

Die Rolle von CpG-Dinukleotiden bei der Regulation der Transgenexpression am Beispiel verschiedener Reportergene

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES
DER NATURWISSENSCHAFTEN (DR. RER. NAT)
DER NATURWISSENSCHAFTLICHEN FAKULTÄT III –
BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG



vorgelegt von
Doris Leikam aus Straubing
April 2006

Die vorliegende Dissertation wurde zwischen Juni 2002 und Mai 2006 am Institut für Medizinische Mikrobiologie und Hygiene der Universität Regensburg durchgeführt.

Promotionsgesuch eingereicht am: 19. April 2006

Die Arbeit wurde von Prof. Dr. Ralf Wagner angeleitet.

Prüfungsvorsitz:	Prof. Dr. R. Wirth
Prüfungsausschuss:	Prof. Dr. Dr. H.R. Kalbitzer
	Prof. Dr. R. Wagner
	Prof. Dr. R. Warth
Prüfungstermin:	26. Juli 2006

Die vorliegende Arbeit wurde selbstständig, ohne unzulässige Hilfe angefertigt.

.....
Doris Leikam

Für meinen Vater

Ein Optimist ist ein Mensch, der ein
Dutzend Austern bestellt, in der
Hoffnung, sie mit der Perle, die er
darin findet, bezahlen zu können

Theodor Fontane

Inhaltsverzeichnis

Zusammenfassung	1
A Einleitung	3
A.1 Ausgangsüberlegungen	3
A.2 CpG-Motive	4
A.2.1 Historie und Klassifizierung.....	4
A.2.2 Wirkung auf das Immunsystem.....	5
A.2.3 Molekularer Mechanismus der Erkennung von CpG-Motiven.....	6
A.2.4 Therapeutische Applikationen.....	7
A.3 CpG-Dinukleotide.....	9
A.3.1 Vorkommen	9
A.3.2 CpG-kontrollierte Regulationsvorgänge	10
A.3.3 CpG-Dinukleotide bindende Proteine.....	11
A.3.4 Molekulare Mechanismen der epigenetischen Regulation via CpGs.....	13
A.4 Komponenten eukaryontischer Expressionsplasmide.....	15
A.4.1 CMV-Promotor	15
A.4.2 BGH-Polyadenylierungssignal	17
A.4.3 pUC <i>origin of replication</i>	18
A.5 Gensynthese.....	19
A.6 Zielsetzung	20
B Material und Methoden	22
B.1 Klonierungen und gentechnische Arbeiten	22
B.1.1 Bakterien und Nährmedien	22
B.1.2 Oligonukleotide	23
B.1.3 Plasmide	23
B.1.3.1 Das Flp-In TM -System (Invitrogen).....	23
B.1.3.1.1 pFRT/lacZeo Vektoren	23
B.1.3.1.2 pOG44-Expressionsplasmid.....	24
B.1.3.1.3 pcDNA5/FRT	24
B.1.3.1.4 Schematische Darstellung.....	25
B.1.4 PCR-Techniken (<i>Polymerase chain reaction</i>)	26
B.1.4.1 Standard PCR-Bedingungen.....	26
B.1.4.2 Gradienten-PCR.....	26
B.1.4.3 Fusions-PCR.....	27
B.1.5 Optimierung und Herstellung synthetischer Gene	28
B.1.6 Isolierung und Analyse genomischer DNA aus eukaryontischen Zellen	29
B.2 Zellkulturtechniken und Infektionen.....	29
B.2.1 Kultivierung verwendeter Zelllinien	29

B.2.2	Bestimmung der Lebendzellzahl	30
B.2.3	Transiente Transfektion eukaryontischer Zellen	30
B.2.4	Herstellung stabiler eukaryontischer Zellen (Flp-In™ System) ..	30
B.2.5	Etablierung monoclonaler CHO-Zellklone.....	31
B.2.6	Langzeitkultur stabil transfizierter CHO-Zellklone	31
B.3	Analyse der transienten und stabilen Proteinexpression	32
B.3.1	Quantitative Bestimmung der Proteinkonzentration der Zellysate	32
B.3.2	Quantifizierung des HIV-1 Kapsidantigens im p24-ELISA	32
B.3.3	Analyse der transienten Proteinexpression im Western Blot	33
B.3.4	FACS-Analysen	33
B.3.4.1.	Lebendfärbung	33
B.3.5	Fluoreszenzmikroskopie	34
B.4	RNA-Arbeitstechniken.....	34
B.4.1	Isolierung von RNA aus eukaryontischen Zellen	34
B.4.2	Photometrische Konzentrationsbestimmung der RNA	34
B.4.3	Beseitigung von DNA-Kontaminationen in der RNA	35
B.4.4	Reverse Transkription (RT).....	35
B.4.5	Qualitative Untersuchung revers transkribierter genomischer RNA.....	35
B.4.6	Quantifizierung von revers transkribierter RNA via realtime-PCR	36
B.4.6.1	Prinzip des LightCyclers™	36
B.4.6.2	Wahl der Parameter für realtime-PCR-Analysen.....	37
B.4.6.3	Quantifizierung von cDNA via LightCycler™- Analysen.....	37
B.4.6.3.1	Relative Quantifizierung von cDNA	37
B.4.6.3.2	Etablierung eines externen Standards	38
B.4.6.3.3	Absolute Quantifizierung von cDNA.....	39
B.4.7	„Nuclear Run On“ Assay	39
B.4.8	Untersuchung der RNA-Stabilität mittels <i>ActinomycinD</i>	40
C	Ergebnisse	42
C.1	Gesamtkonzept.....	42
C.1.1	Anforderungen an den Vektor.....	42
C.1.2	Wahl und Modifizierung der Transgene	42
C.2	Etablierung von CpG-Modifikationen im Vektorrückgrat	43
C.2.1	Design und Herstellung einzelner modularer Vektorkomponenten	43
C.2.1.1	Entwicklung von modularen CpG-reduzierten Vektorkomponenten	44
C.2.1.1.1	Prokaryontische Resistenzkassetten.....	44
C.2.1.1.2	Eukaryontische Selektionskassette	46
C.2.1.1.3	Multipler Klonierungsbereich (<i>MCS</i>)	48
C.2.1.1.4	CMV_TG-Promotor.....	49
C.2.1.1.5	CMV_DH-Promotorvarianten.....	49

	C.2.1.1.6 BGH_WDW-Polyadenylierungssignal.....	51
	C.2.1.2 Etablierung von modularen PCR-basierten Komponenten	52
	C.2.1.2.1 CMV-Promotor.....	52
	C.2.1.2.2 BGH Polyadenylierungssignal	52
	C.2.1.2.3 <i>pUC origin of replication</i>	52
	C.2.1.3 Zusammenfassung	52
	C.2.2 Aufbau und Herstellung verwendeter Vektorkonstrukte.....	54
	C.2.2.1 Basisvektoren zur transienten Expressionsanalyse	54
	C.2.2.1.1 Referenzkonstrukt pRS.....	54
	C.2.2.1.2 Basiskonstrukt pΔS.....	54
	C.2.2.1.3 Modifizierte Basisvektoren (pΔS110', pΔSC, pΔS kana, pΔSB WDW und pΔSC A-F).....	55
	C.2.2.1.4 Übersicht	57
	C.2.2.2 Basisvektoren zur stabilen Expressionsanalyse	58
	C.2.2.2.1 Ausgangsvektor pcDNA5/FRT.....	58
	C.2.2.2.2 pc5-Varianten mit CpG-freiem BGH Polyadenylierungssignal.....	59
	C.2.2.2.3 SYN_84	60
	C.2.2.2.4 SYN_187	60
	C.2.2.2.5 SYN_226	61
	C.2.2.2.6 SYN226-Varianten mit CpG-freiem CMV-Promotor	62
	C.2.2.2.7 Übersicht	62
C.3	Etablierung von CpG-Modifikationen innerhalb des Transgens	64
	C.3.1 Optimierung und Modifizierung der verwendeten Reporter	64
	C.3.1.1 GFP als Transgen.....	64
	C.3.1.2 p24 als Transgen	65
	C.3.2 Herstellung der verwendeten Reportergerätenkonstrukte	65
	C.3.2.1 GFP (<i>green-fluorescent-protein</i>) als Transgen.....	65
	C.3.2.2 p24-Kapsidprotein von HIV-1 als Transgen	66
	C.3.3 Zusammenfassung	66
C.4	Transiente Expressionsstudien.....	67
	C.4.1 Vergleich der CpG-modifizierten Reportergerätenvarianten	67
	C.4.2 Vergleich der Basisvektoren hinsichtlich ihrer Expressionsstärke	69
	C.4.3 Funktionalität weiterer CpG-freier Vektorelemente.....	71
	C.4.3.1 pΔS mit Kanamycinresistenzkassette	71
	C.4.3.2 Expressionsstärke verschiedener CpG-freier BGH P(A)- Signale in pΔS.....	71
	C.4.3.3 Expressionsstärke verschiedener CpG-freier CMV- Promotoren in pΔS	72
	C.4.4 Zusammenfassung	73
C.5	Stabile Expressionsstudien.....	74
	C.5.1 Vergleich der CpG-freien GFP-Reportervariante mit dem Ausgangskonstrukt	74

C.5.2	Langzeitexpressionsanalyse der GFP-Reporterkonstrukte.....	76
C.5.3	Etablierung eines CpG-reduzierten modularen Vektorrückgrates für stabile Expressionsanalysen	77
C.5.4	Funktionalität weiterer CpG-modifizierter Vektorelemente.....	79
C.5.4.1	Expression einer CpG-maximierten GFP-Variante	79
C.5.4.2	Expressionsstärke verschiedener CpG-freier BGH P(A)- Signale im pc5.....	80
C.5.4.3	Expressionsstärke verschiedener CpG-freier CMV- Promotoren im SYN226	81
C.5.5	Zusammenfassung	82
C.6	Untersuchung der GFP-Reportergergenvarianten auf RNA-Ebene	83
C.6.1	Relative Quantifizierung der zellulären RNA (CHO und 293T Zellen).....	83
C.6.1.1	CHO-Zellen	83
C.6.1.2	293T-Zellen.....	85
C.6.2	Absolute Quantifizierung zellulärer RNA in CHO Zellen	86
C.6.2.1	Cytoplasmatische und nukleäre RNA	86
C.6.2.2	Verhältnis nukleärer zu cytoplasmatischer RNA	87
C.6.3	Untersuchung der RNA-Stabilität mittels ActinomycinD.....	88
C.6.4	Untersuchung der RNA-Transkripte hinsichtlich alternativer Spleissprodukte	89
C.6.5	Quantitative Analyse <i>de novo</i> synthetisierter RNA	90
C.6.6	Zusammenfassung	91
D	Diskussion	92
D.1	Neuartige Strategien zur Steigerung der Expressionsstärke	92
D.1.1	Übersicht über vorhandene Systeme zur Erhöhung der Transgenexpression	92
D.1.2	Eigenes Konzept zur Verbesserung der Transgenexpression...	93
D.2	CpG-reduzierte Vektormodule sind funktionell.....	95
D.2.1	Funktionalität einer deletierten Form des pUC <i>origin of</i> <i>replication</i> (pUC 110´ <i>origin</i>).....	95
D.2.2	Funktionalität CpG-freier Varianten des CMV-Promotors	96
D.2.3	Funktionalität CpG-freier BGH Polyadenylierungssignale	98
D.2.4	Funktionalität CpG-freier Resistenzgene	99
D.3	CpG-Reduktion im Vektorrückgrat führt zu erhöhter Transgenexpression...	100
D.3.1	Erhöhung der Expression in transienten Studien.....	100
D.3.2	Erhöhung der Transgenexpression in stabilen Flp-In Zelllinien	101
D.4	CpG-maximierte Transgene weisen eine erhöhte Proteinproduktion auf.....	102
D.4.1	CpG-maximierte Reportergene vermitteln erhöhte Aktivität und Proteinmenge	103
D.4.2	CpG-haltige Transgene weisen erhöhte RNA-Mengen auf.....	104
D.4.3	CpG-optimierte Transgene zeigen keine Unterschiede hinsichtlich Kernexportraten, RNA-Stabilität oder alternativer Spleissprodukte	105

D.4.4 CpG-haltige Transgene besitzen eine erhöhte Menge <i>de novo</i> synthetisierter, transgenspezifischer RNA	106
D.5 Hypothese der CpG-basierten Genreaktion und Ausblicke	107
D.5.1 Arbeitshypothese der CpG-basierten Genregulation	107
D.5.2 Ausblick	109
Literatur	111
Veröffentlichungen	118
F Anhang	119
F.1 Oligonukleotide	119
F.1.1 Oligonukleotide zum Aufbau der Vektoren.....	119
F.1.2 Oligonukleotide zur Generation CpG-freier Promotorkonstrukte	120
F.1.3 Für <i>realtime</i> PCR-Analysen verwendete Oligonukleotide	120
F.2 Sequenzvergleiche	121
F.2.1 Nukleinsäuresequenzvergleiche	121
F.2.2 Aminosäuresequenzvergleiche	133
F.3 Sequenzen.....	136
F.3.1 Sequenzen der Vektorelemente.....	136
F.3.2 Sequenzen eukaryontischer Expressionsplasmide.....	138
Abkürzungsverzeichnis	141
Danksagung	145

Zusammenfassung

DNA-basierte, therapeutische Strategien wie z.B. die DNA-Vakzinierung oder die Gentherapie werden häufig durch ein *Silencen* der Antigenexpression nachhaltig beeinträchtigt, da der rekombinante Wirkstoff nachfolgend nicht mehr in physiologisch ausreichender Konzentration vorliegt. So ist bekannt, dass die Methylierung von CpG-Dinukleotiden in promotornahen Bereichen einen epigenetischen Regulationsmechanismus darstellt, der in vielen Fällen zu einem Abschalten der Genexpression führt. CpG-freie Genvarianten sollten folglich vor Methylierung und einer damit verbundenen Repression der Transkription geschützt sein. Im Gegensatz dazu sollten CpG-haltige Ausgangskonstrukte dem Methylierungsprozess unterliegen und eine verringerte Antigenproduktion zeigen. Anders als erwartet, wurde in dieser Arbeit für eine CpG-freie Version verschiedener Transgene *in vitro* jedoch eine schwächere Reporteraktivität festgestellt als für das CpG-haltige Ausgangskonstrukt. Darüber hinaus erreichte ein CpG-maximiertes Reporter-gen gegenüber dem Ausgangskonstrukt eine gesteigerte Proteinexpression. Der positive Zusammenhang von CpG-Gehalt im Transgen und resultierender Reporteraktivität manifestierte sich auch auf RNA-Ebene. So wurde sowohl in der nukleären als auch in der cytoplasmatischen RNA-Fraktion eine positive Korrelation von CpG-Zahl im kodierenden Bereich und zugehöriger RNA-Transkriptmenge festgestellt. Diese Verringerung transgenspezifischer RNA-Transkripte für das CpG-depletierte Konstrukt konnte jedoch weder auf differentielle Kernexportraten noch auf eine Instabilität CpG-freier RNA oder das Vorliegen alternativer Spleissprodukte zurückgeführt werden. Vielmehr wurde im Verlauf der Arbeit nachgewiesen, dass CpG-haltige Transgene sich bereits durch erhöhte Mengen *de novo* synthetisierter RNA-Transkripte auszeichnen. Das beschriebene CpG-Phänomen wurde somit für virale sowie nicht-virale Transgene unabhängig von Zelltyp und Expressionsdauer nachgewiesen, was einen generellen Mechanismus der CpG-basierten Genaktivierung auf transkriptioneller Ebene impliziert.

Im Bereich der modernen Medizin sowie in der Biotechnologie werden fremde Transgene meist unter Verwendung bakterieller Plasmidvektoren in das jeweilige Expressionssystem eingebracht. Im Vektorrückgrat lokalisierte CpG-Motive, die in ihrer Sequenz ein zentrales CpG-Dinukleotid enthalten, beeinflussen jedoch in erheblichem Maße Immunogenität und Höhe der Transgenexpression *in vivo*. Um unerwünschte Immunreaktionen zu unterbinden und eine stabile Transgenexpression zu gewährleisten, wurde in dieser Arbeit eine partiell CpG-reduzierte Vektorplattform mit modularem Aufbau etabliert. In transienten sowie in stabilen Expressionsexperimenten zeigte sich für CpG-reduzierte Plasmide eine erhöhte Expressionsleistung, verglichen mit einem CpG-haltigen Referenzkonstrukt. Daher wurde zur weiteren Reduzierung des CpG-Gehaltes im Vektorrückgrat eine CpG-freie Version des CMV-Promotors entwickelt, welche *in vitro* nahezu gleiche Reporteraktivität vermittelte wie der SV40-Promotor. Darüber hinaus konnte eine CpG-depletierte Version des BGH Polyadenylierungssignals generiert werden, welche zu einer höheren Transgenexpression führte, als das konventionelle BGH Polyadenylierungssignal.

Somit konnte in dieser Arbeit gezeigt werden, dass durch Modifizierung des CpG-Gehaltes in Vektorrückgrat und Transgen eine erhebliche Steigerung der Transgenexpression erreicht werden kann. Folglich stellt diese Strategie einen Ansatzpunkt zur Optimierung herkömmlicher Vektorkonstrukte dar, und sollte bei der rationalen Konzeption DNA-basierter Therapeutika sowie bei der effektiven Produktion rekombinanter Proteine in Säugerzellen in Betracht gezogen werden. Darüber hinaus wurde in der Arbeit die Grundlage für einen immun-neutralen Vektor geschaffen, der durch Integration spezifischer, immunmodulierender Sequenzen je nach Applikationszweck modifiziert werden kann.

A Einleitung

A.1 Ausgangsüberlegungen

Eine hohe und lang andauernde Expression des rekombinanten Transgens stellt die Grundlage für die Wirksamkeit und den Erfolg DNA-basierter Therapieformen, wie z.B. der DNA-Vakzinierung oder der Gentherapie, dar. Nur somit wird gewährleistet, dass der Wirkstoff in therapeutisch ausreichender Konzentration vorliegt und Einfluss auf physiologische Vorgänge nehmen kann. Bei DNA-basierten Impfstoffen z.B., korreliert die Stärke der generierten Immunantworten mit der Expressionseffizienz des rekombinanten Antigens (Deml *et al.*, 2001). Neben einer ausreichenden Proteinproduktion spielen bei der Gestaltung moderner Therapeutika jedoch noch weitere Faktoren eine Rolle. Um optimalen Impfschutz zu generieren, sollten DNA-Impfstoffe darüber hinaus in der Lage sein, das Immunsystem „direkt“ anzusprechen und eine humorale sowie eine zelluläre Immunantwort zu induzieren (Ulmer *et al.*, 1993; Girard *et al.*, 2006). Anders verhält es sich bei gentherapeutischen Anwendungen. In diesem Fall sind immunaktivierende Einflüsse des DNA-basierten Therapeutikums zu vermeiden, da mit therapeutischen Genen transfizierte Zielzellen ansonsten verstärkt vom Immunsystem erkannt und anschließend eliminiert würden. Weiterhin sollten gentherapeutische Ansätze dahingehend verbessert werden, dass ein Abschalten der Transgenexpression *in vivo*, wie es in der Literatur beschrieben ist (Bestor, 2000; Chevalier-Mariette *et al.*, 2003), verringert bzw. verhindert wird. Nicht nur für Applikationen *in vivo* sind Expressionsleistung des Transgens, und immunmodulatorische Eigenschaften des Vektors von hoher Relevanz, auch für eine effektive und kostengünstige Produktion rekombinanter Antikörper *in vitro* ist eine hohe Proteinausbeute erforderlich (Vernon, 2004). Als molekularbiologisches „Werkzeug“, um DNA-Impfstoffe, gentherapeutische Konstrukte sowie Expressionsvektoren hinsichtlich ihrer Transgenproduktion bzw. Immunogenität zu modifizieren, bieten sich bestimmte, in der DNA vorliegende CpG-Sequenzen an (Zhang *et al.*, 2005). So wird sogenannten CpG-Dinukleotiden (siehe A.3), bei denen es sich um in der Basenpaarabfolge der DNA auftretende Cytidin-Guanosin-Dinukleotide (CpGs) handelt, eine regulatorische Funktion bei der Determinierung der Genexpression zugeschrieben (Jones *et al.*, 1999; Robertson *et al.*, 2001; Yaneva *et al.*, 2002). Darüber hinaus weisen in einem bestimmten Sequenzkontext vorliegende CpG-Dinukleotide, welche als CpG-Motive (siehe A.2) bezeichnet werden, immunmodulatorische Eigenschaften auf (Krieg *et al.*, 1995; Yi *et al.*, 1998). Nachfolgend werden CpG-Motive sowie –Dinukleotide hinsichtlich ihres Vorkommens und ihrer Wirkungsweise näher charakterisiert.

A.2 CpG-Motive

A.2.1 Historie und Klassifizierung

Bereits vor über 100 Jahren entdeckte der New Yorker Chirurg William Coley, dass sich eine bakterielle Infektion im Bereich eines Tumors positiv auf den Krankheitsverlauf auswirkte. Daraufhin applizierte er Patienten mit inoperablen Sarkomen Bakterien oder bakterielle Lysate aus Streptokokkus und Serratia in den Tumorbereich. Diese Behandlung führte bei einem Teil der Patienten zu einer vorübergehenden Rückbildung der Sarkomere und stellt die erste Immuntherapie einer Tumorerkrankung dar (Coley, 1991). Erst im Jahre 1984 wurde bei der Untersuchung verschiedener Fraktionen der Tuberkelbakterien (*Bacillus-Calmette-Guerin*) in Tumormodellen die DNA-Fraktion als verantwortliches Agens für die Immunstimulierung identifiziert (Tokunaga *et al.*, 1984). Wenig später gewann man die Einsicht, dass nur bakterielle DNA, nicht aber die von Wirbeltieren, eine Immunantwort generierte (Yamamoto *et al.*, 1992). Schließlich konnte die Wirksamkeit der bakteriellen DNA-Fragmente auf das Vorliegen unmethylierter CpG-Dinukleotide zurückgeführt werden, die in einem bestimmten Sequenzkontext auftraten (Krieg *et al.*, 1995). Bei den sogenannten CpG-Motiven handelt es sich um oft palindromisch angeordnete Hexanukleotide, die in ihrem Zentrum ein CpG-Dinukleotid besitzen. In der Regel sind murine CpG-Motive nach folgender Gesetzmäßigkeit aufgebaut: Zwei Purinbasen flankieren das zentrale CpG-Dinukleotid im 5' Bereich, während zwei Pyrimidinbasen das 3' Ende des Dinukleotides umgeben. Je nach flankierender Sequenz weisen diese in mikrobieller DNA vorkommenden CpG-Motive unterschiedliche immunmodulatorische Eigenschaften auf (Krieg, 2002).

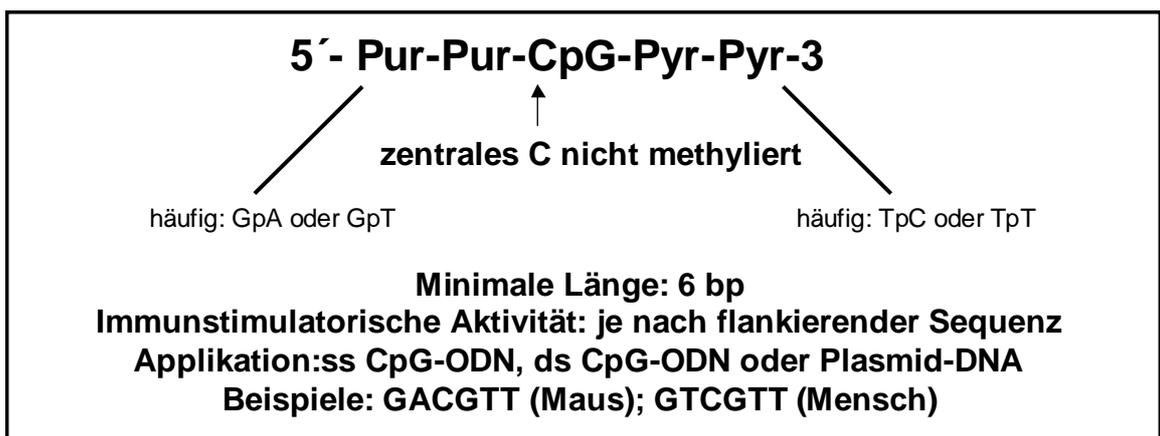


Abbildung A.1: Aufbau eines immunstimulierenden CpG-Motivs. Im allgemeinen flankieren zwei Purinbasen (A oder G) im 5'-Bereich das zentrale CpG-Dinukleotid, während zwei Pyrimidinbasen (C oder T) das 3'-Ende des Dinukleotides flankieren. Weitere Informationen können dem Text entnommen werden.

Aufgrund der von ihnen hervorgerufenen Immunantwort lassen sich murine CpG-Motive in zwei Klassen einteilen. CpG-Motive, bei denen das zentrale Dinukleotid

von den Basen A und T flankiert wird, weisen in der Regel immunstimulierende Eigenschaften auf (Lipford *et al.*, 1997; Yamamoto *et al.*, 1992). Sie werden daher als stimulierende Motive bezeichnet. Liegt das zentrale CpG-Dinukleotid im Kontext der Basen C und G vor, so werden Reaktionen des Immunsystems meist unterbunden bzw. neutralisiert, man spricht von supprimierenden bzw. neutralisierenden CpG-Motiven (Sun *et al.*, 1997; Krieg *et al.*, 1998). Dennoch zeigen CpG-Motive oft eine von Spezies zu Spezies variierende Aktivität. So ist ein klassisches immunstimulatorisches CpG-Motiv in der Maus, welches am 5'-Ende von zwei Purinen (GpA) und am 3'-Ende von zwei Pyrimidinen (TpT) flankiert wird, im humanen System inaktiv. Ein optimales humanes CpG-Motiv, dem die Basensequenz GTCGTT zugrunde liegt, ruft dagegen im murinen System keinerlei Immunaktivierung hervor (Yi *et al.*, 1998; Hartmann *et al.*, 2000).

Eine Untersuchung der von CpG-Motiven generierten Immunreaktionen, erfolgte unter Verwendung sogenannter CpG-ODN, kurzer, einzelsträngiger Desoxyribonukleotide mit spezifischen CpG-Sequenzen (Krieg *et al.*, 1998). Je nach resultierender immunogener Aktivität sowie beteiligter Komponenten des Immunsystems können verschiedene Typen immunstimulierender CpG-ODN unterschieden werden. CpG-ODN vom Typ A stellen durch Induktion des Cytokins IFN α eine Virusinfektion nach. Dem Typ B zugeordnete CpG-ODN fördern die Synthese von Interleukin-12 (IL12) und imitieren die Gegenwart intrazellulärer Bakterien. Der Klasse C zugehörige CpG-ODN rufen sowohl für CpG-ODN-A als auch für CpG-ODN-B beschriebene Immunreaktionen hervor. (Vollmer *et al.*, 2004).

A.2.2 Wirkung auf das Immunsystem

Im Gegensatz zur DNA von Vertebraten, in welcher CpG-Dinukleotide statistisch unterrepräsentiert vorliegen, beeinflusst die mikrobielle, CpG-Motive enthaltende DNA das angeborene Immunsystem von Säugern. CpG-Dinukleotide in Säuger-DNA treten zudem vorwiegend im Sequenzkontext bestimmter Basen auf, so dass ihre immunstimulatorische Wirkung geringer ausfällt, als bei einer zufälligen Basenverteilung zu erwarten wäre (Han *et al.*, 1994; Krieg *et al.*, 1998). Daher liegt die Annahme nahe, dass in bakterieller DNA vorhandene CpG-Motive als molekulare Muster fungieren, anhand derer das Immunsystem von Vertebraten fremde DNA erkennt. Im Rahmen dieses Erkennungsmechanismus initiiert das Immunsystem von Säugern bestimmte Immunreaktionen, welche unter anderem eine Eliminierung der fremden DNA zur Folge haben (Hartmann *et al.*, 1999; Brazolot *et al.*, 1998). Dabei beruht eine CpG-vermittelte Immunmodulation auf folgenden Zusammenhängen:

Dendritische Zellen (DCs) des angeborenen Immunsystems erkennen CpG-Motive direkt und werden von diesen aktiviert (Sparwasser *et al.*, 1998). Dabei sezernieren aktivierte DCs Cytokine (IL12) und Interferone (IFN α), die ihrerseits natürliche Killerzellen (NK-Zellen) zur Produktion von IFN γ anregen. Auch T-Zellen

werden sekundär über produzierte Cytokine wie z.B. IL12 und IFN α induziert, wobei in der Folge eine Th1-vermittelte Immunantwort generiert wird. Andere Zellen des angeborenen Immunsystems wie Monozyten oder Makrophagen werden ebenfalls durch CpG-Motive stimuliert, was in der vermehrten Produktion von Cytokinen (IFN α) und NF κ B resultiert (Stacey *et al.*, 1996; Klinman *et al.*, 2002). Zudem werden B-Zellen als Teil des spezifischen Immunsystems durch CpG-ODN oder bakterielle DNA aktiviert. So regen bestimmte CpG-Motive die Proliferation von B-Zellen an und stimulieren synergistisch mit dem Antigen die Antikörperproduktion (Krieg *et al.*, 1995). Da CpG-Motive neben einer humoralen auch eine zelluläre Immunantwort verstärken können, stellt die Integration dieser Elemente einen geeigneten Ansatzpunkt zur Verbesserung herkömmlicher DNA-Vakzinekonstrukte dar (Silvera *et al.*, 2004; Zhang *et al.*, 2005). In den letzten Jahren wurde in zahlreichen Studien untersucht, anhand welcher molekularer Mechanismen das Immunsystem von Säugern bakterielle, CpG-haltige DNA erkennt.

A.2.3 Molekularer Mechanismus der Erkennung von CpG-Motiven

Bei der Erkennung von unmethylierten CpG-Motiven (CpG-DNA) durch das Immunsystem spielt der *Toll like receptor 9* (TLR9) eine entscheidende Rolle. Erstmals wurde der Namen gebende Rezeptor Toll bei *Drosophila* beschrieben, wo er in der Embryonalentwicklung an der Ausbildung der dorso-ventralen Achse beteiligt ist (Anderson *et al.*, 1985). Da die cytoplasmatische Domäne des Toll-Proteins aus *Drosophila* und der Interleukin-1-Rezeptor (IL-1R) von Säugetieren stark konservierte Strukturen aufweisen, wurde für den Toll-Rezeptor ebenfalls eine Beteiligung an immunmodulierenden Signalkaskaden diskutiert (Belvin *et al.*, 1996). Weitere Rezeptoren, die in den konservierten Bereichen mit Toll bzw. IL-1R übereinstimmen, wurden nachfolgend als Toll-ähnliche Rezeptoren (*Toll like receptors*, TLRs) bezeichnet (Medzhitov *et al.*, 1997). Derzeit sind neben dem TLR9 elf weitere TLRs im Säugersystem bekannt, welche im Wesentlichen in ihrem Aufbau übereinstimmen. So beinhaltet die extrazelluläre Domäne (EZD) Leucin-reiche Wiederholungen (*leucine rich repeats*, LRR) und geht nach einem kurzen, Cystein-reichen Abschnitt in die Transmembrandomäne (TMD) über. Der cytoplasmatische Bereich weist starke Homologie zum IL-1-Rezeptor auf und wird daher als Toll/IL-1 homologe Region (TIR) bezeichnet (Rock *et al.*, 1998). Anders als viele Mitglieder der TLR-Familie wird der TLR9 jedoch nicht auf der Oberfläche der Zelle, sondern in intrazellulären Kompartimenten exprimiert (Kandimalla *et al.*, 2003).

Eine Beteiligung des TLR9 an der Erkennung von CpG-DNA wurde zweifelsfrei nachgewiesen, da TLR9-defiziente Mäuse keine Reaktion auf letale Dosen stimulierender CpG-DNA zeigten. Sowohl im murinen als auch im humanen System stellt der TLR9 die erste molekulare Komponente bei der Erkennung von CpG-DNA dar (Hemmi *et al.*, 2000; Bauer *et al.*, 2001). Von der Zelle unspezifisch aufgenommene CpG-DNA gelangt zunächst in die Endolysosomen, wo bei pH-

Wert 5,5 eine Interaktion mit dem TLR9 erfolgt. Unklar ist derzeit, ob der Rezeptor an CpG-DNA bindet, es gibt aber Hinweise, dass der TLR9 Motive besitzt, über die eine spezifische Bindung an CpG-DNA möglich ist (Krieg, 2002). Die Interaktion des Rezeptors mit der fremden DNA löst schließlich eine MAPK-Signalkaskade (*mitogen activated protein kinases pathway*) aus, im Zuge derer durch Rekrutierung des Adapterproteins MyD88 die Transkriptionsfaktoren AP-1 und NF κ B vermehrt gebildet werden (siehe Abbildung A.2). Als Folge werden spezifische Immunreaktionen auslösende Cytokine produziert und sezerniert (Yi *et al.*, 1998; Akira *et al.*, 2001). Zelltypen, die den TLR-9 exprimieren, werden durch CpG-DNA anhand dieses Signalweges direkt aktiviert. Zu dieser Gruppe gehören beim Menschen B-Zellen und plasmacytoide DCs, während in der Maus zusätzlich myeloide DCs, Monozyten und Makrophagen TLR9 aufweisen (Bauer *et al.*, 2001). Eine Anregung anderer Zelltypen wie z.B. NK-Zellen erfolgt indirekt durch die von TLR9 exprimierenden Zellen produzierten Cytokine (Krug *et al.*, 2001; Kadowaki *et al.*, 2001). Die Kenntnis molekularer Zusammenhänge, auf denen immunstimulatorische Eigenschaften spezifischer CpG-Motive beruhen, ermöglicht den gezielten Einsatz dieser bakteriellen DNA-Komponenten in der modernen Medizin.

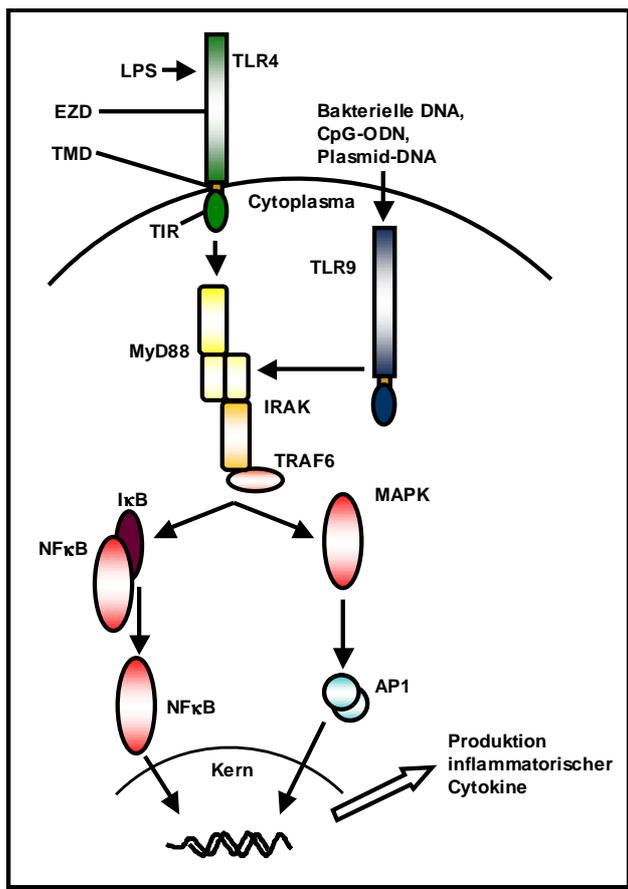


Abbildung A.2: Übersicht über zentrale Stationen des TLR4- bzw. TLR9 vermittelten Signaltransduktionsweges. AP1: Aktivierendes Protein 1; EZD: Extrazelluläre Domäne; I κ B: Inhibitor- κ B; IRAK: IL-1R-assoziierte Kinase; LPS: Lipopolysaccharide; MAPK: *mitogen activated protein kinase*; MyD88: myeloider Differenzierungsfaktor 88; NF- κ B: Nukleärer Faktor- κ B; TIR: Toll/IL-1 homologe Region; TMD: Transmembrandomäne; TLR9: Toll-like Rezeptor 9; TLR4: Toll-like Rezeptor 4; TRAF6: TNF-Rezeptor assoziierter Faktor 6; (Modifiziert nach Kandimalla *et al.*, 2003).

A.2.4 Therapeutische Applikationen

Aufgrund ihres immunstimulatorischen Potentials bietet sich die Verwendung von CpG-DNA als modernes Th1-Adjuvans an. Darüber hinaus ermöglicht die Integra-

tion spezifischer CpG-Motive eine gezielte Veränderung der Immunogenität DNA-basierter Therapeutika wie z.B. von DNA-Impfstoffen oder Gentherapievektoren (Klinmann *et al.*, 2004, Zhang *et al.*, 2005). Zudem ist eine mit bestehenden Therapeutika kombinierte Verabreichung immunmodulierender CpG-DNA derzeit Gegenstand der immuntherapeutischen Forschung (Krieg *et al.*, 2002; Silvera *et al.*, 2004). Dabei zeigt CpG-DNA im Tiermodell ein breites Anwendungsspektrum.

So wurde bei Verwendung von CpG-ODN als Adjuvans nach einer Vakzinierung gegen das Hepatitis-B-Virus (HBV) in Mäusen eine gesteigerte Antikörperproduktion nachgewiesen (Davis *et al.*, 1998). Im Rahmen eines Not-Impfprogramms, das in Indonesien aufgrund einer Epidemie durchgeführt wurde, bestätigten sich diese Ergebnisse auch für Orang-Utans (Davis *et al.*, 2000). Weiter resultierte die zusätzliche Applikation von CpG-ODN in einer Verbesserung eines Malariaimpfstoffes bei Aotus-Affen (Jones *et al.*, 1999). Auch bei der Behandlung von Allergien wie z.B. Asthma war ein positiver Effekt von CpG-ODN zu verzeichnen. So zeigte sich im Mausmodell, dass eine Asthmaentwicklung verhindert wird, wenn CpG-ODN zusammen mit dem Allergen gegeben werden. Zudem lassen sich bestimmte Atemwegserkrankungen mit den CpG-haltigen Molekülen therapieren (Horner *et al.*, 2001; Kline *et al.*, 1998, 2002). Die Injektion von CpG-ODN führte bei bestimmten Tumorerkrankungen im murinen System zudem zu einer Reversion der Zellentartung (Heckelsmiller *et al.*, 2002). Eine prophylaktische Gabe stimulierender CpG-ODN induzierte schließlich einen Schutz vor Infektionen wie z.B. *Leishmania major*-vermittelten Erkrankungen (Lipford *et al.*, 2000). Ebenso wie die herkömmlichen Wirkstoffe konnte auch ein DNA-basierter Impfstoff gegen HIV-1 durch Verwendung immunstimulatorischer CpG-Motive im murinen Modell in seiner Immunogenität verbessert werden (Kojima *et al.*, 2002;).

Aufgrund der erhaltenen Daten werden derzeit in humanen Studien CpG-ODN in Kombination mit herkömmlichen HBV- bzw. Grippe-Impfstoffen getestet. Erste Ergebnisse der HBV-Studie zeigten, dass die Probanden bei Verwendung von CpG-ODN bereits nach zwei Wochen einen Impfschutz aufwiesen. Bei der Kontrollgruppe hingegen wurde erst nach der Auffrischimpfung ein schützender Antikörpertiter detektiert (Coley Pharmaceutical, Langenfeld). Weitere auf CpG-ODN basierende, therapeutische Behandlungen gegen Asthma, verschiedene Atemwegserkrankungen und einige Tumore werden derzeit in klinischen Studien der Phasen I bis III am Menschen erprobt (<http://www.coleypharma.com>).

Unter Berücksichtigung dieser Daten ergibt sich, dass auch in bakteriellen oder viralen Vektoren vorhandene CpG-Motive einen Einfluss auf das Immunsystem von Vertebraten ausüben (Kojima *et al.*, 2002; Ma *et al.*, 2002; Zhang *et al.*, 2005). Zudem wurde in mehreren Studien nachgewiesen, dass im Vektorrückgrat befindliche CpG-Motive bei Eukaryonten die Transgenexpression beeinträchtigen (Reyes-Sandoval *et al.*, 2004). Darüber hinaus sind CpG-Dinukleotide, welche in Kapitel A.3 näher charakterisiert werden, an der epigenetischen Regulation vieler Entwicklungsprozesse beteiligt.

A.3 CpG-Dinukleotide

A.3.1 Vorkommen

Folgt in der Nukleotidsequenz der DNA ein Guanosin (G) auf ein Cytidin (C), so spricht man von einem CpG-Dinukleotid. In bakterieller DNA treten CpG-Dinukleotide gemäß der statistischen Erwartung mit einer Wahrscheinlichkeit von 1:16 auf und liegen vorwiegend unmethyliert vor (Bird, 1986). Im Genom von Eukaryonten dagegen sind CpG-Dinukleotide (CpGs) mit einer Häufigkeit von 1:60 statistisch unterrepräsentiert und tragen am 5C-Atom des Cytosinrestes meist eine Methylgruppe (Krieg *et al.*, 1995; Bestor, 2000). Dabei ist die statistische Unterrepräsentation von CpGs in der DNA von Vertebraten durch die chemische Natur des 5-Methylcytosins (^{5m}C) bedingt (Fryxell *et al.*, 2005). Durch Deaminierung (siehe Abbildung A.3) werden methylierte Cytosine häufig zu Thyminen modifiziert (Shen *et al.*, 1994). Folglich machen die Basen Cytosin und Guanin nur etwa 40% aller Basen in der Säuger-DNA aus, und CpG-Dinukleotide weisen nur etwa 20% der zu erwartenden Häufigkeit auf (Sved *et al.*, 1990; Takai *et al.*, 2002).

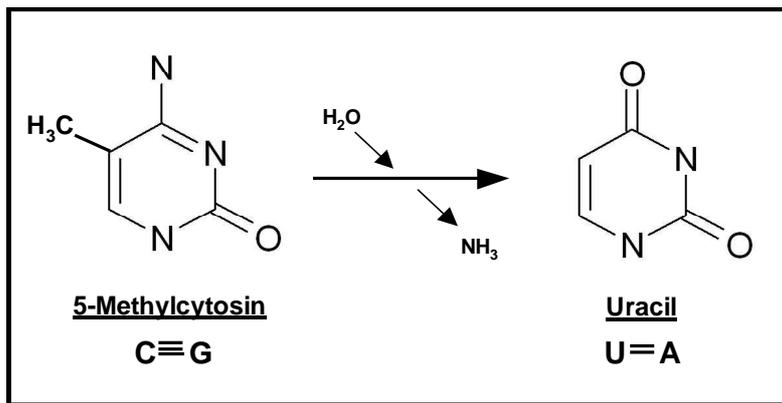


Abbildung A.3: Deaminierung von 5-Methylcytosin. Durch Abspaltung einer Aminogruppe wird methyliertes Cytosin zu Uracil modifiziert. Während der Replikation kommt es zu Fehlpaarungen, da Uracil mit der komplementären Base Adenin paart. Daraus resultiert der Einbau von Thyminid in die DNA der Tochterzellen.

Darüber hinaus beinhaltet das eukaryontische Genom jedoch auch definierte Bereiche, die eine überdurchschnittlich hohe Dichte an CpG-Dinukleotiden besitzen. Diese sogenannten CpG-Inseln weisen im Gegensatz zu den umgebenden Bereichen eine etwa 5- bis 10-fach höhere CpG-Häufigkeit auf und sind in der Regel durch das Vorliegen unmethylierter CpG-Dinukleotide gekennzeichnet (Antequera *et al.*, 1993; Bird *et al.*, 1995). Bei CpG-Inseln handelt es sich um 200-bp-große bzw. bis hin zu mehreren Kilobasen umfassende Sequenzabschnitte, welche 1-2% des humanen Genoms ausmachen (Clark *et al.*, 2002). Meist sind CpG-Inseln mit Promotoren assoziiert, sie können aber auch in die transkribierte Region reichen oder sogar innerhalb von Exons liegen (Gardiner-Garden *et al.*, 1987). Das menschliche Genom enthält circa 45 000 CpG-Inseln, und 72% aller charakterisierten humanen Promotoren besitzen CpG-Inseln in ihrer Umgebung (Saxonov *et al.*, 2006). Neben Haushaltsgenen sind auch die Hälfte aller gewebsspezifischen Gene mit Bereichen hoher CpG-Dichte assoziiert (Larsen *et al.*, 1992; Antequera *et al.*, 1993). Im allgemeinen werden CpG-Inseln mit geringem CpG-Methylierungsgrad, sog. hypomethylierte Inseln, mit einer aktiven Genexpression des entsprechenden Bereiches in Verbindung gebracht

(Cross *et al.*, 1995), während seltener vorkommende hypermethylierte CpG-Inseln einen transkriptionsinaktiven Zustand zur Folge haben (Meehan *et al.*, 1992; De Smet *et al.*, 1999). Durch Modulation des jeweiligen, für Zelltyp und Entwicklungsstadium charakteristischen Methylierungsgrades werden viele zelluläre Prozesse (siehe A.3.2) kontrolliert und reguliert (Azhikina *et al.*, 2005; Berger *et al.*, 2005).

A.3.2 CpG-kontrollierte Regulationsvorgänge

Die Methylierung von CpG-Dinukleotiden stellt einen epigenetischen Regulationsmechanismus von Eukaryonten dar, auf dessen Grundlage viele ontogenetische Vorgänge gesteuert werden. So spielt die CpG-Methylierung zum Beispiel bei der elternspezifischen Ausprägung einer Erbanlage, dem sog. *Genetic Imprinting*, eine Rolle. Während der Keimzellentwicklung wird das mütterliche bzw. väterliche Allel eines Merkmals durch Methylierung von promotornahen CpG-Inseln modifiziert, was eine veränderte Genexpression determiniert (Ferguson-Smith *et al.*, 1993). Nach der Befruchtung wird die elterliche DNA zunächst vollständig demethyliert. Obwohl die meisten CpG-Inseln im frühen Embryonalstadium vor Methylierung geschützt sind, erfolgt zwischen Morula und Blastula-Stadium eine *de novo* Methylierung bestimmter CpG-Inseln (Bestor *et al.*, 1992; Clark *et al.*, 2002). Je nach dem, ob der maternale oder paternale Methylierungsstatus eines Allels wiederhergestellt wird, kommt die mütterliche oder väterliche Version eines Merkmals zur Ausprägung. Auch die Inaktivierung des zweiten X-Chromosoms bei der Frau geht mit einer Hypermethylierung der auf diesem Chromosom lokalisierten CpG-Inseln einher. (Riggs *et al.*, 1992). Im Embryonalstadium festgelegte Methylierungsmuster werden während der Replikation vom Parenteralstrang auf den Einzelstrang der Tochterzelle weitergegeben. Das resultierende Methylierungsprofil stellt jedoch kein starres System dar, sondern ist reversibel und kann je nach Entwicklungsstufe und Zelltyp variieren (Gundersen *et al.*, 1992; Warnecke *et al.*, 1999; Azhikina *et al.*, 2005).

Neben Entwicklungsprozessen werden auch zelluläre Vorgänge durch Modifikation von CpG-Dinukleotiden kontrolliert. So unterscheidet sich das Methylierungsmuster entarteter Zellen erheblich von dem intakter Zellen (Baylin *et al.*, 2001; Jones *et al.*, 1999). Dabei gibt es verschiedene Möglichkeiten, wie veränderte Methylierungsprofile die Entstehung von Krebs bedingen können. Zum einen führt die Hypermethylierung von Promotoren, welche mit CpG-Inseln assoziiert sind, zu einem Abschalten der resultierenden Genexpression. So können z.B. dem Schutz der Zelle dienende Tumorsuppressorgene ihre Funktion nicht mehr ausüben und es kommt zur Zellentartung (Karpf *et al.*, 2002; Yaneva *et al.*, 2002). Eine Vielzahl an Studien hat belegt, dass ein sogenanntes *Silencen* der Genexpression an der Entstehung von Brustkrebs, Leukämie oder Prostatakrebs beteiligt ist (Lu *et al.*, 2001; Agirre *et al.*, 2003; Nelson *et al.*, 2003). Seltener bedingt auch eine Hypomethylierung von inaktiven, methylierten Promotorbereichen eine abnormale Genexpression und führt so zur Tumorbildung (De Smet *et al.*, 1996; Fang *et al.*, 1996).

Auch bei Pflanzen wurde ein CpG-basierter Regulationsmechanismus nachgewiesen, allerdings sind die Zusammenhänge weitaus weniger gut aufgeklärt als bei Vertebraten. So weisen *Arabidopsis*-Pflanzen, bei denen der Methylierungsgrad des Genoms experimentell verringert wurde, eine verfrühte und verkümmerte Blütenbildung auf (Finnegan *et al.*, 2000). Darüber hinaus wurde für Weizen gezeigt, dass nach der sog. Vernalisation, einer Kälteeinwirkung, die die Blütenbildung beschleunigt und diese im Winter verhindert, eine Demethylierung des Genoms eintritt (Sherman *et al.*, 2002). Wie der CpG-basierte, epigenetische Regulationsmechanismus im Einzelnen funktioniert, ist noch nicht vollständig aufgeklärt. Erwiesen ist aber, dass unter anderem bestimmte, an CpG-Dinukleotide bindende Proteine (siehe A.3.3) eine wichtige Rolle spielen.

A.3.3 CpG-Dinukleotide bindende Proteine

Sobald ein CpG-Dinukleotid methyliert (mCpG) vorliegt, stellt es eine potentielle Bindestelle für bestimmte Proteinkandidaten, wie z.B. MBD-Proteine (*methyl-CpG binding domain proteins*) dar. Bei Säugern sind derzeit 5 ubiquitär vorliegende mCpG-Bindeproteine (siehe Abbildung A.4) bekannt, wobei meist ein mCpG für eine Bindung ausreichend ist (Prokhortchouk *et al.*, 2002). Den vier Proteinen MeCP2, MBD1, MBD2 sowie MBD4 ist eine MBD gemeinsam, welche eine Bindung an die mCpG-DNA vermittelt. MBD3, welches mehr als 70% Sequenzhomologie zu MBD2 aufweist, besitzt ebenfalls eine MBD, hat jedoch die Fähigkeit verloren, an methylierte DNA zu binden (Hendrich *et al.*, 1998; Ballestar *et al.*, 2001). Dennoch gehen alle bis dato erwähnten mCpG-Bindeproteine auf einen gemeinsamen evolutorischen Ursprung zurück, wie die Position eines hoch konservierten Introns erkennen lässt. Das fünfte derzeit bekannte mCpG-Bindeprotein Kaiso hingegen weist eine andere Struktur auf und bindet methylierte DNA anhand eines Zinkfinger-Motivs (Prokhortchouk *et al.*, 2001).

Während MeCP2, MBD1, MBD2 und Kaiso *in vitro* sowie *in vivo* als Transkriptionsrepressoren fungieren, ist MBD3 Teil eines Co-Repressorkomplexes (Prokhortchouk *et al.*, 2002; Berger *et al.*, 2005). MBD4 dagegen, stellt eine Komponente des zelleigenen DNA-Reparatur-Mechanismus dar. Das Protein greift bei Vorliegen einer CpG/TpG-Fehlpaarung ein, welche durch Deaminierung von ^{5m}C verursacht wird. Somit wird die CpG-Dichte im jeweiligen DNA-Abschnitt beibehalten, und der Einbau von Punktmutationen verhindert (Carlone *et al.*, 2001; Hendrich *et al.*, 1999). Die essentielle Funktion dieser Proteine spiegelt sich z.B. in der Tatsache wieder, dass *Knock-out*-Mäuse für MBD3 letal sind (Wade *et al.*, 2001). Zudem führen Mutationen im MeCP2-Gen beim Menschen zum Rett-Syndrom, einer tiefgreifenden Entwicklungsstörung. (Amir *et al.*, 1999).

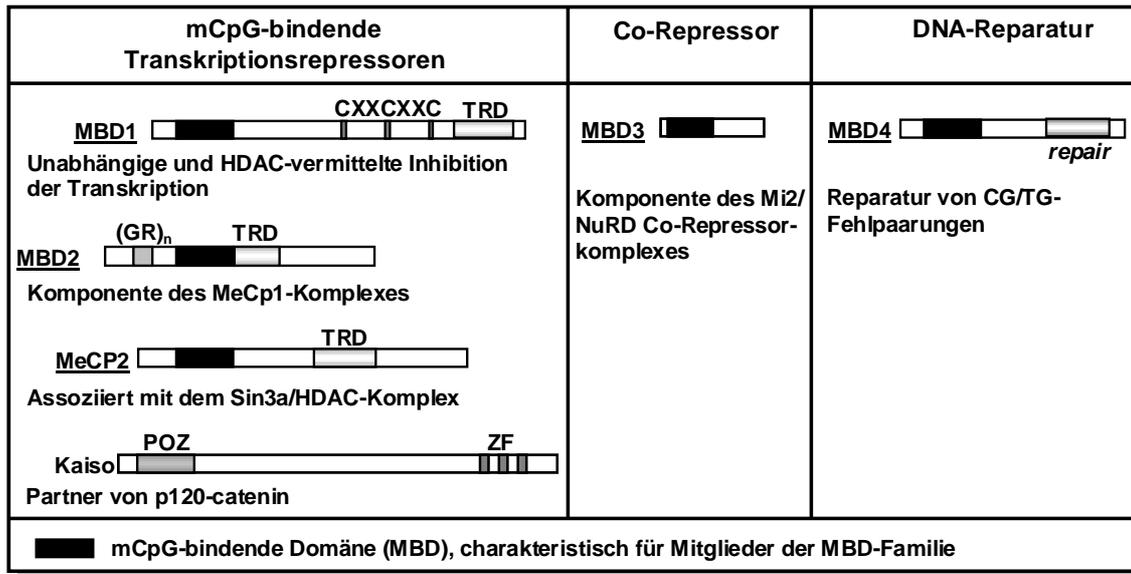


Abbildung A.4: Übersicht über mCpG-bindende Proteine und ihre Funktion. Die Namen der zur Familie der MBD-Proteine gehörenden Vertreter sind unterstrichen. CXXCXXC: Cystein-X-X-Cystein-Motive; (GR)_n: Glycin-Arginin-Wiederholung; TRD: *Transcriptional repressor domain*; POZ: Poxviurs-Zinkfinger-Domäne; ZF: Zinkfinger. Nähere Erläuterungen können dem Text entnommen werden. (Modifiziert nach Prokhortchouk *et al.*, 2002; Klose *et al.*, 2006).

Neben mehreren Transkriptionsrepressoren, wird in der Literatur für Mensch und Maus auch ein die Transkription förderndes Protein beschrieben, das jedoch an unmethylierte CpG-Dinukleotide (CpGs) bindet. Dabei handelt es sich um das CpG-Bindeprotein (CGBP), welches im Kern lokalisiert ist und in zahlreichen embryonalen sowie adulten Zell- und Gewebetypen nachgewiesen wurde (siehe Abbildung A.5). Das 88-kDa Protein besitzt, ebenso wie der Transkriptionsrepressor MBD1 und viele andere Proteine, ein hoch-konserviertes CXXC-Motiv (Bestor *et al.*, 1994; Voo *et al.*, 2000). Dieses Cystein-X-X-Cystein-Motiv vermittelt in Anwesenheit von Zink eine Bindung von CGBP an dsDNA, welche CpGs enthält. Dabei ist bereits ein CpG ausreichend für eine Protein-DNA-Bindung. Die Affinität von CGBP zur DNA steigt jedoch mit der Zahl der in der Sequenz vorliegenden CpGs (Lee *et al.*, 2001).

Darüber hinaus enthält das Protein zwei pflanzliche Homeodomänen (PHDs), welche für Chromatin-assoziierte Proteine oder Regulatoren der Genexpression charakteristisch sind. Welche Rolle PHDs bei der CGBP-vermittelten Transaktivierung der Genexpression spielen, ist allerdings derzeit noch unklar. Dennoch ist erwiesen, dass das murine CpG-bindende Aktivatorprotein an entwicklungspezifischen Vorgängen beteiligt ist. So sind z.B. CGBP-/- *Knock-out*-Embryonen bereits kurz nach Befruchtung nicht mehr lebensfähig. Diskutiert wird für dieses Protein daher eine Regulation ontogenetischer Mechanismen, möglicherweise durch Aktivierung CpG-Insel-assoziiierter Gene (Carlone *et al.*, 2001). Der epigenetische Regulationsmechanismus via CpGs unterliegt somit einem komplexem Zusammenspiel von Transkriptionsrepressoren und aktivatoren auf molekularer Ebene.

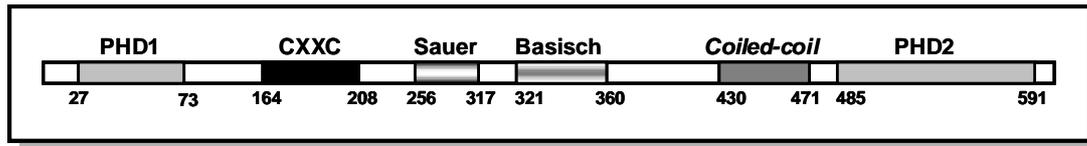


Abbildung A.5: Übersicht über bekannte Domänen des Transkriptionsaktivators hCGBP (humanes CpG-Bindeprotein). PHD: pflanzliche Homeodomäne; CXXC: Cystein-X-X-Cystein-Motiv. (Modifiziert nach Lee *et al.*, 2001)

A.3.4 Molekulare Mechanismen der epigenetischen Regulation via CpGs

Die Erhaltung und Neueinführung der DNA-Methylierung erfordert verschiedene Methyltransferase-Klassen, deren Aufgabenteilung aber nicht absolut strikt ist (Bestor *et al.*, 2000). In Säugerzellen liegen drei aktive DNA-Methyltransferasen (DNMTs) vor, welche jeweils die Übertragung einer CH₃-Gruppe auf das 5'-Atom des Cytosinrestes katalysieren (siehe Abbildung A.6). So obliegt dem Enzym DNMT1 vor allem die Erhaltungs-Methylierung von CpG-Dinukleotiden. Sein bevorzugtes Substrat ist hemimethylierte DNA, wie sie nach der DNA-Replikation in den Tochtersträngen vorliegt. DNMT2 zeigt nur sehr schwache Methyltransferaseaktivität und *Knock-out*-Mäuse für dieses Protein weisen keine wesentlichen Veränderungen in ihrem Methylierungsmusters auf (Okano *et al.*, 1998). Den Enzyme DNMT3a und DNMT3b hingegen wird eine *de novo* Methylierung von CpG-Zielsequenzen zugeordnet, wie sie in der Embryonalentwicklung erfolgt (Okano *et al.*, 1999). Die essentielle Funktion der Enzyme DNMT1, DNMT3a und DNMT3b ergibt sich aus der Tatsache, dass für keines der drei Enzyme lebensfähige *Knock-out*-Mäuse generiert werden konnten. Zudem führen Defekte der *de novo* Methyltransferase DNMT3b beim Menschen zum ICF-Syndrom, welches sich durch Immundefekte, Centromer-Instabilität und Faciale Dysmorphie auszeichnet (Xu *et al.*, 1999; Clark *et al.*, 2002).

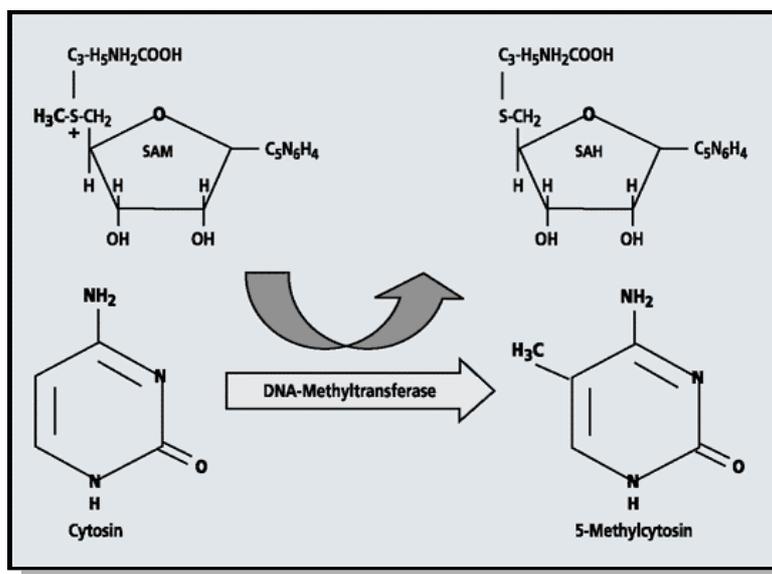


Abbildung A. 6: Entstehung von 5-Methylcytosin. Durch das Enzym DNA-Methyltransferase wird eine Methylgruppe von S-Adenosylmethionin (SAM) auf die Base Cytosin übertragen. Es entsteht 5-Methylcytosin und S-Adosylhomocystin (SAH). (Modifiziert nach Attwood *et al.*, 2002)

Neben der Übertragung von Methylresten auf CpG-Dinukleotide beeinflussen die DNA-Methyltransferasen DNMT1, DNMT3a und DNMT3b auch auf direktem Weg die Regulation der Transkription. Mittels spezifischer Domänen können die Enzyme an die DNA binden und somit potentielle Bindestellen für Proteine oder Transkriptionsfaktoren sterisch blockieren (Karpf *et al.*, 2002; Bachman *et al.*, 2001; Deng *et al.*, 2001). Diesen negativen Regulationsmechanismus nutzen auch die mCpG-Bindeproteine MeCP2, MBD1 und MBD2, welche nach Anlagerung an die DNA eine erfolgreiche Transkription verhindern. Weiter können sowohl DNMTs als auch MBD-Proteine zusätzliche Inhibitoren der Transkription rekrutieren bzw. stellen selbst Komponenten eines Co-Repressor Komplexes dar (Rountree *et al.*, 2000, 2001; Robertson *et al.*, 2000). So besitzt MBD2 die Fähigkeiten mit dem Mi-2-Komplex zu interagieren, während MBD3 eine Komponente dieses Repressor Komplexes darstellt (Zhang *et al.*, 1999; Carlone *et al.*, 2001). Zudem können sowohl DNMTs als auch bestimmte mCpG-Bindeproteine wie z.B. MeCP2, MBD1, MBD2 und MBD3 Histon-Deacetylasen (HDACs) bzw. Histon-Methyltransferasen rekrutieren bzw. sind mit diesen assoziiert (Fuks *et al.*, 2003; Sarraf *et al.*, 2004). Diese Enzyme katalysieren die Deacetylierung und Methylierung der Histone H3 und H4, was zu einer festeren Bindung der Nukleosomen an die DNA und somit zu einer dichteren Packung des Chromatins führt. Eine dicht gepackte Chromatinstruktur weist eine schlechtere Zugänglichkeit der DNA für die Transkriptionsmaschinerie auf und vermindert somit die Genexpression in diesem Bereich (Jenuwein *et al.*, 2001). Darüber hinaus können auch mCpGs alleine aufgrund ihrer sterischen Beschaffenheit die Anlagerung spezifischer Transkriptionsfaktoren wie z.B. USF (*upstream stimulatory factor*) blockieren (Iguchi-Arigo *et al.*, 1989; Fujii *et al.*, 2006). Eine Übersicht über Mechanismen, anhand derer nach Methylierung von CpGs eine Initiation der Transkription verhindert wird, ist in Abbildung A.7 dargestellt.

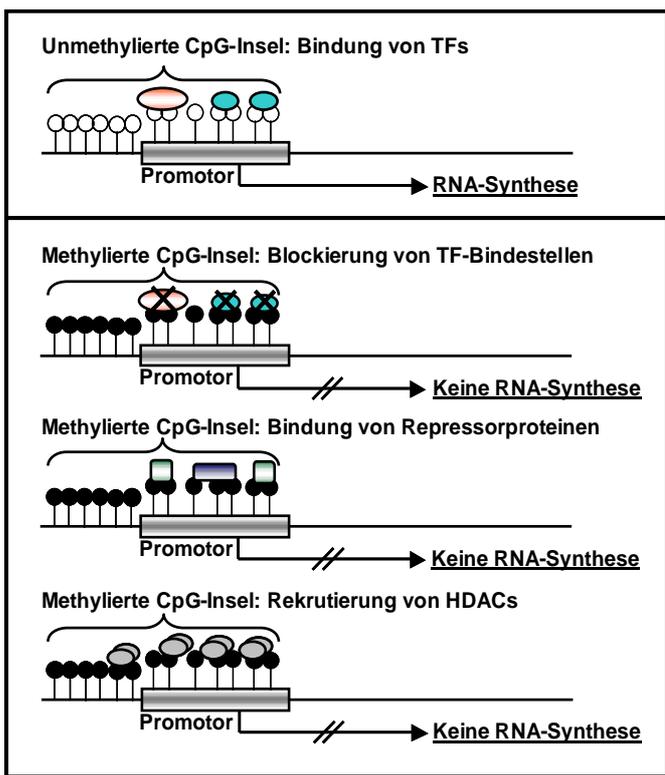


Abbildung A.7: Mechanismen der mCpG-vermittelten Inhibition der Transkription. Unmethylierte CpG-Dinukleotide sind als weiße Kreise dargestellt, methylierte CpG-Dinukleotide werden durch schwarze Kreise repräsentiert. Der Startpunkt der Transkription ist durch einen Pfeil gekennzeichnet. TF: Transkriptionsfaktor; HDAC: Histondeacetylase;

Unklar ist allerdings derzeit noch, ob ein Abschalten der Genexpression durch die abnormale Methylierung von CpG-Inseln bedingt wird, oder, ob das *Silencen* der Genexpression ein verändertes Methylierungsmuster zur Folge hat (Prokhortchouk *et al.*, 2002). So erfolgt eine Methylierung des zweiten X-Chromosoms bei der Frau erst nach Inaktivierung der Genexpression (Clark *et al.*, 2002). Zudem können umgekehrt auch Histon-Methylasen DNMTs in heterochromatische Regionen leiten und somit dort die DNA-Methylierung auslösen. Andere Daten belegen, dass erst die Histon-Acetylierung zur aktiven Demethylierung des betroffenen Genabschnittes führt (Tamaru *et al.*, 2001; Cervoni *et al.*, 2001). Diese Fragestellung erinnert an das Henne-Ei-Szenario und kann möglicherweise nur unter Annahme eines wechselseitigen Mechanismus, über den alle epigenetischen Modifikationen des Genoms verbunden sind, erklärt werden (Sarraf *et al.*, 2004; Klose *et al.*, 2006). Bis *dato* ist zudem unbekannt, auf welchen molekularen Vorgängen eine Transkriptionsaktivierung durch das CGBP beruht. Ungeachtet des zugrunde liegenden Mechanismus, spielen CpG-Dinukleotiden bei der Regulation der Genexpression in Eukaryonten eine wichtige Rolle.

Eine hohe Transgenexpression ist Grundvoraussetzung für gentherapeutische Anwendungen, die DNA-Impfstoffentwicklung und eine effektive, biotechnologische Produktion therapeutischer Wirkstoffe im Säugersystem. Um Expression und Immunogenität von Vektoren durch Einbau von CpG-Dinukleotiden (siehe A.3) bzw. CpG-Motiven (siehe A.2) gezielt verbessern zu können, bedarf es daher zunächst eines möglichst CpG-freien Ausgangsvektors. Eine CpG-Reduzierung funktioneller Vektorkomponenten erfordert jedoch genaue Kenntnis der Bereiche, die für den zugrundeliegenden molekularen Wirkungsmechanismus essentiell sind. Modifikationen in diesem Bereich könnten die Funktionalität der regulatorischen Elemente beeinträchtigen oder aufheben. Daher werden zunächst Aufbau und Funktionsweise ausgewählter, regulatorischer Vektorkomponenten beschrieben.

A.4 Komponenten eukaryontischer Expressionsplasmide

A.4.1 CMV-Promotor

Um eine hohe Transgenproduktion zu gewährleisten, werden in eukaryontischen Expressionsvektoren starke virale Promotoren verwendet. Der *Cytomegalovirus major immediate early*-Promotor/Enhancer-Komplex (CMV-Promotor) stellt dabei derzeit den stärksten bekannten eukaryontischen Promotorbereich dar (Loser *et al.*, 1998). Nahezu unabhängig von Spezies und Zellart vermittelt er eine hohe Expression *in vitro* sowie *in vivo* (Boshart *et al.*, 1985; Brooks *et al.*, 2004). Neben den für eukaryontische Promotoren typischen Elementen wie z.B. der TATA-Box, an welcher die Ausbildung des Transkriptionsinitiationskomplexes erfolgt, enthält der CMV-Promotor mehrere Wiederholungselemente (WH-Elemente) sowie multiple Bindestellen für spezifische Transkriptionsfaktoren (TFs). So binden die Transkriptionsaktivatoren NF κ B und CREB/ATF an die 18bp- und 19bp-Wiederholungsbereiche (Sambucetti *et al.*, 1989; Moens *et al.*, 2001). Das 21bp-

WH-Element hingegen beinhaltet einen negativen Regulator und hemmt die CMV-Promotor abhängige Transkription (Kothari *et al.*, 1991). Eine Zuordnung der einzelnen WH-Elemente und Transkriptionsfaktoren zur Sequenz des CMV-Promotors kann Abbildung A.8 entnommen werden. Die spezifischen Erkennungssequenzen der einzelnen, in der Promotorregion lokalisierten TFs sind in Abbildung D.2 aufgeführt.

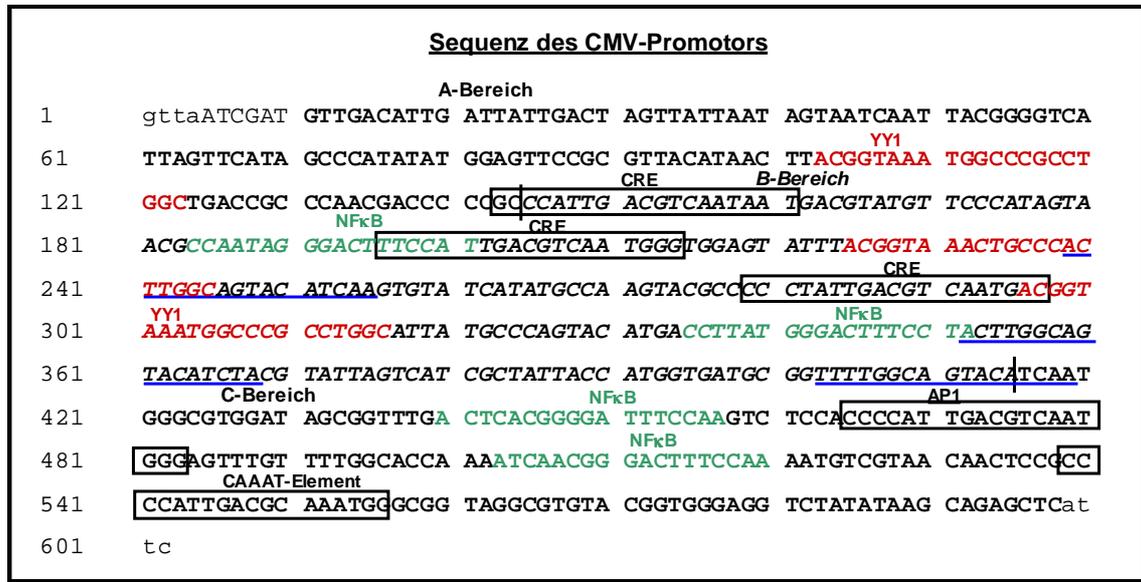


Abbildung A.8: Wiederholungselemente in der Sequenz des CMV-Promotors. In der Promotorregion lokalisierte WH-Elemente sind wie folgt gekennzeichnet bzw. farblich markiert: 17bp-WH-Element - blau unterstrichen, 18bp-WH-Element – grün markiert; 19bp-WH-Element - eingerahmt, 21bp-WH-Element – rot markiert. Die in der jeweiligen Region bindenden TFs sind in der Farbe des zugehörigen WH-Elementes angegeben. Aus Übersichtsgründen werden die Erkennungssequenzen für die einzelnen Transkriptionsfaktoren nicht aufgeführt, sondern können Abbildung D.2 entnommen werden. (Boshart *et al.*, 1985, Meier *et al.*, 1996). Der vordere Sequenzabschnitt (A-Bereich) und die 3'-randständige Sequenz (C-Bereich) sind normal gedruckt, während die Sequenz des mittleren B-Bereichs kursiv dargestellt ist.

