

**Analyse der transkriptionellen Regulation
des humanen endogenen Retrovirus HTDV/HERV-K**

DISSERTATION
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Biologie und Informatik
der Johann Wolfgang Goethe-Universität in Frankfurt am Main

von
Martin Kraft
aus Würzburg

Frankfurt am Main, 2004

DF1

vom Fachbereich Biologie und Informatik der Johann Wolfgang Goethe-Universität
als Dissertation angenommen.

Dekan: Prof. Dr. Heinz Dieter Osiewacz

Gutachter: Prof. Dr. Johannes Löwer
Prof. Dr. Anna Starzinski-Powitz

Datum der Disputation:

Inhaltsverzeichnis:

1	EINLEITUNG	1
1.1	Retroviren	1
1.1.1	Allgemein	1
1.1.2	Struktur	2
1.1.3	Exogene versus Endogene Retroviren.....	6
1.1.4	HERV-K.....	7
1.2	Transkription.....	9
1.2.1	RNA-Polymerasen	9
1.2.2	Der Promotor.....	12
1.2.2.1	Core Promotor	12
1.2.2.2	Regulatorischer Promotor	13
1.2.2.3	Enhancer	14
1.2.3	Transkriptionskontrolle durch epigenetische Mechanismen.....	14
1.2.4	Transkription von Retroviren	14
1.2.4.1	Allgemein	14
1.2.4.2	HERV-K	16
1.3	Zielsetzung der Arbeit	17
2	ERGEBNISSE	18
2.1	Analyse von HERV-K LTRs	18
2.1.1	Charakterisierung der R-Region.....	22
2.1.1.1	3'-Deletionen	23
2.1.1.2	Funktionalität der TATA-Box	25
2.1.1.3	Identifikation des Transkriptionsstarts durch eine 5'-RACE	26
2.2	Die Initiation der Transkription	30
2.2.1	Sp1 und Sp3 binden an GC/GT-Boxen stromaufwärts vom Inr	31
2.2.2	EMSA-Analyse der ersten Sp-Bindungsstelle	34
2.2.3	EMSA-Analyse der zweiten Sp-Bindungsstelle	36
2.2.4	Western Blot Analyse von Sp1 und Sp3	38
2.2.5	Analyse der Funktionalität der Sp-Bindungsstellen <i>in-vivo</i>	40
2.2.6	Sp-Bindungsstellen-Screening	42
2.3	Analyse der U3-, R- und U5-Region	42
2.3.1	Mutationsanalyse U3.....	43

2.3.2	Detailanalyse der Punktmutationscluster M0 – M5	45
2.3.3	Mutationsanalyse R-U5.....	47
2.4	Transkription durch Testis-spezifische Faktoren	49
2.4.1	SRY	50
2.5	Epigenetische Mechanismen	52
3	DISKUSSION	54
3.1	Die HERV-K LTR.....	54
3.1.1	Allgemein	54
3.1.2	Aktiv oder Inaktiv?.....	54
3.2	Die Transkriptionsregulation von HERV-K	57
3.2.1	TATA-Box versus Inr	57
3.2.2	Die Rekrutierung von TFIID.....	58
3.2.3	Regulation durch Sp1/Sp3.....	60
3.2.4	Regulatorischer Promotor und Enhancer	65
3.2.4.1	U3	65
3.2.4.2	R.....	66
3.2.4.3	U5	67
3.3	Die Testis-spezifische Expression.....	67
3.3.1	Differentielle Sp1-Genexpression	68
3.3.2	SRY	69
3.3.3	Epigenetische Mechanismen	72
3.4	Modell zur Transkription von HERV-K.....	74
3.5	Ausblick.....	77
4	MATERIAL UND METHODEN	80
4.1	Häufig benutzte Geräte und Materialien.....	80
4.2	Zellkultur	81
4.2.1	Geräte und Materialien.....	81
4.2.2	Medien und Chemikalien	81
4.2.3	Zellkultur.....	82
4.2.4	Grundhaltung von Zellen	82
4.2.5	Transfektion eukaryotischer Zellen.....	82
4.3	Bakterienkultur	83
4.3.1	Geräte	83

4.3.2	Medien und Chemikalien	83
4.3.3	Bakterienstämme	83
4.3.4	Präparation kompetenter Bakterienzellen	84
4.3.4.1	Transformation rekombinanter DNA	84
4.3.5	Plasmid Präparation aus Bakterien.....	84
4.3.5.1	Plasmid Mini Präparation	84
4.3.5.2	Plasmid MIDI Präparation.....	85
4.4	Nukleinsäure-Methoden	85
4.4.1	Präparation von RNA aus eukaryotischen Zellen	85
4.4.1.1	Präparation von Gesamt-RNA.....	85
4.4.1.2	Präparation von Poly(A)-RNA	85
4.4.2	Polymerase Chain Reaction (PCR)	86
4.4.3	Oligonukleotid-Hybridisierung	87
4.4.4	Restriktion von DNA	87
4.4.5	Generierung von Blunt Ends	87
4.4.6	Dephosphorylierung des Vektors	88
4.4.7	Phenol-Chloroform-Extraktion	88
4.4.8	Fällung von DNA	89
4.4.9	Agarosegel-Elektrophorese	89
4.4.9.1	Isolierung von DNA aus Agarose.....	90
4.4.9.2	Reinigung und Isolierung von DNA aus Low Melting Point (LMP) Agarose Gelen.....	90
4.4.10	Ligation.....	90
4.4.11	Rapid Amplification of 5'-cDNA ends (5'-RACE).....	91
4.4.12	Radioaktive Markierung von DNA	92
4.4.13	Methylierung von DNA.....	93
4.4.14	DNA-Sequenzierung.....	93
4.5	Protein-Methoden	93
4.5.1	Präparation von Protein-Extrakten aus eukaryotischen Zellen	93
4.5.2	Aufreinigung rekombinanter Proteine aus eukaryotischen Zellen.....	95
4.5.3	SDS-Polyacrylamid Gel Elektrophorese (SDS-PAGE)	95
4.5.4	Western-Transfer-Analyse	96
4.5.5	Dualer Luziferase-Assay	97

4.6	Electrophoretic Mobility Shift Assay (EMSA) und Supershift-Assay	98
4.6.1	Protein-Nukleinsäure Bindungs-Reaktion.....	98
4.6.2	Native Polyacrylamid Gel Elektrophorese	98
4.7	Rekombinante Plasmide	99
4.7.1	Klonierungsstrategien.....	99
4.7.1.1	Verwendete Oligonukleotide	99
4.7.1.2	Konstruktion der rekombinanten Plasmide	100
5	ZUSAMMENFASSUNG	106
6	LITERATURVERZEICHNIS	108
7	ANHANG	122
7.1	Sequenzvergleich der HERV-K LTRs pcK30, LTR21, K113 und K115	122
7.2	Konsensussequenzen	123
7.3	Auszüge aus der Uniprot-SwissProt-Datenbank.....	124
7.3.1	Sp1.....	124
7.3.2	Sp3.....	126
7.3.3	SRY	129
8	ABKÜRZUNGSVERZEICHNIS.....	133
	DANKSAGUNG	136
	LEBENS LAUF	138

1 Einleitung

1.1 Retroviren

1.1.1 Allgemein

Retroviren erlangten ihre größte Bekanntheit in den achtziger Jahren durch die weltweite Verbreitung des humanen Immundefizienzvirus (HIV) und das damit verbundene erworbene Immundefizienz-Syndrom (AIDS) (Gallo 1988). Seit Beginn des 20. Jahrhunderts wurden jedoch bereits wichtige Entdeckungen gemacht, die Retroviren mit verschiedenen Tumoren in Verbindung brachten. In den siebziger Jahren schließlich brachte die Entdeckung der Reversen Transkriptase (RT) die Aufklärung des Infektionszyklus von Retroviren einen entscheidenden Schritt weiter.

Retroviren wurden Anfang des 20. Jahrhunderts entdeckt. 1908 konnten Ellermann und Bang zeigen, dass die Hühner-Leukämie durch ein Virus verursacht wird (Ellermann und Bang 1908). Drei Jahre später, im Jahr 1911, demonstrierte Rous die zellfreie Übertragung von Hühnersarkomen durch gefilterte Extrakte auf gesunde Hühner. Die Übertragbarkeit der Krankheit durch ein filtrierbares Agens unterschied sie klar von bakteriellen Infektionen. Das Virus wurde nach seinem Entdecker Rous-Sarkom-Virus (RSV) genannt (Rous 1911). Der Entdeckung von Ellermann und Bang konnte man das aviäre Leukämie-Virus (ALV) zuordnen. Beide Viren zusammen bilden das Virusgenus der aviären Sarkom-Leukämie-Viren (ASLV). Man nannte diese Viren wegen ihrer Verbindung zu Tumorerkrankungen auch Onkoviren. 1936 konnte man den ersten virusinduzierten Tumor bei Säugetieren feststellen (Bittner 1936). Bittner zeigte, dass eine Form von Brustkrebs der Maus durch das Maus-Mamma-Tumor-Virus (MMTV) hervorgerufen wird. 1955 zeigte dann Mühlbock bei MMTV erstmals, dass Retroviren sich nicht nur horizontal über Infektion in Form von Viruspartikeln verbreiten können, sondern dass sie auch als endogener Bestandteil des Wirtsgenoms von den Eltern zu den Nachkommen vertikal übertragen werden können (Mühlbock 1955).

Ein retrovirales Viruspartikel ist etwa 80-100 nm im Durchmesser groß und besitzt zwei Kopien eines RNA-Genoms mit einer Größe von 7-12 kB, die vom Viruscapsid umhüllt sind. Das Genom in Form einer mRNA ist linear, einzelsträngig, nicht segmentiert und von positiver Polarität. Eine Lipiddoppelmembran, die sich von der Membran der Wirtszelle ableitet, bildet die abschließende Virushülle. In dieser Membran sind die viralen Hüllproteine eingebettet.

Retroviren infizieren Zellen, indem sie über Oberflächen-Glykoproteine mit spezifischen Rezeptoren auf der Plasmamembran der Wirtszelle interagieren. Es wird eine Fusion mit der Plasmamembran der Zelle eingeleitet und das genetische Material des Virus wird in das Cytoplasma entlassen. Nach diesem initialen Schritt, der vielen Viren gemein ist, tritt ein Ereignis ein, das die Retroviren deutlich von anderen Virusgruppen unterscheidet und sie zu etwas Besonderem macht: Die Reverse Transkription. Mit der Entdeckung der Reversen Transkription wurde klar, dass ein zentrales Dogma der Zellbiologie ins Wanken geriet. Ging man bis Anfang der siebziger Jahre davon aus, dass die zentralen Vorgänge Transkription und Translation, die die genetische Information vom Gen über die RNA in Protein umwandeln, von der Zelle nur in einer Richtung besritten werden können, zeigte die Erforschung der Retroviren ein neues Enzym auf. Die RT wurde 1970 von Temin, Mizutani und Baltimore entdeckt (Baltimore 1970; Temin und Mizutani 1970). Dieses Enzym ist in der Lage die genetische Information des Virus, die als Einzelstrang-RNA vorliegt, in doppelsträngige DNA umzuschreiben. Diese kann dann in das Wirtsgenom integriert werden, das Virus liegt nun als Provirus im zellulären Genom vor. Von hier aus kann die genetische Information wieder abgelesen werden. Die transkribierten Virusprodukte führen zur Bildung neuer Viruspartikel, die schließlich wieder in den extrazellulären Raum entlassen werden. Eine erneute Infektion kann stattfinden.

1.1.2 Struktur

Retrovirusgenome können einen einfachen oder komplexen Aufbau haben. Die Genome von einfachen Retroviren wie RSV oder MMTV werden in ihrer proviralen Transkription ausschließlich vom viralen Promotor, der LTR, und den zellulären Faktoren der Transkriptionsmaschinerie gesteuert. Komplexe Retroviren wie HIV oder wie das humane T-Zell Leukämie Virus (HTLV) besitzen zusätzlich offene Leserahmen (ORF) für Transaktivatoren. Diese entstehen durch einfache bzw. mehrfache Spleißereignisse. Viruscodierte Proteine wie Tat oder Rev bei HIV können die transkriptionelle und posttranskriptionelle Regulation des Virusgenoms direkt beeinflussen. Tat bindet an eine spezielle 5'-Region am 5'-Ende des viralen Transkripts, stabilisiert den Initiationskomplex bei Beginn der Transkription und ermöglicht die Elongation. Rev ist posttranskriptionell wirksam. Es interagiert mit partiell- und ungespleißten viralen mRNA-Spezies und begünstigt ihren Export ins Cytoplasma (Cullen 1992; Rosen 1992). Das Virus hat also eine gewisse Kontrolle über seine Expression.

Im Folgenden wird auf den Aufbau eines einfachen Retrovirus eingegangen. Das Genom eines prototypischen Retrovirus ist in Abbildung 1-1 dargestellt:

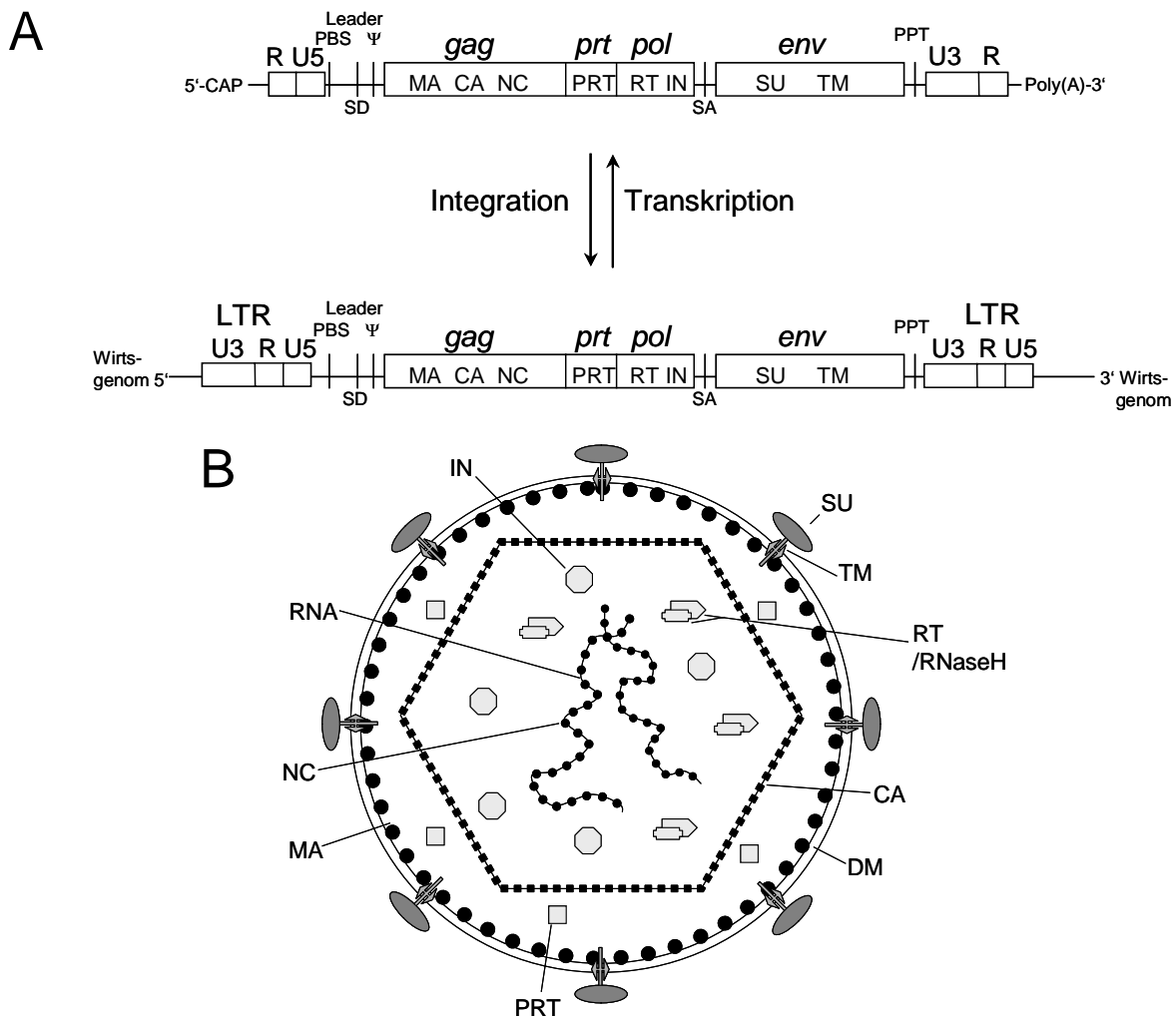


Abbildung 1-1: Schematischer Aufbau eines Retrovirus. A Genomischer Aufbau. Dargestellt sind das retrovirale Genom und die provirale Organisation im Zellgenom. **B Aufbau eines Viruspartikels.** Dargestellt ist das Virion mit allen zugehörigen Proteinen und dem Genom. **gag** Gruppenspezifische Antigene, **pol** Polymerase, **env** Envelope, **MA** Matrix, **CA** Capsid, **NC** Nukleocapsid, **PRT** Protease, **RT** Reverse Transkriptase, **IN** Integrase, **SU** Surface-Protein, **TM** Transmembranprotein, **SA** Spleißakzeptor, **SD** Spleißdonor, **PBS** Primer Binding Site, **PPT** Poly-Purin Tract, ψ Verpackungssignal, **DM** Lipiddoppelmembran, **LTR** Long Terminal Repeat, **R** Terminal redundancy, **U5** Unique 5'-end, **U3** Unique 3'-end.

R (Terminal redundancy)

Die R-Region umfasst redundante Sequenzen am 5'- und 3'-Ende des viralen Genoms. Bei der retroviralen Replikation dient die 5'-R-Region zusammen mit der U5-Region auf dem RNA-Genom als Template zur Generierung eines DNA-Primers durch die RT. Die Redundanz der R-Region ermöglicht die komplementäre Hybridisierung des Primers an die

3'-R-Region und initiiert somit die Elongation des DNA-Erststranges über das gesamte Virusgenom hinweg.

Der Beginn der R-Region wird auf der integrierten Provirus-DNA durch den Transkriptionsstart definiert. Sie enthält Bindungsstellen für die RNA-Polymerase II (PolII), evtl. ein „downstream promoter element“ (DPE) und oft auch das Polyadenylierungssignal (Poly(A)-Signal). Das Ende der R-Region wird durch die Polyadenylierungsstelle (Poly(A)-Stelle) festgelegt. Das Poly(A)-Signal unterscheidet sich nur selten von dem zellulären Signal mit der Konsensussequenz AATAAA. Es stellt das Erkennungssignal für die Poly(A)-Polymerase dar. Die Polymerase hängt ca. 250 Adenosine ab der Poly(A)-Stelle an das 3'-Ende des Transkripts. Die Poly(A)-Stelle liegt 10-30 Basenpaare (bps) stromabwärts vom Poly(A)-Signal und hat oft ein CA-Dinukleotid als Sequenz. Das Dinukleotid scheint aber im Gegensatz zur absoluten Notwendigkeit des Poly(A)-Signals nicht ausschlaggebend für eine funktionale Polyadenylierung zu sein.

U5 (Unique 5'-end)

Die U5-Region liegt stromabwärts von der R-Region und am 5'-Ende des RNA-Genoms (Unique 5'-end). Der R-U5-Übergang wird durch die Poly(A)-Stelle der integrierten Provirus-DNA definiert. Das 3'-Ende dieser Region wird durch den Übergang zur Primer Binding Site (PBS; siehe unten) festgelegt.

Im U5-Bereich können weitere für die Transkription wichtige Faktoren binden. Dies können aktivierende, wie auch reprimierende Transkriptionsfaktoren sein.

U3 (Unique 3'-end)

Analog zu U5 ist die U3-Region am 3'-Ende des RNA-Genoms gelegen (Unique 3'-end). An U3 schließt sich, definiert durch den Transkriptionsstart auf der proviralen DNA, die R-Region an. Die U3-Region wird am 5'-Ende durch den Polypurintrakt (PPT; siehe unten) begrenzt.

Die U3-Region enthält alle cis-aktiven Elemente inklusive regulatorischer Promotor, 5'-Enhancer und Teile des Core Promotors. Das Binden von Transkriptionsfaktoren auf diesen Sequenzen führt ultimativ zur Transkription der viralen Gene.

LTR (long terminal repeat)

LTRs sind die langen terminalen Sequenzwiederholungen (Redundanzen), die bei der Reversen Transkription entstehen. Sie befinden sich am 5'- und 3'-Ende integrierter Proviren.

Die LTR gliedert sich in die Untereinheiten U3, R und U5. Im RNA-Genom fehlt am 5'-Ende die U3 Region und am 3'-Ende die U5-Region. Die LTRs beinhalten alle Promotor-relevanten Bestandteile, die für die Transkription notwendig sind.

PBS (primer binding site)

Die PBS schließt sich an die U5-Region an und ist 18 Nukleotide lang. Sie ist die Bindungsstelle für eine spezifische tRNA, die als RNA-Primer bei der Initiation der DNA-Erststrangsynthese während der Reversen Transkription dient. Verschiedene Retroviren haben eine Spezifität für verschiedene tRNAs.

Leader

Die Leadersequenz ist die 5'-untranslatierte Region zwischen der PBS und den ORFs. Sie enthält zum einen den Spleißdonor, der bei der Prozessierung aller gespleißten viralen Transkripte verwendet wird, zum anderen liegt das Verpackungssignal ψ stromabwärts vom Spleißdonor. Über diese Signalsequenz können die RNA-Genome mit den Nucleocapsidproteinen des sich bildenden Viruspartikels interagieren.

gag-prt-pol

Das gag-prt-pol-Gen codiert für drei Proteine, deren ORFs gewöhnlich überlappen:

- **gag (group-specific-antigen)**: Das gag-Gen codiert ein Vorläuferprotein, das durch die Protease (prt) in die gruppenspezifischen Antigene gespalten wird. In Reihenfolge vom N- zum C-Terminus findet man ORFs für das Matrixprotein (MA), das Capsidprotein (CA) und das Nucleocapsidprotein (NC). Das Matrixprotein lagert sich an der Innenseite der Virushülle an, das Capsidprotein bildet die Hülle, in der die RNA-Genome eingelagert werden. Diese interagieren mit den Nucleocapsidproteinen.
- **prt (protease)**: Die Protease ist für die Prozessierung der gag-, gag/prt- und gag/pol-Vorläuferproteine verantwortlich.
- **pol (polymerase)**: Dieses Gen codiert für drei enzymatische Aktivitäten. Im 5'-Bereich des pol-Gens ist die RT codiert und im 3'-Bereich die Integrase. Die RT besitzt RNA- und DNA-abhängige DNA-Polymerase-Funktion. Im aminoterminalen Bereich des Enzyms sind die Domänen lokalisiert, die für die Polymerasefunktion verantwortlich sind. Im carboxyterminalen Bereich befinden sich Domänen, die eine RNaseH-Funktion ausüben. Diese enzymatische Aktivität ist wichtig beim Abbau des RNA-Strangs bei der Erst- bzw. Zweitstrangsynthese der proviralen DNA.

Die Integrase kann doppelsträngige DNA schneiden. So wird ein Locus im Wirtsgenom geöffnet, in den das Provirus mit Hilfe der Ligase-Funktion des Enzyms inseriert werden kann.

Env (envelope)

Das env-Gen codiert ein Polyprotein, das durch eine zelluläre Protease prozessiert wird. Das **Surface-Protein (SU)** wird auf der Oberfläche der Virusmembran angelagert, das **Transmembranprotein (TM)** ist in der Membran lokalisiert. Beide Proteine bilden Trimere, die miteinander wechselwirken.

PPT (poly-purin tract)

Der PPT liegt im 3'-Bereich des Virusgenoms. Er ist resistent gegenüber dem RNaseH-Verdau und dient während der DNA-Zweitstrangsynthese als Primer.

1.1.3 Exogene versus Endogene Retroviren

Die unter 1.1.2 beschriebenen viralen Genprodukte führen nach Transkription und Translation zum Zusammenbau eines neuen Viruspartikels. Retroviren besitzen also eine exogene Zustandsform in Form eines viralen Partikels, der in der Lage ist, neue Zellen bzw. bei Entlassen in die Umwelt neue Wirte zu infizieren. Die andere Zustandsform ist die endogenisierte provirale Form. Diese provirale Form kann nach Infektion von Keimbahnzellen von Generation zu Generation weitergegeben werden. Man spricht dann von endogenen Retroviren (ERV).

ERVs haben meist ihre exogene Zustandsform im Zuge der Evolution verloren. Eine Koexistenz exogener und endogener Zustandsformen ist nur für MMTV (Moore et al. 1987) und für das Jaagsiekt-Schaf-Retrovirus (JSRV) (Palmarini et al. 1996; York et al. 1992) nachgewiesen. Beim Menschen sind bisher keine endogenen Viren bekannt, die auch exogen vorliegen (Löwer et al. 1993a; Löwer et al. 1996). Endogene Retroviren werden nach der tRNA benannt, die als RT-Primer dient. Zum Beispiel steht das Suffix K bei HERV-K für eine Lysin-spezifische tRNA. Das Anreichern von Punktmutationen führte in fast allen Fällen zur Zerstörung der ORFs der endogenen viralen Gene, als Beispiel seien HERV-I (Maeda 1985), ERV9 (La Mantia et al. 1991; La Mantia et al. 1992) und HERV-E (Martin et al. 1981) genannt, die keine Partikel mehr bilden können. Intakte ORFs für alle viralen Proteine sind nur für HERV-K, dem humanen endogenen Retrovirus K, bekannt (Löwer et al. 1995). Aber auch diese Virusgruppe ist nicht mehr infektiös (Löwer 1999).

1.1.4 HERV-K

Die ersten HERV-K Sequenzen wurden durch Sequenzähnlichkeiten mit dem pol-Gen von intrazisternalen A-Partikeln des Hamsters (Ono 1986) und dem pol-Gen von MMTV (Deen und Sweet 1986) entdeckt. Die erste vollständige Nukleotidsequenz (Ono et al. 1986) zeigte ein defektes Virusgenom. 1984 sind in Teratokarzinomzellen mit Hilfe des Elektronenmikroskops Viruspartikel entdeckt worden (Löwer et al. 1984). Sie wurden zunächst HTDV (human teratocarcinoma derived virus) genannt, um ihrer Expression in Teratokarzinomen Rechnung zu tragen. 1993 konnte gezeigt werden, dass HERV-K Genome für die HTDV-Partikel codieren (Löwer et al. 1993b).

HERV-K ist mit 30-50 Kopien im haploiden Genom vertreten. Daneben hat man ca. 10.000 sLTRs beschrieben. sLTRs sind provirale Loci, die durch homologe LTR-LTR Rekombination alle dazwischen liegende virale Sequenzen verloren haben. Das menschliche Genom besteht zu 1-2 % aus HERV-K Sequenzen.

Abbildung 1-2 zeigt den Prototyp eines HERV-K Genoms. HERV-K exprimiert die Proteine Gag, Prt, Pol, Env und Rec. Ob alle Proteine von einem Provirus codiert werden oder ob verschiedene Proviren sich zur Bildung von Viruspartikeln komplementieren, ist nicht bekannt.

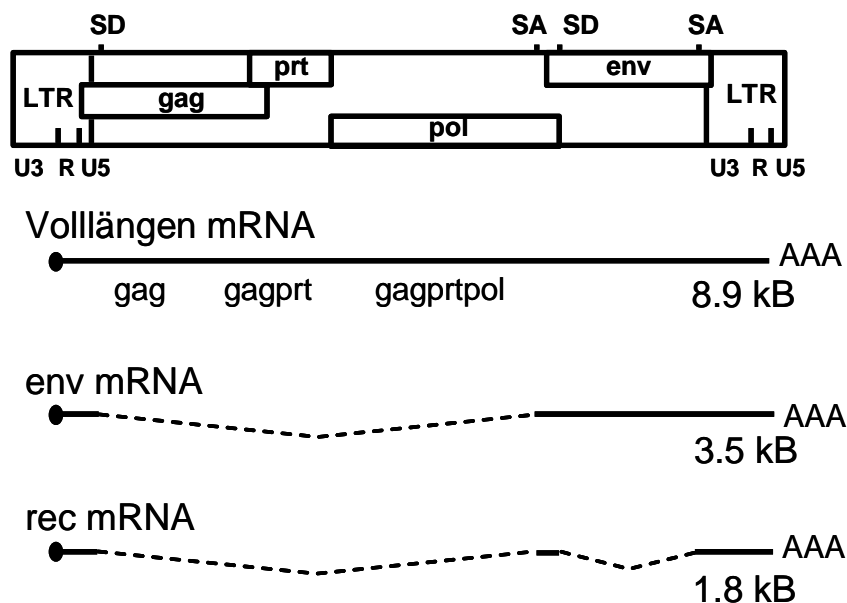


Abbildung 1-2: Darstellung des HERV-K Genoms und der Transkripte. gag Gruppenspezifische Antigene, prt Protease, pol Polymerase, env Envelope, SA Spleißakzeptor, SD Spleißdonor, LTR Long Terminal Repeat. (verändert nach Löwer et al. 1996)

Die Gag-, Prt-, und Pol-Proteine werden aus zwei Vorläuferproteinen mit einem Molekulargewicht von 76 kDa und 200 kDa prozessiert. Beide Vorläufer werden von der Vollängen-mRNA mit einer Größe von 8,9 kB translatiert. Gag spaltet sich in das Matrix- (20 kDa), Capsid- (30 kDa) und das Nucleocapsidprotein (16 kDa) auf (Boller et al. 1993; Mueller-Lantzsch et al. 1993). Die Polymerase (160 kDa) wird aus dem 200 kDa großen gag-prt-pol-Vorläufer prozessiert und ist in der Lage RNA revers zu transkribieren und DNA zu integrieren (Löwer et al. 1984). Die HERV-K Protease (17 kDa) katalysiert die Prozessierung der Vorläuferproteine (Schommer et al. 1996). Das Env-Protein (76 kDa) entsteht aus einer einfach gespleißten mRNA mit einer Größe von 3,5 kB, ist jedoch in Teratokarzinomzellen nur sehr schwach nachzuweisen (Denner et al. 1995). Die schwache Expression von Env ist in einem starken Spleißdonor innerhalb der env-mRNA begründet, der den Großteil des env ORFs als Intron entfernt und den ORF von rec mit einer Größe von 1,8 kB generiert. Die doppelt gespleißte mRNA wird zum akzessorischen Protein Rec (12 kDa) translatiert (Löwer et al. 1995). Dieses Protein besitzt Kernimport- wie auch Kernexportsignale und ist in der Lage, die Vollängen-mRNA und die env-mRNA über eine spezifische Sequenz im 3'-U3-Bereich der RNA zu binden. So wird ein erhöhter Export der viralen RNA aus dem Kern bewirkt.

HERV-K Proviren untergliedern sich in zwei Subtypen. Die genannten Genprodukte beschreiben ein Typ 2 Provirus. Zu den Typ 1 Proviren zählt man HERV-K Sequenzen, die eine 292 bp Deletion am Beginn des env-Gens haben, die zum Verlust des rec-Gens, des Spleißakzeptors und des Spleißdonors führt. Allerdings zeigen neuere Publikationen, dass einige der Typ 1 Proviren für ein gespleißtes env-Transkript codieren (Wang-Johanning et al. 2003a; Wang-Johanning et al. 2003b).

Interessanterweise wurde kürzlich ein von Typ 1 Proviren codierter Leserahmen beschrieben, der ein weiteres akzessorisches Protein zu exprimieren scheint. Ein Typ 1 spezifischer Spleißakzeptor führt zum Spleißprodukt np9, das in Keimzelltumoren entdeckt wurde. Es wurde vermutet, dass die tumorspezifische Expression von np9 maßgeblich an der Onkogenese beteiligt sein könnte (Armbrüster et al. 2002). Die oben genannten HERV-K Genprodukte wurden durch RT-PCR, Elektronenmikroskopie, Northern Blot und Western Blot Analyse nachgewiesen (Boller et al. 1997; Boller et al. 1993; Löwer et al. 1993b; Tönjes et al. 1997).

In einer umfangreichen Untersuchung mit ca. 2.000 Sera von gesunden und kranken Blutspendern zeigte sich eine spezifische humorale Immunantwort gegen HERV-K. Es konnte durch Western Blot Analyse, Immunfluoreszenz und immunocytochemische Verfahren

gezeigt werden, dass über 60 % der Hodentumorpatienten Antikörperreaktivitäten gegen HERV-K Proteine aufwiesen. Dieser Wert fiel nach mindestens einem Jahr nach Tumorentfernung auf ca. 10 %. Erhöhte Reaktivitäten konnte man auch bei schwangeren Frauen mit ca. 9 % feststellen, bei gesunden Blutspendern konnte man nur bei 4 % der Proben Antikörperreaktivitäten gegen HERV-K feststellen (Boller et al. 1997). Die Ergebnisse zeigten deutlich eine Assoziation von HTDV/HERV-K mit Keimzelltumoren.

Diese Ergebnisse reflektieren das Schicksal eines endogenen Retrovirus, das in seiner Expression bestimmten zellulären Gegebenheiten unterworfen ist. Das bedeutet, dass anscheinend eine HERV-K Expression nur in bestimmten Geweben unter bestimmten Bedingungen möglich ist. Dies trifft für Keimzelltumore zu und, wie in einer aktuellen Veröffentlichung von Muster et. al. gezeigt wurde, auch für maligne Melanome (Muster et al. 2003). Transkriptionsprodukte hingegen konnte man mit der sehr sensitiven RT-PCR in fast jedem Gewebe nachweisen. Dazu gehört auch das plazentale Gewebe, in dem auch andere endogene Proviren bevorzugt exprimiert werden. Dies wurde durch eine leicht erhöhte Immunantwort gegen HERV-K bei schwangeren Frauen bestätigt (siehe oben).

1.2 Transkription

Transkription ist das Umschreiben zellulärer Gene in komplementäre RNA. Es gibt drei verschiedene RNA-Spezies, die von drei verschiedenen RNA-Polymerasen transkribiert werden.

1.2.1 RNA-Polymerasen

Die RNA-Polymerase I transkribiert die Gene, die für die großen ribosomalen RNAs (rRNA) codieren. Diese sind Bestandteil der Ribosomen, den ausführenden Enzymkomplexen der Translation. Die RNA-Polymerase III synthetisiert die tRNA, 5S RNA, 7SL RNA, U6 snRNA und einige andere kleine stabile RNAs, von denen viele bei der RNA-Prozessierung eine Rolle spielen. tRNAs haben eine Länge von ungefähr 70-90 Nukleotiden. Sie dienen als Adapter für jeweils eine der zwanzig Aminosäuren, die während der Kettenverlängerung bei der Translation dem wachsenden Aminosäurestrang zugeführt werden. Bemerkenswert ist, dass 80 % der totalen RNA-Produktion von diesen Polymerasen geleistet wird. Dem gegenüber steht die sehr begrenzte Anzahl der Gene, die von der RNA-Polymerase I und III transkribiert werden.

Die RNA-Polymerase II (PolII) transkribiert Gene, die für Proteine codieren. Das Produkt ist die mRNA. mRNA wird während und nach der Transkription prozessiert. Am 5'-Ende wird sie mit einer Kappe aus einem methylierten G-Nukleotid versehen, das 3'-Ende wird polyadenyliert. Da die transkribierten Gene meist aus Exons und Introns bestehen, müssen die Introns von der Spleißmaschinerie herausgeschnitten werden. Die Exons besitzen letztendlich die Information für die Aminosäuresequenz des Proteins. Diese Information ist in den Codons festgelegt, Nukleinsäuretripletts, die für eine bestimmte Aminosäure stehen. An den Ribosomen findet schließlich die Translation der mRNA in ein Protein statt.

Dem sehr kleinen Anteil der zellulären Transkription, der von der PolIII geleistet wird, steht die sehr große Anzahl von verschiedenen Genen gegenüber, die von ihr transkribiert werden. Ersten Schätzungen nach codieren 30.000 Gene für Proteine. Diese Daten konnten durch die Entschlüsselung des menschlichen Genoms erhoben werden (Lander et al. 2001; Venter et al. 2001). Durch alternatives Spleißen wird allerdings diese sehr geringe Anzahl von möglichen Proteinen stark erhöht (Graveley 2001).

Die enorme Vielfalt der Proteine, die in verschiedenen Geweben zu verschiedenen Entwicklungszuständen in einem genau definierten Expressionsmuster der Zelle zur Verfügung stehen müssen, machen eine sehr genaue Kontrolle der Transkription durch Core Promotor und regulatorische Sequenzen nötig (siehe 1.2.2). Da nur die PolIII für die HERV-K Expression von Bedeutung ist, wird nur dieses Enzym ausführlicher beschrieben.

Die PolIII hat mehr als zwölf Untereinheiten und besitzt ein Molekulargewicht von mehr als 500 kDa. Diese Größe gestaltete die Untersuchung der RNA-Polymerase sehr schwierig und erst die Analyse der Kristallstruktur durch Kornberg und Mitarbeiter im Jahr 2000 / 2001 machte die Struktur des Polymerasekomplexes in freier wie in DNA-gebundener Form klar (Bushnell et al. 2002; Cramer et al. 2000; Cramer et al. 2001; Gnatt et al. 2001). Dadurch konnte geklärt werden, wie die Polymerase die DNA bindet und in welcher Form sie in den Elongationsstatus übergeht.

Um die Transkription zu starten, wird die Polymerase von verschiedenen Transkriptionsfaktoren an der Initiationsstelle festgehalten. Nach dem Aufbau des Initiationskomplexes ist ein Ablauf von Konformationsänderungen dafür verantwortlich, dass die Polymerase am DNA-Strang entlang gleiten kann und anhand der DNA-Matrize ein RNA-Strang synthetisiert wird (Elongation). Es konnte gezeigt werden, dass innerhalb des Komplexes ein DNA-RNA-Hybrid existiert, das ca. 9 bps lang ist. Vor dem wachsenden 3'-OH-Ende des neu synthetisierten RNA-Strangs existiert ein ca. 4 bps langes entwundenes DNA-Stück, das durch eine Brückenhelix stabilisiert wird.

Die neu synthetisierte RNA wird entlang einer Wand innerhalb der PolII über eine Ausgangsöffnung in der Nähe der carboxyterminalen Domäne (CTD) aus dem Proteinkomplex entlassen. Diese Trennung der RNA von der DNA-Matrize wird ca. 10 bps stromaufwärts vom 3'-Ende des Transkripts vollführt. Abbildung 1-3 zeigt einen Ausschnitt aus diesen Ereignissen.

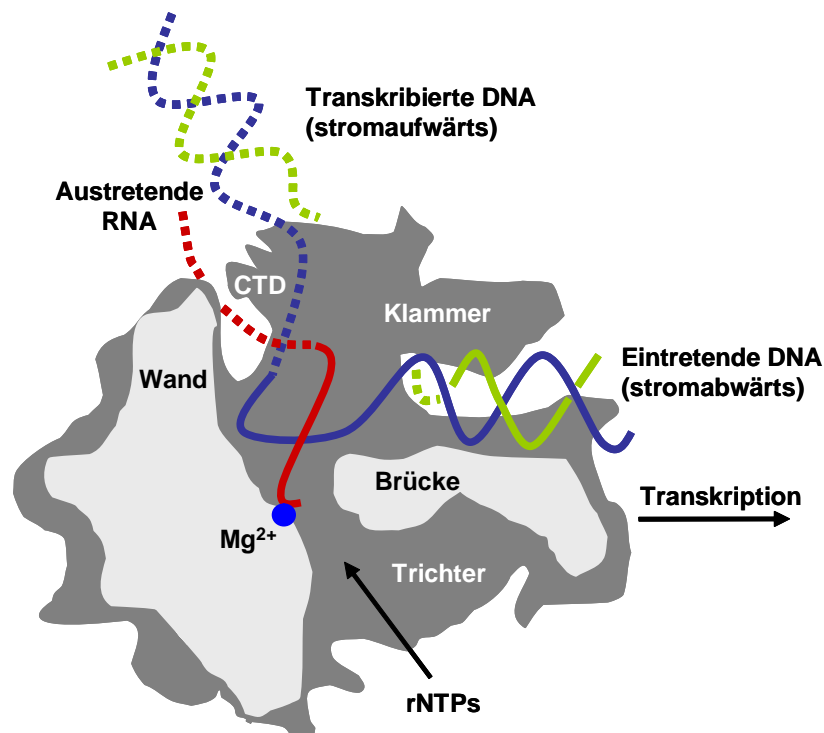


Abbildung 1-3: Schnitt durch den RNA-Polymerase II Elongationskomplex. CTD carboxyterminale Domäne, rNTPs Ribonukleosid-Triphosphate Mg^{2+} katalytisches Zentrum (blauer Punkt) mit einem Magnesium-Ion (Abbildung verändert nach Bushnell et al. 2002).

Die PolII besitzt eine Art molekulare Klammer, die in der DNA-ungebundenen Form geöffnet ist. Hier wird die zu transkribierende DNA-Helix eingelagert. Dadurch kommt das katalytische Zentrum, in dem ein Mg^{2+} -Ion eingelagert ist, in räumliche Nähe zur DNA und die ersten rNTPs können an der DNA polymerisiert werden. Der größer werdende Kontakt des RNA-DNA-Hybrids mit der Klammer bewirkt eine Konformationsänderung, die zum Zuklappen der Klammer führt. Die PolII geht in den Elongationkomplex über. Über eine trichterähnliche Struktur im Enzym werden die rNTPs dem katalytischen Zentrum zugeleitet und die RNA verlässt über eine Öffnung an der CTD der PolII den Enzymkomplex.

1.2.2 Der Promotor

Der Promotor ist das Sequenzelement eines Gens, das für die Regulation der Transkription verantwortlich ist. Das basale Element, das verantwortlich für das Binden der PolIII ist, nennt man Core Promotor. Es konnte *in-vitro* gezeigt werden, dass der Core Promotor ausreichend ist, um die Transkription zu starten. *In-vivo* finden sich jedoch neben dem Core Promotor weitere Kontrollelemente, die die Transkription eines bestimmten Gens kontrollieren. In Abbildung 1-4 ist die Unterteilung in Promotor-proximale Elemente (regulatorischer Promotor), Promotor-distale Elemente (Enhancer) sowie Core Promotor dargestellt.

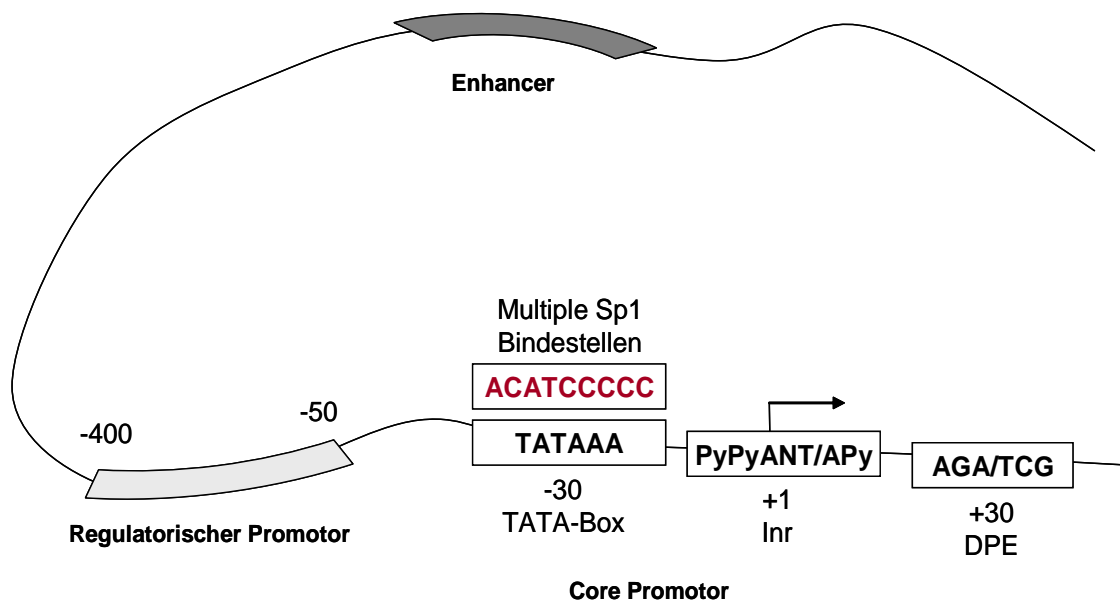


Abbildung 1-4: Schema eines eukaryotischen Promotors. Py Pyrimidin, Inr Initiatorelement, DPE Downstream Promoter Element.

1.2.2.1 Core Promotor

Der Core Promotor ist die Bindungsstelle für die PolIII-Transkriptionsmaschinerie und legt damit Start und Richtung der Transkription fest. Bis heute identifizierte Elemente des Core Promotors umfassen die TATA-Box (25-30 Nukleotide stromaufwärts des Transkriptionsstartes), das Initiatorelement (Inr), das die Initiationsstelle flankiert, und das „downstream promoter element“ (DPE), ca. 30 Nukleotide stromabwärts der Startstelle. Diese drei Elemente können in Core Promotoren in Kombination oder auch alleine auftreten. Das DPE übernimmt beim Aufbau des Initiationskomplexes eine unterstützende Funktion, wurde bis jetzt aber noch nicht alleine in Promotoren gefunden. Multiple Bindungsstellen für den Transkriptionsfaktor Sp1 können ebenfalls die Assemblierung des Initiationskomplexes

verstärken. Sp1-Bindungsstellen finden sich häufig in TATA-Box-losen Promotoren (siehe 1.2.4).

An den Core Promotor bindet der Transkriptionsfaktor TFIID (Burke et al. 1998; Javahery et al. 1994; Kaufmann und Smale 1994; Smale et al. 1998). Dieser Multiproteinkomplex enthält das „TATA binding protein“ (TBP), und das Binden weiterer TBP-assoziiierter Faktoren (TAF) führt zum Aufbau des Initiationskomplexes. Die verschiedenen Core Promotor Elemente scheinen für verschiedene Gene in unterschiedlicher Art und Weise kombiniert zu sein. Die Genregulation kann also zu einem gewissen Grad aktiv durch die basale Transkriptionsmaschinerie beeinflusst werden (Smale 2001). Man kann sich gut vorstellen, dass diese Varietät der Core Promotoren zu einer größeren Diversität von Regulationsmechanismen führt.

Dies könnte auch erklären, wieso distale Enhancer spezifisch Gene in ihrer Transkription verstärken können. Die teilweise sehr große Entfernung von Enhancern zum Core Promotor könnte theoretisch eine Verstärkung aller vom Enhancer erreichbaren Gene bewirken. Dies könnte fatale Folgen für den Organismus haben. Die Proteinkomplexe, die an Enhancer binden, können eine erhöhte Affinität für bestimmte Core Promotor-Kombinationen (TATA, TATA-Inr, Inr, Inr-DPE) aufweisen. Die Genregulation kann so deutlich präzisiert werden.

Sind die verschiedenen Proteinkomplexe und die Reihenfolge der Assemblierung am Core Promotor für TATA-Box gelenkte Gene weitgehend bekannt, liegen für die Inr-gesteuerte Transkription noch nicht genügend Daten vor. Inwieweit und in welcher Weise die besprochenen Elemente zur kombinatorischen Genkontrolle beitragen, ist zurzeit Hauptaufgabe der Transkriptionsforschung.

1.2.2.2 Regulatorischer Promotor

Regulatorische Faktoren, die in naher Umgebung zum Core Promotor binden, werden als Promotor-proximale Elemente bezeichnet und befinden sich meistens in einem DNA-Bereich, der sich von den bps -400 bis -50 relativ zur Initiationsstelle erstreckt. Dieser Bereich wird als regulatorischer Promotor bezeichnet. Hier binden Aktivatoren sowie Repressoren, die über weitere Coaktivatoren oder über einen Mediator das Binden der PolIII am Transkriptionsstart beeinflussen.

1.2.2.3 Enhancer

Enhancer sind regulatorische Sequenzen, die aktivierende Transkriptionsfaktoren binden können. Der Enhancer kann unabhängig vom Promotor stromaufwärts bzw. -abwärts vom Core Promotor lokalisiert sein. Dies kann auch mehrere tausend bps vom Promotor entfernt stattfinden. Die gebundenen Transkriptionsfaktoren können über Wechselwirkungen mit anderen Transkriptionsfaktoren innerhalb des regulatorischen Promotors oder über das Binden von Coaktivatoren die Zusammenführung verschiedener Faktoren aus regulatorischem und Core Promotor erleichtern. Diese Aktivatoren werden auch Enhancerprotein genannt.

Der verstärkende Effekt eines Enhancers lässt sich dadurch erklären, dass das Wechselspiel zwischen Aktivatoren, generellen Transkriptionsfaktoren (GTF) und TAFs am Core Promotor zum Aufbau des Initiationskomplexes erleichtert wird.

1.2.3 Transkriptionskontrolle durch epigenetische Mechanismen

Eine weitere Ebene der transkriptionellen Kontrolle wird durch die chromosomale Verpackung der DNA erreicht. Transkriptionell aktive und inaktive Bereiche lassen sich schon durch mikroskopische Betrachtung unterscheiden. Die Verpackung der DNA in Heterochromatin führt zur transkriptionellen Inaktivität (Brown 1966). Erreicht wir dies durch das Winden der DNA um Histonproteine, was die dichte Packung des Genoms zu Chromosomen ermöglicht (Richmond et al. 1984; Smith 1991). Die teilweise Entwindung der Chromosomen führt zum mikroskopisch erkennbaren Euchromatin (Shirley und Anderson 1977). Bereiche der DNA werden von den Histonen gelöst, die PolIII bekommt Zugang zu den Promotor-relevanten Bereichen und das Gen kann abgelesen werden. Die Verpackung der DNA wird durch Mechanismen, wie die Acetylierung der Histone und die Methylierung der Cytosine in CpG-Dinukleotiden, erreicht (Struhl 1999).

Die zelluläre Transkription ist also die Summe aus epigenetischen Mechanismen, dem Status der DNA-Kondensation im Bereich des Promotors und den im zellulären Kontext vorhandenen Transkriptionsfaktoren.

