

**Gezielte genetische Manipulation durch
Lambda Integrasen in embryonalen
Stammzellen der Maus und in anderen
Säugetierzellen**

Inaugural-Dissertation
zur
Erlangung des Doktorgrades
an der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Nicole Christ
aus Boppard

Copy Team, Köln
2002

Berichterstatter: Priv.-Doz. Dr. Peter Dröge
Prof. Dr. Thomas Langer

Tag der mündlichen Prüfung: 13. Mai 2002

Für meine Eltern

Inhaltsverzeichnis

ABKÜRZUNGSVERZEICHNIS

A	EINLEITUNG	1
A1	Sequenz-spezifische Rekombinasen: Ein wertvolles Werkzeug zur gezielten DNA-Manipulation.....	1
A2	Sequenz-spezifische Rekombinasen in eukaryotischen Zellen.....	3
A2.1	Das Cre/loxP-System.....	3
A2.2	Das FLP/FRT-System.....	4
A2.3	Limitierte Anwendungsmöglichkeiten etablierter Rekombinationssysteme.....	5
A3	Das sequenz-spezifische Rekombinationssystem des Bakteriophagen <i>Lambda</i>.....	7
A3.1	Die integrative λ -Rekombination.....	9
A3.2	Die exzisive λ -Rekombination.....	11
A3.3	Die <i>Lambda</i> Integrase.....	12
A3.4	Die Hilfsproteine der <i>Lambda</i> Integrase: IHF, XIS und FIS.....	15
A4	Das Ziel dieser Studie.....	17
B	MATERIAL	19
B1	Bakterienstämme	19
B2	Chemikalien.....	19
B3	Enzyme	19
B4	Synthetische Oligonucleotide.....	19
B5	Medien für bakterielle Kulturen.....	21
B6	Plasmide.....	21
B7	Medien und Puffer für die eukaryotische Zellkultur	23
B8	Eukaryotische Zelllinien	24
C	METHODEN	25
C1	Molekularbiologische Methoden.....	25
C1.1	Agarosegelelektrophorese	25
C1.2	DNA-Reinigung aus Agarosegelen.....	25
C1.3	Plasmid-Isolierung aus Bakterien	25
C1.4	DNA-Isolierung aus eukaryotischen Zellen.....	26
C1.5	Polymerasekettenreaktion (PCR)	26
C1.6	RT-PCR Analyse von MCT5 mRNA-Ausprägung in HeLa und 293 Zellen	27
C1.7	DNA Sequenzierung	28
C1.8	<i>Southern Blot</i> Analyse	28
C1.9	Konstruktion der Expressionsvektoren	29
C1.10	Konstruktion der Substratvektoren	33

C1.11	Konstruktion der <i>Target</i> Vektoren (<i>Gene Targeting</i>)	36
C2	Biochemische Methoden	37
C2.1	SDS-Polyacrylamid-Gelelektrophorese	37
C2.2	<i>Western Blot</i> Analyse.....	38
C3	Zellbiologische Methoden	39
C3.1	Kultur von Säugetierzellen	39
C3.2	Konstruktion stabiler HeLa-Reporter-Zelllinien	39
C3.3	Kultur von embryonalen Stammzellen der Maus (MES).....	40
C3.4	Durchflusszytometrie	41
C3.5	Fluoreszenz Mikroskopie	41
C4	Rekombinationsanalysen	41
C4.1	Rekombinationsanalysen in <i>E.coli</i>	41
C4.2	Episomale Rekombinationsanalysen in eukaryotischen Zellen.....	42
C4.3	Rekombinationsanalysen in HeLa-Reporter-Zelllinien	44
C4.4	Rekombinationsanalysen in ES-Reporter-Zelllinien	44
D	ERGEBNISSE	46
D1	Ausprägung der <i>Lambda</i> Integrase in Säugetierzellen	46
D2	Episomale sequenz-spezifische Rekombination in Säugetierzellen durch Integrase Mutanten	47
D2.1	Analyse der enzymatischen Aktivitäten von Int und den Derivaten Int-h und Int-h/218 in intra- und intermolekularen Rekombinationreaktionen	48
D2.2	Die DNA Arm-Bindungssequenzen der <i>att</i> -Regionen stimulieren die Rekombination; sie sind jedoch nicht notwendig	51
D2.3	Ausschluss einer alternativen Verknüpfung der <i>att</i> -Sequenzen.....	53
D2.4	Episomale sequenz-spezifische λ -Rekombination in Säugetierzellen	54
D3	Sequenz-spezifische Rekombination in Säugetiergenomen durch Int Mutanten	55
D3.1	Genomische Inversion in HeLa Reporter-Zelllinien	55
D3.2	Sequenz-spezifische Deletion eines Neomycin-Resistenzmarkers durch die Integrase in embryonalen Stammzellen der Maus	58
D4	Die Transfektion von IHF-DNA Komplexen stimuliert die Rekombinationsaktivität der wild-typ Integrase in Säugetierzellen	63
D4.1	Die Ausprägung der IHF Untereinheiten in HeLa Zellen stimuliert nicht die Aktivität der wild-typ Integrase	64
D5	Optimierung des prokaryotischen Hilfsproteins IHF für den Einsatz in eukaryotischen Zellen	65
D5.1	oIHF-Ausprägung in <i>E.coli</i>	66
D5.2	Analyse von enzymatischen Aktivitäten der vier unterschiedlichen oIHF Varianten in <i>E.coli</i>	68
D5.3	Stimulierung der integrativen Rekombinationsaktivität von Int-h/218 durch Ausprägung von oIHF2 in ROSA-IR	70
D6	Identifizierung einer <i>attB</i>-ähnlichen Sequenz (<i>attH2</i>) im menschlichen Genom	73
D6.1	Sequenz-spezifische Rekombination von <i>attH2</i> und <i>attP</i> ^{*2} in <i>E.coli</i>	74
D6.2	Episomale Rekombination von <i>attH2</i> und <i>attP</i> ^{*2} in HeLa durch die Mutante Int-h und Int-h/218.....	75

D6.3	Untersuchung der MCT5 mRNA-Ausprägung in einer menschlichen Nieren- und Gebärmutterhalszelllinie	76
E	DISKUSSION.....	78
E1	Effiziente sequenz-spezifische DNA Rekombination in Säugetierzellen durch <i>Lambda</i> Integrase Mutanten	78
E2	Die Rekombinationsaktivität der wild-typ Integrase in Säugetierzellen.....	82
E3	Die Reaktivität von episomalen DNA-Segmenten in HeLa Zellen während der <i>Lamda</i> Rekombination.....	84
E4	Hinweise für die Bildung eines synaptischen Komplexes in der von Integrase Mutanten katalysierten Rekombination	85
E5	Genetische Manipulation von embryonalen Stammzellen der Maus durch Int-h/218	87
E6	Optimierung des prokaryotischen <i>Integration Host Factors</i> (IHF) zum Einsatz in eukaryotischen Zellen.....	89
E7	Perspektive: Der Einsatz des <i>Lamda</i> Systems zur gezielten Integration von Fremd-DNA ins Genom von Säugetieren.....	92
F	LITERATURVERZEICHNIS.....	96
G	ANHANG	106
H	ZUSAMMENFASSUNG.....	112
I	SUMMARY	113

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>
Bp	DNA-Basenpaar(e)
bpA	<i>bovine growth hormone</i> (Rinder-Wachstumshormon) Polyadenylierungs-Sequenz
BSA	<i>Bovine Serum Albumine</i> ; Rinderserum-Albumin
bzw.	beziehungsweise
C	Cytosin
cDNA	<i>complementary DNA</i>
CMV	Cytomegalovirus (Promoter)
DMSO	Dimethylsulfoxid
DNA	<i>deoxyribonucleic acid</i> ; Desoxyribonucleinsäure
dNTP	Desoxynucleotide
DTT	Dithiothreitol
(d)YT	<i>(double) yeast/tryptone</i> ; (doppelt) Hefe/Trypton
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraessigsäure
EF	Embryonale Fibroblastenzellen
EGFP	<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
FIS	<i>factor for inversion stimulation</i>
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
IHF	<i>Integration Host Factor</i>
Int	Integrase
IPTG	Isopropylthiogalactosid
KDa	Kilo-Dalton
Kb	Kilobasenpaare
LIF	<i>leukemia inhibitory factor</i>

Abkürzungsverzeichnis

MCT	<i>monocarboxylate transporter</i> ; Monokarboxylat-Transporter
MES	<i>mouse embryonic stem cells</i> ; embryonale Stammzellen der Maus
mRNA	<i>messenger RNA</i> , Boten-RNA
NA	Nucleotidaustausche
Neo	Neomycin-Resistenz-Gen
NLS	Nukleare Lokalisationssequenz
OD	optische Dichte
oIHf	<i>one-chain Integration Host Factor</i>
ORF	<i>open reading frame</i> ; offenes Leseraster
PAGE	Polyacrylamid Gel Elektrophorese
PBS	<i>Phosphat Buffered Saline</i> ; Phosphat-gepufferte Salzlösung
PCR	<i>polymerase chain reaction</i> ; Polymerasekettenreaktion
PGK	Phosphoglycerat-Kinase (Promoter)
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

A Einleitung

Die Erforschung der Funktion eines Gens und die Korrektur genetischer Defekte in intakten Organismen setzt die Entwicklung von Werkzeugen voraus, die eine präzise Veränderung von genomischer DNA in höheren Eukaryoten ermöglichen. Dabei gilt es vor allem, neue Strategien zu entwickeln, die eine effiziente und kontrollierbare Manipulation *in vivo* erlauben. Für die Korrektur pathogener genetischer Defekte im menschlichen Genom ist vor allem eine dauerhafte Übertragung genetischen Materials von besonderem Interesse (Kay et al., 1997). Deshalb steht neben der Frage des Gentransfers, d.h. dem Einbringen genetischer Information in eine Zelle ohne diese zu schädigen, auch das Problem der sicheren und effizienten Integration neuer genetischer Information in das menschliche Genom im Vordergrund (Morgan und Anderson, 1993; Wolff, 1994).

A1 Sequenz-spezifische Rekombinasen: Ein wertvolles Werkzeug zur gezielten DNA-Manipulation

Die bedeutendste Beschränkung der gegenwärtigen Methoden des Gentransfers in Säuger-Zellen liegt darin, dass die fremde DNA nach der Aufnahme in die Zelle häufig verloren geht. In den seltenen Fällen, in denen sie stabil in das Wirtszellgenom integriert wird, geschieht dies durch heterologe Rekombination an zufälligen Positionen im Genom oder eines Chromosomenabschnitts. Die Effizienz des stabilen Gentransfers lässt sich durch virale Vektoren steigern. Doch erfolgt die Integration auch in diesem Fall eher zufällig, was insbesondere bei der Gentherapie problematisch ist. Die zufällige Integration viraler Vektoren, die häufig auch mit chromosomalen Deletionen und Rearrangements einhergeht (Miller et al., 2002), kann im ungünstigsten Fall zur Aktivierung von Proto-Onkogenen oder zur Inaktivierung von Tumorsuppressorgenen führen (Verma und Somia, 1997).

Eine Möglichkeit fremde DNA gezielt ins Genom von Säuger-Zellen einzubringen ist durch homologe Rekombination. Ein Beispiel hierfür ist das *Gene Targeting* in embryonalen Stammzellen von Mäusen. Bei dieser Technik wird die

homologe Integration jedoch von häufigen unkontrollierten heterologen Integrationsereignissen überlagert (Marrow und Kucherlapati, 1993). In somatischen Zellen ist das Verhältnis von homologer zu nicht-hologer Rekombination sogar noch ungünstiger als in embryonalen Stammzellen. Obwohl zusätzliche Selektionsstrategien zur Anreicherung der homologen Integrierten existieren, ist die genetische Manipulation durch homologe Rekombination in Säuger-Zellen ineffizient und weitestgehend auf embryonale Stammzellen beschränkt (Vasquez et al., 2001). Um hier korrigierend eingreifen zu können, müsste man zunächst den Mechanismus der homologen Rekombination in höheren Eukaryoten besser verstehen. Dies ist zurzeit jedoch aufgrund von dessen Komplexität nicht absehbar.

Als Alternative zur homologen Rekombination rücken bei der Manipulation eukaryotischer Genome zunehmend sequenz-spezifische Rekombinationssysteme in den Vordergrund (Kilby et al., 1993). Sie besitzen gegenüber der homologen Rekombination oder der viralen Integration den Vorteil, dass der Rekombinationsprozess sehr präzise ist und bei der stabilen Neuverknüpfung

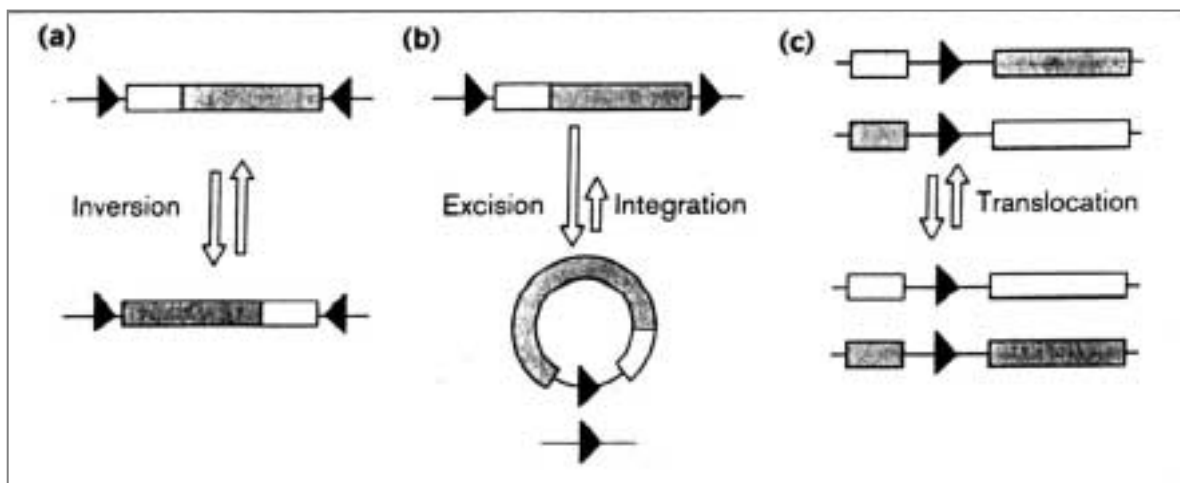


Abbildung A.1: Schematische Darstellung von sequenz-spezifischen Rekombinationsmöglichkeiten zwischen unterschiedlich orientierten Rekombinationsregionen. Die zur Integrase-Familie gehörenden Rekombinasen Cre, FLP oder Int führen die sequenz-spezifische Rekombination zwischen zwei *loxP*-, *FRT*- oder *att*-Regionen, in dieser Reihenfolge, durch (schwarze Dreiecke). **(a)** Die intramolekulare Rekombination zwischen entgegengesetzt orientierten Rekombinationsregionen führt zur Inversion des dazwischenliegenden DNA-Fragmentes, während in **(b)** die Rekombination zwischen gleich orientierten Sequenzen 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 weißen Pfeil symbolisiert. Die Translokation, ebenfalls eine intermolekulare Rekombinationsreaktion, ist in **(c)** dargestellt. Zur Kennzeichnung bestimmter DNA-Abschnitte sind diese grau unterlegt. (modifiziert nach Metzger & Feil, 1999)

zweier Nucleotidsequenzen keine genetische Information verloren geht. Da die sequenz-spezifische Rekombination in der Regel sehr effizient ist -bis zu 90% der Partner werden *in vitro* tatsächlich rekombiniert-, könnte die unkontrollierte Integration genetischen Materials vermindert werden. Ein weiterer Vorteil der sequenz-spezifischen Rekombinationssysteme liegt in deren grundsätzlicher Regulierbarkeit. Werden sequenz-spezifische Rekombinasen gezielt ausgeprägt, können durch Rekombinationsprozesse Gene an- oder abgeschaltet und deren Funktion in einem bestimmten Zelltyp oder in einem bestimmten Entwicklungsstadium untersucht werden (Kilby *et al.*, 1993; Gu *et al.*, 1994). Die vielseitigen Einsatzmöglichkeiten vieler sequenz-spezifischer Rekombinasen kommt auch dadurch zustande, dass sie durch entsprechende Positionierung ihrer Zielsequenzen grundsätzlich zur Insertion, Deletion, Inversion oder zum Versetzen von genetischen Elementen eingesetzt werden können (Abb. A.1; Metzger & Feil, 1999). Das macht sie zu einem vielversprechenden Werkzeug für die Manipulation von genomischer DNA in Eukaryoten und hat in den letzten Jahren vor allem die Erprobung von sequenz-spezifischen Rekombinasen wie die Cre-Rekombinase des Bakteriophagen P1 und die *FLP*-Rekombinase aus Hefe vorangetrieben (Kilby *et al.*, 1993).

A2 Sequenz-spezifische Rekombinasen in eukaryotischen Zellen

A2.1 Das Cre/*loxP*-System

Zweifellos am weitesten verbreitet und entwickelt ist die Anwendung des sequenz-spezifischen Rekombinationssystems des Phagen P1, das sogenannte Cre/*loxP*-System (Nagy, 2000; Christ & Dröge, 2001). Das Cre/*loxP*-System besteht aus der Cre-Rekombinase und zwei 34 Basenpaaren langen *loxP*-Sequenzen. Es funktioniert nicht nur effizient in Bakterien, sondern auch in eukaryotischen Zellen (Sauer, 1993). Die Cre-Rekombinase kann zum Beispiel DNA aus einem Chromosom ausschneiden, wenn dieses von den aus dem Bakteriophagen P1 stammenden *loxP*-Sequenzen eingerahmt ist, und die Rekombinase in eine Zelle eingebracht wird. Diese Form der Manipulation chromosomaler DNA findet bereits Anwendung bei der gezielten Entfernung genetischer Elemente aus einem Chromosom oder bei der Kontrolle der

Genexpression in Eukaryoten (Sauer und Henderson, 1989). Die wohl bedeutendste Anwendung des Cre/loxP-Systems ist dessen Einsatz in Mäusen. Häufig spielen die zu untersuchenden Gene auch in der embryonalen Entwicklung eine wichtige Rolle, wodurch das konstitutive Ausschalten eines Gens in vielen Fällen letal ist. Durch Cre-vermittelte Deletion bestimmter Gene oder Genabschnitte ist es jedoch gelungen, sogenannte konditionale *knock-out* Mäuse zu entwickeln (Gu et al., 1993), die auch als Tiermodell für verschiedene Erkrankungen dienen können. Die gezielte Induktion von Cre in den Mäusen erlaubt es, die Funktion eines Gens ausschließlich in bestimmten Entwicklungsstadien oder in bestimmten Zelltypen zu untersuchen (Orban et al., 1992; Kühn et al., 1995; Sauer, 1998; Metzger & Feil, 1999). Darüber hinaus kann das Cre/loxP-System die Integration eines loxP-tragenden Vektors in eine in das Genom integrierte loxP Sequenz katalysieren (Baubonis und Sauer, 1993). Allerdings muss diese Erkennungssequenz zuvor durch andere Methoden (z.B. homologe Rekombination) als stabile Sequenz in das Genom eingebracht werden (Baubonis und Sauer, 1993). Von großem Interesse ist es daher Techniken mit Cre zu entwickeln, die es erlauben therapeutische Gene in eine im Genom „natürlich“ vorkommende, eine sogenannte loxP-analoge Sequenz, zu integrieren. Wünschenswert wäre hier die gezielte Integration in einem einzelnen Schritt in einen gut charakterisierten Locus. Calos und Kollegen konnten in der Tat loxP-analoge Sequenzen im Maus und im menschlichen Genom identifizieren, die sich in episomalen Rekombinationstests als rekombinationsfähig erwiesen (Thyagarajan et al., 2000). Da die Sequenzen zum Teil stark von der wild-typ loxP-Sequenz abweichen, scheint die Cre-Rekombinase eine hohe Toleranz gegenüber Abweichungen von der Konsensus-Sequenz zu besitzen. Die Integration von Fremd-DNA in solche genomische loxP-analoge Sequenzen durch die Cre-Rekombinase wurde allerdings bisher nur in Hefe gezeigt (Sauer, 1996).

A2.2 Das FLP/FRT-System

Ein weiteres Beispiel für eine erfolgreiche Anwendung sequenz-spezifischer Rekombinationssysteme bei der Manipulation eukaryotischer Genome ist das FLP/FRT-System der Hefe *Saccharomyces cerevisiae*. Auch die FLP-Rekombinase kann, wie zuvor für das Cre/loxP-System beschrieben, zwei künstlich ins Genom von Drosophila oder Säugetierzellen integrierte Rekombinationssequenzen (*FRT*)

effizient rekombinieren und dadurch beispielsweise *FRT*-flankierte Sequenz-Abschnitte ausschneiden. Gleichfalls konnte gezeigt werden, dass die Rückreaktion, also die Integration in einen genomischen Locus, in den zuvor eine *FRT*-Sequenz integriert wurde, durch intermolekulare Rekombination katalysiert werden kann (Golic und Lindquist, 1989; O’Gorman *et al.*, 1991). Sequenz-spezifische Rekombination durch die FLP-Rekombinase konnte nicht nur in Zelllinien, sondern auch in transgenen Mäusen nachgewiesen werden (Dymecki, 1996).

A2.3 Limitierte Anwendungsmöglichkeiten etablierter Rekombinationssysteme

Der Anwendung der etablierten sequenz-spezifischen Rekombinationssysteme Cre und FLP in höheren Eukaryoten sind aufgrund ihrer enzymatischen Eigenschaften Grenzen gesetzt. Die FLP-Rekombinase arbeitet bei einer Temperatur von 30°C optimal. Die Rekombinationsaktivität bei 37°C, der gebräuchlichen Kultivierungstemperatur von Säugerzellen, ist hingegen gering. Um dieses Problem zu lösen, entwickelten Buchholz *et al.* (1996) die thermostabile Mutante FLPe, für die im Vergleich zu wild-typ FLP in Maus ES-Zellen jedoch nur eine 3- bis 5-fach höhere Rekombinationsaktivität bei 37°C nachgewiesen werden konnte (Buchholz *et al.*, 1998).

Für die Cre-Rekombinase wurde in Zell-Kultur eine potentiell toxische Wirkung festgestellt, die zu einer verminderten Teilungsrate der Zellen führen kann. Der zytotoxische Effekt, dem Chromosomen-Abberationen vorausgehen, steigt mit der enzymatischen Aktivität von Cre. Die Chromosomen-Abberationen werden allerdings nicht durch sequenz-spezifische Rekombination verursacht, sondern ihnen liegt eine Endonuklease-Aktivität von Cre zugrunde, der Cre-unabhängige illegitime Rekombinationsereignisse folgen (Loonstra *et al.*, 2001, Silver & Livingston, 2001).

Ein gemeinsames Merkmal der Cre- und FLP-katalysierten sequenz-spezifischen Rekombinationsreaktion ist die Identität ihrer Produkte und Substrate. Das ermöglicht Cre und FLP sowohl die Hin- als auch die Rückreaktion zu katalysieren, wodurch sich insgesamt ein Reaktionsgleichgewicht zwischen den Substraten und Produkten einstellt. Diese, als bidirektional bezeichnete Rekombination (Abb. A.2-A), hat zur Folge, dass Cre und FLP nicht nur die

Integration von DNA-Elementen katalysieren, sondern auch ihre Exzision (Abb. A.1-B). Die Anwendung dieser Rekombinationssysteme zur Integration von DNA-Elementen birgt daher den Nachteil, dass in Gegenwart der Rekombinasen einmal integriertes genetisches Material nicht stabil ist. Es kann durch die Rekombinasen wieder ausgeschnitten werden. Um dies zu vermeiden, sind oftmals Strategien erforderlich, die ein nur temporäres Vorhandensein der Rekombinase sicherstellen. Beispiele hierfür sind die transiente Expression der Rekombinase von einem nicht-replizierenden Vektor (O’Gorman et al., 1991), die Tetracyclin-regulierte Ausprägung der Rekombinase durch ein heterologes Expressionssystem (Gossen et al., 1995) oder der Einsatz von Fusionsrekombinasen, deren katalytische Aktivität abhängig von der Gegenwart eines spezifischen Liganden ist (Logie & Stewart, 1995; Metzger et al., 1995). Die niedrige Effizienz der Integrationsreaktionen hat bislang jedoch den routinemäßigen Einsatz dieser Methoden zur gezielten Integration von Fremd-DNA verhindert.

Die Verwendung von asymmetrischen Rekombinationssequenzen, wie es für das Cre/loxP-System von Araki *et al.* (1997) gezeigt wurde, ermöglicht es, das Reaktionsgleichgewicht zur Hinreaktion zu verschieben. Der Nachteil dieser Methode ist jedoch, dass später keine weiteren Rekombinationsschritte mehr stattfinden können.

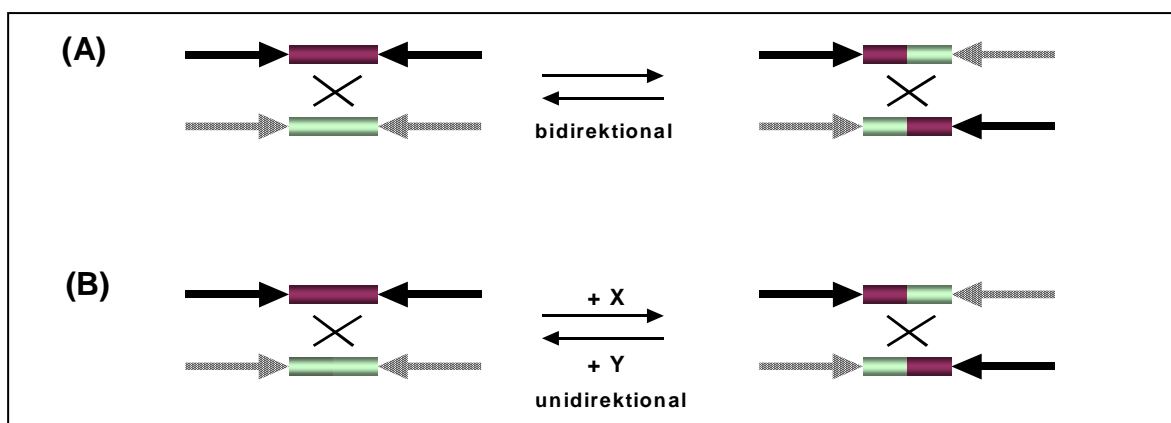


Abbildung A.2: Schematische Darstellung von bi- und unidirektionaler sequenz-spezifischer Rekombination. Dargestellt sind Rekombinationsregionen, die aus zwei entgegengesetzt orientierten Sequenzwiederholungen (Pfeile) bestehen, die wiederum durch eine *overlap*-Sequenz (Rechteck) getrennt werden. In (A) ist die bidirektionale Rekombination gezeigt, die ein Equilibrium zwischen Hin- und Rückreaktionen darstellt. (B) Als unidirektional bezeichnet man die Rekombinationsreaktion, bei der die Hinreaktion in Gegenwart eines Faktors X stattfindet und die Rückreaktion nur in Gegenwart eines Faktors Y katalysiert wird.

Um die Mängel der etablierten sequenz-spezifischen Rekombinasen Cre und FLP zu umgehen, müsste man in der Natur existierende sequenz-spezifische Rekombinationssysteme, bei denen sich Hin- und Rückreaktion durch Kofaktoren steuern lassen (Abb. A.2-B), für den Einsatz in höheren Eukaryoten optimieren. Ein Kandidat für eine solche regulierte Rekombinase ist die vom Phagen *Lambda* kodierte Integrase.

A3 Das sequenz-spezifische Rekombinationssystem des Bakteriophagen *Lambda*

Der dem Rekombinationssystem des Bakteriophagen *Lambda* zugrunde liegende Mechanismus zur Neuverknüpfung von DNA-Strängen wird allgemein als konservative sequenz-spezifische Rekombination bezeichnet. Der Mechanismus zeichnet sich dadurch aus, dass während des Rekombinationsverlaufs zwei DNA-Stränge an genau definierten Stellen geschnitten und die ausgetauschten Stränge

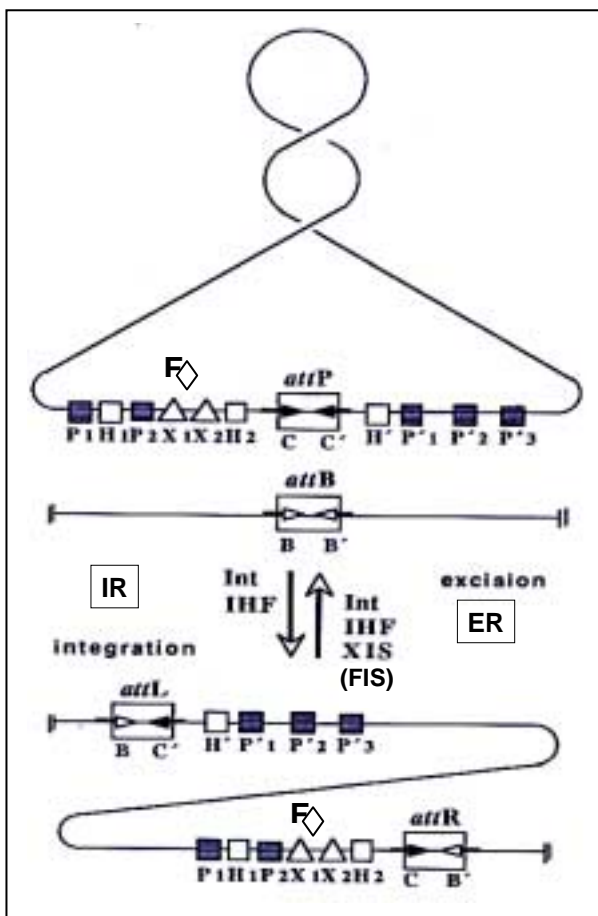


Abbildung A.3: Schematische Darstellung des integrativen (↓) und exzisen (↑) Reaktionswegs der λ -Rekombination. Dargestellt ist eine negativ superhelikal gespannte *attP*-Sequenz, die in Gegenwart von IHF und Int in das lineare *attB*-Element integriert. Dabei entstehen die Produkte *attL* und *attR*, 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- und P'-Arm bezeichnet, die weitere Bindungsstellen der Int (■ P) und Bindungsstellen für IHF (□ H), Xis (Δ X) und Fis (◇ F) besitzen. Int und IHF sind für beide Reaktionswege und XIS ist für die Exzision unerlässlich. FIS stimuliert die Exzision. (Christ & Dröge, 2001)

unverändert, d.h. ohne Abbau oder Neusynthese von DNA, wieder verknüpft werden. Zusätzliche Faktoren, die Energie bereitstellen, sind bei dieser Neuverknüpfung nicht erforderlich (Übersichtsartikel: Landy, 1989).

Ein weiteres Merkmal des *Lambda* Rekombinationssystems ist dessen Fähigkeit, sowohl Neuverknüpfungen von DNA-Strängen auf demselben als auch zwischen unterschiedlichen DNA-Fragmenten zu katalysieren. Befinden sich die Rekombinationssequenzen auf demselben Molekül, bestimmt deren relative Orientierung zueinander, ob der dazwischenliegende DNA-Abschnitt ausgeschnitten oder invertiert wird. Befinden sich die Rekombinationsregionen dagegen auf unterschiedlichen Molekülen, so ist auch die Integration oder Translokation von DNA-Fragmenten durch λ -Rekombination möglich (Abb. A.1).

Somit birgt die sequenz-spezifische Integrase (Int) des Phagen *Lambda*, vergleichbar mit den ihr verwandten Rekombinasen Cre und FLP (Nunes-Düby et al., 1998), vielseitige Anwendungsmöglichkeiten für gezielte DNA-Manipulationen. Im Gegensatz zu dem Cre- oder FLP-System setzt sich das *Lambda* sequenz-spezifische Rekombinationssystem jedoch aus zwei unterschiedlichen Rekombinationsreaktionen zusammen, die als **integrative** und **exzisiv** **Rekombination** bezeichnet werden (Abb. A.3; Landy, 1989). Die Produkte der integrativen Rekombination stellen dabei die Substrate der exzisiv Rekombination dar und umgekehrt. Die Komplexität des λ -Systems resultiert vor allem daraus, dass die eigentliche DNA-Rekombinase, die Phagen-kodierte Integrase (Int), unterschiedliche Hilfsfaktoren zur Durchführung der integrativen oder exzisiv Rekombination benötigt. Die An- oder Abwesenheit dieser Faktoren ermöglichen oder verhindern den Aufbau von funktionellen Nukleoproteinstrukturen an den strukturell unterschiedlichen Rekombinationssequenzen der beiden Rekombinationswege. Dadurch kann die Rekombination gezielt nur in eine Richtung ablaufen und ist somit im Gegensatz zum Cre- oder FLP-System unidirektional (Abb. A.2-B). Diese Eigenschaft macht sich der Bakteriophage *Lambda* zu Nutze, um sich präzise in eine bestimmte Stelle des *E.coli* Genoms zu integrieren (die sogenannte integrative Rekombination). Als Bestandteil des Bakterienchromosoms vermehrt er sich als sogenannter Prophage in einem Zustand, der als Lysogenie bezeichnet wird, bei fortlaufender Bakterienteilung. Dabei schädigt er den Wirt nicht und kann sich unter physiologisch günstigen

Bedingungen passiv von *E.coli* verbreiten lassen. Gerät das Wirts-Bakterium in eine Krise, beispielsweise durch Schädigung der genomischen DNA durch UV- oder ionisierende Strahlung, so kann sich der Prophage durch das Einleiten der exzisen Rekombination wieder aus dem *E.coli* Genom ausschneiden. Extrachromosomal vermehrt sich der Phage wieder aktiv in der Bakterienzelle. Die Freisetzung des Bakteriophagen und seine Verbreitung auf andere Bakterien erfolgt im letzten Schritt durch die Lyse der Bakterienzelle (Landy, 1989).

A3.1 Die integrative λ -Rekombination

Die integrative Rekombinationsreaktion des Bakteriophagen *Lambda* bedarf zweier unterschiedlicher Rekombinationssequenzen, der *bacterial attachment site* (*attB*) im *E.coli* Genom und der *phage attachment site* (*attP*) im Phagengenom. *AttB* ist von allen *att*-Regionen die Kürzeste und umfasst nur 21 Nucleotide. Ein Bestandteil von *attB* ist das sogenannte 15bp *common core*, dass allen *att*-Regionen gemeinsam ist. Das *common core* wird im *attB* als BOB` bezeichnet und besteht aus einer 7 Nucleotide langen sogenannten *overlap*-Region, die von zwei gegenläufigen Int-Bindungssequenzen flankiert ist (Abb. A.4). Die *overlap*-Sequenz ist in allen *att*-Regionen identisch. Entscheidend ist hierbei die Identität dieses Bereiches, und nicht unbedingt die Sequenz selbst, da in diesem Abschnitt der

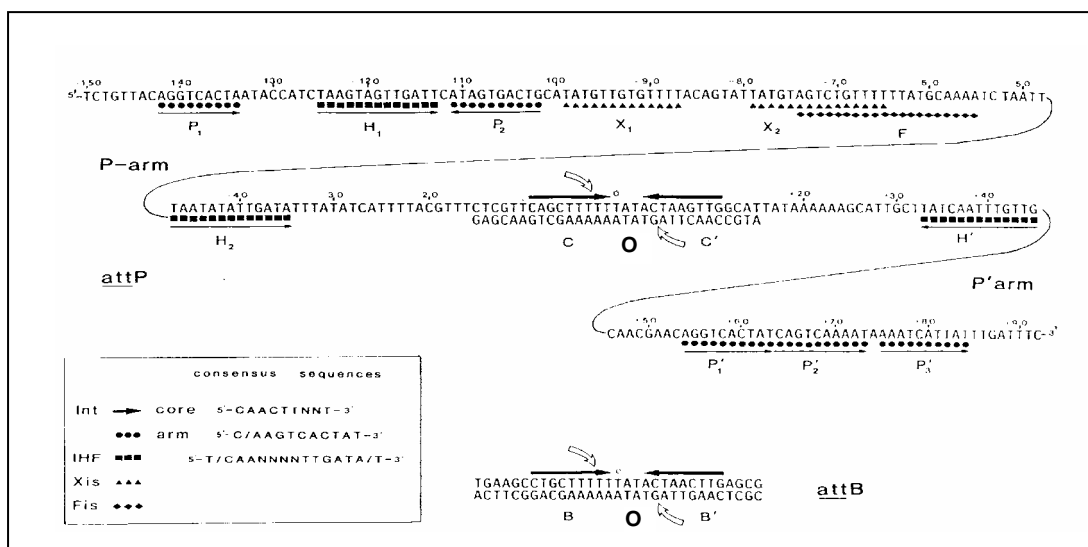


Abbildung A.4: DNA-Sequenz und Proteinbindungsstellen der attP- und attB-Regionen. Die Erkennungssequenzen der jeweiligen Proteine sind durch die im Kasten angezeigten Symbole markiert. Die relativen Orientierungen der Bindungsstellen sind durch Pfeile (→) angezeigt. Die Positionen der Einzelstrangaustausche sind durch gekrümmte Pfeile (↷) gekennzeichnet. (modifiziert nach Landy, 1989).

DNA-Sequenz der Strangaustausch zwischen den Reaktionspartnern durch die Integrase stattfindet (Bauer et al., 1985; Nunes-Düby et al., 1997). Die *overlap*-Sequenz, an dessen 5'- und 3'-Ende die Integrase nukleophil angreift, ist asymmetrisch aufgebaut und bestimmt dadurch die Richtung des koordinierten DNA-Strangaustausches an den *att*-Regionen.

attP ist 243 Nucleotide lang und besteht aus dem *common core* COC' (in Abbildung A.5 auch: POP'), das 5' von dem sogenannten P-Arm und 3' vom P'-Arm flankiert wird (Abb. A.4). Neben den Arm-Regionen, den Bindungsstellen für die Arm-Bindungsdomäne der Integrase (Abschn. A3.3), befinden sich auch die Bindungsstellen für die Hilfsproteine auf diesen Abschnitten. Auf dem P-Arm befinden sich 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 (F). Im Gegensatz dazu trägt der P'-Arm drei Integrase- (P'_1 , P'_2 , P'_3) und eine IHF-Bindungsstelle (H') (Abb. A.4).

Die Phagen-kodierte Integrase benötigt zur Durchführung der integrativen Rekombination ein Bakterien-kodiertes Hilfsprotein, den *Integration Host Factor* (IHF) aus *E.coli* (Abschn. A3.4). Während Int den DNA-Strangaustausch katalysiert, ist die Aufgabe von IHF als Strukturkomponente den Aufbau eines funktionellen, rekombinogenen Nukleoproteinkomplexes zwischen Int und *attP* zu unterstützen. Diesen Komplex bestehend aus IHF, Int und *attP*, nennt man auch Intasom (Echols, 1989). Er kann nur mit Hilfe von negativ superhelikal gespanntem *attP* gebildet werden (Richet et al., 1986). Die Rekombination zwischen *attB* und *attP* kann –sehr vereinfacht– nach heutigem Wissensstand folgendermaßen beschrieben werden: Int und IHF bilden zunächst mit dem ca. 240 Nucleotiden langen superhelikal gespannten *attP* das komplexe und sehr stabile Intasom. Dieses Intasom kann mit dem 21 Nucleotide langen *attB*, welches zwei Bindungsstellen für Int besitzt, interagieren und eine Neuverknüpfung der Polynucleotidstränge zwischen *attP* und *attB* katalysieren. Die Rekombination wird offensichtlich durch ein zufälliges Zusammentreffen des Intasoms mit der proteinfreien *attB* Sequenz im synaptischen Komplex eingeleitet. (Richet et al., 1988) Es erfolgt also kein *Scannen* der DNA durch die Integrase oder durch das Intasom mittels Diffusion entlang des DNA-Strangs. Dies ist wichtig für den Einsatz des λ -Rekombinationssystems in Eukaryoten, da hier die nukleosomal verpackte DNA ein Scannen aus sterischen Gründen verhindern könnte. Die eigentliche

Rekombination erfolgt durch einen koordinierten Strangaustausch (Abb. A.5; Kitts & Nash, 1988; Burgin & Nash, 1992) über ein sogenanntes *Holliday*-Struktur-Intermediat zu den Produkten *attL* und *attR*. Dazu kommt es zunächst nach Ausbildung des synaptischen Komplexes zu einem Einzelstrangaustausch der *top*-Stränge am 5'-Ende der gepaarten *overlap*-Regionen zweier Partner (Abb. A.5-A). Es wird eine Holliday-Struktur (Holliday, 1964) ausgebildet, in der es nach einem aktuellen Modell von Nunes-Düby et al., 1997 durch das sogenannte „*strand swapping*“ zur Kreuzung der Stränge in der Mitte des *overlap*-Bereiches kommt (Abb. A.5-B). Die *top*-Stränge werden anschließend religiert. Durch Isomerisierung der Holliday-Struktur wird das Schneiden der *bottom*-Stränge eingeleitet (Abb. A.5-C). Und im letzten Schritt wird durch den Austausch der *bottom*-Stränge am 3'-Ende der gepaarten *overlap*-Regionen das Holliday-Intermediat aufgelöst und es entstehen die Produkte *attL* und *attR* (Abb. A.5-D).