Die Promotorstärke wird jedoch nicht nur durch *cis*-regulatorische Elemente beeinflusst, sondern kann auch in *trans* reguliert werden. So transaktivieren bestimmte virale Proteine wie z.B. das SV40 T-Antigen die Aktivität des CMV-Promotors (Moens *et al.*, 2001). Durch entzündliche Reaktionen hervorgerufene Cytokine reprimieren dagegen eine CMV-Promotor-vermittelte Expression in *trans* und können ein Abschaltend der Genexpression bewirken (Brooks *et al.*, 2004). Allgemeine und spezifische Transkriptionsfaktoren lagern sich schließlich im Promotorbereich zu mehreren spezifischen Nukleoproteinkomplexen zusammen (Niller *et al.*, 1991). Die Funktion bzw. das komplexe Zusammenspiel der einzelnen WH-Elemente und Transkriptionsfaktoren bei der Regulation der CMV-Promotoraktivität ist jedoch noch nicht vollständig aufgeklärt. Aus diesem Grund wurde in dieser Arbeit ein kombinatorischer Ansatz zur Generierung CpG-freier Promotorvarianten gewählt.

A.4.2 BGH-Polyadenylierungssignal

In eukaryontischen Zellen werden die Transkriptionsprodukte der RNA-Polymerase II (prä-RNA) in mehreren Prozessierungsschritten zu reifen RNA-Transkripten umgewandelt. So wird zunächst an das 5' Ende des Transkriptes eine Kappe aus methyliertem Guanin angehängt. Anschließend kommt es zu einer Entfernung von Introns durch Spleissen, sowie einer 3' Polyadenylierung der Vorläufertranskripte. Der im letzten Teilschritt angefügte Poly(A)-Schwanz ist notwendig für eine Initiation der Translation und schützt die messenger-RNA (mRNA) vor Degradation (Lodish *et al.*, 2001). In eukaryontischen Expressionsvektoren wird zur Terminierung der Transkription zumeist das BGH- (*Bovine Growth Hormone*) Polyadenylierungssignal eingesetzt, welches eine hohe Expressionsleistung vermittelt (Yew *et al.*, 1997; Xu *et al.*, 2002). Im Folgenden wird der Ablauf der Spaltung und Polyadenylierung von prä-RNA in Säugetierzellen näher erläutert (siehe Abbildung A.9).

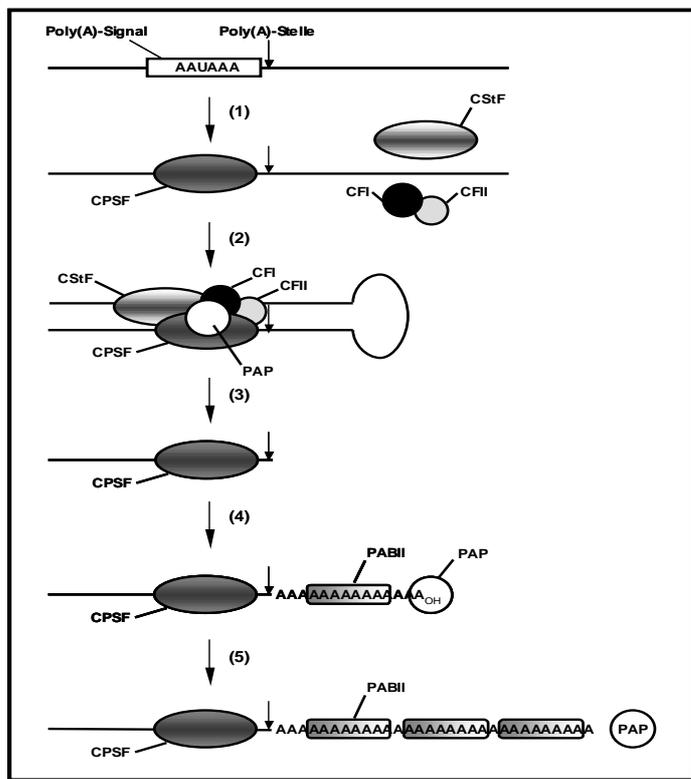


Abbildung A.9: Modell für Spaltung und Polyadenylierung von Prä-mRNA in Säugetierzellen. CPSF: *Cleavage and Polyadenylation Specific Factor*; CStF: *Cleavage Stimulatory Factor*; CFI: *Cleavage Factor I*; CFII: *Cleavage Factor II*; PAP: *Poly(A)-Polymerase*; PABII: *Poly(A)-Bindungsprotein II*. Die Poly(A)-Spaltstelle ist durch einen Pfeil gekennzeichnet. Nähere Erläuterungen können dem Text entnommen werden. (Modifiziert nach Lodish *et al.*, 2001).

Die komplexe Prozessierungsmaschinerie besteht aus wenigstens einem Dutzend Polypeptiden, wobei dem CPS-Faktor (*Cleavage and Polyadenylation Specific Factor*) eine zentrale Rolle zukommt. Das CPSF-Heterotetramer bindet im Bereich des Polyadenylierungssignals spezifisch an die AAUAAA-Sequenz (1, Abbildung A.9) der RNA (Proudfoot, 1989). In Anwesenheit des CSt-Faktors (*Cleavage Stimulatory Factor*), der Spaltungsfaktoren CF I und CF II sowie der Poly(A)-Polymerase (PAP) erfolgt zunächst eine endonukleolytische Spaltung, 10 bis 35 Nukleotide stromabwärts vom AAUAAA-Signal (2,3 Abbildung A.9). Im zweiten Schritt der 3' Prozessierung katalysiert die PAP das Anhängen eines Poly(A)-Schwanzes an die mRNA. Während dieses Prozesses gewährleisten CPSF und

ein Poly(A)-Bindungsprotein (PAB II) den Kontakt der PAP mit der zu verlängern- den RNA (4, Abbildung A.9). Erst nach Erreichen der spezifischen P(A)-Schwanz- länge von circa 250 Nukleotiden wird die Reaktion terminiert (5, Abbildung A.9).

A.4.3 pUC origin of replication

Neben starken Promotoren und Terminationssignalen zum Erhalt einer hohen Transgenexpression, besitzen eukaryontische Expressionsvektoren meist einen bakteriellen Replikationsursprung. Der sogenannte *origin of replication* ermöglicht eine Vermehrung des gewünschten Plasmides in Bakterien, unabhängig von der Replikation der genomischen bakteriellen DNA (Lodish *et al.*, 2001). Häufig beinhalten Vektoren den etwa 600 Basenpaare umfassenden pUC (ColE1) *origin of replication*, dessen Funktion nachfolgend erläutert wird (siehe Abbildung A.10). Der Replikationsprozess beginnt mit der Transkription der 555-bp-langen RNAII, welche unter der Kontrolle des vorgeschalteten RNA_{II}-Promotors steht. Die RNAII bildet eine spezifische Sekundärstruktur aus, deren G-reiche Schleife an eine Cytidin-reiche Region der *origin*-DNA (C-reiche Region) bindet. Das Enzym RNaseH erkennt derartige DNA/RNA-Hybride und schneidet den DNA/RNAII-Komplex so, dass im Bereich des *origins* ein an die Plasmid-DNA angelagerter RNA-Primer entsteht (Itoh *et al.*, 1979). Die freie 3'-OH-Gruppe des RNA-Primers bildet anschließend die Ansatzstelle für die DNA-Polymerase, welche nun eine Replikation des Plasmides vermittelt (Jung *et al.*, 1995).

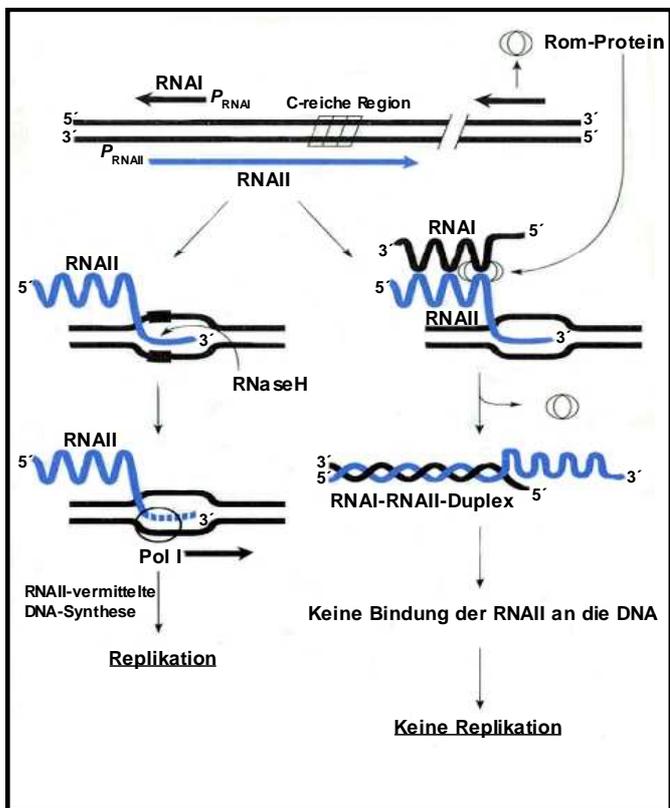


Abbildung A10: Regulation der Plasmidreplikation.

(Nach <http://www.mun.ca/biochem/courses/4103>)

Dieser Replikationsmechanismus wird über das 108-bp-lange RNAI-Transkript und das Rom-Protein reguliert. Dabei ist die Sequenz der RNAI revers komplementär zum 5'-Bereich der RNAII (siehe Abbildung D.1). Nach Stabilisierung durch das Rom-Protein bilden die RNAI und der 5'-Bereich der RNAII eine stabile RNA-RNA-Duplex aus, wodurch schließlich die Plasmidreplikation inhibiert wird (Kues *et al.*, 1989; Cesareni *et al.*, 1982). Dieser negative Regulationsmechanismus sorgt bei *Low-copy*-Plasmiden für eine Kontrolle der vorliegenden Zahl an Plasmidkopien. Ist das RNAII-Transkript jedoch länger als 360 Basenpaare, wie es beim pUC *origin of replication* (siehe Abbildung D.1, grüne Sequenz) der Fall ist, geht die inhibitorische Funktion der RNAI verloren. Plasmide mit einem derartigen Replikationsursprung vermitteln in Bakterien eine hohe Kopienzahl und werden als *High-copy*-Plasmide bezeichnet (Merlin *et al.*, 1995).

A.5 Gensynthese

Unter Berücksichtigung funktioneller Mechanismen wurden für kodierende und regulatorische Vektorkomponenten sowie für verschiedene Transgene CpG-modifizierte Sequenzen konzipiert. Um diese theoretisch konstruierten DNA-Fragmente in der Praxis herzustellen, wurde die Gensynthese als „molekularbiologisches Werkzeug“ gewählt. Neben der klassischen Gensynthese wurde ein Teil der Konstrukte mittels kombinatorischer Verfahren hergestellt. Dabei impliziert der Begriff Gensynthese die Herstellung natürlich vorkommender sowie künstlich veränderter dsDNA-Fragmente aus einzelsträngigen Oligonukleotiden. Derzeit können in polymerase- bzw. ligasebasierten Verfahren bis zu mehreren Kilobasen große dsDNA-Fragmente sowie komplette Plasmide synthetisch generiert werden (Wagner, 2001).

Förderlich ist die Verwendung synthetischer DNA-Sequenzen z.B. bei der Gewinnung rekombinanter Proteine im heterologen Expressionssystem. Der genetische Code besitzt zwar nahezu universelle Gültigkeit, die für eine Aminosäure synonymen Kodons werden jedoch von verschiedenen Spezies im Verhältnis zueinander unterschiedlich häufig genutzt. Zudem besteht eine starke Korrelation zwischen dem sogenannten Kodongebrauch eines Organismus und den entsprechenden t-RNA-Häufigkeiten (Ikemura, 1982, 1985). Daher ist die Expression eines Transgens in einem anderen System als dem Ursprungsorganismus oftmals limitiert oder gar nicht möglich. So schlugen erste Versuche fehl, das *Green-fluorescent-protein* aus der Qualle *Aequorea victoria* in humanen Zellen zu exprimieren. Erst nach Anpassung der Ausgangssequenz an den Kodongebrauch des Menschen wurde eine zufriedenstellende Expression des fluoreszierenden Reporterproteins erzielt (Zolotukhin *et al.*, 1996). Um eine effiziente Produktion rekombinanter therapeutischer sowie diagnostischer Proteine in gängigen pharmazeutischen Expressionssystemen wie z.B. *Pichia pastoris* oder CHO-Zellen (Vernon, 2004) zu gewährleisten, ist somit eine Anpassung der Ausgangssequenz an die im jeweiligen Zielorganismus vorherrschende

Kodonwahl nötig (Graf *et al.*, 2004, Zhou *et al.*, 2004). Im Rahmen dieser sogenannten „Optimierung“ (siehe B.1.5) können weitere Parameter, wie z.B. Spleissstellen, vorzeitige Polyadenylierungssignale, Sequenzrepetitionen sowie CpG-Motive und CpG-Dinukleotide eliminiert oder in die Sequenz integriert werden, um eine verbesserte Expression zu erreichen bzw. wissenschaftliche Fragestellungen zu adressieren.

Neben signifikanten Kosteneinsparungen bei der Produktion bestehender pharmazeutischer Therapeutika trägt die Gensynthese durch Bereitstellung von randomisierten Genvarianten bzw. Genbibliotheken aber auch zur Identifikation neuer Wirkstoffklassen bei. Auch im Bereich der DNA-Vakzinierung haben sich synthetische Gene bewährt, da sie neben einer erhöhten Transgenexpression verbesserte Sicherheitsaspekte aufweisen. So vermindert die verringerte Sequenzhomologie zu Wildtyp-Sequenzen die Wahrscheinlichkeit von Rekombinationsereignissen und somit die Generation chimärer, potentiell pathogener Konstrukte. Des Weiteren spielt die Gensynthese eine bedeutende Rolle im Rahmen der Gentherapie. Die künstliche Konstruktion von Genfähen mit möglichst geringer Immunogenität erhöht dabei die Chance einer erfolgreichen Behandlung eines Gendefektes, da Immunreaktionen, welche zur Eliminierung positiv transfizierter Zellen führen, vermindert werden (Bestor *et al.*, 2000; Reyes-Sandoval *et al.*, 2004). Zudem stellt die Gensynthese derzeit eine Alternative zur Isolierung seltener oder schwer isolierbarer Genvarianten dar. Im Rahmen dieser Arbeit sollte mittels synthetisch hergestellter Vektorkomponenten die Rolle von CpG-Dinukleotiden bei der Regulation der Transgenexpression untersucht werden.

A.6 Zielsetzung

Im Bereich der modernen Medizin sowie in der Biotechnologie werden fremde Transgene meist unter Verwendung bakterieller Plasmidvektoren in das jeweilige Expressionssystem eingebracht. In bakterieller DNA gehäuft auftretende stimulierende CpG-Motive (siehe A.2) beeinflussen in erheblichem Maße die Immunogenität und Transgenexpression eines Vektorkonstruktes. Um unerwünschte Immunreaktionen zu unterbinden und eine stabile Transgenexpression zu gewährleisten, war es Ziel dieser Arbeit, eine modular aufgebaute, CpG-reduzierte Vektorplattform zu entwickeln. Neben kodierenden Bereichen sollten in kombinatorischen Ansätzen auch regulatorische Elemente wie z.B. der CMV-Promotor oder das BGH P(A)-Signal hinsichtlich ihres CpG-Gehaltes minimiert werden. Die erhaltenen CpG-reduzierten Vektorkonstrukte sollten anschließend in transienten und stabilen Zellkulturexperimenten auf ihre Funktionalität und Expressionsleistung hin untersucht werden. Ein „immun-neutrales“ Ausgangsplasmid kann schließlich durch gezielte Integration immunmodulierender Sequenzen je nach Anwendungszweck in seiner Immunwirkung verändert werden.

Darüber hinaus sind CpG-Dinukleotide (siehe A.3) wesentlich an der epigenetischen Regulation der Genexpression beteiligt. Unter anderem wird das Vorliegen

methylierter CpG-Dinukleotide im Bereich des Transgens für ein Abschalten der zugehörigen Genexpression verantwortlich gemacht. Da eine hohe und lang andauernde Expression des Transgens jedoch für viele Plasmid-Applikationen wie z.B. die DNA-Vakzinierung Grundvoraussetzung ist, sollte im Rahmen dieser Arbeit der Zusammenhang von CpG-Dinukleotid-Gehalt im Leserahmen und Höhe und Dauer der resultierenden Transgenexpression näher untersucht werden. Dazu sollten CpG-freie sowie CpG-maximierte Varianten der Reportergene GFP und p24 synthetisch hergestellt werden und in stabilen und transienten Analysen hinsichtlich ihrer Reporteraktivität und Proteinproduktion mit dem jeweiligen Ausgangskonstrukt verglichen werden. Exemplarisch sollten die GFP-Reportervarianten auf transkriptioneller Ebene bezüglich verschiedener Faktoren, wie z.B. vorliegender RNA-Transkriptmenge, Kernexportrate, RNA-Stabilität sowie der Existenz alternativer Spleissprodukte charakterisiert werden. Dadurch sollten detaillierte Erkenntnisse über die Beteiligung von CpG-Dinukleotiden an der Regulation der Transgenexpression gewonnen werden. Eine Kombination von Vektoren und kodierenden Bereichen unterschiedlichen CpG-Gehaltes soll schließlich die Evaluierung und Definition applikationsspezifischer Sequenzanforderungen ermöglichen.

B Material und Methoden

B.1 Klonierungen und gentechnische Arbeiten

Soweit nicht anders vermerkt, wurden für Klonierungen in *Escherichia coli* (*E. coli*) Standardmethoden angewendet (Sambrook *et al.*, 1989). Restriktionsenzyme und T4 DNA-Ligase wurden von Roche (Mannheim) und New England Biolabs (Schwalbach, Taunus) bezogen. Für die Amplifikation von DNA-Abschnitten aus Plasmid-DNA sowie zum Reamplifizieren von PCR-Produkten wurden die *TaqPlus*[®] Precision DNA-Polymerase (Stratagene, Heidelberg) und Nukleotide von Applied Biosystems (Weiterstadt) verwendet. Größere Mengen an Plasmid-DNA wurden über *Nucleobond-Tip AX500* bzw. *AX100* Säulen (Macherey & Nagel, Düren) nach Herstellerangaben aufgereinigt. Zur Aufreinigung von PCR-Produkten wurde der *QIAquick Purification Kit* (Qiagen, Hilden) verwendet. Die Isolierung von DNA-Fragmenten aus Agarosegelen wurde mit Hilfe des *QIAquick Gel Extraction Kits Mini* (Qiagen, Hilden) nach Herstellerangaben durchgeführt. Sequenzierungen zur Verifizierung von DNA-Abschnitten erfolgten nach der Methode von Sanger *et al.* (1977) unter Verwendung eines automatischen Sequenziergerätes der Firma Geneart (Regensburg). Alle übrigen Chemikalien waren kommerzielle Reagenzien von höchster Reinheit.

B.1.1 Bakterien und Nährmedien

Zu Klonierungszwecken und zur Amplifikation von *High-copy*-Plasmiden wurde hauptsächlich der *E. coli* K-12 Stamm **DH5 α** (Life Technologies, Karlsruhe) verwendet. Bei Verwendung von *dam*-sensitiven Enzymen wie z.B. *Cla*I wurden die jeweiligen Plasmide im Stamm **GM2163** (New England Biolabs, Schwalbach, Taunus) amplifiziert. Chemisch transformationskompetente Bakterien wurden nach der RbCl-Methode hergestellt. Die verwendeten Bakterienstämme weisen den wie folgt angegebenen Genotyp auf:

DH5 α : F⁻ *supE44* Δ *lacU169* (ϕ 80 *lacZ* Δ M15) *hsdR1* *recA1* *endA1* *gyrA96* *thi-1* *relA1* (Dieser Stamm ist zur Blau-Weiß Selektion positiver Transformanten geeignet; Hanahan, 1983.)

GM2163: F⁻ *dam-13::Tn9* *dcm-6* *hsdR2* *leuB6* *his-4* *thi-1* *ara-14* *lacY1* *galK2* *galT22* *xyl-5* *mtl-1* *rpsL136* *tonA31* *tsc-78* *supE44* McrA⁻ McrB⁻ (Dieser Stamm wird für Klonierungsarbeiten mit *dam*- und *dcm*-sensitiven Restriktionsenzymen verwendet.)

Die Anzucht der Bakterien erfolgte in *Luria Bertani* (LB-) oder *Terrific Broth* (TB-) Flüssigmedium bzw. auf LB-Agarplatten über Nacht bei 37°C. Zur Selektion der positiven Transformanten wurde entsprechend dem verwendeten Plasmid Ampicillin (100 μ g/ml) oder Kanamycin (25 μ g/ml) zugesetzt.

B.1.2 Oligonukleotide

Sämtliche Oligonukleotidsequenzen sind im Anhang (siehe F.1) angegeben. Oligonukleotide wurden von den Firmen Metabion (Martinsried), Invitrogen (Leek, Niederlande) bzw. Geneart (Regensburg) bezogen. Degenerierte Oligonukleotide sowie Oligonukleotide zur Herstellung synthetischer Gene wurden von der Firma Geneart (Regensburg) hergestellt. Wegen ihrer großen Anzahl werden die für Gensynthesen verwendeten Oligonukleotide nicht aufgeführt, sondern können den kodonoptimierten Sequenzen entnommen werden (siehe Anhang F.2.1). Alle Oligonukleotide für die im Rahmen dieser Arbeit synthetisch hergestellten Gene wurden mit dem Programm *GeneOptimizerTM* (Geneart, Regensburg) unter Verwendung der Genbank-Sequenz entworfen. Zudem wurden an den endständigen Oligonukleotiden für die nachfolgenden Klonierungen passende singuläre Schnittstellen sowie vier zusätzliche, beliebige Nukleotide angefügt, um die Effektivität der Restriktionsenzyme bei den nachfolgenden Klonierungsarbeiten zu gewährleisten.

B.1.3 Plasmide

Detaillierte Gensequenzen sind im Anhang unter Punkt F.3 angegeben.

Zur Herstellung der Reporterplasmide wurde das jeweiligen Transgen unter Verwendung der 5' *HindIII*- sowie der 3' *BamHI*-Restriktionsschnittstellen in den multiplen Klonierungsbereich (MCS) der eukaryontischen Expressionsvektoren kloniert.

B.1.3.1 Das Flp-InTM-System (Invitrogen)

Dieses DNA-Rekombinationssystem macht sich die Fähigkeit der Flp-Rekombinase aus *Saccharomyces cerevisiae* zunutze. So ermöglicht die Flp-Rekombinase eine zielgerichtete Integration des gewünschten Gens in einfacher Ausführung ins Genom der Wirtszelle (Craig, 1988; Sauer, 1994). Damit können rekombinante Zelllinien generiert werden, deren Expressionsleistung direkt verglichen werden kann.

B.1.3.1.1 pFRT/lacZeo Vektoren

Das Flp-In System besteht aus drei verschiedenen Vektoren. Die erste Komponente, der Vektor pFRT/*lacZeo* bzw. pFRT/*lacZeo2*, ist in das Genom der Wirtszelllinie integriert. Dieser Vektor trägt unter der Kontrolle des *SV40 early promoters* (293T-Zellen) bzw. eines veränderten *SV40 early promoters* mit verringerter Aktivität (CHO-Zellen) ein *lacZ*-Zeocin Fusionsgen. Eine FRT-Signal (*Flp Recombination Target*-Signal) nach dem jeweiligen *SV40*-Promotor und dem ATG-Startkodon fungiert als Bindungs- und Insertionsstelle für die Flp-Rekombinase. Die Zeocin-Resistenz im Vektor erlaubt dabei eine Selektion auf positive Klone und somit die Aufrechterhaltung der Wirtszelllinie. Außerdem besitzt

der Vektor einen *pUC origin* und ein Ampicillin-Resistenzgen für das Wachstum und die Selektion in *E. coli*. Bei Flp-In CHO-Zellen sorgt der veränderte *SV40 early promotor* mit verringerter Aktivität für die Integration des Vektors pFRT/lacZeo2 in eine transkriptionell hoch aktive Region im Genom, da nur so die für eine Zeocinresistenz erforderliche Zeocin Expression gewährleistet werden kann.

B.1.3.1.2 pOG44-Expressionsplasmid

Das 5785-bp-große Plasmid pOG44 stellt eine weitere Komponente des Flp-In-Systems dar. Die Transkription des Flp-Gens erfolgt unter der Kontrolle des hCMV (humanes Cytomegalievirus) *immediate early*-Promotors, welcher im Folgenden als CMV-Promotor bezeichnet wird. Zur Erhöhung der RNA-Stabilität enthält der Vektor das SV40-Polyadenylierungssignal (*SV40 late polyadenylation signal*), welches auch eine Transkriptionsterminations-Sequenz enthält. Zwischen CMV-Promotor und Flp-Gen befindet sich ein synthetisches Intron, das für eine gesteigerte Expression des Flp-Gens sorgt. Das Flp-Gen kodiert für eine temperaturabhängige Flp-Rekombinase, die eine konservative Rekombination zwischen den FRT-Signalen (*Flp-In recombination target site*) vermittelt. Für die Replikation und das Wachstum in *E. coli* besitzt der Vektor einen *pUC origin* und für die Selektion positiver bakterieller Transformanten eine Ampicillinresistenzkassette.

B.1.3.1.3 pcDNA5/FRT

Der 5.1-kb-große Expressionsvektor ist Träger des gewünschten Transgens sowie eines weiteren FRT-Signals. Eine Kotransfektion dieses Vektors mit dem pOG44-Expressionsplasmid, das für die Flp-Rekombinase kodiert, führt in Flp-In Säugerzellen zu einer homologen Rekombination der FRT-Stellen im Genom und auf dem Expressionsvektor. Damit kommt es zu einer gezielten Integration des gewünschten Gens in das Genom der Wirtszelle. Die Expression der integrierten Gensequenzen erfolgt durch die Promotorregion aus dem *immediate early gene* des hCMV (CMV-Promoter). Zudem enthält der Vektor die *T7 promotor/priming*-Region für die *in vitro* Transkription von *sense* RNA und die Sequenzierung des Transgens. Weiterhin enthält der Expressionsvektor pcDNA5/FRT ein Polyadenylierungssignal und eine Transkriptionsterminations-Sequenz aus dem *bovine growth hormone* (BGH) (Goodwin *et al.*, 1992) sowie das Hygromycinresistenzgen (ohne ATG) zur Selektion stabiler, positiv transfizierter Zellklone, das von einem SV40-Polyadenylierungssignal terminiert wird. Zusätzlich beinhaltet der Vektor ein Ampicillinresistenzgen (*amp*) für die Selektion positiver, bakterieller Transformanten, sowie einen *pUC origin* für die Replikation in *E. coli*.

B.1.3.1.4 Schematische Darstellung

Die nachfolgende Abbildung zeigt das Zusammenspiel der einzelnen Komponenten des Flp-In Systems. In der jeweiligen Wirtszelllinie befindet sich der Vektor pFRT/*lacZeo* (293T) bzw. pFRT/*lacZeo*2 (CHO) stabil in das Genom integriert. Daraus resultiert die stabile Expression des *lacZ*-Zeocin-Fusionsgens, das die Aufrechterhaltung der gewünschten Zelllinie sowie eine X-Gal-Färbung ermöglicht (1). Im nächsten Schritt wird der pcDNA5/FRT-Expressionsvektor, der das Transgen (GOI; *gene of interest*) enthält, zusammen mit dem für die Flp-Rekombinase kodierenden Plasmid pOG44 im Verhältnis 1:9 in die Wirtszelllinie kotransfiziert (2). Nach Expression der Flp-Rekombinase katalysiert diese eine homologe Rekombination zwischen den FRT-Signalen im Genom der Wirtszelllinie und auf dem Expressionsvektor (3). Bei erfolgreicher Integration des pcDNA5/FRT-Expressionskonstruktes erhält das Hygromycinresistenzgen ein Start-ATG und wird unter der Kontrolle des jeweiligen SV40-Promoters exprimiert (4). Stabil transfizierte Zellen weisen daher Hygromycinresistenz auf und können somit positiv selektioniert werden. Das *lacZ*-Zeocin-Fusionsgen dagegen steht nach homologer Flp-Rekombination nicht mehr unter der Kontrolle des SV40-Promotors (4), weswegen stabil transfizierte Zellen Zeocin-Sensitivität aufweisen und die Aktivität des *lacZ*-Operons verlieren.

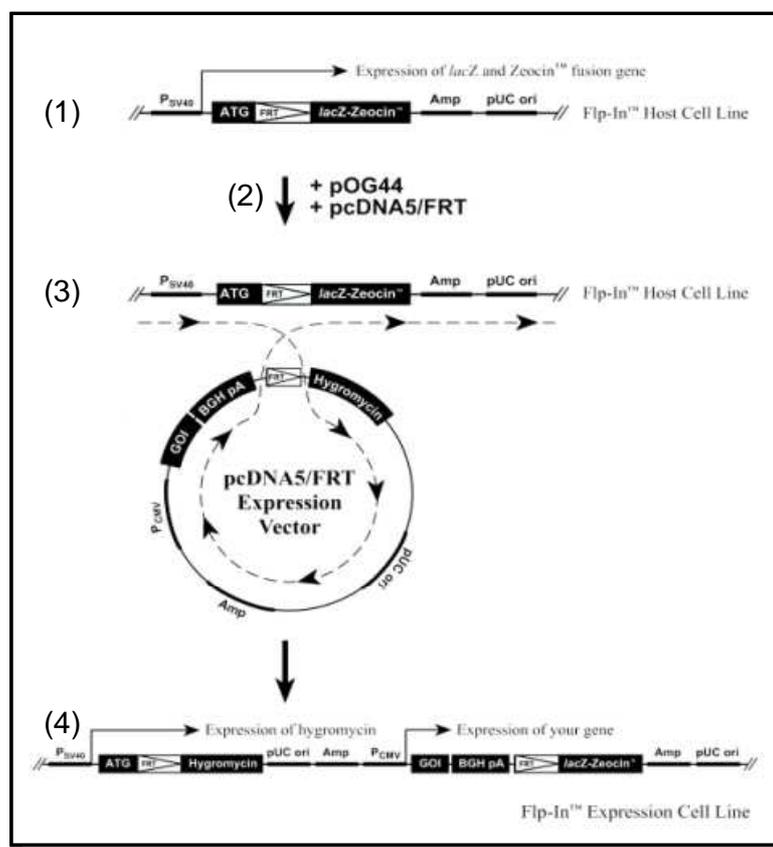


Abbildung B.1: Schematische Darstellung des Flp™-In Systems. Nähere Informationen können dem Text entnommen werden.

B.1.4 PCR-Techniken (Polymerase chain reaction)

Mit Hilfe von Startermolekülen (Primern) wurden gezielt DNA-Fragmente amplifiziert, gewünschte Mutationen in Plasmid-DNA eingeführt sowie *E. coli*-Kolonien nach der Transformation auf positive Transformanten getestet. Die Durchführung der PCR-Analysen erfolgte in einem programmierbaren Thermocycler (Mastercycler Gradient Eppendorf) bzw. einem *iCycler* der Firma BioRad (München). Für alle in der Arbeit durchgeführten PCR-Analysen wurde die *TaqPlus*[®] Precision DNA Polymerase (Stratagene, Heidelberg) verwendet.

B.1.4.1 Standard PCR-Bedingungen

Für Plasmid-DNA wurden folgende PCR-Standardbedingungen zur Amplifizierung von Fragmenten sowie zum Anfügen von Restriktionsstellen an den 5' bzw. 3' Bereichen der gewünschten Sequenz verwendet.

Tabelle B.1: Standard-PCR-Protokoll (750-bp Fragment)

Komponenten	Volumen	PCR-Bedingungen:	Temperatur	Zeit
H ₂ O _{dest}	42,0 µl	1. Denaturierung	95°C	30 sec
dNTPs (10 mM ;je 2,5 mM)	1 µl	2. Amplifizierung*	95°C	30 sec
5'Primer pf (100 µM)	0,5 µl		55°C	30 sec
3'Primer pb (100 µM)	0,5 µl		72°C	45 sec
Matrize (250 ng/µl); bei Kontrolle H ₂ O _{dest}	0,5 µl			
10-fach Puffer	5 µl	3. Elongation	72°C	4min
<i>TaqPlus</i> TM Precision Polymerase (5 U/µl)	0,5 µl			
Endvolumen	50 µl			

* Schritt 2: 30 Zyklen

Der Test auf positive Transformanten wurde gemäß dem oben beschriebenen Standard-PCR-Ansatz durchgeführt. Es wurde jedoch der Qiagen-Mastermix (Qiagen, Hilden) verwendet und das Endvolumen halbiert (25 µl). Zusätzlich wurde der erste Denaturierungsschritt auf vier Minuten verlängert.

B.1.4.2 Gradienten-PCR

Die Analyse größerer Fragmente genomischer DNA (> 1,2 kb) mit zwei spezifischen Primern erfolgte mittels Gradienten-PCR-Analysen. Dabei wurde die *Primer-Annealing*-Temperatur bei jedem Zyklus um 0,5 °C verringert und somit der geeignete Temperaturbereich zur Amplifikation des jeweiligen Fragmentes genomischer DNA ermittelt. Diese Methode ermöglichte den spezifischen Nachweis von in einfacher Ausführung im Genom vorliegender DNA-Sequenzen.

Tabelle B.2: Gradienten-PCR-Protokoll (750-bp Fragment)

Komponenten	Volumen	PCR-Bedingungen:	Temperatur	Zeit
H ₂ O _{dest}	42,0 µl	1. Denaturierung	95°C	4 min
dNTPs (10 mM ;je 2,5 mM)	1 µl	2. „Gradienten-PCR“	95°C	30 sec
5'Primer pf (100 µM)	0,5 µl	Amplifizierung*	60°C-50°C	30 sec
3'Primer pb (100 µM)	0,5 µl		72°C	30 sec
Matrize (250 ng/µl); bei Kontrolle H ₂ O _{dest}	0,5 µl	3. Amplifizierung*	95°C	30 sec
10-fach Puffer	5 µl		45°C	30 sec
TaqPlus™ Precision Polymerase (5 U/µl)	0,5 µl		72°C	30 sec
Endvolumen	50 µl	4. Elongation	72°C	7 min

* Schritt 2: 30 Zyklen; Schritt 3: 10 Zyklen

B.1.4.3 Fusions-PCR

Eine Veränderung der Plasmid-DNA an definierten Stellen sowie die gezielte Kombination von DNA-Fragmenten erfolgte mittels Fusions-PCR. So wurden mit dieser Methode Restriktionsschnittstellen eingefügt bzw. CpG-Dinukleotide durch Verwendung degenerierter Oligonukleotide alternativ ausgetauscht. Zudem wurden partiell CpG-freie Promotorchimären zu vollständig CpG-freien Varianten zusammengefügt. Eine schematische Darstellung der ortsgerichteten Mutagenese durch Fusions-PCR zeigt Abbildung B.2.

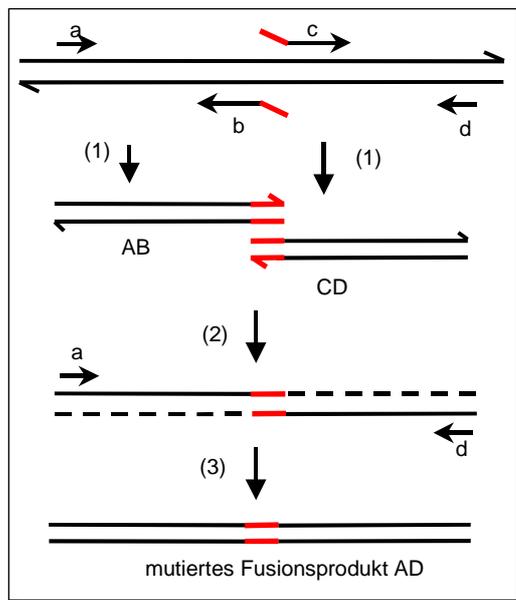


Abbildung B.2: Schematische Darstellung der ortsgerichteten Mutagenese durch Fusions-PCR. Zunächst werden in separaten PCR-Ansätzen zwei Teilfragmente (AB und CD) der zu mutierenden Sequenz generiert (1). Jeder Ansatz enthält dabei ein flankierendes Oligonukleotid, das an das eine Ende der Zielsequenz bindet (Amplifikationsprimer a oder d) und ein internes Oligonukleotid (b oder c), das sich neben der Mutationsstelle anlagert und über sein nicht paarendes 5'-Ende die gewünschten Mutationen einführt. Durch die Verwendung von teilweise komplementären, internen Mutageneseoligonukleotiden erhalten die beiden Teilfragmente AB und CD überlappende Enden, über die sie in einer zweiten PCR fusioniert werden können. Nach Denaturierung können die Einzel-

stränge der benachbarten Teilfragmente über ihre komplementären Enden paaren und so zum Teil als Primer für die Taq-Polymerase fungieren (2). Das Resultat dieser Primer-Verlängerung ist ein mutiertes Fusionsprodukt (3), welches anschließend in Anwesenheit des Oligonukleotidpaares a/d amplifiziert wird. Diese Methode ermöglicht die simultane Fusion von bis zu fünf DNA-Fragmenten durch Paarung der jeweils komplementären Enden.

Für einen Fusions-PCR-Ansatz wurden Standardbedingungen verwendet (B.1.4.1), wobei die Amplifikationsprimer erst nach 10 Zyklen zugegeben wurden. Die jeweiligen *Annealing*-Temperaturen wurden entsprechend der Schmelz-

temperatur der überlappenden Primerbereiche gewählt. Als Matritze dienten bis zu fünf für eine Fusion generierte Teilfragmente (je 250 ng/ μ l).

B.1.5 Optimierung und Herstellung synthetischer Gene

Die Optimierung der Ausgangssequenzen erfolgte unter Verwendung der Software *GeneOptimizer*TM (Geneart, Regensburg). Dabei wurde unter Beibehaltung der Aminosäuresequenz des Ausgangskonstruktes eine Nukleotidsequenz entwickelt, die hinsichtlich verschiedener unterschiedlich gewichteter Parameter verändert war. Neben der Modulation des CpG-Gehaltes und der Eliminierung bzw. Integration bestimmter singulärer Restriktionsschnittstellen wurde bei dem Design der synthetischen Konstrukte unter anderem darauf geachtet, keine negativ wirkenden *cis*-aktiven Elemente wie Spleisstellen, Polyadenylierungssignalstellen, interne TATA-Boxen oder Ribosomen-Bindestellen zu generieren. Kodierende Regionen wurden unter Beibehaltung der Kodonwahl des jeweiligen Organismus optimiert, um die resultierende Transgenexpression mit der des jeweiligen Ausgangskonstruktes vergleichen zu können.

Für jedes Konstrukt wurde deshalb der ermittelte CAI (*codon adaption index*) angegeben, der ein Maß für die Güte der verwendeten Kodons darstellt (Sharp *et al.*, 1987). Bei Verwendung ausschließlich optimaler Kodons wird im Idealfall ein CAI von 1 erreicht. Ein CAI größer als 0,9 führt aber bereits zu einer sehr hohen Transgenexpression, da zwischen dem Kodongebrauch eines Organismus und den entsprechenden tRNA-Frequenzen eine starke Korrelation besteht (Ikemura, 1985; Zhou *et al.*, 2004). Nicht nur der CAI als Durchschnittswert wurde während des Optimierungsprozesses weitestgehend aufrechterhalten, sondern auch die Häufigkeit der selten bzw. oft verwendeten Kodons blieb gleich. Zudem wurden alle generierten Sequenzen hinsichtlich ihres GC-Gehaltes, des Anteils der Nukleotide Cytidin und Guanosin an der Gesamtsequenz, mit ihrer nicht optimierten Version verglichen. Der GC-Gehalt eines synthetischen Gens sollte dem im jeweiligen Zielorganismus vorherrschendem GC-Gehalt angepasst sein. Zudem sollten extrem AT-reiche Sequenzen vermieden werden, da AT-reiche RNA-Transkripte instabil sind und zu einer verringerten Expression führen. Neben der Zahl an CpG-Dinukleotiden wird für die einzelnen Konstrukte auch der Gehalt an TpA-Dinukleotiden angegeben, da deren Beteiligung an der Regulation der Expression in der Literatur diskutiert wird (Duan *et al.*, 2003).

Die konstruierten Sequenzen wurden anhand einer PCR-basierten Gensynthesemethode künstlich hergestellt (Geneart, Regensburg) und unter Verwendung der Restriktionsenzyme *KpnI* und *SacI* in die Klonierungsvektoren pPCR-Script (Stratagene, Heidelberg) bzw. pUC19 (New England Biolabs, Schwalbach, Taunus) eingesetzt.

B.1.6 Isolierung und Analyse genomischer DNA aus eukaryontischen Zellen

Zur Untersuchung genomisch integrierter Plasmid-DNA wurde die genomische DNA aus stabil transfizierten Säugerzellen (siehe B.2.1) mit dem *QIAamp Blood Mini Kit* (Qiagen, Düren) nach Herstellerangaben isoliert. Dabei wurde das Protokoll für kultivierte Zellen verwendet. Anschließend wurde der gesamte Bereich des integrierten Transgens mit den Oligonukleotiden huGFP_pf/huGFP_pb für h_GFP 60, E019pf/E019pb für h_GFP 0 bzw. GFPmax_pf/GFPmax_pb für h_GFP 93 (siehe F.1.1) in einer Standard PCR-Analyse mit 30 Zyklen amplifiziert und die erhaltenen 740-bp PCR-Fragmente bezüglich ihrer Sequenz verifiziert. Um die Integration des Expressionsvektors in die gewünschte genomische Region nachzuweisen wurden weitere Gradienten-PCR-Produkte generiert. Dabei wurde ein Oligonukleotid im Bereich des SV40-Promoters der Zelllinie (FRT_pf) sowie ein Oligonukleotid im *lacZ*-Zeocin-Fusionsgen (FRT_pb) verwendet (72°C – 68°C *Annealing*; 7 min *Elongation*). Bei gewünschter Integration resultiert ein circa 7-kb großes DNA-Fragment aus der Kombination der Primer FRT_pf und FRT_pb. Zur Optimierung der PCR Parameter wurden mehrere Gradienten-PCR-Analysen durchgeführt. In alle PCR-Analysen zum Nachweis genomischer DNA wurde jeweils 1 µg genomischer DNA als Matrize eingesetzt.

B.2 Zellkulturtechniken und Infektionen

Sämtliche Zellkulturmedien und Zusätze wurden von PAN Biotech GmbH (Aidenbach), Invitrogen (GIBCO™, Karlsruhe) oder SIGMA (Deisenhofen) bezogen. Alle verwendeten Chemikalien waren speziell für Zellkulturen getestet. Kulturgefäße für Zellen wurden von BD Biosciences (Heidelberg) bezogen.

B.2.1 Kultivierung verwendeter Zelllinien

Alle Säugerzelllinien wurden bei 37°C und einem CO₂-Anteil von 5% kultiviert sowie nach Erreichen der Konfluenz im Verhältnis 1 : 10 subkultiviert (Ausubel *et al.*, 1987).

Die adhärent wachsende, p53-negative humane Lungenkarzinomzelllinie **H1299** (Mitsudomi *et al.*, 1992) wurde in Dulbecco's Modifiziertem Eagle Medium (*DMEM*) mit L-Glutamin, D-Glucose (4,5 mg/ml), Natriumpyruvat, 10% inaktiviertem fötalem Kälberserum (FKS), Penicillin (100 U/ml) und Streptomycin (100 µg/ml) gezogen. Weiter wurden **Flp-In™ 293T-Zellen** (Invitrogen, R750-07) verwendet. Diese adhärent wachsenden Ad5/SV40-transformierten, humanen Nierenepithelzellen besitzen stabil integriert in ihr Genom eine Kopie des Vektors pFRT//*lacZeo* (siehe B.1.3.1.1.) und exprimieren stabil das *lacZ*-Zeocin-Fusionsgen. Die Kultivierung der Flp-In™ 293T-Zelllinie erfolgte in Dulbecco's Modifiziertem Eagle Medium (*DMEM*) mit L-Glutamin, D-Glucose (4,5 mg/ml), Natriumpyruvat, 10% inaktiviertem fötalem Kälberserum (FKS), Penicillin (100 U/ml) und Streptomycin (100

µg/ml). Als Selektionsmarker wurde dem Medium ZeocinTM (Invitrogen, Leek, Niederlande) mit einer Konzentration von 100 µg/ml zugegeben. Bei den **Flp-InTM CHO-Zellen** (Invitrogen, R758-07) handelt es sich um Zellen aus dem Ovar eines Chinesischen Hamsterweibchens (*Cricetulus griseus*), in deren Genom stabil eine Kopie des Vektors pFRT/*lacZeo2* verankert ist (siehe B.1.3.1.1). Die adhärent wachsende Zelllinie wurde unter Verwendung des HAM's F12 Mediums mit L-Glutamin, 10% inaktiviertem fötalem Kälberserum (FKS), Penicillin (100 U/ml), Streptomycin (100 µg/ml) sowie mit ZeocinTM (100 µg/ml) als Selektionsmarker gezogen.

B.2.2 Bestimmung der Lebendzellzahl

Die Lebendzellzahl wurde mit Hilfe einer Neubauer-Zählkammer bestimmt. Hierzu wurden 50 µl einer gründlich resuspendierten Zellkultur mit der gleichen Menge einer 0,5%-igen Trypanblau-Lösung vermischt. Tote Zellen waren bei der nachfolgenden Mikroskopie durch eine deutliche Blaufärbung erkennbar und wurden nicht in die Berechnung der Lebendzellzahl einbezogen.

B.2.3 Transiente Transfektion eukaryontischer Zellen

$1,5 \times 10^6$ bzw. $2,5 \times 10^5$ H1299-Zellen wurden in Petrischalen (Fläche: 58,95 cm²) bzw. 6-Well-Schalen (Fläche: 9,6 cm²) ausgesät und 24h später durch Calciumphosphat-Präzipitation (Graham *et al.*, 1973) mit max. 45 µg (Petrischalen) bzw. mit max. 15 µg (6-Well-Schalen) Plasmid-DNA transfiziert. Zellen und Kulturüberstände wurden 48 bzw. 72 h *post transfectionem* geerntet.

B.2.4 Herstellung stabiler eukaryontischer Zellen (Flp-InTM System)

Zunächst wurden $1,5 \times 10^5$ Zellen (für CHO Flp-InTM) bzw. $2,5 \times 10^5$ Zellen (für 293T Flp-InTM) in 6-Well-Schalen (Fläche: 9,6 cm²) ausgesät und die Zellen in dem jeweiligen Medium mit Zeocin-Zusatz (100 µg/ml) über Nacht kultiviert. Eine Stunde vor Transfektion wurde bei den Zellen das Zeocin enthaltende Kulturmedium abgesaugt und gegen Zeocin-freies Medium ausgetauscht. 24h nach dem Aussäen wurden die ca. 80% konfluenten Zellen durch Calciumphosphat-Präzipitation mit 15 µg Plasmid-DNA transfiziert. Dabei wurden das Flp-Rekombinase-Expressionsplasmid pOG44 und der das Transgen beinhaltende Vektor pcDNA5/FRT im Verhältnis 9:1 (13,5 µg: 1,5 µg) in die Transfektion eingesetzt, um einen Überschuss an Flp-Rekombinase zu gewährleisten. Da eine erfolgreiche Integration des Expressionsvektors pcDNA5/FRT ins Genom der Wirtszelllinie den Funktionalitätsverlust des *lacZ*Zeocin-Fusionsgens zur Folge hat (siehe B.1.3.1.4), wurde nachfolgend nur **Zeocin-freies** Medium verwendet. 12 bis 14h (Flp-InTM CHO-Zellen) bzw. 8h (Flp-InTM 293T-Zellen) nach Transfektion erfolgte ein weiterer Medienwechsel mit Zeocin-freiem Kulturmedium.

Nach Erreichen der Konfluenz wurden die stabil transfizierten Zellen parallel in ZeocinTM-freiem Medium sowie in Standardmedium mit HygromycinB (PAA, Pasching, Austria) kultiviert. Dabei wurde für Flp-InTM CHO-Zellen eine Anfangskonzentration von 100 µg/ml HygromycinB gewählt, für Flp-InTM 293T-Zellen wurde dem Medium anfangs 12,5 µg/ml HygromycinB zugegeben. Die Zellen wurden erstmals mit dem Hygromycinmedium überschichtet, nachdem die Zellen sich am Boden der 6-Well-Schale abgesetzt hatten. Durch schrittweises Erhöhen der Hygromycinkonzentration (um 100 µg/ml für CHO- bzw. 12,5 µg/ml für Flp-InTM 293T-Zellen) wurden die Zellen auf über 90% GFP-positive Zellen selektioniert. Dabei wurde für Flp-InTM CHO-Zellen eine Endkonzentration von 500 µg/ml HygromycinB im Medium ermittelt während Flp-InTM 293T-Zellen bei einer Hygromycinkonzentration von 100 µg/ml stabil kultiviert wurden. Die unterschiedliche Hygromycintoleranz der beiden Zelllinien ist dabei laut Hersteller auf das Vorliegen des Resistenzgens in transkriptionell unterschiedlich aktiven Bereichen des Genoms zurückzuführen (siehe B.1.3.1).

B.2.5 Etablierung monoclonaler CHO-Zellklone

Um zu untersuchen, ob nach stabiler Transfektion und Selektion auf Hygromycin die Gesamtheit der erhaltenen hygromycinresistenten Zellen isogenisch ist, d. h. den Expressionsvektor an der gleichen Stelle im Genom integriert hat, wurden monoclonale CHO-Zellklone etabliert. Zur Etablierung dieser stabilen Zellklone wurden 1×10^3 stabil transfizierte CHO-Zellen nach Selektion auf 500 µg/ml Hygromycin in einer Petrischale ausgesät. 72h später wurden die zu Zellhaufen herangewachsen Einzelzellen unter dem Fluoreszenzmikroskop vereinzelt und in die Wells einer 96-Well-Platte überführt. Nach Erreichen der Konfluenz wurden die Zellen in 6-Well-Platten überführt und stabil kultiviert. Die einzelnen CHO-Zellklone wurden in FACS-Analysen auf die stabile Expression des gewünschten Transgens hin untersucht.

B.2.6 Langzeitkultur stabil transfizierter CHO-Zellklone

Um das Expressionsmuster der stabil transfizierten CHO- sowie 293T-Zelllinien über einen längeren Zeitraum untersuchen zu können, wurden die monoklonalen sowie die polyklonalen Ansätze über eine Dauer von 24 Monaten stabil kultiviert. Dabei wurden die Zelllinien in dem jeweiligen Standardmedium mit 500 µg/ml bzw. 100 µg/ml Hygromycin gezogen und nach Erreichen der Konfluenz (konstant über den untersuchten Zeitraum zweimal pro Woche) im Verhältnis 1 : 10 subkultiviert.

B.3 Analyse der transienten und stabilen Proteinexpression

B.3.1 Quantitative Bestimmung der Proteinkonzentration der Zelllysate

Die Bestimmung der Gesamtproteinmenge in den aufbereiteten Zelllysaten erfolgte nach der Methode von Bradford (Bradford, 1976) und wurde unter Verwendung des Bio-Rad Protein Assay Reagenz (Bio-Rad) nach Angaben des Herstellers durchgeführt. Die Proteinmenge wurde mit Hilfe einer BSA-Standardkurve bestimmt.

B.3.2 Quantifizierung des HIV-1 Kapsidantigens im p24-ELISA

Um die Menge an HIV-1 Kapsidprotein in Zelllysaten zu bestimmen, wurden transfizierte H1299-Zellen zweimal mit eiskaltem PBS (10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, 137 mM NaCl, 2,7 mM KCl) gewaschen, in eiskaltem PBS abgeschabt, 10 min bei 300 × g abzentrifugiert, in PBS resuspendiert und in den HIV-1 ELISA (siehe unten) eingesetzt. Alle Proben wurden so verdünnt, dass die Messwerte im linearen Bereich der Eichkurve lagen. Um die Menge an HIV-1 Kapsidprotein im Kulturmedium zu quantifizieren, wurden die Überstände 10 min bei 300 × g vorzentrifugiert, filtriert (Porengröße: 0,45 µm) und in den HIV-1 p24-ELISA eingesetzt.

Zur Quantifizierung des Kapsidantigens in Zelllysaten und Kulturüberständen wurden 96-Well *MaxiSorb* Mikrotiterplatten (Nunc, Wiesbaden) mit je 100 µl einer 1:300 Verdünnung des p24-spezifischen Antikörpers M01 in 0,1M Carbonatpuffer (pH 9.5) über Nacht (4°C) beschichtet. Anschließend wurden die Proben sowie der p24-Standard seriell (10-0 ng/ml) in Verdünnungspuffer (1% BSA in PBS) verdünnt und je 100 µl dieser Verdünnungen nach dreimaligem Waschen der Mikrotiterplatte mit Waschpuffer (0,05% Tween-20 in PBS) auf die Platte gegeben und 1h bei 37°C inkubiert. Nach wiederholtem Waschen (6×) wurden je 100 µl einer 1:20000 Verdünnung des zweiten, biotinylierten p24-spezifischen Antikörpers 37G12 in Verdünnungspuffer zugegeben und 1h bei Raumtemperatur (RT) inkubiert. Nach erneutem Waschen (10×) wurden 100 µl einer 1:10000 Verdünnung des Streptavidin-HRP-Konjugats (Roche, Mannheim) zugefügt. Nach einer Inkubation von 30 min bei RT wurden die Mikrotiterplatten abermals gewaschen (10×). Die Antikörper-Antigen-Komplexe wurden unter Verwendung von je 100 µl TMB-Substratlösung (BD, Heidelberg) nach Angaben des Herstellers detektiert. Die Adsorption wurde im Photometer bei einer Wellenlänge von 450 nm bestimmt. Die verwendeten p24-spezifischen Antikörper M01 und 37G12 (biotinyliert) sowie der p24-Standard wurden von PolyImm (Wien, Österreich) bezogen.

B.3.3 Analyse der transienten Proteinexpression im Western Blot

Transfizierte H1299-Zellen wurden zweimal mit eiskaltem PBS (10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, 137 mM NaCl, 2,7 mM KCl) gewaschen, in eiskaltem PBS abgekratzt, 10 min bei 300×g abzentrifugiert und in RIPA-Puffer (50 mM Tris-HCl, pH 8,0, 150 mM NaCl, 0,1% SDS (w/v), 1% Nonidet P-40 (w/v), 0,5% Natriumdesoxycholat (w/v)) nach Zusatz von Proteaseinhibitoren (*Complete Minis*, Roche, Mannheim) 20 min auf Eis lysiert. Unlösliche Bestandteile des Zelllysates wurden 30 min bei 10000×g und 4°C abzentrifugiert. Die Proben wurden mit einem Volumen 2× Probenpuffer (Laemmli, 1970) versetzt und 5 min auf 95°C erhitzt. 50 - 150 µg Gesamtprotein wurden über ein SDS-Polyacrylamidgel aufgetrennt (Laemmli, 1970), auf eine Nitrocellulose-Membran (Schleicher & Schuell, Dassel) transferiert und mit den entsprechenden Antikörpern analysiert. Dabei wurde für die Detektion von GFP (*green fluorescence protein*) ein polyklonaler A.v. *Peptide Antibody* (BD Biosciences Clontech, Heidelberg) aus Kaninchen in einer Verdünnung von 1:500 eingesetzt, der Nachweis des HIV-1 Kapsidantigens (p24) erfolgte unter Verwendung der monoklonalen Mausantikörper A13/5 und 1621 (Wolf *et al.*, 1990) ebenfalls in einer 1:500 Verdünnung. Als sekundärer Antikörper wurde jeweils ein AP-gekoppelter Anti-Kaninchen (Dako, Hamburg) bzw. Anti-Maus (BioRad, München) Antikörper (1:2000) verwendet. Als Proteinstandard wurde der Marker *Precision Plus Protein*TM der Firma BioRad (München) aufgetragen.

B.3.4 FACS-Analysen

Zum Nachweis des autofluoreszierenden (Exc₄₇₁/Em₅₀₃) 27-kDa Proteins GFP (*green fluorescence protein*) wurde mit den Zelllinien eine *fluorescence activated cell sorting* (FACS)-Analyse durchgeführt. Dazu wurden jeweils 5×10⁶ Zellen pelletiert (300 x g, 5 min) und in 1ml FACS-Puffer NaN₃ (PBS, 1% FKS, 1 mg/ml NaN₃) aufgenommen. Die Experimente wurden an einem FACS Calibur Gerät der Firma BD Bioscience Clontech (Heidelberg) durchgeführt, wobei innerhalb der ausgewählten Population 10 000 Zellen gemessen wurden. Die resultierenden Daten wurden mit Hilfe des Programmes WinMDI, Version 2.8 ausgewertet. Statistische Erhebungen wurden mit dem Programm *Excel 2000 (9.0.6926 SP3)* durchgeführt, und die erhaltenen Daten wurden graphisch unter Verwendung des Programms Graphprism dargestellt.

B.3.4.1 Lebendfärbung

Zunächst wurden 5×10⁶ Zellen pelletiert (300 x g, 5 min) und in 1ml FACS-Puffer (PBS, 1% FKS) aufgenommen. Der Ansatz wurde geteilt und 0,5 ml der Zellsuspension mit 1 ml Pri-Arbeitslösung (1 µg/ml Pri) versetzt. Nach erneuter Zentrifugation (300 x g, 2 min) wurde der Überstand abgesaugt, und die Propidiumiodid gefärbten Zellen in 0,5 ml FACS-Puffer resuspendiert. Als Vergleich dienten 0,5 ml der ungefärbten, im FACS-Puffer gelösten Zellen. In nachfolgenden FACS-Analysen konnte mit Hilfe Pri gefärbter toter Zellen die Population lebender Zellen bestimmt werden.

B.3.5 Fluoreszenzmikroskopie

Zum Nachweis des autofluoreszierenden (Exc₄₇₁/Em₅₀₃) 27-kDa Proteins GFP wurden die Zellen direkt auf sterile Deckgläschen in 6-Well-Schalen ausgesät. Nach 72h wurden sie mit dem Zellrasen nach unten luftblasenfrei mit 20 µl Eindeckelmedium (24% (w/v) Glycerin, 9,6% (w/v) Polyvinylalkohol, 2,5% (w/v) DABCO, 50 mM Tris-HCl, pH 8.0) auf Objekträger aufgelegt und getrocknet. Die Fluoreszenzmikroskopie wurde unter Verwendung der entsprechenden Anregungs- und Sperrfilter an einem *LEICA-DMR* Mikroskop (Leica Mikrosysteme Vertrieb GmbH, Bensheim) mit einer *Canon PowerShot S45* Kamera (Canon, Amsterdam) durchgeführt (Vergrößerung: 40-100fach). Dabei wurde eine Belichtungszeit von 0,1 sec und eine Blende von 8,0 gewählt.

B.4 RNA-Arbeitstechniken

Für alle Arbeiten mit RNA wurden gestopfte Spitzen sowie RNase-freie Reagenzien verwendet, Geräte und Arbeitsplatz wurden mit einer 1%-igen SDS-Lösung gesäubert. Zudem wurde ausschließlich mit Handschuhen gearbeitet.

B.4.1 Isolierung von RNA aus eukaryontischen Zellen

Stabil transfizierte Zellen wurden abtrypsinisiert und 2x mit eiskaltem PBS gewaschen. 1×10^7 Zellen wurden mit 175 µl Lyse-Puffer (50mM Tris, 140mM NaCl, 1.5mM MgCl₂, 0.5% NP-40, pH8.0) auf Eis 5 min partiell lysiert. Die cytoplasmatische Fraktion wurden von der Kernfraktion durch Zentrifugation abgetrennt (300 x g, 2 min) und auf Eis gestellt. Die Kerne wurden mit Lyse Puffer vorsichtig gewaschen, erneut zentrifugiert (300 x g, 2 min) und mittels *QIAshredder* (Qiagen) Säulen homogenisiert. Aus den Kernen und der cytoplasmatischen Fraktion wurde jeweils die Gesamt-RNA unter Verwendung des *RNeasy-Kits* (Qiagen, Hilden) nach Herstellerangaben präpariert. Alle RNA-Präparationen wurden in RNase-freiem Wasser aufgenommen und bei -80°C bis zur weiteren Verwendung aufbewahrt.

B.4.2 Photometrische Konzentrationsbestimmung der RNA

Die Konzentration und Reinheit der präparierten RNA wurde photometrisch mit Hilfe eines Zweistrahl-Spektrophotometers (Varian) in Quarzküvetten bestimmt. Dazu wurde zunächst ein Aliquot der jeweiligen RNA-Probe 1:50 mit 0,1%-iger SDS-Lösung verdünnt. Anschließend wurde ein UV-Absorptionsspektrum im Wellenlängenbereich von 260 nm und 280 nm aufgenommen. Über den Extinktionswert bei 260 nm konnte nach dem Lambert-Beerschen Gesetz die RNA-Konzentration bestimmt werden ($1 A_{260} = 40 \mu\text{g ssRNA/ml}$). Über das Verhältnis A_{260}/A_{280} ließ sich die Reinheit bestimmen (reine RNA: $A_{260}/A_{280} \geq 2$).

B.4.3 Beseitigung von DNA-Kontaminationen in der RNA

Um in den aufgereinigten RNA-Proben enthaltene DNA-Kontaminationen zu entfernen, wurde die vermessene RNA einem DNase-Verdau unterzogen. Je 20 µg RNA wurden dazu mit 5 µl RNase-freier DNase (10 U/µl; Roche, Mannheim) in einem Gesamtvolumen von 50 µl 1h bei 37°C inkubiert. Anschließend wurde die DNase 5 min bei 95°C hitzeinaktiviert, oder die RNA unter Verwendung des *Rneasy-Mini-Kits* (*RNA-Cleanup* Protokoll) aufgereinigt. Je ein Aliquot der aufgereinigten RNA wurde spektrophotometrisch vermessen (siehe B.4.2) und zur Kontrolle auf ein 3%-iges Agarosegel aufgetragen. Zur Detektion verbliebener DNA-Kontaminationen diente eine PCR mit spezifischen Primern unter Standardbedingungen (siehe Tabelle B.1) wobei 1 µg der DNase-verdauten RNA bzw. 1 µg genomische DNA (als Positivkontrolle) als Matrize eingesetzt wurde. Wurden in der PCR spezifische Banden für die RNA-Proben detektiert, so erfolgte nach Aufreinigung der RNA ein neuerlicher DNase-Verdau. Bei Abwesenheit spezifischer Banden wurde die Reinheit und Konzentration der RNA erneut spektrophotometrisch bestimmt und die RNA für eine Reverse Transkription verwendet.

B.4.4 Reverse Transkription (RT)

Die Umschreibung der erhaltenen DNA-freien RNA in cDNA erfolgte unter Verwendung des *1st strand cDNA Synthesis Kit for RT-PCR (AMV⁺)* von Roche (Mannheim). Dazu wurde zunächst 1 µg DNase-verdauter RNA vorgelegt und mit RNase-freiem H₂O auf 8,2 µl aufgefüllt. Anschließend wurden 11,8 µl des vorbereiteten Master-Mixes (20 U *AMV Reverse Transcriptase*, 1x *Reaction Buffer*, 5 mM MgCl₂, 1 mM dNTP-Mix, *Random Primer P(dN)₆* 3,2 µg, 50 U RNase Inhibitor) zupipettiert. Während des ersten Inkubationsschrittes (25°C, 10 min) fand das Binden der Startermoleküle statt, im nachfolgenden Schritt (42°C, 1h) wurde die RNA sukzessive revers transkribiert, was in der Synthese einzelsträngiger cDNA resultierte. Abschließend erfolgte eine Enzyminaktivierung bei 99°C für 5 min, die gewonnene cDNA wurde bei -20 °C gelagert. Als Kontrollen wurden in der Reversen Transkription jeweils Ansätze ohne *AMV Reverse Transcriptase* bzw. ohne cDNA als Matrize mitgeführt.

B.4.5 Qualitative Untersuchung revers transkribierter genomischer RNA

Um spezifische Genprodukte in stabil transfizierten Zellen auf RNA-Ebene qualitativ nachzuweisen, wurde die über Reverse Transkription erhaltene cDNA in eine PCR mit spezifischen Primern eingesetzt. Die PCR-Analyse erfolgte dabei unter Standardbedingungen (siehe Tabelle B.1).

B.4.6 Quantifizierung von revers transkribierter RNA via realtime-PCR

Für die Durchführung der *realtime* PCR-Analysen wurden ausschließlich Geräte und Reagenzien der Firma Roche (Mannheim) verwendet. Die Quantifizierung revers transkribierter genomischer RNA via *realtime*-PCR erfolgte mit dem Gerät *Lightcycler*TM unter Verwendung des *Lightcycler FastStart DNA Master SYBR Green I kits*. Erhaltene Daten wurden mit dem Programm *Lightcycler Software 3.5* bzw. *Lightcycler Software 4.0* (Roche, Mannheim) ausgewertet. Statistische Erhebungen wurden in dem Programm *Excel 2000 (9.0.6926 SP3)* unternommen, die graphische Darstellung der erhaltenen Werte erfolgt unter Verwendung des Programms *Graphprism*.

B.4.6.1 Prinzip des *LightCyclers*TM

Das *LightCycler*TM System der Firma Roche ermöglicht eine relative und „absolute“ Quantifizierung transgenspezifischer Transkripte (Eleaume *et al.*, 2004). Dabei erfolgt die Quantifizierung über die Detektion von Fluoreszenzsignalen mit Hilfe des Fluoreszenzfarbstoffes SYBR Green I, der sich in die kleine Grube der DNA einlagert. Der dsDNA-bindende Farbstoff ist hoch sensitiv, kann schon wenige Kopien/Reaktion detektieren (Freeman *et al.*, 1999) und liefert ein stärkeres Signal als EtBr. In ungebundenem Zustand weist der Farbstoff nur schwache vernachlässigbare Fluoreszenz auf, während er gebunden an DNA eine starke Fluoreszenz einer Wellenlänge von 530 nm zeigt (<http://www.roche-applied-science.com>). Sobald die DNA-Menge während der *realtime*-PCR ansteigt, erhöht sich proportional auch die Fluoreszenz. Der *LightCycler*TM misst nach jedem Zyklus die Fluoreszenz der in den Kapillaren befindlichen Proben, wobei die Fluoreszenzdaten sich am Monitor verfolgen lassen. Quantitative Aussagen über die Menge des Ausgangstemplates sind jedoch nur in der logarithmisch-linearen Phase möglich, die mit einer Verdopplung der Ausgangsmoleküle einhergeht. Daher wird der erste Zyklus identifiziert, in dem das logarithmisch ansteigende Fluoreszenzsignal signifikant detektierbar ist. Dieser Wert wird als *crossing point number* angegeben und lässt auf die Ausgangskonzentration des zugehörigen Amplikons schließen. Da SYBER Green I auch an unspezifische Produkte wie Oligonukleotiddimere bindet, muss zusätzlich die spezifische Schmelztemperatur der einzelnen PCR-Produkte untersucht werden. Dazu wird am Ende jedes Laufs die Temperatur in der Kammer langsam angehoben, und die Fluoreszenz in den einzelnen Kapillaren kontinuierlich gemessen. Sobald die dsDNA mit ansteigender Temperatur zu denaturieren beginnt, wird SYBR Green I freigesetzt was in einer plötzlichen Abnahme der Fluoreszenz resultiert. Die Schmelztemperatur wird dann aus der ersten Ableitung dieser Daten gewonnen. Da jedes dsDNA-Produkt seinen spezifischen Schmelzpunkt besitzt, können durch die Schmelzpunktsanalyse spezifische und unspezifische PCR-Produkte, wie z.B. Primerdimere unterschieden werden.

B.4.6.2 Wahl der Parameter für *realtime*-PCR-Analysen

Um die Fluoreszenz der durch *realtime*-PCR entstehenden dsDNA-Produkte optimal zu detektieren, wurden Amplifikationsprimer (siehe F.1.1) gewählt, die 200- bis 400-bp-große DNA-Fragmente generierten. Weiterhin wiesen gewählte Primerpaare einen ähnlichen, zwischen 40 und 70% liegenden GC-Gehalt auf und besaßen eine Länge von 18 bis 24 Nukleotiden. Um die Bildung von Primer-dimeren zu verhindern, wurden sehr spezifische Primer verwendet und die *Annealing*-Temperatur entsprechend ihrer Schmelztemperatur gewählt. Die Elongationszeit errechnete sich aus der Länge des Amplikons laut: $t = (\text{bp}:25) \text{ sec}$.

B.4.6.3 Quantifizierung von cDNA via *LightCycler*TM-Analysen

Für alle *LightCycler*TM-Analysen wurde unter Verwendung des *Lightcycler FastStart DNA Master SYBR Green I kits* zunächst ein Mastermix laut Herstellerangaben vorbereitet (siehe Tabelle B.3), um Pipettierfehler zu minimieren. Ein *LightCycler*TM Rotor wurde mit der benötigten Anzahl an *Lightcycler Capillaries* (20 μl) bestückt und jede Kapillare mit 19 μl Mastermix gefüllt. Danach wurde jeder Kapillare 1 μl durch Reverse Transkription erhaltene cDNA zugegeben. Als Negativkontrolle diente eine Probe, die anstelle der cDNA 1 μl H₂O enthält. Nach Verschließen der Kapillaren und Zentrifugation (3000 rpm; 1 min) des LC-Rotors in einer *LC Carousel Centrifuge* wurde der jeweilige Lauf gestartet. Durch Blockierung der Taq-DNA-Polymerase mit spezifischen Antikörpern wurden bis zum *Hot-Start* unspezifische Primerreaktionen der Polymerase unterbunden (Kellogg *et al.*, 1994). Für alle *Lightcycler*-Analysen wurde nachfolgendes Protokoll verwendet:

Tabelle B.3: *LightCycler*TM-Protokoll

Komponenten	Volumen	PCR-Bedingungen:	Temperatur	Zeit
H ₂ O _{dest}	13,6 μl	1. Denaturierung	95°C	10 min
25 mM MgCl ₂	2,4 μl	2. Amplifizierung*	95°C	15 sec
5'Primer pf (100 μM)	0,5 μl		50°C	20 sec
3'Primer pb (100 μM)	0,5 μl		72°C	30 sec
<i>LC Fast Start Reaction Mix SYBR Green I</i>	2,0 μl	3. Schmelzkurve	95°C	60 sec
Revers transkribierte RNA (Matrize)	1,0 μl		40°C	60 sec
Endvolumen	20 μl	4. Kühlen	40°C	30 sec

* Schritt 2: 50 Zyklen

Zur weiteren Fehlerminimierung wurden automatische Pipetten (Eppendorf, Hamburg) verwendet sowie Messungen in Doppel- bzw. Dreifachansätzen durchgeführt.