1.2.4 Transkription von Retroviren

1.2.4.1 Allgemein

Da retrovirale Proviren im zellulären Genom integriert sind, ist die virale Transkription grundsätzlich der zellulären Umgebung der Wirtszelle unterworfen, und epigenetische Mechanismen können die virale Transkription beeinflussen. Die zelluläre

Transkriptionsmaschinerie bindet auf dem viralen Promotor, um die Transkription zu initiieren. Dieser befindet sich in der proviralen LTR. Der Core Promotor und der regulatorische Promotor liegen zum größten Teil in der U3-Region. Die Bindungsstelle für die PolIII und den zugehörigen GTFs liegt im U3-R-Übergang. Das erste synthetisierte Nukleotid definiert den Anfang der R-Region. Die Transkripte unterliegen den gleichen Modifikationen wie die zelluläre mRNA. Das heißt, die virale mRNA wird mit einer methylierten 5'-G-Nukleotid-Kappe versehen und das 3'-Ende wird polyadenyliert (Furuichi und Shatkin 2000; Shatkin und Manley 2000; Wickens 1990). Viren haben sehr unterschiedliche Wirtszelltropismen und benutzen Transkriptionsinitiationsstrategien, die auch bei zellulären Genen zur gewebespezifischen Expression führen. Dabei können Viren auf ihrer LTR sowohl TATA-vermittelte wie auch Inr-vermittelte Initiationskomplexe bilden. Bei HIV ist ein TATA-Signal auf der LTR zu finden, das auch von TBP zur Initiation benutzt wird (Jones und Peterlin 1994). Die Transkription ist aber stark abhängig von Faktoren wie Sp1 (Jones et al. 1986) und NF- κ B, wobei die NF- κ B-Proteine für die Aktivierung von proviralen HIV-Genomen in T-Zellen verantwortlich sind (Nabel und Baltimore 1987). MMTV-Transkription wird ebenfalls über eine TATA-Box initiiert, kann aber über Steroid Hormone aktiviert werden. So wird eine effiziente Expression in der Brustdrüse erzielt. Die verstärkte Virusreplikation durch den steigenden Hormonspiegel während der Stillzeit gewährleistet die Übertragung des Virus auf die Nachkommen (Bittner 1936; Günzburg und Salmons 1992).

Interessanterweise weisen endogene Retroviren oft eine Abweichung von der TATA-vermittelten Transkription auf und benutzen eine Sp1-vermittelte Inr-abhängige Transkriptionsinitiation. Das humane endogene ERV9 besitzt zwar eine TATA-Box, doch der Transkriptionsstart wird über ein Inr vermittelt. Relativ nah am Inr liegt stromaufwärts eine Sp1-Bindungsstelle, die essentiell für die Transkriptionsinitiation ist (La Mantia et al. 1992; Strazzullo et al. 1994). ERV9 wird bevorzugt in undifferenzierten embryonalen Tumorzellen exprimiert (La Mantia et al. 1991).

Ein ähnliches Bild zeigt sich beim humanen endogenen Retrovirus H (HERV-H). Auch hier hängt die Transkription entscheidend von Sp-Proteinen ab, die stromaufwärts von einem Inr binden können. Ein TATA-Box-Signal, das in der Nähe vom Transkriptionsstart zu finden ist, scheint keine Funktion zu haben (Sjøttem et al. 1996).

Die oben genannten HERVs weisen eine basale bis starke Expression in plazentalem bzw. teratokarzinogenem Gewebe auf (Harris 1998). Die enge Assoziation von HERV-Expression mit Keimzell-verwandtem Gewebe scheint sich in den zellulären Mechanismen der

Transkriptionsregulation in diesem Gewebe zu erklären. Beispiele für Sp-Protein-regulierte Gene finden sich bei SRY, einem Y-Chromosom-codierten Protein, das vermehrt im Testis exprimiert wird (Desclozeaux et al. 1998), oder beim Testis-spezifischen Histon H1t Gen, das ebenfalls deutlich in seiner transkriptionellen Aktivität von Sp-Proteinen abhängig ist (Grimes et al. 2003; Wilkerson et al. 2002). Untersuchungen am porcinen Endometrium während der Schwangerschaft zeigten eine deutliche Abhängigkeit der zellulären Transkription vom Transkriptionsfaktor Sp1. Dem transaktivierenden Potential von Sp1 konnte die erhöhte Expression des epithel-spezifischen Proteins Uteroferrin zugeordnet werden. Dieses Protein ist für den transplazentalen Eisen-Transport im Endometrium verantwortlich (Simmen et al. 2000).

1.2.4.2 HERV-K

In einer vorangegangenen Dissertation (Hasenmaier 1995) wurden HERV-K LTRs aus der genomischen DNA von GH-Zellen, einer HERV-K produzierenden Teratokarzinomzelllinie, isoliert und im Luziferase-Assay auf transkriptionelle Aktivität überprüft. Dabei stellte sich heraus, dass im Reporter-Assay aktive LTRs sich von den inaktiven deutlich durch Anhäufungen von Punktmutationen unterscheiden. Die Akkumulation von Punktmutationen führt möglicherweise durch Zerstörung von Transkriptionsfaktorbindungsstellen zu inaktiven LTRs.

Interessanterweise isolierten Turner et al. (Turner et al. 2001) zwei HERV-K Genome, HERV-K 113 und 115, die vor relativ kurzer Zeit in das menschliche Genom integriert wurden. Das Alter dieser Virusgenome konnte auf ca. 200.000-450.000 Jahre bestimmt werden. Dabei wird davon ausgegangen, dass der direkte Vergleich der Mutationsereignisse in den eigentlich identischen 5'- und 3'-LTRs ein Marker für die Verweildauer eines Provirus im Genom ist (Johnson und Coffin 1999). Die ältesten bekannten HERV-K Sequenzen stammen wahrscheinlich von Keimbahninfektionen, die vor ca. 30 Millionen Jahren stattfanden (Steinhuber et al. 1995).

In der Literatur wurde die Bestimmung der HERV-K-LTR Bereiche auf die Analyse der Sequenz der LTR10 gestützt (Ono 1986). Eine TATA-Box-Konsensussequenz wurde bei bp 532 gefunden. Der Beginn der R-Region und damit der putative Transkriptionsstart (R^{put}) wurde auf das Guanosin bei Position 562 festgesetzt (Breathnach und Chambon 1981; Ono 1986). Der R-U5-Übergang wurde auf Grund eines typischen Poly(A)-Signals bei bp 794 und einer putativen Poly(A)-Stelle (CA) bei bp 813 vorhergesagt ($U5^{put}$) (Breathnach und Chambon 1981; Ono 1986). Eine Analyse der HERV-K mRNAs, die in Teratokarzinomzellen

gebildet werden, zeigte allerdings bei Sequenzierung der cDNAs einen R-U5-Übergang, definiert durch die Poly(A)-Stelle bei bp 878 (U5^{exp}) (Löwer et al. 1993a).

Um den U3-R-Übergang experimentell zu überprüfen, wurden bereits verschiedene 5'-Deletionen der aktiven LTR3 im Luziferase-Assay untersucht (Thelen 1995). Es konnte gezeigt werden, dass die Deletion der ersten 55 bps zu einem Abfall der Promotoraktivität führte. Später wurde der Enhancer-relevante Bereich bis zum bp 150 ausgedehnt. Es konnte im Bereich zwischen den Nukleotiden 62 - 83 eine Bindungsstelle für den Transkriptionsfaktor YY1 identifiziert werden. YY1 kann als 5'-Enhancer-bindendes Protein angesehen werden (Knöbl 1998). Eine schrittweise Deletion des 5'-Bereichs zeigte erst mit Deletion der TATA-Box bei bp 532 einen Rückgang der Aktivität auf 1 %. Ein längeres Konstrukt, das die bps 460-988 enthielt, zeigte eine Restaktivität von 20 %. Der Core Promotor wurde deshalb durch dieses Konstrukt definiert (Thelen 1995). Der tatsächliche Transkriptionsstart (R^{exp}) sowie die 3'-Grenze des Core Promotors wurden jedoch noch nicht bestimmt.

1.3 Zielsetzung der Arbeit

HERV-K wird in seiner Expression durch die oben beschriebenen zellulären Mechanismen beeinflusst. Neben dem bereits identifizierten Enhancerprotein YY1 sollen in dieser Arbeit weitere Bereiche des LTRs analysiert werden. Ein besonderes Augenmerk liegt dabei auf der Identifikation eines Testis-spezifischen Transkriptionsfaktors, der eine vorwiegende Expression von HERV-K in Teratokarzinomzellen erklären könnte.

Hierfür sollen aktive und inaktive LTRs herangezogen werden, um Promotor-relevante Regionen zu charakterisieren. Die potentiellen Transkriptionsfaktor-Bindungsstellen sollen näher untersucht werden. Anschließend sollen LTR-bindende Proteine von HERV-K exprimierenden Zellen und nicht-exprimierenden Zellen identifiziert werden. Weiterhin soll überprüft werden, in welchem Maße die TATA-Box für die Transkriptionsinitiation verantwortlich ist. Als letztes soll untersucht werden, ob nicht-exprimierende Zellen durch Testis-spezifische Proteine in ihrer Fähigkeit, HERV-K zu exprimieren, beeinflusst werden können und ob der Methylierungszustand eines Provirus Rückwirkungen auf die HERV-K Expression haben könnte.

2 Ergebnisse

2.1 Analyse von HERV-K LTRs

Die von B. Hasenmaier (Hasenmaier 1995) isolierten LTRs konnten bereits in einem Funktions-Assay, dem Luziferase-Assay (siehe 4.5.5), auf ihre Aktivität getestet werden. Hierzu wurden eukaryotische Zellen mit einem Konstrukt, in das der zu testende Promotor vor das Luziferasereportergen kloniert wurde, transfiziert. Die Menge der gebildeten Luziferase ist ein Maß für die Promotorstärke und wird als relative Luziferase-Aktivität (RLU) angegeben. Die einzelnen LTRs zeigten klare Unterschiede in ihrer transkriptionellen Aktivität. In dieser Arbeit sind je zwei aktive (pcK30, LTR10) und inaktive (LTR26 und LTR21) LTRs für die Analysen herangezogen worden.

In Abbildung 2-1 ist der Sequenzvergleich von pcK30, LTR10, LTR26 und LTR21 dargestellt. Man sieht deutlich die verschiedenen Cluster (Anhäufungen) von Punktmutationen, die die aktiven von den inaktiven LTRs unterscheiden (Abb. 2-1, z.B. bps 301, 303, 310). Insgesamt unterscheiden sich die aktiven LTRs von den inaktiven in über 40 Punktmutationen. Aktive wie auch inaktive LTRs unterscheiden sich untereinander in nur ca. 12 Punktmutationen (Abb. 2-1, z.B. bps 626, 656).

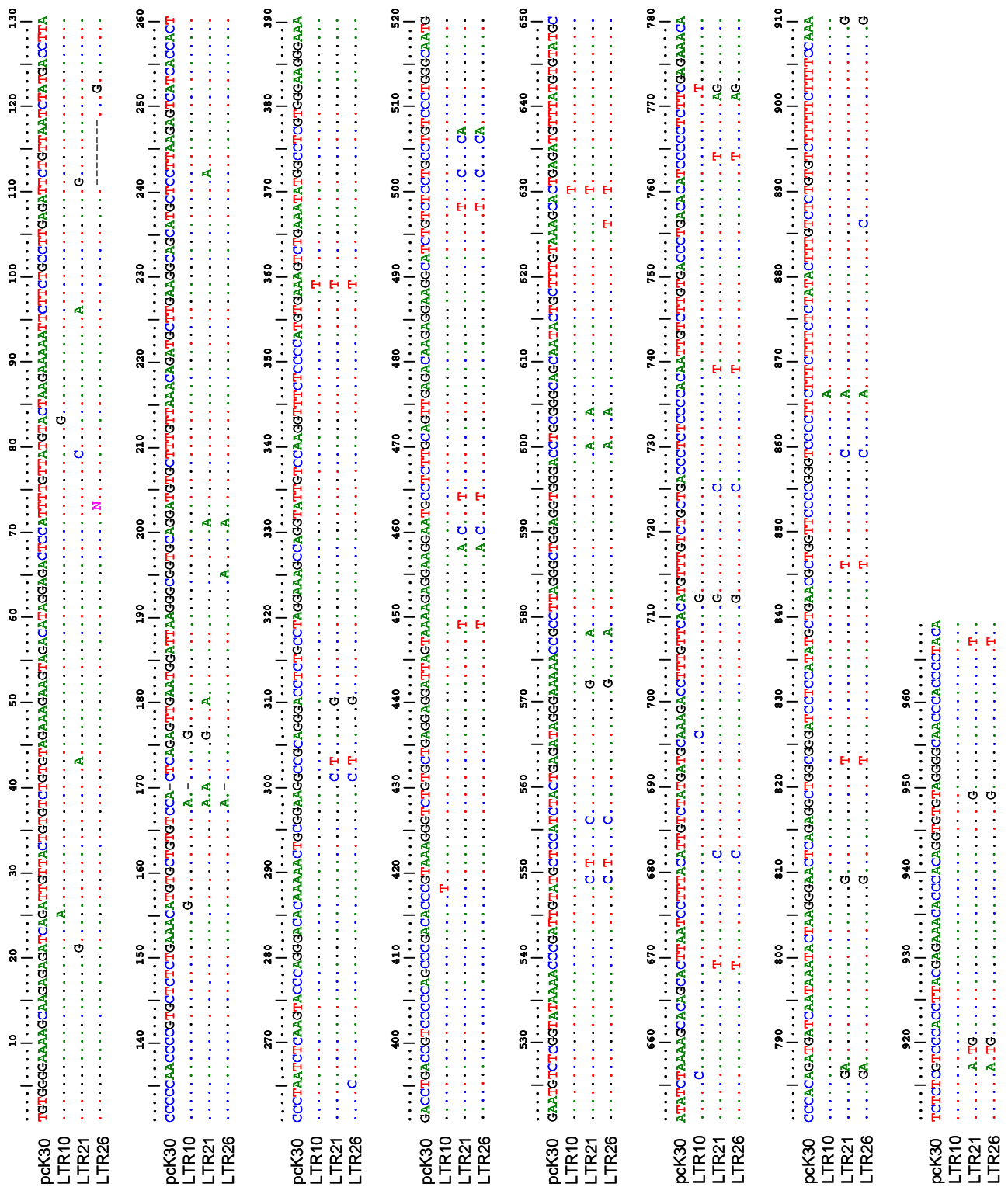


Abbildung 2-1: Sequenzvergleich der verwendeten HERV-K LTRs. Übereinstimmungen mit der aktiven pck30-Sequenz sind als Punkt dargestellt, Abweichungen durch die Abkürzung der entsprechenden Base. Striche kennzeichnen das Fehlen von Basenpositionen, die mindestens in einer anderen Sequenz vorkommen.

In einer phylogenetischen Analyse wurden zusätzlich weitere in unserer Arbeitsgruppe isolierte LTRs (LTR01, LTR18) sowie die kürzlich beschriebenen LTRs K 113 und K 115 (siehe 1.2.4.2, Turner et al. 2001) aufgenommen. Die den letztgenannten LTRs entsprechenden Sequenzen PEI1/2-K113 und PEI1-K115 wurden in unserer Arbeitsgruppe mit Hilfe der veröffentlichten Sequenzen von Turner et al. aus genomischer DNA durch PCR amplifiziert.

Im phylogenetischen Baum (Felsenstein 1993) sind zwei Hauptzweige zu erkennen: Im Luziferase-Assay inaktive LTRs gruppieren sich zueinander (Abb. 2-2, LTR18, LTR21, LTR26) und auch die aktiven LTRs pcK30, LTR10 und LTR01 sind zusammen in einer Gruppe zu finden. Die evolutionär jüngsten LTR-Vertreter der HERV-K Proviren K 113 und K 115 gruppieren sich zu den aktiven LTRs. Geht man davon aus, dass die Anhäufung von Punktmutationen im Zuge der Evolution bei aktiven LTRs zur transkriptionellen Inaktivität führt, stellen sich die inaktiven LTRs als die evolutionär ältere Gruppe dar.

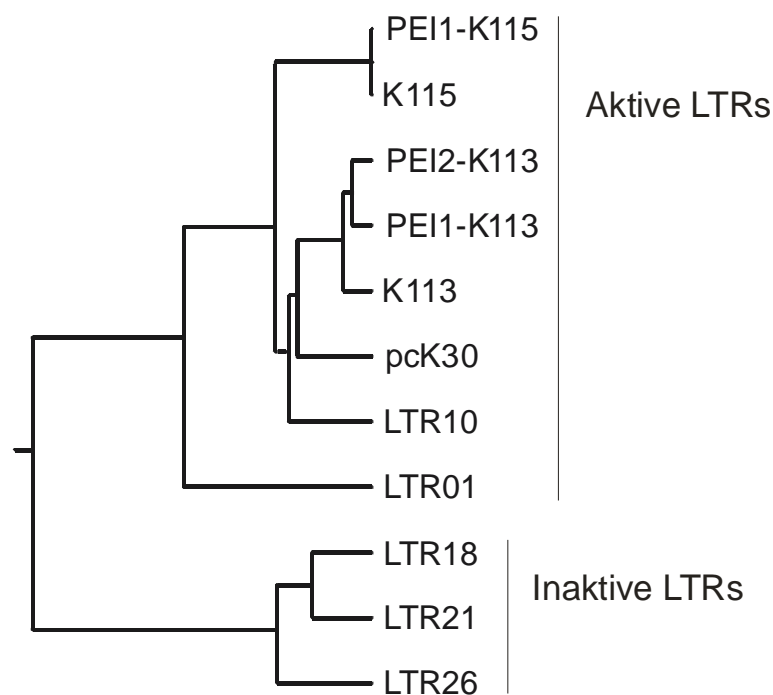


Abbildung 2-2: Phylogenetischer Stammbaum verschiedener HERV-K LTRs. Aktive und inaktive LTRs gruppieren sich abhängig von der Anzahl der sie unterscheidenden Punktmutationen.

In den HERV-K exprimierenden GH-Zellen (Löwer et al. 1984) zeigten die aktiven LTRs im Luziferase-Assay eine ähnliche Aktivität wie die Positivkontrolle, ein durch einen SV40-Promotor (Simian Virus 40) gesteuertes Luziferasekonstrukt (Abb. 2-3, Säulen 2 bis 7). Die inaktiven LTRs hatten kaum mehr Reporteraktivität als die Negativkontrolle, ein

Luziferasevektor ohne Promotor (Abb. 2-3, Säulen 1, 8, 9). Hier wurden die normierten RLU-Werte verglichen.

Die HERV-K 113 und 115 LTRs (PEI1/PEI2-K113, PEI1-K115) zeigten im Vergleich zu den inaktiven LTRs eine sehr starke Aktivität. Dies deckt sich mit dem Ergebnis der phylogenetischen Stammbaumanalyse (vergleiche Abb. 2-2 und Abb. 2-3, Säulen 3 bis 7). Die K113 und K115 LTRs weisen nur geringfügige Abweichungen gegenüber der Sequenz der aktiven pcK30 auf (Turner et al. 2001) (siehe Anhang, 7.1).

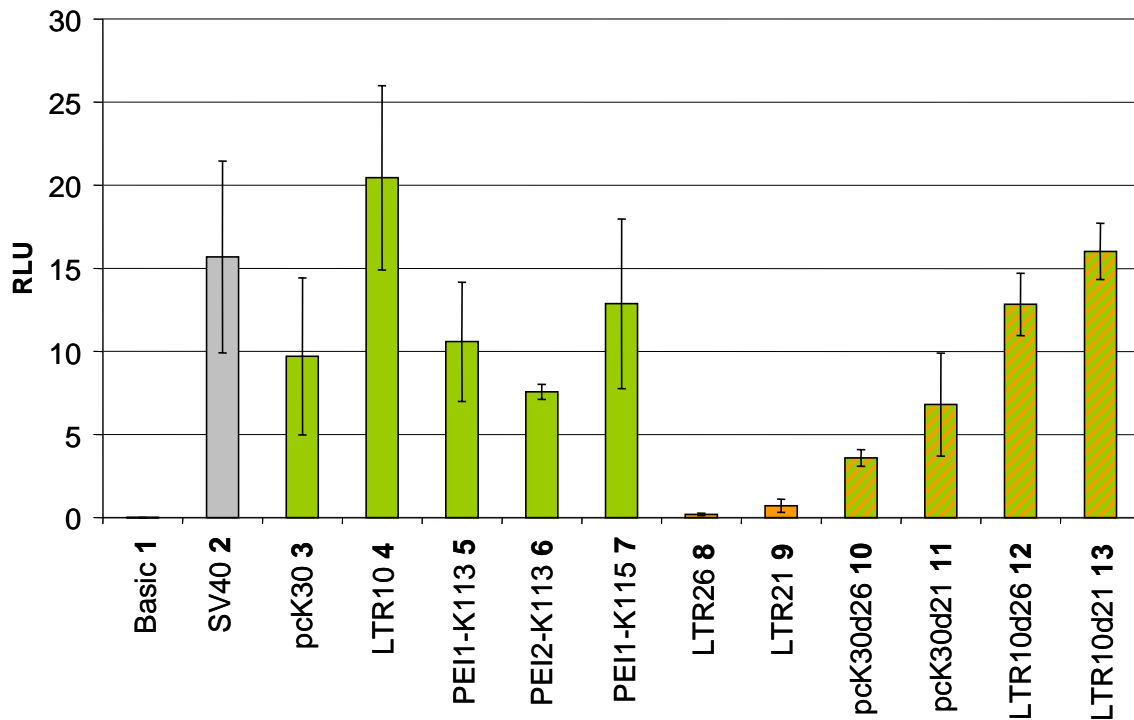


Abbildung 2-3: Relative Luziferase-Aktivität (RLU) verschiedener LTRs. Die Werte sind durch duale Cotransfektion normalisierte Aktivitäten. Erklärung zu den Konstrukten im Text.

Die großen Unterschiede in den Aktivitäten weisen darauf hin, dass die Punktmutationen, die die aktiven von den inaktiven LTRs trennen, transkriptionell sensible Bereiche der LTRs beeinflussen könnten.

Aufbauend auf diesen Befunden und den Daten vorhergehender Dissertationen (Hasenmaier 1995; Knöbl 1998; Thelen 1995) wurden verschiedene Konstrukte kloniert und im Reporter-Assay untersucht. Hierzu wurden chimäre Konstrukte aus aktiven und inaktiven LTRs hergestellt und die Reporteraktivität der Mutanten mit der Wildtypsequenz verglichen.

Chimäre Konstrukte aus den aktiven LTRs pcK30 oder LTR10 und den inaktiven LTR26 oder LTR21 zeigten ein vergleichbares Verhalten im Luziferase-Assay. Die Substitution von aktiven Sequenzelementen durch inaktive LTR-Bereiche führte immer zur Reduktion der

transkriptionellen Aktivität. Das heißt, eine Chimäre aus pcK30 und LTR21 zeigte einen ähnlichen Verlust in transkriptioneller Aktivität wie eine Chimäre aus pcK30 und LTR26. In den Konstrukten pcK30d26 und pcK30d21 wurden die bps 238-580 der aktiven LTR pcK30 durch inaktive Sequenzen der LTR26 und LTR21 substituiert (analog Seite 44, Abb. 2-14 A, Konstrukt 6). In Abbildung 2-3, Säulen 10 und 11 sieht man den Verlust der Promotoraktivität dieser Chimären. Das gleiche Resultat, nämlich Aktivitätsverlust, erhielt man bei Chimären aus der aktiven LTR10 (LTR10d26, LTR10d21; Abb. 2-3, Säulen 12 und 13). Es wird deshalb im Folgenden nur auf die pcK30 und die LTR21 und deren Folgekonstrukte eingegangen. Um mögliche Unterschiede zwischen HERV-K exprimierenden Zellen und nicht-exprimierenden Zellen erkennen zu können, wurde, neben der GH-Zelllinie, die nicht-produzierende Cervixkarzinomzelllinie HeLa (ATCC) gewählt.

Die gezielte Substitution verschiedener Bereiche aktiver LTRs aus U3, R und U5 durch inaktive Sequenzbereiche, wie in den folgenden Kapiteln beschrieben, sollte Promotorrelevante Regionen aufzeigen. Dabei wurde angenommen, dass die punktuellen Differenzen zwischen der aktiven LTR pcK30 und der inaktiven LTR21 für die transkriptionelle Aktivität verantwortlich sind. In den inaktiven LTRs könnten essentielle Transkriptionsfaktorkonsensussequenzen durch Mutationen zerstört worden sein. Im Umkehrschluss könnte die transkriptionelle Inaktivität durch eingebrachte aktive Sequenzen wieder aufgehoben werden. Neben Substitutionen eigneten sich Deletionen innerhalb des Promotors zur Eingrenzung der basalen Promotorbereiche. Auffällige Promotorbereiche wurden dann in einem *in-vitro*-Assay, dem elektrophoretischen Mobilitäts-Assay (EMSA, siehe 4.6), auf das Bindungsverhalten von Transkriptionsfaktoren überprüft. Hierzu wurden DNA-Sonden der entsprechenden Sequenzbereiche radioaktiv markiert und mit nukleären Proteinextrakten inkubiert. Eine Änderung im Laufverhalten der DNA-Banden im Polyacrylamidgel lässt Rückschlüsse auf spezifisch bindende Transkriptionsfaktoren zu.

2.1.1 Charakterisierung der R-Region

Um die Ergebnisse der Mutationsanalysen in einen Gesamtkontext einfügen zu können, ist es von großer Wichtigkeit, die basalen Promotorelemente zu kennen. Transkriptionsfaktoren sind in ihrer Funktion stark abhängig von ihrer Position zum Core Promotor (Naar et al. 2001). Die Charakterisierung der Grenzen der verschiedenen Regionen der LTR ist daher von elementarer Bedeutung. Durch die Sequenzierung der cDNA, die aus HERV-K mRNA aus GH-Zellen generiert wurde, wurde der R-U5-Übergang (U5^{exp}) definiert (Löwer et al. 1993b).

Der U3-R-Übergang (R^{put}), der durch die Transkriptionsinitiationsstelle definiert ist, wird im Folgenden experimentell behandelt.

2.1.1.1 3'-Deletionen

Aufbauend auf den Ergebnissen vorangegangener Dissertationen (Knöbl 1998; Thelen 1995) wurden sukzessiv 3'-deletierte Konstrukte von pcK30 hergestellt (Abb. 2-4 A, 3-7), um das 3'-Ende des Core Promotor zu bestimmen. Die Konstrukte wurden im Luziferase-Assay eingesetzt, um die Promotoraktivitäten in GH- bzw. HeLa-Zellen zu überprüfen. Die transkriptionellen Aktivitäten der verschiedenen Konstrukte sind auf die aktive pcK30 bezogen, die auf 100 % gesetzt wurde. Die Präfixe pcK oder LTR in den Konstruktbezeichnungen (z.B pcK30dK, LTR21d30) werden ab hier weggelassen.

In Abb. 2-4 B Säulen 1 und 2 sieht man den Unterschied der Promotoraktivität zwischen aktiver und inaktiver LTR in GH-Zellen. Die Summe aller Punktmutationen in der LTR21 führte zu einer transkriptionellen Restaktivität von 6 %. Die sukzessive Deletion des 3'-Endes führte bei den Konstrukten 30dG und 30dH zu einem Aktivitätsanstieg auf 130 - 140 %, der durch verlorene Repressor-Bindungsstellen erklärt werden könnte (Abb. 2-4 B, Säulen 3 und 4). Bei Konstrukt 30dG wurde die U5-Region deletiert. Ein Aktivitätsanstieg bei deletierter U5-Region wurde auch schon bei HERV-K sLTRs von Domansky et al. beobachtet (Domansky et al. 2000). Eine weitere Deletion des 3'-Bereichs der R-Region bei Konstrukt 30dH führte zu keinem signifikanten Aktivitätsunterschied zu Konstrukt 30dG. Dieser Bereich scheint also transkriptionell unbedeutend zu sein.

Mit dem Konstrukt 30dI (Abb. 2-4 B, Säule 5), hier fehlt der 3'-Bereich ab bp 793, fiel die transkriptionelle Aktivität auf 41 % und verdeutlicht den Verlust transkriptionell wichtiger Bereiche. Konstrukte 30dK und 30dA (Abb. 2-4 B, Säulen 6 und 7) führten schließlich mit einer Restaktivität von 8 % und 2 % zu fast vollständiger Promotorinaktivität. Diese weiteren Deletionen zeigten, dass keine Transkription mehr möglich war, obwohl die TATA-Box in den Konstrukten vorhanden gewesen ist. Bei einem basalen Promotor, der von der TATA-Box bei bp 532 abhängig ist, wäre dieser drastische Abfall nicht zu erwarten.

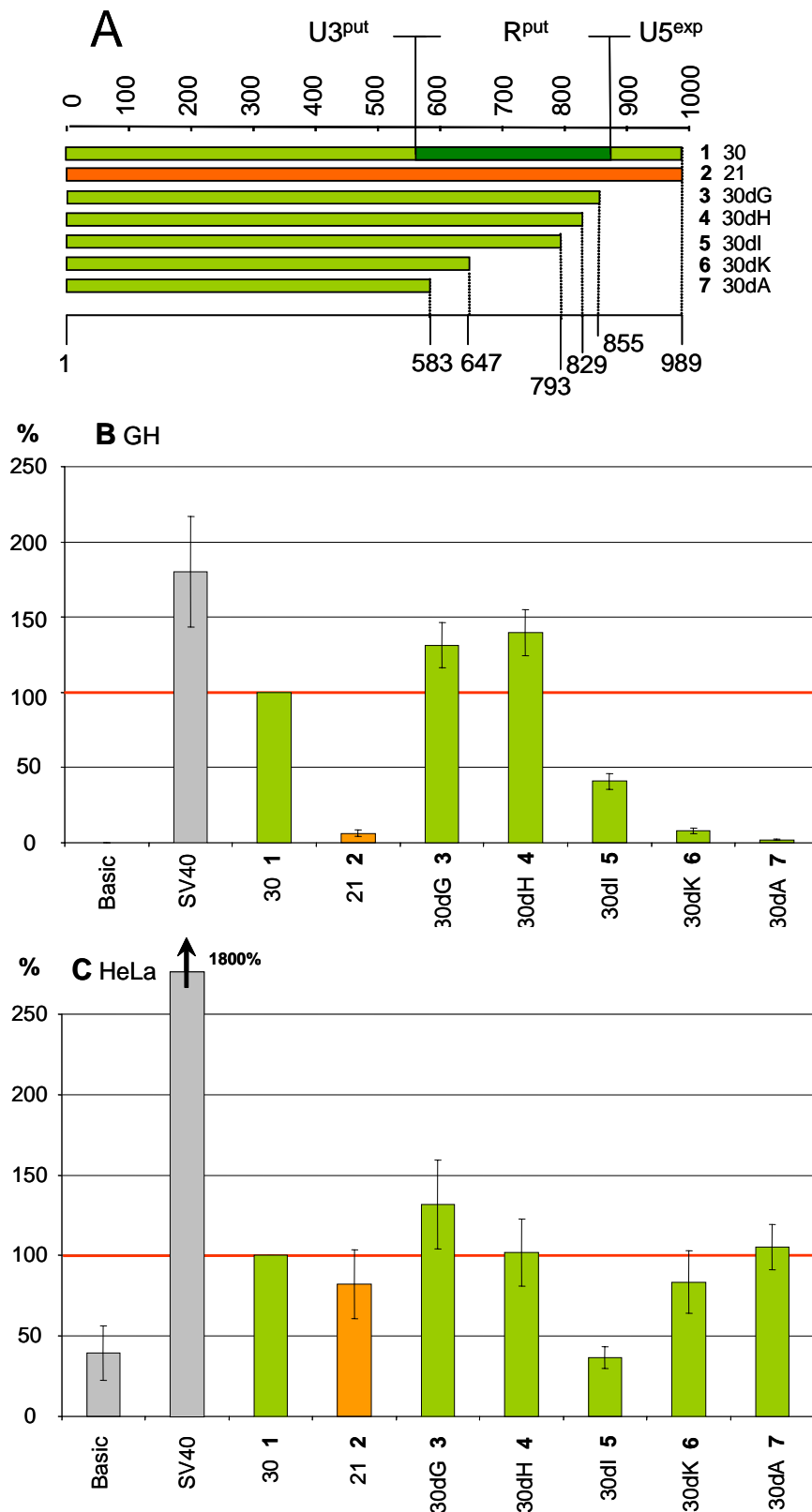


Abbildung 2-4: Relative Luziferase-Aktivität der LTR-Konstrukte. **A** Konstruktübersicht. Zahlen geben die Position im HERV-K LTR in bp an. put = putativ, exp = experimentell ermittelt. **B, C** Relative Luziferase-Aktivität verschiedener LTR-Konstrukte. Die Aktivität in % ist auf die aktive LTR pck30 (Spur 1, 100 %) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle **B** Werte für GH-Zellen. **C** Werte für HeLa-Zellen.

Die Promotoraktivität der LTR-Konstrukte in HeLa-Zellen schwankte im Vergleich zur Negativkontrolle um einen basalen Wert (Abb. 2-4 C). Ein Unterschied zwischen aktivem und inaktivem Promotor war nicht zu erkennen (Abb. 2-4 C, Säulen 1 und 2). Der aktive HERV-K Promotor pcK30 erzeugte in GH-Zellen das 1.000-fache der Aktivität, die ein leerer Reportervektor hervorrufen konnte (Abb. 2-4 B, Säulen Basic und 1). In HeLa-Zellen zeigte sich nur eine 2,5-fache Erhöhung der Aktivität des HERV-K Promotors gegenüber der Negativkontrolle (Abb. 2-4 C, Säulen Basic und 1). Auch der Vergleich zur Positivkontrolle (Abb. 2-4 B, C, Säulen SV40 und 1) macht deutlich, dass die Transkription der HERV-K LTRs in GH-Zellen um ein Vielfaches höher war als in HeLa-Zellen. In GH-Zellen besaß die pcK30-LTR im Vergleich zum starken SV40-Promotor die halbe Promotoraktivität, in HeLa-Zellen erzeugte das Kontrollplasmid eine 18-fach höhere Promotoraktivität als die pcK30-LTR.

2.1.1.2 Funktionalität der TATA-Box

Die oben beschriebenen Ergebnisse legen nahe, dass die TATA-Box nicht als Transkriptionsinitiationssequenz fungiert. Um die Funktionalität der TATA-Box weiter zu untersuchen, ist die Konsensussequenz TATAAAA im Konstrukt 30dT (Abb. 2-5 A, 3) durch Thymidine ersetzt worden. Es wurde mit der Transkriptionsfaktorsuchmaschine TESS (Overton 1997) ausgeschlossen, dass eine Restbindungsaktivität von TBP bestehen könnte. Dieses computerunterstützte Ergebnis wurde auch durch eine statistische Untersuchung von Penotti (Penotti 1990) untermauert, in der verschiedene Sequenzvariationen von TATA-Boxen auf Promotoraktivität geprüft wurden.

Die Mutation der TATA-Box erzeugte in GH-Zellen keinen negativen Effekt auf die Aktivität im Luziferase-Assay (Abb. 2-5 B, Säule 3). Die Aktivität stieg leicht auf 120 % an, unterschied sich aber nicht signifikant von der Wildtyp-LTR pcK30. Die Bindung von TFIID scheint also über ein anderes Sequenzelement auf der HERV-K LTR ermöglicht zu werden.

Die Ergebnisse dieser Mutationsanalyse weisen stark auf eine TATA-Box unabhängige Transkriptionskontrolle hin. Eine Analyse des 5'-Endes, der in GH-Zellen vorkommenden HERV-K mRNA, sollte weitere Hinweise auf den Core Promotor geben.

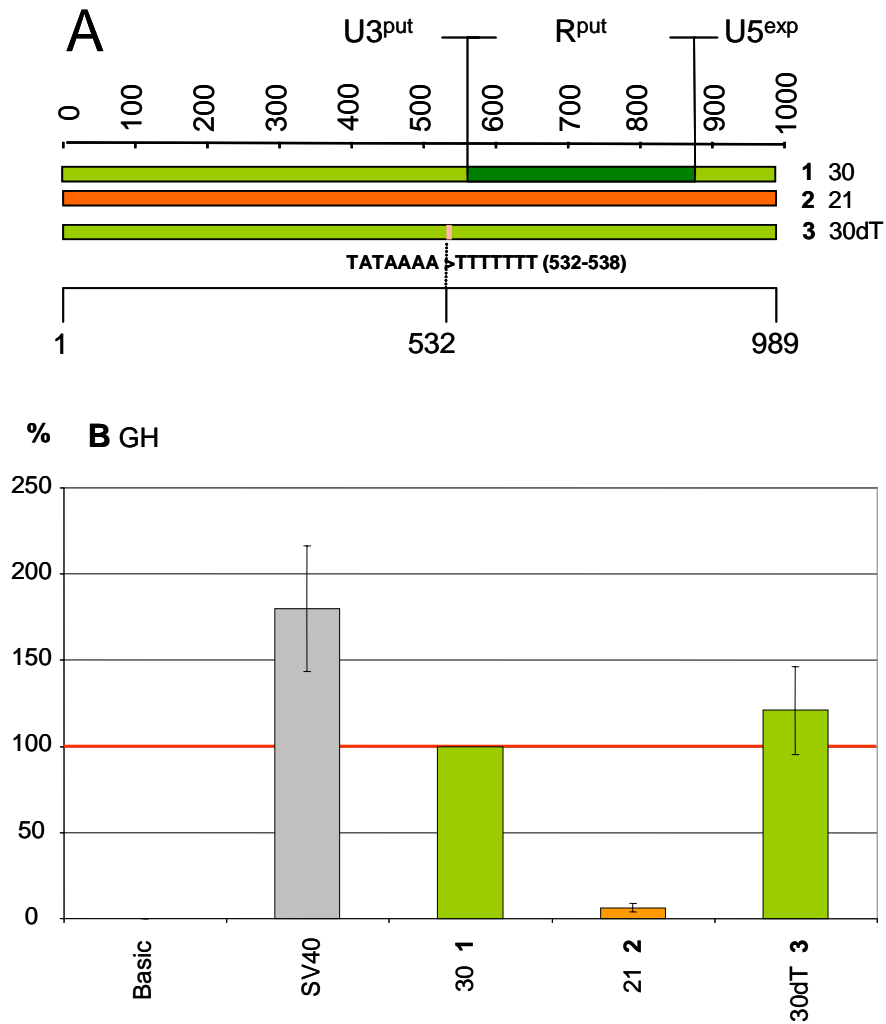


Abbildung 2-5: Relative Luziferase-Aktivität der LTR-Konstrukte. **A** Konstruktübersicht. Zahlen geben die Position im HERV-K LTR in bp an. put = putativ, exp = experimentell ermittelt. **B** Relative Luziferase-Aktivität verschiedener LTR-Konstrukte in GH-Zellen. Die Aktivität in % ist auf die aktive LTR pcK30 (Spur 1, 100 %, rote Linie) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle.

2.1.1.3 Identifikation des Transkriptionsstarts durch eine 5'-RACE

Die Ergebnisse der Mutationsanalyse warfen die Frage nach dem tatsächlichen Transkriptionsstart von HERV-K auf. Dieser kann mit Hilfe der 5'-RACE Technik bestimmt werden. Bei diesem Verfahren wird mRNA einer bestimmten Zelllinie durch eine RT in cDNA umgeschrieben. Durch genspezifische PCRs kann nun der 5'-Bereich einer bestimmten cDNA amplifiziert werden. Durch anschließende Sequenzierung kann man so die Transkriptionsinitiationsstelle bestimmen. Bei den üblichen Verfahren wird nach cDNA-Synthese ein DNA-Oligo an alle vorhandenen 5'-Enden ligiert. Ein Sense Primer, der auf diesem Oligo binden kann, und ein genspezifischer Antisense Primer stromabwärts vom

putativen Transkriptionsstart ermöglichen das Amplifizieren der 5'-cDNA-Enden. Es besteht jedoch immer die Gefahr, dass von unvollständigen Transkripten stammende cDNA amplifiziert wird. Ein besonders kritischer Schritt ist hierbei das Umschreiben der RNA in cDNA durch die RT. Sekundärstrukturen auf der RNA könnten einen vorzeitigen Abbruch der RT-Reaktion herbeiführen. Auch unvollständig transkribierte RNA-Fragmente werden umgeschrieben und werden der RT- und der anschließenden PCR-Reaktion unterworfen. Das korrekte 5'-Ende eines Transkripts bliebe in beiden Fällen verborgen.

Bei dem verwendeten 5'-RACE System wurde versucht, diese Fehlerquellen zu vermeiden (siehe 4.4.11): Statt eines DNA-Oligos wird bereits auf RNA-Ebene ein RNA-Oligo an die Transkripte ligiert. Er wird bei der cDNA-Synthese ebenfalls revers transkribiert und dient dann als Bindestelle für den Sense Primer bei der 5'-PCR. Um sicher zugehen, dass nur vollständige Transkripte der Reaktion zur Verfügung stehen, wird die RNA verschiedenen enzymatischen Behandlungen unterzogen. Eine Phosphatase dephosphoryliert jede unvollständige RNA-Spezies, deren 5'-Ende nicht durch eine Kappe aus einem methylierten G-Nukleotid geschützt ist. Im nächsten Schritt entfernt eine Pyrophosphatase die Kappe der vollständigen Transkripte und entblößt ein Phosphat am 5'-Ende der RNA. Der Ligation des RNA-Oligos stehen nun nur die Phosphate dieser RNAs zur Verfügung. Nach cDNA-Synthese führt die 5'-RACE mit sehr hoher Wahrscheinlichkeit zu dem korrekten 5'-Ende des Transkripts und damit zum Transkriptionsstart.

Um HERV-K Sequenzen zu amplifizieren, wurde folgende Versuchsanordnung befolgt:

Für die 5'-RACE wurden zum einen RNA-Präparationen aus untransfizierten GH- bzw. HeLa-Zellen verwendet, zum anderen Zellen, die mit dem pcK30-Luziferasekonstrukt transfiziert worden sind. Um spezifisch HERV-K cDNA amplifizieren zu können, wurde eine 5'-PCR und sukzessiv eine Nested PCR mit genspezifischen Primern durchgeführt. Die Nested PCR erhöht die Spezifität der PCR-Produkte. Wie in Abbildung 2-6 A zu sehen ist, wurden die Sense Primer für beide PCRs so gewählt, dass sie auf der Sequenz des RNA-Oligos binden können. Der Sense Nested Primer muss hierbei stromabwärts vom Sense Primer der 5'-PCR liegen. Für cDNA, die aus transfizierten Zellen stammte, wurden Antisense Primer generiert, die auf dem Luziferasegen binden können. In diesem Fall muss der Antisense Nested Primer stromaufwärts vom 5'-PCR Antisense Primer liegen. Bei der cDNA aus untransfizierten Zellen konnte der Antisense Primer für die 5'-PCR im HERV-K gag-Gen binden, der Antisense Nested Primer stromaufwärts davon in der PBS.

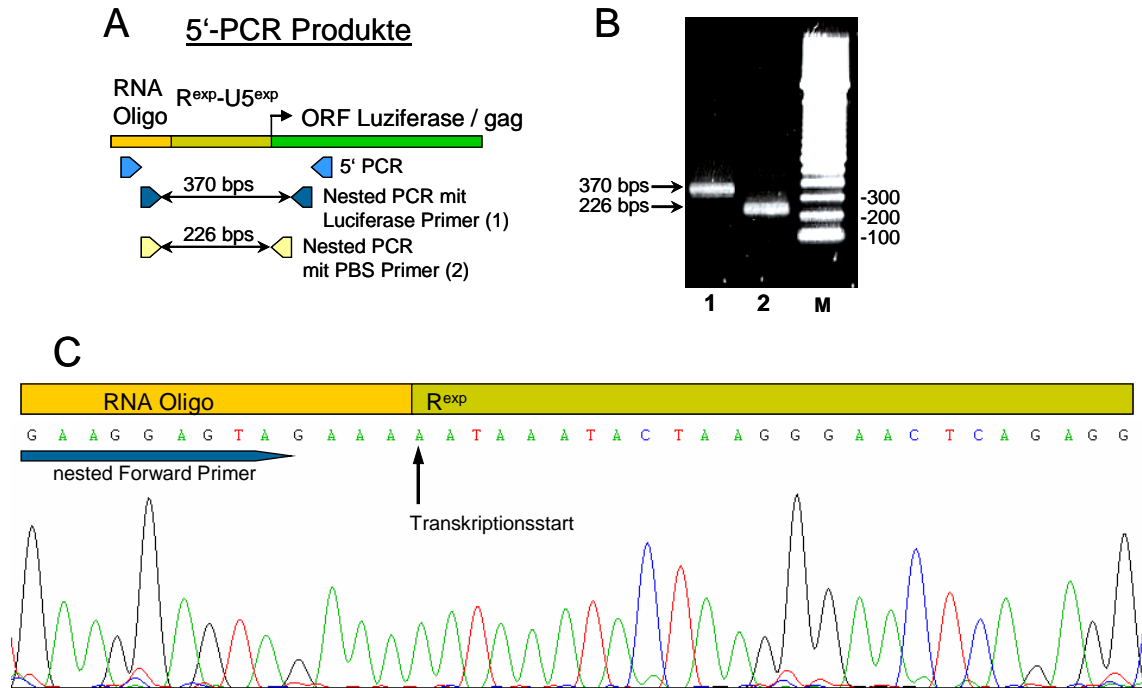


Abbildung 2-6: 5'-RACE. **A** Übersicht über die gewählten Primer zur cDNA-Amplifikation. **B** Agarosegel-Analyse der erhaltenen PCR-Banden. 1 Nested PCR mit Luciferase Primer, 2 Nested PCR mit PBS Primer. M Marker, Größenangabe in bps. **C** Sequenziertes PCR-Produkt. Der Übergang von RNA-Oligo zu HERV-K LTR Sequenz ist gekennzeichnet. exp = experimentell ermittelt.

Der PBS Primer bindet in einem Sequenzbereich, der bei der Analyse der zuvor isolierten LTRs (Hasenmaier 1995) keine Polymorphismen aufwies. Eine Amplifikation aller bekannten transkribierten R-U5 Sequenzen wurde so ermöglicht. Der Vergleich der erhaltenen Sequenzen aus transfizierten und untransfizierten Zellen sollte überprüfen, ob die Ergebnisse des Luciferase-Assays auf die *in-vivo* Situation übertragbar sind. Eine starke Abweichung des ermittelten Transkriptionsstarts bei Plasmid-generierter cDNA von dem Ergebnis der genomischen HERV-K cDNA würde die Aussage des Luciferase-Assays in Frage stellen.

Es zeigte sich für die GH-Präparation ein PCR-Produkt mit einer Größe von 370 bps (Abb. 2-6 B, Spur 1) für die mRNA, die vom Plasmid stammte, und für genomisch transkribierte HERV-K mRNA ein Produkt von 226 bps (Abb. 2-6 B, Spur 2).

Die Sequenzierung der beiden Produkte und der Vergleich mit der pcK30 Sequenz zeigte übereinstimmend einen Beginn der Transkripte bei bp 794 (Abb. 2-6, C). Dieses Ergebnis erklärt den starken Abfall der Promotoraktivität der Deletionsmutanten 30dK und 30dA in GH-Zellen, da hier offenbar Transkriptionsinitiationssequenzen deletiert wurden (Abb. 2-4, Säulen 5 bis 7).

Die 5'-RACE der HeLa-RNA führte zu keinen Produkten (Daten nicht gezeigt). Auch dieses Ergebnis stimmt mit den Beobachtungen aus den Luciferasepromotorstudien überein. Eine

anders gruppieren (siehe Abb. 2-2), würden sich in mehreren Punktmutationen vom RACE-Produkt unterscheiden. Die Tatsache, dass in GH-Zellen vorzugsweise ein bestimmtes Provirus exprimiert wird, weist zusätzlich auf epigenetische Regulationsmechanismen hin.

Auf den Transkriptionsstart fällt gleichzeitig das Poly(A)-Signal, das gefolgt von der Poly(A)-Stelle bei bp 878 ($U5^{exp}$) (Löwer et al. 1993a) den R-U5 Übergang festlegt (Abb. 2-7 A).

In Abbildung 2-7 B sind die ermittelten LTR-Regionen schematisch dargestellt. Der U3-Bereich von HERV-K reicht von bp 1 - 793, der R-Bereich beginnt bei bp 794 und endet bei bp 879 und hat damit eine Länge von 86 bps. Der U5-Bereich schließt die LTR mit den bps 880 - 968 ab.

Eine computergestützte Datenbankanalyse mit dem Transkriptionsfaktorsuchprogramm TESS ergab für den Bereich der bps 750 - 800 der pcK30 zwei Sp1-Konsensussequenzen unterschiedlicher Nukleotidsequenz direkt stromaufwärts vom Transkriptionsstart (Abb. 2-7 A, Sp1; auch GC/GT-Box: Terminus, der sich auf die ersten beschriebenen GC/GT-reichen Sp-Protein-Konsensussequenzen bezieht). Zusätzlich überlappt der Transkriptionsstart mit einer Inr-Konsensussequenz. Ob diese Elemente funktionale Relevanz aufweisen, wurde durch weitere Mutationskonstrukte im Luziferase-Assay überprüft. Ob Sp-Proteine an der Transkription von HERV-K beteiligt sind, wurde in EMSA-Analysen untersucht (siehe 2.2.1). Eine Übersicht über die Sp1-Konsensussequenzen auf der pcK30 ist im Anhang zu finden (siehe 7.2).

2.2 Die Initiation der Transkription

Promotoren, die nicht durch ein TATA-Signal initiiert werden, können durch ein Inr die notwendigen Sequenzen für eine Initiation bereitstellen. Ein Inr hat die Konsensussequenz Py Py A⁺¹ N T Py Py.

Für pcK30 liegt die Sequenz $TCA^{+1}ATAA$ vor (Abb. 2-7, bp 792). Die wichtigsten Kriterien für ein funktionales Inr sind die Nukleotide A und T an den Positionen +1 und +3 der Konsensussequenz (Lo und Smale 1996). Das erste transkribierte Nukleotid ist das Adenosin. Eine Abweichung von der Konsensussequenz ist an Position +4 und +5 zu sehen. Hier weist die LTR zwei Purine statt Pyrimidine auf (Abb. 2-7, bps 797/798).

Inrs können mit Hilfe von Sp-Proteinen die Transkription aktivieren. Mit der TESS-Datenbankanalyse konnten zwei GC/GT-Boxen in der Nähe des Transkriptionsstarts gefunden werden (Abb. 2-7 A, Sp1).

Sp-Proteine gehören zu den Zink-Finger-Proteinen. Neben Sp1, dem ersten isolierten Mitglied dieser Proteinfamilie (Dyanan und Tjian 1983b), wurden weitere Proteine dieser Familie zugeordnet. Sp3 und Sp4 besitzen eine hochkonservierte DNA-Bindungs-Domäne zu Sp1 (Hagen et al. 1992). Sp2 hat eine weniger konservierte DNA-Bindungs-Domäne (Kingsley und Winoto 1992). Dementsprechend binden Sp1, Sp3 und Sp4 die klassische GC/GT-Box mit identischer Affinität (Hagen et al. 1992, 1994). Neben Sp1 - Sp4 sind BTEB1 und BTEB2 isoliert worden, die eine ähnliche Bindungsspezifität aufweisen, sich aber in den aminoterminalen Aktivierungsdomänen stark unterscheiden (Sogawa et al. 1993).

