

Funktionelle Untersuchungen zur sequenz-spezifischen Rekombination durch die Integrase des Bakteriophagen Lambda in eukaryotischen Zellen

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

an der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Teresa Corona

aus Mönchengladbach

Köln, September 2003

Berichtersteller: Prof. Dr. Peter Dröge
Prof. Dr. Helmut W.Klein

Tag der mündlichen Prüfung: 10. November 2003

Für meine Eltern

Inhaltsverzeichnis

Abkürzungsverzeichnis

1	Einleitung	1
1.1	Sequenz-spezifische Rekombinationssysteme und ihr Einsatz in der Biotechnologie.....	2
1.1.1	Das Cre/loxP- und das Flp/FRT-System	4
1.1.2	Der Einsatz sequenz-spezifischer Rekombinationssysteme am Beispiel von Cre und Flp: Möglichkeiten und Einschränkungen.....	4
1.2	Die sequenz-spezifische Rekombination des Phagen <i>Lambda</i>	7
1.2.1	Die integrative λ -Rekombination	8
1.2.2	Die exzitive λ -Rekombination	10
1.2.3	Die λ -Integrase.....	11
1.2.4	Die Protein-Kofaktoren der λ -Integrase: IHF, XIS und FIS.....	15
1.3	Erzeugung transgener Mäuse durch Pronukleus-Injektion	17
1.4	Zielsetzung der Arbeit.....	19
2	Material	20
2.1	Bakterienstämme	20
2.2	Chemikalien	20
2.3	Enzyme	20
2.4	Synthetische Oligonucleotide.....	20
2.5	Medien für bakterielle Kulturen.....	22
2.6	Plasmide	23
2.7	Medien und Puffer für die eukaryotische Zellkultur	24
2.8	Eukaryotische Zelllinien	25
3	Methoden.....	26
3.1	Molekularbiologische Methoden.....	26
3.1.1	Agarosegelelektrophorese	26
3.1.2	DNA-Extraktion aus Agarosegelen.....	26
3.1.3	Plasmid-Isolierung aus Bakterien.....	26
3.1.4	DNA-Isolierung aus eukaryotischen Zellen und Schwanzspitzen der Maus.....	27
3.1.5	Polymerasekettenreaktion (PCR).....	27
3.1.6	RNA-Isolierung und RT-PCR Analyse der rIHF mRNA-Ausprägung in Mausorganen und HeLa-Zellen.....	28
3.1.7	DNA-Sequenzierung	29
3.1.8	<i>Southern Blot</i> Analyse.....	29

3.1.9	Konstruktion und Beschreibung der eukaryotischen Expressionsvektoren	31
3.1.10	Konstruktion der Substratvektoren	33
3.1.11	Konstruktion der Integrations-Vektoren (<i>Targeting</i> genomischer Loci)	35
3.1.12	Konstruktion der Vektoren zur Generierung transgener Zelllinien und Mäuse	35
3.2	Biochemische Methoden	36
3.2.1	SDS-Polyacrylamid-Gelelektrophorese	36
3.2.2	<i>Western Blot</i> Analyse	37
3.3	Zellbiologische Methoden	38
3.3.1	Kultur von Säugetierzellen	38
3.3.2	Konstruktion stabiler HeLa-Zelllinien	38
3.3.3	Durchflusszytometrie	38
3.3.4	Immunofluoreszenz	39
3.4	Rekombinationsanalysen	40
3.4.1	Episomale Rekombinationsanalysen in eukaryotischen Zellen	40
3.4.2	Chromosomale Rekombinationsanalysen	40
3.4.2.1	Genomische Inversion in H8B-Zelllinien	40
3.4.2.2	Genomisches <i>Targeting attB</i> -stabiler HeLa-Zelllinien	41
3.4.2.3	Genomisches <i>Targeting</i> von <i>attH2</i> in HeLa-Zellen	42
4	Ergebnisse	43
4.1	Intermolekulare integrative Rekombination von episomalen Substraten durch die Lambda-Integrase und ihre Mutanten in eukaryotischen Zellen	43
4.1.1	Das Reportersystem für die integrative intermolekulare Rekombination	43
4.1.2	Kinetik der integrativen intermolekularen Rekombination	44
4.1.3	Einfluss der Arm-Bindung auf die integrative Rekombination	46
4.1.3.1	Vergleich der Rekombinationseffizienzen zwischen den <i>att</i> -Regionen der Integrativen Rekombination	46
4.1.3.2	Expression und Aktivität der Integrase-C-Domäne	48
4.1.4	Der Einsatz des rekombinanten IHFs (rIHF) in eukaryotischen Zellen	50
4.1.4.1	Aufbau des rekombinanten <i>Integration Host Factors</i> (rIHF)	51
4.1.4.2	Stabile Expression von rIHF in HeLa-Zelllinien	53
4.1.4.3	Stabile rIHF-Zelllinie gegenüber parentaler HeLa-Zelllinie: Vergleich der intramolekularen episomalen Rekombination der wild-typ Integrase und ihrer Mutanten .	55
4.2	Intermolekulare Rekombination zwischen genomischen und episomalen Substraten durch die Lambda-Integrase und ihre Mutanten in eukaryotischen Zellen	57
4.2.1	<i>Targeting</i> Strategie für Zelllinien mit genomisch stabil integrierter <i>attB</i> -Region	57
4.2.1.1	Generierung <i>attB</i> -stabiler HeLa-Zelllinien	59
4.2.1.2	<i>Targeting attB</i> -stabiler HeLa-Zelllinien mit pCMVSS <i>attP</i> _{mut}	61
4.2.2	Einfluss von rIHF auf die genomische Rekombination	66
4.2.2.1	Generierung von HeLa-Zelllinien mit stabil integriertem rIHF und <i>attB</i>	66
4.2.2.2	<i>Targeting</i> einer rIHF/ <i>attB</i> -stabilen HeLa Zelllinie mit pCMVSS <i>attP</i> _{mut}	68
4.2.2.3	Genomische Inversion in HeLa-Reporterzelllinie in An- und Abwesenheit von rIHF	68

4.2.3	<i>Targeting</i> Strategie für eine <i>attB</i> -ähnliche Sequenz (<i>attH2</i>) im menschlichen Genom....	70
4.2.3.1	Episomale Untersuchung zur Funktionalität der <i>attP</i> *ns/ <i>attH2</i> Rekombination.....	72
4.2.3.2	<i>Targeting</i> der genomischen <i>attH2</i> -Region mit und ohne rIHF.....	73
4.3	Generierung einer rIHF transgenen Maus	75
4.3.1	Analyse der rIHF-transgenen Mäuse	77
5	Diskussion.....	81
5.1	Intemolekulare Rekombination episomaler Substrate	81
5.1.1	<i>Lambda</i> Integrase-Mutanten katalysieren in HeLa-Zellen die intermolekulare Rekombination episomaler Substrate mit vergleichbarer Effizienz wie intramolekulare Reaktionen	81
5.1.2	Einfluss der Arm-Bindung und Rückschlüsse für die Synapsenbildung	83
5.1.3	Der rekombinante <i>integration host factor</i> (rIHF) steigert die episomale Rekombinationsaktivität des Integrase wild-typ-Proteins in HeLa-Zellen	87
5.2	Rekombination mit genomischen Substraten durch <i>Lambda</i>-Integrase und ihre Mutanten in eukaryotischen Zellen.....	90
5.2.1	In HeLa-Zellen ist die sequenz-spezifische Integration von Fremd-DNA in genomische Loci nur bedingt nachzuweisen	90
5.2.2	Eine Steigerung der genomischen Rekombinationsaktivität durch rIHF kann unter den gewählten Bedingungen in HeLa-Zellen nicht detektiert werden	94
5.3	Transgene rIHF-Mäuse.....	97
6	Literaturverzeichnis.....	99
7	Anhang	112
8	Zusammenfassung	121
9	Summary	122

Abkürzungsverzeichnis

Maße und Konzentrationen werden in den gültigen SI-Einheiten angegeben. Enzyme sind wie in der Fachliteratur allgemein üblich abgekürzt und des weiteren werden die folgende Abkürzungen verwendet:

A	Adenin
Abb.	Abbildung
Abschn.	Abschnitt
AS	Aminosäure
<i>att</i>	<i>attachment site</i> ; Rekombinationsregionen
<i>attB</i>	<i>bacterial attachment site</i>
<i>attP</i>	<i>phage attachment site</i>
<i>attL</i>	<i>attachment site left</i>
<i>attR</i>	<i>attachment site right</i>
<i>attP_{mut}</i>	<i>mutant phage attachment site</i> ohne ATGs in P-Arm
<i>attP[*]_{ns}</i>	<i>mutant phage attachment site</i> ohne Stoppcodons in P'-Arm
<i>attH2</i>	<i>human attachment site2</i>
bp	DNA-Basenpaar(e)
bPA	<i>bovine growth hormone</i> (Rinder-Wachstumshormon)
BSA	<i>Bovine Serum Albumine</i> ; Rinderserum-Albumin
bzw.	beziehungsweise
C	Cytosin
Caggs	CMV <i>immediate ealy enhancer</i> fusioniert an den β -Aktin-Promoter aus Huhn
CD	C-Domäne
cDNA	<i>complementary DNA</i>
CMV	Cytomegalovirus (Promoter)
DAPI	4',6-Diamidino-2-Phenylindol
DMSO	Dimethylsulfoxid
DNA	<i>desoxyribonucleic acid</i> ; Desoxyribonucleinsäure
dNTP	Desoxynucleotide
DTT	Dithiothreitol
(d)YT	(<i>double</i>) yeast/tryptone; (doppelt) Hefe/Trypton
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
eGFP, GFP	<i>enhanced green fluorescence protein</i> ; verstärkt grün fluoreszierendes Protein
ES	embryonale Stammzellen
FACS	fluorescence activated cell sorter
FCS	<i>fetal calf serum</i> ; fötales Kälberserum
FITC	Fluorescein-Isothiocyanat
FIS	<i>factor for inversion stimulation</i>
FRT	Rekombinationssequenz der Flp-Rekombinase
G	Guanin
GFP	<i>green fluorescence protein</i> ; grün fluoreszierendes Protein
himA	Gen der Untereinheit α von IHF
himD	Gen der Untereinheit β von IHF

Hyg	Hygromycin Resistenzgen
HPRT	Hypoxanthin-Phosphorybosyl-Transferase-Gen
HSV-TK	Herpes-Simplex-Virus Thimidinkinase
IHF	<i>Integration Host Factor</i>
Int	Integrase
IRES	<i>Internal ribosomal entry site</i>
IPTG	Isopropylthiogalactosid
KDa	Kilo-Dalton
Kb	Kilobasenpaare
MCT	<i>monocarboxylate transporter</i> , Monokarboxylat-Transporter
<i>loxP</i>	Rekombinationssequenz der Cre-Rekombinase
LSM	<i>laser scan microscope</i>
mRNA	<i>messenger RNA</i> , Boten-RNA
NA	Nucleotidaustausche
Neo	Neomycin-Resistenz-Gen
OD	optische Dichte
ORF	<i>open reading frame</i> ; offenes Leseraster
PA	Polyadenylierungs-Sequenz
PAGE	Polyacrylamid Gel Elektrophorese
PBS	<i>Phosphat Buffered Saline</i> ; Phosphat-gepufferte Salzlösung
PCR	<i>polymerase chain reaction</i> ; Polymeraseketten-reaktion
PGK	Phosphoglycerat-Kinase (Promoter)
Puro	Puromycin Resistenzgen
rIHF	<i>one-chain Integration Host Factor</i>
RNA	<i>ribonucleic acid</i> , Ribonukleinsäure
RT-PCR	<i>reverse-transcriptase PCR</i>
SA	<i>splice acceptor</i> ; 3'-Spleißstelle
SDS	Natriumdodecylsulfat
SS	Spleißsubstrat
T	Thymin
TBE	Tris-Borat-EDTA
TE	Tris-EDTA
ÜNK	Übernachtkultur
UV-Licht	Ultraviolett-Licht
v/v	<i>volume/volume</i> ; Volumen/Volumen
wt	wild-typ
w/v	<i>weight/volume</i> ; Gewicht/Volumen
YT	<i>yeast/tryptone</i> ; Hefe-Trypton
z.B.	zum Beispiel

1 Einleitung

Die moderne biotechnologische Forschung verwendet gezielte Mutationen und sequenz-spezifische Rekombination zur Veränderung prokaryotischer und eukaryotischer Genome.

In der Natur gewährleisten Mutationen und Rekombination die genetische Vielfalt einer Spezies. Mutationen verändern einzelne Gene oder kleine Gruppen von Genen in Individuen. Rekombination führt dagegen im Rahmen der Vermehrung zur Neuverteilung genomischer Information unter mehreren Individuen. In der klassischen Biologie ist die Rekombination das Resultat des *crossing overs* von gepaarten Schwesterchromosomen während der Meiose in Eukaryoten. Genau genommen handelt es sich jedoch bei jedem Prozess, der zur Neuverknüpfung von zwei unterschiedlichen DNA-Molekülen führt, um Rekombination. Die genetische Rekombination kann in drei Kategorien eingeteilt werden: Homologe, illegitime und sequenz-spezifische Rekombination. Ausschlaggebend dafür sind Unterschiede in der Sequenz der DNA-Substrate, in den beteiligten Proteinen und im Mechanismus.

Die **homologe Rekombination** vermittelt den Austausch von DNA-Segmenten, die über große Bereiche in ihrer Sequenz übereinstimmen (Smith, 1988; Thaler & Stahl 1988; Übersichtsartikel: Bollag et al., 1989). Der Austausch kann dabei an jeder Position der homologen Region stattfinden. Proteine die diese Art von Rekombination vermitteln sind z.B. rec- und ruv-Proteine in *E. coli* (Smith, 1988; Kowalczykowski et al., 1994) oder Rad-Proteine in Eukaryoten (Baumann & West, 1998).

Der Umlagerung von transponierbaren Elementen (Insertionssequenzen, Transposons) liegt der Mechanismus der **illegitimen Rekombination** zugrunde. Eine Homologie der Rekombinationsregionen ist dabei nicht erforderlich. Die Transposasen werden von den transponierbaren Elementen selbst kodiert. Sie katalysieren diese Reaktion, bei der es auch zur Neusynthese von DNA kommt. Die Neusynthese ist je nach Transposon unterschiedlich umfangreich und kann das gesamte mobile Segment umfassen oder sich nur auf einige Basenpaare beschränken (Grindley & Reed, 1985; Derbyshire & Grindley, 1987; Mizuuchi, 1992).

Die **sequenz-spezifische** Rekombination findet an definierten DNA-Abschnitten

statt, in denen nur ein kurzes Stück homolog ist. Sie ist ein konservativer Prozess, d.h. sie erfolgt ohne DNA-Verlust oder -Neusynthese durch präzise Einführung von Strangbrüchen und Neuverknüpfungen (Craig, 1988; Stark et al., 1992; Nash, 1996).

Im Folgenden wird auf einige sequenz-spezifische Rekombinationssysteme und ihre Anwendungen in der Biotechnologie eingegangen, besondere Aufmerksamkeit gilt dabei dem regulierten System des Bakteriophagen *Lambda* und den daran beteiligten Proteinen.

1.1 Sequenz-spezifische Rekombinationssysteme und ihr Einsatz in der Biotechnologie

Um das Genom von Säuger- oder Pflanzen-Zellen gezielt zu verändern, kann man sich der homologen Rekombination zwischen eingeführter DNA und entsprechender chromosomaler Sequenzen bedienen. Diese Methode, die zum Beispiel beim so genannten *Gene Targeting* in embryonalen Stammzellen von Mäusen angewendet wird, ist jedoch ineffizient und bedarf zusätzlicher Selektionsstrategien (Bollag et al., 1989). Ebenso erfordert sie eine Überprüfung, um zufällige und illegitime Integrationsereignisse herauszufiltern, die häufig die homologe Rekombination überlagern (Marrow und Kucherlapati, 1993; Vasquez et al., 2001). In somatischen Zellen ist das Verhältnis der homologen zur nicht-homologen Rekombination noch ungünstiger als in embryonalen Stammzellen. Für die gezielte genomische Manipulation in somatischen Zellen ist das *Gene Targeting* daher eher ungeeignet.

Alternativen zur homologen Rekombination bieten sequenz-spezifische Rekombinationssysteme (Kilby et al., 1993). Aus Bakteriophagen und Hefen wurden in den letzten Jahren eine Vielzahl sequenz-spezifischer Rekombinasen untersucht, weiterentwickelt und auf eukaryotische Zellen übertragen. Sie stehen damit als Werkzeuge zur gezielten DNA-Manipulation *in vitro* und *in vivo* zur Verfügung (Metzger, D. & Feil, R., 1999; Yu & Bradley, 2001; Lynzik et al., 2003; Olivares et al., 2001; Thyagarajan et al., 2001). Diese Enzyme schneiden DNA an einer spezifischen Stelle auf und verknüpfen sie mit einer zweiten spezifisch geschnittenen DNA-Stelle ohne Verlust oder Neusynthese von DNA. Diese einfache und elegante Reaktion ist sehr effizient und führt zur präzisen Rekombination zwischen zwei geeigneten Partnersequenzen. Von Vorteil bei der Anwendung von sequenz-spezifischen Rekombinationssystemen ist auch ihre Regulierbarkeit. Beispielsweise

können durch gezielte Ausprägung sequenz-spezifischer Rekombinasen Gene an- oder abgeschaltet und deren Funktion in bestimmten Zelltypen oder in bestimmten Entwicklungsstadien untersucht werden (Kilby *et al.*, 1993; Gu *et al.*, 1994). Die

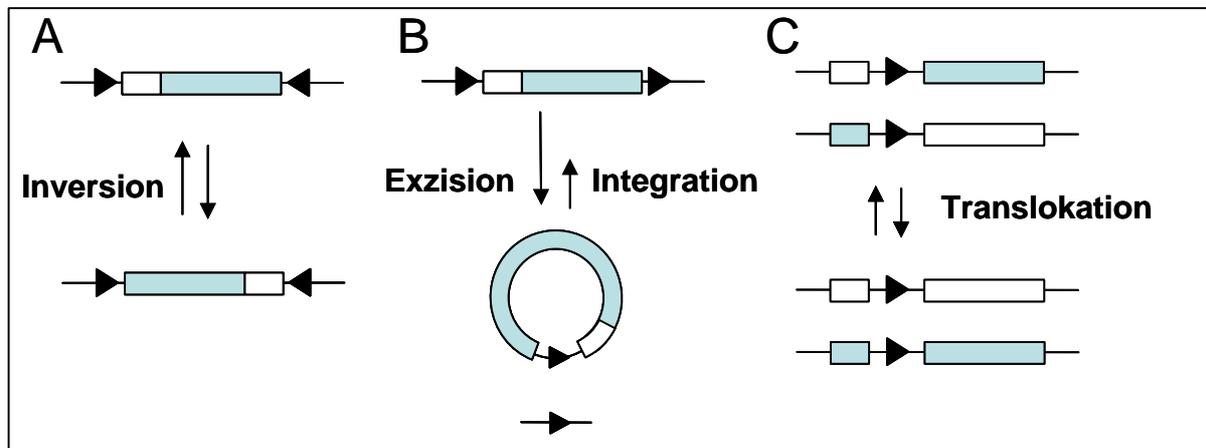


Abbildung 1-1: Schematische Darstellung von sequenz-spezifischen Rekombinationsmöglichkeiten zwischen unterschiedlich orientierten Rekombinationsregionen. Die zur Integrase-Familie gehörenden Rekombinasen Cre, FLP oder Int führen die sequenz-spezifische Rekombination zwischen zwei *loxP*-, *FRT*- oder *att*-Regionen, in dieser Reihenfolge, durch (schwarze Dreiecke). Die intramolekulare Rekombination zwischen entgegengesetzt orientierten Rekombinationsregionen (**A**) führt zur Inversion des dazwischenliegenden DNA-Fragmentes, während die Rekombination zwischen gleich orientierten Sequenzen (**B**) zur Exzision des flankierten DNA-Abschnitts führt (zirkuläres Produkt). Die intermolekulare Rückreaktion, die Integration, ist nicht so effizient und in der Abbildung durch einen kleineren Pfeil symbolisiert. Die Translokation, ebenfalls eine intermolekulare Rekombinationsreaktion, ist in (**C**) dargestellt. Zur Kennzeichnung bestimmter DNA-Abschnitte sind diese grau unterlegt. (nach Metzger & Feil, 1999).

Positionierung der Zielsequenzen bestimmt dabei, ob dies durch Insertion, Deletion, Inversion oder Versetzen von genetischen Elementen geschieht. Beispielsweise führt die intramolekulare Rekombination zwischen entgegengesetzt orientierten Rekombinationsregionen zur Inversion des dazwischenliegenden DNA-Fragments. Die Rekombination zwischen gleich orientierten Sequenzen führt dagegen zur Exzision des flankierten DNA-Abschnitts (Abb. 1-1; Metzger & Feil, 1999). Die Komplexität der einzelnen Rekombinationssysteme kann erheblich variieren, was ihre Abhängigkeit von Kofaktoren und die Größe ihrer Rekombinations-Regionen betrifft.

In den letzten Jahren wurden vor allem Rekombinasen, die nicht von Protein-Kofaktoren abhängen, zur Manipulation eukaryotischer Genome eingesetzt (Metzger, D. & Feil, R., 1999; Yu & Bradley, 2001; Lynzik *et al.*, 2003; Olivares *et al.*, 2001; Thyagarajan *et al.*, 2001). Im Vordergrund standen und stehen bis heute die Cre-Rekombinase des Bakteriophagen P1 (Sauer, 1993) und die *FLP*-Rekombinase aus Hefe (O’Gorman *et al.*, 1991).

1.1.1 Das Cre/*loxP*- und das Flp/FRT-System

Das sequenz-spezifische **Cre/*loxP***-Rekombinationssystem des Phagen P1, besteht aus drei Komponenten: der Cre-Rekombinase (38 KDa) und zwei identischen, 34 Basenpaare langen *loxP*-Sequenzen, an denen die Rekombination stattfindet. Die *loxP*-Sequenz wird von zwei 13 bp langen invertierten Sequenzen gebildet, die durch eine 8 bp lange, asymmetrische Sequenz getrennt werden (Sternberg & Hamilton, 1981, Hamilton & Abremski, 1984). Diese Asymmetrie verleiht der *loxP*-Sequenz eine Richtung und, wie in Abb. 1-1 dargestellt, bestimmt die Orientierung zweier *loxP*-Sequenzen das Resultat der Rekombinationsreaktion: Flankieren zwei *loxP*-Sequenzen ein DNA-Fragment, so kann dieses in Anwesenheit der Rekombinase ausgeschnitten (gleiche Orientierung) oder invertiert werden (entgegengesetzte Orientierung) (Hoess et al., 1986). Gleiches gilt auch für das **Flp/FRT-System**. Hier wird die Rekombination von FRT-Sequenzen durch die Flp-Rekombinase vermittelt. Die Flp-Rekombinase ist im Gegensatz zu Cre eukaryotischen Ursprungs und wird vom 2- μ m-Plasmid aus *Saccharomyces cerevisiae* kodiert. Ihre *Target*-sequenz wird FRT genannt und besteht in ihrer minimalen Form wie *loxP*, aus zwei 13 bp langen invertierten Sequenzen, die durch einen 8 bp langen asymmetrischen Bereich getrennt werden (Jayram, 1985; Gronostajski & Sadowski, 1985; Sadowski, 1995). Beide Systeme benötigen keine Protein-Kofaktoren und sind unabhängig von der Topologie ihrer Substrate (Abremski et al., 1983; Sadowski, 1995). Neben Prokaryoten und Hefen haben sich Cre bzw. Flp auch in Zellen höherer Eukaryoten als funktionsfähig erwiesen (Sauer, 1994; Sauer & Henderson, 1989; 1990; O’Gorman et al., 1991; Golic & Lindquist, 1989, Lyznik et al., 2003). Sie gehören mittlerweile zu den wichtigsten Werkzeugen zur gezielten Manipulation eukaryotischer Genome.

1.1.2 Der Einsatz sequenz-spezifischer Rekombinationssysteme am Beispiel von Cre und Flp: Möglichkeiten und Einschränkungen

Cre und Flp finden ihre bedeutendste Anwendung bei der Generierung von transgenen Zelllinien oder transgenen Tieren wie z.B. Mäusen. Durch ihren Einsatz können komplette Gene oder Genabschnitte, die vorher durch entsprechende *target*-Sequenzen flankiert wurden, deletiert oder invertiert werden (Dymecki, 1996; Werdien et al., 2001; Araki et al., 1995; Lasko et al., 1992; Orban et al., 1992; Farley

et al., 2000). Die Deletion oder Inversion führt zur Ausschaltung (*knock-out*) des anvisierten Gens und kann im Tiermodell bei der Aufklärung des molekularen Mechanismus verschiedener Erkrankungen helfen. Geschieht dies früh in der Entwicklung, wird diese Methode als konstitutives *gene targeting* bezeichnet. Sie generiert modifizierte Allele in allen Zellen des transgenen Tieres vom Zeitpunkt der Befruchtung an und kann während der Entwicklung und im ausgewachsenen Stadium zur Untersuchung der Genfunktion dienen. Das konstitutive Ausschalten eines Gens kann jedoch in einigen Fällen letal sein, z.B. wenn das zu untersuchende Gen für die embryonale Entwicklung wichtig ist. Außerdem kann beim universellen Ausschalten eines Gens zwischen direkten Effekten, die durch das Abschalten in einem bestimmten Gewebe entstehen, und indirekten Effekten, die aus der Ausschaltung in sämtlichen Geweben hervorgehen können, nicht differenziert werden. Um diese Nachteile zu überwinden, wurde mit Hilfe der sequenzspezifischen Rekombinationssysteme das so genannte konditionale *gene targeting* entwickelt. Hier ist die Modifikation eines Gens auf einen bestimmten Zelltyp (gewebespezifisch) oder auf einen bestimmten Zeitpunkt (entwicklungsspezifisch) oder beides beschränkt (Orban et al., 1992; Kühn et al., 1995; Sauer, 1998; Metzger & Feil, 1999; Lasko et al., 1992; Gu et al., 1994; Porter, 1998; Lewandoski, 2001; Wilson & Kola, 2001; Le & Sauer 2000; Takeuchi et al., 2002). Neben Deletion und Inversion können Cre und Flp die Integration eines Vektors mit einer entsprechenden target-Sequenz in eine zuvor in das Genom integrierte *loxP* bzw. FRT-Sequenz katalysieren (Baubonis und Sauer, 1993; O’Gorman et al., 1991; Golic & Lindquist, 1989). Dazu muss die Erkennungssequenz zuvor durch andere Methoden (z.B. homologe oder illegitime Rekombination) als stabile Sequenz in das Genom eingebracht werden. Eine wesentliche Vereinfachung würde die Integration in eine natürlich vorkommende Sequenz des Genoms darstellen, in welche therapeutische Gene ohne vorherige zeitaufwendige Modifikation inseriert werden könnten. Beispielsweise konnten *loxP*-analoge Sequenzen schon im Maus Genom und im menschlichen Genom identifiziert werden und erwiesen sich außerdem in episomalen Tests als rekombinationsfähige Substrate (Thyagarajan et al., 2000). Die Cre-Rekombinase scheint eine hohe Toleranz gegenüber Abweichungen von der Konsensus-*loxP*-Sequenz zu besitzen, da die Sequenzen teilweise stark von der wild-typ-Sequenz abweichen. Die Integration von Fremd-DNA in solche genomische *loxP*-analoge Sequenzen durch Cre wurde bisher nur in Hefe gezeigt (Sauer, 1996).

Trotz seiner weit verbreiteten Anwendung unterliegt der Einsatz von Cre und Flp gewissen Einschränkungen. Bei der Rekombination ihrer *target*-Sequenzen entstehen Produkte, die mit den Ausgangssequenzen identisch sind. Sie können sofort wieder als Substrate für die Rekombination dienen, zumal keine Faktoren involviert sind, die die Reaktionsrichtung beeinflussen. Eine in eine *loxP*- oder FRT-Region integrierte Sequenz kann also in Anwesenheit der Rekombinase direkt wieder ausgeschnitten werden. Zudem konnte erst kürzlich gezeigt werden, dass es bei längerer Anwesenheit von Cre zu Chromosomen-Aberrationen in Spermatischen und möglicherweise anderen Zellen kommen kann (Schmidt et al., 2000). In Pflanzen führt die Cre-Ausprägung während der Organogenese möglicherweise zu abnormen Phänotypen (Coppoolse et al., 2003). Für die Cre-Rekombinase wurde außerdem ein zytotoxischer Effekt festgestellt, dem Chromosomen-Aberrationen vorausgehen. Der Effekt wird nicht durch sequenz-spezifische Rekombination verursacht, sondern beruht auf der Endonuklease-Aktivität von Cre, wodurch Cre-unabhängige, illegitime Rekombinationsereignisse begünstigt werden (Loonstra et al., 2001, Silver & Livingston, 2001).

Einige dieser Probleme könnten durch eine zeitlich limitierte Applikation von Cre bzw. Flp umgangen werden. Diese wird heute z.B. schon mit transienter Expression der Rekombinase von einem nicht-replizierenden Vektor (O’Gorman et al., 1991), mit der Tetracyclin-regulierten Ausprägung der Rekombinase durch ein heterologes Expressionssystem (Gossen & Bujard, 1995) oder mit Fusionsrekombinasen, deren katalytische Aktivität abhängig von der Gegenwart eines spezifischen Liganden ist (Logie & Stewart, 1995; Metzger et al., 1995), erreicht. Einen Weg, die Richtung der Reaktion zu beeinflussen, stellt dagegen der Einsatz modifizierter *target*-Sequenzen dar (Araki et al., 1997).

Möglichkeiten, die Mängel der etablierten sequenz-spezifischen Rekombinasen Cre und FLP zu umgehen, bieten Systeme, bei denen sich die Richtung von Hin- und Rückreaktion durch Kofaktoren steuern lässt. Solche in der Natur vorkommenden Systeme müssen zuvor für den Einsatz in höheren Eukaryoten optimiert werden. Ein Kandidat für eine derartig regulierte Rekombinase ist die vom Phagen *Lambda* kodierte Integrase.

1.2 Die sequenz-spezifische Rekombination des Phagen *Lambda*

Temperente Phagen, wie der Bakteriophage *Lambda*, können nach dem Eindringen in eine Wirtszelle zwei Wege einschlagen. Erstens den virulenten Weg, bei dem sie sich rasant vermehren und schließlich die Wirtszelle durch Lyse zerstören oder zweitens den temperenten Weg. Beim zweiten Weg wird die Phagen-DNA in das Wirtszellgenom eingebaut (Integration) und liegt dort als so genannter Prophage vor. In diesem als Lysogenie bezeichnetem Stadium nutzt der Phage das Replikationssystem von *E. coli* und vermehrt sich passiv, ohne den Wirt zu schädigen. Wie ein Bestandteil der bakteriellen DNA wird er dabei mit jeder Zellteilung an die nächste Generation weitergegeben. Ein Prophage bleibt jedoch für seinen Wirt latent gefährlich, denn der lytische Reaktionsweg kann jederzeit durch eine SOS-Reaktion der Zelle induziert werden. Diese kann als Folge ungünstiger Umweltbedingungen, z.B. einer Schädigung durch UV-Strahlung, ausgelöst werden und führt letztendlich zur Exzision des Phagen, seiner aktiven Vermehrung und zur Lyse der Wirtszelle.

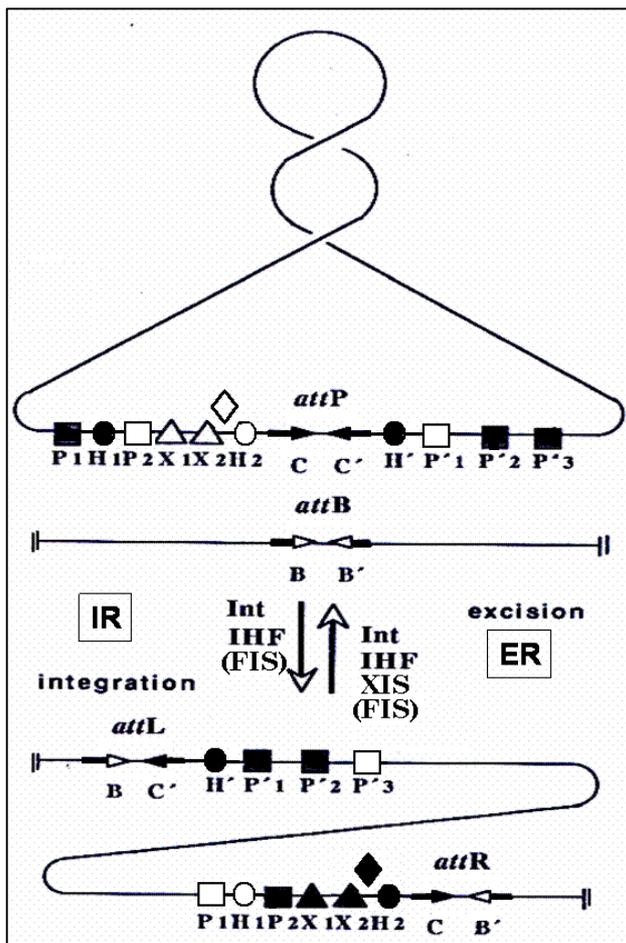


Abbildung 1-2: Schematische Darstellung der integrativen (IR) und exzisen (ER, -) I-Rekombination. Dargestellt ist eine negativ superhelikal gespannte attP-Sequenz, die in Gegenwart von IHF und Int in das lineare attB-Element integriert wird. Dabei entstehen die Produkte attL und attR, die Substrate für die Exzision. Für die Exzision benötigt Int neben IHF zusätzlich die Hilfe von XIS. Alle vier attachment-Regionen besitzen eine gemeinsame core-Region die aus zwei gegenläufigen core-Bindungsstellen besteht: C und C' oder B und B'. Die DNA-Abschnitte links und rechts neben der core-Region in der Phagen att-Region werden als P- bzw. P'-Arm bezeichnet. Sie enthalten weitere Bindungsstellen der Int (P) und Bindungsstellen für IHF (O, H), XIS (Δ, X) und FIS (◇, F). Int und IHF sind für beide Reaktionswege und XIS ist für die Exzision unerlässlich. FIS stimuliert die Exzision. Während der Integration besetzte Proteinbindungsstellen in attP, und bei der Exzision besetzte Bindungsstellen in attLs und attR, sind als gefüllte Symbole dargestellt.

(verändert nach Christ & Dröge, 2001).

Die konservative sequenz-spezifische Rekombination des Bakteriophagen *Lambda* vermittelt die Integration und Exzision der Phagen-DNA in das Genom

seines Wirtes, dem Bakterium *Escherischia coli* (Campbell, 1962; Übersichtsartikel: Landy, 1989; Stark et al., 1992). Die Reaktionen finden zwischen spezifischen Sequenzen statt, den *attachment*-Regionen (*att*) (Abb. 1-2). Bei der Integration sind dies der 243 bp lange DNA-Abschnitt auf dem Phagen-Genom (*attP*) und die 21 bp Zielsequenz im Genom des Wirtes (*attB*). Als Produkte dieser integrativen Rekombination entstehen zwei hybride Regionen, die den inserierten Prophagen links (*attL*) und rechts (*attR*) flankieren und als Substrate für die exzusive Rekombination dienen. Beide Reaktionswege basieren auf einem koordinierten Strangaustausch und werden durch die phagen-kodierte Integrase (Int) in Anwesenheit von Hilfsproteinen katalysiert. Die Art der Hilfsproteine bestimmt dabei die Richtung der Reaktion (Abb. 1-2).

1.2.1 Die integrative λ -Rekombination

Die integrative Rekombination des Phagen *Lambda* findet an den Regionen *attB* des Bakteriengenoms und *attP* des Phagen-genoms statt. Sie benötigt neben der phagenkodierte Integrase (Int) den bakterienkodierte *integration host factor* (IHF) als Hilfsprotein. *attB* ist mit 21 bp die kürzeste *attachment*-Region und besteht

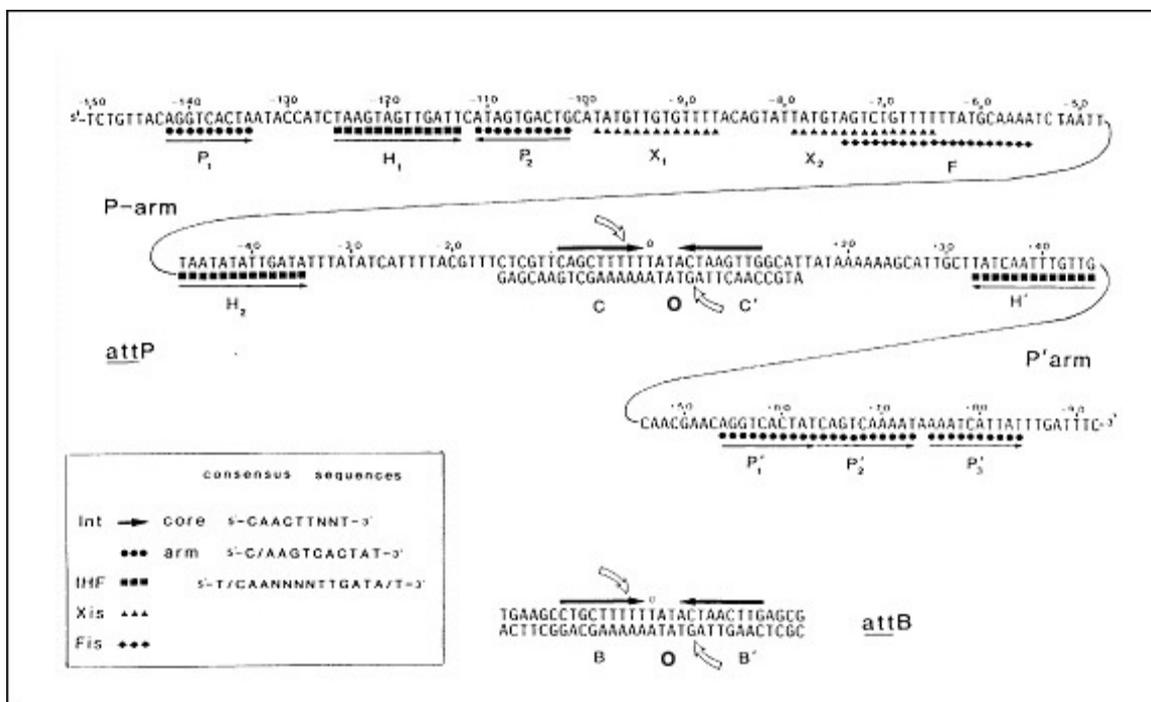


Abb. 1-3: DNA-Sequenz und Proteinbindungsstellen der attP- und attB-Regionen. Die Erkennungssequenzen der jeweiligen Proteine sind durch die im Kasten angezeigten Symbole markiert. Die relativen Orientierungen der Bindungsstellen sind durch Pfeile (→) angezeigt. Die Positionen der Einzelstrangaustausche sind durch gekrümmte Pfeile (↷) gekennzeichnet. (modifiziert nach Landy, 1989).

ausschließlich aus dem so genannten *common core*. Allen *att*-Regionen sind 15 Nukleotide dieses *common core* gemeinsam, das in *attB* als BOB' und in *attP* als COC' oder auch als POP' bezeichnet wird. Es umfasst zwei invertierte Int-Bindungsstellen des *core* Typs die eine 7 bp lange *overlap*-Region (O) flankieren (Abb. 1-3). In der *overlap*-Region findet der Strangaustausch statt und sein asymmetrischer Aufbau bestimmt dabei die Richtung. Mehrere Untersuchungen haben ergeben, dass die Homologie und weniger die Sequenz der *att*-Regionen in diesem Bereich für den koordinierten Strangaustausch essentiell ist (Bauer et al., 1985; Kitts & Nash, 1987; Nunes-Düby et al., 1997).

attP ist mit 243 bp die längste *att*-Region. Die Sequenzen, die ihren *common core* flankieren, werden als Arme bezeichnet und beherbergen diverse Proteinbindungsstellen. Auf dem am 5'-Ende angrenzenden P-Arm liegen zwei Integrase (P_1 und P_2), zwei IHF- (H_1 und H_2), zwei XIS- (X_1 und X_2) und eine FIS-Bindungsstelle

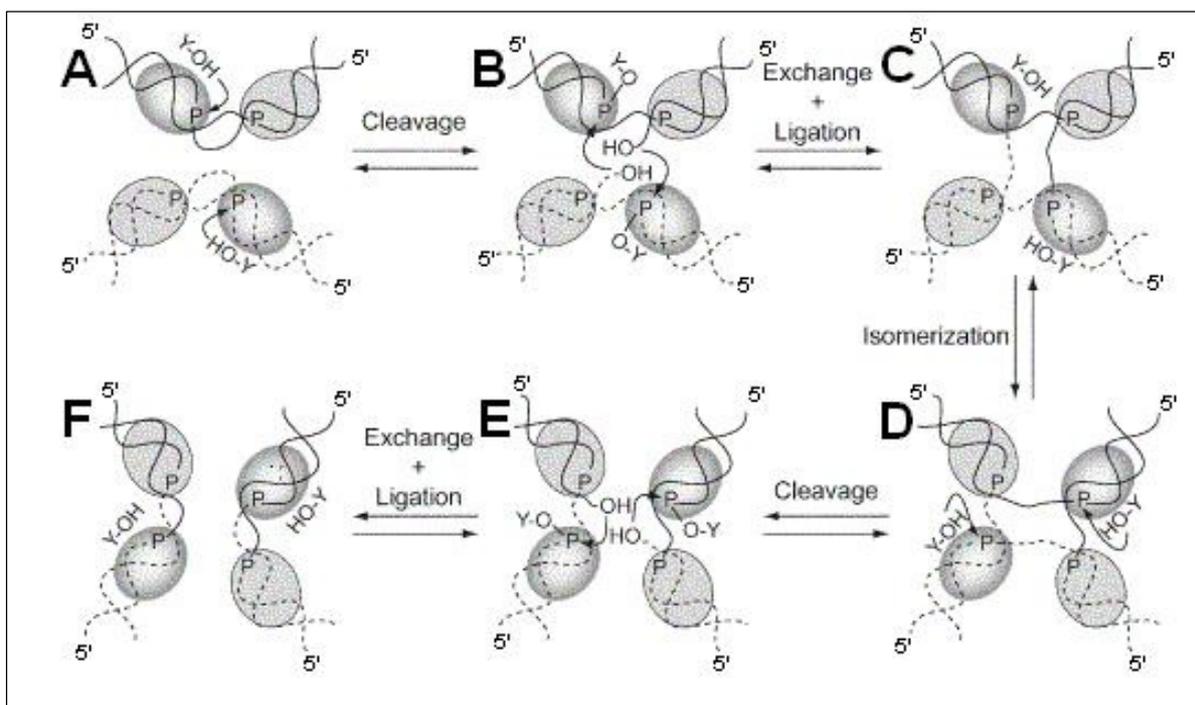


Abb. 1-4: Mechanismus der λ -Integrase (Int) vermittelten sequenz-spezifischen Rekombination (nach Nunes-Düby et al., 1995a). Die Int-Protomere erkennen Paare von invertierten Sequenzen an den Rekombinationsregionen des Phagen (gestrichelte Linien) und des Wirtes (durchgezogene Linien). Das aktivierte Paar von Int-Protomeren, welches das Öffnen der DNA-Stränge katalysiert, ist jeweils hervorgehoben. Die Öffnung wird durch den nukleophilen Angriff eines konservierten Tyrosinrestes (Y-OH) auf das „Ziel-Phosphat“ der DNA durchgeführt (A). Dabei entsteht als Intermediat ein Protein-DNA-Komplex mit einem Tyrosin, das kovalent mit dem 3'-Ende des geöffneten Stranges verbunden ist (Y-O-P) (B). Als nächster Schritt folgen der reziproke Austausch und die Neuverknüpfung der DNA-Stränge unter Bildung eines Holliday-Struktur-Intermediats (C). Proteinwechselwirkungen führen zur Isomerisierung der Struktur, einem kritischen Schritt in diesem Mechanismus. Das zweite Int-Protomer-Paar wird aktiviert (D). Der Isomerisierung folgt eine zweite Runde von Öffnung, Austausch und Neuverknüpfung, die zu dem vollständig integrierten (oder ausgeschnittenen) Endprodukt führt (D-F). (modifiziert aus Subramaniam et al., 2003).

(F). Der P'-Arm am 3'-Ende trägt drei Integrase- (P'_1 , P'_2 , P'_3) und eine IHF-Bindungsstelle (H') (Abb. 1-2 und 1-3). Bei der integrativen Rekombination bindet die Integrase gleichzeitig an Arm- und *core*-Bindungsstellen von *attP*, welches in negativ superhelikal gespannter Form vorliegen muss (Richet et al., 1986; Nash, 1990). Die relativ große Entfernung zwischen den Sequenzen wird dabei mit Hilfe des DNA-biegenden Proteins IHF überwunden. Der entstehende Int-IHF-*attP*-Nukleoproteinkomplex, auch als Intasom bezeichnet (Echols, 1989), interagiert mit den *core*-Regionen des bis dahin proteinfreien *attB*s und bildet den so genannten synaptischen Komplex (Richet et al., 1988). Das Zusammentreffen des Intasoms mit *attB* geschieht zufällig und ist nicht das Ergebnis eines Suchvorgangs durch Diffusion entlang eines DNA-Stranges (Griffith & Nash, 1985; Sprengler et al., 1985; Richet et al., 1988). Die Rekombination erfolgt über Bildung, Isomerisierung und Auflösung eines so genannten Holliday-Intermediats (Holliday, 1964; Nunes-Düby et al., 1987; Kitts & Nash, 1988; Burgin & Nash, 1992; schematische Darstellung, Abb. 1-4). An den Enden der *overlap*-Sequenzen der an Int-gebundenen *core*-Regionen findet der koordinierte Strangaustausch statt. Er wird in zwei sequenziellen Schritten durchgeführt: der erste führt zur Bildung, der zweite zur Auflösung der Holliday-Struktur. Weder die Öffnung noch die Neuverknüpfung der DNA-Stränge erfordern ATP oder DNA-Synthese. Die Energie wird in einem kovalenten Protein-DNA-Intermediat gespeichert, das durch den Phosphoryl-Transfer auf ein konserviertes Tyrosin (Y342) der Integrase gebildet wird (siehe auch Abb.1-4). Produkte der erfolgreichen Rekombination zwischen *attP* und *attB* sind *attL* und *attR*, die ihrerseits als Substrate für die exzitive λ -Rekombination dienen können.

1.2.2 Die exzitive λ -Rekombination

Die exzitive λ -Rekombination findet an den Regionen *attL* und *attR* statt, die natürlicherweise den inserierten Prophagen flankieren (Kap. 1.2; Abb. 1-2). Im Gegensatz zur integrativen Rekombination ist keine superhelikale Spannung der DNA erforderlich (Abremski & Gottesmann, 1979). An beiden involvierten *attachment*-Regionen wird ein Nukleoproteinkomplex (Intasom) mit der phagenkodierten Int und dem bakterienkodierten IHF ausgebildet (Kim und Landy, 1992). *attR* braucht zur Ausbildung eines funktionellen Komplexes jedoch noch die Bindung des phagenkodierten Faktors XIS (Bushman et al., 1984). FIS (*factor for inversion stimulation*), ist ein Protein, das vom Bakterium kodiert wird. Es ist *in vivo*

für eine effiziente exzusive Rekombination essentiell (Ball & Johnson, 1991), *in vitro* wird sie durch FIS gesteigert (Thompson et al., 1987; Landy, 1989). Wie bei der integrativen Reaktion führen auch bei der exzisiven Reaktion zwei koordinierte Einzelstrangaustausche zur Rekombination (siehe auch Abb. 1-4), wobei sich der synaptische Komplex ebenfalls nach zufälligem Zusammentreffen der Intasome bildet.

1.2.3 Die λ -Integrase

Die λ -Integrase (Int) ist der Prototyp der großen Familie der konservativen sequenz-spezifischen Tyrosin-Rekombinasen (Nunes-Düby et al., 1998). Sie katalysiert die Öffnung, den Strangaustausch und die Neuverknüpfung der DNA bei der sequenz-spezifischen Rekombination des Phagen λ . Im Gegensatz zu den einfachen Mitgliedern dieser Familie, besteht die λ -Integrase aus drei Protein-Domänen mit insgesamt zwei autonomen DNA-Bindungsstellen für spezifische Sequenzen der *attachment*-Regionen (Moitoso de Vargas et al., 1988; Tirumalai et

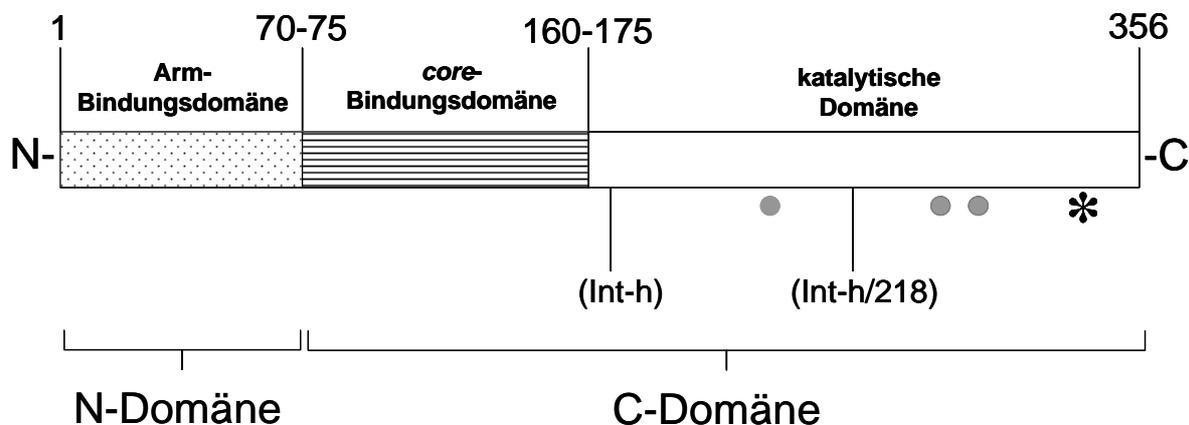


Abbildung 1-5: Schematische Darstellung der *Lambda*-Integrase. Die drei in DNA-Arm-Bindung, *core*-Bindung und Katalyse involvierten Domänen (Tirumalai et al., 1997) sind durch unterschiedliche Musterung hervorgehoben. Die Zahlen beziehen sich auf die Aminosäurereste (AS). AS70-75 und AS166-175 sind Bestandteil einer *Linker*-Region zwischen N- und C-Domäne, bzw. zwischen *core*-Bindungs- und katalytischer Domäne (Sarkar et al. 2002; Aihara et al., 2003). Die grauen Punkte markieren die Position der drei innerhalb der Integrase-Familie hochkonservierten Aminosäurereste Arg212, His308 und Arg311 (Nunes-Düby et al., 1998). Mit einem Stern ist der aktive Aminosäurerest Tyr342 und mit Int-h oder Int-h/218 sind die entsprechenden Aminosäureaustausche in den Integrase-Mutanten markiert. Mit N- ist der N-Terminus, mit C- der C-Terminus des Proteins markiert. (aus Christ, 2002 Dissertation und angepasst nach Sarkar et al., 2002 und Aihara et al., 2003)

al., 1997; Abb. 1-5). Ihre kleinste Domäne ist die Arm-Bindungsdomäne, im Folgenden auch N-Domäne genannt. Sie umfasst die Aminosäurereste 1-70 (Sakar et al., 2002; Abb.1-5) und bindet an fünf Bindungssequenzen des Arm-Typs (von *attP*, *attL* und *attR*), die 70-150 Basenpaare vom Ort des Strangaustausches entfernt

liegen. Für jeden Reaktionsweg werden dabei nur drei besetzt: P1, P'2 und P'3 bei der integrativen bzw. P2, P'1 und P'2 bei der exzisivem Rekombination (Thompson et al., 1987b; Numrych et al., 1990; Abb. 1-2). Obwohl sie für die von der λ -Integrase katalysierten Reaktionen essentiell ist, besitzen viele Tyrosin-Rekombinasen so eine zusätzliche Bindungsdomäne nicht. Die Struktur der N-Domäne von Int wurde 2002 von Wojciak und Mitarbeitern aufgeklärt. Nach dem von ihnen vorgestellten Modell besteht die N-Domäne aus einer 3 strängigen β -Faltblatt-Struktur und einer α -Helix. Mit der β -Faltblatt-Struktur erkennt und bindet sie die Arm-Regionen auf der DNA. Dagegen scheint die α -Helix sowohl für die homomere Interaktion zwischen zwei Int-Molekülen (Jessop et al., 2000; Warren et al., 2003) als auch für die Interaktion mit dem C-Terminus von XIS wichtig zu sein (Cho et al., 2002; Warren et al., 2003). Neben ihren DNA- und Proteinbindungseigenschaften besitzt die N-Domäne auch eine kontext-sensitive, modulierende Funktion auf die DNA-Binde- und Schneideaktivität der C-Domäne (AS 70-356). Diese Aussage beruht auf Beobachtungen, die bei Abspaltung der N-Domäne von Int gemacht wurden. Dabei wurde gezeigt, dass eine verkürzte Integrase ohne N-Domäne das komplette Protein in Topoisomerasefunktion, in sequenz-spezifischer Schneideaktivität und bei der Bindung an den *core*-Regionen übertrifft (Sarkar et al., 2001). Fügt man die N-Domäne als separates Molekül (in *trans*) hinzu, werden die Funktionen der C-Domäne zusätzlich gesteigert. Die N-Domäne behindert demzufolge die intrinsischen Funktionen der Integrase nur in *cis* als Bestandteil des kompletten Proteins. Diese Ergebnisse und die Tatsache, dass diese Behinderung in der Gegenwart von Oligonukleotiden mit Arm-Sequenzen überwunden werden kann, ließen Sarkar und Mitarbeiter 2001 folgendes für die Funktionen der N-Domäne postulieren: Die Integrase wird durch die N-Domäne in einem inaktiven Zustand gehalten, bis diese an eine Arm-Sequenz bindet. Dabei wird durch die Bindung entweder eine Konformationsänderung der N-Domäne von einer inhibierenden in eine aktivierende Position eingeleitet oder die Dimerisierung der Integrase gefördert, die ihrerseits die simultane Bindung der *core*-Region favorisiert.

