# Untersuchungen zur Funktion des CRISPR/Cas Systems von *E. coli*: Isolierung und Charakterisierung spezifischer Komponenten des Cascade-Komplexes und Aufbau eines Systems zur Spacer Integration

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Veronica Hermanns aus Düsseldorf

Düsseldorf, Mai 2014

aus dem Institut für Physikalische Biologie der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Heinrich-Heine-Universität Düsseldorf

Referent:

Prof. Dr. Rolf Wagner

Korreferent:

Prof. Dr. William Martin

Tag der mündlichen Prüfung: 18.06.2014

### Zusammenfassung

Eine besonders wirkungsvolle Art der Immunabwehr in Prokaryoten ist das als CRISPR/Cas bezeichnete System, was für *Clustered Regularly Interspaced Short Palindromic Repeats* und den *CRISPR associated proteins* steht. Es verleiht Bakterien eine adaptive und vererbbare Immunität gegenüber Infektionen durch Fremd-DNA oder –RNA in Form von Viren oder Phagen. Aufgebaut ist es aus Subtyp-spezifischen Cas-Proteinen und einem nicht-kodierenden DNA-Bereich, dem CRISPR-Array. Dieser umfasst Regionen im Wirts-Genom, die aus palindromischen Sequenzwiederholungen (Repeats) bestehen und von variablen Spacer-Bereichen unterbrochen werden. Nach einer Erstinfektion der Bakterienzelle wird die fremde DNA als solche erkannt und von den Cas-Proteinen in kleine DNA-Fragmente gespalten. Mit dem Einbau dieser Fremd-DNA als neuer Spacer in den eigenen CRISPR-Array erfolgt die Immunisierung der Zelle. Für diesen, auch als Adaptation bezeichneten Mechanismus, sind neben den zwei ubiquitären Cas-Nukleasen Cas1 und Cas2 die Präsenz einer Leaderregion, sowie einer Repeat-Einheit essentielle Bestandteile.

In dieser Arbeit wurden Untersuchungen zur Aufklärung des Mechanismus der Spacerintegration in *E. coli* durchgeführt. Der verwendete Bakterienstamm BL21 AI besitzt einen CRISPR II Array, dem die für die Immunisierung essentiellen cas-Gene fehlen. Für eine detaillierte Charakterisierung des Adaptationsmechanismus stand die Entwicklung eines Plasmid-basierten Integrations-Assays im Vordergrund. Dazu musste zunächst ein Plasmid konstruiert werden, dass neben den fehlenden *cas*-Genen, einen synthetischen CRISPR-Array trägt und als Grundlage für den Aufbau eines Plasmid-Systems zur Spacerintegration dient. Die erfolgreiche Etablierung des Plasmid-Systems konnte durch PCR-Amplifikationsstudien, die in vivo den Einbau von neuen Spacern in den plasmidären CRISPR-Array zeigten, verifiziert werden. Zudem konnten neu integrierte Spacersequenzen mittels *single-colony*-PCR (sc-PCR) identifiziert und analysiert werden. Die Einführung gezielter Punktmutationen innerhalb der ersten Repeatsequenz zeigte zum ersten Mal eine sequenzspezifische Spaltung durch die Cas1 und Cas2 Proteine. In weiterführenden Southern Blot Analysen konnte die Hypothese eines sequenzspezifisch-, versetzten Schnittes an beiden Strängen der ersten Repeatsequenz in vivo bestätigt werden und die Bildung von Reaktions-Intermediaten (Zwischenstufen) detektiert werden. Die Analyse dieser Intermediate weist auf eine gekoppelte Spaltungs-Ligations-Reaktion während der Integration hin.

In einem weiteren Teil dieser Arbeit wurden drei der fünf den Cascade-Komplex bildenden Cas-Proteine (Cse1, Cse2 und Cas7) affinitätschromatographisch aufgereinigt und durch *in vitro* Bindestudien mit doppel- und einzelsträngigen Nukleinsäuren analysiert. Dabei konnte eine Bindung des Cse1-Proteins nicht nur an kurze doppelsträngige und einzelsträngige DNA, sondern auch mit einer hohen Affinität an pre-crRNA von *E. coli* gezeigt werden. Das Cse2 Protein hingegen zeigte eine selektive Bindung an Einzelstrang-DNA.

### Summary

A particularly effective type of immune system in prokaryotes is known as CRISPR/Cas system, which stands for *Clustered Regularly Spaced Interspaced Short Palindromic Repeats* and the CRISPR *associated proteins*. It provides prokaryotes with an adaptive and heritable immunity against infection by foreign DNA or RNA (e.g. phages or plasmids). The system consists of subtype-specific Cas proteins and a non-coding DNA region, termed CRISPR array consisting of palindromic repeats that are interspaced by variable spacer sequences. After initial infection of the bacterial cell the invading DNA is recognized and cleaved by Cas proteins into small DNA fragments. Integration of these foreign DNA-derived fragments into the CRISPR array as new spacers results in the immunization of the cell. This process, also referred as adaptation, depends on the presence of the two ubiquitous nucleases Cas1 and Cas2 but also on the presence of a leader region followed by at least one repeat sequence.

To elucidate the mechanism of spacer integration in *E. coli* in more detail, *in vivo* spacer acquisition assays were performed using *E. coli* BL21 AI. This bacterial strain contains a CRISPR II array but lacks all cas genes, including cas1 and cas2, which are required for the immunization. In order to study the mechanism of the adaptation stage, it was necessary to develop a plasmid-based integration assay. Therefore, a plasmid was constructed, which contains the essential cas1 and cas2 genes and a synthetic CRISPR array. The acquisition of new spacers into the plasmid-located array could be verified by in vivo assays. Sequences of the newly aquired spacers were determined by *single-colony* PCR reactions (sc-PCR). Selective point mutations within the first repeat sequence were introduced and analyzed for their effects on spacer aquisition. The obtained results demonstrated a sequence-specific cleavage by the Cas1 and Cas2 proteins for the first time. By subsequent Southern blot analyses, it could be demonstrated that during the integration of new spacers reaction intermediates are formed. A thorough analyses of these spacer integration intermediates revealed that Cas1 and Cas2 introduce staggered nicks specifically at the 5'-ends of the first repeat. Moreover, this staggered nicking is accompanied by the ligation of a new spacer DNA, suggesting a coupled cleavage-ligation (integrase) reaction performed by Cas1 and Cas2 proteins.

In the second part of this work, the three Cas proteins Cse1, Cse2 and Cas7, which are constituents of the Cascade complex, were purified by affinity chromatography and analyzed for their ability to bind double- or single-stranded nucleic acids by mobility shift assays. The results showed a binding of the Cse1 protein to short double-stranded and single-stranded DNA. Moreover, a specific binding of Cse1 to the pre-crRNA with high affinity could be demonstrated. In contrast, the Cse2 protein interacted selectively with single-stranded DNA.

## Inhaltsverzeichnis

1 Einleitung	1
1.1 Die Entdeckung des CRISPR/Cas-Systems	2
1.1.1 Die Struktur des CRISPR-Locus	2
1.2 Der CRISPR-Wirkungsmechanismus – die drei Phasen der Immunabwehr	4
1.2.1 Immunisierungsphase – Integration neuer Spacersequenzen	5
1.2.2 Expression von Komponenten für die Immunabwehr	6
1.2.3 Erkennung und Degradation von Fremd-Nukleinsäuren	8
1.3 Fremd-DNA Erkennung durch Cas-Protein Komplexe	9
1.3.1 Der <i>E. coli</i> Cascade-Komplex	
1.3.2 Erkennung invasiver Fremd-DNA	12
1.4 CRISPR Adaptation	13
1.4.1 Spacer-uptake in <i>E. coli</i> (Typ I-E Systeme)	13
1.4.2 Spacer-uptake in Typ II-Systemen	15
1.5 Industrielle Anwendung: CRISPR-basierte Technologie	17
1.6 Fragestellung und Konzeption der Arbeit	20
2 Ergebnisse	22
2.1 Aufbau eines CRISPR-basierten Plasmid-Systems in <i>E. coli</i>	22
2.1.1 Klonierung des pCR003 WT-Plasmids	23
2.1.2 Detektion der Spaceraufnahme mittels <i>single-colony</i> PCR (sc-PCR)	24
2.1.3 Analyse von neu erworbenen Spacersequenzen	26
2.2 Mutagenese der ersten Repeatsequenz des Plasmid-basierten CRISPR-Arra	ays 29
2.2.1 Positionswahl und Einführung von Mutation RM1	30
2.2.1.1 Funktionelle Analyse der pCR003 RM1 Mutante <i>in vivo</i>	
2.2.2 Einführung und <i>in vivo</i> Analyse von Mutation RM1-R	
2.2.3 Einführung und <i>in vivo</i> Analyse der Mutation RM2	35
2.3 Charakterisierung des Adaptations-Mechanismus	
2.3.1 Detektion von Spacerintegrations-Intermediaten des Plasmid-basierten O	CRISPR-
Arrays	
<ul><li>2.3.2 Verifizierung der <i>in vivo</i> Intermediate durch Southern Blot Analysen</li><li>2.3.3 Bestimmung der DNA-Intermediate durch Primer-Extension mit Sequenase</li></ul>	39 2.045
2.4 Aufbau eines heterologen Integrationssystems in <i>E. coli</i>	

	2.4.1 Klonierung eines Plasmid-basierten Vektors für eine heterologe Spacerintegrati	on
:	2.4.2 Analytische Proteinexpression von <i>S. agalactiae</i> Cas1 und Cas2 in BL21 AI	49
2.5	Expression und Isolation von Komponenten des E. coli Cascade-Komplexes	54
:	2.5.1 Klonierung und analytische Überexpression der fünf Cas-Proteine	54
:	2.5.2 Aufreinigung von CasC (Cas7)/pWUR400 mit fusioniertem StrepII-Tag	56
:	2.5.3 Aufreinigung von CasC (Cas7) mit fusioniertem StrepII-Tag	58
:	2.5.4 Aufreinigung von CasA (Cse1) mit fusioniertem StrepII-Tag	61
	2.5.5 Aufreinigung von CasB (Cse2) mit fusioniertem StrepII-Tag	63
	2.5.6 Ist die assoziierte crRNA im isolierten Cascade-Komplex nachweisbar?	65
:	2.5.7 In vitro Retardierungsanalysen	66
	2.5.7.1 Bindeanlaysen von CasA, CasB und CasC an doppelsträngige DNA-Fragment	e 66
	2.5.7.2 Interaktion von CasA, B und C an Einzelstrang-DNA	69
	2.5.7.3 Bindeanlaysen von CasA, CasB und CasC mit einzel- und doppelsträngigen R Fragmenten	(NA- 71
2	Distance	75
3	DISKUSSION	.75
3.1	Etablierung eines Plasmid-basierten Integrations-Assays	75
3.2	Charakterisierung der Spacerintegration in Typ I-E CRISPR/Cas-Systemen	78
	3.2.1 Rolle der ersten Leader-proximalen Repeatsequenz während der Integration no	euer
	Spacer	79
:	3.2.2 Aufklärung des Adaptationsmechanismus in <i>E. coli</i>	81
	3.2.2.1 Verifizierung der Cas1/Cas2-vermittelten Spaltungs-Ligations-Reaktion	81
	3.2.2.2 Sequenzierung der detektierten DNA-Intermediate	84
3.3	Charakterisierung der Spaceraufnahme in Typ II-A Systemen	84
	3.3.1 Aufbau eines heterologen Integrationssystems in <i>E. coli</i>	85
3.4	Expression und Aufreinigung der <i>E. coli cascade</i> Gene	87
:	3.4.1 Analytische Expression der <i>cascade</i> Gene	87
	3.4.2 Präparative Aufreinigung der Cascade-Proteine	88
3.5	Interaktion von CasA, CasB und CasC mit Nukleinsäuren	90
	3.5.1 Bindung an Einzel- und Doppelstrang-DNA	90
	3.5.2 Ist RNA ein potentieller Bindungspartner?	91
3.6	Ausblick	93
4	Material	.95
4.1	Allgemeines	95
4.2	Verwendete Bakterienstämme und Plasmide	95

4.2.1 Escherichia coli Stämme	95
4.2.2 Plasmide	96
43 DNA-Fragmente	98
	00
4.3.1 Fragmente für Retardierungen	
4.4 Nukleinsäuren	
4.4.1 Oligonukleotide	
4.4.2 Nukleotide	106
4.4.3 Molekulargewichtsmarker	106
4.5 Proteine	106
451 Restriktionsondonuklasson	106
4.5.2 Polymerasen	
4.5.3 Enzyme und sonstige Proteine	107
	100
4.6 Puffer und Medien	108
4.6.1 Standardpuffer	108
4.6.2 Medien	109
4.7 Feinchemikalien	110
48 Verschiedenes	111
+.0 verschleuenes	
5 Methoden	112
<ul><li>5 Methoden</li><li>5.1 Mikrobiologische Methoden</li></ul>	<b>112</b> 112
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> </ul>	<b>112</b> 112
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li></ul>	<b>112</b> 112 112 112
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li></ul>	<b>112</b> 112 112 112 112
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> <li>5.1.2 Kultivierung von Bakterienstämmen</li> <li>5.1.2.1 Anzucht auf Agarplatten</li> <li>5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)</li> </ul>	<b>112</b> 112 112 112 112 112
<ul> <li>5 Methoden</li></ul>	<b>112</b> 112 112 112 112 112 113
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> <li>5.1.2 Kultivierung von Bakterienstämmen</li> <li>5.1.2.1 Anzucht auf Agarplatten</li> <li>5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)</li> <li>5.1.2.3 Anlegen von Glycerinkulturen (Glycerinstocks)</li> <li>5.1.2.4 Erstellung von Wachstumskurven und Berechnung von Wachstumsrate</li> </ul>	<b>112</b> 112 112 112 112 113 en113
<ul> <li>5 Methoden</li></ul>	<b>112</b> 112 112 112 112 113 en113 114
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 113 en113 114 114
<ul> <li>5 Methoden</li></ul>	<b>112</b> 112 112 112 112 113 113 114 114
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> <li>5.1.2 Kultivierung von Bakterienstämmen</li> <li>5.1.2.1 Anzucht auf Agarplatten</li> <li>5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)</li> <li>5.1.2.3 Anlegen von Glycerinkulturen (Glycerinstocks)</li> <li>5.1.2.4 Erstellung von Wachstumskurven und Berechnung von Wachstumsrate</li> <li>5.1.3 Herstellung chemisch kompetenter <i>E. coli</i> Zellen für die Transformation</li> <li>5.1.4 Transformation kompetenter Zellen mit Vektor-DNA durch Hitzeschock</li> <li>5.2 Molekularbiologische Methoden</li> </ul>	<b>112</b> 112 112 112 112 113 113 114 114 114
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> <li>5.1.2 Kultivierung von Bakterienstämmen</li> <li>5.1.2.1 Anzucht auf Agarplatten</li> <li>5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)</li> <li>5.1.2.3 Anlegen von Glycerinkulturen (Glycerinstocks)</li> <li>5.1.2.4 Erstellung von Wachstumskurven und Berechnung von Wachstumsrate</li> <li>5.1.3 Herstellung chemisch kompetenter <i>E. coli</i> Zellen für die Transformation</li> <li>5.1.4 Transformation kompetenter Zellen mit Vektor-DNA durch Hitzeschock</li> <li>5.1.5 Sequenzierung von Plasmid-DNA</li> <li>5.2 Molekularbiologische Methoden</li> <li>5.2.1 Konzentrationsmessungen</li> </ul>	112 112 112 112 112 112 113 en113 en113 114 114 114
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 113 113 113 114 114 115 115
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 113 en113 en113 en114 114 114 115 115 115
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 112 112 112 113 113 114 114 115 115 115 115 116
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 113 en113 en113 en114 114 114 115 115 115 115 116 116
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 112 112 112 113 en113 en113 114 114 115 115 115 115 115 116 116 prep-Kit
<ul> <li>5 Methoden</li></ul>	112 112 112 112 112 112 113 en113 en113 en113 114 114 115 115 115 115 115 116 prep-Kit 116
<ul> <li>5 Methoden</li> <li>5.1 Mikrobiologische Methoden</li> <li>5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien</li> <li>5.1.2 Kultivierung von Bakterienstämmen</li> <li>5.1.2.1 Anzucht auf Agarplatten</li> <li>5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)</li> <li>5.1.2.3 Anlegen von Glycerinkulturen (Glycerinstocks)</li> <li>5.1.2.4 Erstellung von Wachstumskurven und Berechnung von Wachstumsrate</li> <li>5.1.3 Herstellung chemisch kompetenter <i>E. coli</i> Zellen für die Transformation</li> <li>5.1.4 Transformation kompetenter Zellen mit Vektor-DNA durch Hitzeschock</li> <li>5.1.5 Sequenzierung von Plasmid-DNA</li> <li>5.2 Molekularbiologische Methoden</li> <li>5.2.1.1 Konzentrationsbestimmung von Nukleinsäuren</li> <li>5.2.1.2 Bestimmung der Zelldichte von Flüssigkulturen</li> <li>5.2.1.3 Konzentrationsbestimmung von Proteinen</li> <li>5.2.2.1 Isolation von Plasmid-DNA mit Hilfe des GeneJET Plasmid Mini ("Minipräp")</li> <li>5.2.2.2 Plasmidisolation mittels dem Qiagen Plasmid Midi Kit ("Midipräp")</li> </ul>	112 112 112 112 112 112 112 112 112 113 113 114 114 115 115 115 115 115 116 prep-Kit 116 117

5.2.3 Reini	gung und Konzentrierung von Nukleinsäuren	118
5.2.3.1 P	henol/Chloroform-Extraktion	118
5.2.3.2 E	thanol-Präzipitation von Nukleinsäuren	119
5.2.3.3 D	ialyse von Nukleinsäuren	119
5.2.3.4 Is	olation von Nukleinsäuren aus Agarosegelen durch Glaswoll-Elution	120
5.2.3.5 Is	solation von DNA-Fragmenten aus präparativen Agarose-gelen mit Hilfe	des
Ν	/izard® SV Gel and PCR Clean-up Systems	120
5.2.3.6 A	ufreinigung von DNA mit dem QIAquick PCR Purification Kit	121
5.2.4 Gelelo	ektrophoresen	121
5.2.4.1 A	garose-Gelelektrophorese	121
5.2.4.2 N	ative Polyacrylamidgelelektrophorese (PAGE)	122
5.2.4.3 D	enaturierende Polyacrylamidgelelektrophorese (dPAGE)	122
5.2.5 Nach	weismethoden von Nukleinsäuren	123
5.2.5.1 E	thidiumbromidfärbung von Nukleinsäuren	123
5.2.5.2 N	achweis über Autoradiographie	123
5.2.5.3 Si	ilberfärbung von Nukleinsäuren	123
5.2.6 Enzyı	matische Reaktionen	124
5.2.6.1 R	estriktionshydrolysen von DNA	124
5.2.6.2 P	olymerasekettenreaktion (PCR)	124
5.2.6.3 L	igation von DNA-Fragmenten	125
5.2.6.4 5	-Phosphorylierung von DNA-Fragmenten	126
5.2.6.5 0	rtsgerichtete <i>in vitro</i> Mutagenese	126
5.2.6.6 St	ingle-colony PCR (sc-PCR)	126
5.2.6.7 P	hosphorylierung von DNA-Oligonukleotiden	127
([	5´-Endmarkierung mit [γ- <sup>32</sup> P]-ATP)	127
5.2.6.8 R	adioaktive 3'-Endmarkierung von RNA	128
5.2.6.9 R	adioaktive Markierung von DNA-Fragmenten	128
E 2 Drotoinio	alation und Analusa	100
5.3 Proteinise	Diation und Analyse	128
5.3.1 Prote	inisolation im analytischen Maßstab	128
5.3.2 Prote	inisolation im präparativen Maßstab (Aufreinigung der E. coli Cas-Prot	eine
Cse1,	Cse2, Cas7, Cas5 und Cas6e)	129
5.3.3 Affini	tätschromatographie	130
5.3.3.1 Sa	äulenchromatographie mittels Strep-Tactin Resin	130
5.3.3.2 Ir	nmobilisierte Affinitätschromatographie mit fusioniertem StrepII-Tag	132
5.3.4 Aufre	inigung und Konzentrierung von Proteinen	133
5.3.4.1 D	ialyse von Proteinen	133
5.3.4.2 K	onzentrierung von aufgereinigten Proteinen mittels Amicon-Zentrifugatio	n
		133
5.3.5 Gelel	ektrophorese von Proteinen	133
5.3.5.1 D	iskontinuierliche SDS-PAGE	133
5.3.6 Nach	weis von Proteinen in Polyacrylamidgelen	134
5.3.6.1 C	oomassie-Färbung	134
5.4 Spezielle	molekularbiologische Methoden	135
5.4.1 In vit	ro Transkription zur präparativen Gewinnung von pre-crRNA (RiboMAX) .	135
5.4.2 South	nern Blot Analysen	136
5.4.3 Seque	enzierungsreaktion mit Sequenase 2.0 Polymerase	137

	5.4.4 Retardierungsanalysen (Gelshift-Experimente)	
6	Anhang	140
7	Literaturverzeichnis	
8	Abkürzungsverzeichnis	
9	Eidesstattliche Erklärung	

### 1 Einleitung

In der Natur herrscht ein reges Zusammenleben von Eukaryoten, Bakterien und Archaeen, die unterschiedliche Lebensräume besiedeln und ständigen Veränderungen ausgesetzt sind. Diese gegenseitige Wechselbeziehung zwischen den Organismen ist zum einen von großer Bedeutung für das Verständnis von grundlegenden Mechanismen in den Lebenswissenschaften, ist aber auch wegen ihrer direkten Auswirkung in der Biotechnologie und Lebensmittelindustrie, wichtiger Bestandteil der heutigen Forschung.

Prokaryoten unterliegen wie Eukaryoten einer ständigen Bedrohung durch Viren oder Phagen. Daher stellt die Anwesenheit der Viren eine anhaltende Gefahr für das Überleben von Bakterien und Archaeen dar, wodurch sie verschiedene Abwehrmechanismen, die auf mehreren Stufen einer Infektion wirksam sind, entwickelt haben. So führte eine rege und wechselseitige Anpassung von Angriff- und Abwehrmechanismen zu einer -wenn auch nicht friedlichen- Koadaptation von Prokaryoten und Viren (Marchfelder, Maier et al. 2013).

Bakterien und Archaeen haben im Laufe der Evolution mehrere Mechanismen entwickelt, um sich gegen parasitäre Eindringlinge, wie Fremd-DNA, zu schützen. Veränderungen der Oberflächenstrukturen (Rezeptoren) erschweren das Andocken, sowie das Eindringen von Fremd-DNA der Phagen in die Wirtszelle. Durch Restriktions-Modifikations-Systeme können Bakterien bereits eingedrungene Fremd-DNA zerstören oder sich selbst opfern (bakterieller Suizid), um die Vermehrung parasitärer Viren und damit die Zerstörung der eigenen Population zu verhindern (Westra, Swarts et al. 2012). Auch verfügen 40% der Bakterien und über 90% der Archaeen über RNA-basierte adaptive Schutzmechanismen, wie sie auch in Eukaryoten zu finden sind (Al-Attar, Westra et al. 2011). In Prokaryoten wird diese besonders wirkungsvolle Art der Immunabwehr als CRISPR-Cas bezeichnet, was für *Clustered Regularly Interspaced Short Palindromic Repeats* und den CRISPR *associated* proteins steht. Organismen, die solch ein CRISPR System enthalten, können durch Anwesenheit von einem oder mehreren dieser CRISPR-Loci auf ihrem Genom gekennzeichnet sein. Die Funktionsweise des Abwehrsystems ähnelt im Prinzip dem eines eukaryotischen Immunsystems (Goren, Yosef et al. 2012) und verleiht Resistenz gegenüber Bakterien, die eine Begegnung mit Phagen oder anderen invasiven genetischen Elementen überlebt haben (Wagner and Pul 2012). Das System umfasst Regionen im Wirts-Genom, die aus kurzen, oft palindromischen Sequenzwiederholungen (Repeats) bestehen, die wiederum von variablen Spacersequenzen unterbrochen werden (Barrangou, Fremaux et al. 2007, Brouns, Jore et al. 2008) und den Zellen nach Infektion mit Fremd-DNA eine vererbbare Immunität gegenüber dieser DNA verleihen (Sorek, Kunin et al. 2008, van der Oost, Jore et al. 2009). Wird die Wirtszelle mit Fremd-DNA infiziert und als neuer Spacer dem CRISPR-Array hinzugefügt, vermittelt diese bei erneuter Infektion die Erkennung der fremden, homologen Sequenzen und initiiert die Neutralisierung der parasitären DNA. Die erworbene Fremd-DNA bleibt im Wirtsgenom enthalten und fungiert somit als

"immunologisches Gedächtnis", bei dem die Resistenz auch an die Nachkommen vererbt wird (Marchfelder, Maier et al. 2013).

### 1.1 Die Entdeckung des CRISPR/Cas-Systems

Die Existenz der nicht kodierenden DNA-Abschnitte die heute als CRISPR bekannt sind, wurden bereits 1987 bei der Sequenzierung des Genoms von *E. coli* entdeckt (Ishino, Shinagawa et al. 1987). Bis heute konnten ähnliche Muster in sequenzierten Genomen anderer Bakterien und Archaeen identifiziert werden. Als dann 2005 aufgedeckt wurde, dass viele Spacer-Sequenzen Homologien zu Viren- und Plasmid-DNA aufweisen, führte dies zur Hypothese, dass die Funktion von CRISPR darin besteht, den Organismus gegen Fremd-DNA zu verteidigen (Mojica and Garrett 2013). Experimentell bestätigt wurde dies dann anhand von Beobachtungen, die in dem Milchsäurebakterium *Streptococcus thermophilus* gemacht wurden. Nach Vireninfektion zeigte ein kleiner Teil der Zellen den Erwerb neuer, viren-spezifischer Spacersequenzen im eigenen Genom, was zu einer Resistenz gegenüber diesen Viren bei einer Zweitinfektion führte. Eine Deletion dieser neu erworbenen Fremd-Sequenzen wiederum resultierte in einem Verlust der Resistenz, was den Zusammenhang zwischen den integrierten Spacern und der Resistenz gegenüber Phagen erstmals experimentell belegte (Barrangou, Fremaux et al. 2007).

### 1.1.1 Die Struktur des CRISPR-Locus

Der typische CRISPR-Locus besteht aus palindromischen Repeatsequenzen, die durch unterschiedliche Spacersequenzen (21-72) unterbrochen sind und aus der Fremd-DNA stammen (Abbildung 1.1). Je nach CRISPR-System und Organismus kann dabei die Anzahl der einzelnen Repeat-Spacer Einheiten von einigen wenigen bis hin zu mehreren hunderten variieren. Auch die Länge der palindromischen CRISPR Repeats kann zwischen 23 bis 55 Nukleotiden differieren (Mojica and Garrett 2013).

#### Abbildung 1.1: Schematische Anordnung eines CRISPR-Locus.

cas-Gene

Zu sehen ist eine schematische Darstellung eines CRISPR-Locus, bestehend aus einer chromosomalen Anordnung unterschiedlicher Spacersequenzen (S, farbige Quadrate), flankiert von identischen Repeatsequenzen (R, graue Rauten), deren Anzahl variieren kann. Dem CRISPR-Array vorangestellt ist die charakteristische AT-reiche Leader-Region (schwarz), an der die Transkription startet. In braun nachgeschaltet sind die CRISPR-assoziierten *cas* Gene gezeigt, die für die Cas-Proteine codieren und eine wichtige Rolle in der CRISPR-Aktivität spielen.

Leader

Während innerhalb derselben CRISPR Struktur die sich wiederholende Repeatsequenz erhalten bleibt, variiert sie in verschiedenen CRISPR/Cas-Systemen stark (Bolotin, Quinquis et al. 2005, Mojica, Diez-Villasenor et al. 2005). Ein weiteres wichtiges Charakteristikum stellt der bis zu 500 bp große, meist AT-reiche Leaderbereich innerhalb des CRISPR-Locus dar (Jansen, Embden et al. 2002, Pul, Wurm et al. 2010). Dieser fungiert als Promotor und bewirkt die Transkription in eine lange Vorläufer-RNA (precursor-crRNA, pre-crRNA), die dann mit Hilfe der Cas-Proteine prozessiert wird. Die hochkonservierten und GC-reichen CRISPR Repeats bilden sogenannte RNA Haarnadel-Sekundärstrukturen aus, die wiederum als Erkennungssequenzen für eine endoribonukleolytische Prozessierung der pre-crRNA zu den kleinen, reifen CRISPR RNAs (crRNAs) dienen (Kunin, Sorek et al. 2007, Bhaya, Davison et al. 2011). Die eigentliche Prozessierung erfolgt durch einen Multiproteinkomplex, der durch Assemblierung mit den reifen crRNAs essenzieller Bestandteil bei der Erkennung und Degradierung von Fremd-DNA bzw. -RNA ist. Gleichzeitig enthält die Leaderregion Bindestellen für Regulations- und Rekombinationsproteine und ist unmittelbar an der Integration neuer Spacersequenzen beteiligt (Yosef, Goren et al. 2012). Ein weiterer Promotor upstream der cas- (CRISPR-assoziierten) Gene lokalisiert, ist für die Transkription dieser spezifischen Proteingene zuständig, die meist in Operons organisiert sind. Diese Cas-Proteine bestehen häufig aus funktionellen Domänen und weisen Nuklease-, Helikase- und/oder Polymeraseaktivitäten auf. Die Genprodukte werden als Core-Cas Proteine Cas1-6 bzw. als subtypspezifische Genprodukte klassifiziert (Haft, Selengut et al. 2005). Bisher sind mehr als 40 cas-Genfamilien bekannt, die für Proteinkomponenten des CRISPR/Cas-Systems kodieren und in unterschiedlicher Art und Weise Einfluss auf die Immunabwehr nehmen (Makarova, Haft et al. 2011). Die in verschiedene CRISPR/Cas-Subtypen erfolgte Unterteilung 2011 durch Homologievergleiche. Danach wurden zehn unterschiedliche Subtypen, die in drei Hauptklassen zusammengefasst werden vorgeschlagen: Typ I, II und III. Ein typisches Merkmal aller CRISPR/Cas-Systeme ist das Vorhandensein der zwei hochkonservierten *cas* Gene *cas1* und *cas2*, die als einzige *cas* Gene Bestandteil in jedem dieser CRISPR/Cas-Subtypen sind. Während Typ I und III CRISPR/Cas-Systeme sowohl in Bakterien als auch Archaeen vertreten sind, konnte Typ II bisher nur in Bakterien identifiziert werden (Makarova, Haft et al. 2011).

### 1.2 Der CRISPR-Wirkungsmechanismus – die drei Phasen der Immunabwehr

Die CRISPR-vermittelte Immunabwehr von Prokaryoten gegenüber Fremd-DNA wird in drei Phasen eingeteilt: (1) Immunisierung, (2) Expression und Prozessierung, sowie (3) Interferenz (Abbildung 1.2). Während der ersten Phase werden neue Spacersequenzen aus der injizierten, parasitären Fremd-DNA in den eigenen CRISPR-Array eingebaut. Der zweite Schritt beinhaltet die Expression der Cas-Proteine und des CRISPR-Arrays, sowie die Prozessierung der transkribierten Vorläufer-RNA (pre-crRNA). In Folge entstehen die kleinen, reifen crRNAs, die komplementär zu Fremd-DNA Abschnitten sind. Im dritten und letzten Schritt werden bekannte Fremd-DNA Abschnitte erkannt und durch einen crRNAvermittelten Mechanismus degradiert.



#### Abbildung 1.2: Übersicht über die unterschiedlichen Phasen der CRISPR/Cas-Abwehr.

(1) Zeigt die Immunisierungsphase in der Spacersequenzen aus Nukleinsäuren von Viren oder konjugativen Plasmiden am Leader-proximalen Ende in den CRISPR-Array integriert und von den Repeats flankiert werden. In der zweiten Phase (2) wird die Vorläufer-precursor RNA (pre-crRNA) transkribiert und durch die Cas-Proteine zu reifen crRNAs prozessiert. In der dritten Phase (3) findet eine RNA-vermittelte DNA-Interferenz durch den mit crRNA beladenen Cas-Proteinkomplex statt und die Fremd-DNA wird degradiert (entnommen aus (Bhaya, Davison et al. 2011)).

### 1.2.1 Immunisierungsphase – Integration neuer Spacersequenzen

Kommt es durch Viren oder Phagen zu einer Infektion der Zelle, wird die erste Phase des CRISPR/Cas-Abwehrsystems eingeleitet. Während dieser Phase nehmen Archaeen und Bakterien neue Fremd-DNA als neue Spacersequenzen in den eigenen CRISPR-Locus auf. Obwohl die Aufnahme neuer Spacer von Phagen oder Plasmiden in mehreren Organismen bereits experimentell gezeigt werden konnte, ist der eigentliche Mechanismus der Erkennung der Fremd-DNA noch recht unbekannt. Welche Bereiche der Wirtszelle zur Speicherung als neu integrierte Spacer dienen und welche Komponenten an der Immunisierung beteiligt sind, konnte erst kürzlich identifiziert werden. Zu diesen zählen die beiden cas-Gene cas1 und cas2, die in fast allen CRISPR-Typen hoch-konserviert sind und Nukleaseaktivität zeigen (Yosef, Goren et al. 2012). So konnte für Cas1 eine DNA-spezifische Endonukleaseaktivität gezeigt werden (Babu, Beloglazova et al. 2011, Wiedenheft, van Duijn et al. 2011), wohingegen Cas2 eine ssRNA-spezifische Endonukleaseaktivität besitzt. In Adaptationsstudien konnte belegt werden, dass bei Anwesenheit von nur einem der beiden Cas-Proteine keine Fremd-DNA in Form von neuen Spacern in den CRISPR-Array integriert werden kann. Erst bei Expression beider cas-Gene konnte eine Aufnahme von Spacern beobachtet werden, was die Aktivität dieser Gene für die Integration essenziell macht (Yosef, Goren et al. 2012). Des Weiteren werden in einigen Subtypen zusätzlich andere Cas-Proteine (Cas4 und Csn2) in der Immunisierungsphase benötigt (Makarova, Haft et al. 2011).

Die Aufnahme neuer Spacer-Einheiten in den CRISPR-Array erfolgt meist am Leader-nahen Ende, was auch dem Leader-Bereich eine wichtige Rolle in der Adaptation von Fremd-DNA zuteilt (Barrangou, Fremaux et al. 2007). Somit stammen die Leader-nahen Spacersequenzen aus viralen Infektionen jüngerer Zeiten, während die distalen Spacer Auskunft über länger zurückliegende Infektionen geben (Yosef, Goren et al. 2012, Marchfelder, Maier et al. 2013). Eine weitere Beobachtung war, dass der Vorgang der Integration mit einer gleichzeitigen Verdoppelung der ersten Repeatsequenz einhergeht. Ein erster Hinweis wie die Integration von neuen Spacern erfolgt, kam mit der Entdeckung der sogenannten Protospacer in Streptococcus thermophilus auf (Bolotin, Quinquis et al. 2005, Deveau, Barrangou et al. 2008, Horvath, Romero et al. 2008). Die Protospacer umfassen eine Region innerhalb des Phagengenoms, die identisch zu der Sequenz der integrierten Spacer ist. Flankierend zu ihnen wurden kleine, konservierte Regionen gefunden (3-5 bp), die als Protospacer adjacent motifs (PAM) bezeichnet werden und wichtig für die Auswahl künftiger Spacer sind (Mojica and Garrett 2013). Sie dienen nicht nur als CRISPRtypische Motive innerhalb der meisten CRISPR Systeme, sondern nehmen auch Einfluss auf die Orientierung des Einbaus von neu erworbenen Spacern in den Repeat-Array. Darüber hinaus dienen die kurzen PAM-Motiven den beiden universell vertretenden Cas-Proteinen Cas1 und/oder Cas2 als Erkennungssequenz und vermitteln somit die Auswahl der Vorläufer-Spacer (Yosef, Goren et al. 2012). Gleichzeitig sind sie auch wichtiger Bestandteil der Interferenzphase und verhindern eine Autoimmunisierung der eigenen Zelle, indem intakte PAM-Sequenzen als Erkennungssignale nur in der Fremd-DNA vorkommen.

#### 1.2.2 Expression von Komponenten für die Immunabwehr

In der Interferenzphase werden Fremd-DNA Abschnitte erkannt und abgebaut. Um diesen Prozess in der Immunabwehr von Bakterien einleiten zu können, muss zunächst der CRISPR-Locus transkribiert und die crRNAs prozessiert werden. Dabei werden die Cas-Proteine exprimiert und der CRISPR-Array in ein primäres Transkript, die precursor-crRNA transkribiert.

Schon bevor bekannt wurde, welche Möglichkeiten das CRISPR/Cas-Abwehrsystem der Zelle gegen das Eindringen von Fremd-DNA bringt, konnte in Archaeen eine Transkription des CRISPR-Locus und Prozessierung der pre-crRNA beobachtet werden (Tang, Bachellerie et al. 2002, Tang, Polacek et al. 2005). Wo hingegen in einigen Organismen eine ständige Transkription der CRISPR-Loci stattfindet, wird diese in anderen erst durch eine akute Infektion ausgelöst (Marchfelder, Maier et al. 2013). Ein Paradebeispiel dafür stellt der E. coli K12 Stamm dar, bei dem unter normalen Wachstumsbedingungen das CRISPR-Abwehrsystem stark reprimiert ist und erst nach Auflösung dieser Blockade aktiv wird (Pul, Wurm et al. 2010). Grund dafür ist eine Stilllegung der zwei für die Transkription und Expression benötigten Promotoren Pcrispr1 und Pcas, die innerhalb des Leader-Bereiches und der upstream Region der cas Gene lokalisiert sind. Es ist bekannt, dass beide Promotoren auf sehr niedrigem Level konstitutiv aktiv sind (Pougach, Semenova et al. 2010) und durch den globalen Transkriptionsfaktor H-NS reprimiert werden (Pul, Wurm et al. 2010). Eine Aufhebung dieser Repression erfolgt durch den Transkriptionsfaktor LeuO, der als H-NS Antagonist als Aktivator im CRISPR-System die Transkription der cascade Gene induziert (Westra, Pul et al. 2010). Des Weiteren konnte gezeigt werden, dass der Pcas Promotor durch eine konstitutive Expression des Regulators BglJ induziert werden kann. Diese Aktivierung ist LeuO-abhängig und wird durch die Heterodimere BglJ-RcsB aktiviert (Stratmann, Pul et al. 2012).

Die Transkription der langen Vorläufer-RNA (pre-crRNA) erfolgt durch Ablesen des gesamten CRISPR-Locus. Im Anschluss daran kommt es zu einer endonukleolytischen Prozessierung der pre-crRNA in kleinere RNAs (crRNAs) mit einer Länge von 40-60 Nukleotiden. Die Bildung dieser crRNAs geschieht abhängig vom jeweiligen CRISPR/Cas-Typ nach unterschiedlichen Mustern (Abbildung 1.3). Es wird angenommen, dass die dafür verantwortlichen Cas-Proteine dabei konservierte Bereiche (Stem-Loop-Strukturen), die sich aufgrund der palindromischen Struktur der Repeats beiderseits der Spacer ausbilden, als Erkennungsstellen nutzen (Mojica and Garrett 2013). In Typ I CRISPR/Cas-Systemen wie *E. coli* wird die transkribierte pre-crRNA spezifisch durch die Endonuklease Cas6e (CasE), die Bestandteil eines Multiproteinkomplexes ist, in reife, 61 nt lange crRNAs hydrolysiert (Brouns, Jore et al. 2008, Carte, Wang et al. 2008). Nach der Prozessierung beinhalten die reifen

#### **1** Einleitung

crRNAs den antiviralen Spacer, flankiert von kurzen Repeat-RNA-Sequenzen. Dem als Cascade bezeichneten Multiproteinkomplex, bestehend aus den fünf Proteinen Cse1 (CasA), Cse2 (CasB), Cas7 (CasC), Cas5 (CasD) und Cas6e (CasE) konnte neben der Funktion der Prozessierung auch eine Beteiligung an der Erkennung von Fremd-Nukleinsäuren zugeteilt werden (Brouns, Jore et al. 2008). Bei Typ II Systemen wie *Streptococcus pyogenes* wird die Prozessierung der pre-crRNA durch das Multidomain Protein Cas9 und zwei weiteren Komponenten, einer RNaseIII und der tracrRNA (*trans*-encoded crRNA) katalysiert. Durch Komplementarität zu den Repeats bildet die tracrRNA einen kurzen RNA-Doppelstrang mit der pre-crRNA aus, welcher als Prozessierungssignal von Cas9 und RNaseIII erkannt wird (Deltcheva, Chylinski et al. 2011). Typ III CRISPR/Cas-Organismen nutzen Cmr/Csm-Komplexe für die abschließende Prozessierung der reifen crRNAs. Die erste Reifung erfolgt mit Hilfe des Cas6-Proteins (Haurwitz, Jinek et al. 2010, Hatoum-Aslan, Maniv et al. 2011).



#### Abbildung 1.3: Expression und Prozessierung von CRISPR RNAs.

Die Einteilung der verschiedenen CRISPR/Cas-Systeme erfolgt in drei Hauptklassen, die sich in der Bildung der crRNA und Degradierung von Fremd-DNA unterscheiden (Makarova, Haft et al. 2011). Nach Transkription des CRISPR-Locus und Translation der Cas-Proteine, erfolgt die Prozessierung der pre-crRNA in Typ I-Systemen durch die Cas6e-Untereinheit des Cascade-Komplexes. In Typ II-Systemen hingegen, sind für die Reifung die tracrRNA, das Cas9-Protein und die RNaseIII notwendig. In Typ III-Systemen erfolgt die initiale Prozessierung durch das Cas6-Protein. Die weitere Reaktion wird anschließend durch den Csm- bzw. Cmr-Komplex katalysiert (entnommen aus (Wagner and Pul 2012)).

### 1.2.3 Erkennung und Degradation von Fremd-Nukleinsäuren

Die Entdeckung des RNA-vermittelten CRISPR-Schutzmechanismus in Archaeen und Bakterien weist durch analoge Funktionen, wie RNA-Bindung oder benötigte Nuklease- und Helikaseaktivitäten starke Parallelen zum eukaryotischen RNAi-System auf (Lillestol, Redder et al. 2006, Sorek, Kunin et al. 2008). Dieser in Eukaryoten vorkommende Abwehrmechanismus erfolgt über kurze RNAs (siRNAs), die aus Viren stammen und in Kombination mit anderen Zellproteinen gegen Fremd-RNAs zielen. Die Hemmung der viralen Genen wird durch Abbau oder andere Mechanismen bewirkt (Jaskiewicz and Filipowicz 2008).

In Bakterien erfolgt die Erkennung von Fremd-DNA spezifisch durch eine crRNA-vermittelte Interferenz. Dabei binden crRNA-beladene Effektorkomplexe an die zu degradierende Fremd-Nukleinsäure und suchen diese nach passenden Protospacer-Sequenzen ab. Der nicht komplementäre Strang der Fremd-DNA wird verdrängt und es entsteht eine sogenannte R-Schleife (RNA-induzierte DNA-Schleife) (Abbildung 1.4) (Jore, Lundgren et al. 2011, Marchfelder, Maier et al. 2013). Die Degradierung der Ziel-DNA erfolgt anschließend durch Hydrolyse verschiedener Cas-Proteine nahe der PAM-Sequenz.



**Abbildung 1.4: Modell einer ausgebildeten R-Loop Struktur bei Bindung an dsDNA.** Die Basenpaarung der crRNA mit der Protospacer-Sequenz der Fremd-DNA führt zur Verdrängung des nicht-komplementären Stranges unter Ausbildung eines R-Loops (RNA-induzierte DNA-Schleife) (entnommen aus (Westra, Semenova et al. 2013)).

die viraler In Ε. coli beruht Identifizierung von fremder, DNA auf Basenkomplementarität der im Cascade-Komplex gebundenen crRNA mit der **Protospacer-Region** der doppelsträngigen Fremd-DNA. Dabei bilden die prozessierten 61 nt kurzen crRNAs einen 405 kDa großen Ribonukleoproteinkomplex (crRNP), den sogenannten Cascade-Komplex, bestehend aus den fünf Cas-Proteinen CasA, B, C, D und E aus (Jore, Lundgren et al. 2011). Kommt es zur Erkennung von fremden Nukleinsäureabschnitten, wird die Endonuklease Cas3 rekrutiert, welche die Fremd-DNA durch ihre Helikase-/Nukleaseaktivität endonukleolytisch spaltet und exonukleolytisch degradiert (Westra, Nilges et al. 2012). Damit die Zelle zwischen Fremd- und Eigen-DNA unterscheiden kann und es nicht zur Auto-Immunität der

eigenen Zelle kommt, werden die PAM-Motive (Protospacer Adjacent Motifs) zur Hilfe genommen. Diese 2-3 bp kurzen und konservierten Sequenz-Bereiche sind essentiell für die CRISPR-vermittelte Abwehr, können je nach CRISPR/Cas-Typ in ihrer Sequenz variieren und sind in unmittelbarer Nähe der Protospacer lokalisiert (Bolotin, Quinquis et al. 2005, Mojica, Diez-Villasenor et al. 2009, Shah, Hansen et al. 2009). In Experimenten konnte gezeigt werden, dass mismatches innerhalb dieser kurzen Sequenzen zum Erliegen der Immunabwehr führten (Deveau, Barrangou et al. 2008). Um die Beteiligung der PAM-Motive an der Abwehr gegenüber Fremd-DNA näher zu untersuchen und aufzudecken, welche Positionen in den Protospacern für die Interferenz verantwortlich sind, wurden Analysen von Nukleotidpositionen in den Protospacern und PAM-Motiven durchgeführt (Jore 2010). Dabei konnte wie auch schon zuvor in Staphylococcus epidermidis und E. coli gezeigt werden, dass die CRISPR-vermittelte Abwehr die Transformation von Plasmiden inhibieren kann (Marraffini and Sontheimer 2008). Dies bedeutet, dass die CRISPR-vermittelte Abwehr nur erfolgen kann, wenn die Spacersequenz der crRNA mit der Protospacer-Sequenz der Ziel-DNA (Fremd-DNA) übereinstimmt und stellt somit die Basis für eine Sequenz-spezifische Abwehr dar.

### 1.3 Fremd-DNA Erkennung durch Cas-Protein Komplexe

Das prokaryotische CRISPR/Cas-Immunsystem nutzt kleine RNAs (crRNAs) um eindringende Nukleinsäuren von Viren oder konjugativen Plasmiden auszuschalten. Für diese Art von Interferenzmechanismus, in der Fremd-DNA spezifisch durch Effektor-Komplexe (Cas-Protein Komplexe) degradiert wird, werden zusätzliche Proteine benötigt. Je nach CRISPR System-Typ unterscheiden sich diese nicht nur in ihrer Struktur, sondern auch in der Anzahl an individuell beteiligten Cas-Proteinen (Tabelle 1.1). Neben den typisch vorkommenden Cas-Proteinen, wie sie in den meisten CRISPR/Cas-Systemen zu finden sind, können auch sogenannte nicht Cas-Proteine wie die RNaseIII von *Streptococcus pyogenes*, die für die Prozessierung der crRNAs essentiell ist, an der funktionellen CRISPR-Abwehr beteiligt sein (Deltcheva, Chylinski et al. 2011).

In Typ I Systemen erfolgt die Prozessierung durch einen als Cascade (<u>C</u>RISPR-<u>associated</u> <u>c</u>omplex for <u>a</u>ntiviral <u>de</u>fense) bezeichneten Protein-Komplex. Die anschließende Degradierung der Fremd-DNA wird durch das Cas3 Protein vermittelt. In Kombination mit dem Cas3-Protein erkennt und inaktiviert der mit crRNA beladene Cascade-Komplex Fremd-Nukleinsäuren (Brouns, Jore et al. 2008). In Typ II Systemen hingegen fehlen diese Komponenten. Hier wird die Prozessierung der pre-crRNA durch das Cas9-Protein, der RNaseIII und der tracrRNA (<u>trans</u>-encoded RNA) katalysiert (Jinek, Chylinski et al. 2012). Während der Interferenzphase wird die Ziel-DNA von einem Cas9-crRNA:tracrRNA bestehenden Komplex innerhalb der Protospacer-Region durch Basenkomplementariät geschnitten. Für Typ III Systeme sind zwei verschiedene Subtypen (Typ III-A und Typ III-B) und Komplexarten (Csmbzw. Crm-Komplexe) bekannt. Der Hauptunterschied dieser beiden Subtypen liegt in der Natur ihrer Ziel-Nukleinsäure. Während Typ III-A Prokaryoten RNA als Ziel-DNA degradieren, ist es in Typ III-B Systemen DNA (Marraffini and Sontheimer 2008, Hale, Zhao et al. 2009).

**Tabelle 1.1: Klassifizierung beteiligter Effektor-Komplex Proteine der CRISPR/Cas-Typen.** Gezeigt ist eine Tabelle der klassifizierten CRISPR/Cas-Typen und den dazugehörigen, spezifischen Effektor Cas-Proteinen (unterstrichen sind die am Komplex beteiligten Proteine). Die Referenz-Organismen der einzelnen Subtypen sind angegeben.

		Effektor Cas-Proteine	Organismus
Туре І	I-A	<u>Csa5</u> , <u>Cas8a1</u> , <u>Cas7</u> , <u>Cas5</u> , <u>Cas8a2</u> , <u>Cas6</u> , Cas3´- Cas3´´	Aeropyrum pernix
	I-B	<u>Csa6</u> , <u>Cas8b</u> , <u>Cas7</u> , <u>Cas5</u> , Cas3´-Cas3´´	Thermotoga neapolitana
	I-C	<u>Cas5</u> , <u>Cas8c</u> , <u>Cas7</u> , Cas3	Desulfovibrio vulgaris
	I-D	<u>Cas10d</u> , <u>Csc2</u> , <u>Csc1</u> , <u>Cas6</u> , Cas3	Synechocystis sp
	I-E	Cse1, Cse2, Cas7, Cas5, Cas6e, Cas3	Escherichia coli
	I-F	<u>Csy1</u> , <u>Csy2</u> , <u>Csy3</u> , <u>Cas6f</u> , Cas3	Yersinia pestis
Type II	II-A	Cas9	Streptococcus pyogenes
	II-B	Cas9	Wolinella succinogenes
Type III	III-A	<u>Cas6</u> , <u>Cas10</u> , <u>Csm2</u> , <u>Csm3</u> , <u>Csm4</u> , <u>Csm5</u> , Csm6	Mycobacterium tuberculosis
	III-B	Cmr1, Cas10, Cmr3, Cmr4, Cmr5, Cmr6, Cas6	Pyrococcus furiosus

### 1.3.1 Der E. coli Cascade-Komplex

Nach Transkription des CRISPR-Arrays katalysiert eine Metall-unabhängige Endoribonuklease (Cas6e) die Spaltung der langen Vorläufer-RNA in reife, 61 nt kurze crRNAs, bestehend aus einer Spacersequenz, die am 5'-Ende mit 8 nt und am 3'-Ende mit 16-17 nt Überhängen der Repeatsequenzen flankiert sind (Westra, Swarts et al. 2012). In *E. coli* K12 bilden die mit 5'-hydroxyl und 2',3'-zyklischen Phosphat-Enden versehenden reifen crRNAs mit den fünf Cas-Proteinen Cse1, Cse2, Cas7, Cas5 und Cas6e den für die Immunabwehr wesentlichen 405 kDa großen Cascade Multiproteinkomplex (Abbildung 1.5). Erst durch Assoziierung der crRNA erlangt er seine volle Funktionsfähigkeit (Jore, Lundgren et al. 2011). Drei dieser Proteine (Cas5, Cas6 und Cas7) gehören zu der sogenannten RAMP-Proteinklasse (*repeat-associated mysterious protein*) und besitzen ein RNA-Erkennungsmotiv. Die Struktur des *E. coli* Cascade-Komplexes konnte bereits erfolgreich mit Hilfe von Elektronenmikroskopie bestimmt werden (Wiedenheft, Lander et al. 2011). Dabei wurde eine ungleichmäßige Anordnung der einzelnen Protein-Untereinheiten mit einer Stöchiometrie von Cse1<sub>1</sub>Cse2<sub>2</sub>Cas7<sub>6</sub>Cas5<sub>1</sub>Cas6e<sub>1</sub> aufgedeckt (Jore, Lundgren et

#### **1** Einleitung

al. 2011). Der Kern des Komplexes wird aus sechs Cas7 Untereinheiten gebildet, die als Rückgrat für die anderen Cas-Proteinkomponenten und die assoziierte crRNA dienen. Das Cas5 Protein mit nur einer Untereinheit bildet zusammen mit den sechs Cas7 Untereinheiten und einer weiteren Untereinheit der Endoribonuklease Cas6e, einen kleinen Cascade-Komplex, der eine reife crRNA beherbergt. Die strukturelle Anordnung der kleinen crRNA innerhalb des Komplexes bietet auf diese Weise Schutz vor RNA-Degradation. Durch Anlagerung eines Cse2 Dimers und einer weiteren Untereinheit von Cse1 wird der Ribonukleoproteinkomplex vervollständigt und bildet eine Seepferd-ähnliche Struktur aus (Jore 2010). So zusammen gelagert, fungiert er als eine Art Überwachungseinheit, die Fremd-DNA erkennen kann. Für eine CRISPR/Cas-spezifische Abwehrreaktion wird zusätzlich die HD-Nuklease Domäne des Cas3 Proteins benötigt.



#### Abbildung 1.5: Modell des Cascade-Komplexes aus E. coli K12.

Zu sehen sind die einzelnen Cascade-Untereinheiten Cse1, Cse2, Cas7, Cas5 und Cas6e, die den für die Immunabwehr in Typ I-E Systemen charakteristischen Ribonukleoproteinkomplex bilden. Die Cas7 Untereinheit formt das Rückgrat für die übrigen Cascade-Komponenten. In schwarz gezeigt ist die mit dem Komplex assoziierte crRNA.

Die Erkennung von doppelsträngiger, invasiver DNA durch den Cascade-Komplex erfolgt sequenzspezifisch durch Bindung der assoziierten crRNA an die komplementäre Sequenz der DNA. Dabei durchläuft der Cascade-Komplex eine Konformationsänderung, wodurch der nicht komplementäre Strang verdrängt und Bildung angeregt wird. Dass alle Proteinkomponenten zur R-Loop des Cascade-Komplexes essentiell für die erworbene Resistenz in E. coli K12 sind, konnte anhand von Phagenexperimenten (Plaque-Assays) beobachtet werden. Auch konnte neben einer Bindung an einzelsträngige DNA durch Elektrophoretische Mobility Shift Assays (EMSAs) gezeigt werden, dass für die Interaktion der Ziel-DNA keine weiteren Ko-Faktoren wie ATP benötigt werden (Jore, Lundgren et al. 2011). Zudem ist das CRISPR/Cas-System von E. coli in der Lage verschiedene Arten viraler DNA wie Phagen, Plasmid-Transformationen oder die Induzierung mit Prophagen zu neutralisieren (Edgar and Qimron 2010, Semenova, Jore et al. 2011).

### 1.3.2 Erkennung invasiver Fremd-DNA

Die Erkenntnis, dass der Cascade-Komplex an komplementäre, doppelsträngige Protospacer-Regionen ohne Kofaktoren wie ATP binden kann, erlaubt es den Zellen kontinuierlich wie eine Art 'Scanner', ohne den Verbrauch von Energie, nach passenden Ziel-Sequenzen zu suchen und diese zu eliminieren. Essentiell für eine CRISPR-vermittelte Interferenz ist das Vorhandensein eines PAM-Motivs an der Fremd-DNA (Abbildung 1.6). Der mit crRNA-beladene Cascade-Komplex sucht die Ziel-DNA systematisch nach passenden PAM-Motiven ab. Dabei kommt es zur Bindung der 'seed'-Region, einer 7 nt langen Basenpaarabfolge am 3'-Ende des komplementären DNA-Stranges, mit der crRNA. Dies führt zu einer Interaktion des Cse1 Proteins, der größten Untereinheit innerhalb des Cascade-Komplexes, mit der DNA und zur Ausbildung eines sogenannten R-Loops (R-Schleife) (Sashital, Wiedenheft et al. 2012). Es wird angenommen, dass die Ausbildung dieser spezifischen RNA-Schleife anderen Cas-Proteinen, wie dem Cas3 Protein in *E. coli*, als Erkennungssignal für weitere Interferenzprozesse dient (Jore, Lundgren et al. 2011).