Überlappend mit der Inr-Sequenz findet sich mit dem TESS-Suchprogramm eine Konsensussequenz für TFIID. Inrs besitzen normalerweise keine Konsensussequenz für TFIID oder TBP, sondern die Rekrutierung von TFIID erfolgt vermutlich indirekt über Wechselwirkungen mit TAF_{II}135 und CIF150 direkt am Inr (Kaufmann et al. 1998). Die Abweichungen an Position +4 und +5 des HERV-K Inr gegenüber der Konsensussequenz erzeugen an dieser Stelle eine TFIID-Bindungsstelle. Da nicht klar ist, ob TFIID und damit auch die PolII bei HERV-K Promotoren über Bindung zu anderen Transkriptionsfaktoren oder direkt über eine Bindungssequenz rekrutiert wird, wäre eine Bindungsanalyse dieser Konsensussequenz in weiterführenden Experimenten sinnvoll.

2.2.1 Sp1 und Sp3 binden an GC/GT-Boxen stromaufwärts vom Inr

Um die Hypothese zu überprüfen, dass Sp1 oder andere Mitglieder der Sp-Familie an den HERV-K Promotor binden können, wurde ein beide GC/GT-Boxen enthaltendes DNA-Stück radioaktiv markiert und mit nukleären Proteinextrakten behandelt. Die Position der Sonde im Vergleich zur LTR ist in Abbildung 2-8 C, 30-1 dargestellt. Die Sonde wurde mit GH- und HeLa-Proteinextrakten inkubiert, um gewebespezifische Unterschiede bzw. Gemeinsamkeiten zu bestimmen. Das Laufverhalten der resultierenden DNA-Proteinkomplexe wurde im EMSA überprüft.

Die Inkubation der Sonde mit nukleären GH-Extrakten generierte im EMSA zwei prominente Banden (Abb. 2-8 A, Spur 2 Pfeile 1 und 2), die mit Antikörpern gegen Sp1 und Sp3 einen Supershift erzeugten (Abb. 2-8 A, Spur 3 Pfeil S1, Spur 4 Pfeil S2). Anti-Sp1 konnte den oberen Protein-DNA-Komplex 1 in seiner Migrationsgeschwindigkeit hemmen, anti-Sp3 verlangsamte den unteren Komplex 2. Bei Inkubation mit Antikörpern gegen Sp2, Sp4, BTEB1 und BTEB2 konnte keine Supershift detektiert werden (Daten nicht gezeigt). Die Ergebnisse deuten darauf hin, dass Sp1 und Sp3 die Transkriptionsfaktoren sind, die an die GC/GT-Boxen der Sonde binden können.

Die Spezifität der Protein-DNA-Komplexe wurde mit Kompetitionen überprüft. 50-facher molarer Überschuss kalter Sonde 30-1 führte zur Konkurrenz beider Komplexe (Abb. 2-8 A, Spur 5). Die Inkubation mit dem entsprechenden DNA-Stück aus der LTR21 führte zu einer sehr viel schwächeren Konkurrenz (Abb. 2-8 A, Spur 6). Das bedeutet, dass die Punktmutationen der bps 764, 771/772 und 786/787 eine starke Bindung der Sp-Proteine verhindern können. Dies wurde durch die radioaktive Markierung der Sonde 21-1 und anschließender Inkubation mit Proteinextrakten bestätigt. Die oben beschriebenen Banden konnten nur sehr schwach nach zehnmal längerer Expositionzeit des Films detektiert werden (Daten nicht gezeigt).

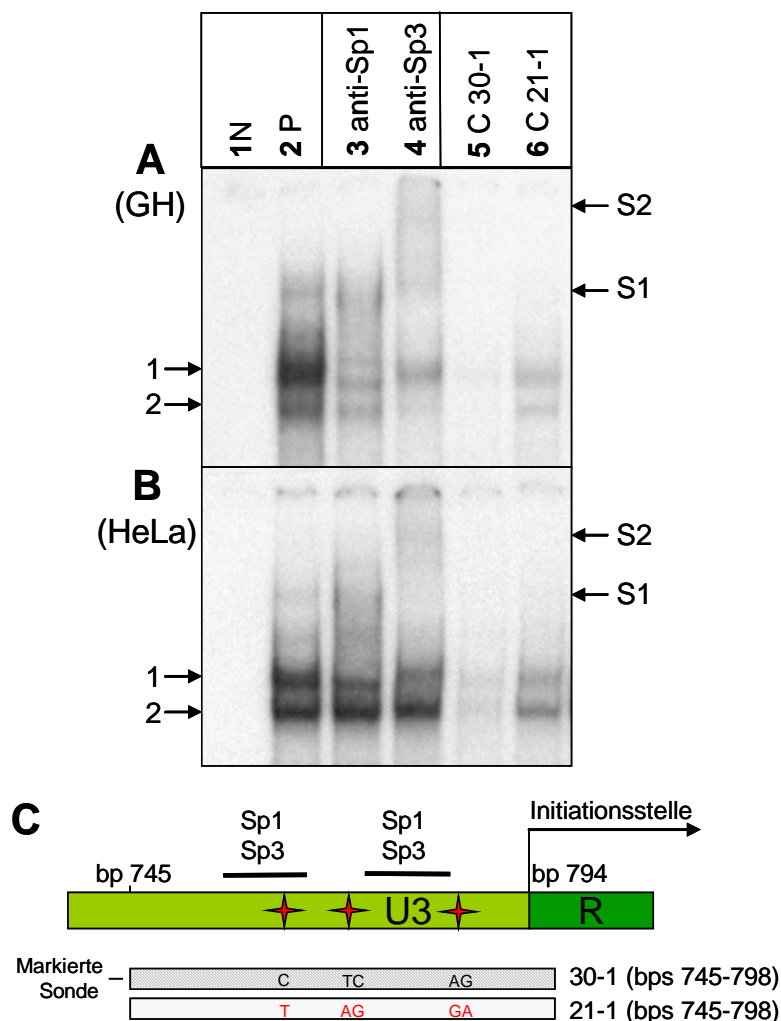


Abbildung 2-8: EMSA mit pck30 Subfragment 30-1. **A** nukleärer GH-Proteinextrakt **B** nukleärer HeLa-Proteinextrakt **C** Übersicht über pck30- und LTR21-Subfragmente im Vergleich zum HERV-K LTR. Zahlen geben die Position im LTR in bp an. Sp-Protein Konsensussequenzen sind durch – gekennzeichnet. **1 N** Negativkontrolle; markierte DNA-Sonde. **2 P** Positivkontrolle; DNA-Sonde inkubiert mit nukleärem Proteinextrakt. **3, 4** Supershiftanalyse mit den Antikörpern anti-Sp1 bzw. anti-Sp3. **5, 6 C** Konkurrenz mit 50-fachem molarem Überschuss der unmarkierten Fragmente 30-1 bzw. 21-1. **Pfeil 1:** Sp1/DNA-Komplex; **Pfeil 2:** Sp3/DNA-Komplex; **Pfeil S1:** Supershift Sp1/DNA-Komplex; **Pfeil S2:** Supershift Sp3/DNA-Komplex.

Für HeLa-Extrakte zeigte sich ein sehr ähnliches Bild. Allerdings erfolgte durch anti-Sp3-Inkubation keine vollständige Verschiebung der unteren Sp3-haltigen Bande (Abb. 2-8 B, Spur 4 Pfeile 2 und S2). Das kann zum einen für eine erhöhte Sp3-Konzentration in HeLa-Extrakten sprechen, zum anderen für einen weiteren Protein-DNA-Komplex unbekannter Zusammensetzung, der die gleiche Migrationsgeschwindigkeit hat wie der Sp3-DNA-Komplex.

Die Spezifität des Protein-DNA-Komplexes und der Antikörper wurde mit rekombinanten Sp1- und Sp3-Proteinen überprüft. Die rekombinanten Proteine erzeugten Banden, die die gleiche Migrationsgeschwindigkeit hatten wie die Banden, die durch Inkubation der Sonden mit nukleären Extrakten hervorgerufen wurden (Abb. 2-9, vergleiche Spur 1 mit Spuren 2 und 4).

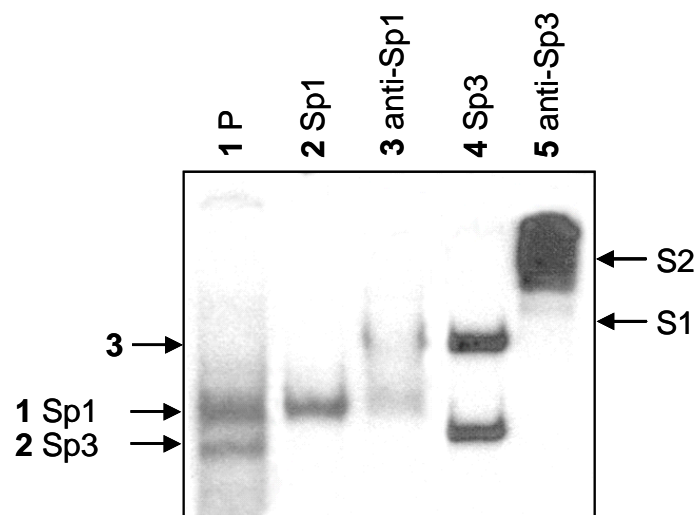


Abbildung 2-9 EMSA mit pck30 Subfragment 30-1. 1 P DNA-Sonde inkubiert mit nukleärem GH-Proteinextrakt. 2 DNA-Sonde inkubiert mit rekombinantem Sp1. 3 DNA-Sonde inkubiert mit rekombinantem Sp1 und anti-Sp1. 4 DNA-Sonde inkubiert mit rekombinantem Sp3 und anti-Sp3. 5 DNA-Sonde inkubiert mit rekombinantem Sp3 und anti-Sp3. **Pfeil 1:** Sp1/DNA-Komplex; **Pfeil 2:** Sp3/DNA-Komplex; **Pfeil 3:** Sp-Protein/DNA-Komplex **Pfeil S1:** Supershift Sp1/DNA-Komplex; **Pfeil S2:** Supershift Sp3/DNA-Komplex.

Das rekombinante Sp3 erzeugte mit der Sonde 30-1 eine weitere langsamer migrierende Bande, die mit nukleären Extrakten nur schwach zu sehen war (Abb. 2-9, Spur 4 Pfeil 3). Dies hängt mit der Anwesenheit von zwei GC/GT-Boxen auf der Sonde zusammen, die beide für Sp3 zugänglich sind. Es können offenbar zwei Arten von Proteinkomplexen entstehen, bei der schnell wandernden Art ist nur eine GC/GT-Box durch Sp3 besetzt, bei der langsam wandernden sind beide Boxen durch Sp3-Proteine besetzt. In nukleären Extrakten, in denen Sp1 und Sp3 um die Bindung konkurrieren, war die Bande nur sehr schwach zu erkennen.

Inkubation von Sonden, die nur eine GC/GT-Box enthielten (z.B. Sonde 30-3, siehe 2.2.2), mit Sp3 führte zu der unteren singulären Bande (Daten nicht gezeigt). Inkubation mit Antikörpern gegen Sp1 und Sp3 führte zur Extinktion bzw. zum Supershift der oben besprochenen Banden (Abb. 2-9, Spur 3 Pfeil S1, Spur 5 Pfeil S2). Die Supershifts glichen den Shifts, die mit nukleären Extrakten beobachtet wurden, fielen aber deutlicher aus (vergleiche Abb. 2-8, Spuren 2-4 mit Abb. 2-9, Spuren 2-5). Bei Inkubation mit nukleären Extrakten, in denen beide Transkriptionsfaktoren, Sp1 sowie Sp3, vorhanden sein können, konkurrieren die Proteine um die Bindungsstellen, was zu einer Übereinanderlagerung der verschiedenen Protein-DNA-Komplexe führen kann. Wie bereits bei Lania et al. und Suske beschrieben (Lania et al. 1997; Suske 1999), besitzen Sp1- und Sp3-DNA-Komplexe eine ähnliche Migrationsgeschwindigkeit. Die resultierende Doppelbande ist im elektrischen Feld nur schwer voneinander zu trennen. Zusätzlich könnte die Auftrennung durch weitere potentielle Bindungspartner für die verwendete DNA-Sonde erschwert werden.

Zusammenfassend waren dies starke Hinweise auf eine TATA-unabhängige Transkription, die durch ein Inr und Sp-Proteine initiiert werden kann. Eine starke Bindung der Sp-Proteine an die GC/GT-Boxen erfolgte nur bei aktiven LTRs. Ob beide GC/GT-Boxen gleichermaßen von den Transkriptionsfaktoren gebunden werden können und in welcher Weise sich die Punktmutationen auf die einzelnen GC/GT-Boxen ausüben, wird in den nächsten Kapiteln behandelt. Die Sonden 30-3 und 30-5 enthielten jeweils nur eine GC/GT-Box und sollten diesen Sachverhalt weiter aufklären.

2.2.2 EMSA-Analyse der ersten Sp-Bindungsstelle

Sonde 30-3 umfasst die erste GC/GT-Box. In Abbildung 2-10 C ist die Position der Sonde und der abgeleiteten für die Kompetition verwendeten Subfragmente im Vergleich zur LTR dargestellt. Bei Inkubation mit den bereits in Kapitel 2.2.1 verwendeten Proteinextrakten erkennt man im EMSA drei Banden (Abb. 2-10, A und B, Spur 2 Pfeile 1-3). Die Banden 1 und 2 ließen sich mit Antikörpern gegen Sp1 und Sp3 in ihrer Migrationsgeschwindigkeit verlangsamen. Bande 1 reagierte mit anti-Sp1 und Bande 2 mit anti-Sp3 (Abb. 2-10, A und B, Spur 2, Spur 3 S1, Spur 4 S2). Bande 3 reagierte nicht mit den Antikörpern gegen Sp-Proteine. Sie entspricht dem bereits unter 2.2.1 erwähnten Protein-DNA-Komplex unbekannter Zusammensetzung und ist bei HeLa-Extrakten sehr viel stärker ausgeprägt als bei GH-Extrakten (vergleiche Abb. 2-10, A und B, Spur 2 Pfeil 3). Da die Sonde nur eine GC/GT-Box enthielt, kann man hier die drei Protein-DNA-Komplexe voneinander unterscheiden.

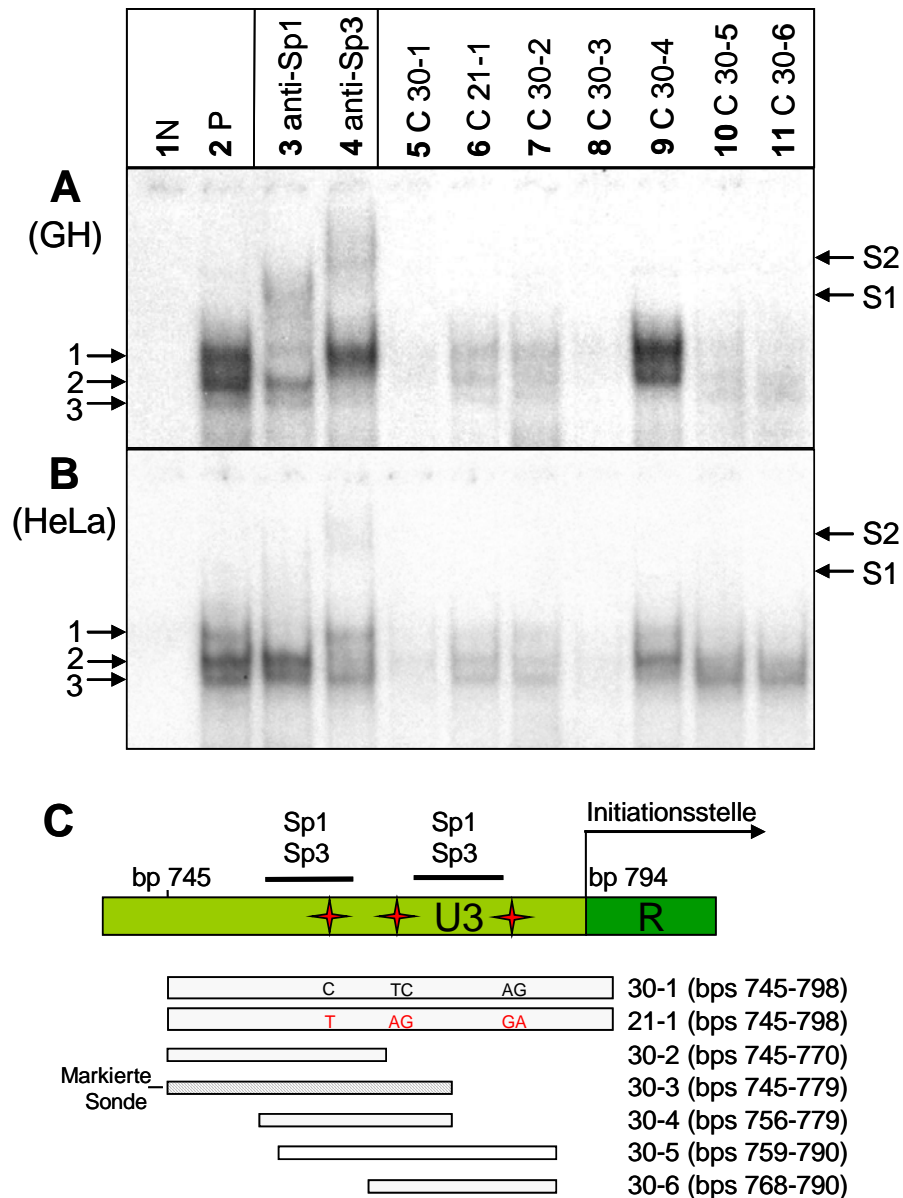


Abbildung 2-10 EMSA mit pcK30 Subfragment 30-3. A nukleäre GH-Proteinextrakte. **B** nukleäre HeLa-Proteinextrakte. **C** Übersicht über pcK30- und LTR21-Subfragmente im Vergleich zum HERV-K LTR. Zahlen geben die Position im LTR in bp an. Sp-Protein Konsensussequenzen sind durch – gekennzeichnet. **1 N** Negativkontrolle; markierte DNA-Sonde. **2 P** Positivkontrolle; DNA-Sonde inkubiert mit nukleärem Proteinextrakt. **3, 4** Supershiftanalyse mit Antikörpern anti-Sp1 und anti-Sp3. **5-11 C** Konkurrenz mit 50-fachem molarem Überschuss der unmarkierten Subfragmente 30-1 bis 30-6 und 21-1. **Pfeil 1:** Sp1/DNA-Komplex; **Pfeil 2:** Sp3/DNA-Komplex; **Pfeil 3:** unbekannter Protein-DNA-Komplex; **Pfeil S1:** Supershift Sp1/DNA-Komplex; **Pfeil S2:** Supershift Sp3/DNA-Komplex.

Wie zu erwarten, ließen sich alle Banden spezifisch mit dem beide GC/GT-Boxen enthaltenden LTR-Fragment 30-1 kompetitieren (Abb. 2-10, A und B, Spur 5). Das Pendant aus der LTR21 konnte, wie auch bei Sonde 30-1, nur schwach kompetitieren (Abb. 2-10, A und B, Spur 6).

Weitere Konkurrenzexperimente mit Subfragmenten sollten aufzeigen, welche Bereiche des beide GC/GT-Boxen enthaltenden Fragments 30-1 für die Konkurrenz verantwortlich sind. Die Subfragmente wurden so gewählt, dass sie die GC/GT-Boxen jeweils einzeln und in verschiedenen Überlappungsbereichen abdeckten (Abb. 2-10, C).

Das Subfragment 30-2 besaß zwar die Sp1-Konsensussequenz der ersten GC/GT-Box, war aber nicht in der Lage, im gleichem Maße wie 30-1 zu konkurrieren (Abb. 2-10, A und B, Spur 7). Dies könnte daran liegen, dass sich 3' von der GC/GT-Box nur weitere 3 bps befanden, was eine stabile Bindung von Sp1 oder Sp3 an den Konkurrenten verhindern könnte. Erst eine Verlängerung des Abschnitts um 9 bps im Subfragment 30-3 ermöglichte die vollständige Konkurrenz der drei Banden (Abb. 2-10, A und B, Spur 8).

Interessanterweise ließ sich die Bande 3 mit Subfragment 30-4 konkurrieren (Abb. 2-10, A und B, Spur 9 Pfeil 3). Dieses Subfragment beinhaltet die erste GC/GT-Box, jedoch nur mit einem Puffer von einem bp 5' von der Konsensussequenz. Es konnte die Sp1- und Sp3-haltigen Banden nicht konkurrieren. Auch konnten die Subfragmente 30-5 und 30-6 zwar die Sp-Proteinbanden 1 und 2 konkurrieren (beide Fragmente beinhalten die zweite GC/GT-Box), jedoch nicht Bande 3 (Abb. 2-10, A und B, Spuren 10 und 11). Die Konkurrenz der Sp-Protein-haltigen Banden war nicht so stark ausgeprägt wie bei den Subfragmenten 30-1 und 30-3 (Abb. 2-10, A und B, vergleiche Spuren 5 und 8 mit Spuren 10 und 11). Dies ist ein Hinweis darauf, dass die Sp-Proteine eine stärkere Bindung zur ersten GC/GT-Box aufbauen können.

Die Ergebnisse weisen zum einen darauf hin, dass beide GC/GT-Boxen Sp1 und Sp3 binden können, es werden jedoch einige bps stromaufwärts wie auch –abwärts benötigt, um eine stabile Bindung zu etablieren. Zum anderen haben sich Hinweise auf die Bindung eines weiteren noch nicht identifizierten Proteinkomplexes, Bande 3, ergeben. Diese Bindungsstelle wird durch das Subfragment 30-4 abgedeckt.

2.2.3 EMSA-Analyse der zweiten Sp-Bindungsstelle

Die Sonde 30-5 umfasst die zweite GC/GT-Box. Die Sequenz der ersten GC/GT-Box ist nur partiell vorhanden. Eine stabile Bindung an dieser unvollständigen Konsensussequenz kann nicht erfolgen. In Abbildung 2-11 C ist die markierte Sonde relativ zur LTR dargestellt.

Die Inkubation mit den Proteinextrakten aus GH- und HeLa-Zellen erzeugte zwei Protein-DNA-Komplexe (Abb. 2-11, A und B, Spur 2 Pfeile 1 und 2). Bande 1 ließ sich durch anti-Sp1 verzögern (Abb. 2-11, A und B, Spur 3 Pfeil S1). Bande 2 reagierte mit anti-Sp3 (Abb. 2-11, A und B, Spur 4 Pfeil S2). Analog der Sonde 30-3 waren die Subfragmente 30-1,

30-3, 30-5 und 30-6 in der Lage, die Sp-DNA-Komplexe zu kompetitieren (Abb. 2-11, A und B, Spuren 5, 8, 10 und 11). Subfragmente 21-1 und 30-2 konnten nur unvollständig (Abb. 2-11, A und B, Spuren 6 und 7) und 30-4 konnte nicht kompetitieren (Abb. 2-11, A und B, Spur 9). Die unvollständige Competition des Subfragments 21-1 kommt, wie bei

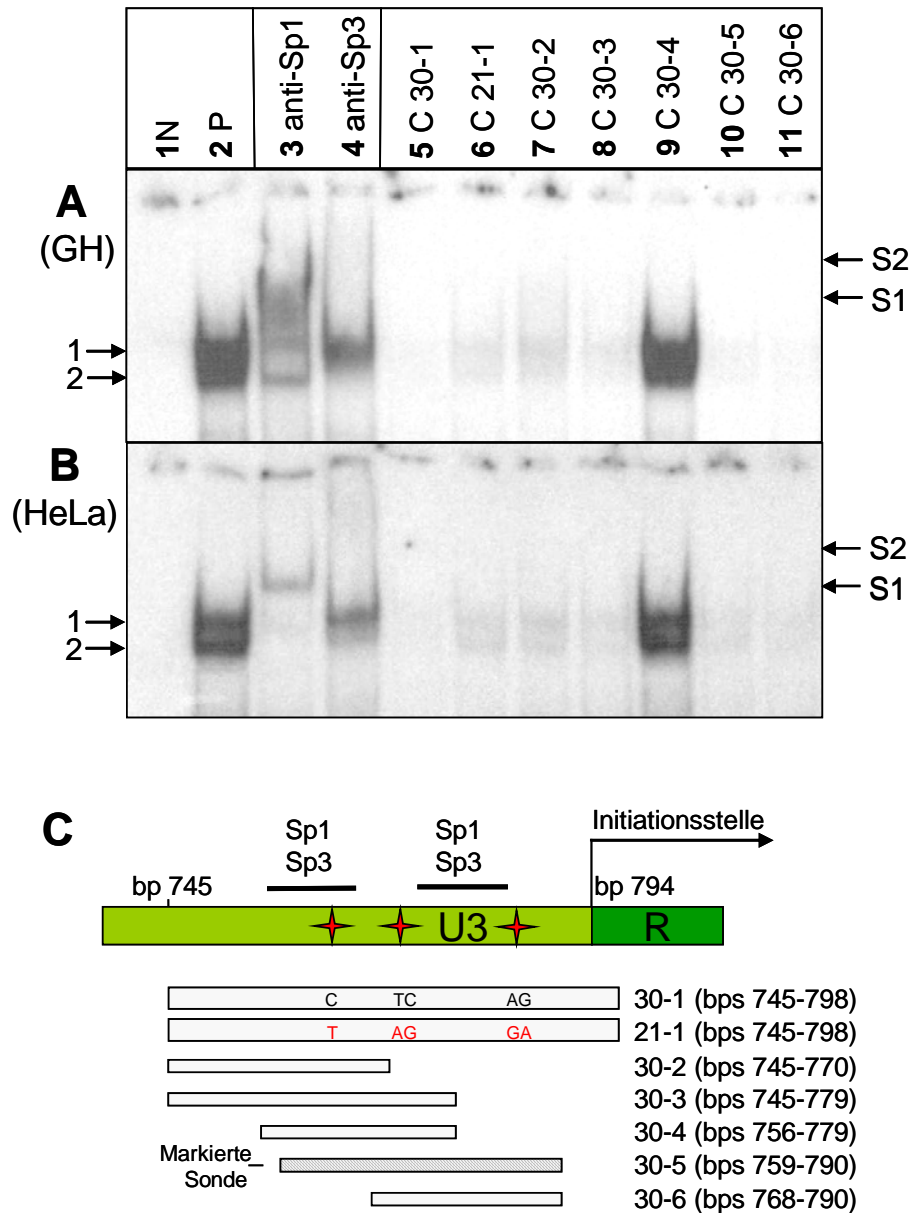


Abbildung 2-11 EMSA mit pck30 Subfragment 30-5. **A** nukleäre GH-Proteinextrakte. **B** nukleäre HeLa-Proteinextrakte. **C** Übersicht über pck30- und LTR21-Subfragmente im Vergleich zum HERV-K LTR. Zahlen geben die Position im LTR in bp an. Sp-Protein Konsensussequenzen sind durch – gekennzeichnet. **1 N** Negativkontrolle; markierte DNA-Sonde. **2 P** Positivkontrolle; DNA-Sonde inkubiert mit nukleärem Proteinextrakt. **3, 4** Supershiftanalyse mit Antikörpern anti-Sp1 und anti-Sp3. **5-11 C** Competition mit 50-fachem molarem Überschuss der unmarkierten Subfragmente 30-1 bis 30-6 und 21-1. **Pfeil 1:** Sp1/DNA-Komplex; **Pfeil 2:** Sp3/DNA-Komplex; **Pfeil S1:** Supershift Sp1/DNA-Komplex; **Pfeil S2:** Supershift Sp3/DNA-Komplex.

2.2.1 besprochen, durch Punktmutationen zustande, die eine stabile Bindung der Sp-Proteine verhindern. Die Bindungskapazität des Kompetitors wurde dadurch herabgesetzt. Das gleiche gilt für Subfragment 30-2, hier fehlen für die Bindung wichtige Sequenzen vom 3'-Ende der Konsensussequenz. Sobald mindestens eine funktionale GC/GT-Box auf dem Kompetitor vorhanden war, konnte eine vollständige Konkurrenz erreicht werden. Bei dem Fragment 30-4 zeigte sich bereits bei der Sonde 30-3, dass die Sp-Proteinkomplexe nicht kompetitiert werden konnten. Die kompetitierbare Bande 3 (Abb. 2-10, A und B, Spur 9 Pfeil 3) zeigte sich auf dieser Sonde nicht.

Zusammenfassend kann man sagen, dass zwischen den bps 745 und 798 zwei GC/GT-Boxen lokalisiert sind, die beide in der Lage sind, Sp1 und Sp3 zu binden. Darüber hinaus kann aus den Konkurrenzstudien mit Subfragment 30-4 geschlossen werden, dass ein noch nicht identifiziertes Protein auf den bps 756-779 binden kann. Nimmt man die Beobachtung mit in Betracht, dass immer einige Nukleotide stromaufwärts oder -abwärts notwendig sind, um eine stabile Protein-DNA Bindung zu erzeugen, engt das den Bereich für eine Konsensussequenz weiter ein. Eine TESS-Datenbankabfrage zeigte in diesem Bereich Konsensussequenzen mit hoher logarithmischer Wahrscheinlichkeit für die Transkriptionsfaktoren c-Ets-1 und für den Glucocorticoid Rezeptor an, doch liegen für beide Proteine die Konsensussequenzen sehr dicht am Ende des Subfragments. Ob eines dieser Proteine tatsächlich binden kann, müsste noch untersucht werden.

Nach dem *in-vitro* Beweis, dass Sp1 und Sp3 an die GC/GT-Boxen binden können, sollte nun eine Western Blot Analyse die intrazelluläre Verteilung von Sp1 und Sp3 aufklären, um Rückschlüsse auf die transkriptionelle Regulation *in-vivo* ziehen zu können.

2.2.4 Western Blot Analyse von Sp1 und Sp3

Es sind bereits viele Gene charakterisiert worden, die von Sp-Proteinen und Inrs reguliert werden (Kiela et al. 2002; O'Leary et al. 1996; Sjøttem et al. 1996; Yajima et al. 1998). Aktivierende Effekte auf die Transkription sind für Sp1 beschrieben worden. Sp3 kann aktivierende, wie auch repressorische Wirkung auf einen Promotor haben (Hagen et al. 1994; Kennett et al. 1997). Nach Entdeckung von verschiedenen Sp3-Isoformen (Kennett et al. 1997) schrieb man vor allem dem Zusammenspiel dieser Isoformen und dem Verhältnis der Konzentrationen von Sp1 und Sp3 im Nukleus aktivierende bzw. reprimierende Effekte auf die Transkription zu (Suske 1999). Die Analyse dieser Proteine in nukleären Extrakten mit Hilfe der Western Blot Methode sollte Anhaltspunkte für die Regulation der HERV-K LTR geben.

Sp1 und Sp3 wurden bereits ausgiebig in Western Blot Analysen untersucht. Dabei zeigten sich für Sp1 und Sp3 unterschiedliche Molekulargewichte. Man schreibt dieses unterschiedliche Laufverhalten posttranslationalen Modifikationen wie z.B. Glykosylierung zu.

Sp1 besteht aus 785 Aminosäuren und hat ein Molekulargewicht von 80,7 kDa (Accession number: P08047). In der Western Blot Analyse kann man eine 95 kDa Isoform und eine phosphorylierte ca. 105 kDa große Isoform detektieren (Jackson et al. 1990).

Sp3 besteht aus 781 Aminosäuren und hat ein Molekulargewicht von 81,9 kDa (Accession number: Q02447). Darüberhinaus existieren zwei kleinere Isoformen, die über interne Startcodons entstehen. Die intern initiierte Isoform 1 besteht aus 496 Aminosäuren und hat ein Molekulargewicht von 53,7 kDa (Accession number: Q8WWU3), die intern initiierte Isoform 2 besteht aus 479 Aminosäuren und hat eine Größe von 52,1 kDa (Accession number: Q8WWU2). Alle Isoformen werden von einer mRNA generiert (Hagen et al. 1992; Kingsley und Winoto 1992). Zusätzliche Informationen zu den Sp-Proteinen sind im Anhang zu finden (siehe 7.3).

In Abbildung 2-12 wurden die Konzentrationen von Sp1 und Sp3 in den nukleären Extrakten verglichen. Um zu gewährleisten, dass gleiche Mengen an Ausgangsprotein verglichen wurden, wurde mit anti-Actin normiert. Die gleichzeitige Analyse der cytoplasmatischen Fraktionen gewährleistete eine Überprüfung der Güte der Extraktpräparationen. Transkriptionsfaktoren sind ausschließlich im Kern zu finden. Das Ausbleiben von Signalen in den cytoplasmatischen Fraktionen zeigte, dass während der Präparation der Extrakte die Kerne nicht zu früh lysiert worden sind.

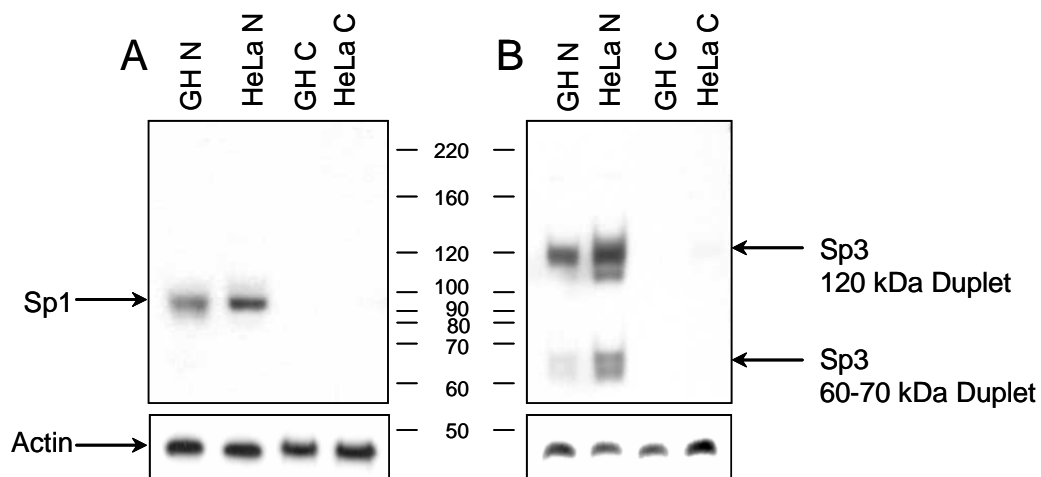


Abbildung 2-12 Western Blot Analyse von Sp1 und Sp3. Je 10 µg GH- bzw. HeLa-Zellextrakt aufgetrennt im 8 %igem SDS-PAGE Gel. Normierung durch anti-Actin 1:1000. Markergrößenangabe in kDa. N = nukleärer Proteinextrakt, C = cytoplasmatischer Proteinextrakt. **A** Erst-Antikörper anti-Sp1, 1:1000. **B** Erst-Antikörper anti-Sp3, 1:3000.

Sp1 wurde als singuläre Bande detektiert (Abb. 2-12 A). Die phosphorylierte Isoform von 105 kDa war nicht zu sehen. Die Sp1-Konzentration bei GH-Zellen war leicht geringer als in Extrakten aus HeLa-Zellen (Abb. 2-12 A, vergleiche GH N/HeLa N).

Im Western Blot konnte man für Sp3 ein Proteinduplet bei 120 kDa detektieren, ein weiteres Proteinduplet konnte man bei 60-70 kDa erkennen. Auch hier wird das vom berechneten Molekulargewicht abweichende Laufverhalten im elektrischen Feld posttranslationalen Modifikationen zugeschrieben (Hagen et al. 1994; Kennett et al. 1997).

In HeLa-Zellen waren alle in der Literatur beschriebenen Sp3-Isoformen vorhanden. In Abbildung 2-12 B kann man die charakteristischen zwei Proteinduplets erkennen. Das Volllängenprotein ist bei 120 kDa zu finden, die kleineren Isoformen bei ca. 60-70 kDa (Abb. 2-12 B, HeLa N).

Es fällt auf, dass das Volllängenprotein sehr viel stärker in HeLa-Zellen vertreten war, als in GH-Zellen (Abb. 2-12 B, vergleiche GH N und HeLa N). Alle anderen Isoformen waren in GH-Zellen nicht oder nur sehr schwach vorhanden. Dies zeigt einen gewebespezifischen Unterschied für Sp3 auf, der regulatorische Bedeutung haben könnte.

Da beide GC/GT-Boxen fähig sind, Sp1 und Sp3 zu binden und beide Proteine in unterschiedlichen Mengen in den getesteten Proteinextrakten vorhanden sind, sollten Luziferasekonstrukte mit mutierten GC/GT-Boxen die funktionale Relevanz der Transkriptionsfaktorbindestellen überprüfen.

2.2.5 Analyse der Funktionalität der Sp-Bindungsstellen *in-vivo*

Die EMSAs haben gezeigt, dass die Sp-Konsensussequenzen stromaufwärts des Transkriptionstartes *in-vitro* von Sp1 und Sp3 gebunden werden können. Im Reporter-Assay wurde deswegen überprüft, ob die Mutation und Deletion der Bindungssequenzen die transkriptionelle Aktivität beeinflussen können. In Abbildung 2-13 A wird dargestellt, in welcher Weise die Punktmutationen bzw. Deletionen in die Sp-Konsensussequenzen von pCK30 eingefügt worden sind.

Es zeigte sich, dass alle Mutationen die Promotoraktivität drastisch senkten (Abb. 2-13 B, Säulen 3 bis 7). Bei den Konstrukten, die eine mutierte erste GC/GT-Box besitzen, fiel die Aktivität auf 23 % (30dZ I) und 31 % (30dZ III) (Abb. 2-13 B, Säulen 3 und 5). Die Mutation bzw. Deletion der zweiten GC/GT-Box führte zu einem Aktivitätsrückgang auf 49 % (30dZ II) bzw. 64 % (30dZ IV) (Abb. 2-13 B, Säulen 4 und 6). Eine Deletion beider GC/GT-Boxen (30dZ V) führte zu einem Aktivitätsverlust von 57% (Abb. 2-13 B, Säule 7). Interessanterweise führten Veränderungen an der ersten GC/GT-Box zu einem stärkeren

Verlust in der Promotoraktivität als Mutationen der zweiten GC/GT-Box (Abb. 2-13 B, vergleiche Säulen 3 und 5 mit Säulen 4 und 6).

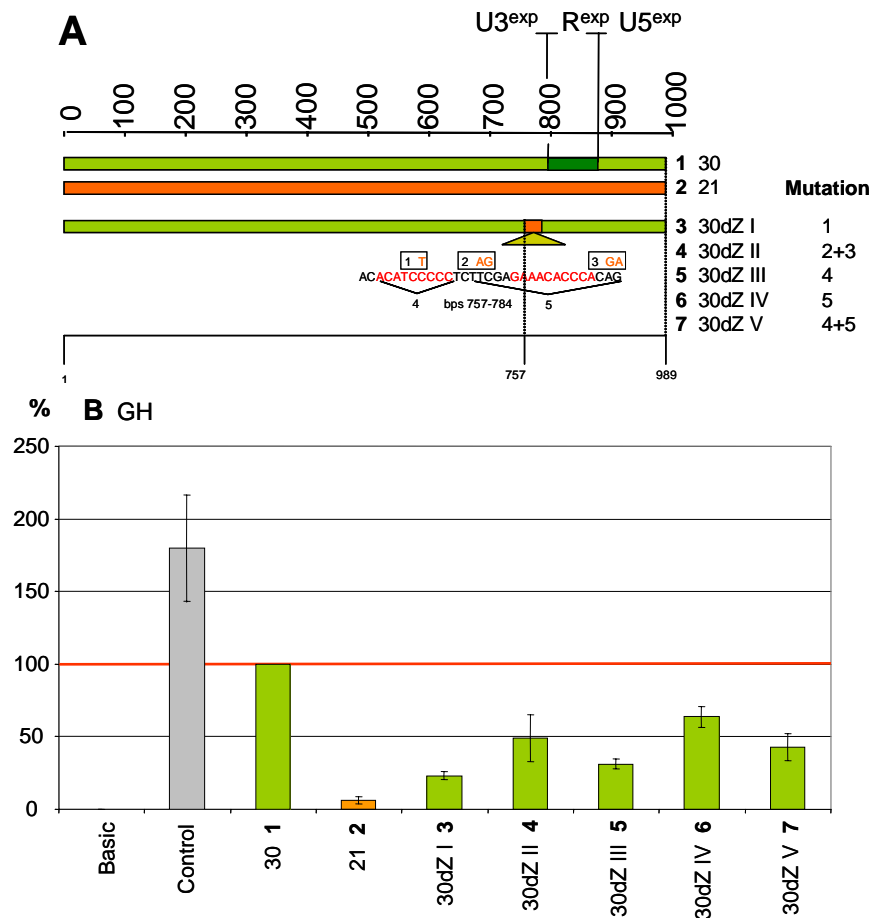


Abbildung 2-13 Relative Luziferase-Aktivität der LTR-Konstrukte. A Konstruktübersicht. Den Konstruktbezeichnungen sind Zahlen zugeordnet, die den Mutationen und Deletionen im vergrößerten Sequenzbereich entsprechen. Die Sequenzen der GC/GT-Boxen sind rot dargestellt, die eingefügten LTR21 Mutationen sind orange eingefärbt. Deletierte Sequenzbereiche sind durch waagrechte Klammern angedeutet. Zahlen geben die Position im HERV-K LTR in bp an. exp = experimentell ermittelt. **B** Relative Luziferaseaktivität verschiedener LTR-Konstrukte in GH-Zellen. Die Aktivität in % ist auf die aktive LTR pck30 (Spur 1, 100 %, rote Linie) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle.

Diese Ergebnisse legen nahe, dass die erste GC/GT-Box eine stärkere Rolle für die Transkription spielt. Mutationen in der zweiten GC/GT-Box führten zu einem geringeren Abfall der Aktivität. Ob ein synergistischer Effekt zwischen beiden GC/GT-Box besteht, müsste noch geklärt werden.

2.2.6 Sp-Bindungsstellen-Screening

Die Analyse der Beteiligung der Sp-Proteine an der Transkription von HERV-K beschränkte sich bis jetzt auf die proximalen Bindungsstellen stromaufwärts des Transkriptionsstarts. Sp1 kann die Transkription jedoch nicht nur als proximales Promotorelement beeinflussen, sondern die Transkription kann auch synergistisch von distal und proximal bindenden Sp-Proteinen stimuliert werden (Courey et al. 1989). *In-vitro* Experimente zeigten bereits, dass eine Aktivierung eines Promotors synergistisch über proximal und distal bindende Sp1-Proteine möglich ist. Die Interaktion der Transkriptionsfaktoren wird durch die Bildung eines DNA-Loops ermöglicht (Mastrangelo et al. 1991).

In einem ersten Versuch, weitere Sp-Bindungsstellen auf dem HERV-K Promotor zu identifizieren, sind im EMSA weitere DNA-Bereiche getestet worden. Ausgehend von der putativen TATA-Box bei bp 532 sind verschiedene Sonden zwischen den bps 515-833 mit

LTR Position	Sp1	Sp3
bps 515-588	-	-
bps 579-652	-	-
bps 647-754	-	-
bps 745-798	+	+
bps 745-833	+	+

Tabelle 1: Zusammenfassung der EMSA-Analyse mit rekombinantem Sp1/Sp3-Protein.

rekombinantem Sp1 und Sp3 inkubiert worden (Tabelle 1). In dem getesteten DNA-Bereich ergaben sich jedoch keine weiteren Sp-Bindungsstellen. Das Binden von Sp1 und Sp3 auf der Sonde bps 745-798 und bps 745-833 erfolgte über die bereits beschriebenen Sp-Bindungsstellen (siehe 2.2.1).

Die Bereiche bps 1-514 und bps 834-989 müssten in weiteren EMSAs überprüft werden. Auch eventuelle Bindestellen in den Übergangsbereichen der verschiedenen Sonden müssten noch untersucht werden. Die Kompetitionsergebnisse aus den Abschnitten 2.2.3 und 2.2.4 zeigten bereits, dass für die Bindung eines Proteins an einer Konsensussequenz zusätzliche Basen stromaufwärts bzw. -abwärts benötigt werden. Das bedeutet, dass putative Bindungsstellen in der Nähe des 5'- oder 3'-Endes einer Sonde zusätzliche bps benötigen. Sonden, die die Bereiche um die bps 580, 650 und 750 umfassen, sollten im EMSA getestet werden. Eine Übersicht über alle putativen GC/GT-Boxen auf der pcK30-LTR ist im Anhang zu finden (siehe 7.2).

2.3 Analyse der U3-, R- und U5-Region

Der Core Promotor von HERV-K initiiert die Transkription in einer TATA-unabhängigen Weise durch ein Inr und aktivierende bzw. reprimierende Sp-Proteine. Ob neben dem bereits

identifizierten 5'-Enhancer YY1 (Knöbl 1998) weitere regulatorische Proteine binden können, sollte eine Mutationsanalyse der U3-, R- und U5-Region zeigen.

2.3.1 Mutationsanalyse U3

Um Promotor-relevante Bereiche zu bestimmen, sind aus pcK30- und LTR21-DNA bestehende chimäre Konstrukte hergestellt worden. Dabei sind zunächst größere Promotorsegmente ausgetauscht worden, um festzustellen, inwieweit sich Punktmutationen in U3, R und U5 auf die Promotoraktivität auswirken. Es wurden immer chimäre Konstrukte mit Deletionskonstrukten verglichen, in denen der substituierte Bereich deletiert worden ist. Alle im Folgenden beschriebenen Konstrukte wurden wieder im transienten Luziferase-Assay auf ihre Promotoraktivität untersucht. Ihre Aktivität wurde prozentual auf die Aktivität der pcK30-LTR bezogen.

Der Promotor von HERV-K ist der Einführung von Punktmutationen gegenüber sehr empfindlich. Beim Austausch großer DNA-Bereiche muss man bedenken, dass man Aktivator- und Repressor-Bindungsstellen gleichermaßen deletiert, mutiert oder generiert. Die Aktivität eines Reportergens ist also immer die Summe aller Effekte, die die Mutationen auf die transkriptionelle Regulierbarkeit haben. Um einen groben Überblick über die relevanten Regionen zu bekommen, ist der Promotor in drei DNA-Segmente unterteilt worden. Über die Schnittstellen SphI und Bsu36I sind die Bereiche I und II ausgetauscht bzw. deletiert worden (Abb. 2-14 A). Der Bereich III, der die R- und U5-Region enthält, wird im Abschnitt 2.3.3 ausführlicher behandelt.

Im Konstrukt 30dd21 sind die ersten 238 bps der pcK30-LTR gegen LTR21-Sequenz ausgetauscht worden. Im Luziferase-Assays fiel dadurch die Promotoraktivität auf 41 % (Abb. 2-14 B, Säule 3). Bei Deletion der ersten 238 bp, wie in Konstrukt 30dB, fiel die Promotoraktivität auf 17 % (Abb. 2-14 B, Säule 5). Vergleicht man Konstrukt 30dd21 mit dem deletierten Pendant, sieht man einen sehr viel stärkeren Aktivitätsverlust bei dem Deletionskonstrukt 30dB. Die transkriptionelle Inaktivierung ist durch eine Abschwächung der YY1-Enhancer Bindungskapazität (Knöbl 1998) zu erklären. YY1 bindet im Bereich bps 62-83 der HERV-K LTR. In der LTR21-Sequenz befindet sich eine Punktmutation bei bp 79. In der Chimäre 30dd21 könnte noch eine verminderte Bindungsaktivität des Faktors YY1 bestehen. Der stärkere Aktivitätsverlust bei der Deletion lässt sich dadurch erklären, dass der YY1-Enhancer nicht mehr binden konnte. Auch könnten noch weitere Proteine im DNA-Bereich bps 1-238 binden, die in der Substitutionsmutanten, im Gegensatz zum

Deletionskonstrukt, einen positiven Effekt auf die transkriptionelle Aktivität ausüben könnten.

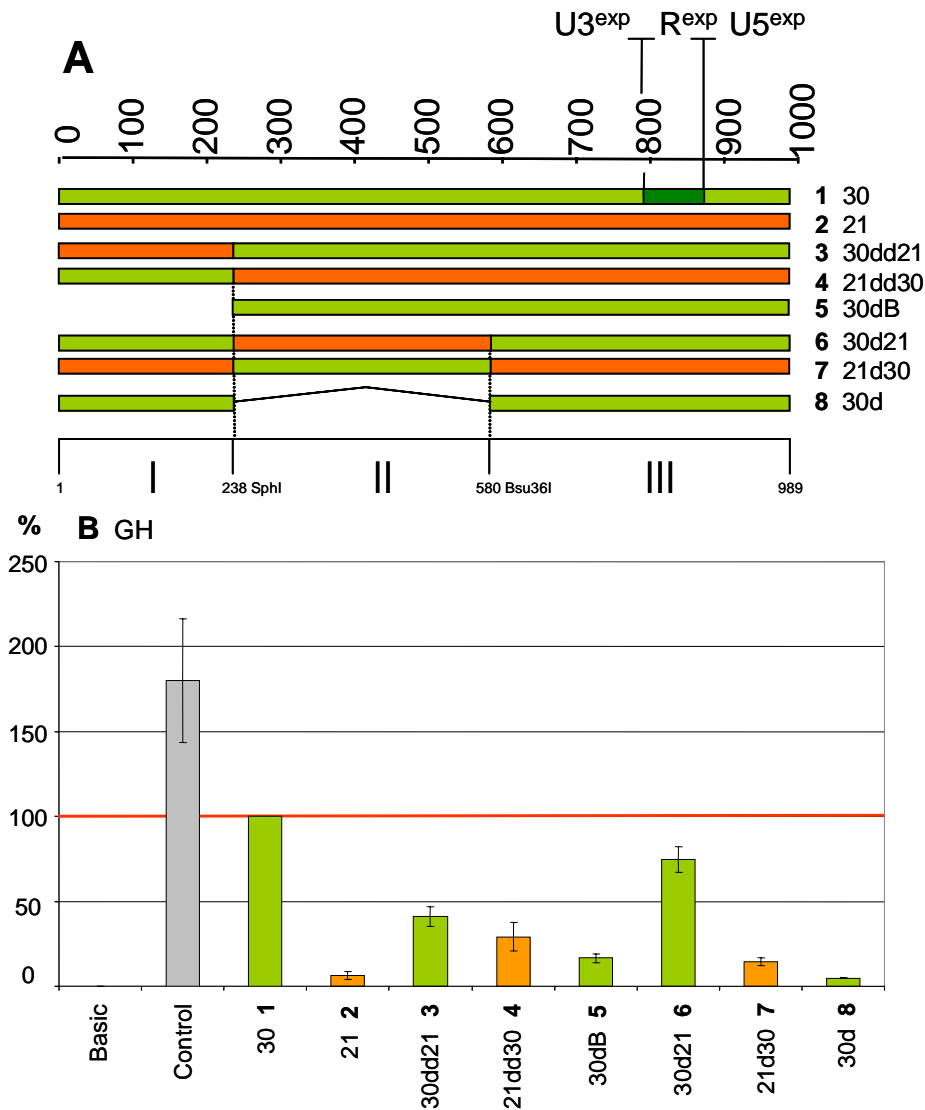


Abbildung 2-14 Relative Luziferase-Aktivität der LTR-Konstrukte. **A** Konstruktübersicht. Zahlen geben die Position im HERV-K LTR in bp an. exp = experimentell ermittelt. **B** Relative Luziferase-Aktivität verschiedener LTR-Konstrukte in GH-Zellen. Die Aktivität in % ist auf die aktive LTR pck30 (Spur 1, 100 %, rote Linie) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle.

