Regulation of *Sleeping Beauty* Transposition in Vertebrate Cells

vorgelegt von

Diplom-Ingenieur

Oliver Walisko

aus Berlin

Von der Fakultät III- Prozesswissenschaften

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

- Dr.-Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Helmut Görisch

Berichter: Prof. Dr. Ulf Stahl

Berichter: Dr. Zoltán Ivics

Tag der wissenschaftlichen Aussprache: 26. Februar 2004

Berlin 2004

D83

To Katya

Acknowledgement

Many people have made their contribution to the work presented in this thesis, and I would like to use these few lines to thank all of them for their help, support and ideas.

Above all I want to thank my advisor Dr. Zoltán Ivics, the father of *Sleeping Beauty* and initiator of this project presented in this thesis. Thank you very much Zoltán for welcoming me into your lab and for involving me in this challenging project. Whenever possible, you let me the freedom to follow and to test my own ideas and to develop as an independent and responsible scientist. You were always prepared for valuable discussions and good advice. I appreciate it very much that you took all of us to the Keystone-Meeting in Santa Fe to give us the possibility to meet with other leading scientists on the transposition field. It was a very good time I had in your lab.

Many thanks also to Dr. Zsuzsanna Izsvák for sharing the gene chip data with me and giving good practical advice throughout the whole project.

I would also like to thank my supervisor Prof. Ulf Stahl who supported me throughout my study at the Technical University of Berlin and whose support and advice is continued in this work.

Many thanks to my colleaques and friends Dr. Christopher Kaufman and Csaba Miskey. You made work enjoyable for me and I hope for many more beers we will have together.

For excellent technical assistance I would like to thank Dora Fiedler, Eva Stüwe and Andrea Katzer. Dora has died on July 2003 following her fight with cancer but she will be always remembered.

Finally, I would like to thank my wife Ekaterina Walisko and for her love, care and absolute support this work is dedicated to her.

Abstract

In this study I used the *Sleeping Beauty* (SB) transposable element as a tool to probe transposon-host interactions in vertebrates. SB transposes efficiently in human cells and I have conducted a yeast two-hybrid screen to search for human proteins, which interact with the SB transposase.

Two human proteins could be identified to be specific interactors of the SB transposase: 1. The cell-cycle regulatory protein Miz-1 (Myc-interacting zinc finger protein 1) and 2. the HMG2L1 (high mobility group protein 2-like 1).

Interaction between SB transposase and Miz-1 results in decreased cyclin D1 promoter activity as revealed by transient transfection experiments. Downregulation of cyclin D1 expression causes retarded growth of transposase-expressing cells, due to a prolonged G1 phase of the cell-cycle. Transposase induced G1 arrest is associated with a decrease of cyclin D1/cdk4-specific phosphorylation of the retinoblastoma protein which is a central component of the cell-cycle machinery. Transposase mediated G1 arrest in proliferating cells is of importance for transposition of the element because a temporary arrest of cells in G1 by serum starvation enhances transposition. Hence, my findings support a model which links Sleeping Beauty transposition to the G1 phase of the cell-cycle, where the non-homologous end-joining pathway of DNA repair is preferentially active.

The relevance of the interaction between the SB transposase and HMG2L1 is not evident because the function of HMG2L1 is unknown. HMG2L1 shares structural similarities with the closely related transcriptional regulators SRY, Sox-4 and LEF-1 and I hypothesize that HMG2L1 is a transcription factor which regulates transcription of yet unidentified target genes. Transcriptional regulation mediated by SB transposase/HMG2L1 I expect to be beneficial for both the transposable element and the host.

Zusammenfassung

In dieser Studie wurde das *Sleeping Beauty* Transposon (SB) als Modell-Transposon verwendet, um Interaktionen zwischen Transposon und Wirtszelle in Wirbeltierzellen zu untersuchen. SB transponiert effizient in menschlichen Zellen. Um Proteinfaktoren zu identifizieren, die mit der SB Transposase in Wechselwirkung treten, wurde mittels der Hefe „Two-Hybrid" Technik eine menschliche cDNA-Bank mit der Transposase als Köderprotein untersucht.

Es konnten zwei menschliche Proteine identifiziert werden, die spezifisch mit der SB Transposase interagieren: Das Zellzyklus regulierende Protein Miz-1 (Myc-interacting zinc finger protein 1) und HMG2L1 (high mobility group protein 2-like 1).

Die Interaktion zwischen der SB Transposase und Miz-1 resultiert in verminderter Cyclin D1-Promoter Aktivität und führt somit zu einer verlängerten G1-Phase des Zellzyklus. Der Transposase induzierte G1-Arrest ist begleitet von einer reduzierten Cyclin D1/Cdk4-spezifischen Phosphorylierung des Retinoblastomaproteins. SB Transposase induzierter G1-Arrest in proliferierenden Zellen ist von Bedeutung für den Transpositionsvorgang, da ein vorübergehender G1-Arrest durch Serumentzug zu einer Erhöhung der Transpositionsrate führt. Dieser Mechanismus erlaubt es der Zelle zusätzliche Zeit zu gewinnen, um DNA Schäden, welche durch das Transposon hervorgerufen wurden zu reparieren, bevor sie in die nächste Zellgeneration übertragen werden.

Die Bedeutung der Wechselwirkung von der SB Transposase mit HMG2L1 ist spekulativ, da die biochemische Funktion von HMG2L1 unbekannt ist. Aufgrund der strukturellen Ähnlichkeit mit den Transkriptionfaktoren SRY, Sox-4 und Lef-1 wird vermutet, dass es sich bei HMG2L1 ebenfalls um einen Transkriptionsfaktor handelt.

Abbreviations

Ac	activator
AD	activation domain
AdML	adenovirus major late
AS	antisense
AT	ataxia-telangiectasia
ATP	adenosine 5´-triphosphate
*att*Tn*7*	attachment site of Tn*7*
BD	binding domain
bleo	bleomycin
BSA	bovine serum albumin
BTB/POZ	broad-complex, tram track, and Bric-a-brac/pox virus and zinc finger
Cdk	cyclin-dependent kinase
cDNA	complementary DNA
CDTA	trans-1,2-diaminocyclohexane-N,N,N´,N´-tetraacetic acid
ChIP	chromatin immunoprecipitation
CHO	Chinese hamster ovary
CIP	cyclin-dependent kinase interacting protein
CKI	cyclin-dependent kinase inhibitor
CMV	cytomegalovirus
CPRG	cholorophenol red-β-D-galactopyranoside
DDD	amino acid sequence of catalytic domain containing of three aspartic acids
DDE	amino acid sequence of catalytic domain containing two aspartic acids and one glutamic acid
DMEM	Dulbecco´s modified Eagle medium
DNA	deoxyribonucleic acid

DNA-PK_{CS}	catalytic subunit of DNA-dependent protein kinase
DR	direct repeat
Ds	dissociation
DSB	double-strand break
DTT	DL-Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EN	endonuclease
EST	expressed sequence tags
FACS	fluorescence-activated cell sorting
G1	gap phase 1
G2	gap phase 2
gag	group-specific antigen
gal	galactose
GAPDH	glyceraldehyde-phosphate dehydrogenase
GFP	green fluorescent protein
HIV	human immunodeficiency virus
HMG	high mobilty group protein
HOXD	homeo box D
HR	homologous recombination
HRP	horseradish peroxidase
HTH	helix-turn-helix
IN	integrase
INK	inhibitor of cyclin-dependent kinase
IR	inverted repeat
IRES	internal ribosome entry site

IS	insertion sequence
kan	kanamycin
kb	kilo bases
kDa	kilo Daltons
L1	LINE1
LEF	lymphoid enhancer-binding factor
LINE	long interspersed nuclear element
LTR	long terminal repeat
luc	luciferase
M	molar
MBP	maltose-binding protein
Miz-1	myc-interacting zinc finger protein 1
MLTF	major late transcription factor
mRNA	messenger ribonucleic acid
MU	mutator
n	sample size
N123	N-terminal region of the SB transposase encompassing the first 123 amino acids
NHEJ	non-homologous end-joining
NLS	nuclear localization signal
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
PR	protease
RAG	recombination-activating gene
Rb	retinoblastoma

RH	gene encoding for ribonuclease H
RLU	relative light units
RNA	ribonucleic acid
RNAi	RNA interference
RT	reverse transcriptase
S	synthesis phase
SB	Sleeping Beauty
SC	synaptic complex
SDS	sodium dodecyl sulfate
PAGE	polyacrylamide gel electrophoresis
sem	standard error of the mean
Ser	serine
SINE	short interspersed nuclear element
SOX	SRY-related HMG-box gene
SRY	sex-determining region Y
stdv	standard deviation
str	streptomycin
Tc1	transposon of *Caenorhabditis elegans* 1
TCF	T-cell factor
Tn	transposon
Tpase	transposase
tRNA	transfer ribonucleic acid
TSD	target-site duplication
Ty	yeast transposon
UTR	untranslated region
UV	ultraviolet

V(D)J	variable(diversity)joining
wt	wildtype

Table of contents

Acknowledgement.....1

Abstract.....2

Zusammenfassung.....3

Abbreviations.....4

1. Introduction.....13

1.1 The discovery of transposable elements.....13

1.2 Transposable elements.....14

1.2.1 DNA transposable elements.....14

1.2.2 Retrotransposons.....16

1.3 The mechanism of transposition reactions.....18

1.3.1 Conservative transposition (cut-and-paste).....19

1.3.2 Replicative transposition (copy-and-paste).....19

1.4 The biochemistry of transposition reactions.....21

1.5 Impact of transposable elements on the host genome.....22

1.6 Regulation of transposition.....23

1.6.1 Transposable elements are self-restraining.....24

1.6.2 Target site specificity in transposition.....26

1.6.2.1 Target site selection by element encoded proteins.....26

1.6.2.2 Target site selection by host encoded proteins.....27

1.6.3 Linking transposition to the host cell metabolism.....27

1.7 Tc1/*mariner* transposable elements and the *Sleeping Beauty* transposon.....28

1.7.1 The *Sleeping Beauty* transposable element.....29

1.7.1.1 Structure of the SB transposase.....30

1.7.1.2 The structure of the inverted repeats.....31

1.7.1.3 Mechanism of *Sleeping Beauty* transposition.....31

1.7.1.4 Applications of the *Sleeping Beauty* transposon system in genomics.....32

1.7.1.5 The *Sleeping Beauty* transposon as a vector system for gene therapy.....33

1.8 Goal of the project.....34

2. Materials and Methods.....35

2.1 Recombinant DNA technology.....35

2.2 Plasmids created in this work.....35

2.3 Cloning of the human HMG2L1 cDNA.....36

2.4 Primer sequences.....37

2.5 Plasmids that are described elsewhere.....38

2.6 Two-hybrid screen and interaction domain mapping.....38

2.7 MBP pull-down assay.....39

2.8 Cell culture and transfection.....39

2.9 Generation of HuH7[IRES-SB] and HuH7[IRES-K] cell lines.....40

2.10 *Sleeping Beauty* transposition assay in temporarily G1 arrested cells.....40

2.11 Reporter gene assays.....40

2.12 Cell-cycle and growth curve analysis.....41

2.13 FACS analysis.....42

2.14 Retinoblastoma protein phosphorylation.....42

2.15 Affymetrix gene chip hybridizations.....43

2.16 RNA isolation and Northern blotting.....43

2.17 Statistical analysis.....44

3. Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1.....45

3.1 Results.....46

3.1.1 Sleeping Beauty transposase downregulates cyclin D1 expression.....46

3.1.2 Cells expressing the Sleeping Beauty transposase exhibit retarded growth due to a prolonged G1 phase of the cell-cycle.....48

3.1.3 G1 arrest is associated with decreased levels of phosphorylated retinoblastoma protein 50

3.1.4 Sleeping Beauty transposase is not sufficient to alter cyclin D1 promoter activity.....51

3.1.5 A yeast two-hybrid screen identifies the Myc-interacting zinc finger protein 1 as a specific interactor of the Sleeping Beauty transposase.....52

3.1.6 Regulation of cyclin D1 and $p15^{INK4B}$ promoters by Miz-1 and Sleeping Beauty transposase.....55

3.1.7 A temporary G1 arrest enhances *Sleeping Beauty* transposition.....57

3.2 Discussion.....59

3.3 Future work.....69

4. The Sleeping Beauty transposase interacts with the human high mobility group 2-like 1 protein, an HMG-box protein with unknown function.....71

4.1 Results.....72

4.1.1 Yeast two-hybrid screen.....72

4.1.2. Cloning of the full-length HMG2L1 cDNA.....73

4.1.3 Detection of HMG2L1 transcripts in HeLa cells.....74

4.1.4 *In vitro* interaction between SB transposase and HMG2L1.....75

4.2 Discussion.....76

4.3 Future work.....79

5. Concluding thoughts.....81

Reference List.....82

Appendix.....94

cDNA and amino acid sequence of the Sleeping Beauty transposase.....94

cDNA and amino acid sequence of human Miz-1.....95

cDNA and amino acid sequence of human HMG2L1.....97

Raw data.....99

1. Introduction

1.1 The discovery of transposable elements

Pioneering work which lead to the concept of transposable genetic elements was done by the maize geneticist Barbara McClintock in the 1940s.

In corn (*Zea Maize*), McClintock identified a genetic factor as a regular site of chromosome breakage (McClintock 1945 and 1946) and she termed it the *Ds* (dissociation) locus. Breakage in the *Ds* locus uncovers recessive alleles that are present in the homologous chromosome (McClintock 1947), which lacks the *Ds* locus, and thereby leads to the loss of the acentric segment from the nucleus during mitosis (Fig. 1-1A).

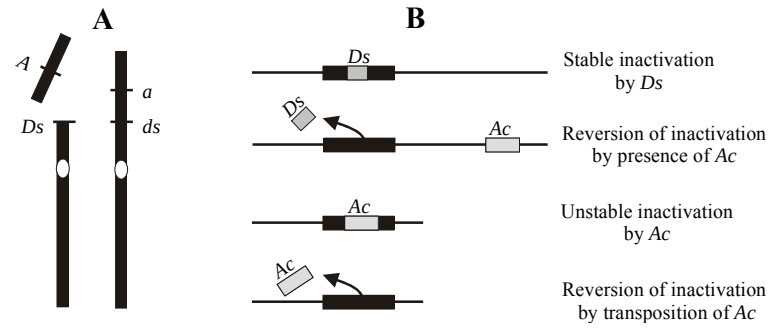


Fig. 1-1: Mutations caused by the *Ac*/*Ds* elements in *Zea Maize*.
(A) Loss of an acentric chromosomal fragment during mitosis. Acentric loss of a chromosomal fragment that carries the dominant allele *A* uncovers the recessive allele *a*. Chromosomal breakage is induced by *Ac* at the site of a *Ds* element (*ds* indicates the lack of *Ds*). The centromeres are shown as white circles.
(B) The effects of *Ac* and *Ds* on gene function. Shown is a hypothetical gene (black box). For details see text.

However, she recognized that the presence of the *Ds*-locus alone is not sufficient to generate dissociation mutations, and she postulated another genetically separate locus required to activate chromosome breakage at *Ds*. This she designated *Ac* (activator) (McClintock 1947). Her attempts to map the *Ac* locus were not successful, and thus she concluded that *Ac* locus may be able to change its position, a process which is now generally called transposition (McClintock, 1948). Similar findings she obtained for the *Ds* locus (McClintock 1948). Further genetic studies revealed that both *Ac* and *Ds* can insert into other genetic loci, thereby causing mutations (McClintock 1948, 1949 and 1951). However, mutations caused by *Ac* turned out to be unstable, whereas those caused by *Ds* were stable in the absence of *Ac* (Fig. 1-1 B). In the presence of *Ac*, *Ds*-induced mutations became unstable, suggesting again that the activity of *Ds* is regulated by *Ac*. Unstable insertion mutations of *Ac* or *Ds* into genes

revert during plant and kernel development by excision of the element, which generally restores gene function and results in revertant cells on a mutant background (variegation).

McClintock designated these (transposable) elements *controlling elements* since they modify gene expression as a result from insertion at or near a gene (McClintock 1956). The *controlling elements*, which McClintock studied for many years and finally could demonstrate their transposition (McClintock 1948), are now designated the *Activator-Dissociation* (*Ac-Ds*) family of maize transposable elements.

McClintock´s discovery of transposable controlling elements in maize, and the fact that she has realized that these elements can induce mutations and alter gene expression, was recognized with the award of the Nobel Prize.

1.2 Transposable elements

Transposable elements can be defined as discrete DNA segments that can move between many non-homologous sites within a genome. This movement is generally called transposition, and can contribute to altered gene expression and recombination. In virtually any organism investigated to date, including bacteria, lower eukaryotes and higher eukaryotes such as vertebrates and plants a great diversity of transposable elements was found.

Transposable elements can be divided into two major groups based on the transposition intermediates: DNA transposable elements transpose through a DNA intermediate which clearly distinguishes them from retrotransposons that employ RNA intermediates during transposition.

1.2.1 DNA transposable elements

Prokaryotic transposable elements were discovered as agents causing highly polar mutations in *E. coli* (Saedler and Starlinger, 1967a, Saedler and Starlinger, 1967b), which were found to be due to insertions of DNA (Jordan *et al*., 1968). These inserted pieces of

DNA were called insertion sequences (ISs). ISs represent the structurally most compact type of DNA transposable elements, and are composed of short terminal inverted repeats flanking the transposase gene (Fig. 1-2 A). Usually no function other than the transposase is encoded by the element, which distinguishes the ISs from transposons.

Transposons encode, in addition to the transposase, another function, which is often conferring antibiotic resistance. Transposons are flanked by inverted or direct repeats which are in many cases structurally similar to IS elements (Fig. 1-2 B).

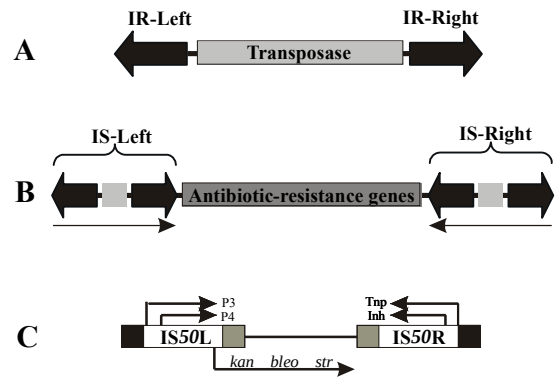


Fig. 1-2: The structure of DNA transposable elements.
(A) Simplified structure of a hypothetical insertion sequence. The central transposase gene is flanked by terminal inverted repeats (IRs are indicated by black arrows). The IRs contain the binding-sites for the transposase and terminal base pairs that are required for transposase-mediated cleavage. **(B) Simplified structure of a hypothetical transposon.** The transposon consists of antibiotic-resistance genes (dark grey box) that are flanked by two copies of an insertion sequence (IS-Left and IS-Right) which carry a gene encoding for the transposase (grey boxes). The thin lined black arrows indicate the inverted orientation of the insertion sequences. **(C) Structure of the bacterial Tn*5* transposon.** Two IS*50* elements flank three antibiotic resistance genes [*kanamycin* (*kan*), *bleomycin* (*bleo*) and *streptomycin* (*str*)]. IS*50* is bordered by two 19-bp sequences (grey and black boxes) that contain the transposase binding sites. IS*50*R encodes the transposase (Tnp) and an inhibitor (Inh). IS*50*L contains a premature stop codon which results in inactive transposase and inhibitor proteins (P3 and P4). The terminal inverted repeats of the IS elements are indicated by grey and black boxes (adapted from Reznikoff *et al*., 1999).

For example, the bacterial transposon Tn*5* (Fig. 1-2 C) is bounded by two, nearly identical IS*50* insertion sequences: IS*50*L (left) and IS*50*R (right) (Rothstein *et al*., 1980). IS*50*R encodes the transposase which catalyses transposition and a truncated transposase, which is an inhibitor of transposition (Krebs and Reznikoff, 1986). IS*50*L contains a premature stop codon, which results in inactive tranpsosase and inhibitor proteins (Rothstein and Reznikoff, 1981). Flanked by the two insertion sequences are three antibiotic resistance genes conferring resistance to kanamycin, bleomycin and streptomycin (Mazodier *et al*., 1985). Although DNA transposons were initially discovered in prokaryotes, they can be also found in eukaryotes. For example, 2 to 3% of the human genome is composed DNA transposons (Lander *et al.*, 2001), but there is no evidence for recent activity.

1.2.2 Retrotransposons

Until the early 1970s, the generally excepted model of information flow within a cell was exclusively from DNA to RNA to protein. This had to be reconsidered when Baltimore (1970) and Temin (1970) reported a RNA-dependent DNA polymerase activity that is present in the virions of retroviruses. Reverse transcription was regarded to be a unique feature of retroviruses until the early 1980´s, when Nishioka *et al*. (1980) and Hollis *et al*. (1982) described the presence of partial duplicates of structural genes lacking intervening sequences (introns). Such genes are generally termed processed pseudogenes (Fig-1-3).

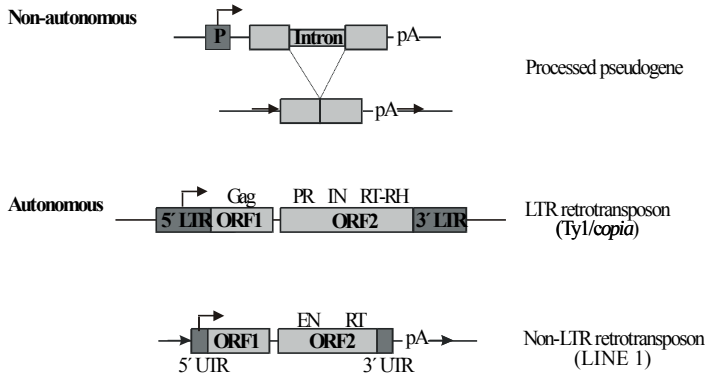


Fig. 1-3: Structures of some retrotransposon types.
Processed pseudogenes are nonfunctional coding sequences, which originate from retrotransposition events of cellular, polyadenylated (pA) mRNA. **LTR retrotransposons** consist of long terminal repeats (LTR) similar to those of retroviruses. The LTRs flank two open reading frames (ORF). ORF1 encodes the group-specific antigen and ORF2 encodes for protease (PR), integrase (IN), reverse transcriptase and RNaseH (RT-RH). **Non-LTR retrotransposons** consist of a 5´ untranslated region that has promoter activity, which is required to drive transcription of element encoded genes. ORF1 encodes a nucleic acid binding protein of unknown function. ORF 2 encodes for an endonuclease (EN) and a reverse transcriptase (RT).
Target site duplications that result from integration of the elements into genomic DNA are shown as arrows

Removal of introns was found to strictly follow the G-T/A-G splicing rule of RNA, and their origin must be from reverse transcription resulting in a cDNA copy of the processed RNA. Investigation of the sequence of human pseudogenes revealed that they are flanked by direct repeats (Van Arsdell *et al*., 1981), a hallmark of DNA transposition (Grindley and Sherratt, 1978 and references therein). This suggests that processed pseudogenes represent a class of mobile elements that transpose through an RNA intermediate, a hallmark component of a process termed retrotransposition.

Retrotransposons can be classified based on their structural properties and their ability to catalyze their own movement (autonomous elements) or dependence on a catalytic activity supplied *in trans* by another element (non-autonomous elements).

Autonomous LTR retrotransposons (Fig. 1-3) bear long terminal repeats (LTRs) which

flank at least two ORFs similar to the *gag* and *pol* genes of retroviruses. LTR-retrotransposons comprise two major subfamilies: Ty1/*copia* and Ty3/*gypsy*, based on the founder elements, found in the yeast *Saccharomyces cerevisiae* (Ty1 and Ty3) and in *Drososopila* (*copia* and *gypsy*) (Boeke *et al*., 1985; Hansen *et al*., 1988; Mount and Rubin, 1985; Marlor *et al*., 1986). Elements of both subfamilies can be distinguished based on the structure of ORFs within the element (Xiong and Eickbush, 1990). LTR-retrotransposons are widely distributed in eukaryotes, and can be found in yeast (Boeke *et al*., 1985; Hansen *et al*., 1988), fungi (Capello *et al*., 1985), insects (Mount and Rubin, 1985; Marlor *et al*., 1986), plants (Voytas and Asubel, 1988) and humans where LTR elements make up 8.3 % of the genome (Lander *et al.*, 2001). Besides LTR retrotransposons, there is another class of autonomous retrotransposons, which do not possess LTRs (non-LTR retrotransposons). Two types of non-LTR retrotransposons can be destinguished: SINEs (short interspersed nuclear elements) and LINEs (long interspersed nuclear elements).

SINEs are relatively small, and highly repeated in genomes. Most SINEs are thought to be derived from tRNAs (Daniels and Deininger, 1985). One dominant form of SINEs in the human genome are the *Alu* elements, which were originally identified as repetitive DNA elements in human DNA renaturation curves (Houck *et al*., 1979). *Alu*s contain a recognition site for the restriction enzyme *Alu*I (hence the name) (Houck *et al*., 1979) and are estimated to represent 10.6% of the total human genome size (Lander *et al.*, 2001).

As suggested by the name, LINEs are longer than SINEs. Full-length LINEs contain two open reading frames (Fig. 1-3). One appears to be equivalent to the *pol* (reverse transcriptase) gene of retroviruses. The major non-LTR element in mammals, LINE-1 (or L1) element, is represented in the human genome with 500 x 10^3 copies, making 16.9 % of the genome (Lander *et al.*, 2001).

Alu elements and the L1 retrotransposons are currently active and can contribute to disease.

1.3. The mechanisms of transposition reactions

The movement of a transposable element can be accomplished in two different ways with regard to the fate of the original element (Fig. 1-4).

During conservative (cut-and-paste) transposition, the element gets excised from the donor locus and is subsequently reinserted elsewhere. This does not result in the amplification of the element. Elements which employ this strategy include Tn*5* (Goryshin and Reznikoff, 1998), Tn*7* (Bainton *et al*., 1991), Tn*10* (Bender and Kleckner, 1986), the *Drosophila* P element (Kaufman and Rio, 1992) and members of the Tc1/*mariner* family (Plasterk *et al*., 1999). The double-strand break (DSB) caused by the excision of the element will be repaired by the host DNA repair machinery. This occurs in the simplest case by rejoining the ends (see also section 1.3.1), which leaves behind a typical footprint (Fig. 1-4). However, DSBs can also be repaired by homologous recombination, which was shown for the *Drosophila* P element (Engels *et al*., 1990) and also for Tc1 from *C. elegans* (Plasterk, 1991).

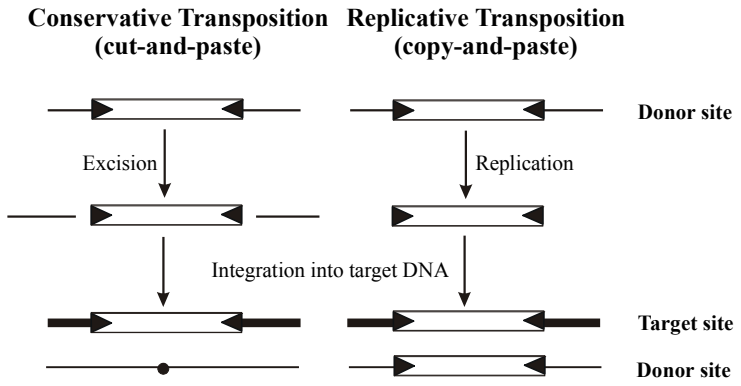


Fig. 1-4: Schematic depiction of the two modes of transposition. During conservative transposition the element is excised from the donor site (thin line) and integrated into a new target DNA (thick line). The broken donor site is repaired by host factors which results in a typical footprint (black circle).
Replicative transposition requires amplification of the element by transcription followed by reverse transcription. The amplified element gets inserted elsewhere in the genome (thick line).

Another mode of transposition is replicative (copy-and-paste) transposition (Fig. 1-4), in which the transposon does not get excised from its donor locus, but instead a copy is produced that inserts elsewhere in the genome. Thus, replicative transposition leads to an increase in the copy number of the transposon within a genome. Elements that use replicative transposition, include retrotransposons (Boeke and Chapman, 1991) and the bacteriophage Mu (Mizuuchi and Craigie, 1986).

1.3.1 Conservative transposition (cut-and-paste)

Physical separation of the transposable element from the donor locus is a hallmark of conservative (cut-and-paste) transposition. As shown in Fig. 1-5, three distinct functional steps can be distinguished during conservative transposition: 1) cleavage of the donor DNA (Excision), 2) joining of the transposon to target DNA (Integration) and 3) host-mediated gap repair.