Die *core*-Bindungsdomäne von Int (AS 75-159) und die katalytische Domäne (AS 170-356) sind über einen flexiblen Aminosäure-*Linker* (I160-R176) miteinander verbunden (Aihara et al., 2003; Abb. 1-6). Dieser ermöglicht es den beiden Domänen, die DNA wie eine Zange zu umfassen. Dabei interagiert die *core*-Bindungsdomäne mit der großen Furche der DNA und die katalytische Domäne auf

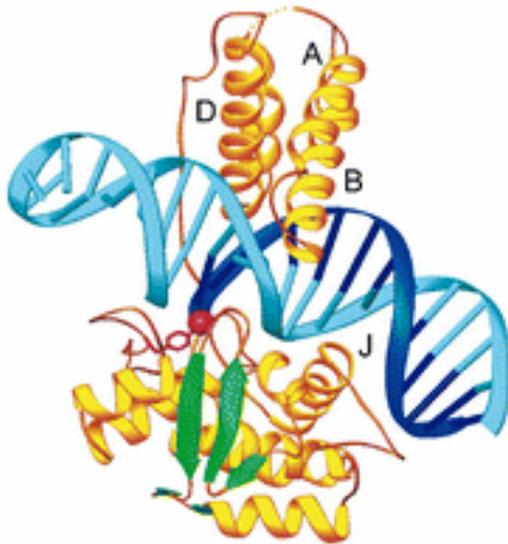


Abb. 1-6: Die C-Domäne umgreift die DNA. Das Diagramm zeigt die *core*-Bindungsdomäne (AS 75-160; über der DNA) und die katalytische Domäne (AS 170-356; unter der DNA) der λ -Integrase und ihre Interaktionen mit den kleinen und großen Furchen auf den gegenüberliegenden Seiten der DNA. Ein langer, ausgedehnter Linker (I160-R176) verbindet die beiden Untereinheiten. Das mit Y342 kovalent verbundene Phosphat ist als rote Kugel dargestellt. Die *core*-Bindungsdomäne ragt in die große Furche nahe der DNA-Schnittstelle. Die katalytische Domäne interagiert mit der großen und kleinen Furche auf der gegenüberliegenden Seite und spreizt dabei die DNA-Schnittstelle (modifiziert nach Aihara et al., 2003).

der gegenüberliegenden Seite mit der kleinen und großen Furche der DNA. Zusammen werden *core*-Bindungs- und katalytische Domäne auch als C-Domäne bezeichnet (Abb. 1-5). Die *core*-Bindungsdomäne vermittelt hauptsächlich die Bindungsaffinität für die Sequenzen des *core*-typs (Tirumalai et al., 1997, 1998). Sie ist wiederum über einen mindestens vier Aminosäurereste umfassenden *Linker* mit der N-Domäne von Int verbunden (Sarkar, et al., 2002) und besitzt eine funktionelle und strukturelle Analogie zur N-terminalen Domäne einfacherer Tyrosin-Rekombinasen (Swalla et al., 2003; Aihara et al., 2003; Subramanya et al., 1997; Guo et al., 1997). Diese besitzen keine Arm-Bindungsdomäne dafür aber eine höhere Affinität zu ihren Sequenzen des Strangaustausches. Die katalytische Domäne beherbergt die enzymatische Aktivität der Integrase und ist auch ohne die anderen Domänen der Integrase noch fähig die DNA zu schneiden und zu ligieren (Tirumalai et al., 1997). Sie besitzt vier konservierte Aminosäurereste, die die Familie der Tyrosin-Rekombinasen auszeichnen (Argos et al., 1986). Eine katalytische Triade mit R212, H308 und R311 unterstützt das aktive Tyrosin 342 bei seinem nukleophilen Angriff auf die DNA (Esposito & Scocca, 1997). Das Tyr342 liegt in Abwesenheit von DNA sehr weit von der katalytischen Triade der eigenen katalytischen Domäne entfernt (Kwon et al., 1997). Ursprünglich ließ dies vermuten, dass Int die DNA in einem *trans*-Mechanismus schneidet, wobei im Rekombinationskomplex das Tyr342 einer Integrase-Untereinheit in das aktive Zentrum einer benachbarten Untereinheit hineinragt. Die Aufklärung der an DNA gebundenen Struktur zeigt jedoch, dass die C-Domäne eine Konformationsänderung vollzieht. Dabei werden das Tyr342 in das aktive Zentrum der eigenen katalytischen Domäne positioniert und die letzten acht C-terminalen Reste für eine Interaktion mit

benachbarten Int-Molekülen exponiert (Aihara et al., 2003). Die gesamte Integrase besitzt eine Größe von ~ 40 KDa.

Int-h ist eine klassische Mutante der Integrase (Miller et al., 1980; Kikuchi et al., 1985), die sich in nur einer Aminosäure vom wild-typ-Protein unterscheidet. Der Austausch befindet sich im *Linker* zwischen *core*-Bindungsdomäne und katalytischer Domäne an Position 174. An dieser Stelle ist statt eines Glutamats ein Lysinrest in die Polypeptidkette eingebaut (Bear et al., 1987). Im Vergleich zu Int führt dieser Austausch zu einer höheren Affinität für *core*-Sequenzen (Patsey & Bruist 1995). Die positiv geladene ϵ -Aminogruppe des Lysins könnte eine ionische Bindung mit dem Phosphatrückgrat der DNA ausbilden und so diesen Effekt hervorrufen. Ebenso wäre es möglich, dass die Mutation die Flexibilität des *Linkers* einschränkt und somit die benötigte Bindungsenergie für die DNA verringert (Aihara et al., 2003). Die höhere Affinität von Int-h zu *core*-Sequenzen führt *in vivo* und *in vitro* bei Anwesenheit von IHF zu einer häufigeren Integration in *att*-analoge Sequenzen (Miller et al., 1980; Patsey & Bruist, 1995). Als *att*-analog bezeichnet man Sequenzen, die Ähnlichkeit mit *attB* haben und deswegen von Int als Integrationsort benutzt werden. Beim wild-typ-Protein ist dies jedoch eher selten der Fall, da Sequenzunterschiede vor allem im *overlap*-Bereich die Isomerisierung der Holliday-Struktur stark hemmen (Kitts & Nash, 1987; Nunes-Düby et al., 1997). Int-h weist dagegen in Gegenwart von IHF für diese Sequenzunterschiede eine höhere Toleranz auf (Patsey & Bruist, 1995; Christ & Dröge 1999). *In vivo* und *in vitro* Untersuchungen haben ergeben, dass der Aminosäureaustausch Int-h in Abwesenheit von IHF zu einer höheren Rekombinationsaktivität gegenüber dem wild-typ-Protein führt (Christ & Dröge 1999; Lange-Gustafson & Nash, 1984). Während sich *in vitro* und in Abwesenheit von IHF die Aktivität von Int um das 500fache reduziert (Landy, 1989), verringert sich die Rekombinationsaktivität von Int-h nur um den Faktor 10. Auch nicht superhelikal gespannte Substrate können durch Int-h besser rekombiniert werden als durch Int (Lange-Gustafson & Nash, 1984). Für Int-h ergibt sich demnach eine teilweise Unabhängigkeit von Kofaktoren.

Int-h/218 ist eine Doppel-Mutante der λ -Integrase (Christ & Dröge 1999). Sie ist abgeleitet von Int-h und besitzt neben E174K einen weiteren Aminosäureaustausch in Position 218 der katalytischen Domäne, wo ebenfalls ein Glutamat gegen ein Lysin ausgetauscht wurde. Wu et al. (1997) vermuteten, dass der E218K Austausch die Affinität für die *core*-Sequenzen der *attachment*-Regionen steigert. Sie entdeckten

die Mutation E218K erstmalig in einer Revertanten, die mit diesem Austausch eine andere Mutation in der *core*-Bindungsdomäne kompensierte, die zum Verlust der Bindungsspezifität für *core*-Sequenzen führte. Die E218K Mutation wurde in Int-h eingeführt, um die Affinität für *core*-Sequenzen zu verbessern und die Rekombinationsaktivität in *att*-analoge Sequenzen, die auch im menschlichen Genom vorkommen (Lorbach, 2000 Dissertation; Christ, 2002 Dissertation) zu erhöhen. In Abwesenheit von Hilfsproteinen besitzt Int-h/218 bei der integrativen und bei der exzisen Rekombination eine höhere Aktivität als Int-h (Christ & Dröge, 1999). Die Mutanten sind jedoch nur teilweise von Kofaktoren unabhängig, denn wie bei der wild-typ-Integrase ist ihre Aktivität durch Hilfsproteine stimulierbar. IHF steigert die integrative Rekombination der Mutanten (Christ & Dröge, 1999, Lange-Gustafson & Nash, 1984). Liegen XIS und IHF vor, so wurde bisher nur für Int-h die Steigerung der exzisen Rekombination untersucht und nachgewiesen (Kikuchi et al., 1985).

1.2.4 Die Protein-Kofaktoren der λ -Integrase: IHF, XIS und FIS

Der ***integration host factor (IHF)*** ist ein ca. 20 kDa großes heterodimeres Protein aus *Escherichia coli*, das eine wichtige Rolle in einer Vielfalt von zellulären Prozessen spielt wie in der transkriptionellen Regulation, Replikation, DNA-Komprimierung und bei verschiedenen sequenz-spezifischen Rekombinationen (Übersichtsartikel: Nash, 1996 und Goosen & van de Putte, 1995). Beim integrativen und beim exzisen Reaktionsweg der λ -Rekombination ist IHF als Strukturprotein ein wichtiger Kofaktor für die Ausbildung von definierten Nukleoprotein-Komplexen (Übersichtsartikel: Azaro & Landy, 2002). Seine biologischen Funktionen vermittelt IHF hauptsächlich durch spezifische Bindung an DNA und die resultierende Biegung (Goodrich et al., 1990). Die Kristallstruktur des an die H'-Sequenz des λ -Phagen gebundenen Proteins (Abb. 4-5) verdeutlicht die Krümmung, die durch seine Bindung an die DNA verursacht wird (Rice et al., 1996). Mit über 160° gilt sie bis heute als stärkste Krümmung, die durch ein sequenz-spezifisches Protein in eine DNA eingeführt wird (Übersichtsartikel: Travers, 1997 und Ellenberger & Landy, 1997). IHF wird von zwei homologen und ähnlich großen α - und β -Untereinheiten gebildet, deren Sequenzen zu 30% Identität aufweisen. Sie umschlingen sich gegenseitig und bilden in der tertiären Struktur einen Proteinkern aus dem zwei Schleifen mit β -Faltblatt-Struktur herausragen. Zur Biegung der DNA schieben sich beide Arme des

Heterodimers mit ihren konservierten Prolinresten in die kleine Furche der DNA. Die Krümmung um das Protein wird zusätzlich durch elektrostatische Interaktionen beider Untereinheiten mit der DNA stabilisiert. Die Gene, die für diese Untereinheiten kodieren heißen *himA* (Miller et al., 1980; IHF- α) und *himD* (Kikuchi et al., 1985; IHF- β). Die Überexpression dieser Untereinheiten in *E. coli* führt zu unlöslichen oder instabilen Polypeptiden (Nash et al., 1987; Wang & Chong, 2003). Auch der Versuch, IHF in eukaryotischen Zellen auszuprägen, war bisher nicht erfolgreich (Nicole Christ, 2002 Dissertation). Diese Tatsache erschwert seinen Einsatz als regulatorischen Faktor bei der Übertragung des λ -Systems auf eukaryotische Zellen.

Während IHF der einzige Protein-Kofaktor ist, der zur Integration des Phagen λ erforderlich ist, benötigt die exzusive Rekombination noch zusätzlich das **XIS**-Protein. XIS ist ein 9 kDa kleines, phagenkodiertes Protein dessen Struktur erst kürzlich aufgeklärt wurde (Sam et al., 2003). Es benutzt ein ungewöhnliches „winged“-Helix Motiv um mit der DNA zu interagieren. Eine einzelne α -Helix kommt dabei in der großen Furche der DNA zu liegen, während eine Schlaufenstruktur („wing“) die kleine Furche kontaktiert. Für XIS wurde keine katalytische Aktivität nachgewiesen aber seine kooperative Bindung an X1 und X2 in *attR* (Bushman et al., 1984; Yin et al., 1985) führt zur Krümmung der DNA (Thompson & Landy, 1988) und ist für die Ausbildung des exzisiven Intasoms essentiell (Bushman et al., 1985; Kim et al., 1990; Kim & Landy, 1992; Better et al., 1983; Better et al., 1982). Die Integration dagegen wird durch XIS gehemmt, da *attP* bei dessen Bindung eine katalytisch inaktive Struktur annimmt (Bushman et al., 1984; Franz & Landy 1995; Moitoso de Vargas & Landy, 1991). Neben der richtungweisenden strukturellen Funktion unterstützt XIS durch die Interaktion seines C-Terminus (Nymrich et al., 1992; Wu et al., 1998) mit der N-Domäne von Int (Cho et al., 2002; Sarkar et al., 2002) die kooperative Bindung der Integrase an P2 (Bushman et al., 1984).

FIS (factor for inversion stimulation) ist ein bakterienkodiertes homodimeres Protein mit 98 AS, das man in *E. coli* und vielen anderen Prokaryoten findet. Es bindet spezifisch und unspezifisch an DNA und ist in der Lage die DNA zu krümmen ($\sim 90^\circ$). FIS ist in viele zelluläre Prozesse involviert, darunter z.B. die Inhibierung (Ball et al., 1992; Koch et al., 1991; Xu & Johnson, 1995) oder die transkriptionelle Aktivierung (Nilsson et al., 1990; Ross et al., 1990; Xu & Johnson, 1995b) einer Vielzahl von Promotoren und die Beteiligung als Kofaktor bei der DNA- Replikation (Filutowicz et al., 1992), der Zellteilung und der Chromosomentrennung (Paull &

Johnson, 1995). Bei der sequenz-spezifischen Rekombination des Phagen *Lambda* ist FIS *in vivo* für eine effiziente exzusive Rekombination essentiell (Ball & Johnson, 1991) *in vitro* wird diese bei suboptimalen XIS-Mengen durch FIS gesteigert (Thompson et al., 1987; Landy, 1989). Auch die integrative Rekombination wird *in vivo* durch FIS gesteigert (Ball & Johnson, 1991b); *in vitro* konnte ein Effekt jedoch erst 2003 durch Esposito und Gerard nachgewiesen werden. Sie konnten zeigen, dass bei niedrigen Int-Konzentrationen, die eher der Situation *in vivo* entsprechen, die integrative Reaktion um bis zu siebenfach gesteigert wird. Auch die Rekombination von *attP*- bzw. *attB*-Substraten, die nicht die optimale Standardtopologie, d.h. negative superhelikale Spannung, aufweisen, ist *in vitro* durch FIS erhöht.

1.3 Erzeugung transgener Mäuse durch Pronukleus-Injektion

In der Biotechnologie werden für Genfunktionsstudien häufig transgene Tiere eingesetzt. Da in der vorliegenden Arbeit eine transgene Maus generiert und analysiert werden sollte, wird an dieser Stelle kurz auf die Erzeugung dieser Tiere eingegangen. Die Einführung genetischen Materials in die Mauskeimbahn wird heutzutage routinemäßig angewandt und als kommerzieller Service von Firmen zur Verfügung gestellt. Die Bezeichnung transgen beschreibt Organismen, denen fremde DNA-Sequenzen stabil in ihr Genom integriert wurde. Transgene Tiere können auf verschiedenen Wegen hergestellt werden: Durch Mikroinjektion, durch Vireninfektion von Embryonen oder durch die Manipulation von embryonalen Stammzellen, die später in sich entwickelnde Blastozysten transferiert werden. Hier soll nur auf die Methode der Pronukleus-Injektion näher eingegangen werden (Abb. 1-7), bei der die DNA direkt in eine befruchtete Eizelle eingebracht wird. Das erste Tier, das aus dieser manipulierten Eizelle hervorgeht, wird *Founder* genannt. Auch wenn mehrere Eizellen mit dem gleichen DNA-Fragment injiziert wurden sind die Integrationsorte in jedem *Founder* –Tier verschieden. Mit gegenwärtigen Protokollen für die Generierung transgener Mäuse durch Embryo-Mikroinjektion kann der Integrationsort der eingeführten DNA nicht vorherbestimmt werden und muss daher als zufällig erachtet werden. Die Methode erlaubt nur das Hinzufügen, nicht aber die Entfernung genetischen Materials. Wird beispielsweise ein mutiertes Mausgen in das Genom eingeführt, so ist das neue Allel zusätzlich zum normalen diploiden Paar anwesend.

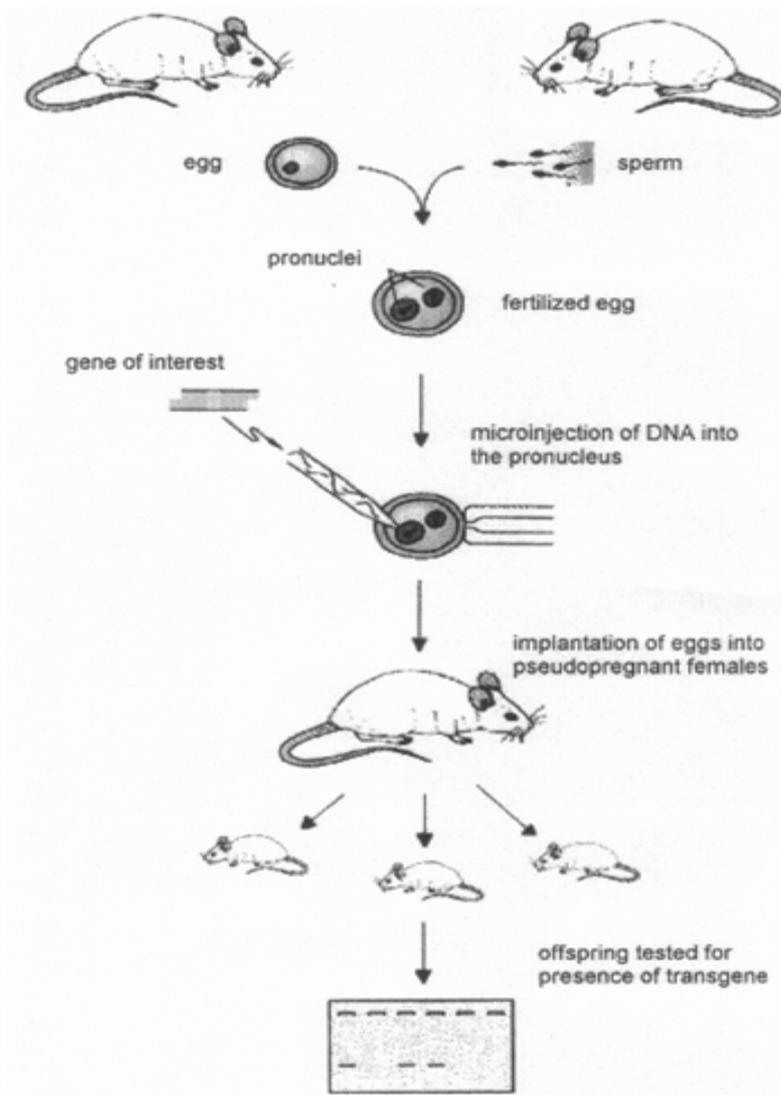


Abb. 1-7: Pronukleus Injektion. Schematische Darstellung der Erzeugung transgener Mäuse durch Pronukleus Injektion. Das DNA-Fragment, das das relevante Gen trägt, wird in einen der Pronuclei einer befruchteten Eizelle injiziert. Durch zufällige Integration wird das Fragment Bestandteil des Genoms der Maus. Die befruchteten, das Fragment enthaltenden, Eizellen werden in den Eileiter von scheinsschwangeren Weibchen eingepflanzt. Nach einer Tragzeit von ca. drei Wochen werden die ersten Tiere geboren. Weitere drei bis vier Wochen später wird der Wurf von der Mutter getrennt. Um transgene *Founder*-Tiere zu identifizieren wird die DNA aus den Schwanzspitzen mittels PCR und *Southern Blot* analysiert. Positive transgene Tiere können ab einem Alter von ca. sechs Wochen mit wild-typ-Mäusen weiter verpaart werden. In der Regel erfolgt dabei die Verteilung des eingeführten Gens nach den Mendelschen Gesetzen. Ist das injizierte Fragment erst im zwei-Zellstadium oder später integriert, kann es zu Mosaik-Mäusen kommen, bei denen die Anzahl positiver Nachkommen geringer ist, als nach Mendel erwartet.

Infolgedessen können nur dominante oder ko-dominante Formen den Phänotypen beeinflussen. Einschränkungen bei Experimenten mit transgenen Tieren können dann auftreten, wenn z.B. der genomische Integrationsort des eingeführten Gens seine Ausprägung verhindert. Ebenso kann die Ausprägung des eingeführten Gens letal sein oder zu einem abnormen Phänotyp führen. In der Mehrzahl der Fälle, die bis jetzt analysiert wurden, hat die Unterbrechung der endogenen Reihenfolge, die durch Integration von Fremd-DNA verursacht werden, keinen offensichtlichen Effekt auf den Phänotyp. Jedoch bedeutet das Fehlen eines nachweisbaren Phänotypus nicht notwendigerweise, dass die Integration in eine unfunktionale Region des Genoms erfolgt ist.

1.4 Zielsetzung der Arbeit

Das *Lambda*-System der sequenz-spezifischen Rekombination wurde erst kürzlich erfolgreich auf eukaryotische Zellen übertragen (Lorbach et al., 2000, Christ & Dröge, 2001). Dabei konnte gezeigt werden, dass die in Kapitel 1.2.3 beschriebenen Integrase-Mutanten Int-h und Int-h/218 die exzusive und integrative Rekombination in Abwesenheit von Hilfsproteinen und negativer superhelikaler Spannung des Substrats katalysieren. Es wird vermutet, dass diese Eigenschaft auf eine erhöhte Affinität von Int-h und Int-h/218 für die so genannten *core*-Sequenzen (Kapitel 1.2.1) beruht. Direkte biochemische Indizien oder ein detailliertes Verständnis der Rekombinationswege von Int-h und Int-h/218 im eukaryotischen Milieu fehlen jedoch bislang.

Mit dieser Arbeit wurde unter anderem versucht, weitere Einblicke in den Mechanismus der von Int-h und Int-h/218 vermittelten Rekombination zu erlangen. Insbesondere wurde dabei die Bedeutung der Arm-Bindungssequenzen (Kapitel 1.2.1) der *att*-Regionen bei der Rekombination in Abwesenheit von natürlichen Kofaktoren untersucht. Neben verschiedenen *att*-Regionen, die als episomale Substrate angeboten wurden, wurde auch die von der N-Domäne getrennte C-Domäne auf Rekombinationsaktivität in menschlichen Zellen getestet.

Nachdem die Integrase-vermittelte sequenz-spezifische Inversion und Deletion von DNA-Fragmenten auf dem Genom eukaryotischer Zellen nachgewiesen werden konnte (Lorbach et al., 2000; Christ & Dröge 2002), war ein weiteres Ziel dieser Arbeit die Integration von Fremd-DNA in eine zuvor im menschlichen Genom platzierte oder in eine natürlich vorkommende *attB*-analoge Sequenz.

Die Direktionalität der λ -Rekombination geht durch den Einsatz der Kofaktor unabhängigen Mutanten Int-h und Int-h/218 verloren. Der Einsatz von Hilfsproteinen (Kapitel 1.2.4) und wild-typ-Integrase bietet jedoch die Möglichkeit, erneut Kontrolle über die Richtung der Reaktion zu gewinnen. Vor diesem Hintergrund wird eine rekombinante monomere Form des *integration host factors* (Kapitel 1.2.4) auf seine Eigenschaften in HeLa-Zellen getestet. Für Untersuchungen im Tiermodell soll eine transgene Maus mit diesem rekombinanten Protein generiert werden.

2 Material

2.1 Bakterienstämme

2.1.1.1 Folgende *E.coli*-Bakterienstämme wurden für Klonierungs- und Plasmidpräparationen eingesetzt:

XL1-Blue	F ⁺ ::Tn10 proA ⁺ B ⁺ lacI ^q Δ(lacZ)M15/recA1 endA1 gyrA96(Nal ^r) thi hsdR17 (r _k ⁻ m _k ⁺)supE44 reA1 lac (Bullock et al.,1987)
DH5a	F ⁻ supE44 ΔlacU169 (80lacZφΔ M15) hsdR17 recA1, endA1, gyrA96 thi-I relA1 (Hanahan, 1985)

2.2 Chemikalien

Wenn nicht anders angegeben, wurden in dieser Arbeit Chemikalien von folgenden Firmen benutzt: Amersham Pharmacia (Freiburg); Fluka AG (Buchs Schweiz); GibcoBRL Life Technologies (Eggenstein); ICN Biomedicals (Meckenheim); Merck (Darmstadt); peqlab GmbH (Erlangen); Carl Roth GmbH + Co. (Karlsruhe); Serva Feinbiochemie (Heidelberg); Sigma Chemie (Deisenhofen).

2.3 Enzyme

Enzyme wurden bezogen von: New England Biolabs GmbH (Schwalbach); Amersham/Pharmacia (Freiburg), Roche/Boehringer (Mannheim); GibcoBRL Life Technologies (Eggenstein) und Takara (Otsu, Shiga, Japan).

2.4 Synthetische Oligonucleotide

Oligonucleotide wurden von der Firma ARK/Sigma (Darmstadt) bezogen. Die in dieser Arbeit verwendeten Oligonucleotide und ihr Verwendungszweck sind in der nachfolgenden Tabelle aufgeführt:

PRIMER	SEQUENZ	BEMERKUNG
attP*ns-S	5'-CAGCTTTCTTATACAAAAGTTGGCATTATACAAAAGCAT-TGC-3'	Klonierung von pCMVattPmut und pWSP*nsGFP
attP4	5'-CGTGACGTAGCCTTCGGGCATGGC-3'	
attP*ns-AS	5'-GCAATGCTTTTGTATAATGCCAACTTTGTATAAGAAA-GCTG-3'	
attP2	5'-TAACGCTTACAATTTACGCGT-3'	
attP*ns-Not	5'-ATAAGAATGCGGCCGCCAAATAATGATTTGATTTTGACT-GATATG-3'	
attP-ATC-1	5'-TTTGGATAAAAAACAGACTAGATAAATACTGTAAAACACAA-GATATGCAGTCACTA-3'	
attP-ATC-3	5'-CTGCATATCTTGTGTTTTACAGTATTATCTAGTCTGTTTT-TATCCAAAATCTAA-3'	
attP-Xbal	5'-GACTGTCTAGAGAAATCAAATAATGAT-3'	
attP-PstI	5'-GACTGCTGCAGCTCTGTTACAGGTCAC-3'	
Int-N-C-Dom-PstI	5'-AACTGCAGCCACCATGGGCACAGCGAGAATCAACAGTG-3'	Klonierung von pCMVSSInt-CD und pCMVSSInt-h/218-CD
Int-C-Eu	5'-GCTCTAGAGCGGCCGCTCATTATTTGATT-3'	
Caggs-5-EcoRI	5'-CCGGAATTCTTCTTCTTCGTATTCCTTCTCT-3'	Klonierung von pRIHF2P
Caggs-3-NheI	5'-CTAGCTAGCAGGCAGCGTCGCAGCGACTCC-3'	
SS-5-NheI	5'-CCCAAGCTTGCTAGCAATTCGCTGTCTGCGAGG-3'	
bPARP-BamHI	5'-CGGGATCCCCAGCTGGTTCTTTCCG-3'	
attH2-Pst-Xba-S	5'-GTTCTGCTTTCTTATACCAAGTGGCT-3'	Klonierung von pCMVattH2
attH2-Pst-Xba-AS	5'-CTAGAGCCACTTGGTATAAGAAAGCAGAACTGCA-3'	
attB-Pstl-Xbal-1	5'-GAAGCCTGCTTTTTTATACTAACTTGAGCGT-3'	Klonierung von pCMVattH2
attB-Pstl-Xbal-2	5'-CTAGACGCTCAAGTTAGTATAAAAAAGCAGGCTTCTGCA-3'	
bPA-HindIII	5'-AGTGCCAAGCTTCCCCAGCTGGTTCTTTC-3'	Klonierung von pTKHattB
attB-Mfel	5'-CAACAACAATTGCTGCTTTTTTATACTAACTTGTGCAG-CCAATATGGGATCG-3'	
5'PA-Not I	5'-ATAAGAATGCGGCCGCTCTAGAGCTCGCTGATCAG-3'	

3'RP-PGK-Not I	5'-ATAAGAATGCGGCCGCGGATGTGGAATGTGTGCGA-3'	Klonierung von prIHF2GFP
attB-5' (P1)	5'-CTGCAATAAACAAGTTAACAACAAC-3'	
PE (P3)	5'-GGGGATCCTCTGTTACAGGTCACTAATAC-3'	
Neo505R (P2)	5'-CAGAAGAACTCGTCAAGAAGGCGA-3'	
Neo504 (P4)	5'-GCCCCGGTTCTTTTTGTCAAGACCG-3'	
MCTII	5'-CGCGGATCCGGCCACAAGTATTCCTT-3'	PCR-Analyse für korrektes attH2- Targeting
eGFP-E-Xba	5'-GCTCTAGATCATTACTTGTACAGCTCGTCC-3'	
IHF5-1	5'-AACTGCAGCCACCATGGGCACCAAGTCAGAATTGATA-GAAAGACT-3'	PCR-Analyse zur Identifizierung rIHF transgener Zelllinien und Mäuse
IHF3-His	5'-GCTCTAGAGAATTCTTATCAGTGGTGGTGGTGGTGG-TGGCCTGATCCACCGTAGATATTGGCGCGATCG-3'	
SS5'	5'-CCCAAGCTTGCTAGCAATTCGCTGTCTGCGAGG-3'	rIHF RT-PCR
IHF3-2	5'-ACGCTCGCCACCAGCGTTTTTCGACCCGGCTTTTT-AAC-3'	
HPRTse	5'-GCTGGTGAAAAGGACCTCT-3'	Kontrolle RT-PCR
HPRTas	5'-CACAGGACTAGAACACCTGC-3'	

2.5 Medien für bakterielle Kulturen

Bakterien wurden nach Sambrook *et al.* (1989) in flüssigem dYT/YT oder auf YT-Agarplatten kultiviert. Zur Selektion von Plasmiden in *E. coli* mit Hilfe von plasmidkodierten Resistenzgenen wurden Antibiotika in den folgenden Konzentrationen zugegeben:

Ampicillin 200 µg/ml, Kanamycin 30 µg/ml, Tetracyclin 10 µg/ml

2.6 Plasmide

Folgende kommerziell erhältliche, von genannten Personen/Firmen zur Verfügung gestellten oder in dieser Arbeit generierte Plasmide wurden eingesetzt:

Expressionsvektoren

pCMV	negativ Kontrolle in eukaryotischen Rekombinationsanalysen (Nicole Christ, Lorbach et al., 2000)
pCMVSSInt	eukaryotische Ausprägung der wild-typ-Integrase (Nicole Christ, Lorbach et al., 2000)
pCMVSSInt-h	eukaryotische Ausprägung der Mutante Int-h (Nicole Christ, Lorbach et al., 2000)
pCMVSSInt-h/218	eukaryotische Ausprägung der Mutante Int-h/218 (Nicole Christ, Lorbach et al., 2000)
pCMVSSInt-CD	eukaryotische Ausprägung der C-Domäne der wild-typ-Integrase (Teresa Corona, diese Arbeit)
pCMVSSInt-h/218-CD	eukaryotische Ausprägung der C-Domäne der Mutante Int-h/218 (Teresa Corona, diese Arbeit)
pRIHF2P	rIHF Expressionsvektor unter Kontrolle des Caggs-Promoters mit Puromycin Resistenzgen zur Generierung stabiler Zelllinien (Teresa Corona, diese Arbeit)

Substratvektoren

pGFP $attB/attP$	Inversionssubstrat zur Bestimmung der <i>Lambda</i> Rekombination zwischen $attB/attP$ (Nicole Christ, Lorbach et al., 2000)
p λ IR	Deletionssubstrat zur Bestimmung der intramolekularen <i>Lambda</i> Rekombination zwischen $attB/attP$ (Nicole Christ, Christ et al., 2002)
p λ ER	Deletionssubstrat zur Bestimmung der intramolekularen <i>Lambda</i> Rekombination zwischen $attL/attR$ (Nicole Christ, Christ et al., 2002)
pCMVSS $attP$ mut	$attP$ -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter und Intron (Teresa Corona, diese Arbeit)
pCMVSS $attB$	$attB$ -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter und Intron (Teresa Corona, diese Arbeit)
pCMVSS $attH2$	$attH2$ -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter und Intron (Teresa Corona, diese Arbeit)
pWS $attB$ GFP	$attB$ -tragendes promoterloses intermolekulares Rekombinations-substrat (Christ et al., 2002)
pWS $attP$ GFP	$attP$ -tragendes promoterloses intermolekulares Rekombinations-substrat (Christ et al., 2002)

pWSP*nsGFP attP*-tragendes promoterloses intermolekulares Rekombinations-substrat (Teresa Corona, diese Arbeit)

Integrations-Vektoren

pCMVSSattPmut attP-tragender *Targetvektor* mit CMV Promoter und Intron für das *Targeting* einer genomischen attB-Region (Teresa Corona, diese Arbeit)

pWSP*nsGFP attP*-tragender *Targetvektor* mit GFP-Reportergen für das *Targeting* der humanen attH2-Region (Teresa Corona, diese Arbeit)

Vektoren zur Generierung stabiler transgener Zelllinien und transgener Mäuse

pTKHatB zur Generierung stabiler Zelllinien, promoterloses neo-Gen hinter attB-Sequenz positioniert, als genomisches Rekombinationssubstrat (Teresa Corona, diese Arbeit)

prIHF2GFP rIHF Expressionsvektor mit Caggs und IRES-GFP Reporter zur Generierung transgener Mäuse (Teresa Corona, diese Arbeit)

prIHF2P rIHF Expressionsvektor mit Caggs-Promoters und Puromycin Resistenzgen zur Generierung stabiler Zelllinien (Teresa Corona, diese Arbeit)

Sonstige Plasmide

pGFP-C1 zur Erfassung der Transfektionseffizienzen (Clontech)

pGEM[®]4Z zur Klonierung von Deletionssubstraten und *Targetvektoren* (Promega)

pGEM[®]7+ zur Klonierung von Deletionssubstraten und *Targetvektoren* (Promega)

pTKHyg zur Ausgangsvektor für die promoterlosen intermolekularen Substratvektoren (Clontech)

pTriEx als Ursprungsvektor für Caggs-Promoter (Novagen)

pPGKSSneo zur Klonierung von prIHF2P (Nicole Christ, Dissertation 2002)

pPGKPuro zur Klonierung von prIHF2P (Schwikardi, unpublizierte Daten)

pTeasy zur Klonierung und Sequenzierung von PCR-Fragmenten (Promega)

p-FRT-IRES/GFP-FRT zur Klonierung von prIHF2GFP (Nathalie Uyttersprot & Stefano Casola, nicht publiziert)

2.7 Medien und Puffer für die eukaryotische Zellkultur

PBS: 137 mM NaCl, 3 mM KCl, 8 mM Na₂HPO₄ (2H₂O), 1,5 mM KH₂PO₄

Einfriermedium: 10 % DMSO, 90 % FCS

Kulturmedium für HeLa:

Dulbecco's Modified Eagle Medium (D-MEM) (GibcoBRL) modifiziert durch Zusatz von: 10 % FCS (PAA), Penicillin (100 U/ml), Streptomycin (100 µg)

Selektionsmedium:

- Kulturmedium für HeLa-Zellen versetzt mit 500-750 µg/ml G418 (GibcoBRL)
- Kulturmedium für HeLa-Zellen versetzt mit 250 µg/ml Hygromycin (Serva)
- Kulturmedium für HeLa-Zellen versetzt mit 5-20 µg/ml Puromycin (Sigma)

Elektroporationsmedium:

RPMI 1640 ohne Phenolrot, ohne Glutamin (GibcoBRL)

2.8 Eukaryotische Zelllinien

HeLa	menschliche cervix carcinoma Zelllinie (Scherer et al., 1953)
293T	menschliche embryonale Nierenzelllinie (Graham et al., 1977), in der die Onkogene Adenovirus E1a und SV40 großes T Antigen ausgeprägt werden
H8B	HeLa Reporterzelllinie, in der das Substrat pGFP attB/attP (Lorbach et al., 2000) in ca. 8 Kopien im Genom an zufälligen Positionen integriert ist (Nicole Christ).

3 Methoden

3.1 Molekularbiologische Methoden

Wenn nicht anders angegeben, wurden in dieser Arbeit molekularbiologische Methoden gemäß Sambrook et al. (1989) oder wie vom jeweiligen Hersteller empfohlen durchgeführt. Beschreibungen von Standard-Methoden wurden zum Teil von Nicole Christ übernommen (Christ, Dissertation 2002).

3.1.1 Agarosegelelektrophorese

Die Auftrennung negativ geladener DNA-Moleküle erfolgte durch die elektrophoretische Auftrennung in einem Agarosegel mit TBE (90 mM Tris-Borat, pH 8,3; 2,5 mM EDTA) als Elektrophoresepuffer. Die DNA-haltigen Proben wurden mit 1/10 Volumen Ladepuffer (0,25 % Bromphenolblau (w/v); 0,25 % Xylencyanol FF (v/v); 30 % Glycerol (v/v); 30 % TBE (v/v); 30 % SDS (v/v)) versetzt. Die Färbung der DNA erfolgte in einem Ethidiumbromid-Bad und wurde durch UV-Licht mit der Wellenlänge 302 nm sichtbar gemacht. Fotografien der gefärbten Agarosegele entstanden mit Hilfe des *Image Master⁰ Systems* der Firma Pharmacia (Freiburg).

Als Größenmarker für die aufgetrennten linearen DNA-Fragmente wurde die 1Kb DNA-Leiter der Firma GibcoBRL oder die 100 Basenpaar-Leiter der Firma Pharmacia (Freiburg) verwendet.

3.1.2 DNA-Extraktion aus Agarosegelen

Die Extraktion von DNA aus Agarosegelen wurde mit dem *Qiaquick[®] Gel Extraction Kit* der Firma Qiagen (Hilden) durchgeführt.

3.1.3 Plasmid-Isolierung aus Bakterien

Die Isolierung von Plasmid-DNA aus Bakterien-Kulturen erfolgte nach dem Prinzip der alkalischen Lyse und anschließender Affinitätschromatographie. Für geringere Mengen (2 ml Übernachtskultur) wurde das *QIAprep[®] Spin Miniprep Kit* der Firma Qiagen (Hilden) verwendet.

Für größere Mengen erfolgte die Isolierung aus 100 ml Bakterien-Kulturen mit

dem *EndoFree[®] Plasmid Maxi Kit*, ebenfalls von der Firma Qiagen (Hilden).

Die DNA-Konzentration wurde bei 260 nm photometrisch bestimmt, wobei die optische Dichte von 1 einer Konzentration von 50 µg/ml entspricht.

3.1.4 DNA-Isolierung aus eukaryotischen Zellen und Schwanzspitzen der Maus

Die Isolierung von DNA aus HeLa-Zellen erfolgte mit Hilfe des *QIAmp[®] DNA Blood Mini Kit* der Firma Qiagen (Hilden).

Zur Gewinnung von genomischer Maus-Schwanz-DNA für die Genotypisierung der transgenen Mäuse, wurde diesen ca. 1cm Schwanzspitze kupiert und diese dann über Nacht bei 56°C mit 600µl Lysis Puffer (100 mM Tris-Cl pH ~8,5, 5 mM EDTA, 0,2 % SDS, 200 mM NaCl, 100-400 µg/ml Proteinase K) inkubiert. Zelltrümmer und Haarreste wurden bei 13.000rpm für 10 min abzentrifugiert. Aus dem Überstand wurde die DNA mit 700 µl 100-%igem Isopropanol gefällt. Zum Waschen wurde die ausgefallene DNA mit einer sauberen Pipettenspitze in ein neues Gefäß mit 1 ml 70-%igem Ethanol überführt und 3 min bei 13.000 rpm zentrifugiert. Dann wurde das Ethanol entfernt und die DNA ca. 15 min bei Raumtemperatur getrocknet. Die Resuspension der DNA erfolgte in 150-400 µl TE. Für *Southern*-Analysen wurden hiervon 40 µl oder 10-15 µg eingesetzt.

3.1.5 Polymerasekettenreaktion (PCR)

Die Polymerasekettenreaktion (Saiki et al., 1988) wurde eingesetzt zum Nachweis von: sequenz-spezifischer Rekombination durch die *Lambda* Integrase in eukaryotischen Zellen, zum Einführen von gezielten Mutationen in Vektoren, zur Herstellung von Proben für die *Southern Blot* Analyse, zur Konstruktion von Substrat- und Expressionsvektoren und zur Genotypisierung von transgenen Zelllinien und Mäusen.

Alle PCR-Reaktionen wurden mit dem *Taq PCR-Core-Kit* (mit *Q-Solution*) der Firma Qiagen (Hilden) durchgeführt. Eine **Standard-PCR-Reaktion** enthielt in der Regel 50 pmol Primer, 50 pmol reversen Primer, 1x Master Mix Q (Taq-Polymerase, dNTPs, Reaktionspuffer und *Q-Solution*), zwischen 100 ng und 1 µg maximal jedoch 5 µl *Template* (Matrize) und destilliertes Wasser ad 50 µl Endvolumen. Die Proben wurden zum Schutz vor Verdampfung mit Paraffin überschichtet und die PCR im Triothermoblock V2.23 der Firma Biometra durchgeführt. In einer Standard-PCR

wurde die DNA in der Regel wie folgt amplifiziert: 30 Zyklen; 45 s bei 94 °C, 45 s bei 58 °C und 1 min / 1 Kb *Template* bei 72 °C. Die PCR-Produkte aus präparativen PCR-Reaktionen, sowie PCR-Reaktionen für den Nachweis von genetischen Veränderungen, wurden elektrophoretisch aufgetrennt, aus den Agarosegelen aufgereinigt und die Integrität durch Sequenzierung überprüft.

Bei der **PCR-Mutagenese** wurden durch die Verwendung von synthetisch hergestellten Oligonucleotid-*Primern* spezifische Nukleotidaustausche in das zu amplifizierende Produkt eingeführt. Hierzu wurde zunächst der DNA-Bereich, in dem die gewünschten Nukleotide ausgetauscht werden sollten, in zwei separaten Standard-PCR-Reaktionen amplifiziert (Abb. 3.1-a). Die Region in der die Nukleotidaustausche eingeführt werden sollten, überlappte sich in beiden PCR-Produkten. In einer abschließenden dritten PCR-Reaktion wurden die zuvor generierten PCR-Fragmente als Matrize eingesetzt (Abb. 3.1-b) und das gesamte zu amplifizierende Produkt einschließlich der spezifisch eingeführten Nukleotide vervielfältigt (Abb. 3.1-c).

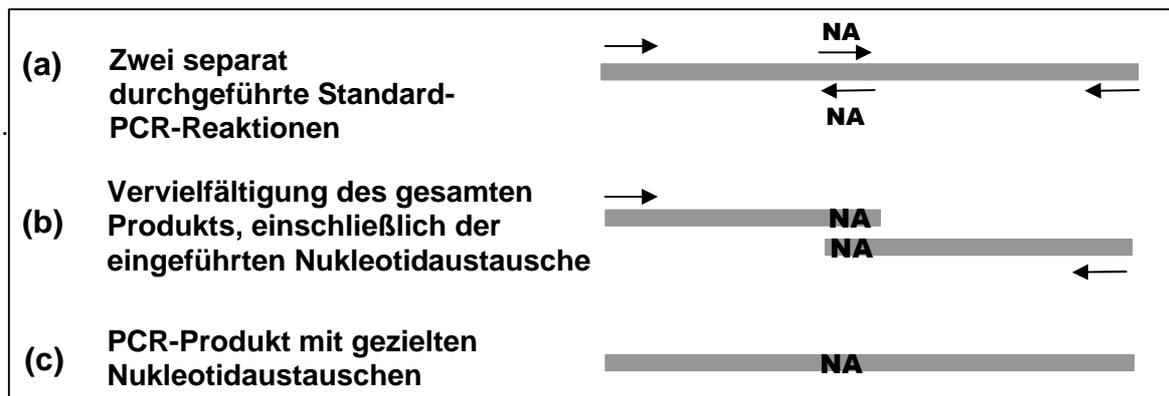


Abbildung 3-1: Schematische Darstellung der PCR-Mutagenese. (nach Christ 2002, Dissertation). Der als Matrize dienende DNA-Bereich ist als graues Rechteck dargestellt. Pfeile (®) stellen die synthetisch hergestellten Oligonucleotid-*Primer* dar; die mit **NA** markierten *Primer* und DNA-Bereiche enthalten die gewünschten Nukleotidaustausche (Christ 2002, Dissertation, Details sind im Text angegeben).

3.1.6 RNA-Isolierung und RT-PCR-Analyse der rHF-mRNA-Ausprägung in Mausorganen und HeLa-Zellen

Zur Analyse der Transkription von rHF in transgenen Mäusen wurde zunächst die RNA mit Hilfe des *Rneasy-Mini-Kits* der Firma Qiagen (Hilden) nach Herstellerprotokoll isoliert. Die zu analysierenden Proben wurden zu diesem Zweck zuvor mit einem *Qiashreder* (Qiagen, Hilden) homogenisiert. Geerntete HeLa-Zellen konnten direkt auf den *Qiashreder* gegeben werden. Schockgefrorene oder in *RNAlater*

(Qiagen, Hilden) aufbewahrte Mausorgane wurden erst im Lysispuffer des *Rneasy-Mini-Kits* mit einem Micropistill (Eppendorf) vorhomogenisiert. Die RNA wurde anschließend mit *RNase-Free DNase* (Promega) behandelt, um eventuell vorhandene DNA-Reste zu entfernen. Die mRNA wurde anschließend mit dem *SuperScript™ First-Strand Synthesis System for RT-PCR* der Firma Invitrogen nach Protokoll revers transkribiert unter Benutzung der Oligo-d(T)₁₂₋₁₈-Primer. Es wurden in der Regel maximal 5 µg DNase behandelte RNA eingesetzt.

Der Nachweis der reversen rIHF-Transkripte erfolgte in einer Standard-PCR mit den Intron-übergreifenden Primern (SS5') und (IHF3-2) (siehe Kapitel 2.4) und den zuvor synthetisierten komplementären DNA-Strängen als Matrize. Dabei wurde 1/10 des cDNA-Ansatzes als *Template* eingesetzt. Zur Kontrolle ob in den mit DNase I behandelten RNA-Präparationen noch Reste von DNA vorhanden sind, wurde parallel eine Standard-PCR mit den zuvor genannten Primern durchgeführt. Zur Analyse der PCR-Produkte wurden diese abschließend in einem Agarosegel elektrophoretisch aufgetrennt. Als positive Kontrolle für die erfolgreiche RNA-Präparation und c-DNA-Synthese wurde ein Teil der mRNA des HPRT (Hypoxanthin-Phosphorybosyl-Transferase)-Gens mit den Primern HPRTse und HPRTas amplifiziert. Diese Primer ergeben bei einer Amplifikationszeit von 30 sek nur bei Anwesenheit der c-DNA eines korrekten Spleißproduktes ein ca. 250 bp großes Fragment. Zur Analyse der PCR-Produkte wurden diese anschließend in einem Agarosegel elektrophoretisch aufgetrennt.

3.1.7 DNA-Sequenzierung

DNA-Sequenzanalyse von Plasmiden oder PCR-Produkten wurde nach der auf Didesoxy-Kettenabbruch-Reaktion beruhenden Methode von Sanger et al., 1977, durchgeführt. Verwendet wurde das *BigDye™ Terminator Cycle Sequencing Kit* (Applied Biosystems), das auf den Einbau von Fluoreszenz-Farbstoff-markierten Didesoxynukleotiden beruht. Markierte DNA wurde mit einem automatischen DNA-Sequencer (ABIPrism™377, Applied Biosystems) aufgetrennt und analysiert. Der DNA-Sequencer wurde von Karin Otto bedient.

3.1.8 Southern Blot Analyse

Genomische *Southern Blot* Analyse (Sambrock et al., 1989) wurde zur Genotypisierung und Charakterisierung von stabilen HeLa-Zelllinien und transgenen

Mäusen und zum Nachweis von Integrase-vermittelter genomischer Rekombination in HeLa-Zellen eingesetzt.

