2433  
/gene="pncA"  
CDS 1798..2433  
/gene="pncA"  
/codon\_start=1  
/transl\_table=11  
/product="putative pyrazinamidase/nicotinamidase PncA"  
/protein\_id="AAF76877.1"  
/db\_xref="GI:8571421"

```

/translation="MADAARPDLEAMADEALIVIDMQNDFCPGGALAVEGGDEIVPA
VNRLIDASPHVVLTDWHPAGHSSFASTHPGKAPFQTVAMPYGEQTLWPEHCVQGSAG
ADFHPSLRWTSAELVIRKGFRRREIDSYSAFFENDHRTPTGLAGYLRERGI RSVTLCGL
gene      ATDFCVAFSALDAVAKGFSTSVVLGACRGIDLNGSLAAMTLRMRDAGVRLI "
          2549..3178
          /gene="pncB"
CDS       2549..3178
          /gene="pncB"
          /codon_start=1
          /transl_table=11
          /product="putative nicotinate phosphoribosyltransferase
          PncB"
          /protein_id="AAF76878.1"
          /db_xref="GI:8571422"

```

```

/translation="MPKTDIARRVYNHTWKLDPIVRSLLDTDYKLLMLQMIWQLYPD
VDATFSLINRKTVRLADEIDEQELRDQLDHARGLRFTTKKEMIWLAGNSFYGR TQIF S
PEFLAWLARFRLPEYELSRRDGFELTFRGRWAETTXWEIPALAIINELRSRAAIKGL
          GPF'LDVLYARAKAKMWSKVEQLREHPNLRISHFG'TRRRHSFLWQRLCV"

```

```

ORIGIN
1   ttttatcccc   tgcgctgtag   tagtgagcgt   ttaactcata   gtcgctagct   ttcaaccgcc
61  aagggctaga   tatacgtctc   ctgcatccca   tccaatacaa   ttacggcgac   ggttcgtaac
121 caatatcgag   gcggtatggc   cgaagagcgc   cagtctcgat   tattttgccg   gctgtcgcgg
181 ggaggtggct   tgttgtggga   cggcctttgt   gcttttcaga   gaccggcgga   ttccctgccc
241 gaaaaatgcg   ggcagagata   ctgtatggaa   tcattgctgt   gccgaacacg   gagtgctgta
301 gtgaaacgag   tgactgccct   tctgctttgt   accgcgcttg   cgggatgccg   gcccgcgccc
361 ggcgagggcc   ctctgacaac   cgatatcggt   tcggatgccg   gccagtcgag   ctccgagatc
421 ggcagaagga   acgcgacggt   attgatatt   gtatagtgct   atggccagtc   ggcgcggtc
481 gtttccgaat   atgtttccac   aaccctcagc   cgcgcgcttc   gcatcggcgg   cgtcgtcggg
541 caggtggtca   tcggcattgg   cgaccagctg   aaggtcacga   tcttcgaggg   gggcagcgac
601 ggcctgtttt   ccacaactga   atcgaagcag   acgagcatcg   atctcgtcgt   ccagcccgat
661 ggcaaggcgg   caattcctta   tgtcggatcg   gtcggttttg   ccggcctgac   gctcagagcag
721 gcgcgcaggg   cgatcctcga   tgcgttgaag   tcgaaggccg   tcgagccgga   cgtgatcgtg
781 acgtcgcagc   cgacggcgtc   gcgcacgttc   acggtgctcg   gtgcccgtcg   cgtcctctct
841 gtcgtgcctc   tcaatctcgt   cagcgagacg   atcaacgagg   tcatcgcgaa   ggcggggcgg
901 ccacggcgcg   agccctacga   gacctatgtg   acgctggtgc   gcggcaagaa   gaccggcacg
961 gtcctgctga   agtcgatcat   cgagagcccc   tcggaaaaca   ttcacgtaaa   gccggggcag
1021 cagatattcg   tgacgcgcga   tccgcgcacc   ttcacgggtg   ttggccaggt   ttgggcaaat
1081 cagcgcgctc   agttcggcgc   caacgatctc   aacctgctcg   aggccgtggc   gctcgcgggc
1141 ggcggctcgg   atcggacggt   ggatgccaaa   ggttacttcc   tgttccgcta   cgaggagccc
1201 gatatcgta   tgagcctgct   cgggcaggag   cgttttcaca   agctgctgag   caagggcatg
1261 aagcccgacc   gcgtcggggc   ctaccgggtc   gtctaaccgt   tcgacatgac   gaaaccggac
1321 agcctcatcg   tcggtcagac   cttcccggtc   aagaatcgcg   acgtcatcta   cgcctcgcgc
1381 catccctctg   tcgacatcac   gaaattcctc   gatttcgtcg   cacggccgat   cggtatcgtg
1441 aatccggcgg   tgaacgtcgc   cgacaatctc   aacgatttca   acaactgacg   ttggagtggg
1501 ccgacaccag   gcgtggggcg   tgtcgcgcgc   gttgccggcg   gaatgctttc   ggcgggcatg
1561 cattaacgctc   ctatagtcaa   tcacctttat   aaggttcgca   ggttcgcggg   gattcgcgct
1621 gcgagcgtcg   aacatgcttg   ccgcgcgcgc   cgttctatgg   ttgatctggt   tttagagcgg
1681 atgaggaaaa   gtgtatgcgg   ttttcgccc   gcatcccgtc   ctaacttatt   agaatcgatc
1741 acgttcatga   cttaggtcga   tccgacctga   atcatcgtga   tcgggatcga   tttgttgatg
1801 gccgatgcgg   ctccggccgga   tttgcgagag   gcaatggctg   acgaagccct   gatcgttatc
1861 gatatgcaga   acgacttctg   ccccggcggc   gctctcgcgg   tcgagggcgg   cgatgaaatc
1921 gtaccgcgcg   tcaatgcctc   catcgacgcg   tcgcccgatg   tcgtcctgac   gcaggactgg
1981 caccggcgcc   gccattccag   tttcgctcgc   acgcatccgg   gcaaggctcc   gtttcagacc
2041 gtggcgatgc   cttacggaga   gcagacgctc   tggcccagac   attgctgcca   ggggagcggc
2101 ggtgcccatt   ttcattccgtc   cctgcgatgg   acctcggcgg   agctcgtcat   acgcaagggc
2161 ttcggtcgcg   agatcgacag   ttattccgcc   tttttcgaaa   acgaccaccg   gaccgccacc
2221 gggcttgccg   gttacctgcg   cgagcgcggc   ataagaagcg   tgacgctttg   cggcctggct
2281 accgatttct   gcgttgctct   ctccgcactc   gatgcggctc   ccaagggttt   ttcgacgtcg
2341 gtggtgctcg   gcgcctgccc   tggcatcgat   ctgaaacggc   cgcttgccgg   catgaccctg
2401 cgcgatgcgc   atgccggcgt   ccggctgatc   tagtcgcccc   cgggacgggg   aagcggggcg
2461 tcggcgcgat   cgccgctcgg   acttaggata   tcgaccacga   tcgctcgcgg   tcagagcggg
2521 tcatgacggt   ccaagaggag   cgtattgcat   gccaaagacc   gatatcgcgc   gacgggtcta
2581 taaccatacc   tggaaactcg   atccgatcgt   ccgcagctcg   ctcgatagcg   atttctacaa

```

```

2641 gcttttgatg ctgcagatga tctggcagct ctatccggac gtggacgca ccttttcgct
2701 gatcaaccgc acgaagacgg tgcgccttgc cgacagatc gacgaacagg aactgvcgca
2761 ccagctcgac cacgcgcgag gcctgcgctt caccaagaaa gaaatgatat ggcttgcggg
2821 taacagcttc tatggccgca cgcagatcct ttcgcgcaa tttctcgct ggctcgcaag
2881 attccgcctt ccggaatacg agctgtcgcg tctgtgacggc caattcgaac tgacctttcg
2941 tggccgctgg gccgagacga ccatntggga aatccctgca ctcgcatca tcaacgagtt
3001 gcgctcgcgc gcggcgataa aagggtcgcg cccattcacg ctggacgtcc tctatgccg
3061 cgccaaggcc aagatgtggt cgaaggctga gcagttgctg gagcatccga acctgvcgat
3121 ttcccacttc ggtacgcgcc ggcggcacag cttcctatgg cagcgttgt gcgtctaggg
3181 gcttaaggaa ggcatcggct cgtcctttac aggctcgagc aacgtgctgc tggcaatgga
3241 caatgatctg gaggcggtcg gcaccaatgc gcacgagctg ccgatggtgg ctgvcgvcgt
3301 tgcccgaac gaccgcgagc ttgcggccgc gcctacaag gtctcgcagg actggaacca
3361 gctctatggc ggcaacctcc tgatcgtgct cccggacgcg ttcgggacgg ccgvcgttct
3421 cagggatgca cccgactggg tggcggactg gaccgcttc cggccggaca gtgccccgcc
3481 catcgaggga ggcgagaaga tcatcgcctg gtggaagaag atgggcccag acccgcggga
3541 gaagctt

```

//

### 7.3 Genbankeintrag kpsF

```

LOCUS       AF247711                1325 bp    DNA    linear    BCT 18-JUN-
2000
DEFINITION  Sinorhizobium meliloti capsule expression protein (kpsF) gene,
complete cds.
ACCESSION  AF247711
VERSION    AF247711.1  GI:8571423
KEYWORDS   .
SOURCE     Sinorhizobium meliloti (Rhizobium meliloti)
ORGANISM   Sinorhizobium meliloti
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
REFERENCE  1 (bases 1 to 1325)
AUTHORS   Epple,G. and Geiger,O.
TITLE     New genes in the chromosomal rkpI-region involved in the
biosynthesis of capsular polysaccharide
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 1325)
AUTHORS   Epple,G. and Geiger,O.
TITLE     Direct Submission
JOURNAL   Submitted (22-MAR-2000) Biotechnology, TU Berlin, Seestrasse 13,
Berlin 13353, Germany
FEATURES   Location/Qualifiers
            source          1..1325
                        /organism="Sinorhizobium meliloti"
                        /mol_type="genomic DNA"
                        /strain="41"
                        /db_xref="taxon:382"
            gene           242..1321
                        /gene="kpsF"
            CDS            242..1321
                        /gene="kpsF"
                        /codon_start=1
                        /transl_table=11
                        /product="capsule expression protein"
                        /protein_id="AAF76879.1"
                        /db_xref="GI:8571424"

/translation="MMGVRHATADAHAGGTVLESIGRTLTTATNGIKALADHLTSDQD
FAGALVDAVELMGDGDGRVVVSGVGKSGHIGRKIAATLASTGTSAYFVHPTEASHGDL
GMITAQDALVLLSWSGETAELANMLTYAKRFKVPVIVSICSNRETLARNSEIALVLPK
VPEACP HGLAPTT SAMLQLAVGDALAIALLERRGFS AEDFKTFHPGKLG AQLRLVQE
LAHGTGQMP LLSVGRPMSEAVIEMSAKGFVVGITDESGLIGVITDGD LRRHMAGDL
LAQVPVQEVMSRNP RVIKGDVLSAAMEFMQDHKVHRAVPRRRRRPGRHPAHSRSAAR

```

```
RRRLIFAEFHRLLAGAAGRRRAPVH"
ORIGIN
  1 aattcgccgc cggcgggata cttcctccgg agacctcggg cgaaaagcgc gtcgggtgag
 61 catgacgggt cgaaacgtcg ccggtttggt ataaatggcc gtcgaaagcc ttggtttgtc
121 catgaaaaga gcctatctgg cgtctatctt gctgcggccc gctgacgagg cgttttttgg
181 ctacgcactt ccggcggaaa accgtgacgc agttttcccg gaggtgccct aaagggggtc
241 gatgatgggt gtgagacacg cgacggcgga tgcgcatgcc ggtggcaccg tgctggagtc
301 gatcggcagg acgctgacga cggcgacgaa cggcatcaag gcccttgacg accatctgac
361 cagtgaccag gacttcgctg gcgccctggt cgatgagggt gaactgatgg gcgacggcga
421 cggccgcgtc gtcgtttccg gggtcggcaa gagcgggcat atcggccgca agattgcagc
481 caccctcgca tcgaccggca cctcggccta tttcgtccat ccgaccgagg caagccacgg
541 cgacctcggc atgatcaccg ctcaggatgc gctggtgctg ctctcctggt cgggcgaaac
601 ggcggaactc gccaacatgc tgacctacgc caagcgtttc aagggtgccg tcgtctcgat
661 ctgctcgaac cgcgagagca cgctcgcacg gaattccgag atcgcgctcg tgctgccgaa
721 ggtgccggag gcctgtccgc acggcctggc gccgacaacc tcggccatgc tgcagcttgc
781 cgtcggcgat gctctggcga tcgcgctcct ggagcggcgc ggcttctcgg ccgaggattt
841 caagaccttc catccgggcg gcaagctggg cgcgcagctg cggctgggtc aggagctggc
901 gcatggcacc ggacagatgc cgttactctc cgtcggccgt ccgatgagcg aggcgggtcat
961 cgaaatgtcg gccaaaggct tcggcgtcgt cggcataaca gatgaaagcg gcaagctgat
1021 cggcgtcatc accgacggcg acctgcgcgc tcacatggcg ggcgacctgt tggcacaacc
1081 ggtccaggag gtgatgtcgc gcaaccgcgc ggtgatcaag ggtgacgtcc ttgccagcgc
1141 cgccatggaa ttcatgcagg accacaaggt ccaccgtgct gttcctcgtc gacgaggccg
1201 gcgccccggt cggcatcctg cacattcacg atctgctgcg cgccggcgtc gcctgatctt
1261 cgctgagttt caccggcttt tggccggggc ggcgggtcga cgaaagggcc cggttcattg
1321 acgtc
//
```

## 7.4 Sequenzen und Restriktionschnittstellen von pPP428

Länge von pPP428 : 29875 bp

```

GAATTCGTAG GCCCCGATGG CGCCGACCGG AATGTCGATG CTGACCTCCG GCCACCGCGC CTCCGGCCCGG
 10          20          30          40          50          60          70
ACCCAGCCGC CGTCGGTGCA TGTGACCTCG AAGTCGTAGC CGTGCATGTG GATCGGGTGG TTGGTCATGG
 80          90         100         110         120         130         140
TGAGGTTGCC GACCCGTACG CGCACGCGGT CGTTCTTGGA AACGACGAGC GGGCTGATGT CCGGGAATAT
150         160         170         180         190         200         210
GCGGCTGTTC CAGCACCACA TATTGAAGTT GGTCATTTCC ATGACGCGCG GCACGTAGGA GCCCGGCTCG
220         230         240         250         260         270         280
ATGTCTGTAG GCGTTGAGCA GGAACACGAA ATCCCGGTCC ACCGGCATGA ACTTCGGGTC CTTGGGATGG
290         300         310         320         330         340         350
ATCACGAAGA AGCCCATCAT CCCCATCGCC ATCTGCACCA TTTTCGTCGGA ATGCGGGTGG TACATGAAGG
360         370         380         390         400         410         420
TGCCGGATTT CACGAGATCG AACTCGTAGA CGAAGTCTT TCCGGGCGGG ATGTGCGGCT GCGACAGCCC
430         440         450         460         470         480         490
GCCGACCCCA TCCATGCCGG ACGGCAGGAT CATGCCGTGC CAGTGGATCG TCGTCGGCTC CGGCAGCTTG
500         510         520         530         540         550         560
TTGGTGACGA AGATGCGGAC CCGATCGCCC TCGACCGCTT CGATCGTCGG ACCCGGCGAC TGGCCGTAT
570         580         590         600         610         620         630
AGCCCCAGAG ATAGGCGGTC ATGCCCTCGG CCATCTCGCG CTCGACCGGC TCGGCGACGA GATGGAATC
640         650         660         670         680         690         700
CTTGACACCG CTGTTTCATCC GGTGAGGCGC GGTCCAGCCG TTGACGGTCA CCACCGGTT GTAGTCGGGG
710         720         730         740         750         760         770
CCGGAGGCCG GACGGACGGG AGCTTGCCTT TCCGCCGTCT CCATGATTGC CGCTTCCGGC AGGCCATAT
780         790         800         810         820         830         840
TCGACGTTTT CGCCAGGCG GCGGTGGATA CAGTGCGGC GCTCGCGCCG AGTAACTGTC GTCTGTTGAA
850         860         870         880         890         900         910
CATGTTCCGT TCCTTTCTCA GTTGCCGCTT TCGGCGGCCG CCGCGACCTC GGTTCGCGG GAGGCTCCCC
920         930         940         950         960         970         980
CGCCGTAAAT CGCCGGAGCC AGATCCGCCT CGGCCAGCCA GAAGTCGCGC TTGGCGTTGA TCGCGAGCAG
990         1000        1010        1020        1030        1040        1050
GATCGAATTG ACCTTTTCGC GGCTGTCGGC GAGCAGTTCG AAGGTGCTGG TGATCATCCC GTTATAGGTG
1060        1070        1080        1090        1100        1110        1120
AGGAGGGACT GCTCCTCGAT CTTGGTGCGC AGCGGCACGA CGCTGTTGCG GTAGTGCCGG GCGATGTCGT
1130        1140        1150        1160        1170        1180        1190
AGTTGGCGCG GTAGGCCTGA TAGGCCGAGC GTGCTTCCGA GCGGACATTG ACGGCCTTCT CGGCGAGAAG
1200        1210        1220        1230        1240        1250        1260
GTTCCCGGCC CGCATATAGG CGAGCTCGGC CTTGCGCATG CGCGCCTTGC CGCTGTGCGAA GATGGGAATG
1270        1280        1290        1300        1310        1320        1330
GCGAATTCGA GATCGACCCG GCCGGTCGTC TCCGACGAAA TGTGCGCGTC CTCCTTTGCC CGTCCGGTCT
1340        1350        1360        1370        1380        1390        1400
CGAAGCCGGC GAGGAGGTTG AGGTCGGTGA CGTAACGGGT GGCCTCGGTC AGATTGTAGG ATCTGGCCGT
1410        1420        1430        1440        1450        1460        1470
CGCCTCCAGG TCGAGCTTCG CCATCTGGAG ATCGACACGG CGCCGAGGG CTTCGCCTC GATCTGGTCG
1480        1490        1500        1510        1520        1530        1540
CGCTTCATCA GGCCTTTCGG CAGGGACGCC AGCCGGTTCG GCACCTGGAA TTTGACGTCC GAACCCAGGA
1550        1560        1570        1580        1590        1600        1610
GGCCCATCAG CCGCGTAGCT CCTCCTTGCC CAGACGCGCC TCGAGCCGG CTTGGCCCC TTCTCCCGCC
1620        1630        1640        1650        1660        1670        1680
AGTTCGGCGT AGAATACATG CTCGCGGGCC TCGACCCCT TCGTCATGGC CCCGCTCTCG CCGAGCTTTT
1690        1700        1710        1720        1730        1740        1750
CGGCAAGCTC GGACGCGCGC TCGGCGCGCC CCTGCGCCTG GTTGTAGCTGG CCCACCGTTT CCCAAGCCGC
1760        1770        1780        1790        1800        1810        1820
TACGGCATTG ATCCAGGCAC GCCGCGTCTC GGAAGCAAGC TGCAACGTGC GCGAAGCGGC ATCCAGCTGC
1830        1840        1850        1860        1870        1880        1890
GCCTTCCGGA ACTCGGTGTC GCGGAGCGCG ATATTCCTGT CGCGGGTGAC GAGCGGAGA ATGTTGGAGA
1900        1910        1920        1930        1940        1950        1960
CGATCGCGCC CTCGACCGTC CGGAACGCC TAAGTCCAGG CGTGCCGATF CCGTAAGCC CGACACCGAC
1970        1980        1990        2000        2010        2020        2030
GGTCCGGTTG ACGAGCATTG TCGCCTGCCA GGCCTCGGCC GCCGACTCGC CGAGATCGGC ATAGGCCGCC
2040        2050        2060        2070        2080        2090        2100
TGCAGCCCCT TGTTGTTGAG CAGCGCGACC TGGACCGCCG TCTCGACGTC GATGGCTTTC TTCGCCATCA
2110        2120        2130        2140        2150        2160        2170
GCGCCTTCAC CCGCTCGGAG ACGGAGCGGG CCTGTTCTG GTTCTGCACC CATAAGGTCT GCTTGCCGGC
2180        2190        2200        2210        2220        2230        2240