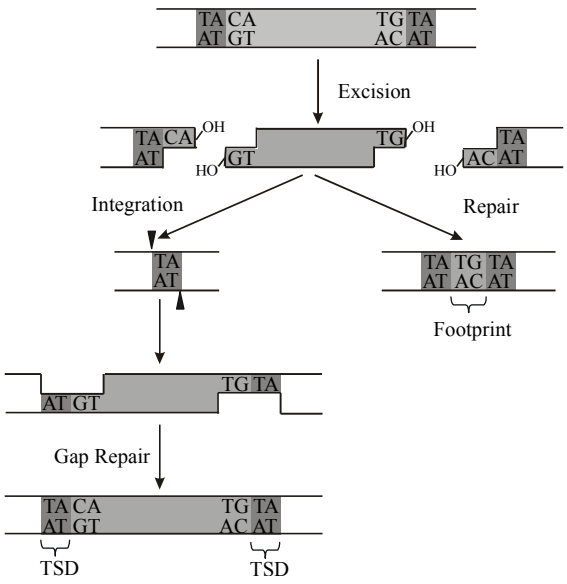


Fig. 1-5: Model for Tc1/*mariner* transposition. The element (grey box) is excised from its genomic location by staggered cuts, which are introduced by the transposase. Second-strand cleavage occurs two nucleotides inwards to the element (for *Sleeping Beauty* three nucleotides). Transposon excision leaves behind two-nucleotide long 3´-overhangs. Repair of the broken DNA ends by host factors generates transposon "footprints". The transposon integrates into a TA-target site, which generates single-stranded gaps. The single-stranded gaps are subjected to DNA repair, which generates target site duplication (TSD).

During donor cleavage the transposon DNA is completely separated from its flanking DNA. The transposon ends are then connected to the target DNA by nucleophilic attack of the exposed 3´-OH ends at staggered positions of the target DNA. This results in an inserted transposon, which is flanked by short, single stranded gaps. In the last step, the single-stranded gaps are repaired by the host repair machinery, resulting in short direct repeats that differ in their length and are characteristic for the element. Repair of the broken DNA at the donor site by rejoining the ends creates a footprint.

1.3.2 Replicative transposition (copy-and-paste)

During replicative transposition, the transposon remains attached to the donor DNA molecule and a copy of the transposon is inserted into target DNA. For both integrated retroviruses and retrotransposons, the event that separates the element from the donor

molecule is transcription. The RNA is reverse transcribed by an element-encoded reverse transcriptase into a cDNA copy that is blunt-ended and can be integrated into the target DNA.

Prior to integration of the cDNA, retroviral integrases remove base pairs (usually two) from the cDNA ends that creates single-stranded 5´-overhangs. This step is called „processing“, and can be regarded in analogy to the initial cut introduced by transposases. This step is not required for LTR retrotransposons and the reason for this difference is not known.

In the integration step, the exposed terminal 3´-OHs act as nucleophiles that cleave the phosphodiester bonds of the target DNA in staggered positions and simultaneously links the 3´ end of the viral (transposon) DNA to the 5´-end of the target DNA. The integration process is completed by the removal of the 5´-overhangs (viral DNA), and the single stranded gaps are repaired by host repair enzymes (Brin *et al*., 2000, Downs and Jackson, 1999; Yoder and Bushman, 2000), which results in short target site duplications.

Taken together, cut-and-paste DNA transposons, LTR-retrotransposons and retroviruses share a common mechanism that enables them to transpose within a genome. This mechanism essentially includes cleavage of the donor DNA, joining to target DNA and as a last step repair of newly formed transposon-target DNA junctions by the host machinery. Whereas the target joining step and the repair of the single-stranded gaps are very similar among these elements, major differences can be seen in the donor cleavage step. DNA cut-and-paste transposons require two cuts to get excised from the donor locus, just a single cut at the 3´-end is sufficient for retroviruses and no cut is required for LTR retrotransposons.

Some DNA transposable elements have adopted a strategy similar to retroviruses, which influences the final transposition product. For example, during phage Mu transposition only the 3´-ends are cleaved, while the donor DNA remains attached to the 5´-end of the transposon. Subsequent nucleophilic attack of the 3´ends on the target DNA creates a transposition intermediate in which the donor DNA is covalently linked to the target DNA,

creating a potential replication fork (Shapiro intermediate) (Shapiro, 1979). Replication results in two directly repeated copies of the transposon that are separated by donor and target DNA (cointegrate). As a result, the element has transposed through replicative transposition. However this mechanism seems to be an exception rather than a rule.

1.4. The biochemistry of transposition reactions

During transposition, a polynucleotide chain is transferred, and hence this reaction is a polynucleotidyl transfer reaction.

In the first step, a polynucleotide 5′-phoshosphate of an intact DNA strand is attacked by an incoming water molecule, which results in hydrolysis of the phosphodiester bond and generates a free 3′OH end. In the strand-transfer reaction step, the cleaved 3′OH ends of the transposon act as nucleophiles that attack phophodiester bonds located in the target DNA, thereby creating a new transposon-target junction (Mizuuchi and Craigie, 1986). Both DNA cleavage and strand-transfer are accompanied by a stereo chemical inversion of a chiral phosphothioate that substitutes regular phosphodiester bonds in the target DNA (Engelman *et al*., 1991). This suggests that the reaction occurs by an in-line transesterification reaction.

Neither of the steps in transpositional recombination requires hydrolysis of high-energy cofactors such as ATP (Morisato and Kleckner, 1987; Vos *et al*., 1996; Bushman and Craigie, 1991). The energy, which is required to form the new phosphodiester bond, comes from the exergonic cleavage reaction of the target DNA.

A straightforward reaction mechanism for phosphoryl transfer reactions is general acid/base catalysis. A proton must be removed from the incoming nucleophile and transferred to the leaving group. Alternatively, electron-starved atoms such as divalent metal ions can play a role in phosphoryl transfer reactions. Metal ions polarize the H-O bond in a water molecule, and thereby render it into a potent nucleophile. Hence, metal-bound water can function as a general acid/base and promote general acid/base catalysis.

Comparison of the crystal structures of the resolvase RuvC, HIV-1 integrase and RNase H (Dyda *et al*., 1994; Ariyoshi *et al*., 1994) revealed a remarkable similarity in the topology of the active sites which include a set of three acidic amino acids at essentially identical positions relative to the tertiary structure. These three invariant acidic amino acids (two aspartic acids and a glutamic acid), also known as the DDE triad (Kulkosky *et al*., 1992), form the active site of a large group of recombinases, including transposases and integrases (Doak *et al.*, 1994; Landree *et al*., 1999; Polard and Chandler, 1995; Kulkosky *et al.,* 1992). The DDE motif is required for both the DNA cleavage and the strand-transfer step and mutation of any one of the three residues results in a severe reduction in activity (Kulkosky *et al*., 1992; Engelman and Craigie, 1992; Bolland and Kleckner, 1996). The crystal structure of retroviral HIV-1 integrase revealed that the carboxylate residues of the two aspartates are important for positioning magnesium ions in the active site (Goldgur *et al*., 1998).

1.5 Impact of transposable elements on the host genome

Mobile elements can cause altered gene expression and genetic recombination. Insertion of a transposable element into the protein-coding region of a gene (exon) can disrupt gene function. For example, bacterial insertion sequences were identified as DNA insertions in the *E. coli gal*-operon (Jordan *et al*., 1968). The *Drosophila* P element was found to cause mutations at the *white* locus (Rubin *et al*., 1982) and the DNA transposon Tc1 is the main cause for mutations in the nematode *C. elegans* (Emmons *et al*., 1983; Moerman and Waterston, 1984). In humans, insertions of the L1 retrotransposon in exon 14 of factor VIII gene were shown to result in hemophilia A (Kazazian *et al*., 1988). Furthermore, L1 can retrotranspose DNA flanking the 3´ end of the element (Holmes *et al*., 1994; Moran *et al*., 1999), which leads to segmental duplications of portions of genomic sequences.

In contrast to insertion of a transposable element into exon sequence of a gene, the element can be inserted into upstream regulatory sequences and thereby alter the regulation

and expression of flanking genes. For example, insertion between the core promoter and adjacent enhancer regions would increase the distance between these regions and thus negatively affect promoter activity. In contrast, Ty elements in *Saccharomyces cerevisiae* were shown to act as enhancers when inserted into the 5´ regulatory region of a gene (Errede *et al*., 1987) and thereby activate transcription of adjacent genes. This behaviour is explained by sequences within the LTR region, which show homology to the SV40 enhancer. In the *Drosophila ananassae*, insertions of the *tom* retrotransposon into numerous different genes has led to mutant phenotypes in the eye (Tanda and Corces, 1991), which suggests the presence of a tissue-specific enhancer in the transposable element.

Taken together, random insertion of transposable elements into genomic sites can alter gene expression.

Another damaging aspect of transposable elements is that repeated, dispersed copies of homologous sequences can promote secondary rearrangements, which can result in deletions, duplications and inversions. This potential of dispersed transposon insertions to promote homologous recombination can be even more damaging to the genome than a *de novo* insertion which can be tolerated if it does not provide a major negative selection. Newly inserted elements provide a source for unequal homologous recombination for millions of years if this threat is not eliminated during evolution. Furthermore, recombination between two elements will affect multiple host genes, which is likely to be more deleterious than a *de novo* insertion which affects only one gene.

1.6 Regulation of transposition

Because transposition is potentially deleterious to the host, it is advantageous for the survival of both the host and the transposable element to develop strategies that reduce the negative effects on the host but ensure proliferation of the element. In principal, there are three strategies that can be employed to reduce the damaging impact on the host genome: 1)

transposable elements are self-restraining and transpositional activity is maintained at low levels; 2) transposable elements are targeted to "safe" places in the genome where insertions are non-deleterious, and 3) the transposition process is linked to a metabolic state of the host organism in which the host can better cope with the consequences of transposition.

1.6.1 Transposable elements are self-restraining

It is in the interest of transposable elements to evolve mechanisms that limit their own amplification and therefore minimize deleterious impacts on the host. In other words, there is no positive selection for highly active transpositional machineries in nature. The fact that it is relatively easy to isolate hyperactive transposase mutants by relatively simple genetic screens (Wiegand and Reznikoff, 1992; Goryshin and Reznikoff, 1998; Beall *et al*., 2002; Lampe *et al*., 1999) reveals that wildtype transposases have a suboptimal structure which can not catalyze transposition at high efficiencies.

Another example for self-restraint transposition comes from the bacterial Tn*5* transposon, in which the transposase acts primarily on the DNA that expresses the transposase (*cis* preference) (Johnson *et al*., 1982). *Cis* preference limits transposition to the transposon copy, which is expressing the transposase, whereas other copies remain unaffected. *Cis* preference could also be observed for Tn*10* and Tn*903* (Morisato *et al*., 1983; Derbyshire *et al*., 1990). Self-restraint transposition is achieved by some elements, which encode an inhibitor protein in addition to the transposase. For example, Tn*5* encodes a truncated transposase version that inhibits transposase activity by forming inactive heterodimers (de La Cruz *et al*., 1993). Similarly, the *Drosophila* P element encodes an inhibitory protein, which is also a truncated form of the transposase arising from alternative splicing (Misra and Rio, 1990) and that represses P element transposition in somatic cells and thus limits P element transposition to the germ line where the transposon to be passed to the next generation and damage to somatic cells can be avoided.

The functional transposase protein can also act as an inhibitor of transposition, a phenomenon that is called overproduction inhibition. Overproduction inhibition has been described for the *mariner* elements *Himar*1 (Lampe *et al*., 1998) and *Mos*1 (Lohe and Hartl, 1996), but also for Tn*5* (Wiegand and Reznikoff, 1992), the *Ac* element of maize (Kunze *et al*., 1993) and the Tc1-like element *Sleeping Beauty* (Yant *et al*., 2000). For Tn5, the formation of catalytically inactive transposase multimers was suggested to mediate this effect (Weinreich *et al*., 1994), whereas for the *Ac* element insolubilty of the transposase at high concentrations could play a role (Kunze *et al*., 1993). As the copy number of these elements increases with time, overproduction inhibition results in a decrease in net transposase activity and thus limits further amplification.

Most Tc1/*mariner* elements found in nature are inactive due to the accumulation of mutations (Maruyama *et al*., 1991; Robertson, 1993), a process called vertical inactivation and which can be explained by the lack of positive selection pressure for functional transposases. Alternatively, it seems to be possible that there is a positive selection pressure for the accumulation of defective elements because they do not damage the host. There are two possible explanations for how these defective transposases could negatively affect transpositional activity: 1) these mutant transposases might be defective in transposition but still could mediate overproduction inhibition or 2) the mutant transposases compete with wild-type transposases for binding sites located in the transposon ends.

Although elements in natural populations accumulated inactivating mutations, their transposase binding sites are often retained. Defective elements bind the transposases but are incompetent for transposition, a process termed transposase titration.

Taken together, there is a great variety of mechanisms that limit the transpositional activity to a minimal level. All these regulatory mechanisms have in common that regulation appears to be an intrinsic property of the element itself. However, for some elements low basal activity can be greatly enhanced by host-encoded factors.

For example, the host protein HU plays a role in bacteriophage Mu transposition in *E. coli* by bending the transposon DNA and bringing together the transposase binding sites (Lavoie and Chaconas, 1990). Bacteriophage λ also requires a host protein, integration host factor (IHF), for efficient recombination (Goodman and Nash, 1989). One transposition-like process in vertebrates that has been extensively studied is immunoglobulin gene rearrangement, also termed V(D)J recombination (Fugmann *et al*., 2000). V(D)J recombination is mediated by the RAG1/2 recombinase (transposase) complex and requires the host-encoded DNA-bending protein HMGB1/2 (Aidinis *et al*., 1999; van Gent *et al*., 1997). Similarly, HMGB1 was greatly enhances transposition efficiency of the *Sleeping Beauty* transposon (Zayed *et al*., 2003).

1.6.2 Target site specificity in transposition

Another level of regulation for some transposable elements is to insert into "safe" places in the genome. Safe places can be: 1) upstream promoter regions of transcribed genes, which reduces the risk for the host to suffer from insertions into essential genes; 2) genes that are represented in the genome with many copies and insertion will not result in complete inactivation of gene function; and 3) regions of transcriptionally silenced chromatin.

However, most elements exhibit low target site specificity and their distribution in the genome is essentially random.

1.6.2.1 Target site selection by element encoded proteins

The Tn*7* transposon is highly specialized to insert into a single sequence motif in the *E.coli* genome called the attachment site of Tn*7*, *att*Tn*7* (Lichtenstein and Brenner, 1982), which is not deleterious to the host. A particular transposase subunit is responsible for targeting (Bainton *et al*., 1993) and excision of the element only occurs if there is an *att*Tn*7* site available for insertion (Bainton *et al*., 1993).

Furthermore, Tn*7* does not insert into DNA that already contains Tn*7* (Arciszewska *et al*., 1989; Stellwagen and Craig, 1997), a phenomenon that is called target immunity. Target immunity helps to avoid multiple copies of the element in the same DNA molecule, which might result in deleterious recombination between the elements. Target immunity was also observed for the bacterial transposons Tn*3* and Mu (Maekawa *et al*., 1996; Adzuma and Mizuuchi, 1989).

1.6.2.2 Target site selection by host-encoded proteins

For some elements, such as the yeast retrotransposons Ty1, Ty3 and Ty5, host-encoded proteins direct the transposable element to a particular target area in the genome where integration can be well tolerated by the host.

For example, the Ty3 element always inserts a few nucleotides upstream of the transcriptional start site of genes which are transcribed by RNA polymerase III (Chalker and Sandmeyer, 1992), and targeting is mediated by RNA polymerase III transcription factors (Kirchner *et al*., 1995). This mechanism helps to avoid insertion into essential genes by directing integration to the upstream region of genes encoding for tRNA, which are present in the genome in many copies.

The yeast retrotransposon Ty5 targets transcriptionally silenced heterochromatin in the yeast genome, including telomeres and silent mating type loci (Zou *et al*., 1996). Targeting is achieved by a domain in the integrase protein that interacts with Sir4p, a structural component of silent chromatin (Xie *et al*., 2001).

1.6.3. Linking transposition to the host cell metabolism

Transposition can be linked to the host cell metabolism by which certain host functions can be exploited to modulate transposition. For example, V(D)J recombination requires the host-encoded components of the non-homologous end-joining (NHEJ) pathway

of DSB repair (Jackson and Jeggo, 1995). Selective degradation of the RAG2 recombinase has been proposed to confine V(D)J recombination to the G1 phase of the cell-cycle (Li *et al*., 1996; Lin and Desiderio, 1993) where NHEJ is preferably active (Takata *et al*., 1998).

1.7 Tc1/*mariner* transposable elements and the *Sleeping Beauty* transposon

Members of the Tc1/*mariner* superfamily of transposable elements are probably the most widespread DNA transposons in nature, and have been found in many species including vertebrates (for review see Plasterk *et al*., 1999). The founding element, Tc1, was initially detected in the nematode *C. elegans* (Emmons *et al*., 1983) where it was found to be the main cause of spontaneous mutations (Eide and Anderson, 1985). Several Tc1-related elements were discovered in other species including the *mariner* element, which was first described genetically in *Drosophila mauritiana* (Jacobson and Hartl, 1985).

Tc1/*mariner* elements are DNA based transposons that 1.3 to 2.4 kb in size and contain a transposase gene that is flanked by terminal inverted repeats. Conservation in amino acid sequence between different transposases is modest (approximately 15% identity) but all contain the typical DDE (Tc1-like elements) or DDD (*mariner*-like elements) amino acid triad, which forms the catalytic center. Transposable elements belonging to the Tc1/*mariner* class all transpose via a cut-and-paste mechanism as shown in Fig. 1-5 (page 19) and integration always occurs at a TA dinucleotide (Van Luenen *et al*., 1994; Ivics *et al*., 1997; Lampe *et al*., 1996; Vos *et al*., 1996).

It was found that no species-specific protein other than the transposase is required for transposition (Vos *et al*., 1996; Lampe *et al*., 1996) suggesting that both excision and integration are catalysed solely by the transposase. These simple reaction requirements explain why Tc1/*mariner* elements have such a broad distribution. Various Tc1/*mariner* elements were shown to transpose in a wide variety of species (Ivics *et al*., 1997; Fadool *et al*., 1998; Sherman *et al*., 1998; Schouten *et al*., 1998).

However, actual transposition of these elements has in most cases not been detected, and although the human genome contains Tc1/*mariner* transposons, all of them are inactive due to the accumulation of mutations (Auge-Gouillou *et al*., 1995; Oosumi *et al*., 1995) Due to this not a single active element has been isolated from vertebrates, hindering studies on the basic mechanism of DNA transposition and transpositional regulation in vertebrate organisms.

1.7.1 The *Sleeping Beauty* transposable element

In an attempt to derive an active Tc1-like transposon from vertebrates, sequence alignment of inactive Tc1-like sequences found in different fish species resulted in a consensus sequence for a likely functional Tc1-like transposable element (Ivics *et al*., 1997). Successive removal of inactivating mutations generated an active element, which was named the *Sleeping Beauty* transposon (Ivics *et al*., 1997). Therefore, *Sleeping Beauty* (SB) can be best characterized as a synthetic, Tc1-like transposon that was reconstructed from multiple dead copies of a once active fish transposable element. The overall structure of the *Sleeping Beauty* transposon is shown in Fig. 1-6A.

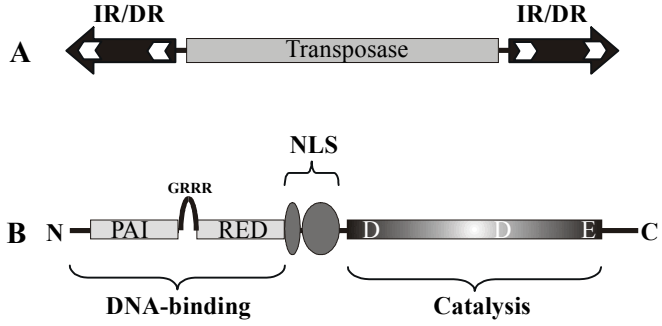


Fig. 1-6: Structure of the *Sleeping Beauty* transposable element.
(A) Structure of the *Sleeping Beauty* transposon. Nearly identical terminal inverted repeats (black arrows) flank a gene encoding the transposase. The inverted repeats (IR) contain two binding sites (white arrows) for the transposase as direct repeats (DR) resulting in the characteristic IR/DR structure. **(B) Structure of the Sleeping Beauty transposase.** The transposase carries an N-terminal, paired-like DNA-binding domain consisting of two helix-turn-helix motifs, PAI and RED, with a GRRR-motif (AT-hook) between them. The RED subdomain overlaps with a bipartite nuclear localization signal (NLS), which is followed by the catalytic DDE domain responsible for DNA cleavage and joining reactions

A central synthetic gene encoding the transposase is flanked by terminal inverted repeats. Each inverted repeat of SB contains two transposase binding sites that form short, 15-20 bp direct repeats.

SB catalyzes efficient cut-and paste transposition (Fig. 1-5; page 19) in cells of different vertebrate classes in tissue culture (Izsvák *et al*., 2000) and in somatic (Yant *et al*., 2000) as

well as germline tissues of the mouse *in vivo* (Fischer *et al*., 2001). Therefore, SB is a valuable tool for functional genomics in several model organisms (Fischer *et al*., 2001; Plasterk *et al*., 1999), and a potential vector for gene therapy in humans (Yant *et al*., 2000). In addition, this element is very useful to study DNA transposition in vertebrate cells.

1.7.1.1 Structure of the SB transposase

In order to catalyze excision and integration, Tc1/*mariner* transposases must include two functional domains that are conserved in the entire Tc1/*mariner* superfamily (Plasterk *et al*., 1999): 1) an N-terminal DNA-binding domain which specifically recognizes the transposon ends, and 2) the catalytic center (Fig. 1-6 B).

The N-terminal region of the transposase that mediates specific recognition of the transposon ends consists of a bipartite DNA-binding domain that encompasses two helix-turn-helix (HTH) motifs (Pietrokovski and Henikoff, 1997; Vos and Plasterk, 1994; Schukkink and Plasterk, 1990, Ivics *et al*., 1997, Izsvák *et al*., 2002). The first of these HTH-motifs is similar in both structure and amino acid sequence to the paired domain of some transcription factors (Vos and Plasterk, 1994; Franz *et al*., 1994; Ivics *et al*., 1996). The two HTH motifs are separated by a GRPR-like motif (GRRR) that shares similarity to an AT-hook (Izsvák *et al*., 2002) which was found to be responsible for minor groove contacts in the Hin invertase of *Salmonella* (Feng *et al*, 1994) and the RAG1 recombinase (Spanopoulou *et al*., 1996). The second HTH motif is followed by a bipartite nuclear localization signal (NLS) that enables the transposase to take advantage of the receptor-mediated nuclear-uptake machinery of the host cell. The NLS is followed by the catalytic domain of the transposase that is characterised by the conserved DDE domain that marks the active site of so many transposase proteins (see section 1.4.).

1.7.1.2 The structure of the inverted repeats

The SB transposase gene is flanked by inverted repeats that contain the binding sites for the SB transposase (Fig. 1-6A; page 29). The IRs of *Sleeping Beauty* are 200-250 bp long and each contains short, 15-20 bp direct repeats (DR), which are the transposase binding sites (Ivics *et al*., 1997). This organization of inverted repeats, termed IR/DR, is conserved among several Tc1-like elements, with the exception of Tc1 itself and the *mariner*-like elements, which contain only a single binding site (Izsvák *et al*., 1995, Plasterk *et al*., 1999). Efficient *Sleeping Beauty* transposition is dependent on the presence and integrity of all four transposase binding sites (Izsvák *et al*., 2000). However, the four binding sites differ from each other in both size and sequence (Ivics *et al*., 1997). Furthermore, the IRs are not perfect inverted repeats since the left one contains, in addition to the two binding sites, an extra transposase binding "half-site" (Izsvák *et al*., 2002), which acts as a transpositional enhancer, suggesting that the IRs are functionally distinct.

1.7.1.3 Mechanism of *Sleeping Beauty* transposition

A complete transposition event encompassing excision from the donor site and integration into a target site can be subdevided into the following steps: 1) recognition and binding of the substrate (IR/DRs of the transposon) by the transposase; 2) formation of a stable synaptic complex (SC) in which the transposon ends are held together by the transposase subunits; 3) excision of the element by the transposase; 3) binding of the SC to the target site; 4) nucleophilic attack of the liberated 3´-OH group on a phosphodiester bond located in the site of integration and 5) repair of transposon-induced DNA lesions by the host-repair machinery.

A hallmark of all transposition reactions is a strict requirement of a coordinated assembly of a synaptic complex, which is competent in executing a complete recombination reaction. The components of a SC are required in stoichiometric quantities and incomplete

assembly results in a transpositional inactivity. Since genome rearrangements are dangerous to the host cell, the demand for a stable protein-DNA complex, which includes all proper DNA sites and protein components before activation of the protein's DNA-cleavage activity reduces the danger of incomplete reactions.

Therefore, similar to other transposable elements, *Sleeping Beauty* transposition is probably controlled at the level of SC formation (Izsvák *et al*., 2002), and the necessary components are the two IRs containing the four transposase-binding sites and four transposase molecules. The DNA-binding domains of the transposase molecules form tetramers together with the transposase-binding sites (Izsvák *et al*., 2002), by which the ends of the element are paired and held together by the transposase subunits. Donor cleavage and joining to target DNA probably occur within the context of this SC.

Because recombination takes place within a stable SC, the process of SC assembly can be a target of physiological control. For example, the high mobility group protein HMGB1 was found to be required for efficient *Sleeping Beauty* transposition *in vivo* (Zayed *et al*., 2003). HMGB1 enhances the binding of the transposase to all four DRs, but its effect is significantly more pronounced at the inner DR sites. Therefore, it appears that at a very early phase in transposition, the transposase first binds to the inner DRs, and then to the outer DRs, suggesting that HMGB1 ensures ordered assembly of a catalytically active synaptic complex.

1.7.1.4 Applications of the *Sleeping Beauty* transposon system in genomics

After completion of the initial sequencing, the human genome appears to contain about 30.000-40.000 protein-coding genes (Lander et al, 2001). However, the exploitation of the obtained nucleic acid sequence for biomedical research is hampered by the lack of knowledge of gene function. Because of the remarkable conservation of the genetic information between organisms, it is a powerful approach to study gene function in vertebrate model organisms such as the zebrafish (*Danio rerio*). Such studies require large-scale gene

identification systems which allow the introduction of random mutations and cloning of affected genes for further studies. Large-scale chemical mutagenesis screens were successfully employed to isolate large numbers of mutations that affect vertebrate development (Haffter *et al*., 1996; Driever *et al*., 1996) but identification of effected genes by positional cloning is difficult. An alternative to chemical mutagenesis is insertional mutagenesis in which exogenous DNA (such as transposon DNA) is used as a mutagen, which serves as a tag for subsequent isolation of mutated genes (transposon tagging). Large-scale transposon mutagenesis screens have been applied to a number of organisms in which controlled transposition has been possible (Hamer *et al*., 2001). However, the requirement of species-specific host factors for the mobilization of the element hampers the application of transposable elements in vertebrates (Gibbs *et al*., 1994; Rio *et al*., 1988). This problem can be circumvented using *Sleeping Beauty* because it transposes with high efficiency in a variety of vertebrate cells (Izsvák *et al*., 2000), and thus has great potential as a molecular tool for functional genomics.

1.7.1.5 The *Sleeping Beauty* transposon as a vector system for gene therapy

Gene therapy aims at functionally replacing a mutated gene by the delivery of a functional transgene that has to be stably integrated into chromosomal DNA to ensure prolonged expression of the transgene. Gene therapy approaches traditionally employ both viral and non-viral systems for the delivery of the transgene into the cell (Kay *et al*., 2001; Nishikawa *et al*., 2001). However, major obstacles in viral gene therapy are size restrictions of the transgene, possible immune responses from the host immune system triggered by viral components, safety issues for the production of the virus, and difficult genetic manipulation. In contrast, non-viral systems are not immunogenic but frequently lack efficient genomic integration, which precludes long-term expression. The *Sleeping Beauty* transposon may provide a major step forward in overcoming these difficulties. The *Sleeping Beauty* vector

system consists of a transgene flanked by the IR/DR sequences which can be mobilized and integrate into genomic DNA when the transposase is supplied either in form of plasmid DNA, mRNA or protein. Hence the *Sleeping Beauty* transposon system helps to surmount a number of obstacles encountered with viral vectors.

1.8 Goal of the project

It is thought that during vertebrate evolution mechanisms have developed that allowed the host to tolerate transposable elements. With the extinction of active DNA transposable elements in vertebrate genomes, such mechanisms might also have been doomed to die out and therefore might not be accessible for investigation. However, it is also possible that some mechanisms have survived and are now involved in other cellular processes than transposition. Hence, investigations of transposon-host interactions gives insight into the evolution of vertebrate genomes, and the impact that transposable elements had on directing this evolution.

Sleeping Beauty is an active DNA transposable element of vertebrate origin and hence offers the possibility to investigate transposon-host interactions in vertebrate cells.