10-20 µg genomische DNA wurden über Nacht mit 100 U des jeweils angegebenen Restriktionsenzym geschnitten und dann in einem 0,8-%igem Agarosegel elektrophoretisch aufgetrennt. Das Gel wurde mit Ethidium-Bromid gefärbt, unter UV-Licht fotografiert, 15 min in 0,25 M HCl depuriniert, mit H₂O gewaschen und anschließend 60 min im Transfer Puffer (0,4 M NaOH, 1 M NaCl) denaturiert. Die DNA wurde über Nacht durch kapillaren Transfer auf eine „Genescreenplus-Nylon-Membran“ (NEN, Belgien) übertragen, in 0,5 M Tris-HCl / 1 M NaCl 30 min neutralisiert und 1h bei 80 °C gebacken. Die Membran wurde vor der Hybridisierung mindestens 3 h in Hybridisierungslösung (50 mM Tris pH 7,5; 1 M NaCl; 1 % SDS; 10 % Dextran Sulfat; 0,5 mg/ml denaturierte Heringssperma-DNA) bei 65 °C in einem Biometra® CompactLineOV4 Ofen prähybridisiert.

Für die Charakterisierung der *attB*-neo stabilen HeLa-Reporterzelllinien wurde eine interne Probe gewählt, die aus dem gesamten neo-Gen bestand. Der Restriktionsverdau der genomischen DNA erfolgte mit HindIII.

Die Genotypisierung der rHF-transgenen Mäuse und der rHF-stabilen Zelllinien wurde mit einer internen Probe, die das gesamte rHF-Gen umfasst, durchgeführt. Die Maus-DNA wurde mit Bgl II und die HeLa-DNA mit BamHI verdaut.

Der Nachweis von Int-vermittelter Rekombination in stabilen HeLa-Reporterzelllinien (H₈B) erfolgte mit einer internen Probe, die das gesamte eGFP-Gen umfasste. Die genomische DNA wurde mit NcoI verdaut.

Alle Proben wurden durch PCR generiert. Die radioaktive Markierung erfolgte durch *Random-primed Labelling* mit Hilfe der thermostabilen Bca-Polymerase (TaKaRa) bei 56 °C. Nicht-eingebaute Radionukleotide wurden mit S-200HR Säulen (Pharmacia) abgetrennt und die Proben 5 min bei 95 °C denaturiert, bevor sie zur Hybridisierungslösung gegeben wurden. Die Hybridisierung erfolgte in allen Experimenten bei 65 °C über Nacht. Nach der Hybridisierung wurde die Membran in allen Experimenten zweimal in 2x SSC / 0,1 % SDS bei 65 °C gewaschen. Je nach gewünschter Stringenz wurden weitere Waschschrte mit abfallender Salzkonzentration durchgeführt. Die Membran wurde in Plastikfolie eingeschweißt und die Signale durch Autoradiographie (Biomax-MS Filme, Kodak) sichtbar gemacht oder mit dem Phosphoimager-System (Fuji Bas 1000, Raytest, Straubenhardt; Software: TINA 2.09, Raytest, Straubenhardt) quantifiziert.

3.1.9 Konstruktion und Beschreibung der eukaryotischen Expressionsvektoren

Zur Ausprägung der Integrasen und des Kofaktors rIHF2His in Eukaryoten wurden Vektoren eingesetzt, auf denen die Expression der Gene durch den Cytomegalovirus-Promoter (CMV; Boshart et al., 1985) bzw. durch den sogenannten *Caggs*-Promoter reguliert wurde. Beim *Caggs*-Promoter handelt es sich um einen hybriden Promoter, der aus dem CMV *immediate early enhancer* fusioniert an den β -Aktin-Promoter aus Huhn besteht (Niwa et al., 1991). **pCMV** und die eukaryotischen Expressionsvektoren für die Integrasen **pCMVSSInt**, **pCMVSSInt-h** und **pCMVSSInt-h/218** (Lorbach et al., 2000) wurden freundlicherweise von Nicole Christ aus unserem Labor zur Verfügung gestellt. Es handelt sich dabei um Derivate vom pPGKCrebpa (Fellenberg, 1998) der neben dem PGK Promoter auch die Polyadenylierungs-Sequenz des Rinder-Wachstumshormons (*bovine growth hormone*; Pfarr et al., 1986; Abb. 3.2) trägt. Zur Steigerung der Transkription wurde das pPGKCrebpa dahingehend modifiziert, dass zwischen die Schnittstellen NheI und PstI ein heterologes Intron (Choi et al., 1991) eingesetzt wurde. Der Phosphoglycerat-Kinase I (PGK) Promoter, zwischen der EcoRI- und NheI-Restriktionsschnittstelle, wurde durch ein CMV-Promoter-Fragment ersetzt, das mittels PCR aus pGFP-C1 (Clontech) amplifiziert wurde. Die bakterielle Selektion dieser auf einem pUC-Replikationsursprung basierenden Expressions-Plasmide erfolgte mit Ampicillin.

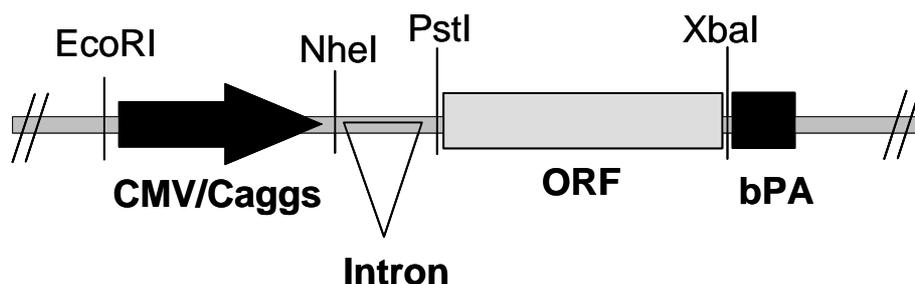


Abbildung 3-2: Schematische Darstellung eukaryotischer Expressionsvektoren. Hervorgehoben sind die genetisch relevanten Elemente der Expressionsvektoren: CMV, Cytomegalovirus Promoter; Caggs-Promoter, CMV *early enhancer* fusioniert an den β -Aktin-Promoter aus Huhn; Intron, heterologes Intron (Choi et al., 1991); ORF, *open reading frame*; bpA, Polyadenylierungs-Sequenz. Die Expression des Gens wird entweder vom CMV oder vom Caggs Promoter reguliert. Das Spleißsubstrat zwischen Promoter und ORF und die Polyadenylierungs-Sequenz vom Rinder-Wachstumshormon sollen den Transport der mRNA vom Kern ins Cytosol fördern (Pfarr et al., 1986; Choi et al., 1991). Gezeigt sind außerdem die für die Klonierung von Derivaten relevanten Schnittstellen EcoRI, NheI, PstI und XbaI. (modifiziert nach Christ, 2002 Dissertation).

Im Folgenden sind die Klonierungen weiterer Derivate beschrieben. Die Nukleotidsequenzen der aufgeführten Primer sind in 2.4 aufgeführt.

pCMVSSInt-CD und **pCMVSSInt-h/218-CD** (siehe Anhang) sind Expressionsvektoren für die C-Domäne der wild-typ- λ -Integrase bzw. ihrer Mutante h/218. Die C-Domänen unterscheiden sich von den kompletten Proteinen durch eine Deletion von 64 Basenpaaren am N-Terminus. Zum Schutz vor Abbau der resultierenden Proteine wurde am 3'-Ende des jeweiligen ATG-Startcodons ein Glycinrest kloniert. Die kodierenden Sequenz der C-Domänen wurden mittels PCR mit den Primern Int-N-C-Dom-PstI und Int-C-Eu generiert und in die PstI / BamHI Schnittstelle des Expressionsvektor pCMVSSInt bzw. pCMVSSInt-h/218 kloniert, wobei das Vollängen-Protein jeweils ersetzt wurde.

prIHF2P (siehe Anhang) ist ein Expressionsvektor für den rekombinanten *Integration Host Factor*, der eine 6-fache Histidin-Extension am N-Terminus besitzt (rIHF, Kap. 4.1.3.1, in der Dissertation von Nicole Christ 2002 noch als oIHF2 bezeichnet). Dieser modifizierte Kofaktor der Integrase wurde von Nicole Christ in unserem Labor kloniert und ist ein Hybrid aus den beiden IHF-Untereinheiten IHF- α und IHF- β , in dem die Untereinheit α (AS L3 bis A94) zwischen die AS Q39 und G40 der Untereinheit β inseriert ist (Kap. 4.1.3.1, Abb. 4-5D). Die Übergänge zwischen der Untereinheit β und der integrierten Untereinheit α wurden durch Aminosäure-*Linker* verbunden (Kap. 4.1.3.1, Abb. 4-5E; Nukleotidsequenz). An der zweiten AS-Position wurde zusätzlich ein Glycinrest eingefügt (Kap. 4.1.3.1, Abb. 4-5E) (Corona et al., 2003). Im prIHF2P Vektor befindet sich das Gen für das rIHF2His unter der Kontrolle des *Caggs*-Promoters (s.o., Niwa et al., 1991). Dieser wurde aus dem pTriEx Vektor (Novagen) mit den Primern CAGGS-5-EcoRI und CAGGS-3-NheI amplifiziert. Das SS-rIHF-His-bpA-Fragment wurde durch PCR mit den Primern SS-5-NheI und bPARP-BamHI aus dem pCMVSSoIHF2His Vektor (Nicole Christ, 2002) gewonnen. Die Enden der Fragmente wurden mit EcoRI / NheI bzw. NheI / BamHI verdaut und in den EcoRI / BamHI geschnittenen pGEM7+ Vektor (Promega) kloniert. Als Selektionsmarker für die Generierung stabiler Ziellinien wurde ein Puromycin Resistenzgen unter der Kontrolle des PGK- (Phosphoglycerat Kinase) Promoter gewählt. Das PGK-SS-Puromycin-bpA-Fragment wurde in die SapI Schnittstelle des resultierenden Vektors kloniert nachdem es mit EcoRI / NotI aus pPGKSSPuro gewonnen wurde. pPGKSSPuro wurde generiert, indem in die Pst I / Xba I Schnittstelle des pPGKSSneo-Vektor (Nicole Christ, Dissertation 2002)

das Puromycin-Resistenzgen aus pPGKPuro (Schwikardi, unpublizierte Daten) kloniert wurde.

3.1.10 Konstruktion der Substratvektoren

Intramolekulare Substratvektoren

Die Deletionssubstratvektoren **p λ IR** und **p λ ER** sind Derivate des pGEM[®]4Z und wurden von Nicole Christ zur Verfügung gestellt. Auf ihnen sind die *att*-Regionen in direkter Wiederholung angeordnet. Die Int-vermittelte sequenz-spezifische Rekombination führt dazu, dass das dazwischenliegende Fragment, einschließlich eines Transkriptionsstopps, ausgeschnitten wird und es zur Expression des Reportergens eGFP (*enhanced green fluorescent protein*) kommt. Auf **p λ IR**, dem Substrat für die Integrative Rekombination, liegen *attB* und *attP*. **p λ ER** wurde wie p λ IR generiert mit dem Unterschied, dass *attL* anstelle von *attB* und statt der *attP*-Region die *attR*-Region eingefügt wurden.

Intermolekulare Substratvektoren

Die Substratvektoren mit dem CMV-Promoter und einem heterologen Intron (Choi et al., 1991) vor einer Rekombinationsregion sind im Folgenden beschrieben:

pCMV*attP*mut: (siehe Anhang) *attP*mut enthält drei G→C Austausche im P-Arm, die durch PCR-Mutagenese, wie unter 3.1.11 beschrieben, erzeugt wurden. Diese Austausche wurden eingeführt, um vor dem eGFP-Gen bzw. dem Neo-Resistenzgen ATG- Startcodons zu entfernen, die nach erfolgter Rekombination eine ordnungsgemäße Translation beeinträchtigen könnten. Diese Mutationen betreffen ausschließlich XIS bzw. FIS Bindungsregionen im P-Arm und sollten auf die integrative Rekombination keinen Einfluss haben. Sie wurden eingeführt, indem zwei überlappende PCR-Produkte generiert wurden. Das erste mit dem Primerpaar *attP*-ATC-1 / *attP*2 und das zweite mit dem Primerpaar *attP*-ATC-3 / *attP*4. Als Matrize diente pGFP*attB/attP* (Lorbach et al., 2000). Anschließend wurden die beiden PCR-Produkte aufgereinigt und als Template in einer dritten PCR mit dem Primerpaar *attP*-PstI / *attP*-XbaI eingesetzt. Das resultierende, *attP* tragende PCR-Fragment wurde mit PstI / XbaI geschnitten und in den pCMVSSInt Vektor (siehe Kap. 3.1.9) kloniert. Der Vektor pCMV*attP*mut wurde auch als *Targetvektor* für *attB*-stabile HeLa-Zelllinien eingesetzt.

pCMV*attB* (siehe Anhang) wurde konstruiert, indem *attP*mut zwischen der PstI-

und XbaI-Schnittstelle im pCMV*attP*mut durch ein doppelsträngiges *attB*-Oligonucleotid (*attB*-PstI-XbaI-1 / *attB*-PstI-XbaI-2) ersetzt wurde (Christ et al., 2002).

pCMV*attH2* (siehe Anhang) ist ebenfalls ein Derivat von pCMV*attP*mut in dem *attP*mut jedoch diesmal durch ein doppelsträngiges *attH2*-Oligonucleotid (*attH2*-PstI-XbaI-S und *attH2*-PstI-XbaI-AS) ersetzt wurde.

Im Folgenden sind Substratvektoren beschrieben, auf denen sich ein Transkriptions-Stoppsignal gefolgt von einer *att*-Region befindet. Das 3'-Ende der *att*-Region grenzt jeweils an ein promoterloses eGFP-Reportergen:

pWS*attB*GFP wurde von Nicole Christ konstruiert (Christ et al., 2002) und ist ein *attB*-tragendes promoterloses intermolekulares Rekombinationssubstrat. Die Rekombination mit einem CMV-Promoter tragendem Substratvektor führt zur Positionierung des Promoters am 5'-Ende des eGFP-Gens und damit zu dessen Expression. Der Transkriptionsstop am 5'-Ende soll bei illegitimer Rekombination dagegen eine Expression des Reporters verhindern.

pWS*attP*GFP ist ein Derivat von pWS*attB*GFP (Christ et al., 2002), auf dem *attB* durch *attP* ausgetauscht wurde. *attP* wurde durch PCR vom λ IR vervielfältigt und zwischen die BamHI- und NotI-Schnittstelle kloniert.

pWSP*_{ns}GFP (siehe Anhang) ist ebenfalls ein Derivat von pWS*attB*GFP. Hier wurde *attB* durch *attP**_{ns} ausgetauscht, das vorher durch PCR-Mutagenese mit den Primern *attP**_{ns}-AS / *attP*2 (1.PCR), *attP**_{ns}-S / *attP*4 (2.PCR) und *attP**_{ns}-Not / *attP*-Pst (3.PCR) generiert wurde. Als Ausgangsmatrize diente die wild-typ-*attP*-Sequenz im pGFP*attB/attP* Vektor (Lorbach et al., 2000). Das Pst I / Not I geschnittene Fragment wurde in einer drei-Fragmente-Ligation mit den 4148 bp großen Not I / EcoR I und dem 1373 bp großen EcoR I / Pst I zu pWSP*_{ns}GFP zusammengefügt. pWSP*_{ns}GFP wurde sowohl in episomalen Rekombinationsanalysen als auch als Integrations-Vektor für das *Targeting* des genomischen *attH2*-Lokus in HeLa-Zellen eingesetzt. Um die Homologie der *overlap*-Region zu bewahren und den Strangaustausch zu erleichtern, wurde das erste T im *overlap* von *attP* gegen ein C ausgetauscht und *attP** generiert. Die für die Bindung der Integrase wichtigen Nukleotide der rechten *core*-Sequenz wurden nicht verändert. Da für die Expression eines Fusionsproteins kein Stoppcodon hinter der *overlap*-Region folgen darf, wurden sie in der ersten Position der rechten *core*-Bindungsstelle (Abb. 4-19) und im gesamten P'-Arm mutiert.

Die Integrität aller relevanten genetischen Elemente auf den Substratvektoren

wurde durch DNA-Sequenzierung überprüft.

3.1.11 Konstruktion der Integrations-Vektoren (*Targeting* genomischer Loci)

pCMVattPmut und **pWSP*_{ns}GFP** (siehe Anhang) werden in dieser Arbeit als Integrations- oder auch *Target*vektoren bezeichnet. Zur Konstruktion siehe in Kapitel 3.1.10 unter intermolekulare Substratvektoren.

3.1.12 Konstruktion der Vektoren zur Generierung transgener Zelllinien und Mäuse

Mit **pTKHattB** (siehe Anhang) wurde eine wild-typ-*attB*-Sequenz gefolgt von einem promoterlosen Resistenzgen stabil in zufällige Positionen des HeLa-Genoms integriert. pTKHattB wurde kloniert, indem das Neo-Gen mit Polyadenylierungs-Sequenz (PA) aus dem p2TVneo⁺ Vektor (von Nicole Christ und Peter Dröge zur Verfügung gestellt) amplifiziert wurde. Hierfür wurde am 5'-Ende das *attB*-enthaltende Oligonukleotid attB-MfeI-5' und am 3'-Ende der Primer bPA-HindIII eingesetzt. Das resultierende *attB*-neo-PA-Fragment wurde in den Mfe I / Hind III geschnittenen pTKHyg Vektor (Clontech) kloniert. pTKHyg enthält ein Hygromycin-Resistenzgen, das über einen HSV-Thymidinkinase Promoter ausgeprägt wird und für die Selektion stabiler Zelllinien eingesetzt wird.

prIHF2P (siehe Anhang) wurde zur Generierung von rIHF transgenen Zelllinien benutzt. Zur Klonierung siehe in Kapitel 3.1.9 Konstruktion und Beschreibung der eukaryotischen Expressionsvektoren.

prIHF2GFP (siehe Anhang) wurde zur Generierung von rIHF transgenen Mäusen benutzt. Wie schon in prIHF2P erfolgt die Ausprägung des rIHF-Gens über den Caggs-Promoter, der in Mäusen eine ubiquitäre Expression vermittelt (Okabe et al., 1997). Zwischen dem 3'-Ende des Gens und der Polyadenylierungs-Sequenz wurde eine von FRT-Sequenzen eingerahmte IRES-eGFP-Reporter-kassette eingeführt. Die IRES (*internal ribosomal entry site*, Clontech; Jackson et al., 1990; Jang et al., 1988) ermöglicht eine zeitgleiche Expression des eGFP-Gens mit rIHF. Die Kassette soll später die Identifizierung transgener, rIHF-ausprägender Tiere erleichtern und könnte, wenn experimentell erforderlich, mit der FLP-Rekombinase über die FRT-Sequenzen deletiert werden. Das 2092 bp große EcoR I Fragment des prIHF2P Vektors enthält die Caggs-Intron-rIHFHis-Sequenz. Für die Konstruktion von prIHF2GFP wurde im ersten Schritt dieses EcoR I Fragment kurz vor die erste FRT-

Sequenz des pFRT-IRES/GFP-FRT Vektors (von Nathalie Uyttersprot zur Verfügung gestellt) kloniert. Im einem zweiten Schritt wurde die Polyadenylierungs-Sequenz (PA) mit den Primern 5'PA-Not I und 3'RP-PGK-Not I aus prIHF2P amplifiziert und hinter die zweite FRT-Sequenz in die Not I Schnittstelle des im ersten Schritt generierten Zwischenproduktes eingesetzt. Bei der Generierung der transgenen Mäuse seitens der Memorec Stoffel GmbH wurde für die Pronucleus-Injektion das 3468 bp große Mlu I / Esp3 I Fragment von prIHF2GFP benutzt.

3.2 Biochemische Methoden

3.2.1 SDS-Polyacrylamid-Gelelektrophorese

Zur Analyse der Integrase und der rIHF-Proteine in HeLa-Zellen oder in Zellen aus Maus-Organen wurden zunächst die Zell-Lysate in SDS-Polyacrylamid-Gelen (Laemmli, 1977) elektrophoretisch aufgetrennt. Die Acrylamid/Bisacrylamid-Konzentration im Sammelgel betrug in der Regel 5 % und sowohl die Auftrennung von Integrase-Proteinen, als auch die Auftrennung von rIHF-Proteinen erfolgte anschließend in einem 12,5 % Trenngel. Nach einer Einlaufzeit der Proteine im Sammelgel bei 15 mA, erfolgte Elektrophorese bei 30 mA.

Die Zell-Lysate aus HeLa-Zellen wurden hergestellt, indem transfizierte Zellen oder stabil transfizierte Zelllinien geerntet, gezählt, in einem definierten Volumen von 3x SDS *Sample Buffer* und 3x DTT (New England Biolabs) aufgenommen und 5 min bei 95 °C erhitzt wurden. In der Regel wurde dann das Lysat von 8×10^5 Zellen für die SDS-PAGE eingesetzt.

Um Rohlysate aus Mausorganen zu isolieren wurden die schockgefrorenen oder in *RNAlater* (Qiagen, Hilden) aufbewahrten Organe mit einem Micropistill in eiskaltem RIPA Puffer (1xPBS; 1 % Igepal CA-630; 0,5 % sodium deoxycholate; 0,1 % SDS) (Eppendorf) zerkleinert. Der Puffer wurde kurz vor der Lyse mit Proteaseinhibitoren (10 mg/ml PMSF; 100 mM Natrium-Orthovanadat; 20 µg/ml Leupeptin; 20 µg/ml Aprotinin; 10 µg/ml α -Antitrypsin) versetzt. Die Proben wurden zweimal bei 13.000 rpm, 4 °C zentrifugiert und jeweils der klare Überstand (Rohlysate) abgenommen. Die Proteinbestimmung erfolgte photometrisch nach der Warburgformel. Für die SDS-Page wurden zwischen 40 und 180 µg Protein in 3x SDS *Sample Buffer* und 3x DTT (New England Biolabs) aufgenommen und vor dem Auftrag 5 min bei 95 °C erhitzt.

Nach der Elektrophorese wurden die Gele entweder mit Coomassie-Blau (Sambrock et al., 1989) gefärbt oder für die anschließende *Western Blot* Analyse in einer Naßblotkammer (Bio-Rad) mit 14 V (Transferpuffer: Glycin/Tris) auf PVDF-Membranen (ImmobilonP, Millipore) übertragen. Das Molekulargewicht der Coomassie-Blau gefärbten Proteine einzelner Banden wurde durch den Vergleich mit dem Größenstandard *Protein Marker Broad Range* (New England Biolabs) abgeschätzt und mit dem *Image Master[®]* System dokumentiert.

3.2.2 *Western Blot* Analyse

Zum Nachweis von Integrase-Proteinen in eukaryotischen Zellen durch *Western Blot* Analyse wurden Zell-Lysate von 8×10^5 transient transfizierten HeLa-Zellen durch fünfminütiges Kochen in SDS *Sample buffer* (New England Biolabs) erzeugt. Die Proteine wurden im 12,5 % (w/v) Polyacrylamidgel aufgetrennt und auf eine PVDF-Membran transferiert. Die PVDF-Membran mit den übertragenen Proteinen aus dem Polyacrylamidgel wurde mit 1 % *blocking solution* (*Western Blot Kit* der Firma Boehringer, Mannheim/Roche) gesättigt und dann zunächst in 0,5 % *blocking solution* mit 1:50000 (α -Integrase, polyklonaler Antikörper gegen *Lambda*-Integrase, A. Landy, Providence USA) 1 h inkubiert. Die Membran wurde nach Angaben des Herstellers gewaschen und der gebundene primäre Antikörper mit dem zweiten Antikörper α -Maus / α -Kaninchen IgG-POD (Meerretich-Peroxidase) gebunden. Der Nachweis der Integrase-Proteine auf der Membran erfolgte über eine Chemolumineszenz-Reaktion, ebenfalls nach Angaben des Herstellers.

In *Western Blot* Analysen, bei denen die in HeLa-Zellen transient ausgeprägte Int-Menge über einen Zeitraum von 72 h untersucht wurde, wurden nach 24 h, 48 h und 72 h Zell-Lysate von 8×10^5 Zellen erzeugt und analysiert.

In *Western Blot* Analysen, bei denen die Langzeitexpression von rIHF in den stabilen Zelllinien untersucht werden sollte, wurden die Zellen über die genannten Zeiträume (4 und 8 Monate nach der Generierung) ständig in Kultur gehalten und die Lysate zu den entsprechenden Zeitpunkten hergestellt (s.o.).

Der Nachweis von rIHF-Proteinen in Lysaten von transient oder stabil transfizierten HeLa-Zellen erfolgte wie für die Integrase beschrieben, mit dem Unterschied, dass der primäre Antikörper (Anti-Serum gegen IHF aus Kaninchen, S.D. Goodman, Los Angeles) 1:1000 in 0,5 % *blocking solution* verdünnt wurde.

Als positive Kontrolle wurde ~3-10 ng gereinigtes rIHF-Protein (Peter Dröge)

eingesetzt.

Um die Beladung der Spuren untereinander zu prüfen, wurden die Membranen nach Herstellerprotokoll gestrippt und nochmals mit anti- α -Aktin 4700 (Sigma) in einer 1:5000 Verdünnung als 1. Antikörper inkubiert.

3.3 Zellbiologische Methoden

Wenn nicht anders angegeben, wurden in dieser Arbeit zellbiologische Methoden gemäß Spector et al., 1998 oder wie vom jeweiligen Hersteller empfohlen durchgeführt.

3.3.1 Kultur von Säugetierzellen

Die in dieser Arbeit verwendeten adhärenenten Säugetierzelllinien wurden in DMEM bei 37 °C und 7,5 % CO₂ kultiviert. Vor Erreichen der Konfluenz, bei einer Zelldichte von ca. 80 %, wurden die Zellen mit Trypsin geerntet und 1:4 mit frischem Medium verdünnt auf neue Kulturschalen gegeben. Eingefroren wurden alle Säugetierzellen in 90 % FCS / 10 % DMSO bei -80 °C und nach einigen Tagen zur Langzeit-Aufbewahrung in flüssigen Stickstoff überführt.

3.3.2 Konstruktion stabiler HeLa-Zelllinien

Für die Generierung von HeLa-Zelllinien mit stabil im Genom integriertem pTKHattB oder prLHF2P wurden 40 μ g des jeweiligen Vektors über Elektroporation (250 V, 960 μ F) mit Hilfe des Gene Pulser (Bio-Rad) in ca. 8×10^6 Zellen eingeschleust. Stabile Zellen wurden in Kulturmedium für HeLa-Zellen, versetzt mit 250 μ g/ml Hygromycin (Serva, Heidelberg) bzw. 20 μ g/ml Puromycin (Sigma), selektioniert. Überlebende Einzelklone wurden durch picken vereinzelt und dann expandiert. Die Charakterisierung der Zelllinien erfolgte durch PCR, DNA-Sequenzierung und genomische *Southern Blot* Analyse.

3.3.3 Durchflusszytometrie

Zur Analyse der eGFP-Ausprägung wurden HeLa-Zellen geerntet, in Einzelzellsuspension gebracht und mit einem FACScan (Becton Dickinson) analysiert oder im FACStar+ (Becton Dickinson) sortiert. Der FACStar+ wurde von

Christoph Göttlinger bedient. Im FACScan wurden tote Zellen mit Hilfe von 7-Amino-Actinomycin D (0,1 µg/ml, Sigma) während der Analyse ausgeschlossen. Im FACStar+ wurden die toten Zellen mit Propidiumiodid (1 µg/ml, Sigma) angefärbt. Die FACS-Daten wurden mit Hilfe der Software CellQuest™ (Becton Dickinson) ausgewertet und entweder als einparametrisches Histogramm oder als zweiparametrisches *Dot Plot* Diagramm dargestellt.

3.3.4 Immunofluoreszenz

Zur Lokalisation des rIHF in den stabil transfizierten HeLa-Zelllinien wurden die Zellen auf Deckgläsern bis zu einer Dichte von 50-70 % kultiviert. Die Präparate wurden mit PBS gewaschen und in 4 % Formaldehyd / PBS für 15 min fixiert. Nach der Fixierung wurden sie erneut gewaschen, für 10 min mit 50 mM Ammoniumchlorid / PBS inkubiert, gewaschen und dann permeabilisiert. Hierfür war eine 5-minütige Behandlung mit 0,4 % TritonX-100 / PBS erforderlich. Die Blockierung erfolgte 30 min in 1 % Milchpulver / 0,5 % Saponin / PBS gefolgt von einer 1,5-stündigen Inkubation mit der 1. Antikörperlösung (α-IHF (1:1000) / 1 % Milchpulver / 0,5 % Saponin / PBS). Vor und nach der Behandlung mit der 2. Antikörperlösung (α-Kaninchen-IgG-Fluorescein-Isothiocyanat (FITC, Sigma) / 1 % Milchpulver / 0,5 % Saponin / PBS, 1 Stunde) wurden je vier Waschschrte mit 0,5 % Saponin / PBS durchgeführt. Für die Kernfärbung wurde 4',6-Diamidino-2-Phenylindol (DAPI) zu einer Endkonzentration von 2 µg/ml der 2. Antikörperlösung zugegeben.

Die Präparate wurden dann auf Objektträger überführt. Dabei wurde das Deckglas mit der Zellseite nach unten auf einen Tropfen *SlowFadeLight* Einbettmedium (Molecular Probes, Leiden) gleiten gelassen.

Fluoreszenz-Aufnahmen wurden mit einem *Zeiss laser scan microscope* (LSM4) durchgeführt.

3.4 Rekombinationsanalysen

3.4.1 Episomale Rekombinationsanalysen in eukaryotischen Zellen

Zur Untersuchung von Rekombinationsaktivitäten der Integrasen auf episomalen Substraten in eukaryotischen Zellen, wurden diese mit verschiedenen kombinierten Expressions- und Substratvektoren kotransfiziert. Die Vektoren wurden durch Elektroporation in die Zellen geschleust. HeLa und stabil transfizierte, rIHF exprimierende Zelllinien wurden zweimal mit PBS gewaschen bevor sie in 700 µl Elektroporationsmedium (siehe Kap. 2.7) aufgenommen wurden. In der Regel wurden insgesamt 60 µg Expressions- und Substratvektor in einem molaren Verhältnis von 1:1 bei 300 V und 960 µF (*Gene pulser*, Bio-Rad) in ungefähr 1×10^7 Zellen transfiziert. Nach der Elektroporation wurden die Zellen aus jedem Ansatz auf unterschiedliche 10 cm Kulturschalen plattiert und zu den in den einzelnen Experimenten angegebenen Zeitpunkten (zwischen 24 und 72 h) durchflusszytometrisch untersucht (Siehe Kap. 3.3.2). Die Integrität aller Produkte wurde durch Sequenzierung überprüft.

Die Transfektionseffizienzen wurden in der Regel durch die Elektroporation von pGFP-C1 in die jeweiligen Zellen und einer anschließenden FACS-Analyse nach 48 h bestimmt und lagen in der Regel über 90 %.

3.4.2 Chromosomale Rekombinationsanalysen

3.4.2.1 Genomische Inversion in H8B-Zelllinien

HeLa-Reporterzellen, die pGFP*attB/attP* stabil im Genom integriert haben, wurden, wenn nicht anders angegeben, mit je 30 µg eines Expressionsvektors für die Integrase und für den Kofaktor rIHF transfiziert. Dazu wurde zirkuläre DNA in etwa 1×10^7 Zellen / 700 µl durch Elektroporation bei 250 V und 960 µF (Bio-Rad *gene pulser*) in die Reporterzelllinien eingeschleust. In einigen Fällen wurde erst 60 µg des Expressionsvektor für rIHF in die Zellen vortransfiziert und die Transfektion mit dem Expressionsvektor für die Integrase bzw. einer ihrer Mutanten erfolgte nach 24 h. Nachdem die Zellen nach der ersten Transfektion mittels Elektroporation in geeigneter Verdünnung (ca. 1×10^5 Zellen) auf 6-well-Platten ausgesät wurden, erfolgte die zweite Transfektion durch Lipofektion (FuGENE™6, Roche) nach Angaben des Herstellers.

Der Nachweis von Integrase-vermittelter genomischer Inversion in den HeLa-Reporterzellen **H8B**, die **pGFPattB/attP** stabil im Genom integriert haben, erfolgte, wenn nicht anders angegeben, 72 h nach der Transfektion durch *Southern Blot* Analyse. Als Sonde wurde das komplette eGFP-Gen eingesetzt. Die Transfektionseffizienz wurde in jedem Experiment in einem separaten Ansatz mit pGFP-C1 ermittelt und betrug im Schnitt etwa 90 % bei der Elektroporation und 50 % bei den Transfektionen mit Fugene. Die eingesetzte Menge von pGFP-C1 richtete sich jeweils nach der Menge des zuvor benutzten Expressionsvektors.

Eine Int-vermittelte alternative Verknüpfung zwischen den *att*-Sequenzen (BOP statt BOP', Christ & Dröge, 1999), die zur Deletion der *att*-flankierten DNA-Sequenzen führen würde, wurde in früheren Experimenten bereits von Nicole Christ ausgeschlossen (Christ, 2002 Dissertation).

3.4.2.2 Genomisches *Targeting attB*-stabiler HeLa-Zelllinien

Für die *Targeting*-Ansätze wurden jeweils 40 µg Expressionsvektor für Int-h/218 oder Mock-Vektor (als negative Kontrolle) zusammen mit 20 µg *Targetvektor* (pCMVattP_{mut}) in die jeweiligen Zelllinien eingebracht. Beim Titrationsexperiment in pTKHattB-Linien wurde FuGENE™6 für die Transfektion des *Targetvektors* benutzt nachdem am Vortag der Expressionsvektor für Int-h/218 mit Elektroporation in die Zellen eingebracht worden war. Nach zwei Wochen Doppelselektion mit Hygromycin [250 µg/ml] und Geneticin [750 µg/ml] wurden überlebende Zellklone entweder vereint (gepoolt) oder vereinzelt und expandiert. Der Nachweis des Integrase-vermittelten genomischen *Targetings* in den HeLa-Reporterzellen, die **pTKHattB** und /oder **prlHF2P** stabil im Genom integriert haben, sollte mittels *Southern Blot* und PCR-Analyse erfolgen. In der PCR wurde die korrekte Integration des *Targetvektors* in die *attB*-Region mit dem Primerpaar P3/P2 (~994 bp; Kap. 4.1.2, Abb. 4-10) nachgewiesen. Der Primer P3 setzt im P-Arm der *attP*-Region an und gibt in Verbindung mit dem Primer P2, der am Ende des Neo-Gens ansetzt, nur dann ein Produkt, wenn nach Integration *attR* vor dem Neo-Gen entsteht (Kap. 4.2.1, Abb. 4-10). Die Amplifikation der *attB*-Neo-Region wurde als Kontrolle mit dem Primerpaar P1/P2 (~850 bp; Kap. 4.2.1, Abb. 4-11-A und 4-10) durchgeführt. In der *Southern Blot* Analyse wurde das gesamte Neo-Gen als Sonde eingesetzt. Korrekte Rekombination des *Targetvektors* in die *attB*-Region der Zelllinie müsste eine 1362 bp Bande ergeben (siehe auch Abb. 4-10, Strategie).

3.4.2.3 Genomisches *Targeting* von *atfH2* in HeLa-Zellen

Für das *Targeting* der natürlich im Genom von HeLa-Zellen vorkommenden *atfH2*-Sequenz, wurde der *Targetvektor* pWSP*_{ns}GFP (20 µg) und der Expressionsvektor für Int-h/218 (40 µg) bzw. pCMV (als Negativ-Kontrolle) eingesetzt. Die Vektoren wurden mittels Elektroporation in HeLa-Zellen bzw. in rHF-stabile HeLa-Zelllinien eingebracht. Nach einer Woche erfolgte die FACS-Analyse oder die FACS-Sortierung der transfizierten Zellen. Bei korrekter Integration des zuvor beschriebenen pWSP*_{ns}GFP-*Targetvektors* soll ein MCT5-eGFP-Fusionsprotein entstehen, welches über den endogenen MCT5-eigenen Promoter ausgeprägt werden soll (Kap. 4.2.3.1 und Abb. 4-18, Strategie). Die PCR-Analyse sollte mit den Primern MCTII und eGFP-End durchgeführt werden und ein Produkt von ~1320 bp ergeben.

4 Ergebnisse

4.1 Intermolekulare integrative Rekombination von episomalen Substraten durch die *Lambda*-Integrase und ihre Mutanten in eukaryotischen Zellen

Bereits 1999 hat unsere Arbeitsgruppe begonnen, das System des Bakteriophagen *Lambda* für die sequenz-spezifische Rekombination in eukaryotischen Zellen zu testen. In vorangegangenen Arbeiten wurden Substrate benutzt, die beide Rekombinationssequenzen auf einem Vektor tragen (intramolekular) (Lorbach et al., 2000; Christ & Dröge 2002). In der vorliegenden Arbeit war die Etablierung der intermolekularen Rekombination von Substraten durch die *Lambda*-Integrase (Int) und ihrer Mutanten Int-h und Int-h/218 von vorrangigem Interesse. Dabei sollten neue Erkenntnisse über den Mechanismus der Rekombination in eukaryotischen Zellen gewonnen und das *Lambda*-System für den biotechnologischen Einsatz weiter entwickelt werden. Im Folgenden steht die intermolekulare Rekombination im Vordergrund, wobei vornehmlich der integrative Reaktionsweg mit *attB* und *attP* in HeLa-Zellen betrachtet wird.

4.1.1 Das Reportersystem für die integrative intermolekulare Rekombination

Die Substratvektoren für die intermolekulare integrative Rekombination wurden so konstruiert, dass *attP* und *attB* auf zwei unterschiedlichen Plasmiden lokalisiert sind (Abb. 4-1, oben). Der eine Vektor trägt dabei den Promoter und der andere einen promoterlosen Reporter, das eGFP-Gen. Die Expression nach illegitimer, zufälliger Rekombination soll durch ein transkriptionelles Stoppsignal verhindert werden, welches direkt vor dem 5'-Ende des eGFP-Gens liegt. Die korrekte Rekombination zwischen den beiden Vektoren führt zur Umlagerung des CMV-Promoters vor das offene Leseraster des eGFP-Gens und somit zu dessen Expression. Die Rekombinationsfrequenz nach Integrase-Transfektion wurde nach durchflusszytometrischer Messung als Prozentzahl der eGFP-positiven Zellen angegeben.

4.1.2 Kinetik der integrativen intermolekularen Rekombination

Nach der Herstellung der Vektoren wurde die generelle Funktionsfähigkeit des Systems getestet. Dabei wurden auch Untersuchungen zur Kinetik der intermolekularen Rekombination durchgeführt. HeLa-Zellen wurden hierfür mit den Substratvektoren (Abb. 4-1, oberer Teil) und den jeweiligen Expressionsvektoren für Int, Int-h oder Int-h/218 kotransfiziert. Als negative Kontrolle diente der Vektor pCMV, der keine Integrase-codierende Sequenz enthält (zur Beschreibung der eukaryotischen Expressionsvektoren siehe auch Kap. 3.1.9). Zu den angegebenen Zeitpunkten (24, 48 und 72 Stunden nach Transfektion) wurden die Zellen geerntet und im FACScan analysiert (Abb. 4-1, Diagramm). Die Transfektionseffizienz wurde für jedes einzelne Experiment durch die Transfektion von peGFPC1 bestimmt und betrug im Allgemeinen 90 bis 98 % (Daten nicht gezeigt).

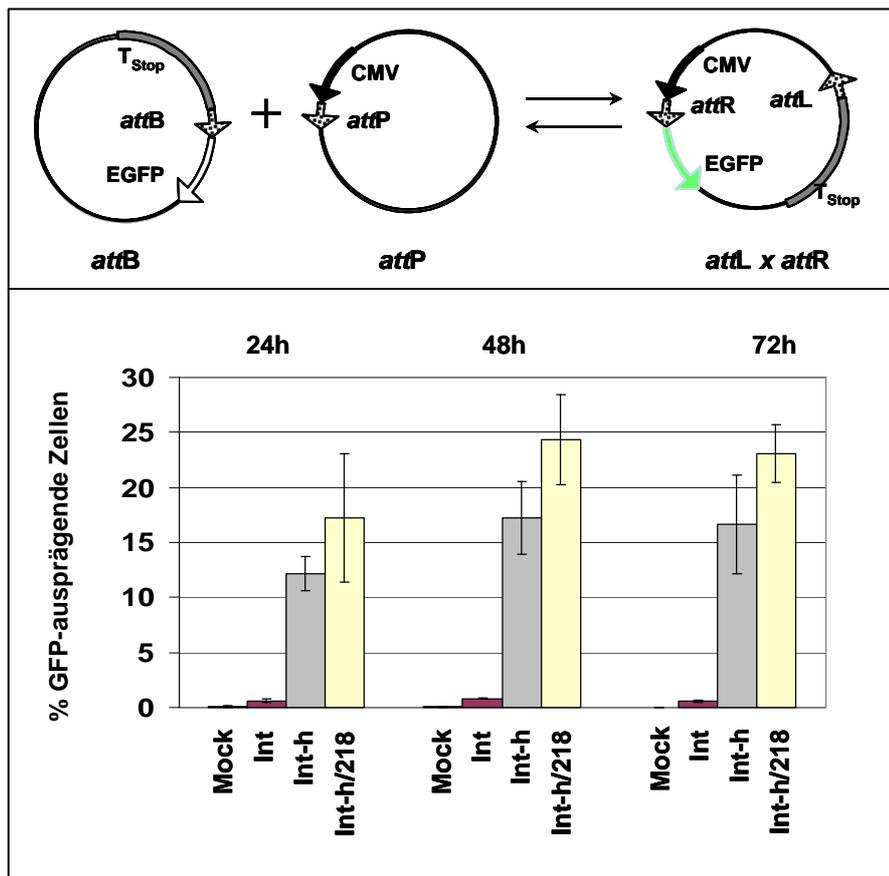


Abb. 4-1: Kinetik der integrativen intermolekularen λ -Rekombination. Die Substratvektoren und ihre Produkte nach erfolgter Rekombination sind schematisch über dem Diagramm dargestellt. Der Anteil an eGFP-ausprägenden Zellen wurde durchflusszytometrisch zu den jeweils angegebenen Zeitpunkten nach Ko-Transfektion der Substrat- und Expressionsvektoren bestimmt. Die Balken zeigen die Mittelwerte aus drei unabhängigen Analysen mit den zugehörigen Standardabweichungen. CMV, Cytomegalovirus Promoter; T_{Stop}, transkriptionelles Stoppsignal; eGFP, *enhanced green fluorescent protein*; Mock, negative Kontrolle nach Ko-Transfektion von pCMV.

Das Ergebnis dreier unabhängiger Experimente ist im Diagramm in Abb. 4-1 dargestellt. Es konnte deutlich gezeigt werden, dass die Integrasen auch im eukaryotischen Milieu in der Lage sind, die Rekombination zwischen Substraten zu katalysieren, die auf unterschiedlichen Molekülen lokalisiert sind. Dabei war die enzymatische Aktivität der Mutanten erheblich höher als die der wild-typ-Integrase. Im Wesentlichen ist dies auf die größere Abhängigkeit der wild-typ-Integrase von der Anwesenheit von Hilfsproteinen wie IHF und XIS, die in eukaryotischen Zellen nicht vorkommen, zurückzuführen. Int-h erreichte hier eine Rekombinationseffizienz von ~21 %; Int-h/218 sogar eine von ~28 %. Im Durchschnitt gelangte der Prozentsatz an eGFP-ausprägenden Zellen für alle getesteten Integrasen zwischen 24 und 48 h nach Transfektion an sein Maximum und blieb dann für den restlichen Verlauf des Experimentes konstant. Demzufolge muss der Großteil der Rekombinationsereignisse in den ersten 24 h nach Transfektion erfolgt sein. Dieses Ergebnis stimmt mit den Expressionsprofilen der Integrasen in den transfizierten Zellen überein (Abb. 4-3B). Über den Zeitbereich der Kinetik wurde repräsentativ für alle verwendeten Integrasen die Int-h/218-Ausprägung in einer *Western Blot* Analyse verifiziert (Abb. 4-3B). Die Analysen zeigen, dass die Int-Ausprägung innerhalb der ersten 24 h nach Transfektion ihr Maximum erreicht und danach kontinuierlich abnimmt. Neben der Detektion der Fluoreszenz wurde auch das Rekombinationsprodukt *attR* durch PCR-Analyse nachgewiesen. Diese wurde mit dem Primer PE, welcher am 5'-Ende der *attP*-Sequenz ansetzt, und dem Primer GFP-End-XbaI, der am 3'-Ende des eGFP ansetzt, durchgeführt. Die nachfolgende Sequenzierung bestätigte die Integrität von *attR* (Daten nicht gezeigt).

Im Vergleich mit den Daten, die in Zusammenarbeit mit Nicole Christ in unserem Labor durchgeführt wurden, unterschied sich die Kinetik der untersuchten integrativen intermolekularen Rekombination nicht von der der integrativen intramolekularen Reaktion (Christ et al., 2002; das System ist unter 4.1.3.3 genauer beschrieben). Die exzise inter- und intramolekulare Rekombination wurde mit den gleichen experimentellen Ansätzen untersucht. Dabei wurden anstelle von *attB* und *attP*, *attL* und *attR* als Rekombinationssequenzen eingesetzt. Auch hierbei waren keine wesentlichen Unterschiede zwischen der inter- und intramolekularen Reaktion zu beobachten. Allerdings war die Effizienz der exzise Rekombinationsreaktion etwas niedriger als die der integrativen (Christ et al., 2002).

4.1.3 Einfluss der Arm-Bindung auf die integrative Rekombination

Die Untersuchungen zur Kinetik der inter- und intramolekularen Rekombination von episomalen Substraten hat sowohl für den integrativen als auch für den exzisen Reaktionsweg gezeigt, dass die Integrase-Mutanten im Durchschnitt in bis zu einem Viertel der transfizierten Zellen Rekombinationsereignisse herbeiführen. Dagegen liegt die enzymatische Aktivität der wild-typIntegrase, die stark von der Anwesenheit von Hilfsproteinen abhängt, nur etwas über dem Hintergrund. Eukaryotische Proteine, die die Funktionen von IHF und XIS übernehmen könnten, scheinen daher in HeLa-Zellen zu fehlen. Es ist bekannt, dass episomale Substrate kurz nach der Transfektion in relaxierter topologischer Form vorliegen (Schwikardi & Dröge 2000). Ein weiterer für Int wichtiger Faktor, die negative Torsionsspannung der DNA-Substrate, ist demzufolge ebenfalls nicht gegeben. Dies lässt vermuten, dass die Integrase-Mutanten die Rekombinationsreaktionen ohne die Bildung eines definierten Nukleoprotein-Komplexes, wie das Intasom an *attP*, katalysieren. Insbesondere die Unabhängigkeit der Mutanten von Protein-Kofaktoren lässt die Frage nach der Bedeutung der Arm-Bindungssequenzen aufkommen, an denen die Hilfsproteine und die N-Domäne der Integrasen binden.

4.1.3.1 Vergleich der Rekombinationseffizienzen zwischen den *att*-Regionen der integrativen Rekombination

Bisher war an den Untersuchungen zur episomalen Rekombination mindestens eine *att*-Region mit Arm-Bindungsregionen beteiligt. Um der Frage nach der Funktion dieser Regionen nachzugehen, wurden intermolekulare Rekombinationsreaktionen durchgeführt, bei denen Paare von Substratvektoren eingesetzt wurden, die *attB* oder *attP* in unterschiedlichen Kombinationen tragen (Abb. 4-2A). Expressionsvektoren für die Integrasen wurden mit diesen Vektoren in HeLa-Zellen kotransfiziert. Der Anteil der eGFP-ausprägenden Zellen wurde 48 h nach Transfektion mittels FACScan ermittelt. Die Effizienz wurde für jedes Experiment einzeln bestimmt und betrug jedes Mal über 90% (Daten nicht gezeigt). In der Abbildung 4-2B sind die Ergebnisse dreier Experimente dargestellt. Demnach war die Effizienz der katalysierten Rekombination zwischen zwei *attP*-Regionen für jede einzelne Integrase vergleichbar mit ihrer Rekombination von *attB* und *attP*. Lagen jedoch nur *attB*-Regionen als Substrate vor, so zeigte sich ein deutlicher Unterschied: nur Int-h/218 war in der Lage, diese effizient zu rekombinieren. Der Anteil der eGFP-

46

ausprägenden Zellen betrug dabei 4-5 %, während er für die wild-typ-Integrase nur bei 0,08 % lag. Diese Ergebnisse verdeutlichen, dass die Gegenwart von Arm-Bindungssequenzen in *att*-Regionen für die Rekombination durch Int-h/218 nicht unbedingt notwendig ist, jedoch die Effizienz der Reaktion in ihrer Anwesenheit deutlich erhöht wird. Diese Stimulation ist noch ausgeprägter, wenn man die Aktivität von Int-h betrachtet. Hier ergab sich eine Steigerung um das Achtfache. Die beobachtete minimale Rekombination durch die wild-typ-Integrase scheint dagegen ausschließlich vom Vorhandensein der Arm-Bindungssequenzen abzuhängen (Abb. 4-2B). Alle Rekombinationsprodukte wurden durch Sequenzierung bestätigt.

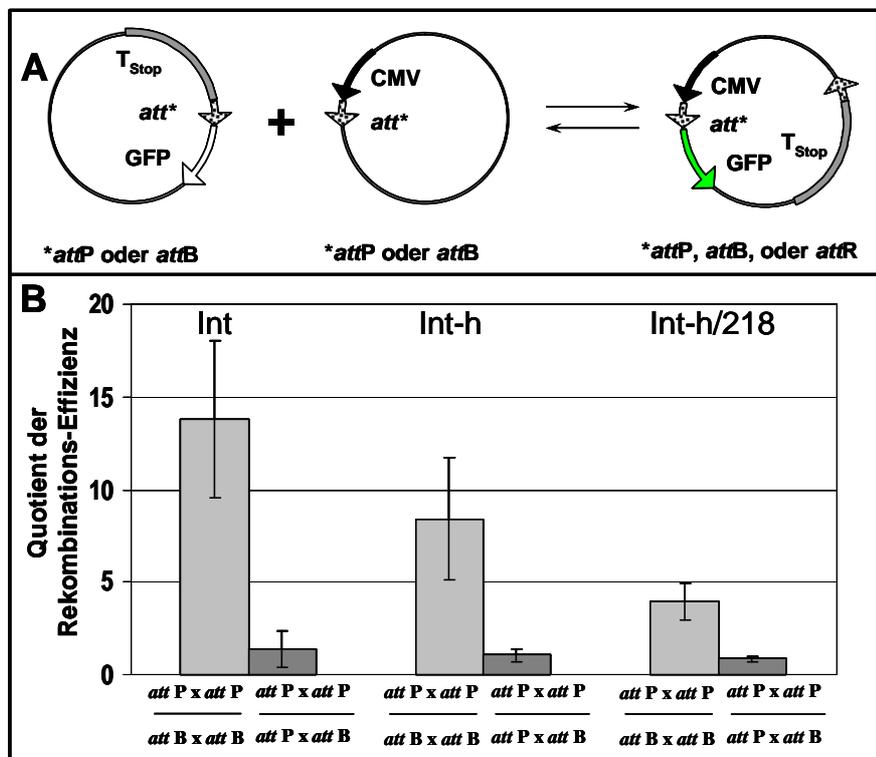


Abb. 4-2: Int-DNA Arm-Bindungssequenzen in *att*-Regionen stimulieren die intermolekulare Rekombination. (A) Substratvektoren für die intermolekulare Rekombination, die entweder *attB* oder *attP* tragen und in verschiedenen Kombinationen zu Produkten rekombiniert werden können. Die resultierende Positionierung des *CMV*-Promoter vor das eGFP-Gen führt zu dessen Ausprägung. **(B)** Verschiedene Kombinationen von Substratvektoren wurden mit den *CMV*-Expressionsvektoren für Int, Int-h oder Int-h/218 in HeLa-Zellen kotransfiziert. 48 h nach der Elektroporation wurden die Zellen durch FACS analysiert und der Quotient der eGFP-ausprägenden Zellen für jede Substrat-Kombination ermittelt. Die Rekombination zwischen *attP* und *attB* wurde als Referenz-Wert eingesetzt. Die Balken zeigen die Mittelwerte aus drei Analysen mit den Standardabweichungen, die als vertikale Linien dargestellt sind. Die entsprechenden Mittelwerte des Anteils an eGFP-ausprägenden Zellen [%] betrug für Int 0,08 (B x B), 1,24 (P x P) und 0,81 (P x B); für Int-h 1,15 (B x B), 8,07 (P x P) und 9,90 (P x B); für Int-h/218 4,01 (B x B), 17,62 (P x P) und 16,45 (P x B).