#### Abbildung 1.6: Schematische Darstellung der Cascade-vermittelten R-Loop Bildung.

(A) Zeigt eine 61/62 nt kurze crRNA nach Prozessierung, bestehend aus flankierenden Repeatsequenzen (rot) und der Spacerregion (grün). (B) Die an den Cascade-Komplex gebundene crRNA bindet durch Basenpaarkomplementariät an die Protospacer-Region doppelsträngiger DNA. Es kommt zur Verdrängung des nicht-komplementären Stranges unter Ausbildung eines R-Loops. In gelb markiert ist das PAM-Motiv, das der Protospacer-Region vorangestellt ist (entnommen aus (Westra, Swarts et al. 2012)).

Die CRISPR Interferenz ist ein zusammenhängender Prozess, der aus mehreren Schritten besteht. Es beginnt mit der Suche des Cascade-Komplexes nach passenden PAM-Motiven entlang doppelsträngiger Fremd-DNA Abschnitten und wird von der Bindung der Ziel-Sequenz durch Basenpaarung der im Komplex assoziierten crRNA mit komplementären Sequenzen gefolgt. Daraufhin erfolgt die Rekrutierung des Cas3-Proteins. Die Ziel DNA wird zuerst spezifisch durch die HD-Nuklease Domäne des Cas3-Proteins geschnitten und dann in Kombination mit der C-terminalen DExD/H-Box Helikase Domäne in 3'- zu 5'-Richtung degradiert (Westra, van Erp et al. 2012).

### **1.4 CRISPR Adaptation**

Bei der CRISPR/Cas-vermittelten Adaptation gegen invasive Nukleinsäuren, wobei fremde DNA-Sequenzen als neu erworbene Spacer in das eigene Genom integriert werden, handelt es sich um einen bisher noch recht unerforschten Mechanismus. Zwar konnte bereits 2007 die Aufnahme von Fremd-DNA in Typ II Systemen beobachtet werden (Barrangou, Fremaux et al. 2007, Horvath, Romero et al. 2008), aber erst in kürzlich veröffentlichten Studien wurden erste Hinweise in Bezug auf den Ablauf dieses speziellen Integrationsprozesses aufgedeckt (Diez-Villasenor, Guzman et al. 2013, Yosef, Shitrit et al. 2013, Li, Wang et al. 2014).

### 1.4.1 Spacer-uptake in *E. coli* (Typ I-E Systeme)

Für Typ I-E Systeme wie *E. coli* wurde die Aufnahme von fremden Nukleinsäuren, als Spacer, bisher von mehreren, unabhängigen Forschergruppen beschrieben (Datsenko, Pougach et al. 2012, Swarts, Mosterd et al. 2012, Yosef, Goren et al. 2012). Dabei konnten *cas1* und *cas2*, zwei der wesentlichen Core-*cas*-Gene, die universell und konserviert in allen CRISPR/Cas-Typen vertreten sind, für die Integration von neuen Spacersequenzen als essentielle Gene identifiziert werden. Im CRISPR-Locus von *E. coli* sind die *cas1/cas2*-Gene *downstream* der *cascade* Gene lokalisiert und werden von einem polycistronischen Transkript translatiert. Unter normalen Wachstumsbedingungen ist die Fähigkeit neue Spacer in den CRISPR-Array zu integrieren aufgrund einer reprimierten Expression der *cas* Gene stark erniedrigt. Wird das Expressionslevel der *cas1* und *cas2* Gene jedoch erhöht, können Fremd-DNA Abschnitte aufgenommen werden. Eine einseitige Überexpression von nur einem der beiden Gene resultiert wiederum in einem Verlust zur Spacer Aufnahme, was zeigt, dass beide Gene essenziell für den Prozess der CRISPR Adaptation sind (Yosef, Goren et al. 2012).

Die Repeat-Spacer Einheit innerhalb eines CRISPR-Locus ist meist einer AT-reichen Leaderregion benachbart, die den Promotor für die Transkription des Arrays beherbergt (Pul, Wurm et al. 2010). Das Einfügen von neuen Spacern in unmittelbarer Nähe der Leaderregion, weist auf eine direkte Beteiligung der Leadersequenz an diesem Prozess hin (Swarts, Mosterd et al. 2012, Yosef, Goren et al. 2012). Insbesondere für *E. coli* konnten vor kurzem wesentliche, an dem Mechanismus der Integration beteiligte Komponenten aufgedeckt werden. Neben einer Rolle als Integrase während der Spacerintegration (Arslan, Hermanns et al., in Revision) konnte dem Cas1 Protein in durchgeführten *in vitro* Analysen eine Metall-abhängige Nukleaseaktivität nachgewiesen werden, die Einzel- oder Doppelstrang-DNA in einer sequenzspezifischen Weise schneidet (Wiedenheft, Zhou et al. 2009). Dem Cas2 Protein kann eine Metall-abhängige, RNA-spezifische Endoribonukleaseaktivität nachgewiesen werden (Beloglazova, Brown et al. 2008). Wie und an welchem Schritt des Integrationsmechanismus die beiden Gene *cas1* und

#### **1** Einleitung

*cas2* beteiligt sind, konnte aktuell anhand von *in vivo* Experimenten, die das Vorhandensein von sogenannten 'Integrations-Zwischenstufen' (Intermediaten) aufdeckten, beschrieben werden (Arslan, Hermanns et al., in Revision). Neben diesen für die Adaptation charakteristischen Proteinen wird für den als "primed adaptation" bezeichneten Prozess zusätzlich die Anwesenheit des Cascade-Komplexes und des Cas3 Proteins benötigt. Bei dieser Art von Adaptation wird die Aufnahme neuer Spacer von Viren oder Phagen begünstigt, die bereits als integrierte Spacer Bestandteil des CRISPR-Arrays sind (Datsenko, Pougach et al. 2012, Swarts, Mosterd et al. 2012). Während *E. coli* auch fremde, unbekannte Nukleinsäuren aufnimmt, ist die Aufnahme in Typ I-B Systemen wie *Haloarcula hispanica* abhängig von Spacern, die zum Teil identische DNA-Sequenzen zu der Fremd-DNA aufweisen und von für die Interferenzphase benötigten Cas Proteinen (Li, Wang et al. 2014).



# Abbildung 1.7: Schematische Darstellung der Aufnahme neuer Spacer während der CRISPR Adaptation.

Nach aktuellem Modell zur Folge wird die Fremd-DNA, bestehend aus einem Protospacer (hell-blau) und einem PAM-Motiv (orange), nach Infizierung der Bakterienzelle von den beiden Nukleasen Cas1 und Cas2 erkannt und hydrolysiert. Der Einbau des neuen Spacers in den eigenen CRISPR-Array erfolgt immer am Leader-proximalen Ende (grün) unter Duplikation der ersten Repeatsequenz (grau).

In Anlehnung an die aktuellen Modelle zur Spacerintegration wird während eines Virusangriffes der Bakterienzelle die virale Fremd-DNA bestehend aus Protospacer und PAM-Motiv von den universell vertretenen Cas Proteinen Cas1 und Cas2 erkannt und in einer sequenzspezifischen Weise geschnitten. Wichtige Integrationsbereiche, an denen die Cas1-spezifischen Schnitte erfolgen, wurden erst kürzlich anhand von Southern Blot Analysen im chromosomalen CRISPR-Array von *E. coli* BL21 AI aufgedeckt. Dabei konnte für den Integrationsprozess ein gekoppelter Spaltungs-Ligations-Mechanismus nachgewiesen werden (Arslan, Hermanns et al., in Revision). Die Integration in den eigenen CRISPR-Array erfolgt unter Verdoppelung des Leader-proximalen Repeats, so dass der neue Spacer von zwei identischen Repeatsequenzen flankiert ist (Abbildung 1.7) (Goren, Yosef et al. 2012). Die erste Repeatsequenz des CRISPR-Arrays fungiert demnach als Matrize für die Duplikation. Neben den Genprodukten *cas1* und *cas2* sind in anderen CRISPR-Typen zusätzlich die subtypspezifischen Proteine Cas4 und Csn2 an dem Immunisierungsprozess beteiligt (Ellinger, Arslan et al. 2012).

### 1.4.2 Spacer-uptake in Typ II-Systemen

Typ II CRISPR/Cas-Systeme, wie sie hauptsächlich in pathogenen Bakterien zu finden sind, beinhalten einen CRISPR-Array und cas Gene, die für die Proteine Cas9, Cas1, Cas2 und Csn2 (Typ II-A) oder Cas4 (Typ II-B) codieren. Ein weiteres Gen, welches für eine *trans*-encoded crRNA, kurz tracrRNA codiert und komplementäre Sequenzen zum Repeat der pre-crRNA enthält, ist zusätzlich im CRISPR-Locus vorhanden (Deltcheva, Chylinski et al. 2011). Bei Basenpaarung der tracrRNA mit komplementären Sequenzen der pre-crRNA, kommt es zur Prozessierung der kleinen crRNAs durch das Cas9 Protein und einem zusätzlichen 'house-keeping' Faktor, der RNaseIII (Gasiunas, Barrangou et al. 2012). Im Anschluss daran werden durch das Cas9 Protein in Kombination mit dem tracrRNA:crRNA Komplex Doppelstrangbrüche an Ziel-Protospacer Sequenzen induziert (Jinek, Chylinski et al. 2012). Eine Beteiligung der Proteine Cas1, Cas2 und Csn2 an der Prozessierung von crRNAs und der Interferenz von Fremd-DNA ist nicht bekannt und führte zu der Annahme, dass sie eine Rolle in der Immunisierungsphase spielen (Deltcheva, Chylinski et al. 2011). Die Fähigkeit zur Adaptation neuer Spacer in Typ II-Systemen wurde erstmals 2007 in *Streptococcus thermophilus* beschrieben. Bakterienpopulationen, die eine Phageninfektion überlebt hatten, erweiterten ihren CRISPR-Array mit Spacern, die identisch zu kleinen DNA-Regionen, den Protospacern der infizierenden Phagen waren (Barrangou, Fremaux et al. 2007). Anhand von Knock-out Experimenten in anderen Studien konnte indirekt gezeigt werden, dass das csn2 Gen aus Typ II-Systemen essentiell an der Aufnahme neuer Spacer beteiligt ist (Deveau, Barrangou et al. 2008). Dennoch ist die Rolle, die Csn2 während der Adaptation von neuen Spacern spielt noch unbekannt. Die erst vor kurzem aufgelöste Struktur von Streptococcus agalactiae Csn2 zeigt die Ausbildung einer stabilen, ringähnlichen Struktur aus Tetrameren, die bedeutende Konformationsänderungen vornehmen kann und konservierte Lysin-Reste besitzt, die Doppelstrang (ds-) DNA bindet (Nam, Kurinov et al. 2011, Ellinger, Arslan et al. 2012). Kurz darauf konnte ein DNA-bindender Mechanismus des Csn2 Proteins, bevorzugt an freie DNA-Enden zu binden, mit Hilfe verschiedener biochemischer und biophysikalischer Techniken (Atomic Force Mikroskopie (AMF), Molecular Dynamics) aufgedeckt werden. Dies verstärkte die bestehende Annahme, dass Csn2 als weiteres *cas* Genprodukt an der Integration von neuen Spacer in den CRISPR-Array beteiligt ist (Abbildung 1.8) (Arslan, Wurm et al. 2013). An welche DNA-Sequenzen das Csn2 Protein bindet, und ob es sich dabei tatsächlich um CRISPR spezifische Sequenzen handelt, konnte bisher noch nicht aufgeklärt werden. Mitaufgereinigte DNA von *E. faecalis* Csn2 zeigte keine CRISPR spezifischen DNA-Sequenzen, was auf eine unspezifische Bindung von Csn2 an Doppelstrang-DNA schließen lässt (Nam, Kurinov et al. 2011).



# Abbildung 1.8: Hypothetisches Modell zur Funktion von Csn2 DNA-Endbindung während der Spacerintegration.

Der DNA-Doppelstrang wird an verschiedenen Positionen (Leader-Repeat-Grenze und Repeat-Spacer-Grenze) (gestrichelte Pfeile) durch die an der Integration beteiligten Proteine Cas1 und Cas2 gespalten und die entstandenen Lücken durch eine Polymerasereaktion unter gleichzeitiger Duplikation des Leader-proximalen Repeats aufgefüllt. Die dabei resultierenden, freien DNA-Enden (blaue Ringe) werden von Csn2 gebunden. Eine solche Bindung schützt nicht nur die freien DNA-Enden vor einer Degradierung, sondern könnte auch gleichzeitig als Signal für die Rekrutierung anderer DNA-Reparaturproteine dienen (entnommen aus (Arslan, Wurm et al. 2013).

Das Cas4 Protein, ein weniger konserviertes Core-Cas Protein, ist homolog zum Csn2 Protein in Typ II-B Systemen vertreten und zeigt Sequenzähnlichkeit zu Exonukleasen der RecB-Familie (Haft, Selengut et al. 2005, Makarova, Grishin et al. 2006). Die Kristallstruktur von *S. solfataricus* Cas4 zeigt wie bei Csn2, eine

#### **1** Einleitung

ringförmige Anordnung des Proteins. Die in anderen Studien beobachtete 5'-3'-DNA-Exonukleaseaktivität führte zu der Annahme, dass Cas4 für DNA End-Modifikationen während der Aufnahme neuer Spacersequenzen verantwortlich sein könnte (Zhang, Kasciukovic et al. 2012). Eine häufige Assoziierung der Gene von *cas4* mit *cas1* und *cas2* oder sogar die Bildung eines als "Cascis" bezeichneten Multiproteinkomplexes (<u>C</u>RISPR-<u>as</u>sociated <u>c</u>omplex for the <u>integration of spacers</u>), die in *Thermoproteus tenax* beobachtet wurde, verstärkten die Annahme, dass auch Cas4 an der CRISPR-vermittelten Immunisierung beteiligt ist (Plagens, Tjaden et al. 2012). In Typ I-E CRISPR/Cas-Systemen wie *E. coli* fehlt das Cas4 Protein. Es kam die hypothetische Annahme auf, dass durch eine Interaktion von Cas1 mit dem house-keeping RecB Nukleaseprotein das Fehlen von Cas4 kompensiert und so ein Beitrag zur CRISPR-Interferenz geleistet wird. Neben dem Part in der rekombinanten DNA-Reparatur spielt der *E. coli* RecBCD-Komplex in der Zell-eigenen Abwehr gegen virale Angriffe eine wesentliche Rolle, die ähnlich zu der CRISPR-vermittelten Abwehr ist.

Nicht nur die Struktur des Cas4 Proteins, sondern auch die Sequenzähnlichkeit zu Exonukleasen der RecB-Familie, lassen Rückschlüsse auf einen katalytischen Mechanismus durch die Nuklease Cas4 zu. Dabei kann dem Protein neben einer Funktion zur Bildung von 3'-Überhängen für die Generierung neuer Spacer, auch eine Rolle in der Zerstörung von fremd DNA-Nukleinsäuren durch eine Umwindung der DNA mit gleichzeitiger Nukleaseaktivität zugesprochen werden (Lemak, Beloglazova et al. 2013). Allerdings müsse dieser Punkt durch zukünftige Studien verifiziert werden, die neue Einblick in die Funktion und den molekularen Mechanismus des Cas4 Proteins innerhalb des CRISPR/Cas-Systems geben könnten.

### 1.5 Industrielle Anwendung: CRISPR-basierte Technologie

Seit Entdeckung des prokaryotischen CRISPR/Cas-Systems und dem Verständnis der Wirkungsweise des immunbasierten Schutzmechanismus, wird diese Art von Abwehrsystem, dass gegen parasitäre Eindringlinge zielt, heutzutage in mehrfacher Hinsicht genutzt. Neben der Anwendung in Sektoren wie der Forschung und Entwicklung, findet es auch in biotechnologischen Prozessen und zur genetischen Manipulation intensive Verwendung. Ob in der Lebensmittelindustrie bei der Herstellung von Milchprodukten oder wichtiger Produkte wie Biokraftstoffen, Grundchemikalien und Pharmazeutika, Viren stellen eine ständige und anhaltende Bedrohung der oft mit erheblichen Aufwand konstruierten Kulturen dar (Terns and Terns 2014). Dank des CRISPR/Cas-Systems kann diesem, nicht zu unterschätzendem Risikofaktor, durch Erzeugung von resistenten Stämmen ein Ende gesetzt werden. Dazu benötigt es lediglich die Infektion von Bakterien mit verschiedenen Phagen und die anschließende Selektion der daraus resultierenden, resistenten Bakterien.

Aufgrund von Phagen-bedingten Zerstörungen bakterieller Populationen kam schon recht früh die Idee auf, Phagen für die Behandlung von Infektionen mit pathogenen Bakterien zu verwenden (Merril, Scholl et al. 2003). Insbesondere für die Lebensmittelindustrie ist die Herstellung von multiresistenten Stammkulturen von großem Interesse, da es sich bei diesen um nicht gentechnisch veränderte Organismen handelt (Marchfelder, Maier et al. 2013). Die Manipulation von CRISPR-Systemen zur Generierung von Phagen-resistenten Stämmen hat neben ökonomischen Vorteilen auch ein kommerziellen Nutzen, da dadurch Verbesserungen der Produktqualitäten und Verringerungen von wirtschaftlichen Verlusten bei Infektionen von Kulturen erzielt werden können (Barrangou and Horvath 2012).

Die Fähigkeit des CRISPR/Cas-Systems, DNA-Moleküle in einer sequenz-spezifischen Weise leiten zu können, hat vor kurzem zur Entwicklung neuer Technologien mit dem Ziel zur Genmanipulationen von Bakterien und Eukarvoten geführt. Dabei werden die Komponenten des Typ II-Systems zur Hilfe genommen um Mutationen zu erzeugen. Π CRISPR/Cas Effektor-Komplexe beinhalten eine **RNA-spezifische** Tvp Endonuklease, das Cas9-Protein, welches die Spaltung von 'blunt-end' dsDNA-Molekülen katalysiert. Dabei wird neben einer crRNA, die das Cas9-Protein zu den Ziel-DNA Sequenzen leitet, eine zweite RNA, die *trans*-encoded crRNA (tracrRNA) rekrutiert, die wiederum mit der crRNA basenpaart und essentiell am Degradierungsprozess von Fremd-DNA beteiligt ist (Deltcheva, Chylinski et al. 2011, Jinek, Chylinski et al. 2012). Beide RNAs (crRNA und tracrRNA) können experimentell über eine Art Verbindungsschleife zu einer einzelnen Guide-RNA gekoppelt (sgRNA, small guide-RNA) und dann in Kombination mit dem Cas9-Protein genutzt werden, um Mutationen zu erzeugen (Charpentier and Doudna 2013) (Abbildung 1.9).



Abbildung 1.9: Modell einer CRISPR/Cas9-vermittelten Genom-Editierung.

Das Cas9-Protein wird zur Ziel-DNA durch die Guide-RNA geleitet und bildet Basenpaare mit der Ziel-DNA aus. Anschließend erfolgt ein spezifischer Schnitt 3 bp oberhalb der PAM-Sequenz an beiden DNA-Strängen (rotes X) (entnommen aus mirusbio.com).

Durch Modifikation der crRNA-Sequenzen können die Sequenzspezifitäten der Effektorkomplexe des Typ II CRISPR/Cas-Systems relativ kostengünstig und einfach gesteuert werden. Lediglich die Beladung des Proteins mit eigens entworfenen crRNA:tracrRNA Molekülen muss erfolgen (Cong, Ran et al. 2013). So entstehen an gewünschten Positionen der Ziel-DNA Doppelstrangbrüche, die anschließend durch

den relativ fehlerbehafteten NHEJ-Mechanismus (Non-Homologous End Joining) nicht ganz korrekt repariert werden und als Resultat eine Mutation begünstigen. Eine bisherige Alternative zu der neu entwickelten Cas9 Genom-Editierung stellten in der Vergangenheit die Zink-Finger-Nukleasen (ZFNs) oder Transkriptionsaktivatorähnlichen Effektornukleasen, kurz TALEN's (Transcription Activator-Like Effector Nucleases), dar (Horvath and Barrangou 2013). Diese können spezifisch an veränderte DNA-Sequenzen binden und spalten, sind aber im Vergleich zum angewandten Typ II System durch Neuanpassung der Nukleasen für jede neue Zielzelle im Genom zeit- und kostenintensiv. Mit der neuen Cas9-Editierungs-Technologie konnten bereits in kürzester Zeit erfolgreich gezielte Genommanipulationen in humanen Zellen und Modellorganismen wie Zebrafisch, Drosophila und anderen Eukaryoten durchgeführt werden (Chang, Sun et al. 2013, Gasiunas and Siksnys 2013, Jinek, East et al. 2013). Zudem kann durch Inhibierung der Nuklease-Aktivität des Cas9 Proteins die Transkription reguliert werden. Der Effektorkomplex bindet an die Ziel-DNA, kann sie aber nicht zerstören und fungiert somit als Transkriptionsregulator, was eine Repression der Zielgene zur Folge hat (Larson, Gilbert et al. 2013). Dieser, als CRISPRi (CRISPR *interference*) bezeichnete Mechanismus ermöglicht es, in Prokaryoten die Transkription gezielt und reguliert zu kontrollieren.

### 1.6 Fragestellung und Konzeption der Arbeit

Die vorliegende Arbeit befasst sich mit dem prokaryotischen CRISPR/Cas-Abwehrmechanismus aus *E. coli* und ist in zwei Teile gegliedert.

Im ersten Teil soll ein Plasmid-System zur Spacerintegration in E. coli aufgebaut und die Fähigkeit neue Spacer aus Fremd-DNA aufnehmen zu können, in vivo durch Adaptationsstudien untersucht werden. Dabei sollen vor allem neue Hinweise über den Adaptationsmechanismus in Bakterien gesammelt werden und bisher noch nicht untersuchte Bereiche innerhalb der ersten Repeatsequenz analysiert werden. Der E. coli BL21 AI Stamm besitzt einen CRISPR II Locus im Genom der aus einem CRISPR-Array mit 13 Repeat-Sequenzen, 12 variablen Spacern und einer AT-reichen Leader-Region besteht. Durch den Verlust des ersten CRISPR-Locus, der auch ein Fehlen der cas-Gene zur Folge hat, ist die Integration von neuen Spacern in den CRISPR-Array unterbunden. Des Weiteren kann auch die CRISPR-vermittelte Abwehr nicht erfolgen, da die Prozessierung der pre-crRNA in die funktionellen reifen crRNAs nicht stattfindet. Diese reifen crRNAs, bestehend aus einer Spacersequenz mit flankierenden Repeats, bilden die Grundlage für die Abwehr von Fremd-DNA. Daraus ergibt sich die Notwendigkeit eines plasmidären Systems, dass die für die Immunisierungsreaktion essentiellen cas-Gene enthält. Hierzu soll zunächst ein Plasmid konstruiert werden das die Grundlage für weitere Experimente bietet. Bei diesem pCR003 Plasmid handelt es sich um ein pACYC-Duet1 Derivat mit plasmid-codiertem cas1-cas2-Operon, in das zusätzlich ein künstlich synthetisierter CRISPR-Array kloniert wurde. Dieser CRISPR-Array besteht aus einer Leaderregion und vier identischen Repeat-Einheiten, die zwischen dem ersten und zweiten Repeat durch eine Spacersequenz unterbrochen sind. Nach erfolgreicher Etablierung des Plasmid-Systems soll in vivo ein Einbau von Spacern in den plasmidären CRISPR-Array untersucht werden und die Adaptation anhand von Langzeitstudien analysiert werden. Im Anschluss daran soll der ersten Repeat durch gezielte Punktmutationen im pCR003 Plasmid verändert werden, um wichtige Sequenzen innerhalb der ersten Repeatsequenz zu identifizieren. Bereits bekannt ist, dass die Integration neuer Spacer meist am Leader-nahen Ende des CRISPR-Arrays erfolgt. Anhand von Southern Blot Analysen und einer Sequenzierungsreaktion soll in vivo der Mechanismus, wie und an welchen Bereichen des CRISPR-Arrays die Integration erfolgt, näher charakterisiert werden. Im weiteren Verlauf soll das zuvor in Typ I-E etablierte Plasmid-basierte Integrations-Assay mit Protein-Komponenten anderer CRISPR Typen aus Typ II-A Systemen komplementiert werden. Die Fähigkeit zur Spaceraufnahme soll anhand von Integrationsstudien analysiert werden.

In einem weiteren Teil der Arbeit ist vorgesehen, die fünf Cas-Proteine CasA (Cse1), CasB (Cse2), CasC (Cas7), CasD (Cas5) und CasE (Cas6e) aus *E. coli* zu isolieren und durch *in vitro* Bindestudien zu analysieren. Diese fünf Proteine bilden den für Typ I-E CRISPR/Cas-Systemen charakteristischen Cascade-Komplex, der mit einer kurzen crRNA assoziiert, Fremd-DNA erkennt und mit Hilfe des Cas3 Proteins degradiert. Zunächst werden dazu die fünf Proteine in einen geeigneten Expressionsvektor der

#### Einleitung

mit einem Affinitäts-Tag (StrepII-Tag) ausgestattet ist kloniert, um sie anschließend mittels Affinitätschromatografie aufreinigen zu können. Die so isolierten Proteine sollen anschließend durch *in vitro* Bindestudien mit einzel-und doppelsträngigen DNA-und RNA-Fragmenten charakterisiert werden.

### 2 Ergebnisse

Bei der Adaptation gegen Fremd-Nukleinsäuren durch den CRISPR/Cas-vermittelten Mechanismus wird fremde, invasive DNA erkannt und als neuer Spacer in den eigenen CRISPR-Array aufgenommen. Bereits bekannt ist, dass dabei die zwei Cas-Nukleasen *cas1* und *cas2* als essentielle Gene Bestandteil sind. Neben diesen beiden universell vertretenden Genen werden zusätzlich eine konservierte Leaderregion und eine palindromische Repeatsequenz für die Immunisierung gebraucht, da die Integration neuer Spacer in der Regel am Leader-nahen Ende des CRISPR-Arrays unter gleichzeitiger Verdoppelung des ersten Repeats erfolgt.

Im ersten Schritt dieser Arbeit stand die Entwicklung eines Plasmid-basierten Integrations-Assays in *E. coli* im Vordergrund. Nach erfolgreicher Etablierung des Assays sollte durch Einführung von Punktmutationen der Einfluss der ersten Repeatsequenz auf die Spacerintegration untersucht werden. Anhand von *in vivo* Studien (Southern Blot Analysen, Sequenzierungsreaktion) war es angedacht, den ablaufenden Mechanismus der Integration näher zu analysieren und wichtige Bereiche, die an der Integration beteiligt sind aufzudecken. In einem weiteren Punkt war geplant ein heterologes Integrationssystem in *E. coli* mit Adaptations-Komponenten des pathogenen Bakterienstammes *Streptococcus agalactiae* zu entwickeln.

### 2.1 Aufbau eines CRISPR-basierten Plasmid-Systems in *E. coli*

In Anlehnung an ein von Qimron und Mitarbeiter (Yosef, Goren et al. 2012) entwickeltes System zur Spacerintegration, konnte in unserem Labor bereits *in vivo* die Aufnahme von neuen Spacern in das Genom von *E. coli* BL21 AI durch Überexpression der für die Integration benötigten *cas*-Gene *cas1* und *cas2* beobachtet werden. Um neben der Beteiligung von *cas1* und *cas2* an der Erkennung der Leaderregion und Spaltung der ersten Repeatsequenz auch die Sequenzen innerhalb des ersten Repeats modifizieren und neu hinzugefügte Spacersequenzen analysieren zu können, war es daher empfehlenswert ein Plasmid-basiertes System zu entwickeln. Dafür musste zunächst ein synthetischer CRISPR-Array in einen geeigneten Expressionsvektor, der die für die Integration essentiellen *cas1* und *cas2* Gene enthält kloniert werden.

### 2.1.1 Klonierung des pCR003 WT-Plasmids

Für die Klonierung des pCR003 Plasmids wurde als Grundbaustein das Plasmid pCR001 verwendet. Dieses pACYC-Duet1 Derivat besitzt die durch PCR amplifizierten *cas1* und *cas2* Gene aus genomischer DNA von MG1655 in die zweite *'multiple cloning site'* (*MCS 2*) hinter den T7-Promotor des Vektors kloniert (Arslan, Wurm et al. 2013).

Um ein Plasmid-basiertes System in E. coli aufbauen zu können, musste in einem ersten Schritt das Ausgangsplasmid pCR001 durch Restriktionshydrolyse mit dem Enzym *Ecl136II* (4.5.1) in der ersten *MCS* 'blunt' linearisiert werden. Darauf folgte die Klonierung des synthetischen CRISPR-Arrays in den geöffneten Vektor, der zuvor durch eine PCR-Reaktion aus dem Plasmid pLS1(+) (4.2.2) mit Hilfe der Oligonukleotide pMA-NcoI-fw und pMA-XhoI-rv (4.4.1) gewonnen wurde. Der synthetisch hergestellte CRISPR-Array besteht aus einer Leaderregion und vier identischen Repeatsequenzen (29 bp), die zwischen dem ersten und zweiten Repeat von einer 32 bp großen, variablen Spacersequenz unterbrochen sind (siehe Anhang Abb. 6.1). Um die Richtigkeit der Klonierung sicherstellen zu können wurde das pCR003 Plasmid zunächst in kompetente XL-1 Zellen transformiert, die Plasmid-DNA durch Minipräparation (5.2.2.1) isoliert und anschließend zur Sequenzierung geschickt (5.1.5). Die Sequenzierungsergebnisse zeigten eine korrekte Klonierung des CRISPR-Arrays mit einer divergenten Orientierung zum cat Gen. Nach Transformation des Plasmides in kompetente E. coli BL21 AI/pUC18-Kan Zellen konnten die Zellen für Integrationsstudien genutzt werden (Abbildung 2.1). Das zuvor in die *E. coli* BL21 AI Zellen transformierte pUC18-Kan Plasmid dient hierbei als Fremd-DNA Träger.



#### Abbildung 2.1: Klonierungsschema des Plasmid-basierten CRISPR-Systems in E. coli.

Gezeigt ist das Schema zur Konstruktion des pCR003 WT Plasmids, der nach Klonierung des synthetischen *E. coli* CRISPR-Arrays (rot) in den linearisierten Ausgangsvektor und Transformation in kompetente *E. coli* BL21 AI Zellen als Basis für das Plasmid-basierte Integrationssystem dient.

Ob eine Aufnahme von Fremd-DNA Sequenzen als neue Spacer, sowohl in den einklonierten CRISPR-Array des Plasmids als auch in das Genom von BL21 AI erfolgt, wurde durch Induktion der Plasmid-codierten *cas1* und *cas2* Gene mit IPTG und L-Arabinose und Wachstum der Zellen für 18-24 Stunden durch eine PCR-Reaktion kontrolliert (5.2.6.6).

# 2.1.2 Detektion der Spaceraufnahme mittels *single-colony* PCR (sc-PCR)

Ob das entwickelte Plasmid-basierte System zur Spacerintegration in *E. coli* aktiv ist, wurde anhand von Adaptationsstudien über mehrere Tage kontrolliert. Dabei wurden nach Induktion der *cas1* und *cas2* Gene die Zellen über Nacht (üN) bei 37°C inkubiert und der Einbau von neuen, noch unbekannten Spacersequenzen sowohl im Genom als auch im Plasmid durch PCR-Amplifikation der entsprechenden CRISPR-Bereiche kontrolliert.

Nach Transformation des pCR003 Plasmides in BL21 AI/pUC18-Kan Zellen wurde eine Kolonie der angewachsenen Agarplatte als üN-Kultur angeimpft (5.1.2.2) und am nächsten Morgen 1/100 in frisches YT-Medium überführt. Zum starten eines ersten Zyklustages wurde das Medium zuvor mit IPTG und L-Arabinose versetzt, um die Induktion der cas-Gene nach Animpfen mit der üN-Kultur zu starten. Ein neuer, weiterer Zyklus wurde nach 18-24-stündiger Inkubationszeit der Zellen durch das Versetzen von frischen YT-Medien mit 1:100 der Kulturlösung des vorhergehenden Zyklus gestartet. Verschiedene Verdünnungen (1:50 und 1:500) der Kultur des ersten Zyklus dienten als template für eine PCR-Reaktion. Neben dem Einbau in den CRISPR-Array des pCR003 Plasmids wurde auch der Einbau von neuen Spacern in die genomische CRISPR-Region der BL21 AI Zellen als Positivkontrolle untersucht. Entsprechend der unterschiedlichen Spacer im Genom von BL21 AI und des pCR003 Plasmids mussten für die Amplifizierung der CRISPR-Bereiche verschiedene Oligonukleotide in die PCR-Reaktion eingesetzt werden (4.4.1). Dabei lag der zu amplifizierende Bereich dieser Oligonukleotide innerhalb der Leaderregion und der zweiten Spacer-Einheit. So war es möglich anhand der DNA-Fragmentenlängen, Bereiche mit neu hinzugefügten Spacern von denen ohne Integration zu unterschieden. Erfolgte kein Einbau in den CRISPR-Array, ist nur eine Bande auf dem Agarosegel zu sehen. Sind zwei oder sogar drei übereinanderlaufende Banden sichtbar, ist dies ein Indiz dafür, dass neue Spacersequenzen eingebaut wurden. Nach Einbau eines neuen Spacers wird der zu amplifizierende Bereich um 61 bp verlängert und kann mit Hilfe der gewählten Vernetzung (Porengröße) des Agarosegels detektiert werden. Das Ergebnis der PCR-Amplifikationsreaktion wurde auf einem 1,2% igen Agarosegel überprüft und ist in Abbildung 2.2 zu sehen.



#### Abbildung 2.2: PCR-Analyse der Spacerintegration beim pCR003 WT Plasmid.

Gezeigt sind die PCR-Produkte von drei Zyklen des pCR003 WT Plasmids auf einem 1,2%igen Agarosegel getrennt. Nach 18-24 stündigem Wachstum der Zellen wurden Verdünnungen der Kulturen (1:50 und 1:500 in YT-Medium) als *template* in eine PCR-Reaktion eingesetzt. Der Einbau von neuen Spacersequenzen wurde mit verschiedenen Oligonukleotiden (4.4.1) sowohl im genomischen Array (links), als auch im synthetischen Array des WT Plasmids (rechts) kontrolliert. Es sind zwei unabhängige Experimente geteigt, bei denen unterschiedliche Kolonien für die Analyse gepickt wurden. Parentale PCR-Fragmente sind mit "parental" gekennzeichnet. Darüber liegende Banden sind als neu erworbene Spacersequenzen mit "extended" gekennzeichnet (1x bzw. 2x extended).

Die Abbildung 2.2 zeigt die Ergebnisse der PCR-Reaktion von je drei Zykluskontrollen nach Induktion und Inkubation für 18-24 Std. Es handelt es sich um eine Doppelbestimmung, bei der unterschiedliche Kolonien einer frisch transformierten Agarplatte gepickt wurden. Diese Art der Bestimmung wurde mehrfach über einige Wochen hinweg mit immer wieder frisch transformierten Zellen wiederholt, um die Reproduzierbarkeit der Ergebnisse verifizieren zu können.

Eine Aufnahme von neuen Spacern sowohl ins Genom als auch ins Plasmid ist in beiden Fällen deutlich zu erkennen. Betrachtet man zunächst den oberen Teil der Abbildung, ist eine Zunahme des genomischen Einbaus in den CRISPR-Array von BL21 AI von Zyklus 1 zu Zyklus 2 (Spuren 1 und 2) zu beobachten. Doch auch der plasmidäre Einbau lässt eine Zunahme erkennen. Eine Erweiterung des Plasmid-basierten CRISPR-Arrays um 61 nt durch Hinzufügen einer neuen Spacer-Repeat Einheit kann nach Amplifizierung des CRISPR-Bereichs mit ausgewählten Primern gezeigt werden. Im Vergleich zu Zyklus 1 und 2 ist im dritten Zyklus sogar eine zweite Bande oberhalb der Parentalen (Spur 3) zu sehen, was auf eine Aufnahme eines weiteren, zweiten Spacers hindeutet. Bei Betrachtung der unteren Zykluskontrolle überwiegt der genomische Einbau gegenüber dem plasmidären. Hier ist eine deutliche Zunahme bei der Aufnahme von neuen Spacern ins Genom von Zyklus 1 bis hin zu Zyklus 3 zu beobachten (Spuren 1 und 2). Doch auch der Einbau in den Array des pCR003 Plasmids lässt eine stetige Zunahme von erkennen (Spuren 3 und 4). Diese Art von Ergebnissen, bei denen das Hinzufügen von neuen Spacern sowohl in den CRISPR-Array des BL21 AI Genoms, als auch in den CRISPR-Array des konstruierten pCR003 Plasmids erfolgt, konnte auch in anderen reproduzierten Experimenten bestätigt werden. Dabei fiel auf, dass der genomische Einbau gegenüber dem plasmidären überwiegte. Während den Zellen die Aufnahme ins Genom mit teilweise bis zu vier Spacern gelingt, schafft es das Plasmid nur auf maximal drei Spacer (Ergebnisse nicht gezeigt). Auch ist eine Abnahme der Fähigkeit neue Spacer aufnehmen zu können, nach Zyklus vier zu sehen. In mehreren, unabhängigen Integrationsstudien konnte beobachtet werden, dass die Zellen nach Zyklustag 4 in ihrem Einbau reduziert waren. Bei einigen wenigen war sogar nach Tag 5 nur noch die parentale Bande sichtbar. Ein möglicher Grund für diese Reduzierung könnte eine verringerte bzw. stagnierte Zellzahl durch ein limitiertes Nährstoffangebot im Medium sein und wird nachfolgend in der Diskussion näher erläutert.

### 2.1.3 Analyse von neu erworbenen Spacersequenzen

Zur Bestätigung der Funktionsfähigkeit des aufgebauten Integrationssystems war es notwendig, die neu eingebauten Spacer-Fragmente auf ihre Größe und Ursprung hin analysieren zu können. Die Schwierigkeit dieser Identifizierung bestand jedoch bei der Isolierung der Plasmid-DNA der neu inserierten Fragmente, genau die Zellen zu detektieren, die nur einen einzigen Typ von Spacern in den Plasmid-basierten CRISPR-Array aufgenommen hatten. Durch Verwendung des pACYC-Duet1 Derivats bei der Klonierung des pCR003 Plasmids konnte durch die geringe Plasmidkopienzahl des Vektors mit nur ca. 10 Kopien pro Zellen die Suche nach richtigen Klonen ein wenig vereinfacht werden. In mehreren Transformations- und Analyseschritten war es möglich, diejenigen Klone, die nur einen 'single' Typ des erweiterten Plasmid-Arrays in der Zelle enthielten zu isolieren (siehe Anhang Abb. 6.2).

Zeigten die Zellen während den Zykluskontrollen neben den parentalen auch zusätzliche Banden, wurden Verdünnungen der Zellkulturen auf selektiven Agarplatten ausgestrichen und am nächsten Tag Kolonien dieser Platten durch colony-PCR auf das Vorhandensein von Spacer-Banden untersucht (5.2.6.6). Bei Einbau bzw. Beobachtung einer Doppelbande (parental und extended) auf dem Agarosegelbild, wurde die entsprechende Kolonie von der angelegten Sicherheitsplatte gepickt und die Plasmid-DNA isoliert (5.2.2.1). Es folgte die Transfomation in kompetente XL-1 oder Top10 Zellen und das Ausplattieren auf Agarplatten, die nur Chloramphenicol als selektives Antibiotikum enthielten. Dadurch konnte sichergestellt werden, dass nur solche Zellen die auch das pCR003 Plasmid tragen wachsen können. In einem weiteren Analyseschritt wurden anschließend Klone die ihren Plasmid-Array durch Hinzufügen neuer Spacer erweitert hatten, mit
Hilfe der *colony*-PCR identifiziert und nach erneuter Präparation der Plasmid-DNA zur Sequenzierung geschickt.

In Abbildung 2.3 ist ein solches Sequenzierungsergebnis als Beispiel dargestellt. Dass der Einbau der neuen Spacersequenz (grün) in unmittelbarer Nähe der Leaderregion erfolgt, wurde zuvor bereits schon in mehreren Studien gezeigt und deutet auf eine direkte Beteiligung des Leaders an der Aufnahme von neuen Spacern hin (Swarts, Mosterd et al. 2012, Yosef, Goren et al. 2012). Gleichzeitig wird deutlich, dass die Integration in Kombination mit einer Duplikation der ersten Repeatsequenz einhergeht, denn das Ergebnis der Sequenzierung zeigt neben der neu integrierten Spacer-Region auch eine weitere Repeatsequenz, wodurch der CRISPR-Array um exakt 61 bp erweitert wird (32 bp Spacer + 29 bp Repeat). Diese Tatsache stützt die bisherige Hypothese das bei der Aufnahme von Fremd-DNA die beiden komplementären Repeat-Stränge durch eine sequenzspezifische Spaltung so gegeneinander verschoben werden, dass sie einer DNA-Polymerase als *template* für eine Duplikation dienen (Abbildung 1.7).

#### Sequenzierung K10-V22

1 GGCTTCTTAT ATACGATGTT ACATTAAGGT TGGTGGGTTG TTTTTATGGG AAAAAATGCT TTAAGAACAA ATGTATACTT CCGAAGAATA TATGCTACAA TGTAATTCCA ACCACCCAAC AAAAATACCC TTTTTTACGA AATTCTTGTT TACATATGAA +1 81 TTAGAGAGTT CCCCGCGCCA GCGGGGATAA ACCTTTATAA GCAGCTCCTG GCCCGCCGAT AAACGAGAGT TCCCCGCGCCC AATCTCTCAA GGGGCGCGGT CGCCCCTATT TGGAAATATT CGTCGAGGAC CGGGCGGCTA TTTGCTCTCA AGGGGCGCGGG Repeat Spacer neu 161 AGCGGGGATA AACCGGCGAC CACACCCGTC CTGTGGATCC TCTACGCGAG TTCCCCGCGC CAGCGGGGAT AAACCGGTCG TCGCCCCTAT TTGGCCGCTG GTGTGGGCAG GACACCTAGG AGATGCGCTC AAGGGGCGCG GTCGCCCCTA TTTGGCCAGC Repeat Spacer alt Repeat 241 ACGAGTTCCC CGCGCCAGCG GGGATAAACC GGGTACCGAG TTCCCCGCGC CAGCGGGGAT AAACCGAAGC TTAATTAACA TGCTCAAGGG GCGCGGTCGC CCCTATTTGG CCCATGGCTC AAGGGGCGCG GTCGCCCCTA TTTGGCTTCG AATTAATTGT Repeat Repeat 321 ACCTGGGCCT CATGGGCCTT CCGCTCACTG TGGACCCGGA GTACCCGGAA GGCGAGTGAC

# Abbildung 2.3: Sequenzierungsergebnis von Plasmid-DNA des pCR003 Plasmids nach Aufnahme eines neuen Spacers.

Zu sehen ist die Sequenz des synthetischen CRISPR-Arrays, der nach Induktion der *cas1* und *cas2* Expression während den Zykluskontrollen einen neuen Spacer aufgenommen hatte. Die palindromischen Repeatsequenzen (rot) weisen eine spezifische Länge von 29 bp auf. In blau ist die bereits enthaltene Spacersequenz mit einer Länge von 32 bp gezeigt. Die neu erworbene Spacersequenz in grün markiert hat ebenfalls eine Länge von 32 bp und stammt aus dem Genom von BL21 AI (Tabelle 2.1).

Nachfolgend sind alle neu erworbenen Spacersequenzen, die mit der Hilfe der *colony*-PCR Methode identifiziert werden konnten, in einer Tabelle zusammengefasst (Tabelle 2.1). Neben der *upstream* Sequenz des Repeats und der Protospacer-Sequenz mit je 30 Basenpaarüberhängen ist auch der Ursprungsort des Protospacers angegeben.

Es kann festgehalten werden, dass die Repeats identisch sind und eine für *E. coli* charakteristische Länge von 29 bp aufweisen. Die Länge der Spacersquenzen

ist ebenfalls wie erwartet mit 32 bp typisch für Typ I-E CRISPR/Cas-Systeme. Die bereits in anderen Studien aufgedeckte Tatsache, dass nach Integration von neuen Spacern am Leader-proximalen Ende eines CRISPR-Locus das letzte Nukleotid der Repeatsequenz von der Sequenz des Protospacers abstammt, konnte ebenfalls bestätigt werden (Datsenko, Pougach et al. 2012). Die Ergebnisse in Abbildung 2.2, sowie die sequenzierten, neu erworbenen Spacer in Tabelle 2.1 belegen die Funktionalität des entwickelten Plasmid-basierten Integrationssystems. Somit war die Grundlage für weiterführende Experimente, wie die Modifizierung der ersten Repeatsequenz innerhalb des Plasmid-Arrays geschaffen.

### Tabelle 2.1: Neu erworbene Spacersequenzen des Plasmid-basierten CRISPR-Arrays.

Die Tabelle enthält identifizierte Spacersequenzen des pCR003 Plasmid-basierten synthetischen CRISPR-Arrays. Neben den *upstream* Repeatsequenzen, sind die Protospacer-Region (Spacersequenz rot markiert) und der Ursprungsort des Protospacers angegeben. Schwarz unterstrichen ist das letzte Nukleotid der Repeatsequenz, welches von der Protospacer-Sequenz abstammt. In blau ist ein *mismatch* innerhalb der Spacersequenz markiert.

		Upstream Repeat sequence	Acquired Spacer	PS-
				origin
1	K17-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>G</u>	AGGCCTAAATCGTCGTGGGTATGTACGGAAATAA	BL21-AI
	V2		TGGCTTTGTCGATGTTAGGCACGCGTTCATACAG	
			GCCGCTGATGATTCCGGCGAACTC	
2	K32-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>G</u>	CTGCCGCATAACGAATCCCTGTTCCAGAAGGGTGA	BL21-AI
	V4		ACTTCCTGATGCAGGCGGCGATGGGTTTTATCAA	
			CATCGAGCAGAACCGCATCATCAA	
3	K10-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>T</u>	ATCGCCAGCAATACCCAGCCCAGCCCGTA <u>T</u> TTAT	BL21-AI
	V22		AAGCAGCTCCTGGCCCGCCGATAAACGAACTGGC	
			ACTGATATAGGTCGCGGTGAGCGT	
4	K78-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>G</u>	GATGAGGATCGTTTCGCATGATTGAACAA <u>G</u> ATGG	pUC18-
	V28		ATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAG	Kan
			AGGCTATTCGGCTATGACTGGGCA	
5	K2-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>A</u>	CTGGACGAAGAGCATCAGGGGCTCGCGCC <u>A</u> GCC	pUC18-
	AP7		GAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCG	Kan
			ACGGCGAGGATCTCGTCGTGACCCA	
6	К44-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>G</u>	TCAGGTACCAGTTGAAGTCCGTCCGGTTC <u>G</u> TCGT	BL21-AI
	AP23		AATGCTCTGGCAATGCGTTGGATCGTTGAAGCTG	
			CTCGTAAACGCGGTGATAAATCCA	
7	K33-	GAGTTCCCCGCGCCAGCGGGGATAAACCC	TCCTTGACGGACAGTTTTCAAAAGATTATGAA	N.D.
	V50			
8	К6-	GAGTTCCCCGCGCCAGCGGGGATAAACC <b>G</b>	CCGCGCATTTGCGCACGTTACTGGGTTAAGGAAG	BL21 AI
	V53		GAGAAGGACAGCGTGGAAATTTACGAGAACGAA	
	•55		AACGACCAGGTAGAAGCGGTTAAAC	
9	K3-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>C</u>	CGCCGGGCTGGCACCTGACTGTGCAATAA <mark>C</mark> CGTA	BL21 AI
	V51		CATACCCAGGCAGATACGTCCAGTCCCGATCCGT	
			TATTTAATCCTCTAAAAACTGGCG	
10	K12-	GAGTTCCCCGCGCCAGCGGGGATAAACC <u>G</u>	CTGATGGTGTCGGCCATCATCGCCGCGTAGTTCC	BL21 AI
			CCTGATTATGCGTCACGGTATGGGTCAGTTCATA	

	V52 **		CGCGGCAAGCTGGTGGTTGATCGC				
11		GAGTTCCCCGCGCCAGCGGGGATAAACC <u>A</u>	CTGCGGCAGATCACTTACCGGGAACATCA <u>A</u> CATA CGGCCTGCCTGAGTGATTGCCAGCAGCATATCGG AAGCATCTTCAATCACCACCGGCG	BL21 AI			
12	K4- V54	GAGTTCCCCGCGCCAGCGGGGATAAACCC	AGTGGATGAATCCCGATAATCGGCCTCCTTCC	N.D.			
**: Integration von zwei Spacern							

N.D.: Keine Übereinstimmung zu (<u>http://blast.ncbi.nlm.nih.gov/Blast.cgi</u>).

## bekannten Sequenzen gemäß

## 2.2 Mutagenese der ersten Repeatsequenz des Plasmidbasierten CRISPR-Arrays

In Studien von Babu und Mitarbeitern konnte eine Nuklease-Beteiligung des *E. coli* Cas1 Proteins gegen verzweigte DNA-Moleküle, die sogenannte 'Holliday-junction'-Strukturen ausbilden, nachgewiesen werden (Babu, Beloglazova et al. 2011). Diese 'Holliday-Struktur' ist wichtiger Bestandteile des *crossing-overs* und begünstigt den Austausch von genetischen Informationen zwischen homologen und nicht-homologen DNA-Strängen.

Dass die Aufnahme neuer Spacer meist am Leader-nahen Ende eines CRISPR-Arrays unter gleichzeitiger Duplikation des ersten Repeats erfolgt zeigt, dass eine Spaltung der Repeatsequenz sehr wahrscheinlich irgendwo an der Leader-Repeat-Grenze stattfindet. Bei näherer Betrachtung der identischen Repeatsequenzen innerhalb eines CRISPR-Array wird deutlich, dass solche 'Kruziform'-ähnlichen Strukturen bedingt durch die palindromische Natur der Repeats ausgebildet werden könnten. Dies führte zu der Hypothese, dass eine so ausgebildete Repeatsequenz während des Integrationsprozesses in der Bakterienzelle von der Cas1 Nuklease erkannt und an bestimmten Positionen geschnitten wird (Abbildung 2.4). Eine Spaltung dieser Region würde somit durch die DNA-Struktur begünstigt werden, könnte aber auch ausschließlich sequenzspezifisch erfolgen. Um dieser Annahme auf den Grund zu gehen, bestand der nächste Schritt in einer Modifizierung des ersten, für die Integration essentiellen Repeats. Anhand verschiedener Punktmutationen, die innerhalb des ersten Repeats eingeführt wurden, sollte der Frage nach eines Struktur-spezifischen oder Sequenz-abhängigen Spaltungsmechanismus nachgegangen werden. Nachfolgend sind die Ergebnisse dieser Studie dargestellt.

BLASTn



**Abbildung 2.4: Putative 'cruciform'-Struktur während der Spacerintegration.** Gezeigt ist ein Schema der ersten Repeatsequenz, die bedingt durch ihre palindromische Natur während des Integrationsprozesses eine sogenannte 'Kreuzförmige-Struktur' ausbilden könnte. Die blauen Pfeile kennzeichnen die Positionen, an denen der Struktur- bzw. Sequenz-spezifische

### Schnitt erfolgt. Unterstrichen sind diejenigen Basen, die nachfolgend modifiziert wurden.

### 2.2.1 Positionswahl und Einführung von Mutation RM1

Die in dieser Arbeit vorgesehene Modifizierung des ersten Repeats des pCR003 Plasmids zielte auf das innerhalb der Sequenz gelegene erste C-Cluster (CCCC, Abb. 2.4), also auf die Position 6-9 des Repeats. Der Mutagenese ging die Fragestellung voraus, ob eine Veränderung dieser Positionen Auswirkungen auf den Einbau von neuen Spacern hat. Durch Modifizierung dieser Region ist die Ausbildung einer hypothetischen 'cruciform-Struktur', wie sie in Abbildung 2.4 gezeigt ist, aufgehoben. Es stellt sich die Frage, ob die Integration von neuen Spacersequenzen bei veränderter Primärsequenz noch erfolgen kann. Durch Modifizierung des Plasmid-basierten CRISPR-Arrays ergibt sich gleichzeitig eine Positivkontrolle für die durchgeführten Integrationsstudien. Wird Spacereinbau im Genom beobachtet, im Plasmid jedoch nicht, lässt dies Rückschlüsse auf die eingebaute Mutationen zu. Für die erste *in vitro* Mutation wurde das Cytosin-Cluster an Position 6-9 zu GGTT (Guanin und Thymin) mittels des QuickChange<sup>®</sup> II Site-Directed Mutagenesis Kit (5.2.6.5) der Firma Stratagene ausgetauscht.

Das Prinzip der Mutagenese beruht auf einer unterschiedlichen Methylierung der parentalen und der mutierten Plasmid-DNA. Daher musste zunächst das Ausgangsplasmid pCR003 aus einem *dam*<sup>+</sup>-Stamm isoliert werden, damit die DNA methyliert vorlag. Durch eine PCR-Reaktion mit den entsprechenden Mutagenese-Primer-Paaren (4.4.1) wurde das Plasmid mit den mutierten Sequenzbasen innerhalb des ersten Repeats amplifiziert. Damit sich die Oligonukleotide trotz des mutations-bedingten *mismatch* stabil an den *template* DNA-Strang anlagern (annealing) konnten wurden Primer mit einer Länge von 36-40 Nukleotiden, einem GC-Gehalt von >50 % und einer Schmelztemperatur ( $T_m$ ) von >77°C konstruiert. Im Anschluss an die PCR-Reaktion liegen in dem Ansatz sowohl Kopien des parentalen Ansatzes, als auch Plasmide, die die Mutation enthalten vor. Die parentale, methylierte Plasmid-DNA wurde anschließend durch das *dam*-Methylierungsabhängige Restriktionsenzym *Dpn*I abgebaut und es blieb die mutierte, doppelsträngige DNA (dsDNA) zurück. Diese lag trotz Einzelstrangbrüchen ('nicks') zirkulär vor und konnte direkt transformiert werden. Die Korrektheit der konstruierten Repeat-Mutante wurde nach Isolierung der Plasmid-DNA (5.2.2.1) durch Sequenzierung bestätigt und ist im Anhang in Abbildung 6.3 dargestellt. Im Anschluss an die Sequenzierung wurde das pCR003 RM1 Plasmid für die anstehenden 'spacer-uptake' Studien in den BL21AI/pUC18-Kan Stamm transformiert und der Einbau von neuen Spacern anhand von PCR-Amplifikationsreaktionen überprüft.

### 2.2.1.1 Funktionelle Analyse der pCR003 RM1 Mutante in vivo

Aufgrund der Tatsache, dass die Mutagenese der ersten Repeatsequenz in dem synthetischen CRISPR-Array des pCR003 Plasmids durchgeführt wurde, konnte eine gleichzeitige Charakterisierung des mutierten pCR003 Plasmids mit dem WT Plasmid erfolgen. Dieser diente durch Analyse sowohl des genomischen, als auch des plasmidären Einbaus zeitgleich als Positivkontrolle in den durchgeführten Integrationsstudien. Wird nur Spacereinbau im Genom beobachtet und im Plasmid nicht, lässt dies Rückschlüsse auf die eingebaute Mutationen zu.

In Abbildung 2.5 sind die Ergebnisse der Integrationsstudien der Mutante RM1 im Vergleich zum Wildtyp (WT-) Plasmid gezeigt.



#### Abbildung 2.5: PCR-Analyse der Spacerintegration in der Mutante RM1.

In (A) sind die Repeatsequenzen des pCR003 WT Plasmids (links) und der RM1 (rechts) gezeigt. Die modifizierten Basen sind in rot gekennzeichnet.

(B) zeigt drei Zyklen (1 bis 3) der Mutation RM1 im pCR003 Plasmid nach PCR-Reaktion auf einem 1,2%igen Agarosegel. Links sind die Positivkontrollen des pCR003 WT Plasmids gezeigt. Rechts sind die Agarosegelbilder der PCR-Ansätze aus den Zellen, die das mutierte Plasmid tragen dargestellt. Der Einbau von neuen Spacersequenzen wurde mit verschiedenen Oligonukleotiden (4.4.1) sowohl im genomischen Array (links), als auch im synthetischen Array der Mutante (rechts) kontrolliert. Parentale PCR-Fragmente sind mit "parental" gekennzeichnet. Darüber liegende Banden sind als neu erworbene Spacersequenzen mit "extended" gekennzeichnet (1x bzw. 2x extended).