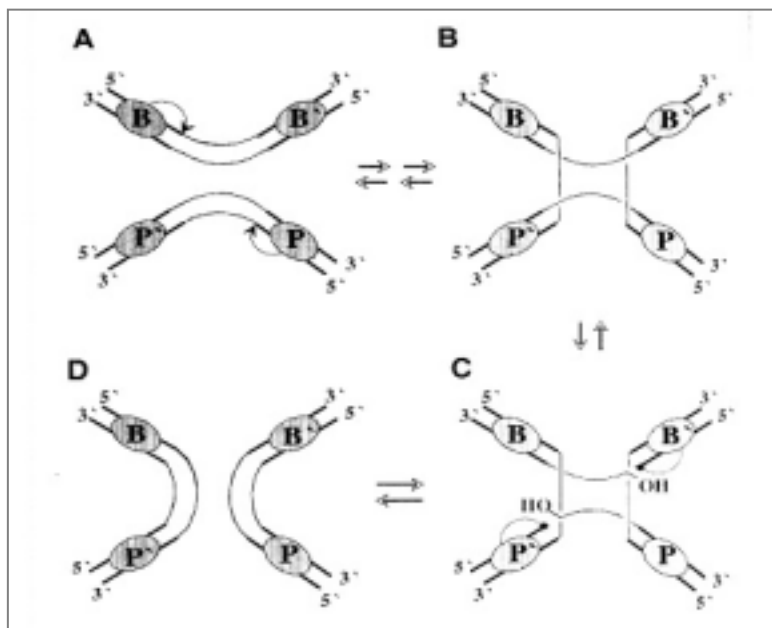


Abbildung A.5 Schematische Darstellung des konventionellen Strangaustauschs bei der integrativen λ -Rekombination. (A) Antiparallele Anordnung von *attB* (BOB') und *attP* (POP'). Nukleo-philer Angriff auf die *top*-Stränge durch Integrasemonomere (ovale Symbole), die an die *core*-Sequenzen des B- (B) und P-Arms (P) der *att*-Regionen gebunden sind. (B) Holliday-Struktur-Intermediat nach dem reziproken Strangaustausch zwischen den beiden *top*-Strängen. (C) Schneiden der *bottom*-Stränge durch Integrasemonomere, die an die *core*-Sequenzen des B'- (B')

und P'-Arms (P') der *att*-Regionen binden führt zur Auflösung des Holliday-Intermediats (D) Nach der Ligation der reziprok ausgetauschten *bottom*-Stränge entstehen die Produkte *attL* (BOP') und *attR* ($B'OP$), die ihrerseits als Substrate für die exzisive Rekombination dienen können. (Christ & Dröge, 1999)

A3.2 Die exzisive λ -Rekombination

Die exzisive Rekombinationsreaktion des Bakteriophagen *Lambda* verwendet die zwei unterschiedlichen Rekombinationssequenzen *attL* und *attR* als Substrate, die, wie zuvor beschrieben, aus der Neuverknüpfung von *attB* und *attP*

hervorgehen. Im Gegensatz zur Rekombination von *attB* und *attP* in der integrativen Rekombinationreaktion, bildet sowohl *attL* als auch *attR* zusammen mit Int und IHF Nukleoprotein-Komplexe von unterschiedlicher Struktur aus (Kim & Landy, 1992). Dabei muss die DNA nicht superhelikal gespannt sein (Abremski & Gottesmann, 1979). *attR* benötigt zur Ausbildung eines funktionellen, rekombinogenen Komplexes neben Int und IHF zusätzlich den phagen-kodierten Faktor XIS (Bushman et al., 1984). Die exzusive Rekombination kann außerdem noch durch den wirtseigenen *factor for inversion stimulation* (FIS) begünstigt werden (Landy, 1989).

Der synaptische Komplex entsteht wie bei der integrativen Rekombination durch zufälliges Zusammentreffen der beiden Reaktionspartner. Die eigentliche Rekombination erfolgt anschließend, wie zuvor für die integrative Rekombination beschrieben, durch zwei koordinierte Einzelstrangausstausche.

A3.3 Die *Lambda* Integrase

Die Phagen-kodierte Integrase (Int) ist ein Mitglied der sogenannten Integrase-Familie (Nunes-Düby et al., 1998). Sie ist ein basisches Protein mit einem Molekulargewicht von ca. 40 KDa, das sich aus drei funktionellen Domänen zusammensetzt (Tirumalai et al., 1997): der katalytischen Region und zwei unterschiedlichen DNA-Bindungsdomänen mit unterschiedlicher Bindungsspezifität (Moitoso de Vargas et al., 1988) (Abb. A.6).

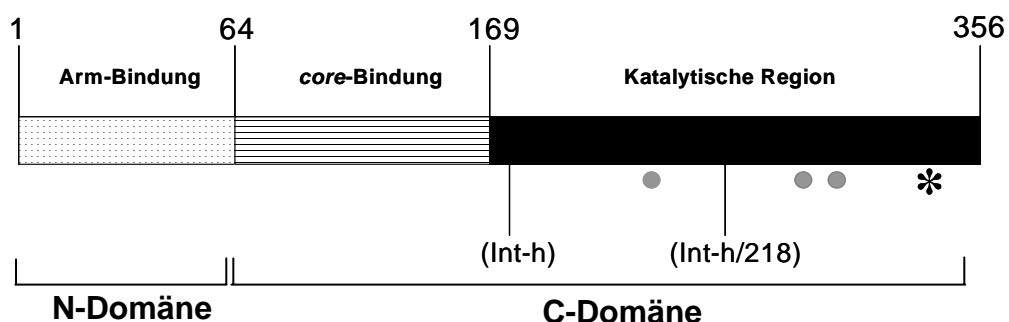


Abbildung A.6: Schematische Darstellung der Lambda Integrase. Die drei in DNA Arm-Bindung, *core*-Bindung und Katalyse involvierten Domänen (Tirumalai et al., 1997) sind durch unterschiedliche Musterung hervorgehoben. Die Zahlen beziehen sich auf die Aminosäurereste. Die grauen Punkte markieren die Position 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 der *core*-Bindungsdomäne (Aminosäurereste 65-169) erkennt die Int spezifische Sequenzen in den *core*-Regionen (Abschn. A3.1), wogegen sie mit der Arm-Bindungsdomäne (Aminosäurereste 1-64) spezifische Sequenzen in den Arm-Regionen (Abschn. A3.1 und A3.2) von *attP*, *attL* und *attR* bindet. Die gleichzeitige Bindung der Integrase an *core*- und Arm-Sequenzen, die zum Aufbau der Nukleoprotein-Komplexen notwendig ist, findet nur mit Unterstützung der Hilfsproteine statt (Abschn. A3.4). Neben einer strukturellen Rolle besitzt die Arm-Bindungsdomäne (im Folgenden als N-Domäne bezeichnet, siehe Abb. A.6) auch regulatorische Funktionen innerhalb der Integrase. Das vollständige Integrase Protein besitzt nicht die Fähigkeit, an isolierte *core*-Regionen zu binden. Spaltet man jedoch die N-Domäne durch Proteolyse ab, kann die C-Domäne (Abb. A.6) *in vitro* nicht nur an die *core*-Regionen binden, sondern auch die DNA an dieser Stelle schneiden und religieren (Landy, 1989). Ob sie diese auch rekombinieren kann, ist bisher nicht bekannt. Ist die N-Domäne mit der C-Domäne verknüpft, dann kann dieser Effekt auch durch sequenz-spezifische Bindung von Oligonucleotiden an die Arm-Bindungsdomäne erzeugt werden (Dibyendu et al., 2001). Dies lässt darauf schließen, dass die DNA-Bindung an der N-Domäne zu einer Konformationsänderung im Integrase-Protein führt, so dass die Integrase in Gegenwart der Hilfsproteine nicht nur die *core*-Sequenzen der *att*-Regionen binden kann, sondern auch das Schneiden der Einzelstränge im synaptischen Komplex einleiten kann. Neben der Bindung an die *core*-Regionen ist die *core*-Bindungsdomäne zusammen mit der Arm-Bindungsdomäne wahrscheinlich auch an der Protein-Protein Interaktion des Dimer-*Interface* beteiligt (Tirumalai et al., 1997; Jessop et al., 2000).

Die katalytische Domäne wird von den Aminosäureresten 170-356 gebildet (Abb. A.6). Dieser innerhalb der Integrase-Familie hochkonservierte C-terminale Abschnitt wird auch als „minimale katalytische Domäne“ (Intc170, Tirumalai et al., 1997) bezeichnet. Dieser Abschnitt besitzt auch noch vom Rest des Proteins isoliert die Eigenschaft, DNA schneiden und ligieren zu können. Die katalytische Aktivität der Domäne basiert auf einer Aminosäure-Tetrade (Nunes-Düby et al., 1998), die bei allen Integrasen konserviert ist. Bei der λ -Integrase sind es die Aminosäurereste Arg 212, His 308, Arg 311, die das katalytische Tyrosin 342 (Abb. A.6; Esposito & Scocca, 1997) beim nukleophilen Angriff auf die DNA unterstützen

und kovalente Protein-DNA Intermediate bilden. Die Energie wird dabei in Form einer 3'-Phosphotyrosin Verknüpfung gespeichert und macht dadurch die Rekombination unabhängig von energiereichen Kofaktoren. Die Auflösung der Protein-DNA Intermediate erfolgt durch die nukleophilen freien 5'-OH DNA-Enden und führt zur rekombinanten Religation der DNA.

Int-h ist eine klassische Integrase Mutante (Miller et al., 1980; Kikuchi et al., 1985), die biochemisch sehr gut charakterisiert ist. Sie unterscheidet sich von der wild-typ Integrase in nur einem Aminosäurerest (Bear et al., 1987). Anstelle von Glutamat befindet sich an der Position 174 von Int-h ein Lysin (E174K, Abb. A.6). Eine genauere Charakterisierung von Int-h *in vitro* ergab, dass in Abwesenheit von IHF im Vergleich zur der um 500-fach reduzierten Aktivität der wild-typ Integrase (Landy, 1989), die Rekombinationsaktivität von Int-h nur um ein 10-faches reduziert ist (Lange-Gustafson & Nash, 1984). Außerdem konnte *in vitro* gezeigt werden, dass die Int-h-vermittelte Rekombination nicht mehr von der Topologie des Substrats abhängig ist (Lange-Gustafson & Nash, 1984). Dadurch ist die integrative Rekombination auch mit einer *attP*-Sequenz möglich, die nicht superhelikal gespannt ist. Int-h zeichnet sich dadurch aus, dass die Affinität zu den *core*-Regionen höher ist als bei der wild-typ Integrase (Patsey & Bruist, 1995). Dies beruht vermutlich auf der positiv geladenen ϵ -Aminogruppe des Lysins, die eine ionische Bindung an das negativ-geladene Phosphatrückgrat ausbilden kann. Die höhere Affinität von Int-h zu den *core*-Regionen führt dazu, dass in Gegenwart von IHF die Integration in *att*-analoge Sequenzen in *E.coli* und *in vitro* häufiger vorkommen kann (Miller, 1980; Patsey und Bruist, 1995). Als *att*-analoge Sequenzen werden Nucleotidsequenzen bezeichnet, die mehr oder weniger Ähnlichkeit mit der wild-typ *attB*-Sequenz besitzen und von der *Lambda* Integrase erkannt und rekombiniert werden können. Sequenzabweichungen von *attB* haben jedoch zur Folge, dass diese Sequenzen von der wild-typ Integrase sehr ineffizient rekombiniert werden. Vor allem durch Sequenzunterschiede im *overlap*-Bereich zwischen zwei Reaktionspartnern wird die Isomerisierung der Holliday-Struktur stark gehemmt (Kitts und Nash, 1987; Nunes-Düby et al., 1997). Int-h dagegen zeigt in Gegenwart von IHF eine größere Toleranz gegenüber diesen Sequenzunterschieden (Patsey und Bruist, 1995, Christ & Dröge, 1999).

Int-h/218 ist ein Derivat von Int-h mit einem zusätzlichen Aminosäurerest-Austausch an Position 218 in der katalytischen Domäne (Christ & Dröge, 1999). Das Glutamat in Int-h ist durch ein Lysin ersetzt worden (E218K, Abb. A.6). Die Mutation E218K wurde von Wu *et al.* (1997) in einer Revertante gefunden, die einen Aminosäureaustausch in der *core*-Bindungsdomäne kompensiert, der zu einem Verlust von spezifischer Int-Bindung an die *core*-Sequenzen führt. Wu *et al.* (1997) vermuten, dass dieser Aminosäureaustausch, wahrscheinlich durch unspezifische Kontakte zwischen dem positiven Lysinrest und der negativ geladenen Phosphatdiesterkette der DNA, zu einer stärkeren Bindung der Revertante an die *core*-Sequenzen führt. Die Mutation E218K wurde in Int-h eingeführt, um die Affinität für die *core*-Sequenzen zu steigern und die Rekombination mit *att*-analogen Sequenzen, die sich auch im menschlichen Genom befinden (Lorbach 2000; diese Arbeit), zu erleichtern. Durch den E218K-Austausch wurde auch die Rekombinationsaktivität von Int-h in Abwesenheit von Hilfsproteinen sowohl in der integrativen sowie der exzisivem Rekombination in *E.coli* gesteigert (Christ & Dröge, 1999).

Abschließend ist zu den Mutanten Int-h und Int-h/218 zu bemerken, dass sie zwar verglichen zur wild-typ Integrase eine höhere enzymatische Aktivität in der Abwesenheit von Kofaktoren besitzen. Die Aktivität der Mutanten wird jedoch genau wie bei der wild-typ Integrase durch die Gegenwart der Hilfsproteine stimuliert und sie katalysieren in Gegenwart von IHF die integrative, jedoch nicht die exzisive Rekombination (Christ & Dröge, 1999; Lange-Gustafson & Nash, 1984). Sind dagegen XIS und IHF in *E.coli* vorhanden, so katalysiert Int-h, wie die wild-typ Integrase, effizient die exzisive Rekombination (Kikuchi *et al.*, 1985). Obwohl nicht gezeigt, ist das gleiche für Int-h/218 anzunehmen.

A3.4 Die Hilfsproteine der *Lambda* Integrase: IHF, XIS und FIS

Der ***Integration Host Factor (IHF)*** wurde als zellulärer Faktor von *E.coli* identifiziert, der sowohl für die λ sequenz-spezifische Integration als auch für die Exzision unerlässlich ist (Miller *et al.*, 1980). Während die exzisive Rekombination durch höhere Konzentrationen von IHF gehemmt wird, ist dies bei der integrativen Rekombination nicht der Fall. (Bushman *et al.*, 1994). IHF ist auch an einer Vielzahl von zellulären Prozessen wie beispielsweise der Replikation, der

transkriptionellen Regulation und an verschiedenen anderen sequenz-spezifischen Rekombinationssystemen beteiligt (Übersichtsartikel: Friedman, 1988; Freundlich et al., 1992). In erster Linie scheint IHF eine Rolle als Strukturprotein in Nukleoprotein-Komplexen zu spielen, indem es durch Bindung an spezifische Sequenzen (Craig & Nash, 1984) eine starke Krümmung in der DNA verursacht (Rice et al., 1996). Diese starke Krümmung der DNA hat zur Folge, dass sich kurze DNA-Schleifen bilden, durch die an die DNA gebundene Proteine in räumliche Nähe gebracht werden um die Nukleoprotein-Komplexe zu stabilisieren.

IHF ist ein Dimer, das sich aus den beiden unterschiedlichen Polypeptiden IHF- α , mit einem Molekulargewicht von ca. 11,2 KDa, und IHF- β von ca. 10,6 KDa (Nash & Robertson, 1981) zusammensetzt. Die Kristall-Struktur eines IHF-DNA Komplexes (Rice et al., 1996) zeigt, dass die beiden Untereinheiten ineinander umschlungen sind und einen kompakten Proteinkern bilden (Abb. D.13-B). Aus diesem Proteinkern erstrecken sich zwei lange β -Schleifen in die kleine Furche der doppelsträngigen DNA, die im 180° Winkel um das IHF Protein gewickelt ist. Die beiden Proteinuntereinheiten von IHF ähneln in ihren Sequenzen dem DNA biegenden Protein HU von *E.coli* und anderen Bakterien (Oberto & Rouviere-Yaniv, 1996). IHF bindet jedoch im Gegensatz zu HU an spezifische DNA-Sequenzen und verursacht zusätzlich eine stärkere Krümmung der DNA. Die beiden Untereinheiten von IHF werden von den Genen *himA* (Miller et al., 1980; IHF- α) und *himD* (Kikuchi et al., 1985; IHF- β) kodiert. Die Überproduktion von *himA* in *E.coli* führt zur Bildung von unstabilem IHF- α , wogegen die Überproduktion von *himD* dazu führt, dass sich unlösliche IHF- β -Komplexe bilden (Nash et al., 1987). Sind beide IHF-Untereinheiten in stöchiometrischen Mengen vorhanden, dann bildet sie ein funktionelles IHF Dimer und IHF- α hält so das Polypeptid IHF- β in Lösung und umgekehrt bewahrt IHF- β das IHF- α Polypeptid vor dem Abbau in *E.coli*.

XIS ist ein basisches Phagen-kodiertes Protein von ca. 9 KDa, das neben IHF in der exzisen Rekombination der Lambda Integrase eine wichtige Rolle spielt. Es besitzt keinerlei katalytische Aktivität, bindet jedoch kooperativ an zwei benachbarte spezifische Sequenzen im P-Arm (A.4, Bushman et al., 1984; Yin et al., 1985). Sind zwei XIS Proteine an *attR* gebunden, dann wird die DNA so gekrümmt, dass sich der spezifische Nukleoprotein-Komplex an *attR* ausbilden

kann. Bindet XIS dagegen an *attP*, wird der Aufbau des zur integrativen Rekombination notwendigen Intasoms (Abschn. A3.1) gehemmt (Moitoso de Vargas & Landy, 1991).

FIS (*factor for inversion stimulation*) ist ein von *E.coli* kodiertes Protein, das die Int-vermittelte exzisive Rekombination stimulieren kann, wenn die XIS Menge limitiert ist (Thompson et al., 1987). Ist XIS in ausreichender Menge vorhanden, dann hat FIS keinen Einfluss auf die exzisive Rekombination. FIS ist wie IHF ein kleines basisches DNA-bindendes Protein, das die DNA um ca. 90° krümmt. Es ist jedoch im Unterschied zu IHF ein Homodimer mit einem Molekulargewicht von 25 KDa, das vor allem in den sequenz-spezifischen Inversionssystemen Hin, Gin und Cin eine wichtige Rolle spielt (Plasterk & Van de Putte, 1984).

A4 Das Ziel dieser Studie

Das Ziel des hier vorgestellten Projekts ist die Entwicklung eines Protokolls für den Einsatz des sequenz-spezifischen Rekombinationssystems des Bakteriophagen *Lambda* zur Manipulation genomischer DNA in höheren Eukaryoten. Hierbei steht die stabile Integration von fremdem genetischem Material in einen definierten Locus einer Zielzelle im Mittelpunkt. Das Rekombinationssystem des Phagen *Lambda* bietet gute Voraussetzungen dafür, da es in der Natur darauf spezialisiert ist, den Phagen stabil in das Genom von *E.coli* zu integrieren. Ohne einen zusätzlichen vom Phagen kodierten Faktor ist die Reaktion irreversibel und das integrierte genetische Material bleibt somit auch in Gegenwart der Integrase erhalten. Die Integration von Fremd-DNA mit der *Lambda* Integrase in eine natürlich vorkommende Sequenz im menschlichen Genom wäre zur Entwicklung von gentherapeutischen Techniken wünschenswert. Daher soll im Rahmen dieser Arbeit auch nach *attB*-analogen Sequenzen im menschlichen Genom gesucht werden und diese auf ihre Rekombinationsfähigkeit durch die Integrase getestet werden.

Neben dem Einsatz zur stabilen Integration, könnte das *Lambda*-System zur Entfernung von DNA-Sequenzen aus einem Chromosom oder zur Kontrolle der Genexpression verwendet werden. Für den Aufbau von komplexen genetischen Systemen, bei denen das Entfernen verschiedener DNA-Segmente erforderlich wird, ist eine Alternative zu den Rekombinasen Cre und FLP von Vorteil. Durch

unterschiedliche Positionierung der Erkennungssequenzen lassen sich verschiedene Rekombinationsreaktionen mit dem *Lambda*-System durchführen. In dieser Arbeit wird neben der Integrase-vermittelten sequenz-spezifischen Inversion von DNA-Fragmenten in einer humanen Zelllinie auch die sequenz-spezifische Deletion eines Neomycin-Resistenzmarkers durch die Integrase in embryonalen Stammzellen der Maus untersucht.

Der Einsatz der Integrase Mutanten Int-h und Int-h/218 soll in Eukaryoten die sequenz-spezifische *Lambda* Rekombination auch in der Abwesenheit von natürlich vorkommenden Kofaktoren ermöglichen. Im Gegensatz zu Cre und FLP lässt sich die Richtung der Rekombinationsreaktion sowohl von der wild-typ Int als auch von den Mutanten Int-h und Int-h/218 durch spezifische Kofaktoren steuern. Kofaktoren wie IHF sollen im Rahmen dieser Arbeit für die Anwendung in Säugerzellen optimiert werden. Die Entwicklung eines Protokolls für die Anwendung der *Lambda* Rekombinase in Säugetierzellen soll zudem von der Untersuchung ihres Reaktionsmechanismus in eukaryotischen Zellen unterstützt werden. Im Mittelpunkt dieser Untersuchungen steht die Frage, welchen Einfluss die Arm-Bindungssequenzen in den *att*-Regionen auf die Rekombination der verschiedenen Int Varianten in der Abwesenheit der natürlichen Kofaktoren besitzen.

B Material

B1 Bakterienstämme

Für Klonierungs- und Plasmidpräparationen

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)

Für *in vivo* Rekombinationsanalysen in *E.coli*

CSH 26ΔIHF F⁻ araΔ (lac pro) thi himA Δ82::Tn10(Tc^R) himD Δ3::Cm^R
 (B. Rak, Freiburg)

CSH 26 F⁻ araΔ (lac pro) thi (Miller, 1972)

B2 Chemikalien

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

Radioaktive Chemikalien wurden von Amersham Pharmacia geliefert.

B3 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).

B4 Synthetische Oligonucleotide

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

Primer	Sequenz	Bemerkung
MCT II	5'-CGCGGATCCGGCCACAAGTATTCCTT-3'	RT-PCR und Klonierung des Substrats pACH2
MCT 3'	5'-TGCAGGACGTCCAGAACGGGCAATAGCTGTAGA-3'	
MCT5'	5'-CGCGGATCCCTTGGCTCTTACAATGCTCACTT-3'	
Int-N-EU	5'-AACTGCAGCTCGAGGTCCACCATGGGAAGAAGG-3'	Amplifikation des Int-Gens
Int-C-EU	5'-GCTCTAGAGCGGCCGCTCATTATTTGATT-3'	
Int-C/NLS-EU	5'-GCTCTAGATCATTAGACCTTACGCTTCTTCT-3'	Amplifikation des Int-Gens mit NLS am C-Terminus
NLS1-C	5'-GCTCTAGATTATCATACCTTTCTTCTTCTTAGGGC-CGCCTCCGGATTTGATTTCAATTTTGTCCCACT-3'	
attB53	5'-CGATTCAGCTAGCTGAAGCCTGCTTTTTTATACTAAC-TTGAGCGAACGCAATCC-3'	Klonierung von Deletions- und Target Vektoren
attB35	5'-AATTGAATTGCGTTCGCTCAAGTTAGTATAAAAAAGC-AGGCTTCAGCTAGCTGAAT-3'	
attP01	5'-GTCACTATCAGTCAAAATACAATCA-3'	
attP02	5'-TGATTGTATTTTGACTGATAGTGAC-3'	
PFP-Nsil	5'-CCAATGCATCCTCTGTTACAGGTCACTAATAC-3'	
P'RP-EcoRV-NotI	5'-ATAAGAATGCGGCCGCAGATATCAGGGAGTGGGAC-AAAATTGAA-3'	
Trc5-N	5'-CGGAATTCATAATTCGTGTCGCTCAAGGCG-3'	
Trc3-N	5'-CGGAATTCGCAAAAAGGCCATCCGTCAG-3'	
himA5-EU	5'-AACTGCAGCCACCATGGCGCTTACAAAAGCTG-3'	Klonierung von pPGKSShimA
himA3-EU	5'-TAGTCTAGATTATCACTCGTCTTTGGCGAAGCG-3'	
himD5-EU	5'-AACTGCAGCCACCATGACCAAGTCAGAATTGATAGA-3'	Klonierung von pPGKSShimD
himD3-EU	5'-TAGTCTAGATTATCAACCGTAAATATTGGCGCGATC-3'	
attL-5P	5'-AAAAGTGCAGTAGAAGCCTGCTTTTTTATACTAAG-3'	Klonierung von pCMVSSattL
attL-3X	5'-TGCTCTAGACCGGGAGGGAGTGGGACAA-3'	
p10	5'-TCCACAGCTCGCGGTTGAG-3'	Rekombinationsanalysen in ES-Zellen
p9	5'-GCTCTAGATCATTACTTGTACAGCTCGTCCATGC-3'	
IHF3-1	5'-GTAAGGCCGCTTCCACCCTGCGCAAGAGTTCGAGGCC-ATATGCT-3'	Klonierung von pTrc/olHF1
IHF5-2	5'-GCAGGGTGAAGCGGCCTTACAAAAGCTGAAATGTCA-GAATAT-3'	
IHF3-2	5'-ACGCTCGCCACCAGCGTTTTTCGACCCGGCTTTTT-AAC-3'	
IHF5-3	5'-CGAAAACGCTGGTGGCGAGCGTATTGAAATCCGCG-3'	
IHF5-1	5'-AACTGCAGCCACCATGGGCACCAAGTCAGAATTGATA-GAAAGACT-3'	Klonierung von pTrc/olHF1 und pTrc/olHF2
IHF3-3	5'-CTAGTCTAGATTATCAACCTGAGATATTGGCGCG-3'	
IHF3-3*	5'-GTAAGACCGCCGCTTCCACCCTGCGCAAGAGTCGAG-GCCATATGCT-3'	Klonierung von pTrc/olHF2
IHF5-4	5'-GGGTGAAGCGGCGGTCTTACAAAAGCTGAAATGTC-AGAATAT-3'	
IHF3-4	5'-ACGCTCGCCCCACCAGCGTTTTTCGACCCGGCTTTTT-AAC-3'	
IHF5-5	5'-AACGCTGGTGGGGGCGAGCGTATTGAAATCCGCG-3'	
IHFHIS5	5'-CATGCCATGGGGGCTAGCACCAAGTCAGAATTGATAG-AAAGACT-3'	
IHFHIS3	5'-GCTCTAGAGAATTCTTATCAGTGGTGGTGGTGGTGG-TGGCCTGATCCACCGTAGATATTGGCGCGATCG-3'	
IHFnl3	5'-GCTCTAGATTATCATACCTTTCTTCTTCTTAGGGCC-TGATCCACCGTAGATATTGGCGCGATCG-3'	Klonierung von olHF2NLS

Primer	Sequenz	Bemerkung
PE	5'-GGGGATCCTCTGTTACAGGTCCTAATAC-3'	PCR-Analysen von Rekombinationsprodukten
p5	5'-TACCATGGTGATGCGGTTTTG-3'	
PGK-RP	5'-TCTGAGCCCAGAAAGCGAAG-3'	
p3	5'-AATCCCAAGCTTCCACCATGGTGAGCAAGGG-3'	
p4	5'-AACCTCTACAAATGTGGTATGG-3'	
PI	5'-GGCAGCACTGCATAATTCTCT-3'	Klonierung von attP ²
PFP-BanI	5'-TACGGTGCCTCTGTTACAGGTCCTAATAC-3'	
attx06	5'-CAACTTAGTATAAGAAAGCTGAAC-3'	
attx05	5'-GTTTCAGCTTTCTTATACTAAGTTG-3'	
P'RP-PstI	5'-AACTGCAGCAGGGAGTGGGACAAAATTGA-3'	

B5 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* und/oder Bakterienstämmen mit Hilfe von Resistenzgenen wurden Antibiotika in den folgenden Konzentrationen zugegeben:

- Plasmidkodierte-Resistenzgene

Ampicillin 200 µg/ml, Chloramphenicol 50 µg/ml, Kanamycin 30 µg/ml, Spectinomycin 50 µg/ml, Tetracyclin 10 µg/ml

- Chromosomale Resistenzgene

Chloramphenicol 10 µg/ml, Kanamycin 30 µg/ml, Tetracyclin 5µg/ml

B6 Plasmide

Folgende kommerziell erhältliche, von genannten Personen zu Verfügung gestellten oder in dieser Arbeit generierten Plasmide wurden eingesetzt:

Expressionsvektoren

p(CMV o. PGK)	negativ Kontrolle in eukaryotischen Rekombinationsanalysen (diese Arbeit)
p(CMV o. PGK)SSInt	eukaryotische Ausprägung der wild-typ Integrase (diese Arbeit)
p(CMV o. PGK)SSInt-h	eukaryotische Ausprägung der Mutante Int-h (diese Arbeit)
p(CMV o. PGK)SSInt-h/218	eukaryotische Ausprägung der Mutante Int-h/218 (diese Arbeit)
pPGKSSInt-hNLS	eukaryotische Ausprägung der Mutante Int-h mit C-terminaler NLS (diese Arbeit)
pPGKSShimA	eukaryotische Ausprägung von HimA (diese Arbeit)

pPGKSShimD	eukaryotische Ausprägung von HimD (diese Arbeit)
pCMVSSoIHF2	eukaryotische Ausprägung oIHF2 (diese Arbeit)
pCMVSSoIHF2NLS	eukaryotische Ausprägung oIHF2NLS (diese Arbeit)
pCMVSSoIHF2HIS	eukaryotische Ausprägung oIHF2HIS (diese Arbeit)
pTrcInt-h	prokaryotische Ausprägung von Int-h (Christ & Dröge, 1999)
pTrc/oIHF1(2; 3 oder 4)	prokaryotische Ausprägung der oIHF Varianten 1-4 (diese Arbeit)
pTrc/oIHF2HIS	prokaryotische Ausprägung von oIHF2HIS (diese Arbeit)
pTrc99a	negative Kontrolle in prokaryotischen Rekombinationsanalysen (Pharmacia Biotech, Freiburg)

Substratvektoren

pGFPattB/attP	Inversionssubstrat zur Bestimmung der <i>Lambda</i> Rekombination zwischen <i>attB/attP</i> (Lorbach et al., 2000)
pGFPattL/attR	Inversionssubstrat zur Bestimmung der <i>Lambda</i> Rekombination zwischen <i>attL/attR</i> (Lorbach et al., 2000)
pλIR	Deletionssubstrat zur Bestimmung der intramolekularen <i>Lambda</i> Rekombination zwischen <i>attB/attP</i> (diese Arbeit)
pλER	Deletionssubstrat zur Bestimmung der intramolekularen <i>Lambda</i> Rekombination zwischen <i>attL/attR</i> (diese Arbeit)
pCMVSSattPmut	<i>attP</i> -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter (Christ et al., 2002)
pCMVSSattL	<i>attL</i> -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter (diese Arbeit)
pCMVSSattB	<i>attB</i> -tragendes intermolekulares Rekombinationssubstrat mit CMV Promoter (Christ et al., 2002)
pWSattBGFP	<i>attB</i> -tragendes promoterloses intermolekulares Rekombinationssubstrat (diese Arbeit)
pWSattPGFP	<i>attP</i> -tragendes promoterloses intermolekulares Rekombinationssubstrat (Christ et al., 2002)
pWSattRGFP	<i>attR</i> -tragendes promoterloses intermolekulares Rekombinationssubstrat (diese Arbeit)
pNC9WT	Derivat von pNC9 (Christ & Dröge, 1999) auf dem sich eine Int-Expressionskassette befindet (diese Arbeit)
pACH2	Inversionssubstrat mit <i>attH2</i> und <i>attP*2</i> (diese Arbeit)

Target Vektoren

pROSA26-1	ROSA26 <i>Target</i> Vektor (Soriano, 1999)
pSAβgeo	3'-Spleißstelle (Friedrich & Soriano, 1991)
pROSA-IR	ROSA26 <i>attP/attB</i> -Reporter <i>Target</i> Vektor (diese Arbeit)

pROSA-ER	ROSA26 <i>attL/attR</i> -Reporter <i>Target</i> Vektor (diese Arbeit)
pA-R4	5' außerhalb der <i>Target</i> Vektor-Sequenz gelegene ROSA26 Sonde (Mao et al., 1999)

Sonstige Plasmide

pGFP-C1	Erfassung der Transfektionseffizienzen (Clontech)
pPGKCrepA	Klonierung von Expressionsvektoren (Fellenberg, 1998)
pGEM [®] 4Z	Klonierung von Deletionssubstraten und <i>Target</i> Vektoren (Promega)
pBS302	Transkriptions-Stoppkassette (Gibco/BRL)
pACYC177	Klonierungsvektor (Rose, 1988)
pTKHyg	Ausgangsvektor für die promoterlosen intermolekularen Substratvektoren (Clontech)

B7 Medien und Puffer für die eukaryotische Zellkultur

PBS: 137mM NaCl, 3mM KCl, 8mMNa₂HPO₄ (2H₂O), 1,5mM KH₂PO₄

Einfrriermedium: 10% DMSO, 90% FCS

Kulturmedien:

für HeLa und NIH/3T3: Dulbecco's Modified Eagle Medium (D-MEM) (GibcoBRL) modifiziert durch Zusatz von: 10% FCS (PAA), Penicillin (100 U/ml), Streptomycin (100 µg)

für IB10 embryonale Maus-Stammzellen (ES-Medium): D-MEM 4,5 g Glucose/l modifiziert durch Zusatz von: 15% FCS (PAN), 1mM Natriumpyruvat (GibcoBRL), 2 mM L-Glutamin (GibcoBRL), 1% nicht-essentielle Aminosäuren (GibcoBRL), 0,1% LIF-haltiger Überstand, 0,1 mM β-Mercaptoethanol (Merck)

für embryonale Maus-Fibroblasten (EF-Medium): D-MEM mit Glutamax (GibcoBRL) modifiziert durch Zusatz von: 10% FCS (PAA), 1mM Natriumpyruvat (GibcoBRL), 1% nicht-essentielle Aminosäuren (GibcoBRL)

Selektionsmedum:

-Kulturmedium für ES-Zellen versetzt mit 200µg/ml G418 (GibcoBRL)

-Kulturmedium für HeLa-Zellen versetzt mit 500µg/ml G418 (GibcoBRL)

Elektroporationsmedium:

RPMI 1640 ohne Phenolrot, ohne Glutamin (GibcoBRL)

B8 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
- NIH/3T3 *Swiss mouse embryo fibroblasts* (Andersson et al., 1979)
- IB10 ein Subklon der embryonalen Maus-Stammzelllinie E14.1 (A. Berns, Niederlande)
- CHO *Chinese Hamster Ovary cells* (Puck et al., 1958)

C Methoden

C1 Molekularbiologische Methoden

Wenn nicht anders angegeben, wurden in dieser Arbeit molekularbiologische Methoden gemäß Sambrook et al. (1989) oder wie vom jeweiligen Hersteller empfohlen durchgeführt.

C1.1 Agarosegelelektrophorese

Die Auftrennung negativ geladener DNA-Moleküle erfolgte durch die elektrophoretische Auftrennung in einem Agarosegel mit TBE (90mM Tris-Borat, pH 8,3; 2,5mM 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 302nm sichtbar gemacht. Fotografien der gefärbten Agarosegele entstanden mit Hilfe des *Image Master[®] Systems* (Pharmacia, Freiburg).

Zur Abschätzung der Größe von aufgetrennten linearen DNA-Fragmenten wurde die 1Kb DNA-Leiter der Firma GibcoBRL oder der 100 Basenpaar-Marker der Firma Pharmacia verwendet.

C1.2 DNA-Reinigung aus Agarosegelen

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

C1.3 Plasmid-Isolierung aus Bakterien

Die Isolierung von Plasmid-DNA aus Bakterien im kleinen Maßstab erfolgte aus 2ml Bakterien-Kulturen durch Affinitätschromatographie mit Hilfe des *QIAprep[®] Spin Miniprep Kit* der Firma Qiagen (Hilden).

Die Aufreinigung von Plasmid-DNA zur Transfektion von eukaryotischen Zellen erfolgte aus 100ml Bakterien-Kulturen mit dem *EndoFree[®] Plasmid Maxi Kit* der Firma Qiagen (Hilden).

C1.4 DNA-Isolierung aus eukaryotischen Zellen

Die Isolierung von DNA aus HeLa Zellen erfolgte mit Hilfe des *QIAmp[®] DNA Blood Mini Kit* der Firma Qiagen (Hilden). Zur Gewinnung von genomischer ES-Zell-DNA wurden diese auf Gelatin-behandelten Petri-Schalen bis zur Konfluenz kultiviert und dann über Nacht bei 56°C in Lysis Puffer (10mM NaCl; 10mM Tris-HCl pH7,5; 10mM EDTA; 0,5% Sacrosyl; 0,5mg/ml Proteinase K) lysiert. Die DNA wurde durch Phenol/Chloroform-Extraktion von Proteinen getrennt und mit Ethanol gefällt. Die Resuspension der DNA erfolgte in TE/10µg/ml RNase A.

C1.5 Polymerasekettenreaktion (PCR)

Die Polymerasekettenreaktion (Saiki et al., 1988) wurde zum Nachweis von sequenz-spezifischer Rekombination durch die *Lambda* Integrase in eukaryotischen Zellen, zum Einführen von gezielten Mutationen in Vektoren und zur Konstruktion von Substrat- und Expressionsvektoren eingesetzt.

Alle PCR Reaktionen wurden mit Hilfe des *Taq PCR-Master Mix Kit* der Firma Qiagen durchgeführt. Eine **Standard-PCR** Reaktion enthielt in der Regel 50pmol Primer, 50pmol reverser Primer, 1x *Master Mix* (Taq-Polymerase, dNTPs, Reaktionspuffer) und zwischen 100ng bis 1µg *Template* (Matrize). 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; 45s bei 94°C, 45s bei 58°C und 1min/1Kb *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-*Primer* spezifische Nucleotidaustausche in das zu amplifizierende Produkt eingeführt. Hierzu wurde zunächst der DNA-Bereich, in dem die gewünschten Nucleotide ausgetauscht werden sollten, in zwei separaten Standard-PCR Reaktionen amplifiziert (Abb. C.1-a). Die Region in der die Nucleotidaustausche eingeführt werden sollten, überlappten sich in beiden PCR-Produkten. In einer abschließenden PCR-Reaktion wurden die zuvor generierten PCR-Fragmente als Matrize eingesetzt (Abb. C.1-b) und das gesamte zu

amplifizierende Produkt einschließlich der spezifisch eingeführten Nucleotide vervielfältigt (Abb. C.1-c).

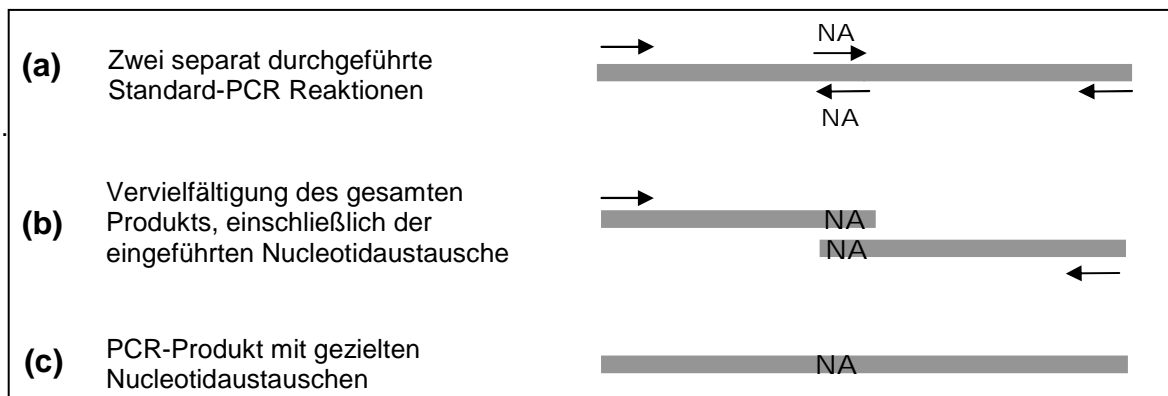


Abbildung C.1: Schematische Darstellung der PCR-Mutagenese. Der als Matrize dienende DNA-Bereich ist als graues Rechteck dargestellt. Pfeile (\rightarrow) stellen die synthetisch hergestellten Oligonucleotid-*Primer* dar; die mit NA markierten *Primer* und DNA-Bereiche enthalten die gewünschten Nucleotidaustausche. (Details sind im Text angegeben.)