```

GGCTTCGGCC GTTGTGCGCCG CGACCAGGGA AAACCCGGCA TTCTTAACCG CGTACTTCGC CAGCGGTAAC  
2250 2260 2270 2280 2290 2300 2310  
GCAGCCGCCG AGTACCAGCG GCAGTGCGAG CGTCGCCGCC AGCTTCATCC TGGCTTTCAT CATGATGCGC  
2320 2330 2340 2350 2360 2370 2380  
CTCCTTTCGG AGCCTGCGCA TCATTCTGGT CGCGCCATGG TTTGGGATCG ACGGCCCTGC GGTCCGTATA  
2390 2400 2410 2420 2430 2440 2450  
GCCCCGACA GGACTCGTAT AGGCAACGGG GCGCACGTCG GCATCGTTGG CCGGAGCTC CCCCAGCG  
2460 2470 2480 2490 2500 2510 2520  
ATCACATCGG GAGGCAGAGT GGCGGCGCAA CCGCTCGCAA GAAGCGGCG GACGCCACG AAAAAACAGT  
2530 2540 2550 2560 2570 2580 2590  
GTTTCATGGA TTTATCCGAA TAGGAGATTG CGGCAGCGCT CGAGAGCAAG GCTTGCTCGC GAACACCGGT  
2600 2610 2620 2630 2640 2650 2660  
TCGACCCGTC AATCGGACGG GTGCATCCTA TTCAGGCGCT TGGGGGCCG TGCAGCAGCG GCATTTCCGC  
2670 2680 2690 2700 2710 2720 2730  
CGTCGTCCGG GCGTCGTCGA CGAATTCCCG GACGGGGGAT ACGCGGGGTC CGCCAATAGC GTCGGTCCGC  
2740 2750 2760 2770 2780 2790 2800  
GCGAGAATCG CCATGCCGGC ACAGAAGTCT TTGCAGCATT GCTGCTTCGC GGGCTTTTGG TCGCCTTCGG  
2810 2820 2830 2840 2850 2860 2870  
CAGAGGAGTC TTGATCGCCG TGCGAATGCA CGCCGCTCGC CATATCGTGA TGGTCGCTCG ACTGCGCGAT  
2880 2890 2900 2910 2920 2930 2940  
TTCAGATTTG GACCAGCCGC TGTGCATGGC TGCGGAGGAA CTCGGCAATG CGTATACCGC CAGTGACACG  
2950 2960 2970 2980 2990 3000 3010  
AAAATCACCA AGCGGAACAG AATCAACAGC TTTTGCCTCG TGATTCGCGC CATGCCCATG CTCAATGCCT  
3020 3030 3040 3050 3060 3070 3080  
ATCATCGGGC TCAAAATTTG CAATCTTCAG CTCACATCGA CACGCAAGGC GCATCGACGT GTGGCGCGAC  
3090 3100 3110 3120 3130 3140 3150  
TACAACGGGC GAACCTCGCC GCTGCAGCCC GATTTCGAGT TCTAGGGCTT ACAACGATGG CAACGTCAAT  
3160 3170 3180 3190 3200 3210 3220  
CGGAAAAAAT CGGGGACGAA GAAAAAAGTC GAAAGTCTCA GGTGCCGTGG CATGCCCGCC ACGGTCAGGA  
3230 3240 3250 3260 3270 3280 3290  
AATTGTTAAC CATATCTCC GTTCTCGGG GAACAGCCCG GCCGCGCCA CAGGAATAAC TCTATTCCGC  
3300 3310 3320 3330 3340 3350 3360  
GCTCTTTCGA ACGCGGAAGC AAAACAACAG GTCGAGAAAG ATGGACATGA TCAAATCGGG TCTCGGCGTC  
3370 3380 3390 3400 3410 3420 3430  
ACCGTCGCGG TTCTCGCCCT GTCCATCACA GTCAGCGGCA TCCTGCATTT TCGCTGTCC AGCTACCGCC  
3440 3450 3460 3470 3480 3490 3500  
CGGCGCGGAA TGCCGCTGT CGCACGAAAT CACCTCATCG ATACCACGAT GAAAGAGGCC GCCTTGTCAG  
3510 3520 3530 3540 3550 3560 3570  
GCGCACAAAG ATCGCGCGAG GATATGCGTC AGGAAACGCG GCGACACTTT TCCGGGCCGA AATCTCCAG  
3580 3590 3600 3610 3620 3630 3640  
CTGAGCCAGA AGTTCGTCGC AGGAGCACTT CTTCGTGTCC ATCAAGTCA GAATGCGCTT GCGGCCCTG  
3650 3660 3670 3680 3690 3700 3710  
TTGGCCTTCA CACCATCCCT GCCGACATTG TGCAGAGCGC ACCAAATATC GAGCGCTCCG CGCATGACGT  
3720 3730 3740 3750 3760 3770 3780  
CCACGTATC CGGAGAATAG GCCTGACGCA CGAACCCTC CAAAGGTGTA CCTCAACGAC CTCATCATGG  
3790 3800 3810 3820 3830 3840 3850  
GGCAATCGGG GGGAAATTCA AGGAAATAGT GAGGTACCGC AATAAGTCAT ATGATTGAGG GATTGCAAGC  
3860 3870 3880 3890 3900 3910 3920  
CGCAGCCGCC GGTTCCTTTT CTCTCGTCAC ACCCGTACGG ATTGCTCCAC GAAATCCTCG CTGCTCGAAA  
3930 3940 3950 3960 3970 3980 3990  
AATTTTCAGT AGCGCTCGAC CTCGGCTGGA TCCCCGCTCG CCTTGCCTGG GTTGTCCGAG AGCTTGACGG  
4000 4010 4020 4030 4040 4050 4060  
CCGGACGGCC ATTTGCTTCG CTGACCTTGC ATACTATGGA GATCGGGTTG AGGCCGTTGA TTTCTGTCCG  
4070 4080 4090 4100 4110 4120 4130  
GGCGCAGCCG GCGAAATCAT TCGTCAGGTT CGTTCCCCAG CCCAAGCTTA TGCGCACGCG CCCTTAGAAA  
4140 4150 4160 4170 4180 4190 4200  
TGGCAGTAGG TCCTGATGAT CGTGTGACG TCGAGCCGTC GGAGAAGATC AGAAGCTTCT CCCGCGGGTC  
4210 4220 4230 4240 4250 4260 4270  
TCGGCCCATC TTCTTCCACC AGGCGATGAT CTTCCTCGCT CCCTCGATGG GCGGGGCACT GTCCGGCCGG  
4280 4290 4300 4310 4320 4330 4340  
AAGCCGGTCC AGTCCGCCAC CCAGTCGGGT GCATCCCTGA GGAACGCGGC CGTCCCGAAC CCGTCCGGGA  
4350 4360 4370 4380 4390 4400 4410  
GCACGATCAG GAGGTGCGC CCATAGAGCT GGTTCAGTCT CTGCAGAACC TTGTAGGGCG CGGCCGCAAG  
4420 4430 4440 4450 4460 4470 4480  
CTCGCGTTCG TTGCGGGCAA GCGCCGACG CACCATCGGC AGCTCGTGC CATTTGGTGC GACCGCTCC  
4490 4500 4510 4520 4530 4540 4550  
AGATCATTTG CCATTGCCAG CAGCACGTTG CTCGAGCCTG TAAAGGACGA GCCGATGCCT TCCTTAAGCG  
4560 4570 4580 4590 4600 4610 4620  
CCTAGACGCA CAAGCGCTGC CATAGGAAGC TGTGCCCGCG GCGCGTACCG AAGTGGGAAA TGCGCAGGTT )  
4630 4640 4650 4660 4670 4680 4690  
CGGATGCTCC CGAACTGCT CGACCTTCGA CCACATCTTG GCCTTGGCGC GGCATAGAG GACGTCCAGC  
4700 4710 4720 4730 4740 4750 4760

GTGAATGGGC CGAGCCCTTT TATCGCCGCG CGCGAGCGCA ACTCGTTGAT GATCGCGAGT GCAGGGATTT  
 4770 4780 4790 4800 4810 4820 4830  
 CCCAATGGTC GTCTCGGCC AGCGGCCACG AAAGGTCAGT TCGAATTGGC CGTCACGACG CGACAGCTCG  
 4840 4850 4860 4870 4880 4890 4900  
 TATTCCGGAA GCGGGAATCT TCGGAGCCAG GCGAGAAAT CCGGCGAAAA GATCTGCGTG CGGCCATAGA  
 4910 4920 4930 4940 4950 4960 4970  
 AGCTGTTACC CGCAAGCCAT ATCATTTCTT TCTTGGTGAA GCGCAGGCCT CGCGCGTGGT CGAGCTGGTC  
 4980 4990 5000 5010 5020 5030 5040  
 GCGCAGTTCC TGTTTCGTCG TCTCGTCGGC AAGGCGCACG GTCTTCGTGC GGTGATCAG CGAAAAGGTC  
 5050 5060 5070 5080 5090 5100 5110  
 GCGTCCACGT CCGGATAGAG CTGCCAGATC ATCTGCAGCA TCAAAAGCTT GTAGAAATCC GTATCGAGCA  
 5120 5130 5140 5150 5160 5170 5180  
 GACTGCGGAC GATCGGATCG AGCTTCCAGG TATGTTATA GACCCGTCCG GCGATATCGG TCTTTGGCAT  
 5190 5200 5210 5220 5230 5240 5250  
 GCAATACGCT CCTCTTGGAC CGTCATGACC CGCTCTGACG CGAGGCGATC GTGGTCGATA TCCTAAGTCC  
 5260 5270 5280 5290 5300 5310 5320  
 GAGCGGCGAT GCGGCCGACG CCGCGCTTCC CCGTCCCGGG GCGGACTAGA TCAGCCGGAC GCCGGCATCG  
 5330 5340 5350 5360 5370 5380 5390  
 CGCATGCGCA GGGTCATGGC CGCAAGCGAG CCGTTCAGAT CGATGCCACG GCAGGCGCCG AGCACCACCG  
 5400 5410 5420 5430 5440 5450 5460  
 ACGTCGAAAA ACCCTTGGCG ACCGCATCGA GTGCCGAGAA GGCAACGCAG AAATCGGTAG CCAGGCCGCA  
 5470 5480 5490 5500 5510 5520 5530  
 AAGCGTCACG CTTCTTATGC CGCGTCGCG CAGGTAACCG GCAAGCCCGG TCGGGTCCG GTGGTCGTTT  
 5540 5550 5560 5570 5580 5590 5600  
 TCGAAAAAGG CGGAATAACT GTCGATCTCG CGACGGAAGC CCTTGCCTAT GACGAGTCC GCCGAGTCC  
 5610 5620 5630 5640 5650 5660 5670  
 ATCGCAGGGA CGGATGAAAA TCGGCACCGG CGCTCCCTG GACGCAATGC TCGGGCCAGA CCGTCTGCTC  
 5680 5690 5700 5710 5720 5730 5740  
 TCCGTAAGGC ATGCCACGG TCTGAAACGG AGCCTTGGCC GGATGCGTCG AGGCGAAACT GGAATGGCCG  
 5750 5760 5770 5780 5790 5800 5810  
 GCCGGGTGCC AGTCTGCGT CAGGACGACA TGCGGCGACG CGTCGATGAG GCGATTGACG GCGGGTACGA  
 5820 5830 5840 5850 5860 5870 5880  
 TTTCATCGCC GCCCTCGACC GCGAGAGCGC CGCCGGGGCA GAAGTCGTTT TGCATATCGA TAACGATCAG  
 5890 5900 5910 5920 5930 5940 5950  
 GGCTTCGTCA GCCATTGCCT CTCGCAAATC CGGCCGAGCC GCATCGGCCA TCAACAAATC GATCCCGATC  
 5960 5970 5980 5990 6000 6010 6020  
 ACGATGATTC AGGTCGGATC GACCTAAGTC ATGAACGTGA TCGATTCTAA TAAGTTAGAG CGGGATGCGG  
 6030 6040 6050 6060 6070 6080 6090  
 GCGGAAAACC GCATACACTT TTCTCATCC CGCTCTAAC CAGATCAACC ATAGAACCGC CGCGGCGGCA  
 6100 6110 6120 6130 6140 6150 6160  
 AGCATGTTTCG ACGCTCGCGA CGCGAATCCC GGCGAACCTG CGAACCTTAT AAAGGTGATT GACTATAGGA  
 6170 6180 6190 6200 6210 6220 6230  
 CGTTAATGCA TGCCGCGCCA AAGCATTCCG CCGGCAACCG CGTCGACACC GCCACGCCT GGTGTCGCGC  
 6240 6250 6260 6270 6280 6290 6300  
 CACTCCAACG TCAGTTGTTG AAATCGTTGA GATTGTCGGC GACGTTACG CCGGAATTCG CGATACCGAT  
 6310 6320 6330 6340 6350 6360 6370  
 CGGCCGTGCG ACGAAATCGA GGAATTTCTG GATGTCGACA GAGGGATGGC GCGAGGCGTA GATGACGTCG  
 6380 6390 6400 6410 6420 6430 6440  
 CGATTCTTGA CCGGGAAGGT CTGACCAGC ATGAGGCTGT CCGGTTTCTG CATGTCGAAC CCGTAGACGA  
 6450 6460 6470 6480 6490 6500 6510  
 CCGGGTAGCG CCCGACGCGG TCGGCCTTCA TGCCCTTGCT CAGCAGCTTG TGAAAACGCT CCTGCCCGAG  
 6520 6530 6540 6550 6560 6570 6580  
 CAGGCTCATG ACGATATCCG GCTCCTCGTA GCGGAACACG AAGTAACCTT TGGCATCCAC CGTCCGATCC  
 6590 6600 6610 6620 6630 6640 6650  
 GAGCCGCGC CGGCGAGCGC CACGGCTCG AGCAGGTTGA GATCGTTGGC GCCGAACTCG ACGCGCTGAT  
 6660 6670 6680 6690 6700 6710 6720  
 TTGCCCGCAC CTGGCCAAGC ACCGTGAAGG TGCGCGGATC GCGCGTCACG AATATCTGGT CGCCCGGCTT  
 6730 6740 6750 6760 6770 6780 6790  
 TACGTGAATG TTTTCCGAGG GGCTCTCGAT GATCGACTTC AGCAGGACCG TGCCGGTCTT CTTGCCGCGC  
 6800 6810 6820 6830 6840 6850 6860  
 ACCAGCGTCA CATAGTCTC GTAGGCTGC GCCGATGGG CGCCGCCTT CGCGATGACC TC GTTGTGATCG  
 6870 6880 6890 6900 6910 6920 6930  
 TCTCGCTGAC GAGATTGAGA GGCACGACAG AAGGACGGCC GACGGCACCG GACACCGTGA CGATGCGCGA  
 6940 6950 6960 6970 6980 6990 7000  
 CGCCGTCGAC GTCGACGTCA CGATCACGTC CGGCTCGACG GCCTTCTGTT TCAACGCATC GAGGATCGCC  
 7010 7020 7030 7040 7050 7060 7070  
 TGGCGCGCCT GCTCGAGCGT CAGGCCGGCA AAACGGACCG ATCCGACATA AGGAATTGCC GCCTTGCCAT  
 7080 7090 7100 7110 7120 7130 7140  
 CGGGCTGGAC GACGAGATCG ATGCTCGTCT GCTTCGATTC AGTTGTGGAA AACAGCCCGT CGCTGCCCGC  
 7150 7160 7170 7180 7190 7200 7210  
 CTCGAAGATC GTGACCTTCA GCTGGTCGCC AATGCCGATG ACCACCTGAC CGACGACGCC GCCGATGCGG  
 7220 7230 7240 7250 7260 7270 7280

AAGCGCGGC TGAGGGTTGT GGAAACATAT TCGGAAACGA GCCGCGCCGA CTGGCCATCG ACATCTACAA  
 7290 7300 7310 7320 7330 7340 7350  
 TATCAAATAC CGTCGCGTTC CTTCTGCCGA TCTCCGAGCC GGACTGGCCG GCATCCGAAA CGATATCGGT  
 7360 7370 7380 7390 7400 7410 7420  
 TGTCAGAGGG CCCTCGCCCG GCACGGCCTG GCATCCCGCA AGCGCGGTAC AAAGCAGAAG GGCAGTCACT  
 7430 7440 7450 7460 7470 7480 7490  
 CGTTTCACTA CGACACTCCG TGTTCGGCAC AGCAATGATT CCATACAGAT ATCTGCGCCG CATTTTCCCG  
 7500 7510 7520 7530 7540 7550 7560  
 GCAGGGAATC CGCCGGTCTC TGAAAAGCAC AAAGGCCGTC CCACAACAAG CCACCTCCCC GCGACAGCCG  
 7570 7580 7590 7600 7610 7620 7630  
 GCAAAAATAA CGAGACTGCG CTCTTCCGGC CATAACCCCT CGATATTGGT TACGAACCGT CGCCGTAATT  
 7640 7650 7660 7670 7680 7690 7700  
 GTATTGGATG GATCGCAGGA GACGTATATC TAGCCCTTCG GGGTTGAAAG CTAGGCACTA TGAGTTAAAA  
 7710 7720 7730 7740 7750 7760 7770  
 GCTCACTAAT ACAGCGCAGG GGATAAAACC GGGGATACCC GCAGAGCAGG AATGTCGCCG ATTGCAATTT  
 7780 7790 7800 7810 7820 7830 7840  
 ACGTTAGACC TTGCAATGGG CTAAGACAAT TTTAACGGCA GTTCAGAGCG TTCGATCTTA CATGACGGTT  
 7850 7860 7870 7880 7890 7900 7910  
 GAAATCATAG GCGGTGCATG CCTTGCCCCG GGAGCGAAAT CTCCGCAGGC GCTGTTCAAA ATCCTTCGTC  
 7920 7930 7940 7950 7960 7970 7980  
 AGGGCAAGTG CACCGTGACC CGCGTCCCTT CGGACCGCTG GGACCTTGGC CGTTCCTGGC ACCCGGTGAT  
 7990 8000 8010 8020 8030 8040 8050  
 GGGCACGCCG GGCAAGACCT ATTCGTTTGC CGCCGGCGTC CTCGACCATA TCTACGATTT CGATCCCGCC  
 8060 8070 8080 8090 8100 8110 8120  
 GTTTTCGGGA TGTCGCGAGC CGAGGCCATG TATATGGACC CGCAGCAGCG GGTTCCTGTTG CAGCTCGCCT  
 8130 8140 8150 8160 8170 8180 8190  
 GGCGAGCGCT CGAGGACGCC AATATTTCCG TGGCGTCGCT GCACGGTGAA AATGTCGGCG TCTATGTCGG  
 8200 8210 8220 8230 8240 8250 8260  
 CGCCTCCAGC CTCGACCACG CCAATCTGAC GGTGACGAT CCGGCTGCCG CCGGGCCCTA CTTTATGACG  
 8270 8280 8290 8300 8310 8320 8330  
 GGCAATACGC TGTCGATCGT CTCCAACCGT ATTTCCACG TCTTCGGCCT GAGCGGCCCG AGCATGACCG  
 8340 8350 8360 8370 8380 8390 8400  
 TCGACACGGC CTGTTCTCTC TCGCTCGTGG CCCTCGATCA GGCATGCGCG GCGTGAACG CCGGAGAGAT  
 8410 8420 8430 8440 8450 8460 8470  
 CGACACGGCC ATCGTCGGCG CCGTCAATAT CCTCGCCAC CCGCTTCCCT TCCTCGGCTT CGCCAGGCG  
 8480 8490 8500 8510 8520 8530 8540  
 CGCATGCTCT CGCCCGAAGG GCTCTGCCG GCCTATGACA ATGACGGCGC CGGCTATGTT CGCGCCGAGG  
 8550 8560 8570 8580 8590 8600 8610  
 GCGGCGTCGT CTTCTGCTCT CGCCGACGG ACCGGGCGCA ACGCGAGCG GACCGGAGCT ATGCCCGGAT  
 8620 8630 8640 8650 8660 8670 8680  
 CGTCGCCTCC GGCCTCAATT CGGCCGGCG CACCAACGGC ATCTCGCTGC CCTCGCGCGA GCGCAGGCC  
 8690 8700 8710 8720 8730 8740 8750  
 AATCTGCTGA GGATGATCTA CGAAGGCAAC GGCATCGACG CCAACCAGGT TGCTTCGTA GAAGGCCATG  
 8760 8770 8780 8790 8800 8810 8820  
 GTACCGGCAC GAAGGTCGGC GACCCGGCGG AGGTCTGGTC CATCGGCACG GTGATCGGCG CGAAGCGCCG  
 8830 8840 8850 8860 8870 8880 8890  
 GGCTCCGGTG CCGATCGGCT CGATAAAGTC CAATATCGGC CATAACGAGC CGGCCTCCGG CCTGTTCCGG  
 8900 8910 8920 8930 8940 8950 8960  
 ATGATGAAGG CGGTTCTGGC TCTGGAGAAC AATATCTGCT CCGCCTCGCT GCATTTGAG ACGCCGAACG  
 8970 8980 8990 9000 9010 9020 9030  
 AGCAGATCGA CTTGACGGC CTCAATGTCC GTGTGACTGC CAATCCGATC GAGCTTCTGA AGGGCAAGCG  
 9040 9050 9060 9070 9080 9090 9100  
 GCGCGGCTC GCCGGCATT ATTCTTTCGG CTTGCGCGGC GCCAACGCC ATGTCGTCAT CAGCGATCCG  
 9110 9120 9130 9140 9150 9160 9170  
 GATCCGGCAC AGGCGGAAAA GGCTGCGACC TCGCCGGCGG GCCACGTGTT TCTGGCAAGC GCCCATACCG  
 9180 9190 9200 9210 9220 9230 9240  
 CGTCGAGCCT CGAAAACCTG CTCGAGAGAC TATAAGGCCG CCTTTGCGGC GCGACGAAGG AAGAGACGCG  
 9250 9260 9270 9280 9290 9300 9310  
 CGCGATCGTC GCCCCTCCG GCGCCAATCG GACGCATATG CGCCATCGGT TCGCGGCGCG CAGCGATCAT  
 9320 9330 9340 9350 9360 9370 9380  
 CCCGAAGATA TCGTCCGGG GATTGCCAAC CACCTGGAAA AGCCCGGCTC CGATATCGGG GAGACGGGCG  
 9390 9400 9410 9420 9430 9440 9450  
 AAGCGGCCGC GAAAGAAGCA AAGGTCGCCCT TCGTCTTCTC CGGCAACGGC TCGCAATGGG CCGGCATGGG  
 9460 9470 9480 9490 9500 9510 9520  
 TGTCGAAGCC TTCCGCGAGA ATCTGCACCT CCGTCAGTCG TTCACTTCGG TCAGCGCCCT TTTCAGGTTT  
 9530 9540 9550 9560 9570 9580 9590  
 CATTGCGATA TCGTCTCAC CGATCTCCTG ACCGACCCCG GACTCGACAG GAAGCTCGCG GATACGAAGG  
 9600 9610 9620 9630 9640 9650 9660  
 TGGCCAGCC GCTGCTTTTC GCGATACAGG CCGCGCTCTC CCACTCGCTC GTGGCCATGG GCATCAAGCC  
 9670 9680 9690 9700 9710 9720 9730  
 GACCGCCGTC TTCGGTCATT CGGTGCGGA AATCGCTGCG GCTTATGCAG CGGGGGCGCT GTCGCTCGTC  
 9740 9750 9760 9770 9780 9790 9800