Die Abbildung 2.5 zeigt die PCR-Analysen des pCR003 RM1 Plasmids. Es sind je drei Zyklen der Zellen nach Induktion der Plasmid-codierten *cas1-cas2* Gene gezeigt. Um eine Verifizierung der Ergebnisse bestätigen zu können, wurden die Experimente über mehrere Wochen hinweg mit frisch transformierten Zellen reproduziert. Zu sehen ist exemplarisch eine dieser Zykluskontrollen nach Einführung der Mutation RM1 in das WT Plasmid. Links dargestellt ist das pCR003 WT Plasmid, welches neben einer Positivkontrolle auch als Vergleich zum mutierten Plasmid dient. Des Weiteren wurde wie zuvor auch schon, neben dem plasmidären Einbau, der genomische Einbau

kontrolliert. Es wird deutlich, dass der Einbau von neuen Spacern ins Plasmid der Mutante im Vergleich zum Wildtyp reduziert ist (Vergleich der Spuren 3 und 4 mit 8 und 9). Der genomische Einbau in der Mutante erfolgt weiterhin, was zeigt, dass die eingefügte Mutation innerhalb des Plasmid-basierten Systems Auswirkungen auf die Spacerintegration hat (Spuren 6 und 7). Die Modifizierung dieses ersten C-Clusters innerhalb der ersten, Leader-nahen Repeatsequenz führt zum Verlust der Basenpaarung, wodurch folglich auch die Ausbildung der hypothetischen 'Kreuzstruktur' nicht mehr erfolgen kann. Vergleicht man die Intensität des Spacereinbaus in beiden Zellen, kann eine Abnahme nach Zyklustag 3 beobachtet werden. Dies konnte auch schon in vorherigen Studien mit dem Wildtyp Plasmid gezeigt werden und deutet auf eine verringerte Zellzahl nach mehreren Zyklen hin. Ein Wachstumsvergleich der Zellen über mehrere Zyklen zeigte jedoch keine signifikanten Unterschiede im Wachstumsverlauf (Daten nicht gezeigt).

Nach Transformation des mutierten pCR003 Plasmids in BL21 AI/pUC18-Kan Zellen konnte ein Verlust der Fähigkeit zur Aufnahme von neuen Spacersequenzen nachgewiesen werden. Als nächstes galt es zu überprüfen, ob dieser Verlust Strukturoder Sequenz-spezifisch erfolgt. Dazu wurden in einem weiteren Schritt dieser Arbeit erneut Modifizierungen in das bereits mutierte pCR003 RM1 Plasmid eingefügt, so dass eine Basenkomplementarität zwischen den einzelnen Purinen und Pyrimidinen in Form der Wildtyp-Situation wiederhergestellt wurde.

### 2.2.2 Einführung und *in vivo* Analyse von Mutation RM1-R

Bei der Einführung der Rückmutation RM1-R wurde die gleiche Herangehensweise und Durchführung wie zu der bereits beschriebenen Mutation RM1 in Kapitel 2.2.1 gewählt. Es mussten lediglich andere Oligonukleotide in die Mutagenesereaktion eingesetzt werden. Das pCR003 RM1 Plasmid diente dabei als Ausgangsvektor für die PCR-Reaktion.

Währenddessen die erste Mutation (RM1) den Verlust der Strukturausbildung innerhalb der ersten Repeatsequenz zur Folge hatte, sollte die hier eingeführte Mutation als Rückmutation der vorherigen dienen, um eine mögliche 'cruciform' Ausbildung wiederherstellen zu können. Dafür mussten das gegenüberliegende G-Cluster der zuvor eingefügten Basenaustausche modifiziert werden (Abbildung 2.6 (A)). Nach Einführung der Nukleinbasen AACC anstelle von GGGG wurde die Mutagenesereaktion durch Sequenzierung bestätigt (siehe Anhang Abb. 6.4) und die Aufnahme neuer Spacer über mehrere Tage hinweg analysiert. Das Ergebnis dieser Analysen ist Abbildung 2.6 gezeigt.



#### Abbildung 2.6: PCR-Analyse der Spacerintegration in der Mutante RM1-R.

In (A) sind die Repeatsequenzen des pCR003 WT Plasmids, sowie der RM1 und RM1-R Plasmide gezeigt. Die modifizierten Basen sind in rot und blau gekennzeichnet.

(B) zeigt drei Zyklen (2 bis 4) der Mutation RM1-R im pCR003 Plasmid nach PCR-Reaktion auf einem 1,2%igen Agarosegel. Als Positivkontrollen wurden neben dem pCR003 WT Plasmid, auch die Mutante RM1 mit untersucht (Links und Mitte). Rechts dargestellt sind die Agarosegelbilder der PCR-Ansätze aus den Zellen, die das mutierte Plasmid tragen. Der Einbau von neuen Spacersequenzen wurde mit verschiedenen Oligonukleotiden (4.4.1) sowohl im Genom, als auch im Plasmid kontrolliert. Parentale PCR-Fragmente sind mit "parental" gekennzeichnet. Darüber liegende Banden sind als neu erworbene Spacersequenzen mit "extended" gekennzeichnet (1x bzw. 2x extended).

Zu sehen sind drei Zyklen der PCR-Analysen nach Einführung der Rückmutation des pCR003 RM1 Plasmids. Wie schon zuvor handelt es sich hierbei um eine repräsentative Auswahl der reproduzierten Adaptationsstudien die exemplarisch genommen wurde. Das pCR003 WT Plasmid wurde als Positivkontrolle in den Spuren 1-4 mit untersucht. Die Spuren 6-9 zeigen die vorherige Mutante RM1 im direkten Vergleich zur Rückmutante RM1-R (Spuren 11-14). In allen drei Zelltypen wurde neben dem genomischen, auch der plasmidäre Einbau kontrolliert. Vergleicht man die Agarosegelbilder der einzelnen Zyklus-kontrollen, ist der Verlust des plasmidären Einbaus von neuen Spacern auch in der Rückmutante RM1-R erkennbar. Der Einbau ins Genom erfolgt dagegen in allen drei Zelltypen. Nach starker Zunahme an

Zyklustag 2, lässt sich im WT und der Mutante RM1 eine Abnahme bereits am dritten Tag beobachten. Der genomische Einbau der Rückmutante RM1-R hingegen zeigt eine gleichbleibende Aufnahme (Spuren 11 und 12). Der Verlust des Spacereinbaus in beiden Mutanten deutet auf einen nicht-strukturspezifischen Mechanismus während der Integration hin. Denn der Modifizierung dieser Nukleinbasen ging die Fragestellung voraus, ob eine Veränderung an dieser Position, ähnlich zu der Wildtyp-Situation, Auswirkung auf den Einbau hat. Nach Wiederherstellung der DNA-Struktur des Repeats bleibt die Aufnahme von neuen Spacern dennoch unterbunden, was eher auf eine sequenzabhänge Spaltung während des Integrationsmechanismuses hindeutet.

Der bestehende Verlust der Rückmutante RM1-R zur Aufnahme von neuen Spacersequenzen weist auf eine Sequenzspezifität statt einer Sekundärstrukturspezifität innerhalb der ersten Repeatsequenz hin. Daher galt es zu überprüfen, ob diese Sequenzspezifität Nukleinbasen-abhängig erfolgt. Dazu wurde in einer weiteren Mutagenesereaktion das pCR003 RM1-R Plasmid erneut modifiziert.

### 2.2.3 Einführung und in vivo Analyse der Mutation RM2

Die Einführung der letzten Mutation RM2 erfolgte vergleichbar zu den zwei vorherigen Mutationen RM1 und RM1-R und zielte auf eine direkte Sequenzveränderung der ersten Repeatsequenz im Vergleich zur Wildtyp-Situation hin, jedoch ohne Verlust der DNA-Strukturausbildung. Hierbei diente das pCR003 RM1-R Plasmid als Ausgangsvektor für die Mutagenesereaktion.

Die Mutation trägt eine Veränderung im C-Cluster bezogen auf den Wildtyp durch den Austausch der vier aufeinanderfolgenden Cytosin Basen zu vier Guanin Basen (Abbildung 2.7 (A)) und wurde durch Sequenzierung bestätigt (siehe Anhang Abb. 6.5). Wird in der Wildtyp-Situation im Plasmid Spacereinbau beobachtet, in dem mutierten Vektor unter gleichen Bedingungen jedoch nicht, zeigt dies den Einfluss der Modifizierung auf die Spacerintegration. Einzelne Sequenzbasen innerhalb der ersten Repeatsequenz, die essentiell an der Integration beteiligt sind, könnten so aufdeckt werden. Das Ergebnis einer dieser Integrationsstudien ist in Abbildung 2.7 dargestellt.



#### Abbildung 2.7: PCR-Analyse der Spacerintegration in der Mutante RM2.

In (A) sind die Repeatsequenzen des pCR003 WT Plasmids und der Mutante RM2 gezeigt. Die modifizierten Basen sind in rot und blau gekennzeichnet.

(B) zeigt drei Zyklen (1 bis 3) der Mutation RM2 im pCR003 Plasmid nach PCR-Reaktion auf einem 1,2%igen Agarosegel. Als Positivkontrolle wurde das pCR003 WT Plasmid (links) ebenfalls untersucht. Rechts dargestellt sind die Agarosegelbilder der PCR-Ansätze aus den Zellen, die das mutierte Plasmid tragen. Die Kontrolle des Einbaus von neuen Spacersequenzen erfolgte mit verschiedenen Oligonukleotiden (4.4.1) sowohl im Genom, als auch im Plasmid. Parentale PCR-Fragmente sind mit "parental" gekennzeichnet. Darüber liegende Banden sind als neu erworbene Spacersequenzen mit "extended" gekennzeichnet (1x bzw. 2x extended).

Bei der Kontrolle der Spacerintegration wurde deutlich, dass auch diese Mutation Einfluss auf den Einbau neuer Spacer nimmt. In drei untersuchten Zyklen konnten lediglich die parentalen Banden in den Zellen, die das mutierte Plasmid tragen, beobachtet werden (Spuren 8 und 9). Genomisch hingegen erfolgt der Einbau neuer Spacersequenzen konstant über alle drei Zyklen hinweg. Auch die plasmidäre Kontrolle des Wildtyps zeigt, dass der Einbau in den Plasmid-basierten synthetischen CRISPR-Array stattfindet. Eine Abnahme der Spaceraufnahme ist diesmal weder im Wildtyp, noch in der Mutante zu sehen. Da alle Zellen derselben Behandlung unterlagen, kann die Reduktion der Spacerintegration auf die eingebaute Mutation zurückgeführt werden. Die modifizierten Sequenzbasen sind somit entscheidend für den Einbau neuer Spacer in das Plasmidsystem.

Bei der vorliegenden Mutation wurde das C-Cluster von CCCC zu GGGG mutiert und der Einfluss auf den Einbau neuer Spacer durch mehrere Zykluskontrollen getestet. Bei allen durchgeführten Zyklen waren lediglich die parentalen Banden im Agarosegel sichtbar. Da die Positivkontrollen eine Überexpression der *cas*-Gene belegen und Spacereinbau sowohl ins Genom, als auch ins Wildtyp-Plasmid bei denselben Zyklen beobachtet wurde, lässt sich daraus ein Einfluss der Mutation auf die Spacerintegration erschließen. Der Austausch dieser vier Basenpaare invers zur Ausgangssituation verhindert somit den Einbau von neuen Spacern und lässt die Vermutung zu, dass das C-Cluster innerhalb des ersten Repeats essentiell für die Integration ist. Im Gegensatz dazu scheint in den beiden anderen Mutationen RM1 und RM1-R, in denen eine DNA-Strukturabhängige Integration getestet wurde, die DNA-Struktur alleine nicht für eine aktive Integration verantwortlich zu sein. Der Ausschluss dieser Strukturspezifität deutet auf eine Sequenz-abhängige Spaltung während der Integration von neuen Spacern hin.

Die Charakterisierung des Leader-proximalen Repeats anhand der drei verschiedenen Mutationen deckte erste Hinweise in Bezug auf die Frage nach einer DNA-Struktur- oder Sequenz-spezifischen Spaltung während der Integration neuer Spacer auf. Nach Einführung der Mutationen konnte in den durchgeführten Integrationsstudien in allen mutierten Plasmiden eine signifikante Reduzierung der Aufnahme von neuen Spacern gegenüber den Positivkontrollen in Form des pCR003 WT Plasmids beobachtet werden. Durch die Rückmutation des pCR003 RM1 Plasmids, die eine 'cruciform'-Strukturausbildung zwischen den einzelnen Nukleinbasen innerhalb des Repeats wieder ermöglicht, konnte gleichzeitig eine rein strukturabhängige Integration ausgeschlossen werden.

Das derzeitige Modell zur Spacerintegration schlägt einen sequenzspezifisch-, versetzten Schnitt an beiden DNA-Strängen des Leader-nahen Repeats vor (Abbildung 1.7). Nach dieser Spaltung können so die freien 5'-Repeat-Enden mit den 3'-Enden des neuen Spacers verknüpft werden und die entstandenen einzelsträngigen Repeat-Lücken mittels DNA-Polymerase aufgefüllt und somit dupliziert werden. Anhand von Southern Blot Analysen mit genomischer DNA aus BL21 AI konnten bereits kurze DNA-Fragmente, sogenannte Zwischenstufen der Integration detektiert werden, die diese Art von Mechanismus begünstigen (Arslan, Hermanns et al., in Revision). Daher galt es als nächstes den Mechanismus der CRISPR-Adaptation anhand des Plasmid-basierte CRISPR-System näher zu analysieren.

### 2.3 Charakterisierung des Adaptations-Mechanismus

Die CRISPR/Cas-vermittelte Adaptation gegen fremde Nukleinsäuren ist abhängig von der Insertion unbekannter DNA-Sequenzen als neue Spacer in den eigenen CRISPR-Array. Bekannt ist, dass der Einbau dieser Fremd-DNA Sequenzen meist am Leader-nahen Ende unter Duplikation der ersten Repeatsequenz einhergeht, was auf eine direkte Beteiligung beider DNA-Stränge des Repeats als *template* für eine Auffüllreaktion der Polymerase hinweist. Die Tatsache, dass einem DNA-Integrationsprozess ein Spaltungsereignis vorangehen muss, führte zu der Annahme, dass auch in *E. coli* während der CRISPR-Adaptation eine solche Spaltungsreaktion erfolgt. Die dabei entstehenden Zwischenstufen (Intermediate) würden Aufschluss über die Art und Weise des Integrationsmechanismuses geben.

In ersten Untersuchungen des genomischen CRISPR-Arrays von *E. coli* BL21 AI konnten bereits Zwischenprodukte dieser Spaltungsreaktion detektiert werden und basierend darauf ein gekoppelter Spaltungs-Ligationsmechanismus aufgedeckt werden (Arslan, Hermanns et al., in Revision). Um die Bildung der Zwischenstufen auch im Plasmid-basierten CRISPR-System zu bestätigen und eine Repeat-abhängige Spaltung mit gekoppelter Ligation des neuen Spacers testen zu können, wurde mit Hilfe von *in vivo* Analysen der Existenz solcher Zwischenprodukte nachgegangen.

### 2.3.1 Detektion von Spacerintegrations-Intermediaten des Plasmidbasierten CRISPR-Arrays

Für die Aufdeckung dieser vorübergehenden Zwischenstufen eignet sich die Methode der Southern Blot Analyse (5.4.2). Hierbei erfolgt der Nachweis von Fragmenten mit Hilfe von radioaktiv markierten Sonden direkt an Plasmid-DNA, die zuvor einem Restriktionsdau unterzogen wurde. Die Wahl der Restriktionsenzyme und DNA-Sonden orientiert sich dabei an der Fragestellung, welche DNA-Sequenzen untersucht werden sollen.

Zur Detektion von *in vivo* Intermediaten während der Integration neuer Spacer-DNA wurden Southern Blot Analysen mit dem pCR003 WT Plasmid, sowie mit mutierten Versionen des WT Plasmids (RM1 und RM2) durchgeführt. Dafür wurde die Plasmid-DNA der Zellen nach 18-stündiger Inkubation (Zyklus 1) mit oder ohne Induktion der *cas1-cas2*-Expression isoliert (5.2.2.2) und mit zwei verschiedenen Restriktionsenzymen (*EcoRI* oder *KpnI*) hydrolysiert. Nach Hitzeinaktivierung des jeweiligen Enzyms wurde je 1 µg Plasmid-DNA auf einer 10% dPAGE getrennt (5.2.4.3) und anschließend auf eine Nitrocellulose-Membran übertragen. Es folgten ein 1 stündiger Elektroblot bei 600 mA und die Fixierung der nach dem Transfer auf die Membran übertragenen DNA-Fragmente durch Bestrahlung mit UV-Licht (UV-Crosslinking). Mit Hilfe von radioaktiv markierten Oligonukleotiden, die komplementär zu beiden DNA-Strängen *upstream* der Leaderregion oder der

Spacersequenz des synthetischen CRISPR-Arrays hybridisierten, konnte so der gesamte CRISPR-Array auf die Existenz möglicher Zwischenstufen getestet werden (Abbildung 2.8).



#### Abbildung 2.8: Schema der Southern Blot Analyse des synthetischen CRISPR-Arrays.

Gezeigt ist eine schematische Darstellung der Versuchsdurchführung für die Southern Blot Analysen des synthetischen Plasmid-basierten CRISPR-Arrays nach Induktion der *cas1-cas2* Induktion. Während der Spacerintegration wird der synthetische CRISPR-Array geöffnet und um eine neue Spacer-Repeat Einheit erweitert. Nach Isolierung der Plasmid-DNA und Hydrolyse mit *EcoRI* oder *KpnI* wird in Southern Blot Experimenten nach den während des Integrationsprozesses entstehenden Zwischenprodukten mit Hilfe verschiedener Oligonukleotide (1 und 2) gesucht.

### 2.3.2 Verifizierung der *in vivo* Intermediate durch Southern Blot Analysen

Nach der Transformation der Plasmide pCR003 WT, sowie der Mutanten RM1 und RM2 in kompetente *E. coli* BL21 AI Zellen war die Analyse der isolierten Plasmide nach Restriktion mit verschiedenen Restriktionsenzymen notwendig, um die Entstehung von *in vivo* Intermediaten zu prüfen. Diese postulierten Intermediate sollten bei der Integration von neuen Spacersequenzen in den synthetischen CRISPR-Array der Plasmide durch eine sequenzspezifische Öffnung des Arrays vor Einbau des neuen Spacers entstehen. Dem aktuellen Modell zur Folge, wird der erste Leader-proximale Repeat an zwei Positionen, der Leader-Repeat-Grenze und unmittelbar nach der ersten Repeatsequenz versetzt an beiden DNA-Strängen gespalten. Die neu erworbene Spacersequenz wird in die entstandene Lücke eingebaut und die beiden einzelsträngigen, identischen Repeat-Bereiche zu beiden Seiten des neuen Spacers in einer Polymerase-Reaktion aufgefüllt und dupliziert. In Studien von Southern Blot Analysen mit genomischer DNA von BL21 AI konnten bereits Zwischenstufen der Spacerintegration von unterschiedlicher Längen detektiert werden (Arslan, Hermanns et al., in Revision). Anhand des entwickelten Plasmid-basierten CRISPR-Systems wurde die Bildung solcher Zwischenstufen untersucht und dabei das Modell eines gekoppelten Spaltungs-Ligations-Mechanismus in Typ I-E CRISPR/Cas-Systemen bestätigt.

Für die Detektion wurde Plasmid-DNA der Stämme BL21 AI/pCR003 WT, BL21 AI/pCR003 RM1 und BL21 AI/pCR003 RM2 genutzt. Die Analyse der Zwischenstufen erfolgte durch Hybridisierung mit radioaktiv markierten Oligonukleotiden (4.4.1), die komplementär zu verschiedenen Sequenzen innerhalb des synthetischen CRISPR-Arrays sind. Die Restriktion der isolierten Plasmid-DNA Proben mit unterschiedlichen Restriktionsenzymen erlaubte es, potentiell wichtige Bereiche des CRISPR-Arrays, an denen eine Spaltung der Nukleinbasen vermutet wird, mit einer höheren Auflösung untersuchen zu können. Bei einer Hydrolyse mit *EcoRI* erfolgte der Schnitt *upstream* der Leaderregion. Eine Behandlung mit *KpnI* führte zu einer Spaltung zwischen dem dritten und vierten Repeat. Nach Hybridisierung verschiedener Oligonukleotide in unmittelbarer Nähe der Restriktionsschnitte konnte so der komplette synthetische CRISPR-Array analysiert und entstehende Intermediate detektiert werden.

Gleichzeitig sollte untersucht werden, ob die Bildung der Intermediate abhängig von der Expression der Plasmid-codierten *cas1* und *cas2* Gene sind, welche essentielle Bestandteile der CRISPR-Adaptation sind. Dafür wurden wildtypische BL21 AI/pCR003 Zellen ohne die Zugabe von IPTG und L-Arabinose angezogen und analysiert.

Die Ergebnisse der Southern Blot Analysen des pCR003 WT Plasmids, sowie der modifizierten RM1 und RM2 Plasmide nach Hydrolyse mit dem Restriktionsenzym *EcoRI* zur Analyse der Leaderregion sind in Abbildung 2.9 dargestellt.



#### Abbildung 2.9: Southern Blot Analyse der Leaderregion des synthetischen CRISPR-Arrays.

Oberhalb der Ergebnisse ist ein Schema der erwarteten Fragmentlängen im Fall einer *cas1/cas2-*induzierten Spaltung und Ligation eines neuen Spacers (SO) gezeigt. Darunter sind die Autoradiogramme der Southern Blot Analysen von 1 µg *EcoRI* hydrolysierter Plasmid-DNA nach dPAGE zu sehen. Die Zellen wurden für die Isolierung der Plasmid-DNA nach 18 stündigem Wachstum mit und ohne Induktion der *cas1-cas2-*Expression geerntet. Die Spuren 1 und 5 zeigen radioaktiv markierte Marker (100 bp Leiter, pUC19/MspI), die als Längenstandards benutzt wurden. Die Hybridisierung erfolgte mit radioaktiv markierten Oligonukleotiden (pCR001-Blot-EcoRI-rv (1) und pCR001-Blot-EcoRI-fw (2)) die komplementär zu beiden DNA-Strängen der Leaderregion sind. Die detektierten Fragmentlängen sind durch Pfeile gekennzeichnet.

Anhand der Daten wird ersichtlich, dass die Spaltungsintermediate nur in den *cas1-cas2* induzierten Plasmiden detektierbar sind (Spuren 3, 4 und 6). Im pCR003 WT Plasmid ohne Induktion der *cas*-Genexpression können keine Spaltungsprodukte beobachtet werden (Spur 2), was die Notwendigkeit der *cas1* und *cas2* Gene in der Integrationsphase des CRISPR/Cas-Systems anzeigt. Bei der Charakterisierung der Leaderregion des synthetischen CRISPR-Arrays konnten mehrere Fragmentbanden detektiert werden. Wird das Oligonukleotid 1 komplementär zum unteren Strang eingesetzt, sind Banden mit einer ungefähren Länge von 120 Nukleotiden zu erkennen, was auf einen direkten Schnitt an der Leader-Repeat-Grenze hindeutet (Primer 1 in dem Schema). Bei Überprüfung des Gegenstranges lässt sich ein Signal von 180 nt detektieren (Spur 3). Die Ermittlung der Fragmentlänge anhand der Sequenz der Plasmid-DNA zeigt, dass die detektierte Bande nicht nur durch Spaltung des ersten Repeats zu erklären ist, sondern vermutlich das Ergebnis eines

gekoppelten Mechanismus sein muss, in dem die Spaltung des Repeats in Kombination mit der Ligation einer neuen Spacersequenz einhergeht. Die Länge der Bande in Spur 3 ist durch die Aufnahme einer neuen Spacersequenz mit bereits dupliziertem Repeat zu erklären. Die eingeführten Mutationen innerhalb der ersten Repeatsequenz (RM1 und RM2) haben ebenso einen Einfluss auf die Bildung der Intermediate. Wo hingegen die Mutante RM1 in Spur 3 bei Hybridisierung mit dem ersten Primer eine schwächere Bande bei 120 nt gleich dem Wildtyp Plasmid zeigt, ist keine Bande im Gegenstrang sichtbar. Eine Spaltung des oberen Stranges an der Leader-Repeat-Grenze scheint zu erfolgen, aber die Integration eines neuen Spacers nicht. Das Ergebnis passt zu den vorherigen Integrationsstudien, in denen eine Aufnahme von neuen Spacern nach Modifizierung des WT Plasmids reduziert ist (Abbildung 2.5). Bei der Mutante RM2 hingegen sind Signale mit einer Größe von 140 nt (Primer 1 Spur 6) und 330 nt (Primer 2 Spur 6) zu beobachten. Diese Fragmentlängen passen jedoch nicht zu dem bisher vermuteten Integrations-Modell. Es handelt sich hierbei wahrscheinlich um unspezifische Spaltungsprodukte, die während des Cas1 spezifischen Schnittes erfolgt sind. Deutlich ist jedoch, dass die eingeführten Modifizierungen einen direkten Einfluss auf die Spezifität der Spaltungs-Ligations-Reaktion haben, was auf eine Sequenzspezifität von Cas1 und Cas2 hinweist, die durch die erste Leader-nahe Repeatsequenz vermittelt wird. Im Vergleich zum Wildtyp ist die Bildung von Intermediaten in den mutierten pCR003 Plasmiden reduziert bzw. unterbunden.

Die Charakterisierung der Repeat-Spacer-Einheit mit Kpnl hydrolysierter DNA der jeweiligen Plasmide, ist in Abbildung 2.10 gezeigt. Die Ergebnisse nach Hybridisierung der Primer drei und vier, die komplementär zu beiden DNA-Strängen der synthetischen Spacersequenz sind, lassen ähnliche Signale wie bei der Leaderregion zuvor erkennen. Auch hier sind die Spaltungsintermediate erst nach Induktion der Expression der cas1-cas2 Gene sichtbar (Spuren 3, 4 und 6). Das uninduzierte WT Plasmid lässt keine Banden erkennen (Spur 2). Bei der Analyse der **Repeat-Spacer-Region** des synthetischen **CRISPR-Arrays** können Banden unterschiedlicher Längen nachgewiesen werden. Nach Hybridisierung des dritten Primers komplementär zum unteren Strang der Spacersequenz sind Banden mit einer Länge von 160 Nukleotiden zu erkennen. Schaut man sich das Schema in Abbildung 2.10 (oben) der erwarteten Fragmentlängen an, wird deutlich, dass die detektierten Banden in Spur 3 und 4 durch die Aufnahme einer neuen Spacersequenz mit bereits dupliziertem Repeat zu erklären sind. Bei Überprüfung des Gegenstranges lässt sich ein Signal von 97 nt detektieren (Spur 3), was auf einen direkten Schnitt des 3'-Endes der ersten Repeatsequenz hindeutet (Primer 4 Schema). Des Weiteren zeigen auch hier die Analysen der mutierten Plasmide eine reduzierte bzw. unterbundene Bildung von Zwischenstufen (Spuren 4 und 6). Schaut man auf den oberen Strang, ist eine Reduzierung in der Mutante RM1 im Gegensatz zum Wildtyp Plasmid erkennbar (Spuren 3 und 4). Keine Signale sind wiederum im Gegenstrang zu beobachten (Vergleich der Spuren 3 und 4). Der Effekt des Verlustes der Entstehung von Intermediaten ist besonders in der RM2 Mutante erkennbar. Wo hingegen nach

Hybridisierung mit Primer 3 ein deutliches Signal weiter *upstream* der eigentlichen Schnittstelle bei ca. 242 nt zu sehen ist, wurde keine Bande im Gegenstrang mit Primer 4 detektiert (Spuren 6). Dieses veränderte Spaltungsmuster, welches auch schon zuvor bei Analysen der Leaderregion von Mutante RM2 (Abb. 2.9 Spuren 6) beobachtet wurde, zeigt nochmals die Notwendigkeit einer korrekten Repeatsequenz während der Integration.



# Abbildung 2.10: Southern Blot Analyse der Repeat-Spacer-Region des synthetischen CRISPR-Arrays.

Oberhalb der Ergebnisse ist ein Schema der erwarteten Fragmentlängen im Fall einer cas1/cas2-induzierten Spaltung und Ligation eines neuen Spacers (SO) gezeigt. Darunter sind die Autoradiogramme der Southern Blot Analysen von 1 µg *Kpnl* hydrolysierter Plasmid-DNA nach dPAGE zu sehen. Die Zellen wurden für die Isolierung der Plasmid-DNA nach 18 stündigem Wachstum mit und ohne Induktion der *cas1-cas2*-Expression geerntet. Die Spuren 1 und 5 zeigen radioaktiv markierte Marker (100 bp Leiter, Mspl/pUC19), die als Längenstandards benutzt wurden. Die Hybridisierung erfolgte mit radioaktiv markierten Oligonukleotiden (pCR001-Blot-C-S1-rv (3) und pCR001-Blot-S1-fw (4)) die komplementär zu beiden DNA-Strängen der synthetischen Spacersequenz sind. Die detektierten Fragmentlängen sind durch Pfeile gekennzeichnet.

Die Auswertung der detektierten Signale der dargestellten Southern Blot Analysen zeigt, dass die gebildeten Zwischenprodukte nicht das Produkt einer einzelnen Spaltungsreaktion sein können. Anhand des Schemas in Abbildung 2.11 wird ersichtlich, dass es sich um eine gekoppelte Integrasereaktion des Cas1 Proteins handelt, in der die Spaltung der Repeatsequenz an der Leader- bzw. Spacerregion mit einer gleichzeitigen Verknüpfung der freien Enden von Repeat und Spacer einhergeht. Der *non-template* Strang wird an der Leader-Repeat Grenze geschnitten und das freie 5'-Ende des duplizierten Repeats mit dem 3'-Ende des neu aufgenommenen Spacers verknüpft (Abb. 2.9 120 nt und Abb. 2.10 160 nt). Genau anders herum verhält es sich mit dem *template*-Strang, der in der ersten Repeat-Spacer Einheit gespalten wird und das freie 5'-Ende des Repeats mit dem 3'-Ende des neuen Spacers verknüpft (Abb. 2.9 182 nt und Abb. 2.10 97 nt).



# Abbildung 2.11: Schema der gekoppelten Integrasereaktion während der Integration eines neuen Spacers.

Zu sehen ist ein Schema der erwarteten Fragmentlängen nach Hybridisierung der Oligonukleotide 1-4 an die hydrolysierte Plasmid-DNA der pCR003 Plasmide. Die Spaltung der Repeatsequenz an der Leader- bzw. Spacerregion geht mit einer gleichzeitigen Verknüpfung der freien Enden von Repeat und Spacer einher. Der CRISPR-Array wird um eine neue Spacersequenz (S0) erweitert.

Zur Verifizierung dieser Ergebnisse wurde eine bereits geblottete Membran, die mit EcoRI hydrolysierten Plasmid-Proben des pCR003 WT und der Mutante RM1 behandelt wurde, mit Sonden der Kpnl gedauten Repeat-Spacer-Region inkubiert. Dafür mussten zunächst die bisherigen Sonden (pCR001-Blot-EcoRI-rv und pCR001-Blot-EcoRI-fw) durch mehrmaliges Waschen (3 x 30 Min.) mit Aqua dest. bei 80°C heruntergewaschen werden. Anschließend konnte die Membran für eine erneute Hybridisierung mit Primern, die komplementär zu beiden Strängen der synthetischen Spacersequenz des CRISPR-Arrays sind behandelt werden. Das Ergebnis dieser Analyse ist im Anhang in Abbildung 6.6 dargestellt. Es konnten keine signifikanten Signale beobachtet werden. Dieses Ergebnis ist zu erwarten gewesen, da nach Hydrolyse mit EcoRI nur der Leaderbereich analysiert werden kann, aber für eine neue Hybridisierung der bereits geblottenen Membran Primer verwendet wurden, die an der synthetischen Spacersequenz binden. Eine Detektion von DNA-Intermediaten ist mit den gewählten Primern nicht möglich und bestätigt abermals das Modell in Abbildung 2.11. Ähnliches konnte in Southern Blot Analysen mit genomischer DNA von E. coli BL21 AI gezeigt werden (Arslan, Hermanns et al., in Revision).

### 2.3.3 Bestimmung der DNA-Intermediate durch Primer-Extension mit Sequenase 2.0

Um parallel zu den durchgeführten Southern Blot Analysen eine weitere Methode zur Verifizierung der detektierten Zwischenstufen vorweisen zu können, wurde in einem nächsten Schritt versucht, die gebildeten Intermediate in einer Sequenzierungsreaktion mit Sequenase 2.0 zu bestätigen. Für die quantitative Analyse von Tanskriptmengen eignet sich die Methode der Primer-Extension Reaktion.

Für die Charakterisierung der gebildeten DNA-Intermediate wurde die Sequenase 2.0 DNA-Polymerase der Firma USB (4.5.2) für eine Primer-Extension Reaktion verwendet. Hierbei wird die zu analysierende DNA von der T7 DNA-Polymerase als *template* genutzt und ein komplementärer Gegenstrang synthetisiert. Wird der DNA-Doppelstrang während des Integrationsprozesses an den zuvor durch Southern Blot detektierten Sequenzen geschnitten, kommt es zur Termination der Polymerase an der entsprechenden Stelle. Die entstandenen Produkte geben durch ihre Länge, abhängig vom verwendeten Oligonukleotid, indirekt die Fragmentlänge der DNA wieder und können über dPAGE aufgetrennt werden (5.2.4.3).

Für die Bestimmung der Längen der Sequenzierungsprodukte wurde Plasmid-DNA von Zellen des Stammes BL21 AI/pCR003 WT mit und ohne Induktion der Cas1/Cas2 Proteine nach 18 stündiger Inkubation bei 37°C durch Midi-Präparation (5.2.2.2) isoliert. Als *template* für die darauffolgende Reaktion mit der Sequenase 2.0 Polymerase dienten die nicht denaturierten ccc-Plasmide des pCR003 Wildtyps. Die Hybridisierungsreaktion an die zu untersuchende DNA wurde mit  $\gamma^{-32}$ [P]ATP markierten Oligonukleotiden (4.4.1) durchgeführt und die gebildeten Produkte über ein 10 %iges PAA-Gel aufgetrennt. Die Ergebnisse dieser Sequenzierungsreaktion wurden anschließend mittels Autoradiographie sichtbar gemacht und sind in Abbildung 2.12 zu sehen. Die Abbildung zeigt zwei unabhängige Experimente, bei denen unterschiedliche DNA-Konzentrationen eingesetzt wurden. Wo hingegen in (A) 2 µg der Plasmide in die Reaktion eingesetzt wurden, zeigt Abbildung 2.12 (B) das Ergebnis der fünffach eingesetzt wurden, folgte eine Phenol/Chloroform Aufreinigung mit anschließender Vereinigung der Proben und der elektrophoretischen Trennung.

**(A)** 



Abbildung 2.12: Primer-Extension Reaktion mit Sequenase 2.0 und pCR003 WT Plasmid.

Zu sehen sind die Autoradiogramme einer Sequenzierungsreaktion mit Sequenase 2.0 nach dPAGE. In (A) wurden je 2 µg des pCR003 WT Plasmids mit und ohne Induktion der *cas1/cas2*-Genxpression als *template* eingesetzt und mit je 1 µl der radioaktiv markierten Oligonukleotide (PE-Leader-pCR001-fw und pCR001-Blot-C-S1-rv) hybridisiert. Zur genauen Größenbestimmung wurde in den Spuren 2 und 6 ein radioaktiv markierter Marker (pUC19/MspI) aufgetragen. Die Spuren 1 und 5 (O) enthalten das jeweilig verwendete Oligonukleotid alleine. Die durch Spaltung des Cas1 entstandenen spezifischen Banden sind durch rote Sterne gekennzeichnet.

(B) zeigt dieselben Ansätze mit insgesamt 10  $\mu$ g nicht denaturierter Plasmid-DNA in An- und Abwesenheit der induzierten Cas1/Cas2 Proteine.

Wie zu erwarten ist beim Vergleich der Intensitäten der Produktbanden in Abbildung 2.12 die Ausbeute in (B), in der insgesamt 10 µg Plasmid-DNA in die Reaktion eingesetzt wurde wesentlich höher als bei (A). Auffällig ist, dass beide Autoradiogramme starke Hintergrundbanden aufweisen, die auf unspezifische Produkte der parentalen, nicht denaturierten Plasmid-DNA zurückzuführen sind. Ein Induktionseffekt in Bezug auf die Bildung der Produktlängen der jeweiligen Plasmide ist gut zu erkennen. In Abwesenheit der Cas1/Cas2 Induktion sind keine Banden im nicht induzierten pCR003 WT Plasmid zu beobachten (Spuren 3 und 7). Im Vergleich

zeigen die Spuren 4 und 8, in denen die Zellen vor der Isolation der Plasmid-DNA mit IPTG und L-Arabinose behandelt wurden, spezifische Bandengrößen von 87 nt und 100 nt. In vorherigen Southern Blot Analysen konnte die Bildung von DNA-Intermediaten bereits gezeigt werden (2.3.2). Auch konnten die exakten Schnittstellen anhand der detektierten Signale mit Hilfe eines Sequenzschemas bestimmt werden (Abb. 2.9 und Abb. 2.10). Schaut man sich in diesem Fall die Sequenz des pCR003 WT Plasmids nach Hybridisierung mit den jeweiligen Oligonukleotiden an, kann eine Übereinstimmung der beobachteten Produkte mit den zuvor durch Southern Blot detektierten Signalen festgestellt werden. Bei Verwendung des PE-Leader-pCR001-fw Primers, welcher an den oberen Strang downstream des Leaders hybridisiert, kann eine Bande bei ungefähr 100 nt beobachtet werden (Spur 4). Die Länge dieser Bande ist durch eine Duplikation der ersten, Leader-nahen Repeatsequenz und des Einbaus eines neuen Spacers zu erklären. Im Gegensatz dazu kann eine Bande von ungefähr 87 nt im Gegenstrang bei Nutzung des Primers pCR001-Blot-C-S1-rv, der innerhalb der vorhandenen Spacersequenz bindet, detektiert werden (Spur 8). Ein Vergleich mit der Sequenz zeigt, dass die Länge dieser Bande aus der Erweiterung des synthetischen CRISPR-Arrays durch Addition eines neuen Spacers resultieren kann.

Zusammenfassend kann festgehalten werden, dass das Ergebnis der gezeigten Analysen der pCR003 WT Plasmide mit und ohne Induktion der cas1/cas2-Genexpression mit den zuvor in vivo detektierten Southern Blot Signalen übereinstimmt. Eine exakte Sequenzierung der beobachteten Produktbanden konnte jedoch aufgrund zu geringer Konzentrationsausbeuten nicht erfolgen. Selbst in Abbildung 2.12 (B), in der für die Erzielung höherer Produktausbeuten die fünffache Menge der Plasmide in die Hybridisierung eingesetzt wurden, ist eine starke Hintergrundpräsenz der parentalen, ungeschnittenen Plasmide zu erkennen. Der Versuch eine Reduzierung der parentalen Plasmidbanden der isolierten Plasmide durch einen vorherigen Denaturierungsschritt mit NAOH vor der Hybridisierungsreaktion zu erzielen, brachte leider nicht den gewünschten Effekt.

In einem weiteren Teil dieser Arbeit sollte untersucht werden, ob das entwickelte CRISPR-basierte Plasmid-System in *E. coli*, neben diesem Bakterium auch mit Integrationselementen aus anderen Organismen wie *Streptococcus agalactiae* funktionell ist. Dafür galt es als erstes, die für die Spacerintegration essentiellen Proteine dieses Bakterienstammes als Grundlage für den Aufbau eines heterologen Systems zu nutzen. Die Vorgehensweise für den Aufbau eines solchen Systems, sowie erste *in vivo* Analysen, sind im nachfolgenden Kapitel erläutert.

### 2.4 Aufbau eines heterologen Integrationssystems in E. coli

Die erfolgreiche Etablierung des Plasmid-basierten Integrationssystems in *E. coli* führte zur Idee mit Hilfe des entwickelten Systems neue Erkenntnisse in Bezug auf Mechanismen der CRISPR-Adaptation anderer CRISPR/Cas-Typen aufzudecken.

Das grampositive Bakterium *Streptococcus agalactiae* ist ein humanpathogenes Bakterium, das Infektionskrankheiten auslöst und als Typ II CRISPR/Cas-Systemen klassifiziert wurde. Diese Art von CRISPR-Klasse besitzt zusätzlich neben den zwei universell vertretenden Cas1 und Cas2 Proteinen, weitere Cas Proteine (Cas4 oder Csn2), die für die Aufnahme neuer Spacersequenzen benötigt werden (Abbildung 2.13).



**Abbildung 2.13: Schematische Anordnung eines CRISPR-Locus in Typ II-Systemen.** Zu sehen ist eine schematische Darstellung eines CRISPR-Locus von *S. agalactiae*, bestehend aus einem CRISPR-Array der die unterschiedlichen Spacersequenzen (S, farbige Quadrate), flankiert von identischen Repeatsequenzen (R, graue Rauten) beinhaltet. Dem CRISPR-Array vorangestellt ist eine Leader-Region (blau), an der die Transkription startet. In rot nachgeschaltet sind die CRISPR-assoziierten *cas* Gene gezeigt, die für die Cas-Proteine codieren und eine wichtige Rolle in der CRISPR-Aktivität spielen.

Um den Mechanismus der Spacerintegration in diesem CRISPR-Subtyp besser verstehen zu können, wurde in unterschiedlichen Experimenten die Aktivität der Cas1 und Cas2 Proteine von *E. coli* in Kombination mit der des Csn2 Proteins von *Streptococcus agalactiae* in einem *E. coli* BL21 AI Hintergrund getestet. Diese Analysen gingen der Fragestellung voraus, ob nach Einführung von Typ II spezifischen Proteinen (Cas1, Cas2 und Csn2) in das bereits etablierte Plasmid-System die Aufnahme von neuen Spacern erfolgreich nachgewiesen werden kann. Dafür musste parallel zum pCR003 WT Plasmid ein neuer Vektor konstruiert werden, der neben den für die Spacerintegration benötigten Subtyp II-A spezifischen Cas-Proteinen auch einen CRISPR-Array von *S. agalactiae* besitzt.

### 2.4.1 Klonierung eines Plasmid-basierten Vektors für eine heterologe Spacerintegration

Der Aufbau eines Plasmid-Systems mit Subtyp II-A spezifischen Elementen erfolgte in Anlehnung an die in Kapitel 2.1.1 beschriebene Klonierung des pCR003 WT Plasmids. Dabei dienten für die geplanten *in vivo* Experimente zwei bereits konstruierte Vektoren (pSagCas1Cas2Csn2 und pSagCas1Cas2K132) als Grundbausteine. Der Unterschied zwischen beiden Vektoren liegt in einer eingefügten Mutation innerhalb des *csn2* Gens, wodurch die Funktion des Gens gestört ist. Die Amplifizierung des Subtyp II-A spezifischen CRISPR Arrays erfolgte aus genomischer DNA von *S agalactiae* mittels der Primer VK8035-fw und VK8035-rv. Nach Hydrolyse der Vektoren mit *Ecl136II* wurde der isolierte CRISPR-Array in die zuvor 'blunt' linearisierten Plasmide ligiert und der korrekte Einbau durch Sequenzierung bestätigt. Die nachfolgend als pCR004 und pCR005 bezeichneten Plasmide wurden anschließend in kompetente BL21 AI/pUC18-Kan Zellen transformiert und die Fähigkeit zur Aufnahme von Fremd-DNA als neue Spacer in den CRISPR-Array von *S. agalactiae* anhand von Integrationsstudien untersucht (2.1.2). Ein exemplarisches Ergebnis dieser Analysen ist in Abbildung 2.14 dargestellt.



## Abbildung 2.14: PCR-Analyse der Spacerintegration der Plasmide pCR004 und pCR005 im Vergleich zum pCR003 WT Plasmid.

Gezeigt sind die PCR-Produkte von drei Zyklen der Plasmide pCR004, pCR005 und des pCR003 WT Plasmids auf einem 1,2%igen Agarosegel getrennt. Nach 18-24 stündigem Wachstum der Zellen wurden Verdünnungen der Kulturen (1:50 und 1:500 in YT-Medium) als *template* in eine PCR-Reaktion eingesetzt. Der Einbau von neuen Spacersequenzen wurde mit verschiedenen Oligonukleotiden (4.4.1) sowohl im Genom, als auch im Plasmid kontrolliert. Die unterschiedliche Höhe der einzelnen Banden ist durch verschiedene Primer zu erklären. Als Positivkontrolle diente das pCR003 WT Plasmid (links). Parentale PCR-Fragmente sind mit "parental" gekennzeichnet. Darüber liegende Banden sind als neu erworbene Spacersequenzen mit "extended" gekennzeichnet (1x extended).

Eine Aufnahme von neuen Spacern in das Genom von BL21 AI und dem CRISPR-Array der pCR004 bzw. pCR005 Plasmide kann nicht bestätigt werden. Das pCR003 WT Plasmid hingegen zeigt sowohl genomischen als auch plasmidären Einbau, was eine fehlerhafte Vorgehensweise in Bezug auf die Durchführung der Integrationsstudien ausschließt. Ähnlich zum pCR003 WT Plasmid wurden für die PCR-Kontrollreaktionen Oligonukleotide, die zu der Leaderregion bzw. Repeat-Einheit des CRISPR-Arrays von *S. agalactiae* komplementär sind eingesetzt. Um die Reproduzierbarkeit des gezeigten Ergebnisses verifizieren zu können, wurden weitere Kontrollzyklen über mehrere Wochen hinweg analysiert. Doch auch diese zeigten keine Aufnahme in den heterologen CRISPR-Array.

Es stellte sich die Frage, ob diese Reduzierung durch eine inaktive Genexpression der *cas1-cas2* Gene von *S. agalactiae* in *E. coli* BL21 AI Zellen verursacht wird. Zur Klärung wurde dafür im weiteren Verlauf die Expression der *cas1-cas2* Gene nach Induktion mit IPTG und L-Arabinose getestet.

### 2.4.2 Analytische Proteinexpression von *S. agalactiae* Cas1 und Cas2 in BL21 AI

Eine Kontrolle der Genexpression sollte ausschließen, dass die auf den pCR004 und pCR005 Plasmid kodierten *cas1* und *cas2* Gene nach Transformation in einen *E. coli* Stamm in ihrer Expression beeinträchtigt waren. Eine inaktive Expression könnte die Reduzierung der Spaceraufnahme in 2.4.1 erklären.

Für die analytische *in vivo* Expression der *cas*-Gene wurden Zellen des Stammes BL21 AI transformiert mit den Plasmiden pCR004 und pCR005 nach Erreichen der exponentiellen Phase mit IPTG und L-Arabinose induziert. Zur Kontrolle wurde das pCR003 WT Plasmid transformiert in BL21 AI mitgezogen. Nach mehrstündigem Wachstum der Zellen wurden diese geerntet, aufgeschlossen und die Proteine extrahiert (5.3.1). Die Trennung der Proteine erfolgte anschließend über Nacht (üN) auf einem denaturierenden SDS-Gel (5.3.5.1). Das Ergebnis der denaturierenden Trennung ist in Abbildung 2.15 dargestellt.



**Abbildung 2.15: Nachweis der** *cas1-cas2* **Genexpression von pCR004/-005 in BL21 AI.** Gezeigt ist die Proteinauftrennung der BL21 AI/pCR003 bzw. –pCR004 und –pCR005 Zellen nach Induktion der Cas1/Cas2 Proteine auf einem 12% SDS-Gel. Die SDS-Gelelektrophorese erfolgte über Nacht bei 60 Volt. Als Positivkontrolle ist das pCR003 WT Plasmid in Spur 1 zu sehen. Ein Molekulargewichtsmarker wurde als Größenstandard mit aufgetragen (M, Spur 4). Das exprimierte 33 kDa große Cas1 Protein ist durch einen Pfeil gekennzeichnet.

Die analytische Expression der Plasmid-codierten Cas1/Cas2 Proteine von *S. agalactiae* zeigt nach Induktion der Zellen keine Cas1 spezifischen Proteinbanden. Das pCR003 WT Plasmid, welches als Kontrolle mitgezogen wurde zeigt hingegen eine stark exprimierte Cas1 Bande in Höhe von 35 kDa (Spur 1). Eine Expression von Cas2 kann nicht nachgewiesen werden, da dass Protein mit einer Größe von ca. 10 kDa recht klein ist und bereits aus dem SDS-Gel heraus gelaufen ist.

Das Ergebnis der Proteinexpression passt zu den vorherigen der Spacerintegrationsstudien, in denen keine Aufnahme von neuen Spacern beobachtet werden konnte (Abbildung 2.14 Spuren 6-14). Ein erneuter Blick in die Sequenzierungsergebnisse zeigt, dass die Proteine auf den klonierten pCR004 und pCR005 Plasmiden nicht wie beim pCR003 WT Plasmid in *E. coli* überlappend-codiert vorlagen. In anfänglichen Integrationsexperimenten des genomischen CRISPR-Arrays aus BL21 AI konnte ein Verlust des Einbaus von neuen Spacern vermerkt werden, wenn die beiden *cas1* und *cas2* Gene nicht hintereinander auf dem pCR001 Plasmid codiert vorlagen (mündliche Aussage Dr. Ümit Pul). Als nächstes galt es daher einen neuen Vektor mit einer aus *E. coli* charakteristischen Überlappung zwischen den Cas1/Cas2 Proteinen zu klonieren, der als Grundlage für weitere Experimente diente.

Für die Konstruktion dieses neuen Vektors wurden die für *S. agalactiae* spezifischen Cas1 und Cas2 Proteine aus dem pCR005 Plasmid mittels PCR und geeigneten

Primern (4.4.1) isoliert und in ein pACYC-Duet1 Derivat kloniert. Der Einbau der *E. coli* spezifischen Basen zwischen den *cas1-cas2* Genen von *S. agalactiae* erfolgte anschließend mit Hilfe von phosphorylierten Primern, die in einer erneuten PCR-Reaktion das als pACYC-SagCas1Cas2 bezeichnete Konstrukt linearisierten und gleichzeitig mit 8 nt langen Überhängen aus *E. coli* versahen. Um die methylierte parentale DNA von der durch PCR amplifizierten linearisierten DNA zu trennen, folgte eine Hydrolyse mit dem Methylierungs-abhängigen Restriktionsenzym *Dpnl.* Nach Rezirkularisierung des sp veränderten Vektors und Transformation in XL-1 wurde die Insertion der acht Basen durch Sequenzierung bestätigt (siehe Anhang Abb. 6.7). Im Anschluss daran wurde erneut die Expression der *cas1-cas2* Gene nach Transformation in BL21 AI getestet. Das SDS-Gel dieser Genexpression ist in Abbildung 2.16 gezeigt.



## Abbildung 2.16: Analyse der *cas1* Genexpression nach Insertion von 8 *E. coli*-spezifischen Basen.

Zu sehen sind die aufgetrennten Proteine verschiedener Plasmide nach Induktion der *cas1-cas2* Gene auf einem 10% SDS-Gel. Die SDS-Gelelektrophorese erfolgte über Nacht bei 60 Volt. Die Spuren 2-4 enthalten das neue Konstrukt mit 8 eingefügten Nukleotiden aus *E. coli.* Als Positivkontrolle ist das pCR003 WT Plasmid in Spur 5 gezeigt. Die Spuren 6, 7 und 8 zeigen den leeren pACYC-Duet1 Vektor transformiert in BL21 AI, sowie die zuvor klonierten Plasmide pACYC-SagCas1Cas2 und pCR004. Als Größenstandard wurde ein Molekulargewichtsmarker in Spur 1 aufgetragen (M). Das exprimierte 33 kDa große Cas1 Protein ist durch einen Pfeil markiert.

Die Abbildung 2.16 der analytischen Expression der Cas1/Cas2 Proteine von *S. agalactiae* nach Klonierung in einen pACYC-Duet1 Hintergrund und Insertion von 8 Nukleotiden aus *E. coli* lässt nach Induktion der Zellen keine Cas1 Expression erkennen. Das in Spur 5 exprimierte Cas1 Protein des pCR003 WT Plasmids kann deutlich von den anderen Proteinbanden im Gel unterschieden werden und lässt eine gestörte Expression der *cas1-cas2* Gene von *S. agalactiae* in *E. coli* BL21 AI vermuten. Es scheint, als wären die für den Spacereinbau benötigten Proteine in *E. coli* nicht vorhanden, was wiederum zu den Spacerintegrationsstudien in Abbildung 2.14 passt und gleichzeitig auf zusätzliche Protein-Komponenten für ein funktionelles heterologes Integrationssystem hindeutet.

In nachfolgenden Experimenten wurde daraufhin versucht durch Klonierung verschiedener Plasmidvarianten, die zum einen für *E. coli* charakteristische Cas-Proteine enthielten und zum anderen beispielsweise den CRISPR-Array von *S. agalactiae* vorangestellt hatten, die Gründe für das inaktiven heterologen Plasmid-System aufdecken zu können. Eine schematische Darstellung verschiedener Klonierungsstrategien mit konstruierten Plasmiden ist im Anhang in Abbildung 6.8 dargestellt.

Die durchgeführten PCR-Reaktionen zur Kontrolle des Spacereinbaus der im Anhang aufgezeigten Plasmide zeigten auch nach weiteren Zykluskontrollen über mehrere Tage hinweg mit unterschiedlichen Plasmid-Varianten in *E. coli* BL21 AI Zellen keine Integration von neuen Spacern. Eine Aufklärung der fehlenden Integrationselemente für die Funktion eines aktiven heterologen Systems in E. coli konnte somit auch nicht mit Hilfe von unterschiedlichen Klonierungsstrategien erreicht werden. Aus den gezeigten Ergebnissen wird ersichtlich, dass für den Aufbau eines funktionellen heterologen Systems in Typ I-E CRISPR/Cas-Systemen wie E. coli und Integrations-Komponenten des pathogenen Bakterienstammes S. agalactiae zusätzliche Proteine neben Cas1, Cas2 und Csn2 für die Immunisierung fehlen müssen. Ein erster Hinweis welche zusätzlichen Komponenten neben denen für die Adaptation schon vorhandenen Proteinen benötigt werden könnten, kam mit der Aufdeckung des als "primed adaptation" bezeichneten Prozesses auf. Bei dieser Art von Adaptation wird zusätzlich die Anwesenheit des als Cascade bezeichneten Interferenz-Komplexes und des Cas3 Proteins im Falle von E. coli benötigt und die Aufnahme neuer Spacer, die bereits als integrierte Spacer Bestandteil des CRISPR-Arrays sind, begünstigt (Datsenko, Pougach et al. 2012, Swarts, Mosterd et al. 2012). Durch Bereitstellung des für Typ II-A Systeme charakteristischen Cas9 Proteins und der tracrRNA könnte eine Aktivierung des heterologen Integrationssystems in E. coli eventuell begünstigt werden und ist nachfolgend Bestandteil der Diskussion.

## 2.5 Expression und Isolation von Komponenten des *E. coli* Cascade-Komplexes

Im zweiten Teil dieser Arbeit war eine präparative Isolierung und Aufreinigung der aus dem Cascade-Komplex bestehenden fünf Cas-Proteine A, B, C, D und E (Cse1, Cse2, Cas7, Cas5 und Cas6e) für *in vitro* Analysen vorgesehen. Die präparative Isolierung sollte nach Klonierung der einzelnen Proteine in einen geeigneten Expressionsvektor durch Affinitätschromatographie über verschiedene Säulen durchgeführt werden. Für die *in vitro* Analysen sind sowohl Bindestudien zwischen den aufgereinigten Cas-Proteinen und einzel- sowie doppelsträngigen DNA-, als auch RNA-Fragmenten beabsichtigt.

## 2.5.1 Klonierung und analytische Überexpression der fünf Cas-Proteine

Die Isolation der Cascade-Protein-Gene aus *E. coli* erfolgte mittels PCR und geeigneten Primern (4.4.1) aus dem Plasmid pWUR400 (Brouns, Jore et al. 2008). Für die präparative Aufreinigung nach Expression mittels Affinitätschromatographie wurden die amplifizierten Protein-Gene anschließend in den zuvor linearisierten Expressionsvektor pET-52(b)+ ligiert und in kompetente BL21 (DE3) transformiert. Dieser Vektor ist neben einem starken induzierbaren T7-Promotor, *upstream* der *'MCS'* (*multiple cloning site*) mit einem N-terminalen StrepII-Tag ausgestattet. So wird das Genprodukt des an dieser Stelle einklonierten Genes bei der Translation an den StrepII-Tag fusioniert und kann in einem weiteren Schritt aufgereinigt werden. Die Vektorkarten der klonierten Expressionsplasmide sind im Anhang unter Abbildung 6.9 zu finden.

Die Überexpression und Isolierung der Cascade-Proteine erfolgte nach erfolgreicher Klonierung in die '*MCS*' und Fusionierung des N-terminalen StrepII-Tags des Expressionsvektors unter der in 5.3 beschriebenen Methode. Es wurden 100 ml YT-Medien (4.6.2) versetzt mit 50 µg/ml Ampicillin mit je einem ml der üN-Kultur der Stämme BL21 (DE3)/pET-CasA, -CasB, -CasC, -CasD und –CasE angeimpft. Nach Erreichen einer OD<sub>600</sub> von 0,5-0,6 wurde die Induktion der *cas*-Genexpression durch die Zugabe von 0,1-0,2 mM IPTG ausgelöst. Je nachdem welches Cas-Protein angezogen wurde, erfolgte ein Wechsel der Zellkulturen von 37°C auf Raumtemperatur (20-25°C). Des Weiteren wurde unmittelbar vor der Induktion mit IPTG, sowie nach Induktion mehrere Stundenwerte (1, 3, 5 Stunden und über Nacht (üN)) Proben von den Zellkulturen abgenommen und die Überexpression der Cas-Proteine analytisch überprüft (5.3.1). Ein Ergebnis dieser Proteinauftrennungen ist exemplarisch in Abbildung 2.17 dokumentiert.