4.1.3.2 Expression und Aktivität der Integrase-C-Domäne

Die Bindung der Rekombinase und die Öffnung der DNA-Stränge an den *core*-Regionen gehört zu den ersten Schritten einer jeden sequenz-spezifischen Rekombinationsreaktion. Es wurde gezeigt, dass die Arm-Bindungsregionen zwar die Rekombinationsaktivität der Int-h/218 steigern, aber nicht essentiell sind. Wie wichtig ist also die an Arm-Sequenzen bindende N-Domäne (AS 1-64) für die Aktivität von Int-h/218 in eukaryotischen Zellen? Landy und Mitarbeiter konnten *in vitro* zeigen, dass eine N-terminale Deletion von 64 Aminosäuren (N-Domäne) bei der wild-typ-Integrase zu einer so genannten C-Domäne führt, welche eine höhere Affinität zur *core*-Region besitzt und eine gesteigerte Schneideaktivität an derselben zeigt (Sarkar et al., 2001). Um zu untersuchen, ob diese verkürzten Integrasen ohne Arm-Bindungsstellen die Rekombination in HeLa-Zellen katalysieren können, wurden Expressionsvektoren für die C-Domänen von Int und Int-h/218 generiert und getestet (Abb. 4-3A).

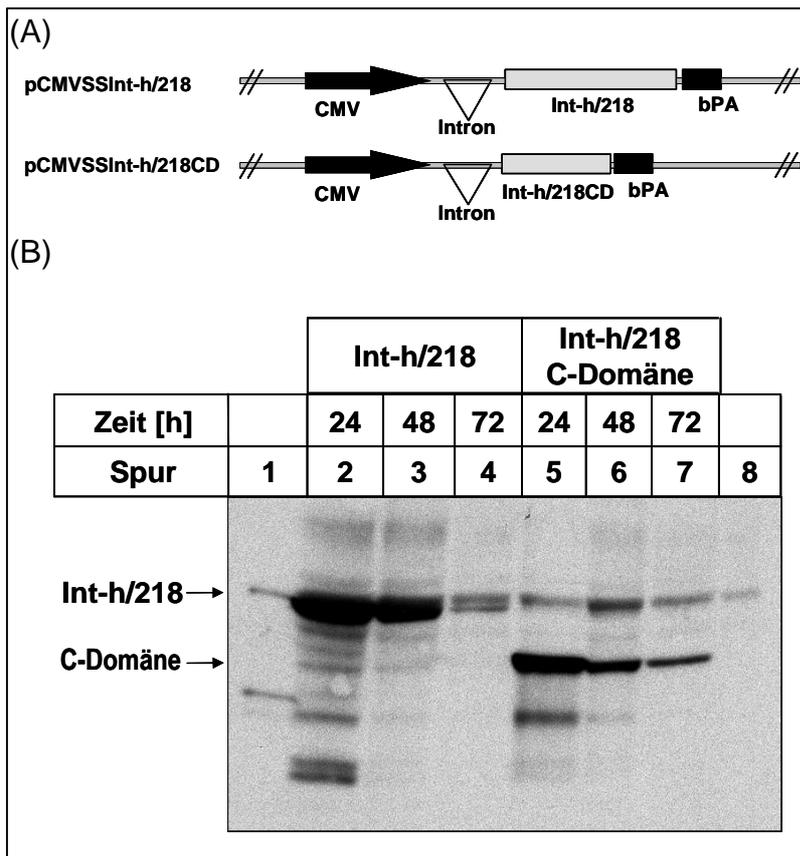


Abb. 4-3: Ausprägung der Int-h/218 Integrase-Mutante und ihrer C-Domäne in HeLa-Zellen. (A) HeLa-Zellen wurden mit einem CMV-Expressionsvektor zur Ausprägung der Int-h/218- Mutante bzw. ihrer C-Domäne transfiziert. CMV, Cytomegalovirus Promoter; bPA, bovine Polyadenylierungs-Sequenz.

(B) Zu den angegebenen Zeitpunkten nach der Elektroporation wurden Zell-Lysate von 5×10^5 Zellen hergestellt und in einer 12,5 % SDS-PAGE analysiert. Zum Nachweis der Integrase-Mutante wurden in der *Western Blot* Analyse polyklonale Maus-Antikörper gegen wild-typ-Int eingesetzt. Spur 1, *E.coli* Zell-Extrakt mit wild-typ-Integrase. Spur 2 bis 4, HeLa-Zell-Lysat mit Int-h/218. Spur 5-7, HeLa-Zell-Lysate mit Int-h/218-C-Domäne, Spur 8 Zell-Lysat von nicht-transfizierten HeLa-Zellen.

Zunächst wurde der Expressionsvektor der C-Domäne von Int-h/218 in HeLa-Zellen transfiziert und durch *Western Blot* Analyse die Ausprägung im Vergleich zum Vollängenprotein betrachtet (Abb. 4-3B). Über einen Zeitraum von 72 h zeigte sich für die C-Domäne ein ähnlicher Verlauf wie für Int-h/218, d.h. die Expression

erreichte innerhalb der ersten 24 h nach Transfektion ihr Maximum und nahm danach kontinuierlich ab.

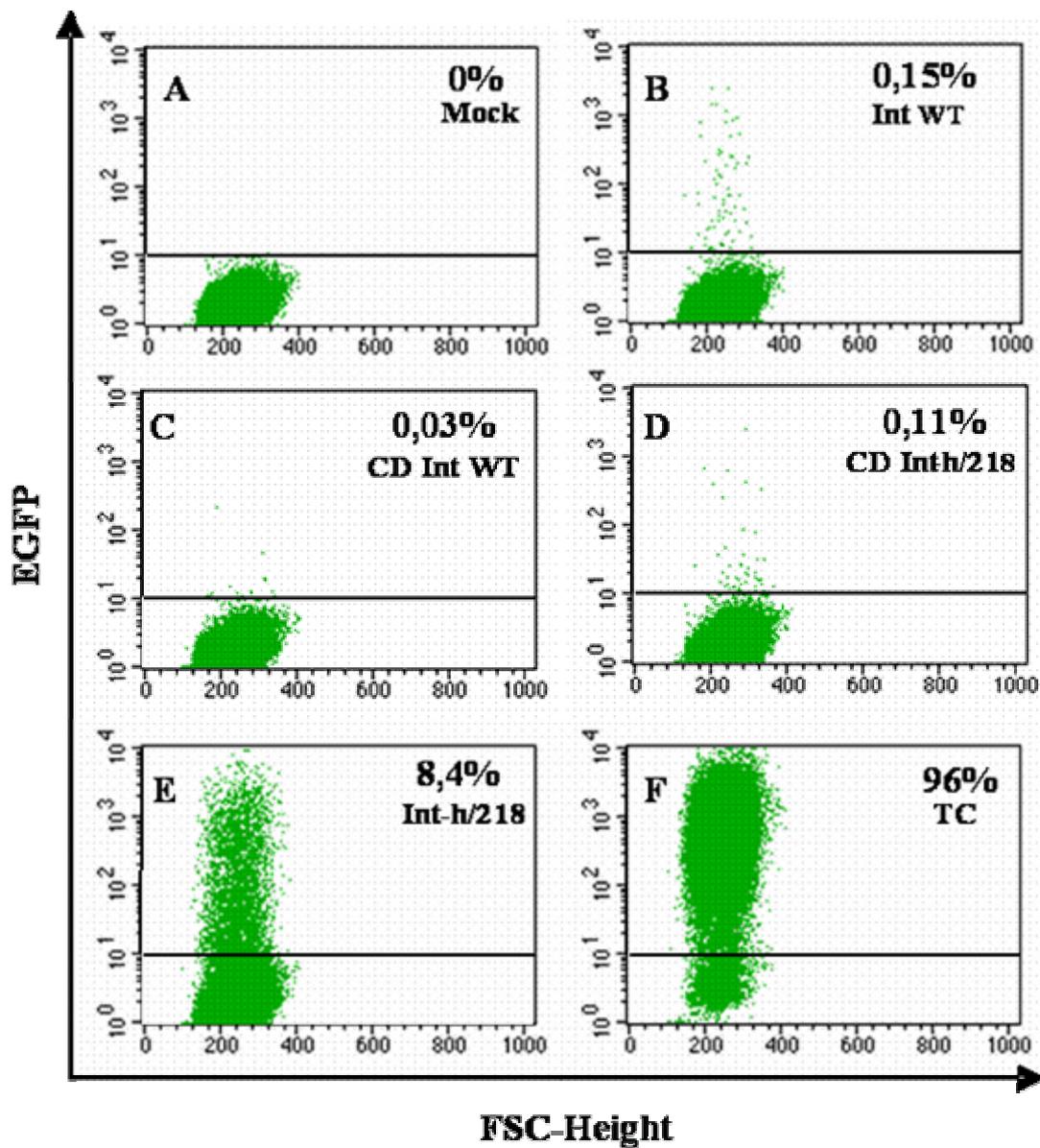


Abb. 4-4: Dot-Plot FACS-Analyse der integrativen intermolekularen Rekombination der C-Domäne von Int-h/218 zeigt eine sehr niedrige Aktivität des verkürzten Proteins. Substratvektoren für die intermolekulare Rekombination (vergleiche Abb. 4-1) und Mock-Vektor ohne Integrase (A), Expressionsvektor für wild-typ-Int (B), für die C-Domäne von wild-typ-Int (C), für die C-Domäne von Int-h/218(D), für das komplette Int-h/218 Protein (E) oder für EGFP als Transfektionskontrolle (TC) (F) wurden in HeLa eingebracht. 48 Stunden nach Kotransfektion wurde der Prozentsatz der GFP-ausprägenden Zellen mittels FACS-Analyse bestimmt. Die Mittelwerte des Anteils an eGFP-ausprägenden Zellen [%] für die inter- und die intramolekulare Rekombination aus je zwei unterschiedlichen Experimenten betragen für Int WT (0,065/0,44); für CD Int WT (0,01/0,02); für CD Int-h/218 (0,055/0,07); für Int-h/218 (7,91/7,61). FSC, forward scatter (Maß für die relative Zellarösse).

Im Anschluss wurde die enzymatische Aktivität der Int und Int-h/218 C-Domänen

untersucht. Hierfür wurden die Expressionsvektoren der C-Domänen jeweils mit Substratvektoren für die inter- oder intramolekulare integrative Rekombination kotransfiziert, wie schon unter 4.1.2 beschrieben. 48 h nach Transfektion wurde die Anzahl der eGFP-ausprägenden Zellen im FACScan gezählt. Eine repräsentative Analyse der intermolekularen integrativen Rekombination ist als *Dot-Plot* in Abb. 4-4 gezeigt. Es wird deutlich, dass die Int-h/218 C-Domäne einen geringen Anteil an eGFP-ausprägenden Zellen zu generieren vermag (vergleiche Abb. 4-4A und D). Dieser ist vergleichbar mit der Menge, die mit dem Volllängen-Protein der wild-typ-Integrase detektiert wird (Abb. 4-4B). Im Vergleich zur Aktivität der Volllängen Int-h/218 ist die ihrer C-Domäne deutlich geringer (Abb. 4-4E). Die C-Domäne der wild-typ-Integrase zeigt dagegen keine relevante Menge an eGFP-ausprägenden Zellen (vergleiche Abb. 4-4A und C). Diese Ergebnisse wurden in zwei weiteren unabhängigen Experimenten mit Substratvektoren für die intramolekulare Rekombination bestätigt (siehe Daten in Legende zu Abb. 4-4).

4.1.4 Der Einsatz des rekombinanten IHFs (rIHF) in eukaryotischen Zellen

Es ist bekannt, dass die wild-typ-Integrase für eine effiziente Rekombination *in vitro* und in *E. coli* die Anwesenheit von Hilfsproteinen benötigt (Landy, 1989, 1993; Azaro & Landy 2002; Christ & Dröge, 1999). Zusätzlich erfordert der integrative Reaktionsweg eine negative Torsionsspannung an der *attP*-DNA (Richet et al, 1986). Auch im eukaryotischen Milieu der HeLa-Zellen scheint dies nicht anders zu sein. In allen hier getesteten Rekombinationsexperimenten war die enzymatische Aktivität der wild-typ-Integrase im Gegensatz zu der ihrer Mutanten eher minimal. In Versuchen zur inter- und intramolekularen Rekombination, in denen Substratvektoren vor dem Einbringen in HeLa-Zellen mit gereinigtem IHF-Protein inkubiert wurden, konnte die Aktivität der wild-typ-Intergase stimuliert werden (Christ et al., 2002). Eine größere Menge an IHF kann allerdings mit dieser Vorgehensweise nicht über einen längeren Zeitraum in eukaryotischen Zellen aufrechterhalten werden. Versuche, das heterodimere IHF über Expressionsvektoren für die beiden Untereinheiten IHF- α und IHF- β in eukaryotischen Zellen zu exprimieren, scheiterten (Nicole Christ, Dissertation 2002). Dabei wurde vermutet, dass es den beiden Untereinheiten nicht gelingt, ein funktionelles IHF-Protein zu bilden. In Zusammenarbeit mit Thomas Schwartz (Rockefeller University, New York) generierte

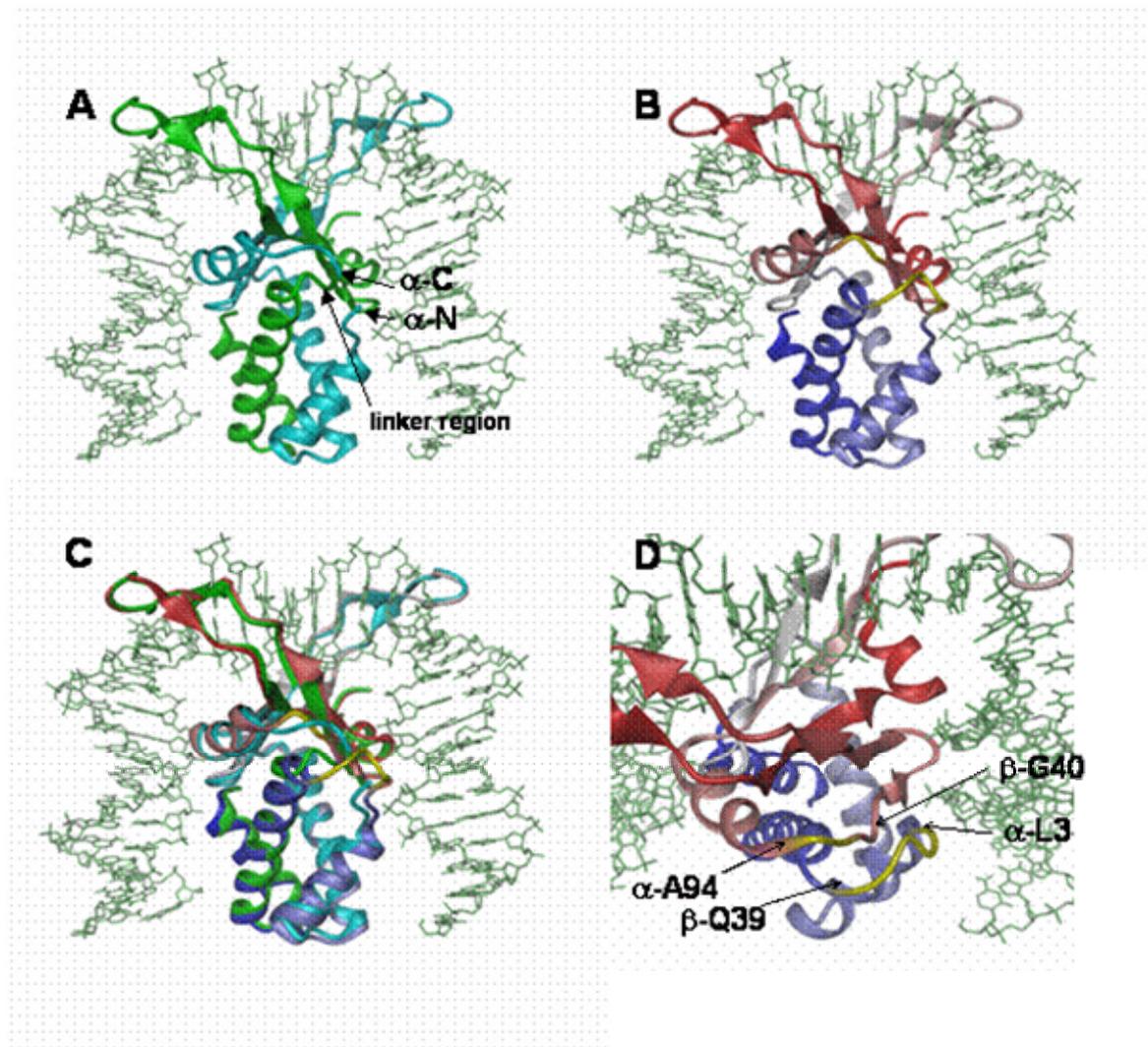
daher Nicole Christ, ein monomeres rekombinantes IHF-ähnliches Protein, im Folgenden als rIHF bezeichnet. Dabei wurde versucht, die dreidimensionale Struktur des heterodimeren IHFs nachzuahmen, um so ein funktionelles Protein zu erhalten.

Nachstehend werden die Langzeit-Expression und die Lokalisation des rIHF in HeLa-Zellen beschrieben. Außerdem wird die Rolle des rIHF in der inter- und intramolekularen Rekombination diskutiert. Für die Analysen wurden Zelllinien generiert, die das rIHF stabil exprimieren.

4.1.4.1 Aufbau des rekombinanten *Integration Host Factors* (rIHF)

Wenn man die Ko-Kristallstruktur des an die H'-Sequenz der $\lambda attL$ -Region gebundenen wild-typ-IHFs betrachtet, so ergibt sich, dass N- und C-Terminus der α -Untereinheit genau entgegengesetzt zu den Enden der β -Untereinheit liegen. Der Einsatz von traditionellen N- oder C-terminalen *Linkern* zur Bildung eines monomeren Proteins war somit ausgeschlossen. Genauere Prüfung der Struktur zeigt jedoch, dass die zwei Enden der α -Untereinheit sehr nah an einer bestimmten Region der β -Untereinheit (β -Q39 bis β -E44) liegen. Diese Region verbindet in IHF- β die zweite α -Helix mit dem ersten β -Faltblatt (Abb. 4-5A). Der rekombinante *integration host factor* (rIHF) wurde daher als ein Hybrid aus den beiden IHF-Untereinheiten IHF- α und IHF- β generiert, indem IHF- α komplett in diese Region von IHF- β eingefügt wurde (Corona et al., 2003). Hierfür wurde der N-Terminus von IHF- α an Position α -L3 mit der Position β -Q39 von IHF- β verbunden. Um den Abstand von ca. 11Å zu überbrücken wurde ein *Aminosäuren-Linker* aus fünf Resten (GGSGG) gewählt (Abb. 4-5B, D und E). Das nun freiliegende Ende β -G40 von IHF- β wurde mit α -A94 am C-Terminus von IHF- α verbunden wobei die 5Å Entfernung mit einem zweiten *Linker* (GG) gefüllt wurde (Abb. 4-5B, D, E). Die Überlagerung der IHF- Ko-Kristallstruktur und des energetisch günstigsten Modells des rIHF ist in Abb. 4-5-C dargestellt. Sie zeigte, dass die zur Verknüpfung der Untereinheiten eingeführten *Aminosäure-Linker* keine bedeutenden Störungen in der dreidimensionalen Struktur verursachen. Für eine eventuelle Reinigung des rIHF wurden am C-Terminus 6 zusätzliche Histidinreste angehängen (Abb. 4-5E). Da der C-Terminus relativ weit von der DNA entfernt ist, wurde eine Behinderung der DNA-Bindung und -Biegung nicht erwartet. Ebenso wurde in der zweiten Position ein Glycinrest eingefügt, um das rIHF-Protein in HeLa-Zellen vor dem Abbau zu

schützten (Bachmair et al., 1986). Das errechnete Molekulargewicht des rIHF beläuft sich demnach auf 22,536 KDa.



E

MGTKS	ELIER	LATQQ	SHIPA	KTVED	AVKEM	LEHMA	STLAQ	<u>GGSGG</u>	45
LTKAE	MSEYL	FDKLG	LSKRD	AKELV	ELFFE	EIRRA	LENGE	QVKLS	90
GFGNF	DLRDK	NQRPG	RNPKT	GEDIP	ITARR	VVTFR	PGQKL	KSRVE	135
<u>NAGGG</u>	ERIEI	RGFGS	FSLHY	RAPRT	GRNPK	TGDKV	ELEGK	YVPHF	180
KPGKE	LRDRA	NIYGG	SGHHH	HHH					203

Abb. 4-5: Strukturen von IHF und rIHF. (A) Ko-Kristallstruktur von IHF und der H' Region von *attL* (nach Rice et al., 1996). α - und β -Untereinheit sind als Bänder in cyan bzw. grün dargestellt. N- und C-Terminus der α -Untereinheit sind markiert. Gleiches gilt für die kurze *Linker*-Region in der β -Untereinheit, die als Insertionsstelle für die α -Untereinheit ausgewählt wurde, um rIHF zu generieren. (B) rIHF-H' Modell. Die zwei *Linker*, die die Verbindung zwischen den beiden Untereinheiten darstellen, sind in gelb dargestellt. Den Verlauf vom N- zum C-Terminus des Proteins verdeutlicht ein Farbgradient von blau (N-Terminus) über weiß bis zu rot (C-Terminus). (C) Überlagerung der IHF-Struktur und des rIHF-Modells. (D) vergleichbar mit (B), mit dem Unterschied, dass die Abbildung vergrößert und um die horizontale Raum-Achse gedreht worden ist. Die zwei *Linker* (gelb) und die entsprechenden Aminosäurereste, die für die Verbindung der beiden Untereinheiten ausgewählt wurden, sind somit in den Vordergrund der Abbildung gerückt. (E) Aminosäuresequenz von rIHF. Die zwei *Linker* sind kursiv und unterstrichen dargestellt. An der zweiten Position wurde zusätzlich ein Glycinrest eingefügt (modifiziert nach Corona et al., 2003; weitere Details siehe Text).

4.1.4.2 Stabile Expression von rIHF in HeLa-Zelllinien

Das rIHF sollte nun in eukaryotischen Systemen getestet werden. Dazu wurde ein geeigneter Expressionsvektor (*prIHF2P*, Abb. 4-6-A) generiert. Das rIHF-Gen wird in diesem Vektor von dem β -Aktin-Promoter aus Huhn, der an den Cytomegalovirus *early enhancer* fusioniert ist (Caggs, Niwa et al., 1991), transkribiert. Um die folgenden Rekombinationsexperimente (4.1.3.3) zu vereinfachen und über einen längeren Zeitraum eine konstante Menge von rIHF in den Zellen zu gewährleisten, wurden stabile Zelllinien generiert. Die Selektion erfolgte mittels Puromycin, dessen Resistenzgen ebenfalls auf *prIHF2P* lokalisiert ist und über den Phosphoglycerat-Kinase (PGK) Promoter exprimiert wird. Resistente Zellklone wurden vereinzelt und expandiert. Die stabile Integration wurde durch PCR (Daten nicht gezeigt) und *Southern Blot* Analyse (4-6-B) nachgewiesen. Als Sonde wurde die Sequenz des kompletten rIHF-Gens eingesetzt. In Abbildung 4-6-B ist die Analyse von drei der generierten Zelllinien dargestellt. Die Zelllinie H/IHF19 enthält mit großer Wahrscheinlichkeit nur eine Kopie des rIHF. H/IHF6 besitzt dagegen zwei Banden im höhermolekularen Bereich (Spur H/IHF6: schwache Banden am oberen Ende des *Blots*) und H/IHF17 zeigt eine Bande in Höhe der Einheitslänge.

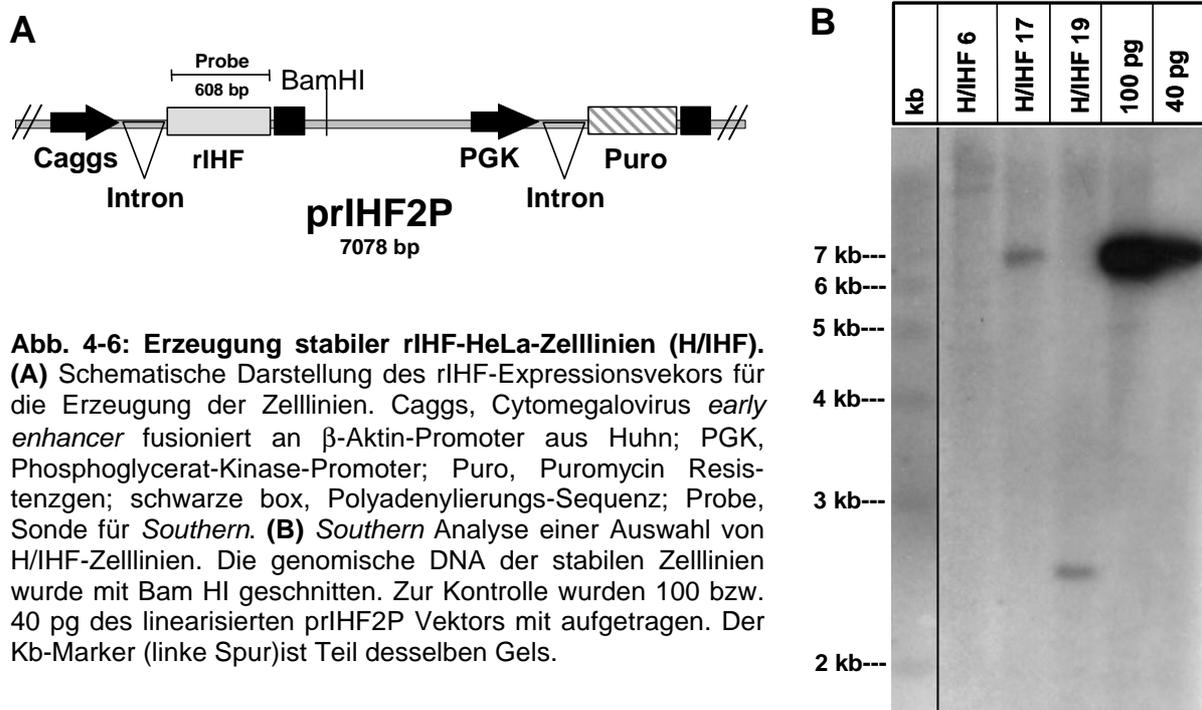


Abb. 4-6: Erzeugung stabiler rIHF-HeLa-Zelllinien (H/IHF). (A) Schematische Darstellung des rIHF-Expressionsvektors für die Erzeugung der Zelllinien. Caggs, Cytomegalovirus *early enhancer* fusioniert an β -Aktin-Promoter aus Huhn; PGK, Phosphoglycerat-Kinase-Promoter; Puro, Puromycin Resistenzgen; schwarze box, Polyadenylierungs-Sequenz; Probe, Sonde für *Southern*. (B) *Southern* Analyse einer Auswahl von H/IHF-Zelllinien. Die genomische DNA der stabilen Zelllinien wurde mit Bam HI geschnitten. Zur Kontrolle wurden 100 bzw. 40 pg des linearisierten *prIHF2P* Vektors mit aufgetragen. Der Kb-Marker (linke Spur) ist Teil desselben Gels.

In einer Langzeitstudie wurde die Ausprägung des rIHF in den stabilen Zelllinien untersucht. Die *Western* Analyse von Rohlysaten, hergestellt aus den kontinuierlich kultivierten transgenen Zellen H/IHF6 und H/IHF17 (Abb. 4-7), zeigt eine hohe und stabile Expression des Proteins über einen Zeitraum von 8 Monaten. Abbauprodukte des rIHF's waren im *Blot* nicht zu erkennen. Die Ergebnisse deuten an, dass das rekombinante Protein von den HeLa-Zellen toleriert wird und somit für eukaryotische Zellen in Kultur offenbar nicht toxisch ist.

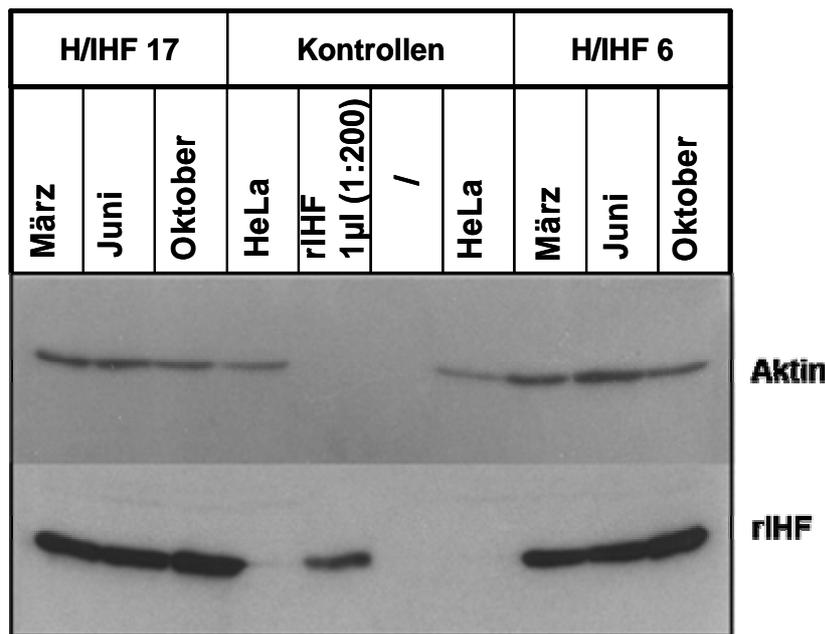


Abb. 4-7: Langzeitexpression des rIHF in den stabilen HeLa-Zelllinien (H/IHF). Rohlysate der transgenen Zelllinien H/IHF6 und H/IHF17 wurden einen, vier und acht Monaten nach Generierung und kontinuierlicher Kultur gewonnen. Die Analyse erfolgte mittels *Western* Blot. Extrakte der parentalen HeLa-Zelllinie dienten als negative Kontrolle. Gereinigtes rIHF (ca. 10ng) wurde als positive Kontrolle eingesetzt. Antikörper gegen Aktin wurde benutzt um eine eventuell unterschiedliche Proteinauftragung zwischen den einzelnen Spuren darzustellen.

Für die Lokalisierung des rIHF wurde eine immunologische Anfärbung der transgenen Zellen durchgeführt. Hierfür mussten die Zellen zunächst fixiert und anschließend permeabilisiert werden, damit die Antikörper in die Zellen eindringen konnten. Als erster Antikörper wurde der polyklonale Antikörper gegen das wild-typ-IHF (S. D. Goodman, Los Angeles) eingesetzt. Durch einen an Fluorescein (FITC) gekoppelten zweiten Antikörper konnte die Detektion im Laser-Scan-Mikroskop (LSM4, Zeiss) erfolgen. So wurde 20 h nach Aussaat der Zellen rIHF vornehmlich im Nukleus der Zellen nachgewiesen, nach 48 h lag rIHF hauptsächlich im Zytosol vor. Die parentalen HeLa-Zellen der Negativ-Kontrolle zeigten zu beiden untersuchten Zeitpunkten nur Hintergrund-Fluoreszenz (Abb. 4-8). Das Ergebnis deutet darauf hin, dass das ~23 KDa große DNA-bindende rIHF (Corona et al., 2003) höchstwahrscheinlich durch intrazelluläre Diffusion in den Zellkern eindringen kann. Die Tatsache, dass es nach 48 h im Zytosol vorliegt lässt jedoch vermuten, dass diese Diffusion in den Kern möglicherweise vom Zellzyklus abhängt.

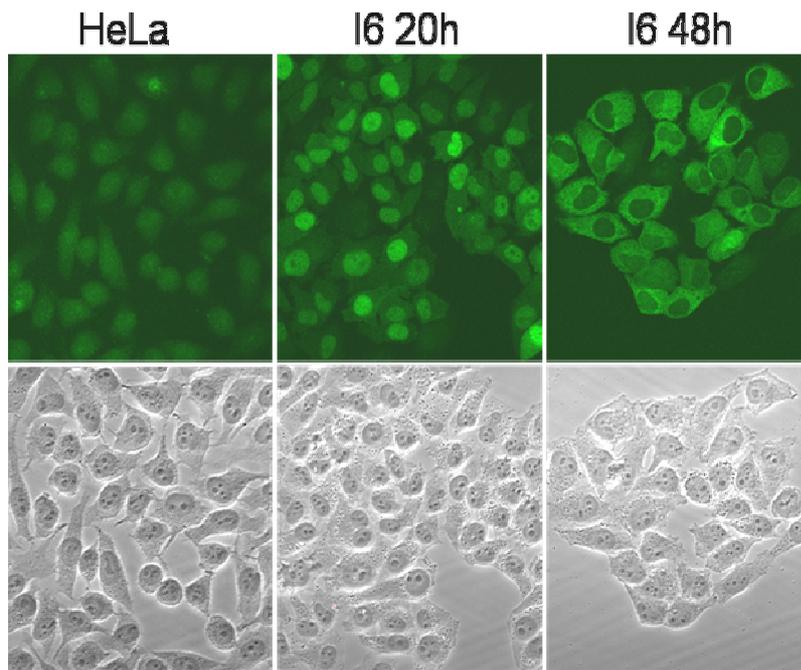


Abb. 4-8: Lokalisierung des rIHF-Proteins in transgenen Zelllinien (H/IHF). Transgene H/IHF 6 und parentale HeLa-Zellen wurden auf Deckgläschen ausplattiert und 20 bzw. 48 Stunden später mit polyklonalen Antikörpern gegen wild-typ-IHF angefärbt (obere Reihe). Die Detektion erfolgte mit fluoreszierenden, FITC-gekoppelten sekundären Antikörpern unter einem LSM (LaserScan Mikroskop, Zeiss). Zum Vergleich sind in der unteren Reihe die zugehörigen Phasenkontrast-Bilder gezeigt. Die Lokalisation des rIHF wurde auch für die Zelllinien (17 und 19) untersucht und führte zu einem mit I6 vergleichbaren Ergebnis (Daten nicht gezeigt).

4.1.4.3 Stabile rIHF-Zelllinie gegenüber parentaler HeLa-Zelllinie: Vergleich der intramolekularen episomalen Rekombination der wild-typ-Integrase und ihrer Mutanten

Untersuchungen des gereinigten rIHF an der Nanyang *Technological University* (Singapur), *in vitro* und in *E. coli*, zeigten keinen großen Unterschied zum wild-typ-Protein. Dabei waren die DNA-Bindung und -Biegung, der Einfluss auf die λ -Rekombination und die Unterstützung bei der Initiierung der Replikation am pSC101 *origin* in *E. coli* betrachtet worden (Bao & Dröge, persönliche Mitteilung; Corona et al., 2003). Zuvor konnte bereits gezeigt werden, dass rIHF in HeLa stabil exprimiert wird und in den Zellkern eindringen kann. Interessant war die Frage, ob rIHF auch in diesen eukaryotischen Zellen in der Lage ist, Einfluss auf die λ -Rekombination zu nehmen. Zur Klärung wurde ein Testsystem für die intramolekulare integrative ($p\lambda$ IR) oder exzisive ($p\lambda$ ER) Rekombination eingesetzt (Abb. 4-9, obere Grafik). Die zu rekombinierenden Regionen, *attP/attB* (integrative Rekombination) bzw. *attL/attR* (exzisive Rekombination), sind jeweils auf demselben Vektor lokalisiert und durch ein transkriptionelles Stoppsignal getrennt. Am 5'-Ende dieser Rekombinationskassette befindet sich jeweils ein CMV-Promoter und ein Intron, am 3'-Ende ein eGFP-Reportergen. Die korrekte Rekombination führt zur Deletion des Stoppsignals und somit zur Ausprägung des eGFPs, die durchflusszytometrisch quantifiziert werden kann. Expressionsvektoren für Int, Int-h, Int-h/218 oder Mock (pCMV als negative

Kontrolle) wurden jeweils mit einem der Substratvektoren in die stabil rIHF exprimierenden Zellen (H/IHF17) und in die parentalen HeLa-Zellen kotransfiziert. 48 h nach Transfektion wurde der Anteil an eGFP ausprägenden Zellen bestimmt. Die Transfektionseffizienzen betragen in der Regel >90 % (Daten nicht gezeigt). Vergleicht man die Werte der beiden Zelllinien miteinander, so zeigt sich, dass in Anwesenheit des rekombinanten IHFs die integrative Rekombination durch die wild-typ-Integrase um das 4- bis 8-fache gesteigert wurde (Abb. 4-9, Diagramm unten links). Diese Stimulierung beschränkte sich jedoch auf den integrativen Reaktionsweg; für die exzisive Rekombination wurde eine vergleichbare Steigerung der Effizienz nicht bestätigt (Abb. 4-9, Diagramm unten rechts).

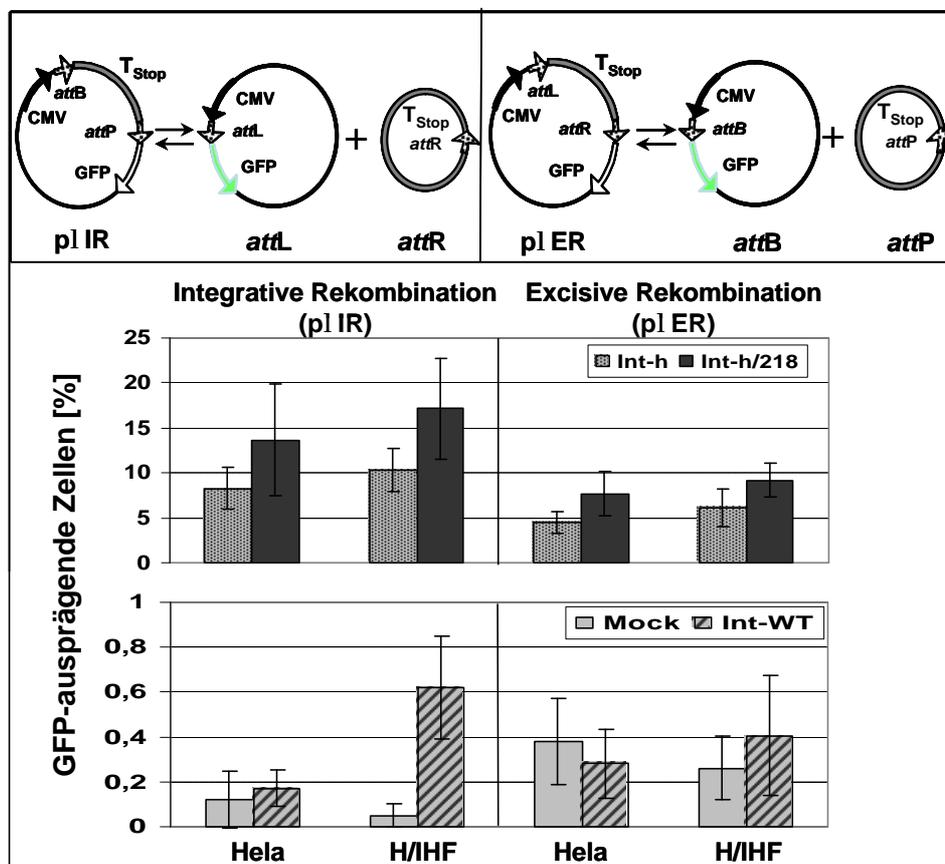


Abb. 4-9: Stimulierung der integrativen Rekombination durch wild-typ-Int in rIHF2His transgenen Zellen. In den oberen Abbildungen sind die Substrate für die integrative (pλIR) und für die exzisive (pλER) Rekombination mit den jeweils erwarteten Produkten schematisch dargestellt. Durch Elektroporation wurden diese Substrate gemeinsam mit Expressionsvektor für wild-typ-Int, Int-h, Int-h/218 oder Mock Vektor (als negativ Kontrolle) in die Zellen eingebracht. Um den eventuellen Einfluss des rIHF messen zu können, wurde neben der transgenen Zelllinie (H/IHF17) zum Vergleich auch die parentale HeLa-Zelllinie transfiziert. 48 Stunden nach Ko-Transfektion von Substrat- und Expressionsvektoren wurde der Anteil an GFP-ausprägenden Zellen durchflusszytometrisch bestimmt. Die Balken zeigen die Mittelwerte aus 5 bis 6 unabhängigen Analysen mit den zugehörigen Standardabweichungen. CMV, Cytomegalovirus Promoter; T_{Stop}, transkriptionelles Stoppsignal; GFP, green fluorescent protein.

Die Aktivität der Integrase-Mutanten Int-h und Int-h/218, die auch in Abwesenheit des wild-typ-IHFs schon sehr hoch ist, konnte durch die Anwesenheit des rIHFs ebenfalls nicht wesentlich verstärkt werden (Abb. 4-9, obere Diagramme). Insgesamt betrachtet reicht jedoch die beobachtete Steigerung der integrativen Rekombinationsaktivität durch die wild-typ-Integrase nicht aus, um das Niveau ihrer Mutanten zu erreichen.

4.2 Intermolekulare Rekombination zwischen genomischen und episomalen Substraten durch die *Lambda*-Integrase und ihre Mutanten in eukaryotischen Zellen

Ein potenzielles Anwendungsgebiet des *Lambda*-Systems ist die gezielte Integration von therapeutischer DNA in das menschliche Genom. Die nachgehend vorgestellten *Targeting*-Strategien sollen zwei mögliche Ansätze darstellen, um sich diesem Ziel zu nähern.

4.2.1 *Targeting*-Strategie für Zelllinien mit genomisch stabil integrierter *attB*-Region

Die intramolekulare Rekombination von zwei, auf demselben Chromosom liegenden, *att*-Regionen durch die λ -Integrase-Mutanten, konnte in unserem Labor schon erfolgreich nachgewiesen werden. Dabei wurde sowohl die Inversion in humanen Zellen als auch die Deletion eines Marker-Gens in embryonalen Stammzellen der Maus untersucht (Lorbach et al., 2000; bzw. Christ et al., 2002). Nachstehend wird eine Strategie beschrieben, nach der eukaryotische Genome unter Verwendung des intermolekularen integrativen Reaktionswegs gezielt manipuliert werden sollen (Abb. 4-10). Um *attB*-stabile Zelllinien zu generieren, wurde der pTKH*attB* Vektor entwickelt (Abb. 4-11A). Dieser enthält ein promoterloses Resistenzgen für Neomycin, welches direkt hinter einer *attB*-Sequenz lokalisiert ist. In entgegengesetzter Orientierung befindet sich ein Resistenzgen für Hygromycin, welches über den HSV-TK-Promoter (Herpes-Simplex-Virus Thymidinkinase Promoter) ausgeprägt wird. Zellen, die den Vektor stabil tragen, werden somit resistent gegenüber Hygromycin. Der minimale *Targetvektor* pCMVSS*attP*_{mut} (Kap. 3.1.10 und 4-1), der in die genomische *attB*-Zielsequenz integrieren soll, trägt einen

CMV (Cytomegalovirus) Promoter, gefolgt von einem Intron und einer *attP*-Sequenz. Diese *attP*-Sequenz ist dabei an drei Positionen so verändert, dass vorhandene ATGs im P-Arm zu ATCs mutiert sind. Die Mutationen betreffen ausschließlich XIS- bzw. FIS-Bindungsregionen und sollten daher auf die integrative Rekombination keinen Einfluss haben.

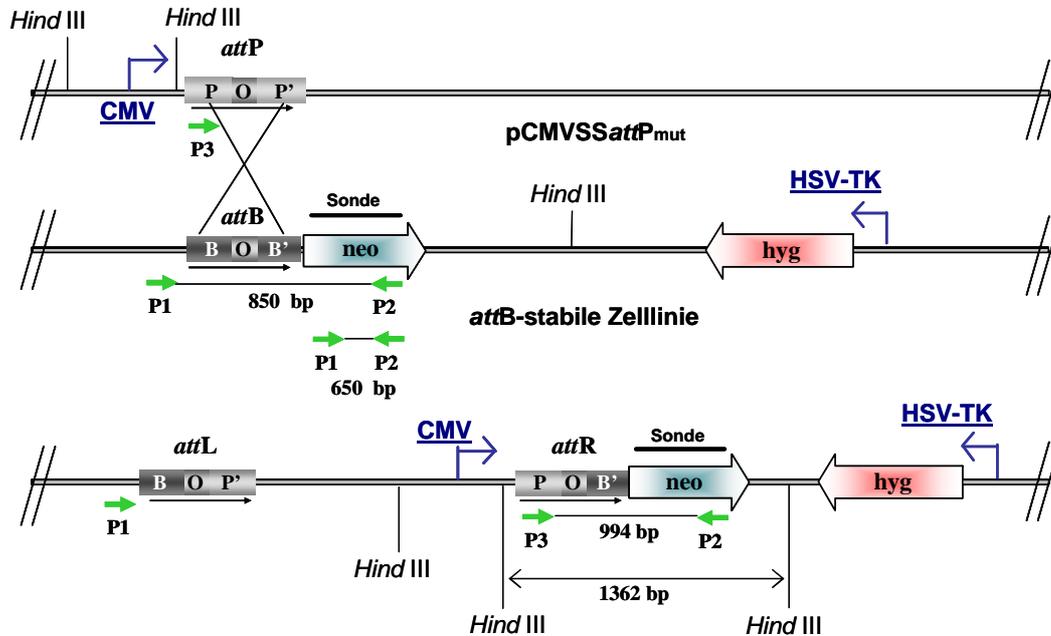


Abb. 4-10: Targeting-Strategie für Zelllinien in denen *attB* stabil integriert ist. Relevante genetische Elemente des *Targetvektors*, der Zielsequenz und des Integrationsproduktes sind schematisch dargestellt. Die Rekombinations-Kassette wird mittels *CMV*-Promoter exprimiert. Erfolgreiche Integration bewirkt eine Expression des *neo*-Gens und führt somit zur Geneticin-Resistenz der Zellen. *attP* bzw. *attB*, Phagen bzw. Bakterien codierte *attachment* Region; *attL* bzw. *attR*, resultierende linke bzw. rechte *attachment* Region nach Integration; *CMV*, Cytomegalovirus Promoter; *HSV-TK*, Herpes-Simplex-Virus Thymidinkinase-Promotor; *hyg*, Hygromycin Resistenzgen; *neo*, Neomycin Resistenzgen. Gekrümmte Pfeile stehen für Promotoren: \curvearrowright , gibt die Orientierung der zugehörigen *attachment* Region an; \rightarrow , gibt die Position und Orientierung der Primer für nachfolgende PCR-Analysen an.

Die Integrase-Mutante Int-h/218 hat sich im Vorfeld als enzymatisch aktivstes Protein erwiesen und soll daher für die *Targeting*-Experimente verwendet werden. Nach Ko-Transfektion mit dem Expressionsvektor für Int-h218 soll die korrekte Integration des *Targetvektors* den *CMV*-Promoter vor das Neomycin Resistenzgen rekombinieren und so zu dessen Expression führen. Selektion mit Geneticin-haltigem Medium soll Zellklone hervorbringen, die ausschließlich positiv für dieses Ereignis sind. Vor der Generierung stabiler Ziellinien mit *pTKHattB* wurde in Ko-Transfektionsexperimenten die Funktionsfähigkeit der Vektoren getestet. Dazu wurde in eine menschliche Zelllinie (293T) *pTKHattB*, *Targetvektor* und Expressionsvektor für Int-h/218 (bzw. *pCMV*-Vektor als negative Kontrolle) ko-

transfiziert. Nach drei Wochen Selektion mit 750 µg/ml Geneticin-haltigem Medium wurden die überlebenden Zellklone ausgezählt. Der Ansatz mit Int-h218 ergab 935, der mit pCMV transfizierte Ansatz nur 756 Neo-resistente Klone. Im Kontrollansatz entstanden die Klone vermutlich durch illegitime Integration in transkribierte genomische Regionen. Dies kann zu einer Ausprägung des Resistenzgens über einen endogenen Promoter als eigenständiges Protein oder als Fusionsprotein führen. Die ca. 23 % höhere Anzahl resistenter Klone, die bei der Ko-Transfektion von Int-h/218 beobachtet wurde, beruht höchstwahrscheinlich auf episodale intermolekulare Rekombinationsereignisse, deren Produkte erst im Anschluss an die Rekombination illegitim in das Genom integrieren. Das Ergebnis deckt sich mit den Untersuchungen in Kap. 4.1.2, in denen die Int-h/218-vermittelte episodale intermolekulare Rekombination in 20 bis 29 % der transfizierten Zellen nachgewiesen wurde. Die Klone jedes Ansatzes wurden vereint und mittels PCR analysiert. Nur bei der Ko-Transfektion von Int-h/218 konnte mit dem Primerpaar P3/P2 (~994 bp; Abb. 4-10) das erwartete Rekombinationsprodukt nachgewiesen werden (Daten nicht gezeigt).