C1.6 RT-PCR Analyse von MCT5 mRNA-Ausprägung in HeLa und 293 Zellen

Zur Analyse der Transkription von MCT5 in humanen HeLa, 293 und 293T Zellen wurde zunächst deren RNA mit Hilfe des *Rneasy-Mini-Kits* der Firma Qiagen isoliert. Die zu analysierenden Zellen wurden zu diesem Zweck zuvor mit einem *Qiasredder* (Qiagen) homogenisiert. Die RNA wurde anschließend mit DNaseI behandelt, um eventuell vorhandene DNA-Reste zu entfernen. In der Regel wurden 5µg RNA mit 15U DNaseI 1h bei 37°C inkubiert und anschließend durch eine Phenol/Chloroform Extraktion und eine Ethanolfällung gereinigt. Die mRNA wurde anschließend mit Hilfe des *First-Strand cDNA Synthese Kits* der Firma Pharmacia reverse-transkribiert. In der Regel setzte sich ein cDNA-Synthese-Ansatz wie folgt zusammen:

4µg RNA bzw. 2,5µg DNaseI behandelte RNA gelöst in DEPC-H₂O in einem Gesamtvolumen von 8µl; 5µl Reaktionsmix (reverse Transkriptase und dNTPs), 1µl NotI-d(T)₁₈-Primer (1:25 verdünnt) und 1µl DTT (200mM Stock). Die cDNA-Synthese erfolgte innerhalb 1h bei 37°C.

Der Nachweis der reversen MCT5-Transkripte erfolgte in einer Standard-PCR mit den Primern (MCT5') und (MCT3') (siehe B4 und Anhang) und den zuvor synthetisierten komplementären DNA-Stränge als Matrize. Dazu wurde der cDNA-Ansatz mit Rnase-freiem H₂O auf 40µl gebracht und ¼ des Ansatzes als *Template* eingesetzt. Zur Kontrolle ob in den mit DNase I behandelten RNA-Präparationen

noch Reste von DNA vorhanden sind, wurde parallel zur cDNA-Synthese mit diesen eine Standard-PCR mit den zuvor genannten Primern durchgeführt. Zur Analyse der PCR-Produkte wurden diese abschließend in einem Agarosegel elektrophoretisch aufgetrennt

C1.7 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 Didesoxynucleotiden beruht. Markierte DNA wurde mit einem automatischen DNA-Sequencer (ABIPrism™377, Applied Biosystems) aufgetrennt und analysiert. Der DNA-Sequencer wurde von Karin Otto bedient.

C1.8 Southern Blot Analyse

Genomische *Southern Blot* Analyse (Sambrock et al., 1989) wurde zur Identifizierung homologer ES-Zell-Rekombinanten, Charakterisierung von HeLa Reporter-Zelllinien und zum Nachweis von Integrase-vermittelter genomischer Rekombination in HeLa und ES-Zellen eingesetzt.

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

Der Nachweis des Int-spezifischen Rekombinationssubstrats auf dem modifizierten ROSA26 Allel und der Int-vermittelten genomischen Rekombination, erfolgte mit einem EcoRI-Verdau mit einer externen Probe. Die externe Probe ist

ein 600bp-PacI/Spel-Fragment aus dem Vektor A-R4 (Mao et al., 1999). Illegitime Integration des *Target* Vektors pROSA-IR oder pROSA-ER wurde mit dem 1Kb-SacII/XbaI-Fragment aus dem Vektor pROSA26-1 (Soriano, 1999) als interne Probe in Kombination mit einem EcoRI-Verdau ausgeschlossen.

Der Nachweis von Int-vermittelter Rekombination in stabilen HeLa Reporter-Zelllinien erfolgte mit einer internen Probe, die das gesamte EGFP-Gen umfasste und durch PCR generiert wurde. Die genomische DNA wurde mit NcoI verdaut. Die Proben wurden durch *Random-primed Labelling* mit Hilfe der thermostabilen Bca-Polymerase (TaKaRa) bei 56°C radioaktiv markiert. Nicht-eingebaute Radionucleotide wurden mit S-200HR Säulen (Pharmacia) abgetrennt und die Proben 5min bei 95°C denaturiert, bevor sie zur Hybridisierungslösung gegeben wurden. Die Hybridisierung erfolgte in allen Experimenten bei 65°C übernacht. Nach der Hybridisierung wurde die Membran in allen Experimenten zweimal in 2xSSC/0,1% SDS bei 65°C gewaschen. Je nach gewünschter Stringenz wurden weitere Waschrinne mit abfallender Salzkonzentration durchgeführt. Die Membran wurde in Plastikfolie eingeschweißt und die Signale durch Autoradiographie (Biomax-MS Filme, Kodak) sichtbar gemacht oder mit dem Phosphorimager-System (Fuji Bas 1000, Raytest, Straubenhardt; Software: TINA 2.09, Raytest, Straubenhardt) quantifiziert.

C1.9 Konstruktion der Expressionsvektoren

C1.9.1 Eukaryotische Expressionsvektoren

Zur Ausprägung der Integrasen, der Kofaktoren und des Neomycin-Resistenzgens in Eukaryoten wurden Vektoren eingesetzt, auf denen die entsprechenden Gene entweder vom Phosphoglycerat-Kinase I Promoter (PGK; Adra et al., 1987) oder dem Cytomegalovirus-Promoter (CMV; Boshart et al., 1985) transkribiert wurden. Die eukaryotischen Expressionsvektoren sind 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.C.2) trägt. Zur Steigerung der Transkription wurde das pPGKCrebpa zunächst so modifiziert, dass zwischen die Schnittstellen NheI und PstI ein heterologes Intron (Choi et al., 1991) eingesetzt wurde. Die bakterielle

Selektion dieser auf einem pUC-Replikationsursprung basierenden Expressions-Plasmide erfolgte mit Ampicillin.

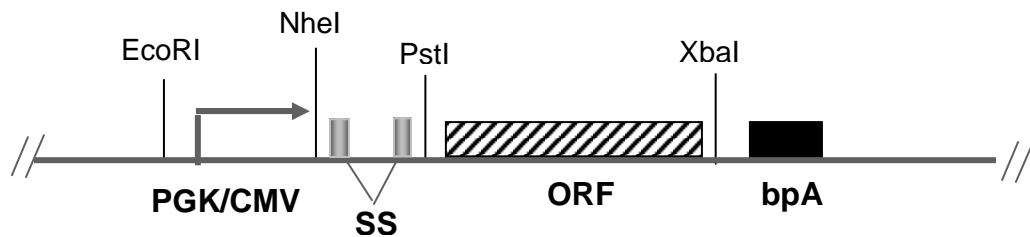


Abbildung C.2: Schematische Darstellung eukaryotischer Expressionsvektoren. Hervorgehoben sind die genetisch relevanten Elemente der Expressionsvektoren: PGK, Phosphoglycerat-Kinase I Promoter; CMV, Cytomegalovirus Promoter; SS, 5' Spleiß-Substrat; ORF, *open reading frame*; bpA, Polyadenylierungs-Sequenz. Das zu exprimierende Gen wird entweder vom PGK oder CMV Promoter transkribiert. 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.

Im folgenden sind die Klonierungen der einzelnen Derivate beschrieben. Die Nucleotidsequenzen der aufgeführten Primer sind in B4 detailliert aufgelistet.

PGK-Expressionsvektoren:

Zur Konstruktion der Expressionsvektoren **pPGKSSInt**, **pPGKSSInt-h** und **pPGKSSInt-h/218** wurde die wild-typ Integrase aus pTrcInt (Christ & Dröge, 1999); Int-h aus pKEXInt-h (Lorbach et al., 2000) und Int-h/218 aus pTrcInt-h/218 (Christ & Dröge, 1999) mit den Primern (Int-N-EU) und (Int-C-EU) in einer Standard-PCR amplifiziert. Die PCR-Fragmente wurden anschließend zwischen die Schnittstellen PstI und XbaI (Abb. C.2) in das pPGKCrebpA kloniert, wodurch das für Cre codierende Fragment durch ein Int-Gen ersetzt wurde.

Im **pPGK** wurde das Int-Gen zusammen mit dem Intron durch einen Restriktionsverdau mit NheI/XbaI deletiert.

Im **pPGKSSInt-hNLS** ist der C-Terminus von Int-h mit der SV40 *large tumor antigen* nuklearen Lokalisationssequenz (NLS; Kalderon et al., 1984) fusioniert. Der Vektor wurde wie pPGKInt-h konstruiert mit der Ausnahme, dass pKEXInt-hNLS als Matrize diente und statt des Primers (Int-C-EU) der Primer (Int-C/NLS-EU) zur Amplifikation eingesetzt wurde.

pPGKSSneo ist ein Derivat von pPGKSSInt-h, in dem das Int-h-Gen zwischen den Schnittstellen PstI und XbaI durch das Neomycin-Resistenz-Gen ersetzt

wurde, dass durch PCR vom pGFP-C1 amplifiziert wurde. Dabei wurde die XbaI-Schnittstelle zerstört.

pPGKSShimA und **pPGKSShimD** sind Derivate von pPGKSSInt, in denen das zwischen den Schnittstellen PstI und XbaI (Abb. C.2) liegende Int-Gen entweder durch das himA- oder das himD-Gen ersetzt wurde. Die Gene der beiden IHF Untereinheiten sind bakteriellen Ursprungs und wurden in einer Standard-PCR mit dem *E.coli*-Genom als Matrize vervielfältigt. himA wurde mit den Primern (himA5-EU)/(himA3-EU) und himD mit den Primern (himD5-EU)/(himD3-EU) erzeugt.

CMV-Expressionsvektoren:

Zur Konstruktion der CMV-Expressionsvektoren, wurde der PGK Promoter zwischen der EcoRI- und NheI-Restriktionsschnittstelle durch ein CMV Promoter-Fragment ersetzt, das mit PCR aus pGFP-C1 (Clontech) amplifiziert wurde.

pCMVSSInt, **pCMVSSInt-h** und **pCMVSSInt-h/218** wurden konstruiert wie pPGKSSInt, pPGKSSInt-h und pPGKSSInt-h/218, außer dass die PCR-Fragmente in den modifizierten Ausgangsvektor pPGKCrebpA, in dem der PGK Promoter durch einen CMV Promoter ersetzt ist, kloniert wurden.

pCMVSSoIHF2 wurde erzeugt, indem das oIHF2-Gen durch PCR mit den Primern (IHF5-1) und (IHF3-3) aus pTrc/oIHF2 kloniert und zwischen die Schnittstellen PstI und XbaI im pCMVSSInt-h kloniert wurde. Dabei wurde das Int-h-Gen durch das oIHF2-Gen ersetzt.

pCMVSSoIHF2NLS wurde wie pCMVSSoIHF2 generiert mit dem Unterschied, dass in der PCR statt des Primers (IHF3-3) der Primer (IHFnl3) eingesetzt wurde. Der C-Terminus des oIHF2-Gens ist in diesem Vektor mit der SV40 *large tumor antigen* nuklearen Lokalisationssequenz fusioniert (Kalderon et al., 1984).

pCMVSSoIHF2HIS wurde kloniert, indem das oIHF2HIS Gen mit NcoI und XbaI aus dem Vektor pTrc/oIHF2HIS geschnitten und in den NcoI/XbaI-geschnittenen Vektor pCMVSSoIHF2 kloniert wurde. Am C-Terminus des oIHF2-Gens sind in diesem Vektor sechs Histidin-Reste angehängt.

Im **pCMV** wurde das Int-Gen zusammen mit dem Intron durch einen Restriktionsverdau mit NheI/XbaI deletiert.

pCMVSSGFP wurde generiert, indem das Int-h-Gen im pCMVSSInt-h mit PstI und XbaI ausgeschnitten und durch ein PCR-Fragment des EGFP-Gens,

amplifiziert vom pGFP-C1 (Clontech), ersetzt wurde. Dabei wurde die XbaI-Schnittstelle zerstört.

C1.9.2 Prokaryotische Expressionsvektoren

Zur Ausprägung der *one-chain Integration Host Factor* (oIHF) Varianten in *E.coli*, wurden jeweils oIHF1-4 in den Expressionsvektor pTrc99a der Firma Pharmacia Biotech (Freiburg) inseriert. oIHF 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 (Abb. D.13-A). Die Übergänge zwischen der Untereinheit β und der integrierten Untereinheit α wurden durch Aminosäure-Linker verbunden, die in den einzelnen oIHF Varianten eine unterschiedliche Länge besitzen (Abschn. D5, Abb. D.13-A; die Nucleotidsequenz ist im Anhang beschrieben). Alle oIHF Varianten wurde durch PCR-Mutagenese (C1.5) generiert, indem zunächst der N-Terminus der Untereinheit β (AS M1 bis Q39), die Untereinheit α (AS L3 bis A94) und der C-terminale Abschnitt der Untereinheit β (AS G40 bis G94) in drei separaten Standard-PCR Reaktionen amplifiziert wurden. Die Primer wurden so konstruiert, dass sie die entsprechenden AS-Linker enthalten, die sich überlappten. Als *Template* dienten pPGKSShimA für die Untereinheit α und pPGKSShimD für die Untereinheit β . An der zweiten AS Position in allen oIHF Proteinen wurde zusätzlich ein Glycinrest eingefügt (Abb. D.13-A).

pTrc/oIHF1 wurde kloniert, indem zunächst die zuvor aufgeführten Proteinabschnitte in drei separaten Ansätzen mit den Primern (IHF5-1)/(IHF3-1), (IHF5-2)/(IHF3-2) und (IHF5-3)/(IHF3-3), in dieser Reihenfolge, vervielfältigt wurden. In einer abschließenden PCR wurde dann oIHF1 mit den Primern (IHF5-1)/(IHF3-3) amplifiziert und zwischen die NcoI- und XbaI-Schnittstelle im pTrc99a (Pharmacia) inseriert.

pTrc/oIHF2 wurde wie pTrc/oIHF1 generiert mit dem Unterschied, dass die Primerpaare (IHF5-1)/(IHF3-3*), (IHF5-4)/(IHF3-4) und (IHF5-5)/(IHF3-3), in dieser Reihenfolge, verwendet wurden.

pTrc/oIHF3 wurde konstruiert, indem der C-terminale Abschnitt von oIHF2, inklusive des 2-AS-Linkers, mit HindIII aus pTrc/oIHF2 geschnitten und zwischen

die HindIII-Schnittstellen im pTrc/oIHF1 kloniert wurde. Dadurch wurde der C-Terminus von oIHF1 durch den von oIHF2 ersetzt.

pTrc/oIHF4 wurde konstruiert, indem der C-terminale Abschnitt von oIHF1, inklusive des 1-AS-Linkers, mit HindIII aus pTrc/oIHF1 geschnitten und zwischen die HindIII-Schnittstellen im pTrc/oIHF2 kloniert wurde. Dabei wurde der C-Terminus von oIHF2 durch den von oIHF1 ersetzt.

pTrc/oIHF2HIS wurde kloniert, indem durch eine Standard-PCR mit den Primern (IHFHIS5) und (IHFHIS3) oIHF2 vervielfältigt wurde und das PCR-Produkt zwischen die Schnittstellen NcoI und XbaI im Vektor pTrc99a kloniert wurde. Am C-Terminus des oIHF2-Gens sind in diesem Vektor sechs Histidin-Reste angehängt.

Die Integrität aller relevanten genetischen Elemente auf den Expressionsvektoren wurde durch DNA-Sequenzierung überprüft.

C1.10 Konstruktion der Substratvektoren

Intramolekulare Substratvektoren

Auf den von Elke Lorbach konstruierten Inversionssubstratvektoren **pGFPattB/attP** und **pGFPattL/attR** (Lorbach et al., 2000) sind die entsprechenden Rekombinationsregionen (*att*) in gegenläufiger Wiederholung angeordnet, sodass durch Rekombination das dazwischenliegende DNA-Fragment umgedreht wird.

Die Deletionssubstratvektoren **pλIR** und **pλER** sind Derivate des pGEM[®]4Z. Auf ihnen sind die *att*-Regionen in direkter Wiederholung angeordnet, sodass die Intermittente sequenz-spezifische Rekombination das dazwischenliegende Fragment ausschneidet. **pλIR** wurde konstruiert, indem *attB* als doppelstängiges Oligonucleotid ((attB53) und (attB35); B4) zwischen die Schnittstellen ClaI und EcoRI im pPGKSSneo (siehe PGK-Expressionsvektoren) kloniert wurde. Der CMV Promoter inklusive des Introns wurde vom pCMVSSInt durch PCR vervielfältigt und zwischen die KpnI- und ClaI-Schnittstelle des *attB*-tragenden Vektors inseriert. Die entstandene CMV-*attB*-neo-Expressionskassette wurde im nächsten Schritt durch PCR amplifiziert und in das BamHI-geschnittene pGEM[®]4Z kloniert. Die *attP*-Region, die einen A→C Nucleotidaustausch im P'-Arm enthält, wurde wie in C1.11 beschrieben durch PCR-Mutagenese erzeugt und mit NsiI und NotI geschnitten. Das *attP*-tragende PCR-Fragment wurden an das 3'-Ende eines BamHI/PstI-

geschnittenen Fragments aus dem pBS302 (Gibco/BRL) ligiert, auf dem sich eine Stoppsignal für die Transkription befindet. Das EGFP Gen wurde wie in C1.11 beschrieben kloniert und zusammen mit dem Stopp/*attP*-Fragment in den BamHI/XbaI-geschnittenen Vektor mit dem CMV Promoter, dem *attB* und der neo-Expressions-kassette kloniert. **pλER** wurde wie pλIR generiert mit dem Unterschied, dass *attL*, erzeugt durch PCR vom pGFP*attL/attR*, anstelle von *attB* in das ClaI/EcoRI-geschnittene pPGKneo eingefügt wurde. Außerdem wurde statt der *attP*-Region die *attR*-Region, die durch PCR vom pGFP*attL/attR* erzeugt wurde, in den NsiI/NotI-geschnittenen Vektor integriert.

pNC9WT wurde zur Untersuchung der Funktionalität von oIHF Proteinen in *E.coli* eingesetzt. Zum einen nimmt es in Rekombinationsassays die Funktion als Substratvektor und zum anderen als Int-Expressionsvektor ein. Es trägt die Rekombinationsregionen *attB* und *attP* in direkter Wiederholung, die eine Int-Expressionskassette flankieren. Durch Rekombination wird diese deletiert. pNC9WT ist ein Derivat von pNC9 (Christ & Dröge, 1999) und wurde konstruiert, indem zunächst das Int-Gen mitsamt dem Trc-Promoter vom pTrcInt (Christ & Dröge, 1999) mit den Primern (Trc5-N) und (Trc3-N) vervielfältigt wurde. Das Trc-tragende PCR-Fragment wurde anschließend zwischen die EcoRI-Schnittstellen im pNC9 kloniert (Abb. D.15-A grauer Pfeil), wobei eine *res*-Region ersetzt wurde.

pACH2 ist ein Derivat von pACYC177 (Rose, 1988). Es enthält eine aus humaner DNA (HeLa) amplifizierte Sequenz (accession no. AC025987, EMBL-Datenbank; Sequenz siehe Anhang), die in entgegengesetzter Orientierung zu einer modifizierten *attP*-Sequenz (*attP*^{*2}) angeordnet ist. Dieser im folgenden als *attH2* betitelte Sequenz wurde mit den Primern (MCTII) und (MCT3') (siehe Anhang und Abschn. B4) in einer Standard-PCR (C1.5) von genomischer HeLa-DNA vervielfältigt. Die PCR-Produkte mit einer Größe von ~660bp wurden zwischen BamHI und AatII eingefügt. *attH2* ist eine natürlich vorkommende *attB*-analoge Sequenz, die sich in vier Nucleotiden im Vergleich zur *attB*-Sequenz unterscheidet (Abb. D.18). Drei Nucleotidaustausche befinden sich in den *core*-Sequenzen und eine in der *overlap*-Region. Um die Homologie zwischen *attH2* und *attP* im *overlap* zu wahren, wurde eine modifizierte *attP*-Region (*attP*^{*2}) generiert. Hierzu wurde das erste Nucleotid im *overlap* von *attP*, ein T, durch C ausgetauscht. Der Nucleotidaustausch wurde mittels einer PCR-Mutagenese

(C1.5) mit den Primern (PFP-BanI)/(attx06) und (attx05)/(P'RP-PstI) eingeführt. Als *Template* diente pGFPattB/attP. Die PCR-Produkte wurden im letzten Schritt in den attH2-tragenden Vektor zwischen die Schnittstellen PstI und BanI kloniert.

Intermolekulare Substratvektoren

Die Substratvektoren mit dem CMV Promoter vor einer Rekombinationsregion sind im Folgenden beschrieben:

Das **pCMVattPmut** enthält drei G→C Austausche im P-Arm, die durch PCR-Mutagenese von Teresa Corona kloniert wurden (Christ et al., 2002). Diese Austausche wurden eingeführt, um ATG Startcodons vor dem EGFP-Gen zu entfernen, die eine ordnungsgemäße Translation von EGFP beeinträchtigen könnten.

pCMVattB wurde ebenfalls von Teresa Corona konstruiert, indem attPmut zwischen der PstI- und XbaI-Schnittstelle im pCMVattPmut durch ein doppelsträngiges attB-Oligonucleotid ersetzt wurde (Christ et al., 2002).

pCMVattL ist ein Derivat von pCMVattPmut in dem attPmut durch attL ersetzt wurde. attL wurde durch PCR vom pλER mit den Primern (attL-5P) und (attL-3X) vervielfältigt und zwischen die PstI und XbaI Schnittstellen im pCMVattPmut eingefügt.

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

pWSattBGFP wurde konstruiert, indem zunächst ein Teil des Hygromycin-Resistenzgens vom pTKHyg (Clontech) mit Aval und NdeI ausgeschnitten wurde und im zweiten Schritt die Schnittstellen mit der Klenow Polymerase wieder aufgefüllt und ligiert wurden. Anschließend wurde ein attB-EGFP PCR-Fragment in die MfeI und HindIII Schnittstellen eingefügt und dabei gleichzeitig 5' von attB eine NheI-Schnittstelle erzeugt. Im letzten Schritt wurde dann das Transkriptions-Stoppsignal (pBS302) zwischen die EcoRI- und NheI-Schnittstelle eingefügt.

pWSattRGFP ist ein Derivat von pWSattBGFP. Das Transkriptions-Stoppsignal wurde durch ein Transkriptions-Stoppsignal-attR Fragment, das mit BamHI und NotI aus pλER geschnitten wurde, ersetzt.

pWSattPGFP ist ebenfalls ein Derivat von pWSattBGFP, auf dem *attB* durch *attP* ausgetauscht wurde. *attP* wurde durch PCR vom pλIR vervielfältigt und zwischen die BamHI- und NotI-Schnittstelle kloniert.

Die Integrität aller relevanten genetischen Elemente auf den Substratvektoren wurde durch DNA-Sequenzierung überprüft.

C1.11 Konstruktion der *Target* Vektoren (*Gene Targeting*)

Die *Target* Vektoren zur Insertion der Integrase Rekombinationssubstrate in das erste Intron des endogenen ROSA26 Locus in ES-Zellen der Maus (Abb. D.8) wurden wie folgt hergestellt:

Zur Konstruktion des *Target* Vektors **pROSA-IR**, auf dem *attB* und *attP* in direkter Wiederholung angeordnet sind, wurde zunächst *attB* als doppelsträngiges Oligonukleotid ((*attB*53) und (*attB*35); B4) in das ClaI/EcoRI-geschnittenen pPGKSSneo (siehe PGK-Expressionsvektoren) kloniert. Anschließend wurde durch die Insertion eines zweiten Oligonucleotides in den ClaI/KpnI-geschnittenen *attB*-tragenden Vektor eine BamHI-Schnittstelle erzeugt. Die 3'-Spleißstelle (SA) wurde aus dem pSAβgeo (Friedrich & Soriano, 1991) ausgeschnitten und zwischen die BamHI- und ClaI-Schnittstelle kloniert. Diese SA-*attB*-neo-Expressionskassette wurde dann durch PCR amplifiziert und in pGEM[®]4Z (Promega) kloniert, wodurch 5' von der SA-Stelle eine XbaI-Schnittstelle inseriert und die BamHI-Schnittstelle zerstört wurde. Der A→C Nucleotidaustausch im P'-Arm der *attP* Region zerstört ein Terminationssignal für die Translation und wurde durch PCR-Mutagenese (C1.5) erzeugt (Primer: (*attP*01), (*attP*02), (PFP-Nsil) und (P'RP-EcoRV-NotI)). Als Matrize diente das pGFP*attB/attP* (Lorbach et al., 2000). Das *attP*-tragende PCR-Produkt wurde mit Nsil und NotI geschnitten und an das 3'-Ende eines BamHI/PstI-geschnittenen Fragments ligiert, auf dem sich ein Stoppsignal für die Transkription befindet. Dieses Fragment stammt vom pBS302 (Gibco/BRL). Das EGFP-Gen und das Polyadenylierungs-Signal wurden durch PCR vom pCMVSSGFP amplifiziert und mit NotI und XbaI geschnitten. Dieses Fragment wurde dann zusammen mit einem BamHI/NotI-geschnittenen Fragment, das den Transkriptions-Stopp und die *attP*-Region umfasst, in den zuvor klonierten Vektor mit der 3'-Spleißstelle, der *attB*-Region und die Neomycinresistenz ligiert. Im letzten Schritt wurden die

generierte Rekombinationskassette in den XbaI-linearisierten *Target* Vektor pROSA26-1 (Soriano, 1999) eingesetzt.

pROSA-ER, der *attL* und *attR* in direkter Wiederholung trägt, wurde wie pROSA-IR kloniert mit dem Unterschied, dass *attL* durch PCR vom pGFP*attL/attR* (Lorbach et al., 2000) vervielfältigt und anschließend in das ClaI/EcoRI-geschnittenen pPGKSSneo kloniert wurde. Die *attR*-Region wurde ebenfalls vom pGFP*attL/attR* durch PCR generiert und wie *attP* zwischen die NsiI und NotI kloniert.

Die Integrität aller relevanten genetischen Elemente auf den *Target* Vektoren wurde durch DNA-Sequenzierung überprüft.

C2 Biochemische Methoden

C2.1 SDS-Polyacrylamid-Gelelektrophorese

Zur Analyse der Integrase und oIHf Proteine in Säugetierzellen oder *E.coli* 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 oIHf Proteinen erfolgte anschließend in einem 12,5% Trenngel. Nach einer Einlaufzeit der Proteine im Sammelgel bei 15mA, erfolgte Elektrophorese bei 30mA.

Die Zell-Lysate wurden hergestellt, indem transfizierte Zellen oder transformierte *E.coli* in 1x SDS *Sample Buffer* (New England Biolabs) aufgenommen und 5min bei 95°C erhitzt wurden. Viskose denaturierte DNA in den Lysaten wurde durch kurzes sonifizieren zerkleinert und für die SDS-PAGE die Ansätze in der Regel in einem Volumen von 80µl aufgetragen.

Nach der Elektrophorese wurden die Gele entweder mit Coomassie-Blau (Sambrook et al., 1989) gefärbt oder für die anschließende *Western Blot* Analyse in einer Naßblotkammer (Bio-Rad) mit 14V (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.

C2.2 Western Blot Analyse

Der Nachweis von Integrase Proteinen in eukaryotischen Zellen durch *Western Blot* Analyse wurde wie folgt durchgeführt: 72h nach der Elektroporation von 60µg Expressionsvektor wurden Zell-Lysate von $1-2 \times 10^6$ transfizierten Zellen durch fünfminütiges Kochen von geernteten HeLa oder NIH/3T3 Zellen 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 verdünntem primären Antikörper (Anti-Integrase, polyklonaler Antikörper gegen *Lambda* Integrase, A. Landy, Providence USA) 1h inkubiert. Die Membran wurde nach Angaben des Herstellers gewaschen und der gebundene primäre Antikörper mit dem zweiten Antikörper Anti-Maus/Anti-Kaninchen IgG-POD (Meerretich-Peroxidase) gebunden. Der Nachweis der Integrase-Proteine auf der Membran erfolgte über eine Chemiluminiszenz-Reaktion, ebenfalls nach Angaben des Herstellers.

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

Zum Nachweis von oIHF Proteinen in Bakterien wurde der *E.coli* Stamm CSH26ΔIHF mit oIHF Expressionsvektoren transformiert. Resistente Klone wurden gepickt und über Nacht bei 37°C kultiviert. Die ÜNK wurden 1:150 verdünnt, und bei einer OD₆₀₀ von ~0,3 wurden die Ansätze geteilt und eine Hälfte mit 2mM IPTG versetzt, um die Expression von oIHF zu induzieren. Nach siebenstündiger Kultur wurde 1ml Kultur pro Ansatz entnommen, die Bakterien durch Zentrifugation pelletiert, die Bakterienpellets in 50µl Tris-HCl aufgenommen, mit 50µl SDS *sample buffer* (New England Biolabs) versetzt und 5min gekocht. Anschließend wurde die genomische DNA durch Sonifizierung zerkleinert. Für die SDS-PAGE wurden 5µl Lysat pro Spur aufgetragen. Die PVDF-Membran mit den übertragenen Proteinen aus dem Polyacrylamidgel wurde wie zuvor beschrieben behandelt mit dem Unterschied, dass der primären Antikörper (Anti-Serum gegen IHF, S.D. Goodman, Los Angeles) 1:1000 in 0,5% *blocking solution* verdünnt wurde.

Der Nachweis von oIHF Proteinen in Lysaten von ES-Reporterzellen, die mit Expressionsvektoren für oIHF und Int-h218 transfiziert wurden (Abschn. C4.4), erfolgte 24h nach der Elektroporation mit dem Anti-Serum gegen IHF.

Zum Nachweis von HimA und HimD in eukaryotischen Zellen, wurden HeLa Zellen mit 60µg pPGKSShimA und/oder pPGKSShimD transfiziert und Zell-Lysate 24h nach der Elektroporation mit dem Anti-Serum gegen IHF im *Western Blot* analysiert.

Als positive Kontrolle wurde ~50ng gereinigtes IHF Protein (Invitrogen) eingesetzt.

C3 Zellbiologische Methoden

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

C3.1 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.

C3.2 Konstruktion stabiler HeLa-Reporter-Zelllinien

HeLa Reporter-Zelllinien mit stabil im Genom integriertem pGFPattB/attP oder pGFPattL/attR wurden von Micha Schwikardi (Lorbach et al., 2000) wie folgt generiert: Zunächst wurden 10 µg pGFPattB/attP oder pGFPattL/attR mit dem Restriktionsenzym ApaI linearisiert und anschließend in getrennten Ansätzen durch Lipofektion mit FuGENE™6 (Roche) in HeLa Zellen transfiziert. Stabile Zelllinien wurden in Kulturmedium für HeLa-Zellen, versetzt mit 500 µg/ml G418 (GibcoBRL), selektioniert. Die Charakterisierung der Reporter-Zelllinien erfolgte durch PCR, DNA-Sequenzierung und genomische *Southern Blot* Analyse.

C3.3 Kultur von embryonalen Stammzellen der Maus (MES)

Für alle Experimente wurde die Mycoplasma-freie ES-Zelllinie IB10 (zur Verfügung gestellt von K. Rajewsky, Institut für Genetik, Universität Köln) verwendet. Diese Linie ist ein Subklon der ES-Zelllinie E14.1 (Kühn et al., 1993), die aus Blastozysten des Maus-Stamms 129/ola generiert wurde. Die ES-Zellkultur wurde wie beschrieben durchgeführt (Pasparakis und Kollias, 1995). Die ES-Zellen wurden in der Regel auf einem mit Mitomycin C mitotisch inaktivierten Zellrasen von Maus-Embryonalen-Fibroblasten (MEF) in der Anwesenheit von rekombinantem *Leukemia Inhibitory Factor* (LIF) kultiviert. LIF wurde freundlicherweise von Angela Egert zur Verfügung gestellt und die Herstellung von Neomycin/Puromycin-resistenten MEF (Mauslinie Dr-4, Tucker et al., 1997) wurde nach Torres und Kühn (1997) durchgeführt. Die ES-Zellen wurden bei 37°C, 10% CO₂ kultiviert und vor Erreichen der Konfluenz passagiert.

Die verwendete FCS-Charge ist zuvor auf ihre Fähigkeit hin getestet worden, das undifferenzierte Wachstum von ES-Zellen zu gewährleisten (auf EF-Zellrasen 50% Plattierungseffizienz, 75% der Kolonien mit undifferenzierter Morphologie). In der Regel wurden die ES-Zellen in einer Dichte von 2×10^4 Zellen/cm² ausplattiert und 2 Tage später geerntet.

Eingefroren wurden die ES-Zellen in 90%FCS/10%DMSO bei -80°C und nach einigen Tagen zur Langzeit-Aufbewahrung in flüssigen Stickstoff überführt. Um den EF-Zellrasen herzustellen, wurde ein Röhrchen präparierter EF-Zellen aufgetaut und auf Kulturschalen (\varnothing 15 cm) ausplattiert. Bei Bedarf wurden die EF-Zellen durch maximal zwei weitere Passagen expandiert und zur mitotischen Inaktivierung bis zur Konfluenz in EF-Medium kultiviert. Die konfluenten EF-Zellen wurden zur mitotischen Inaktivierung für 2-4h mit Mitomycin C-haltigem (10µg/ml) EF-Medium bei 37°C inkubiert, zweimal mit PBS gewaschen, mit Trypsin abgelöst, in EF-Medium aufgenommen und in einer Dichte von 6 bis 8×10^4 Zellen/cm² ausplattiert. Nach 4 bis 8 Stunden wurden die ES-Zellen auf den konfluenten, mitotisch inaktivierten, adhärenen Zellrasen ausplattiert.

Homolog rekombinante ES-Zell Klone wurde wie folgt generiert: 30µg *Target* Vektor (pROSA-IR oder pROSA-ER) wurden mit KpnI linearisiert und durch Elektroporation (500µF, 230V, in 0,4mm Küvetten (peqlab)) in IB10-Zellen transfiziert, die 48h zuvor passagiert wurden. Dazu wurden 1×10^7 ES-Zellen in

800µl RPMI1640 (ohne Phenolrot) resuspendiert und mit dem linearisierten *Target* Vektor gemischt. Die transfizierten ES-Zellen wurden dann in einer Dichte von $2-5 \times 10^5$ /ml auf einen MEF-Rasen in Kulturschalen ausplattiert und nach 24h wurde mit der G418-Selektion (0,2mg/ml) begonnen. Die rekombinanten Klone wurden zwischen Tag 8 und 10 gepickt, in *96-well* Kulturplatten expandiert, drei Replika-Platten bei -80°C eingefroren und mit *Southern Blot* (C1.8) im *96-well* Format (Pasparakis & Kollias, 1995) auf homologe Rekombinante getestet. Positive homologe Rekombinanten wurden dann in 2-3 Passagen expandiert und zur Langzeit-Aufbewahrung bei -80°C eingefroren und in flüssigem Stickstoff gelagert.

C3.4 Durchflusszytometrie

Zur Analyse von EGFP-Ausprägung wurden HeLa oder ES-Zellen geerntet, in Einzelzellsuspension gebracht und mit einem FACScan (Becton Dickinson) analysiert oder einem FACStar+ (Becton Dickinson) sortiert. Der FACStar+ wurde von Christoph Göttlinger bedient. Im FACScan wurden tote Zellen mit Hilfe von 7-Amino-Actinomycin D (0,1µg/ml, Sigma) während der Analyse ausgeschlossen, während im FACStar+ die toten Zellen mit Propidiumiodid (1µg/ml, Sigma) angefärbt wurden. Die FACS-Daten wurden mit Hilfe der Software CellQuest™ (Becton Dickinson) ausgewertet und entweder als einparametrisches Histogramm oder als zweiparametrisches *Dot Plot* Diagramm dargestellt.

C3.5 Fluoreszenz Mikroskopie

Zur Analyse von EGFP-Ausprägung unter dem Fluoreszenz Mikroskop wurden ES-Zellen in geringer Dichte auf Mitomycin C-inaktivierten MEF ausplattiert und 48h in LIF-haltigem Medium kultiviert. Fluoreszenz-Aufnahmen wurden mit einem *Zeiss laser scan microscope* (LSM4) durchgeführt.

C4 Rekombinationsanalysen

C4.1 Rekombinationsanalysen in *E.coli*

Zur Aktivitätsbestimmung verschiedener Integrasen, der klonierten humanen *attH2* und der unterschiedlichen *olHF*-Proteine wurden *in vivo* Rekombinationsanalysen in *E.coli* durchgeführt (Christ & Dröge, 1999). Dazu wurde entweder der

E.coli Stamm CSH26 oder CSH26 Δ IHF mit prokaryotischen Substrat- (C1.10) und Expressionsvektoren (C1.9) nach Hanahan (1983 und 1985) kotransformiert. Einzelne transformierte Klone wurden in dYT-Medium mit den entsprechenden Antibiotikazusätzen über Nacht bei 37°C kultiviert und die Plasmid-DNA mit Hilfe des *QIAprep Spin Miniprep Kit* (Qiagen) isoliert. Die Substratvektoren sind so konstruiert, dass abhängig von der relativen Orientierung der *att*-Regionen zueinander, die sequenz-spezifische Rekombination durch die Integrasen die Substrate in unterschiedlicher Weise verändert. Liegen die *att*-Regionen in direkter Wiederholung vor, führt die sequenz-spezifische Rekombination zur Deletion des dazwischenliegenden DNA-Fragments. Die Wiederholung der *att*-Regionen in gegenläufiger Form dagegen, führt zu einer Inversion des dazwischenliegenden DNA-Fragments. Die veränderten Substrate wurden durch Restriktionsanalysen analysiert, da sich auch das Schnittmuster bestimmter Restriktasen änderte.

Zur Charakterisierung der verschiedenen oIHF-Proteine wurde, wie zuvor beschrieben, *E.coli* CSH26 Δ IHF mit dem entsprechenden oIHF-Expressionsvektor und dem pNC9WT kotransformiert. Klone wurden bis zu einer OD₆₀₀ von 0,4-0,6 kultiviert, der Kulturansatz geteilt und eine Hälfte mit 1mM IPTG versetzt, um die Expression von oIHF zu steigern. Plasmid-DNA wurde ca. 5-6h nach Induktion isoliert, zur Analyse mit Aval geschnitten und in einem Agarosegel elektrophoretisch aufgetrennt. Die Quantifizierung der elektrophoretisch aufgetrennten Aval-Fragmente erfolgte mit der Software: TINA 2.09 (Raytest, Straubenhardt).

C4.2 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, NIH/3T3, CHO und embryonale Mausstammzellen wurden zweimal mit PBS gewaschen bevor sie in 700 μ l Elektroporationsmedium (B7) aufgenommen wurden. In der Regel wurden insgesamt 60 μ g Expressions- und Substratvektor in einem molaren Verhältnis von 1:1 bei 300V und 960 μ F (*Gene pulser*, Bio-Rad) in ungefähr 1×10^7 Zellen transfiziert. Nach der Elektroporation wurden die Zellen aus jedem Ansatz auf

unterschiedliche 10cm Kulturschalen plattiert und zu den in den einzelnen Experimenten angegebenen Zeitpunkten (zwischen 24 und 72h) durchflusszytometrisch untersucht (C3.2) oder die isolierte DNA mit PCR analysiert. Die Produkt der episomalen, intramolekularen Rekombinationen mit den Substraten $\rho\lambda$ IR und $\rho\lambda$ ER (Abschn. D2.1) wurden in einer Standard-PCR mit den Primerpaaren (p5)/(p9) und (PE)/(PGK-RP) und die Rekombinationsprodukte der episomalen, intermolekularen Rekombinationen mit dem Primerpaar (PE)/(p9) amplifiziert. Die Integrität aller Produkte wurde durch Sequenzierung überprüft.

Die Produkte der episomalen Inversion mit den Substraten ρ GFPattB/attP und ρ GFPattL/attR wurden in einer Standard-PCR mit dem Primerpaar p3/p4 (Lorbach et al., 2000) analysiert und die Integrität der Sequenz durch Sequenzierung überprüft.

Die Produkte der episomalen Inversion mit dem Substrat ρ ACH2 in HeLa wurden in einer Standard-PCR mit dem Primerpaar PI/PE (Abb. D.20-A) analysiert und die Integrität der Sequenz durch Sequenzierung überprüft.