Therefore, I used the *Sleeping Beauty* transposable element as a model system to uncover cellular pathways that are involved in transposition. Furthermore, discovering cellular responses to transposition may help to improve the *Sleeping Beauty* transposon system as a tool for functional genomics and gene therapy.

2. Materials and Methods

2.1 Recombinant DNA technology

All methods used to generate and isolate recombinant plasmid DNA were performed as described in Sambrook *et al*., 1989.

2.2 Plasmids created in this work.

Designation	Insert preparation	Vector preparation	Sequenced
pGBKT7 -SB	Template: CMV-SB Primer: FTC-Stop, SB-Start Restriction: *Eco*RI, *Nde*I	pGBKT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGBKT7 -SB/Gln	Template: CMV-SB Primer: SB-Start/*Nde*I/Gln, FTC-Stop Restriction: *Eco*RI, *Nde*I	pGBKT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGBKT7 -N123/Gln	Template: CMV-SB Primer: SB-Start/*Nde*I/Gln, N123-Stop Restriction: *Eco*RI, *Nde*I	pGBKT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGBKT7 -DDE/Gln	Template: CMV-SB Primer: DDE-Start/*Nde*I/Gln, DDE-Stop/*Eco*RI Restriction: *Eco*RI, *Nde*I	pGBKT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGBKT7 -Cterm/Gln	Template: CMV-SB Primer: FTC-Stop, Cterm-Start/*Nde*I/Gln Restriction: *Eco*RI, *Nde*I	pGBKT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGADT7 -Miz1	Template: pUHD10.1-Miz Primer: Miz1-1, Miz1-2 Restriction: *Eco*RI, *Nde*I	pGADT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGADT7 -Miz1 (269-803)	Template: pUHD10.1-Miz Primer: Miz1-Glu(269)/*Nde*I, Miz1-2 Restriction: *Eco*RI, *Nde*I	pGADT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
pGADT7 -Miz1 (1-268)	Template: pUHD10.1-Miz Primer: Miz1-1, Miz1-Asn(268)/*Eco*RI Restriction: *Eco*RI, *Nde*I	pGADT7 (Clontech) Restriction: *Eco*RI, *Nde*I	Yes
CMV -Miz-1 wt	Template: pUHD10.1-Miz Primer: Miz1-Start/*Hind*III, Miz1-Stop/*Xho*I Restriction: *Hind*III, *Xho*I	pcDNA3.1/Zeo(+) (Invitrogen) Restriction: *Hind*III/*Xho*I	No, but functionally tested

Designation	Insert preparation	Vector preparation	Sequenced
CMV -Miz-1 (269-803)	Template: pUHD10.1-Miz Primer: Miz1-Start/*Hind*III, Miz1-Stop/*Xho*I Restriction: *Hind*III, *Xho*I	pcDNA3.1/Zeo(+) (Invitrogen) Restriction: *Hind*III/*Xho*I	No, but functionally tested
CMV-HMG2L1	Described below		Yes

2.3 Cloning of the human HMG2L1 cDNA

A partial HMG2L1 cDNA (encompassing amino acids 110 to 600) was generated from overlapping cDNA clones; one was recovered from the two-hybrid screen and two commercially available EST clones (IMAGp998J2210102Q2 and IMAGp958N021240Q2), which were purchased from the "Deutschen Ressourcenzentrum für Genomforschung GmbH". The overlapping cDNAs were fused together using *Avr*II and *Bst*XI restriction sites and the resulting partial HMG2L1 cDNA was amplified by PCR (primers: HMG2L1-Nterm/Start, HMG2L1-Cterm/Stop) and cloned into *Nde*I/*Eco*RI sites of pUC19. The sequence was confirmed by sequencing.

The N-terminal region of HMG2L1 (encompassing amino acids 1 to 110) was assembled by a single-step oligo-assembly method described by Stemmer *et al*., 1995. Briefly, 16 oligodeoxyribonucleotides (N1-N16) were synthesized (BioTeZ, Berlin-Buch GmbH) which collectively encode both strands of the N-terminal region of the *HMG2L1* gene. The oligos were assembled in a PCR reaction using *pwo* DNA polymerase (Roche). The generated PCR product was gel-purified and cloned into *Hind*III/*Ava*II sites of pcDNA3.1/Zeo(+)-HMG2L1(MDLL-IMPGL) creating the full-length HMG2L1 cDNA pcDNA3.1/Zeo(+)-HMG2L1(MAYDDS-IMPGL). Sequence was confirmed by sequencing.

2.4 Primer sequences

Primer designation	Sequence 5′→ 3′
Cterm-Start/*Nde*I/Gln	GAT ATA CAT ATG GGT GGT GGT GGT GGT GGT GGT GGT GGT GGT ACA AAC CTG ACT CAG TTA CAC CAG C
DDE-Start/*Nde*I/Gln	GAT ATA CAT ATG GGT GGT GGT GGT GGT GGT GGT GGT GGT GGT CAA AAC CGA CAT AAG AAA GCC AGA CTA CGG
DDE-Stop/*Eco*RI	TCT AGA ATT CTA AGG CCT CCT TGC TCG CAC ACG C
FTC-Stop	TCT AGA ATT CTA GTA TTT GGT AGC ATT G
HMG2L1(EEKD)	GAA GAG AAG GAC AAA GAG AGA GAG AG
HMG2L1-Cterm/*Bst*X1	GCA GAA CCA CAC AGA TGG GCA TAG TG
HMG2L1-Cterm/Stop	TCT AGA ATT CTC ACA GTC CCG GCA TGA TG
HMG2L1-mut/*Bst*XI	GCA GAA CCA CAC AGA TGG GCA TAG TGA AAA GAA GAA GAA GAA GGA AGA GAA GGA CAA AGA GAG AGA GAG AG
HMG2L1-Nterm/*Avr*II	CAT ATA CCT AGG AGT AAA ACA CCC TGC
HMG2L1-Nterm/Start	TAT ACA TAT GGA CCT GTT GAA AGC TAT CAC
HMG2L1-Start(MAYDDS)/*Hind*III	ATA GCT AAG CTT GCC ACC ATG GCT TAT GAT GAC TCC GTG AAG
HMG2L1-Start(MDLL)/*Hind*III	ATA GCT AAG CTT GCC ACC ATG GAC CTGNTTG AAA GCT ATC ACT TC
HMG2L1-Seq3	GGT AAT TGC TTC CAC ACC TCA
HMG2L1-Stop/*Xho*I	CTA GTA CTC GAG TCA CAG TCC CGG CAT GAT GTA AGC
Miz1-1	TAT ACA TAT GGA CTT TCC CCA GCA CAG CCA
Miz1-2	TCT AGA ATT CTC ACT CGG CAG GCG GGG GAC ATT C
Miz1-Asn(268)/*Eco*RI	TCT AGA ATT CTC AAT TCT CGT TCT CCT CGG GG
Miz1-Glu(269)/*Nde*I	TAT ACA TAT GGA GGA GTC AGC GGG CAC AGA C
Miz1-Glu(269)-start/*Hind*III	ATA GCT AAG CTT GCC ACC ATG GAG GAG TCA GCG GGC ACA GAC
Miz1-Start/*Hind*III	ATA GCT AAG CTT GCC ACC ATG GAC TTT CCC CAG CAC AGC CAG
Miz1-Stop/*Xho*I	GCT ATA CTC GAG TCA CTC GGC AGG CGG GGG ACA TTC
N1/*Nde*I	TAT ACA TAT GGC TTA TGA TGA CTC CGT GAA GAA AGA AGA TTG TTT TG
N10	TCT TGT GTG TGT CCG TCC CCA AGA AGT AAA GTT CAC TAT
N11	GGG GAC GGA CAC ACA CAA GAA GAA GAG GAA GCA CTC CTC T
N12	TCT CCA TAG TAG TAA TCA TCA GAG GAG TGC TTC CTC TTC T
N123-Stop	TCT AGA ATT CTA GAG CAG TGG CTT CTT CCT TGC
N13	GAT GAT TAC TAC TAT GGA GAT ATT TCG TCT TTG GAA TCG T
N14	CTT TTT CTT TTT CTT CTG TGA CGA TTC CAA AGA CGA AAT A
N15	CAC AGA AGA AAA AGA AAA AGT CCA GCC CAC AGT CTA CTG A
N16/*Hind*III	TCT CTA AGC TTT CAA CAG GTC CAT AGC TGT ATC AGT AGA CTG TGG GCT GGA
N2	CTC AAA GGTATG ATC ACC ATC AAA ACA ATC TTC TTT CTT C
N3	ATG GTG ATC ATA CCT TTG AGG ACA TAG GAC TTG CAG CTG G
N4	TTT TCT CTC GTT GGC TTC GGC CAG CTG CAA GTC CTA TGT C
N5	CCG AAG CCA ACG AGA GAA AAA ACG TTC TTA CAA AGA TTT T
N6	ATT TCT TCC TCT TCC CTT AAA AAA TCT TTG TAA GAA CGT T
N7	TTA AGG GAA GAG GAA GAA ATT GCT GCT CAG GTC AGG AAT T
N8	CTT CAA CTT CTT CTT GGA AGA ATT CTT GAC CTG AGC AGC A
N9	CTT CCA AGA AGA AGT TGA AGG ATA GTG AAC TTT ACT TCT T
SB-Start	TAT ACA TAT GGG AAA ATC AAA AGA AAT CAG
SB-Start/*Nde*I/Gln	TAT ACA TAT GGG TGG TGG TGG TGG TGG TGG TGG TGG TGG TAT GGG AAA ATC AAA AGA AAT CAG CC

2.5 Plasmids that are described elsewhere

The pCMV-SB transposase-expressing helper plasmid, the pCMV-SB/AS control plasmid, which contains the SB transposase gene in antisense orientation, and the pT/neo donor constructs were described earlier (Ivics *et al*, 1997). The plasmid pIRES-SB, which was derived from pIRES-hyg (Clontech), was described earlier (Izsvak *et al.,* 2000). Plasmid pT/GFP was generated by cloning an *Ase*I/*Mlu*I fragment of pEGFP-C1 (Clontech) into the *Msc*I/*Stu*I sites of pT (Ivics *et al.,* 1997). The cyclin D1- and $p15^{INK4b}$-luciferase reporter constructs [pD1luc and luc-hp15(-2500/+160)] and pCMV-Myc were described previously (Muller *et al*., 1994; Stall *et al*., 2001; Philipp *et al*.,1994). The plasmid expressing the maltose-binding protein (MBP)-SB transposase fusion was described earlier (Zayed *et al.*, 2003).

2.6 Two-hybrid screen and interaction domain mapping

In total, $2x10^7$ independent transformants of a pretransformed HeLa cDNA library (Matchmaker, Clontech) were screened according to the manufacturer´s instructions, using medium stringency selection for protein-protein interactions. Library plasmids were rescued from positive yeast colonies by transformation of *E. coli* and sequenced.

A series of Miz-1 deletion constructs were cotransformed together with the bait construct into the yeast reporter strain AH109 (Clontech). Similarly, the recovered truncated HMG2L1 cDNA was cotransfected with the bait construct to verify the interaction. Yeast clones containing both plasmids were selected and subsequently plated onto medium selecting for protein-protein interaction and incubated at 30˚C. After 12 days, plates were scored for growth by visual inspection.

2.7 MBP pull-down assay

Maltose binding protein (MBP)-SB transposase fusion protein and MBP protein expression was induced in *E. coli* strain BL21 (0.4 mM IPTG, 30ºC for 5 hours). Soluble MBP protein was isolated from cells using BugBuster reagent (Novagen) according to the manufacturers instructions, and bound directly to amylose resin (New England BioLabs). Insoluble MBP-SB was isolated from purified inclusion bodies using sarcosyl (Burgess, 1996), followed by dialysis and binding to resin. Protein concentrations were estimated by gel electrophoresis and coomassie-staining and were adjusted by dilution with unbound beads. Protein-loaded beads were equilibrated for 4 hours at 4˚C on a rotary shaker in binding buffer (20 mM HEPES, pH 7.5, 0.4 M KCl, 25% (v/v) glycerol, 1 mM EDTA, 2 mM $MgCl_2$ and 5 mM DTT). *In vitro* translation of Miz-1 and HMG2L1 was performed with rabbit reticulocyte lysate (Promega) and [^{35}S]-methionine (Amersham). 10 µl samples of radiolabelled protein were added to the MBP-SB-, and MBP-bound resin, and incubated for 4 hours as above. Beads were collected by centrifugation, and washed 10-times in 1 ml binding buffer with 60 mM KCl and without BSA. The proteins were resolved on a 12.5% SDS-polyacrylamide gel. Bands were visualized by autoradiography.

2.8 Cell culture and transfection

Animal cells were cultured at 37˚C and 5% CO_2 in DMEM (Gibco) supplemented with 10% fetal bovine serum (PAA). One day prior to transfection, cells were seeded onto 6-well plates, and incubated for 16-20 hours until 50-80% confluence was reached. Cells were transfected with purified plasmid DNA (Qiagen) using Fugene6 tranfection reagent (Roche).

2.9 Generation of HuH7[IRES-SB] and HuH7[IRES-K] cell lines

The pIRES-SB and pIRES-hyg constructs were transfected into HuH7 cells using Fugene6 transfection reagent. After selection in the presence of 150 μg/ml hygromycin B (Invitrogen), clones were pooled and checked for expression of the SB transposase by Western hybridization, using a polyclonal rabbit antiserum against the SB transposase (created in this lab).

2.10 *Sleeping Beauty* transposition assay in temporarily G1 arrested cells

CHO-K1 cells were transfected as described above with the transposon substrate pT/neo and the helper plasmid pCMV-SB. In a control experiment the helper plasmid was substituted by pCMV-β. After 6 hours of incubation with the DNA-lipid complexes, cells were harvested by trypsinization and transferred onto 10 cm plates containing DMEM plus 10% fetal bovine serum. Plates were incubated for 24 hours to allow uptake of plasmid DNA into the nucleus. G1 arrest was induced by exchanging the medium for DMEM supplemented with 0.5% fetal bovine serum and incubation for 48 hours. To select for transgene integration, cells were grown in DMEM supplemented with 10% fetal bovine serum and 1000 μg/ml G418 (BRL). After 10 days of selection, colonies were fixed on the plates with 10% formaldehyde in PBS for 10 min, stained with methylene blue in PBS for 1 hour, washed extensively with water, air dried and counted**.** Transposition efficiency was determined as the ratio of colonies formed in the presence of the transposase versus colony numbers in the absence of transposase.

2.11 Reporter gene assays

HeLa cells were transfected as described above with 50 ng luciferase reporter construct, 500 ng pCMV-SB or pCMV-SB/AS (as control) and 50 ng pCMV-β as an internal

control for transfection efficiency. Two days post-transfection, cells were washed with PBS and harvested from 6-ml plates using 400 µl of a buffer containing 25 mM Tris-phosphate (pH 7.8), 2 mM DTT, 2 mM CDTA, 10% glycerol and 1% Triton X-100. Light intensity resulting from luciferase was measured in a Lumat LB 9507 luminometer with a 10-second integration period. Briefly, 100 µl of the luciferase assay reagent containing 20 mM Tricine, pH 7.8, 1.07 mM $(MgCO_3)_4Mg(OH)_2$ x $5H_2O$, 2.67 mM $MgSO_4$, 0.1 mM EDTA, pH 7.8, 33.3 mM DTT, 270 µM Coenzyme A, 470 µM luciferin, 530 µM ATP, was mixed with 30 µl cell extract in a luminometer cuvette and measured. The resulting readings were expressed as relative light units (RLU). For the β-galactosidase assay, 30 µl of protein extract was mixed with 970 µl buffer containing 100 mM HEPES (pH 7.3), 150 mM NaCl, 4.5 mM L-Aspartate [hemi-Mg salt], 1% BSA, 0.05% Tween 20 and 1.6 mM CPRG (chlorphenol red-beta-D-galactopyranosid). The reaction was stopped with 0.5 ml 3 mM $ZnCl_2$ and the reaction time was recorded. The OD_{578} of the samples was measured with a spectrophotometer and background from untransfected cells was subtracted. In parallel, the amount of total protein extracted from the cells was determined by Bradford Assay (BioRad) and A_{-595} readings were corrected for the buffer blank. 1 unit of β-galactosedase is defined as the amount which hydrolyzes 1 µmol of CPRG per min per total protein (Yeast protocol handbook; Clontech) and the number of β-galactosedae units was calculated using the formula:

$$\text{units } \beta\text{-galactosidase} = 1000 \times A_{-578}/(\text{reaction time} \times A_{-595})$$

As a measure for promoter activity, light intensities (RLU) were normalized to units of β-galactosidase .

2.12 Cell-cycle and growth curve analysis

Growth curves were determined in triplicate by seeding out 50 x 10^3 cells per well of a 6-well plate containing DMEM supplemented with 10% fetal bovine serum. Cells were incubated at 37˚C (5% CO_2) and at various time points, cells were harvested by trypsinization

and pelleted by centrifugation (4˚C, 200xg, 6 min). In order to obtain a cell density appropriate for counting, the cell pellet was resuspended in an appropriate volume of medium (resuspension volume) and submitted to counting with a Neubauer hematocytometer. The number of cells in each of the four corner squares was counted and the average cell number per corner square was calculated for each repetition (n). The number of cells per well was determined using the following calculation:

cells/well = 10.000 x Average number of cells per corner square x Resuspension volume (ml)

2.13 FACS analysis

Cell-cycle stages were determined by fluorescence-activated cell sorting (FACS), using a FACSCalibur (Becton Dickinson); data were analysed using CELLQuest^{TM} Version 3.1 (Becton Dickinson). Briefly, cells were seeded out in low density onto 10-cm dishes and incubated for 48 hours. For propidium iodide staining, cells were incubated in 1.5 ml 70% ethanol (-20˚C) on ice for 30 minutes. A total number of 300 x 10^3 cells was washed with 2 ml PBS and resuspended in 0.4 ml staining solution containing 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4, 0.1% Triton X-100, 0.1 mM EDTA, 250 µg/ml RNase A (DNase-free) and 56,3 µg/ml propidium iodide. Samples were incubated for 30 minutes at room temperature in the dark and then submitted to analysis.

2.14 Retinoblastoma protein phosphorylation

HuH7[IRES-SB] and HuH7[IRES-K] cells were seeded in equal density onto 10-cm tissue culture plates and incubated for 48 hours at 37 ˚C (5% CO_2). Total protein was isolated from the cells using RIPA buffer (1x PBS, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, and protease inhibitors). 20 µg of total protein was loaded and separated on a 10% acrylamide gel and transferred to PROTRAN BA85 (Schleicher & Schuell) membrane. For

detection of phosphorylated Rb, anti-phospho-Rb (Ser780) and anti-phospho-Rb (Ser807/811) antibodies were used (Cell Signaling Technology, Inc). For loading control, anti-actin Mab-5 antibody (Neomarkers) was used.

2.15 Affymetrix gene chip hybridizations

10^6 HeLa cells expressing the Sleeping Beauty transposase from the IRES-SB construct (Izsvak *et al*., 2000) as well as the control line (IRES-K) were transfected with 2.5 µg pT/GFP. Transfection was performed using Fugene6 transfection reagent (Roche) or by nucleofection (Amaxa). After 48 hours incubation, approximately 60-70 % of the cells expressed GFP. The cells were harvested, and total cellular RNA was isolated (Qiagen, RNeasy Total RNA isolation kit). Sample processing and labelling was according to Lockhart *et al.,* 1996). Briefly, 20 µg of total RNA was reverse transcribed using a cDNA Synthesis System (Roche). The reverse transcription was primed with a dT24 primer containing a T7 RNA polymerase promoter (Roche). *In vitro* transcription was performed in the presence of biotin-labelled CTP and UTP (Enzo Diagnostics) on double-stranded cDNAs, resulting in labelled cRNA that was used as the probe in the hybridization experiment. Affymetrix gene chips (HGU-95AV2) were hybridized overnight, washed, and stained with streptavidin-phycoerythrin. Data were collected by a laser scanning technique. Background analysis and normalization of transcript intensities was done according to the manufacturer's instructions using the MICROARRAY SUITE 5.0 software (Affymetrix). Changes in transcript levels were calculated using absolute signal intensities. The fold change values derived from the background noise were excluded from the analysis.

2.16 RNA isolation and Northern blotting

Total RNA was isolated from HeLa cells using RNeasy® Mini Kit (QIAGEN) according to the manufacturers instructions. Total RNA was separated on a 1% denaturing

agarose gel and transferred to a membrane (Nytran®N, Schleicher & Schuell) using standard protocols and UV cross-linked. For the preparation of radiolabeled cDNA probes, three different domains of HMG2L1 were amplified by PCR using N1/*Nde*I and N16/*Hind*III primers (N-terminal fragment), HMG2L1-START(MDLL)/*Hind*III and HMG2L1-Seq3 primers (central fragment), and HMG2L1(EEKD) and HMG2L1-STOP/*Xho*I primers (C-terminal fragment). The PCR product was purified from excess nucleotides using QIAquick® Nucleotide Removal Kit (QIAGEN) according to the manufacturer´s instructions. PCR fragments were radiolabeled using standard protocols for random primed labeling and excess label was removed using QIAquick® Nucleotide Removal Kit (QIAGEN) according to the manufacturer´s instructions.

For hybridization, the membranes were equilibrated with hybridization buffer (0.5 M $NaPO_4$, 7% SDS, 1mM EDTA (pH 7.2)) for 30 min. at 68ºC before radiolabeled single-stranded probes were added. Hybridization was performed for 16 hrs at 68ºC and membranes were washed with washing buffer (40 mM $NaPO_4$, 1% SDS, 1mM EDTA (pH 7.2)) in four changes for 20 min. each at 68ºC. The damp membranes were sealed in Saran Wrap® and exposed to X-ray film at -80°C for two weeks.

2.17 Statistical analysis

All experiments were carried out at least three times and results are reported as the mean (average) +/- sem (standard error of the mean). The sem was calculated by dividing the standard deviation (stdv) by the square root of the sample size (n). Standard rules were applied to determine the error propagation during mathematical operations.

3. Sleeping Beauty transposase modulates cell-cycle progression through interaction with Miz-1

3.1 Results

3.1.1 Sleeping Beauty transposase downregulates cyclin D1 expression

The process of transposition probably induces a network of cellular pathways including cell-cycle checkpoints, apoptotic checkpoints and DNA damage responses. To identify such pathways, I investigated differential gene expression in cells undergoing transposition, using gene chip hybridizations (Izsvak and Szabo, unpublished). The experiments were done in human HeLa cells, because SB transposition is very efficient in this cell line (Izsvak *et al*., 2000), and therefore cellular responses to transposition are expected to be measurable. A HeLa-derived transgenic cell line (IRES-SB) stably expressing the SB transposase at high levels from the CMV enhancer/promoter has been described earlier (Izsvak *et al.*, 2000). The transposase is expressed from a bicistronic expression vector, which allows for selection of transgene expression by antibiotic selection. Control cells (IRES-K) carry the integrated empty expression vector only, and therefore do not express the transposase. A large number of individual antibiotic-resistant clones were pooled and expanded. Both cell lines were transfected with a plasmid carrying a GFP-tagged transposon (pT/GFP), which induces transposition in IRES-SB cells. In order to distinguish changes in cellular gene expression that are specific for transposition events from those that arose due to the presence of the transposase, I also included untransfected IRES-SB cells in the analysis.

Affymetrix HGU95 gene chips were used for hybridization, using probes prepared from total cellular RNA obtained in two independent experiments. Fold changes in transcript levels between transposase-expressing versus control, as well as transposing versus non-transposing experimental samples were calculated. A fold change was considered biologically relevant when significantly higher than that of the housekeeping gene glyceraldehyde-phosphate dehydrogenase (GAPDH).

Out of the 12.000 genes analyzed, approximately 400 genes showed a marked change in transcript levels. Among these were genes encoding for chromatin factors and DNA repair factors and will not be further presented in this work.

Fig. 3-1 shows average changes in transcript levels of selected cell-cycle genes in the presence of either the transposase only or the transposase plus transpososon DNA (pT/GFP), as compared to samples without the transposase.

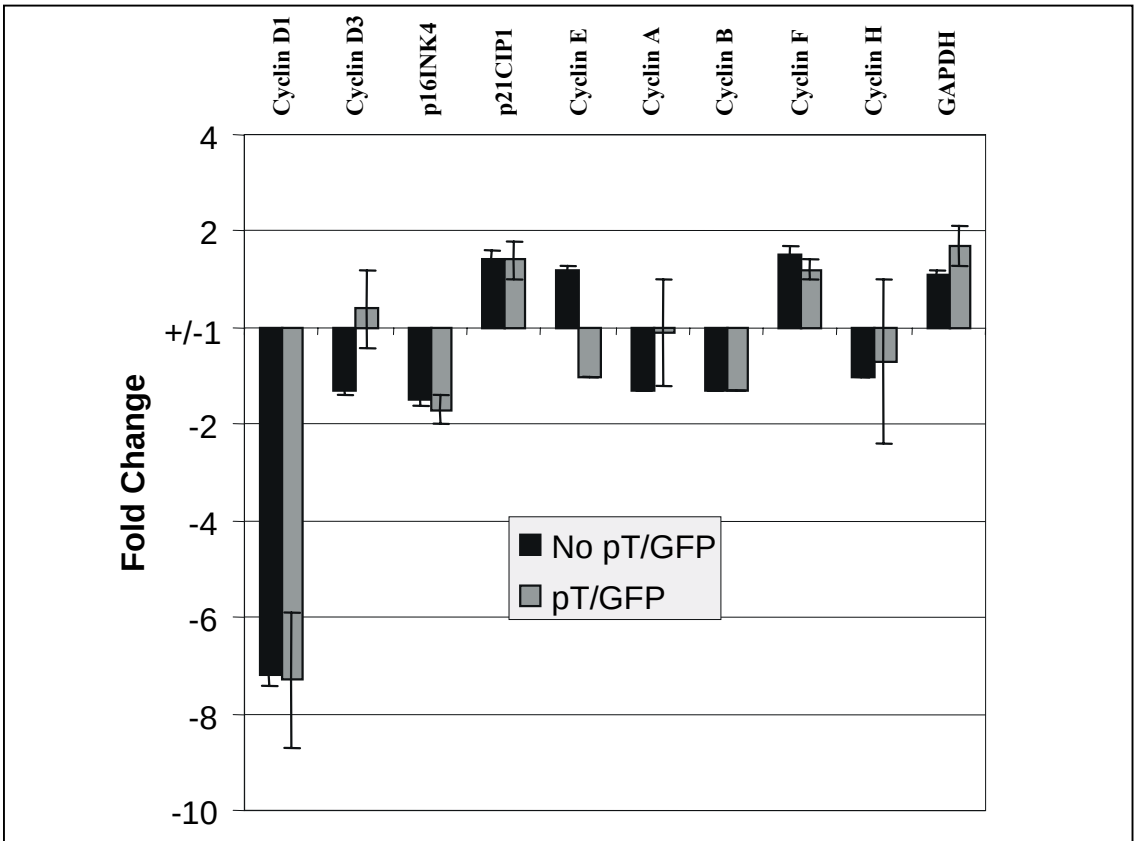


Fig. 3-1: Differential expression of selected cell-cycle regulatory genes in human cells in response to the presence of Sleeping Beauty transposase. Shown are relative changes of transcript levels in a transgenic HeLa cell line stably expressing the SB transposase compared to wild-type cells, in the presence (pT/GFP) and absence (No pT/GFP) of a plasmid containing a GFP-tagged transposon. A fold change of +/-1 indicates no change.

Selected genes produced significant signals above background noise, and included various cyclins and cyclin-dependent kinase (cdk) inhibitors (Fig. 3-1). As shown in Fig. 3-1, transcriptional changes of the majority of these cell-cycle regulatory genes were within the range of GAPDH, and thus are not considered biologically significant. However, in the presence of the transposase, there is an about 6- to 8-fold decrease in cyclin D1 mRNA levels.

The HGU95 chips contains different targets for the cyclin D1 gene, and I consistently observed similar changes for all targets, providing an instant validation of the result. Changes in cyclin D1 transcript levels do not seem to depend on the presence of transposon DNA, and are apparently due to the mere presence of the transposase.

These findings suggest that Sleeping Beauty transposase can, either directly or indirectly, downregulate cyclin D1 promoter activity, and thereby has the potential to modulate cell-cycle progression.

3.1.2 Cells expressing the Sleeping Beauty transposase exhibit retarded growth due to a prolonged G1 phase of the cell-cycle

D-type cyclins are required for progression through the G1 phase of the cell-cycle [for review see Sherr, 1995. Decreased cellular levels of cyclin D1 prevent cells from entering the S phase, resulting in cell-cycle arrest in the G1 phase (Baldin *et al.,* 1993). Thus, because cells expressing the SB transposase have decreased levels of cyclin D1, they are expected to exhibit slower growth and an increase in the G1 cell population.