Abbildung 2.17: Analytische Überprüfung der Cascade-Protein Induktion.

Vor der Induktion mit IPTG (v) 1, 3 und 5 Stunden danach bzw. über Nacht (üN) wurden Kontrollproben von den Zellkulturen entnommen und mittels Ultraschall aufgeschlossen. Auf den dargestellten 12% SDS-Gelen wurden zwei Molekulargewichtsmarker (M, Spuren 1 und 19) als Größenstandards aufgetragen. Die überexprimierten Cas-Proteine sind mit einem Stern (rot) gekennzeichnet und die Größen der Proteine für eine Zuordnung neben dem SDS-Gel dargestellt. Eine Expression der jeweiligen Cas-Proteine nach Induktion mit IPTG ist bei allen Proteinen bis auf CasD (Cas5, Spuren 5-7) zu erkennen.

Das analytische SDS-Gel zur Überprüfung der Cascade-Genexpression zeigt bei vier von fünf Proteinen eine deutlich exprimierte Proteinbande. Eine Überexpression von CasD (Cas5) ist nach Induktion mit IPTG und Wachstum der Zellen für 3 Stunden und üN nicht zu erkennen. Wie die Expression der Cascade-Gene in *E. coli* BL21 (DE3) reguliert ist und welche Faktoren daran beteiligt sind, konnte bereits in Experimenten mit dem für die fünf cascade-Gene exprimierenden Vektor pWUR400 beobachtet werden (Spreiz 2011). Nach Transformation des Plasmides pWUR400 in wildtypische und  $\Delta hns$ -Zellen und Induktion mit IPTG wurde eine ungleichmäßige Expression der einzelnen Cascade-Proteine CasA, CasB, CasC, CasD und CasE beobachtet. Darüber hinaus konnte eine deutlich höhere Expression der Cas-Proteine im  $\Delta hns$ -Stamm als im Wildtypstamm BL21 (DE3) aufgezeigt werden. Diese unterschiedlich starke Expression der einzelnen Proteine kann auf die ungleiche, stöchiometrische der Cas-Protein-Untereinheiten Verteilung innerhalb des Cascade-Komplexes zurückgeführt werden (siehe Kapitel 1.3.1) und stimmt mit dem gezeigten Expressionsverhalten der einzelnen Proteine überein. Mit nur einer Untereinheit im Cascade-Komplex werden die Cas-Proteine CasA, CasD und CasE geringer exprimiert als CasB und CasC mit zwei bzw. sechs Untereinheiten (Spreiz 2011).

Für den weiteren Verlauf der präparativen Isolierung der einzelnen Cascade-Proteine wurde deshalb neben den angezogenen Kulturen der Stämme BL21 (DE3)/pET-CasA, -CasB, -CasC, -CasD und -CasE auch Zellen des Stammes BL21 (DE3)/pWUR400/pET-CasC angezogen. Es sollte untersucht werden, ob die Anwesenheit der *cascade* Gene als stabilisierende Effektoren Einfluss auf die Expression und Isolierung des mit StrepII-Tag markierten CasC-Proteins nehmen. Dafür wurde nach erfolgreicher Klonierung des CasC-Proteins in den Vektor pET-52(b)+ die Expression des Proteins in Gegenwart des Cascade-Komplexes untersucht. Nach Transformation in kompetente BL21 (DE3)/pWUR400 Zellen und Induktion der Kulturen mittels IPTG wurde die Expression der Cas-Proteine analytisch auf einem Proteingel überprüft (5.3.1). Das Ergebnis dieser Analyse ist im Anhang in Abbildung 6.10 dargestellt und lässt im Vergleich zu den Zellen ohne den Cascade-Komplex eine abgeschwächte Proteinexpression erkennen. Als nächstes galt es den Cascade-Komplex durch das am N-Terminus mit StrepII-Tag markierte CasC-Protein über Affinitätschromatographie zu isolieren.

Die Anzucht und Ernte der einzelnen Expressionsstämme nach Transformation der Plasmide erfolgte unter verschiedenen Wachstumstemperaturen und -zeiten. Nach Ernte der Zellen wurde versucht, die Proteine über Affinitätschromatographie mittels einer Glassäule oder der Äkta Prime Plus Anlage zu isolieren. Dabei konnte eine Aufreinigung des CasD Proteins im Gegensatz zu den übrigen Cas-Proteinen auch nach mehreren Anläufen und veränderten Versuchsbedingungen bei der Anzucht der Zellen (Temperatur, IPTG-Konzentrationen und Stundenwerte) oder den Pufferzusammensetzungen während der Aufreinigung nicht erzielt werden. Eine Expression von CasD konnte bereits zuvor schon analytisch kaum beobachtet werden (Abb. 2.17). Das CasE Protein hingegen war durch die Ausbildung von 'inclusion bodies' während der Aufreinigung schwer löslich und konnte in nur sehr geringen Konzentrationen isoliert werden. Die Schwierigkeiten, die während der Isolation dieser zwei Proteine auftraten sind Bestandteile der Diskussion und werden dort näher erläutert. Im nachfolgenden Kapitel sind daher nur die Aufreinigungen der drei Cas-Proteine CasA (Cse1), CasB (Cse2) und CasC (Cas7) mit und ohne den Cascade-Komplex näher beschrieben.

### 2.5.2 Aufreinigung von CasC (Cas7)/pWUR400 mit fusioniertem StrepII-Tag

Die Aufreinigung des Cascade-Komplexes erfolgte nach Transformation von pET-CasC in BL21 (DE3)/pWUR400 Zellen unter der 5.3.2 beschriebenen Methode. Dafür wurden 800 ml YT-Medium (4.6.2) 1/100 aus einer vorherigen üN-Kultur des Stammes BL21 (DE3)/pWUR400/pET-CasC angeimpft und nach Erreichen einer OD<sub>600</sub> von 0,5-0,6 die Induktion durch Zugabe von 0,1 mM IPTG ausgelöst. Im Anschluss an die Induktion wurde die Wachstumstemperatur der Kulturkolben von 37°C auf Raumtemperatur (20-25°C) reduziert und die Zellen nach 16-18 Stunden am nächsten Morgen geerntet. Dabei konnte 4,5 g Zellmaterial gewonnen werden, dass in 50 ml Aufschlusspuffer, versetzt mit 0,2 mM PMSF aufgenommen und mit Ultraschall aufgeschlossen wurde. In einem Zentrifugationsschritt (30 Min. 14000 rpm 4°C) wurden die Zellbestandteile vom Protein entfernt und der proteinhaltige Überstand durch Affinitätschromatographie über eine mit Strep-Tactin Resin gefüllte Glassäule von anderen Proteinen abgetrennt. Dieses Trägermaterial weist eine starke Bindungsaffinität an mit StrepII-Tag markierte Proteine auf und fixiert sie an der Säule. Für die affinitätschromatographische Aufreinigung musste die gepackte Säule zuerst mit Waschpuffer (5.3.3.1) äquilibriert werden. Im Anschluss wurde der proteinhaltige CasC/Cascade-Überstand auf die Säule gepumpt und der Durchfluss aufgefangen. Es folgte ein Waschschritt bis die anhand des Schreiberprofils zu verfolgende Absorption der Probe wieder die Basislinie erreicht hatte. Um zunächst Fremdproteine von der Säule zu waschen, wurde dem Waschpuffer eine geringe Konzentration an Desthiobiotin (0,25 mM) hinzugegeben. Für die langsame Elution des StrepII-tag markierten CasC/Cascade-Proteins wurde ein Stufengradient von 0,5 bis 2,5 mM Desthiobiotin eingesetzt, wobei das Protein mit 2,5 mM Desthiobiotin von 2.18 der Säule eluiert wurde. In Abbildung ist das Ergebnis der Affinitätschromatographie von pET-CasC/Cascade nach Trennung auf einem 10% SDS-Gel und Färbung mit Coomassie-Blue (5.3.6.1) gezeigt.



## Abbildung 2.18: SDS-Gel der Säulenchromatographie des CasC/Cascade-Proteins mittels Strep-Tactin Resin.

Zu sehen ist ein 10 %iges SDS-Gel (Coomassie gefärbt) auf dem 20  $\mu$ l der Fraktionen 1 bis 20 aufgetragen wurden. Zur Orientierung wurde ein Molekulargewichtsmarker M in Spur 23 mit aufgetragen. Die Spuren 1-4 zeigen Proben nach dem Aufschluss (A), nach dem Zentrifugieren (Ü), vom Pellet (P) und vom Durchbruch (D). Das CasC/Cascade-Protein wurde mit einem Desthiobiotingradient von 0,25-2,5 mM von der Säule eluiert. Die einzelnen Cas-Proteinbanden sind markiert. Links dargestellt sind vergrößerte Ausschnitte der Proteinbanden von CasA, CasB, CasC und CasE.

Die Abbildung 2.18 zeigt die Isolierung des mit StrepII-Tag markierten CasC-Proteins in Gegenwart des Cascade-Komplexes. Neben dem isolierten CasC-Protein sind die einzelnen Proteine des Cascade-Komplexes gut zu erkennen. Die Spuren 1 und 2 nach Aufschluss und Zentrifugation zeigen deutliche Mengen an exprimierten Cas-Proteinen. Anhand Spur 3 ist erkennbar, dass einiges an Protein nach dem Aufschluss im Pellet verbleibt. Die Isolation des Cascade-Komplexes in Gegenwart des markierten CasC-Proteins kann durch die Zugabe von Desthiobiotin erreicht werden. Bereits ab Fraktion 6 wurde dem Waschpuffer eine geringe Menge an Desthiobiotin (0,25 mM) hinzugegeben, was zu einer leichten Elution der Cas-Proteine von der Säule führte (Spuren 7 und 8). Um die Elution zu verstärken, erfolgte daher ein Wechsel zum Elutionspuffer mit einer höheren Desthiobiotinkonzentration von 2,5 mM. Dabei konnte das an der Säule fixierte CasC-Protein, sowie der Cascade-Komplex bestehend aus den fünf Proteinen CasA, CasB, CasC, CasD und CasE isoliert werden. Die auf dem Proteingel detektierten Doppelbanden des CasC-Proteins sind durch den fusionierten StrepII-Tag zu erklären. Wobei die untere Proteinbande das CasC-Protein des Cascade-Komplex darstellt, weist die obere (CasC\*) nach der N-terminalen Fusionierung mit dem StrepII-Tag ein höheres Molekulargewicht auf und wird langsamer aufgetrennt. Eine Expression des CasD-Proteins, welche zuvor in analytischen Experimenten nicht detektierbar war, kann bei der Aufreinigung des Cascade-Komplexes gezeigt werden und weist auf eine Stabilisierung durch den Komplex hin.

Nach erfolgreicher Isolierung des am CasC-Protein gekoppelten Cascade-Komplexes galt es in einem nächsten Schritt das N-terminal markierte CasC-Protein in Abwesenheit des für die Cas-Protein codierenden pWUR400 Plasmids zu isolieren.

### 2.5.3 Aufreinigung von CasC (Cas7) mit fusioniertem StrepII-Tag

Nach Klonierung des durch PCR-amplifizierten *casC* Gens in den pET-52(b)+ Expressionsvektor und erfolgreicher analytischer Überexpression (Abb. 2.17) folgte die präparative Aufreinigung des Proteins nach der unter 5.3.2 beschriebenen Methode. Dafür wurde 800 ml YT-Medium (4.6.2) mit 8 ml aus einer vorherigen üN-Kultur des Stammes BL21 (DE3)/pET-CasC angeimpft und nach Erreichen einer OD<sub>600</sub> von 0,5-0,6 durch Zugabe von 0,1 mM IPTG induziert. Die weitere Inkubation der Zellen nach Induktion erfolgte bei Raumtemperatur (RT) über Nacht. Nach Ernte der Zellen am nächsten Morgen konnte 5,2 g Zellmaterial gewonnen werden, dass in einem nächsten Schritt in 100 ml Aufschlusspuffer, versetzt mit 0,2 mM PMSF aufgenommen wurde. Es folgte ein Ultraschallaufschluss der Zellen, wobei die Zelltrümmer in einem anschließenden Zentrifugationsschritt entfernt wurden. Nach einem weiteren Zentrifugationsschritt in der Ultrazentrifuge (20 Min. 15000 rpm 4°C) wurde die Hälfte des proteinhaltigen Überstandes von CasC auf eine zuvor äquilibrierte StrepTrap<sup>Tm</sup>-Säule (GE Healthcare) gegeben. Bei dieser immobilisierten Affinitätschromatographie mit Hilfe der Äkta Prime Plus Anlage werden die im Aufschlusspuffer befindlichen Proteine über ein Schlauchsystem in die zuvor mit Waschpuffer A (5.3.3.2) äquilibrierte Säule aufgezogen. Erreicht die Probe das Säulenmaterial erfolgt der Durchbruch von Proteinen, wobei das StrepII-Tag markierte Protein an der Säulenmatrix fixiert wird. Anhand eines Absorptionsprofils kann die affinitätschromatographische Aufreinigung verfolgt werden. Zeigt das Profil einen Abfall des Durchbruchs an, wird die Säule mit Waschpuffer A und B (5.3.3.2) gewaschen, um die Basislinie wieder einzustellen. Die Elution des an der Säule gebundenen Proteins wird anschließend in einem Schritt durch eine sogenannte *'step elution'* (Step-Elution) mit einer in Puffer C enthaltenen Desthiobiotinkonzentration von 2,5 mM erreicht. Das Elutionsprofil einer CasC-Aufreinigung ist exemplarisch in Abbildung 2.19 gezeigt.



**Abbildung 2.19: Elutionsprofil der Affinitätschromatographie von BL21 (DE3)/pET-CasC.** Der Graph entspricht der gemessenen Absorption der proteinhaltigen Lösung bei 280 nm und zeigt die einzelnen Phasen der Aufreinigung. Auf der X-Achse ist die Zeit in Minuten, sowie die Nummer der Fraktion (rot) dargestellt. Auf der Y-Achse sind verschiedene Messwerte wie die UV-Absorption in mAu (blau), der Druck MPa (braun), die Temperatur in °C (türkis) und die Konzentration in % des Elutionspuffers C (grün) gezeigt. Das gebundene CasC-Protein wurde mittels 350 und 75 mM NaCl der Waschpuffer A und B gewaschen und anschließend in einem Schritt mit Elutionspuffer C mit 2,5 mM Desthiobiotin eluiert. Der rote Pfeil markiert den Elutionspeak von CasC.

Anhand des Profils der Proteinaufreinigung kann der Durchbruch des CasC-Proteins bereits nach 30 min. beobachtet werden. In zwei darauffolgenden Waschschritten mit Puffer A und B wurde die Basislinie erneut eingestellt. Nach Zugabe des Puffers C mit einer Desthiobiotinkonzentration von 2,5 mM erfolgt die Elution des an der Säule gebundenen Proteins, was durch einen Peak bei 190 min (roter Pfeil) angezeigt wird. Proben der gesammelten Fraktionen wurden anschließend auf einem 12% SDS-Gel aufgetragen und über Nacht getrennt. Das Ergebnis dieser Trennung ist in Abbildung 2.20 zu sehen.



## Abbildung 2.20: SDS-Gel der Affinitätschromatographie des CasC Proteins über eine StrepTrap<sup>™</sup>-Säule (GE Healthcare).

Zu sehen ist ein 12 %iges SDS-Gel (Coomassie gefärbt) auf dem je 20 µl der gesammelten Fraktionen 6 bis 38 aufgetragen wurden. Zur Orientierung wurde ein Molekulargewichtsmarker M in Spur 1 aufgetragen. Die Spuren 2-5 zeigen Proben nach Aufschluss und Zentrifugation (Ü1), nach einem weiteren Ultrazentrifugationsschritt (Ü2) und von den Pellets (P1 und P2). Die Elution des CasC-Proteins von der StrepTrap<sup>TM</sup>-Säule erfolgte in einem Schritt mit einer Desthiobiotinkonzentration von 2,5 mM. Das eluierte CasC-Protein ist mit einem Pfeil gekennzeichnet.

Das SDS-Gel der präparativen Aufreinigung des CasC-Proteins mit fusioniertem StrepII-Tag zeigt nach Ultraschallaufschluss der Zellen, dass viel Protein im Pellet verblieben ist (Spuren 1 und 2). Nach einem zweiten Zentrifugationsschritt in der Ultrazentrifuge ist verhältnismäßig wenig Protein im Überstand (Ü2) im Vergleich zum Pellet lokalisiert (P2, Spur 5). Nachdem die Säule zunächst mit Puffer A äquilibriert und der proteinhaltige Überstand über das Schlauchsystem aufgezogen wurde, erfolgt der Durchbruch des Proteins. In zwei darauffolgenden Waschschritten mit Puffer A und B (Spuren 8-11) werden zunächst unspezifisch gebundene Proteine von der Säule gewaschen und freie Proteine entfernt. Die Elution des CasC-Proteins wurde in einer step-Elution und einer Desthiobiotinkonzentration von 2,5 mM ab Fraktion 32 erreicht. Dabei kann eine Elution des Proteins von der Säule bis hin zu Fraktion 36 beobachtet werden, die auch anhand des Peaks im Elutionsprofil in Abbildung 2.20 deutlich zu erkennen ist. Im Anschluss an die Elution wurden die Fraktionen 32 bis 36 vereinigt und über Nacht dialysiert (5.3.4.1), um das isolierte CasC-Protein von dem gebundenen Desthiobiotin zu lösen. Eine anschließende Konzentrationsbestimmung mittels Bradford-Reagenz (5.2.1.3) zeigte ausreichende Ausbeuten an aufgereinigtem CasC-Protein. Nach Regeneration der Säule mit *Aqua dest.* und NaOH wurde in einem neuen Lauf die restliche Probe des nach Ultraschallaufschluss und Zentrifugation erhaltenen Überstandes auf die Säule gegeben. Das Ergebnis des SDS-Gels dieser aufgetragenen Fraktionen lässt wie zum vorherigen in Abb. 2.20 eine gute Elution des CasC-Proteins erkennen (Abbildung nicht gezeigt). Für folgende *in vitro* Analysen wurde das aufgereinigte CasC-Protein nach Dialyse aliquotiert und bei -20°C gelagert.

Neben der affinitätschromatographischen Aufreinigung von CasC über eine StrepTrap<sup>Tm</sup>-Säule mit Hilfe der Äkta Prime Plus Anlage der Firma GE Healthcare wurde das Protein auch über das wie in Kapitel 5.3.3.1 beschriebene Glassäulensystem mit Strep-Tactin Resin als Trägermaterial isoliert. Die Anzucht der BL21 (DE3)/pET-CasC Zellen erfolgte dabei gleich wie in 2.5.4 und lieferte im direkten Vergleich höhere Proteinkonzentrationen als zuvor. Ein Ergebnis der fraktionierten Proben auf einem 12% SDS aufgetragen ist im Anhang in Abbildung 6.11 gezeigt.

### 2.5.4 Aufreinigung von CasA (Cse1) mit fusioniertem StrepII-Tag

Für die Aufreinigung des am N-Terminus mit einem StrepII-Tag versehenen CasA-Proteins erfolgte die Anzucht der pET-CasA/BL21 (DE3) Zellen vergleichbar mit der in Kapitel 2.5.2. Es wurden vier 800 ml YT-Medien (4.6.2) nach Induktion mit 0,2 mM IPTG bereits nach 5 Stunden geerntet und dabei 12 g Zellmaterial gewonnen. Das Zellpellet von zwei 800 ml Kulturen (6 g Zellen) wurde in 60 ml Aufschlusspuffer, versetzt mit 0,2 mM PMSF aufgenommen und nach einem ersten Ultraschallaufschluss und Zentrifugationsschritt (30 Min. 14000 rpm 4°C) erneut in 40 ml Aufschlusspuffer aufgenommen und aufgeschlossen. Der proteinhaltige Überstand wurde anschließend auf die mit Strep-Tactin Resin gefüllte Glassäule gegeben und mittels Affinitätschromatographie von anderen Proteinen abgetrennt. Zum Wiedereinstellen der Basislinie wurde die Säule zunächst mit Waschpuffer ohne die Zugabe von Desthiobiotin gewaschen. Für die Ablösung von eventuell schwach bindenden Fremdproteinen diente dann ein Waschschritt mit der Zugabe von 0,25 mM Desthiobiotin. In einem letzten Elutionsschritt wurde das Protein mit Hilfe eines Stufengradienten von 0,5 bis 2,5 mM Desthiobiotin von der Säule eluiert. In Abbildung 2.21 ist das Ergebnis der Affinitätschromatographie von CasA nach Trennung auf einem 12% SDS-Gel und Färbung mit Coomassie-Blue (5.3.6.1) gezeigt.



#### Abbildung 2.21: SDS-Gel der Säulenchromatographie des CasA Proteins mittels Strep-Tactin Resin.

Zu sehen ist ein 12 %iges SDS-Gel (Coomassie gefärbt) auf dem 20-40  $\mu$ l der Fraktionen 1 bis 24 aufgetragen wurden. Die Fraktionen 1 bis 11 enthalten 20  $\mu$ l Probe und die Fraktionen 12 bis 24 40  $\mu$ l. Zur Orientierung wurde ein Molekulargewichtsmarker M in Spur 25 aufgetragen. Die Spuren 1-6 zeigen Proben nach dem Aufschluss (A1, A2), nach dem Zentrifugieren (Ü1, Ü2), vom Durchbruch (D) und Pellet (P). Das Protein wurde mit einem Desthiobiotingradienten von der Strep-Tactin Säule eluiert (Fraktionen 9 bis 16 0,5 mM Desthiobiotin, ab Fraktionen 17 2,5 mM Desthiobiotin). Die CasA-Banden sind mit einem Pfeil markiert.

Es ist erkennbar, dass im Vergleich zu den Spuren 1 und 2 die Proben nach dem zweiten Ultraschallaufschluss weniger CasA-Protein im Überstand als zuvor (Spur 3 und 4) enthalten. Anhand von Spur 6 wird deutlich, dass viel Protein im Pellet verblieben ist, so dass insgesamt nur ein Bruchteil des überexprimierten CasA-Proteins auf die Säule gegeben wurde. Nachdem die Säule zunächst gewaschen wurde (Fraktionen 1 bis 8), wurde dem Puffer ab Fraktion 9 eine geringe Konzentration an Desthiobiotin (0,25 mM) hinzugegeben, um unspezifisch gebundene Proteine von der Säule zu waschen und freie Proteine zu entfernen. Die Elution des CasA-Proteins erfolgte durch einen Desthiobiotingradienten von 0,5-2,5 mM in den Fraktionen 9 bis 24. Es wird deutlich, dass die Elution des Proteins bereits in den Fraktionen 13-17 bei einer Desthiobiotinkonzentration von nur 0,5 mM ihr Maximum hat. Nach Wechsel des Elutionspuffers mit 2,5 mM Desthiobiotin ab Fraktion 17 sind abgeschwächte Proteinbanden zu erkennen (Spuren 19-24), was darauf hindeutet, dass das gebundene CasA-Protein bereits in den vorherigen Fraktionen von der Säule eluiert wurde. Nach Beendigung der Säulenchromatographie wurden die Fraktionen 13 bis 18 vereinigt und über Nacht dialysiert, um das gebundene Desthiobiotin wieder herauszuwaschen (5.3.4.1).
Insgesamt lieferte die Aufreinigung des CasA-Proteins nur eine geringe Ausbeute. Bereits in den analytischen Expressionskontrollen konnte nur eine schwache Expression von CasA nach Zugabe von IPTG beobachtet werden, die auch nach mehreren Versuchen, in denen die Versuchsbedingungen wie IPTG-Konzentration, Temperatur oder die Wachstumszeit der Zellen verändert wurden, nicht erhöht werden. Für nachfolgende *in vitro* Bindestudien wurde das dialysierte CasA-Protein deshalb aufkonzentriert (5.3.4.2) und bis zur weiteren Verwendung bei -20°C gelagert.

### 2.5.5 Aufreinigung von CasB (Cse2) mit fusioniertem StrepII-Tag

Für die Aufreinigung des CasB-Proteins wurde ein 800 ml YT-Kolben (4.6.2) mit 8 ml aus einer vorherigen üN-Kultur des Stammes BL21 (DE3)/pET-CasB angeimpft und nach Erreichen einer OD<sub>600</sub> von 0,5-0,6 die Überexpression des Proteins durch Zugabe von 0,1 mM IPTG ausgelöst. Nach Induktion wurde die Temperatur des Kulturkolben von 37°C auf Raumtemperatur (RT) reduziert und über Nacht schüttelnd weiter inkubiert. Am nächsten Morgen wurden die Zellen geerntet, wobei 4,8 g Zellmaterial gewonnen werden konnte. Das Zellmaterial wurde in 50 ml Aufschlusspuffer, versetzt mit 0,2 mM PMSF aufgenommen und die Zellen mittels Ultraschall aufgeschlossen. Die Zelltrümmer wurden durch Zentrifugation entfernt und die Hälfte des proteinhaltigen Überstandes auf eine mit Strep-Tactin Resin gefüllte Glassäule gegeben. Die Vorgehensweise der Affinitätschromatographie erfolgte ähnlich zu der in 2.5.2 und 2.5.4 und ist im Methodenteil unter 5.3.3.1 näher erläutert. Nach Fraktionierung der Proben wurden 20-40 µl der Fraktionen auf einem 12% SDS-Gel über Nacht getrennt. Die Abbildung 2.22 zeigt exemplarisch ein Ergebnis der Affinitätschromatographie des CasB-Proteins nach Färbung mit Coomassie-Blue (5.3.6.1).



#### Abbildung 2.22: SDS-Gel der Affinitätschromatographie des CasB Proteins mittels Strep-Tactin Resin.

Zu sehen ist ein 12 %iges SDS-Gel (Coomassie gefärbt) auf dem je 20-40  $\mu$ l der gesammelten Fraktionen 1-24 aufgetragen wurden. Zur Orientierung wurde ein Molekulargewichtsmarker M in Spur 29 mit aufgetragen. Die Spuren 1-4 zeigen Proben nach dem Aufschluss (A), nach dem Zentrifugieren (Ü), vom Pellet (P) und vom Durchbruch (D). Die Spuren 5-11 enthalten 20  $\mu$ l, die Spuren 12-24 40  $\mu$ l Probe der aufgetragenen Fraktionen. Die Elution des CasB-Proteins von der Strep-Tactin Säule erfolgte mittels eines Stufengradienten von 0,25-2,5 mM Desthiobiotin (Fraktionen 8 bis 10 mit 0,25 mM Desthiobiotin, ab Fraktion 11 bis 14 mit 0,5 mM Desthiobiotin und ab Fraktion 15 bis 21 mit 2,5 mM Desthiobiotin). Das eluierte CasB-Protein ist mit einem Pfeil gekennzeichnet.

Das SDS-Gel in Abbildung 2.22 zeigt, dass die präparative Aufreinigung des **CasB-Proteins** mit fusioniertem StrepII-Tag gut funktioniert hat. Nach Ultraschallaufschluss lässt sich eine gleichbleibende Konzentration an CasB in Spur 1 und im Überstand nach Zentrifugation in Spur 2 erkennen. Das Pellet in Spur 3 zeigt im Vergleich eine schwächere Proteinbande, was ein Verbleiben des Proteins im Pellet ausschließt. Anhand der Spuren 5-7 wird jedoch deutlich, dass bereits viel Protein beim Waschen von der Säule gelöst wurde. Durch Verwendung eines Desthiobiotingradienten von 0,25-2,5 mM konnte eine stufenweise Elution des Proteins erreicht werden. Wo hingegen in den Fraktionen 4 bis 7 zunächst nur eine schwache Proteinbande erkennbar ist, kann ab Fraktion 8 nach Zugabe von 0,25 mM Desthiobiotin zum Waschpuffer eine Zunahme der Bandenintensität beobachten werden. Die Spuren 14-20 nach Erhöhung der Desthiobiotinkonzentration auf 0,5 bzw. 2,5 mM (ab Fraktion 15) zeigen eine deutliche Elution des CasB-Proteins. Um eine möglichst hohe Konzentration an isoliertes CasB-Protein zu erhalten, wurden die Fraktionen 18 bis 23 im Anschluss an die Aufreinigung vereinigt und über Nacht dialysiert (5.3.4.1). In einer anschließenden Konzentrationsbestimmung mittels

Bradford-Reagenz (5.2.1.3) konnte jedoch ähnlich zu CasA eine nur geringe Ausbeute an isoliertes Protein erzielt werden. Für nachfolgende *in vitro* Analysen wurden deshalb die dialysierten Proben in Amicon-Zentrifugenröhrchen aufkonzentriert (5.3.4.2) und bis zur weiteren Verwendung bei -20 °C gelagert.

# 2.5.6 Ist die assoziierte crRNA im isolierten Cascade-Komplex nachweisbar?

In *E. coli* bilden die zuvor prozessierten, 61 nt kurzen crRNAs mit den fünf Cas-Proteinen Cse1, Cse2, Cas7, Cas5 und Cas6e den für die Immunabwehr wesentlichen Cascade Multiproteinkomplex (Abbildung 1.5), der erst durch Assoziierung der crRNA volle Funktionsfähigkeit erlangt (Jore, Lundgren et al. 2011). Der Kern des Komplexes wird aus sechs Cas7 Untereinheiten gebildet und dient als Rückgrat für die anderen Cas-Protein-Untereinheiten und die assoziierte crRNA. Die strukturelle Anordnung der kleinen crRNA innerhalb des Komplexes bietet auf diese Weise Schutz vor RNA-Degradation.

In weiteren Verlauf dieser Arbeit sollte der in 2.5.2 isolierten CasC/Cascade-Komplex auf das Vorhandensein einer assoziierten crRNA untersucht werden. Für die Isolierung der Nukleinsäure wurden 2 x 4 ml der zuvor dialysierten CasC/Cascade Proben mittels Phenol/Chloroform aufgereinigt (5.2.3.1), mit Ethanol präzipitiert (5.2.3.2) und nach dem lösen in TE-Puffer (4.6.1) auf einem Agarosegel analysiert (5.2.4.1). Das Ergebnis dieser elektrophoretischen Trennung ist in Abbildung 2.23 (A) gezeigt.



#### Abbildung 2.23: Nachweis der assoziierten crRNA.

In (A) ist ein 0,8% Agarosegel mit 1  $\mu$ g der nach Phenol/Chloroform aufgereinigten CasC/Cascade Probe gezeigt (Spur 2). Als Marker wurde in Spur 1 eine 100 bp Leiter mit aufgetragen. (B) zeigt die Proben nach Behandlung mit DNaseI (Spur 2), mit RNAse A (Spur 3) und alleine als Kontrolle (K) in Spur 4 aufgetragen.

Das Agarosegelbild des aufgereinigten, isolierten CasC/Cascade Proteins in Abbildung 2.23 (A) zeigt eine vermutliche RNA-Bande in Spur 2. Um die detektierte Bande als RNA bestätigen zu können, wurde in einem weiteren Versuch die isolierte Nukleinsäure mit DNaseI und RNAse A behandelt (Abb. 2.23 (B)). Das Gelbild in (B) zeigt, dass nach Behandlung mit RNAse A keine Bande mehr nachgewiesen werden kann, was die Existenz einer RNA nach Phenol/Chloroform-Aufreinigung bestätigt. Die im Cascade-Komplex assoziierte crRNA konnte bei der affinitäts-chromatographischen Aufreinigung in 2.5.2 mit isoliert werden.

Eine Assoziierung der aufgereinigten CasA, CasB und CasC Proteine mit Nukleinsäuren konnte nach Behandlung mit Phenol/Chloroform und Analyse auf einem Agarosegel nicht bestätigt werden (Daten nicht gezeigt). Für eine nähere Charakterisierung der isolierten Cas-Proteine wurden deshalb im weiteren Verlauf dieser Arbeit *in vitro* Analysen mit verschiedenen Nukleinsäuren als Interaktionspartner durchgeführt.

### 2.5.7 In vitro Retardierungsanalysen

Im Folgenden war es das Ziel, mittels Gelshift-Experimenten (5.4.4) Bindestudien mit den isolierten CasA (Cse1), CasB (Cse2) und CasC (Cas7) Proteinen aus *E. coli* und verschiedenen DNA- bzw. RNA-Fragmenten durchzuführen. Mit Hilfe dieser Gelverzögerungsanalysen kann eine Verringerung der Laufgeschwindigkeit eines Proteinkomplexes unter nicht-denaturierenden Bedingungen gegenüber der freien Nukleinsäure nachgewiesen werden. Ursache dieser Mobilitätsverzögerung ist das erhöhte Molekulargewicht des Komplexes gegenüber der freien DNA.

# 2.5.7.1 Bindeanlaysen von CasA, CasB und CasC an doppelsträngige DNA-Fragmente

Um eine potentielle Bindung der isolierten *E. coli* Cas-Proteine A, B und C an doppelsträngige DNA zu testen, wurde das Leader BL21 Fragment aus *E. coli* (4.3.1) verwendet. Hierfür wurde das Fragment zunächst über Restriktionshydrolyse mit den Enzymen *EcoRI* und *HincII* geschnitten, um 5'-überhängende Enden zu erhalten. In einer anschließenden Auffüllreaktion wurden die entstandenen 5'-überhängenden Enden des Leader BL21 Fragmentes mit Hilfe des Klenow-Enzyms der DNA Polymerase I durch den Einbau von [ $\alpha^{32}$ P]dNTPs radioaktiv markiert (5.2.6.9) und für Bindeanalysen mit den jeweiligen Cas-Proteinen genutzt.

In Abbildung 2.24 sind die Retardierungsanalysen der isolierten Cas-Proteine A, B und C mit dem *EcoRI/HincII* geschnittenen Leader BL21 Fragment gezeigt. Die Konzentration des Leaderfragmentes wurde konstant bei 2 nM gehalten und die getesteten Proteine in steigenden Konzentrationen hinzutitriert (0, 0,5, 1, 2 und  $4 \mu M$ ).



# Abbildung 2.24: Bindung von CasA, CasB und CasC an das *EcoRI/HincII* geschnittene Leader BL21 Fragment.

Gezeigt sind die Autoradiogramme von Retardierungen mit den Cas-Proteinen A, B und C auf einer 5%igen, nativen PAGE getrennt. In einem Volumen von 15  $\mu$ l wurden 2 nM radioaktiv markiertes Leader BL21 Fragment mit steigenden Konzentrationen der verschiedenen Proteine für 10 Min. bei Raumtemperatur (RT) inkubiert (0, 0,5, 1, 2 und 4  $\mu$ M). Um elektrostatische Wechselwirkungen zu unterdrücken, wurde den Ansätzen Heparin (20 ng/  $\mu$ l) hinzugegeben. Die jeweils letzten beiden Spuren zeigen Proben ohne die Zugabe von Heparin. Die freie dsDNA 1 und 2, sowie der CasA:Freie dsDNA 2 Komplex sind durch Pfeile markiert.

Die Autoradiogramme des *EcoRI/HincII* hydrolysierten Leaderfragments zeigen eine Bindung von CasA an das Fragment, die bereits bei einer Konzentration von 0,5 µM CasA durch Abnahme der freien dsDNA anhand eines Shifts zu erkennen ist (Spur 2). Mit zunehmenden CasA Konzentrationen (1 und 2 µM) nimmt auch die Bindung an das Leaderfragment zu (Spur 3 und 4). Eine fast vollständige Bindung kann bereits in Spur 4 bei 2 µM beobachtet werden, die in Spur 5 bei 4 µM in Form einer schwachen Komplexbande endet. Die Menge an freier DNA hingegen nimmt mit steigenden CasA Konzentrationen ab (Spuren 1 bis 5). Eine Bindung der Proteine CasB und CasC an das *EcoRI/HincII* geschnittene Leaderfragment kann nicht beobachtet werden. Nach Restriktionshydrolyse liegt das Leader BL21 Fragment als ein 322 bp großes doppelsträngiges Fragment vor. Auf den oben dargestellten Retardierungsanalysen ist jedoch mehr als nur eine Fragmentbande zu sehen. Eine mögliche Ursache für das Auftreten dieser zusätzlichen Fragmentbanden kann zum einen die sogenannte 'Star-Aktivität' des EcoRI Restriktionsenzyms sein. Unter bestimmten Bedingungen zeigt *EcoRI* eine verminderte Selektivität für seine Erkennungssequenz und schneidet DNA-Sequenzen außerhalb der eigentlichen Erkennungssequenz. Des Weiteren kann es sich bei den zusätzlich detektierten Fragmenten des Leader BL21 Fragmentes um aufgeschmolzene Einzelstrangbereiche handeln, die bei der Phenol/Chloroform-Aufreinigung während der radioaktiven Markierung entstanden sind.

Für die Aufklärung dieser zusätzlichen Fragmentbanden wurde das zuvor radioaktiv markierte EcoRI/HincII geschnittene Leaderfragment nach der Klenow-Reaktion in 1x KGlu80-Puffer für 3 Min. bei 96°C denaturiert und zur Renaturierung langsam im Heizblock auf Raumtemperatur abgekühlt. Dieser Denaturierungs-Renaturierungsschritt des Fragments führt dazu, dass bei der Klenow-Reaktion entstandene einzelsträngige Fragmentanteile wieder in doppelsträngige überführt werden. Die im Anschluss durchgeführte Retardierungsanalyse mit steigenden Konzentrationen des CasA Proteins zeigt jedoch gleiche Ergebnisse wie zuvor in Abbildung 2.24 (Ergebnisse nicht gezeigt). Dies belegt, dass es sich bei den zusätzlich auftretenden Banden nicht um einzelsträngige Fragmente handeln kann, sondern stützt die Vermutung, dass es sich um kurze doppelsträngige Fragmente handelt, die während der Restriktionshydrolyse entstanden sein müssen. Um dies zu bestätigen wurde das hydrolysierte Leaderfragment auf einem 15% Harnstoffgel denaturierend getrennt (5.2.4.3) und anschließend mit Silber gefärbt (5.2.5.3). Nach der Silberfärbung kann zusätzlich zu der 300 bp Bande, der Größe des EcoRI/HincII hydrolysierten Fragments, eine weitere Bande oberhalb von 200 bp (siehe Anhang Abb. 6.12) detektiert werden, was die Existenz eines weiteren kurzen DNA-Fragmentes anzeigt.

Um die Bindung von CasA an kurze doppelsträngige DNA-Fragmente, wie sie in Abbildung 2.24 beobachtet wurde verifizieren zu können, wurde im weiteren Verlauf das isolierte Leader BL21 Fragment zusätzlich mit *BamHI/SacI* hydrolysiert. Nach Restriktion und Klenow-Markierung wurde hierbei der obere (*non-template*) Strang des 298 bp großen Fragments radioaktiv markiert. Die Ergebnisse der Retardierungsanalysen sind nachfolgend in Abbildung 2.25 gezeigt.



# Abbildung 2.25: Bindung von CasA, CasB und CasC an das *BamHI/SacI* geschnittene Leader BL21 Fragment.

Gezeigt sind die Autoradiogramme von Retardierungen der Cas-Proteinen CasA, B und C mit einer 5% igen, nativen PAGE getrennt. In einem Volumen von 15  $\mu$ l wurden 2 nM radioaktiv-markiertes Leader BL21 Fragment mit steigenden Konzentrationen der Proteine bei Raumtemperatur (RT) für 10 Min. inkubiert (0, 0,5, 1, 2 und 4  $\mu$ M). Um elektrostatische Wechselwirkungen zu unterdrücken, wurde den Ansätzen Heparin (20 ng/  $\mu$ l) hinzugegeben. Die jeweils letzten beiden Spuren zeigen Proben ohne die Zugabe von Heparin.

Wie schon zuvor wurde die Konzentration des Leaderfragments konstant bei 2 nM gehalten und die getesteten Proteine in steigenden Konzentrationen hinzutitriert (0, 0,5, 1, 2 und 4  $\mu$ M). Eine Bindung an das 298 bp große *BamHI/SacI* hydrolysierte Leaderfragment kann bei allen drei Proteinen nicht gezeigt werden und belegt die zuvor beobachtete Bindung von CasA an das zweite freie *EcoRI/HincII* geschnittene Leaderfragment (Abb. 2.24 Spuren 1 bis 5). Somit kann eine Bindung von doppelsträngigen DNA-Fragmenten die kleiner als 298 bp sind für *E. coli* CasA festgehalten werden. Für eine Verifizierung dieser Aussage und nähere Charakterisierung in Bezug auf die Länge der bevorzugten dsDNA, sind jedoch weiterführende Experimente notwendig. Aufgrund von Zeitmangel war dies jedoch kein Bestandteil mehr dieser Arbeit und wird deswegen im Ausblick nochmal behandelt.

### 2.5.7.2 Interaktion von CasA, B und C an Einzelstrang-DNA

Als nächstes galt es das Bindungsverhalten der isolierten Cas-Proteine an einzelsträngige DNA zu testen. Hierfür wurde ein synthetisches Oligonukleotid (LeaderI-Rep1-58nt) verwendet und mit Hilfe der T4-DNA Kinase am unphosphorylierten 5'-Ende mit [ $\gamma$ -<sup>32</sup>P]-ATP radioaktiv markiert (5.2.6.7). Nach einer 10 minütigen Inkubation des Primers mit unterschiedlichen Konzentrationen der Cas-Proteine bei Raumtemperatur wurde das Bindungsverhalten auf einem 5% nativen PAA-Gel analysiert (5.4.4). Das Ergebnis dieser Gelretardierung ist in Abbildung 2.26 gezeigt.



Abbildung 2.26: Analyse der Bindung von CasA, CasB und CasC an einzelsträngige DNA.

Gezeigt sind die Autoradiogramme nach Gelretardierung der Cas-Proteinen A, B und C mit dem zuvor radioaktiv markierten LeaderI-Rep1-58nt Oligonukleotid (4.4.1). In einem Volumen von 15  $\mu$ l wurden 10 nM radioaktiv-markiertes Oligonukleotid mit steigenden Konzentrationen der verschiedenen Proteine bei Raumtemperatur inkubiert (0, 0,5, 1, 2 und 4  $\mu$ M). Als Kompetitor wurde den Ansätzen Heparin (20 ng/  $\mu$ l) hinzugegeben. Die jeweils letzte Spur zeigt einen Ansatz ohne die Zugabe von Heparin. Die Trennung der Komplexe erfolgte auf einer 5%igen, nativen PAGE.

Bei zwei der drei getesteten Proteine kann eine Bindung an das radioaktiv markierte DNA-Oligonukleotid beobachtet werden. Nach Titration der Proteine CasA und CasB ist eine schwache Bindung an das 58 nt lange DNA-Oligonukleotid zu erkennen. Bei CasA kann diese Bindung bereits bei einer Konzentration von 1  $\mu$ M beobachtet werden und nimmt mit steigenden CasA Konzentrationen zu. Nach Zugabe von 4  $\mu$ M ist die Bildung eines CasA/ssDNA Komplexes zu erkennen (Spuren 3 bis 5). Eine Bindung von CasB an die einzelsträngige DNA ist erst bei 4  $\mu$ M durch die Bildung einer Komplexbande zu sehen (Spur 5). Eine Abnahme der freien DNA hingegen ist aufgrund der hohen Oligonukleotidkonzentration nicht zu beobachten. Die Bindung von CasC an einzelsträngige DNA konnte nicht gezeigt werden.

Als Ergebnis dieser Retardierungsanalyse kann eine schwache Bindung von CasA und CasB an einzelsträngige DNA festgehalten werden (Abb. 2.26). Vergleicht man die Zunahme der Komplexbande nach Titration beider Proteine kann eine stärkere Bindung von CasA an das radioaktiv markierte Oligonukleotid als von CasB beobachtet werden (Vergleich der Spuren 1 bis 5). Im direkten Vergleich zu den gezeigten Retardierungsanalysen in 2.5.7.1 wird jedoch deutlich, dass CasA mit einer wesentlich höheren Affinität an das verkürzte doppelsträngige Leader BL21 Fragment bindet (Abb. 2.24). Eine Bindung des CasC Proteins an ssDNA kann nicht gezeigt werden.

# 2.5.7.3 Bindeanlaysen von CasA, CasB und CasC mit einzel- und doppelsträngigen RNA-Fragmenten

Für die Untersuchung einer Interaktion der isolierten *E. coli* Cas-Proteine A, B und C mit einzel-und doppelsträngigen RNA-Fragmenten wurden verschiedene RNAs genutzt. Die 3'-Endmarkierung von RNA-Fragmenten wurde dabei durch die T4-RNA-Ligase mit radioaktiv markiertem Cytosin ([5'-<sup>32</sup>P]-pCp) erreicht (5.2.6.8).

Als erstes wurde die Bindung der Cas-Proteine an eine zuvor transkribierte pre-crRNA getestet. Die Herstellung dieser 257 nt langen einzelsträngigen RNA (ssRNA) erfolgte aus dem Plasmid pZA-T7S1+ durch eine *in vitro* Synthese mit Hilfe des RiboMAX-Systems von Promega (5.4.1) und beinhaltet einen CRISPR-Array bestehend aus Repeat 1 bis 4. Bedingt durch die palindromische Natur kann es innerhalb der Repeatsequenzen zur Ausbildung sogenannter 'Loop-Strukturen' kommen, wodurch doppelsträngige RNA-Bereiche entstehen. Für die Analyse einer Komplexbildung wurden 10 nM radioaktiv markierte pre-crRNA mit ansteigenden Mengen der Proteine CasA, CasB und CasC für 10 Min. bei Raumtemperatur (RT) inkubiert und anschließend 20 ng/µl Heparin als Kompetitor hinzugegeben, um unspezifische Bindungen zu vermeiden. Nach einer weiteren Inkubation für 5 Min. bei RT wurden die Proben auf einem 5%igen nativen PAA-Gel getrennt. In Abbildung 2.27 ist das Ergebnis dieser Gelverzögerungsanalyse gezeigt.



#### Abbildung 2.27: Retardierungsanalyse von CasA, CasB und CasC an pre-crRNA.

Gezeigt sind die Autoradiogramme der Bindungsanalysen der Proteine CasA, CasB und CasC und einer 257 nt langen pre-crRNA. In einem Volumen von 15  $\mu$ l wurden 10 nM radioaktiv markierte pre-crRNA mit steigenden Konzentrationen der verschiedenen Proteine bei Raumtemperatur inkubiert (0, 0,5, 1, 1,5, 2, 4 und 8  $\mu$ M). Als Kompetitor wurde den Ansätzen Heparin (20 ng/ $\mu$ l) hinzugegeben. Die ersten drei Abbildungen zeigen Analysen der pre-crRNA nach Titration der jeweiligen Proteine CasA, B und C. In der letzten Abbildung erfolgte die Inkubation der pre-crRNA zusammen mit zwei Proteinen (CasA und CasB). Die Trennung der Komplexe erfolgte auf einer 5%igen, nativen PAGE. Die durch frühzeitige Termination entstandenen pre-crRNA Transkripte, sowie die freie vollständige pre-crRNA und der CasA:pre-crRNA Komplex sind durch Pfeile markiert.

Neben dem freien Volllängen-Transkript der pre-crRNA sind weitere Banden in der Abbildung 2.27 zu sehen. Bei diesen zusätzlichen Banden handelt es sich wahrscheinlich um kurze pre-crRNA Transkripte, die während der in vitro RiboMAX Synthese durch Termination an den ausgebildeten 'Loop-Strukturen' der Repeatsequenzen entstanden sind. Eine starke Bindung von CasA an das Volllängen-Transkript der pre-crRNA kann durch eine Feintitration in 0,5 µM Schritten nachgewiesen werden (links, Spuren 1 bis 6). Vergleicht man die Spuren 1 und 2 kann bereits bei 0,5 µM CasA ein Shift der freien RNA beobachtet werden, der mit steigenden CasA Konzentrationen immer weiter zunimmt (Spuren 3 bis 5) und in Spur 6 bei 4 µM als Komplexbande erscheint. Gleichzeitig ist eine Abnahme an freier pre-crRNA mit steigender Komplexbildung zu erkennen. Die Titrationsreihen der Proteine CasB und CasC mit unterschiedlich hohen Konzentrationen der Proteine hingegen lassen keine Bindung an die pre-crRNA erkennen. Aufgrund der Tatsache, dass die beiden Proteine CasA und CasB innerhalb des E. coli Cascade-Komplexes

#### 2 Ergebnisse

miteinander interagieren wurde überprüft, ob eine gleichzeitige Inkubation beider Proteine mit der pre-crRNA zu einem Supershift führt. Das letzte Autoradiogramm auf der rechten Seite zeigt die Analyse nach Inkubation, bei der die Konzentrationen des CasA Proteins konstant bei 0,5  $\mu$ M gehalten und CasB in unterschiedlichen Konzentrationen hinzutitriert (0, 0,5, 1, 2 und 4  $\mu$ M) wurde. Ein Supershift der pre-crRNA nach steigenden CasB Konzentrationen kann in den Spuren 1 bis 5 nicht beobachtet werden.

Anhand der gezeigten Ergebnisse kann eine starke Bindung des aus *E. coli* isolierten CasA-Proteins an die precursor-crRNA festgehalten werden. Gleichzeitig weisen die Bindungsanalysen auf eine hohe Affinität gegenüber dieser Art von CRISPR-RNA hin. Bei der Bildung der pre-crRNAs wird der gesamten CRISPR-Locus abgelesen und im Anschluss das transkribierte Primärtranskript (pre-crRNA) endonukleolytisch in kleine crRNAs unter Beteiligung verschiedener Cas-Proteinen prozessiert. In *E. coli* erfolgt diese Prozessierung spezifisch durch die Endonuklease Cas6e, die Bestandteil des für die Immunabwehr essentiellen Cascade-Komplexes ist. Die starke Bindung von CasA in Abbildung 2.27 lässt vermuten, dass auch CasA an der Spaltung der Vorläufer-RNA in die kleinen crRNAs beteiligt ist und wird nachfolgend näher im Diskussionsteil erläutert.

Im weiteren Verlauf für die Charakterisierung potentieller Bindepartner der isolierten Cas-Proteine wurden Retardierungsanalysen mit der 6S RNA aus *E. coli* und einem pRNA-Oligonukleotid durchgeführt. Die Interaktion des pRNA-Oligonukleotids wurde mittels einer "kalten" Retardierung in Gegenwart der isolierten Cas-Proteine getestet. Diese pRNA ist ein 20 nt kurzes Transkript der 6S RNA und durch ihre knappe Länge nur wenig strukturiert. Für die Analysen wurden 200 µM der pRNA mit verschiedenen Konzentrationen der Proteine CasA, -B und –C mit 1x Bindungspuffer und 20 ng/µl Heparin für 15 Min. bei RT inkubiert. Die Ansätze wurden anschließend auf einem 5%igen nativen PAA-Gel getrennt und mit Ethidiumbromid gefärbt (5.2.5.1). Das Ergebnis dieser Analyse lässt keine Bindung der Proteine an die pRNA erkennen und ist im Anhang in Abbildung 6.13 dargestellt.

Die 6S RNA aus *E. coli* zählt zu den stabilen, regulatorischen RNAs, ist 184 Nukleotide lang und durch ihre spezifische Sekundärstruktur charakterisiert. Um eine Bindung der Proteine an die meist doppelsträngigen Bereiche der 6S RNA untersuchen zu können, wurden 15 nM radioaktiv markierte 6S RNA mit steigenden Konzentrationen der Proteine CasA, -B und -C (0, 0,5, 1, 2 und 4  $\mu$ M) für 10 Min. bei Raumtemperatur (RT) inkubiert und nach Zugabe von 20 ng/ $\mu$ l Heparin, um unspezifische Bindungen zu vermeiden auf einem 5%igen PAA-Gel getrennt.

Die Abbildung 2.28 zeigt das Ergebnis der Retardierungsanalysen nach Autoradiographie. Eine Bindung der Cas-Proteine A, B und C nach Inkubation mit 6S RNA kann auch hier nicht beobachtet werden.



#### Abbildung 2.28: Analyse der Bindung von CasA, CasB und CasC an 6S RNA.

Die dargestellten Autoradiogramme zeigen Gelverzögerungsanalysen der Proteine CasA, CasB und CasC und der 184 nt langen 6S RNA aus *E. coli*. In einem Volumen von 15  $\mu$ l wurden 15 nM radioaktiv-markierte 6S RNA mit steigenden Konzentrationen der verschiedenen Proteine bei Raumtemperatur inkubiert (0, 0,5, 1, 2, und 4  $\mu$ M). Als Kompetitor wurde den Ansätzen Heparin (20 ng/  $\mu$ l) hinzugegeben. Die Analyse der Gelretardierung erfolgte auf einer 5%igen, nativen PAGE.

Anhand der gezeigten Abbildungen (2.27, 2.28 und 6.13) kann eine bevorzugte Bindung des *E. coli* CasA Proteins mit hoher Affinität an pre-crRNA festgehalten werden. Eine Bindung an die 6S RNA aus *E. coli* oder einem kurzen pRNA Oligonukleotid kann im Vergleich nicht nachgewiesen werden und zeigt, dass die Bindung von CasA an die pre-crRNA nicht auf eine simple Interaktion mit RNA zurückzuführen ist, sondern eher als eine sequenzspezifische Bindung angesehen werden kann. Des Weiteren kann keine Bindung der anderen zwei Cas-Proteine CasB und CasC an eine der drei RNAs gezeigt werden. Ob dabei die Sequenzen der 6S RNA oder der pRNA entscheidend für eine Bindung sind müsste in weiterführenden Analysen näher untersucht werden.

### 3 Diskussion

Diese Arbeit befasste sich mit dem in Prokaryoten adaptiv erworbenem CRISPR/Cas-Abwehrsystem von *E. coli* und war in zwei Hauptbereiche unterteilt. Zum einen sollte ein Plasmid-System zur Untersuchung der Spacerintegration in Typ I-E Systemen aufgebaut und nach erfolgreicher Etablierung anhand von *in vivo* Integrationsstudien analysiert werden. Zum anderen sollten die den Cascade-Komplex bildenden fünf Cas-Proteine Cse1 (CasA), Cse2 (CasB), Cas7 (CasC), Cas5 (CasD) und Cas6e (CasE) aus *E. coli* über Affinitätschromatographie isoliert und durch *in vitro* Analysen näher charakterisiert werden.

Insbesondere der bisher noch weitestgehend unerforschte Mechanismus der Adaptation war dabei von Interesse. Nach Modifizierung einzelner Basen innerhalb der ersten Leader-nahen Repeatsequenz wurde der Einfluss der Sekundärstruktur und der Primärsequenz auf die Fähigkeit, neue Spacer aus Fremd-DNA aufnehmen zu können untersucht. Dabei konnte gezeigt werden, dass definierte Sequenzbereiche wichtig für die Integration der neuen Spacer sind. Durch Komplementation einzelner Adaptationsbereiche mit Integrationskomponenten aus Typ II-A Systemen wurde zudem versucht, ein heterologes System zur Spaceraufnahme in *E. coli* zu schaffen. Bei der Charakterisierung der Proteinuntereinheiten des *E. coli* Cascade-Komplexes konnten drei der fünf Cas-Proteine erfolgreich isoliert und Interaktionspartner durch *in vitro* Bindestudien mit verschiedenen Nukleinsäuren identifiziert werden.

Im nachfolgenden Abschnitt wird die Bedeutung und der Zusammenhang der beobachteten Ergebnisse, die bei der Analyse des Adaptationsmechanismus und Charakterisierung einzelner Cas-Proteine erzielt werden konnten, mit bereits bekannten Daten verglichen und genauer diskutiert.

### 3.1 Etablierung eines Plasmid-basierten Integrations-Assays

Der Adaptationsprozess des CRISPR/Cas-Systems, bei dem neue Spacersequenzen aus Fremd-DNA in das eigene Genom integriert werden, ist ein bis heute noch nicht vollständig aufgeklärter Mechanismus. Der Grundstein für eine Aufklärung wurde bereits 2007 mit der Beobachtung über eine Aufnahme von Fremd-Nukleinsäuren in den CRISPR-Array des Milchsäurebakteriums *S. thermophilus* gelegt und war fortan Bestandteil der nachfolgenden Forschungen (Barrangou, Fremaux et al. 2007). Es konnte gezeigt werden, dass die Aufnahme von neuen Spacern meist am Leader-nahen Ende eines CRISPR-Arrays erfolgt. Den beiden universell vertretenen Cas-Proteinen Cas1 und Cas2 wurde durch ihre Entbehrlichkeit während der Interferenzphase und ihrer vorhandenen Nukleaseaktivität ebenfalls eine Rolle bei der Integration zugesprochen. Der erste handfeste Beweis einer aktiven Beteiligung von Cas1 und Cas2 an der Spaceraufnahme wurde 2012 in *E. coli* experimentell belegt. Dabei wurden gleichzeitig neben einer essentiellen Notwendigkeit dieser beiden Proteine, weitere wichtige DNA-Elemente für eine funktionelle Integration aufgedeckt (Yosef, Goren et al. 2012). Zu diesen wichtigen Elementen wird der bereits zuvor schon erwähnte Leaderbereich gezählt, aber auch die erste Leader-proximale Repeatsequenz, die bei dem Integrationsprozess dupliziert wird.