B.4.6.3.1 Relative Quantifizierung von cDNA

Theoretisch erfolgt in der logarithmisch-linearen Phase der PCR in jedem Zyklus eine Verdopplung der DNA-Kopienzahl. Eine PCR-Effizienz von zwei wird jedoch nur im Idealfall erreicht. Um verschiedene cDNA-Proben gleicher bp-Länge mittels *realtime*-PCR hinsichtlich ihrer Kopienzahl vergleichen zu können, muss zunächst die für jedes PCR-Produkt spezifische Effizienz berechnet werden. Dazu wurden

verschiedene Verdünnungen der zu untersuchenden cDNA-Probe hergestellt und als Doppelansätze in *Lightcycler*-Analysen, wie unter B.4.6.3 beschrieben, eingesetzt. Die errechnete *crossing point number*, die Zahl der Zyklen nach der das Fluoreszenzsignal der jeweiligen Probe signifikant detektiert wird, wird nach dem Lauf logarithmisch gegen die Verdünnungsstufen aufgetragen. Aus der sich ergebenden Regressionsgerade ermittelt sich der sogenannte *slope*, die negative Steigung der Geraden. Unter Verwendung der negativen Steigung errechnet sich die PCR-Effizienz E wie folgt:

$$E = 10^{-1/\text{slope}}$$

Mit dieser Information können im selben Lauf amplifizierte cDNA-Proben unterschiedlicher Konzentration relativ zueinander verglichen werden, da die Amplifikation der PCR-Produkte in der linear-logarithmischen Phase exponentiellen Gesetzen folgt. Dabei fungiert die errechnete PCR-Effizienz E als Basis, die ermittelte *crossing point number* n wird exponential eingesetzt. N repräsentiert die Menge an dsDNA welche zu einer spezifischen Fluoreszenzdetektion durch den *Lightcycler* führt.

$$N = N_0 \times E^n$$

Da N, die Menge an dsDNA welche zu einer spezifischen Fluoreszenzdetektion durch den *Lightcycler* führt, für Proben verschiedener Ausgangskonzentration gleichgesetzt werden kann, lässt sich ein relativer Vergleich zweier unbekannter Proben wie folgt durchführen:

$$N_{0A}/N_{0B} = E_B^{nB}/E_A^{nA}$$

Dabei wurde N in drei unabhängigen Experimenten bestimmt und der Mittelwert in die Formel eingesetzt. Es wurden jeweils Dreifachansätze der Verdünnungsstufen 1:10, 1:100 und 1:1000 untersucht.

B.4.6.3.2 Etablierung eines externen Standards

Für die Herstellung reproduzierbarer Standardkurven wurde ein PCR-Produkt des gewünschten Amplikons spektrophotometrisch vermessen und nach Ermittlung des Molekulargewichts die Kopienzahl wie folgt berechnet (NEB Katalog 2005/06, S. 281):

$$\text{MW dsDNA} = (\text{Zahl der bp dsDNA}) \times (650/\text{g/mol})$$

Unter Verwendung der Avogadro'schen Konstante $A = 6,023 \times 10^{23}$ Teilchen/mol ergibt sich die Zahl der Kopien/ μl :

$$\text{Kopienzahl}/\mu\text{l} = [(6,023 \times 10^{23})/\text{MW dsDNA}] \times \text{Konz. des PCR-Produktes pro } \mu\text{l}$$

Anschließend wurden verschiedene Verdünnungen des PCR-Produktes mit bekannter Kopienkonzentration (10^9 Kopien/ μl – 10^2 Kopien/ μl) angelegt und in Doppelansätzen im *LightCycler*TM amplifiziert, wie unter B.4.6.3 beschrieben. Mit Hilfe der LC-Software wurden die erhaltenen *crossing point numbers* (Zykluszahlen) gegen die logarithmische Konzentration (Kopienzahl/ μl) aufgetragen. Aus der sich ergebenden Standardgerade ermittelt sich der sogenannte *slope*, die negative Steigung der Geraden. Daraus kann unter Verwendung der negativen Steigung, wie unter B.6.3.2 beschrieben, die PCR-Effizienz *E* berechnet werden. Ein Maß für die Güte der Standardgeraden repräsentiert die Abweichung von der Geradenfunktion *error*, deren Wert für eine lineare Abhängigkeit der einzelnen Verdünnungen 0,2 nicht überschreiten soll. Bei Vorliegen einer linearen Abhängigkeit kann die Kopienzahl unbekannter cDNA-Proben pro μl unter Verwendung der Standardkurve extrapoliert werden. Nach Ermittlung der Standardgeraden wurde diese extern separat gespeichert.

B.4.6.3.3 Absolute Quantifizierung von cDNA

Zur absoluten Quantifizierung revers transkribierter RNA werden ein Kalibrator, eine Probe mit definierter Kopienkonzentration, die auch zur Etablierung der Standardgeraden verwendet wurde, sowie die zu untersuchenden cDNA-Proben im selben Lauf im *Lightcycler* amplifiziert (siehe B.4.6.3). Nach Beendigung des Laufs kann mit Hilfe der extern gespeicherten Standardgerade (siehe B.4.6.3.2) der Kalibrator normalisiert werden. Die Ausgangskonzentration der unbekanntenen Proben wird schließlich durch Extrapolation aus der Standardkurve ermittelt. Als Negativkontrolle wurde stets eine Probe mit Wasser als Matrize mitgeführt. Auf diese Weise erfolgte eine Bestimmung der spezifischen Transgenkopienkonzentration pro μg eingesetzter Gesamt-RNA für cytoplasmatische und nukleäre RNA-Fractionen sowie für *de novo* synthetisierte RNA (siehe B.4.7). Auch die Ermittlung der jeweiligen Transgenkopienkonzentration revers transkribierter Gesamt-RNA nach *ActinomycinD*-Zugabe (siehe B.4.8) erfolgte mit dieser Methode.

B.4.7 „Nuclear Run On” Assay

Der „Nuclear Run On“-Assay wurde nach einem von Heap *et al.*, 2000 beschriebenen Verfahren unter Verwendung von Streptavidin gekoppelten *Dynabeads M-280* (DynaL, Oslo) durchgeführt. Zunächst wurden 3×10^7 stabil transfizierte CHO-Zellen geerntet und zweimal mit eiskaltem PBS gewaschen. Das resultierende Pellet wurde in 4ml Lysepuffer (0,5% NP-40, 10 mM Tris-HCl pH 7,4, 3 mM MgCl_2 , 10 mM NaCl und 150 mM Sucrose) resuspendiert und anschließend 5 min auf Eis inkubiert. Anschließend wurden die Kerne vom Überstand durch Zentrifugation bei 4°C abgetrennt ($170 \times g$, 2 min) und mit Lysepuffer ohne NP-40 gewaschen. Die erneut sedimentierten Kerne ($170 \times g$, 2 min) wurden in 100 μl Glycerinpuffer (50 mM Tris/HCl pH 8.3, 5 mM MgCl_2 , 0.1 mM EDTA, 40% Glycerin) aufgenommen und auf Eis gehalten. Weiter wurden die Kernpräparationen mit einem Volumen 2xTranskriptionspuffer (200 mM KCl, 20 mM Tris-HCl pH 8,0, 5mM MgCl_2 , 4mM DTT, je 4 mM ATP, GTP and CTP, 200 mM Sucrose and

20% Glycerin) versetzt und 8 µl biotin-16-UTPs eines 10 mM Tetralithiums Salzes (Roche; Mannheim) zu der Mischung gegeben. Nach 30-minütiger Inkubation des Reaktionsgemisches bei 29°C wurden 6 µl 250 mM CaCl₂ sowie 6 µl RNase-freier DNase I (10 U/µl; Roche, Mannheim) zugegeben (10 min, 29°C), um die Reaktion zu stoppen. Die folgende Isolation der Gesamt-RNA erfolgte mittels des *RNeasy-Kits* (Qiagen, Hilden) nach Herstellerangaben. Die erhaltene RNA wurde in 50 µl RNase-freiem H₂O aufgenommen. 50 µl der „Run on“-RNA wurden mit dem gleichen Volumen in Bindepuffer (10 mM Tris-HCl pH 7,5, 1 mM EDTA and 2 M NaCl) gelöster *Dynabeads M-280* (Dyna, Oslo) versetzt und 20 min bei 42°C sowie 2h bei Raumtemperatur inkubiert. Anschließend wurden die Dynabeads mittels der magnetischen Apparatur (Dyna, Oslo) separiert und zweimal 15 min mit 500 µl 15% Formamid und 2xSSC gewaschen. Nach einem weiteren Waschschriff mit 1 ml 2xSSC (5 min) wurden die Dynabeads in RNase-freiem H₂O gelöst. Je 3 µl der so erhaltenen RNA wurden nach DNase-Behandlung (siehe B.4.3) als Matrize in eine Reverse Transkription (siehe B.4.4) eingesetzt und die resultierende cDNA in *Lightcycler*-Analysen absolut quantifiziert (siehe B.4.6.3.3).

B.4.8 Untersuchung der RNA-Stabilität mittels *ActinomycinD*

Die Untersuchung der RNA-Halbwertszeit erfolgte unter Verwendung des Transkriptioninhibitors *ActinomycinD* (Biocat, Heidelberg) nach einer Methode von Leclerc *et al.*, 2002. 72h vor Beginn des Experimentes wurden stabil transfierte Zellen unterschiedlicher Konzentrationen in 6-Well-Schalen ausgesät. Dabei wurden die Konzentrationen so gewählt, dass die Zellen zum Zeitpunkt der *ActinomycinD*-Gabe 90% Konfluenz aufwiesen. Für eine *ActinomycinD*-Behandlung von 24h bzw. 12h wurden je 6x10⁵ Zellen pro Schale ausgesät, für eine Inhibition der RNA-Synthese von 6h, 3h, 1,5h und 0h wurden je 3,5x10⁵ Zellen in eine Schale vorgelegt. 24h, 12h, 6h, 3h, 1,5h und 0h vor Ernte der Zellen wurde dem jeweils eine Stunde zuvor gewechselten Medium 2,4µM *ActinomycinD* zugesetzt. Daraus resultierte eine Inhibition der RNA-Synthese über einen unterschiedlichen Zeitraum hinweg. Zum Zeitpunkt 0 wurden die verschiedenen Ansätze mit PBS gewaschen und direkt in der 6-Well Platte lysiert. Anschließend wurde für alle Proben die Gesamt-RNA unter Verwendung des *RNeasy-Kit* (Qiagen, Hilden) nach Herstellerangaben präpariert. Je Ansatz wurde 1 µg der erhaltenen RNA nach DNase-Behandlung (siehe B.4.3) als Matrize in eine Reverse Transkription (siehe B.4.4) eingesetzt und die resultierende cDNA in *Lightcycler*-Analysen absolut quantifiziert (siehe B.4.6.3.3). Bei exponentieller Auftragung der für die verschiedenen Zeitpunkte erhaltenen RNA-Kopienzahlen gegen die jeweiligen Zeitpunkte kann die Zerfallskonstante *k* bestimmt werden, welche für die Ermittlung der Halbwertszeit essentiell ist. Dabei berechnet sich die Halbwertszeit einer RNA wie folgt:

$$t_{1/2} = -\ln 2/k$$

Zur Ermittlung der RNA-Halbwertszeit für die jeweiligen Transgene wurden zwei unabhängige Experimente durchgeführt, die Halbwertszeit des Haushaltsgens wurde in einem Experiment bestimmt.

C Ergebnisse

C.1 Gesamtkonzept

C.1.1 Anforderungen an den Vektor

Da CpG-Motive (siehe A.2) einen erheblichen Einfluss auf Komponenten des humoralen und zellulären Immunsystems ausüben (Krieg *et al.*, 1995; Hartmann *et al.*, 2000; Bauer *et al.*, 2001) bzw. CpG-Dinukleotide (siehe A.3) an der Regulation der Genexpression beteiligt sind (Hsieh, 1994; Kudo, 1998), war es Ziel dieser Arbeit, ein CpG-minimiertes Vektorrückgrat zu etablieren. Eine modulare Komposition des Vektors soll es ermöglichen, gezielt CpG-Motive zu integrieren und deren immunmodulierende Eigenschaften zu untersuchen. Ein modularer Aufbau ermöglicht darüber hinaus die Eliminierung bestimmter Vektorelemente wie z.B. Antibiotikaresistenzkassetten, die für den jeweiligen Verwendungszweck nicht relevant sind. Auch ein Austausch einzelner funktioneller Bereiche (z.B. Promotoren oder Polyadenylierungssignale) sowie die Kombination unterschiedlicher Komponenten soll gewährleistet sein. Einzelne Vektorelemente und ihre funktionellen Teilbereiche müssen dazu flankierende singuläre Restriktionsschnittstellen aufweisen. Die Verwendung gängiger Restriktionsenzyme soll eine einfache Handhabung und Herstellung der Vektorplattform ermöglichen. Weiterhin soll eine Modifizierung des Ausgangsvektors unter Beibehaltung der Transkriptionsfaktorbindestellen sowie der Kodonwahl in den kodierenden Regionen erfolgen. Auch eine Etablierung unerwünschter DNA-Motive (z.B. TATA-Boxen, Ribosomeneintrittsstellen oder Spleissstellen) soll vermieden werden, um die Funktionalität des Vektors nicht zu beeinträchtigen. Die generierte Vektorplattform soll schließlich sowohl für transiente als auch für stabile Expressionsstudien einsetzbar sein. Für die Überprüfung dieses Vektorkonzeptes muss auch das zu exprimierende Transgen definierte Eigenschaften erfüllen.

C.1.2 Wahl und Modifizierung der Transgene

Eine Methylierung von in der kodierenden Region lokalisierten CpG-Dinukleotiden (siehe A.3.4) führt laut Literatur (Hisano *et al.*, 2003; Chevalier-Mariette *et al.*, 2003) zu einer Verringerung bzw. dem Abschalten der Genexpression in eukaryontischen Zellen. Sowohl für eine DNA-Vakzinierung als auch für gentherapeutische Zwecke und die Erzeugung rekombinanter Wirkstoffe ist eine hohe und lang andauernde Expression der gewünschten Transgene jedoch Voraussetzung. Daher stellte sich zu Beginn dieser Arbeit die Frage, ob eine Eliminierung aller CpG-Dinukleotide aus dem offenen Leserahmen eine Expressionssteigerung bzw. –verlängerung zur Folge hätte und *vice versa*. Um den Einfluss von CpG-Dinukleotiden auf die Regulation der Genexpression untersuchen zu können, sollte ausgehend von einem Reporter gen unter Beibehaltung der Aminosäuresequenz je eine CpG-freie sowie eine CpG-maximierte Variante etabliert werden.

Als Reporter gen wurde zum einen das auf humane Kodonwahl optimierte *green-fluorescent-protein*-Gen (h_GFP 60) verwendet, dessen 740-bp-lange Sequenz (siehe F.2.1) 60 CpG-Dinukleotide aufweist (Zolotukhin *et al.*, 1996). Da Sequenzabschnitte dieser bp-Länge sich erfahrungsgemäß leicht klonieren lassen, und eine Eliminierung aller CpG-Dinukleotide aus der kodierenden Region möglich ist (siehe C.3.1.1), eignet sich der Reporter zur Untersuchung dieser Fragestellung. Durch Messen der GFP-vermittelten, grünen Autofluoreszenz in einem Durchflusszytometer (FACS-Analyse) kann sowohl die mittlere Fluoreszenzintensität der Zellen als Maß für die Transgenexpression als auch der Prozentsatz fluoreszierender und damit das Transgen exprimierender Zellen bestimmt werden. Verschiedene Reporter genvarianten können damit sowohl hinsichtlich ihrer Expressionshöhe als auch ihrer Transfektionseffizienz verglichen werden. Weiterhin lässt sich GFP über einen längeren Zeitraum stabil in eukaryontischen Zellen exprimieren, ohne toxische Effekte aufzuweisen (Liu *et al.*, 1999). So können auch Langzeitexpressionsstudien mit diesem Reporter gen durchgeführt werden. Zudem besteht die Möglichkeit, die Transgenexpression in zeitlich effektiven und preisgünstigen FACS-Analysen qualitativ und quantitativ nachzuweisen. Da es sich bei GFP um ein etabliertes Reporter gen handelt, ist eine breite Auswahl an funktionellen Antikörpern und Reagenzien käuflich verfügbar.

Als weiteres Reporter gen wurde das von HIV-1 abgeleitete und bereits früher auf humane Kodonwahl optimierte Kapsidproteing *p24* gewählt (Graf *et al.*, 2000), dessen Sequenz (h_p24 38) unter F.2.1 angegeben ist. Dabei handelt es sich um ein anwendungsbezogenes Transgen, da *p24* Teil einer bereits in der klinischen Phase befindlichen DNA-Vakzine ist (Deml *et al.*, 2001). Auch Handhabung und Expressionsnachweis dieses Reporters sind in der Arbeitsgruppe etabliert. Im *p24*-ELISA ist jedoch nur eine Bestimmung der Expressionshöhe und nicht der Transfektionseffizienz möglich.

C.2 Etablierung von CpG-Modifikationen im Vektorrückgrat

C.2.1 Design und Herstellung einzelner modularer Vektorkomponenten

Als Basis für die Herstellung einer modular aufgebauten Vektorplattform mit verringertem Anteil an CpG-Dinukleotiden wurde die Nukleinsäuresequenz des Plasmides *pcDNA5/FRT* von Invitrogen herangezogen. Basierend auf dem käuflichen Vektor *pcDNA5/FRT* (siehe C.2.2.2.1) wurden verschiedene Elemente definiert, die teilweise optimiert und synthetisch produziert oder PCR-basiert, hergestellt wurden. Die Resistenzkassetten sowie der multiple Klonierungsbereich, jeweils mit angrenzenden Sequenzen, wurden unter Verwendung der Software *Geneoptimizer*TM einer Multiparameter-Optimierung unterzogen und anschließend synthetisch hergestellt (siehe B.1.5). Funktionelle Bereiche des Vektors wie der *pUC origin of replication*, der eukaryontische CMV-Promotor und das BGH Poly-

adenylierungssignal zur Termination der Transkription wurden über PCR generiert. Im folgenden werden Aufbau und Herstellung der einzelnen Elemente näher erläutert.

C.2.1.1 Entwicklung von modularen CpG-reduzierten Vektorkomponenten

C.2.1.1.1 Prokaryontische Resistenzkassetten

1. Ampicillinresistenzkassette (Amp)

Basierend auf dem käuflichen Vektor pcDNA5/FRT (siehe C.2.2.2.1) wurden die Ampicillinresistenzkassette sowie angrenzende Füllsequenzen zu dem Vektorelement Amp zusammengefasst. Dabei enthält Amp neben der für das Ampicillinresistenzgen (*amp*) kodierenden Sequenz auch den *bla*-Promotor sowie nicht-kodierende Füllsequenzen. Im pcDNA5/FRT besitzt das Element 72 CpG-Dinukleotide und wird daher als Amp₇₂ bezeichnet. Eine Optimierung hinsichtlich verschiedener Parameter ergab das Element Amp₁, welches nur noch 1 CpG-Dinukleotid aufweist. Tabelle C.1 bietet eine Übersicht über alle Modifizierungen, die in die Sequenz von Amp₁ im Rahmen der Multi-Parameter-Optimierung (siehe B.1.5) eingeführt wurden. Nachfolgende Positionsangaben beziehen sich dabei auf den unter F.2.1 angegebenen Nukleinsäurevergleich der Vektorelemente Amp₇₂ aus pc5 und Amp₁.

Tabelle C.1: Vergleich der Elemente Amp₁ und Amp₇₂ hinsichtlich ihres CpG-Gehaltes und ihrer singulären Restriktionsschnittstellen.

<i>Ampicillinkassette</i>	<i>Amp_1</i> (Nr.: 02-774)	<i>Amp_72</i> (pc5)
Eingefügte singuläre RE-Schnittstellen	<i>ClaI</i> (5) <i>NsiI</i> (279) <i>ScaI</i> (384) <i>HpaI</i> (1250) <i>BglII</i> (1400)	-
Deletierte RE-Schnittstellen	-	<i>BglII</i> (225) <i>ScaI</i> (694)
CpG-Zahl <i>bla</i> -Promotor	1	1
CpG-Zahl <i>amp</i> (ORF)	0	49
CpG-Zahl <i>Füllsequenz</i>	1 (<i>ClaI</i>)	22
CpG-Zahl Gesamt	2	72

Für die beiden Konstrukte ist jeweils die Gesamt-CpG-Zahl angegeben. Zusätzlich wird der CpG-Gehalt der Promotorregion, des Leserahmens und angrenzender Füllsequenzen aufgeführt. Die Positionen der jeweiligen Restriktionsschnittstellen werden in Klammern angegeben und beziehen sich auf den unter F.2.1 angegebenen Nukleinsäuresequenzvergleich der beiden Amp-Konstrukte.

Aus klonierungstechnischen Gründen wurde das Fragment Amp₁ durch eine 5' *ClaI*-Schnittstelle (5) sowie eine 3' *BglII*-Schnittstelle (1400) ergänzt. Der *bla*-Promotor wird im synthetisch generierten Fragment von den singulären Schnittstellen *NsiI* (279) und *ScaI* (384) flankiert, und die kodierende Region (*amp*) kann unter Verwendung der singulären Enzyme *ScaI* (384) und *HpaI* (1250) subkloniert werden. Eine Deletion der internen *BglII*- (225) sowie *ScaI*- (694) Schnittstellen ermöglichte eine singuläre Verwendung dieser Enzyme im Element Amp₁. Zudem konnten durch Optimierung im Leserahmen von *amp* sowie in den angrenzenden Füllsequenzen alle CpG-Dinukleotide entfernt werden.

Ausschließlich die Sequenz des *bla*-Promotors, die 1 CpG enthält, wurde nicht verändert, um seine Funktionalität nicht zu beeinträchtigen. Ein weiteres CpG-Dinukleotid ist in der flankierenden Schnittstelle *ClaI* (ATCGAT) enthalten. Der modulare Aufbau der Vektorplattform ermöglicht jedoch später einen problemlosen Austausch dieser Schnittstelle.

Tabelle C.2: Die kodierenden Regionen *amp_0* und *amp_49*. Für den jeweiligen Leserahmen wird die Anzahl an CpG- sowie TpA-Dinukleotiden, der prozentuale Anteil der Nukleotide Cytidin und Guanosin an der Gesamtsequenz (GC-Gehalt) sowie der *codon adaption index* (CAI) angegeben. Zudem werden flankierende Klonierungsschnittstellen erwähnt.

<i>amp(ORF)</i>	<i>amp_0</i>	<i>amp_49</i>
CpG-Anzahl	0	49
CAI	0,55	0,61
GC-Gehalt	44 %	49 %
TpA-Anzahl	42	44
Klonierung	<i>ScaI/HpaI</i>	<i>ScaI/HpaI</i>

Parameter, die durch Optimierung des *amp*-Leserahmens verändert wurden, sind in Tabelle C.2 zusammengefasst. Die CpG-Zahl innerhalb der kodierenden Region konnte von 49 (*amp_49*) auf 0 (*amp_0*) reduziert werden. Dabei erfolgte die Eliminierung der CpGs unter Verwendung der im Gen vorhandenen Kodonhäufigkeit. Da im allgemeinen Transgenexpression und Kodonwahl positiv korrelieren (Zhou *et al.*, 2004; Graf *et al.*, 2004), eine verbesserte Expression des Ampicillinresistenzgens jedoch zu einer Verringerung der produzierten Plasmidkopienzahl und damit DNA-Ausbeute führen könnte, wurden zudem einige „schlechtere“ Kodons innerhalb des *amp*-Leserahmens eingefügt. So beträgt der *codon adaption index* (CAI), ein Maß für die Güte der Kodonwahl, für *amp_49* aus pc5 0,61, während die optimierte Version *amp_0* einen CAI von 0,55 aufweist. Beide Versionen des Ampicillinresistenzgens sind jedoch hinsichtlich ihres GC-Gehaltes (*amp_49*: 49% vs. *amp_0*: 44%) sowie ihrer TpA-Anzahl (*amp_49*: 44 vs. *amp_0*: 42) vergleichbar und können *ScaI/HpaI* subkloniert werden. Ein AS-Sequenzvergleich der beiden kodierenden Bereiche ist unter Kapitel F.2.2 aufgeführt. Die entworfene Sequenz für *Amp_1* wurde anschließend synthetisch hergestellt (Auftragsnummer: 02-774) und *KpnI/SacI* in den Vektor PCR-Script kloniert.

2. Kanamycinresistenzgen (Kana)

Basierend auf der Kanamycinresistenz-vermittelnden Sequenz (*kan*) aus dem Vektor pet28 (Novagen, Madison, Wisconsin), die 42 CpG-Dinukleotide enthält, wurde das Element *kan_0* entwickelt, welches kein CpG-Dinukleotid mehr aufweist. Eine Übersicht über alle Modifizierungen, die im Rahmen der Multi-Parameter-Optimierung (siehe B.1.5) in den Leserahmen von *kan* eingeführt wurden, bietet Tabelle C.3. Nachfolgende Positionsangaben beziehen sich dabei auf den unter F.2.1 angegebenen Nukleinsäurevergleich der Kanamycinresistenzgene *kan_42* aus pET28 und *kan_0*.

Wie aus Tabelle C.3 hervorgeht, wurde die CpG-Zahl innerhalb der kodierenden Region von 42 auf 0 (*kan_0*) reduziert. Dabei erfolgte die Eliminierung der CpGs

unter Verwendung der im *kan*-Gen vorhandenen Kodonhäufigkeit. Der CAI-Wert wurde durch die Depletierung der CpGs nur geringfügig verändert (*kan_42*: 0,64 vs. *kan_0*: 0,61). Hinsichtlich des GC-Gehaltes, des Anteils der Nukleotide Guanosin und Cytidin an der Gesamtsequenz, (*kan_42*: 43% vs. *kan_0*: 40%) sowie ihrer TpA-Anzahl (*kan_41*: vs. *kan_0*: 41) sind die beiden Versionen des *kan*-Gens vergleichbar und können *ScaI/HpaI* subkloniert werden. Ein AS-Sequenzvergleich beider Konstrukte ist unter Kapitel F.2.2 aufgeführt.

<i>kan</i> (ORF)	<i>kan_0</i>	<i>kan_42</i>
CpG-Anzahl	0	42
CAI	0,61	0,64
GC-Gehalt	40 %	43 %
TpA-Anzahl	41	41
Klonierung	<i>ScaI/HpaI</i>	<i>ScaI/HpaI</i>

Tabelle C.3: Die kodierenden Regionen *kan_0* und *kan_42*. Für den jeweiligen Leserahmen wird die Anzahl an CpG- sowie TpA-Dinukleotiden, der prozentuale Anteil der Nukleotide Guanosin und Cytidin an der Gesamtsequenz (GC-Gehalt) sowie der *codon adaption index* (CAI) angegeben. Weiter werden flankierende Klonierungsschnittstellen aufgeführt.

Die entworfene Sequenz für *kan_0* wurde anschließend synthetisch hergestellt (Auftragsnummer: 0401196) und *ScaI/HpaI* in das Fragment Amp_1 kloniert. Die so erhaltene Kanamycinresistenzkassette wird als Kana_1 bezeichnet, der funktionelle sowie der kodierende Bereich werden analog wie in Tabelle C.1 für Amp_1 beschrieben von singulären Schnittstellen flankiert.

C.2.1.1.2 Eukaryontische Selektionskassette

Die Hygromycinresistenzkassette sowie angrenzende, nicht-kodierende Sequenzen aus dem Vektor pcDNA5/FRT (siehe C.2.2.2.1) wurden zu dem Vektorelement Hygr zusammengefasst. Somit setzt sich das Element Hygr aus der kodierenden, Resistenz-vermittelnden Sequenz *hygr*, dem BGH-Polyadenylierungssignal, einem FRT-Signal zur Vermittlung homologer Rekombination (siehe B.1.3.1) und angrenzenden Füllsequenzen zusammen. Im pcDNA5/FRT besitzt das Element Hygr 174 CpG-Dinukleotide und wird daher als Hygr_174 bezeichnet. Eine Optimierung hinsichtlich verschiedener Parameter ergab das Element Hygr_1, welches nur noch 1 CpG-Dinukleotid aufweist. Tabelle C.4 bietet eine Übersicht über alle Modifizierungen, die in die Sequenz von Hygr_1 im Rahmen der Multi-Parameter-Optimierung (siehe B.1.5) eingeführt wurden. Nachfolgende Positionsangaben beziehen sich dabei auf den unter F.2.1 angegebenen Nukleinsäurevergleich der Vektorelemente Hygr_174 aus pc5 und dem optimierten Fragment Hygr_1.

Zur Vereinfachung nachfolgender Klonierungen wurde das Fragment Hygr_1 durch eine 5' *XmaI*-Schnittstelle (5) sowie eine 3' *EcoRI*-Schnittstelle (2046) ergänzt. Zudem wurde die kodierende Region (*hygr*) im synthetisch generierten Fragment von den singulären Schnittstellen für *AflI* (348) und *SwaI* (1377) flankiert. Das SV40-Polyadenylierungssignal konnte unter Verwendung der singulären Enzyme *PshAI* (1514) und *PacI* (1655) kloniert werden. Eine Deletion

der internen *EcoRI*-Schnittstelle (594) ermöglichte dabei die singuläre Verwendung dieses Enzyms im Element *Amp_1*. Weiter konnten durch Optimierung im Leserahmen von *hygr* sowie in den angrenzenden Füllsequenzen alle CpG-Dinukleotide entfernt werden. Ausschließlich die Sequenz des FRT-Signals, die 1 CpG enthält, wurde nicht verändert, um ihre Funktionalität aufrechtzuerhalten. Ein weiteres CpG-Dinukleotid ist in der flankierenden Schnittstelle *XmaI* (CCCGGG) enthalten. Ein modularer Aufbau der Vektorplattform ermöglicht jedoch später einen problemlosen Austausch dieser Schnittstelle.

Tabelle C.4: Vergleich der Elemente Hygr_1 und Hygr_174 hinsichtlich CpG-Gehalt und singulärer Restriktions-

<i>Hygromycinkassette</i>	<i>Hygr_1</i> (Nr.: 02-772)	<i>Hygr_174</i> (pc5)
Eingefügte singuläre RE-Schnittstellen	<i>XmaI</i> (5) <i>AflII</i> (348) <i>SwaI</i> (1377) <i>PshAI</i> (1514) <i>PacI</i> (1655) <i>EcoRI</i> (2046)	-
Deletierte RE-Schnittstellen	-	<i>EcoRI</i> (594)
CpG-Zahl <i>FRT-site</i>	1	1
CpG-Zahl <i>hygr</i> (ORF)	0	103
CpG-Zahl <i>SV40 P(A)</i>	0	0
CpG-Zahl <i>Füllsequenz</i>	1 (<i>XmaI</i>)	70
CpG-Zahl Gesamt	2	174

Restriktions-schnittstellen. Für die beiden Konstrukte ist jeweils die Gesamtzahl an CpG-Dinukleotiden sowie der CpG-Gehalt des FRT-Signals, des Leserahmens, des SV40 Polyadenylierungssignals und angrenzender Füllsequenzen aufgeführt. Die Positionen der jeweiligen Restriktionsschnittstellen werden in Klammern angegeben und beziehen sich auf den unter F.2.1 angegebenen Sequenzvergleich der beiden Hygr-Konstrukte.

Parameter, die durch Optimierung des *hygr*-Leserahmens verändert wurden, sind für *hygr_103* aus pc5 und *hygr_0* in Tabelle C.5 aufgelistet. Die CpG-Zahl innerhalb der kodierenden Region konnte von 103 auf 0 reduziert werden. Dabei erfolgte die Eliminierung der CpGs unter Verwendung der im Hygromycin-resistenzgen vorhandenen Kodonhäufigkeit. Der CAI wurde durch Depletierung der CpGs geringfügig verbessert, was zur Steigerung der Transgenexpression, und damit zu einer Vereinfachung der Selektion, beitragen sollte (*hygr_103*: 0,70 vs. *hygr_0*: 0,76). Während das Gen im pc5 einen GC-Gehalt von 58% aufweist, besitzt die optimierte Variante einen GC-Gehalt von 48%. Auch hinsichtlich ihres TpA-Gehaltes unterscheiden sich die beiden Versionen von *hygr* (*hygr_103*: 25 vs. *hygr_0*: 32). Eine Subklonierung des Leserahmens ist unter Verwendung der Enzyme *AflII* und *SwaI* möglich. Die entworfene Sequenz für *Hygr_1* wurde anschließend synthetisch hergestellt (Auftragsnummer: 02-772) und *KpnI/SacI* in den Vektor PCR-Script kloniert.

<i>hygr(ORE)</i>	<i>hygr_0</i>	<i>hygr_103</i>
CpG-Anzahl	0	103
CAI	0,76	0,70
GC-Gehalt	48 %	58 %
TpA-Anzahl	32	25
Klonierung	<i>AflIII/SwaI</i>	<i>AflIII/SwaI</i>

Tabelle C.5: Die kodierenden Regionen *hygr_0* und *hygr_103*. Für den jeweiligen Leserahmen wird die Anzahl an CpG- sowie TpA-Dinukleotiden, der prozentuale Anteil der Nukleotide Guanosin und Cytidin an der Gesamtsequenz (GC-Gehalt) sowie der *codon adaption index* (CAI) angegeben. Weiter werden flankierende Klonierungsschnittstellen aufgeführt.

C.2.1.1.3 Multipler Klonierungsbereich (MCS)

Gemäß der Sequenz im pc5 setzt sich der multiple Klonierungsbereich (MCS) aus dem T7-Promotor, der anschließenden Akkumulation von singulären Schnittstellen (*Polylinker*) und den angrenzenden Füllsequenzen zusammen. In dieser 208-bp-langen Sequenz befinden sich 11 CpGs (*MCS_11*). Nach Optimierung und Verkürzung des Elements auf 143 bp konnte die CpG-Zahl für das Element *MCS_1* auf 1 reduziert werden. Tabelle C.6 bietet eine Übersicht über alle Modifizierungen, die in die Sequenz von *MCS_1* im Rahmen der Multi-Parameter-Optimierung (siehe B.1.5) eingeführt wurden. Nachfolgende Positionsangaben beziehen sich dabei auf den unter F.2.1 angegebenen Nukleinsäurevergleich der multiplen Klonierungsbereiche.

Tabelle C.6: Vergleich der multiplen Klonierungsbereiche *MCS_1* und *MCS_11*. Angegeben sind jeweils der CpG- sowie TpA-Gehalt, singuläre Restriktionsschnittstellen und der prozentuale Anteil der Nukleotide Cytidin und Guanosin an der Gesamtsequenz (GC-Gehalt). Zusätzlich wird der CpG-Gehalt des Promotorbereichs, des eigentlichen Klonierungsbereichs (*Polylinker*) sowie angrenzender Füllsequenzen (Rest) aufgeführt. Die Positionen der jeweiligen Schnittstellen werden in Klammern angegeben und beziehen sich auf den unter F.2.1 angegebenen Nukleinsäuresequenzvergleich der beiden MCS-Konstrukte.

<i>MCS</i>	<i>MCS_0 (02-773)</i>	<i>MCS_11 (pc5)</i>
Eingefügte singuläre RE-Schnittstellen	<i>SacI</i> (5) <i>PmeI</i> (219)	-
Deletierte RE-Schnittstellen	-	<i>PmeI</i> (92) <i>AflIII</i> (99) <i>KpnI</i> (108) <i>SacI</i> (114) <i>EcoRI</i> (143)
CpG-Zahl <i>T7 Prom.</i>	1	1
CpG-Zahl <i>Polylinker</i>	0	9
CpG-Zahl Rest	0	1
CpG-Anzahl	0	11
GC-Gehalt	44 %	51 %
TA-Gehalt	15	16

Das verbleibende CpG im Element *MCS_1* befindet sich im funktionellen Bereich des T7-Promotors, dessen Sequenz unverändert blieb. Mehrere im multiplen Klonierungsbereich *MCS_11* befindliche Restriktionsenzym-schnittstellen wurden deletiert (*PmeI* 92, *AflIII* 99, *KpnI* 108, *SacI* 114, *EcoRI* 143), um diese für weitere Vektorelemente als flankierende singuläre Schnittstellen einsetzen zu können. So wurde zur Klonierung des *MCS_1*-Elements eine 5' *SacI*-Schnittstelle (5) sowie eine 3' *PmeI*-Schnittstelle (219) angefügt, die Restriktionsschnittstellen *HindIII* (102), *EcoRV* (112) und *BamHI* (120) wurden als singuläre Schnittstellen zur Klonierung von Transgenen beibehalten. Hinsichtlich des GC-Gehaltes, des

Anteils der Nukleotide Guanosin und Cytidin an der Gesamtsequenz, (*MCS_11*: 51% vs. *MCS_1*: 44%) sowie ihrer TpA-Anzahl (*MCS_11*: 16 vs. *MCS_1*: 15) sind die beiden Versionen des multiplen Klonierungsbereichs vergleichbar. Beide Versionen des Elements können mittels *SacI/PmeI* subkloniert werden. Die entworfene Sequenz für *MCS_1* wurde anschließend synthetisch hergestellt (Auftragsnummer: 02-773) und *KpnI/SacI* in den Vektor PCR-Script kloniert.

C.2.1.1.4 CMV_TG-Promotor

Zur Etablierung des CMV_TG-Promotors (Yew *et al.*, 2002) wurden in der Sequenz des CMV-Promotors alle CpG-Dinukleotide durch die Basenabfolge „TpG“ ersetzt. Zudem wurde für nachfolgende Klonierungen 5´ eine *ClaI* sowie 3´ eine *SacI*-Schnittstelle angefügt (siehe F.3.1). Die so erhaltene Sequenz wurde anschließend synthetisch produziert (Auftragsnummer: 03-551) und in den Vektor PCR-Script kloniert. Zur Herstellung des Konstruktes SYN_CMV_TG wurde die 5´ *ClaI*-Schnittstelle unter Verwendung des Oligonukleotides CMV_Kpn_pf (siehe F.1.1) später gegen eine *KpnI*-Schnittstelle ausgetauscht.

C.2.1.1.5 CMV_DH-Promotorvarianten

Unter Verwendung degenerierter Oligos wurden im Rahmen einer Diplomarbeit (Carola Paintner, Regensburg) partiell CpG-freie CMV-Promotorchimären entwickelt und auf ihre Funktionalität hin getestet. Bei diesen Konstrukten wurde das vordere Drittel (A), der mittlere Teil (B) oder das endständige Drittel (C) der CMV-Promotorsequenz CpG-depletiert. Nach Eliminierung aller CpGs innerhalb des Teilfragments A wurden in FACS-Analysen zwei unterschiedlich stark GFP exprimierende Klone ermittelt. Nach Sequenzanalyse dieser Konstrukte zeigte sich, dass die untersuchten Promotoren A1BC und A3BC jedoch in ihrer Sequenz übereinstimmten. Somit konnte für den Teilbereich A nur eine funktionelle Promotorchimäre etabliert werden, welche im Folgenden als ABC bezeichnet wird. Die Depletion der CpGs im mittleren (B-Bereich) bzw. endständigen (C-Bereich) Sequenzabschnitt führte dagegen zur Etablierung zweier (AB1C, AB2C) bzw. fünf (ABC3, ABC4, ABC5, ABC6, ABC7) funktioneller Promotorbereiche. Eine Übersicht über die Expressionsstärke der jeweiligen partiell CpG-freien CMV-Promotorchimären gibt Abbildung C.1.

Zur Konstruktion verschiedener vollständig CpG-depletierter CMV_DH-Promotorvarianten (D = nicht C, H = nicht G) wurden die CpG-freien Teilbereiche der in der Diplomarbeit entwickelten Konstrukte kombiniert. Dabei wurden nur CpG-freie Teilstücke der am höchsten exprimierenden Promotorkonstrukte verwendet (ABC, AB1C, AB2C sowie ABC3, ABC4, ABC7, siehe Abbildung C.1). In F-PCR-Experimenten (siehe B.1.4.3) wurden die jeweiligen Teilfragmente unter Verwendung spezifischer Primer zusammengefügt. Mittels der *Amplifikationsprimer* wurde zudem 5´ eine *ClaI*- sowie 3´ eine *SacI*-Schnittstelle eingefügt. Eine Übersicht über die für das jeweilige Konstrukt verwendeten Primer gibt Tabelle C.7, die Sequenzen der Oligonukleotide sind unter F.1.2 angegeben.

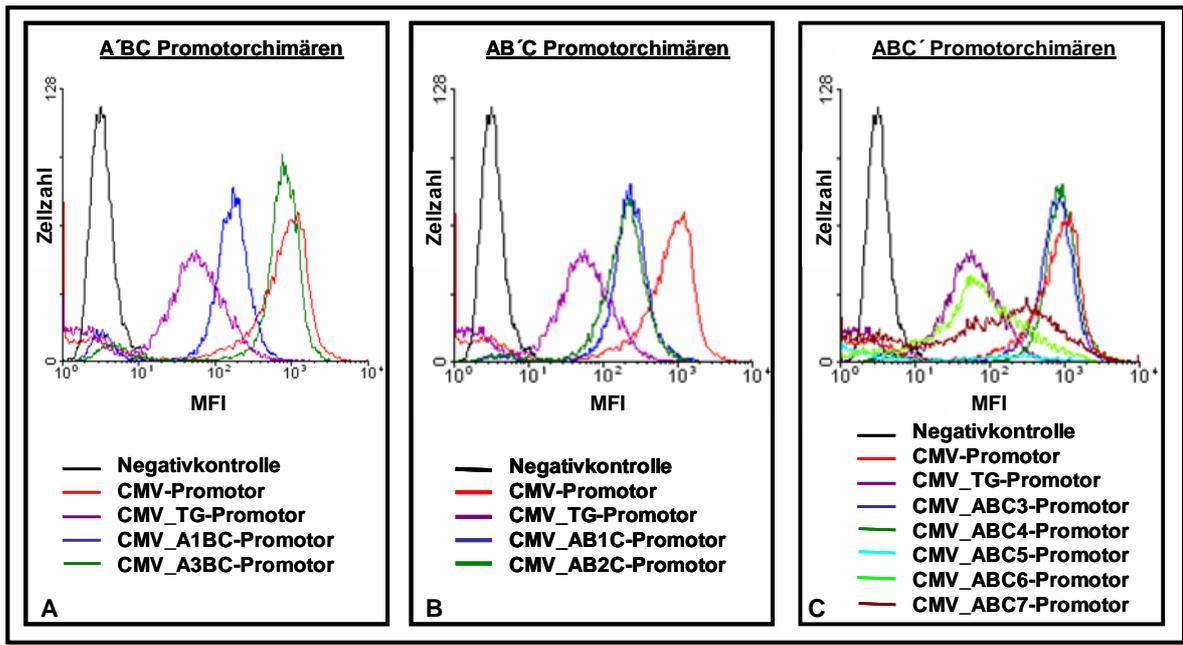


Abbildung C.1: GFP-Expression verschiedener CpG-freier rekombinanter CMV-Promotorchimären. Die partiell CpG-freien Promotorkonstrukte wurden in stabilen H1299-Zellen hinsichtlich ihrer h_GFP 60-Reporterexpression untersucht (FACS-Analyse). Der Abbildung sind die mittleren Fluoreszenzintensitäten (MFI) von GFP als Maß für die Promotorstärke (x-Achse) und die Zahl fluoreszierender Zellen (y-Achse) zu entnehmen. Das jeweilige CpG-freie Teilstück wird als Fragment A', B' bzw. C' bezeichnet. Die Zuordnung der einzelnen Konstrukte zu den jeweiligen Kurven kann der Legende entnommen werden. Dargestellt ist jeweils ein Vergleich verschiedener Promotorvarianten, bei denen das vordere Drittel der Sequenz (A), der mittlere Bereich (B) oder das 3' proximale Drittel der Sequenz (C) CpG-depletiert wurde. Als Negativkontrolle wurden *mock*-transfizierte Zellen eingesetzt. Der herkömmliche CMV-Promotor sowie eine in der Literatur beschriebene CpG-freie CMV_TG-Version dienen als Positivkontrollen.

Die sechs so erhaltenen CMV_DH-Promotorvarianten wurden wie in Tabelle C.7 beschrieben mit CMV_A bis CMV_F bezeichnet. Ein Nukleinsäurevergleich dieser Promotorkonstrukte ist unter F.2.1 angegeben. Um eine nachfolgende Klonierung der Konstrukte zu erleichtern, wurde zudem die 5' *Clal*-Schnittstelle unter Verwendung des Oligonukleotides CMV_Kpn_pf (siehe F.1.1) gegen eine *KpnI*-Schnittstelle ausgetauscht.

<i>CMV_DH</i> Promotorvarianten	Zur Fusion verwendete Primer (A' B')	Zur Fusion verwendete Primer (B' C')	Amplifikations- primer	Bezeichnung der Promotorvarianten
<i>AB1C3</i>	AB1_Fpf/ AB1_Fpb	B1C3_Fpf/ B1C3_Fpb	CMV_pf/CMV_pb	CMV_A
<i>AB1C4</i>	AB1_Fpf/ AB1_Fpb	B1C4_Fpf/ B1C4_Fpb	CMV_pf/CMV_pb	CMV_B
<i>AB1C7</i>	AB1_Fpf/ AB1_Fpb	B1C7_Fpf/ B1C7_Fpb	CMV_pf/CMV_pb	CMV_C
<i>AB2C3</i>	AB2_Fpf/ AB2_Fpb	B2C3_Fpf/ B2C3_Fpb	CMV_pf/CMV_pb	CMV_D
<i>AB2C4</i>	AB2_Fpf/ AB2_Fpb	B2C4_Fpf/ B2C4_Fpb	CMV_pf/CMV_pb	CMV_E
<i>AB2C7</i>	AB2_Fpf/ AB2_Fpb	B2C7_Fpf/ B2C7_Fpb	CMV_pf/CMV_pb	CMV_F

Tabelle C.7: Übersicht über die CpG-depletierten CMV-Promotorchimären. Aus den Teilfragmenten A, B1, B2, C3, C4 und C7 wurden sechs CpG-freie CMV_DH-Promotorvarianten kombiniert. Angegeben sind jeweils die zur Fusion der Teilbereiche A' und B' bzw. B' und C' verwendeten Fusionsprimer sowie die eingesetzten Amplifikationsprimer. Die resultierenden Konstrukte werden im Folgenden als CMV_A bis CMV_F bezeichnet.

C.2.1.1.6 BGH_WDW-Polyadenylierungssignal

Ausgehend von der Sequenz für das BGH-Polyadenylierungssignal im pc5 (A, Abbildung C.2), das 3 CpG-Dinukleotide besitzt, wurden mit Hilfe degenerierter Oligonukleotide mehrere CpG-freie BGH P(A)-Varianten generiert. Dabei wurde das 3' gelegene CpG-Dinukleotid durch den Einsatz von BGH_CWpb als rückwärtigen Amplifikationsprimer zu CW umgewandelt. Zusätzlich vermittelte der degenerierte vorwärtige Amplifikationsprimer BGH_DGintpf in der PCR einen Austausch des internen CpGs gegen die Basenabfolge DG. Das resultierende PCR-Produkt (B, Abbildung C.2) diente anschließend als Matrize für eine PCR mit den randomisierten Oligonukleotiden BGH_WGpf/ BGH_CWpb, wodurch das im 5' Bereich liegende CpG durch WG ersetzt wurde, und somit CpG-freie BGH P(A)-Varianten (C, Abbildung C.2) entstanden. Zusätzlich wurde anhand der Amplifikationsprimer 5' eine *PmeI* sowie 3' eine *XmaI*-Schnittstelle eingefügt. Anschließend wurden die CpG-depletierten Konstrukte *PmeI/XmaI* in den Vektor pΔS (siehe C.2.2.1.2) kloniert, und deren Sequenz anhand der Primer CMV_intL1pf und ori_mL1L2pb (siehe F.1.1) analysiert. Dabei konnten fünf verschiedene CpG-freie BGH P(A)-Varianten etabliert werden. Diese werden im Folgenden als BGH_WDW bezeichnet, je nach Austausch des jeweiligen CpGs durch die Basenabfolge WG, DG oder CW, beginnend am 5' Ende der Sequenz. Ein Nukleinsäurevergleich der erhaltenen CpG-freien BGH P(A)-Varianten BGH_AAA, BGH_AGA, BGH_AGT, BGH_ATA, BGH_ATT sowie der BGH P(A)-Ausgangssequenz ist unter F.2.1 gegeben.

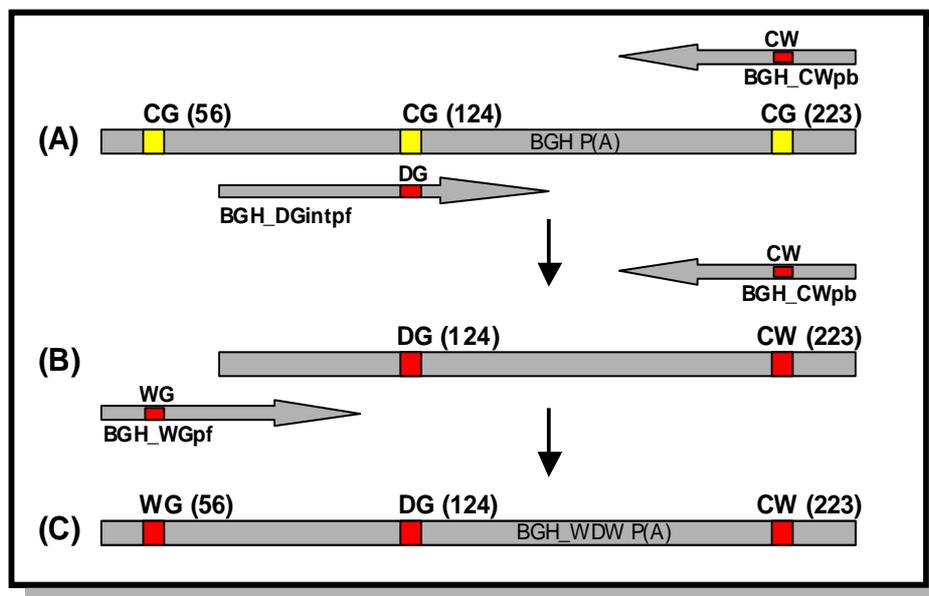


Abbildung C.2: Schematische Darstellung der Herstellung CpG-freier BGH Polyadenylierungssignale. Im BGH P(A)-Signal bestehende CpG-Dinukleotide sind gelb markiert, durch randomisierte Primer eingeführte, degenerierte Positionen sind rot angezeichnet. Angegebene Positionsangaben beziehen sich dabei auf den unter F.2.1 angegebenen Nukleinsäuresequenzvergleich der erhaltenen CpG-freien BGH_WDW P(A)-Varianten mit dem herkömmlichen BGH P(A)-Signal.

C.2.1.2 Etablierung von modularen PCR-basierten Komponenten

C.2.1.2.1 CMV-Promotor

Im Gegensatz zu den optimierten und synthetisch hergestellten Vektorkomponenten wurde die Sequenz des CMV-Promotors über PCR generiert. Dabei diente die CMV-Promotorsequenz aus pc5 als Matrize für die Reaktion. Zur weiteren Klonierung wurde mittels der verwendeten Amplifikationsprimer CMV_pf/CMV_pb (siehe F.1.1) 5´ eine *KpnI*- sowie eine *Clal*- und 3´ eine *SacI*-Schnittstelle eingefügt. Insgesamt besitzt die Sequenz des CMV-Promotors (siehe F.3) 31 CpG-Dinukleotide und ist 594-bp-lang.

C.2.1.2.2 BGH Polyadenylierungssignal

Um das BGH-Polyadenylierungssignal zu erhalten, wurde eine PCR mit den Oligonukleotiden BGH_pf/BGH_pb (siehe F.1.1) durchgeführt. Dabei diente das BGH P(A)-Signal aus dem Vektor pc5 als Matrize. Durch die Amplifikationsprimer wurde die Sequenz 5´ um eine singuläre *PmeI*-, sowie 3´ um eine singuläre *XmaI*- sowie eine *KpnI*-Schnittstelle verlängert. In der 235-bp-langen Sequenz (siehe F.3.1) befinden sich 3 CpG-Dinukleotide. Durch Anfügen einer *XmaI*-Schnittstelle mit der Erkennungssequenz CCCGGG wird ein weiteres CpG-Dinukleotid geschaffen. Der modulare Aufbau der Vektoren ermöglicht jedoch einen späteren Austausch bzw. eine Eliminierung der *XmaI*-Schnittstelle.

C.2.1.2.3 pUC *origin of replication*

Die Sequenz des pUC *origin of replication* wurde nicht verändert, um seine Funktionsfähigkeit aufrechtzuerhalten. Daher wurde das Element mit den Oligonukleotiden oriXE_pf/oriCB_pb (siehe F.1.1) und unter Verwendung des Vektors pc5 als Matrize über PCR hergestellt. Dabei fügten die Primer 5´ die Erkennungssequenz für die Enzyme *XmaI* und *EcoRI* sowie 3´ eine *Clal*- sowie eine *BglII*-Schnittstelle an (siehe F.3.1). Dies ermöglicht zum einen eine *XmaI/BglII*-Klonierung dieses Fragmentes in den Vektor pΔS (C.2.2.1.2), zum anderen den Einbau des *origins* in den Vektor SYN_84 (C.2.2.2.3) bzw. SYN_226 (C.2.2.2.5) unter Verwendung der Enzyme *EcoRI/BglII*. Das Element besitzt 45 CpG-Dinukleotide, durch die Einführung der *XmaI*-Schnittstelle wurde ein weiteres CpG generiert.

C.2.1.3 Zusammenfassung

Genoptimierung und synthetische Herstellung (siehe B.1.5) ermöglichten es, den CpG-Gehalt bestimmter Vektormodule erheblich zu reduzieren. So wurde die CpG-Zahl innerhalb des Amp-Fragments von 72 auf 1 CpG und im Leserahmen des Kanamycinresistenzgens von 42 auf 0 CpG vermindert (siehe C.2.1.1.1). Im Hygr-Element erfolgte eine Reduktion von 174 CpGs auf 1 CpG (siehe C.2.1.1.2), und der multiple Klonierungsbereich MCS (11 CpGs) wies nach Optimierung noch

1 CpG auf (siehe C.2.1.1.3). Zudem wurde eine in der Literatur (Yew *et al.*, 2002) beschriebene CpG-freie Version des CMV-Promotors (siehe C.2.1.1.4) künstlich produziert. Unter Verwendung randomisierter Oligonukleotide konnten auch eigene CpG-freie CMV-Promotorvarianten (siehe C.2.1.1.5) sowie BGH-Varianten (siehe C.2.1.1.6) entwickelt werden.

Im Vergleich dazu besitzen über PCR generierte Fragmente wie der CMV-Promotor mit 31CpG (siehe C.2.1.2.1), das BGH P(A)-Signal mit 3 CpG (C.2.1.2.2) und der pUC *origin of replication* mit 45 CpG (C.2.1.2.3) eine relativ hohe CpG-Dichte. Eine verkürzte Version des pUC *origin of replication* beinhaltet noch 39 CpGs (siehe C.2.2.1.3). Tabelle C.8 gibt eine Übersicht über verwendete Klonierungsschnittstellen, den CpG-Gehalt sowie damit verbundene murine bzw. humane CpG-Motive (siehe A.2) der einzelnen Vektormodule. Dabei werden murine CpG-Motive erwähnt, die der Basenpaarabfolge RRCpGYT entsprechen, sowie humane CpG-Motive, denen die Sequenz GTCGTT zu Grunde liegt.

<i>Elemente</i>	<i>Singuläre Schnittstellen</i>	<i>CpG Dinukleotide</i>	<i>Murine CpG-Motive</i>	<i>Humane CpG-Motive</i>
<i>CMV-Prom.</i>	<i>ClaI/SacI</i>	31	5	-
<i>CMV-Prom._TG</i>	<i>ClaI/SacI</i>	0	-	-
<i>CMV-Prom._A-F</i>	<i>ClaI/SacI</i> <i>KpnI/SacI</i>	0	-	-
<i>MCS_11</i>	<i>SacI/PmeI</i>	11	1	-
<i>MCS_1</i>	<i>SacI/PmeI</i>	1	-	-
<i>BGH P(A)</i>	<i>PmeI/XmaI</i>	3	-	-
<i>BGH_WDW P(A)</i>	<i>PmeI/XmaI</i>	0	-	-
<i>Hygr_174</i>	<i>XmaI/EcoRI</i>	174	7	-
<i>Hygr_1</i>	<i>XmaI/EcoRI</i>	1	-	-
<i>origin</i>	<i>EcoRI/BglIII</i> <i>XmaI/BglIII</i>	45	4	-
<i>110_origin</i>	<i>XmaI/BglIII</i>	39	4	-
<i>Amp_72</i>	<i>BglII/ClaI</i>	72	7	1
<i>Amp_1</i>	<i>BglII/ClaI</i>	1	-	-
<i>kan_42</i>	<i>ScaI/HpaI</i>	42	2	-
<i>kan_0</i>	<i>ScaI/HpaI</i>	0	-	-

Tabelle C.8: Übersicht der einzelnen Vektorelemente. Aufgeführt werden die flankierenden singulären Schnittstellen, der CpG-Gehalt sowie die Anzahl der humanen bzw. murinen CpG-Motive für das jeweilige Vektormodul.

C.2.2 Aufbau und Herstellung verwendeter Vektorkonstrukte

C.2.2.1 Basisvektoren zur transienten Expressionsanalyse

Im Folgenden werden Aufbau und Klonierung der für transiente Expressionsanalysen konstruierten Vektoren erläutert.

C.2.2.1.1 Referenzkonstrukt pRS

Als Grundlage für das Referenzkonstrukt pRS (siehe Abbildung C.3) diente die Sequenz des Vektors pc5 (siehe C.2.2.2.1). Da die eukaryontische Selektionsmarkerkassette Hygr (C.2.1.1.2) aus pc5 für transiente Studien nicht von Bedeutung ist, wurde diese in einer PCR mit dem Oligonukleotidpaar referenz_oripf/referenz_BGHpb (siehe F.1.1) sowie pc5 als Matrize eliminiert. Statt dessen wurde durch die Amplifikationsprimer zwischen BGH P(A)-Signal und pUC *origin of replication* eine singuläre *NsiI*- sowie *HpaI*-Schnittstelle eingefügt. Das PCR-Produkt wurde nach Spaltung mit *NsiI* religiert, und das resultierende Plasmid als pRS (3085 bp) bezeichnet. In Übereinstimmung mit dem pc5 (5070 bp) besitzt das Referenzkonstrukt neben dem CMV-Promotor, einem multiplen Klonierungsbereich, dem BGH Polyadenylierungssignal auch einen pUC *origin of replication* sowie eine Ampicillinresistenzkassette. Die eingefügten singulären Schnittstellen (*NsiI*; *HpaI*) erleichtern später eine Modifikation des Referenzkonstruktes pRS. Die Sequenz des Konstruktes ist im Anhang unter F.3.2 angegeben und weist 164 CpG-Dinukleotide auf.

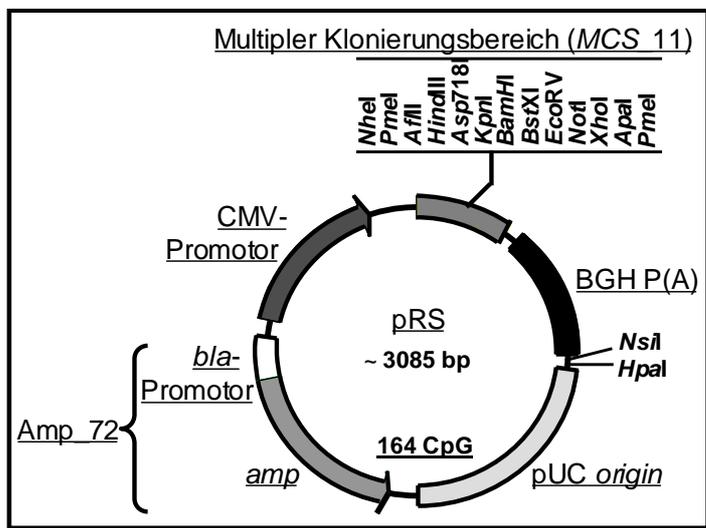


Abbildung C.3: Aufbau des für transiente Studien verwendeten Referenzkonstruktes pRS. Der Vektor enthält neben einer Transgenexpressionskassette aus CMV-Promotor, multiplen Klonierungsbereich (MCS_11) und BGH P(A)-Signal auch einen pUC *origin*, der eine Replikation in Bakterien ermöglicht. Das Ampicillinresistenzvermittelnde Gen (*amp*), dessen vorgeschaltener Promotor (*bla*-Promotor) und dessen angrenzende Füllsequenzen werden zu dem Element Amp_72 zusammengefasst. Das 3085-bp-große Plasmid enthält 164 CpG-Dinukleotide.

C.2.2.1.2 Basiskonstrukt pΔS

Zur Konstruktion des CpG-modifizierten Vektors pΔS wurden zunächst die Elemente CMV-Promotor (*KpnI/SacI*; C.2.1.2.1), *MCS_1* (*SacI/PmeI*; C.2.1.1.3) sowie das BGH P(A)-Signal (*PmeI/KpnI*; C.2.1.2.2) unter Verwendung geeigneter Schnittstellen in den Klonierungsvektor PCR-Script gesetzt. Das resultierende

Konstrukt (I, Abbildung C.4) enthält die Transgenexpressionskassette des p Δ S (960 pb) und wird als PSC_CMB bezeichnet. In einer anschließenden Dreifachligation wurden die *Clal/Xmal*-verdaute Transgenexpressionskassette, das *Clal/BglII*-gespaltene, CpG-reduzierte Amp_1-Element (II, Abbildung C.4) sowie der über PCR generierte, *BglII/Xmal*-geschnittene pUC *origin of replication* (III, Abbildung C.4) zu dem Vektor p Δ S (IV, Abbildung C.4) zusammengefügt. Der erhaltene Vektor wurde unter Verwendung der Primer CMV_pf, CMV_pb, BGH_pb, ori_mL1L2pb, ampA_m4 sowie ampC_L9 vollständig sequenziert. Insgesamt besitzt p Δ S 83 CpG-Dinukleotide, im Vergleich zu pRS konnte der CpG-Gehalt damit auf 51% reduziert werden.

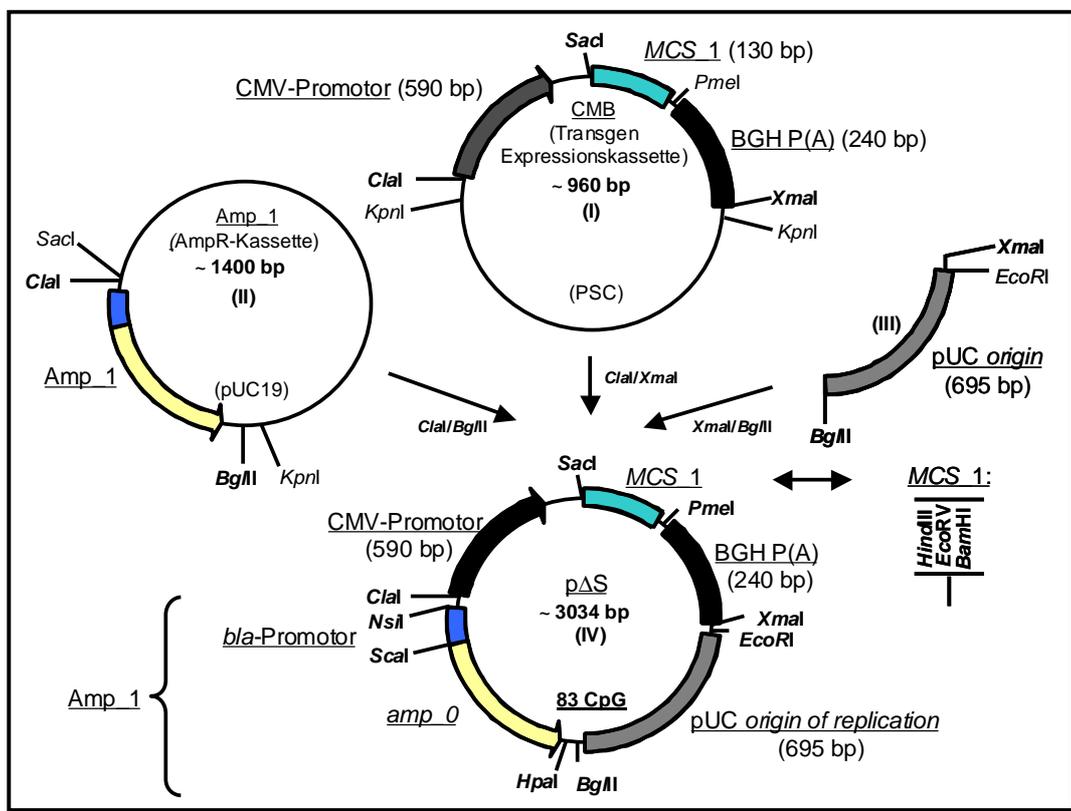


Abbildung C.4: Klonierungsschema zum Erhalt des CpG-reduzierten Basisvektors p Δ S. Synthetisch hergestellte, CpG-minimierte Elemente sind farbig dargestellt, während unmodifizierte, anhand von PCR generierte Komponenten grau bzw. schwarz unterlegt sind. Die einzelnen Vektor-komponenten werden von singulären Schnittstellen flankiert, innerhalb des Elementes Amp_1 sind zudem funktionelle Bereiche von singulären Schnittstellen umgeben. Im multiplen Klonierungsbereich (MCS) befinden sich Schnittstellen für *HindIII*, *EcoRV* sowie *BamHI*. Nähere Informationen zu Aufbau und Herstellung des CpG-reduzierten Basisvektors p Δ S können dem Text entnommen werden.

C.2.2.1.3 Modifizierte Basisvektoren (p Δ S110', p Δ SC, p Δ S kana, p Δ SB WDW und p Δ SC A-F)

Bei der Herstellung des Vektors p Δ S (C.2.2.1.2) trat nach der Dreifachligation eine Version des Vektors p Δ S auf, die eine 110-bp-lange Deletion am 5' Ende des pUC *origin of replication* aufwies. Zudem wurde ein Defekt in der am 5' Ende des pUC *origin of replication* lokalisierten *EcoRI*-Schnittstelle festgestellt. Unter Ver-

wendung der Oligonukleotide 110_X/E_pf und 110_X/E_pb (siehe F.1.1) wurde in einer PCR die zerstörte Schnittstelle wiederhergestellt, so dass das entstehende Konstrukt p Δ S110' (2924 bp; siehe F.3.2) bis auf eine 110-bp-lange Deletion im *origin* analog zum Vektor p Δ S (3034 bp; siehe F.3.2) aufgebaut ist. Wie im Verlauf der Arbeit gezeigt werden konnte, hat die Deletion keinen Einfluss auf die Funktionalität des Replikationsursprungs. Durch Entfernen der 110 Basenpaare wurden 6 CpG-Dinukleotide aus dem *pUC origin of replication* entfernt, die Sequenz der verkürzten Replikationskomponente 110'pUC *origin* (39 CpG) ist unter F.3.1 beschrieben. Im Vergleich zu pRS (164 CpG) konnte der CpG-Gehalt des Vektors p Δ S110' (77 CpG) damit um 53% vermindert werden. Zudem wurde ausgehend vom Vektor p Δ S (83 CG; C.2.2.1.2) der CMV-Promotor (31 CG; C.2.1.2.1) durch den CpG-freien CMV_TG-Promotor (C.2.1.1.4) ersetzt (1; Abbildung C.5). Dabei gewährleisteten flankierende singuläre 5' *Clal*- bzw. 3' *SacI*-Schnittstellen diesen Austausch der Promotorbereiche. Das resultierende Plasmid p Δ SC (52 CG) weist bezogen auf pRS (164 CG; C.2.2.1.1) einen CpG-Gehalt von 32% auf.

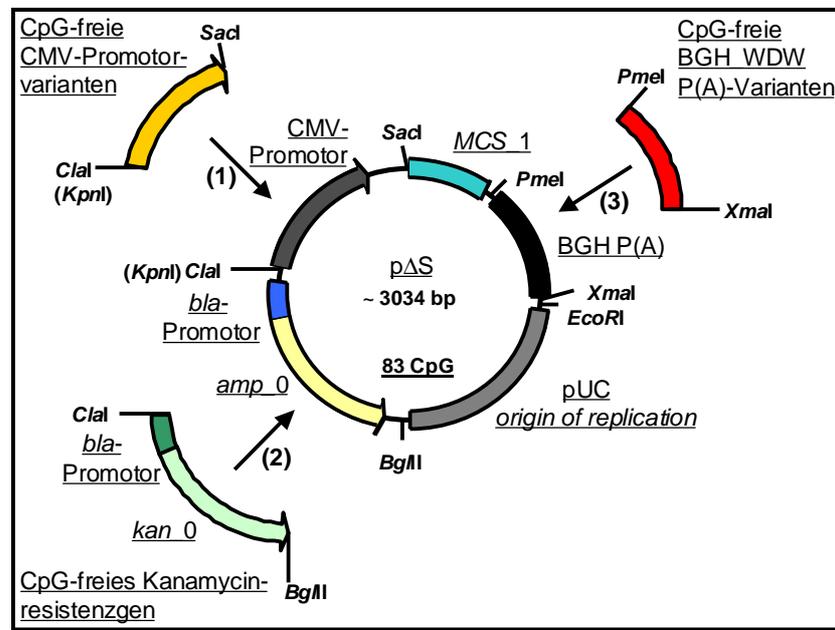


Abbildung C.5: Klonierungsschema zum Erhalt weiterer, CpG-reduzierter Konstrukte.

Synthetisch hergestellte, CpG-minimierte Elemente sind farbig dargestellt, während unmodifizierte, anhand von PCR generierte Komponenten grau bzw. schwarz unterlegt sind. (1) Der Austausch des herkömmlichen CMV-Promotors durch den CMV_TG-Promotor (*Clal*/*SacI*) bzw. durch die im Rahmen dieser Arbeit etablierten CMV_A bis CMV_F-Promotorvarianten (*KpnI*/*SacI*) resultierte in der Generierung der Vektoren p Δ SC bzw. p Δ SC_A bis p Δ SC_F, die 52 CpGs besitzen. (2) Mittels *PmeI*/*XmaI*-Klonierung wurde das konventionelle BGH P(A)-Signal durch die CpG-freien BGH_WDW P(A)-Varianten ersetzt. Die entstandenen Konstrukte werden im Folgenden als p Δ SB_WDW bezeichnet und weisen 80 CpGs auf. (3) Weiter führte die Integration eines CpG-freien Kanamycinresistenzgens (*kan_0*) mit vorgeschaltetem *bla*-Promotor in den p Δ S zum Erhalt des Vektors p Δ S_kana (83 CpG).

Um nachfolgende Klonierungen zu erleichtern, wurde die das 5' Ende des Promotorbereichs flankierende *Clal*-Schnittstelle im p Δ SC unter Verwendung der Fusionsprimer CMVsyn_ *bla*_Fpf und CMVsyn_ *bla*_Fpb gegen *KpnI* ausgetauscht.

Das Enzym *KpnI* weist im Gegensatz zu *Clal* keine *dam*-Sensitivität auf, d.h. kann auch nach Transformation in den Dam-Methylasen enthaltenden Stamm DH5 α verwendet werden. Analog wie für p Δ SC beschrieben (1, Abbildung C.5) wurden die generierten, CpG-freien CMV-Promotorvarianten CMV_A bis CMV_F (siehe C.2.1.1.5) mittels *KpnI/SacI*-Klonierung in den Vektor p Δ S eingebracht. Diese Vektorkonstrukte werden mit p Δ SC_A bis p Δ SC_F bezeichnet und besitzen bezogen auf pRS 52 CpG-Dinukleotide bzw. einen Gehalt von 32%.