To test this hypothesis, I compared the growth rates of cells expressing the SB transposase to those of their wildtype counterparts (Fig. 3-2). The HeLa-derived cell line described above is probably not suitable for this experiment, because the presence of papillomavirus E6 and E7 proteins in HeLa cells (Goodwin and DiMaio, 2000) can abrogate G1/S checkpoints. Therefore, stable transgenic derivatives of the human hepatoma cell line HuH7 were made by transfection of the same IRES-SB and IRES-K constructs as above, followed by antibiotic selection and expansion of pooled colonies. Stable expression of the SB transposase in the IRES-SB cells was verified by Western analysis using an anti-SB transposase antibody (Fig. 3-2 A).

To compare the kinetics of growth in the two cell lines, equal numbers of cells of each cell line were seeded onto plates, harvested at different time points and counted using

hematocytometry. As shown in Fig. 3-2 A, cells expressing the SB transposase proliferate slower than control cells.

To verify whether retarded growth in the presence of the transposase is due to changes in cell-cycle progression, cells were analysed for DNA content by flow cytometry (Fig. 3-2B). Cells expressing the transposase showed a 5% increase in the G0/G1-phase cell population compared to control cells. Accumulation of cells in G0/G1 resulted in an equal decrease in S-phase cells, whereas the G2/M-population remained unaffected (Fig. 3-2 B).

These data suggest that Sleeping Beauty transposase can induce a moderate cell-cycle arrest in G1, presumably due to downregulation of the cyclin D1 gene.

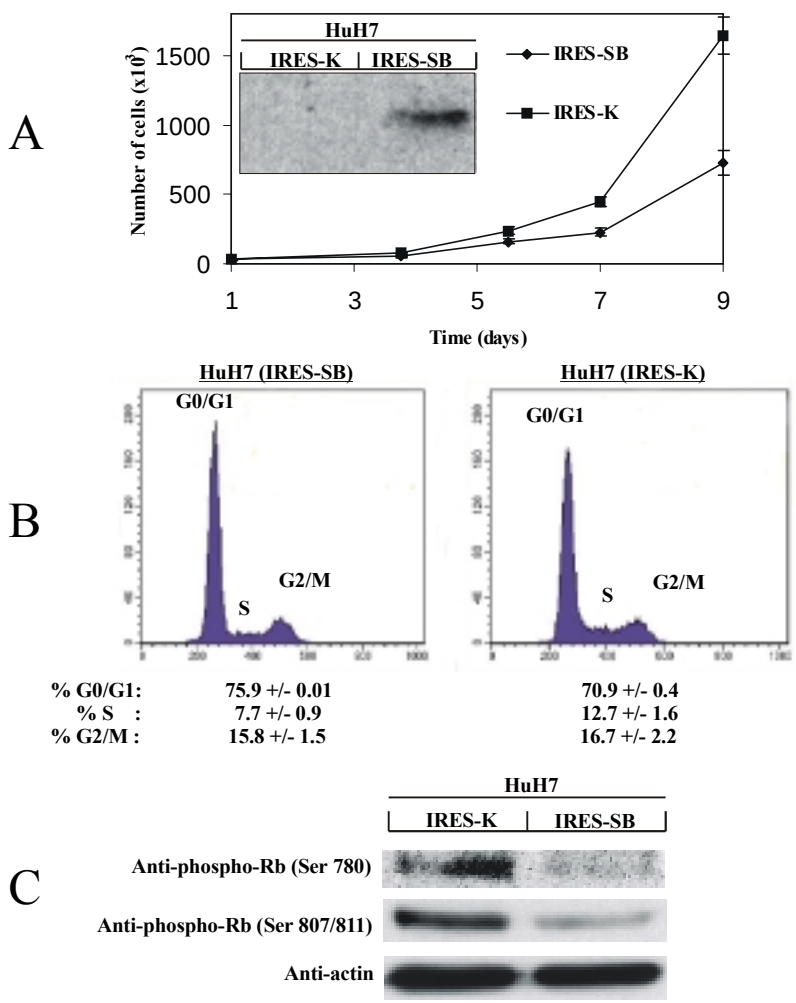


Fig. 3-2: Induction of G1 arrest by Sleeping Beauty transposase in human cells. (A) Growth curves. Growth curves of pooled clones of human hepatoma HuH7 cells expressing SB transposase (IRES-SB) compared to HuH7 cells containing the empty vector (IRES-K). Stable expression of SB transposase in IRES-SB cells was verified by Western analysis using an anti-SB transposase antibody. (**B**) **FACS analysis.** FACScan profiles of exponentially growing, HuH7-derived IRES-SB and IRES-K cells. Indicated are the percentages of cells in different cell-cycle stages. (**C**) **RB phosphorylation.** Levels of phosphorylated retinoblastoma protein in IRES-K and IRES-SB cells were analysed by Western hybridization with antibodies raised against phospho-Rb (Ser780) and phospho-Rb (Ser 807/811). An anti-actin antibody was used to control equal loading .

3.1.3 G1 arrest is associated with decreased levels of phosphorylated retinoblastoma protein

A downstream target of the different cyclin-cdk complexes acting in G1 is the retinoblastoma (Rb) tumor suppressor protein. Rb is expressed throughout the cell-cycle at a fairly steady level, but its phosphorylation state changes in a phase-specific manner (Buchkovich *et al.,* 1989). During G1, phosporylation of Rb by specific cyclin-cdk complexes inhibits the growth-suppressive function of Rb, which allows the cell to enter the S phase (Hinds *et al.,* 1992; Lundberg and Weinberg, 1998).

In order to determine if transposase-induced G1 arrest in HuH7 cells is mediated by cyclin D1, I examined levels of phosphorylated Rb in transposase-expressing and control cells. Western hybridizations with antibodies raised against phospho-Rb (Ser780) and phospho-Rb (Ser807/811) were done. These serine residues are specifically phosphorylated by cyclin D1-cdk4 but not by cyclin E-cdk2 and cyclin A-cdk2 (Zarkowska and Mittnacht, 1997), and therefore serve as useful indicators of cyclin D1-mediated cdk activity. Phosphorylation at the Ser780 site disrupts binding of Rb to the E2F family of transcription factors, thereby enabling the G1/S transition (Kitagawa *et al*., 1996). Fig. 3-2 C shows a significant decrease in Rb phosphorylated at both the Ser780 and the Ser807/811 sites in HuH7 cells expressing the SB transposase.

The decreased levels of hyperphosphorylated (i.e. growth-promoting) Rb in cells expressing the transposase provide strong evidence that transposase-induced G1 arrest is mediated by the Rb/E2F pathway. The results are consistent with these events being triggered by downregulation of the cyclin D1 gene in response to the presence of transposase.

3.1.4 Sleeping Beauty transposase is not sufficient to alter cyclin D1 promoter activity

The above results establish an impact of the SB transposase on cell-cycle progression, and indicate that this impact is brought about, at least in part, by a decrease in cyclin D1 gene activity. I considered that the transposase either has a direct effect on cyclin D1 transcription, or its effect is mediated by other cellular proteins.

To investigate potential direct effects of the SB transposase on the cyclin D1 promoter, a transient reporter gene expression assay was employed. A cyclin D1 promoter-luciferase reporter construct was cotransfected into HeLa cells together with a plasmid expressing the SB transposase (pCMV-SB) or with a control plasmid that contains the transposase gene in an antisense orientation (pCMV-SB/AS). Fig. 3-3 shows that transient overexpression of the SB transposase did not lead to a significant change in the activity of the luciferase reporter gene under the control of the cyclin D1 promoter.

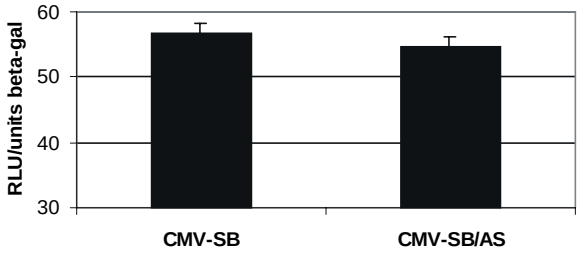


Fig. 3-3: Sleeping Beauty transposase does not alter cyclin D1 promoter activity in a transient transfection experiment. A cyclin D1 promoter-luciferase reporter construct was cotransfected into HeLa cells together with a plasmid expressing the SB transposase (CMV-SB) or with a control plasmid that contains the transposase gene in an antisense orientation (CMV-SB/AS). CMV-β, expressing β-galactosidase, served as an internal control for transfection. Forty-eight hours post-transfection, luciferase activity was measured and normalized with β-galactosidase activity. Numbers represent ratios of luciferase activity (light units per reaction time) and β-galactosidase activity.

I conclude that the SB transposase alone is not sufficient for altering the activity of the cyclin D1 gene, and that a yet unidentified host-encoded factor might be required for mediating transposase-induced repression of the cyclin D1 promoter.

3.1.5 A yeast two-hybrid screen identifies the Myc-interacting zinc finger protein 1 as a specific interactor of the Sleeping Beauty transposase

It is likely that the SB transposase interacts with a number of cellular proteins, and that these interactions affect the transposition reaction as well as endogenous cellular processes. It has been previously established that SB transposase interacts with the DNA-bending protein HMGB1 *in vivo* (Zayed *et al*., 2003).

In order to identify host proteins with the potential to mediate the cyclin D1 promoter response via interaction with the SB transposase, I used the yeast two-hybrid system (Fields *et al*., 1989). A human cDNA library prepared from HeLa cells was chosen for the screen, because our initial observation for reduced cyclin D1 expression in the gene chip experiment was also made in this cell line. This, together with the potential of the *Sleeping Beauty* system as a powerful vector for human gene therapy (Yant *et al*., 2000), drew our attention to the human genome and putative SB interactors therein.

A total of $2x10^7$ independent transformants of a HeLa cell cDNA library fused to the GAL4 activation domain (AD) were screened using the full-length SB transposase gene fused to the GAL4 DNA-binding domain (BD) as bait. Plasmid DNA was isolated from clones that conferred activation of yeast reporter genes, and the cDNAs were identified by sequencing and comparison to sequences deposited in GenBank. One of the isolated cDNAs encoded a truncated version (encompassing amino acids 269 to 788) of the human Myc-interacting zinc finger protein 1 (Miz-1) (Fig. 3-4 A, experiment 1). Another cDNA that conferred activation of the reporter genes encodes for the human high mobility group 2-like 1 protein and will be described in section 4. of this thesis.

Miz-1 is a Myc-associated transcription factor (Peukert *et al*., 1997) and a potent cell-cycle regulator (Staller *et al*., 2001). To verify the SB transposase/Miz-1 interaction, I tested both the full-length Miz-1 protein and a truncated version lacking the N-terminal 268 amino acids for interaction by cotransformation with the SB bait construct into yeast cells. Protein-

protein interaction was detected between the transposase and both the full-length Miz-1 gene product and the truncated version (Fig. 3-4A, experiments 2 and 3). Thus, similar to the interaction of Miz-1 to Myc (Peukert *et al.,* 1997), the N-terminal region of Miz-1 missing in the clone recovered from the library (amino acids 1 to 268) does not contribute to the binding of Miz-1 to the transposase. Indeed, the N-terminal 268 amino acids including the BTB/POZ domain of Miz-1 does not interact with the SB transposase (Fig. 3-4A, experiment 4). To demonstrate further the specificity of the interaction, the GAL4-BD/SB transposase was coexpressed with the GAL4-AD protein (Fig. 3-4A, experiment 5), and the GAL4-AD/Miz-1 protein was coexpressed with the GAL4-BD protein (Fig. 3-4A, experiment 6) in yeast. No reporter gene activation was detected in either case, demonstrating that SB transposase does not interact with the GAL4 activation domain, and that Miz-1 does not interact with the GAL4 DNA-binding domain. Consequently, both Miz-1 and SB transposase are required for interaction.

I carried out a rough mapping of the region of the transposase responsible for interaction with Miz-1 (Fig. 3-4B). Three regions of the transposase polypeptide corresponding to the DNA-binding domain (N123), the catalytic domain containing the DD(35)E signature and the C-terminal end of unknown function were fused to GAL4-BD and tested for interaction with full-length Miz-1 in a two-hybrid cotransformation experiment in yeast (Fig. 3-4B). As with the full-length transposase (Fig. 3-4B, experiment 1), Miz-1 was found to interact with the catalytic DD(35)E domain of the SB transposase (Fig. 3-4B, experiment 3), and no binding was found for the other domains (Fig. 3-4B, experiments 2 and 4).

A physical interaction between the SB transposase and Miz-1 was confirmed *in vitro* using a pull-down assay (Fig. 3-4C). A maltose-binding protein (MBP)-SB transposase fusion protein (MBP-SB) was expressed and purified from *E. coli*, and the protein was immobilized on amylose beads. Immobilized MBP served as a control. Miz-1 was *in vitro* translated in the

presence of [^{35}S]-methionine. Radioactively labeled Miz-1 was incubated with equal amounts of either resin-bound MBP-SB or MBP. The resins were washed extensively, and bound labelled proteins visualized by SDS-PAGE. As shown in Fig. 3-4C, Miz-1 was specifically retained on the MBP-SB beads (lane 2), but not on the MBP beads (lane 3), confirming the interaction between SB transposase and Miz-1.

Taken together, the results establish a specific interaction between Sleeping Beauty transposase and Miz-1, and suggest that SB's effect on the cell-cycle may be realized through this interaction.

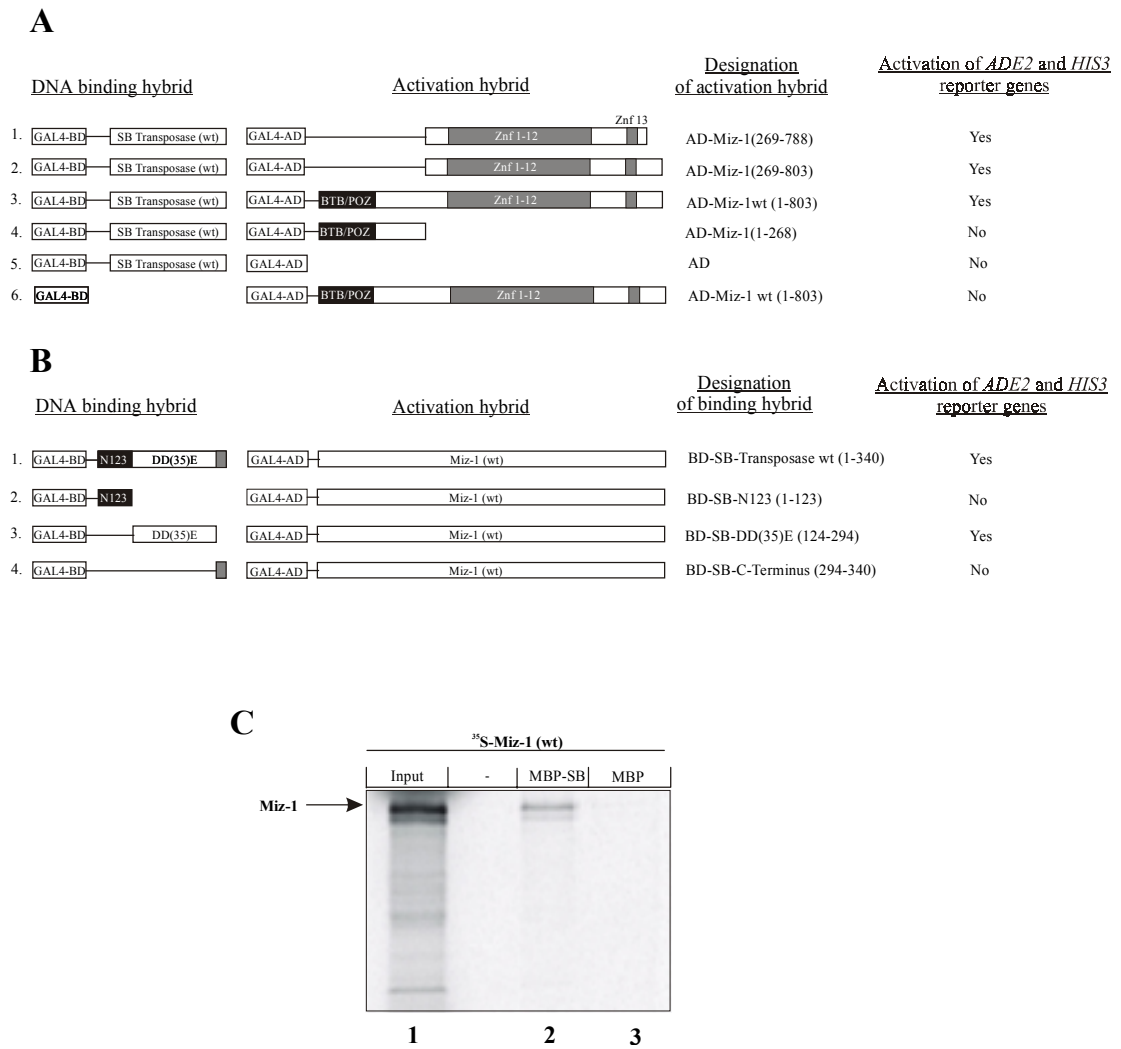


Fig. 3-4: Sleeping Beauty transposase interacts with Miz-1. (**A**) **SB transposase interaction *in vivo*.** SB transposase/Miz-1 interaction is shown in a two-hybrid assay by cotransforming full-length (wt) Miz-1 or its deletion derivatives together with full-length transposase. Protein-protein interaction is assessed by activation of the *ADE2* and *HIS3* reporter genes. (**B**) **Rough mapping of the SB transposase domain responsible for interaction with Miz-1 in a two-hybrid assay.** Full-length (wt) Miz-1 was cotransformed together with full-length SB transposase or its deletion derivatives. (**C**) ***In vitro* interaction between SB transposase and Miz-1.** ^{35}S-labelled, *in vitro*-synthesized Miz-1 was incubated with immobilized maltose binding protein (MBP) alone or with equal amounts of MBP-SB transposase fusion protein. Beads were extensively washed, the recovered material was separated by SDS-PAGE and visualized by autoradiography.

3.1.6 Regulation of the cyclin D1 and p15^{INK4B} promoters by Miz-1 and Sleeping Beauty transposase

Miz-1 was described previously as a transcriptional regulatory protein capable of activating the adenovirus major late (AdML), cyclin D1, p15^{INK4B}, p21^{CIP1}, Mad4 and Nramp1 promoters (Peukert*et al*., 1997; Staller *et al*., 2001; Herold *et al*., 2002; Kime and Wright, 2003; Bowen *et al*., 2003). However, impaired transactivation of the cyclin D1 and p15^{INK4B} promoters by Miz-1 was observed when coexpressed with Myc (Peukert *et al*., 1997; Staller *et al*., 2001). These findings led me to hypothesize that SB transposase might exert an effect similar to Myc, namely, that through its interaction with Miz-1, SB might mediate repression of cell-cycle regulatory genes such as cyclin D1 or p15^{INK4B}.

To substantiate a role of SB transposase in the regulation of cell-cycle genes through association with Miz-1, I investigated the activities of the cyclin D1 and p15^{INK4B} promoters in transient transfection experiments. HeLa cells were cotransfected with cyclin D1- and p15^{INK4B} promoter-driven luciferase reporter constructs together with the pCMV-Miz-1 expression plasmid alone, or either in combination with a construct expressing Myc or the transposase-expressing plasmid pCMV-SB. The pCMV-SB/AS plasmid was used as a control for the transposase, and pCMV-β (expressing β-galactosidase) as an internal control for transfection. Fig. 3-5A shows that Miz-1 activates both the cyclin D1 and p15^{INK4B} promoters, consistent with published results (Peukert *et al*., 1997; Staller *et al*., 2001). An N-terminally truncated Miz-1 protein (containing amino acids 269-803), that retains its ability to interact with SB transposase (Fig. 3-4A; experiment 2), transactivates the cyclin D1 promoter at a level comparable to the full-length Miz-1 protein (Fig. 3-5A).

Next I examined the ability of SB transposase to repress Miz-1-mediated transcriptional activation. As expected, coexpression of Myc antagonizes the activating effect of Miz-1 at the cyclin D1 promoter (Fig. 3-5B, lane 1). Significantly, coexpression of the SB transposase with Miz-1 led to a 3-fold decrease of luciferase expression from the cyclin D1

promoter (lane 2). However, only minimal repression of cyclin D1 promoter activity was observed in the presence of Miz-1(269-803) (lane 3).

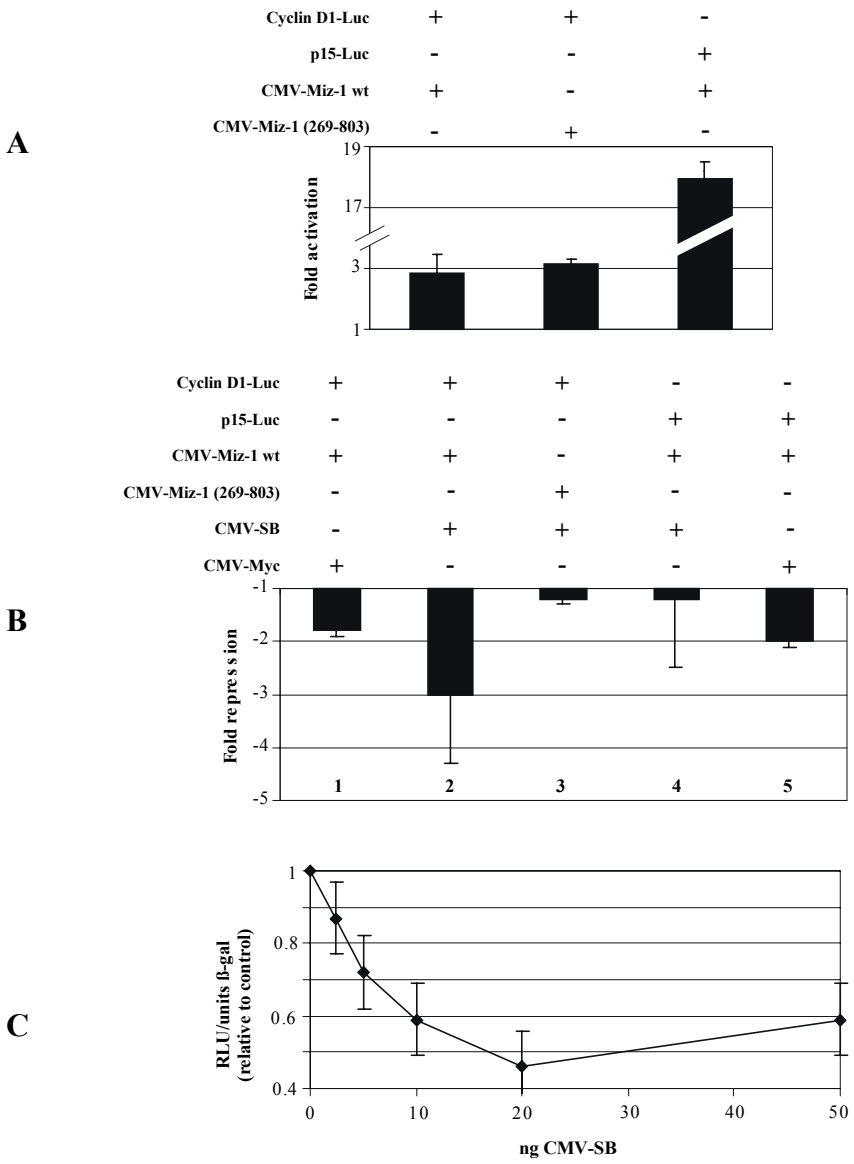


Fig. 3-5: Regulation of the cyclin D1 and p15^{INK4B} promoters by Miz-1, Myc and Sleeping Beauty transposase in transient transfections in HeLa cells. (**A**) Activation of cyclin D1 and p15^{INK4B} promoters by wildtype and truncated Miz-1 (containing amino acids 269 to 803). Data represent fold activation of reporter gene expression. Data were normalized with transfection efficiencies, determined from a cotransfected CMV-β-galactosidase standard. (**B**) Repression of cyclin D1 and p15^{INK4B} promoters by Myc and SB transposase in the presence of wildtype and truncated Miz-1 (containing amino acids 269 to 803). Coexpression of SB transposase with Miz-1 leads to a 3-fold reduction of cyclin D1 promoter activity whereas p15^{INK4b} promoter activity remains unaffected. Minimal repression of cyclin D1 promoter activity can be observed in the presence of Miz-1(269-803). Coexpression of Myc antagonizes the activating effect of Miz-1 at the cyclin D1 and p15^{INK4B} promoters. (**C**) The response of the cyclin D1 promoter to increasing amounts of transfected CMV-SB plasmid (in the presence of a constant amount of CMV-Miz-1) is dose-dependent. Changes of +/- 1indicate no change.

This suggests that cyclin D1 repression by the SB transposase requires the intact Miz-1 protein. Furthermore, the response of the cyclin D1 promoter to increasing amounts of transfected CMV-SB plasmid (in the presence of a constant amount of CMV-Miz-1) was found to be dose-dependent (Fig. 3-5C), providing a further confirmation that the effect on

promoter activity requires the transposase. In contrast to the cyclin D1 promoter, the activity of the $p15^{INK4B}$ promoter remained unaffected in the presence of the transposase (Fig. 3-5B, lane 4), although the reported repressing ability of Myc on the $p15^{INK4B}$ promoter (Staller *et al*., 2001) was clearly reproduced (lane 5).

These data show that downregulation of the cyclin D1 promoter requires both the transposase and Miz-1, and that, unlike Myc, the SB transposase does not affect expression of the $p15^{INK4B}$ gene.

3.1.7 A temporary G1 arrest enhances *Sleeping Beauty* transposition

The above results show that, mediated by the cell-cycle regulatory factor Miz-1, SB transposase downregulates the cyclin D1 promoter, which results in slower cell growth. What is the biological relevance of this process? A reasonable assumption is that a temporary block in the G1 phase of the cell-cycle is beneficial for the transposable element.

To investigate whether a transposase-induced G1 arrest has an impact on the efficiency of transposition, a quantitative transposition assay was performed in serum-starved CHO-K1 cells. Serum-starvation was chosen for the induction of the G1 arrest, because this treatment has a reversible effect: by refeeding the cells with serum, the cells resume the cell-cycle. CHO-K1 cells were used for this experiment, because they were shown to readily respond to serum-starvation (Chou and Chou, 1999), and because SB transposition is efficient in this cell line (Izsvak *et al*., 2000). The *in vivo* transposition assay is based on cotransfection of a donor plasmid carrying an antibiotic resistance gene (*neo*)-marked transposon and a transposase-expressing helper plasmid (pCMV-SB) into cultured cells. In control experiments pCMV-β substitutes for the transposase helper plasmid. Cells are placed under antibiotic selection, and the numbers of resistant colonies are counted. The ratio between numbers obtained in the presence versus the absence of transposase provides a measure of the efficiency of transposition. CHO-K1 cells were subjected to serum withdrawal one day post-

transfection, which resulted in an enrichment of a population of cells in the G1 phase, as determined by FACS analysis (Fig. 3-6A). Cells were maintained under G1 arrest for two days; released from arrest by readdition of serum, and kept under G-418 selection until colonies appeared on the plates. As shown in Fig. 3-6B, there was an approximately 30% increase in transposition efficiency in serum-starved cells.

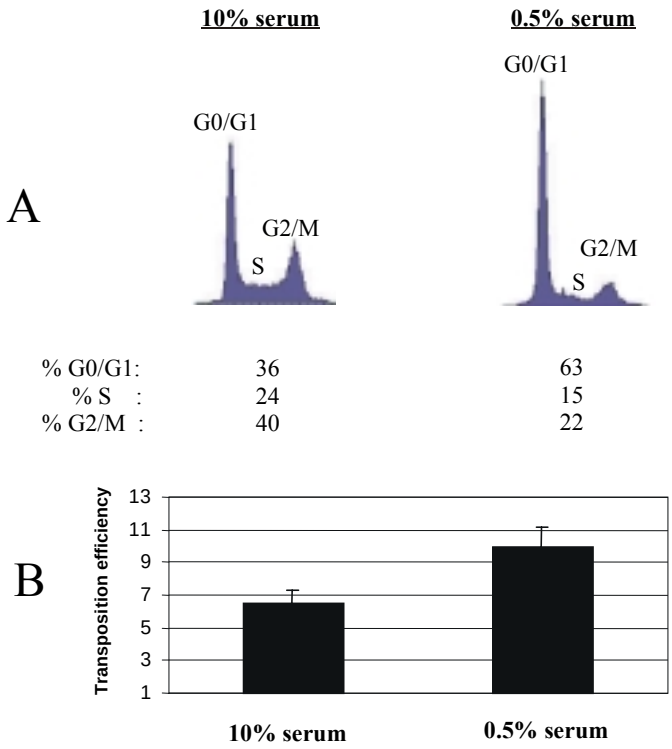


Fig. 3-6: A temporary G1 arrest by serum-starvation enhances *Sleeping Beauty* transposition in CHO-K1 cells. (**A**) CHO-K1 cells readily respond to serum withdrawal with a marked G1 arrest. (**B**) *In vivo* transposition assay in G1 arrested (serum-starved) and exponentially growing CHO-K1 cells. Transposition efficiency represents fold increase in numbers of antibiotic-resistant cell clones in the presence versus in the absence of the transposase.