GACGCCGTGT CGATCGTCGC CAAACGATCG CTGCACCAGG ACCTTCTGGC CGGCCAGGGC ACGATGGCGG  
 9810 9820 9830 9840 9850 9860 9870  
 CCGTTATGCT GGGCGAGGAG GCAGCCAAGG CCTTCGCGGC GGAGCGCGGT CTCGACGATG TCTGCGTCGC  
 9880 9890 9900 9910 9920 9930 9940  
 AGCAATCAAC GCTCACAATT CGGTGACGAT CTCCGGCCCC GCGAACGAGA TTTCGGCGTT CCGCGACGCT  
 9950 9960 9970 9980 9990 10000 10010  
 GCGCGCAAGG CGAAAAATCC GGTTCAGATC CTCGACATCA ATTATCCGTT CCACCATCCG ATCATCGACC  
 10020 10030 10040 10050 10060 10070 10080  
 GGGCGAAAGA GCGGTCCTC TCCGATATCC CTGATATCGC CCCCCGCCG ACGGAGCTGG CCTATCTGTC  
 10090 10100 10110 10120 10130 10140 10150  
 GACCGTCACG GGTGCCGCGC TCGACGGGAC TGCTCTCGAT CCCGACTACT GGTGGAAGAA CGTTCGCGAG  
 10160 10170 10180 10190 10200 10210 10220  
 CCCGTGCGTT TCCAGGCCGC TGCCGAGGCC GCGCTCGATC TCGGCTGCAT GCTGTTCATC GAGATCTCGC  
 10230 10240 10250 10260 10270 10280 10290  
 CGCGTCCGAT CCTTGGCTCC TACGTCAAGG AGACGATCAA GCAGGCGGCA GTGCCGGTT CGGTGGTGGC  
 10300 10310 10320 10330 10340 10350 10360  
 GACATTGCTG CGCGACGCCG GCGCGAACGG CCACGATCCG ATCTCCGCGT CGATGGCGCG CGCAGTCGCC  
 10370 10380 10390 10400 10410 10420 10430  
 CACGGTGGCC CCGTCGACCG GTCGCGTGTG TATGGCAAGC GGGACGCATT CATGGAATC CCCGCGCTGC  
 10440 10450 10460 10470 10480 10490 10500  
 CCTTCGAGCC GGTGGAATG CGGCCGGCGG CCACGACCGA TGCGACGGAT CTCTTCGGAC GTTCCGCAAA  
 10510 10520 10530 10540 10550 10560 10570  
 GCCCTATCGG CTCAACGGCT GCGCGGCCGA CCCC AATGCG GGAAGCTGGA AAAACCATGT CGACGCGCAT  
 10580 10590 10600 10610 10620 10630 10640  
 CTCTTCCCGG ATCTCGCCGA ACACGTGGTC GACGGCAAGG CGATCCTCCC GGGCAGCGGT TTCATCGAAA  
 10650 10660 10670 10680 10690 10700 10710  
 TCGCCGTTTC CGCCGCCAG CAATATCAGC GCAGCGACGA GGTGAGATC ACCAATCTGG AGATCGTCCG  
 10720 10730 10740 10750 10760 10770 10780  
 CCCGCTCGAA CTGAGCGACA GCCGCATCAT GGAACCTCTG ACCATCCTGT CGCCGGAGAC CGGCGACATC  
 10790 10800 10810 10820 10830 10840 10850  
 GAGATCCGAT CGCGCGAACG TCTCTCGGAA GACGATTGGA CGGTGCATGC CGTCGCGGGA AGCCGCAAGC  
 10860 10870 10880 10890 10900 10910 10920  
 CCTTGCCGAC GACGAGCAAC GACTGTCCGT CCACCAAAGG CTTGGCGAAG ACGGCGACCG TGACGGCGGC )  
 10930 10940 10950 10960 10970 10980 10990  
 CAAGGCCTAT GAAACGGCGA AGCAGTTCGG GCTGGACTAC GGCCCGCAAT TCCAGCTTCT CTCCAAAGCC  
 11000 11010 11020 11030 11040 11050 11060  
 GTGGCCTATG GCGACCGGCT TATCGACGTC GAGTTGAAGC ACCCCAGGC GTCGGGGCAC CCCACGTCA  
 11070 11080 11090 11100 11110 11120 11130  
 CCTACGCCCT TCATCCGATC TCGGTCGACG CGACCTTCCA CGGCCTGGTC GCACTCTTCG ACCGCTTCC  
 11140 11150 11160 11170 11180 11190 11200  
 CGGCGACAGG GCGCGGCCT ATATTCTGT CCGTTTCGGC TCCGTGCGCG TCGTCAATGC CGGCCGGCCG  
 11210 11220 11230 11240 11250 11260 11270  
 ATCGCCCGCG CGACGATCGA GATCGAGCGG GTCACGCGCA ACTCGATCAA GGCACGGTTC CATTTCTTCG )  
 11280 11290 11300 11310 11320 11330 11340  
 ACGAGACCGG CGTCGCGATC GCACACTTCG AGGATTGCCG CTTCCGCCGT ACCTATCTGC GCAAGCACAA  
 11350 11360 11370 11380 11390 11400 11410  
 GACGCTTGAA GGCCTTTCCT TCCACTATGA GGCAGTTCG TCCGACGTGG TCATGCCCGG CACGAAAGTT  
 11420 11430 11440 11450 11460 11470 11480  
 GCGACGGCCA TGCCCTGCGAC CCTGGCCGAG CCGATCGATG ACAATGGCGT CGACAACGCG ACCCTGTCT  
 11490 11500 11510 11520 11530 11540 11550  
 TCAACGCGC CATCTACCGC GCCTGCCAGG AAATTGCCCT GAAGCTCGGC AGGGGACCG CCACAGTGCG  
 11560 11570 11580 11590 11600 11610 11620  
 AGGCGACGCG CTGCCCGGG ACTTCGCGTT CCAGTGCTTC CTGACGAGCT GCCTGCTGAC GCTCGAGGAT  
 11630 11640 11650 11660 11670 11680 11690  
 GCCCGCCTTT GCGAGACCA GAACGGCAAC TGGAAGGTCG CACGGGAGTT CTCGCTTCCG ACCGTTCCGG  
 11700 11710 11720 11730 11740 11750 11760  
 AGATCCTTGG CGAGCTCTAT GGCGACCGCC CCGATCGTGC GGTGGAAGCG GTGCTCGTCA ACAACGCCTA  
 11770 11780 11790 11800 11810 11820 11830  
 TGCGGAAGCA CTCAGCCGTC TCGACCGCT GCTTGGGCCG GTGCGTGCAG GCGACGACG GGCTTCGTT  
 11840 11850 11860 11870 11880 11890 11900  
 AATGCCGGTT TCATCAGCGA AGCGACGCTC GACCATCAGG CCGTGCATTC CGTGGCGAGC CGTTCGCGCA  
 11910 11920 11930 11940 11950 11960 11970  
 TGGACCAGGT GCTTCATACG GTCGAACACG TTCTCGCTGC GCAGCCGGCA GCGCCTGCG GCTCATCGAA  
 11980 11990 12000 12010 12020 12030 12040  
 CTCGGCAGCG TCTCCGGCTT CAGCCGCCGG CTGGCCGATC TGGCAGCGCG CAATGGTGCA GCATTTTCCA  
 12050 12060 12070 12080 12090 12100 12110  
 TTTTCGAGCC GCGCGACAAT GCACAACGGA ATCTGGAAAT CGCCTTCGAG GAGGATGAAT GTGCGCGTTC  
 12120 12130 12140 12150 12160 12170 12180  
 TGAAGAAGAG CGAATTCGCG GATATCGGGC CGTTCGACCT CGCCGTCAGT GCTTCGGACA ATCTCTACCA  
 12190 12200 12210 12220 12230 12240 12250  
 GCTGATCGAA GAGGAGAGCG GCGTCCGCAC GGCCCTTCGC TCCATGCCCC CCGGAAGCGG GCTCGTCGCG  
 12260 12270 12280 12290 12300 12310 12320

GCCGTCAGCG CGCCGTCGAT CTTCCGCGAT TTCGCCCTGG GTCTCTCGGG CGGATGGTTC GAGCGCAGCC  
 12330 12340 12350 12360 12370 12380 12390  
 AGACGGCAGA ATTCCCGATC GGCAGGATCG CGGCCGTACC CCATTGGCAG AAGTGTCTCG CAGATCTCGA  
 12400 12410 12420 12430 12440 12450 12460  
 CCTCGGCGAC GTGGTCGTGA CGGATCGCGA ATGCCCGCAT GGGAAACACCA TCCTGATCGA GGCCCGGGGC  
 12470 12480 12490 12500 12510 12520 12530  
 ACAGCGGCGG CAGGCGCGGG CGCCGAGCCG GGCAGAACCG GAACACCATA TCTGCTGATC GAGACTGCCA  
 12540 12550 12560 12570 12580 12590 12600  
 CTGCCAGAA ACCGGTCCGG GCCGGTGGAA GGAGTGCCCG TCCCGTCCGT GTGACAGGGC ACCTTGCCGC  
 12610 12620 12630 12640 12650 12660 12670  
 AGATATGGCG ACGCTCCGCG CAGCCCTCGA AGCCATGGGC GACAGACCGT TCGCGCGGT TTTCATGTCC  
 12680 12690 12700 12710 12720 12730 12740  
 GAGGCGGAGC GCGCCGTCGG CGACGGTTCG GCCCTGCTGC AGACGCGTGT ATTGGCGCTC AGCCCTTTG  
 12750 12760 12770 12780 12790 12800 12810  
 CCGAGGCATT GAAACAGCAT CTCGCCGGCG TCGAGCCGAG GGCAGACACC GCGCGCCCTT GTGCTCGTGG  
 12820 12830 12840 12850 12860 12870 12880  
 CCCCAGGCGG CGCTCCCTTA TCACCTCGGC AGGCTCCGGG CTGAATTCCG GATTGTGGGC CTTCCGCCGC  
 12890 12900 12910 12920 12930 12940 12950  
 GTCCTTCAGA ACGAGTATGA GTTTATCGAC GTTCACACGC TCGATCTTGC CGGTGATGCC GACATCGGCC  
 12960 12970 12980 12990 13000 13010 13020  
 GCAAGGATAT GCTTGGCCGA GCGCTFCCCC TGCTGACGGT TTCGGGACAG AACCGGGAAT GGCTGCTCGA  
 13030 13040 13050 13060 13070 13080 13090  
 CCGGAAGACG GGCCTGCTGT CGGAGCTTCG GGCCGTACCC GGCACAACGA ACAAGGTTGA CGGGAAAACC  
 13100 13110 13120 13130 13140 13150 13160  
 AACGCCTTCG CGGCCCGCAC CATTCCGCAG CGCGTCAGCT CGCAGGTTGC AAGCATCGCC TGGGAAGAGG  
 13170 13180 13190 13200 13210 13220 13230  
 CGAGCGTACC GGAGCCGGGC CCCAACGAGG TCGTCGTTGC CGTCGCGGCA ACCGGCTTGA ATTTCCGCGA  
 13240 13250 13260 13270 13280 13290 13300  
 CGTGATGTGG GCCATGGGCC TGTTGCCGGA AGAGCGTTG GAAGACGGTT CGCCGGGGCA ACCATCGGCA  
 13310 13320 13330 13340 13350 13360 13370  
 TGGAACTGTC CGGCCACGTG GTCGCGGTGC GCGAAGAGAT CGACGATCTC GCGGTCCGGC ATGCGGTGAT  
 13380 13390 13400 13410 13420 13430 13440  
 GGCCATCGCC TCCGCCGCTT TCTCGACCCA TGCGGTCTGC GCGCGCGCCG CATCGCCAAG TGCCCGGAGC  
 13450 13460 13470 13480 13490 13500 13510  
 GTTACGCCC GGGCCCGGC CACATTCGCG GTCGCTTCC TTACCGCCTA CTACGCCATT GTCGAGCTCG  
 13520 13530 13540 13550 13560 13570 13580  
 GCCGCATTCA GGCTGGAGAG ACGATTTTGA TCCACGGTGG CTCGGCGGCG GTCGGCCTTG CCGCGCTCCA  
 13590 13600 13610 13620 13630 13640 13650  
 GGTCCGAAAG CTCGAGGGC CCAAGGTCAT CGCCACGGCG GGAAGCGCCG AAAAGCGGCG CTTCTGACG  
 13660 13670 13680 13690 13700 13710 13720  
 ATGCTCGGCG CCGACCATGT CTTGATTCG CGTTCGCTCG GCTTCGTCGA GGATGTCCGT TCGGTACCG  
 13730 13740 13750 13760 13770 13780 13790  
 GCGGGGAAGG CGTCGACCTC GTGCTCAATT CGCTTTTCGC CGAAGCGATG GAGCAGAGCC TCTCGCTGGT  
 13800 13810 13820 13830 13840 13850 13860  
 GAAACCTTTT GGCCGGTTCC TGGAACTCGG CAAGCGCGAC TATTACGCCG ACAGCAAGAT CGGGCTCAGG  
 13870 13880 13890 13900 13910 13920 13930  
 CCGTTCGCGC GCAATGTCAG CTATTTCCGG ATCGACGCCG ATCAGTTGCT GGTAATGCG CCGGATCTGA  
 13940 13950 13960 13970 13980 13990 14000  
 CCAGGCGCAT CTTTCATGAC ATCGGAGCGC TCTTCGAGGA AGGCAACGTC ACGCCGCTTC CCTATCGCGC  
 14010 14020 14030 14040 14050 14060 14070  
 CTTTCGATTAC GGATGAAATC GGGCAACTTG TCCGGCTGAT GCAGAATGCC GGCATATCG GCAAGATCGT  
 14080 14090 14100 14110 14120 14130 14140  
 GGTGCTGCCC CCCGTTGCCG GCAGGCACGA GGTGACGGCG AAGGCTGTGA GGGCATGAA GTTCGACCCG  
 14150 14160 14170 14180 14190 14200 14210  
 GACGGCGTCC ATCTCGTCTG CGGCGGCATC GGCGGCTTCG GCCTCGTCCG CGCCAACCTGG CTCGTCGAGA  
 14220 14230 14240 14250 14260 14270 14280  
 AAGGCGCACG CCGTATTCGG CTTTGCTCCC GCCGAGGCCA GCCGGATGCG GAAACGCGCG CGATGATCGA  
 14290 14300 14310 14320 14330 14340 14350  
 GCTGCGGGCA GAATGGCGG GTTGGCGGCTT CGGTCCATGC TTGCGACATC ACGGATGCCG CGGCCGTGGA  
 14360 14370 14380 14390 14400 14410 14420  
 GACGCTTCTG GCGACGCTTC GTTCGGAAGC ACCGTTTCGCT CCGTCGTTCA CGCGGCGATG GTGCTCGAGC  
 14430 14440 14450 14460 14470 14480 14490  
 ATGCGCTGAT CAGCAATCTG AGCCGCGAAC GCAACCGGCC GGTATCGAA ACGAAAGCCA AGGGCGCGCG  
 14500 14510 14520 14530 14540 14550 14560  
 AATTCTCGAT CGCCTTACCC GTGGTGACAG GCTCGACAAC TTCATCTGT TCTCCTCGCG GACGACGCTG  
 14570 14580 14590 14600 14610 14620 14630  
 GTCGGCAATC CCGGCCAGGC GAACTATGTC GCCGCAACG GCTATCTCGA AGGCCTTGCC CGTGCACGGC  
 14640 14650 14660 14670 14680 14690 14700  
 GCCAGGAGGG CCTGGCAGGG CTCGCCATCG GCTTCGGCGC GATTGCCGAT GCCGGCTACC TGACGCAGAA  
 14710 14720 14730 14740 14750 14760 14770  
 CGCCGATGTC AACGACTGCG TCGCCAAGCG CATCGGCAAA ACGGCGTCTGA AGGCCAGGT GCGCTCGAC  
 14780 14790 14800 14810 14820 14830 14840

ATGGTGGAAA ACCACGTCGC CGCGGATCCG GGCACGGTTCG ACGCGGCCGT CGTGATGATT TCGGAGATCG  
 14850 14860 14870 14880 14890 14900 14910  
 ACTGGACGGC TGC CGCGGAAC CTGCCGGTCG CCGCAACGCG CTCTTCGAGG TGATCCTGCG CAGCGCCGAT  
 14920 14930 14940 14950 14960 14970 14980  
 CAGCATGCCG CAGGCGCCGA GGGCACGACG ATGGATCTCG TCGCCATGAT CGAAGGCAAA TCGCCGCAGG  
 14990 15000 15010 15020 15030 15040 15050  
 AGGCCGAGGA CATCCTCTTC GATCTGGTGG CCGGCGAGAT CGCTGCCATC CTTCGTGTCT CGAAGGATAC  
 15060 15070 15080 15090 15100 15110 15120  
 GGTCACTCGC GGAAAGATAT TGAAGGAGAT CGGCCTCGAC AGCCTCATGG CCGTTGAACT CGGCATGAGT  
 15130 15140 15150 15160 15170 15180 15190  
 TTCCAGCAGA ACACCGGCTT CGACATGCCT TTGAGCGGCG TCGCGGACAA TACCACGGTC GGCACATCG  
 15200 15210 15220 15230 15240 15250 15260  
 CGCGCAAGCT CCATGAAAAA GTAAGCAAGC CGTGACCAGG GCAACGAGAA CGAGGCCGGT GACGACAAGC  
 15270 15280 15290 15300 15310 15320 15330  
 TTGTAACAGA ACTCACGCAG CGTCATGTGC GAGCGGACAA GAAAAAGCA CTGAGTCAAT GAGCAACGAT  
 15340 15350 15360 15370 15380 15390 15400  
 GGAAACGGCC AGACCTTCCC GAAAATGAAC AGCAGTCTCA AGGAGAGCCT GTTGACAGG ATGCGGAACA  
 15410 15420 15430 15440 15450 15460 15470  
 CGCATCAATC GTCCGAGCGC AACCGGATGG CGCGCTCGGA GCGGGAGTTG CCGCCGGCTC CGCGTCGCCA  
 15480 15490 15500 15510 15520 15530 15540  
 GCAGGCGCGT TTCGAGGATC TGCCCGAGTA TAAGCAGGTC CTTACGCAGA AGATCGCCAG CGAGCAGCTC  
 15550 15560 15570 15580 15590 15600 15610  
 GGCATCGCCA ACCCGTTCTA CCGGCGCAT CAGACGGCAG CCGGTGCGAC GACGATGATC GATGGCCCGA  
 15620 15630 15640 15650 15660 15670 15680  
 AGCTGATCAA CTTCGCTCC TATGACTATC TCGGGCTCAA CCGCATGCT CACGTCCTCG AGCGGGCCGC  
 15690 15700 15710 15720 15730 15740 15750  
 CGAGACGATC GCCGGCTTCG GGATTTCGCG GTCGGCGAGC CGCCTCGTTG CCGGCGAACG GCCCGAGCAT  
 15760 15770 15780 15790 15800 15810 15820  
 GTGGAGCTCG AAGAAAAGAT CGCGCAATTC TACGGCGTTG ATGCCGCCGT CTGTTTCGTA AGCGGCTACC  
 15830 15840 15850 15860 15870 15880 15890  
 TGACGAATGT CGCCCGGATC AGCTGCCTGA TGGGGCCGAA GGACCTCGTC ATCCACGACG AGTTCATCCA  
 15900 15910 15920 15930 15940 15950 15960  
 CAACAGCGCG CTCGCGGGA TCAAGCTTTC AGGTGCGACA CGCCGCTTCT TCAAGCACA CGACACCGCG  
 15970 15980 15990 16000 16010 16020 16030  
 GATCTCGAGC ATGTCTGCG CACCGTCGCC GGTGACTACC GGCATCCTT GGTGATCGTC GAGGGCATCT  
 16040 16050 16060 16070 16080 16090 16100  
 ACTCGATGGA CGGCGACGTC GCCAATCTGC CGGCGCTTCT GAAGCTCAGG GCCGAATACG GCTTCTGGCT  
 16110 16120 16130 16140 16150 16160 16170  
 GATGGTCGAC GAGGCGCATT CGCTCGGGT CCTCGGCCG CACGGCAGG GCCTCGCCGA ACATTTCCGG  
 16180 16190 16200 16210 16220 16230 16240  
 GCCGATCCGC ACGAAGTCGA CATCTGGATG GGAACATGTT CGAAGACGAC GTCGAGCTGC GCGGCTACA  
 16250 16260 16270 16280 16290 16300 16310  
 TAGCAGGAAG CGCCGCCCTG GCGGCCGTG TAAAGGCATC TGCCGCGCGC TTCGTCTATA CGCTCGCCCT  
 16320 16330 16340 16350 16360 16370 16380  
 CGCGCCCGTG CTTCGCCGAT CCGCGGTCG CAGCCTCGAC ATCCTGGCAA GCGAACCAGA ACGCACGGCG  
 16390 16400 16410 16420 16430 16440 16450  
 GCGGTGAGAC GCAACGGCAG CCTGTTCCTG AAGCTTGCCA AGGAGGCCGG CCTCGATACC GGGCTGAGCG  
 16460 16470 16480 16490 16500 16510 16520  
 GCGGCTTTTC GGTCTGTGCT GTCATCTCG GGGATTCGCT GCGCGCCGTC CAGCTTTTCA ACGACCTGCT  
 16530 16540 16550 16560 16570 16580 16590  
 CGCCGCGGC AACGTCTGCG CGATCATCCA CCCGCGGTG CCGGAAGGAC AGGCGCGCCT TCGCGTCTT  
 16600 16610 16620 16630 16640 16650 16660  
 CATCACTGC GACCACACCG AGGAACAGAT CCGCGCACG GTCACATTGG AGCTGAACGC CTGAAGGATC  
 16670 16680 16690 16700 16710 16720 16730  
 TCACCAGCCG CAATTTCCGG CTCGGGCGCG TCGATATCGA GAAGATGCTG AGTGCCTCT CGGCATGCTG  
 16740 16750 16760 16770 16780 16790 16800  
 ACCACCAGCG GCCGCACGGC CGCCAACGGG TTCGTCCATG ACACATGTGA TCATACCAGG CGGCTCGAGC  
 16810 16820 16830 16840 16850 16860 16870  
 GGAATTGGCC TCGCGGTCG CTCGATCTAC GCGGCCCGCA GCCCGTCTC TCTGTGCGCA CGCTCGGCCG  
 16880 16890 16900 16910 16920 16930 16940  
 ATCTTCTGGA GAAGGCAGCA CAGAAGCTGG TAGCCGAACA CGGTCTCGAG GCCGGAGCGG TCCGCATAGA  
 16950 16960 16970 16980 16990 17000 17010  
 GGCCGCCGAT GTTTCCAAAG GAGAGGAGAT CGAGGCCGCG GTCTTCCGCT GTGTGGATGC GTTCGGACCC  
 17020 17030 17040 17050 17060 17070 17080  
 TGCGACGTTT TCGTGACCTC GGCAGGCGTC GTCGAACCGG CGCCATTGCA GCGATGACG GGATCGGCCT  
 17090 17100 17110 17120 17130 17140 17150  
 TCCACCAGCA GATGGAACG AACTTTTCGG GCACGGTGCA TGCCGTGCGC GCAGTTTATC CGGACATGAA  
 17160 17170 17180 17190 17200 17210 17220  
 GAACAGGCCG CGGCATATTC TGATGGTTTC TTCGGGTGCC GGCCTGATCG GCATCTACGG ATATACCGCC  
 17230 17240 17250 17260 17270 17280 17290  
 TACTGCGCCT CGAAGTTCGC GCTTAACGGC TTCGCCAGG GCTGCGCAGC GAACGCACGC CCAATGTC  
 17300 17310 17320 17330 17340 17350 17360