Einen ähnlichen Effekt konnte man bei den Konstrukten 30d21 und 30d sehen. Hier ist der Bereich bps 238-580 mutiert worden. Es zeigte sich eine noch größere Diskrepanz zwischen der Chimären und der Deletion (Abb. 2-14 B, Säulen 6 und 8). Die Chimäre hatte noch 75 % Promotoraktivität, während die Deletionsmutante so inaktiv wie die LTR21 war. Die pcK30-LTR unterscheidet sich in diesem Bereich durch fünf Cluster von Punktmutationen von der LTR21. Die Mutationen im Bereich bps 238 bis 580 wurden deshalb detaillierter in ihrer Auswirkung auf die Promotoraktivität untersucht (siehe 2.3.2).

Durch Substitution von inaktiven LTR21-Sequenzen durch pcK30-DNA wurde eine erhöhte Aktivität induziert. Austausch des Bereichs bps 1-238 führte zu einem Anstieg von 6 % auf 29 %. Der Austausch des mittleren Promotorbereichs bps 238-580 führte zu einem Anstieg auf 14 % (Abb. 2-14 B, Säulen 4 und 7). Die Punktmutationen der LTR21 in diesen Bereichen scheinen also zum Verlust von Aktivator-Bindungsstellen zu führen.

2.3.2 Detailanalyse der Punktmutationscluster M0 – M5

Der minimale Aktivitätsverlust des chimären Konstrukts 30d21 (siehe 2.3.1; Abb. 2-14 B, Säule 6) und der hohe Aktivitätsverlust des Deletionskonstrukts 30d (siehe 2.3.1; Abb. 2-14 B, Säule 8) legten die Vermutung nahe, dass die Mutationen der LTR21 zwischen den bps 238-580 nicht nur repressorische Auswirkungen haben könnten. Die Punktmutationen der LTR21 sind in diesem Bereich in fünf Clustern (Abb. 2-15 A, M0/M1-M5) konzentriert. Um Auswirkungen einzelner Cluster auf die Transkription zu überprüfen, wurden die jeweiligen Bereiche der pcK30-LTR durch die entsprechenden Sequenzen der LTR21 ersetzt. Die Promotoraktivität dieser Konstrukte wurde wieder im Luziferase-Assay untersucht. In Abb. 2-15 A ist dargestellt, in welcher Weise die verschiedenen Cluster M0-M5 in der pcK30 LTR ausgetauscht wurden.

Die Aktivität des Konstrukts 30dQ stieg im Assay auf 160 %. Die Mutationen M0 und M1 übten einen positiven Effekt auf die Transkription aus (Abb. 2-15 B, Säule 4). Das M2-M5 umfassende Konstrukt 30dP führte zur Erniedrigung der Aktivität auf 38 % (Abb. 2-15 B, Säule 5). Dieser Unterschied zwischen dem aktivierenden und reprimierenden Potential der Cluster M0-M5 spiegelte sich als additiver Effekt im Konstrukt 30d21 wieder, das zu einer um 25 % verminderten transkriptionellen Aktivität führte (Abb. 2-15 B, Säule 3). Hier liegt also eine Summe verschiedener Effekte vor. In diesem Konstrukt wurden möglicherweise Aktivator / Repressor-Bindungsstellen generiert bzw. mutiert. Das Deletionskonstrukt 30d (siehe 2.3.1; Abb. 2-14 B, Säule 8) verdeutlichte diesen Effekt. Bei Deletion potentieller Aktivator / Repressor-Bindungsstellen fiel die Aktivität auf 5 %.

Mit den Konstrukten 30dHy, 30dR und 30dS konnten die reprimierenden Effekte der LTR21-Sequenzen dem Bereich bps 449-578 zugeordnet werden (Abb. 2-15 B, Säulen 6 bis 8). Konstrukt 30dHy verdeutlichte, dass Cluster M2 einen transkriptionell negativen Effekt ausübt. Die Aktivität fiel auf 33 % (Abb. 2-15 B, Säule 6). Niedrige Werte erhielt man auch bei den Konstrukten 30dR (34 %) und 30dS (21 %) (Abb. 2-15 B, Säulen 7 und 8). 30dS besitzt nur die Punktmutationen des Clusters M5 bei bp 572 und 578.

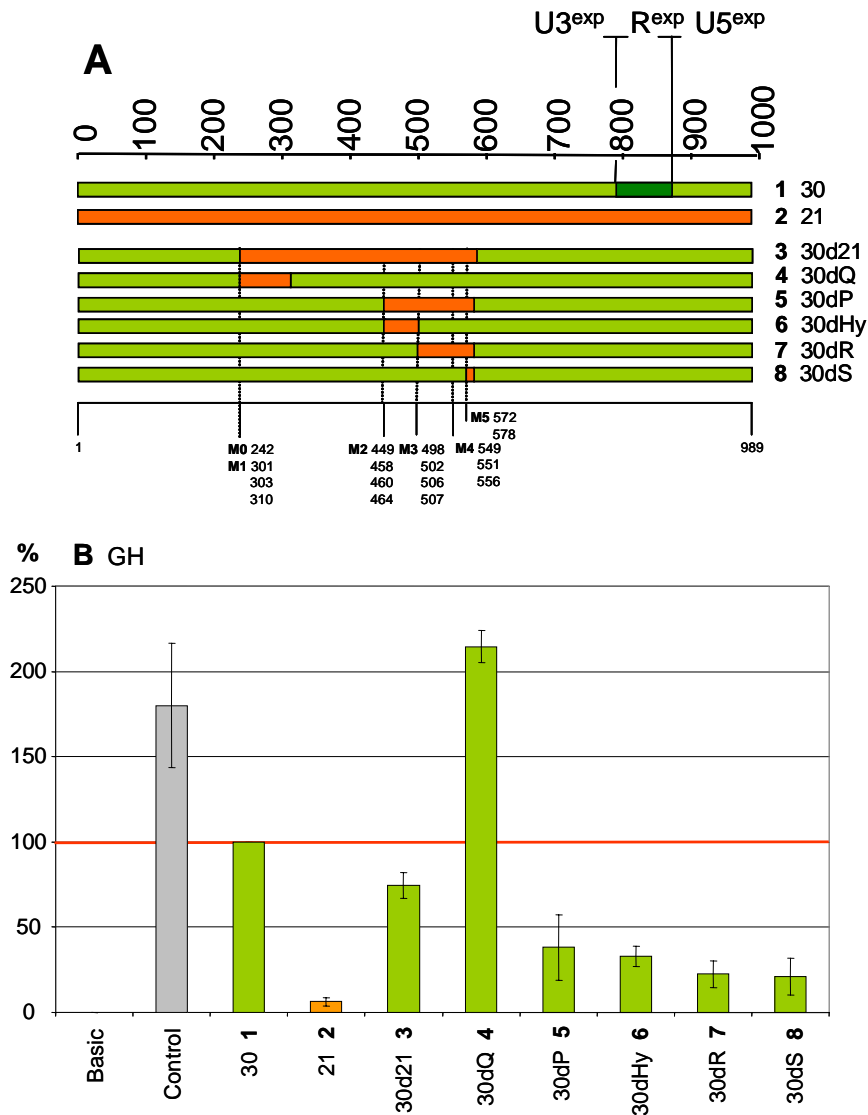


Abbildung 2-15 Relative Luziferase-Aktivität der LTR-Konstrukte. **A** Konstruktübersicht. Zahlen geben die Position im HERV-K LTR in bp an. M0 - M5 = mutagene Cluster. exp = experimentell ermittelt. **B** Relative Luziferase-Aktivität verschiedener LTR-Konstrukte in GH-Zellen. Die Aktivität in % ist auf die aktive LTR pcK30 (Spur 1, 100 %, rote Linie) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle.

Die zusätzlichen Punktmutationen von Cluster M3 und M4, die in Konstrukt 30dR vorhanden sind, führten zu keinem signifikanten Unterschied gegenüber dem Konstrukt 30dS. Da diese zwei Punktmutationen einen transkriptionell sensiblen Bereich zu beeinflussen scheinen, wurde die DNA-Region des Clusters M5 einer Analyse in dem Transkriptionsfaktor-suchprogramm TESS unterzogen. Die Suche ergab mit hoher logarithmischer Wahrscheinlichkeit eine Konsensussequenz für Sp1 bei den bps 576-581. Im EMSA konnte an dieser Stelle mit den verwendeten Sonden noch keine Bindung gezeigt werden (siehe 2.2.6).

Zusammenfassend kann man sagen, dass nicht alle LTR21 Punktmutationen

aktivitätsvermindernd sind. Die Inaktivität der LTR21 beruht vielmehr auf der Summe aller Punktmutationen.

2.3.3 Mutationsanalyse R-U5

Der Bereich stromabwärts vom Transkriptionsstart ist ebenfalls von transkriptionellem Interesse. Dieser Bereich wird von der PolIII als Template verwendet und wird daher regulatorisch weniger benutzt. Am Aufbau des Initiationskomplexes beteiligte Repressoren und Aktivatoren können Bindungsstellen in diesem Bereich aufweisen, werden dann jedoch von der PolIII verdrängt. Ein DNA Bereich, der stromabwärts des Initiationspunktes liegen kann, ist das DPE. Es ist gewöhnlich in einem Bereich 28-33 bps stromabwärts vom Transkriptionsstart lokalisiert und hat eine relativ lockere Konsensussequenz: A/G-G-A/T-C/T-G/A/C. An Position +24 relativ zum Transkriptionsstart ist vorzugsweise ein Guanodin lokalisiert (Burke und Kadonaga 1997). Diese Untersuchungen sind an *Drosophila melanogaster* gemacht worden, doch ist die Konsensussequenz von *Drosophila melanogaster* zu *Homo sapiens* konserviert (Burke und Kadonaga 1997; Zhou und Chiang 2001). Das DPE kann mit einem Inr kooperieren, um eine korrekte Bindung von TFIID zu gewährleisten und um für eine punktgenaue Initiation bei TATA-losen Promotoren zu sorgen (Kutach und Kadonaga 2000). Bei dem HERV-K Promotor liegt 28 bps stromabwärts vom Transkriptionsstart für die aktiven LTRs die Sequenz GGCGG (siehe Abb. 2-1, bps 821-825) vor, es findet sich eine Abweichung zur Konsensussequenz an Position 3 und 4 des DPE. Die inaktiven LTRs besitzen die Sequenz GGTGG. Ein Thymin an Position 3 der Konsensussequenz bedeutet eine geringere Abweichung von der DPE-Konsensussequenz. An Position +24 besitzen aktive wie auch inaktive LTRs das von der Konsensussequenz geforderte Guanodin.

pck30 und LTR21 unterscheiden sich im R-U5-Bereich in elf Punktmutationen. Ob diese Mutationen, vor allem jene, die das DPE betreffen, Auswirkungen auf die transkriptionelle Aktivität haben, wurde mit einer Reihe von chimären Konstrukten getestet (Abb. 2-16 A).

Bei der R-U5-Chimären 30dM ist der LTR-Bereich ab bp 798 durch LTR21-DNA ersetzt. 30dM zeigte einen mäßigen Abfall der Transkriptionsaktivität auf 58 % (Abb. 2-16 B, Säule 3). Konstrukt 30dL, das im Gegensatz zu 30dM die Punktmutationen der bps 809 und 823 nicht besitzt (vergleiche Abb. 2-16 A und 2-1), zeigte einen leichten Abfall der Promotoraktivität auf 82% (Abb. 2-16 B, Säule 5). Konstrukt 30dO besitzt nur die Punktmutation bei bp 846 und rief keinen signifikanten Verlust der Reporteraktivität hervor (Abb. 2-16 B, Säule 7). Die Ursache des Aktivitätsverlusts muss also bei den bps 809, 823

oder den restlichen acht Punktmutationen stromabwärts von Mutation bp 846 liegen. Die Konstrukte 30dG und 30dH (siehe 2.1.1.1, Abb. 2-4 B, Säulen 3 und 4) zeigten bereits, dass die Deletion der Bereiche bps 829-989 eine Steigerung in der Aktivität hervorrief.

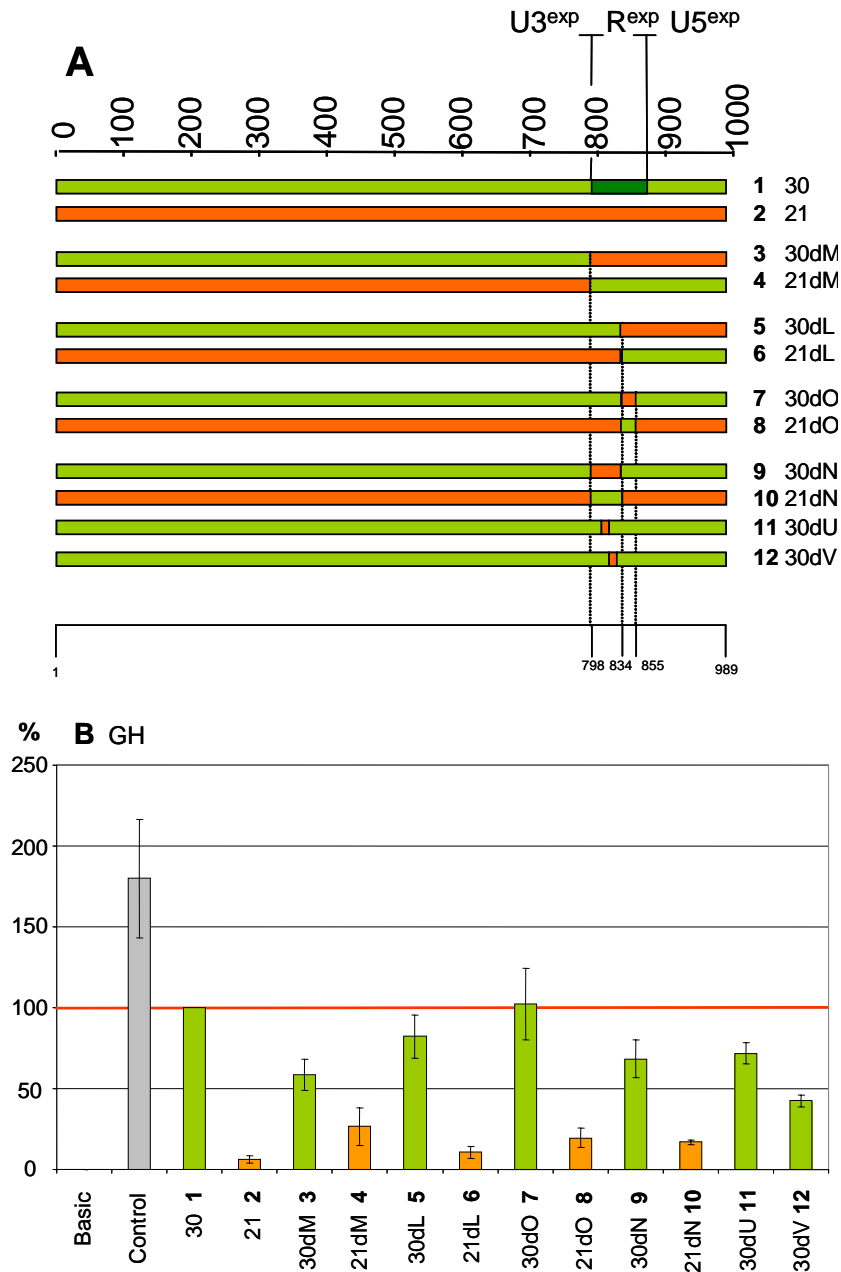


Abbildung 2-16 Relative Luziferase-Aktivität der LTR-Konstrukte. **A** Konstruktübersicht. Zahlen geben die Position im HERV-K LTR in bp an. exp = experimentell ermittelt. **B** Relative Luziferase-Aktivität verschiedener LTR-Konstrukte in GH-Zellen. Die Aktivität in % ist auf die aktive LTR pck30 (Spur 1, 100 %, rote Linie) bezogen. Erklärungen zu den Konstrukten im Text. Basic = Negativkontrolle, SV40 = Positivkontrolle.

Sollten also Transkriptionsfaktoren in diesem Bereich binden, führen die Punktmutationen der LTR21 zu keiner Senkung der Promotoraktivität. Der beobachtete Rückgang der Aktivität von Konstrukt 30dM beschränkt sich also auf die Punktmutationen 809 und 823. Der Austausch beider Nukleotide (30dN) sowie singuläre Substitution des Nukleotids 809 (30dU) induzierte eine Erniedrigung der Aktivität um ca. 30 % (Abb. 2-16 B, Säulen 9 und 11). Einzelaustausch des bp 823 (30dV) hingegen senkte die Reporteraktivität auf 48 % (Abb. 2-16 B, Säule 12). Diese Base liegt im putativen DPE mit der Sequenz GGCGG, die damit zu GGTGG mutiert wurde.

Die chimären Konstrukte, bei denen die entsprechende pcK30-DNA in die LTR21 Sequenz eingeführt wurde, belegen die Wichtigkeit dieser Sequenzbereiche. Konstrukte 21dM und 21dN besitzen ein pcK-DPE. Die Aktivität der inaktiven LTR wurde von 6 % auf 26 % bzw. 17 % angehoben (Abb. 2-16 B, Säulen 4 und 10). Konstrukte 21dL und 21dO zeigten einen Anstieg der Aktivität auf 11 % bzw. 20 % (Abb. 2-16 B, Säulen 6 und 8). Die ausgetauschten Bereiche schienen im pcK30-Kontext (30dL, 30dO) von eher untergeordneter Wichtigkeit zu sein. Es war ein relativ schwacher Rückgang in der Aktivität zu beobachten. Der Anstieg der Aktivität bei Konstrukt 21dO auf 20 % legt jedoch den Schluss nahe, dass dieser Bereich doch regulatorische Bedeutung haben könnte. Dieser Sachverhalt müsste in einer EMSA-Analyse näher untersucht werden.

Dies sind erste Hinweise auf ein funktionales DPE im HERV-K Promotor. Die genauen Grenzen des DPE müssten durch weitere Mutationsanalysen festgelegt werden.

2.4 Transkription durch Testis-spezifische Faktoren

Die Analyse des HERV-K Promotors zeigte Bereiche auf, die für die Transkription von besonderer Bedeutung sind. Der Aktivitätsunterschied zwischen aktiven und inaktiven LTRs ist durch Punktmutationen in Konsensussequenzen von Transkriptionsfaktoren zu erklären. Das Ausmaß der Transkription eines bestimmten Gens ist sehr von den exprimierten Transkriptionsfaktoren in einem bestimmten Gewebe abhängig. Es ist nicht auszuschließen, dass die Hochregulation eines Testis-spezifischen Transkriptionsfaktors in Teratokarzinomen zur Expression von HERV-K führt. Der Vergleich mit anderen Karzinomzelllinien zeigte keine oder nur eine basale Expression von HERV-K. Viruspartikel sind bis heute nur in Teratokarzinomzellen nachgewiesen worden (Löwer et al. 1984). Da die Regulation über Sp1 und Sp3 nicht die alleinige Expression im Testis erklärt, ist es wahrscheinlich, dass ein weiterer exklusiv im Testis vorkommender Transkriptionsfaktor für die Expression von HERV-K mit verantwortlich ist.

2.4.1 SRY

SRY ist ein HMG-Protein (high mobility group-Protein), das vermehrt im Testis exprimiert wird. Es besitzt eine DNA-Bindungsdomäne (HMG-Box), jedoch keine Aktivierungsdomäne. Die DNA-Bindungsdomäne hat die Fähigkeit, DNA in einem stärkeren Winkel zu biegen als durch die Sequenz vorgegeben ist (Tang und Nilsson 1998). SRY wird auf dem Y-Chromosom codiert und bestimmt das männliche Geschlecht in Säugetieren (Goodfellow und Lovell-Badge 1993).

Initiale Untersuchungen zeigten, dass SRY als mRNA (RT-PCR) sowie als Protein (Western Blot Analyse) in GH-Zellen nachzuweisen ist. HeLa-Zellen, die aus einem Cervixkarzinom kultiviert wurden, zeigten kein Signal für das Y-Chromosom-codierte SRY (Daten nicht gezeigt). Um den Effekt von SRY auf die Transkription von HERV-K zu untersuchen, wurde ein SRY-Expressionsplasmid eingesetzt. Die Kotransfektion des SRY-Expressionsplasmids mit dem Reporterkonstrukt pcK30 führte im Luziferase-Assay zu einer Steigerung der transkriptionellen Aktivität der pcK30-LTR. In GH-Zellen konnte eine Steigerung der pcK30-Aktivität auf 174 % detektiert werden (Abb 2-17 A, Säulen 1 und 2). Für die LTR21 zeigte sich kein signifikanter Anstieg der Aktivität (Abb 2-17 A, Säulen 3 und 4).

Interessanterweise zeigte sich in HeLa-Zellen, dass die Aktivität der aktiven LTR durch SRY gesteigert werden konnte. Ohne SRY schwankte die gemessene Aktivität der aktiven und inaktiven LTRs im Vergleich zur Negativkontrolle um einen basalen Wert. Es zeigte sich kein Unterschied zwischen aktiven und inaktiven LTRs (siehe 2.1.1.1, Abb. 2-4 C, Säulen Basic, 1 und 2). Kotransfektion von SRY jedoch induzierte einen deutlichen Aktivitätsunterschied zwischen pcK30 und LTR21. Die aktive pcK30 wird in ihrer Fähigkeit, den Reporter zu exprimieren, verstärkt. Die Aktivität stieg auf 365 %. Die inaktive LTR21 hingegen ließ sich nicht signifikant in der Aktivität steigern. Es stellte sich der in GH-Zellen beobachtete Aktivitätsunterschied zwischen aktiver und inaktiver LTR ein (Abb. 2-17 B, Säulen 2 und 4). In GH-Zellen ist die pcK30-LTR um den Faktor 16 aktiver als die LTR21. In HeLa-Zellen zeigte sich der Unterschied in einem Faktor von 3,5. Das bedeutet, dass SRY nicht ausreichend ist, um den Promotor in HeLa-Zellen vollständig zu aktivieren.

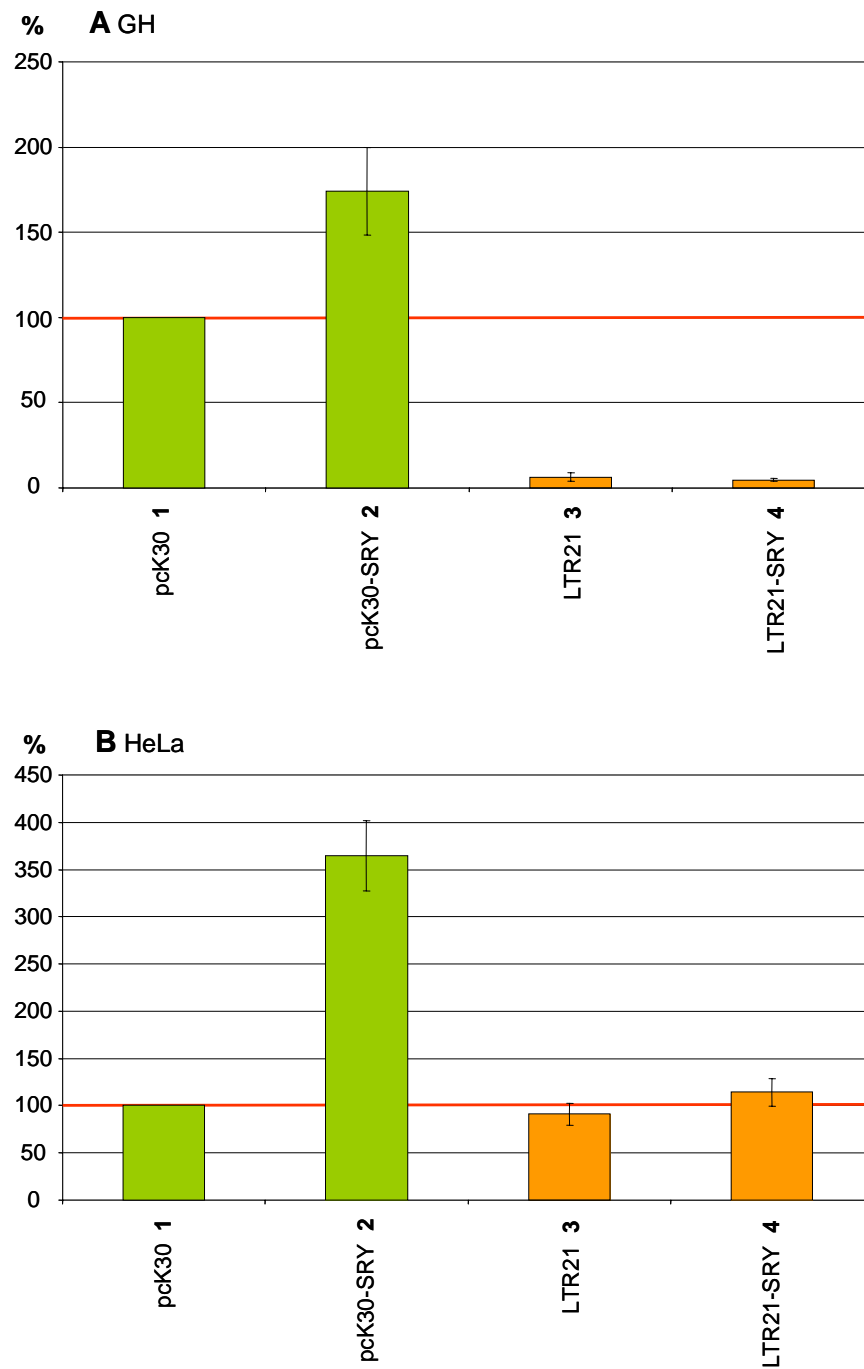


Abbildung 2-17 Relative Luziferase-Aktivität der LTRs pck30 und LTR21. Die Aktivität in % ist auf die aktive LTR pck30 (Spur 1, 100 %, rote Linie) bezogen. **A** Werte für GH-Zellen **B** Werte für HeLa-Zellen. **1** pck30; **2** pck30 cotransfiziert mit einem SRY-Expressionsplasmid; **3** LTR21; **4** LTR21 cotransfiziert mit einem SRY-Expressionsplasmid.

Ungeklärt bleibt auch, ob SRY in *cis* auf der LTR selbst wirkt, oder ob in *trans* ein Gen aktiviert wird, dessen Genprodukt zur gesteigerten Transkription von HERV-K führt.

LTR Position	SRY
bps 515-588	-
bps 579-652	-
bps 647-754	-
bps 745-798	-
bps 745-833	-

Tabelle 2: Zusammenfassung der EMSA-Analyse mit rekombinantem SRY-Protein.

In einem ersten Versuch, Bindungsstellen für SRY auf dem HERV-K Promotor zu identifizieren, sind die bereits beschriebenen Sonden (siehe 2.2.6) mit rekombinantem SRY im EMSA untersucht worden (Tabelle 2). Es konnte für alle getesteten Sonden keine Bindung von rekombinantem SRY gezeigt werden. Auch hier müssten die übrigen DNA-Bereiche des Promotors noch getestet werden.

Dies sind erste Hinweise auf eine Beteiligung des Testis-spezifischen Transkriptionsfaktors SRY an der Transkription von HERV-K.

2.5 Epigenetische Mechanismen

Neben der direkten Beeinflussung der Transkription eines Promotors durch Transkriptionsfaktoren mit aktivierendem oder repressorischem Potential gibt es immer mehr Hinweise auf eine Transkriptionskontrolle durch epigenetische Mechanismen. Eine zentrale Rolle wird der DNA-Methylierung und dem Chromatin zugeordnet. Die Hypermethylierung von CpG-Inseln innerhalb eines Promotors führt zusammen mit Histon-Deacetylierung zur transkriptionellen Inaktivierung durch Kondensierung der DNA (Heterochromatin). Transkriptionell aktive Gene befinden sich im Euchromatin. Diese DNA-Bereiche sind hypomethyliert und die Hyperacetylierung der Histone führt zum Entwinden des Heterochromatins (Rice und Allis 2001; Strahl und Allis 2000). Auf molekularer Ebene können Gene durch DNA-Methyltransferasen (Robertson et al. 2000), Methyl-CpG Bindungsproteine (Jones et al. 1998), Histon-modifizierende Enzyme, „chromatin remodeling factors“, Transkriptionsfaktoren und chromosomale Proteine aktiviert bzw. inaktiviert werden.

In einem ersten Versuch, den Einfluss der Methylierung auf die HERV-K Transkription zu analysieren, wurde das Reporterkonstrukt der LTR pcK30 mit der CpG-Methylase M. Sss I behandelt. Es wurden so alle CpG-Inseln auf dem Plasmid methyliert. Die Methylierung der CpG-Inseln innerhalb der ORFs jedoch hat keinen Effekt auf die transkribierende PolIII. Man kann so den Effekt der Methylierung der LTR auf die Transkription im Reporter-Assay messen.

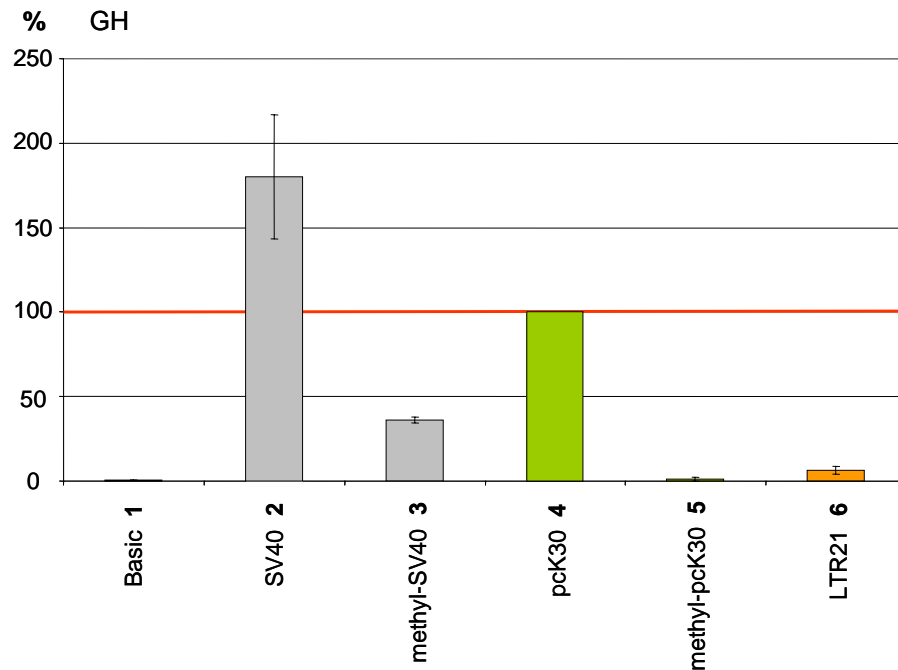


Abbildung 2-18 Relative Luziferase-Aktivität der LTRs pcK30 und LTR21. Die Aktivität in % ist auf die aktive LTR pcK30 (Spur 3, 100 %, rote Linie) bezogen. Die Werte wurden in GH-Zellen ermittelt. **1** Basic; **2** SV40; **3** methylierter SV40-Vektor; **4** pcK30; **5** methylierter pcK30-Vektor; **6** LTR21.

Die Promotoraktivität des aktiven pcK30 fiel durch die Methylierung von 100 % auf 1 % (Abb. 2-18, Säulen 4 und 5). Das gleiche gilt für die Positivkontrolle, einem SV40-Promotor gesteuerten Luziferasekonstrukt. Im Vergleich zu pcK30 fiel die Aktivität von 180 % auf 36 % (Abb. 2-18, Säulen 2 und 3). Epigenetische Faktoren könnten also eine Rolle bei der Regulierung der HERV-K Expression spielen. Auf welche Weise die Methylierung den Rückgang der Promotoraktivität hervorrief, ist noch nicht untersucht worden.

3 Diskussion

3.1 Die HERV-K LTR

3.1.1 Allgemein

Das humane endogene Retrovirus HERV-K hat sich im menschlichen Genom mit ca. 50 Volllängenkopien manifestiert. Die inserierten Virusgenome resultieren aus Keimbahninfektionen exogener Retroviren vor ca. 3×10^6 Jahren und den daraus abgeleiteten Retrotranspositionseignissen. Als endogenes Retrovirus kann die genetische Information vertikal verbreitet werden. Exogene Retroviren hingegen sind bei ihrer Verbreitung auf horizontale Infektionswege angewiesen. Als integraler Bestandteil des Wirtsgenoms ist das Provirus, genauso wie zelluläre DNA, Mutationsereignissen unterworfen. Durch die evolutionäre Mutationsrate bei mangelnder Selektion haben ERVs im Allgemeinen die Fähigkeit zur Expression verloren. Zu viele akkumulierte Punktmutationen haben Promotor und ORFs zerstört. HERV-K ist als einziges Provirus noch in der Lage, alle viralen Gene zu exprimieren. Es besitzt intakte ORFs für die für Retroviren typischen Proteine Gag, Prt, Pol, Env und einen zusätzlichen ORF für das akzessorische Protein Rec (siehe 1.1.4). Es gibt drei dominante mRNA-Spezies, aus denen die Virusproteine gebildet werden: eine 8,9 kB, eine 3,5 kB und eine 1,8 kB große mRNA. Eine basale Transkription der mRNA ist zwar in vielen Geweben beobachtet worden, doch Viruspartikel sind bis jetzt nur in Teratokarzinomzellen nachgewiesen worden. Die exklusive Expression von Viruspartikeln in männlichen Teratokarzinomen macht eine Analyse des HERV-K Promotors sinnvoll.

In dieser Arbeit wurde die Regulation der Transkription von HERV-K untersucht. Die U3-, R- und U5-Regionen eines HERV-K LTRs wurden experimentell bestimmt und der Core und der regulatorische Promotor wurden neu definiert.

3.1.2 Aktiv oder Inaktiv?

Die Analyse der HERV-K LTR-Sequenzen zeigte, dass die LTRs über die Zeit verschiedene Punktmutationen angesammelt haben. Es stellte sich heraus, dass bestimmte Mutationen sich in charakteristischen Clustern angehäuft haben (siehe 2.3.2, 2-15 A, z.B. M2). Diese Cluster lassen sich bei verschiedenen LTRs wiederfinden. Den inaktiven LTRs konnten die Cluster von Punktmutationen zugeordnet werden. Es ist anzunehmen, dass die LTRs Inaktivität durch Anhäufung von Punktmutationen erworben haben. Der Vergleich der LTR-Gruppierung im phylogenetischen Stammbaum mit den Promotoraktivitäten im Luziferase-Assay zeigte, dass

die Akkumulation von Punktmutationen zu transkriptioneller Inaktivität führt. Die untersuchten LTRs pcK30 und LTR10 gruppieren sich im phylogenetischen Baum im gleichen Zweig wie die LTRs K 113 und K 115. Die inaktiven LTRs, ausgelöst durch die Cluster von Punktmutationen, finden sich in einer anderen Gruppe (siehe 2.1, Abb. 2-2). Eine hohe Anzahl von Punktmutationen spricht für eine längere Verweildauer im menschlichen Genom. Man nimmt an, dass endogene Retroviren Mutationen mit einer ungefähren Rate von $2,3 \times 10^{-9}$ bis 5×10^{-9} pro Stelle und Jahr anhäufen. Das entspricht ungefähr einer Mutation alle 200.000-450.000 Jahre. Bei Integration liegt das Provirus mit zwei identischen LTRs im Genom vor. Die Punktmutationen, die die 5'-LTR von der 3'-LTR unterscheiden, sind ein Maß für das Alter eines spezifischen Provirus. Zusätzlich kann man in einer PCR-Analyse die Verteilung von HERV-K Loci beim Menschen, Schimpansen, Gorilla und Orang-Utan untersuchen, um Integrationsereignisse anhand des Stammbaums der menschlichen Entwicklung festzusetzen. Aus diesen Daten kann man das Alter eines Provirus berechnen und es in den Kontext der menschlichen Entwicklung stellen, zum anderen kann so eine Präzisierung des evolutionären Stammbaums anhand molekularbiologischer Daten vorgenommen werden (Johnson und Coffin 1999).

Es konnte so gezeigt werden, dass HERV-K die Keimbahn der Primaten vor ca. 35×10^6 Jahren das erste Mal infiziert hatte, also nach der Trennung der Entwicklungsstammbäume der Platyrrhini (Neuweltaffen) und Catarrhini (Altweltaffen). Das Provirus einer inaktiven LTR, wie der LTR18, konnte in Menschen, Schimpansen und Gorillas nachgewiesen werden, nicht aber in Orang-Utans. Die Integration fand also nach Abspaltung der Entwicklungslinie des Orang-Utans statt, aber bevor sich die Entwicklung von Mensch, Schimpanse und Gorilla vor ca. 5×10^6 Jahren trennte (Johnson und Coffin 1999). Mit 36 Unterschieden zwischen 5'- und 3'-LTR konnte der LTR ein Alter von $5,7 \times 10^6$ Jahren zugeordnet werden. Eine aktive LTR, wie die LTR10, stammt von einem Provirus, der nur in Menschen nachgewiesen werden konnte. Diese Proviren sind also jünger als 5×10^6 Jahren, dem Zeitpunkt an dem die Entwicklung von Mensch und Schimpanse auseinander ging. Bei Sequenzierung konnte man zwei Punktmutationen zwischen 5'- und 3'-LTR feststellen. Das Alter dieser LTR wird also auf ca. 1×10^6 Jahre geschätzt (Barbulescu et al. 1999).

Die im Jahr 2001 veröffentlichten Sequenzen der HERV-K Proviren K 113 und K 115 (Turner et al. 2001) zeigten auf, dass die evolutionär jüngste LTR des Provirus HERV-K113 eine vergleichbare Aktivität besitzt wie die evolutionär älteren LTRs pcK30 und LTR10 (siehe 2.1, Abb. 2-2, Abb. 2-3). Da sich die 5'- und 3'-LTRs von HERV-K 113 nicht durch Punktmutationen unterscheiden, schätzte man das Alter des Provirus K 113 auf maximal

200.000-450.000 Jahren. Das Alter von HERV-K 115 konnte nicht genau bestimmt werden. Zwar konnte man vierzehn Unterschiede zwischen beiden LTRs ausmachen, doch neun dieser Mutationen wurden schon in LTRs mehrerer HERV-K Proviren beobachtet (Barbulescu et al. 1999). Diese neun Mutationen könnten deshalb durch Rekombination oder durch Gen-Konversion mit einem anderen HERV-K Locus entstanden sein. Die übrigen fünf Mutationen geben deshalb das Maximalalter des Provirus mit 1×10^6 - $2,25 \times 10^6$ Jahren vor, da auch hier die Möglichkeit der Substitution durch Gen-Konversion besteht.

Die Sequenzierung des Genoms des Provirus K 113 zeigte, dass intakte ORFs für jedes Virusgen existieren. Das Provirus hebt sich jedoch von anderen ab, da es keinen Aminosäureaustausch in einer konservierten Domäne der RT aufweist. Proviren, wie HERV-K(HLM-2.HOM) (Mayer et al. 1999) oder HERV-K(C7) (Tönjes et al. 1999), besitzen in dieser Domäne anstelle des Standard YIDD-Motivs ein CIDD-Motiv. Die Autoren (Turner et al. 2001) vermuteten, dass HERV-K in der Lage war, die Keimbahn des Menschen noch in evolutionär jüngster Zeit zu infizieren und dass das Provirus K 113 Menschen in der heutigen Zeit reinfizieren könnte. HERV-K 113 konnte in 29 %, HERV-K 115 in 16 % aller getesteten Personen nachgewiesen werden. Dies deckt sich mit Untersuchungen unserer Arbeitsgruppe (R. Löwer, persönliche Mitteilung).

Es stellt sich daher die Frage, ob HERV-K Loci im menschlichen Genom noch aktiv sind, und wenn ja, welche Konsequenzen ergeben sich daraus für den Organismus? Geht man vom evolutionären Selektionsdruck aus, sollte eigentlich ein Provirus, das keinen essentiellen Vorteil für den Wirt besitzt, über die Jahrtausende viele Mutationen akkumulieren und damit inaktiv werden. Tatsächlich sind die ORFs von fast allen ERVs durch Mutationen oder Deletionen zerstört worden (O'Connell et al. 1984; Repaske et al. 1983) und eine Expression aller viraler Proteine konnte nur für HERV-K nachgewiesen werden (Löwer et al. 1996). Ob die Endogenisierung von Retroviren einen Nutzen für den Wirt hat, war und ist Ausgangspunkt zahlreicher Forschungsvorhaben. Die Vor- und Nachteile, die ein Provirus auf seinen Wirt haben kann, sind in verschiedenen Veröffentlichungen diskutiert worden:

Bis heute konnte noch keine exogene Variante der endogenen Retroviren identifiziert werden. Alle untersuchten Viren verloren nach Integration durch Anhäufung von Mutationen ihre Infektiosität. Auch wenn nicht alle Virusproteine exprimiert werden können, eine Bedingung für Infektiosität, könnte die teilweise Expression von Virusproteinen eine Infektion durch ihre exogenen Verwandten verhindern. Das Provirus würde so einen positiven Effekt für den Wirt erzeugen, indem es sich zu einem schützenden zellulären Gen entwickelt (Weiss 1993).

Auch Enhancer-Effekte in zellulären Genen, ausgelöst durch retrovirale Sequenzen, konnten beobachtet werden. Für HERV-E ist bekannt, dass die Insertion mehrerer proviraler Sequenzen zu einer Erweiterung der Pankreas-spezifischen Expression des Amylasegens in der Speicheldrüse führte (Samuelson et al. 1990).

Insertion durch Retrotransposition oder Infektion von Retroviren kann aber natürlich auch schwere Folgen für den Organismus haben. Die Insertion in das Genom des Wirtes kann zelluläre Gene stark beeinflussen. Neben den schon erwähnten Promotor- und Enhancer-Effekten auf die Expression von Genen, können die ORFs der betroffenen Gene verändert bzw. zerstört werden (Hung 1994). Dies kann essentielle Gene, Proto-Oncogene oder Tumor-Supressor-Gene betreffen. Alle erwähnten Veränderungen können zur Tumorbildung führen.

Die Infektion von Keimbahnzellen ermöglicht es dem Virus, sich vertikal zu verbreiten. Die Fähigkeit des Virus zu replizieren, beschränkt sich also nicht auf die horizontale infektiöse Verbreitung, sondern die Erbinformation des Virus wird vom Wirt zum Nachkommen weitergegeben. Zur Verbreitung eines Virus über die Keimbahn bei gleichzeitigem Erhalt der Infektiösität ist die Erhaltung der Expression in Keimbahnzellen notwendig. Nur so kann es zu Neuinfektionen kommen und damit zur genetischen Erneuerung des proviralen Gen-Pools. Die Expression von HERV-K scheint aber heutzutage auf Keimbahntumorzellen beschränkt zu sein. Das könnte für Veränderungen im Transkriptionsfaktormuster in Keimbahnzellen seit der letzten Neuinfektion von HERV-K sprechen, sodass eine Expression ausgeschlossen ist. Eine andere Möglichkeit wäre eine zeitlich begrenzte Expression von für HERV-K essentiellen Transkriptionsfaktoren, z.B. während der Embryogenese, sodass in adulten Keimbahnzellen keine Expression von HERV-K mehr möglich wäre (siehe 3.3.1).

Die Analyse des HERV-K Promotors kann dazu beitragen, die essentiellen Faktoren der Expression zu determinieren. Diese Arbeit leistet einen Beitrag zum Verständnis der Mechanismen der HERV-K Expression in Keimbahntumoren.

3.2 Die Transkriptionsregulation von HERV-K

3.2.1 TATA-Box versus Inr

Die Analyse eines Promotors beginnt mit der Identifikation des minimalen Promotors. Die dafür in Frage kommenden Mittel sind der Luziferase-Assay, mit dem Promotor-relevante Regionen eingegrenzt werden können, und eine 5'-RACE, die den tatsächlichen Transkriptionsstart eines Gens bestimmen kann. Die schrittweise Deletion eines Promotors

vom 5'- bzw. 3'- Ende kann Hinweise auf den minimalen Promotor geben, doch muss man beachten, dass Deletionen den Gesamtkontext des Promotors verändern. Die Deletion von Aktivator-Bindungsstellen kann zu starkem Abfall der Aktivität führen, obwohl der Core Promotor noch intakt ist. Eine Übereinstimmung der Ergebnisse der Deletionsmutanten, der Vollängenmutationen und der 5'-RACE erhöhen die Zuverlässigkeit der Hypothese. Anfängliche Untersuchungen zeigten bei einer Deletion der ersten 460 bps eines aktiven LTRs einen sehr starken Abfall der Aktivität auf 20 %. Eine weitere Deletion von 120 bps reduzierte die Aktivität auf 1 %. In diesen 120 bps war die putative TATA-Box enthalten, sodass man die bps 460-988 vorsichtig als Core Promotor definiert hatte (Hasenmaier 1995). Ergänzend wurden in der vorliegenden Arbeit 3'-Deletionsmutanten angefertigt. Es wurde schnell klar, dass schon eine 3'-Deletion von 188 bps zu einem drastischen Verlust der Reporteraktivität führte. Geht man von einem Transkriptionsinitiationskomplex aus, der auf der TATA-Box bei bp 532 basiert, müsste eine größere Restaktivität bei Konstrukt 30dI zu erwarten sein (siehe 2.1.1.1, Abb. 2-4 B, Säule 5). Das Ausbleiben einer Aktivitätsreduktion bei einer Substitution der TATA-Box durch Thymidine brachte Gewissheit über einen TATA-Box unabhängigen Regulationsweg in GH-Zellen (siehe 2.1.1.2, Abb. 2-5). Es ist jedoch nicht auszuschließen, dass in anderen Geweben zu bestimmten Entwicklungszuständen eine alternative TATA-vermittelte Transkription stattfinden könnte. Die Expression von HERV-K in malignen Melanomzellen (Muster et al. 2003, siehe 1.1.4) wirft in diesem Zusammenhang die Frage nach einem alternativen Promotor auf. Für MMTV wurde ein alternativer Promotor beschrieben, der in der U3 Region lokalisiert ist. Dieser Promotor wird nur in B-Zellen benutzt (Günzburg et al. 1993).

3.2.2 Die Rekrutierung von TFIID

Die Regulation von TATA-losen Promotoren kann z.B. über Sp1 und Sp3 erfolgen. Sp-Proteine können stromaufwärts vom Transkriptionsstart binden und ermöglichen wahrscheinlich durch aktive Interaktion mit Untereinheiten von TFIID die Initiation (Courey et al. 1989; Hagen et al. 1994; Ihn und Trojanowska 1997; Udvadia et al. 1995). Sp1 trägt also dazu bei, dass TFIID korrekt an die Initiationsstelle geführt wird. Bei TATA-Box-haltigen Promotoren dirigiert das 30 bp vom Initiationspunkt bindende TBP alle TAFs sowie die PolIII in eine Position, die eine exakte Initiation ermöglicht (Matsui et al. 1980; Parker und Topol 1984). Für Promotoren ohne TATA-Box ist diskutiert worden, ob eine TATA-Box äquivalente Stelle 30 bps stromaufwärts vom Inr für die Bindung von TBP verantwortlich sein könnte (Zenzie-Gregory et al. 1993). Die Untersuchungen ergaben, dass

eine stärkere Inr-Aktivität zu verzeichnen war, wenn die -30-Region sehr A/T-reich war. Untersuchungen mit TFIID, das eine mutierte TBP-Untereinheit enthielt, zeigten jedoch, dass Inr-Promotoren über TBP-Bindung initiiert werden können, dass dies aber auch nicht für jeden Promotor zutrifft (Martinez et al. 1995). Das bedeutet, dass die Rekrutierung von TFIID und damit der initiale Schritt beim Aufbau des Transkriptionsinitiationskomplexes über verschiedene Strategien vermittelt werden kann.

Das Binden von TBP an der TATA-Box bewirkt das Ausbilden einer Sattelstruktur, die zur teilweisen Entwindung und Biegung der DNA kurz vor dem Transkriptionsstart führt und somit das Binden der PolIII ermöglicht. Auch bei der Inr-vermittelten Transkription kann TBP das Binden von TFIID ermöglichen (Smale 1997). Die korrekte Positionierung kann über TAFs wie TAFII135 und CIF150 erfolgen (Kaufmann et al. 1998). Doch ist die Anwesenheit von TBP nicht zwingend. Das DPE ist ein weiteres Sequenzelement, das im Zusammenhang mit TATA-losen Promotoren gefunden wurde. Zusammen mit einem Inr kann das DPE den Multiproteinkomplex TFIID zum Transkriptionsstart dirigieren. Dabei hängt die korrekte Initiation sehr stark vom Abstand des DPEs vom Inr ab. In Untersuchungen, in denen der Abstand des DPEs zum Inr verkürzt bzw. verlängert wurde, konnten starke Verluste in der transkriptionellen Aktivität beobachtet werden (Kutach und Kadonaga 2000).

Die vorliegenden Ergebnisse erfordern eine neue Definition des Core Promotors von HERV-K. Durch die 5'-RACE wurde der Transkriptionsstart bei bp 794 gefunden. Das deckt sich mit dem starken Verlust der Aktivität der Deletionskonstrukte 30dI, 39dK, 30dA (siehe 2.1.1.1, Abb. 2-4 B, Säule 5 bis 7). Ausgehend vom Transkriptionsstart findet man direkt stromaufwärts zwei Sp1-Bindungsstellen (TESS) und stromabwärts eine DPE-Konsensussequenz (siehe 2.1.1.3, Abb. 2-7, 2.3.3).