I conclude that an artificially induced block in the G1 phase of the cell-cycle enhances *Sleeping Beauty* transposition, suggesting that SB's natural ability to slow down the cell-cycle is beneficial for the transposition process.

3.2 Discussion

The results I have obtained in this study strongly suggest that the Sleeping Beauty transposase can regulate cell-cycle progression of proliferating cells.

The cell-cycle is an ordered set of processes through which the cells devide and produce new daughter cells. The cell-cycle can be divided into four phases (Fig. 3-7): The gap phase before DNA replication (G1), the phase in which DNA synthesis takes place (S), a second gap phase after DNA replication (G2) and the mitotic phase in which the cells devide (M). Regulatory mechanisms (checkpoints) assure the ordered progression through the cell-cycle and monitor the internal amount of DNA, and possible DNA damage (Murray, 1994). The core of the cell-cycle machinery are the cyclin-dependent kinases (cdks). The cdk protein kinase holoenzymes are composed of a catalytic subunit (cdk) that, through association with cyclins allow initiation of subsequent cell-cycle events. Throughout the cell-cycle cdk levels remain constant and initiation of the next cell-cycle stage is achieved by association with cyclins. Cyclins are defined as proteins whose expression levels oscillate in a specific manner throughout the cell-cycle and by that bind and activate cdks. Each cdk interacts with a specific subset of cyclins. For example, D-type cyclins (D1, D2 and D3) associate with cdk4/6 during the G1 phase of the cell-cycle, and thus allow progression into S phase (for review see Sherr, 1995). Low cellular levels of cyclin D1 result in an extended G1 phase and a delayed entry into S phase (Baldin *et al.*, 1993), whereas overexpression of D-type cyclins results in a shortened G1 interval (Quelle *et al.*, 1993). cdk4/6 activity can also be

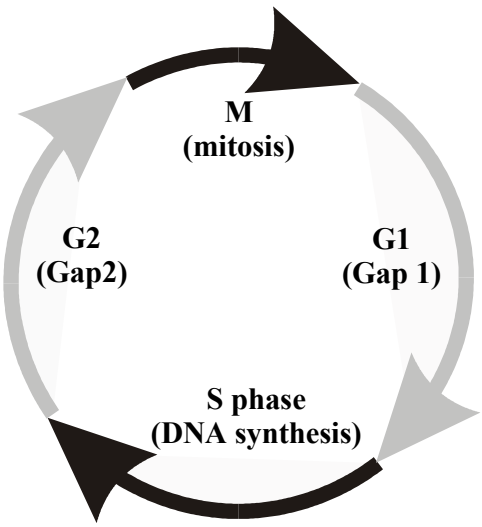


Fig. 3-7: The cell-cycle of proliferating cells. The cell-cycle of proliferating cells consist of alternating DNA synthetic (S) and mitotic (M) phases , separated by gap phases (G1 and G2).

negatively regulated by cyclin-dependent kinase inhibitors (CkIs), which selectively compete with D-type cyclins for binding to cdk4/6 and thus inhibit cdk4/6 function. These are p15, p16, p18 and p19 (Hannon and Beach, 1994; Serrano *et al*., 1993; Hirai *et al*.,1995). Active cyclin-cdk complexes promote positive growth control by inactivating negative regulators such as Rb (Fig. 3-8). Hypophosphorylated Rb during G1 blocks transition into S phase, whereas phosphorylation of Rb by specific cyclin-cdk complexes inhibits the growth-suppressive function of Rb, which allows the cells to enter S phase (Hinds *et al*., 1992; Lundberg and Weinberg, 1998).

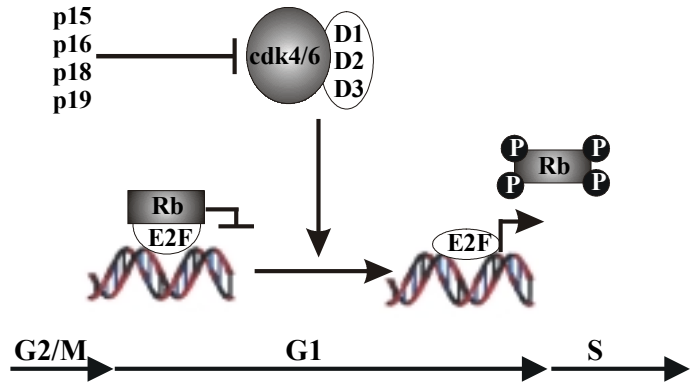


Fig. 3-8: Regulation of G1/S transition by the Rb/E2F pathway. Progression through G1 phase is controlled by the phosporylation-state of the Rb protein. Hypophosphorylated Rb exhibits a growth-suppressive function by binding to and inactivating (perpendicular line) E2F, a transcription factor required for transcription of S phase-specific genes. Phosphorylation of Rb by cyclin D-cdk4/6 complexes leads to a release of E2F and allows the transcription of E2F-controlled S phase-specific genes (rectangular arrow). The activity of cdk4/6 is negatively regulated by the INK4 family if Cdk-inhibitors (perpendicular line) and positively controlled by D-type cyclins.

Here I establish that the Sleeping Beauty transposase induces retarded growth (Fig. 3-2A; page 49), and G1 arrest in proliferating cells (Fig. 3-2B; page 49). Several lines of observation suggest that SB's effect on the cell-cycle is mediated by cyclin D1: 1) gene chip analysis shows markedly reduced cyclin D1 transcript levels in cells expressing the transposase (Fig. 3-1; page 47); 2) SB transposase specifically represses the cyclin D1 promoter in transient reporter assays (Fig. 3-5B; page 56); 3) cells expressing the transposase have reduced levels of retinoblastoma protein phoshorylated at cyclin D1/cdk4 target sites (Fig. 3-2C; page 49); and 4) reduced cellular levels of cyclin D1 and phosphorylated retinoblastoma protein can cause an arrest of the cell-cycle in the G1 phase (Baldin *et al*., 1993; Hinds *et al*., 1992; Lundberg and Weinberg, 1998). My results are therefore consistent with the suggestion that G1 arrest by the transposase is due to cyclin D1 downregulation, although I can not exclude the involvement of other cell-cycle regulatory genes. Furthermore,

my results indicate that presence of the SB transposase alone in cells is sufficient to trigger cyclin D1 downregulation, because levels of cyclin D1 transcript in the presence and absence of transposon DNA were not different (Fig. 3-1; page 47). This result indicates that downregulation of the cyclin D1 gene is not due to transposition-induced DNA damage.

I demonstrate further that the human Myc-interacting zinc finger protein 1 (Miz-1) is an interactor of the Sleeping Beauty transposase (Fig. 3-4; page 54), and is required for cyclin D1 downregulation by the SB transposase (Fig. 3-5; page 56). Miz-1 is a transcription factor containing 13 zinc fingers (Fig. 3-4 A; page 54), of which 12 are clustered in the central portion of the protein (Peukert *et al*., 1997). Miz-1 contains an N-terminal, conserved POZ (pox virus and zinc finger) domain found in many transcription factors with transcriptional repression activity (Deltour *et al*., 1999; Dhordain *et al*., 1997; Huynh *et al*., 2000). The POZ domain has been shown to target the protein to subnuclear foci (Dhordain *et al*., 1997; Peukert *et al*., 1997) and thereby inhibiting DNA binding of the attached zinc finger (Dhordain *et al*., 1995; Bardwell and Treisman, 1994). Miz-1 was first identified by its association with the Myc protein (Peukert *et al*., 1997), which is a transcription factor that can both activate and repress transcription of a large set of target genes (for review see Eisenman 2001; Amati *et al*., 2001; Menssen and Hermeking, 2002). Dimerization of Myc with Max forms a complex that can activate transcription (Kretzner *et al*., 1992), whereas dimerization of Myc with Miz-1 represses transcription (Peukert *et al*., 1997; Staller *et al*., 2001).

Miz-1 activates the AdML and cyclin D1 promoters in transient transfection experiments (Peukert *et al*., 1997). Impaired transactivation of these promoters by Miz-1 was observed when coexpressed with Myc (Peukert *et al*., 1997), suggesting a role of Miz-1 in gene repression by Myc. Subsequently, Staller *et al*. (2001) demonstrated that Miz-1 strongly enhances transcription of the cdk inhibitor gene $p15^{INK4B}$, and that transcriptional activation of $p15^{INK4B}$ is inhibited when Myc forms a complex with Miz-1 at the $p15^{INK4B}$ promoter. Based on these previous studies, I hypothesize that the Sleeping Beauty transposase might exert a

function similar to Myc, namely, transcriptional regulation of cell-cycle regulatory genes (i.e. cyclin D1 and $p15^{INK4B}$). However, I found that interaction between Sleeping Beauty transposase and Miz-1 specifically downregulated cyclin D1 promoter activity, whereas $p15^{INK4B}$ expression remained unchanged (Fig. 3-5B; page 56). This suggests that the transposase/Miz-1 interaction is distinct from the Myc/Miz-1 interaction with regard to the target gene affected. Overexpression of the transposase alone did not alter cyclin D1 promoter activity (Fig. 3-3; page 51), suggesting that endogenous levels of Miz-1 in HeLa cells are not sufficient to mediate SB's repressing effect in transient reporter assays. Consistent with this explanation, Miz-1 was found to be of low abundance in HeLa cells (Peukert *et al*., 1997), and its exogenous overexpression was required to show transcriptional effects on the AdML, cyclin D1 and $p15^{INK4B}$ promoters (Peukert *et al*., 1997; Staller *et al*., 2001). Although the mechanism by which SB transposase antagonizes the effect of Miz-1 at the cyclin D1 promoter was not addressed in this study in detail, my data allow me to draw some conclusions. Namely, an N-terminally truncated version of Miz-1, lacking the POZ domain, interacts with the transposase (Fig. 3-4A; page 54), activates the cyclin D1 promoter (Fig. 3-5A; page 56), but does not appear to mediate repression by the transposase (Fig. 3-5B; page 56). Thus, interaction between Miz-1 and the SB transposase is necessary but not sufficient for repression. A similar phenomenon has been reported for Myc's repressing activity on Miz-1-mediated AdML transcription, which requires the integrity of the POZ domain (Peukert *et al*., 1997). Based on these results, a model has been proposed in which association of Miz-1 with Myc renders Miz-1 insoluble, by inducing a latent insolubilizing function of the POZ domain (Peukert *et al*., 1997). Further work will be required to address if a similar mechanism is at work in transposase-mediated repression.

Cyclin D1 initiates cell cycle progression towards S phase by binding to and activating cdk4 and cdk6 [for review see Sherr, 1995]. Cyclin D1 also appears to be an essential component of the G1 DNA damage checkpoint which transiently delays progression of G1-

phase to allow repair of damaged DNA prior to replication and prevents segregation of damaged chromosomes.

DNA damage induced G1 arrest through the ATM (Ataxia-telangiectasia mutated)-p53/$p21^{CIP1}$ pathway is well described. p53 plays a central role in this pathway by turning on transcription of p21 (EL-Deiry *et al*., 1993). p21 is a potent inhibitor of G1 cyclin-dependent kinases and thus leads to G1 arrest (Harper *et al*., 1993).

However, this pathway requires multiple steps of phosphorylation and transcription, and thus takes several hours to achieve cell cycle arrest (Agami and Bernards, 2000). In contrast to the ATM-p53-/$p21^{CIP1}$ pathway, downregulation of cellular cyclin D1 was found to trigger a faster response to ionizing radiation induced genotoxic stress. (Agami and Bernards, 2000).

Mismatch repair-dependent attenuation of endogenous cyclin D1 followed by treatment with the DNA damaging agent cisplatin was also shown to induce G1 arrest (Lan *et al*., 2002). In murine macrophage cells, cyclin D1 was found to be downregulated after UV-irradiation at both the mRNA level and by proteolysis (Miyakawa and Matsushime, 2001). This immediate G1 arrest is p53-independent, but at a later stage p53 is required to maintain the initial cyclin D1-mediated arrest (Agami and Bernards, 2000), by transcriptional activation of the cdk inhibitor $p21^{CIP1}$ [for review see Levine, 1997 and references therein]. Delay in the G1/S transition and S phase progression by cell-cycle checkpoints is thought to facilitate DNA repair, and to avoid replication and subsequent propagation of potentially hazardous mutations.

My results are consistent with the proposal that the Sleeping Beauty transposase acts as other genotoxic agents (i.e. IR or UV), and triggers an immediate damage response by downregulation of cyclin D1 in an attempt to halt progression of the cell-cycle and thereby allowing time for efficient DNA repair prior to replication. It cannot be excluded that the p53-$p21^{CIP1}$-mediated G1 checkpoint can also be activated, at a later stage when chromosomal

breaks were introduced due to transposition. A p53-p21^{CIP1}-mediated effect could not have been revealed in the gene chip hybridization experiment (Fig. 3-1; page 47), because this pathway is corrupted in HeLa cells (Goodwin and DiMaio, 2000). Transposase-induced G1 arrest is expected to be beneficial for transposition, because increased transpositional activity was observed in G1-arrested cells (Fig. 3-6B, page 58). I argue that the G1-preference of the *Sleeping Beauty* transposable element has evolved because transposon induced DNA lesions display a major threat to the vitality of the host and hence a preventative activation of DNA damage checkpoints by the transposase could provide an evolutionary advantage over other elements. As it is shown in Fig. 1-5 (page 19), transposition of the *Sleeping Beauty* element induces two types of DNA damage: double-strand breaks (DSBs) during excision and single-stranded gaps after integration of the element into target DNA. Whereas single-stranded gaps are relatively harmless and can be easily repaired by host replication or repair enzymes, DSBs are a major threat to the integrity of the genome and, if left unrepaired, they can lead to chromosome loss and death of the host.

DSBs can be repaired in eukaryotic cells by at least two pathways, homologous recombination (HR) and non-homologous end-joining (NHEJ). Both pathways are complementary, but act at different stages of the cell-cycle. Whereas the HR pathway is preferentially used in the late S and G2 phase, the NHEJ pathway plays a dominant role in G1 and early S phase (Takata *et al*., 1998). Yeast cells repair the majority of their DSBs by HR, by essentially copying the missing information from one homologue to the other, preventing loss of genetic information. However, in mammalian cells repair of DSBs seems to occur preferably by NHEJ because genomes of higher eukaryotes have a substantial fraction of repetitive DNA that might result in problems during the homology search for repair by HR. In contrast to repair by HR, repair of DSBs by NHEJ does not require sequence homology for the repair process. During NHEJ, the broken DNA ends are rejoined without the use of extensive sequence homology (for review see Critschlow and Jackson, 1998).

However, when DSBs occur in DNA the broken ends are usually incompatible and thus a simple ligation step would be inefficient. Therefore, the broken ends must be processed by the repair machinery to create a terminal microhomology (1-4 nucleotides) between both ends that is sufficient for efficient ligation. Hence processing of the broken ends leads to the loss of genetic information but prevents the genome from more severe damage that can be caused by DSBs.

The dependence of mammalian cells on the NHEJ pathway to repair DSBs suggests that the DNA damage introduced by excision of the *Sleeping Beauty* transposon will be repaired by this pathway. Indeed, it has been shown that *Sleeping Beauty* transposition is severely reduced in cells which lack essential components of the NHEJ pathway (Izsvák *et al*., in press). Hence, it is reasonable to propose a correlation between the particular pathway used for the repair of transposon-induced DNA damage and the cell-cycle stage where recombination occurs. This is nicely illustrated by gene rearrangements through V(D)J recombination, a transposition-like process in vertebrates that is mediated by the RAG1/2 recombinase (transposase) complex. V(D)J recombination requires the NHEJ pathway (Jackson and Jeggo, 1995). Selective degradation of the RAG2 protein has been proposed to contribute to confining V(D)J recombination to the G1 phase of the cell-cycle (Li *et al*., 1996; Lin and Desiderio, 1993), in which NHEJ is preferentially active (Takata *et al*., 1998). RAG2 accumulates in G1, declines before the cell enters the S phase, and remains low throughout the rest of the cell-cycle (Lin and Desiderio, 1994). Phosphorylation of RAG2 by the cyclin A-cdk2 complex shortens the half-life of the protein at the G1/S boundary, whereas overexpression of $p27^{Kip1}$, a negative regulator of cyclin A-cdk2, results in elevated levels of RAG2, G1 arrest and increased recombination (Lee and Desiderio, 1999). Thus, V(D)J recombination and *Sleeping Beauty* transposition share a common feature of regulation that directs transpositional activity to the G1 phase of the cell-cycle where transposon-induced DNA damage can be most efficiently repaired, and damage to the host is thus minimized.

However, the way in which this regulation is achieved differs significantly. Whereas in V(D)J recombination it is the recombinase (transposase) that is regulated by the cell-cycle machinery, in transposition it is the Sleeping Beauty transposase that regulates the cell-cycle machinery.

The consequences of cyclin D1 downregulation on the cell-cycle machinery and the effect of the Sleeping Beauty transposase are summarized in Fig. 3-9.

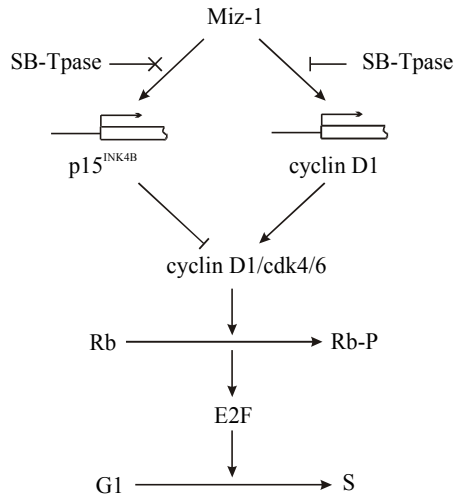


Fig. 3-9: The effect of the Sleeping Beauty transposase on molecular events leading to the G1/S transition in cycling cells. Schematic representation of some of the molecular processes prior to DNA replication in the S phase. Cyclin D1 promotes (arrow), p15 inhibits (perpendicular line) the G1/S transition. Cyclin D1 associates with the cyclin-dependent kinases cdk4 and cdk6, p15 is a cdk inhibitor. Cdks phosphorylate the retinoblastoma protein. Phosphorylation inhibits binding of the E2F transcription factors, thereby enabling the G1/S transition. SB transposase downregulates expression of the cyclin D1 gene in a Miz-1-dependent manner, but does not affect expression of the p15 gene (crossed line). Reduced levels of cyclin D1 result in G1 arrest, which promotes transposition.

Based on my findings, I propose a model in which Sleeping Beauty transposase induces cyclin D1-dependent G1 arrest in proliferating cells through interaction with Miz-1, in order to allow efficient repair of transposon induced DNA double-strand breaks, preferentially by NHEJ. Furthermore, I propose that transposase induced G1 arrest has no effect on single-stranded gap repair because it was demonstrated for retroviruses that a G1 arrest does not increase efficiency of gap repair after retroviral integration (Roe *et al*., 1997).

Can this proposed model, as it is shown in Fig. 3-9 be extended to other members of the Tc1/*mariner* superfamily of transposable elements or is it a unique feature of the *Sleeping Beauty* transposon? An experimental approach to this question is hampered by the fact that there are currently only two Tc1/*mariner* transposable elements active in vertebrate cells: the *Sleeping Beauty* transposon (Ivics *et al*., 1997) and the *Frog Prince* (FP) transposon, which was recently isolated from the frog *Rana pipiens* (Miskey *et al*., 2003).

Preliminary experiments that address the impact of Tc1, Himar1 and the FP transposase on cyclin D1 promoter activity suggest that this is a specialized pathway that is used solely by the SB transposon (data not shown). However, the transpositional activity of Tc1 and Himar1 in human cells is much lower than that of the SB transposon (Fischer *et al*., 2000), which in turn suggests that some cellular requirements for these transposons are not met in human cells. Support for this assumption comes from the fact that the Miz-1 protein is restricted to vertebrate genomes (personal communication, Martin Eilers), and hence interaction of Tc1 and Himar transposases with Miz-1 does not occur in their natural hosts. However, there might be functional homologues in these organisms which could mediate effects similar to the human Miz-1 protein in conjunction with the Sleeping Beauty transposase. This is in sharp contrast to the *Sleeping Beauty* transposon which was isolated from a vertebrate organism (fish) containing the evolutionary highly conserved Miz-1 protein (Fig. 3-10). Therefore, interaction of human Miz-1 protein with the SB transposase is expected to reflect a similar situation in its natural host, fish.

In order to substantiate this assumption, a zebrafish cDNA library was screened with the Sleeping Beauty transposase as a bait protein to attempt recovery of the zebrafish Miz-1 protein as an interactor with the transposase (diploma thesis, Mareike Becker). However, this attempt was inconclusive since not a single interacting protein could be identified. Whether this is due to leakiness of the screen or due to other technical difficulties remains to be elucidated. The same might be true for the FP transposable element, which was isolated from the frog *Rana pipiens* (Miskey *et al*., 2003). However, transient transfection experiments suggest that the FP transposase does not significantly affect cyclin D1 promoter-activity. This might be due to the fact that FP is evolutionary a younger transposon than *Sleeping Beauty* (Miskey *et al*., 2003). Hence, regulatory mechanisms that limit the amplification of the element and thus damage to the host might be poorly developed.

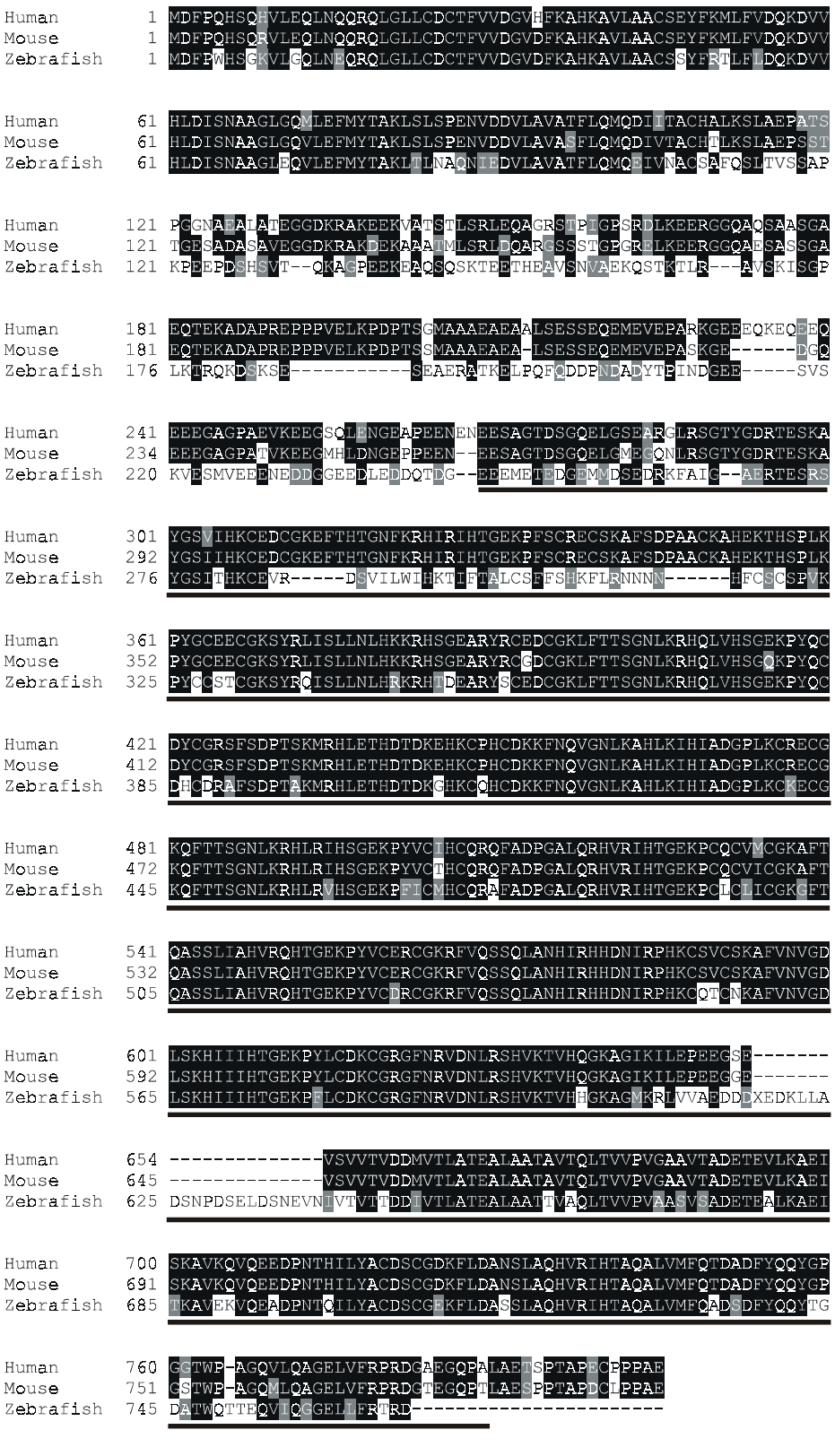


Fig. 3-10: Alignment of the amino acid sequences of Miz-1 proteins from human, mouse and zebrafish. Identical residues are shaded and dashes denote gaps. Underline indicates the region corresponding to the fragment which was recovered from the yeast two-hybrid screen.

Support for this assumption comes from two lines of evidences: 1) Transpositional activity of the FP transposable element in vertebrate cells is significantly higher than the SB transposon

(Miskey *et al*., 2003) and 2) in contrast to Tc1, *Himar*1 and *Sleeping Beauty*, the FP transposable element does not seem to be subjected to overproduction inhibition (Miskey, unpublished observation). This might suggest that the FP transposase has not evolved to a point which allows the employment of a G1 arrest pathway mediated by the frog Miz-1 homologe.

3.3 Future work

The proposed model suggests that *Sleeping Beauty* transposition occurs preferentially in the G1 phase of the cell-cycle where transposon-induced DNA damage can be repaired most efficiently by NHEJ.

Although my own work clearly demonstrates impact of the transposase on the cell-cycle by downregulation of cyclin D1, the link to the NHEJ pathway remains speculative, and requires further investigation.

Therefore, I suggest to perform a transposition assay in G1-arrested and proliferating cells which lack essential components of the NHEJ repair machinery. If the model is true, then a G1 arrest in NHEJ-mutant cells would not result in an increase in transpositional activity, because the broken DNA can not be repaired, and hence a G1 arrest will not provide an advantage.

Another experimental approach to this question is to investigate the resulting transposition footprints in cells which proliferate normally and in those which have a shortened G1 phase and hence impaired repair. If the proposed model is true, then increased cell proliferation (shortened G1) will lead to incomplete DNA repair and thus to altered footprints in the donor DNA (see Fig. 1-5; page 19).

Another aspect which will have to be addressed is whether the proposed model can be extended to other members of the Tc1/*mariner* superfamily of transposable elements. This would include the isolation of functional Miz-1 homologues from *C. elegans* and *Drosophila*

and their subsequent analysis with regard to cell-cycle and transposition. Alternatively, dead copies of human Tc1/*mariner* elements can be brought back to life and tested for interaction with Miz-1.

Another unanswered question is whether the described interaction between the human Miz-1 protein and the fish Sleeping Beauty transposase reflects a similar process in the natural host fish. Functional characterization of interaction between SB transposase and the zebrafish Miz-1 protein will clarify this question.

4. The Sleeping Beauty transposase interacts with the human high mobility group 2-like 1 protein, an HMG-box protein with unknown function

4.1 Results

A second human protein identified in a yeast two-hybrid screen to be a specific interactor of SB transposase is the human high mobility group protein 2-like 1 (HMG2L1). Seroussi *et al*. (1999) have previously predicted the amino acid sequence of this protein, a domain of which shares 39% amino acid sequence with the HMG box of bovine HMG2 (hence the name). However, the cellular function and biochemical activities of HMG2L1 are unknown.

4.1.1 Yeast two-hybrid screen

I isolated a cDNA that encodes a truncated version (encompassing amino acids 138 to 393) of the previously predicted human high mobility group protein 2-like 1 (HMG2L1) (Serroussi *et al*., 1999). To demonstrate the reproducibility of SB transposase/HMG2L1 interaction, I cotransfected the isolated HMG2L1 cDNA [fused to the GAL4 activation domain (GAL4-AD)] together with the SB transposase [fused to the GAL4 DNA-binding domain (GAL4-BD)] into the yeast reporter strain AH109 (Fig. 4-1, experiment 1) that resulted again in reporter gene activation.