4.2.1.1 Generierung *attB*-stabiler HeLa-Zelllinien

Abb. 4-11 zeigt den Ablauf der Generierung und die Identifizierung *attB*-stabiler Zelllinien. Im ersten Schritt wurden HeLa-Zellen mit ungeschnittenem (U) oder linearisiertem (L) pTK*HattB* transfiziert (Abb. 4-11A). Zwei Tage nach Transfektion wurde mit der Hygromycin-Selektion begonnen, um Zellklone zu erhalten, welche den Vektor an einer zufälligen Stelle des Genoms integriert hatten. Nach ca. zwei bis vier Wochen Selektion konnten dann Klone vereinzelt und expandiert werden. In einer *Southern* Analyse wurde die Anzahl der integrierten Vektorkopien untersucht (Abb. 4-11B). Dabei wurde das komplette Neo-Gen als Sonde eingesetzt (Abb. 4-11A). Interessanterweise besaßen fast alle untersuchten Zelllinien nur eine Kopie. Lediglich die Zelllinie *HattB*3, die aus einem anderen Ansatz mit ungeschnittenem pTK*HattB* stammt, zeigt zwei Banden. Bei den Zelllinien U13, U26 und L6 konnten keine Banden detektiert werden, allerdings waren die positiven Signale auf diesem *Blot* ohnehin sehr schwach. Um sicher zu gehen, dass keine zufällige Ausprägung des Resistenzgens für Neomycin durch einen endogenen Promoter vorhanden war, wurde jede Zelllinie auf Sensitivität gegenüber Geneticin [750 µg/µl] getestet. Nur Zelllinien, die nach ein bis zwei Woche starben, wurden weiter verwendet. Die

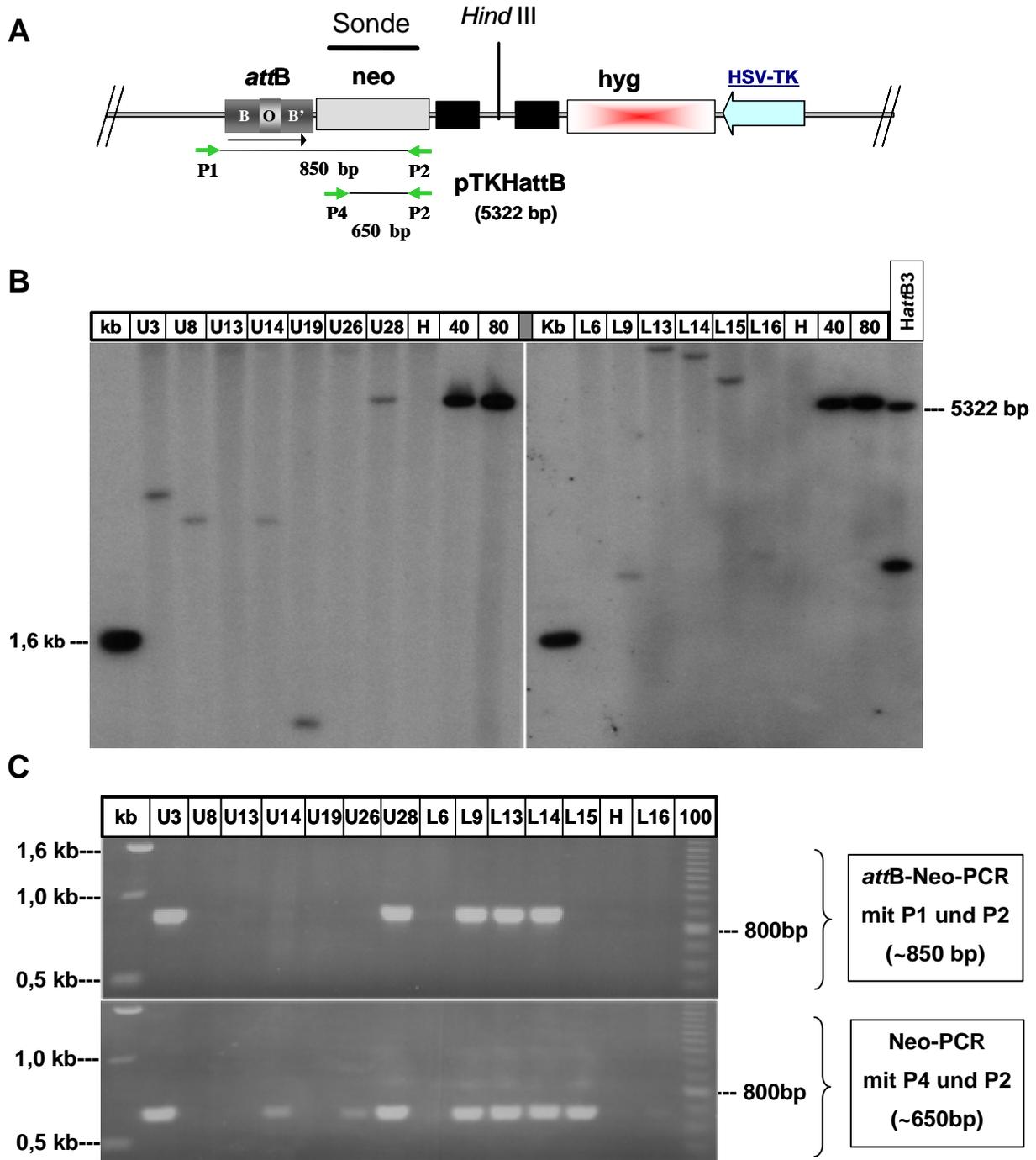


Abb. 4-11: Identifizierung attB-stabiler Zelllinien. (A) Vektor für die Generierung der attB-stabilen Zelllinien. Die schwarzen Boxen stehen für Polyadenylierungs-Sequenz; HSV-TK, Herpes Simplex Virus Thymidin-Kinase-Promotor; hyg, Hygromycin Resistenzgen; neo, Neomycin Resistenzgen; \longrightarrow gibt die Orientierung von attB an; \rightarrow , gibt Position und Orientierung der Primer für die PCR-Analysen in (C) an. **(B)** Southern Analyse. HeLa-Zellen wurden nach Transfektion mit linearisiertem (L) oder ungeschnittenem (U) Vektor transfiziert und unter Selektion gebracht (250 μ g/ml Hygromycin). Nach zwei bis vier Wochen wurden resistente Klone vereinzelt und expandiert. Genomische DNA der HeLa-Zellklone wurde mit dem Restriktionsenzym Hind III, welches nur einmal im pTKHattB-Vektor schneidet, behandelt. Die Integration des Vektors ins Genom wurde mit einer Neo-Sonde nachgewiesen. Auf dem Blot sind nur die Klone gezeigt, die sich als sensitiv gegenüber Geneticin [750 μ g/ml] erwiesen haben. Kb, Kb-Marker; H, untransfizierte HeLa-Zellen; 40 bzw. 80, Auftrag von 40 bzw. 80 pg linearisiertem Vektor (Einheitslänge), HattB3, Zelllinie aus einem anderen Transfektionsansatz mit ungeschnittenem pTKHattB. **(C)** PCR-Analyse der Zellklone. Nicht alle in der Southern Analyse positiven Klone besitzen die für das Targeting wichtige attB-Neo Sequenz. H, untransfizierte HeLa-Zellen; Kb, Kb-Marker; 100, 100 bp-Marker.

genomische *attB*-Neo-Region und das Resistenzgen für Neomycin wurden durch PCRs mit den Primerpaaren P1/P2 (erwartete Größe: ca. 850 bp) bzw. P4/P2 (erwartete Größe: ca. 650 bp) amplifiziert (Abb. 4-11A und C) und durch Sequenzanalyse überprüft (Daten nicht gezeigt). Die PCR-Analyse der *HattB3*-Linie ist in Abb. 4-12 (rechte Gelhälfte) gezeigt. Nicht alle in der *Southern* Analyse positiven Klone besaßen die für das *Targeting* wichtige *attB*-Neo Sequenz. Nur U3, U28, L9 und *HattB3* kamen letztendlich für weitere Versuche in Frage. Die Zelllinien L13 und L14 zeigten bei der Neo-Selektion zwar eine Sensitivität, diese betraf jedoch nicht alle Zellen, da nach zwei und mehr Wochen noch resistente Klone gesehen wurden. Diese Tatsache wurde als Indiz für nicht klonale Zelllinien betrachtet, und führte daher zum Ausschluss von Folgeexperimenten.

4.2.1.2 *Targeting attB*-stabiler HeLa-Zelllinien mit pCMVSS*attP*_{mut}

Es wurden mehrere *Targeting*-Experimente durchgeführt. Dabei wurden jeweils 40 µg Expressionsvektor für In-h/218 oder Mock-Vektor (als negative Kontrolle) zusammen mit 20 µg *Targetvektor* (pCMVSS*attP*_{mut}) in die jeweiligen Zelllinien eingebracht. Nach zwei Wochen Doppelselektion mit Hygromycin [250 µg/ml] und Geneticin [750 µg/ml] starben entweder alle Zellen oder es wuchsen auch Klone auf negativen Kontrollansatz aus. Dabei variierte die Anzahl. Für die *HattB3*-Linie lag sie z.B. zwischen ca. 150 bis 500 Klonen pro Transfektionsansatz. Insgesamt betrachtet gab es jedoch in den einzelnen Experimenten keinen großen Unterschied zwischen den Ansätzen mit In-h/218 und denen ohne Integrase (Daten nicht gezeigt). Es schien folglich vorerst nicht sinnvoll, die Klone zu vereinzeln und getrennt zu untersuchen. Stattdessen wurden die Neo-resistenten Klone vereint und als Pool mittels PCR geprüft. Für die Amplifikation der *attB*-Neo-Region wurde wieder das Primerpaar P1/P2 (~850 bp; Abb. 4-11A und 4-10) eingesetzt. Die korrekte Integration des *Targetvektors* in die *attB*-Region wurde mit dem Primerpaar P3/P2 (~994 bp; Abb. 4-10) nachgewiesen. Der Primer P3 setzt im P-Arm der *attP*-Region an und gibt in Verbindung mit dem Primer P2, der am Ende des Neo-Gens ansetzt, nur dann ein Produkt, wenn nach Integration *attR* vor dem Neo-Gen entsteht (Abb. 4-10). In Abbildung 4-12 ist eine PCR-Analyse des *Targetings* der *HattB3*-Zelllinie dargestellt. In dem Diagramm über dem Gel sind die transfizierten Vektoren angegeben. Die rechte Gelhälfte zeigt die PCR für die *attB*-Neo-Sequenz, die linke die für die korrekte Integration. Nur die Ko-Transfektion des *attP*-tragenden Vektors

und des Expressionsvektors für Int-h/218 führte zur Detektion des erwarteten Integrationsproduktes (Spur 3), dessen Integrität durch Sequenzierung bestätigt wurde (Daten nicht gezeigt). Als negative Kontrollen dienten die untransfizierte Zelllinie und mit Mock-Vektor bzw. mit Mock- und *Targetvektor* transfizierte Zellen. Dieses Ergebnis wurde in drei unabhängigen Transfektions-Experimenten verifiziert.

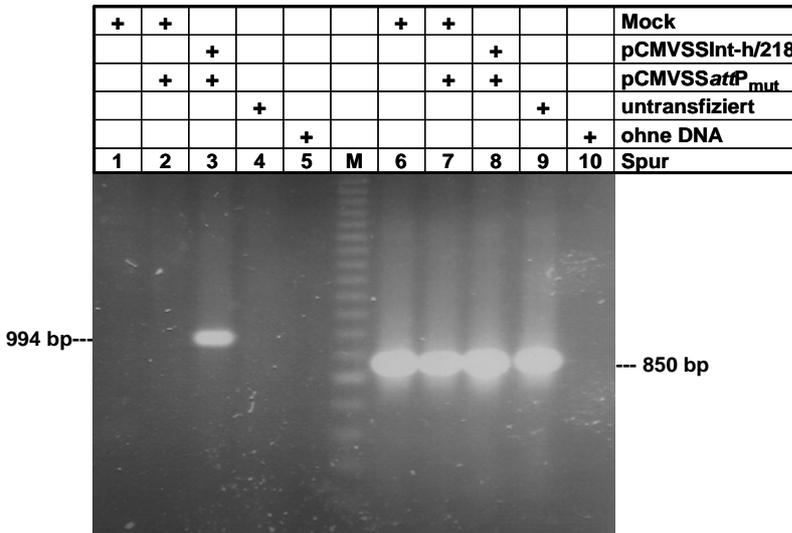


Abb. 4-12: PCR-Analyse der Integrase vermittelten sequenz-spezifischen Integration in die genomische attB-Region einer stabilen Zelllinie. Zellen der *HattB3*-Linie, die 1-2 Kopien der attB-Region stabil im Genom integriert haben, wurden mit den angegebenen Vektoren transfiziert. Nach zwei Wochen Selektion mit Geneticin [750 µg/ml] wurden die resistenten Zellklone gepoolt und die genomische DNA isoliert. Die PCR zur Amplifikation des korrekten Integrationsproduktes wurde mit dem Primerpaar P3 und P2 durchgeführt (Spuren 1-5). Um die genomische attB-Region zu amplifizieren wurden die Primer P1 und P2 (Spuren 6-10) benutzt (siehe auch Abb. 4-10, Strategie). Nur die Ko-Transfektion des attP-tragenden Vektors und des Expressionsvektors für Int-h/218 führte zur Detektion des erwarteten Integrationsproduktes (Spur 3). Dieses Ergebnis wurde in 3 unabhängigen Experimenten verifiziert. M, 100bp-Marker.

Nur die Ko-Transfektion des attP-tragenden Vektors und des Expressionsvektors für Int-h/218 führte zur Detektion des erwarteten Integrationsproduktes (Spur 3). Dieses Ergebnis wurde in 3 unabhängigen Experimenten verifiziert. M, 100bp-Marker.

Die Ursache für die hohe Anzahl an Neo-resistenten Klonen in den negativen Kontrollen blieb zunächst unklar. Nachdem in einem weiteren Experiment mit der *HattB3*-Zelllinie auch untransfizierte Zellen unter Selektionsdruck resistente Klone ergaben, die sogar eine Konzentration von 3000 µg/ml Geneticin überlebten, wurden die Klone erneut für jeden Ansatz vereint und als „resistente Pools“ analysiert. Eine *Southern* Analyse (Abb. 4-13) verhalf zur vorläufigen Klärung des Problems. Es stellte sich heraus, dass nach Selektion eine dritte Integrationsstelle des Neo-Gens im *Blot* auftauchte (Abb. 4-13, Spuren 2-8), die bei der ursprünglichen Zelllinie nicht gesehen wurde (Abb. 4-13, Spur 9). Dieses dritte Neo-Gen war vermutlich für die Neo-Resistenz der Zellen verantwortlich, da es möglicherweise in eine transkribierte Region des Genoms integriert ist und über einen endogenen Promoter eigenständig oder als Fusionsprotein ausgeprägt wird. Korrekte Rekombination des *Targetvektors* in die attB-Region der Zelllinie müsste eine 1362 bp Bande ergeben (Abb. 4-10, Strategie). Diese war jedoch bei keinem Ansatz erkennbar.

vorhanden war, ist bei den Subklonen verschwunden (Spuren 1-6).

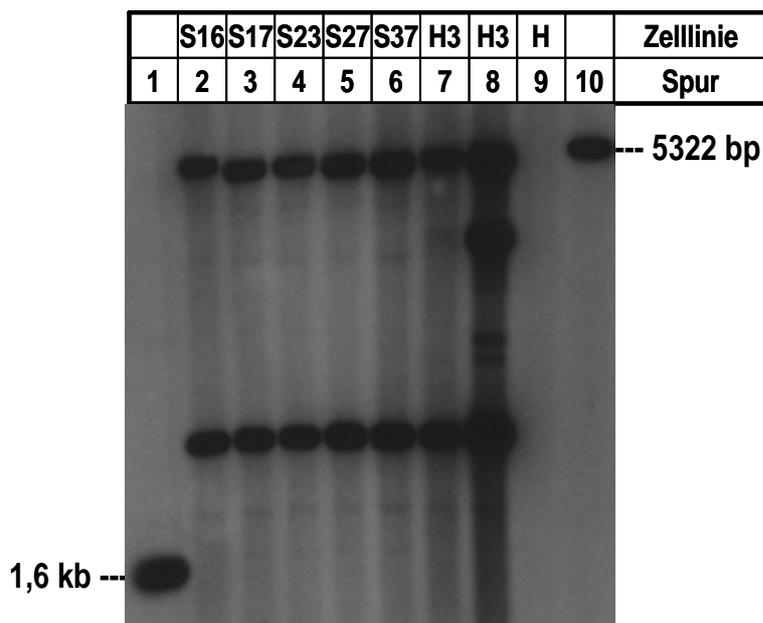


Abb. 4-14: Southern Analyse nach Subklonierung von *HattB3*. Zellen der *HattB3*-Linie wurden dünn ausgesät um Einzelklone zu erhalten. Nach zwei bis vier Wochen unter ständigem Selektionsdruck (250 µg/ml Hygromycin) wurden die Klone vereinzelt und expandiert. Genomische DNA der *HattB3*-Subklone (S..) wurde mit dem Restriktionsenzym Hind III, welches nur einmal im pTK*HattB*-Vektor schneidet, behandelt. Als Sonde diente das gesamte Neo-Gen. Der *Blot* zeigt nur die Klone, die sich als sensitiv gegenüber Geneticin [750 µg/ml] erwiesen haben. H3, *HattB3* Ausgangszelllinie; H, HeLa. Spur 1, Kb-Marker, Spur 7 bzw. 8, *HattB3* Zellen vor bzw. nach Geneticin-Selektion. Spur 10, Auftrag von 40 pg linearisiertem Vektor.

Im Anschluss wurden diverse *Targeting*-Experimente mit den neuen Subklonen der *HattB3*-Linie durchgeführt. Der Hintergrund der Experimente war deutlich reduziert. Es wurden nun wesentlich weniger neoresistente Klone gesehen. Statt 150-500 für die *HattB3*-Linie waren es nur noch 3-10 für ihre Subklone. Dennoch zeigten die negativen Kontrollen noch immer keinen bedeutenden Unterschied zu Ansätzen, in denen die Integrase mit transfiziert wurde. In einigen Experimenten befanden sich auf den Kontrollen sogar mehr Klone (Daten nicht gezeigt). Bei den anderen *attB*-stabilen Zelllinien U3, U28 und L9 (siehe Abb. 4-11) starben jedoch meist alle Zellen. Ein Grund hierfür konnte die Konzentration des für die Selektion eingesetzten Geneticins sein. In einem Titrations-Experiment sollte daher für jede Zelllinie die Schwellenkonzentration für Geneticin bestimmt werden, bei der nach einem *Targeting* noch Zellklone überlebten.

Für das Titrations-Experiment wurden die Zellen der jeweiligen Linien erst mit Expressionsvektor für Int-h/218 transfiziert und auf 6-well-Platten in vergleichbarer Dichte vorgelegt. Der *attP*-tragende *Targetvektor* wurde am darauf folgenden Tag durch Lipofektion mit Fugene™ in die jeweiligen Zellen gebracht. Die Transfektionseffizienz der Elektroporation betrug 50 bis 80 %, die der Lipofektion nur zwischen nur 16 bis 32 %. Für S17 wurde sogar nur eine Effizienz von ca. 5 % nach Lipofektion gemessen. Zwei Tage nach Einbringen des *Targetvektors* wurde die

Selektion mit fünf Geneticin-Konzentrationen zwischen 200 und 600 µg/ml begonnen (Abb. 4-15, Tabelle). Nach zwei Wochen wurden die resistenten Zellklone vereint und die genomische DNA isoliert. Die Sensitivität für Geneticin war unterschiedlich und

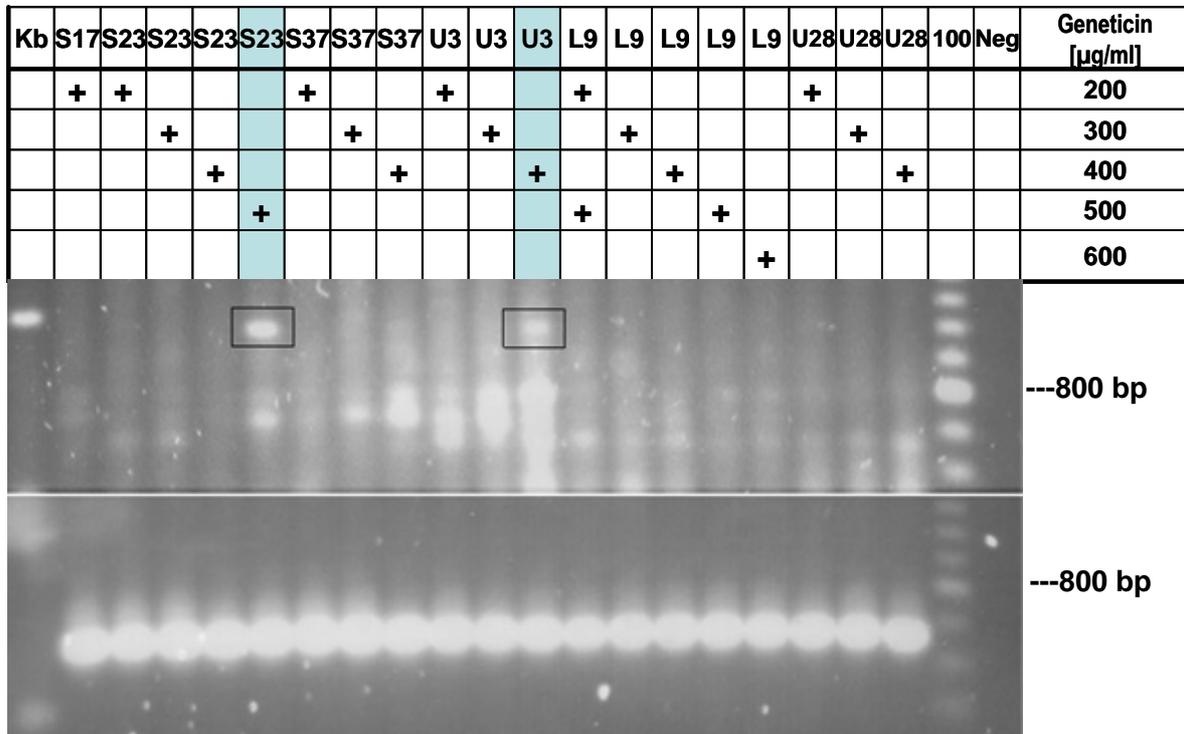


Abb. 4-15: PCR Analyse nach Targeting verschiedener attB-stabiler Zelllinien und nachfolgender Selektions-Titration. Zellen verschiedener attB-stabiler Zelllinien, die 1 oder 1-2 Kopien der attB-Region stabil im Genom integriert haben (siehe Abb. 4-11C), wurden mit dem Expressionsvektor für Int-h/218 und dem Targetvektor pCMVSSattP_{mut} transfiziert. Nach zwei Wochen Selektion mit verschiedenen Geneticin Konzentrationen [200-600 µg/ml] wurden die resistenten Zellklone vereint und die genomische DNA isoliert, wobei es nicht bei jeder Konzentration überlebende Zellen gab. Die PCR zur Amplifikation des korrekten Integrationsproduktes wurde mit dem Primerpaar P3 und P2 durchgeführt (oberes Gel). Um die genomische attB-Region zu amplifizieren, wurden die Primer P1 und P2 benutzt (unteres Gel, siehe auch Abb. 4-10, Strategie). Nur bei S23/500 und bei U3/400 wurde das erwartete Integrationsprodukt detektiert (schattierte Spalten in Tabelle). Kb, Kb-Marker; 100, 100 bp-Marker; neg., Negativ-Kontrolle ohne DNA.

es gab nicht bei jeder Konzentration überlebende Zellen. Der Schwellenwert, schien bei den meisten Zelllinien zwischen 400 und 500 µg/ml zu liegen. Nur die Zelllinie L9 zeigte auch bei 600 µg/ml noch überlebende Klone. Bei S17 dagegen war schon bei 300 µg/ml kein einziger resistenter Zellklon mehr vorhanden. Die PCR zur Amplifikation des korrekten Integrationsproduktes wurde mit dem Primerpaar P3/P2 durchgeführt (Abb. 4-15, oberes Gel). Um die genomische attB-Region zu amplifizieren wurden die Primer P1 und P2 benutzt (Abb. 4-15, unteres Gel; siehe auch Abb. 4-10, Strategie). Nur bei S23/500 und bei U3/400 wurde das erwartete

Integrationsprodukt detektiert und konnte durch Sequenzierung bestätigt werden.

Mit der S23 Zelllinie wurden weitere *Targeting*-Experimente durchgeführt, bei denen die Selektion nur noch mit 500 µg/ml Geneticin erfolgte. Die Anzahl der Neomycin-resistenten Klone stieg wieder an auf bis zu 200 und wurde auch bei den Kontrollen weiter beobachtet (Daten nicht gezeigt). Vereinzelt man einige dieser im Int-h/218 Ansatz resistenten Klone, so zeigten sich weder in PCR- noch in *Southern* Analyse das richtige Rekombinationsprodukt. Auch gab der *Southern* kein Indiz für eine Neuordnung der genomischen Neo-Sequenz, die eventuell zu einer Resistenz geführt haben könnte (Daten nicht gezeigt).

4.2.2 Einfluss von rIHF auf die genomische Rekombination

Die vorherigen Experimente zeigen, dass die Integration des *attP*-tragenden Vektors in die, stabil im HeLa-Genom eingeführte, *attB*-Region prinzipiell funktioniert. Dennoch ist die Effizienz, mit der Int-h/218 die Reaktion katalysiert, nicht hoch genug, um sich deutlich vom Hintergrund abzuheben. Eine Möglichkeit diese Effizienz zu steigern bot der Einsatz des rekombinanten IHFs (Kap. 4.1.3). Das rIHF hatte in den episomalen Tests zwar keinen elementaren Einfluss auf die Aktivität der Mutanten, dennoch war eine geringe Steigerung erkennbar (Abb. 4-9). Verglichen mit der wild-typ-Integrase in den episomalen Tests, katalysierte Int-h/218 die intermolekulare chromosomale Rekombination mit geringerer Effizienz. Der positive Effekt des rIHF's könnte daher auf genomischer Ebene deutlicher ausfallen als zuvor in den transienten Experimenten beobachtet wurde.

4.2.2.1 Generierung von HeLa-Zelllinien mit stabil integriertem rIHF und *attB*

Ausgehend von der S23 Zelllinie mit der stabil integrierten *attB*-Neo-Region sollten Zelllinien generiert werden, die zusätzlich rIHF kontinuierlich ausprägten. Hierfür wurden S23 Zellen mit prIHF2P Vektor (Abb. 4-16A und Kap. 4.1.3.2, Abb. 4-6A) transfiziert. Die Selektion erfolgte ausschließlich mit 20 µg/ml Puromycin, da nach versuchter Doppelselektion mit 250 µg/ml Hygromycin (an die *attB*-Region gekoppeltes Resistenzgen) keine Klone erzeugt werden konnten. Resistente Klone wurden vereinzelt und expandiert. Die Identifizierung rIHF stabiler S23 Zelllinien ist in Abb. 4-16 dargestellt. Die Ausprägung des rIHF's wurde mittels *Western Blot* Analyse untersucht und bestätigte sich für Si8, Si12, Si14, Si16, Si17 und Si29. Positive Ergebnisse sind als „+“ in der Tabelle in Abb. 4-16-B gekennzeichnet. In der Tabelle

sind ebenfalls die Ergebnisse für die Neo- und die *attB*-Neo-PCR angegeben (vergleiche auch Abb. 4-11A). Alle Si Zelllinien zeigen ein positives Signal für das Neomycin-Resistenzgen, bis auf Si29 haben jedoch alle die *attB*-Neo-Region verloren. Mit diesem Verlust geht auch der Verlust der Hygromycin-Resistenz einher. Bis auf Zelllinie Si29 starben alle Si-Linien, wenn sie einem Selektionsdruck mit Hygromycin [250 µg/ml] ausgesetzt wurden. Gegenüber Neomycin [500 µg/ml] zeigten sich alle Si-Linien sensitiv (Daten nicht gezeigt). Betrachtete man ihre genomische DNA mit einer Neo-Sonde im *Southern Blot*, zeigte nur Si29 das Muster der S23-Ausgangszelllinie. Bei zwei Zelllinien, Si12 und Si23, konnte eine Umlagerung des Neo-Gens detektiert werden. Bei den anderen war, trotz des positiven Signals in der PCR, keine Neo-Bande im Bereich zwischen 1,6 Kb und 6 Kb mehr zu erkennen.

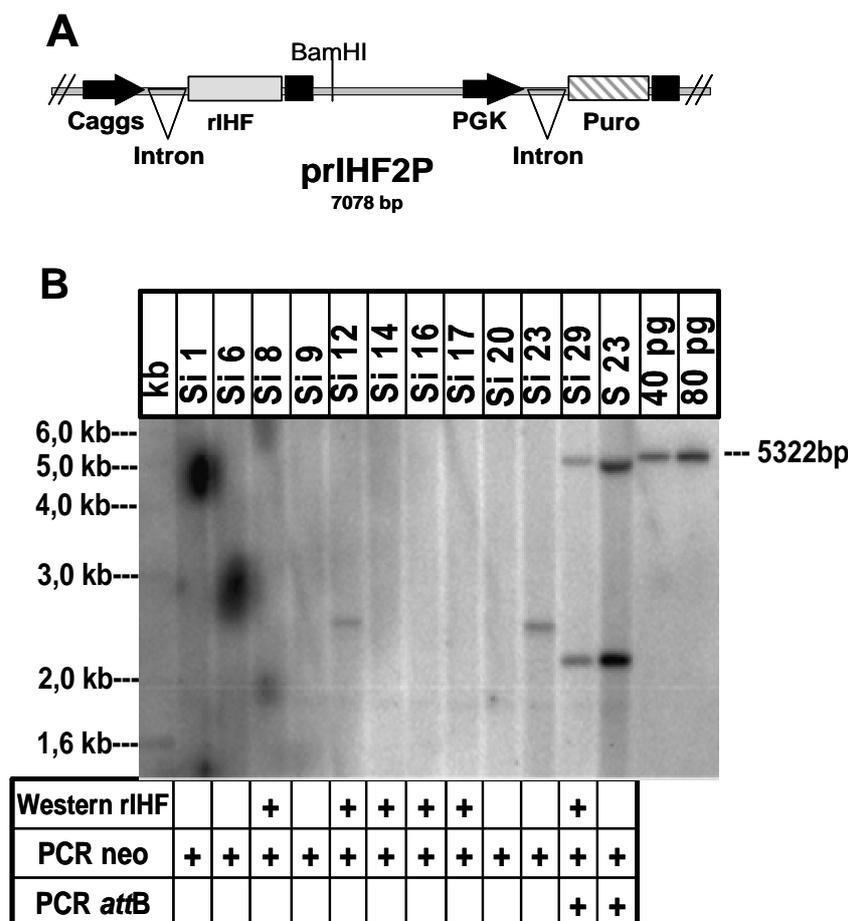


Abb. 4-16: Identifizierung rIHF-stabiler S23-Zelllinien.

(A) Schematische Darstellung des Vektors für die Generierung der rIHF-stabilen S23-Zelllinien, welche schon 1-2 Kopien der *attB*-Region tragen (siehe auch Abb. 4-14 und 4-10). Caggs, Cytomegalovirus *early enhancer* fusioniert an den β -Aktin-Promoter aus Huhn; PGK, Phosphoglycerat-Kinase-Promoter; Puro, Puromycin Resistenzgen. Die schwarzen Boxen stehen für Polyadenylierungs-Sequenzen **(B)** Analyse der rIHF-stabilen S23-Zelllinien (Si..). S23-Zellen wurden mit prIHF2P-Vektor transfiziert und unter Selektionsdruck gebracht (20 µg/ml Puromycin). Nach zwei bis vier Wochen wurden resistente Klone vereinzelt und expandiert. In einer *Western* Analyse wurde erst die Expression des rekombinanten rIHF untersucht. Diverse PCRs wurden durchgeführt, um das rIHF-Gen,

das Neo-Gen und die *attB*-Region nachzuweisen (siehe Tabelle unter der Abbildung). Positive Signale sind für die entsprechenden Zelllinien mit einer „+“ dargestellt. Die *Southern* Analyse wurde wie in Abb. 4-10 und 4-14 beschrieben durchgeführt. Bis auf Si29 haben alle neuen Zelllinien die ursprünglichen *attB*-Regionen der S23 Ausgangszelllinie verloren. Kb, Kb-Marker; 40 bzw. 80, Auftrag von 40 bzw. 80 pg linearisiertem pTK*HattB*-Vektor (Einheitslänge).

4.2.2.2 *Targeting* einer rHF/*attB*-stabilen HeLa-Zelllinie mit pCMVSS*attP*_{mut}

Aus der im vorhergegangenen Kapitel beschriebenen Generierung von HeLa-Zelllinien mit stabil integriertem rHF und *attB* ist nur eine Zelllinie, die Si29, hervorgegangen (Abb. 4-16). Diese Zelllinie wurde für *Targeting*-Experimente mit Inth-218 und pCMVSS*attP*_{mut} eingesetzt. Ziel war die Steigerung der intermolekularen chromosomalen Rekombinationsaktivität dieser Integrase-Mutante durch die konstitutive Anwesenheit von rHF. Es wurden drei *Targeting*-Experimente durchgeführt. Dabei wurden jeweils 40 µg Expressionsvektor für In-h/218 oder Mock-Vektor (als negative Kontrolle) zusammen mit 20 µg *Targetvektor* (pCMVSS*attP*_{mut}) in die jeweiligen Zelllinien eingebracht. Nach ca. zwei Wochen Doppelselektion mit Puromycin [5 µg/ml] und Geneticin [500 µg/ml] starben in zwei von vier Experimenten alle Zellen. In den anderen zwei Experimenten wuchsen auch Klone auf den Platten, die mit dem negativen Kontrollansatz ausgesät wurden. Die Anzahl der Klone war sehr gering (zwischen 1-8). Die Vereinzlung und Expansion für eine weitergehende Untersuchung schlug fehl. Die Zellen waren sehr empfindlich und starben nach der Vereinzlung oder wuchsen nicht.

4.2.2.3 Genomische Inversion in HeLa-Reporterzelllinie in An- und Abwesenheit von rHF

Der Einfluss von rHF auf die integrative Rekombination zwischen *attB* und *attP* wurde in der HeLa-Reporterzelllinie H₈B untersucht, welche von Nicole Christ zur Verfügung gestellt wurde. In dieser Zelllinie sind ca. acht Kopien des Substrats pGFP*attB/attP* an zufälligen Positionen im Genom integriert (Daten nicht gezeigt). Die korrekte Rekombination zwischen den beiden *att*-Regionen führt zur Inversion des dazwischen liegenden eGFP-Gens und zu einem veränderten Nco I-Restriktionsmuster. Das rekombinierte Produkt mit einer Größe von zwei Kb lässt sich so vom nicht-rekombinierten Substrat unterscheiden und mit dem eGFP-Gen als Sonde im *Southern Blot* nachweisen (Abb. 4-17A und B). Zellen der H₈B-Linie wurden zuerst mit 60 µg des Expressionsvektors für rHF oder des Kontrollvektors pCMV elektroporiert und in geeigneter Dichte auf 6-Loch Platten ausgesät. Die Transfektionseffizienz betrug dabei ~90 % (Daten nicht gezeigt). Nach 24 h erfolgte durch Lipofektion das Einbringen von 1,5 µg Expressionsvektor der jeweiligen Integrase oder von pCMV als negative Kontrolle (Abb. 4-17C, Tabelle). Hierbei

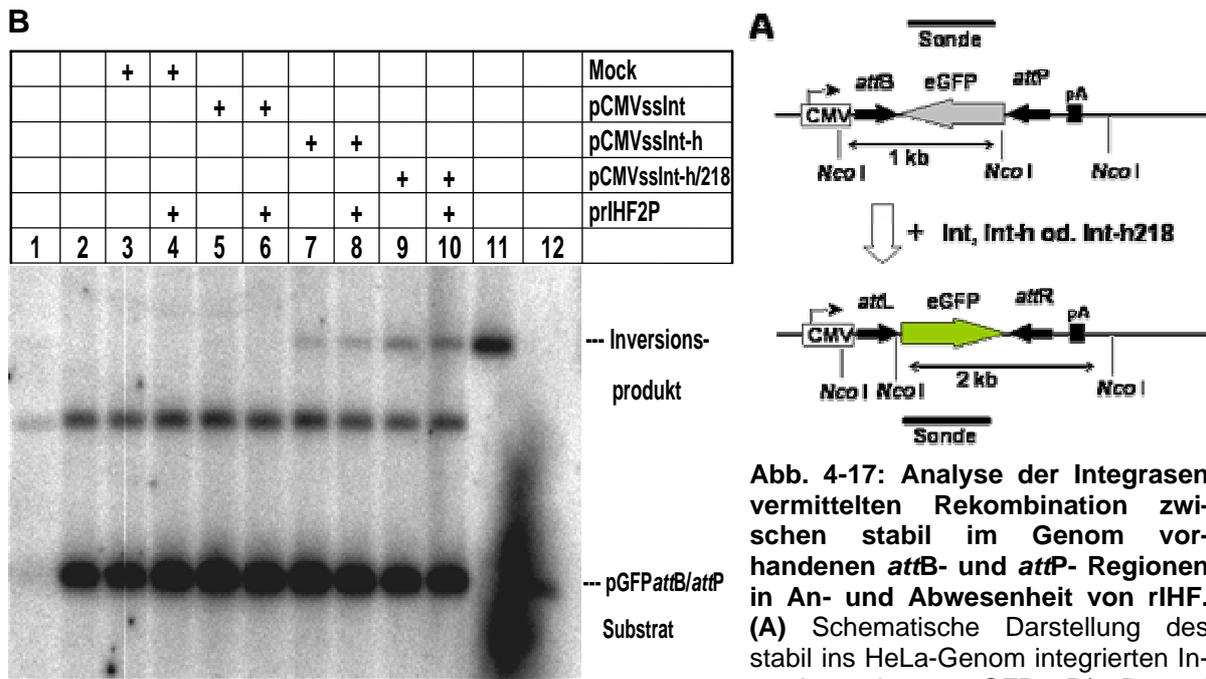


Abb. 4-17: Analyse der Integrasen vermittelten Rekombination zwischen stabil im Genom vorhandenen *attB*- und *attP*-Regionen in An- und Abwesenheit von rIHF. (A) Schematische Darstellung des stabil ins HeLa-Genom integrierten Inversionssubstrats pGFPattP/*attB* und

des entsprechenden Rekombinationsprodukts nach Integrase-Expression. Der CMV-Promoter ist durch ein Rechteck und die Rekombinationssequenzen *attB* und *attP* sind durch dicke schwarze Pfeile dargestellt. Das eGFP-Gen ist durch einen grau bzw. grün schattierten Pfeil dargestellt. pA, Polyadenylierungs-Sequenz. NcoI: Restriktionsschnittstelle. **(B)** *Southern* Analyse der Rekombination in der Reporterzelllinie H₈B sieben Tage nach Transfektion der angegebenen Expressionsvektoren für die Integrasen (Details siehe Text). Die genomische DNA wurde isoliert, mit NcoI geschnitten und die Rekombination mit einer eGFP-Sonde analysiert. Spur 1 Kb-Marker; Spur 2, NcoI geschnittene DNA aus nicht transfizierten H₈B-Zellen; Spur 11 und 12, mit NcoI geschnittenes, rekombiniertes und nicht-rekombiniertes Plasmid pGFPattB/*attP*.

betrug die Effizienz ca. 50 % (Daten nicht gezeigt). Nach weiteren 7 Tagen, in denen die Zellen einmal passagiert wurden, erfolgte die *Southern* Analyse (Abb. 4-17C). Nur die Ko-Transfektion der Expressionsvektoren für Int-h und Int-h/218 ergab eine detektierbare Bande des Rekombinationsproduktes (Abb. 4-17C, Spuren 7-9). Im Durchschnitt betrug die mit dem Phosphoimager gemessene Intensität des Signals für Int-h ~4 % und für Int-h-218 ~8 % der Gesamtradioaktivität (Daten nicht gezeigt). Berücksichtigt man die Transfektionseffizienz von ca. 50 % bei der Lipofektion so wurden zwischen 8 und 16 % der genomischen Substrate rekombiniert. Ein bedeutender Unterschied zwischen der An- oder Abwesenheit von rIHF konnte nicht festgestellt werden.

Die zusätzliche Bande, die in allen Ansätzen der H₈B-Linie auftaucht, beruht wahrscheinlich auf der Integration eines verkürzten Substrats. Sie wurde für die Quantifizierung mit einbezogen, da sie ebenfalls noch *attB* und *attP* enthalten und als Substrat für die Rekombination dienen könnte. Das erzeugte Produkt könnte eventuell nicht vom Produkt des vollständigen Substrates unterschieden werden. Die Integrität der Produkte, die durch die Int-h bzw. Int-h/218 vermittelte Rekombination entstehen, wurde durch PCR und Sequenzierung in früheren Experimenten bestätigt

und wurde daher nicht noch einmal untersucht (Lorbach et al., 2000). Ebenso konnte durch PCR-Analyse die Rekombination durch die wild-typ-Integrase nachgewiesen werden (Nicole Christ, Dissertation 2002). Die Sensität des *Southern Blots* war dafür nicht ausreichend (Abb. 4-17C, Spuren 5 und 6).

4.2.3 Targeting-Strategie für eine *attB*-ähnliche Sequenz (*attH2*) im menschlichen Genom

Der Ansatz des genomischen *Targetings* einer zuvor im Genom stabil integrierten *attB*-Sequenz (Kap. 4.2.1) erfordert die aufwendige Generierung von Zelllinien, die, wie die vorhergehenden Kapitel zeigen, nicht immer unproblematisch ist. Instabilität von Chromosomen, welche zu Genomumordnungen führen, oder negative Positionseffekte (Magliaccio et al., 2000) können die Zuverlässigkeit des Systems entscheidend beeinflussen. Für eine eventuelle gentherapeutische Anwendung des *Lambda*-Systems wäre daher die Integration von Fremd-DNA in eine natürlich vorkommende Sequenz des menschlichen Genoms von Vorteil. Solche pseudo Rekombinationsregionen wurden schon für die Cre-Rekombinase (Thyagarajan et al., 2000), und für die Integrasen ϕ C31 und R4 entdeckt (Thyagarajan et al., 2001;

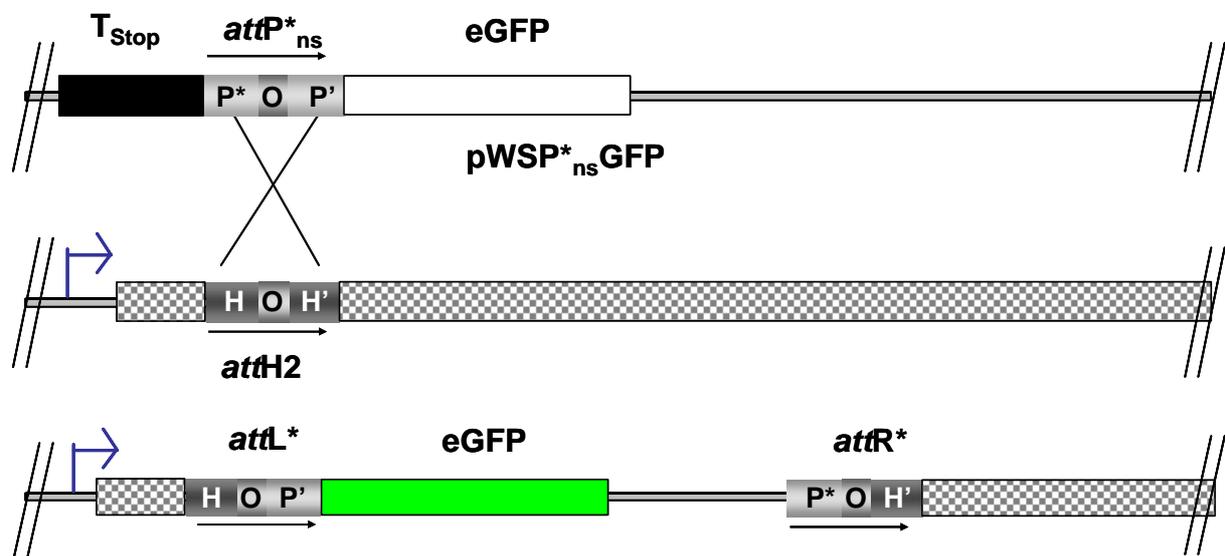


Abb. 4-18: Targeting-Strategie für die *attB*-ähnliche *attH2*-Region der HeLa-Zellen. Relevante genetische Elemente des *Targetvektors*, der Zielsequenz und des Integrationsproduktes sind schematisch dargestellt. Die Rekombinations-Kassette wird durch den endogenen MCT5-Promoter exprimiert. Erfolgreiche Integration soll eine Ausprägung des MCT5-eGFP Fusionsproteins und somit detektierbare Fluoreszenz bewirken. *attP*ns*, modifizierte Phagen-*attachment*-Region, die Mutationen in der *overlap*-Region (O) besitzt und in der im P'-Arm die Stopcodons entfernt wurden (siehe auch Abb. 4-19); *attH2*, im menschlichen MCT5-Gen natürlich vorkommende *attB*-ähnliche *attachment*-Region; *attL** bzw. *attR**, resultierende linke bzw. rechte *attachment* Region nach Integration; MCT5, Monocarboxylat-Transporter 5; eGFP, Reporter gen für das *enhanced green fluorescent Protein*; T_{Stop}, transkriptionelles Stoppsignal; der gekrümmte Pfeil steht für den endogenen MCT5-Promoter; der gerade Pfeil gibt die Orientierung der zugehörigen *attachment*-Region an; gemusteter Balken, genomischer MCT5 Lokus.

Olivares et al., 2001). Die relative kurze Sequenz der *attB*-Region bot sich an, um nach ähnlichen Sequenzen im menschlichen Genom zu suchen. Es wurden keine identischen Regionen gefunden aber ähnliche (Peter Dröge, unveröffentlichte Daten). Unter diesen scheint die sogenannte *attH2*-Region (Nicole Christ, Dissertation 2002; Abb. 4-19) sehr interessant zu sein. Sie ist Bestandteil einer in diversen Organen gebildeten menschlichen mRNA, die für den Monocarboxylat-Transporter 5 (MCT5) kodiert (Halestrap & Price, 1999, wurde in Price et al., 1998) noch als MCT4 bezeichnet). Auch in menschlichen Gebärmutterhals- und Nierenzelllinien (HeLa bzw. 293T) konnte die *attH2*-tragende mRNA nachgewiesen werden (Christ, Dissertation 2002). Das MCT5-Gen liegt auf dem Chromosom 1 von Homo sapiens (*Accession*-Nummer AC025987 in der EMBL-Datenbank). *attH2* ist dabei Bestandteil eines Exons, welches aufgrund unvollständiger Sequenzdaten des betreffenden Chromosomenabschnittes nicht genau identifiziert werden konnte. Vergleicht man die Sequenzen von *attH2* und *attB*, so ergeben sich vier Nukleotidunterschiede, wovon drei in der rechten *core*-Bindungsstelle und eine am 5'-Ende der *overlap*-Region liegen. Die in Abb. 4-18 dargestellte *Targeting*-Strategie für die menschliche *attH2*-Region hat Ähnlichkeit mit so genannten *promoter-trap*-Strategien (Friedrich & Soriano, 1991; Soriano et al., 1991). Das gewählte Reportersystem beruht auf einem MCT5-eGFP-Fusionsprotein, welches nach korrekter Integration des *Targetvektors* über den endogenen MCT5-eigenen Promoter exprimiert werden soll. Mittels FACS-Analyse oder Fluoreszenz-Mikroskopie könnten dann die Rekombinationsereignisse analysiert werden.

	→										←											
<i>attP</i>	C	A	G	C	T	T	T	T	T	T	A	T	A	C	T	A	A	G	T	T	G	
<i>attB</i>	C	T	G	C	T	T	T	T	T	T	A	T	A	C	T	A	A	C	T	T	G	
<i>attH2</i>	C	T	G	C	T	T	T	<u>C</u>	T	T	A	T	A	C	<u>C</u>	A	A	G	T	<u>G</u>	G	
<i>attP</i> ^{*_{ns}}	C	A	G	C	T	T	T	C	T	T	A	T	A	C	A	A	A	G	T	T	G	
		*	*	*	*	*	*		*				*			*	*	*	*	*	*	*

Abb. 4-19: Sequenzvergleich des common core von *attP*, *attB*, *attH2* und *attP*^{*_{ns}}. Die *overlap*-Region ist grau unterlegt. Die von *attP* und *attB* abweichenden Nukleotide in *attH2* sind unterstrichen und fett gedruckt. Um *attP* an *attH2* anzupassen, sind zwei Mutationen eingeführt worden, die zur sogenannten *attP*^{*_{ns}}-Region geführt haben. Diese Mutationen sind in *attP*^{*_{ns}} fett gedruckt und betreffen nur Nukleotide, die für die Bindung der wild-typ-Integrase (Int) nicht so wichtig sind. Die horizontalen Pfeile markieren die Bindungssequenzen für Int an die *core*-Region. Für die Bindung besonders wichtige Nukleotide (Ross & Landy, 1983) sind mit (*) gekennzeichnet.

Es ist bekannt, dass die *Lambda*-Integrase einige Austausche in der *core*-Region toleriert (Nash, 1981). Um die Homologie der *overlap*-Region zu bewahren und den

Strangtausch zu erleichtern, wurde das erste T im *overlap* von *attP* gegen ein C ausgetauscht und *attP** generiert. Die für die Bindung der Integrase wichtigen Nukleotide der rechten *core*-Sequenz wurden nicht verändert. Allerdings darf für die Expression des Fusionsproteins kein Stoppcodon hinter der *overlap*-Region folgen, daher wurden sie in der ersten Position der rechten *core*-Bindungsstelle (Abb. 4-19) und im restlichen P'-Arm (Daten nicht gezeigt) mutiert. Die veränderte *attP*-Region wurde *attP**_{ns} genannt. Der *Targetvektor* (Abb. 4-18, oben) trägt ein promoterloses eGFP-Gen, welches direkt an das *attP**_{ns} anschließt. Die korrekte Integration des Vektors positioniert das eGFP-Gen in das offene Leseraster des MCT5-Gens. Im Gegensatz zum *promoter-trap* soll nur die Integration in die *attH2*-Region des MCT5-Gens ein Signal ergeben. Um den Hintergrund möglichst gering zu halten, wurde direkt vor das 5'-Ende der *attP**_{ns}-Region ein transkriptionelles Stoppsignal kloniert.

4.2.3.1 Episomale Untersuchung zur Funktionalität der *attP**_{ns}/*attH2*

Rekombination

Das *attP**_{ns}/*attH2*-System wurde zunächst episomal getestet. Dabei sollte auch untersucht werden, ob und in wie weit die Unterschiede in der Sequenz der *core*-Region Einfluss auf die Rekombination haben. Die verwendeten Vektoren pWSPGFP bzw. pWSP*_{ns}GFP und pCMVSS*attB* bzw. pCMVSS*attH2* basieren auf dem unter 4.1.3.1 vorgestellten Reportersystem für die integrative intermolekulare Rekombination. Die *att*-Regionen sind auf zwei unterschiedlichen Plasmiden lokalisiert (Abb. 4-20A). Der Vektor mit der *attB*- bzw. der *attH2*-Region trägt jeweils den Promoter. Auf dem anderen Vektor befindet sich hinter einem Transkriptionsstopp die *attP*- bzw. die *attP**_{ns}-Region, gefolgt von dem promoterlosen Reporter gen für eGFP. Die korrekte Rekombination zwischen beiden Molekülen führt zur Umlagerung des CMV-Promoters vor das offene Leseraster des eGFP-Gens und somit zu dessen Expression. Diese kann mit Hilfe des FACScan quantifiziert werden und wird als Prozent grüne Zellen in der Tabelle in Abb. 4-20 angegeben. Die Substratvektoren wurden, wie in der ersten Zeile der Tabelle angegeben (Abb. 4-20B), in unterschiedlichen Kombinationen mit dem Expressionsvektor für Int-h/218 oder pCMV (Mock) als negative Kontrolle in HeLa-Zellen kotransfiziert. Der Anteil der eGFP-ausprägenden Zellen wurde 48 h nach Transfektion mittels FACScan ermittelt. Die Effizienz wurde für jedes Experiment einzeln bestimmt und lag im Durchschnitt bei 85 % (Daten nicht gezeigt).