Ein weiterer Ansatz zur Verminderung der CpG-Zahl im Vektorrückgrat war die Einführung eines CpG-freien BGH Polyadenylierungssignals in p Δ S. Dazu wurden fünf degenerierte BGH_WDW-Varianten (siehe C.2.1.1.6) unter Verwendung der Restriktionsenzyme *PmeI* und *XmaI* je in ein p Δ S-Vektorrückgrat kloniert (3, Abbildung C.5). Die resultierenden Konstrukte werden im Folgenden je nach Austausch des jeweiligen CpGs durch die Basenabfolge WpG, DpG oder CpW als p Δ SB_WDW bezeichnet. Alle p Δ SB_WDW-Konstrukte weisen einen CpG-Gehalt von 49% auf und besitzen 80 CpG-Dinukleotide innerhalb ihrer Sequenz.

Eine CpG-freie Version des Kanamycinresistenzgens *kan_0* (C.2.1.1.1) wurde unter Verwendung der singulären Schnittstellen (*Scal/HpaI*) in den CpG-reduzierten p Δ S eingefügt (2, Abbildung C.5). Da Kanamycin nicht als therapeutisches Antibiotikum eingesetzt wird, bieten sich derartige Vektoren für humane Applikationen an. Das entstandene Konstrukt p Δ S_kana besitzt ebenso wie p Δ S 83 CpGs bzw. verglichen mit pRS einen CpG-Gehalt von 51% (siehe C.2.2.1.4).

C.2.2.1.4 Übersicht

Unter Verwendung künstlich generierter bzw. anhand randomisierter Oligonukleotide hergestellter Fragmente sowie PCR-basierter Elemente wurden verschiedene Plasmide für transiente Expressionsanalysen hergestellt. Dabei setzen sich alle Vektoren aus einer Transgenexpressionskassette, einer Version des pUC *origin of replication* sowie einer prokaryontischen Resistenzkassette zusammen. Ausgehend von einem pc5-abgeleiteten Referenzkonstrukt pRS (164 CpG; C.2.2.1.1) wurde der modular aufgebaute Basisvektor p Δ S (83 CpG; C.2.2.1.2) etabliert. Ein Ersatz des *amp_0*-Elements (C.2.1.1.1) durch das *kan_0*-Element (C.2.1.1.1) im p Δ S führte zu dem für humane Applikationen anwendbaren Vektor p Δ S_kana (83 CpG; C.2.2.1.3). Weiter führte eine Deletion im 5' Bereich des pUC *origin of replication* zum Erhalt des Plasmids p Δ S110' (C.2.2.1.3), das 77 CpGs beinhaltet. Die Integration von CpG-freien BGH_WDW P(A)-Varianten (80 CpG; C.2.1.1.6) führte zur Etablierung der Konstrukte p Δ SB_AAA, p Δ SB_AGA, p Δ SB_AGT, p Δ SB_ATA sowie p Δ SB_ATT (80 CpG; C.2.2.1.3), die bezogen auf pRS einen CpG-Gehalt von 49% besitzen. Durch Einbau des CMV_TG-Promotors (C.2.1.1.4) in den p Δ S wurde der stärker CpG-reduzierte Vektor p Δ SC (52 CpG; C.2.2.1.3) generiert, dessen CpG-Gehalt bei 32% liegt. Analog wie für p Δ SC beschrieben, wurden die Vektoren p Δ SC_A bis p Δ SC_F (52 CpG; C.2.2.1.3) generiert, welche CpG-freie CMV_DH-Promotorvarianten enthielten. Der CpG-Gehalt dieser Plasmide wurde auf 32% vermindert. In Tabelle C.9

werden die jeweiligen Konstrukte hinsichtlich der CpG-Zahl ihrer Komponenten, der Gesamtzahl an CpG-Dinukleotiden sowie ihres prozentualen CpG-Gehaltes verglichen. Auch die Zahl der im jeweiligen Vektor auftretenden murinen sowie humanen Motive wird Tabelle C.9 aufgeführt.

Vektor	CMV-Prom. (CpG)	MCS (CpG)	BGH P(A) (CpG)	Amp (CpG)	origin (CpG)	Sonstige (CpG)	Gesamt (CpG)	% CpG von pRS
<i>pRS</i>	31	11	3	72	45	2	164	100
<i>pAS</i>	31	1	3	1	45	2	83	51
<i>pAS110'</i>	31	1	3	1	39	2	77	47
<i>pASC</i>	0	1	3	1	45	2	52	32
<i>pASkana</i>	31	1	3	1	45	2	83	51
<i>pASB_WDW</i>	31	1	0	1	45	2	80	49
<i>pASC_A-F</i>	0	1	3	1	45	2	52	32

Tabelle C.9: Übersicht über die für transiente Expressionsstudien etablierten Vektorkonstrukte. Für das jeweilige Konstrukt wird die CpG-Zahl der einzelnen Vektorelemente und der absolute sowie prozentuale CpG-Gehalt des Vektors angegeben. Unter „Sonstige“ werden CpG-Dinukleotide aufgeführt, welche klonierungsbedingt zwischen den Elementen eingefügt wurden bzw. durch Angrenzen zweier funktioneller Elemente entstanden sind.

C.2.2.2 Basisvektoren zur stabilen Expressionsanalyse

In diesem Kapitel wird der Aufbau und die Herstellung der für stabile Expressionsanalysen verwendeten Konstrukte näher beschrieben.

C.2.2.2.1 Ausgangsvektor pcDNA5/FRT

Als Ausgangskonstrukt auf dessen Grundlage die modular aufgebaute Vektorplattform entwickelt wurde, diente der käufliche Vektor pcDNA5/FRT (Invitrogen). Dieses Plasmid wurde gewählt, da der pc5 (C.2.2.2.1) neben funktionellen Elementen wie der Transgenexpressionskassette (A,B,C, Abbildung C.6) mit multiplem Klonierungsbereich, dem pUC *origin of replication* (F, Abbildung C.6) und einer prokaryontischen Resistenzkassette (G, Abbildung C.6) auch ein FRT-Signal (D, Abbildung C.6) zur Vermittlung homologer Rekombination besitzt. Als weitere Komponente enthält der Vektor pc5 einen eukaryontischen Selektionsmarker (E, Abbildung C.6), der die Generierung stabiler Zelllinien gewährleistet. Dadurch ermöglicht der Vektor unter Verwendung des Flp-InTM Systems die gezielte Integration einer Kopie des Transgens in einen definierten Locus der Zielzelle (siehe B.1.3.1.3). Das 5070-bp-große Plasmid besitzt 337 CpG-Dinukleotide in seiner Sequenz und diente als Referenzkonstrukt für weitere aus synthetischen sowie PCR-basierten Elementen aufgebaute Plasmide zur stabilen Expressionsanalyse. Zudem beinhaltet der Vektor 24 murine und 1 humanes CpG-Motiv (siehe A.2).

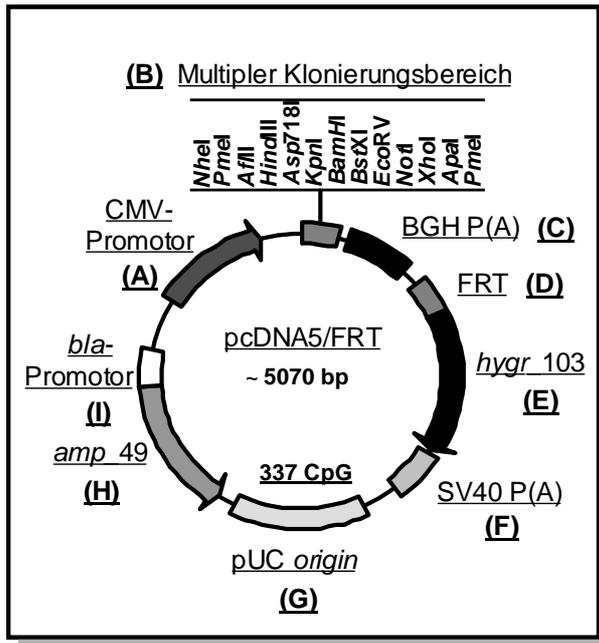


Abbildung C.6: Aufbau des für stabile Expressionsstudien verwendeten Ausgangskonstruktes pc5. Neben einer eukaryontischen Transgenexpressionskassette bestehend aus CMV-Promotor (A), multiplen Klonierungsbereich (B) und BGH P(A)-Signal (C) besitzt der Vektor einen pUC *origin of replication* (G) sowie eine prokaryontische Resistenzkassette *Amp*_174 (H und I). Das FRT-Signal (D), die kodierende Sequenz für das Hygromycinresistenzgen (E) und das SV40 P(A)-Signal (F) werden zu dem Element *Hygr*_174 zusammengefasst. Die Sequenz des Plasmids (5070 bp) enthält 337 CpG-Dinukleotide.

C.2.2.2.2 pc5-Varianten mit CpG-freiem BGH Polyadenylierungssignal

Zwei CpG-freie Versionen des BGH Polyadenylierungssignals, BGH_ATA und BGH_AGA (C.2.1.1.6) wurden in den Vektor pc5 integriert, um in stabilen Zellkulturexperimenten ihren Einfluss auf die Höhe der Transgenexpression untersuchen zu können. Dazu wurde über F-PCR ein Fragment generiert, das im amplifizierten Bereich analog zum pc5 ist, allerdings eine CpG-depletierte Version des BGH Polyadenylierungssignals besitzt. Zudem wurde über die zur Fusion verwendeten Oligonukleotide BGH_F1pf/BGH_F1pb bzw. BGH_F2pf/BGH_F2pb (siehe F.1.1) eine 5' flankierende *PmeI*- bzw. 3' eine flankierende *XmaI*-Erkennungssequenz eingefügt. Anschließend wurde das Fragment unter Verwendung der Primer 773A_pf und ori_mL1L2 amplifiziert und das PCR-Produkt *Bam*HI/*MscI* verdaut (A, Abbildung C.7). Aus der Integration des jeweiligen F-PCR-Produktes in den *Bam*HI/*MscI*-geschnittenen pc5 (B, Abbildung C.7) resultierten die Konstrukte pc5_AGA sowie pc5_ATA (C, Abbildung C.7). Das SV40 Polyadenylierungssignal wurde aus dem Vektor pc5 (Invitrogen) mit den Primern SV40P(A)_pf/SVP(A)_pb amplifiziert und unter Verwendung der singulären Erkennungssequenzen für die Enzyme *PmeI*/*XmaI* in den oben beschriebenen Vektor kloniert. Die *XmaI*-Schnittstelle (CCCGGG) besitzt ein CpG-Dinukleotid, daher weisen die Konstrukte pc5_AGA, pc5_ATA und pc5_SV40 je 335 CpG-Dinukleotide auf, was einem CpG-Gehalt von 99% entspricht.

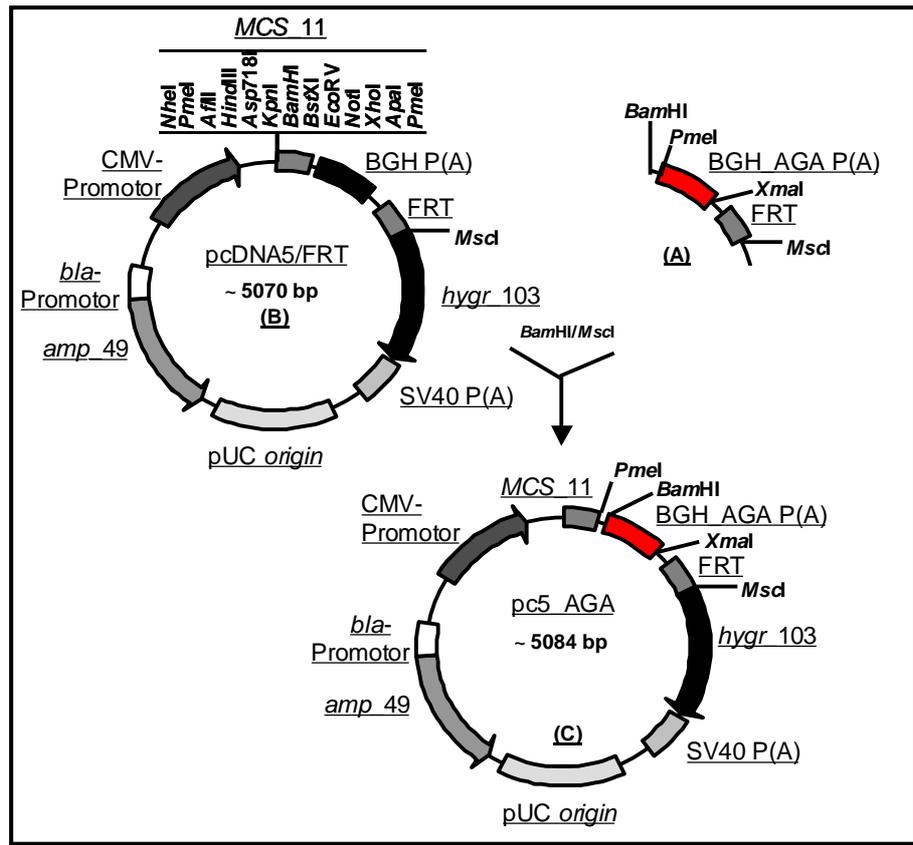


Abbildung C.7: Klonierungsschema zum Erhalt von pc5-Varianten mit CpG-freiem P(A)-Signal. Synthetisch hergestellte, CpG-minimierte Elemente sind farbig dargestellt, während unmodifizierte, anhand von PCR generierte Komponenten in Grau- bzw. Schwarztönen ausgefüllt sind. Zunächst wurde über F-PCR ein Fragment (A) generiert, welches in seiner Sequenz dem pc5 entsprach, jedoch das veränderte Polyadenylierungssignal BGH_AGA sowie zwei flankierende Schnittstellen enthielt. Unter Verwendung der singulären Schnittstellen *BamHI* und *MscI* wurde dieses Fragment in den pc5 (B) kloniert, woraus der Vektor pc5_AGA resultierte (C). Die CpG-freie Variante des BGH P(A)-Signals BGH_ATA sowie das SV40 (A)-Signal wurden anschließend mittels *PmeI/XmaI*-Klonierung in den Vektor pc5_AGA eingesetzt, woraus sich die Konstrukte pc5_ATA und pc5_SV40 ergaben (nicht abgebildet).

C.2.2.2.3 SYN_84

Der modular aufgebaute Vektor p Δ S (I, Abbildung C.8) wurde unter Verwendung der singulären Schnittstellen *EcoRI/XmaI* um die synthetisch hergestellte eukaryontische Selektionsmarkerkassette Hygr_1 (II, Abbildung C.8) erweitert. Das entstandene 5065-bp-große Konstrukt SYN_84 (III, Abbildung C.8) entsprach in seinem Aufbau dem Ausgangsvektor pc5 und wies noch 84 CpG-Dinukleotide auf. Folglich wurde der CpG-Gehalt in SYN_84 auf 25% des ursprünglichen Wertes reduziert. Weiter wurde die Zahl der murinen CpG-Motive im Plasmid auf 9 vermindert, und das humane CpG-Motiv im Vektorrückgrat wurde eliminiert (siehe C.2.2.2.7).

C.2.2.2.4 SYN_187

Um den Vektor SYN_187 zu erhalten, wurde ausgehend vom Plasmid SYN_84 die für die Hygromycinresistenz kodierende Sequenz *hygr_0* (C.2.1.1.2) anhand der

Restriktionsschnittstellen *Afl*II/*Swa*I gegen die Sequenz des Hygromycinresistenzgens aus pc5 *hygr_103* (siehe C.2.1.1.2) ausgetauscht (V, Abbildung C.8). Das resultierende 5065-bp-große Konstrukt beinhaltet 187 CpG-Dinukleotide und verglichen mit pc5 einen CpG-Gehalt von 55%. Die Sequenz des Vektors enthält 12 murine, jedoch kein humanes CpG-Motiv (siehe C.2.2.2.7).

C.2.2.2.5 SYN_226

Zur Konstruktion des CpG-modifizierten Vektors SYN_226 wurde die eukaryontische Selektionsmarkerkassette Hygr_174 (IV, Abbildung C.8) *Xma*I/*Pac*I anstelle der CpG-freien Hygr_1-Komponente (C. 2.1.1.2) in den Vektor SYN_84 integriert. Der 5042-bp-große Vektor SYN_226 enthält 226 CpG-Dinukleotide, 15 murine jedoch kein humanes CpG-Motiv (siehe A.2). Bezogen auf pc5 besitzt der Vektor einen CpG-Gehalt von 67% (siehe C.2.2.2.7).

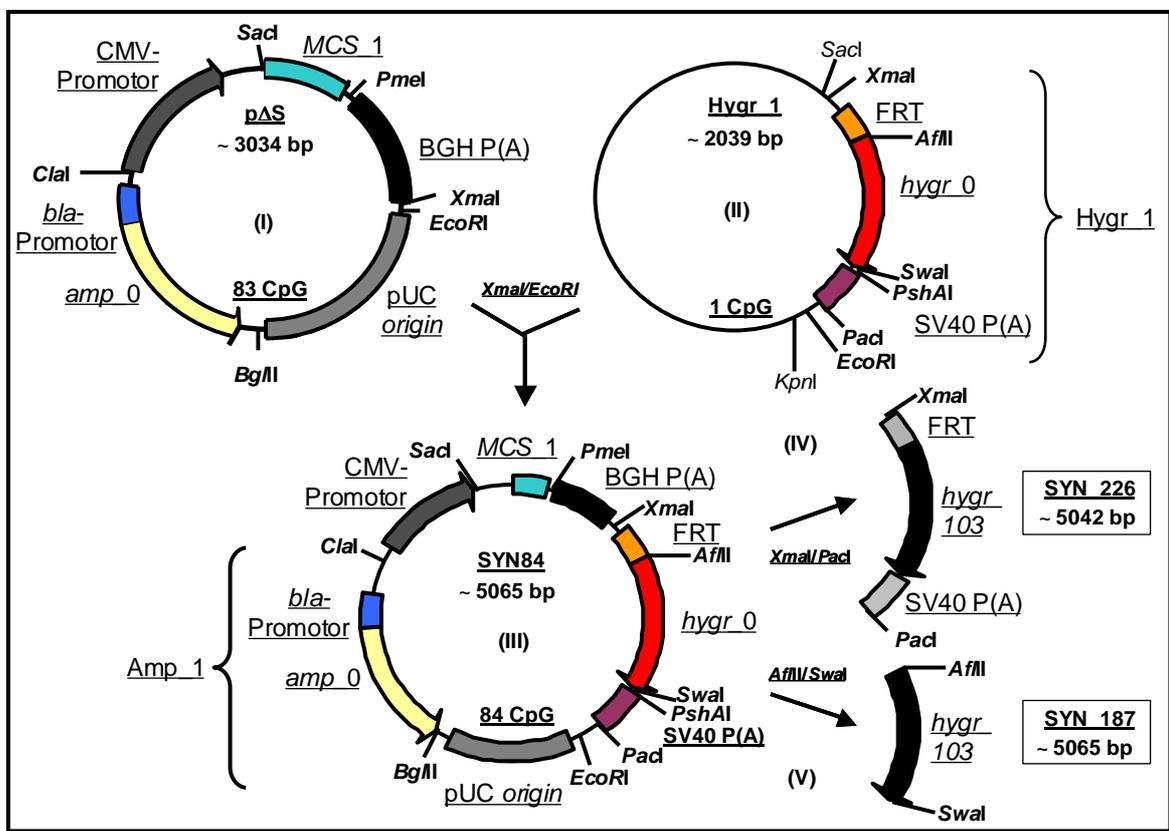
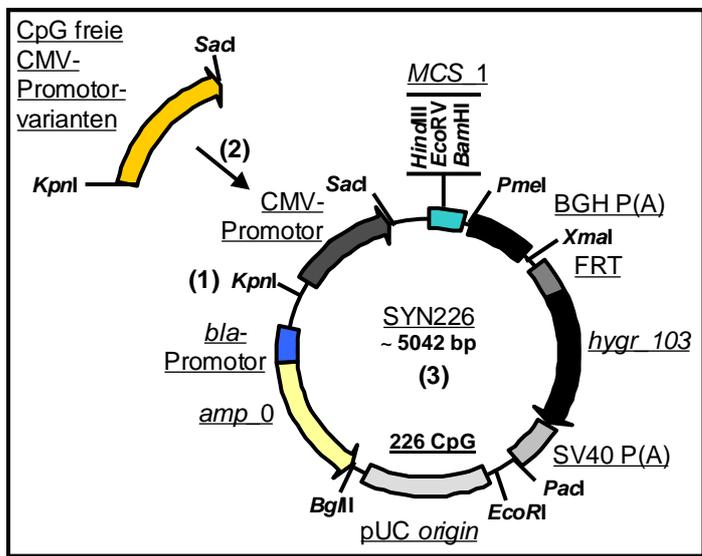


Abbildung C.8: Klonierungsschema zur Herstellung der CpG-reduzierten Vektoren für stabile Expressionsanalysen. Synthetisch hergestellte, CpG-minimierte Elemente sind farbig dargestellt, während unmodifizierte, anhand von PCR generierte Komponenten grau bzw. schwarz unterlegt sind. Ausgehend vom Vektor pΔS (I) wurde durch Integration der Hygr_1-Kassette (II) mittels der Enzyme *Xma*I/*Eco*RI das Konstrukt SYN84 (III) geschaffen. Dabei setzt sich die Hygr_1-Kassette aus dem FRT-Rekombinationssignal, dem kodierenden Bereich für das Hygromycinresistenzgen sowie dem SV40 P(A)-Signal zusammen. Die Amp^r-Kassette beinhaltet neben dem ORF auch einen vorgeschalteten *bla*-Promotor. Ein Austausch des 5' *Xma*I- und 3' *Pac*I-flankierten Bereiches der Hygromycinkassette (Hygr_1) gegen den analogen, CpG-haltigen Bereich aus dem pc5 resultierte im Erhalt des Vektors SYN226 (IV). Die Integration der herkömmlichen für die Hygromycinresistenz kodierenden Sequenz (*hygr_103*) in den SYN84 ergab schließlich das Plasmid SYN187 (V).

C.2.2.2.6 SYN226-Varianten mit CpG-freiem CMV Promotor

Zunächst wurde im Vektor SNYN226 (226 CG; C.2.2.2.5), analog wie für p Δ SC beschrieben (siehe C.2.2.1.3), die 5' den CMV-Promotorbereich flankierende *Cla*I-Schnittstelle gegen die Erkennungssequenz des Enzyms *Kpn*I ausgetauscht (1, Abbildung C.9). Die CpG-depletierten Promotorchimären CMV_A bis CMV_F (siehe C.2.1.1.5) sowie der CMV_TG-Promotor (siehe C.2.1.1.4) wurden anschließend *Kpn*I/*Sac*I in den Vektor SYN_226 eingesetzt (2, Abbildung C.9). Der SV40-Promotor wurde aus dem Vektor pFRTlac/Zeo2 (Invitrogen) mit den Primern SV40prom_pf/ SV40prom_pb (siehe F.1.1) amplifiziert und ebenso in den Vektor kloniert. Die resultierenden Konstrukte (3, Abbildung C.9) werden als SYN226_TG und SYN_SV40 sowie SYN226_A bis SYN226_F bezeichnet. Ihr CpG-Gehalt wurde auf 57% des pc5 vermindert.

Abbildung C.9: Herstellung der SYN226-Varianten



mit CpG-freiem CMV-Promotor. Synthetisch hergestellte, CpG-minimierte Elemente sind farbig dargestellt, während unmodifizierte, anhand von PCR generierte Komponenten in Grau- bzw. Schwarztönen ausgefüllt sind. Zunächst wurde die den 5' Bereich des CMV-Promotors flankierende *Cla*I-Schnittstelle im SYN226 anhand einer Fusions-PCR durch eine *Kpn*I-Erkennungssequenz ersetzt (1). Anschließend wurden CpG-freie CMV_DH-Promotorvarianten sowie der CMV_TG-Promotor (2) unter Verwendung der Enzyme *Kpn*I/*Sac*I in den CpG-reduzierten Vektor SYN226 eingesetzt (3). Nähere Informationen zur Entstehung dieser Konstrukte können dem Text entnommen werden.

C.2.2.2.7 Übersicht

Ausgehend von der Sequenz des Vektors pcDNA5/FRT (337 CpG; C.2.2.2.1) wurden verschiedene CpG-reduzierte Konstrukte für stabile Expressionsanalysen generiert. Dabei setzten sich alle Plasmide aus funktionellen Elementen wie der Transgenexpressionskassette (A,B,C, Abbildung C.6), dem pUC *origin of replication* (F, Abbildung C.6) und einer prokaryontischen Resistenzkassette (G, Abbildung C.6) zusammen. Zudem enthalten die Vektoren Komponenten, welche durch Vermittlung homologer Rekombination (FRT-Signal; D, Abbildung C.6) die Etablierung stabiler Zelllinien ermöglichen (E, Abbildung C.6).

Um den Einfluss der CpG-depletierten BGH P(A)-Signale BGH_AGA und BGH_ATA auf die Höhe der Transgenexpression direkt vergleichen zu können, wurden die Konstrukte pc5_AGA sowie pc5_ATA (siehe C. 2.2.2.2) hergestellt.

Diese entsprechen in ihrer Sequenz bis auf die CpG-freie BGH-Polyadenylierungsstelle dem pc5 und enthalten 335 CpG-Dinukleotide. Bezogen auf pc5 ergibt sich ein CpG-Gehalt von 99%. Der modular aufgebaute CpG-reduzierte Vektor SYN_84 (84 CpG; C.2.2.2.3) besitzt neben PCR generierten Elementen wie dem pUC *origin of replication*, dem eukaryontischen CMV-Promotor und der Polyadenylierungsstelle auch künstlich hergestellte Elemente wie die Ampicillinresistenzkassette Amp_1, die eukaryontische Selektionsmarkerkassette Hygr_1 und den multiplen Klonierungsbereich MCS_1. Insgesamt wurde der CpG-Gehalt in diesem Vektor verglichen mit pc5 auf 25% reduziert (siehe Tabelle C.9). Durch Integration der CpG-haltigen kodierenden Region des Hygromycinresistenzgens *hygr_103* bzw. der kompletten Hygromycinresistenzkassette Hygr_174 in den Vektor SYN_84 wurden die Vektoren SYN_187 (187 CpG; C.2.2.2.4) bzw. SYN_226 (226 CpG; C.2.2.2.5) generiert. Dabei konnte der CpG-Gehalt bezogen auf pc5 für SYN_187 auf 55% und der von SYN_226 auf 67% vermindert werden. Zum Vergleich verschiedener CpG-depletierter Promotorchimären (siehe C.2.1.1.5) bezüglich ihrer Expressionsleistung in stabilen Zellen wurden die Konstrukte SYN226_CMV_A bis SYN226_CMV_F (194 CpG; C.2.2.2.6) etabliert. Als Positivkontrolle wurde das Konstrukt SYN226_CMV_TG (194 CpG; C.2.2.2.6) generiert. Bei den Konstrukten wurde der CMV-Promotor jeweils durch eine CpG-freie CMV-Version ersetzt. Hierdurch reduzierte sich der CpG-Gehalt der Konstrukte auf 58%, bezogen auf pc5. Tabelle C.9 gibt eine Übersicht über die für stabile Expressionsexperimente verwendeten Plasmide.

<i>Vektor</i>	<i>CMV-Prom.</i> (CpG)	<i>MCS</i> (CpG)	<i>BGH P(A)</i> (CpG)	<i>Hygr</i> (CpG)	<i>Origin</i> (CpG)	<i>Amp</i> (CpG)	<i>Sonstige</i> (CpG)	<i>Gesamt</i> (CpG)	<i>% CpG</i> <i>von pc5</i>
pc5	31	11	3	174	45	72	1	337	100
pc5_AGA	31	11	0	174	45	72	2	335	99
pc5_ATA	31	11	0	174	45	72	2	335	99
pc5_SV40	31	11	0	174	45	72	2	335	99
SYN_84	31	1	3	1	45	1	2	84	25
SYN_187	31	1	3	104	45	1	2	187	55
SYN_226	31	1	3	143	45	1	2	226	67
SYN_226 CMVp.	0	1	3	143	45	1	1	194	58

Tabelle C.9: Übersicht über die für stabile Expressionsstudien etablierten Vektorkonstrukte.

Für das jeweilige Konstrukt wird die CpG-Zahl der einzelnen Vektorelemente und der absolute sowie der prozentuale CpG-Gehalt des Vektors angegeben. Unter „Sonstige“ werden CpG-Dinukleotide aufgeführt, welche klonierungsbedingt zwischen den Elementen eingefügt wurden bzw. durch Angrenzen zweier funktioneller Elemente entstanden sind.

C.3 Etablierung von CpG-Modifikationen innerhalb des Transgens

C.3.1 Optimierung und Modifizierung der verwendeten Reportergene

C.3.1.1 GFP als Transgen

Um den Einfluss im offenen Leserahmen vorkommender CpG-Dinukleotide (siehe A.3) auf Höhe und Dauer der Genexpression untersuchen zu können, wurden verschiedene CpG-modifizierte Versionen des GFP-Reportergens (siehe C.1.2) entwickelt. Veränderungen der Ausgangssequenz erfolgten dabei unter Beibehaltung der Aminosäureabfolge des *green-fluorescent-proteins*. Ausgehend von der Sequenz für h_GFP 60 (siehe F.2.1), die 60 CpGs innerhalb der kodierenden Region aufweist (Zolotukhin *et al.*, 1996), wurde eine ebenfalls auf humane Kodonwahl optimierte Version des GFP-Gens (h_GFP 0) generiert, deren ORF kein CpG-Dinukleotid mehr enthält. Durch Maximierung des CpG-Gehaltes innerhalb der kodierenden Region von GFP entstand somit das Konstrukt h_GFP 93 (siehe F.2.1), welches 93 CpG-Dinukleotide aufweist. Ein Nukleinsäurevergleich der drei GFP-Reportergensequenzen ist unter F.2.1 angegeben, ein Aminosäuresequenzvergleich der drei Konstrukte in F.2.2 dargestellt.

Die Modifizierung der Leserahmen hinsichtlich des CpG-Gehaltes erfolgte unter Verwendung der Software Geneoptimizer™ (Geneart, Regensburg) im Rahmen einer Multiparameter-Optimierung (siehe B.1.5). Eine Übersicht über weitere sequenzspezifische Parameter ist in Tabelle C.10 dargestellt. So wurde der *codon adaption index* (CAI), ein Maß für die Güte der Kodonwahl (siehe B.1.5), durch die Depletierung der CpGs nur geringfügig verändert (h_GFP 60: 0,95; h_GFP 0: 0,94), während eine Maximierung des CpG-Gehaltes bei h_GFP 93 (CAI: 0,84) eine Verschlechterung des CAI um 0,11 zur Folge hatte (siehe Tabelle C.3). Die drei gewählten Varianten des GFP-Gens sind hinsichtlich ihres GC-Gehalts (h_GFP 60: 61%; h_GFP 0: 55%; h_GFP 93: 61%) vergleichbar und stimmen in ihrer TpA-Zahl (h_GFP 60: 15; h_GFP 0: 15; h_GFP 93: 15) überein (siehe Tabelle C.10). Zur weiteren Klonierung wurden eine 5' *HindIII*- sowie eine 3' *BamHI*-Schnittstelle an die kodierende Region angefügt. Eine Zusammenstellung der einzelnen Parameter ist in Tabelle C.10 gegeben.

<i>GFP (ORF)</i>	<i>h_GFP_0</i>	<i>h_GFP_60</i>	<i>h_GFP_93</i>
CpG-Anzahl	0	60	93
CAI	0,93	0,96	0,84
GC-Gehalt	55 %	61 %	61 %
TpA-Anzahl	15	15	15
Klonierung	<i>HindIII/BamHI</i>	<i>HindIII/BamHI</i>	<i>HindIII/BamHI</i>

Tabelle C.10: Vergleich der CpG-modifizierten GFP-Reportervarianten. Für den kodierenden Bereich des jeweiligen Reporterkonstruktes wird die CpG-Zahl, der CAI-Wert sowie der prozentuale Anteil der Nukleotide Cytidin und Guanodin an der

Gesamtsequenz (GC-Gehalt) angegeben. Zudem werden die Anzahl der TpA-Dinukleotide sowie die flankierenden Restriktionsschnittstellen der Reportervarianten aufgeführt.

C.3.1.2 p24 als Transgen

Als weiteres Transgen wurde das von HIV-1 abgeleitete p24-Kapsidprotein verwendet, das bereits früher auf eine Expression in humanen Zellen optimiert worden war (Graf *et al.*, 2000). Analog zu den GFP-Reportergenvarianten (siehe C.3.1.1) wurde ausgehend von der humanisierten Sequenz des p24 (h_p24 38) eine CpG-reduzierte Version (h_p24 0) sowie eine Variante mit maximiertem CpG-Gehalt (h_p24 103) konstruiert. Die Depletierung aller CpG-Dinukleotide für h_p24 0 sowie die Erhöhung der CpG-Zahl auf 104 CpGs erfolgte dabei unter Aufrechterhaltung der Aminosäuresequenz in dem Kapsidprotein (siehe F.2.2). Ein Nukleinsäurevergleich der drei p24-Reportergenvarianten ist unter F.2.1 aufgeführt.

Eine Übersicht über weitere Parameter, die im Zuge der CpG-Modifikation verändert wurden, gibt Tabelle C.11. Während die CpG-freie p24-Variante h_p24 0 (CAI: 0,95) einen vergleichbaren CAI-Wert (siehe B.1.5) wie das Ausgangskonstrukt h_p24 38 (CAI: 0,99) aufwies, führt die Maximierung des CpG-Gehaltes bei h_p24 103 zu einer erheblichen Verschlechterung der Kodonwahl (CAI: 0,76). Dabei wurde eine Erhöhung der CpG-Zahl gezielt auf Kosten der Kodonwahl unternommen. So wurde die Frage adressiert, ob CpGs den Einfluss der Kodonwahl auf die Expression überkompensieren können. Hinsichtlich des prozentualen GC-Gehaltes (h_p24 0: 60, h_p24 38: 65, h_p24 104: 65) sowie der TpA-Anzahl (h_p24 0: 6, h_p24 38: 7, h_p24 104: 6) weisen die drei Konstrukte vergleichbare Werte auf. Für nachfolgende Klonierungsschritte wurde das 5' Ende der jeweiligen p24-Sequenz um eine *HindIII*-, das 3' Ende um eine *BamHI*-Schnittstelle ergänzt (siehe Tabelle C.11).

<i>p24 (ORF)</i>	<i>h_p24_0</i>	<i>h_p24_38</i>	<i>h_p24_104</i>
CpG-Anzahl	0	38	104
CAI	0,95	0,99	0,76
GC-Gehalt	60 %	65 %	65 %
TpA-Anzahl	6	7	6
Klonierung	<i>HindIII/BamHI</i>	<i>HindIII/BamHI</i>	<i>HindIII/BamHI</i>

Tabelle C.11: Vergleich der CpG-modifizierten p24-Reportervarianten. Für den kodierenden Bereich des jeweiligen Reporterkonstruktes wird die CpG-Zahl, der CAI-Wert sowie der prozentuale Anteil der Nukleotide Cytidin und Guanodin an der Gesamt-

sequenz (GC-Gehalt) angegeben. Zudem werden die Anzahl der TpA-Dinukleotide sowie die flankierenden Restriktionsschnittstellen der Reportervarianten aufgeführt.

C.3.2 Herstellung der verwendeten Reportergenkonstrukte

C.3.2.1 GFP (*green-fluorescent-protein*) als Transgen

Das CpG-freie Reportergen h_GFP 0 (Auftrag: E019) sowie die CpG-maximierte Variante h_GFP 93 (Auftragsnummer: 040308) wurden synthetisch hergestellt (siehe B.1.5) und unter Verwendung der Schnittstellen *HindIII/BamHI* in die

jeweiligen Expressionsvektoren integriert. Das Ausgangskonstrukt h_GFP 60 wurde von Helga Hofmann (AG Wagner, Regensburg) zur Verfügung gestellt und via PCR anhand der Amplifikationsprimer huGFP_pf/huGFP_pb (siehe F.1.1) mit einer 5' *HindIII*- sowie einer 3' *BamHI*-Schnittstelle versehen. Das erhaltene PCR-Produkt wurde ebenfalls *HindIII/BamHI* in die jeweiligen Expressionsvektoren kloniert. Dabei wurden die GFP-Konstrukte für transiente Expressionsanalysen in die Basisvektoren pRS und pΔS sowie in die modifizierten Basisvektoren pΔS110', pΔSC, pΔS_kana, pΔSC_DH und pΔSB_WDW eingesetzt (siehe C.2.2.1). Für stabile Expressionsexperimente wurden die Plasmide pc5, pc5_ATA, pc5_AGA sowie die Vektoren SYN_84, SYN_187, SYN_226 sowie SYN_CMV_DH (siehe C.2.2.2) verwendet.

C.3.2.2 p24-Kapsidprotein von HIV-1 als Transgen

Die CpG-modifizierte Konstrukte h_p24 0 (Auftragsnummer: 03-501) sowie h_p24 104 (Auftragsnummer: 040308) wurden synthetisch hergestellt, während die Ausgangssequenz h_p24 38 unter Verwendung der Oligonukleotide p24_pf und p24_pb mittels PCR aus einem HIV-1 *syngag* Konstrukt (Graf *et al.*, 2000) amplifiziert wurde. Für nachfolgende Klonierungen wurde zudem durch die Amplifikationsprimer eine 5' *HindIII*- sowie eine 3' *BamHI*-Schnittstelle an die Sequenz angehängt. Zur transienten Expressionsanalyse wurden die Konstrukte h_p24 0, h_p24 38 und h_p24 104 unter Verwendung der flankierenden Schnittstellen in die Vektoren pRS, pΔS, pΔS110' sowie pΔSC (siehe C.2.2.1) eingefügt.

C.3.3 Zusammenfassung

Um den Einfluss von CpG-Dinukleotiden innerhalb des Leserahmens auf die Regulation der Genexpression untersuchen zu können, wurden zwei verschiedene Reportergene hinsichtlich ihres CpG-Gehaltes modifiziert. Zum einen diente die auf humane Kodonwahl optimierte Version des *green-fluorescent-protein*-Gens (Zolotukhin *et al.*, 1996) aus der Qualle *Aequorea victoria* als Transgen (60 CpG; C.3.1.1). Zum anderen wurde die kodierende Region des bereits früher auf Expression in humanen Zellen optimierten p24-Kapsidproteins aus HIV-1 (Graf *et al.*, 2000) als Reporter verwendet (38 CpG; C.3.1.2). Multiparameter-Optimierung und künstlicher Genaufbau ermöglichten die Herstellung je einer im Vergleich zum Ausgangskonstrukt (h_GFP 60 bzw. h_p24 38) CpG-reduzierten (h_GFP 0 bzw. h_p24 0) sowie einer CpG-maximierten Version des Transgens (h_GFP 93 bzw. h_p24 104). Eine Depletion des CpG-Gehaltes bei h_GFP 0 bzw. h_p24 0 führte zu einer Verringerung der CpG-Zahl um 100%. Die Varianten h_GFP 93 (93 CpG) bzw. h_p24 104 (104 CpG) wiesen dagegen im Vergleich zum Ausgangskonstrukt einen um 55% bzw. 173% gesteigerten CpG-Gehalt auf. Dabei wurde die Aminosäuresequenz des jeweiligen Proteins durch die CpG-Modifizierungen, wie der unter F.2.2 dargestellte Aminosäuresequenzvergleich für GFP bzw. p24 zeigt, nicht verändert.

Für h_GFP 0 (CAI: 0,93) bzw. für h_p24 0 (CAI: 0,95) führte die Entfernung aller CpG-Dinukleotide aus der kodierenden Region zu keiner erheblichen Änderung der Kodongüte. Eine Maximierung der CpGs innerhalb des Leserahmens erfolgte jedoch sowohl für h_GFP 93 (CAI: 0,84) als auch für h_p24 104 (CAI: 0,76) auf Kosten der Kodonwahl (siehe Tabelle C.10 bzw. C.11). Die Ausgangstransgene h_GFP 60 (CAI: 0,96) bzw. h_p24 38 (CAI: 0,99) wurden PCR basiert generiert und analog zu den modifizierten Varianten mit flankierenden singulären 5' *HindIII*- sowie 3' *BamHI*-Schnittstellen versehen. Anschließend wurden die jeweiligen Transgene unter Verwendung eben dieser Enzyme in die jeweiligen Expressionsvektoren eingesetzt. Ein Nukleinsäurevergleich der generierten Reportertransgenvarianten ist unter F.2.1 gegeben. Tabelle C.12. charakterisiert die einzelnen Transgenvarianten hinsichtlich ihrer verwendeten Klonierungsschnittstellen, ihres absoluten sowie auf das Ausgangstransgen bezogenen relativen CpG-Gehalts. Zudem wird die Zahl muriner CpG-Motive aufgeführt, die der Basenpaarabfolge RRCpGYT entsprechen. Humane CpG-Motive, denen die Sequenz GTCGTT zu Grunde liegt, sind in keiner der Transgenvarianten vorhanden.

<i>Reportertransgenvarianten</i>	<i>Singuläre Schnittstellen</i>	<i>CpG-Dinukleotide</i>	<i>% CpG-Dinukleotide</i>	<i>Murine CpG-Motive</i>
<i>h_GFP 0</i>	<i>HindIII/BamHI</i>	0	0	-
<i>h_GFP 60</i>	<i>HindIII/BamHI</i>	60	100	1
<i>h_GFP 93</i>	<i>HindIII/BamHI</i>	93	155	3
<i>h_p24 0</i>	<i>HindIII/BamHI</i>	0	0	-
<i>h_p24 38</i>	<i>HindIII/BamHI</i>	38	100	3
<i>h_p24 104</i>	<i>HindIII/BamHI</i>	104	273	12

Tabelle C.12: Übersicht über CpG-Gehalt bzw. CpG-Motive der verwendeten Reportertransgenvarianten. Weitere Informationen können dem Text zu entnehmen werden.

C.4 Transiente Expressionsstudien

In transienten Expressionsstudien wurden verschiedene Reportertransgene mit modifiziertem CpG-Gehalt (siehe C.3.1) bzw. Vektoren unterschiedlicher CpG-Zahl (siehe C.2.2.1) anhand ihrer Expressionsleistung verglichen. Dazu wurden die Konstrukte in Doppelansätzen, wie in Kapitel B.2.3 beschrieben, in transiente Transfektionen eingesetzt. Anschließend wurde die Reporteraktivität des verwendeten Transgens als Maß für die Höhe der Genexpression bestimmt.

C.4.1 Vergleich der CpG-modifizierten Reportertransgenvarianten

Zu Beginn der Arbeit gab es Hinweise, dass die Expression eines Transgens durch in der kodierenden Region gelegene CpG-Dinukleotide negativ beeinflusst würde (Yew *et al.*, 2002). Um diese Frage zu adressieren, wurden die drei GFP-Reportertransgenvarianten (siehe C.3.1.1) h_GFP 0, h_GFP 60 sowie h_GFP 93, welche 0 CpG-, 60 CpG- bzw. 93 CpG-Dinukleotide aufwiesen, in den Vektor pAS

(siehe C.2.2.1.2) kloniert. In einem weiteren Ansatz zur Untersuchung dieser Fragestellung wurden drei CpG-modifizierte p24-Reporterkonstrukte (siehe C.3.1.2) h_p24 0, h_p24 38 und h_p24 104 in den CpG-reduzierten Vektor p Δ S eingebracht. Nach transienter Transfektion von humanen Lungenkarzinomzellen (H1299) mit den Reporterkonstrukten wurde die resultierende GFP-Aktivität im Durchflusszytometer (siehe B.3.3) bzw. die p24-Kapsidantigenproduktion im p24-ELISA (siehe B.3.2) bestimmt. Für GFP sind die Angaben der mittleren Fluoreszenzintensitäten (MFI) aus drei unabhängigen Transfektionen in Abbildung C.10A zusammengefasst, die p24-Konzentrationen drei unabhängiger Experimente sind in Abbildung C.10B dargestellt. Die für die Ausgangskonstrukte h_GFP 60 bzw. h_p24 38 ermittelte Reporteraktivität wurde als 100% festgesetzt.

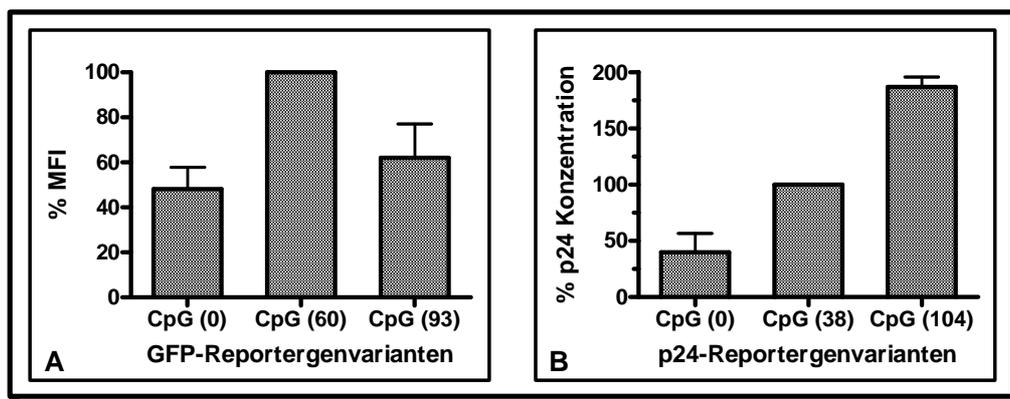


Abbildung C.10: Vergleich der CpG-modifizierten Reportergergenvarianten hinsichtlich ihrer Transgenexpression. H1299-Zellen wurden mit GFP- bzw. p24-Reportervarianten transient transfiziert. Anschließend wurde im Durchflusszytometer die GFP-vermittelte Autofluoreszenz (A) bzw. im p24-ELISA die Menge des produzierten Kapsidantigens (B) bestimmt. Die jeweiligen MFIs bzw. p24-Konzentrationen wurden dabei um die für *mock*-transfizierte Zellen ermittelten Werte bereinigt. Die gezeigten Daten sind Mittelwerte aus drei unabhängigen Transfektionen, die mit zwei unterschiedlichen DNA-Präparationen durchgeführt wurden. Für die transfizierten Konstrukte ist die CpG-Zahl im Transgen sowie die resultierende Reporteraktivität angegeben. (A) Expression der GFP-Reporterkonstrukte h_GFP 0 (CpG 0), h_GFP 60 (CpG 60) und h_GFP 93 (CpG 93) im Vektor p Δ S. (B) Expression der hinsichtlich ihres CpG-Gehaltes variierenden Reporterkonstrukte h_p24 0 (CpG 0), h_p24 38 (CpG 38) sowie h_p24 104 (CpG 104) im Vektor p Δ S.

Wie aus Abbildung C.10A hervorgeht, führte eine Depletion der CpG-Dinukleotide bei dem Konstrukt h_GFP 0 nicht zu einer Verbesserung der Reporterexpression, sondern resultierte anders als erwartet in einer 50%-igen Verringerung der MFI. Eine Maximierung des CpG-Gehaltes im Leserahmen der GFP-Variante h_GFP 93 bewirkte jedoch keine weitere Erhöhung der Transgenexpression, sondern hatte eine Reduktion der MFI auf 62% zur Folge (Abbildung C.10A). Damit lag die vom CpG-maximierten Konstrukt h_GFP 93 hervorgerufene Reporteraktivität zwischen der des CpG-freien (h_GFP 0) und der des Ausgangskonstruktes mit 60 CpG-Dinukleotiden (h_GFP 60). Analog zu GFP kam es bei Verwendung von p24 als Reporter für die CpG-reduzierte Version h_p24 0 zu einer Verringerung der Reporteraktivität auf 40% (linker Balken, Abbildung C.10B). Eine Maximierung des CpG-Gehaltes hingegen resultierte bei h_p24 104 in einer 87%-igen Steigerung der Antigenexpression auf 187% (rechter Balken, Abbildung C.10B). Im Gegensatz zu GFP war bei Verwendung von p24 als

Reporter eine klare Korrelation der im Leserahmen lokalisierten CpG-Anzahl und der Höhe der gemessenen Reporterexpression erkennbar (siehe Abbildung C.10B). Inwieweit diese Korrelation durch das verwendete Vektorrückgrat beeinflusst wird, wurde durch eine vergleichende Analyse unterschiedlicher Kombinationen von Plasmid und Antigen untersucht.

C.4.2 Vergleich der Basisvektoren hinsichtlich ihrer Expressionsstärke

Aus der Literatur ist bekannt, dass im Promotorbereich (Deng *et al.*, 2001; Agirre *et al.*, 2003) bzw. im Vektorrückgrat (Yew *et al.*, 2002; Reyes-Sandoval *et al.*, 2004) lokalisierte CpG-Dinukleotide die Expression sowie die Immunogenität eines Vektors beeinflussen können. Daher wurden drei Plasmide mit verringertem CpG-Gehalt generiert und hinsichtlich ihrer Expressionsstärke *in vitro* mit einem Referenzvektor pRS (164 CpG) verglichen. In die CpG-modifizierten Vektoren p Δ S (83 CpG), p Δ S110' (77 CpG) und p Δ SC (52 CpG) sowie in das Referenzkonstrukt pRS (164 CpG) wurde das CpG-freie GFP-Reportergen h_GFP 0 einkloniert. Um eine breitere Datengrundlage zu schaffen, wurde das 38 CpGs beinhaltende h_p24 38-Konstrukt ebenfalls in die Vektorauswahl eingebaut. Mit den erhaltenen Konstrukten wurden H1299-Zellen transient transfiziert. Anschließend wurde die resultierende Reporteraktivität für GFP bzw. p24 quantitativ bestimmt. In Abbildung C.11 werden die mittleren Fluoreszenzintensitäten (MFI) von GFP für die jeweiligen Vektorkonstrukte dargestellt, wobei die Mittelwerte aus 4 unabhängigen Transfektionen gezeigt werden. Als Transgen wurde das CpG-freie GFP-Konstrukt verwendet, da dies eine relativ geringe Reporterexpression zeigt (siehe Abbildung C.10A) und daher selbst nach Steigerung der Expressionsraten nicht in toxischer Konzentration vorliegen sollte. Die vom Referenzkonstrukt pRS hervorgerufene MFI wurde dabei auf 100% festgesetzt. Wie in Abbildung C.11 erkennbar ist, sind alle generierten Vektoren funktionell und exprimieren GFP in vergleichbarem Maße wie das Referenzkonstrukt pRS. Innerhalb der CpG-reduzierten Vektoren lassen sich nur geringe Unterschiede in der Transgenexpression feststellen, was durch das geringe Expressionsniveau des gewählten Transgens bedingt sein kann. Dennoch ist bei Verwendung von h_GFP 0 die Tendenz erkennbar, dass die künstlich hergestellten Plasmide p Δ S und p Δ S110' im Vergleich zu pRS eine erhöhte Transgenexpression aufweisen.

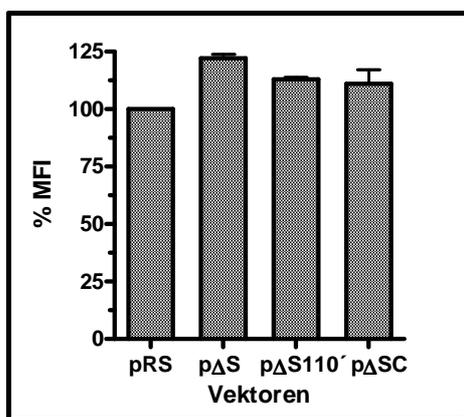


Abbildung C.11: Vergleich der CpG-reduzierten Vektorrückgrate hinsichtlich ihrer h_GFP 0 Expression. H1299-Zellen wurden mit den im Text beschriebenen Konstrukten (x-Achse) transient transfiziert. Anschließend wurde im Durchflusszytometer die GFP-vermittelte Autofluoreszenz (y-Achse) bestimmt. Dargestellt sind die Mittelwerte aus vier unabhängigen Transfektionen. Die MFI des Referenzkonstruktes wurde mit 100% angenommen, zudem wurden die für die Vektorkonstrukte ermittelten MFI-Werte um die Autofluoreszenz *mock*-transfizierter Zellen bereinigt.

Ein analoger Zusammenhang wurde auch für das zweite untersuchte Testantigen festgestellt (siehe Abbildung C.12). In der Abbildung ist für die einzelnen Vektor-konstrukte eine Übersicht der produzierten p24-Konzentrationen dargestellt. Angegeben sind die Mittelwerte aus 8 unabhängigen Experimenten. Da H1299-Zellen ausgehend von h_p24 38 noch eine mehr als 100%-ige Expressionssteigerung des Kapsidantigens tolerieren (siehe Abbildung C.10B), wurde bei diesem Ansatz h_p24 38 als Reporter gen eingesetzt. Wie Abbildung C.12 zu entnehmen ist, wurde bei allen untersuchten Vektoren eine Expression des p24-Kapsidantigens nachgewiesen. Das Plasmid p Δ SC (52 CpG) besitzt einen modifizierten CMV-Promotor, der generell eine schwächere Transgenexpression bedingt. Daher erklärt sich die relativ geringe p24-Expression dieses Konstruktes. Analog zu GFP wiesen die CpG-reduzierten Plasmide p Δ S (83 CpG) und p Δ S110' (77 CpG) die höchsten p24-Konzentrationen auf. Dies gibt einen Hinweis auf eine mögliche Beteiligung der im Vektor gelegenen CpG-Dinukleotide an der Regulation der Expression. In der Abbildung treten jedoch relativ große Standardabweichungen auf, daher kann keine signifikante Aussage getroffen werden. Genauere Untersuchungen des CpG-Einflusses auf die Transgenexpression wurden daher in stabilen Expressionsanalysen durchgeführt (siehe C.5).

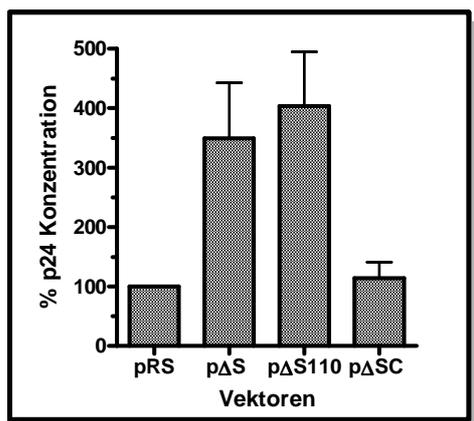


Abbildung C.12: Vergleich der CpG-reduzierten Vektorrückgrate hinsichtlich ihrer h_p24 38-Transgenexpression. H1299-Zellen wurden mit den im Text beschriebenen Konstrukten (x-Achse) transient transfiziert. Anschließend wurde im p24-ELISA die Menge des p24-Kapsidantigens (y-Achse) bestimmt. Dargestellt sind die Mittelwerte aus drei unabhängigen Transfektionen. Die p24-Konzentration des Referenzkonstruktes wurde dabei als 100% festgesetzt. Zudem wurden die für die Vektorstrukturen ermittelten p24-Konzentrationen um die für *mock*-transfizierte Zellen detektierten Werte bereinigt.

Zusätzlich zu den verwendeten Genvarianten (h_GFP 0 bzw. h_p24 38) wurden h_GFP 60 bzw. h_p24 0 als Reporter in die jeweiligen Vektoren eingesetzt. In transienten Experimenten wurden die GFP- bzw. p24-Reporterkonstrukte anschließend hinsichtlich ihrer Transgenexpression verglichen. In Übereinstimmung mit den unter C.4.1 generierten Daten war für alle CpG-freien Konstrukte unabhängig vom Vektorrückgrat eine schwächere Reporteraktivität als für ihre CpG-haltigen Pendanten festzustellen (Daten nicht gezeigt). Da für unabhängige Antigene ein Zusammenhang zwischen CpG-Gehalt des Vektors und Expressionsstärke zumindest im Ansatz gezeigt werden konnte, ist es naheliegend, den Einfluss weiterer CpG-Reduktionen im Vektorrückgrat zu untersuchen. Die Veränderung der Plasmidsequenz in regulatorischen Vektorbereichen kann allerdings deren Funktionalität erheblich beeinträchtigen. Daher wurden weitere, im Rahmen dieser Arbeit generierte CpG-freie Vektorelemente zunächst auf ihre Funktionalität hin untersucht.

C.4.3 Funktionalität weiterer CpG-freier Vektorelemente

C.4.3.1 p Δ S mit Kanamycinresistenzkassette

Ausgehend vom Vektor p Δ S (83 CpG) wurde das Plasmid p Δ S_kana (83 CpG) entwickelt, welches Kanamycinresistenz vermittelt. Das Konstrukt ist ein geeigneter Kandidat für humane Applikationen, da die Bildung von Ampicillinresistenzen, welche die Wirksamkeit bestimmter Medikamente einschränken, verhindert wird. In transienten Experimenten wurde das Plasmid p Δ S_kana hinsichtlich Funktionalität und resultierender Expressionshöhe analysiert. Dabei wurde das CpG-depletierte h_GFP 0-Konstrukt als Reporter gewählt, weil es wie aus Abbildung C.10A hervorgeht im Vergleich zu h_GFP 60 eine geringe mittlere Fluoreszenzintensität (MFI) induziert. So wurde gewährleistet, dass eine Steigerung der Expression möglich war, und keine toxischen Effekte vermittelt wurden. In Abbildung C.13 werden die MFIs der beiden Vektorkonstrukte gegenübergestellt, wobei die Mittelwerte aus 3 unabhängigen Transfektionen gezeigt werden. Demnach ist der Vektor p Δ S_kana funktionell und exprimiert das gewünschte Transgen doppelt so gut wie das Plasmid p Δ S, dessen GFP-Reportergenaktivität mit 100% angesetzt wird.

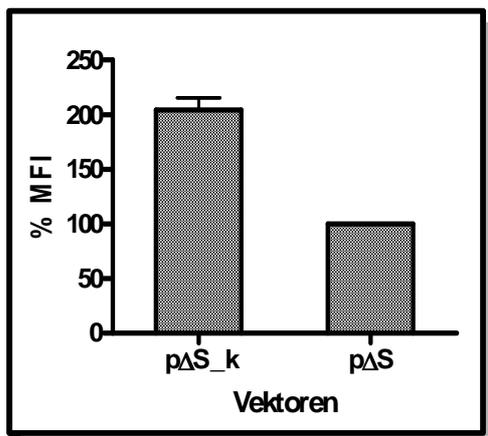


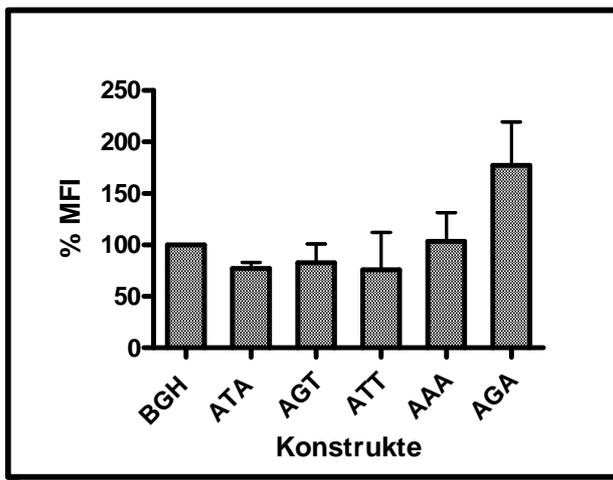
Abbildung C.13: Funktionalität des CpG-reduzierten Vektors p Δ S_kana. H1299-Zellen wurden mit p Δ S- bzw. p Δ S_kana-Expressionskonstrukten transient transfiziert (x-Achse). Anschließend wurde in FACS-Analysen die GFP-vermittelte Autofluoreszenz (MFI) des h_GFP 0-Reporters bestimmt (y-Achse). Die für die Vektoren ermittelten MFI-Werte wurden um die Autofluoreszenz *mock*-transfizierter Zellen bereinigt. In der Graphik sind die Mittelwerte aus drei unabhängigen Transfektionen dargestellt, wobei eine DNA-Charge verwendet wurde. Die Reporteraktivität des Vektors p Δ S wurde dabei als 100% angenommen.

C.4.3.2 Expressionsstärke verschiedener CpG-freier BGH P(A)-Signale in p Δ S

Ausgehend vom Vektor p Δ S (83 CpG) wurden 5 Plasmide generiert, deren Sequenz analog zu p Δ S ist, aber anstelle des herkömmlichen Polyadenylierungssignals (3 CpG) je eine CpG-freie Variante des BGH P(A)-Signals (siehe C.2.1.1.6) enthält. Um die Funktionsfähigkeit dieser Konstrukte zu überprüfen bzw. sie hinsichtlich ihrer Antigenproduktion einstufen zu können, wurden p Δ S_ATA, p Δ S_AGT, p Δ S_ATT, p Δ S_AAA sowie p Δ S_AGA (80 CpG) in transienten Transfektionsstudien mit H1299-Zellen untersucht. Dabei steht in der Dreierkombination XXX der Buchstabe X für die Einführung der Base X anstelle eines C oder G, was im P(A)-Signal die Zerstörung des jeweiligen CpG-Dinukleotides zur Folge hat. Als Reporter diente das im Durchflusszytometer nachweisbare GFP, wobei aus unter C.4.2.1 beschriebenen Gründen die h_GFP 0-Variante verwendet wurde. Abbildung C.12 zeigt die mittlere Fluoreszenzintensität (MFI) für die 5 Konstrukte,

wobei der aus 3 Transfektionen errechnete Mittelwert dargestellt ist. Die MFI des als Positivkontrolle dienenden Vektors p Δ S wurde mit 100% angesetzt. Wie aus der Darstellung entnommen werden kann, wiesen alle fünf Konstrukte Reporteraktivität auf. Die CpG-freien BGH P(A)-Versionen ATA, AGT sowie ATT exprimierten GFP in geringerem Maße als das konventionelle BGH P(A)-Signal, während die Variante AAA eine vergleichbare GFP-Produktion zeigte wie die Positivkontrolle. Die Verwendung von BGH_AGA führte in transienten Experimenten zu einer etwa 70%-igen Erhöhung der Expressionsleistung (siehe Abbildung C.14). Die Ergebnisse transienter Studien werden jedoch von einer Vielzahl an Faktoren wie z.B. der verwendeten DNA-Charge, den Transfektionsbedingungen und dem Zustand der transfizierten Zellen beeinflusst. Aus diesem Grund wurden die zwei CpG-freien Varianten BGH_ATA und BGH_AGA in stabilen Expressionsstudien (siehe C.5) weiter untersucht, in denen die erwähnten Faktoren keine Rolle spielen. Zusätzlich zu CpG-freien BGH P(A)-Signalen wurden auch CpG-depletierte CMV-Promotorvarianten in transienten Experimenten analysiert (siehe C.4.3.3).

Abbildung C.14: Expressionsstärke verschiedener CpG-freier BGH P(A)-Signale bei Verwendung von h_GFP 0 als Transgen.



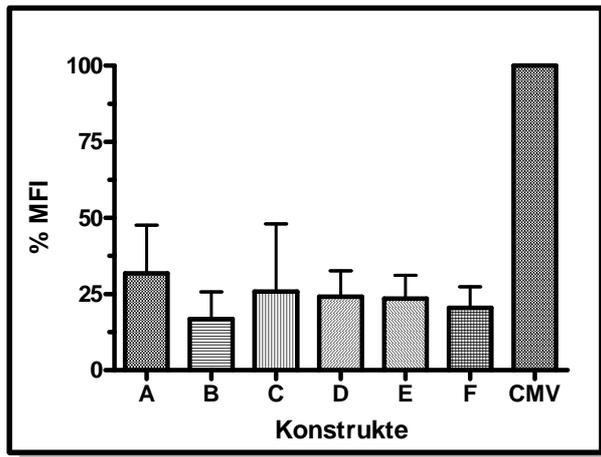
H1299-Zellen wurden mit den Konstrukten p Δ SB_ATA, p Δ SB_AGT, p Δ SB_ATT, p Δ SB_AAA sowie p Δ SB_AGA transient transfiziert (x-Achse). Anschließend wurde in FACS-Analysen die GFP-vermittelte MFI des h_GFP 0-Reporters bestimmt (y-Achse). Alle ermittelten MFI-Werte wurden dabei um die Autofluoreszenz *mock*-transfizierter Zellen bereinigt. Die Expressionsleistung des Vektors p Δ S, welcher das herkömmliche BGH P(A)-Signal beinhaltet, wurde als 100% angenommen. Dargestellt sind die Mittelwerte aus drei unabhängigen Transfektionen, wobei zwei verschiedene DNA-Chargen verwendet wurden.

C.4.3.3 Expressionsstärke verschiedener CpG-freier CMV-Promotoren in p Δ S

Die CpG-depletierten Promotoren CMV_A bis CMV_F (siehe C.2.1.1.5) wurden in den Vektor p Δ S (83 CpG) eingesetzt, welcher h_GFP 0 als Reporter gen beinhaltet. Als Positivkontrolle für transiente Expressionsstudien wurde der herkömmliche CMV-Promotor, der zu den stärksten, derzeit bekannten viralen Promotoren zählt (Löser *et al.*, 1998), in das Plasmid p Δ S integriert. Die so erhaltenen Konstrukte wurden, wie in Kapitel C.4.3.2 beschrieben, in transienten Transfektionsexperimenten mit H1299-Zellen hinsichtlich ihrer Expressionsleistung verglichen. Abbildung C.15 gibt eine Übersicht über die von den einzelnen Promotoren hervorgerufene Reporteraktivität (MFI). Es ist für jedes Konstrukt der Mittelwert aus 4 unabhängigen Transfektionen angegeben, wobei die für den konventionellen CMV-Promotor ermittelte MFI gleich 100% gesetzt wurde. Aus Abbildung C.15 geht hervor, dass alle CpG-depletierten CMV-Promotoren (CMV_A bis

CMV_ F) GFP exprimierten, wenn auch in erheblich geringerem Maße als die Positivkontrolle (CMV). Schwankungen in den einzelnen Transfektionsexperimenten (siehe C.4.3.2) erschwerten jedoch einen direkten Vergleich der Promotorvarianten. Daher wurde eine quantitative Analyse der Promotorstärken in stabilen Expressionsexperimenten (siehe C.5) durchgeführt.

Abbildung C.15: Expressionsstärke verschiedener CpG-freier CMV-Promotoren bei Verwendung von h_GFP 0 als Transgen.



H1299-Zellen wurden mit den im Text beschriebenen Konstrukten transient transfiziert (x-Achse). Anschließend wurden die Plasmide p Δ SC_A (A) bis p Δ SC_F (F), welche CpG-depletierte Versionen des CMV-Promotors beinhalten, hinsichtlich ihrer h_GFP 0 Reporterexpression mit dem Vektor p Δ S (CMV) verglichen, der den herkömmlichen CMV-Promotor enthält. Alle ermittelten MFI-Werte wurden um die Autofluoreszenz *mock*-transfizierter Zellen bereinigt (y-Achse). Dargestellt sind die Mittelwerte aus vier unabhängigen Transfektionen, wobei zwei verschiedene DNA-Chargen verwendet wurden.

C.4.4 Zusammenfassung

Um den Einfluss von im Transgen lokalisierten CpG-Dinukleotiden auf die Antigenexpression zu untersuchen, wurden je drei sich hinsichtlich ihres CpG-Gehaltes unterscheidende GFP- bzw. p24-Genvarianten generiert und in transienten Experimenten analysiert. Anders als erwartet wurde eine erhöhte Expression des CpG-haltigen Konstruktes im Vergleich zum CpG-freien Transgen festgestellt. Eine Maximierung der CpG-Zahl auf 104 CpGs führte bei dem Konstrukt h_p24 104 zu einer weiteren Steigerung der p24-Konzentration. Somit korrelierte die Zahl der in der kodierenden Region gelegenen CpG-Dinukleotide positiv mit der Höhe der gemessenen Reporteraktivität (siehe C.4.1). Darüber hinaus wurde der Einfluss von im Vektorrückgrat vorkommenden CpG-Dinukleotiden auf die GFP- bzw. p24-Reporteraktivität untersucht. Dazu wurden die drei CpG-reduzierten Plasmide p Δ S (83 CpG), p Δ S_110' (77 CpG) sowie p Δ SC (52 CpG) hinsichtlich ihrer Funktionalität und Expressionsleistung mit einem Referenzvektor pRS (164 CpG) verglichen. Alle untersuchten Plasmidkonstrukte exprimierten das gewünschte Transgen, wobei die CpG-reduzierten Vektoren p Δ S und p Δ S_110' tendenziell eine stärkere p24- bzw. GFP-Expression aufwiesen (siehe C.4.2). Zudem wurden mehrere CpG-freie Vektorelemente im Kontext des Vektors p Δ S auf Funktionsfähigkeit und Expressionsstärke untersucht. Dabei zeigte sich, dass das Plasmid p Δ S_kana, das ein CpG-freies Kanamycinresistenzgen besitzt, zu einer höheren Transgenproduktion führte als p Δ S (siehe C.4.3.1). Alle unter C.2.1.1.7 beschriebenen CpG-freien BGH P(A)-Versionen exprimierten im p Δ S-Vektorkontext das gewünschte GFP-Reportergen, wobei die Variante BGH_AGA

die beste Expression zeigte (siehe C.4.3.2). Auch die CpG-depletierten CMV-Promotorvarianten CMV_A bis CMV_F wurden in transienten Promotorstudien hinsichtlich ihrer GFP-Reporteraktivität verglichen. Dabei stellte sich heraus, dass alle CpG-freien CMV-Versionen das *green-fluorescent-protein* exprimierten, wenn auch in erheblich geringerem Maße als der herkömmliche CMV-Promotor (siehe C.4.3.3). Zur Verifizierung der in transienten Studien generierten Daten sowie zum Erhalt quantitativer Vergleichswerte wurden nachfolgend stabile Expressionsanalysen (siehe C.5) durchgeführt.

C.5 Stabile Expressionsstudien

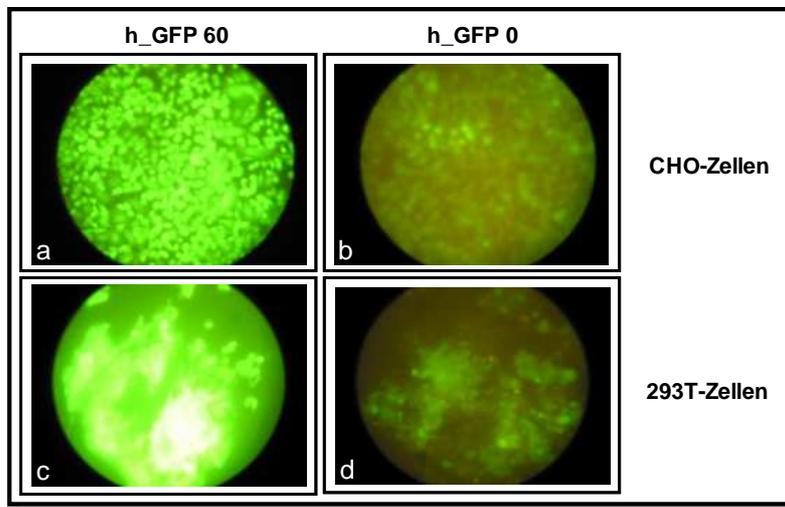
Für eine schnelle Etablierung und Selektion von stabilen rekombinanten Zellen wurde das unter B.1.3.1 erläuterte Flp-In System (Invitrogen) verwendet. Ein wesentlicher Vorteil dieses Systems besteht in einer gerichteten Integration einer Kopie des Transgens in einen definierten Lokus der Zielzelle. Diese Technologie bietet damit die besten Voraussetzungen für einen quantitativen Vergleich verschiedener Transgenvarianten hinsichtlich ihrer Expressionsleistung, da physiologische und genetische Faktoren der Zielzelle weitestgehend identisch sind. So lassen sich im Gegensatz zu transienten Studien auch geringfügige Unterschiede in der Expressionshöhe einzelner Konstrukte reproduzierbar nachweisen. Die Bildung von Spontanresistenzen gegenüber dem eukaryontischen Selektionsmarker Hygromycin führte zur Etablierung stabiler *mock*-transfizierter Zelllinien, die als Negativkontrollen dienten.