Zwar muss die Funktionalität des DPEs noch mit gezielten Mutationen überprüft werden, doch zeigten pcK30/LTR21-Basenpaarsubstitutionen bereits einen transkriptionell sensitiven Bereich im putativen DPE auf (siehe 2.3.3, Abb. 2-16 B, Säule 12). Ergänzend könnte die Empfindlichkeit des DPEs gegenüber Abstandsveränderungen zwischen DPE und Inr, wie oben erwähnt (Kutach und Kadonaga 2000), durch Basenpaarinsertionen bei HERV-K überprüft werden.

Ein DNaseI-Footprint mit rekombinanten TFIID könnte weitere Hinweise darauf geben, ob der TFIID-Komplex das HERV-K DPE utiliziert. Untersuchungen von Burke et al. zeigten bereits, dass bei einem von TFIID rekrutierten DPE sich der Proteinkomplex bis zum bp +35 relativ zum Transkriptionsstart ausdehnt und der Bereich bei bp -30 im Gegensatz zu TATA-Box-haltigen Promotoren nicht zur Bindung benutzt wird (Burke und Kadonaga 1996).

Die Bindung von TFIID auf dem HERV-K LTR scheint synergistisch über das Inr, das DPE und über Interaktionen mit Sp-Proteinen vermittelt zu werden. Der Core Promotor fängt nach dieser Theorie stromaufwärts von der ersten Sp-Bindungsstelle an und endet stromabwärts vom DPE. In diesem DNA-Stück wären demnach alle für die Initiation notwendigen Elemente enthalten. Der Core Promotor umfasst also ungefähr die bps 740 – 830.

3.2.3 Regulation durch Sp1/Sp3

Die Entdeckung des Transkriptionsfaktors Sp1 als regulatorisches Element des SV40-Promotors begründete eine immer noch wachsende Familie von über zwanzig Zink-Finger-Proteinen, zu denen Sp1 - Sp4 und die Krüppel-ähnlichen Proteine gehören. Ihnen allen sind eine Zink-Finger-DNA-Bindungsdomäne und Aktivierungs- bzw. Repressor-Domänen gemein (Abb. 3-1). Die hohe Konservierung der DNA-Bindungsdomäne führt dazu, dass Mitglieder dieser Proteinfamilie GC/GT-Boxen mit variierender Affinität binden können. Die Proteine weisen jedoch sehr unterschiedliche transkriptionelle Eigenschaften auf und tragen so zur kombinatorischen Genregulation bei. Die Erkenntnisse über die multiplen Effekte dieser Proteinfamilie deuten auf eine sehr komplexe und differenzierte Kontrolle von Sp-Protein gelenkten Genen hin (Bouwman und Philipsen 2002; Suske 1999).

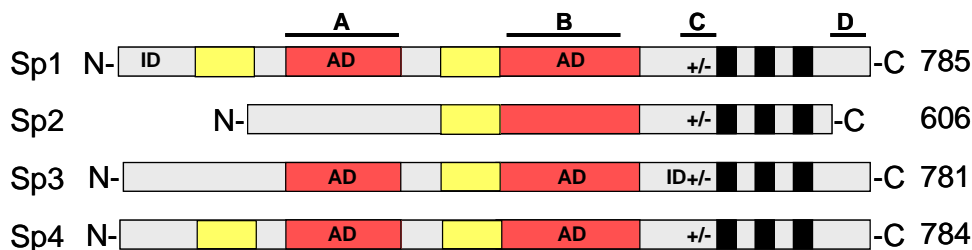


Abbildung 3-1: Struktur von Sp-Proteinen. Schema der vier Proteine Sp1 – Sp4 aus dem Menschen. Zahlen geben Positionen in Aminosäuren an. Die Längen von Sp1 und Sp3 beziehen sich auf die Vollängenproteine nach Jackson et al. (Jackson et al. 1990) und Hagen et al. (Hagen et al. 1992). Farbige Kästchen zeigen Regionen an, die reich an Glutamin- (rot) und Serin / Threonin- (gelb) Aminosäureresten sind. (+/-) Region, reich an geladenen Aminosäuren. Schwarze Kästchen geben die Zink-Finger-Domänen an. AD Aktivierungsdomäne. ID Inhibitorische Domäne. A, B, C und D geben Regionen an, die zu den transkriptionellen Eigenschaften beitragen (Courey und Tjian 1988). (Abbildung verändert nach Suske 1999).

Die EMSA-Analyse der stromaufwärts vom Inr gelegenen GC/GT-Boxen zeigte, dass Sp1 und Sp3 an diese binden können (siehe 2.2.2, Abb. 2-10 und 2.2.3, Abb. 2-11). Sp2 und Sp4 zeigten im Supershift-Assay mit nukleären Extrakten aus GH- bzw. HeLa-Zellen keine Änderung im Migrationsverhalten der Protein-DNA-Komplexe. Es kann deshalb ausgeschlossen werden, dass diese Proteine einen Einfluss auf die basale Transkription von HERV-K haben. Die Überprüfung des Bindungsverhaltens der anderen Familienmitglieder

gestaltet sich schwierig. Fehlende kommerziell hergestellte Antikörper bzw. rekombinante Proteine ermöglichen keinen schnellen Zugang zur Beantwortung der Frage, welche Sp/Krüppel-ähnlichen Proteine an der Regulation beteiligt sein könnten. Die Biotinylierung der Sp-Bindungsstellen-haltigen Sonde und anschließende Inkubation mit nukleären Zellextrakten und Isolierung der gebundenen Proteine über eine Streptavidin-haltige Matrix würde es ermöglichen, die relevanten GC/GT-Box-bindenden Proteine zu identifizieren. Ein Vergleich zwischen GH- und HeLa-Zellen könnte weitere Hinweise auf regulatorische Mechanismen des HERV-K Promotors beantworten. Dieser Sachverhalt müsste jedoch in weiterführenden Arbeiten untersucht werden.

Es ist aber anzunehmen, dass das gezeigte starke Bindungsverhalten von Sp1 und Sp3 im EMSA (siehe 2.2.1, Abb. 2-9) auch ausschlaggebend für die *in-vivo* Situation ist. Da im EMSA unter Verwendung nukleärer Extrakte keine weiteren prominenten Banden detektiert werden konnten, kann man annehmen, dass andere schwach bindende Transkriptionsfaktoren kaum einen Einfluss auf die Transkription nehmen. Die Region zwischen der ersten und zweiten Sp-Bindungsstelle zeigte im EMSA einen Proteinkomplex unbekannter Zusammensetzung. Dieser Komplex konnte mit dem Subfragment 30-4 kompetitiert werden, nicht aber die Protein/DNA-Komplexe, die als Sp1 und Sp3 identifiziert worden sind (siehe 2.2.2, Abb. 2-10, Spur 9). Der Proteinkomplex kommt also als ein weiteres regulatorisches Mitglied der Sp-Familie nicht in Frage, da die Bindung stromabwärts der ersten GC/GT-Box stattzufinden scheint. Für den Glucocorticoid Rezeptor und c-Ets-1 sind zwar Konsensussequenzen in diesem Bereich gefunden worden, doch die Bindung ist noch nicht bewiesen worden. Auch liegen in der Literatur keine Berichte über die Beteiligung dieser zwei Faktoren am Aufbau von Transkriptionsinitiationskomplexen vor. Eine ausschließliche Bindung von Sp-Proteinen in diesem Bereich *in-vivo* ist wahrscheinlich.

Die Mutation und Deletion der GC/GT-Boxen und die folgende Analyse im Luziferase-Assay zeigte, dass auch *in-vivo* diese Bindungsstellen von Bedeutung sind. Die Reporteraktivität fiel bei Mutationen der ersten GC/GT-Box um mehr als 70 %. Die zweite GC/GT-Box führte zur Aktivitätsreduktion um ca. 50 %. Die Restaktivitäten sprechen für eine verminderte Bindungskapazität der Sp-Proteine. Das unvollständige Kompetitionsvermögen der Sonde 21-1 (siehe 2.2.1, Abb. 2-8, Spur 6) belegte diesen Befund. Die Punktmutationen der LTR21 schränken also die Bindung der Sp-Proteine ein, führen aber nicht zum kompletten Verlust der Bindung. Die transkriptionelle Inaktivität der LTR21 erklärt sich durch weitere Punktmutationen im Enhancer und im regulatorischen Promotor (siehe 3.2.4). Der stärkere Rückgang der Promotoraktivität bei Mutation der ersten GC/GT-Box deckt sich mit der

in-vitro Beobachtung, dass im EMSA Subfragmente, die die erste GC/GT-Box enthalten, besser kompetitieren können (siehe 2.2.2, Abb. 2-10, vergleiche Spur 8 und 10). Darüberhinaus könnte bei Besetzung der zweiten GC/GT-Box eine sterische Hinderung mit TFIID und anderen Komponenten des Initiationskomplexes auftreten. Diese Frage könnte mit dem schon besprochenen DNaseI-Footprint für TFIID beantwortet werden.

Sp1 wurde als erstes kloniertes Mitglied der Sp-Proteinfamilie lange eine ausschließliche Rolle in der Regulation von GC/GT-Box gesteuerten Genen zugeordnet. Erst die Klonierung weiterer Mitglieder ließ die Komplexität dieser Transkriptionsfaktoren erahnen (Hagen et al. 1992; Kingsley und Winoto 1992). Sp1–Sp4 besitzen eine konservierte Zink-Finger-Domäne, die zur DNA-Bindung benutzt wird. In Abbildung 3-1 sind die Proteine schematisch dargestellt. Die hohe Konservierung der Zink-Finger-Domäne von Sp1 und Sp3 (90 %) führt zu sehr ähnlichen Bindungsaffinitäten für Konsensus-GC/GT-Boxen. Sp1, Sp3 und Sp4 weisen zwei glutaminreiche Transaktivierungsdomänen A und B auf, die für die transkriptionelle Aktivität wichtig sind. In der Nähe dieser Domänen finden sich Serin/Threonin-reiche Sequenzen, die Ziel posttranslationaler Modifikationen sind (Abb. 3-1). Sp2 besitzt nur eine glutaminreiche Domäne, hat aber mit den anderen Faktoren einen geladenen Sequenzbereich aminoterminal von der Zink-Finger-Domäne gemein. Die Bindungsaffinität von Sp2 zu GC/GT-Boxen unterscheidet sich von den anderen Faktoren, da ein kritisches Histidin im ersten Zink-Finger durch Leucin substituiert ist (Kingsley und Winoto 1992).

Es konnte sehr bald nach Entdeckung weiterer Proteine aus der Sp-Proteinfamilie gezeigt werden, dass Sp1 nicht alleine an der Regulation der Transkription über GC/GT-Boxen beteiligt ist. Das Zusammenspiel von Sp1 und Sp3 als regulatorische Grundeinheit ist bereits in zahlreichen Publikationen behandelt worden (Galvagni et al. 2001; Hagen et al. 1994; Kaczynski et al. 2003; Sjøttem et al. 1996). Auch zeigte sich, dass retrovirale Elemente bevorzugt eine Sp-Protein-vermittelte Transkription zur Expression nutzen. Dies konnte für ERV9 (Strazzullo et al. 1994), HIV (Harrich et al. 1989) und HERV-H (Sjøttem et al. 1996) gezeigt werden. Auch HERV-H nutzt einen Regulationsweg, der ohne TATA-Box auskommt, aber abhängig von Sp1/Sp3 und einem Inr ist.

Initiale Untersuchungen zeigten, dass Sp3, im Gegensatz zu Sp1, repressorisches Potential besitzen kann (Hagen et al. 1994). Es wurde jedoch schnell klar, dass Sp3 nicht nur die klassischen Bedingungen eines Repressors erfüllt, sondern auch aktivierendes Potential bei z.B. dem humanen alpha2(I) Kollagen Promotor ausüben konnte (Ihn und Trojanowska

1997). Verschiedene Theorien sind zu diesem Sachverhalt aufgestellt worden, die folgend erläutert werden.

Eine genauere Betrachtung der Sequenzmotive von Sp1 und Sp3 zeigte bei beiden Proteinen zwei konservierte aminoterminal glutaminreiche Regionen und eine carboxyterminale Zink-Finger-Domäne (Abb. 3-1). Bei Sp1 ist direkt am Aminoterminus eine inhibitorische Domäne lokalisiert, die durch das Binden eines stark konservierten nukleären Proteins p74 das Aktivierungspotential des Transkriptionsfaktors erniedrigen kann (Murata et al. 1994). Für Sp3 ist am Übergang von den glutaminreichen Domänen zur Zink-Finger-Domäne ein inhibitorisches Aminosäuresequenzmotiv IKEE identifiziert worden, das durch posttranslationale Modifikation die repressorische Funktion von Sp3 aktivieren könnte. Mutation des Aminosäuretriplets KEE führte zum Verlust der repressorischen Eigenschaften und zu einem verstärkten Transaktivierungspotential in Promotoren, die mehrere GC/GT-Boxen besitzen (Dennig et al. 1996). Diese inhibitorische Domäne könnte das fehlende Aktivierungspotential von Sp3 erklären, da auch die Aktivierungsdomänen von Sp1 und Sp3 hoch konserviert sind. Diesem Motiv konnten verschiedene Modifikationen zugeordnet werden. Zum einen konnte gezeigt werden, dass Sp3 *in-vivo* acetyliert sein kann (Braun et al. 2001), zum anderen konnte der inhibitorischen Domäne kurz darauf eine weitere Modifikation zugeordnet werden. Die kovalente Bindung des Proteins SUMO-1 an das IKEE-Motiv wurde von Sapetschnig et al. analysiert (Sapetschnig et al. 2002).

SUMO-1 ist ein kleines Protein, das 18 % Homologie zu Ubiquitin besitzt. Zusammen mit SUMO-2 und SUMO-3 formt SUMO-1 eine kleine Proteinfamilie (Melchior 2000; Yeh et al. 2000). Alle drei Proteine können kovalent an die ϵ -Aminogruppe von Lysinresten gebunden werden. Als Substrate für die kovalente Modifikation durch SUMO-1 konnte man z.B. RanGAP, p53 und Steroidhormonrezeptoren identifizieren. Die Auswirkungen der SUMO-1-Modifikation auf die zellulären Prozesse stellten sich jedoch sehr unterschiedlich dar. RanGAP konnte durch die SUMO-1-Modifikation mit RanBP2 (Ran-binding protein 2) dimerisieren und führte so zur Akkumulation am nukleären Porenkomplex (Mahajan et al. 1997). Bei p53 konnte die kovalente Bindung von SUMO-1 die transkriptionelle Aktivität erhöhen (Rodriguez et al. 1999), im Falle der Steroidrezeptoren führte die Bindung von SUMO-1 zu zurückgehender transkriptioneller Aktivität (Poukka et al. 2000). Für HSF1 und HSF2 zeigte sich, dass die DNA-Bindungsaktivität durch SUMO-1 erhöht werden kann (Goodson et al. 2001; Hong et al. 2001). Für Sp3 konnte gezeigt werden, dass die SUMOylierung die Repressor-Funktion aktiviert (Sapetschnig et al. 2002). Mutationen, die eine SUMOylierung verhinderten, zeigten auch in Transfektionsexperimenten mit

Sp3-Expressionsplasmiden keine repressorische Wirkung mehr. Inwieweit die Acetylierung und SUMOylierung ein Mechanismus zur Aktivierung und Inaktivierung von Sp3 darstellt, wird zurzeit noch untersucht.

Die Entdeckung von verschiedenen Sp3-Isoformen eröffnete eine weitere Möglichkeit der Regulation. Nach dieser Hypothese wurde den verschiedenen Sp3-Isoformen die Aktivator bzw. Repressor-Funktion zugeordnet. Es existieren mindestens drei Isoformen, die sich in der Western Blot Analyse als 120 kDa Proteinduplet und als 60-70 kDa Proteinduplet zeigten (Kennett et al. 1997). Bei den kleineren Isoformen wird die Translation über interne Startcodons zwischen den zwei glutaminreichen Domänen initiiert. Das bedeutet, dass ein großer Teil der Transaktivierungsdomäne verloren geht. Kennett et al. konnten zeigen, dass die intern initiierten Isoformen repressorische Wirkung aufweisen und dass das Vollängenprotein als Transaktivator fungieren kann (Kennett et al. 1997). Dies steht im Kontrast zu Beobachtungen von Braun et al. und Sapetschnig et al., die die inhibitorische Domäne dem IKEE Motiv zuordneten (Braun et al. 2001; Sapetschnig et al. 2002). Dieses Motiv bleibt von den aminoterminal trunkierten Isoformen unberührt. Inwieweit ein Zusammenspiel der verschiedenen beschriebenen Modifikationen mit den verschiedenen Isoformen stattfindet, und inwieweit sich dies auf die Sp1-gelenkte Transkriptionsinitiation auswirkt, konnte noch nicht gezeigt werden.

Ein weiterer Erklärungsansatz zur transkriptionellen Wirkung von Sp1 und Sp3 nimmt das molekulare Verhältnis der Proteine im Nukleus mit in Betracht. Da Sp1 und Sp3 ähnliche Bindungsaffinitäten aufweisen, könnte eine hohe Sp1-Konzentration das notwendige Aktivierungspotential für die Initiation der Transkription liefern. Eine höhere Sp3-Konzentration hingegen könnte zur Repression der Transkription führen. Ein Beispiel hierfür findet sich bei der aktiven Transkription des Endothel-spezifischen *KDR/flk-1* Promotors. In Endothelzellen ist das Sp1/Sp3 Verhältnis sehr viel höher als in Nicht-Endothelzellen (Hata et al. 1998).

Betrachtet man die intranukleäre Verteilung von Sp3 in GH- und HeLa-Zellen, fällt auf, dass alle detektierten Isoformen sehr viel stärker in HeLa-Extrakten vorhanden sind. Die kleineren Isoformen sind in GH-Extrakten nur sehr schwach zu detektieren. Sollte sich die repressorische Wirkung von Sp3 tatsächlich auf die kleinen Isoformen beziehen, würde das die ausbleibende Aktivität der HERV-K LTRs in HeLa-Zellen zumindest teilweise erklären.

Das Vollängenprotein ist in HeLa-Zellen sehr viel stärker vorhanden. Das bedeutet, dass in GH-Zellen ein höheres Verhältnis von Sp1 zu Sp3 vorliegt als in HeLa-Zellen (siehe 2.2.4,

Abb. 2-12). Auch hier könnte die Anwesenheit von hohen Sp3-Konzentrationen in HeLa-Zellen zur Repression führen.

In der Western Blot Analyse von Sp3 in GH- und HeLa-Zellextrakten zeigten sich mit einem SUMO-1-Antikörper keine übereinstimmenden Signale für Sp3 und SUMO-1. Die SUMOylierung der Sp3-Isoformen konnte nicht gezeigt werden (Daten nicht gezeigt). Die von Sapetschnig et al. beobachtete repressorische Wirkung der SUMO-1-Modifikation scheint für HERV-K nicht relevant zu sein. Fehlerquellen, wie das Abspalten von SUMO-1 durch Hydrolasen während der Extraktpräparation, sind jedoch beschrieben worden (Sapetschnig et al. 2002).

3.2.4 Regulatorischer Promotor und Enhancer

Im regulatorischen Promotor liegen proximale Elemente, die sich aktiv an der Transkriptionsregulation beteiligen. Nach Definition liegen diese Elemente meistens in einem Bereich, der sich 50 - 400 bps stromaufwärts vom Transkriptionsstart befindet. Ein Enhancerelement kann sich distal vom Core Promotor befinden, dies kann sogar mehrere tausend bps vom Core Promotor entfernt sein. Bei Retroviren liegt der regulatorische Promotor in U3, das Enhancerelement ist unabhängig von der LTR lokalisiert.

3.2.4.1 U3

Bei HERV-K befindet sich der stark aktivierende Transkriptionsfaktor YY1 am 5'-Ende des LTRs (Knöbl et al. 1999). Nach der Definition liegt dieser Aktivator außerhalb des als regulatorischen Promotor bezeichneten DNA-Bereichs und wird deshalb als Enhancer bezeichnet.

Wie die Mutationsanalyse des LTRs pcK30 zeigte, befinden sich stromaufwärts vom regulatorischen Promotor Bereiche, die bei Deletion oder Mutation die transkriptionelle Aktivität beeinflussen. Die Konstrukte 30dd21 und 30dB zeigten, dass sowohl die Deletion der bps 1-238 als auch die Substitution durch LTR21-DNA zum starken Verlust der transkriptionellen Aktivität führten (siehe 2.3.1, Abb. 2-14 B, Säulen 3, 5). Neben dem bereits analysierten Enhancerprotein YY1 könnten hier weitere Transkriptionsfaktoren von Wichtigkeit sein. Erste EMSA-Ergebnisse früherer Dissertationen zeigten einen größeren YY1-haltigen Enhancer-Komplex, der die bps 1-150 benutzt. Die YY1-Bindung wurde mittels DNaseI-Footprinting auf den DNA-Bereich der bps 62-83 eingegrenzt. Die Funktion der anderen Proteine ist noch nicht untersucht worden.

Die Mutation des Bereichs, der als regulatorischer Promotor bezeichnet wird, warf einige Widersprüche auf. Die Substitution der bps 238-580 durch LTR21-DNA zeigte einen geringen Verlust der Aktivität von 25 %. Die Deletion hingegen erniedrigte die Aktivität um 95 %. Erst die sukzessive Mutation der Cluster M0 – M5 zeigte, dass auf dem DNA-Bereich (bps 238-580) aktivierende wie auch repressorische Faktoren binden können (siehe 2.3.2, Abb. 2-15). Repressorisch wirkende Elemente binden wahrscheinlich in einem Bereich, der durch die mutierten Cluster M0 und M1 beeinflusst wird. In diesem Bereich erhöhte die DNA des inaktiven LTR21 die transkriptionelle Aktivität. Denkbar wäre natürlich auch das Generieren einer neuen Aktivator-Bindungsstelle, die sich im pK30-Kontext positiv auf die Transkription auswirkt. Die Bindungsstelle besitzt anscheinend auf dem LTR21 keine Relevanz. Der gesamte Promotor ist durch zu viele Mutationen stillgelegt worden.

Sehr interessant sind die Veränderungen der Promotoraktivität bei Mutationen der Cluster M2 und M5. Beide bewirken einen starken Abfall der Aktivität in GH-Zellen. Ein besonderes Augenmerk liegt dabei bei Cluster M5. Die TESS-Datenbankanalyse ergab den Verlust einer GC/GT-Box durch die eingeführte LTR21-DNA. Zwar ergaben EMSA-Studien mit der Sonde bps 515-588 keine aktive Bindung von Sp1 oder Sp3 (siehe 2.2.6, Tab. 1), doch muss man bedenken, dass die Bindungsstelle, die die bps 576-581 abdeckt, sehr dicht am 3'-Ende der Sonde liegt. Es ist bereits beobachtet worden, dass für eine starke Proteinbindung einige bps stromauf- bzw. abwärts der Bindungsstelle wichtig sind (siehe 2.2.2). Eine 3'-erweiterte Sonde sollte Klarheit zu dieser Frage bringen.

Sollte sich die Sp1-Bindung bewahrheiten, könnte eine synergistische Wirkung zwischen Sp-Proteinen am Transkriptionsstart und im regulatorischen Promotor stattfinden. *In-vitro* Experimente zeigten bereits, dass Sp1-Tetramere über distale Bindungsstellen synergistisch transaktivieren können. Die DNA bildet dabei einen Loop, um die Bindungsstellen in räumliche Nähe zu bringen (Li et al. 1991; Mastrangelo et al. 1991; Pascal und Tjian 1991). Die Inkubation von LTR-weiten Sonden mit rekombinatem Sp1- bzw. Sp3-Protein sollte eine Antwort auf diese Frage geben. Es bleibt zu untersuchen, ob distale Elemente außerhalb des LTRs regulatorische Bedeutung haben könnten.

3.2.4.2 R

Die R-Region wird durch den Transkriptionsstart und die Poly(A)-Stelle begrenzt. Die Core Promotor-Elemente Inr und DPE liegen beide in der R-Region. Bei Transkription ist sie somit größtenteils durch die basale Transkriptionsmaschinerie belegt. Der Bereich zwischen DPE und Poly(A)-Stelle könnte jedoch von regulatorischen Transkriptionsfaktoren benutzt werden,

die Nähe zur Bindungsstelle des Multiproteinkomplexes der PolIII macht dies jedoch unwahrscheinlich. Das Konstrukt 30dO, das eine Punktmutation bei bp 846 trägt, bestätigt dies zu einem gewissen Grad. Hier ist kein Abfall in der Aktivität zu detektieren (siehe 2.3.3, Abb. 2-16).

3.2.4.3 U5

Die Analyse der U5-Region erschloss einen weiteren Bereich, der transkriptionell von Bedeutung sein könnte. Die schon erwähnten 3'-Deletionen 30dG und 30dH (siehe 2.1.1.1, Abb. 2-4 B, Säulen 3, 4) zeigten einen Anstieg der transkriptionellen Aktivität auf 140 % bei Deletion der letzten 60 bps des LTRs. Untersuchungen an einem aktiven HERV-K sLTR von Domansky et al. (Domansky et al. 2000) zeigten ein ähnliches Ergebnis. Hier wurden 70 bps deletiert und es wurde ebenfalls ein Anstieg der transkriptionellen Aktivität gemessen. Es scheint also am 3'-Ende des Promotors ein regulierender Faktor binden zu können, der sich negativ auf die Transkription auswirkt. Dies scheint unabhängig von aktiven oder inaktiven LTRs zu sein. Bei dem chimären Konstrukt 30dL, indem der zuvor deletierte DNA-Abschnitt durch LTR21-DNA ersetzt wurde, konnte kein Anstieg in der Aktivität festgestellt werden (siehe 2.3.3, Abb. 2-16 B, Spur 6). Ähnliche repressorische Elemente in der U5 Region wurden bereits für das humane Foamy Virus (HFV) (Yang et al. 1997) und HTLV (Okumura et al. 1996; Seiki et al. 1990) beschrieben. Interessanterweise konnten Okumura et al. (Okumura et al. 1996) der negativ regulierenden Region das Binden von Sp1 und Sp3 zuordnen. Diese Bindung geschieht über eine GC/GT-Box, die das DNA-Motiv CACCC enthält. Für HERV-K finden sich am 3'-Ende von U5 drei GC/GT-Boxen (TESS), die alle das gleiche DNA-Motiv CACCC enthalten (siehe Anhang, 7.2). Auch hier könnten synergistische Effekte zwischen Sp1 und Sp3 und den anderen stromaufwärts gelegenen Bindungsstellen stattfinden. Die funktionale Relevanz dieser Bindungsstellen müsste in weiteren Experimenten belegt werden.

3.3 Die Testis-spezifische Expression

Die Kernfrage nach den Gründen, die eine Expression von HERV-K in männlichen Teratokarzinomzellen bewirken, stellt sich auf verschiedenen Ebenen:

Stellt die Transkriptionskontrolle durch Sp1 und Sp3 den Schlüssel zur Gewebespezifität dar?

Ist ein Testis-spezifischer Faktor, der exklusiv in männlichen Teratokarzinomen

überexprimiert wird, verantwortlich für die Aktivierung? Inwiefern spielen diese Elemente zusammen und was unterscheidet nicht-maligne Testiszellen von den Teratokarzinomen?

Die bereits besprochenen Ergebnisse deuten auf eine Multikomponenten Transkriptionskontrolle hin. Die Analyse stützte sich auf die klassischen Methoden wie Reporter-Assay und EMSA. Diese Methoden können gewiss die Basis der transkriptionellen Mechanismen aufdecken, es bleibt jedoch abzuwarten, inwieweit distale Elemente außerhalb des LTRs, gewebespezifische Expressionsunterschiede der essentiellen Transkriptionsfaktoren und chromosomaler Aufbau der HERV-K Loci im Genom Einfluss auf die Transkription nehmen.

3.3.1 Differentielle Sp1-Genexpression

Viele Gene werden gewebespezifisch angeschaltet, indem die Transkription durch einen gewebespezifischen Transkriptionsfaktor limitiert wird. Sp1 erfüllt diese Bedingungen nur bedingt. Sp1 ist ein ubiquitär exprimierter Transkriptionsfaktor, der in den Nuklei aller Säugerzellen gefunden wird (Briggs et al. 1986). Doch konnte gezeigt werden, dass die Transkription von Sp-gelenkten Genen von der Verfügbarkeit dieses Faktors abhängt. Untersuchungen am SV40-Promotor zeigten, dass die artifizielle Erhöhung von Sp1 *in-vitro* (Dyanan et al. 1985; Dyanan und Tjian 1983a), wie auch in *Drosophila*-Zellen (Courey und Tjian 1988), zu einer erhöhten Promotoraktivität führten. Darüber hinaus zeigte sich, dass durch SV40 eine Erhöhung des endogenen Sp1-Gehalts herbeigeführt werden konnte und dadurch eine erhöhte Transkription GC/GT-Box gesteuerter Gene in kultivierten Zellen beobachtet wurde (Saffer et al. 1990a). Das bedeutet, dass die Expression der Gene, die von ubiquitären Faktoren wie Sp1 abhängen, durch die Menge an verfügbaren Sp1 reguliert wird. Zusätzlich konnte demonstriert werden, dass die Sp1-Bindungsstellen von verschiedenen Promotoren verschiedene Bindungsaffinitäten aufweisen, die sich in einem Faktor von 10-20 unterscheiden (Kadonaga et al. 1986). Verschiedene Gene mit verschiedenen Affinitäten für Sp1 sind also abhängig von dem Sp1-Gehalt in dem spezifischen Gewebe, indem das Gen exprimiert werden soll. Das bedeutet, dass z.B. bei einem geringen Sp1-Gehalt in einem bestimmten Gewebe nur Gene mit starker Bindungsaffinität aktiviert werden können und umgekehrt Gene mit schwachen Bindungsaffinitäten sehr sensitiv gegenüber Schwankungen im Sp1-Gehalt sein können.

Untersuchungen an der Maus belegten, dass der Gehalt an Sp1-mRNA und übereinstimmend der Gehalt an Sp1-Protein verschiedener Gewebe stark entwicklungsabhängig ist. Saffer et al. (Saffer et al. 1991) konnten einen 100-fachen Unterschied im Sp1-Gehalt in verschiedenen

Gewebe zu verschiedenen Entwicklungszuständen feststellen. Dies begründet sich in der differentiellen Genkontrolle des Sp1 Single-Locus-Gens auf Chromosom 15 (Saffer et al. 1990b). Die Autoren konnten zeigen, dass Sp1 nicht nur ein basaler Aktivator für eine Reihe von Haushaltsgenen ist, sondern dass der Gehalt von Sp1 maßgeblich an der regulatorischen Kontrolle von exprimierten Genen während Entwicklung und Differenzierung beteiligt ist. Interessanterweise zeigte sich ein sehr hoher Sp1-Gehalt im Testis und speziell in den Spermatiden. Fortschreitende Differenzierung der Zellen führte zu niedrigem Sp1-Gehalt, ein Hinweis dafür, dass in ausdifferenzierten Geweben nur ein minimaler Sp1-Gehalt für die basale Expression von Haushaltsgenen notwendig ist.

Kann es also sein, dass ein Provirus eine transkriptionelle Regulation evolutionär bevorzugt, die während der Entwicklung des Wirts gewisse Zeitfenster eröffnet, die zur Expression und damit zur Verbreitung benutzt werden könnten? Sp1 ist ein Transkriptionsfaktor, der genau diese Voraussetzungen erfüllt. Angewiesen auf Verbreitung über Keimzellen könnte das Virus die zeitlich begrenzten Anstiege des Sp1-Gehalts in Spermatozyten und embryonalen Zellen zur Expression nutzen. Interessanterweise korreliert die Aktivität eines HERV-K LTR-gelenkten Reportergens in der Maus mit dem Expressionsmuster von Sp1 in differenziertem und undifferenziertem Gewebe. Casau et al. (Casau et al. 1999) zeigten, dass die transkriptionelle Aktivität eines transgenen aktiven HERV-K LTRs in undifferenzierten Spermatozyten und im Testis am höchsten ist, dort wo sich auch eine hohe Sp1-Konzentration findet (Saffer et al. 1991). Die Analysen von Saffer et al. deuten an, dass die Expression von Sp1 mit zunehmender Differenzierung der Zellen nachlässt.

Die Western Blot Analyse von Sp1 und Sp3 zeigte bereits einen gewebespezifischen Unterschied zwischen GH- und HeLa-Zellen. Das Mengenverhältnis von Sp1 zu Sp3 war in GH-Zellen größer als in HeLa-Zellen und Sp3 war mit allen Isoformen stärker in HeLa-Zellen vertreten. Da die Frage noch nicht klar beantwortet werden kann, welche Sp3-Modifikationen oder welche Sp3-Isoformen an repressorischen Wirkungen auf die Transkription beteiligt sind, wäre es sehr lohnenswert, die Sp1/Sp3-Verteilung in verschiedenen Geweben zu untersuchen. Am Mausmodell, das bereits erfolgreich für eine HERV-K LTR Analyse herangezogen wurde (Casau et al. 1999), könnte man diesen Sachverhalt an differenzierten und undifferenzierten Geweben überprüfen.

3.3.2 SRY

Die oben diskutierten Befunde über die differentielle Sp1-Genexpression im Zusammenhang mit der Entwicklung eines Organismus und die Testis-spezifische Expression von HERV-K

legen die Vermutung nahe, dass Transkriptionsfaktoren, die an der Gonadogenese des Mannes beteiligt sind, auch bei der Transkription von HERV-K von Bedeutung sein könnten. Zu diesen Proteinen zählen unter anderem Sox-Proteine, Sf1, Dax1 und SRY, deren Expressionsfenster während der Vertebraten Gonadogenese überlappen (O'Neill und O'Neill 1999). SRY ist hier von besonderem Interesse, da es auf dem Y-Chromosom codiert ist und damit exklusiv bei männlichen Vertebraten zu finden ist. Dem Protein wird die Funktion zugeordnet, die Entwicklung eines Organismus zum männlichen Geschlecht zu lenken.

SRY (Accession number: Q05066) hat hohe Homologien zu den HMG-Box-Proteinen, die architektonische Aufgaben bei der Organisation der Chromosomen übernehmen. Es wird von einem Intron-losen Gen von 3,8 kB Größe codiert. Das Protein ist 204 Aminosäuren groß und weist ein Molekulargewicht von 24 kDa auf (Su und Lau 1993).

Die HMG-Box ist in der Lage, DNA bei Bindung in einem größeren Maße zu biegen, als durch die Sequenz vorgegeben ist. Der größte Teil der SRY-Aminosäuresequenz besteht aus einer HMG-Box Domäne, so dass dem Protein eine hauptsächliche Rolle bei der Verwindung der DNA zugesprochen wird (Ukiyama et al. 2001). Darüber hinaus gibt es Beobachtungen, dass die carboxyterminalen sieben Aminosäuren mit der PDZ-Domäne von Sip-1 interagieren können. Dies könnte als Brücke zu weiteren Transkriptionsfaktoren genutzt werden (Poulat et al. 1997). Die Funktion von SRY scheint also größtenteils darin zu bestehen, bei spezifischen Genen essentielle Transkriptionsfaktoren in räumliche Nähe zu bringen. Ob dies über Proteinbrücken oder allein durch DNA Biegung geschieht, muss noch im Detail geklärt werden. Eindrucksvoll belegt wurde die Wichtigkeit dieses Proteins bei der Gonadogenese von Berta et al. und Jager et al. (Berta et al. 1990; Jager et al. 1990). Sie konnten zeigen, dass geschlechtsumgekehrte XY-chromosomale Frauen Mutationen im *SRY*-Gen aufweisen. Im Tierversuch entwickelten sich XX-Mäuse zu Männchen, wenn sie ein *SRY*-Transgen besaßen (Koopman et al. 1991).

Sehr interessant waren in diesem Zusammenhang die Ergebnisse, die mit dem *SRY*-Expressionsplasmid erhalten wurden. In GH-Zellen konnte die Überexpression von *SRY* die Aktivität von pcK30 im Luziferase-Assay auf 174 % steigern. Die LTR21 blieb weiterhin inaktiv (siehe 2.4.1, Abb. 2-17 A). *SRY* scheint also für aktive LTRs transkriptionell wirksam zu sein. Umso erstaunlicher waren die Ergebnisse, die bei HeLa-Zellen erhalten wurden. Diese Zellen, entstanden aus einem Cervixkarzinom, können *SRY per definitionem* nicht exprimieren. Die Überexpression von *SRY* in HeLa-Zellen und die anschließende Analyse der transkriptionellen Aktivität der LTR-Reporterkonstrukte im Luziferase-Assay ergaben einen *SRY*-abhängigen Anstieg der Reporteraktivität der aktiven

LTR pcK30 auf 365 %. Die Aktivität der LTR21 stieg nicht signifikant an (siehe 2.4.1, Abb. 2-17 B). SRY könnte also ein für die Transkription von HERV-K essentieller Transkriptionsfaktor sein, denn es wurde in GH- wie auch in HeLa-Zellen ein Anstieg der transkriptionellen Aktivität der aktiven LTR bewirkt. Die inaktiven LTRs konnten scheinbar durch einen fehlenden aktiven Transkriptionsinitiationskomplex nicht in ihrer Aktivität gesteigert werden.

Darüberhinaus scheint die Transkription von weiteren Faktoren, die *cis* oder *trans* wirken könnten, abhängig zu sein. In GH-Zellen wurde die aktive LTR bei Überexpression 16x stärker transkribiert als die inaktive, in HeLa-Zellen war eine 3,5-fach verstärkte Transkription zu messen (vergleiche 2.4.1, Abb. 2-17 A und B, Säulen 2 und 4). Der Unterschied in der Transkriptionsaktivität zwischen der aktiven und der inaktiven LTR war in HeLa-Zellen im Vergleich zu GH-Zellen weniger groß. Die Ergebnisse weisen jedoch klar auf eine wichtige Rolle von SRY bei der Transkription von HERV-K hin.

Der Befund relativiert zu einem gewissen Grad die Hypothese über die negativ regulierende Funktion von Sp3 in HeLa-Zellen. Es könnte auch sein, dass der essentielle Faktor SRY in HeLa-Zellen nicht verfügbar ist, und deswegen eine gerichtete Transkription der LTRs nicht möglich war. Das Zusammenspiel dieser Faktoren müsste in weiteren Experimenten untersucht werden.

Die bereits zur Analyse der Transkriptionsinitiation verwendeten EMSA-Sonden wurden daraufhin mit rekombinantem SRY inkubiert. Es konnte jedoch in dem LTR-Bereich bps 515-833 keine Bindung nachgewiesen werden (siehe 2.4.1, Tabelle 2). Es ist daher von großer Wichtigkeit, die übrigen LTR-Bereiche im EMSA zu untersuchen. Die Transkriptionssuchmaschine TESS gibt im Bereich bps 515-833 acht potentielle Bindungsstellen für SRY an (siehe Anhang, 7.2). Es ist zu beachten, dass zwei dieser Bindungsstellen sehr dicht am 3'-Ende der Sonden lokalisiert sind. Auch hier ist nicht auszuschließen, dass dadurch eine stabile SRY-Bindung verhindert wurde. Diese Bereiche bieten sich auch im Zusammenhang mit der putativen, im Cluster M5 lokalisierten Sp1-Bindungsstelle zur Untersuchung an. Da die zwei putativen SRY-Bindungsstellen zwischen den aktivierenden Sp-Bindungsstellen am Transkriptionsstart und der putativen Bindungsstelle bei den bps 576-581 lokalisiert sind, wäre denkbar, dass eine dieser SRY-Bindungsstellen die synergistische Bindung der Sp-Proteine zwischen den zwei entfernten GC/GT-Boxen im regulatorischen bzw. im Core Promotor durch DNA-Biegung erleichtern könnte.

3.3.3 Epigenetische Mechanismen

Die aktive Transkription eines Gens hängt nicht nur vom Core und vom regulatorischen Promotor ab. Schon bei mikroskopischen Untersuchungen von genomischer DNA *in-vivo* fiel auf, dass es DNA-Regionen gibt, die stark kondensiert sind und Bereiche, die entwunden zu sein scheinen. Man nannte diese transkriptionell aktiven Bereiche Euchromatin. Die stark kondensierte durch Histone verpackte DNA nannte man Heterochromatin (Brown 1966).

Die Gliederung in Eu- und Heterochromatin deckte sich später mit Untersuchungen über Methylierungszustände der betroffenen DNA-Abschnitte. Starke Methylierung führte zur Kondensation der DNA und damit zur transkriptionellen Inaktivität (Rice und Allis 2001; Strahl und Allis 2000). In Vertebraten wird bevorzugt der C-5 Kohlenstoff im Cytosin des 5'-CpG-3'-Dinukleotids methyliert. Die antisense 3'-GpC-5' Sequenz kann ebenfalls methyliert werden. Beide Methylgruppen bilden eine dreidimensionale Struktur in der großen Furche der DNA (Ohki et al. 2001). Es zeigte sich, dass ca. 60-90 % der CpG-Sequenzen im Genom methyliert sind. CpG-Dinukleotide können in größeren Anhäufungen vorkommen und werden dann CpG-Inseln genannt. Diese sind meistens mit Promotoren assoziiert. Hier kann die Hypermethylierung zur Repression der Genexpression führen (Ng und Bird 1999).

Epigenetik ist die erbbare Veränderung in der Genexpression, die ohne eine Änderung der DNA-Sequenz einhergeht (Wolffe und Matzke 1999). Sie stellt einen der Transkription übergeordneten Mechanismus der Genkontrolle dar. Durch Faktoren, wie DNA Methyltransferasen, Methyl-CpG Bindungsproteine, Histon-modifizierende Enzyme, „chromatin remodeling factors“, Transkriptionsfaktoren und chromosomale Proteine kann die Genexpression beeinflusst werden. All diese Faktoren sind an der Strukturierung des Genoms beteiligt. Methylierungszustände entscheiden über den Verpackungsgrad der Chromosomen. Diese Modifikationen sind vererbbar. Bei der Replikation mit der anschließenden somatischen Zellteilung werden die Methylierungszustände von der paternalen Zelle auf die Tochterzellen übertragen, da die DNA-Stränge bei der Mitose erhalten bleiben. Methylierungszustände bilden also ein molekulares Gedächtnis der Expressionskontrolle.

Interessante Aspekte der Transkriptionskontrolle von HERV-K zeigte der Versuch, bei dem ein Reportervektor, der die aktive LTR pck30 trug, nach Methylierung aller CpG-Dinukleotide im Reporter-Assay nahezu keine transkriptionelle Aktivität mehr aufwies (siehe 2.5, Abb. 2-18). Auch wenn ein Reportervektor weit davon entfernt ist, ein natürliches Umfeld wie ein Chromosom zu imitieren, zeigte es doch, dass Methylierung einen drastischen Effekt auf die Transkription ausüben kann. Im Promotor finden sich 59 potentielle CpG-Dinukleotide, die auf chromosomaler Ebene zur Stilllegung des HERV-K Promotors

führen könnten. Da eine Histonverpackung des Vektors sehr unwahrscheinlich ist, könnten jedoch Methyl-CpG-Bindungsproteine die Transkription empfindlich stören. Auch könnten methylierungssensitive Transkriptionsfaktoren in ihrer Bindungskapazität inhibiert werden. Die Transkription würde also durch fehlende Transaktivatoren drastisch eingeschränkt werden. Bemerkenswert ist die Beobachtung, dass Sp1 wie auch YY1 zu den methylierungsinsensitiven Transkriptionsfaktoren gehören (Nakao 2001). Das bedeutet, dass Enhancer wie auch Initiationsaktivatoren durch Methylierungszustände nicht beeinflusst werden können. Der Rückgang der pcK30-Aktivität bei Methylierung kann also entweder durch weitere bis jetzt unbekannte bindungsinhibierte Transkriptionsfaktoren zustande kommen oder Methyl-CpG-Bindungsproteine verhindern den Zugang der für die Transkription sehr wichtigen Faktoren Sp1 und YY1. Dieser Sachverhalt müsste durch weitere *in-vivo* Methoden überprüft werden. Methylierungszustände könnten *in-vivo* mit der Bisulfidsequenzierung identifiziert werden. Die Beteiligung von Histonen an der Inaktivierung des HERV-K Promotors könnte mit Chromatin Immunpräzipitation (ChIP) überprüft werden. *In-vivo* kommt Sp1 evtl. noch eine ganz andere Bedeutung zu. Als GC/GT-Box bindendes Protein besetzt Sp1 CpG-Dinukleotide. Diese Dinukleotide könnten durch Sp1 vor Methylierung geschützt werden. Im sich entwickelnden Embryo könnte das von entscheidender Bedeutung sein. Untersuchungen an der Maus haben gezeigt, dass während der embryonalen Entwicklung einer Welle globaler Demethylierung des Genoms eine Welle von *de-novo* Methylierungen folgt (Monk et al. 1987). Die CpG-Inseln bleiben dabei demethyliert, eine Voraussetzung für die Regulation der Genexpression im sich entwickelnden Embryo. Sp1 konnte die Funktion zugeordnet werden, dass es CpG-Dinukleotide vor der Methylierung schützt (Brandeis et al. 1994; Macleod et al. 1994). Es wird vermutet, dass der Expressionslevel von Sp1 während der embryonalen Entwicklung mit den variierenden Methylierungszuständen der exprimierten Gene korreliert (Worrad und Schultz 1997). Ob dies von Relevanz für den HERV-K Promotor ist, müsste noch gezeigt werden.

Abschließend ist zu sagen, dass es sich bei Sp1 um einen versatilen Transkriptionsfaktor handelt, der bei der Expressionentwicklungsspezifischer Gene eine große Rolle spielt. Die Beteiligung dieses Faktors auf transkriptioneller wie auch auf epigenetischer Ebene könnte vom Provirus benutzt werden, um die eigene Transkription zu gewährleisten. In diesem Zuge wäre es sehr interessant, die Methylierungszustände eines aktiven HERV-K Locus in Spermatozyten, Oocyten und embryonalen Zellen zu bestimmen. Dies könnte am Mausmodell mit transgenen Stämmen untersucht werden. Es gibt bereits Hinweise, dass HERV-K Loci in

Teratokarzinomzellen hypomethyliert sind und deswegen eine Expression möglich ist (Florl et al. 1999; Löwer et al. 1995).

Daraus resultierend könnte sich der Kreis schließen, sollte man feststellen, dass nicht nur in Teratokarzinomen die CpGs eigentlich stillgelegter DNA-Regionen hypomethyliert sind, sondern auch in undifferenzierten Keimzellen und embryonalem Gewebe zumindest eine gesteigerte Transkription von HERV-K möglich ist.

3.4 Modell zur Transkription von HERV-K

Die HERV-K LTR weist typische retrovirale Promotorelemente auf. Retroviren sind auf die Transkriptionsmaschinerie der Wirtszelle angewiesen. Die gefundenen Promotorelemente sind deshalb auch typisch für eukaryotische Zellen.

Der HERV-K Promotor untergliedert sich in die Bereiche Core Promotor, regulatorischer Promotor und Enhancer. Der Core Promotor ist die Region, die alle für die basale Transkription notwendigen Elemente enthält. Er besteht aus einem Inr, dem DPE und zwei GC/GT-Boxen.

Ähnliche Strukturen findet man auch bei anderen retroviralen Promotoren wie ERV-9 (La Mantia et al. 1992) und HERV-H (Sjøttem et al. 1996). Der Promotor des DNA-Polymerase β (β -Pol)-Gens ist ein Beispiel für ein Haushaltsgen, das TATA-unabhängig durch Sp1 und ein Inr transkribiert wird (Weis und Reinberg 1997). Allen genannten Promotoren sind ein Inr und stromaufwärts gelegene GC/GT-Boxen gemein. Die GC/GT-Boxen können von Sp-Proteinen gebunden werden. Den Sp-Proteinen wird eine aktivierende Funktion bei der Transkription zugesprochen, die ultimativ das Binden von TFIID ermöglicht.

Inrs besitzen die Konsensussequenz Py Py A⁺¹ N T Py Py. Jedoch ist die Sequenz nicht zwingend für eine aktive Transkriptionsinitiation. Wie in Tabelle 3 dargestellt, weichen die

TCA <u>A</u> TAA	HERV-K
TG <u>A</u> CTCT	HERV-H
TGG <u>G</u> TCT	HIV-1
CC <u>A</u> TTGT	β -Pol
TC <u>A</u> TTCT	TdT

Tabelle 3: Vergleich der Inr-Sequenzen verschiedener Promotoren.

Sequenzen der Inrs, bis auf die des terminalen Desoxynukleotidyl Transferase (TdT) Promotors, von den geforderten Pyrimidinen an den Positionen -1, -2, +4 und +5 relativ zum Transkriptionsstart ab. Selbst bei dem durch das konservierte A belegten Transkriptionsstart kann eine Abweichung bei HIV-1 beobachtet werden. Ein G kann als Startstelle genutzt werden, sollte die -1 Position nicht durch

ein Pyrimidin, sondern durch das Purin G belegt sein (Smale et al. 1990). Bei HERV-K finden sich an Position +4 und +5 zwei Purine. Diese Abweichungen sind im Vergleich zu anderen Inrs jedoch nicht ungewöhnlich und das HERV-K Inr kann als funktional angesehen werden. Die Spezifität der Transkriptionsinitiation kommt vermutlich durch die exakte Positionierung des TFIID-Komplexes durch zusätzliche Aktivatoren, wie Sp1 oder DPE-bindende Proteine, zustande.

Bei HERV-K fällt das Poly(A)-Signal auf das Inr. Bei Transkription des retroviralen Genoms entstehen durch die redundanten LTR-Regionen am 5'- bzw- 3'-Ende der mRNA zwei Poly(A)-Signale. Deswegen muss sichergestellt werden, dass die 5'-Poly(A)-Stelle nicht von der Poly(A)-Polymerase verwendet werden kann. Dies hätte den Verlust sämtlicher ORFs zur Folge, bevor die Translation beginnen könnte. In Abbildung 3-2 ist der Aufbau der LTRs einiger Retroviren in Zusammenhang zur Poly(A)-Stelle gesetzt worden.