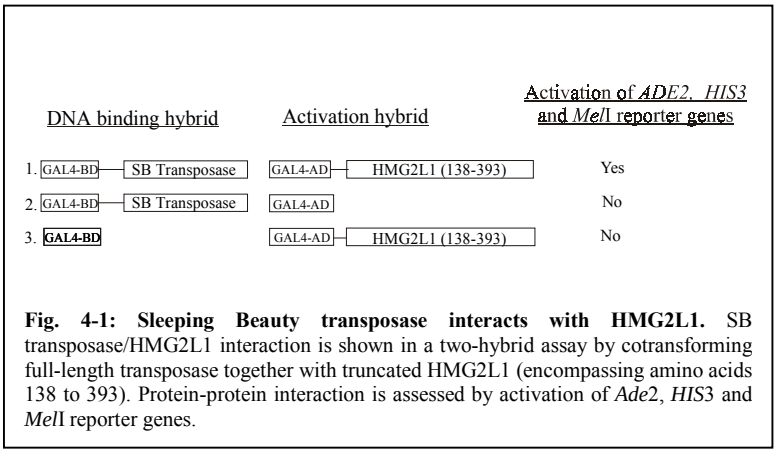


Fig. 4-1: Sleeping Beauty transposase interacts with HMG2L1. SB transposase/HMG2L1 interaction is shown in a two-hybrid assay by cotransforming full-length transposase together with truncated HMG2L1 (encompassing amino acids 138 to 393). Protein-protein interaction is assessed by activation of *Ade*2, *HIS*3 and *Mel*I reporter genes.

To further demonstrate the specificity of the interaction, the GAL4-BD/SB transposase was coexpressed with the GAL4-AD protein (Fig. 4-1, experiment 2) and the GAL4-AD/HMG2L1 (truncated) was coexpressed with the GAL4-BD protein (Fig. 4-1, experiment 3). No reporter gene activation was detected in either case, demonstrating that SB transposase does not interact with the GAL4 activation domain, and that the truncated HMG2L1 protein version does not interact with the GAL4 DNA-binding domain.

4.1.2 Cloning of the full-length HMG2L1 cDNA

The cDNA recovered from the two-hybrid screen encodes the central region of the predicted full-length HMG2L1 protein, and does not include the HMG box (Fig. 4-2).

Therefore, I searched for commercially available EST clones which contain the missing parts. Two EST clones were obtained from the ´´Deutsches Ressourcenzentrum für Genomforschung GmbH´´, one partially encoding the N-terminus of HMG2L1 (IMAGp998J2210102Q2), and the other encoding the C-terminus (IMAGp958N021240Q2), including the HMG box (Fig. 4-2).

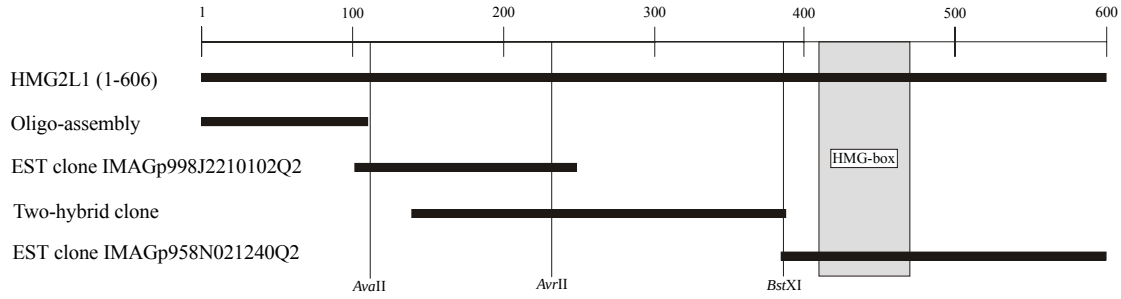


Fig. 4-2: Cloning of the human full-length HMG2L1 gene. The gene was assembled from the cDNA clone recovered from the two-hybrid screen, two commercially available EST clones and oligonucleotides (for details see text).

The cDNA-inserts of the EST clones were sequenced using vector-specific primers and after sequence comparison, *Avr*II and *Bst*XI restriction sites were chosen to assemble partial cDNA inserts together (Fig. 4-2). The resulting cDNA (encompassing amino acids 110 to 606 of the predicted full-length HMG2L1 protein) was cloned into *Eco*RI/*Nde*I sites of pUC19 and confirmed by sequencing. The missing cDNA encoding for the N-terminal 110 amino acids was synthesized using an oligo-assembly approach described by Stemmer *et al*., 1995, fused to an *Ava*II site and the full-length HMG2L1 cDNA (amino acids 1 to 606) was cloned into pcDNA3.1/Zeo (+) (Invitrogen) using *Hind*III/*Xho*I sites.

4.1.3 Detection of HMG2L1 transcripts in HeLa cells

In order to verify that the cloned, full-length HMG2L1 cDNA is represented in HeLa cells by the corresponding mRNA, I conducted a Northern blot analysis using total RNA isolated from HeLa cells.

Three different HMG2L1-specific probes were hybridised to membrane bound mRNA (Fig. 4-3A): one representing the N-terminus of the HMG2L1 protein (probe 1), another representing the central region which contains the interaction domain responsible for protein-protein interaction with the SB transposase (probe 2) and a third spanning the C-terminal region which contains the HMG box (probe 3). The membranes were washed extensively and bound radioactive DNA was visualized by autoradiography (Fig. 3-3B). All three probes used produced a signal representing a mRNA of approximately 2.0 kb in size. Both the unity in size of the detected transcripts by the three different probes, and the fact that its size very well corresponds to the one expected for HMG2L1, strongly suggest that the cDNA I cloned represents endogenous transcripts, and therefore can be used for further experiments.

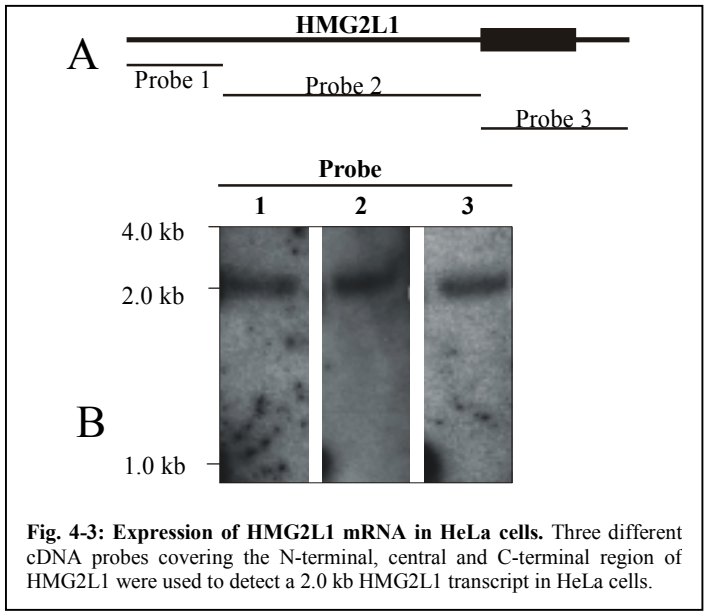


Fig. 4-3: Expression of HMG2L1 mRNA in HeLa cells. Three different cDNA probes covering the N-terminal, central and C-terminal region of HMG2L1 were used to detect a 2.0 kb HMG2L1 transcript in HeLa cells.

4.1.4 *In vitro* interaction between SB transposase and HMG2L1

A physical interaction between the SB transposase and full-length HMG2L1 was confirmed *in vitro* using a pull-down assay (Fig. 4-4).

A maltose-binding protein (MBP)-SB transposase fusion protein (MBP-SB) was expressed and purified from *E. coli*, and the protein was immobilized on amylose beads. Immobilized MBP served as a control. Full-length HMG2L1 was *in vitro* translated in the presence of radioactive methionine. Radioactively labeled HMG2L1 was incubated with equal amounts of resin-bound MBP-SB and MBP. The resins were washed extensively, and bound radioactive proteins were visualized by SDS-PAGE. As shown in Fig. 4-4, HMG2L1 was specifically retained on the MBP-SB beads (lane 2), but not on the MBP beads (lane 3), confirming the interaction between SB transposase and HMG2L1.

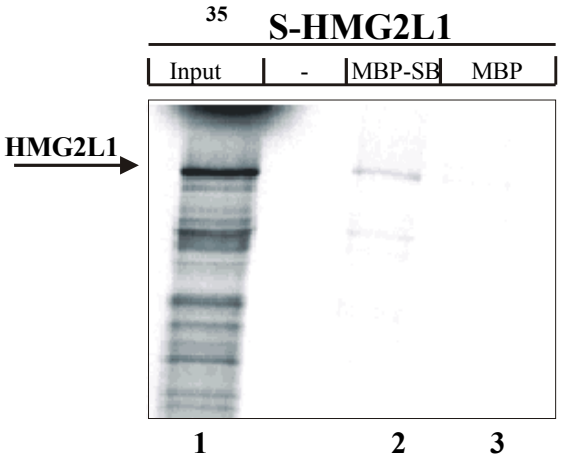


Fig. 4-4: *In vitro* interaction between SB transposase and HMG2L1. ^{35}S-labelled, *in vitro*-synthesized full-length HMG2L1 was incubated with immobilized maltose binding protein (MBP) alone or with equal amounts of MBP-SB transposase fusion protein. Beads were extensively washed, the recovered material was separated by SDS-PAGE and visualized by autoradiograpy.

4.2 Discussion

I identified the high mobility group 2-like 1 (HMG2L1) protein to be a specific interactor of the Sleeping Beauty transposase. I cloned the full-length HMG2L1 cDNA, verified the interaction by a yeast two-hybrid assay and an *in vitro* pull-down assay. Furthermore, I demonstrated that the cloned HMG2L1 cDNA is represented in HeLa cells by the corresponding mRNA using Northern blot analysis.

High mobility group (HMG) proteins were discovered as abundant heterogeneous non-histone components of chromatin, which have high electrophoretic mobility (Goodwin *et al*., 1973). HMG proteins contact the minor groove of DNA through a conserved DNA-binding domain of approximately 85 amino acids called the HMG box and binding is accompanied by a sharp bend in the helix (Giese *et al*., 1992; Ferrari *et al*, 1992; Fisher *et al*., 1992; Zayed *et al*., 2003). Two types of HMG proteins can be distinguished according to their ability for sequence discrimination and the number of HMG boxes (for review see Grosschedl *et al*., 1994). The canonical type contains multiple HMG boxes and binds to DNA in a sequence-independent fashion. This type includes the HMGB1/2 proteins, which lack intrinsic transcriptional activity but were shown to stimulate transcription from RNA polymerase II promoters (Shykind *et al*., 1995). Furthermore, HMGB1/2 proteins are able to modulate the activity of other transcriptional activators such as Oct-1, Oct-2, Oct-6, HOXD9 (Dailey *et al*., 1994), p53 (Jayaraman *et al*., 1998) and MLTF (Watt and Molloy, 1988). The primary role of HMGB1/2 proteins in transcriptional regulation is that they bend DNA, and facilitate the binding of various factors to their cognate DNA sequences (Watt and Molloy, 1988).

The second type of HMG proteins is represented by those which have a single HMG box embedded into their primary sequence, and bind to DNA in a sequence-specific fashion through minor groove contacts that alter DNA conformation (for review see Bewley *et al*., 1998). This type of HMG proteins is represented by the Sox and TCF (T-cell factor)/LEF1

(lymphoid enhancer-binding factor 1) family of proteins. The prototypic member of the Sox family is the male sex determining factor SRY (sex-determining region Y).

SRY is expressed specifically in testis, but not in other tissues suggesting a crucial role in testis development (Sinclair *et al*., 1990; Gubbay *et al*., 1990). Like HMGB1, SRY binds DNA four-way junctions in a sequence-independent fashion, but can also bind and subsequently bend linear duplex DNA containing a heptameric target sequence (Ferrari *et al*., 1992). Similarly to SRY, LEF-1 is a developmentally regulated sequence-specific DNA binding protein, which is specifically expressed in pre-B and T lymphocytes (Travis *et al*., 1991). LEF-1 binds to a defined target sequence (Giese *et al*., 1991; Travis *et al*., 1991) through minor groove contacts and bends DNA (Giese *et al*., 1992). For both SRY and LEF-1, bending of DNA was suggested to play a major role in the mechanism by which transcription of target genes is regulated (Ferrari *et al*., 1992; Giese *et al*., 1992). Bending of the DNA helix might facilitate the interaction between proteins bound at distant sites in the enhancer region and thus activate transcription of the target gene containing the HMG-specific sequence.

A member of the Sox protein family, which is related by homology to the HMG box region of SRY, is the Sox-4 protein (SRY-related HMG-box protein 4). Sox-4 is expressed primarily in T and pre-B lymphocyte cell lines, and binds in a sequence-specific fashion to T-cell specific enhancer regions (van de Wetering *et al*., 1993). In contrast to LEF-1 and SRY, Sox-4 contains a transactivation domain (van de Wetering *et al*., 1993), which can mediate transcriptional activation in the absence of the HMG box.

Interestingly, SRY, LEF-1 and Sox-4 share some similarities with HMG2L1 (see Fig. 4-5). First, these four proteins do not contain any detectable functional domain other than the HMG box, and second, there is very little (or even no) sequence homology to other proteins.

Therefore, I think that the HMG2L1 protein and the previously described Sox, SRY and LEF-1 proteins share some biochemical and functional characteristics, namely: 1) binding to a

specific target sequence located in promoter/enhancer regions of genes, 2) architectural changes in target DNA through bending and 3) transcriptional regulation of these target genes.

In the light of *Sleeping Beauty* transposition, interaction between the transposase and HMG2L1 possibly regulates transcription of unknown target genes in such a way that it affects transposition. Support for this assumption comes from my work on the interaction between Miz-1 and the SB transposase (see section 2). I clearly demonstrated that interaction between these proteins modulates the promoter activity of the Miz-1 target gene cyclin D1, and that this change in promoter activity is beneficial for transposition.

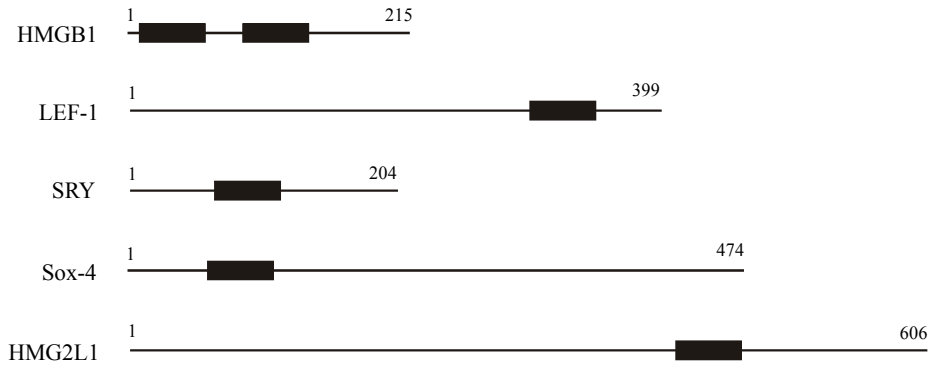


Fig. 4-5: HMG boxes in nonsequence-specific and sequence-specific DNA-binding proteins. HMGB1 represents a class of HMG box (black boxes) proteins with multiple HMG boxes which bind to DNA in sequence-independent fashion. The other class of HMG box proteins which binds to DNA in a sequence-specific fashion includes LEF-1 , SRY, Sox-4 and the recently isolated HMG2L1 protein. Numbers denote amino acid positions.

4.3 Future work

Nothing is known about the cellular function of HMG2L1, but based on structural similarity to the Sox-4, LEF-1 and SRY proteins (Fig. 4-5), I predict a role of HMG2L1 in transcriptional regulation and development. Therefore, I suggest to investigate: 1) target genes and target sequences recognized by HMG2L1, 2) the mechanism of target gene regulation, and 3) the role of HMG2L1 in SB transposition.

A common approach to identify target genes of a particular transcription factor is to overexpress or knock-down that factor, and to investigate genome-wide changes in transcript levels by DNA microarray analysis (Müller *et al*., 2001). This approach will help to identify possible HMG2L1 target genes and binding of the HMG2L1 protein to the promoter of these target genes can be analysed by chromatin immunoprecipitation (ChIP) (Zeller *et al*., 2001; Haggerty *et al*., 2003).

Putative HMG2L1 target sequences recovered by the ChIP method need to be verified by electromobility-shift assays, DNaseI footprinting and circular permutation analysis (Ferrari *et al*., 1992; Giese *et al*., 1992).

The impact of HMG2L1 protein on the efficiency of SB transposition *in vivo* can be determined using a transpositional assay in tissue culture cells. This method allows the quantification of transpositional activity (Ivics *et al*., 1997) when HMG2L1 is overexpressed or endogenous HMG2L1 is knocked-down using RNA interference (RNAi). It is expected that either overexpression or knock-down (or both) of HMG2L2 expression will result in a measurable change in transposition efficiency.

There are at least two different mechanisms how regulation of transposition by HMG2L1 could be achieved. First, HMG2L1 might regulate another host-gene function(s) and association with SB transposase leads to changes in gene expression thereby affecting transposition. I demonstrated such a mechanism for the SB transposase/Miz-1 interaction (Walisko *et al*., unpublished). Second, HMG2L1 could bind to the inverted repeats of the

Sleeping Beauty transposon, and hence assist synaptic complex formation at the transposon ends. Such a mechanism was suggested by Zayed *et al*. (2003) who could show that HMGB1 is a cellular cofactor in SB transposition.

The first mechanism will be addressed by the chromatin immunoprecipitation and microarray analyses described above. The second mechanism will be addressed by electrophoretic mobility-shift experiments, in which we will compare the mobility of the inverted repeats only, inverted repeats plus HMG2L1 and inverted repeats plus HMG2L1 and transposase. This experimental set-up would tell us whether HMG2L1 is possibly involved in synaptic complex formation as it was suggested for HMGB1 (Zayed *et al*., 2003). If it turns out to be so, additional experiments would be necessary to address this question in detail.

5. Concluding thoughts

The present study is to my knowledge the only one so far which systematically aimed at investigating transposon-host interactions of the Tc1-like element *Sleeping Beauty* in human cells. I succesfully identified two human proteins which interact with the SB transposase: the cell-cycle regulatory protein Miz-1, and the HMG-box protein HMG2L1.

Interaction of the transposase with Miz-1 was demonstrated to regulate the host cell-cycle in a way which is beneficial for both transposition of the element and survival of the host. This illustrates the intimate relationship between transposable elements and the host cell that has evolved and rendered a "selfish" element into a tame, domesticated inhabitant of vertebrate genomes. Furthermore, my findings have profound practical consequences with regard to possible applications of the *Sleeping Beauty* system in insertional mutagenesis. Insertional mutagenesis screens rely on the fact that the observed phenotype is due to a transposon insertion and not due to chromosomal rearrangements caused by transposon-induced DNA DSBs. In this regard, my results clearly suggest that mechanisms exist (namely the transposase/Miz-1 interaction) which support accurate DNA repair, and thus reduce (or even eliminate) the chance of chromosomal rearrangements which may lead to high background mutations during the screen. This clearly supports the potential use of *Sleeping Beauty* in vertebrate functional genomics. In the light of *Sleeping Beauty* as a non-viral vector system in gene therapy, my results suggest that transposition possibly can occur in non-dividing cells such as T-cells, brain tissue and nerve cells which would extend the use of *Sleeping Beauty* to treat genetic disorders in both proliferating and non-proliferating cells.

At the current stage of my research, it is difficult to predict what the impact of the transposase/HMG2L1 interaction on both the element itself and the host is. This requires detailed molecular studies of the cellular function of the HMG2L1 protein as briefly proposed in section 3.3.

Reference List

Adzuma K, Mizuuchi K (1989) Interaction of proteins located at a distance along DNA: mechanism of target immunity in the Mu DNA strand-transfer reaction. Cell 57: 41-47

Agami R, Bernards R (2000) Distinct initiation and maintenance mechanisms cooperate to induce G1 cell cycle arrest in response to DNA damage. Cell 102: 55-66

Aidinis V, Bonaldi T, Beltrame M, Santagata S, Bianchi ME, Spanopoulou E (1999) The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol.Cell Biol. 19: 6532-6542

Amati B, Frank SR, Donjerkovic D, Taubert S (2001) Function of the c-Myc oncoprotein in chromatin remodeling and transcription. Biochim.Biophys.Acta 1471: M135-M145

Arciszewska LK, Drake D, Craig NL (1989) Transposon Tn7. cis-Acting sequences in transposition and transposition immunity. J.Mol.Biol. 207: 35-52

Ariyoshi M, Vassylyev DG, Iwasaki H, Nakamura H, Shinagawa H, Morikawa K (1994) Atomic structure of the RuvC resolvase: a holliday junction-specific endonuclease from E. coli. Cell 78: 1063-1072

Auge-Gouillou C, Bigot Y, Pollet N, Hamelin MH, Meunier-Rotival M, Periquet G (1995) Human and other mammalian genomes contain transposons of the mariner family. FEBS Lett. 368: 541-546

Bainton R, Gamas P, Craig NL (1991) Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DNA. Cell 65: 805-816

Bainton RJ, Kubo KM, Feng JN, Craig NL (1993) Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell 72: 931-943

Baldin V, Lukas J, Marcote MJ, Pagano M, Draetta G (1993) Cyclin D1 is a nuclear protein required for cell cycle progression in G1. Genes Dev. 7: 812-821

Baltimore D (1970) RNA-dependent DNA polymerase in virions of RNA tumour viruses. Nature 226: 1209-1211

Bardwell VJ, Treisman R (1994) The POZ domain: a conserved protein-protein interaction motif. Genes Dev. 8: 1664-1677

Beall EL, Mahoney MB, Rio DC (2002) Identification and analysis of a hyperactive mutant form of Drosophila P-element transposase. Genetics 162: 217-227

Bender J, Kleckner N (1986) Genetic evidence that Tn10 transposes by a nonreplicative mechanism. Cell 45: 801-815

Bewley CA, Gronenborn AM, Clore GM (1998) Minor groove-binding architectural proteins: structure, function, and DNA recognition. Annu.Rev.Biophys.Biomol.Struct. 27: 105-131

Boeke JD, Garfinkel DJ, Styles CA, Fink GR (1985) Ty elements transpose through an RNA intermediate. Cell 40: 491-500

Boeke JD, Chapman KB (1991) Retrotransposition mechanisms. Curr.Opin.Cell Biol. 3: 502-507

Bolland S, Kleckner N (1996) The three chemical steps of Tn10/IS10 transposition involve repeated utilization of a single active site. Cell 84: 223-233

Bowen H et al. (2003) Characterization of the murine Nramp1 promoter: requirements for transactivation by Miz-1. J.Biol.Chem. 278: 36017-36026

Brin E, Yi J, Skalka AM, Leis J (2000) Modeling the late steps in HIV-1 retroviral integrase-catalyzed DNA integration. J.Biol.Chem. 275: 39287-39295

Buchkovich K, Duffy LA, Harlow E (1989) The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. Cell 58: 1097-1105

Burgess RR (1996) Purification of overproduced Escherichia coli RNA polymerase sigma factors by solubilizing inclusion bodies and refolding from Sarkosyl. Methods Enzymol. 273: 145-149

Bushman FD, Craigie R (1991) Activities of human immunodeficiency virus (HIV) integration protein in vitro: specific cleavage and integration of HIV DNA. Proc.Natl.Acad.Sci.U.S.A 88: 1339-1343

Bustin M (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Mol.Cell Biol. 19: 5237-5246

Cappello J, Handelsman K, Lodish HF (1985) Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. Cell 43: 105-115

Chalker DL, Sandmeyer SB (1992) Ty3 integrates within the region of RNA polymerase III transcription initiation. Genes Dev. 6: 117-128

Chou LF, Chou WG (1999) Dna-end binding activity of Ku in synchronized cells. Cell Biol.Int. 23: 663-670

Critchlow SE, Jackson SP (1998) DNA end-joining: from yeast to man. Trends Biochem.Sci. 23: 394-398

Dailey L, Yuan H, Basilico C (1994) Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. Mol.Cell Biol. 14: 7758-7769

Daniels GR, Deininger PL (1985) Repeat sequence families derived from mammalian tRNA genes. Nature 317: 819-822

de la Cruz NB, Weinreich MD, Wiegand TW, Krebs MP, Reznikoff WS (1993) Characterization of the Tn5 transposase and inhibitor proteins: a model for the inhibition of transposition. J.Bacteriol. 175: 6932-6938

Deltour S, Guerardel C, Leprince D (1999) Recruitment of SMRT/N-CoR-mSin3A-HDAC-repressing complexes is not a general mechanism for BTB/POZ transcriptional repressors: the case of HIC-1 and gammaFBP-B. Proc.Natl.Acad.Sci.U.S.A 96: 14831-14836

Derbyshire KM, Kramer M, Grindley ND (1990) Role of instability in the cis action of the insertion sequence IS903 transposase. Proc.Natl.Acad.Sci.U.S.A 87: 4048-4052

Dhordain P et al. (1995) The BTB/POZ domain targets the LAZ3/BCL6 oncoprotein to nuclear dots and mediates homomerisation in vivo. Oncogene 11: 2689-2697

Dhordain P et al. (1997) Corepressor SMRT binds the BTB/POZ repressing domain of the LAZ3/BCL6 oncoprotein. Proc.Natl.Acad.Sci.U.S.A 94: 10762-10767

Doak TG, Doerder FP, Jahn CL, Herrick G (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc.Natl.Acad.Sci.U.S.A 91: 942-946

Downs JA, Jackson SP (1999) Involvement of DNA end-binding protein Ku in Ty element retrotransposition. Mol.Cell Biol. 19: 6260-6268

Driever W et al. (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. Development 123: 37-46

Dyda F, Hickman AB, Jenkins TM, Engelman A, Craigie R, Davies DR (1994) Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science 266: 1981-1986

Eide D, Anderson P (1985) Transposition of Tc1 in the nematode Caenorhabditis elegans. Proc.Natl.Acad.Sci.U.S.A 82: 1756-1760

Eisenman RN (2001) Deconstructing myc. Genes Dev. 15:2023-2030

el Deiry WS et al. (1993) WAF1, a potential mediator of p53 tumor suppression. Cell 75: 817-825

Emmons SW, Yesner L, Ruan KS, Katzenberg D (1983) Evidence for a transposon in Caenorhabditis elegans. Cell 32: 55-65

Engelman A, Mizuuchi K, Craigie R (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67: 1211-1221

Engelman A, Craigie R (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J.Virol. 66: 6361-6369

Engels W et al. (1990) Ischemia and reperfusion induced formation of eicosanoids in isolated rat hearts. Am.J.Physiol 258: H1865-H1871

Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J (1990) High-frequency P element loss in Drosophila is homolog dependent. Cell 62: 515-525

Errede B, Company M, Hutchison CA, III (1987) Ty1 sequence with enhancer and mating-type-dependent regulatory activities. Mol.Cell Biol. 7: 258-265

Fadool JM, Hartl DL, Dowling JE (1998) Transposition of the mariner element from Drosophila mauritiana in zebrafish. Proc.Natl.Acad.Sci.U.S.A 95: 5182-5186

Feng JA, Johnson RC, Dickerson RE (1994) Hin recombinase bound to DNA: the origin of specificity in major and minor groove interactions. Science 263: 348-355

Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, Bianchi ME (1992) SRY, like HMG1, recognizes sharp angles in DNA. EMBO J. 11: 4497-4506

Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. Nature 340: 245-246

Fischer SE, Wienholds E, Plasterk RH (2001) Regulated transposition of a fish transposon in the mouse germ line. Proc.Natl.Acad.Sci.U.S.A 98:6759-6764

Fisher RP, Lisowsky T, Parisi MA, Clayton DA (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. J.Biol.Chem. 267: 3358-3367

Franz G, Loukeris TG, Dialektaki G, Thompson CR, Savakis C (1994) Mobile Minos elements from Drosophila hydei encode a two-exon transposase with similarity to the paired DNA-binding domain. Proc.Natl.Acad.Sci.U.S.A 91: 4746-4750

Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG (2000) The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu.Rev.Immunol.18: 495-527

Gibbs PD, Gray A, Thorgaard G (1994) Inheritance of P element and reporter gene sequences in zebrafish. Mol.Mar.Biol.Biotechnol. 3: 317-326

Giese K, Amsterdam A, Grosschedl R (1991) DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. Genes Dev. 5: 2567-2578

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

Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR (1998) Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. Proc.Natl.Acad.Sci.U.S.A 95: 9150-9154

Goodman SD, Nash HA (1989) Functional replacement of a protein-induced bend in a DNA recombination site. Nature 341: 251-254

Goodwin EC, DiMaio D (2000) Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways. Proc.Natl.Acad.Sci.U.S.A 97: 12513-12518

Goodwin GH, Sanders C, Johns EW (1973) A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. Eur.J.Biochem. 38: 14-19

Goryshin IY, Reznikoff WS (1998) Tn5 in vitro transposition. J.Biol.Chem. 273: 7367-7374

Grindley ND, Sherratt DJ (1979) Sequence analysis at IS1 insertion sites: models for transposition. Cold Spring Harb.Symp.Quant.Biol. 43 Pt 2: 1257-1261