Eine Int-vermittelte alternative Verknüpfung zwischen den att-Sequenzen (BOP statt BOP', Christ & Dröge, 1999), die bei den Inversionssubstraten ρ GFPattB/attP und ρ GFPattL/attR zur Deletion der att-flankierten DNA-Sequenzen führen würde, wurde in einer Standard-PCR mit dem Primerpaar (p5)/(p6) (Lorbach, 2000) ausgeschlossen. In Rekombinationsanalysen mit den Substraten $\rho\lambda$ IR und $\rho\lambda$ ER wurde eine alternative Verknüpfung mit dem Primerpaar (p5)/(PE) ausgeschlossen.

Die Transfektionseffizienzen wurden in der Regel durch die Elektroporation von ρ GFP-C1 in die jeweiligen Zellen und einer anschließenden FACS-Analyse nach 48h bestimmt.

Episomale Rekombinationsanalysen in denen **gereinigtes IHF** (Invitrogen) in HeLa kotransfiziert wurde sind im Folgenden beschrieben: Zunächst wurden 30 μ g eines Int-Expressionsvektors, wie zuvor beschrieben, durch Elektroporation in ungefähr 6×10^6 HeLa Zellen transfiziert. Dann wurden nach 3 bis 4h ungefähr 1×10^5 Zellen mit 2 μ g Substratvektor für die intramolekulare Rekombination und 2 μ g Substratvektoren, in einem molaren Verhältnis von 1:1, für die intermolekulare Rekombination transfiziert. Die Substratvektoren wurden dazu mit dem gereinigten IHF mindestens 30 Minuten bei Raumtemperatur vorinkubiert und die Transfektion

der IHF-DNA Komplexe erfolgte dann durch Lipofektion (FuGENE™6, Roche). Die Analyse der doppelt transfizierten Zellen erfolgte 48h nach der zweiten Transfektion durch Durchflusszytometrie (C3.2). Die Transfektionseffizienzen wurden in der Regel durch die Transfektion von pGFP-C1 in einem separaten Ansatz und einer anschließenden FACS-Analyse nach 48h bestimmt. Durch Elektroporation hatten durchschnittlich 90% der Zellen die Plasmid-DNA aufgenommen, wogegen durch Lipofektion in der Regel ca. 80% der Zellen transfiziert wurden.

C4.3 Rekombinationsanalysen in HeLa-Reporter-Zelllinien

HeLa Reporterzellen, die entweder pGFP $_{attB}$ / $attP$ oder pGFP $_{attB}$ / $attP$ stabil im Genom integriert haben, wurden mit 20 μ g eines Expressionsvektors der Integrase transfiziert. Dazu wurde zirkuläre Expressionsvektor-DNA in etwa 1×10^7 Zellen/700 μ l durch Elektroporation bei 280V und 960 μ F (Bio-Rad *gene pulser*) in die Reporterzelllinien eingeschleust. Der Nachweis von Integrase-vermittelter genomischer Rekombination in den HeLa Reporterzellen erfolgte 72h nach der Transfektion durch *Southern Blot* und PCR-Analyse mit den Primern (p4) und (p3) (B4; Lorbach et al., 2000). Die PCR-Produkte wurden durch Sequenzanalyse auf ihre Richtigkeit hin überprüft. Die Transfektionseffizienz wurde in jedem Experiment in einem separaten Ansatz mit 20 μ g pGFP-C1 ermittelt und betrug im Schnitt etwa 50%.

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 einer Standard-PCR mit dem Primerpaar (p5)/(p4) (Lorbach, 2000) ausgeschlossen.

C4.4 Rekombinationsanalysen in ES-Reporter-Zelllinien

3×10^6 ROSA-IR oder ROSA-ER Reporterzellen wurden in der Regel 24h nach ihrer Passage mit insgesamt 60 μ g Expressionsvektor bei 290V und 500 μ F transfiziert. In Rekombinationsanalysen mit oIHF wurden von jedem der Expressionsvektoren 40 μ g transfiziert. Nach der Elektroporation wurden je 1/3 der Zellen auf Mitomycin C-behandelten MEF Zellrasen in 10 \varnothing cm Kulturschalen in LIF-haltigem ES-Medium (B7) kultiviert. EGFP-exprimierende ES-Zellen wurden

entweder 72h nach Transfektion mit Hilfe der Durchflußzytometrie analysiert oder nach 48h sortiert (C3.4). Sortierte ES-Zellen wurden in geringer Dichte weiter auf einem Mitomycin C-inaktivierten MEF Zellrasen in doppelt konzentriertem LIF-haltigem ES-Medium kultiviert und einzelne ES-Zell Kolonien gepickt und expandiert. Der Nachweis von Integrase-vermittelter Rekombination auf genomischer Ebene, die Deletion des Neomycin-Resistenzgens in den ES-Zelllinien ROSA-B und ROSA-L, erfolgte zum einen durch PCR mit dem Primerpaar (p10)/(p9) (B4, Abb. D.8-C) und durch *Southern Blot* Analyse (C1.8). Die Integrase-vermittelte Expression von EGFP wurde mit Durchflusszytometrie (C3.4) oder durch Fluoreszenz Mikroskopie (C3.5) analysiert und dokumentiert.

Die Transfektionseffizienz wurde in jedem Experiment in einem separaten Ansatz mit 60µg pGFP-C1 ermittelt und betrug zwischen 60 und 70%.

D Ergebnisse

D1 Ausprägung der *Lambda* Integrase in Säugetierzellen

Zu Beginn dieser Arbeit wurde gezeigt, dass die *Lambda* Integrase (Int), die prokaryotischen Ursprungs ist, auch in eukaryotischen Zellen ausgeprägt werden kann. Dazu wurden die eukaryotischen Expressionsvektoren pPGKSSInt-h und pPGKSSInt-hNLS (Abb. D.1-A) konstruiert, die Int-h und Int-hNLS vom Phosphoglycerat-Kinase (PGK) Promoter ausprägen. Int-hNLS ist ein Derivat von Int-h, an dessen 3'-Ende des offenen Leserasters eine nukleare Lokalisationssequenz angehängt ist. Dem Kontrollplasmid pPGK fehlt eine Integrase kodierende Region. Der Nachweis der Integrase Proteine erfolgte durch *Western Blot* Analyse von HeLa Zell-Lysaten, die 72h nach der Elektroporation der Expressionsvektoren erzeugt wurden. Abbildung D.1-B zeigt einen repräsentativen *Western Blot* mit Zell-Lysaten der transfizierten HeLa Zellen. Ein Protein mit dem

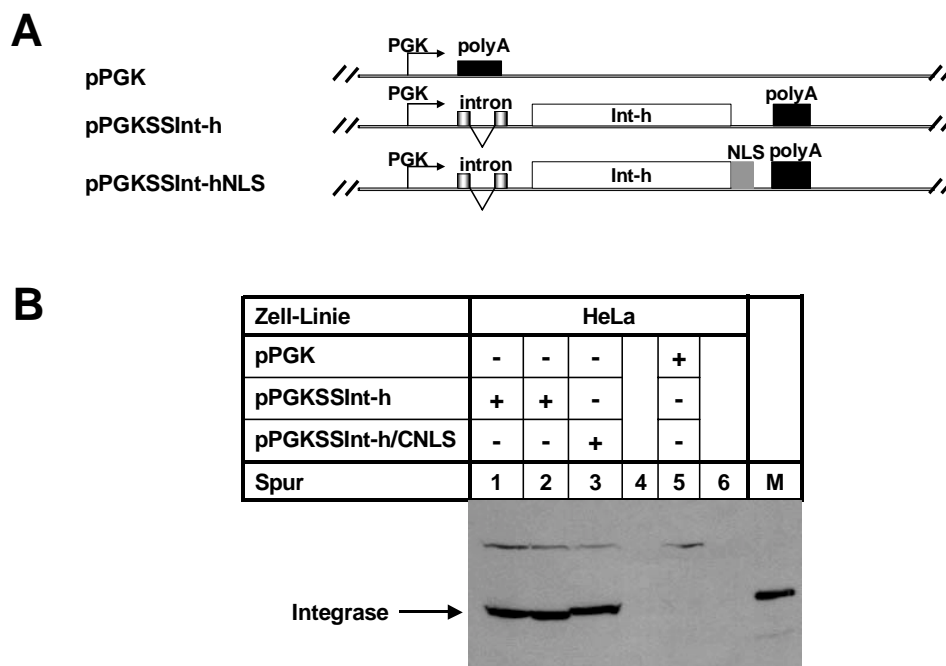


Abbildung D.1: Ausprägung der λ -Integrase in HeLa. In (A) sind Expressionsvektoren für Int-h und Int-hNLS dargestellt, die zur Ausprägung der Integrasen, in HeLa transfiziert wurden. PGK, Phosphoglycerat-Kinase Promoter; NLS, SV40 *large tumor antigen* nukleare Lokalisationssequenz; polyA, Polyadenylierungs-Sequenz. (B) 72h nach der Elektroporation wurde die Integrase durch eine *Western Blot* Analyse in Zell-Lysaten von transfizierten HeLa Zellen mit polyklonalen Int-Antikörpern nachgewiesen. M, Marker: Lysat einer Int-h/NLS ausprägenden Bakterienkultur.

erwarteten molekularen Gewicht von ca. 40 KDA wurde mit einem Anti-Serum gegen die wild-typ Integrase (A. Landy, Providence, USA) nur in Zellen nachgewiesen, die den Int-h kodierenden Vektor enthalten (Abb. D.1-B Spur 1 und 2). Das etwas langsamere Laufverhalten des Proteins in den pPGKSSInt-hNLS transfizierten HeLa Zellen (Spur 3) ist auf 7 zusätzliche Aminosäurereste der NLS zurückzuführen. In Zellen, die mit pPGK transfiziert wurden, ist dagegen kein Protein mit diesem Molekulargewicht nachweisbar (Spur 5). Als Kontrolle und Größenmarker wurde ein Lysat von Bakterienzellen eingesetzt, die Int-hNLS ausprägten. Auch in NIH/3T3 Zellen, die mit Int-Expressionsvektoren transfiziert wurden, konnten Integrase Proteine durch *Western Blot* Analyse nachgewiesen werden (Daten nicht gezeigt). Diese Experimente zeigen, dass die prokaryotische Integrase sowohl in humanen als auch in Mauszellen ausgeprägt wird. Im folgenden wurde untersucht, ob es sich dabei auch um eine funktionsfähige Rekombinase handelt.

D2 Episomale sequenz-spezifische Rekombination in Säugetierzellen durch Integrase Mutanten

Die zentrale Frage dieser Arbeit ist: Funktioniert das sequenz-spezifische Rekombinationssystem des Phagen *Lambda* auch in Eukaryoten? Nach den Kenntnissen über das Int-System muß bei der Übertragung in Eukaryoten folgendes berücksichtigt werden: Die Bildung des Intasoms in der integrativen Rekombination erfordert eine negative Torsionsspannung der *attP*-tragenden DNA-Sequenz. Durch das Einschleusen einer *attP*-tragenden Fremd-DNA in eine eukaryotische Zelle ist jedoch zu erwarten, dass diese relaxiert wird, da die Zelle zwar Topoisomerasen zur Relaxierung, jedoch keine prokaryotische Gyrase für das Einführen superhelikaler Windungen besitzt. Dieses Problem kann mit den Integrase Mutanten Int-h (Miller *et al.*, 1980) und dessen Derivat Int-h/218 (Christ & Dröge, 1999) gelöst werden. Lange-Gustafson und Nash zeigten 1984, dass es für die Aktivität von Int-h geringe Bedeutung hat, ob das Substrat linear oder unter Torsionsspannung steht. Im Gegensatz zu der integrativen Rekombination ist die exzusive Rekombination zwar unabhängig von negativer Torsionsspannung der DNA, sie wird jedoch durch sie stimuliert. Dieser Punkt ist deshalb wichtig, da auch

die Effizienz der Rekombinationsreaktion eine wichtige Rolle in der Übertragung des Systems in Eukaryoten spielt. Um zunächst die Funktionalität der eukaryotisch ausgeprägten Integrasen in Säugetierzellen zu untersuchen, ist auch die höhere enzymatische Aktivität der Mutanten in der Abwesenheit von Hilfsproteinen von Vorteil, da weder IHF noch XIS in eukaryotischen Zellen vorkommen.

D2.1 Analyse der enzymatischen Aktivitäten von Int und den Derivaten Int-h und Int-h/218 in intra- und intermolekularen Rekombinationreaktionen

Um die enzymatische Aktivität der Int und den Mutanten Int-h und Int-h/218 in eukaryotischen Zellen zu untersuchen, wurden Substratvektoren für die integrative und die exzise Rekombination konstruiert. Eine interessante Frage in diesem Zusammenhang ist, wie *att*-Sequenzen, die sich auf demselben oder auf unterschiedlichen episomalen DNA-Segmenten befinden von den Integrasen rekombiniert werden können. Dazu wurde zunächst ihre enzymatische Aktivität sowohl in der **intra-** als auch in der **intermolekularen**, integrativen Rekombination untersucht.

Die Substratvektoren zur Analyse der **intramolekularen** Rekombination wurden so konstruiert, dass *attB* und *attP* als direkte Wiederholungen ein transkriptionelles Stoppsignal flankieren (Abb. D.2-A). Diese Rekombinationskassette wiederum ist am 5'-Ende vom CMV Promoter und am 3'-Ende vom EGFP offenen Leseraster eingerahmt. Rekombination zwischen *attB* und *attP* erzeugt die Produkte *attL* und *attR* und führt zum Ausschneiden des Stoppsignals. Die daraus resultierende Ausprägung von EGFP erlaubt es, Rekombinationsereignisse mit Durchflusszytometrie (FACS) nachzuweisen. Expressionsvektoren für Int (pCMVSSInt), Int-h (pCMVSSInt-h) oder Int-h/218 (pCMVSSInt-h/218), siehe CMV-Expressionsvektoren unter C1.9.1, wurden mit den Substratvektoren in HeLa Zellen kotransfiziert. Der Vektor pCMV ohne das offene Leseraster einer Integrase diente als Negativ-Kontrolle. Die Transfektionseffizienz wurde für jedes einzelne Experiment bestimmt und betrug in der Regel 95 bis 98% (Daten nicht gezeigt). Die FACS Analysen von drei Experimenten zeigen, dass beide Mutanten die Substrate effizient rekombinieren. In einigen Experimenten prägten bis zu 30% der Zellen EGFP aus (Abb. D.2-A, Diagramm). Die Integrität der erwarteten Rekombinationsprodukte wurde indirekt durch Sequenzierung von PCR-Fragmenten bestätigt (Daten nicht gezeigt). Die Analysen der episomalen,

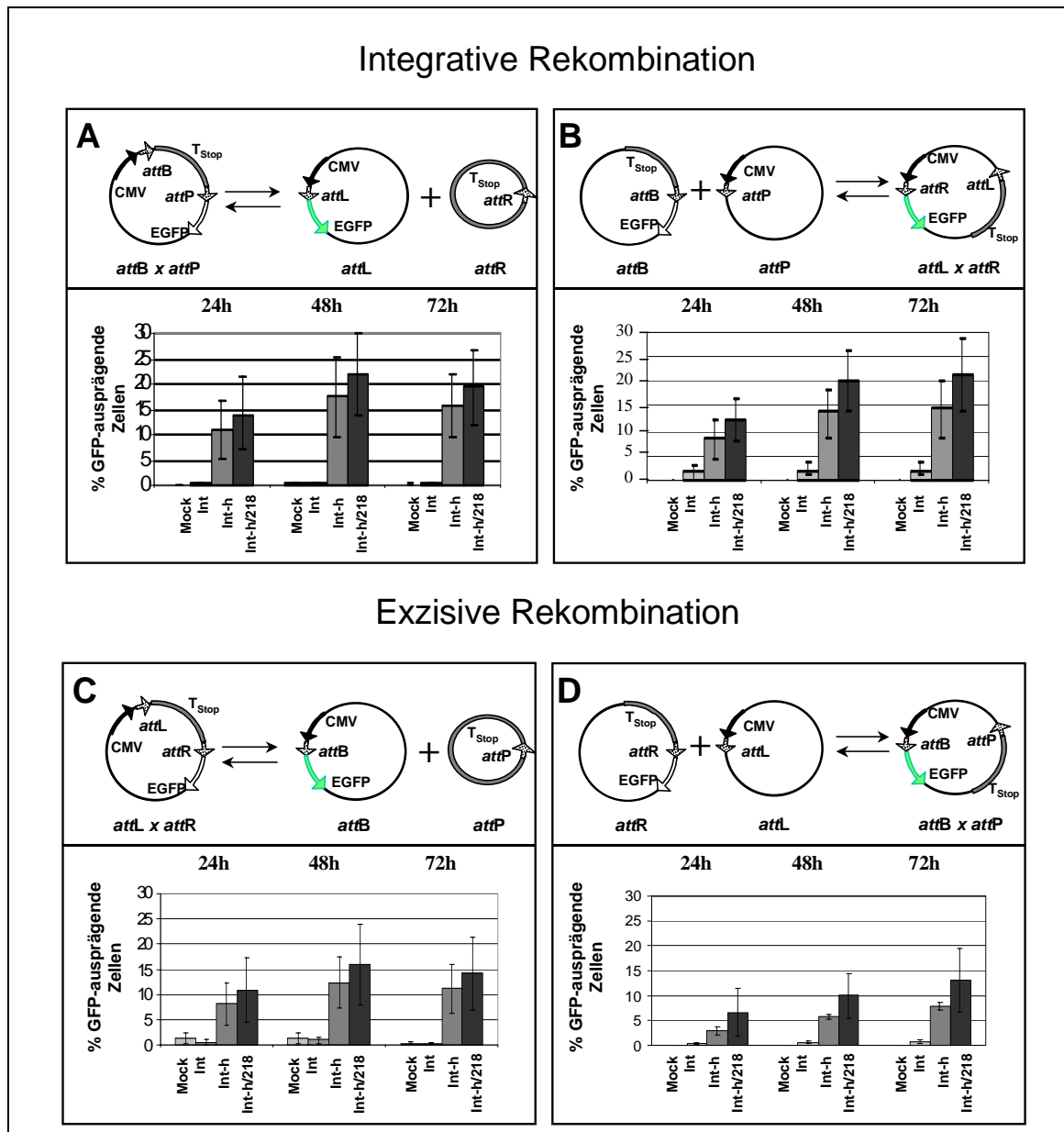


Abbildung D.2: Intra- und Intermolekulare *Lambda* Rekombinationsreaktionen. (A) Intramolekulare integrative (*attB*x*attP*) Rekombination. (B) Intermolekulare integrative (*attB*x*attP*) Rekombination. (C) Intramolekulare exzusive (*attL*x*attR*) Rekombination. (D) Intermolekulare exzusive (*attL*x*attR*) Rekombination. Die Substratvektoren und ihre rekombinanten Produkte sind schematisch über jedem Diagramm dargestellt. Der Anteil an EGFP-ausprägenden HeLa Zellen wurde durchflusszytometrisch zu drei Zeitpunkten nach der Kotransfektion von Substrat und Expressionsvektoren bestimmt. Die Balken zeigen die Mittelwerte aus drei Analysen mit den Standardabweichungen, die als vertikale Linien dargestellt sind. CMV, Cytomegalovirus Promoter; T_{Stop}, transkriptionelles Stoppsignal; EGFP, *enhanced green fluorescence protein* kodierendes Gen.

integrativen Rekombination zeigen auch, dass die Doppel-mutante Int-h/218 enzymatisch aktiver ist als Int-h. Im Gegensatz dazu ist die wild-typ Integrase kaum aktiv. Der Anteil an EGFP-ausprägenden Zellen stieg innerhalb der ersten 48h nach der Transfektion an und blieb dann die folgenden 24h konstant. Die

Messungen der Rekombinationsreaktionen nach 24, 48 und 72h zeigen zudem, dass die Mehrheit der Rekombinationsereignisse innerhalb der ersten 24h stattgefunden haben muss. Dies wird zusätzlich durch eine *Western Blot* Analyse von Int-h/218-ausprägenden HeLa Zellen unterstützt, die in Zusammenarbeit mit Teresa Corona durchgeführt wurde. Die *Western Blot* Analyse zeigt, dass über einen Zeitraum von 72h nach der Elektroporation, die Int-h/218 Menge nach 24h, oder zu einem früheren Zeitpunkt, am höchsten war. Die folgenden 48h nahm sie dann kontinuierlich ab (Abb. D.3, Spur 2 bis 4). In episomalen Rekombinationsanalysen mit diesen intramolekularen Substratvektoren kann nicht ausgeschlossen werden, dass einem Teil der EGFP-ausprägenden Zellen auch inter- anstelle von intramolekularer, integrativer Rekombination zu Grunde liegt. Jedoch auch wenn dies der Fall sein sollte, können die Daten als Referenz für die nachfolgenden intermolekularen Rekombinationsanalysen herangezogen werden, um eine Aussage über die Effizienz der intermolekularen Rekombination im Vergleich zur intramolekularen Reaktion zu treffen.

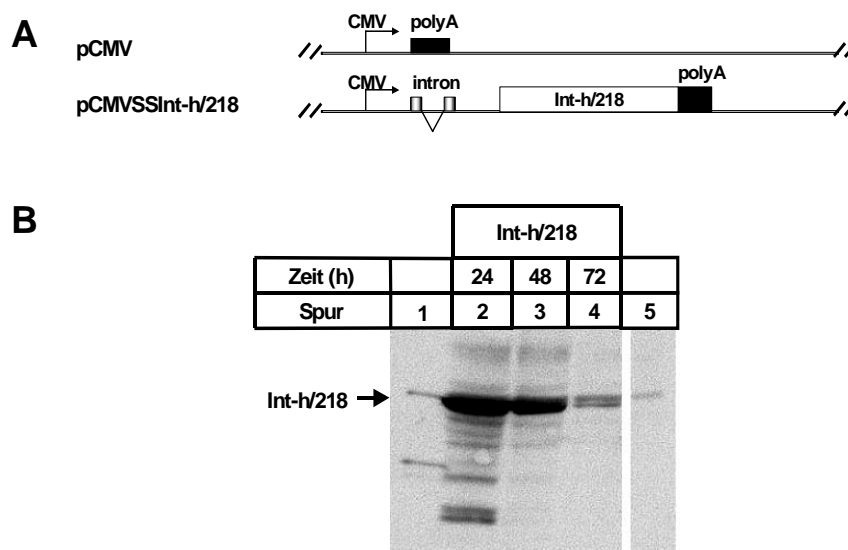


Abbildung D.3: Int-h/218-Ausprägung in HeLa Zellen über einen Zeitraum von 72h nach der Elektroporation. (A) HeLa Zellen wurden mit einem CMV-Expressionsvektor zur Ausprägung der Int-h/218 Mutante transfiziert. CMV, Cytomegalovirus Promoter; polyA, Polyadenylierungs-Sequenz. (B) Zu den angegebenen Zeitpunkten 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 Antikörper gegen Wildtyp-Int eingesetzt. Spur 1, *E.coli* Zell-Extrakt mit Wildtyp-Integrase. Spur 2 bis 4, HeLa Zell-Lysat mit Int-h/218. Spur 5, HeLa Zell-Lysat von nicht-transfizierten HeLa Zellen vom selben Gel.

Die **intermolekulare**, integrative Rekombination wurde untersucht, indem *attB* und *attP* auf verschiedene Plasmide kloniert wurden. Die sequenz-spezifische

Rekombination von *attB* und *attP* versetzt den CMV Promoter an das 5'-Ende des offenen Leserasters von EGFP (Abb. D.2-B), wodurch EGFP nur nach der intermolekularen Rekombination ausprägt werden kann. Die Ergebnisse der FACS Analysen von HeLa Zellen, die mit den beiden Substratvektoren der intermolekularen Rekombination und den Int-Expressionsvektoren kotransfiziert wurden, waren vergleichbar mit denen der intramolekularen Rekombinationsanalysen (Abb. D.2-B, Diagramm). Auch bei der intermolekularen Rekombination muss die Mehrzahl der Reaktionseignisse innerhalb der ersten 24h nach der Elektroporation stattgefunden haben und die enzymatische Aktivität von Int-h/218 war ebenfalls höher als die von Int-h. Ebenso katalysiert die wild-typ Int, im Vergleich zu den Mutanten, in diesen Analysen kaum die intermolekulare Rekombination in HeLa. Diese Ergebnisse zeigen, dass über einen Zeitraum von 72h, die intermolekulare Rekombination mindestens genauso effizient ist wie die entsprechende intramolekulare Rekombination.

Mit der gleichen experimentellen Strategie wurde auch die **intra-** und **intermolekulare**, exzusive Rekombination verglichen. Wie schon zuvor mit *attB* und *attP* beobachtet, bestätigen auch die Rekombinationsanalysen zwischen *attL* und *attR*, dass sich die Effizienz der intermolekularen Rekombination nicht signifikant von der Effizienz der intramolekularen Rekombination unterscheidet (Abb. D.2-C und -D). Die Ergebnisse zeigen jedoch auch, dass die Effizienz der exzisiven Rekombinationsreaktionen im Allgemeinen etwas niedriger ist verglichen mit denen der integrativen Rekombination.

Abhängig von der Orientierung der *att*-Sequenzen zueinander kann die dazwischen liegende DNA-Sequenz durch die Int entweder ausgeschnitten oder invertiert werden. Um zu prüfen, ob die Mutanten auch in der Lage sind in eukaryotischen Zellen DNA-Segmente zu invertieren, wurden die Vektoren pGFP*attB/attP* und pGFP*attL/attR* als Substrate eingesetzt. Auf ihnen sind die *att*-Sequenzen in invertierter Wiederholung angeordnet und flankieren dabei das EGFP-Gen. Diese Rekombinationskassette grenzt ihrerseits am 5'-Ende an den CMV Promoter und am 3'-Ende an eine Polyadenylierungssequenz (Abb. D.6-A und D.7-A; Lorbach et al., 2000). Die Substratvektoren wurden zusammen mit dem Expressionsvektor pPGKInt-h (Abb. D.1-A) in HeLa transfiziert und 72h später die DNA mit PCR analysiert. pPGK wurde als Negativ-Kontrolle eingesetzt. Die Ergebnisse einer PCR-Analyse mit dem Primerpaar (p3)/(p4) zeigen, dass die

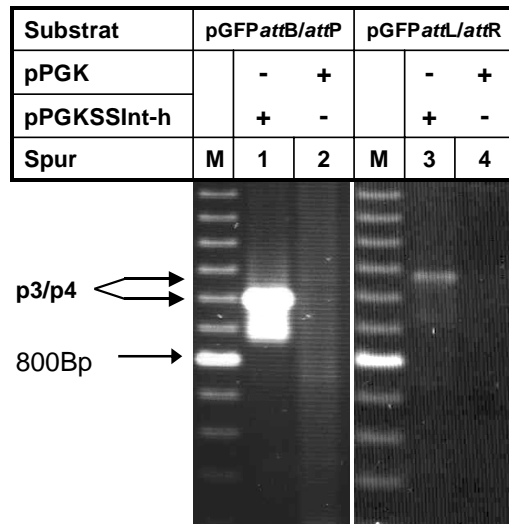


Abbildung D.4: PCR-Analyse von integrativer und exziser Rekombination in transienten Kotransfektions Experimenten mit Inversionsubstraten. Die Substratvektoren pGFPattB/attP und pGFPattL/attR wurden mit den PGK-Expressionsvektor für Int-h in HeLa Zellen ko-transfiziert. 72 Stunden nach der Transfektion wurde die DNA isoliert und durch PCR analysiert. Nur nach Inversion des att-flankierten EGFP-Gens durch Int-h entstehen PCR Produkte dem Primerpaar p3/p4 (siehe Abb. D.6-A und D.7-A). M: 100Bp-Marker

Mutante Int-h sowohl die *attB* und *attP*- als auch die *attL* und *attR*-flankierten episomalen DNA-Abschnitte in HeLa Zellen invertieren können (Abb. D.4 Spur 1 und 3).

D2.2 Die DNA Arm-Bindungssequenzen der *att*-Regionen stimulieren die Rekombination; sie sind jedoch nicht notwendig

Die Ergebnisse der episomalen Rekombinationsanalysen in HeLa zeigen, dass die Mutanten in der Lage sind, effizient die integrative und exzise Rekombination in eukaryotischen Zellen zu katalysieren. Die enzymatische Aktivität der wild-typ Integrase ist dagegen marginal. Dies deutet daraufhin, dass Hilfsproteine, die die natürlichen Hilfsproteine IHF und XIS ersetzen können, nicht oder zumindest nicht in ausreichender Menge in den menschlichen Zellen vorhanden sind. Episomale Substrate liegen schon bald nach ihrer Transfektion in den Zellen in relaxierter topologischer Form vor (Schwikardi & Dröge, 2000). Das legt den Schluss nahe, dass die Int Mutanten die Rekombination katalysieren, ohne zuvor einen definierten Nukleoprotein-Komplex, vergleichbar mit dem Intasom an *attP*, zu bilden. Das wiederum wirft die Frage auf, welche funktionelle Rolle die DNA Arm-Bindungsregionen unter diesen Umständen in der *Lambda* Rekombination spielen. In den bisher durchgeführten Analysen waren die Arm-Bindungsregionen immer in mindestens einem Rekombinationspartner vorhanden. Um die funktionelle Rolle der Arm-Bindungssequenzen besser untersuchen zu können, wurden in Zusammenarbeit mit Teresa Corona (Institut für Genetik, Universität Köln; Christ et al., 2002) intermolekulare Rekombinationsanalysen mit verschiedenen kombinierten

Paaren von Substratvektoren, die entweder *attB* oder *attP* tragen, in HeLa Zellen durchgeführt (Abb. D.5-A). 48h nach der Kotransfektion der Substratvektoren mit einem Expressionsvektor für die wild-typ Integrase, Int-h oder Int-h/218 wurde der Anteil an EGFP-ausprägenden Zellen durch FACS ermittelt. Die Transfektions-effizienzen betragen in der Regel >90% (Daten nicht gezeigt). Die Ergebnisse von drei Experimenten zeigen, dass die intermolekulare Rekombination zwischen *attP*-Paaren genauso effizient ist wie die Rekombination zwischen *attB* und *attP* (Abb. D.5-B). Die Rekombination zwischen zwei *attB*-Regionen konnte dagegen nur von der Mutante Int-h/218 katalysiert werden. Die Effizienz dieser Reaktion war durchschnittlich ca. vierfach niedriger als die Effizienz der Rekombination zwischen *attP* und *attB* (Abb. D.5-B). Dies zeigte sich darin, dass der Anteil an EGFP-ausprägenden Zellen, die von Rekombinationsereignissen zwischen zwei *attB*-Regionen stammen, auf 4 bis 5% sank. Diese Ergebnisse machen deutlich, dass Int-h/218 die Arm-Bindungssequenzen in den *att*-Regionen zur Rekombination nicht benötigt. Sie zeigen jedoch auch, dass durch die Anwesenheit der Arm-

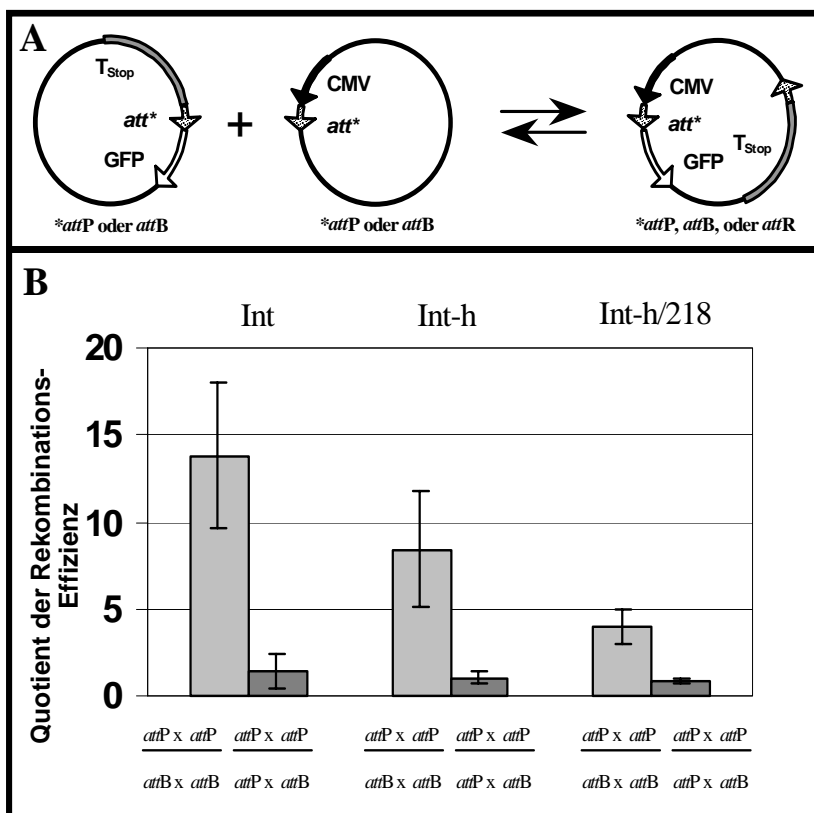


Abbildung D.5: Int DNA Arm-Bindungssequenzen in *att*-Regionen stimulieren die intermolekulare Rekombination. (A) Substratvektoren für die intermolekulare Rekombination, die entweder *attB* oder *attP* tragen und in verschiedenen Kombination zu Produkten rekombiniert werden können, die EGFP vom CMV Promoter ausprägen. (B) Verschiedene Kombinationen von Substratvektoren wurden mit den CMV-Expressionsvektoren für Int, Int-h oder Int-h/218 in HeLa Zellen kotransfiziert. 48h nach der Elektroporation wurden die Zellen durch FACS analysiert und der Quotient der GFP-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.

Bindungssequenzen in den *att*-Regionen die Rekombination durch Int-h/218 signifikant gesteigert werden kann. Der stimulatorische Effekt der Arm-Bindungssequenzen in den *att*-Regionen ist sogar noch stärker (ungefähr 8-fach), wenn anstelle von Int-h/218 die Mutante Int-h in den Zellen ausgeprägt wurde. Die marginale Rekombinationsaktivität der wild-typ Integrase scheint dagegen nahezu ausschließlich von der Gegenwart der Arm-Bindungssequenzen in den *att*-Regionen abhängig zu sein.

D2.3 Ausschluss einer alternativen Verknüpfung der *att*-Sequenzen

Die Mutanten Int-h und Int-h/218 produzieren unter bestimmten Bedingungen in *E.coli* zusätzlich neben den konventionellen Produkten ein alternatives Produkt (Christ & Dröge, 1999). Hierbei handelt es sich um eine ungewöhnliche Verknüpfung der *att*-Sequenzen. Anstelle von BOP' entsteht an den *att*-Regionen die Verknüpfung BOP. Bei den Substratvektoren pGFP*attP/attB* und pGFP*attL/attR* führt dies zu einer Deletion anstelle von Inversion der *att*-flankierten DNA-Sequenzen. In Rekombinationsanalysen mit den Substratvektoren pGFP*attP/attB* und pGFP*attL/attR* wurde durch PCR-Analysen eine solche Deletion dadurch ausgeschlossen, dass mit einem entsprechenden Primerpaar (Abschn. C4.2) kein PCR-Produkt der erwarteten Größe nachgewiesen werden konnte (Daten nicht gezeigt). Auch in den restlichen episomalen Rekombinationsanalysen konnte mit geeigneten PCR-Analysen, zumindest im Rahmen der Nachweisgrenze für PCR, eine alternative Verknüpfung der *att*-Sequenzen durch die Mutanten ausgeschlossen werden (Daten nicht gezeigt).

D2.4 Episomale sequenz-spezifische λ -Rekombination in Säugetierzellen

Die enzymatische Aktivität der Integrase Mutanten konnte außer in der menschlichen Zelllinie HeLa auch in anderen eukaryotischen Zellen nachgewiesen werden: In der Mausfibroblasten-Zelllinie NIH/3T3 wurde in episomalen Rekombinationsanalysen mit den Inversionssubstraten pGFP*attB/attP* und pGFP*attL/attR* durch PCR-Analysen gezeigt, dass die Mutanten integrative und exzusive Rekombinationsreaktionen katalysieren (Daten nicht gezeigt). Episomale Rekombinationsanalysen mit den Deletionssubstraten p λ IR und p λ ER bestätigten

die enzymatische Aktivität der Mutanten auch in der Hamsterzelllinie CHO und in embryonalen Stammzellen der Maus (IB10) (Daten nicht gezeigt).

D3 Sequenz-spezifische Rekombination in Säugetiergenomen durch Integrase Mutanten

Bisher wurde die enzymatische Aktivität der Integrase Mutanten in Säugetierzellen ausschließlich in episomalen Rekombinationsanalysen bestätigt und charakterisiert. Chromosomale DNA in Eukaryoten unterscheidet sich von genomischer DNA in Prokaryoten unter anderem durch ihre nukleosomale Organisation und fehlender freier Torsionsspannung. Mit den folgenden Experimenten wurde untersucht, ob auch stabil ins eukaryotische Genom integrierte Substrate von den Integrase Mutanten rekombiniert werden können.

D3.1 Genomische Inversion in HeLa Reporter-Zelllinien

Die integrative und exzusive Inversion von genomischen DNA-Abschnitten durch Int im Säugetiergenom wurde in HeLa Zellen untersucht, die pGFP $attB/attP$ oder pGFP $attL/attR$ stabil im Genom integriert haben. Zur Ausprägung der Integrasen in diesen Reporterzellen wurden CMV-Expressionsvektoren (C1.9.1) für Int, Int-h oder Int-h/218 in die Zellen transfiziert. Als negative Kontrolle diente der Vektor pCMV. 72h nach der Elektroporation wurde die genomische DNA isoliert, mit NcoI geschnitten und durch *Southern Blot* Analyse untersucht. Die sequenz-spezifische Inversion des EGFP-Gens im Genom ändert das Restriktionsmuster von NcoI und erzeugt ein Restriktionsfragment von ca. 2Kb, während das entsprechende Fragment im nicht-rekombinierten Substrat eine Größe von 1Kb besitzt. Die Substrat- und Produkt-spezifischen Restriktionsfragmente wurden mit dem EGFP-Gen als Sonde (Abb. D.6-A) nachgewiesen.

Die integrative Rekombination zwischen $attB$ und $attP$ durch die Integrasen wurden in der Reporterzelllinie H₈B untersucht. In dieser Reporterzelllinie sind ca. 8 Kopien des Substrats pGFP $attB/attP$ an zufälligen Positionen im Genom integriert (Daten nicht gezeigt). Die Ergebnisse einer *Southern Blot* Analyse in dieser Reporterzelllinie zeigen, dass ein durch Rekombination entstandenes genomisches Restriktionsfragment von ca. 2Kb nur dann in den Zellen nachweis-

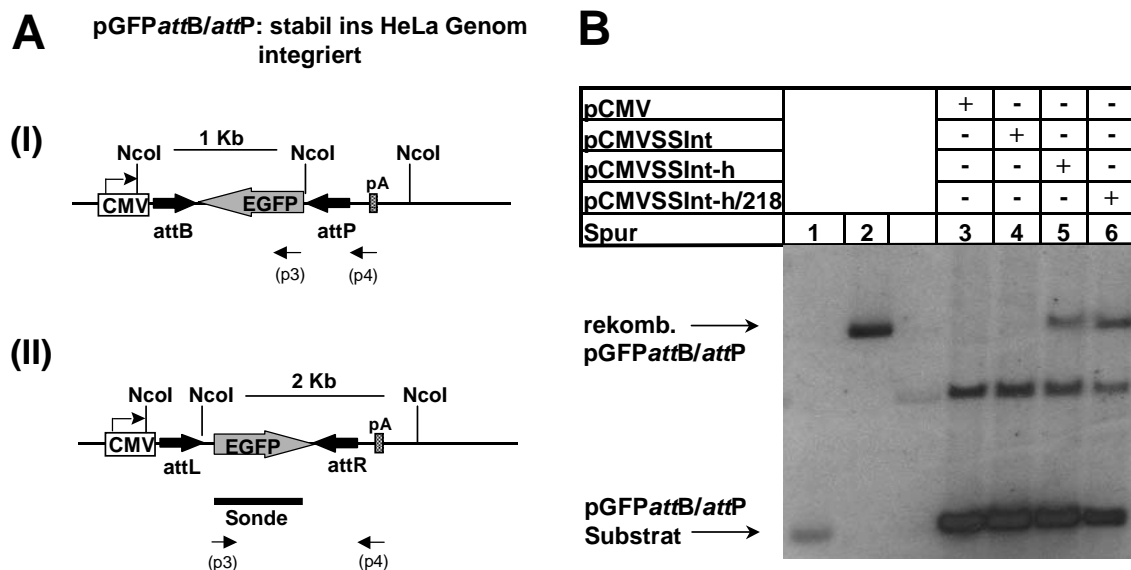


Abbildung D.6: Analyse der *Lambda* Mutanten-vermittelten genomischen Inversion zwischen *attP* und *attB*. (A) Schematische Darstellung des stabil ins HeLa Genom integrierten Inversionssubstrats pGFPattP/attB (I) und des entsprechenden Rekombinationsprodukts (II). Der CMV Promoter ist durch ein Rechteck und die Rekombinationssequenzen attB und attP sind durch dicke schwarze Pfeile dargestellt. Das EGFP Gen ist durch einen grau schattierten Pfeil dargestellt. pA, Polyadenylierungs-Sequenz. Primer für die PCR-Analyse sind als kleine Pfeile dargestellt. NcoI: Restriktionsschnittstelle. (B) Southern Analyse von integrativer Rekombination in der Reporterzelllinie H₈B 72h nach der Elektroporation der Int-Expressionsvektoren. Genomische DNA der Int-ausprägenden Reporterzelllinie wurde mit NcoI geschnitten und Rekombination mit einer EGFP-Sonde (A-(II)) analysiert. Spur 1 und 2, mit NcoI geschnittenes, nicht-rekombiniertes und rekombiniertes Plasmid pGFPattB/attP; Spur 3, NcoI geschnittene DNA aus pCMV transfizierten H₈B Zellen; Spur 4, Wildtyp-Int-ausprägende H₈B Zellen; Spur 5, Int-h-ausprägende H₈B Zellen; Spur 6, Int-h/218-ausprägende H₈B Zellen.

bar ist, wenn die Vektoren pCMVSSInt-h oder pCMVSSInt-h/218 (Abb. D.6-B Spur 5 und 6) transfiziert wurde. Eine quantitative Analyse mit dem Phosphoimager-System ergab, dass unter Berücksichtigung der Transfektionseffizienz von ca. 50%, 16% der integrierten pGFPattB/attP Moleküle durch Int-h rekombiniert wurden. In Int-h/218-ausprägenden H₈B-Zellen wurden sogar 30% der Moleküle rekombiniert, wogegen Rekombination durch die wild-typ Integrase zumindest im Rahmen der Nachweisgrenze für Southern Blot Analyse nicht nachgewiesen werden konnte (Spur 4). PCR Analysen mit den Primern (p3)/(p4) (Abb. D.6-A) zeigten jedoch, dass auch die wild-typ Integrase in der Lage ist wenige pGFPattB/attP Moleküle im Genom der Zelllinie H₈B zu rekombinieren (Daten nicht gezeigt).