GGCATTTCGA TCTGCTTCA CCCGATACGG AGACCCGCA ATTCAAGAGG GAGCTTGCCG CCCGGCCAGT  
17370 17380 17390 17400 17410 17420 17430  
GGAGGCAAGC GTAATCATGG GAACCGTCCG CCCATGGACG GCCGAGGCCG TCGCGCGAAA GATCGTCAAG  
17440 17450 17460 17470 17480 17490 17500  
GGCATCGACC GCGCGCGCTT CGAGATTTAT TTTGGCGCGG CCTTGTATCT TCTGCCGCTT CGGGCGCGCG  
17510 17520 17530 17540 17550 17560 17570  
GTCAGGCCTT TCCTCAATTG GTGGTTCGAC CGGGCGATCG CGCGCAACAG TGGACTGTGA CCGCGCGGCC  
17580 17590 17600 17610 17620 17630 17640  
CGCGGCCATG CTTGTGCGAG CCAATATTAG GCTTTAAGTG TCGGGAGAAG CCTTCTGCCG AGGCGCGGAC  
17650 17660 17670 17680 17690 17700 17710  
GGTGAATGG CTCAAAGGTA TCGGCTTGGC GCTCATTCCC TTCAACCTGC ATGACTACCC GATAGCCTTG  
17720 17730 17740 17750 17760 17770 17780  
ACCCCTCGGT GCTACCTGCT GTCCTGCGCG GTGATCTTCT ATACGGACCG GTTCGCCCTT CCGCGCGCGC  
17790 17800 17810 17820 17830 17840 17850  
AACGCAGGAA GAACGCACCG AAATACGGGC GCAATCGGGA CAGGATCGAC CATCTCGCCC GGCTGCCGGT  
17860 17870 17880 17890 17900 17910 17920  
CATCGCGCTC GTCTTTGCCG GCTTCTTCGC GATCTCGTGG CGGCCCTCT ATGCCGCCGC CGGCACGATG  
17930 17940 17950 17960 17970 17980 17990  
AGTTTCTTCA TCATTTTCAC CGGCATTTCC CGCGGAAGT TCAAGTTCAT CCGTGAGCCG CTGGTCTTTT  
18000 18010 18020 18030 18040 18050 18060  
CCGATATCGC CCTGGTCGCG GATGTCTTCA AGTACAAGAC GATCTTCTAT GCGAGTTGCG TGAACATCGT  
18070 18080 18090 18100 18110 18120 18130  
CTTCTGGATC GTCGCCTTCC TTTACGTGTT CGCGTATCG GCACTTTACA TGTATTTTCA GCCGACGATC  
18140 18150 18160 18170 18180 18190 18200  
CTGCCGAAA GAAGCAGGCT CTTCTGGGTT CTGGTGATGG TGGGGATCGC CGCCGACCC TGGGGACTGT  
18210 18220 18230 18240 18250 18260 18270  
TGTTTTACGG GCCCGTGAAC CGCCCGACGG CCGCTCTCGT GCAGAGGCTG GTGAAGGCGA TCAACGTCOA  
18280 18290 18300 18310 18320 18330 18340  
GATGAACACG GTGCGTTTCG GTACCTTCGC TTCCGTCGTC TTCCATTTCA TCATCTGGCT CCGCGTCAAG  
18350 18360 18370 18380 18390 18400 18410  
CGCGACAAGA TCGTCGCCGA ATTTGTCGAA ATACTGCGCG CCGCGGTGCA CGACCTCADC GCCACGAGG  
18420 18430 18440 18450 18460 18470 18480  
AGGCCCGCT GATCATCGTA TGGCAGTCGG AGTCCTTTCAT CGACATGCCG CATTTCGGTG TCGATTCCGAT  
18490 18500 18510 18520 18530 18540 18550  
CAAGCTTCCG ACGGTCGACC GGTTCGCGAA GCAGCGGGTG CAATGGGGGC GCTTGAGCAA TGTCTTCGAG  
18560 18570 18580 18590 18600 18610 18620  
GGCGGCTACA CGCTGCGCAC CGAATTTGCC GTCTGAGCG GCCTTGTGCC CGACGACATC CACGTCGATG  
18630 18640 18650 18660 18670 18680 18690  
CCAGCTATCC TTATCTCCGT GCCGCGCACT ATGCCGATAT CGTCTGGCCC GAAAGCTCA AGCGCGGCTG  
18700 18710 18720 18730 18740 18750 18760  
GCGCAGCAT TTCATTCATC CCTATGACCG GACGTTCTTC CTGCGGCACA AGCGGATGCC CCTGCTCGGC  
18770 18780 18790 18800 18810 18820 18830  
TTCGAGAAGC TTACCATGCT CGATGCCTTC GACCATAAAC CGGAGCGGGA CCGGCTTTAT GTCTCCGACG  
18840 18850 18860 18870 18880 18890 18900  
CCACGCTGGC AGCCCGCGTG CTCGCCGAG TCGAGAAGCT GCCGGAAGAG GAGAGCGGCT TTTTCTTCGT  
18910 18920 18930 18940 18950 18960 18970  
CGCCTCGATG GCAAAATCAC GCCCTGGGA ACCGGGGCGC GTCGGTACGC TCACCAATCC CGTCGATATC  
18980 18990 19000 19010 19020 19030 19040  
TACCTGGCCA TCCTCGAGCA GTCTGACGCT GCGTTGAAGC AGCTGATCGA CGGCCTCAAC AAGCTCGACC  
19050 19060 19070 19080 19090 19100 19110  
GGCCGTCTG GTTCGTCTTT TACGGCGACC ACGGCCCTT TCTGAAATCT TTCGCGGATC CCTTCCCGGA  
19120 19130 19140 19150 19160 19170 19180  
TCCCCGCACG GATTATTTCA TCGTGCCGCT CGCCAAGGCC CGCGCCGCGG CCCACAGCCC GAAGCGGCGG  
19190 19200 19210 19220 19230 19240 19250  
AAGGACGAGG ATCCTTGGAA CCTGCTCGGC TCCATGCTGA AGCAGCCAA TCTGCACAAG GACGCGCTGC  
19260 19270 19280 19290 19300 19310 19320  
AATAGCGCTG ACTTCCGAA ATCCGCCGA TGACAGCAGT AACGACCATA GTACCCGCGC CGCCAAGAAC  
19330 19340 19350 19360 19370 19380 19390  
CTTCTGTTT TTGCAGGGC CCGGTCGCG CGTCTTCGCC AAGGTCGCCG GGCATCTCGA AGTGTGGGC  
19400 19410 19420 19430 19440 19450 19460  
CATTGCTGTC TGCGCATCAA CCTCAATCCC GCGCACCAGA TATTCTGGCG GCGCCGGGCG GCCATAACTA  
19470 19480 19490 19500 19510 19520 19530  
CCGTGGACTC GTTGGCGCGG TGGCCGGCCT ATGTGGAGGC CTTTATCCGC CGCCACGCGG TCACCGACCT  
19540 19550 19560 19570 19580 19590 19600  
CGTCTGCTC GCGCAAGAGC GCCCTATCA CCGGATCGCC ACAGCGGCCG CGCGAAAAGC CCGCGTCCGC  
19610 19620 19630 19640 19650 19660 19670  
GTCTTTGTGG TCGAAATGGG CTATCTCAGG CCGGACTGGT TGACGCTCGA GCGCGGCGGC ATGTGCTCCA  
19680 19690 19700 19710 19720 19730 19740  
ACTCGCACTT TCCGGTCGAT CCCGCGCAA TCCTGCGAGC GGCAATGGGA CTGCCCGAGC CCGACTGGCA  
19750 19760 19770 19780 19790 19800 19810  
ACGGCTCTAC CGGCAGAGCT TTCTTGCGGA GCGGCATAC GACCTCCTCT ACAACCTGCC GAACGTGTTT  
19820 19830 19840 19850 19860 19870 19880

TTCTGCTTCC TCTTTTCGCTA CCCGCGGCAC GGCATCTTTC ATCCGCTCGC CGAATATGCC GGGTGGATCC  
 19890 19900 19910 19920 19930 19940 19950  
 GGATCCGGCG GCTGGCGGCC GGCAAACGCG AGCAGCGGGC CGCGGATGCG TTGATCCGCT CGTTTTCCGC  
 19960 19970 19980 19990 20000 20010 20020  
 GAAAAATGCG CCGTACTTCG TCTATCCGCT TCAACTTGAA ACGGACTATC AATTGCGTGC GCATTACCCC  
 20030 20040 20050 20060 20070 20080 20090  
 TTCGGCAGCC AGCGGGAAGC GATCGACGGC ATCCTTGCCT CTTTCGCGCA CCGTGCGCCG GCAGGAACCA  
 20100 20110 20120 20130 20140 20150 20160  
 GGCTCGCCGT CAAGGTGCAC CCGCTCGACA ACAGCCTGAT CTCCTGGCGC AGGATCGTGC CGGACAAGGC  
 20170 20180 20190 20200 20210 20220 20230  
 CGCCTCACTC GGTGTGCGCG ACAGGTCCCT CTATCTCGAC GCGGCAATC TCGATTTCGCT CACGGAGCAC  
 20240 20250 20260 20270 20280 20290 20300  
 AGTTGCGGAA TGGTACCGT CAACTCGACC GCCGACTGC ATGCTCTGAA GCAGGGCAAG CCCGTGAAGG  
 20310 20320 20330 20340 20350 20360 20370  
 TCCTGGGAAA GGCCGTCTTC GACATTGCCG GTCTCACCGA CCAGCAGCCG CTCGACGGCT TCTGGACGGC  
 20380 20390 20400 20410 20420 20430 20440  
 GCCCGAGGCC CCGGATCCGG TGCTGAGCGG GGCAATGTTC AGGCTGATGG CCGCTTCGAT CCAGGTGCGC  
 20450 20460 20470 20480 20490 20500 20510  
 GGGAACTTCT ATTCCGTGCG GGTACGGATG CCGGCCGCAA GGCCATCGCC GCAGGCTGCA CCGCAACGAC  
 20520 20530 20540 20550 20560 20570 20580  
 GTCAATGAAC CGGGGCCCTT CGTCGACCCG CCGCCCCGGC CAAAAGCCGG TGAAACTCAG CGAAGATCAG  
 20590 20600 20610 20620 20630 20640 20650  
 GCGACGCCG CGCGCAGCAG ATCGTGAATG TGCAGGATGC CGACCGGGC GCCGGCCTCG TCGACGAGGA  
 20660 20670 20680 20690 20700 20710 20720  
 ACAGCACGGT GGACCTTGTG GTCCTGCATG AATTCATGG CGGCGCTGGC AAGGACGTCA CCCTTGATCA  
 20730 20740 20750 20760 20770 20780 20790  
 CCCGCGGGTT GCGCGACATC ACCTCCTGGA CCGGTTGTGC CAACAGGTCG CCCGCCATGT GACGGCGCAG  
 20800 20810 20820 20830 20840 20850 20860  
 GTCGCCGTCT GTGATGACGC CGATCAGCTT GCCGCTTCA TCTGTTATGC CGACGACGCC GAAGCCCTTG  
 20870 20880 20890 20900 20910 20920 20930  
 GCCGACATTT CGATGACCGC CTCGCTCATC GGACGGCCGA CGGAGAGTAA CGGCATCTGT CCGGTGCCAT  
 20940 20950 20960 20970 20980 20990 21000  
 GCGCCAGCTC CTGAACCAGC CGCAGCTGCG GCCTCAGCTT GCCGCCGGA TGGAAGGTCT TGAATCCTC  
 21010 21020 21030 21040 21050 21060 21070  
 GGCCGAGAAG CCGCGCCGCT CCAGGAGCGC GATCGCCAGA GCATCGCCGA CGGCAAGCTG CAGCATGGCC  
 21080 21090 21100 21110 21120 21130 21140  
 GAGGTTGTCT GCGCCAGGCC GTGCGGACAG GCCTCCGGCA CCTTCGGCAG CACGAGCGCG ATCTCGGAAT  
 21150 21160 21170 21180 21190 21200 21210  
 TCCGTGCGAG CGTGCTCTCG CGGTTGAGC AGATCGAGAC GATCGGCACC TTGAAACGCT TGGCGTAGGT  
 21220 21230 21240 21250 21260 21270 21280  
 CAGCATGTTG GCGAGTTCG CCGTTTCGCG CGACCAGGAG AGCAGCACCA GCGCATCCTG AGCGGTGATC  
 21290 21300 21310 21320 21330 21340 21350  
 ATGCCAGGT CGCCGTGGCT TGCCTCGGTC GGATGGACGA AATAGGCCGA GGTGCCGGTC GATGCGAGGG  
 21360 21370 21380 21390 21400 21410 21420  
 TGGCTGCAAT CTTGCGGCCG ATATGCCCGC TCTTGCCGAC CCCGGAAACG ACGACGCGGC CGTCGCCGTC  
 21430 21440 21450 21460 21470 21480 21490  
 GCCCATCAGT TCGACCGCAT CGACCAGGCC GCCCGGAAG TCCTGGTCAC TGGTCAGATG GTCTGCAAGG  
 21500 21510 21520 21530 21540 21550 21560  
 GCCTTGATGC CGTTCGTCG CGTCGTCAGC GTCCTGCCGA TCGACTCCAG CACGGTGCCA CCGCATGCG  
 21570 21580 21590 21600 21610 21620 21630  
 CATCCGCCGT CGCGTGTCTC ACACCATCA TCGACCCCTT TTAGGGCACC TCCGGGAAA CTGCGTCACG  
 21640 21650 21660 21670 21680 21690 21700  
 GTTTTCCGCC GGAAGTGCCT AGCCAAAAA CGCCTCGTCA GCGGGCCGCA GCAAGATAGA CGCCAGATAG  
 21710 21720 21730 21740 21750 21760 21770  
 GCTCTTTTCA TGGACAAACC AAGCCTTTTCG ACGGCCATTT ATACCAAACC GCGGACGTTT CGAACCCTCA  
 21780 21790 21800 21810 21820 21830 21840  
 TGCTCAGCCG ACGCGCTTTT CGACCAGGT CTCCGAGGA AGGATCCCGC CGGCGCGGAA TTCGGCGAGC  
 21850 21860 21870 21880 21890 21900 21910  
 GTCATCGAAT CGAGGATGGA GGACATCGCG TCACGCACC ATGTCATCGA GATGCGCACC CGGCAGGTCT  
 21920 21930 21940 21950 21960 21970 21980  
 CCGGATCGGT CGAGTCGTCG CACATTTTCA AAGCCGTGCG GCTGGCGCAG CGTATCGGCG CCAGCGGGCC  
 21990 22000 22010 22020 22030 22040 22050  
 GTCGAGCGTC CTTATGACAT GGCCGATGCG AATCTCGGAG GCCGGTCGCG ACAGGGAATA ACCACCGCCC  
 22060 22070 22080 22090 22100 22110 22120  
 GGTCCTTTCT TCGAGCGCAA CACGCCGGCG TTGCGCAGTT CCAGGAGTAT AGTGTCTAGA AACTTTTTTCG  
 22130 22140 22150 22160 22170 22180 22190  
 GAATGTTGTT GCGCTGTGCG ATTTCCGGTGA TGAAGGCAGT CTGCCGGGT GGAAGCCGTG CCAGGTGCGC  
 22200 22210 22220 22230 22240 22250 22260  
 CAATGCTTTC AGTCCGTATT TTCCCTTCTT CGTCAGCATT CGAGTCAATT CTCCCCGATC CCGCAGAAAC )  
 22270 22280 22290 22300 22310 22320 22330  
 GCCGGATCCG ACCTCGGAAA CAGTCTGTGC CCGCATAGC ACAAGAATGC GTCGCAATA CACTGATATA  
 22340 22350 22360 22370 22380 22390 22400