Grosschedl R, Giese K, Pagel J (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. Trends Genet. 10: 94-100

Gubbay J et al. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature 346: 245-250

Haffter P et al. (1996) The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. Development 123: 1-36

Haggerty TJ, Zeller KI, Osthus RC, Wonsey DR, Dang CV (2003) A strategy for identifying transcription factor binding sites reveals two classes of genomic c-Myc target sites. Proc.Natl.Acad.Sci.U.S.A 100:5313-5318

Hamer L, DeZwaan TM, Montenegro-Chamorro MV, Frank SA, Hamer JE (2001) Recent advances in large-scale transposon mutagenesis. Curr.Opin.Chem.Biol. 5: 67-73

Hannon GJ, Beach D (1994) p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. Nature 371: 257-261

Hansen LJ, Chalker DL, Sandmeyer SB (1988) Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. Mol.Cell Biol. 8: 5245-5256

Herold S et al. (2002) Negative regulation of the mammalian UV response by Myc through association with Miz-1. Mol.Cell 10: 509-521

Hinds PW, Mittnacht S, Dulic V, Arnold A, Reed SI, Weinberg RA (1992) Regulation of retinoblastoma protein functions by ectopic expression of human cyclins. Cell 70: 993-1006

Hirai H, Roussel MF, Kato JY, Ashmun RA, Sherr CJ (1995) Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. Mol.Cell Biol. 15: 2672-2681

Hollis GF, Hieter PA, McBride OW, Swan D, Leder P (1982) Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. Nature 296: 321-325

Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH, Jr. (1994) A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat.Genet. 7: 143-148

Houck CM, Rinehart FP, Schmid CW (1979) A ubiquitous family of repeated DNA sequences in the human genome. J.Mol.Biol. 132: 289-306

Huynh KD, Fischle W, Verdin E, Bardwell VJ (2000) BCoR, a novel corepressor involved in BCL-6 repression. Genes Dev.14:1810-1823

Ivics Z, Izsvak Z, Minter A, Hackett PB (1996) Identification of functional domains and evolution of Tc1-like transposable elements. Proc.Natl.Acad.Sci.U.S.A 93: 5008-5013

Ivics Z, Hackett PB, Plasterk RH, Izsvak Z (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91: 501-510

Izsvak Z, Ivics Z, Hackett PB (1995) Characterization of a Tc1-like transposable element in zebrafish (Danio rerio). Mol.Gen.Genet. 247: 312-322

Izsvak Z, Ivics Z, Plasterk RH (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J. Mol. Biol. 302: 93-102

Izsvak Z, Khare D, Behlke J, Heinemann U, Plasterk RH, Ivics Z (2002) Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. J.Biol.Chem. 277: 34581-34588

Jackson SP, Jeggo PA (1995) DNA double-strand break repair and V(D)J recombination: involvement of DNA-PK. Trends Biochem.Sci. 20: 412-415

Jacobson JW, Hartl DL (1985) Coupled instability of two X-linked genes in Drosophila mauritiana: germinal and somatic mutability. Genetics 111: 57-65

Jayaraman L, Moorthy NC, Murthy KG, Manley JL, Bustin M, Prives C (1998) High mobility group protein-1 (HMG-1) is a unique activator of p53. Genes Dev. 12: 462-472

Johnson RC, Yin JC, Reznikoff WS (1982) Control of Tn5 transposition in Escherichia coli is mediated by protein from the right repeat. Cell 30: 873-882

Jordan E, Saedler H, Starlinger P (1968) O0 and strong-polar mutations in the gal operon are insertions. Mol.Gen.Genet. 102: 353-363

Kaufman PD, Rio DC (1992) P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell 69: 27-39

Kay MA, Glorioso JC, Naldini L (2001) Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. Nat.Med. 7: 33-40

Kazazian HH, Jr., Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE (1988) Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332: 164-166

Kime L, Wright SC (2003) Mad4 is regulated by a transcriptional repressor complex that contains Miz-1 and c-Myc. Biochem.J. 370: 291-298

Kirchner J, Connolly CM, Sandmeyer SB (1995) Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. Science 267: 1488-1491

Kitagawa M et al. (1996) The consensus motif for phosphorylation by cyclin D1-Cdk4 is different from that for phosphorylation by cyclin A/E-Cdk2. EMBO J. 15: 7060-7069

Krebs MP, Reznikoff WS (1986) Transcriptional and translational initiation sites of IS50. Control of transposase and inhibitor expression. J.Mol.Biol. 192: 781-791

Kretzner L, Blackwood EM, Eisenman RN (1992) Myc and Max proteins possess distinct transcriptional activities. Nature 359: 426-429

Kulkosky J, Jones KS, Katz RA, Mack JP, Skalka AM (1992) Residues critical for retroviral integrative recombination in a region that is highly conserved among

retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol.Cell Biol. 12:2331-2338

Kunze R, Behrens U, Courage-Franzkowiak U, Feldmar S, Kuhn S, Lutticke R (1993) Dominant transposition-deficient mutants of maize Activator (Ac) transposase. Proc.Natl.Acad.Sci.U.S.A 90: 7094-7098

Lampe DJ, Churchill ME, Robertson HM (1996) A purified mariner transposase is sufficient to mediate transposition in vitro. EMBO J. 15: 5470-5479

Lampe DJ, Grant TE, Robertson HM (1998) Factors affecting transposition of the Himar1 mariner transposon in vitro. Genetics 149: 179-187

Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ, Robertson HM (1999) Hyperactive transposase mutants of the Himar1 mariner transposon. Proc.Natl.Acad.Sci.U.S.A 96: 11428-11433

Lan Z et al. (2002) DNA damage invokes mismatch repair-dependent cyclin D1 attenuation and retinoblastoma signaling pathways to inhibit CDK2. J.Biol.Chem.277: 8372-8381

Lander ES et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921

Landree MA, Wibbenmeyer JA, Roth DB (1999) Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. Genes Dev. 13: 3059-3069

Lavoie BD, Chaconas G (1990) Immunoelectron microscopic analysis of the A, B, and HU protein content of bacteriophage Mu transpososomes. J.Biol.Chem. 265: 1623-1627

Lee J, Desiderio S (1999) Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. Immunity. 11: 771-781

Levine AJ (1997) p53, the cellular gatekeeper for growth and division. Cell 88: 323-331

Li Z, Dordai DI, Lee J, Desiderio S (1996) A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. Immunity. 5: 575-589

Lichtenstein C, Brenner S (1982) Unique insertion site of Tn7 in the E. coli chromosome. Nature 297: 601-603

Lin WC, Desiderio S (1993) Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. Science 260: 953-959

Lohe AR, Hartl DL (1996) Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Mol.Biol.Evol. 13: 549-555

Lundberg AS, Weinberg RA (1998) Functional inactivation of the retinoblastoma protein requires sequential modification by at least two distinct cyclin-cdk complexes. Mol.Cell Biol. 18: 753-761

Maekawa T, Yanagihara K, Ohtsubo E (1996) A cell-free system of Tn3 transposition and transposition immunity. Genes Cells 1: 1007-1016

Marlor RL, Parkhurst SM, Corces VG (1986) The Drosophila melanogaster gypsy transposable element encodes putative gene products homologous to retroviral proteins. Mol.Cell Biol. 6: 1129-1134

Maruyama K, Schoor KD, Hartl DL (1991) Identification of nucleotide substitutions necessary for trans-activation of mariner transposable elements in Drosophila: analysis of naturally occurring elements. Genetics 128: 777-784

Mazodier P, Cossart P, Giraud E, Gasser F (1985) Completion of the nucleotide sequence of the central region of Tn5 confirms the presence of three resistance genes. Nucleic Acids Res. 13: 195-205

McClintock B (1945) Cytogenetic studies of maize and Neurospora. Carnegie Inst. of Wash. Year Book 44: 108-112

McClintock B (1946) Maize genetics. Carnegie Inst. of Wash. Year Book 45: 176-186

McClintock B (1947) Cytogenetic studies of maize and Neurospora. Carnegie Inst. of Wash. Year Book 46: 146-152

McClintock B (1948) Mutable loci in maize. Carnegie Inst. of Wash. Year Book 47: 155-169

McClintock B (1949) Mutable loci in maize. Carnegie Inst. of Wash. Year Book 48: 142-154

McClintock B (1951) Chromosome organization and genetic expression. Cold Spring Habor Symp. Quant. Biol. 16: 13-47

McClintock B (1956) Controlling elements and the gene. Cold Spring Habor Symp. Quant. Biol. 21: 197-216

Menssen A, Hermeking H (2002) Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. Proc.Natl.Acad.Sci.U.S.A 99:6274-6279

Miskey C, Izsvak Z, Plasterk RH, Ivics Z (2003) The Frog Prince: a reconstructed transposon from Rana pipiens with high transpositional activity in vertebrate cells. Nucleic Acids Res. 31: 6873-6881

Misra S, Rio DC (1990) Cytotype control of Drosophila P element transposition: the 66 kd protein is a repressor of transposase activity. Cell 62: 269-284

Miyakawa Y, Matsushime H (2001) Rapid downregulation of cyclin D1 mRNA and protein levels by ultraviolet irradiation in murine macrophage cells. Biochem.Biophys.Res.Commun. 284: 71-76

Mizuuchi K, Craigie R (1986) Mechanism of bacteriophage mu transposition. Annu.Rev.Genet. 20: 385-429

Mizuuchi K, Craigie R (1986) Mechanism of bacteriophage mu transposition. Annu.Rev.Genet. 20: 385-429

Moerman DG, Waterston RH (1984) Spontaneous unstable unc-22 IV mutations in C. elegans var. Bergerac. Genetics 108: 859-877

Moran JV, DeBerardinis RJ, Kazazian HH, Jr. (1999) Exon shuffling by L1 retrotransposition. Science 283: 1530-1534

Morisato D, Way JC, Kim HJ, Kleckner N (1983) Tn10 transposase acts preferentially on nearby transposon ends in vivo. Cell 32: 799-807

Morisato D, Kleckner N (1987) Tn10 transposition and circle formation in vitro. Cell 51: 101-111

Mount SM, Rubin GM (1985) Complete nucleotide sequence of the Drosophila transposable element copia: homology between copia and retroviral proteins. Mol.Cell Biol. 5: 1630-1638

Muller H et al. (1994) Cyclin D1 expression is regulated by the retinoblastoma protein. Proc.Natl.Acad.Sci.U.S.A 91: 2945-2949

Muller H et al. (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. Genes Dev. 15: 267-285

Murray A (1994) Cell cycle checkpoints. Curr.Opin.Cell Biol. 6: 872-876

Nishikawa M, Huang L (2001) Nonviral vectors in the new millennium: delivery barriers in gene transfer. Hum.Gene Ther. 12: 861-870

Nishioka Y, Leder A, Leder P (1980) Unusual alpha-globin-like gene that has cleanly lost both globin intervening sequences. Proc.Natl.Acad.Sci.U.S.A 77: 2806-2809

Oosumi T, Belknap WR, Garlick B (1995) Mariner transposons in humans. Nature 378: 672

Peukert K, Staller P, Schneider A, Carmichael G, Hanel F, Eilers M (1997) An alternative pathway for gene regulation by Myc. EMBO J. 16: 5672-5686

Philipp A et al. (1994) Repression of cyclin D1: a novel function of MYC. Mol.Cell Biol. 14: 4032-4043

Pietrokovski S, Henikoff S (1997) A helix-turn-helix DNA-binding motif predicted for transposases of DNA transposons. Mol.Gen.Genet. 254: 689-695

Plasterk RH (1991) The origin of footprints of the Tc1 transposon of Caenorhabditis elegans. EMBO J. 10: 1919-1925

Plasterk RH, Izsvak Z, Ivics Z (1999) Resident aliens: the Tc1/mariner superfamily of transposable elements. Trends Genet. 15: 326-332

Polard P, Chandler M (1995) Bacterial transposases and retroviral integrases. Mol.Microbiol. 15: 13-23

Quelle DE et al. (1993) Overexpression of mouse D-type cyclins accelerates G1 phase in rodent fibroblasts. Genes Dev. 7: 1559-1571

Reznikoff WS et al. (1999) Tn5: A molecular window on transposition. Biochem.Biophys.Res.Commun. 266: 729-734

Rio DC, Barnes G, Laski FA, Rine J, Rubin GM (1988) Evidence for Drosophila P element transposase activity in mammalian cells and yeast. J.Mol.Biol. 200: 411-415

Robertson HM (1993) The mariner transposable element is widespread in insects. Nature 362: 241-245

Roe T, Chow SA, Brown PO (1997) 3'-end processing and kinetics of 5'-end joining during retroviral integration in vivo. J.Virol. 71: 1334-1340

Rothstein SJ, Jorgensen RA, Postle K, Reznikoff WS (1980) The inverted repeats of Tn5 are functionally different. Cell 19: 795-805

Rothstein SJ, Reznikoff WS (1981) The functional differences in the inverted repeats of Tn5 are caused by a single base pair nonhomology. Cell 23: 191-199

Rubin GM, Kidwell MG, Bingham PM (1982) The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell 29: 987-994

Saedler H, Starlinger P (1967a) 0 degree mutations in the galactose operon in E. coli. II. Physiological characterization. Mol.Gen.Genet. 100: 190-202

Saedler H, Starlinger P (1967b) 0 degree mutations in the galactose operon in E. coli. I. Genetic characterization. Mol.Gen.Genet. 100: 178-189

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. Second edition, vol. 1-3, Cold Spring Harbor Laboratory Press

Schouten GJ, van Luenen HG, Verra NC, Valerio D, Plasterk RH (1998) Transposon Tc1 of the nematode Caenorhabditis elegans jumps in human cells. Nucleic Acids Res. 26: 3013-3017

Schukkink RF, Plasterk RH (1990) TcA, the putative transposase of the C. elegans Tc1 transposon, has an N-terminal DNA binding domain. Nucleic Acids Res. 18: 895-900

Seroussi E et al. (1999) TOM1 genes map to human chromosome 22q13.1 and mouse chromosome 8C1 and encode proteins similar to the endosomal proteins HGS and STAM. Genomics 57: 380-388

Serrano M, Hannon GJ, Beach D (1993) A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature 366: 704-707

Shapiro JA (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc.Natl.Acad.Sci.U.S.A 76: 1933-1937

Sherman A et al. (1998) Transposition of the Drosophila element mariner into the chicken germ line. Nat.Biotechnol. 16: 1050-1053

Sherr CJ (1995) D-type cyclins. Trends Biochem.Sci. 20: 187-190

Shykind BM, Kim J, Sharp PA (1995) Activation of the TFIID-TFIIA complex with HMG-2. Genes Dev. 9: 1354-1365

Sinclair AH et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346: 240-244

Spanopoulou E, Zaitseva F, Wang FH, Santagata S, Baltimore D, Panayotou G (1996) The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell 87: 263-276

Staller P et al. (2001) Repression of p15INK4b expression by Myc through association with Miz-1. Nat.Cell Biol. 3:392-399

Stellwagen AE, Craig NL (1997) Avoiding self: two Tn7-encoded proteins mediate target immunity in Tn7 transposition. EMBO J. 16: 6823-6834

Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 164: 49-53

Sved JA, Eggleston WB, Engels WR (1990) Germ-line and somatic recombination induced by in vitro modified P elements in Drosophila melanogaster. Genetics 124: 331-337

Takata M et al. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J. 17: 5497-5508

Tanda S, Corces VG (1991) Retrotransposon-induced overexpression of a homeobox gene causes defects in eye morphogenesis in Drosophila. EMBO J. 10: 407-417

Temin HM, Mizutani S (1970) RNA-dependent DNA polymerase in virions of Rous sarcoma virus. Nature 226: 1211-1213

Travis A, Amsterdam A, Belanger C, Grosschedl R (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function [corrected]. Genes Dev. 5: 880-894

Van Arsdell SW, Denison RA, Bernstein LB, Weiner AM, Manser T, Gesteland RF (1981) Direct repeats flank three small nuclear RNA pseudogenes in the human genome. Cell 26: 11-17

van de WM, Oosterwegel M, van Norren K, Clevers H (1993) Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. EMBO J. 12: 3847-3854

van Gent DC, Hiom K, Paull TT, Gellert M (1997) Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J. 16: 2665-2670

van Luenen HG, Colloms SD, Plasterk RH (1994) The mechanism of transposition of Tc3 in C. elegans. Cell 79: 293-301

Vos JC, Plasterk RH (1994) Tc1 transposase of Caenorhabditis elegans is an endonuclease with a bipartite DNA binding domain. EMBO J. 13: 6125-6132

Vos JC, De B, I, Plasterk RH (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev. 10: 755-761

Vos JC, De B, I, Plasterk RH (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev. 10: 755-761

Voytas DF, Ausubel FM (1988) A copia-like transposable element family in Arabidopsis thaliana. Nature 336: 242-244

Watt F, Molloy PL (1988) High mobility group proteins 1 and 2 stimulate binding of a specific transcription factor to the adenovirus major late promoter. Nucleic Acids Res. 16: 1471-1486

Weinreich MD, Gasch A, Reznikoff WS (1994) Evidence that the cis preference of the Tn5 transposase is caused by nonproductive multimerization. Genes Dev. 8: 2363-2374

Wiegand TW, Reznikoff WS (1992) Characterization of two hypertransposing Tn5 mutants. J.Bacteriol. 174: 1229-1239

Xie W, Gai X, Zhu Y, Zappulla DC, Sternglanz R, Voytas DF (2001) Targeting of the yeast Ty5 retrotransposon to silent chromatin is mediated by interactions between integrase and Sir4p. Mol.Cell Biol. 21: 6606-6614

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based on their structure and reverse transcriptase sequence. The EMBO J. 9: 3353-3362

Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak Z, Kay MA (2000) Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. Nat.Genet.25: 35-41

Yoder KE, Bushman FD (2000) Repair of gaps in retroviral DNA integration intermediates. J.Virol. 74: 11191-11200

Zarkowska T, Mittnacht S (1997) Differential phosphorylation of the retinoblastoma protein by G1/S cyclin-dependent kinases. J.Biol.Chem. 272: 12738-12746

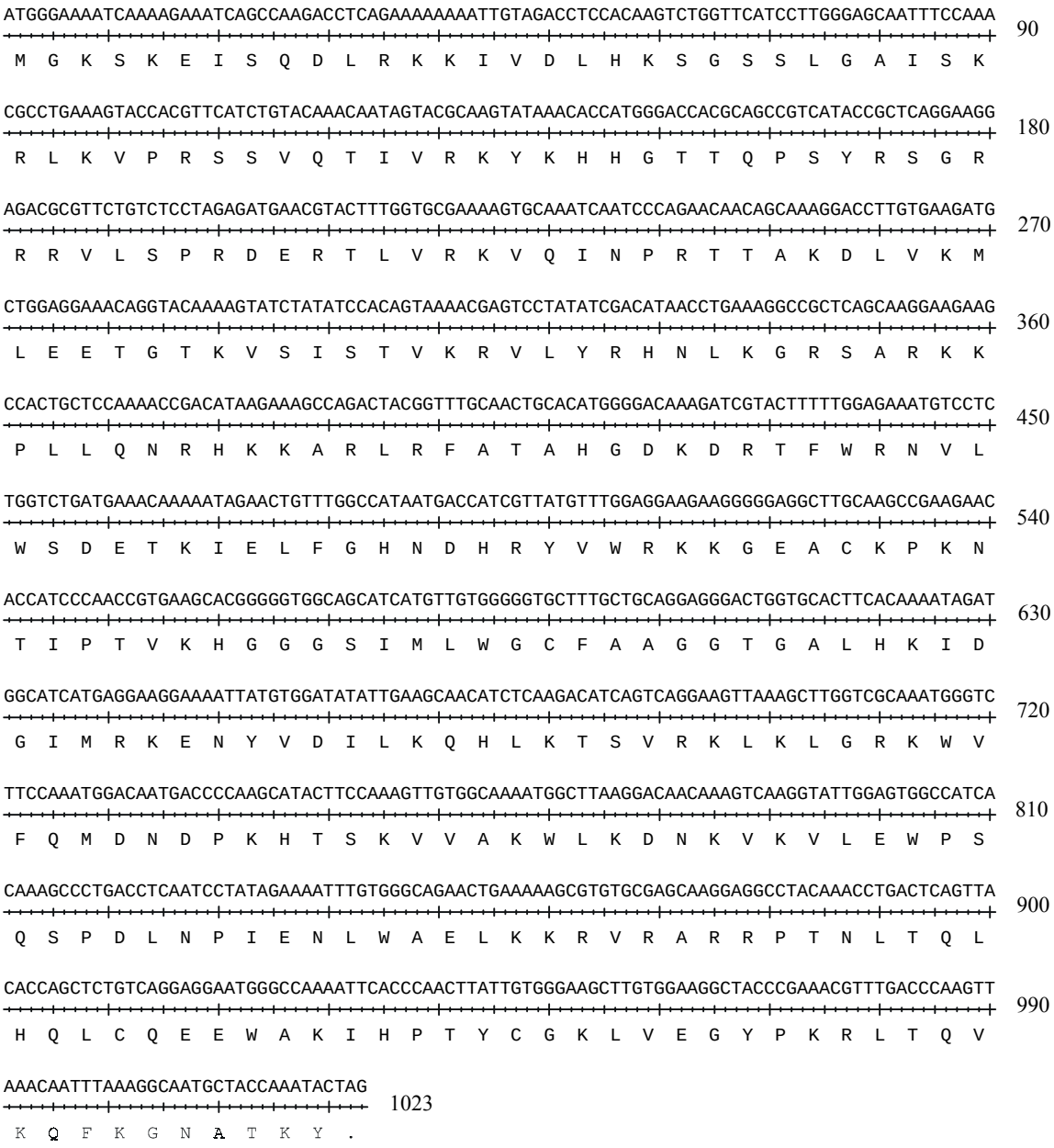
Zayed H, Izsvak Z, Khare D, Heinemann U, Ivics Z (2003) The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. Nucleic Acids Res. 31: 2313-2322

Zeller KI, Haggerty TJ, Barrett JF, Guo Q, Wonsey DR, Dang CV (2001) Characterization of nucleophosmin (B23) as a Myc target by scanning chromatin immunoprecipitation. J.Biol.Chem. 276: 48285-48291

Zou S, Ke N, Kim JM, Voytas DF (1996) The Saccharomyces retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev. 10: 634-645

Appendix

cDNA and amino acid sequence of the Sleeping Beauty transposase.