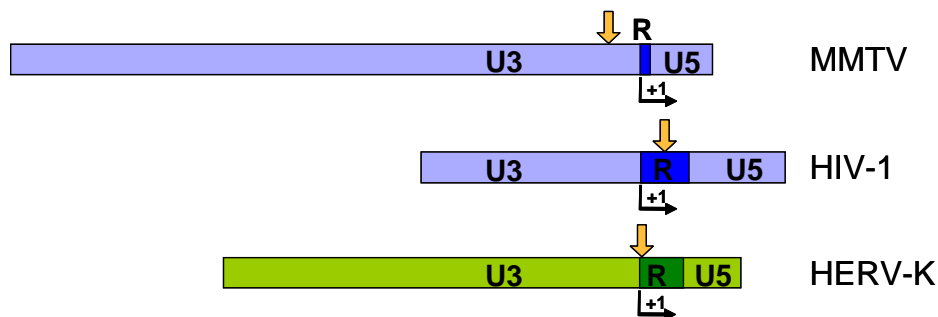


Abbildung 3-2: Vergleich der LTRs verschiedener Retroviren. Dargestellt sind die verschiedenen Regionen der LTR. **U3** Unique 3'-end, **R** Terminal redundancy, **U5** Unique 5'-end, \downarrow Poly(A)-Signal, $\xrightarrow{+1}$ Transkriptionsstart. **MMTV** Maus-Mamma-Tumor-Virus, **HIV-1** Humanes Immundefizienzvirus Typ 1, **HERV-K** humanes endogenes Retrovirus K.

Bei einigen Retroviren, wie MMTV, HTLV-1 oder ASLV liegt das Poly(A)-Signal stromaufwärts der Transkriptionsinitiationsstelle. Das bedeutet, dass nach Transkription nur ein 3'-Poly(A)-Signal vorhanden ist. Eine komplexere Unterdrückung des 5'-Poly(A)-Signals muss bei Retroviren erfolgen, bei denen das Signal stromabwärts von der U3-R-Grenze liegt. Hier sind in der mRNA zwei potentiell funktionale Poly(A)-Signale vorhanden. Bei HIV-1 zeigte sich, dass *in-vivo* beide Signale aktiv sind (Cherrington und Ganem 1992). Man konnte jedoch beobachten, dass das 3'-Poly(A)-Signal dominant ist. Man geht deshalb davon aus, dass eine korrekte Positionierung der Poly(A)-Polymerase nur durch weitere Elemente in U3 stattfinden kann. Ebenfalls denkbar ist ein negativ wirkender Mechanismus, der nur das 5'-Poly(A)-Signal betrifft. Untersuchungen von Iwasaki und Temin zeigten, dass eine

effiziente Polyadenylierung nur möglich ist, wenn das Signal mehr als 500 bps vom 5'-Ende der mRNA entfernt ist (Iwasaki und Temin 1990).

Das Poly(A)-Signal überlappt bei HERV-K mit dem Inr und gruppiert es damit zu den Viren, die eine Poly(A)-Signal-haltige R-Region aufweisen. Dies präsentiert der Poly(A)-Polymerase zwar zwei potentielle Poly(A)-Stellen auf der mRNA, doch ist nicht anzunehmen, dass die extreme Nähe der Konsensussequenz zum 5'-Ende der mRNA eine Prozessierung des 5'-Poly(A)-Signals zulässt.

Zusammenfassend kann man ein erstes grobes Modell zur Transkription von HERV-K aufstellen. Die besprochenen Elemente sind in der Abbildung 3-3 grafisch dargestellt.

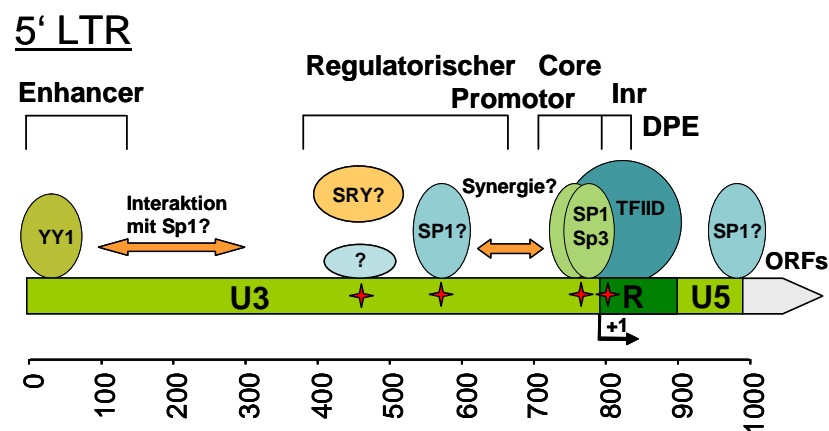


Abbildung 3-3: Modell zur Transkription von HERV-K. Inr Initiatorelement, DPE Downstream Promotor Element, ORF Open Reading Frame. LTR Long Terminal Repeat, R Terminal redundancy, U5 Unique 5'-end, U3 Unique 3'-end, \rightarrow Transkriptionsstart, \star Punktmutationen, die sich stark auf die Promotoraktivität auswirken, \leftrightarrow Putative Interaktionen zwischen Proteinen. Experimentell ermittelte Transkriptionsfaktoren sind in grün, putative in blau dargestellt.

Der Core Promotor, definiert durch Inr, DPE und GC/GT-Boxen ermöglicht das Binden des basalen Faktors TFIID. Das Inr legt das erste transkribierte Nukleotid fest, das bei bp 794 liegt. Dem DPE kommt eine Rolle bei der Positionierung von TFIID bei TATA-losen Promotoren zu. Im regulatorischen Promotor, der sich meistens 50-400 bps stromaufwärts vom Transkriptionsstart befindet, konnten zwei Regionen bestimmt werden, die regulatorische Bedeutung zu haben scheinen. Eine putative GC/GT-Box bei den bps 576-581 ist sehr anfällig gegenüber Mutationen. Eine synergistische Wechselwirkung zwischen den Sp-Proteinen muss noch bewiesen werden.

Ein zweiter Bereich im regulatorischen Promotor scheint von transkriptioneller Bedeutung zu sein. Eine Mutation von 4 bps in der Region 242-310 führte zu einem Anstieg der transkriptionellen Aktivität.

Der Enhancer schließlich wird vom Transkriptionsfaktor YY1 besetzt. Interaktionen mit Sp1 sind beschrieben worden (Seto et al. 1993), doch genauere Untersuchungen müssten zeigen, welche Elemente des HERV-K Promotors von diesem Faktor beeinflusst werden.

Der Einfluss und die genaue Lokalisierung von SRY muss noch untersucht werden, doch scheint eine Aktivierung der HERV-K LTRs durch SRY möglich zu sein. Am 3'-Ende des LTRs ist ein negativ wirkendes Element identifiziert worden, das nach einer Datenbankanalyse eine putative Sp1-Bindungsstelle beinhaltet.

3.5 Ausblick

Die in dieser Arbeit dargestellten Ergebnisse zeigten eindeutig, dass der Promotor von HERV-K in einer TATA-unabhängigen Weise reguliert wird. Es konnte gezeigt werden, dass der basale Promotor durch Sp1 und Sp3 gesteuert wird und dass SRY die Transkription des HERV-K LTRs positiv beeinflussen kann.

Diese Befunde werfen einige Fragen für weiterführende Experimente auf:

- Gibt es weitere Sp1/Sp3-Bindungsstellen auf der HERV-K LTR?
- Wenn ja, gibt es synergistische Effekte zwischen den Sp-Proteinen oder zwischen den Sp-Proteinen und anderen Transkriptionsfaktoren, wie z.B. YY1?
- Welche Rolle spielt SRY? Bindet es direkt auf der HERV-K LTR, oder wirkt es in *trans* durch Aktivierung eines Gens, das einen für HERV-K essentiellen Transkriptionsfaktor codiert?

Diese Fragestellungen könnten durch eine erweiterte EMSA-Analyse geklärt werden. Überlappende Sonden, die den gesamten LTR-Bereich abdecken, könnten mit rekombinantem Sp1, Sp3 und SRY inkubiert werden, um weitere Bindungsstellen zu determinieren. Bindungspartner der Transkriptionsfaktoren könnten z.B. durch ein Biotin-Streptavidin-System isoliert werden. Dabei wird die LTR-Sonde, die die Konsensussequenz (z.B. eine GC/GT-Box) enthält, mit nukleären Extrakten inkubiert. Die starke Biotin-Streptavidinbindung wird benutzt, um die an der biotinylierten Sonde anhaftenden Transkriptionsfaktoren und deren Bindungspartner über die Streptavidin-Matrix zu isolieren.

Die Transkription der HERV-K LTR wurde durch Methylierung der LTR stark eingeschränkt. Auch hier ergeben sich Fragestellungen für weiter Experimente:

- In welcher Art wirkt sich die die Methylierung auf die Transkription aus?
- Ist Sp1, als GC/GT-Box bindender Transkriptionsfaktor, an der Regulierung der Methylierung der CpG-Dinukleotide der HERV-K LTR beteiligt?
- Werden durch Methylierung Transkriptionsfaktoren daran gehindert, auf der LTR zu binden, oder üben Methylgruppen-bindende Proteine einen negativen Effekt auf die Transkription aus?
- Werden auf genomischer Ebene HERV-K Loci durch Hypermethylierung und der damit verbundenen lokalen Chromatinkondensation stillgelegt?
- Ist die Hypomethylierung und damit die transkriptionelle Aktivität eines spezifischen HERV-K Locus gewebespezifisch, und wenn ja, welche Bedingungen führen zur Demethylierung der LTR?
- Korreliert das Expressionsmuster von Sp1 mit den Methylierungszuständen auf dem HERV-K Locus?
- Gibt es während der Embryogenese Zeitfenster, die durch eine erhöhte Sp1-Expression HERV-K Expression zulassen?

Diese Fragen könnten mit Hilfe der Bisulfit-Modifikation bearbeitet werden. Bei dieser Methode wird genomische DNA eines bestimmten Gewebes isoliert und einer Bisulfit-Modifikation unterzogen. Alle unmethylierten Cytosine werden hierbei zu Thyminen modifiziert. Amplifikation eines spezifischen HERV-K Locus und anschließende Sequenzierung erlauben die Determinierung der methylierten CpG-Dinukleotide auf der LTR. Eine Chromatin-Immunpräzipitation (ChIP) könnte den Status der DNA-Kondensation eines HERV-K Locus bestimmen. Dabei werden *in-vivo* die auf der genomischen DNA anhaftenden Proteine chemisch vernetzt (crosslinking). Nach Isolierung des Chromatins wird die genomische DNA durch Scherkräfte in kleine Fragmente zerlegt. Durch einen Antikörper gegen deacylierte Histone können Histon-haltige DNA-Stücke präzipitiert werden. Aufhebung der Vernetzung und Verdau der Proteine erlauben eine sequenzspezifische Amplifikation von DNA-Fragmenten, an die Histone gebunden waren. Eine positive Amplifikation mit HERV-K-spezifischen Primern und anschließende Sequenzierung erlauben eine Aussage über HERV-K Loci, die von Histonen besetzt waren. Auf diese Weise könnte

man Rückschlüsse auf den Kondensierungszustand eines spezifischen HERV-K Locus gewinnen.

Eine Untersuchung der Methylierungszustände von HERV-K Loci im Vergleich zum Expressionsmuster von Sp1 und Sp3 durch z.B. eine Western Blot Analyse in verschiedenen Geweben sollte weitere Einblicke in die Regulation des Testis-spezifischen HERV-K Promotors gestatten. Es sollte klar werden, welche Rolle die transkriptionellen Schlüsselkomponenten Sp1, Sp3 und SRY bei der gewebespezifischen Expression von HERV-K spielen und inwieweit die Methylierung Gewebespezifität unterstützt oder sogar bedingt.

4 Material und Methoden

Für die molekularbiologischen Arbeiten wurden, soweit nicht anders angegeben, Standard-Methoden verwendet (Sambrook 1989). Die Zentrifugationen sind in rpm (rounds per minute) angegeben. Zentrifugation von größeren Volumina (50 ml Gefäße) sind in den Zentrifugen Varifuge 3.0R und Minifuge RF (Heraeus), kleine Volumina (1,5 ml - 2,2 ml Gefäße) sind in den Zentrifugen Biofuge pico, A und 15R (Heraeus) und Centrifuge 5402 (Eppendorf) ausgeführt worden. Ultrazentrifugiert wurde in der TL-100 Ultracentrifuge (Beckman) mit einem TLA 100.1 Rotor. Soweit nicht anders angegeben, erfolgten alle Arbeiten bei RT.

4.1 Häufig benutzte Geräte und Materialien

Autoklaven:	Hochdruckdampfsterilisator LS6/6/6-2FD-GT KT 77.169 (BHS Medizintechnik), Bioclav (Schütt)
Chemikalien:	soweit nicht anders angegeben von den Firmen Amersham Biosciences, Boehringer Mannheim, Invitrogen, Merck, New England Biolabs, Serva, Sigma, Roth
Glasgeräte:	Bechergläser, Glaspipetten, Messzylinder, Erlenmeyerkolben etc. (Schott, Brand)
Handschuhe:	Manufix (Braun)
Kühlschrank (4 °C) :	Liebherr
Magnetrührer:	RET GS (Ikamag)
pH-Meter:	Mikroprozessor-pH-Meter (Knick)
Photometer:	Gene Quant Pro (Amersham Pharmacia Biotech)
Pipetten und -spitzen:	Eppendorf
Plastikgefäße:	0,2 ml Gefäße (Eppendorf), 1,5 ml, 2 ml Gefäße (Eppendorf), 12 ml, 20 ml, 25 ml, 50 ml Röhrchen (Greiner)
Thermomixer:	Modell 5436 und 5437 (Eppendorf)
Tiefkühlschränke:	-20 °C (Liebherr), -85 °C Ultra Low Freezer (New Brunswick Scientific)
Transformator:	Electrophoresis Power Supply-EPS 301 (Amersham); Power Pac 300 (Bio Rad)
Vakuumpumpe:	Vakuubrand
Vortexer:	Heidolph Reax 2000

Waagen:	Excellence (Satorius); 2024 MP (Satorius); Feinwaage: 2024 MP (Satorius)
Wasserbad:	Köttermann
Wippschüttler:	Biometra WT12
Zentrifugen:	Tischzentrifugen: Picofuge (Stratagene), Biofuge (Modelle 15R, pico oder A; Heraeus), Centrifuge 5402 (Eppendorf) Untertischzentrifuge: Varifuge 3.0R, Minifuge RF (Heraeus) Ultrazentrifuge: TL-100 (Beckman)

4.2 Zellkultur

4.2.1 Geräte und Materialien

Brutschrank:	Cytoperm (Heraeus)
Mikroskop:	IDO 3 (Zeiss)
Mundschutz:	Coldex
Sterilbank:	Lamin Air HLB2448 GS (Heraeus)
Stickstofftonne:	BT55/BT40 (L'air liquide)
Zellkulturflaschen:	80 cm ² , straight neck, Nunclon Flasche (Nalge Nunc International)

4.2.2 Medien und Chemikalien

DMEM:	Dulbecco's Minimal Essential Medium; (Biochrom KG)
DMSO:	Dimethylsulfoxid (Sigma)
FKS:	Fötale Kälberserum (PAA)
L-Glutamin:	100x, 205 mM
NSP:	250x, Neomycin, Streptomycin, Penicillin; je 10.000 U/ml in Hank's Lösung
PBS (ohne Ca ²⁺ , Mg ²⁺):	137 mM NaCl, 3 mM KCl, 16,5 mM Na ₂ HPO ₄ , 1,5 mM KH ₂ PO ₄
Trypsin/Versen:	10 ml 2,5 % Trypsin, 40 ml Versen, 50 ml PBS
Versen:	159 mM NaCl, 2,7 mM KCl, 6,5 mM Na ₂ HPO ₄ , 1,5 mM KH ₂ PO ₄ , 2,7 mM Versen (=EDTA)

4.2.3 Zellkultur

Folgende adhärenzte Zelllinien wurden verwendet:

GH	Humanes Teratokarzinom	(Löwer et al. 1984)
HeLa	Humanes Cervixkarzinom	(American Tissue Culture Collection, ATCC)

4.2.4 Grundhaltung von Zellen

In der Grundhaltung wurden alle Zellen in 80 cm² Kulturflaschen (Nunc) in ca. 20 ml DMEM-Medium, 10 % FKS, 50 U/ml NSP, 2,3 mM Glutamin, im Cytoperm-Brutschrank bei 37 °C, 5 % CO₂ und 80 % Luftfeuchtigkeit kultiviert. Alternierend wurde pro Woche je einmal das Medium gewechselt und die Zellen im Verhältnis 1:2-1:10 (GH) passagiert, HeLa-Zellen wurden zweimal pro Woche passagiert: Zur Teilung wurden die Zellen mit Trypsin-Versen gewaschen, mit 1 ml Trypsin-Versen kurz bei 37 °C abgelöst und in ca. 20 ml frischem Medium auf neue, mit CO₂ vorbegeaste Zellkulturflaschen aufgeteilt.

4.2.5 Transfektion eukaryotischer Zellen

Lysispuffer:	Passive Lysis Buffer, 5x (Promega)
Pipettenspitzen:	Biozym budget line Safe Seal Tips (Biozym)
Transfektionsreagenz:	Lipofectamine (Invitrogen)
Trypanblau:	0,5 % in PBS (ohne Ca ²⁺ und Mg ²⁺) (Merck)
Zählkammer:	incl. geschliffenem Deckglas (Fuchs Rosenthal)
Zellkulturflaschen:	80 cm ² , straight neck, Nunclon Flasche (Nalge Nunc International)
96-Loch-Platten:	Nunclon MicroWell Platte (Nalge Nunc International)

Zur transienten Transfektion eukaryotischer Zellen wurden ca. 3 (HeLa) - 6 (GH) x10⁴ Zellen (gezählt in einer Rosenthal-Kammer) in 96-Loch-Platten ausgesät. Nach 24 (HeLa) – 48 (GH) Stunden wurden die Zellen mit rekombinantem Plasmid (100 ng Reporter Luziferasevektor, 20 ng Renilla-Vektor) mit Lipofectamine nach Herstellerangaben transfiziert und in FKS-freiem Medium inkubiert. Nach 3 Stunden erfolgte ein Mediumwechsel mit FKS-haltigem Medium. 27 Stunden nach Transfektion wurden die Zellen im dualen Luziferase-Assay weiterverarbeitet (siehe 4.5.5).

Für Transfektionen, die anschließender RNA-Präparation dienten, wurden 5 (HeLa) – 20 (GH) $\times 10^6$ Zellen in 145 cm² Platten ausgesät. Nach 24 (HeLa) - 48 (GH) Stunden wurden die Zellen mit 8 μ g rekombinantem Plasmid transfiziert. 27 Stunden nach Transfektion wurde die RNA der Zellen präpariert.

4.3 Bakterienkultur

4.3.1 Geräte

Bakterienschüttler:	New Brunswick Scientific Edison N.J. USA
Brutschrank:	Typ B5050E (Heraeus)
Sterilbank:	Lamin Air HLB2448 GS (Heraeus)

4.3.2 Medien und Chemikalien

LB-Medium:	Luria-Bertani-Medium, 1 % Bacto-Trypton, 0,5 % Bacto Hefe-Extrakt, 1 % NaCl, pH 7,0; autoklaviert
LB-Amp-Medium:	LB-Medium mit 100 μ g/ml Ampicillin
LB-Agar-Platten:	1,5 % Bacto-Agar in LB-Medium, pH 7,0; autoklaviert
LB-Amp-Platten:	LB-Agar-Platten mit 100 μ g/ml Ampicillin
LB-Amp-X-Gal-Platten:	LB-Agar-Platten mit 100 μ g/ml Ampicillin und 80 μ g/ml X-Galactosidase
SOC-Medium:	2 % Bacto-Trypton, 0,5 % Bacto-Yeast Extract, 8,5 mM NaCl, 2,5 mM KCl, 10 mM MgCl ₂ , 20 mM Glucose; pH 7,0; sterilfiltriert
TFB1-Puffer:	30 mM K-Acetat, 100 mM RbCl, 50 mM Mn(II)Cl ₂ , 10 mM CaCl ₂ , 15 % Glycerin, pH 5,8; sterilfiltriert
TFB2-Puffer:	10 mM MOPS, 75 mM CaCl ₂ , 10 mM RbCl, 15 % Glycerin, pH 6,5; sterilfiltriert

4.3.3 Bakterienstämme

Escherichia coli:	DH5 α (Invitrogen) XL-1 Blue (Stratagene)
-------------------	---

4.3.4 Präparation kompetenter Bakterienzellen

Für die Transformation rekombinanter Plasmide wurde *Escherichia coli* DH5 α verwendet, die modifiziert nach Hanahan (Hanahan 1983) kompetent gemacht wurden.

Eine 100 ml Kultur in LB-Medium wuchs im Schüttelinkubator bei 37 °C, 225 rpm zu einer OD₅₅₀ von 0,5-0,55 und wurde anschließend für 5 Minuten auf Eis gekühlt. Die Bakterien wurden für 10 Minuten bei 3.000 rpm, 4 °C zentrifugiert (Minifuge), in 20 ml TFB1 Puffer resuspendiert und für 5 Minuten auf Eis inkubiert. Nach erneuter Zentrifugation wurden die Bakterien in 2 ml TFB2 Puffer aufgenommen, 15 Minuten auf Eis inkubiert und in 100 μ l Aliquots aufgeteilt. Die Aliquots wurden sofort in flüssigem Stickstoff eingefroren und bei -70 °C gelagert.

4.3.4.1 Transformation rekombinanter DNA

Pro Transformation wurden 50 μ l auf Eis aufgetaute kompetente Zellen mit bis zu 10 μ l Ligationsansatz versetzt und 10-30 Minuten auf Eis inkubiert. Nach einem 42 °C Hitzeschock für 90 Sekunden wurden die Bakterien für 2 Minuten auf Eis gekühlt und nach Zugabe von 500 μ l LB-Medium für maximal 1 Stunde bei 37 °C, 225 rpm regenerieren gelassen. Die Bakterien wurden für 5 Minuten bei 3500 rpm zentrifugiert (Biofuge); 500 μ l Medium wurde verworfen und die in den restlichen 50 μ l SOC-Medium resuspendierten Zellen auf geeignetem Selektionsmedium (LB-Amp-Platten) ausplattiert. Die Agar-Platten wurden dann invertiert über Nacht bei 37 °C inkubiert.

4.3.5 Plasmid Präparation aus Bakterien

Plasmid DNA wurde durch modifizierte alkalische Lyse (Birnboim 1983; Birnboim und Doly 1979) aus Bakterien isoliert. Die erhaltenen Plasmide wurden mittels Restriktions-Fragment-Längen-Analyse überprüft und mutmaßlich richtige Plasmide anschließend sequenziert. Erwünschte Klone wurden eingefroren: Aus einer Übernacht-Kultur wurden 3 ml Bakterien Suspension entnommen, mit 1 ml 80 % Glycerin versetzt und bei -70 °C gelagert.

4.3.5.1 Plasmid Mini Präparation

Von einer 3 ml Übernacht-Kultur (37 °C, 225 rpm, LB-Amp-Medium) einzelner Bakterienkolonien aus den Transformationen wurden 2 ml für 5 Minuten bei 5.000 rpm zentrifugiert (Biofuge). Für die Präparation wurde das „Qiaprep MiniPrep Kit“ (Qiagen) nach

Herstellerangaben verwendet. Die DNA-Konzentrationen wurden spektrophotometrisch bestimmt.

4.3.5.2 Plasmid MIDI Präparation

Plasmide aus einer 100 ml (high copy number Plasmide; z.B. pBluescript) bzw. 250 ml (low copy number Plasmide; z.B. pSVOAL) Übernacht-Kultur (37 °C, 225 rpm, LB-Amp-Medium) wurden mit dem „Qiagen Plasmid Purification Kit“ (Qiagen) entsprechend den Herstellerangaben isoliert. Die DNA-Konzentrationen wurden spektrophotometrisch bestimmt.

4.4 Nukleinsäure-Methoden

4.4.1 Präparation von RNA aus eukaryotischen Zellen

DEPC-H₂O: 0,1 % DEPC in Aq. bidest.; autoklaviert

Die RNA wurde mit Hilfe der RNA-Kits von Qiagen und Promega präpariert. Es wurden nur sterilisierte Plastikpipetten (Greiner) und Reaktionsgefäße (Eppendorf, Nunc) verwendet, um RNase Kontaminationen zu verhindern. Zum Lösen von RNA und bei den spektrophotometrischen Messungen wurde ausschließlich DEPC-H₂O benutzt.

4.4.1.1 Präparation von Gesamt-RNA

Die Zellen wurden auf Eis dreimal mit eiskaltem PBS (ohne Ca²⁺, Mg²⁺) gewaschen und in 2x 5 ml PBS (ohne Ca²⁺, Mg²⁺) pro 145 cm² Kulturplatte mit einem Zellschaber mechanisch vom Boden gelöst. Die Zellsuspensionen wurden vereinigt und für 10 Minuten bei 1.000 rpm, 4 °C, zentrifugiert. Um die Gesamt-RNA zu gewinnen, wurde das Pellet entsprechend der Zellmenge mit dem „RNeasy Mini Kit“ oder dem „RNeasy MIDI/Maxi Kit“ (Qiagen) nach Herstellerangaben präpariert. Die RNA-Konzentrationen wurden spektrophotometrisch bestimmt.

4.4.1.2 Präparation von Poly(A)-RNA

5 mg Gesamt-RNA wurden mit dem „polyATtract mRNA Isolation System“ (Promega) entsprechend den Spezifikationen des Herstellers aufgereinigt.

4.4.2 Polymerase Chain Reaction (PCR)

Cycler:	T3 Thermocycler (Biometra) iCycler Thermal Cycler (Bio-Rad)
PCR Reagenzien:	GeneAmp 10 mM dNTPs (Applied Biosystems) AmpliTaq DNA Polymerase, 10x PCR-Puffer (Perkin-Elmer) AmpliTaq Gold DNA Polymerase, 10x AmpliTaq Gold-Puffer (Perkin-Elmer) Pfu Turbo DNA Polymerase DpnI Restriktionsenzym (Stratagene)
Primersynthese:	Thermo Hybaid
Sterilbank im PCR-Labor:	Lamin Air HLB 2448 GS (Heraeus)

PCR mit AmpliTaq DNA Polymerase: Eine 50 µl Reaktion bestand aus: 1-20 ng Template DNA, je 0,2-1 µM Primer 1 und 2, 1x PCR-Puffer (incl. 1,5 mM MgCl₂), je 0,25 mM dATP, dTTP, dGTP, dCTP und 1-2,5 U AmpliTaq DNA Polymerase.

PCR mit AmpliTaq Gold DNA Polymerase: Eine 50 µl Reaktion bestand aus: 1-20 ng Template DNA, je 0,5-1,0 µM Primer 1 und 2, 1x AmpliTaq Gold-Puffer (incl. 1,5 mM MgCl₂), je 0,25 mM dATP, dTTP, dGTP, dCTP und 2,5 U AmpliTaq Gold DNA Polymerase. Gegebenfalls wurden die Mg²⁺-Konzentrationen zwischen 0,5-4 mM variiert.

Reaktionsbedingungen im Thermocycler: Denaturierung: 30-60 Sekunden, 94 °C; Hybridisierung: 30 Sekunden, 55 °C; Polymerisation: 60-120 Sekunden, 72 °C; erste Denaturierung 2 Minuten, 94 °C; letzte Polymerisation: 10 Minuten, 72 °C; 20-30 Zyklen. Die Hybridisierungstemperatur von 55 °C sollte 5-10 °C unterhalb der Schmelztemperatur des Primerpaars liegen, die mit der Base-Stacking-Methode überprüft und ggf. verändert wurde. Konnte mit diesen Bedingungen keine DNA Amplifikation erzielt werden, wurde die Polymerisationszeit verlängert. Zur Identität der verwendeten Template DNA und Primer siehe 4.7.

Zur Einführung von Mutationen in ein rekombinantes Plasmid wurde das „QuickChange-“ oder „ExSite PCR-Based Site-Directed-Mutagenesis Kit“ (Stratagene) nach Angaben des Herstellers verwendet. Hier wurde das ganze Plasmid von der Pfu Turbo DNA Polymerase synthetisiert und die Mutationen wurden über mutagene Primer eingefügt. Unabdingbar für diese Methode waren eine verlängerte Polymerisierungszeit von 1 Minute/kB Plasmid, eine erhöhte Templatmenge von 100 ng und eine verminderte Zykluszahl von 15-20.

Transformationen sollten nur mit superkompetenten Zellen mit einer Effizienz von mindestens $1 \times 10^9/\mu\text{g}$ Plasmid (XL1-blue, Stratagene) erfolgen.

4.4.3 Oligonukleotid-Hybridisierung

Hybridisierungspuffer: 10 mM Tris, pH 7,5 – 8,0, 50 mM NaCl, 1 mM EDTA

Gleiche Mengen zweier in Hybridisierungspuffer gelöster und komplementärer Oligonukleotide ($100 \mu\text{M}$) wurden für 2 Minuten bei $95 \text{ }^\circ\text{C}$ im Thermocycler denaturiert und mit einer „Ramping“ Zeit von 30 Minuten auf RT abgekühlt. Lagerung der Oligos erfolgte kurzfristig bei $4 \text{ }^\circ\text{C}$, langfristig bei $-70 \text{ }^\circ\text{C}$.

4.4.4 Restriktion von DNA

Restriktionsenzyme: Boehringer Mannheim, NEB

Puffer: 10x, vom Hersteller mitgeliefert

BSA: 100x, 10 mg/ml bovines Serum-Albumin, vom Hersteller mitgeliefert (NEB)

Pro μg DNA wurden 1-2 Units (U) Enzym zugesetzt. Die Reaktionsbedingungen (Temperatur, Puffer, eventuell BSA-Zusatz) entsprachen den Angaben des Herstellers. Kontrollrestriktionen von Mini-, Midi- und Maxipräparationen wurden ca. 2 Stunden in 10-20 μl Gesamtvolumen inkubiert, PCR-Produkte sowie Vektoren, die für Klonierungen benötigt wurden, über Nacht in 100-200 μl . Wenn die Pufferbedingungen dies zuließen, wurden teilweise mehrere Enzyme in einer Reaktion eingesetzt. Nach dem Verdau folgte eine Phenol-Chloroform-Extraktion mit anschließender Fällung der DNA, um das Enzym und BSA (falls vorhanden) zu eliminieren oder die Fragmentlängenanalyse im Agarose-Gel.

4.4.5 Generierung von Blunt Ends

Enzym: Klenow-Fragment der DNA-Polymerase I (NEB)

Puffer: 10x für Klenow (NEB)

Nukleotide: dNTPs: dCTP, dGTP, dATP, dTTP-Lösung (je 2,5 mM; Perkin-Elmer), $-20 \text{ }^\circ\text{C}$

Zum Auffüllen überhängender 5'-Enden wurden pro 50 µl Restriktionsansatz je 1 µl der zum Auffüllen notwendigen dNTPs zugegeben. Sollten überhängende 3'-Enden entfernt werden, wurden keine Nukleotide zugegeben, um die 3'-5' Exonuklease des Klenow-Fragments zu aktivieren. Pro µg DNA wurde 1 U Klenow-Fragment der DNA-Polymerase I eingesetzt und der Ansatz bei Raumtemperatur für 30 Minuten inkubiert. Zum Denaturieren des Enzyms wurde 10 Minuten auf 75 °C erhitzt. Anschließend wurde eine Phenol-Chloroform-Extraktion zum Entfernen des Enzyms durchgeführt, sowie eine Aufreinigung über ein präparatives Agarosegel.

4.4.6 Dephosphorylierung des Vektors

CIP: Calf Intestinal Phosphatase, 1 U/µl (Promega)
Puffer: 10x (Promega)

Die Reaktion entfernt am 5'-Ende restringierter DNA die Phosphatgruppe und verhindert so die intermolekulare und die Selbst-Ligation des Vektors. Nur das Phosphatgruppen-haltige Insert kann an die Enden des Plasmids angefügt werden.

Die Dephosphorylierung wurde für Schnittstellen mit 5' Überhang mit 1 U Enzym pro µg DNA für 1 Stunde bei 37 °C durchgeführt. Bei blunt ends und 3' Überhängen wurde 15 Minuten bei 37 °C und 15 Minuten bei 56 °C inkubiert. Nach Zugabe einer weiteren U Enzym wurden die Inkubationszeiten wiederholt. Um die Reaktion zu beenden, wurde das Enzym durch 10-minütiges Erhitzen auf 65 °C inaktiviert. Im Anschluss folgte eine Phenol-Chloroform-Extraktion, um die Alkalische Phosphatase zu entfernen und eventuell eine Aufreinigung der DNA über ein präparatives Gel.

4.4.7 Phenol-Chloroform-Extraktion

Phenol: Roti-Phenol (Roth)
Phenol-Chloroform: Phenol, Chloroform, Isoamylalkohol; Verhältnis 50:50:1
Chloroform-Isoamylalkohol: Chloroform, Isoamylalkohol; Verhältnis 50:1

Die wässrige Lösung der DNA wurde mit dem gleichen Volumen Phenol versetzt, eine Minute lang von Hand geschüttelt und dann 1 Minute bei 10.000 rpm zentrifugiert (Biofuge pico). Die obere, wässrige Phase wurde in ein neues Eppendorf-Gefäß überführt. Dieser Vorgang wurde zunächst mit einer Phenol-Chloroform-Lösung und dann mit einer

Chloroform-Isoamlyalkohol-Lösung wiederholt. Die DNA wurde dann über ein präparatives Gel aufgereinigt oder gefällt.

4.4.8 Fällung von DNA

Speed Vac: SVC 100 (Savant)
 Chemikalien: 3 M Natriumacetat-Lösung,
 Ethanol abs.,
 Ethanol 70 %

Der wässrigen DNA-Lösung wurden 0,1 Volumenanteil einer 3 M Natriumacetat-Lösung sowie ca. 2-2,5 Volumenanteile Ethanol abs. zugesetzt und geschüttelt. Anschließend wurde 20 Minuten bei -20 °C präzipitiert und dann 30 Minuten bei 13.500 rpm und 4 °C zentrifugiert (Biofuge 15R). Das Pellet wurde mit 500 µl 70 %-igem Ethanol gewaschen und erneut für 5 Minuten zentrifugiert. Danach wurde das Pellet in der Speed Vac getrocknet und in einem geeigneten Volumen Aqua bidest bei 4 °C über Nacht gelöst.

4.4.9 Agarosegel-Elektrophorese

Agarose: für analytische Gele: Standard EEO research grade (Serva); für präparative Gele: Seakem LE Agarose FMC (Biozym Diagnostics)
 Agarosegelkammern: diverse horizontale Gelapparaturen (NEB; Owl Scientific, Inc.)
 DNA-Lade-Puffer: 10x, 0,25 % Bromphenolblau, 0,25 % XylenCyanol, 15 % Ficoll
 Größenstandard: 1 kb-Marker oder 100 bp ladder (Invitrogen)
 TAE-Puffer: 40 mM Tris-Acetat, pH 7,2, 1 mM EDTA
 Transformator: Power Pac 300 (Bio-Rad)
 Ethidiumbromid-Bad: TAE-Puffer mit 10 mg/ml Ethidiumbromid
 Geldokumentation: UV-Transilluminator TI2, Biodoc II (Biometra)

Die Konzentration des Gels wurde im Rahmen von 0,8 % bis 3 % um so höher gewählt, je kleiner die zu trennenden Fragmente waren. Die DNA-Lösung wurde vor dem Auftragen mit DNA-Ladepuffer versetzt. Die an der Elektrophoresekammer angelegte Spannung betrug je nach Größe des Gels zwischen 80-120 Volt. Die DNA-Fragmente wurden durch ihre Wanderung im elektrischen Feld der Größe nach aufgetrennt. Anschließend wurde die DNA

im Gel im Ethidiumbromidbad angefärbt und unter UV-Licht sichtbar gemacht. Das Gel konnte jetzt bildgebend dokumentiert werden, die gewünschten Banden wurden möglichst eng ausgeschnitten.

4.4.9.1 Isolierung von DNA aus Agarose

Die DNA wurde aus Seakem-Agarose wiedergewonnen, indem die Agarosestücke mit der gewünschten Bande mit dem „Qiaquick Gel Extraction Kit“ (Qiagen) extrahiert wurden, wobei die Herstellerangaben befolgt wurden. Die Elution wird allerdings immer durch Zugabe von 30 µl 70 °C warmes Aqua bidest und folgender Inkubation bei 37 °C für 10 Minuten durchgeführt. Die DNA-Konzentrationsbestimmung des Eluats erfolgte photometrisch.

4.4.9.2 Reinigung und Isolierung von DNA aus Low Melting Point (LMP) Agarose Gelen

Agarose: NuSieve GTG Agarose (Cambrex)

DNA-Fragmente (in DNA-Lade-Puffer) wurden auf einem 4 % (500-1.000 bp Fragmente) bzw. 6 % (30-500 bp Fragmente) GTG Agarosegel unter Standard Elektrophorese Bedingungen in TAE-Puffer in einer horizontalen Gelapparatur aufgetrennt. Die mit Ethidiumbromid gefärbten, auf einem UV-Transilluminator sichtbar gemachten DNA-Banden wurden eng aus dem Gel ausgeschnitten. Die Gelstücke wurden bei 60 °C geschmolzen und einmal mit Phenol extrahiert und die Phenol-Phase mit TE pH 8,0 reextrahiert. Die vereinigten wässrigen Phasen wurden einmal mit Phenol/Chloroform und zweimal mit Chloroform extrahiert. Anschließend wurde die DNA mit 2 Volumen 96 % Ethanol und 0,1 Volumen 3 M Natriumacetat-Lösung präzipitiert, mit 70 % Ethanol gewaschen und in Aq. bidest. aufgenommen.

4.4.10 Ligation

Ligase: T4 Ligase (NEB)

Puffer: 10x (NEB)

Kit: pGEM-T Easy Vector System (Promega)

Zur Ligation wurden 100 ng dephosphorylierter Vektor (Calf Intestinal Phosphatase; Promega, nach Herstellerangaben) und das zu ligierende DNA-Fragment im molaren Verhältnis 1:3 eingesetzt. Die Ligation erfolgte in einem Gesamtansatz von 20 µl in 1x T4 DNA Ligase-Puffer mit 400 U (NEB units!) T4 DNA Ligase (NEB) für 10 Minuten bei RT (sticky end ligation). Für blunt end Ligationen wurde die Vektormenge auf bis zu 1 µg erhöht, 2.000 U Ligase verwendet und die Inkubation erfolgte bei 14 °C über Nacht.

PCR-Produkte für die Sondengenerierung beim EMSA und 5'-RACE PCR-Produkte wurden in den pGEM-T Easy Vektor nach Herstellerangaben (Promega) ligiert.

4.4.11 Rapid Amplification of 5'-cDNA ends (5'-RACE)

Dnase:	DNase I, Amplification Grade, 1 U/µl (Invitrogen)
Puffer:	10x (Invitrogen)
Phenol:	Phenol, TE-Puffer pH 8,0 gesättigt, 0,1 % Hydrochinon
DEPC-H ₂ O:	0,1 % DEPC in Aq. bidest.; autoklaviert
DMSO:	Dimethylsulfoxid (Sigma)
Kit:	GeneRacer Kit (Invitrogen)

Die 5'-RACE Technik ermöglicht es, ausgehend von Bereichen bekannter Nukleotidsequenz, unbekannte Bereiche am 5'-Ende der cDNA gezielt zu amplifizieren und deren Sequenz zu bestimmen (Frohman et al. 1988). Dafür wurde das „GeneRacer Kit“ (Invitrogen) nach Angaben des Herstellers mit folgenden Modifikationen verwendet: Bei allen enzymatischen Schritten wurde DEPC-H₂O benutzt. Die RNA wurde vor Beginn einem DNase I Verdau unterzogen, um genomische DNA Kontaminationen zu unterbinden. Dafür wurde die RNA (250 ng mRNA bzw. 1 µg Gesamt-RNA) in einem 10 µl Ansatz mit 1 U Enzym und 1x DNaseI-Puffer für 15 Minuten bei RT inkubiert und anschließend wurde das Enzym für 10 Minuten bei 65 °C deaktiviert und dann auf Eis gestellt. Zur Auflösung von Sekundärstrukturen wurde die RNA vor jedem enzymatischen Schritt für 5 Minuten auf 80 °C erhitzt und danach sofort auf Eis gestellt. Zur Reduzierung der Sekundärstrukturen in GC-reichen RNA-Regionen wurde der RT-Reaktion sowie der anschließenden PCR 5 % DMSO zugesetzt. Es wurde die „SUPERSCRIPT II Reverse Transcriptase“ verwendet.

Etwa 250 ng mRNA wurde mit einer Phosphatase (CIP) behandelt, wodurch die Phosphatreste am 5'-Ende von verkürzter mRNA, rRNA und tRNA entfernt wurden. Davon ausgenommen waren lediglich die Phosphatreste intakter mRNA-Moleküle, welche durch die 5'-Cap-Struktur vor Dephosphorylierung geschützt waren. Im zweiten Schritt wurden die

5'-Cap-Strukturen von der intakten mRNA durch eine Pyrophosphatase (total acid phosphatase, TAP) entfernt, wodurch diese Moleküle einen freien Phosphatrest am 5'-Ende erhielten. Mit Hilfe dieses Restes konnte durch das Enzym T4-RNA-Ligase ein aus 44 Basen bestehendes synthetisches RNA-Oligonukleotid (GeneRacer-Oligo: 5'-CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA-3') an den 5'-Bereich intakter mRNA-Moleküle angehängt werden. Durch den Einsatz reverser genspezifischer Primer wurde durch reverse Transkription mit dem Enzym „SUPERSCRIPT II Reverse Transkriptase“ spezifisch die modifizierte mRNA in einzelsträngige cDNA umgeschrieben. In einem letzten Schritt konnte mit einem zweiten reversen genspezifischen Primer und einem zu dem synthetischen RNA-Oligo-Nukleotid komplementären sense-Primer (GeneRacer 5'-Primer oder GeneRacer 5'-Nested-Primer) die fehlende Sequenz durch PCR amplifiziert werden. Die Größe der erhaltenen PCR-Produkte wurden durch Agarosegelelektrophorese ermittelt und die DNA aus dem Gel extrahiert und anschließend sequenziert.

Für die PCR wurden folgende Primer benutzt (für Sequenz siehe 4.7.1.1):

Untransfizierte RNA-Präparation:

5'-PCR	GeneRacer 5' Primer / CH40
5'-Nested PCR	GeneRacer 5' Nested Primer / B2

pGL3-pcK30 transfizierte RNA-Präparation:

5'-PCR	GeneRacer 5' Primer / Mkr17
5'-Nested PCR	GeneRacer 5' Nested Primer / Mkr18

4.4.12 Radioaktive Markierung von DNA

Kinase:	T4 Polynukleotid Kinase 10 U/μl (NEB)
Puffer:	10x (NEB)
Radioaktivität:	Adenosin 5'-[α 32P]-triphosphat, Triethylammoniumsalz (spez. Aktivität: 3.000 Ci/mmol, 10 μCi/μl; Amersham)
Gelfiltration:	Nick Column, Pharmacia
Szintillationszähler:	LS 1800; Beckman
TE-Puffer:	10 mM Tris-HCl, pH 7,6 oder 8,0, 1 mM EDTA

Eine 10 μl Reaktion aus 3,5 pmol DNA (dephosphoryliert), 1 μl 10x T4-PNK-Puffer, 10 U T4 Polynukleotid Kinase und 10 μCi [α 32P]ATP wurde für 30 Minuten bei 37 °C inkubiert. Die markierte DNA wurde über eine Nick Column entsprechend den Herstellerangaben gereinigt.

Die Cerenkov-Strahlung jeder erhaltenen Fraktion wurde im Szintillationszähler bestimmt. Eine spezifische Aktivität von $1-2 \times 10^6$ cpm/pmol DNA wurde routinemäßig erhalten.

4.4.13 Methylierung von DNA

Methylase: Sss I Methylase (CpG Methylase) 4 U/ μ l (NEB)
 Puffer: 10x (NEB); 160 μ M S-adenosylmethionine (SAM)

Die CpG Methylase methyliert alle Cytosinreste (C⁵) innerhalb der Doppelstrang-Dinukleotid-Erkennungssequenz 5'-CG-3'.

Für eine 20 μ l Reaktion wurden pro μ g DNA 4 U Enzym eingesetzt und für 1 Stunde bei 37 °C inkubiert. Im Anschluss folgte eine Aufreinigung der DNA über ein präparatives Gel.

4.4.14 DNA-Sequenzierung

Die Sequenzierreaktion wurde nach Angaben des Herstellers (Applied Biosystems) durchgeführt. Es wurden jeweils 250 ng DNA als Template und 10 pmol Primer eingesetzt. Die Sequenzierung von DNA erfolgt nach dem Prinzip des durch Didesoxynukleotide hervorgerufenen Kettenabbruchs bei der enzymatischen Synthese von DNA-Doppelsträngen (Sanger et al. 1977). Zur Sequenzanalyse wurde der Reaktionsansatz, nach Aufreinigung mit „Dye Ex Spin Columns“ (Qiagen), auf ein hochauflösendes, denaturierendes Polyacryamidgel aufgetragen. Mittels eines Lasers können die unterschiedlich fluoreszierenden 3'-Enden der verschiedenen langen Nukleinsäureketten detektiert werden. Aus der Fluoreszenzabfolge kann die Sequenz des DNA-Fragments abgeleitet werden.

4.5 Protein-Methoden

4.5.1 Präparation von Protein-Extrakten aus eukaryotischen Zellen

PE-Puffer A: GH: 10 mM HEPES pH 7,9, 0,3 M Sucrose, 1,5 mM MgCl₂, 10 mM KCl, 0,1 mM EGTA, 0,5 mM DTT (DTT Zugabe unmittelbar vor Gebrauch), 1/25 Proteinaseinhibitoren, 50 mM NaCl, 0,2 % NP40
HeLa: 10 mM HEPES pH 7,9, 0,3 M Sucrose, 1,5 mM MgCl₂, 10 mM KCl, 0,1 mM EGTA, 0,5 mM DTT (DTT Zugabe unmittelbar vor Gebrauch), 1/25 Proteinaseinhibitoren

PE-Puffer B:	20 mM HEPES pH 7,9, 420 mM NaCl, 1,5 mM MgCl ₂ , 0,1 mM EGTA, 25 % Glycerin (v/v), 1/25 Proteinaseinhibitoren, 0,5 mM DTT (Proteinaseinhibitoren und DTT Zugabe unmittelbar vor Gebrauch)
PE-Puffer C:	300 mM HEPES pH 7,9, 1,4 M KCl, 30 mM MgCl ₂
PE-Puffer D:	20 mM HEPES pH 7,9, 75 mM NaCl, 0,1 mM EDTA, 20 % Glycerin (v/v), 1/25 Proteinaseinhibitoren, 0,5 mM DTT (Proteinaseinhibitoren und DTT Zugabe unmittelbar vor Gebrauch)
Proteinaseinhibitor:	complete Protease Inhibitor Cocktail Tablets (Roche)
Photometer:	Spectronic 1201 (Milton Roy)

Die Protein-Extrakte wurden modifiziert nach Dignam et al. (Dignam et al. 1983) hergestellt. Die in 80 cm² Flaschen (Nunc) herangezogenen Zellen wurden dreimal mit PBS auf Eis gewaschen und mit einem Zellschaber in 2x 5 ml PBS pro Kulturflaschen vom Flaschenboden mechanisch gelöst. Die Zellen wurden dann für 10 Minuten bei 1500 rpm (Minifuge RF), 4 °C zentrifugiert und im 5-fachen Zellvolumen PBS (4 °C) resuspendiert. Nach erneuter Zentrifugation wurden sie im 3 (GH) - 5 (HeLa)-fachen Zellvolumen PE-Puffer A (4 °C) resuspendiert. Unter mikroskopischer Kontrolle (!) wurden die Zellen dann im Glas-Homogenisator (Braun) mechanisch mit mindestens 10 (HeLa) - 20 (GH) Stempelhüben aufgeschlossen.

Die intakten Nuklei wurden für 10 Minuten bei 2500 rpm pelletiert.

Zur Präparation des zytoplasmatischen Protein-Extraktes wurde der Überstand mit 0,11 Volumen PE-Puffer C versetzt und für 1 Stunde bei 50.000 rpm (TL-100), 4 °C zentrifugiert. Dieser Überstand wurde über Nacht gegen das 20-fache Volumen PE-Puffer D bei 4 °C dialysiert. Der zytoplasmatische Protein-Extrakt (= Dialysat) wurde aliquotiert, mit Stickstoff schockgefroren und bei -70 °C gelagert.

Zur Präparation des nukleären Protein-Extraktes wurde das Nuklei-Pellet im 2,5-fachen Pelletvolumen PE-Puffer B aufgenommen und im Glas-Homogenisator mit mindestens 20 Stempelhüben mechanisch aufgeschlossen (mikroskopische Kontrolle!). Nach leichter Agitation auf einem Magnetrührer für 30 Minuten, 4 °C wurden die Nuklei Reste für 30 Minuten bei 50.000 rpm (TL-100), 4 °C abzentrifugiert. Der Überstand wurde über Nacht gegen das 50-fache Volumen PE-Puffer D bei 4 °C dialysiert. Das Dialysat wurde für 20

Minuten bei 14.000 rpm (Biofuge 15R), 4 °C zentrifugiert und der nukleäre Protein-Extrakt (= Überstand) wurde aliquotiert, mit Stickstoff schockgefroren und bei -70 °C gelagert.

Der Proteingehalt der zytoplasmatischen und nukleären Extrakte wurde mit dem BioRad Protein-Assay (BioRad) entsprechend den Herstellerangaben quantifiziert.

4.5.2 Aufreinigung rekombinanter Proteine aus eukaryotischen Zellen

Rollrad: Heidolph
 Proteinaseinhibitor: complete Protease Inhibitor Cocktail Tablets (Roche) ohne EDTA

Zur Proteinaufreinigung wurde das Gen des gewünschten Proteins in den pcDNA3.1/myc-His(-)-Vektor kloniert und in HeLa-Zellen transfiziert (siehe 4.2.5).

Zur Isolierung von rekombinantem Protein wurde das „ProBond Purification System“ von Invitrogen unter nativen Bedingungen mit folgenden Modifikationen verwendet: Es wurden Proteinaseinhibitoren ohne EDTA verwendet, 4 ml Zellysate wurden auf eine Säule gegeben, die 1 ml Slurry enthielt. Die Bindungsreaktion fand auf einem Rollrad bei 4 °C für 1 Stunde statt.