GTCGACAACG GAGAAGGATC GGTAACCGAG GATGGAAGAG TGCAGGAGCG CGTCAGGGTT TCAATGTATC  
 22410 22420 22430 22440 22450 22460 22470  
 ACTGACGCGC CCCAGAGAAT TTCTTGTCCT GGAAGAGAAT TTCTTGTCCT GGAACGGTTC GCACGTCACG  
 22480 22490 22500 22510 22520 22530 22540  
 CGGCGGCAAT GCCGGGCGTC CGATTGGTGA GGTTCGGCCC GCCCTGGGCG GAAAGGAGCC GGCCAATGGC  
 22550 22560 22570 22580 22590 22600 22610  
 CGCGCGTGCT TCATCGACCG AACGGAAGCG TTGCGCGTCC AGGTCCCAGA CGTGATACTT GACAGCCAGG  
 22620 22630 22640 22650 22660 22670 22680  
 AACTTCAGAC GGTCTTGGTC GGAACGACG ACTCCGACGG CTTACCACC ATATTCGATA ACTTGCTTGC  
 22690 22700 22710 22720 22730 22740 22750  
 TCATGTAAAT AACGTCCTC CGCACCCACG GGCGCGGATA ATCAATTCT CGGCAAAGGA TGAACCTGGT  
 22760 22770 22780 22790 22800 22810 22820  
 AAGTTTGAAC GGTACATTC GCGAAGTTG ACGAATGACA CCTTTAACCG TTGCACGACA ATGGGAACGT  
 22830 22840 22850 22860 22870 22880 22890  
 AGCCGCAGAA AGCGGCGGTC GAGAATTCGA ACATGTCACT CTCCTTTGGC TGGCCCCGCA TAGGACGCCA  
 22900 22910 22920 22930 22940 22950 22960  
 GGCGCTCATG AACAAATCGA AGCGGACTCC TTCCGCTTCC CACAGTGGCA AACTAAAATA ATCTACCAAT  
 22970 22980 22990 23000 23010 23020 23030  
 TTCGTCAACT GCAGAAACGC AAACGCGAAA AATAATTCTC ACTACTCGTT TCCGCAGCTG GCGCGTCGGC  
 23040 23050 23060 23070 23080 23090 23100  
 AAATCCGTTT GCCTGTCCGG GCTATCTCGC TATCTAGGTA GCGCCGAATC GCCGTGCAAG GGACGGGTCG  
 23110 23120 23130 23140 23150 23160 23170  
 AATATACGAG CAGGATTTTC CGACGAAAAA GGGGCGATAC CTCGTAGAGG CTGCCCCTTT GCCGCGATCG  
 23180 23190 23200 23210 23220 23230 23240  
 ACACGGTGGG ATCAGGAAAG GCGCGGACGG TTTTCCGCCC GCATCCGTCA ACGGCGCGAG CCCGTCAGGC  
 23250 23260 23270 23280 23290 23300 23310  
 CGTCGGCGTT FCCCGCTGCT GCCCGGCGGA ACGAGCGGCT TCCCGCTCCT TCCACCATGC GGCTCAGGAA  
 23320 23330 23340 23350 23360 23370 23380  
 CAGGAGGATC GGACCGCCGA TATAGATGGA CGAGGTCGTC GCGACGATGA CGCCGAAGAT CATCGGCCAG  
 23390 23400 23410 23420 23430 23440 23450  
 GCGAAGTCTT GCACCGCCTC GCCGCCCCAG ATCGCCATCG GCACAAGGGA AAGCGCCGTT GCACCGGAGG  
 23460 23470 23480 23490 23500 23510 23520  
 TGAAGATGCA TTGCGCGATC ACCTGGTTGA TGCTCATGTC GATGAGATCG GAGAACGGCA TCGACTTGTA  
 23530 23540 23550 23560 23570 23580 23590  
 CTGCGCGAGG TTTTCCGCGA TGCGGTCTTA AACCACCACC TTGTCGTTCA CCGAATAGCC GATCATCGTC  
 23600 23610 23620 23630 23640 23650 23660  
 AGCAGCGCTG CGATCGCGGT CAGGTTGAAG TCGATGCCCG TGAGTGCAA GAACCCGACG GTCTTGGTGA  
 23670 23680 23690 23700 23710 23720 23730  
 TGTCGAGCAA GAGCACGGC ATCGCGCCGA CGGCGAAGTG CCATTCGAAG CGGAACCAGA TGTAGAGCAG  
 23740 23750 23760 23770 23780 23790 23800  
 GATCGCCAAC ATGGCGAGGG CGACGCGGAG GAAGCCGGAT CGCGCCAGTT CGCCACTGAC CGTCGGACCG  
 23810 23820 23830 23840 23850 23860 23870  
 ACGACCTCGG TCCGTTCCAT GCTGGCACCG GGGATGGGCG TCGTCACCCG ATCCTTGACC TTGTTCACTG  
 23880 23890 23900 23910 23920 23930 23940  
 CCACCGTCTG CGCCTGCTCG CCGCCCGGCT GCCGCTGCAC GCGGATCAGC ACCGATTGGC CGCCGCCGAA  
 23950 23960 23970 23980 23990 24000 24010  
 GTCCTGGAGC GCCACCTCGC CGAGATTGAG CTCTCGAGG CTTTGACGCA GTGTCGGCAG ATTGATCTTC  
 24020 24030 24040 24050 24060 24070 24080  
 TCCTTCGAAA CCGCCTCGAC CTGAATGCCA CCGATGAAAT CGATGCCGTA GTTGAAGACC GGCGTGAAGA  
 24090 24100 24110 24120 24130 24140 24150  
 ACAGGATGAC GGAGCTGATC GACAGGAATG CCGAAAAGCC GATCGCGAGG AAGCGGCCCT TCATGAAGGA  
 24160 24170 24180 24190 24200 24210 24220  
 GAAGTGGGC AGCTGCGGCA CCTTGCCAAA AAGCGACGGG ATTTCAGCT TCTTCATCTT GCGGCGCACG  
 24230 24240 24250 24260 24270 24280 24290  
 ACGACTTCGC GCATCAGCAG GCGCACGACC GTAACCGATG TGAACATCGA AATGCCGATG CCCAGCATCA  
 24300 24310 24320 24330 24340 24350 24360  
 TGGTGACGGC AAAGCCCTTC ACCGGCCCGC TGCCGAACCA GAAGAGAAGC ACCGTGCCCG CCAAGGTCGT  
 24370 24380 24390 24400 24410 24420 24430  
 GACGTTGAAG TCGATGATCG TCGCATAAGC CTTGTTGAAG CCGATATCCA ACGCTTTCAT CGCGCCGGCG  
 24440 24450 24460 24470 24480 24490 24500  
 CCCGCCTCGG TCTCTTCGCG AATACGCGCA TTGATCAGGA TGTTCGCGTC CACGGCCATG CCGATGCCGA  
 24510 24520 24530 24540 24550 24560 24570  
 GAATGATACC GGCATACCG GGCAAGGTCA GCGTCAAGCC TATGAGGCCG AGTACGCCGA TCGTCATGAT  
 24580 24590 24600 24610 24620 24630 24640  
 CGTGTGCAAC ACCAGGCCGA CATTGGCGAT CATACCCCAG GCCCGTAGA GGACGACCAT GAGGACCACG  
 24650 24660 24670 24680 24690 24700 24710  
 ACGAGGCCGA AACCGCGGAG ACCGGTAAAG AGGCCCATGC GAATCGAGTC GCTGCCGAGA TTCGGGCCGA  
 24720 24730 24740 24750 24760 24770 24780  
 CCGAGCGCTC TTCGATGATC GTCAGCGGTG CCGGCAGCGC ACCGGAGCGC AGGAGTGGCG AAAGAACCCT  
 24790 24800 24810 24820 24830 24840 24850  
 CGCCTCCTCG ACCGTGAAGC TGCCGCTGAT CTGGCCGCGG CCACCGAGAA TGGGCTCGTT GATGACCGGC  
 24860 24870 24880 24890 24900 24910 24920

GCCGTCAGCA CCTTGCCGTC AAGCACGATC GCAAACGGCC GGCCGACGTT TTAGCGTGTG ATGTCGGCGA  
 24930 24940 24950 24960 24970 24980 24990  
 ACTGGCGGGC ACCGAGCGAG TCGAAGCTGA AATCGACGAC CGGCTGGTTG GTGCGCTGAT CGAAGCCGAC  
 25000 25010 25020 25030 25040 25050 25060  
 CTTGGCGTCG TCCAGCCGCT CGCCGAAAT GGCAGCGCG CTTTCGACCG GATAGCGGTT GCCGTCATTG  
 25070 25080 25090 25100 25110 25120 25130  
 GCGCCCGGCA GAATGTCGAC CCCACGCGGC GCCGGCTGTG TGACGTCGAC GCTCTGATCG AGCATGTGGA  
 25140 25150 25160 25170 25180 25190 25200  
 AGCTCATCTG TGCCGTCGAA CCCAGAAGCT TGGCAAGGG GTTCGGATTC CTTAAAGGCC GGGCAAGCTG  
 25210 25220 25230 25240 25250 25260 25270  
 AACCAGAATG CGGGTTGGAA CCGATGCGTT GAATCAGCGG TTCGGGGACG CCGACCTGGT CAACGCGGTT  
 25280 25290 25300 25310 25320 25330 25340  
 GCGGGATAAT TTCCAGGCTC TGCTCGACCG CTTTGGTTCAT GCGGTCGGCA AAGACCCGCC TCGGTCATCG  
 25350 25360 25370 25380 25390 25400 25410  
 CCAGGGTGAT GACATCGCCC GCCCGTCGTT ACCTCGACCT CAGGAGCCGG GCGGGCGAAA CCCACCGTGC  
 25420 25430 25440 25450 25460 25470 25480  
 TCACGGGCGT GCACAGTTCC TGCAGCTTCG GCAGAACCTT GTCCCGGTCC TGAGCATCGG GGATGCTTAC  
 25490 25500 25510 25520 25530 25540 25550  
 GGTACGCTG TTGCCGCTGA TGCGCGCCGA GGAAGCGGAA ACGCGCTCGC CGCGCAGGAC GCGACGGGTA  
 25560 25570 25580 25590 25600 25610 25620  
 TCGTCAGAAA GCGAGTTGAG GCGAGCCTTC TGCAGCCCGC CCCCGTCCAC CTCGAGAACG AGGTGCGAAC  
 25630 25640 25650 25660 25670 25680 25690  
 CGCCCTCAG ATCGAGCCG AGCGTGACCG GCTCGAACGG GAGAACGGAA GCGACCTGCT CCCTCATGGC  
 25700 25710 25720 25730 25740 25750 25760  
 CGGCGTCAGA ACGCTCGGCA GCGCCGCAAG GCATCCGAAG ATGATGATCG CGACATAGGC GAGGACCGCC  
 25770 25780 25790 25800 25810 25820 25830  
 CATCTGGATG TACGCATGTG GGAATFCCAT AAACGCCTGA ACGCGCCTTG GGCAGCGTCC GGGCTGGGTT  
 25840 25850 25860 25870 25880 25890 25900  
 TTCATCAATG GCTGAGGCGG GATCGAAAGC CCTGCCGAA TCAAACGATG CTGCCGCCGG TGCGGCGCGC  
 25910 25920 25930 25940 25950 25960 25970  
 GGATGCACTG GTTCTGGCAC GGCAGCCAC GTGCGATTTT GGTCTCGCGC AGAGATGGAG AAGGAGAAGC  
 25980 25990 26000 26010 26020 26030 26040  
 CACGGGCGCA GCGAGCGCCA CGCGTTGAAA TTGGATGCAT GGCGCCGGAT CGCCGCCGGG CATCGACTTG  
 26050 26060 26070 26080 26090 26100 26110  
 CCGTCAGCGC GCTCGCCGCT GAAACGGAGG TCGGGCAGTG CAACCGCCCT GCAAATCTGG CGGGCGACCG  
 26120 26130 26140 26150 26160 26170 26180  
 GCCGGTCGGA GGAACCTGTA TCGGGTTGAG CGACATTGCC GGCCGGCGCT CCGGCCGCGC GGCCTGTGTC  
 26190 26200 26210 26220 26230 26240 26250  
 ACCCTGGCCG ACCGCGATGA CCTGCATGGA AAGAGAGGCG CAGAACAGCC ACAGCACGGC GAGAAGCCGG  
 26260 26270 26280 26290 26300 26310 26320  
 CTGGCGTGCA TGCTCTGCT GCCACCACC ATGAGGTCCA TCGAGGACGC GGTCCCTTTC TCTTTGCATC  
 26330 26340 26350 26360 26370 26380 26390  
 CGGTCAATCC GCGATCGAAC GCAAGCGAGA CGAATTCAGC CCGAGCACCA AACTAAACGA AGAACGTGGC  
 26400 26410 26420 26430 26440 26450 26460  
 CGGATTCGAG CATGTGCGCT CTTAACATCA CTCCGGGCCG GAAAAGTCAA ATCCGGGCCG TCATGCAGTG  
 26470 26480 26490 26500 26510 26520 26530  
 GCGCCGGCGG TTCTTTCTTG ATGGATTTTC ATGAGCGAAG GCCCCGGGCG CGACGCCGCA ACGGCGAGCG  
 26540 26550 26560 26570 26580 26590 26600  
 CGCGCGCAG ATCCCTTACG GTCGCGAGC GCTCAGCGG CAGGCCGTAG GCTTCGGCGG TCAGCAGGAA  
 26610 26620 26630 26640 26650 26660 26670  
 ATCGGGTGGG GAAGGTTTGA CGCCCTCCGG CGTGATCCCG GATTGACCA TGTAGGATTC GATTTCCCGG  
 26680 26690 26700 26710 26720 26730 26740  
 TATCCGTCAT TGTTCAGAC GAGGAAGACG ACCCTGGCAT TCGCGTCCAC CGCCGAGCCG ATCTCGGCCA  
 26750 26760 26770 26780 26790 26800 26810  
 GTGAAAATG GAACCCGCCA TCGCCGACGA GGCATATGAC CGGCCGCCCG CGATCGGCGA CCGCAGCGCC  
 26820 26830 26840 26850 26860 26870 26880  
 CACTGCAAGC CGGAGGCGCA TAGCCAGCG AGCCGTAGCC CGTCGCCGAA TTGAACCAGG CCCTTTGCCG  
 26890 26900 26910 26920 26930 26940 26950  
 CGGCGCGTCG CAATAAAGGT TGCCGGCATA GACGGCCTGG GTCGAGTCCC CGACGATGGT GCAGTCGGGG  
 26960 26970 26980 26990 27000 27010 27020  
 AGCGTCCGGT AGATCGCCTC GACGACGCC ATTTCCGCAA GCATTTTGTG GTCAGCTCCT TAAGCGCCCG  
 27030 27040 27050 27060 27070 27080 27090  
 TTGCGCGCGG CCTTCGCCCG TAAGCTTCCA TCCTTGGCCG TCGCATGGTT CGGCAGGAAG CCGAGAATGC  
 27100 27110 27120 27130 27140 27150 27160  
 CCGCAGTCTC AGCCTTGGCG CGGAGAGGAT CGAAAGCGCA GCTTGGGGC CACGCGCGAG CTGCGCCGCA  
 27170 27180 27190 27200 27210 27220 27230  
 TCGATGTCGG TGCATAAAG GTTCTTTCAGC AGCGGAAAGC CGCCGTCCGC ATAGAGATCG TAATCGGTCT  
 27240 27250 27260 27270 27280 27290 27300  
 GGCCATCTC GGTGCCGAGC GCCAGCACGA GATCCGCCTC CCGCAGCAGC GCGCGGACCG CCTTGAGGCT  
 27310 27320 27330 27340 27350 27360 27370  
 CGGGCTTGGC GGCACCCGGA GCGGATGGCC GGCAAGCATG CCGCGGGCGT TGACCGTGTG GACCACCGGC  
 27380 27390 27400 27410 27420 27430 27440

GCGCCGATCC GCTCGGCGAG TTCGCGCACT TCCGCTTCCG CGGTCAGGGC GCCGCCGCCG CACAGGATGA  
27450 27460 27470 27480 27490 27500 27510  
CCGGCCGCGC GGCTTCCGCG CAGAGGATCG CCGCCTTCTG CAGCGTTTTCG CTGTCCGGAGC GCGGACGTGT  
27520 27530 27540 27550 27560 27570 27580  
CGGGCCGCCG GGCTTTGCGG GCCGGCCCTC GATCTTCAGG GCCATCACGT CGGTCGGAAT TTCGATATGG  
27590 27600 27610 27620 27630 27640 27650  
ACCGGGCCTG GCCGGCCGGA AAGCAGAACG GCAAAGGCGC GGTTCGACGAC GAGCGGCAGA TCGGGGGGT  
27660 27670 27680 27690 27700 27710 27720  
TGAGAAGCGT GTGCGAATAG AGCGCCAGCG TCTTCATCAT TCCATTCTGA TCGGGCAGCT CGTGCAGCAG  
27730 27740 27750 27760 27770 27780 27790  
CCCCCGCCA TGCCCGAGGG AATCCCGCGC GTTGACGCGG GAGATCACCA GCATCGGGAT CGAGTCTTGC  
27800 27810 27820 27830 27840 27850 27860  
CGGGCTTGGC CCATGGCGGT GATCGTGTG GTAAGGCCCG GACCGGTGAT CACGAGCACC ACGCCGGGCT  
27870 27880 27890 27900 27910 27920 27930  
TGCCCGTGAC GCGCGCATAA CCGTCGGCCA TGAAGCCGGC ACCCTGTTCG TGGCGCGGAG TGACGTGGCG  
27940 27950 27960 27970 27980 27990 28000  
AATTTTCGAT GTGGCAAGGC CGCGGTAAAG TTCGACCGTG TGCACGCCCG GAATGCCGAA TACGACCTCC  
28010 28020 28030 28040 28050 28060 28070  
ACGCCGTTGG CTTCCAGGAG ATCGACGAGG GTCTCGCCGA CGGTTTTCGA TTCGACGGTC ATGCGCGGCG  
28080 28090 28100 28110 28120 28130 28140  
CTCCTTGGCG ATGGTGCAGC GTTCCGCGAG CGCGCGCATG CGGCGGCATG CCTCCTCGAT GCGGCGATCC  
28150 28160 28170 28180 28190 28200 28210  
GGAACGGTCA GGCTCACGCG CAGGAAGTTG CGTGCCCTGT CGCCGAAGGA GGACCCCGGC ATCACCGCAA  
28220 28230 28240 28250 28260 28270 28280  
CGCCCTCCTC TTCCAGGAGT GCCCAGGCGA AGGCTTCGCC GGAGAGGCC GTGCCGAGA CGTCGATCAG  
28290 28300 28310 28320 28330 28340 28350  
CGAAACATA CCGGCTCCG GCGGCAGCAC CGAAACCCCG GACGCCCCCG CAAGACCGTC GATGATCCGC  
28360 28370 28380 28390 28400 28410 28420  
CGGGCGCGCC GGCTGTAGGA TTCCCGCATC TGCCTCGCCG TATCGATGTC ATGCGTCAGC GCGTAAGCGG  
28430 28440 28450 28460 28470 28480 28490  
TCATATCGGC GATGAACGGC TGCTGGCCGA AAAGCATGGT CTCGGAATC GGCAGCAGC GCTCGGTGAA  
28500 28510 28520 28530 28540 28550 28560  
CTCGGCGGG CCGACGGCCC AGCCGCTGCG AAAGCCGGGT GCTGCATGCG ACTTCGAGAT GGACGAGACG  
28570 28580 28590 28600 28610 28620 28630  
ACGACCGTGC GTTCGCAAG ATCCGGATTG TCGAATGGCG AGGCGAAGAG CGCATCGAAG ACGAGCTCCT  
28640 28650 28660 28670 28680 28690 28700  
CATATACCTC GTCGCAGACG ATCCAGAGAT CGTGCCGCCG CGCGACTTCA CCGATCGCCG CAATTTCTTC  
28710 28720 28730 28740 28750 28760 28770  
CGCCGTCAA ACACGCGCGT GGGGTGTGCG GCGTATTAAG AACAGCACGC GGCATTCGGG CGTGACCGCC  
28780 28790 28800 28810 28820 28830 28840  
TTTTCCAGAT CCTCCGCCG CATGTGAAAG CCGTTTTCG GGTTCAGCGG TACGAATACC GGATGGGCC  
28850 28860 28870 28880 28890 28900 28910  
CCGTGGACCG GATAGCCCT TCATAGGTCG CGTAGAGGGG GTCCTCCGAG AGCACGCCGT CTTCCGCCTC  
28920 28930 28940 28950 28960 28970 28980  
GGCAAGCGCG AACATGACGG CGAAGAGCGC CGTCTGCGTT CCCGGGAAGC AAAGAATGTT CTCGCGGTC  
28990 29000 29010 29020 29030 29040 29050  
ACGGCGGCAC GCGGCGCAGC GTACTTTTCG GTAAGCGCCG CGACGACCGC CGGCTCCCCG CGGCCGTTG  
29060 29070 29080 29090 29100 29110 29120  
AATAGCGGTA CCGTCCCGCA TTCATGGCCC GCTGGCACTC TTCCAGAAGC GCGCGGTCCG GCGGCAGATC  
29130 29140 29150 29160 29170 29180 29190  
CGGTTCCGCC ATGGTCACTC CGATGAGCTC GCGCCACTC GCCTTCAACT GACGTGCGCG GGTGTGCAGC  
29200 29210 29220 29230 29240 29250 29260  
GTCCATTGTC CGGAACCGAG CTCGCAAGA CGTGATGTGA TCGACGCGTA GCGCAAGAAT TCACTCCCTC  
29270 29280 29290 29300 29310 29320 29330  
ATTGATTTT TATTTTTCGA TCCGAAGGCC GAGCAGGGCT TCGACCGAGG CCATGGCGAT GCCGACCAAGT  
29340 29350 29360 29370 29380 29390 29400  
TCGCCCTCGC GAAAAAGATC GCCGGCGAGG CAGCCTTCGA CCCAGAGCCC GTCCACGAGG CCGTTGATCG  
29410 29420 29430 29440 29450 29460 29470  
CGATTGCGTA GCGGCGGCAC TCCTCGGGAC CGGCAGGCCG GCCCTTGCC GAAAGAAAAT CGTGAGCAA  
29480 29490 29500 29510 29520 29530 29540  
CTCCTGCAAC GCGTTGCGGA AGGCGAGGTA GCCGTCGCGA TGGATTTCG CGAGACCGG ATCGACCGG  
29550 29560 29570 29580 29590 29600 29610  
ACCTGACTGA TGAAGGCGGC CCAGAGAGAG ATGCTGCGGC TGTCGGCGAC CGGTTCCGTC AGATTGATGG  
29620 29630 29640 29650 29660 29670 29680  
CGATGAAGTC CTTGAGACGC GTCGCGGGAT CGCCCTCGAC GTTCTTCGCC TTCTCCGTA GAGAGGCAAT  
29690 29700 29710 29720 29730 29740 29750  
GACGGCGCG TAAGCCTCCG CAACCATCTG GTCTTTGGAT TCGAAATAAT GCGGACGAG CCCTTGCGGT  
29760 29770 29780 29790 29800 29810 29820  
AACGCCCGGC GCGAACGGC ATCTGGCGGA CGGTCGCTCC CTTGAGGCCG AATTC  
29830 29840 29850 29860 29870

## Restriction Enzyme Report for pPP428

page 1

Length of pPP428: 29875 bp, +1 at: 1; Simple Restriction from: 1 to: 29875;  
 Mon, 2. Aug 2004 7:20 Uhr