C.5.1 Vergleich der CpG-freien GFP-Reportervariante mit dem Ausgangskonstrukt

In transienten Expressionsexperimenten wurde für die Reporterkonstrukte h_GFP 0 und h_GFP 60 eine positive Korrelation von CpG-Gehalt im ORF und gemessener Reporteraktivität festgestellt (siehe Abbildung C.10A). Eine quantitative Untersuchung der beiden Transgene hinsichtlich ihrer Antigenproduktion erfolgte anschließend in stabilen Expressionsanalysen unter Verwendung des Flp-In SystemsTM (siehe B.1.3.1). Dabei wurden zwei unterschiedliche Säugetierzelllinien verwendet, um eine fundierte Datengrundlage zu schaffen. Wie unter B.2.4 beschrieben, wurden stabil transfizierter CHO- sowie 293T-Zelllinien hergestellt und auf mehr als 99% positive Zellen selektioniert. In diesen Zellen befand sich je eine Kopie des Transgens h_GFP 60 bzw. h_GFP 0 in genetisch identischer Umgebung unter der Kontrolle des eukaryontischen CMV-Promotors. Unterschiede in der Expression der GFP-Reporterkonstrukte waren daher auf Veränderungen in der kodierenden Region zurückzuführen. Zunächst wurden polyklonale Ansätze der stabilen CHO- bzw. 293T-Zelllinien im Fluoreszenzmikroskop betrachtet. Analog zu den transienten Daten war bei beiden Zelltypen für das CpG-freie Konstrukt h_GFP 0 eine erhebliche Reduktion der GFP-vermittelten Fluoreszenz erkennbar (Abbildung C.16 b und d). In 293T-Zellen war für die

beiden Reporter-genvarianten zudem ein größerer Unterschied der Autofluoreszenz erkennbar als in CHO-Zellen.

Abbildung C.16: Fluoreszenzanalyse der stabil GFP exprimierenden CHO- sowie 293T-Zelllinien.



Wie unter B.2.4 beschrieben, wurden stabil h_GFP 60 (a und c) bzw. h_GFP 0 (b und d) exprimierende CHO- sowie 293T-Zellen etabliert. Nach Selektion wiesen beiden Zellpopulationen mehr als 99% positive, GFP exprimierenden Zellen auf. Aufnahmen der rekombinanten CHO- bzw. 293T-Zellen sind in Teilabbildung a und b bzw. c und d dargestellt. Für die Aufnahmen wurde eine Belichtungszeit von 2 sec und eine Blende von 2,8 gewählt.

Da für einen Abfall der GFP-Reporteraktivität auch die Instabilität des vermittelnden Proteins verantwortlich sein kann, wurden zusätzlich Western Blot-Analysen mit den polyklonalen Konstrukten h_GFP 60 und h_GFP 0 sowie je zwei monoklonalen Ansätzen (siehe unten) dieser Konstrukte durchgeführt (siehe Abbildung C.17). Der Proteinnachweis per Western Blot bestätigte die Daten der Fluoreszenzmikroskopieaufnahmen (siehe Abbildung C.16), wonach das CpG-haltige Transgen eine erhöhte Reporteraktivität zeigte. Für beide Genvarianten wurde in stabilen CHO-Zellen GFP (27 kDa) in seiner vollen Länge nachgewiesen, Unterschiede in der Prozessierung oder proteolytischen Degradation waren nicht detektierbar.

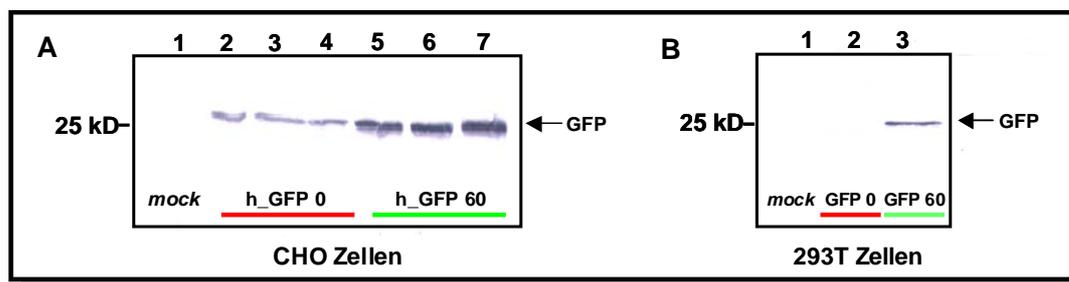


Abbildung C.17: Western Blot-Analyse der stabil GFP exprimierenden CHO- bzw. 293T-Zelllinien. Wie unter B.2.4 beschrieben wurden stabil h_GFP 60 bzw. h_GFP 0 exprimierende CHO- (A) sowie 293T-Zellen (B) etabliert. Nach Selektion wiesen beiden Zellpopulationen mehr als 99% positive, GFP exprimierenden Zellen auf. In Western Blot-Analysen wurde das *green fluorescent protein* anschließend mit dem polyklonalen *A.v. Peptide Antibody* aus Kaninchenserum nachgewiesen und über AP-Färbung detektiert. (A) Aufgetragen wurde Gesamtprotein *mock*-transfizierter Zellen (Spur 1), monoklonaler (Spur 2 und 3) sowie polyklonaler h_GFP 0 exprimierender Zellen (Spur 4) sowie monoklonaler (Spur 5 und 6) und polyklonaler (Spur 7) h_GFP 60 produzierender Zellen. (B) Aufgetragen wurde Gesamtprotein *mock*-transfizierter Zellen (Spur 1), polyklonaler h_GFP 0 exprimierender Zellen (Spur 2) sowie polyklonaler h_GFP 60 produzierender Zellen (Spur 3).

Zur Überprüfung der Funktionalität und Verlässlichkeit des Flp-In Systems wurden je 20 monoklonale CHO-Zellklone für die untersuchten Konstrukte etabliert und deren genomische DNA in Gradienten-PCR-Experimenten (siehe B.1.4.2) mit den Primern FRT_pf und FRT_pb (siehe F.1) analysiert. Zusätzlich wurde die Proteinexpression der Klone in FACS-Experimenten und im Fluoreszenzmikroskop untersucht. Dabei zeigten alle Klone den selben genetischen Hintergrund und wiesen identische Reporteraktivitäten auf (Daten nicht gezeigt). Diese Ergebnisse bestätigten die Anwendungsrelevanz und die hohe Reproduzierbarkeit dieses Rekombinationssystems. Daher wurde das System nachfolgend für quantitative Untersuchungen der Expressionsleistung verwendet.

C.5.2 Langzeitexpressionsanalyse der GFP-Reporterkonstrukte

In der Literatur wird beschrieben, dass eine Methylierung von CpG-Dinukleotiden einen reprimierenden Effekt auf die Genexpression in Vertebraten hat und zu einem Abschalten der Genexpression führen kann (Bestor, 2000; Hisano *et al.*, 2003). Analog dazu sollten CpG-freie Konstrukte länger und besser exprimiert werden (Chevalier-Mariette *et al.*, 2003). In stabilen Langzeitexpressionsanalysen wurde diese Fragestellung anhand der Reporterkonstrukte h_GFP 60 und h_GFP 0 adressiert. Nach Etablierung von rekombinanten 293T- bzw. CHO-Zelllinien (siehe C.5.1) wurde die Expression monoklonaler sowie polyklonaler Reporterkonstrukte über einen Zeitraum von 16 Monaten durch regelmäßige Messung der GFP-vermittelten grünen Autofluoreszenz bestimmt. Die Angaben der mittleren Fluoreszenzintensitäten sind in Abbildung C.18A (293T-Zellen) und C.18B (CHO-Zellen) zusammengefasst. Die h_GFP 60 Expression konnte über den gesamten Messzeitraum in beiden Zelllinien relativ konstant mit einer mittleren Fluoreszenzintensität von 800 (293T) bzw. 700 (CHO) nachgewiesen werden. Das CpG-reduzierte h_GFP 0 Konstrukt zeigte ebenfalls eine über den gesamten Zeitraum nachweisbare Reporteraktivität. Die mittlere Fluoreszenzintensität war jedoch verglichen mit dem h_GFP 60 Konstrukt um den Faktor 10-20 (293T) bzw. 6-9 (CHO) verringert. Diese Ergebnisse stimmen mit den transienten (siehe C.4.1) sowie den unter C.5.1 generierten Daten überein. Auch bei den Fluoreszenzmikroskopieaufnahmen (C.5.1, Abbildung C.16) war für 293T Zellen ein größerer Unterschied in der GFP-Autofluoreszenz der Konstrukte zu erkennen. Somit resultierte eine Eliminierung der in der kodierenden Region lokalisierten CpG-Dinukleotide in dieser Arbeit nicht in einer erhöhten oder dauerhafteren Proteinexpression, sondern hatte eine Reduktion der Reporteraktivität zur Folge. Welche Auswirkung darüber hinaus eine CpG-Reduktion des Vektorrückgrates auf die Transgenexpression bewirkt, wurde nachfolgend im stabilen Expressionssystem untersucht.

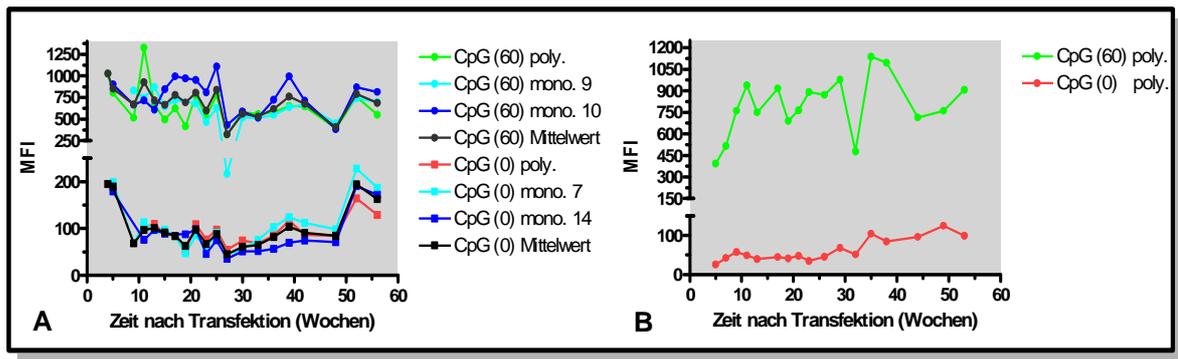


Abbildung C.18: Langzeitexpressionsanalyse der stabil GFP exprimierenden CHO- bzw. 293T-Zellen. Wie unter B.2.4 beschrieben wurden stabil h_GFP 60 bzw. h_GFP 0 exprimierende CHO- sowie 293T-Zellen etabliert. Zusätzlich zu den polyklonalen CHO-Zellpopulationen wurden mehrere monoklonale CHO-Zelllinien generiert. Die h_GFP 60 bzw. h_GFP 0 Expression der polyklonalen sowie zweier monoklonaler Zelllinien wurde über einen Zeitraum von 15 Monaten durch regelmäßige Messung der GFP-vermittelten Autofluoreszenz in einem Durchflusszytometer bestimmt. Die Angaben der mittleren Fluoreszenzintensität (MFI) für die jeweiligen Konstrukte sind in Abbildung A (CHO-Zellen) und B (293T-Zellen) zusammengefasst. Eine Zuordnung der Konstrukte kann der Legende entnommen werden. (A) Neben polyklonalen h_GFP 60 bzw. h_GFP 0 CHO-Zellklonen (grüner bzw. roter Kurvenverlauf) wurden auch je zwei Zelllinien monoklonalen Ursprungs (türkiser bzw. blauer Kurvenverlauf) exemplarisch im Langzeitexperiment untersucht. (B) Dargestellt ist der Verlauf der h_GFP 60 (grüner Kurvenverlauf) bzw. h_GFP 0 (roter Kurvenverlauf) Langzeitexpression in 293T-Zellen.

C.5.3 Etablierung eines CpG-reduzierten modularen Vektorrückgrates für stabile Expressionsanalysen

Zur Klärung der Frage, ob im Vektorrückgrat befindliche CpG-Dinukleotide einen Einfluss auf die stabile Expressionsleistung eines Vektors ausüben, wurden ausgehend von dem käuflichen Vektor pc5 mehrere CpG-reduzierte Plasmide generiert (siehe C.2.2.2). Der Vektor SYN_84, der eine CpG-freie Version der Hygromycinresistenzkassette beinhaltet, das Plasmid SYN_187, bei dem nur der für die Resistenz kodierende Bereich CpG-depletiert ist, sowie der Vektor SYN_226, dessen Hygromycinresistenzkassette analog zum pc5 ist, wurden zunächst in stabilen Expressionsanalysen auf ihre Funktionalität untersucht. Als Reporterkonstrukt diente dabei die CpG-freie Version des Reportergens h_GFP 0. Nach Selektion und Etablierung stabiler CHO-Zelllinien (siehe B.2.4) wurde die Reporteraktivität der einzelnen Vektorstrukture im Durchflusszytometer gemessen. Als signifikant positive GFP Expression wurde dabei eine MFI von mehr als 10 gewertet. Wie aus der Darstellung C.19A hervorgeht, vermittelten die Plasmide SYN_85 und SYN_187 keine GFP-Expression, obwohl sie sich in transienten Experimenten als funktionell erwiesen (Daten nicht gezeigt). Das Konstrukt SYN_226 hingegen erreichte eine GFP-Expression von 142 MFI. Verglichen mit dem Ausgangsvektor pc5 (MFI: 93) führte die Verwendung von SYN_226 zu einer Expressionssteigerung von ca. 50% (siehe Abbildung C.19B).

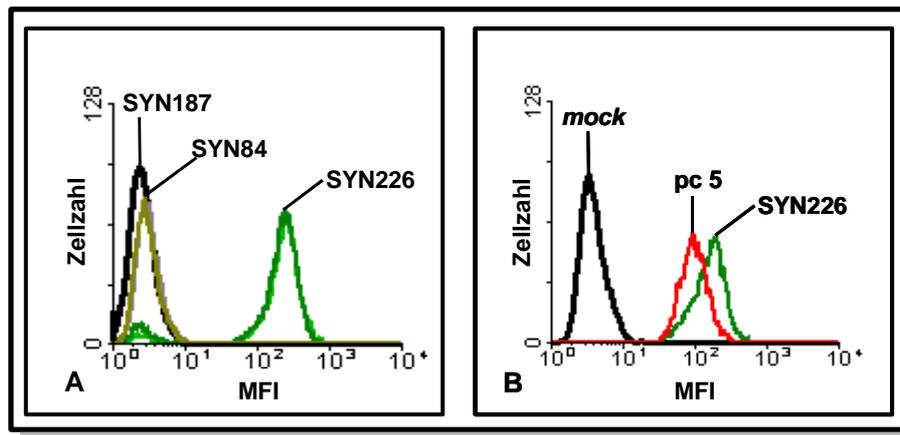
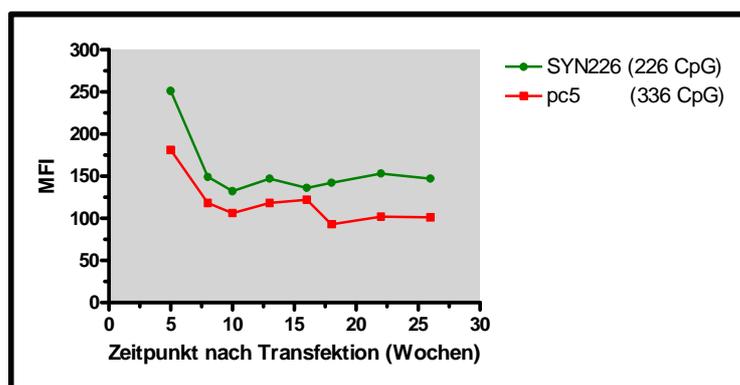


Abbildung C.19: Vergleich verschiedener CpG-reduzierter Vektorrückgrate hinsichtlich ihrer h_GFP 0 Expression im stabilen Expressionssystem. Für stabil h_GFP 0 exprimierende CHO-Zellen wurde nach 4-wöchiger Selektion die für das jeweilige Vektorrückgrat erreichte h_GFP 0-Reporteraktivität detektiert. (A) Dargestellt ist die h_GFP 0 Expression des Vektors SYN84 (olivgrüne Kurve), des Vektors SYN187 (schwarze Kurve) sowie des Vektors SYN226 (hellgrüne und dunkelgrüne Kurve). (B) Gezeigt wird die h_GFP 0 Expressionsleistung des Plasmides SYN226 (dunkelgrüne Kurve) im Vergleich zum Vektor pc5 (rote Kurve). Mock-transfizierte Zellen, welche spontane Hygromycinresistenz entwickelten, dienten als Negativkontrolle (schwarze Kurve). Auf der y-Achse ist die Zellzahl, auf der x-Achse die für das jeweilige Konstrukt ermittelte MFI dargestellt.

Um zu untersuchen, ob das CpG-reduzierte Vektorrückgrat auch über längere Zeit hinweg zu einer verbesserten Transgenexpression führt, wurden die Konstrukte SYN_226 sowie pc5 über einen Zeitraum von 26 Wochen hinsichtlich ihrer h_GFP 0 Reporteraktivität untersucht. Über den gesamten Zeitraum konnte für SYN_226 die GFP-vermittelte Reporteraktivität mit einer durchschnittlichen MFI von 156 gemessen werden, während sich für den pc5 ein MFI-Mittelwert von 118 ergab (siehe Abbildung C.20). Dies entspricht einer um 32% erhöhten Expression gegenüber dem käuflichen Vektor pc5. In Übereinstimmung mit Daten aus transienten Studien resultierte auch im stabilen Expressionssystem die Verwendung eines CpG-reduzierten Vektors in einer Expressionssteigerung im Bezug auf das Ausgangskonstrukt. Um den für Vektorrückgrat und Transgen auf Proteinebene beobachteten CpG-Effekt detaillierter zu untersuchen, wurden weitere CpG-modifizierte Vektorelemente generiert und hinsichtlich ihrer Funktionalität und Expressionsleistung charakterisiert.

Abbildung C.20: GFP Langzeitexpression des Vektors SYN226 (226 CpG) und des Plasmides pcDNA5/FRT (337 CpG).



Unter Verwendung der im Text beschriebenen Plasmide wurden stabile h_GFP 0 exprimierende CHO-Zellen etabliert. Die mehr als 99% positiven, GFP exprimierenden Zellen wurden über einen Zeitraum von 26 Wochen hinsichtlich ihrer Transgenexpression (MFI) im Durchflusszytometer untersucht. Eine Zuordnung der Konstrukte kann der Legende entnommen werden.

C.5.4 Funktionalität weiterer CpG-modifizierter Vektorelemente

Da der CpG-Effekt sowohl in CHO-Zellen als auch in 293T-Zellen nachgewiesen wurde (siehe Abbildung C.16, C.17 und C.18), wurden zur Untersuchung weiterer CpG-freier Vektorkomponenten ausschließlich CHO-Zellen verwendet, welche schnell und effizient transfiziert werden konnten.

C.5.4.1 Expression einer CpG-maximierten GFP-Variante

Zur Erweiterung des existierenden Konstruktpanels h_GFP 0 und h_GFP 60 wurde eine CpG-maximierte Version h_GFP 93 in stabile Expressionsexperimente eingesetzt. In rekombinanten CHO-Zellen wurden die drei GFP-Konstrukte schließlich hinsichtlich ihrer resultierenden Reporteraktivität im Fluoreszenzmikroskop (Abbildung C.21) bzw. im Durchflusszytometer (Abbildung C.22) analysiert. Analog zu den unter C.5.1 generierten Daten war im FM für die Konstrukte h_GFP 0 und h_GFP 60 eine Korrelation von CpG-vermittelter Autofluoreszenz und CpG-Gehalt im ORF erkennbar. Auch in stabilen Zellen führte die Verwendung von h_GFP 93 als Transgen jedoch verglichen mit dem Ausgangskonstrukt nicht zu einer weiteren Expressionssteigerung (siehe Abbildung C.21).

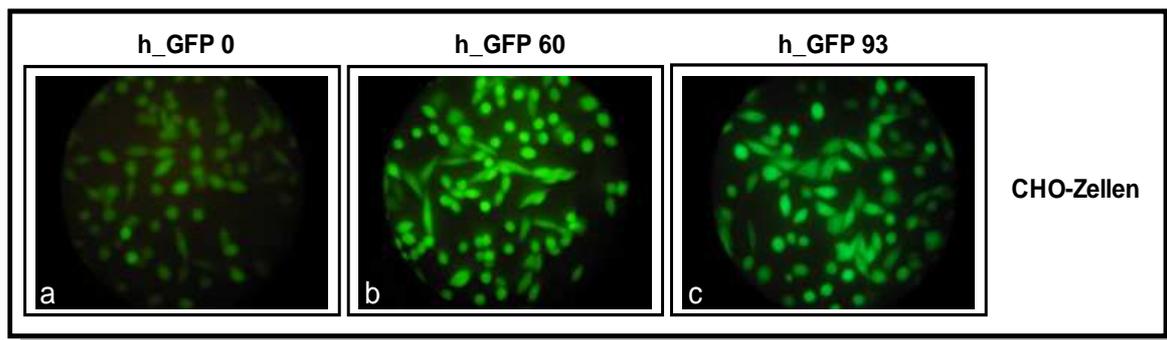


Abbildung C.21: Fluoreszenzanalyse der GFP Expression verschiedener CpG-modifizierter Reporterkonstrukte in stabilen CHO-Zelllinien. Wie unter B.2.4 beschrieben wurden stabil h_GFP 0 (a), h_GFP 60 (b) bzw. h_GFP 93 (c) exprimierende CHO-Zelllinien etabliert. Nach Selektion wiesen alle Zellpopulationen mehr als 99% positive, GFP exprimierende Zellen auf. Für die Fluoreszenzmikroskopieaufnahmen wurde eine Belichtungszeit von 0,1 sec und eine Blende von 8,0 gewählt.

Diesen Zusammenhang spiegelten auch die Daten der FACS-Analyse wieder (siehe Abbildung C.22). Für das 60 CpGs enthaltende Transgen (MFI: 745) war dabei eine 8-fache Expressionssteigerung im Vergleich zur CpG-freien Version (MFI: 93) feststellbar. Bezogen auf h_GFP 60 wies das Konstrukt h_GFP 93 (MFI: 494) dagegen nur zwei Drittel der Reporteraktivität auf. Diese reproduzierbar generierten Daten stimmen mit den unter C.4 erhaltenen Ergebnissen aus transienten Expressionsstudien überein und implizieren einen generellen Mechanismus der CpG-basierten Modulation der Transgenexpression. Da auch für den CpG-Gehalt im Vektorrückgrat ein Einfluss auf die Antigenproduktion gezeigt werden

konnte (siehe C.5.3), sollten auch regulatorische Vektorkomponenten wie z.B. das Polyadenylierungssignal hinsichtlich ihres CpG-Gehaltes minimiert werden.

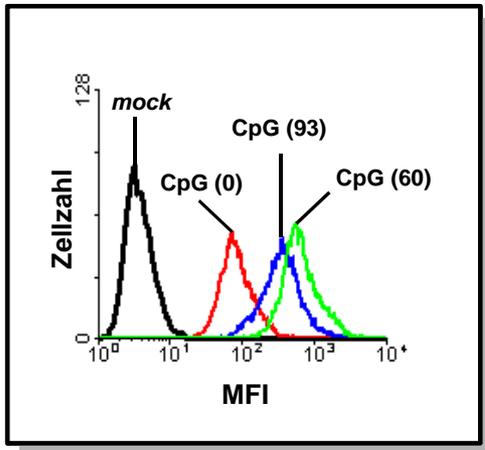
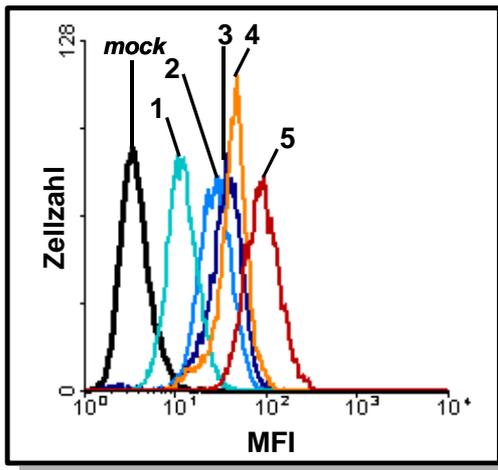


Abbildung C.22: Vergleich der GFP Expression CpG-modifizierter Reporterkonstrukte im stabilen Expressionssystem (FACS-Daten). Wie unter B.2.4 beschrieben wurden stabil h_GFP 0 (roter Graph), h_GFP 60 (grüner Graph) bzw. h_GFP 93 (blauer Graph) exprimierende CHO-Zelllinien etabliert. Anschließend wurde für das jeweilige Transgen die h_GFP 0-Reporteraktivität im FACS gemessen. Als Negativkontrolle wurden *mock*-transfizierte Zellen (schwarzer Graph) verwendet, welche spontane Hygromycinresistenz entwickelten. Auf der y-Achse ist die Zellzahl, auf der x-Achse die für das jeweilige Konstrukt ermittelte MFI dargestellt.

C.5.4.2 Expressionsstärke verschiedener CpG-freier BGH P(A)-Signale im pc5

Zwei in transienten Experimenten auf ihre Funktionsfähigkeit getestete CpG-freie P(A)-Signalvarianten BGH_AGA und BGH_ATA (siehe Abbildung C.14) wurden in stabilen Expressionsstudien hinsichtlich ihrer h_GFP 0 Expression untersucht. Ziel war es dabei, die modifizierten BGH P(A)-Varianten sowie das herkömmliche SV40-Polyadenylierungssignal hinsichtlich ihrer resultierenden Transgenexpression zu vergleichen. Aus diesem Grund wurden BGH_AGA und BGH_ATA sowie das SV40-Signal wie unter C.2.2.2.2 beschrieben in den Vektor pc5 kloniert, der h_GFP 0 als Transgen beinhaltet. Als Positivkontrolle diente dabei der Vektor pc5/h_GFP 0, der das unveränderte BGH-Polyadenylierungssignal enthält. Mit den beschriebenen Konstrukten wurden rekombinante CHO-Zelllinien generiert, deren MFI als Maß für die Genexpression im Durchflusszytometer bestimmt wurde. Wie aus Abbildung C.23 hervorgeht, exprimierte die Variante BGH_ATA (Verteilung 3) mit einer MFI von 94 das Transgen mindestens so gut wie die Positivkontrolle (MFI: 93, Verteilung 4). Eine Verwendung des BGH_AGA-Konstruktes (MFI: 71, Verteilung 2) hingegen, resultierte in einer Verringerung der MFI auf 76%. Die vom SV40 P(A)-Signal (MFI: 38, Verteilung 1) hervorgerufene Reporteraktivität betrug nur noch 41% der Positivkontrolle. Analog zu den unter C.5.3 gewonnenen Daten zeigte das herkömmliche BGH-Polyadenylierungssignal im Kontext des CpG-reduzierten Vektors SYN_226 (MFI: 142, Verteilung 5) eine Expressionssteigerung von 50% bezogen auf die Positivkontrolle. Vergleichbare Ergebnisse wurden unter Verwendung von h_GFP 93 als Transgen erzielt (Daten nicht gezeigt). Das Konstrukt BGH_ATA kann daher in die nächste Generation CpG-reduzierter Vektoren integriert werden. Zur weiteren Minimierung des CpG-Gehaltes im Vektorrückgrat wurden darüber hinaus CpG-freie CMV-Promotorkonstrukte etabliert.

Abbildung C.23: Einfluss verschiedener CpG-freier Polyadenylierungssignale auf die h_GFP 0 Expression in stabilen CHO-Zelllinien (FACS-Daten).

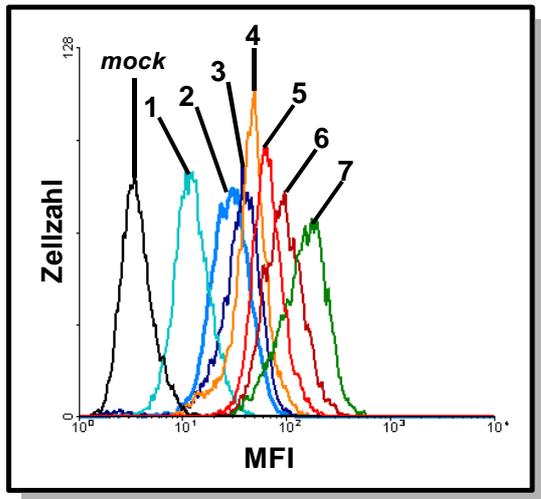


Die im Rahmen dieser Arbeit generierten CpG-freien Polyadenylierungssignale BGH_AGA (2, hellblaue Verteilung) bzw. BGH_ATA (3, dunkelblaue Verteilung), sowie das SV40 P(A)-Signal (1, türkise Verteilung) und das herkömmliche BGH P(A)-Signal (4, orange Verteilung) wurden im Sequenzkontext des Vektors pc5 hinsichtlich ihrer Expressionsleistung verglichen. Zudem wurde die MFI des konventionellen BGH P(A)-Signals im Plasmid SYN226 (5, rote Verteilung) bestimmt. Mock-transfizierte Zellen, welche spontane Hygromycinresistenz entwickelten, dienen als Negativkontrolle (schwarze Verteilung). Auf der y-Achse ist die Zellzahl, auf der x-Achse die für das jeweilige Konstrukt ermittelte MFI dargestellt.

C.5.4.3 Expressionsstärke verschiedener CpG-freier CMV-Promotoren im SYN226

In transienten Expressionsstudien wurden sechs funktionelle CpG-freie CMV-Promotorvarianten etabliert (siehe C.4.3.3). Die Untersuchung dieser Konstrukte CMV_A bis CMV_F in stabilen Expressionsanalysen sollte eine quantitative Analyse ihrer Promotorstärken liefern sowie einen Vergleich mit dem CpG-freien CMV_TG-Promotor (Yew *et al.*, 2002) ermöglichen (siehe C.2.1.1.5). Als Positivkontrollen dienten der in eukaryontischen Expressionsvektoren häufig verwendete CMV-Promotor sowie der SV40-Promotor. Die CpG-depletierten Promotorvarianten CMV_A bis CMV_F sowie der CMV_TG-Promotor und das SV40-Konstrukt wurden wie unter C.2.2.2.6 beschrieben in den Vektor SYN226 kloniert, der h_GFP 0 als Transgen enthält. Im Durchflusszytometer wurde nach Selektion und Etablierung stabiler CHO-Zelllinien anhand der GFP-Reporteraktivität die Promotorstärke des jeweiligen Konstruktes festgestellt. Abbildung C.24 gibt eine Übersicht über die von den einzelnen Promotoren hervorgerufenen, mittleren Fluoreszenzintensitäten (MFIs) als Maß für die Promotorstärke. So ermittelt sich für das Konstrukt CMV_F (Verteilung 1) eine MFI von 13, was bezogen auf den herkömmlichen CMV-Promotor (MFI: 93, Verteilung 6) einer Expression von 14% entspricht. Die Varianten CMV_D (MFI: 29, Verteilung 2) und CMV_E (MFI: 37, Verteilung 3) erreichen 31% bzw. 40% der Reporteraktivität verglichen mit dem unveränderten CMV-Konstrukt. Stärker als die CpG-freien CMV-Varianten CMV_A bis CMV_C exprimiert der SV40-Promotor (MFI: 44, Verteilung 4), dessen Expression 47% des CMV-Promotors beträgt sowie der CMV_TG-Promotor (74, Verteilung 5) der 80% Reporteraktivität aufweist. Wie bereits unter C.5.4.2 erwähnt, vermittelt der herkömmliche CMV-Promotor im Kontext des CpG-reduzierten Vektors SYN_226 (MFI: 142, Verteilung 7) eine 50%-ige Expressionssteigerung bezogen auf die Positivkontrolle (MFI: 93; Verteilung 6). Zusammenfassend lässt sich feststellen, dass die CMV-Variante CMV_E, die in vergleichbarem Maße exprimiert wie der SV40-Promotor, in die nächste Generation CpG-reduzierter Plasmide integriert werden kann. Allerdings erreicht das CMV_E-Promotorkonstrukt nicht die Expressionsstärke des herkömmlichen CMV-Promotors.

Abbildung C.24: Vergleich verschiedener CpG-freier Promotorvarianten hinsichtlich ihrer h_GFP 0 Expression in stabilen CHO-Zelllinien (FACS-Daten). Im Vektor SYN226 wurde der konventionelle CMV-Promotor durch CpG-freie Promotorvarianten ersetzt bzw. durch den SV40-Promotor ausgetauscht. Die Promotorstärken der somit generierten Konstrukte CMV_F (1, türkiser Verteilung), CMV_D (2, hellblaue Verteilung), CMV_E (3, dunkelblaue Verteilung) sowie des SV40-Promotorkonstruktes (4, orange Verteilung) und eines in der Literatur beschriebenen CMV_TG-Promotors (5, hellrote Verteilung) wurden mit der des herkömmlichen CMV-Promotors im Kontext des Vektors pc5 (6, dunkelrote Verteilung) bzw. SYN226 (7, grüne Verteilung) verglichen. *Mock*-transfizierte Zellen, welche spontane Hygromycin-resistenz entwickelten, dienten als Negativkontrolle (schwarze Verteilung). Auf der y-Achse ist die Zellzahl, auf der x-Achse die für das jeweilige Konstrukt ermittelte MFI dargestellt.



Die Konstrukte CMV_A, CMV_B und CMV_C wiesen reproduzierbar keinerlei GFP-Autofluoreszenz auf, was ein unerwartetes Ergebnis ist, da sie sich in transienten Experimenten funktionell erwiesen (siehe Abbildung C.15). Möglicherweise exprimieren diese Konstrukte das Transgen so schwach, dass die GFP-vermittelte Autofluoreszenz nicht signifikant von der Eigenfluoreszenz *mock*-transfizierter Zellen, die als Negativkontrolle dienen, unterschieden werden kann.

Die Konstrukte CMV_A, CMV_B und CMV_C wiesen reproduzierbar keinerlei GFP-Autofluoreszenz auf, was ein unerwartetes Ergebnis ist, da sie sich in transienten Experimenten funktionell erwiesen (siehe Abbildung C.15). Möglicherweise exprimieren diese Konstrukte das Transgen so schwach, dass die GFP-vermittelte Autofluoreszenz nicht signifikant von der Eigenfluoreszenz *mock*-transfizierter Zellen, die als Negativkontrolle dienen, unterschieden werden kann.

C.5.5 Zusammenfassung

Analog zu den p24-Reportergen (siehe C.4.1) war für die Konstrukte h_GFP 0 und h_GFP 60 im stabilen Expressionssystem eine positive Korrelation von CpG-Gehalt und gemessener Reporteraktivität feststellbar (siehe C.5.1). Zudem führte eine Maximierung der CpG-Zahl bei dem Konstrukt h_GFP 93 nicht zu einer weiteren Expressionssteigerung, sondern resultierte bezogen auf das Ausgangskonstrukt h_GFP 60 in einer 33%-igen Verringerung der Reporteraktivität (siehe C.5.4.1). Diese reproduzierbar generierten Daten stimmen mit den in Kapitel C.4 erhaltenen Ergebnissen aus transienten Expressionsstudien überein und implizieren eine Beteiligung der CpG-Dinukleotide an der Regulation der Genexpression. Darüber hinaus wurden die CpG-freien Vektorrückgrate SYN84, SYN187 sowie SYN226 in stabilen Expressionsanalysen auf ihre Funktionsfähigkeit und Expressionshöhe getestet. Dabei ermöglichte nur das Plasmid SYN226 die Etablierung stabiler Zelllinien. Zudem wies der CpG-reduzierte Vektor SYN226 in stabilen Zellkulturexperimenten verglichen mit dem Referenzkonstrukt pc5 eine um 32% erhöhte Transgenproduktion auf. Analog zu transienten Experimenten bestätigte diese Beobachtung die Tendenz, dass CpG-minimierte Plasmide eine verbesserte Antigenproduktion aufweisen. Auch in einer Langzeitstudie über 26 Wochen zeigte sich dieser Trend.

Zur weiteren CpG-Reduzierung funktioneller Vektorelemente wurden CpG-freie Varianten des BGH P(A)-Signals sowie des CMV-Promotors hinsichtlich ihrer Funktion und Expressionshöhe quantitativ untersucht. Dabei wurde die Variante BGH_ATA etabliert, die eine mit dem BGH-Polyadenylierungssignal vergleichbare Expression zeigte. Zudem wurde eine CpG-freie CMV-Promotorvariante CMV_E etabliert, die eine geringere Expression aufwies als der herkömmliche CMV-Promotor, aber in ihrer Expressionsstärke dem SV40-Promotor vergleichbar war.

Für verschiedene Transgene konnte, unabhängig vom verwendeten Vektorsystem und der exprimierenden Zelllinie, eine positive Korrelation von CpG-Zahl und resultierender Reporterexpression festgestellt werden. Um zu untersuchen, ob dieser auf Proteinebene beobachtete Effekt auch auf transkriptioneller Ebene auftritt, wurden die drei GFP-Reporterkonstrukte h_GFP 0, h_GFP 60 und h_GFP 93 in RNA-Studien detaillierter untersucht.

C.6 Untersuchung der GFP-Reportervarianten auf RNA-Ebene

Der in transienten und stabilen Experimenten beobachtete Abfall GFP-vermittelter Reporteraktivität kann auf RNA-Ebene durch verschiedene Faktoren wie z.B. reduzierten Kernexport, Instabilität CpG-freier RNA-Transkripte oder geringere Transkriptionsraten bedingt sein. Um diese Fragestellung zu adressieren, wurden zunächst die Konstrukte h_GFP 60 und h_GFP 0 in rekombinanten CHO- sowie 293T-Zelllinien auf transkriptioneller Ebene analysiert.

C.6.1 Relative Quantifizierung der zellulären RNA (CHO und 293T Zellen)

C.6.1.1 CHO-Zellen

Für eine Aufklärung der Transkriptionsaktivität wurde zunächst wie in B.4.6.3.1 beschrieben eine quantitative *realtime*-PCR für die stabil GFP exprimierenden CHO- sowie 293T-Zellen (siehe C.5.1) durchgeführt. Aus stabilen CHO- bzw. 293T-Zellen wurde dazu cytoplasmatische sowie nukleäre RNA präpariert (siehe B.4.1), welche nach Beseitigung von DNA-Kontaminationen (siehe B.4.3) als Matrize in eine Reverse Transkription (siehe B.4.4) mit willkürlich bindenden Primern eingesetzt wurde. Die resultierenden cDNA-Proben für das CpG-freie Konstrukt und sein Pendant wurden verdünnt und in einer anschließenden *realtime*-PCR mit den Oligonukleotiden GFPint142_pf und GFPint516_pb (siehe F.1.3) relativ zueinander quantifiziert. Als interne Kontrolle wurde die RNA-Menge des ebenfalls in das Zellgenom integrierten Hygromycinresistenzgens anhand der Primer hygrint385_pf und hygrint754_pb (siehe F.1.3) quantitativ bestimmt.

Die Ergebnisse der relativen Quantifizierung sind in Abbildung C.25 für die cytoplasmatische bzw. die nukleäre RNA-Fraktion aus CHO-Zellen dargestellt. Wie aus Abbildung C.25A bzw. C.25C hervorgeht, zeigten die gemessenen Konstrukte keinen Unterschied hinsichtlich ihrer Menge cytoplasmatischer bzw. nukleärer Hygromycin RNA-Transkripte. Die Zahlen der cytoplasmatischen (Abbildung C.25B) bzw. nukleären (Abbildung C.25D) GFP-spezifischen RNA-Transkripte hingegen korrelierten sehr gut mit den Daten der transienten und stabilen Proteinexpression. So war für das CpG-depletierte Konstrukt nach relativer Quantifizierung im *Lightcycler* im Vergleich zum Ausgangskonstrukt eine etwa 7-fach geringere cytoplasmatische sowie eine etwa 13-fach geringere nukleäre RNA-Menge nachzuweisen. Zudem kann der Schmelzpunktcurve entnommen werden, dass bei beiden rekombinanten Zelllinien nur ein Hygromycin-spezifisches PCR-Produkt detektiert wurde (siehe Teilabbildung C.25a). Den beiden GFP-Konstrukten h_GFP 0 und h_GFP 60 hingegen liegt eine unterschiedliche Nukleinsäuresequenz zugrunde. Daher besitzen die resultierenden PCR-Produkte unterschiedliche spezifische Schmelztemperaturen wie der Schmelzpunktanalyse entnommen werden kann (siehe Teilabbildung C25b).

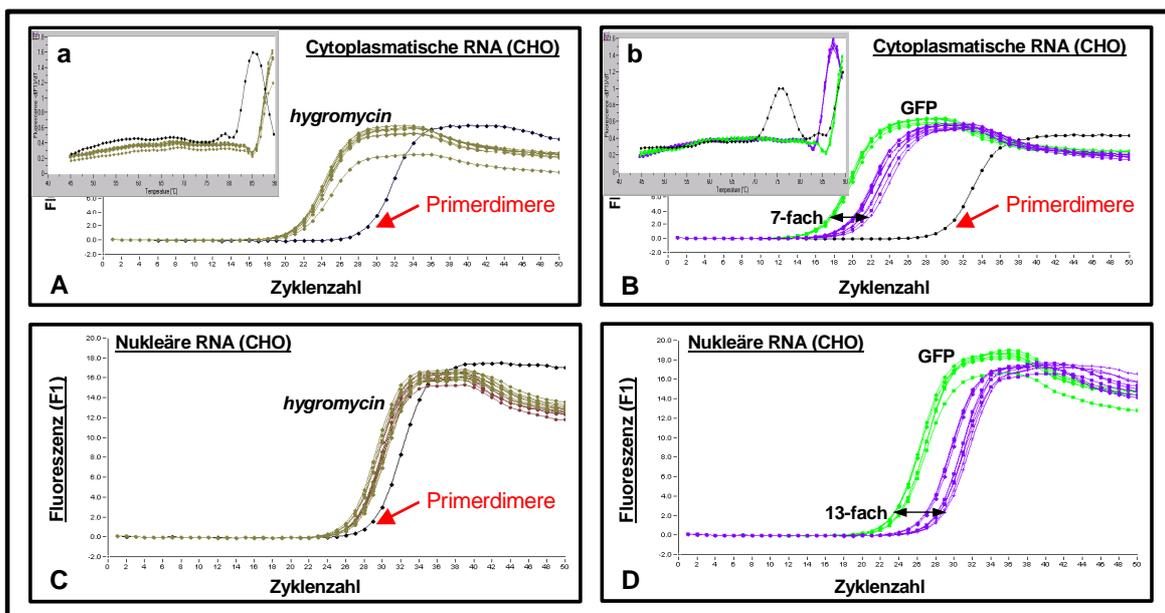


Abbildung C.25: Relative Quantifizierung der cytoplasmatischen sowie nukleären RNA-Transkripte von CHO-Zellen via *realtime*-PCR. Aus stabil h_GFP 60 bzw. h_GFP 0 exprimierenden CHO-Zellen wurde die cytoplasmatische sowie die nukleäre RNA präpariert und in LC-Analysen quantifiziert. Die für eine signifikante Detektion des Fluoreszenzsignals (y-Achse) nötige Zyklenzahl (x-Achse) stellt ein indirektes Maß für die Menge zugrunde liegender RNA-Transkripte dar. (A) Relative Quantifizierung der cytoplasmatischen Hygromycin Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (olivgrüne Kurven). (B) Relative Quantifizierung der cytoplasmatischen GFP-Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (hellgrüne bzw. violette Kurven). (C) Relative Quantifizierung der nukleären Hygromycin-Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (olivgrüne Kurven). (D) Relative Quantifizierung der nukleären GFP-Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (hellgrüne bzw. violette Kurven). Exemplarisch sind in den Abbildungen Zyklenzahl und Schmelztemperatur der gebildeten Primerdimere (schwarze Kurve) verzeichnet. In den Teilabbildungen a bzw. b ist für das Hygromycin-basierte PCR-Produkt (olivegrüne Kurven) bzw. für die GFP-spezifischen PCR-Produkte (h_GFP 0: violette Kurven; h_GFP 60: hellgrüne Kurven) der Verlauf einer Schmelzpunktanalyse gezeigt.

C.6.1.2 293T-Zellen

Da für CHO-Zellen sowohl in der cytoplasmatischen als auch in der nukleären Fraktion eine Korrelation von CpG-Gehalt und detektierter RNA-Menge festzustellen war, wurde für 293T-Zellen ausschließlich das cytoplasmatische RNA-Niveau der Reporterkonstrukte quantitativ analysiert. Dabei wurde analog, wie unter C.6.1.1 für die relative Quantifizierung cytoplasmatischer RNA-Transkripte von CHO-Zellen beschrieben, vorgegangen. Als interne Kontrolle wurde wiederum die cDNA-Kopienzahl des im Zellgenom lokalisierten Hygromycinresistenzgens quantitativ bestimmt. Die Resultate dieser Untersuchung sind in Abbildung C.26 dargestellt. Gemäß Abbildung C.26A wiesen alle untersuchten Konstrukte einen vergleichbaren Hygromycin RNA-Status auf. Bei den untersuchten GFP-Konstrukten h_GFP 0 und h_GFP 60 hingegen war ein 29-facher Unterschied der RNA-Menge zu detektieren (siehe Abbildung C.26B). Auch bei 293T-Zellen stimmten die für CpG-freies Transgen und Ausgangskonstrukt generierten Daten mit den Ergebnissen der Proteinexpression überein.

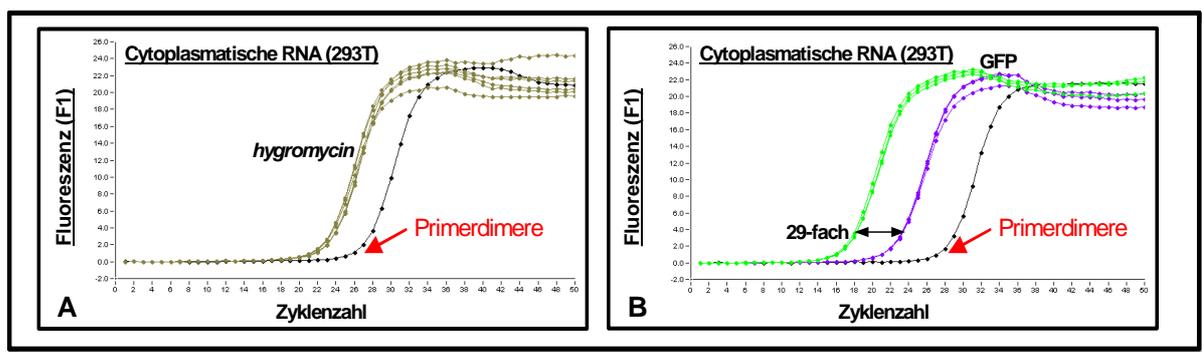


Abbildung C.26: Relative Quantifizierung der cytoplasmatischen RNA aus 293T-Zellen via *realtime*-PCR. Aus stabil h_GFP 60 bzw. h_GFP 0 exprimierenden 293T-Zellen wurde die cytoplasmatische RNA präpariert und in LC-Analysen quantifiziert. Die für eine signifikante Detektion des Fluoreszenzsignals (y-Achse) nötige Zyklenzahl (x-Achse) stellt ein indirektes Maß für die Menge zugrunde liegender RNA-Transkripte dar. (A) Relative Quantifizierung der cytoplasmatischen Hygromycin-Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (olivgrüne Kurven). (B) Relative Quantifizierung der cytoplasmatischen GFP-Transkripte für h_GFP 60 bzw. h_GFP 0 exprimierende Zelllinien (hellgrüne bzw. violette Kurven). Zudem kann beiden Teilabbildungen entnommen werden, nach welcher Zyklenzahl die Fluoreszenz gebildeter Primerdimere signifikant nachweisbar ist (schwarze Kurve).

Alle im *Lightcycler* erhaltenen PCR-Produkte wurden zudem auf einem Agarosegel hinsichtlich ihrer Größe analysiert (Daten nicht gezeigt) und ihre Sequenz verifiziert. Zusätzlich wurde die spezifische Schmelztemperatur für die jeweiligen PCR-Produkte sowie für die Oligonukleotiddimere bestimmt. Wie in Abbildung C.25A und C.25B zu erkennen ist, unterscheiden sich die generierten Primerdimere sowie die jeweiligen spezifisch gebildeten PCR-Produkte deutlich hinsichtlich ihrer Schmelztemperatur (siehe Abbildung C.25). Nach relativer Quantifizierung der transgenspezifischen RNA-Transkripte für h_GFP 60 und h_GFP 0 sollte für die CpG-modifizierten GFP-Konstrukte auch die absolute RNA-Menge bestimmt werden.

C.6.2 Absolute Quantifizierung zellulärer RNA in CHO Zellen

C.6.2.1 Cytoplasmatische und nukleäre RNA

In rekombinanten CHO- sowie 293T-Zellen wurde der auf Proteinebene beobachtete CpG-Effekt auch auf RNA-Ebene detektiert. Für beide Zelltypen korrelierte die relativ quantifizierte, transgenspezifische RNA-Menge mit der CpG-Zahl im Leseraster des Transgens. Für weiterführende Experimente wurden daher ausschließlich rekombinante CHO-Zelllinien verwendet, welche sich durch schnelleres Wachstum und einfachere Handhabung auszeichneten. Zudem wurde bei der absoluten Quantifizierung der RNA-Transkripte das bestehende Konstruktpanell um die CpG-maximierte Version h_GFP 93 (siehe C.5.4.1) erweitert, welche mit den Oligonukleotiden GFPmax165_pf und GFPmax520_pb (siehe F.1.3) analysiert wurde. Nach Etablierung externer Standardgeraden, wie unter B.4.6.3.2 beschrieben, wurden die GFP-spezifischen RNA-Transkripte in absoluten *Lightcycler*-Experimenten quantifiziert. Dabei wurde analog zur relativen Quantifizierung (siehe C.6.1.1) vorgegangen, als Maß für die vorliegende RNA-Menge wurde jedoch die absolute cDNA-Kopienzahl detektiert. Zudem wurde als interne Kontrolle die absolute Zahl der Hygromycinresistenz-vermittelnden Transkripte nachgewiesen. Eine Übersicht über die cytoplasmatischen (Abbildung C.27A) bzw. nukleären, transgenspezifischen cDNA-Kopienzahlen (Abbildung C.27B), welche auf die zugrunde liegenden RNA-Mengen schließen lassen, gibt Abbildung C.27.

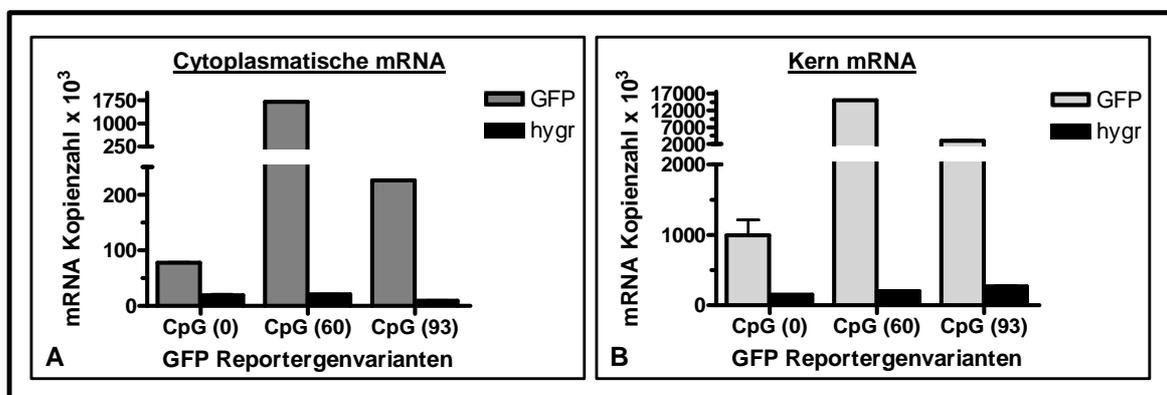


Abbildung C.27: Absolute Quantifizierung der cytoplasmatischen sowie nukleären RNA-Transkripte von CHO-Zellen via *realtime*-PCR. Aus stabil h_GFP 60 bzw. h_GFP 0 exprimierenden CHO-Zellen wurde die cytoplasmatische sowie die nukleäre RNA präpariert und in LC-Analysen quantifiziert. Die für das jeweilige Konstrukt, unter Verwendung externer Standardgeraden, ermittelte cDNA-Kopienzahl gibt Aufschluss über die Menge ursprünglich vorhandener RNA-Transkripte. Angegeben wird für das jeweilige GFP-Konstrukt (GFP) h_GFP 0 (CpG 0), h_GFP 60 (CpG 60) sowie h_GFP 93 (CpG 93) die cytoplasmatische (A) sowie die nukleäre (B) RNA-Kopienzahl als Mittelwert aus 3 unabhängigen Analysen. Als interne Kontrolle wurde für Hygromycin (hygr) die Zahl der cytoplasmatischen (A) sowie der nukleären (B) RNA-Transkripte bestimmt.

Sowohl in der cytoplasmatischen als auch in der nukleären Fraktion zeigt sich für die untersuchten Reporterkonstrukte ein ähnliches Bild wie in transienten Proteinexpressionsstudien (siehe C.4.1). So weist das CpG-freie Konstrukt im *Lightcycler*

eine 20-fach geringere cytoplasmatische (siehe Abbildung C.27A) sowie eine 14-fach geringere nukleäre RNA-Kopienzahl (siehe Abbildung C.27B) auf als das Ausgangskonstrukt. Für das CpG-maximierte Konstrukt h_GFP 93 wurde eine 8-fach geringere cDNA-Kopienzahl im Cytoplasma bzw. eine 5-fach geringere cDNA-Menge im Kern detektiert. Wie aus Abbildung C.27A und C.27B weiterhin hervorgeht, weisen alle untersuchten Konstrukte vergleichbare Mengen des Hygromycin-Transkriptes auf. Somit stellt sich die Frage, ob möglicherweise ein differentieller Kernexport für das Vorliegen unterschiedlicher Mengen GFP-spezifischer RNA-Transkripte im Cytoplasma verantwortlich ist.

C.6.2.2 Verhältnis nukleärer zu cytoplasmatischer RNA

Eine andere Darstellung der unter C. 6.2.1 generierten Daten ist in Abbildung C.28 gewählt. Hier werden absolute Kopienzahlen cytoplasmatischer sowie nukleärer GFP-spezifischer RNA-Kopien für die einzelnen GFP-Reportergene gegenübergestellt. Dabei errechnete sich für die Konstrukte h_GFP 0 sowie h_GFP 93 eine 9-fach höhere Kopienzahl im Kern, während sich bei dem Ausgangskonstrukt eine 12-fach höhere RNA-Menge im Nukleus ergab. Da alle untersuchten Konstrukte in etwa gleiche Verhältnisse von nukleärer zu cytoplasmatischer RNA zeigten, kann ausgeschlossen werden, dass ein Kernexportproblem für den beobachteten CpG-Effekt verantwortlich ist. In weiteren Experimenten sollte daher geklärt werden, ob das Vorliegen einer verringerten RNA-Menge in Cytoplasma und Kern durch Instabilität des CpG-freien RNA-Transkriptes bedingt ist.

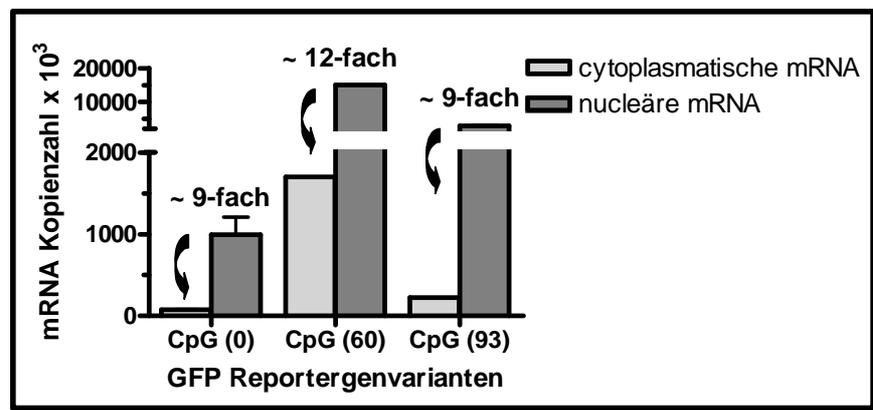


Abbildung C.28: Gegenüberstellung nukleärer und cytoplasmatischer RNA-Kopienzahlen für die GFP-spezifischen Transkripte. Die für die GFP-Konstrukte h_GFP 0 (CpG 0), h_GFP 60 (CpG 60) sowie h_GFP 93 (CpG 93) ermittelten cytoplasmatischen (Abbildung C.27A) bzw. nukleären (Abbildung C.27B) GFP-Kopienzahlen sind in dieser Abbildung gegenübergestellt. Zudem wird für jedes Konstrukt das Verhältnis nukleärer zu cytoplasmatischer GFP-Transkripte über dem Balkenpaar angegeben.

C.6.3 Untersuchung der RNA-Stabilität mittels ActinomycinD

In Inhibitorstudien mit *ActinomycinD* (AMD) wurden die verschiedenen GFP-Transkripte bezüglich ihrer RNA-Stabilität bzw. resultierender Halbwertszeit analysiert. Dazu wurden die stabil GFP exprimierenden CHO-Zelllinien (siehe C.5.4.1) mit 2,4 μM AMD behandelt. Zu den Zeitpunkten 0, 1,5, 3, 6, 12 und 24h nach dem Transkriptionsblock wurde dann die Gesamt-RNA aus den Zellen präpariert. Die absolute Menge an GFP kodierender RNA wurde in *realtime*-PCR-Analysen wie unter B.4.6.3 beschrieben quantifiziert. Ebenso wurde bei der Bestimmung β -actin kodierender Transkripte vorgegangen, die bei diesem Experiment als interne Kontrolle dienten. Zu den jeweiligen Zeitpunkten wurde für die Konstrukte h_GFP 0, h_GFP 60 sowie h_GFP 93 die Menge an GFP bzw. β -actin-spezifischer RNA ermittelt und in Abbildung C.29A bzw. C.29B als zeitlicher Verlauf dargestellt.

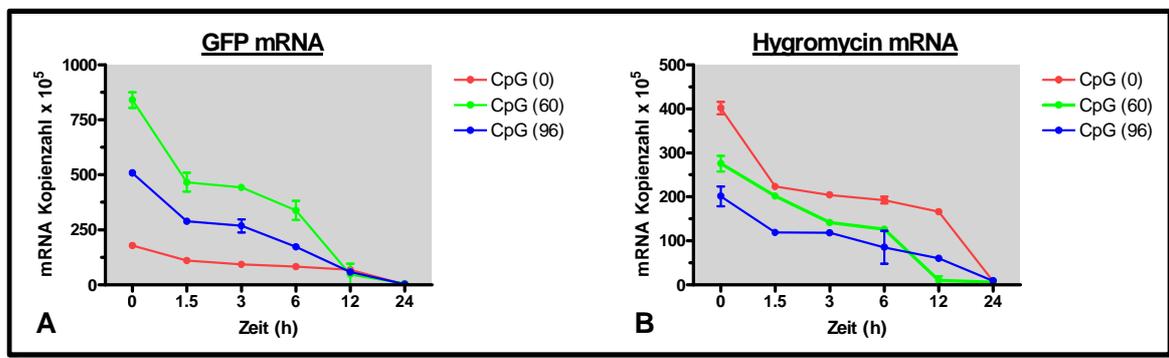


Abbildung C.29: RNA-Stabilität verschiedener CpG-modifizierter GFP-Konstrukte. (A) Für die Konstrukte h_GFP 0 (rote Linie), h_GFP 60 (grüne Linie) und h_GFP 93 (blaue Linie) ist der zeitliche Verlauf der GFP RNA-Degradation dargestellt. (B) Für die Konstrukte h_GFP 0 (rote Linie), h_GFP 60 (grüne Linie) und h_GFP 93 (blaue Linie) ist der zeitliche Verlauf der Hygromycin RNA-Degradation dargestellt.

Wie aus dieser Darstellung hervorgeht, stimmten die Mengenverhältnisse an GFP RNA zum Zeitpunkt null mit den in C.6.2.1 beschriebenen *Lightcycler*-Daten überein. Drei Stunden nach dem AMD-vermittelten transkriptionellen Block war von den GFP-Transkripten nur noch ca. 50% der ursprünglichen Menge nachweisbar. Zwar unterschieden sich die absoluten Mengen der CpG-modifizierten Transkripte deutlich voneinander, die Kinetik des RNA-Abbaus erwies sich jedoch unabhängig von ihrem CpG-Gehalt. Nach Ermittlung der Zerfallskonstante k (siehe B.4.8) ergaben sich für alle untersuchten GFP-Konstrukte in AMD-Studien vergleichbare Halbwertszeiten, welche bei ca. 3 Stunden lagen. Die Halbwertszeit von β -actin betrug etwa 4 Stunden für die Konstrukte mit 0 CpG und 60 CpG bzw. 5,5 Stunden für das CpG-maximierte Konstrukt (siehe Abbildung C.30). Dieser Wert ist mit den in der Literatur (Leclerc *et al.*, 2002) für β -actin beschriebenen Halbwertszeiten von 6,6 bis 13 h vereinbar. Der beobachtete CpG-Effekt resultiert also nicht aus einer erhöhten Stabilität des CpG-reichen RNA-Transkriptes. Daher wurden die RNA-Transkripte der

verschiedenen GFP-Reportergenvarianten zudem auf die Existenz alternativer Spleissprodukte hin untersucht.

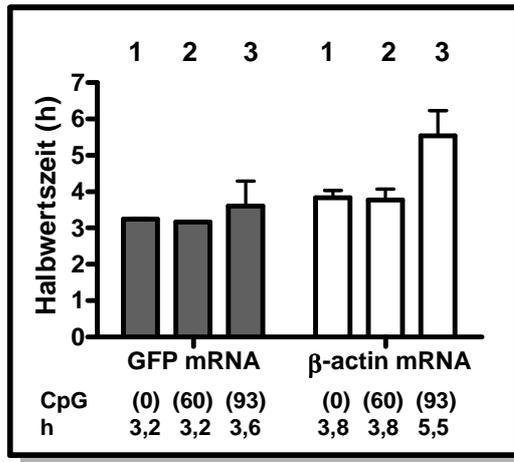


Abbildung C.30: Halbwertszeit verschiedener GFP- bzw. β-actin kodierender Transkripte. Ausgehend von den in Abbildung C.29A und C.29B dargestellten Daten wurde für die Konstrukte h_GFP 0 (1), h_GFP 60 (2) bzw. h_GFP 93 (3) die Zerfallskonstante für das jeweilige GFP-Transkripte sowie für das Hygromycin-Transkript errechnet (siehe B.4.8). Die daraus ermittelte RNA-Halbwertszeit für GFP (graue Balken) bzw. Hygromycin (weiße Balken) ist in der Abbildung dargestellt. Zudem ist unter dem Balken für das jeweilige Konstrukt die Zahl der CpGs und die Halbwertszeit in Stunden angegeben.

C.6.4 Untersuchung der RNA-Transkripte hinsichtlich alternativer Spleissprodukte

Obwohl der Literatur zu entnehmen ist, dass die humanisierte Form von GFP nicht gespleisst wird (Zolotukhin *et al.*, 1996), kann eine Modifizierung der Nukleotidsequenz ungewollt zu einem abweichenden Spleissverhalten führen. Aus diesem Grund wurden die RNA-Transkripte der verschiedenen GFP-Reportergenvarianten auf die Existenz alternativer Spleissprodukte hin untersucht. Dazu wurde die wie unter Punkt B.4.4 beschrieben generierte cDNA der jeweiligen Konstrukte als Matrize in eine PCR eingesetzt. Anhand der Primer hu/E019_pf und hu/E019_pb (für die Konstrukte h_GFP 0 und h_GFP 60) sowie der Oligonukleotide GFPmax_pf und GFPmax_pb (für die CpG-maximierte Version h_GFP 93) wurde die komplette NS-Sequenz der jeweiligen GFP-Version, beginnend bei ATG und abschließend mit TAG, amplifiziert. Bei Vorliegen alternativer Spleissprodukte sollten in der PCR-Analyse neben dem Vollängenkonstrukt mit 740 bp auch kleinere Fragmente detektiert werden. In LC-Analysen wurde für h_GFP 0 und h_GFP 60 ein internes 363-bp-langes Fragment des RNA-Transkriptes analysiert (siehe B.4.6.3). Für das CpG-reduzierte Konstrukt ergab sich dabei eine Reduktion der GFP-spezifischen Transkripte (siehe Abbildung C.27). Würde das 363-bp-große Fragment durch Spleissen gezielt aus dem Transkript entfernt, so wäre in der Vollängen-PCR zusätzlich zu dem ungespleissten Konstrukt ein 377-bp-großes Produkt zu erwarten. Gleiches gilt für die CpG-maximierte Variante, deren im *Lightcycler* analysiertes Fragment 375-bp-groß ist.

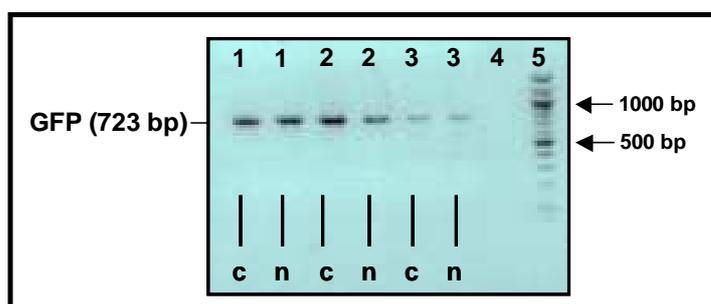


Abbildung C.31: Vollängen-PCR über die cDNA CpG-modifizierter GFP-Reporterkonstrukte. Abgebildet sind die erhaltenen PCR-Produkte der einzelnen Konstrukte sowie eine mit H₂O durchgeführte PCR als Negativkontrolle (4) und der DNA-Marker (5) auf einem 1%-igem Agarosegel.

Die erhaltenen Ergebnisse für cytoplasmatische sowie nukleäre RNA sind in Abbildung C.31 dargestellt. Wie aus dieser Darstellung hervorgeht wurden für die einzelnen Konstrukte keinerlei alternative Spleissprodukte detektiert. Auch durch Sequenzierung der erhaltenen PCR-Produkte konnten keine Abweichungen von den Ausgangssequenzen nachgewiesen werden. Zusammenfassend lässt sich feststellen, dass das auf RNA-Ebene beobachtete CpG-Phänomen weder durch veränderten Kernexport, noch durch unterschiedliche RNA-Stabilitäten oder das Vorliegen differentieller Spleissprodukte erklärt werden konnte. Möglich war jedoch, dass für das CpG-depletierte Reporterkonstrukte bereits *de novo* weniger RNA synthetisiert wurde.

C.6.5 Quantitative Analyse *de novo* synthetisierter RNA

Um zu analysieren, ob bereits für neu synthetisierte RNA eine Korrelation von CpG-Gehalt und vorliegender Transkriptmenge gegeben ist, wurde die transkriptionelle Aktivität der stabil GFP exprimierenden CHO-Zellen (siehe C.5.4.1) in einem anderem Assay-System („*nuclear run-on*“) bestimmt (siehe B.4.7). Dazu wurden die isolierten Kerne, wie unter Kapitel B.4.7 beschrieben, mit Biotin-16-UTP markierten Nukleotiden supplementiert. Nach einer sehr kurzen Inkubationsphase wurde die Gesamt-RNA aus den Kernen isoliert und mit Streptavidin gekoppelten *Dynabeads M-280* versetzt. Mit Hilfe einer magnetischen Apparatur wurde anschließend die an *magnetic beads* gebundene, *de novo* synthetisierte RNA separiert und in *realtime*-PCR-Analysen quantifiziert. Dabei spiegelt dieser Assay den aktuellen Status der nukleären Transkription während des 30-minütigen Pulses wieder. Als interne Kontrolle diente das Haushaltsgen β -actin, dessen Menge neu synthetisierter Transkripte unter Verwendung der Primer *actin_pf* und *actin_pb* (siehe F.1.3) in *Lightcycler*-Analysen bestimmt wurde.

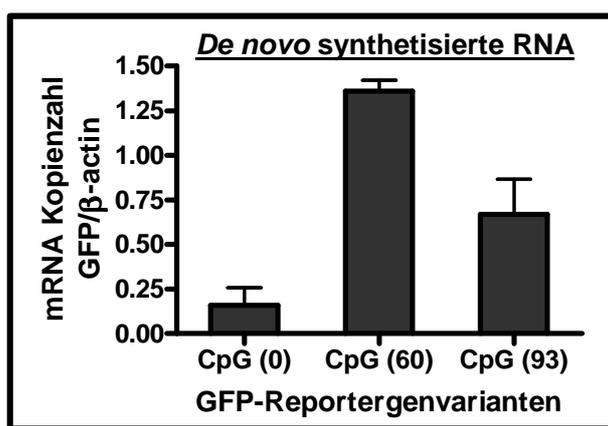


Abbildung C.32: Quantifizierung *de novo* synthetisierter RNA im „*nuclear run-on*“-Assay. Aus stabil h_GFP 0, h_GFP 60 bzw. h_GFP 93 exprimierenden CHO-Zellen wurden 3×10^7 Kerne präpariert. Die Kerne wurden mit Biotin-markierten Nukleotiden supplementiert, und nach einem kurzem Puls wurde die Gesamt-RNA präpariert. Nach Aufreinigung der markierten RNA durch magnetische *Beads*, anschließendem DNase-Verdau sowie Reverser Transkription wurden die resultierenden cDNA-Proben im *Lightcycler* quantifiziert. Für alle CpG-modifizierten GFP-Versionen wurde die Menge GFP-spezifischer RNA-Transkripte

detektiert. Zudem wurde für alle Konstrukte die *de novo* synthetisierte RNA-Menge des Haushaltsgens β -actin bestimmt. In dieser Darstellung ist die jeweilige neu synthetisierte GFP-Transkriptmenge normalisiert auf β -actin angegeben.

Eine Übersicht über die Menge neu synthetisierter GFP-Transkripte, normalisiert auf die jeweilige β -actin Menge, gibt Abbildung C.32. So ist in Abbildung C.32 für die Konstrukte h_GFP 0 und h_GFP 60 wiederum eine positive Korrelation des

CpG-Gehaltes und der Menge *de novo* synthetisierter RNA-Transkripte erkennbar. Das CpG-depletierte Reporterkonstrukt wies dabei im Vergleich zum Ausgangskonstrukt h_GFP 60 eine um den Faktor 7 verringerte Zahl an Transkripten auf. Wie aus Abbildung C.32 ebenfalls hervorgeht, ist die Menge der Transkripte für das CpG-maximierte Konstrukt h_GFP 93 im Vergleich zum Ausgangskonstrukt um die Hälfte reduziert. Für die transkriptionelle Aktivität der untersuchten Reporterkonstrukte ergab sich also im „*nuclear run-on*“ ein ähnliches Muster wie in vorhergegangenen RNA-Experimenten (siehe C.6.1.1) und Proteinexpressionsstudien (siehe C.5.4.1). *Realtime*-PCR-Analysen weisen hohe Reproduzierbarkeit, Sensitivität sowie Signifikanz auf und stellen daher ein äußerst verlässliches Testsystem zur Evaluierung vorliegender RNA-Transkriptmengen dar. Somit konnte für das beobachtete CpG-Phänomen folgende Hypothese aufgestellt werden: Der CpG-Gehalt im Transgen stellt einen intragenomischen Regulationsmechanismus dar, anhand dessen Transgenexpression und damit verbundene Reporteraktivität auf transkriptioneller Ebene gesteuert werden.