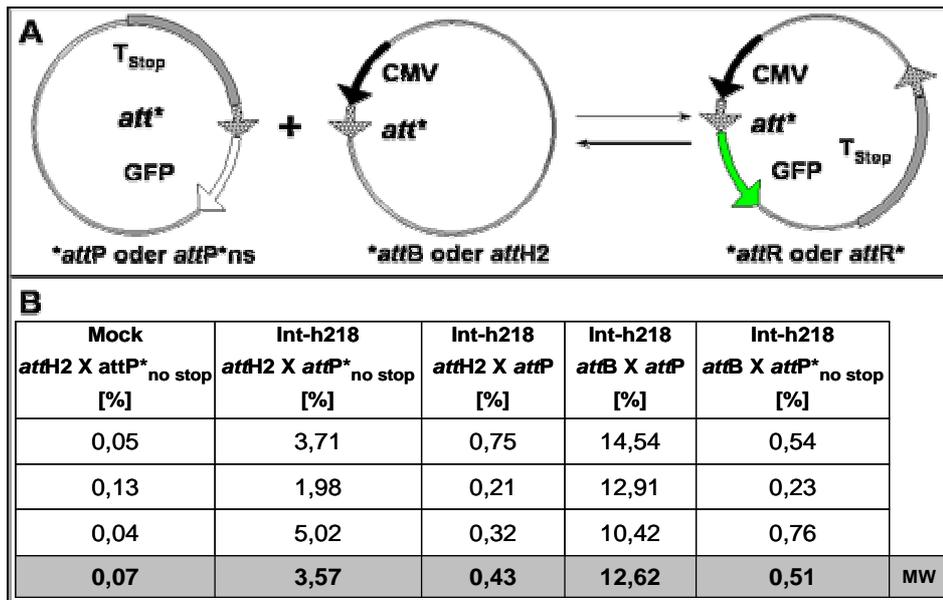


Abb. 4-20: Spezifität der att-Regionen. (A) Substratvektoren für die intermolekulare Rekombination, die entweder *attB* bzw.: *attH2*, oder *attP* bzw. *attP*_{ns}* tragen und in verschiedenen Kombinationen zur Rekombination eingesetzt werden. Der CMV-Promoter wird nach erfolgreicher Rekombination vor das eGFP-Gen positioniert, was zu dessen Ausprägung führt. CMV, Cytomegalovirus Promoter; *T_{Stop}*, transkriptionelles Stoppsignal; eGFP, *enhanced green florescent protein*. **(B)** Verschiedene Kombinationen von Substratvektoren wurden mit dem Expressionsvektor für Int-h/218 oder mit pCMV (Mock), als negative Kontrolle, in HeLa-Zellen kotransfiziert. 48h nach der Elektroporation wurden die Zellen durch FACS analysiert. Der Anteil an GFP-ausprägenden Zellen ist für jeden Ansatz und jedes Experiment einzeln angegeben. Die Mittelwerte (MW) aus diesen drei Versuchen sind in der grau unterlegten Zeile dargestellt.

In der Abbildung 4-20B sind die Ergebnisse dreier Experimente und die entsprechenden Mittelwerte dargestellt. Man erkennt, dass nur aufeinander zugeschnittene *att*-Paare, *attH2/attP*_{ns}* und *attB/attP* effizient rekombiniert wurden (3,57 % bzw. 12,62 %). Dennoch war sie für *attB/attP* ungefähr um das vierfache höher als für *attH2 / attP*_{ns}*. Der Unterschied von zwei (bei *attP / attP**) bzw. drei (bei *attB / attH2*) Nukleotiden in der *core*-Region führte zu einer deutlichen Reduktion der Rekombinationsereignisse auf ein marginales Level (0,4-0,5 %). Die Integrität aller Rekombinationsprodukte wurde durch Sequenzierung bestätigt.

4.2.3.2 Targeting der genomischen *attH2*-Region mit und ohne rHF

Mit den folgenden Versuchen sollte getestet werden, ob die Integrase-Mutante Int-h/218 in der Lage ist, in die in menschlichen Zellen natürlich vorkommende genomische *attH2*-Region des Monocarboxylat Transporter 5 zu integrieren. Vorab wurde mit episomalen Substratvektoren gezeigt, dass eine angepasste *attP**-Region durch Int-h218 mit *attH2* rekombiniert wird. Allerdings erreicht diese Reaktion nur ein Viertel der Effizienz verglichen mit *attB* und *attP* (Kap. 4.2.3.1, Abb. 4-20B). Als

Reportersystem für das chromosomale *Targeting* dient ein MCT5-eGFP-Fusionsprotein, welches, nach korrekter Integration des zuvor beschriebenen pWSP*_{ns}GFP-*Targetvektors*, über den endogenen MCT5-eigenen Promoter exprimiert werden soll (Kap. 4.2.3.1 und Abb. 4-18, Strategie). *Targetvektor* (20 µg) und Expressionsvektor für Int-h/218 (40 µg) bzw. pCMV (als Negativ-Kontrolle) wurden mittels Elektroporation in HeLa-Zellen eingebracht. Nach einer Woche erfolgte die FACS-Analyse oder die FACS-Sortierung der transfizierten Zellen.

Abb. 4-21 stellt ein repräsentatives Beispiel für die erhaltenen Ergebnisse dar. Der detektierte Anteil an fluoreszierenden Zellen in Ansätzen mit Integrase lag meist bei marginalen 0,02 % (einmalig 0,03 %).

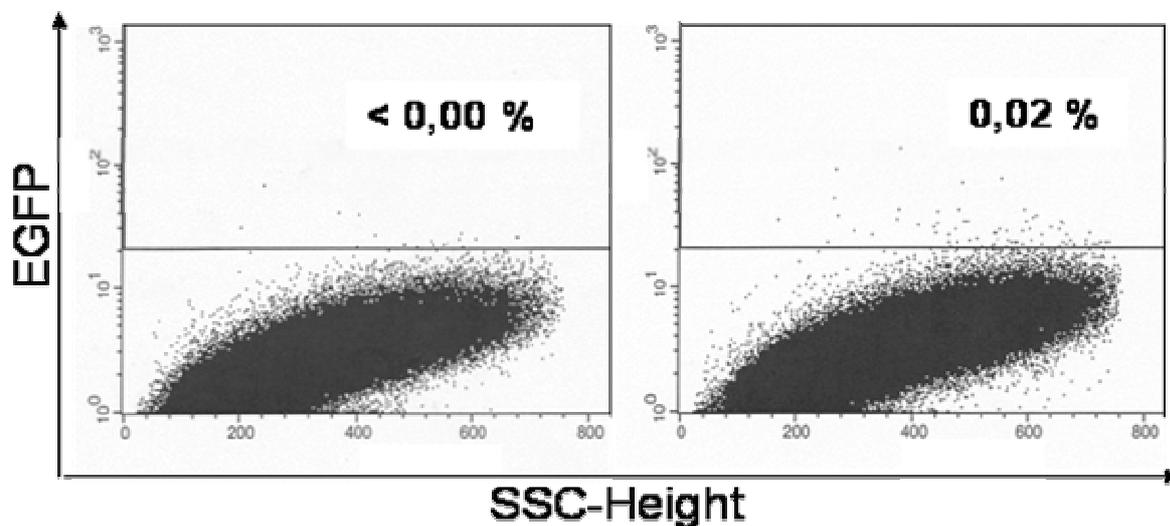


Abb. 4-21: Repräsentative Dot-Plot FACS-Analyse der Integration in die genomische *atfH2*-Region. Mock-Vektor ohne Integrase (linke Grafik) oder Expressionsvektor für Int-h/218 (rechte Grafik) wurden mit dem *Targetvektor* pWSP*_{ns}GFP in HeLa-Zellen eingebracht. Eine Woche nach Transfektion wurde der Prozentsatz der GFP-ausprägenden Zellen mittels FACS-Analyse oder FACS-Sortierung bestimmt. SSC, *sideward scatter*. Der Anteil an eGFP-ausprägenden Zellen ist in jeder Grafik jeweils in % angegeben.

Auch in Anwesenheit von rIHF konnte keine messbare Steigerung erreicht werden. Dabei wurde das *Targeting* sowohl bei Ko-Transfektion des Expressionsvektors prIHF2P (Abb. 4-6A) als auch in stabil rIHF-ausprägenden Zelllinien (Kap. 4.1.3.2) analysiert (Daten nicht gezeigt). Die transfizierten Zellen wurden sortiert und expandiert. Unter dem Mikroskop konnten danach keine fluoreszierenden Zellen beobachtet werden. Die PCR-Analyse sortierter Zellen ergab nur ein sehr schwaches Signal der erwarteten Größe (Daten nicht gezeigt). Das PCR-Fragment wurde in den T-easy Vektor (Promega) umklontiert und die resultierenden Vektoren aus mehreren Bakterienklonen sequenziert. Ein richtiges Rekombinationsprodukt konnte jedoch nicht nachgewiesen werden (Daten nicht gezeigt). In Stammzellen der Maus erreicht die intramolekulare integrative

attB/attP-Rekombination durch Int-h/218 eine Effizienz von ca. 0,01 bis 0,1%. Dabei ist je eine Kopie der beiden *att*-Regionen in einem Abstand von ca. 3 Kb im *ROSA26*-Lokus integriert (Christ & Dröge, 2002). Die durchgeführten *Targeting*-Experimente könnten daher die um ein Viertel, oder auch höher, reduzierte Rekombinationseffizienz von *attH2/attP*_{ns}* widerspiegeln. Ob die Integration in die *attH2*-Region tatsächlich erfolgt, ist jedoch nicht eindeutig erwiesen.

4.3 Generierung einer rIHF transgenen Maus

In Kapitel 4.1.3 wurde der Einsatz des rekombinanten IHFs in HeLa-Zellen beschrieben. Dabei wurden die Langzeit-Expression, die Lokalisation und der Einfluss auf die Integrase vermittelte Rekombination betrachtet. IHF spielt jedoch nicht nur eine Rolle in der Rekombination. Auch an anderen zellulären Prozessen, wie Transkription, Replikation und DNA-Verdichtung, ist IHF beteiligt (Übersichtsartikel: Nash, 1996; Goosen & van de Putte, 1995; Freundlich et al., 1992; Friedman, 1988). Dabei beruht die Ausübung seiner biologischen Funktion auf die Bindung und Biegung von DNA an spezifischen Sequenzen. Gleichwertige Proteine in Eukaryoten, die wie IHF in Transkription und Rekombination eingebunden sind, sind die so genannten HMG-Proteine. Im Gegensatz zu IHF ist in den meisten Fällen eine Spezifität für bestimmte Sequenzen nicht oder nur sehr gering vorhanden (Travers, 1997; Giese et al., 1992). Zudem ist ihre Fähigkeit zur DNA-Biegung deutlich geringer (<120°) als die von IHF (>160°) (Giese et al., 1992). Bioinformatische Untersuchungen führten zur Entdeckung von potenziellen IHF-Bindungsstellen in Genomen diverser Organismen, wie Mensch, Maus und HIV1-Virus (Peter Dröge, unpublizierte Ergebnisse). Um weitere *in vivo* Studien mit dem rekombinanten IHF durchzuführen, sollte neben rIHF-transgenen HeLa-Zelllinien eine transgene Maus generiert werden. Die Tiere wurden mittels Pronukleusinjektion von der Firma Memorec hergestellt (Abb. 1-7). Das Fragment für die Injektion ist in Abb. 4-22A schematisch abgebildet. Wie schon für prIHF2P dargestellt, erfolgt die Ausprägung des rIHF-Gens über den Caggs-Promoter, der in Mäusen eine ubiquitäre Expression vermittelt (Okabe et al., 1997). Zwischen dem 3'-Ende des Gens und der Polyadenylierungs-Sequenz wurde eine von FRT-Sequenzen flankierte IRES-eGFP-Reporter-kassette eingeführt. Die IRES (*internal ribosomal entry site*, Clontech; Jackson, et al., 1990; Jang et al., 1988) ermöglicht eine

zeitgleiche Translation des eGFP- und des rIHF-Gens von einer mRNA. Die Kasette sollte später die Identifizierung transgener, rIHF-ausprägender Tiere erleichtern und

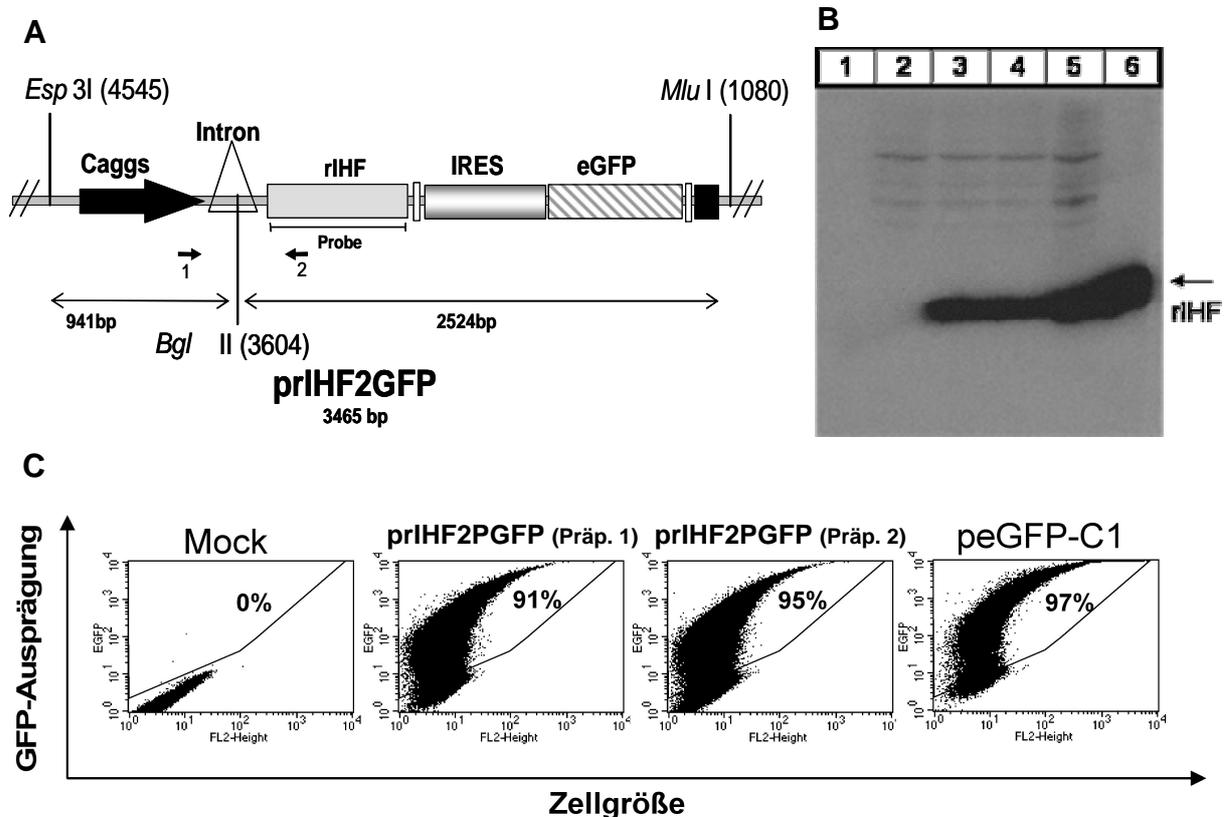


Abb. 4-22: Expressionsvektor zur Generierung der transgenen rIHF-Mäuse. (A) Schematische Darstellung der relevanten genetischen Elemente des Expressionsvektors für rIHF. Caggs, Cytomegalovirus *early enhancer* fusioniert an β -Aktin-Promoter aus Huhn; rIHF, codierende Sequenz des rekombinanten IHF; IRES, *internal ribosomal entry site*; Probe, Sonde für *Southern*; eGFP, *enhanced green fluorescent* Protein; schwarze Box, Polyadenylierungs-Sequenz; weiße schmale Box, Erkennungs-Sequenzen der FLP-Rekombinase (FRT), für die eventuelle Entfernung der IRES-eGFP-Kasette in den transgenen Tieren. Die mit 1 und 2 beschrifteten Pfeile geben die Position der Intron-übergreifenden Primer für die nachfolgende RT-PCR an. Zur Generierung der Tiere wurde nur das dargestellte *Esp* 3I / *Mlu* I Fragment benutzt. **(B)** Zum Nachweis der rIHF-Expression in HeLa-Zellen wurden HeLa-Zellen mit den angegebenen Vektoren transfiziert. Nach 48 h wurden Zelllysate von ca. 8×10^5 Zellen hergestellt und in einer 12,5 %igen SDS-Page aufgetrennt. Für die Detektion des rIHF wurde in der *Western Blot* Analyse ein Antikörper gegen wild-typ-IHF (S.D. Goodman, Los Angeles) eingesetzt. Spur 1, Marker; Spur 2, mit pCMV-Vektor transfizierte HeLa-Zellen (negative Kontrolle), Spur 3, mit Klon 5 von prIHF2GFP transfizierte Zellen; Spur 4 mit Klon 10 von prIHF2GFP transfizierte Zellen (positive Kontrolle); Spur 5, mit prIHF2P (s. Abb. 4-6) transfizierte Zellen; Spur 6, gereinigtes rIHF (~10 ng). **(C)** Dot-Plot FACS-Analyse zur Kontrolle der eGFP-Expression. 48 h nach Transfektion mit den angegebenen Vektoren, wurden die HeLa-Zellen geerntet im FACScan analysiert. Die y-Achse zeigt die Anzahl der eGFP-ausprägenden Zellen, die x-Achse den FSC, *forward scatter* (als Maß für die relative Zellgröße). Die jeweiligen prozentualen Werte sind in den einzelnen *Dot-Plots* angegeben.

könnte, wenn experimentell erforderlich, mit der FLP-Rekombinase über die FRT-Sequenzen deletiert werden. Vor der Generierung der transgenen Mäuse wurde die Funktionsfähigkeit des Vektors in HeLa-Zellen getestet. Der Nachweis der rIHF-Expression wurde durch *Western Blot* Analyse erbracht (Abb. 4-22B). HeLa-Zellen wurden hierfür mit den angegebenen Vektoren transfiziert. Nach 48 h wurden

Zelllysate von von ca. 8×10^5 Zellen hergestellt und in einer SDS-Page aufgetrennt. Für die Detektion des rIHF wurde in der *Western Blot* Analyse Antikörper gegen wildtyp-IHF (S.D. Goodman, Los Angeles) eingesetzt. Als negative Kontrolle fungierten mit pCMV-Vektor (Mock) als positive Kontrollen mit prIHF2P transfizierte HeLa-Zellen und gereinigtes rIHF (~10ng). Die Ausprägung von rIHF in HeLa-Zellen wurde für zwei Vektor-Präparationen aus verschiedenen Bakterienklonen (Präparationen 1 und 2) des generierten prIHF2GFP gezeigt (Abb. 4-22B, Spuren 4 und 5) und war vergleichbar mit der von prIHF2P (Spur 6). Zur Kontrolle der IRES-vermittelten eGFP-Expression wurde 48 h nach Transfektion mit den angegebenen Vektoren eine FACS-Analyse durchgeführt. Abb. 4-22C zeigt die dazugehörigen *Dot-Plots* wobei die Anzahl der eGFP-ausprägenden Zellen jeweils als prozentualer Wert angegeben ist. Sie unterschied sich für die mit prIHF2GFP transfizierten Zellen (91 % und 95 %) nicht wesentlich von der mit peGFP-C1 (positive Kontrolle, 97 %) transfizierten. Allerdings schienen die eGFP-ausprägenden Zellen bei der Expression über die IRES-Sequenz weniger hell zu sein als über den CMV-Promoter in peGFP-C1. Vermutlich ist das ein Indiz dafür, dass die Menge an eGFP pro Zelle etwas geringer war. Die Ausprägung des eGFP konnte auch nachgewiesen werden, wenn nur das *Esp 3I / Mlu I* Fragment in die Zellen transfiziert wurde (Daten nicht gezeigt).

4.3.1 Analyse der rIHF-transgenen Mäuse

Zur Identifizierung transgener Elterntiere bzw. ihrer Nachkommen wurde sowohl über PCR-Analyse (Daten nicht gezeigt) als auch über *Southern Blot* (Abb. 4-23) die Integration des *Esp 3I / Mlu I* Fragmentes (Abb. 4-22A) ins Mausgenom nachgewiesen. Unter den 83 Tieren der F0 Generation waren nur 12 transgen für rIHF und wurden weiter verpaart. Ein Elterntier verstarb und drei gaben keine transgenen F1 Tiere. Die Nachkommen der restlichen 8 transgenen Elterntiere zeigten in der *Southern* Analyse das gleiche Bandenmuster wie die entsprechenden Eltern (Daten nicht gezeigt). Die *Southern* Analyse der untersuchten Nachkommen (für jedes Elterntier mindestens einer) ist in Abb. 4-23 dargestellt. Die genomische DNA der Tiere wurde aus den Schwanzspitzen isoliert und mit *Bgl II*, einem Restriktionsenzym, welches nur einmal in der eingeführten DNA Sequenz schneidet, behandelt (siehe auch Abb. 4-22). Die Anzahl der integrierten Kopien des rIHF ist unterschiedlich. Berücksichtigt man die Intensität der Mehrfachintegrationsbande, so liegt sie zwischen einer (Abb. 4-23, Spur 7) und ca. sieben oder acht Kopien (Abb.

4-23, Spuren 8 und 9).

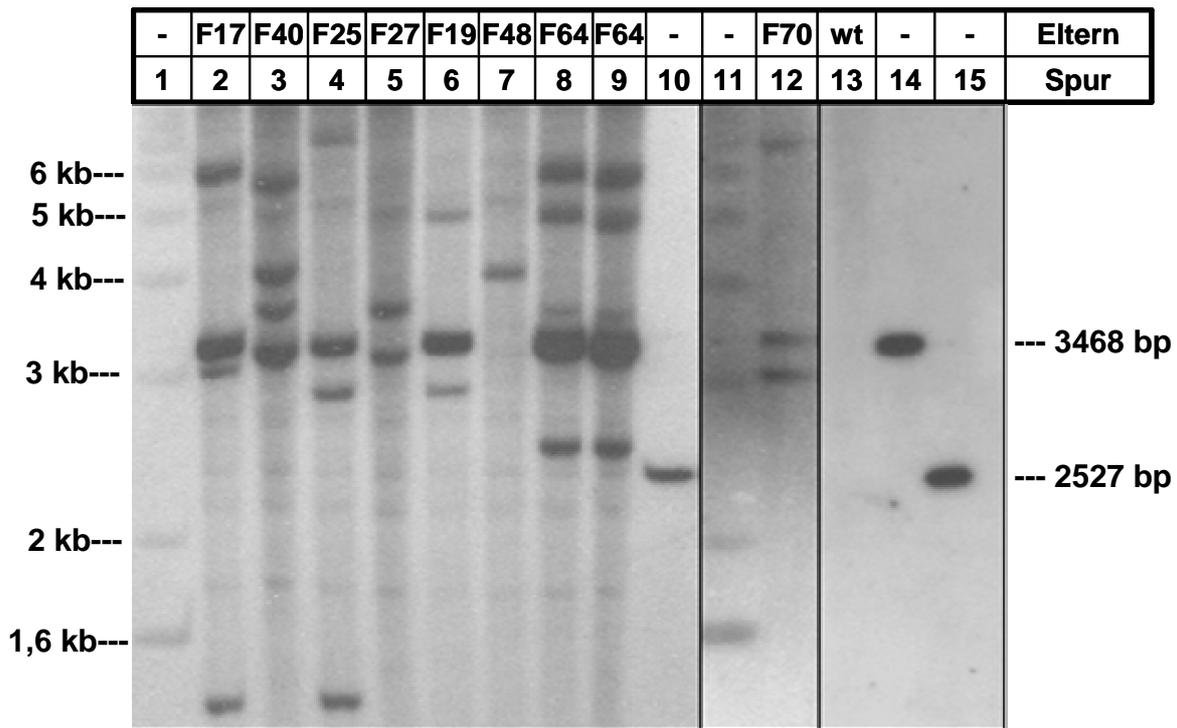


Abb. 4-23: Southern Analyse der untersuchten Nachkommen. Die genomische DNA der Tiere wurde aus den Schwanzspitzen isoliert und mit *Bgl* II, einem Restriktionsenzym, welches nur einmal in der eingeführten DNA Sequenz schneidet, behandelt (siehe auch Abb. 4-22). Spuren 1 und 11, Kb-Marker; Spur 14, Auftrag von 40 pg prIHF2GFP-Fragment; Spuren 10 und 15, Auftrag von 40 pg mit *Bgl* II verdautem prIHF2GFP-Fragment (zu den resultierenden Banden siehe Abb. 4-22); f, founder (Elterntier); wt, wild-typ-Maus; in Spur 8 und 9 ist die DNA eines männlichen und eines weiblichen Nachkommens vom Elterntier 64 aufgetragen; Spuren 1-10 sind von einem Gel, Spuren 11-15 sind ebenfalls von einem Gel wobei zwischen 12 und 13 für die Abbildung Spuren ausgelassen wurden.

Unter dem Fluoreszenzbinokular zeigten weder Schwanzspitzen noch Organe, der untersuchten Mäuse detektierbare eGFP-Expression. Auch die FACS Analyse isolierter Milzzellen bestätigte dieses negative Ergebnis (Daten nicht gezeigt).

Die Organe der F1 Nachkommen wurden daraufhin auf die Ausprägung von rIHF untersucht. Ein repräsentativer *Western Blot* eines weiblichen Nachkommens ist in Abb. 4-24 gezeigt. Pro Spur wurden ca. 180 µg Rohlysate aufgetragen, als positive Kontrolle dienten ~3 ng gereinigtes rIHF. Die Detektion erfolgte mit einem polyklonalen Antikörper gegen wild-typ-IHF (S.D. Goodman, Los Angeles). Ein Signal in der Größe der Kontrolle konnte in keinem der untersuchten Organe nachgewiesen werden. Die verhältnismäßig starke Kreuzreaktion mit einem unbekanntem Protein in der Leber, wurde auch bei wild-typ-Mäusen beobachtet (Daten nicht gezeigt). Um die Beladung der Spuren gegeneinander zu prüfen, wurden die Membranen nach Herstellerprotokoll *gestrippt* und nochmals mit Antikörper gegen α -Aktin 4700 (Sigma) inkubiert.

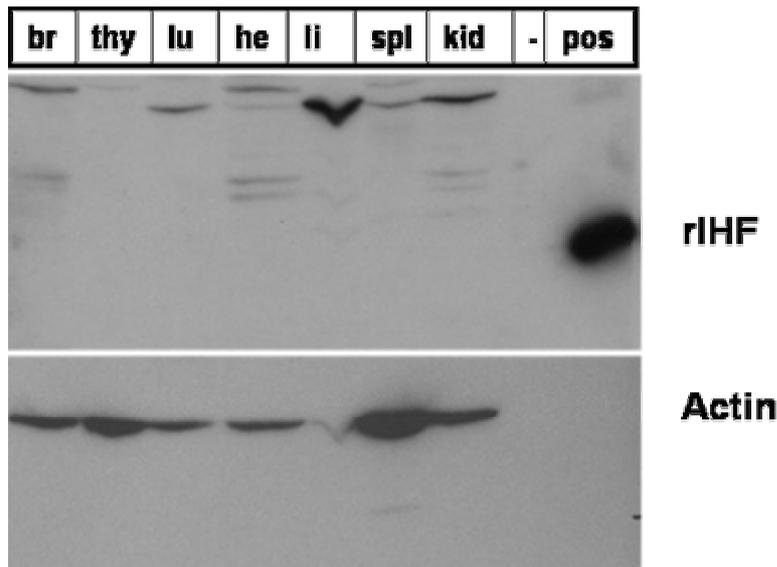


Abb. 4-24: Die *Western Blot* Analyse einer transgenen Maus zeigt keine detektierbare Menge des rIHF Proteins und unterscheidet sich nicht von der Analyse einer wild-typ-Maus (Daten nicht gezeigt). Hier wurden Rohlysate aus den Organen eines weiblichen Nachkommens von *Founder 40* gewonnen. Pro Spur ca. 180 µg aufgetragen und in einer SDS-PAGE aufgetrennt. Für den Nachweis des rekombinanten IHFs wurde in der *Western Blot* Analyse ein Antikörper gegen wild-typ-IHF (S.D. Goodman, Los Angeles) eingesetzt. Um die Beladung der Spuren untereinander zu prüfen, wurden die

Membranen nach Herstellerprotokoll *gestrippt* und nochmals mit Antikörper gegen α -Aktin 4700 (Sigma) inkubiert. Br, Gehirn; thy, Thymus; lu, Lunge; he, Herz; li, Leber; spl, Milz; kid, Niere; pos, positive Kontrolle (gereinigtes rIHF ~3 ng, (von P. Dröge)).

Um die Möglichkeit auszuschließen, dass aufgrund der Nachweisgrenze des *Western Blots* kein positives Ergebnis für die rIHF-Expression erhalten wurde, wurde eine RT-PCR Analyse zum Nachweis von rIHF-RNA in den Organen der transgenen Tiere durchgeführt. Ein Beispiel ist in Abb. 4-25 wiedergegeben. Aus den Organen eines männlichen F1 Nachkommens wurde die RNA isoliert und ein c-DNA-Synthese durchgeführt. Mit den Intron-übergreifenden Primern SS5' und IHF3-2 wurde versucht, einen Teil der Sequenz des rekombinanten IHFs zu amplifizieren (oberes Gel, für die Positionen der Primer siehe Abb. 4-22, mit 1 bzw. 2 beschriftete Pfeile). Eine korrekt *gespleisste* RNA würde ein Produkt von 490 bp ergeben, während eine Kontamination mit genomischer DNA ein ca. 720 bp großes Produkt ergäbe. Die cDNA aus HeLa-Zellen, die mit prIHF2GFP-Fragment transfiziert wurden, wurde als positive Kontrolle für erfolgreich *gespleisstes* rIHF eingesetzt. Keines der untersuchten Organe zeigt jedoch ein Signal der erwarteten Größe. Die Qualität der c-DNA wurde mit HPRT-Primern überprüft, die unter den gewählten Bedingungen nur bei Anwesenheit des korrekten Spleißproduktes eine ca. 250 bp große Bande ergeben. Diese wurde bei allen eingesetzten c-DNAs gesehen (Abb. 4-25, unteres Gel). Als negative Kontrolle wurde gezeigt, dass die Primer auf genomischer DNA nicht funktionieren (Abb. 4-25, Spur 11, unteres Gel).

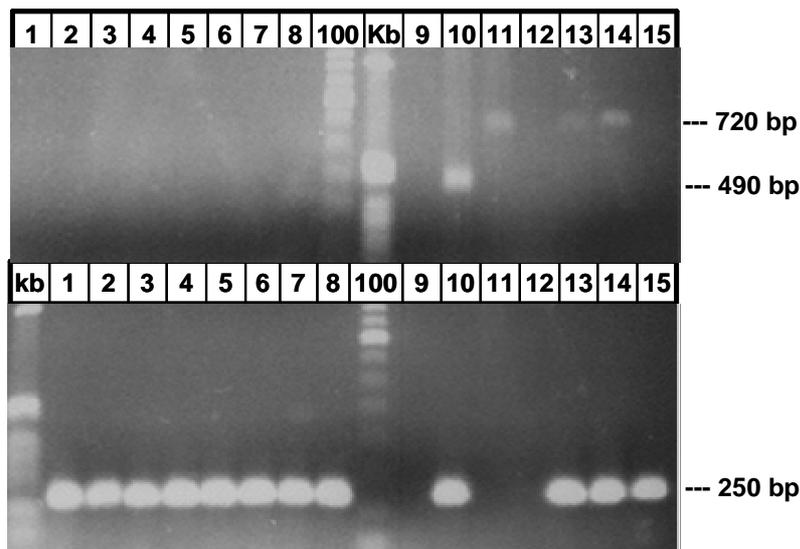


Abb. 4-25: RT-PCR Analyse zum Nachweis von rIHf RNA in den Organen einer transgenen Maus zeigt kein detektierbares Signal. Aus den Organen eines männlichen Nachkommens von Elterntier 70 wurde RNA isoliert. Nach erfolgter c-DNA Synthese wurde versucht, einen Teil der Sequenz des rekombinanten IHFs mit den Intron-übergreifenden Primern SS5' und IHF3-2 zu amplifizieren (oberes Gel, für die Positionen siehe Abb. 4-22, mit 1 bzw. 2 beschriftete Pfeile). Eine korrekt gespleisste RNA würde ein Produkt von 490 bp ergeben, während eine Kontamination mit genomischer DNA ein größeres

Produkt von ca. 720 bp ergibt (siehe Markierungen am rechten Rand, oberes Gel). Das untere Gel zeigt die Amplifizierung eines Teils der c-DNA des HPRT-Gens als positive Kontrolle für die Qualität der benutzten c-DNAs. Unter den gewählten Bedingungen ergeben die Primer HPRTse und HPRTas nur bei Anwesenheit des korrekten Spleißproduktes eine ca. 250bp große Bande. Spuren 1-10 mit DNase behandelte c-DNA aus folgenden Organen: Spur 1, Lunge; Spur 2, Thymus; Spur 3, Herz; Spur 4, Leber; Spur 5, Niere; Spur 6, Milz; Spur 7, Hoden; Spur 8, Gehirn; Spur 9, negative Kontrollreaktion ohne c-DNA; Spur 10 oberes Gel, mit prIHf2GFP-Fragment transfizierte HeLa-Zellen 48 h nach Transfektion geerntet; Spur 10 unteres Gel, c-DNA der Milz eines Nachkommen von Elterntier 48; Spuren 13-15, C-DNA von Lunge, bzw. Thymus, bzw. Herz bei der keine DNase-Behandlung durchgeführt wurde; Spur 11, genomische DNA aus der Schwanzspitze; Spur 12, kein Auftrag; Kb, Kb-Marker; 100, 100 bp-Marker.

5 Diskussion

5.1 Intermolekulare Rekombination episomaler Substrate in eukaryotischen Zellen

Aus früheren Arbeiten ist bekannt, dass die *Lambda*-Integrase-Mutanten Int-h (E174K) und Int-h/218 (E174K/E218K) die intramolekulare Rekombination episomaler Substrate auch ohne natürliche Kofaktoren in *E.coli* und in menschlichen Zellen katalysieren (Christ & Dröge 1999; Lorbach et al., 2000; Christ et al., 2002). Dabei ist unter solchen Bedingungen Int-h/218 aktiver als Int-h und die Aktivität des wild-typ-Proteins eher marginal. Als Ursache für die erhöhte Aktivität der Mutanten vermutet man eine stärkere Affinität zu den *core*-Sequenzen, an denen der Strangaustausch stattfindet (Patsey & Bruist, 1995). Ein weiterer interessanter Aspekt für die Reaktivität episomaler Substrate und für den späteren Einsatz der Integrase und ihrer Mutanten bei der gezielten Manipulation eukaryotischer Genome liegt in ihrer Fähigkeit, intermolekulare Rekombinationsreaktionen durchzuführen.

5.1.1 *Lambda* Integrase-Mutanten katalysieren in HeLa-Zellen die intermolekulare Rekombination episomaler Substrate mit vergleichbarer Effizienz wie intramolekulare Reaktionen

Zu Beginn dieser Arbeit wurde ein System entwickelt (Kap. 4.1.1), mit dem die intermolekulare Rekombination zwischen *attB* und *attP* in eukaryotischen Zellen getestet werden kann. Es sollte dazu beitragen, weitere Informationen über den Reaktionsmechanismus der Integrasen im eukaryotischen Milieu zu erlangen. Unter intermolekularen Rekombinationsreaktionen versteht man Reaktionen zwischen Rekombinationssequenzen, die auf verschiedenen DNA Molekülen (*in trans*) lokalisiert sind. Die Ergebnisse in Abb. 4-1 zeigen, dass die Integrasen die intermolekulare Rekombination in HeLa-Zellen katalysieren. Erwartungsgemäß ist Int-h/218 aktiver als Int-h und erreicht in einigen Experimenten sogar einen Anteil von 29 % eGFP ausprägenden Zellen. Die Aktivität des wild-typ-Proteins ist mit 2 bis 3 % eher marginal. Da dies auch bei Untersuchungen der exzisen Rekombination, bei der keine superhelikale Spannung der Substrate erforderlich ist, beobachtet wurde (Christ et al., 2002), liegt die Schlussfolgerung nahe, dass in eukaryotischen

Zellen Proteine nicht oder zumindest nicht in ausreichender Menge vorhanden sind, die die natürlichen Kofaktoren der *Lambda*-Integrase ersetzen können. Vergleicht man die Daten mit den Untersuchungen, die in Zusammenarbeit mit Nicole Christ durchgeführt wurden (Christ et al., 2002), so ergibt sich kein wesentlicher Unterschied zwischen den Effizienzen und Kinetiken von entsprechenden inter- und intramolekularen Reaktionen. Dies deutet darauf hin, dass in den Zellen der Großteil der produktiven Begegnungen zwischen zwei *attachment*-Regionen auf einem zufälligen Zusammentreffen beruht, eine Beobachtung, die im Einklang mit frühen *in vitro* Studien zur Synapsen-Bildung mit wild-typ-Integrase steht (Mizuuchi et al., 1980; Spengler et al., 1985). Gleiche Effizienzen zwischen episomalen inter- und intramolekularen Rekombinationen wurden auch in Studien nachgewiesen, in denen die verwandte Cre-Rekombinase in Pflanzenzellen eingesetzt wurde (Dale & Ow, 1990) oder in denen Wege der homologen Rekombination in menschlichen Zellen untersucht wurden (Rouet et al., 1994). Eine wichtige Schlussfolgerung aus den Ergebnissen der vorliegenden Arbeit und aus schon publizierten Ergebnissen ist demnach, dass unter Gleichgewichtsbedingungen in eukaryotischen Zellen keine offensichtliche Präferenz für episomale Rekombinationsreaktionen in *cis* oder *trans* existiert.

Der Zeitverlauf der untersuchten Kinetik zeigt, dass der Großteil der Rekombinationsereignisse innerhalb der ersten 24 h nach Transfektion erfolgt. Dies kann ein Hinweis dafür sein, dass episomale DNA-Fragmente innerhalb einer eukaryotischen Zelle ziemlich beweglich sind, vergleichbar mit bestimmten Proteinen (Phair & Misteli, 2000). Die Einlagerung von transfizierter DNA in kleine zelluläre Kompartimente und somit eine lokale Konzentrationserhöhung bietet eine andere mögliche Erklärung. Ob die Rekombination im Zytoplasma oder im Kern der Zellen erfolgt, ist jedoch noch nicht geklärt. Fakt ist, dass die mit der Substrat-DNA ko-transfizierten Expressionsvektoren für die Integrase zur Transkription erst in den Kern gelangen müssen. Daher ist es wahrscheinlich, dass zumindest einige Rekombinationsereignisse im Nukleus stattfinden. Für Rekombinationen, die auf genomischer Ebene stattfinden (siehe auch Kap. 5.2), ist dies jedenfalls unbestritten (Lorbach et al., 2000).

5.1.2 Einfluss der Arm-Bindung und Rückschlüsse für die Synapsenbildung

Die *Lambda*-Integrase katalysiert in ihrem natürlichen Umfeld die Integration und die Exzision des Bakteriophagen *Lambda* in bzw. aus dem Genom seines Wirtes, *Escherichia coli*. In beide Reaktionswege sind jeweils unterschiedliche Paare so genannter *attachment*-Sequenzen (*att*) involviert. Bei der Integration sind es *attB/attP* bei der Exzision *attL/attR*. Jede *att*-Sequenz enthält zwei invertierte Integrase-Bindungsstellen des *core*-Typs, die von einer 7 bp langen *overlap*-Region getrennt werden. In den Regionen, die die Bindungsstellen des *core*-Typs flankieren, besitzen alle *att*-Sequenzen noch Integrase-Bindungsstellen des Arm-Typs und weitere Bindungsstellen für die Hilfsproteine IHF, XIS und FIS (Kap. 1.2.4). Eine Ausnahme bildet das aus 21 Basenpaaren bestehende *attB*, dem zusätzliche Bindungsstellen jeder Art fehlen. Die Integrase besitzt als heterobivalentes Protein die Fähigkeit sowohl an *core*- als auch an Arm-Sequenzen zu binden. Während die C-Domäne der Integrase eine Affinität zu den *core*-Sequenzen aufweist, wird die Bindung an Arm-Sequenzen durch die N-Domäne des Proteins vermittelt (Kap. 1.2.3, Abb. 1-5). Für die simultane Bindung ist jedoch die Anwesenheit von Hilfsproteinen notwendig und beim integrativen Reaktionsweg die negative superhelikale Spannung der *attP*-DNA. An *attP*, *attL* und *attR* formen sich übergeordnete Nukleoproteinkomplexe, die so genannten Intasome. *attB* bleibt zunächst frei von Proteinen und wird erst durch das Intasom an *attP* eingefangen. Die in dieser Arbeit eingesetzten Mutanten der *Lambda*-Integrase, Int-h (E174K) bzw. Int-h/218 (E174K/E218K), können in *E. coli* und in eukaryotischen Zellen auch in Abwesenheit von Kofaktoren Rekombination katalysieren (Christ & Dröge 1999; Lorbach et al., 2000). Als Ursache wird eine höhere Affinität zu *core*-Sequenzen vermutet (Patsey & Bruist, 1995), wofür es bis heute jedoch keinen biochemischen Nachweis gibt. Welchen Einfluss die Arm-Bindung seitens der Integrase oder ihrer Mutanten auf die von ihnen vermittelte Rekombination hat, wurde in dieser Arbeit in zwei unterschiedlichen Ansätzen getestet. Erstens wurden intermolekulare Rekombinationsreaktionen mit verschiedenen *attB / attP*-Paarungen durchgeführt und zweitens wurden von Int und Int-h/218 verkürzte Proteine, denen die Fähigkeit zur Arm-Bindung fehlt, hergestellt und getestet.

Die Ergebnisse in Kap. 4.1.3.1, in dem verschiedene *attachment*-Paarungen eingesetzt wurden, zeigen, dass *attP*-Paare mindestens so effizient rekombiniert werden wie *attP / attB*-Paare. Die Rekombination von *attB*-Paaren wird jedoch nur von Int-h/218 hinreichend katalysiert (Abb. 4-2). Da *attB* nur aus *core*-Sequenzen besteht,

könnte dies ein Hinweis auf eine erhöhte Affinität für diese Sequenzen seitens Int-h/218 sein. Zwar kann auch Int-h in geringem Maße *attB*-Paare rekombinieren (Abb. 4-2), dennoch führt vermutlich erst die noch höhere Affinität von Int-h/218 zu einer effizienten Reaktion. Im Hinblick auf den Reaktionsmechanismus bietet sich eine Erklärung an, die an den Mechanismus von Cre und Flp (Landy, 1993, Voziyanov, et al., 1999) erinnert: vier Int-h/218-Monomere binden demnach an zwei *attB*-Sequenzen, und es wird durch zufälliges Zusammentreffen von zwei *attB*-gebundenen Int-Dimeren ein synaptischer Komplex ausgebildet, der den Strangaustausch katalysiert. Je schlechter die Reaktion in Abwesenheit von Arm-Sequenzen abläuft, desto mehr wird sie in ihrer Anwesenheit stimuliert. Für Int-h und Int-h/218 ergibt sich eine Steigerung der Reaktion um das vier- bis achtfache während die Randaktivität des wild-typ-Proteins vollständig von der Gegenwart der Arm-Sequenzen abhängt.

Was lässt sich daraus für die funktionelle Rolle der Arm-Sequenzen bei der Integrase vermittelten Rekombination in Eukaryoten schließen? Es konnte gezeigt werden, dass die wild-typ-Integrase in den untersuchten Rekombinationsreaktionen eine Aktivität besitzt, die nur gering über dem Hintergrund liegt. In eukaryotischen Zellen sind demnach Proteine, die die natürlichen Kofaktoren der Integrase in ihren Funktionen ersetzen könnten, nicht oder zumindest nicht in ausreichender Menge vorhanden. Daher ist es unwahrscheinlich, dass die Architektur der von Int-h und Int-h/218 an den *attachment*-Regionen gebildeten Nukleoproteinkomplexe mit der identisch ist, die vom wild-typ und den mutierten Proteinen in *E. coli* und *in vitro* mit Hilfe der natürlichen Kofaktoren erzeugt wird. Für die Stimulierung durch die *attachment*-Regionen in HeLa-Zellen bleiben zumindest drei mögliche Erklärungen: Erstens existieren in Säugetierzellen Proteine, wie z.B. HMG1 und 2, die unspezifisch an DNA binden und diese dadurch krümmen können. Bei Rekombinationsreaktionen mit wild-typ-Integrase konnte schon gezeigt werden, dass sie IHF *in vitro* teilweise ersetzen können (Segall et al., 1994). In HeLa-Zellen ist es daher denkbar, dass die simultane Bindung eines Integrase Protomers an Arm- und *core*-Sequenzen derselben *attachment*-Region nur im Zusammenspiel mit höherer Affinität der Integrase-Mutanten für *core*-Sequenzen und mit unspezifischen, Protein-induzierten DNA-Krümmungen erreicht wird. Die simultane Bindung führt wiederum zur Stabilisierung der teilweise rekombinationsaktiven Nukleoproteinkomplexe. Zweitens könnte die Aktivierung durch die *att*-Sequenzen auch in *trans* erreicht werden. Landy und Mitarbeiter konnten dies erst kürzlich *in vitro* nachweisen. Dabei stellten sie fest, dass in freier Lösung Effizienz

und Genauigkeit der Auflösung des Holliday-Intermediats durch die Zugabe von Arm-DNA gesteigert wird. Diese Steigerung ist unabhängig von Protein-Kofaktoren. Sie vermuten, dass die Arm-DNA eine intrinsische Komponente des übergeordneten Rekombinationskomplexes darstellt, die die Integraseaktivität koordiniert (Radman-Livaja et al., 2003). Drittens wird durch die starke Bindung der Integrase an Arm-Sequenzen möglicherweise die lokale Konzentration der Integrase-Mutanten in der Nähe der *core*-Regionen erhöht. Dies kann zu einer Steigerung der Rekombination zwischen zwei *attachment*-Regionen führen, da an den *core*-Regionen letztendlich der Strangaustausch stattfindet. Die Erhöhung der lokalen Proteinkonzentration in der Nähe von Aktionsorten ist ein allgemein gültiges Konzept in der Kontrolle von DNA-Transaktionen von bivalenten DNA-bindenden Proteinen (Dröge & Müller-Hill, 2001).

Neben der Bedeutung der Arm-Sequenzen wurde auch die Rolle der N-terminalen Arm-Bindungsdomäne (N-Domäne) bei der von wild-typ-Int und Int-h/218 vermittelten Rekombination in HeLa-Zellen untersucht. Hierfür wurden Integrase-Proteine mit deletierter N-Domäne hergestellt und getestet. Die Daten in Kap. 4.1.3.2 (Abb. 4-3) zeigen, dass in HeLa-Zellen die Expression der verkürzten Proteine (C-Domänen) vergleichbar mit denen der Vollängen-Proteine ist. Zur Analyse der Rekombinationsaktivität wurden inter- und intramolekulare integrative Reaktionswege betrachtet (Kap. 4.1.3.2). Die Ergebnisse in Abb. 4-4 zeigen, dass nur bei der Expression der Int-h/218-C-Domäne ein Anteil an eGFP-ausprägenden Zellen detektiert werden kann. Dieser Anteil entsprach der Menge, die mit der Expression der wild-typ-Vollängen-Integrase gesehen wurde. Wäre die Rolle der N-Domäne auf die verstärkte Bindung von Arm-Sequenzen beschränkt, hätte man eine Rekombinationseffizienz erwartet, die vergleichbar ist mit der Rekombination von *attB*-Paaren durch Vollängen-Int-h/218, d.h. nicht unter 4-5 %. Tatsächlich ist sie jedoch erheblich niedriger. Dies kann als Indiz dafür angesehen werden, dass die N-Domäne noch andere Funktionen bei der Integrase vermittelten Rekombination innehat. Segall und Mitarbeiter zeigten als erste, dass die N-Domäne der wild-typ-Integrase scheinbar in Protein-Protein Interaktionen involviert ist, die für die Rekombination notwendig sind (Jessop et al., 2000). Ihre Ergebnisse werden durch neuere Studien (Cho et al., 2002; Warren et al., 2003) bestärkt und durch die Aufklärung der Kristall-Struktur der N-Domäne (Wojciak et al., 2002) die zeigt, dass die N-Domäne aus einer 3 strängigen β -Faltblattstruktur und einer α -Helix besteht. Mit der β -Faltblatt-Struktur erkennt und bindet sie die Arm-Regionen auf der DNA. Dagegen scheint die α -Helix unter anderem für die homophile

Interaktion zwischen zwei Int-Molekülen wichtig zu sein (Jessop et al., 2000; Warren et al., 2003). Auch in menschlichen Zellen ist diese Funktion der N-Domäne denkbar. Durch die Dimerisierung von Int-Molekülen wird vermutlich die Zusammenlagerung und Stabilisierung der Nukleoproteinkomplexe erreicht, die sich beispielsweise zwischen zwei *attB*-Sequenzen ausbilden.

Für die C-Domäne der wild-typ-Integrase wurde bereits *in vitro* nachgewiesen, dass sie die Fähigkeit besitzt DNA an *core*-Sequenzen zu schneiden (Sarkar et al., 2001). Da man beim Einsatz der C-Domäne von Int-h/218 einen (wenn auch sehr niedrigen) Anteil eGFP-ausprägender Zellen beobachtet (Abb. 4-4) ist es sehr wahrscheinlich, dass auch sie eine solche Schneideaktivität besitzt. Wie bereits im vorangegangenen Abschnitt diskutiert, wird diese Schneideaktivität wahrscheinlich ohne die Ausbildung eines stabilen synaptischen Komplexes zwischen zwei *att*-Sequenzen durchgeführt. Die Aufklärung der Ko-Kristallstruktur mit der linken *core*-Bindungssequenz von *attP* (Aihara et al., 2003) zeigt, dass die C-Domäne in der Lage ist an diese zu binden. Sie umfasst die DNA wie eine Zange wobei die *core*-Bindungsdomäne mit der kleinen Furche und die katalytische Domäne auf der gegenüberliegenden Seite mit der kleinen und großen Furche der DNA interagiert. Diese *cis*-Interaktion der katalytischen Domäne könnte zur Einführung von Einzel- oder Doppelstrangbrüchen führen, die wiederum in einer homologen oder illegitimen Rekombination zwischen *attB* und *attP* und somit in einer Ausprägung des eGFP-Reporters enden. Für die C-Termini der wild-typ-C-Domäne wird eine *trans*-Interaktion mit benachbarten Int-Molekülen vermutet (Aihara et al., 2003, Tekle et al., 2002). Alternativ könnten daher durch die C-Domäne auch transiente synaptische Komplexe gebildet werden, die gelegentlich stabil genug sind, um die ersten Strangaustausche einzuleiten. Die Auflösung der resultierenden Holliday-Strukturen könnten durch noch nicht identifizierte Faktoren, die es in Säugetier-Zellen zu geben scheint, erfolgen (Costantinou et al., 2001). Um den tatsächlichen Mechanismus aufzuklären, müssen jedoch noch ausführlichere *in vitro* Charakterisierungen der Int-h/218 C-Domäne erfolgen.

Insgesamt lässt sich aus den vorliegenden Ergebnissen festhalten, dass eine durch Integrase-Mutanten vermittelte effiziente Rekombination in HeLa-Zellen sowohl Arm-Sequenzen als auch Vollängen-Proteine erfordert, die in der Lage sind, diese Sequenzen zu binden.