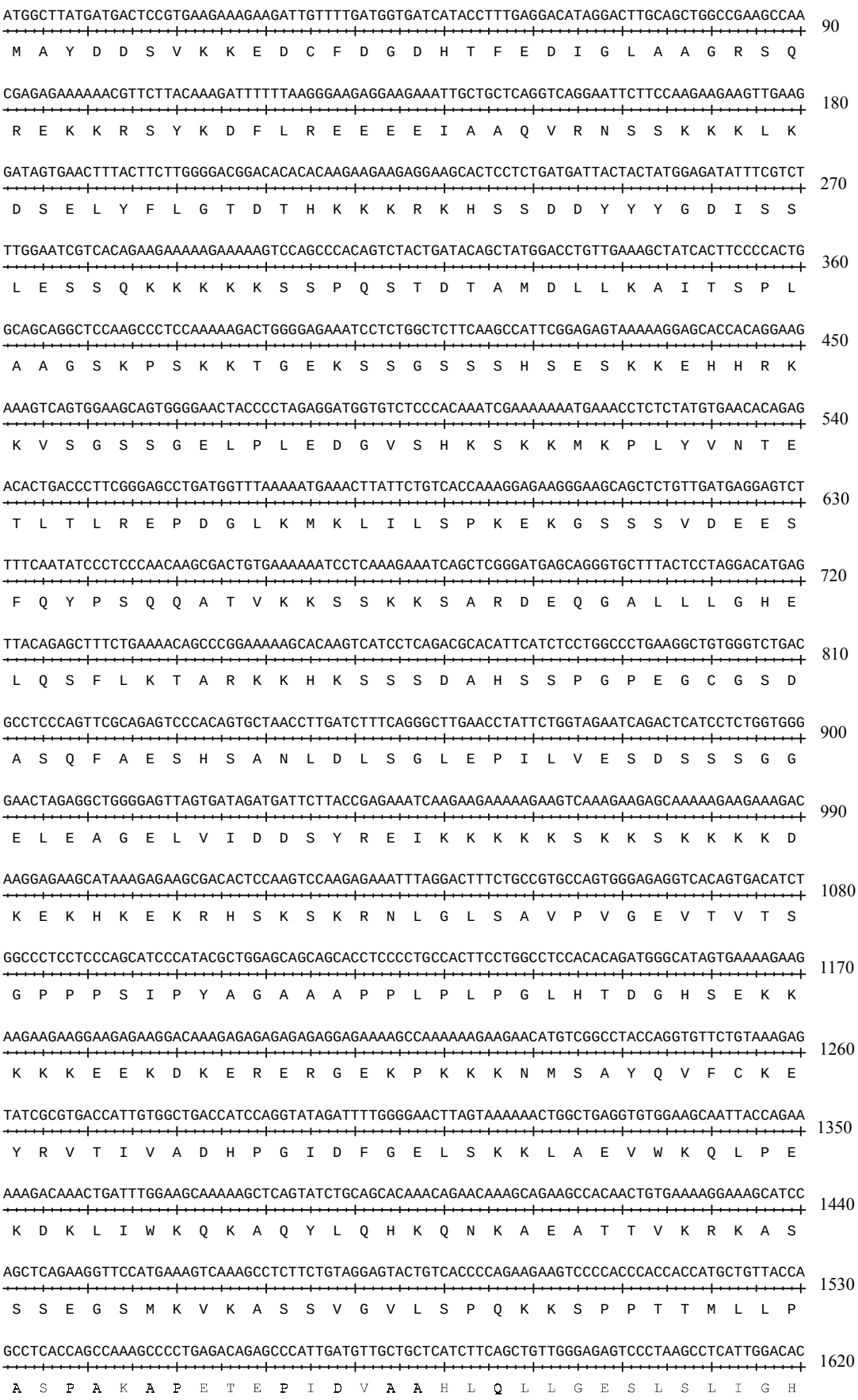


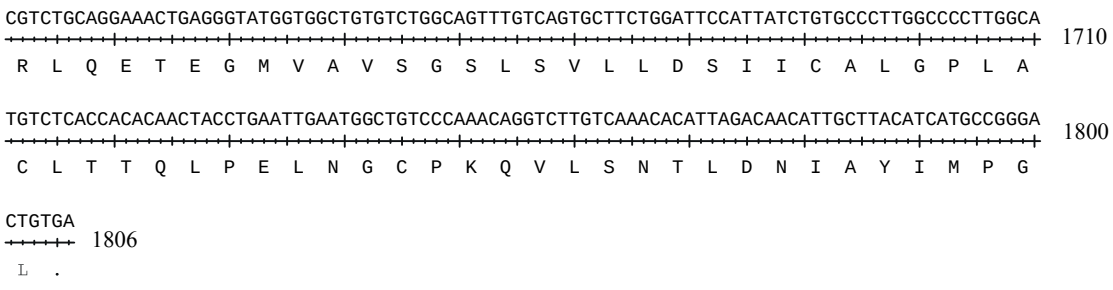
cDNA and amino acid sequence of human Miz-1.

```
ATGGACTTTCCCCAGCACAGCCAGCATGTCTTGGAACAGCTGAACCAGCAGCGGCAGCTGGGGCTTCTCTGTGACTGCACCTTTGTGGTG      90
 M  D  F  P  Q  H  S  Q  H  V  L  E  Q  L  N  Q  Q  R  Q  L  G  L  L  C  D  C  T  F  V  V

GACGGTGTTCACTTTAAGGCTCATAAAGCAGTGCTGGCGGCCTGCAGCGAGTACTTCAAGATGCTCTTCGTGGACCAGAAGGACGTGGTG     180
 D  G  V  H  F  K  A  H  K  A  V  L  A  A  C  S  E  Y  F  K  M  L  F  V  D  Q  K  D  V  V

CACCTGGACATCAGTAACGCGGCAGGCCTGGGGCAGATGCTGGAGTTTATGTACACGGCCAAGCTGAGCCTGAGCCCTGAGAACGTGGAT     270
 H  L  D  I  S  N  A  A  G  L  G  Q  M  L  E  F  M  Y  T  A  K  L  S  L  S  P  E  N  V  D

GATGTGCTGGCCGTGGCCACTTTCCTCCAAATGCAGGACATCATCACGGCCTGCCATGCCCTCAAGTCACTTGCTGAGCCGGCTACCAGC     360
 D  V  L  A  V  A  T  F  L  Q  M  Q  D  I  I  T  A  C  H  A  L  K  S  L  A  E  P  A  T  S

CCTGGGGGAAATGCGGAGGCCTTGGCCACAGAAGGAGGGGACAAGAGAGCCAAAGAGGAGAAGGTGGCCACCAGCACGCTGAGCAGGCTG     450
 P  G  G  N  A  E  A  L  A  T  E  G  G  D  K  R  A  K  E  E  K  V  A  T  S  T  L  S  R  L

GAGCAGGCAGGACGCAGCACACCCATAGGCCCCAGCAGGGACCTCAAGGAGGAGCGCGGCGGTCAGGCCCAGAGTGCGGCCAGCGGTGCA     540
 E  Q  A  G  R  S  T  P  I  G  P  S  R  D  L  K  E  E  R  G  G  Q  A  Q  S  A  A  S  G  A

GAGCAGACAGAGAAAGCCGATGCGCCCCGGGAGCCGCCGCCTGTGGAGCTCAAGCCAGACCCCACGAGTGGCATGGCTGCCGCAGAAGCT     630
 E  Q  T  E  K  A  D  A  P  R  E  P  P  P  V  E  L  K  P  D  P  T  S  G  M  A  A  A  E  A

GAGGCCGCTTTGTCCGAGAGCTCGGAGCAAGAAATGGAGGTGGAGCCCGCCCGGAAAGGGGAAGAGGAGCAAAAGGAGCAAGAGGAGCAA     720
 E  A  A  L  S  E  S  S  E  Q  E  M  E  V  E  P  A  R  K  G  E  E  E  Q  K  E  Q  E  E  Q

GAGGAGGAGGGCGCAGGGCCAGCTGAGGTCAAGGAGGAGGGTTCCCAGCTGGAGAACGGAGAGGCCCCCGAGGAGAACGAGAATGAGGAG     810
 E  E  E  G  A  G  P  A  E  V  K  E  E  G  S  Q  L  E  N  G  E  A  P  E  E  N  E  N  E  E

TCAGCGGGCACAGACTCGGGGCAGGAGCTCGGCTCCGAGGCCCGGGGCCTGCGCTCAGGCACCTACGGCGACCGCACGGAGTCCAAGGCC     900
 S  A  G  T  D  S  G  Q  E  L  G  S  E  A  R  G  L  R  S  G  T  Y  G  D  R  T  E  S  K  A

TACGGCTCCGTCATCCACAAGTGCGAGGACTGTGGGAAGGAGTTCACGCACACGGGGAACTTCAAGCGGCACATCCGCATCCACACGGGG     990
 Y  G  S  V  I  H  K  C  E  D  C  G  K  E  F  T  H  T  G  N  F  K  R  H  I  R  I  H  T  G

GAGAAGCCCTTCTCGTGCCGGGAGTGCAGCAAGGCCTTTTCCGACCCGGCCGCGTGCAAGGCCCATGAGAAGACGCACAGCCCTCTGAAG    1080
 E  K  P  F  S  C  R  E  C  S  K  A  F  S  D  P  A  A  C  K  A  H  E  K  T  H  S  P  L  K

CCCTACGGCTGCGAGGAGTGCGGGAAGAGCTACCGCCTCATCAGCCTGCTGAACCTGCACAAGAAGCGGCACTCGGGCGAGGCGCGCTAC    1170
 P  Y  G  C  E  E  C  G  K  S  Y  R  L  I  S  L  L  N  L  H  K  K  R  H  S  G  E  A  R  Y

CGCTGCGAGGACTGCGGCAAGCTCTTCACCACCTCGGGCAACCTCAAGCGCCACCAGCTGGTGCACAGCGGCGAGAAGCCCTACCAGTGC    1260
 R  C  E  D  C  G  K  L  F  T  T  S  G  N  L  K  R  H  Q  L  V  H  S  G  E  K  P  Y  Q  C

GACTACTGCGGCCGCTCCTTCTCCGACCCCACTTCCAAGATGCGCCACCTGGAGACCCACGACACGGACAAGGAGCACAAGTGCCCACAC    1350
 D  Y  C  G  R  S  F  S  D  P  T  S  K  M  R  H  L  E  T  H  D  T  D  K  E  H  K  C  P  H

TGCGACAAGAAGTTCAACCAGGTAGGGAACCTGAAGGCCCACCTGAAGATCCACATCGCTGACGGGCCCCTCAAGTGCCGAGAGTGTGGG    1440
 C  D  K  K  F  N  Q  V  G  N  L  K  A  H  L  K  I  H  I  A  D  G  P  L  K  C  R  E  C  G

AAGCAGTTCACCACCTCAGGGAACCTGAAGCGGCACCTTCGGATCCACAGCGGGGAGAAGCCCTACGTGTGCATCCACTGCCAGCGACAG    1530
 K  Q  F  T  T  S  G  N  L  K  R  H  L  R  I  H  S  G  E  K  P  Y  V  C  I  H  C  Q  R  Q

TTTGCAGACCCCGGCGCTCTGCAGCGGCACGTCCGCATTCACACAGGTGAGAAGCCATGCCAGTGTGTGATGTGCGGTAAGGCCTTCACC    1620
 F  A  D  P  G  A  L  Q  R  H  V  R  I  H  T  G  E  K  P  C  Q  C  V  M  C  G  K  A  F  T
```


CAGGCCAGCTCCCTCATCGCCCACGTGCGCCAGCACACCGGGGAGAAGCCCTACGTCTGCGAGCGCTGCGGCAAGAGATTCGTCCAGTCC
1710
Q A S S L I A H V R Q H T G E K P Y V C E R C G K R F V Q S
AGCCAGTTGGCCAATCATATTCGCCACCACGACAACATCCGCCCACACAAGTGCAGCGTGTGCAGCAAGGCCTTCGTGAACGTGGGGGAC
1800
S Q L A N H I R H H D N I R P H K C S V C S K A F V N V G D
CTGTCCAAGCACATCATCATTCACACTGGAGAGAAGCCTTACCTGTGTGATAAGTGTGGGCGTGGCTTCAACCGGGTAGACAACCTGCGC
1890
L S K H I I I H T G E K P Y L C D K C G R G F N R V D N L R
TCCCACGTGAAGACCGTGCACCAGGGCAAGGCAGGCATCAAGATCCTGGAGCCCGAGGAGGGCAGTGAGGTCAGCGTGGTCACTGTGGAT
1980
S H V K T V H Q G K A G I K I L E P E E G S E V S V V T V D
GACATGGTCACGCTGGCTACCGAGGCACTGGCAGCGACAGCCGTCACTCAGCTCACAGTGGTGCCGGTGGGAGCTGCAGTGACAGCCGAT
2070
D M V T L A T E A L A A T A V T Q L T V V P V G A A V T A D
GAGACGGAAGTCCTGAAGGCCGAGATCAGCAAAGCTGTGAAGCAAGTGCAGGAAGAAGACCCCAACACTCACATCCTCTACGCCTGTGAC
2160
E T E V L K A E I S K A V K Q V Q E E D P N T H I L Y A C D
TCCTGTGGGGACAAGTTTCTGGATGCCAACAGCCTGGCTCAGCATGTGCGAATCCACACAGCCCAGGCACTGGTCATGTTCCAGACAGAC
2250
S C G D K F L D A N S L A Q H V R I H T A Q A L V M F Q T D
GCGGACTTCTATCAGCAGTATGGGCCAGGTGGCACGTGGCCTGCCGGGCAGGTGCTGCAGGCTGGGGAGCTGGTCTTCCGCCCTCGCGAC
2340
A D F Y Q Q Y G P G G T W P A G Q V L Q A G E L V F R P R D
GGGGCTGAGGGCCAGCCCGCACTGGCAGAGACCTCCCCTACAGCTCCTGAATGTCCCCCGCCTGCCGAGTGA
2412
G A E G Q P A L A E T S P T A P E C P P P A E .

cDNA and amino acid sequence of human HMG2L1.




CGTCTGCAGGAAACTGAGGGTATGGTGGCTGTGTCTGGCAGTTTGTCAGTGCTTCTGGATTCCATTATCTGTGCCCTTGGCCCCTTGGCA
1710
R L Q E T E G M V A V S G S L S V L L D S I I C A L G P L A
TGTCTCACCACACAACTACCTGAATTGAATGGCTGTCCCAAACAGGTCTTGTCAAACACATTAGACAACATTGCTTACATCATGCCGGGA
1800
C L T T Q L P E L N G C P K Q V L S N T L D N I A Y I M P G
CTGTGA
1806
L .

Raw data: Gene chip analysis

			pT/GFP									
			Fugene transfected			Nucleofected (Amaxa)			Statistics			
Gene	GeneBank Acc-#	Affymetrix-ID	Signal (SB)	Signal (K)	Fold change	Signal (SB)	Signal (K)	Fold change	n	Average Fold change	stdv	sem
Cyclin D1	M64339	2017_s_at	103.5	1120.8	-10.8	276	1218	-4.4	4	-7.3	2.7	1.4
		38418_at	725.1	4447	-6.1	539	4327	-8				
Cyclin D3	M92287	1795_g_at	4218	3956	1.1	5511	3690	1.5	4	0.4	1.6	0.8
		1794_at	4395	4609	1	2301	4691	-2				
p16INK4	U26727	1713_s_at	2606	3772	-1.4	2368	4529	-1.9	2	-1.7	0.4	0.3
p21CIP1	U03106	2031_s_at	3765	2046	1.8	2785	2703	1	2	1.4	0.6	0.4
Cyclin E	M73812	2021_s_at	425	435	1	538	551	1	2	1	0.0	0.0
Cyclin A	X51688	1943_at	3645	4544	-1.2	4554	4492	1	2	-0.1	1.6	1.1
Cyclin B	M25753	34736_at	4390	5682	-1.3	4160	5463	-1.3	2	-1.3	0.0	0.0
Cyclin F	U17105	1996_s_at	540	386	1.4	523	538	1	2	1.2	0.3	0.2
Cyclin H	U11791	1924_at	2286	2293	1	945	2274	-2.4	2	-0.7	2.4	1.7
GAPDH	M33197		23229	18527	1.3	41010	20668	2	2	1.7	0.5	0.4

			No pT/GFP									
			Fugene transfected			Nucleofected (Amaxa)			Statistics			
Gene	GeneBank Acc-#	Affymetrix-ID	Signal (SB)	Signal (K)	Fold change	Signal (SB)	Signal (K)	Fold change	n	Average Fold change	stdv	sem
Cyclin D1	M64339	2017_s_at	168.1	1120.8	-6.7	168.1	1218.4	-7.3	4	-7.2	0.3	0.2
		38418_at	595.3	4447.4	-7.4	595.3	4326.8	-7.3				
Cyclin D3	M92287	1795_g_at	3159,6	3956.3	-1.3	3159,6	3689.9	-1.2	4	-1.3	0.1	0.1
		1794_at	3318.1	4609.2	-1.4	3318.1	4691	-1.4				
p16INK4	U26727	1713_s_at	2787.5	3771.9	-1.4	2787.5	4528.8	-1.6	2	-1.5	0.1	0.1
p21CIP1	U03106	2031_s_at	3180.2	2046.4	1.6	3180.2	2703.3	1.2	2	1.4	0.3	0.2
Cyclin E	M73812	2021_s_at	558.2	434.9	1.3	558.2	550.5	1	2	1.2	0.2	0.1
Cyclin A	X51688	1943_at	3481.3	4544	-1.3	3481.3	4491.9	-1.3	2	-1.3	0.0	0
Cyclin B	M25753	34736_at	4225.4	5681.7	-1.3	4225.4	5462.7	-1.3	2	-1.3	0.0	0
Cyclin F	U17105	1996_s_at	641.6	386.2	1.7	641.6	537.7	1.2	2	1.5	0.4	0.2
Cyclin H	U11791	1924_at	2178	2192.5	1	2178	2273.6	1	2	1	0.0	0
GAPDH	M33197		20730.4	18526.7	1.1	20730.4	20668.4	1	2	1.1	0.1	0.1

Raw data: Growth curve analysis

Cell line	Days post inocculation	n	Resuspension volume in ml	Cells per corner square (4 squares counted)	Average cells per square	Cells per well	Average cells per well	stdv	sem
HuH7 (IRES-SB)	1	1	0.05	48; 49; 44; 50	47,8	23900	30900	6596	3808
		2		50; 82; 70; 52	63.5	31800			
		3		78; 76; 82; 60	74	37000			
	3.75	1	0.125	40; 52; 52; 40	46	57500	51400	8032	4637
		2		47; 32; 48; 47	43.5	54400			
		3		38; 38; 21; 38	33.8	42300			
	5.5	1	1	16; 18; 16; 16	16.5	165000	160333	25325	14621
		2		17; 13; 10; 13	13.3	133000			
		3		20; 18; 19; 16	18.3	183000			
	7	1	1.5	24; 10; 19; 21	18.5	277500	229000	46629	26921
		2		18; 13; 13; 16	15	225000			
		3		11; 7; 22; 11	12.3	184500			
	9	1	2	30; 23; 33; 33	29.8	596000	727333	163173	94208
		2		50; 41; 36; 55	45.5	910000			
		3		41; 31; 30; 33	33.8	676000			
HuH7 (IRES-K)	1	1	0.05	59; 53; 57; 54	55.8	27900	30300	2150.19	1300
		2		65; 63; 53; 65	61.5	30800			
		3		64; 58; 77; 58	64.3	32100			
	3.75	1	0.125	53; 48; 56; 72	57.3	71600	73200	14214	8200
		2		72; 64; 75; 71	70.5	88100			
		3		58; 32; 52; 49	47.8	59800			
	5.5	1	1	23; 21; 16; 15	18.8	188000	237000	47571	27500
		2		32; 24; 28; 29	28.3	283000			
		3		19; 31; 23; 23	24	240000			
	7	1	1.5	33; 25; 24; 25	26.8	402000	446800	53695	30900
		2		28; 35; 38; 34	33.8	506300			
		3		29; 25; 31; 30	28.8	432000			
	9	1	2	99; 82; 108; 88	94.3	1886000	1648700	233121	134500
		2		78; 96; 69; 85	82	1640000			
		3		71; 70; 78; 65	71	1420000			

Raw data: Cyclin D1 promoter response to SB transposase in the absence of Miz-1

	Reaction time: 10 min					Average			Average minus blank					Average		
Sample	**Trans-fection**	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**stdv**	**sem**
+ SB-Tpase	1	1	37876	0.547	0.607	39333	0.546	0.604	39164	0.481	0.068	707.4	55.4	56.6	2.8	1.6
		2	41690	0.547	0.603											
		3	38434	0.544	0.601											
	2	1	53310	0.650	0.598	53874	0.648	0.601	53705	0.583	0.065	896.9	59.8			
		2	53573	0.647	0.601											
		3	54741	0.648	0.605											
	3	1	37353	0.531	0.605	38433	0.533	0.603	38264	0.469	0.067	700.0	54.7			
		2	38943	0.531	0.603											
		3	39004	0.539	0.601											
- SB-Tpase	1	1	43299	0.619	0.597	44131	0.618	0.602	43962	0.553	0.066	837.8	52.5	54.7	1.9	1.1
		2	44766	0.617	0.608											
		3	44329	0.618	0.601											
	2	1	55507	0.683	0.596	54802	0.688	0.600	54633	0.623	0.064	973	56.1			
		2	55064	0.696	0.602											
		3	53836	0.684	0.601											
	3	1	46529	0.623	0.597	48498	0.614	0.599	48329	0.548	0.063	869.8	55.6			
		2	46828	0.615	0.596											
		3	52138	0.603	0.605											
untrans-fected		1	177	0.066	0.596	169	0.066	0.599	0	0	NA					
		2	196	0.066	0.600											
		3	134	0.066	0.600											
Buffer blank		1		0.065	0.537	NA	0.065	0.536	NA	NA	0.000					
		2		0.065	0.533											
		3		0.065	0.537											

Raw data: Cyclin D1 promoter response to SB transposase in the presence of wildtype Miz-1

	Reaction time: 5.5 min					Average			Average minus blank					Average		
Sample	**Trans-fection**	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**Fold change**	**sem**
+ SB-Tpase	1	1	77417	1.672	0.459	79081	1.699	0.475	78783	1.629	0.475	623.7	126.3	111.6 (stdv= 15.3)	- 3.0	1.3
		2	80745	1.734	0.487											
		3		1.692	0.479											
	2	1	44349	1.088	0.471	44476	1.109	0.482	44178	1.039	0.482	391.8	112.8			
		2	43782	1.120	0.478											
		3	45296	1.120	0.498											
	3	1	29454	0.859	0.476	29111	0.883	0.491	28813	0.813	0.491	301.0	95.7			
		2	29173	0.897	0.489											
		3	28706	0.894	0.509											
- SB-Tpase	1	1	32662	0.400	0.477	33003	0.416	0.512	32705	0.346	0.512	122.8	266.4	338.2 (stdv= 110.3)		
		2	32674	0.423	0.526											
		3	33672	0.424	0.533											
	2	1	1441	0.076	0.482	1463	0.077	0.508	1165	0.007	0.508	2.5	465.3			
		2	1420	0.078	0.503											
		3	1528	0.077	0.540											
	3	1	8887	0.152	0.488	8819	0.155	0.514	8521	0.085	0.514	30.1	283.4			
		2	8750	0.158	0.521											
		3			0.533											
untrans-fected		1	263	0.069	0.512	298	0.070	0.524	0	0	NA					
		2	333	0.070	0.527											
		3		0.070	0.532											
Buffer blank		1		0.069		NA	0.067	0.000	NA	NA	0.000					
		2		0.066	0.000											
		3		0.067												

Raw data: p15^{INK4b} promoter response to SB transposase in the presence of wildtype Miz-1

Reaction time: 5.5 min						Average			Average minus blank					Average		
Sample	**Trans-fection**	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**Fold change**	**sem**
+ SB-Tpase	1	1	3660	0.215	0.491	3657	0.206	0.517	3359	0.136	0.517	47.7	70.4	187.2 (stdv= 111.0)	- 1.2	1.3
		2	3654	0.215	0.527											
		3		0.187	0.534											
	2	1	21586	0.274	0.484	22172	0.282	0.513	21874	0.212	0.513	75.1	291.4			
		2	22330	0.284	0.511											
		3	22601	0.287	0.543											
	3	1	17699	0.326	0.523	17997	0.332	0.538	17699	0.262	0.538	88.6	199.8			
		2	18087	0.340	0.548											
		3	18205	0.331	0.544											
- SB-Tpase	1	1	513	0.071	0.522	500	0.072	0.541	202	0.002	0.541	0.6	361.2	232.0 (stdv= 114.0)		
		2	532	0.072	0.547											
		3	456	0.072	0.554											
	2	1	4446	0.171	0.515	4629	0.161	0.555	4331	0.091	0.555	29.7	145.7			
		2	4785	0.156	0.573											
		3	4655	0.155	0.576											
	3	1	635	0.075	0.510	636	0.075	0.542	338	0.005	0.542	1.8	189.0			
		2	632	0.075	0.554											
		3	642	0.076	0.561											
untrans-fected		1	263	0.069	0.512	298	0.070	0.524	0	0.000	0.524					
		2	333	0.070	0.527											
		3		0.070	0.532											
Buffer blank		1		0.069			0.067	0.000	NA	NA	0.000					
		2		0.066	0.000											
		3		0.067												

Raw data: Cyclin D1 promoter response to SB transposase and truncated Miz.1

Reaction time = 7.5 min					Minus blank					Average	Statistics			
Sample	n	RLU	A-578	A-595	RLU	A-578	A-595	units ß-gal	RLU/ units ß-gal	RLU/ units ß-gal	stdv	sem	Fold change	sem of fold change
+ Miz-1 (269-803)	1	17508	0.530	0.974	17276	0.452	0.439	137.3	125.8	126.8	1.5	0.9	+ 3.1	0.2
	2	17498	0.549	0.993	17266	0.471	0.458	137.1	125.9					
	3	17627	0.546	0.996	17395	0.468	0.461	135.4	128.5					
- Miz-1 (269-803)	1	4892	0.485	1.035	4660	0.407	0.500	108.5	42.9	41.0	3.6	2.1		
	2	4941	0.494	1.043	4709	0.416	0.508	109.2	43.1					
	3	4423	0.518	1.051	4191	0.440	0.516	113.7	36.9					
+ Miz-1 (269-803) SB-Tpase	1	12788	0.415	1.025	12556	0.337	0.490	91.7	136.9	126.4	14.4	8.3	- 1.2	0.1
	2	12982	0.438	1.033	12750	0.360	0.498	96.4	132.3					
	3	13665	0.518	1.015	13433	0.440	0.480	122.2	109.9					
+Miz-1 (269-803) - SB-Tpase	1	10060	0.324	1.066	9828	0.246	0.531	61.8	159.1	157.3	1.8	1.0		
	2	10193	0.332	1.064	9961	0.254	0.529	64.0	155.6					
	3	10169	0.339	1.086	9937	0.261	0.551	63.2	157.3					
untransfected	1	253	0.078	1.07	0	0.000	0.535							
	2	219	0.079	1.105			0.570							
	3	225	0.078	1.112			0.577							
Buffer blank	1			0.531			0.000							
	2			0.538										
	3			0.537										

Raw data: Cyclin D1 promoter response to wildtype Miz-1 and Myc

Reaction time: 8 min					Minus blank					Average	Statistics			
Sample	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**stdv**	**sem**	**Fold change**	**sem of fold change**
+ Miz-1	1	47194	0.537	0.772	46658	0.471	0.772	76.3	611.8	618.9	20.5	11.8	+ 2.9	0.6
	2	45897	0.534	0.828	45361	0.468	0.828	70.7	642.0					
	3	45759	0.543	0.795	45223	0.477	0.795	75.0	603.0					
- Miz-1	1	11035	0.498	0.856	10499	0.432	0.856	63.1	166.4	210.1	74.5	43.0		
	2	10793	0.501	0.890	10257	0.435	0.890	61.1	167.9					
	3	19731	0.513	0.862	19195	0.447	0.862	64.8	296.1					
+Myc	1	29257	0.493	0.865	28721	0.427	0.865	61.7	465.5	484.7	16.7	9.6	- 1.8	0.1
	2	30970	0.500	0.883	30434	0.434	0.883	61.4	495.4					
	3	30527	0.499	0.890	29991	0.433	0.890	60.8	493.2					
- Myc	1	110029	1.033	0.914	109493	0.967	0.914	132.2	827.9	869.8	72.3	41.8		
	2	108660	1.049	0.941	108124	0.983	0.941	130.6	828.0					
	3	113999	0.982	0.962	113463	0.916	0.962	119.0	953.3					
untransfected	1	594	0.066	1.026	0	0.000								
	2	551	0.066	1.073										
	3	463	0.066	1.06										
Buffer blank	1		0.062	0.000			0.000							
	2		0.062											
	3		0.062											

Raw data: $p15^{INK4b}$ promoter response to wildtype Miz-1 and Myc

Reaction time: 6 min					Minus blank					Average				
Sample	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**stdv**	**sem**	**Fold change**	**sem of fold change**
+ Miz-1	1	266203	0.777	0.352	265879	0.706	0.352	334.3	795.4	804.3	12.4	7.2	+ 17.9	0.3
	2	277636	0.815	0.366	277312	0.744	0.366	338.8	818.5					
	3	281475	0.831	0.360	281151	0.760	0.360	351.9	799.1					
- Miz-1	1	17843	0.946	0.369	17519	0.875	0.369	395.2	44.3	45.1	0.7	0.4		
	2	18038	0.964	0.384	17714	0.893	0.384	387.6	45.7					
	3	18398	0.962	0.373	18074	0.891	0.373	398.1	45.4					
+Myc	1	49929	0.485	0.378	49605	0.414	0.378	182.5	271.7	277.0	10.6	6.1	- 2.0	0.1
	2	49569	0.499	0.419	49245	0.428	0.419	170.2	289.3					
	3	48619	0.498	0.398	48295	0.427	0.398	178.8	270.1					
- Myc	1	151597	0.766	0.416	151273	0.695	0.416	278.4	543.3	542.2	9.2	5.3		
	2	155786	0.833	0.435	155462	0.762	0.435	292.0	532.5					
	3	151690	0.813	0.450	151366	0.742	0.450	274.8	550.8					
untransfected	1	236	0.070	0.319	0	0.000								
	2	393	0.071	0.337										
	3	344	0.071	0.342										
Buffer blank	1		0.068											
	2		0.068	0.000										
	3		0.068											

Raw data: Cyclin D1 promoter response to increasing amounts of SB transposase

Reaction time: 4.5 min						Average			Average minus blank					Average			RLU/ units ß-gal (normalized)	
ng CMV-SB	Trans-fection	n	RLU	A-578	A-595	RLU	A-578	A-595	RLU	A-578	A-595	units ß-gal	RLU/ units ß-gal	RLU/ units ß-gal	stdv	sem		sem
0 ng	1	1	557762	0.739	0.817	560656	0.740	0.819	560420	0.669	0.292	509.1	1100.7	999.1	143.3	82.7	1.00	NA
		2	563549	0.741	0.821													
	2	1	564675	0.741	0.740	559895	0.747	0.751	559659	0.676	0.224	670.1	835.1					
		2	555115	0.752	0.762													
	3	1	582183	0.791	0.814	571836	0.809	0.832	571600	0.738	0.305	538.6	1061.3					
		2	561488	0.827	0.849													
2.5 ng	1	1	329513	0.600	0.848	329241	0.603	0.858	329005	0.532	0.331	357.2	921.2	868.2	47.1	27.2	0.87	0.10
		2	328968	0.606	0.868													
	2	1	378319	0.677	0.817	377579	0.683	0.827	377343	0.612	0.300	454.1	831.0					
		2	376839	0.689	0.836													
	3	1	395562	0.658	0.808	392568	0.669	0.816	392332	0.598	0.289	460.2	852.5					
		2	389574	0.679	0.823													
5 ng	1	1	233930	0.535	0.809	231835	0.544	0.820	231599	0.473	0.293	358.7	645.6	715.6	69.3	40.0	0.72	0.10
		2	229739	0.553	0.831													
	2	1	242712	0.546	0.862	242133	0.552	0.874	241897	0.481	0.347	308.5	784.2					
		2	241554	0.558	0.885													
	3	1	247768	0.534	0.817	250248	0.548	0.831	250012	0.477	0.304	348.7	717.0					
		2	252727	0.562	0.845													
10 ng	1	1	128785	0.401	0.912	131046	0.410	0.912	130810	0.339	0.385	195.4	669.5	590.1	79.9	46.1	0.59	0.09
		2	133306	0.418	0.912													
	2	1	146320	0.431	0.839	146373	0.431	0.851	146137	0.360	0.324	247.3	590.9					
		2	146425	0.431	0.862													
	3	1	136722	0.406	0.823	136931	0.413	0.810	136695	0.342	0.283	268.2	509.8					
		2	137139	0.419	0.797													
20 ng	1	1	104887	0.381	0.819	104656	0.388	0.839	104420	0.317	0.312	225.8	462.5	459.8	35.7	20.6	0.46	0.06
		2	104424	0.394	0.858													
	2	1	101502	0.383	0.797	103801	0.390	0.816	103565	0.319	0.289	244.9	422.9					
		2	106100	0.396	0.835													
	3	1	108614	0.398	0.867	107983	0.409	0.871	107747	0.338	0.344	218.0	494.2					
		2	107351	0.419	0.875													

						Average			Average minus blank					Average			RLU/ units ß-gal (normalized)	
ng CMV-SB	**Transf-ection**	**n**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**RLU**	**A-578**	**A-595**	**units ß-gal**	**RLU/ units ß-gal**	**RLU/ units ß-gal**	**stdv**	**sem**		**sem**
50 ng	1	1	283942	0.703	0.815	286969	0.744	0.826	286733	0.673	0.299	500.7	572.7	587.1	31.7	18.3	0.59	0.07
		2	289996	0.784	0.836													
	2	1	302401	0.742	0.810	301442	0.767	0.817	301206	0.696	0.290	533.0	565.2					
		2	300483	0.791	0.824													
	3	1	310581	0.830	0.846	310653	0.826	0.864	310417	0.755	0.337	497.9	623.5					
		2	310725	0.822	0.882													
untransfected		1	236	0.071	0.854	236	0.071	0.854	0	0.000	0.327	0.0						
Buffer blank		1		0.078	0.527						0.000							

Raw data: Transposition assay in serum-starved CHO-K1 cells

	1% serum containing medium						0.5% serum containing medium					
	Colonies per plate		Transposition Efficiency	Average Transposition Efficiency	stdv	sem	Colonies per plate		Transposition Efficiency	Average Transposition Efficiency	stdv	sem
n	+ Tpase	- Tpase					+ Tpase	- Tpase				
1	363	50	7.3	6.5	2.2	0.8	108	10	10.8	9.9	3.5	1.3
2	479	86	5.6				180	12	15.0			
3	111	16	6.9				104	7	14.9			
4	196	28	7.0				276	34	8.1			
5	164	19	8.6				159	19	8.4			
6	180	19	9.5				158	16	9.9			
7	80	22	3.6				14	2	7.0			
8	110	32	3.4				57	11	5.2			