Die zusätzliche Bande im Southern Blot (Abb. D.6-B), die in allen Ansätzen mit der H₈B-Zelllinie auftritt, stammt wahrscheinlich von der Integration eines unvollständigen Substrats in das Genom. Sie wurde in die Quantifizierungen als

Substratbande mit einbezogen, da sich auch auf diesem Fragment noch *attB* und *attP* befinden könnten und durch Rekombination von Int das gleiche NcoI-Fragment entstünde wie bei der Rekombination des vollständigen Substrats. Die Rekombinationsprodukte wurden zusätzlich durch PCR amplifiziert und ihre Integrität durch Sequenzierung überprüft (Daten nicht gezeigt). Die Rekombinationsanalysen in der Reporterzelllinie H₈B zeigen daher, dass mutierte Int Proteine in der Lage sind, die intramolekulare, integrative Rekombinationsreaktion auch an stabil integrierten genomischen Erkennungssequenzen zu katalysieren.

Die exzitive Rekombination zwischen *attL* und *attR* wurde in Reporterzelllinien untersucht, in denen das Substrat pGFP*attL/attR* an zufälligen Loci im HeLa Genom integriert ist. Die Ergebnisse der Rekombinationsanalysen in zwei verschiedenen Reporterzelllinien mit nur einer ins Genom integrierten Substratkopie zeigen, dass 72 Stunden nach der Elektroporation des Int-h-Expressionsvektors keine exzitive Rekombination durch PCR nachzuweisen war (Reporterzelllinie H₂L: Abb. D.7, Spur 2). Die Inversion des EGFP-Gens konnte erst nach weiteren 72 Stunden Int-h-Ausprägung durch Isolierung von DNA aus

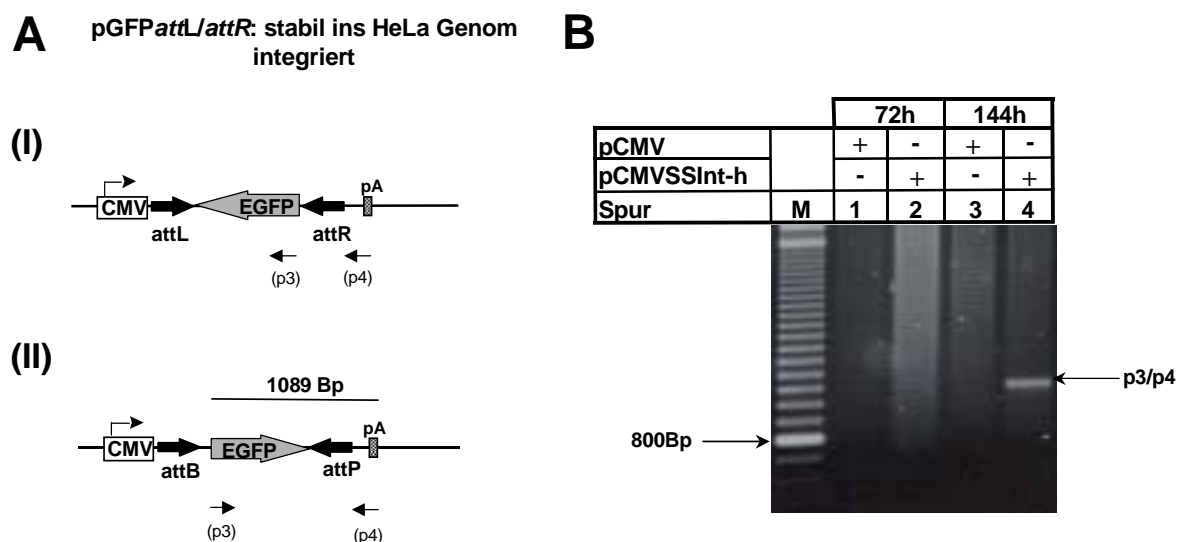


Abbildung D.7: Analyse der *Lambda* Mutanten-vermittelten genomischen Inversion zwischen *attL* und *attR*. (A) Schematische Darstellung des stabil ins HeLa Genom integrierten Inversions-Substrats pGFP*attL/attR* (I) und des entsprechenden Rekombinationsprodukts (II). (B) PCR-Analyse von exzistiver Rekombination in der Int-h ausprägenden Reporterzelllinie H₂L nach 72 und 144 Stunden (h). Mit den Primern p3 und p4 wird nur dann ein 1089Bp Fragment amplifiziert, wenn das EGFP-Gen invertiert vorliegt (A-(II)). Spur M, 100Bp Marker; Spur 1, Kontrolle ohne Int-h und Spur 2, Int-h-ausprägende H₂L Zellen nach 72 Stunden; Spur 3, Kontroll-Ansatz und Spur 4, Int-h Ansatz nach 144 Stunden.

der Reporterzelllinie H₂L, in Form einer schwachen PCR-Produktbande (Abb. D.7, Spur 4) nachgewiesen werden. Diese Ergebnisse zeigen, dass auch die genomische Rekombination zwischen *attL* und *attR* durch die Mutante Int-h katalysiert wird. Im Vergleich zur Int-h-katalysierten integrativen Rekombinationsreaktion in H₈B Zellen, ist die Effizienz der exzisen Rekombination in den beiden untersuchten Reporterzelllinien, die nur eine einzige Kopie des Substrats im Genom besitzen, deutlich niedriger. Obwohl nicht gezeigt, ist ähnliches für die Mutante Int-h/218 anzunehmen.

Die Inversion des EGFP-Gens durch integrative oder exzise Rekombination sollte zu dessen Ausprägung vom CMV Promoter (Abb. D.6-A (II) und D.7-A (II)) führen. Die HeLa Reporterzellen für die integrative und exzise Rekombination wurden mit den Expressionsvektoren für die Integrasen transfiziert und 72 Stunden nach der Elektroporation durchflusszytometrisch analysiert. Trotz mehrfacher Wiederholung dieser Experimente konnten keine EGFP-ausprägenden Zellen nachgewiesen werden (Daten nicht gezeigt). Die alternative Verknüpfung BOP an den *att*-Regionen wurde durch die gleichen PCR-Analysen, die in Abschn. D2.3 beschrieben sind, ausgeschlossen.

D3.2 Sequenz-spezifische Deletion eines Neomycin-Resistenzmarkers durch die Integrase in embryonalen Stammzellen der Maus

In den folgenden Experimenten wurde gezeigt, dass die Mutante Int-h/218 auch in embryonalen Stammzellen (ES) der Maus eingesetzt werden kann, um *att*-flankierte Resistenzmarkergene durch integrative oder exzise Rekombination aus dem Genom zu schneiden.

Dazu wurden entsprechende *Target* Vektoren mit *attB/attP* oder *attL/attR* kloniert, die durch homologe Rekombination (*Gene Targeting*) in den ubiquitär ausgeprägten ROSA26 Locus integriert wurden. Der ROSA26 Locus wurde als Integrationsort der Vektoren gewählt, weil er schon erfolgreich für die Generierung von Reporter-Linien zum Nachweis von Cre-Rekombination verwendet wurde (Soriano, 1999; Mao et al., 1999; 2001). Zur Konstruktion der *Target* Vektoren wurde ein Neomycin-Resistenzgen, das vom Phosphoglycerat-Kinase (PGK) Promoter transkribiert wird, entweder von *attB* und *attP* oder *attL* und *attR* in direkter Orientierung flankiert. Um die Transkription vom PGK Promoter auf das Neomycin-Resistenzgen zu beschränken wurde ein Stoppsignal am 3'-Ende des

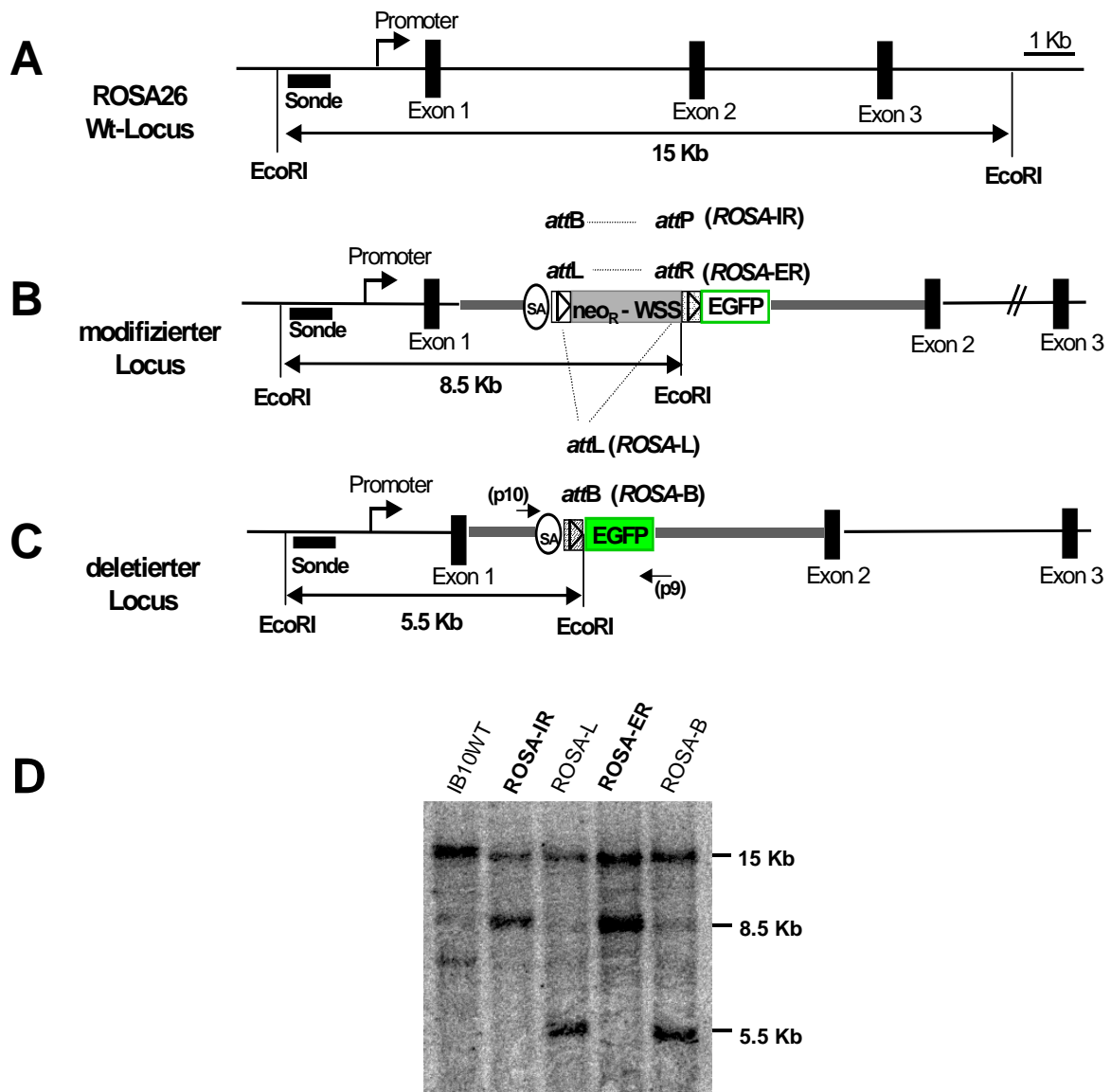


Abbildung D.8: Genetische Manipulation in embryonalen Stammzellen der Maus durch mutierte *Lambda* Integrasen. (A, B) Erzeugung von Reporter ES-Zelllinien für die integrative (*attBxattP*) und die exzusive (*attLxattR*) Rekombination durch *Gene Targeting*. Eine Rekombinationskassette mit dem Neomycinresistenz-Gen (*neo_R*) und einem Transkriptions-Stopp (WSS), die entweder von *attB* und *attP* oder *attL* und *attR* in direkter Wiederholung flankiert sind, wurde in das erste Intron des ROSA26 Locus von IB10 Zellen integriert. Gleichzeitig mit der Rekombinationskassette wurde an dessen 5'-Ende eine 3'-Spleißstelle (SA) und an dessen 3'-Ende ein promoterloses EGFP integriert. Die homologe Sequenz des *Target* Vektors im ersten Intron ist durch eine dicke graue Linie hervorgehoben. Die Reporterzelllinien für die integrative Rekombination werden als ROSA-IR bezeichnet und die für die exzusive als ROSA-ER. (C) Rekombination durch die Int Mutanten schneidet das Neomycin-Resistenzgen mit dem Transkriptions-Stopp aus, wodurch EGFP vom endogenen ROSA26 Promoter ausgeprägt wird. Ebenso markiert sind die Positionen der Schnittstellen für das Restriktionsenzym EcoRI, die Länge der erwarteten genomischen DNA-Fragmente nach Inkubation mit EcoRI und die Position der Sonde, die in Southern Analysen verwendet wurde. Primer für die PCR-Analyse sind als kleine Pfeile dargestellt (D) *Southern Blot* Analyse von genomischer DNA. Die DNA wurde mit EcoRI inkubiert, in einem 0,8% Agarosegel aufgetrennt und durch *Southern blotting* mit einer 5' externen Sonde (siehe A, B und C) analysiert. IB10WT, genomische DNA aus Wildtyp IB10 Zellen als Kontrolle; ROSA-IR und ROSA-ER, genomische DNA von Reporter-Zelllinien für integrative und exzusive Rekombinationsreaktionen; ROSA-L und ROSA-B, genomische DNA aus Zellen, die aufgrund ihrer GFP-Ausprägung durch FACS sortiert wurden.

Neomycingens plaziert. Auf beiden *Target* Vektoren befindet sich zudem am 3'-Ende der Rekombinationskassetten ein offenes Leseraster für EGFP. Die *Target* Vektoren mit der Rekombinationskassette für die integrative oder exzusive Rekombinationsreaktion wurden jeweils durch homologe Rekombination in das erste Intron des ROSA26 Locus integriert (Abb. D.8-A und B; Abschn. C3.3). Im Genom der daraus hervorgegangenen Reporter ES-Zellen befindet sich daher das *att*-Sequenzen flankierte Neomycin-Resistenzgen gefolgt von dem EGFP-Gen. (Abb. D.8-B). Die Reporter ES-Zelllinien für die integrative Rekombination wurden als ROSA-IR und die für die exzusive Rekombination als ROSA-ER bezeichnet. Rekombination durch Int-h/218 schneidet das Resistenzgen aus dem Genom in beiden Reporterzelllinien aus und führt zur Ausprägung des EGFP-Gens (Abb. D.8-C). Die Ausprägung von EGFP erfolgt von einer durch den ROSA26 Promoter ubiquitär erzeugten mRNA (Soriano, 1999; Mao et al., 1999), auf der die künstlich integrierte 3'-Spleißstelle (SA) mit der 5'-Spleißstelle des ersten Exons verknüpft ist (Abb. D.8-C). Homolog rekombinante Reporter-ES Zelllinien wurden mit G418 selektioniert, deren genomische DNA mit EcoRI fragmentiert und durch *Southern Blotting* mit einer 5' externen Sonde analysiert (Abb. D.8; Mao et al., 1999).

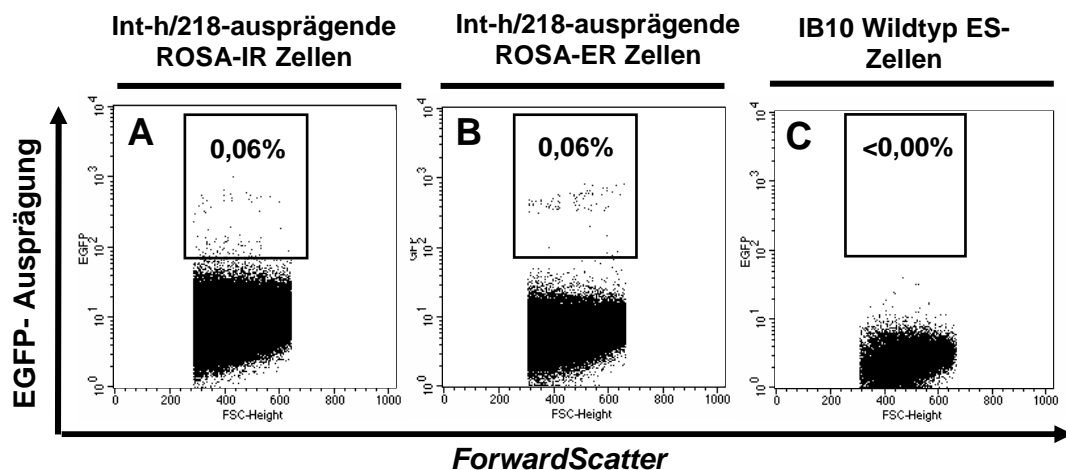


Abbildung D.9: Isolierung von EGFP-ausprägenden ROSA-IR oder ROSA-ER Reporter ES-Zellen mit FACS nach Transfektion von pCMVSSInt-h/218. Reporter ES-Zellen wurden mit dem CMV-Expressionsvektor für Int-h/218 transfiziert. 48h nach der Transfektion wurden die EGFP-ausprägenden ES-Zellen mit FACS isoliert und in geringer Dichte auf einem Mitomycin C-inaktivierten MEF Zellrasen in LIF-haltigem ES-Medium kultiviert. Von diesen wurden einzelne ES-Zell Kolonien gepickt und expandiert. **(A)** Isolierung von EGFP-ausprägenden ROSA-IR und **(B)** ROSA-ER ES-Zellen mit den in der Abbildung angegebenen *gates*. Die Reinheit der sortierten Zellen betrug der Reanalyse zufolge in (A) 91% und in (B) 98%. In beiden Sortierungen wurden tote Zellen mit dem Farbstoff Propidium Iodid ausgeschlossen. **(C)** FACS Analyse von IB10 Wildtyp ES-Zellen.

Ein repräsentativer *Southern Blot* in Abbildung D.8-D zeigt jeweils ein Beispiel für die Analyse der ROSA-IR und ROSA-ER Zelllinien. In beiden Zelllinien wurden die erwarteten genomischen Restriktionsfragmente nachgewiesen: ein 15Kb-Fragment das vom wild-typ ROSA26 Locus des einen Allels und ein 8,5Kb-Fragment, das vom homolog rekombinierten ROSA-Locus des zweiten Allels stammt (vergleiche Abb. D.8-A und -B). Die Integrationsfrequenz der Vektoren durch homologe Rekombination lag zwischen 6,5 und 12%. Illegitime Integration der *Target* Vektoren wurde mit einem 1Kb-Fragment aus dem Vektor pROSA26-1 (Abschn. C1.8) als interne Probe in Kombination mit einem EcoRI-Verdau ausgeschlossen (Daten nicht gezeigt).

Zur Deletion des Resistenzmarkergens wurden ungefähr 3×10^6 Zellen von jeder Reporterzelllinie mit dem CMV-Expressionsvektor für Int-h/218 (Abschn. C1.9.1) transfiziert. Die Transfektionseffizienz, die für jedes Experiment einzeln ermittelt wurde, betrug zwischen 60 und 70% (Daten nicht gezeigt). 48 Stunden nach der Transfektion wurden die EGFP-ausprägenden Zellen mit FACS isoliert. Der Anteil an sortierten Zellen betrug zu diesem Zeitpunkt in beiden Ansätzen ungefähr 0,06% von der gesamten analysierten Zellpopulation (Abb. D.9). Auch in Rekombinationsanalysen, in denen keine Sortierung durchgeführt wurden, lag der Anteil an EGFP-ausprägenden Zellen immer zwischen 0,01% und 0,1% (Daten

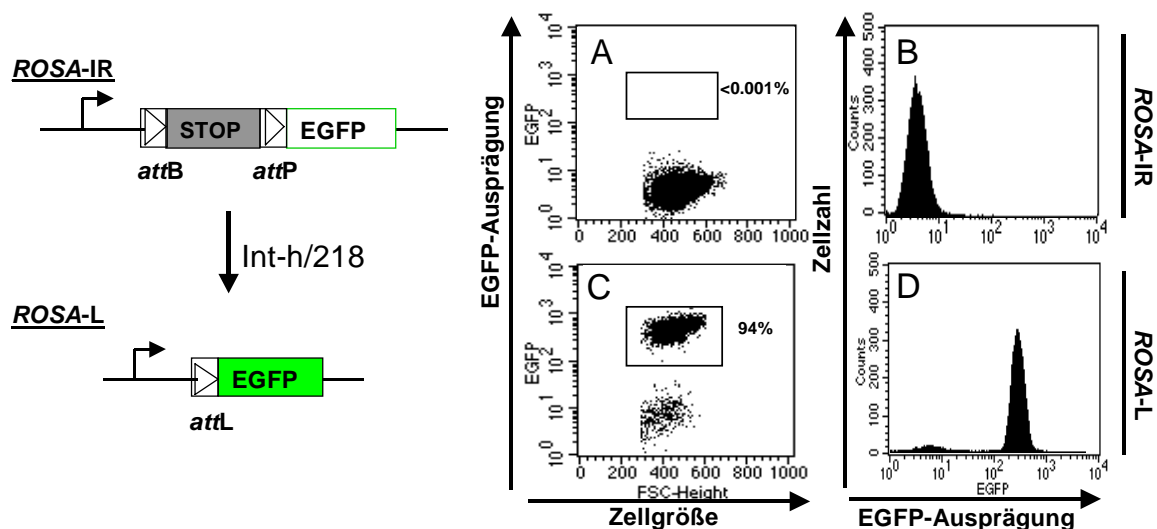


Abbildung D.10: FACS Analyse der EGFP-Ausprägung. (A, B) 5×10^4 Zellen der Reporter-Zelllinie ROSA-IR für integrative Rekombination wurden auf ihre EGFP-Ausprägung hin untersucht. (C, D) Die gleich Anzahl an Zellen wurden von einer expandierten EGFP-ausprägenden ES Zelle analysiert, die von einer einzelnen durch Int-h/218 rekombinierten ROSA-IR Zelle abstammt. Der geringe Anteil an Zellen die keine EGFP-Ausprägung in (C, D) aufweisen, stammen wahrscheinlich vom MEF Zellrasen auf dem die ES-Zellen kultiviert wurden.

nicht gezeigt). Die isolierten ES-Zellen wurden in geringer Dichte weiter auf einem MEF-Zellrasen kultiviert und einzelne ES-Zell Kolonien gepickt und expandiert. Eine vergleichende FACS-Analyse zwischen nicht-transfizierten ROSA-IR Zellen (Abb. D.10-A und -B) und expandierten ROSA-L Zellen, die von einer einzelnen transfizierten und sortierten ROSA-IR Zelle abstammen (Abb. D.10-C und -D), zeigt, dass die zuletzt genannten Zellen im hohen Maß EGFP ausprägen.

Auch vergleichende Fluoreszenz-Mikroskopische Analysen bestätigen, dass die Ausprägung von EGFP nur in expandierten ROSA-B Zellen nachzuweisen war (Abb. D.11-C und -D), die von einer einzelnen transfizierten ROSA-ER Zelle abstammen (Abb. D.11-A und -B).

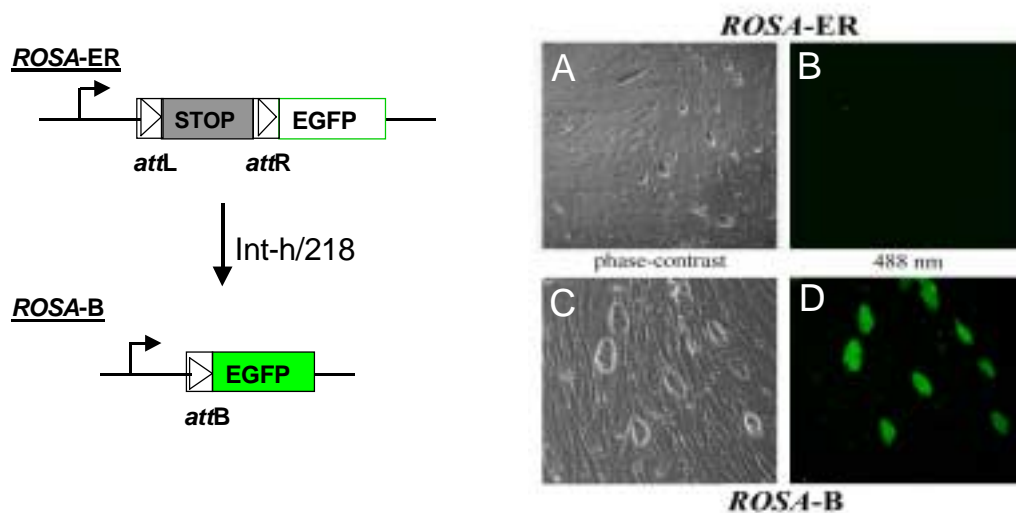


Abbildung D.11: Fluoreszenz-Mikroskopie vor und nach Rekombination. (A, B) Individuelle Zell-Klone von der Reporter-Zelllinie ROSA-ER. (C, D) Zell-Klone von ROSA-ER Zellen nach sequenz-spezifischer Rekombination und anschließender Sortierung von EGFP-ausprägenden Zellen.

Diese Ergebnisse zeigen, dass die Ausprägung von Int-h/218 in ROSA-IR und ROSA-ER zur Synthese von EGFP führt und sequenz-spezifische Rekombination durch die Integrase Mutante stattgefunden hat. Durch *Southern Blot* Analyse von genomischer DNA konnte die Deletion des Resistenz-Markergens auch direkt im Genom nachgewiesen werden. Dazu wurde genomische DNA von den expandierten ES Zell Klonen isoliert, mit EcoRI inkubiert und, wie zuvor beschrieben, durch *Southern Blotting* analysiert (Abb. D.8-A bis -C).

Der *Southern Blot* in Abbildung D.8-D zeigt, dass ROSA-L und ROSA-B Zellen die beiden erwarteten genomischen EcoRI-Resistriktionsfragmente aufweisen: ein 15Kb-Fragment, das aus dem wild-typ Locus stammt, und ein 5,5Kb-Fragment das

durch Deletion des Resistenzgens entstanden ist. Die Sequenzierung von PCR-Produkten, die mit genomischer DNA von diesen Zellen erzeugt wurden, bestätigten, dass die Rekombination durch Int-h/218 sequenz-spezifisch war und die Produkte die erwarteten Nucleotidsequenzen hatten.

D4 Die Transfektion von IHF-DNA Komplexen stimuliert die Rekombinationsaktivität der wild-typ Integrase in Säugetierzellen

Für eine effiziente, durch die wild-typ Integrase katalysierte, integrative Rekombination *in vitro* und in *E.coli* ist sowohl der Protein Kofaktor IHF als auch eine negativ gespannte *attP*-tragende DNA Sequenz unerlässlich (Better et al., 1982; Richet et al., 1986; Christ & Dröge, 1999). Die wild-typ Integrase besitzt im Gegensatz zu den Mutanten Int-h und Int-h/218 nur eine marginale Rekombinationsaktivität in eukaryotischen Zellen. Dies lässt darauf schließen, dass in eukaryotischen Zellen keine oder nur ungenügende Mengen an Kofaktoren vorhanden sind. Eine interessante Frage in diesem Zusammenhang ist, ob die minimal vorhandene integrative Rekombinationsaktivität der wild-typ Integrase in HeLa Zellen durch das prokaryotische Hilfsprotein IHF gesteigert werden kann.

Es wurde zuerst untersucht, ob die minimale Rekombinationsaktivität der wild-typ Integrase durch die Transfektion von negativ gespanntem, und mit gereinigtem IHF Protein vorinkubierten, Substratvektor stimuliert wird.

Dazu wurden HeLa Zellen zunächst entweder mit dem CMV-Expressionsvektor für die wild-typ Integrase oder für Int-h transfiziert. 3 bis 4 Stunden nach der Elektroporation der Expressionsvektoren wurde dann das Deletionssubstrat für die intramolekulare Rekombinationen zwischen *attP* und *attB*, oder die entsprechenden Substrate für die intermolekularen, integrativen Rekombinationsreaktionen (Abb. D.12-A), eine ½ Stunde in einem Ansatz mit und in einem zweiten ohne gereinigtem IHF Protein vorinkubiert. Anschließend wurden sowohl die Protein-DNA Gemische als auch die Protein-freien Kontrollansätze mit Hilfe von FuGENE™6 (Roche) in die Int-transfizierten HeLa Zellen gebracht. Der Anteil an Zellen, die EGFP ausprägten, wurde 48 Stunden später ermittelt und zwischen den Ansätzen verglichen (Abb. D.12-B). Die Ergebnisse von drei Experimenten zeigen, dass die intramolekulare Rekombination durch die wild-typ Integrase in Gegenwart

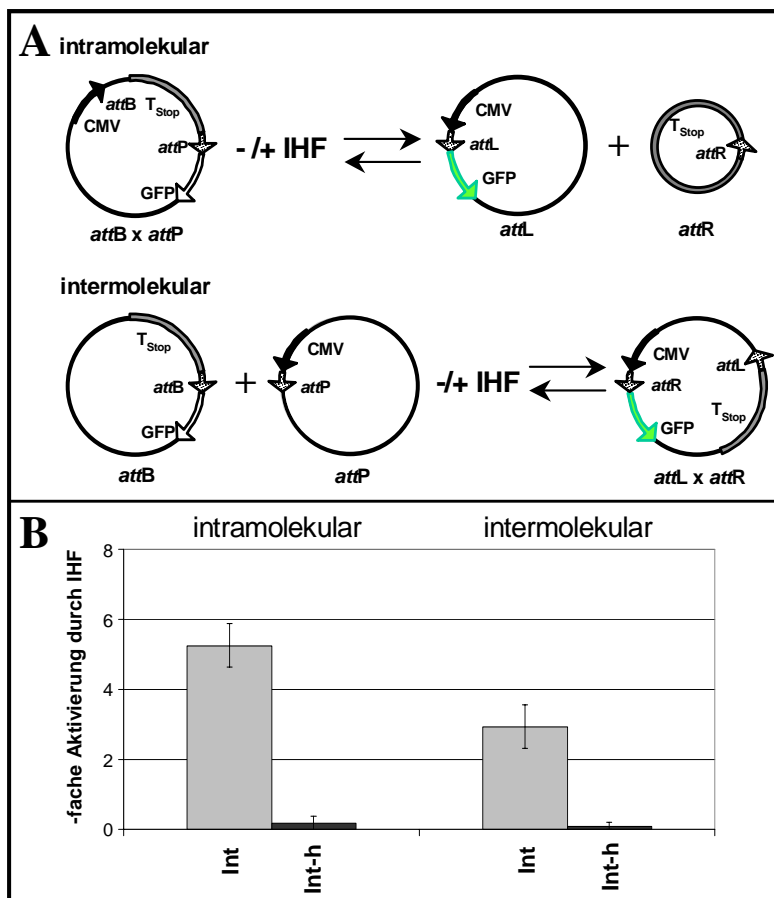


Abbildung D.12: Gereinigtes IHF Protein stimuliert die intra- und intermolekularen integrativen Rekombinationsreaktion der wild-typ Int in HeLa Zellen. (A) Schematische Darstellung von Substratvektoren, die vor der Transfektion in wild-typ Int- oder Int-h-ausprägenden HeLa Zellen, zusammen mit oder ohne IHF Protein vorinkubiert wurden. **(B)** 48 Stunden nach der Transfektion wurde der Anteil an EGFP-ausprägenden Zellen mit FACS analysiert. Das Verhältnis dieser Anteile ist in Form von Aktivierung der Rekombinationsaktivität durch IHF dargestellt. Die Balken zeigen die Mittelwerte aus drei Analysen, die vertikalen Linien innerhalb der Balken die zugehörigen Standardabweichungen.

des IHF Proteins im Durchschnitt bis zu 5-fach stimuliert wurde. Der Anteil an EGFP-positiven Zellen stieg beispielsweise in einem Experiment von 1% in Abwesenheit von IHF auf bis zu 6% in dessen Gegenwart. Die Stimulierung des gereinigten IHF Proteins auf die intermolekulare Rekombination war etwas niedriger, etwa 3-fach.

Die Stimulation der integrativen Rekombination durch IHF über einen Zeitraum von 48 Stunden wurde nur in Ansätzen mit der wild-typ Integrase beobachtet. Die Rekombinationsaktivität von Int-h über diesen Zeitraum (Abb. D.12-B), sowie die Transfektionseffizienzen (Daten nicht gezeigt), wurde dagegen von IHF nicht beeinflusst.

D4.1 Die Transfektion von Expressionsvektoren für die IHF Untereinheiten in HeLa Zellen stimuliert nicht die Aktivität der wild-typ Integrase

Die Transfektion von gereinigtem IHF Protein ermöglicht nur ein begrenztes Vorhandensein des Hilfsproteins in eukaryotischen Zellen. Um größere Mengen an IHF auch über einen längeren Zeitraum in den Zellen zur Verfügung zu stellen,

wurden PGK-Expressionsvektoren für die Peptide HimA (\cong IHF- α) und HimD (\cong IHF- β) konstruiert (Abschn. C1.9.1). Mit episomalen Rekombinationsanalysen (Abschn. D2.1) wurde untersucht, ob die intramolekulare, integrative Rekombinationsreaktion der Integrasen auch durch die gleichzeitige Kotransfektion der Expressionsvektoren für HimA und HimD stimuliert werden kann. Die Ergebnisse zeigten jedoch, dass über einen Zeitraum von 48h weder eine Stimulierung der Rekombinationsaktivität durch die wild-typ Integrase noch durch die Mutante Int-h nach der Kotransfektion der Him-Expressionsvektoren nachgewiesen werden konnte (Daten nicht gezeigt). Die Ergebnisse einer *Western Blot* Analyse von HeLa Zellen, die mit den Expressionsvektoren für HimA und HimD transfiziert wurden, ergaben außerdem, dass keine Proteine mit den erwarteten Molekulargewichten nachgewiesen werden konnten (Daten nicht gezeigt).

D5 Optimierung des prokaryotischen Hilfsproteins IHF zum Einsatz in eukaryotischen Zellen

Die Ergebnisse der im Abschnitt D4.1 beschriebenen Experimente machen deutlich, dass die Kotransfektion von eukaryotischen Zellen mit Expressionsvektoren für die IHF Untereinheiten HimA und HimD zu keiner Stimulierung der Int-vermittelten sequenz-spezifischen Rekombination führt. Da auch die Ausprägung des HimA (IHF- α) und/oder HimD (IHF- β) Proteins in HeLa nicht durch *Western Blot* Analysen nachgewiesen werden konnte, liegt die Vermutung nahe, dass mit dieser experimentellen Strategie HimA und HimD kein funktionelles IHF bilden. Der Einsatz von IHF in eukaryotischen Zellen erforderte daher eine Optimierung des prokaryotischen Proteins, beispielsweise durch Ausprägung eines IHF-ähnlichen Proteins in Form eines Monomers. Dazu wurden vier verschiedene sogenannten *one-chain Integration Host Factors* (oIHF) generiert. Die oIHF Proteine sind Hybride aus den beiden IHF-Untereinheiten IHF- α und IHF- β . Sie wurden erzeugt, indem die Untereinheit α (AS L3 bis A94) zwischen die AS Q39 und G40 der Untereinheit β inseriert wurde (Abb. D.13-A; Abschn. C1.9.2). Die Übergänge zwischen der Untereinheit β und der integrierten Untereinheit α sind durch Aminosäure-Linker verbunden, die in den einzelnen oIHF Varianten eine unterschiedliche Länge besitzen (Abb. D.13-A; die Nucleotidsequenz ist im Anhang beschrieben). Der AS-Rest L3 am N-Terminus

A

1	M	G	T	K	S	E	L	I	E	R	L	A	T	Q	Q	S	H	I	P	A	K	T	V	E	D	A	V	K	E	M	L	E	H	M	A	S	T	L	A	Q
	AS-Linker 2																																							
41	G	G	S	G	G	L	T	K	A	E	M	S	E	Y	L	F	D	K	L	G	L	S	K	R	D	A	K	K	L	V	E	L	F	F	G	G	I	R	R	A
	AS-Linker 1																																							
81	L	E	N	G	E	Q	V	K	L	S	G	F	G	N	F	D	L	R	D	K	N	Q	R	P	G	R	N	P	K	T	G	E	D	I	P	I	T	A	R	R
	AS-Linker 4																																							
121	V	V	T	F	R	P	G	Q	K	L	K	S	R	V	E	N	G	G	G	E	R	I	E	I	R	G	F	G	S	F	S	L	H	Y	R	A	P	R	T	
	AS-Linker 3																																							
161	G	R	N	P	K	T	G	D	K	V	E	L	E	G	K	Y	V	P	H	F	K	P	G	K	E	L	R	D	R	A	N	I	Y	G	*					

oIHF1 \equiv AS-Linker 1 und 3	oIHF3 \equiv AS-Linker 1 und 4
oIHF2 \equiv AS-Linker 2 und 4	oIHF4 \equiv AS-Linker 2 und 3

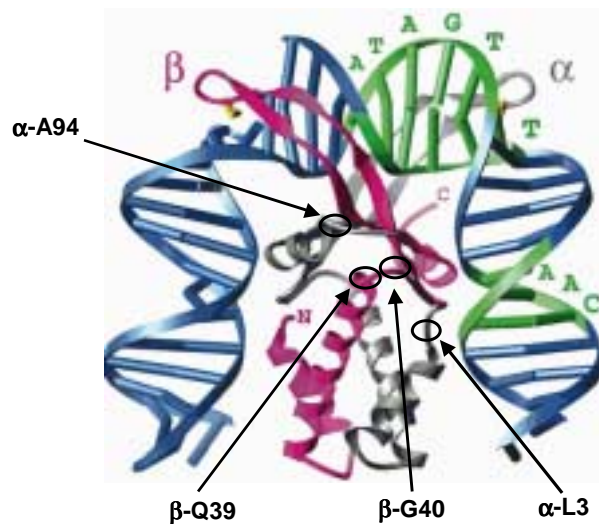
B

Abbildung D.13: Die *one-chain Integration Host Factor* (oIHF) Varianten 1-4. In (A) ist die AS-Sequenz der vier oIHF Varianten dargestellt. Die fettgedruckten AS stellen die Linker zwischen der in IHF- β (pink) integrierten IHF Untereinheit α (grau) dar. Die unterschiedliche Länge der AS-Linker ist durch farbige Rechtecke angegeben. Die unterstrichenen AS kennzeichnen die verknüpften AS-Reste der beiden Untereinheiten. Das schwarz gedruckte G2 wurde durch die PCR-Klonierung in allen Varianten erzeugt. (B) Die Kristall-Struktur eines wild-typ IHF-DNA Komplexes (modifiziert nach Rice et al., 1996). Die Untereinheit

IHF- α ist in grau dargestellt, IHF- β in pink. Die Konsensus DNA-Bindungssequenz für IHF ist grün hervorgehoben und besitzt fast ausschließlich Kontakte zur langen flexiblen β -Schleife von IHF- α und zum Proteinkern von IHF- β . Die Pfeile und Kreise markieren die AS-Reste, die im oIHF durch die AS-Linker verknüpft wurden und entsprechen den unterstrichenen in (A). Der Abstand zwischen dem Q39 der β -Kette zum L3 der α -Kette beträgt 11,1 Angström (AS-Linker 1 und 2). Zwischen A94 der Kette α und dem G40 der Kette- β besteht ein 5,0 Angström Abstand (AS-Linker 3 und 4).

von IHF- β ist mit dem AS-Rest Q39 von IHF- α im oIHF1 durch 4 AS-Reste (AS-Linker 1) und A94 am C-Terminus von IHF- β ist mit G40 von IHF- α durch ein Glycin (AS-Linker 3) verbunden (Abb. D.13). Im oIHF2 befindet sich anstelle des Linkers 1 mit den vier Resten, der AS-Linker 2 mit fünf Resten und der Linker 3 ist durch den AS-Linker 4 mit zwei Glycinresten ersetzt. Im oIHF3 befinden sich an den entsprechenden Positionen die AS-Linker 1 und 4 und im oIHF4 die AS-Linker 2 und 3 (Abb. D.13-A).