Beim Aufbau des Plasmid-basierten Integrations-Assays, der in Anlehnung an ein von Qimron und Mitarbeitern (Yosef, Goren et al. 2012) entwickeltes System erfolgte, konnten nach einer erfolgreichen Etablierung neben bisher noch unbekannten Ergebnissen zur Aufklärung des Adaptationsmechanismuses, auch bereits gezeigte Daten bestätigt werden.

Die Funktionalität des synthetischen CRISPR-Arrays der für das Plasmid-System genutzt wurde, konnte nach Klonierung in den Ausgangsvektor pCR001 durch die in Kapitel 2.1.2 gezeigten Ergebnisse anhand von in vivo Adaptationsstudien bestätigt werden. Dabei konnte eine Erweiterung des Plasmid-basierten CRISPR-Arrays des pCR003 WT Plasmids um eine neue Repeat-Spacer Einheit nach Amplifizierung des CRISPR-Bereichs anhand von sc-PCR Reaktionen (single-colony PCR) gezeigt werden (siehe Abb. 2.2). Gleichzeitig kann eine Zunahme der aufgenommenen Spacer von Zyklus zu Zyklus beobachtet werden, die im Vergleich zum plasmidären Einbau eine bevorzugte Aufnahme ins Genom von BL21 AI aufzeigt. Während den Zellen der Einbau ins Genom mit teilweise bis zu drei Spacern gelingt, lassen sich für das pCR003 WT Plasmid-System nur maximal zwei zusätzliche Spacer beobachten. Eine Abnahme der Spacerintegration hingegen kann in beiden Fällen meist nach Zyklustag vier vermerkt werden. Die Zykluskontrollen nach Tag fünf ließen sogar zum Teil nur noch die parentalen Banden erkennen. Da während der Adaptationsstudien das Wachstumsverhalten der Zellen nicht konstant über alle Zyklustage hinweg kontrolliert wurde, beziehen sich die nachfolgenden Aussagen nur auf stichprobenartige Wachstumswerte. Wobei die Zellen an den Zyklustagen eins bis drei ein exponentielles Wachstumsverhalten zeigten, konnte nach Zyklustag vier eine verringerte Zellzahl nach Messung der optischen Dichte vermerkt werden. Dieser Wachstumsnachteil spiegelt sich auch anhand des reduzierten Einbaus wieder, der bei Zyklustag vier durch die parentale Bande sichtbar wird. Einer der Gründe für das verringerte Zellwachstum kann die erhöhte Expression der plasmid-codierten cas1 und cas2 Gene bedingt durch die zusätzliche Induktion mit IPTG und L-Arabinose sein. Unter optimalen Bedingungen werden Gene in der Zelle sehr effizient produziert. Die Zugabe der Induktoren führt dabei zu einer vermehrten Translation der Gene in hohen Konzentrationen und verbraucht somit einen signifikanten Anteil der Energie der Zelle. Der Energieverbrauch der Zelle macht sich wiederum durch einen Wachstumsnachteil bemerkbar. Die Analyse der Spacerintegration erfolgte immer über mehrere Zyklen. Dabei kann die erhöhte Proteinmenge der Cas1/Cas2 Nukleasen einen negativen Effekt auf die Zellen ausüben. Zudem kann es sein, dass eine zu starke Aufnahme von Spacern bedingt durch die Induktion der cas1 und cas2 Gene, die DNA instabilisieren, so dass die Zellen einen Wachstumsnachteil erfahren. Um die getroffenen Aussagen in Bezug auf das Wachstumsverhalten der Zellen jedoch näher beurteilen zu können, müssten die Adaptationsstudien reproduziert und anhand von Wachstumskurven funktionell analysiert werden.

Der Nachweis und die Analyse der neu erworbenen Spacersequenzen waren jeweils sehr zeitaufwendig und bestanden aus mehreren Schritten, wie Ausplattieren, PCR-Amplifikationen, Plasmid Isolierungen, Transformationsreaktionen, erneute PCR-Analyse (siehe exemplarisch Abb. 3.1 einer *single-colony* PCR-Analyse) und schließlich Sequenzierungen. Ein genaues Schema dieses Verfahren ist im Anhang in Abbildung 6.2. dargestellt.



Abbildung 3.1: *Single-colony* PCR-Analysen (sc-PCR) des Plasmid-basierten CRISPR-Array.

Nach Isolierung der Plasmid-DNA von Klonen die einen neuen Spacer integriert haben (A), erfolgt die Transformation in kompetente XL-1 Zellen. Anschließenden werden 20 Einzelkolonien der transformierten Agar-Platte gepickt und in einer zweiten PCR-Reaktion nach Klonen, die nur den 'single' Typ des erweiterten Plasmid-Arrays zeigen für eine darauffolgende Sequenzierungsreaktion analysiert (B). Parentale PCR-Fragmente sind mit "parental" gekennzeichnet und neu erworbene Spacersequenzen mit "extended" (entnommen aus Arslan, Hermanns et al., in Revision).

Die Ergebnisse der Sequenzierungsanalysen sind in Übereinstimmung mit den Befunden anderer Autoren und zeigen den Einbau des neuen Spacers in unmittelbarer Nähe der Leaderregion, sowie die Verdoppelung der ersten Repeatsequenz (siehe Abb. 2.3). Gleichzeitig stützten die Daten die bisherige Hypothese, dass bei der Aufnahme von Fremd-DNA die beiden komplementären Repeatstränge durch eine sequenzspezifische Spaltung so gegeneinander verschoben werden, dass sie einer DNA-Polymerase als *template* für die Verdoppelung dienen (siehe Abb. 1.7). Die Tabelle in der alle neu erworbenen Spacersequenzen zusammengefasst wurden zeigt wiederholt die Funktionalität des aufgebauten Plasmid-basierten Integrationssystems (siehe Tabelle 2.1). Neben einer für *E. coli* charakteristischen Repeatlänge von 29 bp konnte die Länge der Spacersequenz, wie zu erwarten war, mit einer Länge von 32 bp detektiert werden. Die in vorherigen Studien aufgedeckte Tatsache, dass nach Integration von neuen Spacern am Leader-nahen Ende eines CRISPR-Locus die letzte Base der *upstream* Repeatsequenz identisch zum letzten Nukleotid der Protospacer-Sequenz ist konnte ebenfalls bestätigt werden (Datsenko, Pougach et al. 2012). Als Herkunftsort der neu erworbenen Spacer konnte sowohl die genomische DNA des BL21 AI Stammes, als auch das als Fremd-DNA Träger fungierenden pUC18-Kan Plasmid experimentell belegt werden.

### 3.2 Charakterisierung der Spacerintegration in Typ I-E CRISPR/Cas-Systemen

Obwohl der konkrete Mechanismus der Adaptation bis heute nicht vollständig verstanden ist, wurde bereits festgestellt, dass der Leaderbereich als Bindeplattform für rekombinante Proteine dient und neben den für die Integration ubiquitären Cas1 und Cas2 Proteinen, auch das erste Leader-nahe Repeat essentiell ist. Deletionen der kompletten Leaderregion führen zu einem Verlust der Spacerintegration (Yosef, Goren et al. 2012). Inwiefern der Leaderbereich Einfluss auf die Integration neuer Spacer nimmt, wurde im Rahmen einer Bachelorarbeit in unserer Arbeitsgruppe untersucht (Popowitsch 2013). Dabei konnte anhand von durchgeführten Mutationsanalysen an unterschiedlichen Sequenzbereichen der Leaderregion gezeigt werden, dass neben dem mittig lokalisierten AT-Cluster-Bereich auch die Leader-Repeat-Grenze ausschlaggebend den Einbau von neuen Spacern beeinflusst. Das Ergebnis der in vivo Integrationsstudien zeigte eine deutliche Reduktion des Spacereinbaus der modifizierten Plasmide. Des Weiteren zeigte eines der Plasmide nach der Mutagenese eine Deletion der ersten drei Repeatsequenzen des synthetischen CRISPR-Arrays. Anhand einer Rückmutation konnte bestätigt werden, dass nur eine Repeatsequenz für eine funktionelle Spacerintegration ausreichend ist.

Die erfolgreiche Etablierung des Plasmid-basierten Integrationssystems in Kapitel 2.1 erlaubte einfache Modifizierung **CRISPR-Sequenzen** eine der durch Mutationsanalysen. Dabei stand die Charakterisierung der DNA-Sequenzspezifität des ersten Repeats im Vordergrund. In zurückliegenden in vitro Studien konnte gezeigt werden, dass das Cas1 Protein von E. coli sogenannte 'cruciform-' DNA-Strukturen spaltet (Babu, Beloglazova et al. 2011). So entstand die Hypothese, dass Cas1 in einer DNA-Strukturspezifischen Art und Weise an die Repeatsequenz bindet, die bedingt durch die palindromische Natur solche 'cruciform-' Strukturen ausbildet (siehe Abb. 2.4). Ein weiterer Aspekt war die gezeigte Tatsache, dass solche palindromisch bedingt ausgebildeten DNA-Strukturen spezifisch von Nukleasen oder Integrasen während einer Rekombinations- und Integrationsreaktion gebunden werden können (Cote and Lewis 2008). Ein ähnlicher Mechanismus könnte auch während des Adaptationsprozesses der CRISPR/Cas-vermittelten Abwehr in *E. coli* genutzt werden und wurde durch gezielt eingeführte Basenmodifikationen der ersten Leader-nahen Repeatsequenz untersucht (siehe Modell Abb. 3.2).



# Abbildung 3.2: Hypothetisches Modell der kreuzförmigen Strukturausbildung der ersten Repeatsequenz.

Es konnte gezeigt werden, dass das *E. coli* Cas1 Protein Nukleaseaktivität gegenüber verzweigten DNA-Strukturen, sogenannten 'Holliday junctions' aufweist (Babu, Beloglazova et al. 2011). Aufgrund dessen wurde angenommen, dass es während des Integrationsprozesses zur Ausbildung solcher 'cruciform' DNA-Strukturen kommen kann, die vom Cas1 Protein in einer Sequenz-unabhängigen, aber DNA-Struktur-spezifischen Weise gespalten werden. Nach Modifizierung der ersten Leader-nahen Repeatsequenz ist die Strukturausbildung gehemmt. Die Repeats (R) sind durch graue Rauten bzw. kreuzförmige Loops gezeigt und die Spacer (S) durch farbige Kästchen bzw. Striche markiert.

### 3.2.1 Rolle der ersten Leader-proximalen Repeatsequenz während der Integration neuer Spacer

Bei der Charakterisierung der ersten Leader-nahen Repeatsequenz konnten durch die Einführung verschiedener Punktmutationen an zuvor ausgewählten Sequenzbereichen erste Hinweise zur Beantwortung der Frage nach einer DNA-Struktur- oder Sequenz-spezifischen Spaltung während der Aufnahme von neuen Spacersequenzen aufgedeckt werden. Nach Einführung der drei unterschiedlichen Mutationen, die zu einer veränderten Sequenzabfolge des ersten an Position 6 bis 9 gelegenen C-Clusters führten, zeigten die in Kapitel 2.2 durchgeführten Integrationsstudien in allen drei mutierten Plasmiden eine signifikante Reduzierung der Spaceraufnahme gegenüber den Positivkontrollen. Des Weiteren konnte anhand einer Rückmutante (RM1-R) eine rein DNA-strukturabhängige Integration ausgeschlossen werden.

Der ersten eingeführten Mutation RM1 ging die Fragestellung voraus, ob eine Modifizierung dieser Positionsbasen Auswirkungen auf den Einbau von neuen Spacern hat, da nach Mutationseinführung die Ausbildung der hypothetischen 'cruciform-Struktur' unterbunden ist. Bei den Analysen der in vivo Integrationsstudien (2.2.1.1) stellte sich heraus, dass die Aufnahme von neuen Spacern in der Mutante im Vergleich zum Wildtyp-Vektor reduziert ist (siehe Abb. 2.5). Die parallel durchgeführte genomische und plasmidäre Analyse der Spacerintegration des pCR003 WT Vektors diente als Positivkontrolle und lässt zeitgleich Rückschlüsse auf die eingebaute Mutation zu. Die Tatsache, dass der genomische Einbau in der Mutante RM1 im Vergleich zum plasmidären nicht reduziert ist zeigt, dass die eingeführte Mutation innerhalb des synthetischen CRISPR-Arrays Einfluss auf den Einbau neuer Spacer nimmt. Die unterbundene Basenpaarung der einzelnen Nukleinbasen und der daraus resultierende Verlust der 'Kreuzstruktur'-Ausbildung weisen zunächst auf eine nicht-sequenzspezifische, DNA-strukturabhängige Spaltung der ersten Repeatsequenz hin. Für die Klärung dieser Frage wurde in einem nächsten Schritt eine Rückmutante des bereits mutierten pCR003 RM1 Plasmid erstellt, in der die Basenkomplementarität zwischen den einzelnen Nukleinbasen wiederhergestellt wurde.

Während die vorherige RM1 Mutation den Verlust der Strukturausbildung innerhalb der Repeatsequenz zur Folge hatte, zielte die eingeführte Mutation in Kapitel 2.2.2 auf eine Wiederherstellung der Ausbildung mit einer veränderten Primärsequenz. Die Ergebnisse dieser Rückmutation zeigen im direkten Vergleich zur Wildtyp- und RM1-Situation weiterhin eine reduzierte Aufnahme in den Plasmid-basierten CRISPR-Array (siehe Abb. 2.6). Der Einbau ins Genom dagegen erfolgt in allen drei Zelltypen. Die mutierten Plasmiden lassen lediglich die parentalen Banden erkennen und deuten auf einen nicht-strukturspezifischen Mechanismus während der Integration hin. Um die erneut auftretende Frage nach einer Sequenzspezifität statt einer Sekundärstrukturspezifität innerhalb der ersten Repeatsequenz zu klären, wurde in einer weiteren Mutagenesereaktion das pCR003 RM1-R Plasmid modifiziert.

Mit der eingeführten letzten Mutation, einer direkten Sequenzveränderung ohne jedoch die Strukturausbildung zu unterbinden galt es, die Frage nach einer sequenzspezifischen Nukleinbasen-abhängigen Integration zu klären. Diese Mutation führte zum Verlust der Spacerintegration (siehe Abb. 2.7). Da die parallel mitgezogene Positivkontrolle in Form des pCR003 Wildtyp-Vektors eine Überexpression der *cas* Gene durch einen aktiven Einbau sowohl ins Genom als auch Plasmid belegt und auch die Mutanten genomischen Einbau zeigten, kann daraus geschlossen werden, dass der Austausch dieser vier Basenpaare ohne Sekundärstrukturverlust den Einbau von neuen Spacern verhindert. Dies zeigt, dass das C-Cluster innerhalb des ersten Repeats essentiell für den Integrationsprozess ist. Eine Veränderung der DNA-Struktur wie es in den beiden vorherigen Mutationen RM1 und RM1-R der Fall war, scheint hingegen nicht alleine für einen aktiven Einbau verantwortlich zu sein. Die durchgeführten Mutationsanalysen weisen auf die Bedeutung der ersten Leaderproximalen Repeateinheit bei der Adaptation neuer Spacer hin. Nach Einführung der Mutationen in die untersuchten Plasmide konnte eine signifikante Reduzierung des Einbaus neuer Spacer beobachtet werden. Die Rückmutation des pCR003 RM1 Plasmid zeigte dabei gleichzeitig, dass eine rein strukturabhängige Integration ausgeschlossen werden kann und die Spaltung der ersten Repeatsequenz eher sequenzspezifisch erfolgt.

### 3.2.2 Aufklärung des Adaptationsmechanismus in E. coli

Nachdem die Mutagenesanalysen einen ersten Hinweis auf eine sequenzspezifische Spaltung der ersten Leader-proximalen Repeatsequenz ergab, konnte dieser durch weiterführende *in vivo* Analysen verifiziert werden.

Die CRISPR/Cas-vermittelte Adaptation gegen fremde Nukleinsäuren ist abhängig von der Insertion unbekannter DNA-Sequenzen als neue Spacer in den eigenen CRISPR-Array. Der Einbau dieser Fremd-DNA Sequenzen erfolgt dabei meist am Leader-nahen Ende unter Duplikation der ersten Repeatsequenz und zeigt eine direkte Beteiligung beider DNA-Stränge des Repeats als template für die anschließende Auffüllreaktion der Polymerase. Demnach muss dem Spacer Integrationsprozess ein Spaltungsereignis vorangehen. In ersten Untersuchungen des genomischen CRISPR-Arrays von E. coli BL21 AI konnten bereits Zwischenprodukte dieser Spaltungsreaktion detektiert und basierend darauf ein gekoppelter Spaltungs-Ligationsmechanismus aufgedeckt werden (Arslan, Hermanns et al., in Revision). Um die Bildung solcher Zwischenstufen auch im Plasmid-basierten CRISPR-System zu bestätigen und eine Repeat-abhängige Spaltung mit gekoppelter Ligation des neuen Spacers aufzeigen zu können, wurde versucht existierende Zwischenprodukte mit Hilfe von Southern Blot Analysen nachzuweisen. Die postulierten Intermediate sollten bei der Integration von neuen Spacersequenzen in den synthetischen CRISPR-Array der Plasmide durch eine sequenzspezifische Öffnung des Arrays vor Einbau des neuen Spacers entstehen. Der Nachweis erfolgte mit Hilfe von radioaktiv markierten Sonden an bestimmten Bereichen des synthetischen CRISPR-Arrays von zuvor isolierter und hydrolysierter Plasmid-DNA (siehe Abb. 2.8).

### 3.2.2.1 Verifizierung der Cas1/Cas2-vermittelten Spaltungs-Ligations-Reaktion

Für die Detektion wurden die Plasmide pCR003 WT, RM1 und RM2 in BL21 AI transformiert und nach Induktion der *cas1-cas2* Expression die zuvor isolierte Plasmid-DNA mit den Restriktionsenzymen *EcoRI* bzw. *KpnI* hydrolysiert, um potentiell wichtige Bereiche des CRISPR-Arrays, an denen eine Spaltung der

82

Nukleinbasen vermutet wird, mit einer höheren Auflösung untersuchen zu können. Die Hydrolyse mit *EcoRI* resultierte in einem Schnitt *upstream* der Leaderregion, wobei die Behandlung mit *KpnI* zu einer Spaltung zwischen dem dritten und vierten Repeat führte. Nach Hybridisierung verschiedener Oligonukleotide in unmittelbarer Nähe der Restriktionsschnitte wurde so der komplette synthetische CRISPR-Array analysiert und entstehende Intermediate detektiert. Ob die Bildung der Intermediate dabei abhängig von der Expression der Plasmid-codierten *cas1* und *cas2* Gene war, wurde parallel durch eine Positivkontrolle in Form des pCR003 WT ohne die Zugabe von IPTG und L-Arabinose während der Zellanzucht analysiert.

Die mit dem pCR003 WT Plasmid erhaltenen Bandenlängen der Southern Blot Analysen in Kapitel 2.3.2 bestätigen eine gekoppelte Spaltungs-Ligations-Reaktion an der ersten Leader-nahen Repeatsequenz, wie sie schon zuvor mit genomischer DNA aus BL21 AI gezeigt werden konnte (Arslan, Hermanns et al., in Revision). Die Ergebnisse der Analyse der Leaderregion lassen einen Schnitt an der Leader-Repeat-Grenze des oberen Stranges mit einer Länge von 120 nt (siehe Abb. 2.9, Spur 3) erkennen. Gleichzeitig wird der untere Strang der KpnI hydrolysierten Wildtypprobe am 3'-Ende der Repeatsequenz bei ungefähr 100 nt gespalten (siehe Abb. 2.10, Spur 3) und eine neue Spacersequenz mit den einzelsträngigen 5'-überhängenden Repeatenden ligiert (siehe Abb. 2.9, 182 nt und Abb. 2.10, 160 nt). Auch wird deutlich, dass die eingeführten Mutationen innerhalb des Repeats die Bildung der Intermediate beeinflussen. Wo hingegen die Signale der oberen Stränge beider Hydrolysen eine reduzierte Bande der pCR003 RM1 Mutante zeigen, wurden keine Signale am komplementären Strang detektiert. Somit scheint zwar eine Spaltung des oberen Stranges an der Leader-Repeat-Grenze zu erfolgen, aber der Einbau bzw. die Ligation eines neuen Spacers ist unterbunden. Diese Daten passen zu den vorherigen in vivo Adaptationsstudien aus Kapitel 2.2.1.1 in denen ebenfalls eine reduzierte Aufnahme von Spacern nach Einführung der Mutation RM1 sichtbar war (siehe Abb. 2.5). Einen weitaus deutlicheren Effekt ließ die Mutante RM2 erkennen. Hier wurden Signale unterschiedlicher Länge weiter downstream der eigentlichen Schnittstellen detektiert, die mit ihren Längen nicht mit dem bisherigen Integrations-Modell übereinstimmten. Diese Änderung im Spaltungsmusters bestärkt die Spezifität der zuvor beobachteten Spaltungsreaktion und zeigt, dass neben einer Protein-abhängigen Expression, auch die Integrationsstelle Cas1/Cas2 der Repeatsequenz von Bedeutung ist. In den nicht-induzierten pCR003 WT Plasmiden können keine Spaltungsprodukte beobachtet werden, was die Notwendigkeit der cas1 und cas2 Gene bei der Integration der CRISPR/Cas-vermittelten Abwehr anzeigt. Die veränderten Spaltungsmuster der mutierten Plasmide zeigen nicht nur, dass die eingeführten Modifizierungen eine deutliche Auswirkung auf die Spezifität der Spaltungs-Ligations-Reaktion haben, sondern weisen zusätzlich auch auf eine Sequenzspezifität der Cas1/Cas2 Proteine hin, die durch den ersten Repeat vermittelt wird.

#### **3 Diskussion**

Die Ergebnisse der Southern Blot Analysen zeigen, dass die gebildeten Zwischenprodukte nicht das Produkt einer einzelnen Spaltungsreaktion sein können, sondern bestätigen eine gekoppelte Integrasereaktion des Cas1 Proteins, in der die Spaltung der Repeatsequenz an der Leader- bzw. Spacerregion mit einer gleichzeitigen Verknüpfung der freien Enden von Repeat und Spacer einhergeht. Das nachfolgende Integrations-Modell zur Aufnahme von neuen Spacern (Abbildung 3.3) zeigt zusammengefasst die erzielten Ergebnisse der Southern Blot Analysen.



#### Abbildung 3.3: Modell der Immunisierung des CRISPR/Cas-Abwehrsystems.

Zu sehen ist eine schematische Darstellung des in unserer Arbeitsgruppe erarbeiteten Mechanismus zur Adaptation neuer Repeat-Spacer-Einheiten. In einem ersten Schritt wird die erste Leader-nahe Repeatsequenz an beiden DNA-Strängen versetzt gespalten und der neue Spacer in die entstandene Lücke integriert. Die einzelsträngigen Repeatüberhänge werden in einem nächsten Schritt in einer Polymerasereaktion aufgefüllt und die freien Enden mittels DNA-Ligase miteinander verknüpft. Es konnte *in vivo* gezeigt werden, dass die Spaltung des ersten Repeats sequenzspezifisch, versetzt erfolgt und mit der Integration einer neuen Spacersequenz in einer simultan gekoppelten Reaktion abläuft (eckige Klammer).

#### 3.2.2.2 Sequenzierung der detektierten DNA-Intermediate

Eine Verifizierung der durch Southern Blot detektierten DNA-Zwischenstufen durch eine zusätzliche Sequenzierungsreaktion, konnte mit dem in 2.3.3 durchgeführten Experimenten nicht zweifelsfrei gezeigt werden. Die durchgeführten Primer-Extension Reaktionen mit Sequenase 2.0 und den nicht denaturierten ccc-Plasmiden des pCR003 Wildtyp mit und ohne Induktion der cas1-cas2 Genexpression als *template*, zeigten zwar eine Übereinstimmung mit den zuvor *in* vivo detektierten Southern Blot Signalen, aber eine exakte Sequenzierung der beobachteten Produktbanden der Primer-Extension Reaktion war nicht erfolgreich. Zum einen lag dies an den zu schwachen Ausbeuten der Intermediate, zum anderen erschwerte eine zu starke Hintergrundpräsenz der parentalen, ungeschnittenen Wildtyp-Plasmide eine klare Unterscheidung zwischen Intermediatund Hintergrundbanden (siehe Abb. 2.12). Selbst nachdem die fünffache Menge der isolierten ccc-Plasmide in die Hybridisierungsreaktion eingesetzt wurde, konnte keine Verbesserung der Ausbeute und Qualität der Intermediate erzielt werden. Für eine Reduzierung der parentalen Plasmidbanden wurden die ccc-Plasmide in einem vorherigen Denaturierungsschritt mit NAOH behandelt, um die doppelsträngigen DNA-Stränge aufzubrechen und eine anschließende Hybridisierung der verwendeten Oligonukleotide zu erleichtern. Das Ergebnis nach Primer-Extension Reaktion erzielte jedoch leider nicht den gewünschten Effekt einer Reduzierung. Dennoch kann als Ergebnis dieser Analysen eine Übereinstimmung der beobachteten Produkte mit den zuvor durch Southern Blot detektierten Signalen nach Verwendung verschiedener Primer festgehalten werden. Bei Verwendung des PE-Leader-pCR001-fw Primers wurde eine Bande bei ca. 100 nt beobachtet, die durch eine Duplikation der ersten Repeatsequenz und den simultanen Einbau eines neuen Spacers zu erklären ist. Der Gegenstrang zeigte ein Signal bei ca. 87 nt, dessen Länge aus einer Erweiterung des synthetischen CRISPR-Arrays durch Addition eines neuen Spacers resultiert. Eine Bestätigung des vorgeschlagenen Modells zur Spacerintegration kann somit auch mit den durch Primer-Extension Reaktion gezeigten Daten festgehalten werden.

### 3.3 Charakterisierung der Spaceraufnahme in Typ II-A Systemen

Der Aufbau eines in *E. coli* funktionellen heterologen Adaptationssystems mit Integrationselementen aus Typ II-A Systemen von *S. agalactiae* konnte mit den in Kapitel 2.4 durchgeführten Klonierungsexperimenten nicht erzielt werden. Dabei wurde anhand verschiedener Experimente die Aktivität der universell vertretenden Cas1 und Cas2 Proteine in Kombination mit für in *S. agalactiae* essentiellen Integrationsproteinen in einem *E. coli* BL21 AI Hintergrund getestet.

Für die Komplementation wurde der Bakterienstamm *S. agalactiae* verwendet, einem früher zum Nmeni Subtyp klassifizierten HNH-Typ der CRISPR/Cas-Systeme.

In diesem ist ein als Cas9 bezeichnetes Protein für die Prozessierung der reifen crRNAs und Degradation der invasiven Fremd-DNA zuständig. Die Spaltung der pre-crRNA wird in Gegenwart des Cas9 Proteins und zwei weiteren Faktoren, einer 'house-keeping' RNA (RNaseIII) und tracrRNA (*trans*-encoded crRNA) katalysiert. Eine Beteiligung der Proteine Cas1 und Cas2 an der Prozessierung und Interferenz von Fremd-DNA ist dabei nicht bekannt und führte zu der Annahme, dass sie eine Rolle in der Immunisierungsphase spielen (Deltcheva, Chylinski et al. 2011). Zusätzlich zu Cas1 und Cas2 ist das für Typ II-A Systeme charakteristische Csn2 Protein ein essentieller Bestandteil für die Integration von neuen Spacern (Deveau, Barrangou et al. 2008). Eine Rolle bei DNA-Endmodifikationen konnte ihm bereits zugeschrieben werden (Arslan, Wurm et al. 2013).

#### 3.3.1 Aufbau eines heterologen Integrationssystems in E. coli

Die funktionelle Aufnahme von neuen Spacern ins Genom von BL21 AI oder dem **CRISPR-Array** der pCR004/pCR005 Plasmide konnte durch in vivo Adaptationsstudien nicht gezeigt werden. Anhand des als Positivkontrolle mitgezogenen pCR003 WT Plasmids kann eine fehlerhafte Vorgehensweise bei der Durchführung der Integrationsstudien ausgeschlossen werden. Ob eine inaktive Genexpression der cas1-cas2 Gene von S. agalactiae in E. coli BL21 AI Zellen eine mögliche Ursache der nicht beobachteten Spaceraufnahme sein könnte, wurde daraufhin in Expressionexperimenten der sagcas1-sagcas2 Gene nach Induktion mit IPTG und L-Arabinose getestet. Die analytische Expression dieser Gene konnte nach Induktion nicht erzielt werden und bestärkt die Vermutung einer inaktiven Genexpression in einem E. coli Hintergrund. Mit dem Hinweis, dass in E. coli eine funktionelle Spacerintegration nur erreicht werden kann, wenn beide Gene unmittelbar hintereinander auf dem Plasmid als ein Operon kodiert vorliegen (mündliche Aussage Dr. Ümit Pul), wurde ein erster Anhaltspunkt für die Gründe der nicht beobachteten Genexpression geliefert. Die Plasmidsequenzen der Vektoren pCR004 bzw. pCR005 zeigten im Vergleich mit dem pCR001 Plasmid keine direkte Überlappung der *cas1* und *cas2* Gene. Daher wurde in einem nächsten Schritt ein neuer Vektor kloniert, der zwischen den *sagcas1-sagcas2* Genen eine acht Basenpaar lange Insertion aus *E. coli* stammend enthielt. Doch eine Expression der Gene kann nach Induktion nicht beobachtet werden (siehe Abb. 2.16). In nachfolgenden Experimenten wurde daraufhin versucht mit Hilfe von verschieden-klonierten Plasmidvarianten, die die für *E. coli* oder *S. agalactia*e charakteristischen Integrationsgene cas1, cas2 und csn2 enthielten, die Gründe für das inaktive heterologe Plasmid-System aufzudecken (siehe Abb. 6.8 im Anhang). Die durchgeführten Integrationsstudien ließen keinen Einbau erkennen. Lediglich eine erfolgreiche Expression des aus *S. agalactiae* stammenden Cas1 Proteins konnte nach Umklonierung des *cas1-cas2* Operons in ein pET-52(b)+ Derivat erreicht werden. Die Nutzung dieses pET-Vektors als Derivat für weiterführende Klonierungsexperimente,

wie sie im Anhang in Abbildung 6.8 näher erläutert sind, führte auch nicht zur gewünschten Expression der fehlenden Proteinkomponenten für den Aufbau eines funktionellen heterologen System in *E. coli.* 

Die gezeigten Ergebnisse deuten darauf hin, dass für eine funktionelle Adaptation in Typ I-E CRISPR/Cas-Systemen zusätzliche Integrationsproteine neben den bereits vorhandenen gebraucht werden. Erste Hinweise dazu gehen auf die Aufdeckung des als "primed adaptation" bezeichneten Prozesses zurück. Bei dieser Art von Adaptation wird im Falle von *E. coli* zusätzlich die Anwesenheit des Cascade-Komplexes und des Cas3 Proteins aus der Interferenzphase benötigt (Datsenko, Pougach et al. 2012, Swarts, Mosterd et al. 2012). Ein weiteres Beispiel ist die Spacerintegration in Typ I-B Systemen wie Haloarcula hispanica die nur über diesen 'priming'-Prozess erfolgt. Hier werden nicht nur die Proteine Cas1 und Cas2, sondern zusätzlich das Cas4 Protein und andere Cas-Proteine die an der Reifung der crRNA und/oder der CRISPR-vermittelten Interferenz beteiligt sind benötigt (Li, Wang et al. 2014). Gleichzeitig ist die Präsenz eines bereits integrierten Spacers der Komplementarität zu viralen Sequenzen aufweist von großer Bedeutung. Diese Tatsachen weisen darauf hin, dass eine solche Art der Adaptation auch für S. agalactiae gelten könnte. Anstelle des Cascade-Komplexes wird in Typ II-Systemen die CRISPR-Interferenz vom Cas9 Protein, einer nicht-kodierende RNA (tracrRNA) und einer house-keeping RNaseIII übernommen. Diese drei Komponenten sind an der Prozessierung der pre-crRNA zu den reifen crRNAs beteiligt, die nach Hydrolyse am Cas9 Effektorprotein gebunden bleiben und als Effektorkomplex die Erkennung der Ziel-Sequenz innerhalb der Fremd-DNA vermitteln. Eine Bereitstellung des Cas9 Proteins und der tracrRNA als zusätzliche Integrationsproteine der Spacerintegration könnte möglicherweise eine Aktivierung des heterologen Adaptationssystems in *E coli* begünstigen.

Neben dem zum Typ II-A zählenden Integrationsprotein Csn2 ist das Cas4 Protein, ein Homolog in Typ II-B Systemen vertreten (Makarova, Grishin et al. 2006). Erst kürzlich konnte für das Csn2 Protein eine ringähnliche Struktur aus Tetrameren und ein DNA-bindender Mechanismus bevorzugt an freie doppelsträngige DNA-Enden veröffentlicht werden (Nam, Kurinov et al. 2011, Arslan, Wurm et al. 2013). Die Struktur des Cas4 Proteins aus Typ II-B Systemen zeigt Sequenzähnlichkeiten zu Exonukleasen der RecB-Familie aus E. coli, was ebenfalls auf einen katalytischen Mechanismus für DNA End-Modifikationen während der Aufnahme von neuen Spacer hindeutet. Ebenso weist die vorhandene 5'-3'-DNA-Exonukleaseaktivität auf eine Funktion bei der Bildung von 3'-Überhängen für die Generierung neuer Spacer hin (Zhang, Kasciukovic et al. 2012, Lemak, Beloglazova et al. 2013). Die Exonukleaseaktivität des Cas4 Proteins und eine bekannte Interaktion mit den Proteinen Cas1 und Cas2 unterstreichen die Beteiligung von Cas4 an der CRISPR-vermittelten Immunisierung. Mit der Beobachtung das in Typ II CRISPR-Systemen die das Csn2 Protein tragen ein AddAB-Reparatursystem vorhanden ist und umgekehrt in den Systemen die das Cas4 Protein tragen ein RecBCD-System, kam die Vermutung auf, dass in E. coli das Fehlen dieser Proteine

durch eine Interaktion von Cas1 und dem house-keeping RecB Protein kompensiert werden könnte. In *E. coli* ist das RecBCD-Reparatursystem bestehend aus den Proteinen RecB, RecC und RecD für die Reparatur von DNA-Doppelstrangbrüchen und die Wiederaufnahme der Replikation zuständig. Ein als AddAB bezeichnetes Helikase-/Nuklease-System ist in *Bacillus subtilis* für die Reparatur von Doppelstrangbrüchen zuständig und weist zum RecBCD-System funktionell homologe Proteine auf. In ersten Integrationsstudien mit genomischer DNA aus BL21 AI und dem pCR001 Wildtyp-Plasmid konnte bereits eine signifikante Reduzierung der Aufnahme von neuen Spacern in RecB-defizienten *E. coli* Stämmen gezeigt werden. Eine Komplementation mit dem redundanten AddAB Protein führte zum Aufheben der Reduzierung (unveröffentlichte Daten, Zihni Arslan). Diese Tatsache unterstützt die Hypothese, dass eine zusätzliche Beteiligung der Proteine AddAB oder RecB möglicherweise zur Aktivierung des bisher inaktiven heterologen Systems in *E. coli* führen könnten. Allerdings müsste diese Hypothese durch zukünftige Studien näher untersucht werden.

### 3.4 Expression und Aufreinigung der E. coli cascade Gene

In einem zweiten Teil dieser Arbeit sollten die den Cascade-Komplex bildenden fünf Cas-Proteine Cse1 (CasA), Cse2 (CasB), Cas7 (Casc), Cas5 (CasD) und Cas6e (CasE) zunächst analytisch exprimiert und anschließend präparativ isoliert werden, um sie mittels *in vitro* Bindestudien näher charakterisieren zu können.

Die Expression von Genen bezeichnet den Vorgang bei dem die genetische Information umgesetzt und für die Zelle nutzbar gemacht wird. Es ist der intrazelluläre Weg vom Gen zum Genprodukt. Dabei sind verschiedene Stufen der Genexpression, wie z.B. die Transkription, RNA-Prozessierung, Translation oder die Proteinreifung zu unterscheiden.

#### 3.4.1 Analytische Expression der cascade Gene

Für die Überexpression und Aufreinigung der *E. coli cascade* Gene wurde das pET-52(b)+ Expressionssystem der Firma Novagen (4.2.2) verwendet, dass sich wegen seiner sehr hohen Transkriptionseffizienz hervorragend für die Klonierung und Expression rekombinanter Proteine aus *E. coli* eignet. Das Expressionssystem des Vektors basiert auf dem Bakteriophagen T7, der sehr starke Promotoren besitzt, die von der Phagen-eigenen T7-RNA-Polymerase erkannt werden. Nach Klonierung der Zielgene in die '*MCS*' des Vektors stehen diese unter Kontrolle des durch IPTG induzierbaren T7 Promotors. Darüber hinaus ist der Vektor mit zwei Affinitätstags ausgestattet, einem N-terminal gelegenen StrepII-Tag und einem C-terminalen His-Tag, die für die duale Aufreinigung von Volllängen-Fusionsproteinen geeignet

sind. Nach Transformation der klonierten pET-Konstrukte in BL21 (DE3) Zellen wurden analytische Expressionsversuche mit IPTG durchgeführt. Die anschließende Trennung über eine SDS-Gel PAGE ließ unterschiedlich hohe Konzentrationen der einzelnen Cas-Proteine erkennen (siehe Abb. 2.17). Wo hingegen bei CasA, CasB und CasC eine starke Expression nach Induktion und Wachstum der Zellen über Nacht zu erkennen war, zeigte das CasE-Protein im Vergleich schwächere Proteinbanden. Eine Expression des CasD-Proteins konnte nicht gezeigt werden. Die Expression der einzelnen Gene scheint somit unterschiedlich stark zu erfolgen und deutet auf mögliche Stabilitätsunterschiede hin. In Studien konnte gezeigt werden, dass die Expression von Genen vor allem auf Translationsebene kontrolliert wird. Dabei ist die Translationseffizienz universeller Bestandteil der Genkontrolle in Bakterien und erlaubt eine unterschiedliche Expression von Operon-codierten Proteinkomplexen mit mehreren Untereinheiten und ungleicher Stöchiometrie (Quax, Wolf et al. 2013). Das die Expression der cascade-Gene in BL21 (DE3) unterschiedlich reguliert ist konnte bereits in vorangehenden Expressionsstudien mit dem Plasmid pWUR400 gezeigt werden (Spreiz 2011). Die analytische Überexpression des pWUR400 Plasmides in wildtypischen und  $\Delta hns^2$  Zellen zeigte, dass nicht alle der in dem Cascade-Komplex enthaltenen fünf Cas-Proteine mit gleicher Intensität exprimiert werden und im Vergleich zum Wildtypstamm die Überproduktion im  $\Delta hns^{-}$ -Stamm stärker war. Diese unterschiedlich starke Expression entspricht der ungleichen stöchiometrischen Verteilung der einzelnen Protein-Untereinheiten innerhalb des Cascade-Komplexes mit einer Anordnung von CasA<sub>1</sub>B<sub>2</sub>C<sub>6</sub>D<sub>1</sub>E<sub>1</sub> (Jore 2010, Quax, Wolf et al. 2013). Ob die Anwesenheit des Holo-Cascade-Komplexes einen stabilisierenden Effekt auf die Expression einzelner Proteinuntereinheiten des Komplexes hat wurde daraufhin in weiterführenden Expressionsversuchen nach Transformation der klonierten pET-Plasmide in BL21 (DE3)/pWUR400 Zellen getestet. Die Ergebnisse dieser analytischen Genexpressionen bestätigten den stabilisierenden Effekt des Cascade-Plasmids auf die Expression einzelner Cas-Proteine (Daten nicht gezeigt). Im Gegensatz zu den einzeln exprimierten Cas-Proteinen, ließen die Bandenintensitäten in Gegenwart des Cascade-Plasmids im direkten Vergleich eine deutlich stärkere Expression erkennen.

#### 3.4.2 Präparative Aufreinigung der Cascade-Proteine

Nach Klonierung in das pET-System entstehen durch mRNA Synthese und anschließende Translation Fusionsproteine aus dem N-terminalen StrepII-Tag und einem der fünf Cas-Proteine. Die Affinität des StrepII-Tag Proteins zu seinem Substrat Desthiobiotin kann man sich anschließend in einem Affinitätstag-basiertem Reinigungsverfahren zu nutze machen. Für die Aufreinigung wurden stets Zellen verwendet, die frisch ausgestrichen worden waren, da bei älteren Zellen eine reduzierte Induktion der Proteine beobachtet wurde (Daten nicht gezeigt). Die Induktion der Expression der StrepII-Fusionsproteine erfolgte bei einer OD<sub>600</sub> von 0,5-0,6 mit niedrigen IPTG-Konzentrationen (0,1-0,2 mM). Anschließend wurden die Wachstumstemperaturen der Zellen von 37°C auf Raumtemperatur reduziert, um eine Denaturierung und den Abbau der Proteine zu minimieren. Die Ernte der Zellen erfolgte je nach Protein nach 5 Stunden oder Wachstum über Nacht. Die so enthaltenen Proteine, wurden über eine mit Strep-Tactin Resin als Trägermaterial gepackte Glassäule oder der Äkta Prime Plus Anlage und einer StrepTrap<sup>™</sup> Säule aufgereinigt. Die SDS-Gelbilder der präparativen Aufreinigungen zeigten häufig, dass ein Großteil der Proteine im unlöslichen Pellet verblieben war (siehe Abb. 2.20, 2.21 und 2.22). Sowohl die geringen Konzentrationen an IPTG, als auch die Senkung der Expressionstemperatur nach Induktion führten nicht zu einer erhöhten Löslichkeit des Proteins, sondern nur zu einer Verringerung der Gesamtexpression. Dennoch konnten ausreichende Ausbeuten der Proteine CasA, CasB und CasC für die nachfolgenden *in vitro* Analysen erzielt werden.

Die Expression des StrepII-CasE Fusionsproteins erfolgte nur in sehr geringen Mengen. Es wurde deutlich, dass der Großteil der CasE-Proteine nicht im proteinhaltigen Überstand, sondern im Pellet als unlösliche inclusion bodies gewonnen werden konnte. Die Effizienz der Expression von StrepII-Fusionsproteinen hängt dabei häufig von der Faltung und Stabilität des Zielproteins ab. Die Bildung der inclusion bodies kann zum einen durch eine Verringerung der Gesamtexpression verhindert werden und zum anderen durch Unterstützung der korrekten Faltung des Proteins, beispielsweise durch eine Coexpression mit Chaperonen (Woestenenk, Hammarstrom et al. 2004). Wie schon zuvor zeigte eine Senkung der Expressionstemperatur zur Reduzierung der in vivo Aggregation von CasE und eine geringe Induktorkonzentration keine Erhöhung des Anteils an löslichem Protein. Es konnte nur eine Verringerung der Gesamtexpression beobachtet werden. Eine Umklonierung in den Vektor pET100-His6-TEV mit C-terminalem His6-Tag und Aufreinigung über eine Ni-NTA-Agarose Säule zeigte eine nur schwache Affinität des His<sub>6</sub>-CasE Proteins an das Trägermaterial und brachte ebenfalls keine Erhöhung der Ausbeute und Löslichkeit des Proteins. Eine andere Möglichkeit unlösliche Proteine zugänglich zu machen und eine Rückfaltung aus den inclusion bodies zu erreichen, bietet die Behandlung mit Guanidiniumhydrochlorid (GHCL). Dabei wurde versucht, die unlöslichen Proteine aus den Pellets der Proteinaufschlüsse schrittweise mit Hilfe von unterschiedlichen GHCL-Konzentrationen in einem Denaturierungund Renaturierungsschritt in Lösung zu halten. Im Fall des CasE-Proteins führte die Renaturierung zu einem erneuten ausfallen des Proteins durch Fehlrückfaltung, wodurch es wieder unlöslich vorlag. Es wurde deutlich, dass für eine Aufreinigung von CasE möglicherweise eine Fusion mit einem anderen gut löslichen Protein wie z.B. dem Maltosebindeprotein notwendig ist. So könnte zusätzlich die Löslichkeit und Stabilität von CasE erhöht werden. Eine weitere Möglichkeit für eine bessere Zugänglichkeit der exprimierten Proteine wäre ein erleichterter Zellaufschluss der mit Hilfe des Zellstammes E. coli BL21 (DE3)/pLysS erreicht wird. Neben geringen Mengen T7 Lysozym das zur Spaltung der Zellwand genutzt wird, unterdrückt das Plasmid pLysS auch die basale Expression des T7-Promotors durch Bindung an die T7 RNA-Polymerase. So wird die Transkription des Zielproteins bis zur Expression durch IPTG unterdrückt, was wiederum zu einer besseren Löslichkeit des Proteins führen könnte.

Eine präparative Aufreinigung des CasD StrepII-Fusionsproteins aus E. coli wurde im Gegensatz zu allen anderen Cas-Proteinen nicht erreicht und passt zu der bereits schon zuvor beobachteten geringen analytischen Expression von CasD (siehe Abb. 2.17). Dass eine Aufreinigung von E. coli CasD möglich ist, zeigt die präparative Aufreinigung von CasC/Cascade, in der durch Co-Expression und immobilisierte CasC-Affinitätschromatographie CasD als Bestandteil des Cascade-Komplexes mit (siehe Abb. 2.18). Die aufgereinigt wurde Gegenwart des gesamten Cascade-Komplexes deutet auf einen stabilisierenden Effekt auf die Expression des einzelnen CasD-Proteins hin und konnte schon zuvor in Expressionsversuchen beobachtet werden (Daten nicht gezeigt). Der Grund für die niedrige Expression von CasD ist auf der Ebene der Proteinstabilität zu suchen und resultiert vermutlich aus einer Instabilität durch proteolytischen Abbau.

### 3.5 Interaktion von CasA, CasB und CasC mit Nukleinsäuren

### 3.5.1 Bindung an Einzel- und Doppelstrang-DNA

Nachdem drei der fünf Cascade-Proteine durch Affinitätschromatographie erfolgreich isoliert werden konnten, wurde in nächsten Versuchen nach potentiellen Bindepartner für eine nähere Charakterisierung der Proteine CasA, CasB und CasC gesucht. Für die Interaktion mit DNA wurde neben einem aus BL21 (DE3) isolierten Leader-Fragment, ein 58 nt langes einzelsträngiges Oligonukleotid (LeaderI-Repeat1-58nt) als Interaktionspartner angeboten. Die Ergebnisse der Retardierungsanalysen zeigten, dass E. coli CasA mit einer hohen Affinität an kurze doppelsträngige DNA-Fragmente bindet, aber eine im Vergleich schwache Bindung an das radioaktiv DNA-Oligonukleotid aufweist. Bereits mit Veröffentlichung markierte der Kristallstruktur des CasA-Proteins von T. thermophilus wurde die Rolle von CasA als essentieller Bestandteil für eine Bindung des Cascade-Komplexes an Ziel-DNA Bereiche aufgedeckt. Eine alleinige Bindung des CasA-Proteins konnte in diesen Studien nicht gezeigt werden (Mulepati, Orr et al. 2012). Eine 2012 gezeigte sequenzunspezifische Bindung von isolierten CasB-Proteinen aus den Bakterienstämmen E. coli, Thermobifida fusca und T. thermophilus an doppelsträngige DNA konnte mit den in dieser Arbeit durchgeführten Bindestudien und dem Leader-BL21 Fragment nicht bestätigt werden (Nam, Huang et al. 2012). Das CasB-Protein zeigte keine Bindung an das Leader-BL21 Fragment, ließ aber wie bei CasA eine verringerte Bindung an das einzelsträngige DNA-Oligonukleotid erkennen. Für das CasC-Protein konnte eine Bindung an keines der beiden DNA-Typen gezeigt werden.

Das für die Bindestudien genutzte Leader-Fragment umfasst den upstream Bereich des CRISPR II Array aus dem Genom von BL21 (DE3). Durch in vivo Untersuchungen konnte bereits ein aktiver Promotor (Pcr2) innerhalb der Leaderregion des CRISPR II Arrays identifiziert werden und anhand von Retardierungsanalysen eine Bindung von H-NS und RNAP an diese Region gezeigt werden (Spreiz 2011). Bei diesen Untersuchungen wurde analysiert, ob die Leaderregion auch als Plattform für die Bindung aufgreinigter Cas-Proteine aus E. coli fungiert. Insgesamt ließen die Bindeanlaysen eine nur geringe Bindung an die konservierten Bereiche des Leader-Fragmentes erkennen. Die Ergebnisse in Kapitel 2.5.7.1 zeigten eine Bindung von CasA an ein verkürztes Leader-Fragment, kleiner als das eigentlich 322 bp EcoRI/HincII geschnittene Fragment (siehe Abb. 2.24). Dass es sich dabei nicht um aufgeschmolzene Einzelstrangbereiche handelt, die während der Klenow-Markierung des Fragmentes entstanden sind, wurde durch einen Denaturierungs- und Renaturierungsschritt mit 1 x KGlu80-Puffer ausgeschlossen. Die Ursache dieses verkürzten Leader-Fragmentes könnte eine auftretende 'Star-Aktivität' des für die Generierung der 5'-Überhänge genutzten EcoRI Restriktionsenzyms sein. Eine denaturierende Auftrennung über ein Harnstoffgel verifizierte das Vorhandensein einer zusätzlichen Bande bei 200 bp (siehe Abb. 6.12 im Anhang). Dass es sich bei der Bindung von CasA nicht um eine sequenzspezifische Bindung an bevorzugte Bereiche des Leader-Fragmentes aus BL21 (DE3) handelt, konnte weiterhin anhand einer Retardierungsanalysen mit demselben Leader-Fragment, hydrolysiert mit BamHI/Sacl, gezeigt werden. Die Ergebnisse dieser Analysen in Kapitel 2.5.7.2 zeigen keine Bindung und können als Indizien für eine Bindung von CasA an kurze DNA-Fragmente kleiner als 298 bp genommen werden.

### 3.5.2 Ist RNA ein potentieller Bindungspartner?

Die Untersuchungen einer Interaktion mit verschiedenen RNAs in Kapitel 2.5.7.3 zeigten eine starke Bindung des isolierten CasA Proteins an die durch *in vitro* Synthese transkribierte precursor-crRNA aus *E. coli*. Dabei wurde eine hohe Affinität gegenüber dieser Art von CRISPR-RNA mit einer Dissoziationskonstante im mikromolaren Bereich festgestellt. Bereits nach Zugabe von 0,5 µM CasA-Protein kann ein Shift der freien RNA beobachtet werden, der mit steigenden CasA Konzentrationen zunimmt und bei 4 µM CasA gesättigt ist (siehe Abb. 2.27). Die Titrationsreihen der Proteine CasB und CasC ließen im Vergleich keine derartige Bindung erkennen. Das CasA Protein als größte Untereinheit des Cascade-Komplexes ist ein wichtiger Bestandteil für die CRISPR/Cas-vermittelte Abwehr. Bindestudien zeigten die Notwendigkeit von CasA für die Bindung des Cascade-Komplexes an die Ziel-DNA (Mulepati, Orr et al. 2012, Sashital, Wiedenheft et al. 2012). In Elektronenmikroskopischen Aufnahmen (cryo-EM) des *E. coli* Cascade-Komplexes wurde eine Interaktion der großen Untereinheit von CasA mit der Untereinheit des CasB-Proteins gezeigt (Wiedenheft, Lander et al. 2011). Diese Beobachtung führte zu

der Annahme, dass eine gleichzeitige Inkubation beider Proteine mit der pre-crRNA eventuell zu einem Supershift führen könnte. Das Ergebnis dieser Gelverzögerungsanlyse ließ jedoch keine Bindung an die pre-crRNA erkennen.

Die durchgeführten Bindestudien zeigen, dass E. coli CasA mit einer hohen Affinität bevorzugt an pre-crRNA bindet. Eine Bindung an die 6S RNA, einer kleinen regulatorischen RNA aus E. coli oder einem kurzen pRNA Oligonukleotid konnte im Vergleich nicht gezeigt werden. Die Bindung von CasA an die pre-crRNA ist somit nicht auf eine unspezifische Interaktion generell mit RNA zurück zuführen, sondern scheint sequenzspezifisch zu erfolgen. Eine sequenz-unspezifische Bindung von CasB an einzelsträngige RNA wie sie schon 2012 gezeigt wurde, konnte mit den in dieser Arbeit durchgeführten Bindestudien nicht übereinstimmend festgestellt werden (Nam, Huang et al. 2012). Eine Bindung des CasC-Proteins, welches mit sechs Kopien im Cascade-Komplex vertreten ist und das Rückgrat für die Assoziierung der prozessierten crRNA bildet, ist weder an die pre-crRNA, noch an die 6S RNA oder die pRNA zu erkennen. Die Sequenzspezifität des CasA-Proteins für die verwendete pre-crRNA weist auf eine Rolle von CasA während der Prozessierungsphase hin. In E. coli erfolgt die Prozessierung der Vorläufer-RNA zu den reifen kurzen crRNAs endonukleolytisch durch das Cas6e-Protein und Ausbildung einer Loop-Struktur, die als Erkennungssequenz für die Spaltung dient. Erst kürzlich konnte gezeigt werden, dass die große CasA Untereinheit des Cascade-Komplexes eine Rolle als essentieller Partner des Cas3-Proteins bei der Erkennung der Ziel-DNA und der korrekten Positionierung von Cas3 zum PAM-Motiv spielt (Hochstrasser, Taylor et al. 2014).

### 3.6 Ausblick

Die hier durchgeführten Experimente ermöglichen einen Einblick in den CRISPR-basierten Adaptationsmechanismus von *E. coli* und die Charakterisierung einzelner Proteinkomponenten des Cascade-Komplexes. Sie zeigen aber auch, dass einiges über die Funktion und Rolle der an der Integration beteiligten Proteine und Sequenzbereiche bis jetzt noch nicht ganz verstanden ist und für ein besseres Verständnis weitere Versuche notwendig sind. Insbesondere das Bindungsverhalten der isolierten Cas-Proteine CasA, CasB und CasC müsste in weiteren Untersuchungen näher charakterisiert werden.

Der Aufbau eines Plasmid-basierten Integrations-Assays stand im Vordergrund dieser Arbeit und bot nach einer erfolgreichen Etablierung die Grundlage für weiterführende Analysen. Mit Hilfe von *single-colony* PCR-Reaktionen konnte *in vivo* nicht nur der Einbau von neuen Spacersequenzen in den synthetischen CRISPR-Array gezeigt, sondern diese durch Sequenzierung auch identifiziert und analysiert werden. Nach Einführung von gezielten Punkmutationen innerhalb des für die Integration essentiellen ersten Repeats, konnten erste Hinweise in Bezug auf eine DNA-sequenzspezifische Spaltung während des Einbaus neuer Spacer aufgedeckt werden. In allen mutierten Plasmiden wurde nach Mutagenese eine signifikante Reduzierung der Spacerintegration gezeigt und gleichzeitig eine rein strukturabhängige Integration ausgeschlossen. Für eine basengenaue Bestimmung des für die Integration notwendigen Sequenzbereiches innerhalb des ersten Repeats, sollten künftig weitere Modifizierungen eingeführt werden, um darüber hinaus das Modell zum Adaptationsmechanismus weiter aufklären zu können. Die bisherige Hypothese eines sequenzspezifisch-, versetzten Schnittes an beiden Strängen des ersten Repeats wurde anhand von Southern Blot Analysen bestätigt und dabei die Existenz von Zwischenstufen (DNA-Intermediaten), die während der Integration entstanden sind detektiert. Zudem wurde gekoppelter ein Spaltungs-Ligationsmechanismus verifiziert, dessen Spaltungsereignis mit einer gleichzeitigen Verknüpfung der freien Enden von Repeat und Spacer einhergeht. Als weitere Kontrolle des vorgeschlagenen Mechanismus-Modells könnten Southern Blot Analysen mit bereits klonierten Leader-Mutanten durchgeführt werden, die eine Reduzierung des Spacereinbau in Integrationsstudien zeigten (Popowitsch 2013). Weiterhin sollte für den Aufbau eines aktiven heterologen Integrationssystems in E. coli BL21 AI mit Integrations-Komponenten von S. agalactiae nachgeprüft werden, ob eine Aktivität von Cas9, oder der tracrRNA die Funktionalität des Integrationssystems positiv beeinflussen. Ein weiterer interessanter Punkt stellt die Bereitstellung der Proteine RecB oder AddAB in *E. coli* dar.