5.1.3 Der rekombinante *Integration Host Factor* (rIHF) steigert die episodale Rekombinationsaktivität des Integrase wild-typ-Proteins in HeLa-Zellen

Wie schon in den Kapiteln zuvor erwähnt, können die Integrase-Mutanten Int-h und Int-h/218 sowohl exzisive als auch integrative Rekombinationsreaktionen in Abwesenheit von Hilfsproteinen effizient katalysieren (Christ & Dröge, 1999; Lorbach et al., 2000; Christ et al., 2002). Die Regulierbarkeit der Reaktion, die einen entscheidenden Vorteil gegenüber den alternativen Rekombinationssystemen wie z.B. *Cre/loxP* und *Flp/FRT* darstellt, geht jedoch mit dem Einsatz der Mutanten in eukaryotischen Zellen verloren. Versuche mit integrativen Rekombinationssubstraten, die im Vorfeld mit gereinigtem IHF inkubiert wurden zeigten, dass die Aktivität der wild-typ-Integrase auch in eukaryotischen Zellen stimuliert werden kann (Christ et al., 2002). Diese Ergebnisse offenbaren die Möglichkeit durch die Kombination von wild-typ-Integrase und entsprechenden Kofaktoren wieder Kontrolle über die Reaktionsrichtung zu erlangen, ein Aspekt der vor allem für die stabile sequenz-spezifische Integration von Fremd-DNA in das Genom von Eukaryoten wichtig ist. Der Einsatz von gereinigtem Protein erlaubt es jedoch nicht, größere Menge an IHF über einen längeren Zeitraum in eukaryotischen Zellen bereitzustellen. Versuche, das heterodimere IHF über Expressionsvektoren für die beiden Untereinheiten IHF- α und IHF- β in eukaryotischen Zellen zu exprimieren scheiterten (Nicole Christ, Dissertation 2002), da vermutlich kein funktionelles IHF-Protein gebildet wird. Untersuchungen in *E. coli* zeigen, dass eine Überexpression der beiden IHF α - und β -Untereinheiten zu instabilen Polypeptiden und unlöslichen Aggregaten führt (Nash et al., 1987; Wang et al., 2003). Um diese technischen Probleme bei der stöchiometrischen Ausprägung der IHF-Untereinheiten zu überwinden, generierte Nicole Christ in Zusammenarbeit mit Thomas Schwartz (Rockefeller University, New York) ein monomeres rekombinantes IHF-ähnliches Protein, im Folgenden als rIHF bezeichnet (Corona et al., 2003). Dabei wurde versucht, die dreidimensionale Struktur des heterodimeren IHFs beizubehalten, um so ein funktionelles Protein zu erhalten (siehe Kap. 4.1.4). Ein Schwerpunkt dieser Arbeit war die Untersuchung dieses neuartigen Proteins im eukaryotischen Milieu. Neben funktionellen Tests, die den Einfluss von rIHF auf die episodale und chromosomale (Kap. 5.2.2) Rekombination der *Lambda*-Integrase und ihrer Mutanten ermitteln sollten, wurden auch die Langzeitexpression und die Lokalisation des Proteins analysiert. Es wurden transgene Zelllinien generiert, die mindestens

eine Kopie des rIHF-Gens stabil im Genom integriert haben (Abb. 4-6). Die *Western* Analyse in Abb. 4-7 zeigt eine konstante und relativ hohe Expression des rekombinanten Proteins über einen Zeitraum von acht Monaten, bei der keine nennenswerten Abbauprodukte erkennbar sind. Zumindest für HeLa-Zellen in Kultur scheint demnach rIHF nicht toxisch zu sein. Immunfluoreszenz-Anfärbung *in vivo* ermöglichte den Nachweis von rIHF im Zellkern (Abb. 4-8). Die Präsenz im Nukleus ist höchstwahrscheinlich auf intrazelluläre Diffusion des 23 kDa kleinen Proteins zurückzuführen. Eine vorsichtige Schätzung, die auf mehreren *Western* Analysen beruht (Daten nicht gezeigt), lässt auf eine Anzahl von 20.000 bis 200.000 Molekülen pro Zelle schließen. Vorläufige bioinformatische Analysen haben zu der Identifizierung von ca. 90.000 consensus IHF-Bindungssequenzen geführt, die mehr oder weniger zufällig im menschlichen Genom verteilt sind. Ungefähr 15.000 von diesen besitzen auch eine passende poly(dTA)-Sequenz in einer geeigneter Position und kommen daher für eine starke Bindung und Krümmung seitens rIHF in Frage (Peter Dröge, persönliche Mitteilung). Die teilweise starke Anfärbung der Zellkerne die bei einigen Zellen in Abb. 4-8 beobachtet werden kann, deutet darauf hin, dass zumindest einige dieser Sequenzen für rIHF zugänglich sind.

Wie zuvor schon erwähnt, ist das *Lambda*-Rekombinationssystem wesentlich komplexer als z.B. das *Cre/loxP*-, das *Flp/FRT*- oder auch das *phiC31* Integrase System (Azaro & Landy, 2002; Lewis & Hatfull, 2001). Dieser hohe Grad an Komplexität führt zu einem hohen Maß an Kontrolle über die Richtung der Rekombination. Auch die Zuverlässigkeit in der Wahl der Partner-Sequenzen und bei der Bildung der Rekombinationsprodukte wird dadurch gesteigert (Azaro & Landy, 2002; Radman-Livaja et al., 2003). In *E. coli*-Stämmen, die IHF exprimieren, ist z.B. die Int vermittelte Integration von *attP* in eine andere Sequenz als die 21 bp umfassende *attB*-Region ein sehr seltenes Ereignis (Miller & Friedmann, 1981). Um den Einfluss von rIHF auf die episomale Rekombination in eukaryotischen Zellen zu untersuchen, wurden intramolekulare Rekombinationsanalysen in einer rIHF-transgenen Zelllinie durchgeführt und mit Analysen in der parentalen HeLa-Zelllinie verglichen. Die Ergebnisse in Abb. 4-9 zeigen, dass die Aktivität der Integrase-Mutanten Int-h und Int-h/218, die auch in Abwesenheit des wild-typ-IHFs schon sehr hoch ist, durch die Anwesenheit des rIHF nur geringfügig verstärkt wurde (Abb. 4-9, obere Diagramme). Die Effizienz der integrativen Rekombination durch die wild-typ Integrase konnte jedoch mit dem rekombinanten IHF um das vier- bis achtfache

gesteigert werden. Diese Stimulierung beschränkte sich jedoch auf den integrativen Reaktionsweg; für die exzusive Rekombination wurde eine vergleichbare Steigerung der Effizienz nicht bestätigt (Abb. 4-9, Diagramm unten rechts). Vermutlich ist dies darauf zurückzuführen, dass der für den exzisiven Rekombinationsweg benötigte Kofaktor XIS in HeLa-Zellen nicht vorhanden ist. Eine genauere Betrachtung der unteren Diagramme in Abb. 4-9 zeigt auch, dass in HeLa-Zellen die exzusive Rekombination durch wild-typ-Int gegenüber der integrativen Rekombination um das zwei- bis dreifache höher ist. Möglicherweise stimulieren in Eukaryoten die HMG1- oder die HMG2-Proteine diese Reaktion auch in Abwesenheit von IHF und XIS. Zumindest *in vitro* konnte die Stimulierung der exzisiven Rekombination durch diese Proteine schon gezeigt werden (Segall et al., 1994). Der relativ hohe Hintergrund, der mit dem negativen Kontrollvektor erreicht wird, erschwert die Interpretation.

Scheinbar werden funktionsfähige Intasome nur an *attP* und vielleicht auch an *attL* gebildet, nicht aber an der hybriden *attR*-Sequenz, weil vermutlich hierfür XIS notwendig ist. Eine potenzielle Anwendungsmöglichkeit für das wild-typ- λ -Int-System zusammen mit rIHF könnte die kontrollierte und gezielte Integration in eine der ca. 20 *attB*-ähnlichen Sequenzen sein, die im humanen Genom identifiziert wurden (Jinming Lee und Peter Dröge, unpublizierte Ergebnisse, siehe auch Kap. 5.1.4 und 5.2). Die hohe Zuverlässigkeit des λ -Systems könnte die Anzahl ungewollter Integrationsereignisse ins Genom einschränken. Auch die illegitime intragenomische Rekombination, die bei einigen anderen sequenz-spezifischen Rekombinationssystemen beobachtet wurde (Schmidt et al., 2000; Thyagarajan et al., 2001), könnte mit dem λ -System limitiert werden. Möglicherweise könnte es für die sichere genomische Manipulation, beispielsweise bei der Gentherapie in menschlichen Stammzellen wichtig werden. Die hohe Zuverlässigkeit des λ -Systems könnte dabei auch die in Anwesenheit von rIHF beobachtete niedrige Effizienz aufwiegen.

Mit rIHF wurden in unserer Arbeitsgruppe noch weitere Untersuchungen *in vitro* und in *E. coli* durchgeführt. Bis auf eine etwas niedrigere Affinität seitens rIHF für die H', eine der am besten untersuchten IHF-Bindungssequenzen aus *attP* (Yang & Nash, 1995), konnten hinsichtlich der biochemischen Eigenschaften und der Funktionalität keine Unterschiede zum wild-typ-Protein festgestellt werden (Corona et al., 2003). Es ist also wahrscheinlich, dass die meisten, wenn nicht sogar alle Seitenketten, die direkt oder indirekt mit DNA interagieren, im rekombinanten Protein korrekt positioniert sind. Nach bisherigem Kenntnisstand ist dies das erste Beispiel

eines rekombinanten Proteins, das durch die komplette Insertion einer Proteinuntereinheit in einer anderen generiert wurde. Bei bekannter dreidimensionaler Struktur könnte diese Strategie ebenso für andere heteromere Proteinkomplexe Anwendung finden.

5.2 Rekombination mit genomischen Substraten durch *Lambda*-Integrase und ihre Mutanten in eukaryotischen Zellen

Die bisher diskutierten Experimente wurden mit episomalen Substraten in HeLa-Zellen durchgeführt. Ein weiteres Thema dieser Arbeit war aber die Untersuchung der genomischen Rekombinationsaktivität der Integrase und ihrer Mutanten. Daher wurden in diesem Projekt Strategien für die sequenz-spezifische Integration von Fremd-DNA ins Genom eukaryotischer Zellen entwickelt und getestet. Für diesen Zweck wurden Zelllinien generiert, die eine wild-typ-*attB*-Sequenz stabil im Genom integriert haben. Außerdem wurde eine *attB* ähnliche Sequenz, die natürlich im Genom von HeLa-Zellen vorkommt, auf ihre episomale und genomische Rekombinationsfähigkeit untersucht. Zusätzlich wurde der Effekt des rHF auf die genomische Rekombination untersucht. Dabei wurde neben der intermolekularen Integration auch die Inversion in einer bereits publizierten H₈B-Zelllinie (Lorbach et al., 2000) geprüft.

5.2.1 In HeLa-Zellen ist die sequenz-spezifische Integration von Fremd-DNA in genomische Loci nur bedingt nachzuweisen

In der vorliegenden Arbeit wurden zwei Strategien entwickelt und getestet, nach der eukaryotische Genome unter Verwendung des intermolekularen integrativen Reaktionswegs gezielt manipuliert werden sollen (Abb. 4-10; Abb. 4-18).

Die zuerst vorgestellte Strategie (Abb. 4-10) beinhaltet die Herstellung von HeLa-Zelllinien (Kap. 4.2.1.1, Abb. 4-11), die *attB* gefolgt von einem promoterlosen Resistenzgen für Neomycin stabil in ihrem Genom integriert haben (Abb. 4-11A). In diese genomische *attB*-Zielsequenz soll der minimale *Targetvektor* pCMVSS*attP*_{mut} (Kap. 3.1.10 und Abb. 4-1) nach Integrase-Expression integrieren. Durch die Neu-positionierung des CMV-Promoters vor das Neo-Gen sollte die Resistenz angeschaltet werden und die Selektion mit Neomycin-haltigem Medium sollte

Zellklone hervorbringen, die ausschließlich positiv für dieses Ereignis waren. Da die Integrase-Mutante Int-h/218 sich im Vorfeld als enzymatisch aktivstes Protein erwiesen hatte, wurde sie für die in Kap. 4.2.1.2 beschriebenen *Targeting*-Experimente in verschiedene Zelllinien eingesetzt. Die korrekte durch Int-h/218 vermittelte Integration konnte in mehreren Fällen durch PCR (Abb. 4-12 und Abb. 4-15) und nachfolgender Sequenz-Analyse (Daten nicht gezeigt) nachgewiesen werden. Der Nachweis der Integration durch *Southern* Analyse (Abb. 4-13) und die Isolierung und Expansion von Zellklonen, die für dieses Ereignis positiv waren, gelang jedoch nicht.

Ein wiederholt auftretendes Problem war der hohe Hintergrund, der sich auch bei der Transfektion des negativen Kontrollvektors in einer erheblichen Anzahl Neo-resistenter Klone darstellte. Die anfänglich in den negativen Kontrollen der HattB3-Zelllinie beobachteten Neo-resistenten Klone beruhten darauf, dass die Zelllinie nicht klonal war. Eine zusätzliche Kopie des Neo-Gens war zufällig in einigen Zellen integriert und führte zur Resistenz und Anreicherung dieser Zellen nach Selektion (Abb. 4-13). Durch Subklonierung (Abb. 4-14) konnte der Hintergrund drastisch reduziert werden, dennoch wurden im Vergleich zu den Negativ-Kontrollen in den *Targeting*-Experimenten mit Int-h/218 keine Unterschiede in der Anzahl resistenter Klone beobachtet (Daten nicht gezeigt). Da bei den anderen Zelllinien meist keine resistenter Klone auftraten wurde vermutet, dass zur Selektion eine zu hohe Geneticin-Konzentration verwendet wurde. Die Schwellenwert-Konzentration des für die Selektion eingesetzten Geneticins wurde daher für jede Zelllinie einzeln mit einem *Targeting*-/Titrations-Experiment bestimmt (Abb. 4-15). Bei den Zelllinien S23 und bei U3 wurde das erwartete Integrationsprodukt detektiert und konnte durch Sequenzierung bestätigt werden. Mit der S23-Zelllinie, die ein Subklon von HattB3 ist, wurden weitere *Targeting*-Experimente durchgeführt, bei denen die Selektion nur noch mit 500 µg/ml anstatt mit 750 µg/ml Geneticin erfolgte. Die Anzahl der Neomycin-resistenten Klone stieg wieder an auf bis zu 200 und wurde auch bei den Kontrollen weiter beobachtet (Daten nicht gezeigt). Vereinzelt man einige dieser im Int-h/218 Ansatz resistenten Klone, so zeigten sich weder in PCR- noch in *Southern* Analyse das richtige Rekombinationsprodukt. Die Ursache dieser immer wieder spontan auftretenden Neo-Resistenz ist nicht ganz klar. Wie die Expressionsvektoren der Integrase enthält der Kontrollvektor pCMV (Mock) einen CMV-Promoter. Ihm fehlen jedoch das Intron und ein offenes Leseraster für eine Integrase. Vorstellbar

wäre, dass durch illegitime Integration der pCMV-Promoter vor das Neo-Gen positioniert wird und somit seine Ausprägung bewirkt. Unter Berücksichtigung der Größe des menschlichen Genoms erscheint diese Vorstellung eher unwahrscheinlich und bietet nur eine schwache Erklärung für die zum Teil hohe Anzahl von Klonen, die in den negativen Kontrollen beobachtet wurde. Denkbar wäre eine Neuordnung der genomischen Neo-Sequenz, die zur Expression des Gens führt, allerdings existiert hierfür im *Southern* der analysierten Klone kein Indiz (Daten nicht gezeigt). Auch eine spontane Resistenz-Ausbildung, wie sie in HeLa-Zellen für Geneticin (Kirch et al., 2003) und BL60-Zellen für Hygromycin (Lorbach, Dissertation 2000) schon beobachtet wurde, wäre eine mögliche Erklärung für dieses Phänomen.

Der Ansatz des schon beschriebenen genomischen *Targetings* einer zuvor im Genom stabil integrierten *attB*-Sequenz (Kap. 4.2.1) erfordert die aufwendige Generierung von Zelllinien, die wie die vorhergehenden Kapitel zeigen, nicht immer unproblematisch ist. Die zweite in dieser Arbeit vorgestellte Strategie beinhaltet daher die Integration in die sogenannte *attH2*-Region (Kap. 4.2.3; Abb. 4-18 und Abb. 4-19), eine der ca. zwanzig identifizierten *attB*-ähnlichen Sequenzen im menschlichen Genom (Peter Dröge, unveröffentlichte Daten). *attH2* ist Bestandteil einer in diversen Organen gebildeten menschlichen mRNA, die für den Monocarboxylat-Transporter 5 (MCT5) kodiert (Halestrap & Price, 1999, wurde in Price et al., 1998 noch als MCT4 bezeichnet). Auch in menschlichen Gebärmutterhals- und Nierenzelllinien (HeLa bzw. 293T) konnte die *attH2*-tragende mRNA nachgewiesen werden (Christ, Dissertation 2002). Die in Abb. 4-18 dargestellte *Targeting*-Strategie für die menschliche *attH2*-Region hat Ähnlichkeit mit so genannten *promoter-trap*-Strategien (Friedrich & Soriano, 1991; Soriano et al., 1991). Das gewählte Reportersystem beruht auf einem MCT5-eGFP-Fusionsprotein, welches nach korrekter Integration des *Targetvektors* über den endogenen MCT5-eigenen Promoter exprimiert werden soll. *attH2* unterscheidet sich in vier Nukleotiden von der wild-typ-*attB*-Sequenz. Drei dieser Nukleotidunterschiede befinden sich in einer *core*-Sequenz und werden von der *Lambda*-Integrase toleriert (Nash, 1981). Um die Homologie zwischen *attH2* und *attP* in der *overlap*-Region zu bewahren, musste die *attP*-Sequenz des *Targetvektors* an *attH2* angepasst werden (Abb. 4-19). Außerdem mussten im P'-Arm alle Stopcodons mutiert werden, um die Ausprägung des Fusionsproteins zu gewährleisten. In episomalen Voruntersuchungen mit Int-h-

218 wurde die Kompatibilität von *attH2* mit *attP* und mit der angepassten *attP*_{ns}*-Sequenz getestet. Die Ergebnisse in Kap. 4.2.3.1 zeigen deutlich, dass die Int-h/218 vermittelte Rekombination einer hohen Spezifität für Partnersequenzen unterliegt. Nur zusammengehörige Sequenzen, d.h. *attB/attP* und *attH2/attP*_{ns}* werden effizient rekombiniert. Dieser Aspekt ist bei einem eventuellen gentherapeutischen Einsatz des *Lambda*-Systems für die Sicherheit sehr wichtig. Allerdings zeigen die Ergebnisse auch, dass insgesamt die Rekombination von *attB/attP* um ein vierfaches besser ist als die von *attH2/ attP*_{ns}* (Abb. 4-20).

Im Anschluß an die episomalen Tests des *attH2/attP*_{ns}*-Systems wurde versucht, mit *attP*_{ns}* in die genomische *attH2*-Region zu integrieren. *Targetvektor* und Expressionsvektor für Int-h/218 bzw. pCMV (als Negativ-Kontrolle) wurden mittels Elektroporation in HeLa-Zellen eingebracht. Eine Woche später erfolgten die FACS-Analyse oder die FACS-Sortierung der transfizierten Zellen (Abb. 4-22). Der detektierte Anteil an fluoreszierenden Zellen in Ansätzen mit Integrase lag meist bei marginalen 0,02 % (einmalig 0,03 %). Die transfizierten Zellen wurden sortiert und expandiert. Unter dem Mikroskop konnten danach keine fluoreszierenden Zellen mehr beobachtet werden; auch ein korrektes Rekombinationsprodukt konnte nicht nachgewiesen werden (Daten nicht gezeigt). Auch andere Gruppen konnten die Abnahme und das Verschwinden von Fluoreszenz in stabilen Zelllinien detektieren und führten dieses Phänomen zum Teil auf die Eliminierung des Gens aus dem Genom und zum Teil auf Hypermethylierung der fremden DNA zurück (Migliaccio et al., 2000; Kirch et al., 2003). Kirch und Mitarbeiter beobachteten für das stabil im Genom von HeLa integrierte hrGFP aus *Renilla reniformis* (Tsien et al., 1998) eine unvollständige Abstellung des Gens. Sie konnten in Zellen, in denen die Fluoreszenz erloschen war, die codierende Sequenz im Genom und die Bildung der mRNA aber kein Protein nachweisen. Als mögliche Erklärung geben sie Modifikationen an die nach der Transkription auftreten und so die Proteinausprägung abstellen.

Nach den vorliegenden Ergebnissen scheint die Integration in den genomischen *attH2* demnach sehr ineffizient zu sein. In Stammzellen der Maus erreicht die intramolekulare integrative *attB/attP*-Rekombination durch Int-h/218 eine Effizienz von ca. 0,01 bis 0,1%. Dabei ist je eine Kopie der beiden *att*-Regionen in einem Abstand von ca. 3 kb im *ROSA26*-Lokus integriert (Christ & Dröge, 2002). Die durchgeführten *Targeting*-Experimente könnten daher die mindestens um ein Viertel schlechtere Rekombination von *attH2/attP*_{ns}* widerspiegeln. Ob die Integration in die

attH2-Region tatsächlich erfolgt, ist jedoch nicht eindeutig erwiesen. Möglich wäre, dass die gewählte *attB*-ähnliche Sequenz unter den bisher identifizierten Sequenzen für den vorgestellten experimentellen Ansatz nicht die beste ist. Elke Lorbach konnte in ihrer Dissertation die Integration in eine andere *attH*-Sequenz durch PCR nachweisen. Die Isolierung eines positiven Klons gelang ihr jedoch ebenfalls nicht. Eine Schwäche der in der vorliegenden Arbeit angewandten Strategie ist der fehlende Nachweis eines funktionellen MCT5-eGFP-Fusionsproteins. Zwar wurde die erfolgreiche Fusion eines Gens an eGFP schon häufig gezeigt (Elliott & O'Hare, 1999; Wang & Hazelrigg, 1994; Marshall et al., 1995; Sheridan et al., 2002), dennoch kann die beobachtete Fluoreszenz auch auf zufällige Integration des eGFP hinter einen aktiven genomischen Promoter zurückzuführen sein. Es wäre möglich, dass die Integration in die genomische *attH2*-Region keine oder keine ausreichende Fluoreszenz ergibt. In diesem Fall werden bei der FACS-Analyse und Sortierung die korrekten Integrierten nicht erfasst und gehen verloren.

Zusammenfassend belegen die Ergebnisse, dass die Int-h/218 vermittelte sequenz-spezifische Integration fremder DNA ins Genom menschlicher Zellen prinzipiell funktioniert. Für den Einsatz in der Forschung oder sogar für therapeutische Zwecke müssten jedoch System und Strategie verbessert werden, um die Effizienz der Integration und die Stabilität der Integrate zu erhöhen. Die Wahl der Zelllinie ist ebenfalls sorgsam zu überdenken. Stabilere Zelllinien oder sogar Stammzellen wären für zukünftige Experimente vielleicht besser geeignet.

5.2.2 Eine Steigerung der genomischen Rekombinationsaktivität durch rIHF kann unter den gewählten Bedingungen in HeLa-Zellen nicht detektiert werden

Unter 5.1.3 wurde ein rekombinantes IHF-Protein beschrieben, dass die durch wild-typ-Integrase vermittelte integrative Rekombination episomaler Substrate in HeLa-Zellen stimuliert. Für eine mögliche Anwendung bei der Manipulation eukaryotischer Genome war es von Interesse, ob rIHF auch die Rekombination genomischer Substrate verbessert. Zur Klärung dieser Fragestellung wurden die Integration in und die Inversion von genomischen Loci in An- und Abwesenheit von rIHF untersucht. Die intramolekulare Rekombination von zwei, auf demselben Chromosom liegenden, *att*-Regionen durch die λ -Integrase-Mutanten wurde in

humanen Zellen durch genomische Inversion (Lorbach et al., 2000) und in embryonalen Stammzellen der Maus durch die Deletion eines Marker-Gens gezeigt (Christ et al., 2002). Die genomische Inversion in H₈B-Zellen sollte in der vorliegenden Arbeit als Reporter für eine mögliche Stimulierung durch rIHF dienen. H₈B-Zellen sind HeLa-Zellen, die ca. 8 Kopien eines *attP/attB* Inversionssubstrates stabil an zufälligen Positionen im Genom integriert haben (Kap. 4.2.2.3, Lorbach et al., 2000). Eine quantitative *Southern* Analyse (Abb. 4-17) zeigt, dass nur die Integrase-Mutanten Int-h und Int-h/218 effizient die genomischen *attP*- und *attB*-Sequenzen rekombinieren. Im Durchschnitt betrug die mit dem Phosphoimager gemessene Intensität des Signals für Int-h ~4 % und für Int-h-218 ~8 % der Gesamtradioaktivität (Daten nicht gezeigt). Berücksichtigt man die Transfektionseffizienz von ca. 50 % bei der Lipofektion so wurden zwischen 8 % und 16 % der genomischen Substrate rekombiniert. Ein bedeutender Unterschied zwischen der An- oder Abwesenheit von rIHF wurde jedoch nicht festgestellt. Die Rekombination durch die wild-typ-Integrase konnte auch in früheren Arbeiten lediglich durch PCR-Analyse nachgewiesen werden (Nicole Christ, Dissertation 2002). Auch in den hier durchgeführten Experimenten ist die Aktivität von wild-typ-Int sehr gering und kann mittels *Southern* Analyse nicht detektiert werden (Abb. 4-17C, Spuren 5 und 6). Eine Aussage über die Stimulierung der Int vermittelten Rekombination genomischer Substrate durch rIHF ist in diesem Fall daher nicht möglich. Dass die Bindung von rIHF an genomische Sequenzen erfolgen kann, ist jedoch zumindest indirekt durch die Lokalisierung des Proteins im Kern nachgewiesen. Für eine Stimulierung der Rekombination müssten sich Integrase und rIHF gleichzeitig im Kern befinden. Zumindest für rIHF wurde in dieser Arbeit jedoch gezeigt, dass die Lokalisation im Zellkern zeitlichen Schwankungen unterliegt (Abb. 4-8). Ob das für die Integrase ebenso der Fall ist und ob diese Schwankungen synchron mit denen des rIHF sind, ist nicht bekannt. Zur Klärung dieser Frage ist eine Untersuchung der zeitlichen zellulären Lokalisation der Integrase nötig, wie in Kap. 4.1.4.2 für rIHF beschrieben. Die fehlende Stimulierung der genomischen Rekombination durch rIHF könnte auch damit begründet werden, dass die Ausbildung funktioneller Nukleoproteinkomplexe durch bestimmte Kernproteine verhindert wird. Für das Histondimer H2A-H2B konnte beispielsweise nachgewiesen werden, dass es an *attL* bindet (Segall et al., 1994).

Der Einfluss von rIHF auf die Integration fremder DNA in genomische Loci wurde

wegen der schon zuvor beobachteten niedrigen Effizienz der genomischen Integration nur für Int-h/218 untersucht. Für diese Integrase-Mutante wurde in den episomalen Tests lediglich eine geringe Steigerung der Rekombinationseffizienz durch rIHF nachgewiesen (Abb. 4-9). Verglichen mit der wild-typ-Integrase in den episomalen Tests katalysierte Int-h/218 die intermolekulare chromosomale Rekombination mit geringerer Effizienz. Der positive Effekt des rIHF hätte auf genomischer Ebene daher deutlicher ausfallen können als zuvor in den transienten Experimenten beobachtet wurde. Die Stimulierung der Integration fremder DNA in genomische *attB*- (Kap. 4.2.2.2) oder *attH2*-Sequenzen (Kap. 4.2.3.2) konnte jedoch ebenfalls nicht nachgewiesen werden. Sie wurde in stabilen Zelllinien, die rIHF (Kap. 4.2.3.2) oder rIHF und *attB* (Abb. 4-16) stabil im Genom integriert haben, untersucht und durch Kotransfektion des rIHF-Expressionsvektors. Ob tatsächlich keine Stimulierung erfolgt, kann aufgrund der beobachteten niedrigen Effizienz nicht gesagt werden, da minimale Veränderungen möglicherweise nicht detektiert werden können. Zumindest in embryonalen Stammzellen der Maus konnte Nicole Christ während ihrer Dissertation nachweisen, dass die intramolekulare Rekombination von *attB/attP* durch die gleichzeitige Ausprägung von rIHF gesteigert werden kann.

Die Möglichkeiten für die Anwendung des rIHF beschränken sich jedoch nicht nur auf eine Rolle in der Rekombination. Denkbar wäre etwa, dass dieses rekombinante Protein bei der Erforschung von Struktur und Funktion eukaryotischer Nukleoproteinkomplexe in lebenden Zellen helfen könnte. Zusätzlich könnte rIHF in biopharmazeutischen Produktionstechniken eingesetzt werden. Durch seine DNA-Bindungs- und Krümmungseigenschaften könnte es dabei beispielsweise einen Promoter frei von Nukleosomen halten und dadurch der Transkriptions-Maschinerie der Zelle den Zugang erleichtern. Die gewünschte hohe Expression eines Gens könnte so erhalten oder gesteigert werden. Eine weitere Aussicht bietet der Einsatz von rIHF-transgenen Zelllinien zur Lenkung der retroviralen Integration in spezifische Regionen des Genoms, zumal für wild-typ-IHF gezeigt wurde, dass es *in vitro* die Integration der HIV-1 cDNA in protein-induzierte Krümmungen fördert (Bor et al., 1995).

5.3 Transgene rIHF-Mäuse

Den abschließenden Teil dieser Arbeit stellt die Generierung einer rIHF-transgenen Maus dar. Um weitere *in vivo* Studien mit dem rekombinanten IHF durchzuführen, sollte neben rIHF-transgenen HeLa-Zelllinien eine transgene Maus generiert werden. Das Fragment für die Injektion ist in Abb. 4-22A schematisch abgebildet. Neben dem rIHF-Gen, dessen Ausprägung vom Caggs-Promoter kontrolliert wird, enthält das Fragment ein IRES-eGFP-Konstrukt, das später die Identifizierung transgener, rIHF-ausprägender Tiere erleichtern sollte. Die generelle Funktionalität des Fragmentes, d.h. die Expression des rIHF- und des eGFP-Gens, wurde nur in HeLa-Zellen nachgewiesen (Abb. 4-22). Dennoch ist davon auszugehen, dass es auch in Mauszellen funktioniert, da dies zumindest für die vier relevanten Elemente des Fragmentes schon gezeigt wurde: 1) Der Caggs-Promoter ist in Mäusen ubiquitär aktiv (Okabe et al., 1997). 2) rIHF wurde schon in embryonalen Stammzellen der Maus eingesetzt (Nicole Christ, Dissertation 2002). Langzeitstudien diesbezüglich fehlen jedoch. 3) Die benutzte IRES Sequenz hat sich in B-Zellen von transgenen Mäusen ebenfalls als funktional erwiesen (Nathalie Uyttersprot, unpublizierte Daten). 4) Das eGFP-Gen wurde schon in diversen Mauslinien als Reporter gen verwendet (Okabe et al., 1997, Novak et al., 2000; Kawamoto et al., 2000; Übersichtartikel: Hadjantonakis & Nagy, 2001). Obwohl also alle verwendeten Elemente nachweislich funktionell sind, wurden keine positiven Nachkommen gefunden, die eGFP oder rIHF (Abb. 4-24 und 4-25) ausprägten. Die fehlende Expression des rIHF in den untersuchten Nachkommen kann mehrere Ursachen haben. Zum einen besteht die Möglichkeit, dass die Expression des rIHF Gens in Mäusen generell oder nur während eines bestimmten Stadiums (z.B. der embryonalen Entwicklung) nicht toleriert wird und daher bei der Herstellung der Mäuse nur transgene Tiere erhalten wurden, in denen das Gen ausgestellt ist. Um diese Frage genauer zu klären, wäre beispielsweise die Herstellung transgener Mäuse hilfreich, in denen das rIHF nur konditional ausgeprägt wird, also nur zu gewissen Zeiten und in bestimmten oder in allen Geweben. Dazu könnte ein Fragment durch *gene targeting* (homologe Rekombination) in einen ubiquitär ausgeprägten Locus wie z.B. ROSA26 (Soriano, 1999; Mao et al., 1999; 2001) eingebracht werden. Die Anschaltung des Gens könnte dann durch die sequenz-spezifische Deletion eines vorgelagerten und mit Rekombinationssequenzen

flankierten Transkriptionsstopps erfolgen, beispielsweise nach Verpaarung mit einer Maus, die die entsprechende Rekombinase unter der Kontrolle des gewünschten zeit- und/oder orts-spezifischen Promoters ausprägt.

Ein weiterer Grund fehlender rIHF ausprägender Nachkommen, könnte darin liegen, dass zu wenige transgene Elterntiere erzeugt wurden: von 83 Tieren der F0 Generation waren nur 12 transgen für rIHF und wurden weiter verpaart. Ein Elterntier verstarb und 3 gaben keine transgenen Nachkommen. Es wurden demnach nur Nachkommen von 8 Eltern untersucht, die nach Injektion das DNA-Fragment lediglich zufällig in ungünstigen Positionen im Genom integriert haben, sodass eine Ausprägung nicht stattfinden konnte. Die Bandenmuster der rIHF *Southern* Analysen von transgenen Nachkommen (Abb. 4-23) und entsprechenden Eltern (Daten nicht gezeigt) unterschieden sich nicht. Dies könnte bedeuten, dass, wenn in einer Maus mehrere Integrationsorte vorhanden waren, diese sich doch in so räumlicher Nähe (vermutlich nahe zusammen auf einem Chromosom) befanden, dass sie bei der Vererbung nicht segregierten. Mehrere inserierte Fragment-Kopien könnten daher ähnlichen oder sogar gleichen Positionseffekten unterliegen, die die Ausprägung der eingeführten fremden Gene unterdrücken. Die erfolgreiche Generierung einer transgenen Mauslinie erfordert nicht selten die Untersuchung von Nachkommen von über 50 Elterntieren (Thomas Wunderlich und Nathalie Uyttersprot, unpublizierte Daten).

Unter dem Fluoreszenzbinokular zeigten weder Schwanzspitzen noch Organe der untersuchten Mäuse detektierbare eGFP-Expression. Auch die FACS-Analyse isolierter Milzzellen bestätigte dieses negative Ergebnis (Daten nicht gezeigt). Da die eGFP-Translation über die IRES-Sequenz von der Transkription des rIHF-Gens abhängig ist, ist die nicht detektierte Fluoreszenz der Mäuse in erster Linie auf die fehlende rIHF-Expression zurückzuführen.

6 Literaturverzeichnis

Abremski, K. & Gottesman, S. (1979). The form of the DNA substrate required for excisive recombination of bacteriophage λ . *J. Mol. Biol.*, **131**, 637-649.

Abremski, K., Hoess, R.H., & Sternberg, N. (1983). Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell*, **32**, 1301-1311.

Aihara, H., Kwon, H.J., Nunes-Düby, S.E., Landy, A. & Ellenberger, T. (2003). A conformational switch controls the DNA cleavage activity of λ Integrase. *Molecular Cell*, **12**, 187-198.

Araki, K., Araki, M., Miyazaki, J. & Vasalli, P. (1995). Site-specific recombination of a transgene in fertilized eggs by transient expression of Cre recombinase. *Proc. Natl Acad. Sci. USA*, **92**, 160-164.

Araki, K., Araki, M. & Yamamura, K. (1997). Targeted integration of DNA using mutant lox sites in embryonic stem cells. *Nucl. Acids Res.*, **25**, 868-872.

Argos, P., Landy, A., Abremski, K., Egan, J.B., Haggard-Ljungquist, E., Hoess, R.H., Kahn, M.L., Kalionis, B., Narayana, S.V. & Pierson, L.S. (1986). The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J.*, **5**, 433-440.

Azaro, M.A., & Landy, A. (2002). λ Int and the λ Int family. In Craig, N.L., Craigie, R., Gellert, M. and Lambowitz, A (eds), *Mobile DNA II*. ASM Press, Washington, DC, pp. 118-148.

Ball, C.A. & Johnson, R.C. (1991). Efficient excision of phage λ from the *Escherichia coli* chromosome requires the Fis protein. *J. Bacteriol.*, **173**, 4032-4038.

Ball, C.A. & Johnson, R.C. (1991b). Multiple effects of Fis on integration and the control of lysogeny in phage λ . *J. Bacteriol.*, **173**, 4027-4031.

Ball, C.A., Osuna, R., Ferguson, K.C. & Johnson, R.C. (1992). Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J. Bacteriol.*, **174**, 8043-8056.

Baubonis, W. & Sauer, B. (1993) Genomic targeting with purified Cre recombinase. *Nucl. Acids Res.*, **17**, 147-161.

Bauer, C.E., Gardner, J.F. & Gumpert, R.I. (1985). Extent of sequence homology requires for bacteriophage lambda site-specific recombination. *J. Mol. Biol.*, **181**, 187-197.

Baumann, P. & West, S.C. (1998). Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *TIBS*, **23**, 247-251.

- Bachmair, A., Finley, D. & Varshavsky, A. (1986) *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science*, **234**, 179-186.
- Bear, S.E., Clemens, J.B., Enquist, L.W. & Zagursky, R.J. (1987) Mutational analysis of the lambda *int* gene: DNA sequence of dominant mutations. *J. Bacteriol.*, **169**, 5880-5883.
- Better, M., Lu, C., Williams, R.C. & Echols, H. (1982). Site-specific DNA condensation and pairing mediated by the *int* protein of bacteriophage λ . *Proc. Natl Acad. Sci. USA*, **79**, 5837-5841.
- Better, M., Wickner, S., Auerbach, J. & Echols, H. (1983). Role of the Xis protein of bacteriophage λ in a specific reactive complex at the *attR* prophage attachment site. *Cell*, **32**, 161-168.
- Bollag, R.J., Waldman, A.S., Liskay, R.M. (1989). Homologous recombination in mammalian cells. *Annu. Rev. Gen.*, **23**, 199-225.
- Bullock, W.O., Fernandez, J.M., Shot, J.M. (1987). XL1: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Biotechniques*, **5**, 376-378.
- Burgin, A.B. & Nash, H.A. (1992). Symmetry in the mechanism of bacteriophage λ integrative recombination. *Proc. Natl Acad. Sci. USA*, **89**, 9642-9646.
- Bushman, W., Yin, S., Thio, L.L. & Landy, A. (1984). Determinants of directionality in lambda site-specific recombination. *Cell*, **39**, 699-706.
- Bushman, W., Thompson, J.F., Vargas, L. & Landy, A. (1985). Control of directionality in λ site-specific recombination. *Science*, **230**, 906-911.
- Cho, E.H., Gumpert, R.I. & Gardner, J.F. (2002). Interactions between integrase and excisionase in the phage λ excisive nucleoprotein complex. *J. Bacteriol.*, **184**, 5200-5203.
- Choi, T., Huang, M., Gorman, C. & Jaenisch, R. (1991). A generic intron increases gene expression in transgenic mice. *Mol. Cell. Biol.*, **11**, 3070-3074.
- Christ, N. & Dröge, P. (1999). Alteration in the directionality of λ site-specific recombination catalyzed by mutant integrases *in vivo*. *J. Mol. Biol.*, **288**, 825-836.
- Christ, N. & Dröge, P. (2001). Site-specific DNA recombination: a promising technique for revealing gene functions in eukaryotes. *Boehringer Ingelheim Fonds FUTURA*, **16** (1), 24-31.
- Christ, N. & Dröge, P. (2002). Genetic manipulation of mouse embryonic stem cells by mutant λ integrase. *Genesis*, **32**, 203-208.
- Christ, N., Corona, T. & Dröge, P. (2002). Site-specific recombination in human cells mediated by mutant λ integrases: implications for synaptic complex formation and the reactivity of episomal DNA segments. *J. Mol. Biol.*, **319**, 305-314.

- Christ, N. (2002). Gezielte genetische Manipulation durch *Lambda* Integrasen in embryonalen Stammzellen der Maus und in anderen Säugetierzellen. Doktorarbeit am Institut für Genetik der Universität zu Köln.
- Corona, T., Bao, Q., Christ, N., Schwartz, T., Li, J., & Dröge, P. (2003). Activation of site-specific DNA integration in human cells by a single chain integration host factor *Nucl. Acids Res.*, **31**, 5440-5448.
- Constantinou, A., Davies, A.A. & West, S.C. (2001). Branch migration and Holliday junction resolution catalyzed by activities in mammalian cells. *Cell*, **104**, 259-268.
- Craig, N. L. (1988). The mechanism of conservative site-specific recombination. *Annu. Rev. Genet.*, **22**, 77-105.
- Coppoolse, E.R., de Vroomen, M.J., Roelofs, D., Smit, J., van Gennip, F., Hersmus, B.J., Nijkamp, H.J. & van Haaren, M.J. (2003). Cre recombinase expression can result in phenotypic aberrations in plants. *Plant Mol. Biol.*, **51**, 263-279.
- Dale, E.C. & Ow, D.W. (1990). Intra- and intermolecular site-specific recombination in plant cells mediated by bacteriophage P1 recombinase. *Gene*, **91**, 79-85.
- Derbyshire, K.M. & Grindley N.D.F. (1987). Replicative and conservative transposition in bacteria. *Cell*, **47**, 325-327.
- Dröge, P. & Müller-Hill, B. (2001). High local protein concentration at promoters: strategies in prokaryotic and eukaryotic cells. *BioEssays*, **23**, 179-183.
- Dymecki, S.M. (1996). FLP recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl. Acad. Sci. USA*, **93**, 6191-6196.
- Echols, H. (1989). Multiple DNA-protein interaction governing high-precision DNA transactions. *Science*, **233**, 1055-1056.
- Elliott, G. & O'Hare, P. (1999). Intercellular trafficking of VP22-GFP fusion proteins. *Gene Ther.*, **6**, 149-151.
- Ellenberger, T. and Landy, A. (1997). A good turn for DNA: the structure of integration host factor bound to DNA. *Structure*, **5**, 153-157.
- Esposito, D. & Scocca, J.J (1997). The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucl. Acids Res.*, **25**, 3605-3614.
- Esposito, D. & Gerard, G.F. (2003). The *Escherichia coli* Fis protein stimulates bacteriophage λ integrative recombination *in vitro*. *J. Bacteriol.*, **185**, 3076-3080.
- Farley, F.W., Soriano, P., Steffen, L.S. & Dymecki, S.M. (2000). Widespread recombinase expression using FLPeR (flipper) mice. *Genesis.*, **28**, 106-110.

Fellenberg, K. (1998). Optimierung von Cre-Oestrogenrezeptor-Fusionsproteinen zur lokalen Geninaktivierung in der Haut. Diplomarbeit im Fachbereich Biologie der Universität zu Köln.

Filutowicz, M., Ross, W., Wild, J. & Gourse, R.L. (1992). Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J. Bacteriol.*, **174**, 398-407.

Franz, B. & Landy, A. (1995). The Holliday junction intermediates of λ integrative and excisive recombination respond differently to the bending proteins integration host factor and excisionase. *EMBO J.*, **14**, 397-406.

Friedrich, G. & Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.*, **5**, 1513-1523.

Freundlich, M., Ramani, N., Mathew, E., Sirko, A. & Tsui, P. (1992). The role of integration host factor in gene expression in *Escherichia coli*. *Mol. Microbiol.*, **6**, 2557-2563.

Friedman, D.I. (1988). Integration host factor: a protein for all reasons. *Cell*, **55**, 545-554.

Giese, K., Cox, J., & Grosschedl, R. (1992). The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell*, **69**, 185-195.

Golic, K.G. & Lindquist, S. (1989) The FLP Recombinase of Yeast Catalyses Site-Specific Recombination in the Drosophila Genome. *Cell*, **59**, 499-509.

Goodrich, J.A., Schwartz, M.L. & McClure, W.R. (1990). Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor (IHF). *Nucleic Acids Res.*, **18**, 4993-5000.

Gossen, M. & Bujard, H. (1995). Efficacy of tetracycline-controlled gene expression is influenced by cell type: commentary. *Biotechniques*, **19**, 213-216.

Goosen, N. & van de Putte, P. (1995). The regulation of transcription initiation by integration host factor. *Mol. Microbiol.*, **16**, 1-7.

Graham, F.L., Smiley, J., Russell, W.C. & Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.*, **36**, 59-74.

Grindley, N.D.F. & Reed, R.R. (1985). Transpositional recombination in prokaryotes. *Annu. Rev. Biochem.*, **54**, 863-896.

Griffith, J.D. & Nash, H.A. (1985). Genetic rearrangement of DNA induces knots with unique topology: Implications for the mechanism of synapsis and crossing-over. *Procl. Natl. Acad. Sci. USA*, **82**, 3124-3128.

- Gronostajski, R.M. & Sadowski, P.D. (1985). Determination of DNA sequences essential for FLP-mediated recombination by a novel method. *J. Biol. Chem.*, **260**, 12320-12327.
- Guo, F., Gopaul, D.N. & van Duyne, G.D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination. *Nature*, **389**, 40–46.
- Hadjantonakis, A.K. & Nagy, A. (2001). The color of mice: in the light of GFP-variant reporters. *Histochem. Cell Biol.*, **115**, 49-58.
- Hamilton, D.L. & Abremski, K. (1984). Site-specific recombination by the bacteriophage P1 lox-cre system. Cre mediated synapsis of two lox sites. *J. Mol. Biol.*, **178**, 481-486.
- Halestrap, A.P. & Price, N.T. (1999). The proton-linked monocarboxylate transporter (MCT) family: structure, function and regulation. *Biochem. J.*, **343**, 281-299.
- Hoess, R., Wierzbicki, A. & Abremski K. (1986). The role of the spacer region in P1 site-specific recombination. *Nucleic Acids Res.*, **14**, 2287-2300.
- Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet. Res.*, **5**, 282-304.
- Jackson, R.J, Howell, M.T. & Kaminski, A. (1990). The novel mechanism of picornavirus RNA translation. *Trends Biochem. Sci.*, **15**, 477-483.
- Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C. & Wimmer, E. (1988). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. *J. Virol.*, **62**, 2636-2643.
- Jayram, M. (1985). Two-micrometer circle site-specific recombination: The minimal Substrate and the possible role of flanking sequences. *Procl. Natl. Acad. Sci. USA*, **82**, 5875-5879.
- Jessop, L., Bankhead, T. Wong, D. & Segall, A.M. (2000). The amino terminus of bacteriophage λ Integrase is involved in protein-protein interactions during recombination. *J. Bacteriol.*, **182**, 1024-1034.
- Kawamoto, S., Niwa, H., Tashiro, F., Sano, S., Kondoh, G., Takeda, J., Tabayashi, K. & Miyazaki, J. (2000). A novel reporter mouse strain that expresses enhanced green fluorescent protein upon Cre-mediated recombination. *FEBS Lett.*, **410**, 263-268.
- Kikuchi, A., Flamm, E. & Weisberg, R.A. (1985). An *E.coli* mutant unable to support site-specific recombination of bacteriophage *lambda*. *J. Mol. Biol.*, **183**, 129-140.
- Kilby, N. J., Snaith, M. R. & Murray, J. A. (1993). Site-specific recombinases: tools for genome engineering. *Trends Genet.* **9**, 413-421.

- Kim, S., Moitoso de Vargas, L., Nunes-Düby, S.E. & Landy, A. (1990). Mapping of a higher order protein-DNA complex: two kinds of long-range interactions in λ *attL*. *Cell*, **63**, 773-781.
- Kim, S. & Landy, A. (1992). Lambda Int protein bridges between higher order complexes at two distant chromosomal loci *attL* and *attR*. *Science*, **256**, 198-203.
- Kirch, P., Hafner, M., Zentgraf, H. & Schilling, L. (2003). Time course of fluorescence intensity and protein expression in HeLa cells stably transfected with hrGFP. *Mol. Cells*, **15**, 341-348.
- Kitts, P.A. & Nash, H.A. (1987). Homology-dependent interactions in phage λ site-specific recombination. *Nature*, **329**, 346-348.
- Kitts, P.A., & Nash, H.A. (1988). An intermediate in the phage λ site-specific recombination reaction is revealed by phosphorothioate substitution in DNA. *Nucleic Acids Res.*, **16**, 6839-6856.
- Koch, C., Ninneman, O., Fuss, H. & Kahmann, R. (1991). The N-terminal part of the *E. coli* DNA binding protein FIS is essential for stimulating site-specific DNA inversion but is not required for specific DNA binding. *Nucl. Acids Res.*, **19**, 5915-5922.
- Kowalczykowski, S.S., Dixon, D.A., Eggleston, A.K., Luder, S.D., & Rehrauer, W.M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.*, **58**, 401-465.
- Kühn, R., Schwenk, F., Aguet, M., Rajewsky, K. & Müller, W. (1995) Inducible Targeting in Mice. *Science*, **269**, 1427-1429.
- Kwon, H.J., Tirumalai, R., Landy, A. & Ellenberger, T. (1997). Flexibility in DNA recombination: structure of the lambda integrase catalytic core. *Science*, **276**, 126–131.
- Laemmli, U.K. (1977). Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature*, **227**, 680-685.
- Landy, A. (1989) Dynamic, structural and regulatory aspects of lambda site-specific recombination. *Annu. Rev. Bioch.*, **58**, 913-949.
- Landy, A. (1993). Mechanistic and structural complexity in the site-specific recombination pathway of Int and FLP. *Curr. Opin. Genet. Dev.*, **3**, 699-707.
- Lange-Gustafson, B.J. & Nash, H.A. (1984). Purification and properties of Int-h, a variant protein involved in site-specific recombination by bacteriophage λ . *J. Biol. Chem.*, **259**, 12724-12732.
- Lakso, M., Sauer, B., Mosinger, B., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L. & Westphal, H. (1992). Targeted oncogene activation by site specific recombination in transgenic mice. *PNAS*, **89**, 6232-6236.

- Le, Y. & Sauer, B. (2000). Conditional gene knockout using Cre recombinase. *Methods Mol. Biol.*, **136**, 477-485.
- Lewandoski, M. (2001). Conditional control of gene expression in mouse. *Nat. Rev. Genet.*, **2**, 743-755.
- Lewis, J.A. & Hatfull, G.F. (2001). Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. *Nucleic Acids Res.*, **29**, 2205-2216.
- Logie, C. & Stewart, A.F. (1995). Ligand-regulated site-specific recombination. *Proc. Natl Acad. Sci. USA*, **92**, 5940-5944.
- Lorbach, E., Christ, N., Schwikardi, M. & Dröge, P. (2000). Site-specific recombination in human cells catalyzed by phage λ integrase mutants. *J. Mol. Biol.*, **296**, 1175-1181.
- Lorbach, E. (2000). Sequenz-spezifische DNA Rekombination durch λ Integrase-Mutanten in humanen Zellen. Diplomarbeit im Fachbereich Biologie der Universität zu Köln.
- Loonstra, A., Vooijs, M., Beverloo, H.B., Al Allak, B., van Drunen, E., Kanaar, R. & Berns, A. (2001). Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc. Natl Acad. Sci. USA*, **98**, 9209-9214.
- Lyznik, L.A., Gordon-Kamm, W.J. & Tao, Y. (2003). Site-specific recombination for genetic engineering in plants. *Plant Cell Rep.*, **21**, 825-932.
- Mao X., Fujiwara Y., Chapdelaine, A., Yang, H. & Orkin S.H. (2001). Activation of EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse strain. *Blood*, **97**, 324-326.
- Marshall, J., Molloy, R., Moss, G.W., Howe, J.R. & Huges, T.E. (1995). The jellyfish green fluorescent protein: a new tool for studying ion channel expression and function. *Neuron*, **14**, 211-215.
- Metzger, D. Clifford, J., Chiba, H. & Chambon, P. (1995). Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc. Natl Acad. Sci. USA*, **92**, 6991-6995.
- Metzger, D. & Feil, R. (1999). Engineering the mouse genome by site-specific recombination. *Curr. Opin. Biotechnol.*, **10**, 470-476.
- Migliaccio, A.R., Bengra, C., Ling, J., Pi, W., Li, C., Zeng, S., Keskinetepe, M., Whitney, B., Sanchez, M., Migliaccio, G. & Tuan, D. (2000). Stable and unstable transgene integration sites in the human genome: extinction of the green fluorescent protein transgene in K562 cells. *Gene*, **256**, 197-214.
- Miller, H.I., Mozola, M.A. & Friedman, D.I. (1980). *int-h*: an *int* mutation of phage λ that enhances site-specific recombination. *Cell*, **20**, 721-729.