D5.1 oIHF-Ausprägung in *E.coli*

Zum Nachweis von oIHF Proteinen in Bakterien wurde der *E.coli* Stamm CSH26 Δ IHF, bei dem die Gene *himA* und *himD* durch Transposons zerstört sind (Abschn. B1; Abb. D.14-B Spur 1 bis 8), mit prokaryotischen Expressionsvektoren für oIHF1 bzw. oIHF2 transformiert. Die Expressionsvektoren wurden konstruiert, indem oIHF1 oder oIHF2 (Abb. D.13-A) durch PCR in den Expressionsvektor pTrc99a der Firma Pharmacia Biotech (Freiburg) inseriert wurden. Die Transkription der Gene erfolgt vom *trc* Promoter und wird durch den lac Repressor reguliert, der von einem *lacI^q* Gen auf demselben Vektor kodiert wird (Abb. D.14-A). Das negative Kontroll-Plasmid pTrc entspricht dem Plasmid pTrc99a, dem eine oIHF-kodierende Region fehlt. Von den transformierten Bakterien wurden resistente Klone gepickt und über Nacht bei 37°C kultiviert. Die Ansätze wurden 1:150 verdünnt, bei einer OD₆₀₀ von ~0,3 geteilt und in einer Hälfte die Ausprägung

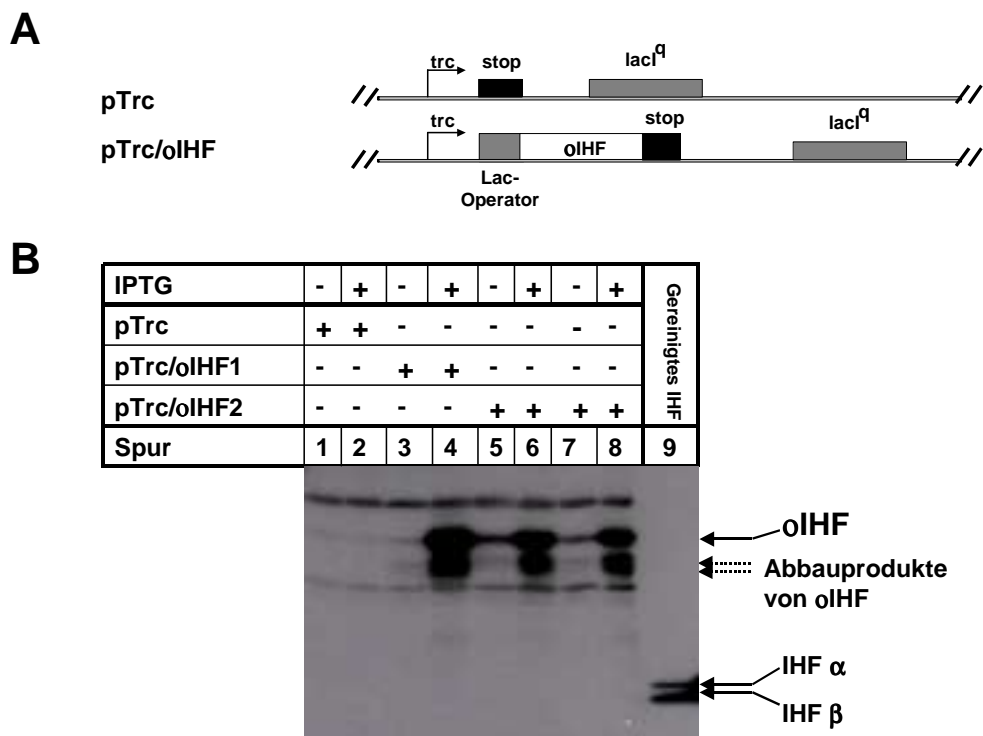


Abbildung D.14: Ausprägung von oIHF in CSH26 Δ IHF. In **(A)** sind prokaryotische Expressionsvektoren für oIHF dargestellt. *Trc*, *trc*-Promoter (Amann et al., 1988); oIHF, *one-chain Integration Host Factor*; *lacI^q*, Lac Repressor. **(B)** *E.coli* CSH26 Δ IHF wurde mit den oIHF-Expressionsvektoren transformiert und übernacht kultiviert. Aus den ÜNK wurden Bakterien überimpft, bis zu einer OD₆₀₀ von ~0,3 kultiviert, die Ansätze geteilt und mit (+) oder ohne (-) IPTG 7h bei 37°C inkubiert. oIHF1 und oIHF2 wurden dann durch *Western Blot* Analyse mit einem Anti-Serum gegen IHF den in Zell-Lysaten der induzierten und nicht-induzierten Bakterien nachgewiesen. Spur 7 und 8 ist eine unabhängige Reproduktion von Spur 5 und 6. IHF α , Untereinheit α des IHF Proteins; IHF β , Untereinheit β des IHF Proteins.

von oIHF mit 2mM IPTG induziert. Der Nachweis der oIHF Proteine erfolgte 7h nach der Induktion mit IPTG durch *Western Blot* Analysen. Dazu wurden Zell-Lysate mit einem Anti-Serum gegen IHF (S.D. Goodman, Los Angeles, USA) untersucht. Ein repräsentativer *Western Blot* in Abbildung D.14-B zeigt eine Analyse mit Zell-Lysaten von induzierten und nicht-induzierten transformierten CSH26 Δ IHF Bakterien. Ein Protein mit dem erwarteten molekularen Gewicht von ca. 20 KDa konnte in den Bakterien nachgewiesen werden, die einen der oIHF kodierenden Vektoren enthalten (Abb. D.14 Spur 3 bis 8). Dabei war die Ausprägung des 20 KDa Proteins in Bakterien, in denen die Transkription mit IPTG induziert wurde, deutlich höher (Spur 4, 6 und 8) als in den entsprechenden nicht-induzierten Ansätzen (Spur 3, 5 und 7). In *E.coli* CSH26 Δ IHF Bakterien, die mit dem Vektor pTrc transformiert wurden, konnte dagegen weder in induzierten (Spur 2) noch in nicht-induzierten (Spur 1) Ansätzen ein Protein mit diesem Molekulargewicht nachgewiesen werden. Als Kontrolle und Größenmarker wurde gereinigtes IHF Protein (Invitrogen) eingesetzt (Spur 9).

Diese Experimente zeigen, dass sowohl oIHF1 als auch oIHF2 in CSH26 Δ IHF Bakterien ausgeprägt werden. Im folgenden wurde untersucht, ob diese prokaryotisch ausgeprägten IHF-ähnlichen Proteine auch die integrative Rekombinationreaktion der wild-typ Integrase in *E.coli* stimulieren können.

D5.2 Analyse von enzymatischen Aktivitäten der vier unterschiedlichen oIHF Varianten in *E.coli*

Zur Ausprägung der oIHF Varianten 3 und 4 in *E.coli*, wurden entsprechende prokaryotische Expressionsvektoren kloniert (Abschn. C1.9.2). Um die Funktionalität der oIHF1-4 Proteine in *E.coli* zu testen, wurden Rekombinationsanalysen im wild-typ IHF defizienten *E.coli* Stamm CSH26 Δ IHF mit der wild-typ Integrase durchgeführt. Dazu wurde der Vektor pNC9WT konstruiert (Abschn. C1.10). Auf ihm sind *attB* und *attP* in direkter Wiederholung angeordnet. Dieser Vektor diente gleichzeitig auch als Expressionsvektor für die wild-typ Integrase, dessen offenes Leseraster sich auf dem Vektor befindet und einschließlich ihres *trc*-Promoters von den *att*-Sequenzen flankiert wird (Abb. D.15-A). Rekombination durch die wild-typ Integrase findet nur in Gegenwart von IHF oder eines IHF-

ähnlichen Hilfsproteins statt und führt zum Ausschneiden der Int-Expressionskassette, was wiederum das Restriktionsmuster der Restriktase Aval ändert.

CSH26 Δ IHF wurde mit pNC9WT und jeweils einem der vier oIHF-Expressionsvektoren (Abb. D.14-A) kotransformiert, einzelne transformierte Klone bis zu einer OD₆₀₀ von 0,4-0,6 kultiviert, der Ansatz geteilt und in einer Hälfte die oIHF-Ausprägung mit 1 mM IPTG induziert. Die Plasmid-DNA wurde ca. 5-6h nach Induktion isoliert, zur Analyse mit Aval geschnitten und in einem Agarosegel elektrophoretisch aufgetrennt. Eine repräsentative Restriktions-Analyse von isolierter Plasmid-DNA in Abbildung D.15-B zeigt, dass ein durch integrative

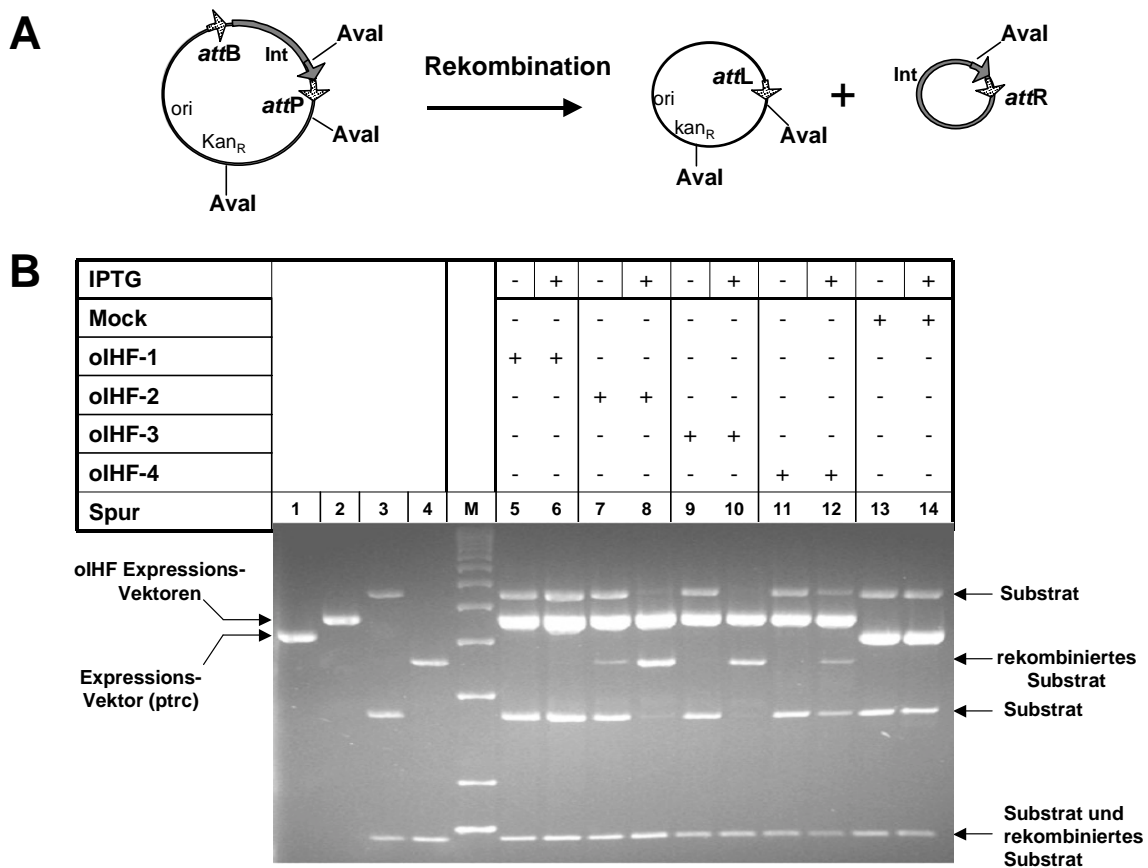


Abbildung D.15: Stimulierung intramolekularer integrativer Rekombination der wild-typ Integrase in CSH26 Δ IHF durch verschiedene oIHF Varianten. In (A) ist schematisch der Vektor pNC9WT dargestellt, der zugleich Substrat ist für die integrative Rekombination (*attB* und *attP* in direkter Wiederholung sind als helle Pfeile dargestellt) und Expressionsvektor für die wild-typ Integrase (grauer Pfeil). Rekombination durch Int führt zur Deletion der Int-Expressions-Kassette. Aval: Restriktionsschnittstelle. *E.coli* CSH26 Δ IHF wurde mit pNC9WT und einem Expressionsvektor für oIHF 1, 2, 3 oder 4 ko-transformiert. pTrc diente als negative Kontrolle (Mock). Einzelne transformierte Klone wurden bis zu einer OD₆₀₀ von 0,4-0,6 kultiviert, der Kulturansatz geteilt und eine Hälfte mit 1 mM IPTG inkubiert. 5-6h nach Induktion wurde die Plasmid-DNA isoliert, mit Aval geschnitten und in einem Agarosegel aufgetrennt (B). Spur M: Größenmarker; Spur 1: mit Aval geschnittener pTrc Vektor; Spur 2: geschnittener Expressionsvektor für oIHF; Spur 3: Aval-Fragmente von pNC9WT und Spur 4: von rekombiniertem pNC9WT.

Rekombination entstandenes Aval-Fragment nur dann in CSH26 Δ IHF nachzuweisen ist, wenn diese mit einem Expressionsvektor für die oIHF-Varianten 2-4 kotransfiziert wurden (Spur 7, 8, 10 und 12). In *E.coli*, die mit dem Kontrollvektor pTrc kotransformiert wurden oder oIHF1 ausprägten, konnte dagegen keine Rekombination nachgewiesen werden (Spur 13-14 und 5-6). Während es in oIHF2-ausprägenden Bakterien auch ohne Induktion zur Rekombination kommt (Spur 7), wurde durch die Zugabe von IPTG das Substrat fast vollständig rekombiniert (Spur 8). Quantitative Analysen der Aval-Fragmente ergab, dass ~17% des Substrates vor der Induktion rekombiniert wurden. Die annähernd vollständige Rekombination des Substrates nach IPTG-Zugabe ist auf eine gesteigerte oIHF2-Ausprägung zurückzuführen (siehe *Western Blot* in Abb. D.14-B: vergleiche Spur 5 mit 6 und Spur 7 mit 8). Im Gegensatz dazu wurden in oIHF3- und oIHF4-ausprägenden *E.coli* das Substrat nur nach Zugabe von IPTG durch die wild-typ Integrase rekombiniert (vergleiche Spur 9 mit 10 und Spur 11 mit 12). Während in induzierten oIHF3-ausprägenden Bakterien das Substrat, vergleichbar mit dem induzierten oIHF2 Ansatz, fast komplett rekombiniert wurde, stimuliert induziertes oIHF4 die Rekombination nur zu 44,4%.

Diese Ergebnisse zeigen, dass nur die oIHF-Varianten 2-4 die integrative Rekombinationreaktion der wild-typ Integrase in einem IHF-defizienten *E.coli* Stamm stimulieren. Die größte Aktivität zeigte dabei oIHF2, das offensichtlich auch schon in Gegenwart von geringen Mengen die Rekombination stimuliert. oIHF1 ist dagegen inaktiv.

D5.3 Stimulierung der integrativen Rekombinationsaktivität von Int-h/218 durch Ausprägung von oIHF2 in ROSA-IR

Das oIHF Protein wurde zum Einsatz in eukaryotischen Zellen generiert. Bisherige Analysen zeigen, dass es in *E.coli* ausgeprägt wird und die integrative Rekombinationsaktivität der wild-typ Integrase stimuliert. Die Ergebnisse der Rekombinationsanalysen zeigen außerdem, dass das oIHF2 Protein am stärksten die Rekombination stimuliert. Um dessen Funktionalität auch in eukaryotischen Zellen zu untersuchen wurde ein CMV-Expressionsvektor für oIHF2 und oIHF2HIS kloniert (Abb. D.16-A). Das oIHF2HIS Protein besitzt am C-Terminus 6 zusätzliche Histidin-Reste, die zur Aufreinigung dieses Proteins aus *E.coli* angehängt wurden.

Diese Expressionsvektoren wurden jeweils mit dem CMV-Expressionsvektor für Int-h/218 in ROSA-IR, der Reporter-ES-Zelllinie für genomische integrative Rekombinationsreaktionen, kotransfiziert. Als negative Kontrolle wurde der Vektor pCMV (Abb. D.16-A) kotransfiziert. 24 Stunden nach der Elektroporation wurden Zell-Lysate von den transfizierten ES-Zellen hergestellt und die oIHF Proteine in *Western Blot* Analysen mit einem Anti-Serum gegen wild-typ IHF analysiert. Ein Protein mit dem erwarteten molekularen Gewicht von ca. 20 KDA wurde in den Zellen nachgewiesen, die den oIHF2 kodierenden Vektor enthielten (Abb. D.16-B Spur 3 und 4). Das etwas langsamere Laufverhalten des Proteins in den pCMVSSoIHF2HIS kotransfizierten ROSA-IR Zellen (Spur 4) ist auf die 6 zusätzliche Histidin-Aminosäurereste zurückzuführen. In Zellen, die mit pCMV kotransfiziert wurden, ist kein Protein mit diesem Molekulargewicht nachweisbar (Spur 2 und 5). Als Kontrolle und Größenmarker wurden außerdem aus *E.coli* aufgereinigtes oIHF2HIS Protein (Spur 1; von P. Dröge zur Verfügung gestellt) und IHF Protein von Invitrogen (Spur 6) eingesetzt.

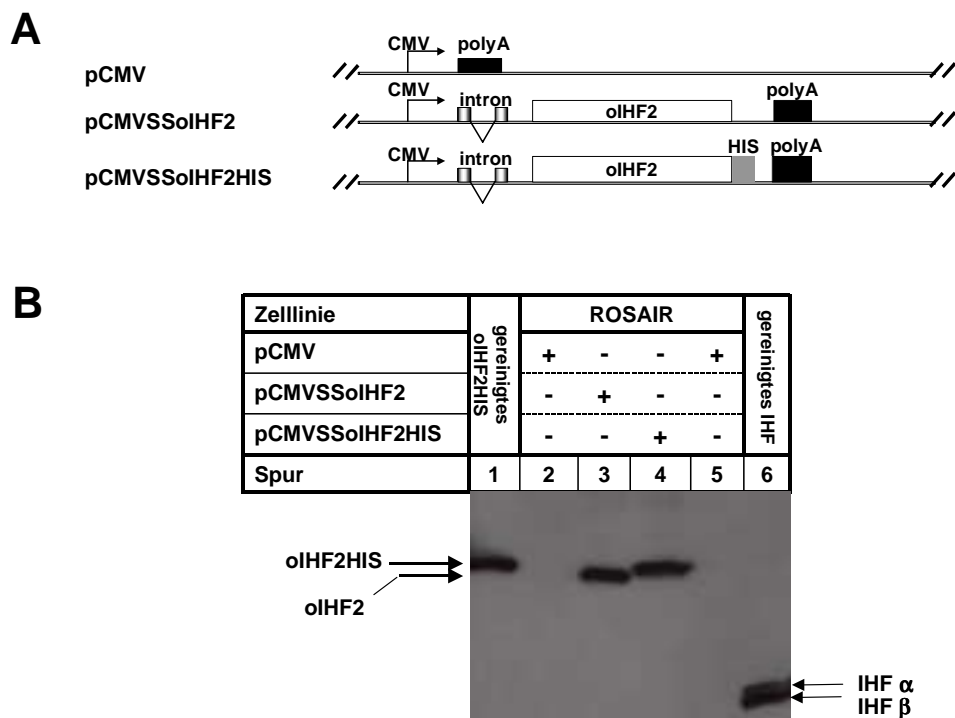
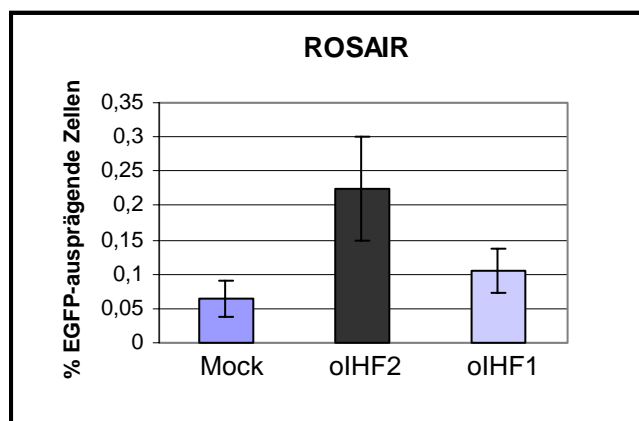


Abbildung D.16: Ausprägung von oIHF2 in ROSA-IR Zellen. In (A) sind schematisch die eukaryotischen Expressionsvektoren für oIHF2 und oIHF2HIS dargestellt. Sie wurden zusammen mit pCMVSSInt-h/218 in ROSA-IR Zellen kotransfiziert. CMV, Cytomegalovirus Promoter; HIS 6xHistidin Tag; polyA, Polyadenylierungs-Sequenz. (B) 24h nach der Elektroporation wurde die oIHF2 Proteine durch eine *Western Blot* Analyse in Zell-Lysaten der transfizierten ROSA-IR Zellen mit Anti-Serum gegen IHF nachgewiesen. Spur 1: aus *E.coli* gereinigtes oIHF2HIS Protein (P.Dröge); Spur 6: wild-typ IHF Protein (Invitrogen).

Die Ergebnisse der *Western Blot* Analysen zeigen, dass oIHF2 und oIHF2HIS auch in embryonalen Stammzellen der Maus ausgeprägt werden. Im Gegensatz zu Analysen mit Zell-Lysaten von pTrc/oIHF-transformierten *E.colis* (Absch. D5.1) waren in den Zell-Lysaten der kotransfizierten ROSA-IR Zellen keine Abbauprodukte des oIHF2 Proteins nachzuweisen (Vergleiche Abb. D.16-B mit Abb. D.14-B).

Um die Funktionalität des in ROSA-IR Zellen ausgeprägten oIHF2 Proteins zu untersuchen, wurden nach weiteren 48 Stunden die transfizierten ES-Zellen durchflusszytometrisch untersucht. Die Ergebnisse von vier unabhängigen Experimenten zeigen, dass in ROSA-IR Zellen, die mit pCMVSSoIHF2 kotransfiziert wurden, die integrative Rekombinationsaktivität von Int-h/218 im Durchschnitt 3,5-fach erhöht war (Abb. D.17). In einigen Experimenten prägten bis zu 0,28% der Zellen EGFP aus. Ansätze in denen der Expressionsvektor für oIHF2HIS anstelle von pCMVSSoIHF2 kotransfiziert wurde zeigten, dass die 6 Histidinreste am C-terminalen Abschnitt von oIHF2 keinen Einfluss auf die Stimulierung haben (Daten nicht gezeigt). Dies konnte in integrativen Rekombinationsanalysen in *E.coli* bestätigt werden (Daten nicht gezeigt). Ebenso zeigte auch eine nukleare Lokalisationssequenz am C-Terminus von oIHF2 keinen Einfluss auf die stimulatorische Wirkung der Int-h/218-vermittelten integrativen Rekombination in ROSA-IR Zellen.

Abbildung D.17: oIHF2 stimuliert die genomische Rekombinationsaktivität von Int-h/218 in ROSA-IR Reporter ES-Zellen. ROSA-IR Zellen, als Reporter für genomische Rekombinationsreaktionen zwischen *attB* und *attP*, wurden mit CMV-Expressionsvektoren für Int-h/218 und für oIHF2 oder oIHF1 kotransfiziert. 72h nach der Elektroporation wurde der Anteil an EGFP-ausprägenden ROSA-IR Zellen bestimmt. Die Balken zeigen die Mittelwerte aus vier (oIHF2) und zwei (oIHF1) Experimenten mit den Standard-abweichungen, die als vertikale Linien dargestellt sind. Als negative Kontrolle wurde der Vektor pCMV anstelle der oIHF-Expressionsvektoren kotransfiziert (Mock).



Im Gegensatz dazu zeigen die Ergebnisse von zwei unabhängigen Experimenten, in denen oIHF1 anstelle von oIHF2 in ROSA-IR ausgeprägt wurde,

wie erwartet keine signifikante Stimulierung der Int-h/218-vermittelten Rekombination. Die Transfektionseffizienz wurde für jedes einzelne Experiment bestimmt und betrug zwischen 68 und 85% (Daten nicht gezeigt).

Diese Ergebnisse zeigen, dass das zum Einsatz in Eukaryoten optimierte IHF-ähnliche Hilfsprotein oIHF2 die integrative Rekombinationsaktivität von Int-h/218 im Genom von Reporter ES-Zellen der Maus stimuliert und das fehlende wild-typ IHF zum Teil ersetzen kann.

D6 Identifizierung einer *attB*-ähnlichen Sequenz (*attH2*) im menschlichen Genom

Die Ergebnisse der in den Abschnitten D1 bis D3.1 beschriebenen Experimente zeigen, dass die Mutanten Int-h und Int-h/218 in menschlichen Zellen die integrative und exzusive Rekombinationsreaktion katalysieren. Es wurde außerdem gezeigt, dass die Ausprägung des für eukaryotische Zellen optimierte oIHF2 die integrative Rekombinationsaktivität von Int-h/218 in Reporter-ES-Zellen stimuliert (Abschn. D5.4). Zur Entwicklung von gentherapeutischen Techniken wäre die Integration von Fremd-DNA mit der *Lambda* Integrase in eine natürlich vorkommende Sequenz im menschlichen Genom wünschenswert. Daher wurde im Rahmen dieser Arbeit auch nach *attB*-analogen Sequenzen (Abschn. A3.3) im menschlichen Genom gesucht. Da *Lambda* Integrase einige Austausche in der *core*-Region toleriert (Nash, 1981), wurde mit dem Programm SeqLab-*FindPatterns* (*WisconsinPackage Version 10.2-UNIX; Genetics Computer Group, Madison, WI, USA*) in der Datenbank des *Human Genomes* (EMBL) nach einer geeigneten „*attH*“ gesucht. Im Rahmen dieser Suche wurde eine Sequenz (*attH2* in

	core-Bindungssequenz							core-Bindungssequenz													
<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	C	T	T	A	T	A	C	C	A	A	G	T	G	G	
	*	*	*	*	*	*	*	*				*		*	*	*	*	*	*	*	

Abbildung D.18: Sequenzvergleich von *attP*, *attB* und *attH2*. Horizontale Pfeile markieren die *core*-Sequenzen. Die *overlap*-Region ist grau unterlegt und die in *attH2* von *attP* und *attB* abweichenden Nucleotide sind fettgedruckt. Mit (*) markierte Nucleotide der *attB*- und *attP*-Sequenzen sind wichtig für die Bindung der Wildtyp-Integrase an die *core*-Sequenzen (Ross & Landy, 1983).

Abb. D.18) gefunden, die als Bestandteil einer mRNA aus der menschlichen Placenta, den Monocarboxylat-Transporter (MCT)5 kodiert (MCT5 (Halestrap & Price, 1999) wurde in Price et al. (1998) noch als MCT4 bezeichnet). Die komplette mRNA Sequenz ist unter der Kennzeichnung SLC16A4 oder der *Accession*-Nummer U59185 in der EMBL-Datenbank vermerkt und zum Teil (Basen 351 bis 750) im Anhang aufgeführt.

Ein Vergleich mit dem entsprechenden Gen, das sich auf dem Chromosom 1 von *Homo sapiens* befindet (*Accession*-Nummer AC025987 in der EMBL-Datenbank), ergab, dass *attH2* Bestandteil eines MCT5 Exons ist (siehe Anhang). Um welches Exon es sich dabei handelt, konnte aufgrund einer unvollständigen DNA-Sequenz dieses Chromosomen-Abschnitts nicht festgestellt werden. Der Vergleich der *attH2*-Sequenz mit der von *attB* oder *attP* ergab 3 Nucleotidaustausche. Zwei Nucleotidaustausche befinden sich in der rechten *core*-Bindungssequenz und eine am 5'-Ende der *overlap*-Sequenz (Abb. D.18). In den folgenden Experimenten sollte die *attH2*-Sequenz auf ihre Fähigkeit hin untersucht werden, als Substrat für die *Lambda* Integrase zu dienen.

D6.1 Sequenz-spezifische Rekombination von *attH2* und *attP*2* in *E.coli*

Zunächst wurde *attH2* im *E.coli* Stamm CSH26, der den Ko-Faktor IHF ausprägt, auf ihre Eignung als Rekombinationssubstrat untersucht. Dazu wurde der Vektor pACH2 konstruiert, der die *att*-Regionen *attH* und *attP*2* in entgegengesetzter Orientierung zueinander trägt (Abb. D.19-A). Die *attH2*-Sequenz wurde aus humaner DNA (HeLa) amplifiziert (Abschn. C1.10). *attP*2* ist eine modifizierte *attP*-Sequenz, in der das erste Nucleotid der *overlap*-Sequenz, ein T, durch ein C ersetzt wurde (Abschn. C1.10). Hierdurch sollte die Homologie im *overlap* zwischen *attH2* und *attP* gewahrt werden, die entscheidend für eine Rekombinationsreaktion ist (Abschn. A3.1). Rekombination führt zur Inversion des *att*-flankierten DNA-Abschnitts (Abb. D.19-A) und zu einer Änderung des AatII/AvaI-Restriktionsmusters der Plasmid-DNA.

Die Ergebnisse der Rekombinationsanalysen in CSH26 mit dem Substrat pACH2 und dem prokaryotischen Expressionsvektor pTrcInt-h (Christ & Dröge, 1999) zeigen, dass *attH2* und *attP*2* vollständig durch Int-h in Gegenwart von IHF zu *attL** und *attR** (D.19-A) rekombiniert wurden (Abb. D.19-B Spur 1 und 2).

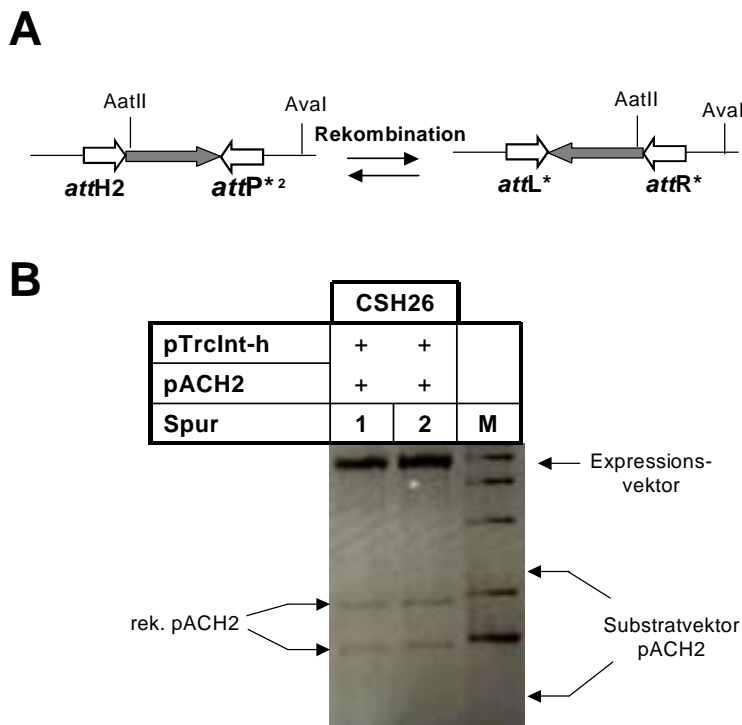


Abbildung D.19: Rekombinations-Analysen von *attH2* und *attP*2* in *E.coli* CSH26. In (A) ist der Substratvektor pACH2 dargestellt, auf dem *attH2* und *attP*2* in invertierter Orientierung zueinander angeordnet sind (helle Pfeile). Der graue Pfeil gibt die Orientierung des *att*-flankierten DNA-Segments an. AatII und Aval: Restriktions-schnittstellen (B) *E.coli* CSH26 wurden mit pACH2 und dem prokaryotischen Int-h-Expressionsvektor, pTrcInt-h, ko-transformiert. Die isolierte Plasmid-DNA aus den ÜNK wurde mit AatII und Aval geschnitten und in einem Agrosegel elektrophoretisch aufgetrennt. Spur 1: Plasmid-DNA aus ko-transformierten *E.coli* CSH26; Spur 2: unabhängige Reproduktion von Spur 1; M: 1Kb DNA-Leiter. Die Positionen der Aval/AatII-geschnittenen DNA-Fragmente des Substratvektors pACH2 sind ausschließlich durch Pfeile angezeigt.

D6.2 Episomale Rekombination von *attH2* und *attP*2* in HeLa durch die Mutante Int-h und Int-h/218

Um die Rekombination zwischen *attH2* und *attP*2* durch die Mutanten Int-h und Int-h/218 in eukaryotische Zellen zu untersuchen, wurden der Substratvektor pACH2 und der CMV-Expressionsvektor für Int-h oder Int-h/218 in HeLa kotransfiziert. 72h nach der Transfektion wurde die DNA isoliert und mit PCR analysiert. Das Primerpaar PI/PE amplifiziert nur dann ein Produkt von ~700Bp, wenn es durch Rekombination zwischen *attH2* und *attP*2* zur Inversion des *att*-flankierten DNA-Segments gekommen ist (Abb. D.20-A).

Die Ergebnisse einer repräsentativen PCR-Analyse zeigen, dass nur dann ein Produkt von ~700Bp entsteht, wenn die HeLa Zellen Int-h (Abb. D.20-B Spur 5) oder Int-h/218 (Spur 6) ausprägten. In PCR-Analysen von DNA aus Zellen, die mit dem Kontroll-Vektor pCMV kotransfiziert wurden, war kein Produkt nachweisbar (Spur 4). Diese Ergebnisse zeigen, dass *attH2* auch in eukaryotischen Zellen durch die Integrase Mutanten episodale rekombiniert werden kann.

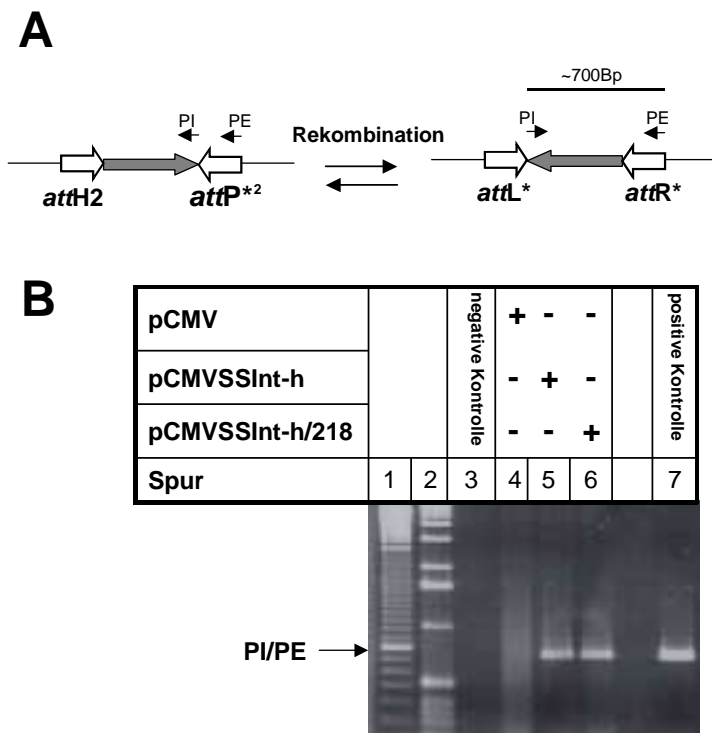


Abbildung D.20: Episomale Rekombination von *attH2* und *attP*2* durch Integrase Mutanten in HeLa. In (A) ist der Substratvektor pACH2 dargestellt, auf dem *attH2* und *attP*2* in invertierter Orientierung zueinander angeordnet sind (helle Pfeile). Der graue Pfeil gibt die Orientierung des *att*-flankierten DNA-Segementes an. Die kleinen Pfeile stellen die Primer dar. (B) PCR-Analyse von DNA, die 72h nach der Elektroporation aus HeLa Zellen isoliert wurde, die mit pACH2 und dem Expressionsvektoren für Int-h oder Int-h/218 transfiziert wurden. Die Primer sind so positioniert, dass nur nach der Rekombination ein PCR-Produkt von ~700Bp entsteht. Spur 1: 100Bp-Marker; Spur 2: 1Kb DNA-Leiter; Spur 3: PCR-Kontrolle ohne *Template*; Spur 7: in *E.coli* durch Int-h rekombiniertes pACH2.

D6.3 Untersuchung der MCT5 mRNA-Ausprägung in einer menschlichen Nieren- und Gebärmutterhalszelllinie

Die Integration von Fremd-DNA in natürlich vorkommende humane Sequenzen durch die *Lambda* Integrase erfordert Strategien zur Selektion von sequenzspezifisch rekombinanten Zellen. Dazu ist eine Charakterisierung des *attH* Locus in einer als Modellsystem dienenden Zelllinie von Vorteil.

Die humane *attH2*-Sequenz ist Bestandteil einer mRNA, die aus der menschlichen Placenta stammt (Price et al., 1998, in dieser Arbeit als MCT4 bezeichnet). Außer in der Placenta wird die Isoform 5 des MCT noch in Geweben von Herz, Niere, Eierstock, Dick- und Dünndarm ausgeprägt (siehe Abb. 2: MCT4 in Price et al., 1998). Auch die Transkription von MCT5 in der menschlichen Nierenzelllinie 293T als mögliches Modellsystem für die Int-vermittelte Integration von Fremd-DNA ins menschliche Genom wurde im Rahmen dieser Arbeit durch RT-PCR bestätigt. Dazu wurde die RNA aus 293T Zellen isoliert und eine cDNA-Synthese mit Hilfe von Oligo-dT-*Primer* durchgeführt (Abschn. C1.6). Anschließend wurde ein 650Bp langes DNA-Fragmente mit der *attH2* Sequenz durch PCR mit dem Primerpaar (MCT5')/(MCT3') (siehe Anhang) vervielfältigt. Als Kontrolle für die

Qualität der RNA Präparation und die cDNA-Synthese wurde in einem Ansatz die isolierte RNA vor der PCR zusätzlich mit DNase I behandelt.

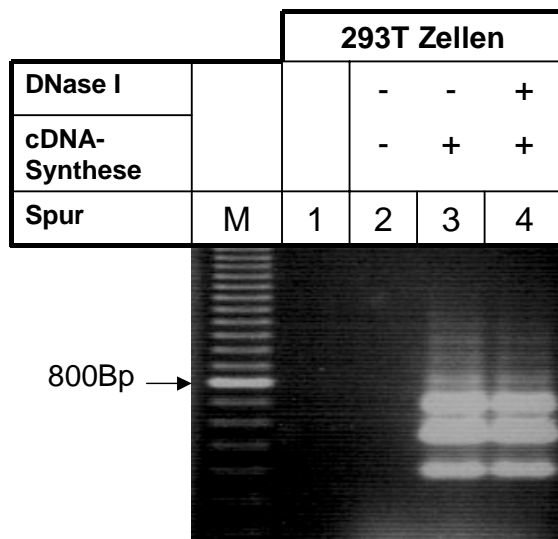


Abbildung D.21: RT-PCR-Analyse von MCT5 in humanen 293T Zellen. Aus der transformierten embryonalen Nierenzelllinie 293T wurde die RNA isoliert und nach einer cDNA-Synthese mit den Primern (MCT5') und (MCT3') amplifiziert. Zur Kontrolle der aufgereinigten RNA wurde ein Ansatz (Spur 4) vor der PCR zusätzlich mit DNase I behandelt. Die Abbildung zeigt die in einem Agarosegel aufgetrennten PCR-Produkte nach der Anfärbung mit Ethidiumbromid. Das erwartete MCT5 Transkript ist ~650Bp lang (1). Spur M: 100Bp Marker, Spur 1: PCR ohne *Template*; Spur 2: Ansatz ohne cDNA-Synthese, Spur 3: PCR von cDNA aus 293T Zellen; Spur 4: PCR der DNase I-behandelten RNA. 2 und 3 markieren zusätzliche RT-PCR-Produkte.