Von besonderer Bedeutung ist auch die Untersuchung der für die Immunabwehrreaktion essentiellen Cas-Proteine Cas1 und Cas2. Diese sind innerhalb der verschiedenen CRISPR-Systeme hoch konserviert und ließen in ersten *in vitro* Analysen erste Hinweise einer gemeinsamen Interaktion vermuten (Schwabroch 2012). Eine aktuelle Studie bestätigt die Bildung eines für die Spacerintegration wichtigen Cas1/Cas2-Komplexes (Nuñez, Kranzusch et al. 2014). Der Einfluss und die Aktivität dieses Komplexes auf die Integration neuer Spacer sollte durch zukünftige biochemische Untersuchungen *in vitro* näher analysiert werden.

Die Expression und affinitätschromatographische Aufreinigung von Protein-Untereinheiten des *E. coli* Cascade-Komplexes war für eine Charakterisierung einzelner Cas-Proteine vorgesehen. Dabei konnte für das Cse1-Protein *in vitro* eine bisher noch nicht gezeigte Bindung an einzelsträngige DNA und an eine CRISPR-RNA nachgewiesen werden, deren Bedeutung noch völlig offen ist. Eine Expression von CasD (Cas5) ohne die Gegenwart des Cascade-Plasmids konnte in dieser Arbeit nicht erzielt werden und verlangt zukünftige Untersuchungen. Die Problematik der geringen Ausbeute und Bildung von *inclusion bodies* des CasE-Proteins könnte versucht werden durch Co-Faktoren oder einer Co-Expression mit einem anderen gut löslichen Protein zu umgehen. Hier gibt es noch viele offene Fragen und damit Bedarf für weitere Untersuchungen.

Zusammenfassend kann festgehalten werden, dass diese Arbeit neue, bisher noch nicht gezeigte Einblicke in den Adaptationsprozess von *E. coli* aufdecken konnte, aber zum jetzigen Zeitpunkt nicht alle Sachverhalte vollständig geklärt und analysiert wurden. Die bisherigen Ergebnisse bereiten einen Weg für weiterführende Untersuchungen, in denen die postulierte Integraseaktivität der Cas1/Cas2 Proteine im gekoppelten Spaltungs-Ligationsmechanismus durch *in vitro* Studien näher untersucht werden sollte.

### 4 Material

### 4.1 Allgemeines

Alle Chemikalien die für diese Arbeit verwendet wurden besaßen, sofern nicht anders vermerkt den Reinheitsgrad *pro analysis*. Für Puffer und Lösungen wurde ausschließlich hochreines Milli-Q-Wasser aus der hauseigenen Anlage (mit nachgeschaltetem water purification system EPA Est. 41237-MA-1 Millipore GmbH, Neu Isenburg) verwendet, dass nachfolgend als *Aqua dest*. bezeichnet wird.

### 4.2 Verwendete Bakterienstämme und Plasmide

### 4.2.1 Escherichia coli Stämme

**Tabelle 4.1: Übersicht über die verwendeten Stämme aus** *Escherichia coli* In der Tabelle sind die Genotypen und die Referenzen der jeweiligen Stämme zusammengefasst.

Stamm	Genotyp	Referenz/Herkunft
BL21 (DE3)	(F <sup>-</sup> , r <sup>-</sup> <sub>B</sub> m <sup>-</sup> <sub>B</sub> ), rif S, λ-lysogen mit T7	Wood <i>et al.</i> J. Mol. Biol.
	Gen1 (Polymerase) durch Met-	<u>16</u> : 118-133 (1966)
	Transduktion des Stammes B834	
BL21 AI	F-, ompT, hsdSB (r-B mB), gal, dcm,	Invitrogen, USA
	araB::T7RNAP-tetA	
HB 101	F <sup>-</sup> , <i>rec</i> A13, <i>hsd</i> S20 (r <sup>-</sup> <sub>B</sub> m <sup>-</sup> <sub>B</sub> ), <i>ara</i> -14,	Boyer and Roulland-
	proA2, layY1, galK2, rpsL20 (Str <sup>r</sup> ), xyl-	Dussoix, 1969
	5, mtl-1, supE44, mcrB	
TOP10	F <sup>-</sup> mcrA Δ(mrr-hsdRMS-mcrBC)	Invitrogen, USA
	φ80 <i>lac</i> ZΔM15 Δ <i>lac</i> X74 recA1 araD139	
	$\Delta(ara-leu)$ 7697 galU galK rpsL (Str <sup>R</sup> )	
	endA1 nupG λ-	
XL-1 blue	recA1, lac <sup>-</sup> , endA1, gyr A96, thi, hsd	Stratagene
	R17, supE44, relA1, [F': pro AB, lac Iq,	
	<i>lac</i> Z∆M15, Tn10]	
DH5a	SupE44, Δ <i>lac</i> U169, (φ80 <i>lac</i> ZΔM15),	Hanahan, D. (1983)
	hsdR17, (rK <sup>-</sup> mK <sup>+</sup> ), recA1, endA1,	J. Mol. Biol. <u>166</u> :557 ff.
	gyrA96, thi-1, relA1 Δ(arg F-lac ZYA)	

### 4.2.2 Plasmide

Tabelle 4.2: Übersicht über die erstellten und verwendeten Plasmide.

Die Tabelle zeigt die Beschreibungen und Referenzen der in dieser Arbeit verwendeten Plasmide.

Plasmid	Beschreibung	Referenz/Herkunft
pUC18	Derivat des pBR322-M13mp18	Yanisch-Peron et al.,
	(2686 Bp), high-copy Plasmid, mit	1985
	'multiple cloning site'(MCS) Poly-	
	linker, Gen für Ampicillinresistenz	
pACYC Duet-1	Low-Copy-Expressionsvektor mit	Novagen
	T7 RNAP (4008 bp), Cam <sup>R</sup> , <i>lacl</i>	
	Repressor und zwei Multiple-	
	Cloning-Sites (MCS)	
pWUR400	pCDF-1b Derivat mit T7 RNAP und	(Brouns, Jore et al.
	Str <sup>R</sup> ; enthält die <i>cas</i> -Gene <i>cas</i> -	2008)
	casB-casC-casD-casE	
pET-52b(+)	pET-52 Derivat mit MCS die unter	Novagen
	Kontrolle des T7 Promotors und	
	T7 Terminators steht (5227 bp),	
	trägt N-terminal StrepII-Tag und C-	
	terminal His-Tag, Amp <sup>R</sup>	
pET-CasA	pET-52b(+) Derivat, trägt E. coli	Diese Arbeit
	CasA mit N-terminalem StrepII-Tag	
	in MCS unter Kontrolle des T7	
	Promotors und T7 Terminators	
	(6703 bp), Amp <sup>R</sup>	
pET-CasB	pET-52b(+) Derivat, trägt E. coli	Diese Arbeit
	CasB mit N-terminalem StrepII-Tag	
	in MCS unter Kontrolle des T7	
	Promotors und T7 Terminators	
	(5686 bp), Amp <sup>R</sup>	
pET-CasC	pET-52b(+) Derivat, trägt <i>E. coli</i>	Diese Arbeit
	CasC mit N-terminalem StrepII-Tag	
	in MCS unter Kontrolle des T7	
	Promotors und T7 Terminators	
	(6307 bp), Amp <sup>R</sup>	
pET-CasD	pET-52b(+) Derivat, trägt <i>E. coli</i>	Diese Arbeit
	CasD mit N-terminalem StrepII-Tag	
	in MCS unter Kontrolle des T7	
	Promotors und T7 Terminators	
	(5866 bp), Amp <sup>R</sup>	
pET-CasE	pET-52b(+) Derivat, trägt <i>E. coli</i>	Diese Arbeit
	<i>CasE</i> mit N-terminalem StrepII-Tag	

	in MCS unter Kontrolle des T7	
	Promotors und T7 Terminators	
	(5803 bp), Amp <sup>R</sup>	
pCR001	pACYC-Duet-1 Derivat (Novagen),	(Arslan, Wurm et al.
	trägt <i>cas1-cas2</i> aus genomischen	2013)
	MG1655 Hintergrund	,
pLS1(+)	pACYC-Duet-1 Derivat mit	(Arslan, Stratmann et
	Spacersequenz (GCGACCACACCC	al. 2013)
	GTCCTGTGGATCCTCTACGC)	
	zwischen den ersten beiden	
	Repeats des pLS0 Plasmides	
pZA-T7S1+	pUC18-T7 Derivat, trägt <i>E. coli</i>	Promotion Zihni
p=======	CRISPR R1-R4 vor <i>kan</i> -Gen	Arslan 2014
	geschaltet	
pUC18-Kan	pUC18 Derivat. Substitution von	(Arslan, Stratmann et
	<i>bla</i> mit <i>neo</i> von pKD4 zwischen	al. 2013)
	SsnI-Dral	
nCR003 WT	pACYC-Duet-1 Derivat mit cas1-	Diese Arbeit
	cas2 aus genomischer DNA von	
	MG1655 (2. MCS) mit CRISPR	
	Array R1-R4 und Spacer gegen	
	Tet-Gen aus nLS1-TET+. Cam <sup>R</sup>	
pCR003 RM1	pCR003 Derivat. trägt Mutation	Diese Arbeit
	innerhalb des ersten Repeats an	
	Position 242 bis 245. Austausch	
	von CCCC zu GGTT	
pCR003 RM1-R	pCR003 RM1 Derivat. trägt eine	Diese Arbeit
r	Mutation innerhalb des ersten	
	Repeats an Position 254 bis 257.	
	Austausch von GGGG zu AACC.	
	Rückmutation zur WT Situation	
pCR003 RM2	pCR003 RM1-R Derivat, trägt	Diese Arbeit
r	Mutationen innerhalb des ersten	
	Repeats an Position 244 bis 245.	
	Austausch von TT zu GG und an	
	Position 254 bis 255. Austausch	
	von AA zu CC	
pSagCas1Cas2Csn2	pACYC-Duet-1 Derivat. trägt cas1-	Promotion Zihni
Fendense	<i>cas2-csn2</i> Gene aus genomischen	Arslan 2014
	Hintergrund von <i>S. agalactiae</i>	
pSagCas1Cas2K132	pACYC-Duet-1 Derivat. trägt <i>cas1-</i>	Promotion Zihni
1	<i>cas2</i> sowie eine mutierte Form von	Arslan 2014
	<i>csn2</i> (K132) aus genomischen	
	Hintorgrund von S agalactiao	

pCR004	pSagCas1Cas2Csn2 Derivat mit	Diese Arbeit
	CRISPR-Array (337 bp) aus	
	S agalactiae	
pCR005	pSagCas1Cas2K132 Derivat mit	Diese Arbeit
	CRISPR-Array (337 bp) aus	
	S agalactiae	
pACYC-SagCas1	pACYC-Duet-1 Derivat, trägt cas1-	Diese Arbeit
Cas2	cas2 von S. agalactiae aus dem	
	pCR005 Plasmid amplifiziert	
pCR010	pET-52(b)+ Derivat, trägt cas1-	Diese Arbeit
	cas2 Gene von S. agalactiae aus	
	pACYC-SagCas1Cas2 Plasmid	
	mittels PCR amplifiziert	
pCR0012	pACYC-SagCas1Cas2 Plasmid mit	Diese Arbeit
	Einbau von 8 Basen aus <i>E. coli</i>	
	zwischen den <i>cas1-cas2</i> Genen	
pCR014	pACYC-Duet1 Derivat, trägt	Diese Arbeit
	CRISPR-Array von <i>S. agalactiae</i>	
pCR014+Csn2	pACYC-Duet1 Derivat, trägt	Diese Arbeit
	CRISPR-Array und <i>csn2</i> Gen von	
	S agalactiae	
pET-SagCas1Cas2	pET-52(b)+ Derivat, trägt cas1-	Diese Arbeit
Csn2	cas2-csn2 Gene von S. agalactiae	
pEcoSagCas1Cas2	pSagCas1Cas2Csn2 bzwK132	Diese Arbeit
Csn2	Plasmid mit synthetischem	
	CRISPR-Array E. coli	
pET-EcoCas1Cas2	pET-52(b)+ Derivat, trägt cas1-	Diese Arbeit
	<i>cas2</i> Gene von <i>E. coli</i>	

## 4.3 DNA-Fragmente

### 4.3.1 Fragmente für Retardierungen

Leader BL21-Fragment (E/H)	Hergestellt durch <i>Eco</i> RI und <i>HincII</i> Restriktion aus pUC18-LeaderBL21 anti, Länge des oberen Stranges 322 bp
Leader BL21-Fragment (B/S)	Hergestellt durch <i>BamHI</i> und <i>SacI</i> Restriktion aus pUC18-LeaderBL21 anti, Länge des unteren Stranges 289 bp
## 4.4 Nukleinsäuren

#### 4.4.1 Oligonukleotide

Alle in dieser Arbeit verwendeten Oligonukleotide wurden von der Firma Thermo Fisher Scientific GmbH aus Ulm bezogen.

T7 II-fw	5'-TAATACGACTCACTATAGG-3'	
	Diente als Primer für Sequenzierungen.	
pET100-T7-rv	5'-TAGTTATTGCTCAGCGGTGG-3'	
	Diente als Primer für Sequenzierungen.	
pUC18-MCSup	5'-GCGGATAACAATTTCACAC-3'	
	Bindet <i>upstream</i> der multiple cloning site (MCS) von pUC18 und seinen Derivaten und wurde für Sequenzierungen benutzt.	
pMA-rv	5'-CCCATGAGGCCCAGGTTG-3'	
	Bindet an die Position 409 bp bis 426 bp (17 bp) des pCR003 Plasmids und wurde als Primer für die Integration neuer Spacer oder Sequenzierungen eingesetzt.	
lacI_oligo	5'-GCGGTTGGGAATGTAATTCA-3'	
	Wurde als Primer für die Identifizierung positiver Klone des pCR003 WT Plasmids für <i>single-colony</i> PCR- Reaktionen eingesetzt.	
Duet-Down-Primer	5'-GATTATGCGGCCGTGTACAA-3'	
	Oligo bindet an die Position 577 bp bis 596 bp (20 bp) des pCR003 Plasmids und diente als Primer für Sequenzierungen, PCR-Reaktionen und <i>single-colony</i> PCR.	

Duet-UP2-Primer	5'-TTGTACACGGCCGCATAATC-3'	
	Oligo bindet an die Position 577 bp bis 596 bp (20 bp) des pCR003 Plasmids und diente als Primer für Sequenzierungen, PCR-Reaktionen und <i>single-colony</i> PCR.	
pET-Fwd	5'-GCGGCCGCAGAGCTC-3'	
	Oligo bindet an die Position 5047 bp bis 5061 bp (15 bp) des pET-52b(+) Vektors und wurde als Primer für die Linearisierung des Vektors mittels PCR-Reaktion eingesetzt.	
pET-Rev neu	5'-AGGATCCTGGTACCCGGGT-3'	
	Oligo bindet an die Position 5020 bp bis 5038 bp (19 bp) des pET-52b(+) Plasmids und wurde als reverse Primer für die Linearisierung des Vektors mittels PCR-Reaktion eingesetzt.	
CasA-fw	5'-GGACCCGGGATGGCTAATTTGCTTATTGATAACT GG-3'	
	Oligo bindet an die Position 71 bp bis 97 bp (25 bp) des pWUR400 Plasmids und diente als Primer für die Amplifizierung von CasA mit Einbau einer <i>Xma</i> I- Schnittstelle.	
CasA-rv	5'-TCAGCCATTTGATGGCCC-3'	
	Oligo bindet an die Position 1565 bp bis 1582 bp (27 bp) des pWUR400 Plasmids und diente als reverse Primer für die Amplifizierung von CasA.	
CasB-fw	5'-CCAGGATCCTATGGCTGATGAAATTGATGCA-3'	
	Oligo bindet an die Position 1575 bp bis 1595 bp (21 bp) des pWUR400 Plasmids und diente als Primer für die Amplifizierung von CasB mit Einbau einer <i>BamH</i> I- Schnittstelle.	
CasB-rv neu	5'-TTACGCATTTTTGTTTGTGGTCAA-3'	
	Oligo bindet an die Position 2034 bp bis 2057 bp (24 bp) des pWUR400 Plasmids und diente als reverse Primer für die Amplifizierung von CasB.	

4 Material	101
CasC-Fwd	5'-GGGTACCAGGATCCTATGTCTAACTTTATCAATAT TCATG-3'
	Oligo bindet an die Position 2070 bp bis 2094 bp (25 bp) des pWUR400 Plasmids und diente als Primer für die Amplifizierung von CasC mit 15 bp Verlängerung homolog zum pET-52b(+) Vektor.
CasC-Rev	5'-GAGCTCTGCGGCCGCTCACGCCTCGCCATTATTAC GA-3'
	Oligo bindet an die Position 3140 bp bis 3161 bp (22 bp) des pWUR400 Plasmids und diente als reverse Primer für die Amplifizierung von CasC mit 15 bp Verlängerung homolog zum pET-52b(+) Vektor.
CasD-fw	5'-GGACCCGGGATGAGATCTTATTTGATCTTGCGG-3'
	Oligo bindet an die Position 3164 bp bis 3187 bp (24 bp) des pWUR400 Plasmids und diente als Primer für die Amplifizierung von CasD mit Einbau einer <i>Xma</i> I- Schnittstelle.
CasD-rv	5'-TTACTGAGATACATCCATACCTCCTTT-3'
	Oligo bindet an die Position 3812 bp bis 3838 bp (27 bp) des pWUR400 Plasmids und diente als reverse Primer für die Amplifizierung von CasD.
CasE-fw	5'-CCAGGATCCTATGTATCTCAGTAAAGTCATCATTG CC-3'
	Oligo bindet an die Position 3825 bp bis 3851 bp (27 bp) des pWUR400 Plasmids und diente als Primer für die Amplifizierung von CasE mit Einbau einer <i>BamH</i> I- Schnittstelle.
CasE-rv	5'-TCACAGTGGAGCCAAAGATAGCAA-3'
	Oligo bindet an die Position 4401 bp bis 4424 bp (24 bp) des pWUR400 Plasmids und diente als reverse Primer für die Amplifizierung von CasE.
pMA-NcoI-fw	5'-GCGTACCATGGAAACAAAGAATTAGCTGATC-3'
	Forward PCR-Primer für die Amplifizierung des synthetischen CRISPR Arrays aus dem pLS1-Tet+ Plasmid.

4 Material	102
pMA-XhoI-rv	5'-GGTCCCTCGAGACAGCTATGACCATGTTAAT-3'
	Reverse PCR-Primer für die Amplifizierung des synthetischen CRISPR Arrays aus dem pLS1-Tet+ Plasmid.
Sc-BL21_Lead_fwNeU	5'-AGTTGGTAGATTGTGACTGGCT-3'
	Oligo bindet an die Position 1002850 bp bis 1002871 bp des BL21 AI Genoms und diente als Primer für PCR- Reaktionen, um zu kontrollieren, ob Spacer ins Genom eingebaut wurden.
Sc-BL21_2.2_rvNeU	5'-TATAGAGATCGTTTTTGGAATTTAC-3'
	Oligo bindet an die Position 1003005 bp bis 1003029 bp des BL21 AI Genoms. Reverse Primer für PCR-Reaktionen, um Spacereinbau ins Genom zu kontrollieren.
pCR001-S1-fw	5'-CGTAGAGGATCCACAGGACGG-3'
	Oligo bindet an die Position 276 bp bis 296 bp des pCR003 WT Plasmids und diente als Primer für PCR-Reaktionen, um Spacereinbau im Plasmid nachzuweisen.
pCR001-Leader-rv	5'-GCGTACCATGGAAACAAAGAATTA-3'
	Oligo zu bindet an die Position 121 bp bis 144 bp des pCR003 WT Plasmids und diente als Primer für PCR- Reaktionen, um Spacereinbau im Plasmid nachzuweisen.
pCR001-Blot-S1-fw	5'-GACCACACCCGTCCTGTGGATCCT-3'
	Ist gegen die Spacersequenz (S1) des pCR003 WT Plasmides und der Mutanten gerichtet und wurde für die Hybridisierung der Southern Blots zur Detektion von Cas1 spezifischen 'nicks' verwendet.
pCR001-Blot-C-S1-rv	5'-AGGATCCACAGGACGGGTGTGGTC-3'
	Komplementäres Oligo zu pCR001-Blot-S1-fw wurde für die Hybridisierung der Southern Blots zur Detektion von Cas1 spezifischen 'nicks' und für Sequenzierungs- reaktionen mittels Primer-Extension verwendet.
pCR001-Blot-EcoRI-fw	5'-AATTCGAGGCGTACCATGGA-3'
	Bindet innerhalb der Leadersequenz des pCR003 WT Plasmids und der Mutanten und wurde für die Hybridisierung der Southern Blots zur Detektion von Cas1 spezifischen 'nicks' verwendet.

pCR001-Blot-EcoRI-rv	5'-TCCATGGTACGCCTCGAATT-3'
	Komplementäres Oligo zu pCR001-Blot-EcoRI-fw das für die Hybridisierung der Southern Blots zur Detektion von Cas1 spezifischen 'nicks' verwendet wurde.
PE-Leader-pCR001-fw	5'-ATGGGAAAAAATGCTTTAAGAACA-3'
	Bindet an die Position 197 bp bis 220 bp des pCR003 WT Plasmids und wurde als Oligo für die Sequenzierungs- reaktion mittels Primer-Extension Analyse mit Sequenase 2.0 genutzt.
pCR1/Repeat_Mut.2-fw	5′-GTTTATCCCCGCTGGCGC <u>AACC</u> AACTCTCTAAAAGTAT AC-3′
	Mutagenese Primer, der die Basen AACC (unterstrichen) anstelle von GGGG in den ersten Repeat des pCR003 WT Plasmids mittels PCR-Reaktion einführt.
pCR1/Repeat_Mut.2-rv	5´-GTATACTTTTAGAGAGTT <u>GGTT</u> GCGCCAGCGGGGATAA AC-3´
	Komplementärer Mutagenese Primer zu pCR1/Repeat_ Mut.2-fw.
pCR1/Repeat_Mut.2_ rück-fw	5′-GTGTGGTCGCCGGTTTAT <u>GGTT</u> GCTGGCGCAACCAACTC T-3′
	Mutagenese Primer, der die Basen GGTT (unterstrichen) anstelle von CCCC in den ersten Repeat des pCR003 RM1 Plasmids mittels PCR-Reaktion einführt.
pCR1/Repeat_Mut.2_ rück-rv	5′-AGAGTTGGTTGCGCCAGC <u>AACC</u> ATAAACCGGCGACCACA C-3′
	Komplementärer Mutagenese Primer zu pCR1/Repeat_ Mut.2_rück-fw.
pCR1/Repeat_Mut.3-fw	5′- GCCGGTTTATGG <u>GG</u> GCTGGCGC <u>CC</u> CCAACTCTCTAA -3′
	Mutagenese Primer, der die Basen GG und CC (unterstrichen) anstelle von TT und AA in den ersten Repeat des pCR003 RM1-R Plasmids mittels PCR- Reaktion einführt.
pCR1/Repeat_Mut.3-rv	5′-TTAGAGAGTTGG <u>GG</u> GCGCCAGC <u>CC</u> CCATAAACCGGC-3′
	Komplementärer Mutagenese Primer zu pCR1/Repeat_

	Mut.3-fw.
pRNA (de novo RNA)	5'-AUCGGCUCAGGGGACUGGCC-3'
	RNA-Oligo das für Bindungsanalysen der CasA, CasB und CasC Proteine genutzt wurde (20 mer).
LeaderI-Rep1-58nt	5′-TGCTTTAAGAACAAATGTATACTTTTAGAGAGTTCCCCG CGCCAGCGGGGATAAACCG-3′
	DNA-Oligo das für Bindungsanalysen der CasA, CasB und CasC Proteine genutzt wurde.
VK8035_fw	5'-GATAAGGATTATTTCTTGATAGGCG-3'
	Primer für die Amplifizierung des CRISPR-Array aus genomischer DNA von <i>Streptococcus agalactiae</i> .
VK8035_rv	5'-AGGTTCGCGATGGTGTCATC-3'
	Reverse Primer für die Amplifizierung des CRISPR-Array aus genomischer DNA von <i>Streptococcus agalactiae</i> .
pCR004/005-S1	5'-CTTGCAAGCTATTTAGGTCAGTGT-3'
	Bindet an Position 298 bp bis 321 bp der Plasmide pCR004/-005 und diente als Primer für PCR-Reaktionen und um Spacereinbau in den Plasmiden zu kontrollieren.
pCR004/005-fw	5′-ATGGTATAATATTAGTAAAAGCACAG-3′
	Bindet an Position 151 bp bis 176 bp der Plasmide pCR004/-005 und wurde für die Integrationsstudien der Plasmide eingesetzt.
pCR004/005-S2	5'-AGCGTGATAACATATTTTCTTT-3'
	Bindet an Position 366 bp bis 387 bp der Plasmide pCR004/-005 und wurde ebenfalls für die Integrations- studien der Plasmide eingesetzt.
S.a.gCas1+2_NdeI-fw	5′-GAATACCATATGGCAGGTTGGCGAACTGTT-3′
	Diente als Primer für die Amplifizierung <i>cas1</i> und <i>cas2</i> Gene von <i>S. agalactiae</i> aus pCR005 Plasmid.

S.a.gCas1+2_PacI-rv	5′-CCGGCCTTAATTAATCTTGATCATAAGAATCTCCTA GG-3′
	Diente als reverse Primer für die Amplifizierung der <i>cas1</i> und <i>cas2</i> Gene von <i>S. agalactiae</i> aus pCR005 Plasmid und wurde für Sequenzierungen genutzt.
SagCas1-BamHI-fw	5'-CCAGGATCCTATGGCAGGTTGGCGAACTGTT-3'
	Diente als Primer für die Amplifizierung der Gene <i>cas1</i> sowie <i>cas1cas2csn2</i> aus <i>S. agalactiae</i> aus den Plasmiden pSagCas1Cas2Csn2, pACYC-SagCas1Cas2 und pACYC-SagCas1Cas2 mit <i>E. coli</i> Insertion.
SagCas1-rv	5'-TCATATCCTAAACTCCGGAATTCCATT-3'
	Reverse Primer für die Amplifizierung des <i>cas1</i> Gens aus <i>S. agalactiae</i> aus dem Plasmid pSagCas1Cas2Csn2.
SagCas2-fw	5'-ATGAGTTATCGGTATATGCGAATGATT-3'
	Oligo für die Amplifizierung von <i>cas2 (S. agalactiae</i> ) aus pACYC-SagCas1Cas2.
SagCas2-rv	5′-TTAATCTTGATCATAAGAATCTCCTAGGAAAAC-3′
	Reverse Oligo für die Amplifizierung von <i>cas2</i> ( <i>S. agalactiae</i> ) aus pSagCas1Cas2Csn2, pACYC-SagCas1Cas2 und pACYC-SagCas1Cas2 mit <i>E. coli</i> Insertion.
SagCas1_Insertion-rv	5'-TTCAGCTACTCCGGAATTCCATTCCCA-3'
	Oligo für die Amplifizierung von <i>cas1 (S. agalactiae</i> ) aus pACYC-SagCas1Cas2 mit 8 bp Insertion aus <i>E. coli</i> pCR001 Plasmid (8 Basen nach RBS).
SagCsn2-rv	5′-TTATACCATATTTTCGCCTATCAAGAAATAATCC-3′
	Reverse Oligo für die Amplifizierung von <i>cas1cas2csn2</i> aus <i>S. agalactiae</i> aus dem Plasmid pSagCas1Cas2Csn2.
pET-52_lin-fw	5'-GGTGCACTTGAAGTCCTCTTTCAG-3'
	Primer für die Entfernung des N-terminalen StrepII-Tag des pET-52b(+) Plasmids.

pET-52\_lin-rv

#### 5'-GCTTGCCATGGTATATCTCCTTCT-3'

Reverse Primer für die Entfernung des N-terminalen StrepII-Tag des pET-52b(+) Plasmids.

#### 4.4.2 Nukleotide

Adenosin-5'-triphosphat Cytidin-5'-triphosphat Guanosin-5'-triphosphat Uridin-5'-triphosphat (FPLC ultra pure) Adenosin-5'-triphosphat (FPLC ultra pure) Guanosin-5'-triphosphat (FPLC ultra pure) Uridin-5'-triphosphat (FPLC ultra pure) Adenosin-5'- $[\gamma^{-32}P]$ -triphosphat [5'-32P]pCp (cytidine-3',5'-bis-phosphate) 2'-Desoxyadenosin-5'-triphosphat 2'-Desoxycytidin-5'-triphosphat 2'-Desoxythymidin-5'-triphosphat 2'-Desoxythymidin-5'-triphosphat

Roche, Mannheim Roche, Mannheim Roche, Mannheim Roche, Mannheim Pharmacia, München Pharmacia, München Pharmacia, München Pharmacia, München Hartmann, Braunschweig Hartmann, Braunschweig Roche, Mannheim Roche, Mannheim Roche, Mannheim

#### 4.4.3 Molekulargewichtsmarker

1 Kb-Leiter
100 bp-Leiter
pUC19/MspI
PageRuler <sup>™</sup> Unstrained Protein Leiter
Unstrained Protein Molecular Weight Marker

New England Biolabs, USA New England Biolabs, USA New England Biolabs, USA Fermentas, St. Leon-Rot Fermentas, St. Leon-Rot

#### 4.5 Proteine

#### 4.5.1 Restriktionsendonukleasen

BamHI	MBI Fermentas, St. Leon-Rot
DpnI	MBI Fermentas, St. Leon-Rot
Ecl136II	MBI Fermentas, St. Leon-Rot
EcoRI	MBI Fermentas, St. Leon-Rot

HincII	MBI Fermentas, St. Leon-Rot
HindIII	MBI Fermentas, St. Leon-Rot
KpnI	MBI Fermentas, St. Leon-Rot
NcoI	MBI Fermentas, St. Leon-Rot
NdeI	MBI Fermentas, St. Leon-Rot
NotI	MBI Fermentas, St. Leon-Rot
PacI	MBI Fermentas, St. Leon-Rot
PvuI	MBI Fermentas, St. Leon-Rot
SacI	MBI Fermentas, St. Leon-Rot
Smal	New England Biolabs, USA
SspI	MBI Fermentas, St. Leon-Rot
Xbal	MBI Fermentas, St. Leon-Rot

#### 4.5.2 Polymerasen

Klenow-Fragment der DNA-Polymerase I	Promega, Madison, USA
Phusion High-Fidelity DNA-Polymerase	Finnzymes, Espoo, Finnland
T7-RNA-Polymerase	Promega, Madison, USA
Sequenase Version 2.0	USB, Braunschweig
<i>E. coli</i> RNA-Polymerase, Core- und Holo-Enzym	freundliche Gabe von R. Wurm,
	nach (Burgess and Jendrisak
	1075) vorändort nach

freundliche Gabe von R. Wurm, nach (Burgess and Jendrisak 1975) verändert nach (Gonzalez, Wiggs et al. 1977), charakterisiert nach (Chamberlin, Nierman et al. 1979)

#### 4.5.3 Enzyme und sonstige Proteine

DNaseI (RNase-frei) Lysozymchlorid Pyrophosphatase Proteinase K Ribonuklease A (RNase A) Rinderserumalbumin (BSA) Rinderserumalbumin (BSA), acetyliert RNasin RNase Inhibitor

T4-Polynukleotidkinase (PNK) T4-DNA-Ligase Roche, Mannheim Sigma, St. Louis, USA Sigma, St. Louis, USA Sigma, St. Louis, USA Sigma, St. Louis, USA New England Biolabs, USA New England Biolabs, USA MBI Fermentas, St. Leon-Rot Thermo Scientific, Rockford USA New England Biolabs, USA Roche, Mannheim T4 RNA LigaseThermo Scientific, Rockford<br/>USAQuick T4 DNA LigaseNew England Biolabs, USA

## 4.6 Puffer und Medien

## 4.6.1 Standardpuffer

Formamid-Probenpuffer:	95 % (v/v) deionisiertes Formamid 25 mM EDTA, pH 8,0 0,1 % (w/v) Bromphenolblau 0,1 % (w/v) Xylencyanol
1 x SDS-Laufpuffer:	28,8 g/l Glycin 6 g/l Tris 1 g/l SDS
5 x SDS-Puffer:	150 g Tris 720 g Glycin 25 g SDS ad 5 l mit <i>Aqua dest.</i>
4 x SDS-Probenpuffer	0,5 M Tris-HCl, pH 6,8 8 % SDS 80 % Glycerin 0,04 % (w/v) Bromphenolblau
50 x TAE-Puffer:	2 M Tris-Acetat, pH 7,5 50 mM EDTA, pH 8,0
5 x TAE-Probenpuffer:	0,025 % (w/v) Bromphenolblau 0,025 % (w/v) Xylencyanol 30 % (v/v) Glycerin in 5 x TAE-Puffer
10 x TBE-Puffer:	0,89 M Tris-Borat, pH 8,3 25 mM EDTA, pH 8,0
2 x TBE-Probenpuffer:	0,025 % (w/v) Bromphenolblau 0,025 % (w/v) Xylencyanol 30 % (v/v) Glycerin

in 2 x TBE
10 mM Tris-HCl, pH 8,0
1 mM EDTA, pH 8,0
0,5 M Tris-Acetat, pH 8,0
0,1 M Mg-Acetat
0,8 M Kaliumglutamat
10 mM Dithiothreitol (DTT)
1 mM EDTA, pH 8.0
10 μg/ml acetyliertes BSA
0,8 g NaCl
0,02g KCl
0,01 ml 1 M Na2HPO4
in 100 ml <i>Aqua dest.,</i> pH 7,4 mit HCL

### 4.6.2 Medien

YT-Medium:	8 g/l Trypton 5 g/l Hefe-Extrakt 5 g/l NaCl in <i>Aqua dest.</i> auf pH 7,4 einstellen und autoklavieren.
YT-Festmedium (YT-Agarplatten):	YT-Medium mit 15 g/l Agar
M9-Salz:	30 g/l Na <sub>2</sub> HPO <sub>4</sub> 15 g/l KH <sub>2</sub> PO <sub>4</sub> 5 g/l NH <sub>4</sub> Cl 2,5 g/l NaCl
M9-Minimal-Medium:	Für 100 ml: 20 ml 5x M9 Salz 2 mM MgSO <sub>4</sub> 0,1 mM CaCl <sub>2</sub> 0,4 % C-Quelle (20 % Glycerin) 0,2 % Casaminoacid (20 %) 5 μg/ml Thiamin

#### 4.7 Feinchemikalien

Acrylamid Agar Agarose Agarose, (ultrapure SeaKem) Amberlite MB-1 (Ionenaustauscher) Ammoniumacetat Ammoniumperoxodisulfat (APS) Ampicillin Borsäure **Bradford-Reagenz** Brij 35 Bromphenolblau Calciumchlorid Chloramphenicol Chloroform **Coomassie Brilliant Blue R 250** di-Natriumhydrogenphosphat (Na<sub>2</sub>HPO<sub>4</sub>) Diethyl-Pyrocarbonat (DEPC) Dithiothreitol (DTT) Essigsäure Ethanol abs. Ethidiumbromid Ethylendiamintetraacetat (EDTA) Formaldehyd Formamid Glucose Glycerin Glycin Harnstoff Hefe-Extrakt Heparin HEPES Isopropyl-ß-thiogalactosid (IPTG) Kaliumacetat Kaliumchlorid Kaliumdihydrogenphosphat (KH<sub>2</sub>PO<sub>4</sub>) Kaliumglutamat L-Arabinose Magnesiumchlorid Magnesiumsulfat β-Mercaptoethanol N,N'-Methylenbisacrylamid

Serva, Heidelberg Sigma, St. Louis, USA Sigma, St. Louis, USA Seakem, Hamburg ICN Biomedicals, USA Acros Organics, Geel, Belgien Merck, Darmstadt Sigma, St. Louis, USA Grüssing, Filsum Biorad Aldrich, Steinheim Merck, Darmstadt Acros Organics, Geel, Belgien Roche, Mannheim Roth, Karlsruhe Serva, Heidelberg VWR, Prolabo Alfa Aesar Fluka Chemie AG, Schweiz Roth, Karlsruhe Riedel-De-Haen, Seelze Roche, Mannheim Merck, Darmstadt Merck, Darmstadt Roth, Karlsruhe Merck, Darmstadt Merck, Darmstadt Roth, Karlsruhe J.T. Baker, Gross-Gerau Gibco BRL, Eggenstein Sigma, St. Louis, USA Roth, Karlsruhe Roche, Mannheim J.T Baker, Gross-Gerau J.T Baker, Gross-Gerau J.T. Baker, Gross-Gerau Aldrich Chemie GmbH Steinheim Roth. Karlsruhe Merck, Darmstadt Chimica, Geel Belgien Roth, Karlsruhe Serva, Heidelberg

N,N,N',N'-Tetramethylethylendiamin (TEMED) Merck, Darmstadt Natriumacetat Natriumchlorid Natriumdeoxycholat Natriumdodecylsulfat (SDS) Natriumhydroxid (Plätzchen) Natriumhydrogenphosphat (Na<sub>2</sub>HPO<sub>4</sub>) Natriumthiosulfat Phenol Phenylmethylsulfonylflourid (PMSF) Polyethylenglykol (PEG<sub>6000</sub>) Saccharose Salzsäure (HCl) Streptomycin Tris-(hydroxymethyl)-aminomethan (Tris) Trypton **Xylencyanol** 

Riedel-de-Haan, Seelze J.T. Baker, Gross-Gerau Roth, Karlsruhe Serva, Heidelberg Merck, Darmstadt Merck, Darmstadt Merck, Darmstadt Roth, Karlsruhe Serva, Heidelberg Sigma, St. Louis, USA Merck, Darmstadt Roth, Karlsruhe Roche, Mannheim Difco, Detroit, USA Difco, Detroit, USA Serva, Heidelberg

#### 4.8 Verschiedenes

Amberlite Mischbettionenaustauscher Amersham Hybond<sup>™</sup>-N<sup>+</sup> **Centrifugal Filter** Chromatographiepapier 3MM **DE 81-Papier** Dialyseschläuche, SpectraPor Expositionskassetten E1000 Entwickler F1000 Fixierer GeneJET Plasmid Miniprep Kit Glaswolle, silanisiert MN Faltenfilter Qiagen Plasmid Midi Kit **QIAquick PCR Purification Kit Röntgenfilme RX** Sterilfilter FB 030/3 (0,2 µm) Verstärkerfolie Du Pont Cronex

Roth, Karlsruhe GE Healthcare, UK Millipore, USA Whatman, England Whatman, England Spectrum Medical Industries, USA Siemens Christiansen GmbH, München Christiansen GmbH, München **Thermo Scientific** ICN Biochemicals, USA Macherey-Nagel, Düren Qiagen, Hilden Qiagen, Hilden Fuji Photo Film, Japan Schleicher und Schuell, Dassel **DuPont**, Bad Homburg

## 5 Methoden

## 5.1 Mikrobiologische Methoden

Bei der Handhabung mit Bakterienstämmen wurde stets unter sterilen Bedingungen gearbeitet. Die Vorschriften des Gentechnikgesetzes für das Arbeiten mit genetisch veränderten Organismen (GvO's) (S1) wurden dabei eingehalten und anfallender mikrobiologischer Abfall vor der Entsorgung autoklaviert.

### 5.1.1 Sterilisation von Lösungen, Glasgeräten und sonstigen Arbeitsmaterialien

Verwendete temperaturstabile Lösungen, Puffer, Medien und Laborutensilien wurden durch Dampfdrucksterilisation (50 min, 120°C, 1-2 bar) autoklaviert. Temperatursensitive Lösungen wurden mit Hilfe eines 0,2 µm Membranfilters (Schleicher und Schuell) sterilfiltriert und verwendete Glasgeräte, wie Glaspipetten und Zentrifugenröhrchen, für 4h bei 210°C hitzesterilisiert.

#### 5.1.2 Kultivierung von Bakterienstämmen

#### 5.1.2.1 Anzucht auf Agarplatten

Die Anzucht von Bakterienstämmen erfolgte aus Glycerinstocks (5.1.2.3) mit Hilfe einer sterilen Impföse auf YT-Agarplatten (4.6.2), die gegebenenfalls mit einem selektiven Antibiotikum versehen waren. Die Inkubation der Agarplatten erfolgte im Brutschrank bei 37°C bis zum sichtbaren Wachstum von Einzelkolonien. Anschließend konnten Einzelkolonien auf eine neue Agarplatte (Masterplatte) überführt oder in Flüssigkultur übernommen werden. Die Lagerung der Zellen für 2-4 Wochen konnte auf YT-Platten bei 4°C erfolgen. Für eine längerfristige Sicherung von Bakterienstämmen dienten Glycerinkulturen.

## 5.1.2.2 Anzucht von Flüssigkulturen über Nacht (üN) oder über Tag (üT)

Für die Anzucht einer Flüssigkultur wurde eine Einzelkolonie mit einem sterilen Zahnstocher von einer Agarplatte gepickt und in 3 ml bzw. 25 ml YT-Medium (4.6.2)

überführt. Das Medium wurde, wenn nötig, zuvor mit einem selektiven Antibiotikum versetzt. Es wurden folgende Antibiotika-(End-) Konzentrationen eingesetzt:

Ampicillin	100 μg/ml bzw. 50 μg/ml
Chloramphenicol	34 μg/ml bzw. 25 μg/ml
Kanamycin	25 μg/ml
Tetrazyklin	12,5 μg/ml
Streptomycin	50 μg/ml

Die Inkubation der Flüssigkulturen erfolgte bei 37°C im Wasserbad des Rundschüttlers (New Brunswick Scientific Gio Gyrotory, ca. 200 rpm).

#### 5.1.2.3 Anlegen von Glycerinkulturen (Glycerinstocks)

Zur langfristigen Sicherung der verwendeten Bakterienstämme wurden 0,6 ml einer üN-Kultur des gewünschten Stammes in einem sterilen Stockgläschen mit 0,4 g 100 % (400 µl) Glycerin versetzt und 1 bis 2 Stunden bei Raumtemperatur auf einem Horizontalschüttler (Gerhardt, LS 10) inkubiert. Die Lagerung der Bakterienkultur erfolgte nach kurzer Behandlung mit flüssigem Stickstoff bei -70°C.

# 5.1.2.4 Erstellung von Wachstumskurven und Berechnung von Wachstumsraten

Für die Erstellung von Wachstumskurven wurde eine 100 ml Flüssigkultur 1:100 mit einer üN-Kultur des entsprechnenden Bakterienstamm angeimpft und alle 30 min eine Probe zur Streumessung (5.2.1.2) am Photometer (Beckmann DU 640 Spectrophotometer) bei einer Wellenlänge von 600 nm entnommen. Die Messwerte wurden halblogarithmisch gegen die Zeit aufgetragen und die Wachstumsrate (hier:  $\mu$ = Verdoppelungen pro Stunde) aus der Geradengleichung der linearisierten Steigung der Wachstumskurve während des exponentiellen Wachstums (maximale Teilungsrate der Zelle) nach folgender Formel berechnet:

 $\mu = \lg (M_2) - \lg (M_1) / \lg 2 (t_2 - t_1)$ 

mit M<sub>1</sub> = OD<sub>600</sub> zum Anfang des exponentiellen Wachstums

 $M_2 = OD_{600}$  zum Ende des exponentiellen Wachstums

 $t_1$  = Zeitpunkt zu M<sub>1</sub>

 $t_2$  = Zeitpunkt zu M<sub>2</sub>.

# 5.1.3 Herstellung chemisch kompetenter *E. coli* Zellen für die Transformation

Die Herstellung chemisch kompetenter *E. coli* Zellen erfolgte nach einer von Dagert & Ehrlicher, (1979) veröffentlichten Methode. Dazu wurden 100 ml YT-Medium mit 1/100 Volumen einer üN-Kultur des entsprechenden Bakterienstammes angeimpft und das Wachstum mittels Streumessung (5.2.1.2) verfolgt. Nach Erreichen einer optischen Dichte ( $OD_{600}$ ) von 0,5 bis 0,6 (exponentielle Wachstumsphase) wurden je 50 ml der Kultur in sterile, vorgekühlte Corex-Zentrifugenröhrchen überführt und 20 Minuten auf Eis inkubiert. Anschließend folgte ein Zentrifugationsschritt bei 4500 rpm, 4°C für 5 Minuten (Heraeus Megafuge 1.0R). Der klare Überstand wurde abgenommen, das Zellpellet vorsichtig in 20 ml eiskalter 0,1 M CaCl<sub>2</sub>-Lösung resuspendiert und eine Stunde auf Eis inkubiert. Nach erneuter Sedimentation der Zellen (siehe oben) wurde der Überstand verworfen und das Zellpellet in 2 ml eiskalter 85 mM CaCl<sub>2</sub>-Lösung mit 15 % Glycerin aufgenommen. Die Zelllösung wurde dann zu je 200 µl in sterile Eppendorfgefäße aliquotiert. Nach Inkubation für 1-16 Stunden konnten die Zellen entweder direkt für Transformationsexperimente verwendet oder kurz mit Stickstoff behandelt bei -70°C gelagert werden.

#### 5.1.4 Transformation kompetenter Zellen mit Vektor-DNA durch Hitzeschock

Für die Übertragung von Vektor-DNA in kompetente Bakterienzellen wurde eine von Hanahan publizierte Methode (Hanahan 1985) angewendet. Ein 200 μl Aliquot kompetenter Zellen wurde langsam auf Eis aufgetaut, mit der gewünschten Menge an Plasmid-DNA (5–50 ng) vorsichtig gemischt und 1h auf Eis inkubiert. Nach Hitzeschock für 3 min bei 42°C im Wasserbad, wurden die Zellen kurz auf Eis gestellt, um die Transformationseffizienz zu erhöhen. Anschließend erfolgte die Zugabe von 800 μl vorgewärmten YT-Medium (4.6.2) und eine einstündige Inkubation bei 37°C unter Schütteln bei 550 rpm (Haep Labor Consultant, HTM 130L). Zum Schluss wurden 100–200 μl der Zellen auf selektive und vorgewärmte Agarplatten (4.6.2) ausgestrichen und über Nacht bei 37°C bis zum Anwachsen von Einzelkolonien inkubiert. Durch Auszählen der Kolonien konnte unter Berücksichtigung der eingesetzten DNA-Menge und des ausplattierten Zellvolumens die Transformationsrate der kompetenten Zellen pro μg Plasmid-DNA bestimmt werden. Diese lag abhängig von Stamm und Plasmid bei 10<sup>4</sup> bis 10<sup>5</sup> Transformanden pro μg DNA.

#### 5.1.5 Sequenzierung von Plasmid-DNA

Sequenzierungen wurden bei der Firma StarSEQ (Mainz) in Auftrag gegeben. Dazu wurden Plasmid- und Primerkonzentration den jeweiligen Erfordernissen angepasst.

Die Elektropherogramme wurden mit dem Programm FinchTV Version 1.4.0 ausgewertet. Die Sequenzen wurden mit dem Programm ClustalX Version 2.0 verifiziert.

## 5.2 Molekularbiologische Methoden

#### 5.2.1 Konzentrationsmessungen

#### 5.2.1.1 Konzentrationsbestimmung von Nukleinsäuren

Die Bestimmung der Konzentration und der Reinheit von Nukleinsäuren erfolgte durch die Messung der optischen Dichte (OD) mit Hilfe eines Messgerätes der Firma Thermo Scientific (NanoDrop2000). Die Konzentrationsbestimmung erfolgte durch Absorptionsmessung bei einer Wellenlänge  $\lambda$  von 260 nm. Zunächst wurde hierzu der NanoDrop2000 mit 1 µl der Lösung, in der auch die Nukleinsäure gelöst wurde, meist TE-Puffer (4.6.1) oder *Aqua dest.*, kalibriert. Für die anschließende Messung wurde 1 µl der Probe auf die Messfläche pipettiert. Der Reinheitsgrad einer Nukleinsäure wurde über den Quotienten OD<sub>260</sub>/OD<sub>280</sub> berechnet. Für eine reine Probe sollte der Wert zwischen 1,8 und 2,0 liegen. Deutlich verringerte Werte deuten auf Verunreinigungen durch Proteine oder Phenolreste hin.

#### 5.2.1.2 Bestimmung der Zelldichte von Flüssigkulturen

Zur Bestimmung des Wachstum von Bakterienzellen in Flüssigkulturen wurde die optische Dichte (OD) bei 600 nm in regelmäßigen Abständen im Spektralphotometer (Beckmann, Typ DU 64) verfolgt, die sich annähernd proportional zur Zellzahl verhält. Hierzu wurden 100 ml YT-Medien (4.2.6) mit 1 ml einer 3 ml üN-Kultur (5.1.2.2) der entsprechenden Zellen angeimpft und bei konstanter Temperatur geschüttelt. Zu den angegebenen Zeiten wurde ein Aliquot entnommen und die Absorption der Probe im Spektralphotometer (Beckmann, Typ DU 64) bei 600 nm gemessen. Die Kalibrierung erfolgte mit dem jeweiligen Medium, in welchem die Bakterien wuchsen. Damit in einem linearen Bereich gemessen werden konnte (OD<sub>600</sub>-Werte zwischen 0,1 und 0,9), mussten die Proben gegebenenfalls mit sterilem YT-Medium (4.6.2) verdünnt werden. Für die Auswertung wurden die gemessenen Werte logarithmisch gegen die Zeit aufgetragen.

Aus der Steigung m =  $\Delta \log OD_{600}/\Delta t$  [min<sup>-1</sup>] in der exponentiellen Wachstumsphase ließen sich Wachstumsrate und Verdopplungszeit errechnen.

Die Wachstumsrate  $\mu$  errechnet sich nach folgender Formel:  $\mu = In10/In2 \cdot m \cdot 60 [h^{-1}]$ .

Die Verdopplungszeit td berechnet sich nach der Formel: td =  $1/\mu \cdot 60$  [min].

#### 5.2.1.3 Konzentrationsbestimmung von Proteinen

Die Konzentrationsbestimmung von Proteinen erfolgte mittels des "Bradford-Microassays" der Firma Bio Rad. Dieser basiert auf einer quantitativen photometrischen Messung der Absorption bei einer Wellenlänge von 595 nm. Es wurden 2–20 µl der Proteinlösung mit 780–798 µl *Aqua dest.* und 200 µl Bradford-Reagenz (4.7) gemischt und wenige Minuten bei Raumtemperatur (RT) inkubiert. Die vorherige Kalibrierung des Photometers erfolgte gegen 780–798 µl *Aqua dest.,* 2-20 µl des sich in der Proteinlösung befindlichen Puffers und 200 µl Bradford-Reagenz. Anhand einer zuvor erstellten Eichgerade mit bekannten Proteinkonzentrationen eines BSA-Standards (4.5.3), konnte so die Konzentration der untersuchten Probe bestimmt werden.

#### 5.2.2 Isolation von Nukleinsäuren aus Bakterien

### 5.2.2.1 Isolation von Plasmid-DNA mit Hilfe des GeneJET Plasmid Miniprep-Kit ("Minipräp")

Die Plasmidpräparation durch das Kit von Thermo Scientific erfolgte nach den Angaben des Herstellers. Es wurden 3-12 ml der üN gewachsenen Zellen (high oder low-copy Plasmid) bei 8000 rpm für 2 Minuten pelletiert. Das Pellet wurde in 250  $\mu$ l Resuspensions-Puffer (RNase-haltig, 4°C) aufgenommen und nach Zugabe von 250  $\mu$ l Lyse-Puffer durch mehrmaliges invertieren gemischt. Anschließend wurde die Probe mit 350  $\mu$ l Neutralisations-Puffer versetzt und erneut mehrmals invertiert. Nach Ausfallen der chromosomalen DNA erfolgte eine 5 minütige Zentrifugation bei 12000 rpm und der klare Überstand wurde auf eine im Kit enthaltene Säule gegeben. Auf eine 1 minütige Zentrifugation bei 12000 rpm folgten zwei Waschschritte mit jeweils 500  $\mu$ l Wasch-Puffer. Um Reste des Wasch-Puffers von der Säule zu entfernen wurde abermals für 1 Minute zentrifugiert. Die Elution der Plasmid-DNA erfolgte mit 30–50  $\mu$ l Elutions-Puffer und die Bestimmung der Konzentration mithilfe des Nano-Drop2000 (5.2.1.1). Zum Schluss wurde die Qualität der Präparation auf einem 0,8%igen Agarosegel (5.2.4.1) überprüft.

# 5.2.2.2 Plasmidisolation mittels dem Qiagen Plasmid Midi Kit ("Midipräp")

Die Midi-Präparation von Plasmid-DNA beruht wie in 5.2.2.1 auf dem Prinzip der alkalischen Lyse und wurde angewendet, wenn ein erfolgreich kloniertes Plasmid durch Restriktionshydrolyse (5.2.6.1) bereits identifiziert wurde und im Anschluss daran die Sequenz des Plasmids überprüft werden sollte. Die Präparation folgte dabei weitgehend den Anweisungen des Kits von Qiagen. 50 ml einer üN-Kultur wurden dafür in Corex Glasgefäße überführt und mittels Zentrifugation pelletiert (6000 rpm, 15 min, 4°C; Heraeus Megafuge 1.0 R). Das Zellpellet wurde in 4 ml Puffer P1 (RNasehaltig, 4°C) resuspendiert und nach Zugabe eines zweiten Puffers (Puffer P2) durch mehrmaliges Invertieren gemischt. Nach 5 minütiger Inkubation bei Raumtemperatur sollte die Probe eine Blaufärbung, verursacht durch LyseBlue (im Puffer enthalten), vorweisen. Anschließend wurden 4 ml eiskalter Puffer P3 zugegeben, die Probe mehrfach invertiert und für 15 Minuten auf Eis inkubiert. Ein Farbwechsel von blau nach milchig zeigt dass die Lyse der Probe funktioniert hat. Durch Zentrifugation (6000 rpm, 15 min, 4°C; Heraeus Megafuge 1.0 R) wurden Zelltrümmer und chromosomale DNA weitestgehend entfernt. Der Überstand wurde auf eine Quiagentip 100 Säule pipettiert, welche zuvor mit 4 ml QBT Puffer äquilibriert wurde. Nach zweimaligen Waschen mit 10 ml QC Waschpuffer, wurde die DNA mittels 5 ml QF Puffer von der Säule eluiert. Im Anschluss daran wurde die gelöste DNA mit 3,5 ml Isopropanol (RT) vermischt und bei 6000 rpm, 4°C für 30 Minuten zentrifugiert (Heraeus Megafuge 1.0 R). Nach Abnahme des Überstandes wurde das Pellet mit 2 ml 70% (v/v) Ethanol gewaschen (10 Minuten, 6000 rpm), luftgetrocknet und in 40 μl TE-Puffer (4.6.1) aufgenommen.

# 5.2.2.3 Isolation von Plasmid-DNA im präparativen Maßstab ("Maxipräp")

Für die Gewinnung von größeren Mengen an Plasmid-DNA wurde die von Hillen beschriebenen Methode angewendet. Dazu wurden zunächst 3 ml YT-Medium (4.6.2) mit einer Einzelkolonie des entsprechenden Bakterienstammes angeimpft und üN bei 37°C geschüttelt. Am nächsten Morgen wurde ein 25 ml YT-Kolben im Verhältnis 1:100 mit dieser üN-Kultur (5.1.2.2) versetzt und für mehrere Stunden bei 37°C im Wasserbad geschüttelt (über-Tag-Kultur, üT). Am Abend wurden dann 800 ml YT-Medium mit 8 ml der üT-Kultur angeimpft und bei 37°C über Nacht auf einem Schüttler inkubiert (New Brunswick Scientific, Model G25). Die verwendeten YT-Medien wurden stets mit einem geeigneten Antibiotikum zur Selektion supplementiert. Am Tag darauf wurde die Zellkultur auf zwei JA-10 Becher verteilt und für 10 min bei 8000 rpm und 4°C (Beckman Zentrifuge J2-21, JA-10-Rotor) zentrifugiert. Der Überstand wurde verworfen und die Pellets in 8 ml Saccharoselösung bei 4°C resuspendiert. Zum Zellaufschluss wurde die Suspension auf zwei Ti55.2-Polycarbonatröhrchen aufgeteilt, mit jeweils 3 ml 0,5 M EDTA, pH 8,0 und 3 ml Lysozymlösung versetzt, mehrmals invertiert und dann für 15-30 min auf Eis inkubiert. Darauf folgte die Zugabe von 2 ml Brij-Doc-Lösung und eine 10minütige Inkubation auf Eis. Nach Zentrifugation für 45 min bei 44000 rpm und 4°C (Beckman L8-55 Ultrazentrifuge, Rotor Ti55-2) wurde der klare, plasmidhaltige Überstand mit 2,5 µl RNase A (20 mg/ml) pro ml Überstand versetzt und für 30 min bei 37°C inkubiert. Direkt im Anschluss folgte die Zugabe von 2,5 µl Proteinase K (20 mg/ml) pro ml Überstand und eine erneute Inkubation für 30 min bei 37°C. Um die Plasmid-DNA zu fällen wurden die Proben mit ½ Volumen PEG-Lösung für 30 min auf Eis inkubiert und anschließend für 45 min (4000 rpm und 4°C, Beckman Zentrifuge J2-21, JA-17-Rotor) zentrifugiert. Das Pellet wurde in 3 ml TE-Puffer (4.6.1) aufgenommen und Phenol/Chloroform extrahiert (5.2.3.1). Nach Ethanolfällung (5.2.3.2) wurde das lyophilisierte Pellet in einem geeigneten Volumen TE-Puffer gelöst, die Konzentration durch Messung am Nanodrop (5.2.1.1) bestimmt und die Qualität der Präparation auf einem analytischen Agarosegel (5.2.4.1) überprüft.