- Miller, H.I. & Friedman, D.I. (1981). An *E. coli* gene product required for λ site-specific recombination. *Cell*, **20**, 711-721.
- Mizuuchi, K., Gellert, M., Weisberg, R.A. & Nash, H.A. (1980). Catenation and supercoiling in the products of bacteriophage lambda integrase recombination in vitro. *J. Mol. Biol.*, **141**, 485-494.
- Mizuuchi, K. (1992). Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.*, **61**, 1011-1051.
- Moitoso de Vargas, L., Pargellis, C.A., Hasan, N.M., Bushman, E.W. & Landy, A. (1988). Autonomous DNA binding domains of λ integrase recognize two different sequence families. *Cell*, **54**, 923-929.
- Moitoso de Vargas, L. & Landy, A. (1991). A switch in the formation of alternative DNA loops modulates λ site-specific recombination. *Proc. Natl Acad. Sci. USA*, **88**, 588-592.
- Nash, H.A. (1981). Integration and excision of bacteriophage lambda: the mechanism of conservative site-specific recombination. *Annu. Rev. Genet.*, **15**, 143-167.
- Nash, H.A., Robertson, C.A., Flamm, E., Weisberg, R.A. and Miller, H.I. (1987). Overproduction of Escherichia coli integration host factor, a protein with nonidentical subunits. *J. Bacteriol.*, **169**, 4124-4127.
- Nash, H.A. (1996a). Site-specific recombination: integration, excision, resolution and inversion of defined DNA segment. In: Neidhardt, F.C. et al. (eds) *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, Washington, DC pp. 2363-2376.
- Nash, H.A. (1996b). The HU and IHF proteins: accessory factors for complex protein-DNA assemblies. In Lin, E.C.C. and Lynch, A.S., (eds), *Regulation of Gene Expression in E. coli*. pp. 149-179.
- Nilsson, L., Vanet, A., Vijgenboom, E. & Bosch, L. (1990). The role of Fis in *trans* activation of stable RNA operons of *E. coli*. *EMBO J.*, **9**, 727-734.
- Niwa, H., Yamamura, K. & Miyazaki, J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, **108**, 193-199.
- Novak, A., Guo, C., Yang, W. & Nagy, A. (2000). Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis*, **28**, 147-155.
- Nunes-Düby, S.E., Matsumoto, L. & Landy, A. (1987). Site-specific recombination intermediates trapped with suicide substrates. *Cell*, **50**, 779-788.
- Nunes-Düby, S.E., Azaro, M.A. & Landy, A. (1995a). Swapping DNA strands and sensing homology without branch migration in lambda site-specific recombination. *Curr. Biol.*, **5**, 139-148.

- Nunes-Düby, S.E., Smith-Mungo, L.I. & Landy, A. (1995b). Single base-pair precision and structural rigidity in a small IHF-induced DNA loop. *J. Mol. Biol.*, **253**, 228-242.
- Nunes-Düby, S.E.; Yu, D. & Landy, A. (1997). Sensing homology at the strand-swapping step in lambda excisive recombination. *J. Mol. Biol.*, **272**, 493-508.
- Nunes-Düby, S.E., Tirumalai, R.S., Kwon, H.J., Ellenberger, T., Landy, A. (1998). Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucl. Acids Res.*, **26**, 391-406.
- Numrych, T.E., Gumport, R.I. & Gardner, J.F. (1990). A comparison of the effects of single-base and triple-base changes in the integrase arm-type binding sites on the site-specific recombination of bacteriophage lambda. *Nucl. Acids Res.*, **18**, 3953-3959.
- Numrych, T.E., Gumport, R.I. & Gardner, J.F. (1992). Characterization of the bacteriophage λ excisionase (Xis) protein: the C-terminus is required 'wing' within the minor groove. *Nature Struct. Biol.*, **8**, 84-90.
- O'Gorman, S., Fox, D.T. & Wahl, G.M. (1991). Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science*, **251**, 1351-1355.
- Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. (1997). 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.*, **407**, 313-319.
- Olivares, E.C., Hollis, R.P., and Calos, M.P. (2001). Phage R4 integrase mediates efficient integration in mammalian cells. *Gene*, **278**, 167-176.
- Orban, P.C., Chui, D. & Marth, J.D. (1992). Tissue and site-specific DNA recombination in transgenic mice. *Proc. Natl Acad. Sci. USA*, **89**, 6861-6865.
- Patsey, R.L. & Bruist, M.F. (1995). Characterization of the interaction between the lambda intasome and attB. *J. Mol. Biol.*, **252**, 47-58.
- Paull, T.T. & Johnson, R.C. (1995). DNA looping by *Saccharomyces cerevisiae* high mobility group proteins NHP6A/B. Consequences for nucleoprotein complex assembly and chromatin condensation. *J. Biol. Chem.*, **270**, 8744-8754.
- Phair, R.D. & Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature*, **354**, 604-609.
- Porter, A. (1998). Controlling your losses: conditional gene silencing in mammals. *Trends Genet.*, **14**, 73-79.
- Price, N.T., Jackson, V.N. & Halestrap A.P. (1998). Cloning and sequencing of four new mammalian monocarboxylate transporter (MCT) homologues confirms the existence of a transporter family with ancient past. *Biochem. J.*, **329**, 321-328.
- Radman-Livaja, M., Shaw, C., Azaro, M.A., Biswas, T., Ellenberger, T. & Landy, A. (2003). Arm sequences contribute to the architecture and catalytic function of a λ integrase-Holliday junction complex. *Mol. Cell*, **11**, 783-794.

- Rice, P.A., Yang, S., Mizuuchi, K. & Nash, H.A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295-1306.
- Richet, E., Abcarian, P. & Nash, H. (1986). The interaction of recombination proteins with supercoiled DNA: defining the role of supercoiling in lambda integrative recombination. *Cell*, **46**, 1011-1021.
- Richet, E., Abcarian, P. & Nash, H. (1988). Synapsis of Attachment sites during lambda integrative recombination involves capture of a naked DNA by a protein-DNA complex. *Cell*, **52**, 9-17.
- Ross, W., Thompson, J.F., Newlands, J.T. & Gourse, R.L. (1990). E. coli Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J.*, **9**, 3733-3742.
- Rouet, P., Smith, F. & Jasin, M. (1994). Expression of a site-specific endonuklease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. USA*, **91**, 6064-6068.
- Sadowski, P.D (1995). The FLP recombinase of the 2 µm plasmid of *Saccharomyces cerevisiae*. *Prog. Nucleic Acid Res. Mol. Biol.*, **51**, 53-91.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.
- Sam, M.D., Papagiannis, C.V., Connolly, K.M., Corselli, L., Iwahara, J., Lee, J., Phillips, M., Wojciak, J.M., Johnson, R.C. & Clubb, R.T. (2002). Regulation of directionality in bacteriophage λ site-specific recombination: structure of the Xis protein. *J. Mol. Biol.*, **324**, 791-805.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular Cloning, A laboratory manual. *Cold Spring Harbor Laboratory Press (2nd Edition)*.
- Sarkar, D., Radman-Livaja, M., and Landy, A. (2001). The small DNA binding domain of λ Int is a context-sensitive modulator of recombinase functions. *EMBO J.*, **20**, 1203-1212.
- Sarkar, D., Azaro, M.A., Aihara, H., Papagiannis, C., Tirumalai, R.S., Nunes-Düby, S.E., Johnson, R.C., Ellenberger, T., and Landy, A. (2002). Differential affinity and cooperativity functions of the amino-terminal 70 residues of λ integrase. *J. Mol. Biol.*, **324**, 775-789.
- Sauer, B. (1994). Site-Specific Recombination: developments and applications. *Curr. Opin. Biotechnol.*, **5**, 521-527.
- Sauer, B. (1996). Multiplex Cre/loxP recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. *Nucl. Acids Res.*, **24**, 4608-4613.
- Sauer, B. (1998). Inducible gene targeting in mice using cre/lox system. *Methods*, **14**, 381-392.

- Sauer, B. & Henderson, N. (1989). Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. *Nucl. Acids Res.*, **17**, 147-161.
- Scherer et al., (1953). *J. Exp. Med.*, **97**, 695.
- Schmidt, E.E., Taylor, D.S., Prigge, J.R., Batnett, S. & Capecchi, M. (2000). Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc. Natl Acad. Sci. USA*, **97**, 13702-13707.
- Schwikardi, M. & Dröge, P. (2001). Use of site-specific recombination as a probe of nucleoprotein complex formation in chromatin. *Eur. J. Biochem.*, **268**, 6256-6262.
- Schwikardi, M. & Dröge, P. (2000). Site-specific recombination in mammalian cells catalyzed by $\gamma\delta$ resolvase mutants: implications for the topology of episomal DNA. *FEBS Lett.*, **471**, 147-150.
- Segall, A.M., Goodman, S.D. & Nash, H.A. (1994). Architectural elements in nucleoprotein complexes: interchangeability of specific and non-specific DNA binding proteins. *EMBO J.*, **13**, 4536-4548.
- Sheridan, D.L., Berlot, C.H., Robert, A., Inglis, F.M., Jakobsdottir, K.B., Howe, J.R., & Huges, T.E (2002). A new way to rapidly create functional, fluorescent fusion proteins: random insertion of GFP with an *in vitro* transposition reaction. *BMC Neuroscience*, **3**, 7-17.
- Silver, D.P. & Livingston, D.M. (2001). Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol. Cell*, **8**, 233-243.
- Smith, G.R. (1988). Homologous recombination in procaryotes. *Microbiol. Rev.*, **52**, 1-28.
- Soriano, P., Friedrich, G. & Lawinger, P. (1991). Promoter interactions in retrovirus vectors introduced into fibroblasts and embryonic stem cells. *J. Virol.*, **65**, 2314-2319.
- Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genetics*, **21**, 70-71.
- Sprengler, S.J., Stasiak, A. & Cozzarelli, N.R. (1985). The stereostructure of knots and catenans produced by phage λ integrative recombination: implications for mechanism and DNA structure. *Cell*, **42**, 325-334.
- Stark, W.M., Boocock, M.R., & Sherratt, D.J. (1992). Catalysis by site-specific recombinases. *Trends Genet.*, **8**, 432-439.
- Sternberg, N. & Hamilton, D. (1981). Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *J. Mol. Biol.*, **150**, 467-486.
- Subramaniam, S., Tewari, A.K., Nunes-Düby, S.E. & Foster, M.P. (2003). Dynamics and DNA substrate recognition by the catalytic domain of lambda integrase. *J. Mol. Biol.*, **329**, 423-39.

- Subramanya, H.S., Arciszewska, L.K., Baker, R.A., Bird, L.E., Sherratt, D.J. & Wigley, D.B. (1997). Crystal structure of the site-specific recombinase, XerD. *EMBO J.*, **16**, 5178-5187.
- Swalla, B.M., Gumpert, R.I. & Gardner, J.F. (2003). Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda Integrase core-binding domain. *Nucl. Acids Res.*, **31**, 805-818.
- Takeuchi, T., Nomura, T., Tsujita, M., Suzuki, M., Fuse, T., Mori, H. & Missina, M. (2002). Flp recombinase transgenic mice of C57BL/6 strain for conditional gene targeting. *Biochem. Biophys. Res. Commun.*, **293**, 953-957.
- Tekle, M., Warren, D.J., Biswas, T., Ellenberger, T., Landy, A. & Nunes-Duby, S.E. (2002). Attenuating functions of the C Terminus of lambda integrase. *J. Mol. Biol.*, **324**, 649-665.
- Thaler, D.S. & Stahl, F. (1988). *Annu. Rev. Gen.*, **22**, 169-197.
- Thompson, J.F., de Vargas, L., Koch, C., Kahmann, R. & Landy A. (1987). Cellular factors couple recombination with growth phase: characterization of a new component in the lambda site-specific recombination pathway. *Cell*, **50**, 901-908.
- Thompson, J.F., Moitoso de Vargas, L., Skinner, S.E. & Landy, A. (1987, b). Protein-protein interactions in a higher-order structure direct lambda site-specific recombination. *J. Mol. Biol.*, **195**, 481-493.
- Thompson, J.F. & Landy, A. (1988). Empirical estimation of protein-induced DNA bending angles: applications to λ site-specific recombination complexes. *Nucl. Acids Res.*, **16**, 9687-9705.
- Thyagarajan, B., Guimaraes, M.J., Groth, A.C. and Calos, M.P. (2000). Mammalian genomes contain active recombinase recognition sites. *Gene*, **244**, 47-54.
- Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S. and Calos, M.P. (2001). Site-specific genomic integration in mammalian cells mediated by phage ϕ C31 integrase. *Mol. Cell. Biol.*, **21**, 3926-3934.
- Tirumalai, R.S., Kwon, H.J., Healey, E., Landy, A. (1997). The catalytic domain of λ site-specific recombinase. *Proc. Natl Acad. Sci. USA*, **94**, 6104-6109.
- Tirumalai, R.S., Kwon, H.J., Cardente, E.H., Ellenberger, T. & Landy, A. (1998). Recognition of core-type DNA sites by lambda integrase. *J. Mol. Biol.*, **279**, 513-527.
- Travers, A. (1997). DNA-protein interactions: IHF – the master bender. *Curr. Biol.*, **7**, R252-R254.
- Tsien, R.Y. (1998). The green fluorescent protein. *Annu. Rev. Biochem.*, **67**, 509-544.
- Voziyanov, Y. Pathania, S. & Jayaram, M. (1999). A general model for site-specific recombination by the integrase family recombinases. *Nucl. Acids Res.*, **27**, 930-941.

- Wang, S. & Hazelrigg, T. (1994). Implication for bcd mRNA localization from spatial distribution of exu protein in *Drosophila oogenesis*. *Nature*, **369**, 400-403.
- Wang, H. & Chong, S. (2003). Visualization of coupled protein folding and binding in bacteria and purification of the heterodimeric complex. *Proc. Natl. Acad. Sci., USA*, **100**, 478-483.
- Warren, D., Sam, M.D., Manley, K., Sarkar, D., Lee, S.Y., Abbani, M., Wojciak, J.M., Clubb, R.T. & Landy, A. (2003). Identification of the lambda Integrase surface that interacts with Xis reveals a residue that is also critical for Int dimer formation. *Proc. Natl Acad. Sci., USA*, **100**, 8176–8181.
- Werdien, D., Peiler, G. & Ryffel, U. (1999). FLP and Cre recombinase functions in *Xenopus* embryos. *Nucl. Acids Res.*, **29**, 147-161.
- Wilson, T. & Kola, I. (2001). The loxP/Cre system and genome modifications. *Methods Mol. Biol.*, **158**, 83-94.
- Wojciak, J.M., Sarkar, D., Landy, A. & Clubb, R.T. (2002). Arm-site binding by the lambda Integrase protein: solution structure and functional characterization of its amino-terminal domain. *Proc. Natl Acad. Sci. USA*, **99**, 3434-3439.
- Wu, Z., Gumport, R.I. & Gardner, J.F. (1998). Defining the structural and functional roles of the carboxyl region of the bacteriophage λ excisionase (Xis) protein. *J. Mol. Biol.*, **281**, 651-661.
- Xu, J. & Johnson, R.C. (1995). Identification of genes negatively regulated by Fis: Fis and RpoS comodulate growth-phase-dependent gene. *J. Bacteriol.*, **177**, 938-947.
- Xu, J. & Johnson, R. C. (1995, b). Fis activates the RpoS-dependent stationary phase expression of *proP* in *Escherichia coli*. *J. Bacteriol.*, **177**, 5222-5231.
- Yang, S-W. & Nash, H.A. (1995). Comparison of protein binding to DNA *in vivo* and *in vitro*: defining an effective intracellular target. *EMBO J.*, **14**, 6292-6300.
- Yin, S., Bushman, W. & Landy, A. (1985). Interaction of the λ site-specific recombination protein Xis with attachment site DNA. *Proc. Natl Acad. Sci., USA*, **82**, 1040-1044.
- Yu, Y. & Bradley, A. (2001). Engineering chromosomal rearrangements in mice. *Nature Rev. Genet.*, **2**, 780-790.

7 Anhang

Komplette Sequenz um den *attH2*-Locus (*human putative monocarboxylate transporter (MCT5)* auf dem Chromosom 1: Nucleotide 27181 bis 28060:
 Sequenzinformation aus der EMBL-Datenbank (*accession no.* AC025987)

```

27181  5'-AAGAATTGCC TTTCTTTCTT CTCTTAAGGT CCCCTGGTTG CTATTATTTG
      3'-TTCTTAACGG AAAGAAAGAA GAGAATTCCA GGGGACCAAC GATAATAAAC

27231      TGACATACTT GGAGAGAAAA CTACCTCCAT TCTTGGGGCT TTCGTTGTTA
      ACTGTATGAA CCTCTCTTTT GATGGAGGTA AGAACCCCGA AAGCAACAAT

27281      CTGGTGGATA TCTGATCAGC AGCTGGGCCA CAAGTATTCC TTTTCTTTGT
      GACCACCTAT AGACTAGTCG TCGACCCGGT GTTCATAAAG AAAAGAAACA
      MCTII →

27331      GTGACTATGG GACTTCTACC CGGTGAGTCC ATTGTAATTA TGTGACCATA
      CACTGATACC CTGAAGATGG GCCACTCAGG TAACATTAAT ACACTGGTAT

27381      GAAGGCCAGG AGTGGTGGCT CACACTTGTA ATCCTAGCAC TTTGGGAGGC
      CTTCCGGTCC TCACCACCGA GTGTGAACAT TAGGATCGTG AAACCCTCCG

27431      CAAAGTGGGT GGATCGCTTG AGTCCAGTTC GAGACCAGCC TGGGCAACAT
      GTTTCACCCA CCTAGCGAAC TCAGGTCAAG CTCTGGTCGG ACCCGTTGTA

27481      GGTGAAACCC TGTCTCTACT AAAAATACAA AAATTAGCCA GGCCAGTGGT
      CCACTTTGGG ACAGAGATGA TTTTTATGTT TTTAATCGGT CCGGTCACCA

27531      GGCACATGCC TGTAGTCCCA GCTATTTGGG AAGCTGAGGT GGGAGGATTG
      CCGTGTACGG ACATCAGGGT CGATAAACCC TTCGACTCCA CCCTCCTAAC

27581      CTTGAGCCCA GGAAGCTGAG GCTGTTGTGA GCCGAGATTG TGCCACTGCA
      GAACTCGGGT CCTTCGACTC CGACAACACT CGGCTCTAAC ACGGTGACGT

27631      CTCCAGTCTG GGTGACAAAG TGAGACCCTG TCTCAAATA TGTGTGTGTG
      GAGGTCAGAC CCACTGTTTC ACTCTGGGAC AGAGTTTTAT ACACACACAC

27681      TGTGTGTATG TATATATGCA TCCATCCATA GAAGATAGAA GATGGAACAG
      ACACACATAC ATATATACGT AGGTAGGTAT CTTCTATCTT CTACCTTGTC

27731      GACTCTTCCC TTCTACTGTT CCTGGGTCCA AAAGGCCTAG CTGTATTATC
      CTGAGAAGGG AAGATGACAA GGACCCAGGT TTTCCGGATC GACATAATAG

27781      TCGGATCCAG GAGATTATCT TTTCTGTTAC CAAGCCACTT AATGTAAGAG
      AGCCTAGGTC CTCTAATAGA AAAGACAATG GTTCGGTGAA TTACATTCTC
    
```


Komplette Nukleinsäure-Sequenz des rHF:

```

1  5'-ATGGGCACCA AGTCAGAATT GATAGAAAGA CTTGCCACCC AGCAATCGCA
   3'-TACCCGTGGT TCAGTCTTAA CTATCTTTCT GAACGGTGGG TCGTTAGCGT

51  CATTCCCGCC AAGACGGTTG AAGATGCAGT AAAAGAGATG CTGGAGCATA
   GTAAGGGCGG TTCTGCCAAC TTCTACGTCA TTTTCTCTAC GACCTCGTAT

101 TGGCCTCGAC TCTTGCGCAG GGTGGAAGCG GCGGTCTTAC AAAAGCTGAA
    ACCGGAGCTG AGAACGCGTC CCACCTTCGC CGCCAGAATG TTTTCGACTT

151 ATGTCAGAAT ATCTGTTTGA TAAGCTTGGG CTTAGCAAGC GGGATGCCAA
    TACAGTCTTA TAGACAAACT ATTTCGAACCC GAATCGTTTC CCCTACGGTT

201 AGAACTGGTT GAACTGTTTT TCGAAGAGAT CCGTCGCGCT CTGGA AAAACG
    TCTTGACCAA CTTGACAAAA AGCTTCTCTA GGCAGCGCGA GACCTTTTGC

251 GCGAACAGGT GAAACTCTCT GGTTTTGGTA ACTTCGATCT GCGTGATAAG
    CGCTTGTTCCA CTTTGAGAGA CCAA AACCAT TGAAGCTAGA CGCACTATTC

301 AATCAACGCC CGGGACGTAA CCCGAAAACG GGCGAGGATA TTCCATTAC
    TTAGTTGCGG GCCCTGCATT GGGCTTTTGC CCGCTCCTAT AAGGGTAATG

351 AGCACGGCGC GTGGTGACCT TCAGACCCGG GCAGAAGTTA AAAAGCCGGG
    TCGTGCCGCG CACCACTGGA AGTCTGGGCC CGTCTTCAAT TTTTCGGCCC

401 TCGAAAACGC TGGTGGGGGC GAGCGTATTG AAATCCGCGG TTTCGGCAGT
    AGCTTTTTCG ACCACCCCGC CTCGCATAAC TTTAGGCGCC AAAGCCGTCA

451 TTCTCTTTGC ACTACCGCGC ACCACGTACC GGACGTAATC CGAAGACTGG
    AAGAGAAACG TGATGGCGCG TGGTGCATGG CCTGCATTAG GCTTCTGACC

501 CGATAAAGTA GAACTGGAAG GAAAATACGT TCCTCACTTT AAACCTGGTA
    GCTATTTTCAT CTTGACCTTC CTTTTATGCA AGGAGTGAAA TTTGGACCAT

551 AAGAACTGCG CGATCGCGCC AATATCTACG GTTGATAA-3'
    TTCTTGACGC GCTAGCGCGG TTATAGATGC CAACTATT-5'

```

1. Linker

GGTGGAAGCG GCGGTCTTAC
CCACCTTCGC CGCCAGAATG

2. Linker

TGGTGGGGGC
ACCACCCCGC

Komplette Nukleinsäure-Sequenz der Lambda Integrase Mutanten:

Int-h(E174K) und Int-h/218(E174K/E218K)

```

1  5' -ATGGGAAGAA GCGAAGTCA TGAGCGCCGG GATTTACCCC CTAACCTTTA
   3' -TACCCTTCTT CCGCTTCAGT ACTCGCGGCC CTAAATGGGG GATTGGAAAT

51  TATAAGAAAC AATGGATATT ACTGCTACAG GGACCCAAGG ACGGGTAAAG
   ATATTCTTTG TTACCTATAA TGACGATGTC CCTGGGTTCC TGCCCATTTT

101 AGTTTGGATT AGGCAGAGAC AGGCGAATCG CAATCACTGA AGCTATACAG
   TCAAACCTAA TCCGTCTCTG TCCGCTTAGC GTTAGTGACT TCGATATGTC

151 GCCAACATTG AGTTATTTTC AGGACACAAA CACAAGCCTC TGACAGCGAG
   CGGTTGTAAC TCAATAAAAG TCCTGTGTTT GTGTTCGGAG ACTGTCGCTC

201 AATCAACAGT GATAATTCCG TTACGTTACA TTCATGGCTT GATCGCTACG
   TTAGTTGTCA CTATTAAGGC AATGCAATGT AAGTACCGAA CTAGCGATGC

251 AAAAAATCCT GGCCAGCAGA GGAATCAAGC AGAAGACACT CATAAATTAC
   TTTTTTAGGA CCGGTCGTCT CCTTAGTTTCG TCTTCTGTGA GTATTTAATG

301 ATGAGCAAAA TTAAAGCAAT AAGGAGGGGT CTGCCTGATG CTCCACTTGA
   TACTCGTTTT AATTTCTGTTA TTCCTCCCCA GACGGACTAC GAGGTGAACT

351 AGACATCACC ACAAAGAAA TTGCGGCAAT GCTCAATGGA TACATAGACG
   TCTGTAGTGG TGTTTTCTTT AACGCCGTTA CGAGTTACCT ATGTATCTGC

401 AGGGCAAGGC GGCCTCAGCC AAGTTAATCA GATCAACACT GAGCGATGCA
   TCCCGTTCCG CCGCAGTCGG TTCAATTAGT CTAGTTGTGA CTCGCTACGT

451 TTCCGAGAGG CAATAGCTGA AGGCCATATA ACAACAAACC ATGTCTGCTGC
   AAGGCTCTCC GTTATCGACT TCCGGTATAT TGTTGTTTGG TACAGCGACG
  
```

AAG (E174K)

```

501 CACTCGCGCA GCAAATCAG AGGTAAGGAG ATCAAGACTT ACGGCTGACG
   GTGAGCGCGT CGTTTTAGT TCATTTCCTC TAGTTCTGAA TGCCGACTGC
  
```

E174

```

551 AATACCTGAA AATTTATCAA GCAGCAGAAT CATCACCATG TTGGCTCAGA
   TTATGGACTT TTAAATAGTT CGTCGTCTTA GTAGTGGTAC AACCGAGTCT
  
```

```

601 CTTGCAATGG AACTGGCTGT TGTTACCGGG CAACGAGTTG GTGATTTATG
   GAACGTTACC TTGACCGACA ACAATGGCCC GTTGCTCAAC CACTAAATAC
  
```

AAA (E218K)

```

651 CGAAATGAAG TGGTCTGATA TCGTAGATGG ATATCTTTAT GTCGAGCAAA
   GCTTTACTTC ACCAGACTAT AGCATCTACC TATAGAAATA CAGCTCGTTT
  
```

E218

ANHANG

701 GCAAAACAGG CGTAAAAATT GCCATCCCAA CAGCATTGCA TATTGATGCT
CGTTTTGTCC GCATTTTAA CGGTAGGGTT GTCGTAACGT ATAACTACGA

751 CTCGGAATAT CAATGAAGGA AACACTTGAT AAATGCAAAG AGATTCTTGG
GAGCCTTATA GTTACTTCCT TTGTGAACTA TTTACGTTTC TCTAAGAACC

801 CGGAGAAACC ATAATTGCAT CTACTCGTCG CGAACCGCTT TCATCCGGCA
GCCTCTTTGG TATTAACGTA GATGAGCAGC GCTTGGCGAA AGTAGGCCGT

851 CAGTATCAAG GTATTTTATG CGCGCACGAA AAGCATCAGG TCTTTCCTTC
GTCATAGTTC CATAAAATAC GCGCGTGCTT TTCGTAGTCC AGAAAGGAAG

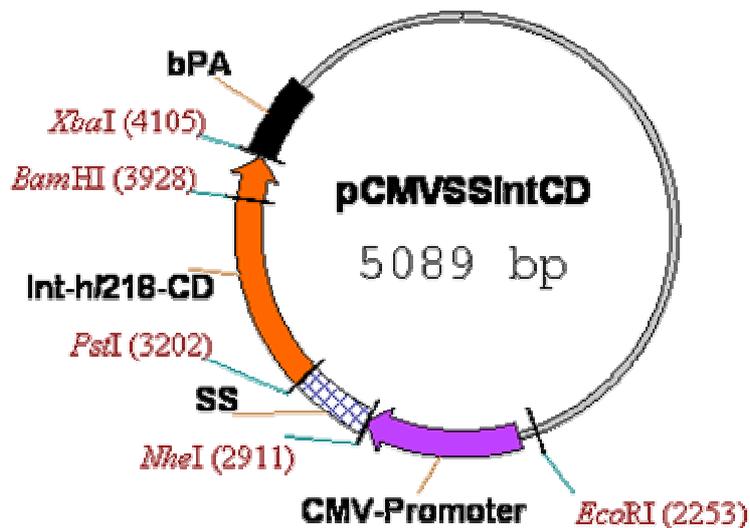
901 GAAGGGGATC CGCCTACCTT TCACGAGTTG CGCAGTTTGT CTGCAAGACT
CTTCCCCTAG GCGGATGGAA AGTGCTCAAC GCGTCAAACA GACGTTCTGA

951 CTATGAGAAG CAGATAAGCG ATAAGTTTGC TCAACATCTT CTCGGGCATA
GATACTCTTC GTCTATTCGC TATTCAAACG AGTTGTAGAA GAGCCCGTAT

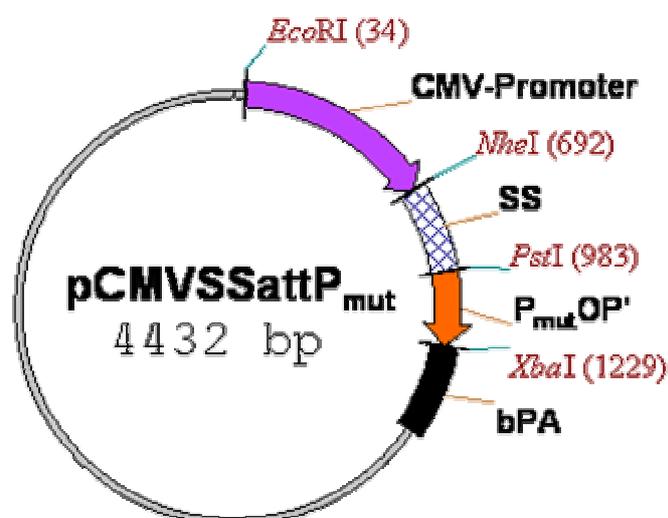
1001 AGTCGGACAC CATGGCATCA CAGTATCGTG ATGACAGAGG CAGGGAGTGG
TCAGCCTGTG GTACCGTAGT GTCATAGCAC TACTGTCTCC GTCCCTCACC

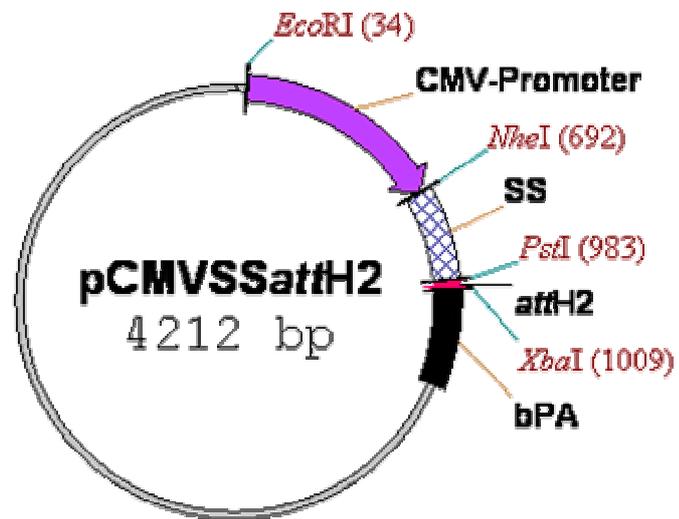
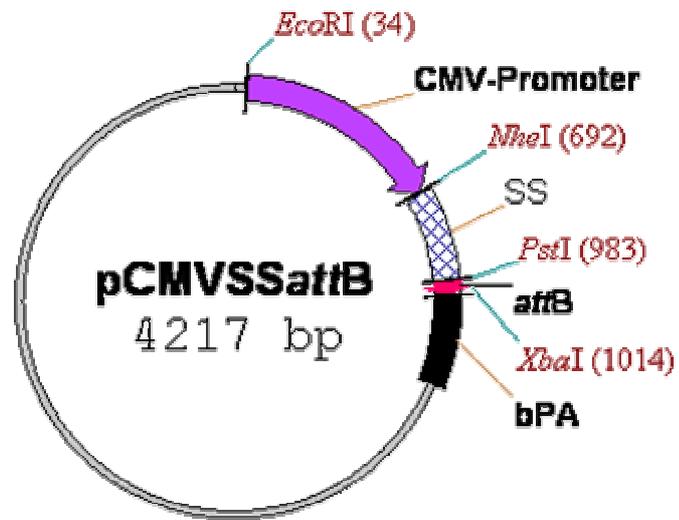
1051 GACAAAATTG AAATCAAATA ATGA-3'
CTGTTTTAAC TTTAGTTTAT TACT-5'

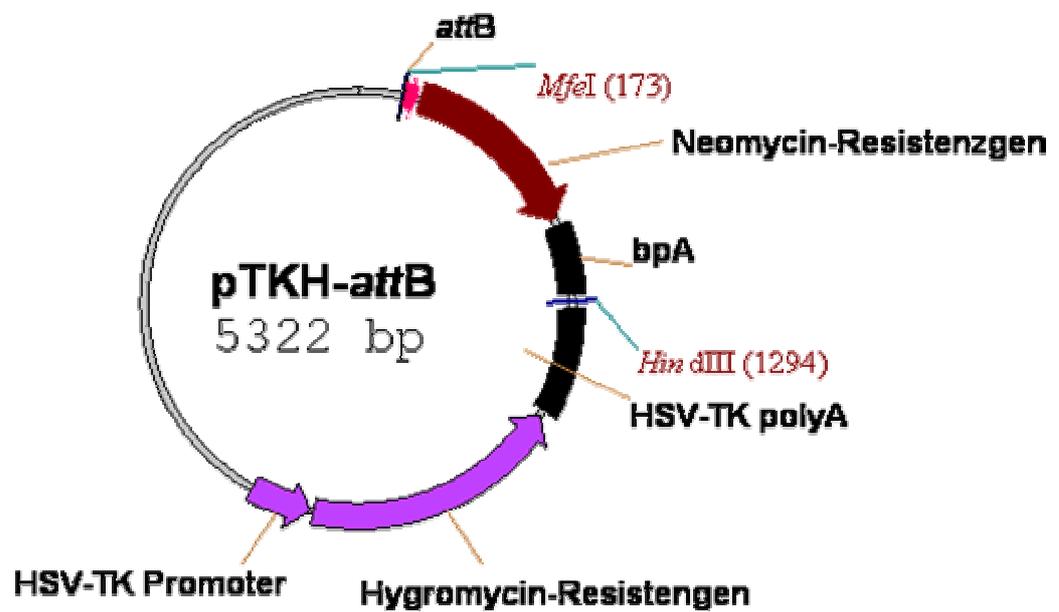
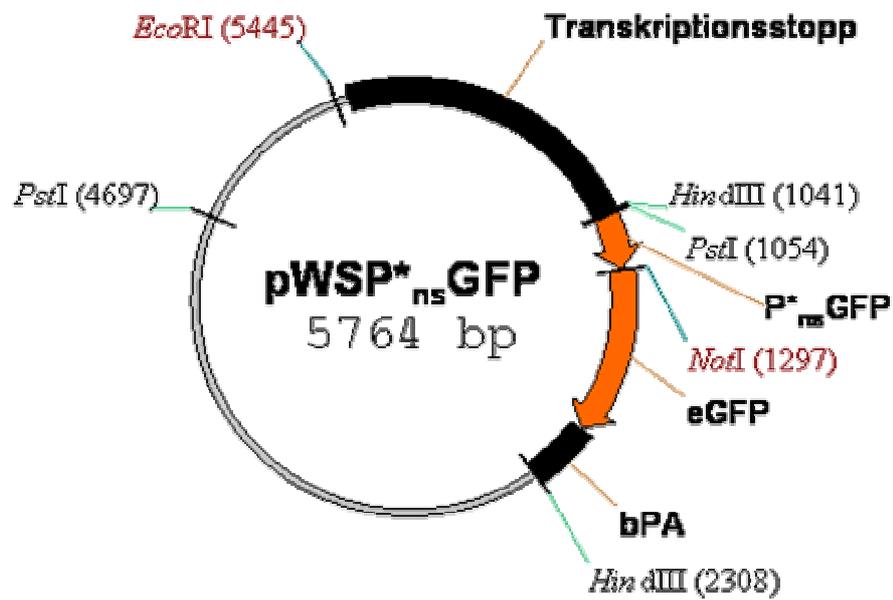
Schematische Darstellung, der in dieser Arbeit generierten Vektoren. Genauere Angaben zur Klonierung und zu den verwendeten Bezeichnungen und Abkürzungen finden sich in den Kapiteln 3.1.9 bis 3.1.12 und im Abkürzungsverzeichnis.

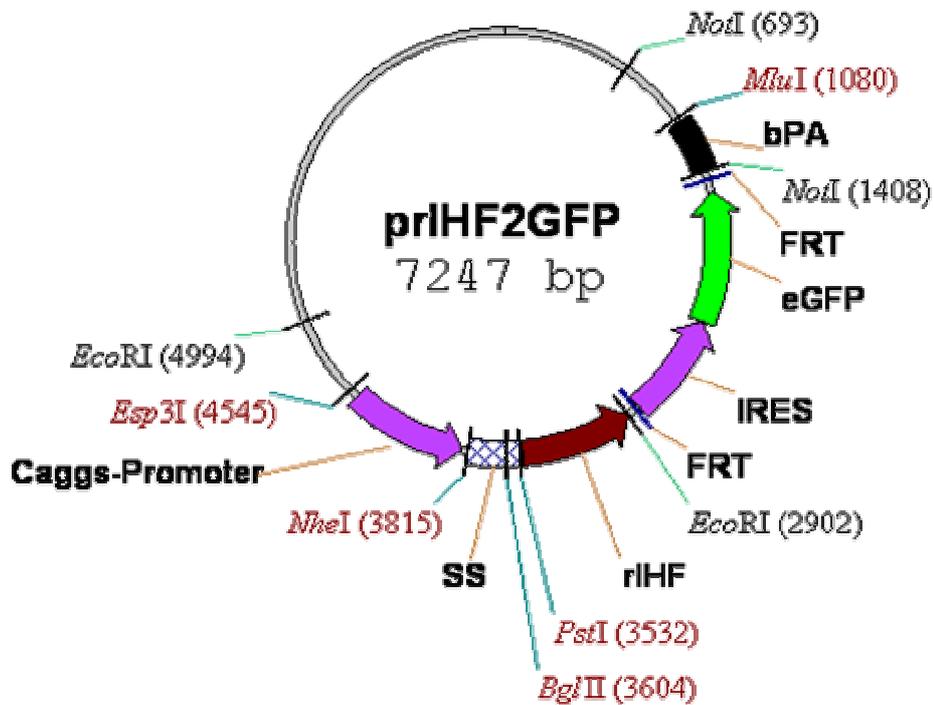
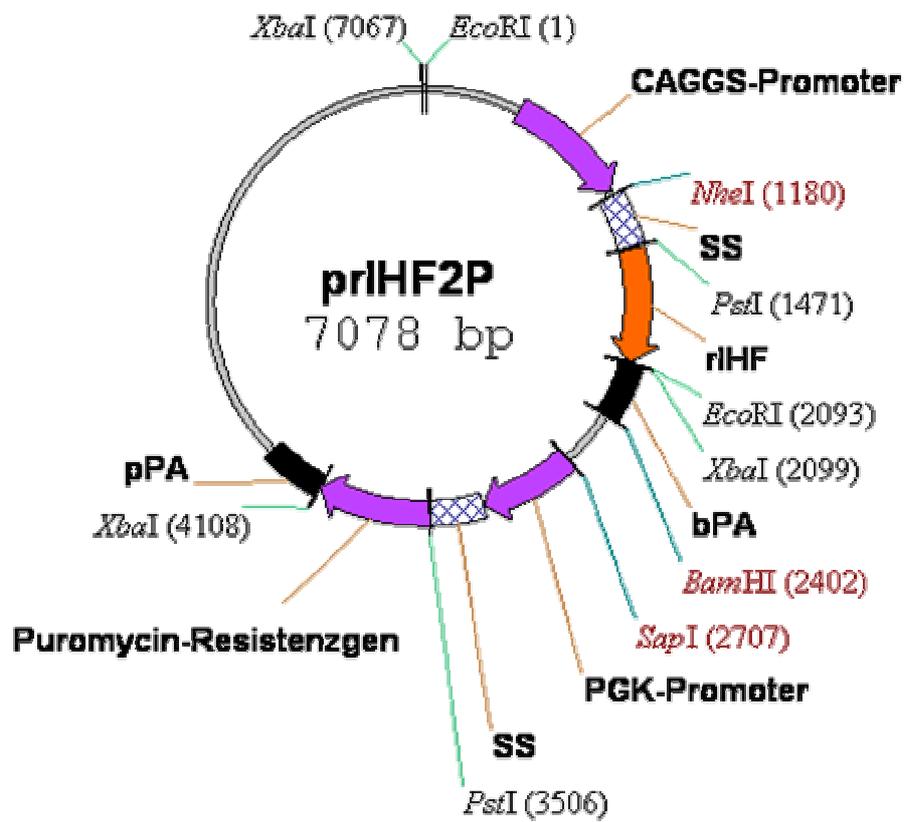


pCMVSSInt-h/218 entspricht im Aufbau pCMVSSIntCD. Lediglich die Sequenz der Integrase enthält die beschriebenen Mutationen E174K und E218K (siehe auch S. 115).









8 Zusammenfassung

Im Gegensatz zur wild-typ-Integrase des Phagen *Lambda* katalysieren die Mutanten Int-h und Int-h/218 die sequenz-spezifische Rekombination zwischen zwei *attachment*-Regionen auch in Abwesenheit von Kofaktoren wie IHF, XIS und negativer gespannter DNA. In dieser Arbeit wurde untersucht, welchen Effekt das Milieu menschlicher Zellen auf diese Reaktionen hat. Der Vergleich der hier durchgeführten intermolekularen Rekombinationsreaktionen zwischen episomalen Substraten mit den Kinetiken von intramolekularen Reaktionen zeigt, dass offensichtlich keine Präferenz für einen der beiden Reaktionstypen besteht. Es konnte gezeigt werden, dass zwei 21 bp umfassende Rekombinationsregionen, die nur DNA-Sequenz des *core*-Typs enthalten, durch Integrase Mutanten intermolekular rekombiniert werden können. Die Effizienz dieser Reaktion wurde durch die Anwesenheit von Sequenzen des *Arm*-Typs mehrfach gesteigert. Integrase-Proteine mit deletierter N-terminaler Domäne, die ebenfalls in dieser Arbeit generiert und getestet wurden, zeigen dagegen kaum Rekombinationsaktivität in eukaryotischen Zellen. Neben der *Arm*-Bindungs-Funktion der N-Domäne wurde somit auch ein Hinweis auf eine funktionelle Rolle bei der Rekombination erbracht.

Bei der Analyse der chromosomalen Rekombinationsaktivität wurde für Int-h/218 nachgewiesen, dass neben der Inversion eines mit *attB/attP* flankierten Reportergens prinzipiell auch die Integration eines *attP*-tragenden Vektors in eine im Genom von HeLa-Zellen befindliche *attB*-Sequenz katalysiert wird.

Das heterodimere IHF (Integration Host Factor) ist ein sequenz-spezifisches DNA-bindendes und -biegendes Protein aus *E.coli* und Kofaktor der *Lambda*-Integrase. Es spielt eine wichtige Rolle in einer Vielzahl von DNA-Transaktionen einschließlich Rekombination, Transkription und Replikation. In eukaryotischen Zellen führt die Ausprägung seiner beiden Untereinheiten nicht zur Ausbildung einer nachweisbaren Proteinmenge. Für den Einsatz in Eukaryoten wurde daher ein rekombinantes IHF Protein (rIHF) untersucht, dass durch die Insertion der nahezu kompletten α -Untereinheit von IHF in seine β -Untereinheit erstellt wurde. Dieses rIHF wurde in HeLa-Zellen stabil ausgeprägt und toleriert. Es wurde vornehmlich im Zellkern lokalisiert und stimulierte in rIHF-transgenen HeLa-Zelllinien die integrative Rekombination episomaler Substrate durch die wild-typ Integrase. Für weitere *in vivo* Studien mit rIHF wurden transgene Mäuse generiert und analysiert, in denen eine Expression des rekombinanten Proteins jedoch nicht nachgewiesen werden konnte.

9 Summary

In contrast to wild-type Integrase of bacteriophage *Lambda* the mutant proteins Int-h and Int-h/218 catalyze site-specific recombination between two attachment sites in the absence of accessory factors such as IHF, XIS and negative DNA supercoiling. In this thesis I examined which effect the human cellular environment exerts on these reactions. The comparison of intermolecular recombination reactions between episomal substrates with kinetics of intramolecular reactions, shows that apparently no preference for one of the two reaction types exists. In addition it was shown that two copies of recombination sites containing only the 21 bp comprising core-type DNA sequence could be recombined by mutant Integrases. The efficiency of this reaction was increased several-fold by the presence of arm-type DNA sequences. It could further be demonstrated that an N-terminal truncated mutant Integrase exhibited only a very weak recombinogenic activity in a eukaryotic background. This result provides a strong hint for a functional role of the N-domain in addition to its arm-type DNA binding.

The analysis of chromosomal recombination activity showed for Int-h/218 that it catalyzes not only the inversion of an *attB/attP* flanked reporter gene but also the integration of an *attP* containing vector into an *attB*-site placed in the HeLa-genome.

The heterodimeric IHF (integration host factor) is a sequence-specific DNA-binding and DNA-bending protein from *E.coli* that functions as cofactor for the Integrase of bacteriophage *Lambda*. It plays an important role in a variety of DNA transactions including recombination, transcription and replication. In eukaryotic cells the expression of the two subunits comprising IHF does not lead to the assembly of a detectable amount of protein. A recombinant IHF protein (rIHF), generated by the insertion of the almost complete α -subunit of IHF into the β -subunit, was therefore tested for the use in eukaryotes. This rIHF was stably expressed and tolerated by HeLa cells. It is localized primarily in the cell nucleus and triggers the integrative recombination by wild-type Int in rIHF transgenic HeLa cell lines. For further in vivo studies with rIHF transgenic mice were generated and analyzed. However, the expression of the recombinant protein could not be detected in these mice.

Danksagungen

Das in dieser Dissertation beschriebene Projekt wurde zwischen Mai 2000 und September 2003 am Institut für Genetik der Universität zu Köln unter Anleitung von Prof. Dr. Peter Dröge durchgeführt. Das Projekt wurde durch das Zentrum für molekulare Medizin Köln finanziert.

An dieser Stelle möchte ich mich besonders bei meinem Betreuer Peter Dröge für seine ständige Unterstützung, Förderung und die gewährten experimentellen Freiheiten während meiner Doktorarbeit bedanken.

Ich möchte mich auch bei meiner Arbeitsgruppe bedanken, für die durchgehend angenehme Atmosphäre und die fachliche und menschliche Unterstützung, die entscheidend zur Entstehung dieser Arbeit beigetragen hat. Insbesondere seien hier Micha Schwikardi und Nicole Christ erwähnt, die jederzeit ein offenes Ohr für Fragen und Diskussionen hatten und mit denen die Arbeit immer viel Spaß gemacht hat.

Benno Müller-Hill (Universität Köln), Klaus Rajewsky (Universität Köln), Ari Waismann (Universität Köln) und Walter Dörfler (Universität Köln) danke ich für die Bereitstellung von Räumlichkeiten, Geräten und Reagenzien.

Bei allen anderen Mitarbeitern des Hauses möchte ich mich für die geleistete Hilfe in diversen Angelegenheiten bedanken und hoffe sie nehmen es mir nicht übel, wenn sie hier nicht einzeln aufgezählt sind.

Bei Karin Otto, Christoph Göttlinger und Dennis Webb bedanke ich mich für ihre technische Unterstützung. Bei Angela Egert, Sigrid Irlenbusch und Brigitte Hampel für die Hilfestellung in der Zellkultur. Brigitte Hampel danke ich besonders für die Unterweisung in der Maus-Präparation. Nathalie Uyttersprot und Stefano Casola danke ich für die Bereitstellung des p-FRT-IRES/GFP-FRT-Vektors.

Allen Mitarbeitern der 2. Etage danke ich außerdem für die herzliche Kollegialität, die mir entgegengebracht wurde. Sie hat besonders in den letzten anderthalb Jahren meine wissenschaftliche Motivation innerhalb und außerhalb des Instituts bestärkt. Kristina, Gloria, Nadine, Sonja, Nathalie, Casi, Ari und allen anderen: Danke! Mein besonderer Dank gilt hier Thomas Wunderlich, der mit freundschaftlicher Zusammenarbeit und fachlichem Rat weite Strecken meiner Arbeit begleitet hat.

Mein Dank gilt außerdem meinen Freunden außerhalb der Universität, die mir in dieser Zeit sehr viel Verständnis und Geduld entgegengebracht und mich mit viel Liebe moralisch unterstützt haben. Insbesondere Rainer musste einiges ertragen und ist mit Sicherheit die wichtigste Stütze in dieser Zeit gewesen, dafür danke ich dir.

Zum Schluss sei aber vor allem meiner Familie gedankt, meinen Eltern, E und Pi ohne die ich nie soweit gekommen wäre.

Lebenslauf

Persönliche Angaben

Name: Teresa Corona
Geburtsdatum: 17.12.71
Geburtsort: Mönchengladbach
Familienstand: ledig
Eltern: Pietro Corona und Vittoria Corona geb. Corona
Staatsangehörigkeit: italienisch
Wohnort: Mühlenbach 28, 50676 Köln

Ausbildung

Mai 2000 - Mai 2003 Dissertation am Institut für Genetik der Universität zu Köln unter Anleitung von Prof. Dr. Peter Dröge.
Thema: *Funktionelle Untersuchungen zur sequenz-spezifischen Rekombination durch die Integrase des Bakteriophagen Lambda in eukaryotischen Zellen*

Aug. 1998 - Jul. 1999 Diplomarbeit am Institut für Biochemie der Universität zu Köln unter Anleitung von Prof. Dr. H. W. Klein.
Thema: *Klonierung und Expression modifizierter löslicher Insulinrezeptorkinasen und C-terminaler Untereinheiten.*

Juli 1998 Diplom-Hauptprüfung in den Fächern: Biochemie, Genetik und Pharmakologie

Januar 1996 Diplom-Vorprüfung

WS 1993/1994 Immatrikulation im Fach Biologie an der Universität zu Köln

1991 - 1993 Ausbildung zur Bankkauffrau bei der Stadtsparkasse Mönchengladbach

1991 Abitur

1982 - 1991 Bischöfliche Marienschule, Mönchengladbach

1978 - 1982 Montessori Grundschule, Mönchengladbach

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von den unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. P. Dröge betreut worden.

(Teresa Corona)

Teilpublikationen

Corona T, Bao Q, Christ N, Schwarz T, Li J, Dröge P.

Activation of site-specific DNA integration in human cells by a single chain integration host factor

Nucl. Acids Res. 2003, Sept 1;31(17):5440-8

PMID: 12930965 [PubMed - indexed for MEDLINE]

Christ N, **Corona T**, Dröge P.

Site-specific recombination in eukaryotic cells mediated by mutant lambda integrases: implications for synaptic complex formation and the reactivity of episomal DNA segments.

J Mol Biol. 2002 May 31;319(2):305-14.

PMID: 12051908 [PubMed - indexed for MEDLINE]