Abbildung D.21 zeigt die Produkte einer RT-PCR von isolierter 293T-RNA aufgetrennt in einem Agarosegel. Das erwartete Fragment von ~650Bp (1) konnte nur in den Ansätzen mit cDNA-Synthese nachgewiesen werden (Spur 3 und 4). In Ansätzen mit isolierter RNA ohne cDNA-Synthese wurden in der PCR-Analyse keine Produkte erzeugt (Spur 2). Die Integrität der DNA-Sequenz des 650Bp-Fragments wurde durch Sequenzierung betätigt (Daten nicht gezeigt). Neben dem zu erwarteten 650Bp-Fragment wurden zwei weitere RT-PCR-Fragmente in der Größe zwischen 500-600Bp (2) und 300-400Bp (3) vervielfältigt. Die Sequenzierung dieser Fragmente ergab, dass es sich ebenfalls um MCT5 Transkripte handelte, denen jeweils ein Sequenzabschnitt fehlt. Im 500Bp-Fragment fehlt Nucleotid 73 bis 404 und im 320Bp-Fragment Nucleotid 73 bis 548 der kodierenden MCT5-Sequenz (siehe Anhang). Hierbei könnte es sich möglicherweise um alternativ gespleißte Transkripte handeln. Auch in den beiden kürzeren Fragmenten ist die komplette *attH2*-Sequenz noch vorhanden. RT-PCR-Analysen in 293 und HeLa Zellen zeigten die gleichen Ergebnisse (Daten nicht gezeigt).

E Diskussion

E1 Effiziente sequenz-spezifische DNA Rekombination in Säugetierzellen durch *Lambda* Integrase Mutanten

In dieser Arbeit wurde ein Protokoll für einen einfachen und effizienten Einsatz des *Lambda* Rekombinationssystems in Eukaryoten entwickelt. Dabei wurden neben der wild-typ Integrase auch die Integrase Mutanten Int-h und Int-h/218 eingesetzt, deren enzymatische Eigenschaften entscheidend für den Einsatz des *Lambda* Systems in eukaryotischen Zellen waren. Int-h ist wie sein Derivat Int-h/218 enzymatisch aktiver als die wild-typ Integrase. Jedoch lässt sich auch bei ihnen, wie bei der wild-typ Integrase, die Richtung der Rekombinationsreaktion *in vitro* und in *E.coli* durch den natürlich vorkommenden Kofaktor *Integration Host Factor* (IHF) steuern (Christ & Dröge, 1999; Lange-Gustafson & Nash, 1984). Dies ist auch für den Kofaktor XIS anzunehmen, da in seiner Gegenwart die Mutante Int-h die exzusive Rekombination in *E.coli* genauso effizient katalysiert wie die wild-typ Integrase (Kikuchi et al., 1985). Ein Vorteil der Integrase Mutanten gegenüber der wild-typ Integrase ist deren Fähigkeit, auch ohne die natürlich vorkommenden Protein-Kofaktoren IHF und XIS effizient die integrative und exzusive Rekombination katalysieren zu können (Landy, 1989; Christ & Dröge, 1999). Die Topologie des Substrats spielt dabei keine Rolle (Lange-Gustafson & Nash, 1984).

Die *Lambda* Integrase ist, wie auch die Cre Rekombinase, prokaryotischen Ursprungs. Aus diesem Grund wurde zunächst durch *Western Blot* Analysen sichergestellt, dass das Integrase-Protein in Säugetierzellen ausgeprägt wird. Das vollständige Integrase-Protein mit dem erwarteten Molekulargewicht von ca. 40 KDA wurde mit einem Anti-Serum gegen die wild-typ Integrase in einer Reihe von Säugetierzellen nachgewiesen, die mit einem Expressionsvektor für Int-h transfiziert waren; darunter menschliche HeLa Zellen (Abb. D.1), NIH/3T3 Fibroblastenzellen der Maus (Abschn. D.1) und die humane B-Zelllinie BL60 (Lorbach et al., 2000). Dass es sich dabei tatsächlich um die Mutante Int-h handelt, wurde anschließend durch den Nachweis von sequenz-spezifischer integrativer und exziver Rekombinationsaktivität in diesen Zellen bestätigt. In nachfolgenden Experimenten konnte gezeigt werden, dass Int-h auch in anderen Zelllinien wie

embryonalen Stammzelle der Maus und CHO Hamster-Zellen rekombinationsaktiv ist (Abschn. D2.1 und D2.4).

Eine genauere Untersuchung enzymatischer Aktivitäten der Integrase Mutante Int-h sowie der ihres Derivates Int-h/218 erfolgte in transienten Kotransfektions Experimenten. Die Ergebnisse zeigen, dass beide Mutanten sowohl *att*-Sequenzen rekombinieren können, die in invertierter Orientierung zueinander vorliegen, als auch solche, die in direkter Orientierung zueinander angeordnet sind (Abschn. D.2). Dadurch war es mit den Integrasen Int-h und Int-h/218 in Säugetierzellen möglich, gezielt *att*-flankierte DNA Abschnitte auf episomalen DNA-Segmenten zu invertiert bzw. zu deletiert. Eine alternative Verknüpfung zwischen den *att*-Sequenzen, wie es mit den Mutanten in *E. coli* beobachtet wurde (Christ & Dröge 1999), konnte nicht nachgewiesen werden. Eine mögliche Erklärung hierfür könnte das Fehlen von dazu notwendigen Kofaktoren oder eine zu niedrige Effizienz der Rekombinationsreaktion in eukaryotischen Zellen sein. Int-h/218 erwies sich wie erwartet in den episomalen Analysen als aktivste Integrase (Abb. D.2). Die Fähigkeit der Mutante Int-h, *att*-Sequenzen auch ohne Kofaktoren sequenzspezifisch zu rekombinieren, wird im allgemeinen auf eine stärkere Affinität der Mutante für die *core*-Sequenzen zurückgeführt (Patsey & Bruist, 1995). Möglicherweise wird diese Affinität durch den zusätzlichen Lysin-Rest in der katalytischen Domäne der Doppelmutante Int-h/218 verstärkt, was die höhere enzymatische Aktivität der Doppelmutante erklären würde. Auch Untersuchungen in *E.coli* zeigen, dass der zusätzliche E218K-Austausch in Int-h/218 vor allem zu einer gesteigerten Rekombinationsaktivität von Int-h führt (Christ & Dröge, 1999).

Dennoch ist der genaue Rekombinationsmechanismus der Integrase Mutanten Int-h und In-h/218 in Abwesenheit von Kofaktoren zum gegenwärtigen Zeitpunkt noch nicht vollständig geklärt. Im Rahmen dieser Arbeit wurden daher Untersuchungen durchgeführt, die zeigen, dass die Arm-Bindungssequenzen in den *att*-Regionen eine wichtige Rolle bei der Integrase Mutanten-vermittelten Rekombination in eukaryotischen Zellen spielen (siehe Abschn. E4). Das gleiche gilt auch für den N-terminale Abschnitt der Integrase. Außerdem konnte gezeigt werden, dass *att*-Sequenzen, die auf zwei unterschiedlichen episomalen DNA-Segmenten liegen, mindestens so effizient durch die Integrase Mutanten katalysiert werden wie solche, die sich auf dem gleichen DNA-Abschnitt befinden (Abschn. E3).

In Säugetierzellen wurde die enzymatische Aktivität der Integrase Mutanten Int-h und Int-h/218 in der integrativen und exzisen Rekombination zunächst nur in transienten Kotransfektions-Experimenten demonstriert. Ein vorrangiges Ziel dieser Arbeit war aber der Einsatz der Integrase Mutanten zur gezielten genetischen Veränderung im Genom von Eukaryoten. Daher stellte sich die Frage, ob, trotz möglicher nukleosomaler Verpackung, auch stabil integrierte Substrate von den Integrase Mutanten erkannt und rekombiniert werden können. Dazu wurden Untersuchungen in HeLa Reporterzelllinien durchgeführt, die entweder eine oder mehrere Kopien des Substrats pGFP $attB/attP$ oder pGFP $attL/attR$ an zufälligen Positionen stabil im Genom integriert haben (Abschn. D3). Durch sequenzspezifische Rekombination der *att*-Sequenzen wurde ein DNA-Abschnitt so invertiert, dass es zur Ausprägung von EGFP kommen sollte. Dabei wurde auch das Restriktionsmuster des NcoI-Restriktionsenzym im Genom verändert (Abb. D.6-A).

Eine quantitative *Southern Blot* Analyse zeigte, dass die Integrase Mutanten Int-h und Int-h/218 effizient *attB* und *attP* im Genom von menschlichen Zellen rekombinieren können. Über einen Zeitraum von 72 Stunden wurden in der Reporterzelllinie H₈B von Int-h bis zu 16% der Substratmoleküle rekombiniert und von der Doppelmutante Int-h/218 sogar bis zu 30% (Absch. D3.1). In den beiden Reporterzelllinien für die exzise Rekombination dagegen konnte über den gleichen Zeitraum keine Inversion des EGFP-Gens nachgewiesen werden. Erst 72 Stunden später gelang es die exzise Rekombination durch Int-h mittels PCR nachzuweisen (Abb. D.7-B). Eine einfache Erklärung dafür könnte sein, dass im Genom der Reporterzelllinien für die exzise Rekombination nur eine Kopie des Substrats integriert ist. In der Reporterzelllinien für integrative Rekombinationsreaktionen H₈B sind dagegen mehrere Kopien des Substrats integriert. Dadurch steigt die Wahrscheinlichkeit der Rekombinationereignisse vor allem in den Zellen, die größere Mengen an Integrase Mutanten ausprägen. In den Reporterzellen für die exzise Rekombination kann dagegen trotz größerer Mengen an Integrase Mutante immer nur eine Kopie des Substrates rekombiniert werden. Zudem könnten auch die das Substrat flankierenden genomischen DNA-Sequenzen oder an diese Sequenzen bindende Proteine in den beiden Reporterzelllinien für exzise Rekombinationsreaktionen einen negativen Einfluss auf die Rekombination ausüben. Obwohl die mit der *Lambda* Integrase verwandte Cre-

Rekombinase durch die Chromatinstruktur des eukaryotischen Genoms nicht behindert wird (Kilby et al., 1993), könnten die Integrase Mutanten beispielsweise durch das Histondimer H2A-H2B, das durch die Bindung an *attL* die Ausbildung eines synaptischen Komplexes verhindert (Segall et al., 1994), in ihrer Aktivität gestört werden. Auch in drei unabhängigen BL60 Reporterzelllinien, in denen jeweils eine Kopie des Substrates pGFP*attL/attR* an zufälligen Positionen im Genom integriert sind, konnte über einen Zeitraum von 72 Stunden keine Rekombination durch Int-h nachgewiesen werden (Lorbach et al., 2000). In BL60 Reporterzellen mit nur einer stabil ins Genom integrierten Kopie des Substrats pGFP*attB/attP* wurde dagegen über den gleichen Zeitraum durch PCR-Analysen die integrative Rekombination durch Int-h nachgewiesen (Lorbach et al., 2000). Dies würde ebenfalls für eine höhere Effizienz der Int-h-vermittelten integrativen Rekombinationsreaktion an stabil integrierten *attB*- und *attP*-Sequenzen sprechen. Auch die exzisive Rekombinationsreaktionen der Integrase Mutanten an *att*-Sequenzen, die sich auf dem gleichen episomalen DNA-Sequenzen befinden, waren durchschnittlich ca. 20% niedriger als die entsprechenden integrativen Rekombinationsreaktionen (Abb. D.2 vergleiche A und C). Eine mögliche Erklärung dafür könnten unterschiedliche Auswirkungen der starken Transkriptionsaktivität des CMV-Promoters auf die Integrasebindung an *attL* und *attB* und/oder auf die Synapsen-Bildung der integrativen und exzisiven Rekombinationsreaktionen in den Zellen sein. Ob sich auch die integrative Rekombinationsreaktion im Genom von der entsprechenden exzisiven Reaktion durch die Integrase Mutanten unterscheidet, bleibt weiterhin zu untersuchen. Sinnvoll wäre in diesem Zusammenhang die Untersuchung einer größeren Anzahl an Reporterzelllinien mit nur einer Substratkopie an verschiedenen Positionen im Genom. Hilfreich wäre dabei auch der Vergleich von exzisiver und integrativer Rekombinationsreaktion an gleichen Positionen im Genom wie beispielsweise in Rekombinationsanalysen mit den Reporter ES-Zellen ROSA-IR und ROSA-ER (Abschn. E5).

Eine Ausprägung von EGFP konnte in den HeLa Reporterzellen trotz durch *Southern Blot* nachgewiesener effizienter Rekombination in durchflusszytometrischen Analysen nicht ermittelt werden. Die transiente Transfektion von rekombinierten pGFP*attB/attP* und pGFP*attL/attR* Vektoren zeigt, dass die Inversion des EGFP-Gens im Prinzip zur effizienten Ausprägung des EGFP Proteins führt (Daten nicht gezeigt). Die Ursache für das Ausbleiben des EGFP

Proteins in rekombinierten H₈B Zellen könnte eine Hybridisierung von *antisense*-RNA mit der mRNA des EGFP-Gens sein. Die *antisense*-RNA wird vom konstitutiv aktiven CMV Promoter nicht-rekombinierter Substratmoleküle der gleichen Zellen ausgeprägt. Die Technik der *antisense*-RNA Hybridisierung mit der komplementären mRNA zur Ausschaltung von Genen in Zellkultur (Nemir et al., 2000; Paddison et al., 2002) und *in vivo* als Alternative zu konditionalen „*knock out*“-Mäusen (Helene & Toulme, 1990; Erickson, 1999; Varga et al., 1999) ist weit verbreitet. Aufgrund der fehlenden EGFP-Ausprägung war es nicht möglich, den Anteil an Zellen mit rekombiniertem Substrat zu bestimmen. Um das Problem der *antisense*-RNA zu vermeiden, könnte der Substratvektor in Zukunft ähnlich wie in den Reporter ES-Zellen konstruiert werden (Abschn. E5). Anstelle der Inversion des EGFP-Gens würde dann ein Transkriptions-Stopp zwischen Promoter und EGFP-Gen deletiert werden, so dass das Reportergen nur nach Rekombination ausgeprägt wird.

Die sequenz-spezifischen Rekombinationsereignisse durch *Lambda* Integrasen und anderer Rekombinasen (Nagy, 2000; Olivares et al., 2001; Thyagarajan et al., 2001; Diaz et al., 1999; Schwikardi & Dröge, 2001) im Genom von Säugetierzellen machen deutlich, dass auch Rekombinasen prokaryotischen Ursprungs chromosomale DNA trotz ihrer nukleosomalen Organisation als Rekombinations-substrate erkennen und umsetzen können.

E2 Die Rekombinationsaktivität der wild-typ Integrase in Säugetierzellen

Im Gegensatz zu den Integrase Mutanten rekombiniert die wild-typ Integrase in eukaryotischen Zellen nur geringfügig *att*-Sequenzen, die sich auf episomalen DNA-Segmenten befinden, oder solche, die stabil ins Genom integriert sind (Abschn. D2 und D3). Durch die Transfektion von IHF-DNA Komplexen (Abb. D.12) konnte ihre enzymatische Aktivität in der episomalen, integrativen Rekombination deutlich gesteigert werden. Diese Stimulierung zeigt, dass die wild-typ Integrase das Potenzial zu einer effizienten Rekombination besitzt. Offenbar fehlen in Säugetierzellen jedoch Faktoren, die eine ähnliche Funktion wie IHF oder XIS besitzen, und/oder negativ gespannte DNA ist nicht in ausreichenden Mengen vorhanden. Diese Aussage unterscheidet sich von der einer anderen Studie, in der

das Gegenteil postuliert wird. Die Annahme, dass Kofaktoren in Säugetierzellen vorhanden sein könnten, stützt sich in der anderen Studie auf die Ergebnisse von nicht-quantitativen PCR-Analysen. Diese zeigen, dass die wild-typ Integrase des *Lambda* verwandten Bakteriophagen HK022 sowohl die integrative als auch die exzusive Rekombinationsreaktion auf episomalen DNA-Segmenten in Maus- und Affenzellen ausführt (Kolot et al., 1999). Im Gegensatz zu der vorliegenden Arbeit wurden jedoch mit der HK022 Integrase keine quantitativen Untersuchungen der episomalen Rekombinationsreaktionen durchgeführt. Es ist jedoch anzunehmen, dass auch die HK022 Integrase wie die wild-typ *Lambda* Integrase nur eine geringfügige Rekombinationsaktivität in Abwesenheit von Kofaktoren besitzt.

Die Steigerung der integrativen Rekombinationsereignisse über einen Zeitraum von 48 Stunden durch die Transfektion von IHF-DNA Komplexen macht deutlich, dass die Aktivität der *Lambda* wild-typ Integrase in Säugetierzellen erhöht werden kann (Abb. D.12). Zusammen mit einer zeitlich limitierten und regulierten Ausprägung der Integrase in den Zellen eröffnet sich dadurch die Möglichkeit, mit geeigneten Kofaktoren eine stärkere Regulation der sequenz-spezifischen Rekombinationsreaktionen zu erreichen. Dies ist vor allem für die zukünftige Entwicklung von gentherapeutischen Techniken mit dem *Lambda* System von Interesse, in denen Vektoren mit einer *attP*-Sequenz durch integrative Rekombination in eine *attB*-ähnliche Sequenz im menschlichen Genom integriert werden sollen (Abschn. E7).

Eine Steigerung der integrativen Rekombination durch die Mutante Int-h konnte dagegen über den gleichen Zeitraum mit dieser Methode nicht gemessen werden (Abb. D.12). Eine mögliche Erklärung dafür liefert die ohnehin signifikant höhere enzymatische Aktivität der Mutante Int-h gegenüber der wild-typ Integrase. Dies würde einen vergleichbar geringen stimulatorischen Effekt von IHF, der vor allem in der Anfangsphase vorherrschen könnte, überdecken. Da sich die Richtung der Rekombinationreaktion mit den Integrase Mutanten sowohl in *E.coli* als auch *in vitro* durch spezifische Kofaktoren steuern und die Effizienz der Rekombination stimulieren lässt, wäre auch anzunehmen, dass Int-h durch IHF in Säugetierzellen aktiviert wird. Gleiches würde man auch für die Mutante Int-h/218 erwarten.

Um eine messbare Stimulierung der Integrase Mutanten über den Zeitraum von 48 Stunden zu erzeugen, müsste vermutlich der Kofaktor IHF in größerer Menge vorhanden sein. Diese Vermutung wird vor allem durch Untersuchungen in den

Reporter-ES-Zellen unterstützt, die zeigen, dass der Kofaktor IHF als Monomer die enzymatische Aktivität von Int-h/218 stimuliert (Abschn. E6).

E3 Die Reaktivität von episomalen DNA-Segmenten in HeLa Zellen während der *Lambda* Rekombination

Zum besseren Verständnis der *Lambda* Integrase Mutanten katalysierten Rekombination in eukaryotischen Zellen wurden in Abschnitt D2 Studien durchgeführt, die den Mechanismus der Rekombination im zellulären Umfeld von Eukaryoten untersuchen, in dem die natürlich vorkommenden Kofaktoren IHF, XIS und negativ gespannte DNA nicht vorkommen. Die Ergebnisse von episomalen Rekombinationsanalysen in HeLa Zellen (Abschn. D2.1) zeigen, dass Int-h und Int-h/218 sowohl die integrative als auch die exzusive, intermolekulare Rekombinationsreaktion mindestens genauso effizient katalysieren wie die entsprechende intramolekulare Reaktion. Dies deutet darauf hin, dass die produktive Umsetzung der *att*-Sequenzen in den Zellen hauptsächlich durch ein zufälliges Zusammentreffen erfolgt. Dies steht im Einklang mit früheren Studien, in denen die Synapsen-Bildung der Sequenzen durch die wild-typ Integrase untersucht wurde (Mizzuuchi et al., 1980; Spengler et al., 1985). Diese Ergebnisse werden von Untersuchungen unterstützt, die mit der verwandten Cre Rekombinase durchgeführt wurden. Auch in ihnen wurden gleiche Effizienzen von inter- und intramolekularen, episomalen Rekombinationsreaktionen in Pflanzenzellen (Dale & Ow, 1990) und Säugetierzellen (Rouet et al., 1994) beobachtet. Angesichts dieser Ergebnisse kann aus den in dieser Studie durchgeführten Analysen die allgemein gültige wichtige Schlussfolgerung gezogen werden, dass im Gleichgewicht von Rekombinationsreaktionen in eukaryotischen Zellen keine eindeutige Präferenz für *cis* Interaktionen gegenüber *trans* Interaktionen zwischen transfizierter episomaler DNA Segmente besteht.

Die episomalen Rekombinationsanalysen über einen Zeitraum von 72 Stunden zeigen auch, dass die Mehrzahl der produktiven Interaktionen zwischen den *att*-Sequenzen in den ersten 24 Stunden nach der Transfektion stattgefunden haben müssen. Das kann bedeuten, dass die episomale DNA innerhalb der Zelle sehr beweglich ist, wie es auch bei einigen Proteinen der Fall ist (Phair & Misteli, 2000). Eine andere mögliche Erklärung für die effiziente Rekombination innerhalb der

ersten 24 Stunden könnte auch eine erhöhte lokale Konzentration von transfizierten DNA Segmenten sein, die sich in kleinen Kompartimenten der Zelle befinden könnten. Ob die Rekombination im Zytoplasma oder im Nukleus der Zelle abläuft, ist nicht bekannt. Die Tatsache, dass der mit dem Substrat kotransfizierte Expressionsvektor für die Integrase zunächst den Nukleus erreichen muss, um Transkripte der *Int* zu erzeugen, könnte darauf hinweisen, dass zumindest einige Rekombinationsreaktionen im Nukleus stattgefunden haben. Dies ist auch dann der Fall, wenn genomische *att*-Sequenzen als Substrate dienen (Lorbach et al., 2000).

E4 Hinweise für die Bildung eines synaptischen Komplexes in der von Integrase Mutanten katalysierten Rekombination

Die Analysen in Abschn. D 2.2 zeigen, dass von allen Integrasen vor allem die Mutante *Int-h/218* effizient die intermolekulare Rekombination zwischen zwei *attB* Sequenzen katalysiert (Abb. D.4). *attB* besitzt im Gegensatz zu *attP*, *attL* oder *attR* keine Arm-Bindungssequenzen. Dass die Rekombination trotzdem stattfindet, kann durch einen einfachen Reaktionsablauf erklärt werden, der dem der Cre und FLP Rekombinase ähnlich ist (Landy, 1993; Voziyanov et al., 1999): Vier *Int-h/218* Monomere binden an die beiden *core*-Sequenzen der *attB*-Regionen. Durch zufälliges Zusammentreffen zweier *attB*-gebundenen *Int* Dimere entsteht ein synaptischer Komplex, in dem der Strang austausch stattfindet. Sind dagegen die Arm-Bindungssequenzen gegenwärtig, wird die Rekombinationsaktivität der *Int* Mutanten vier- bis achtfach stimuliert (Abb. D.5). Diese Ergebnisse werfen die folgende Frage auf: Welche funktionelle Rolle spielen die Arm-Bindungssequenzen in der durch Integrase Mutanten-vermittelten Rekombination?

Da in eukaryotischen Zellen keine Hilfsproteine vorkommen, die IHF oder XIS ersetzen könnten (Abschn. E2), kann man daraus die folgende wichtige Schlussfolgerung ziehen: Die Nucleoprotein Strukturen, die durch die Bindung der Integrase Mutanten an den *att*-Sequenzen zustande kommen, unterscheiden sich wahrscheinlich in ihrer Architektur von denen, die mit den natürlich vorkommenden Kofaktoren und der wild-typ Integrase (oder den Mutanten) in *E.coli* und *in vitro* gebildet werden. Für den stimulatorischen Effekt der Arm-Bindungssequenzen auf die Rekombination ohne die Hilfe von Kofaktoren gibt es zwei mögliche

Erklärungen: Zum einen wäre es denkbar, dass DNA-krümmende Proteine aus Säugetierzellen, wie beispielsweise HMG1 und HMG2, unspezifisch an Positionen zwischen Arm- und *core*-Bindungssequenzen binden und diese krümmen. Für HMG1 und HMG2 konnte *in vitro* gezeigt werden, dass sie teilweise den Ko-Faktor IHF in integrativen und exzisen Rekombinationsreaktionen der wild-typ Integrase ersetzen können (Segall et al., 1994). Zusammen mit der höheren Bindungsaffinität der Integrase Mutanten für die *core*-Bindungssequenzen, könnten die durch diese Proteine induzierten DNA-Krümmungen die simultane Bindung eines Integrase Protomers an die Arm- und *core*-Bindungssequenzen innerhalb derselben *att*-Region stabilisieren. Dies wiederum könnte zur Bildung von stabilen Nukleoprotein-Komplexen führen, die teilweise rekombinationsaktiv sind. Andererseits ist es auch möglich, dass die Gegenwart der Arm-Bindungssequenzen in den *att*-Regionen, die von der Integrase mit hoher Affinität gebunden werden, eine lokale Konzentrationserhöhung von Integrase Mutanten in der Nähe der *core*-Sequenzen bewirken. An den *core*-Sequenzen findet der Strangaustausch statt, und eine Erhöhung der Konzentration an Integrase-Mutanten in diesen Regionen würde die Häufigkeit von Rekombinationsreaktionen zwischen zwei *att*-Sequenzen steigern. Die Erhöhung der lokalen Protein Konzentration in der Nähe von Aktionsorten ist ein allgemein gültiges Konzept in der Kontrolle von DNA Transaktionen von bivalent DNA-bindenden Proteinen (Dröge & Müller-Hill, 2001).

Zusätzliche Informationen zum Reaktionsmechanismus der Integrase Mutanten in menschlichen Zellen liefern Studien, die von Teresa Corona (Universität Köln) durchgeführt wurden. Sie zeigen, dass neben den Arm-Bindungssequenzen auch der N-terminalen Abschnitt der wild-typ Integrase und der Mutante Int-h/218 einen Einfluss auf die Rekombinationsaktivität von Int-h/218 besitzen (Christ et al., 2002). Dieser Abschnitt der Integrase bindet an die Arm-Bindungssequenzen der *att*-Regionen. Segall und Mitarbeiter konnten nachweisen, dass der N-Terminus der wild-typ Integrase mit großer Wahrscheinlichkeit an Protein-Protein Interaktionen beteiligt ist, die für die Rekombination notwendig sind (Jessop et al., 2000). Es wäre daher denkbar, dass die N-terminale Domäne von Int-h/218 in menschlichen Zellen eine wichtige Rolle in der Bildung oder Stabilisierung eines synaptischen Komplexes spielt, wie er beispielsweise zwischen zwei *attB*-Sequenzen gebildet wird. Kürzlich wurde gezeigt, dass die isolierte C-Domäne der wild-typ Integrase

(Abb. A.6) die DNA *in vitro* an den *core*-Regionen schneidet (Sarkar et al., 2001). Von dieser Eigenschaft ist auch bei der isolierten C-Domäne der Mutante Int-h/218 auszugehen, die ein schwaches Potenzial zur Rekombination besitzt (Christ et al., 2002). Wie diese Rekombinationsereignisse allerdings ablaufen ist nicht klar. Wie schon zuvor diskutiert, könnten die DNA ohne die Bildung von stabilen synaptischen Komplexen zwischen den *att*-Sequenzen geschnitten werden. Diese Einzel- oder Doppelstrangbrüche an den *core*-Regionen könnten dann homologe oder illegitime Rekombinationsereignisse zwischen den *att*-Sequenzen auslösen, die zur Ausprägung des Reporters EGFP führen. Eine andere mögliche Erklärung könnte auch die kurzzeitige Bildung von synaptischen Komplexen durch die C-Domäne sein, die in einigen Fällen stabil genug sind, um die beiden Einzelstrangaustausche einzuleiten. Die daraus hervorgehende Holliday-Struktur könnte dann durch noch nicht identifizierte Faktoren, die es in Säugetierzellen zu geben scheint, aufgelöst werden (Constantinou et al., 2001). Um diese beiden Möglichkeiten zu überprüfen, müsste jedoch die C-Domäne *in vitro* ausführlicher charakterisiert werden.

Zusammenfassend gilt festzustellen, dass für eine effiziente sequenzspezifische Rekombination durch *Lambda* Integrase Mutanten in Säuger-Zellen sowohl die Arm-Bindungssequenzen in den *att*-Regionen als auch die Integrase Proteine in ihrer vollen Länge notwendig sind. Außerdem sind in diesen Zellen keine Faktoren vorhanden, die die Funktion der natürlichen Kofaktoren aus *E. coli* ersetzen können.

E5 Genetische Manipulation von embryonalen Stammzellen der Maus durch Int-h/218

Die Studien in Abschnitt D3.2 zeigen, dass Int-h/218 in embryonalen Stammzellen der Maus eingesetzt werden kann, um nach der Integration von DNA-Fragmenten durch homologe Rekombination gezielt Resistenzmarker-Gene aus dem Genom auszuschneiden. Das Ausschneiden kann sowohl durch die Int-h/218-vermittelte integrative als auch exzisive Rekombinationsreaktion katalysiert werden. Die Ausprägung von EGFP in den Reporter ES-Zellen ROSA-IR und ROSA-ER wurde zum einen als Reporter für die Int-h/218-vermittelte Rekombination eingesetzt und zum anderen zur Selektion von rekombinierten

Zellen. Obwohl die Effizienz für beide Rekombinationsreaktionen ziemlich gering ist -in der Regel prägten zwischen 0,01% und 0,1% der transfizierten Reporterzellen EGFP aus (Abb. D.9 und Daten nicht gezeigt)- machen diese Ergebnisse jedoch deutlich, dass Int-h/218 präzise und zuverlässig die genomischen Erkennungssequenzen beider Rekombinationswege in den ES-Zellen rekombiniert. Dabei konnte kein signifikanter Unterschied zwischen der Int-h/218-vermittelten integrativen und exzisivem Rekombinationsreaktion festgestellt werden. Aufgrund der geringen Effizienz wären minimale Unterschiede auch nicht messbar. Der Grund für die limitierte Rekombinationsaktivität der Mutante Int-h/218 an den ins Mausgenom integrierten natürlichen *att*-Regionen im Vergleich zu denen, die ins Genom von HeLa Zellen integriert sind, könnte durch eine geringere Menge an Integrase Protein in den ES-Zellen verursacht werden. Dies wiederum könnte die Folge von einer geringen Aufnahme des Expressionsvektors pro Zelle sein, was durch die Beobachtung unterstützt wird, dass die Transfektionsrate der ES-Zellen deutlich niedriger war als die für HeLa Zellen (Daten nicht gezeigt). Eine alternative Erklärung wäre die hohe Teilungsrate der ES-Zellen, die vermutlich zur schnellen Verdünnung der intrazellulären Int Konzentration führt. Obwohl Soriano (1999) gezeigt hat, dass der ROSA26 Locus der Maus im Allgemeinen zugänglich für Cre ist, könnte sich auch die Chromatin Struktur und die Häufigkeit, mit der die Replikationsgabel durch den genomischen Locus läuft, negativ auf die sequenz-spezifische Rekombinationsfrequenz in ES-Zellen auswirken. In den ES-Reporterzellen ist nur eine Kopie des Substrates in das Genom integriert. In der HeLa Reporterzelllinie H₈B, in der Int-h/218 bis zu 30% der Substratmoleküle rekombiniert (Abschn. D 3.1), sind dagegen ca. 8 Kopien von pGFP*attB*/*attP* an zufälligen Positionen stabil ins Genom integriert. Dadurch könnte sich die Wahrscheinlichkeit der Rekombinationsereignisse pro Zelle erhöhen und zu einer höheren Rekombinationseffizienz in den HeLa Reporter-Zellen beitragen.

Die geringe Effizienz der Rekombination in ES Reporterzellen könnte jedoch auf eine hohe Spezifität der Integrase Mutante Int-h/218 für natürliche *att*-Sequenzen, die künstlich in das Genom von Maus ES-Zellen integriert wurden, hinweisen. Angesichts der kürzlich entdeckten, aberranten Cre-abhängigen Rekombinationsreaktionen in Säugetier- und Bakterienzellen (Loonstra et al., 2001; Silver & Livingston, 2001; Aranda et al., 2001), könnte daher ein verbessertes Int-System zu einer sinnvollen und sicheren Ergänzung der schon existierenden sequenz-

spezifischen Rekombinationssysteme für Säugetierzellen werden. Dies gilt insbesondere für die Entfernung von Resistenzmarkern aus dem Genom.

E6 Optimierung des prokaryotischen *Integration Host Factors* (IHF) für den Einsatz in eukaryotischen Zellen

Wie schon in der Einleitung beschrieben, lässt sich, im Gegensatz zu Cre und FLP, sowohl die Richtung der wild-typ Integrase-vermittelten Rekombinationsreaktion als auch die der Mutanten Int-h und Int-h/218 durch spezifische Kofaktoren *in vitro* und in *E.coli* steuern. Dies ist ein entscheidender Vorteil des *Lambda* Rekombiantionssystem gegenüber dem von Cre und FLP dar. Ist die Richtung der Rekombinationsreaktion steuerbar, kann man selbst in Gegenwart der Rekombinase eine Stabilität der Rekombinationsprodukte erreichen, was wiederum zu einer höheren Effizienz führt. Dies ist vor allem für eine stabile Integration von Fremd-DNA ins Säugetiergenom durch sequenz-spezifische Rekombinasen wünschenswert. Eine kontrollierte sequenz-spezifische Integration ins menschliche Genom durch die *Lambda* Integrase könnte beispielsweise eine Anwendung in gentherapeutischen Techniken finden.

Die Integrase Mutanten Int-h und Int-h/218 sind in der Lage, sowohl die integrative als auch die exzusive Rekombination effizient in Säugetierzellen zu katalysieren (Abschn. E1), während die wild-typ Integrase aufgrund fehlender Kofaktoren wie IHF und XIS in eukaryotischen Zellen nicht signifikant aktiv ist (Abschn. E2). Kofaktoren wie IHF könnten dazu genutzt werden, die enzymatische Aktivität der wild-typ Integrase zu stimulieren (Abschn. E2) und die Richtung der von Int-Mutanten vermittelten Rekombinationsreaktion in Säugetierzellen zu steuern.

Die Kotransfektion von Expressionsvektoren für die beiden IHF Untereinheiten HimA (IHF α) und HimD (IHF β) zeigt nicht die gewünschte Stimulierung der wild-typ Integrase (D4.1). Eine *Western Blot* Analyse mit einem Anti-Serum gegen wild-typ IHF der kotransfizierten Zellen ergab außerdem, dass keine Proteine mit den erwarteten Molekulargewichten in den Zellen nachgewiesen werden konnte (Daten nicht gezeigt). Eine mögliche Erklärung dafür könnte eine Instabilität der beiden IHF-Untereinheiten sein, die durch fehlende oder fehlerhafte Dimerisierung schneller abgebaut würden. Untersuchungen in *E.coli* zeigen, dass nur

stöchiometrische Mengen IHF α intrazelluläres IHF β löslich machen und umgekehrt nur durch stöchiometrische Mengen IHF β die Untereinheit IHF α stabilisieren können (Nash et al., 1987). Um den Dimerisierungsschritt der beiden IHF-Untereinheiten in eukaryotischen Zellen zu umgehen und um gleichzeitig die funktionelle dreidimensionale Struktur des Dimers beizubehalten, wurde das IHF Protein im Rahmen dieser Arbeit als Monomer konstruiert. Eine Möglichkeit, die beiden IHF-Untereinheiten in Form einer einzigen Polypeptidkette zu synthetisieren, wäre den C-Terminus der einen Untereinheit mit dem N-Terminus der anderen Untereinheit durch einen Aminosäure(AS)-Linker zu verbinden. Die Kristall-Struktur (Abb. D.13-B) macht jedoch deutlich, dass ein AS-Linker die Bindung an die DNA beeinträchtigen würde und somit die Krümmung der DNA nicht mehr gewährleistet wäre. Um dies zu vermeiden, wurden die ersten beiden Helices der Untereinheit β an den N-Terminus der Untereinheit α fusioniert. Der C-terminale Abschnitt von α wiederum wurde mit dem verbliebenen Teil der Untereinheit β verbunden (Abb. D.13). Um die dreidimensionale Struktur des IHF-ähnlichen Proteins zu erhalten wurde an den Verknüpfungspunkten der beiden Untereinheiten zusätzliche AS-Reste eingeführt. Insgesamt wurden vier verschiedene AS-Linker in vier Varianten des *one chain* IHF (oIHF) Proteins untersucht. Zunächst wurde sichergestellt, dass eine Ausprägung dieses künstlich erzeugten Proteins in *E.coli* möglich ist. In der Tat wurden durch Western Blot Analysen oIHF1 und oIHF2 (Abb. D.13-A) in *E.coli* nachgewiesen (Abb. D.14-B). In oIHF2 ist die Untereinheiten α am 5'-Ende mit 5 Aminosäureresten und am 3'-Ende mit zwei Resten an die Untereinheit β verknüpft. In oIHF1 dagegen befinden sich an den entsprechenden Positionen 4 und 1 AS-Reste.

Rekombinationsanalysen in *E.coli* zeigten, dass oIHF2, nicht aber oIHF1, die integrative Rekombinationsreaktion durch die wild-typ Integrase stimuliert (Abb. D.15-B). Eine Instabilität des oIHF1 Proteins kann als Ursache für die fehlende Aktivität in *E.coli* durch die *Western Blot* Analysen ausgeschlossen werden (Abb. D.14-B). Sie lässt sich eher dadurch erklären, dass durch die kürzeren Linker in oIHF1 die dreidimensionale Struktur des Proteins so verändert sein kann, dass keine spezifische Bindung des Proteins an die IHF-Bindungssequenz mehr zustande kommt und/oder die oIHF1 Struktur keine ausreichende Krümmung der

DNA zulässt, die für die Ausbildung eines funktionellen Intasoms notwendig ist (Nunes-Düby et al., 1995).

Um den Grund für die Inaktivität von oIHF1 näher zu untersuchen, müssten *in vitro*-DNA-Bindungsstudien mit gereinigtem oIHF1 und oIHF2 durchgeführt werden. Die Bindungseigenschaft der beiden Proteine an H', eine der am besten untersuchten IHF-Bindungssequenzen aus *attP* (Yang & Nash 1995), könnte in Gelretardierungsanalysen verglichen werden. Auch die DNA-krümmenden Eigenschaften der beiden Proteine könnten in zukünftigen Experimenten untersucht werden und mit der von wild-typ IHF verglichen werden.

Weiteren Aufschluss über die Funktionsweise der oIHF Proteine könnten auch Untersuchungen in dieser Arbeit mit den Varianten oIHF3 und oIHF4 geben. oIHF3 und oIHF4 sind Hybride aus oIHF1 und oIHF2 und unterscheiden sich von den zuletzt genannten dadurch, dass die AS-Linker untereinander ausgetauscht wurden (Abb. D.13-A). oIHF3 und oIHF4 stimulieren beide in *E.coli* die enzymatische Aktivität der wild-typ Integrase; oIHF3 jedoch stärker als oIHF4 (Abb. D.15-B). Das könnte bedeuten, dass durch die Länge der AS-Linker auch die für die Stimulierung der Rekombination notwendige dreidimensionale Struktur entscheidend beeinflusst wird. Ob dies die spezifische Bindung an die IHF-Bindungssequenzen oder aber die Art der DNA-Krümmung beeinflusst, könnte wiederum nur durch entsprechende *in vitro* Analysen untersucht werden.

Die Ausprägung des rekombinationsstimulierenden oIHF2 Proteins in *E.coli* zeigt, dass funktionsfähiges IHF grundsätzlich in Form eines einzigen Proteins ausgeprägt werden kann. Das einsträngige IHF-ähnliche Protein erlaubt es sogar, die ersten beiden Helices von IHF β vom Rest der Untereinheit zu trennen und in einem neuen Kontext bzw. in einer neuen Abfolge zu falten, wie es auch schon für andere Proteine gezeigt wurde (Hahn et al., 1994). Die Funktionsfähigkeit des oIHF2 Proteins lässt vermuten, dass dessen dreidimensionale Struktur der von wild-typ IHF ähnlich ist. Dies könnte jedoch nur durch den Vergleich der Kristallstrukturen festgestellt werden, die es für oIHF2 noch nicht gibt.