4.5.3 SDS-Polyacrylamid Gel Elektrophorese (SDS-PAGE)

Gelapparatur: Mighty small II SE 250 (Hoefer Scientific Instruments)
 Transformator: Electrophoresis Power Supply-EPS 301 (Amersham Pharmacia Biotech)
 DTT: 30x; Dithiothreitol (NEB)
 Glycingel-Laufpuffer: 1x; 25 mM Tris-HCl, 250 mM Glycin, 0,1 % SDS, pH 8,3
 Auftragspuffer: 3x SDS-Auftragspuffer (NEB)
 Marker: MultiMark Multi-Colored Standard (Novex)
 Glycingel: Acrylamid/Bisacrylamid (Roth)
 Tricinegel: 10-20 % (Invitrogen)
 Tricinegel-Laufpuffer: 1x; 100 mM Tris-Base, 100 mM Tricine, 0,1 % SDS, pH 8,3
 Tricinegel-Auftragspuffer: 1x; 450 mM Tris-HCl, 12 % Glycerin, 4 % SDS, 0,0025 % Coomassie Blue, 0,0025 % Phenolred, pH 8,45
 APS: 10 % Ammonium Persulfat (Sigma)
 TEMED: Tetramethylethyldiamin (SERVA)

Die diskontinuierliche SDS-PAGE wurde nach Laemmli (Laemmli 1970) durchgeführt:

Sammelgel: 5 % Acrylamid/Bisacrylamid (30:0,8), 0,125 M Tris-HCl pH 6,8, 0,1 % SDS;

Trenngel: 8-12,5 % Acrylamid/Bisacrylamid (30:0,8), 0,375 M Tris-HCl pH 8,8, 0,1 % SDS;

Polymerisation durch: 0,075 % APS und 0,0625 % TEMED; Gel-Größe: 200 x 160 x 1 mm.

Die Proteinproben (nukleäre GH-, HeLa-Zell-Extrakte) wurden in 1x Proben-Puffer für 3 Minuten bei 85 °C erhitzt und auf das vertikale Gel aufgetragen. Die Elektrophorese fand bei konstanter Stromstärke von 125 V, max. A mit 1x Laufpuffer statt.

4.5.4 Western-Transfer-Analyse

Blotkammer:	TE Series Transpher Electrophoresis Unit (Hoefer Scientific Instruments)
Blotmembran:	Protran; Nitrocellulose Transfer Membrane (Schleicher und Schuell)
Blotpuffer:	25 mM Tris-HCl, 192 mM Glycin, 10 % Methanol, pH 8,6
Filterpapier:	3 MM (Whatman)
Färbelösung:	Ponceau S (Sigma)
Blockierlösung:	6 % Magermilchpulver (Merck) in PBS mit 0,05 % Tween 20 (Sigma)
Magermilchlösung:	3 % Magermilchpulver (Merck) in PBS mit 0,05 % Tween 20 (Sigma)
Waschlösung:	0,05 % Tween 20 (Sigma) in PBS
1. Antikörper :	anti-Sp1: Maus (Santa Cruz) 1:1.000 in 3 % Magermilchlösung verdünnt. anti-Sp3: Kaninchen (Santa Cruz) 1:3.000 in 3 % Magermilchlösung verdünnt. anti-SRY: Kaninchen (Santa Cruz) 1:1.000 in 3 % Magermilchlösung verdünnt. anti-Actin (C-11): Ziege (Santa Cruz) 1:1.000 in 3 % Magermilchlösung verdünnt.
2. Antikörper :	anti-Kaninchen: Peroxidase gekoppelt (Sigma) 1:5.000 in 3 % Magermilchlösung verdünnt. anti-Ziege: Peroxidase gekoppelt (Sigma) 1:15.000 in 3 % Magermilchlösung verdünnt.

	anti-Maus: Peroxidase gekoppelt (Sigma)
	1:10.000 in 3 % Magermilchlösung verdünnt.
Detektion:	ECL plus Western Blotting Detection System (Amersham)
Röntgenfilme:	BioMax MS von Kodak

Die über SDS-PAGE aufgetrennten Proteine wurden elektrophoretisch in 1x Blot-Puffer bei konstanter Stromstärke von 50 V, max. A auf eine Protran Membran übertragen, die zuvor 15 Sekunden in Blot-Puffer äquilibriert wurde. Die anschließende Behandlung der Membran fand unter Schütteln bei RT statt. Die Membran wurde für 30 Minuten in 20 ml Blockierlösung inkubiert. Anschließend wurde die Membran für eine Stunde mit dem 1. Antikörper auf einem Wippschüttler inkubiert, 3x kurz gewaschen und mit dem 2. Antikörper für eine weitere Stunde inkubiert. Nach erneutem Waschen wurde die Membran mit dem "ECL plus Western Blotting Detection System" (Amersham) nach Herstellerangaben entwickelt.

4.5.5 Dualer Luziferase-Assay

Luminometer:	MicroLumat Plus (EG&G Berthold)
Reporter-Assay:	Dual Luciferase Assay (Promega)
96-Loch-Platten (weiß):	LumiNunc Platte (Nalge Nunc International)

Mit rekombinantem pGL3-Plasmid (codiert für die Firefly Luziferase) und dem Co-Reporter Vektor phRL-null (codiert für die Renilla Luziferase) transfizierte eukaryotische Zellen (s. 4.2.5) wurden ca. 27 Stunden nach der Transfektion mit 20 µl Passive-Lysis-Puffer (Promega) für 20 Minuten lysiert. Die Messung erfolgte im Luminometer. Inkubation mit 50 µl LARII (Promega) für 2 Sekunden und Messung für 10 Sekunden ergaben den Wert für die Firefly Luziferase, Zugabe von 50 µl Stop&Glow Reagent (Promega; 6 Sekunden Inkubation, 10 Sekunden Messung) ergaben den Wert für die cotransfizierte Renilla Luziferase. Der Quotient aus beiden Werten ergab einen normalisierten Wert, der Fehlerquellen, wie unterschiedliche Zellzahlen oder unterschiedliche Transfektionseffizienzen, minimiert.

4.6 Electrophoretic Mobility Shift Assay (EMSA) und Supershift-Assay

4.6.1 Protein-Nukleinsäure Bindungs-Reaktion

EMSA-Puffer, 10x:	100 mM Tris-HCl pH 7,5, 500 mM NaCl, 10 mM MgCl ₂ , 5 mM EDTA, 5 mM DTT, 40 % Glycerin, 0,5 µg/µl Poly(dI:dC), 5 mM ZnCl ₂
Gel-Loading-Buffer:	10x; 250 mM Tris-HCl pH 7,5, 0,2 % Bromphenolblau, 0,2 % Xylen Cyanol, 40 % Glycerin
Rekombinante Proteine:	humanes Sp1 und Sp3 (Alexis Biochemicals)

Eine 20 µl Reaktion bestand aus 8-20 fmol Sonden-DNA (10.000-20.000 cpm), 1x EMSA-Puffer sowie 2-5 µg nukleärem Protein-Extrakt bzw. 10 ng rekombinantem Protein und wurde für 30 Minuten bei RT inkubiert. Kompetitive EMSA-Ansätze enthielten zusätzlich unmarkierte spezifische (identisch oder variiert zur markierten Sonde) Kompetitor-DNA in variablen Mengen, meist 50-100x molarer Überschuss. Diese wurden vor Zugabe der Sonden-DNA 20 Minuten bei RT mit den nukleären Extrakten vorinkubiert. Die Reaktion wurde mit 2 µl 10x Gel Loading Buffer gestoppt und auf ein natives Polyacrylamid-Gel aufgetragen (Fried und Crothers 1981; Garner und Revzin 1981; Lane et al. 1992).

Für Supershift-Assays wurde die Reaktion nach Zugabe der markierten Sonden-DNA mit einem Antiserum (z.B. 2 µg anti-Sp1 (Santa Cruz Biotechnology)) versetzt, 20 Minuten bei RT inkubiert und mit einem nativen 4 % Acrylamid Gel analysiert.

4.6.2 Native Polyacrylamid Gel Elektrophorese

Geltrockner:	Gel Dryer Model 583 (Biorad)
Kühlgerät für	
Gelapparaturen:	MGW LAUDA RMT / RM6
TBE-Puffer, 1x:	90 mM Tris-Borat pH 7,2, 2 mM EDTA

Die vertikale Polyacrylamid Gel Elektrophorese wurde mit dem Protean II xi System (BioRad) durchgeführt. Ein 5 % Polyacrylamid Gel (200 x 160 x 1 mm; 5 % Acrylamid, 0,26 % Bisacrylamid [Gel 40, 19:1; Roth], 0,5x TBE, 2,5 % Glycerin; Polymerisation durch:

0,075 % APS und 0,0625 % TEMED) wurde vor dem Probenauftrag etwa 20 Minuten bei 15 V/cm mit 0,5x TBE als Laufpuffer vorlaufen gelassen. Die Elektrophorese fand bei 10 V/cm und 4 °C statt. Das Gel wurde auf Filterpapier (3 MM, Whatman) überführt, bei 80 °C für mindestens 45 Minuten getrocknet und bei -70 °C über Nacht autoradiographiert (Biomax MR-Film; Kodak).

4.7 Rekombinante Plasmide

4.7.1 Klonierungsstrategien

Als Vektoren wurden die Plasmide pBluescript II KS- (Stratagene), pGL3 (Promega) und pcDNA3.1 verwendet. Die HERV-K LTRs pcK30 und LTR21 wurden aus dem pSVOAL-Vektor über die Schnittstelle HindIII entnommen (Hasenmaier 1995; Thelen 1995). Die verwendeten pHIVgagΔU3RIIIHyRcRE-Plasmide stammen von Ch. Magin (Magin 1998). Das Expressionsplasmid pcDNA3.1-hSRY wurde freundlicherweise von Xin Yuan zur Verfügung gestellt (Yuan et al. 2001).

Die zur Insertion verwendeten DNA-Fragmente wurden entweder durch PCR oder durch Restriktionsverdau aus HERV-K LTR-Plasmiden und anschließender LMP Agarose Gel Aufreinigung der gewünschten verdauten PCR-Produkte oder Restriktionsfragmente hergestellt. Diese Insertionsfragmente wurden in, mit geeigneten Restriktionsenzymen (NEB) restringierte, 5' dephosphorylierte (Calf Intestinal Phosphatase; Promega) Vektoren ligiert. Richtige Konstrukte wurden nach Plasmid Mini Präparation transformierter E.coli DH5α Bakterien durch Restriktionsfragmentlängenanalyse und Sequenzierung identifiziert, durch Plasmid Maxi Präparation isoliert und durch Sequenzierung erneut bestätigt.

4.7.1.1 Verwendete Oligonukleotide

MKr4: 5'-GAGATAGGGAAAAACCGCCTTAGGGCTGGAGG-3'

MKr9: 5'-CATATGGAGGATCCCACCAGCCTC-3'

MKr10: 5'-CATATGGAGGATCCCGCCAGCCTGTCAGTCCCCTTAG-3'

MKr11: 5'-GGAATGTCTCGGTTTTTTTCCCGATTGTATGC-3'

MKr12: 5'-GCATACAATCGGGAAAAAAACCGAGACATTCC-3'

MKr17: 5'-CGGGCGCAACTGCAACTCCGATAAATAACG-3'

MKr18: 5'-GGAACCAGGGCGTATCTCTTCATAGCCTTATGC-3'

MKr21: 5'-AAGCTTTTATTGATC-3'

MKr27: 5'-AAGCTTCTTGTGACC-3'

MKr43: 5'-TCTTCGAGAAACACCCACAGATG-3'
MKr44: 5'-CATCTGTGGGTGTTTCTCGAAGA-3'
MKr47: 5'-CTTGTGACCCTCACACATCCCCCTCT-3'
MKr48: 5'-AGAGGGGGATGTGTCAGGGTCACAAG-3'
MKr51: 5'-CTTGTGACCCTGACACATCCCCCTCTTCGAGAAAC-3'
MKr52: 5'-GTTTCTCGAAGAGGGGGATGTGTCAGGGTCACAAG-3'
MKr53: 5'-GACACATCCCCCTCTTCGAGAAAC-3'
MKr54: 5'-GTTTCTCGAAGAGGGGGATGTGTC-3'
MKr55: 5'-ACATCCCCCTCTTCGAGAAACACCCACAGATG-3'
MKr56: 5'-CATCTGTGGGTGTTTCTCGAAGAGGGGGATGT-3'
MKr57: 5'-ATGATCAATAAATACTAAGGGA ACTCAGAGG-3'
MKr58: 5'-AGAGGGGGATGTGTCAGGG-3'
MKr59: 5'-CCCACGAATGATCAATAAATACTAAGGG-3'
MKr60: 5'-TGTTTCTCCTAGAGGGGGATGTGTCAG-3'
MKr62: 5'-CAGGGTCACAAGACAATTGTGGGGAGAGGG-3'
MKr63: 5'-GGAGATGTGTCAGGGTCACAAGACAATTGTGG-3'
MKr65: 5'-CTCTATGATCAATAAATACTAAGGGA ACTCAGAGGC-3'
B2: 5'-AGCTCTCTAGAAAGCTTCCTCCACGTTGGGCACCA-3'
CH40: 5'-TCTTCCGAGCGCGCAAGCTTACCG-3'
GeneRacer 5' Primer: 5'-CGACTGGAGCACGAGGACACTGA-3'
GeneRacer 5' Nested Primer: 5'-GGACACTGACATGGACTGAAGGAGTA-3'

4.7.1.2 Konstruktion der rekombinanten Plasmide

pGL3-pcK30

Die pcK30 LTR wurde über HindIII aus dem pSVOAL Vektor (Hasenmaier 1995) geschnitten und in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21

Die LTR21 wurde über HindIII aus dem pSVOAL Vektor (Hasenmaier 1995) geschnitten und in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pBS-pcK30

Die pcK30 LTR wurde über HindIII aus dem pSVOAL Vektor (Hasenmaier 1995) geschnitten und in den ebenso restringierten 2961 bp großen pBS II KS- Vektor ligiert.

pBS-LTR21

Die LTR21 wurde über HindIII aus dem pSVOAL Vektor (Hasenmaier 1995) geschnitten und in den ebenso restringierten 2961 bp großen pBS II KS- Vektor ligiert.

pGL3-pcK30dG

pBS-pcK30 wurde mit KpnI und SmaI geschnitten und das Insert in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dH

pBS-pcK30 wurde mit KpnI und BamHI geschnitten, das Insert blunt geendet und in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dI

pBS-pcK30 wurde mit KpnI und BclI geschnitten, das Insert blunt geendet und in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dK

pBS-pcK30 wurde mit KpnI und NsiI geschnitten, das Insert blunt geendet und in den mit SmaI restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dA

pBS-pcK30 wurde mit Bsu36I und EcoRI geschnitten, das Insert entfernt und der Vektor mit den restlichen LTR-Sequenzen wurde blunt geendet, religiert und über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dT

pGL3-pcK30dT wurde mit den Primern MKr11 und MKr12 und pGL-pcK30 als Template unter Verwendung des „QuickChange Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGEM-T-pcK30-1

Das auf pGL-pcK30 mit den Primern MKr21 und MKr27 amplifizierte PCR Produkt wurde nach HindIII Restriktion in den ebenso restringierten 3015 bp langen pGEM-T Easy Vektor ligiert. Zur Sondengenerierung wurde das Plasmid mit HindIII restringiert, das Insert dephosphoryliert und radioaktiv kinasiert.

pGEM-T-LTR21-1

Das auf pGL-LTR21 mit den Primern MKr21 und MKr27 amplifizierte PCR Produkt wurde nach HindIII Restriktion in den ebenso restringierten 3015 bp langen pGEM-T Easy Vektor ligiert. Zur Sondengenerierung wurde das Plasmid mit HindIII restringiert, das Insert dephosphoryliert und radioaktiv kinasiert.

Subfragmente zur Konkurrenz:

30-2

Hybridisierung der Primer MKr47 und MKr48.

30-3

Hybridisierung der Primer MKr51 und MKr52.

30-4

Hybridisierung der Primer MKr53 und MKr54.

30-5

Hybridisierung der Primer MKr55 und MKr56.

30-6

Hybridisierung der Primer MKr43 und MKr44.

pGL3-pcK30dZ I

pGL3-pcK30dZ I wurde mit den Primern MKr61 und MKr63 und pGL-pcK30 als Template unter Verwendung des „ExSite PCR-Based Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGL3-pcK30dZ II

pGL3-pcK30dZ II wurde mit den Primern MKr59 und MKr60 und pGL-pcK30 als Template unter Verwendung des „ExSite PCR-Based Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGL3-pcK30dZ III

pGL3-pcK30dZ III wurde mit den Primern MKr61 und MKr62 und pGL-pcK30 als Template unter Verwendung des „ExSite PCR-Based Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGL3-pcK30dZ IV

pGL3-pcK30dZ IV wurde mit den Primern MKr57 und MKr58 und pGL-pcK30 als Template unter Verwendung des „ExSite PCR-Based Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGL3-pcK30dZ V

pGL3-pcK30dZ V wurde mit den Primern MKr62 und MKr65 und pGL-pcK30dZ IV als Template unter Verwendung des „ExSite PCR-Based Site-Directed Mutagenesis Kit“ (Stratagene) hergestellt.

pGL3-pcK30d21

pBS-pcK30 und pBS-LTR21 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21d30

pBS-pcK30 und pBS-LTR21 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dd21

pBS-pcK30 und pBS-LTR21 wurden jeweils mit KpnI und SphI restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dd30

pBS-pcK30 und pBS-LTR21 wurden jeweils mit KpnI und SphI restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dB

pBS-pcK30 wurde mit HincII und SphI restringiert, blunt geendet und religiert. Über KpnI und SacI wurde das Insert in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dB

pBS-LTR21 wurde mit HincII und SphI restringiert, blunt geendet und religiert. Über KpnI und SacI wurde das Insert in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30d

pBS-pcK30 wurde mit SphI und Bsu36I restringiert, blunt geendet und religiert. Über KpnI und SacI wurde das Insert in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21d

pBS-LTR21 wurde mit SphI und Bsu36I restringiert, blunt geendet und religiert. Über KpnI und SacI wurde das Insert in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dQ

pBS-pcK30 und pHIVgag Δ U3RIIIHyRcRE6 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pHIVgag Δ U3RIIIHyRcRE6 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dP

pBS-pcK30 und pHIVgag Δ U3RIIIHyRcRE1 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pHIVgag Δ U3RIIIHyRcRE1 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dHy

pBS-pcK30 und pHIVgag Δ U3RIIIHyRcRE15 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pHIVgag Δ U3RIIIHyRcRE15 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dR

pBS-pcK30 und pHIVgag Δ U3RIIIHyRcRE7 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pHIVgag Δ U3RIIIHyRcRE7 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dS

pBS-pcK30 und pHIVgag Δ U3RIIIHyRcRE10 wurden jeweils mit SphI und Bsu36I restringiert. Das Insert aus pHIVgag Δ U3RIIIHyRcRE10 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dM

pBS-pcK30 und pBS-LTR21 wurden jeweils mit BclI und EcoRI restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dM

pBS-pcK30 und pBS-LTR21 wurden jeweils mit BclI und EcoRI restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dL

pBS-pcK30 und pBS-LTR21 wurden jeweils mit NdeI und EcoRI restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dL

pBS-pcK30 und pBS-LTR21 wurden jeweils mit NdeI und EcoRI restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21 ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dO

pBS-pcK30dL und pBS-LTR21 wurden mit SmaI restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30dL ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dO

pBS-LTR21dL und pBS-pcK30 wurden mit SmaI restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21dL ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dN

pBS-pcK30dM und pBS-LTR21 wurden jeweils mit NdeI und EcoRI restringiert. Das Insert aus pBS-LTR21 wurde in den restringierten pBS-pcK30dM ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-LTR21dN

pBS-LTR21dM und pBS-pcK30 wurden jeweils mit NdeI und EcoRI restringiert. Das Insert aus pBS-pcK30 wurde in den restringierten pBS-LTR21dM ligiert. Anschließend wurde die LTR über HindIII in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dU

Das auf pBS-pcK30 mit den Primern MKr4 und MKr10 amplifizierte PCR Produkt wurde nach HindIII Restriktion in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

pGL3-pcK30dV

Das auf pBS-pcK30 mit den Primern MKr4 und MKr9 amplifizierte PCR Produkt wurde nach HindIII Restriktion in den ebenso restringierten 4818 bp großen pGL3-Basic Vektor ligiert.

5 Zusammenfassung

Das humane Genom besteht zu etwa 8% aus retroviralen Sequenzen. Davon sind ca. 1-2% dem humanen endogenen Retrovirus K (HERV-K) zuzuordnen. Das Virus ist mit ca. 30-50 Proviren und ca. 10.000 sLTRs im humanen Genom vertreten. HERV-K besitzt intakte ORFs für alle retroviralen Proteine und zusätzlich ein ORF für das akzessorische Protein Rec. Obwohl eine basale Transkription in verschiedenen Geweben nachgewiesen werden konnte, ist eine Expression von HERV-K Proteinen und Viruspartikeln nur in Keimzelltumoren nachgewiesen worden.

In dieser Arbeit wurde die transkriptionelle Regulation des gewebespezifischen Promotors von HERV-K näher charakterisiert. Hierbei wurden regulatorisch wichtige Sequenzen mit Hilfe des Luziferase-Assays eingegrenzt. Transkriptionell sensitive Regionen wurden daraufhin im EMSA auf die Bindung von Transkriptionsfaktoren untersucht.

Durch transiente Transfektionen von Luziferasereporterkonstrukten in verschiedenen Zelllinien stellte sich heraus, dass HERV-K LTRs in Keimzelltumorzellen aktiv sind, dass es aber auch inaktive LTRs gibt, die im Zuge der Evolution durch Punktmutationen ihre transkriptionelle Aktivität verloren haben. Aktive und inaktive LTRs unterscheiden sich durch Punktmutationen, die sich in verschiedenen Sequenzabschnitten häufen. Chimäre Konstrukte aus aktiven und inaktiven Sequenzabschnitten und gezielte Mutationen und Deletionen in aktiven HERV-K LTRs enthüllten verschiedene DNA-Bereiche, die transkriptionelle Sensitivität aufweisen. Die LTR-Bereiche bps 572-578, bps 757-798 und bps 809-823 waren im Aktivitäts-Assay besonders empfindlich gegenüber Mutation. In diesen Bereichen liegen zum einen GC/GT-Boxen, also Konsensussequenzen für die Transkriptionsfaktoren Sp1 und Sp3, zum anderen Konsensussequenzen für ein Inr und ein DPE. Das Binden von Sp1 und Sp3 auf den GC/GT-Boxen der bps 757-798 konnte durch eine EMSA-Analyse bewiesen werden. Der transkriptionell sensitive Bereich der bps 572-578 weist eine weitere putative GC/GT-Box auf.

Die TATA-Box bei bp 532 zeigte sich unempfindlich gegenüber Mutation. Eine Beteiligung dieser Konsensussequenz am basalen Promotor wurde deshalb ausgeschlossen. Unterstützt durch die Ergebnisse einer 5'-RACE, ein Verfahren, das den Transkriptionsstart eines bestimmten Gens bestimmen kann, konnte gezeigt werden, dass der basale HERV-K Promotor aus Konsensussequenzen für die Transkriptionsfaktoren Sp1 und Sp3, einem Inr und einem DPE besteht. Sp1 und Sp3 sind ubiquitäre Transkriptionsfaktoren, die durch ihr Mengenverhältnis in einem bestimmten Gewebe oder durch posttranslationale Modifikationen unterschiedliche Auswirkungen auf die Transkription haben können. Diese Eigenschaften

könnten zur Gewebespezifität von HERV-K beitragen. Die Core Promotor Komponenten Inr und DPE sind für die korrekte Positionierung von TFIID und damit der PolII verantwortlich. Des weiteren wurde eine starke Beteiligung des Testis-spezifischen Transkriptionsfaktors SRY an der Transkription von HERV-K belegt. Im Luziferase-Assay konnte ein SRY-Expressionsplasmid die Transkription eines aktiven LTRs in GH- wie auch in HeLa-Zellen steigern. Ein Indiz dafür, dass es sich um einen für die Transkription von HERV-K essentiellen Faktor handelt. Verschiedene SRY-Konsensussequenzen auf der HERV-K LTR machen eine Wirkung in *cis* wahrscheinlich, doch könnte auch in *trans* ein für die Transkription wichtiger Transkriptionsfaktor in seiner Expression verstärkt werden. Ebenfalls konnte bewiesen werden, dass epigenetische Mechanismen eine starke Rolle bei der Transkription von HERV-K spielen. Die Methylierung eines LTR-haltigen Luziferase-Reportervektors führte zum fast vollständigen Verlust der transkriptionellen Aktivität. Es bleibt abzuwarten, in welcher Weise die Schlüsselkomponenten Sp1, Sp3, SRY und Hypermethylierung des HERV-K Genoms zur Gewebespezifität von HERV-K beitragen.

6 Literaturverzeichnis

- Armbrüster, V., Sauter, M., Krautkrämer, E., Meese, E., Kleiman, A., Best, B., Roemer, K., Mueller-Lantzsch, N. (2002).** "A novel gene from the human endogenous retrovirus K expressed in transformed cells." Clin Cancer Res **8**(6): 1800-7.
- Baltimore, D. (1970).** "RNA-dependent DNA polymerase in virions of RNA tumour viruses." Nature **226**(252): 1209-11.
- Barbulescu, M., Turner, G., Seaman, M. I., Deinard, A. S., Kidd, K. K., Lenz, J. (1999).** "Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans." Curr Biol **9**(16): 861-8.
- Berta, P., Hawkins, J. R., Sinclair, A. H., Taylor, A., Griffiths, B. L., Goodfellow, P. N., Fellous, M. (1990).** "Genetic evidence equating SRY and the testis-determining factor." Nature **348**(6300): 448-50.
- Birnboim, H. C. (1983).** "A rapid alkaline extraction method for the isolation of plasmid DNA." Methods Enzymol **100**: 243-55.
- Birnboim, H. C., Doly, J. (1979).** "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." Nucleic Acids Res **7**(6): 1513-23.
- Bittner, J. J. (1936).** "Some possible effects of nursing on the mammary gland tumour incidence in mice." Science **84**: 162.
- Boller, K., Janssen, O., Schuldes, H., Tönjes, R. R., Kurth, R. (1997).** "Characterization of the antibody response specific for the human endogenous retrovirus HTDV/HERV-K." J Virol **71**(6): 4581-8.
- Boller, K., König, H., Sauter, M., Mueller-Lantzsch, N., Löwer, R., Löwer, J., Kurth, R. (1993).** "Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV." Virology **196**(1): 349-53.
- Bouwman, P., Philipsen, S. (2002).** "Regulation of the activity of Sp1-related transcription factors." Mol Cell Endocrinol **195**(1-2): 27-38.
- Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., Cedar, H. (1994).** "Sp1 elements protect a CpG island from de novo methylation." Nature **371**(6496): 435-8.
- Braun, H., Koop, R., Ertmer, A., Nacht, S., Suske, G. (2001).** "Transcription factor Sp3 is regulated by acetylation." Nucleic Acids Res **29**(24): 4994-5000.
- Breathnach, R., Chambon, P. (1981).** "Organization and expression of eucaryotic split genes coding for proteins." Annu Rev Biochem **50**: 349-83.
- Briggs, M. R., Kadonaga, J. T., Bell, S. P., Tjian, R. (1986).** "Purification and biochemical characterization of the promoter-specific transcription factor, Sp1." Science **234**(4772): 47-52.

- Brown, S. W. (1966).** "Heterochromatin." Science **151**(709): 417-25.
- Burke, T. W., Kadonaga, J. T. (1996).** "Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters." Genes Dev **10**(6): 711-24.
- Burke, T. W., Kadonaga, J. T. (1997).** "The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila." Genes Dev **11**(22): 3020-31.
- Burke, T. W., Willy, P. J., Kutach, A. K., Butler, J. E., Kadonaga, J. T. (1998).** "The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box." Cold Spring Harb Symp Quant Biol **63**: 75-82.
- Bushnell, D. A., Cramer, P., Kornberg, R. D. (2002).** "Structural basis of transcription: alpha-amanitin-RNA polymerase II cocystal at 2.8 A resolution." Proc Natl Acad Sci U S A **99**(3): 1218-22.
- Casau, A. E., Vaughan, J. E., Lozano, G., Levine, A. J. (1999).** "Germ cell expression of an isolated human endogenous retroviral long terminal repeat of the HERV-K/HTDV family in transgenic mice." J Virol **73**(12): 9976-83.
- Cherrington, J., Ganem, D. (1992).** "Regulation of polyadenylation in human immunodeficiency virus (HIV): contributions of promoter proximity and upstream sequences." Embo J **11**(4): 1513-24.
- Courey, A. J., Holtzman, D. A., Jackson, S. P., Tjian, R. (1989).** "Synergistic activation by the glutamine-rich domains of human transcription factor Sp1." Cell **59**(5): 827-36.
- Courey, A. J., Tjian, R. (1988).** "Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif." Cell **55**(5): 887-98.
- Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E., Burgess, R. R., Edwards, A. M., David, P. R., Kornberg, R. D. (2000).** "Architecture of RNA polymerase II and implications for the transcription mechanism." Science **288**(5466): 640-9.
- Cramer, P., Bushnell, D. A., Kornberg, R. D. (2001).** "Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution." Science **292**(5523): 1863-76.
- Cullen, B. R. (1992).** "Mechanism of action of regulatory proteins encoded by complex retroviruses." Microbiol Rev **56**(3): 375-94.
- Deen, K. C., Sweet, R. W. (1986).** "Murine mammary tumor virus pol-related sequences in human DNA: characterization and sequence comparison with the complete murine mammary tumor virus pol gene." J Virol **57**(2): 422-32.
- Denner, J., Phelps, R. C., Löwer, J., Löwer, R., Kurth, R. (1995).** J. Cancer Res. Clin. Oncol. **121**(Suppl. 1, 5).

- Dennig, J., Beato, M., Suske, G. (1996).** "An inhibitor domain in Sp3 regulates its glutamine-rich activation domains." Embo J **15**(20): 5659-67.
- Desclozeaux, M., Poulat, F., de Santa Barbara, P., Soullier, S., Jay, P., Berta, P., Boizet-Bonhoure, B. (1998).** "Characterization of two Sp1 binding sites of the human sex determining SRY promoter." Biochim Biophys Acta **1397**(3): 247-52.
- Dignam, J. D., Lebovitz, R. M., Roeder, R. G. (1983).** "Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei." Nucleic Acids Res **11**(5): 1475-89.
- Domansky, A. N., Kopantzev, E. P., Snezhkov, E. V., Lebedev, Y. B., Leib-Mosch, C., Sverdlov, E. D. (2000).** "Solitary HERV-K LTRs possess bi-directional promoter activity and contain a negative regulatory element in the U5 region." FEBS Lett **472**(2-3): 191-5.
- Dynan, W. S., Saffer, J. D., Lee, W. S., Tjian, R. (1985).** "Transcription factor Sp1 recognizes promoter sequences from the monkey genome that are simian virus 40 promoter." Proc Natl Acad Sci U S A **82**(15): 4915-9.
- Dynan, W. S., Tjian, R. (1983a).** "Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II." Cell **32**(3): 669-80.
- Dynan, W. S., Tjian, R. (1983b).** "The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter." Cell **35**(1): 79-87.
- Ellermann, V., Bang, O. (1908).** "Experimentelle Leukämie bei Hühnern." Zentralbl Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. Orig. **46**: 595-609.
- Felsenstein, J. (1993).** PHYLIP (Phylogeny Inference Package) version 3.5c., Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Florl, A. R., Lower, R., Schmitz-Drager, B. J., Schulz, W. A. (1999).** "DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas." Br J Cancer **80**(9): 1312-21.
- Fried, M., Crothers, D. M. (1981).** "Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis." Nucleic Acids Res **9**(23): 6505-25.
- Frohman, M. A., Dush, M. K., Martin, G. R. (1988).** "Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer." Proc Natl Acad Sci U S A **85**(23): 8998-9002.
- Furuichi, Y., Shatkin, A. J. (2000).** "Viral and cellular mRNA capping: past and prospects." Adv Virus Res **55**: 135-84.
- Gallo, R. C. (1988).** "HIV--the cause of AIDS: an overview on its biology, mechanisms of disease induction, and our attempts to control it." J Acquir Immune Defic Syndr **1**(6): 521-35.
- Galvagni, F., Capo, S., Oliviero, S. (2001).** "Sp1 and Sp3 physically interact and co-operate with GABP for the activation of the utrophin promoter." J Mol Biol **306**(5): 985-96.

- Garner, M. M., Revzin, A. (1981).** "A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system." Nucleic Acids Res **9**(13): 3047-60.
- Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., Kornberg, R. D. (2001).** "Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution." Science **292**(5523): 1876-82.
- Goodfellow, P. N., Lovell-Badge, R. (1993).** "SRY and sex determination in mammals." Annu Rev Genet **27**: 71-92.
- Goodson, M. L., Hong, Y., Rogers, R., Matunis, M. J., Park-Sarge, O. K., Sarge, K. D. (2001).** "Sumo-1 modification regulates the DNA binding activity of heat shock transcription factor 2, a promyelocytic leukemia nuclear body associated transcription factor." J Biol Chem **276**(21): 18513-8.
- Graveley, B. R. (2001).** "Alternative splicing: increasing diversity in the proteomic world." Trends Genet **17**(2): 100-7.
- Grimes, S. R., Wilkerson, D. C., Noss, K. R., Wolfe, S. A. (2003).** "Transcriptional control of the testis-specific histone H1t gene." Gene **304**: 13-21.
- Günzburg, W. H., Heinemann, F., Wintersperger, S., Miethke, T., Wagner, H., Erfle, V., Salmons, B. (1993).** "Endogenous superantigen expression controlled by a novel promoter in the MMTV long terminal repeat." Nature **364**(6433): 154-8.
- Günzburg, W. H., Salmons, B. (1992).** "Factors controlling the expression of mouse mammary tumour virus." Biochem J **283** (Pt 3): 625-32.
- Hagen, G., Muller, S., Beato, M., Suske, G. (1992).** "Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes." Nucleic Acids Res **20**(21): 5519-25.
- Hagen, G., Muller, S., Beato, M., Suske, G. (1994).** "Sp1-mediated transcriptional activation is repressed by Sp3." Embo J **13**(16): 3843-51.
- Hanahan, D. (1983).** "Studies on transformation of Escherichia coli with plasmids." J Mol Biol **166**(4): 557-80.
- Harrich, D., Garcia, J., Wu, F., Mitsuyasu, R., Gonazalez, J., Gaynor, R. (1989).** "Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat." J Virol **63**(6): 2585-91.
- Harris, J. R. (1998).** "Placental endogenous retrovirus (ERV): structural, functional, and evolutionary significance." Bioessays **20**(4): 307-16.
- Hasenmaier, B. (1995).** Untersuchung zur Regulation der Expression der humanen endogenen retroviralen Sequenz HERV-K durch zelluläre und virale Faktoren. Frankfurt, Johann Wolfgang Goethe-Universität.

- Hata, Y., Duh, E., Zhang, K., Robinson, G. S., Aiello, L. P. (1998).** "Transcription factors Sp1 and Sp3 alter vascular endothelial growth factor receptor expression through a novel recognition sequence." J Biol Chem **273**(30): 19294-303.
- Hong, Y., Rogers, R., Matunis, M. J., Mayhew, C. N., Goodson, M. L., Park-Sarge, O. K., Sarge, K. D., Goodson, M. (2001).** "Regulation of heat shock transcription factor 1 by stress-induced SUMO-1 modification." J Biol Chem **276**(43): 40263-7.
- Hung, F. (1994).** The Retroviridae, Volume 3. Levy, J. A., Plenum, New York: 313-362.
- Ihn, H., Trojanowska, M. (1997).** "Sp3 is a transcriptional activator of the human alpha2(I) collagen gene." Nucleic Acids Res **25**(18): 3712-7.
- Iwasaki, K., Temin, H. M. (1990).** "The efficiency of RNA 3'-end formation is determined by the distance between the cap site and the poly(A) site in spleen necrosis virus." Genes Dev **4**(12B): 2299-307.
- Jackson, S. P., MacDonald, J. J., Lees-Miller, S., Tjian, R. (1990).** "GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase." Cell **63**(1): 155-65.
- Jager, R. J., Anvret, M., Hall, K., Scherer, G. (1990).** "A human XY female with a frame shift mutation in the candidate testis-determining gene SRY." Nature **348**(6300): 452-4.
- Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., Smale, S. T. (1994).** "DNA sequence requirements for transcriptional initiator activity in mammalian cells." Mol Cell Biol **14**(1): 116-27.
- Johnson, W. E., Coffin, J. M. (1999).** "Constructing primate phylogenies from ancient retrovirus sequences." Proc Natl Acad Sci U S A **96**(18): 10254-60.
- Jones, K. A., Kadonaga, J. T., Luciw, P. A., Tjian, R. (1986).** "Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1." Science **232**(4751): 755-9.
- Jones, K. A., Peterlin, B. M. (1994).** "Control of RNA initiation and elongation at the HIV-1 promoter." Annu Rev Biochem **63**: 717-43.
- Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., Wolffe, A. P. (1998).** "Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription." Nat Genet **19**(2): 187-91.
- Kaczynski, J., Cook, T., Urrutia, R. (2003).** "Sp1- and Kruppel-like transcription factors." Genome Biol **4**(2): 206.
- Kadonaga, J. T., Jones, K. A., Tjian, R. (1986).** "Promoter-specific activation of RNA polymerase II transcription by Sp1." Trends Biochem. Sci. **11**: 20-23.
- Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T., Muller, R. (1998).** "CIF150, a human cofactor for transcription factor IID-dependent initiator function." Mol Cell Biol **18**(1): 233-9.

- Kaufmann, J., Smale, S. T. (1994).** "Direct recognition of initiator elements by a component of the transcription factor IID complex." Genes Dev **8**(7): 821-9.
- Kennett, S. B., Udvardi, A. J., Horowitz, J. M. (1997).** "Sp3 encodes multiple proteins that differ in their capacity to stimulate or repress transcription." Nucleic Acids Res **25**(15): 3110-7.
- Kiela, P. R., LeSueur, J., Collins, J. F., Ghishan, F. K. (2002).** "Transcriptional regulation of the Rat NHE3 Gene: Functional interactions between GATA-5 and Sp family transcription factors." J Biol Chem.
- Kingsley, C., Winoto, A. (1992).** "Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression." Mol Cell Biol **12**(10): 4251-61.
- Knöbl, M. (1998).** Charakterisierung zellspezifischer Faktoren in der Regulation der terminalen Steuereinheiten (LTR) des humanen endogenen Retrovirus HERV-K. Tübingen, Eberhard-Karls-Universität.
- Knöbl, M., Löwer, R., Löwer, J. (1999).** "Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1." J Virol **73**(2): 1254-61.
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., Lovell-Badge, R. (1991).** "Male development of chromosomally female mice transgenic for Sry." Nature **351**(6322): 117-21.
- Kutach, A. K., Kadonaga, J. T. (2000).** "The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters." Mol Cell Biol **20**(13): 4754-64.
- La Mantia, G., Maglione, D., Pengue, G., Di Cristofano, A., Simeone, A., Lanfrancone, L., Lania, L. (1991).** "Identification and characterization of novel human endogenous retroviral sequences preferentially expressed in undifferentiated embryonal carcinoma cells." Nucleic Acids Res **19**(7): 1513-20.
- La Mantia, G., Majello, B., Di Cristofano, A., Strazzullo, M., Minchiotti, G., Lania, L. (1992).** "Identification of regulatory elements within the minimal promoter region of the human endogenous ERV9 proviruses: accurate transcription initiation is controlled by an Inr-like element." Nucleic Acids Res **20**(16): 4129-36.
- Laemmli, U. K. (1970).** "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(259): 680-5.
- Lander, E. S., Linton, L. M., et al. (2001).** "Initial sequencing and analysis of the human genome." Nature **409**(6822): 860-921.
- Lane, D., Prentki, P., Chandler, M. (1992).** "Use of gel retardation to analyze protein-nucleic acid interactions." Microbiol Rev **56**(4): 509-28.

- Lania, L., Majello, B., De Luca, P. (1997).** "Transcriptional regulation by the Sp family proteins." Int J Biochem Cell Biol **29**(12): 1313-23.
- Li, R., Knight, J. D., Jackson, S. P., Tjian, R., Botchan, M. R. (1991).** "Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription." Cell **65**(3): 493-505.
- Lo, K., Smale, S. T. (1996).** "Generality of a functional initiator consensus sequence." Gene **182**(1-2): 13-22.
- Löwer, R. (1999).** "The pathogenic potential of endogenous retroviruses: facts and fantasies." Trends Microbiol **7**(9): 350-6.
- Löwer, R., Boller, K., Hasenmaier, B., Korbmacher, C., Müller-Lantzsch, N., Löwer, J., Kurth, R. (1993a).** "Identification of human endogenous retroviruses with complex mRNA expression and particle formation." Proc Natl Acad Sci U S A **90**(10): 4480-4.
- Löwer, R., Löwer, J., Frank, H., Harzmann, R., Kurth, R. (1984).** "Human teratocarcinomas cultured in vitro produce unique retrovirus-like viruses." J Gen Virol **65** (Pt 5): 887-98.
- Löwer, R., Löwer, J., Kurth, R. (1996).** "The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences." Proc Natl Acad Sci U S A **93**(11): 5177-84.
- Löwer, R., Löwer, J., Tondera-Koch, C., Kurth, R. (1993b).** "A general method for the identification of transcribed retrovirus sequences (R-U5 PCR) reveals the expression of the human endogenous retrovirus loci HERV-H and HERV-K in teratocarcinoma cells." Virology **192**(2): 501-11.
- Löwer, R., Tönjes, R. R., Korbmacher, C., Kurth, R., Löwer, J. (1995).** "Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K." J Virol **69**(1): 141-9.
- Macleod, D., Charlton, J., Mullins, J., Bird, A. P. (1994).** "Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island." Genes Dev **8**(19): 2282-92.
- Maeda, N. (1985).** "Nucleotide sequence of the haptoglobin and haptoglobin-related gene pair. The haptoglobin-related gene contains a retrovirus-like element." J Biol Chem **260**(11): 6698-709.
- Magin, C. (1998).** "Untersuchungen zur Steuerung der Expression der humanen endogenen Retrovirus-Familie HERV-K." Dissertation, Johann Wolfgang Goethe-Universität, Frankfurt.
- Mahajan, R., Delphin, C., Guan, T., Gerace, L., Melchior, F. (1997).** "A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2." Cell **88**(1): 97-107.

- Martin, M. A., Bryan, T., Rasheed, S., Khan, A. S. (1981).** "Identification and cloning of endogenous retroviral sequences present in human DNA." Proc Natl Acad Sci U S A **78**(8): 4892-6.
- Martinez, E., Zhou, Q., L'Etoile, N. D., Oelgeschlager, T., Berk, A. J., Roeder, R. G. (1995).** "Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding." Proc Natl Acad Sci U S A **92**(25): 11864-8.
- Mastrangelo, I. A., Courey, A. J., Wall, J. S., Jackson, S. P., Hough, P. V. (1991).** "DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement." Proc Natl Acad Sci U S A **88**(13): 5670-4.
- Matsui, T., Segall, J., Weil, P. A., Roeder, R. G. (1980).** "Multiple factors required for accurate initiation of transcription by purified RNA polymerase II." J Biol Chem **255**(24): 11992-6.
- Mayer, J., Sauter, M., Racz, A., Scherer, D., Mueller-Lantzsch, N., Meese, E. (1999).** "An almost-intact human endogenous retrovirus K on human chromosome 7." Nat Genet **21**(3): 257-8.
- Melchior, F. (2000).** "SUMO--nonclassical ubiquitin." Annu Rev Cell Dev Biol **16**: 591-626.
- Monk, M., Boubelik, M., Lehnert, S. (1987).** "Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development." Development **99**(3): 371-82.
- Moore, R., Dixon, M., Smith, R., Peters, G., Dickson, C. (1987).** "Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of gag and pol." J Virol **61**(2): 480-90.
- Mueller-Lantzsch, N., Sauter, M., Weiskircher, A., Kramer, K., Best, B., Buck, M., Grasser, F. (1993).** "Human endogenous retroviral element K10 (HERV-K10) encodes a full-length gag homologous 73-kDa protein and a functional protease." AIDS Res Hum Retroviruses **9**(4): 343-50.
- Mühlbock, O. (1955).** "Note on a new inbred mouse strain GR/A." Eur. J. Cancer. **1**: 123-124.
- Murata, Y., Kim, H. G., Rogers, K. T., Udvardi, A. J., Horowitz, J. M. (1994).** "Negative regulation of Sp1 trans-activation is correlated with the binding of cellular proteins to the amino terminus of the Sp1 trans-activation domain." J Biol Chem **269**(32): 20674-81.
- Muster, T., Waltenberger, A., Grassauer, A., Hirschl, S., Caucig, P., Romirer, I., Fodinger, D., Seppel, H., Schanab, O., Magin-Lachmann, C., Lower, R., Jansen, B., Pehamberger, H., Wolff, K. (2003).** "An endogenous retrovirus derived from human melanoma cells." Cancer Res **63**(24): 8735-41.
- Naar, A. M., Lemon, B. D., Tjian, R. (2001).** "Transcriptional coactivator complexes." Annu Rev Biochem **70**: 475-501.

- Nabel, G., Baltimore, D. (1987).** "An inducible transcription factor activates expression of human immunodeficiency virus in T cells." Nature **326**(6114): 711-3.
- Nakao, M. (2001).** "Epigenetics: interaction of DNA methylation and chromatin." Gene **278**(1-2): 25-31.
- Ng, H. H., Bird, A. (1999).** "DNA methylation and chromatin modification." Curr Opin Genet Dev **9**(2): 158-63.
- O'Connell, C., O'Brien, S., Nash, W. G., Cohen, M. (1984).** "ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationships." Virology **138**(2): 225-35.
- Ohki, I., Shimotake, N., Fujita, N., Jee, J., Ikegami, T., Nakao, M., Shirakawa, M. (2001).** "Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA." Cell **105**(4): 487-97.
- Okumura, K., Sakaguchi, G., Takagi, S., Naito, K., Mimori, T., Igarashi, H. (1996).** "Sp1 family proteins recognize the U5 repressive element of the long terminal repeat of human T cell leukemia virus type I through binding to the CACCC core motif." J Biol Chem **271**(22): 12944-50.
- O'Leary, K. A., McQuiddy, P., Kasper, C. B. (1996).** "Transcriptional regulation of the TATA-less NADPH cytochrome P-450 oxidoreductase gene." Arch Biochem Biophys **330**(2): 271-80.
- O'Neill, M. J., O'Neill, R. J. (1999).** "Whatever happened to SRY?" Cell Mol Life Sci **56**(11-12): 883-93.
- Ono, M. (1986).** "Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes." J Virol **58**(3): 937-44.
- Ono, M., Yasunaga, T., Miyata, T., Ushikubo, H. (1986).** "Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome." J Virol **60**(2): 589-98.
- Overton, J. S. a. G. C. (1997).** TESS: Transcription Element Search Software on the WWW', Technical Report CBIL-TR-1997-1001-v0.0, Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania, URL:<http://www.cbil.upenn.edu/tess>.
- Palmarini, M., Coutsens, C., Dalziel, R. G., Bai, J., Stedman, K., DeMartini, J. C., Sharp, J. M. (1996).** "The exogenous form of Jaagsiekte retrovirus is specifically associated with a contagious lung cancer of sheep." J Virol **70**(3): 1618-23.
- Parker, C. S., Topol, J. (1984).** "A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity." Cell **36**(2): 357-69.
- Pascal, E., Tjian, R. (1991).** "Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism." Genes Dev **5**(9): 1646-56.

- Penotti, F. E. (1990).** "Human DNA TATA boxes and transcription initiation sites. A statistical study." J Mol Biol **213**(1): 37-52.
- Poukka, H., Karvonen, U., Janne, O. A., Palvimo, J. J. (2000).** "Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1)." Proc Natl Acad Sci U S A **97**(26): 14145-50.
- Poulat, F., Barbara, P. S., Desclozeaux, M., Soullier, S., Moniot, B., Bonneaud, N., Boizet, B., Berta, P. (1997).** "The human testis determining factor SRY binds a nuclear factor containing PDZ protein interaction domains." J Biol Chem **272**(11): 7167-72.
- Repaske, R., O'Neill, R. R., Steele, P. E., Martin, M. A. (1983).** "Characterization and partial nucleotide sequence of endogenous type C retrovirus segments in human chromosomal DNA." Proc Natl Acad Sci U S A **80**(3): 678-82.
- Rice, J. C., Allis, C. D. (2001).** "Histone methylation versus histone acetylation: new insights into epigenetic regulation." Curr Opin Cell Biol **13**(3): 263-73.
- Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., Klug, A. (1984).** "Structure of the nucleosome core particle at 7 Å resolution." Nature **311**(5986): 532-7.
- Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L., Wolffe, A. P. (2000).** "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nat Genet **25**(3): 338-42.
- Rodriguez, M. S., Desterro, J. M., Lain, S., Midgley, C. A., Lane, D. P., Hay, R. T. (1999).** "SUMO-1 modification activates the transcriptional response of p53." Embo J **18**(22): 6455-61.
- Rosen, C. A. (1992).** "HIV regulatory proteins: potential targets for therapeutic intervention." AIDS Res Hum Retroviruses **8**(2): 175-81.
- Rous, P. (1911).** "A sarcoma of the fowl transmissible by an agent separable from the tumor cells." J. Exp. Med. **13**: 397-411.
- Saffer, J. D., Jackson, S. P., Annarella, M. B. (1991).** "Developmental expression of Sp1 in the mouse." Mol Cell Biol **11**(4): 2189-99.
- Saffer, J. D., Jackson, S. P., Thurston, S. J. (1990a).** "SV40 stimulates expression of the transacting factor Sp1 at the mRNA level." Genes Dev **4**(4): 659-66.
- Saffer, J. D., Thurston, S. J., Annarella, M. B., Compton, J. G. (1990b).** "Localization of the gene for the trans-acting transcription factor Sp1 to the distal end of mouse chromosome 15." Genomics **8**(3): 571-4.
- Sambrook, J., Fritsch, E.F., Maniatis, T. (1989).** "Molecular cloning; A laboratory manual." 2. Ed., 3 Bde.. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, USA.