Saccharose-Lösung:	25 % (w/v) Saccharose 50 mM Tris-HCL nH 8 0
Lysozym-Lösung:	20 mg/ml Lysozymchlorid 50 mM Tris-HCl, pH 8,0
Brij-Doc-Lösung:	2 Volumen 10 % (w/v) Brij 35 1 Volumen 10 % (w/v) Na-Deoxycholat in <i>Aqua dest.</i> pH 8,0
RNase A:	20 mg/ml in 10 mM Tris-HCl, pH 8,0 (nach Maniatis <i>et al.</i> , 1982) von DNase- Aktivität befreit
Proteinase K:	20 mg/ml in <i>Aqua dest.</i>
PEG-Lösung:	30 % (w/v) PEG <sub>6000</sub> in 1,5 M NaCl

#### 5.2.3 Reinigung und Konzentrierung von Nukleinsäuren

#### 5.2.3.1 Phenol/Chloroform-Extraktion

Diese Art der Extraktion wurde angewendet, um wässrige Nukleinsäurelösungen von Proteinen oder Enzymen/Puffern nach erfolgter Restriktionshydrolyse zu entfernen. Aufgrund der unterschiedlichen Löslichkeit der zu extrahierenden Substanzen, können diese in einem Zweiphasen-System voneinander getrennt werden. Dafür wurde die Lösung mit 1 Volumen Phenol/Chloroform (Verhältnis 1:1) versetzt, für eine Minute gevortext und zur besseren Trennung der wässrigen von der organischen Phase für 5 min zentrifugiert (13000 rpm, Hettich Mikro22). Die obere, nukleinsäurehaltige Phase (wässrige Phase) wurde abgenommen, in ein neues Eppi überführt und mit 1 Volumen Chloroform versetzt, um restliche Phenolreste zu entfernen. Die Interphase und die organische Phase wurden verworfen. Nach einminütigem Vortexen wurde erneut für 5 min zentrifugiert. Der nun protein- und phenolfreie Überstand wurde abgehoben, in ein neues Eppendorfgefäß überführt und die Nukleinsäure durch eine Ethanolfällung (5.2.3.2) aufkonzentriert.

#### 5.2.3.2 Ethanol-Präzipitation von Nukleinsäuren

Die Fällung von wässrigen, nukleinsäurehaltigen Lösungen erfolgte mit 1/10 Volumen 3 M Natriumacetat (pH 5,5) und 2-3 Volumen eiskaltem, absolutem Ethanol durch mehrmaliges Invertieren und anschließender Inkubation bei -20°C für 30-60 Minuten oder alternativ für 20-30 Minuten bei -70°C. Für Oligonukleotide (4.4.1) oder Sequenzierreaktionen (5.4.3) diente 1/5 Volumen Glykogen (20 µg/ml) und 1/10 Volumen 7,5 M Ammoniumacetat, pH 5,7 als Fällungshilfe. Die so gefällten Nukleinsäuren wurden anschließend durch Zentrifugation für 30–45 Minuten (12000 rpm, Eppendorfgefäß, Hettich Mikro22 oder 6000 rpm, Greinerröhrchen, Heraeus Megafuge 1.0R) pelletiert. Der Überstand wurde abgenommen, das Pellet mit einem Volumen 80%igen Ethanol versetzt und durch eine erneute 5minütige Zentrifugation von Salzresten befreit. RNA-Proben und markierte Oligonukleotide wurden zweimal mit 80% Ethanol gewaschen. Der Überstand wurde verworfen und das Pellet im Vakuumkonzentrator (Heto SpeedVac VR-1) lyophilisierte oder an der Luft getrocknet. Zum Schluss erfolgte die Aufnahme des Pellet in einem geeigneten Volumen *Aqua dest.* oder TE-Puffer (4.6.1).

#### 5.2.3.3 Dialyse von Nukleinsäuren

Um bei späteren Experimenten störende Salze in den präparierten Nukleinsäureproben zu entfernen, konnten Nukleinsäurelösungen bis zu einem Volumen von 100  $\mu$ l mit Hilfe der Mikrotropfendialyse entsalzt werden. Dafür wurde eine sterile Petrischale mit 30-40 ml TE-Puffer (4.6.1) gefüllt und ein VS Millipore Membranfilter (Durchmesser 2,5 cm, Porengröße 0,025  $\mu$ m) mit der glänzenden Seite nach oben auf die Oberfläche der Flüssigkeit gelegt. Die zu dialysierende Probe wurde dann auf den Membranfilter pipettiert und für 1-2 Stunden bei Raumtemperatur inkubiert. Anschließend wurde die Probe vorsichtig vom Filter abgenommen, in ein neues Reaktionsgefäß überführt und der Membranfilter 2 x mit je 5  $\mu$ l TE-Puffer gespült, um auf dem Filter verbliebene Nukleinsäure aufzunehmen. Zum Schluss wurde die Konzentration der Nukleinsäureprobe spektralphotometrisch (5.2.1.1) bestimmt.

# 5.2.3.4 Isolation von Nukleinsäuren aus Agarosegelen durch Glaswoll-Elution

Um DNA- oder RNA-Fragmente bestimmter Größen zu isolieren, wurden diese über ein präparatives Agarosegel (5.2.4.1) elektrophoretisch getrennt. Unter UV-Licht wurde das entsprechende Fragment sichtbar gemacht und mit einem sterilen Skalpell aus dem Gel ausgeschnitten. Die Elution des DNA-Fragmentes erfolgte anschließend mittels Zentrifugation über silikonisierte Glaswolle. Hierzu wurde der Boden eines 0,5 ml PCR-Reaktionsgefäßes mit einer Kanüle, die über dem Bunsenbrenner erhitzt wurde, durchbohrt und der untere Teil des Gefäßes mit einer sterilen Pinzette mit silikonisierter Glaswolle dicht gepackt. Das mit Glaswolle gepackte 0,5 ml Reaktionsgefäß wurde dann in ein 1,5 ml Eppendorfgefäß gestellt und das Agarosegelstück mit dem zu eluierenden Nukleinsäure-Fragment über die Glaswolle gepackt. Die Elution der DNA erfolgt anschließend durch mehrfaches Zentrifugieren über die Glaswolle (einmal 5 min bei 6000 rpm und anschließend bei 12000 rpm, Hettich Mikro22). Das wässrige Eluat wurde Phenol/Chloroform extrahiert (5.2.3.1), mit Ethanol präzipitiert (5.2.3.2) und in TE-Puffer (4.6.1) gelöst. Die Reinheit und Konzentration wurden auf einem 0,8% Agarosegel (5.2.4.1) und mittels Spektralphotometer (5.2.1.1) analytisch überprüft.

## 5.2.3.5 Isolation von DNA-Fragmenten aus präparativen Agarosegelen mit Hilfe des Wizard<sup>®</sup> SV Gel and PCR Clean-up Systems

Alternativ zur Isolation von Nukleinsäuren über Glaswoll-Elution konnten diese auch mittels des Wizard<sup>®</sup> SV Gel and PCR Clean-up System nach Anleitung des Herstellers Promega aufgereinigt werden. Nach Auftrennung des Nukleinsäure-Fragmentes über ein präparatives Agarosegel (5.2.4.1) wurde dieses unter UV-Licht mit einem sterilen Skalpell ausgeschnitten und in ein steriles 1,5 ml Eppendorfgefäß überführt. Durch Zugabe von 10 µl Membrane Binding Solution pro 10 mg Gel und Inkubation bei 50-65°C im Heatblock wurden die Gelstücke aufgelöst. Die gewonnene Gel-Lösung wurde anschließend auf eine SV Minicolumn mit dazugehörigem Eppendorfgefäß überführt, für 1 Minute bei RT inkubiert und dann für 1 Minute bei 14.000 rpm zentrifugiert. Der aufgefangene Durchfluss wurde verworfen und die Säule in zwei Waschritten (1x 700 µl und 1x 500 µl Wash Solution) gewaschen. Für die Elution der an der Membran gebunden Nukleinsäure wurde die Säule in ein neues, steriles Eppendorfgefäß überführt, 30-50 µl TE-Puffer hinzugegeben und für 1 Minute bei RT inkubiert. Durch eine darauffolgende Zentrifugation für 2 Minuten bei 14.000 rpm wurde die Nukleinsäure von der Säule eluiert und die Konzentration konnte am NanoDrop2000 (5.2.1.1) bestimmt werden.

#### 5.2.3.6 Aufreinigung von DNA mit dem QIAquick PCR Purification Kit

Um DNA-Fragmente oder Plasmid-DNA nach erfolgter PCR-, Mutagenesereaktion oder Restriktion aufzureinigen, wurde das QIAquick PCR Purification Kit der Firma Qiagen genutzt. Dafür wurde zum Volumen der eingesetzten Probe das 5fache Volumen an PB Puffer hinzugegeben, welcher zuvor mit einem pH Indikator (1:250) versetzt wurde. Verfärbte sich die Lösung orange oder violett mussten 10  $\mu$ l 3M Natriumacetat zum neutralisieren hinzugegeben werden. Die Probe wurde anschließend auf eine im Kit enthaltene Säule gegeben und 30-60 Sekunden bei 13000 rpm zentrifugiert. Der Durchfluss wurde verworfen und die Säule mit 750  $\mu$ l Puffer PE gewaschen. Darauf folgte ein erneuter Zentrifugationsschritt (30-60 Sekunden bei 13000 rpm) und der Durchfluss wurde abermals verworfen. Um die Säule von Waschpufferresten zu befreien wurde nochmals zentrifugiert. Dann wurde die Säule in ein 1,5 ml steriles Eppendorfgefäß überführt und mit 30–50  $\mu$ l EB Puffer behandelt. Nach erfolgter Inkubation für 2 Minuten bei RT, konnte die Probe durch einen abschließenden Zentrifugationsschritt von der Säule eluiert werden und die Konzentration mittels NanoDrop2000 (5.2.1.1) bestimmt werden.

#### 5.2.4 Gelelektrophoresen

#### 5.2.4.1 Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese dient der analytischen und präparativen Auftrennung von Nukleinsäure-Fragmenten nach ihrer Größe durch Verwendung eines elektrischen Feldes. Soweit nicht anders angegeben wurden 0,8 bis 1,2 %ige (w/v) Agarosegele verwendet (Maniatis *et al.*, 1982). Dazu wurde die gewünschte Menge an Agarose in 1 x TAE-Puffer (4.6.1) auf 100 ml mit destilliertem Wasser aufgefüllt und unter ständigem Rühren bis zum vollständigen Lösen der Agarose aufgekocht. Nach Abkühlen der Lösung auf unter 60°C, wurde sie mit 0,5 µg/ml Ethidiumbromid versetzt und als ein Flachbettgel (14 x 11 cm) gegossen. Nach Polymerisation der Gelmatrix wurde das Flachbettgel in eine Elektrophoresekammer gelegt und mit 1 x TAE-Laufpuffer, der ebenfalls 0,5  $\mu$ g/ml Ethidiumbromid enthielt, überschichtet. Die DNA-Proben wurden mit einem Volumen 2 x TAE-Probenpuffer versetzt und bei 90 bis 120 V für 45 bis 90 min elektrophoretisch aufgetrennt. Das Ethidiumbromid interkaliert in die DNA und die Nukleinsäuren können durch Verwendung von UV-Licht in einem Dokumentationssystem (Syngene, G:BOX) sichtbar gemacht werden. Für präparative Agarosegele wurde 'ultrapure' Agarose verwendet. Um RNA aufzutrennen wurde die Elektrophoresekammer, der Gelschlitten und der Gelkamm vorher für 1-2 h mit 0,1 % DEPC-Lösung behandelt um eventuell vorhandene Nukleasen zu inaktivieren.

#### 5.2.4.2 Native Polyacrylamidgelelektrophorese (PAGE)

Bei Polyacrylamidgelen handelt es sich um Polymerisationsprodukte bestehend aus Acrylamid und N,N'-Methylenbisacrylamid, wobei letzteres als Quervernetzer für die zweidimensionale Kettenbildung aus Acrylamid dient. Die Porengröße und damit der Auftrennungsbereich des Polyacrylamidgels werden durch die Konzentration des Acrylamids und den Vernetzungsgrad bestimmt. Die Polymerisation ist eine Radikalkettenreaktion, die durch Zugabe von N,N,N,N'-Tetramethylethylendiamin (TEMED, Polymerisationskatalysator) katalysiert und Ammoniumperoxodisulfat (APS, Radikalbildner) gestartet wird.

Diese Art von Gel wurde für Retardierungsgelelektrophoresen (5.4.4) verwendet und 5%ige Gellösungen mit einer Vernetzung von 46:1 (Acrylamid:Bisacrylamid) angesetzt. Die Polymerisation wurde durch Zugabe von 0,05 % (v/v) TEMED und 0,5 % (v/v) 10 % APS gestartet. Nach gießen der Gellösung zwischen zwei zuvor mit Ethanol gesäuberten Glasplatten wurde das auspolymerisierte Gel (330 x 250 x 1 mm) in eine Elektrophoresekammer eingespannt und als Laufpuffer 0,5 x TBE-Puffer (5.6.1) verwendet. Die DNA-Proben wurden vor dem Auftragen ebenfalls mit 1/3 Volumen 2 x TBE-Probenpuffer versetzt und bei 20-30 mA für 2 bis 4 Stunden aufgetrennt. Die Visualisierung der aufgetrennten Nukleinsäuren erfolgte mittels Autoradiographie (5.2.5.2).

#### 5.2.4.3 Denaturierende Polyacrylamidgelelektrophorese (dPAGE)

Die denaturierende PAGE wurde zur Produktanalyse von DNA-Sequenzierungen (5.3.1) verwendet und trennt Nukleinsäuren durch Verwendung des denaturierenden Agens Harnstoff und hohen Geltemperaturen strukturunabhängig aufgrund ihres Molekulargewichts auf. Es wurden 10, 12- oder 15% ige Gele mit einer Vernetzung von 19:1 (Acrylamid:Bisacrylamid) verwendet, die Watt-abhängig in drei Stufen vorgeheizt wurden (10 min 25 W, 10 min 50 W, 10 min 90 W) und mit 1 x TBE-Puffer (4.6.1) liefen. Nachdem sich der Harnstoff vollständig gelöst hatte wurde, die Gellösung für einige Minuten entgast, die Polymerisation durch die Zugabe von TEMED und APS gestartet und zwischen zwei vorbereitete Glasplatten gegossen (Dimension des Gels 310 x 395 x 0,5 mm). Nach Auspolymerisierung des Gels wurde die Elektrophoresekammer eingesetzt und die zu analysierenden es in Nukleinsäureproben, versetzt mit Formamid-Probenpuffer, für 6 min bei 96°C denaturiert und bei 90 W aufgetrennt. Um bei radioaktiven Proben eine Kontamination des Anodenpuffers so gering wie möglich zu halten, wurde die Anodenseite der Glasplatten zuvor mit DE-81 Ionenaustauschpapier umwickelt.

### 5.2.5 Nachweismethoden von Nukleinsäuren

#### 5.2.5.1 Ethidiumbromidfärbung von Nukleinsäuren

Die Nachweisgrenze von DNA und RNA beträgt bei der Ethidiumbromid-Färbung 10-20 ng pro Bande in Polyacrylamidgelen (5.2.4.2) und 20-30 ng pro Bande in Agarosegelen (5.2.4.1). Bei präparativen Gelen ist die Färbung durch 0,1%-ige Essigsäure reversibel. Für die Färbung wird das Gel für 10 min in einer EtBr-Färbelösung (2  $\mu$ g/ml) unter schütteln inkubiert (Gerhardt, Schüttelmaschine LS10) und anschließend am UV-Transilluminator (Herolab UVT 2035) ausgewertet.

#### 5.2.5.2 Nachweis über Autoradiographie

Der Nachweis von radioaktiv markierten Proben erfolgte durch die Belichtung eines Röntgenfilms. Nach erfolgter Elektrophorese wurde das Gel auf einen alten Röntgenfilm gezogen, mit Polyethylenfolie bedeckt und in eine Expositionskassette gelegt. Bei Rotlicht konnte anschließend in der Dunkelkammer ein Röntgenfilm (Kodak X-OMAT AR oder FUJI RX) auf das Gel gelegt und dieser für 1 bis 2 Tage bei -20°C exponiert werden. Falls nötig wurde zusätzlich eine Verstärkerfolie (DuPont Cronex) aufgelegt und der Film bei -70°C exponiert Die Entwicklung der Röntgenfilme erfolgte in der Entwicklermaschine AGFA Curix 60.

#### 5.2.5.3 Silberfärbung von Nukleinsäuren

Nukleinsäuren, die in nativen PAA-Gelen aufgetrennt wurden, konnten mit der von Beidler beschriebenen Methode mit Silber angefärbt werden (Beidler, Hilliard et al. 1982). Diese Methode erlaubt den Nachweis von Nukleinsäuren im ng-Bereich. Hierzu wurden die Gele nach der Elektrophorese in Fixierlösung I unter leichtem Schütteln (Gerhardt, Schüttelmaschine LS10) mindestens 10 min fixiert und anschließend 10 min in der Silbernitratlösung gefärbt. Danach wurden die Gele dreimal für jeweils 20–60 Sek. in *Aqua dest.* gewaschen und dann solange in der Entwicklungslösung inkubiert, bis die Banden gut sichtbar wurden. Nach Fixierung für 10 min in Fixierlösung II wurde das Gel in Folie eingeschweißt und zur Dokumentation gescannt (Hewlett Packard, Scan Jet 4c/T).

Fixierer I: 10% Ethanol 0.5% Essigsäure

Färbelösung: 0.19% (w/v) Silbernitrat

Entwickler:	15 g/l NaOH 0.08 g/l Natriumborhydrid (NaBH4) 0.148% (v/v) Formaldehyd (CH2O) immer frisch angesetzt
Fixierer II:	0.75% (w/v) Natriumcarbonat (Na2CO3)

#### 5.2.6 Enzymatische Reaktionen

#### 5.2.6.1 Restriktionshydrolysen von DNA

Restriktionsenzyme schneiden DNA spezifisch oder unspezifisch und spielen eine wichtige Rolle bei der Analyse von DNA-Fragmenten, Plasmiden und bei der Herstellung und Identifizierung von Klonierungsvektoren. Die in dieser Arbeit verwendeten Restriktionsenzyme gehören zum Typ II der Endonukleasen und schneiden innerhalb ihrer Erkennungssequenz. Je nach Restriktionsendonuklease entstehen hierbei DNA-Stränge mit 3'-oder 5'-über-hängenden ('sticky') oder glatten ('blunt') Enden. Die Optimalbedingungen der verwendeten Enzyme, wie Temperatur oder Pufferbedingung wurden den Angaben des Herstellers entnommen. Je nach eingesetzter Nukleinsäuremenge variierte die Inkubationszeit der Ansätze von über Nacht, bis hin zu nur wenigen Stunden meist bei 37°C im Wärmeschrank. Zur Kontrolle der Restriktion wurde nach Inkubation ein Aliquot des Ansatzes mittels Agarosegelelektrophorese (5.2.4.1) aufgetrennt. Wurden die erhaltenen Fragmente für weitere enzymatische Reaktionen benötigt, folgte eine Phenol/Chloroform Extraktion (5.2.3.1) mit anschließender Ethanolfällung (5.2.3.2) und Analyse auf einem Agarosegel (0,8 %). Andernfalls konnten sie nach Herstellerangaben mit vorgegebenen Temperaturen hitzeinaktiviert werden.

#### 5.2.6.2 Polymerasekettenreaktion (PCR)

Die Methodik der PCR (*Polymerase Chain Reaction*) ermöglicht neben einer sequenzgenauen Amplifizierung von Nukleinsäurefragmenten *in vitro* auch die präparative DNA Gewinnung aus kleinsten Probenmengen (Mullis and Faloona 1987)und wurde 1987 erstmals von Mullis angewendet. Für die in dieser Arbeit durchgeführten Amplifizierungsreaktion wurde die *Phusion® High-Fidelity* DNA Polymerase (Biolabs), eine künstlich hergestellte "Chimäre" bestehend aus einer *Pyrococcus*-ähnlichen Polymerase und einer zusätzlich eingefügten dsDNA-bindenen Proteindomäne genutzt. Diese Art von DNA Polymerase ist besonders schnell und besitzt eine 3' $\rightarrow$ 5'-Exonukleaseaktivität, die im Vergleich zu anderen Polymerasen wie der *Taq*-Polymerase, eine um den Faktor 12 höhere Genauigkeit der Polymerase zur Folge hat. Mit Hilfe von zwei synthetischen Oligonukleotiden (Primern), die an komplementäre

#### 5 Methoden

Bereiche der Template-DNA binden (*annealing*), kann so ein definierter Sequenzbereich nach mehreren Zyklen von *Melting*, *Annealing* und *Elongation* millionenfach amplifiziert werden. Durch Denaturierung werden die gebildeten Doppelstränge aufgeschmolzen und es kommt durch wiederholte Zyklen von Hybridisierung und Verlängerung der Oligonukleotide durch Temperaturveränderung zu einer exponentiellen Vervielfältigung der Ziel-DNA.

Die Reaktion erfolgte in 1 x Phusion HF-Puffer (mit 1,5 mM MgCl<sub>2</sub>, Finnzymes) mit je 0,2 mM dNTP, 5-10 ng *template*-DNA, 0,5  $\mu$ M von jedem Oligonukleotid und 0,4 U *Phusion High-Fidelity* DNA Polymerase pro 20  $\mu$ l Ansatz.

Die Schmelztemperaturen der Oligonukleotide wurden mit folgender Formel ermittelt:

 $T_m = 69.3 + 0.41(\% GC) - 650/L$ L = Länge des Oligonukleotids (bp)

Die gewählte Annealingtemperatur der verwendeten Oligonukleotide ist dabei abhängig von der Schmelztemperatur und sollte in der Regel 5-10°C unterhalb des  $T_m$ -Wertes liegen. Anhand eines Temperaturgradienten konnte die optimale Annealingtemperatur bei der die Produktausbeute am besten ist ermittelt und für weitere Reaktionen gewählt werden.

Für die Reaktion wurde ein Thermocycler (Eppendorf, Mastercycler epgradient S) benutzt und das PCR-Programm entsprechend den Versuchsbedingungen angepasst:

Initiale Denaturierung:	30 sec 98°C
Amplifikationszyklen (25-30x):	5-10 sec 98°C (Denaturierung)
	10-30 sec x°C (Annealing)
	15-30 sec/kb 72°C (Verlängerung)
Finale Produktverlängerung:	5 min 72°C.

Die Qualität der PCR-Reaktion wurde anschließend auf einem analytischen Agarosegel (5.2.4.1) überprüft und die PCR-Produkte für die weitere Verwendung mit Hilfe des QIAquick PCR Purification Kits (5.2.3.6) aufgereinigt oder Phenol/Chlorform extrahiert (5.2.3.1).

#### 5.2.6.3 Ligation von DNA-Fragmenten

Bei der Ligation zweier DNA-Fragmente werden freie 5'-Phosphat- und 3'-OH-Gruppen von DNA-Molekülen kovalent miteinander verknüpft. Als Enzym wurde hierbei entweder die T4-DNA Ligase von Roche, oder die Quick T4-DNA Ligase von Biolabs verwendet. Das Reaktionsvolumen betrug in beiden Fällen 10  $\mu$ l mit 1 x T4-DNA Ligase Puffer, 10-75 ng eines durch Restriktionshydrolyse geöffneten Vektors und das einzuklonierende Fragment (Insert) in verschiedenen Vektor- zu Insertverhältnissen (1:2, 1:5, 1:10), die durch folgende Formel berechnet wurden:

ng Vektor  $\cdot$  Kbp Insert / Kbp Vektor

Die Ligase wurde immer zuletzt mit 1 U hinzugegeben, um die Reaktion nicht frühzeitig zu starten. Bei Verwendung der T4-DNA Ligase von Roche musste über Nacht bei 4°C ('sticky end'-Ligation) oder bei Raumtemperatur ('blunt end'-Ligation) inkubiert werden. Bei der Quick T4-DNA Ligase von Biolabs reichte eine 5minütige Inkubation bei Raumtemperatur aus, um die enzymatische Reaktion zu starten. Anschließend konnte der Ligationsansatz direkt in eine Transformationsreaktion eingesetzt werden (5.1.4).

#### 5.2.6.4 5'-Phosphorylierung von DNA-Fragmenten

PCR-Fragmente, die unmittelbar nach einer PCR-Reaktion in eine Ligation eingesetzt werden besitzen 'blunt' Enden und somit keine 5'-Phosphatgruppen. Mit Hilfe der T4-Polynukleotid-Kinase (PNK), einem Enzym, welches die γ-Phosphatgruppe von ATP an das 5'-Ende von DNA-Fragmenten überträgt, lassen sich so fehlende Phosphate anhängen. Für die Reaktion wurden 300 pMol DNA, 10 U Enzym und 1x T4-DNA-Ligase-Puffer (NEB), welcher 1 mM ATP enthält, eingesetzt. Nach einer 20minütigen Hitzeinaktivierung des Enzyms bei 65°C wurde die DNA durch Phenol/Chloroform-Extraktion (5.2.3.1) und Ethanolfällung (5.2.3.2) aufgereinigt.

#### 5.2.6.5 Ortsgerichtete in vitro Mutagenese

Die in dieser Arbeit durchgeführten *in vitro* Mutagenesen erfolgten an Anlehnung an das QuickChange<sup>®</sup> II Site-Directed Mutagenesis Kit der Firma Stratagene. Die einzelnen enzymatischen Reaktionen wurden entsprechend dem Protokoll des Herstellers unter Verwendung eigener Enzyme durchgeführt.

#### 5.2.6.6 *Single-colony* PCR (sc-PCR)

Um die Aufnahme neuer Spacer in das Plasmid-basierte CRISPR System nachweisen zu können, wurden *single-colony* PCR's aus induzierten Flüssigkulturen durchgeführt. Hierfür wurden die entsprechenden Plasmide (pCR003 WT und pCR003 Mutanten) in kompetente BL21AI/pUC-Kan18 Zellen transformiert und Kolonien von dieser Platte als üN-Kulturen angeimpft. Am nächsten Morgen wurden die Zellen 1:100 verdünnt in frisches 50 ml YT-Medium überführt, dass zuvor für die Induktion der *cas*-Gene mit IPTG und L-Arabinose versetzt wurde. Nach 18-24stündiger Wachstumszeit wurden Verdünnungen (1:50 und 1:500) dieser Kultur als *template* für die PCR Reaktion mit spezifischen Oligonukleotiden eingesetzt, um den genomischen bzw. plasmidären Einbau neuer Spacer im CRISPR-Array zu analysieren. Zeitgleich wurde ein neuer Zyklus gestartet, indem frisches YT-Medium 1:100 mit Zellen des vorhergehenden Zyklus versetzt wurde. Nach PCR-Amplifikation konnten die Ergebnisse auf einem 1,2%igen Agarosegel (5.2.4.1) überprüft werden.

Um den Einbau von inserierten Fragmenten nach Klonierungen zu bestätigen oder Klone für die Sequenzierung neuer Spacersequenzen zu detektieren, wurden alternativ Kolonien einer Agarplatte mit einem sterilen Zahnstocher angepickt, auf einer Masterplatte gesichert und in ein 25  $\mu$ l PCR-Ansatz getaucht. Anschließend wurde eine PCR-Reaktion gestartet (5.2.6.2) und die Produkte auf einem 1,2%igen Agarosegel (5.2.4.1) überprüft. Um die Plasmid-DNA zugängig zu machen bzw. das Annealing zu verbessern, wurde der Iniationsschritt bei dieser Art von PCR auf 10 Minuten erhöht.

## 5.2.6.7 Phosphorylierung von DNA-Oligonukleotiden (5´-Endmarkierung mit [γ-<sup>32</sup>P]-ATP)

Aufgrund ihres Herstellungsprozesses sind synthetische Oligonukleotide am 5'-Ende unphosphoryliert und können in einer Kinasierungsreaktion durch die T4-Polynukleotid-Kinase mit  $[\gamma^{-32}P]$ -ATP am 5'-Ende radioaktiv markiert und dann als Sonde eingesetzt werden. Hierzu wurden in einem 20 µl Ansatz 10 pmol des zu markierenden Oligonukleotids mit 1 U T4-Polynukleotid-Kinase 20-50  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]-ATP in 1 x Kinase-Puffer für 45 min bei 37°C inkubiert. Durch eine 10 minütige Hitzeinaktivierung bei 68°C wurde die T4-Polynukleotidkinase inaktiviert und das Oligonukleotid nach Zugabe von 1 µl Glykogen (20 µg/ml), 1/10 Volumen 7,5 M Ammoniumacetat (pH 5,0) und 100 µl eiskaltem, absoluten Ethanol (5.2.3.2) gefällt. Nach 45 minütiger Zentrifugation bei 12000 rpm und 4°C (Heraeus Biofuge A) wurde das erhaltene Pellet zweimal mit 80%igem Ethanol gewaschen, luftgetrocknet und in 20 µl TE-Puffer aufgenommen. Zum besseren Lösen des Pellets wurden die Proben anschließend für 1 h bei RT geschüttelt (Eppendorf Mixer 5432) und bis zur weiteren Verwendung bei 4°C gelagert.

1 x Kinase-Puffer:

50 mM Tris-HCl, pH 7,6 5 mM DTT 10 mM MgCl2 0,1 mM EDTA, pH 8,0

## 5.2.6.8 Radioaktive 3'-Endmarkierung von RNA

Bei der 3'-Endmarkierung von RNA wird durch die T4-RNA-Ligase RNA mit radioaktiv markiertem Cytosin ( $[5'-^{32}P]-pCp$ ) ligiert. Dabei wird in 20 µl 1 µg RNA mit 20 U T4-RNA-Ligase und 20 µCi  $[5'-^{32}P]-pCp$  in 10 % DMSO und 1 x Ligase-Puffer über Nacht bei 4°C inkubiert. Nach Zugabe von 1 µl Glykogen (20 µg/ml), 2 µl 3 M Natriumacetat (NaOAc, pH 5,3) und 80 µl eiskaltem, absolutem Ethanol wurde die RNA für 30 min auf Eis gefällt und durch Zentrifugation (4°C, 12000 rpm, 30-45 min) pelletiert. Danach wurde das Pellet zweimal mit 70 % Ethanol gewaschen, an der Luft getrocknet und in 100 µl TE-Puffer (4.6.1) gelöst.

#### 5.2.6.9 Radioaktive Markierung von DNA-Fragmenten

Die radioaktive Markierung von 5'-überhängenden Enden von doppelsträngigen DNA-Fragmenten erfolgte durch den Einbau von  $[\alpha^{32}P]$  dNTPs, mit Hilfe des Klenow-Enzyms der DNA Polymerase I. Ebenso ist auch eine radioaktive Markierung durch ein entsprechendes Nukleotid möglich.

Dafür musste folgender Reaktionsansatz angesetzt werden:

Reaktionsansatz:x  $\mu$ l DNA-Fragment (1  $\mu$ g)2.5  $\mu$ l 10x Klenow-Pufferje 1  $\mu$ l dCTP, dGTP, dTTP (10 mM)1  $\mu$ l [ $\alpha^{32}$ P] dATP (10  $\mu$ Ci/ $\mu$ l; 3000 Ci/mmol)1  $\mu$ l Klenow-Enzym (5U/ $\mu$ l)ad 25  $\mu$ l Aqua dest.

Nach 30 minütiger Inkubation des Reaktionsansatzes bei RT wurde dieser mit TE-Puffer (4.6.1) auf ein Gesamtvolumen von 50  $\mu$ l aufgefüllt, Phenol/Chloroform extrahiert (5.2.3.1) und mit Ethanol gefällt (5.2.3.2). Das luftgetrocknete Pellet wurde in 100  $\mu$ l TE-Puffer aufgenommen und zum besseren Lösen für 1-2 h bei RT geschüttelt (Eppendorf Mixer 5432). Bis zur weiteren Verwendung konnte es bei 4°C gelagert werden.

## 5.3 Proteinisolation und Analyse

## 5.3.1 Proteinisolation im analytischen Maßstab

Um die Überexpression eines Proteins zu testen wurden 100 ml YT-Medien (4.6.2) mit 1/100 einer üN-Kultur der gewünschten Probe in Gegenwart des selektiven

Antibiotikums angeimpft und bei 37°C auf einem Schüttler (New Brunswick Scientific Gio Gyrotory) inkubiert. Das Wachstum der Zellen wurde spektralphotometrisch überprüft (5.2.1.2) und eine Wachstumskurve erstellt. Beim Erreichen einer OD<sub>600</sub> von 0,5–0,6 wurde die Überexpression des Proteins durch die Zugabe von 0,1-0,5 mM IPTG bei BL21(DE3) Zellen induziert. Wurde der *E. coli* Stamm BL21 AI verwendet, musste zusätzlich noch L-Arabinose zum Medium gegeben werden. Unmittelbar vor IPTG Zugabe wurden 10 ml Zellkultur als Nullprobe entnommen und nach Induktion weitere Stundenwerte (1, 3, 5 Stunden oder üN-Werte) als Referenzen abgenommen. Damit alle Proben in etwa die gleichen Zellzahlen enthielten, wurden die Volumina der abgenommenen Proben den OD-Werten angepasst. Die Kulturen wurden bei 5000 rpm für 10 Minuten, 4°C sedimentiert (Heraeus Megafuge 1.0), der Überstand verworfen, das Zellpellet einmal mit 1 ml PBS-Puffer gewaschen, zentrifugiert und anschließend in 500  $\mu$  250 mM Tris-HCL, pH 7,8 aufgenommen. Der Zellaufschluss erfolgte am Ultraschallgerät (Labsonic U) mit folgenden Einstellungen:

Sonde:	Nadelsonde 40T
Leistung:	40 W
Repeating Duty Cycle:	0,9
Power Range Switch:	low

30 Sekunden Ultraschall 30 Sekunden auf Eis 3 Zyklen

Die Zelltrümmer wurden durch Zentrifugation (10 Minuten, 12000 rpm, Heraeus Biofuge 15) entfernt und die proteinhaltigen Überstände in neue Eppendorfgefäße überführt. Von diesen Proben wurden anschließend zwischen 10 und 20  $\mu$ l auf ein denaturierendes SDS-Gel (5.2.8.1) getrennt.

PBS-Puffer:

8 g NaCl 0,2 g KCl 10,1 ml 1M Na<sub>2</sub>HPO<sub>4</sub> mit HCl auf pH 7,4 einstellen ad 1 l mit Aqua dest.

#### 5.3.2 Proteinisolation im präparativen Maßstab (Aufreinigung der *E. coli* Cas-Proteine Cse1, Cse2, Cas7, Cas5 und Cas6e)

Für die Präparation der einzelnen Cas-Proteine CasA (Cse1), CasB (Cse2), CasC (Cas7), CasD (Cas5) und CasE (Cas6e) aus *E. coli* wurden die Bakterienstämme auf selektiven YT-Platten (4.6.2) ausgestrichen, davon 25 ml YT-Medien angeimpft und über Nacht bei 37 °C inkubiert. Anschließend wurden 2-3 800 ml YT-Medien 1:100 mit der entsprechenden üN-Kultur angeimpft und bei 37°C auf einem Rundschüttler

(New Brunswick Scientific Gio Gyrotory, ca. 200 rpm) bis zu einer OD<sub>600</sub> von 0,5-0,6 inkubiert. Nach Entnahme von 10 ml Zellkultur als Referenz vor Induktion, wurde die Expression der Proteine durch Zugabe von 0,1-0,2 mM IPTG induziert. Je nachdem welches Cas-Protein isoliert wurde, erfolgte ein Wechsel der Zellkulturen von 37°C auf Raumtemperatur (20-25°C). Anschließend wurden die Zellen für weitere 3 bzw. 5 Stunden oder üN bei der entsprechenden Temperatur inkubiert. Für die Zellernte wurden die Kulturen auf JA-10 Becher aufgeteilt und für 15 min bei 8000 rpm und 4°C sedimentiert (Beckmann Zentrifuge J2-21, JA-10 Rotor). Die Zellpellets wurden entweder bis zur weiteren Verwendung bei -20°C gelagert, oder in 20-50 ml Aufschlusspuffer, versetzt mit 0,2-0,5 mM PMSF als Protease-Inhibitor aufgenommen und mittels Ultraschall (Labsonic U, Braun Biotech Int.) aufgeschlossen.

Ultraschallaufschluss der Zellen:

Sonde:	Konische Sonde 50 T
Leistung:	160-180 W
Repeating Duty Cycle:	0,9
Power Range:	high

10 Sekunden Ultraschall 20 Sekunden auf Eis 10 Zyklen

Die Zelltrümmer wurden anschließend abzentrifugiert (30 min, 14000 rpm, 4 °C, Beckman Zentrifuge J2-21, JA17-Rotor), der proteinhaltige Überstand abgenommen und mittels Affinitätschromatographie (5.3.3) aufgereinigt.

Aufschlusspuffer (StrepII-Tag):	350 mM NaCl
	20 mM HEPES
	1 mM DTT, pH 7,5
	+ 0,2-0,5 mM PMSF

#### 5.3.3 Affinitätschromatographie

#### 5.3.3.1 Säulenchromatographie mittels Strep-Tactin Resin

Für die Proteinaufreinigung der einzelnen Cas-Proteine (Cse1, Cse2, Cas7, Cas5 und Cas6e) wurde das Econo-System der Firma Bio-Rad verwendet, dieses System besteht aus einer Pumpe (Pharmacia Biotech, Pump P-1), einem Fraktionssammler (Pharmacia LKB, Super Frac), einem Durchflussspektralphotometer (Pharmacia LKB Uvicord SII) sowie einem Schreiber (Gleitmann, Flatbed Recorder).

Das dabei verwendete Trägermaterial (Qiagen, Strep-Tactin Superflow Plus) enthält Strep-Tactin Resin, was eine starke Bindungsaffinität an Strep-tag, einem synthetischen Peptid des Biotins das aus 8 Aminosäuren (WSHPQFEK) besteht aufweist und die N-terminal markierten, rekombinanten Cas-Proteine bindet und sie somit an der Säule fixiert. Das Säulenmaterial wurde nach Herstellerangaben gepackt und die Säule mit Waschpuffer A äquilibriert.

Die bei der präparativen Proteinisolation (5.3.2) gewonnenen, proteinhaltigen Überstände wurden auf die Säule gepumpt (Flussrate 0,6 ml/min.) und der Durchfluss in 3-6 ml Fraktionen aufgefangen. Anschließend wurde die Säule mit Waschpuffer B gewaschen, bis die anhand des Schreiberprofils zu verfolgende Absorption der Probe bei 280 nm wieder die Basislinie erreicht hatte. Durch Zugabe des Waschpuffers B mit einer geringen Desthiobiotinkonzentration von 0,25 mM werden zunächst unspezifisch gebundene Proteine von der Säule gewaschen und freie Proteine entfernt. Dann erfolgte ein weiterer Waschschritt mit 0,5 mM Desthiobiotin. Diese stufenweise Erhöhung der Desthiobiotinkonzentration im Waschpuffer bewirkte die langsame Elution der StrepII-Tag markierten Proteine von der Säule. Zuletzt wurde Elutionspuffer C mit 2,5 mM Desthiobiotin auf die Säule gepumpt, um die noch auf der Säule verbliebenen Proteine zu eluieren.

Für die Regenerierung der Säule für weitere Nutzungen wurden zunächst 3x 50 ml *Aqua dest.* auf die Säule gepumpt. Dann wurde die Säule mit 0,5 M NAOH gewaschen und erneut mit *Aqua dest.* gewaschen. Abschließend musste die Säule für eine Wiederverwendung mit Aufschlusspuffer ohne die Zugabe von PMSF äquilibriert werden. Die Lagerung der Säule erfolgte bei 4°C. Von den gesammelten Fraktionen wurden je 20-40  $\mu$ l auf einem SDS-Gel (5.3.5.1) analysiert. Fraktionen die eine saubere Bande und hohe Proteinkonzentration aufwiesen wurden vereinigt, dialysiert und gegebenenfalls ankonzentriert (5.3.4).

Aufschlusspuffer:	350 mM NaCl 20 mM HEPES 1 mM DTT, pH 7,5 0,2-0,5 mM PMSF
Waschpuffer:	75 mM NaCl 20 mM HEPES 1 mM DTT, pH 8
Elutionspuffer:	75 mM NaCl 20 mM HEPES 1 mM DTT 2,5 mM Desthiobiotin, pH 8

### 5.3.3.2 Immobilisierte Affinitätschromatographie mit fusioniertem StrepII-Tag

Eine Alternative zur Säulenchromatographie mit dem Econo-System der Firma Bio-Rad stellte die Aufreinigung der StrepII-markierten Proteine über StrepTrap<sup>TM</sup>-Säulen mit Hilfe der ÄKTA Prime Plus der Firma GE Healthcare dar. Bei dieser Art der Aufreinigung wurde die Säule zunächst mit Waschpuffer A bis zum Einstellen der Basislinie äquilibriert und dann die im Aufschlusspuffer befindlichen Proteine über ein Schlauchsystem aufgezogen (0,5 MPa, Flussrate 0,3 ml/min). Erreicht die Probe über das Schlauchsystem die Säule, erfolgt der Durchbruch des Proteins und der Fraktionszähler, zum sammeln einzelner Fraktionen, wird eingeschaltet (Flussrate 3 ml/min.). Wird anhand des Computerprofils ein Abfall des Durchbruchs angezeigt, wird die Säule mit Waschpuffer B gewaschen. Um das an der Säule gebundene Protein zu eluieren, wird Elutionspuffer C in einem Schritt (*step elution*) mit einer Desthiobiotinkonzentration von 2,5 mM auf die Säule gepumpt. Um die Säule für weitere Aufreinigungen nutzen zu können, wurde diese erst mit *Aqua dest.* (5-10 ml), dann mit einer 0,5 M NAOH-Lösung (10 ml) und erneut mit *Aqua dest.* (15 ml) zum regenerieren gewaschen.

350 mM NaCl
20 mM HEPES
1 mM DTT, pH 7,5
0,2-0,5 mM PMSF
350 mM NaCl
20 mM HEPES
1 mM DTT, pH 8
75 mM NaCl
20 mM HEPES
1 mM DTT, pH 8
75 mM NaCl
20 mM HEPES
1 mM DTT
2,5 mM Desthiobiotin, pH 8

#### 5.3.4 Aufreinigung und Konzentrierung von Proteinen

#### 5.3.4.1 Dialyse von Proteinen

Zur Entsalzung der eluierten Proteine und um das Desthiobiotin aus den Proben zu entfernen, wurde eine Dialyse mit Dialyseschläuchen (SpectraPor, Porengröße 6000-8000, 20,4 mm) durchgeführt. Diese wurden vor Gebrauch 30 min in Na<sub>2</sub>CO<sub>3</sub>-Lösung gekocht, anschließend mit *Aqua dest.* gespült und erneut für 15 min gekocht. Bis zur Nutzung konnten sie bei 4 °C, in 50 % Ethanol gelagert werden und mussten vor Verwendung nochmals mit *Aqua dest.* gespült werden.

Die Dialyse der jeweiligen Cas-Proteine erfolgte über Nacht (üN) bei 4°C gegen 1 Liter Dialysepuffer.

Dialysepuffer:

75 mM NaCl 20 mM HEPES 1 mM DTT, pH 7,5 + 0,5% Glycerin

### 5.3.4.2 Konzentrierung von aufgereinigten Proteinen mittels Amicon-Zentrifugation

Um zuvor mittels Affinitätschromatographie aufgereinigte Proteine für weitere Analysen nicht-denaturierend aufzukonzentrieren, konnten diese mit Hilfe von Amicon Ultrafiltern durch Zentrifugation aufgrund ihrer Molekülgröße aufgetrennt werden. Hierzu wurden die Ultrafilter zunächst mit dem entsprechenden Bindepuffer (500  $\mu$ l) der Probe für 30 Min. bei 4000 rpm, 4°C durch Zentrifugation äquilibriert. Anschließend konnte die gewünschte Probe in den Zentrifugalfilter gegeben und zentrifugiert werden (90 Min, 4000 rpm, 4°C).

#### 5.3.5 Gelelektrophorese von Proteinen

#### 5.3.5.1 Diskontinuierliche SDS-PAGE

Die diskontinuierliche Gelelektrophorese wurde zur Auftrennung von Proteinen in Abhängigkeit ihres Molekulargewichtes angewendet. Bei dieser Methode wird ein engporiges Trenngel (312 mm x 180 mm x 1 mm) mit einem weitporigem Sammelgel (310 mm x 50 mm x 1 mm) (B x H x T) miteinander kombiniert und als denaturierendes Agens Natriumdodecylsulfat (SDS) eingesetzt. Durch hohe Affinität als anionisches Tensid bindet das SDS an die Proteine, überdeckt ihre Eigenladung im Gel, so dass sie negativ geladen sind und im elektrischen Feld umgekehrt proportional zu ihrer Molekülmasse wandern.

Vor dem Auftragen wurden die Proteinproben mit ¼ Volumen 4 x SDS-Probenpuffer (4.6.1) sowie 2  $\mu$ l  $\beta$ -Mercaptoethanol versetzt, 5 min. bei 96 °C denaturiert und anschließend auf Eis abgekühlt. Bei großen SDS-Gelen erfolgte die Trennung der Proben bei 120 V, bei kleinen Gelen bei 60 V üN. Als Laufpuffer diente 1x SDS-Puffer und Visualisierung wurde das Gel anschließend mit Coomassie (5.3.6.1).

Lösung A:	30 % Acrylamid-/ Bisacrylamid-Lösung (30:1)
Lösung B:	1,5 M Tris-HCl, pH 8,8
Lösung C:	10 % (w/v) SDS
Lösung D:	0,5 M Tris-HCl, pH 6,8
Trenngel (12 %):	28 ml Lösung A 17,5 ml Lösung B 0,7 ml Lösung C ad 70 ml mit <i>Aqua dest.</i> 70 μl TEMED 700 μl 10 % (w/v) APS
Sammelgel (6%):	4 ml Lösung A 5 ml Lösung B 0,2 ml Lösung C ad 20 ml mit <i>Aqua dest.</i> 20 μl TEMED 200 μl 10 % (w/v) APS

#### 5.3.6 Nachweis von Proteinen in Polyacrylamidgelen

#### 5.3.6.1 Coomassie-Färbung

Für einen quantitativen Nachweis von Proteinen in Polyacrylamidgelen wurde eine Coomassie-Färbung mit dem Farbstoff "Coomassie Brillant Blue R250" für 30 bis 60 Minuten auf einem Horizontalschüttler (Gerhardt, Schüttelmaschine LS10) bei RT durchgeführt. Danach wurde das Gel solange in Entfärbelösung geschüttelt, bis der Hintergrund nahezu entfärbt und die Proteinbanden deutlich zu erkennen waren. Die Gele wurden anschließend luftdicht in Folie eingeschweißt und gescannt (UMAX, Astra 4000U). Mit dieser Färbemethode können Proteinbanden ab 300 ng
nachgewiesen werden, zudem ermöglicht es diese Methode eine quantitative Aussage über die Proteinkonzentration im Gel zu treffen. Wenn die Coomassie-Färbung nicht sensitiv genug war, um Proteine sichtbar zumachen, wurden die Gele mit Silber nachgefärbt (5.2.5.3).

Färbelösung:	50 % (v/v) Methanol
	10 % (v/v) Essigsaure
	0,1 % (w/v) Coomassie Brillant Blue R250
Entfärbelösung:	10 % (v/v) Methanol
	10 % (v/v) Essigsäure

### 5.4 Spezielle molekularbiologische Methoden

# 5.4.1 *In vitro* Transkription zur präparativen Gewinnung von pre-crRNA (RiboMAX)

Die Herstellung von pre-crRNA im präparativen Maßstab erfolgte durch eine *in vitro* Synthese mit Hilfe des RiboMAX-Systems von Promega. Die dafür verwendete Phagenspezifische T7-RNA-Polymerase benötigt für die Transkriptionsinitiation einen T7-Promoter. Dieser befindet sich auf dem Plasmid pZA-T7S1+, der den *E. coli* CRISPR Array R1-R4 vor dem *kan*-Gen kodiert besitzt. Durch Linearisierung des Vektors mit dem Restriktionsenzym *Sspl downstream* des pre-crRNA Transkripts kann die RNA-Polymerase von dem *template*-Plasmid abfallen und eine neue Transkription initiieren. Diese Art der *in vitro* Transkription wird *"run-off"*- oder *"multiple-round run off"*- Transkription genannt.

Für die Transkription wurden jeweils 5-10 Ansätze mit einem Reaktionsvolumen von je 30  $\mu$ l angesetzt. In jeden Ansatz wurden 2-3 pmol zuvor linearisiertes Plasmid, 3 mM 4 NTP-Mix, 1,5  $\mu$ l Pyrophosphatase (0,1U/ $\mu$ l), 15 U RNasin und 288 U T7-RNA-Polymerase in 1x RiboMAX-Puffer sowie 6 mM DTT pipettiert. Die Ansätze wurden bei 37°C für eine Stunde inkubiert und anschließend auf einem analytischen Agarosegel (5.2.4.1) kontrolliert. War die Transkription erfolgreich wurden die Proben vereinigt, auf einem präparativen Agarosegel aufgetrennt und über mit Glaswolle gefüllte Eppis (5.2.3.4) eluiert. Das Eluat wurde Phenol/Chloroform (5.2.3.1) extrahiert und Ethanol gefällt (5.2.3.2).

5x RiboMAX-Puffer:

400 mM Hepes-KOH, pH 7,5 60 mM MgCl<sub>2</sub> 10 mM Spermidin 200 mM DTT

### 5.4.2 Southern Blot Analysen

Zur Detektion von in vivo Intermediaten während der Integration neuer Spacer-DNA wurden Southern Blot Analysen mit dem pCR003 WT Plasmid, sowie mit mutierten Versionen des WT Plasmides nach Induktion der Cas1/Cas2 Expression durchgeführt. Dafür wurden je 5 µg der durch Midi-Präp. (5.2.2.2) isolierten Plasmide mit verschiedenen Restriktionsenzymen (EcoRI und KpnI) hydrolysiert und auf einer 10% dPAGE getrennt (5.2.4.3). Die elektrophoretisch aufgetrennten Fragmente wurde anschließend auf eine Nitrocellulose-Membran (GE, Amersham Hybond<sup>™</sup>-N<sup>+</sup>) übertragen und in eine Elektroblot-Apparatur (Owi, Panther™ Semidry Electroblotter, Model HEP-1) überführt. Diese Apparatur nutzt die negative Ladung der DNA, wobei das Gel auf 2 Lagen, mit 1x TBE Laufpuffer befeuchtetem Whatman 3 MM Papier auf der Seite der Kathode liegt. Darüber befindet sich die Membran und weitere 2 Lagen feuchtes Whatman 3 MM Papier, auf welche die Anode festgeschraubt wurde. Der elektrophoretische Blot erfolgte für 60 min bei 600 mA. Nach dem Transfer wurden die auf die Membran übertragenen DNA-Fragmente durch Bestrahlung mit UV-Licht in einem UV-Crosslinker (Stratagene, UV Stratalinker<sup>™</sup> 1800) kovalent auf der Nylonmembran fixiert und für 2 Std. im Hybridisierungsofen bei 80°C inkubiert. Anschließend konnten sie für die Hybridisierung mit <sup>32</sup>[P]-markierten Sonden eingesetzt werden. Dafür wurde die Membran mit der DNA-haltigen Seite nach innen in eine Glasröhre überführt und 2 h mit 20 ml einer Prähybridisierungslösung bei der gewählten Hybridisierungstemperatur (5-10°C unterhalb der Schmelztemperatur des Oligonukleotids) im Hybridisierungsofen vorinkubiert. Im Anschluss daran erfolgte die Hybridisierung mit 20 ml Hybridisierungslösung und 5 µl des radioaktiv markierten Oligonukleotids über Nacht. Zum Entfernen unspezifisch gebundener Sonden wurde die Membran am nächsten Morgen 3 x 30 Minuten mit je 50 ml Waschlösung gewaschen.

Um nacheinander verschiedene Sonden gegen die gleiche DNA hybridisieren zu können, konnte die alte Sonde durch dreimaliges Waschen für 1 Std. mit 50 ml *Aqua dest.* bei 80°C entfernt und anschließend mit einer neuen Sonde hybridisiert werden.

Hybridisierungslösung:	5 x SSC-Puffer 7,5 % SDS 50 μg/ml Heringsperm-DNA
Waschlösung:	5 x SSC-Puffer 0,1 % SDS
20 x SSC-Puffer:	300 mM Natriumcitrat 3 M NaCl, pH 7,0

### 5.4.3 Sequenzierungsreaktion mit Sequenase 2.0 Polymerase

Zur Verifizierung der durch Southern Blot Analysen (5.4.2) detektierten Zwischenstufen wurden Sequenzierungsreaktionen mit der Sequenase 2.0 der Firma USB (4.5.2) durchgeführt. Dabei eignet sich für die quantitative Analyse von Transkriptmengen die Methode der Primer-Extension Reaktion, die in dieser Arbeit in Anlehnung an die von Stern beschriebenen Methode durchgeführt wurde (Stern, Moazed et al. 1988). Hierbei wird die zu analysierende DNA von der T7 DNA-Polymerase als *template* genutzt und ein komplementärer Gegenstrang synthetisiert. Die dabei entstehenden Produkte geben durch ihre Länge indirekt die Fragmentlänge der DNA wieder und können über eine dPAGE (5.2.4.3) aufgetrennt werden.

Die Sequenzierungsreaktionen wurden mit aufgereinigter, nicht denaturierter Plasmid-DNA des pCR003 WT Plasmides und unter Verwendung des Sequenase Version 2.0 DNA Sequencing Kit der Firma USB durchgeführt. Bei dieser DNA-Polymerase handelt es sich um eine genetisch veränderte Form der T7 DNA-Polymerase mit hoher Prozessivität und fehlender 3'-5'-Exonukleaseaktivität. Als Oligonukleotide wurden das PE-Leader-pCR001-fw und das pCR001-Blot-C-S1-rv Oligo eingesetzt. Für die Reaktion musste zunächst eine Hybridisierungsreaktion an die zu untersuchende Plasmid-DNA mit y-<sup>32</sup>[P]ATP markiertem Oligonukleotid durchgeführt werden. Dafür wurden in einem Volumen von 10 µl, 2 µg Plasmid-DNA mit 1 µl radioaktiv markiertem Oligonukleotid (5.2.6.7) und 2 µl 5-fach konzentriertem Sequenase-Reaktionspuffer erst für 2 Min bei 98°C inkubiert, anschließend für weitere 2 Min bei 65°C und dann über Nacht langsam im Heatblock auf Raumtemperatur abgekühlt. Am nächsten Morgen wurde der Hybridisierungsansatz zur Fortsetzung der Sequenzierungsreaktion mit 1 µl 0,1 M DTT, 1,65 µl eines 300 µM konzentrierten dNTP-Mixes und 2 µl der Sequenase 2.0 Polymerase (1:8 verdünnt in eiskaltem Enzym-Dilution Puffer) versetzt. Nach 10minütiger Inkubation bei 37°C wurde die Reaktion durch Zugabe von 5 µl Stopp-Lösung gestoppt. Die so entstandenen Fragmente geben durch ihre Länge, abhängig vom verwendeten Oligonukleotid, die Länge der spezifischen Cas1 Schnittstelle wieder und können über dPAGE aufgetrennt werden. Als Kontrolle wurden je 1 µl der markierten Oligonukleotide anstelle von Plasmid-DNA mit TE-Puffer aufgefüllt eingesetzt. Vor dem auftragen auf eine 10 % dPAGE (5.2.4.3) wurden die Proben 5 min bei 96°C denaturiert. Die Visualisierung der Banden erfolgte mittels Autoradiographie (5.2.5.2).