ENZYME	SITE	(fragment length) position			
AgeI	A'CCGGT	( 6499)	6499	( 3948)	10447
		( 2164)	12611	( 5216)	17827
		( 741)	18568	( 2252)	20820
		( 3911)	24731	( 3171)	27902
		( 1757)	29659	( 216)	
BamHI	G'GATCC	( 4018)	4018	( 5152)	9170
		( 5694)	14864	( 4302)	19166
		( 12)	19178	( 81)	19259
		( 686)	19945	( 6)	19951
		( 502)	20453	( 1429)	21882
( 452)	22334	( 7541)			
BclI	T'GATCA	( 1101)	1101	( 2307)	3408
		( 1686)	5094	( 9403)	14497
		( 1187)	15684	( 1164)	16848
		( 1642)	18490	( 2295)	20785
		( 561)	21346	( 3186)	24532
( 3375)	27907	( 1968)			
BglIII	A'GATCT	( 4950)	4950	( 5332)	10282
		( 2170)	12452	( 17423)	
ClaI	AT'CGAT	( 3538)	3538	( 1892)	5430
		( 507)	5937	( 72)	6009
		( 52)	6061	( 1097)	7158
		( 4357)	11515	( 4154)	15669
		( 8451)	24120	( 3111)	27231
( 1232)	28463	( 1412)			
EcoRI	G'AATTC	( 1)	1	( 1332)	1333
		( 1419)	2752	( 3602)	6354
		( 5838)	12192	( 207)	12399
		( 524)	12923	( 7827)	20750
		( 457)	21207	( 691)	21898
		( 1015)	22913	( 2939)	25852
		( 570)	26422	( 2895)	29317
( 553)	29870	( 5)			
EcoRV	GAT'ATC	( 5235)	5235	( 74)	5309
		( 1286)	6595	( 819)	7414
		( 126)	7540	( 1849)	9389
		( 45)	9434	( 165)	9599
		( 507)	10106	( 9)	10115
		( 2088)	12203	( 4562)	16765
		( 1300)	18065	( 663)	18728
		( 309)	19037	( 5438)	24475
( 5400)					
FspI	TGC'GCA	( 1148)	1148	( 147)	1295
		( 1102)	2397	( 1786)	4183
		( 346)	4529	( 154)	4683
		( 714)	5397	( 6003)	11400
		( 610)	12010	( 2959)	14969
		( 1080)	16049	( 1286)	17335
		( 1241)	18576	( 60)	18636
		( 837)	19473	( 607)	20080
		( 1549)	21629	( 336)	21965
		( 189)	22154	( 1441)	23595
( 6280)					
HindIII	A'AGCTT	( 4174)	4174	( 79)	4253

## Restriction Enzyme Report for pPP428

page 2

		( 902)	5155	( 10172)	15327
		( 656)	15983	( 498)	16481
		( 2071)	18552	( 285)	18837
		( 6389)	25226	( 1886)	27112
		( 2763)			
KpnI	GGTAC'C	( 3887)	3887	( 4937)	8824
		( 9540)	18364	( 11511)	
NcoI	C'CATGG	( 2415)	2415	( 6401)	8816
		( 899)	9715	( 2988)	12703
		( 609)	13312	( 4150)	17462
		( 3293)	20755	( 7116)	27871
		( 1510)	29381	( 494)	
NdeI	CA'TATG	( 3899)	3899	( 5447)	9346
		( 17498)	26844	( 3031)	
NotI	GC'GGCCGC	( 945)	945	( 3526)	4471
		( 4983)	9454	( 3717)	13171
		( 6474)	19645	( 10230)	
PstI	CTGCA'G	( 2104)	2104	( 1072)	3176
		( 1279)	4455	( 692)	5147
		( 7634)	12781	( 8351)	21132
		( 1911)	23043	( 2461)	25504
		( 150)	25654	( 1898)	27552
		( 2323)			
SacI	GAGCT'C	( 1286)	1286	( 1223)	2509
		( 3149)	5658	( 6118)	11776
		( 1802)	13578	( 2250)	15828
		( 8214)	24042	( 4655)	28697
		( 522)	29219	( 63)	29282
		( 593)			
SmaI	CCC'GGG	( 5357)	5357	( 2582)	7939
		( 2751)	10690	( 946)	11636
		( 889)	12525	( 359)	12884
		( 9352)	22236	( 4339)	26575
		( 2448)	29023	( 852)	
SphI	GCATG'C	( 1300)	1300	( 1764)	3064
		( 210)	3274	( 1977)	5251
		( 145)	5396	( 846)	6242
		( 1688)	7930	( 616)	8546
		( 1725)	10271	( 628)	10899
		( 4088)	14987	( 741)	15728
		( 1069)	16797	( 395)	17192
		( 3151)	20343	( 1285)	21628
		( 1984)	23612	( 2720)	26332
		( 1078)	27410	( 780)	28190
		( 418)	28608	( 1267)	
XbaI	T'CTAGA	( 22175)	22175	( 7700)	

## 7.5 Vergleich von RkpABCDEF von *Sinorhizobium meliloti* AK631 mit RkpA aus dem sequenzierten Genom von *S. meliloti* 1021

### 7.5.1 Proteinalignment

```

RkpA1021      MTVEIIGRACLAPGAKSPQALFKILRQKCTVTRVPSDRWDLARFWHPVMGTPGKTYSF
RkpAK631      MTVEIIGRACLAPGAKSPQALFKILRQKCTVTRVPSDRWDLARFWHPVMGTPGKTYSF
1.....10.....20.....30.....40.....50.....

RkpA1021      AGVL DHIYDFDPAVFGMSQREAMYMDPQQRVLLQLAWRALEDANISVASLHGENVGVYV
RkpAK631      AGVL DHIYDFDPAVFGMSQREAMYMDPQQRVLLQLAWRALEDANISVASLHGENVGVYV
61.....70.....80.....90.....100.....110.....

RkpA1021      ASLDHANLTVDDPAAAGPYFMTGNTLSIVSNRISHVFGLSGSPSMTVDTACSSSLVALDQ
RkpAK631      ASLDHANLTVDDPAAAGPYFMTGNTLSIVSNRISHVFGLSGSPSMTVDTACSSSLVALDQ
121.....130.....140.....150.....160.....170.....

RkpA1021      AMRALNAGEIDTAIVGGVNI LAHPLPFVGFQAARML SPEGLC RAYDNDGAGYVRAEGGV
RkpAK631      AMRALNAGEIDTAIVGGVNI LAHPLPFVGFQAARML SPEGLC RAYDNDGAGYVRAEGGV
181.....190.....200.....210.....220.....230.....

RkpA1021      FVLRRTDRAQREDRSYARIVASGVNSAGRTNGISLPSREAOANLLRMIYEGNGIDANQV
RkpAK631      FVLRRTDRAQREDRSYARIVASGVNSAGRTNGISLPSREAOANLLRMIYEGNGIDANQV
241.....250.....260.....270.....280.....290.....

RkpA1021      AFVEGHGTGTVKVGDPAEVWSIGTVIGAKRRAPVPIGSIKSNIGHTEPASGLFGMMKAVLA
RkpAK631      AFVEGHGTGTVKVGDPAEVWSIGTVIGAKRRAPVPIGSIKSNIGHTEPASGLFGMMKAVLA
301.....310.....320.....330.....340.....350.....

RkpA1021      LENNYLPASLHFETPNEQIDFDGLNVRVTANPIELLKGRARLAGINSFGFGGANAHVVI
RkpAK631      LENNYLPASLHFETPNEQIDFDGLNVRVTANPIELLKGRARLAGINSFGFGGANAHVVI
361.....370.....380.....390.....400.....410.....

RkpA1021      SDPDPAAQAEKAATSPAGHVFLASAHTASSLENLLRDYKAAAFAGATKEETRAIVAASGANR
RkpAK631      SDPDPAAQAEKAATSPAGHVFLASAHTASSLENLLERLAAA-----
421.....430.....440.....450.....460.....470.....

RkpA1021      THMRHRFAARSDHPEDIVRAIANHLEKPGSDIGETGEEAAAKEAKVAFVFSNGSQWAGMG
RkpAK631      --MRHRFAARSDHPEDIVRAIANHLEKPGSDIGETGEEAAAKEAKVAFVFSNGSQWAGMG
481.....490.....500.....510.....520.....530.....

RkpA1021      VEFRENLFHQSFSTSVSALFRFHSDIVLTDLLTDELDLDRKLADTKVAQPLLFAIQAA
RkpAK631      VEFRENLFHQSFSTSVSALFRFHSDIVLTDLLTDEGLDRKLADTKVAQPLLFAIQAA
541.....550.....560.....570.....580.....590.....

RkpA1021      DSVLVAMGKPTAVFGHVSVEIAAAYAAGALSLVDVAVSIVAKRSLHQDLLAGQGTMAAVML
RkpAK631      DSVLVAMGKPTAVFGHVSVEIAAAYAAGALSLVDVAVSIVAKRSLHQDLLAGQGTMAAVML
601.....610.....620.....630.....640.....650.....

RkpA1021      GEEAAKAFAAERGLDDVCVAIAINAHNSVTISGPANEISAFRDAARKAKIPVQILDINYPF
RkpAK631      GEEAAKAFAAERGLDDVCVAIAINAHNSVTISGPANEISAFRDAARKAKIPVQILDINYPF
661.....670.....680.....690.....700.....710.....

RkpA1021      HHPIIDRAKEAFLSDIPDIAPRRTLAYLSVTVGAAIDGTALDPDYWWKNVREPVRFOAA
RkpAK631      HHPIIDRAKEAFLSDIPDIAPRRTLAYLSVTVGAAIDGTALDPDYWWKNVREPVRFOAA
721.....730.....740.....750.....760.....770.....

RkpA1021      AEAAIDLGCMLFIEISPRPILGSYVKETIKQAAVPA SVVATLLRDAGANGHDPISASMAR
RkpAK631      AEAAIDLGCMLFIEISPRPILGSYVKETIKQAAVPA SVVATLLRDAGANGHDPISASMAR
781.....790.....800.....810.....820.....830.....

RkpA1021      AVAHGAAVDRSRVYGKRDAFIELPALPFEPVELRPAATDATDLFGRSAKPYRLNGWRGD
RkpAK631      AVAHGAAVDRSRVYGKRDAFIELPALPFEPVELRPAATDATDLFGRSAKPYRLNGWRGD
841.....850.....860.....870.....880.....890.....

RkpA1021      PNAGSWKNHVD AHLFPDLAEHVVDGKAILPGSGFIEIAVSAAQQYHGSDEVEITNLEIVR
RkpAK631      PNAGSWKNHVD AHLFPDLAEHVVDGKAILPGSGFIEIAVSAAQQYHGSDEVEITNLEIVR
901.....910.....920.....930.....940.....950.....

RkpA1021      PLELSDSRIMELSTILSPETGDIEIRSRLS EDDWTVHAVARSRKPLPTTSNDPSTKG
RkpAK631      PLELSDSRIMELSTILSPETGDIEIRSRLS EDDWTVHAVARSRKPLPTTSNDPSTKG
961.....970.....980.....990.....1000.....1010.....

RkpA1021      LAKTATVTAAKAYETAKQFGLDYGPFQQLSKAVAYGDRLLIDVELKHPRA SGHPHVTYAL
RkpAK631      LAKTATVTAAKAYETAKQFGLDYGPFQQLSKAVAYGDRLLIDVELKHPRA SGHPHVTYAL
1021.....1030.....1040.....1050.....1060.....1070.....

RkpA1021      HPISVDATFHGLVALFDRFTGDRGGAPYIPVRFSGSVRVVNAGRPIARATIEIERVSANSI
RkpAK631      HPISVDATFHGLVALFDRFTGDRGGA-YIPVRFSGSVRVVNAGRPIARATIEIERVSANSI
1081.....1090.....1100.....1110.....1120.....1130.....

```

RkpA1021 K A R F H F F D E T G V A I A H F E D C R F R R T Y L R K H K T L E G L S F H Y E A V L S D V V M P G T K V A T A M P A  
RkpAK631 K A R F H F F D E T G V A I A H F E D C R F R R T Y L R K H K T L E G L S F H Y E A V L S D V V M P G T K V A T A M P A  
1141 .....1150.....1160.....1170.....1180.....1190.....

RkpA1021 T L A E P I D D N G V D N A T L L F N A A I Y R A C Q E I A L K L G R G T A T V R G D A L P G D F A F Q C F L T S C L L  
RkpAK631 T L A E P I D D N G V D N A T L L F N A A I Y R A C Q E I A L K L G R G T A T V R G D A L P G D F A F Q C F L T S C L L  
1201 .....1210.....1220.....1230.....1240.....1250.....

RkpA1021 T L E D A G L C E H Q N G N W K V A R E F S L P T V P E I L G E L Y G D R P D R A V E A V L V N N A Y A E A L S R L D A  
RkpAK631 T L E D A G L C E H Q N G N W K V A R E F S L P T V P E I L G E L Y G D R P D R A V E A V L V N N A Y A E A L S R L D A  
1261 .....1270.....1280.....1290.....1300.....1310.....

RkpA1021 L L G P V R A G D D A A S F N A G F I S E A T L D H Q A V H S V A S R S R M D Q V L H T V E H V L A A Q P A G A R L R L  
RkpAK631 L L G P V R A G D D A A S F N A G F I S E A T L D H Q A V H S V A S R S R M D Q V L H T V E H V L A A Q P A G A C G --  
1321 .....1330.....1340.....1350.....1360.....1370.....

RkpA1021 T E L G S V S A G F S R R L A D L A A R N G A A L S I F E P R D N A Q R N L E I A F E E D A H V R V L K K S E F A D I G  
RkpAK631 - S S N S A A S P A S A A G W P I W Q R A M V Q H F P F S S R A T M H N G I W K S P S R R M N V R V L K K S E F A D I G  
1381 .....1390.....1400.....1410.....1420.....1430.....

RkpA1021 P F D L A V S A S D N L Y Q L I E E E S G V R T A L R S M P P G S G L V A A V S A P S I F A D F A L G L S G G W F E R S  
RkpAK631 P F D L A V S A S D N L Y Q L I E E E S G V R T A L R S M P P G S G L V A A V S A P S I F A D F A L G L S G G W F E R S  
1441 .....1450.....1460.....1470.....1480.....1490.....

RkpA1021 Q T A E F P I G R I A A V P H W Q K C L A D L D L G D V V T D R E C P H G N T I L I E A R G T A A A G A G A E P G E T  
RkpAK631 Q T A E F P I G R I A A V P H W Q K C L A D L D L G D V V T D R E C P H G N T I L I E A R G T A A A G A G A E P G E T  
1501 .....1510.....1520.....1530.....1540.....1550.....

RkpA1021 T T P Y L L I E T A T A Q K P V A A G G R S A A P V R V T G D L A A D M A T L R A A L E A M G D R P V R A V F M S E A E  
RkpAK631 G T P Y L L I E T A T A Q K P V A A G G R S A A P V R V T G D L A A D M A T L R A A L E A M G D R P L R A V F M S E A E  
1561 .....1570.....1580.....1590.....1600.....1610.....

RkpA1021 R A V G D S A L L Q T R V L A L S A F A E A L K Q H L A G V E P E G D D - R P R L V L V A P G G A P L S A S A G S G L  
RkpAK631 R A V G D S A L L Q T R V L A L S A F A E A L K Q H L A G V E P R A T T G A P - - C A R G P G R R S L I T S A G S G L  
1621 .....1630.....1640.....1650.....1660.....1670.....

RkpA1021 N S G L W A F A R V L Q N E Y E F I D V H T L D L A G D A D I G R K D M L A A A L P L L T V S G Q N R G W L L D R K T G  
RkpAK631 N S G L W A F A R V L Q N E Y E F I D V H T L D L A G D A D I G R K D M L A A A L P L L T V S G Q N R E W L L D R K T G  
1681 .....1690.....1700.....1710.....1720.....1730.....

RkpA1021 L L S E L R A V P G A T N K A D G K T N A F A A A T I R Q R V S Q V A S I A W E E A S V P E P G P N E V V V A A T  
RkpAK631 L L S E L R A V P G A T N K V D G K T N A F A A A T I R Q R V S Q V A S I A W E E A S V P E P G P N E V V V A A T  
1741 .....1750.....1760.....1770.....1780.....1790.....

RkpA1021 G L N F R D V M W A M G L L P E E A L E D G F A G - - - - - A T I G - - - - - M E L S G H  
RkpAK631 G L N F R D V M W A M G L L P E E A L E D G S P G O P S A W N C P A T W S R S A K R S T I S R S A M R A A M E L S G H  
1801 .....1810.....1820.....1830.....1840.....1850.....

RkpA1021 V V A V G E E I D D L A V G D A V M A I A S A A F S T H A V V A R A G I A K L P R A L S P V A A A T L P V A F L T A Y Y  
RkpAK631 V V A V G E E I D D L A V G D A V M A I A S A A F S T H A V V A R A - A S P S A A S V E P V A A A T L P V A F L T A Y Y  
1861 .....1870.....1880.....1890.....1900.....1910.....

RkpA1021 A I V E L G R I Q A G E T I L I H G A A G G V G L A A L Q V A K L K G A K V I A T A G T R E K R R F L T M L G A D H V F  
RkpAK631 A I V E L G R I Q A G E T I L I H G S G G V G L A A L Q V A K L K G A K V I A T A G S A E K R R F L T M L G A D H V F  
1921 .....1930.....1940.....1950.....1960.....1970.....

RkpA1021 D S R S L G F V E D V R S V T G G E G V D L V L N S L F A E A M E Q S L S L V K P F G R F L E L G K R D Y Y A D S K I G  
RkpAK631 D S R S L G F V E D V R S V T G G E G V D L V L N S L F A E A M E Q S L S L V K P F G R F L E L G K R D Y Y A D S K I G  
1981 .....1990.....2000.....2010.....2020.....2030.....

RkpA1021 L R P F R R N V S Y F G I D A D Q L L V N A P D L T R R I F M E I G A L F E E G K L T P L P Y R A F D Y D E I G N A F R  
RkpAK631 L R P F R R N V S Y F G I D A D Q L L V N A P D L T R R I F M Q I G A L F E E G N V T P L P Y R A F D Y G A A - - A - -  
2041 .....2050.....2060.....2070.....2080.....2090.....

RkpA1021 L M Q N A G H I G K I V V L P P V A G R H E V T A K A V R G M K V D P D G V H L V V G G I G G F L V A A N W L V E K G  
RkpAK631 - M Q N A G H I G K I V V L P P V A G R H E V T A K A V R G M K V D P D G V H L V V G G I G G F L V A A N W L V E K G  
2101 .....2110.....2120.....2130.....2140.....2150.....

RkpA1021 A R R I A L C S R R G Q P D A E T R A M I E R - - - W Q N G G V A A S V - - H A C D I T D A A A V E T L L A T L R S E A  
RkpAK631 A R R I A L C S R R G Q P D A E T R A M I E L R A E W R R C G F G P C L R H H G C R G R G D A S G D A S F G - - - - S  
2161 .....2170.....2180.....2190.....2200.....2210.....

RkpA1021 P L R S V V H A A M V L D D A L I G N L S R E R N R P V I E T K A K G A A I L D R L T R G D R L D N F I L F S S A T T L  
RkpAK631 T V R S V V H A A M V L D D A L I S N L S R E R N R P V I E T K A K G A A I L D R L T R G D R L D N F I L F S S A T T L  
2221 .....2230.....2240.....2250.....2260.....2270.....

```

RkpA1021  VGNPQQANYVAANGYLEGLARARRQEGLAGLAVGFGAIADAGYLTQNADVNDLLAKRIGK
RkpAK631  VGNPQQANYVAANGYLEGLARARRQEGLAGLAVGFGAIADAGYLTQNADVNDLLAKRIGK
          2281.....2290.....2300.....2310.....2320.....2330.....

RkpA1021  TALKAQVALDMVENHVAADPGTVDAAVVMISEIDWTAARNLPVARNALFEVILRSADQHA
RkpAK631  TALKAQVALDMVENHVAADPGTVDAAVVMISEIDWTAARNLEVA-----TRSS--RA
          2341.....2350.....2360.....2370.....2380.....2390.....

RkpA1021  AGAEGTMDLVAMIEGKSPQEAEDILFDLVAGEIAAILRVSKDTVTRGKILKEIGLDSL
RkpAK631  A-----AMDLVAMIEGKSPQEAEDILFDLVAGEIAAILRVSKDTVTRGKILKEIGLDSL
          2401.....2410.....2420.....2430.....2440.....2450.....

RkpA1021  AVELGMSFQQNTGFDMPPLSGVADNTTVGDIARKLHEKVSKEVDQGNENENEAGDDKLVTEL
RkpAK631  AVELGMSFQQNTGFDMPPLSGVADNTTVGDIARKLHEKVSKEVDQGNENENEAGDDKLVTEL
          2461.....2470.....2480.....2490.....2500.....2510.....

RkpA1021  TQRHVGADKEKALSQ
RkpAK631  -----
          2521.....2530..