C.6.6 Zusammenfassung

Die auf Proteinebene für verschiedene Transgene beobachtete Korrelation von Reporteraktivität und CpG-Gehalt wurde in verschiedenen Säugetierzellen auch auf RNA-Ebene nachgewiesen. Für das CpG-freie GFP-Reporterkonstrukt war sowohl die Zahl der stationär (siehe C.6.1 und C.6.2.1) vorliegenden RNA-Transkripte als auch die Menge *de novo* synthetisierter RNA (siehe C.6.5) erheblich verringert. Hinsichtlich ihrer Kernexportrate (siehe C.6.2.2), der Stabilität ihrer RNA-Transkripte (siehe C.6.3) sowie alternativer Spleissprodukte (siehe C.6.4) zeigten die untersuchten Reportervarianten keine signifikanten Unterschiede, was für eine Beteiligung der CpG-Dinukleotide an der Regulation der Genexpression auf transkriptioneller Ebene spricht.

D Diskussion

D.1 Neuartige Strategien zur Steigerung der Expressionsstärke

D.1.1 Übersicht über vorhandene Systeme zur Erhöhung der Transgenexpression

Für die moderne Medizin mit ihren Bereichen Gentherapie und DNA-Vakzinierung, die molekularbiologische Grundlagenforschung sowie die pharmazeutische Industrie gewinnt die Herstellung rekombinanter therapeutischer bzw. diagnostischer Proteine zunehmend an Bedeutung. Grundvoraussetzung für eine effektive Produktion heterologer Therapeutika bzw. Basis für den Erfolg gentherapeutischer Maßnahmen sowie DNA-basierter Impfstrategien ist dabei eine hohe und lang andauernde Expression des rekombinanten Genproduktes. Da zur Einbringung des gewünschten Transgens in heterologe Expressionssysteme zumeist DNA-Vektoren verwendet werden, setzen viele Ansätze zur Optimierung der Proteinausbeute bei der Gestaltung dieser Komponente an.

So bietet die Firma Lonza (Basel, Schweiz) ein käufliches System (GS-System) zur Etablierung stabiler, rekombinanter Zellklone mit hoher Expressionsleistung an. Das System nutzt die Tatsache, dass die Aminosäure Glutamin, deren Bildung aus Glutamat und Ammoniak durch das Enzym Glutamat-Synthetase katalysiert wird, für viele Säugerzellen essentiell ist. Da die meisten Säugerzellen dieses Enzym nicht oder nicht in ausreichendem Maße produzieren, verwendet das System einen Expressionsvektor, der für die Glutamat-Synthetase kodiert. Die Integration dieses Vektors in eine transkriptionell hoch aktive Region im Genom gewährleistet zum einen die für die Zelle notwendige Glutamin-Herstellung und zum anderen eine stabile und effektive Expression des gewünschten Transgens (Bebington *et al.*, 1992). Auf diese Weise werden derzeit viele als Therapeutika eingesetzte bzw. in klinischen Studien evaluierte rekombinante Antikörper großtechnisch produziert (Vernon, 2004).

Ein weiterer Ansatz zur Erhöhung der stabilen Transgenproduktion beinhaltet die Integration von MARs (*matrix attached regions*) in einen spezifischen Expressionsvektor (Zahn-Zabal *et al.*, 2001). Dabei handelt es sich um Ankersequenzen, über welche Chromatindomänen mit dem Proteinskelett des Zellkerns verbunden werden. Diese Sequenzen ermöglichen extrachromosomal vorliegenden DNA-Vehikeln die Nutzung des zellulären Vervielfältigungsapparats und führen nach Öffnung der Chromatinstruktur und Reduktion superhelikaler Spannung zur Anhebung transkriptioneller Initiationsraten (Phi-Van *et al.*, 1990). Diese Erkenntnis ist vor allem für die Konstruktion retroviraler Vektoren von großer Bedeutung, welche ausgeprägten Mechanismen zur Repression der Genexpression unterliegen (Dang *et al.*, 2000).

Eine andere Strategie zur Erhöhung der Antigenexpression besteht in der Verwendung Intron-haltiger Expressionsvektoren (Yew *et al.*, 1997). Auf diese Weise wird *in vitro* eine verbesserte Polyadenylierung des RNA-Transkriptes erreicht, was zu einer erhöhten Transgenproduktion führt (Niwa *et al.*, 1990). So ist auch das in dieser Arbeit verwendete Plasmid pOG44, das für die Flp-Rekombinase kodiert, Träger eines synthetischen Introns (siehe B.1.3.1.2). Darüber hinaus hat der Einsatz starker, ubiquitär funktioneller Hybridpromotoren einen förderlichen Einfluss auf Transkription und Proteinproduktion eines Gens (Xu *et al.*, 2002; Yew *et al.*, 2001). Auch das in dieser Arbeit verwendete Flp-In System (Invitrogen) bietet eine Möglichkeit zur Optimierung der stabilen Transgenexpression. Dabei erfolgt, wie unter Punkt B.1.3.1 beschrieben, die Flp-Rekombinase vermittelte Integration des Expressionsvektors in einen transkriptionell hoch aktiven Bereich des Genoms der Wirtszelle. Daraus resultieren sowohl für den Selektionsmarker HygromycinB als auch für das rekombinante Genprodukt hohe Transkriptionsraten sowie Proteinausbeuten (Craig, 1988; Sauer, 1994). Während alle in diesem Kapitel erwähnten Strategien zur Expressionssteigerung auf einem Austausch einzelner Vektorkomponenten bzw. der Integration zusätzlicher Sequenzinformation beruhen, wurde in der vorliegenden Arbeit ein anderer Ansatz gewählt.

D.1.2 Eigenes Konzept zur Verbesserung der Transgenexpression

Für viele DNA-Applikationen wird eine hohe und stabile Transgenexpression benötigt. Um diese zu gewährleisten, wurde in der vorliegenden Arbeit der gesamte Sequenzkontext eines funktionsfähigen Plasmids inklusive kodierender Region hinsichtlich Expressions-modulierender Faktoren modifiziert. Als Expressions-modulierende Komponente bieten sich CpG-Dinukleotide an, da CpG-haltige Sequenzen anhand multipler Mechanismen Dauer und Höhe der Transgenexpression beeinflussen. Nachfolgend wird zunächst die Beteiligung der CpG-Dinukleotide an der Regulation der Genexpression näher erläutert.

Aus der Literatur ist bekannt, dass im Promotorbereich liegende CpG-Dinukleotide (Jones *et al.*, 1999), aber auch im 5´ bzw. 3´ Bereich der kodierenden Region lokalisierte CpG-Dinukleotide (Chevalier-Mariette *et al.*, 2003; Hisano *et al.*, 2003) durch DNA-Methyltransferasen (DNMTs) methyliert werden. Dabei katalysieren diese Enzyme die Übertragung einer Methylgruppe auf das 5´ Kohlenstoffatom des Cytidinrests. Der Methylierungsprozess geht meist mit einer Abnahme der Genexpression einher (Deng *et al.*, 2001; Jones *et al.*, 1999). So werden durch Einfügen der CH₃-Gruppe in die Promotorregion zum einen Transkriptionsfaktorbindestellen blockiert (Molloy *et al.*, 1990; Fujii *et al.*, 2006), zum anderen können Transkriptionsrepressoren wie z.B. MeCP2 an 5´mCpG-Dinukleotide binden (Nan *et al.*, 1998; Ng *et al.*, 1999), woraus eine verringerte Genaktivität resultiert. DNMTs besitzen ebenfalls Repressorfunktion und können analog zu mCpG-Bindeproteinen Histon-Deacetylasen (HDACs) rekrutieren, welche im Genom eine Kondensation der Chromatinstruktur und somit ein Abschalten der Transkription (Bachman *et al.*, 2001; Fuks *et al.*, 2001; Sarraf *et al.*, 2004)

bewirken. Ob auch extrachromosomal in Säugerzellen vorliegende DNA diesem negativen Regulationsmechanismus unterliegt, ist bisher noch nicht bekannt. Dennoch gibt es in der Literatur Hinweise, dass eukaryontische Expressionsplasmide, deren CpG-Gehalt verringert wurde *in vivo* zu einer hohen und lang andauernden Transgenproduktion führen (Yew *et al.*, 2002; Hodges *et al.*, 2004). CpG-Dinukleotide sind jedoch nicht nur an einer Steuerung der Genexpression beteiligt, sondern spielen bei weiteren zellulären Prozessen eine Rolle.

So üben in einem bestimmten Sequenzkontext vorliegende CpG-Dinukleotide, sogenannte CpG-Motive (siehe A.2), *in vivo* einen Einfluss auf spezifische Komponenten des humoralen und zellulären Immunsystems aus (Krieg *et al.*, 1995; Krieg *et al.*, 1998; Brazolot *et al.*, 1998). CpG-Dinukleotide und somit auch CpG-Motive kommen im Genom von Vertebraten mit einer wesentlich geringeren Häufigkeit vor als in bakterieller DNA. Auf diese Weise fungieren CpG-Motive als Komponente eines zellulären Erkennungsmechanismus, anhand dessen das Immunsystem von Vertebraten bakterielle DNA als fremd erkennt und anschließend den bakteriellen Erreger eliminiert. In der Literatur gibt es weiterhin Hinweise, dass die 5' Cytidin-Methylierung eines *in vivo* applizierten Plasmids in einer verlängerten Transgenexpression resultiert, da CpG-Motive nach Methylierung vom Immunsystem nicht mehr als solche erkannt werden. In der Folge bleiben entzündliche Immunreaktionen des Organismus aus, welche die Genexpression negativ beeinflussen (Reyes-Sandoval *et al.*, 2004; Hodges *et al.*, 2004). Für bestimmte Applikationsformen z.B. im Bereich der Gentherapie sind immunstimulierende Eigenschaften des Vektorrückgrates unerwünscht, um eine Erkennung transfizierter Zielzellen durch das Immunsystem zu vermeiden. Anders verhält es sich bei DNA-Vakzinierungen. Hier wird auf eine Aktivierung bestimmter Komponenten des humoralen und zellulären Immunsystems abgezielt (Horner *et al.*, 2001; Kojima *et al.*, 2002). Somit sind spezifisch das Immunsystem stimulierende CpG-Motive als vektoreigene Adjuvanskomponente erwünscht (Zhang *et al.*, 2005).

Um die von CpG-Motiven hervorgerufenen Effekte gezielt einsetzen zu können, ist jedoch zunächst ein „immun-neutrales“ Plasmid nötig, welches keine unspezifisch immunmodulierenden CpG-Motive im Vektorrückgrat enthält. Durch Integration ausgesuchter CpG-Motive kann dieses Ausgangsplasmid schließlich hinsichtlich seiner gewünschten Immunwirkung modifiziert werden. Da immunmodulatorische Eigenschaften sogenannter CpG-Motive *in vivo* zudem die Transgenexpression beeinträchtigen, und CpG-Dinukleotide nach Methylierung laut Literatur zu einer Repression der Transkription führen, wurde in dieser Arbeit versucht, ein möglichst CpG-reduziertes, funktionelles Vektorrückgrat zu generieren. Für zukünftige Anwendungen ist eine Kombination dieses CpG-minimierten Vektors mit den unter Punkt D.1.1 diskutierten Strategien zur Expressionsverbesserung denkbar. Zunächst sollte jedoch anhand verschiedener, partiell CpG-reduzierter und damit verringerter Methylierung unterliegender Vektoren analysiert werden, welche Konsequenzen sich *in vitro* für die resultierende Transgenproduktion ergeben. Eine Modifizierung des CpG-Gehaltes im Vektorrückgrat erfolgte dabei unter Einbeziehung von Füllsequenzen, kodierenden Regionen sowie regulatori-

schen Bereichen. Vor allem im Bereich funktioneller und regulatorischer Vektorelemente erweist sich eine CpG-Reduktion als problematisch, da deren Funktion beeinträchtigt bzw. zerstört werden kann. Aus diesem Grund wurde zunächst die Funktionalität der einzelnen CpG-minimierten Vektormodule untersucht.

D.2 CpG-reduzierte Vektormodule sind funktionell

D.2.1 Funktionalität einer deletierten Form des pUC origin of replication (pUC 110'origin)

Im Verlauf dieser Arbeit wurde gezeigt, dass eine CpG-reduzierte Version des pUC *origin of replication* Replikationsfähigkeit in Bakterien vermittelte (siehe C.2.2.1.3). Erreicht wurde diese Verringerung der CpG-Zahl im Vergleich zum Ausgangskonstrukt durch eine Deletion von 110 Basenpaaren im 5' Bereich des *origins*. Trotz der 110-bp-langen Deletion im *origin* wies das zugehörige Vektorkonstrukt im Vergleich zu Plasmiden mit dem herkömmlichen *origin of replication* keine offensichtlichen Unterschiede in Wachstumsverhalten oder erhaltener Plasmidkopienzahl auf. Es liegen jedoch keine vergleichenden Daten aus Wachstumskinetiken vor.

In Bakterien unterliegt die Vervielfältigung extrachromosomal vorliegender Plasmide einem Regulationsmechanismus, an dem im wesentlichen die RNA-Transkripte RNAI und RNAII des *origins* beteiligt sind (siehe Abbildung A.10). Dabei verhindert das kürzere RNAI-Fragment nach Bindung an das RNAII-Transkript die DNA-Replikation (Cesareni *et al.*, 1982; Kues *et al.*, 1989). Ist das Transkript der RNAII jedoch länger als 360 Basenpaare, wie es beim pUC *origin of replication* der Fall ist (siehe Abbildung D.1, schwarzer Pfeil), geht die inhibitorische Funktion der RNAI (siehe Abbildung D.1, roter Pfeil) verloren. Eine Bindung der RNAI an die RNAII bewirkt in diesem Fall zwar eine Konformationsänderung im 5' Bereich der RNAII, die Sekundärstruktur des an die DNA bindenden Bereichs (siehe Abbildung D.1, grüne Sequenz) bleibt jedoch unverändert (Merlin *et al.*, 1995). Auf diese Weise wird eine funktionelle Replikation gewährleistet. Eine Deletion im 5' Bereich der RNAII, wie sie bei dem CpG-reduzierten *origin* vorliegt (siehe Abbildung D.1, blaue Sequenz), sollte daher ebenfalls keinen Einfluss auf die Replikationsfähigkeit eines Plasmids haben, da der an die DNA bindende, funktionelle Bereich des RNAII-Transkriptes nicht modifiziert wird. Folglich übernimmt der 5' Bereich der RNAII anscheinend keine essentielle Funktion bei der Replikation von Plasmiden. Neben einem bakteriellen Replikationsursprung besitzen eukaryontische Expressionsvektoren zumeist starke virale Promotoren, wie z.B den CMV-Promotor, dessen Elemente und Funktionen nachfolgend diskutiert werden.

Sequenz des pUC origin of replication	
1	gtcaCCCGGG TAGTGAATTC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
61	AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
121	GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATAACAG GCGTTTCCCC
181	CTGGAAGCTC CCTCGTGGC TCTCCTGTTT CGACCCTGCC GCTTACCGGA TACCTGTCCG
241	CCTTCTCCC TTCGGAAGC GTGGCGCTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT
301	CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCGTT CAGCCCAGCC
361	GCTGCGCCTT ATCCGGTAACT TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
421	CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
481	AGTTCCTGAA GTGGTGGCCT AACTACGGCT AACTAGAAG GACAGTATTT GGTATCTGCC
541	CTCTGCTGAA GCCAGTTACC TTCGGAAGAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
601	CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
661	GATCTCAAGA AGATCCTTG ATCTTTTCTA CGGGAGATCT GTC AATCGAT agct

Abbildung D.1: DNA-Sequenz des pUC origin of replication (Umgekehrt komplementär). Fett gedruckt erscheint der funktionelle Bereich des origins laut pcDNA5/FRT. Ein schwarzer Pfeil kennzeichnet den Sequenzabschnitt, der für das 555-bp lange RNAI-Transkript kodiert. Die dem RNAI-Transkript (108 bp) zugrunde liegende DNA-Sequenz wird durch einen roten Pfeil hervorgehoben. Grün ist der an die DNA-bindende Bereich der RNAI markiert, und die 110-bp umfassende Deletion im 110_origin ist blau gekennzeichnet (Cesareni *et al.*, 1982; Kues *et al.*, 1989).

D.2.2 Funktionalität CpG-freier Varianten des CMV-Promotors

Für verschiedene Promotoren wurde bereits nachgewiesen, dass die Methylierung von CpG-Dinukleotiden im Promotorbereich zu einer drastischen Abnahme der Genexpression führt (Deng *et al.*, 2001; Agirre *et al.*, 2003). Auch bei dem konstitutiv und hoch exprimierenden, eukaryontischen CMV-Promotor (siehe A.4.1) wurde dieses Phänomen in *in vivo*-Expressionsstudien und in Zellkulturexperimenten eingehend beschrieben (Brooks *et al.*, 2004; <http://www.invitrogen.com>). Aus diesem Grund sollte in der vorliegenden Arbeit ein CpG-freier CMV-Promotor entwickelt werden, welcher keiner methylierungsbedingten Stilllegung unterliegen kann. Als problematisch erwies sich dabei die Tatsache, dass die Sequenz des CMV-Promotors multiple Bindestellen für Transkriptionsfaktoren enthält. Ein Austausch von Nukleotiden zur Eliminierung von CpG-Dinukleotiden generiert so veränderte Transkriptionsfaktorbindestellen, die möglicherweise nicht mehr funktionell bzw. in ihrer Funktionalität beeinträchtigt sind (Zhang *et al.*, 1995; Deng *et al.*, 2001; Tanaka *et al.*, 2005). Daher wurde ein kombinatorischer Ansatz zur Generierung CpG-freier CMV-Promotorvarianten gewählt (siehe C.2.1.1.5).

Als Maß für die Promotorstärke der funktionellen, CpG-freien Promotorvarianten wurde in transienten und stabilen Expressionsstudien die Autofluoreszenz der GFP-Reporterkonstrukte bestimmt. Dabei zeigte sich für alle CpG-depletierten Promotoren verglichen mit dem konventionellen CMV-Konstrukt eine deutliche Verringerung der Expressionsleistung. Dennoch entsprach das am besten exprimierende CpG-freie Promotorkonstrukt hinsichtlich seiner vermittelten Transgenexpression dem SV40-Promotor, welcher eine Komponente vieler eukaryontischer Expressionsplasmide darstellt. In Übereinstimmung mit Daten anderer Arbeitsgruppen, zeigte sich, dass das SV40-abgeleitete Promotorkonstrukt im Vergleich zu dem herkömmlichen CMV-Promotor nur etwa die Hälfte der

Expressionsleistung erbrachte (Foecking *et al.*, 1986; Ponder *et al.*, 1991; Guo *et al.*, 1996).

Zur Generierung komplett CpG-freier Promotorvarianten wurden zunächst partiell CpG-freie Promotorchimären etabliert, bei denen der vordere 5´ Bereich (A-Bereich), der mittlere Teil (B-Bereich) oder das endständige 3´ Drittel (C-Bereich) der Promotorsequenz CpG-depletiert wurde (siehe C.2.1.1.5). Bei der Charakterisierung dieser chimären Konstrukte bezüglich ihrer Promotorleistung zeigte sich, dass im mittleren Drittel (B-Abschnitt) der Promotorsequenz eingeführte Nukleotidaustausche zu einer Verringerung der Promotorstärke führten. Die Eliminierung von CpG-Dinukleotiden im 5´ oder 3´ proximalen Sequenzbereich bedingte jedoch keine Veränderung der Reporteraktivität. Diese Beobachtung impliziert eine wichtige Funktion des mittleren B-Bereichs, welche durch die angrenzenden Sequenzabschnitte nicht kompensiert werden kann. Tatsächlich befinden sich in dem internen B-Abschnitt mehrere der 19bp-Wiederholungselemente (siehe Abbildung D.2 und A.8), welche laut Literaturangaben ausschlaggebend für eine starke, konstitutive Expression sind (Boshart *et al.*, 1985; Stinski *et al.*, 1985). Denkbar ist zudem, dass durch Modifikation des mittleren Sequenzabschnittes wichtige TF-Erkennungssequenzen zerstört bzw. in ihrer Affinität zu dem jeweiligen Bindungspartner beeinträchtigt werden. Durch CpG-Depletion im A- oder C-Bereich wurde die Promotoraktivität des jeweiligen chimären Ausgangskonstruktes jedoch nicht beeinflusst. Möglicherweise übernehmen in diesem Fall in anderen Bereichen der Promotorregion lokalisierte TF-Erkennungssequenzen die Funktion der zerstörten Elemente und Bindestellen im 5´ bzw. 3´ proximalen Bereich. Eine Übersicht über die multiplen, im Bereich des CMV-Promotors auftretenden Erkennungssequenzen für Transkriptionsfaktoren gibt Abbildung D.2. Die Anzahl und Position der in der Promotorregion vorhandenen Wiederholungselemente kann Abbildung A.8 entnommen werden.

Sequenz des CMV-Promotors	
	A-Bereich
1	gttaATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA
61	TTAGTTCATA GCCCATATAT GGAGTTCGCG GTTACATAAC TTACG GTAAA TGGCCCGCCT
121	GGCTGAC CCGC CCAACGACCC CGGCCATTG ACGTCAATAA T GACGTATGT TCCCATAGTA
181	ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGG TGGAGT ATTTACGGTA AACTGCCCA C
241	TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCC CC CTATT GACGT CAATGAC GGT
301	AAATGGCCCG CCTGGCATT A TGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
361	TACATCT ACG TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT
421	GGCGTG GAT AGCGGTTGA CTCAC GGGGA TTTCC AAGTC TCC CCCCAT TGACGTCAAT
481	GGG AGTTTGT TTTGGCACCA AAATCAAC GG GACTTTCCAA AATGTCGTAA CAACTCC CC
541	CCATTGACGC AAATGG CCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTC at
601	tc

Abbildung D.2: Transkriptionsfaktorbindestellen im CMV-Promotor. In der Promotorregion lokalisierte Elemente und Erkennungssequenzen für Transkriptionsfaktoren sind strukturell gekennzeichnet bzw. farblich markiert. Sp1: hellblau, *non-consensus*-Bindestelle für Sp1(GT-Box): dunkelblau, CREB: rot, NFkB: orange, YY1: türkis; AP1: unterstrichen; CAAT-Element und TATA-

Box: grün. Kopien des 19-bp-Wiederholungselements sind eingerahmt. Aus Übersichtsgründen sind die 17-, 18- und 21-bp-Wiederholungselemente nicht aufgeführt, sondern ihre Anordnung kann Abbildung A.8 entnommen werden (Boshart *et al.*, 1985; Moens *et al.*, 2001; Meier *et al.*, 1996). Der vordere Sequenzabschnitt (A-Bereich) und die 3'-randständige Sequenz (C-Bereich) sind normal gedruckt, während die Sequenz des mittleren B-Bereichs kursiv dargestellt ist.

Auch ein in der Literatur beschriebener CpG-freier CMV-Promotor, dessen CpG-Dinukleotide systematisch durch TpG ersetzt wurden, weist eine geringere Promotorstärke auf als die herkömmliche CMV-Version (Yew *et al.*, 2002). Der CpG-depletierte CMV-Promotor erreichte im stabilen Expressionssystem jedoch eine höhere Reporteraktivität als die in dieser Arbeit etablierten CpG-freien Promotorvarianten. Möglicherweise wurde die regulatorische Funktion des Elements durch den systematischen Austausch eines Nukleotides (C zu T) weniger beeinträchtigt als durch den Einsatz randomisierter Oligonukleotide mit zwei degenerierten Positionen. Dennoch wurde in dieser Arbeit eine CpG-freie Promotorvariante etabliert, welche eine um den Faktor zwei geringere Expressionsleistung zeigte wie der CMV-Promotor, der zu den stärksten derzeit bekannten viralen Promotoren zählt (Löser *et al.*, 1998).

Bei der Diskussion der generierten Daten muss zudem berücksichtigt werden, dass ein möglicher positiver CpG-Effekt durch Beeinträchtigung *cis*-aktiver regulatorischer Elemente überlagert werden kann. Weder für den herkömmlichen CMV-Promotor noch für die CpG-freien Konstrukte konnte aber in dieser Arbeit eine Verringerung bzw. Abschaltung der Genexpression in Zellkultur beobachtet werden. Möglicherweise tritt dieser Effekt *in vivo* bzw. bei Verwendung anderer Zelllinien verstärkt auf.

D.2.3 Funktionalität CpG-freier BGH Polyadenylierungssignale

Im Rahmen der Arbeit wurden neben CpG-freien CMV-Promotorvarianten (siehe D.2.2) und einer CpG-reduzierten Form des puC *origin of replication* (D.2.1) auch CpG-depletierte Versionen des BGH Polyadenylierungssignals entwickelt. Unter Verwendung degenerierter Oligonukleotide wurden mehrere CpG-freie BGH P(A)-Signale hergestellt, und anschließend in transienten und stabilen Zellkulturexperimenten hinsichtlich ihrer Funktionsfähigkeit analysiert. Dabei erwiesen sich alle CpG-depletierten P(A)-Varianten als funktionell und zeigten eine mit dem herkömmlichen BGH Polyadenylierungssignal vergleichbare Transgenexpression. Darüber hinaus konnte in transienten und stabilen Zellkulturexperimenten je eine CpG-freie P(A)-Version ermittelt werden, welche mit einer höheren Expressionsleistung assoziiert war als das unveränderte BGH P(A)-Signal.

In transienten Expressionsanalysen generierte Daten sind abhängig von multiplen Faktoren wie z.B. der eingesetzte DNA-Charge, dem Zustand bzw. der Passagenzahl der Zellen, der Zusammensetzung verwendeter Reagenzien sowie dem Experimentator. Um ausgewählte P(A)-Signale quantitativ hinsichtlich ihrer erreichten Expressionsleistung zu analysieren, wurden daher stabile Zellkulturexperimente durchgeführt. Dabei erreichte eine CpG-freie BGH P(A)-Variante eine bessere Transgenexpression als das herkömmliche BGH-Signal.

Zudem vermittelte sowohl das CpG-freie BGH P(A)-Signal als auch sein unverändertes BGH P(A)-Pendent eine etwa doppelt so hohe Reporteraktivität wie ein Referenzkonstrukt mit SV40 P(A)-Signal. Diese Beobachtung stimmt mit Daten aus der Literatur überein, wonach das BGH P(A)-Signal im Vergleich zum SV40-Konstrukt eine mindestens doppelt so starke Expressionsleistung *in vitro* sowie *in vivo* vermittelt (Yew *et al.*, 1997; Xu *et al.*, 2002).

Aus dem Nukleinsäurevergleich der CpG-depletierten P(A)-Signale (siehe F.2.1) geht hervor, dass das AAUAAA-Signal (Proudfoot, 1989), welches für eine 3' Prozessierung und Polyadenylierung der RNA von erheblicher Bedeutung ist (siehe A.2.2) bei den einzelnen Konstrukten nicht verändert wurde. Daher ist es naheliegend, dass alle untersuchten Polyadenylierungsvarianten sich als funktionell erwiesen und im Vektorkontext eine dem unveränderten BGH P(A)-Konstrukt vergleichbare Reporterexpression gewährleisteten. Im Rahmen dieser Arbeit wurden nicht nur CpG-minimierte, regulatorische Komponenten wie das BGH P(A)-Signal etabliert, sondern auch kodierende Vektorelemente hinsichtlich ihres CpG-Gehalts minimiert.

D.2.4 Funktionalität CpG-freier Resistenzgene

Für die bakteriellen Resistenzmarker Ampicillin bzw. Kanamycin wurden CpG-depletierte Genvarianten hergestellt und auf ihre Funktionalität hin überprüft. Sowohl die für Amp- als auch die für Kana-kodierende, CpG-freie Genversion ermöglichte im Vektorkontext eine funktionelle Selektion des Plasmids in Bakterien. Quantitative Vergleichsaussagen hinsichtlich produzierter Proteinmenge bzw. vorliegender Wachstumskinetik wurden im Laufe dieser Arbeit nicht generiert, da die Funktionalität als essentielles Kriterium bereits nachgewiesen wurde.

Zusätzlich zu den zwei bakteriellen Resistenzgenen wurde eine CpG-depletierte Form des eukaryontischen Hygromycinresistenzgens produziert, das in vielen Vektorsystemen die Selektion stabil transfizierter Zellklone gewährleistet. Auch das CpG-freie, eukaryontische Resistenzgen wurde im CpG-reduzierten Vektorkontext auf seine Funktionalität hin in Zellkultur untersucht. Dabei zeigte sich, dass bei Verwendung der CpG-freien Hygromycin-Variante keine stabilen Zelllinien generiert werden konnten. Ein Austausch des CpG-freien Hygromycinresistenzgens durch das herkömmliche Pendent ermöglichte ebenfalls keine Etablierung von rekombinanten Zellen. Aus klonierungstechnischen Gründen wurde am 5' Bereich der für Hygromycin kodierenden Sequenz jedoch eine zusätzliche Schnittstelle (6 bp) angefügt, die in herkömmlichen Plasmiden nicht vorhanden ist. Daher stellte sich die Frage, ob durch die zusätzlich angefügte Schnittstelle der Mechanismus der FRT-vermittelten Integration (siehe B.1.3.1) an sich beeinträchtigt sei. Zur Klärung dieser Frage wurde im verwendeten, CpG-reduzierten Vektor die komplette Hygromycinresistenzkassette gegen den analogen Sequenzbereich des käuflichen Ausgangsvektors ausgetauscht, welcher im 5' Bereich der kodierenden Region keine Schnittstelle aufweist. Mit dem so erhaltenen Vektorkonstrukt konnten anschließend stabile Hygromycinresistenz vermittelnde Zellen generiert werden. Um zu analysieren, ob sich auch der

anfangs untersuchte CpG-reduzierte Vektor nach Entfernen der zusätzlichen Schnittstelle als funktionell erweisen würde, wurde die den 5' Bereich des Hygromycingens flankierende Schnittstelle via F-PCR aus dem Vektorkontext eliminiert. Die Verwendung des resultierenden Vektorkonstruktes ermöglichte jedoch keine Etablierung rekombinanter, stabiler Zelllinien.

Somit lässt sich über die Funktionalität des CpG-freien Hygromycinresistenzgens keine eindeutige Aussage treffen. Wie im Laufe dieser Arbeit jedoch gezeigt werden konnte, weisen CpG-depletierte Transgene in eukaryontischen Expressionssystemen eine weitaus geringere Antigenproduktion auf als ihre CpG-haltigen Pendants (siehe C.4.1). Somit liegt die Vermutung nahe, dass auch die CpG-freie Version des Hygromycingens in Zellkultur ein geringeres Expressionsniveau bedingt als die CpG-haltige Genvariante. Folglich ist denkbar, dass das synthetisch generierte Hygromycinkonstrukt an sich funktionell ist, aber die in der Zelle vorliegende Hygromycinkonzentration zu gering ist, um eine Resistenz zu vermitteln. Welchen Einfluss darüber hinaus CpG-reduzierte Vektoren auf die jeweilige Transgenexpression ausüben, wurde in transienten sowie stabilen Expressionsanalysen untersucht.

D.3 CpG-Reduktion im Vektorrückgrat führt zu erhöhter Transgenexpression

D.3.1 Erhöhung der Expression in transienten Studien

Bekannt ist, dass im Vektorrückgrat lokalisierte CpG-Dinukleotide durch Auslösen entzündlicher Reaktionen (Reyes-Sandoval *et al.*, 2004) die Expression eingebrachter Transgene *in vivo* beeinträchtigen (Yew *et al.*, 2002; Hodges *et al.*, 2004). In der Literatur gibt es jedoch keinerlei Hinweise darauf, dass im Plasmid befindliche CpG-Dinukleotide auch *in vitro* die Expressionsstärke eines Transgens beeinflussen. Zur Untersuchung dieser Fragestellung wurden die einzelnen unter C.2.1 beschriebenen, funktionellen sowie regulatorischen Vektorkomponenten zu verschiedenen Plasmiden mit unterschiedlichem CpG-Gehalt kombiniert. Anschließend wurde die transiente Expressionsleistung der generierten CpG-reduzierten Konstrukte im Vergleich zu einem Referenzplasmid mit konventionellen Vektorkomponenten in Zellkulturexperimenten analysiert.

In Studien mit humanen Lungenkarzinomzellen zeigte sich bei Verwendung verschiedener Reportergene eine tendenziell stärkere Expressionsleistung der CpG-reduzierten Vektoren im Vergleich zu einem CpG-haltigem Referenzkonstrukt. Ein CpG-reduziertes Plasmid, welches einen in der Literatur beschriebenen, CpG-freien CMV-Promotor (Yew *et al.*, 2002) enthält (siehe C.4.2), erreichte bei Verwendung von GFP als Reporter ebenfalls ein höheres Expressionsniveau als das CpG-haltige Referenzkonstrukt. Im Gegensatz dazu vermittelte das Vektorrückgrat mit dem CpG-freien CMV-Promotor eine geringere p24-Konzentration in Zellkulturexperimenten als die übrigen, CpG-reduzierten

Vektoren (siehe C.4.2). Letztere Beobachtung stimmt mit Daten aus der Literatur überein, wonach für den CpG-freien CMV-Promotor ein geringeres Expressionslevel detektiert wird als für das konventionelle CMV-Promotorkonstrukt (Yew *et al.*, 2002). Da die verwendeten Transgene sich jedoch hinsichtlich ihres CpG-Gehaltes wesentlich unterscheiden, spielen möglicherweise transgenspezifische Faktoren bei der unterschiedlichen Expressionsleistung dieses Plasmids eine Rolle. Darüber hinaus kann ein positiver CpG-Effekt durch die verringerte Promotorleistung des CpG-freien CMV-Promotors in mehr oder weniger starkem Ausmaß kompensiert werden.

Die Ergebnisse transienter Expressionsanalysen werden wie bereits erwähnt (siehe D.2.3) von multiplen Faktoren beeinflusst wie z.B. dem physiologischen Zustand der verwendeten Zellen, der Zusammensetzung eingesetzter Reagenzien, der DNA-Charge und nicht zuletzt der durchführenden Person. Im Rahmen der dadurch verursachten Schwankungen muss eine Vielzahl von Experimenten durchgeführt werden, um statistische Erhebungen zu ermöglichen und signifikante Daten zu generieren. Dennoch erweisen sich transiente Analysen als geeignetes Auswahlverfahren, um mehrere Kandidaten schnell, effizient und ökonomisch günstig hinsichtlich ihrer Expressionshöhe zu vergleichen. Nach Auswahl geeigneter Vektoren können diese in einem stabilen Expressionssystem, wie z. B dem Flp-In System, quantitativ untersucht werden.

D.3.2 Erhöhung der Transgenexpression in stabilen Flp-In Zelllinien

Im Gegensatz zur Durchführung transienter Experimente (siehe D.3.1) sind stabile Expressionsanalysen weitaus zeitaufwendiger bzw. auf bestimmte Zelltypen beschränkt. Da bei Verwendung des Flp-In Systems jedoch genau eine Kopie des jeweiligen Vektorkonstruktes in einen definierten genomischen Locus der Zielzelle integriert, sind die generierten Daten reproduzierbar und unabhängig von äußeren Faktoren wie DNA-Chargen oder Experimentator. Somit ist ein quantitativer Vergleich der einzelnen Vektor- oder Transgenkandidaten hinsichtlich ihrer resultierenden Transgenexpression möglich (Craig, 1988; Sauer, 1994). Eine einmalige Etablierung rekombinanter Zellklone gewährleistet die dauerhafte Produktion des gewünschten Transgens, statistische Erhebungen sind daher nicht erforderlich. Wie aus Kapitel C.5.1 hervorgeht, eignet sich das Flp-In System, um verschiedene Konstrukte hinsichtlich ihres Expressionsniveaus zu vergleichen, da in dieser Arbeit generierte Klone eines Konstruktes reproduzierbar einen identischen genetischen Hintergrund sowie eine unveränderte Expressionsstärke aufwiesen.

In stabilen Experimenten bestätigte sich die in transienten Studien beobachtete Tendenz einer positiven Korrelation von CpG-Gehalt im Vektorrückgrat und resultierender Transgenexpression (siehe C.5.3). Im Vergleich zum Referenzkonstrukt war für den CpG-reduzierten Vektor eine etwa 33%-ige Expressionssteigerung nachweisbar, welche über einen Zeitraum von 26 Wochen konstant gemessen wurde. Auffallend ist dabei, dass der transient nur als Tendenz beobachtete CpG-Effekt in stabilen Zelllinien verstärkt auftritt. Eine verbesserte

Expressionsleistung CpG-reduzierter Plasmide in Zellkultur ist von besonderer Relevanz, da therapeutisch eingesetzte Proteine und Antikörper wie Cytokine, Gerinnungsfaktoren (VIII & IX) oder Erythropoietin in zunehmendem Maße mittels stabiler Zelllinien hergestellt werden (Rader, 2003). Zudem stellen die in dieser Arbeit verwendeten, hoch exprimierenden CHO-Zellen ein weit verbreitetes eukaryontisches Produktionssystem zur heterologen Expression rekombinanter Antigene dar (Vernon *et al.*, 2004). Um wirtschaftlich zu arbeiten, d.h. Zeit, Kosten und Aufwand zu minimieren, wird dabei eine maximale Proteinausbeute pro Zelle angestrebt. Eine Erhöhung der stabil exprimierten Proteinmenge, die auch über einen längeren Zeitraum erhalten bleibt, ist daher für die Antikörper- bzw. Therapeutikaherstellung von großer Relevanz. Somit konnte in dieser Arbeit gezeigt werden, dass eine CpG-Reduktion im Vektorrückgrat eine Strategie zur Verbesserung der stabilen Transgenproduktion darstellt. Darüber hinaus sollte untersucht werden, ob auch CpG-Modifizierungen in der kodierenden Region des Transgens eine veränderte Antigenexpression zur Folge haben.

D.4 CpG-maximierte Transgene weisen eine erhöhte Proteinproduktion auf

Zusätzlich zu einem für Expressionszwecke optimierten Vektorrückgrat sollte im Rahmen dieser Arbeit auch eine Strategie zur Verbesserung der transgenspezifischen Proteinexpression entwickelt werden. Eine Vielzahl von Studien in Vertebraten hat belegt, dass eine Methylierung von CpG-Dinukleotiden in promotornahen Regionen (Kudo *et al.*, 1998; Deng *et al.*, 2001; Li *et al.*, 2004) einen reprimierenden Effekt auf die Genexpression ausübt. Darüber hinaus gibt es in der Literatur Hinweise, dass auch eine CpG-Methylierung in der kodierenden Region in bestimmten Geweben zu einem Abschalten der Genexpression führt (Hisano *et al.*, 2003; Chevalier-Mariette *et al.*, 2003; Bestor, 2000). Die „methylierungsbedingte“ Repression der Genaktivität beruht auf einer Veränderung der Chromatinstruktur, welche nachfolgend das Binden der Transkriptionsmaschinerie erschwert und eine Initiation der Transkription verhindert (Deng *et al.*, 2001). Auch Transkriptionsrepressoren, welche sich spezifisch an methylierte CpG-Dinukleotide anlagern, sind an diesem *Silencing*-Prozess beteiligt (Jones *et al.*, 1998). Darüber hinaus können TF-Bindestellen durch Methylierung von CpG-Dinukleotiden sterisch blockiert werden, was ebenfalls zu einer Stilllegung der Transkription führen kann (Clark *et al.*, 1997; Umezawa *et al.*, 1997; Fujii *et al.*, 2006).

Da CpG-Dinukleotide anhand multipler Effekte Einfluss auf die Genexpression nehmen, war es Ziel dieser Arbeit, auch in der kodierenden Sequenz eines Transgens den CpG-Gehalt zu verändern, und anschließend die Expressionsleistung der CpG-modifizierten Transgenvarianten mit dem Ausgangskonstrukt zu vergleichen. Erwartet wurde, dass bei der CpG-depletierten Variante eine CpG-Methylierung sowie ein damit verbundenes Abschalten der Genexpression unterbleibt und dadurch eine effektive Proteinexpression gewährleistet wird. Im Gegensatz dazu sollten das Ausgangskonstrukt und CpG-maximierte Transgen-

versionen dem Methylierungsprozess unterliegen und eine Abnahme der Antigenproduktion zeigen.

D.4.1 CpG-maximierte Reportergene vermitteln erhöhte Aktivität und Proteinmenge

Anders als erwartet, wies eine CpG-freie Version verschiedener Transgene *in vitro* eine schwächere Reporteraktivität auf als das CpG-haltige Ausgangskonstrukt (siehe C. 4.1). Dieses Ergebnis wurde sowohl in transienten als auch in stabilen Expressionsstudien erzielt, wobei unterschiedliche humane Zelllinien sowie CHO-Zellen verwendet wurden. Im transienten Expressionssystem resultierte die Eliminierung aller CpG-Dinukleotide aus der kodierenden Region für GFP in einer 50%-igen Verringerung der GFP-vermittelten Fluoreszenz, während bei Verwendung von p24 als Antigen die CpG-freie Variante sogar 60% der Reporteraktivität einbüßte (siehe C.4.1). Da für beide Reportergene eine Korrelation von CpG-Gehalt und detektierter Reporteraktivität festgestellt werden konnte, sollte am Beispiel von GFP dieser Zusammenhang näher untersucht werden. Wie für das Vektorrückgrat beschrieben (siehe D.3), zeigte sich auch bei der Untersuchung des Transgens, dass der transient beobachtete CpG-Effekt in stabilen Expressionsexperimenten verstärkt wurde. So war die mittlere Fluoreszenzintensität für das CpG-depletierte GFP-Konstrukt in stabilen Zelllinien um den Faktor 6-9 für CHO-Zellen, bzw. 10-20 für 293T-Zellen reduziert. Diese Beobachtungen implizierten die Frage, ob erhöhte MFI-Werte des CpG-haltigen GFP-Konstruktes durch eine verbesserte Fluoreszenzaktivität begründet sind oder auf dem Vorliegen einer höheren Proteinmenge basieren. Wie Western Blot-Analysen zeigten, wurde für die CpG-haltige GFP-Version eine erhöhte Proteinmenge detektiert. Zudem wurde für beide GFP-Varianten ausschließlich das Voll-längenprotein mit 27 kDa nachgewiesen. Daher kann ausgeschlossen werden, dass Degradationsprodukte für die verringerten MFI-Werte des CpG-freien Konstrukts verantwortlich sind (siehe C.5.1). Auch eine andere Arbeitsgruppe beschreibt bei Verwendung einer CpG-freien GFP-Variante im Vergleich zum Ausgangsprodukt einen Abfall der spezifischen Fluoreszenzaktivität *in vivo* bzw. *in vitro* (Dalle *et al.*, 2005).

Eine Erhöhung der CpG-Zahl im Transgen führte bei p24 zu einer erheblichen Steigerung der transienten Antigenexpression verglichen mit dem unmodifizierten Ausgangskonstrukt. Im Unterschied dazu resultierte eine Maximierung des CpG-Gehaltes in der kodierenden Region von GFP nicht in einer weiteren Erhöhung der Proteinproduktion, sondern bedingte eine Reduktion des transienten bzw. stabilen GFP-Expressionsniveaus. Auch für die CpG-maximierten Genversionen wurden in Western Blot-Analysen jedoch keine Degradationsprodukte nachgewiesen (Daten nicht gezeigt), welche diese Beobachtung erklären könnten. Möglich wäre aber, dass bei der maximierten GFP-Variante ein positiver CpG-Effekt durch die Verwendung seltener Kodons und einer damit verbundenen, verringerten Translation kompensiert wird. Denkbar ist aber auch eine zellulär-vermittelte Reduktion der Transgenexpression, da GFP in zu hoher Konzentration toxisch für

eukaryontische Zellen ist (Liu *et al.*, 1999). Somit wurde eine positive Korrelation von CpG-Gehalt im kodierenden Bereich und detektierter Reporteraktivität festgestellt. Ob dieser Zusammenhang auch auf RNA-Ebene gegeben ist, wurde in *realtime* PCR-Analysen exemplarisch für die GFP-Konstrukte untersucht.

D.4.2 CpG-haltige Transgene weisen erhöhte RNA-Mengen auf

Analog zu den Proteindaten zeichnete sich das CpG-haltige GFP-Konstrukt auch auf RNA-Ebene durch erhöhte Mengen cytoplasmatischer sowie nukleärer, transgenspezifischer RNA-Transkripte aus. So wurde für cytoplasmatische RNA in relativen Quantifizierungen via *Lightcycler* eine 7-fach (CHO-Zellen) bzw. 29-fach (293T-Zellen) erhöhte cytoplasmatische RNA-Menge der CpG-haltigen Transgenversion detektiert. Darüber hinaus wies das CpG-haltige GFP-Konstrukt in CHO-Zellen im Kern eine 13-fach höhere Menge GFP-spezifischer RNA auf. Da sich in beiden untersuchten Zelllinien das CpG-haltige GFP-Konstrukt durch erhöhte relative RNA-Mengen auszeichnete, wurde am Beispiel stabiler CHO-Zellen die absolute Menge transgenspezifischer RNA-Transkripte exemplarisch bestimmt. Auch hier zeichnete sich ein CpG-haltiges Konstrukt verglichen mit dem CpG-depletierten Pendant durch das Vorliegen vermehrter RNA-Transkripte aus. Allerdings wurde für das CpG-haltige Konstrukt eine 20-fach höhere cytoplasmatische sowie eine 14-fach höhere, nukleäre RNA-Kopienzahl festgestellt als für das Ausgangskonstrukt (siehe C.6.1 und C.6.1). Dass die relativen und absoluten RNA-Daten nicht völlig übereinstimmen, lässt sich folgendermaßen erklären:

Bei absoluten Quantifizierungsexperimenten wird die transgenspezifische RNA-Kopienzahl unter Verwendung eines externen Standards errechnet (siehe B.4.6.3.3). Bei der relativen Quantifizierung hingegen wird für das jeweilige Konstrukt die für eine signifikante Detektion durch den LC benötigte Zyklenzahl als Maß für die zugrunde liegende RNA-Menge ermittelt (siehe B.4.6.3.1). Da in relativen *realtime* PCR-Analysen ein Unterschied in der Zyklenzahl von 3,2 bereits dem Faktor 10 errechneter RNA-Transkripte entspricht, resultieren bereits minimale Unterschiede in der detektierten Zyklenzahl in erheblich veränderten RNA-Kopienzahlen (Freeman *et al.*, 1999). So kann erklärt werden, warum in relativen Quantifizierungsexperimenten für die CpG-haltige GFP-Version eine um den Faktor 7 bzw. 13 gesteigerte cytoplasmatische bzw. nukleäre RNA-Menge detektiert wurde, während in absoluten Experimenten ein cytoplasmatischer Faktor von 20 sowie ein nukleärer Faktor von 14 berechnet wurde. Zusammenfassend lässt sich feststellen, dass die relative Quantifizierung ein schnelles und effizientes Verfahren darstellt, um Kopienzahlen desselben Amplikons zu ermitteln. Für den Vergleich verschiedener Transgene jedoch stellt die absolute Quantifizierungsmethode mit dem Einsatz externer Standards ein äußerst sensitives und valides Nachweisverfahren dar. Unabhängig von der untersuchten Transgenvariante befanden sich in der nukleären RNA-Fraktion etwa 10-mal mehr transgenspezifische RNA-Kopien als im cytoplasmatischen Anteil.

Auch das CpG-maximierte GFP-Konstrukt wurde in *realtime* PCR-Analysen hinsichtlich seiner RNA-Kopienzahl untersucht. Die Ergebnisse korrelieren mit den in stabilen Expressionsstudien generierten Daten. So wurde für die maximierte Version von GFP im Vergleich zum Ausgangskonstrukt eine 8-fach geringere cytoplasmatische Transkriptmenge sowie eine 5-fach geringere RNA-Menge im Kern detektiert. Zusammenfassend lässt sich feststellen, dass die Reduktion der Proteinexpression für das CpG-freie Konstrukt auf das Vorliegen verringerter nukleärer sowie cytoplasmatischer RNA-Mengen zurückgeführt werden kann.

D.4.3 CpG-optimierte Transgene zeigen keine Unterschiede hinsichtlich Kernexportraten, RNA-Stabilität oder alternativer Spleissprodukte

Die beobachtete Reduktion GFP-kodierender RNA-Transkripte für das CpG-depletierte Konstrukt kann durch verschiedene Faktoren wie z.B. reduzierten Kernexport, Instabilität CpG-freier RNA, Vorliegen alternativer Spleissprodukte oder geringere Transkriptionsraten bedingt sein. Wie aus Kapitel C.6.2.2 hervorgeht, weisen alle untersuchten GFP-Konstrukte in etwa gleiche Verhältnisse von nukleärer zu cytoplasmatischer, GFP-spezifischer RNA auf. Alle drei GFP-Transkripte werden somit mit der gleichen Effizienz aus dem Kern ins Cytoplasma ausgeschleust. Daher kann ausgeschlossen werden, dass dem beobachteten CpG-Effekt ein Kernexportproblem zugrunde liegt (siehe C.6.2.2). Trotz gleicher Exportrate können die GFP-spezifischen Transkripte dennoch unterschiedliche RNA-Stabilitäten aufweisen und somit eine Veränderung der GFP-vermittelten Autofluoreszenz bedingen.

Um diese Fragestellung zu adressieren, wurde im Rahmen dieser Arbeit die Halbwertszeit der jeweiligen GFP- sowie Hygromycin-spezifischen RNA-Transkripte bestimmt (Leclerc *et al.*, 2002). Auch hierbei ergab sich kein Unterschied für die transgenspezifischen RNA-Stabilitäten der einzelnen GFP-Konstrukte, wodurch das Vorliegen unterschiedlicher RNA-Transkriptmengen erklärt werden könnte. Allerdings gibt es in der Literatur Hinweise, wonach eine verbesserte Expression CpG-haltiger Transgene durch die Existenz eines CpG-reichen RNA-Transkriptes mit stabiler, langsam degradierender Sekundärstruktur erklärt wird. Zudem wird das Fehlen von UpA-Dinukleotiden in der RNA, welche von Endoribonukleasen gespalten werden, in dieser Veröffentlichung für die erhöhte Stabilität CpG-reicher RNA verantwortlich gemacht (Duan *et al.*, 2003). Alle in der vorliegenden Arbeit verwendeten GFP-Reportergenvarianten weisen jedoch einen vergleichbaren TpA-Gehalt auf (siehe C.3.1). Daraus resultiert, dass das beobachtete CpG-Phänomen unabhängig vom TpA-Gehalt der kodierenden Region auftritt. Somit konnten für die untersuchten GFP-Konstrukte weder differentielle Kernexportraten noch unterschiedliche RNA-Stabilitäten detektiert werden. Eine andere Erklärung für das Vorliegen unterschiedlicher GFP-spezifischer RNA-Mengen wären differenzierte Spleissvorgänge bei den verschiedenen GFP-Transkripten.

Die humanisierte Form von GFP, welche in dieser Arbeit als Ausgangskonstrukt verwendet wurde, wird laut Literatur nicht gespleisst und enthält in ihrer Sequenz

kein Intron (Zolotukhin *et al.*, 1996). Im Zuge der CpG-Modifizierung könnte die Sequenz der CpG-depletierten bzw. der CpG-maximierten Transgenvariante jedoch derart verändert worden sein, dass das resultierende RNA-Transkript einem Spleissvorgang unterliegt. Daher wurde für alle drei analysierten GFP-Reportervarianten das Vorliegen alternativer Spleissprodukte untersucht. Jedoch konnten im Rahmen dieser Arbeit für keines der CpG-modifizierten Transgene oder das Ausgangsprodukt Spleissprodukte detektiert werden. Da weder Kern-exportrate, noch RNA-Stabilität oder unterschiedliches Spleissverhalten eine Erklärung für den beobachteten CpG-Effekt liefert, stellt sich die Fragen, ob für CpG-haltige Transgene bereits eine erhöhte *de novo* synthetisierte RNA-Transkriptmenge nachgewiesen werden kann.

D.4.4 CpG-haltige Transgene besitzen eine erhöhte Menge de novo synthetisierter, transgenspezifischer RNA

Um zu analysieren, ob bereits für neu synthetisierte RNA ein Zusammenhang von CpG-Gehalt und vorliegender Transkriptmenge gegeben ist, wurde die transkriptionelle Aktivität der stabil GFP exprimierenden Zellen in einem „*nuclear run-on*“-Assay untersucht. Dabei spiegelt dieser Assay den aktuellen Status der nukleären Transkription während eines 30-minütigen Pulses mit Biotin-markierten Nukleotiden wieder (Heap *et al.*, 2000). Biotin-markierte Nukleotide ermöglichen eine Isolierung der *de novo* synthetisierten RNA, welche anschließend im *Lightcycler* bezüglich transgenspezifischer RNA-Transkripte quantifiziert wurden. Da *realtime* PCR-Analysen äußerst sensitiv sind, eine hohe Reproduzierbarkeit aufweisen und innerhalb einer Versuchsreihe geringen statistischen Schwankungen unterliegen, liefert das System sehr fundierte und relevante Daten. So zeigten die resultierenden Ergebnisse für das CpG-haltige Konstrukt eine erhöhte *de novo* RNA-Synthese im Vergleich zur CpG-freien Variante. Analog zu den auf Expressions- bzw. RNA-Ebene generierten Daten wies die CpG-maximierte Version von GFP eine zwischen den RNA-Mengen des CpG-depletierten bzw. unveränderten Konstruktes befindliche Transkriptkopienzahl auf. Am Beispiel von GFP wurde somit nachgewiesen, dass CpG-haltige Konstrukte bereits auf transkriptionellem Niveau eine erhöhte Menge transgenspezifischer RNA-Transkripte besitzen. Diese Beobachtung impliziert, dass eine CpG-vermittelte Regulation der Genexpression bereits auf der Ebene der Transkription erfolgt. Unklar ist derzeit allerdings noch, auf welche Weise der CpG-Effekt vermittelt wird. Eine mögliche Hypothese, wie die CpG-vermittelte Genregulation ablaufen könnte, wird in Kapitel D.5 näher erläutert.

D.5 Hypothese der CpG-basierten Genregulation und Ausblick

D.5.1 Arbeitshypothese der CpG-basierten Genregulation

Im Rahmen dieser Arbeit wurde das beschriebene CpG-Phänomen bei stabil transfizierten 293T-Zellen humanen Ursprungs sowie stabilen rekombinanten CHO-Zellen, welche vom chinesischen Hamster (*Cricetulus griseus*) stammen, beobachtet. Auch in transienten Expressionsstudien mit humanen Lungenkarzinomzellen (H1299) zeigte sich für verschiedene Transgene eine Korrelation von CpG-Gehalt im Leserahmen und resultierender Proteinexpression. In weiteren Zellkulturexperimenten der Arbeitsgruppe wurde eben dieser CpG-Effekt auch bei Verwendung nicht-viraler Konstrukte wie z.B. Cytokinen oder *Erythropoietin* im transienten sowie im stabilen Expressionssystem detektiert. Am Beispiel von *Erythropoietin* wurde im murinen System zudem ein Zusammenhang von CpG-Gehalt im Transgen und *in vivo* hervorgerufener Reporteraktivität bzw. Proteinexpression festgestellt (Asli Bauer, Susann Meißner, Denijal Kosovac, unveröffentlichte Daten). Die Applikation viraler sowie nicht-viraler Transgene sowie der Einsatz verschiedenster Zelllinien sprechen dabei für einen generellen Mechanismus der CpG-basierten Genregulation auf transkriptioneller Ebene.

Bei den jeweiligen Expressionskonstrukten wurde unter Beibehaltung der Aminosäuresequenz ausschließlich die kodierende Region des Transgens modifiziert, der Sequenzkontext des jeweiligen Vektorrückgrates wurde nicht verändert. Der beobachtete CpG-Effekt trat unabhängig vom verwendeten Vektorrückgrat auf und ist somit auf Modifikationen des Transgens zurückzuführen. Ein erhöhter Methylierungsgrad, wie er in CpG-maximierten Genen vorliegen könnte, wird in der Literatur mit einer Repression der Translation und einer damit verbundenen Verringerung der Expression in Verbindung gebracht (Bird *et al.*, 1987; Chevalier-Mariette *et al.*, 2003; Hisano *et al.*, 2003). Mit diesem Modell kann die verstärkte Expressionsleistung CpG-reicher Gene daher nicht erklärt werden. In einer anderen Veröffentlichung (Duan *et al.*, 2003) wird nicht das vermehrte Auftreten von CpG-Dinukleotiden, sondern das Fehlen von TpA-Dinukleotiden in der kodierenden Region für eine höhere Expression CpG-reicher Gene verantwortlich gemacht (siehe D.4.3). Auch mit dieser Theorie stimmen die im Rahmen der Arbeit generierten Ergebnisse nicht überein, da die jeweiligen Transgene trotz vergleichbarem TpA-Gehalt (siehe C.3.1) unterschiedliche Expressionshöhen aufwiesen. Neben Repressorproteinen, die an methylierte CpG-Dinukleotide binden, wird in der Literatur jedoch auch ein Transaktivatorprotein beschrieben, das an unmethylierte CpGs bindet und die Expression auf transkriptioneller Ebene erhöht (Voo *et al.*, 2000).

Dabei handelt es sich um das 88-kDa große, im Kern lokalisierte CGBP (CpG-Bindeprotein), welches in der Literatur auch als CFP1 (CXXC finger protein 1) bezeichnet wird (Lee *et al.*, 2005). Neben einer CXXC-Domäne zur Bindung von

CpG-haltiger DNA besitzt das Protein auch zwei pflanzliche Homeodomänen, welche charakteristisch für Chromatin-assoziierte Proteine und/oder Regulatoren der Genexpression sind (Carlone *et al.*, 2001; Lee *et al.*, 2001). Das Protein wird ubiquitär exprimiert, liegt im Genom von Mensch, Maus, Hamster und Ratte sehr konserviert vor und bindet spezifisch an doppelsträngige-DNA mit mindestens einem CpG-Dinukleotid. Anhand welcher molekularer Mechanismen eine Erhöhung der Transkriptionsrate durch CGBP im Einzelnen erfolgt, ist derzeit noch nicht bekannt. Für murine, embryonale Stammzellen des Genotyps *CGBP*^{-/-} konnte jedoch im Vergleich zur unveränderten Zelllinie ein verringerter CpG-Methylierungsgrad, sowie eine reduzierte Expression und Aktivität von DNMTs nachgewiesen werden. Zudem wurde unlängst gezeigt, dass CGBP eine Komponente des Histon-H3-Methyltransferase-Komplexes von Säugetieren darstellt. Somit wird für CGBP in der Literatur die Rolle eines epigenetischen Regulatorproteins diskutiert, welches die Methylierung von Cytosinresten sowie die Modifikation von Histonen kontrolliert (Carlone *et al.*, 2005; Lee *et al.*, 2005). Aufgrund dieser Beobachtungen könnte das in dieser Arbeit beobachtete CpG-Phänomen durch die Existenz des Transaktivators CGBP erklärt werden. Unter Berücksichtigung der vorhandenen Daten (Voo *et al.*, 2000; Carlone *et al.*, 2000) würden CpG-reiche Transgene mehrere Aktivatoren binden und so, wie beobachtet, eine erhöhte Transkription und damit eine verstärkte Expressionsleistung hervorrufen. CGBP besitzt (Lee *et al.*, 2001) eine optimale Bindesequenz der Abfolge (A/C)CpG(A/C), welche für p24 im CpG-maximierten Gen etwa dreimal öfter vorkommt, als im Ausgangskonstrukt und im CpG-depletierten Konstrukt völlig fehlt. Dass die murine Form von CGBP an unmethylierte Plasmid-DNA, nicht aber an *in vitro* methylierte DNA bindet, konnte in der Arbeitsgruppe Wagner bereits gezeigt werden (Asli Bauer, unveröffentlichte Daten).

Auch eine verbesserte Expressionsleistung CpG-reduzierter Plasmide könnte mit dieser Arbeitshypothese in Einklang gebracht werden. So würden CpG-minimierte Vektoren eine verminderte Anzahl dieses Aktivators binden, während CpG-haltige Expressionsplasmide in ihrer Sequenz multiple Bindestellen für CGBP besitzen. Bei Verwendung CpG-haltiger Plasmide würden die im kodierenden Bereich benötigten Transaktivatoren somit verstärkt vom Vektorrückgrat gebunden und stünden nicht für eine Aktivierung der Transkription zur Verfügung. Voraussetzung dabei wäre, dass das Protein in der Zelle in physiologischer Konzentration und nicht im Überschuss vorliegt. Da bei transienten Transfektionen das Expressionsniveau durch eine Vielzahl von Plasmiden pro Zelle bedingt wird, ist eine Konkurrenz von Vektorrückgrat und Transgen hinsichtlich limitiert vorliegendem Transaktivator durchaus denkbar. Ob jedoch auch das Vorliegen einer Vektorkopie, wie es bei stabilen (Flp-In) Zelllinien der Fall ist, ausreichend ist, um dem Transgen die benötigten Aktivatorproteine kompetitiv zu entziehen, ist jedoch fraglich.

Wie bei den meisten allgemein gültigen Mechanismen gibt es auch hier Ausnahmen. So zeigt die CpG-maximierte Version von GFP im Vergleich zu den anderen in der Arbeitsgruppe untersuchten Transgenen keine verbesserte Transgenexpression. Allerdings wurde bei diesem Konstrukt der CpG-Gehalt

inklusive optimaler CGBP-Erkennungssequenzen (25) im Vergleich zum Ausgangskonstrukt (18) nur um den Faktor 0,5 gesteigert, während bei den anderen untersuchten Kandidaten die CpG-Zahl und damit auch die Zahl der optimalen Bindestellen für CGBP um den Faktor 3-6 erhöht wurde. Eine andere Erklärung wäre, dass bei dem CpG-haltigen Ausgangskonstrukt das „CpG-Optimum“ bereits erreicht ist, und eine darüber hinaus gehende Erhöhung des CpG-Gehaltes die Antigenproduktion negativ beeinflusst. Denkbar ist auch, dass neben einer möglichen Beteiligung des Aktivatorproteins noch andere Faktoren bei der CpG-regulierten Beeinflussung des Transkriptionsniveaus eine Rolle spielen.

D.5.2 Ausblick

Um die Theorie eines an unmethylierte CpG-Dinukleotide bindenden Transkriptionsaktivators zu erhärten, müsste zunächst der im Vektorrückgrat sowie im Transgen vorherrschende Methylierungsstatus ermittelt werden, um zu zeigen, dass die Bindungsvoraussetzungen für den Transkriptionsaktivator gegeben sind. Am Beispiel stabiler Zelllinien kann dies mit Hilfe der Bisulfit-Sequenzierungsmethode erfolgen. Unmethylierte Cytosine werden dabei zu Uracil deaminiert, während methylierte Cytosine unverändert bleiben (Frommer *et al.*, 1992). Auch durch Einsatz von 5-aza-2'-Deoxycytidin, welches Methyltransferasen inhibiert (Luo *et al.*, 2006), kann eine Beteiligung des Methylierungsprozesses an der beobachteten CpG-basierten Transkriptionsregulation ausgeschlossen werden. Weiterhin sollte durch Ausschalten dieses Transaktivators z.B. durch si-RNAs der beobachtete CpG-Effekt gemildert bzw. ausgeschaltet werden können. Darüber hinaus könnte die Berücksichtigung optimaler Bindemotive für CGBP (Lee *et al.*, 2001) bei der Optimierung von Transgensequenzen Aufschluss über eine Beteiligung dieses Proteins an der Modifizierung der Expressionshöhe geben. Auch eine Titration des CpG-Gehaltes in der kodierenden Region verschiedener Transgene kann Informationen über mögliche Sättigungseffekte der Transkriptionsaktivierung via CGBP liefern. Neben genetischen Untersuchungen können auch biochemische Experimente zur Aufklärung der Physiologie des Transaktivatorproteins beitragen.

Darüber hinaus sollte eine weitere CpG-Minimierung der Vektorkonstrukte angestrebt werden. Dazu könnte z.B. die CpG-freie Version des Polyadenylierungssignals BGH_ATA in das Vektormodul eingesetzt werden. Die in dieser Arbeit etablierten, CpG-freien Promotorvarianten könnten mit Elementen des in der Literatur beschriebenen CpG-freien Promotors (Yew *et al.*, 2002) kombiniert werden, um die Promotorstärke zu erhöhen. Anschließend könnte das resultierende Promotorkonstrukt ebenfalls in die Vektorplattform integriert werden. Auch eine chimäre Promotorvariante aus partiell CpG-depletierten CMV-Konstrukten, die eine dem herkömmlichen Konstrukt vergleichbare Expression hervorrufen, und Teilen des etwas schwächer exprimierenden CMV_TG-Promotors wäre denkbar. Zudem könnten die bestehenden Vektormodule zur

Steigerung der Proteinausbeuten mit konventionellen, expressionssteigernden Elementen, wie unter D.1.1 beschrieben, versehen werden.

Ungeachtet des zugrunde liegenden Mechanismus stellt die CpG-Modifizierung in Vektorrückgrat und Transgen eine effektive Strategie zur Verbesserung der Transgenexpression dar. Daher sollte dieses Vorgehen bei der rationalen Konzeption DNA-basierter Therapeutika sowie bei der effektiven Produktion rekombinanter Proteine zukünftig in Betracht gezogen werden. Durch Kombination CpG-reduzierter Plasmide mit bestimmten immunmodulatorischen Elementen können schließlich anwendungsspezifische Sequenzanforderungen an das Vektorrückgrat evaluiert und definiert werden.