5 x Sequenase Reaction Buffer:	200mM Tris-HCl (pH 7.5) 100mM MgCl <sub>2</sub> 250mM NaCl
Sequenase Dilution Buffer:	10 mM Tris-HCL (pH 7.5) 5 mM DTT 0,1 mM EDTA

4 dNTP-Lösung:

### 5.4.4 Retardierungsanalysen (Gelshift-Experimente)

Die Methode der Verzögerungsgelelektrophorese oder Retardierungsgelelektrophorese kann zur qualitativen und quantitativen Darstellung von Komplexen aus DNA bzw. RNA und Proteinen herangezogen werden und beruht auf einer geringeren Mobilität von Protein-DNA bzw. RNA-Komplexen im Vergleich zu freier Nukleinsäure. Ursachen für diese Mobilitätsverzögerung sind das erhöhte Molekulargewicht des Komplexes gegenüber der freien DNA bzw. RNA, sowie eine Veränderung der Konformation der Nukleinsäure durch die Einführung von proteininduzierten Krümmungen.

Für die Analyse der Bindung der einzelnen Cas-Proteinen CasA, CasB und CasC wurden die Retardierungsanalysen mit radioaktiv markierten Nukleinsäure-Fragmenten (5.2.6.7, 5.2.6.8, 5.2.6.9) durchgeführt. Dafür wurden in einem Volumen von 15  $\mu$ l 2–15 nM radioaktiv markierte DNA oder RNA mit 10 x Bindungspuffer und steigenden Konzentrationen der Cas-Proteine (0,2-8  $\mu$ M) inkubiert. Verdünnungen einzelner Cas-Proteine erfolgte in Waschpuffer (5.3.3.1). Um elektrostatische Wechselwirkungen zu unterdrücken, enthielten die Ansätze zusätzlich 20-50 ng/ $\mu$ l Heparin. Für die Komplexbildung wurden die Ansätze 10 min. bei RT inkubiert, dann Heparin als Inhibitor hinzugegeben und anschließend mit 5  $\mu$ l 2 x TBE-Probenpuffer (4.6.1) versehen. Die Trennung erfolgte für 2-3 Stunden auf einem nativen 5 %igen Polyacrylamidgel (5.2.4.2) bei 30 mA. Als Elektrophoresepuffer wurde 0,5 x TBE-Puffer verwendet. Nach der Elektrophorese wurden die aufgetrennten Proben autoradiographiert (5.2.5.2).

Nach der Klenow-Reaktion wurden die radioaktiv markierten DNA-Fragmente in 1x KGlu80-Puffer 3 min bei 96°C denaturiert und anschließend zur Renaturierung langsam im Heizblock auf RT abgekühlt. Durch diese anschließende Denaturierung und Renaturierung der Fragmente konnte der bei der Klenow-Reaktion entstehende Anteil an einzelsträngigem Fragment in doppelsträngiges Fragment überführt werden.

10x Klenow-Puffer:	500 mM Tris/HCl, pH 7.2
	100 mM MgSO4
	1 mM DTT

Im Gegensatz zu den Verzögerungsanalysen mit radioaktiv markierten Nukleinsäure-Fragmenten wurden auch das Bindungsverhalten von unmarkierter, kalter pRNA an die jeweiligen Cas-Proteine untersucht. Dabei wurden aufgrund der geringeren Sensitivität im Vergleich zu den radioaktiven Bindestudien, höhere Mengen an RNA und Protein eingesetzt. 1,3 μg pRNA (de novo RNA-Oligo, 20 nt)1,2-3 μg der jeweiligen Cas-Proteine (*A*, *B* und *C*)

in 20  $\mu$ l Ansätzen mit 1x Bindungspuffer, 1 $\mu$ g Heparin und Waschpuffer. Die Ansätze wurden für 15 min bei Raumtemperatur (RT) inkubiert, mit 5  $\mu$ l 2x TBE-Probenpuffer versehen und auf einem 5% natives PAA-Gel getrennt. Anschließend wurde das Polyacrylamidgel mit Ethidiumbromid gefärbt (5.2.5.1) und das Ergebnis am UV-Transilluminator (Herolab UVT 2035) ausgewertet.

10 x Bindungspuffer:

500 mM Tris-HCl, pH 7,4 700 mM KCl 150 mM NaCl 10 mM EDTA 100 mM β-Mercaptoethanol

### 6 Anhang

1	gcg	gta	ccat	cg	gaa	aca	aaga	ati	cago	tgat	ct	tta	ata	at	aag	gaa	atgt	tac	att	aago	; ti	tgg	tgg	gtt	gt	ttt	at	gg	gaa	aaa	atg	c	ttt	aag	aac	a
	cgc	at	ggta	ac	ctt	tgt	ttct	taa	atco	acta	ga.	aat	tat	ta	tte	ctt	taca	ato	taa	atte	aa	+1	acc	caa	ca	aaaa	ata	cc	ctt	ttt	tac	g	aaa	ttc	ttg	It
	>>.													svi	the	tis	cher	CRI	SPR	Ari	av.															>
	а	1	y I	n	g	n	k	e	1	a c	1	1	-3	-	-	g	n	v	τ	1 1		1	v	g	c	Í	У	1	g	k	k	C	Í	k	n	L.
101	aat	gt	ata	ct	ttt	aga	gagt	tco		cgco	ag	cgg	gga	ta	aac	cgg	cgac	cad	acc	cgto	t ct	tgt	gga	tee	to	tace	gcg	ag	tto		gcg	C	cag	cgg	gga	t
	tta	aca	tat	ga	aaa	tct	ctca	ag	gggd	gegg	R1	gcc	cct	at	ttg	gcc	gctg	gto	tgg	ggcag	ga Si	aca 1	cct	agg	ag	atgo	cgc	tc	aag	1999	R2	g	gtc	gaa	cct	a
	>													syr	the	tis	cher	CRI	SPE	A-Ari	ay.															>
	k	C	i		1	1	e s		s p	r	q	r	g		- 1	τ	g d	ł	t	r	I	p	v	d	p	1	r	e	Í	f r	a	1	p	a	g	
201	aaa	acc	ggt	cg	acg	agt	tccc	cgo	egeo	ageg	1 99	gat	aaa	CC	gggt	tac	cgag	tto	:000	gege	: 08	ago	ggg	gat	aa	acco	aa	gc	tta	att	aac	a	acc	tgg	gcc	t
	ttt	gg	cca	gc	tgc	tca	aggg	gc	gegg	togo R3	c c c	cta	ttt	gg	ccca	atg	gctc	aag	ggg	acaca	gt R4	tcg 4	ccc	cta	tt	tgga	tt	cg	aat	taa	ittg	t	tgg	acc	cgg	a
	>													svi	the	tis	cher	CRI	SPR	A-Ari	av.															>
	i	n	r	3	τ	3	3	p	r	q	r	g	-	t	g	У	r	v	p	r	a	9	g	d		k I	þ	k	1	n	-		d 1	p	g	p
301	cat	gg	gaat	tt	ccg	ctc	acto	cco	gct	ttco	ag	tcg	gga	aa	cct	gtc	gtgc	cag	cto	gcatt	aa	aca	tgg	tca	ta	gctq	gtc	tc	gag	idds	acc					
	gta	acc	cgga	aa	ggc	gag	tgac	ggg	goga	aagg	to	agc	cct	tt	gga	cag	cacg	gto	gad	gtaa	t t	tgt	acc	agt	at	cgad	cag	ag	cto	cct	gg					
	>											syn	the	tis	sche	r C	RISP	R-Ar	ray	1											>>					
	h	1	g 1	p	3	a	h	C	p	1 5	5	3	r	e	τ	C	r	a	3	c i		n	m	V	i	a	V	2	3	r	d					

### Abbildung 6.1: Sequenz des synthetischen CRISPR-Arrays aus E. coli.

Gezeigt ist die Sequenz des synthetischen CRISPR-Arrays die mittels PCR und den Oligonukleotiden pMA-NcoI-fw und pMA-XhoI-rv (4.4.1) aus dem Plasmid pLS1(+) amplifiziert wurde. Der Transkriptionsstart ist mit +1 angegeben, die identischen Repeatsequenzen rot markiert und die 32 bp große Spacersequenz in blau markiert.



## Abbildung 6.2: Schema des Verfahrens zur Isolierung von *E. coli* Zellen mit einem ´single´ Typ des Plasmid-erweiterten CRISPR-Arrays.

Zeigen die *E. coli* BL21 AI/pUC18-Kan Zellen nach Transformation mit dem pCR003 Plasmid (schwarze Kreise) während der PCR-Analyse den Einbau von Spacern, werden sie auf selektive Agarplatten ausgestrichen und in einer ersten PCR-Reaktion nach Klonen, die neue Spacer in den CRISPR-Array des Plasmids aufgenommen haben gesucht (rote Kreise). Die Plasmid-DNA dieser Klone wird isoliert und in kompetente XL-1 Zellen transformiert, um in einer zweiten PCR-Reaktion, nur Klone mit 'single' Typ des erweiterten Plasmid-Arrays für eine anschließende Sequenzierung zu erhalten (entnommen aus (Arslan, Hermanns et al., in Revision)).

gcgtaccatg	gaaacaaaga	attagctgat (	ctttaataat	aaggaaatgt	tacattaagg	ttggtgggtt	gtttttatgg	gaaaaaatgc	tttaagaaca
cgcatggtac	ctttgtttct	taatcgacta (	gaaattatta	ttcctttaca	atgtaattcc	aaccacccaa +1	caaaaatacc	ctttttacg	aaattettgt
>>				E. coli CR	ISPR-Array.				
a y h	g n k	e l a d	1	- g n	v t l r	l v g	c f y	g k k c	f k n
aatgtatact	tttagagagt	tggttgcgcc	agcggggata	aaccggcgac	cacaccogto	ctgtggatcc	tctacgcgag	ttccccgcgc	cagcggggat
ttacatatga	aaatctctca	accaacgcgg 1	tegecectat	ttggccgctg	gtgtgggcag	gacacctagg S1	agatgcgctc	aaggggcgcg R2	gtcgccccta
>				E. coli CR	ISPR-Array.				
k c i	1 1 e s	w l r	q r g	- t g d	h t r	p v d	p l r e	fpa	p a g
aaaccggtcg	acgagtteec	cgcgccagcg (	gggataaacc	gggtaccgag	tteccegege	cagcggggat	aaaccgaagc	ttaattaaca	acctgggcct
tttggccagc	tgctcaaggg	gcgcggtcgc ( R3	ccctatttgg	cccatggete	aaggggcgcg	gtcgccccta R4	tttggcttcg	aattaattgt	tggacccgga
>				E. coli CR	ISPR-Array.				>
i n r s	t s s	prq :	rg-t	g y r	vpr	a s g d	k p k	1 n -	d b d b
catgggcctt	ccgctcactg	cccgctttcc a	agtcgggaaa	cctgtcgtgc	cagctgcatt	aacatggtca	tagctgtctc	gagggacc	
gtacccggaa	ggcgagtgac	gggcgaaagg 1	tcagcccttt	ggacagcacg	gtcgacgtaa	ttgtaccagt	atcgacagag	ctccctgg	
>			E. col	Li CRISPR-A:	rray			>>	
h g p	s a h	c p l s	s r e	tcr	a s c i	n m v	i a v	s r d	

#### Abbildung 6.3: Sequenzierungsergebnis der Mutante pCR003 RM1.

Zu sehen ist ein Sequenzausschnitt nach Sequenzierung des CRISPR-Arrays des pCR003 RM1 Plasmids mit der eingebauten Mutation im C-Cluster (blaue Basen) des ersten Repeats. Die anderen Repeats (R2-R4 rot), sowie die bereits vorhandene Spacersequenz (S1, grün) sind auch zu sehen.

```
gcgtaccatg gaaacaaaga attagctgat ctttaataat aaggaaatgt tacattaagg ttggtgggtt gtttttatgg gaaaaaatgc tttaagaaca
cgcatggtac ctttgtttct taatcgacta gaaattatta ttcctttaca atgtaattcc aaccacccaa caaaaatacc cttttttacg aaattcttgt
                                                +1
               .....E. coli CRISPR-Array.....
                                                                      ....
<<....
 t g h f c l i l q d k i i l f h
                                       - ml t p p n n k h s f i s - s c
aatgtatact tttagagagt tggttgcgcc agcaaccata aaccggcgac cacacccgtc ctgtggatcc tctacgcgag ttcccccgcgc cagcggggat
ttacatatga aaatototoa accaacgogg togttggtat ttggcogotg gtgtgggcag gacacotagg agatgogoto aagggggogg gtogocoota
                      R1
                                               S1
                                                                    R2
<.....E. coli CRISPR-Array.....
 iyv kls ntag avm frr gcgd qpd evr tgra
                                                                        l p s
aaaccggtcg acgagtteee egegecageg gggataaace gggtacegag tteeeegege cageggggat aaacegaage ttaattaaca acetgggeet
tttggccage tgeteaaggg gegeggtege ecetatttgg eecatggete aaggggegeg gtegeeeeta tttggetteg aattaattgt tggaecegga
                   R3
                                               R4
 .....E. coli CRISPR-Array.....
                                                                          .....
lgt ssng aga pifrtgl egr
                                               wrp yvsa
                                                               - n v
                                                                       v q a
catgggcett cegetcaetg eccgetttee agtegggaaa ectgtegtge cagetgeatt aacatggtea tagetgtete gagggaee
gtacccggaa ggcgagtgac gggcgaaagg tcagcccttt ggacagcacg gtcgacgtaa ttgtaccagt atcgacagag ctccctgg
```

#### Abbildung 6.4: Sequenzierungsergebnis der Rückmutante pCR003 RM1-R.

Zu sehen ist ein Sequenzausschnitt nach Sequenzierung des CRISPR-Arrays des pCR003 RM1-R Plasmids mit der eingebauten Mutation an den Positionen 6-9 des ersten Repeats (blaue Basen). Die anderen Repeats (R2-R4 rot), sowie die bereits vorhandene Spacersequenz (S1, grün) sind auch zu sehen.

gcgtaccatg gaaac	aaaga attagctgat	c ctttaataat aaggaaatgt	tacattaagg ttggtgggtt	gtttttatgg g	jaaaaaatgc tttaagaaca
cgcatggtac ctttg	tttct taatcgacta	a gaaattatta ttootttaca	atgtaattcc aaccacccaa +1	caaaaatacc c	sttttttacg aaattcttgt
>>		E. coli CR	ISPR Array		
a y h g n	k e l a d	1 1 g n	v t l r l v g	c f y g	j k k c f k n
aatgtatact tttag	agagt tggggggcgcc	agcocccata aaccggcgac	cacaccogto ctgtggatco	tctacgcgag t	tccccgcgc cagcggggat
ttacatatga aaatc	tetca accecegegg	g tegggggtat ttggeegetg R1	gtgtgggcag gacacctagg S1	agatgcgctc a	agggggggg gtegeeeta R2
>		E. coli CR	USPR Array		
k c i l l	e s w g r	q p p - t g d	i h t r p v d	p l r e	fpapag
aaaccggtcg acgag	ttccc cgcgccagcg	g gggataaacc gggtaccgag	ttccccgcgc cagcggggat	aaaccgaagc t	taattaaca acctgggcct
tttggccagc tgctc	aaggg gcgcggtcgc R3	ccctatttgg cccatggete	aaggggggggg gtcgcccta R4	tttggcttcg a	attaattgt tggaccogga
>		E. coli CR	ISPR Array		
inrst	s s p r q	rg-tgyr	v prasgd	k p k	ln-qpgp
catgggcctt ccgct	cactg cccgctttcc	c agtogggaaa cotgtogtgo	cagctgcatt aacatggtca	tagctgtctc g	jagggacc
gtacccggaa ggcga	gtgac gggcgaaagg	g tcagcccttt ggacagcacg	gtcgacgtaa ttgtaccagt	atcgacagag c	tccctgg
>		E. coli CRISPR A	rray		>>
h g p s a	h c p l s	s s r e t c r	a s c i n m v	i a v s	s r d

#### Abbildung 6.5: Sequenzierungsergebnis der Mutante pCR003 RM2.

Zu sehen ist ein Sequenzausschnitt nach Sequenzierung des CRISPR-Arrays des pCR003 RM2 Plasmids mit der eingebauten Mutation an den Positionen 6-9 und 18-21 des ersten Repeats (blaue Basen). Die anderen Repeats (R2-R4 rot), sowie die bereits vorhandene Spacersequenz (S1, grün) sind auch dargestellt.



## Abbildung 6.6: Southern Blot Analyse der Leaderregion des synthetischen CRISPR-Arrays mit Sonden gegen die Spacersequenz.

Gezeigt sind die Autoradiogramme der Southern Blot Analysen von 1 µg *EcoRI* hydrolysierter Plasmid-DNA des pCR003 WT und der Mutante RM1 nach dPAGE. Die Zellen wurden für die Isolierung der Plasmid-DNA nach 18 stündigem Wachstum mit und ohne Induktion der *cas1/cas2*-Expression geerntet. Die zuvor hybridisierten Primer (pCR001-Blot-EcoRI-rv und pCR001-Blot-EcoRI-fw) komplementär zu beiden DNA-Strängen der Leaderregion, wurden durch mehrmaliges Waschen (3 x 30 Min.) bei 80°C entfernt. Anschließend wurde die Membran mit neuen Sonden die gegen die Spacersequenz des synthetischen CRISPR-Arrays gerichtet sind hybridisiert (pCR001-Blot-S1-fw und pCR001-Blot-C-S1-rv). Die Spuren 4 zeigen einen radioaktiv markierten Marker (pUC19/MspI), der als Längenstandard dient.

at	gtt	tt	cgg	gga	cat	acaa	Cta	acaa	atgg	jt	aaag	aaa	atgt	ato	ttt	caaa	ta	ttg	tca	gc	gact	tata	icca	aaa	aag	ttat
ta	caa	aaa	gcc	cct	gta	tgtt	gat	tgtt	taco	ca	ttto	ttt	taca	tag	aaa	gttt	at	aac	agt	cg	ctga	atat	ggt	ttt	ttc	aata
>.								all and					.Sag	Cas1						-		1212				>
1	m	f	3	g	t	У	n	Y	n	a	k	e	m	Y	1	3	n	î	v	3	d	Y	t	k	k	v
ta	agt	ccg	cta	aat	agt	gatg	gga	aatç	ggaa	at	tccg	ga	gtag	ctg	aaa	tgag	tt	atc	ggt	at	atgo	gaa	atga	ttt	taa	tgtt
at	tca	agc	gat	tta	tca	ctac	CC	tad	cctt	a	aggo	ct	catc	gac	ttt	actc	aa	tag	cca	ta	taco	gett	act	aaa	att	acaa
													Ins	erti	on											
>.						.Sag	Casi	1					>>													
i	k	3	1	n	3	d	g	n	g		i p	2 6	- 19													
															>	>				.Sa	agCas	32.				>
																m	3	Y	r	Y	m	r	m	î	1	m
tg	ata	atg	cct	act	gaa	acag	ca	gaaq	gaad	g	gaag	gco	gtat	cgt	aag	ttta	ga	aag	ttt	ct	ctt	gago	gaa	ggo	ttt	atca
ac	tat	ac	gga	tga	ctt	tgtc	gt	ctto	cttg	c	ctto	cg	cata	gca	ttc	aaat	Ct	ttc	aaa	iga	gaad	tco	ctt	cco	jaaa	tagt
>.													.Sag	Cas2	1994					0.00						>

Abbildung 6.7: Sequenzierungsergebnis nach Einführung der 8 Basen aus *E. coli* in das pACYC-SagCas1Cas2 Plasmid.

Zu sehen ist ein Sequenzausschnitt der *cas1-cas2* Gene aus *S. agalactiae* nach Sequenzierung und Einführung der 8 Basen aus *E. coli* (rote Basen), die in das pACYC-SagCas1Cas2 Plasmid durch eine PCR-Reaktion eingeführt wurden.



## Abbildung 6.8: Schematische Darstellung der Klonierungsstrategien mit unterschiedlichen Plasmidvarianten in BL21 AI.

Dargestellt sind die unterschiedlich konstruierten Plasmide in BL21 AI transformiert, die für das heterologe Integrationssystem in *E. coli* kloniert wurden. Welche Integrations-Komponenten aus *S. agalactiae* verwendet wurden, ist nachfolgend erläutert.

**1)** pCR004/-5 Plasmid  $\rightarrow$  pSagCas1Cas2Csn2 Vektor mit CRISPR Array aus *S. agalactiae.* Transformiert mit dem pUCKan-18 Plasmid in BL21 AI.

**2)** pCR010 Plasmid  $\rightarrow$  pET-52(b)+ Derivat mit den *cas1-cas2* Genen aus *S. agalactiae.* Transformiert mit dem pCR003 WT Plasmid in BL21 AI.

**3)** pCR014+Csn2 Plasmid  $\rightarrow$  pACYC-Duet1 Derivat mit dem CRISPR Array und *csn2* Gen aus *S. agalactiae.* Transformiert mit dem pCR010 Plasmid in BL21 AI.

**4)** pCR012 Plasmid  $\rightarrow$  pET-52(b)+ Derivat mit den *cas1-cas2* Genen aus *S. agalactiae* nach Insertion der 8 Basen aus *E. coli*. Transformiert mit dem pCR014+Csn2 Plasmid in BL21 AI.

**5)** pEcoSagCas1Cas2Csn2/-K132 Plasmid  $\rightarrow$  pSagCas1Cas2Csn2/-K132 Plasmide mit kloniertem CRISPR Array aus *E. coli*. In BL21 AI transformiert mit dem Vektor pET-EcoCas1Cas2, einem pET-52(b)+ Derivat der *cas1-cas2* aus *E. coli* kloniert enthält.

**6)** pCR014 Plasmid  $\rightarrow$  pACYC-Duet1 Derivat mit CRISPR-Array aus *S. agalactiae.* In BL21 AI transformiert mit dem Vektor pET-SagCas1Cas2Csn2, einem pET-52(b)+ Derivat der *cas1-cas2-csn2* aus *S. agalactiae* kloniert enthält.

**6)** pCR014 Plasmid mit dem Vektor pET-SagCas1Cas2Csn2 nach Deletion des StrepII-Tags in BL21 AI transformiert.



## Abbildung 6.9: Plasmidkarten der klonierten pET-CasA, -CasB, -CasC, -CasD und -CasE Vektoren.

Die Abbildung zeigt die Expressionsvektoren, die für die präparative Aufreinigung der einzelnen Cas-Proteine kloniert wurden. Neben dem klonierten *cas*-Gen sind das *lacI*-Gen, sowie das Ampicillin-Resistenzgen *bla* durch einen roten Pfeil angegeben. Des Weiteren sind der T7 Promotor und –Terminator in blau gezeigt. Der Replikationsursprung (ori) ist in grün angegeben.



#### Abbildung 6.10: Analytische Überprüfung der CasC/Cascade Expression.

Vor der Induktion mit IPTG (v), sowie 1 und 3 Stunden danach bzw. über Nacht (üN) wurden Kontrollproben von den Zellkulturen entnommen und mittels Ultraschall aufgeschlossen. Auf den dargestellten 12% SDS-Gel wurden zwei Molekulargewichtsmarker (M, Spuren 1 und 14) als Größenstandards aufgetragen. Die Spuren 2-5 enthalten den leeren pET-52(b)+ Vektor in BL21 (DE3) transformiert als externe Kontrolle. Die Spuren 6-9 zeigen den BL21 (DE3)/pET-CasC Stamm und die Spuren 10-13 den BL21 (DE3)/pET-CasC/pWUR400 Stamm. Überexprimierte Proteine sind gekennzeichnet.



#### Abbildung 6.11: SDS-Gel der Affinitätschromatographie des CasC Proteins mittels Strep-Tactin Resin.

Zu sehen ist ein 12 %iges SDS-Gel (Coomassie gefärbt), auf dem je 20-40  $\mu$ l der gesammelten Fraktionen 1-26 aufgetragen wurden. Zur Orientierung wurde ein Molekulargewichtsmarker M in Spur 29 mit aufgetragen. Die Spuren 1-4 zeigen Proben nach dem Aufschluss (A), nach dem Zentrifugieren (Ü) und vom Durchbruch (D, D2). Die Spuren 1-10 enthalten 20  $\mu$ l, die Spuren 11-26 40  $\mu$ l Probe der aufgetragenen Fraktionen. Die Elution des Cas-Proteins von der Strep-Tactin Säule erfolgte mittels eines Gradienten von 0,25-2,5 mM Desthiobiotin (Fraktionen 1 bis 4 mit 0,25 mM Desthiobiotin, ab Fraktion 5 bis 8 mit 0,5 mM Desthiobiotin und ab Fraktion 9 mit 2,5 mM Desthiobiotin). Das eluierte CasC-Protein ist mit einem Pfeil gekennzeichnet.



## Abbildung 6.12: Qualitätskontrolle des hydrolysierten *EcoRI/HincII* Leader BL21 Fragments.

Gezeigt ist ein 15% iges denaturierendes PAA-Gel nach Silberfärbung, auf dem 300 bzw. 500 ng des hydrolysierten *EcoRI/HincII* Leader BL21 Fragments aufgetragen wurden. Als Kontrolle wurden zusätzlich 500 ng des *BamHI/SacI* geschnittenen Leader BL21 Fragmentes mit aufgetragen. Zur Orientierung wurde als Marker der 100 bp DNA Leiter (NEB) in Spur 1 aufgetragen.



#### Abbildung 6.13: Analyse der Bindung von CasA, CasB und CasC an pRNA.

Gezeigt ist ein 5% iges natives PAA-Gel nach Färbung mit Ethidiumbromid. Zur Analyse der Bindung an ein 20 nt langes pRNA Oligonukleotid wurden verschiedene Konzentrationen der Cas-Proteine A, B und C (1-4  $\mu$ g) mit 200  $\mu$ M der pRNA für 15 Min. bei RT inkubiert. Um elektrostatische Wechselwirkungen zu unterdrücken, wurde den Ansätzen in den Spuren 1 bis 4 Heparin (20 ng/  $\mu$ l) hinzugegeben. Als Kontrolle wurde in Spur 1 pRNA alleine aufgetragen.

## 7 Literaturverzeichnis

Al-Attar, S., E. R. Westra, J. van der Oost and S. J. Brouns (2011). "Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes." <u>Biol Chem</u> **392**(4): 277-289.

Arslan, Z., T. Stratmann, R. Wurm, R. Wagner, K. Schnetz and U. Pul (2013). "RcsB-BglJ-mediated activation of Cascade operon does not induce the maturation of CRISPR RNAs in E. coli K12." <u>RNA Biol</u> **10**(5): 708-715.

Arslan, Z., R. Wurm, O. Brener, P. Ellinger, L. Nagel-Steger, F. Oesterhelt, L. Schmitt, D. Willbold, R. Wagner, H. Gohlke, S. H. Smits and U. Pul (2013). "Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2." <u>Nucleic Acids Res</u> **41**(12): 6347-6359.

Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., Pul, Ü., (2014). "Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system." <u>Nucleic Acids Res (in Revision)</u>

Babu, M., N. Beloglazova, R. Flick, C. Graham, T. Skarina, B. Nocek, A. Gagarinova, O. Pogoutse, G. Brown, A. Binkowski, S. Phanse, A. Joachimiak, E. V. Koonin, A. Savchenko, A. Emili, J. Greenblatt, A. M. Edwards and A. F. Yakunin (2011). "A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair." <u>Mol Microbiol</u> **79**(2): 484-502.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero and P. Horvath (2007). "CRISPR provides acquired resistance against viruses in prokaryotes." <u>Science</u> **315**(5819): 1709-1712.

Barrangou, R. and P. Horvath (2012). "CRISPR: new horizons in phage resistance and strain identification." <u>Annu Rev Food Sci Technol</u> **3**: 143-162.

Beidler, J. L., P. R. Hilliard and R. L. Rill (1982). "Ultrasensitive staining of nucleic acids with silver." <u>Anal Biochem</u> **126**(2): 374-380.

Beloglazova, N., G. Brown, M. D. Zimmerman, M. Proudfoot, K. S. Makarova, M. Kudritska, S. Kochinyan, S. Wang, M. Chruszcz, W. Minor, E. V. Koonin, A. M. Edwards, A. Savchenko and A. F. Yakunin (2008). "A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats." J Biol Chem **283** (29): 20361-20371.

Bhaya, D., M. Davison and R. Barrangou (2011). "CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation." <u>Annu Rev Genet</u> **45**: 273-297.

Bolotin, A., B. Quinquis, A. Sorokin and S. D. Ehrlich (2005). "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." <u>Microbiology</u> **151**(Pt 8): 2551-2561.

Brouns, S. J., M. M. Jore, M. Lundgren, E. R. Westra, R. J. Slijkhuis, A. P. Snijders, M. J. Dickman, K. S. Makarova, E. V. Koonin and J. van der Oost (2008). "Small CRISPR RNAs guide antiviral defense in prokaryotes." <u>Science</u> **321**(5891): 960-964.

Burgess, R. R. and J. J. Jendrisak (1975). "A procedure for the rapid, large-scall purification of Escherichia coli DNA-dependent RNA polymerase involving Polymin P precipitation and DNA-cellulose chromatography." <u>Biochemistry</u> **14**(21): 4634-4638.

Carte, J., R. Wang, H. Li, R. M. Terns and M. P. Terns (2008). "Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes." <u>Genes Dev</u> **22**(24): 3489-3496.

Chamberlin, M. J., W. C. Nierman, J. Wiggs and N. Neff (1979). "A quantitative assay for bacterial RNA polymerases." Journal of Biological Chemistry **254**(20): 10061-10069.

Chang, N., C. Sun, L. Gao, D. Zhu, X. Xu, X. Zhu, J. W. Xiong and J. J. Xi (2013). "Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos." <u>Cell Res</u> **23**(4): 465-472.

Charpentier, E. and J. A. Doudna (2013). "Biotechnology: Rewriting a genome." <u>Nature</u> **495**(7439): 50-51.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." <u>Science</u> **339**(6121): 819-823.

Cote, A. G. and S. M. Lewis (2008). "Mus81-dependent double-strand DNA breaks at in vivo-generated cruciform structures in S. cerevisiae." <u>Mol Cell</u> **31**(6): 800-812.

Datsenko, K. A., K. Pougach, A. Tikhonov, B. L. Wanner, K. Severinov and E. Semenova (2012). "Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system." <u>Nat Commun</u> **3**: 945.

Deltcheva, E., K. Chylinski, C. M. Sharma, K. Gonzales, Y. Chao, Z. A. Pirzada, M. R. Eckert, J. Vogel and E. Charpentier (2011). "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." <u>Nature</u> **471**(7340): 602-607.

Deveau, H., R. Barrangou, J. E. Garneau, J. Labonte, C. Fremaux, P. Boyaval, D. A. Romero, P. Horvath and S. Moineau (2008). "Phage response to CRISPR-encoded resistance in Streptococcus thermophilus." <u>J Bacteriol</u> **190**(4): 1390-1400.

Diez-Villasenor, C., N. M. Guzman, C. Almendros, J. Garcia-Martinez and F. J. Mojica (2013). "CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of Escherichia coli." <u>RNA Biol</u> **10**(5): 792-802.

Edgar, R. and U. Qimron (2010). "The Escherichia coli CRISPR system protects from lambda lysogenization, lysogens, and prophage induction." <u>J Bacteriol</u> **192**(23): 6291-6294.

Ellinger, P., Z. Arslan, R. Wurm, B. Tschapek, C. MacKenzie, K. Pfeffer, S. Panjikar, R. Wagner, L. Schmitt, H. Gohlke, U. Pul and S. H. Smits (2012). "The crystal structure of

the CRISPR-associated protein Csn2 from Streptococcus agalactiae." <u>J Struct Biol</u> **178**(3): 350-362.

Gasiunas, G., R. Barrangou, P. Horvath and V. Siksnys (2012). "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria." <u>Proc Natl Acad Sci U S A</u> **109**(39): E2579-2586.

Gasiunas, G. and V. Siksnys (2013). "RNA-dependent DNA endonuclease Cas9 of the CRISPR system: Holy Grail of genome editing?" <u>Trends Microbiol</u> **21**(11): 562-567.

Gonzalez, N., J. Wiggs and M. J. Chamberlin (1977). "A simple procedure for resolution of Escherichia coli RNA polymerase holoenzyme from core polymerase." <u>Arch Biochem Biophys</u> **182**(2): 404-408.

Goren, M., I. Yosef, R. Edgar and U. Qimron (2012). "The bacterial CRISPR/Cas system as analog of the mammalian adaptive immune system." <u>RNA Biol</u> **9**(5): 549-554.

Goren, M. G., I. Yosef, O. Auster and U. Qimron (2012). "Experimental definition of a clustered regularly interspaced short palindromic duplicon in Escherichia coli." <u>J Mol</u> <u>Biol</u> **423**(1): 14-16.

Haft, D. H., J. Selengut, E. F. Mongodin and K. E. Nelson (2005). "A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes." <u>PLoS Comput Biol</u> **1**(6): e60.

Hale, C. R., P. Zhao, S. Olson, M. O. Duff, B. R. Graveley, L. Wells, R. M. Terns and M. P. Terns (2009). "RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex." <u>Cell</u> **139**(5): 945-956.

Hanahan, D. (1985). Techniques for transformation of E. coli. <u>DNA cloning: a practical approach</u>. D. M. Glover. Oxford, United Kingdom, IRL Press. **1:** pp 109-135.

Hatoum-Aslan, A., I. Maniv and L. A. Marraffini (2011). "Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site." <u>Proc Natl Acad Sci U S A</u> **108**(52): 21218-21222.

Haurwitz, R. E., M. Jinek, B. Wiedenheft, K. Zhou and J. A. Doudna (2010). "Sequenceand structure-specific RNA processing by a CRISPR endonuclease." <u>Science</u> **329**(5997): 1355-1358.

Hochstrasser, M. L., D. W. Taylor, P. Bhat, C. K. Guegler, S. H. Sternberg, E. Nogales and J. A. Doudna (2014). "CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference." <u>Proc Natl Acad Sci U S A</u>.

Horvath, P. and R. Barrangou (2013). "RNA-guided genome editing a la carte." <u>Cell</u> <u>Res</u> **23**(6): 733-734.

Horvath, P., D. A. Romero, A. C. Coute-Monvoisin, M. Richards, H. Deveau, S. Moineau, P. Boyaval, C. Fremaux and R. Barrangou (2008). "Diversity, activity, and evolution of CRISPR loci in Streptococcus thermophilus." <u>J Bacteriol</u> **190**(4): 1401-1412.

Ishino, Y., H. Shinagawa, K. Makino, M. Amemura and A. Nakata (1987). "Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product." <u>J Bacteriol</u> **169**(12): 5429-5433.

Jansen, R., J. D. Embden, W. Gaastra and L. M. Schouls (2002). "Identification of genes that are associated with DNA repeats in prokaryotes." <u>Mol Microbiol</u> **43**(6): 1565-1575.

Jaskiewicz, L. and W. Filipowicz (2008). "Role of Dicer in posttranscriptional RNA silencing." <u>Curr Top Microbiol Immunol</u> **320**: 77-97.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier (2012). "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." <u>Science</u> **337**(6096): 816-821.

Jinek, M., A. East, A. Cheng, S. Lin, E. Ma and J. Doudna (2013). "RNA-programmed genome editing in human cells." <u>Elife</u> **2**: e00471.

Jore, M. (2010). <u>CRISPR-mediated Antiviral Defence in Prokaryotes</u>, publisher not identified.

Jore, M. M., M. Lundgren, E. van Duijn, J. B. Bultema, E. R. Westra, S. P. Waghmare, B. Wiedenheft, U. Pul, R. Wurm, R. Wagner, M. R. Beijer, A. Barendregt, K. Zhou, A. P. Snijders, M. J. Dickman, J. A. Doudna, E. J. Boekema, A. J. Heck, J. van der Oost and S. J. Brouns (2011). "Structural basis for CRISPR RNA-guided DNA recognition by Cascade." <u>Nat Struct Mol Biol</u> **18**(5): 529-536.

Kunin, V., R. Sorek and P. Hugenholtz (2007). "Evolutionary conservation of sequence and secondary structures in CRISPR repeats." <u>Genome Biol</u> **8**(4): R61.

Larson, M. H., L. A. Gilbert, X. Wang, W. A. Lim, J. S. Weissman and L. S. Qi (2013). "CRISPR interference (CRISPRi) for sequence-specific control of gene expression." <u>Nat. Protocols</u> **8**(11): 2180-2196.

Lemak, S., N. Beloglazova, B. Nocek, T. Skarina, R. Flick, G. Brown, A. Popovic, A. Joachimiak, A. Savchenko and A. F. Yakunin (2013). "Toroidal Structure and DNA Cleavage by the CRISPR-Associated [4Fe-4S] Cluster Containing Cas4 Nuclease SSO0001 from Sulfolobus solfataricus." Journal of the American Chemical Society **135**(46): 17476-17487.

Li, M., R. Wang, D. Zhao and H. Xiang (2014). "Adaptation of the Haloarcula hispanica CRISPR-Cas system to a purified virus strictly requires a priming process." <u>Nucleic Acids Res</u> **42**(4): 2483-2492.

Lillestol, R. K., P. Redder, R. A. Garrett and K. Brugger (2006). "A putative viral defence mechanism in archaeal cells." <u>Archaea</u> **2**(1): 59-72.

Makarova, K. S., N. V. Grishin, S. A. Shabalina, Y. I. Wolf and E. V. Koonin (2006). "A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action." <u>Biol Direct</u> **1**: 7.

Makarova, K. S., D. H. Haft, R. Barrangou, S. J. Brouns, E. Charpentier, P. Horvath, S. Moineau, F. J. Mojica, Y. I. Wolf, A. F. Yakunin, J. van der Oost and E. V. Koonin (2011). "Evolution and classification of the CRISPR-Cas systems." <u>Nat Rev Microbiol</u> **9**(6): 467-477.

Marchfelder, A., L.-K. Maier, N. Heidrich and Ü. Pul (2013). "CRISPR-Cas." <u>Biologie in</u> <u>unserer Zeit</u> **43**(3): 158-165.

Marraffini, L. A. and E. J. Sontheimer (2008). "CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA." <u>Science</u> **322**(5909): 1843-1845.

Merril, C. R., D. Scholl and S. L. Adhya (2003). "The prospect for bacteriophage therapy in Western medicine." <u>Nat Rev Drug Discov</u> **2**(6): 489-497.

Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez and C. Almendros (2009). "Short motif sequences determine the targets of the prokaryotic CRISPR defence system." <u>Microbiology</u> **155**(Pt 3): 733-740.

Mojica, F. J., C. Diez-Villasenor, J. Garcia-Martinez and E. Soria (2005). "Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements." <u>J Mol Evol</u> **60**(2): 174-182.

Mojica, F. M. and R. Garrett (2013). Discovery and Seminal Developments in the CRISPR Field. <u>CRISPR-Cas Systems</u>. R. Barrangou and J. van der Oost, Springer Berlin Heidelberg: 1-31.

Mulepati, S., A. Orr and S. Bailey (2012). "Crystal structure of the largest subunit of a bacterial RNA-guided immune complex and its role in DNA target binding." <u>J Biol</u> <u>Chem</u> **287**(27): 22445-22449.

Mullis, K. B. and F. A. Faloona (1987). "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." <u>Methods Enzymol</u> **155**: 335-350.

Nam, K. H., Q. Huang and A. Ke (2012). "Nucleic acid binding surface and dimer interface revealed by CRISPR-associated CasB protein structures." <u>FEBS Lett</u> **586**(22): 3956-3961.

Nam, K. H., I. Kurinov and A. Ke (2011). "Crystal structure of clustered regularly interspaced short palindromic repeats (CRISPR)-associated Csn2 protein revealed Ca2+-dependent double-stranded DNA binding activity." <u>J Biol Chem</u> **286**(35): 30759-30768.

Nuñez, J. K., P. J. Kranzusch, J. Noeske, A. V. Wright, C. W. Davies and J. A. Doudna (2014). "Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity." <u>Nat Struct Mol Biol</u> **advance online publication**.

Plagens, A., B. Tjaden, A. Hagemann, L. Randau and R. Hensel (2012). "Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon Thermoproteus tenax." <u>J Bacteriol</u> **194**(10): 2491-2500.

Popowitsch, A. (2013). <u>"Bestimmung von kritischen CRISPR Leader-DNA-Sequenzen</u> <u>für die Immunisierung gegen Fremd-DNA"</u> Bachelorarbeit, Heinrich-Heine-Universität. Pougach, K., E. Semenova, E. Bogdanova, K. A. Datsenko, M. Djordjevic, B. L. Wanner and K. Severinov (2010). "Transcription, processing and function of CRISPR cassettes in Escherichia coli." <u>Mol Microbiol</u> **77**(6): 1367-1379.

Pul, U., R. Wurm, Z. Arslan, R. Geissen, N. Hofmann and R. Wagner (2010). "Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS." <u>Mol Microbiol</u> **75**(6): 1495-1512.

Quax, Tessa E. F., Yuri I. Wolf, Jasper J. Koehorst, O. Wurtzel, R. van der Oost, W. Ran, F. Blombach, Kira S. Makarova, Stan J. J. Brouns, Anthony C. Forster, E. Gerhart H. Wagner, R. Sorek, Eugene V. Koonin and J. van der Oost (2013). "Differential Translation Tunes Uneven Production of Operon-Encoded Proteins." <u>Cell Reports</u> **4**(5): 938-944.

Sashital, D. G., B. Wiedenheft and J. A. Doudna (2012). "Mechanism of foreign DNA selection in a bacterial adaptive immune system." <u>Mol Cell</u> **46**(5): 606-615.

Schwabroch, T. (2012). <u>"Charakterisierung von CRISPR/Cas Proteinen und die Rolle</u> <u>der CRISPR Leader-DNA bei der bakteriellen Immunisierung gegen Fremd-DNA"</u> Masterarbeit, Heinrich-Heine-Universität.

Semenova, E., M. M. Jore, K. A. Datsenko, A. Semenova, E. R. Westra, B. Wanner, J. van der Oost, S. J. Brouns and K. Severinov (2011). "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence." <u>Proc Natl Acad Sci U S A</u> **108**(25): 10098-10103.

Shah, S. A., N. R. Hansen and R. A. Garrett (2009). "Distribution of CRISPR spacer matches in viruses and plasmids of crenarchaeal acidothermophiles and implications for their inhibitory mechanism." <u>Biochem Soc Trans</u> **37**(Pt 1): 23-28.

Sorek, R., V. Kunin and P. Hugenholtz (2008). "CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea." <u>Nat Rev</u> <u>Microbiol</u> **6**(3): 181-186.

Spreiz, V. (2011). <u>"Klonierung und Charakterisierung der CRISPR-Gene in *E. coli* <u>BL21"</u>, Heinrich-Heine-Universität.</u>

Stern, S., D. Moazed and H. F. Noller (1988). "Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension." <u>Methods Enzymol</u> **164**: 481-489.

Stratmann, T., U. Pul, R. Wurm, R. Wagner and K. Schnetz (2012). "RcsB-BglJ activates the Escherichia coli leuO gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants." <u>Mol Microbiol</u> **83**(6): 1109-1123.

Swarts, D. C., C. Mosterd, M. W. van Passel and S. J. Brouns (2012). "CRISPR interference directs strand specific spacer acquisition." <u>PLoS One</u> **7**(4): e35888.

Tang, T. H., J. P. Bachellerie, T. Rozhdestvensky, M. L. Bortolin, H. Huber, M. Drungowski, T. Elge, J. Brosius and A. Huttenhofer (2002). "Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus." <u>Proc Natl Acad Sci U S A</u> **99**(11): 7536-7541.

Tang, T. H., N. Polacek, M. Zywicki, H. Huber, K. Brugger, R. Garrett, J. P. Bachellerie and A. Huttenhofer (2005). "Identification of novel non-coding RNAs as potential antisense regulators in the archaeon Sulfolobus solfataricus." <u>Mol Microbiol</u> **55**(2): 469-481.

Terns, R. M. and M. P. Terns (2014). "CRISPR-based technologies: prokaryotic defense weapons repurposed." <u>Trends Genet</u> **30**(3): 111-118.

van der Oost, J., M. M. Jore, E. R. Westra, M. Lundgren and S. J. Brouns (2009). "CRISPR-based adaptive and heritable immunity in prokaryotes." <u>Trends Biochem Sci</u> **34**(8): 401-407.

Wagner, R. and Ü. Pul (2012). CRISPR: A Bacterial Immunity System Based on Small RNAs. <u>From Nucleic Acids Sequences to Molecular Medicine</u>. V. A. Erdmann and J. Barciszewski, Springer Berlin Heidelberg: 121-143.

Westra, E. R., B. Nilges, P. B. van Erp, J. van der Oost, R. T. Dame and S. J. Brouns (2012). "Cascade-mediated binding and bending of negatively supercoiled DNA." <u>RNA</u> <u>Biol</u> **9**(9): 1134-1138.

Westra, E. R., U. Pul, N. Heidrich, M. M. Jore, M. Lundgren, T. Stratmann, R. Wurm, A. Raine, M. Mescher, L. Van Heereveld, M. Mastop, E. G. Wagner, K. Schnetz, J. Van Der Oost, R. Wagner and S. J. Brouns (2010). "H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO." <u>Mol Microbiol</u> **77**(6): 1380-1393.

Westra, E. R., E. Semenova, K. A. Datsenko, R. N. Jackson, B. Wiedenheft, K. Severinov and S. J. Brouns (2013). "Type I-E CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition." <u>PLoS Genet</u> **9**(9): e1003742.

Westra, E. R., D. C. Swarts, R. H. Staals, M. M. Jore, S. J. Brouns and J. van der Oost (2012). "The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity." <u>Annu Rev Genet</u> **46**: 311-339.

Westra, E. R., P. B. van Erp, T. Kunne, S. P. Wong, R. H. Staals, C. L. Seegers, S. Bollen, M. M. Jore, E. Semenova, K. Severinov, W. M. de Vos, R. T. Dame, R. de Vries, S. J. Brouns and J. van der Oost (2012). "CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3." <u>Mol Cell</u> **46**(5): 595-605.

Wiedenheft, B., G. C. Lander, K. Zhou, M. M. Jore, S. J. J. Brouns, J. van der Oost, J. A. Doudna and E. Nogales (2011). "Structures of the RNA-guided surveillance complex from a bacterial immune system." <u>Nature</u> **477**(7365): 486-489.

Wiedenheft, B., E. van Duijn, J. B. Bultema, S. P. Waghmare, K. Zhou, A. Barendregt, W. Westphal, A. J. Heck, E. J. Boekema, M. J. Dickman and J. A. Doudna (2011). "RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions." <u>Proc Natl Acad Sci U S A</u> **108**(25): 10092-10097.

Wiedenheft, B., K. Zhou, M. Jinek, S. M. Coyle, W. Ma and J. A. Doudna (2009). "Structural basis for DNase activity of a conserved protein implicated in CRISPRmediated genome defense." <u>Structure</u> **17**(6): 904-912.

Woestenenk, E. A., M. Hammarstrom, S. van den Berg, T. Hard and H. Berglund (2004). "His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors." <u>J Struct Funct Genomics</u> **5**(3): 217-229.

Yosef, I., M. G. Goren and U. Qimron (2012). "Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli." <u>Nucleic Acids Res</u> **40**(12): 5569-5576.

Yosef, I., D. Shitrit, M. G. Goren, D. Burstein, T. Pupko and U. Qimron (2013). "DNA motifs determining the efficiency of adaptation into the Escherichia coli CRISPR array." <u>Proc Natl Acad Sci U S A</u> **110**(35): 14396-14401.

Zhang, J., T. Kasciukovic and M. F. White (2012). "The CRISPR associated protein Cas4 Is a 5' to 3' DNA exonuclease with an iron-sulfur cluster." <u>PLoS One</u> **7**(10): e47232.

## 8 Abkürzungsverzeichnis

α	alpha
А	Adenosin
A <sub>260</sub>	Absorption bei 260 nm
A <sub>280</sub>	Absorption bei 280 nm
Å	Ångström
Abb.	Abbildung
abs.	absolut
Amp.	Ampicillin
APS	Ammoniumperoxodisulfat
ATP	Adenosintriphosphat
bp	Basenpaare
BSA	Rinderserumalbumin
bzw.	beziehungsweise
С	Cytidin
ca.	circa
cat	Chloramphenicolacetyltransferase
ссс	superhelikale DNA
°C	Grad Celsius
cas	CRISPR associated
Ci	Curie (2,22 x 1012 dpm)
cm	Zentimeter
cpm	"counts per minute"
CRISPR	clustered regularly interspaced short
	palindromic repeats
crRNA	CRISPR-RNA
СТР	Cytidintriphosphat
d	Durchmesser
Da	Dalton
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytidintriphosphat
dGTP	Desoxyguanosintriphosphat
Δ	Deletion (Mutation)
ΔG (dG)	freie Energie
DEPC	Diethylpyrocarbonat
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DNase	Desoxyribonuklease
dNTP	Desoxynukleosidtriphosphat
dpm	Zerfälle pro Minute
dsDNA	doppelsträngige Desoxyribonukleinsäure
DTT	Dithiothreitol
dTTP	Desoxythymidintriphosphat

E <sub>260</sub> , E <sub>280</sub>	Extinktion bei 260 oder 280 nm
E. coli	Escherichia coli
EDTA	Ethylendiamintetraacetat
e.g	for example
et al.	et alii (und andere)
EtBr	Ethidiumbromid
evtl.	eventuell
f.c.	final concentration, Endkonzentration
f	femto (10 <sup>-15</sup> )
g	Gramm
G	Guanosin
GTP	Guanosintriphosphat
GvO	gentechnisch veränderter Organismus
h	Stunde(n)
HEPES	2-[4-(2-Hydroxyethyl)-1-piperazinyl]-
	ethansulfonsäure
H-NS	"histone-like nucleoid structuring protein"
i.d.R.	in der Regel
k	kilo (10 <sup>3</sup> )
kb	Kilobasen
kDa	Kilodalton
KGlu	Kaliumglutamat
Km	Michaelis-Konstante
1	Liter
λ	lambda (Wellenlänge)
m	milli (10 <sup>-3</sup> )
М	Mol pro Liter
mA	Milliampere
MCS	multiple Klonierungsstelle
mg	Milligramm
min	Minute
mm	Millimeter
μ	Mikro (10 <sup>-6</sup> )
μg	Mikrogramm
μl	Mikroliter
μΜ	Mikromolar
mM	Millimolar
mol	Stoffmenge: (1 mol = 6 x 10 <sup>23</sup> Teilchen)
mRNA	messenger RNA
MW	Molekulargewichtsmarker
n	nano (10 <sup>-9</sup> )
ng	Nanogramm
nm	Nanometer
NaOAc	Natriumacetat

	Tukielijuure
Nt	Nukleotid(e)
NTP	Nukleosidtriphosphat
OD <sub>600</sub>	optische Dichte bei 600 nm
ORF	open reading frame, offener Leserahmen
р	pico (10 <sup>-12</sup> )
[ <sup>32</sup> P]	radioaktives Phosphorisotop
Р	Promotor
PAGE	Polyacrylamidgelelektrophorese
PAA	Polyacrylamid
PCR	Polymerasekettenreaktion
PE	Primer-Extension
PEG	Polyethylenglykol
PMSF	Phenylmethansulfonylflourid
%	Prozent
Pre-crRNA	precursor crRNA
рН	pH-Wert
RNA	Ribonukleinsäure
RNAi	RNA-Interferenz
RNAP	RNA-Polymerase
RNase	Ribonuklease
rpm	Umdrehungen pro Minute
RT	Raumtemperatur
S	Svedberg (Sedimentationskoeffizient)
S SDS	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat
S SDS sec	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde
S SDS sec ssDNA	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure
S SDS sec ssDNA t	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit
S SDS sec ssDNA t T	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin
S SDS sec ssDNA t T Tab.	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle
S SDS sec ssDNA t T Tab. TEMED	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin
S SDS sec ssDNA t T Tab. TEMED Tris	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan
S SDS sec ssDNA t T Tab. TEMED Tris U	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N´,N´-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan Uridin
S SDS sec ssDNA t T Tab. TEMED Tris U	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan Uridin Unit(s) (Enzymeinheit)
S SDS sec ssDNA t T Tab. TEMED Tris U U	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan Uridin Unit(s) (Enzymeinheit) unter anderem
S SDS sec ssDNA t T Tab. TEMED Tris U U U U	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheit
S SDS sec ssDNA t T Tab. TEMED Tris U U U U U U U U	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheitüber Nacht
S SDS sec ssDNA t T Tab. TEMED Tris U U U U U U U U U U U U U U U U U U U	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan Uridin Unit(s) (Enzymeinheit) unter anderem Untereinheit über Nacht
S SDS sec ssDNA t T Tab. TEMED Tris U U U U U U U U U U U U U U U U U U U	Svedberg (Sedimentationskoeffizient) Natriumdodecylsulfat Sekunde einzelsträngige Desoxyribonukleinsäure Zeit Thymidin Tabelle N,N,N',N'-Tetramethylethylendiamin Tris(hydroxymethyl)-aminomethan Uridin Unit(s) (Enzymeinheit) unter anderem Untereinheit über Nacht über Tag Uridintriphosphat
S SDS sec ssDNA t T Tab. TEMED Tris U U U u.a UE üN ü T U U U	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheitüber Nachtüber TagUridintriphosphatultraviolett
S SDS sec ssDNA t T Tab. TEMED Tris U U U U U U U U U U U U U U U U U U U	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheitüber Nachtüber TagUridintriphosphatultraviolettVolt
S SDS sec ssDNA t T Tab. TEMED Tris U U U U U U U U U U U U U U U U U U U	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheitüber Nachtüber TagUridintriphosphatultraviolettVoltvergleiche
S         SDS         sec         ssDNA         t         T         Tab.         TEMED         Tris         U         U         U         U         UU         UU         UV         VU         UF         UN         V         V         VIS	Svedberg (Sedimentationskoeffizient)NatriumdodecylsulfatSekundeeinzelsträngige DesoxyribonukleinsäureZeitThymidinTabelleN,N,N',N'-TetramethylethylendiaminTris(hydroxymethyl)-aminomethanUridinUnit(s) (Enzymeinheit)unter anderemUntereinheitüber Nachtüber TagUridintriphosphatultraviolettVoltvergleichesichtbares Licht ("visible")

W	Watt
WT	Wildtyp
w/v	"weight per volume" (Gewichtsprozent)
YT	yeast tryptone
z.B.	zum Beispiel

## 9 Eidesstattliche Erklärung

Angefertigt im	Institut für Physikalische Biologie
	Arbeitsgruppe Molekularbiologie der
	Prokaryoten
	Heinrich-Heine-Universität Düsseldorf

Tag der Einreichung

12.Mai 2014

Erster Gutachter Zweiter Gutachter Univ.-Prof. Dr. Rolf Wagner Univ.-Prof. Dr. William Martin

Hiermit erkläre ich an Eides statt, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

(Ort, Datum, Unterschrift)

## Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich bei der Anfertigung dieser Arbeit unterstützt haben!!!

Mein besonderer Dank geht an Herrn Prof. Dr. Rolf Wagner für die Möglichkeit diese Arbeit in seiner Arbeitsgruppe anfertigen zu dürfen und für die hervorragende wissenschaftliche Betreuung. Danke für die stetige Ansprechbarkeit, Beantwortung jeglicher Fragen, den vielen nützlichen Ratschlägen und Hilfestellungen, sowie die Korrektur meiner Arbeit. Die letzten drei Jahre waren für mich sehr lehrreich!

Herrn Prof. Dr. William Martin danke ich herzlich für die Übernahme der Koreferenz.

Besonders danken möchte ich auch Ümit für die Bereitstellung des interessanten Themas, seiner stetigen Ansprechbarkeit, das Korrekturlesen und die konstruktiven Verbesserungsvorschläge.

Außerdem danke ich allen früheren und jetzigen Mitarbeitern der AG Wagner (Zihni, Tobi, Anna, Ben, Sabine, René, Tommy, Lena, Lilli, Mariam und Julian) für ihre Hilfsbereitschaft und Anregungen bei Problemen. Danke für die super tolle Arbeitsatmosphäre und die gemeinsamen Mittagspausen.

Ein großer Dank geht auch an Barbara für die viele Arbeit im Hintergrund, die Bereitstellung der Labormaterialien und Entlastung im Labor.

Reini dir danke ich für dein immer offenes Ohr, die tollen Gespräche, deine Fürsorglichkeit, den vielen Ratschlägen, der steten Hilfsbereitschaft, den aufgereinigten Proteinen und den vielen anderen Dingen, die Du unbemerkt von uns erledigt hast. Du wirst mir sehr fehlen!

Meiner Familie danke ich für die Unterstützung während der gesamten Studien- und Promotionszeit und das anerzogene Durchhaltevermögen. Tausend Dank an meinen Mann für die aufmunternden Worte immer weiterzumachen und die Rücksichtnahme in all den Jahren. Ihr seit mein Ein und Alles!