```

## 7.5.2 Nucleotidalignment

### Multiple Alignment

page 1

```

rkpA1021      1 ATGACGGTTGAAATCATAGGGCGTGCATGCCTTGCCCGGGAGCGAAATCTCCGCAGGCG
rkpA-F        1 ATGACGGTTGAAATCATAGGGCGTGCATGCCTTGCCCGGGAGCGAAATCTCCGCAGGCG

rkpA1021      61 CTGTTCAAAATCCTTCGTGAGGCAAGTGCACCGTGACCCGCGTCCCTTCGGACCGCTGG
rkpA-F        61 CTGTTCAAAATCCTTCGTGAGGCAAGTGCACCGTGACCCGCGTCCCTTCGGACCGCTGG

rkpA1021      121 GACCTTGCGCGTPTCTGGCACCCGGTGCATGGGCACGCCCGGCAAGACCTATTCTGTTGGCC
rkpA-F        121 GACCTTGCGCGTPTCTGGCACCCGGTGCATGGGCACGCCCGGCAAGACCTATTCTGTTGGCC

rkpA1021      181 GCCGGCGTTCGACCATATCTACGATTTGATCCCGCGTTTTTCGGGATGTCGACGGC
rkpA-F        181 GCCGGCGTTCGACCATATCTACGATTTGATCCCGCGTTTTTCGGGATGTCGACGGC

rkpA1021      241 GAGGCCATGTATATGGACCCGACGACGGGTTCTGTTGCAGCTCGCCTGGCGAGCGCTC
rkpA-F        241 GAGGCCATGTATATGGACCCGACGACGGGTTCTGTTGCAGCTCGCCTGGCGAGCGCTC

rkpA1021      301 GAGGACGCCAATATTTCCGTGGCGTCGCTGCACGGAGAAAAATGTCGGCGTCTATGTCGGC
rkpA-F        301 GAGGACGCCAATATTTCCGTGGCGTCGCTGCACGGAGAAAAATGTCGGCGTCTATGTCGGC

rkpA1021      361 GCCTCCAGCCTCGACACGCCAATCTGACGGTCGACGATCCGGCTGCCGCGGGCCCTAC
rkpA-F        361 GCCTCCAGCCTCGACACGCCAATCTGACGGTCGACGATCCGGCTGCCGCGGGCCCTAC

rkpA1021      421 TTCATGACGGGCAACACGCTGTCGATCGTCTCCAACCGTATTTCCACGCTCTTCGGCCTG
rkpA-F        421 TTCATGACGGGCAACACGCTGTCGATCGTCTCCAACCGTATTTCCACGCTCTTCGGCCTG

rkpA1021      481 AGCGGCCCGAGCATGACCCGTCGACACGGCCTGTTCTCTTCGCTCGTGGCCCTCGATCAG
rkpA-F        481 AGCGGCCCGAGCATGACCCGTCGACACGGCCTGTTCTCTTCGCTCGTGGCCCTCGATCAG

rkpA1021      541 GCGATGCGGGCGCTGAACCCCGAGAGATCGACACGGCCATCGTCGGCGCGTCAATATC
rkpA-F        541 GCGATGCGGGCGCTGAACCCCGAGAGATCGACACGGCCATCGTCGGCGCGTCAATATC

rkpA1021      601 CTCGCCACCCGCTTCCTTCGTGCGCTTCGCCAGGCGCGCATGCTCTCGCCGAAGGG
rkpA-F        601 CTCGCCACCCGCTTCCTTCGTGCGCTTCGCCAGGCGCGCATGCTCTCGCCGAAGGG

rkpA1021      661 CTCTGCCCGCCTATGACAATGACGGCGCCGGCTATGTTCCGCGCGAGGGCGCGCTGCTC
rkpA-F        661 CTCTGCCCGCCTATGACAATGACGGCGCCGGCTATGTTCCGCGCGAGGGCGCGCTGCTC

rkpA1021      721 TTCGTCCTGCGCCGACCGACCGGGCGCAACGCGAGCGCGACCGGAGCTATGCCGGATC
rkpA-F        721 TTCGTCCTGCGCCGACCGACCGGGCGCAACGCGAGCGCGACCGGAGCTATGCCGGATC

rkpA1021      781 GTCGCCTCCGGCGTCAATTCGGCCGGGCGCAACGGCATCTCGCTGCCCTCGCGGAG
rkpA-F        781 GTCGCCTCCGGCGTCAATTCGGCCGGGCGCAACGGCATCTCGCTGCCCTCGCGGAG

rkpA1021      841 GCGCAGGCCAATCTGCTGAGGATGATCTACGAAGGCAACGGCATCGACGCCAACAGGTC
rkpA-F        841 GCGCAGGCCAATCTGCTGAGGATGATCTACGAAGGCAACGGCATCGACGCCAACAGGTC

rkpA1021      901 GCCTTCGTAGAAGGCCATGGTACCGGCACGAAGTTCGGCGACCCGGCGGAGGTATGGTCC
rkpA-F        901 GCCTTCGTAGAAGGCCATGGTACCGGCACGAAGTTCGGCGACCCGGCGGAGGTATGGTCC

rkpA1021      961 ATAGGCACGGTGATCGGCGCAAGCGCCGGGCCCGGTGCCGATCGGCTCGATAAAGTCC
rkpA-F        961 ATAGGCACGGTGATCGGCGCAAGCGCCGGGCCCGGTGCCGATCGGCTCGATAAAGTCC

rkpA1021      1021 AATATCGGCATACCGAGCCGGCTCCGGCCTGTTTCGGGATGATGAAGGCGGTTCTGGCT
rkpA-F        1021 AATATCGGCATACCGAGCCGGCTCCGGCCTGTTTCGGGATGATGAAGGCGGTTCTGGCT

rkpA1021      1081 CTGGAGAACAATTATCTGCCGGCTTCGTCGATTTTCGAGACGCCGAACGAGCAGATCGAC
rkpA-F        1081 CTGGAGAACAATTATCTGCCGGCTTCGTCGATTTTCGAGACGCCGAACGAGCAGATCGAC

rkpA1021      1141 TTCGACGGCTCAATGTCCGTGTGACTGCCAATCCGATCGAGCTTCTGAAGGGCAAGCGG
rkpA-F        1141 TTCGACGGCTCAATGTCCGTGTGACTGCCAATCCGATCGAGCTTCTGAAGGGCAAGCGG

rkpA1021      1201 GCGGGCTCGCCGCATCAATTCCTTCGGCTTCGGCGGCGCAACGCCATGTCGTCATC
rkpA-F        1201 GCGGGCTCGCCGCATCAATTCCTTCGGCTTCGGCGGCGCAACGCCATGTCGTCATC

rkpA1021      1261 AGCGATCCGGATCCGGCACAGGCGAAAAGGCTGCGACCTCGCCGGCGGGCCACGTTT
rkpA-F        1261 AGCGATCCGGATCCGGCACAGGCGAAAAGGCTGCGACCTCGCCGGCGGGCCACGTTT

```

## Multiple Alignment

page 2

```
rkpA1021 1321 CTGGCAAGCGCCCATACCGCGTCGAGCCTCGAAAACCTGCTC_AGAGACTATAAGGCCGC
rkpA-F 1321 CTGGCAAGCGCCCATACCGCGTCGAGCCTCGAAAACCTGCTC_GAGAGACTATAAGGCCGC

rkpA1021 1380 CTTTGC GGCGCGCAGGAAGAGACGCGCGGATCGTCGCCCTCCGGCGCAATCG
rkpA-F 1381 CTTTGC GG_CGCGCAGGAAGAGACGCGCGGATCGTCGCCCTCCGGCGCAATCG

rkpA1021 1440 GACGCATATGCGCCATCGGTTTCGCGGCAGCAGCGATCATCCGAAGATATCGTCGGGC
rkpA-F 1440 GACGCATATGCGCCATCGGTTTCGCGGCAGCAGCGATCATCCGAAGATATCGTCGGGC

rkpA1021 1500 GATTGCCAACCACTGGAAAAGCCCGGCTCCGATATCGGGGAGACGGCGAAGCGGCCG
rkpA-F 1500 GATTGCCAACCACTGGAAAAGCCCGGCTCCGATATCGGGGAGACGGCGAAGCGGCCG

rkpA1021 1560 GAAAGAAGCAAAGTTCGCCTTCGTCTTCTCCGGCAACGGCTCGCAATGGCCCGCATGGG
rkpA-F 1560 GAAAGAAGCAAAGTTCGCCTTCGTCTTCTCCGGCAACGGCTCGCAATGGCCCGCATGGG

rkpA1021 1620 TGTCGAGGCCCTTCGCGGAGAATCTGCACTTCCGTGAGTTCGTTCACTTCGGTCAGCGCCCT
rkpA-F 1620 TGTCGAGGCCCTTCGCGGAGAATCTGCACTTCCGTGAGTTCGTTCACTTCGGTCAGCGCCCT

rkpA1021 1680 TTTGAGTTCCATTCGGATATCGTCTCACCATCTCCTGACCGACCCCGAAGTTCGACAG
rkpA-F 1680 TTTGAGTTCCATTCGGATATCGTCTCACCATCTCCTGACCGACCCCGAAGTTCGACAG

rkpA1021 1740 GAAGCTCGCGGATACGAAGTGGCCAGCCGCTGCTTTTCGCGATACAGCCCGCTCTC
rkpA-F 1740 GAAGCTCGCGGATACGAAGTGGCCAGCCGCTGCTTTTCGCGATACAGCCCGCTCTC

rkpA1021 1800 CGACTCGCTCGTGGCCATGGGCATCAAGCCGACCCGCTTTCGGTTCATTCGGTCGGCGA
rkpA-F 1800 CGACTCGCTCGTGGCCATGGGCATCAAGCCGACCCGCTTTCGGTTCATTCGGTCGGCGA

rkpA1021 1860 AATCGCTGCGGCTTATGACGCGGGGGCTGTCGCTCGTCGACCGCGTTCGATCGTCGC
rkpA-F 1860 AATCGCTGCGGCTTATGACGCGGGGGCTGTCGCTCGTCGACCGCGTTCGATCGTCGC

rkpA1021 1920 CAAACGATCGTGCACCAGGACCTTCTGGCCGGCCAGGGCACGATGGCGCCGTTATGCT
rkpA-F 1920 CAAACGATCGTGCACCAGGACCTTCTGGCCGGCCAGGGCACGATGGCGCCGTTATGCT

rkpA1021 1980 GGGCGAGGAGGCAGCCAAGGCCCTTCGCGCGGAGCGCGGTCTCGACGATGTCGCGTCGC
rkpA-F 1980 GGGCGAGGAGGCAGCCAAGGCCCTTCGCGCGGAGCGCGGTCTCGACGATGTCGCGTCGC

rkpA1021 2040 AGCAATCAACGCTCACAATTCGGTGACGATCTCCGGCCCGCGAAGGAGATTTCCGGCGTT
rkpA-F 2040 AGCAATCAACGCTCACAATTCGGTGACGATCTCCGGCCCGCGAAGGAGATTTCCGGCGTT

rkpA1021 2100 CCGCGACGCTGCGCGCAAGGCCAAAATCCCGTTTCAGATCCTCGACATCAATTTATCCGTT
rkpA-F 2100 CCGCGACGCTGCGCGCAAGGCCAAAATCCCGTTTCAGATCCTCGACATCAATTTATCCGTT

rkpA1021 2160 CCACCATCCGATCATCGACCGGGCGAAAGAGGCGTTCCTCTCCGATATCCTGATATCGC
rkpA-F 2160 CCACCATCCGATCATCGACCGGGCGAAAGAGGCGTTCCTCTCCGATATCCTGATATCGC

rkpA1021 2220 CCCCCGCGGACGGAGCTGGCCTATCTGTCGACCGTTCACGGGTGCGCGCTCGACGGGAC
rkpA-F 2220 CCCCCGCGGACGGAGCTGGCCTATCTGTCGACCGTTCACGGGTGCGCGCTCGACGGGAC

rkpA1021 2280 TGCTCTCGATCCCGACTACTGGTGAAGAACGTTTCGCGAGCCCGTTCGTTTCAGGCGCG
rkpA-F 2280 TGCTCTCGATCCCGACTACTGGTGAAGAACGTTTCGCGAGCCCGTTCGTTTCAGGCGCG

rkpA1021 2340 TGCCGAGGCCGCGCTCGATCTCGGCTGCATGCTGTTTCATCGAGATCTCGCCGCTCCGAT
rkpA-F 2340 TGCCGAGGCCGCGCTCGATCTCGGCTGCATGCTGTTTCATCGAGATCTCGCCGCTCCGAT

rkpA1021 2400 CCTTGGTCTCTACGTCAAGGAGACGATCAAGCAGGCGGAGTCCCGGCTTCGGTGGTGGC
rkpA-F 2400 CCTTGGTCTCTACGTCAAGGAGACGATCAAGCAGGCGGAGTCCCGGCTTCGGTGGTGGC

rkpA1021 2460 GACATTGCTGCGCGACGCCGCGCGAAGCGCCACGATCCGATCTCCGCTCGATGGCGCG
rkpA-F 2460 GACATTGCTGCGCGACGCCGCGCGAAGCGCCACGATCCGATCTCCGCTCGATGGCGCG

rkpA1021 2520 CGCAGTCGCCACGGTGC CGCGTTCGACCGGTCGCGTGTCTATGGCAAGCGGGACGCATT
rkpA-F 2520 CGCAGTCGCCACGGTGC CGCGTTCGACCGGTCGCGTGTCTATGGCAAGCGGGACGCATT

rkpA1021 2580 CATCGAACTCCCCGCGCTGCCTTTCGAGCCGGTGGAACTGCGGCCGCGGCCACGACCGA
rkpA-F 2580 CATCGAACTCCCCGCGCTGCCTTTCGAGCCGGTGGAACTGCGGCCGCGGCCACGACCGA
```

## Multiple Alignment

page 3

```

rkpA1021 2640 TGCACGGATCTCTTCGGACGTTCCGCAAAGCCCTATCGGCTCAACGGCTGGCGCGGCGA
rkpA-F 2640 TGCACGGATCTCTTCGGACGTTCCGCAAAGCCCTATCGGCTCAACGGCTGGCGCGGCGA

rkpA1021 2700 CCCCAATGCGGGAAGCTGGAAAAACCACGTCGACGCGCATCTCTCCCTGATCTCGCCGA
rkpA-F 2700 CCCCAATGCGGGAAGCTGGAAAAACCATGTCGACGCGCATCTCTCCCGATCTCGCCGA

rkpA1021 2760 ACACGTGGTCGACGGCAAGGCGATCTCCCGGGCAGCGGTTTCATCGAAATCGCCGTTTC
rkpA-F 2760 ACACGTGGTCGACGGCAAGGCGATCTCCCGGGCAGCGGTTTCATCGAAATCGCCGTTTC

rkpA1021 2820 CGCCGCCAGCAATATCACGGCAGCGACGAGTCGAGATCACCAATCTGGAGATCGTCCG
rkpA-F 2820 CGCCGCCAGCAATATCACGGCAGCGACGAGTCGAGATCACCAATCTGGAGATCGTCCG

rkpA1021 2880 CCCGCTCGAACTGAGCGACAGCCGCATCATGGAATCTCGACCATCTGTGCGCCGAGAC
rkpA-F 2880 CCCGCTCGAACTGAGCGACAGCCGCATCATGGAATCTCGACCATCTGTGCGCCGAGAC

rkpA1021 2940 CGGCGACATCGAGATCCGGTCGCGCAACGTCTCTCGGAAGACGATTGGACGGTGCATGC
rkpA-F 2940 CGGCGACATCGAGATCCGGTCGCGCAACGTCTCTCGGAAGACGATTGGACGGTGCATGC

rkpA1021 3000 CGTCGCGCAAGCCGCAAGCCCTTACCGACGACGAGCAACGACTGTCCGTCCACCAAAGG
rkpA-F 3000 CGTCGCGCAAGCCGCAAGCCCTTACCGACGACGAGCAACGACTGTCCGTCCACCAAAGG

rkpA1021 3060 CTTGGCGAAGACGGCGACCGTGACGGCGCCAAAGCCTATGAAACGGCGAAGCAGTTCGG
rkpA-F 3060 CTTGGCGAAGACGGCGACCGTGACGGCGCCAAAGCCTATGAAACGGCGAAGCAGTTCGG

rkpA1021 3120 GCTGGACTACGGCCCGCAATTCAGCTTCTCTCCAAAGCCGTGGCCTATGGCGACCGGCT
rkpA-F 3120 GCTGGACTACGGCCCGCAATTCAGCTTCTCTCCAAAGCCGTGGCCTATGGCGACCGGCT

rkpA1021 3180 CATCGACGTGAGTTGAAGCACCCCGGGCATCGGGGCACCCACGTCACCTACGCCCT
rkpA-F 3180 TATCGACGTGAGTTGAAGCACCCCGGGCATCGGGGCACCCACGTCACCTACGCCCT

rkpA1021 3240 TCATCCGATCTCGGTCGACGCGACCTTCCACGGCCTGGTCGCACTCTTCGACCGCTTCAC
rkpA-F 3240 TCATCCGATCTCGGTCGACGCGACCTTCCACGGCCTGGTCGCACTCTTCGACCGCTTCAC

rkpA1021 3300 GGGCGACAGGGGGCGCTCCCTATATTCCTGTCCGTTTCGGCTCCGTGCGCGTCGTCAA
rkpA-F 3300 GGGCGACAGGGGGCGCG __CCTATATTCCTGTCCGTTTCGGCTCCGTGCGCGTCGTCAA

rkpA1021 3360 TGCCGGCCGGCCGATCGCCCGCGGACGATCGAGATCGAGCGGGTCAGCGCCAACTCGAT
rkpA-F 3357 TGCCGGCCGGCCGATCGCCCGCGGACGATCGAGATCGAGCGGGTCAGCGCCAACTCGAT

rkpA1021 3420 CAAGGCACGGTTCCATTTCTTCGACGAGACCGGCGTCGGATCGCACACTTCGAGGATTG
rkpA-F 3417 CAAGGCACGGTTCCATTTCTTCGACGAGACCGGCGTCGGATCGCACACTTCGAGGATTG

rkpA1021 3480 CCGCTTCGGCCGTACCTATCTGCGCAAGCACAAAGCGCTTGAAGGCCTTTCCCTTCCACTA
rkpA-F 3477 CCGCTTCGGCCGTACCTATCTGCGCAAGCACAAAGCGCTTGAAGGCCTTTCCCTTCCACTA

rkpA1021 3540 TGAGGCAGTTCTGTCCGACGTGGTTCATGCCCGGCACGAAAGTTGCGACGGCCATGCCTGC
rkpA-F 3537 TGAGGCAGTTCTGTCCGACGTGGTTCATGCCCGGCACGAAAGTTGCGACGGCCATGCCTGC

rkpA1021 3600 GACGCTGGCCGAGCCGATCGATGACAATGGCGTCGACAACGCGACCTGTCTTCAACGC
rkpA-F 3597 GACCTGGCCGAGCCGATCGATGACAATGGCGTCGACAACGCGACCTGTCTTCAACGC

rkpA1021 3660 CGCCATCTACCGCCTGCCAGGAAATGGCCCTGAAGCTCGGCAGGGGGACCGCCACAGT
rkpA-F 3657 CGCCATCTACCGCCTGCCAGGAAATGGCCCTGAAGCTCGGCAGGGGGACCGCCACAGT

rkpA1021 3720 GCGAGGCGACGCGCTGCCCGGGGACTTCGCGTTCAGTGTCTTCTGACGAGCTGCCCTGCT
rkpA-F 3717 GCGAGGCGACGCGCTGCCCGGGGACTTCGCGTTCAGTGTCTTCTGACGAGCTGCCCTGCT

rkpA1021 3780 GACGCTCGAGGATGCCGGCCTTTGCGAGCACCAGAACGGCAACTGGAAGGTCGCACGGGA
rkpA-F 3777 GACGCTCGAGGATGCCGGCCTTTGCGAGCACCAGAACGGCAACTGGAAGGTCGCACGGGA

rkpA1021 3840 GTTCTCGCTTCCGACCGTTCCGGAGATCCTTGGCGAGCTCTATGGCGACCGCCCGATCG
rkpA-F 3837 GTTCTCGCTTCCGACCGTTCCGGAGATCCTTGGCGAGCTCTATGGCGACCGCCCGATCG

rkpA1021 3900 TGCGGTGGAAGCGGTGCTCGTCAACAACGCCATGCGGAAGCACTCAGCCGTCTCGACGC
rkpA-F 3897 TGCGGTGGAAGCGGTGCTCGTCAACAACGCCATGCGGAAGCACTCAGCCGTCTCGACGC