E Literaturverzeichnis

- Agirre, X., Vizmanos, J.L., Calasanz, M.J., Garcia-Delgado, M., Larrayoz, M.J., and Novo, F.J.** (2003): Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients, *Oncogene* 22: 1070-1072
- Akira, S., Takeda, K., and Kaisho, T.** (2001): Toll-like receptors: critical proteins linking innate and acquired immunity, *Nat.Immunol.* 2: 675-680
- Amir, R.E., Van, d.V., I, Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y.** (1999): Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2, *Nat.Genet.* 23: 185-188
- Anderson, K.V., Bokla, L., and Nusslein-Volhard, C.** (1985): Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product, *Cell* 42: 791-798
- Antequera, F. and Bird, A.** (1993): CpG islands, *EXS* 64: 169-185
- Antequera, F. and Bird, A.** (1993): Number of CpG islands and genes in human and mouse, *Proc.Natl.Acad.Sci.U.S.A* 90: 11995-11999
- Attwood, J.T., Yung, R.L., and Richardson, B.C.** (2002): DNA methylation and the regulation of gene transcription, *Cell Mol.Life Sci.* 59: 241-257
- Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. & Struhl, K.** (1987): Current protocols in molecular biology, John Wiley & Sons, Inc., New York
- Azhikina, T.L. and Sverdlov, E.D.** (2005): Study of tissue-specific CpG methylation of DNA in extended genomic loci, *Biochemistry (Mosc.)* 70: 596-603
- Bachman, K.E., Rountree, M.R., and Baylin, S.B.** (2001): Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin, *J.Biol.Chem.* 276: 32282-32287
- Ballestar, E., Pile, L.A., Wassarman, D.A., Wolffe, A.P., and Wade, P.A.** (2001): A *Drosophila* MBD family member is a transcriptional corepressor associated with specific genes, *Eur.J.Biochem.* 268: 5397-5406
- Ballestar, E. and Wolffe, A.P.** (2001): Methyl-CpG-binding proteins. Targeting specific gene repression, *Eur.J.Biochem.* 268: 1-6
- Bauer, S., Kirschning, C.J., Hacker, H., Redecke, V., Hausmann, S., Akira, S., Wagner, H., and Lipford, G.B.** (2001): Human TLR9 confers responsiveness to bacterial DNA via species-specific CpG motif recognition, *Proc.Natl.Acad.Sci.U.S.A* 98: 9237-9242
- Baylin, S.B., Esteller, M., Rountree, M.R., Bachman, K.E., Schuebel, K., and Herman, J.G.** (2001): Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer, *Hum.Mol.Genet.* 10: 687-692
- Bebbington, C.R., Renner, G., Thomson, S., King, D., Abrams, D., and Yarranton, G.T.** (1992): High-level expression of a recombinant antibody from myeloma cells using a glutamine synthetase gene as an amplifiable selectable marker, *Biotechnology (N.Y.)* 10: 169-175
- Belvin, M.P. and Anderson, K.V.** (1996): A conserved signaling pathway: the *Drosophila* toll-dorsal pathway, *Annu.Rev.Cell Dev.Biol.* 12: 393-416
- Berger, J. and Bird, A.** (2005): Role of MBD2 in gene regulation and tumorigenesis, *Biochem.Soc.Trans.* 33: 1537-1540
- Bestor, T.H., Gunderson, G., Kolsto, A.B., and Prydz, H.** (1992): CpG islands in mammalian gene promoters are inherently resistant to de novo methylation, *Genet.Anal.Tech.Appl.* 9: 48-53
- Bestor, T.H. and Verdine, G.L.** (1994): DNA methyltransferases, *Curr.Opin.Cell Biol.* 6: 380-389
- Bestor, T.H.** (2000): The DNA methyltransferases of mammals, *Hum.Mol.Genet.* 9: 2395-2402
- Bestor, T.H.** (2000): Gene silencing as a threat to the success of gene therapy, *J.Clin.Invest* 105: 409-411
- Bird, A.P.** (1980): DNA methylation and the frequency of CpG in animal DNA, *Nucleic Acids Res.* 8: 1499-1504
- Bird, A.P.** (1986): CpG-rich islands and the function of DNA methylation, *Nature* 321: 209-213
- Bird, A.P., Taggart, M.H., Nicholls, R.D., and Higgs, D.R.** (1987): Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene, *EMBO J.* 6: 999-1004
- Bird, A.P.** (1995): Gene number, noise reduction and biological complexity, *Trends Genet.* 11: 94-100
- Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B., and Schaffner, W.** (1985): A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus, *Cell* 41: 521-530
- Bradford, M.M.** (1976): A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal.Biochem.* 72: 248-254
- Brazolot Millan, C.L., Weeratna, R., Krieg, A.M., Siegrist, C.A., and Davis, H.L.** (1998): CpG DNA can induce strong Th1 humoral and cell-mediated immune responses against hepatitis B surface antigen in young mice, *Proc.Natl.Acad.Sci.U.S.A* 95: 15553-15558
- Brooks, A.R., Harkins, R.N., Wang, P., Qian, H.S., Liu, P., and Rubanyi, G.M.** (2004): Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle, *J.Gene Med.* 6: 395-404
- Carlone, D.L. and Skalnik, D.G.** (2001): CpG binding protein is crucial for early embryonic development, *Mol.Cell Biol.* 21: 7601-7606
- Carlone, D.L., Lee, J.H., Young, S.R., Dobrota, E., Butler, J.S., Ruiz, J., and Skalnik, D.G.** (2005): Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein, *Mol.Cell Biol.* 25: 4881-4891

- Cervoni, N. and Szyf, M.** (2001): Demethylase activity is directed by histone acetylation, *J.Biol.Chem.* 276: 40778-40787
- Cesareni, G., Muesing, M.A., and Polisky, B.** (1982): Control of ColE1 DNA replication: the rop gene product negatively affects transcription from the replication primer promoter, *Proc.Natl.Acad.Sci.U.S.A* 79: 6313-6317
- Chevalier-Mariette, C., Henry, I., Montfort, L., Capgras, S., Forlani, S., Muschler, J., and Nicolas, J.F.** (2003): CpG content affects gene silencing in mice: evidence from novel transgenes, *Genome Biol.* 4: R53
- Clark, S.J., Harrison, J., and Molloy, P.L.** (1997): Sp1 binding is inhibited by (m)Cp(m)CpG methylation, *Gene* 195: 67-71
- Clark, S.J. and Melki, J.** (2002): DNA methylation and gene silencing in cancer: which is the guilty party?, *Oncogene* 21: 5380-5387
- Coley, W.B.** (1991): The treatment of malignant tumors by repeated inoculations of erysipelas. With a report of ten original cases. 1893, *Clin.Orthop.Relat Res.* 3-11
- Craig, N.L.** (1988): The mechanism of conservative site-specific recombination, *Annu.Rev.Genet.* 22: 77-105
- Cross, S.H. and Bird, A.P.** (1995): CpG islands and genes, *Curr.Opin.Genet.Dev.* 5: 309-314
- Dalle, B., Rubin, J.E., Alkan, O., Sukonnik, T., Pasceri, P., Yao, S., Pawliuk, R., Leboulch, P., and Ellis, J.** (2005): eGFP reporter genes silence LCRbeta-globin transgene expression via CpG dinucleotides, *Mol.Ther.* 11: 591-599
- Dang, Q., Auten, J., and Plavec, I.** (2000): Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector, *J.Virol.* 74: 2671-2678
- Davis, H.L., Weeratna, R., Waldschmidt, T.J., Tygrett, L., Schorr, J., and Krieg, A.M.** (1998): CpG DNA is a potent enhancer of specific immunity in mice immunized with recombinant hepatitis B surface antigen, *J.Immunol.* 160: 870-876
- Davis, H.L., Suparto, I.I., Weeratna, R.R., Jumintarto, Iskandriati, D.D., Chamzah, S.S., Ma'ruf, A.A., Nente, C.C., Pawitri, D.D., Krieg, A.M., Heriyanto, Smits, W., and Sajuthi, D.D.** (2000): CpG DNA overcomes hyporesponsiveness to hepatitis B vaccine in orangutans, *Vaccine* 18: 1920-1924
- De Smet, C., De Backer, O., Faraoni, I., Lurquin, C., Bresseur, F., and Boon, T.** (1996): The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation, *Proc.Natl.Acad.Sci.U.S.A* 93: 7149-7153
- De Smet, C., Lurquin, C., Lethe, B., Martelange, V., and Boon, T.** (1999): DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter, *Mol.Cell Biol.* 19: 7327-7335
- Deml, L., Bojak, A., Steck, S., Graf, M., Wild, J., Schirmbeck, R., Wolf, H., and Wagner, R.** (2001): Multiple effects of codon usage optimization on expression and immunogenicity of dna candidate vaccines encoding the human immunodeficiency virus type 1 gag protein, *J.Virol.* 75: 10991-11001
- Deng, G., Chen, A., Pong, E., and Kim, Y.S.** (2001): Methylation in hMLH1 promoter interferes with its binding to transcription factor CBF and inhibits gene expression, *Oncogene* 20: 7120-7127
- Duan, J. and Antezana, M.A.** (2003): Mammalian mutation pressure, synonymous codon choice, and mRNA degradation, *J.Mol.Evol.* 57: 694-701
- Eleaume, H. and Jabbouri, S.** (2004): Comparison of two standardisation methods in real-time quantitative RT-PCR to follow *Staphylococcus aureus* genes expression during in vitro growth, *J.Microbiol.Methods* 59: 363-370
- Fang, J.Y., Zhu, S.S., Xiao, S.D., Jiang, S.J., Shi, Y., Chen, X.Y., Zhou, X.M., and Qian, L.F.** (1996): Studies on the hypomethylation of c-myc, c-Ha-ras oncogenes and histopathological changes in human gastric carcinoma, *J.Gastroenterol.Hepatol.* 11: 1079-1082
- Ferguson-Smith, A.C., Sasaki, H., Cattanaach, B.M., and Surani, M.A.** (1993): Parental-origin-specific epigenetic modification of the mouse H19 gene, *Nature* 362: 751-755
- Finnegan, E.J., Peacock, W.J., and Dennis, E.S.** (2000): DNA methylation, a key regulator of plant development and other processes, *Curr.Opin.Genet.Dev.* 10: 217-223
- Foecking, M.K. and Hofstetter, H.** (1986): Powerful and versatile enhancer-promoter unit for mammalian expression vectors, *Gene* 45: 101-105
- Freeman, W.M., Walker, S.J., and Vrana, K.E.** (1999): Quantitative RT-PCR: pitfalls and potential, *Biotechniques* 26: 112-115
- Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L.** (1992): A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands, *Proc.Natl.Acad.Sci.U.S.A* 89: 1827-1831
- Fryxell, K.J. and Moon, W.J.** (2005): CpG mutation rates in the human genome are highly dependent on local GC content, *Mol.Biol.Evol.* 22: 650-658
- Fujii, G., Nakamura, Y., Tsukamoto, D., Ito, M., Shiba, T., and Takamatsu, N.** (2006): CpG methylation at the USF-binding site is important for the liver-specific transcription of the chipmunk HP-27 gene, *Biochem.J.* 395: 203-209
- Fuks, F., Burgers, W.A., Godin, N., Kasai, M., and Kouzarides, T.** (2001): Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription, *EMBO J.* 20: 2536-2544
- Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T.** (2003): The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation, *J.Biol.Chem.* 278: 4035-4040
- Gardiner-Garden, M. and Frommer, M.** (1987): CpG islands in vertebrate genomes, *J.Mol.Biol.* 196: 261-282
- Girard, M.P., Osmanov, S.K., and Kieny, M.P.** (2006): A review of vaccine research and development: The human immunodeficiency virus (HIV), *Vaccine*
- Goodwin, E.C. and Rottman, F.M.** (1992): The 3'-flanking sequence of the bovine growth hormone gene contains novel elements required for efficient and accurate polyadenylation, *J.Biol.Chem.* 267: 16330-16334

- Graf, M., Bojak, A., Deml, L., Bieler, K., Wolf, H., and Wagner, R.** (2000): Concerted action of multiple cis-acting sequences is required for Rev dependence of late human immunodeficiency virus type 1 gene expression, *J.Virol.* 74: 10822-10826
- Graf, M., Deml, L., and Wagner, R.** (2004): Codon-optimized genes that enable increased heterologous expression in mammalian cells and elicit efficient immune responses in mice after vaccination of naked DNA, *Methods Mol.Med.* 94:197-210.: 197-210
- Graham, F.L. and Eb, A.J.** (1973): A new technique for the assay of infectivity of human adenovirus 5 DNA, *Virology* 52: 456-467
- Gundersen, G., Kolsto, A.B., Larsen, F., and Prydz, H.** (1992): Tissue-specific methylation of a CpG island in transgenic mice, *Gene* 113: 207-214
- Guo, Z.S., Wang, L.H., Eisensmith, R.C., and Woo, S.L.** (1996): Evaluation of promoter strength for hepatic gene expression in vivo following adenovirus-mediated gene transfer, *Gene Ther.* 3: 802-810
- Han, J., Zhu, Z., Hsu, C., and Finley, W.H.** (1994): Selection of antisense oligonucleotides on the basis of genomic frequency of the target sequence, *Antisense Res.Dev.* 4: 53-65
- Hanahan, D.** (1983): Studies on transformation of *Escherichia coli* with plasmids, *J.Mol.Biol.* 166: 557-580
- Hartmann, G., Weiner, G.J., and Krieg, A.M.** (1999): CpG DNA: a potent signal for growth, activation, and maturation of human dendritic cells, *Proc.Natl.Acad.Sci.U.S.A* 96: 9305-9310
- Hartmann, G. and Krieg, A.M.** (2000): Mechanism and function of a newly identified CpG DNA motif in human primary B cells, *J.Immunol.* 164: 944-953
- Heap, D.M., Herrmann, M.G. and Wittwer, C.T.** (2000): Nuclear Run-On assay using biotin labeling, magnetic bead capture and analysis by fluorescence-based RT-PCR, *BioTechniques* 29: 1012-1017
- Heckelsmiller, K., Rall, K., Beck, S., Schlamp, A., Seiderer, J., Jahrsdorfer, B., Krug, A., Rothenfusser, S., Endres, S., and Hartmann, G.** (2002): Peritumoral CpG DNA elicits a coordinated response of CD8 T cells and innate effectors to cure established tumors in a murine colon carcinoma model, *J.Immunol.* 169: 3892-3899
- Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., and Akira, S.** (2000): A Toll-like receptor recognizes bacterial DNA, *Nature* 408: 740-745
- Hendrich, B. and Bird, A.** (1998): Identification and characterization of a family of mammalian methyl-CpG binding proteins, *Mol.Cell Biol.* 18: 6538-6547
- Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J., and Bird, A.** (1999): The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites, *Nature* 401: 301-304
- Hisano, M., Ohta, H., Nishimune, Y., and Nozaki, M.** (2003): Methylation of CpG dinucleotides in the open reading frame of a testicular germ cell-specific intronless gene, *Tact1/Act17b*, represses its expression in somatic cells, *Nucleic Acids Res.* 31: 4797-4804
- Hodges, B.L., Taylor, K.M., Joseph, M.F., Bourgeois, S.A., and Scheule, R.K.** (2004): Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides, *Mol.Ther.* 10: 269-278
- Horner, A.A., Datta, S.K., Takabayashi, K., Belyakov, I.M., Hayashi, T., Cinman, N., Nguyen, M.D., Van Uden, J.H., Berzofsky, J.A., Richman, D.D., and Raz, E.** (2001): Immunostimulatory DNA-based vaccines elicit multifaceted immune responses against HIV at systemic and mucosal sites, *J.Immunol.* 167: 1584-1591
- Horner, A.A., Van Uden, J.H., Zubeldia, J.M., Broide, D., and Raz, E.** (2001): DNA-based immunotherapeutics for the treatment of allergic disease, *Immunol.Rev.* 179: 102-118
- Hsieh, C.L.** (1994): Dependence of transcriptional repression on CpG methylation density, *Mol.Cell Biol.* 14: 5487-5494
- Iguchi-Ariga, S.M. and Schaffner, W.** (1989): CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation, *Genes Dev.* 3: 612-619
- Ikemura, T.** (1982): Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. Differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs, *J.Mol.Biol.* 158: 573-597
- Ikemura, T.** (1985): Codon usage and tRNA content in unicellular and multicellular organisms, *Mol.Biol.Evol.* 2: 13-34
- Itoh, T. and Tomizawa, J.** (1979): Initiation of replication of plasmid ColE1 DNA by RNA polymerase, ribonuclease H, and DNA polymerase I, *Cold Spring Harb.Symp.Quant.Biol.* 43 Pt 1: 409-417
- Jenuwein, T.** (2001): Re-SET-ting heterochromatin by histone methyltransferases, *Trends Cell Biol.* 11: 266-273
- Jones, P.A. and Laird, P.W.** (1999): Cancer epigenetics comes of age, *Nat.Genet.* 21: 163-167
- Jones, P.A.** (1999): The DNA methylation paradox, *Trends Genet.* 15: 34-37
- Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P.** (1998): Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, *Nat.Genet.* 19: 187-191
- Jones, T.R., Obaldia, N., III, Gramzinski, R.A., Charoenvit, Y., Kolodny, N., Kitov, S., Davis, H.L., Krieg, A.M., and Hoffman, S.L.** (1999): Synthetic oligodeoxynucleotides containing CpG motifs enhance immunogenicity of a peptide malaria vaccine in Aotus monkeys, *Vaccine* 17: 3065-3071
- Jung, Y.H. and Lee, Y.** (1995): RNases in ColE1 DNA metabolism, *Mol.Biol.Rep.* 22: 195-200
- Kadowaki, N., Ho, S., Antonenko, S., Malefyt, R.W., Kastelein, R.A., Bazan, F., and Liu, Y.J.** (2001): Subsets of human dendritic cell precursors express different toll-like receptors and respond to different microbial antigens, *J.Exp.Med.* 194: 863-869
- Kandimalla, E.R., Zhu, F.G., Bhagat, L., Yu, D., and Agrawal, S.** (2003): Toll-like receptor 9: modulation of recognition and cytokine induction by novel synthetic CpG DNAs, *Biochem.Soc.Trans.* 31: 654-658

- Karpf, A.R. and Jones, D.A.** (2002): Reactivating the expression of methylation silenced genes in human cancer, *Oncogene* 21: 5496-5503
- Kellogg, D.E., Rybalkin, I., Chen, S., Mukhamedova, N., Vlasik, T., Siebert, P.D., and Chenchik, A.** (1994): TaqStart Antibody: "hot start" PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase, *Biotechniques* 16: 1134-1137
- Kline, J.N., Waldschmidt, T.J., Businga, T.R., Lemish, J.E., Weinstock, J.V., Thorne, P.S., and Krieg, A.M.** (1998): Modulation of airway inflammation by CpG oligodeoxynucleotides in a murine model of asthma, *J.Immunol.* 160: 2555-2559
- Kline, J.N., Kitagaki, K., Businga, T.R., and Jain, V.V.** (2002): Treatment of established asthma in a murine model using CpG oligodeoxynucleotides, *Am.J.Physiol Lung Cell Mol.Physiol* 283: L170-L179
- Kline, J.N.** (2002): DNA therapy for asthma, *Curr.Opin.Allergy Clin.Immunol.* 2: 69-73
- Klinman, D.M., Takeshita, F., Gursel, I., Leifer, C., Ishii, K.J., Verthelyi, D., and Gursel, M.** (2002): CpG DNA: recognition by and activation of monocytes, *Microbes.Infect.* 4: 897-901
- Klinman, D.M., Currie, D., Gursel, I., and Verthelyi, D.** (2004): Use of CpG oligodeoxynucleotides as immune adjuvants, *Immunol.Rev.* 199: 201-216
- Klinman, D.M., Xie, H., Little, S.F., Currie, D., and Ivins, B.E.** (2004): CpG oligonucleotides improve the protective immune response induced by the anthrax vaccination of rhesus macaques, *Vaccine* 22: 2881-2886
- Klinman, D.M.** (2004): Use of CpG oligodeoxynucleotides as immunoprotective agents, *Expert.Opin.Biol.Ther.* 4: 937-946
- Klinman, D.M.** (2004): Immunotherapeutic uses of CpG oligodeoxynucleotides, *Nat.Rev.Immunol.* 4: 249-258
- Klose, R.J. and Bird, A.P.** (2006): Genomic DNA methylation: the mark and its mediators, *Trends Biochem.Sci.* 31: 89-97
- Kojima, Y., Xin, K.Q., Ooki, T., Hamajima, K., Oikawa, T., Shinoda, K., Ozaki, T., Hoshino, Y., Jounai, N., Nakazawa, M., Klinman, D., and Okuda, K.** (2002): Adjuvant effect of multi-CpG motifs on an HIV-1 DNA vaccine, *Vaccine* 20: 2857-2865
- Kothari, S., Baillie, J., Sissons, J.G., and Sinclair, J.H.** (1991): The 21bp repeat element of the human cytomegalovirus major immediate early enhancer is a negative regulator of gene expression in undifferentiated cells, *Nucleic Acids Res.* 19: 1767-1771
- Krieg, A.M., Yi, A.K., Matson, S., Waldschmidt, T.J., Bishop, G.A., Teasdale, R., Koretzky, G.A., and Klinman, D.M.** (1995): CpG motifs in bacterial DNA trigger direct B-cell activation, *Nature* 374: 546-549
- Krieg, A.M., Wu, T., Weeratna, R., Efler, S.M., Love-Homan, L., Yang, L., Yi, A.K., Short, D., and Davis, H.L.** (1998): Sequence motifs in adenoviral DNA block immune activation by stimulatory CpG motifs, *Proc.Natl.Acad.Sci.U.S.A* 95: 12631-12636
- Krieg, A.M., Yi, A.K., Schorr, J., and Davis, H.L.** (1998): The role of CpG dinucleotides in DNA vaccines, *Trends Microbiol.* 6: 23-27
- Krieg, A.M., Love-Homan, L., Yi, A.K., and Harty, J.T.** (1998): CpG DNA induces sustained IL-12 expression in vivo and resistance to *Listeria monocytogenes* challenge, *J.Immunol.* 161: 2428-2434
- Krieg, A.M.** (2002): CpG motifs in bacterial DNA and their immune effects, *Annu.Rev.Immunol.* 20: 709-760
- Krieg, A.M.** (2002): From A to Z on CpG, *Trends Immunol.* 23: 64-65
- Krug, A., Towarowski, A., Britsch, S., Rothenfusser, S., Hornung, V., Bals, R., Giese, T., Engelmann, H., Endres, S., Krieg, A.M., and Hartmann, G.** (2001): Toll-like receptor expression reveals CpG DNA as a unique microbial stimulus for plasmacytoid dendritic cells which synergizes with CD40 ligand to induce high amounts of IL-12, *Eur.J.Immunol.* 31: 3026-3037
- Kudo, S.** (1998): Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated, *Mol.Cell Biol.* 18: 5492-5499
- Kues, U. and Stahl, U.** (1989): Replication of plasmids in gram-negative bacteria, *Microbiol.Rev.* 53: 491-516
- Laemmli, U.K.** (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227: 680-685
- Larsen, F., Gundersen, G., Lopez, R., and Prydz, H.** (1992): CpG islands as gene markers in the human genome, *Genomics* 13: 1095-1107
- Leclerc, G.J., Leclerc, G.M., and Barredo, J.C.** (2002): Real-time RT-PCR analysis of mRNA decay: half-life of Beta-actin mRNA in human leukemia CCRF-CEM and Nalm-6 cell lines, *Cancer Cell Int.* 2: 1
- Lee, J.H., Voo, K.S., and Skalnik, D.G.** (2001): Identification and characterization of the DNA binding domain of CpG-binding protein, *J.Biol.Chem.* 276: 44669-44676
- Lee, J.H. and Skalnik, D.G.** (2005): CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex, *J.Biol.Chem.* 280: 41725-41731
- Li, Q.L., Kim, H.R., Kim, W.J., Choi, J.K., Lee, Y.H., Kim, H.M., Li, L.S., Kim, H., Chang, J., Ito, Y., Youl, L.K., and Bae, S.C.** (2004): Transcriptional silencing of the RUNX3 gene by CpG hypermethylation is associated with lung cancer, *Biochem.Biophys.Res.Commun.* 314: 223-228
- Lipford, G.B., Bauer, M., Blank, C., Reiter, R., Wagner, H., and Heeg, K.** (1997): CpG-containing synthetic oligonucleotides promote B and cytotoxic T cell responses to protein antigen: a new class of vaccine adjuvants, *Eur.J.Immunol.* 27: 2340-2344
- Lipford, G.B., Sparwasser, T., Zimmermann, S., Heeg, K., and Wagner, H.** (2000): CpG-DNA-mediated transient lymphadenopathy is associated with a state of Th1 predisposition to antigen-driven responses, *J.Immunol.* 165: 1228-1235

- Liu, B. and Stinski, M.F.** (1992): Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements, *J.Virol.* 66: 4434-4444
- Liu, H.S., Jan, M.S., Chou, C.K., Chen, P.H., and Ke, N.J.** (1999): Is green fluorescent protein toxic to the living cells?, *Biochem.Biophys.Res.Commun.* 260: 712-717
- Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J.E.** (2001): *Molekulare Zellbiologie*, 4. Auflage, Spektrum Akademischer Verlag, Heidelberg, Berlin
- Löser, P., Jennings, G.S., Strauss, M., and Sandig, V.** (1998): Reactivation of the previously silenced cytomegalovirus major immediate-early promoter in the mouse liver: involvement of NFkappaB, *J.Virol.* 72: 180-190
- Lu, A., Gupta, A., Li, C., Ahlborn, T.E., Ma, Y., Shi, E.Y., and Liu, J.** (2001): Molecular mechanisms for aberrant expression of the human breast cancer specific gene 1 in breast cancer cells: control of transcription by DNA methylation and intronic sequences, *Oncogene* 20: 5173-5185
- Luo, W., Wang, X., Kageshita, T., Wakasugi, S., Karpf, A.R., and Ferrone, S.** (2006): Regulation of high molecular weight-melanoma associated antigen (HMW-MAA) gene expression by promoter DNA methylation in human melanoma cells, *Oncogene*
- Ma, X., Forns, X., Gutierrez, R., Mushahwar, I.K., Wu, T., Payette, P.J., Bukh, J., Purcell, R.H., and Davis, H.L.** (2002): DNA-based vaccination against hepatitis C virus (HCV): effect of expressing different forms of HCV E2 protein and use of CpG-optimized vectors in mice, *Vaccine* 20: 3263-3271
- Medzhitov, R., Preston-Hurlburt, P., and Janeway, C.A., Jr.** (1997): A human homologue of the Drosophila Toll protein signals activation of adaptive immunity, *Nature* 388: 394-397
- Meehan, R., Lewis, J., Cross, S., Nan, X., Jeppesen, P., and Bird, A.** (1992): Transcriptional repression by methylation of CpG, *J.Cell Sci.Suppl* 16: 9-14
- Meehan, R.R., Lewis, J.D., and Bird, A.P.** (1992): Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA, *Nucleic Acids Res.* 20: 5085-5092
- Meier, J.L. and Stinski, M.F.** (1996): Regulation of human cytomegalovirus immediate-early gene expression, *Intervirology* 39: 331-342
- Melki, J.R. and Clark, S.J.** (2002): DNA methylation changes in leukaemia, *Semin.Cancer Biol.* 12: 347-357
- Merlin, S. and Polisky, B.** (1995): Assessment of quantitative models for plasmid ColE1 copy number control, *J.Mol.Biol.* 248: 211-219
- Mitsudomi, T., Viallet, J., Mulshine, J.L., Linnoila, R.I., Minna, J.D., and Gazdar, A.F.** (1991): Mutations of ras genes distinguish a subset of non-small-cell lung cancer cell lines from small-cell lung cancer cell lines, *Oncogene* 6: 1353-1362
- Moens, U., Van Ghelue, M., Kristoffersen, A.K., Johansen, B., Rekvig, O.P., Degre, M., and Rollag, H.** (2001): Simian virus 40 large T-antigen, but not small T-antigen, trans-activates the human cytomegalovirus major immediate early promoter, *Virus Genes* 23: 215-226
- Molloy, P.L. and Watt, F.** (1990): DNA methylation and specific protein-DNA interactions, *Philos.Trans.R.Soc.Lond B Biol.Sci.* 326: 267-275
- Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A.** (1998): Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, *Nature* 393: 386-389
- Nelson, W.G., De Marzo, A.M., and Isaacs, W.B.** (2003): Prostate cancer, *N.Engl.J.Med.* 349: 366-381
- Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A.** (1999): MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex, *Nat.Genet.* 23: 58-61
- Niller, H.H. and Hennighausen, L.** (1991): Formation of several specific nucleoprotein complexes on the human cytomegalovirus immediate early enhancer, *Nucleic Acids Res.* 19: 3715-3721
- Niwa, M., Rose, S.D., and Berget, S.M.** (1990): In vitro polyadenylation is stimulated by the presence of an upstream intron, *Genes Dev.* 4: 1552-1559
- Okano, M., Xie, S., and Li, E.** (1998): Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases, *Nat.Genet.* 19: 219-220
- Okano, M., Bell, D.W., Haber, D.A., and Li, E.** (1999): DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, *Cell* 99: 247-257
- Phi-Van, L., von Kries, J.P., Ostertag, W., and Stratling, W.H.** (1990): The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes, *Mol.Cell Biol.* 10: 2302-2307
- Ponder, K.P., Dunbar, R.P., Wilson, D.R., Darlington, G.J., and Woo, S.L.** (1991): Evaluation of relative promoter strength in primary hepatocytes using optimized lipofection, *Hum.Gene Ther.* 2: 41-52
- Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E.** (2001): The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor, *Genes Dev.* 15: 1613-1618
- Prokhortchouk, E. and Hendrich, B.** (2002): Methyl-CpG binding proteins and cancer: are MeCpGs more important than MBDs?, *Oncogene* 21: 5394-5399
- Proudfoot, N.J.** (1989): How RNA polymerase II terminates transcription in higher eukaryotes, *Trends Biochem.Sci.* 14: 105-110
- Rader, R.A.** (2003): *Biopharma: Biopharmaceutical products in the U.S. market*, 2nd edition, Biotechnology Information Institute, Rockville, MD

- Reyes-Sandoval, A. and Ertl, H.C.** (2004): CpG methylation of a plasmid vector results in extended transgene product expression by circumventing induction of immune responses, *Mol. Ther.* 9: 249-261
- Riggs, A.D. and Pfeifer, G.P.** (1992): X-chromosome inactivation and cell memory, *Trends Genet.* 8: 169-174
- Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P.** (2000): DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters, *Nat. Genet.* 25: 338-342
- Robertson, K.D.** (2001): DNA methylation, methyltransferases, and cancer, *Oncogene* 20: 3139-3155
- Robertson, K.D.** (2002): DNA methylation and chromatin - unraveling the tangled web, *Oncogene* 21: 5361-5379
- Rock, F.L., Hardiman, G., Timans, J.C., Kastelein, R.A., and Bazan, J.F.** (1998): A family of human receptors structurally related to *Drosophila* Toll, *Proc. Natl. Acad. Sci. U.S.A.* 95: 588-593
- Roman, M., Martin-Orozco, E., Goodman, J.S., Nguyen, M.D., Sato, Y., Ronaghy, A., Kornbluth, R.S., Richman, D.D., Carson, D.A., and Raz, E.** (1997): Immunostimulatory DNA sequences function as T helper-1-promoting adjuvants, *Nat. Med.* 3: 849-854
- Rountree, M.R., Bachman, K.E., and Baylin, S.B.** (2000): DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci, *Nat. Genet.* 25: 269-277
- Rountree, M.R., Bachman, K.E., Herman, J.G., and Baylin, S.B.** (2001): DNA methylation, chromatin inheritance, and cancer, *Oncogene* 20: 3156-3165
- Sambrook, J., Fritsch, E.F., and Maniatis, T.** (1989): *Molecular cloning: A laboratory manual*, 2:
- Sambucetti, L.C., Cherrington, J.M., Wilkinson, G.W., and Mocarski, E.S.** (1989): NF-kappa B activation of the cytomegalovirus enhancer is mediated by a viral transactivator and by T cell stimulation, *EMBO J.* 8: 4251-4258
- Sanger, F., Nicklen, S., and Coulson, A.R.** (1977): DNA sequencing with chain-terminating inhibitors, *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463-5467
- Sarraf, S.A. and Stancheva, I.** (2004): Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly, *Mol. Cell* 15: 595-605
- Sato, Y., Roman, M., Tighe, H., Lee, D., Corr, M., Nguyen, M.D., Silverman, G.J., Lotz, M., Carson, D.A., and Raz, E.** (1996): Immunostimulatory DNA sequences necessary for effective intradermal gene immunization, *Science* 273: 352-354
- Sauer, B.** (1994): Site-specific recombination: developments and applications, *Curr. Opin. Biotechnol.* 5: 521-527
- Saxonov, S., Berg, P., and Brutlag, D.L.** (2006): A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters, *Proc. Natl. Acad. Sci. U.S.A.* 103: 1412-1417
- Sharp, P.M. and Li, W.H.** (1987): The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications, *Nucleic Acids Res.* 15: 1281-1295
- Shen, J.C., Rideout, W.M., III, and Jones, P.A.** (1994): The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA, *Nucleic Acids Res.* 22: 972-976
- Sherman, J.D. and Talbert, L.E.** (2002): Vernalization-induced changes of the DNA methylation pattern in winter wheat, *Genome* 45: 253-260
- Silvera, P., Savary, J.R., Livingston, V., White, J., Manson, K.H., Wyand, M.H., Salk, P.L., Moss, R.B., and Lewis, M.G.** (2004): Vaccination with gp120-depleted HIV-1 plus immunostimulatory CpG oligodeoxynucleotides in incomplete Freund's adjuvant stimulates cellular and humoral immunity in rhesus macaques, *Vaccine* 23: 827-839
- Sparwasser, T., Koch, E.S., Vabulas, R.M., Heeg, K., Lipford, G.B., Ellwart, J.W., and Wagner, H.** (1998): Bacterial DNA and immunostimulatory CpG oligonucleotides trigger maturation and activation of murine dendritic cells, *Eur. J. Immunol.* 28: 2045-2054
- Stacey, K.J., Sweet, M.J., and Hume, D.A.** (1996): Macrophages ingest and are activated by bacterial DNA, *J. Immunol.* 157: 2116-2122
- Stinski, M.F. and Roehr, T.J.** (1985): Activation of the major immediate early gene of human cytomegalovirus by cis-acting elements in the promoter-regulatory sequence and by virus-specific trans-acting components, *J. Virol.* 55: 431-441
- Sun, S., Beard, C., Jaenisch, R., Jones, P., and Sprent, J.** (1997): Mitogenicity of DNA from different organisms for murine B cells, *J. Immunol.* 159: 3119-3125
- Sved, J. and Bird, A.** (1990): The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model, *Proc. Natl. Acad. Sci. U.S.A.* 87: 4692-4696
- Takai, D. and Jones, P.A.** (2002): Comprehensive analysis of CpG islands in human chromosomes 21 and 22, *Proc. Natl. Acad. Sci. U.S.A.* 99: 3740-3745
- Tamaru, H. and Selker, E.U.** (2001): A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*, *Nature* 414: 277-283
- Tanaka, Y., Kurumizaka, H., and Yokoyama, S.** (2005): CpG methylation of the CENP-B box reduces human CENP-B binding, *FEBS J.* 272: 282-289
- Tokunaga, T., Yamamoto, H., Shimada, S., Abe, H., Fukuda, T., Fujisawa, Y., Furutani, Y., Yano, O., Kataoka, T., Sudo, T., and .** (1984): Antitumor activity of deoxyribonucleic acid fraction from *Mycobacterium bovis* BCG. I. Isolation, physicochemical characterization, and antitumor activity, *J. Natl. Cancer Inst.* 72: 955-962
- Ulmer, J.B., Donnelly, J.J., Parker, S.E., Rhodes, G.H., Felgner, P.L., Dworki, V.J., Gromkowski, S.H., Deck, R.R., DeWitt, C.M., Friedman, A., and et al** (1993): Heterologous protection against influenza by injection of DNA encoding a viral protein [see comments], *Science.* 259: 1745-1749
- Umezawa, A., Yamamoto, H., Rhodes, K., Klemsz, M.J., Maki, R.A., and Oshima, R.G.** (1997): Methylation of an ETS site in the intron enhancer of the keratin 18 gene participates in tissue-specific repression, *Mol. Cell Biol.* 17: 4885-4894

- Vernon, A. (2004): The GS Gene Expression System, Business Briefing: Pharmatech 2004
- Vollmer, J., Weeratna, R., Payette, P., Jurk, M., Schetter, C., Laucht, M., Wader, T., Tluk, S., Liu, M., Davis, H.L., and Krieg, A.M. (2004): Characterization of three CpG oligodeoxynucleotide classes with distinct immunostimulatory activities, *Eur.J.Immunol.* 34: 251-262
- Voo, K.S., Carlone, D.L., Jacobsen, B.M., Flodin, A., and Skalniak, D.G. (2000): Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1, *Mol.Cell Biol.* 20: 2108-2121
- Wade, P.A. (2001): Methyl CpG-binding proteins and transcriptional repression, *Bioessays* 23: 1131-1137
- Wagner, R. (2001): Neue Impulse für die Biotechnologie durch synthetische Gene, *BIOforum* 9.
- Warnecke, P.M. and Clark, S.J. (1999): DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation, *Mol.Cell Biol.* 19: 164-172
- Warnecke, P.M. and Bestor, T.H. (2000): Cytosine methylation and human cancer, *Curr.Opin.Oncol.* 12: 68-73
- Wolf, H., Modrow, S., Soutschek, E., Motz, M., Grunow, R., and Döbl, H. (1990): Production, mapping and biological characterisation of monoclonal antibodies to the core protein (p24) of the human immunodeficiency virus type 1., *AIFO* 1: 24-29
- Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999): Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene, *Nature* 402: 187-191
- Xu, Z.L., Mizuguchi, H., Ishii-Watabe, A., Uchida, E., Mayumi, T., and Hayakawa, T. (2002): Strength evaluation of transcriptional regulatory elements for transgene expression by adenovirus vector, *J.Control Release* 81: 155-163
- Yamamoto, S., Yamamoto, T., Kataoka, T., Kuramoto, E., Yano, O., and Tokunaga, T. (1992): Unique palindromic sequences in synthetic oligonucleotides are required to induce IFN [correction of INF] and augment IFN-mediated [correction of INF] natural killer activity, *J.Immunol.* 148: 4072-4076
- Yamamoto, S., Yamamoto, T., Shimada, S., Kuramoto, E., Yano, O., Kataoka, T., and Tokunaga, T. (1992): DNA from bacteria, but not from vertebrates, induces interferons, activates natural killer cells and inhibits tumor growth, *Microbiol.Immunol.* 36: 983-997
- Yaneva, J., Zacharieva, S. (2002) DNA methylation and cancer: An overview, *Experimental Pathology and Parasitology*, 5 (10), 45-54
- Yew, N.S., Wysokenski, D.M., Wang, K.X., Ziegler, R.J., Marshall, J., McNeilly, D., Cherry, M., Osburn, W., and Cheng, S.H. (1997): Optimization of plasmid vectors for high-level expression in lung epithelial cells, *Hum.Gene Ther.* 8: 575-584
- Yew, N.S., Zhao, H., Wu, I.H., Song, A., Tousignant, J.D., Przybylska, M., and Cheng, S.H. (2000): Reduced inflammatory response to plasmid DNA vectors by elimination and inhibition of immunostimulatory CpG motifs, *Mol.Ther.* 1: 255-262
- Yew, N.S., Przybylska, M., Ziegler, R.J., Liu, D., and Cheng, S.H. (2001): High and sustained transgene expression in vivo from plasmid vectors containing a hybrid ubiquitin promoter, *Mol.Ther.* 4: 75-82
- Yew, N.S., Zhao, H., Przybylska, M., Wu, I.H., Tousignant, J.D., Scheule, R.K., and Cheng, S.H. (2002): CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo, *Mol.Ther.* 5: 731-738
- Yi, A.K., Chang, M., Peckham, D.W., Krieg, A.M., and Ashman, R.F. (1998): CpG oligodeoxyribonucleotides rescue mature spleen B cells from spontaneous apoptosis and promote cell cycle entry, *J.Immunol.* 160: 5898-5906
- Yi, A.K. and Krieg, A.M. (1998): Rapid induction of mitogen-activated protein kinases by immune stimulatory CpG DNA, *J.Immunol.* 161: 4493-4497
- Zahn-Zabal, M., Kobr, M., Girod, P.A., Imhof, M., Chatellard, P., de Jesus, M., Wurm, F., and Mermod, N. (2001): Development of stable cell lines for production or regulated expression using matrix attachment regions, *J.Biotechnol.* 87: 29-42
- Zhang, A., Jin, H., Zhang, F., Ma, Z., Tu, Y., Ren, Z., Zhang, X., Zhu, K., and Wang, B. (2005): Effects of multiple copies of CpG on DNA vaccination, *DNA Cell Biol.* 24: 292-298
- Zhang, X.Y., Ni, Y.S., Saifudeen, Z., Asiedu, C.K., Supakar, P.C., and Ehrlich, M. (1995): Increasing binding of a transcription factor immediately downstream of the cap site of a cytomegalovirus gene represses expression, *Nucleic Acids Res.* 23: 3026-3033
- Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999): Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation, *Genes Dev.* 13: 1924-1935
- Zhou, Z., Schnake, P., Xiao, L., and Lal, A.A. (2004): Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization, *Protein Expr.Purif.* 34: 87-94
- Zolotukhin, S., Potter, M., Hauswirth, W.W., Guy, J., and Muzyczka, N. (1996): A "humanized" green fluorescent protein cDNA adapted for high-level expression in mammalian cells, *J.Virol.* 70: 4646-4654

Veröffentlichungen

Teile der vorliegenden Arbeit wurden veröffentlicht in:

D.Leikam, F. Notka & R. Wagner: Advancing DNA vaccine development by increasing protein expression via CpG motive modulation of the coding sequence. Jahrestagung der Gesellschaft für Virologie, Tübingen, 2004.

F. Notka, M. Graf, D. Leikam, A. Bauer, D. Kosovac & R. Wagner: Modifying the coding region of genes for enhanced protein production. Jahrestagung der European Society of Gene Therapy (ESGT), Tampere, 2004.

D.Leikam, F. Notka & R. Wagner: Regulation of transgene expression by intragenic CpG dinucleotides. Jahrestagung der Gesellschaft für Virologie, München, 2005.

F. Notka, M. Graf, D. Leikam, R. Wagner, D. Raab: Modulating gene expression by directed insertion and removal of CpG dinucleotides. WO 2006015789.

F Anhang

F.1 Oligonukleotide

Sämtliche Oligonukleotidsequenzen sind in 5'→3'-Orientierung angegeben. Zur Herstellung synthetischer Gene verwendete Oligonukleotide werden aufgrund ihrer großen Anzahl nicht aufgeführt, sondern können den kodonoptimierten Sequenzen (siehe F.2.1) entnommen werden.

F.1.1 Oligonukleotide zum Aufbau der Vektoren

Name:	Sequenz:
CMV_Kpn_pf	GTTAGGTACCGTTGACATTGATTATTGACTAGTTATTAATAG
CMV_pb	GAATGAGCTCTGCTTATATAGACC
BGH_GDintpf	GTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCCTTTCCTAATAAA ATGAGGAAATTGCATDGCATTGTCTGAGTAGG
BGH_WGpf	TACTGTTAACCAGTGTTTAAACCTGTGCCTTCTAGTTGCCAGCCATCTGTTGT TTGCCCTCCCCWGTGCCTTCCCTTGACCCTGG
BGH_CWpb	TAGTGAGCTCGATACCCGGGCCATAGAGCCCACWGCATCCCCAGCATGCC
CMV_intL1_pf	ACTTTCCAAAATGTCGTAACAAC
ori_mL1L2pb	TTCCTGGCCTTTTGCTGGCCT
CMV_pf	CAAAGGTACCGTTAATCGATGTTGACATTGATTATTGACTA
BGH_pf	CTT AGA GCT CCA GTG TTT AAA CCT GTG CCT TCT AGT TGC CAG
BGH_pb	CAA AGG TAC CGA TAC CCG GGC CAT AGA GCC CAC CGC ATC
oriXE_pf	GTCACCCGGGTAGTGAATTCATGTGAGCAAAGGC
OriCB_pb	GATCTTTTCTACGGGAGATCTGTCAATCGATAGCT
referenz_oripf	TCAGATGCATCCGTACGTTAACATGTGAGCAAAGGCCAGCA
referenz_BGHpb	AGTCATGCATCCATAGAGCCCACCGCATCCCCA
AmpA_m4	TCTGCTCCCTGCTTGTGTGTTGGAGGTCAGTACTGAGTA
AmpC_L9	AACCTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT
110_X/E_pf	TACTGAATTCATGTGAGCAAAGGCCAGC
110_X/E_pb	AGTAGAATTCAGTACCCGGGCCATAGAGCCCACCG
BGH_F1pf	GAGTTTAAACCTGTGCCTTCTAG
BGH_F1pb	CTAGAAGGCACAGGTTTAAACTCGAGGCTGATCAGCGGGTTTA
BGH_F2pf	GCTTCTGAGGCGGAAAGAAC
BGH_F2pb	GTTCTTTCCGCCTCAGAAGCCCGGGCCATAGAGCCCCT
773A_pf	GAACGAGCTCTCTGGCTAACT
SV40P(A)_pf	TGAGGATCCACTGGTTTAAACAACCTTGTTTATTGCAGCTTATAATG
SV40P(A)_pb	AGTACCCGGGCAGACATGATAAGATACATTG
CMVsyn_bla_Fpf	GATGTAAGTGGCCAGATATACTAGGTACCGTTGACATTGATTATTG
CMVsyn_bla_Fpb	CAATAATCAATGTCAACGGTACCTAGTATATCTGGCCAGTACATC

SV40prom_pf	GATCGGTACCTGTGGAATGTGTGTCAGTTAGG
SV40prom_pb	CTCAGAGCTCGAGCTTTTTGCAAAGCCTAGG
p24_pf	ACTAGGTACCATCTAAGCTTATGCCCATCGTGCAGAACATCCA
p24_pb	TCAAGAGCTCGACTGGATCCTATTACAGCACCCCTGGCCTTGTGGC
FRT_pf	CGGATCTGATCAGCACGTGATG
FRT_pb	GTTTTCCAGTCACGACGTTG
hu/E019_pf	AAGCTTGCCACCATGGTG
hu/E019_pb	GGATCCTATCACTTGTACAGCTC
GFPmax_pf	ATGGTGTCTGAAGGGCGAGG
GFPmax_pb	CTATCACTTGTACAGCTCGTCC
huGFP_pf	GTACGGTACCTACGAAGCTTGCCACCATGGTGAGCAAGGG
huGFP_pb	ACGAGCTGTACAAGTAATAGGATCCTACTGAGCTCTACT

F.1.2 Oligonukleotide zur Generation CpG-freier Promotorkonstrukte

AB1_Fpf	CCCAATTACCCCTTCCCATTGATATCAATAATGATATATG
AB1_Fpb	CATATATCATTATTGATATCAATGGGAAGGGGTAATTGGG
AB2_Fpf	CCCAATTACCCCTTCCCATTGAGATCAATAATGATATATG
AB2_Fpb	CATATATCATTATTGATCTCAATGGGAAGGGGTAATTGGG
B1C3_Fpf	GATGTAGTTTTGGCAGTACATCAATGGACTGGATAGTGTTTG
B1C3_Fpb	CAAACACTATCCAGTCCATTGATGTACTGCCAAAACACTACATC
B1C4_Fpf	GATGTAGTTTTGGCAGTACATCAATGGGTCTGGATAGGTGTTTG
B1C4_Fpb	CAAACACCTATCCAGACCCATTGATGTACTGCCAAAACACTACATC
B1C7_Fpf	GATGTAGTTTTGGCAGTACATCAATGGGGCTGGATAGTAG
B1C7_Fpb	CTACTATCCAGCCCCATTGATGTACTGCCAAAACACTACATC
B2C3_Fpf	GATGGTGTTTTGGCAGTACATCAATGGACTGGATAGTGTTTG
B2C3_Fpb	CAAACACTATCCAGTCCATTGATGTACTGCCAAAACACCCATC
B2C4_Fpf	GATGGTGTTTTGGCAGTACATCAATGGGTCTGGATAGGTG
B2C4_Fpb	CACCTATCCAGACCCATTGATGTACTGCCAAAACACCCATC
B2C7_Fpf	GATGGTGTTTTGGCAGTACATCAATGGGGCTGGATAG
B2C7_Fpb	CTATCCAGCCCCATTGATGTACTGCCAAAACACCCATC

F.1.3 Für realtime PCR-Analysen verwendete Oligonukleotide

GFPint142_pf	CCCTGAAGTTCATCTGCACC
GFPint516_pb	GATCTTGAAGTTCACCTTGATG
hygrint385_pf	GAGCGGGTTCGGCCATTC
hygrint754_pb	GTATTGGGAATCCCCGAACATCG
GFPmax165_pf	CGGCAAGCTGCCCGTGCC
GFPmax520_pb	CTGCCGTCCTCGATGTTGTG

actin_pf	GGTGGGCATGGGCCAGAAGG
actin_pb	CGCGTAGCCCTCGTAGATGG

F.2 Sequenzvergleiche

Die Sequenzen der einzelnen Vektorelemente sind entsprechend dem finalen PCR-Produkt angegeben. Start- und Stoppkodons sowie funktionelle Bereiche (z.B. Promotoren, Polyadenylierungssignale) sind fettgedruckt, Restriktionsschnittstellen sind eingerahmt. Positionsangaben beziehen sich jeweils auf die nachfolgende Sequenz. CpG-Dinukleotide werden durch graue Schraffur hervorgehoben und die Kozzak-Sequenz kursiv dargestellt. Weiter sind terminale Überhänge durch graue Kleinbuchstaben gekennzeichnet.

F.2.1 Nukleinsäuresequenzvergleiche

Amp-Alignment: Amp_72 (pc5) vs. Amp_1 (02-774)

285-383: *bla*-Promotor;

*Cla*I (5), *Bgl*III (225), *Nsi*I (279), *Sca*I (384), *Sca*I (694), *Hpa*I (1250), *Bgl*III (1400);

	5 15 25 35 45 55
Amp_1	gtcc ATCGAT TAGTATATCT GGCCAGTACA TCTCCAAGCA GTGCAAAAGG CCTAACCCCTA
Amp_72 pc5	----- GCGTATATCT GGCCCGTACA TCGCGAAGCA GCGCAAAACG CCTAACCCCTA

	65 75 85 95 105 115
Amp_1	AGCAGATTCT TCATGCAATT GTAGGTCAAG CCTTGCCTTG TTGTAGCTTA AATTTTGCTA
Amp_72 pc5	AGCAGATTCT TCATGCAATT GTCGGTCAAG CCTTGCCTTG TTGTAGCTTA AATTTTGCTC

	125 135 145 155 165 175
Amp_1	GCCCACTACT CAGTGACCTC CAACACACAA GCAGGGAGCA GATACTGGCT TAACTATGAG
Amp_72 pc5	GCGCACTACT CAGCGACCTC CAACACACAA GCAGGGAGCA GATACTGGCT TAACTATCGG

	185 195 205 215 225 235
Amp_1	GCATCAGAGC AGATTGTACT GAGAGTGCAC CATAGGGGAT TGGGAGAACT CCCTATCAGT
Amp_72 pc5	GCATCAGAGC AGATTGTACT GAGAGTGCAC CATAGGGGAT CCGGAGATCT CCGATCCGT

	245 255 265 275 285 295
Amp_1	CAATTTCAGG TGGCACTTTT CAGGGAAATG TGAGCTGAT GCAT ACCCCT ATTGTTTAT
Amp_72 pc5	CGACGTCAGG TGGCACTTTT CCGGGAAATG TGCGCGA----- ACCCCT ATTGTTTAT

	305 315 325 335 345 355
Amp_1	TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC
Amp_72 pc5	TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC

	365 375 385 395 405 415
Amp_1	AATAATATTG AAAAGGAAG AGTAGTACTA TGAGTATTCA ACATTTCAGA GTGCCCCTTA
Amp_72 pc5	AATAATATTG AAAAGGAAG AGT-----A TGAGTATTCA ACATTTCCGT GTCGCCCTTA

	425 435 445 455 465 475
Amp_1	TTCCCTTTTT TGCTGCCTTT TGCCTTCCTG TTTTGTCTCA CCCAGAAACC CTGGTGAAG
Amp_72 pc5	TTCCCTTTTT TGCGCATTT TGCCTTCCTG TTTTGTCTCA CCCAGAAACG CTGGTGAAG

	485 495 505 515 525 535
Amp_1	TAAAAGATGC TGAAGATCAG TTGGGTGCAA GAGTGGGTTA CATTGAACTG GATCTCAACT
Amp_72 pc5	TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA

```

.....|.....|.....|.....|.....|.....|.....|.....|
545      555      565      575      585      595
Amp_1    CTGGTAAGAT CCTTGAGAGT TTTAGACCAG AAGAAAGATT TC CAATGATG AGCACCTTTA
Amp_72 pc5 GCGGTAAGAT CCTTGAGAGT TTTGCGCCCG AAGAACCCTT TC CAATGATG AGCACCTTTA

.....|.....|.....|.....|.....|.....|.....|.....|
605      615      625      635      645      655
Amp_1    AAGTTCTGCT ATGTGGTGCA GTATTATCCA GAATTGATGC TGGGCAAGAG CAACTGGGTA
Amp_72 pc5 AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CCGGCAAGAG CAACTCGGTC

.....|.....|.....|.....|.....|.....|.....|.....|
665      675      685      695      705      715
Amp_1    GAAGAATACA CTATTCTCAG AATGACTTGG TTGAGTACAG TCCAGTCACA GAAAAGCATC
Amp_72 pc5 GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC

.....|.....|.....|.....|.....|.....|.....|.....|
725      735      745      755      765      775
Amp_1    TTACAGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA
Amp_72 pc5 TTACAGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA

.....|.....|.....|.....|.....|.....|.....|.....|
785      795      805      815      825      835
Amp_1    CTGCTGCCAA CTTACTTTTG ACAACCATTG GAGGACCAA GGAGTTGACT GCTTTTTTGC
Amp_72 pc5 CTGCGGCCAA CTTACTTCTG ACAACCATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC

.....|.....|.....|.....|.....|.....|.....|.....|
845      855      865      875      885      895
Amp_1    ACAACATGGG TGATCATGTA ACTAGGCTTG ATAGATGGGA ACCAGAGCTG AATGAAGCCA
Amp_72 pc5 ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGAGCTG AATGAAGCCA

.....|.....|.....|.....|.....|.....|.....|.....|
905      915      925      935      945      955
Amp_1    TACCAAATGA TGAGAGGGAC ACCACCATGC CTGTAGCAAT GGCAACAACC TTGAGGAAAC
Amp_72 pc5 TACCAAACGA CGAGCGTGAC ACCACCATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC

.....|.....|.....|.....|.....|.....|.....|.....|
965      975      985      995      1005     1015
Amp_1    TATTAAGTGG TGAAGTCTT ACTCTAGCTT CCAGACAACA ATTAATAGAC TGGATGGAGG
Amp_72 pc5 TATTAAGTGG CGAAGTCTT ACTCTAGCTT CCGGCAACA ATTAATAGAC TGGATGGAGG

.....|.....|.....|.....|.....|.....|.....|.....|
1025     1035     1045     1055     1065     1075
Amp_1    CAGATAAAGT TGCAGGACCA CTTCTGAGGT CTGCCCTTCC AGCTGGCTGG TTCATTGCAG
Amp_72 pc5 CCGATAAAGT TGCAGGACCA CTTCTGCGCT CCGCCCTTCC GGCTGGCTGG TTTATTGCTG

.....|.....|.....|.....|.....|.....|.....|.....|
1085     1095     1105     1115     1125     1135
Amp_1    ATAAATCTGG AGCAGGTGAG AGAGGGTCTA GAGGTATCAT TGCAGCACTG GGGCCAGATG
Amp_72 pc5 ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG

.....|.....|.....|.....|.....|.....|.....|.....|
1145     1155     1165     1175     1185     1195
Amp_1    GTAAGCCCTC CAGGATTGTA GTTATCTACA CCACAGGCAG TCAGGCAACC ATGGATGAGA
Amp_72 pc5 GTAAGCCCTC CCGTATCGTA GTTATCTACA CACCGGGGAG TCAGGCAACT ATGGATGAAC

.....|.....|.....|.....|.....|.....|.....|.....|
1205     1215     1225     1235     1245     1255
Amp_1    GAAATAGGCA GATTGCAGAG ATTGGTGCCT CATTGATTAA GCATTGGTAA GTTAACTGCTG
Amp_72 pc5 GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA -----CTGT

.....|.....|.....|.....|.....|.....|.....|.....|
1265     1275     1285     1295     1305     1315
Amp_1    CAGACCAAGT TTAATCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA
Amp_72 pc5 CAGACCAAGT TTAATCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA

.....|.....|.....|.....|.....|.....|.....|.....|
1325     1335     1345     1355     1365     1375
Amp_1    GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CATGAGTTTT
Amp_72 pc5 GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CCGTGAAGTTTT

.....|.....|.....|.....|.....|.....|.....|.....|
1385     1395     1405
Amp_1    GGTTCCTACTG AGAGTCAGAC AGATCTagta
Amp_72 pc5 CGTTCCTACTG AGCTCAGAC -----

```

Hygr-Alignment: Hygr_174 (pc5) vs. Hygr_1 (02-772)

293 - 340: FRT - Element; 1524 - 1654: SV 40 Polyadenylierungssignal

XmaI (5), AflIII (348), SmaI (1377), PshAI (1514), PacI (1655), EcoRI (594, 2046);

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      5       15       25       35       45       55
Hygr_1  tagt[CCCGGG] CTTCTGAGGC TGAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACTCA
Hygr_174 pc5 ----- CTTCTGAGGC GGAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACCGC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      65       75       85       95       105      115
Hygr_1  CCCTGTAGCA GTGCATTAAG CTCAGCAGGT GTGGTGGTTA GGCTCAGTGT GACTGCTACA
Hygr_174 pc5 CCCTGTAGCG GCGCATTAAG CGCGCCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA

      ....|....| ....|....| ....|....| ....|....| ....|....|
      125      135      145      155      165      175
Hygr_1  CTTGCCAGTG CCCTAGCACC AGCTCCTTTC ACTTCTTCC CTTCTTTCT CACCAGATTG
Hygr_174 pc5 CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTCTTCC CTTCTTTCT CGCCACGTTG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      185      195      205      215      225      235
Hygr_1  GCAGGCTTTC CCAGTCAAGC TCTAAATTGG GGTCCCTTT AGGTTTCAGA TTTAGTGCTT
Hygr_174 pc5 GCCGCTTTC CCGTCAAGC TCTAAATCGG GGTCCCTTT AGGTTCCGA TTTAGTGCTT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      245      255      265      275      285      295
Hygr_1  TACTGCACCT GGACCCAAA AAACCTGATT AGGGTGATGG TTCTTGATCC TAGAAGTTCC
Hygr_174 pc5 TACGGCACCT CGACCCAAA AAACCTGATT AGGGTGATGG TTCACGTACC TAGAAGTTCC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      305      315      325      335      345      355
Hygr_1  TATTCGAAG TTCTATTCT CTAGAAAGTA TAGGAACTTC CTTGGCCCTT AAGAAAAAGC
Hygr_174 pc5 TATTCGAAG TTCTATTCT CTAGAAAGTA TAGGAACTTC CTTGGCC--- ---AAAAAGC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      365      375      385      395      405      415
Hygr_1  CTGAACTCAC TGCCACCTCT GTGGAGAAGT TTCTTATGA AAAATTTGAC TCTGTCACTG
Hygr_174 pc5 CTGAACTCAC CGCGACGTCT GTCGAGAAGT TTCTGATCGA AAAGTTGAC AGCGTCTCCG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      425      435      445      455      465      475
Hygr_1  ACTTGATGCA GCTCAGTGAA GGTGAAGAAT CTAGGGCTTT CAGCTTTGAT GTAGGAGGGA
Hygr_174 pc5 ACCTGATGCA GCTCTCGGAG GCGGAAGAAT CTCTGTGCTTT CAGCTTCGAT GTAGGAGGGC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      485      495      505      515      525      535
Hygr_1  GAGGATATGT CCTGAGAGTA AATAGCTGTG CAGATGGTTT CTACAAAGAT AGATATGTGT
Hygr_174 pc5 GTGGATATGT CCTGCGGTA AATAGCTGCG CCGATGGTTT CTACAAAGAT CGTTATGTTT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      545      555      565      575      585      595
Hygr_1  ATAGGCACCT TGCACTGCA GCCCTCCCAA TTCCAGAAGT GTTAGACATT GGGGAATTTT
Hygr_174 pc5 ATCGGCACCT TGCACTCGGC GCGCTCCCGA TTCCGAAGT GCTTGACATT GGGGAATTTA

      ....|....| ....|....| ....|....| ....|....| ....|....|
      605      615      625      635      645      655
Hygr_1  CTGAGAGCTT GACCTATTGC ATCTCCAGAA GAGCACAGGG TGTACCTTG CAAGACCTGC
Hygr_174 pc5 GCGAGAGCCT GACCTATTGC ATCTCCCGCC GTGCACAGGG TGTACCGTTG CAAGACCTGC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      665      675      685      695      705      715
Hygr_1  CTGAAACTGA ACTGCCTGCT GTGCTGCAGC CTGTGGCAGA GGCCATGGAT GCAATTGCTG
Hygr_174 pc5 CTGAAACCGA ACTGCCTGCT GTTCTGCAGC CGGTCCCGGA GGCCATGGAT GCGATCGCTG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      725      735      745      755      765      775
Hygr_1  CAGCAGACCT TAGCCAGACC TCTGGGTTTG GCCATTGG ACCACAAGGA ATTGGTCAAT
Hygr_174 pc5 CGGCCGATCT TAGCCAGACC AGCGGTTTCG GCCATTGG ACCACAAGGA ATCGGTCAAT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      785      795      805      815      825      835
Hygr_1  AACTACATG GAGAGATTTT ATATGTGCCA TTGCAGATCC CCATGTGTAT CATTGGCAAA
Hygr_174 pc5 AACTACATG GCGTATTTT ATATGCGCGA TTGCTGATCC CCATGTGTAT CACTGGCAAA

```


	845	855	865	875	885	895
Hygr_1	CTGTGATGGA	TGACACTGTG	AGTGCAAGTG	TGGCCCAAGC	TCTGGATGAG	CTTATGCTTT
Hygr_174 pc5	CTGTGATGGA	CGACACCGTC	AGTGCGTCCG	TCCGCGCAGG	TCT CGATGAG	CTGATGCTTT

	905	915	925	935	945	955
Hygr_1	GGGCAGAGGA	CTGCCCTGAA	GTGAGGCATC	TGGTTCATGC	AGATTTTGGC	TCCAACAATG
Hygr_174 pc5	GGGC CGAGGA	CTGCC CCGAA	GTC CGGCACC	TCGTGCACGC	GGATTTCCGC	TCCAACAATG

	965	975	985	995	1005	1015
Hygr_1	TCTTAACAGA	CAATGGCAGA	ATAACAGCAG	TGATTGACTG	GAGTGAAGCC	ATGTTTGGTG
Hygr_174 pc5	TCCTG ACCGA	CAATGG CCGC	ATAACAG CCG	TCATTGACTG	GAG CCGAGCCG	ATGTT CCGGG

	1025	1035	1045	1055	1065	1075
Hygr_1	ATTCCAATA	TGAGGTGGCC	AACATATTCT	TCTGGAGGCC	TTGGTTGGCT	TGTATGGAAC
Hygr_174 pc5	ATTCCAATA	CGAGTCCGC	AACATCTTCT	TCTGGAGGCC	GTGGTTGGCT	TGTATGAGAC

	1085	1095	1105	1115	1125	1135
Hygr_1	AGCAGACCAG	ATACTTTGAG	AGAAGGCATC	CTGAGCTGGC	AGGAAGCCCA	AGACTCAGGG
Hygr_174 pc5	AGCAG ACCGC	CTACTT CGAG	CGGAGGCATC	CGGAGCTTGC	AGGAT CCCGC	CGGCTCCGGG

	1145	1155	1165	1175	1185	1195
Hygr_1	CCTATATGCT	CAGAATTGGT	CTTGACCAAC	TGTATCAGAG	CTTGGTTGAT	GGCAATTTTG
Hygr_174 pc5	CGTATATGCT	CCGCATTGGT	CTTGACCAAC	TCATCAGAG	CTTGGTTGAC	GGCAATTTCCG

	1205	1215	1225	1235	1245	1255
Hygr_1	ATGATGCAGC	TTGGGCCCAG	GGAAGGTGTG	ATGCAATTGT	CAGGTCTGGA	GCAGGGACTG
Hygr_174 pc5	ATGATGCAGC	TTGGG CCGAG	GGT CGATCCG	ACGCAATCGT	CCGATCCGGA	GCCGGGACTG

	1265	1275	1285	1295	1305	1315
Hygr_1	TGGGGAGAAC	ACAAATTGCC	AGGAGAAGTG	CAGCTGTTTG	GACAGATGGC	TGTGTAGAGG
Hygr_174 pc5	TCGGGCGTAC	ACAAAT CCGC	CGCAGAAGCG	CGGCCGTCTG	GAC CGATGGC	TGTGTAGAAG

	1325	1335	1345	1355	1365	1375
Hygr_1	TGCTGGCAGA	TAGTGGAAAC	AGAAGGCCCA	GCACTAGACC	AAGGGCAAAG	GAATAGATTT
Hygr_174 pc5	TACT CGCCGA	TAGTGGAAAC	CGACGCCCCA	GCACT CGTCC	GAGGGCAAAG	GAATAG----

	1385	1395	1405	1415	1425	1435
Hygr_1	AAAT CATGTA	CTAAGAGATT	TGGATTCCAC	TGCAGCCTTC	TATGAAAGGT	TGGGCTTCAG
Hygr_174 pc5	---- CACGTA	CTA CGAGATT	T CGATTCCAC	CGCCGCCTTC	TATGAAAGGT	TGGGCTT CCG

	1445	1455	1465	1475	1485	1495
Hygr_1	AATGGTTTTC	CAGGACTCCA	GCTGGATGAT	CCTCCAGCAC	TGGGATCTCA	TGCTGGAGTT
Hygr_174 pc5	AAT CGTTTTTC	CGGGACGCCG	GCTGGATGAT	CCTCCAG CGC	GGGGATCTCA	TGCTGGAGTT

	1505	1515	1525	1535	1545	1555
Hygr_1	CTTCTCCAC	CCA GACACTA	GTC CACTTGT	TTATTGCAGC	TTATAATGGT	TACAAATAAA
Hygr_174 pc5	CTT CGCCAC	CC-----	-- CACTTGT	TTATTGCAGC	TTATAATGGT	TACAAATAAA

	1565	1575	1585	1595	1605	1615
Hygr_1	GCAATAGCAT	CACAAATTTT	ACAATAAAG	CATTTTTTTC	ACTGCATTCT	AGTTGTGGTT
Hygr_174 pc5	GCAATAGCAT	CACAAATTTT	ACAATAAAG	CATTTTTTTC	ACTGCATTCT	AGTTGTGGTT

	1625	1635	1645	1655	1665	1675
Hygr_1	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGTTAATT	AA TATACCAT	TGACCTCTAG
Hygr_174 pc5	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTG-----	--TATAC CGT	CGACCTCTAG

	1685	1695	1705	1715	1725	1735
Hygr_1	CTAGAGCTTG	GAGTAATCAT	GGTCATAGCT	GTTTCCTGTG	TGAAATTGTT	ATCAGCTCAC
Hygr_174 pc5	CTAGAGCTTG	GCGTAATCAT	GGTCATAGCT	GTTTCCTGTG	TGAAATTGTT	ATC CGCTCAC

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      1745      1755      1765      1775      1785      1795
Hygr_1      AATTCACAC AACATACTAG CAGGAAGCAT AAAGTGATA GCCTGGGGTG CCTAATGAGT
Hygr_174 pc5 AATTCACAC AACATACGAG CCGGAAGCAT AAAGTGATA GCCTGGGGTG CCTAATGAGT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      1805      1815      1825      1835      1845      1855
Hygr_1      GAGCTAACTC ACATTAATTG CCTTCTCTCT ACTGCCAGCT TTCCAGTCAG GAAACCTGTA
Hygr_174 pc5 GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      1865      1875      1885      1895      1905      1915
Hygr_1      GTGCCAGCTG CATTAAATGAA TCTGCCAACT CACTGGGAGA GGAGGTTTGC CTATTGGGCA
Hygr_174 pc5 GTGCCAGCTG CATTAAATGAA TCCGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      1925      1935      1945      1955      1965      1975
Hygr_1      CTCTTCAGCT TCCTAGCTCA CTGACTGGCT GCACTCAGTG GTTCAGCTGC TGCTAGCAGT
Hygr_174 pc5 CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      1985      1995      2005      2015      2025      2035
Hygr_1      ATCAGCTCAC TCAAAGGCTG TAATAGGGTT ATCCACAGAA TCAGGGGATA ACTCAGGAAA
Hygr_174 pc5 ATCAGCTCAC TCAAAGGCG TAATACGGTT ATCCACAGAA TCAGGGGATA ACCGAGGAAA

      ....|....| ....|
      2045      2055
Hygr_1      GAACGAATT Caggt
Hygr_174 pc5 GAAC-----

```

MCS-Alignment: MCS_11 (pc5) vs. MCS_1 (02-773)

54-73: T7 Promoter;

SacI (5,114), *PmeI* (92,219), *AflIII* (99), *HindIII* (102), *KpnI* (108), *BamHI* (120), *EcoRV* (112,153), *EcoRI* (143);

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      5          15          25          35          45          55
MCS_11 pc5  ----- TCTGGCTAAC TAGAGAACCC ACTGCTTACT GGCTTATCGA AATTAATACG
MCS_1      tgtaGAGCTC TCTGGCTAAC TAGAGAACCC ACTGCTTACT GGCTTATCTA AATTAATACG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      65          75          85          95          105         115
MCS_11 pc5  ACTCACTATA GGGAGACCCA AGCTGGCTAG CGTTTAAACT TAAGTTTGGT ACCBAGCTCG
MCS_1      ACTCACTATA GGGAGACCCA A----GCT-- -G-----T TAAGCTTGGT AGATATCAGG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      125         135         145         155         165         175
MCS_11 pc5  GATCCACTAG TCCAGTGTGG TGGAATTCTG CGATATCCA GCACAGTGGC GGCCGCTCGA
MCS_1      GATCCACTCA -----

      ....|....| ....|....| ....|....| ....|....|
      185         195         205         215         225
MCS_11 pc5  GTCTAGAGGG CCCGTTTAAA CCCGCTGATC AGCCTCGA-----
MCS_1      ----- --GCTGATC AGCCTCCAGT TTAACtgta

```

Kana-Alignment: kana_42 (pc5) und kana_0 (041196)

ScaI (1), ClaI (128); HpaI (823) ;

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      5       15       25       35       45       55
kana_0  AGTACTATGA GCCATATTCA AAGAGAAACC TCTTGCTCTA GGCCAAGGTT AAATTCCAAC
kana_42 pc5 -----ATGA GCCATATTCA ACGGGAAACG TCTTGCTCTA GGCCGCGATT AAATTCCAAC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      65       75       85       95       105      115
kana_0  ATGGATGCTG ATTTATATGG GTATAAATGG GCTAGAGATA ATGTTGGGCA ATCAGGTGCC
kana_42 pc5 ATGGATGCTG ATTTATATGG GTATAAATGG GCTCGCGATA ATGTCGGGCA ATCAGGTCGG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      125      135      145      155      165      175
kana_0  ACAATCTACA GGTTGTATGG GAAGCCAGAT GCACCAGAGT TGTTTCTGAA ACATGGCAAA
kana_42 pc5 ACAATCTATC GATGTATGG GAAGCCCGAT GCCCCAGAGT TGTTTCTGAA ACATGGCAAA

      ....|....| ....|....| ....|....| ....|....| ....|....|
      185      195      205      215      225      235
kana_0  GGTTCGTGTTG CCAATGATGT TACAGATGAG ATGGTCAGAC TAAACTGGCT GACTGAATTT
kana_42 pc5 GGTAGCGTTG CCAATGATGT TACAGATGAG ATGGTCAGAC TAAACTGGCT GACCGAATTT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      245      255      265      275      285      295
kana_0  ATGCCTCTTC CAACCATCAA GCATTTTATC AGGACTCCTG ATGATGCATG GTTACTCACC
kana_42 pc5 ATGCCTCTTC CGACCATCAA GCATTTTATC CGTACTCCTG ATGATGCATG GTTACTCACC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      305      315      325      335      345      355
kana_0  ACTGCCATCC CAGGGAAAAC AGCATTCCAG GTATTAGAAG AATATCCTGA TTCAGGTGAA
kana_42 pc5 ACTGCGATCC CCGGAAAAC AGCATTCCAG GTATTAGAAG AATATCCTGA TTCAGGTGAA

      ....|....| ....|....| ....|....| ....|....| ....|....|
      365      375      385      395      405      415
kana_0  AATATGTTG ATGCCCTGGC AGTGTTCCTG AGAAGGTTGC ATTCTATFCC TGTTTGTAAAT
kana_42 pc5 AATATGTTG ATGCGCTGGC AGTGTTCCTG CGCCGGTTGC ATTCGATTCC TGTTTGTAAAT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      425      435      445      455      465      475
kana_0  TGTCCTTTTA ACTCTGATAG AGTATTTAGG CTGGCTCAGG CCCAATCAAG AATGAATAAT
kana_42 pc5 TGTCCTTTTA ACAGCGATCG CGTATTTCTG CTCGCTCAGG CGCAATCACG AATGAATAAC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      485      495      505      515      525      535
kana_0  GGTTTGGTTG ATGCCAGTGA TTTTGATGAT GAGAGAAATG GCTGGCCTGT TGAACAAGTC
kana_42 pc5 GGTTTGGTTG ATGCGAGTGA TTTTGATGAT GAGCGTAATG GCTGGCCTGT TGAACAAGTC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      545      555      565      575      585      595
kana_0  TGGAAAGAAA TGCACAAACT TTTGCCATTC TCACCAGATT CAGTGGTCAC TCATGGTGAT
kana_42 pc5 TGGAAAGAAA TGCATAAACT TTTGCCATTC TCACCGGATT CAGTCGTCAC TCATGGTGAT

      ....|....| ....|....| ....|....| ....|....| ....|....|
      605      615      625      635      645      655
kana_0  TTCTCACTTG ATAACCTTAT TTTTGATGAG GGGAAATTAA TAGGTTGTAT TGATGTTGGA
kana_42 pc5 TTCTCACTTG ATAACCTTAT TTTTGATGAG GGGAAATTAA TAGGTTGTAT TGATGTTGGA

      ....|....| ....|....| ....|....| ....|....| ....|....|
      665      675      685      695      705      715
kana_0  AGAGTGGGAA TTGCAGACAG ATACCAGGAT CTTGCCATCC TATGGAAC TGCTGGGTGAG
kana_42 pc5 CGAGTCGGAA TGCAGACCG ATACCAGGAT CTTGCCATCC TATGGAAC TGCTGGGTGAG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      725      735      745      755      765      775
kana_0  TTTTCTCCTT CATTACAGAA AAGGCTTTTT CAAAAATATG GTATTGATAA TCCTGATATG
kana_42 pc5 TTTTCTCCTT CATTACAGAA ACGGCTTTTT CAAAAATATG GTATTGATAA TCCTGATATG

      ....|....| ....|....| ....|....| ....|....|
      785      795      805      815      825
kana_0  AATAAATTGC AGTTTCATTT GATGCTGGAT GAGTTTTTCT AAGTTAAC
kana_42 pc5 AATAAATTGC AGTTTCATTT GATGCTCGAT GAGTTTTTCT AA-----

```

CMVAlignment: CMV vs. CMVp.-Varianten A-F :

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      5      15      25      35      45      55
CMV      gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACCGGGGTCA
CMV_A    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA
CMV_B    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA
CMV_C    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA
CMV_D    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA
CMV_E    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA
CMV_F    gttAATCGAT GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TAATGGGGTCA

```

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      65      75      85      95      105     115
CMV      TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT
CMV_A    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT
CMV_B    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT
CMV_C    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT
CMV_D    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT
CMV_E    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT
CMV_F    TTAGTTCATA GCCCATATAT GGAGTTCATT CTTACATAAC TTAAAGTAA- TGGCCTTCCT

```

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      125     135     145     155     165     175
CMV      GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA
CMV_A    GGCTGACAAC CCAATTACCC CTTCCCATTG ATATCAATAA TGATATATGT TCCCATAGTA
CMV_B    GGCTGACAAC CCAATTACCC CTTCCCATTG ATATCAATAA TGATATATGT TCCCATAGTA
CMV_C    GGCTGACAAC CCAATTACCC CTTCCCATTG ATATCAATAA TGATATATGT TCCCATAGTA
CMV_D    GGCTGACAAC CCAATTACCC CTTCCCATTG AGATCNATAA TGATATATGT TCCCATAGTA
CMV_E    GGCTGACAAC CCAATTACCC CTTCCCATTG AGATCNATAA TGATATATGT TCCCATAGTA
CMV_F    GGCTGACAAC CCAATTACCC CTTCCCATTG AGATCNATAA TGATATATGT TCCCATAGTA

```

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      185     195     205     215     225     235
CMV      ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC
CMV_A    ATTCCAATAG GGACTTTCCA TTGATTTCAA TGGGTGGACT ATTTGGAGTA AACTGCCCCAC
CMV_B    ATTCCAATAG GGACTTTCCA TTGATTTCAA TGGGTGGACT ATTTGGAGTA AACTGCCCCAC
CMV_C    ATTCCAATAG GGACTTTCCA TTGATTTCAA TGGGTGGACT ATTTGGAGTA AACTGCCCCAC
CMV_D    ATACCAATAG GGACTTTCCA TTGAGTTCAA TGGGTGGACT ATTTATAGTA AACTGCCCCAC
CMV_E    ATACCAATAG GGACTTTCCA TTGAGTTCAA TGGGTGGACT ATTTATAGTA AACTGCCCCAC
CMV_F    ATACCAATAG GGACTTTCCA TTGAGTTCAA TGGGTGGACT ATTTATAGTA AACTGCCCCAC