- Samuelson, L. C., Wiebauer, K., Snow, C. M., Meisler, M. H. (1990).** "Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution." Mol Cell Biol **10**(6): 2513-20.
- Sanger, F., Nicklen, S., Coulson, A. R. (1977).** "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A **74**(12): 5463-7.
- Sapetschnig, A., Rischitor, G., Braun, H., Doll, A., Schergaut, M., Melchior, F., Suske, G. (2002).** "Transcription factor Sp3 is silenced through SUMO modification by PIAS1." Embo J **21**(19): 5206-15.
- Schommer, S., Sauter, M., Krausslich, H. G., Best, B., Mueller-Lantzsch, N. (1996).** "Characterization of the human endogenous retrovirus K proteinase." J Gen Virol **77**(Pt 2): 375-9.
- Seiki, M., Hikikoshi, A., Yoshida, M. (1990).** "The U5 sequence is a cis-acting repressive element for genomic RNA expression of human T cell leukemia virus type I." Virology **176**(1): 81-6.
- Seto, E., Lewis, B., Shenk, T. (1993).** "Interaction between transcription factors Sp1 and YY1." Nature **365**(6445): 462-4.
- Shatkin, A. J., Manley, J. L. (2000).** "The ends of the affair: capping and polyadenylation." Nat Struct Biol **7**(10): 838-42.
- Shirley, M. A., Anderson, K. M. (1977).** "Electron-microscopic visualization of transcriptionally active and less active chromatin fractions from the rat ventral prostate and their content of histones." Can J Biochem **55**(1): 9-18.
- Simmen, R. C., Zhang, X. L., Zhang, D., Wang, Y., Michel, F. J., Simmen, F. A. (2000).** "Expression and regulatory function of the transcription factor Sp1 in the uterine endometrium at early pregnancy: implications for epithelial phenotype." Mol Cell Endocrinol **159**(1-2): 159-70.
- Sjøttem, E., Anderssen, S., Johansen, T. (1996).** "The promoter activity of long terminal repeats of the HERV-H family of human retrovirus-like elements is critically dependent on Sp1 family proteins interacting with a GC/GT box located immediately 3' to the TATA box." J Virol **70**(1): 188-98.
- Smale, S. T. (1997).** "Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes." Biochim Biophys Acta **1351**(1-2): 73-88.
- Smale, S. T. (2001).** "Core promoters: active contributors to combinatorial gene regulation." Genes Dev **15**(19): 2503-8.
- Smale, S. T., Jain, A., Kaufmann, J., Emami, K. H., Lo, K., Garraway, I. P. (1998).** "The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes." Cold Spring Harb Symp Quant Biol **63**: 21-31.

- Smale, S. T., Schmidt, M. C., Berk, A. J., Baltimore, D. (1990).** "Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID." Proc Natl Acad Sci U S A **87**(12): 4509-13.
- Smith, M. M. (1991).** "Histone structure and function." Curr Opin Cell Biol **3**(3): 429-37.
- Sogawa, K., Kikuchi, Y., Imataka, H., Fujii-Kuriyama, Y. (1993).** "Comparison of DNA-binding properties between BTEB and Sp1." J Biochem (Tokyo) **114**(4): 605-9.
- Steinhuber, S., Brack, M., Hunsmann, G., Schwelberger, H., Dierich, M. P., Vogetseder, W. (1995).** "Distribution of human endogenous retrovirus HERV-K genomes in humans and different primates." Hum Genet **96**(2): 188-92.
- Strahl, B. D., Allis, C. D. (2000).** "The language of covalent histone modifications." Nature **403**(6765): 41-5.
- Strazzullo, M., Majello, B., Lania, L., La Mantia, G. (1994).** "Mutational analysis of the human endogenous ERV9 proviruses promoter region." Virology **200**(2): 686-95.
- Struhl, K. (1999).** "Fundamentally different logic of gene regulation in eukaryotes and prokaryotes." Cell **98**(1): 1-4.
- Su, H., Lau, Y. F. (1993).** "Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach." Am J Hum Genet **52**(1): 24-38.
- Suske, G. (1999).** "The Sp-family of transcription factors." Gene **238**(2): 291-300.
- Tang, Y., Nilsson, L. (1998).** "Interaction of human SRY protein with DNA: a molecular dynamics study." Proteins **31**(4): 417-33.
- Temin, H. M., Mizutani, S. (1970).** "RNA-dependent DNA polymerase in virions of Rous sarcoma virus." Nature **226**(252): 1211-3.
- Thelen, K. (1995).** Inventarisierung LTR-haltiger mRNA humaner Krebszelllinien und ausführliche Charakterisierung einzelner Spezies. Frankfurt, Johann Wolfgang Goethe-Universität.
- Tönjes, R. R., Boller, K., Limbach, C., Lugert, R., Kurth, R. (1997).** "Characterization of human endogenous retrovirus type K virus-like particles generated from recombinant baculoviruses." Virology **233**(2): 280-91.
- Tönjes, R. R., Czuderna, F., Kurth, R. (1999).** "Genome-wide screening, cloning, chromosomal assignment, and expression of full-length human endogenous retrovirus type K." J Virol **73**(11): 9187-95.
- Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M. I., Kidd, K. K., Lenz, J. (2001).** "Insertional polymorphisms of full-length endogenous retroviruses in humans." Curr Biol **11**(19): 1531-5.

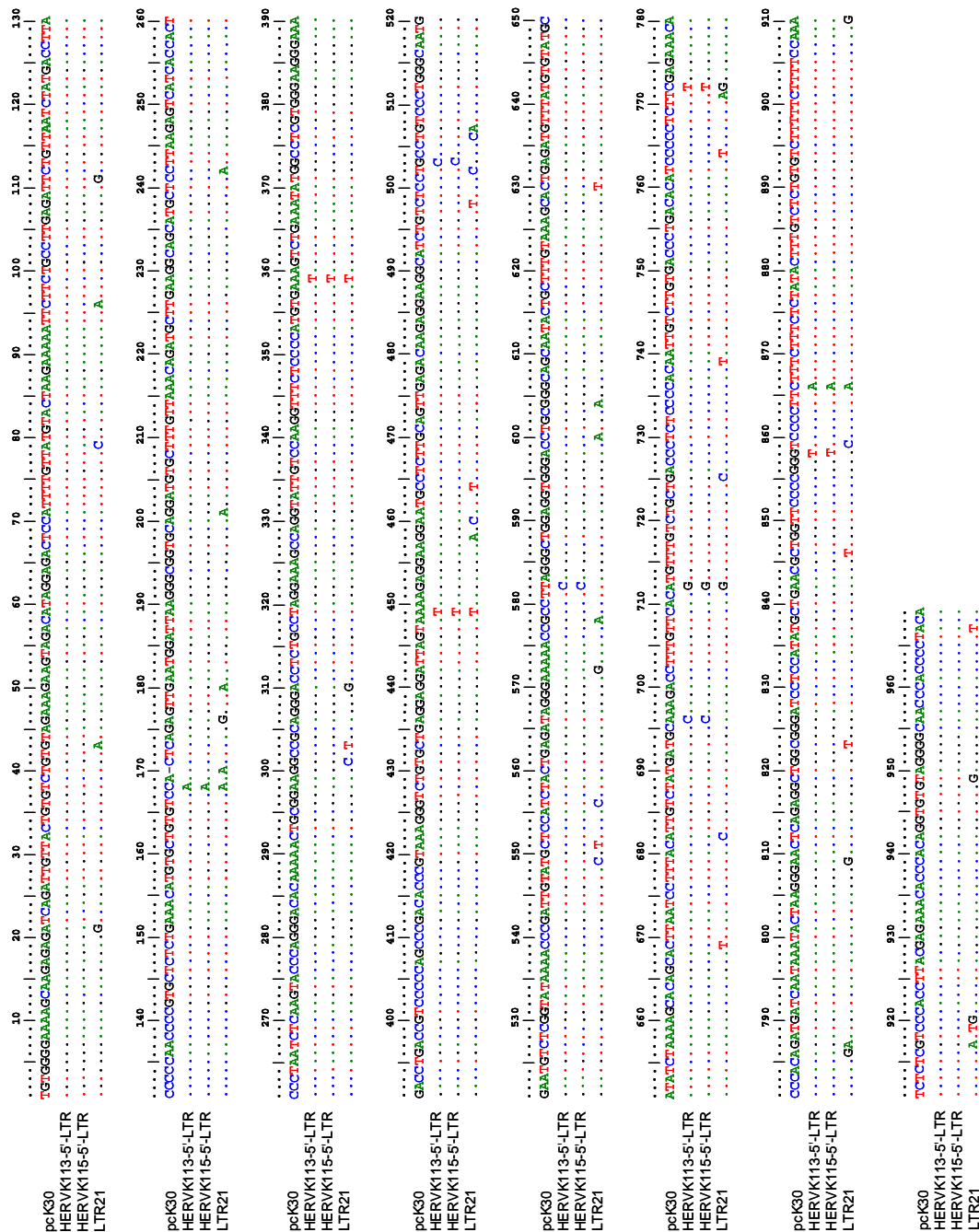
- Udvardia, A. J., Templeton, D. J., Horowitz, J. M. (1995).** "Functional interactions between the retinoblastoma (Rb) protein and Sp-family members: superactivation by Rb requires amino acids necessary for growth suppression." Proc Natl Acad Sci U S A **92**(9): 3953-7.
- Ukiyama, E., Jancso-Radek, A., Li, B., Milos, L., Zhang, W., Phillips, N. B., Morikawa, N., King, C. Y., Chan, G., Haqq, C. M., Radek, J. T., Poulat, F., Donahoe, P. K., Weiss, M. A. (2001).** "SRY and architectural gene regulation: the kinetic stability of a bent protein-DNA complex can regulate its transcriptional potency." Mol Endocrinol **15**(3): 363-77.
- Venter, J. C., Adams, M. D., et al. (2001).** "The sequence of the human genome." Science **291**(5507): 1304-51.
- Wang-Johanning, F., Frost, A. R., Jian, B., Azerou, R., Lu, D. W., Chen, D. T., Johanning, G. L. (2003a).** "Detecting the expression of human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma." Cancer **98**(1): 187-97.
- Wang-Johanning, F., Frost, A. R., Jian, B., Epp, L., Lu, D. W., Johanning, G. L. (2003b).** "Quantitation of HERV-K env gene expression and splicing in human breast cancer." Oncogene **22**(10): 1528-35.
- Weis, L., Reinberg, D. (1997).** "Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins." Mol Cell Biol **17**(6): 2973-84.
- Weiss, R. A. (1993).** The Retroviridae, Volume 2. Levy, J. A., Plenum, New York: 1-108.
- Wickens, M. (1990).** "How the messenger got its tail: addition of poly(A) in the nucleus." Trends Biochem Sci **15**(7): 277-81.
- Wilkerson, D. C., Wolfe, S. A., Grimes, S. R. (2002).** "Sp1 and Sp3 activate the testis-specific histone H1t promoter through the H1t/GC-box." J Cell Biochem **86**(4): 716-25.
- Wolffe, A. P., Matzke, M. A. (1999).** "Epigenetics: regulation through repression." Science **286**(5439): 481-6.
- Worrad, D. M., Schultz, R. M. (1997).** "Regulation of gene expression in the preimplantation mouse embryo: temporal and spatial patterns of expression of the transcription factor Sp1." Mol Reprod Dev **46**(3): 268-77.
- Yajima, S., Lee, S. H., Minowa, T., Mouradian, M. M. (1998).** "Sp family transcription factors regulate expression of rat D2 dopamine receptor gene." DNA Cell Biol **17**(5): 471-9.
- Yang, P., Zemba, M., Aboud, M., Flugel, R. M., Lochelt, M. (1997).** "Deletion analysis of both the long terminal repeat and the internal promoters of the human foamy virus." Virus Genes **15**(1): 17-23.
- Yeh, E. T., Gong, L., Kamitani, T. (2000).** "Ubiquitin-like proteins: new wines in new bottles." Gene **248**(1-2): 1-14.

- York, D. F., Vigne, R., Verwoerd, D. W., Querat, G. (1992).** "Nucleotide sequence of the jaagsiekte retrovirus, an exogenous and endogenous type D and B retrovirus of sheep and goats." J Virol **66**(8): 4930-9.
- Yuan, X., Lu, M. L., Li, T., Balk, S. P. (2001).** "SRY interacts with and negatively regulates androgen receptor transcriptional activity." J Biol Chem **276**(49): 46647-54.
- Zenzie-Gregory, B., Khachi, A., Garraway, I. P., Smale, S. T. (1993).** "Mechanism of initiator-mediated transcription: evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence." Mol Cell Biol **13**(7): 3841-9.
- Zhou, T., Chiang, C. M. (2001).** "The intronless and TATA-less human TAF(II)55 gene contains a functional initiator and a downstream promoter element." J Biol Chem **276**(27): 25503-11.

7 Anhang

7.1 Sequenzvergleich der HERV-K LTRs pcK30, LTR21, K113 und K115

Die 5'-LTR-Sequenzen wurden den veröffentlichten HERV-K Genomen K113 und K115 entnommen (Turner et al. 2001).



Sequenzvergleich der HERV-K LTRs pcK30, LTR21, K113 und K115. Übereinstimmungen mit der aktiven pcK30-Sequenz sind als Punkt dargestellt, Abweichungen durch die Abkürzung der entsprechenden Base. Striche kennzeichnen das Fehlen von Basenpositionen, die mindestens in einer anderen Sequenz vorkommen.

7.2 Konsensussequenzen

Ergebnis der Suche nach potentiellen Transkriptionsfaktorkonsensussequenzen mit Hilfe des Transkriptionsfaktorsuchprogramms TESS (Overton 1997):

Gesucht wurde auf der HERV-K LTR pcK30 nach Bindestellen für die humanen Transkriptionsfaktoren Sp1-Sp4 (GC/GT-Boxen) und SRY. Für GC/GT-Boxen wurden Konsensussequenzen mit einer log-likelihood <10 ausgeschlossen. Die experimentell bewiesenen GC/GT-Boxen im Bereich bps 756-783 haben eine log-likelihood von 18,00.

Bei der Suche nach SRY- Konsensussequenzen wurde eine log-likelihood Ausschlussgrenze von <6 gewählt.

```

1  TGTGGGGAAA AGCAAGAGAG ATCAGATTGT TACTGTGTCT GTGTAGAAAG AAGTAGACAT AGGAGACTCC ATTTTGTAT
      -----
81  GTACTAAGAA AAATTCTTCT GCCTTGAGAT TCTGTTAATC TATGACCTTA CCCCCAACCC CGTGCTCTCT GAAACATGTG
      ←-----
161 CTGTGTCCAC TCAGAGTTGA ATGGATTAAG GCGGTGCAG GATGTGCTTT GTTAAACAGA TGCTTGAAGG CAGCATGCTC
      -----
241 CTTAAGAGTC ATCACCCTC CCTAATCTCA AGTACCCAGG GACACAAAAA CTGCGGAAGG CCGCAGGGAC CTCTGCCTAG
      -----
321 GAAAGCCAGG TATTGTCAA GTTTCTCCC CATGTGAAAG TCTGAAATAT GGCCTCGTGG GAAGGGAAAG ACCTGACCGT
      -----
401 CCCCAGCCC GACACCCGTA AAGGGTCTGT GCTGAGGAGG ATTAGTAAAA GAGGAAGGAA TGCCTCTTGC AGTTGAGACA
      -----
481 AGAGGAAGGC ATCTGTCTCC TGCCTGTCCC TGGGCAATGG AATGTCTCGG TATAAAACCC GATTGTATGC TCCATCTACT
      -----
561 GAGATAGGGA AAAACCGCCT TAGGGCTGGA GGTGGGACCT GCGGGCAGCA ATACTGCTTT GTAAAGCACT GAGATGTTTA
      -----
641 TGTGTATGCA TATCTAAAAG CACAGCACTT AATCCTTTAC ATTGTCTATG ATGCAAAGAC CTTTGTTCAC ATGTTTGTCT
      -----
721 GCTGACCCTC TCCCACAAT TGTCTTGTGA CCCTGACACA TCCCCTCTT CGAGAAACAC CCACAGATGA TCAATAAATA
      -----
801 CTAAGGGAAC TCAGAGGCTG GCGGGATCCT CCATATGCTG AACGCTGGTT CCCCGGTCC CCTTCTTCT TTCTCTATAC
      -----
881 TTTGTCTCTG TGTCTTTTTT TTTTCAAAT CTCTCGTCCC ACCTTACGAG AAACACCCAC AGGTGTGTAG GGGCAACCCA
      -----
961 CCCCTACA
      ←-----

```

——— GC/GT-Box
 - - - - SRY-Konsensussequenz

7.3 Auszüge aus der Uniprot-SwissProt-Datenbank

7.3.1 Sp1

Entry name	SP1_HUMAN
Accession number	P08047 , Q9H3Q5, Q9NR51, Q9NY21, Q9NYE7
Created	Rel. 08, 1-AUG-1988
Sequence update	Rel. 40, 16-OCT-2001
Annotation update	Rel. 43, 15-MAR-2004
Description	Transcription factor Sp1
Gene name(s)	SP1 OR TSFP1
Organism source	Homo sapiens (Human)
Taxonomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

Length: **785 aa**, molecular weight: **80693 Da**.

MSDQDHSMD	MTAVVKIEKG	VGGNNGGNGN	GGGAFSQARS	SSTGSSSSTG	GGQESQPSP	60
LALLAATCSR	IESPNENSNN	SQGPSQSGGT	GELDLTATQL	SQGANGWQII	SSSSGATPTS	120
KEQSGSSTNG	SNGSESSKNR	TVSGGQYVVA	AAPNLQNNQV	LTGLPGVMPN	IQYQVIPQFQ	180
TVDGGQLQFA	ATGAQVQDQD	SGQIQIIPGA	NQQIITNRGS	GGNIIAAMPN	LLQQAVPLQG	240
LANNVLSGQT	QYVTNVPVAL	NGNITLLPVN	SVSAATLTPS	SQAVTISSSG	SQESGSQPVT	300
SGTTISSASL	VSSQASSSSF	FTNANSYSTT	TTTNSMGIMN	FTTSGSSGTN	SQGQTPQRVS	360
GLQGSDALNI	QQNQTSGGSL	QAGQQKEGEQ	NQQTQQQQIL	IQPQLVQGGQ	ALQALQAAPL	420
SGQTFTTQAI	SQETLQNLQL	QAVPNSGPII	IRTPTVGPNG	QVSWQTLQLQ	NLQVQNPQAA	480
TITLAPMQGV	SLGQTSSSNT	TLTPIASAAS	IPAGTVTVNA	AQLSSMPGLQ	TINLSALGTS	540
GIQVHPIQGL	PLAIANAPGD	HGAQLGLHGA	GGDGIHDDTA	GGEEGENSPD	AQPQAGRTR	600
REACTCPYCK	DSEGRGSGDP	GKKKQHICHI	QGCGKVYGKT	SHLRAHLRWH	TGERPFMCTW	660
SYCGKRFTRS	DELQRHKRTH	TGEKKFACPE	CPKRFRMSDH	LSKHIKTHQN	KKGGPGVALS	720
VGTLPLDSGA	GSESGTATP	SALITTNMVA	MEAICPEGIA	RLANSGINVM	QVADLQSINI	780
SGNGF						785

References:

- [1] Haggart,M.H., Ladurner,A.G., Submitted APR-2000 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 4-785 FROM N.A. Comments: TISSUE=Cervical carcinoma
- [2] Takahara,T., Kanazu,S., Yanagisawa,S., Akanuma,H., Heterogeneous Sp1 mRNAs in human HepG2 cells include a product of homotypic trans-splicing. (2000) *J. Biol. Chem.* 275:38067-38072. Position: SEQUENCE OF 1-558 FROM N.A. Medline: 20545561. PubMed: 10973950.
- [3] Kadonaga,J.T., Carner,K.R., Masiarz,F.R., Tjian,R., Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. (1987) *Cell* 51:1079-1090. Position: SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675. Medline: 88080466. PubMed: 3319186.
- [4] Nicolas,M., Noe,V., Ciudad,C.J., Expression of transcription factor Sp1 mRNA in mammalian cells. Submitted APR-2000 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 1-109 FROM N.A.

[5] Handschug,K., Huebner,A., Sequencing of the 5' end of human transcription factor SP1 mRNA. Submitted FEB-2000 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 1-98 FROM N.A.

[6] Jackson,S.P., Tjian,R., O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. (1988) *Cell* 55:125-133 Position O-GLYCOSYLATION. Medline: 89003041. PubMed: 3139301.

[7] Narayan,V.A., Kriwacki,R.W., Caradonna,J.P., Structures of zinc finger domains from transcription factor Sp1. Insights into sequence-specific protein-DNA recognition. (1997) *J. Biol. Chem.* 272:7801-7809. Position: STRUCTURE BY NMR OF 654-684 AND 684-712. Medline: 97218212. PubMed: 9065444.

[8] Parks,C.L., Shenk,T., The serotonin 1a receptor gene contains a TATA-less promoter that responds to MAZ and Sp1. (1996) *J. Biol. Chem.* 271:4417-4430. Position: IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES. Medline: 96224025. PubMed: 8626793.

7.3.2 Sp3

Sp3 (Vollängenprotein)

Entry name	SP3_HUMAN
Accession number	Q02447 , Q8TD56, Q9BQR1
Created	Rel. 31, 1-FEB-1995
Sequence update	Rel. 42, 10-OCT-2003
Annotation update	Rel. 42, 10-OCT-2003
Description	Transcription factor Sp3 (SPR-2)
Gene name(s)	SP3
Organism source	Homo sapiens (Human)
Taxonomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

Length: **781 aa**, molecular weight: **81925 Da**.

MTAPEKPKVKQ	EEMAALDVDS	GGGGGGGGGH	GEYLQQQQQH	GNGAVAAAAA	AQDTQPSPLA	60
LLAATCSKIG	PPSPGDDEEE	AAAAAGAPAA	AGATGDLASA	QLGGAPNRWE	VLSATPTTIK	120
DEAGNLVQIP	SAATSSGQYV	LPLQNLQNLQ	IFSVAPGSDS	SNGTVSSVQY	QVIPQIQSAD	180
GQQVQIGFTG	SSDNGGINQE	SSQIQIIPGS	NQTLLASGTP	SANIQNLIPQ	TGQVQVQGVA	240
IGSSSFPGQT	QVVANVPLGL	PGNITFVPIN	SVDLDSLGLS	GSSQTMTAGI	NADGHLINTG	300
QAMDSSDNSE	RTGERVSPDI	NETNTDSDLF	VPTSSSSQLP	VTIDSTGILQ	QNTNSLTTSS	360
GQVHSSDLQG	NYIQSPVSEE	TQAQNIQVST	AQPVVQHLQL	QESQQPTSQA	QIVQGITPQT	420
IHGVSASGQN	ISQQALQNLQ	LQLNPGTFLI	QAQTVTPSGQ	VTWQTFQVQG	VQNLQNLQIQ	480
NTAAQQITLT	PVQTLTLGQV	AAGGAFTSTP	VSLSTGQLPN	LQTVTVNSID	SAGIQLHPGE	540
NADSPADIRI	KEEEPDPPEW	QLSGDSTLNT	NDLTHLRVQV	VDEEGDQQHQ	EGKRLRRVAC	600
TCPNCKEGGG	RGTNLGKKKQ	HICHIPGCGK	VYGKTSHLRA	HLRWHSGERP	FVCNWMYCGK	660
RFTRSDELQR	HRRTHTEGKK	FVCPECKRF	MRSDDLAKHI	KTHQNKKGIIH	SSSTVLASVE	720
AARDDTLITA	GGTLLILANI	QQGSVSGIGT	VNTSATSNDQ	ILTNTIPLQ	LVTVSGNETM	780
E						781

References:

- [1] Oleksiak,M.F., Crawford,D.L., 5' structure of human Sp3. Submitted MAR-2002 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 1-135 FROM N.A.
- [2] Meyer-Grahle,U., Regulation of hTERT-gene transcription by SP3. Submitted MAR-2001 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 13-112 FROM N.A.
- [3] Kingsley,C., Winoto,A., Submitted MAR-2001 to the EMBL GenBank DDBJ databases. Position: SEQUENCE OF 71-781 FROM N.A., AND REVISIONS. Comments TISSUE=T-cell.
- [4] Kingsley,C., Winoto,A., Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. (1992) *Mol. Cell. Biol.* 12:4251-4261. Position: SEQUENCE OF 129-781 FROM N.A. Medline: 93024366. PubMed: 1341900
- [5] Hagen,G., Mueller,S., Beato,M., Suske,G., Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. (1992) *Nucleic Acids Res.* 20:5519-5525.

Position: SEQUENCE OF 85-781 FROM N.A., AND VARIANT ALA-164. Comments: TISSUE=Uterus.
Medline: 93087156. PubMed: 1454515.

[6] Ihn,H., Trojanowska,M., Sp3 is a transcriptional activator of the human alpha2(I) collagen gene. (1997) *Nucleic Acids Res.* 25:3712-3717. Position: FUNCTION. Medline: 97426517. PubMed: 9278495.

Sp3 (intern initiierte Isoform 1)

Entry name **Q8WWU3**
Accession number Q8WWU3
Created TrEMBLrel. 20, 1-MAR-2002
Sequence update TrEMBLrel. 20, 1-MAR-2002
Annotation update TrEMBLrel. 24, 1-JUN-2003
Description Specificity protein 3 internally initiated isoform 1
Gene name(s) SP3
Organism source Homo sapiens (Human)
Taxonomy Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo

Length: **496 aa**, molecular weight: **53729 Da**

```

MTAGINADGH LINTGQAMDS SDNSERTGER VSPDINETNT DTDLFVPTSS SSQLPVTIDS      60
TGILQQNTNS LTTSSGQVHS SDLQGNIIQS PVSEETQAQN IQVSTAQPVV QHLQLQESQQ      120
PTSQAQIVQG ITPQTIHGVS ASGQNISQQA LQNLQLQLNP GTFLIQAQTV TPSGQVTWQT      180
FQVQGVQNLQ NLQIQNTAAQ QITLTPVQTL TLGQVAAGGA FTSTPVSLSL GQLPNLQTVT      240
VNSIDSAGIQ LHPGENADSP ADIRIKEEEP DPEEWQLSGD STLNTNDLTH LRVQVVDEEG      300
DQQHQEGKRL RRVACTCPNC KEGGGRGTNL GKKKQHICHI PGCGKVGKKT SHLRAHLRWH      360
SGERPFVCNW MYCGKRFRS DELQRHRRTH TGEKKFVCP ESKRFMRSDH LAKHIKTHQN      420
KKGIIHSSSTV LASVEAARD TLITAGGTTL ILAKIQQGSV SGIGTVNTSA TSNQDILTNT      480
EIPLQLVTVS GNETME

```

References:

[1] Hernandez,E.M., Johnson,A., Notario,V., Chen,A., Richert,J.R., AUA as a Translation Initiation Site in Vitro for the Human Transcription Factor Sp3. (2002) *J. Biochem. Mol. Biol.* 0:0-0. Position: SEQUENCE FROM N.A.

Sp3 (intern initiierte Isoform 2)

Entry name **Q8WWU2**
 Accession number **Q8WWU2**
 Created TrEMBLrel. 20, 1-MAR-2002
 Sequence update TrEMBLrel. 20, 1-MAR-2002
 Annotation update TrEMBLrel. 24, 1-JUN-2003
 Description Specificity protein 3 internally initiated isoform 2
 Gene name(s) SP3
 Organism source Homo sapiens (Human)
 Taxonomy Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo

Length: **479 aa**, molecular weight: **52064 Da**

```

MDSSDNSERT  GERVSPDINE  TNTDSDLFVP  TSSSSQLPVT  IDSTGILQQN  TNSLTTSSGQ      60
VHSSDLQGNV  IQSPVSEETQ  AQNIQVSTAQ  PVVQHLQLQE  SQQPTSQAQI  VQGITPQTIH     120
GVQASGQNIS  QQALQNLQLQ  LNPQTFLIQA  QTVTPSGQVT  WQTFQVQGVQ  NLQNLQIQNT     180
AAQQITLTPV  QTLTLGQVAA  GGAFTSTPVS  LSTGQLPNLQ  TVTVNSIDSA  GIQLHPGENA     240
DSPADIRIKE  EEPDPEEWQL  SGDSTLNTND  LTHLRVQVVD  EEGDQQHQEG  KRLRRVACTC     300
PNCKEGGGRG  TNLGKKKQHI  CHIPGCGKVY  GKTSHLRAHL  RWHSGERPFV  CNWMYCGKRF     360
TRSDELQRHR  RTHTGKQKQV  CPECSKRFMR  SDHLAKHIKT  HQNKKGIHSS  STVLASVEAA     420
RDDTLITAGG  TTLILAKIQQ  GSVSGIGTVN  TSATSNQDIL  TNTEIPLQLV  TVSGNETME      479
  
```

References:

[1] Hernandez,E.M., Johnson,A., Notario,V., Chen,A., Richert,J.R., AUA as a Translation Initiation Site in Vitro for the Human Transcription Factor Sp3. (2002) *J. Biochem. Mol. Biol.* 0:0-0. Position: SEQUENCE FROM N.A.

7.3.3 SRY

Entry name **SRY_HUMAN**
 Accession number **Q05066**
 Created Rel. 29, 1-JUN-1994
 Sequence update Rel. 29, 1-JUN-1994
 Annotation update Rel. 43, 15-MAR-2004
 Description Sex-determining region Y protein (Testis-determining factor)
 Gene name(s) SRY OR TDF
 Organism source Homo sapiens (Human)
 Taxonomy Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

Length: **204 aa**, molecular weight: **23884 Da**

MQSYASAMLS	VFNSDDYSPA	VQENIPALRR	SSSFLCTESC	NSKYQCETGE	NSKGNVQDRV	60
KRPMNAFIVW	SRDQRRKMAL	ENPRMRNSEI	SKQLGYQWKM	LTEAEKWPF	QEAQKLQAMH	120
REKYPNYKYR	PRRKAKMLPK	NCSLLPADPA	SVLCSEVQLD	NRLYRDDCTK	ATHSRMEHQL	180
GHLPPINAAS	SPQQRDRYSH	WTKL				204

References:

- [1] Sinclair,A.H., Berta,P., Palmer,M.S., Hawkins,J.R., Griffiths,B.L., Smith,M.J., Foster,J.W., Frischauf,A.-M., Lovell-Badge,R., Goodfellow,P.N., A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. (1990) *Nature* 346:240-244. Position: SEQUENCE FROM N.A. Medline: 90326153. PubMed: 1695712.
- [2] Su,H., Lau,Y.-F.C., Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach. (1993) *Am. J. Hum. Genet.* 52:24-38 Position SEQUENCE FROM N.A. Medline: 93167288. PubMed: 8434602.
- [3] Behlke,M.A., Bogan,J.S., Beer-Romero,P., Page,D.C., Submitted DEC-1992 to the EMBL GenBank DDBJ databases Position SEQUENCE FROM N.A.
- [4] Whitfield,L.S., Hawkins,J.R., Goodfellow,P.N., Sulston,J., 41 kilobases of analyzed sequence from the pseudoautosomal and sex- determining regions of the short arm of the human Y chromosome. (1995) *Genomics* 27:306-311. Position: SEQUENCE FROM N.A. Medline: 96044437. PubMed: 7557997.
- [5] King,C.Y., Weiss,M.A., The SRY high-mobility-group box recognizes DNA by partial intercalation in the minor groove: a topological mechanism of sequence specificity. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:11990-11994 Position CHARACTERIZATION OF DNA-BINDING. Medline: 94089755. PubMed: 8265659.

- [6] Ferrari,S., Harley,V.R., Pontiggia,A., Goodfellow,P.N., Lovell-Badge,R., Bianchi,M.E., SRY, like HMG1, recognizes sharp angles in DNA. (1992) *EMBO J.* 11:4497-4506. Position: CHARACTERIZATION OF DNA-BINDING. Medline: 93049201. PubMed: 1425584.
- [7] Werner,M.H., Huth,J.R., Gronenborn,A.M., Clore,G.M., Molecular basis of human 46X,Y sex reversal revealed from the three- dimensional solution structure of the human SRY-DNA complex. (1995) *Cell* 81:705-714. Position: STRUCTURE BY NMR OF 56-131. Medline: 95292338. PubMed: 7774012.
- [8] Murphy,E.C., Zhurkin,V.B., Louis,J.M., Cornilescu,G., Clore,G.M., Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. (2001) *J. Mol. Biol.* 312:481-499. Position: STRUCTURE BY NMR OF 57-218. Medline: 21448821. PubMed: 11563911.
- [9] Hawkins,J.R., Mutational analysis of SRY in XY females. (1993) *Hum. Mutat.* 2:347-350. Position: REVIEW ON VARIANTS. Medline: 94080248. PubMed: 8257986.
- [10] Cameron,F.J., Sinclair,A.H., Mutations in SRY and SOX9: testis-determining genes. (1997) *Hum. Mutat.* 9:388-395 Position REVIEW ON VARIANTS. Medline: 97288993. PubMed: 9143916.
- [11] Berta,P., Hawkins,J.R., Sinclair,A.H., Taylor,A., Griffiths,B.L., Goodfellow,P.N., Fellous,M., Genetic evidence equating SRY and the testis-determining factor. (1990) *Nature* 348:448-450. Position: VARIANTS GDXY LEU-60 AND ILE-64. Medline: 91061886. PubMed: 2247149.
- [12] Affara,N.A., Chalmers,I.J., Ferguson-Smith,M.A., Analysis of the SRY gene in 22 sex-reversed XY females identifies four new point mutations in the conserved DNA binding domain. (1993) *Hum. Mol. Genet.* 2:785-789. Position: VARIANTS GDXY GLY-62; THR-78 AND TRP-133. Medline: 93357752. PubMed: 8353496.
- [13] Vilain,E., McElreavey,K., Jaubert,F., Raymond,J.-P., Richaud,F., Fellous,M., Familial case with sequence variant in the testis-determining region associated with two sex phenotypes. (1992) *Am. J. Hum. Genet.* 50:1008-1011. Position: VARIANT GDXY LEU-60. Medline: 92238481. PubMed: 1570829.
- [14] Hawkins,J.R., Taylor,A., Goodfellow,P.N., Migeon,C.J., Smith,K.D., Berkovitz,G.D., Evidence for increased prevalence of SRY mutations in XY females with complete rather than partial gonadal dysgenesis. (1992) *Am. J. Hum. Genet.* 51:979-984. Position: VARIANTS GDXY MET-90 AND ILE-106. Medline: 93035394. PubMed: 1415266.
- [15] Hawkins,J.R., Taylor,A., Berta,P., Levilliers,J., van der Auwera,B., Goodfellow,P.N., Mutational analysis of SRY: nonsense and missense mutations in XY sex reversal. (1992) *Hum. Genet.* 88:471-474. Position: VARIANT GDXY ARG-95. Medline: 92155700. PubMed: 1339396.
- [16] Braun,A., Kammerer,S., Cleve,H., Loehrs,U., Schwarz,H.-P., Kuhnle,U., True hermaphroditism in a 46,XY individual, caused by a postzygotic somatic point mutation in the male gonadal sex-determining locus (SRY): molecular genetics and histological findings in a sporadic case. (1993) *Am.*

J. Hum. Genet. 52:578-585. Position: VARIANT GDXY HIS-101. Medline: 93190989. PubMed: 8447323.

[17] Jaeger,R.J., Harley,V.R., Pfeiffer,R.A., Goodfellow,P.N., Scherer,G., A familial mutation in the testis-determining gene SRY shared by both sexes. (1992) *Hum. Genet.* 90:350-355. Position: VARIANT GDXY SER-109. Medline: 93131260. PubMed: 1483689.

[18] Zeng,Y., Ren,Z., Zhang,M., Huang,Y., Zeng,F., Huang,S., A new de novo mutation (A113T) in HMG box of the SRY gene leads to XY gonadal dysgenesis. (1993) *J. Med. Genet.* 30:655-657. Position: VARIANT GDXY THR-113. Medline: 94016377. PubMed: 8105086.

[19] Poulat,F., Soulier,S., Goze,C., Heitz,F., Calas,B., Berta,P., Description and functional implications of a novel mutation in the sex-determining gene SRY. (1994) *Hum. Mutat.* 3:200-204. Position: VARIANT GDXY CYS-127. Medline: 94290489. PubMed: 8019555.

[20] Haqq,C.M., King,C.Y., Ukiyama,E., Falsafi,S., Haqq,T.N., Donahoe,P.K., Weiss,M.A., Molecular basis of mammalian sexual determination: activation of Mullerian inhibiting substance gene expression by SRY. (1994) *Science* 266:1494-1500. Position: VARIANT GDXY THR-68. Medline: 95076269. PubMed: 7985018.

[21] Schmitt-Ney,M., Thiele,H., Kaltwasser,P., Bardoni,B., Cisternino,M., Scherer,G., Two novel SRY missense mutations reducing DNA binding identified in XY females and their mosaic fathers. (1995) *Am. J. Hum. Genet.* 56:862-869. Position: VARIANTS GDXY GLY-91 AND LEU-125. Medline: 95233434. PubMed: 7717397.

[22] Hiort,O., Klaubert,G.T., True hermaphroditism with 46,XY karyotype and a point mutation in the SRY gene. (1995) *J. Pediatr.* 126:1022-1022. Position: VARIANT GDXY ALA-60. Medline: 95294690. PubMed: 7776083.

[23] Lundberg,Y., Ritzen,M., Harlin,J., Wedell,A., Novel missense mutation (P131R) in the HMG box of SRY in XY sex reversal. (1998) *Hum. Mutat. Suppl.* 1:0-S328. Position: VARIANT GDXY ARG-131.

[24] Scherer,G., Held,M., Erdel,M., Meschede,D., Horst,J., Lesniewicz,R., Midro,A.T., Three novel SRY mutations in XY gonadal dysgenesis and the enigma of XY gonadal dysgenesis cases without SRY mutations. (1998) *Cytogenet. Cell Genet.* 80:188-192. Position: VARIANTS GDXY ARG-64 AND VAL-67. Medline: 98341496. PubMed: 9678356.

[25] Domenice,S., Nishi,M.Y., Billerbeck,A.E.C., Latronico,A.C., Medeiros,M.A., Russell,A.J., Vass,K., Carvalho,F.M., Costa Frade,E.M., Arnhold,I.J.P., Mendonca,B.B., A novel missense mutation (S18N) in the 5' non-HMG box region of the SRY gene in a patient with partial gonadal dysgenesis and his normal male relatives. (1998) *Hum. Genet.* 102:213-215. Position: VARIANT GDXY ASN-18. Medline: 98180891. PubMed: 9521592.

[26] Doerk,T., Stuhmann,M., Miller,K., Schmidtke,J., Independent observation of SRY mutation I90M in a patient with complete gonadal dysgenesis. (1998) *Hum. Mutat.* 11:90-91. Position: VARIANT GDXY MET-90. Medline: 98111385. PubMed: 9450909.

- [27] Imai,A., Takagi,A., Tamaya,T., A novel sex-determining region on Y (SRY) missense mutation identified in a 46,XY female and also in the father. (1999) *Endocr. J.* 46:735-739. Position: VARIANT GDXY SER-76. Medline: 20134222 PubMed: 10670762.
- [28] Jakubiczka,S., Bettecken,T., Stumm,M., Neulen,J., Wieacker,P., Another mutation within the HMG-box of the SRY gene associated with Swyer syndrome. (1999) *Hum. Mutat.* 13:85-85. Position: VARIANT GDXY ARG-108.
- [29] Schaeffler,A., Barth,N., Winkler,K., Zietz,B., Ruummele,P., Knuechel,R., Schoelmerich,J., Palitzsch,K.-D., Identification of a new missense mutation (Gly95Glu) in a highly conserved codon within the high-mobility group box of the sex-determining region Y gene: report on a 46,XY female with gonadal dysgenesis and yolk-sac tumor. (2000) *J. Clin. Endocrinol. Metab.* 85:2287-2292. Position: VARIANT GDXY GLU-95. Medline: 20309305. PubMed: 10852465.
- [30] Canto,P., de la Chesnaye,E., Lopez,M., Cervantes,A., Chavez,B., Vilchis,F., Reyes,E., Ulloa-Aguirre,A., Kofman-Alfaro,S., Mendez,J.P., A mutation in the 5' non-high mobility group box region of the SRY gene in patients with Turner syndrome and Y mosaicism. (2000) *J. Clin. Endocrinol. Metab.* 85:1908-1911. Position: VARIANT TURNER SYNDROME ASN-18. Medline: 20300159. PubMed: 10843173.
- [31] Okuhara,K., Tajima,T., Nakae,J., Fujieda,K., A novel missense mutation in the HMG box region of the SRY gene in a Japanese patient with an XY sex reversal. (2000) *J. Hum. Genet.* 45:112-114. Position: VARIANT GDXY TYR-87. Medline: 20186456. PubMed: 10721678.
- [32] Jordan,B.K., Jain,M., Natarajan,S., Frasier,S.D., Vilain,E., Familial mutation in the testis-determining gene SRY shared by an XY female and her normal father. (2002) *J. Clin. Endocrinol. Metab.* 87:3428-3432. Position: VARIANT GDXY PHE-127. Medline: 22103278. PubMed: 12107262.

Datenbank Internet-Adresse und Copyright:

<http://harvester.embl.de/>

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

8 Abkürzungsverzeichnis

ψ	Verpackungssignal
A	Adenosin
AIDS	Acquired Immune Deficiency Syndrome
ALV	Aviäres Leukämie-Virus
ASLV	Aviäres Sarkom-Leukämie-Virus
bp(s)	Basenpaar(e)
C	Cytosin
CA	Capsid Protein
cDNA	komplementäre Desoxyribonukleinsäure
CTD	Carboxyterminale Domäne
DM	Lipiddoppelmembran
DNA	Desoxyribonukleinsäure
DPE	Downstream Promoter Element
EMSA	Electrophoretic Mobility Shift Assay
Env	Envelope Protein
ERV	endogenes Retrovirus
G	Guanosin
Gag	Group Specific Antigen
GC/GT-Box	Sp1/Sp3-Konsensussequenz
GH	humanes Teratokarzinom
GTF	Genereller Transkriptionsfaktor
HeLa	Humanes Cervixkarzinom
HERV	Humanes endogenes Retrovirus
HFV	Humanes Foamy Virus
HIV	Humanes Immundefizienzvirus
HMG	High Mobility Group
HTDV	Human Teratocarcinoma Derived Virus
HTLV	Humanes T-Zell Leukämie Virus
IN	Integrase Protein
Inr	Initiator
JSRV	Jaagsiekte-Schaf-Retrovirus
kDA	Kilodalton

kB	Kilobase
LTR	Long Terminal Repeat
MA	Matrix Protein
MMTV	Maus-Mamma-Tumor-Virus
mRNA	Messenger RNA
N	Nukleotid (A, C, T, G)
NC	Nukleocapsid Protein
NF- κ B	Nukleärer Faktor der Ig κ leichten Kette von B-Zellen
ORF	Open Reading Frame
PBS	Primer Binding Site
PCR	Polymerase Chain Reaction
Pol	Polymerase / Reverse Transkriptase Protein
PolII	RNA-Polymerase II
Poly(A)-	Polyadenylierungs-
PPT	Poly-Purin Tract
PRT	Protease Protein
Py	Pyrimidin
R	Terminal Redundancy
RACE	Rapid Amplification of cDNA Ends
Rec	funktionales HERV-K-Äquivalent zu Rev (HIV)
Rev	Regulator of Expression of Virion Proteins (Protein, HIV)
RNA	Ribonukleinsäure
rNTP	Ribonukleosid-Triphosphat
rRNA	Ribosomale RNA
RSV	Rous-Sarkom-Virus
RT	Reverse Transkriptase Protein
SA	Spleißakzeptor
SD	Spleißdonor
sLTR	Solitary LTR
snRNA	Small Nuclear RNA
Sp1	Specificity Protein 1
SRY	Sex-determining Region Y
SU	Surface Protein
SUMO	Small Ubiquitin-like Modifier

SV40	Simian Virus 40
T	Thymin
TAF	TBP-assoziierter Faktor
Tat	Transaktivator of Transcription (Protein)
TATA-Box	TATA-Binde Protein-Konsensussequenz
TBP	TATA-Binde Protein
TFIID	Transkriptionsfaktor IID
TM	Transmembran Protein
tRNA	Transfer RNA
U	Unit
U3	Unique 3'-End
U5	Unique 5'-End
YY1	Yin-Yang 1

Abkürzungen, die nur einmal verwendet wurden, sind an Ort und Stelle erläutert. Aminosäuren und Nukleotide werden im Einbuchstabencode abgekürzt. Gen-Namen sind kursiv gedruckt. Physikalische, chemische und biologische Größen und gängige Abkürzungen der deutschen Sprache sind hier nicht aufgeführt.

Danksagung

Die vorliegende Dissertation wurde am Paul-Ehrlich-Institut, Bundesamt für Sera und Impfstoffe, Abteilung Retroelemente, in Langen angefertigt. Ich danke Herrn Prof. Dr. Löwer für die Bereitstellung des Arbeitsplatzes und die fachliche Anleitung bei der Durchführung.

Herrn Prof. Löwer und Frau Dr. Löwer danke ich für die Vergabe des Themas und besonders für die stetige Bereitschaft zur Diskussion und zur Erörterung von Problemen.

Frau Prof. Dr. Anna Starzinski-Powitz, Institut der Anthropologie und Humangenetik für Biologen, Johann Wolfgang Goethe-Universität Frankfurt am Main danke ich für die Übernahme des Zweitgutachtens.

Frau Steffanie Liller danke ich für ihre ausgezeichnete Hilfe bei der Klonierung von Expressionsplasmiden und bei der Western Blot Analyse.

Frau Ulrike Held und Frau Heike Strobel danke ich für ihre ständige Hilfsbereitschaft während der Durchführung von verschiedensten Experimenten.

Frau Patricia Bach und Frau Christina Hartz danke ich für ihre Hilfe bei der Klonierung der verschiedenen Luziferase-Konstrukte.

Herrn Herbert Bartel danke ich für die angenehme Hilfe bei den verschiedenen „Last-Minute“ Dia-Entwicklungen und Posterdrucken.

Mein besonderer Dank gilt meinen Kollegen Frau Silvia Hahn und Herrn Pablo Beltran, die mir während der gesamten Zeit immer hilfreich zur Seite standen, um Experimente und Probleme zu lösen. Ebenfalls für fachliche und menschliche Hilfe danke ich Herrn Dr. Michael Hoff und Herrn Oliver Stephan.

Ich danke meiner Freundin Frau Sabine Schlitt für Unterstützung, wenn ich sie brauchte und exzellentes Korrekturlesen.

Ich danke weiterhin folgenden Mitarbeitern und Weggefährten für Hilfsbereitschaft, Anregungen und Diskussionsbereitschaft: Frau Ines Amberg, Herrn Dr. Johannes Blümel, Herrn Dr. Michael Chudy, Herrn PD Dr. Christoph Forreiter, Frau Kim-Cybil Hermann, Herrn Dr. Frank Landgraf, Frau Dr. Christine Magin, Herrn Dr. Micha Nübling, Herrn Sven Sagasser, Herrn Erik Seibold, Herrn Martin Selbert, Herrn Dr. Volker Specke und all denen, die ich trotz ihres Beitrags hier vergessen habe.

Mein spezieller Dank gebührt meiner Kommilitonin und Weggefährtin während der Diplomarbeit Frau Dr. Daniela Löw, die mir die Grundlagen des molekularbiologischen Arbeitens vermittelte. Ihr viel zu früher Tod hinterlässt bei mir und meinen ehemaligen Laborkollegen des Labors 321 Herrn PD Dr. Christoph Forreiter, Herrn Dr. Michael Hoff, Herrn Dr. Frank Landgraf, Frau Siegrid Ranostaj und Herrn Sven Sagasser ein Loch, das nicht mehr zu füllen sein wird.

Den größten Dank gebe ich an meinen verstorbenen Vater, er erweckte das Interesse an den Naturwissenschaften in mir.

Lebenslauf

Name: Martin Kenneth Kraft
Wohnort Rendeler Str. 22
60385 Frankfurt
Geburtsdatum: 28.10.1972
Geburtsort: Würzburg
Eltern: Antje Kraft, Dr. Volker Kraft
Nationalität: deutsch
Familienstand: ledig

1978 / 1979 Vorschule Brüssel

1979 - 1983 Grundschule Taufkirchen bei München

1983 - 1992 Gymnasium Philipp-Reis-Schule Friedrichsdorf, Hessen

1992 **Abitur**

1993 - 1994 Studium der Chemie an der Johann Wolfgang Goethe-Universität

1994 - 1999 Studium der Biologie an der Johann Wolfgang Goethe-Universität

10/1996 **Vordiplomsprüfungen in Biologie**

Prüfer: Prof. Dr. G. Kahl (Botanik), Prof. Dr. G. Fleissner (Zoologie), Prof. Dr. E. W. Grabner (Chemie), Prof. Dr. W. Aßmus (Physik)

26.10.1998 **Diplomprüfungen in Biologie** in den Fächern Botanik (Hauptfach), Mikrobiologie, pharmazeutische Biologie

Prüfer: Prof. Dr. L. Nover, Prof. Dr. K. Weising (Botanik), Prof. Dr. F. W. Pons (Mikrobiologie), Prof. Dr. T. Dingermann (pharmazeutische Biologie)

- 1999 **Diplomarbeit** über das Thema: „Funktionsanalyse von Hsp90 in Pflanzen“; angefertigt an der Johann Wolfgang Goethe-Universität im Institut für Zellbiologie, unter Anleitung von Prof. Dr. L. Nover
- 29.10.1999 **Diplom in Biologie**, Johann Wolfgang Goethe-Universität
- 1999 - 2004 **Dissertation** über das Thema „Transkriptionelle Regulation des humanen endogenen Retrovirus HTDV/HERV-K“; angefertigt am Paul-Ehrlich Institut, Langen unter Anleitung von Herrn Prof. Dr. J. Löwer

Teilnahme an Kongressen und Veröffentlichungen:

M. Kraft, J. Löwer, R. Löwer (2001a). Transkriptionelle Regulation des humanen endogenen Retrovirus HTDV/HERV-K. Jahrestagung der Gesellschaft für Virologie 2001, 14.03. - 17.03. 2001 in Dresden.

M. Kraft, J. Löwer, R. Löwer (2001b). Transcriptional regulation of the human endogenous retrovirus HTDV/HERV-K: mutational analysis of the long terminal repeat. 13th International Workshop on Retroviral Pathogenesis. 26.9. - 29.10. 2001 in Lugano, Schweiz.

M. Kraft, J. Löwer, R. Löwer (2002). Promoter-relevant regions of the endogenous retrovirus HTDV/HERV-K LTR revealed by mutational analysis. Jahrestagung der Gesellschaft für Virologie 2002. 08.04. - 10.04. 2002 in Erlangen.

M. Kraft, J. Löwer, R. Löwer (2004). Transcriptional initiation of HERV-K LTR is mediated by Sp1/Sp3 in a tissue-specific context. (In Vorbereitung)