```

## Multiple Alignment

page 4

```
rkpA1021 3960 GCTGCTTGGGCCGGTGCCTGCAGGCGACGACGCGGTTTCGTTCAATGCCGGTTTCATCAG
rkpA-F 3957 GCTGCTTGGGCCGGTGCCTGCAGGCGACGACGCGGTTTCGTTCAATGCCGGTTTCATCAG

rkpA1021 4020 CGAAGCGACGCTCGACCATCAGGCCGTGCATTCCGTGGCGAGCCGGTTCGCGCATGGACCA
rkpA-F 4017 CGAAGCGACGCTCGACCATCAGGCCGTGCATTCCGTGGCGAGCCGGTTCGCGCATGGACCA

rkpA1021 4080 GGTGCTTCATACGGTTCGAACACGTTCTCGCTGCGCAGCCGGCAGGCCGCCCTGCGGCT
rkpA-F 4077 GGTGCTTCATACGGTTCGAACACGTTCTCGCTGCGCAGCCGGCAGGCCGCCCTGCGGCT

rkpA1021 4140 CATCGAACTCGGCAGCGTCTCCGCCGGTTCAGCCGCCGGTGGCCGATCTGGCAGCGCG
rkpA-F 4133 CATCGAACTCGGCAGCGTCTC_CCGCTTCAGCCGCCGGTGGCCGATCTGGCAGCGCG

rkpA1021 4200 CAATGGTGCAGCACTTTCATTTTCGAGCCGCGGACAATGCACAACGGAATCTGGAAAT
rkpA-F 4190 CAATGGTGCAGCACTTTCATTTTCGAGCCGCGGACAATGCACAACGGAATCTGGAAAT

rkpA1021 4260 CGCCTTCGAGGAGGATGCGCATGTGCGGTTCTGAAGAAGAGCGAATTCGCCGATATCGG
rkpA-F 4250 CGCCTTCGAGGAGGATG_AATGTGCGGTTCTGAAGAAGAGCGAATTCGCCGATATCGG

rkpA1021 4320 CCCGTTTGACCTCGCCGTCAGTGCCTTCGGACAATCTCTACCAGGTGATCGAAGAGGAGAG
rkpA-F 4308 GCCGTTTGACCTCGCCGTCAGTGCCTTCGGACAATCTCTACCAGGTGATCGAAGAGGAGAG

rkpA1021 4380 CGGCGTCCGCACGGCCCTTCGCTCCATGCCCCCGGAAGCGGGCTCGTCGCGCCGTCAG
rkpA-F 4368 CGGCGTCCGCACGGCCCTTCGCTCCATGCCCCCGGAAGCGGGCTCGTCGCGCCGTCAG

rkpA1021 4440 CGCGCCGTCGATCTTCGCGGATTTTCGCCCTGGGTCTCTCGGGCGGATGGTTCGAGCGCAG
rkpA-F 4428 CGCGCCGTCGATCTTCGCGGATTTTCGCCCTGGGTCTCTCGGGCGGATGGTTCGAGCGCAG

rkpA1021 4500 CCAGACGGCAGAATTCGCGATCGGCAGGATCGCGCCGTACCCCATTTGGCAGAAGTGTCT
rkpA-F 4488 CCAGACGGCAGAATTCGCGATCGGCAGGATCGCGCCGTACCCCATTTGGCAGAAGTGTCT

rkpA1021 4560 CGCAGATCTCGACCTCGGCGACGTGGTGTGACGGATCGCGAATGCCCGCATGGGAACAC
rkpA-F 4548 CGCAGATCTCGACCTCGGCGACGTGGTGTGACGGATCGCGAATGCCCGCATGGGAACAC

rkpA1021 4620 CATCCTGATCGAGCCCGTGGCACAGCGCGGGCGGCGCAGGCCCGGAGCCGGCGAAAC
rkpA-F 4608 CATCCTGATCGAGCCCGGGGCACAGCGCGGCGGCGGCGCAGGCCCGGAGCCGGCGAAAC

rkpA1021 4680 CACGACACCATATCTGCTGATCGAGACTGCCACTGCCAGAAACCGGTTCGACGCGGTGG
rkpA-F 4668 CGGAACACCATATCTGCTGATCGAGACTGCCACTGCCAGAAACCGGTTCGACGCGGTGG

rkpA1021 4740 AAGGAGTGCCGCTCCCGTCCGTGTGACAGGCGACCTTGCCCGCAGACATGGCGACGCTCCG
rkpA-F 4728 AAGGAGTGCCGCTCCCGTCCGTGTGACAGGCGACCTTGCCCGCAGATATGGCGACGCTCCG

rkpA1021 4800 CGCAGCCCTCGAAGCCATGGGCGACAGACCGGTGCGCGGGTTTTCATGTCCGAGGCGGA
rkpA-F 4788 CGCAGCCCTCGAAGCCATGGGCGACAGACCGGTGCGCGGGTTTTCATGTCCGAGGCGGA

rkpA1021 4860 GCGCGCCGTGCGGACGCTTCGGCCCTGCTGCAGACGCGTGTATTGGCGCTCAGCGCCTT
rkpA-F 4848 GCGCGCCGTGCGGACGCTTCGGCCCTGCTGCAGACGCGTGTATTGGCGCTCAGCGCCTT

rkpA1021 4920 TGCCGAGGCATTGAAACAGCATCTCGCCGGCTCGAGCCGAGGGCGACACCGGCCGCG
rkpA-F 4908 TGCCGAGGCATTGAAACAGCATCTCGCCGGCTCGAGCCG_AGGGCGACACCGGCCG_GCG

rkpA1021 4980 CCTTGTGCTCGTGGCCCCGGGCGGCTCCCTTATTCAGCCTCGGCAGGCTCCGGGCTGAA
rkpA-F 4966 CCTTGTGCTCGTGGCCCCGGGCGGCTCCCTTATCA_CCTCGGCAGGCTCCGGGCTGAA

rkpA1021 5040 TTCCGGATTGTGGCCCTTCGCCCGGCTCTTCAGAACGAATATGAGTTTATCGACGTTCA
rkpA-F 5025 TTCCGGATTGTGGCCCTTCGCCCGGCTCTTCAGAACGAGTATGAGTTTATCGACGTTCA

rkpA1021 5100 CACGCTCGATCTTGCCGGCATGCCGACATCGGCCGCAAGGATATGCTTGCCCGCAGCGCT
rkpA-F 5085 CACGCTCGATCTTGCCGGTATGCCGACATCGGCCGCAAGGATATGCTTGCCCGCAGCGCT

rkpA1021 5160 TCCCTGCTGACGGTTTCGGGACAGAACCGGGATGGCTGCTCGACCGGAAGACGGGCCCT
rkpA-F 5145 TCCCTGCTGACGGTTTCGGGACAGAACCGGGAATGGCTGCTCGACCGGAAGACGGGCCCT

rkpA1021 5220 GCTGTGCGAGCTTCGGGCGTACCCGGCGCAACGAACAAGGTTGACGGGAAAACCAACGC
rkpA-F 5205 GCTGTGCGAGCTTCGGGCGTACCCGGCGCAACGAACAAGGTTGACGGGAAAACCAACGC
```

## Multiple Alignment

page 5

```
rkpA1021 5280 CTTTCGCGGCCGCGACCATTCGCGCAGCGCTCAGCTCGCAGGTTGCAAGCATCGCCTGGGA
rkpA-F 5265 CTTTCGCGGCCGCGACCATTCGCGCAGCGCTCAGCTCGCAGGTTGCAAGCATCGCCTGGGA

rkpA1021 5340 AGAGGCGAGCGTACCGGAGCCGGGCCCCAACGAGGTCGTCTCGCCGTCGCGGCAACCGG
rkpA-F 5325 AGAGGCGAGCGTACCGGAGCCGGGCCCCAACGAGGTCGTCTCGCCGTCGCGGCAACCGG

rkpA1021 5400 CTTGAATTTTCGCGACGTGATGTGGGCCATGGGCCCTGTTGCCGGAAGAGCGCTTGAAGA
rkpA-F 5385 CTTGAATTTTCGCGACGTGATGTGGGCCATGGGCCCTGTTGCCGGAAGAGCGCTTGAAGA

rkpA1021 5460 CGGCTTCGCGGGGCAACCATCGGCATGGAAGTGTCCGGCCACGTGGTCGCGGTCGCGCA
rkpA-F 5445 CGG_TTCGCGGGGCAACCATCGGCATGGAAGTGTCCGGCCACGTGGTCGCGGTCGCGCA

rkpA1021 5520 AGAGATCGACGATTCGCGGTCGCGCATGCGGTGATGGCGATCGCCTCCGCCGCTTCTC
rkpA-F 5504 AGAGATCGACGATTCGCGGTCGCGCATGCGGTGATGGCGATCGCCTCCGCCGCTTCTC

rkpA1021 5580 GACCCATGCGGTTCGTTCGCGCGCGCCGCGCATCGCCAAGTGTCCGCGAGCGTTGAGCCCGGT
rkpA-F 5564 GACCCATGCGGTTCGTTCGCGCGCGCCG_CATCGCCAAG_TGCCGCGAGCGTTGAGCCCG_T

rkpA1021 5640 GGCCGCGGCCACATTCGCGGTCGCGCTTCTTTACCGCCTACTACGCCATTGTCGAGTCGG
rkpA-F 5621 GGCCGCGGCCACATTCGCGGTCGCGCTTCTTTACCGCCTACTACGCCATTGTCGAGTCGG

rkpA1021 5700 CCGCATTTCAGGCTGGAGAGACGATTTTGTATCCACGGTGTGCGCGCGCGCTCGCCCTTGC
rkpA-F 5681 CCGCATTTCAGGCTGGAGAGACGATTTTGTATCCACGGTGTGCGCGCGCGCTCGCCCTTGC

rkpA1021 5760 CGCGCTCCAGGTCGCAAAGCTCAAGGGCGCCAAGGTCATCGCCACGGCGGAAACGCGC_G
rkpA-F 5741 CGCGCTCCAGGTCGCAAAGCTCAAGGGCGCCAAGGTCATCGCCACGGCGGAA_GCGCCG

rkpA1021 5819 AAAAGCGGCGCTTCTTGACGATGCTCGGCGCGACCATGTCTTCGATTCGCGTTCGCTCG
rkpA-F 5800 AAAAGCGGCGCTTCTTGACGATGCTCGGCGCGACCATGTCTTCGATTCGCGTTCGCTCG

rkpA1021 5879 GCTTCGTCGAGGATGTCGGTTCGGTTCACCGCGGGGAAGGCGTCGACCTCGTCTCAATT
rkpA-F 5860 GCTTCGTCGAGGATGTCGGTTCGGTTCACCGCGGGGAAGGCGTCGACCTCGTCTCAATT

rkpA1021 5939 CGCTTTTCGCGAAGCGATGGAGCAGAGCCTCTCGCTGGTGAACCCCTTTGGCCGGTTCC
rkpA-F 5920 CGCTTTTCGCGAAGCGATGGAGCAGAGCCTCTCGCTGGTGAACCCCTTTGGCCGGTTCC

rkpA1021 5999 TGGAATTCGGCAAGCGGACTATTACGCGACAGCAAGATCGGGCTCAGGCCGTTCCGCC
rkpA-F 5980 TGGAATTCGGCAAGCGGACTATTACGCGACAGCAAGATCGGGCTCAGGCCGTTCCGCC

rkpA1021 6059 GCAATGTCAGCTATTTTCGGCATCGACGCTGATCAGTTGCTGGTGAATGCGCCGGATCTGA
rkpA-F 6040 GCAATGTCAGCTATTTTCGGCATCGACGCGATCAGTTGCTGGTGAATGCGCCGGATCTGA

rkpA1021 6119 CCAGGCGCATCTTCATGGAGATCGGAGCGCTCTTCGAGGAAGGCAAGCTCACGCCGCTTC
rkpA-F 6100 CCAGGCGCATCTTCATGGAGATCGGAGCGCTCTTCGAGGAAGGCAAGCTCACGCCGCTTC

rkpA1021 6179 CCTATCGCGCTTCGATTACG_ATGAAATCGG_CAAATGCTT_CCCGGCTGATGCAGAATG
rkpA-F 6160 CCTATCGCGCTTCGATTACGATTACGAAATCGGGCAA_CTTGTCCGGCTGATGCAGAATG

rkpA1021 6236 CCGGCCATATCGGCAAGATCGTGGTGTGCTGCCGCCGTTGCCGGCAGGCACGAGGTGACGG
rkpA-F 6218 CCGGCCATATCGGCAAGATCGTGGTGTGCTGCCGCCGTTGCCGGCAGGCACGAGGTGACGG

rkpA1021 6296 CGAAGGCTGTGAGGGGCATGAAGGTCGACCCGGACGGCGTCCATCTCGTCTGCGCGGCA
rkpA-F 6278 CGAAGGCTGTGAGGGGCATGAAGGTCGACCCGGACGGCGTCCATCTCGTCTGCGCGGCA

rkpA1021 6356 TCGGTGGCTTCGGCTTCGTTCGCGGCAACTGGCTCGTCGAGAAAGGCGCACGCCGTATTG
rkpA-F 6338 TCGGTGGCTTCGGCTTCGTTCGCGGCAACTGGCTCGTCGAGAAAGGCGCACGCCGTATTG

rkpA1021 6416 CGCTTTGCTCCCGCCGAGGCCAGCCGGATGCGGAGACGCGCGCATGATCGAGC_GCTGG
rkpA-F 6398 CGCTTTGCTCCCGCCGAGGCCAGCCGGATGCGGAAACGCGCGCATGATCGAGCTGCGGG

rkpA1021 6475 CAGAATGGCGCGTTGCGGCTTCGGTCCATGCTTTCGACATCACGGATGCGCGGCCGCTG
rkpA-F 6458 CAGAATGGCGCGTTGCGGCTTCGGTCCATGCTTTCGACATCACGGATGCGCGGCCGCTG

rkpA1021 6535 GAGACGCTTTCGGCGACGCTTCGTTTCGGAAGCACCGCTTCGCTCCGTCGTTACGCGGGC
rkpA-F 6518 GAGACGCTTTCGGCGACGCTTCGTTTCGGAAGCACCG_TTCGCTCCGTCGTTACGCGGGC
```

## Multiple Alignment

page 6

```

rkpA1021 6595 ATGGTGCCTCGACGATGCGCTGATCGGCAATCTGAGCCGCGAACGCAACCGGCCGGTCATC
rkpA-F   6577 ATGGTGCCTCGACGATGCGCTGATCAGCAATCTGAGCCGCGAACGCAACCGGCCGGTCATC

rkpA1021 6655 GAAACGAAAGCCAAGGGCGCGCAATTCTCGATCGCCTTACCCGTGGTGACAGGCTCGAC
rkpA-F   6637 GAAACGAAAGCCAAGGGCGCGCAATTCTCGATCGCCTTACCCGTGGTGACAGGCTCGAC

rkpA1021 6715 AACTTCATCCTGTTCTCCTCGGCCACGACGCTGGTCGGCAACCCCGCCAGGCGAACTAT
rkpA-F   6697 AACTTCATCCTGTTCTCCTCGGCCACGACGCTGGTCGGCAATCCCGCCAGGCGAACTAT

rkpA1021 6775 GTCGCCGCCAACGGCTATCTCGAAGGCCTTGCCCGTGCACGGGCCAGGAGGGCCTGGCA
rkpA-F   6757 GTCGCCGCCAACGGCTATCTCGAAGGCCTTGCCCGTGCACGGGCCAGGAGGGCCTGGCA

rkpA1021 6835 GGGCTCGCCGTCGGCTTCGGCGCGATTGCCGATGCCGGCTACCTGACGCAGAACGCCGAT
rkpA-F   6817 GGGCTCGCCATCGGCTTCGGCGCGATTGCCGATGCCGGCTACCTGACGCAGAACGCCGAT

rkpA1021 6895 GTCAACGACCTGCTCGCCAAGCGCATCGGCAAAACGGCGCTGAAGGCCAGGTGGCGCTC
rkpA-F   6877 GTCAACGACCTGCTCGCCAAGCGCATCGGCAAAACGGCGCTGAAGGCCAGGTGGCGCTC

rkpA1021 6955 GACATGGTGGAAAACACGTCGCCGCGGATCCGGGCACGGTCGACGCGGCCGTCGTGATG
rkpA-F   6937 GACATGGTGGAAAACACGTCGCCGCGGATCCGGGCACGGTCGACGCGGCCGTCGTGATG

rkpA1021 7015 ATTTGCGAGATCGACTGGACGGCTGCGCGGAACCTGCCGGTCGCCCGCAACCGCTCTTC
rkpA-F   6997 ATTTGCGAGATCGACTGGACGGCTGCGCGGAACCTGCCGGTCGCC_GCAACCGCTCTTC

rkpA1021 7075 GAGGTGATCCTGCGCAGCGCCGATCAGCATGCCGAGGCGCCGAGGGCACGACGATGGAT
rkpA-F   7056 GAGGTGATCCTGCGCAGCGCCGATCAGCATGCCGAGGCGCCGAGGGCACGACGATGGAT

rkpA1021 7135 CTCGTCGCCATGATCGAAGGCAAATCGCCGCAGGAGGCCGAGGACATCCTCTTCGATCTG
rkpA-F   7116 CTCGTCGCCATGATCGAAGGCAAATCGCCGCAGGAGGCCGAGGACATCCTCTTCGATCTG

rkpA1021 7195 GTGGCCGGCGAGATCGCTGCCATCCTTCGCGTCTCGAAGGATACGGTCACTCGCGGAAAAG
rkpA-F   7176 GTGGCCGGCGAGATCGCTGCCATCCTTCGCGTCTCGAAGGATACGGTCACTCGCGGAAAAG

rkpA1021 7255 ATATTGAAGGAGATCGGCCTCGACAGCCTCATGGCCGTTGAACTCGGCATGAGTTTCCAG
rkpA-F   7236 ATATTGAAGGAGATCGGCCTCGACAGCCTCATGGCCGTTGAACTCGGCATGAGTTTCCAG

rkpA1021 7315 CAGAACACCGGCTTCGACATGCCTTTGAGCGGCGTCGCGGACAATACCACGGTCGGCGAC
rkpA-F   7296 CAGAACACCGGCTTCGACATGCCTTTGAGCGGCGTCGCGGACAATACCACGGTCGGCGAC

rkpA1021 7375 ATCGCGCGCAAGCTCCATGAAAAAGTAAGCAAGC_GTGACCAGGGCAACGAGAACGAGAA
rkpA-F   7356 ATCGCGCGCAAGCTCCATGAAAAAGTAAGCAAGC_GTGACCAGGGCAACGAGAACGAGAA

rkpA1021 7434 CGAGGCCGGTGACGACAAGCTTGTAAACAGAACTCACGCAGCGTCATGTCGGAGCGGACAA
rkpA-F

rkpA1021 7494 GGAAAAAGCACTGAGTCAATGA
rkpA-F

```

## **7.6 Danksagung**

An erster Stelle möchte ich meiner Familie danken, die mich in all den Jahren immer unterstützt hat. Sie haben meinen Ankündigungen, dass ich die Promotion schon noch beenden werde, immer geglaubt, selbst zu Zeiten zu denen ich kaum noch daran geglaubt habe. Besonders möchte ich meine Frau Julie und unseren Sohn Douglas hervorheben, die mir in den besonders schweren Zeiten immer weiter geholfen haben.

Mein besonderer Dank gilt der ausgezeichneten Arbeitsgruppe Geiger in Berlin. Die Truppe war, sowohl was fachliche als auch menschliche Qualitäten betrifft, die absolute Spitze. Neben Prof. Dr. Otto Geiger, Prof. Dr. Isabel Maria Lopez-Lara, Viola Khodaverdi, Karel deRudder, Dr. Christian „useless“ Sohlenkamp und Frau Dr. Barbara Weissenmayer soll auch das externe Mitglied Dr. Max Schobert genannt werden. In dieser Arbeitsgruppe hab ich sehr viel gelernt und sehr gute Freunde gefunden. Durch die Querelen nach dem Abwandern der Arbeitsgruppe Geiger nach Mexiko geriet meine Person und damit auch die Zukunft dieser Arbeit zwischen die Fronten eines persönlichen Streites. An diesem Punkt wäre die Fertigstellung der Promotion nicht mehr möglich gewesen, wenn nicht Dr. Hans von Döhren eingegriffen hätte. Er hat der DFG gegenüber die Verantwortung für das Projekt übernommen und mir einen neuen Arbeitsplatz zu Verfügung gestellt. Dafür gilt ihm mein besonderer Dank. Mein Dank gilt ebenfalls den Mitarbeitern von Prof. Dr. Ulrich Szewzyk, die mir, wenn immer es möglich war, mit Geräten und Ratschlägen ausgeholfen haben.

Abschließend möchte ich mich noch bei meinen neuen Kollegen der AG Geßner in der Charité bedanken, die mir den nötigen Freiraum gegeben haben diese Arbeit nach langer Zeit doch noch fertigzustellen.

## 7.7 Eigene Veröffentlichungen

### 7.7.1 Teile der Arbeit wurden bereits veröffentlicht

**Epple, G., van der Drift K. M., Thomas-Oates, J. E. & Geiger, O. (1998).** Characterization of a novel acyl carrier protein, RkpF, encoded by an operon involved in capsular polysaccharide biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* **180**, 4950-4954.

### 7.7.2 Sonstige Veröffentlichungen

**Praus, M., Kettelgerdes, G., Baier, M., Holzhütter, H. G., Jungblut, P. R., Maissen, M., Epple, G., Schleuning, W. D., Köttgen, E., Aguzzi, A. & Geßner, R. (2003).** Stimulation of plasminogen activation by recombinant cellular prion protein is conserved in the NH(2)-terminal fragment PrP23-110. *Thromb. Haemost.* **89**, 812-819.

**Epple, G., Langfeld, K., Baier, M., Holzhütter, H. G., Schleuning, W. D., Köttgen, E., Geßner, R. & Praus, M. (2004).** Both Lysine-Clusters of the NH2-Terminal Prion-Protein Fragment PrP23-110 are Essential for t-PA mediated Plasminogen Activation. *Thromb. Haemost.* **91**, 465-472.

**Epple, G., Schleuning, W. D., Kettelgerdes, G., Köttgen, E., Geßner, R. & Praus, M. (2004).** Prion protein stimulates tissue-type plasminogen activator-mediated plasmin generation via a lysine-binding site on kringle 2. *J. Thromb. Haemost.* **2**, 962-968.

**Ney, A., Booms, P., Epple, G., Mörgelin, M., Guo, G., Kettelgerdes, G., Geßner, R. & Robinson, P.N. (2005).** Self-Association of the C-terminal Lectin Domain of Versican. *FEBS Letters* submitted.

## 7.8 Lebenslauf

### Persönliche Daten

Name	Guido Franz Epple
Geburtsdatum und -ort	07.04.1967, Paderborn

### Schulen:

Aug.1973 - Juni 1977	Grundschule Elisabeth-Schule, Paderborn
----------------------	--

---

Aug.1977 - Juni 1986      Gymnasium  
Goerdeler-Gymnasium, Paderborn

**Berufsausbildung:**

Sep.1986 - Jan.1989      Berufsausbildung zum Kälteanlagenbauer  
Firma Dipl.Ing. Bruno Zimmer, Bielefeld

**Berufstätigkeit:**

Febr. - April 1989      Firma Dipl.Ing. Bruno Zimmer, Bielefeld

Mai - Sep. 1989      Firma Polarmatic Kälte und Klimatechnik  
GmbH, Paderborn

**Akademische Laufbahn:**

Okt. 1989 – Juni 1997      Studium der Biotechnologie an der TU Berlin

17.06.1997      Diplom; Note : sehr gut

**Dissertation**

Juli 1997 – März 2000      Wissenschaftlicher Mitarbeiter an der  
TU Berlin im DFG-Projekt „Biosynthese  
kapsulärer Polysaccharide in *Sinorhizobium  
meliloti*“ im FG Technische Biochemie  
(AG Geiger) des Instituts für Biotechnologie

April. 2000 – Juli 2001      Wissenschaftlicher Mitarbeiter an der TU  
Berlin, Fortsetzung des DFG-Projekt  
„Biosynthese kapsulärer Polysaccharide in  
*Sinorhizobium meliloti*“ am Max-Volmer-

Institut für Biophysikalische Chemie und  
Biochemie (AG von Döhren)

**Wissenschaftliche Tätigkeit:**

Seit 01.04.2002      Wissenschaftlicher Mitarbeiter an der Charité  
Berlin, Institut für Laboratoriumsmedizin und  
Pathobiochemie (AG Geßner)