```

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      245     255     265     275     285     295
CMV      TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACCGCCCC CTATTGACCGT CAATGACCGGT
CMV_A    TTGGCAGTAC ATCAAGTGTA TC-TATGCCA AGTAGCCCCC CTAATGAAAT CAATGAAAGT
CMV_B    TTGGCAGTAC ATCAAGTGTA TC-TATGCCA AGTAGCCCCC CTAATGAAAT CAATGAAAGT
CMV_C    TTGGCAGTAC ATCAAGTGTA TC-TATGCCA AGTAGCCCCC CTAATGAAAT CAATGAAAGT
CMV_D    TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTAATCCCC CTATTGATAT CAATGAATGT
CMV_E    TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTAATCCCC CTATTGATAT CAATGAATGT
CMV_F    TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTAATCCCC CTATTGATAT CAATGAATGT

```

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      305     315     325     335     345     355
CMV      AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
CMV_A    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
CMV_B    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
CMV_C    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
CMV_D    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTG-CAG
CMV_E    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTG-CAG
CMV_F    AAATGGCCAC CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTG-CAG

```


	365	375	385	395	405	415
CMV	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATGC	GGTTTTGGCA	GTACATCAAT
CMV_A	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGT	AGTTTTGGCA	GTACATCAAT
CMV_B	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGT	AGTTTTGGCA	GTACATCAAT
CMV_C	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGT	AGTTTTGGCA	GTACATCAAT
CMV_D	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGG	TGTTTTGGCA	GTACATCAAT
CMV_E	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGG	TGTTTTGGCA	GTACATCAAT
CMV_F	TACATCTAGC	TATTAGTCAT	ACCTATTACC	ATGGTGATGG	TGTTTTGGCA	GTACATCAAT

	425	435	445	455	465	475
CMV	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC	TCCACCCCAT	TGACCTCAAT
CMV_A	GGA-CTGGAT	AGT-GTTTGA	CTCATAGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGATCAAT
CMV_B	GGGTCTGGAT	AGGTGTTTGA	CTCAATGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGCTCAAT
CMV_C	GGGGCTGGAT	AGTAGTTTGA	CTCAGTGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGATCAAT
CMV_D	GGA-CTGGAT	AGT-GTTTGA	CTCATAGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGATCAAT
CMV_E	GGGTCTGGAT	AGGTGTTTGA	CTCAATGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGCTCAAT
CMV_F	GGGGCTGGAT	AGTAGTTTGA	CTCAGTGGGA	TTTCCAAGTC	TCCACCCCAT	TGAGATCAAT

	485	495	505	515	525	535
CMV	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA	AATGTCGTAA	CAACTCCGCC
CMV_A	GGGAGTTTGT	TTTGGCACCA	AAATCAATTG	GACTTTCCAA	AATGTTTTAA	CAACTCATCC
CMV_B	GGGAGTTTGT	TTTGGCACCA	AAATCAATTG	GACTTTCCAA	AATGTGTTAA	CAACTCTCCC
CMV_C	GGGAGTTTGT	TTTGGCACCA	AAATCAAAAAG	GACTTTCCAA	AATGTTATAA	CAACTCTTCC
CMV_D	GGGAGTTTGT	TTTGGCACCA	AAATCAATTG	GACTTTCCAA	AATGTTTTAA	CAACTCATCC
CMV_E	GGGAGTTTGT	TTTGGCACCA	AAATCAATTG	GACTTTCCAA	AATGTGTTAA	CAACTCTCCC
CMV_F	GGGAGTTTGT	TTTGGCACCA	AAATCAAAAAG	GACTTTCCAA	AATGTTATAA	CAACTCTTCC

	545	555	565	575	585	595
CMV	CCATTGACGC	AAATGGGCGG	TAGGCGTGTA	CGGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_A	CCATTGATCC	AAATGGGGAG	TAGGCTTGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_B	CCATTGAGCC	AAATGGGATG	TAGGCATGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_C	CCATTGATCC	AAATGGGTAG	TAGGCTTGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_D	CCATTGATCC	AAATGGGGAG	TAGGCTTGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_E	CCATTGAGCC	AAATGGGATG	TAGGCATGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn
CMV_F	CCATTGATCC	AAATGGGTAG	TAGGCTTGTA	TTGTGGGAGG	TCTATATAAG	CAAGAGCTCnn

BGH -Alignment: BGH P (A) versus BGH WDW-Varianten:**103-108: Polyadenylierungssignal (AAUAAA)**

	5 15 25 35 45 55
BGH_pc5	cagtGTTTAA ACCTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCCGTGC
BGH_AAA	nnnnGTTTAA ACCTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCAGTGC
BGH_AGA	nnnnGTTTAA ACCTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCAGTGC
BGH_AGT	nnnnGTTTAA ACCTGTGCCT -CTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCAGTGC
BGH_ATA	nnnnGTTTAA ACCTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCAGTGC
BGH_ATT	nnnnGTTTAA ACCTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTCGCCCC TCCCCAGTGC

	65 75 85 95 105 115
BGH_pc5	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
BGH_AAA	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
BGH_AGA	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
BGH_AGT	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
BGH_ATA	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
BGH_ATT	CTTCCTTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG

	125 135 145 155 165 175
BGH_pc5	CATGGCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
BGH_AAA	CATAGCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
BGH_AGA	CATGGCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
BGH_AGT	CATGGCATTG GCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
BGH_ATA	CATGTCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
BGH_ATT	CATTGCATTG TCTGAGTAGG TGTCAATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA

	185 195 205 215 225 235
BGH_pc5	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGGGTGGGC TCTATGGCCC
BGH_AAA	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCAGTGGGC TCTATGGCCC
BGH_AGA	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCAGTGGGC TCTATGGCCC
BGH_AGT	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCTGTGGGC TCTATGGCCC
BGH_ATA	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCAGTGGGC TCTATGGCCC
BGH_ATT	AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCTGTGGGC TCTATGGCCC

	245
BGH_pc5	GGGtatc
BGH_AAA	GGGnnnn
BGH_AGA	GGGnnnn
BGH_AGT	GGGnnnn
BGH_ATA	GGGnnnn
BGH_ATT	GGGnnnn

GFP-Alignment:

HindIII (1), BamHI (735);

```

      ....|....|....|....|....|....|....|....|
      5      15      25      35      45      55
h_GFP(0)  AAGCTTGGCCA CCATGGTGTC CAAGGGGAG GAGCTGTTCA CAGGGGTGGT GCCCATCCTG
h_GFP(60) AAGCTTGGCCA CCATGGTGAG CAAGGGCGAG GAGCTGTTCA CCGGGGTGGT GCCCATCCTG
h_GFP(93) AAGCTTGGCCA CCATGGTGTC GAAGGGCGAG GAGCTGTTCA CCGCGTCTG GCCGATCCTC

      ....|....|....|....|....|....|....|....|
      65      75      85      95      105     115
h_GFP(0)  GTGGAGCTGG ATGGGGATGT GAATGGCCAC AAGTTCTCTG TGTCTGGGGA GGGGGAGGGG
h_GFP(60) GTCGAGCTGG ACGGCGACGT AAACGGCCAC AAGTTCAGCG TGTCCGGCGA GGGCGAGGGC
h_GFP(93) GTCGAACTCG ACGGCGACGT GAACGGCCAC AAGTTCAGCG TGAGCGGCGA GGGCGAGGGC

      ....|....|....|....|....|....|....|....|
      125     135     145     155     165     175
h_GFP(0)  GATGCCACCT ATGGCAAGCT CACCCTGAAG TTCATCTGCA CCACAGGCAA GCTGCCAGTG
h_GFP(60) GATGCCACCT ACGGCAAGCT GACCCTGAAG TTCATCTGCA CCACCGGCAA GCTGCCCGTG
h_GFP(93) GACGCGACGT ACGGCAAGCT CACCGTGAAG TTCATCTGCA CGACCGGCAA GCTGCCCGTG

      ....|....|....|....|....|....|....|....|
      185     195     205     215     225     235
h_GFP(0)  CCCTGGCCCA CCCTGGTGAC CACCTTACC TATGGGGTGC AGTGCTTCTC CAGATACCCA
h_GFP(60) CCCTGGCCCA CCCTCGTGAC CACCTTACC TACGGCGTGC AGTGCTTCTC CCGCTACCCC
h_GFP(93) CCGTGGCCGA CGCTCGTGAC GACGTTACCG TACGGCGTGC AGTGCTTCTC GCGCTACCCC

      ....|....|....|....|....|....|....|....|
      245     255     265     275     285     295
h_GFP(0)  GACCACATGA AGCAGCATGA CTTCTTCAAG TCTGCCATGC CTGAGGGGCTA TGTGCAGGAG
h_GFP(60) GACCACATGA AGCAGCACGA CTTCTTCAAG TCCGCCATGC CCGAAGGCTA CGTCCAGGAG
h_GFP(93) GACCACATGA AGCAGCACGA CTTCTTCAAG AGCGCGATGC CCGAGGGCTA CGTGCAGGAG

      ....|....|....|....|....|....|....|....|
      305     315     325     335     345     355
h_GFP(0)  AGGACCATCT TCTTCAAGGA TGATGGCAAC TACAAGACCA GGGCTGAGGT GAAGTTTGAG
h_GFP(60) CGCACCATCT TCTTCAAGGA CGACGGCAAC TACAAGACC GCGCCGAGGT GAAGTTTGAG
h_GFP(93) CCGACGATCT TCTTCAAGGA CGACGGCAAC TACAAGACCG GCGCTGAAGT GAAGTTTGAG

      ....|....|....|....|....|....|....|....|
      365     375     385     395     405     415
h_GFP(0)  GGGGATACCC TGGTGAACAG GATTGAGCTG AAGGGCATTG ACTTTAAGGA GGATGGCAAT
h_GFP(60) GGCGACACCC TGGTGAACCG CATCGAGCTG AAGGGCATCG ACTTCAAGGA GGACGGCAAC
h_GFP(93) GGCGATACCG TCGTGAACCG GATCGAGCTG AAGGGCATCG ACTTTAAGGA GGACGGCAAT

      ....|....|....|....|....|....|....|....|
      425     435     445     455     465     475
h_GFP(0)  ATCCTGGGCC ACAAGCTGGA GTACAACCTAC AACAGCCACA ATGTGTACAT CATGGCAGAC
h_GFP(60) ATCCTGGGGC ACAAGCTGGA GTACAACCTAC AACAGCCACA ACGTCTATAT CATGGCCGAC
h_GFP(93) ATCCTCGGCC ACAAGCTCGA GTACAACCTAC AACTCGCACA ACGTGTACAT CATGGCCGAC

      ....|....|....|....|....|....|....|....|
      485     495     505     515     525     535
h_GFP(0)  AAGCAGAAGA ATGGCATCAA GGTGAACTTC AAGATCAGGC ACAACATGTA GGATGGCTCT
h_GFP(60) AAGCAGAAGA ACGGCATCAA GGTGAACTTC AAGATCCGCC ACAACATCGA GGACGGCAGC
h_GFP(93) AAGCAGAAGA ACGGCATCAA AGTGAACTTC AAGATCCGGC ACAACATCGA GGACGGCAGC

      ....|....|....|....|....|....|....|....|
      545     555     565     575     585     595
h_GFP(0)  GTGCAGCTGG CAGACCACTA CCAGCAGAAC ACCCCCATG GAGATGGCCC TGTCTGCTCT
h_GFP(60) GTGCAGCTCG CCGACCACTA CCAGCAGAAC ACCCCCATCG GCGACGGCCC CGTGCTGCTG
h_GFP(93) GTGCAGCTCG CCGACCACTA CCAGCAGAAC ACCCCGATCG GCGACGGCCC CGTCTGCTCT

      ....|....|....|....|....|....|....|....|
      605     615     625     635     645     655
h_GFP(0)  CCAGACAACC ACTACCTGAG CACCCAGTCT GCCCTGAGCA AGGACCCCAA TGAGAAGAGG
h_GFP(60) CCCGACAACC ACTACCTGAG CACCCAGTCC GCCCTGAGCA AAGACCCCAA CGAGAAGCGC
h_GFP(93) CCCGACAACC ACTACCTGTC GACGAGAGC GCCCTGTGCA AGGACCCGAA CGAGAAGCGC

      ....|....|....|....|....|....|....|....|
      665     675     685     695     705     715
h_GFP(0)  GACCACATGG TGCTGCTGGA GTTGTGACA GCTGTGGCA TCACCCCTGG CATGGATGAG
h_GFP(60) GATCACATGG TCCTGCTGGA GTTCGTGACC GCCGCGGGA TCACTCTCGG CATGGACGAG
h_GFP(93) GACCACATGG TGCTGCTCGA GTTCGTGACC GCCGCTGCA TCACGCTCGG CATGGACGAG

```

```
.....|.....| .....|.....|  
      725      735  
h_GFP(0)  CTGTACAAGT GATAGGATCC  
h_GFP(60) CTGTACAAGT AATAGGATCC  
h_GFP(93) CTGTACAAGT GATAGGATCC
```

p24-Alignment:

HindIII (1), BamHI (708);

```

      ....|....| ....|....| ....|....| ....|....| ....|....|
      5      15      25      35      45      55
h_p24 (0)  AAGCTTATGC CCATTGTGCA GAACATCCAG GGCCAGATGG TGCACCAGGC CATCAGCCCC
h_p24 (38) AAGCTTATGC CCATCGTGCA GAACATCCAG GGCCAGATGG TGCACCAGGC CATCAGCCCC
h_p24 (104) AAGCTTATGC CGATCGTGCA GAACATCCAG GGCCAGATGG TGCACCAGGC GATCTCGCCG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      65      75      85      95      105     115
h_p24 (0)  AGGACCCTGA ATGCCTGGGT GAAGGTGGTG GAGGAGAAGG CCTTCAGCCC TGAGGTGATC
h_p24 (38) AGGACCCTGA ACGCCTGGGT GAAGGTGGTG GAGGAGAAGG CCTTCAGCCC CGAGGTGATC
h_p24 (104) CGGACCGTGA ACGCGTGGGT GAAAGTCGTC GAGGAGAAGG CGTTCTCGCC CGAAGTGATC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      125     135     145     155     165     175
h_p24 (0)  CCCATGTTCT CTGCCCTGTC TGAGGGGGCC ACCCCCAGG ACCTGAACAC CATGCTGAAC
h_p24 (38) CCCATGTTCA GCGCCCTGAG CGAGGGAGCC ACCCCCAGG ACCTGAACAC CATGCTGAAC
h_p24 (104) CCGATGTTCA GCGCGCTGAG CGAGGGCGCG ACCCGCAGG ACCTGAACAC GATGCTGAAC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      185     195     205     215     225     235
h_p24 (0)  ACAGTGGGGG GCCACCAGGC TGCCATGCAG ATGCTGAAGG AAACCATCAA TGAGGAGGCT
h_p24 (38) ACCGTGGGCG GCCACCAGGC CGCCATGCAG ATGCTGAAGG AGACCATCAA CGAGGAGGCC
h_p24 (104) ACCGTGCGCG GCCACCAGGC CGCGATGCAG ATGCTGAAGG AAACGATCAA CGAGGAGGCC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      245     255     265     275     285     295
h_p24 (0)  GCTGAGTGGG ACAGAGTGCA CCCTGTGCAT GCTGGCCCCA TTGCCCTTGG CCAGATGAGG
h_p24 (38) GCCGAGTGGG ACAGGGTGCA CCCCGTGCAC GCCGGCCCCA TCGCCCCGGG CCAGATGAGG
h_p24 (104) GCCGAGTGGG ACCGCGTGCA CCCCGTGCAC GCCGGACCGA TCGCGCCTGG CCAGATGCGG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      305     315     325     335     345     355
h_p24 (0)  GAGCCCAGGG GCTCTGACAT TGCTGGCACC ACCTCCACCC TGCAGGAGCA GATGCGCTGG
h_p24 (38) GAGCCCCGCG GCAGCGACAT CGCGGGCACC ACCAGCACCC TGCAGGAGCA GATCGGCTGG
h_p24 (104) GAACCGCGCG GAAGCGACAT CGCGGCACG ACCTCGACCG TGCAGGAGCA GATCGGCTGG

      ....|....| ....|....| ....|....| ....|....| ....|....|
      365     375     385     395     405     415
h_p24 (0)  ATGACCAACA ACCCCCCCAT CCCTGTGGGG GAGATCTACA AGAGATGGAT CATCCTGGGC
h_p24 (38) ATGACCAACA ACCCCCCCAT CCCCGTGGGC GAAATCTACA AGAGGTGGAT CATCTGGGC
h_p24 (104) ATGACGAACA ACCCGCCGAT CCCCGTCGCG GAGATCTACA AGCGGTGGAT CATCCTCGGC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      425     435     445     455     465     475
h_p24 (0)  CTGAACAAGA TTGTGAGGAT GTACAGCCCC ACCTCCATCC TGGACATCAG GCAGGGCCCC
h_p24 (38) CTGAACAAGA TCGTGAGGAT GTACAGCCCC ACCAGCATCC TGGATATCAG GCAGGGCCCC
h_p24 (104) CTGAACAAGA TCGTGGGAT GTACTCGCGG ACCTCGATCC TGGACATCCG GCAGGGCCCC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      485     495     505     515     525     535
h_p24 (0)  AAGGAGCCCT TCAGGGACTA TGTGGACAGG TTCTACAAGA CCCTGAGGGC TGAGCAGGCC
h_p24 (38) AAAGAGCCCT TCAGGGACTA CGTGGACAGG TTCTACAAGA CCCTGCGCGC CGAGCAGGCC
h_p24 (104) AAGGAGCCGT TCCGCGACTA CGTGCACCGG TTCTACAAGA CGCTGCGCGC CGAGCAGGCC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      545     555     565     575     585     595
h_p24 (0)  AGCCAGGAGG TGAAGAACTG GATGACAGAG ACCCTGCTGG TGCAGAATGC CAACCCTGAC
h_p24 (38) AGCCAGGAGG TGAAGAACTG GATGACCGAG ACCCTGCTGG TGCAGAACGC CAACCCCGAC
h_p24 (104) TCGCAGGAAG TGAAGAACTG GATGACCGAA ACCCTGCTCG TGCAGAACGC GAACCCCGAC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      605     615     625     635     645     655
h_p24 (0)  TGCAAGACCA TCCTGAAGGC CCTGGGCCCA GCTGCCACCC TGGAGGAGAT GATGACAGCC
h_p24 (38) TGCAAGACCA TCCTGAAGGC CCTGGGACCC GCCGCCACCC TGGAGGAGAT GATGACCGCC
h_p24 (104) TGCAAGACCA TCCTGAAGGC GCTCGGCCCA GCCCGCACGC TCGAGGAGAT GATGACCGCC

      ....|....| ....|....| ....|....| ....|....| ....|....|
      665     675     685     695     705
h_p24 (0)  TGCCAGGGGG TGGAGGCCCC TGGCCACAAG GCCAGGGTGC TGTAAATAGGA TCC
h_p24 (38) TGCCAGGGCG TGGGCGGCCC CGGCCACAAG GCCAGGGTGC TGTAAATAGGA TCC
h_p24 (104) TGCCAGGGCG TCGGCGGACC CGGCCACAAG GCCGCGTGC TGTAAATAGGA TCC

```

F.2.2 Aminosäuresequenzvergleiche:

Jeweils dargestellt ATG bzw. erste Aminosäure bis TAG

AS-Vergleich: amp_49 (pc5) vs. amp_0 (02-774) ATG (M) bis TAG (*):

	5 15 25 35 45 55
amp_0	MSIQHFRVAL IPFFAAFCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN SGKILESFRP
amp_49 pc5	MSIQHFRVAL IPFFAAFCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN SGKILESFRP

	65 75 85 95 105 115
amp_0	EERFPMSTF KVLCCGAVLS RIDAQEQQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL
amp_49 pc5	EERFPMSTF KVLCCGAVLS RIDAQEQQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL

	125 135 145 155 165 175
amp_0	CSAAITMSDN TAANLLTTI GGPKELTAFI HNMGDHVTRL DRWEPELNEA IPNDERDTM
amp_49 pc5	CSAAITMSDN TAANLLTTI GGPKELTAFI HNMGDHVTRL DRWEPELNEA IPNDERDTM

	185 195 205 215 225 235
amp_0	PVAMATTLRK LLTGELLTLA SRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS
amp_49 pc5	PVAMATTLRK LLTGELLTLA SRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS

	245 255 265 275 285
amp_0	RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW*
amp_49 pc5	RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW*

AS-Vergleich: hygr_103 (pc5) vs. hygr_0 (02-772) von erster AS (K) bis TAG (*):

	5 15 25 35 45 55
hygr_0	KKPELTATS V EKFLIEKFDS VSDLMLQSEG EESRAFSFDV GGRGYVLRVN SCADGFYKDR
hygr_103 pc5	KKPELTATS V EKFLIEKFDS VSDLMLQSEG EESRAFSFDV GGRGYVLRVN SCADGFYKDR

	65 75 85 95 105 115
hygr_0	YVYRHFASAA LPIPEVLDIG EFSESLTYCI SRRAQGVTLQ DLPETELPAV LQPVAEAMDA
hygr_103 pc5	YVYRHFASAA LPIPEVLDIG EFSESLTYCI SRRAQGVTLQ DLPETELPAV LQPVAEAMDA

	125 135 145 155 165 175
hygr_0	IAAADLSQTS GFGPFGPQGI GQYTTWRDFI CAIADPHVYH WQTVMDDTVS ASVAQALDEL
hygr_103 pc5	IAAADLSQTS GFGPFGPQGI GQYTTWRDFI CAIADPHVYH WQTVMDDTVS ASVAQALDEL

	185 195 205 215 225 235
hygr_0	MLWAEDCPEV RHLVHADFGS NNVLTDNNGRI TAVIDWSEAM FGDSQYEVAN IFFWRPWLAC
hygr_103 pc5	MLWAEDCPEV RHLVHADFGS NNVLTDNNGRI TAVIDWSEAM FGDSQYEVAN IFFWRPWLAC

	245 255 265 275 285 295
hygr_0	MEQQTRYFER RHPELAGSPR LRAYMLRIGL DQLYQSLVDG NFDAAWAQG RCDIVRSGA
hygr_103 pc5	MEQQTRYFER RHPELAGSPR LRAYMLRIGL DQLYQSLVDG NFDAAWAQG RCDIVRSGA

	305 315 325 335
hygr_0	GTVGRTQIAR RSAAVWTDGC VEVLADSGNR RPSTRPRAKE *
hygr_103 pc5	GTVGRTQIAR RSAAVWTDGC VEVLADSGNR RPSTRPRAKE *

AS-Vergleich: kana_42 (pc5) & kana_0 (041196) von ATG (M) bis TAG (*):

```

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      5      15      25      35      45      55
kana_0    MSHIQRETSC SRPRLNSNMD ADLYGYKWAR DNVGQSGATI YRLYGKPDAP ELFLKHGKGS
kana_42 pc5 MSHIQRETSC SRPRLNSNMD ADLYGYKWAR DNVGQSGATI YRLYGKPDAP ELFLKHGKGS

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      65      75      85      95      105     115
kana_0    VANDVTDEM VRLNWLTEFMP LPTIKHFIRT PDDAWLLTA IPGKTAFQVL EEPDSEGENI
kana_42 pc5 VANDVTDEM VRLNWLTEFMP LPTIKHFIRT PDDAWLLTA IPGKTAFQVL EEPDSEGENI

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      125     135     145     155     165     175
kana_0    VDALAVFLRR LHSIPVCNCP FNSDRVFRLA QAQSRMNNGL VDASDFDDER NGWPVEQVWK
kana_42 pc5 VDALAVFLRR LHSIPVCNCP FNSDRVFRLA QAQSRMNNGL VDASDFDDER NGWPVEQVWK

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      185     195     205     215     225     235
kana_0    EMHKLLPFSP DSVVTHGDFS LDNLIFDEGK LIGCIDVGRV GIADRYQDLA ILWNCLGEFS
kana_42 pc5 EMHKLLPFSP DSVVTHGDFS LDNLIFDEGK LIGCIDVGRV GIADRYQDLA ILWNCLGEFS

      ....|.....| .....|.....| .....|.....| ..
      245     255     265
kana_0    PSLQKRLFQK YGIDNPDMNK LQFHLMLDEF F*
kana_42 pc5 PSLQKRLFQK YGIDNPDMNK LQFHLMLDEF F*

```

AS-Vergleich der GFP-Varianten von ATG (M) bis Doppelstopp TGA TAG (* *)

```

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      5      15      25      35      45      55
h_GFP(0)  MVSKGEELFT GVPVILVELD GDVNGHKFSV SGEGEGDATY GKLTCLKFICT TGKLPVPWPT
h_GFP(60) MVSKGEELFT GVPVILVELD GDVNGHKFSV SGEGEGDATY GKLTCLKFICT TGKLPVPWPT
h_GFP(93) MVSKGEELFT GVPVILVELD GDVNGHKFSV SGEGEGDATY GKLTCLKFICT TGKLPVPWPT

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      65      75      85      95      105     115
h_GFP(0)  LVTTFTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEEDTL
h_GFP(60) LVTTFTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEEDTL
h_GFP(93) LVTTFTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEEDTL

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      125     135     145     155     165     175
h_GFP(0)  VNRIELKGID FKEDGNILGH KLEYNYNSHN VYIMADKQKN GIKVNFKIRH NIEDGGSVQLA
h_GFP(60) VNRIELKGID FKEDGNILGH KLEYNYNSHN VYIMADKQKN GIKVNFKIRH NIEDGGSVQLA
h_GFP(93) VNRIELKGID FKEDGNILGH KLEYNYNSHN VYIMADKQKN GIKVNFKIRH NIEDGGSVQLA

      ....|.....| .....|.....| .....|.....| .....|.....| .....|.....|
      185     195     205     215     225     235
h_GFP(0)  DHYQQNTPIG DGPVLLPDNH YLSTQSALS DPNEKRDHMV LLEFVTAAGI TLGMDLYK*
h_GFP(60) DHYQQNTPIG DGPVLLPDNH YLSTQSALS DPNEKRDHMV LLEFVTAAGI TLGMDLYK*
h_GFP(93) DHYQQNTPIG DGPVLLPDNH YLSTQSALS DPNEKRDHMV LLEFVTAAGI TLGMDLYK*

.
h_GFP(0)  **
h_GFP(60) **
h_GFP(93) **

```

AS-Vergleich der p24-Varianten von ATG (M) bis Doppelstopp TAA TAG (* *)

	
		5	15	25	35	45	55
h_p24 (0)		MPIVQNIQQQ	MVHQAI SPRT	LNAWVKVVEE	KAFSPEVIPM	FSALSEGATP	QDLNTMLNTV
h_p24 (38)		MPIVQNIQQQ	MVHQAI SPRT	LNAWVKVVEE	KAFSPEVIPM	FSALSEGATP	QDLNTMLNTV
h_p24 (104)		MPIVQNIQQQ	MVHQAI SPRT	LNAWVKVVEE	KAFSPEVIPM	FSALSEGATP	QDLNTMLNTV
	
		65	75	85	95	105	115
h_p24 (0)		GGHQAAMQML	KETINEEAAE	WDRVHPVHAG	PIAPGQMREP	RGSDIAGTTS	TLQEQIGWMT
h_p24 (38)		GGHQAAMQML	KETINEEAAE	WDRVHPVHAG	PIAPGQMREP	RGSDIAGTTS	TLQEQIGWMT
h_p24 (104)		GGHQAAMQML	KETINEEAAE	WDRVHPVHAG	PIAPGQMREP	RGSDIAGTTS	TLQEQIGWMT
	
		125	135	145	155	165	175
h_p24 (0)		NNPPIPVGEEI	YKRWIILGLN	KIVRMYSPTS	ILDIRQGPKE	PFRDYVDRFY	KTLRAEQASQ
h_p24 (38)		NNPPIPVGEEI	YKRWIILGLN	KIVRMYSPTS	ILDIRQGPKE	PFRDYVDRFY	KTLRAEQASQ
h_p24 (104)		NNPPIPVGEEI	YKRWIILGLN	KIVRMYSPTS	ILDIRQGPKE	PFRDYVDRFY	KTLRAEQASQ
	
		185	195	205	215	225	
h_p24 (0)		EVKNWMTETL	LVQNaNPDCK	TILKALGPAA	TLEEMMTACQ	GVGGPGHKAR	VL**
h_p24 (38)		EVKNWMTETL	LVQNaNPDCK	TILKALGPAA	TLEEMMTACQ	GVGGPGHKAR	VL**
h_p24 (104)		EVKNWMTETL	LVQNaNPDCK	TILKALGPAA	TLEEMMTACQ	GVGGPGHKAR	VL**

F.3 Sequenzen:

F.3.1 Sequenzen der Vektorelemente:

Aufgeführt werden jeweils die singulären flankierenden Restriktionsschnittstellen der einzelnen Vektorelemente sowie der CpG Gehalt. Positionsangaben beziehen sich jeweils auf die nachfolgende Sequenz.

CMV_TG-Promotor: 0 + 1 CG (ClaI); ClaI (5), SacI (593);

```

1      gtttaATCGAT GTTGACATTG ATTATGACT AGTTATTAAT AGTAATCAAT TATGGGGTCA
61     TTAGTTCATA GCCCATATAT GGAGTTCTGT GTTACATAAC TTATGGTAAA TGGCCTGCCT
121    GGCTGACTGC CCAATGACCC CTGCCATTG ATGTCAATAA TGATGTATGT TCCCATAGTA
181    ATGCCAATAG GGACTTTCCA TTGATGTCAA TGGGTGGACT ATTTATGGTA AACTGCCCCAC
241    TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTATGCCCC CTATTGATGT CAATGATGGT
301    AAATGGCCTG CCTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
361    TACATCTATG TATTAGTCAT TGCTATTACC ATGGTGATGT GGTTTTGGCA GTACATCAAT
421    GGGTGTGGAT AGTGGTTTGA CTCATGGGGA TTTCCAAGTC TCCACCCCAT TGATGTCAAT
481    GGGAGTTTGT TTTGGCACCA AAATCAATGG GACTTTCCAA AATGTTGTAA CAACTCTGCC
541    CCATTGATGC AAATGGGTGG TAGGTGTGTA TGGTGGGAGG TCTATATAAG CAGAGCTCat
601    tc

```

CMV-Promotor: 31 + 1 CG (ClaI); ClaI (5), SacI (593);

```

1      gtttaATCGAT GTTGACATTG ATTATGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA
61     TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT
121    GGCTGACCGC CCAACGACCC CCGCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA
181    ACGCCAATAG GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCCAC
241    TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT
301    AAATGGCCCG CCTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG
361    TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT
421    GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT
481    GGGAGTTTGT TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC
541    CCATTGACGC AAATGGCCCG TAGGCGTGTG CCGTGGGAGG TCTATATAAG CAGAGCTCat
601    tc

```

BGH P(A): 3 + 1 CG (XmaI); PmeI (5), XmaI (238);

```

1      cagtGTTTAA ACCtTGTGCCT TCTAGTTGCC AGCCATCTGT TGTGTTGCCCC TCCCCCGTGC
61     CTTCCCTGAC CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATG
121    CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA
181    AGGGGGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGGGTGGGC TCTATGGCCC
241    GGGtatc

```

pUC Origin (antisense-Orientierung): 45 + 1 CG (XmaI/BglIII), 45 CG (EcoRI/BglIII);

XmaI (5), EcoRI (15), BglIII (695), ClaI (705)

```

1      gtcaCCCCGG TAGTGAATTC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA
61     AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCT TGACGAGCAT CACAAAAATC
121    GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC
181    CTGGAAGCTC CCTCGTGGC TCTCCTGTTT CGACCCTGCC GCTTACCCGA TACCTGTCCG
241    CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT
301    CCGTGTAGGT CGTTGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCGTT CAGCCCGACC
361    GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
421    CACTGGCAGC AGCCACTGGT AACAGGATTA CGAGAGCGAG GTATGTAGGC GGTGCTACAG
481    AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATT GGTATCTGCG
541    CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
601    CCACCCTGG TAGCGTGGT TTTTGTGTTT GCAAGCAGCA GATTACCGC AGAAAAAAG
661    GATCTCAAGA AGATCCTTG ATCTTTCTA CGGGAGATCT GTCAATCGAT agct

```

110⁺pUC origin (antisense-Orientierung): 39 + 1 CG (*Xma*I); *Xma*I (1), *Bgl*III (580);

```

1      CCCGGG TAGT GAATTC ATGT GAGCAAAGG CCAGCAAAG GCCAGGAACC GTAAAAAGGC
61     CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA AAAATCGAGC
121    CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCTGG
181    AAGTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT
241    FCTCCCTTCG GGAAGCGTGG CGCTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT
301    GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCGTTCAGC CCGACCGCTG
361    CGCCTTATCC GGTAACATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT
421    GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTCACAGATT
481    CTTGAAGTGG TGGCTAACC ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT
541    GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT AGATCT

```

Amp^r 1: 1 + 1 CG (*Cl*aI);

*Cl*aI (5), *Bgl*III (1401)

```

1      gtcc ATCGAT TAGTATATCT GGCCAGTACA TCTCCAAGCA GTGCAAAAGG CCTAACCCCTA
61     AGCAGATTCT TCATGCAATT GTAGGTCAAG CCTTGCCTTG TTGTAGCTTA AATTTTGCTA
121    GCCCACTACT CAGTGACCTC CAACACACAA GCAGGGAGCA GATACTGGCT TAACATGAG
181    GCATCAGAGC AGATTGTACT GAGAGTGCAC CATAGGGGAT TGGGAGAACT CCCTATCAGT
241    CAATTTTCAGG TGGCACTTTT CAGGGAAATG TGAGCTGAT GCA TACCCCT ATTTGTTTAT
301    TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCCTGA TAAATGCTTC
361    AATAATATTG AAAAAAGGAG AGTAGTACTA TGAGTATTCA ACATTCAGA GTTGCCCTTA
421    TTCCCTTTTT TGCTGCCTTT TGCCTTCCTG TTTTGTCTCA CCCAGAAACC CTGGTGAAG
481    TAAAAGATGC TGAAGATCAG TTGGGTGCAA GAGTGGGTTA CATTGAACTG GATCTCAACT
541    CTGGTAAGAT CCTTGAGAGT TTTAGACCAG AAGAAAGATT TCCAATGATG AGCACCTTA
601    AAGTTCTGCT ATGTGGTGCA GTATTATCCA GAATTGATGC TGGGCAAGAG CAACTGGGTA
661    GAAGAATACA CTATTCTCAG AATGACTTGG TTGAGTACAG TCCAGTCACA GAAAAGCATC
721    TTACAGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA
781    CTGCTGCCAA CTTACTTTTG ACAACCATTG GAGGACCAA GGAGTTGACT GCTTTTTTGC
841    ACAACATGGG TGATCATGTA ACTAGGCTTG ATAGATGGGA ACCAGAGCTG AATGAAGCCA
901    TACCAATGTA CAGGATGTA ACCACCATGC CTGTAGCAAT GGCAACAACC TTGAGGAAC
961    TATTAAGTGG TGAAGTCTT ACTCTAGCTT CCAGACAACA ATTAATAGAC TGGATGGAGG
1021   CAGATAAAGT TGCAGGACCA CTTCTGAGGT CTGCCCTTCC AGCTGGCTGG TTCAATGCTG
1081   ATAAATCTGG AGCAGGTGAG AGAGGGTCTA GAGGTATCAT TGCAGCACTG GGGCCAGATG
1141   GTAAGCCCTC CAGGATGTA GTTATCTACA CCACAGGCAG TCAGGCAACC ATGGATGAGA
1201   GAAATAGGCA GATTGACAGG ATTGGTGCCT CATTGATTAA GCATTGGTAA GTTAACTCTG
1261   CAGACCAAGT TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA
1321   GGATCTAGGT GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CATGAGTTTT
1381   GGTTCCTACT AGAGTCAGAC AGATCT agta

```

Hygr^r 1: 1 + 1 CG (*Xma*I);

*Xma*I (5), *Eco*RI (2046)

```

1      tagt CCCGGG CTTCTGAGGC TGAAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACTCA
61     CCCTGTAGCA GTGCATTAAG CTCAGCAGGT GTGTTGGTTA GGCTCAGTGT GACTGCTACA
121    CTTGCCAGTG CCCTAGCACC AGCTCCTTTC ACTTCTTCC CTTCCTTCT CACCACATTC
181    GCAGGCTTTC CCAGTCAAGC TCTAAATTGG GGGTCCCTTT AGGGTTCAGA TTTAGTGCTT
241    TACTGCACCT GGACCCCAA AAACCTGATT AGGGTGATGG TTCTTGTACC TAGAAGTTCC
301    TATTCGAAG TTCCTATTCT CTAGAAAGTA TAGGAACTTC CTTGGCCCTT AAGAAAAGC
361    CTGAACCTAC TGCCACCTCT GTGGAGAAGT TTCTTATTGA AAAATTTGAC TCTGTCACTG
421    AACTTGATGCA CCTCAGTGAA GGTGAAGAAT CTAGGGCTTT CAGCTTTGAT GTAGGAGGGA
481    GAGGATATGT GCTGAGAGTA AATAGCTGTG CAGATGGTTT CTACAAAGAT AGATATGTGT
541    ATAGGCACTT TGCATCTGCA GCCCTCCCAA TTCCAGAAGT GTTAGACATT GGGGAATTTT
601    CTGAGAGCTT GACCTATTGC ATCTCCAGAA GAGCACAGGG TGTACCTTGC CAAGACCTGC
661    CTGAAACTGA ACTGCCTGCT GTGCTGCAGC CTGTGGCAGA GGCCATGGAT GCAATTGCTG
721    CAGCAGACCT TAGCCAGACT TCTGGGTTTTG GCCCATTTGG ACCACAAGGA ATTGGTCAAT
781    AACTACATG GAGAGATTTT ATATGTGCCA TTGCAGATCC CCATGTGTAT CATTGGCAAA
841    CTGTGATGGA TGACACTGTG AGTGCAAGTG TGGCCCAAGC TCTGGATGAG CTTATGCTTT
901    GGGCAGAGGA CTGCCCTGAA GTGAGGCATC TGGTTCATGC AGATTTTGGC TCCAACAATG
961    TCTTAACAGA CAATGGCAGA ATAACAGCAG TGATTGACTG GAGTGAAGCC ATGTTTGGTG
1021   ATTCCCAATA TGAGGTGGCC AACATATTCT TCTGGAGGCC TTGGTTGGCT TGTATGGAAC
1081   AGCAGACCAG ATACTTTGAG AGAAGGCATC CTGAGCTGGC AGGAAGCCCA AGACTCAGGG
1141   CCTATATGCT CAGAATTGGT CTTGACCAAC TGTATCAGAG CTTGGTTGAT GGCAATTTTG
1201   ATGATGCAGC TTGGGCCAGG GGAAGGTGTG ATGCAATTGT CAGGTCTGGA GCAGGGACTG
1261   TGGGGAGAAC ACAAATTGCC AGGAGAAGTG CAGCTGTTTG GACAGATGGC TGTGTAGAGG
1321   TGCTGGCAGA TAGTGAAAC AGAAGGCCCA GCACTAGACC AAGGGCAAAG GAATAGATTT
1381   AAAATCATGTA CTAAGAGATT TGGATTCCAC TGCAGCCTTC TATGAAAGGT TGGCTTCAG
1441   AATGGTTTTT CAGGACTCCA GCTGGATGAT CCTCCAGCA TGGGATCTCA TGCTGGAGTT
1501   CTTCTCCAC CCAAGACACTA GTC AACTTGT TTATTGCAGC TTATAATGGT TACAAATAAA
1561   GCAATAGCAT CACAAATTTT ACAAATAAAG CATTTTTTTC ACTGCATCTCT AGTGTGGTTT
1621   TGTCCAACAT CATCAATGTA TCTTATCATG TCTGTTAATT AAATATACCAT TGACCTCTAG
1681   CTAGAGCTTG GAGTAATCAT GGTTCATAGCT GTTTCCTGTG TGAATTTGTT ATCAGCTCAC
1741   AATCCACAC AACATACTAG CAGGAAGCAT AAAGTGATAA CCCTGGGGTG CTAATGAGT
1801   GAGCTAACTC ACATTAATTG CCTTCTCTC ACTGCCAGCT TTCCAGTCAG GAAACCTGTA

```

```

1861 GTGCCAGCTG CATTAAATGAA TCTGCCAACT CACTGGGAGA GGAGGTTTGC CTATTGGGCA
1921 CTCTTCAGCT TCCTAGCTCA CTGACTGGCT GCACTCAGTG GTTCAGCTGC TGCTAGCAGT
1981 ATCAGCTCAC TCAAAGCTG TAATAGGGT ATCCACAGAA TCAGGGGATA ACTCAGGAAA
2041 GAACGAATT Cagggt

```

MCS_1: 1 CG(T7-Prom.)

SacI (5), *PmeI* (132)

```

1      tgtaGAGCTC TCTGGCTAAC TAGAGAACCC ACTGCTTACT GGCTTATCTA AATTAATACG
61     ACTCACTATA GGGAGACCCA AGCTGTTAAG CTTGGTA GAT ATCAGGGATC CACTCAGCTG
121    ATCAGCCTCC AGTTTAAACtgta

```

kana_0: 0 CG

ScaI (1), *HpaI* (823)

```

1      AGTACTATGA GCCATATTCA AAGAGAAACC TCTTGCTCTA GGCCAAGGTT AAATTCCAAC
61     ATGGATGCTG ATTTATATGG GTATAAATGG GCTAGAGATA ATGTTGGGCA ATCAGGTGCC
121    ACAATCTACA GGTGTATATG GAAGCCAGAT GCACCAGAGT TGTTCCTGAA ACATGGCAAA
181    GGTTCTGTTG CCAATGATGT TACAGATGAG ATGGTCAGAC TAAACTGGCT GACTGAATTT
241    ATGCCCTCTT CAACCATCAA GCATTTTATC AGGACTCCTG ATGATGCATG GTTACTCACC
301    ACTGCCATCC CAGGAAAAAC AGCATTCAGG GTATTAGAAG AATATCCTGA TTCAGGTGAA
361    AATATTGTTG ATGCCCTGGC AGTGTTCCTG AGAAGGTTGC ATTCTATTC TGTTCGTAAT
421    TGTCCCTTTA ACTCTGATAG AGTATTTAGG CTGGCTCAGG CCAATCAAG AATGAATAAT
481    GGTTCGTTG ATGCCAGTGA TTTTGATGAT GAGAGAAATG GCTGGCCTGT TGAACAAGTC
541    TGGAAAGAAA TGCACAAACT TTTGCCATTG TCACCAGATT CAGTGGTCAC TCATGGTGAT
601    TTCTCACTTG ATAACCTTAT TTTTGATGAG GGGAAATTA TAGGTTGTAT TGATGTTGGA
661    AGAGTGGGAA TTGCAGACAG ATACCAGGAT CTTGCCATCC TATGGAACCT CCTGGGTGAG
721    TTTTCTCCTT CATTACAGAA AAGGCTTTTT CAAAAATATG GTATTGATAA TCCTGATATG
781    AATAAATTGC AGTTTCATTT GATGCTGGAT GAGTTTTTCT AAGTTAAC

```

delBGH BamHI/MscI:

BamHI (1), *MscI* (672)

108-332: BGH P(A); 622-669: FRT-Element

```

1      GGATCCACTA GTCCAGTGTG GTGGAATTCT GCAGATATCC AGCACAGTGG CCGCCGCTCG
61     AGTCTAGAGG GCCCGTTTAA ACCCGCTGAT CAGCCTCGAG TTTAAACCTG TGCCTTCTAG
121    TTGCCAGCCA TCTGTGTTT GCCCCTCCCC AGTGCCTTCC TTGACCCCTGG AAGGTGCCAC
181    TCCCACTGTC CTTTCCTAAT AAAATGAGGA AATTGCAGGG CATTGTCTGA GTAGGTGTCA
241    TTCTATCTCG GGGGTGGGG TGGGGCAGGA CAGCAAGGGG GAGGATTGGG AAGACAATAG
301    CAGGCATGCT GGGGATGCAG TGGGCTCTAT GGCCCGGGCT TCTGAGGCGG AAAGAACCAG
361    CTGGGGCTCT AGGGGTATC CCCACGCGCC CTGTAGCGGC CTATTAAGCG CCGCGGGTGT
421    GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCTTTTCCG
481    TTTCTTCCCT TCCTTCTCG CCACGTTCCG CCGCTTTCCC CGTCAAGCTC TAAATCGGGG
541    GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA
601    GGGTGTGGT TCACGTACCT AGAAGTTCC TTTCCGAAGT TCCTATTCTC TAGAAAGTAT
661    AGGAACCTCC TTGGCCA

```

F.3.2 Sequenzen eukaryontischer Expressionsplasmide:

Dargestellte Vektorsequenzen beginnen mit der Sequenz des CMV-Promotors bzw. der vorangehenden singulären Schnittstelle. Aufgeführt werden die Sequenzen der Plasmide pRS, pΔS sowie pΔS110'. Alle weiteren Vektoren können durch Austausch von Vektorelementen, wie unter Kapitel C.2.2 beschrieben, generiert werden.

pRS (3085 bp): 164 CG

```

1      GTTGACATTG ATTATGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA
61     GCCCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCCG
121    CCAACGACCC CCGCCATTG ACGTCAATAA TGACGTATGT TCCATAGTA ACGCCAATAG
181    GGACTTTCCA TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCAC TTGGCAGTAC
241    ATCAAGTGTA TCATATGCCA AGTACGCCCT CTATGACGT CAATGACGGT AAATGGCCCG
301    CTTGGCATTG TGCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG
361    TATTAGTCAT GCCTATTACC ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT
421    AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT
481    TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA CACTCCGCC CCATTGACGC
541    AAATGGGCGG TAGGCGTGTG CGGTGGGAGG TCTATATAAG CAGAGCTCTC TGGCTAACTA
601    GAGAACCAC TGGCTTACTG CTTATCGAAA TTAATACGAC TCACTATAGG GAGACCCAAG
661    CTGGCTAGCG TTTAAACTTA AGCTTGGTAC CGAGCTCGGA TCCACTAGTC CAGTGTGGTG
721    GAATTCGTGA GATATCCAGC ACAGTGGCGG CCGCTCGAGT CTAGAGGGCC CGTTAAACC
781    CGCTGATCAG CCTCGACTGT GCCTTCTAGT TGCCAGCCAT CTGTGTTTGG CCCTCCCCC
841    GTCCCTTCC TACCTTGGG AGGTGCCACT CCCACTGTCC TTTCTAATA AATGAGGAA
901    ATTGCATCGC ATTGTCTGAG TAGGTGTCAT TCTATTCTGG GGGGTGGGGT GGGGCAGGAC

```

961 AGCAAGGGGG AGGATTGGGA AGACAATAGC AGGCATGCTG GGGATGCGGT GGGCTCTATG
 1021 GATGCATCCG TACGTTAACA TGTGAGCAA AGGCCAGCAA AAGGCCAGGA ACCGTA AAAA
 1081 GGCCCGGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCT GACGAGCATC ACAAAAAATCG
 1141 ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATAACCAGG CGTTTCCCCC
 1201 TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCGC CTTACCGGAT ACCTGTCCCG
 1261 CTTTCTCCCT TCGGGAAGCG TGCGCGTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC
 1321 GGTGTGAAG GTTCCGTCCA AGCTGGGCTG TGTGCACGAA CCCCCGTT CAGCCGACCG
 1381 CTGCGCCTTA TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC
 1441 ACTGGCAGCA GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGCGG GTGCTACAGA
 1501 GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC
 1561 TCTGCTGAAG CCGGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAAACAAC
 1621 CACCGCTGGT AGCGGTGGTT TTTTGTGTTG CAAGCAGCAG ATTACGCGCA GAAAAAAGG
 1681 ATCTCAAGAA GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGA ACGAAAATC
 1741 ACGTTAAGGG ATTTTGGTCA TGAGATTATC AAAAAAGGATC TTCACCTAGA TCCTTTTAAA
 1801 TTAAAAATGA GTTTTAAAT CAATCTAAAG TATATATGAG TAAACTTGGT CTGACAGTTA
 1861 CCAATGCTTA ATCAGTGAGG CACCTATCTC AGCGATCTGT CTATTTCTGT CATCCATAGT
 1921 TGCCCTGACTC CCCGTCGTGT AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG
 1981 TGCTGCAATG ATACCCGCGAG ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA
 2041 GCGAGCCGGA AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCCGCT CCATCCAGTC
 2101 TATTAATTGT TGCCGGGAAG CTAGAGTAAG TAGTTCGCCA GTTAATAGTT TGCGCAACGT
 2161 TGTGTCATT GCTACAGGCA TCGTGGTGTG ACGTCTGTCG TTTGGTATGG CTTCAATCAG
 2221 TCCCGGTTCC CAACGATCAA GGCGAGTTAC ATGATCCCC ATGTTGTGCA AAAAAGCGGT
 2281 TAGCTCCTTC GTTCTTCCGA TCGTGTCTAG AAGTAAGTTG CCGCAGTGT TACTACTCAT
 2341 GGTTATGGCA GCACTGCATA ATTCTCTTAC TGTTCATGCCA TCCGTAAGAT GCTTTTCTGT
 2401 GACTGTGAG TACTCAACCA AGTCATTCTG AGAATAGTGT ATGCGGCGAC CGAGTTGCTC
 2461 TTGCCCCGGC TCAATACGGG ATAATACCGC GCCACATAGC AGAACTTTAA AAGTGCTCAT
 2521 CATTGGAAAG GTTCTTCGG GCGGAAAACT CTC AAGGATC TTACCGCTGT TGAGATCCAG
 2581 TTCGATGTAA CCCACTCGTG CACCCAACCTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT
 2641 TTCTGGGTGA GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA GGGCGACACG
 2701 GAAATGTTGA ATACTCATA TCTTCTTTT TCAATATTAT TGAAGCATT ATCAGGGTTA
 2761 TTGTCTCATG AGCGGATACA TATTTGAATG TATTTAGAAA AATAAACAAA TAGGGGTTCC
 2821 GCGCACATTT CCCCAGAAAG TGCCACCTGA CGTCGACGGA TCGGGAGATC TCCGATCCC
 2881 CTATGGTGCA CTCTCAGTAC AATCTGCTCT GATGCCGAT AGTTAAGCCA GTATCTGCTC
 2941 CCTGCTTGTG TGTGGAGGTT CGCTGAGTAG TGCCGAGCA AAATTTAAGC TACAACAAGG
 3001 CAAGGCTTGA CCGACAATTG CATGAAGAAT CTGCTTAGGG TTAGGCGTTT TGCGCTGCTT
 3061 CGCGATGTAC GGGCCAGATA TACGC

pAs (3034 bp): **83** CG

1 ATCGATGTTG ACATTGATTA TTGACTAGTT ATTAATAGTA ATCAATTACG GGGTCATTAG
 61 TTCATAGCCC ATATATGGAG TTCCGCGTTA CATAACTTAC GGTAATAGGC CCGCTGGCT
 121 GACCGCCCAA CGACCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC ATAGTAACGC
 181 CAATAGGGGA TTTCATTTGA CGTCAATGGG TGGAGTATTT ACGTAAACT GCCACTTGG
 241 CAGTACATCA AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT GACGGTAAAT
 301 GGCCCGCCTG GCATTATGCC CAGTACATGA CCTTATGGGA CTTTCTACT TGGCAGTACA
 361 TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC ATCAATGGGC
 421 GTGGATAGCG GTTTGACTCA CCGGGATTTT CAAGTCTCCA CCCCATTGAC GTCAATGGGA
 481 GTTTGTTTTG GCACCAAAAT CAACGGGACT TTCCAAAATG TCGTAACAAC TCCGCCCAT
 541 TGACGCAAAAT GGGCGGTAGG CGTGTACGGT GGGAGGTCTA TATAAGCAGA GCTCTCTGGC
 601 TAACTAGAGA AACCACTGCT TACTGGCTTA TCTAAATTA TACGACTCAC TATAGGGAGA
 661 CCAAAGCTGT TAACTTGGT AGATATCAGG GATCCACTCA GCTGATCAGC CTCCAGTTA
 721 AACCTGTGCC TTCTAGTTGC CAGCCATCTG TTGTTTGCCC CTCCCCGTG CCTTCTTGA
 781 CCCTGGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAAT GCATCGCATT
 841 GTCTGAGTAG GTGTCAATCT ATTCTGGGGG GTGGGGTGGG GCAGGACAGC AAGGGGGAGG
 901 ATTGGGAAGA CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGCC CGGTAGTGA
 961 ATTCATGTGA GCAAAAAGGC AGCAAAAAGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC
 1021 GTTTTCCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG
 1081 GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
 1141 GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CCGATACCTG TCCGCTTTC TCCTTCGGG
 1201 AAGCGTGGCG CTTTCTCATA GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTGCTTCG
 1261 CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG
 1321 TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC
 1381 TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGT
 1441 GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TCGCTCTGC TGAAGCCAGT
 1501 TACCTTCGGA AAAAGAGTTG GTAGCTCTT ATCCGGCAAA CAAACCACCG CTGGTAGCGG
 1561 TGGTTTTTTT GTTTGCAAGC AGCAGATTAC CCGCAGAAAA AAAGGATCTC AAGAAGATCC
 1621 TTTGATCTTT TCTACGGGAG ATCTGTCTGA CTCTCAGTGG AACCAAAACT CATGTTAAG
 1681 GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA ATTAATAATG
 1741 AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGGT TAACCTACCA
 1801 ATGCTTAATC AATGAGGCAC CAATCTCTGC AATCTGCCTA TTTCTCTCAT CCATGGTTGC
 1861 CTGACTGCCT GTGGTGTAGA TAACTACAAT CCTGGAGGGC TTACCATCTG CCCCAGTGC
 1921 TGCAATGATA CCTCTAGACC CTCTCTCACC TGCTCCAGAT TTATCTGCAA TGAACCAGCC
 1981 AGCTGGAAGG GCAGACTCA GAAGTGGTCC TGCAACTTTA TCTGCTCCA TCCAGTCTAT
 2041 TAATTTGTTG CTGGAAGCTA GAGTAAGCAG TTCACAGTT AATAGTTTCC TCAAGGTTGT
 2101 TGCCATTGCT ACAGGCATGG TGGTGTCCCT CTCATCATTT GGTATGGCTT CATTCAGCTC
 2161 TGGTTCCCAT CTATCAAGCC TAGTTACATG ATCACCCTAT TGTGCAAAA AAGCAGTCAA
 2221 CTCCTTTGGT CCTCAATGG TTGTCAAAAG TAAGTTGGCA GCAGTGTAT CACTCATGGT

```

2281 TATGGCAGCA CTGCATAAATT CTCTTACTGT CATGCCATCT GTAAGATGCT TTTCTGTGAC
2341 TGGACTGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATT CTCTACCCA GTTGCTCTTG
2401 CCCAGCATCA ATTCTGGATA ATACTGCACC ACATAGCAGA ACTTTAAAGG TGCTCATCAT
2461 TGGAAATCTT TCTTCTGGTC TAAAACCTCTC AAGGATCTTA CCAGAGTTGA GATCCAGTTC
2521 AATGTAACCC ACTCTTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGGGTTTC
2581 TGGGTGAGCA AAAACAGGAA GGCAAAAAGC AGCAAAAAGG GGAATAAGGG CAACTCTGAA
2641 ATGTTGAATA CTCATAGTAC TACTCTTCCT TTTTCAATAT TATGAAGCA TTTATCAGGG
2701 TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
2761 ATGCATTCAG CTCACATTTT CCTGAAAAGT GCCACCTGAA ATTGACTGAT AGGGAGTTCT
2821 CCCAATCCCC TATGGTGCAC TCTCAGTACA ATCTGCTCTG ATGCCTCATA GTTAAGCCAG
2881 TACTTGCTCC CTGCTTGTGT GTTGGAGGTC ACTGAGTAGT GGGCTAGCAA AATTTAAGCT
2941 ACAACAAGGC AAGGCTTGAC CTACAATTGC ATGAAGAATC TGCTTAGGGT TAGGCCTTTT
3001 GCACTGCTTG GAGATGTACT GGCCAGATAT ACTA

```

pAS110⁻ (2924 bp): 77 CG

```

1 ATCGATGTTG ACATTGATTA TTGACTAGTT ATTAATAGTA ATCAATTACG GGGTCATTAG
61 TTCATAGCCC ATATATGGAG TTCCGCGTTA CATAACTTAC GGTAAATGGC CGCCTGGCT
121 GACCGCCCAA AGACCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC ATAGTAACGC
181 CAATAGGGAC TTTCCATTGA CGTCAATGGG TGGAGTATTT ACGGTAAACT GCCCACTGG
241 CAGTACATCA AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT GACGGTAAAT
301 GGCCCGCCTG GCATTTATGCC CAGTACATGA CCTTATGGGA CTTTCTACT TGGCAGTACA
361 TCTACTGATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC ATCAATGGGC
421 GTGGATAGCG GTTTGACTCA CGGGGATTTT CAAGTCTCCA CCCCATTTGAC GTCAATGGGA
481 GTTTGTTTTG GCACCAAAAT CAACGGGACT TTCAAAAATG TCGTAACAAC TCCGCCCAT
541 TGACGCAAAAT GGGCGGTAGG CGTGTACGGT GGGAGGCTA TATAAGCAGA GCTCTCTGGC
601 TAAC TAGAGA ACCCACTGCT TACTGGCTTA TCTAAATTA TACGACTCAC TATAGGGAGA
661 CCCAAGCTGT TAAGCTTGGT AGATATCAGG GATCCACTCA GCTGATCAGC CTCCAGTTTA
721 AACCTGTGCC TTCTAGTTGC CAGCCATCTG TTGTTTGCC CTCCCCGTG CCTTCTTGA
781 CCCTGGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT
841 GTCTGAGTAG GTGCTACTCT ATTCTGGGGG GTGGGGTGGG GCAGGACAGC AAGGGGGAGG
901 ATTGGGAAGA CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGCC CGGGTAGTGA
961 ATTCATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC
1021 GTTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG
1081 GTGGCGAAGA ACCCACTGCT TATAAAGATA CCAGGCGTTT CCCCTGGAA GCTCCCTCGT
1141 GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCTTTC TCCCTTCGGG
1201 AAGCGTGGCG CTTTTCTCATA GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG
1261 TC CAAAGCTG GGTGTGTGC ACCAAACCCC CGTTCAGCCC GACCGCTGCG CTTATCCGG
1321 TAAC TATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC
1381 TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG
1441 GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TCGGCTCTGC TGAAGCCAGT
1501 TACCTTCGGA AAAAGAGTTG GTAGCTCTAG ATCTGTCTGA CTCTCAGTGG AACCAAACT
1561 CATGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACTTAG ATCTTTTAA
1621 ATTA AAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGGT
1681 TAAC TTACCA ATGCTTAATC AATGAGGCAC CAATCTCTGC AATCTGCCTA TTTCTCTCAT
1741 CCATGGTTGC CTGACTGCCT GTGGTGTAGA TAACTACAAT CCTGGAGGGC TTACCATCTG
1801 GCCCCAGTGC TGCAATGATA CCTCTAGACC CTCTCTCACC TGCTCCAGAT TTATCTGCAA
1861 TGAACCAGCC AGCTGGAAGG GCAGACCTCA GAAGTGGTCC TGCAACTTTA TCTGCCTCCA
1921 FCCAGTCTAT TAATTGTTGT CTGGAAGCTA GAGTAAGCAG TTCACCAGTT AATAGTTTCC
1981 TCAAGGTTGT TGCCATTGCT ACAGGCATGG TGGTGTCCCT CTCATCATTT GGTATGGCTT
2041 CATT CAGCTC TGGTTCCCAT CTATCAAGCC TAGTTACATG ATCACCCTAG TTGTGCAAAA
2101 AAGCAGTCAA CTCCTTTGGT CCTCCAATGG TTGTCAAAAG TAAGTTGGCA GCAGTGTAT
2161 CACTCATGGT TATGGCAGCA CTGCATAAAT CTCTTACTGT CATGCCATCT GTAAGATGCT
2221 TTTCTGTGAC TGGACTGTAC TCAACCAAGT CATTCTGAGA ATAGTGTATT CTCTACCCA
2281 GTTGCTCTTG CCCAGCATCA ATTCTGGATA ATACTGCACC ACATAGCAGA ACTTTAAAGG
2341 TGCTCATCAT TGGAAATCTT TCTTCTGGTC TAAAACCTCTC AAGGATCTTA CCAGAGTTGA
2401 GATCCAGTTC AATGTAACCC ACTCTTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA
2461 CCAGGGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAAAGC AGCAAAAAGG GGAATAAGGG
2521 CAAC TCTGAA ATGTTGAATA CTCATAGTAC TACTCTTCCT TTTTCAATAT TATGAAGCA
2581 TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC
2641 AAATAGGGGT ATGCATTCAG CTCACATTTT CCTGAAAAGT GCCACCTGAA ATGACTGAT
2701 AGGGAGTTCT CCCAATCCCC TATGGTGCAC TCTCAGTACA ATCTGCTCTG ATGCCTCATA
2761 GTTAAGCCAG TATCTGCTCC CTGCTTGTGT GTTGGAGGTC ACTGAGTAGT GGGCTAGCAA
2821 AATTTAAGCT ACAACAAGGC AAGGCTTGAC CTACAATTGC ATGAAGAATC TGCTTAGGGT
2881 TAGGCCTTTT GCACTGCTTG GAGATGTACT GGCCAGATAT ACTA

```

Abkürzungsverzeichnis

A. v.	<i>Aequorea victoria</i>
AG	Arbeitsgruppe
AMD	ActinomycinD
Amp	Ampicillinresistenzkassette
amp	Ampicillinresistenzgen
AP	Alkalische Phosphatase
AP1	Aktivierendes Protein1
AS	Aminosäure
ATP	Adenosin-5'-triphosphat
BGH	<i>bovine growth hormone</i>
bla	β-Laktamat
bp	Basenpaar
BSA	<i>bovine serum albumine</i>
°C	Grad Celsius
^{5m} C	5-Methylcytosin
ca.	circa
CAI	<i>codon adaption index</i>
CF	<i>Cleavage factor</i>
CFB1(=CGBP)	<i>CXXC finger protein1</i>
CGBP	CpG-Bindeprotein
CHO	<i>Chinese hamster ovaries</i>
CMV-Promotor	<i>hCMV immediate early promotor</i>
CpG	Cytidin-phosphat-Guanosin-Dinukleotid
CPSF	<i>Cleavage and polyadenylation specific factor</i>
CREB	<i>Ca²⁺/cAMP-response element binding protein</i>
CStf	<i>Cleavage stimulatory factor</i>
CXXC	Cystein-X-X-Cystein
D	A oder G oder T (nicht C)
Da	Dalton
<i>dam⁻</i>	Methylierungsdefekt in <i>E. coli</i> (<i>dam</i> -Methylase)
DC	Dendritische Zelle
<i>dcm⁻</i>	Methylierungsdefekt in <i>E. coli</i> (<i>dcm</i> -Methylase)
DNA	Desoxyribonukleinsäure (DNS)
DNMT	DNA-Methyltransferase
dNTP	Desoxyribonukleotidtriphosphat
ds	doppelsträngig
DTT	Dithiothreitol
ECL	<i>enzyme chemoluminescence</i>
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraacetat
ELISA	<i>enzyme-linked immunosorbent assay</i>
<i>et al.</i>	und andere
EZD	extrazelluläre Domäne
FACS	<i>fluorescence activated cell sorting</i>
FKS	fötale Kälberserum
FM	Fluoreszenzmikroskop
FRT	<i>Flp recombination target</i>
F-PCR	Fusions-PCR
g	Erdbeschleunigung (9,81 ms ⁻²) oder Gramm
GFP	<i>green fluorescent protein</i>
GmbH	Gesellschaft mit beschränkter Haftung

GOI	<i>Gene of interest</i>
GS	Glutamat-Synthetase
h	Stunden (<i>hours</i>)
H	A oder C oder T (nicht G)
hCMV	humanes Cytomegalie-Virus
HDAC	Histondeacetylase
h_GFP	auf humane Kodonwahl angepasste Version von GFP
HIV-1	Humanes Immundefizienz-Virus Typ-1
H ₂ O _{dest}	destilliertes Wasser
h_p24	auf humane Kodonwahl angepasste Version von p24
HRP	Meerrettich-Peroxidase (<i>horseradish peroxidase</i>)
hygr	Hygromycinresistenzgen
Hygr	Hygromycinresistenzkassette
IFN	Interferon
IκB	Inhibitor-κB
IL	Interleukin
IL-1R	Interleukin-1-Rezeptor
IRAK	IL-1R-assoziierte Kinase
k	kilo (10 ³)
k	Zerfallskonstante
kan	Kanamycinresistenzgen
l	Liter
λ	Wellenlänge
<i>lacZ</i>	β-Galactosidasegen
LB	Luria Bertani
LC	<i>Lightcycler</i>
LRR	<i>leucin rich repeats</i>
M	molar (Mol pro Liter)
μ	mikro (10 ⁻⁶)
m	Meter oder milli (10 ⁻³)
MAPK	<i>mitogen activated protein kinase</i>
MAR	<i>matrix attached region</i>
MBD	mCpG-bindende-Domäne
MBD1	<i>methyl-CpG-binding domain protein1</i>
mCpG	methyliertes CpG
MCS	Multipler Klonierungsbereich
MeCP2	<i>methyl-CpG-binding protein2</i>
min	Minuten
MFI	Mittlere Fluoreszenzintensität
MyD88	myeloider Differenzierungsfaktor 88
n	nano (10 ⁻⁹)
NEB	New England Biolabs
NF-κB	Nukleärer Faktor-κB
NK	Natürliche Killerzelle
Nr	Auftragsnummer
NS	Nukleinsäure
OD	optische Dichte
ODN	Oligodesoxynukleotid
ORF	<i>open reading frame</i>
p	pico (10 ⁻¹²)
P(A)	Poly(A)
PAB	Poly(A)-Bindeprotein
PAGE	Polyacrylamidgelelektrophorese

PAP	Poly(A)-Polymerase
PBS	Phosphat-gepufferte Kochsalzlösung (<i>phosphate buffered saline</i>)
pc5	pcDNA5/FRT
PCR	Polymerase-Kettenreaktion (<i>polymerase chain reaction</i>)
PCR-Script	Klonierungsvektor pPCR-Script
PFA	Paraformaldehyd
<i>Pfu</i>	DNA-abhängige DNA-Polymerase aus <i>Pyrococcus furiosus</i>
pH	Konzentration der H ⁺ -Ionen in einer Lösung
PHD	pflanzliche Homeodomäne
POZ	Poxvirus-Zinkfinger-Domäne
RE	Restriktionsenzym
RNA	Ribonukleinsäure
RT	Reverse Transkriptase
SAH	S-Adenosyl-homocystin
SAM	S-Adenosyl-methionin
SDS	Natriumdodecylsulfat (<i>sodium dodecyl sulfate</i>)
sec	Sekunden
Sp1	<i>Specificity protein 1</i>
ss	einzelsträngig
SV40	Simian Virus 40
SV40-Promotor	SV40 <i>early promoter</i>
SV40 P(A)-Signal	SV40 <i>late polyadenylation signal</i>
<i>Taq</i>	DNA-abhängige DNA-Polymerase aus <i>Thermus aquaticus</i>
TB	<i>Terrific broth</i>
TBS	Tris-gepufferte Kochsalzlösung
TCA	Trichloracetat
TF	Transkriptionsfaktor
TIR	Toll/IL-1 homologe Region
TLR	Toll-like Rezeptor
TM	Trademark
TMD	Transmembrandomäne
TNF	Tumor-Nekrose-Faktor
TRAF	TNF-Rezeptor-assoziiertes Faktor
TRD	<i>Transcriptional repressor domain</i>
Tris	Tris(hydroxymethyl)aminomethan
TTBS	Tris-gepufferte Kochsalzlösung mit Tween-20
TpA	Thymidin-phosphat-Adenosin-Dinukleotid
U	<i>unit</i>
USF	<i>upstream stimulatory factor</i>
UV	ultraviolett
v/v	Volumeneinheit pro Volumeneinheit (<i>volume per volume</i>)
W	A oder T
WH-Element	Wiederholungselement
wt	Wildtyp
w/v	Gewichtseinheit pro Volumeneinheit (<i>weight per volume</i>)
w/w	Gewichtseinheit pro Gewicht (<i>weight per weight</i>)
X	Platzhalter
X-gal	5-Brom-4-Chlor-3-Indolyl-β-D-Galactopyranosid
YY1	Yin-Yang-1
z. B.	zum Beispiel
ZF	Zinkfinger

Kurzschreibweisen der Aminosäuren:

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

Kurzschreibweise der Basen:

Adenin	A
Cytosin	C
Guanin	G
Thymin	T
Uracil	U
Purin	R
Pyrimidin	Y

Danksagung

Zunächst bedanke ich mich ganz herzlich bei Herrn Prof. Dr. Ralf Wagner, der mir diese Promotion ermöglicht hat. Ohne seine kompetente Unterstützung wäre diese Arbeit nicht zustande gekommen. Insbesondere möchte ich mich für die angenehme, freundschaftliche Atmosphäre, die allgegenwärtige Diskussionsbereitschaft und das Verständnis für private Belange bedanken.

Darüber hinaus gilt mein Dank Herrn Prof. Dr. Hans Wolf für die Bereitstellung eines hervorragend ausgestatteten Arbeitsplatzes am Institut für Medizinische Mikrobiologie und Hygiene am Universitätsklinikum Regensburg.

Herrn Prof. Dr. Dr. Hans Robert Kalbitzer danke ich für die bereitwillige und unkomplizierte Übernahme des Erstgutachtens.

Ein ganz liebes Dankeschön geht an die Mitglieder der Diagnostik-Mannschaft von der AG Reischl, die mir in Sachen *realtime*-PCR mit Rat und Tat zur Seite standen und die Benutzung des *Lightcyclers* zu allen möglichen und unmöglichen Zeiten gewährleisteten. Vielen Dank !

Bei Dr. Frank Notka bedanke ich mich ganz besonders für die zuverlässige und kompetente Betreuung dieser Arbeit. Vor allem die konstruktiven Vorschläge zur Konzeption und Gestaltung der schriftlichen Arbeit waren mir eine sehr große Hilfe. Lieber Frank, vielen Dank auch fürs Korrekturlesen !

Ein ganz besonderer Dank gilt meinem „Korrekturleseteam“ Gabi Fiedler und Thomas Schödl, die sich unermüdlich durch meine schriftlichen Entwürfe gekämpft haben. Mein ewiger Dank ist Euch sicher !

Bei meinen Kollegen der AG Wagner möchte ich mich für die Hilfsbereitschaft, Kollegialität und schöne Arbeitsatmosphäre bedanken. Insbesondere die Mittagspausen mit Euch haben auch den stärksten Motivationstiefs den Garaus gemacht. Vor allem Asli Bauer danke ich ganz herzlich für die fachlich und menschlich sehr gute Zusammenarbeit. Vielen Dank fürs ausdauernde Lesen und die aufbauenden Worte !! Außerdem möchte ich mich bei Denijal Kosovac für die tatkräftige Unterstützung im Rahmen seiner Diplomarbeit und die Generierung der p24-Daten bedanken.

Der größte Dank gilt meinen Eltern, die mir mein Studium ermöglicht haben und immer an mich geglaubt haben. Ein liebes Dankeschön geht auch an meine Tante, die mich stets unterstützt hat. Meiner kleinen Schwester rechne ich hoch an, dass sie im fernen Argentinien diese Arbeit durchgelesen hat. Muchas gracias !

Last but not least möchte ich mich bei Sascha bedanken, der mich in schweren Zeiten stets aufgemuntert, motiviert und bekocht hat. So long and thanks for all the fish !