Von vorrangigem Interesse in dieser Arbeit ist jedoch, ob oIHF2 auch integrative Rekombinationreaktionen der Integrase Mutanten in Säugetierzellen stimuliert. Das dies zumindest in der ES Reporterzelllinie ROSA-IR der Fall ist, zeigen Rekombinationsanalysen mit Int-h/218 und oIHF2 (Abschn. D5.3). Die

Rekombination zwischen den integrierten *attB* und *attP* Sequenzen durch Int-h/218 konnte zwar durch die Ausprägung von oIHF2 mehrfach gesteigert werden, jedoch bleibt die Rekombinationseffizienz mit bis zu 0,37% (unter Berücksichtigung der Transfektionseffizienz) deutlich geringer als bei entsprechenden Experimenten mit der Rekombinase Cre (F.Edenhofer, Universität Köln: persönliche Mitteilung). Auch durch eine zusätzliche nukleare Lokalisationssequenz (NLS), die grundsätzlich die Funktionalität des oIHF2 nicht beeinträchtigt, konnte die Rekombinationseffizienz nicht weiter gesteigert werden (Daten nicht gezeigt). In den ES-Zellen wurde das oIHF2 Protein ausschließlich in seiner vollen Länge nachgewiesen. Abbauprodukte wie sie in *E.coli* beobachtet wurden traten dagegen nicht auf (Vergleiche Abb. D.16 mit D.14) Das bedeutet, dass oIHF2 in den ES-Zellen stabil ausgeprägt wird und sehr wahrscheinlich auch ohne eine NLS in den Nukleus gelangen kann. In Zukunft müsste untersucht werden, ob oIHF2 auch in anderen Säugetierzellen die Rekombinationsaktivität von Int-h/218 erhöht. Denkbar wäre auch, dass die Ausprägung von oIHF2 in geeigneten Reporterzellen zur Stimulation der sequenz-spezifischen Integration von *attP*-tragenden Vektoren eingesetzt werden kann. Von den vier oIHF Varianten ist oIHF2 am stärksten aktiv. Es ist jedoch nicht auszuschließen, dass eine Verlängerung der AS-Linker die Aktivität noch steigern könnte.

E7 Perspektive: Der Einsatz des *Lambda* Systems zur gezielten Integration von Fremd-DNA ins Genom von Säugetieren

Wie lässt sich das sequenz-spezifische *Lambda* Rekombinationssystem in Säugetierzellen für die Forschung und für therapeutische Zwecke sinnvoll anwenden?

Die effizienten integrativen Rekombinationsreaktionen der Mutante Int-h/218 in Gegenwart von oIHF, oder auch die exzisen Rekombinationsreaktionen durch die *Lambda* Integrase Mutanten wären hervorragend zur Erzeugung von viralen Vektoren mit geringer Immunogenität geeignet. Virale Vektoren werden in der Gentherapie bevorzugt als Vehikel zum Gentransfer eingesetzt. Herkömmliche virale Vektoren erzeugen jedoch Antigene, die in der Regel im Organismus eine Immunreaktion auslösen. Dadurch werden die transduzierte Zellen und das transgene Produkt zerstört und/oder eine erneute Anwendung der viralen Vektoren

verhindert (Kay et al., 1997). Das wiederum mindert den therapeutischen Effekt der Gentherapie und kann außerdem im Organismus zu starken Nebenwirkungen führen. Um die Immunogenität der Vektoren zu verringern, werden derzeit die viralen Gene durch sequenz-spezifische Rekombinasen entfernt (Parks et al., 1996)

Eine gezielte genetische Veränderung im Genom von Säugetierzellen mit Hilfe des *Lambda* Integrase Systems könnte zum gegenwärtigen Zeitpunkt vor allem eine Anwendung in Studien mit sogenannten Doppelreporter-Systemen (Novak et al., 2000; Werdien et al., 2001) finden. Ein weiteres mögliches Einsatzgebiet ist die Analyse von Genfunktionen, in denen das rekombinante Gen gleichzeitig mit einem Selektionsmarker ausgeprägt wird. Für den Aufbau von komplexen genetischen Systemen in *Gene Targeting* Strategien, die das Entfernen verschiedener DNA-Segmente erfordern, könnten *Lambda* Integrasen auch zusammen mit anderen Rekombinasen, wie Cre und FLP, angewendet werden.

Die wahrscheinlich interessanteste medizinische Anwendung könnte das *Lambda* System in der Gentherapie finden. Die effiziente integrative Rekombinationsreaktion durch Integrase Mutanten könnte hier zur gezielten Integration von Fremd-DNA ins menschliche Genom eingesetzt werden. Das dies im Prinzip möglich ist, wurde bereits von E.Lorbach (Lorbach, 2000) und T.Corona (persönliche Mitteilung) gezeigt. Der Kofaktor oIHF2, oder auch verbesserte Varianten desselbigen, könnten dabei eine entscheidende Rolle spielen. Zum einen würde oIHF die integrative Rekombinationsreaktion von Int stimulieren. Zum anderen würde oIHF gewährleisten, dass das therapeutische Gen auch in Gegenwart der Integrasen im Genom stabil integriert bleibt. Für die Entwicklung von gentherapeutischen Methoden wäre vor allem die Integration von Fremd-DNA mit *Lambda* Integrasen in eine natürlich vorkommende Sequenz des menschlichen Genoms wünschenswert, wie es ansatzweise schon für die R4 Integrase und ϕ C31 Integrase gezeigt werden konnte (Olivares et al., 2001; Thyagarajan et al., 2001). Für die *Lambda* Integrase Mutanten wurde im Rahmen dieser Arbeit *attH2*, eine *attB*-analoge Sequenz, im Genom des Menschen identifiziert, die sich in Säugetierzellen als rekombinationsaktiv erwies (Abschn. D6). Die Voraussetzung dafür, fremde DNA durch Integrasen in *attH2* zu integrieren, ist die Entwicklung geeigneter Strategien zur Selektion sequenz-spezifisch rekombinanter Zellen.

Hierbei könnte die in Abschn. D6.3 beschriebene Ausprägung des *attH2* Locus als Teil des Monokarboxylat-Transporter (MCT)5 in einer menschlichen Nieren- und Gebärmutterhalszelllinie von Vorteil sein. Die Fremd-DNA könnte ins Genom dieser Zellen so integriert werden, dass es zur Fusion des MCT5 offenen Leserasters mit dem eines Selektionsmarkers kommt. Ähnlich wie in sogenannten *promoter-trap*-Strategien (Friedrich & Soriano, 1991; Soriano et al., 1991), würde man sich auch hier die Ausprägung des Fusionsproteins nach dem Integrationsereignis zur Selektion zu Nutze machen.

Die *attH2*-flankierenden DNA-Sequenzen könnten jedoch die Ausprägung eines transgenen Produkts beeinflussen. Es wäre daher von Vorteil, zwischen einer Reihe von gut charakterisierten „natürlichen“ Integrationsorten wählen zu können. Die Rekombinationsaktivität der *attH2*-Sequenz macht deutlich, dass Abweichungen von der *attB*-Sequenz von der *Lambda* Int toleriert werden. Daher ist es wahrscheinlich, dass weitere potenziell aktive, natürliche *attH*-Sequenzen im humanen Genom zu finden sind. Die kürzlich entwickelte Methode der „Substratgebundenen Protein Evolution“ eröffnet dazu noch die Möglichkeit *Lambda* Integrasen zu entwickeln, die neue Erkennungssequenzen besser binden und rekombinieren können (Buchholz & Stewart, 2001). Calos und Kollegen konnten bereits durch „DNA-Shuffling“ die Integrase ϕ C31 so verändern, dass sie mit einer 2-3fach höheren Effizienz eine „natürliche Sequenz“ aus dem Säugetiergenom rekombiniert (Sclimenti et al., 2001).

Anstelle einer im Genom von Säugetieren natürlich vorkommenden *att*-Sequenz könnte man auch die wild-typ *attB*-Sequenz ins Genom integrieren und als Zielsequenz zur sequenz-spezifischen Integration fremder DNA nutzen. Diese Strategie bietet den Vorteil, dass Reporter-Gene in verschiedene Regionen des Genoms integriert werden können, um beispielsweise den Einfluss flankierender DNA-Sequenzen zu untersuchen. Denkbar ist auch eine Anwendung dieser Strategie in humanen Stammzellen. Es könnten beispielsweise therapeutische Gene in gut charakterisierte Loci von etablierten Linien integriert werden, bevor diese in verschiedene Gewebetypen ausdifferenziert würden. Zudem bietet sie die Möglichkeit gezielt *loxP*- bzw. *FRT*-Sequenzen in das Genom einzubringen, um so weitere Manipulation durch andere Rekombinasen durchführen zu können.

Auch die Entwicklung des Kofaktors XIS in Säugetierzellen könnte das Spektrum der Einsatzmöglichkeiten des *Lambda* System erweitern. Mit seiner und der Hilfe des Kofaktors oIHf könnte man beispielsweise *att*-flankierte DNA-Abschnitte kontrolliert invertieren und dadurch Gene an- und abschalten. Diese Vielfalt der Anwendungsmöglichkeiten macht deutlich, dass auch das sequenz-spezifische *Lambda* Rekombinationssystem in Säugetierzellen sowohl für die Forschung als auch in der Medizin von großem Nutzen sein könnte.

F 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.

Adra, C.N., Boer, P.H., McBurney, M.W. (1987). Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene* **60**, 65-74.

Aman, E., Ochs, B., Abel, K.J. (1988). Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *E. coli*. *Gene*, **69**, 301.

Andersson, P., Goldfarb, M.P. & Weinberg, R.A. (1979). A defined subgenomic fragment of in vitro synthesized Moloney sarcoma virus DNA can induce cell transformation upon transfection. *Cell*, **16**, 63-75.

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.

Aranda, M., Kanellopoulou, C., Christ, N., Peitz, M., Rajewsky, K. & Dröge, P. (2001). Altered Directionality in the Cre-loxP Site-specific Recombination Pathway. *J. Mol. Biol.*, **311**, 453-459.

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 Rekombination. *J. Mol. Biol.*, **181**, 187-197.

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.

Boshart, M., Weber, F., Jahn, G., Dorsch-Häsler K., Fleckenstein B. & Schaffner W. (1985). A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell*, **41**, 521-530.

Buchholz, F., Ringrose, L., Angrand, P-O., Rossi, F. & Stewart, A.F. (1996). Different thermostabilities of FLP and Cre recombinases: implications of applied site-specific recombination. *Nucl. Acids Res.*, **24**, 4256-4262.

Buchholz, F., Angrand, P-O. & Stewart, A.F. (1998). Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.*, **16**, 657-662.

- Buchholz, F. & Stewart, A.F. (2001). Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nature Biotechnology*, **19**, 1047-1052.
- 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 Rekombination. *Cell*, **39**, 699-706.
- Bushman, W., Thompson, J.F., Vargas, L. & Landy, A. (1994). Control of Directionality in Lambda Site Specific Recombination. *Science*, **230**, 906-911.
- 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). Genetik manipulation of mouse embryonic stem cells by mutant λ integrase. *Genesis*, **32**, issue 3.
- 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. (submitted)
- 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. & Nash, H.A. (1984). E.coli integration host factor binds to specific sites in DNA. *Cell*, **39**, 707-716.
- 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.
- Diaz, V., Rojo, F., Martinez, A.C., Alonso, J.C. & Bernad, A. (1999). The prokaryotic beta-recombinase catalyzes site-specific recombination in mammalian cells. *J. Biol. Chem.*, **274**, 6634-6640.
- Dibyendu, S., Radman-Livaja, M. & Landy, A. (2001). The small DNA binding domain of λ integrase is a context-sensitive modulator of recombinase functions. *EMBO J.*, **20**, 1203-1212.

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.

Erickson, R.P. (1999). Antisense transgenics in animals. *Methods*, **18**, 304-310.

Esposito, D. & Scocca, J.J (1997). The integrase family of tyrosine recombinases: evolution of a conserve active site domain. *Nucl. Acids Res.*, **25**, 3605-3614.

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

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.

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

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

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.

Gu, H., Zou, J.R., & Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidence through Cre-loxP-mediated gene targeting. *Cell*, **73**, 1155-1164.

Gu, H., Marth, J.D., Orban P.C., Mossmann H. & Rajewsky, K. (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type specific gene targeting. *Science (Wash. DC)*, **265**, 103-106.

- Hahn, M., Piotukh, K., Borriss, R. & Heinemann, U. (1994). Native-like in vivo folding of a circularly permuted jellyroll protein shown by crystal structure analysis. *Proc. Natl Acad. Sci. USA*, **91**, 10417-10421.
- Halestrap, A.P. & Price, N.T. (1999). The proton-linked monocarboxylate transporter (MCT) family: structure, function and regulation. *Biochem. J.*, **343**, 281-299.
- Helene, C. & Toulme, J.J. (1990). Specific regulation of gene expression by anti-sense, sense, and antigene nucleic acids. *Biochim. Biophys. Acta*, **1049**, 99-125.
- Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet. Res.*, **5**, 282-304.
- 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.
- Kalderon, D. Roberts, B.L., Richardson, W.D. & Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. *Cell*, **39**, 499-509.
- Kay, M.A., Liu, D. & Hoogerbrugge, P. (1997). Gene Therapy. *Proc Natl. Acad. Sci., USA*, **94**, 12744-12746.
- 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. & Landy, A. (1992). Lambda Int protein bridges between higher order complexes at two distant chromosomal loci *attL* and *attR*. *Science*, **256**, 198-203.
- 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. (1988a) Bacteriophage lambda site-specific recombination proceeds with a defined order of strand exchange. *J. Mol. Biol.*, **204**, 95-107.
- Kolot, M. Silberstein, N. & Yagil, E. (1999). Site-specific recombination in mammalian cells expressing the Int recombinase of bacteriophage HK022. *Mol. Biol. Rep.*, **26**, 207-213.
- Kühn, R., Löhler, J., Remnick, D., Rajewsky, K. & Müller, W. (1993). Interleukin-10-deficient mice develop chronic enterocolitis. *Cell*, **75**, 263-274.
- Kühn, R., Frieder, S., Aguet, M., Rajewsky K. & Müller, W. (1995) Inducible Targeting in Mice. *Science*, **269**, 1427-1429.

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.

Logie, C. & Stewart, A.F. (1995). Ligand-regulated site-specific recombination. *Proc. Natl Acad. Sci. USA*, **92**, 5940-5944.

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.

Lorbach, E., Christ, N., Schwikardi, M. & Dröge, P. (2000). Site-specific Rekombination 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.

Marrow, B. & Kucherlapati, R. (1993). Targeting in Mammalian Cells by Homologous Recombination. *Curr. Opin. Biotech.*, **4**, 577-582.

Mao X., Fujiwara Y. & Orkin S.H. (1999). Improved reporter strain for monitoring Cre recombinase-mediated DNA excision in mice. *Proc. Natl Acad. Sci. USA*, **96**, 5037-5042.

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.

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.

Miller, D.G., Rutledge, E.A. & Russell, D.W. (2002). Chromosomal effects of adeno-associated virus vector integration. *Nature Genetics*, **30**, 147-148.

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, Jeffrey H. (1972). Experiments in Molecular Genetics. *Cold Spring Harbor Laboratory*

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.

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.

Morgan, R.A. & Anderson, W.F. (1993). Human Gene Therapy. *Annu. Rev. Biochem.*, **62**, 191-217.

Nagy A. (2000). Cre Recombinase: The Universal Reagent for Genome Tailoring. *Genesis*, **26**, 99-109.

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. (1981). Purification and properties of the E.coli protein factor required for λ integrative recombination. *J. Biol. Chem.*, **256**, 9246-9253.

Nash, H.A. Robertson, C.A., Flamm, E., Weisberg, R.A. & Miller, H.I. (1987). Overproduction of Escherichia coli Integration Host Factor, a Protein with Nonidentical Subunits. *J. Bacteriol.*, **169**, 4124-4127.

Nemir, M., Bhattacharyya, D., Li, X., Singh, K., Mukherjee, A.B. & Mukherjee, B.B. (2000). Targeted inhibition of osteopontin expression in the mammary gland causes abnormal morphogenesis and lactation deficiency. *J. Biol. Chem.*, **275**, 969-976.

Novak, A., Guo, C., Yang, W., Nagy, A. & Lobe C.G. (2000). Z/EG, a double reporter mouse line that express enhanced green fluorescence protein upon Cre-mediated excision. *Genesis*, **28**, 147-155.

Nunes-Düby, S.E., Smith-Mungo, L.I. & Landy, A. (1995). 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.

Oberto, J. & Rouviere-Yaniv, J. (1996). Serratia marcescens contains a heterodimeric HU protein like Escherichia coli and Salmonella typhimurium. *J. Bact.* **178**, 293-297.

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.

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.

Olivares, E.C., Roger, P.H. & Calos, M.P. (2001). Phage R4 integrase mediates site-specific integration in human cells. *Gene*, **278**, 167-176.

Paddinson, P.J., Candy, A.A. & Hannon, G.J. (2002). Stable suppression of gene expression by RNAi in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 1443-1448.

Parks, R.J., Chen, L., Anton, M., Sankar, U., Rudnicki, M.A. & Graham, F.L. (1996). A helper-dependent adenovirus vector system: removal of helper virus by Cre-mediated excision of the viral packaging signal. *Proc. Natl Acad. Sci. USA*, **93**, 13565-13570.

Pasparakis, M. & Kollias, G. (1995). Production of cytokine transgenic and knockout mice. *In Cytokines: A practical approach*, F.R. Balkwill, ed. (Oxford: IRL Press), pp. 297-325.

Patsey, R.L. & Bruist, M.F. (1995). Characterization of the interaction between the lambda intasome and attB. *J. Mol. Biol.*, **252**, 47-58.

Pfarr, D.S., Rieser, L.A., Woychik, R.P., Rottman F.M., Rosenberg, M. & Reff, M.E. (1986). Differential effects of polyadenylation regions on gene expression in mammalian cells. *DNA*, **5**, 115-122.

Phair, R.D. & Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature*, **354**, 604-609.

Plasterk, R.H. & Van de Putte, P. (1984). Genetic switches by DNA inversions in prokaryotes. *Biochim. Biophys. Acta*, **782**, 111-119.

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.

Puck et al. (1958). *J. Exp. Med.*, **108**, 945.

- 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 Rekombination Involves Capture of a Naked DNA by a Protein-DNA Complex. *Cell*, **52**, 9-17.
- Rose, R.E. (1988). The nucleotid sequence of pACYC177. *Nucl. Acids Res.*, **16**, 356.
- Ross, W. & Landy, A. (1983). Patterns of λ Int Recognition in the Regions of Strand Exchange. *Cell*, **33**, 261-272.
- 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.
- 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.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular Cloning, A laboratory manual. *Cold Spring Harbor Laboratory Press* (2nd Edition).
- Sanger, F., Nicklen, S., Coulson, A. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463-5467.
- Sarkar, D. Radman-Livaja, M. & Landy, A. (2001). The small DNA binding domain of lambda integrase is a context-sensitive modulator of recombinase function. *EMBO J.*, **20**, 1203-1212.
- Sauer, B. (1993) Manipulation of Transgenes by Site-Specific Recombination: Use of Cre Recombinase. *Methods Enzymol.*, **225**, 890-900.
- 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.
- 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.
- 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.
- Scimienti, C.R., Thyagarajan, B. & Calos, M.P. (2001). Directed evolution of a recombinase for improved genomic integration at a native human sequence. *Nucleic Acid Research*, **29**(24), 5044-5051.
- 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.
- 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.
- 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.
- Spector, D.L., Goldman R.D. & Leinwand, L.A. (1998). *Cells: a laboratory manual*. Cold Spring Harbor Laboratory Press.
- Spengler, S.J., Stasiak, A. & Cozzarelli, N.R. (1985). The stereostructure of knots and catenanes produced by phage lambda integrase recombination: implications for mechanism and DNA structure. *Cell*, **42**, 325-334.
- 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.
- Thyagarajan, B., Guimaraes, M.J., Groth, A.C. & 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. & 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.
- Torres, R.M. & Kühn, R. (1997). *Laboratory protocols for conditional gene targeting* (Oxford: Oxford University Press).

Tucker, K.L., Wang Y., Dausman, J. & Jaenisch, R. (1997). A transgenic mouse strain expressing four drug-selectable marker genes. *Nucleic Acid Research*, **25**(18), 3745-3746.

Varga, L.V., Toth, S., Novak, I. & Falus, A. (1999). Antisense strategies: functions and applications in immunology. *Immunol. Lett.*, **69**, 217-224.

Vasquez, K.M., Marburger, K., Intody, Z. & Wilson, J.H. (2001). Manipulating the mammalian genome by homologous recombination. *Proc. Natl Acad. Sci. USA*, **98**, 8403-8410.

Verma, I.M. & Somia, N. (1997). Gene therapy-promises, problems and prospects. *Natur*, **389**, 239-242.

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.

Werdien, D., Peiler, G. & Ryffel, G.U. (2001). FLP and Cre recombinase function in *Xenopus* embryos. *Nucl. Acids Res.*, **29**, E53-3.

Wolff, J.A. & Griffiths, E. (1994). Gene Therapeutics. *Birkhäuser Boston*. ISBN 0-8176-3650-1.

Wu, Z., Gumpert, R.I. & Gardner, J.F. (1997). Genetic analysis of second-site revertants of bacteriophage lambda integrase mutants. *J. Bacteriol.*, **179**, 4030-4038.

Yang, C.C. & 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). Interactions of the lambda site-specific protein XIS with attachment site DNA. *Proc. Natl Acad. Sci. USA*, **82**, 1040-1044.

G 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 AAACCCCTCCG

27431  CAAAGTGGGT GGATCGCTTG AGTCCAGTTC GAGACCAGCC TGGGCAACAT
      GTTTCACCCA CCTAGCGAAC TCAGGTCAAG CTCTGGTTCGG 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 AGAGTTTTTAT 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

```

```

27831   TGATTCTAAA CTATTACGTT TTTCTTTTAG GTTTGGGTC TGCTTTCTTA
        ACTAAGATTT GATAATGCAA AAAGAAAATC CAAACCCAAG ACGAAAGAAT
                                  ←-----→
27881   TACCAAGTGG CTGCTGTGGT AACTACCAA TACTTCAAAA AACGATTGGC
        ATGGTTCACC GACGACACCA TTGATGGTTT ATGAAGTTTT TTGCTAACCG
        ←-----→
27931   TCTTTCTACA GCTATTGCC GTCCTGGGAT GGGACTGACT TTTCTTTTGG
        AGAAAGATGT CGATAACGGG CAAGACCCTA CCCTGACTGA AAAGAAAACC
        ←-----→
27981   CACCCTTTAC AAAATTCTCG ATAGATCTGT ATGACTGGAC AGGTACGTCA
        GTGGGAAATG TTTTAAGGAC TATCTAGACA TACTGACCTG TCCATGCAGT

28031   TTCTAAGTTT ATTTACTATC AACTGTGCAA-3'
        AAGATTCAA TAAATGATAG TTGACACGTT-5'

```

NNNNNN ≅ Exon X **NNNNNN** ≅ Exon Y

Die Pfeile markieren die Nucleotidsequenz der Primer MCTII und MCT3'.

Komplette kodierende Sequenz um den *atfH2*-Locus (*human putative monocarboxylate transporter (MCT5)*): Nucleotide 1 bis 750: Sequenzinformation aus der EMBL-Datenbank (*accession no.* U59185)

```

                                MCT5'
                                →-----→
1   5' -CTTGGCTCTT ACAATGCTCA CTTGTTTTCA CAATGCAGCA AAATGAAATG
   3' -GAACCGAGAA TGTTACGAGT GAACAAAAGT GTTACGTCGT TTTACTTTAC

51   CCTTAGAAAA AGAGTAACAT TCCAGAAAAC GGTGTAATTT ATTTTCTTTC
   GGAATCTTTT TCTCATTGTA AGGTCTTTTG CCACATTAAA TAAAAAGAAG

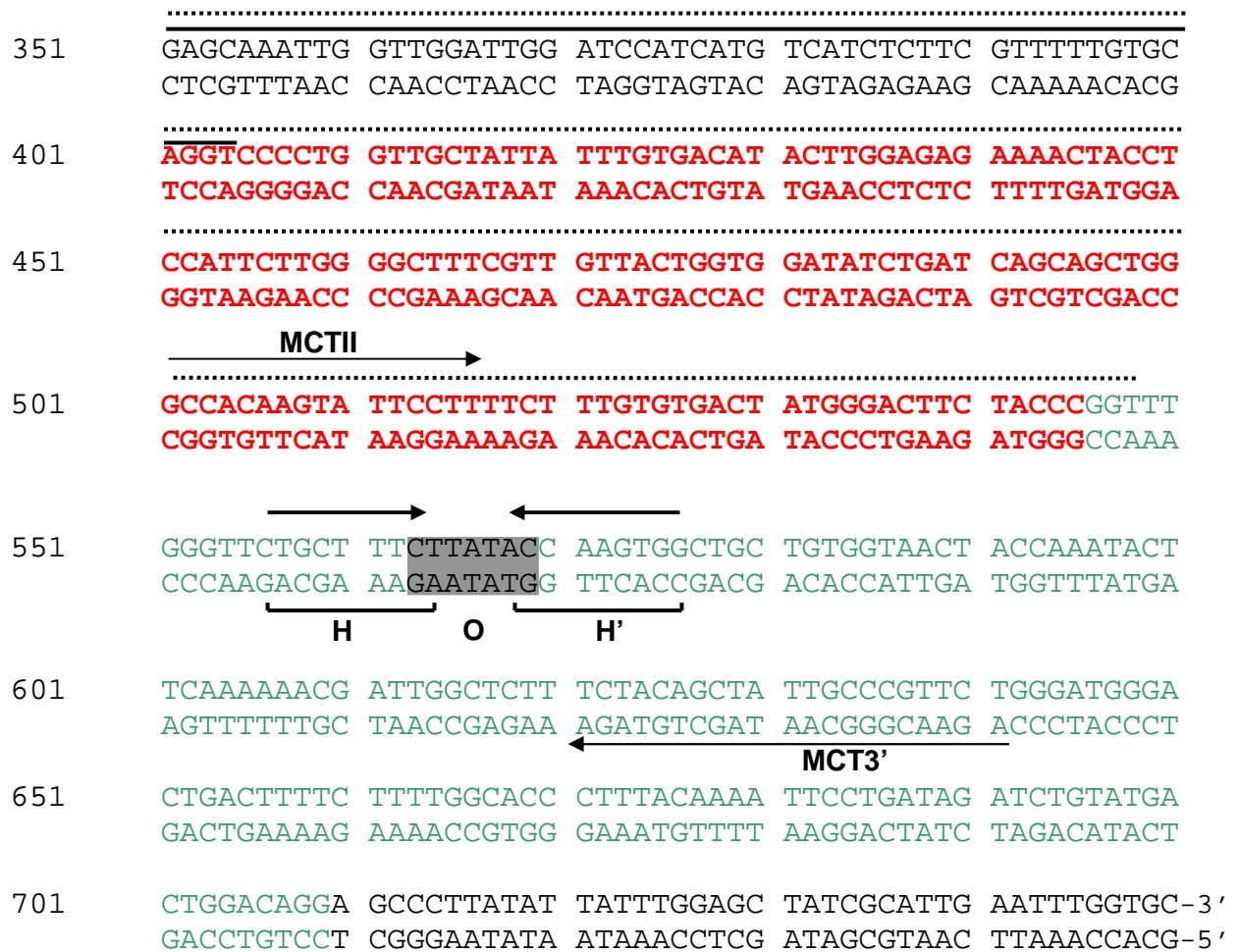
101  CTTAATTGCC CCATCTGTGG AGGATTTCTT TGCTGAACAC CACATCAAAG
   GAATTAACGG GGTAGACACC TCCTAAAGAA ACGACTTGTG GTGTAGTTTC

151  GGATCTTCTG CATTTAAAAT AGAAGAGGCA TCATGCTGAA GAGGGAGGGG
   CCTAGAAGAC GTAAATTTTA TCTTCTCCGT AGTACGACTT CTCCCTCCCC

201  AAGGTCCAAC CTTACACTAA AACCTGGAT GGAGGATGGG GATGGATGAT
   TTCCAGGTTG GAATGTGATT TTGGGACCTA CCTCCTACCC CTACCTACTA

251  TGTGATTCAT TTTTTCCTGG TGAATGTGTT TGTGATGGGG ATGACCAAGA
   AACTAAGTA AAAAAGGACC ACTTACACAA AACTACCCC TACTGGTTCT
   .....
301  CTTTTGCAAT TTTCTTTGTG GTCTTTCAAG AAGAGTTTGA AGGCACCTCA
   GAAAACGTTA AAAGAAACAC CAGAAAGTTC TTCTCAAAC TCCGTGGAGT

```



NNNNNN ≅ Exon X

NNNNNN ≅ Exon Y

————— ≅ ~132 Nucleotide die im RT-PCR Produkt 2 fehlen (Abschn. D6.3)

..... ≅ ~290 Nucleotide die im RT-PCR Produkt 3 fehlen (Abschn. D6.3)

Die Pfeile markieren die Nucleotidsequenz der Primer MCT5', MCTII und MCT3'.

Komplette Nukleinsäure-Sequenz der vier oIHf Varianten:

oIHf1(Linker 1 und Linker 3); oIHf2 (Linker 2 und 4); oIHf3 (Linker 1 und Linker 4);
oIHf4 (Linker 2 und Linker 3)

```

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
                                Linker 2
                                Linker 1

151 ATGTCAGAAT ATCTGTTTGA TAAGCTTGGG CTTAGCAAGC GGGATGCCAA
    TACAGTCTTA TAGACAAACT ATTCGAACCC GAATCGTTCC CCCTACGGTT

201 AGAACTGGTT GAACTGTTTT TCGAAGAGAT CCGTCGCGCT CTGGAAAACG
    TCTTGACCAA CTTGACAAAA AGCTTCTCTA GGCAGCGCGA GACTTTTTGC

251 GCGAACAGGT GAAACTCTCT GGTTTTGGTA ACTTCGATCT GCGTGATAAG
    CGCTTGTCCTA CTTTGAGAGA CCAAACCAT TGAAGCTAGA CGCACTATTC

301 AATCAACGCC CGGGACGTAA CCCGAAAACG GCGGAGGATA TTCCCATTAC
    TTAGTTGCGG GCCCTGCATT GGGCTTTTGC CCGCTCCTAT AAGGGTAATG

351 AGCACGGCGC GTGGTGACCT TCAGACCCGG GCAGAAGTTA AAAAGCCGGG
    TCGTGCCGCG CACCACTGGA AGTCTGGGCC CGTCTTCAAT TTTTCGGCCC

401 TCGAAAACGC TGGTGGG GGC GAGCGTATTG AAATCCGCGG TTTCGGCAGT
    AGCTTTTTCG ACCACCCCG CTCGCATAAC TTTAGGCGCC AAAGCCGTCA
                                Linker 4
                                Linker 3

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'

```

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 GTTAGTGA CTGATATGTC

151 GCCAACATTG AGTTATTTTC AGGACACAAA CACAAGCCTC TGACAGCGAG
 CGGTTGTAAC TCAATAAAAG TCCTGTGTTT GTGTTCCGAG ACTGTCGCTC

201 AATCAACAGT GATAATTCCG TTACGTTACA TTCATGGCTT GATCGCTACG
 TTAGTTGTCA CTATTAAGGC AATGCAATGT AAGTACCGAA CTAGCGATGC

251 AAAAAATCCT GGCCAGCAGA GGAATCAAGC AGAAGACACT CATAAATTAC
 TTTTTTAGGA CCGGTCGTCT CCTTAGTTCG TCTTCTGTGA GTATTTAATG

301 ATGAGCAAAA TTAAAGCAAT AAGGAGGGGT CTGCCTGATG CTCCACTTGA
 TACTCGTTTT AATTTCTGTTA TTCCTCCCA GACGGACTAC GAGGTGAACT

351 AGACATCACC AAAAAAGAAA TTGCGGCAAT GCTCAATGGA TACATAGACG
 TCTGTAGTGG TGTTTTCTTT AACGCCGTTA CGAGTTACCT ATGTATCTGC

401 AGGGCAAGGC GCGTCAGCC AAGTTAATCA GATCAACACT GAGCGATGCA
 TCCCGTTCCG CCGCAGTCGG TTCAATTAGT CTAGTTGTGA CTCGCTACGT

451 TTCCGAGAGG CAATAGCTGA AGGCCATATA ACAACAAACC ATGTCGCTGC
 AAGGCTCTCC GTTATCGACT TCCGGTATAT TGTTGTTTGG TACAGCGACG

AAG (E174K)

501 CACTCGCGCA GCAAATCAG AGGTAAGGAG ATCAAGACTT ACGGCTGACG
 GTGAGCGCGT CGTTTTAGTC TCATTCTC 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

701 GCAAACAGG CGTAAAAATT GCCATCCCAA CAGCATTGCA TATTGATGCT
 CGTTTTGTCC GCATTTTTAA CGGTAGGGTT GTCGTAACGT ATAACCTACGA

751 CTCGGAATAT CAATGAAGGA AACACTTGAT AAATGCAAAG AGATTCTTGG
 GAGCCTTATA GTTACTTCCT TTGTGAACTA TTTACGTTTC TCTAAGAACC

801 CGGAGAAACC ATAATTGCAT CTACTIONGTCG 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 '

H Zusammenfassung

Der Transfer von fremder genetischer Information in kultivierte eukaryotische Zellen und intakte höhere Organismen erfordert die Entwicklung neuer Methoden, die eine gezielte Veränderung im Genom ermöglichen. Sequenz-spezifische Rekombinationssysteme stellen für diesen Zweck eine vielversprechende Alternative zur homologen Rekombination dar. Das hier vorgestellte Projekt hatte zum Ziel, das sequenz-spezifische Rekombinationssystem des Bakteriophagen *Lambda* für den Einsatz in Säugetierzellen zu optimieren. Die Verwendung der Integrase(Int)-Mutanten Int-h und Int-h/218 ermöglichte es, auch ohne die in *E. coli* natürlich vorkommenden Kofaktoren gezielte genetische Veränderungen in Säugetierzellen vorzunehmen. Durch Positionierung der Erkennungssequenzen für die integrative (*attBxattP*) oder exzisive (*attLxattR*) Rekombinationsreaktion wurde mit Hilfe der Int-Mutanten die DNA im Genom von Säuger-Zellen verändert: Erstens konnten DNA-Fragmente in einer humanen Zelllinie invertiert werden. Zweitens gelang es, einen Neomycin-Resistenzmarker, der zuvor durch *Gene Targeting* stabil ins Genom integriert wurde, in embryonalen Stammzellen der Maus zu deletieren. Im Gegensatz zu Cre und FLP lässt sich die Richtung der Rekombinationsreaktion sowohl von der wild-typ Int als auch von den Int-Mutanten in *E.coli* und *in vitro* durch spezifische Kofaktoren steuern. Der Kofaktor IHF wurde im Rahmen dieser Arbeit für die Anwendung in Säuger-Zellen genetisch modifiziert und zur Stimulierung der integrativen Rekombination in embryonalen Stammzellen der Maus eingesetzt. Außerdem ergab die Untersuchung der Int-vermittelten Rekombination folgende Erkenntnisse über den Ablauf der Reaktion in menschlichen Zellen: 1) Die Arm-Bindungssequenzen in den *att*-Regionen wirken stimulierend auf die Rekombination. 2) Die intermolekulare Rekombination von episomalen DNA-Segmenten durch die Integrase ist mindestens ebenso effizient wie die entsprechende intramolekulare Rekombination. 3) Die minimale integrative Rekombinationsaktivität der wild-typ Integrase in menschlichen Zellen kann durch die Transfektion von IHF-DNA Komplexen stimuliert werden. 4) Endogene Faktoren in Säugetierzellen koennen die natürlichen Kofaktoren aus *E.coli* nicht ersetzen. Eine im menschlichen Genom natürlich vorkommende *att*-Sequenz (*attH2*) erwies sich als rekombinationsaktiv. Eine solche Zielsequenz wäre für die Entwicklung von gentherapeutischen Techniken mit dem *Lambda* System von Vorteil.

I Summary

The transfer of genetic material into eukaryotic cells and intact higher organisms requires new methods to perform controlled genetic manipulations. For this, site-specific recombination represents a promising alternative to homologous recombination. The project presented in this thesis had the goal of optimizing the site-specific recombination system of bacteriophage *lambda* for its application in mammalian cells. The mutant integrases Int-h and Int-h/218 catalysed sequence-specific DNA manipulation in the absence of natural accessory factors in mammalian cells. In addition, the specific positioning of target sites for integrative (*attBxattP*) or excisive (*attLxattR*) recombination reactions in the genome of mammalian cells allowed defined genetic manipulations: First, a DNA-fragment was inverted in human cells. Secondly, it was shown that a neomycin resistance marker gene, which had been introduced into the genome by gene targeting, could be deleted in mouse embryonic stem cells. In contrast to Cre and FLP, the direction of recombination reactions catalyzed by wildtype and mutant integrases in *E.coli* and *in vitro* can be controlled by accessory factors. The accessory factor IHF was optimized for its application in mammalian cells and was used to stimulate integrative recombination in mouse embryonic stem cells. Furthermore, the analysis of recombination reactions catalyzed by integrases in human cells revealed the following results: 1) Arm-binding sites in the *att*-regions stimulate recombination. 2) Intermolecular recombination of episomal DNA-segments is at least as efficient as the corresponding intramolecular reaction. 3) Marginal integrative recombination activity of wildtype integrase can be stimulated by transfection of IHF-DNA complexes. 4) Endogenous factors in mammalian cells cannot substitute the natural accessory factors from *E.coli*. An endogenous integrase target sequence in the human genome, *attH2*, was shown to be active in recombination. Such a target sequence could be useful for the development of genetherapeutic techniques with the *lambda* system.

Lebenslauf

Persönliche Daten

Name	Nicole Christ
Geburtsort	Boppard (Bundesrepublik Deutschland)
Geburtsdatum	12. September 1972
Familienstand	ledig
Eltern	Richard Christ; Monika Christ, geb. Schneider
Anschrift	Merowingerstraße 31; 50677 Köln

Ausbildung

1979-1983	Grundschule Halsenbach
1983-1992	Kant-Gymnasium Boppard
Juni 1992	Allgemeine Hochschulreife
August 1992- Dez. 1992	Praktikum im mikrobiologischen Labor der Firma Seba-med in Boppard
Febr. 1993-Juli 1993	Französischstudium an der Universität Sorbonne in Paris, Frankreich
WS 93/94	Immatrikulation im Fach Biologie an der Universität zu Köln
Oktober 1995	Diplom-Vorprüfung
Feb. 1998- Mai 1998	Praktikum im <i>Pharmaceutical Research Center</i> der Firma Bayer in Wuppertal
Dezember 1998	Diplom-Hauptprüfung in den Fächer: Genetik, Biochemie und Pharmakologie
März 1998-Dez.1998	Diplomarbeit am Institut für Genetik, Universität zu Köln unter der Anleitung von Priv.-Doz. Dr. Peter Dröge. Thema: „Untersuchungen einer neuen Form des DNA-Strang austauschs, die durch zwei Mutanten der λ -Integrase katalysiert wird“
März 1999-Febr. 2002	Dissertation am Institut für Genetik, Universität zu Köln unter der Anleitung von Priv.-Doz. Dr. Peter Dröge. Thema: „Gezielte genetische Manipulation durch <i>Lambda</i> Integrasen in embryonalen Stammzellen der Maus und in anderen Säugetierzellen“

Danksagungen

Das in dieser Dissertation beschriebene Projekt wurde zwischen März 1999 und Februar 2002 am Institut für Genetik der Universität zu Köln unter Anleitung von Priv.-Doz. Dr. Peter Dröge durchgeführt. Das Projekt wurde durch den Boehringer Ingelheim Fonds 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 allen Kollegen am Institut für Genetik bedanken, die auf die eine oder andere Weise zum Erfolg dieses Projektes beigetragen haben. Insbesondere danke ich am Institut für Genetik Ivo Lieberam, Elke Lorbach, Angela Egert, Teresa Corona, Micha Schwikardi, Dennis Webb, Thomas Wunderlich und Frank Edenhofer. Karin Otto und Christoph Göttlinger danke ich für technische Unterstützung.

Benno Müller-Hill (Universität Köln), Klaus Rajewsky (Universität Köln), Philippe Soriano (Seattle, USA), Arthur Landy (Providence, USA), Dominik Esposito (Invitrogen), Steven D. Goodman (Los Angeles, USA) und Thomas Schwartz (Rockefeller University, USA) danke ich für die Bereitstellung von Reagenzien und ihren wissenschaftlichen Rat.

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 Priv.-Doz. Dr. P. Dröge betreut worden.

Köln, den 26.02.2002

Nicole Christ

Teilpublikationen:

- 1.) Christ, N. & Dröge, P. (1999). Alterations in the Directionality of λ Site-specific Recombination Catalyzed by Mutant Integrases *in Vivo*. **J. Mol. Biol.**, 288, 825-836.
- 2.) 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.
- 3.) Aranda, M.* , Kanellopoulou, C.* , Christ, N.* , Peitz, M., Rajewsky, K. & Dröge, P. (2001) Altered Directionality in the Cre-loxP Site-specific Recombination Pathway. **J. Mol. Biol.**, 311, 453-459.
- 4.) 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.
- 5.) Christ, N. & Dröge, P. (2002). Genetic manipulation of mouse embryonic stem cells by mutant λ integrase. **Genesis**, 32, 203-208.
- 6.) 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.**, im Druck.

* Gleichstellung der Autoren